Strengthening Data Sovereignty Through Digital Text Watermarking

Dissertation

zur Erlangung des Grades eines

DOKTORS DER INGENIEURSWISSENSCHAFTEN

der Technischen Universität Dortmund
an der Fakultät für Informatik

von

Malte Hellmeier

Dortmund

2026

Tag der mündlichen Prüfung: 17.04.2026
Dekan: Prof. Dr. Jens Teubner
Gutachter:

Prof. Dr. Falk Howar

Prof. Dr.-Ing. Frederik Möller

Abstract

With the continuous shift from the analog to the digital world, data has become a common resource in our everyday lives. Individuals care about protecting their personal data, and companies keep data secure to gain a competitive advantage. Nevertheless, sharing and using data in an interconnected world is necessary to create value through cooperation. Various international regulations increase the complexity of the tension between data sharing and data protection. Keeping control over data, often referred to as data sovereignty, is of utmost importance. However, the concept of data sovereignty is interpreted differently and lacks a precise delimitation and clear conceptualization. Existing technical solutions for maintaining data sovereignty remain limited and specialized.

This cumulative doctoral dissertation uses Design Science Research to address the lack of technical solutions for strengthening data sovereignty. First, we conceptualized and delimited data sovereignty from adjacent terms in the domains of information systems and software engineering. We reviewed the existing literature and conducted interviews with practitioners, thereby identifying data sovereignty challenges and highlighting digital watermarking technologies as a promising solution approach. Second, we designed and developed a digital text watermarking artifact, Innamark, to mitigate the identified data sovereignty challenges. It focuses on protecting text, the most widely shared and used content type, with limited watermarking solutions available. Third, we derived design knowledge in the form of design principles to generalize our findings based on our experience and gathered insights.

We instantiated Innamark as an IT artifact, including a reusable Kotlin multiplatform library. Different demonstrations, in the form of implemented applications and a test in a data space context, supported us during the iterative design process. Further, we evaluated Innamark with respect to embedding capacity, invisibility, and robustness. Using a benchmark evaluation on a dataset of 1 000 000 articles, we compared Innamark with nine related algorithms. To the best of our knowledge, the findings show that Innamark is the first invisible and robust watermarking method that does not increase the length of the watermarked text, despite its limited embedding capacity.

Our three contributions (i) the conceptual model for data sovereignty, (ii) Innamark as a digital watermarking artifact, and (iii) the resulting design principles, support a broad range of applications and help researchers and practitioners to strengthen data sovereignty.

Keywords: Data Sovereignty, Digital Watermarking, Text Watermarking, Information Hiding, Design Science Research, Information Systems, Software Engineering.

List of Contributions

Peer-Reviewed Scientific Publications for this Thesis

I **A Delimitation of Data Sovereignty from Digital and Technological Sovereignty**
by Malte Hellmeier, Franziska von Scherenberg. In *European Conference on Information Systems 2023 Research Papers*, ECIS 2023. 306, pp. 1–19. Kristiansand, Norway, 2023. https://aisel.aisnet.org/ecis2023_rp/306.
(In this document cited as Paper I [100]).

II **Implementing Data Sovereignty: Requirements & Challenges from Practice**
by Malte Hellmeier, Julia Pampus, Haydar Qarawlus, Falk Howar. In *Proceedings of the 18th International Conference on Availability, Reliability and Security*, ARES 2023. 143, pp. 1–9. Benevento, Italy, 2023. DOI: 10.1145/3600160.3604995.
(In this document cited as Paper II [95]).

III **Data Sovereignty in Information Systems**
by Franziska von Scherenberg, Malte Hellmeier, Boris Otto. In *Electronic Markets*, 34:15, pp. 1–11. Springer, 2024. DOI: 10.1007/s12525-024-00693-4.
(In this document cited as Paper III [271]).

IV **A Hidden Digital Text Watermarking Method Using Unicode Whitespace Replacement**
by Malte Hellmeier, Haydar Qarawlus, Hendrik Norkowski, Falk Howar. In *Proceedings of the 58th Hawaii International Conference on System Sciences*, HICSS 2025. pp. 7411–7420. Waikoloa Village, Big Island, Hawaii, USA, 2025. DOI: 10.24251/hicss.2025.886.
(In this document cited as Paper IV [99]).

V **Strengthening Data Sovereignty Through Digital Watermarking in Data Spaces**
by Malte Hellmeier, Haydar Qarawlus. In *Proceedings of the 58th Hawaii International Conference on System Sciences*, HICSS 2025. pp. 4346–4355. Waikoloa Village, Big Island, Hawaii, USA, 2025. DOI: 10.24251/hicss.2025.520.
(In this document cited as Paper V [96]).

VI **Innamark: A Whitespace Replacement Information Hiding Method**
by Malte Hellmeier, Hendrik Norkowski, Ernst-Christoph Schrewe, Haydar Qarawlus,

Falk Howar. In *IEEE Access*, 13, pp. 123120–123135. IEEE, 2025. DOI: 10.1109/ACCESS.2025.3583591.
(In this document cited as Paper VI [94], previously published as preprint in [93]).

VII **A Fragile Watermarking Technique for Integrity Authentication of CSV-Files Using Invisible Line-Ending Control Characters**
by Florian Zimmer, Malte Hellmeier, Motoki Nakamura, Tobias Urbanek. In *Proceedings of the 22nd International Conference on Security and Cryptography*, SECRYPT 2025. pp. 455–466. Bilbao, Spain, 2025. DOI: 10.5220/0013559600003979.
(In this document cited as Paper VII [288]).

Patents and Artifacts for this Thesis

i **German Patent Application**
Verfahren und System zur zeichenanzahlneutralen Einbettung einer digitalen Signatur in ein digitales Dokument
by Malte Hellmeier, Haydar Qarawlus, Hendrik Norkowski. Currently assigned to Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V. (accessed on Feb. 10, 2026). Application number: DE102023125012.4A. Priority: 15.09.2023. Application: 15.09.2023. Publication: 20.03.2025.
(In this document cited as [97]).

ii **International Patent Application**
Method and System for the Character-Number-Neutral Embedding of a Digital Signature in a Digital Document
by Malte Hellmeier, Haydar Qarawlus, Hendrik Norkowski. Currently assigned to Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V. (accessed on Feb. 10, 2026). Application number: PCT/EP2024/075670. Priority: 15.09.2023. Application: 13.09.2024. Publication: 20.03.2025.
(In this document cited as [98]).

iii **Innamark Source Code**
by Fraunhofer ISST. Public GitHub repository. https://github.com/FraunhoferISST/Innamark (accessed on Feb. 10, 2026).
(In this document cited as [75]).

Other Peer-Reviewed Scientific Publications

1. **Sovereign Data Exchange in Cloud-Connected IoT using International Data Spaces**
   by Haydar Qarawlus, Malte Hellmeier, Johannes Pieperbeck, Ronja Quensel, Steffen Biehs, Marc Peschke. In *IEEE Cloud Summit*, 2021. pp. 13–18. Hempstead, NY, USA, 2021. DOI: 10.1109/IEEECloudSummit52029.2021.00010.

2. **Increasing the Business Value Of Free-Floating Carsharing Fleets By Applying Machine-Learning Based Relocations**
   by Christoph Prinz, Malte Hellmeier, Mathias Willnat, Christine Harnischmacher, Lutz M. Kolbe. In *European Conference on Information Systems 2022 Research Papers*, ECIS 2022. 70, pp. 1–15. Timișoara, Romania, 2022. https://aisel.aisnet.org/ecis2022_rp/70.

3. **A Digitization Pipeline for Mixed-Typed Documents Using Machine Learning and Optical Character Recognition**
   by Tizian Matschak, Florian Rampold, Malte Hellmeier, Christoph Prinz, Simon Trang. In *The Transdisciplinary Reach of Design Science Research*, DESRIST 2022. *Lecture Notes in Computer Science*, 13229, pp. 195–207. St Petersburg, FL, USA, 2022. DOI: 10.1007/978-3-031-06516-3_15.

4. **Supporting Changes in Digital Ownership and Data Sovereignty Across the Automotive Value Chain with Catena-X**
   by Marvin Michael Manoury, Theresa Riedelsheimer, Malte Hellmeier, Tom Meyer. In *Procedia Computer Science*, ISM 2024. 253, pp. 374–383. Elsevier B.V., 2025. DOI: 10.1016/j.procs.2025.01.099.

5. **Facilitating Data Usage Control Through IPv6 Extension Headers**
   by Haydar Qarawlus, Malte Hellmeier, Falk Howar. In *Proceedings of the 14th International Conference on Data Science, Technology and Applications*, DATA 2025. pp. 526–535. Bilbao, Spain, 2025. DOI: 10.5220/0013566200003967.

6. **Security and Detectability Analysis of Unicode Text Watermarking Methods Against Large Language Models**
   by Malte Hellmeier. In *12th International Conference on Information Systems Security and Privacy*, ICISSP 2026. pp. 197–204. Marbella, Spain, 2026. DOI: 10.5220/0014268700004061.

Comments on My Participation

I **A Delimitation of Data Sovereignty from Digital and Technological Sovereignty**
As the first author, I am the lead author of all sections of the paper and created the first full draft. The literature review was conducted in close collaboration with the other author, Franziska von Scherenberg. I conducted the final analysis by creating all figures and tables. I presented the paper in person at the conference.

II **Implementing Data Sovereignty: Requirements & Challenges from Practice**
As the first author, I am the lead author of sections two, three, and five and co-author of all other sections. The results are based on interviews conducted and analyzed by the first three authors of the paper. Thus, the results section four was created jointly by the same authors. I presented the paper in person at the conference.

III **Data Sovereignty in Information Systems**
The fundamentals paper builds on the initial results reported in conference Paper I. The conceptual model was developed based on the literature results and discussions among all authors. As the second author, I co-authored all sections of the paper and led the authorship of all technical parts and the exemplified description of the model. The paper received the *Most Cited Paper* award in 2024 from Electronic Markets.

IV **A Hidden Digital Text Watermarking Method Using Unicode Whitespace Replacement**
As the first author, I created the first full draft of all sections of the paper. The presented watermarking algorithm was developed based on discussions from all authors. The presented implementation was managed by me, with the main development work done by Hendrik Norkowski for the library and CLI tool, and the main development work done by me for the web interface. I presented the paper in person at the conference.

V **Strengthening Data Sovereignty Through Digital Watermarking in Data Spaces**
As the first author, I am the lead author of all sections of the paper and created the first full draft. All figures and tables were created based on discussions of all authors. The development process was carried out by all authors with support from Ernst-Christoph Schrewe, as mentioned in the acknowledgments. I presented the paper in person at the conference. The paper was nominated for the best paper award.

VI **Innamark: A Whitespace Replacement Information Hiding Method**
This journal paper is an extended version of Paper IV, while the contributions of the authors are the same. Ernst-Christoph Schrewe joined the author team by supporting the development of the related work for the evaluation, which I executed. As the first author, I led the authorship of all sections and created all figures and tables based on initial versions of the plot figures from Haydar Qarawlus.

VII **A Fragile Watermarking Technique for Integrity Authentication of CSV-Files Using Invisible Line-Ending Control Characters**
As the second author, I am the lead author of the watermarking part in section two and a co-author of all other sections. As stated in the CRediT author statement, my contributions include the conceptualization, methodology, investigation, writing of the original draft, and visualizations. The paper was certified as a best paper candidate.

i **German Patent Application**
Verfahren und System zur zeichenanzahlneutralen Einbettung einer digitalen Signatur in ein digitales Dokument
As the first inventor, I contributed the main idea and concepts behind the application with a share of >50% of the invention. The patent application was written by the patent attorney firm *Pfenning, Meinig & Partner mbB*. The content is the same as the international application in [98].

ii **International Patent Application**
Method and System for the Character-Number-Neutral Embedding of a Digital Signature in a Digital Document
As the first inventor, I contributed the main idea and concepts behind the application with a share of >50% of the invention. The patent application was written by the patent attorney firm *Pfenning, Meinig & Partner mbB*. The content is the same as the German application in [97].

iii **Innamark Source Code**
As the project lead, I coordinated the overall architecture and development of the Innamark reference implementation with its watermarker library, web interface, and command-line tool. The concrete contributions emerge from the GitHub commit history.

Acknowledgements

It feels surreal to finally bring the dissertation to an end by preparing this thesis for publication. This work was developed during my time as a research associate at Fraunhofer ISST in collaboration with TU Dortmund University. While science and research are primarily shaped by communication, team effort, and collaboration, this doctoral dissertation would not have been possible without a strong backbone of various people. I would therefore like to express my gratitude to a number of people for both their professional subject-specific input and their personal support.

First and foremost, I want to thank my supervisor and challenger, Prof. Dr. Falk Howar, for his continuous support. Thank you for various calls, critical questions, deep discussions, consistent trust in me and my work, and the generous freedom to build my own research agenda. Looking back, I am amazed at how small ideas that arose during our discussions developed into new directions and possibilities of my research journey.

I would also like to express my thanks to Prof. Dr.-Ing. Frederik Möller and the other committee members for providing feedback and challenging my thesis.

I am especially grateful for my scientific colleagues from Fraunhofer ISST and beyond. Special thanks go to my coworker Tom and to our department managers, Dr. Jürgen Schmelting in the former Logistics department, and his successor, Dr.-Ing. Joachim Hunker, in the Industrial Manufacturing department. Thank you for your feedback, support, and discussions at the office desk or in the kitchen while enjoying a great French press coffee brewed from freshly ground beans. I would like to thank the current and former student assistants I have worked with, namely Christoph, Carolin, Hendrik, Joris, and Ernst. Your continued interest and work helped bring these ideas and the Innamark project to life and running code.

I am deeply proud to have worked with all my co-authors and would like to express my particular appreciation to my two main research buddies, Franziska and Haydar. Thank you for bringing so many papers to life with me, even in late working hours and productive weekend discussions. I remain thankful for the many interesting people I met at conferences around the globe, including my travel buddies. Thank you for the long talks and the great fun we had together.

Last but not least, I am so grateful to my wife, Benita. Thank you for your deepest trust and moral support. I also thank my family for always believing in me and continuously supporting me. Further, I want to thank all my friends for their understanding when I focused on work, and for reminding me that there is a great life outside of doctoral studies.

Thank you all, let's see what the future brings!

Table of Contents

List of Figures

List of Tables

List of Acronyms

AI	artificial intelligence
API	application programming interface
ASCII	American Standard Code for Information Interchange
CCPA	California Consumer Privacy Act
CLI	command line interface
CRC	cyclic redundancy check
CSV	comma-separated values
DP	design principle
DPP	Digital Product Passport
DSP	Dataspace Protocol
DSR	Design Science Research
EDC	Eclipse Dataspace Components
ERP	enterprise resource planning
ESPR	Ecodesign for Sustainable Products Regulation
GDPR	General Data Protection Regulation
GUI	graphical user interface
HMAC	Hash-based Message Authentication Code
IDSA	International Data Spaces Association
IoT	Internet of Things
KC	knowledge claim
LLM	large language model
LSB	least significant bit
MRQ	main research question
OEM	original equipment manufacturer
PIE	Policy Information Extension
RQ	research question
SMS	Short Message Service
SNOW	Steganographic Nature of Whitespace
UML	Unified Modeling Language
WEE	Watermark Embedding Extension

Chapter 1

Introduction

1.1 Motivation

The rapid increase in digital technologies and the vast amount of shared data explain our societal shift from the analog to a digital world, raising opportunities and tensions around data protection. The total amount of data and information being created, captured, copied, and consumed worldwide reached 173.4 zettabytes in 2024 and is estimated to grow rapidly to 527.5 zettabytes in 2029 [247]. In addition, that data is not only stored locally on storage media but also shared globally, as evidenced by the worldwide amount of mobile traffic, which reached 122.97 exabytes per month in 2024 and is expected to grow to 244.65 exabytes per month in 2029 [65]. Famous and often quoted statements reinforce these assumptions, such as the analogy of information as the oil of the 21st century by Peter Sondergaard from Gartner Research in 2011, or the analogy of data as the world's most important asset by Yuval Noah Harari communicated at the World Economic Forum 2018 [107]. Our interconnected world continues to evolve, with computers and mobile devices becoming faster and more efficient, seemingly limited only by Moore's law [180].

Societal trends toward digitized processes are evident in our daily lives, both in the private and corporate spheres: individuals scan old analog photos printed on photo paper to share and use them digitally, or scan their personal identity documents for verification on the banking portal. Companies initiate data governance initiatives by integrating supplier delivery note information into their enterprise resource planning (ERP) systems as part of their digital transformation [272]. The reasons behind these *digitization* efforts, which involve converting data into digital formats, stem from the need and desire for more *digitalization*, as the use of digital data improves processes and creates value [191], [251], [272]. This technological progress is accompanied by the emergence of new applications and technologies, as evidenced by the recent interest around large language models (LLMs), such as ChatGPT [252]. If we digitize our data, we can share and use it jointly, such as training LLMs for specific tasks, helping many participants. Our society has to find ways to share data without losing control over it, as cooperation enables better growth than competition, where data is used only in isolated systems [119], [211].

Challenges. Despite the benefits of joint data use, various challenges exist regarding regulations and competition, while the technical debt of existing data sharing IT landscapes creates skepticism and mistrust.

First, regulatory compliance both supports and hinders data sharing. One example is the European regulation regarding a Digital Product Passport (DPP), launched in 2024 as part of the Ecodesign for Sustainable Products Regulation (ESPR) [70]. It forces companies to provide a DPP that includes information such as the product's life cycle and sustainability metrics. Calculating sustainability metrics, such as a carbon footprint, as the sum of all greenhouse gas emissions [70], requires data sharing between companies, especially for complex products with extended supply chains, such as vehicles. While the ESPR supports data sharing, the 2018 introduced U.S. CLOUD Act hinders it, as individuals and companies fear losing control over their data since the act allows American authorities to access data stored by American companies [106]. Such regulations pose challenges for data localization [127].

Second, competitive advantages still challenge data sharing. On the one hand, different use cases, such as predictive maintenance or collaborative condition monitoring, in which actors cooperate to improve maintenance efforts, require many contributing companies and shared data [119]. On the other hand, most actors want to use data, but do not want to share it. A German study of 244 companies found that 51% reported strong business success through data sharing, while 30% used data from other companies, but only 17% shared it with external parties [248]. Many companies avoid sharing data because they fear losing control over it [191].

Third, a lack of traceability, security issues, unclear data integrity, or authenticity problems creates skepticism and mistrust in many sectors: "With the transition of the printing industry to digital platforms, the security of documents has become a major challenge" [177, p. 85]. Digital content, such as text documents, papers, or contracts, that requires high protection often faces illegal duplication or contextual tampering [91], [237]. The aforementioned hype surrounding LLMs reinforces skepticism and mistrust, making it increasingly difficult to distinguish a text written by a human from a text generated by a machine [44], [46], [136], [228], [252], [285]. Various risks arise from these LLM-generated texts, such as plagiarism, piracy of copyrighted and rewritten content, or LLM-generated phishing, spam, malware, and misinformation [57], [263], [285]. We need to build robust solutions for copyright protection, authentication, and integrity verification to overcome these challenges, especially in data sharing scenarios among different parties [8], [26], [286].

Existing Solutions. Researchers and practitioners propose approaches to address these challenges, but current solutions remain limited. The missing possibilities of keeping control over data are often referred to as *data sovereignty* [119], [227]. Initial attempts in related fields, such as digital or technological sovereignty, have focused on identifying indirect solutions at the organizational or political level [20], [62]. However, we focus on finding direct technical solutions in the domain of information systems and software engineering.

One existing approach is decentralized networks, like blockchains [42], [105], [224] or data spaces [179], [183]. Blockchains help ensure trust and security through their well-known distributed consensus algorithm [51]. However, they also face many disadvantages, such as scaling and bootstrapping issues, as well as the potential for fraudulent activity due to their pseudonymous nature [51]. In contrast, data spaces as a concept explicitly promise data sovereignty, low entry barriers, interoperability, decentralization, and secure data sharing [179], [183]. Initiatives such as Gaia-X and the International Data Spaces Association (IDSA) support and guide data space development, aiming to establish trust and provide specific infrastructure, services, and standards [192]. However, practice reveals challenges and technical debt because the concepts only create framework conditions rather than working technical solutions [95].

Other approaches that do not explicitly mention data sovereignty but face similar challenges include watermarking and related research areas, such as steganography and digital rights management. These information hiding methods conceal data within cover media to be protected, such as image, audio, video, or text files [5], [8]. Those methods appear promising, since they have been used for many years to protect various assets, such as physical bank notes and passports with their safety features [177], movies, as known from copy-protected DVDs [159], or images, as known from watermarks embedded in stock-photo databases before buying [28]. The ideas are currently being adopted in the LLM domain, for example, by Kirchenbauer et al. [134] as one of the first watermark approaches for LLMs, by SynthID, as Google's watermarking engine for Gemini [54], and by many similar approaches [158]. However, neither decentralized solutions nor existing digital watermarking approaches solve the sovereignty and robustness requirements of shared text data.

Research Gap. A closer review identifies conceptual and technical research gaps in both data sovereignty and digital watermarking, which motivates this thesis.

Regarding data sovereignty, evaluating the number of publications reveals a continued increase in interest in the field, while theoretical and practical gaps remain [100]. From a theoretical perspective, researchers often use related terms, such as digital or technological sovereignty, without clear differentiation, which leads to conflicting definitions [173] and

an unclear relationship between them [48]. A clear academic delimitation and conceptualization of the field is under-researched [2], [20]. This conceptual fragmentation results in different research streams unaware of one another, as the first group publishes its findings under keywords related to sovereignty. In contrast, other research groups use keywords such as security, privacy, or data control while working on very similar aspects [80]. From a practical perspective, practitioners lack concrete, runnable technical solutions [181], [251]. The theoretical basis, political discourses, and concrete interests, as revealed by the mentioned surveys, indicate a high level of interest on one side, alongside a notable gap in technical practice-oriented data sovereignty solutions on the other [52], [181], [290].

Regarding digital watermarking, various solutions for different multimedia covers exist. However, text data represents between 75% and 85% of the Internet [286] and is the most shared content type [30], [219], but has the fewest watermarking solutions available [3], [11]. This lack of digital text watermarking solutions stems from the limited options and high complexity of plain text compared to the greater flexibility of other multimedia types, such as images, audio, or video [23], [26], [30], [126]. Existing LLM watermarking techniques do not transfer to our problem domain because their solutions directly integrate into the training procedure or model itself [158]. These techniques embed watermarks during text generation and do not work on existing text documents, such as scanned documents or digitized delivery notes from our example above. Other linguistic-based methods change the wording of existing text to embed a watermark, leading to problems in application scenarios with high semantic and text-quality standards, such as legal documents, quotes, or poems [11], [121]. For these use cases, staying invisible without interfering with daily use and being robust enough to maintain data sovereignty even when used in different contexts, file types, and applications is crucial [5], [11]. Our evaluation reveals that existing text watermarking solutions that do not alter the semantics of the text are either visible to humans or lack robustness in typical business applications, such as Microsoft Word documents, text messages, or emails [94].

Contribution. To address the research gaps, we make three key contributions in this cumulative doctoral dissertation, structured overall by the Design Science Research (DSR) methodology following Peffers et al. [197].

First, we structure the concept of data sovereignty through literature reviews and derive a conceptual model from scientific and gray literature to clarify the cluttered field of sovereignty terms. To carefully identify the current requirements and challenges for data sovereignty in practical use cases, we conduct a semi-structured interview study using Grounded

Theory. The gathered insights help us guide the design and development of a technical solution.

Second, we design, implement, and evaluate a digital text watermarking artifact to address these data sovereignty challenges. During the iterative DSR development cycles, we design and develop this digital text watermarking artifact using Unicode whitespace replacement to hide any byte-encoded information within a cover text. A reference implementation demonstrates the artifact in different applications, such as a web interface, command line interface (CLI) tool, and a data space context. We further evaluate the proposed artifact in a testbed, comparing it with nine algorithms from related work to demonstrate its strengths and weaknesses.

Third, we generalize the gathered design knowledge as design principles (DPs) for researchers and practitioners. We use established structures and formats from related work for their formulations [39], [87] and evaluate them against five established criteria [112]. The resulting DPs aim to help researchers build alternative digital watermarking artifacts to strengthen data sovereignty.

1.2 Research Questions

We break down the aforementioned challenges into two main research directions on data sovereignty and digital watermarking. First, little attention has been paid to defining and delimiting data sovereignty on a theoretical basis and to developing technical solutions to realize it in practice. Second, existing solutions in the domain of digital watermarking and steganography have been developed mostly in isolation, or lack invisibility or robustness features, and are therefore unsuitable for most data sovereignty use cases. To the best of our knowledge, no publications to date have developed digital watermarking artifacts to strengthen data sovereignty. The following main research question (MRQ) guides our research, which combines both domains:

MRQ: How to strengthen data sovereignty through digital watermarking?

To address this question, we formulate five sub-research questions (RQs), grouped into two main research areas. Since the research design uses DSR [197] as a baseline (see Chapter 3), the set of RQs aims to define the research scope, guide the process, and position our contributions [259]. We apply the guidelines from Recker [216] for scientific research and Thuan, Drechsler, and Antunes [259] for formulating RQs in DSR.

Research Area 1: Data Sovereignty Foundations

The first research area aims to clarify the cluttered field of data sovereignty, as researchers often misinterpret or misuse the concept due to its lack of a foundational conceptualization [20], [202], as summarized in RQ 1.1.

> ***RQ 1.1:*** *How can we define, differentiate, and conceptualize data sovereignty?*

Paper I [100] addresses the issues by differentiating data sovereignty from the related terms of digital and technological sovereignty through a structured literature review. Paper III [271] builds upon it by developing a conceptual model for data sovereignty based on publications from research and practice. After defining, differentiating, and conceptualizing the concept, RQ 1.2 emerged to identify current technical challenges in practice to strengthen data sovereignty and address the aforementioned MRQ.

> ***RQ 1.2:*** *Which specific technical challenges for data sovereignty exist?*

Paper II [95] uses semi-structured interviews based on Grounded Theory to collect these challenges from industry experts. The results show a lack of organizational, technical, and personal aspects for implementing data sovereignty. While this work focuses on the technical aspects, these challenges align with findings from related work in [290] and motivate the second research area: the design and development of a solution.

Research Area 2: Digital Watermarking Design & Development

The second research area builds on the idea identified in Paper II [95] of using digital watermarking to enhance data sovereignty. RQ 2.1 considers the gathered input and focuses on the design and implementation of a digital watermarking artifact for text as the most shared and used content type [30], [219].

> ***RQ 2.1:*** *How can we design and implement a digital watermarking artifact for text-based data to ensure imperceptibility and robustness?*

Paper IV [99] introduces the first version of a digital watermarking artifact with an implemented prototype. It can hide any byte-encoded sequence within a cover text while remaining robust across various applications and imperceptible to humans. Moreover, it is important to identify evaluation criteria and conduct experiments to derive enhancements that increase the effectiveness of the artifact, as stated in RQ 2.2.

> ***RQ 2.2:*** *How can we demonstrate and evaluate the effectiveness of a digital watermarking artifact to ensure data sovereignty improvements?*

Paper V [96] starts by demonstrating the solution in an IT landscape through Connector extensions for data spaces. The gathered knowledge has led to an updated, enhanced version of the artifact through an iterative process, presented in Paper VI [94]. It addresses existing drawbacks and evaluates the updated version alongside existing solutions to position the artifact within the current theoretical and practical landscape. Nevertheless, the current solution focuses on pure text-based data, while other data sovereignty use cases exist, leading to RQ 2.3 regarding generalization.

> ***RQ 2.3:*** *How can we ensure the generalizability of a digital watermarking artifact for various use cases using design principles?*

While Paper VI [94] provides recommendations on which current solutions are suitable for specific application areas, there is no one-size-fits-all solution applicable to every use case. Paper VII [288] presents an alternative fragile watermarking technique for comma-separated values (CSV) files. Therefore, this thesis generalizes the gathered design knowledge into DPs in Section 5.1.

1.3 Outline

We structure this cumulative doctoral dissertation on the publication schema guidelines from Gregor and Hevner [85] and the DSR methodology from Peffers et al. [197].

Besides the main text and the published papers in the appendix, this thesis includes an electronic supplement with our source code. We published the core components as a reference implementation in a GitHub repository in [75] for transparency and reproducibility. We applied the core concept of our digital watermarking artifact for a German and international patent [97], [98]. The rest of this thesis is structured as follows.

Chapter 1 motivates the topic of data sovereignty by outlining the need for technical solutions to improve it. It identifies the research gap and gives an overview of our contributions. Section 1.2 presents five research questions, clustered by two research areas. Chapter 2 introduces the background information relevant to this thesis while discussing related work. It focuses on the concept of sovereignty and adjacent forms in Section 2.1. Similarly, Section 2.2 focuses on watermarking, particularly digital watermarking and its adjacent domains, such as steganography. Lastly, Section 2.3 provides an overview of various character encodings and introduces different Unicode whitespace characters, which are essential for understanding the digital watermarking artifact we designed. Chapter 3 lays the methodological foundation by presenting the applied research design. It shows how the RQs and our published papers integrate into the process while laying the foundation for the

remainder of the thesis. Chapter 4 presents the core results, based on the most relevant contributions of our published papers. It is structured around the first five activities following the DSR methodology of Peffers et al. [197], from identifying and motivating the problem (Section 4.1) to defining the solution objectives (Section 4.2), to its design and development (Section 4.3), with demonstrations (Section 4.4) and a final evaluation (Section 4.5). While the core part focuses on the design of our digital text watermarking artifact, we further divide the design and development in Section 4.3 into the watermark embedding process (Section 4.3.1) and its extraction (Section 4.3.2). All sections in these Chapters 2, 3, and 4 conclude their findings in a highlighted summary box at the end of each section. Chapter 5 discusses our results by deriving design knowledge in the form of four DPs in Section 5.1. We discuss the implications for researchers and practitioners in Section 5.2, and the limitations and future research opportunities in Section 5.3. Chapter 6 concludes this cumulative doctoral dissertation by summarizing our research contributions and answering the initially identified research questions. Appendix A presents the seven most important peer-reviewed publications relevant to this work, with a metadata overview in tabular form.

Chapter 2

Background & Related Work

To provide the background information needed to understand this work, we focus on the two concepts of the MRQ: sovereignty and watermarking. Each concept, along with its background information, adjacent domains, related work, and our contribution to highlight the research gap, is presented below. In addition, we introduce character encodings and provide an overview of whitespace characters, both of which are relevant to the design of our IT artifact.

2.1 Sovereignty

The authority of states in international relations shapes the concept of *sovereignty* [141], [199]. Historically, the first occurrences date back to the 13th century [199]. Due to various wars and conflicting political interests, Bodin [33] defined sovereignty as the authority of a ruler to make absolute decisions. A later historical analysis showed that sovereignty is more than just political authority because it combines the three categories of authority, supremacy, and territoriality [199].

Researchers across different domains have defined the concept in various ways over time, viewing it as a broad concept rather than a single definition [199], [289]. Ongoing technological developments shift assets such as documents from the analog to the digital world, raising new issues in the field of sovereignty. While national borders on maps define territorial authority, this becomes increasingly complex in the digital world, where data is transferred over the internet and passes through different countries via cross-border fiber cables [38], [172]. New concepts arise, such as *cyber sovereignty*, often used by Western authors [20], [289], which focus on the shift of territorial sovereignty into transnational cyberspace [38]. Due to global internationalization, states no longer need exclusive knowledge of technological building processes nor extract all needed resources exclusively within their own territorial border – they can build on strong international relationships to enable a self-determined ability to shape technical development and usage, often referred to as *technological sovereignty* [62], [169]. The shift from pure authority to control gives rise to newer forms of sovereignty [38]. Research refers to control over digital assets as *digital*

sovereignty [89], with France as a pioneer in forming the notion [145], and control over data as *data sovereignty* [119]. Various bodies of literature emerged, such as *indigenous data sovereignty*, which focuses on ownership, collection, and application of data from indigenous peoples [254]. It is particularly relevant for machine learning use cases such as federated learning in bioinformatics, because indigenous people constitute around 7% of the global population, yet are often excluded from genomic datasets and included in fewer than 1% [34]. As this thesis focuses on baseline concepts of data sovereignty in information systems and software engineering, we do not examine these indigenous aspects further. Given the increasing popularity over the last decade [100], we discuss data sovereignty in detail below.

Contribution. The distinction between these sovereignty concepts remains unclear. Researchers have identified conflicting definitions [173], showing that the concept's relationship is not well understood [48]. We address this question in our Paper I [100] through a structured literature review, delimiting the most commonly used concepts: data sovereignty, digital sovereignty, and technological sovereignty. Table 2.1 provides a summarized overview of the different concepts of sovereignty.

Table 2.1 Overview of Sovereignty Concepts

Concept	Description
Sovereignty	Having control and executive authority within a specific territorial region, typically a state [33], [141], [199].
Cyber Sovereignty	Having control within the transnational cyberspace [20], [38].
Technological Sovereignty	Having control over technologies and resources on an international relationship level [62], [169].
Digital Sovereignty	Having control over digital technologies, like software, processes, hardware, infrastructure, and the freedom of their selection on a political and economic level [74], [89], [124].
Data Sovereignty	Having control over data in a self-determined way on an individual level [21], [108], [119], [152].
Indigenous Data Sovereignty	Indigenous people, having control over the collection, application, and ownership of their own data [254].

Data Sovereignty

The ability of a natural person or company to control their own data is commonly summarized as data sovereignty [119], [183], [227]. The origin and reasons for the awareness are often attributed to the Snowden revelations and the USA Patriot Act in 2001, while the latter allows US government employees to access data stored on American servers as a response to the terrorist attacks on September 11, 2001 [89], [202], [264]. After that, international discussions arose, especially within the European Union, on how to regulate and retain control over data when using international cloud providers [32], [89]. The concept of data sovereignty emerged and was adopted over time, while authors see parallels to *autonomy* [108] and often use *self-determination* alongside data sovereignty [21], [52], [119], [183], [267]. Nevertheless, various simultaneous discussions began across politics, research, and practice, often based on different viewpoints. Upon closer inspection, these viewpoints typically fall into three main perspectives: social, legal, and technical (see Figure 2.1).

Social Perspective. Keeping control over data is challenging in our interconnected society. Our motivated shift from the analog to the digital world through digitization and digitalization (see Chapter 1) explains the need for more trust, control, and protection [251]. While the origins of sovereignty relate to the authority and control within a specific territorial border as introduced above [33], [199], Polatin-Reuben and Wright [202] provide an example of regionally focused data sovereignty of the BRICS countries (Brazil, Russia, India, China, and South Africa). Compared to well-defined borders between countries and states on maps, borders in the digital realm, with decentralized structures and distributed systems, are complex [289]. While practitioners apply existing concepts of sovereignty in their own closed systems, it becomes increasingly challenging to remain data sovereign in distributed systems such as the cloud [21], [67]. Researchers discuss those sovereignty challenges in cyberspace as *data localization* [188], [255]. Since the transition from data sovereignty to the adjacent domain of digital sovereignty is fluid, Couture and Toupin [48] and Tan, Chi, and Lam [251] started to investigate both concepts but missed the relation to the other concepts of sovereignty. Aydin and Bensghir [18] initiated a conceptual discussion of *digital data sovereignty*, and Moschko, Blazevic, and Piller [181] defined *organizational data sovereignty*, but the combined concept did not find application later. Only Hummel et al. [108] reviewed adjacent concepts by analyzing 341 publications on sovereignty, but from a systematic theological perspective rather than information systems and software engineering, and they omitted technological sovereignty.

Legal Perspective. Legal regulations, such as the European General Data Protection Regulation (GDPR) or the California Consumer Privacy Act (CCPA), govern the protection of individual data [17]. Both regulations are closely related to the same data sovereignty principles, which enable negotiations to reach agreements, whereas pure *data privacy* focuses solely on protection [119] as a subset of data sovereignty [194]. Besides the GDPR and CCPA, Ryan, Gürtler, and Bogucki [223] presented an overview of existing acts and directives in Europe's legal policy framework, highlighting the field's complexity. Singi et al. [244] propose a governance framework based on knowledge graphs to help organizations identify the regulations they must comply with. The existing reviews in [17], [223], [283] focus on the legal perspective of data sovereignty, while Pohle and Thiel [201] discuss digital sovereignty. Data sovereignty often divides into two main aspects: jurisdiction, supported by legal regulations, and individual control over the data itself [22]. This split leads to several legal and technical issues [264], which we discuss in more detail in the next technical perspective.

Technical Perspective. From a technical perspective, retaining control in dynamic networks of different actors – so-called data ecosystems – is paramount [225], as a closer investigation reveals existing architectures, concepts, and organizations related to data sovereignty [290]. The IDSA, as a non-profit organization, enables companies to build *data spaces* as a decentralized data sharing concept, with data sovereignty as one of its core principles [183]. Other related international projects are the Connected Industries Open Framework from Japan or the Trusted Industrial Data Matrix from China [253]. Besides those projects, solutions like decentralized identities, access control, and policy-compliant computations are examples of existing technologies [66], while others are discussed by Tao, Yang, and Ge [253]. Moreover, researchers have built architectures to enforce sovereignty or enhance data usage control [73], [128], [182], [227], [230], [290], and they derived DPs from them [190], [230]. Others build solutions based on blockchain and distributed ledger technologies [31], [42], [153]. Until today, we have seen different example approaches, aiming to enhance data sovereignty, such as in artificial intelligence (AI) pipelines [12], for mobility and smart city services [15], [68], for manufacturing use cases [82], [150], [168], [217], for patient data in healthcare [16], [31], [200], or for Internet of Things (IoT) use cases [160], [186], [209]. However, researchers only briefly touch on existing challenges and hurdles. Schmidt et al. [231] derived challenges regarding data exchange based on a literature review and an interview study. Further, Moschko, Blazevic, and Piller [181] conducted another interview study. Their findings highlight the need for more technical solutions to support data sovereignty [181], the main goal of this work.

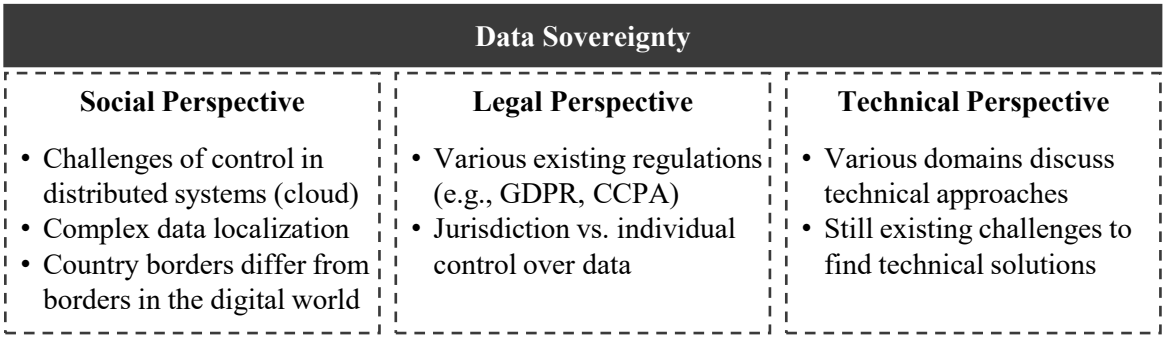


Fig. 2.1 Data Sovereignty Perspectives

Contribution. In the literature, authors use, understand, and interpret data sovereignty differently, which reflects its presence across various domains. Table 2.2 provides an overview of existing definitions, including significant differences and similarities. Comparing these existing definitions of data sovereignty reveals diverse perspectives across domains. A strong legal focus is found in [60], [109], [128], with the ethical perspective discussed in [81], whereas in [21], [119] the focus lies on a self-determined perspective [271]. Several researchers have identified no explicit agreement and a lack of a uniform definition and conceptualization for data sovereignty [20], [167], [171], [173], [202]. In our Paper I [100], we differentiate between adjacent concepts, and in Paper III [271], we develop a conceptual model for data sovereignty (see Figure 4.3) that addresses the gap identified in RQ 1.1. After the publication of our results, we have continued to see strong interest in the subject area, with subsequent reviews and conceptualizations in [1], [22], [194], [223] that build on our work. One example is the review by Pampus and Heisel [194], which used our final literature set of Paper I [100] as input to their review process to distinguish data sovereignty from privacy. Furthermore, we demonstrate the need for more technical solutions to increase data sovereignty through our interview study in Paper II [95], subsequently confirmed by a further interview study in [181].

In summary, sovereignty is a centuries-old political concept regarding authority, power, and political territoriality [199]. Due to the shift into the digital world, new specializations have arisen, such as data sovereignty, which focuses on the control and self-determination of data [119]. The characteristics of data sovereignty are complex and vary across social, legal, and technical perspectives. However, a clear delimitation and conceptualization are missing, and various technical challenges remain unresolved, underscoring the need for more data sovereignty solutions and motivating the RQs addressed in this thesis.

Table 2.2 Existing Definitions of Data Sovereignty

Definition	Year
"The term 'data sovereignty', while lacking a firm definition, refers to a spectrum of approaches adopted by different states to control data generated in or passing through national internet infrastructure. It can be understood as a subset of cyber sovereignty, defined as the subjugation of the cyber domain to local jurisdictions." [202, p. 1]	2014
"Data sovereignty, understood as the responsible shaping of informational freedom, in a manner appropriate to the risks and opportunities presented by big data, is the central ethical and legal goal in confronting the challenges and opportunities presented by big data." [81, p. 30]	2017
"Data sovereignty refers to the self-determination of individuals and organizations with regard to the use of their data." [119, p. 550]	2019
"Consequently, the data sovereignty concept arises, which is defined as the ability of the data owner to decide itself how to share and use its data." [227, p. 101]	2019
"Data sovereignty is the concept that data is subject to laws and regulations of a particular nation." [60, p. 256]	2020
"Our understanding of data sovereignty is the ability to formulate self-defined data-usage rules, influence and trace the data/information flows while being free in the decision of (not) sharing data and migrating data whenever and wherever it is desirable." [152, p. 9]	2021
"Self-determination how, when and at what price others (across the value chain) may use my data" [21, p. 10]	2021
"Data sovereignty is the capability of a natural person or corporate entity for exclusive selfdetermination with regard to its economic data goods." [183, p. 27]	2021
"According to data sovereignty, data is subject to the laws of the country in which it is collected, and the constraints of the data provider who may define how the data can be used, in what context, and by whom, among others." [128, p. 33]	2023
"Ideally, [data sovereignty] provides a legal framework that enables users to effectively control their digital traces. Data sovereignty asserts the rights of nations or regions to govern the flow and management of data within their borders, thereby asserting their sovereignty in the digital realm." [109, p. 6]	2024
"[D]ata sovereignty describes the ability of a data sharing participant (individual or organization) to take actions in the data sharing process autonomously and self-determined." [194, p. 16]	2025

Updated version based on Paper III [271]

2.2 Watermarking

Several security measures protect the integrity of banknotes, including specific watermarks that are visible only at a particular viewing angle or under specific lighting. Those watermarks are mostly invisible to users in typical usage scenarios, while remaining robust and conveying information such as the authenticity of the bank note [50]. Historically, records indicate the first watermarks on paper about 1282 in Italy [50]. Over time, with the global shift toward the digital world in the last century, digital watermarks emerged, differing from traditional watermarks in their focus on protecting digital assets [177]. They span across various multimedia types, as shown by the first patent application for watermarking music in 1954 [50] and the first ideas in text watermarking in 1995 [35], [36]. During the early 90s, researchers published fewer than 10 papers per year on digital watermarking, until the number increased to over 100 in 1998 [198]. To date, over 10 000 articles have been indexed in the IEEE Xplore database, with over 100 000 results on Google Scholar[1]. Due to the increasing interest, various reviews have emerged, structuring the field. They have focused on specific aspects, such as different covers, including watermarking images [26], [72], [234], [274], and watermarking non-media data, such as time-series, sequences, or data streams [245]. Others have focused on their applications, such as IoT and cyber-physical systems [71], multimedia and databases [143], or natural language watermarking [261].

One special form of watermarking is *fingerprinting*, which aims to hide a specific identification mark, such as a serial number, as a label inside a cover [184], [198]. Fingerprinting protects the intellectual property of digital content by allowing the data owner to trace leaks and identify breaches of contracts or license agreements, since it embeds information about the buyer or user in the data [198], [286].

Differentiating existing watermarking techniques is crucial to understanding their similarities and differences. In the following, we introduce a set of criteria and categories commonly used in related work. Figure 2.2 summarizes the current research status based on existing categories and classifications [177], [198], [210], [243].

- *Cover Medium*: Defines the data type or carrier where the watermark is to be embedded, such as text or specific file types like image, audio, or video [11], [198]. Besides the generic types, more specialized covers, such as network protocols [208] or databases [246], exist.

- *Imperceptibility*: Focuses on the visibility of the watermark during viewing and analysis. While a visible watermark clearly displays a copyright notice, an invisible watermark hides inside the cover [11], [177], [210].

[1] Our search was conducted on 12th August 2025 by using “digital watermarking” as a keyword.

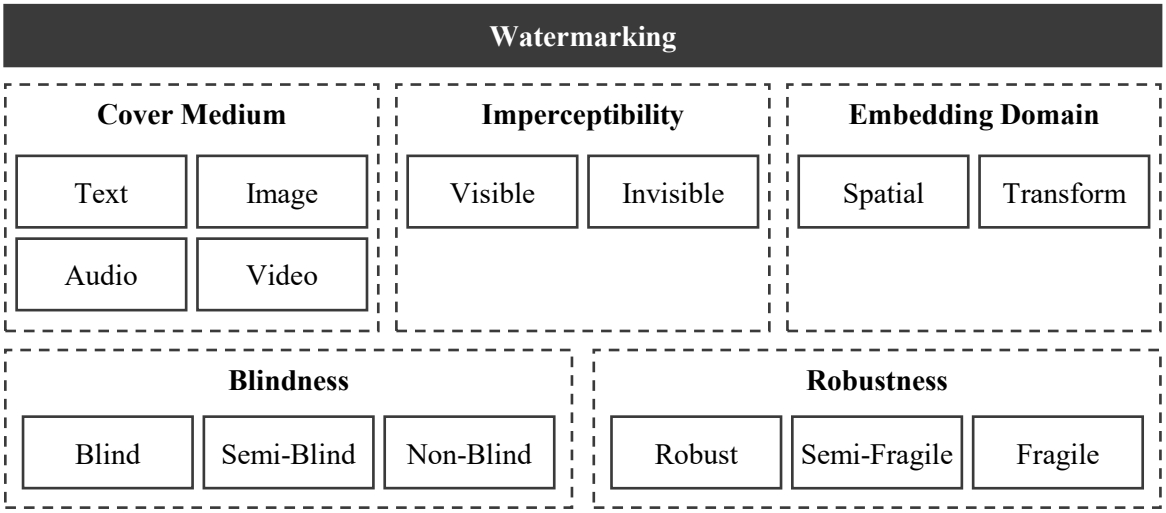


Fig. 2.2 Watermarking Classification

- *Embedding Domain*: Spatial watermarking techniques, such as least significant bit (LSB) methods, directly embed the watermark within the cover through manipulations, while transformation techniques use cover transformations for embedding [210], [243].
- *Blindness*: A blind watermarking technique is independent of the original unwatermarked cover, a semi-blind technique partly depends on the original content, while a non-blind technique needs the original cover and watermark for the extraction process [126], [157], [210], [246].
- *Robustness*: A robust watermark cannot be removed from the cover without destroying it, while a fragile watermark is directly destroyed when modifying the cover, with a semi-fragile watermark lying in between, being robust to specific operations like reformatting but being destroyed under malicious attacks [157], [198], [243].

Due to the wide range of application scenarios and a long history, the literature contains multiple definitions and concepts in the field of watermarking. Table 2.3 presents an overview of existing definitions of digital watermarking before we introduce adjacent domains in the upcoming section.

Table 2.3 Existing Definitions of Digital Watermarking

Definition	Year
"In digital watermarking, relevant information is embedded in an imperceptive way into a digital document. The embedded information is called a watermark." [207, p. 19]	2006
"A digital watermark can be described as a visible or an invisible, preferably the latter, identification code that permanently is embedded in the data." [116, p. 230]	2009
"The definition of digital watermarking is to protect digital data against illegal copying and other attacks through inserting a special digital watermark." [11, p. 6370]	2016
"Digital watermarking aims to hide a message (usually encoded in some form) into a digital signal without disturbing the signal itself." [177, p. 84]	2017
"[D]igital watermarking technology is a typical information hiding method, which covers text, image and video." [210, p. 1311]	2023

Updated version based on Paper IV [99] and Paper VI [94]

In summary, watermarking aims to protect assets [11]. It started with physical assets such as paper or bank notes [50] and moved to the digital world for cover media such as images, audio, video, or text [11], [198]. Existing solutions differ in their focus on imperceptibility, embedding domain, blindness, and robustness [210]. These characteristics are the basis for the subsequent development of our digital watermarking artifact.

2.2.1 Adjacent Domains

This section introduces adjacent domains to watermarking and clarifies their differences and relationships (see Figure 2.3).

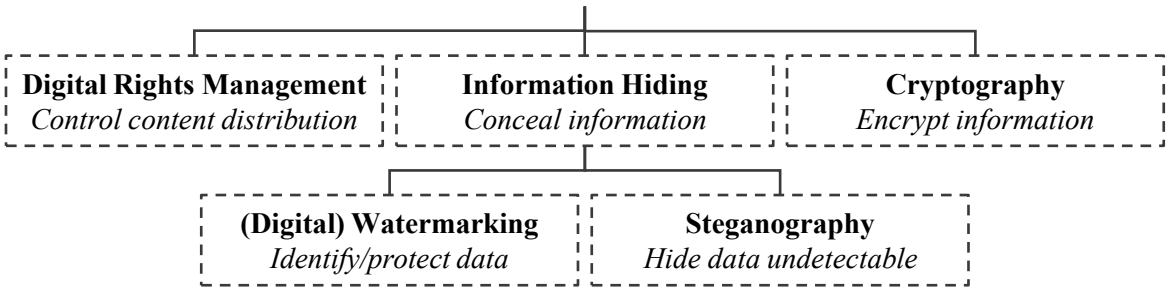
Steganography. Similar to watermarking, steganography aims to hide data inside a cover medium such as an image, audio, video, or text [59], [142], [185]. The origin of steganography comes from the Greek words 'steganos', meaning 'covered', and 'graphia', meaning 'writing' [11], [49], [50], [284]. Different methods have arisen over time, starting historically with Greek tablets covered with wax, people with secret messages tattooed on their heads that were hidden by regrowing hair, to invisible inks used in World War II [49], [122]. Due to the shift in the digital world, electronic steganography techniques use stego-keys to control the data-hiding process so that an unwanted third party cannot detect or recover the

original message [198]. Many techniques emerged, leading to taxonomies that structure the diverse landscape, whereas text steganography or linguistic steganography focuses on textual documents or messages as cover media [277]. In comparison to watermarking, steganography focuses more on concealing data in an undetectable way [11], "so that potential monitors do not even know that a message is being sent" [125, p. 1]. It is often used for secret communication between two parties to prevent detection by a man-in-the-middle [198]. An early motivating example is the *prisoners problem* introduced by Simmons [239]. It is about two people, arrested in separate cells in the same prison, planning an escape. Since a warden analyzes all communication, the prisoners need to find a way to communicate secretly without getting noticed.

Cryptography. The use of encryption techniques to protect data is often referred to as cryptography [185]. While watermarking and steganography embed a watermark directly inside a cover, cryptography describes procedures of changing the state between encrypted and decrypted data. These procedures are mostly substitutions or permutations to transform data into a secure, unreadable format [218], [219]. In the case of text, plain text converts into a cipher text [3]. Thus, it differs from the other domains because it focuses on protecting data during transit and does not persist in the cover after decryption [50]. Typically, the sender, receiver, and an attacker in the middle are aware that the data is encrypted [125].

Information Hiding. The majority of researchers classify watermarking and steganography under the umbrella concept of information hiding [3], [198], [203], while some also classify cryptography in this category [11]. It is defined as the concealment of a watermark inside a cover to enable security use cases such as data integrity, secret communication, or content authentication [8]. Its origin dates back to the first documentation in 1972 [203], while Petitcolas, Anderson, and Kuhn [198] published a comprehensive survey in 1999, analyzing the past on a broader scale across the entire information hiding domain.

Digital Rights Management. Another concept having a similar goal to watermarking is digital rights management, which protects digital assets when being shared. It focuses on controlling the usage and distribution of digital content, such as limiting access or managing usage rights across different types [159]. It is best known in the private sector for protected CDs or DVDs that are only playable on specific players [159]. Thus, the purpose of digital rights management is to prevent piracy, illegal copying of purchased content, and unauthorized redistribution [177].



Based on Paper IV [99]

Fig. 2.3 Interrelationship between Domains

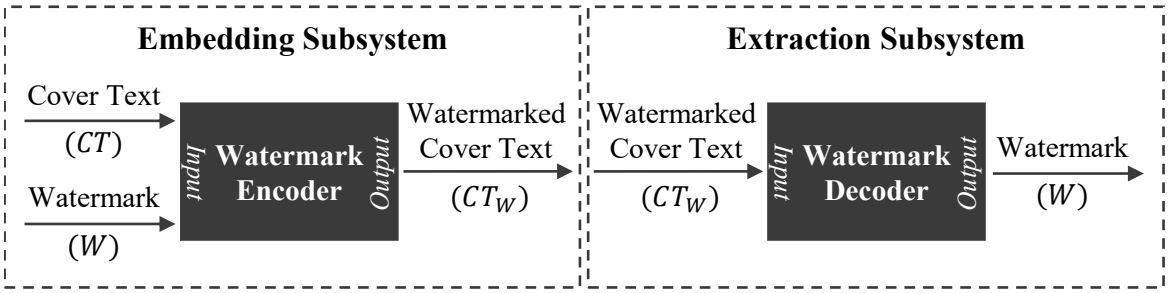
Contribution. Based on the MRQ, this thesis aims to support data sovereignty by enabling interoperable control over data. Thus, it focuses on the watermarking domain under the umbrella of the information hiding category in the solution space, because steganography has a stronger focus on secret communications [198], cryptography on changing data in a decrypted format [185], and digital rights management lacks interoperability due to the focus on specific systems [159]. We focus on text as a cover medium in digital watermarking to encourage protection.

In summary, adjacent domains include digital rights management (for controlling content distribution), cryptography (for information encryption), steganography (for hiding data undetectably), and information hiding (for information concealment) [49]. This thesis builds upon digital watermarking under the umbrella of information hiding because it aims to protect data, in our case, text data.

2.2.2 Digital Text Watermarking

Compared to the long history of watermarking, the first ideas for watermarking text documents emerged only about 30 years ago [35], [36]. These first ideas presented by Brassil et al. [35] still have a very similar motivation today, by creating identification techniques to protect documents from unauthorized distribution. Those text watermarking systems consist of a watermark *embedding* and a watermark *extraction* subsystem (see Figure 2.4). Based on related work [6], [50], [177], the encoder of the embedding subsystem uses a *watermark* $\mathcal{W}$ and a *cover text* CT as an input to transform it into a *watermarked cover text* $CT_{\mathcal{W}}$ as an output. The decoder part of the extraction subsystem is the reverse variant, which uses the watermarked cover $CT_{\mathcal{W}}$ as input to extract the original watermark $\mathcal{W}$, also known as *multi-bit* watermarking [157]. In comparison, a *zero-bit* extraction subsystem only detects

whether a $CT_{\mathcal{W}}$ contains a watermark but cannot extract the original watermark $\mathcal{W}$ [157]. Since embedding and extraction are two separate, interoperable subsystems, they operate independently.



Based on Ahvanooey et al. [6], Cox [50], and Mohanty et al. [177]

Fig. 2.4 Text Watermarking System

When developing such embedding subsystems, researchers distinguish between different techniques for embedding a watermark within a text as a cover medium. No uniform classification exists for these techniques, but common patterns indicate how they are categorized. Figure 2.5 provides a summarized overview of these techniques, introduced below.

Linguistic-based. These techniques focus on the natural language for embedding a watermark [5], [8], [23], [126], [135], [142], [262]. Early techniques classified as *semantic methods* or *linguistic transformations* replace synonyms for watermark embedding [26], [261]. A simple technique could use a synonym encoding table to embed a watermark. For example, the cover text "The people are very smart!" is watermarked by replacing the word 'smart' with the synonym 'intelligent' to encode a '0' or replacing it with the word 'clever' to encode a '1'. Liu et al. [158] further divide the category into *lexical-based watermarking*, for synonym replacement methods, and *syntactic-based watermarking*, focusing on changes in the syntax of the text. Some of the latest techniques use language models to enhance synonym substitution [146].

Format-based. Language-independent techniques use changes to the text format rather than content changes to embed the watermark [23], [135], [158], [286]. *Alignment modification* techniques use shifts between lines and words to insert them [142], [262]. A simple technique could reduce the space between two words to encode a '0' or increase the space to encode a '1'. Other techniques, such as the *structural* method, use formatting options, including different font types and colors [282], to embed the watermark [5], [8]. Unfortunately, those techniques apply to *text* documents (e.g., HTML, PDF, Word documents), but not to

plain text (e.g., UTF-8 without formatting), because the shifting of lines, words, and features requires changes in the display of text rather than being encoded directly in the text itself. Using a specific editor is necessary for these manipulations, as copying the watermarked cover into any plain-text editor often destroys the watermark due to automatic formatting normalization. In this context, Bender et al. [26] presented one of the first extensive reviews of different cover types, also focusing on text and directly addressing the copying limitations of Brassil et al. [35] because "[s]oft-copy text is in many ways the most difficult place to hide data" [26, p. 332]. As a solution for plain text, specific *white spacing* or *open space* methods are also classified under this format-based category, since they make use of adding or manipulating whitespaces at various positions [26], [142], [260], [262]. We classify the digital watermarking artifact presented in this work as a format-based white spacing technique.

Generation-based. These techniques, often called *coverless* [8] or *random & statistical generation* [23], [135], mainly differ from the other categories by generating a watermark text rather than modifying an input cover text [158]. A simple technique could generate a short sentence with fewer than five words to encode a '0', like "The people are smart!" or a longer sentence with more than five words to encode a '1', like "The people here are very smart!". Those techniques are of great interest in the domain of watermarking in LLMs, supported by new regulations. The European Union published the AI Act, which requires providers of LLM systems to implement transparency measures to identify whether a text is generated by a machine or written by a human, with watermarks being a possible solution [69]. In this context, Liu et al. [158] conducted an extensive survey on text watermarking systems for LLMs, and Li, Wang, and Barni [157] surveyed deep neural network watermarking. Most of the presented techniques directly integrate the watermarking mechanism into the LLM-specific aspect, e.g., during the generation of logits, token sampling, or training [158]. One example approach proposed by Kirchenbauer et al. [134] separates the model vocabulary into a green and a red list based on a computed hash of the prompt token, and uses it for text generation without retraining the original model. Statistical null-hypothesis tests help detect a watermarked text, even if only a portion of the generated text is checked [134]. Another example is Google's watermarking engine SynthID, which uses tournament sampling for text generation, as introduced in their Nature article [54].

Contribution. Existing approaches and reviews have specialized focus areas, ranging from different covers like image, audio, or video to LLM-focused approaches, or restrictions to specific languages or characters. This work focuses on designing a format-based watermarking solution for the most widely used content type of plain text, to support data sovereignty

in line with the MRQ. The existing reviews in [4], [158], [166], [286] and evaluations in [5], [135], extended by our own reviews, help identify the most relevant methods, which we introduce below.

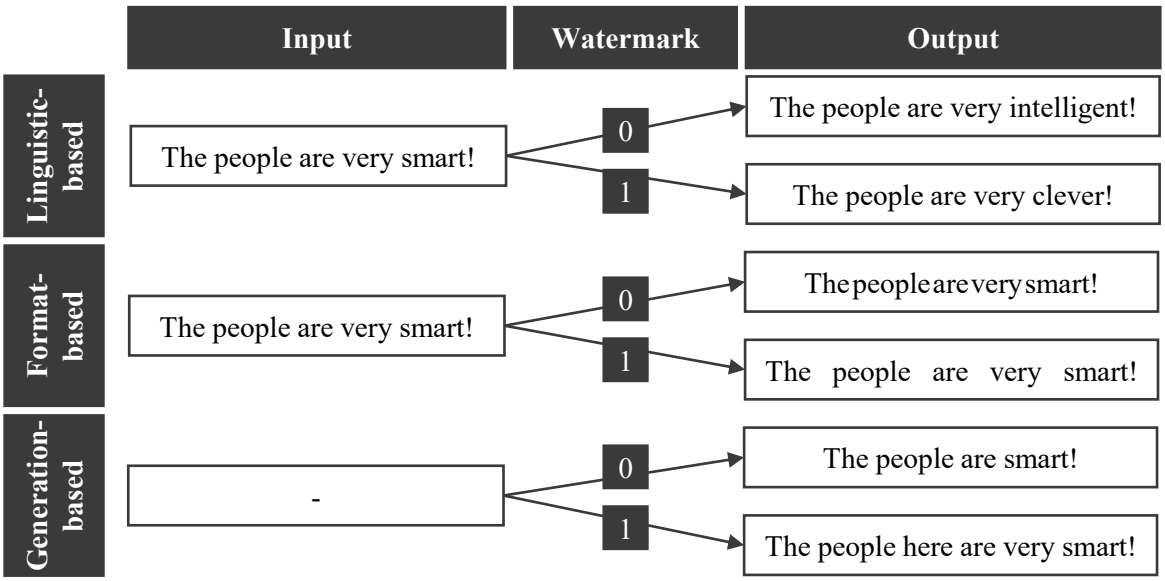


Fig. 2.5 Text Watermarking Classification by Example

In summary, digital text watermarking systems use an embedding mechanism to embed a watermark within a cover text, and an extraction mechanism to read and identify watermarks [6], [50], [177]. Existing solutions differ in their embedding strategy, classified as linguistic-based when using the language (e.g., synonym substitution), format-based when using structural components (e.g., font changes), or generation-based when creating new watermarked texts (e.g., statistical generation) [203]. We classify the watermarking artifact developed in this thesis as a format-based method.

2.2.3 Related Text-based Methods [2]

Researchers have developed various techniques for using text as a cover medium. After presenting existing reviews and techniques, this section introduces the nine most related methods for this thesis and summarizes them in Table 2.4. Due to the seamless transitions between information hiding domains, researchers classify their algorithms as either watermarking or steganography. To maintain uniform terminology throughout this work, we consistently use *watermark* for any data or secret message, and *cover text* for any carrier or corpus.

[2] This section contains verbatim content previously published in Paper VI [94].

Besides the previously introduced initial attempts in [26], [35], Jalil and Mirza [116] presented one of the first reviews in the text watermarking domain and identified in 2009 that "the amount of work done on text watermarking is very limited and specific" [116, p. 233]. Since then, additional techniques have arisen, which have been reviewed for watermarking [11], [126], [286], steganography [4], [142], [166], [262], or the more general information hiding domain [8], [30]. The performance of these existing techniques is of great importance. The advantages and limitations have been evaluated by Ahvanooey et al. [5], focusing on the two categories of existing linguistic-based and format-based techniques. One of the latest evaluation comparisons was conducted by Knöchel and Karius [135], comparing the standard metrics of capacity, imperceptibility, robustness, and complexity.

Other text watermarking approaches specialize in languages or file formats or benefit from specific extensions and are therefore not part of the nine most related techniques. Some authors propose techniques specialized for securing web pages [7] or news articles, combined with blockchain verification [29]. The ANiTW technique proposed by Ahvanooey et al. [6] and later extended by Rofiatunnajah and Barmawi [221] uses invisible signatures as watermarks. Sato et al. [228] propose three text watermarking techniques in their Easymark family focusing on very easy implementation but with limited robustness. Ray et al. [215] embed a secret message using two different whitespaces, but it relies on a special encryption and is related to steganography. Rafat and Sher [212] use whitespace replacement with four characters, but require agreement on a stego key between the sender and receiver. Butler [37] proposed a combination of Unicode variation selectors to hide data within an emoji, as an alternative to plain text. Other specialized techniques as in [184] hide data inside a stego key instead of the cover or focus on the characters or constructs of specific languages [126], as in [165] for English, in [250], [275], [287] for Chinese and in [10], [88], [131] for Arabic. Specializations on file type exist in [132], [154], [229], [260] for PDF, or in [43], [115], [175], [233] for HTML instead of plain text as aimed here.

Since this thesis introduces a new format-based text watermarking artifact to support data sovereignty, we reviewed related work, as outlined in Paper VI [94], and identified a set of nine algorithms introduced in the remainder of this section (see Table 2.4). They all work on unformatted plain text by inserting or substituting specific characters and are therefore classified as *format-based* techniques (see Section 2.2.2) [135]. Some of these identified methods are presented only in academic papers, others only as published reference implementations, and yet others as a mixture of both. Nevertheless, we have implemented all methods and have evaluated them in a benchmark comparison in Section 4.5.

SNOW. One of the oldest whitespace steganography algorithms for ASCII texts is Steganographic Nature of Whitespace (SNOW) [149]. Although the first release dates back to the 20th century, it was made available open-source under the Apache 2.0 license in 2013 [149], with the latest update on GitHub made in 2016 [148]. The source code was written in the C programming language [148], while compiled versions for Java and Windows DOS are available [149]. The embedding process encodes the watermark into tab and space characters and appends it to the cover text, starting with a tab character under the consideration of a predefined line length [147]. Users can optionally enable upstream compression and encryption before the encoding process [147].

UniSpaCh. A well-known algorithm in the field of information hiding for text documents is UniSpaCh, proposed by Por, Wong, and Chee [204]. It is an extended version of WhiteSteg that replaces a single whitespace character between two words or paragraphs with either one or two characters to encode a zero or one [203]. UniSpaCh uses two different methods to embed the watermark in the text. For spaces between words and sentences, regular whitespace characters either remain as they are or are extended by adding a thin, six-per-em, or hair space to encode two bits per embedding location [204]. For end-of-line and inter-paragraph spacings, the remaining space is filled with a combination of hair, six-per-em, punctuation, and thin spaces to encode two bits per character [204].

AITSteg. Ahvanooey et al. [3] proposed a text steganography algorithm for SMS or social media communication. The embedding method creates the watermark by using a Gödel function and the sending/receiving time of the cover message, along with the watermark's length for hashing, and inserts it by replacing every two bits with one zero-width character [3]. Future work, such as CovertSYS [9] published later, extended the baseline concepts as introduced below.

Shiu et al. The data hiding algorithm proposed by Shiu et al. [238] focuses on communication via social media messengers. Due to the narrow width of social media messaging windows, the method relies on a fixed line length and can hide three bits per line of the cover text [238]. After encoding a watermark into a bit stream based on the ASCII mapping, it embeds the first bit by adding a whitespace at the end of a line, changing the length of the line to embed the second bit, and adding a whitespace between two words to embed the third bit [238].

Rizzo et al. A text watermarking algorithm based on replacing Unicode characters with specific confusables, also known as homoglyphs, was initially proposed by Rizzo et al. [218], [220], later extended up to a fine-grain watermarking approach in [219]. The latter method generates a watermark by using a keyed hash function with a watermark and a secret password [219]. Afterward, the watermark is embedded by replacing specific characters with their confusables, or leaving them unchanged to embed one bit in each, and replacing spaces with a set of specific whitespace characters to embed three bits in each [219].

StegCloak. The open-source algorithm StegCloak, published by KuroLabs [144] and described by Mohanasundar [176], is a JavaScript steganography algorithm that is able to hide a watermark within a cover text, with optional password encryption and a Hash-based Message Authentication Code (HMAC). In the embedding process, the watermark is compressed, optionally encrypted, and encoded as a set of zero-width characters to be inserted at a single location after a classical whitespace in the cover text [176].

Lookalikes. Another implementation is the Unicode Lookalikes algorithm by Thompson [258], which is part of the Python package pyUnicodeSteganography. Similar to Rizzo, Bertini, and Montesi [219], the method replaces specific characters with their confusables to encode a watermark inside the cover text [258].

CovertSYS. Ahvanooey et al. [9] presented a multilingual steganography method focusing on short messages in social media networks. As in their previous approach [3], four zero-width characters and a timestamp are used to encode the watermark. Further, a password-based approach using a one-time pad and an XOR operation transforms the watermark into an encrypted bitstream, which is then appended to the cover text [9]. CovertSYS is very similar to the steganography algorithm proposed by Bashir, Li, and Hou [23], which also uses a one-time pad, XOR operations, and zero-width characters, and is thus not analyzed individually in this work to avoid duplication.

Shazzad-Ur-Rahman et al. The data-hiding approach of Shazzad-Ur-Rahman et al. [235] can embed five bits per embeddable location, whereas their updated version [236] can embed six. The main idea of the latter procedure is to encrypt the watermark using AES and convert the resulting binary stream into blocks of six bits [236]. With the help of two lists, specific Unicode characters are replaced with their confusables, and whitespace characters are replaced with a particular combination of smaller whitespace characters to embed the watermark in the cover text [236].

Contribution. All presented approaches either increase the number of characters, or are noticeable by humans, or are not robust in different applications and file formats. To close this gap, the present thesis presents an invisible text watermarking artifact, initially introduced in Paper IV [99] and further developed in Paper VI [94], with the source code published on GitHub [75]. Using certain elements from Rizzo, Bertini, and Montesi [219], our method replaces all whitespaces in a cover text with a specific identified alphabet set of similar-looking Unicode whitespaces. However, it differs from related work due to its imperceptibility and robustness, as shown in the evaluation (see Section 4.5).

In summary, some solutions in the domain of digital watermarking and steganography can invisibly hide information within plain text. Different reviews structure the field, while this section introduces the nine most relevant format-based methods compared to our work. Most of them substitute specific characters or add additional non-visible characters, such as zero-width characters, to embed the watermark. Overall, the presented approaches either increase the number of characters, or are noticeable to humans, or are not robust across different applications and file formats. These nine methods form the basis for our benchmark evaluation in Section 4.5.

Table 2.4 Overview of Related Text-Based Methods

Name	Release	Pub. Type	Techniques
SNOW [147], [148]	Before 1998	Documentation, Software	Appends additional tabs and whitespace characters at the end of the text.
UniSpaCh [204]	2012	Paper	Adds small whitespace characters between words and sentences, and fills up lines and parts between paragraphs with these characters.
AITSteg [3]	2018	Paper	Hides data at the beginning of the cover text by using symmetric-key encoding and a transformation into zero-width characters.
Shiu et al. [238]	2018	Paper	Hides data line-wise by either changing the line length or adding whitespace between words or at the end of the line.
Rizzo et al. [218], [219]	2016/2019	Paper	Replaces whitespace and other Unicode characters with their confusables.
StegCloak [144], [176]	2020	Blog post, Software	Inserts the watermark at one location in the cover text using zero-width characters.
Lookalikes [258]	2021	Software	Replaces a specific set of Latin characters with their Unicode confusables.
CovertSYS [9]	2022	Paper	Adds zero-width characters at the end of the cover text by using the current date and time and a one-time pad.
Shazzad-Ur-Rahman et al. [235], [236]	2021/2023	Paper	Replaces confusables and changes whitespaces to a specific combination of small whitespace characters.

Based on Paper VI [94]

2.3 Character Encoding

Before transferring messages over communication channels, their characters must be encoded into an appropriate representation. One well-known method is Morse code, which maps characters to a set of dashes and dots to transfer them via short and long light flashes or audio sounds of varying length [163]. Today, computers use bits and bytes, but similarly need to encode characters into their binary representations [163]. Before diving deeper, it is important to distinguish between the following terminologies:

- *Character*: A coded representation with a specific meaning, whereas graphic characters have a visual representation [113], [163]. Example: The small Latin character 'a'.
- *Code Point*: A specific integer value, used for internal encoding of characters as numbers [257]. Example: The hexadecimal code point 61.
- *Code Units*: A specific size or length of integers for representation [257]. Example: A code unit of UTF-8 consists of 8 bits.
- *(Coded) Character Set*: A set of rules showing relationships to map a character to its respective bit [113]. Example: The mapping of the small Latin character 'a' to the hexadecimal code point 61.
- *Character Encoding*: Specifies how every code point is expressed by one or more code units [257]. Example: The Unicode Standard defines three encoding forms: UTF-8 (8-bit), UTF-16 (16-bit), and UTF-32 (32-bit).

In practice, developers mostly work with different standardized character encodings described below, while code points or units are primarily conceptual details.

ASCII. One of the oldest standardized encodings still in use is the American Standard Code for Information Interchange (ASCII), approved and published in 1963 by the American Standards Association [14], [163]. It is based on a 7-bit encoding, resulting in $2^7 = 128$ characters, including printable and non-printable control characters [14]. Those printable characters include the most commonly used characters on classical QWERTY keyboards, such as the upper- and lowercase Latin alphabet, numbers, and specific special symbols like punctuation, braces, or the percentage sign. Over time, different variants of ASCII emerged, using eight bits to encode $2^8 = 256$ characters instead of only 128 as in ISO/IEC 8859.

ISO/IEC 8859. A later extension of the 7-bit ASCII encoding is the series of standards from ISO and IEC, which defines an 8-bit encoding, also known as the ISO-8859 family. It extends ASCII by using the same 128 characters to ensure compatibility with existing applications that build on ASCII. The standard is divided into parts, each offering extensions for various languages. The first part, ISO-8859-1, published in 1998, is focused on Latin-1 for western European languages, using the 8th bit to encode characters like 'ä' or '¿' [113], while the newest part, ISO-8859-16, was published in 2001, focusing on south-eastern European languages with characters like 'š' [114].

Unicode. The Unicode Consortium publishes the core specification of the Unicode Standard, currently available in version 17.0, covering 159 801 characters [257]. In this work, we define the set of all Unicode characters as $\mathcal{U} := \{u : u \text{ is a Unicode character}\}$. The standard focuses on character encoding and covers many additional topics, such as line breaking, number formatting, and security [257]. Due to its structure, Unicode is compatible with other encodings because its first 256 code points follow the same arrangement as ISO-8859-1, and thus the first 128 code points of ASCII [257]. It defines the three encoding forms UTF-8, UTF-16, and UTF-32. To be more concrete, Unicode orders its characters into 17 specific groups of 2^{16} code points each, called *planes* [257]. The first plane, 0, is the basic multilingual plane and contains the most relevant characters for modern languages like Latin, African, Asian, and Indonesian, including specific symbols, while planes 4 to 13 are currently empty; other specific characters, like Egyptian hieroglyphs, are located on the other planes [257]. UTF-32 is the simplest one-to-one encoding since it directly represents every code point with a single 32-bit/4-byte code unit, while UTF-16 maps to the basic multilingual plane characters with 2 to 4 bytes, and UTF-8 maps to 1 to 4 bytes, with direct backward compatibility to ASCII [257].

Table 2.5 Character Encoding Example as Hexadecimal Values

Character	Unicode	ASCII	ISO-8859-1	UTF-8	UTF-16	UTF-32
a	U+0061	61	61	61	0061	00000061
ä	U+00E4	N/A	E4	C3 A4	00E4	000000E4
€	U+20AC	N/A	N/A	E2 82 AC	20AC	000020AC
🎓	U+1F393	N/A	N/A	F0 9F 8E 93	D83C DF93	0001F393

Based on American Standards Association [14], ISO/IEC [113], The Unicode Consortium [257]

To illustrate the similarities and differences, Table 2.5 shows the encodings of four characters in five encodings (ASCII, ISO-8859-1, UTF-8, UTF-16, UTF-32). While UTF-8, UTF-16, and UTF-32 support all characters based on the Unicode standard, they all encode the character 'a', while 'ä' is not supported in ASCII, and '€' is not supported in ASCII and ISO-8859-1. UTF-8 shows significant strengths, as it supports the full set of Unicode characters and offers better memory efficiency due to its variable-length encoding based on the fixed size of 8-bit code units. As a result, UTF-8 is seen as a de facto standard, with 98.9% of all websites using it and only 0.9% using ISO-8859-1 [273]. Further, the WHATWG defines UTF-8 as mandatory in their encoding standard [278], while the RFC2277 defines policies for character languages and sets in the international internet and specifies that "[p]rotocols MUST be able to use the UTF-8 charset" [13, p. 2]. Given this prevalence, the methods presented in this thesis focus on Unicode and UTF-8 encoding.

Whitespaces

The Unicode Standard [257] specifies 17 different space characters (see Table 2.6). Let δ be the most common space that is displayed when hitting the space bar on a QWERTY keyboard (U+0020). Most of the other spaces differ in their width depending on the font used, such as the smaller tiny hair space (U+200A) or the larger em quad (U+2003). The 'Width' column in Table 2.6 displays every whitespace character in the *Times New Roman* font between the two arrows, which each whitespace character highlighted in a dark background to show its dimensions. Other spaces, such as the no-break space (U+00A0), are closely related to the default space (U+0020) and share the same width, but differ in their line-breaking behavior [257]. Let s be a space character in the Unicode standard with a real positive width, and let $\mathcal{S} := \{s : s \text{ is a space character} \wedge s \in U\}$ represent the set of these 17 space characters.

Additionally, other spaces like the zero-width space (U+200B) are "although called a "space" in its name, does not actually have any width or visible glyph in display" [257, p. 336] and thus not an element of $\mathcal{S}$. The lower part of Table 2.6 introduces the invisible spaces with zero width, which the Unicode standard classifies as layout control characters [257].

In data hiding and steganography, different techniques use characters like invisible mathematical operators to conceal their information. For example, Mohanasundar [176] uses the invisible times (U+2062), displayed between the 'm' and 'c' in '$E = mc^2$', or the invisible plus (U+2064), displayed between the '3' and '$\frac{1}{4}$' in '$3\frac{1}{4}$' [257]. Since the digital watermarking artifact developed in this thesis builds on regular spaces, other invisible non-space characters are not considered further.

Table 2.6 Unicode Space Overview

Code	Name	Width	UTF-8 Hex	UTF-8 Storage
Regular Space Characters [257, pp. 335–336] (here: $\mathcal{S}$)				
U+0020	Space	→| ←	20	1 Byte
U+00A0	No-break Space	→| ←	C2 A0	2 Bytes
U+1680	Ogham Space Mark	→| ←	E1 9A 80	3 Bytes
U+2000	En Quad	→| ←	E2 80 80	3 Bytes
U+2001	Em Quad	→| ←	E2 80 81	3 Bytes
U+2002	En Space	→| ←	E2 80 82	3 Bytes
U+2003	Em Space	→| ←	E2 80 83	3 Bytes
U+2004	Three-per-em Space	→| ←	E2 80 84	3 Bytes
U+2005	Four-per-em Space	→| ←	E2 80 85	3 Bytes
U+2006	Six-per-em Space	→| ←	E2 80 86	3 Bytes
U+2007	Figure Space	→| ←	E2 80 87	3 Bytes
U+2008	Punctuation Space	→| ←	E2 80 88	3 Bytes
U+2009	Thin Space	→| ←	E2 80 89	3 Bytes
U+200A	Hair Space	→| ←	E2 80 8A	3 Bytes
U+202F	Narrow No-break Space	→| ←	E2 80 AF	3 Bytes
U+205F	Medium Mathematical Space	→| ←	E2 81 9F	3 Bytes
U+3000	Ideographic Space	→| ←	E3 80 80	3 Bytes
Zero Width Spaces/Layout Controls [257, pp. 1137–1139]				
U+2060	Word Joiner	→←	E2 81 A0	3 Bytes
U+FEFF	Zero Width No-break Space	→←	EF BB BF	3 Bytes
U+200B	Zero Width Space	→←	E2 80 8B	3 Bytes
U+200C	Zero Width Non-joiner	→←	E2 80 8C	3 Bytes
U+200D	Zero Width Joiner	→←	E2 80 8D	3 Bytes

Based on The Unicode Consortium [257]

In summary, the ASCII character encoding published in 1963 defines how machines represent a set of 128 characters [14]. It was later extended and is backward compatible with the ISO/IEC 8859 series and UTF-8. The latter is able to encode the full Unicode character set of 159 801 characters [257] and is considered the most widely used standard. This Unicode standard defines various space characters with different characteristics, ranging from larger ones to non-visible zero-width spaces [257]. These spaces serve as the technical foundation for our whitespace-replacement digital text watermarking artifact developed in this thesis.

Chapter 3

Research Design

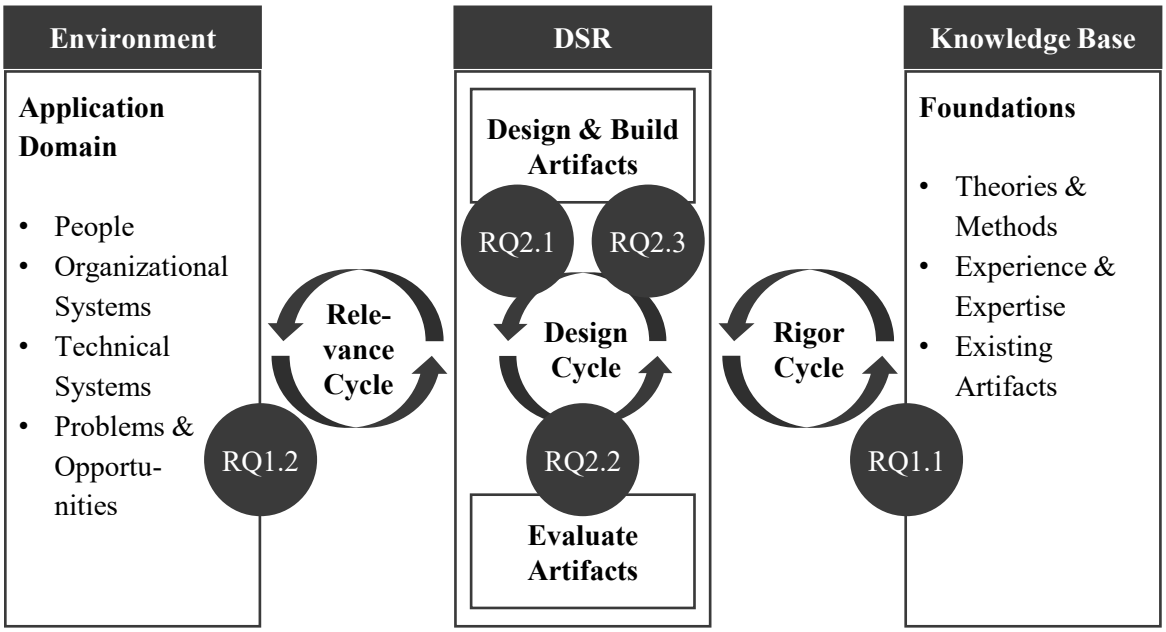
This thesis uses the DSR paradigm as its methodological foundation. It employs DSR to develop a technical solution for the problem domain, because it "creates and evaluates IT artifacts intended to solve identified organizational problems" [103, p. 77]. In our case, the identified problem of missing digital watermarking artifacts to support data sovereignty guides our research as MRQ.

The origin of design science goes back to the 1960s, when Buckminster Fuller introduced the combination of rationalism, science, and technology [87]. Later, March and Smith [170] and Simon [240] laid the foundation for the design and evaluation of artifacts, which, over time, formed DSR [25]. Today, DSR is not a single method but a paradigm with many methods that guide users in building artifacts. Therefore, this chapter outlines the research design by describing the concrete knowledge contribution and process paradigm based on existing guidelines and frameworks [101], [103], [197]. It begins by positioning our RQs and contribution within the context, followed by a detailed explanation of the DSR methodology as outlined by Peffers et al. [197], which structures our work.

3.1 Environment & Contribution

Our goal is to design an artifact that builds on existing knowledge and fits into organizational software environments to address the problem domain outlined in the Introduction (see Chapter 1) [102]. DSR is a well-suited methodology for this objective, often used in information systems and software engineering research [64], [279], as it provides a foundation for solving complex problems through iterative loops of continuous design and evaluation [103].

Environment. Figure 3.1 summarizes how the three-cycle DSR approach helps address our RQs to the application environment and knowledge base [101], [259]. The *rigor cycle* uses existing theories, methods, experience, and artifacts from the knowledge base to integrate them into the design process [101]. At the beginning of our first research area to conceptualize data sovereignty, we conducted a literature review to define, differentiate, and conceptualize data sovereignty for RQ1.1 [100], [271]. Besides the rigorous foundation,



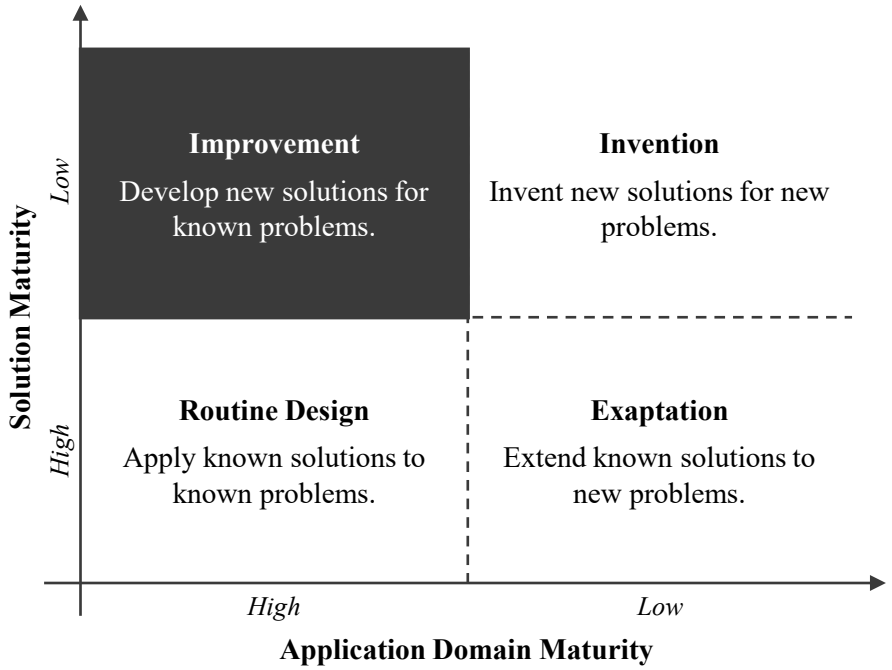
Based on Hevner [101, p. 88]

Fig. 3.1 DSR Cycles with our RQs

the design process actively incorporates the environment through the *relevance cycle*, which considers the application domain, including its requirements, systems, problems, and opportunities [101]. We conducted an interview study with practitioners to identify requirements and challenges related to RQ1.2 [95].

The second research area for designing and building a digital watermarking artifact draws on the theoretical basis of the rigor cycle and the practical insights from the relevance cycle. The *design cycle* iteratively processes both inputs to create rigorous artifacts that researchers continuously evaluate and improve [101]. While RQ2.1 focuses on the design and implementation of an IT watermarking artifact for text-based documents, we made evaluations and improvements addressing RQ2.2 [94], [99], and generalized the findings regarding RQ2.3.

Contribution Type. Researchers apply DSR across various areas beyond information systems and software engineering, such as medicine, management, and biomedical [151], leading to different types of contributions. Gregor and Hevner [85] propose a DSR knowledge contribution framework (see Figure 3.2) that clusters them based on solution and application domain maturity. According to their framework, an *invention* is very rare because it offers a new solution to problems for which "research questions may not even have been raised before" [85, p. 346]. In contrast, an *exaptation* is typical for DSR in information systems because it solves new problems by extending existing solutions. If the solution does not need to



Based on Gregor and Hevner [85, p. 345]

Fig. 3.2 DSR Knowledge Contribution

be extended and applies to a known problem, the contribution is a *routine design*. Our work builds upon existing problems for data sovereignty. It develops a new digital watermarking artifact for text to make an *improvement* contribution, highlighted by the dark background quadrant in Figure 3.2.

Contribution Level. Beyond the contribution type, every research deliverable is further distinguished by its contribution level and domain. DSR researchers hold different views and debate over acceptable abstraction and contribution levels to find a balance between scientific theory and the technological artifact, while every DSR project should reflect on *design knowledge* [24]. Different representations exist for communicating such design knowledge, ranging from artifacts [103], to DPs [39], [40], and up to the most abstract form of design theory [86]. Gregor and Hevner [85] classify those contribution types into three levels (see Figure 3.3), while Lee, Pries-Heje, and Baskerville [155] distinguish between an *instance domain*, with specific solutions for specific problems, and an *abstract domain*, with abstract solutions for abstract problems.

Based on Gregor and Hevner [85], maturity *level 1* includes specific instances for products and processes, while more abstract contributions, such as constructs or methods, are classified as *level 2*, and deeply studied theories as *level 3*. Our research has two contribu-

	Contribution Types	**Example Artifacts**
More abstract, complete, and mature knowledge	**Level 3:** Well developed design theory about embedded phenomena	Design theories (mid-range and grand theories)
↕	**Level 2:** Nascent design theory – knowledge as operational principles/architecture	Constructs, methods, models, design principles, technological rules
More specific, limited, and less mature knowledge	**Level 1:** Situated implementation of artifact	Instantiations (software products or implemented processes)

Based on Gregor and Hevner [85, p. 342]

Fig. 3.3 DSR Contribution Types and Levels

tion levels. First, we designed and implemented an IT artifact for digital watermarking, as a *level 1* knowledge contribution (see Figure 3.3). This digital watermarking artifact belongs to the *instance domain* because it is a specific solution for a specific practical problem [155]. Second, we followed Baskerville et al. [24] by generalizing design knowledge *after* the actual design and evaluation of the IT artifact to address RQ 2.3. We constructed DPs to reach *level 2* as a higher abstraction contribution level (see Figure 3.3). Such abstraction helps create more generalizable knowledge, enabling greater projectability [269] by shifting our contribution from the *instance domain* up to the more general *abstract domain* [155]. It further helps practitioners and theorists by applying the gathered knowledge across different application scenarios [25], [39], [87]. We classify and discuss the resulting contributions as design knowledge at the end of this thesis in Section 5.1 as a standard approach for concrete DSR IT artifacts related to RQ 2.3 [111].

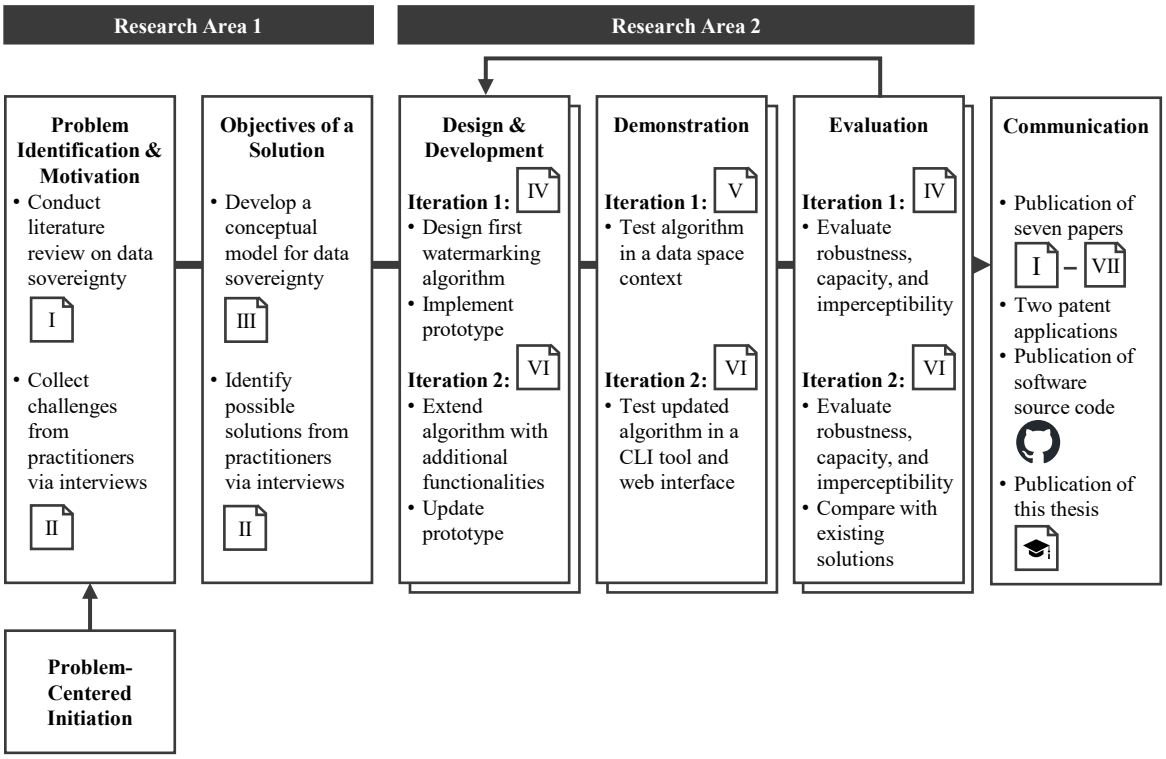
In summary, this thesis uses DSR to design, build, and evaluate a new artifact to address existing problems (contribution [85]) based on knowledge from the rigor cycle and requirements from the relevance cycle [101]. This artifact serves as an instantiated prototype implementation, contributing at *level 1*, while the gathered knowledge abstracts into DPs as a *level 2* contribution [24], [85], [155], addressing RQ 2.1, RQ 2.2, and RQ 2.3.

3.2 Design Science Research Paradigm

Selecting a suitable method for the DSR process is important for guiding rigorous research execution and for communicating results [24], [64]. We follow the guidelines from Venable, Pries-Heje, and Baskerville [266] to select an appropriate research paradigm for structuring our work. Iivari [111] differentiates information systems DSR projects into two groups of strategies. Research under *Strategy 1* builds generic IT artifacts at the meta-level, while research under *Strategy 2* builds concrete IT artifacts for specific problems [111]. As shown in Figure 3.2, the objective of this research is the *improvement* of a known problem [85]. It aims to develop a concrete IT artifact to address the identified practical problem domain of the lack of a technical solution that strengthens data sovereignty, belonging to *Strategy 2* [111]. Based on the different DSR genres described in [87], we classify our work in the *computational genre*, following Rai [214], because we developed a technical solution in the form of an algorithm.

We selected the DSR methodology from Peffers et al. [197], initially introduced as a process model in [196], to structure our research. We did not select one of the two closest alternatives of *Action Design Research* following Sein et al. [232], and *Systems Development Research Methodology* following Nunamaker, Chen, and Purdin [189], since we did not focus on one single client nor aim to create a full design theory as Action Design Research typically does. Instead, we focused on an iterative process for a practical problem rather than a linear one as in the Systems Development Research Methodology. Nevertheless, we created DPs a posteriori in Section 5.1, reflecting RQ 2.3 [24], [178]. Furthermore, we followed and applied the DSR guidelines of Hevner et al. [103] (see Table 3.1) and the principles for DSR projects of vom Brocke et al. [269] (see Table 3.2). Below, we outline how we applied the six activities of the DSR methodology by Peffers et al. [197].

Problem Identification & Motivation. The first activity motivates the DSR process through problematization and identifies a suitable entry point. In our case, we argued for a *problem-centered initiation* [197], as we identified a lack of conceptualization of data sovereignty and existing technical challenges for practical use cases, as outlined in RQ 1.1 and RQ 1.2 for research area one. Paper I [100] establishes and demonstrates the relevance of the problem space by drawing on the literature and highlighting data sovereignty. Based on the second guideline for problem relevance in Hevner et al. [103], the research activity also needs to focus on practitioners to be relevant to their community. Therefore, we additionally conducted semi-structured interviews with experts from different industries in Paper II [95]. We applied Grounded Theory as a methodological foundation, with two coding cycles, to analyze the transcribed interview records.



Based on Peffers et al. [196, p. 93], [197, p. 54]

Fig. 3.4 Applied DSR Process Methodology

Objectives of a Solution. Given the current state of the problem with existing knowledge and solutions, this activity derives the objectives of a solution [197]. In our case, we derived a conceptual model of data sovereignty in Paper III [271] from theory, using a multivocal literature review to include a broad range of literature, both academic and practice-oriented [27], [79]. Additionally, Paper III [271] presents a practical perspective alongside the results from the aforementioned interview study. A closer review and discussion of possible solutions to strengthen data sovereignty revealed digital watermarking as a promising solution space with still-existing open gaps. To ensure validity, we derived and formulated our knowledge in four claims following Larsen et al. [151] and mapped them to the related technical challenges derived from practice and existing watermarking literature from research. These activities concluded the first research area of data sovereignty foundations.

Design & Development. After clearly defining the solution's objective, the next activity is to build the artifact [197]. Our second research area for digital watermarking is based on the results from research area one and therefore integrates into the overall methodology. We designed and implemented a first version of our digital text watermarking artifact as a

Table 3.1 Applied DSR Guidelines

No.	Guideline	Our Application
1	Design as an Artifact	We present an implemented digital text watermarking solution as a library, CLI tool, web interface, and connector extensions (IT artifact) and four DPs.
2	Problem Relevance	We identified a lack of technical solutions for strengthening data sovereignty.
3	Design Evaluation	We conducted an *Experimental – Simulation* evaluation by performing benchmark analysis on the IT artifact and a *Descriptive – Informed Argument &* *Scenario* evaluation by using knowledge base information and scenario demonstrations (data space context).
4	Research Contributions	We made practical contributions through our implemented IT artifact and theoretical contributions in the form of design knowledge represented as DPs.
5	Research Rigor	We used foundational knowledge from existing literature, artifacts, and methods in the domains of data sovereignty and digital watermarking.
6	Design as a Search Process	We apply an iterative process of all six activities with a *problem-centered initiation*, following the DSR methodology of Peffers et al. [197].
7	Communication of Research	We communicated the results through seven peer-reviewed conference and journal papers, two patent applications, the software source code of the IT artifact, and this thesis.

Based on Hevner et al. [103, p. 83]

prototype in the first iteration in Paper IV [99], using the Kotlin programming language to address RQ 2.1. Later, we created an extended and updated version of the artifact in a second iteration in Paper VI [94] to address identified gaps, such as robustness and interoperability.

Demonstration. After initial development, the next activity is to demonstrate that the artifact works as expected to solve the problem [197]. In Paper V [96], we tested the digital text watermarking artifact in a specific data space setup in iteration one. We demonstrated the updated version in iteration two by a CLI tool and a web interface as described in Paper VI [94].

Evaluation. To observe the performance of the artifact, we compare and measure it [197], which is crucial for showing the usefulness of IT artifacts in software engineering [102].

Table 3.2 Applied DSR Principles

No.	Principle	Description	Our Application
1	Positioning	The identified problem domain, solution, and evaluation are presented.	We present the problem in Section 4.1, the solution in Sections 4.2, 4.3, 4.4, and the evaluation in Section 4.5.
2	Grounding	The prior knowledge is clearly stated where the research builds upon, e.g., based on conducted literature reviews.	We conducted literature reviews in Paper I [100] and Paper III [271] and present related work for sovereignty and watermarking in Chapter 2.
3	Aligning	The design process is structured and evolved transparently, e.g., by using a DSR process structure.	We apply the DSR methodology following Peffers et al. [197].
4	Advancing	The extension of prior knowledge is defined in terms of how to position oneself inside the problem and solution space.	We build on related work, e.g., Hummel et al. [108] for data sovereignty and Rizzo, Bertini, and Montesi [219] for text watermarking. Problem and solution space are presented in Figure 6.1.

Based on vom Brocke et al. [269, pp. 531–532]

The specific evaluation technique is highly contingent on the artifact and the domain, as it requires domain-specific knowledge of field-specific metrics [102]. Since our artifact is an algorithm with a reference implementation, we used *benchmarks* due to their key characteristics of high relevance, reproducibility, fairness of test configurations, verifiability, and usability [90]. They are a standard tool for comparing related solutions based on specific characteristics [270], and are also common in digital watermarking [50]. We evaluated the first version of the artifact in iteration one in Paper IV [99] and the updated version in Paper VI [94] in response to RQ 2.2. Following the initial results, which led to a decision to revisit activity three for design and development, we evaluated it and compared it with existing solutions from research and practice. The final evaluation directly maps to the derived claims of our knowledge, verifying their validity and providing evidence [151].

Communication. The last activity communicates the whole process and results, from the initial problem to the final artifact, to audiences such as researchers or practitioners [197]. An appropriate communication consists of describing the identified problem, showing its importance, and explaining how to design and build a solution, rather than presenting the

solution alone [151], [197]. It is essential to consider both dimensions of the application domain's environment, encompassing the relevance cycle and the theoretical foundation of the existing knowledge base from the rigor cycle [101], [104]. Consequently, our communication has four main elements,

(i) we published seven scholarly, peer-reviewed, open-access *papers* (see Appendix A) including a preprint on arXiv (see [93]) [61];

(ii) we filed two *patent* applications due to the high practical usability and novelty;

(iii) we released the *source code on GitHub* in [75] for transparency and reproducibility [61], [104] as a DSR instantiated artifact [85]; and

(iv) we wrote this *cumulative doctoral dissertation*, including further explanations and derived DPs.

One primary goal is to enable researchers and practitioners to use and adopt the findings in future work, even in potential new fields not initially intended for [61], such as medical healthcare documents or legal contracts.

In summary, we structured this problem-centered approach using the DSR methodology following Peffers et al. [197], while considering existing DSR principles [269] and guidelines [103]. It identifies the problems of missing data sovereignty conceptualization and missing digital watermarking artifacts for text. It builds, designs, and develops a technical solution, which is demonstrated and evaluated across two iterations. We communicated the results as papers, patent applications, software source code, and in this thesis.

Chapter 4

Results

We published our findings as conference and journal papers, listed in Appendix A, and present them in an organized, summarized, and further-explained form below. We structure it according to the first five activities of the DSR methodology from Peffers et al. [197]. Since we created the final artifact iteratively, this chapter presents the latest version.

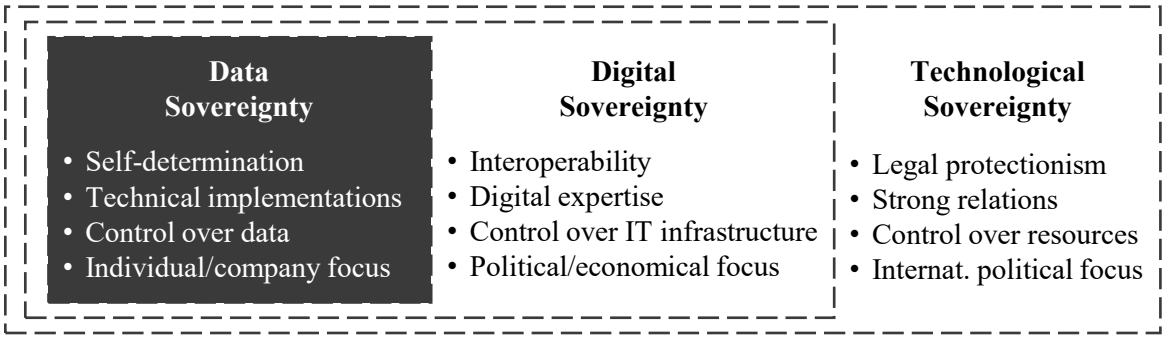
4.1 Problem Identification & Motivation

The findings from Paper I [100] and Paper II [95] form the basis for problem identification and motivation of our research, presented below and structured by the theoretical and practical environment [103].

Theoretical Environment. In the theoretical environment of the knowledge base [103], RQ 1.1 highlights the need for a definition, differentiation, and conceptualization of data sovereignty. Since researchers identified a lack of a missing data sovereignty definition [20], [202], we conducted a structured literature review following Webster and Watson [276] and vom Brocke et al. [268] in Paper I [100]. Due to the high practical relevance of DSR, we included gray literature such as technical reports and political speeches in the review, leading to a *multivocal literature review* [27], [79].

Compared with adjacent concepts such as digital sovereignty and technological sovereignty, *data sovereignty* focuses on self-determination and control over individual and organizational data [21], [108], [119], [152]. The aim and scope of *digital sovereignty* differ, as it focuses on the control of software, processes, hardware, or infrastructure rather than on data and their freedom of selection at the political and economic levels [74], [89], [124]. *Technological sovereignty* is even more abstract because it highlights control over complete technologies and resources at the international relationship level [62], [169]. Since our research aims to solve concrete problems for practitioners, we will focus on data sovereignty, as highlighted in the summary delimitation in Figure 4.1.

In other domains, such as the physical or social sciences, researchers often *replicate* studies to prove their validity – an aspect often missing in information systems [58]. While



Based on Paper I [100]

Fig. 4.1 Delimitation between Data, Digital, and Technological Sovereignty

we conducted our initial literature review of Paper I [100] in April 2022, we replicated the analysis more than three years later in August 2025 for the concept of data sovereignty, using the same search queries. We used the same five databases, IEEE Xplore, AISeL, ProQuest, ACM, and ScienceDirect, by executing a search for the 'data sovereignty' term in the title, abstract, and keywords. We considered all results without date restrictions and removed duplicates. During replication, we identified that the presentation of the search string in Paper I [100] misses further details, as databases like AISeL classify their keywords as subjects or offer an additional peer-reviewed flag. Table 4.1 provides a more detailed overview, showing the same search strings used in 2022 for Paper I [100] and in 2025 for our replicated search as a supplementary correction.

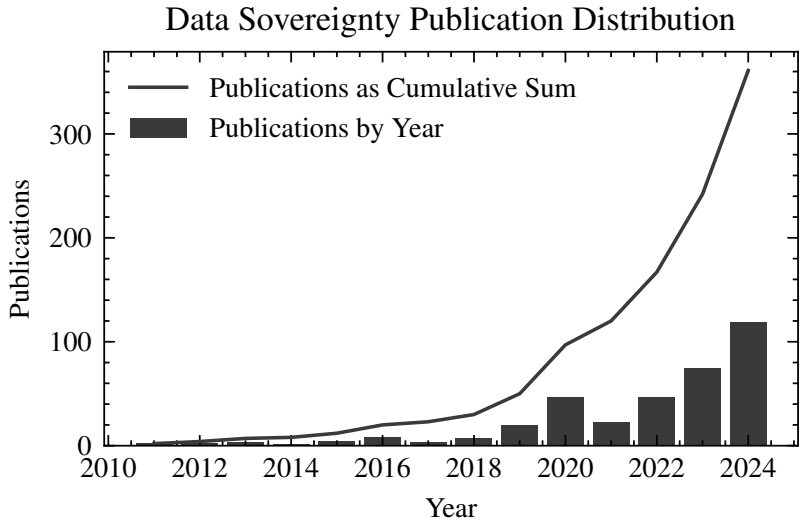
The results showed a steady increase in publication numbers, indicating that data sovereignty still has momentum. The unfiltered total number of publications increased from 142 in April 2022 to 469 in August 2025 (see Table 4.1). We removed 28 entries, mainly duplicates, call-for-papers, or descriptive sections, resulting in a final literature corpus of 441 publications. Figure 4.2 shows the progression over time, clustered by year as bars and the cumulative sum as a line chart. Figure 4.2 ends in 2024, since the search was conducted in late 2025, resulting in 361 considered publications.

Practical Environment. In the practical environment of the application domain [103], RQ 1.2 calls for identifying current challenges in data sovereignty. Since the research activity needs to focus on practitioners to be relevant for their community, we conducted semi-structured expert interviews with 11 practitioners in Paper II [95]. We used Grounded Theory as the methodological foundation following Charmaz [41], utilizing initial coding as a first-cycle method and focus/theoretical coding as a second-cycle method [226], with downstream categorization. The interview participants were industry experts, mostly working in devel-

Table 4.1 Search Results for Literature on Data Sovereignty

Database	Search Query	Results (2022)	Results (2025)
IEEE Xplore	`"Abstract":"data sovereignty" OR "Document Title":"data sovereignty" OR "Author Keywords":"data sovereignty"`	53	158
AISeL	`(abstract:"data sovereignty" OR title:"data sovereignty" OR subject:"data sovereignty") AND` "Peer-reviewed only" checkbox enabled	4	5
ProQuest	`AB("data sovereignty") OR TI("data sovereignty") OR SU("data sovereignty")`	49	160
ACM	`[Title: "data sovereignty"] OR [Abstract: "data sovereignty"] OR [Keywords: "data sovereignty"]`	14	46
ScienceDirect	`tak("data sovereignty")`	22	100
Total		**142**	**469**

Updated version with the results from 2022 of Paper I [100]



Updated version of Paper I [100]

Fig. 4.2 Publication Distribution on Data Sovereignty

opment, IT management, and research and development, with 3 to 40 years of professional experience. We designed the participant selection to be as diverse as possible, including interviewees from start-ups to corporations with headquarters in countries around the world.

The results of Paper II [95] revealed 13 challenges across (i) organizational, (ii) technical, and (iii) personal & emotional aspects. Due to the technical orientation of this thesis, we focus on finding a solution for the four main technical challenges identified in Paper II [95]:

- *Access & Usage Control*: The first technical challenge concerns the enforcement of access and usage control. Based on the interviewees' statements, most existing approaches are only extended access controls. Usage control is more complex with minimal enforcement mechanisms [205].

- *Infrastructure & Landscape*: Current data sovereignty enforcement solutions only work in a limited space of a trusted execution environment [12]. The industry consists of diverse IT architectures and system landscapes. Organizations face challenges in maintaining data sovereignty when sharing data across companies and systems. Based on an interview statement, added value only arises with system-independent data sovereignty.

- *Data Processing Life Cycle*: Data itself has its own life cycle, starting with the creation of a new data asset, through its usage and storage until final deletion [76], [213]. Existing data sovereignty solutions only address individual data life cycle activities, such as when one actor shares data with another. It remains challenging to ensure data sovereignty throughout the entire data life cycle.

- *Identity Management*: The last challenge relates to identification, certification, and verification. This challenge includes questions like: Who can prove that I am really the person or company I claim to be?

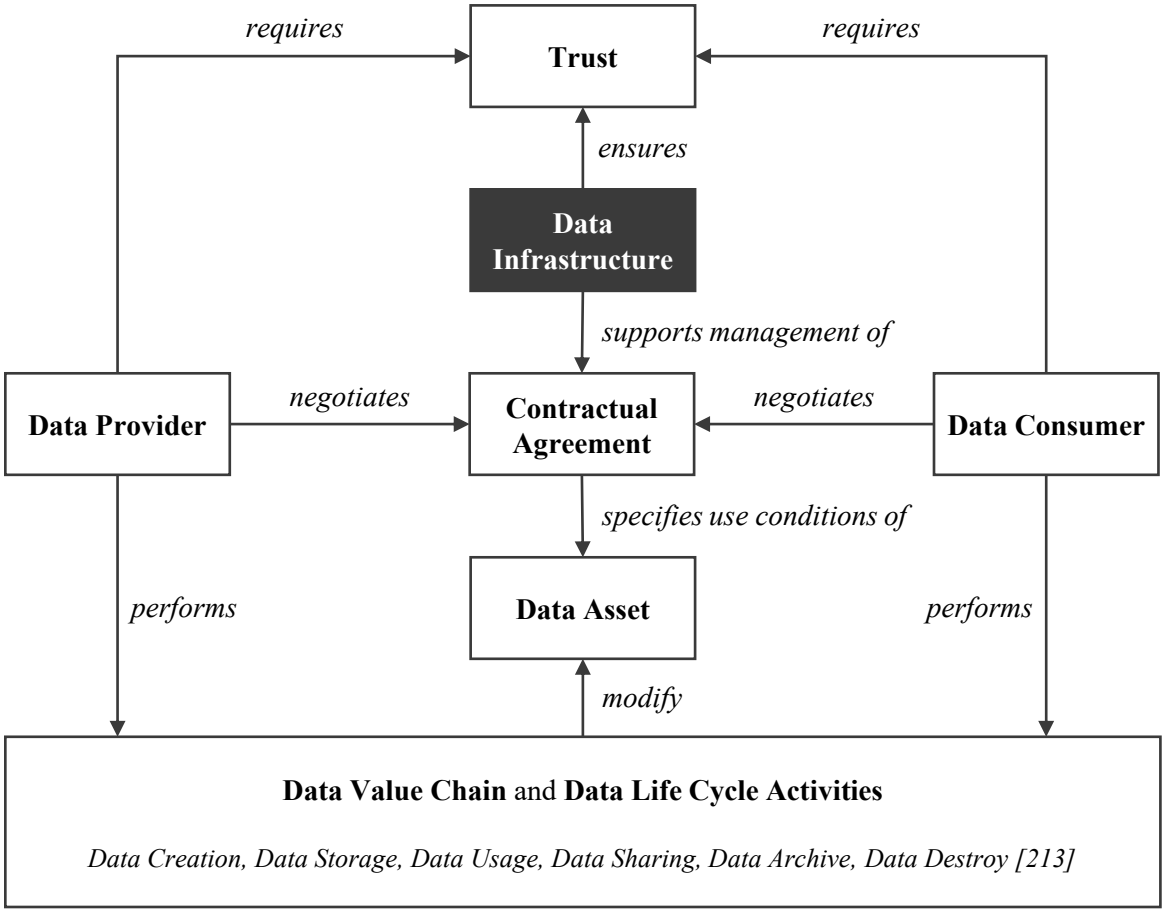
In summary, the topic of data sovereignty attracts significant interest in research and practice, with publication numbers still increasing. However, the academic literature lacks a foundational conceptualization of the concept, which defines its core meaning and distinguishes it from adjacent fields, motivating the first research area on data sovereignty foundations. Also, practitioners are looking for data sovereignty solutions that enable access and usage control enforcement across various IT landscapes, including identity management throughout the whole data life cycle, which introduces the second research area for designing and developing solutions.

4.2 Objectives of a Solution

In this second activity, we use the previously identified problem domain of data sovereignty challenges to derive the main objective of this thesis. We continue with the open points by first conceptualizing data sovereignty to conclude the theoretical research area one, and inferring digital watermarking as a possible solution space as an entry point for research area two.

Data Sovereignty Conceptualization. To address the lack of a clear definition of data sovereignty, we derived a conceptual model from literature in Paper III [271] (see Figure 4.3), which concluded research area one. To illustrate its structure, consider an original equipment manufacturer (OEM) as a *data provider* in the automotive industry currently designing a new car model. The company needs to cooperate with another 3D printing company, the *data consumer*, as a Tier 1 supplier to print a specific part of the new car model as contract work. This part acts as a *data asset*, and the company stores it as a 3D model file. Given the novelty of the new car model, controlling the data asset is crucial for the data-providing OEM to maintain a competitive advantage. Consequently, the OEM negotiates specific use conditions for the 3D printing data asset with the data consumer through a *contractual agreement*. Those conditions include a restriction on using only the 3D model file for the specific printing job, with a strict prohibition on sharing it with other companies. Both parties require a baseline of *trust* in the overall data sharing and usage process across the *data value chain and data life cycle activities*. To be data sovereign in this exemplified usage scenario, practitioners need a *data infrastructure* that supports the management and enforcement of the agreed conditions in the contractual agreement, further ensuring and strengthening overall trust.

In this context, data infrastructure is the central point that determines the strengths and weaknesses of data sovereignty (see Figure 4.3). The four technical challenges presented in the previous problem identification and motivation activity (see Section 4.1) directly map to the characteristics of the data infrastructure component. Stronger access and usage control mechanisms, stronger integration into existing IT landscapes throughout the data life cycle, and stronger identity management ensure greater trust and facilitate the effective management and enforcement of contractual agreements. When operating within an ecosystem of diverse participants, practitioners need technical solutions to enable such guarantees [162]. These observations lead to the question of how to design an artifact, in our case, an IT artifact in the domain of software engineering, to strengthen data sovereignty and address the existing challenges.



Based on Paper III [271]

Fig. 4.3 Conceptual Model for Data Sovereignty

Possible Solutions. A closer investigation of the scientific literature and our expert interviews revealed five main areas of possible technical solutions to strengthen data sovereignty, initially introduced in Paper II [95]:

- *Data Encryption:* Applying encryption techniques helps to secure data during storage and transit [188]. Besides traditional password- or key-based encryption techniques, specific trusted execution environments like Intel SGX or AMD SEV can also secure data during its use within protected enclaves [21], [47], [161].
- *Access Control:* Restricting and controlling access to data via password authentication, endpoint protection, or specific middleware components helps to stay more data sovereign [95]. Other possibilities include sticky policies [174] and connectors in data spaces as described next.

- *Data Spaces:* Sharing and using data in a decentralized architecture under specific boundary conditions and purposes offer additional possibilities to strengthen data sovereignty by defining and partially enforcing specific access and usage policies [140], [179]. Data spaces use a standardized Dataspace Protocol [137] and a technical connector component [119].

- *Metadata Enrichment:* Adding metadata about the contractual agreement inside a data asset helps identify the conditions under which data providers use the data asset [193]. One example is enriching the data with confidential levels (like public, internal, confidential, strictly confidential).

- *Watermarking:* Embedding a robust watermark inside physical assets or digital data helps to increase copyright protection and prevent copying [50]. Examples include existing algorithms for different data types, such as images, audio, video, or text [11], [198].

A closer review of these five possible solutions reveals already ongoing research or only limited solutions for increasing data sovereignty. For the first *data encryption* area, research is currently ongoing, as noted by Lohmöller et al. [161], aiming to strengthen data sovereignty within specific trusted execution environments. The second *access control* area already has working technical authentication solutions, but only helps maintain data sovereignty to a limited extent. After granting access to protected data, a data provider loses control and no longer remains data sovereign. While the third *data space* architecture promises to be a possible solution to current data sovereignty issues, practice shows various problems, such as the lack of enforcement mechanisms within the connector component to ensure control over data [95], [242]. At the same time, mitigations are only possible when combining with other concepts such as trusted execution environments [161]. The fourth *metadata enrichment* is still widely adopted by companies for data classification. Nevertheless, their robustness remains limited because employees can easily remove a visual confidential flag from the data asset. Thus, we can treat the fifth *watermarking* area as a more robust form of metadata enrichment. Big companies in the multimedia industry are successfully using watermarking or related technologies, such as digital rights management, to control and protect their data [159]. Examples include Adobe Stock's visible watermarks above images that get removed after purchase, or invisible watermarks inside movies that identify a pirate's seating position during filming in a cinema [156]. Ideas from other domains include securing sensor data from critical infrastructures via watermarking [222].

Combining data sovereignty with watermarking for industrial data requires further investigation. This research direction is in line with related work because the existing success of

watermarking "encourages the development of more sophisticated watermarking algorithms as part of a larger system for protecting valuable digital documents" [249, p. 694]. On closer analysis, both data sovereignty and digital watermarking share very similar objectives, such as protecting and controlling data and intellectual property [11], [182], [218]. In a direct comparison, watermarking has a long history dating back to the 13th century [50], with various technical and non-technical solutions emerging over time, while data sovereignty is a relatively new concept of the 20th century [89] with a majority of conceptual and information systems-related solutions [100]. Researchers have paid little attention to combining both research domains to address current challenges, which has led to our objective and the primary driver of this work.

Digital Watermarking. To investigate how digital watermarking can strengthen data sovereignty, we discuss which types of data need protection. Our contextual analysis in Paper II [95] showed that text-based data is the most relevant content type for practical use cases. Text data includes file types such as .pdf, .docx, spreadsheets, machine-readable files, and specific engineering design files [95]. The statement aligns with related work that classifies text as the most shared content type [30], [219]. Nevertheless, reviewing existing digital watermarking techniques reveals a large number of solutions for content types such as images, video, or audio, with text-based solutions being an under-researched and highly challenging area [11], [30], [126].

In addition, big tech companies like OpenAI and Google have recently generated broad public interest in AI-based LLM solutions through the release of chatbots such as ChatGPT and Gemini (formerly called Bard). This new era raises questions very similar to those presented so far, such as distinguishing between text generated by an LLM and text created by a human [44], [46], [136], [228], [252], [285]. The core problem and question remain centered on protecting text-based data assets to verify, trace, and maintain control over them. Using digital watermarking to verify the authenticity of text outputs from LLMs remains a key objective for mitigating the risks of generated content, such as plagiarism and intellectual property issues [57], [158]. The EU mandates it at a legal level in the AI Act [69].

To design a watermarking artifact that addresses the aforementioned problem domain and increases data sovereignty, we formulated concrete *knowledge claims (KCs)* to guide our contributions and achieve better outcomes than existing artifacts [151]. Larsen et al. [151] define such KCs as "a proposition about an artifact that asserts its contribution to science and society through a particular form or function" [151, p. 1273]. Table 4.2 lists our derived four KCs from the previously presented technical challenges as a result of the interview study of Paper II [95] under the consideration of current literature.

Table 4.2 Derived Knowledge Claims

#	Name	Description	Related Technical Data Sovereignty Challenges of Paper II [95]	Related Watermarking Literature
KC1	Embedding	A user is able to embed a watermark in a text-based data asset and verify it later.	*Access & Usage Control* and *Identity Management* when using the data provider's name and contractual agreement conditions as a watermark.	[11], [126], [219]
KC2	Invisibility	A user cannot identify whether a digital text document displayed on a computer screen has a watermark inside it or not.	*Infrastructure & Landscape* and *Data Processing Life Cycle* to work inconspicuously for users in various use cases.	[5], [11], [126], [135], [157]
KC3	Modification Robustness	The watermark is robust against text alterations, including adding, changing, or deleting existing sentences, reordering paragraphs, or changing fonts and formatting.	*Data Processing Life Cycle* to work on all data life cycle stages and activities.	[11], [126], [135], [157]
KC4	Usage Robustness	The watermark remains intact when the text is copied, pasted, and used across different business applications and file formats.	*Access & Usage Control* and *Infrastructure & Landscape* to remain intact across diverse IT infrastructure landscapes to maintain control.	[5], [8], [116], [126]

Following Larsen et al. [151]

The objective of this thesis is to design and develop a digital watermarking artifact for text that answers the **MRQ:** ***How to strengthen data sovereignty through digital watermarking?*** The applied DSR methodology from Peffers et al. [197] uses the introduced KCs to guide our iterative artifact design and development process (RQ 2.1), evaluate and enhance the artifact (RQ 2.2), and generalize the findings (RQ 2.3).

In summary, we conceptualized data sovereignty and identified the need to strengthen the data infrastructure component. We discussed possible solutions, including data encryption, access control mechanisms, data spaces, metadata enrichment, and watermarking, while we selected the latter as the most promising approach for text data. To build an objective for the solution, we combined knowledge from the environmental application domain (interviews) with the knowledge base (literature) and derived four claims of our knowledge. These KCs focus on embedding, invisibility, modification robustness, and usage robustness, guiding the design and evaluation of our IT artifact.

4.3 Design & Development

In this third activity, we present the design and development of the IT artifact [197]. We distinguish our solution objective from existing work and introduce two algorithms for watermark embedding and extraction as our developed IT artifact for the practical problem [85].

Artifact Characteristics. RQ 2.1, as an entry point for research area two, targets the design and development of a digital watermarking artifact for plain text that fulfills the derived KCs (see Table 4.2). Even if digital text watermarking algorithms are limited and underresearched [11], [126], some solutions already exist, partly classified in related fields such as steganography or information hiding, as shown by existing reviews [4], [135], [142], [198], [262]. Section 2.2 presents a detailed overview of related work and similar methods, and Section 4.5 compares and evaluates them against our artifact. To the best of our knowledge, none of the existing solutions fulfill all of the derived KCs. For example, solutions by Shazzad-Ur-Rahman et al. [236] or Rizzo, Bertini, and Montesi [219] are not invisible or imperceptible, whereas the latter and others by Ahvanooey et al. [9] exhibit a low robustness when used across different applications [94]. Due to the existing gap, researchers call for new robust solutions [286].

In response, we designed a new digital text watermarking artifact, which we call *Innamark* in the remainder of this thesis, initially developed in Paper IV [99] and extended in a second iteration in Paper VI [94]. Innamark is a portmanteau of the words *invisible* and *watermarking*. It consists of an embedding and extraction algorithm, which we present below in its latest version. Table 4.3 summarizes the nomenclature used in these algorithms and the descriptions we provide. Due to its novelty, we registered Innamark's embedding technique for patent protection in Germany [97] and filed an international patent application under the Patent Cooperation Treaty [98].

Table 4.3 Overview of Nomenclature

Symbol	Meaning
CT	Cover text
$\mathcal{W}$	Watermark as text
$\mathcal{W}_{bytes}$	Byte representation of $\mathcal{W}$ as a sequence of digits
$\mathcal{W}_H$	Hidden whitespace representation of $\mathcal{W}_{bytes}$
$CT_{\mathcal{W}}$	Cover text containing an embedded watermark
$\mathcal{U}$	Set of all Unicode characters
$\mathcal{S}$	Set of all 17 Unicode space characters (see Table 2.6)
δ	Classical most used Unicode whitespace (U+0020)
ϕ	Separator whitespace character
$\mathcal{A}_+$	Whitespace alphabet with ϕ
$\mathcal{A}_-$	Whitespace alphabet without ϕ
θ	Configuration parameter for encryption, compression, etc.
$Emb(CT, \mathcal{W}_{bytes}, \theta)$	Embedding algorithm
$Ext(CT_{\mathcal{W}})$	Extraction algorithm

Following Ahvanooey et al. [6] and based on Paper IV [99] and Paper VI [94]

4.3.1 Watermark Embedding

A watermark encoder, as introduced in the background Section 2.2.2 and in Figure 2.4, transforms a cover text CT and a watermark $\mathcal{W}$ into a watermarked cover text $CT_{\mathcal{W}}$. Our algorithm uses $\mathcal{W}_{bytes}$ instead of $\mathcal{W}$ as input, because it hides any byte-encoded sequence within a cover text rather than text only. The main idea of Innamark's watermarking encoder is to transform the watermark into a hidden set $\mathcal{W}_H$ of similar-looking Unicode whitespaces and embed it in CT by replacing all classical whitespaces δ with our encoded hidden set $\mathcal{W}_H$ to create $CT_{\mathcal{W}}$. The overall embedding process of the encoder consists of three parts and aims to address *KC1: Embedding*. It starts by analyzing and applying the user-defined configuration parameter θ to support specific use cases and preferences, such as compression, hash-based verification, or error correcting codes (Part 1). Next, the watermark in its byte representation $\mathcal{W}_{bytes}$ is encoded into the hidden whitespace set $\mathcal{W}_H$ (Part 2). At the end, $\mathcal{W}_H$ is inserted into CT by repeatedly replacing all whitespaces δ with the elements of $\mathcal{W}_H$ to embed the watermark multiple times, resulting in $CT_{\mathcal{W}}$ (Part 3). We present each part in detail with a concrete embedding example, and summarize the overall process in Algorithm 1.

Algorithm 1 Watermark Embedding $Emb(CT, \mathcal{W}_{bytes}, \theta)$

Require: Cover text (CT), Watermark ($\mathcal{W}_{bytes}$), Configuration parameter (θ)
Ensure: Cover text including hidden watermark ($CT_{\mathcal{W}}$)

function EMB($CT, \mathcal{W}_{bytes}, \theta$)
  ▷ *PART 1: Apply tag based on the config parameter θ*
  $\mathcal{W}_{bytes} \leftarrow applyTag(\mathcal{W}_{bytes}, \theta)$ ▷ Add InnamarkTag config to the watermark

  ▷ *PART 2: Encode watermark into the hidden whitespace alphabet*
  $d \leftarrow \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil$ ▷ Represent each input byte by d alphabet chars
  for each $q \in \mathcal{W}_{bytes}$ **do** ▷ Encode watermark incl. InnamarkTag input bytes
    for $i \leftarrow 1$ **to** d **do** ▷ Apply cascading modulo per input byte
      $r \leftarrow q \bmod |\mathcal{A}_-|$ ▷ Identify correct alphabet char as remainder r
      $q \leftarrow \left\lfloor \frac{q}{|\mathcal{A}_-|} \right\rfloor$ ▷ Update quotient q for the next loop
      $\mathcal{W}_H \leftarrow \mathcal{W}_H + a_{r+1}$ ▷ Append new hidden alphabet whitespace $a_{r+1} \in \mathcal{A}_-$
    end for
  end for

  ▷ *PART 3: Insert hidden watermark multiple times by transforming the cover text*
  $i \leftarrow 0$ ▷ Use i as helper for watermark length
  for each $c \in CT$ **do** ▷ Iterate over all chars of the input cover text
    if $c = \delta$ **then** ▷ Check for replaceable whitespaces
      if $i = 0$ **then** ▷ Check start of insertion
        $CT_{\mathcal{W}} \leftarrow CT_{\mathcal{W}} + \phi$ ▷ Start by appending the separator character
        $i \leftarrow i + 1$ ▷ Increase i to 1 to start watermark insertion
      else if $0 < i \leq |\mathcal{W}_H|$ **then** ▷ Check if the current watermark is not fully inserted
        $w_{H_i} \leftarrow \mathcal{W}_{H_i}$ ▷ Get next char $w_{H_i} \in \mathcal{A}_-$ of the hidden watermark
        $CT_{\mathcal{W}} \leftarrow CT_{\mathcal{W}} + w_{H_i}$ ▷ Append char to the watermarked cover text
        $i \leftarrow i + 1$ ▷ Increase helper variable for watermark length
      else ▷ Watermark fully inserted
        $i \leftarrow 0$ ▷ Restart with another insertion
      end if
    else ▷ Current cover text char is not a classical whitespace
      $CT_{\mathcal{W}} \leftarrow CT_{\mathcal{W}} + c$ ▷ Append cover text char to watermarked cover text
    end if
  end for
  return $CT_{\mathcal{W}}$ ▷ Return watermarked cover text – Done
end function

Based on Paper IV [99] and Paper VI [94]

Part 1: Apply Tag. The first part defines a one-byte *InnamarkTag* to identify the watermark type and structure. It is a similar approach to the headers in IPv6 protocols, where specific sets of bits have well-defined structures and locations [56]. Every bit in our InnamarkTag identifies if a specific well-defined functionality is enabled (1) or not (0). In the current version of our artifact, five optional configurations are possible (see Table 4.4). During every watermark insertion process, users have the possibility to combine functionalities based on the use case and environmental requirements via the configuration parameter θ.

For example, if the user needs to embed a long watermark in a compressed form, the second bit (see Table 4.4) of the tag is flipped to one, resulting in 01000000 as the one-byte InnamarkTag. If a SHA3-256 hash is added in addition to the compression, the fifth bit is flipped as well, resulting in 01001000. The flipping mechanisms, together with the fixed one-byte size, allow the use of any combination of optional functionalities while maintaining a well-defined structure for interoperable extraction. The last three bits of the InnamarkTag (bits 6-8) are currently unused to accommodate future requirements. If specific, customized configurations are needed, flipping the first bit indicates a special custom format that differs from the default implementation.

Table 4.4 Structure of an InnamarkTag

Bit Pos.	Type	Function
1	Custom	Identifies an unknown, customer-specific format.
2	Compressed	Watermark uses compression via zlib [77].
3	Sized	Watermark includes its size/length.
4	CRC32	Watermark includes CRC.
5	SHA3-256	Watermark includes a 256-bit long SHA-hash.
6	*Unused*	-
7	*Unused*	-
8	*Unused*	-

Based on Paper VI [94]

In Algorithm 1, we implement the InnamarkTag application using the `applyTag` method, with the watermark $\mathcal{W}_{bytes}$ and the configuration parameter θ as input parameters. Since the method uses only deterministic evaluation and the application of hashing, compression, etc., depending on θ, we abstracted it into a single method call to maintain clarity in Algorithm 1. The algorithm applies the method directly to the watermark. It returns an updated version of $\mathcal{W}_{bytes}$ because compressing changes the watermark, and the Sized, CRC32, and SHA3-256 (see Table 4.4) parameters extend the watermark with their information as a prefix. If no additional functionalities are needed ($\theta = \varnothing$), the algorithm uses the default tag 00000000.

The first version of the watermarking artifact presented in Paper IV [99] does not include a tag. Since Larsen et al. [151] suggest validating the KCs during artifact creation, we identified a lack regarding *KC3: Modification Robustness* and *KC4: Usage Robustness* after DSR iteration one. Without a tag, related technical challenges, such as maintaining control across different IT infrastructures and landscapes, remain challenging when the watermark format is unclear. Related work also emphasizes the importance of robustness, especially after modifications [11], [286]. This leads to an updated artifact including the InnamarkTag, which we present here and first introduced in Paper VI [94] as part of DSR iteration two.

Part 2: Encode Watermark. After $\mathcal{W}_{bytes}$ is adapted based on the InnamarkTag, the second part transforms $\mathcal{W}_{bytes}$ into $\mathcal{W}_H$ as a sequence of similar-looking whitespaces. While the main idea of Innamark is to replace all classical space characters δ with similar-looking space characters, we analyzed all regular space characters $\mathcal{S}$ (see the first part of Table 2.6). The goals were to identify spaces that are not directly distinguishable by humans and that are supported by different applications to fulfill *KC2: Invisibility* and *KC4: Usage Robustness*. Table 4.5 presents the results of our analysis, initially introduced in Paper IV [99]. The 'Width' column compares the width of the analyzed space with δ to indicate if it is similar, making it unable to distinguish by a human in normal viewing conditions (✓) or if a human is able to identify it (✗). For the other columns, a tick indicates that the space character persists, while a cross indicates that the tested file type or application does not support it. We document edge cases in brackets and discuss them later in Section 4.5.

Besides the classical space δ, only five space characters are supported by all tested file types and applications, and their width is similar to δ. The resulting set of five space characters shown as bold formatted in Table 4.5 form our whitespace alphabet $\mathcal{A}_+ := \{a : a \in S \land a \in U \land a$ only ticks in Table 4.5$\}$, so $\mathcal{A}_+ \subset \mathcal{S} \subset \mathcal{U}$. Since it is highly suggested to embed the watermark multiple times [177], if possible, let ϕ be one of the five whitespaces from the alphabet $\mathcal{A}_+$ that we use as a separator between multiple insertions. Let $\mathcal{A}_-$ be the whitespace alphabet without the separator character ϕ, so that $\mathcal{A}_+$ includes $\mathcal{A}_-$ and ϕ, with $\phi \notin \mathcal{A}_-$, leading to $\mathcal{A}_- \subset \mathcal{A}_+$.

Next, the algorithm transforms every byte of the watermark $\mathcal{W}_{bytes}$ into the whitespace alphabet $\mathcal{A}_-$. We use a cascading modulo operation, as described in Algorithm 1, to perform this transformation. For every input byte,

$$d = \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil \tag{4.1}$$

Table 4.5 Whitespace Evaluation

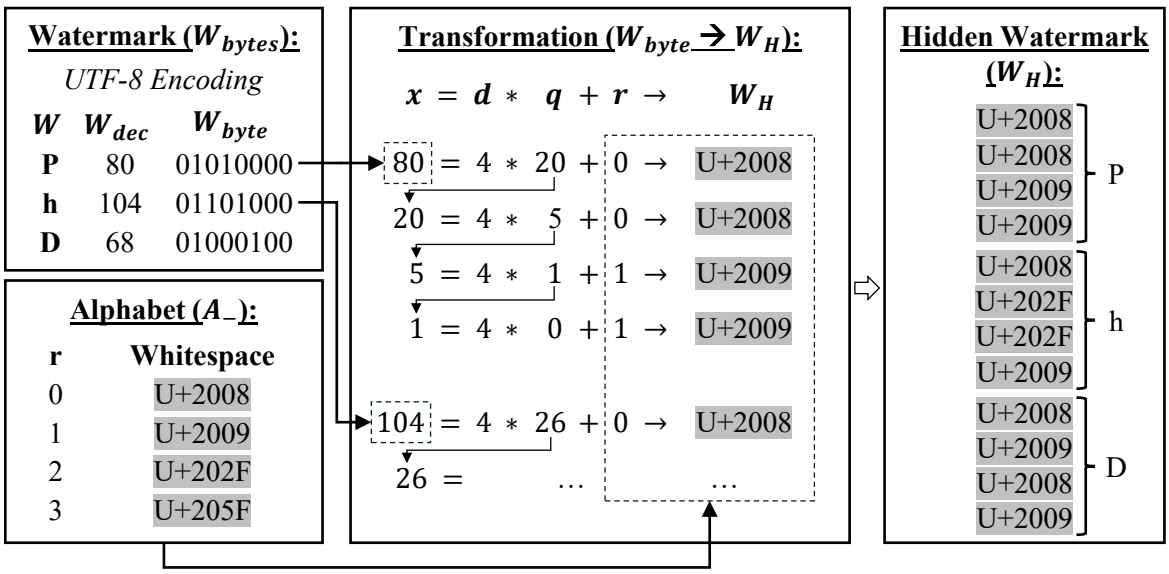
Whitespace Name	Code	Width	.txt	.docx	.pdf	Mail	MS Teams
Space	U+0020	✓	✓	✓	✓	✓	✓
No-break Space	U+00A0	✓	✓	✗	✗	✗	✗
Ogham Space Mark	U+1680	✗	✓	✓	✓	✓	✓
En Quad	U+2000	✗	✓	✓	(✓)	✓	✓
Em Quad	U+2001	✗	✓	✓	(✓)	✓	✓
En Space	U+2002	✗	✓	✓	✗	✓	✓
Em Space	U+2003	✗	✓	✓	✗	✓	✓
Three-per-em Space	**U+2004**	✓	✓	✓	(✓)	✓	✓
Four-per-em Space	U+2005	✓	✓	✗	✗	✗	✓
Six-per-em Space	U+2006	(✗)	✓	✓	(✓)	✓	✓
Figure Space	U+2007	✗	✓	✓	(✓)	✓	✓
Punctuation Space	**U+2008**	✓	✓	✓	(✓)	✓	✓
Thin Space	**U+2009**	✓	✓	✓	(✓)	✓	✓
Hair Space	U+200A	(✗)	✓	✓	(✓)	✓	✓
Narrow No-break Space	**U+202F**	✓	✓	✓	(✓)	✓	✓
Medium Mathematical Space	**U+205F**	✓	✓	✓	(✓)	✓	✓
Ideographic Space	U+3000	✗	✓	✓	✗	✓	✓

Following Korpela [138] and based on Paper IV [99]

whitespaces of $\mathcal{A}_{\_}$ are needed for the encoding, whereas $|\cdot|$ denotes the cardinality. While our implementation uses $|\mathcal{A}_{\_}| = 4$, leading to $d = 4$ with a 1:4 ratio between input bytes and output whitespaces. When using UTF-8 encoding, each Latin letter or number in a watermark is represented by one byte and converted to four whitespace characters. Due to the modularity of our implementation, it supports alternative alphabets $\mathcal{A}_{\_}$ for different use cases, resulting in other encoding ratios.

In Algorithm 1, the outer loop goes through each input byte, while the inner loop depends on the alphabet used and its encoding ratio d. The main goal of this second part is to create the hidden sequence $\mathcal{W}_H$, with $\mathcal{W}_H = \{a : \forall a \in \mathcal{A}_{\_}\}$ by suggestively appending every transformed character to $\mathcal{W}_H$. Figure 4.4 shows an example that transforms the watermark text $\mathcal{W}$ = "PhD" into the sequence $\mathcal{W}_H$ of 12 similar-looking whitespaces, based on UTF-8 encoding.

Part 3: Insert Watermark. In the last part, the algorithm embeds the encoded watermark $\mathcal{W}_H$ multiple times into the cover text. Based on Algorithm 1, it loops over every character c in the cover text CT and searches for all classical whitespaces δ. Next, it replaces the



Based on Paper VI [94]

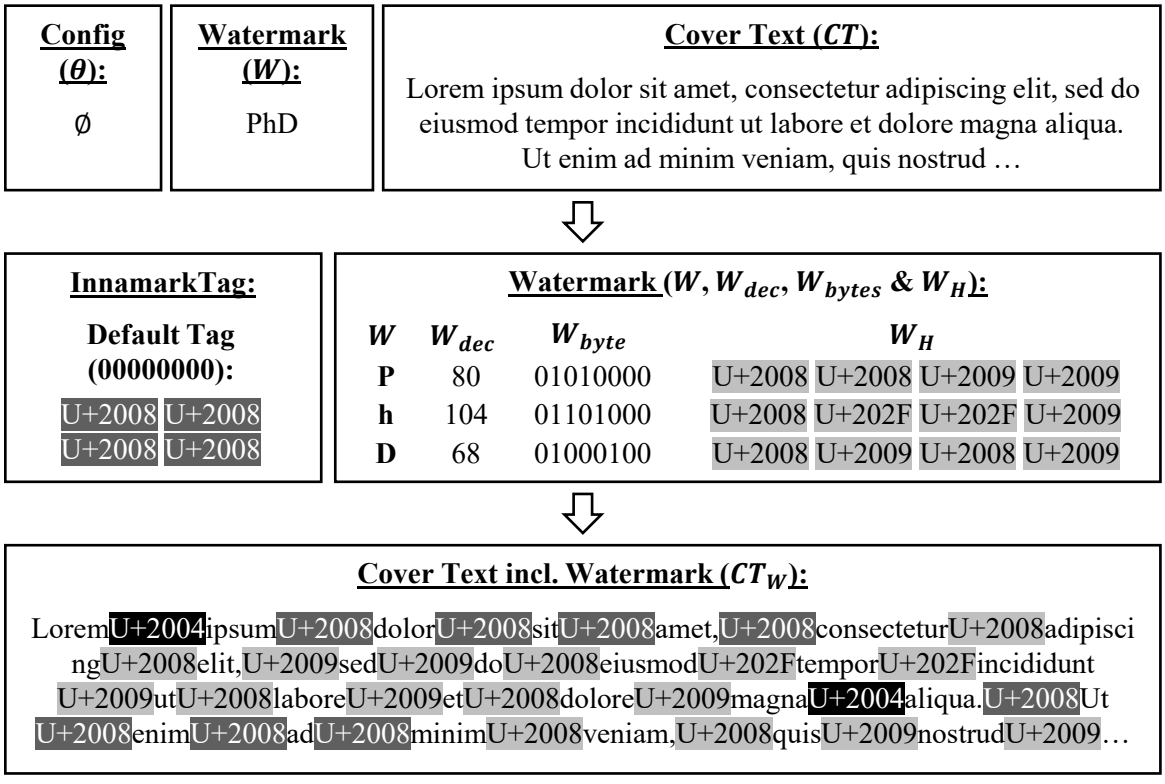
Fig. 4.4 Watermark Alphabet Transformation Example

first δ found by our separator character ϕ to indicate the start of an embedded watermark. Afterward, the algorithm inserts the whitespace sequence $\mathcal{W}_H$ that includes the watermark and optional configuration properties in the InnamarkTag via continuous δ replacements. Thus, it replaces the second δ found by the first similar-looking, hidden whitespace characters w_{H_1} as the first element of $\mathcal{W}_H$; it replaces the third δ found by w_{H_2} and so on until it has $\mathcal{W}_H$ fully inserted once.

Compared to related work, most existing algorithms are highly vulnerable to minor text changes because they embed the watermark only once [3], [9], [144], [147], [148], [176]. Mohanty et al. [177] suggest that watermarks should span a large area or the full document. To fulfill *KC3: Modification Robustness*, we decided to embed the watermark multiple times, as in Rizzo, Bertini, and Montesi [219], to ensure it remains extractable even if parts of the text are deliberately or accidentally altered. Consequently, the loop starts again by inserting the separator ϕ and $\mathcal{W}_H$ multiple times until all δ are replaced. Thus, part three successively builds the watermarked cover text by iteratively creating $CT_{\mathcal{W}}$ through multiple insertions of $\mathcal{W}_H$. At the end, the algorithm returns $CT_{\mathcal{W}}$ as a watermarked cover text that no longer contains any δ.

The final number of multiple insertions of the watermark strongly depends on the length of $\mathcal{W}_H$ relative to the length and number of whitespaces in CT. While a short watermark is inserted many times inside a long cover text, a long watermark might only fit one or two times, or not even one time, inside a very short cover text.

Embedding Example. For explanatory purposes, we demonstrate the described embedding process by watermarking a three-letter watermark inside a Lorem ipsum cover text (see Figure 4.5). We use the simplest version without any special configurations, leading to the three input parameters $\theta = \varnothing$ for the default configuration, $\mathcal{W} =$ "PhD" for the watermark, and $CT =$ "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud ..." for the cover text as shown on the top part of Figure 4.5. Since Algorithm 1 can embed any byte-encoded sequence, it uses the byte representation $\mathcal{W}_{bytes}$ instead of the string $\mathcal{W}$ as input. Based on ASCII and UTF-8, the first letter $\mathcal{W}_P =$ "P" of the watermark has the decimal value $\mathcal{W}_{dec_P} = 80$, which corresponds to the byte representation $\mathcal{W}_{byte_P} = 01010000$.



Based on Paper VI [94]

Fig. 4.5 Watermark Embedding Example

Part 1 starts by applying the configuration parameter θ to build the InnamarkTag. Since no special configurations are specified here ($\theta = \varnothing$), the algorithm uses the default tag 00000000 without any bit flips.

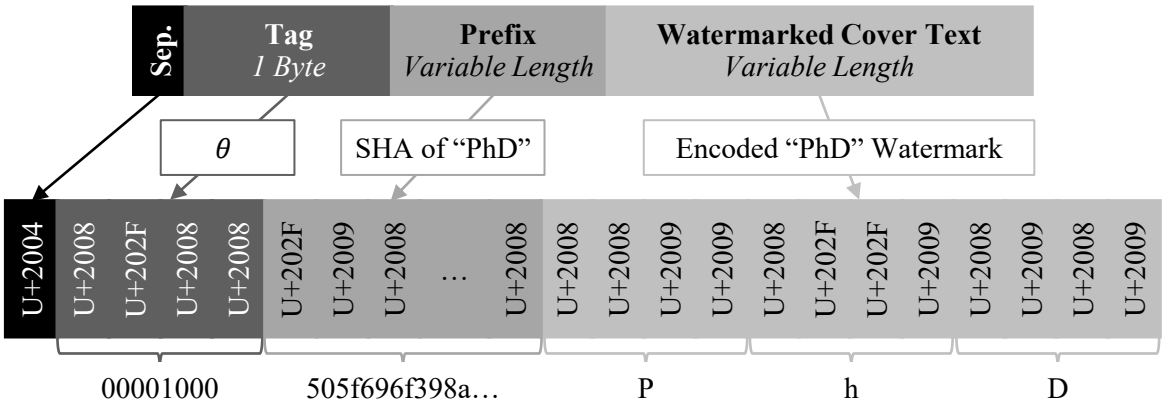
Part 2 continues by transforming the watermark into the sequence of similar-looking Unicode whitespace characters to build the hidden version $\mathcal{W}_H$ (see Figure 4.4). We use

our identified alphabet $\mathcal{A}\_$ with $d = 4$. The cascading modulo operation starts by repeatedly calculating $q \bmod |\mathcal{A}\_| = r$, while updating q with the last quotient and r with the resulting remainder. Skipping the zeros of the default InnamarkTag, this step uses the first byte $\mathcal{W}_{byte_P} = 01010000$ as the decimal value $\mathcal{W}_{dec_P} = 80 = q$, leading to $r = 0$ ($q \bmod d = r \Rightarrow 80 \bmod 4 = 0$) with the quotient $q = 20$. Based on the alphabet mapping on the left part of Figure 4.4, our mapping uses the punctuation space (U+2008) in cases of $r = 0$. The cascading modulo operation continues with the new quotient $q = 20$ and runs $d = 4$ times per byte. Thus, the algorithm transforms the letter $\mathcal{W}_P =$ "P" into the four whitespaces $\mathcal{W}_{H_P} =$ "U+2008 U+2008 U+2009 U+2009" and repeats the process for the other two letters "h" and "D" (see Figure 4.4).

Part 3 replaces all whitespaces to produce the resulting watermarked cover text $CT_{\mathcal{W}}$ as shown at the bottom of Figure 4.5. It replaces the first classical whitespace δ between the words "Lorem ipsum" by our separator character $\phi =$ "U+2004" (three-per-em space) to indicate the start of the watermark. The algorithm replaces the upcoming four whitespaces to embed the InnamarkTag, in our case, the default tag represented by "U+2008 U+2008 U+2008 U+2008". Since this configuration does not enable any special options, no prefix is inserted, while the subsequent replacements embed the transformed watermark itself. After the first watermark insertion, the embedding process restarts by inserting the separator, the InnamarkTag, the prefix (in this example, empty), and the watermark until no whitespaces δ remain. As seen in Figure 4.5, the algorithm embeds the watermark almost twice, while the second insertion aborts after the second $\mathcal{W}_P =$ "P".

If the user chooses another configuration θ, only Part 1 changes the InnamarkTag and $\mathcal{W}_{bytes}$, while the rest of the process remains unchanged. If, for example, the user enables hash verification, the algorithm flips the fifth bit of the InnamarkTag to 1 according to Table 4.4, resulting in 00001000. Further, the `applyTag` method computes the SHA3-256 hash of $\mathcal{W} =$ "PhD" and adds it between the tag and watermarked cover text. Figure 4.6 illustrates this alternative version.

In summary, we designed Innamark's three-part watermark embedding algorithm. It (i) applies a configuration tag to enable possible checks, compression, hashes, and error correction; (ii) encodes the watermark into an alphabet set of similar-looking Unicode whitespaces; (iii) embeds the watermark multiple times into the cover text by replacing all whitespaces with the encoded hidden whitespace sequence of the previous step. We demonstrate the embedding algorithm using a "PhD" watermark within a "Lorem ipsum" cover text with the default configuration.



Based on Paper VI [94]

Fig. 4.6 "PhD" Watermark Example with Enabled Hashing

4.3.2 Watermark Extraction

A watermark decoder, as part of an extraction subsystem, transforms a watermarked cover text $CT_{\mathcal{W}}$ back into the original watermark $\mathcal{W}$ (see Figure 2.4). Since Innamark is a blind watermarking system, it does not need the original cover text CT as an additional input parameter. Similar to the encoder's watermark embedding process, we split the extraction into three parts. It starts by searching for all similar-looking whitespaces based on our alphabet $\mathcal{A}_+$ to extract the hidden watermark $\mathcal{W}_H$ from the input text (Part 1). Next, it analyzes the InnamarkTag and applies specific operations, like verifying a hash if hashing is enabled or decompressing a compressed watermark if compression is enabled (Part 2). Lastly, it decodes the hidden watermark $\mathcal{W}_H$ back in its original byte representation $\mathcal{W}_{bytes}$ (Part 3). Algorithm 2 summarizes the overall process, while we present each part in detail below.

Part 1: Extract Watermark. The first part extracts the relevant characters of the watermark from the input text. It starts by iterating over every character $c \in CT_{\mathcal{W}}$ to check whether it belongs to our alphabet set $\mathcal{A}_+$. While the algorithm skips the first separator character ϕ, which indicates the start of a watermark, it continuously appends all identified $c \in \mathcal{A}_-$ to $\mathcal{W}_H$ to rebuild the set of relevant hidden watermark characters. The second ϕ found marks the end of the first watermark and the start of the second identical watermark and triggers a break operation to finish Part 1, since the algorithm has fully extracted the watermark.

The reason for multiple watermark insertions during watermark embedding is to increase the *KC3: Modification Robustness* and *KC4: Usage Robustness* as recommended in related work [11], [177], [219]. If problems occur during the reconstruction of the watermark due

Algorithm 2 Watermark Extraction $Ext(CT_{\mathcal{W}})$

Require: Cover text including hidden watermark ($CT_{\mathcal{W}}$)
Ensure: Extracted watermark ($\mathcal{W}_{bytes}$) or error

function Ext($CT_{\mathcal{W}}$)
  ▷ *PART 1: Extract watermark alphabet characters from input text*
  for each $c \in CT_{\mathcal{W}}$ **do** ▷ Iterate over all chars of a watermarked cover text
    if $c \in \mathcal{A}_+$ **then** ▷ Check if current char is in alphabet
      if $c = \phi$ **and** $\mathcal{W}_H \neq \varnothing$ **then** ▷ Check if current char is not the first separator
        break ▷ Watermark fully extracted – Break
      else if $c \in \mathcal{A}_-$ **then** ▷ Check if the current char is part of watermark
        $\mathcal{W}_H \leftarrow \mathcal{W}_H + c$ ▷ Append char to hidden watermark
      end if
    end if
  end for

  ▷ *PART 2: Analyze tag and prefix and check for errors*
  $\mathcal{W}_H, error \leftarrow analyzeTag(\mathcal{W}_H)$ ▷ Extract, check, and analyze InnamarkTag
  if error **then** ▷ Check if tag analysis reports error (like an invalid hash)
    return error ▷ Return error – Done
  end if

  ▷ *PART 3: Decode hidden watermark into byte format*
  $d \leftarrow \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil$ ▷ Represent each input byte by d alphabet chars
  for $i \leftarrow 0$ **to** $|\mathcal{W}_H|$ step d **do** ▷ Iterate through watermark in blocks of d
    for $y \leftarrow 0$ **to** $d - 1$ **do** ▷ Iterate through every hidden alphabet whitespace
      $a_{r+1} \leftarrow \mathcal{W}_{H_{i+y+1}}$ ▷ Read current alphabet whitespace $a_{r+1} \in \mathcal{A}_-$
      $b \leftarrow b + r \cdot d^y$ ▷ Reconstruct input byte iteratively with $r \in [0, \ldots, d-1]$
    end for
    $\mathcal{W}_{bytes} \leftarrow \mathcal{W}_{bytes} + b$ ▷ Append reconstructed byte to watermark result
  end for
  return $\mathcal{W}_{bytes}$ ▷ Return watermark in byte format – Done
end function

Based on Paper IV [99] and Paper VI [94]

to intentional or unintentional destruction, the decoder instead uses the second or another identical replica of the watermark. In Algorithm 2, we focus on the simplest version by describing the extraction of the first watermark occurrence only, to keep this work legible and understandable. We demonstrate the full potential in our Kotlin reference implementation published on GitHub [75], including implemented error handling for altered or broken watermarks and frequency analysis to identify intact watermarks in a text with multiple watermark defects.

Part 2: Analyze Tag. The second part reads the InnamarkTag from the extracted watermark $\mathcal{W}_H$ and analyzes it. It uses the fixed one-byte length and bit-mapping structure introduced in Table 4.4. With it, the algorithm understands and interprets the format and type of watermark, without any knowledge about the embedding process or system. It enables interoperability, directly addressing the identified technical challenges in Paper II [95] of diverse IT infrastructures and landscapes. The `analyzeTag` method in Algorithm 2 uses the extracted watermark $\mathcal{W}_H$, including the tag, an optional prefix, and the watermark itself, as input and returns $\mathcal{W}_H$ in its pure format, uncompressed, without the tag and prefix, or an error. An error occurs if the input is almost completely broken, if the algorithm cannot verify the hash or uncompress the watermark, or if the first bit of the InnamarkTag equals one, indicating a custom, unknown format. If no critical errors occur, the algorithm proceeds to the third decoding part.

Part 3: Decode Watermark. The last part decodes the watermark by extracting the sequence of whitespaces $\mathcal{W}_H$ and converting it back to its original byte representation $\mathcal{W}_{bytes}$. We abstracted and generalized Algorithm 2 to support different encoding alphabets, whereas this work uses the alphabet $\mathcal{A}_{\_}$ of four whitespaces that we identified. This leads to a step size of $d = 4$ for the outer loop, since the encoder represents each byte with four whitespaces, as shown earlier (see Figure 4.4). The inner loop iterates over every whitespace to reconstruct every original byte b by reversing the modulo operation of the embedding. After the decoder has processed four whitespaces back into b, it appends to the resulting set $\mathcal{W}_{bytes}$ and continues until it has decoded the full watermark.

Extraction Example. Building on the embedding example shown in Figure 4.5, we demonstrate the extraction of the watermark. We use the input cover text with an embedded watermark $CT_{\mathcal{W}}$ = "Lorem*U+2004*ipsum*U+2008*dolor...", initializing $\mathcal{W}_H = \varnothing$.

Part 1 starts by iterating over all characters c of the input text until it finds the first whitespace of our alphabet $\mathcal{A}_{+}$. The first whitespace U+2004 between the two words "Lorem" and "ipsum" is skipped because it is the separator ϕ. The upcoming U+2008 between "ipsum" and "dolor" is part of our alphabet $\mathcal{A}_{\_}$ and appended to $\mathcal{W}_H$. This procedure continues until it finds the second ϕ between "magna" and "aliqua", resulting in a break and the end of Part 1. Consequently, $|\mathcal{W}_H| = 16$ because it consists of four whitespaces for the InnamarkTag and 12 whitespaces for the watermark itself.

Part 2 analyzes the InnamarkTag by checking and trimming the first four whitespaces of $\mathcal{W}_H$. Here, "U+2008 U+2008 U+2008 U+2008" indicates the default tag 00000000 without any specializations. Therefore, the `analyzeTag` method returns the whitespace wa-

termark sequence $\mathcal{W}_H$ without the tag, so $|\mathcal{W}_H| = 12$ and error = $\varnothing$. In case of another tag, such as an enabled hashing (see Figure 4.6), the `analyzeTag` method checks the hash stored between the tag and the watermarked cover text. It returns an error if it is invalid or the whitespace watermark sequence $\mathcal{W}_H$ without a tag and prefix.

Part 3 decodes the hidden watermark $\mathcal{W}_H$ back into its byte representation. In our example, the outer loop splits $\mathcal{W}_H$ into chunks of $d = 4$, while the inner loop reverses the modulo operation of the embedding for every chunk. Starting with the first chunk $\mathcal{W}_{H_{1-4}}$ = "U+2008 U+2008 U+2009 U+2009", the first iteration of the inner loop uses the first element $\mathcal{W}_{H_1} = a_1$ = "U+2008" to use its index $r = 0$ in the alphabet (see Figure 4.4) to calculate $b = b + r \cdot d^y \Rightarrow 0 + 0 \cdot 4^0 = 0$. The second inner loop iteration continues with the next element $\mathcal{W}_{H_2} = a_1$ = "U+2008" to update $b = b + r \cdot d^y \Rightarrow 0 + 0 \cdot 4^1 = 0$. Consequently, the third iteration leads to $\mathcal{W}_{H_3} = a_2$ = "U+2009" and $b = b + r \cdot d^y \Rightarrow 0 + 1 \cdot 4^2 = 16$ and the fourth to $\mathcal{W}_{H_4} = a_2$ = "U+2009" and $b = b + r \cdot d^y \Rightarrow 16 + 1 \cdot 4^3 = 80$. The algorithm appends the resulting decimal value of $b = 80$ to $\mathcal{W}_{bytes}$. Thus, it continues with the next chunk until it has processed everything, resulting in a returned value of $\mathcal{W}_{bytes} = \{80, 104, 68\}$. Based on the UTF-8 and ASCII mapping, these three numbers represent the watermark $\mathcal{W}$ = "PhD".

In summary, we designed Innamark's three-part watermark extraction algorithm. It (i) extracts the watermark by reading all hidden alphabet whitespace characters; (ii) analyzes and applies the configuration tag by verifying hashes or sizes, handling compression, dealing with encryption, and performing error correction; and (iii) decodes the hidden watermark back into byte format. We demonstrate the extraction algorithm using the text input from the previous "PhD" watermark example.

4.4 Demonstration

We present our developed artifact for RQ 2.1 and RQ 2.2 as a Kotlin multiplatform library. To demonstrate the resolution of current data sovereignty problems, we showcase its functionality in a data space context with connector extensions for diverse IT architectures. We used the results to continuously improve the artifact, as demonstrated by a web interface and a CLI tool in its latest version, showcasing its application in JavaScript and Java projects.

Kotlin Library. We implemented the watermark embedding of Algorithm 1 and the extraction of Algorithm 2 in an extended version as a Kotlin multiplatform library and published it on GitHub [75]. Kotlin is a relatively new programming language, started in 2010 by JetBrains [63]. Kotlin is often known only for Android development, but it offers the possi-

bility of a Java build target for developing server-side backend applications or being compiled to JavaScript for frontend application development, thanks to its multiplatform support [63]. We used both build targets for the web interface and CLI tool to address the challenge of integration into various IT infrastructures and landscapes, as discussed in Paper II [95]. Similar to Java, Kotlin is a statically typed programming language that enables functional and object-oriented development with its core being available free and open source [63], making it an appropriate language for building scientific IT artifacts [187].

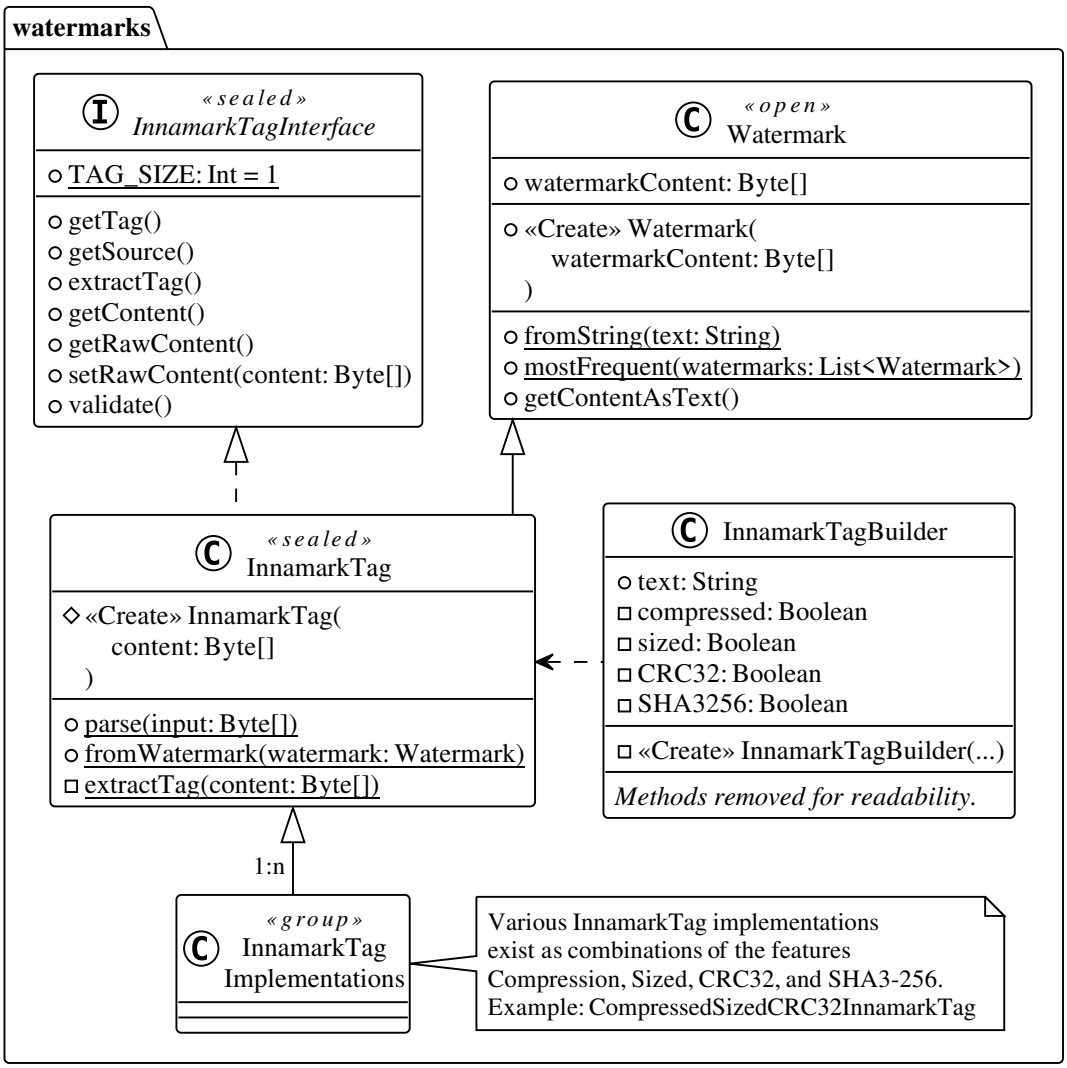


Fig. 4.7 Simplified Excerpt of Innamark's UML Class Diagram for Watermarks

We structure the overall architecture of our Innamark Kotlin library into two core packages. Figure 4.7 illustrates the Unified Modeling Language (UML) class diagram for `watermarks`, and Figure 4.8 illustrates the UML class diagram for `watermarkers`. Both

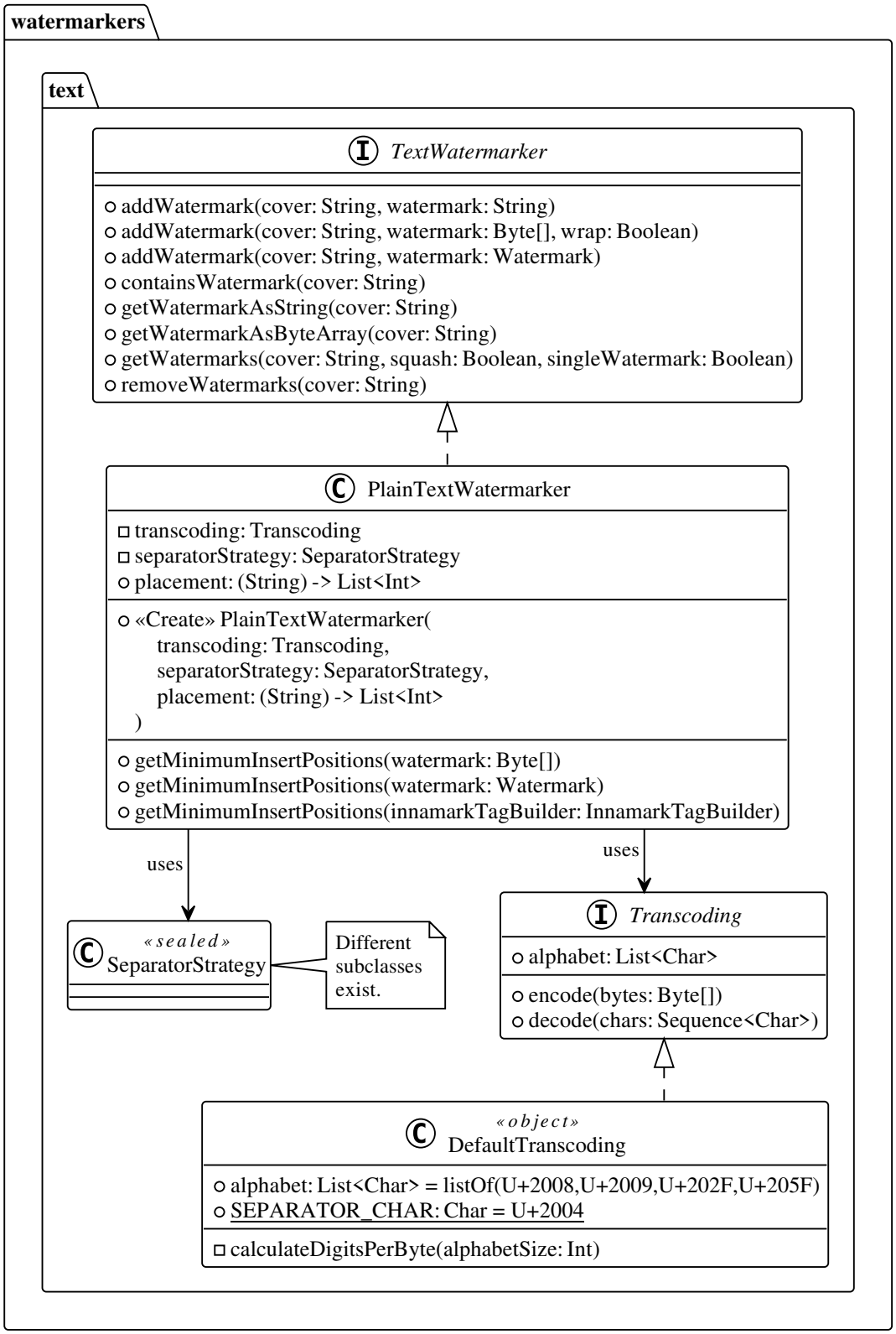


Fig. 4.8 Simplified Excerpt of Innamark's UML Class Diagram for Watermarkers

diagrams focus on the core structures, while we omit specific helper functionalities, return types, error classes, and inherited methods implemented in our reference implementation [75] to maintain readability.

The `watermarks` package in the Innamark library (see Figure 4.7) defines a watermark representation. The highest abstraction level is the `Watermark` class, with one implemented example in the `InnamarkTag` class, which defines our specific structure with its tag and prefix, as described in the previous section. While different options such as compression, error correction, and hashing are available, we follow the builder design pattern to "[s]eparate the construction of a complex object from its representation" [78, p. 97] in `InnamarkTagBuilder`.

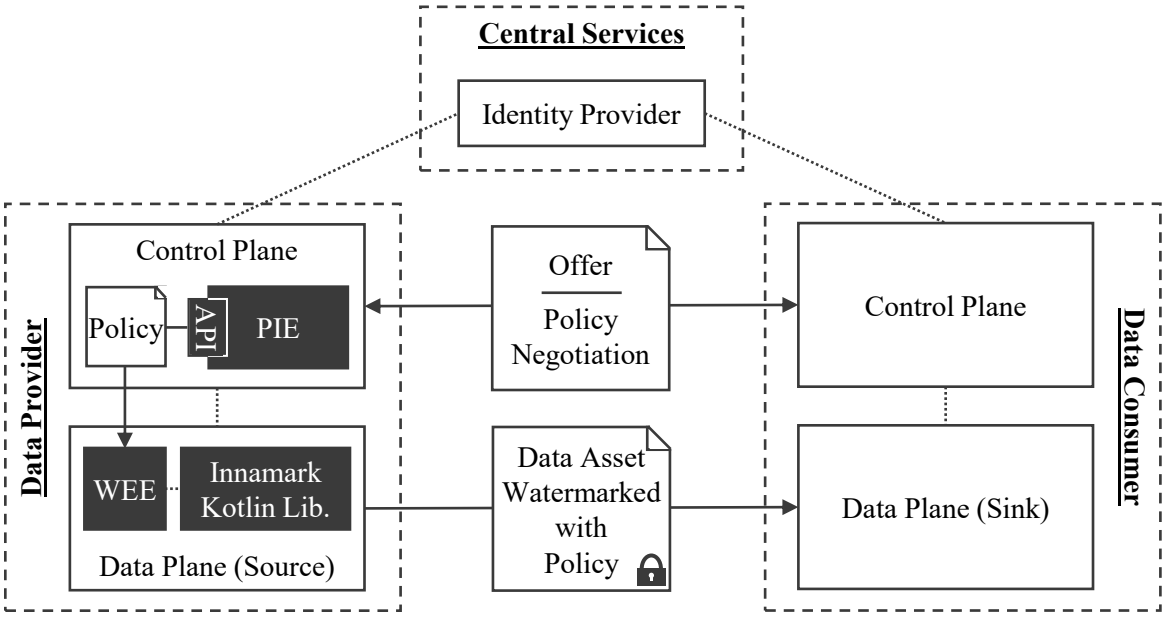
The `watermarkers` package uses watermarks from the Innamark library, which provides the actual embedding functionality. Figure 4.8 shows the `text` package with the `TextWatermarker` interface at the highest abstraction level, allowing the addition, checking, getting, and removal of watermarks from text. The `PlainTextWatermarker` implements this interface and embeds watermarks into plain text. If support for other text-based documents is necessary in the future, developers can implement extensions such as an `HTMLWatermarker` within the same `text` package. Those watermarkers use a `Transcoding` to transform Strings into a specific alphabet set, in our case $\mathcal{A}\_$, and a `SeparatorStrategy`, in our case by starting every watermark with the single separator character ϕ.

The modular architecture allows for easy integration of other watermark types and watermarking methods. If developers need to integrate our CSV watermarking algorithm presented in Paper VII [288], they can implement it using the `csv` package added to the `watermarkers` package.

Connector Extensions. We use our Innamark Kotlin library to develop two data space connector extensions that demonstrate its usage in practical software systems. Such data spaces are defined as "decentralized data infrastructures designed to enable data-sharing scenarios across organizational boundaries by implementing mechanisms for secure and trustworthy data sharing" [179, p. 6]. This federated infrastructure makes it an appropriate demonstration environment for the problem domain introduced initially. The core component of a data space is a *connector*, which enables data sharing possibilities based on the standardized Dataspace Protocol (DSP) [137]. A connector's architecture logically separates it into a *control plane* that coordinates the transfer process and a *data plane* that transfers the actual data [137].

Data spaces and their connectors have different advantages, while we mitigate some disadvantages by extending the concept. On the one hand, a significant benefit of these data

spaces is the ability to define policies that govern when data assets are accessed or used, thereby enabling downstream policy enforcement [290]. On the other hand, a major problem with the concept is that policies are stored *beside* the data, so the data assets themselves are only protected to a limited extent when considered together with the policies. Similar to the concept of sticky policies, which are typically processed over the full data life cycle [129], [174], [195], we extend the data space connector concept in Paper V [96] by watermarking policies directly *inside* the data asset cover. Since watermarking helps mitigate existing limitations in usage control [205], our results demonstrate how to address the *Access & Usage Control* challenge of Paper II [95].



Based on Paper V [96]

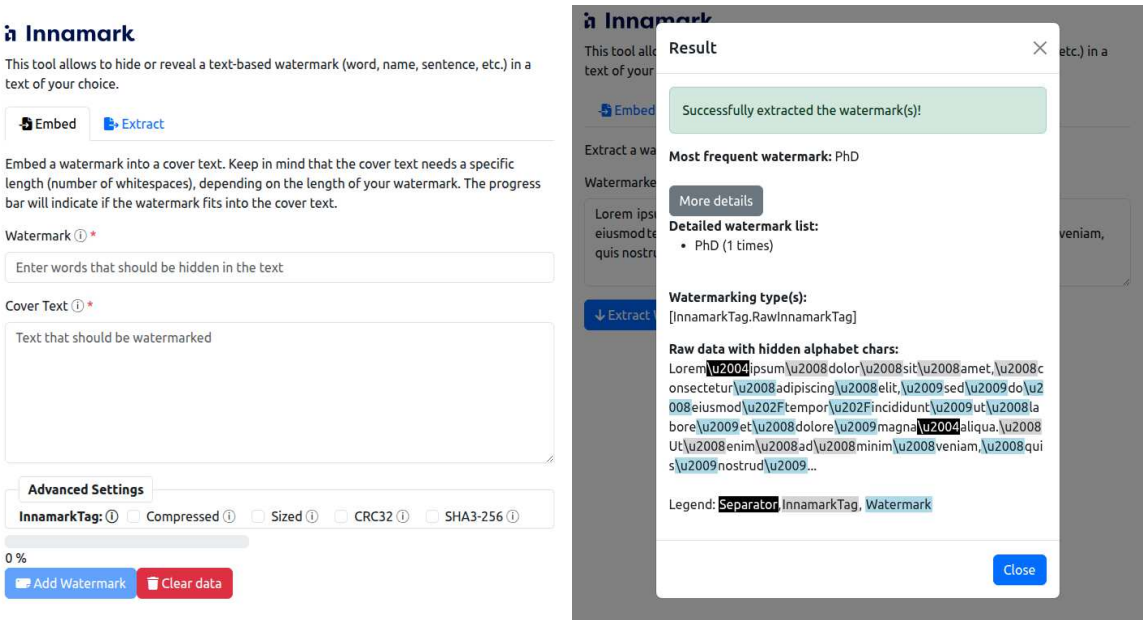
Fig. 4.9 Data Space Connector Architecture with Watermarking Extensions

We introduced two connector extensions for Innamark and tested them in an implementation testbed based on the Eclipse Dataspace Components (EDC) reference implementation version 0.5.1 [256] in Paper V [96]. Figure 4.9 shows the high-level architecture with our contribution displayed in dark boxes. We extend the provider's data plane with our Watermark Embedding Extension (WEE), which embeds the watermark into the data asset before pushing it to the consumer's data sink. This WEE uses our Kotlin library as a Java build target to embed text-based data assets, while it ignores all other media types. Since the data plane and control plane are architecturally separate, policy information is provided and negotiated on the control plane, but this information remains absent on the data plane. Therefore, we developed a Policy Information Extension (PIE) running in the control plane,

that provides an application programming interface (API) for external requests to access policy information. The WEE uses the provided API to retrieve policy information and embed it as a policy watermark within the data asset cover. While Figure 4.9 focuses on the push transfer, it works on the same principle for the pull transfer defined in the DSP [137].

The demonstration shows the integration of Innamark as an IT artifact inside data space environments. The lack of interoperability to strengthen data sovereignty is a major challenge in those ecosystems [55], especially when data is encrypted, and enforceable policies are stored alongside it [195]. By leveraging the extension mechanism, the developed extensions integrate directly into existing connector-based IT environments, mitigating the identified *Infrastructure & Landscape* challenge of Paper II [95]. Due to the integration on the provider side only, our solution has the significant advantage of working independently of the consumer side, making it interoperable in existing data spaces.

Web Interface. We demonstrated Innamark in a user-facing scenario by developing a graphical user interface (GUI) as a web interface. Since users have the possibility to compile the Innamark Kotlin library for Kotlin, Java, and JavaScript build targets, we used the latter to demonstrate its functionality in JavaScript projects. We used the KVision framework for web development in Kotlin and JavaScript for prototype development [120].



(a) Embedding Tab (b) Extraction Tab

Fig. 4.10 Screenshots of Innamark's Web Interface

Figure 4.10 shows two screenshots from a Chrome web browser of our web interface, deployed on an Ubuntu 24.04 server via a Docker compose setup. It consists of two tabs: the watermark-embedding tab for Algorithm 1, and the extraction tab, which provides the functionality of Algorithm 2. Figure 4.10a shows the embedding tab with an input field for the watermark $\mathcal{W}$, a textarea field for the cover text CT, an advanced settings checkbox section to set the configuration parameter θ for the InnamarkTag, and a progress bar that provides feedback on whether $\mathcal{W}$ fits inside CT. When clicking the *Add Watermark* button, a pop-up displays the watermarked cover text $CT_{\mathcal{W}}$. When copying out, the extraction tab provides a textarea field to insert and verify whether a possible text $CT_{\mathcal{W}}$ contains a watermark. Figure 4.10b shows the resulting pop-up of the same "PhD" watermark example inside a "Lorem ipsum" dummy text, introduced above in Figure 4.5. The web interface detects the watermark and provides additional information, such as the type of the InnamarkTag or the raw data, highlighting the special whitespaces. Our Kotlin library executes all watermarking-related tasks in the background.

CLI. After demonstrating the Innamark Kotlin library in a data space architecture and a JavaScript web interface, we also developed a CLI to demonstrate it in Java. Users typically map the resulting .jar file to an alias, such as `Innamark`, in shells like ZSH or Bash to watermark strings or text files directly in a terminal. Similar to the web interface, we make the source code available publicly in the same GitHub repository [75].

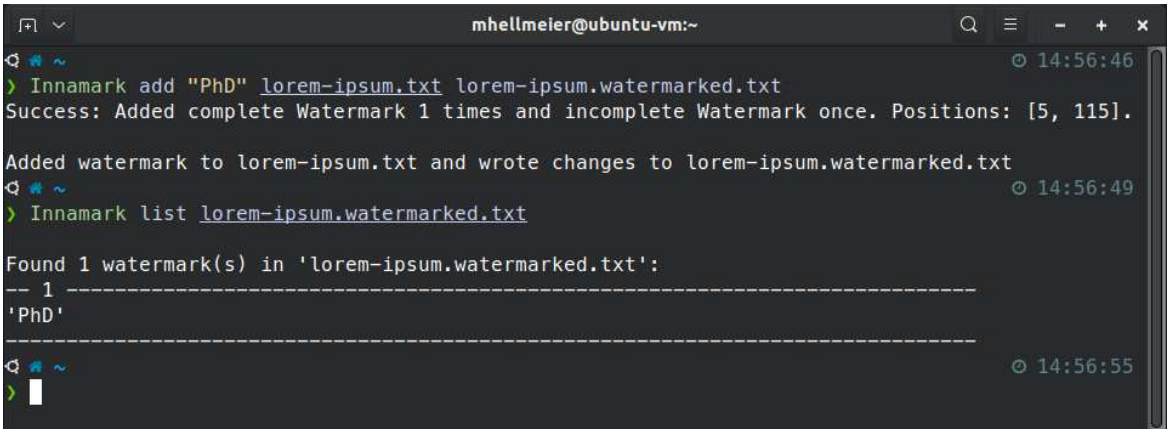


Fig. 4.11 Screenshot of Innamark's CLI Tool

Figure 4.11 shows an example of watermarking the "Lorem ipsum" cover text CT with the "PhD" watermark $\mathcal{W}$ in a terminal on an Ubuntu 24.04 system. By running the command `Innamark add "PhD" lorem-ipsum.txt lorem-ipsum.watermarked.txt`, the CLI tool creates a `lorem-ipsum.watermarked.txt` file as a watermarked output, embedding the "PhD" watermark in the cover text of `lorem-ipsum.txt` using our Kotlin library.

The second command demonstrates the extraction part, listing that one watermark was successfully extracted (see Figure 4.11). The CLI tool concludes the demonstration of a diverse set of relevant applications across various programming languages and environments.

In summary, we implemented and demonstrated Innamark as a Kotlin multiplatform library published on GitHub [75]. This library is used in three other demonstrations: (i) in connector extensions to demonstrate it in a data space context; (ii) in a web interface to test the JavaScript build target; and (iii) in a CLI tool to demonstrate the Java build target. Together, these demonstrations show that Innamark integrates seamlessly into diverse IT infrastructures and landscapes.

4.5 Evaluation

In this section, we present an evaluation of Innamark against the four KCs (see Table 4.2) to verify their validity and provide evidence [151]. We used *specification-based benchmarks* in our evaluation because they align with the initially introduced problem domain and require development before execution [139], [270]. We used typical benchmark criteria common for digital watermarking and information hiding [50], namely robustness, capacity, and imperceptibility [5], [19], [23], [126], [157], [286], which we mapped to our derived KCs:

- Firstly, we evaluated *KC1: Embedding* by checking the possibility of embedding a watermark inside a cover. Additionally, we compared the embedding capacity as the ratio of the watermark length $\mathcal{W}$ to the length of the cover text CT.
- Secondly, we evaluated *KC2: Invisibility* by comparing numerical metrics such as Jaro-Winkler Similarity, number of characters, and file size, as well as by checking for a caret navigation attack.
- Thirdly, we evaluated *KC3: Modification Robustness* by testing insertion, deletion, and replacement of watermarked text parts.
- Fourthly, we evaluated *KC4: Usage Robustness* by testing retyping and reformatting attacks as well as the robustness in various end-user applications.

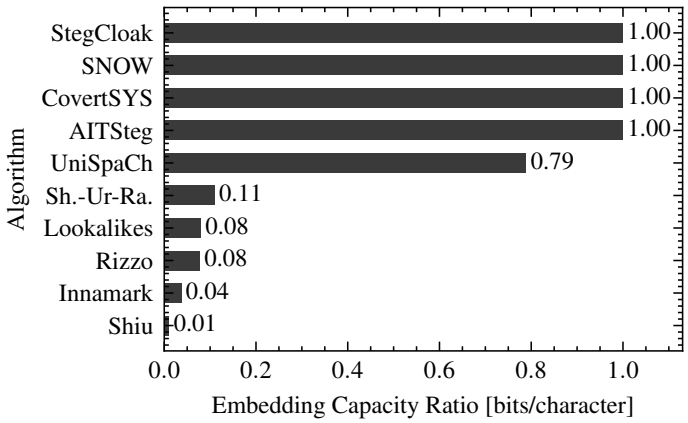
We used feedback from each evaluation iteration to improve the artifact. The development life cycle, which incorporated additional functionalities over time, demonstrated this targeted, improvement-oriented trend. One example is the improvement introduced by our InnamarkTag. During and after the first iteration, the evaluation of the proposed artifact at

that time, with its first Kotlin library version from Paper IV [99] and its first demonstration in a data space context from Paper V [96], identified drawbacks in limited embedding capacity, limited expandable interoperability, and limited robustness. The second iteration helped address those issues by introducing the InnamarkTag in Paper VI [94] for interoperability and additional functionalities, such as compression and hashing, to improve embedding capacity and robustness. In this section, we present the benchmark evaluation of Innamark after iteration two, structured by the existing metrics mapped to our KCs listed above and based on our publication of Paper VI [94].

We compared Innamark with the nine related algorithms from literature and practice (see Table 2.4) using a large-scale benchmark dataset. We implemented Innamark and all nine related algorithms in the Java programming language and used a test dataset of 1 000 000 English Wikipedia articles to provide a diverse test landscape, covering texts of different lengths, structures, and domains. The articles are publicly available from Hugging Face, based on the official Wikimedia Foundation dump in [280]. We randomly chose articles from the set and stored all article IDs for reproducibility. We conducted the benchmark evaluations on a Linux server, running Ubuntu 24.04. After analyzing all four KCs, this evaluation section concludes by summarizing and classifying the results.

KC1: Embedding. The first KC describes the user's ability to embed a watermark in a cover text. To address the technical data sovereignty challenges outlined in Paper II [95], selecting an appropriate watermark based on the application scenario is crucial. Using the policy of a contractual agreement, as demonstrated in the data space scenario in Paper V [96], helps address the *Access & Usage Control* challenge. In the *Identity Management* challenge, using the data provider's name as a watermark helps because the extractor can successfully identify the data owner during watermark extraction.

A sufficient *embedding capacity* is the first benchmark criterion, as we aim to embed policy or data provider names as watermarks [11], [126], [219]. The embedding capacity is calculated as "the average ratio between the number of embedded bits and the number of characters in each document" [219, p. 13]. Some algorithms, namely AITSteg, CovertSYS, StegCloak, and SNOW, use very small or zero-width characters and have theoretically unlimited embedding capacity when no content length restrictions apply, such as in X posts (formerly Tweets on Twitter) or as in Short Message Services (SMSs). We tested the other six algorithms on our dataset of 1 000 000 Wikipedia articles with an average length of ∼2514 characters. We determined the maximum possible embedding in a batch test job by applying each algorithm to the dataset and incrementally increasing the watermark length until an error occurred. In Figure 4.12, we summarize the results and set the unlimited-embedding-



Higher values are better. Based on Paper VI [94]

Fig. 4.12 Embedding Capacity Evaluation

capacity algorithms to 1.00 for maintaining consistency. Innamark embedded on average 93 bits per article, resulting in an embedding capacity ratio of 93/2514 $\approx$ 0.04 bits/character. This capacity is substantially lower than the best ratio for UniSpaCh (see Figure 4.12), but is sufficient to embed short identifiers.

All benchmark evaluations of Innamark use the default InnamarkTag ($\theta = \varnothing$), which resulted in a comparatively low embedding capacity. By enabling compression with zlib [77], Innamark can achieve a significantly higher embedding ratio. Since compression highly depends on the watermark used and applies equally to all algorithms, we did not enable it in our benchmark evaluation. This choice ensures consistency and comparability, and avoids artificially enhancing Innamark's performance.

KC2: Invisibility. The second KC focuses on the non-noticeability or imperceptibility of a watermark hidden inside a cover [5], [11], [126], [135], [157]. Related to the technical data sovereignty challenges described in Paper II [95], users work across diverse *Infrastructure & Landscapes*, whereas text data has different *Data Processing Life Cycles*. A good watermarking system needs to maintain invisibility even under different treatments and approaches. We evaluated this invisibility by analyzing and benchmarking:

- the Jaro-Winkler Similarity;
- the number of characters;
- the file size;
- the caret navigation attack.

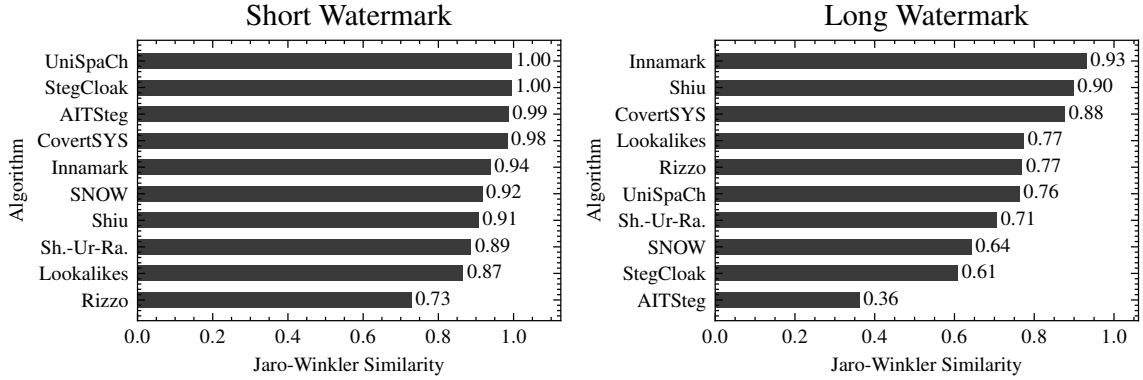
Firstly, the *Jaro-Winkler Similarity* is a widely used String comparison metric, often applied in watermarking and steganography evaluations [4], [6], [9], [19], [164], [165], [184]. Let $|s_1|$ and $|s_2|$ be the length of two strings, c the number of matching characters, τ the number of transpositions based on the characters, and m all matching characters [130], [281]. The Jaro Similarity $0 \leq \Phi \leq 1$ in Equation 4.2 indicates the amount of similarity of those two Strings s_1 and s_2 [281]:

$$\Phi(s_1, s_2) = \begin{cases} 1 & : s_1 = s_2 \\ \frac{1}{3}\left(\frac{c}{|s_1|} + \frac{c}{|s_2|} + \frac{c-\tau}{c}\right) & : m > 0 \\ 0 & : \text{otherwise.} \end{cases} \tag{4.2}$$

The updated Jaro-Winkler Similarity $0 \leq \Phi_n \leq 1$ in Equation 4.3 adds a fixed weighting by the constant value of 0.1 as suggested in [281]:

$$\Phi_n(s_1, s_2) = \Phi(s_1, s_2) + i \cdot 0.1 \cdot (1 - \Phi(s_1, s_2)). \tag{4.3}$$

We conducted two batch test jobs on all 1 000 000 Wikipedia articles. The first job attempts to hide the short, four-character-long watermark $\mathcal{W}$ = "John" within the articles. In contrast, the second job attempts to hide a long 455-character watermark $\mathcal{W}$ = "Lorem ipsum (…) id est laborum." inside the articles. Afterward, we calculated the Jaro-Winkler Similarity Φ_n for all successful watermark operations to compare the similarity of the input article CT with the watermarked output $CT_{\mathcal{W}}$. Figure 4.13 summarizes the results of the average Φ_n for the short- and long-watermark batch jobs. In both cases, Innamark had, on

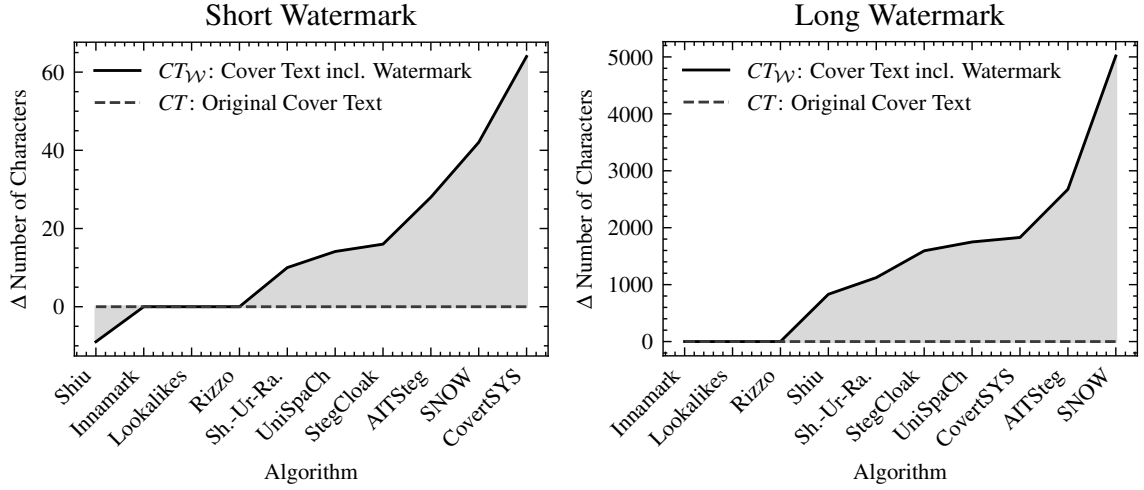


Higher values are better. Based on Paper VI [94]

Fig. 4.13 Jaro-Winkler Similarity Evaluation

average, a similarity of $\Phi_n > 0.9$. For short watermarks, Innamark is in the upper half, and even the best solution with $\Phi_n \approx 0.93$ for long watermarks.

Secondly, an unusual *number of characters* can reveal a watermarked text. For example, suppose a user copies a watermarked sentence into an application such as Microsoft Word. In that case, it raises suspicion when the number of characters displayed in the footer is far too high due to hidden zero-width characters. In Figure 4.14, we compared the average difference Δ in the number of characters between the original cover text CT and its watermarked version $CT_{\mathcal{W}}$. We evaluated all articles, with the first batch job on the short four-character watermark and the second on the long 455-character watermark. The negative value of Shiu et al. indicates a decrease in character size because the algorithm deletes whitespaces and adds a new line separator, which is not considered a character. As seen, only the replacement-based techniques, Innamark, Lookalikes, and Rizzo's fine-grain watermarking, showed the best results by not increasing the number of characters.

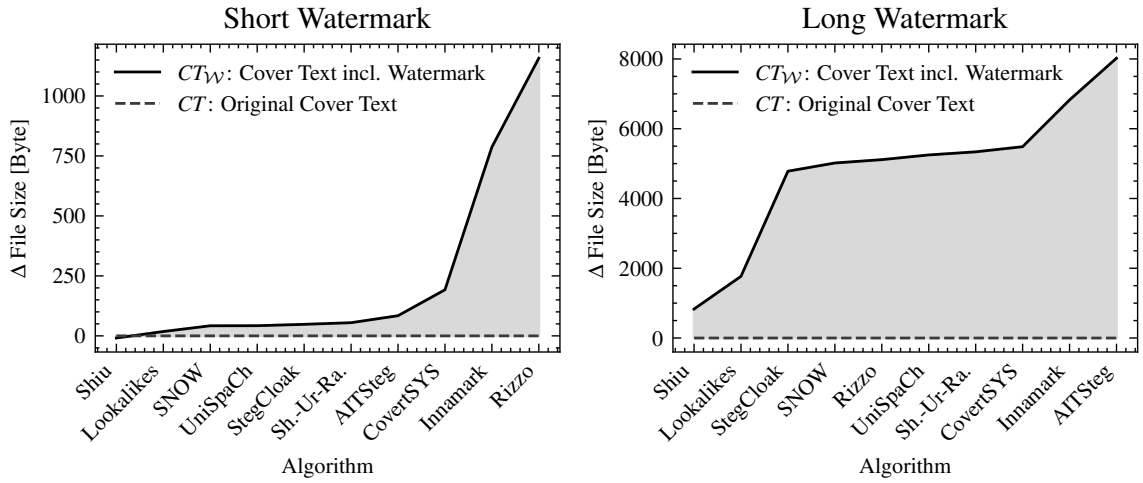


Values close to 0 are better. Based on Paper VI [94]

Fig. 4.14 Character Size Evaluation

Thirdly, a comparatively high *file size* can reveal a watermarked text file. Similar to the previous number of character evaluation, we ran two batch jobs to compare the file sizes of CT and $CT_{\mathcal{W}}$ across all algorithms. Figure 4.15 depicts the results for the short four-character and long 455-character watermark batch jobs. Despite the apparent similarity to the number of character test, the results differ entirely. While algorithms like Innamark and Rizzo et al. performed much better than StegCloak in the previous test, the results in these file-size tests were reversed. This reversed result is related to the different storage requirements for UTF-8-encoded characters, as seen in the examples in Table 2.5, and to the em-

bedding technique used. The classical whitespace δ (U+0020) needs one byte of storage in UTF-8, while a zero-width space like U+200C or spaces like the thin space (U+2009) need three bytes of storage (see Table 2.6). Further, Rizzo et al. and Innamark insert the watermark multiple times to increase robustness, which, however, results in the worst performance in the file-size benchmark for short watermarks (see Figure 4.15). StegCloak performs better in the file-size comparison, since it adds the watermark to only one position, without the additional storage overhead of an InnamarkTag.



Lower values are better. Based on Paper VI [94]

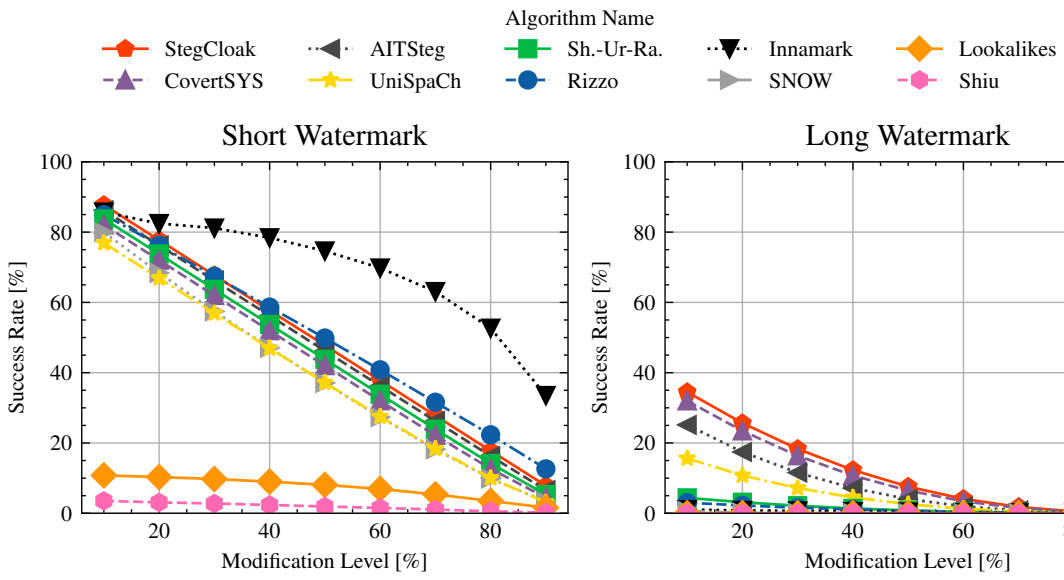
Fig. 4.15 File Size Evaluation

Fourthly, we evaluated the visibility by using the *caret navigation attack*. Digital text documents opened in an editor have a blinking cursor, the caret, indicating the current editing position. End users are able to use the arrow keys on a keyboard to navigate the caret through the text and perform edits. However, those carets remain in place without visibly moving between zero-width characters. While the caret technically moves one character further, the user gets suspicious without noticing the expected cursor movement. We tested all algorithms by manually checking a $CT_{\mathcal{W}}$ against this attack. It results in only Rizzo et al., Lookalikes, and Innamark staying invisible since they all use a one-to-one replacement approach. All other algorithms add extra characters, leading to anomalies such as multiple characters identified or hanging caret positions.

KC3: Modification Robustness. KC3 and KC4 focus on the robustness of a watermarking scheme as one of the three main dimensions, along with embedding capacity and invisibility [5], [157]. Our *KC3: Modification Robustness* focuses on the user's text modifications,

such as adding new sentences, words, or symbols, and deleting or modifying existing content. Every data asset has its life cycle in the so-called *big data value stream* from creation or acquisition through storage and usage to deletion [76]. Maintaining data sovereignty throughout its life cycle is crucial, as identified in Paper III [271], but remains challenging due to the *Data Processing Life cycle* challenge outlined in Paper II [95]. Consequently, the robustness against modifications is one of the most important criteria for our MRQ to strengthen data sovereignty through digital watermarking.

In the literature, modifications to watermarked texts, such as deletions or insertions, are classified and benchmarked as tampering attacks [5], [6], [23], [117]. In our analysis, we used our two watermarked samples of 1 000 000 Wikipedia articles each, one with the short four-character watermark and the other with the long 455-character watermark, for all ten algorithms in the testbed. For each article, we selected a random position in the text, deleted a block corresponding to 10% of the text length, and added the same amount of data by repeating a single ASCII letter. Afterward, we applied the extraction algorithms to determine whether the original watermark $\mathcal{W}$ was successfully recovered from the modified cover text. We repeated the evaluation with text modifications of 20%, 30%, and so on up to 90%. Figure 4.16 summarizes the results of all 18 cases (nine for the short watermark, and nine for the long watermark).



Higher values are better. Based on Paper VI [94]

Fig. 4.16 Modification Robustness Evaluation

Due to continuous improvements during the iterative development of Innamark with multiple watermark insertions and the InnamarkTag, Innamark is by far the most robust method against modifications for short watermarks (see Figure 4.16). Even if 80% of a watermarked cover text $CT_{\mathcal{W}}$ is completely modified, it is still possible to successfully extract the watermark in more than half of the cases. For long watermarks, the results for all compared algorithms are very close together and, in all cases, below 40% success rate.

KC4: Usage Robustness. The last KC focuses on the robustness of a watermarked text in use. The challenges of Paper II [95] indicate a need for more *Access & Usage Control*, even when data is used in a diverse IT *Infrastructure & Landscape*. Taking the data value chain and its life cycle into account [76], data assets in real-world use cases move between different applications and users. The goal of applying a watermarking scheme with high robustness is to ensure persistence across all life cycle and application stages.

Since robustness is a significant factor in this work, we conducted a usage robustness evaluation of existing Unicode whitespace during the development of Innamark, as previously discussed in the *Design & Development* Section 4.3 (see Table 4.5). For evaluation and comparison, we watermarked a Lorem ipsum dummy text with all algorithms. Afterward, we used the copy-and-paste attack [5], [126] to test it across three file formats (.txt, .docx, and .pdf), an Outlook email client, a Microsoft Teams chat message, and three social media applications, following Shazzad-Ur-Rahman et al. [235]. Kamaruddin et al. [126] highlight the importance of such document transformations as a result of their review, since robustness is essential if the file format or content changes. Additionally, we evaluated the usage robustness for a retyping and reproduction attack [5], [116] by manually typing a watermarked cover text $CT_{\mathcal{W}}$ into a new document. Lastly, we tested all algorithms against a reformatting attack [5], [8] by checking whether the watermark is extractable after changing the font style, size, and color in Microsoft Word.

Table 4.6 shows that all tested algorithms fail the retyping attack and resist the reformatting attack. These results relate to the nature of format-based watermarking techniques. In contrast, linguistic-based techniques would withstand a retyping attack because they work, for example, by using synonym replacement rather than character changes. Comparing usage robustness across different file formats and applications, only Innamark remains robust in almost all tested cases. A '(✓)' in Table 4.6 indicates that there are some cases of successful and unsuccessful watermark extractions during our evaluation, consistent with the initial results in Table 4.5. For example, the PDF24 Reader successfully extracts a text watermarked by Innamark from a .pdf document, but Adobe Acrobat Reader destroys it. The results of the other algorithms vary widely because they use different characters that are supported only

Table 4.6 Usage Robustness

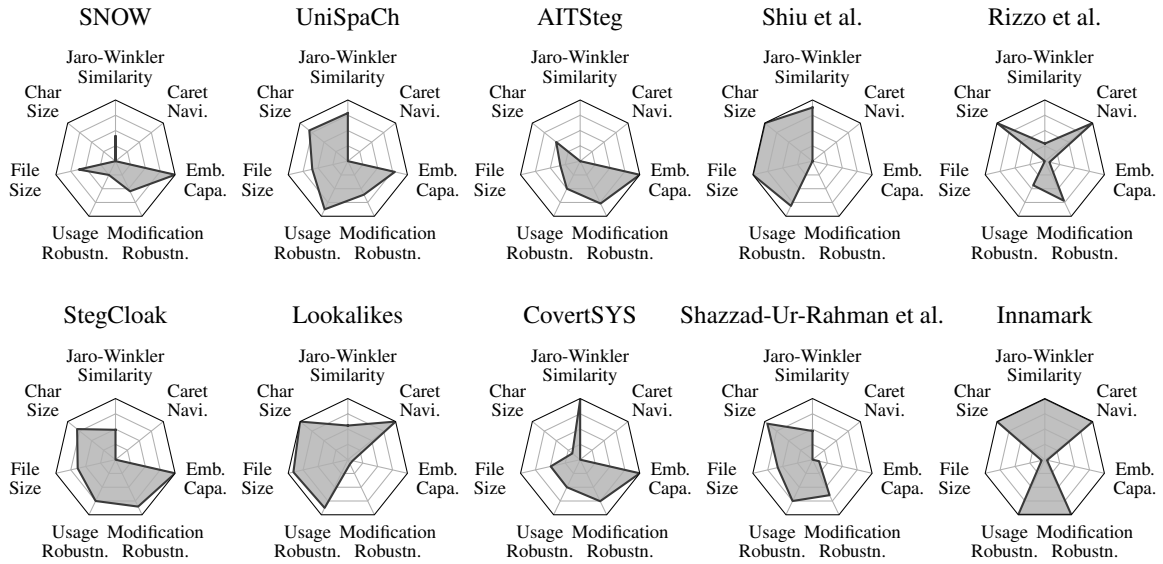
Algorithm	Retyping	Formatting	.txt	.docx	.pdf	Mail	Teams	WhatsApp	Facebook Msg.	X/Twitter
SNOW	✗	✓	✓	✓	✗	✗	✗	✗	✗	✗
UniSpaCh	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓
AITSteg	✗	✓	✓	✗	✗	✗	✓	✓	✓	✗
Shiu et al.	✗	✓	✓	✓	✗	✓	✓	✓	✓	(✓)
Rizzo et al.	✗	✓	✓	✗	✗	✗	(✓)	✓	✓	✗
StegCloak	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓
Lookalikes	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓
CovertSYS	✗	✓	✓	(✓)	✗	✗	✓	✓	✗	✗
Shazzad-Ur-Rahman et al.	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓
Innamark	✗	✓	✓	✓	(✓)	✓	✓	✓	✓	✓

Based on Paper VI [94]

by specific applications. Only the SNOW algorithm does not work in any of the applications tested, as its embedding mechanism relies on tailing whitespaces. Most applications filter and remove those additional spaces [220].

Summarizing Evaluation Results. Figure 4.17 summarizes our evaluation across the four benchmarking KCs: embedding, invisibility, modification robustness, and usage robustness (see Table 4.2). Here, we split the invisibility into our four tested dimensions: Jaro-Winkler Similarity, differences in character size, differences in file size, and the caret navigation attack. Figure 4.17 ranks the results, with a data point at the edge of the radar diagram for the best algorithm and one in the middle for the worst. The caret navigation dimension has only a Boolean value, with no gradation in between.

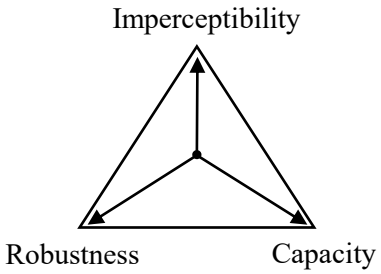
Innamark shows the best results in both robustness dimensions, a comparatively low result for the embedding capacity, and a mixed result for the invisibility dimensions, as indicated by the bottom-right radar chart in Figure 4.17. However, improvements are possible to further increase the embedding capacity, e.g., by enabling compression of the Innamark-Tag. Nevertheless, Innamark will consistently achieve lower results than algorithms with theoretically unlimited embedding capacity, such as StegCloak, SNOW, CovertSYS, and AITSteg. Upon closer inspection of the results in Figure 4.17, it became apparent that no algorithm performs best or nearly best in all areas. These differences reflect the trade-off



Based on Paper VI [94]

Fig. 4.17 Algorithm Comparison

triangle between the counteracting dimensions of imperceptibility, robustness, and capacity (see Figure 4.18) [157]. Related to the effectiveness part of RQ 2.2, the goal of all watermarking algorithms is to find a trade-off between these three opposing dimensions [5], [157]. For this reason, we prioritized the imperceptibility and robustness dimensions over embedding capacity, as stated in RQ 2.1, during the design and development of Innamark. Since our results showed that a perfect solution across all dimensions is currently not possible, the goal is to focus on improving specific dimensions rather than all of them. We discuss these findings and their consequences for research and practice in the upcoming Section 5.



Following Li, Wang, and Barni [157] and based on Paper VI [94]

Fig. 4.18 Trade-off Triangle

In summary, we evaluated Innamark in a testbed comprising ten algorithms (nine from related work and Innamark) by running multiple experiments on a dataset of 1 000 000 Wikipedia articles. We used the derived four KCs extended by related work as benchmark criteria: (i) the embedding capacity; (ii) the invisibility, measured by the Jaro-Winkler Similarity, the number of characters, the file size, and the caret navigation attack; (iii) the modification robustness; and (iv) the usage robustness, measured in different applications and file formats and against a retyping and reformatting attack. Innamark showed the best results across both robustness dimensions with limited embedding capacity, while our comparison chart summarizes the strengths and weaknesses of all tested algorithms and positions Innamark in the current landscape.

Chapter 5

Discussion & Design Knowledge

In this chapter, we analyze and interpret our two main contributions: the conceptual model for data sovereignty and Innamark, a novel digital text-watermarking system addressing current data sovereignty challenges. These interpretations, combined with the design knowledge gathered during iterative development, are generalized and formalized as DPs, constituting our third main contribution. At the same time, we discuss the implications for theory and practice, as well as limitations, and future research opportunities.

Data Sovereignty. We observed an increasing interest and discourse in politics, academia, and practice regarding data sovereignty, with a lack of differentiation among adjacent fields. We followed the advice of existing publications, asking the academic community to structure the field [2], [48]. Consequently, our results complement existing work in two key ways: first, by focusing on the information systems and software engineering domain, where concepts like technological sovereignty are important but not addressed in the review by Hummel et al. [108]; and second, by incorporating practical experiences from an interview study with experts from different companies. These steps were necessary since, at the time of publication from Paper I [100], Paper II [95], and Paper III [271], the topics were relatively new, with limited reviews and discussions.

The findings indicate that current challenges in data sovereignty are rooted in the three clusters of (i) organizational, (ii) technical, and (iii) personal & emotional challenges [95], rather than a single loss of control only [183]. While this thesis provides a first step toward designing an artifact to address the technical challenges, the other two clusters must also be considered and may influence the technical aspects. Our conceptual model clarifies that data sovereignty issues typically involve a settlement of an agreement between two parties: a data provider offering data assets and a data consumer using them. We confirm the importance of trust levels in data sovereignty. Even if we build technical solutions to increase overall trust, which directly leads to greater data sovereignty, a baseline level of trust is always needed. In our case, trust within Innamark and its implementation is an important prerequisite, as we postulate that data sovereignty cannot exist in the face of constant mistrust.

Digital Watermarking. To address the identified data sovereignty challenges, we designed and implemented our technical IT artifact, Innamark, in Paper IV [99], demonstrated it in Paper V [96], and further developed and evaluated it in Paper VI [94]. Innamark aims to increase trust and strengthen data sovereignty among participants. By extending the work by Rizzo, Bertini, and Montesi [219], we contributed a novel watermarking system, focused on text covers, an under-researched and most complex task [26], [30], [126]. The interpretation of Innamark's design decisions highlights its advantages over related work. First, Innamark is more robust and does not increase the number of characters compared to existing zero-width character methods [3], [9], [144], [176] due to our one-to-one replacement technique. Second, by substituting classical whitespaces with a specific set of similar-looking Unicode whitespaces, Innamark outperforms existing replacement-based techniques [218], [219], [235], [236], [258]. Our approach supports multiple applications while remaining robust and language-independent, as it does not rely on specific language constructs, making it imperceptible.

Nevertheless, each application scenario has its own framework conditions and requirements. By adopting Innamark, e.g., by adjusting the whitespace alphabet used or extending the InnamarkTag, practitioners can achieve better outcomes. They can easily implement those specific adjustments, due to our abstract software architecture and the published reference implementation on GitHub [75]. We hypothesize that our ideas of using watermarking to strengthen data sovereignty could also serve as an enabler for other domains, as shown for sensor data by related work [222] and in our Paper VII [288] for CSV files.

In the sections below, we derive design knowledge, discuss the resulting implications for theory and practice, and consider limitations and future research opportunities.

5.1 Design Principles

This section derives and evaluates DPs as design knowledge for a more abstract level, guiding the development of watermarking applications to strengthen data sovereignty. In the upcoming first part, we describe our methodology and the framework conditions for deriving and extracting DPs based on existing models and experiences from related work. This description includes our application of DP dimensions, the formulation format used, and a classification based on an applied DP development taxonomy. In the upcoming second part, we present the derived DPs, each in the same structured format. The final part evaluates the DPs using the reusability framework by Iivari, Hansen, and Haj-Bolouri [112].

Design Principle Methodology & Framework Conditions. A DP is defined as "a statement that prescribes what and how to build an artifact in order to achieve a predefined design goal" [39, p. 4040]. Over time, various concepts and recommendations have emerged on when, how, and with whom to build DPs. We follow the structured framework from Purao, Kruse, and Maedche [206] by considering the four relevant domains of influence (what influences DPs?), temporality (when to generate DPs?), actors (who identify DPs?), and content (how to document DPs?) (see Table 5.1).

Table 5.1 Applied DP Dimensions

Dimension	Our Application
Influences	We created the final version of DPs based on the results of our artifact design effort and benchmark evaluations of related methods.
Temporality	We created the DPs after the artifact's design and implementation, also known as the reflective perspective [24], [178].
Actors	We, as the research team, formulated the DPs.
Content	We used the skeleton format from Chandra, Seidel, and Gregor [39] and the schema by Gregor, Kruse, and Seidel [87] to formulate the DPs.

Following Purao, Kruse, and Maedche [206]

The first *influence* dimension focuses on the key sources [206], specifically the results from our iterative artifact design, combined with the gathered knowledge from the benchmark evaluation metrics and results. The following two dimensions, *temporality* and *actors*, focus on when and by whom DPs are created [206]. In our case, we, as the research team, formulated the DPs as the suggested norm within the DSR community [206]. In contrast, we made the temporal discovery after the artifact's design and evaluation, as suggested by Baskerville et al. [24], a perspective also known as the reflective perspective [178]. The last *content* dimension focuses on the formulation format used to document DPs [206]. Initial formulation formats, such as those in van den Akker [265], were updated by Chandra, Seidel, and Gregor [39], who analyzed 43 DSR articles to derive an effective DP formulation pattern. In the upcoming presentation of our DPs, we applied their suggested formulation structure:

> "Provide the system with **[material property—in terms of form and function]** in order for users to **[activity of user/group of users—in terms of action]**, given that **[boundary conditions—user group's characteristics or implementation settings]**." [39, p. 4045]

To extend the formulation structure, we used and applied the conceptual DP schema by Gregor, Kruse, and Seidel [87] in tabular form. It is even more detailed and includes (i) the four *actors* of implementers, users, enactors, and theorizers; (ii) the *context* with its conditions, settings, and characteristics; (iii) employed *mechanisms*, like forms, activities, or processes; and (iv) an optional *rationale* for empirical or theoretical justification [87]. Due to the wide range of DP development possibilities, we applied the taxonomy introduced by Möller, Guggenberger, and Otto [178] (see Figure 5.1), with our applied characteristics highlighted in a dark background for each dimension.

Dimension (D_n)	**Characteristics (C_{nm})**				
D_1: Perspective	Supportive			Reflective	
D_2: Research Design	DSR	A(D)R		Qualitative	Case Study
D_3: MR Source	Literature	Theory	Interviews	Workshops/ Focus Groups	None
D_4: DP Design	Derived		Extracted		Responsive
D_5: Iterations	Single			Multiple	
D_6: Evaluation	Expert/User Feedback		Instantiation/Field Testing		Argumentation
D_7: Formulation	Free			Based on Template	

Based on Möller, Guggenberger, and Otto [178, p. 211]

Fig. 5.1 Applied DP Development Taxonomy

Derived Design Principles. In the following, we present our derived DPs, each on a single page and following the same structure. It starts with the description in the suggested skeleton format of Chandra, Seidel, and Gregor [39], shown in the gray boxes on top of every page. Every DP includes additional *explanations* for more context; concrete *examples* connected to Innamark and related methods from the literature and practice (see Section 2.2.3); resulting *implications*, related to our KCs (see Table 4.2); and the identified technical data sovereignty challenges of Paper II [95]. Every DP concludes with a table of additional information for the actors, context, mechanism, and rationale based on the schema by Gregor, Kruse, and Seidel [87].

DP1: Principle of Embedding Strategy
Provide the system with a deterministic, blind embedding mechanism based on Unicode characters in order for users to embed a watermark within plain cover text, given that the extraction works interoperably and independently from the embedding system.

Explanations: The first DP relates to the embedding process of a watermark into a cover text and its relation to the extraction. Figure 2.4 shows the differences between an encoder and decoder for the two separate, disconnected subsystems [6], [50], [177]. This non-connectedness stems from the blindness characteristic, which allows watermark extraction without knowledge of the original, unwatermarked cover text [210], [246], which is important for our data sovereignty background and multiple application scenarios across diverse IT infrastructure landscapes.

Examples: Innamark and all evaluated, related methods are deterministic and blind embedding techniques, working with Unicode characters and plain text. Thus, all fulfill the blindness criteria and our DP1.

Implications: The DP addresses our *KC2: Invisibility* for text-based data assets. To overcome the related technical *Access & Usage Control* and *Identity Management* challenges identified in Paper II [95], a watermarking system must be able to embed specific identity information, contractual agreements, or policy information within a text worthy of protection.

Table 5.2 DP1: Principle of Embedding Strategy

Structure	Application
Aim, implementer, and user	For designers, researchers, and developers (implementers) to enable watermarking of plain text (aim) with an embedding and extraction subsystem (enactor) by any party without access or knowledge of the embedding during extraction (user).
Context	In any software component processing Unicode text that needs to be extended by watermarking functionalities.
Mechanism	Identify and determine a deterministic and blind embedding mechanism based on Unicode characters.
Rationale	Because it enables independence from the original watermark and embedding subsystem during extraction, offering application independence and interoperability.

Structure following Gregor, Kruse, and Seidel [87]

DP2: Principle of Alphabet Selection
Provide the system with an alphabet of a fixed set of Unicode characters for watermark embedding in order for users to work with text across existing systems, applications, and IT infrastructure landscapes without limitations, given that the balance between imperceptibility and embedding capacity is suitable.

Explanations: Format-based text watermarking systems need to define an alphabet for embedding. There should be a consciously selected set tailored to the conditions, such as focusing on zero-width characters when embedding capacity is crucial or prioritizing confusables when imperceptibility is vital. They need to consider the support for end-user applications, as not all systems and file types support all Unicode characters (see Table 4.6).

Examples: Innamark uses a set of five similar-looking Unicode whitespaces (see Table 4.5). Related algorithms rely on an alphabet of classical whitespaces [147], [148], [238], Unicode confusables [258], smaller spaces [204], zero-width characters [3], [9], [144], [176], or different combinations [218], [219], [235], [236].

Implications: This DP addresses *KC2: Invisibility* and *KC4: Usage Robustness* because a well-chosen alphabet set addresses the related technical challenges of the watermark's persistence across different IT *Infrastructure & Landscape* and throughout the whole *Data Processing Life Cycle*, as discussed in Paper II [95].

Table 5.3 DP2: Principle of Alphabet Selection

Structure	Application
Aim, implementer, and user	For designers, researchers, and developers (implementers) developing a watermark system (enactor) that needs to determine the balance between imperceptibility and embedding capacity (aim) for employees working with texts (users).
Context	In existing programs and IT landscapes, such as Microsoft Office or pre-installed mail clients, without limitations.
Mechanism	Identify and determine an alphabet of a fixed set of Unicode characters supported by the context.
Rationale	Because the watermark persists in the applications used in a specific IT infrastructure landscape. The selected characters focus either more on a high embedding capacity (e.g., by using zero-width characters) or on high imperceptibility (e.g., by using Unicode confusables).

Structure following Gregor, Kruse, and Seidel [87]

DP3: Principle of Multiple Insertions

Provide the system with multiple watermark insertions, distributed throughout the text during embedding, in order for users to extract the watermark due to a high modification robustness, given that the watermarked text has been edited, such as rewritten, partly copied, or replaced.

Explanations: Text can be easily altered or distributed by any user without specialized knowledge. If a watermark is inserted only once within a cover text, changing a single character destroys it, leading to a corrupt version that cannot be successfully extracted. Instead, the DP suggests inserting the watermark multiple times at different locations to increase the likelihood of finding an unaltered watermark during extraction after modification.

Examples: Only Innamark and the updated version of Rizzo, Bertini, and Montesi [219] focus on multiple watermark insertions. All other tested algorithms only embed the watermark once.

Implications: The DP addresses the *KC3: Modification Robustness* through multiple insertions. It relates to the *Data Processing Life Cycle* challenge in Paper II [95], by keeping the watermark persistent throughout creation, usage, sharing, and storage [76], [213].

Table 5.4 DP3: Principle of Multiple Insertions

Structure	Application
Aim, implementer, and user	For designers and researchers (implementers) to ensure robustness against text modifications (aim) in different text applications like text processors, chat applications, or mail clients (enactor) when end-users work and modify a watermarked text (users).
Context	In different editing scenarios, such as rewriting, copying only parts, or specific replacements.
Mechanism	Ensure to embed the watermark multiple times, distributed in different places throughout the text.
Rationale	Because it offers high modification robustness, allowing successful watermark extraction even if parts of a watermarked text are altered, replaced, or deleted.

Structure following Gregor, Kruse, and Seidel [87]

DP4: Principle of Modularity

Provide the system with a modular, structured tag that stores the configuration used during embedding, in order for users to enable a combination of functionalities such as compression, encryption, or verification, given that the extraction works across diverse environments with different application scenarios and requirements.

Explanations: Different use cases have different requirements for watermarks, ranging from long watermarks that need to be embedded in short cover texts (requiring compression) to highly confidential watermarks (requiring encryption) to integrity covers (requiring verification). Offering modularity via a configuration tag keeps the blindness criterion active because the extraction subsystem identifies the watermark type independently of the embedding subsystem.

Examples: Reviewing related methods, AITSteg [3] builds its algorithm on the sending time, while CovertSYS [9] uses a one-time pad. Only Innamark offers the option to use different functionalities, such as compression or verification, compared to the other analyzed algorithms.

Implications: The DP addresses the *KC4: Usage Robustness*, which is related to the *Access & Usage Control* and *Infrastructure & Landscape* challenges of Paper II [95]. Using a tag to store the configuration of enabled functionalities enables correct interpretation during watermark extraction in another system, thereby facilitating access and usage control enforcement.

Table 5.5 DP4: Principle of Modularity

Structure	Application
Aim, implementer, and user	For designers and researchers (implementers) to extend the watermarking system with a combination of additional functionalities, such as compression, encryption, or verification (aim), for special users with high data requirements (user).
Context	In diverse environments with a need to support different application scenarios with different requirements.
Mechanism	Ensure embedding a structured tag, storing the configuration used during embedding.
Rationale	Because it enables independence between watermark embedding and extraction, making it modular.

Structure following Gregor, Kruse, and Seidel [87]

Evaluation. After deriving our DPs, we evaluated them in a final step. DPs are often formulated inconsistently or incompletely, leading to limited reusability, whereas projectability is desired [25], [39], [112]. As a preventive measure, we evaluated our four DPs against the five criteria of the reusability evaluation framework proposed by Iivari, Hansen, and Haj-Bolouri [112]. Table 5.6 presents the evaluation results, with the criteria focusing on:

- *Accessibility:* The degree of understandable formulation with an easy language, making sense for practitioners [112].
- *Importance:* The relevance in practical problems in specific domains [112].
- *Novelty & Insightfulness:* The innovation potentials and possibility to surprise practitioners with new insights [112].
- *Actability & Guidance:* The practice-readiness with clear guidelines in an actable format [112].
- *Effectiveness:* The resulting intended or unintended positive effects and consequences for usage in practice [112].

The first two DPs of the embedding strategy and alphabet selection are important and highly relevant principles. However, they offer minor novelty due to their basic foundation and already established use. The last two DPs of multiple insertions and modularity, however, exhibit greater novelty in the watermarking domain. In contrast, their importance depends heavily on the conditions of the application scenarios and the IT infrastructure and landscape used.

Table 5.6 DP Evaluation

	DP1: Embedding Strategy	DP2: Alphabet Selection	DP3: Multiple Insertions	DP4: Modularity
Accessibility	Clear and individual, as terms like "deterministic" or "blind" are rather academic and may be unfamiliar to practitioners.	Clear and individual, the "alphabet" term may be unknown to practitioners.	Clear and individual.	Individual, as a "structured tag" is not clear to practitioners.
Importance	High relevance due to the core goal of text watermarking.	High relevance, because an alphabet is a needed factor that must be determined initially.	High relevance for high-robustness scenarios; low relevance for high-embedding scenarios.	High relevance for diverse IT infrastructure and landscapes; high application-scenario relevance; low relevance for highly specific systems and use cases.
Novelty & Insightfulness	Moderate novelty, since strategy selection is always necessary.	Moderate novelty, while characters exist, but their behavior differs.	Moderate to high novelty, as multiple insertions rarely occur.	High novelty, since modularity over tags is a new concept in watermarking.
Actability & Guidance	Clear actionable tasks and goals, low guidance due to missing strategy development insights.	Clear actionable tasks and goals, moderate guidance due to missing Unicode restrictions for alphabet selection.	Clear actionable tasks and goals, moderate guidance due to missing information about the number of insertions.	Moderate actionable tasks and goals, moderate guidance due to missing tag format orientations.
Effectiveness	Moderate effectiveness, since independence and interoperability depend on the strategy selection.	High effectiveness, since alphabet selection controls imperceptibility and embedding capacity.	High effectiveness, since replication strength modification robustness.	High effectiveness, since tags enable additional functionalities.

Following Iivari, Hansen, and Haj-Bolouri [112]

5.2 Theoretical & Practical Implications

This thesis yields implications for theory and practice from our three core contributions: a conceptual model for data sovereignty, the Innamark text watermarking artifact, and the derived DPs. Below, we provide concrete recommendations for researchers and practitioners on how to use our contributions, structured by the two research areas.

Research Area 1: Data Sovereignty Foundations

The first research area delimited data sovereignty from adjacent concepts and developed a conceptual model for it. We clustered and analyzed practical experiences regarding their requirements, challenges, and possible solutions.

First, we invite scientists to use our theoretical knowledge to build future research on sovereignty, while closely reviewing and extending it. We already see studies from various fields that utilize our results. Examples building upon our findings include research in the domains of data spaces [83], [84], [110], smart cities [15], biotechnology and healthcare [133], and network design [118], as well as an alternative data sovereignty conceptualization [1]. Further studies in other domains can assess whether the results are applicable or require adjustments.

Second, this thesis focuses on solutions for the technical challenges derived in Paper II [95]. We expect that researchers and practitioners use the results of the other two challenge categories, *organizational* and *personal & emotional*, as a starting point, not further discussed here [95]. Potential implications include law and legal specialists who need to assess the regulatory influences of various local and international acts, psychologists who validate the domain of trust, and human resources and business development departments that examine the effects of changes in security and privacy, comfort, and overall communication.

Third, we encourage practitioners to use the conceptual model (see Figure 4.3) and map it onto their systems and processes. Applying the model enables practitioners to identify gaps, e.g., if the current infrastructure ensures trust or supports the management of contractual agreements [271]. It enables expanding the field of vision to assess how data sovereignty affects all data value chain activities, rather than focusing solely on the current state of data.

Research Area 2: Digital Watermarking Design & Development

In the second research area, we iteratively designed a watermarking artifact for plain text and demonstrated and evaluated it, including derived DPs as an abstraction level.

First, the generic Kotlin library implementing Innamark, published on GitHub [75], supports the integration into various frontend and backend applications in research and practice for testing and applying our watermarking system. While we designed Innamark to protect and secure data, the implementation could also be misused, since watermarking and steganography have applications in terrorism and organized crime [49]. Examples include dual use for military operations, as already occurred with steganography techniques during World War II [122], or the transportation of malware, as shown in various JavaScript, CSS, and PHP cases [123], [241]. Notwithstanding, due to Innamark's communication and publication, it is less useful for such criminal activities, as it is not a secret technique.

Second, we encourage both researchers and practitioners to check, adjust, and change Innamark to their specific needs. One example relates to the *DP2: Alphabet Selection*, which highlights the need to define a fixed set of Unicode characters that account for the application environment, requirements, and framework conditions. In this thesis, we used a set of five similar-looking whitespaces that provide a suitable balance across the three dimensions of robustness, imperceptibility, and embedding capacity. Nevertheless, every application area is different; an environment that needs high embedding capacity should consider extending the alphabet with additional characters. Therefore, we invite practitioners to carefully review and adjust the alphabet when using our algorithm in practice. Another example is the InnamarkTag related to *DP4: Modularity*, introduced in the second DSR iteration, with configurable parameters such as compression, error-correcting codes, and hashing for verification. As with the alphabet selection, researchers and practitioners should carefully use and adjust the InnamarkTag based on their application scenarios. Thanks to our generic and extendable design, we left three bit positions unused (see Table 4.4) to shift the decision of additional extensions to the user. Practitioners may use these unassigned bit positions for a specific encryption technique for highly confidential data, such as the recently introduced text encryption system specialized for watermarking by Dafik et al. [53], to name one example.

Third, a significant challenge for practitioners is finding and using the best-fitting algorithm to strengthen data sovereignty through watermarking. Innamark is one possible method, whereas zero-width solutions with theoretically unlimited embedding capacity might make more sense in specific use cases. The trade-off triangle (see Figure 4.18) explains the trade-off, arising from the three conflicting goals of imperceptibility, robustness, and capacity, each with opposite effects. We created our resulting algorithm comparison (see Figure 4.17) to serve as a classification aid. It helps practitioners to find the best-fitting algorithm for their use case and researchers to adopt existing solutions and optimize them in their weak areas.

Finally, the derived DPs serve as an enabler for researchers and practitioners to build new IT artifacts or to further generalize our findings within a validated *design theory* at contribution *level 3* in the sense of Gregor and Hevner [85] (see Figure 3.3). One suitable approach is to develop a new watermarking algorithm that adheres to all four DPs. Another approach lies in extending and adopting existing related work based on the DPs, e.g., by replicating the watermark (*DP3: Principle of Multiple Insertions*) or adding additional functionalities (*DP4: Principle of Modularity*). Lastly, design knowledge could serve as an enabler for transferring the findings into adjacent domains such as steganography or digital rights management.

5.3 Limitations & Future Research

Despite iterative development, evaluations, and careful consideration, this thesis has limitations and offers opportunities for future research. We discuss more detailed limitations and future research opportunities in the individually published papers, listed in Appendix A. Below, we examine the most relevant aspects regarding this cumulative doctoral dissertation.

First, Innamark's replacement of Unicode whitespaces with similar-looking whitespaces limits its applicability. It only works on text containing a specific number of Unicode whitespace characters. Consequently, our proposed method does not work on text without or with only a limited number of whitespaces, depending on the length of the watermark to be embedded. For example, the Chinese language generally does not use whitespaces between words, making our solution unusable for it. The same applies to other encodings that do not support our alphabet characters, such as ASCII. Even though nearly 99% of all websites use UTF-8 [273], our solution does not cover a tiny number of cases. Future research should analyze and identify additional embedding and replacement techniques that offer high language flexibility [126]. Related work, such as Rizzo, Bertini, and Montesi [219] or the Lookalikes algorithm [258], replaces Unicode confusables, which, in turn, leads to other limitations, such as reduced imperceptibility and usage robustness. Other concepts, such as the line-ending approach discussed in our Paper VII [288], need further investigation.

Second, ensuring robustness is the most challenging task in text watermarking compared with other multimedia cover media. Our evaluation showed that Innamark achieved the best robustness compared to nine algorithms in the testbed. Nevertheless, an attacker can easily destroy or completely remove the watermark by simply replacing all whitespaces with the classical default Unicode whitespace. These attackers can go one step further by using the latest LLM models to identify and extract a watermark, raising security issues closely related to robustness. Research is ongoing to analyze these aspects further, while our initial results

indicate that the latest LLMs typically identify a watermarked text but struggle to extract it [92]. In this regard, Christ, Gunn, and Zamir argue: "We conclude that no undetectable watermarking scheme can be completely unremovable" [45, p. 1137]. This unremovability is related to the limited possibilities of plain text. Other covers, such as image, audio, or video, have a much larger number of embedding possibilities. We hypothesize that all plain text watermarking algorithms are either fragile or semi-fragile, but never robust. This assumption is in line with the *Impossibility Theorem*, which states that it is always possible to find an erasing function that removes a format-based watermark, introduced and formally proved by Sato et al. [228]. This theorem is why we classify Innamark as semi-fragile, despite its best robustness results. Future research should prove or refute the assumption and continue exploring robustness. One example is system designs, as in [121], where an attacker plays against an encoder/decoder watermarking system to optimize itself.

Third, we evaluated our results using theoretical argumentations, benchmarks, and testbed demonstrations, but we still lack long-term real-world scenarios. We mapped the conceptual model of data sovereignty to two field examples in Paper III [271], but not to a monitored usage scenario. We used Innamark in a web interface and CLI tool as well as integrated it into a data space connector already used in practice, but we still miss long-term feedback from practitioners. At the time of writing, we are collaborating with two industry partners to implement Innamark in their systems and infrastructure, aiming to roll out to thousands of published texts. Research is still in progress to closely monitor the findings and improve the current IT artifact in future studies.

Fourth, the way we applied the DSR methodology following Peffers et al. [197] limits the reuse of design knowledge across projects. We used the *problem-centered initiation* [197] as an entry point for our first research area, given current challenges with data sovereignty. This approach is criticized by vom Brocke et al. [269] because the primary goal of DSR is to use and adopt design knowledge from different projects to reuse and extend what is already there. Due to their focus on a specific problem, DSR projects tend to be monolithic and isolated rather than integrated into the existing research landscape [269]. Although our literature search did not identify any DPs, design knowledge, or design theories related to text watermarking. Even if we evaluated Innamark against related work, the limitation of isolated creation of design knowledge still applies. This isolation is closely related to aspects that enhance projectability rather than generalizability alone [25]. Future research should extend the search for possible DSR outcomes and build integrated design knowledge *across* projects rather than just *within* a project [269].

Chapter 6

Conclusion

This chapter summarizes our research contributions and answers the RQs introduced in Section 1.2, concluding our findings.

6.1 Answers to the Research Questions

After presenting all relevant results, we answer the RQs of both research areas, initially introduced in Section 1.2. One essential aspect of DSR is connecting the design knowledge gathered in the problem space with the solution space [269]. The remainder of this section answers every RQ and maps our contributions and published papers to them. Figure 6.1 concludes the thesis with a more detailed yet comprehensive overview of all relevant aspects of the problem space, such as RQs and KCs, on the left, and the resulting contributions, such as artifacts, DPs, or publications, on the right.

Research Area 1: Data Sovereignty Foundations

The first research area addresses the lack of structured knowledge on data sovereignty by building a foundation (RQ 1.1) and identifying open technical challenges (RQ 1.2).

RQ 1.1: *How can we define, differentiate, and conceptualize data sovereignty?*

We have answered the first RQ with contributions from Paper I [100] and Paper III [271], defining, differentiating, and conceptualizing data sovereignty. We have used the final literature corpus of 81 relevant articles from Paper I [100] to differentiate the concept of *data sovereignty* from the two most commonly used concepts, *digital sovereignty* and *technological sovereignty*, in the information systems and software engineering domain. The results directly align with the limited number of related works, such as Hummel et al. [108], which have reviewed data sovereignty in another domain. In the subsequent fundamentals journal Paper III [271], we have identified that defining data sovereignty in a single sentence fails to encompass all relevant aspects. As a result, we have derived a conceptual model (see Figure 4.3). Due to the intense political and practical interest, Paper III [271] uses a multivocal literature review to include gray literature, such as reports and political speeches, rather

than peer-reviewed articles alone [79]. It describes our main activities in the DSR rigor cycle, including foundational methods, experiences, and artifacts from the existing knowledge base [101].

In sum, we have successfully defined, differentiated, and conceptualized data sovereignty. Building on this, the question regarding the current hurdles and challenges that hinder data sovereignty arises:

> ***RQ 1.2:*** *Which specific technical challenges for data sovereignty exist?*

While our Paper I [100] and Paper III [271] have covered the rigorous literature perspective of data sovereignty, Paper II [95] has focused on the relevance cycle by including environmental experiences from the application domain [101]. Over 400 minutes of transcribed and coded data from semi-structured interviews with eleven experts from companies of different sizes and located in different countries worldwide yielded a set of thirteen challenges [95]. While seven challenges relate to *organizational* aspects and two to *personal & emotional* settings, this thesis has focused on the four *technical* challenges.

In sum, we have identified four technical challenges: missing *Access & Usage Control*, the complex diversity of different IT *Infrastructure & Landscapes*, the consideration of the whole *Data Processing Life Cycle*, and the need for *Identity Management* [95]. These challenges align with related work [290] and serve as a starting point for the second research area.

Research Area 2: Digital Watermarking Design & Development

The expert discussions of Paper II [95] suggested that digital watermarking techniques could address the aforementioned technical challenges. We have formulated four KCs and mapped them to the identified technical challenges and current results from the literature (see Table 4.2). They focus on a watermark *embedding* technique with the requirement of *invisibility* or imperceptibility and the two robustness claims of *modification robustness* and *usage robustness*, summarized in the first RQ 2.1 of the second research area. Our subsequent research has focused on the DSR design cycle [101], using demonstration and evaluation techniques to improve the artifact (RQ 2.2) further and to generalize the results in the form of DPs (RQ 2.3).

> ***RQ 2.1:*** *How can we design and implement a digital watermarking artifact for text-based data to ensure imperceptibility and robustness?*

A digital watermarking artifact, with a focus on imperceptibility and robustness, helps strengthen data sovereignty. During the DSR design cycle [101], we have built Innamark to

answer the RQ. Paper IV [99] has introduced the first version of the artifact, with an updated and extended version in Paper VI [94]. For transparency and reproducibility, we have published the source code of the implemented Kotlin library on GitHub [75]. We have identified an alphabet set of five similar-looking Unicode whitespaces that persist across various applications and file formats to encode the watermark (see Table 4.5). We embed the watermark by replacing every classical whitespace with the encoded, similar-looking whitespace (see Algorithm 1), and extract it by analyzing the whitespaces of a watermarked text (see Algorithm 2). The middle part of Figure 6.1 summarizes the relationships among KCs, RQs, algorithms, published papers, and source code.

In sum, we have designed the first digital watermarking artifact for plain text that does not increase the content length while staying imperceptible and robust. Due to its novelty as the first artifact to fulfill these criteria, a German and international patent application was filed [97], [98]. The design has directly addressed the derived KCs, which are analyzed, demonstrated, and improved in the upcoming RQ:

> ***RQ 2.2:*** *How can we demonstrate and evaluate the effectiveness of a digital watermarking artifact to ensure data sovereignty improvements?*

We have addressed RQ 2.2 by demonstrating and evaluating Innamark in multiple settings that correspond to the fourth *Demonstration* and fifth *Evaluation* activities of the DSR methodology from Peffers et al. [197]. In Paper V [96], we have built two Connector extensions for a data space use case to strengthen access and usage control by watermarking policies inside text data assets. While data sovereignty is one of the key principles of data spaces [183], we directly used the results to improve the artifact during the iterative design and development. In Paper VI [94], we have introduced a web interface and CLI tool to highlight the demonstration for JavaScript (frontend) and Java (backend) applications. Furthermore, we have compared our watermarking artifact Innamark in a controlled benchmark evaluation against nine algorithms in Paper VI [94], using a public dataset of 1 000 000 Wikipedia articles as cover texts. We have built on existing benchmark metrics, namely the embedding capacity ratio to evaluate the *embedding capacity*; the Jaro-Winkler Similarity, the number of characters, the file size, and the persistence toward the caret navigation attack to evaluate the *imperceptibility*; continuous text changes to evaluate the *modification robustness*; a retyping attack, a reformatting attack, and persistence in different file formats and applications to evaluate the *usage robustness*.

In sum, we have demonstrated Innamark in three different ways (see Section 4.4) and have evaluated it against nine methods from related work in a benchmark testbed (see Section 4.5). The upcoming RQ covers the generalization of the instantiated artifact:

> ***RQ 2.3:*** *How can we ensure the generalizability of a digital watermarking artifact for various use cases using design principles?*

The last RQ aims to move the more specific implemented artifact to a more abstract, generalized, and operationalized level [85]. We have extracted and derived DPs of our design effort following the framework by Purao, Kruse, and Maedche [206] and applying the taxonomy of Möller, Guggenberger, and Otto [178] (see Section 5.1). The presented DPs are formulated according to the schema of Chandra, Seidel, and Gregor [39] and structured according to Gregor, Kruse, and Seidel [87]. Thus, the design knowledge consists of *DP1: Principle of Embedding Strategy* to find a deterministic and blind mechanism; *DP2: Principle of Alphabet Selection* to find a suitable character set; *DP3: Principle of Multiple Insertions* to increase robustness via duplicated insertions; and *DP4: Principle of Modularity* to offer additional functionalities using structured tags. Additionally, we have showcased the use case transfer of watermarking capabilities for data sovereignty with CSV files in Paper VII [288] and network packets in [208].

In sum, we have derived four DPs, exemplified each of them based on our designed artifact and related work, and discussed their implications by mapping them to the initial technical data sovereignty challenges of research area one. The focus on text data has been extended to watermarking for CSV files. Taken together, these results allow us to answer the MRQ:

> ***MRQ: How to strengthen data sovereignty through digital watermarking?***

To conclude, strengthening data sovereignty is possible by extending existing concepts and systems with digital watermarking mechanisms. We have delimited and conceptualized data sovereignty to provide researchers and practitioners with tools, such as the conceptual model, to determine where and to what extent expansions are necessary. We have also designed and implemented Innamark, the first character-neutral, imperceptible, and robust digital watermarking artifact for plain text. Due to its modular structure as a published Kotlin library, developers are able to use and extend it easily. Consequently, our findings offer *one* possible digital text watermarking technique for strengthening data sovereignty, not *the* only one, as shown by the proposed algorithm comparison in the evaluation. Therefore, our resulting design knowledge, in the form of DPs, aims to help researchers and practitioners continue addressing the discussed limitations and to enable future research streams to further strengthen data sovereignty.

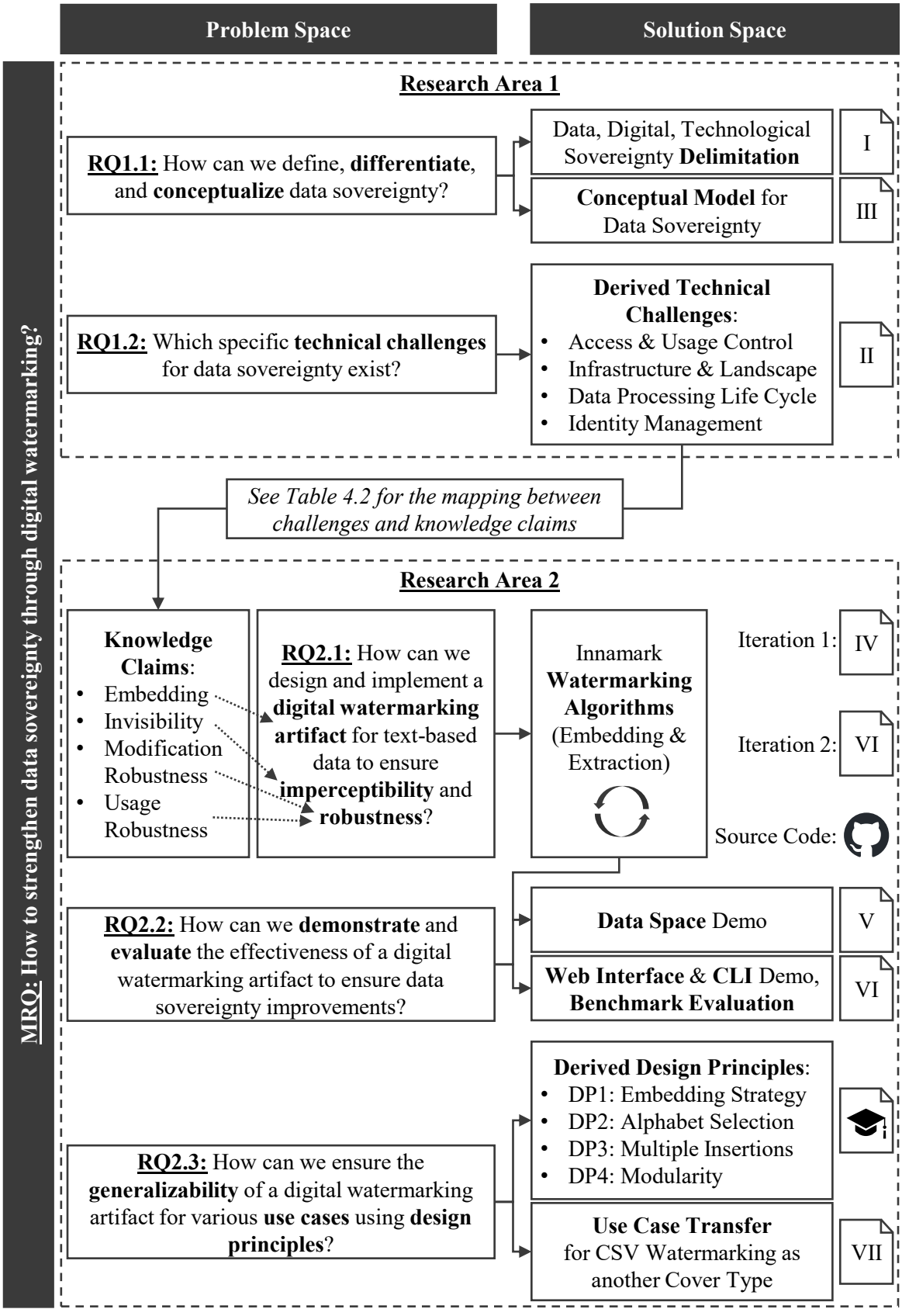


Fig. 6.1 Summarizing Problem and Solution Overview

6.2 Summary

This thesis is motivated by the ongoing shift from the analog to the digital world, which intensifies efforts in digitization and digitalization. Since new technologies and applications, such as LLMs, and process optimizations shape our everyday lives, individuals, organizations, and governments are increasingly interested in protecting our digital world and the data behind it. The need for and use of concrete solutions to remain *data sovereign* are of foremost importance.

Analyzing state-of-the-art through a structured literature review has led us to delimit the field and to construct a conceptual model of data sovereignty. We have communicated with experts from different industries through semi-structured interviews and analyzed the findings using coding techniques and Grounded Theory, identifying four significant technical challenges. The results indicate that digital watermarking techniques are the most promising solution for strengthening data sovereignty, but combining the two research domains remains under-researched. We have derived four KCs for digital watermarking and mapped our derived technical challenges to them.

We have iteratively designed Innamark, a watermark embedding and extraction algorithm for plain text. We confirm that replacing whitespaces with similar-looking Unicode whitespaces directly addresses the derived KCs. We have implemented and demonstrated the system in a data space scenario, a web interface, and a CLI tool. The benchmark evaluation on a dataset of 1 000 000 Wikipedia articles has identified strengths and weaknesses through a direct comparison of ten algorithms (Innamark and nine from related work). Our gathered design knowledge, encapsulated in four DPs, generalizes the findings and paves the way for future watermarking systems.

This cumulative doctoral dissertation demonstrates how digital watermarking strengthens data sovereignty. Researchers and practitioners in information systems and software engineering can use, improve, and benefit from these contributions. Nevertheless, our work has limitations: the conceptual model lacks a monitored usage scenario, whereas Innamark and the set of DPs focus on Unicode encoding, robustness complexities, and project-focused design knowledge. They all lack a long-term evaluation in specific application scenarios with practitioners. Future work is currently in progress to test, adopt, and further improve the results, and to deploy them in large-scale real-world use cases with different industry partners.

Declaration of Resources, Tools, and AI Technologies

During the preparation of this work, I used the following tools, external components, libraries, and resources. The final document was created with XeLaTeX, based on an adapted version of the PhD thesis template for the Cambridge University Engineering Department, developed by Krishna Kumar (license: MIT). The references were created with BibLaTeX and Biber from a Citavi export. The illustrative figures were created in Microsoft PowerPoint; the screenshots were captured directly; the UML diagrams were created using plantUML; and the plots were created in Python using NumPy, pandas, Matplotlib, and SciencePlots. Some figures include external icons, such as an education hat (source: iconfinder.com, license: free for commercial use) and the GitHub logo (source and license: github.com/logos). For language correction, accuracy, grammar, punctuation, spelling checks, and content review, AI-assisted tools such as Grammarly, DeepL, and GPT models were used. I carefully reviewed, edited, and revised all suggested AI corrections and outputs. AI tools were not used to compose full sections, paragraphs, or blocks of content. I take full responsibility for the accuracy and originality of the final content of this thesis.

References

[1] Antragama Ewa Abbas, Thomas van Velzen, Hosea Ofe, Geerten van de Kaa, Anneke Zuiderwijk, and Mark de Reuver, "Beyond control over data: Conceptualizing data sovereignty from a social contract perspective," *Electronic Markets*, vol. 34, art. no. 20, 21 pages, 2024. DOI: 10.1007/s12525-024-00695-2.

[2] Abid A. Adonis, "Critical Engagement on Digital Sovereignty in International Relations: Actor Transformation and Global Hierarchy," *Global: Jurnal Politik Internasional*, vol. 21, no. 2, pp. 262–282, 2019. DOI: 10.7454/global.v21i2.412.

[3] Milad Taleby Ahvanooey, Qianmu Li, Jun Hou, Hassan Dana Mazraeh, and Jing Zhang, "AITSteg: An Innovative Text Steganography Technique for Hidden Transmission of Text Message via Social Media," *IEEE Access*, vol. 6, pp. 65981–65995, 2018. DOI: 10.1109/ACCESS.2018.2866063.

[4] Milad Taleby Ahvanooey, Qianmu Li, Jun Hou, Ahmed Raza Rajput, and Yini Chen, "Modern Text Hiding, Text Steganalysis, and Applications: A Comparative Analysis," *Entropy*, vol. 21, no. 4, art. no. 355, 31 pages, 2019. DOI: 10.3390/e21040355.

[5] Milad Taleby Ahvanooey, Qianmu Li, Hiuk Jae Shim, and Yanyan Huang, "A Comparative Analysis of Information Hiding Techniques for Copyright Protection of Text Documents," *Security and Communication Networks*, vol. 2018, art. no. 5325040, 22 pages, 2018. DOI: 10.1155/2018/5325040.

[6] Milad Taleby Ahvanooey, Qianmu Li, Xuefang Zhu, Mamoun Alazab, and Jing Zhang, "ANiTW: A Novel Intelligent Text Watermarking technique for forensic identification of spurious information on social media," *Computers & Security*, vol. 90, art. no. 101702, 14 pages, 2020. DOI: 10.1016/j.cose.2019.101702.

[7] Milad Taleby Ahvanooey, Hassan Dana Mazraeh, and Seyed Hashem Tabasi, "An innovative technique for web text watermarking (AITW)," *Information Security Journal: A Global Perspective*, vol. 25, no. 4-6, pp. 191–196, 2016. DOI: 10 . 1080 / 19393555.2016.1202356.

[8] Milad Taleby Ahvanooey, Mark Xuefang Zhu, Wojciech Mazurczyk, and Malika Bendechache, "Information Hiding in Digital Textual Contents: Techniques and Current Challenges," *Computer*, vol. 55, no. 6, pp. 56–65, 2022. DOI: 10.1109/MC.2021.3113922.

[9] Milad Taleby Ahvanooey, Mark Xuefang Zhu, Wojciech Mazurczyk, Qianmu Li, Max Kilger, Kim-Kwang Raymond Choo, and Mauro Conti, "CovertSYS: A systematic covert communication approach for providing secure end-to-end conversation via social networks," *Journal of Information Security and Applications*, vol. 71, art. no. 103368, 16 pages, 2022. DOI: 10.1016/j.jisa.2022.103368.

[10] Norah Alanazi, Esam Khan, and Adnan Gutub, "Involving Spaces of Unicode Standard Within Irreversible Arabic Text Steganography for Practical Implementations," *Arabian Journal for Science and Engineering*, vol. 46, no. 9, pp. 8869–8885, 2021. DOI: 10.1007/s13369-021-05605-8.

[11] Mohammed Hazim Alkawaz, Ghazali Sulong, Tanzila Saba, Abdulaziz S. Almazyad, and Amjad Rehman, "Concise analysis of current text automation and watermarking approaches," *Security and Communication Networks*, vol. 9, no. 18, pp. 6365–6378, 2016. DOI: 10.1002/sec.1738.

[12] Marcel Altendeitering, Julia Pampus, Felix Larrinaga, Jon Legaristi, and Falk Howar, "Data Sovereignty for AI Pipelines: Lessons Learned from an Industrial Project at Mondragon Corporation," in *2022 IEEE/ACM 1st International Conference on AI Engineering – Software Engineering for AI (CAIN)*, 2022, pp. 193–204. DOI: 10.1145/3522664.3528593.

[13] Harald Tveit Alvestrand, "IETF Policy on Character Sets and Languages," RFC Editor, RFC 2277, 1998, 9 pages. Available at: https://www.rfc-editor.org/rfc/rfc2277.

[14] American Standards Association, *American Standard Code for Information Interchange*, ASA X3.4-1963, New York, NY, USA, Jun. 17, 1963. Available at: https://www.sensitiveresearch.com/Archive/CharCodeHist/Files/CODES%20standards%20documents%20ASCII%20Sean%20Leonard%20Oct%202015/ASCII%2063,%20X3.4-1963.pdf.

[15] Bokolo Anthony and Sizarta Sarshar, "Enhancing data sovereignty to improve intelligent mobility services in smart cities," *Urban Governance*, vol. 5, no. 1, pp. 20–31, 2025. DOI: 10.1016/j.ugj.2025.02.002.

[16] Arno Appenzeller, Ewald Rode, Erik Krempel, and Jürgen Beyerer, "Enabling data sovereignty for patients through digital consent enforcement," in *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2020, pp. 249–252. DOI: 10.1145/3389189.3393745.

[17] Diego Adetayó C. Assumpção and Alberto Blumenschein-Cruz, "Data Sovereignty and International Sovereignty Principles - The Data Sovereignty Observatory," in *Proceedings of the 17th International Conference on Theory and Practice of Electronic Governance*, 2024, pp. 388–391. DOI: 10.1145/3680127.3680221.

[18] Atilla Aydin and Turksel Kaya Bensghir, "Digital Data Sovereignty: Towards a Conceptual Framework," in *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, 2019, pp. 1–6. DOI: 10.1109/UBMYK48245.2019.8965469.

[19] Salwa Shakir Baawi, Mohd Rosmadi Mokhtar, and Rossilawati Sulaiman, "Enhancement of Text Steganography Technique Using Lempel-Ziv-Welch Algorithm and Two-Letter Word Technique," in *Recent Trends in Data Science and Soft Computing*, ser. Advances in Intelligent Systems and Computing, vol. 843, Faisal Saeed, Nadhmi Gazem, Fathey Mohammed, and Abdelsalam Busalim, Eds., Cham, Switzerland: Springer International Publishing, 2019, pp. 525–537. DOI: 10.1007/978-3-319-99007-1_49.

[20] Marie Baezner and Patrice Robin, "Cyber Sovereignty and Data Sovereignty," 2018. DOI: 10.3929/ethz-b-000314613.

[21] Christian Banse, "Data Sovereignty in the Cloud - Wishful Thinking or Reality?" In *Proceedings of the 2021 on Cloud Computing Security Workshop*, Keynote Talk III, 2021. DOI: 10.1145/3474123.3486792.

[22] Vijon Baraku, Edon Ramadani, Iraklis Paraskakis, Simeon Veloudis, and Poonam Yadav, "Defining personal data Sovereignty: An ontologically-based framework facilitating subject privacy control," *Data and Information Management*, vol. 10, no. 1, art. no. 100108, 12 pages, 2026. DOI: 10.1016/j.dim.2025.100108.

[23] Hafsat Muhammad Bashir, Qianmu Li, and Jun Hou, "A High Capacity Text Steganography Utilizing Unicode Zero-Width Characters," in *2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, 2020, pp. 668–675. DOI: 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics50389.2020.00116.

[24] Richard Baskerville, Abayomi Baiyere, Shirley Gregor, Alan Hevner, and Matti Rossi, "Design Science Research Contributions: Finding a Balance between Artifact and Theory," *Journal of the Association for Information Systems*, vol. 19, no. 5, pp. 358–376, 2018. Available at: https://aisel.aisnet.org/jais/vol19/iss5/3.

[25] Richard Baskerville and Jan Pries-Heje, "Projectability in Design Science Research," *Journal of Information Technology Theory and Application (JITTA)*, vol. 20, no. 1, art. no. 3, pp. 53–76, 2019. Available at: https://aisel.aisnet.org/jitta/vol20/iss1/3.

[26] Walter Bender, Daniel Gruhl, Norishige Morimoto, and Anthony Lu, "Techniques for data hiding," *IBM Systems Journal*, vol. 35, no. 3.4, pp. 313–336, 1996. DOI: 10.1147/sj.353.0313.

[27] Karen M. Benzies, Shahirose Premji, K. Alix Hayden, and Karen Serrett, "State-of-the-evidence reviews: advantages and challenges of including grey literature," *Worldviews on Evidence-Based Nursing*, vol. 3, no. 2, pp. 55–61, 2006. DOI: 10.1111/j.1741-6787.2006.00051.x.

[28] Hal Berghel and Lawrence O'Gorman, "Protecting ownership rights through digital watermarking," *Computer*, vol. 29, no. 7, pp. 101–103, 1996. DOI: 10.1109/2.511977.

[29] Flavio Bertini, Alessandro Benetton, and Danilo Montesi, "Distributed Ledger and Text Watermarking for Fine-Grain Provenance Checking of Textual Content," in *Companion Proceedings of the ACM on Web Conference 2025*, 2025, pp. 2626–2633. DOI: 10.1145/3701716.3717536.

[30] Flavio Bertini, Stefano Giovanni Rizzo, and Danilo Montesi, "Can Information Hiding in Social Media Posts Represent a Threat?" *Computer*, vol. 52, no. 10, pp. 52–60, 2019. DOI: 10.1109/MC.2019.2917199.

[31] Gioele Bigini, Mirko Zichichi, Emanuele Lattanzi, Stefano Ferretti, and Gabriele D'Angelo, "Decentralized Health Data Distribution: A DLT-based Architecture for Data Protection," in *2022 IEEE International Conference on Blockchain (Blockchain)*, 2022, pp. 97–104. DOI: 10.1109/Blockchain55522.2022.00023.

[32] Filippo Gualtiero Blancato, "The cloud sovereignty nexus: How the European Union seeks to reverse strategic dependencies in its digital ecosystem," *Policy & Internet*, vol. 16, no. 1, pp. 12–32, 2023. DOI: 10.1002/poi3.358.

[33] Jean Bodin, *Les six livres de la république*. 1577.

[34] Nima Boscarino, Reed A. Cartwright, Keolu Fox, and Krystal S. Tsosie, "Federated learning and Indigenous genomic data sovereignty," *Nature Machine Intelligence*, vol. 4, no. 11, pp. 909–911, 2022. DOI: 10.1038/s42256-022-00551-y.

[35] Jack T. Brassil, Steven Low, Nicholas F. Maxemchuk, and Lawrence O'Gorman, "Electronic marking and identification techniques to discourage document copying," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 8, pp. 1495–1504, 1995. DOI: 10.1109/49.464718.

[36] Jack T. Brassil, Steven Low, Nicholas F. Maxemchuk, and Lawrence O'Gorman, "Hiding Information in Document Images," in *Proc. Conf. Information Sciences and Systems (CISS-95)*, 1995, pp. 482–489.

[37] Paul Butler. "Smuggling arbitrary data through an emoji." 2025. Available at: https://paulbutler.org/2025/smuggling-arbitrary-data-through-an-emoji/ (accessed on Feb. 10, 2026).

[38] Amaël Cattaruzza, Didier Danet, Stéphane Taillat, and Arthur Laudrain, "Sovereignty in cyberspace: Balkanization or democratization," in *2016 International Conference on Cyber Conflict (CyCon U.S.)*, 2016, pp. 1–9. DOI: 10.1109/CYCONUS.2016.7836628.

[39] Leona Chandra, Stefan Seidel, and Shirley Gregor, "Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions," in *2015 48th Hawaii International Conference on System Sciences*, 2015, pp. 4039–4048. DOI: 10.1109/HICSS.2015.485.

[40] Leona Chandra Kruse, Stefan Seidel, and Sandeep Purao, "Making Use of Design Principles," in *Tackling Society's Grand Challenges with Design Science*, ser. Lecture Notes in Computer Science, vol. 9661, Jeffrey Parsons, Tuure Tuunanen, John Venable, Brian Donnellan, Markus Helfert, and Jim Kenneally, Eds., Cham, Switzerland: Springer International Publishing, 2016, pp. 37–51. DOI: 10.1007/978-3-319-39294-3_3.

[41] Kathy Charmaz, *Constructing Grounded Theory, A Practical Guide Through Qualitative Analysis*. London: Sage Publications Ltd, 2006, 223 pages, ISBN: 9780761973539.

[42] Yaoliang Chen, Shi Chen, Jiao Liang, Lance Warren Feagan, Weili Han, Sheng Huang, and X. Sean Wang, "Decentralized data access control over consortium blockchains," *Information Systems*, vol. 94, art. no. 101590, 15 pages, 2020. DOI: 10.1016/j.is.2020.101590.

[43] Yung-Chen Chou, Chun-Yi Huang, and Hsin-Chi Liao, "A Reversible Data Hiding Scheme Using Cartesian Product for HTML File," in *2012 Sixth International Conference on Genetic and Evolutionary Computing*, 2012, pp. 153–156. DOI: 10.1109/ICGEC.2012.30.

[44] Miranda Christ and Sam Gunn, "Pseudorandom Error-Correcting Codes," in *Advances in Cryptology – CRYPTO 2024*, ser. Lecture Notes in Computer Science, vol. 14925, Leonid Reyzin and Douglas Stebila, Eds., Cham, Switzerland: Springer Nature Switzerland, 2024, pp. 325–347. DOI: 10.1007/978-3-031-68391-6_10.

[45] Miranda Christ, Sam Gunn, and Or Zamir, "Undetectable Watermarks for Language Models," in *Proceedings of Thirty Seventh Conference on Learning Theory*, vol. 247, 2024, pp. 1125–1139. Available at: https://proceedings.mlr.press/v247/christ24a.html.

[46] Elizabeth Clark, Tal August, Sofia Serrano, Nikita Haduong, Suchin Gururangan, and Noah A. Smith, "All That's 'Human' Is Not Gold: Evaluating Human Evaluation of Generated Text," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2021, pp. 7282–7296. DOI: 10.18653/v1/2021.acl-long.565.

[47] Victor Costan and Srinivas Devadas, "Intel SGX Explained," *Cryptology ePrint Archive*, art. no. 2016/086, 118 pages, 2016. Available at: https://eprint.iacr.org/2016/086.

[48] Stephane Couture and Sophie Toupin, "What does the notion of "sovereignty" mean when referring to the digital?" *New Media & Society*, vol. 21, no. 10, pp. 2305–2322, 2019. DOI: 10.1177/1461444819865984.

[49] Ingemar Cox, Matthew Miller, and Jeffrey Bloom, *Digital Watermarking and Steganography*, 2. Aufl. s.l.: Elsevier professional, 2007, 624 pages, ISBN: 978-0-12-372585-1.

[50] Ingemar J. Cox, *Digital watermarking*. San Diego, CA, USA: Academic Press, 2001, 542 pages, ISBN: 1-55860-714-5.

[51] Michael Crosby, Nachiappan, Pradan Pattanayak, Sanjeev Verma, and Vignesh Kalyanaraman, "Blockchain Technology: Beyond Bitcoin," *Applied Innovation Review*, no. 2, pp. 6–19, 2016. Available at: https://scet.berkeley.edu/reports/blockchain/.

[52] Peter Dabrock, "From Data Protection to Data Sovereignty. A multidimensional governance approach for shaping informational freedom in the 'onlife'-era," *Cursor_ Zeitschrift für explorative Theologie*, 2019. DOI: 10.21428/fb61f6aa.f0bf0cc2.

[53] Dafik, Swaminathan Venkatraman, G. Sathyanarayanan, Rifki Ilham Baihaki, Indah Lutfiyatul Mursyidah, and Ika Hesti Agustin, "Enhancing Text Encryption and Secret Document Watermarking through Hyperladder Graph-Based Keystream Construction on Asymmetric Cryptography Technology," *Statistics, Optimization & Information Computing*, vol. 14, no. 1, pp. 247–263, 2025. DOI: 10.19139/soic-2310-5070-2310.

[54] Sumanth Dathathri, Abigail See, Sumedh Ghaisas, Po-Sen Huang, Rob McAdam, Johannes Welbl, Vandana Bachani, Alex Kaskasoli, Robert Stanforth, Tatiana Matejovicova, Jamie Hayes, Nidhi Vyas, Majd Al Merey, Jonah Brown-Cohen, Rudy Bunel, Borja Balle, Taylan Cemgil, Zahra Ahmed, Kitty Stacpoole, Ilia Shumailov, Ciprian Baetu, Sven Gowal, Demis Hassabis, and Pushmeet Kohli, "Scalable watermarking for identifying large language model outputs," *Nature*, vol. 634, no. 8035, pp. 818–823, 2024. DOI: 10.1038/s41586-024-08025-4.

[55] Patrício de Alencar Silva, Reza Fadaie, and Marten van Sinderen, "Towards a Digital Twin for Simulation of Organizational and Semantic Interoperability in IDS Ecosystems," in *Interoperability for Enterprise Systems and Applications Workshops, I-ESA Workshops 2022*, vol. 3214, 2022. Available at: https://ceur-ws.org/Vol-3214/WS6Paper4.pdf.

[56] Stephen E. Deering and Robert M. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," RFC Editor, RFC 8200, 2017, 42 pages. Available at: https://www.rfc-editor.org/rfc/rfc8200.

[57] Matthew DeLorenzo, Phat Tieu, Chen Chen, Vasudev Gohil, and Jeyavijayan Rajendran, "Watermarking LLMs — Challenges and Opportunities in Electronic Design Automation," in *2025 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, 2025, pp. 1–4. DOI: 10.1109/COINS65080.2025.11125763.

[58] Alan Dennis and Joseph Valacich, "A Replication Manifesto," *AIS Transactions on Replication Research*, vol. 1, art. no. 1, pp. 1–4, 2014. DOI: 10.17705/1atrr.00001.

[59] Longjam Velentina Devi and Navanath Saharia, "PepSteg: A Text Steganographic Approach based on Prime Embedding Position," *Arabian Journal for Science and Engineering*, 2025. DOI: 10.1007/s13369-025-10070-8.

[60] Quentin Docter and Cory Fuchs, "Compliance and security in the cloud," in *CompTIA Cloud Essentials+ Study Guide*, Quentin Docter and Cory Fuchs, Eds., 2nd Edition, Wiley, 2020, pp. 253–302. DOI: 10.1002/9781119642138.ch7.

[61] Cathal Doyle, Markus Luczak-Roesch, and Abhinav Mittal, "We Need the Open Artefact: Design Science as a Pathway to Open Science in Information Systems Research," in *Extending the Boundaries of Design Science Theory and Practice*, ser. Lecture Notes in Computer Science, vol. 11491, Bengisu Tulu, Soussan Djamasbi, and Gondy Leroy, Eds., Cham, Switzerland: Springer International Publishing, 2019, pp. 46–60. DOI: 10.1007/978-3-030-19504-5_4.

[62] Jakob Edler, Knut Blind, Henning Kroll, and Torben Schubert, "Technology Sovereignty as an Emerging Frame for Innovation Policy – Defining Rationales, Ends and Means," Karlsruhe, Germany, 70, 2021, 36 pages. Available at: https://publica.fraunhofer.de/dokumente/N-638343.html.

[63] Roman Elizarov, Sebastian Aigner, Svetlana Isakova, and Dmitry Jemerov, *Kotlin in action*, 2nd ed. Shelter Island, NY, USA: Manning, 2024, 529 pages, ISBN: 978-1617299605.

[64] Emelie Engström, Margaret-Anne Storey, Per Runeson, Martin Höst, and Maria Teresa Baldassarre, "How software engineering research aligns with design science: a review," *Empirical Software Engineering*, vol. 25, no. 4, pp. 2630–2660, 2020. DOI: 10.1007/s10664-020-09818-7.

[65] Ericsson. "Annual mobile data traffic worldwide from 2012 to 2031 (in exabytes per month) [Graph]," Statista. 2025. Available at: https://www.statista.com/statistics/630107/annual-mobile-data-usage-worldwide/ (accessed on Feb. 10, 2026).

[66] Jens Ernstberger, Jan Lauinger, Fatima Elsheimy, Liyi Zhou, Sebastian Steinhorst, Ran Canetti, Andrew Miller, Arthur Gervais, and Dawn Song, "SoK: Data Sovereignty," in *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*, 2023, pp. 122–143. DOI: 10.1109/EuroSP57164.2023.00017.

[67] Christian Esposito, Aniello Castiglione, and Kim-Kwang Raymond Choo, "Encryption-Based Solution for Data Sovereignty in Federated Clouds," *IEEE Cloud Computing*, vol. 3, no. 1, pp. 12–17, 2016. DOI: 10.1109/MCC.2016.18.

[68] Christian Esposito, Aniello Castiglione, Flavio Frattini, Marcello Cinque, Yanjiang Yang, and Kim-Kwang Raymond Choo, "On Data Sovereignty in Cloud-Based Computation Offloading for Smart Cities Applications," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4521–4535, 2019. DOI: 10.1109/JIOT.2018.2886410.

[69] European Commission, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)," Brussels, Belgium, 2024/1689, 2024. Available at: https://eur-lex.europa.eu/eli/reg/2024/1689.

[70] European Commission, "Regulation (EU) 2024/1781 of the European Parliament and of the Council of 13 June 2024 establishing a framework for the setting of ecodesign requirements for sustainable products, amending Directive (EU) 2020/1828 and Regulation (EU) 2023/1542 and repealing Directive 2009/125/EC," Brussels, Belgium, 2024/1781, 2024. Available at: https://eur-lex.europa.eu/eli/reg/2024/1781/oj.

[71] Oleg Evsutin, Anna Melman, and Ahmed A. Abd El-Latif, "Overview of Information Hiding Algorithms for Ensuring Security in IoT Based Cyber-Physical Systems," in *Security and Privacy Preserving for IoT and 5G Networks*, ser. Studies in Big Data, vol. 95, Ahmed A. Abd El-Latif, Bassem Abd-El-Atty, Salvador E. Venegas-Andraca, Wojciech Mazurczyk, and Brij B. Gupta, Eds., Cham, Switzerland: Springer International Publishing, 2022, pp. 81–115. DOI: 10.1007/978-3-030-85428-7_5.

[72] Oleg Evsutin, Anna Melman, and Roman Meshcheryakov, "Digital Steganography and Watermarking for Digital Images: A Review of Current Research Directions," *IEEE Access*, vol. 8, pp. 166589–166611, 2020. DOI: 10.1109/ACCESS.2020.3022779.

[73] Danniar Reza Firdausy, Patrício de Alencar Silva, Marten van Sinderen, and Maria-Eugenia Iacob, "Towards a Reference Enterprise Architecture to enforce Digital Sovereignty in International Data Spaces," in *2022 IEEE 24th Conference on Business Informatics (CBI)*, 2022, pp. 117–125. DOI: 10.1109/CBI54897.2022.00020.

[74] Luciano Floridi, "The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU," *Philosophy & Technology*, vol. 33, no. 3, pp. 369–378, 2020. DOI: 10.1007/s13347-020-00423-6.

[75] Fraunhofer ISST, *Innamark*, 2026. Available at: https://github.com/FraunhoferISST/Innamark (accessed on Feb. 10, 2026).

[76] André Freitas and Edward Curry, "Big Data Curation," in *New horizons for a data-driven economy, A roadmap for usage and exploitation of big data in Europe*, José María Cavanillas, Edward Curry, and Wolfgang Wahlster, Eds., Cham, Switzerland: Springer International Publishing AG, 2016, pp. 87–118. DOI: 10.1007/978-3-319-21569-3_6.

[77] Jean-Loup Gailly and Laurence Peter Deutsch, "ZLIB Compressed Data Format Specification version 3.3," RFC Editor, RFC 1950, 1996, 11 pages. Available at: https://www.rfc-editor.org/rfc/rfc1950.

[78] Erich Gamma, Richard Helm, Ralph Johnson, and John M. Vlissides, *Design patterns, Elements of reusable object-oriented software*. Reading, MA, USA: Addison-Wesley, 1995, 395 pages, ISBN: 978-0201633612.

[79] Vahid Garousi, Michael Felderer, and Mika V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," *Information and Software Technology*, vol. 106, pp. 101–121, 2019. DOI: 10.1016/j.infsof.2018.09.006.

[80] Carrie Gates and Jacob Slonim, "Owner-controlled information," in *Proceedings of the 2003 workshop on New security paradigms - NSPW '03*, 2003, pp. 103–111. DOI: 10.1145/986655.986670.

[81] German Ethics Council, "Big Data and Health – Data Sovereignty as the Shaping of Informational Freedom, Executive Summary & Recommendations," Berlin, Germany, 2017, 52 pages. Available at: https://www.ethikrat.org/en/publications/opinions/big-data-and-health/.

[82] Tim Giese and Reiner Anderl, "Maintaining Control over Distributed Data Through a Data Sovereignty Model," in *2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA)*, 2022, pp. 1–7. DOI: 10.1109/ICITDA55840.2022.9971218.

[83] Anna Gieß, Marius Hupperz, Thorsten Schoormann, and Frederik Möller, "What Does it Take to Connect? Unveiling Characteristics of Data Space Connectors," in *Proceedings of the 57th Hawaii International Conference on System Sciences*, 2024. DOI: 10.24251/HICSS.2024.511.

[84] Anna Gieß, Thorsten Schoormann, Frederik Möller, and Inan Gür, "Discovering data spaces: A classification of design options," *Computers in Industry*, vol. 164, art. no. 104212, 15 pages, 2025. DOI: 10.1016/j.compind.2024.104212.

[85] Shirley Gregor and Alan R. Hevner, "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly*, vol. 37, no. 2, pp. 337–355, 2013. DOI: 10.25300/MISQ/2013/37.2.01.

[86] Shirley Gregor and David Jones, "The Anatomy of a Design Theory," *Journal of the Association for Information Systems*, vol. 8, no. 5, pp. 312–335, 2007. DOI: 10.17705/1jais.00129.

[87] Shirley Gregor, Leona Kruse, and Stefan Seidel, "Research Perspectives: The Anatomy of a Design Principle," *Journal of the Association for Information Systems*, vol. 21, pp. 1622–1652, 2020. DOI: 10.17705/1jais.00649.

[88] Adnan Gutub, "Emerging Arabic Text Watermarking Utilizing Combinations of Different Diacritics," *Arabian Journal for Science and Engineering*, 2024. DOI: 10.1007/s13369-023-08629-4.

[89] Dara Hallinan, *Data Protection and Privacy, Enforcing Rights in a Changing World*, in collab. with Ronald Leenes and Paul de Hert. London, UK: Bloomsbury Publishing Plc, 2022, vol. 14, 311 pages, ISBN: 978-1-50995-451-3.

[90] Wilhelm Hasselbring, "Benchmarking as Empirical Standard in Software Engineering Research," in *Evaluation and Assessment in Software Engineering*, 2021, pp. 365–372. DOI: 10.1145/3463274.3463361.

[91] Jiahuan He, Qichao Ying, Zhenxing Qian, Guorui Feng, and Xinpeng Zhang, "Semi-structured data protection scheme based on robust watermarking," *EURASIP Journal on Image and Video Processing*, vol. 2020, art. no. 12, 10 pages, 2020. DOI: 10.1186/s13640-020-00500-y.

[92] Malte Hellmeier, "Security and Detectability Analysis of Unicode Text Watermarking Methods against Large Language Models," in *Proceedings of the 12th International Conference on Information Systems Security and Privacy*, 2026, pp. 197–204. DOI: 10.5220/0014268700004061.

[93] Malte Hellmeier, Hendrik Norkowski, Ernst-Christoph Schrewe, Haydar Qarawlus, and Falk Howar, "Innamark: A Whitespace Replacement Information-Hiding Method," 2025, Preprint. DOI: 10.48550/arXiv.2502.12710.

[94] Malte Hellmeier, Hendrik Norkowski, Ernst-Christoph Schrewe, Haydar Qarawlus, and Falk Howar, "Innamark: A Whitespace Replacement Information-Hiding Method," *IEEE Access*, vol. 13, pp. 123120–123135, 2025. DOI: 10.1109/ACCESS.2025.3583591.

[95] Malte Hellmeier, Julia Pampus, Haydar Qarawlus, and Falk Howar, "Implementing Data Sovereignty: Requirements & Challenges from Practice," in *Proceedings of the 18th International Conference on Availability, Reliability and Security*, 2023, pp. 1–9. DOI: 10.1145/3600160.3604995.

[96] Malte Hellmeier and Haydar Qarawlus, "Strengthening Data Sovereignty Through Digital Watermarking in Data Spaces," in *Proceedings of the 58th Hawaii International Conference on System Sciences*, 2025, pp. 4346–4355. DOI: 10.24251/hicss.2025.520.

[97] Malte Hellmeier, Haydar Qarawlus, and Hendrik Norkowski, "Verfahren und System zur zeichenanzahlneutralen Einbettung einer digitalen Signatur in ein digitales Dokument [Method and System for the Character-Number-Neutral Embedding of a Digital Signature in a Digital Document]," German patent application no. DE102023125012.4A, Sep. 15, 2023. Available at: https://patents.google.com/patent/DE102023125012A1/ (accessed on Feb. 10, 2026).

[98] Malte Hellmeier, Haydar Qarawlus, and Hendrik Norkowski, "Method and System for the Character-Number-Neutral Embedding of a Digital Signature in a Digital Document," International patent application no. PCT/EP2024/075670, Sep. 13, 2024. Available at: https://patents.google.com/patent/WO2025056772A1/ (accessed on Feb. 10, 2026).

[99] Malte Hellmeier, Haydar Qarawlus, Hendrik Norkowski, and Falk Howar, "A Hidden Digital Text Watermarking Method Using Unicode Whitespace Replacement," in *Proceedings of the 58th Hawaii International Conference on System Sciences*, 2025, pp. 7411–7420. DOI: 10.24251/hicss.2025.886.

[100] Malte Hellmeier and Franziska von Scherenberg, "A Delimitation of Data Sovereignty from Digital and Technological Sovereignty," *ECIS 2023 Research Papers*, art. no. 306, 2023. Available at: https://aisel.aisnet.org/ecis2023_rp/306.

[101] Alan R. Hevner, "A Three Cycle View of Design Science Research," *Scandinavian Journal of Information Systems*, vol. 19, no. 2, 2007. Available at: https://aisel.aisnet.org/sjis/vol19/iss2/4/.

[102] Alan R. Hevner and Samir Chatterjee, *Design Research in Information Systems*. Boston, MA, USA: Springer US, 2010, vol. 22. DOI: 10.1007/978-1-4419-5653-8.

[103] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram, "Design Science in Information Systems Research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004. DOI: 10.2307/25148625.

[104] Alan R. Hevner, Jeffrey Parsons, Alfred Benedikt Brendel, Roman Lukyanenko, Verena Tiefenbeck, Monica Chiarini Tremblay, and Jan vom Brocke, "Transparency in design science research," *Decision Support Systems*, vol. 182, art. no. 114236, 11 pages, 2024. DOI: 10.1016/j.dss.2024.114236.

[105] Michael Hofmeier and Wolfgang Hommel, "Application for Electronic Signatures Using Blockchain Technology to Support Trust, Sovereignty and Privacy," in *2024 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics*, 2024, pp. 587–594. DOI: 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics62450.2024.00108.

[106] House of Representatives, *H.R.4943 - CLOUD Act*, 115th Congress, 2d Session, 2018. Available at: https://www.congress.gov/bill/115th-congress/house-bill/4943.

[107] Patrik Hummel, Matthias Braun, Steffen Augsberg, and Peter Dabrock, "Sovereignty and data sharing," *ITU Journal: ICT Discoveries*, vol. 1, no. 2, 2018. Available at: https://www.itu.int/en/journal/002/Pages/11.aspx.

[108] Patrik Hummel, Matthias Braun, Max Tretter, and Peter Dabrock, "Data sovereignty: A review," *Big Data & Society*, vol. 8, no. 1, pp. 1–17, 2021. DOI: 10.1177/2053951720982012.

[109] Jonathon Hutchinson, Milica Stilinovic, and Joanne E. Gray, "Data sovereignty: The next frontier for internet policy?" *Policy & Internet*, vol. 16, no. 1, pp. 6–11, 2024. DOI: 10.1002/poi3.386.

[110] Andreas Hutterer and Barbara Krumay, "Exploring Digital Transformation Opportunities of Data Spaces," in *Proceedings of the 58th Hawaii International Conference on System Sciences*, 2025, pp. 4326–4335. DOI: 10.24251/HICSS.2025.518.

[111] Juhani Iivari, "Distinguishing and contrasting two strategies for design science research," *European Journal of Information Systems*, vol. 24, no. 1, pp. 107–115, 2015. DOI: 10.1057/ejis.2013.35.

[112] Juhani Iivari, Magnus Rotvit Perlt Hansen, and Amir Haj-Bolouri, "A proposal for minimum reusability evaluation of design principles," *European Journal of Information Systems*, vol. 30, no. 3, pp. 286–303, 2021. DOI: 10.1080/0960085X.2020.1793697.

[113] ISO/IEC, *Information technology — 8-bit single-byte coded graphic character sets, Part 1: Latin alphabet No. 1*, ISO/IEC 8859-1:1998(E), ICS 35.040.10, International Organization for Standardization, Geneva, Switzerland, Apr. 15, 1998, 10 pages. Available at: https://www.iso.org/standard/28245.html.

[114] ISO/IEC, *Information technology — 8-bit single-byte coded graphic character sets, Part 16: Latin alphabet No. 10*, ISO/IEC 8859-16:2001(E), ICS 35.040.10, International Organization for Standardization, Geneva, Switzerland, Jul. 15, 2001, 9 pages. Available at: https://www.iso.org/standard/33428.html.

[115] Raj J. Jaiswal and Nitin N. Patil, "Implementation of a new technique for web document protection using unicode," in *2013 International Conference on Information Communication and Embedded Systems (ICICES)*, 2013, pp. 69–72. DOI: 10.1109/ICICES.2013.6508287.

[116] Zunera Jalil and Anwar M. Mirza, "A Review of Digital Watermarking Techniques for Text Documents," in *2009 International Conference on Information and Multimedia Technology*, 2009, pp. 230–234. DOI: 10.1109/ICIMT.2009.11.

[117] Zunera Jalil, Anwar M. Mirza, and Tahir Iqbal, "A zero-watermarking algorithm for text documents based on structural components," in *2010 International Conference on Information and Emerging Technologies*, 2010, pp. 1–5. DOI: 10.1109/ICIET.2010.5625705.

[118] Shakthivelu Janardhanan, Maria Samonaki, Poul Einar Heegaard, Wolfgang Kellerer, and Carmen Mas-Machuca, "Network Sovereignty: A Novel Metric and Its Application on Network Design," *IEEE Transactions on Reliability*, vol. 74, no. 2, pp. 2927–2941, 2025. DOI: 10.1109/TR.2024.3434560.

[119] Matthias Jarke, Boris Otto, and Sudha Ram, "Data Sovereignty and Data Space Ecosystems," *Business & Information Systems Engineering*, vol. 61, no. 5, pp. 549–550, 2019. DOI: 10.1007/s12599-019-00614-2.

[120] Robert Jaros, *KVision*, 2025. Available at: https://github.com/rjaros/kvision (accessed on Feb. 10, 2026).

[121] Zhe Ji, Qiansiqi Hu, Yicheng Zheng, Liyao Xiang, and Xinbing Wang, "A Principled Approach to Natural Language Watermarking," in *Proceedings of the 32nd ACM International Conference on Multimedia*, 2024, pp. 2908–2916. DOI: 10.1145/3664647.3681544.

[122] Neil F. Johnson and Sushil Jajodia, "Exploring steganography: Seeing the unseen," *Computer*, vol. 31, no. 2, pp. 26–34, 1998. DOI: 10.1109/MC.1998.4655281.

[123] jscrambler. "Steganography in a Magecart Attack." 2022. Available at: https://jscrambler.com/blog/steganography-in-magecart-attack (accessed on Feb. 10, 2026).

[124] Henning Kagermann, Karl-Heinz Streibich, and Katrin Suder, *Digital Sovereignty, Status Quo and Perspectives*. Munich, Germany: acatech - National Academy of Science and Engineering, 2021, 29 pages, ISBN: 978-3-96834-011-1.

[125] David Kahn, "The History of Steganography," in *Information Hiding*, vol. 1174, 1996, pp. 1–5. DOI: 10.1007/3-540-61996-8_27.

[126] Nurul Shamimi Kamaruddin, Amirrudin Kamsin, Lip Yee Por, and Hameedur Rahman, "A Review of Text Watermarking: Theory, Methods, and Applications," *IEEE Access*, vol. 6, pp. 8011–8028, 2018. DOI: 10.1109/ACCESS.2018.2796585.

[127] Jaganmohan Reddy Kancharla and S. D. Madhu Kumar, "Advancing Data Sovereignty in Distributed Environments: An In-Depth Exploration of Data Localization Challenges," in *2024 International Conference on Computer, Electronics, Electrical Engineering & their Applications (IC2E3)*, 2024, pp. 1–6. DOI: 10.1109/IC2E362166.2024.10827688.

[128] Vasileios Karagiannis, Astrid Al-Akrawi, and Oliver Hödl, "Data Sovereignty at the Edge of the Network," in *2023 IEEE 7th International Conference on Fog and Edge Computing (ICFEC)*, 2023, pp. 33–39. DOI: 10.1109/ICFEC57925.2023.00013.

[129] Günter Karjoth, Matthias Schunter, and Michael Waidner, "Platform for Enterprise Privacy Practices: Privacy-Enabled Management of Customer Data," in *Privacy Enhancing Technologies, Second International Workshop, PET 2002*, vol. 2482, 2003, pp. 69–84. DOI: 10.1007/3-540-36467-6_6.

[130] Jan Martin Keil, "Efficient Bounded Jaro-Winkler Similarity Based Search," in *BTW - Datenbanksysteme für Business, Technologie und Web*, 2019, pp. 205–214. DOI: 10.18420/btw2019-13.

[131] Esam Ali Khan, "A High Capacity Steganography Method Utilizing Arabic Letters Characteristics and Poetry Binary Structure," in *2025 16th International Conference on Information and Communication Systems (ICICS)*, 2025, pp. 1–6. DOI: 10.1109/ICICS65354.2025.11073116.

[132] Behrooz Khosravi, Behnam Khosravi, Bahman Khosravi, and Khashayar Nazarkardeh, "A new method for pdf steganography in justified texts," *Journal of Information Security and Applications*, vol. 45, pp. 61–70, 2019. DOI: 10.1016/j.jisa.2019.01.003.

[133] Tong Min Kim, Taehoon Ko, Byoung Woo Hwang, Hyung Goo Paek, and Wan Yeon Lee, "Self-sovereign management scheme of personal health record with personal data store and decentralized identifier," *Computational and Structural Biotechnology Journal*, vol. 28, pp. 16–28, 2025. DOI: 10.1016/j.csbj.2024.11.036.

[134] John Kirchenbauer, Jonas Geiping, Yuxin Wen, Jonathan Katz, Ian Miers, and Tom Goldstein, "A Watermark for Large Language Models," in *Proceedings of the 40th International Conference on Machine Learning*, vol. 202, 2023, pp. 17061–17084. Available at: https://proceedings.mlr.press/v202/kirchenbauer23a/kirchenbauer23a.pdf.

[135] Mandy Knöchel and Sebastian Karius, "Text Steganography Methods and their Influence in Malware: A Comprehensive Overview and Evaluation," in *Proceedings of the 2024 ACM Workshop on Information Hiding and Multimedia Security*, 2024, pp. 113–124. DOI: 10.1145/3658664.3659637.

[136] Nils Köbis and Luca D. Mossink, "Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry," *Computers in Human Behavior*, vol. 114, art. no. 106553, 13 pages, 2021. DOI: 10.1016/j.chb.2020.106553.

[137] Peter Koen, Maarten Kollenstart, James Marino, Julia Pampus, Anil Turkmayali, Sebastian Steinbuss, and Arno Weiß. "Dataspace Protocol 2025-1." Sebastian Steinbuss, Ed. 2025. Available at: https://eclipse-dataspace-protocol-base.github.io/DataspaceProtocol/2025-1/ (accessed on Feb. 10, 2026).

[138] Jukka Korpela. "Unicode spaces." 2002. Available at: https://www.jkorpela.fi/chars/spaces.html (accessed on Feb. 10, 2026).

[139] Samuel Kounev, Klaus-Dieter Lange, and Jóakim von Kistowski, *Systems Benchmarking*. Cham, Switzerland: Springer International Publishing, 2020. DOI: 10.1007/978-3-030-41705-5.

[140] Anhelina Kovach, Leticia Montalvillo, Jorge Lanza, Pablo Sotres, and Aitor Urbieta, "Understanding data spaces: A Systematic Mapping Study of foundations, technical building blocks, and sectoral adoption," *Computer Science Review*, vol. 59, art. no. 100819, 27 pages, 2026. DOI: 10.1016/j.cosrev.2025.100819.

[141] Stephen D. Krasner, *Problematic Sovereignty*. Columbia University Press, 2001. DOI: 10.7312/kras12178.

[142] Ramakrishnan Bala Krishnan, Prasanth Kumar Thandra, and Magapu Sai Baba, "An Overview of Text Steganography," in *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, 2017, pp. 1–6. DOI: 10.1109/ICSCN.2017.8085643.

[143] Sanjay Kumar, Binod Kumar Singh, and Mohit Yadav, "A Recent Survey on Multimedia and Database Watermarking," *Multimedia Tools and Applications*, vol. 79, no. 27, pp. 20149–20197, 2020. DOI: 10.1007/s11042-020-08881-y.

[144] KuroLabs, *StegCloak*, 2020. Available at: https://github.com/KuroLabs/stegcloak (accessed on Feb. 10, 2026).

[145] Neal Kushwaha, Przemyslaw Roguski, and Bruce W. Watson, "Up in the Air: Ensuring Government Data Sovereignty in the Cloud," in *2020 12th International Conference on Cyber Conflict (CyCon)*, 2020, pp. 43–61. DOI: 10.23919/CyCon49761.2020.9131718.

[146] Oleksandr Kuznetsov, Emanuele Frontoni, Kyrylo Chernov, Marco Amesano, and Cristian Randieri, "Securing Digital Communications with AI-Enhanced Synonym Substitution in Text," in *Proceedings of the 10th International Conference of Yearly Reports on Informatics, Mathematics, and Engineering*, 2025, pp. 116–124. Available at: https://ceur-ws.org/Vol-3984/p13.pdf.

[147] Matthew Kwan. "How SNOW works." 2013. Available at: https://darkside.com.au/snow/description.html (accessed on Feb. 10, 2026).

[148] Matthew Kwan, *SNOW*, mattkwan-zz, 2016. Available at: https://github.com/mattkwan-zz/snow (accessed on Feb. 10, 2026).

[149] Matthew Kwan. "The SNOW Home Page." 2025. Available at: https://darkside.com.au/snow/ (accessed on Feb. 10, 2026).

[150] Giuseppe Landolfi, Andrea Barni, Gabriele Izzo, Alessandro Fontana, and Andrea Bettoni, "A MaaS platform architecture supporting data sovereignty in sustainability assessment of manufacturing systems," *Procedia Manufacturing*, vol. 38, pp. 548–555, 2019. DOI: 10.1016/j.promfg.2020.01.069.

[151] Kai R. Larsen, Roman Lukyanenko, Roland M. Mueller, Veda C. Storey, Jeffrey Parsons, Debra Vandermeer, and Dirk S. Hovorka, "Validity in Design Science," *MIS Quarterly*, vol. 49, no. 4, pp. 1267–1294, 2025. DOI: 10.25300/MISQ/2024/18064.

[152] Florian Lauf, Simon Scheider, Sven Meister, Marija Radic, Philipp Herrmann, Max Schulze, André T. Nemat, Sarah J. Becker, Marcel Rebbert, Constantin Abate, Ralf Konrad, Jan Bartsch, Tobias Dehling, and Ali Sunyaev, "Data Sovereignty and Data Economy—Two Repulsive Forces?" Dortmund, Germany, 2021, 37 pages. DOI: 10.24406/ISST-N-634865.

[153] Ah Ra Lee, Min Gyu Kim, Kyung Jae Won, Il Kon Kim, and Eunjoo Lee, "Coded Dynamic Consent framework using blockchain for healthcare information exchange," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 1047–1050. DOI: 10.1109/BIBM49941.2020.9313330.

[154] I-Shi Lee and Wen-Hsiang Tsai, "A new approach to covert communication via PDF files," *Signal Processing*, vol. 90, no. 2, pp. 557–565, 2010. DOI: 10.1016/j.sigpro.2009.07.022.

[155] Jong Seok Lee, Jan Pries-Heje, and Richard Baskerville, "Theorizing in Design Science Research," in *Service-Oriented Perspectives in Design Science Research*, ser. Lecture Notes in Computer Science, Hemant Jain, Atish P. Sinha, and Padmal Vitharana, Eds., vol. 6629, Berlin, Heidelberg, Germany: Springer Berlin Heidelberg, 2011, pp. 1–16. DOI: 10.1007/978-3-642-20633-7_1.

[156] Min-Jeong Lee, Kyung-Su Kim, and Heung-Kyu Lee, "Digital Cinema Watermarking for Estimating the Position of the Pirate," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 605–621, 2010. DOI: 10.1109/TMM.2010.2061221.

[157] Yue Li, Hongxia Wang, and Mauro Barni, "A survey of Deep Neural Network watermarking techniques," *Neurocomputing*, vol. 461, pp. 171–193, 2021. DOI: 10.1016/j.neucom.2021.07.051.

[158] Aiwei Liu, Leyi Pan, Yijian Lu, Jingjing Li, Xuming Hu, Xi Zhang, Lijie Wen, Irwin King, Hui Xiong, and Philip Yu, "A Survey of Text Watermarking in the Era of Large Language Models," *ACM Computing Surveys*, vol. 57, no. 2, pp. 1–36, 2025. DOI: 10.1145/3691626.

[159] Qiong Liu, Reihaneh Safavi-Naini, and Nicholas Paul Sheppard, "Digital Rights Management for Content Distribution," in *Proceedings of the Australasian Information Security Workshop Conference on ACSW Frontiers 2003*, Australian Computer Society, Inc., vol. 21, 2003, pp. 49–58. Available at: https://dl.acm.org/doi/10.5555/827987.827994.

[160] Shuwen Liu and George C. Polyzos, "SovereignEdge: A Context-Aware Cryptographic Architecture for Data Sovereignty in Mobile Edge–Fog–Cloud IoT Data Spaces," in *2025 8th World Conference on Computing and Communication Technologies (WCCCT)*, 2025, pp. 166–175. DOI: 10.1109/WCCCT65447.2025.11027965.

[161] Johannes Lohmöller, Roman Matzutt, Joscha Loos, Eduard Vlad, Jan Pennekamp, and Klaus Wehrle, "Complementing Organizational Security in Data Ecosystems with Technical Guarantees," in *2024 Conference on Building a Secure & Empowered Cyberspace (BuildSEC)*, 2024, pp. 49–56. DOI: 10.1109/BuildSEC64048.2024.00016.

[162] Johannes Lohmöller, Jan Pennekamp, Roman Matzutt, Carolin Victoria Schneider, Eduard Vlad, Christian Trautwein, and Klaus Wehrle, "The unresolved need for dependable guarantees on security, sovereignty, and trust in data ecosystems," *Data & Knowledge Engineering*, vol. 151, art. no. 102301, 2024. DOI: 10.1016/j.datak.2024.102301.

[163] Charles E. Mackenzie, *Coded Character Sets, History and Development*. Reading, MA, USA: Addison-Wesley, 1980, 513 pages, ISBN: 0-201-14460-3.

[164] Mohammed Abdul Majeed, Rossilawati Sulaiman, and Zarina Shukur, "New Text Steganography Technique based on Multilayer Encoding with Format-Preserving Encryption and Huffman Coding," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 12, pp. 163–172, 2022. DOI: 10.14569/IJACSA.2022.0131222.

[165] Mohammed Abdul Majeed, Rossilawati Sulaiman, and Zarina Shukur, "New Text Steganography Technique Based on Part-of-Speech Tagging and Format-Preserving Encryption," *KSII Transactions on Internet and Information Systems*, vol. 18, no. 1, pp. 170–191, 2024. DOI: 10.3837/tiis.2024.01.010.

[166] Mohammed Abdul Majeed, Rossilawati Sulaiman, Zarina Shukur, and Mohammad Kamrul Hasan, "A Review on Text Steganography Techniques," *Mathematics*, vol. 9, no. 21, art. no. 2829, 28 pages, 2021. DOI: 10.3390/math9212829.

[167] Felix Mannhardt, Agnes Koschmider, Nathalie Baracaldo, Matthias Weidlich, and Judith Michael, "Privacy-Preserving Process Mining - Differential Privacy for Event Logs," *Business & Information Systems Engineering*, vol. 61, no. 5, pp. 595–614, 2019. DOI: 10.1007/s12599-019-00613-3.

[168] Marvin Manoury, Theresa Riedelsheimer, Malte Hellmeier, and Tom Meyer, "Supporting Changes in Digital Ownership and Data Sovereignty Across the Automotive Value Chain with Catena-X," *Procedia Computer Science*, vol. 253, pp. 374–383, 2025. DOI: 10.1016/j.procs.2025.01.099.

[169] Christoph March and Ina Schieferdecker, "Technological Sovereignty as Ability, Not Autarky," *SSRN Electronic Journal*, pp. 1–39, 2021. DOI: 10.2139/ssrn.3872378.

[170] Salvatore T. March and Gerald F. Smith, "Design and natural science research on information technology," *Decision Support Systems*, vol. 15, no. 4, pp. 251–266, 1995. DOI: 10.1016/0167-9236(94)00041-2.

[171] Katrin Martens and Jana Zscheischler, "The Digital Transformation of the Agricultural Value Chain: Discourses on Opportunities, Challenges and Controversial Perspectives on Governance Approaches," *Sustainability*, vol. 14, no. 7, art. no. 3905, 15 pages, 2022. DOI: 10.3390/su14073905.

[172] Tim Maurer, Isabel Skierka, Robert Morgus, and Mirko Hohmann, "Technological sovereignty: Missing the point?" In *2015 7th International Conference on Cyber Conflict: Architectures in Cyberspace*, 2015, pp. 53–68. DOI: 10.1109/CYCON.2015.7158468.

[173] Marina Micheli, Marisa Ponti, Max Craglia, and Anna Berti Suman, "Emerging models of data governance in the age of datafication," *Big Data & Society*, vol. 7, no. 2, pp. 1–15, 2020. DOI: 10.1177/2053951720948087.

[174] Daniele Miorandi, Alessandra Rizzardi, Sabrina Sicari, and Alberto Coen-Porisini, "Sticky Policies: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 12, pp. 2481–2499, 2020. DOI: 10.1109/TKDE.2019.2936353.

[175] Nighat Mir, "Copyright for web content using invisible text watermarking," *Computers in Human Behavior*, vol. 30, pp. 648–653, 2014. DOI: 10.1016/j.chb.2013.07.040.

[176] Mohanasundar. "How to Hide Secrets in Strings— Modern Text hiding in JavaScript." 2020. Available at: https://blog.bitsrc.io/how-to-hide-secrets-in-strings-modern-text-hiding-in-javascript-613a9faa5787 (accessed on Feb. 10, 2026).

[177] Saraju P. Mohanty, Anirban Sengupta, Parthasarathy Guturu, and Elias Kougianos, "Everything You Want to Know About Watermarking: From Paper Marks to Hardware Protection," *IEEE Consumer Electronics Magazine*, vol. 6, no. 3, pp. 83–91, 2017. DOI: 10.1109/MCE.2017.2684980.

[178] Frederik Möller, Tobias Moritz Guggenberger, and Boris Otto, "Towards a Method for Design Principle Development in Information Systems," in *Designing for Digital Transformation. Co-Creating Services with Citizens and Industry*, ser. Lecture Notes in Computer Science, vol. 12388, Sara Hofmann, Oliver Müller, and Matti Rossi, Eds., Cham, Switzerland: Springer International Publishing, 2020, pp. 208–220. DOI: 10.1007/978-3-030-64823-7_20.

[179] Frederik Möller, Ilka Jussen, Virginia Springer, Anna Gieß, Julia Christina Schweihoff, Joshua Gelhaar, Tobias Guggenberger, and Boris Otto, "Industrial data ecosystems and data spaces," *Electronic Markets*, vol. 34, art. no. 41, 17 pages, 2024. DOI: 10.1007/s12525-024-00724-0.

[180] Gordon E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, pp. 114–117, 1965.

[181] Lukas Moschko, Vera Blazevic, and Frank T. Piller, "Managing Data Sovereignty: An Organizational Competence for Successful Open Value Creation," *R&D Management*, vol. 55, no. 4, pp. 1124–1137, 2024. DOI: 10.1111/radm.12740.

[182] Andres Muñoz-Arcentales, Sonsoles López-Pernas, Alejandro Pozo, Álvaro Alonso, Joaquín Salvachúa, and Gabriel Huecas, "An Architecture for Providing Data Usage and Access Control in Data Sharing Ecosystems," *Procedia Computer Science*, vol. 160, pp. 590–597, 2019. DOI: 10.1016/j.procs.2019.11.042.

[183] Lars Nagel and Douwe Lycklama, "Design Principles for Data Spaces - Position Paper," version 1.0, 2021. DOI: 10.5281/ZENODO.5105744.

[184] Alfin Naharuddin, Adhi Dharma Wibawa, and Surya Sumpeno, "A High Capacity and Imperceptible Text Steganography Using Binary Digit Mapping on ASCII Characters," in *2018 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2018, pp. 287–292. DOI: 10.1109/ISITIA.2018.8711087.

[185] Shreya KL Narasimhan and Bala R. Krishnan, "Text steganography: enhanced character-level embedding algorithm using font attribute with increased resilience to statistical attacks," *Multimedia Tools and Applications*, 2024. DOI: 10.1007/s11042-024-19272-y.

[186] Michael Nast, Benjamin Rother, Frank Golatowski, Dirk Timmermann, Jens Leveling, Christian Olms, and Christian Nissen, "Work-in-Progress: Towards an International Data Spaces Connector for the Internet of Things," in *2020 16th IEEE International Conference on Factory Communication Systems (WFCS)*, 2020, pp. 1–4. DOI: 10.1109/WFCS47810.2020.9114503.

[187] Alexander Nozik, "Kotlin language for science and Kmath library," *AIP Conference Proceedings*, vol. 2163, no. 1, art. no. 040004, 5 pages, 2019. DOI: 10.1063/1.5130103.

[188] Yudhistira Nugraha, Kautsarina, and Ashwin Sasongko Sastrosubroto, "Towards data sovereignty in cyberspace," in *2015 3rd International Conference on Information and Communication Technology (ICoICT)*, 2015, pp. 465–471. DOI: 10.1109/icoict.2015.7231469.

[189] Jay F. Nunamaker, Minder Chen, and Titus D. M. Purdin, "Systems Development in Information Systems Research," *Journal of Management Information Systems*, vol. 7, no. 3, pp. 89–106, 1990. Available at: https://www.jstor.org/stable/40397957.

[190] Sebastian Opriel, Frederik Möller, Gero Strobel, and Boris Otto, "Data Sovereignty in Inter-organizational Information Systems," *Business & Information Systems Engineering*, 2024. DOI: 10.1007/s12599-024-00893-4.

[191] Boris Otto, "Interview with Reinhold Achatz on "Data Sovereignty and Data Ecosystems"," *Business & Information Systems Engineering*, vol. 61, no. 5, pp. 635–636, 2019. DOI: 10.1007/s12599-019-00609-z.

[192] Boris Otto, "GAIA-X and IDS," version 1.0, 2021, 33 pages. DOI: 10 . 5281 / ZENODO.5269077.

[193] Boris Otto and Matthias Jarke, "Designing a multi-sided data platform: findings from the International Data Spaces case," *Electronic Markets*, vol. 29, no. 4, pp. 561–580, 2019. DOI: 10.1007/s12525-019-00362-x.

[194] Julia Pampus and Maritta Heisel, "A Delimitation of Data Sovereignty From Privacy," in *Empowering Digital Sovereignty*, Sanjay Misra, Petter Kvalvik, Kai Morgan Kjølerbakken, and Per-Arne Jørgensen, Eds., IGI Global Scientific Publishing, 2025, pp. 1–24. DOI: 10.4018/979-8-3693-9137-2.ch001.

[195] Siani Pearson and Marco Casassa-Mont, "Sticky Policies: An Approach for Managing Privacy across Multiple Parties," *Computer*, vol. 44, no. 9, pp. 60–68, 2011. DOI: 10.1109/MC.2011.225.

[196] Ken Peffers, Tuure Tuunanen, Charles E. Gengler, Matti Rossi, Wendy Hui, Ville Virtanen, and Johanna Bragge, "The Design Science Research Process: A Model for Producing and Presenting Information Systems Research," in *Proceedings of the 1st International Conference on Design Science Research in Information Systems and Technology (DESRIST)*, 2006, pp. 83–106.

[197] Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee, "A Design Science Research Methodology for Information Systems Research," *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007. DOI: 10.2753/MIS0742-1222240302.

[198] Fabien A. P. Petitcolas, Ross Anderson, and Markus G. Kuhn, "Information Hiding – A Survey," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1062–1078, 1999. DOI: 10.1109/5.771065.

[199] Daniel Philpott, "Sovereignty: An Introduction and Brief History," *Journal of International Affairs*, vol. 48, no. 2, pp. 353–368, 1995. Available at: https://www.jstor.org/stable/24357595.

[200] Aude Plateaux, Patrick Lacharme, Christophe Rosenberger, and Kumar Murty, "A contactless e-health information system with privacy," in *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2013, pp. 1660–1665. DOI: 10.1109/IWCMC.2013.6583805.

[201] Julia Pohle and Thorsten Thiel, "Digital sovereignty," *Internet Policy Review*, vol. 9, no. 4, pp. 1–19, 2020. DOI: 10.14763/2020.4.1532.

[202] Dana Polatin-Reuben and Joss Wright, "An Internet with BRICS Characteristics: Data Sovereignty and the Balkanisation of the Internet," in *4th USENIX Workshop on Free and Open Communications on the Internet (FOCI 14)*, 2014. Available at: https://www.usenix.org/conference/foci14/workshop-program/presentation/polatin-reuben.

[203] Lip Yee Por, Tan Fong Ang, and Beh Mei Yin Delina, "WhiteSteg: A New Scheme in Information Hiding Using Text Steganography," *WSEAS Transactions on Computers*, vol. 7, no. 6, pp. 735–745, 2008. Available at: https://dl.acm.org/doi/10.5555/1458369.1458384.

[204] Lip Yee Por, KokSheik Wong, and Kok Onn Chee, "UniSpaCh: A text-based data hiding method using Unicode space characters," *Journal of Systems and Software*, vol. 85, no. 5, pp. 1075–1082, 2012. DOI: 10.1016/j.jss.2011.12.023.

[205] Alexander Pretschner, Manuel Hilty, Florian Schütz, Christian Schaefer, and Thomas Walter, "Usage Control Enforcement: Present and Future," *IEEE Security & Privacy Magazine*, vol. 6, no. 4, pp. 44–53, 2008. DOI: 10.1109/MSP.2008.101.

[206] Sandeep Purao, Leona Chandra Kruse, and Alexander Maedche, "The Origins of Design Principles: Where do… they all come from?" In *Designing for Digital Transformation. Co-Creating Services with Citizens and Industry*, ser. Lecture Notes in Computer Science, vol. 12388, Sara Hofmann, Oliver Müller, and Matti Rossi, Eds., Cham, Switzerland: Springer International Publishing, 2020, pp. 183–194. DOI: 10.1007/978-3-030-64823-7_17.

[207] Muhammad Qadir and Ishtiaq Ahmad, "Digital Text Watermarking: Secure Content Delivery and Data Hiding in Digital Documents," *IEEE Aerospace and Electronic Systems Magazine*, vol. 21, no. 11, pp. 18–21, 2006. DOI: 10.1109/MAES.2006.284353.

[208] Haydar Qarawlus, Malte Hellmeier, and Falk Howar, "Facilitating Data Usage Control Through IPv6 Extension Headers," in *Proceedings of the 14th International Conference on Data Science, Technology and Applications*, 2025, pp. 526–535. DOI: 10.5220/0013566200003967.

[209] Haydar Qarawlus, Malte Hellmeier, Johannes Pieperbeck, Ronja Quensel, Steffen Biehs, and Marc Peschke, "Sovereign Data Exchange in Cloud-Connected IoT using International Data Spaces," in *2021 IEEE Cloud Summit (Cloud Summit)*, 2021, pp. 13–18. DOI: 10.1109/IEEECloudSummit52029.2021.00010.

[210] Wang Qi, Bei Yue, Chen Wangdu, Pan Xinghao, Cheng Zhipeng, Wang Shaokang, Wang Yizhao, and Wang Chenwei, "An Overview on Digital Content Watermarking," in *Signal and Information Processing, Networking and Computers*, ser. Lecture Notes in Electrical Engineering, vol. 917, Jiande Sun, Yue Wang, Mengyao Huo, and Lexi Xu, Eds., Singapore, Singapore: Springer Nature Singapore, 2023, pp. 1311–1318. DOI: 10.1007/978-981-19-3387-5_157.

[211] Zhining Qin, David W. Johnson, and Roger T. Johnson, "Cooperative versus Competitive Efforts and Problem Solving," *Review of Educational Research*, vol. 65, no. 2, pp. 129–143, 1995. DOI: 10.2307/1170710.

[212] Khan Farhan Rafat and M. Sher, "Secure Digital Steganography for ASCII Text Documents," *Arabian Journal for Science and Engineering*, vol. 38, no. 8, pp. 2079–2094, 2013. DOI: 10.1007/s13369-013-0574-5.

[213] Kumar Rahul and Rohitash Kumar Banyal, "Data Life Cycle Management in Big Data Analytics," *Procedia Computer Science*, vol. 173, pp. 364–371, 2020. DOI: 10.1016/j.procs.2020.06.042.

[214] Arun Rai, "Editor's Comments: Diversity of Design Science Research," *MIS Quarterly*, vol. 41, no. 1, pp. iii–xviii, 2017. Available at: https://aisel.aisnet.org/misq/vol41/iss1/2/.

[215] Rishav Ray, Jeeyan Sanyal, Debanjan Das, and Asoke Nath, "A New Challenge of Hiding any Encrypted Secret Message inside any Text/ASCII File or in MS Word File: RJDA Algorithm," in *2012 International Conference on Communication Systems and Network Technologies*, 2012, pp. 889–893. DOI: 10.1109/CSNT.2012.191.

[216] Jan Recker, *Scientific research in information systems, A beginner's guide*, 2nd ed. Cham, Switzerland: Springer, 2021, 221 pages, ISBN: 978-3-030-85436-2.

[217] Magnus Redeker, Soren Volgmann, Florian Pethig, and Johannes Kalhoff, "Towards Data Sovereignty of Asset Administration Shells across Value Added Chains," in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2020, pp. 1151–1154. DOI: 10.1109/ETFA46521.2020.9211955.

[218] Stefano Giovanni Rizzo, Flavio Bertini, and Danilo Montesi, "Content-preserving Text Watermarking through Unicode Homoglyph Substitution," in *Proceedings of the 20th International Database Engineering & Applications Symposium on - IDEAS '16*, 2016, pp. 97–104. DOI: 10.1145/2938503.2938510.

[219] Stefano Giovanni Rizzo, Flavio Bertini, and Danilo Montesi, "Fine-grain watermarking for intellectual property protection," *EURASIP Journal on Information Security*, vol. 2019, art. no. 10, 20 pages, 2019. DOI: 10.1186/s13635-019-0094-2.

[220] Stefano Giovanni Rizzo, Flavio Bertini, Danilo Montesi, and Carlo Stomeo, "Text Watermarking in Social Media," in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 2017, pp. 208–211. DOI: 10.1145/3110025.3116203.

[221] Nabila B. Rofiatunnajah and Ari M. Barmawi, "Improving ANiTW Performance Using Bigrams Character Encoding and Identity-Based Signature," *IEEE Access*, vol. 11, pp. 24257–24280, 2023. DOI: 10.1109/ACCESS.2023.3254586.

[222] Siddhartha Deb Roy, Ankush Sharma, Saikat Chakrabarti, and Sanjoy Debbarma, "Securing Power System Data in Motion by Timestamped Digital Text Watermarking," *IEEE Transactions on Smart Grid*, vol. 15, no. 5, pp. 4974–4985, 2024. DOI: 10.1109/TSG.2024.3370892.

[223] Mark Ryan, Paula Gürtler, and Artur Bogucki, "Will the real data sovereign please stand up? An EU policy response to sovereignty in data spaces," *International Journal of Law and Information Technology*, vol. 32, art. no. eaae006, 32 pages, 2024. DOI: 10.1093/ijlit/eaae006.

[224] Hyunho Ryu, Hyunsung Kim, Saurabh Agarwal, Dilip Kumar Sharma, Beaton Kapito, and Patrick Ali, "Data Sovereignty Provision Blockchain for Remote Healthcare Service," in *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, 2023, pp. 1–4. DOI: 10.1109/ISCON57294.2023.10112016.

[225] Marcelo Iury S. Oliveira, Glória de Fátima Barros Lima, and Bernadette Farias Lóscio, "Investigations into Data Ecosystems: a systematic mapping study," *Knowledge and Information Systems*, vol. 61, no. 2, pp. 589–630, 2019. DOI: 10.1007/s10115-018-1323-6.

[226] Johnny Saldaña, *The coding manual for qualitative researchers*, 2nd ed. Los Angeles, CA, USA: Sage, 2013, 303 pages, ISBN: 978-1-44624-736-5.

[227] David Sarabia-Jacome, Ignacio Lacalle, Carlos E. Palau, and Manuel Esteve, "Enabling Industrial Data Space Architecture for Seaport Scenario," in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, 2019, pp. 101–106. DOI: 10.1109/WF-IoT.2019.8767216.

[228] Ryoma Sato, Yuki Takezawa, Han Bao, Kenta Niwa, and Makoto Yamada, "Embarrassingly Simple Text Watermarks," 2023, Preprint. DOI: 10.48550/arXiv.2310.08920.

[229] Adem Savaş and Mehmet Zeki Konyar, "A New Reversible Data Hiding Method in PDF Files for Secret Communication," in *2025 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2025, pp. 1–6. DOI: 10.1109/BlackSeaCom65655.2025.11193925.

[230] Simon Scheider, Florian Lauf, Frederik Möller, and Boris Otto, "A Reference System Architecture with Data Sovereignty for Human-Centric Data Ecosystems," *Business & Information Systems Engineering*, vol. 65, pp. 577–595, 2023. DOI: 10.1007/s12599-023-00816-9.

[231] Kaja Schmidt, Gonzalo Munilla Garrido, Alexander Mühle, and Christoph Meinel, "Mitigating Sovereign Data Exchange Challenges: A Mapping to Apply Privacy- and Authenticity-Enhancing Technologies," in *Trust, Privacy and Security in Digital Business*, ser. Lecture Notes in Computer Science, vol. 13582, Sokratis Katsikas and Steven Furnell, Eds., Cham, Switzerland: Springer International Publishing, 2022, pp. 50–65. DOI: 10.1007/978-3-031-17926-6_4.

[232] Maung K. Sein, Ola Henfridsson, Sandeep Purao, Matti Rossi, and Rikard Lindgren, "Action Design Research," *MIS Quarterly*, vol. 35, no. 1, pp. 37, 2011. DOI: 10.2307/23043488.

[233] Mohammad Shirali Shahreza, "A New Method for Steganography in HTML Files," in *Advances in Computer, Information, and Systems Sciences, and Engineering*, Khaled Elleithy, Tarek Sobh, Ausif Mahmood, Magued Iskander, and Mohammad Karim, Eds., Dordrecht, Netherlands: Springer Dordrecht, 2006, pp. 247–252. DOI: 10.1007/1-4020-5261-8_39.

[234] Sunpreet Sharma, Ju Jia Zou, Gu Fang, Pancham Shukla, and Weidong Cai, "A review of image watermarking for identity protection and verification," *Multimedia Tools and Applications*, vol. 83, no. 11, pp. 31829–31891, 2024. DOI: 10.1007/s11042-023-16843-3.

[235] Md. Shazzad-Ur-Rahman, Md. Mahib Hosen Ornob, Amit Singha, Md. Shamim Kaiser, and Nahid Ibne Akhter, "An Effective Text Steganographic Scheme Based on Multilingual Approach for Secure Data Communication," in *2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, 2021, pp. 1–8. DOI: 10.1109/ICIEVicIVPR52578.2021.9564231.

[236] Md. Shazzad-Ur-Rahman, Md. Shamim Kaiser, Meherun Bintey Alam, and Sifat Nawrin Nova, "A Data Hiding Technique Combining Steganography and Cryptography for Secured Communication," in *2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 2023, pp. 432–437. DOI: 10.1109/ICICT4SD59951.2023.10303563.

[237] Zhenhao Shi, Hongxia Wang, Heng Wang, and Xinyi Huang, "Robust Screen-Shooting Document Watermarking for Multiple Fonts," *IEEE Signal Processing Letters*, vol. 31, pp. 2215–2219, 2024. DOI: 10.1109/LSP.2024.3449231.

[238] Hung-Jr Shiu, Bor-Shing Lin, Bor-Shyh Lin, Po-Yang Huang, Chien-Hung Huang, and Chin-Laung Lei, "Data Hiding on Social Media Communications Using Text Steganography," in *Risks and Security of Internet and Systems*, ser. Lecture Notes in Computer Science, vol. 10694, Nora Cuppens, Frédéric Cuppens, Jean-Louis Lanet, Axel Legay, and Joaquin Garcia-Alfaro, Eds., Cham, Switzerland: Springer International Publishing, 2018, pp. 217–224. DOI: 10.1007/978-3-319-76687-4_15.

[239] Gustavus J. Simmons, "The Prisoners' Problem and the Subliminal Channel," in *Advances in Cryptology*, David Chaum, Ed., 1st ed., Boston, MA, USA: Springer US, 1984, pp. 51–67. DOI: 10.1007/978-1-4684-4730-9_5.

[240] Herbert Alexander Simon, *The sciences of the artificial*, 3rd ed. Cambridge, MA, USA: MIT Press, 1996, 231 pages, ISBN: 0-262-19374-4.

[241] Denis Sinegubko. "Whitespace Steganography Conceals Web Shell in PHP Malware." 2021. Available at: https://blog.sucuri.net/2021/02/whitespace-steganography-conceals-web-shell-in-php-malware.html (accessed on Feb. 10, 2026).

[242] Ashneet Khandpur Singh, Marcel Ortiz Sánchez, Marc Garnica Caparros, Mario Reyes de los Mozos, Silvia Castellvi, Simon Dalmolen, Kosmas Tsiakas, Paschalis Itsios, Ioannis Mariolis, Dimitrios Giakoumis, Silvia Rodríguez, Asier Aguayo Velasco, Daniel Cabello, and Borja Perez Lopez, "A Security Analysis of European Data Space Architectures," *Data Science and Engineering*, 2025. DOI: 10.1007/s41019-025-00311-z.

[243] Om Prakash Singh, Amit Kumar Singh, Gautam Srivastava, and Neeraj Kumar, "Image watermarking using soft computing techniques: A comprehensive survey," *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 30367–30398, 2021. DOI: 10.1007/s11042-020-09606-x.

[244] Kapil Singi, Swapnajeet Gon Choudhury, Vikrant Kaulgud, R. Jagadeesh ChandraP. Bose, Sanjay Podder, and Adam P. Burden, "Data Sovereignty Governance Framework," in *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, 2020, pp. 303–306. DOI: 10.1145/3387940.3392212.

[245] Arezou Soltani Panah, Ron van Schyndel, Timos Sellis, and Elisa Bertino, "On the Properties of Non-Media Digital Watermarking: A Review of State of the Art Techniques," *IEEE Access*, vol. 4, pp. 2670–2704, 2016. DOI: 10.1109/ACCESS.2016.2570812.

[246] Erik Sonnleitner, "A robust watermarking approach for large databases," in *2012 IEEE First AESS European Conference on Satellite Telecommunications (ESTEL)*, 2012, pp. 1–6. DOI: 10.1109/ESTEL.2012.6400082.

[247] Statista. "Volume of data or information created, captured, copied, and consumed worldwide from 2010 to 2029 (in zettabytes) [Graph]," Statista. 2025. Available at: https://www.statista.com/statistics/871513/worldwide-data-created/ (accessed on Feb. 10, 2026).

[248] Andreas Streim and David Schönwerth, *Unternehmen wollen Daten nutzen, aber nicht teilen [Companies want to use data, but not share it]*, bitkom, Ed., Berlin, Germany, May 10, 2023. Available at: https://www.bitkom.org/print/pdf/node/18718 (accessed on Feb. 10, 2026).

[249] Jonathan K. Su, Frank Hartung, and Bernd Girod, "Digital watermarking of text, image, and video documents," *Computers & Graphics*, vol. 22, no. 6, pp. 687–695, 1998. DOI: 10.1016/S0097-8493(98)00089-2.

[250] Xingming Sun, Gang Luo, and Huajun Huang, "Component-based digital watermarking of Chinese texts," in *Proceedings of the 3rd international conference on Information security*, 2004, pp. 76–81. DOI: 10.1145/1046290.1046306.

[251] Kheng Leong Tan, Chi-Hung Chi, and Kwok-Yan Lam, "Survey on Digital Sovereignty and Identity: From Digitization to Digitalization," *ACM Computing Surveys*, vol. 56, no. 3, pp. 1–36, 2024. DOI: 10.1145/3616400.

[252] Ruixiang Tang, Yu-Neng Chuang, and Xia Hu, "The Science of Detecting LLM-Generated Text," *Communications of the ACM*, vol. 67, no. 4, pp. 50–59, 2024. DOI: 10.1145/3624725.

[253] Yaodong Tao, Shuai Yang, and Hongmei Ge, "Comparative Study on Data Sovereignty Guarantee Technology," in *2022 IEEE 13th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP)*, 2022, pp. 1–6. DOI: 10.1109/PAAP56126.2022.10010593.

[254] John Taylor and Tahu Kukutai, Eds., *Indigenous Data Sovereignty, Toward an Agenda*, vol. 38, The Australian National University Research Monograph, Acton, ACT, Australia: Australian National University Press, 2016, 318 pages, ISBN: 9781760460303.

[255] Richard D. Taylor, ""Data localization": The internet in the balance," *Telecommunications Policy*, vol. 44, no. 8, art. no. 102003, 15 pages, 2020. DOI: 10.1016/j.telpol.2020.102003.

[256] The Eclipse Foundation, *EDC Connector*, 2024. Available at: https://github.com/eclipse-edc/Connector (accessed on Feb. 10, 2026).

[257] The Unicode Consortium, *The Unicode Standard, Version 17.0.0*. South San Francisco, CA, USA: The Unicode Consortium, 2025, available at: https://unicode.org/versions/Unicode17.0.0/, ISBN: 978-1-936213-35-1.

[258] Graham Thompson, *pyUnicodeSteganography Lookalikes*, 2021. Available at: https://github.com/bunnylab/pyUnicodeSteganography/blob/main/pyUnicodeSteganography/lookalikes.py (accessed on Feb. 10, 2026).

[259] Nguyen Hoang Thuan, Andreas Drechsler, and Pedro Antunes, "Construction of Design Science Research Questions," *Communications of the Association for Information Systems*, vol. 44, pp. 332–363, 2019. DOI: 10.17705/1CAIS.04420.

[260] Rajesh Kumar Tiwari and G. Sahoo, "A Novel Methodology for Data Hiding in PDF Files," *Information Security Journal: A Global Perspective*, vol. 20, no. 1, pp. 45–57, 2011. DOI: 10.1080/19393555.2010.544703.

[261] Mercan Topkara, Cuneyt M. Taskiran, and Edward J. Delp, "Natural language watermarking," in *Security, Steganography, and Watermarking of Multimedia Contents VII*, 2005, pp. 441–452. DOI: 10.1117/12.593790.

[262] Lalit Kumar Tyagi, Anish Gupta, and Aezeden Mohamed, "Unveiling the Invisible an In-Depth Analysis of Text Steganography Techniques, Challenges, and Advancement," in *2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, 2023, pp. 177–183. DOI: 10.1109/ICTACS59847.2023.10390024.

[263] Miroslav Vacura, "Watermark as a Tool to Address Abuse of Large-Scale Language Models," in *IDIMT-2024, Changes to ICT, management and business process through AI : 32nd Interdisciplinary Information Management Talks : Sept. 4-6, 2024, Hradec Králové, Czech Republic*, ser. Schriftenreihe Informatik, vol. 53, Petr Doucek, Michael Sonntag, and Lea Nedomova, Eds., Linz: Trauner Verlag, 2024, pp. 419–426. DOI: 10.35011/IDIMT-2024-419.

[264] David Vaile, "The Cloud and data sovereignty after Snowden," *Australian Journal of Telecommunications and the Digital Economy*, vol. 2, no. 1, pp. 1–59, 2014. DOI: 10.7790/ajtde.v2n1.31.

[265] Jan van den Akker, "Principles and Methods of Development Research," in *Design Approaches and Tools in Education and Training*, Jan van den Akker, Robert Maribe Branch, Kent Gustafson, Nienke Nieveen, and Tjeerd Plomp, Eds., Dordrecht, Netherlands: Springer Netherlands, 1999, pp. 1–14. DOI: 10.1007/978-94-011-4255-7_1.

[266] John R. Venable, Jan Pries-Heje, and Richard Baskerville, "Choosing a Design Science Research Methodology," *ACIS 2017 Proceedings*, art. no. 112, 2017. Available at: https://aisel.aisnet.org/acis2017/112/.

[267] Stefaan G. Verhulst, "Operationalizing digital self-determination," *Data & Policy*, vol. 5, art. no. e14, 17 pages, 2023. DOI: 10.1017/dap.2023.11.

[268] Jan vom Brocke, Alexander Simons, Kai Riemer, Björn Niehaves, Ralf Plattfaut, and Anne Cleven, "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research," *Communications of the Association for Information Systems*, vol. 37, pp. 205–224, 2015. DOI: 10.17705/1CAIS.03709.

[269] Jan vom Brocke, Robert Winter, Alan Hevner, and Alexander Maedche, "Special Issue Editorial –Accumulation and Evolution of Design Knowledge in Design Science Research: A Journey Through Time and Space," *Journal of the Association for Information Systems*, vol. 21, no. 3, pp. 520–544, 2020. DOI: 10.17705/1jais.00611.

[270] Jóakim von Kistowski, Jeremy A. Arnold, Karl Huppler, Klaus-Dieter Lange, John L. Henning, and Paul Cao, "How to Build a Benchmark," in *Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering*, 2015, pp. 333–336. DOI: 10.1145/2668930.2688819.

[271] Franziska von Scherenberg, Malte Hellmeier, and Boris Otto, "Data Sovereignty in Information Systems," *Electronic Markets*, vol. 34, art. no. 15, 11 pages, 2024. DOI: 10.1007/s12525-024-00693-4.

[272] Johannes Vrana and Ripudaman Singh, "Digitization, Digitalization, and Digital Transformation," in *Handbook of Nondestructive Evaluation 4.0*, Norbert Meyendorf, Nathan Ida, Ripudaman Singh, and Johannes Vrana, Eds., Cham, Switzerland: Springer International Publishing, 2022, pp. 107–123. DOI: 10.1007/978-3-030-73206-6_39.

[273] W3Techs. "Usage of character encodings broken down by ranking." 2026. Available at: https://w3techs.com/technologies/history_overview/character_encoding (accessed on Feb. 10, 2026).

[274] Wenbo Wan, Jun Wang, Yunming Zhang, Jing Li, Hui Yu, and Jiande Sun, "A comprehensive survey on robust image watermarking," *Neurocomputing*, vol. 488, pp. 226–247, 2022. DOI: 10.1016/j.neucom.2022.02.083.

[275] Zhi-Hui Wang, Chin-Chen Chang, Chia-Chen Lin, and Ming-Chu Li, "A reversible information hiding scheme using left–right and up–down chinese character representation," *Journal of Systems and Software*, vol. 82, no. 8, pp. 1362–1369, 2009. DOI: 10.1016/j.jss.2009.04.045.

[276] Jane Webster and Richard T. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly*, vol. 26, no. 2, pp. xiii–xxiii, 2002. Available at: https://www.jstor.org/stable/4132319.

[277] Steffen Wendzel, Luca Caviglione, Wojciech Mazurczyk, Aleksandra Mileva, Jana Dittmann, Christian Krätzer, Kevin Lamshöft, Claus Vielhauer, Laura Hartmann, Jörg Keller, Tom Neubert, and Sebastian Zillien, "A Generic Taxonomy for Steganography Methods," *ACM Computing Surveys*, vol. 57, no. 9, pp. 1–37, 2025. DOI: 10.1145/3729165.

[278] WHATWG. "Encoding, Living Standard." 2025. Available at: https://encoding.spec.whatwg.org/ (accessed on Feb. 10, 2026).

[279] Roel J. Wieringa, *Design Science Methodology for Information Systems and Software Engineering*. Berlin, Heidelberg, Germany: Springer Berlin Heidelberg, 2014. DOI: 10.1007/978-3-662-43839-8.

[280] Wikimedia Foundation, *Wikimedia Downloads*, version 20231101.en, Hugging Face, 2023. Available at: https://huggingface.co/datasets/wikimedia/wikipedia (accessed on Feb. 10, 2026).

[281] William E. Winkler, "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage," 1990, 8 pages. Available at: https://eric.ed.gov/?id=ED325505.

[282] Kahim Wong, Jicheng Zhou, Kemou Li, Yain-Whar Si, Xiaowei Wu, and Jiantao Zhou, "FontGuard: A Robust Font Watermarking Approach Leveraging Deep Font Knowledge," *IEEE Transactions on Multimedia*, pp. 1–15, 2025. DOI: 10.1109/TMM.2025.3604908.

[283] Andrew Keane Woods, "Litigating data sovereignty," *The Yale Law Journal*, vol. 128, no. 2, pp. 328–406, 2018. Available at: http://hdl.handle.net/20.500.13051/10358.

[284] Min Wu and Bede Liu, *Multimedia Data Hiding*. New York, NY, USA: Springer New York, 2003. DOI: 10.1007/978-0-387-21754-3.

[285] Zhenyu Xu, Ruoyu Xu, and Victor S. Sheng, "Beyond Binary Classification: Customizable Text Watermark on Large Language Models," in *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024, pp. 1–8. DOI: 10.1109/IJCNN60899.2024.10650062.

[286] Shuguang Yuan, Jing Yu, Ke Yang, Yuyang Wang, Zhaochen Li, Jiabao Qiu, Tengfei Yang, and Chi Chen, "A study of watermarking techniques for data publishing," *Multimedia Tools and Applications*, 2025. DOI: 10.1007/s11042-025-20886-z.

[287] Xinmin Zhou, Sichun Wang, Weidong Zhao, and Rui Peng, "A semi-fragile watermarking scheme for content authentication of chinese text documents," in *2009 2nd IEEE International Conference on Computer Science and Information Technology*, 2009, pp. 439–443. DOI: 10.1109/ICCSIT.2009.5234911.

[288] Florian Zimmer, Malte Hellmeier, Motoki Nakamura, and Tobias Urbanek, "A Fragile Watermarking Technique for Integrity Authentication of CSV-Files Using Invisible Line-Ending Control Characters," in *Proceedings of the 22nd International Conference on Security and Cryptography*, 2025, pp. 455–466. DOI: 10.5220/0013559600003979.

[289] Elena Zinovieva, "Evolution of the Concept "Territorial Sovereignty" in the Digital Age," in *Proceedings of Topical Issues in International Political Geography (TIPG 2022)*, ser. Springer Geography, Radomir Bolgov, Vadim Atnashev, Yury Gladkiy, Art Leete, Alexey Tsyb, Sergey Pogodin, and Andrei Znamenski, Eds., Cham, Switzerland: Springer Nature Switzerland, 2024, pp. 187–195. DOI: 10.1007/978-3-031-50407-5_15.

[290] Johannes Zrenner, Frederik Oliver Möller, Christian Jung, Andreas Eitel, and Boris Otto, "Usage control architecture options for data sovereignty in business ecosystems," *Journal of Enterprise Information Management*, vol. 32, no. 3, pp. 477–495, 2019. DOI: 10.1108/JEIM-03-2018-0058.

Appendix A

Paper

Paper I

Table A.1 Metadata Overview of Paper I

Title	A Delimitation of Data Sovereignty from Digital and Technological Sovereignty
Authors	Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Franziska von Scherenberg *Fraunhofer ISST, Dortmund, Germany*
Publication Year	2023
Publication Type	Conference
Conference Name	31st European Conference on Information Systems (ECIS)
Conference Location	Kristiansand, Norway
Conference Date	11. June 2023 - 16. June 2023
Publisher / Database	AIS
DOI / Link	https://aisel.aisnet.org/ecis2023_rp/306
Status	Published
Ranking	VHB: A (2024 Rating) ICORE: - (2026 Rating) ERA: A (2010 Rating)
Comment	-

A DELIMITATION OF DATA SOVEREIGNTY FROM DIGITAL AND TECHNOLOGICAL SOVEREIGNTY

Research Paper

Malte Hellmeier, Fraunhofer ISST, Dortmund, Germany,
malte.hellmeier@isst.fraunhofer.de

Franziska von Scherenberg, Fraunhofer ISST, Dortmund, Germany,
franziska.von.scherenberg@isst.fraunhofer.de

Abstract

Digital technology significantly impacts our everyday social lives and how we conduct business. This development results in abundant new data generated by people and organizations. Subsequently, future technological instruments must ensure data sovereignty that empowers individuals to maintain control over their data. However, data sovereignty is still blurry and conceptually overlaps with similar terminologies, such as digital and technological sovereignty. From an Information Systems (IS) point of view, delimiting data sovereignty from digital and technological sovereignty is crucial, creating a uniform understanding, especially for data ecosystems. Our study contributes to sharpening data sovereignty with a systematic literature review of 81 articles. It concludes that data sovereignty mainly drives IS activities by protecting data assets on individual and organizational levels. In contrast, digital sovereignty is shaped by digital expertise and interoperability, while technological sovereignty is the broadest concept with regulations and relations on an international level.

Keywords: Data Sovereignty, Digital Sovereignty, Technological Sovereignty, Data Ecosystems, Literature Review

1 Introduction

Over the past years, the concept of sovereignty in the digital realm has increasingly gained attention in the international discourse due to data protection and challenges in climate change. Even 28 % of small and medium companies in the UK confirm that data sovereignty will drive their future decisions over data handling and storage (4D Data Centres, 2018). In Information Systems (IS) research, it is often used in the context of data ecosystems to build software systems and architectures that guarantee control over data. Due to increasing efforts toward digitalization that shape our everyday lives, different actors, amongst them individuals and organizations, recognize the value of data and want to keep control over it to prevent unintended usage on distribution. The importance of using data and technologies in a sovereign way to guarantee a shift into a sustainable society has been recognized (Caravella, Costantini, and Crespi, 2021) and underlined: "With extensive global digitalisation in all areas of society, our data sovereignty becomes a core aspect to ensure economic growth and social justice in Europe and to manage climate change" (AIT, 2022, p. 1). However, companies must face challenges through collaboration and the sharing of data. These companies understand that data is valuable, and sharing it is necessary not only to stay competitive, optimize internal business processes, and create new business opportunities but also to face challenges that single organizations cannot solve independently (Jarke, Otto, and Ram, 2019). Consequently, these developments show the need for more data sovereignty in research and practice.

Researchers have studied data sharing and sovereignty concepts, including their application in systems and organizations. However, the scientific discourse reflects the inaccurate delimitation of data sovereignty from the commonly used notions of digital and technological sovereignty (Micheli et al., 2020). Despite differences in the meaning of data, digital and technological sovereignty exist; past research often uses the concepts as if they were interchangeable or in a wrong way, like introducing data sovereignty with the concepts belonging to digital sovereignty (Lian, 2021). Due to several simultaneous research activities in the last years, IS literature misses an analysis of the near past on how the terms relate and influence each other. Researchers have already pointed to the necessity of an analytic differentiation, identifying and motivating the problem in the context of sovereignty (Couture and Toupin, 2019) to ensure uniform research usage and establish their accurate application in practice.
This study aims to guide future studies in delimitating these terminologies. It reviews the recently strongly rising research field of the quantitatively three most used sovereignty terms in IS research. Therefore, it aims to answer the following key Research Question (RQ) with the help of a literature review:
How can data sovereignty in Information Systems be delimited from digital and technological sovereignty?
Our analysis structures as follows, including concrete contributions to answer the RQ:

(i) Showing the developments of sovereignty in the analog and digital realm and identifying the research gap by reviewing related work (sections two and three).

(ii) Explaining the applied, systematic literature review method with its conditions like search strategy and the conducted procedure to create a final article collection (section four and appendix).

(iii) Visualizing indicators of the relevant literature distribution, showing detailed information of data, digital and technological sovereignty, including their similarities and differences (section five).

(iv) Guiding future research by showing limitations with a discussion on theoretical and practical implications and a conclusion of all findings (sections six and seven).

2 Background

Before diving deeper into the concrete form of different expressions, it is essential to show the foundation of the sovereignty term with its history and meaning in the analog realm.
Bodin (1577) introduced the first definition of sovereignty in the analog sphere that was continuously adapted over time until now (Adonis, 2019). This lead to different characteristics and focus points of sovereignty itself, while it "is generally defined as the supreme authority over a political entity (a polity)" (Couture and Toupin, 2019, p. 2308). In the last decades of the 20th century, the relation of sovereignty with the digital emerged (Grant, 1983). Different notions and conceptualizations have arisen from there, like the dominating concepts of data, digital, and technological sovereignty Not only IS coined these concepts but also related domains from the analog realm formed them over time. That is why the three concepts require a specialized review and a comprehensive analysis to ensure their relevance in a specific domain, in this case: IS research.

3 Related Work

Over time, several publications looked at other sovereignty aspects related to different domains. This chapter presents the most relevant articles on IS-grounded sovereignty terms and their related concepts. It further describes their delimitation to this work.
Starting with the review from Hummel et al. (2021), they analyze the notions of data, digital, and cyber sovereignty, focusing on context, values, and agents of the different terms. The authors selected the articles based on a systematic review and evaluated them, focusing on data sovereignty. In their research, they did not consider further practical implications. Similarly, Pedreira, Barros, and Pinto (2021) conducted a literature review on data, digital, and cyber sovereignty. They concentrate on vulnerabilities and outbreaks in industry scenarios. Whereas Hummel et al. (2021) emphasize the concept of data sovereignty, Pedreira,

Barros, and Pinto (2021) focus on cybersecurity. However, both reviews do not investigate the concept of technological sovereignty, even if the European Commission has positioned it as a political objective (ASD, 2020) and declared it relevant for future technical and non-technical aspects (Maurer et al., 2015). Besides this, Chapdelaine and McLeod Rogers (2021) refer to data and digital sovereignty without researching technological sovereignty. The authors look at technological sovereignty from a legal point of view, especially for media platforms and individuals. Other juridical studies such as Kushwaha, Roguski, and Watson (2020) elaborate on data sovereignty, touching upon the notions of digital and technological sovereignty concerning laws such as the US CLOUD Act and other regulations in different regions such as the UK, Germany, and Poland. Both publications come from the legal field instead of pure IS research. On a national level, Mawere and van Stam (2020) spotlight data sovereignty issues in Africa, especially Zimbabwe, based on the health system to guide local government. Besides information about the concept, the authors also refer to technological sovereignty without referring to digital sovereignty.

In contrast, Asswad and Marx Gómez (2021) discuss the concept of data ownership from an IS point of view. They describe the role of data ownership and point to its advantages and problems, such as the structure of the Internet of Things (IoT) data or missing regulations. Finally, a literature analysis strengthens the results, and the authors use the concept of data ownership relating to data sovereignty in the technical realm.

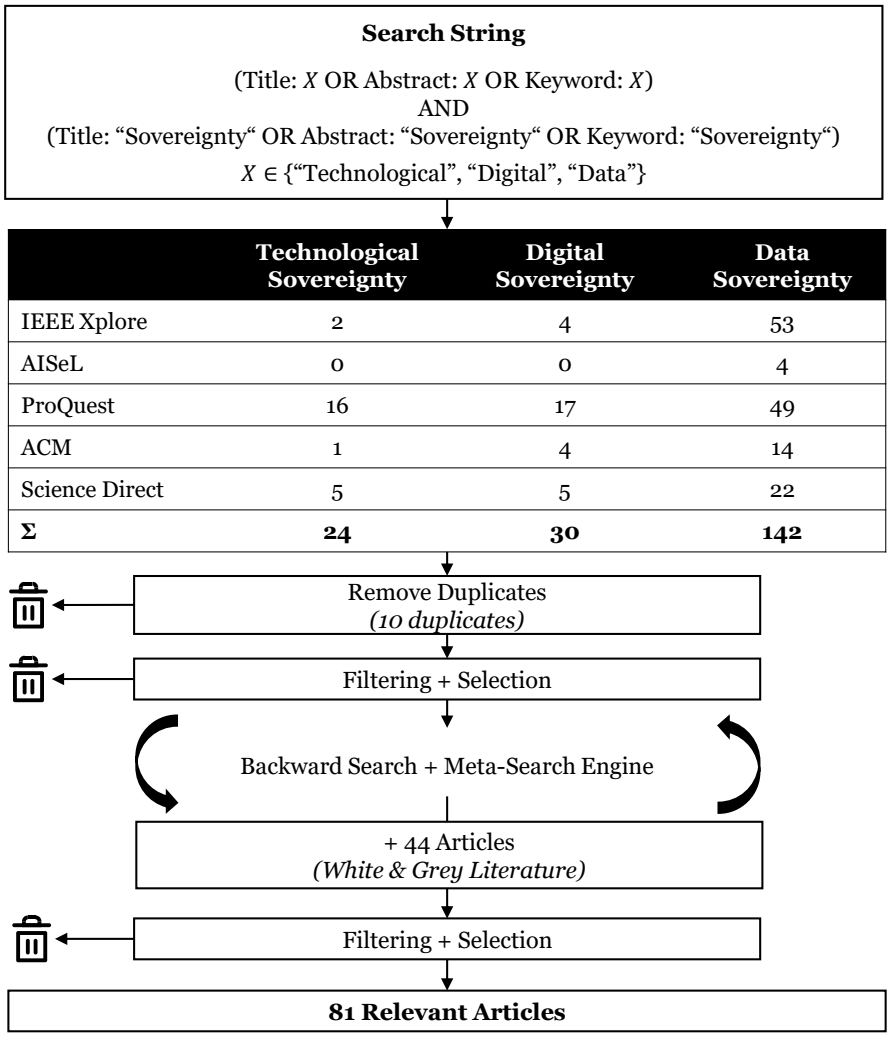
Considering related literature, it becomes clear that from an IS point of view, a detailed analysis of the delimitation of data sovereignty from digital and technological sovereignty does not yet exist. Since cyber sovereignty is also a frequently discussed topic (Hummel et al., 2021; Pedreira, Barros, and Pinto, 2021), it will not be considered further here. Instead, this study analyzes from an IS point of view, and "[t]he fact that the majority of academic literature focuses on the legal implications of sovereignty in cyberspace indicates that the issue of cyber sovereignty is most often framed and understood as a matter of International Law" (Baezner and Robin, 2018, p. 5).

However, literature shows that researchers have made the first steps in describing the terms (Couture and Toupin, 2019). They point to the necessity of the analysis to identify and motivate the problem: "One question that remains open relates to the relationship between different terms" (Couture and Toupin, 2019, p. 2318). In their research, the authors approach a first differentiation of the terms by creating hypotheses on the delimitation of data, digital, and technological sovereignty. They mention that future research must investigate these concepts (Couture and Toupin, 2019). These arguments align with the publication from Mawere and van Stam (2020). Their research cannot constitute a clear differentiation between data and technological sovereignty. In addition, Mawere and van Stam (2020) propose comparing the terms by bringing together different viewpoints to identify their relevance. Therefore, this research extends current work to answer the RQ mentioned above. It closes the gap of the current unspecific delimitation because existing ideas are "[. . .] just hypotheses that would need to be further explored in future works" (Couture and Toupin, 2019, p. 2318). Even though other related concepts are frequently discussed, our quantitative analysis focuses on the IS domains instead of concepts from indigenous people like indigenous data sovereignty (Taylor and Kukutai, 2016) or international law like cyber sovereignty (Baezner and Robin, 2018).

4 Research Design

This study conducts a systematic literature review, analyzing the three most used terms in IS research data sovereignty, digital sovereignty, and technological sovereignty in detail. Since sovereignty aspects are not only relevant to science but also in industry and politics, the present research is based on the recommendations from Webster and Watson (2002) and the guidelines from vom Brocke et al. (2015), extended by a Multivocal Literature Review (MLR). Besides classical reviews on scientific publications (white literature), an MLR increases the scope by including political speeches and technical reports from practitioners (grey literature) to focus on real-world problems due to the combination of academic research and practice (Garousi, Felderer, and Mäntylä, 2019). This study refers to the rationale of Benzies et al.

(2006) to include grey literature, which is, among others, the low quality of evidence. Further, the context for implementing the intervention (Benzies et al., 2006) is vital for sovereignty aspects and strengthens the consideration of publications from practice. Especially in topics around sovereignty whose contexts have policy and industry focus, the inclusion of grey literature provides added value (Benzies et al., 2006). This literature review executes the searches to identify white literature using a keyword-based approach, applied to varying sources to collect journal articles, conference proceedings, and book chapters. Various databases have different key areas and are sometimes limited to specific publishers. This study selected the following five: IEEE Xplore for computer science and technical publications, AISeL and ACM for an IS focus, and ProQuest and Science Direct to include other adjacent domains. In every database, the search strings *technological sovereignty*, *digital sovereignty*, and *data sovereignty* are used, resulting in 15 searches. Every search term is split into its two components (e.g., "data" and "sovereignty") connected with an AND operator to identify literature that divides the words in a text part without using, e.g., "data sovereignty" in one. Here, the title, abstract, and keywords were searched, as Bandara et al. (2015) recommended, without any publication date restriction. The top of Figure 1 shows the resulting numbers based on searches made in April 2022.



	Technological Sovereignty	Digital Sovereignty	Data Sovereignty
IEEE Xplore	2	4	53
AISeL	0	0	4
ProQuest	16	17	49
ACM	1	4	14
Science Direct	5	5	22
Σ	**24**	**30**	**142**

Figure 1. Literature Search Process.

Ten duplicate entries were found in the downstream intermediate step, which led to a reduction from 196 to 186 articles. Here, some entries are listed in two of the five databases, and others are shown two times in

the ProQuest results because this service summarizes 18 databases and can lead to internal duplicates. Due to the precise selection of the databases with mostly disjoint subsets of results, a high number of duplicates is prevented in advance. This first literature set is subject to a full-text scan to check its relevance based on the RQ as the first selection criteria for relevant articles. For this purpose, the texts were searched for the three terms *data sovereignty*, *digital sovereignty*, and *technological sovereignty* and checked if they give a concrete definition, discussion, implementation, or explanation as an inclusion/exclusion criteria. Thus, all articles that do not give new insights about at least one of the three terms, focus on other topics or only mention one term without further details are filtered out. Concerning inter-coder reliability, inconclusive cases are discussed, evaluated, and documented with an exclusion reason in a protocol by the author team to create a final decision and prevent subjectivity.

As described in the recommendations from Webster and Watson (2002), this study applies a backward search of the key literature after the previous filtering step. Here, the authors checked and analyzed all references on the relevant text parts identified in the previous step, whether white or grey literature. At this stage, Kuhrmann, Fernández, and Daneva (2017) suggest extending the primary search with results from a meta-search engine such as Google. This approach supports the inclusion of additional academic publications not listed in the five databases and additional grey literature. It is in line with the MLR guideline seven of Garousi, Felderer, and Mäntylä (2019) that focuses on, e.g., additional web searches. Since many results exist, only 1st tier grey literature with a more known outlet control and expertise is analyzed. It excludes 2nd and 3rd-tier grey literature like Q/A sites, emails, or tweets (Garousi, Felderer, and Mäntylä, 2019). Finally, the authors stopped the process at the theoretical saturation, based on guideline eight of an MLR (Garousi, Felderer, and Mäntylä, 2019). At this stage, this research identifies additional 44 articles for review.

The current white and grey literature collection is subject to the same filtering and selection step described above. The resulting set consists of 81 articles, 52 extracted from the databases and 29 found by the backward search or with the help of the meta-search engine. Besides its content, this study analyzes the collection by tagging every article with metadata such as title, author, publication date, category, source, country of the main authors' institution, and similar facts. The overall process is summarized in Figure 1 to avoid existing concealment problems in IS literature reviews and to ensure replicability (vom Brocke et al., 2009). The final article overview is shown in the Appendix in Table 1 to ensure complete transparency. Upon request, the full list, including all sources and analyzed details like country distribution, is shared to ensure complete repeatability.

5 Results

The following section presents the findings from the systematic literature review with a downstream thematic classification. For this purpose, the authors thoroughly analyzed all relevant articles by extracting the terms' information concerning the RQ.

After a descriptive overview of the literature set, the authors describe data, digital, and technological sovereignty in detail and show their origin and current usages. Based on the RQ, a summary of every term and their relation to data sovereignty can be found in the textbox at the end of every subsection. They are not intended to create a new definition but to conclude the findings from the different descriptions of the articles from the literature review. Subsequently, for the derivation of future action recommendations, all information is analyzed concerning IS research to show the differences between data sovereignty and the other terms depicted in Figure 5.

5.1 Descriptive Findings

Research on sovereignty in the digital realm first occurred in the last decades of the 20th century (Grant, 1983). However, most data, digital, and technological sovereignty publications stem from the last years.

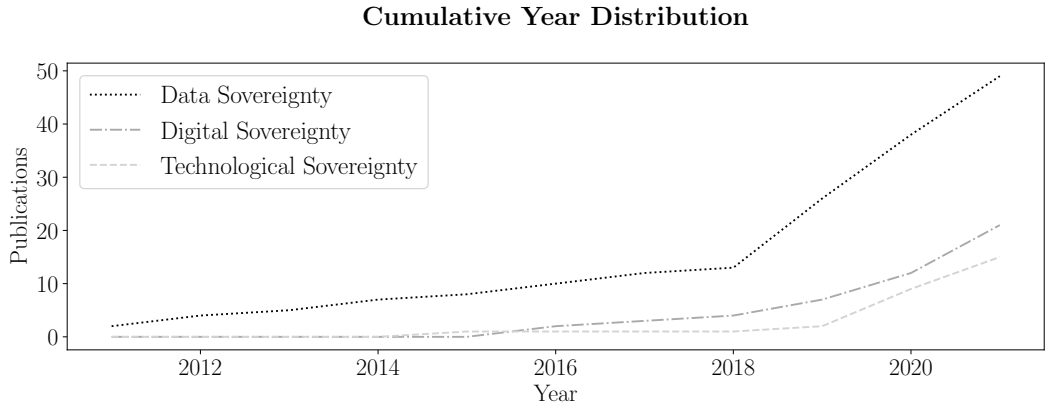


Figure 2. Cumulative Publication Time Distribution of the Past Ten Years on the Relevant Literature Set.

Figure 2 shows the cumulative number of publications for every term in the last ten years based on the relevant literature set from the literature review. Articles are counted in the statistic if they cover one of the three sovereignty aspects. Around 11 % of the scanned literature refers to publications that do not exclusively focus on data, digital, or technological sovereignty but observe more than one term. An example is the research of Pedreira, Barros, and Pinto (2021) that mentions digital and data sovereignty. In this case, the authors counted this paper twice, one time for every category. The graph shows that since 2018, data sovereignty has been the most discussed term, while digital and technological sovereignty are also increasingly gaining attention.

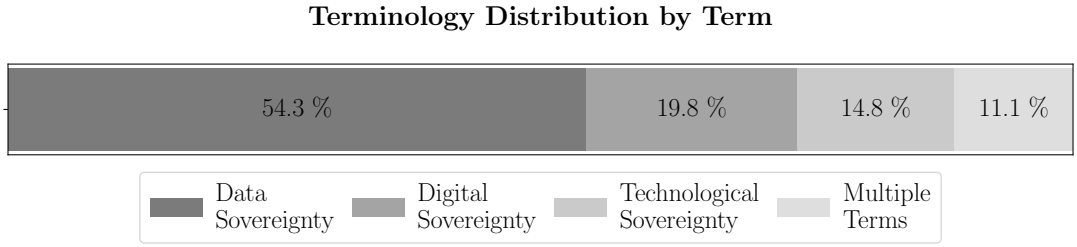


Figure 3. Terminology Distribution on the Relevant Literature Set.

More concretely, 54.3 % of the 81 articles specialize in data sovereignty, while 19.8 % focus on digital sovereignty and 14.8 % cover technological sovereignty, summarized in Figure 3. The rapid simultaneous increase in research further motivates a structured review of past results to guide future research and prevent misunderstandings in the delimitation of the terms.

During the literature search and selection, every article of the final literature set is assigned to a country based on the location of the main authors' institutions to create a location-based distribution of the publications shown in Figure 4. Here, the authors aggregated the countries to their associated continents and rounded the percentage values to whole numbers. While there are no publications from South America and, logically, Antarctica, only one publication from Africa (Mawere and van Stam (2020)) and one from Australia / Oceania (Vaile (2014)) are present. Therefore, the research is mainly conducted in Europe, followed by North American and Asian countries based on quantitative measurement. Most European papers are published by German researchers, followed by publications from France (six papers) and the United Kingdom (six papers). The rest is split up between different European countries like the Netherlands, Belgium, and Italy (three papers each), Finland, and Spain (two papers each), and others with one publication each.

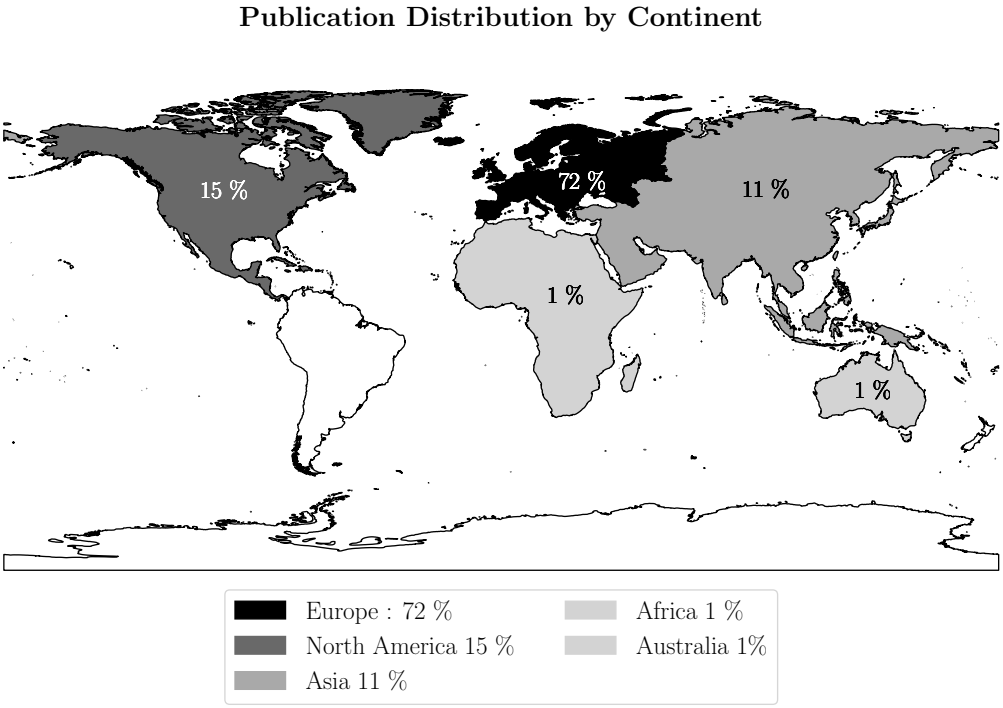


Figure 4. Publication Continent Distribution on the Relevant Literature Set.

5.2 Concepts of Sovereignty

Different conceptualizations of sovereignty terminology in the digital have appeared in recent decades. We focus on the most used notions in IS research data, digital and technological sovereignty. Cyber sovereignty (Couture and Toupin, 2019; Pohle and Thiel, 2020), in contrast to digital sovereignty, comprises the physical dimension of sovereignty in cyberspace (Baezner and Robin, 2018) and constitutes only one aspect of digital sovereignty providing the motivations for the researchers to exclude it from this study' delimitation. The results are formed under the consideration of mixed terms such as digital data sovereignty (Aydin and Bensghir, 2019), personal data sovereignty (Micheli et al., 2020), or urban technological sovereignty (Vadiati, 2022). This study's scope does not include the delimitation of little-used terms such as AI or 5G sovereignty (Floridi, 2020). Besides, the present research does not discuss terms from other domains without a full technical or IS focus, such as indigenous (Taylor and Kukutai, 2016), food, body, or state sovereignty (Couture and Toupin, 2019).

5.2.1 Data Sovereignty

Data sovereignty is the term most used based on the findings of the present literature review (Figure 2). Fifty-one publications of the relevant literature basket form the following results. However, the term lacks a unique definition, and research shows that its usage does not happen consistently (Martens and Zscheischler, 2022; Micheli et al., 2020). It was initially taken up in 2001, with the first ideas mentioned in the US Patriot Act (Hallinan, 2022). When using the term data sovereignty, researchers often refer to what is understood by the concept of (data) self-determination (Banse, 2021; Hummel et al., 2018; Jarke, Otto, and Ram, 2019). The scope of the term is broad and includes different requirements around data, such as confidentiality and data integrity, with the protection against unwanted modifications or data availability (Nugraha, Kautsarina, and Sastrosubroto, 2015). Therefore, it is relevant for individuals and organizations (Jarke, Otto, and Ram, 2019).

For further understanding, the terminology has to be analyzed from an IS and Software Engineering (SE) point of view. In a classical data transfer, a person, organization, or system (referred to as a data provider) shares data with a third party (referred to as a data consumer) and wants to keep control over it (Otto et al., 2019; Zrenner et al., 2019). One example is medical data that belong to a patient (data provider) who can decide and allow access to hospitals or doctors (data consumer) (Plateaux et al., 2013). Since the data have to be secured, data sovereignty means that a data provider can decide and keep control over his or her data. Unfortunately, practice often solves it with written contracts or oral agreements, which regularly fail and can be prevented with technical solutions (Zieglmeier and Pretschner, 2021).

So far, different technical solutions and architectures have arisen that aim to take up data sovereignty ideas. One attempt is the International Data Spaces (IDS) initiative, building up data spaces and data ecosystems based on a reference architecture model to enable sovereign data transfers with the help of different components such as IDS Connectors (Otto et al., 2019). Other ideas are trustworthy architectures (Zieglmeier and Pretschner, 2021), connector-based communication schemes for IoT devices (Qarawlus et al., 2021), blockchain integration (Hong and Kim, 2020), or combinations with other standards such as the Industrie 4.0 Asset Administration Shell (AAS) (Redeker et al., 2020).

Developing systems and architectures or building concepts for technical solutions in IS or SE must meet personal or organizational requirements and comply with the law. In the context of data sovereignty, business and countries formulate requirements in the form of regulations to protect individuals, as seen with the EU General Data Protection Regulation (GDPR), approved in 2016. Such regulations focus on national data sovereignty, a term taken up by Irion (2012), building the transition to digital and technological sovereignty. Therefore, developers, system architects, and researchers must look at data sovereignty and understand other terms and requirements to create compliant data sovereign solutions, even if they are mainly coined by other domains such as economic or political science.

> Concepts of self-determination and the capability of a data provider to keep control over their own data assets form the term **data sovereignty**. Among other things, it is used in the IS and SE domains to create technical solutions to protect individual and company data and highly depends on economic, political, and legal aspects.

5.2.2 Digital Sovereignty

Since the meaning and usage of *digital sovereignty* differ from data sovereignty, the scientific discourse must avoid using both as synonyms. Researchers discuss the concept and its history in detail, shown by the 24 articles of the final literature set. The first ideas go back to the late 1990s. Key events like the Patriot Act in 2001, the Snowden disclosure in 2013, Brexit in 2016, and the COVID pandemic in 2020 formed the term and changed its definitions (Hallinan, 2022). Similar to the other discussed concepts, digital sovereignty has no unique definition because of its changes over time and its dependence on context and stakeholders (Hallinan, 2022).

Due to the formative sociopolitical events during the last years, different nations recognized problems in protecting their citizens in the digital realm that triggered discussions and actions around a nation's digital sovereignty (Pohle and Thiel, 2020). Derived measures include the control and influence of the digital world formed by hardware, software, and infrastructure (Floridi, 2020). However, problems arise because nations can enact local laws and regulations, but cyberspace goes beyond it and covers the whole world. While, for example, governments can regulate the construction of internet cables across borders, the governance of data that flows through it is more complex. Therefore, research coins the attempt to create generalized regulations based on territory and borders regarding data flows as digital sovereignty and accompanying terms such as cyberspace or territorial sovereignty (Cattaruzza et al., 2016; Hallinan, 2022). The study deliberately chooses to analyze digital sovereignty because apart from territorial aspects of cyber sovereignty, it moreover comprises aspects of the digital transformation from a political point of

view (Pohle and Thiel, 2020). Former German Chancellor Angela Merkel further clarified in her speech in 2019: "[...] digital sovereignty does not mean protectionism [...] rather, it describes the ability both of individuals and society to shape the digital transformation in a self-determined way" (Merkel, 2019). Thus, regulations act as an enabler for digital sovereignty because they shape the concept's perception and development.

Besides the political and territorial focus that relates more to control than authority (Cattaruzza et al., 2016), digital sovereignty significantly impacts enterprises and individuals. Businesses and states influence the notion of digital sovereignty in sometimes contrasting ways. On the one hand, companies build, design, and actively participate in the digital realm. In contrast, states use control mechanisms and deploy regulations, sometimes decelerating innovative processes and societal progress (Floridi, 2020). On an individual level, digital sovereignty relates to the concept of interoperability. Another significant element is the freedom to select and use digital assets without being bound to a specific technology (Kagermann, Streibich, and Suder, 2021).

Digital sovereignty focuses on actions, expertise, and control mechanisms in the digital world, while publications often concentrate on territorial borders in the political and economic realm. It helps protect businesses and individuals from selecting and using technology and digital assets in an interoperable way. Regarding data sovereignty, digital regulations, and economic actions are implemented in software systems and directly influence data handling, covered in IS research.

5.2.3 Technological Sovereignty

Regarding the broadest concept, research refers to tech-, technology- or *technological sovereignty* in 17 publications of our gathered literature basket. Its origins go back to the 20th century, with one of the first technological sovereignty definitions described as "the capability and the freedom to select, to generate or acquire and to apply, build upon and exploit commercial technology needed for industrial innovation" (Grant, 1983, p. 240). In these early days, the term referred to the ability of different states to develop, use and produce their technologies and innovations (Couture and Toupin, 2019). The main element that characterizes technological sovereignty is the capability to build techniques and other products, including competencies and licenses. Besides, Grant (1983) describes technological sovereignty as a guarantor of freedom due to reduced dependencies on other states. This assumption is grounded in the fact that licenses and regulations are needed to allow and enable industrial and productive innovations (Grant, 1983).

On closer inspection, these studies assume that states must produce and control all resources and invest in their research as a necessary derivation for building products and infrastructures without external dependencies. However, this feasibility must be questioned because most states can only enable local production and invention with import and export relationships with other countries. For example, rare-earth elements are primarily mined in Asia but are relevant in Europe to build electrical products. European dependencies from the United States in various domains extend a strong international link (Crespi et al., 2021). Therefore, researchers extended the concept of technological sovereignty, which does not mean autarky or complete technological independence (Edler et al., 2021). Cooperations and communication are necessary to build trade relationships and create innovations (March and Schieferdecker, 2021). These aspects can be found in the definition from Edler et al. (2020): "We define technology sovereignty as the ability of a state or a federation of states to provide the technologies it deems critical for welfare, competitiveness, and its ability to act and to be able to develop these or source them from other economic areas without one-sided structural dependency" (Edler et al., 2020, p. 8).

Due to its increasingly linked national and international relations, technological sovereignty became relevant in politics. Since 2019 it has been included in political debates (March and Schieferdecker, 2021) and integrated into political strategies. In her speech in 2020, European Commission's president von der Leyen described it as Europe's capability to make choices based on its values and rules (von der Leyen,

2020). Accordingly, the importance of technological sovereignty can be underlined by its position in the digital strategy of the EU (European Commission, 2020). Concerning the conflict of complete autarky and international trade, technological sovereignty, in the understanding of the European Commission, is a baseline of self-supply extended by solid import and export relationships for resistance against crises such as pandemics, wars or others (ASD, 2020). However, the topic is significant in Europe and other states like Canada, Australia (Couture and Toupin, 2019), Brazil, and China (Maurer et al., 2015). March and Schieferdecker (2021) further summarize these political aspects: "Technological sovereignty is the ability of a polity to self-determinedly shape the development and use of technologies and technology-based innovations which impact its political and economic sovereignty" (March and Schieferdecker, 2021, p. 9). Technological sovereignty includes several trends like data storage location, new undersea cables, localized routing, and others (Maurer et al., 2015).

Technological sovereignty activities focus on a political level with a sometimes national but more international focus. Strategies and regulations influence how sovereign data systems are built. Strong country cooperation relationships support keeping control over resources and influence data sovereignty activities by reducing dependencies.

5.3 Sovereignty in IS Research

As described in the literature overview section, sovereignty has reached momentum in the last few years. The literature review reveals that data sovereignty proves assertiveness in IS research since it describes essential parts of developing technologies.

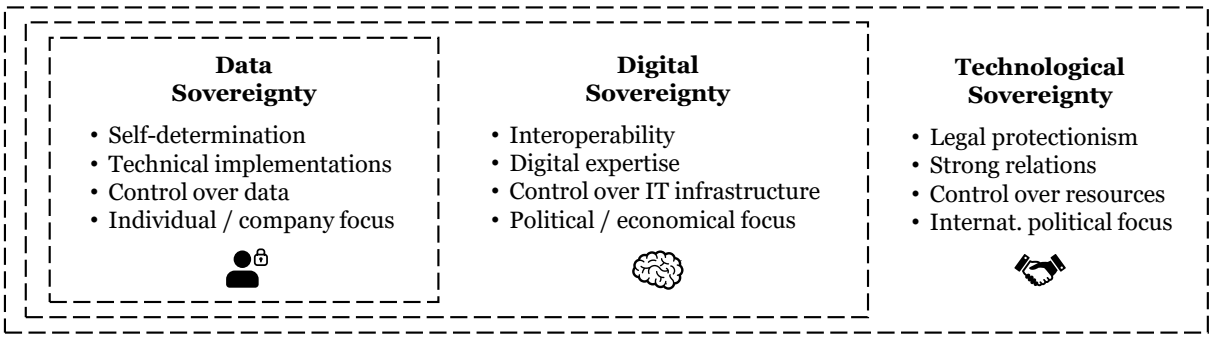


Figure 5. Delimitation of the Analyzed Sovereignty Terms.

A summary of the main characteristics related to IS in the style of Calzada (2021a) is presented in Figure 5. Here, digital sovereignty refers to the digital literacy of a population and the level of development of digitization concepts. Furthermore, it influences IS research through the strong orientation with political and economic focus and their request for more interoperability.

Instead, technological sovereignty, as the broadest concept, is essential in the political discourse, often on an international level, as shown by initiatives from the EU. This concept focuses on the technical usage and sometimes independence of resources from other nations and trade areas with strong relations. It is often associated with freedom and dependencies on political levels.

The review also shows that data sovereignty in IS research, the ability to control individual or organizational data assets, cannot be limited to one research domain. Therefore, observing adjacent fields such as digital or technological sovereignty from an IS point of view is crucial to fully understand how they form and influence the definition of data sovereignty. For example, economic activities and political regulations directly impact system development processes and the definition of data sovereignty. Examples are the recently introduced Data Act on a European level (Perarnaud and Fanni, 2022) or regulations from other

regions like the US CLOUD Act or the Personal Information Protection Law (PIPL) of China. Also, if regulations on an international or national level are discussed for technological or digital sovereignty, they control and shape how data sovereignty can be realized in software and algorithms.

6 Discussion

This IS research-based delimitation of data sovereignty from digital and technological sovereignty, as described in the RQ, serves as a necessity for future research. A mutual understanding of the terms' similarities, differences, and usage is crucial for further developments in the digital economy and SE projects, especially in all data-driven innovations in data spaces and data ecosystems. Since the results from the literature review set a first starting point, implications for theory and practice, as well as limitations and points of contact for future research, have to be discussed.

6.1 Theoretical Implications

Prior research has shown differences between data, digital and cyber sovereignty (Hummel et al., 2021; Pedreira, Barros, and Pinto, 2021). However, the importance of technological sovereignty has increased. It has overtaken cyber sovereignty, as shown by current research activities in the academic domain and the positioning on political levels like the European Commission (ASD, 2020). Especially in IS and SE research, data sovereignty, unlike others, finds a concrete application without a clear delimitation. Since previous work started with assumptions, they motivate future research activities (Couture and Toupin, 2019) concerning a comparison (Mawere and van Stam, 2020).
This study works on the gap by systematically reviewing the past in the academic and practical domains. The results show the theoretical importance, especially since 2018 (Figure 2). Quantitatively, a clear overweighting of data sovereignty with a strong European focus is shown (Figure 4). It strengthens, on the one hand, the need for a clear delimitation, but on the other hand, it poses the question of why some regions are more underrepresented than others. The final thematic classification has implications for future IS activities and other related sovereignty domains, such as political, economic, and legal perspectives.

6.2 Practical Implications

The results show that the different concepts focus on different domains, which leads to direct implications for practice. As stated above, the concepts have strong relationships and influence each other. A geopolitical decision for technological sovereignty acts on digital sovereignty and influences concrete software products and their development process in private and industrial contexts. Therefore, practical key players must understand the differences and evaluate the dependencies and impacts. It is further crucial that these domains work together to prevent isolated concepts and misunderstandings.
In extension to the sustainable aspects mentioned in the introduction, companies and countries have to work together to steer in the direction of carbon neutrality and to achieve the agreed targets in the Paris Agreement (Caravella, Costantini, and Crespi, 2021). Data ecosystems and current research projects such as the IDS or GAIA-X help with this challenge in a practical context by creating systems built on data sovereign principles that follow the strategies and regulations of digital and technological sovereignty. Building on the results of this work, delimitating the term helps to strengthen trust between different parties through the internalization of data-sharing principles needed to accelerate innovation to solve sustainable challenges.

6.3 Limitations and Future Research

Despite careful work and evaluation, this study and the results have some limitations and points of contact for future research, discussed in the following.

Firstly, the search term selection only focuses on the three concepts: data, digital and technological sovereignty. Research has shown several papers discussing similar ideas, such as data protection, ownership, and maintaining usage or access control. One example is shown in the paper from Gates and Slonim (2003), addressing privacy, control, and data aspects without mentioning any sovereignty term. In this case, research might relate to data sovereignty without using the term. Therefore, further research is essential to understand better the relationship between the concepts of data sovereignty and data protection, data ownership, or similar terms.
Secondly, Figure 4 depicts the original distribution of the literature based on the main authors' institution. Notably, a large part of the publications is rooted in Europe, with an apparent underrepresentation of Africa, Australia, and South America ($<= 1\%$). Future research needs to evaluate the reasons and implications of the distribution and possible effects on transnational activities, like impacts on software implementations used in different regions.
Thirdly, this research's results do not clearly show, that the motivation for more sovereignty often lies in achieving more sustainability. Since the implementation of data sovereignty can promote more data sharing, sustainability goals can be achieved because organizations share their data for a greater common good to achieve more climate protection (DSSC, 2023). However, this study's focus did not cover the link between sustainability goals and the implementation of one of the sovereignty terms in detail and therefore needed to be discussed in future research.
Lastly, we argued that technological and digital sovereignty influence data sovereignty in IS research, as shown by several examples like national or international regulations that protect individual or organizational data. It is still unsettled if data sovereign models or systems can be created, even if none of the other two concepts are implemented. Future research has to analyze each concept's realization and how this encourages or prevents others.

7 Conclusion

Our literature has pointed out the delimitation of data sovereignty from digital and technological sovereignty and aims to contribute to a better understanding and coordinated usage of the concepts. Concluding, it is recommendable for IS research not to use data, digital, and technological sovereignty as if they were interchangeable concepts – they do not mean the same. Instead, the results of the RQ can be summarized as follows:
The concept of data sovereignty is embedded in IS research and used in the context of control over data on an individual or organizational level. Instead, technological sovereignty is essential in the political discourse with mostly international targets. This concept focuses on the usage and reduced dependencies of resources from other nations and trade areas and the preservation of solid relations. Therefore, the target of digital sovereignty lies in the political and economic realm. It refers to the digital literacy of a population and the level of development of an organization's digital features with control over infrastructures and aspects of interoperability.
As shown in the discussion, the delimitation of data sovereignty from adjacent fields, such as digital sovereignty and technological sovereignty, is essential to give future research and software architecture development a fundamental groundwork for using the terminologies with unanswered questions. Moreover, as pointed out in the study, data sovereignty cannot only be viewed from an IS research perspective since it is interwoven with other domains that influence it and contribute to the developments in the digital economy.

Acknowledgments

This work was partially funded by the "Silicon Economy Logistics Ecosystem" project. The project "Silicon Economy Logistics Ecosystem" is funded by the German Federal Ministry of Transport and Digital Infrastructure.

Appendix

No.	Source (Author & Year)	*Data Sov.*	*Digital Sov.*	*Techno. Sov.*
1	Adonis (2019)		x	
2	ASD (2020)			x
3	Aydin and Bensghir (2019)	x		
4	Banse (2021)	x		
5	Bauer et al. (2019)	x		
6	Bendiek and Neyer (2020)		x	
7	Braud et al. (2021)		x	
8	Calzada (2021a)	x		
9	Calzada (2021b)			x
10	Caravella, Costantini, and Crespi (2021)			x
11	Cattaruzza et al. (2016)		x	
12	Chapdelaine and McLeod Rogers (2021)	x	x	
13	Chen et al. (2020)	x		
14	Christakis (2020)		x	
15	Corbett and Cochrane (2020)	x		
16	Couture and Toupin (2019)	x	x	x
17	Crespi et al. (2021)		x	x
18	Cuno et al. (2019)	x		
19	Dabrock (2020)	x		
20	Diesen (2021)			x
21	Edler et al. (2020)			x
22	Edler et al. (2021)			x
23	Esposito, Castiglione, and Choo (2016)	x		
24	Esposito et al. (2019)	x		
25	European Commission (2020)			x
26	Filippi and McCarthy (2012)	x		
27	Floridi (2020)		x	
28	Friedrichsen and Bisa (2016)		x	
29	German Ethics Council (2017)	x		
30	Grant (1983)			x
31	Gupta, Lanteigne, and Kingsley (2020)	x		
32	Hallinan (2022)		x	
33	Hartsch et al. (2021)	x		
34	Hong and Kim (2020)	x		
35	Hummel et al. (2018)	x		
36	Hummel et al. (2021)	x	x	
37	Irion (2012)	x		
38	Janardhanan and Mas-Machuca (2022)		x	x
39	Jarke, Otto, and Ram (2019)	x		
40	Kagermann, Streibich, and Suder (2021)		x	
41	Komaitis (2021)		x	
42	Kukkola (2018)		x	
43	Kushwaha, Roguski, and Watson (2020)	x	x	x
44	Lauf et al. (2021)	x		
45	Lian (2021)	x		
46	Litvinenko (2021)		x	
47	Lynch (2020)			x
48	Mannhardt et al. (2019)	x		
49	March and Schieferdecker (2021)			x
50	Mark (2019)	x		
51	Martens and Zscheischler (2022)	x		
52	Maurer et al. (2015)			x
53	Mawere and van Stam (2020)	x		x
54	Merkel (2019)		x	
55	Micheli et al. (2020)	x		
56	Mooy (2017)	x		
57	Munoz-Arcentales et al. (2019)	x		
58	Nagel and Lycklama (2021)	x		
59	Nast et al. (2020)	x		
60	Nugraha, Kautsarina, and Sastrosubroto (2015)	x		
61	Otto (2019)	x		
62	Otto and Burmann (2021)	x		
63	Pedreira, Barros, and Pinto (2021)	x	x	
64	Peterson, Gondree, and Beverly (2011)	x		
65	Plateaux et al. (2013)	x		
66	Pohle and Thiel (2020)		x	
67	Polatin-Reuben and Wright (2014)	x		
68	Posch (2017)		x	
69	Qarawlus et al. (2021)	x		
70	Redeker et al. (2020)	x		
71	Ruohonen (2021)		x	
72	Ruparelia (2016)	x		
73	Sarabia-Jacome et al. (2019)	x		
74	Schleicher et al. (2011)	x		
75	Singi et al. (2020)	x		
76	Tan, Chi, and Lam (2022)	x	x	
77	Taylor (2020)	x		
78	Vaile (2014)	x		
79	von der Leyen (2020)			x
80	Zieglmeier and Pretschner (2021)	x		
81	Zrenner et al. (2019)	x		
∑	*81 publications*	*51*	*24*	*17*

Table 1. Final Literature Set.

References

4D Data Centres (2018). *The state of the UK server room*. URL: https://cdn2.hubspot.net/hubfs/6750926/4D_Data_Centres_December2019/Pdf/4D_DC_UK_Server_Room_Whitepaper.pdf (visited on Mar. 15, 2023).

Adonis, A. A. (2019). "Critical Engagement on Digital Sovereignty in International Relations: Actor Transformation and Global Hierarchy." *Global: Jurnal Politik Internasional* 21 (2), 262–282. ISSN: 1411-5492. DOI: 10.7454/global.v21i2.412.

AIT (2022). *Data Sovereignty for a Sustainable, Competitive Soviety*. Ed. by AIT Austrian Institute of Technology. URL: https://www.ait.ac.at/news-events/single-view/detail/7373?cHash=4391e19a3b1a4ad72aecb5102418e5c3 (visited on Mar. 15, 2023).

ASD (2020). *Industry considerations on Technological Sovereignty*. Ed. by AeroSpace and Defence Industries Association of Europe. Brussels. URL: https://www.asd-europe.org/industry-considerations-on-technological-sovereignty-concept-paper (visited on Mar. 15, 2023).

Asswad, J. and J. Marx Gómez (2021). "Data Ownership: A Survey." *Information* 12 (11), 1–32. DOI: 10.3390/info12110465.

Aydin, A. and T. K. Bensghir (2019). "Digital Data Sovereignty: Towards a Conceptual Framework." In: *2019 1st International Informatics and Software Engineering Conference (UBMYK)*. IEEE, pp. 1–6. ISBN: 978-1-7281-3992-0. DOI: 10.1109/UBMYK48245.2019.8965469.

Baezner, M. and P. Robin (2018). *Cyber Sovereignty and Data Sovereignty*. DOI: 10.3929/ethz-b-000314613.

Bandara, W., E. Furtmueller, E. Gorbacheva, S. Miskon, and J. Beekhuyzen (2015). "Achieving Rigor in Literature Reviews: Insights from Qualitative Data Analysis and Tool-Support." *Communications of the Association for Information Systems* 37, 154–204. DOI: 10.17705/1CAIS.03708.

Banse, C. (2021). "Data Sovereignty in the Cloud - Wishful Thinking or Reality?" In: *Proceedings of the 2021 on Cloud Computing Security Workshop*. Ed. by Y. Zhang and M. van Dijk. New York, USA: ACM, pp. 153–154. ISBN: 9781450386531. DOI: 10.1145/3474123.3486792.

Bauer, J., R. Helmke, A. Bothe, and N. Aschenbruck (2019). "CAN't track us: Adaptable privacy for ISOBUS controller area networks." *Computer Standards & Interfaces* 66, 103344. ISSN: 0920-5489. DOI: 10.1016/j.csi.2019.04.003.

Bendiek, A. and J. Neyer (2020). "Europas digitale Souveränität. Bedingungen und Herausforderungen internationaler politischer Handlungsfähigkeit." In: *Demokratietheorie im Zeitalter der Frühdigitalisierung*. Ed. by M. Oswald and I. Borucki. Wiesbaden and Heidelberg: Springer VS, pp. 103–125. ISBN: 978-3-658-30996-1. DOI: 10.1007/978-3-658-30997-8_6.

Benzies, K. M., S. Premji, K. A. Hayden, and K. Serrett (2006). "State-of-the-evidence reviews: advantages and challenges of including grey literature." *Worldviews on evidence-based nursing* 3 (2), 55–61. ISSN: 1545-102X. DOI: 10.1111/j.1741-6787.2006.00051.x.

Bodin, J. (1577). *Les six liures de la republique*.

Braud, A., G. Fromentoux, B. Radier, and O. Le Grand (2021). "The Road to European Digital Sovereignty with Gaia-X and IDSA." *IEEE Network* 35 (2), 4–5. ISSN: 0890-8044. DOI: 10.1109/MNET.2021.9387709.

Calzada, I. (2021a). "Data Co-Operatives through Data Sovereignty." *Smart Cities* 4 (3), 1158–1172. DOI: 10.3390/smartcities4030062.

Calzada, I. (2021b). "Epilogue. RESETTING smart city citizenship: Amidst the post-COVID-19 hyperconnected-virialised societies." In: *Smart City Citizenship*. Elsevier, pp. 235–244. ISBN: 9780128153000. DOI: 10.1016/B978-0-12-815300-0.09987-1.

Caravella, S., V. Costantini, and F. Crespi (2021). "Mission-Oriented Policies and Technological Sovereignty: The Case of Climate Mitigation Technologies." *Energies* 14 (20), 6854. DOI: 10.3390/en14206854.

Cattaruzza, A., D. Danet, S. Taillat, and A. Laudrain (2016). "Sovereignty in cyberspace: Balkanization or democratization." In: *2016 International Conference on Cyber Conflict (CyCon U.S.)* IEEE, pp. 1–9. ISBN: 978-1-5090-5258-5. DOI: 10.1109/CYCONUS.2016.7836628.

Chapdelaine, P. and J. McLeod Rogers (2021). "Contested Sovereignties: States, Media Platforms, Peoples, and the Regulation of Media Content and Big Data in the Networked Society." *Laws* 10 (3), 66. DOI: 10.3390/laws10030066.

Chen, Y., S. Chen, J. Liang, L. W. Feagan, W. Han, S. Huang, and X. S. Wang (2020). "Decentralized data access control over consortium blockchains." *Information Systems* 94, 101590. ISSN: 0306-4379. DOI: 10.1016/j.is.2020.101590.

Christakis, T. (2020). "'European Digital Sovereignty': Successfully Navigating Between the 'Brussels Effect' and Europe's Quest for Strategic Autonomy." *SSRN Electronic Journal*, 1–103. DOI: 10.2139/ssrn.3748098.

Corbett, J. and L. Cochrane (2020). "Geospatial Web, Participatory." In: *International Encyclopedia of Human Geography*. Elsevier, pp. 131–136. ISBN: 9780081022962. DOI: 10.1016/B978-0-08-102295-5.10604-3.

Couture, S. and S. Toupin (2019). "What does the notion of "sovereignty" mean when referring to the digital?" *New Media & Society* 21 (10), 2305–2322. ISSN: 1461-4448. DOI: 10.1177/1461444819865984.

Crespi, F., S. Caravella, M. Menghini, and C. Salvatori (2021). "European Technological Sovereignty: An Emerging Framework for Policy Strategy." *Inter economics* 56 (6), 348–354. ISSN: 0020-5346. DOI: 10.1007/s10272-021-1013-6.

Cuno, S., L. Bruns, N. Tcholtchev, P. Lämmel, and I. Schieferdecker (2019). "Data Governance and Sovereignty in Urban Data Spaces Based on Standardized ICT Reference Architectures." *Data* 4 (1), 16. DOI: 10.3390/data4010016.

Dabrock, P. (2020). "How to Put the Data Subject's Sovereignty into Practice. Ethical Considerations and Governance Perspectives." In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. Ed. by A. Markham, J. Powles, T. Walsh, and A. L. Washington. New York, USA: ACM, pp. 1–2. ISBN: 9781450371100. DOI: 10.1145/3375627.3377142.

Diesen, G. (2021). *Great Power Politics in the Fourth Industrial Revolution: The Geoeconomics of Technological Sovereignty*. First edition. London: I. B. Tauris & Company Limited. ISBN: 978-0-7556-0701-3. DOI: 10.5040/9780755607037.

DSSC (2023). *Starter Kit for Data Space Designers*. Ed. by Data Spaces Support Centre. URL: https://dssc.eu/download/802/ (visited on Mar. 23, 2023).

Edler, J., K. Blind, R. Frietsch, S. Kimpeler, H. Kroll, C. Lerch, T. Reiss, F. Roth, T. Schubert, J. Schuler, and R. Walz (2020). *Technology sovereignty: From demand to concept*. Ed. by Fraunhofer ISI. Karlsruhe, Germany. URL: https://publica.fraunhofer.de/dokumente/N-599757.html (visited on Mar. 15, 2023).

Edler, J., K. Blind, H. Kroll, and T. Schubert (2021). *Technology Sovereignty as an Emerging Frame for Innovation Policy – Defining Rationales, Ends and Means*. Ed. by Fraunhofer ISI. Karlsruhe, Germany. URL: https://publica.fraunhofer.de/dokumente/N-638343.html (visited on Mar. 15, 2023).

Esposito, C., A. Castiglione, and K.-K. R. Choo (2016). "Encryption-Based Solution for Data Sovereignty in Federated Clouds." *IEEE Cloud Computing* 3 (1), 12–17. DOI: 10.1109/MCC.2016.18.

Esposito, C., A. Castiglione, F. Frattini, M. Cinque, Y. Yang, and K.-K. R. Choo (2019). "On Data Sovereignty in Cloud-Based Computation Offloading for Smart Cities Applications." *IEEE Internet of Things Journal* 6 (3), 4521–4535. DOI: 10.1109/JIOT.2018.2886410.

European Commission (2020). *Shaping Europe's Digital Future*. Ed. by European Union. DOI: 10.2759/091014.

Filippi, P. de and S. McCarthy (2012). "Cloud Computing: Centralization and Data Sovereignty." *European Journal of Law and Technology* 3 (2), 1–18. URL: https://ssrn.com/abstract=2167372 (visited on Mar. 15, 2023).

Floridi, L. (2020). "The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU." *Philosophy & technology* 33 (3), 369–378. ISSN: 2210-5433. DOI: 10.1007/s13347-020-00423-6.

Friedrichsen, M. and P.-J. Bisa, eds. (2016). *Digitale Souveränität: Vertrauen in der Netzwerkgesellschaft*. Wiesbaden: Springer VS. ISBN: 978-3-658-07349-7.

Garousi, V., M. Felderer, and M. V. Mäntylä (2019). "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering." *Information and Software Technology* 106, 101–121. ISSN: 0950-5849. DOI: 10.1016/j.infsof.2018.09.006.

Gates, C. and J. Slonim (2003). "Owner-controlled information." In: *Proceedings of the 2003 workshop on New security paradigms - NSPW '03*. Ed. by O. S. Saydjari, S. Foley, R. Sekar, C. F. Hempelmann, and V. Raskin. New York, USA: ACM Press, pp. 103–111. ISBN: 1581138806. DOI: 10.1145/986655.986670.

German Ethics Council (2017). *Big Data and Health – Data Sovereignty as the Shaping of Informational Freedom: Executive Summary & Recommendations*. Berlin. URL: https://www.ethikrat.org/en/publications/publication-details/?tx_wwt3shop_detail%5Bproduct%5D=4&tx_wwt3shop_detail%5Baction%5D=index&tx_wwt3shop_detail%5Bcontroller%5D=Products&cHash=7bb9aadb656b877f9dbd49a61e39df2f (visited on Mar. 15, 2023).

Grant, P. (1983). "Technological Sovereignty: Forgotten Factor in the 'Hi-Tech' Razzamatazz." *Prometheus* 1 (2), 239–270. ISSN: 0810-9028. DOI: 10.1080/08109028308628930.

Gupta, A., C. Lanteigne, and S. Kingsley (2020). *SECure: A Social and Environmental Certificate for AI Systems*. DOI: 10.48550/arXiv.2006.06217.

Hallinan, D. (2022). *Data Protection and Privacy, Volume 14: Enforcing Rights in a Changing World*. Computers, Privacy and Data Protection Ser. London: Bloomsbury Publishing Plc. ISBN: 9781509954513.

Hartsch, F., J. Kemmerer, E. R. Labelle, D. Jaeger, and T. Wagner (2021). "Integration of Harvester Production Data in German Wood Supply Chains: Legal, Social and Economic Requirements." *Forests* 12 (4), 460. DOI: 10.3390/f12040460.

Hong, S. and H. Kim (2020). "VaultPoint: A Blockchain-Based SSI Model that Complies with OAuth 2.0." *Electronics* 9 (8), 1231. DOI: 10.3390/electronics9081231.

Hummel, P., M. Braun, S. Augsberg, and P. Dabrock (2018). "Sovereignty and data sharing." *ITU Journal: ICT Discoveries* 1 (2). ISSN: 2616-8375.

Hummel, P., M. Braun, M. Tretter, and P. Dabrock (2021). "Data sovereignty: A review." *Big Data & Society* 8 (1), 1–17. ISSN: 2053-9517. DOI: 10.1177/2053951720982012.

Irion, K. (2012). "Government Cloud Computing and National Data Sovereignty." *Policy & Internet* 4 (3-4), 40–71. ISSN: 1944-2866. DOI: 10.1002/poi3.10.

Janardhanan, S. and C. Mas-Machuca (2022). "Modeling and Evaluation of a Data Center Sovereignty." In: *2022 18th International Conference on the Design of Reliable Communication Networks (DRCN)*. IEEE, pp. 1–8. ISBN: 978-1-6654-0987-2. DOI: 10.1109/DRCN53993.2022.9758037.

Jarke, M., B. Otto, and S. Ram (2019). "Data Sovereignty and Data Space Ecosystems." *Business & Information Systems Engineering* 61 (5), 549–550. ISSN: 2363-7005. DOI: 10.1007/s12599-019-00614-2.

Kagermann, H., K.-H. Streibich, and K. Suder (2021). *Digital Sovereignty: Status Quo and Perspectives*. acatech Impuls. Munich: acatech - National Academy of Science and Engineering. ISBN: 978-3-96834-011-1.

Komaitis, K. (2021). "Europe's ambition for digital sovereignty must not undermine the Internet's values." *Computer Fraud & Security* 2021 (1), 11–13. ISSN: 1361-3723. DOI: 10.1016/S1361-3723(21)00008-7.

Kuhrmann, M., D. M. Fernández, and M. Daneva (2017). "On the pragmatic design of literature studies in software engineering: an experience-based guideline." *Empirical Software Engineering* 22 (6), 2852–2891. ISSN: 1382-3256. DOI: 10.1007/s10664-016-9492-y.

Kukkola, J. (2018). "Civilian and military information infrastructure and the control of the Russian segment of Internet." In: *2018 International Conference on Military Communications and Information Systems (ICMCIS)*. IEEE, pp. 1–8. ISBN: 978-1-5386-4559-8. DOI: 10.1109/ICMCIS.2018.8398700.

Kushwaha, N., P. Roguski, and B. W. Watson (2020). "Up in the Air: Ensuring Government Data Sovereignty in the Cloud." In: *2020 12th International Conference on Cyber Conflict (CyCon)*. IEEE, pp. 43–61. ISBN: 9789-949-9904-7-4. DOI: 10.23919/CyCon49761.2020.9131718.

Lauf, F., S. Scheider, S. Meister, M. Radic, P. Herrmann, M. Schulze, A. T. Nemat, S. J. Becker, M. Rebbert, C. Abate, R. Konrad, J. Bartsch, T. Dehling, and A. Sunyaev (2021). *Data Sovereignty and Data Economy—Two Repulsive Forces?* Ed. by Fraunhofer Institute for Software and Systems Engineering ISST. Dortmund. DOI: 10.24406/ISST-N-634865.

Lian, Y. (2021). *Data Rights Law 3.0*. Peter Lang UK. ISBN: 9781789978384.

Litvinenko, A. (2021). "Re-Defining Borders Online: Russia's Strategic Narrative on Internet Sovereignty." *Media and Communication* 9 (4), 5–15. DOI: 10.17645/mac.v9i4.4292.

Lynch, C. R. (2020). "Contesting Digital Futures: Urban Politics, Alternative Economies, and the Movement for Technological Sovereignty in Barcelona." *Antipode* 52 (3), 660–680. ISSN: 0066-4812. DOI: 10.1111/anti.12522.

Mannhardt, F., A. Koschmider, N. Baracaldo, M. Weidlich, and J. Michael (2019). "Privacy-Preserving Process Mining - Differential Privacy for Event Logs." *Business & Information Systems Engineering* 61 (5), 595–614. ISSN: 2363-7005. DOI: 10.1007/s12599-019-00613-3.

March, C. and I. Schieferdecker (2021). "Technological Sovereignty as Ability, Not Autarky." *SSRN Electronic Journal*, 1–39. DOI: 10.2139/ssrn.3872378.

Mark, R. (2019). "Ethics of Public Use of AI and Big Data." *The ORBIT Journal* 2 (2), 1–33. ISSN: 2515-8562. DOI: 10.29297/orbit.v2i1.101.

Martens, K. and J. Zscheischler (2022). "The Digital Transformation of the Agricultural Value Chain: Discourses on Opportunities, Challenges and Controversial Perspectives on Governance Approaches." *Sustainability* 14 (7). DOI: 10.3390/su14073905.

Maurer, T., I. Skierka, R. Morgus, and M. Hohmann (2015). "Technological sovereignty: Missing the point?" In: *2015 7th International Conference on Cyber Conflict: Architectures in Cyberspace*. IEEE, pp. 53–68. ISBN: 978-9-9499-5442-1. DOI: 10.1109/CYCON.2015.7158468.

Mawere, M. and G. van Stam (2020). "Data Sovereignty: A Perspective From Zimbabwe." In: *12th ACM Conference on Web Science Companion*. New York, USA: ACM, pp. 13–19. ISBN: 9781450379946. DOI: 10.1145/3394332.3402823.

Merkel, A. (2019). *Speech by Federal Chancellor Dr Angela Merkel opening the 14th Annual Meeting of the Internet Governance Forum in Berlin on 26 November 2019*. Berlin. URL: https://www.bundeskanzler.de/bk-en/news/speech-by-federal-chancellor-dr-angela-merkel-opening-the-14th-annual-meeting-of-the-internet-governance-forum-in-berlin-on-26-november-2019-1701494 (visited on Mar. 15, 2023).

Micheli, M., M. Ponti, M. Craglia, and A. Berti Suman (2020). "Emerging models of data governance in the age of datafication." *Big Data & Society* 7 (2), 1–15. ISSN: 2053-9517. DOI: 10.1177/2053951720948087.

Mooy, M. de (2017). *Rethinking Privacy Self-Management and Data Sovereignty in the Age of Big Data*. Ed. by Bertelsmann Foundation. DOI: 10.11586/2017009.

Munoz-Arcentales, A., S. López-Pernas, A. Pozo, Á. Alonso, J. Salvachúa, and G. Huecas (2019). "An Architecture for Providing Data Usage and Access Control in Data Sharing Ecosystems." *Procedia Computer Science* 160, 590–597. ISSN: 1877-0509. DOI: 10.1016/j.procs.2019.11.042.

Nagel, L. and D. Lycklama (2021). *Design Principles for Data Spaces - Position Paper*. Ed. by International Data Spaces Association. DOI: 10.5281/ZENODO.5105744.

Nast, M., B. Rother, F. Golatowski, D. Timmermann, J. Leveling, C. Olms, and C. Nissen (2020). "Work-in-Progress: Towards an International Data Spaces Connector for the Internet of Things." In: *2020 16th IEEE International Conference on Factory Communication Systems (WFCS)*. IEEE, pp. 1–4. ISBN: 978-1-7281-5297-4. DOI: 10.1109/WFCS47810.2020.9114503.

Nugraha, Y., Kautsarina, and A. S. Sastrosubroto (2015). "Towards data sovereignty in cyberspace." In: *2015 3rd International Conference on Information and Communication Technology (ICoICT)*. IEEE, pp. 465–471. DOI: 10.1109/icoict.2015.7231469.

Otto, B. (2019). *Interview with Reinhold Achatz on "Data Sovereignty and Data Ecosystems"*. DOI: 10.1007/s12599-019-00609-z.

Otto, B. and A. Burmann (2021). "Europäische Dateninfrastrukturen." *Informatik Spektrum* 44 (4), 283–291. ISSN: 0170-6012. DOI: 10.1007/s00287-021-01386-4.

Otto, B., S. Steinbuss, A. Teuscher, and S. Lohmann (2019). *IDS Reference Architecture Model*. Ed. by International Data Spaces Association. DOI: 10.5281/ZENODO.5105529.

Pedreira, V., D. Barros, and P. Pinto (2021). "A Review of Attacks, Vulnerabilities, and Defenses in Industry 4.0 with New Challenges on Data Sovereignty Ahead." *Sensors* 21 (15), 5189. DOI: 10.3390/s21155189.

Perarnaud, C. and R. Fanni (2022). *The EU Data Act: Towards a new European data revolution?* URL: https://ideas.repec.org/p/eps/cepswp/35693.html (visited on Mar. 15, 2023).

Peterson, Z. N. J., M. Gondree, and R. Beverly (2011). "A Position Paper on Data Sovereignty: The Importance of Geolocating Data in the Cloud." In: *Proceedings of the 3rd USENIX Conference on Hot Topics in Cloud Computing*. Ed. by I. Stoica and J. Wilkes. USENIX Association, pp. 1–5. URL: https://dl.acm.org/doi/10.5555/2170444.2170453 (visited on Mar. 15, 2023).

Plateaux, A., P. Lacharme, C. Rosenberger, and K. Murty (2013). "A contactless e-health information system with privacy." In: *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, pp. 1660–1665. ISBN: 978-1-4673-2480-9. DOI: 10.1109/IWCMC.2013.6583805.

Pohle, J. and T. Thiel (2020). "Digital sovereignty." *Internet Policy Review* 9 (4), 1–19. ISSN: 2197-6775. DOI: 10.14763/2020.4.1532.

Polatin-Reuben, D. and J. Wright (2014). "An Internet with BRICS Characteristics: Data Sovereignty and the Balkanisation of the Internet." *4th USENIX Workshop on Free and Open Communications on the Internet*, 1–10. URL: https://www.usenix.org/conference/foci14/workshop-program/presentation/polatin-reuben (visited on Mar. 15, 2023).

Posch, R. (2017). "Digital Sovereignty and IT-Security for a Prosperous Society." In: *Informatics in the Future*. Ed. by H. Werthner and F. van Harmelen. Cham: Springer International Publishing, pp. 77–86. ISBN: 978-3-319-55734-2. DOI: 10.1007/978-3-319-55735-9_7.

Qarawlus, H., M. Hellmeier, J. Pieperbeck, R. Quensel, S. Biehs, and M. Peschke (2021). "Sovereign Data Exchange in Cloud-Connected IoT using International Data Spaces." In: *2021 IEEE Cloud Summit (Cloud Summit)*. IEEE, pp. 13–18. ISBN: 978-1-6654-2582-7. DOI: 10.1109/IEEECloudSummit52029.2021.00010.

Redeker, M., S. Volgmann, F. Pethig, and J. Kalhoff (2020). "Towards Data Sovereignty of Asset Administration Shells across Value Added Chains." In: *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, pp. 1151–1154. ISBN: 978-1-7281-8956-7. DOI: 10.1109/ETFA46521.2020.9211955.

Ruohonen, J. (2021). "The Treachery of Images in the Digital Sovereignty Debate." *Minds and Machines* 31 (3), 439–456. ISSN: 0924-6495. DOI: 10.1007/s11023-021-09566-7.

Ruparelia, N. B. (2016). *Cloud computing*. The MIT Press essential knowledge series. Cambridge, Massachusetts and London, England: The MIT Press. ISBN: 9780262334129.

Sarabia-Jacome, D., I. Lacalle, C. E. Palau, and M. Esteve (2019). "Enabling Industrial Data Space Architecture for Seaport Scenario." In: *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*. IEEE, pp. 101–106. ISBN: 978-1-5386-4980-0. DOI: 10.1109/WF-IoT.2019.8767216.

Schleicher, D., C. Fehling, S. Grohe, F. Leymann, A. Nowak, P. Schneider, and D. Schumm (2011). "Compliance Domains: A Means to Model Data-Restrictions in Cloud Environments." In: *2011 IEEE 15th International Enterprise Distributed Object Computing Conference*. IEEE, pp. 257–266. ISBN: 978-1-4577-0362-1. DOI: 10.1109/EDOC.2011.22.

Singi, K., S. G. Choudhury, V. Kaulgud, R. J. C. Bose, S. Podder, and A. P. Burden (2020). "Data Sovereignty Governance Framework." In: *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*. New York, USA: ACM, pp. 303–306. ISBN: 9781450379632. DOI: 10.1145/3387940.3392212.

Tan, K.-L., C.-H. Chi, and K.-Y. Lam (2022). *Analysis of Digital Sovereignty and Identity: From Digitization to Digitalization*. arXiv. DOI: 10.48550/arXiv.2202.10069.

Taylor, J. and T. Kukutai, eds. (2016). *Indigenous data sovereignty: Toward an agenda*. Vol. no. 38. Research monograph / Centre for Aboriginal Economic Policy Research, College of Arts and Social Sciences, The Australian National University, Canberra. Acton, ACT, Australia: Australian National University Press. ISBN: 9781760460303.

Taylor, R. D. (2020). ""Data localization": The internet in the balance." *Telecommunications Policy* 44 (8), 102003. ISSN: 0308-5961. DOI: 10.1016/j.telpol.2020.102003.

Vadiati, N. (2022). "Alternatives to smart cities: A call for consideration of grassroots digital urbanism." *Digital Geography and Society* 3, 100030. ISSN: 2666-3783. DOI: 10.1016/j.diggeo.2022.100030.

Vaile, D. (2014). "The Cloud and data sovereignty after Snowden." *Australian Journal of Telecommunications and the Digital Economy* 2 (1), 1–59. ISSN: 2203-1693. DOI: 10.7790/ajtde.v2n1.31.

vom Brocke, J., A. Simons, B. Niehaves, B. Niehaves, K. Reimer, R. Plattfaut, and A. Cleven (2009). "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process." *ECIS 2009 Proceedings*.

vom Brocke, J., A. Simons, K. Riemer, B. Niehaves, R. Plattfaut, and A. Cleven (2015). "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research." *Communications of the Association for Information Systems* 37, 205–224. DOI: 10.17705/1CAIS.03709.

von der Leyen, U. (2020). *Shaping Europe's digital future: op-ed by Ursula von der Leyen, President of the European Commission*. Brussels. URL: https://ec.europa.eu/commission/presscorner/detail/en/ac_20_260 (visited on Mar. 15, 2023).

Webster, J. and R. T. Watson (2002). "Analyzing the Past to Prepare for the Future: Writing a Literature Review." *MIS Quarterly* 26 (2), xiii–xxiii. ISSN: 0276-7783.

Zieglmeier, V. and A. Pretschner (2021). *Trustworthy Transparency by Design*. DOI: 10.48550/arXiv.2103.10769.

Zrenner, J., F. O. Möller, C. Jung, A. Eitel, and B. Otto (2019). "Usage control architecture options for data sovereignty in business ecosystems." *Journal of Enterprise Information Management* 32 (3), 477–495. ISSN: 1741-0398. DOI: 10.1108/JEIM-03-2018-0058.

Paper II

Table A.2 Metadata Overview of Paper II

Title	Implementing Data Sovereignty: Requirements & Challenges from Practice
Authors	Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Julia Pampus *Fraunhofer ISST, Dortmund, Germany* Haydar Qarawlus *Fraunhofer ISST, Dortmund, Germany* Falk Howar *TU Dortmund & Fraunhofer ISST, Dortmund, Germany*
Publication Year	2023
Publication Type	Conference
Conference Name	18th International Conference on Availability, Reliability and Security (ARES)
Conference Location	Benevento, Italy
Conference Date	29. August 2023 - 01. September 2023
Publisher / Database	ACM
DOI / Link	https://doi.org/10.1145/3600160.3604995
Status	Published
Ranking	VHB: - (2024 Rating) ICORE: B (2026 Rating) ERA: B (2010 Rating)
Comment	Presented at the 20th International Workshop on Trust, Privacy and Security in the Digital Society (Trustbus). Trustbus was formerly a separate conference (also ICORE ranked: B) and is now part of ARES.

Implementing Data Sovereignty: Requirements & Challenges from Practice

Malte Hellmeier
malte.hellmeier@isst.fraunhofer.de
Fraunhofer ISST
Dortmund, Germany

Julia Pampus
julia.pampus@isst.fraunhofer.de
Fraunhofer ISST
Dortmund, Germany

Haydar Qarawlus
haydar.qarawlus@isst.fraunhofer.de
Fraunhofer ISST
Dortmund, Germany

Falk Howar*
falk.howar@tu-dortmund.de
TU Dortmund
Dortmund, Germany

ABSTRACT

Data sovereignty, the possibility to keep control over data, is gaining increasing attention in both research and industry. Due to complex supply chains and a strong trend toward digitization, digital assets are essential to be fast and competitive. As a result, companies need to share data while retaining control over it to prevent unwanted leaks of sensitive data. However, implementing effective data governance, access, and usage control mechanisms can be challenging, especially in cross-company data sharing networks and ecosystems like dataspaces. In this paper, we examine the industrial landscape and interview eleven experts from software providers and producing organizations to identify their requirements and challenges of existing data sovereign solutions. Based on Grounded Theory and semi-structured interviews, we explore the motivations and issues behind data sharing from an Information Systems and Software Engineering point of view. The findings include current industrial contexts, use cases, and solutions with data sovereignty's technical and non-technical implementations. Seven requirements and thirteen challenges were observed throughout a qualitative analysis. Clustered by organizational, technical, personal, and emotional viewpoints, they are discussed with initial approaches for mitigation. The results identify current practical needs and will enable the design of future data sovereignty solutions in theory and different practical domains.

CCS CONCEPTS

• **Security and privacy** → **Information flow control**; *Human and societal aspects of security and privacy*; *Security requirements*; Trust frameworks; **Security services**; • **Information systems** → **Data exchange**; *Enterprise information systems*; • **Applied computing** → *Enterprise computing*; • **Social and professional topics** → Network access control.

*Also with Fraunhofer ISST.

ARES 2023, August 29–September 01, 2023, Benevento, Italy

ACM ISBN 979-8-4007-0772-8/23/08.
https://doi.org/10.1145/3600160.3604995

KEYWORDS

Data Sovereignty, Data Sharing, Data Ecosystems, Trust, Semi-structured Interviews

ACM Reference Format:
Malte Hellmeier, Julia Pampus, Haydar Qarawlus, and Falk Howar. 2023. Implementing Data Sovereignty: Requirements & Challenges from Practice. In *The 18th International Conference on Availability, Reliability and Security (ARES 2023), August 29–September 01, 2023, Benevento, Italy.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3600160.3604995

1 INTRODUCTION

The relevance of data as an "essential feature of digitization and data economy" [15, p. 1] is an increasingly important topic. Therefore, researchers and practitioners create mechanisms to protect data in different lifecycles stages, from creation over storage until sharing and deletion [3, 26]. Examples include restricted access to data, control of data flows, prevention of inferences to confidential information, and encryption mechanisms [7]. However, data should not only be transmitted and stored securely; the self-determination of data, often referred to as data sovereignty [14, 19, 22], is highly relevant internationally. Discussions started on a political level based on the US Patriot Act for the USA [13], the European Strategy for Data [9], and the protection regulations of the General Data Protection Regulation (GDPR) [11].

Therefore, the term *data sovereignty* is receiving growing attention, especially in the context of industrial data ecosystems [24]. With these ecosystems' implementation in cross-company networks, technical and non-technical requirements, experiences, and challenges are emerging. In this paper, we look at the industry's current state through semi-structured interviews to derive theoretical and practical implications. Thus, our work is guided by the following Research Question (RQ): *What are the current requirements and challenges in implementing data sovereignty in practice?*

The concrete contributions of this paper are structured as follows: Section 2 introduces background concepts and presents related work for delimitation and classification into current research streams. Section 3 outlines the research method. Next, the definition of requirements for data sharing and challenges for implementing data sovereignty in practice with its contexts, solutions, and approaches is explained in Section 4 and discussed in Section 5. Finally, Section 6 concludes our findings.

2 BACKGROUND

In the following, the essential terms for the present research are introduced and described to create a uniform understanding, including a review of related work.

2.1 Data Sovereignty

The possibility of keeping control over own data assets is often referred to as *data sovereignty* [14]. Its origin goes back to the US Patriot Act in 2001, giving the US Government permission to access data stored on servers in the USA [13]. Over time, discussions in research and practice started on how to stay self-determined with data when sharing it with third parties. Thus, different domains investigate this topic, from technical solutions [23, 37] to regulations and political discussions [18, 20]. This results in various definitions depending on the viewpoint. The three most important ones for the present research are introduced in Table 1. While implementation challenges in practice focus on individual and company levels, related terms with a broader and more political direction, like digital or technological sovereignty, are not considered further [14].

Table 1: Data Sovereignty Definitions

Definition	Source
"Data sovereignty refers to the self determination of individuals and organizations with regard to the use of their data."	[19, p. 550]
"Consequently, the data sovereignty concept arises, which is defined as the ability of the data owner to decide itself how to share and use its data."	[31, p. 101]
"Data sovereignty is the capability of a natural person or corporate entity for exclusive selfdetermination with regard to its economic data goods."	[22, p. 27]

2.2 Grounded Theory

Grounded Theory is a methodology to systematically create a theory based on data [12]. Its origins are rooted in social research based on the concepts of Barney Glaser and Anselm Strauss, published in 1967 [12]. Since then, it has been applied and adapted over time, while the initial method is often referred to as *classical grounded theory*. Along with *interpretive grounded theory* by Juliet Corbin and Anselm Strauss [5] and *constructivist grounded theory* by Kathy Charmaz [4], these are the three most used types. Their usage is not limited to social science but also gets applications in various other research areas, including Software Engineering (SE) [1, 36]. The methodological types' similarities and differences are generally discussed in [34] and with a SE lens in [36].

2.3 Related Work

Before our research, different publications dealt with data sovereignty, its requirements, challenges, and implementations.

In the publication from Akaichi and Kirrane, requirements for usage control are extracted from the literature. Based on a descriptive scenario, the authors studied existing frameworks and concepts in [2] by holistically considering the implementation of data sharing.

Pedreira et al. conducted a literature review on industrial attacks, their defenses, and vulnerabilities in [25]. They focus on different sovereignty terms, namely cyber, digital, and data sovereignty. Future challenges discussed by the authors are, e.g., the enablement of data exchange, control over data and devices, enforcement mechanisms, data analysis, and interoperability [25].

Another systematic literature review was conducted by Hummel et al. and, similarly to [25], focuses on data, digital, and cyber sovereignty [17]. Based on their final literature corpus, the authors identified requirements like data security or legal aspects and four significant challenges (namely *features of data*, *technical designs*, *epistemic issues*, and *legal issues* [17]).

Legal challenges are further discussed by Singi et al.. The main issues are, on the one hand, different laws in different countries which must be complied with by globally operating companies and, on the other hand, the fast changes of rules based on new regulations or government adjustments [35].

Furthermore, businesses and governments face challenges in implementing data sovereignty. The literature discusses those issues from a more legislative perspective on controlling cyberspace data flows in [6].

Recently, Schmidt et al. have elaborated on challenges for sovereign data exchange through a systematic literature review and expert interviews in [32]. In contrast to this work, all potential interviewees have a strong research focus as they "have co-authored journal and conference papers" [32, p. 52]. The researchers identified thirteen challenges structured into three categories:

- *Organizational challenges*: Three challenges are described, focusing on managing jurisdictions, like legal guidelines and different regulations, missing standards for technological solutions, and aversion due to high costs and missing incentives and benefits [32].
- *Data processing and publishing challenges*: The five challenges of this category focus on enabling data privacy through anonymization while preserving its metadata and usability in an interoperable and scalable way with less complexity [32].
- *Infrastructure challenges*: Aspects of trust in the organizations and the infrastructure, as well as elements of data origin, access and usage control, and verifiability, form the five challenges of the infrastructure category [32].

Since data sovereignty requirements and challenges have been viewed predominantly from a theoretical or legislative lens, we aim to close the gap by extending existing research like [2] or [32] with a practical point of view. Thus, this study embeds in the current research stream by extending previously identified facets from literature reviews with requirements and challenges from experts in the industry.

3 RESEARCH METHOD

In order to answer the RQ, we conducted semi-structured interviews in conjunction with Grounded Theory. With these, we followed the interview guidelines from primary literature [21, 28, 33] with concrete experiences in SE [8, 16] to structure our research. Since different strands have arisen over time, as introduced in Section 2.2, we adhere to the principles of the reinterpreted *constructivist grounded theory* [4], as we build our results on a RQ answered from interviews

Table 2: Semi-structured Interview Guideline

Part	ID	Main Questions
Briefing		*Interview conditions and asking for name, company information, job position, and years of experience.*
I. Use Case	Q1	Does your company share data with other companies, or does your company create solutions that help other companies share data?
	Q2	What are the areas/fields/production stages in which your company [shares data]/[supports to share data] or plans to share data in the near future (providing and consuming)?
	Q3	Why does your company or its customers need to share data? What is an example use(-case) description of data your company produces and/or consumes?
II. Requirements	Q4	What are your company's requirements regarding data sharing?
	Q5	How important is the control over data during and after the data sharing process (incl. their confidentiality)?
	Q6	Does your company or used solutions currently ensure control over data during and after the data sharing (technically or non-technically)? If so, how?
III. Data Sovereignty	Q7	How would you describe the "data sovereignty concept" in your own words (one sentence)?
	Q8	*Showing data sovereignty definitions of Table 1 to the participants.*
	Q9	Based on the previously described data sharing cases, (how) does your company deal with data sovereignty?
IV. Challenges	Q10	How does your company implement data sovereignty, and what are the experiences so far (maturity level)?
	Q11	What challenges do you (or your company) encounter in implementing data sovereignty (from project observations, etc.)?
	Q12	How do you (or your company) try to solve those challenges?
De-Briefing		*Additional questions and information of the further procedure.*

that are analyzed with initial coding and a downstream categorization [36]. This methodological process is described in detail in the following subsections.

3.1 Interview Preparation

Based on different types of interviews, identifying concrete, practical challenges need further explanations to have the possibility for follow-up questions to identify the actual practical problems. Thus, fully structured interviews, questionnaires, or group interviews are not eligible. Instead, to focus each interview on the RQ, we applied individual semi-structured interviews using questions and possibilities to direct the discussion to the problem [16].

Interview Design. The design consists of *main questions* that define the overall structure, extended by *follow-up questions* to identify additional information and *probes* to manage the conversation [28]. We developed an interview guideline of main questions, separated into four parts, presented in Table 2. The final version of the interview guideline was created in three iterative cycles. We conducted two pilot interviews a priori internally to test and validate the questions and train the interviewer [21].

As Rapley recommended in the briefing, we asked for permission to record the interview, explained the procedure and intention of the interview, and highlighted the importance of confidentiality and anonymity [27]. Furthermore, demographic data were collected. The questions of the *I. Use Case* part (Q1-Q3) aimed to identify areas, participants, and reasons for data sharing, followed by the *II. Trust & Requirements* part (Q4-Q6) to shift the discussion into control of data sharing. Afterward, in the *III. Data Sovereignty* segment (Q7-Q9), the interviewer presented three definitions of the term (shown in Table 1) to reduce the influence of different interpretations on upcoming questions. Finally, current problems and possible mitigations were discussed in the *IV. Challenges* part (Q10-Q12). A short de-briefing completed all interviews [21].

Participant Selection. Related work like [32] identify challenges by reviewing literature and interviewing participants with research backgrounds. Therefore, one participant selection criteria for the present study are persons currently working for a company instead of a university or research organization. To ensure generality, we identified participants from different organizations of different sizes with headquarters located in different countries. We contacted 15 persons via e-mail and asked for participation, as 15 +/- 10 people are sufficient for interview studies [21].

3.2 Data Collection

All potential candidates, who responded to our inquiry and had the willingness to participate, were contacted to schedule an appointment. Based on the results from the pilot interviews, we scheduled a one-hour online video meeting with every participant individually.

Based on existing research discussions for taking notes or using a tape recorder [16], we decided to use both. As a result, three people took part in each interview: the interviewee, one main interviewer, and a passive interviewer taking notes. Every session consisted of a briefing, the question-based part of the interview, de-briefing [21] and was recorded and transcribed with Microsoft Teams[1], including oral and written consent from the participants to be transparent and GDPR-compliant.

[1] www.microsoft.com/teams (Accessed: 2023-06-28)

3.3 Data Analysis

To analyze the collected data, we used the qualitative research software MaxQDA[2]. In the first step, the transcripts are imported and manually checked against the recording by correcting wrong parts or incorrectly transcribed words. In the second step, coding is used for analysis by associating specific codes to parts of the data. While different coding techniques exist, we used *Initial Coding* (similar to *Open Coding* [5, 29]) as it is a standard first-cycle coding method used for interview data in Grounded Theory [4, 29]. Here, paragraph-by-paragraph coding is used with a downstream categorization as a second-cycle coding method based on focused and theoretical coding [29, 36].

4 RESULTS

The results are built on the analyzed semi-structured interviews conducted in February 2023. Of the 15 requested persons, 11 were willing to participate in the study and were interviewed by a subgroup of the author team. Further interview details, including the job position or department and the duration of the recorded interview itself, are summarized in Table 3. Every participant is employed by a different-sized company, from small-sized with under 100 employees over midsize companies to corporations with over 400 000 employees. Professional experience in their jobs is also widely divided from 3-4 years until over 40 years. Nevertheless, 64 % have more than 18 years of experience. These details are not included in Table 3 to preserve the participants' anonymity and avoid inferences about the persons and employment.

Table 3: Interviewee Details

Participant	Job Position / Department	Duration
Interviewee 1	Development	51 min
Interviewee 2	Development	44 min
Interviewee 3	IT Management	49 min
Interviewee 4	Research & Development	41 min
Interviewee 5	Development	29 min
Interviewee 6	Development	30 min
Interviewee 7	Development	40 min
Interviewee 8	Development	37 min
Interviewee 9	Research & Development	40 min
Interviewee 10	IT Management	44 min
Interviewee 11	IT Management	31 min
$\sum$		**436 min**

In the following, the analysis results are presented by giving details on the context of the use cases tackled by the companies, the requirements for implementing data sovereignty, current possibilities and solutions to ensure control over data, and the challenges encountered, including the first approaches for mitigation. To strengthen the results, direct quotations from the interviews are presented, which were translated carefully from German into English by the authors.

[2] www.maxqda.com (Accessed: 2023-06-28)

4.1 Context

In part *I. Use Case* of the interviews, participants were asked to describe the context and use cases of data sharing within their organizations. Several motivations for data sharing were identified from the discussions. We classify these as follows:

- *Production process optimizations:* Many interviewees cited optimizing production lines and processes as a main motivation for internal and external data sharing. This includes optimizations across all stages of production, e.g., identifying problems in the supply chain and promptly resolving them, the ability to identify and issue recalls, and efficiently handling the recycling process.
- *Sustainability:* Mitigating environmental impacts were named by several interviewees. The main motivation is to reduce the products and services' carbon footprint and increase transparency.
- *Regulatory requirements:* Numerous participants from the production sector named regulatory obligations as a reason for sharing data. Some regulatory bodies require fully transparent data regarding CO_2 tracking. Other examples include obtaining product specifications for certification procedures and ensuring compliance with supply chain laws.
- *Data monetization:* Multiple interviewees named monetization as a motivation for data sharing. This can happen when the data provider has no direct use for the data but is aware of the usage by third parties. In this scenario, data can be sold or given on a pay-per-use basis.

All use cases may involve different types of data, including (1) text or numeric data for direct reading and analysis (e.g., .pdf, .docx files or spreadsheets), (2) machine-readable data for direct processing through various systems, or (3) engineering design files (e.g., computer-aided design).

4.2 Requirements

Part *II. Requirements* of the interview was about defining conditions for data sharing. A classified list is depicted in Table 4, including the number of coded segments in parentheses. Each requirement is mapped to the interviewees (I) mentioning it and presented in the following. The requirements are categorized into two clusters with an organizational and technical focus.

Table 4: Requirements for Data Sharing

No.	Requirement	Mentions (I)
	Organizational Requirements (38 coded segments)	
R1	Law, Legal, Restrictions & Guidelines (24)	1-11
R2	Confidentiality (11)	6,7,9-11
R3	Data/Service Classification (3)	1,7,11
	Technical Requirements (36 coded segments)	
R4	Policy Enforcement (24)	1,3,4-10
R5	Data Security & Integrity (6)	3,4,7,9,11
R6	Data Visibility & Offering (4)	1,3,9
R7	Transparency (2)	2,4

4.2.1 Organizational Requirements.

Law, Legal, Restrictions & Guidelines (R1). As mentioned above, data sharing is often motivated by regulatory requirements like the European Data Act [9] or the GDPR [11]. They describe the technical and non-technical aspects and specify rules, i.a., for data deletion, processing, and persistence. As noted by almost all of the interviewees, it significantly affects the implementation of data sharing solutions: "We have to make sure that every end user in a system where data is collected can agree that there is transparency about the localization of the data, about the processing in this network of many products and cloud services, etc. and that this right to deletion is also supported, from end to end across all layers and all products" (I4). In addition, some interviewees also mentioned domain-specific requirements. These comprise, e.g., encryption techniques, signature mechanisms, or transfer technologies. For instance, I7 explained that specific signatures might be accepted in court proceedings and some not.

Confidentiality (R2). Ensuring the confidential handling of data is a fundamental requirement, particularly in situations where data sharing is restricted to a closed group. It is essential that no unauthorized third party can access any sensitive information to prevent potential privacy violations. Unauthorized access can lead to inferences being made about the data provider, which can have significant consequences (I7).

Data/Service Classification (R3). During the interviews, it was noted that the requirements for data sharing are highly dependent on the type of data being shared. For instance, I8 highlighted that internal company policies often define data categorization on external, internal, confidential, or secret levels. Additionally, companies may determine which data can potentially create or threaten a competitive advantage, making it particularly sensitive (I7, I11). As a result, the services that work with such data need to be classified and restricted.

4.2.2 Technical Requirements.

Policy Enforcement (R4). Policy enforcement was considered particularly important by most interviewees. It covers several aspects: On the one hand, control after the data sharing is supposed to be given. It includes specified data deletion methods or the need to anonymize data. On the other hand, access to the data should be implemented through internal and external user authorizations, which are supplemented by organizational assurances. Here, I4 emphasized the importance of implementing policy enforcement at all levels. I10 identified the need to provide technical interfaces for this purpose. The interviewees generally stressed the importance of utilizing state-of-the-art techniques and standards for data transfer, as this can significantly enhance the integration of disparate systems.

Data Security & Integrity (R5). In addition, data security and integrity are crucial to the employers of the interviewees and their clients during data transfer and storage. It includes, e.g., signature and audit processes, encryption mechanisms, fail-safe systems, and ISO certifications of the used infrastructure and hardware. With this, I7 referred to legal regulations stating that questions like "Which signatures, for example, are recognized by a court and which are not?" must be considered during implementation.

Data Visibility & Offering (R6). In data sharing, the discoverability and identification of data (sources) are essential (I3). This is partly accomplished by identifying "the data and [giving] context to this data on the meta-level" (I3). In accordance with closed-group data sharing scenarios, companies acting as data owner/provider only want to make data available to specific organizations.

Transparency (R7). Complementary to R6, I4 highlighted the importance of transparent processes and information flows in data sharing scenarios to create security and trust. These can be established, e.g., by auditing and verifying log files (I2).

4.3 Solutions

In Q6 and Q9, the interviewees were asked to outline current solutions for implementing the requirements aforementioned in Section 4.2, particularly regarding data sovereignty. Separated into organizational and technical, these can be summarized as follows.

4.3.1 Organizational Solutions.

Interviewees considered signing physical, *legally binding contracts* among participating parties as indispensable requirements before exchanging data. In case legal issues arise, all participants stated that enforcing contractually agreed-upon fines is a safety net for their data.

Next, I3, I7, and I10 mentioned already enforced organizational approaches like assigning *confidentiality levels* with binding obligations to data sets. Each level defines the roles in addition to the persons having rights to access, use, or modify specific data sets. In addition, various interviewees named already applied mixed *changes in internal processes* regarding how data is handled. This includes guidelines, such as the checks done before data is shared, integrity checks, or data classification. Complementary, *training* ensures that the relevant staff is adequately equipped with the knowledge to handle various situations relating to sharing business-critical and sensitive information (I9).

4.3.2 Technical Solutions.

Available technical solutions encompass a wide range of techniques. In general, a *metadata enrichment* of data sets is often used to add additional layers of identifying information, such as confidentiality level or ownership chain. This enables extended checks upon internal or external data sharing.

Next, some interviewees named *data encryption* a way to address security requirements for data during transit (I7, I9). This is meant to act as an additional layer of encryption in addition to the standard Transport Layer Security encryption used during transmission. Complementary, I3 named the use of *watermarking* solutions to partially protect data by tagging it upon creation or retrieval. This allows the quick identification and patching of problem sources.

Furthermore, interviewees stated that various *access control* methods are used to control who can gain access to data (I5, I6, I10). This includes the basic username and password-based authentication methods, API endpoint protection, and standardized and sophisticated access control methods using the eXtensible Access Control Markup Language (XACML) or the Open Digital Rights Language[3].

[3] www.w3.org/TR/odrl-model (Accessed: 2023-06-28)

In data ecosystems, some interviewees shared their experience using *dataspace connectors*, such as the Connector of the Eclipse Dataspace Components[4] for cross-organizational data sharing (I1, I2, I6, I10). Following the principles and architecture guidelines of the International Data Spaces (IDS) and Gaia-X, dataspace connectors enable various data sharing methods concerning techniques like metadata enrichment, data encryption, and access control.

4.4 Challenges

Part *IV. Challenges* of the interviews focused on the experiences and issues of implementing data sovereignty in practice. We distinguish between three types of challenges: organizational, technical, and personal and emotional challenges.

Each coded segment of an interview is associated with a challenge abbreviated by C1 to C13 for thematic clustering. These challenges are bunched into the three aforementioned types. Table 5 gives a summarized overview, including the number of coded segments for every challenge and a mapping to the interviewees I1 to I11 addressing it, sorted according to the frequency. In the following, we describe each challenge in more detail, giving concrete examples from the interviews.

4.4.1 Organizational Challenges.

Law, Legal, Restrictions & Guidelines (C1). A major organizational challenge strongly related to R1 comprises legislative issues. One aspect mentioned in the interviews deals with regulations based on operational areas. Besides specific country rules, examples include regulations like the Data Act on the European level [10]. Data sharing can further be retained by specific company restrictions or over regulations. With this, data sharing companies possess existing paper-based contracts, as stated in Section 4.3. Living digital solutions for strengthening data sovereignty can transform paper-based contracts into machine-readable formats. Nevertheless, I7 suspects that paper contracts will continue to exist since linking information in a source code or payload to its legislative interpretation is always needed.

Realization (C2). Since first technical data sovereignty solutions exist, as shown in Section 4.3, they lack in their realization: "still the challenge that most companies are nowadays not yet ready with their business processes to be able to use something like this correctly at all" (I2). Besides internal process issues, the interviewees mentioned problems like time pressure or a too-early technological stage, as most companies are unfamiliar with such topics while others are still exploring based on test data. Further discussed realization problems cover missing requirements or meaningful use cases.

Staffing (C3). In order to create successful data sovereignty concepts, employees of different domains have to work together to include technical solutions in their systems. In the opinion of I2, "these concepts sell much better to business decision-makers because, for IT decision-makers, it is an insane amount of extra work." People must understand that protecting data is only possible with additional effort. As mentioned by I8, creating awareness for data sharing topics inside a company is challenging. Thus, staffing problems arise: "So I see very few companies that really staff this with their A team because they have understood how important the topic will be in the future. This happens relatively rarely" (I2).

Business & Economics (C4). From a business and economic perspective, controlling data requires additional technical effort. Even if the first solutions are free and open source, their operation is associated with additional costs. In a producing company, these extra costs do not improve the final product directly. Moreover, small and medium-sized enterprises working with big corporations have to adopt their data sharing processes, sometimes against their will, as they are determined by the player with the significantly stronger market position (I7).

Security & Privacy (C5). I10 compared a water hose to a data sharing network with different companies and systems: "If you have a gap somewhere, then water leaks out, and consequently, this complete system chain must actually be secured." Therefore, privacy and security must be observed throughout, as current challenges identified by the interviewees include unwanted data sharing, competitors' access, physical cloning, and data leaks.

Standards (C6). Data Sharing is the focus of numerous initiatives, such as the IDS Association or Gaia-X. Since strong standards are needed, multiple roles and different initiatives can be problematic, as it "only works if you have a solution at the end, a bit like VHS tapes. If you have five different systems in the market, then there are silos again. Nobody really wants this" (I5). Thus, bringing existing academic standards into practice is essential.

Communication (C7). The participants mentioned problems in communication, mainly on an international level. Despite contracts and regularities addressed in C1, sharing data with players from other countries is dangerous. Therefore, bringing together different views can be challenging in an international cross-company data sharing network.

4.4.2 Technical Challenges.

Access & Usage Control (C8). The most quantitative significant technical challenge based on the coded segments deals with control mechanisms. While, among others, access control solutions are currently used as described above, there are open points, such as role-based access control [30] for company roles. More complex in this respect is usage control defined through policies and their enforcement. Some participants argue that full technical usage policy enforcement is impossible and that real humans are always needed for verification. I5 claimed that "there are first approaches for real data sovereignty, while 95 % of the usage control concepts are extended access control methods." The complexity lies in the technical enforcement on the data-consuming side. I10 expects the first solutions in many years, while I7 suspects it in 10-15 years at the latest.

Infrastructure & Landscape (C9). When sharing data between companies, diverse infrastructures become challenging when trying to control data assets. Some participants integrate cloud services into their landscape, while others work with different external apps or old legacy systems. Implementing currently used solutions like

[4] www.github.com/eclipse-edc (Accessed: 2023-06-28)

Table 5: Challenges in Implementing Data Sovereignty

No.	Challenge	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11
	Organizational Challenges (57 coded segments)											
C1	Law, Legal, Restrictions & Guidelines (16)		x	x	x	x	x	x	x	x	x	
C2	Realization (12)		x	x	x	x	x		x		x	
C3	Staffing (7)		x	x					x	x		x
C4	Business & Economic (6)							x		x	x	
C5	Security & Privacy (6)							x	x	x	x	
C6	Standards (6)					x	x	x	x	x		
C7	Communication (4)			x			x					
	Technical Challenges (49 coded segments)											
C8	Access & Usage Control (19)	x			x	x	x	x	x	x	x	x
C9	Infrastructure & Landscape (13)	x			x	x		x	x	x	x	x
C10	Data Processing Life Cycle (11)				x		x		x	x	x	x
C11	Identity Management (6)						x		x			
	Personal & Emotional Challenges (14 coded segments)											
C12	Trust (11)						x		x	x	x	
C13	Comfort (3)	x					x			x		

a dataspace connector based on the IDS principles [23] or a secure execution environment allows processing data in a trusted environment to maintain data sovereignty. However, it is challenging "because the added value only arrives when I can also use the received data in third-party applications" (I7).

Data Processing Life Cycle (C10). Cavanillas et al., and Rahul and Banyal describe different data life cycle steps and activities, from creation, storage, and usage, until deletion [3, 26]. Deleting data becomes challenging in conjunction with C9 and is addressed by I4 and I10 since some systems cannot perform remove actions due to legal retention periods or backup strategies with snapshots. In addition, open building blocks mentioned in the interviews cover detection mechanisms in external systems, data integrity, and data processing concepts.

Identity Management (C11). Problems regarding certificates, identification, and verification are mentioned mainly by I6 and I8. Questions raised by the interviewees are, for example, from I6, "How do I identify companies?" or "How can a Trust Anchor ensure that the core of a claim is really attached to?" Since verification authorities exist, it is still unclear who verifies them since scandals are sometimes uncovered.

4.4.3 Personal & Emotional Challenges.

Trust (C12). Companies and decision-makers fear losing control over data and thus hold data sharing activities. Anxiety exists from the possible misuse of data due to problems of employees that technical measures cannot solve, "even if you have the best data sovereignty [...] you can take a picture, or someone has a pad in front of them and can transcribe it" (I8). Therefore, trust is needed in all stages, from confidence in compliance with the legal agreements (C1) to the processing software systems (C9, C10) until the employees and administrators working with the systems (C3, C5).

Comfort (C13). In order to strengthen data sovereignty, additional organizational and technical solutions are needed, leading to less comfort (I9) and carelessness (I6). Therefore, persuasion is needed because "the biggest hindrance is always to convince people that they need an additional part of software" (I1).

4.5 Approaches

In Q12, the interviewees were asked about approaches to addressing the mentioned challenges. Structured by the previously introduced three challenge types, their ideas, and solutions are presented based on 49 coded segments.

Organizational Approaches. As often mentioned in various interviews, starting small by testing existing solutions and using pilot products can help mitigate C2 since all issues cannot be solved directly by one solution (I3, I6, I8, I10). Therefore, creating a uniform understanding and showing the benefits for a company with workshops, training, and meetings brings awareness among all employees in a company and helps with the challenges C2, C3, and C5. I10 summarizes it: "I think the most important thing, particularly with all these data sovereignty technologies, is that they have to be usable quickly, that you have to make it understandable to the user."

Since most issues cannot be solved by one single company, working together to create uniform guidelines (C1), standards (C6), and clear communication (C7) distribute the work and costs (C4) among several players. I2 suggested including large companies: "where there are dominant players who have the financial resources to guide its implementation in the early days." Network effects occur when working together, which drive the dissemination of data sharing solutions [19]. To get there, I5 and I7 suggest combining existing standards and solutions (C6) instead of developing everything from scratch.

Technical Approaches. From a technical point of view, different approaches such as XACML, dataspace connectors based on IDS [23], encryption techniques, or secure multi-party computation are named concerning the access and usage control challenge (C8). Following the infrastructure (C9) and data processing life cycle (C10), I9 suggests bringing enforcement logic down to network components by "designing edge components a little more intelligent." Other approaches focus more on decentralization in the system architecture, webhooks, and digital watermarking techniques allowing subsequent verifiability and traceability. If no possibilities exist to keep control of the data, e.g., if a system only adds a deletion flag and never deletes data, I4 suggests that "data should not be allowed to flow through such a system on the consumer side" or only with intensive tracking and logging. Since most value creation occurs in the consuming application and not in a secure container, I7 summarizes: "That is the reason for a middle part. I would call it *trust-based policy execution*. I still do not have real enforcement from a provider's perspective, but I can use certified components, verifiable identity organizations, and the software manufacturer to find out if deletion functions exist in the consuming components." Such an identification chain helps mitigate C11.

Personal & Emotional Approaches. Besides organizational and technical challenges, finding approaches for the personal and emotional challenges of C12 and C13 is demanding since they are individual and difficult to measure. For example, I6 and I9 suggest increasing the trust relationship between participants by creating a trust chain with external verifications. It can lead to confidence in the community instead of trusting a single player or system.

5 DISCUSSION

Implementing data sovereignty in different IT landscapes is challenging and drives current research activities and practical realizations. In order to answer the RQ, the presented qualitative-based approach on eleven semi-structured interviews with practitioners has identified a comprehensive set of requirements and challenges. The following sections discuss the implications for theory and practice with their limitations and future research opportunities.

5.1 Theoretical Implications

Unlike existing research with a strong literature review focus [17, 25], our results extend the findings of Schmidt et al. with a remarkable practical point of view and a baseline requirement set. Under their challenge classification in [32], our resulting categories with organizational and technical focus show strong similarities. Therefore, this study confirmed and expanded the current research stream and strengthened the need to work on those challenges. Even though pure SE research is necessary to find technical solutions, considering adjacent domains and non-technical aspects is crucial since they influence implementations on a broader level.

5.2 Practical Implications

Our work reinforces the practical relevance of data sovereignty in industrial environments due to increased digitization in recent years. The proposed requirements and challenges contribute to existing working streams on privacy and security in data sharing networks by pointing out open points. Although the interviewees were asked to outline possible approaches to tackle those challenges, their answers differed. For example, some participants emphasized the importance of one unique standard or solution, while others reinforced their diversity to preserve the freedom and power of selection. Therefore, future research and developments in practical environments are necessary to increase data self-determination.

5.3 Limitations & Future Work

Limitations of this work are partly owed to the nature of interview-based studies. Firstly, the interviews vary in length, and their execution and analysis are carried out by more than one researcher leading to possible personal distinctions. Secondly, since all interviewees are in an employment relationship, they may have been constrained due to the interview situation or company policies. Thirdly, generalizability must be checked since the results are built upon our eleven-participant sample. Fourthly, although our selection of participants includes international companies, all interviewees are based in Europe with a background in data sovereignty and data ecosystems. Thus, there might need to be more diversity to consider the subject of data sovereignty from all perspectives.

Concerning future research opportunities, subjects like watermarking, data integrity, usage control, and enforcement were considered essential and relevant by the interviewees but were not fully addressed by current solutions. Therefore, future work should look at different technical approaches and their integration into existing solutions. For example, most of the answers focused on the application layer but could be considered more diverse from a technical point of view. Moreover, the analysis of the interviews showed that some of the requirements are vague and not very specific, as the topic of data sovereignty encompasses various aspects and is challenging to comprehend. Accordingly, this affords research in requirements engineering and data sharing design.

6 CONCLUSION

In this work, we examined the topic of data sovereignty and its practical implementation in industry. We considered and extended relevant work by extracting information from an interview study. Based on Grounded Theory, our evaluation resulted in the presented context information, requirements, existing solutions, encountered challenges, and suggested approaches.

As mentioned in the discussion, our work significantly contributes to the field of data sovereignty by exploring the current implementation landscape in relevant industries. Furthermore, we show that the chosen topic offers additional research potential, encompassing diverse issues. Despite common initial assumptions, data sovereignty is more than the technical implementation of data security and privacy. Its implementation and related challenges include organizational as well as personal and emotional challenges. Comprehensively, our work shows the importance of trust for data sovereignty and highlights the need to explore how technology solutions can strengthen it in different domains and contexts.

ACKNOWLEDGMENTS

This work was partially supported by the German Federal Ministry for Economic Affairs and Climate Action (funding number: 13IK004).

REFERENCES

[1] Steve Adolph, Wendy Hall, and Philippe Kruchten. 2011. Using grounded theory to study the experience of software development. *Empirical Software Engineering* 16, 4 (2011), 487–513. https://doi.org/10.1007/s10664-010-9152-6

[2] Ines Akaichi and Sabrina Kirrane. 2022. Usage Control Specification, Enforcement, and Robustness: A Survey. https://doi.org/10.48550/arXiv.2203.04800

[3] José María Cavanillas, Edward Curry, and Wolfgang Wahlster (Eds.). 2016. *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe*. Springer International Publishing AG, Cham.

[4] Kathy Charmaz. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage Publications Ltd, London.

[5] Juliet M. Corbin and Anselm L. Strauss. 2015. *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4 ed.). Sage, Los Angeles, Calif.

[6] Jing de Jong-Chen. 2015. Data sovereignty, cybersecurity, and challenges for globalization. *Georgetown Journal of International Affairs* 16 (2015). https://heinonline.org/HOL/LandingPage?handle=hein.journals/geojaf16&div=72&id=&page=

[7] Dorothy E. Denning and Peter J. Denning. 1979. Data Security. *Comput. Surveys* 11, 3 (1979), 227–249. https://doi.org/10.1145/356778.356782

[8] Tore Dybå, Rafael Prikladnicki, Kari Rönkkö, Carolyn Seaman, and Jonathan Sillito. 2011. Qualitative research in software engineering. *Empirical Software Engineering* 16, 4 (2011), 425–429. https://doi.org/10.1007/s10664-011-9163-y

[9] European Commission. 2020. European data strategy: Making the EU a role model for a society empowered by data. https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

[10] European Commission. 2022. Proposal for a Regulation of the European Parliament and of the Council on harmonised Rules on Fair Access to and Use of Data (Data Act). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0068

[11] European Parliament and Council of the European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/eli/reg/2016/679/oj

[12] Barney G. Glaser and Anselm L. Strauss. 1967. *The discovery of grounded theory: Strategies for qualitative research*. Aldine, New York, NY.

[13] Dara Hallinan. 2022. *Data Protection and Privacy, Volume 14: Enforcing Rights in a Changing World*. Bloomsbury Publishing Plc, London.

[14] Malte Hellmeier and Franziska von Scherenberg. 2023. A Delimitation of Data Sovereignty from Digital and Technological Sovereignty. *ECIS 2023 Research Papers* (2023). https://aisel.aisnet.org/ecis2023_rp/306

[15] Arghavan Hosseinzadeh, Andreas Eitel, and Christian Jung. 2020. A Systematic Approach toward Extracting Technically Enforceable Policies from Data Usage Control Requirements. In *Proceedings of the 6th International Conference on Information Systems Security and Privacy*. SCITEPRESS - Science and Technology Publications, 397–405. https://doi.org/10.5220/0008936003970405

[16] Siw Elisabeth Hove and Bente Anda. 2005. Experiences from Conducting Semi-structured Interviews in Empirical Software Engineering Research. In *11th IEEE International Software Metrics Symposium (METRICS'05)*. IEEE. https://doi.org/10.1109/METRICS.2005.24

[17] Patrik Hummel, Matthias Braun, Max Tretter, and Peter Dabrock. 2021. Data sovereignty: A review. *Big Data & Society* 8, 1 (2021), 1–17. https://doi.org/10.1177/2053951720982012

[18] Kristina Irion. 2012. Government Cloud Computing and National Data Sovereignty. *Policy & Internet* 4, 3-4 (2012), 40–71. https://doi.org/10.1002/poi3.10

[19] Matthias Jarke, Boris Otto, and Sudha Ram. 2019. Data Sovereignty and Data Space Ecosystems. *Business & Information Systems Engineering* 61, 5 (2019), 549–550. https://doi.org/10.1007/s12599-019-00614-2

[20] Neal Kushwaha, Przemyslaw Roguski, and Bruce W. Watson. 2020. Up in the Air: Ensuring Government Data Sovereignty in the Cloud. In *2020 12th International Conference on Cyber Conflict (CyCon)*. IEEE, 43–61. https://doi.org/10.23919/CyCon49761.2020.9131718

[21] Steinar Kvale and Svend Brinkmann. 2009. *InterViews: Learning the craft of qualitative research interviewing* (2 ed.). Sage, Los Angeles.

[22] Lars Nagel and Douwe Lycklama. 2021. Design Principles for Data Spaces - Position Paper. https://doi.org/10.5281/ZENODO.5105744

[23] Boris Otto, Sebastian Steinbuss, Andreas Teuscher, and Steffen Lohmann. 2019. IDS Reference Architecture Model. https://doi.org/10.5281/ZENODO.5105529

[24] Boris Otto, Michael ten Hompel, and Stefan Wrobel. 2022. *Designing Data Spaces*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-93975-5

[25] Vítor Pedreira, Daniel Barros, and Pedro Pinto. 2021. A Review of Attacks, Vulnerabilities, and Defenses in Industry 4.0 with New Challenges on Data Sovereignty Ahead. *Sensors* 21, 15 (2021). https://doi.org/10.3390/s21155189

[26] Kumar Rahul and Rohitash Kumar Banyal. 2020. Data Life Cycle Management in Big Data Analytics. *Procedia Computer Science* 173 (2020), 364–371. https://doi.org/10.1016/j.procs.2020.06.042

[27] Tim Rapley. 2004. Interviews. In *Qualitative Research Practice*, Clive Seale, Giampietro Gobo, Jaber Gubrium, and David Silverman (Eds.). Sage Publications Ltd, London, 15–33.

[28] Herbert J. Rubin and Irene S. Rubin. 2012. *Qualitative interviewing: The art of hearing data* (3 ed.). Sage, Los Angeles and London and New Delhi and Singapore and Washington DC.

[29] Johnny Saldaña. 2013. *The coding manual for qualitative researchers* (2 ed.). Sage, Los Angeles, Calif.

[30] Ravi S. Sandhu, Edward J. Coyne, Hal L. Feinstein, and Charles E. Youman. 1996. Role-based access control models. *Computer* 29, 2 (1996), 38–47. https://doi.org/10.1109/2.485845

[31] David Sarabia-Jacome, Ignacio Lacalle, Carlos E. Palau, and Manuel Esteve. 2019. Enabling Industrial Data Space Architecture for Seaport Scenario. In *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*. IEEE, 101–106. https://doi.org/10.1109/WF-IoT.2019.8767216

[32] Kaja Schmidt, Gonzalo Munilla Garrido, Alexander Mühle, and Christoph Meinel. 2022. Mitigating Sovereign Data Exchange Challenges: A Mapping to Apply Privacy- and Authenticity-Enhancing Technologies. In *Trust, Privacy and Security in Digital Business*, Sokratis Katsikas and Steven Furnell (Eds.). Lecture Notes in Computer Science, Vol. 13582. Springer International Publishing, Cham, 50–65. https://doi.org/10.1007/978-3-031-17926-6_4

[33] Clive Seale, Giampietro Gobo, Jaber Gubrium, and David Silverman (Eds.). 2004. *Qualitative Research Practice*. Sage Publications Ltd, London. https://doi.org/10.4135/9781848608191

[34] Kailah Sebastian. 2019. Distinguishing Between the Strains Grounded Theory: Classical, Interpretive and Constructivist. *Journal for Social Thought (JST)* 3, 1 (2019). https://ojs.lib.uwo.ca/index.php/jst/article/view/4116

[35] Kapil Singi, Swapnajeet Gon Choudhury, Vikrant Kaulgud, R. Jagadeesh ChandraP. Bose, Sanjay Podder, and Adam P. Burden. 2020. Data Sovereignty Governance Framework. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*. ACM, New York, 303–306. https://doi.org/10.1145/3387940.3392212

[36] Klaas-Jan Stol, Paul Ralph, and Brian Fitzgerald. 2016. Grounded theory in software engineering research. In *Proceedings of the 38th International Conference on Software Engineering*, Laura Dillon, Willem Visser, and Laurie Williams (Eds.). ACM, New York, NY, 120–131. https://doi.org/10.1145/2884781.2884833

[37] Valentin Zieglmeier and Alexander Pretschner. 2021. Trustworthy Transparency by Design. https://doi.org/10.48550/arXiv.2103.10769

Received 30 April 2023; accepted 11 June 2023; revised 28 June 2023

Paper III

Table A.3 Metadata Overview of Paper III

Title	Data Sovereignty in Information Systems
Authors	Franziska von Scherenberg *Fraunhofer ISST, Dortmund, Germany* Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Boris Otto *Fraunhofer ISST & TU Dortmund, Dortmund, Germany*
Publication Year	2024
Publication Type	Journal (Fundamentals)
Journal Name	Electronic Markets
Publisher / Database	Springer
DOI / Link	https://doi.org/10.1007/s12525-024-00693-4
Status	Published
Ranking	VHB: B (2024 Rating) CORE: A (2020 Rating) ERA: - (2010 Rating) SJR: Q1 (2024 Rating)
Comment	The article was published in the *Fundamentals* section. Compared to research papers, fundamentals papers focus on the structure of currently relevant topics and terms.[1] Furthermore, the paper is awarded with the *Most Cited Paper* award in 2024 by Electronic Markets.

[1] https://electronicmarkets.org/call-for-papers/ccfp/call-for-fundamentals (accessed on Feb. 10, 2026)

Electronic Markets (2024) 34:15
https://doi.org/10.1007/s12525-024-00693-4

FUNDAMENTALS

Data Sovereignty in Information Systems

Franziska von Scherenberg[1] · Malte Hellmeier[1] · Boris Otto[1,2]

Received: 27 January 2023 / Accepted: 12 January 2024

Abstract
Data has become a strategic asset for societal prosperity and economic competitiveness. There has long been an academic consensus that the value of data unfolds during its use. Consequently, many stakeholders have called for expanding the use and reuse of data, including the public and open variety, as well as that from private data providers. However, citizens and organizations want self-determination over their data use, that is, data sovereignty. This fundamentals paper applies a literature review to conceptualize the term in Information Systems (IS) research by summarizing current findings and definitions to add further structure to the field. It contributes to the current research streams by introducing a core conceptual model consisting of seven interacting core aspects, involving trust between data providers and consumers for data assets, supported by data infrastructure and contractual agreements on all data lifecycle stages. We evaluate and discuss this conceptual model through recent field examples and provide an overview of future research opportunities.

Keywords Data sovereignty · Information systems · Literature review · Conceptualization

JEL Classification L86 · M15

Introduction

Data assets are digital goods and the basis for all Information Systems (IS). They have become a strategic asset for societal prosperity and economic competitiveness. Accordingly, studying data as a concept is essential for further developments in IS research (Singi et al., 2020). According to recent estimations, data assets will grow in quantity and increase in importance in the upcoming years (Statista, 2022). More data sharing that further enables data-driven decision-making is one reason for this growth and increase (Munoz-Arcentales et al., 2019). However, those who share their data fear a loss of control and competitive disadvantage, which is why a data economy that protects the individual and organization's interests is vital (Lauf et al., 2021). In this context, data sovereignty becomes a success factor as its implementation strengthens actors to decide on the use of their data as an economic asset (Banse, 2021), thus paving the way to a digital space wherein providers and consumers can control all of their data actions.

Practically speaking, data sovereignty constitutes a key piece in building safe environments where data providers and consumers overcome trust issues while sharing data. Given the importance of handling data according to sovereignty principles, policymakers must ensure "fair data sharing practices" (European Commission, 2022, p. 26) and create secure frameworks. Legislations derived from the European Strategy for Data, such as the Data Governance Act (DGA) or the Data Act (DA), as well as the General Data Protection Regulation (GDPR) that came into force in 2016, regulate the data protection of different actors. They directly influence technological design in order to balance economic opportunities with society's interests in sharing and reusing data (Labadie et al., 2019). In addition, politicians, organizations, and other stakeholders recognize data sovereignty as essential for controlling the data of individuals and organizations; however, when referring to data

Responsible Editor: Christiane Lehrer

✉ Franziska von Scherenberg
franziska.von.scherenberg@isst.fraunhofer.de

Malte Hellmeier
malte.hellmeier@isst.fraunhofer.de

Boris Otto
boris.otto@isst.fraunhofer.de; boris.otto@tu-dortmund.de

[1] Fraunhofer ISST, Speicherstr. 6, 44147 Dortmund, Germany

[2] Chair for Industrial Information Management, TU Dortmund, Joseph-von-Fraunhofer-Str. 2-4, 44227 Dortmund, Germany

Published online: 27 February 2024

sovereignty, it is often unclear whether these actors share the same understanding of the concept.

A deeper understanding of how organizations and individuals technically implement control over data when sharing it is crucial for all research into digital self-determination and motivates the study of data sovereignty in IS. First, academia demands more alignment and less isolation in exploring the core aspects and relations of data sovereignty. Moreover, there is persistent terminological ambiguity in IS research, particularly in studies on indigenous people (Taylor & Kukutai, 2016), data sovereignty in the cloud (Irion, 2012), or data sovereignty of individuals and enterprises (Jarke et al., 2019), to name just a few examples. Additionally, holistic research on data sovereignty that observes the overall concept is either absent or fails to live up to the expectations of exploring the handling of data in a sovereign way within IS (Hummel et al., 2021; Kushwaha et al., 2020).

Moreover, former IS research has faced challenges, provided loose ends, or come to divergent conclusions. This is shown by different data sovereignty definitions with contrasting focuses on law (Docter & Fuchs, 2020), self-determination (Banse, 2021; Jarke et al., 2019; Nagel & Lycklama, 2021), data flows (Lauf et al., 2021), or informational freedom (German Ethics Council, 2017). Further, studies have focused on implementing data sovereignty without clarifying the concept's foundation (Opriel et al., 2021; Plattform Industrie 4.0, 2022). Other research has analyzed the impact of data sovereignty on data sharing without examining the concept itself (Azkan et al., 2022). Previous articles and studies have described sovereignty as a capability (Nagel & Lycklama, 2021) without proving a theoretical approach. This study aims to fill these gaps by analyzing the current state of research and developing a conceptual model that can help researchers and practitioners navigate this cluttered field so as to gain a mutual understanding of the concept.

This research is structured as follows: It begins by describing data sovereignty and contextualizing its background, as well as analyzing previous contributions in IS and adjacent domains from academia and practitioners. As described in detail in the appendix, a Multivocal Literature Review (MLR) is applied to developing a conceptual model (Fig. 1) that specifies the core aspects of data sovereignty (Table 2). It draws on agency theory to support a consistent understanding of the concept within the realm of IS, as well as to form a baseline for further analytical, exploratory, and design-oriented research. Using real-world examples, the proposed model illuminates all core aspects and explains their roles and relations. The paper concludes by discussing theoretical and practical implications while considering limitations and future research opportunities.

Background and related work

In the digital world, the concept of *sovereignty* describes forms of independence, control, and autonomy over digital infrastructures, technologies, data, and digital content

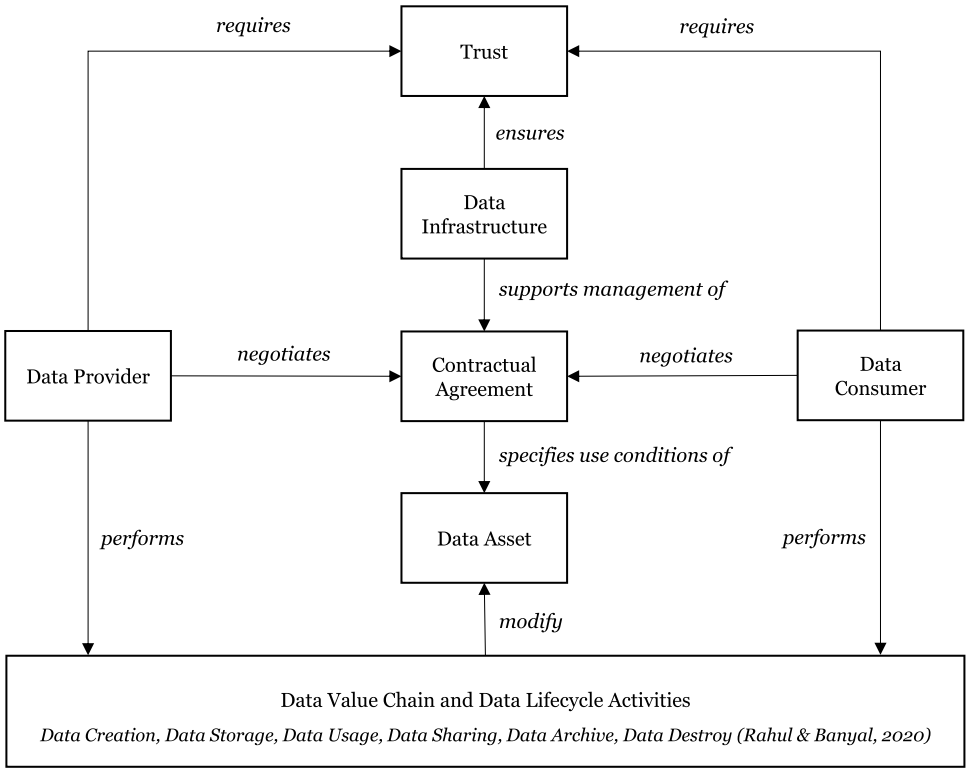


Fig. 1 Conceptualization of data sovereignty in IS

(Pohle & Thiel, 2020). Discussions of sovereignty with a technological focus began in the 1980s (Grant, 1983; Hinsley, 1986) when it extended to various forms and domains, such as technological-, digital-, data-, or cyber sovereignty (Hellmeier & von Scherenberg, 2023). *Data sovereignty* is a relatively new term used in decision-making and data ownership (Hummel et al., 2021). Over time, researchers have shaped its meaning, emphasizing its different nuances. Table 1 summarizes various data sovereignty definitions from different research domains to contextualize the concept.

Direct comparisons reveal different perspectives on the same term. For example, Polatin-Reuben and Wright (2014) mentioned a missing definition and shaped the concept on a national level, while both Jarke et al. (2019) and Nagel and Lycklama (2021) described it for individuals and enterprises. Other publications, such as the German Ethics Council (2017), included such technical aspects as big data, while Docter and Fuchs (2020) introduced the legal perspective. Research has often referred to the notion of digital *self-determination* that exceeds the perspective of data sovereignty as it considers not only one's data but also "data about oneself" (Verhulst, 2023, p. 8) and is related to protecting personal data and user consent. In contrast to data sovereignty, digital self-determination makes no distinctions between data and their actors but sees both as an entity (Verhulst, 2023).

Within IS, current discussions on data sovereignty are increasingly driven by regulations that balance data protection and use before, during, and after the sharing process, such as the European GDPR, the DGA (European Commission, 2020), and the DA (European Commission, 2022). However, control over data is not only a fundamental European principle. Such regulations as China's Personal Information Protection Law (PIPL) or the California Consumer Privacy Act (CCPA) show that it is also gaining increasing attention globally (Chander et al., 2021), as the global rise in data exploitation stems mainly from the market power of monopolistic US and Chinese organizations, thus explaining the increasing demand for new data governance models.

The technical implementation of data sovereignty can initiate beneficial consequences of data sharing since it enables organizations to find a solution for balancing data protection and use. These are, first, *cost sharing*, where actors save money and time when sharing their data under the prerequisites of data sovereignty; second, the *greater common good,* where organizations can, for example, be motivated to share data for the achievement of CO_2 targets; and third, *joint innovation,* which can only occur when actors work together, as most participants are unable to realize the application individually (Data Spaces Support Centre, 2023b). These examples show that value is not created by one player, but

Table 1 Collection of data sovereignty definitions

Authors	Definition	Research domain
Polatin-Reuben and Wright (2014, p. 1)	"The term 'data sovereignty', while lacking a firm definition, refers to a spectrum of approaches adopted by different states to control data generated in or passing through national internet infrastructure. It can be understood as a subset of cyber sovereignty, defined as the subjugation of the cyber domain to local jurisdictions."	IS
German Ethics Council (2017, p. 30)	"Data sovereignty, understood as the responsible shaping of informational freedom, in a manner appropriate to the risks and opportunities presented by big data, is the central ethical and legal goal in confronting the challenges and opportunities presented by big data."	Ethics/Humanities and Social Sciences
Jarke et al. (2019, p. 550)	"Data sovereignty refers to the self-determination of individuals and organizations with regard to the use of their data."	Computer Science
Sarabia-Jacome et al. (2019, p. 101)	"Consequently, the data sovereignty concept arises, which is defined as the ability of the data owner to decide itself how to share and use its data."	Electrical and Electronics Engineers Science
Docter and Fuchs (2020, p. 256)	"Data sovereignty is the concept that data is subject to laws and regulations of a particular nation."	IS
Hong and Kim (2020, p. 19)	"[...] data sovereignty, which refers to the right to use and control one's own information."	Computer Science
Lauf et al. (2021, p. 9)	"Our understanding of data sovereignty is the ability to formulate self-defined data-usage rules, influence and trace the data/information flows while being free in the decision of (not) sharing data and migrating data whenever and wherever it is desirable."	IS
Banse (2021, p. 10)	"Self-determination how, when and at what price others (across the value chain) may use my data"	Computer Science
Nagel and Lycklama (2021, p. 27)	"Data sovereignty is the capability of a natural person or corporate entity for exclusive self-determination with regard to its economic data goods."	IS

Table 2 Specification of the core conceptual data sovereignty aspects and relations

Core aspects	Specification
Data asset	An asset over which control is to be retained. It includes various possibilities, from files over complete batches, databases, or data warehouses to ideas and technologies. Such data assets must be controlled against their access and usage (Munoz-Arcentales et al., 2019).
Data provider	A natural person, company, or organization that has been given the right to control the data asset (Gil et al., 2020; Zrenner et al., 2019).
Data consumer	A natural person, company, or organization interested in using, creating, deleting, or sharing data assets owned or controlled by a data provider (Gil et al., 2020; Zrenner et al., 2019).
Contractual agreement	An agreement, signed by at least the data provider and connected to the data asset, determines their access and usage. The agreement is based on a previously negotiated contract, which can be verbal, written, or digital (Jarke et al., 2019).
Data value chain and data lifecycle activities	A cluster of all activities performed on the data asset throughout its lifecycle, from data creation to storage, analysis, sharing, and deletion (Curry, 2016; Rahul & Banyal, 2020).
Data infrastructure	A system or concept reviewing, documenting, and executing the rules of the contract agreement in the form of policy enforcement (Nagel & Lycklama, 2021), often included in current IT architectures (Hummel et al., 2021) as a data infrastructure (Munoz-Arcentales et al., 2019).
Trust	A fundamental core component for data sovereignty. On the one hand, it is required by all players who want to perform data value chain activities on a data asset (Peterson et al., 2011). On the other hand, the infrastructure helps with its amplification (Nagel & Lycklama, 2021).
Relations	Specification
Data provider and data consumer require trust	Trust is always required by all stakeholders as a baseline (Nagel & Lycklama, 2021).
Data infrastructure ensures trust	A manual or technical infrastructure helps ensure trust (Munoz-Arcentales et al., 2019), for example, through enforcement mechanisms.
Data provider and data consumer negotiate contract agreements	Data providers and consumers have to negotiate a contract to create an agreement that specifies the use conditions of the data asset (Zrenner et al., 2019).
Data infrastructure supports management of contractual agreement	A manual or technical infrastructure supports the management of contract agreements through validation techniques (Munoz-Arcentales et al., 2019).
Contractual agreement specifies use conditions of data asset	A contract agreement specifies the use conditions of the data asset, for example, through policies (Zrenner et al., 2019).
Data provider and data consumer perform data value chain and data lifecycle activities	Data providers or consumers with access to the data asset perform data value chain and data lifecycle activities (Otto et al., 2022).
Data value chain and data lifecycle activities modify data asset	A data asset can reach different statuses and versions because it is modified by data value chain activities (Curry, 2016; Rahul & Banyal, 2020).

through various actors' combinations and data enrichment in *data ecosystems* (Gelhaar et al., 2021).

The *ecosystem* concept originally stems from ecological science and draws on the attention of living organisms that co-exist in a healthy environment (Chapin et al., 2011). Ecosystems and, in this regard, *data ecosystems* do not function with central governance but rather work in balance. They can be open or closed (Oliveira & Lóscio, 2018), and while open data ecosystems are free for everyone to join, the closed variety often enforces technical or legal entry barriers (Capiello et al., 2020; Janssen et al., 2012; van den Homberg & Susha, 2018). Actors in data ecosystems depend on and benefit from each other in equilibrium, without one being dominant. As such, all actors should be equipped with an instrument to control their own data without being controlled by one central instance to create a trusted environment. Consequently, implementing data sovereignty is an essential part of this (Gelhaar et al., 2021; Otto et al., 2022).

Conceptualizing data sovereignty in IS research

This chapter proposes a data sovereignty conceptual model consisting of core conceptual aspects and relations. Conceptual models are critical for simplifying and abstracting

reality, as well as helping researchers and practitioners to understand, organize, and communicate complex or novel concepts (Houy et al., 2012). As described in the appendix, the conceptual model was developed by consulting the IS data sovereignty literature in Tables 1, 2, and Table S1. It is grounded in the agency theory to ensure that the model can fully explain the concept and offer a basis for real-world application (Eisenhardt, 1989).

The core of this theory, developed during the 1960s and 1970s, is to analyze the relationship between two actors (Eisenhardt, 1989). Its underlying assumption is that these two actors pursue their objectives, which often differ, acting in their self-interest. In addition, it implies an information asymmetry between both actors. In order to avoid mistrust, control mechanisms are installed that lead to greater transparency (Eisenhardt, 1989). With the help of this theory, challenges in organizational relationships can be more effectively uncovered, and governance structures more deeply understood (Eisenhardt, 1989).

Through the lens of this theory, data sovereignty can be implemented as an instrument with the central objective of establishing more trust. As outlined in the theory's description, contractual agreements provide the necessary transparency on the actions of both actors (here, data providers and consumers). According to Eisenhardt (1989), this theory can be applied to buyer–supplier and other agency relationships and, therefore, is suited for relations in the context of data sharing that arises in open or closed data ecosystems. With the implementation of data sovereignty, actors have an instrument at hand that paves the way for a more balanced power structure and supports all parties in pursuing their objectives.

The presented conceptual model applies the concept of data sovereignty in IS research, supporting both researchers and practitioners to develop a holistic understanding of the concept and serves to guide those (i.e., practitioners) who seek to implement data sovereignty technically. It aims for a completeness that has, as yet, not been provided by existing IS literature and definitions (see Table 1). In addition, this conceptual model helps all stakeholders better understand and communicate the concept of data sovereignty.

The seven core aspects referenced in Table 2 result from the IS literature's analysis and our experience in this field, using agency theory as the basis for the development of this conceptual model (Creswell, 2009). Details about the MLR search process, including scientific and grey literature, are described in the appendix. The modeling process considered the contributions listed in Table S1, explicitly focusing on data sovereignty in the IS domain. We use examples to explain how we derived the conceptual model when explaining each core aspect. Table 2 summarizes all core aspects and relations and lists their specifications.

With directed arrows, the conceptual model illustrated in Fig. 1 represents the relations of the core conceptual aspects. The model acknowledges the data asset as its central component that must be protected in an organizational or personal context if shared with other parties (Nagel & Lycklama, 2021). During its lifecycle, from creation to sharing and deletion (Rahul & Banyal, 2020), a data asset can reach different statuses and versions because it is *modified* by activities in the data value chain (Curry, 2016). These activities are *performed* by the data provider or the data consumer who gained access to the data asset (Otto et al., 2022). In order to implement data sovereignty, the provider and consumer must *negotiate* a contract that *specifies the use conditions of* the data asset (Zrenner et al., 2019). Access and usage policies are possible examples of such contracts (Gil et al., 2020). Due to frequent mistrust between the parties involved (Lauf et al., 2021), a data provider often seeks to ensure that the consumer only performs data value chain activities described in the contractual agreement. Therefore, a manual or technical data infrastructure helps *ensure* trust because it *supports the management of* contracts through enforcement techniques (Munoz-Arcentales et al., 2019). Nevertheless, trust is always *required* by all stakeholders involved (Nagel & Lycklama, 2021), even if the concept reduces the minimum amount needed to create a data sovereignty solution. The following subsections describe every core aspect in detail.

Data asset

Based on the conceptual model, data sovereignty can be defined as an instrument to keep control over an actor's data asset. Examples of data assets can range from individual files to complete batches and full data streams. Such *data assets* must be controlled in terms of their access and usage (Munoz-Arcentales et al., 2019). Data are defined as assets describing intangible objects that can be reproduced repeatedly (Capiello et al., 2020). However, it is worth noting that there is no single definition of the concept in IS research (McKinney & Yoos, 2010). Data are contextual, and their ownership is difficult to define. They cannot be classified as private or common goods, such as traditional commodities (Jentzsch, 2018), since there are no legally binding concepts regarding their ownership (Bärenfänger, 2017). The data asset has been placed at the base of the model as it is key for each application of data sovereignty. Since the status of the data asset is modified by the data value chain and lifecycle activities, they are directly related to the data asset and positioned at the bottom as a baseline.

Data provider and data consumer

A *data provider* can decide to keep their data private for internal use, share it publicly, or allow access to a restricted number of third parties based on custom rules. For example,

contracts are created and negotiated between the data provider and the *data consumer* to keep control over the data asset. Providers and consumers can be individuals, enterprises, or organizations sharing data assets (Cavanillas et al., 2016; Marfia et al., 2017). In the case of a contractual agreement, the provider can be further divided into the role of a *data owner* that creates and executes control over the data asset and authorizes a data provider to make it available to other parties (Hummel et al., 2021; Otto et al., 2019). In addition, when referring to data consumer, other sources, such as the Data Spaces Support Centre Glossary, use the term *data recipient* (Data Spaces Support Centre, 2023a). Besides contractual arrangements between both partners, data-providing enterprises can share data directly or through existing systems, such as data marketplaces (Nagel & Lycklama, 2021). Here, a data consumer can buy either the data asset itself or limited usage rights. Since both actors are represented as core aspects in the model, they are placed on the left side for the provider part and on the right for the consumer part, as all activities are performed in between them.

Contractual agreement

As stated above, exercising data sovereignty can promote data sharing between organizations. In the traditional sense, written contracts are drawn up to increase trust, which results in a contractual agreement after mutual consent. Due to a lack of control, these agreements are often not fully honored and lack high levels of trustworthiness (Nagel & Lycklama, 2021). IS research has recognized and addressed this problem to enforce *contract agreements* that are negotiated and monitored semi-automatedly with the help of infrastructures and architectures to reduce (un)intentional data misuse (Jarke et al., 2019). Therefore, different systems and processes in various domains focus on smart contracts (Ghazizadeh & Sun, 2021). The data provider and consumer can be two neutral actors creating a contract based on rights and obligations, data usage policies, and terms and conditions (Zrenner et al., 2019), described in more detail in the infrastructure section. They can give or revoke their consent to change access rights and specify conditions of how their data can be accessed and used. The contractual agreement is located in the middle, as it consists of the main conditions for maintaining control over data assets — the main goal of data sovereignty.

Data value chain and data lifecycle activities

As depicted in Fig. 1, the *data value chain* includes different activities in the *data lifecycle* of a data asset: creation, storage, usage, sharing, archiving, and destruction (Rahul & Banyal, 2020). In this context, the implementation of data sovereignty enables an organization or individual to control the data asset throughout the data lifecycle. According to Curry (2016), an information flow consists of different activities that perform transformation steps to turn a data input into a data output. In the context of data sovereignty, the ability to keep control must extend over all data value chain activities, from creation to transformation to deletion, rather than focusing on individual activities (Banse, 2021). The activities must be consistent with the contractual agreements and usage conditions to enable self-determination. Accordingly, the data asset itself in Fig. 1 is not directly linked to the data provider or consumer (Nagel & Lycklama, 2021). Instead, the data provider and data consumer perform value chain activities on the data asset.

Data infrastructure

The data infrastructure component enforces terms and conditions determined in the contractual agreement (Munoz-Arcentales et al., 2019; Nagel & Lycklama, 2021). It is centrally located in the model since it works between the data provider and consumer by validating and executing terms and restrictions (Nagel & Lycklama, 2021), specified in the contractual agreement (see Fig. 1). These terms are divided into access control (AC) and usage control (UC), which protect data assets in almost all activities in the data value chain and lifecycle. As implied by the term AC, the concept focuses on the concrete control of access. Seeing as control is lost once access is granted, UC extends the control over data before and after third-party access (Gil et al., 2020), specifying which aspects of actors in ecosystems can access and use the data (Zrenner et al., 2019). However, AC and UC requirements specified in contractual agreements do not add value if not enforced correctly. Therefore, data infrastructure components, such as software systems, must validate the conditions of the contractual agreement and execute the actions described in the policies (Gil et al., 2020). Concepts based on decentralized identities (Ernstberger et al., 2023) and initiatives, such as the International Data Spaces Association (IDSA) and GAIA-X, operate according to standards and the technical implementation of data infrastructure components to address these problems. Their solutions find application in various domains, such as the cloud, IoT devices (Qarawlus et al., 2021), manufacturing (Landolfi et al., 2019), and many others.

Trust

According to Schilke and Cook (2013), trust has emerged as a central theme in inherently uncertain relationships, with Botsman (2017) defining the term as the "confident relationship with the unknown" (2017, p. 8). While in private and closed scenarios, trust can be established in the first instance because the actors know each other, it is challenging in the

second scenario as the data provider and consumer are partly unknown due to complex supply chain networks with many participants (Gil et al., 2020). In the conceptual model, the relationship of trust needs to be considered from two different angles. In the first step, trust is required by the data provider and consumer (Peterson et al., 2011). In this context, actors in open and closed data ecosystems must establish a fundamental trust relationship in the methods and technologies used to enter a relationship and realize data sovereignty. In the second step, trust can be enhanced as soon as parties, such as data providers and consumers, establish contractual agreements via data-sharing infrastructure to accelerate business transactions (Yang et al., 2021). Thus, the basic trust required by data consumers and providers helps strengthen the overall trust in the data infrastructure that enforces the policies specified in the contract agreements. To make the argument of trust a core aspect for developing a more robust conceptual model, Munoz-Arcentales et al. (2019) stated: "*Trust.* It is the basis for all the relations between different organizations. Thus, being part of trusted environments is a key part of every operation, including data exchange. Data usage control is achieved thanks to this principle" (2019, p. 592), which makes it an essential component of data sovereignty.

Examples from the field

The model was evaluated by concrete examples from the field. Such real-life scenarios can demonstrate its usefulness and possible applications. One example stems from the German automotive industry and deals with data exchange in the supply chain. The case study, its requirements, and the results presented by Opriel et al. (2021) can be mapped to the core aspects. In their study, the data exchange occurs between an original equipment manufacturer (OEM) and a specific supplier (data provider, data consumer). They exchange industrial information on demand and capacity (data assets) at different stages (data value chain and lifecycle activities) based on such current standards as the Electronic Data Interchange (EDI) (data infrastructure). The researchers identified the need for trust and the possibilities of contractual agreements in their problem, barriers, and business requirements analysis: "[Data sovereignty] can foster trust in each other and reduce risks being affected in data breaches (P16) […]. In order to secure legal aspects, the system shall provide functionalities to link usage policies with contractual definitions (R16)" (Opriel et al., 2021, p. 436). Here, the instrument of data sovereignty is implemented to overcome trust issues originating from power imbalances between participating actors and, therefore, serves as an excellent example of agency theory's applicability.

Another concrete example explains the shared use of data in a network of enterprises. In its white paper, Plattform Industrie 4.0 (2022) demonstrated how the technical implementation of data sovereignty plays a crucial role in multilateral data sharing for Collaborative Condition Monitoring (CCM) between such participants as component suppliers and factory operators (data provider, data consumer). They share and use (data value chain and lifecycle activities) datasets, such as sensor data (data asset), to leverage data-driven business models via a decentralized, federated infrastructure (Plattform Industrie 4.0, 2022). Similar to the previous case, the core aspects of the conceptual model can be directly mapped to their results, as summarized in Table 3. Component suppliers, machine suppliers, and factory operators create legally binding concepts to ensure trust between each other. Moreover, this example showcases agency theory's relevance in this actor relationship and highlights that data sovereignty is a suitable instrument with which to overcome mistrust and weaken power imbalances, even if both actors have their own interests.

Discussion and future research opportunities

The presented conceptual model offers a new approach to understanding data sovereignty's implementation in IS research by considering adjacent domains. It contributes to the existing literature by laying the foundation for further research, as well as by filling the above-described research gaps of underlying conflicts and inconsistencies. The following discussion describes the study's practical and theoretical implications and addresses current limitations and future research opportunities.

This study's results lead to direct implications for practice, as they serve to guide and provide a mutual understanding of the concept for individuals and companies. It aims to help users technically implement data sovereignty, e.g., actors in research projects, organizations building data-sharing ecosystems, and stakeholders strengthening the role of data sovereignty through regulatory bodies. In addition, industry and research projects related to IDSA or Gaia-X can help to further communicate and develop this topic by designing systems based on data sovereignty principles. Additionally, individuals and society can play an enhanced role in demanding technology that implements data sovereignty for all data lifecycle stages by design, in line with European values. The conceptual model can further refine this vision and clarify communication.

For theory, this work's conceptual model can be seen as a necessary academic addition to ongoing discussions. The terminological ambiguity, viewpoints of current research streams, and existing definitions were brought together by defining and describing core aspects. We acknowledge that, in IS research, other models have sought to offer a mutual understanding of

Table 3 Examples mapped to the core conceptual data sovereignty aspects

Core aspects	Example 1: Demand and Capacity Management in the Automotive Supply Chain (Opriel et al., 2021)	Example 2: Collaborative Condition Monitoring of Industrial Assets (Plattform Industrie 4.0, 2022)
Data asset	Industrial information for demand and capacity	Datasets, such as sensor data
Data provider and data consumer	OEM (car manufacturer) and their Tier 1 supplier	Component supplier, machine supplier, and factory operator
Contract agreement	Currently used paper-based contracts should be replaced by usage policies linked to contract definitions.	Currently used bilateral contracts should be replaced by data licenses to specify usage restrictions.
Data value chain and data lifecycle activities	All lifecycle activities (from creation to sharing to deletion) with a focus on data exchange.	Data usage/data sharing
Infrastructure	Currently used manual data exchange and EDI standards should be extended by decentralized platforms.	Use of federated infrastructures on a cross-industry basis based on the Asset Administration Shell (IDTA, 2023)
Trust	Since trust is identified as a major attribute, it can be strengthened by the implementation of usage control.	Data space providers and operators must create legally binding concepts to ensure trust between participants.

data sovereignty. However, Ernstberger et al.'s (2023) model has a nearly exclusive technical layer perspective, while Zrenner et al.'s (2019) applies it in the manufacturing domain only, and the model of Otto et al. (2019) is a specific reference architecture. An additional theoretical impact arises from linking different research streams that describe the core conceptual aspects. To the best of our knowledge, some of these (e.g., the data value chain and trust) had not previously been contextualized in this manner, meaning that this study offers an approach with the potential to open up new perspectives. Due to its fundamentality, this IS research's theoretical contributions can be tested and applied in different research areas, e.g., with a legal or political focus.

Despite careful evaluation, this fundamentals study suffers from limitations as it could not cover all essential research strands. Nevertheless, these can provide input for future research opportunities according to different paradigms, namely design science research (DSR), which aims to develop artifacts addressing real-world problems (Hevner et al., 2004), and behavioral research focusing on why groups or individuals act in a specific way and how they can be influenced (Skinner, 1965). To theoretically ground the conceptual model, the agency theory approach was chosen. However, its limitation must be acknowledged, which include, for example, a closer relation to the area of IS, defined "as a system[s] in […] organization[s]" (Davis, 2000, p. 67). Moreover, the literature analyses have limitations since using different databases or searches could lead to different results.

In line with the DSR paradigm, further limitations are addressed in Table 4 and described in the following: First, future research should examine the necessary development of an artifact that supports individuals in controlling their data (*RQ#1*). Moreover, the implementation of data sovereignty according to the model for individuals is valid; nevertheless, its enforcement requires further attention. Research on the enforcement of data sovereignty for individuals exists (Lomotey et al., 2022). However, as this was outside the scope of this study, future research could explore which artifacts need to be developed to enhance individuals' ability to control their data. Additionally, the future development of the instrument of data sovereignty was not covered in this research. Therefore, identifying the capabilities needed to implement data sovereignty as an instrument is critical (*RQ#2*). Building on this, conducting design-oriented studies of maturity models to track and measure data sovereignty's implementation (*RQ#3*) could be a promising research direction. Furthermore, there is a need for IT artifacts in policy management and data spaces, as well as reference models and methods, to establish, develop, improve, and ensure data sovereignty in internal and external data management activities *(RQ#4)*, such as validation, enforcement, signing, watermarking, or data integrity concepts (Hellmeier et al., 2023).

In the context of this study's limitations, the behavioral research paradigm applies to various research opportunities. Due to this research's qualitative literature approach, subjectivity can be seen as a limitation. Even if examples from the field are mapped to the conceptual model (Opriel et al., 2021; Plattform Industrie 4.0, 2022), applying the model in practice, e.g., in the "common European data spaces" (Data Spaces Support Centre, 2023b, p. 5) would prove its utility in various data sharing projects (*RQ#5*). Additionally, this could help validate agency theory's application for reaching an overall understanding of implementing data sovereignty as an instrument. Moreover, the cost of such implementation has not been discussed in this research. The relationship between the value of data and data economics on the one hand, and data sovereignty on the other, acknowledging that data assets may vary in criticality and value, is an exciting research strand. Open questions have to be answered focusing on the maintenance costs of data infrastructure and standards for enforcement (*RQ#6*). Besides, data sovereignty is a prerequisite for enabling more data sharing (Azkan et al., 2022). This study has not explicitly analyzed whether data

Table 4 Summary of future research opportunities

Design science research		
Example	Research question	
Enforcing data sovereignty for individuals	What artifacts need to be developed to enhance individuals' ability to control their data?	*RQ#1*
Capabilities of data sovereignty	What is the design of a model for capabilities needed to implement data sovereignty?	*RQ#2*
Maturity model for data sovereignty	What is the design of a maturity model that measures data sovereignty?	*RQ#3*
Operationalization of data sovereignty	What artifacts need to be developed to support the operationalization of data sovereignty?	*RQ#4*
Behavioral research		
Example	Research question	
Application of the conceptual model to additional practical cases in data spaces	How do data spaces apply this conceptual model?	*RQ#5*
Relationship between the value of data and data economics on the one hand, and data sovereignty on the other	Who builds, maintains, and pays for the data infrastructure? Who sets the standards and enforces the rules?	*RQ#6*
More data sharing as an incentive for the implementation of data sovereignty	How does the implementation of data sovereignty affect data sharing?	*RQ#7*

sovereignty positively or negatively impacts data sharing, thus making it necessary to explore this aspect in the future and re-evaluate the topic's importance *(RQ#7)*.

Summary

As IS research on data sovereignty remains in its infancy, this study has included academic and practical literature in its investigation so as to determine a common understanding of the concept itself (see Fig. 1). As shown by the analysis of the current research stream, data sovereignty is not uniformly defined, with contrasting explanations and definitions having been offered. This fundamentals paper expands IS research's knowledge on data sovereignty by providing a conceptual model following agency theory and validated by documented real-world examples. It emphasizes the specification of the core aspects (derived from the literature) needed to implement data sovereignty. The technological implementation of data sovereignty is essential for guaranteeing trusted data sharing between individuals and organizations of different parties and make innovation happen. However, further practical and theoretical implications have yet to be uncovered, and future research must still evaluate and apply the proposed model.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12525-024-00693-4.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Azkan, C., Gür, I., Hupperz, M., Gelhaar, J., Gieß, A., Groß, T., Frings, S., Kett, H., Kutzias, D., Strauß, O., Büchel, J., Demary, V., Engels, B., Goecke, H., Mertens, A., Röhl, K.-H., Rusche, C., Scheufen, M., Schröder, B., & Valet, S. (2022). *Incentives and economics of data sharing: Fields of action of cross-company data exchange and status quo of the German economy*. https://ieds-projekt.de/wp-content/uploads/2022/08/IEDS-Whitepaper_Englisch.pdf. Accessed 12 Dec 2023

Banse, C. (2021). Data sovereignty in the cloud - Wishful thinking or reality? *Conference on Computer and Communications Security*, 153–154. https://doi.org/10.1145/3474123.3486792

Bärenfänger, R. (2017). *Managing information services in the digital economy*. Difo-Druck GmBh.

Botsman, R. (2017). *Who can you trust? How technology brought us together and why it might drive us apart* (First edition). Public Affairs.

Capiello, C., Gal, A., Jarke, M., & Rehof, J. (2020). *Data ecosystems: Sovereign data exchange among organizations* (Dagstuhl Seminar 19391), pp. 66–134. https://doi.org/10.4230/DagRep.9.9.66

Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New horizons for a data-driven economy*. Springer International Publishing. https://doi.org/10.1007/978-3-319-21569-3

Chander, A., Abraham, M., Chandy, S., Fang, Y., Park, D., & Yu, I. (2021). Achieving privacy. *SMU Law Review, 74*(4), 607–664.

Chapin, F. S., Matson, P. A., & Vitousek, P. M. (2011). *Principles of terrestrial ecosystem ecology*. Springer, New York. https://doi.org/10.1007/978-1-4419-9504-9

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage.

Curry, E. (2016). *The big data* value *chain: Definitions, concepts, and theoretical approaches*. https://doi.org/10.1007/978-3-319-21569-3_3

Data Spaces Support Centre. (2023a). *Blueprint Version 0.5* (No. 1.0). https://dssc.eu/space/BPE/179175433/Data+Spaces+Blueprint+%7C+Version+0.5+%7C+September+2023. Accessed 12 Dec 2023

Data Spaces Support Centre. (2023b). *Starter kit for data space designers* (No. 1.0). https://dssc.eu/space/SK/29523973/Starter+Kit+for+Data+Space+Designers+%7C+Version+1.0+%7C+March+2023. Accessed 12 Dec 2023

Davis, G. B. (2000). Information systems conceptual foundations: Looking backward and forward. In R. Baskerville, J. Stage, & J. I. DeGross (Eds.), *IFIP advances in information and communication technology. Organizational and social perspectives on information technology* (Vol. 41, pp. 61–82). US: Springer.

Docter, Q., & Fuchs, C. (Eds.). (2020). *CompTIA cloud essentials+ study guide*. Wiley. https://doi.org/10.1002/9781119642138

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *The Academy of Management Review, 14*(1), 57. https://doi.org/10.2307/258191

Ernstberger, J., Lauinger, J., Elsheimy, F., Zhou, L., Steinhorst, S., Canetti, R., Miller, A., Gervais, A., & Song, D. (2023). SoK: Data sovereignty. *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)* (pp. 122–143). IEEE. https://doi.org/10.1109/EuroSP57164.2023.00017

Esposito, C., Castiglione, A., & Choo, K.-K.R. (2016). Encryption-based solution for data sovereignty in federated clouds. *IEEE Cloud Computing, 3*(1), 12–17. https://doi.org/10.1109/MCC.2016.18

European Commission. (2020). *Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act)* (COM/2020/767 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0767. Accessed 12 Dec 2023

European Commission. (2022). *Proposal for a regulation of the european parliament and of the council on harmonised rules on fair access to and use of data (Data Act)* (COM/2022/68 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0068. Accessed 12 Dec 2023

Garousi, V., Felderer, M., & Mäntylä, M. V. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology, 106*, 101–121. https://doi.org/10.1016/j.infsof.2018.09.006

Gelhaar, J., Groß, T., & Otto, B. (2021). A taxonomy for data ecosystems. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 6113–6122. https://doi.org/10.24251/HICSS.2021.739

German Ethics Council. (2017). *Big Data and Health – Data Sovereignty as the Shaping of Informational Freedom: Executive Summary & Recommendations*. https://www.ethikrat.org/en/publications/publication-details/?tx_wwt3shop_detail%5Bproduct%5D=4&tx_wwt3shop_detail%5Baction%5D=index&tx_wwt3shop_detail%5Bcontroller%5D=Products&cHash=7bb9aadb656b877f9dbd49a61e39df2f. Accessed 12 Dec 2023

Ghazizadeh, E., & Sun, T. (2021). A systematic literature review of smart contract applications. In K. Arai, S. Kapoor, & R. Bhatia (Eds.), *Advances in intelligent systems and computing: Vol. 1290. Proceedings of the Future Technologies Conference (FTC) 2020, Volume 3* (Vol. 1290, pp. 877–888). Springer International Publishing. https://doi.org/10.1007/978-3-030-63092-8_59

Gil, G., Arnaiz, A., Diez, F. J., & Higuero, M. V. (2020). Evaluation methodology for distributed data usage control solutions. *2020 Global Internet of Things Summit (GIoTS)* (pp. 1–6). IEEE. https://doi.org/10.1109/GIOTS49054.2020.9119565

Grant, P. (1983). Technological sovereignty: Forgotten factor in the "Hi-Tech" Razzamatazz. *Prometheus, 1*(2), 239–270. https://doi.org/10.1080/08109028308628930

Hellmeier, M., Pampus, J., Qarawlus, H., & Howar, F. (2023). Implementing data sovereignty: Requirements & challenges from practice. *Proceedings of the 18th International Conference on Availability,* Reliability *and Security* (pp. 1–9). ACM. https://doi.org/10.1145/3600160.3604995

Hellmeier, M., & von Scherenberg, F. (2023). A delimitation of data sovereignty from digital and technological sovereignty. *ECIS 2023 Research Papers*. https://aisel.aisnet.org/ecis2023_rp/306. Accessed 12 Dec 2023

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105. https://doi.org/10.2307/25148625

Hinsley, F. H. (1986). *Sovereignty* (2. ed.). Cambridge University Press.

Hojati, M., Farmer, C., Feick, R., & Robertson, C. (2021). Decentralized geoprivacy: Leveraging social trust on the distributed web. *International Journal of Geographical Information Science, 35*(12), 2540–2566. https://doi.org/10.1080/13658816.2021.1931236

Hong, S., & Kim, H. (2020). VaultPoint: A blockchain-based SSI model that complies with OAuth 2.0. *Electronics, 9*(8), 1231. https://doi.org/10.3390/electronics9081231

Houy, C., Fettke, P., & Loos, P. (2012). Understanding understandability of conceptual models – What are we actually talking about? In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, P. Atzeni, D. Cheung, & S. Ram (Eds.), *Lecture Notes in Computer Science. Conceptual Modeling* (Vol. 7532, pp. 64–77). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34002-4_5

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society, 8*(1). https://doi.org/10.1177/2053951720982012

IDTA. (2023). *Specification of the asset administration shell - Part 1: Metamodel.* Industrial Digital Twin Association. https://industrialdigitaltwin.org/wp-content/uploads/2023/04/IDTA-01001-3-0_SpecificationAssetAdministrationShell_Part1_Metamodel.pdf. Accessed 12 Dec 2023

Irion, K. (2012). Government cloud computing and the policies of data sovereignty. *Policy and Internet*, *4*(3-4). https://doi.org/10.2139/ssrn.1935859

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258–268. https://doi.org/10.1080/10580530.2012.716740

Jarke, M., Otto, B., & Ram, S. (2019). Data sovereignty and data space ecosystems. *Business & Information Systems Engineering, 61*(5), 549–550. https://doi.org/10.1007/s12599-019-00614-2

Jentzsch, N. (2018). *Dateneingentum - Eine gute Idee für die Datenökonomie?* [Data ownership - A good idea for the data economy?]. https://www.stiftung-nv.de/de/publikation/dateneigentum-eine-gute-idee-fuer-die-datenoekonomie. Accessed 12 Dec 2023

Kuhrmann, M., Fernández, D. M., & Daneva, M. (2017). On the pragmatic design of literature studies in software engineering: An

experience-based guideline. *Empirical Software Engineering, 22*(6), 2852–2891. https://doi.org/10.1007/s10664-016-9492-y

Kushwaha, N., Roguski, P., & Watson, B. W. (2020). Up in the air: Ensuring government data sovereignty in the cloud. *2020 12th International Conference on Cyber Conflict (CyCon)* (pp. 43–61). IEEE. https://doi.org/10.23919/CyCon49761.2020.9131718

Labadie, Clément., & Legner, C. (2019). Understanding data protection regulations from a data management perspective: A capability-based approach to EU-GDPR. *14th International Conference on Wirtschaftsinformatik,*, 1292–1306. https://aisel.aisnet.org/wi2019/track11/papers/3/. Accessed 12 Dec 2023

Landolfi, G., Barni, A., Izzo, G., Fontana, A., & Bettoni, A. (2019). A MaaS platform architecture supporting data sovereignty in sustainability assessment of manufacturing systems. *Procedia Manufacturing, 38*(38), 548–555. https://doi.org/10.1016/j.promfg.2020.01.069

Lauf, F., Scheider, S., Meister, S., Radic, M., Herrmann, P., Schulze, M., Nemat, A. T., Becker, S. J., Rebbert, M., Abate, C., Konrad, R., Bartsch, J., Dehling, T., & Sunyaev, A. (2021). *Data sovereignty and data economy—Two repulsive forces?* https://doi.org/10.24406/isst-n-634865

Lomotey, R. K., Kumi, S., & Deters, R. (2022). Data trusts as a service: Providing a platform for multi-party data sharing. *International Journal of Information Management Data Insights, 2*(1), 100075. https://doi.org/10.1016/j.jjimei.2022.100075

Marfia, F., Fornara, N., & Nguyen, T.-V.T. (2017). A framework for managing data provider and data consumer semantic obligations for access control. *AI Communications, 30*(1), 67–82. https://doi.org/10.3233/AIC-170725

McKinney, E. H., & Yoos, C. J. (2010). Information about information: A taxonomy of views. *MIS Quarterly, 34*(2), 329. https://doi.org/10.2307/20721430

Munoz-Arcentales, A., López-Pernas, S., Pozo, A., Alonso, Á., Salvachúa, J., & Huecas, G. (2019). An architecture for providing data usage and access control in data sharing ecosystems. *Procedia Computer Science, 160*(160), 590–597. https://doi.org/10.1016/j.procs.2019.11.042

Nagel, L., & Lycklama, D. (2021). *Design principles for data spaces* - Position paper. https://doi.org/10.5281/zenodo.5105744

Oliveira, M. I. S., & Lóscio, B. F. (2018). What is a data ecosystem? In M. Janssen, S. A. Chun, V. Weerakkody, A. Zuiderwijk, & C. C. Hinnant (Eds.) *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (pp. 1–9). ACM. https://doi.org/10.1145/3209281.3209335

Opriel, S., Möller, F., Burkhardt, U., & Otto, B. (2021). Requirements for usage control based exchange of sensitive data in automotive supply chains. *Proceedings of the 54th Hawaii International Conference* on *System Sciences*, 431–440. https://doi.org/10.24251/HICSS.2021.051

Otto, B., ten Hompel, M., & Wrobel, S. (2022). Designing data spaces. *Springer International Publishing*. https://doi.org/10.1007/978-3-030-93975-5

Otto, B., Steinbuss, S., Teuscher, A., & Lohmann, S. (2019). *Ids Reference Architecture Model* (No. 3.0). https://doi.org/10.5281/ZENODO.5105529

Peterson, Z. N. J., Gondree, M., & Beverly, R. (2011). A position paper on data sovereignty: The importance of geolocating data in the cloud. *Proceedings of the 3rd USENIX Conference on* Hot *Topics in Cloud Computing*. https://dl.acm.org/doi/10.5555/2170444.2170453

Plattform Industrie 4.0. (2022). *Multilateral data sharing in industry: Concept using "Collaborative Condition Monitoring" as a basis for new business models*. https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Multilateral_Data_Sharing.pdf. Accessed 12 Dec 2023

Pohle, J., & Thiel, T. (2020). Digital sovereignty. *Internet Policy Review*, *9*(4). https://doi.org/10.14763/2020.4.1532

Polatin-Reuben, D., & Wright, J. (2014). An Internet with BRICS characteristics: Data sovereignty and the balkanisation of the Internet. *4th USENIX Workshop on Free and Open Communications on the Internet*. https://www.usenix.org/conference/foci14/workshop-program/presentation/polatin-reuben. Accessed 12 Dec 2023

Qarawlus, H., Hellmeier, M., Pieperbeck, J., Quensel, R., Biehs, S., & Peschke, M. (2021). Sovereign data exchange in cloud-connected IoT using international data spaces. *2021 IEEE Cloud Summit (Cloud Summit)* (pp. 13–18). IEEE. https://doi.org/10.1109/IEEECloudSummit52029.2021.00010

Rahul, K., & Banyal, R. K. (2020). Data life cycle management in big data analytics. *Procedia Computer Science, 173*, 364–371. https://doi.org/10.1016/j.procs.2020.06.042

Sarabia-Jacome, D., Lacalle, I., Palau, C. E., & Esteve, M. (2019). Enabling industrial data space architecture for seaport scenario. *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)* (pp. 101–106). IEEE. https://doi.org/10.1109/WF-IoT.2019.8767216

Schilke, O., & Cook, K. S. (2013). A cross-level process theory of trust development in interorganizational relationships. *Strategic Organization, 11*(3), 281–303. https://doi.org/10.1177/1476127012472096

Schinle, M., Erler, C., & Stork, W. (2021). Data sovereignty in data donation cycles - Requirements and enabling technologies for the data-driven development of health applications. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 3972–3981. https://doi.org/10.24251/HICSS.2021.482

Singi, K., Choudhury, S. G., Kaulgud, V., Bose, R. J. C., Podder, S., & Burden, A. P. (2020). Data sovereignty governance framework. *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops* (pp. 303–306). ACM. https://doi.org/10.1145/3387940.3392212

Skinner, B. F. (1965). *Science and human behavior*. New York, NY: The Free Press.

Statista. (2022). *Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2030*. https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

Tan, K.-L., Chi, C.-H., & Lam, K.-Y. (2022). *Analysis of digital sovereignty and identity: From digitization to digitalization.* https://doi.org/10.48550/arXiv.2202.10069

Taylor, J., & Kukutai, T. (Eds.). (2016). *Research monograph / Centre for Aboriginal Economic Policy Research, College of Arts and Social Sciences, The Australian National University, Canberra: no. 38. Indigenous data sovereignty: Toward an agenda*. Australian National University Press. http://www.jstor.org/stable/10.2307/j.ctt1q1crgf

van den Homberg, M., & Susha, I. (2018). Characterizing data ecosystems to support official statisticswith open mapping data for reporting on sustainable development goals. *ISPRS International Journal of Geo-Information, 7*(12), 456. https://doi.org/10.3390/ijgi7120456

Verhulst, S. G. (2023). Operationalizing digital self-determination. *Data & Policy, 5*, e14. https://doi.org/10.1017/dap.2023.11

Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems, 29*(3), 129–147. https://doi.org/10.1080/12460125.2020.1798591

Yang, R., Liu, N., Pang, Z., Wang, Y., Jia, Q., Lu, W., Li, Z., Li, M., & Wu, L. (2021). The next generation identity platform for digital era based on blockchain. *Lecturer Notes in Electrical Engineering, 677*(677), 1035–1044. https://doi.org/10.1007/978-981-33-4102-9_124

Zrenner, J., Möller, F. O., Jung, C., Eitel, A., & Otto, B. (2019). Usage control architecture options for data sovereignty in business ecosystems. *Journal of Enterprise Information Management, 32*(3), 477–495. https://doi.org/10.1108/JEIM-03-2018-0058

Paper IV

Table A.4 Metadata Overview of Paper IV

Title	A Hidden Digital Text Watermarking Method Using Unicode Whitespace Replacement
Authors	Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Haydar Qarawlus *Fraunhofer ISST, Dortmund, Germany* Hendrik Norkowski *Fraunhofer ISST, Dortmund, Germany* Falk Howar *TU Dortmund & Fraunhofer ISST, Dortmund, Germany*
Publication Year	2025
Publication Type	Conference
Conference Name	58th Hawaii International Conference on System Sciences (HICSS)
Conference Location	Waikoloa Village, Big Island, Hawaii, USA
Conference Date	07. January 2025 - 10. January 2025
Publisher / Database	AIS Affiliated
DOI / Link	https://doi.org/10.24251/hicss.2025.886
Status	Published
Ranking	VHB: B (2024 Rating) CORE: A (2018 Rating)[2] ERA: A (2010 Rating)
Comment	-

[2] The HICSS conference was last ranked by CORE in 2018 (see https://portal.core.edu.au/conf-ranks/575/, accessed on Feb. 10, 2026). It is not listed in the latest ICORE 2026 rating.

Proceedings of the 58th Hawaii International Conference on System Sciences | 2025

A Hidden Digital Text Watermarking Method Using Unicode Whitespace Replacement

Malte Hellmeier
Fraunhofer ISST
malte.hellmeier@isst.fraunhofer.de

Haydar Qarawlus
Fraunhofer ISST
haydar.qarawlus@isst.fraunhofer.de

Hendrik Norkowski
Fraunhofer ISST
hendrik.norkowski@isst.fraunhofer.de

Falk Howar
TU Dortmund & Fraunhofer ISST
falk.howar@tu-dortmund.de

Abstract

The possibility of hiding information inside a digital medium is often referred to as watermarking or steganography. Since various solutions for image, video, and audio files exist, keeping control over text is challenging due to its limited possibilities. In this paper, we present a new digital text watermarking algorithm to hide a byte-encoded sequence inside an unformatted text. By substituting conventional whitespaces with a set of five similar-looking Unicode spaces, the cover text's structure and length stay untouched while remaining imperceptible to humans. We propose a software design and proof-of-concept multiplatform implementation with a downstream experimental evaluation for robustness, capacity, and visibility. Our findings indicate a stronger concealment and application robustness with limited embedding capacity compared to existing solutions utilizing zero-width spaces.

Keywords: Digital Watermarking, Information Hiding, Homoglyphs, Unicode, UTF-8

1. Introduction

Due to the increasing growth of digitization in organizations and the private sector, more and more processes and assets have shifted from the analog to the digital world. The spectrum spans from e-books as an alternative for printed books in the private sector to production documents sent via E-Mail or specific system endpoints for companies instead of being printed out and sent by post. Keeping control over data, often guided by principles of data sovereignty (Hellmeier & von Scherenberg, 2023; Jarke et al., 2019), is crucial to prevent losing control over it. Referring to the first example, "[p]irating of e-books is particularly rampant as they are already digital. It is even easier (and cheaper) to copy and electronically distribute e-books [...]" (Davis & Kazi, 2021, p. 20). Taking the automotive industry as another example, Original Equipment Manufacturers (OEMs) develop components for vehicles that external suppliers construct since they cannot produce all parts in-house independently. The OEMs needs to protect such sensitive documents and only shares them carefully with their suppliers and customers out of concerns of unauthorized reproduction or reverse engineering. Therefore, staying self-determined with data after it leaves its own control boundaries is important but still challenging (Hellmeier et al., 2023).

One possible solution in this area of tension for e-books (Davis & Kazi, 2021) and company data (Hellmeier et al., 2023) lies in the domain of digital watermarking. Reviewing the current landscape of research publications and existing methods, including adjacent fields like steganography, shows a majority of solutions for specific types like images, video, or audio files. Besides it, text is classified as one of the most shared content types (Bertini et al., 2019). Especially different sorts of documents, which are considered industry standards, ranging from Microsoft Word-based specification documents to emails, are based on text (Rizzo et al., 2019). However, digital text watermarking is also described as an underresearched field with only a few proposed techniques (Ahvanooey et al., 2018), mainly due to the missing noise tolerance in text (Bertini et al., 2019).

Current approaches, described in detail in Section 2.2, are often based on changing or generating the cover text semantic, structure, or layout (Askari et al., 2023). Moreover, different whitespace solutions

URI: https://hdl.handle.net/10125/109738
978-0-9981331-8-8

are adding additional spaces to the document (Por et al., 2008, 2012), including non-visible zero-width spaces (ZWSPs) (Ahvanooey et al., 2018, 2020). However, all existing methods are either not robust in different applications or visible to humans due to cover text changes, unfamiliar depicted characters, or increased content sizes. Therefore, existing solutions are unsuitable for the examples of e-books and automotive specification documents since humans can notice them due to the increased content size or the watermark gets destroyed when copied into external applications.

To close this gap, we present a new digital watermarking method for unformatted text documents. It differs from state-of-the-art work by watermarking a cover document without increasing the number of characters or changing the semantics of the original document while keeping visibility abnormalities to a minimum. With the help of different whitespaces defined in the Unicode standard (The Unicode Consortium, 2023), the homoglyph-based approach replaces conventional spaces in the cover document with its supplements to embed a byte-encoded watermark inside the cover. Our implemented prototype helps to demonstrate the actual usability and is used in our preliminary experimental evaluation. To sum up, the following concrete contributions structure our work:

(i) We show the current landscape and an evaluation of existing solutions to identify the research gap in Section 2.

(ii) We introduce an invisible watermark embedding and extraction algorithm with implemented examples for text-based documents in Section 3.

(iii) We evaluate, discuss, and summarize our proposed solution with its limitations and future research opportunities by analyzing its robustness, capacity, and visibility in an experimental evaluation in Section 4, 5 and 6.

2. Background

In the following, digital text watermarking with its relation to cryptography, information hiding, and steganography is introduced to create a uniform understanding. Similar algorithms discussed in the literature are introduced in Related Work, including a delimitation to the present work.

2.1. Digital Text Watermarking

Distinguishing between different terminologies in the domain of *cryptography* and *information hiding* is crucial and has multifacetedly been discussed in literature (Ahvanooey et al., 2018, 2022; Petitcolas et al., 1999; Rizzo et al., 2016, 2019). Starting with the broadest concept: "Information hiding is the science of concealing a secret message or watermark inside a cover media (a host file/message) for providing various security purposes such as content authentication, integrity verification, covert communication, and so on" (Ahvanooey et al., 2022, p. 56). Since information hiding can further be divided into steganography and watermarking, as shown in Figure 1, it differs from cryptography because its root is not about transforming plain text into an encrypted cipher text (Ahvanooey et al., 2018). A detailed survey and classification is published by Petitcolas et al. (1999).

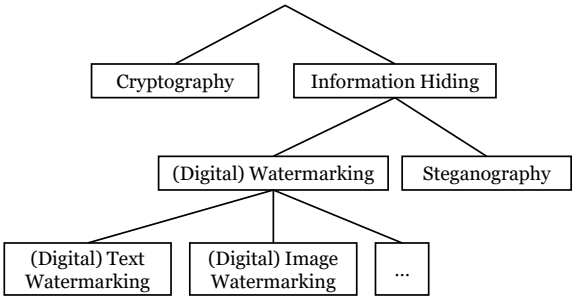


Figure 1. Correlation Between Terminologies.

Methods aiming to exchange secret information in a hidden way are categorized as *steganography*. They can be classified into character-level, bit-level, or a mixture of both categories (Krishnan et al., 2017), leading to the following definition: "Steganography embeds a secret message inside an innocent looking cover medium, stealthily, without creating any attention. The cover medium used can be a text, image, audio, video, network packets, etc." (Krishnan et al., 2017, p. 1).

In contrast, *digital watermarking* focuses on "a visible or an invisible, preferably the latter, identification code that permanently is embedded in the data" (Jalil & Mirza, 2009, p. 230). Those data assets can range from images, audio, and video to the underresearched and most difficult cover medium of plain text (Bertini et al., 2019), often referred to as digital watermarking, text watermarking, or digital text watermarking. An overview of different definitions arisen over time is presented in Table 1.

To bring everything together, Rizzo et al. summarize the concepts as follows: "While cryptography algorithms make unreadable the information by applying a kind of permutation or substitution to the original content, the steganography algorithms provide techniques to hide new information into the carrier,

Table 1. Digital Watermarking Definitions.

#	Definition
I	"A digital watermark can be described as a visible or an invisible, preferably the latter, identification code that permanently is embedded in the data." (Jalil & Mirza, 2009, p. 230)
II	"In digital watermarking, relevant information is embedded in an imperceptive way into a digital document. The embedded information is called a watermark." (Qadir & Ahmad, 2006, p. 19)
III	"[D]igital watermarking technology is a typical information hiding method, which covers text, image and video." (Qi et al., 2023, p. 1311)

that is a readable document. Whereas watermarking algorithms ensure the authentication and the copyright protection by applying a watermark to the digital content" (Rizzo et al., 2016, p. 97). An overview of the interrelationships of the terms is presented in Figure 1.

2.2. Related Work

Over the years, different digital text watermarking algorithms have been presented. Various literature reviews, comparisons, and overviews exist, analyzing the research area, including adjacent fields like steganography (Agarwal, 2013; Ahvanooey et al., 2019, 2022). We discuss related solutions and their delimitation to this work by focusing on approaches modifying the cover document to hide text. Further, related approaches like homoglyph-based attacks are considered.

Por et al. (2008) presented a WhiteSteg algorithm replacing whitespaces between words with one or two spaces to represent a zero or a one for watermark insertion. Further, tab characters are added at the end of paragraphs and sentences. Here, humans can easily recognize a watermarked text due to the different spaces between words.

Following this initial approach, Por et al. (2012) described a new UniSpaCh data hiding method. By splitting a text into two groups, the first group replaces inter-word and inter-sentence spaces with two whitespaces: A classical whitespace and one of three whitespaces with a small width. The second group adds whitespaces for end-of-line and inter-paragraph spacing, including alternative whitespaces with different widths. Due to additional symbols added, the proposed algorithm increases the number of characters in a cover document.

A ZWSP solution, proposed by Ahvanooey et al. (2018), converts a secret message into a protected hidden message using a symmetric key for encoding and decoding. Afterward, the binary message is transformed into one containing only ZWSPs representing two Bits each. This secret message is placed before the first character of the message, leaving the cover text itself untouched. Similar to Por et al. and other ZWSP approaches (Ahvanooey et al., 2020; Por et al., 2012), the solutions increase the number of characters.

Rizzo et al. (2016, 2019) introduced a homoglyph-based watermarking technique using replacement. The hash function SipHash uses a text and password to generate a 64-bit long watermark. It is embedded into a cover text by using and replacing similar-looking Unicode characters. A similar method is described by Shazzad-Ur-Rahman et al. (2023) using AES and six hidden Bits for every whitespace. However, the proposed solutions are noticeable by human eyes and not robust when the watermark text is copied into another system not supporting such symbols like emails as identified in downstream analyses (Ahvanooey et al., 2022).

Such homoglyph-based approaches are also known in the field of phishing attacks with different DNS or domain squatting techniques (Kintis et al., 2017). "[A] homograph is a letter or string that has enough of a visual similarity to a different letter or string that the two may be confused for one another" (Holgers et al., 2006, p. 261). These homograph-based squatting attacks use homographs in widely used domain names to remain unnoticed (Holgers et al., 2006; Kintis et al., 2017). They work in URLs but differ from the solution presented in this work since they are either not robust in other applications or visible to humans outside of browsers.

Other solutions presented in the literature focused on font color (Askari et al., 2023), font glyphs (Xiao et al., 2018), other types with different covers like HTML (Mir, 2014), PDF (Khosravi et al., 2019), specialized Chinese characters (Wang et al., 2009), or bitmap images converted to bit strings as watermarks (Sonnleitner, 2012). Nevertheless, these solutions are either recognizable by human eyes or increase the number of characters on the cover document.

The present research aims to close this gap by introducing a novel digital watermarking approach addressing the issues. Unlike previous research, this is, to our knowledge, the first watermarking approach directly in the cover document that (1) is not recognizable by human eyes (2) without raising the number of characters, while (3) being robust in all common widely used business applications.

3. Proposed Solution

In the following, our proposed solution for digital text watermarking is described in detail by distinguishing between the embedding and extraction process of the watermark in the cover text. Further, details about the implementation as a multiplatform library with two usage examples are provided.

3.1. Watermark Embedding

The proposed embedding method is able to embed a watermark in a Unicode-encoded cover text CT. We define every Unicode symbol as u that is included in the set $\mathcal{U} := \{u : u \text{ is Unicode character}\}$, containing 149 813 characters based on Unicode standard version 15.1 (The Unicode Consortium, 2023). Following the standard, different whitespaces s exist, where the set of all Unicode space characters $\mathcal{S} := \{s : s \text{ is space character} \wedge s \in U\}$ contains 17 elements. We define the classical and most used space character U+0020 as δ with $\delta \in \mathcal{S}$. We evaluated every s as shown in Table 2 to define our own subset of whitespaces as alphabet $\mathcal{A}_+ := \{a : a \in S \wedge a \in U \wedge a \text{ meets criteria}\}$, whereas the criteria are non-noticeability for humans and robustness in different applications and file formats, described in detail in Section 4. $\mathcal{A}_+$ contains five elements shown by the bold-formatted whitespaces in Table 2 and we can summarize that $\mathcal{A}_+ \subset \mathcal{S} \subset \mathcal{U}$. The watermark itself is encoded through four of the five elements of $\mathcal{A}_+$, because one element is used as a separator character $\phi \in A_+$. We define the watermarker alphabet without separator character as $\mathcal{A}_-$, where $\mathcal{A}_+ = \mathcal{A}_- \cup \{\phi\}$ and $\phi \notin \mathcal{A}_-$.

For the description of the embedding method, we base our nomenclature on related work (Ahvanooey et al., 2020) while all elements start at index one. The embedding function $Emb(CT, \mathcal{W})$ begins by transforming the watermark $\mathcal{W}$ into a hidden watermark $\mathcal{W}_H$ by mapping it to the whitespace homoglyph alphabet $\mathcal{A}_-$. Since $|\mathcal{A}_-| = 4$, where $|\cdot|$ denotes the cardinality, every byte of the watermark $w \in \mathcal{W}$ is represented by four elements of $\mathcal{A}_-$ because:

$$\left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil = 4 \tag{1}$$

To fully include $\mathcal{W}_H$ in a cover text CT, the number of normal space characters δ in CT must be at least equal to the number of elements of the hidden watermark $\mathcal{W}_H$:

$$|\{x \in CT : x \in \delta\}| \geq |\mathcal{W}_H| \tag{2}$$

The final creation of the watermarked cover text $CT_\mathcal{W}$ is made by replacing all δ successively with the elements of $\mathcal{W}_H$. If CT has more δ than $|\mathcal{W}_H|$, the next element is replaced with the separator character ϕ, and the insertion process starts again until all δ are replaced to include the watermark multiple times. Since the algorithm replaces all spaces to achieve better robustness on modification attacks on specific parts, the resulting watermarked cover text does not contain normal space characters:

$$\forall x \in CT_\mathcal{W} : x \in \mathcal{U} \wedge x \neq \delta \tag{3}$$

The full embedding algorithm $Emb(CT, \mathcal{W}) = CT_\mathcal{W}$ is shown in Algorithm 1.

Algorithm 1: Watermark Embedding Algorithm $Emb(CT, W)$.

Data: $CT \leftarrow$ Cover text $CT := \{c_1, c_2, \ldots, c_n\}$, $\forall c \in \mathcal{U}$
Data: $\mathcal{W} \leftarrow$ Watermark $\mathcal{W} := \{w_1, w_2, \ldots, w_n\}$, $\forall w \in \{0, 1, \ldots, 255\}$
Result: $CT_\mathcal{W} \leftarrow$ Watermarked cover text

```
// Encode watermark
d ← ⌈log_2 2^8 / log_2 |A_-|⌉;
foreach w ∈ W do
    for i ← 1 to d do
        j ← w mod |A_-|;
        w ← ⌊w / |A_-|⌋;
        W_H ← W_H + a_{j+1};        // a_{j+1} ∈ A_-
// Insert watermark
i ← 1;
foreach c ∈ CT do
    if c = δ then
        if i <= |W_H| then
            w_{H_i} ← W_{H_i};      // w_{H_i} ∈ A_-
            CT_W ← CT_W + w_{H_i};
            i ← i + 1;
        else
            CT_W ← CT_W + φ;
            i ← 1;
    else
        CT_W ← CT_W + c
return CT_W
```

An example of the proposed watermark embedding algorithm is shown in Figure 2. The cover text $CT =$ "Lorem ipsum [...]" should be watermarked with the text $\mathcal{T} =$ "OEM1". First, every character of $\mathcal{T}$ is encoded into $\mathcal{W}$. In this case, the first letter "O" of the watermark is represented by the UTF-8 Hexadecimal value "4F" (U+004F), which equals the decimal value 79. Afterward, every character of the watermark $\mathcal{W}$ is encoded into $\mathcal{W}_H$ by transforming it into the alphabet $\mathcal{A}_-$ with a loop-based modulo operation as described in Algorithm 1. Next, every whitespace δ of CT is replaced with the corresponding space in $\mathcal{W}_H$. Since CT has more whitespaces than needed, the separator char ϕ is added, shown by the black U+2004 in Figure 2.

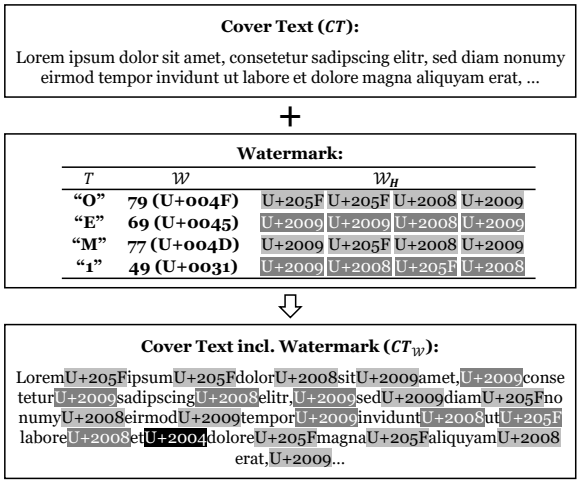


Figure 2. Watermark Embedding Example.

Since the insertion process starts again, the algorithm inserts the first character "O" of the watermark a second time and stops after all whitespaces are replaced.

3.2. Watermark Extraction

In accordance with the previously introduced naming (Ahvanooey et al., 2020) and elements in lists starting with index one, the extraction method $Ext(CT_{\mathcal{W}}) = \mathcal{W}$ is split into two parts for watermark extraction and decoding. The first part starts by iterating over all characters of the input watermarked cover text $CT_{\mathcal{W}}$ until it finds the first occurrence of ϕ as the separator character. Through the filtering of $\mathcal{A}_+$, it allows to extract the hidden watermark $\mathcal{W}_H$ from the cover text.

The second part decodes the hidden watermark $\mathcal{W}_H$ into its byte representation $\mathcal{W}$. The step size for the decoding part depends on the length of the watermarker alphabet without separator character and is defined as

$$d := \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil \tag{4}$$

with $d = 4$ for our alphabet introduced in Section 3.1, because each byte is represented by four whitespaces of $\mathcal{A}_-$. The cascading modulo operation from Algorithm 1 can be transformed back into its byte representation b. All b form the watermark $\mathcal{W}$, which in turn can be converted by the UTF-8 representation into the decoded text $\mathcal{T}$. The overall extraction process is summarized in Algorithm 2.

Algorithm 2: Watermark Extraction Algorithm $Ext(CT_{\mathcal{W}})$.

Data: $CT_{\mathcal{W}} \leftarrow$ Watermarked cover text
Result: $\mathcal{W} \leftarrow$ Extracted watermark
// Extract watermark
foreach $c \in CT_{\mathcal{W}}$ **do**
  if $c \in \mathcal{A}_+$ **then**
    if $c = \phi$ **then**
      break
    else
      $\mathcal{W}_H \leftarrow \mathcal{W}_H + c$
// Decode watermark
$d \leftarrow \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil$;
for $i \leftarrow 0$ **to** $|\mathcal{W}_H|$ *step* d **do**
  for $y \leftarrow 0$ **to** $d - 1$ **do**
    $a_k \leftarrow \mathcal{W}_{H_{i+y+1}}$; // $a_k \in \mathcal{A}_-$
    $b \leftarrow b + (k-1) \cdot d^y$; // $k \in [1, \ldots, d]$
  $\mathcal{W} \leftarrow \mathcal{W} + b$;
return $\mathcal{W}$

3.3. Implementation

Besides the theoretical presentation, both watermark embedding and extraction methods were technically implemented as a generic library in the Kotlin programming language to test and validate our solution. Kotlin was chosen since it is interoperable with the widely used Java programming language while supporting multiplatform targets. Thus, our solution's implemented version can be used in applications supporting the Java Virtual Machine (JVM) and in JavaScript solutions due to both build targets.

To test the library, we developed a Command Line Interface (CLI) tool for the JVM that can embed and extract a byte-encoded String into another String or text-based document, like a *.txt* file. Besides this, a webinterface is implemented as a second usage example for the JavaScript build target. This frontend acts as a graphical user interface and is likewise able to embed and extract watermarks in cover texts.

The source code of the watermarking library with its two usage examples of a CLI tool and a webinterface is made available[1] to ensure full transparency and applicability.

4. Experimental Evaluation

Different criteria must be considered when building watermarking or steganography methods. Therefore, this section builds on existing benchmarks to analyze and evaluate the proposed solution for its robustness, capacity, and visibility (Ahvanooey et al., 2018). For

[1] https://github.com/FraunhoferISST/TREND

each benchmark, the proposed solution is compared against the whitespace-based approach *UniSpaCh* by Por et al. (2012) and the homoglyph-based *Finegrain* approach by Rizzo et al. (2019).

4.1. Robustness Analysis

When considering the algorithm's robustness, different modification attacks of a watermark text and the robustness in different applications through copy and paste operations must be analyzed and compared to related work.

Modification Attacks. The watermarked text $CT_{\mathcal{W}}$ can be changed intentionally by a malicious attacker or unintentionally by an uninvolved third party. This includes the most significant alteration attacks through modification, deletion, and insertion attacks (Sonnleitner, 2012). Since modifications relate to text changes, all other modification options are always a combination of insertion and deletion, analyzed in the following.

Starting with *insertions*, adding additional characters to $CT_{\mathcal{W}}$ does not break the watermark if they are not included in $\mathcal{A}_+$. Due to the selection of the whitespace homoglyph alphabet presented in Table 2, the proposed characters are special characters of the Unicode standard. In conjunction with the problem domain in the automotive industry motivated in the introduction, unintentional insertions by employees of a company will not break the watermark since the characters of $\mathcal{A}_+$ are not included in standard industrial documents like specifications or emails and can not be inserted by mistake, as they are not present on classic computer keyboards. Nevertheless, attackers aware of $\mathcal{A}_+$ can consciously use these characters' insertion or random replacement to break the watermark.

For *deletion* attacks, the watermark could get destroyed, depending on the amount and type of deleted bytes. Removing any amount of Unicode characters $\mathcal{U} \setminus \mathcal{A}_+$ will not break the watermark since the extraction process only searches for elements of $\mathcal{A}_+$. Since the watermark is added multiple times in CT depending on the length of CT and W as discussed in the capacity analysis, the watermark only gets entirely destroyed if whitespaces of every repetition are removed with specialized text normalizations. Similar to the insertion attacks, malicious attackers can use it to destroy the watermark.

Considering *formatting* attacks, they are not relevant for plain text files since they can not be formatted with different font sizes, types, or colors. Therefore, attacks that are successful in related work like tampering (Ahvanooey et al., 2022) do not work with the proposed solutions because it is built upon the characters themselves instead of their formations.

Compared to existing solutions, our proposed method and the Finegrain approach have similar robustness against modification attacks since both solutions repeat the embedding process to insert the watermark multiple times (Rizzo et al., 2019). Since UniSpaCh embeds a significant amount in inter-paragraph spaces (Por et al., 2012), modifying a space between two paragraphs has a high probability of destroying the watermark, leading to more insufficient robustness in direct comparison.

Application Robustness. Our method uses five different whitespace characters in $\mathcal{A}_+$ to encode the watermark, shown by the bold entries in Table 2. They were selected based on an upstream whitespace analysis to stay robust when copied into another application.

All space characters defined by the Unicode Consortium are tested in different tools and file formats, presented in Table 2. Only official space characters with a width > 0 are tested because they are possible candidates for the whitespace homoglyph alphabet $\mathcal{A}_+$. A character like the ZWSP (U+200B) "although called a "space" in its name, does not actually have any width or visible glyph in display [...] and is treated as a format control character, rather than as a space character" (The Unicode Consortium, 2023, p. 267) and therefore excluded from our evaluation.

In our analysis, we checked the *visibility* of the whitespace characters by analyzing their width based on Korpela (2002) in comparison to the commonly used space U+0020 ($\approx \frac{1}{4}$ em). If abnormalities exist causing unusual space, we classify it as different visibility compared to the U+0020 space, depicted as "✗" in Table 2. When the difference is not noticeable, a "✓" is shown ($\approx \frac{1}{3}$ to $\frac{1}{5}$ em) while a "(✗)" indicates that the difference could be noticeable by human eyes (like $\frac{1}{2}$ or $\frac{1}{6}$ em).

Next, we analyzed the robustness in different *applications* and file types, namely plain *.txt* text files tested with the Microsoft Windows default notepad editor, Word's *.docx* format as one example for the Microsoft Office software stack, and the exported version as *.pdf*. To validate the operation system independence, the *.txt* file is also checked with the default editors on Ubuntu 22.04 (gedit) and macOS 13 (TextEdit). Besides the file types, we checked the spaces with mailing over different mail servers displayed with Microsoft Outlook and Thunderbird and the communication and meeting software Microsoft Teams. The selection of file types and programs was

Table 2. Whitespace Evaluation based on Korpela (2002).

Name	Code	Visibility	.txt	.docx	.pdf	Mail	Microsoft Teams
Space	U+0020	✓	✓	✓	✓	✓	✓
No-break Space	U+00A0	✓	✓	✗	✗	✗	✗
Ogham Space Mark	U+1680	✗	✓	✓	✓	✓	✓
En Quad	U+2000	✗	✓	✓	(✓)	✓	✓
Em Quad	U+2001	✗	✓	✓	(✓)	✓	✓
En Space	U+2002	✗	✓	✓	✗	✓	✓
Em Space	U+2003	✗	✓	✓	✗	✓	✓
Three-per-em Space	U+2004	✓	✓	✓	(✓)	✓	✓
Four-per-em Space	U+2005	✓	✓	✗	✗	✗	✓
Six-per-em Space	U+2006	(✗)	✓	✓	(✓)	✓	✓
Figure Space	U+2007	✗	✓	✓	(✓)	✓	✓
Punctuation Space	U+2008	✓	✓	✓	(✓)	✓	✓
Thin Space	U+2009	✓	✓	✓	(✓)	✓	✓
Hair Space	U+200A	(✗)	✓	✓	(✓)	✓	✓
Narrow No-break Space	U+202F	✓	✓	✓	(✓)	✓	✓
Medium Mathematical Space	U+205F	✓	✓	✓	(✓)	✓	✓
Ideographic Space	U+3000	✗	✓	✓	✗	✓	✓

deliberate, as these are considered industry standards for office and collaboration tools in many branches, except phone calls and SMS (DataReportal et al., 2023). This fits the initial introduced use cases for shared e-books in the private sector and specification documents in the automotive industry.

For the *.pdf* format, our analysis showed differences depending on the PDF viewer. For example, the En Quad (U+2000) is replaced with a standard U+0020 space when copying the content out of Adobe Acrobat Reader, while it stays persistent with PDF24 Reader. In such cases, a "(✓)" is shown in Table 2 because the respective whitespace characters remain in the original PDF file, but the robustness depends on the PDF viewer used.

To sum up, only the Three-per-em Space (U+2004), the Punctuation Space (U+2008), the Thin Space (U+2009), the Narrow-no-break Space (U+202F) and the Medium Mathematical Space (U+205F) are not noticed by humans and are robust in most of our tested applications and file formats and therefore form the whitespace homoglyph alphabet $\mathcal{A}_+$.

Compared to existing solutions, UniSpaCh has similar application robustness because it uses Six-per-em Space, Punctuation Space, Thin Space, and Hair Space (Por et al., 2012). These characters stay robust in our tested file formats and applications, as seen in Table 2. Based on our analysis, the Four-per-em Space gets replaced, e.g., in Word documents or emails. Since this space is used in the Finegrain approach (Rizzo et al., 2019), it has the worst application robustness.

4.2. Capacity Analysis

Our proposed solution transforms every one-byte character of a text-based watermark into four whitespaces of $\mathcal{A}_-$. This necessitates a minimum number of whitespace characters varying depending on the length of the desired watermark that can be calculated as follows:

$$|\delta|_{min} = |W| \times 4 \tag{5}$$

This represents the minimum number of space characters needed to embed the watermark in the text at least once, while a bigger cover text leads to further included copies of the watermark. The capacity can be illustrated by watermarking a tweet on the platform X (formally called Twitter). By using the maximum tweet length of 280 characters with an English text with an average word length of 4.79 characters (Norvig, 2012), followed by a space, the proposed solution is able to include a text-based watermark of around 12 characters inside it.

Compared to existing solutions, UniSpaCh has a higher embedding capacity because it uses end-of-line and inter-paragraph spacings beside inter-sentence and inter-word spaces only. To exemplify the difference, $\sim$574 bits can be inserted in an empty line between two paragraphs (Por et al., 2012). Rizzo et al. (2019) analyzed their Finegrain approach and identified a higher embedding capacity than UniSpaCh based on tests with New York Times newspaper paragraphs. Nevertheless, the embedding capacities of all three methods highly depend on the cover text.

4.3. Visibility Analysis

Abnormalities in visibility can be detected visually by a user when they notice something unusual or technically by checking the length or number of characters in a watermarked text.

In our proposed solution, the format and number of characters stay the same due to the pure replacement without adding additional characters, thus $|CT| = |CT_{\mathcal{W}}|$. Therefore, visibility is not noticed visually by a user or technically when comparing the number of characters. Nevertheless, depending on the displayed version, a technical length analysis can detect the slight width differences of some whitespaces in $\mathcal{A}_+$.

Compared to existing approaches, a watermarked text by UniSpaCh (Por et al., 2012) significantly increases the number of characters and thus becomes noticeable technically by comparing character counts. Users can also differentiate a watermarked text due to unusually large distances between words. Compared to the Finegrain approach by Rizzo et al. (2019), users notice differences in character representations and unusual distances between words.

5. Discussion

To our knowledge, the digital watermarking method introduced in this work provides, for the first time, a solution that can hide content inside text documents without increasing the document's length. In doing so, it complements and extends existing methods by focusing on pure Unicode whitespace replacement. We conducted an experimental evaluation based on common techniques like robustness, capacity, and visibility, including insertion, deletion, and formatting attacks, to systematically prove its strengths and weaknesses (Ahvanooey et al., 2018; Kamaruddin et al., 2018). We hypothesize that our proposed algorithm could serve as an enabler for increasing data sovereignty between companies to improve trust in supply chains like the automotive industry. Nevertheless, limitations exist that drive future research opportunities, which are identified and discussed in the following.

5.1. Limitations & Future Reserach

First, the proposed solution is based and tested on the Unicode standard and the UTF-8 scheme. Future research must check the influence of other coding schemes like the UTF-16, ISO-standardized Latin-1, and the potential impacts on the re-coding process to a smaller scheme like ASCII.

Second, the algorithm is robust against standard users who do not know and recognize a watermarked document. However, people familiar with the strategy can use smart attacks to apply targeted destruction of the watermark, e.g., random replacement of $\mathcal{A}_+$ as described in the robustness part. This can be eliminated by using a different random subset of $\mathcal{A}_+$ for every embedding operation. Future research is in progress to increase the robustness further by adding a key-based encryption and compression layer or by further increasing the extraction algorithm with smart analysis of broken watermarks to increase noise tolerance.

Third, print-out and re-scan attacks of watermarked texts using Optical Character Recognition (OCR) can but do not have to destroy the watermark. Since most spaces have a slightly different width that is not recognizable by human eyes, machines can restore the original whitespaces after a scan if configured correctly. This highly depends on the used font and applied OCR technique.

Fourth, selecting an appropriate digital watermarking or steganography algorithm for a specific use case is challenging. Future research is needed to compare the proposed method more extensively with existing solutions already presented in scientific literature and practical tools to highlight use-case-dependent strengths and weaknesses.

Fifth, the capacity depends on the number of whitespaces in the cover text $\mathcal{CT}$, the length of the watermark $\mathcal{W}$, and the elements of the watermarking alphabet $\mathcal{A}_+$ for text-based watermarks. Thus, limitations apply if a long watermark is to be embedded in a comparatively short text. Future research is needed to evaluate approaches to increasing capacity. This includes variations of the proposed solution like increasing $\mathcal{A}_+$ or a hybrid approach by combining it with other homoglyph-based or ZWSP solutions.

6. Conclusion

We have shown a new homoglyph-based digital watermarking solution for pure text documents by replacing whitespaces with similar-looking Unicode whitespaces. Our proposed solution allows the inclusion of a byte-encoded watermark inside a cover text without being noticed by human eyes or increasing the cover document's length. We have designed and implemented the watermarking embedding and extraction as a multiplatform library in the Kotlin programming language to demonstrate practical usability. Our findings will help to increase data sovereignty in practical use cases to prevent losing control when data is shared with external parties.

Based on the study's experimental evaluation, we

build a hidden solution that is robust in different business applications, like Word documents, emails, or Microsoft Teams chats. The comparison of existing solutions, especially UniSpaCh (Por et al., 2012) and Finegrain (Rizzo et al., 2019) indicated strong robustness and non-visibility of the proposed method with limitations on the encoding scheme, advanced robustness, and embedding capacity. Further work is being done to increase these constraints through checksums, error-correcting codes, and compressions.

Acknowledgments

This research was supported by the Center of Excellence Logistics and IT funded by the Fraunhofer-Gesellschaft.

References

Agarwal, M. (2013). Text steganographic approaches: A comparison. *International Journal of Network Security & Its Applications*, *5*(1), 91–106. https://doi.org/10.5121/ijnsa.2013.5107

Ahvanooey, M. T., Li, Q., Hou, J., Dana Mazraeh, H., & Zhang, J. (2018). Aitsteg: An innovative text steganography technique for hidden transmission of text message via social media. *IEEE Access*, *6*, 65981–65995. https://doi.org/10.1109/ACCESS.2018.2866063

Ahvanooey, M. T., Li, Q., Hou, J., Rajput, A. R., & Chen, Y. (2019). Modern text hiding, text steganalysis, and applications: A comparative analysis. *Entropy (Basel, Switzerland)*, *21*(4). https://doi.org/10.3390/e21040355

Ahvanooey, M. T., Li, Q., Zhu, X., Alazab, M., & Zhang, J. (2020). Anitw: A novel intelligent text watermarking technique for forensic identification of spurious information on social media. *Computers & Security*, *90*, 101702. https://doi.org/10.1016/j.cose.2019.101702

Ahvanooey, M. T., Zhu, M. X., Mazurczyk, W., & Bendechache, M. (2022). Information hiding in digital textual contents: Techniques and current challenges. *Computer*, *55*(6), 56–65. https://doi.org/10.1109/MC.2021.3113922

Askari, M., Mahmood, A., & Iqbal, Z. (2023). A novel font color and compression text steganography technique. *2023 International Conference on Communication, Computing and Digital Systems (C-CODE)*, 1–6. https://doi.org/10.1109/C-CODE58145.2023.10139867

Bertini, F., Rizzo, S. G., & Montesi, D. (2019). Can information hiding in social media posts represent a threat? *Computer*, *52*(10), 52–60. https://doi.org/10.1109/MC.2019.2917199

DataReportal, We Are Social, & Meltwater. (2023). Share of professionals worldwide using selected communication channels and digital tools for work as of 3rd quarter 2022, by frequency. Retrieved September 12, 2024, from https://www.statista.com/statistics/1306580/usage-communication-tools-for-work-worldwide-by-frequency/

Davis, C. L., & Kazi, U. (2021). Piracy of books in the digital age. In M. Bogre & N. Wolff (Eds.), *The routledge companion to copyright and creativity in the twenty-first century* (pp. 18–28). Routledge.

Hellmeier, M., Pampus, J., Qarawlus, H., & Howar, F. (2023). Implementing data sovereignty: Requirements & challenges from practice. *Proceedings of the 18th International Conference on Availability, Reliability and Security*, 1–9. https://doi.org/10.1145/3600160.3604995

Hellmeier, M., & von Scherenberg, F. (2023). A delimitation of data sovereignty from digital and technological sovereignty. *ECIS 2023 Research Papers*. Retrieved September 12, 2024, from https://aisel.aisnet.org/ecis2023_rp/306

Holgers, T., Watson, D. E., & Gribble, S. D. (2006). Cutting through the confusion: A measurement study of homograph attacks. *Proceedings of the annual conference on USENIX '06 Annual Technical Conference*, 261–266. Retrieved September 12, 2024, from https://www.usenix.org/legacy/events/usenix06/tech/holgers.html

Jalil, Z., & Mirza, A. M. (2009). A review of digital watermarking techniques for text documents. *2009 International Conference on Information and Multimedia Technology*, 230–234. https://doi.org/10.1109/ICIMT.2009.11

Jarke, M., Otto, B., & Ram, S. (2019). Data sovereignty and data space ecosystems. *Business & Information Systems Engineering*, *61*(5), 549–550. https://doi.org/10.1007/s12599-019-00614-2

Kamaruddin, N. S., Kamsin, A., Por, L. Y., & Rahman, H. (2018). A review of text watermarking: Theory, methods, and applications. *IEEE Access*, *6*, 8011–8028. https://doi.org/10.1109/ACCESS.2018.2796585

Khosravi, B., Khosravi, B., Khosravi, B., & Nazarkardeh, K. (2019). A new method for pdf steganography in justified texts. *Journal of Information Security and Applications*, *45*, 61–70. https://doi.org/10.1016/j.jisa.2019.01.003

Kintis, P., Miramirkhani, N., Lever, C., Chen, Y., Romero-Gómez, R., Pitropakis, N., Nikiforakis, N., & Antonakakis, M. (2017). Hiding in plain sight. In B. Thuraisingham, D. Evans, T. Malkin, & D. Xu (Eds.), *Proceedings of the 2017 acm sigsac conference on computer and communications security* (pp. 569–586). ACM. https://doi.org/10.1145/3133956.3134002

Korpela, J. (2002). Unicode spaces. Retrieved September 12, 2024, from https://www.jkorpela.fi/chars/spaces.html

Krishnan, R. B., Thandra, P. K., & Baba, M. S. (2017). An overview of text steganography. *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, 1–6. https://doi.org/10.1109/ICSCN.2017.8085643

Mir, N. (2014). Copyright for web content using invisible text watermarking. *Computers in Human Behavior*, *30*, 648–653. https://doi.org/10.1016/j.chb.2013.07.040

Norvig, P. (2012). English letter frequency counts: Mayzner revisited or etaoin srhldcu. Retrieved September 12, 2024, from https://norvig.com/mayzner.html

Petitcolas, F. A. P., Anderson, R., & Kuhn, M. G. (1999). Information hiding-a survey. *Proceedings of the IEEE*, *87*(7), 1062–1078. https://doi.org/10.1109/5.771065

Por, L. Y., Ang, T. F., & Delina, B. M. Y. (2008). Whitesteg: A new scheme in information hiding using text steganography. *WSEAS Transactions on Computers*, *7*(6), 735–745. Retrieved September 12, 2024, from https://dl.acm.org/doi/10.5555/1458369.1458384

Por, L. Y., Wong, K., & Chee, K. O. (2012). Unispach: A text-based data hiding method using unicode space characters. *Journal of Systems and Software*, *85*(5), 1075–1082. https://doi.org/10.1016/j.jss.2011.12.023

Qadir, M., & Ahmad, I. (2006). Digital text watermarking: Secure content delivery and data hiding in digital documents. *IEEE Aerospace and Electronic Systems Magazine*, *21*(11), 18–21. https://doi.org/10.1109/MAES.2006.284353

Qi, W., Yue, B., Wangdu, C., Xinghao, P., Zhipeng, C., Shaokang, W., Yizhao, W., & Chenwei, W. (2023). An overview on digital content watermarking. In J. Sun, Y. Wang, M. Huo, & L. Xu (Eds.), *Signal and information processing, networking and computers* (pp. 1311–1318, Vol. 917). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-3387-5_157

Rizzo, S. G., Bertini, F., & Montesi, D. (2016). Content-preserving text watermarking through unicode homoglyph substitution. In B. C. Desai, M. Toyama, J. Bernardino, & E. Desai (Eds.), *Proceedings of the 20th international database engineering & applications symposium on - ideas '16* (pp. 97–104). ACM Press. https://doi.org/10.1145/2938503.2938510

Rizzo, S. G., Bertini, F., & Montesi, D. (2019). Fine-grain watermarking for intellectual property protection. *EURASIP Journal on Information Security*, *2019*(1). https://doi.org/10.1186/s13635-019-0094-2

Shazzad-Ur-Rahman, M., Kaiser, M. S., Alam, M. B., & Nova, S. N. (2023). A data hiding technique combining steganography and cryptography for secured communication. *2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 432–437. https://doi.org/10.1109/ICICT4SD59951.2023.10303563

Sonnleitner, E. (2012). A robust watermarking approach for large databases. *2012 IEEE First AESS European Conference on Satellite Telecommunications (ESTEL)*, 1–6. https://doi.org/10.1109/ESTEL.2012.6400082

The Unicode Consortium. (2023). *The unicode standard, version 15.1.0*. Retrieved September 12, 2024, from https://www.unicode.org/versions/Unicode15.1.0/

Wang, Z.-H., Chang, C.-C., Lin, C.-C., & Li, M.-C. (2009). A reversible information hiding scheme using left–right and up–down chinese character representation. *Journal of Systems and Software*, *82*(8), 1362–1369. https://doi.org/10.1016/j.jss.2009.04.045

Xiao, C., Zhang, C., & Zheng, C. (2018). Fontcode: Embedding information in text documents using glyph perturbation. *ACM Transactions on Graphics*, *37*(2), 1–16. https://doi.org/10.1145/3152823

Paper V

Table A.5 Metadata Overview of Paper V

Title	Strengthening Data Sovereignty Through Digital Watermarking in Data Spaces
Authors	Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Haydar Qarawlus *Fraunhofer ISST, Dortmund, Germany*
Publication Year	2025
Publication Type	Conference
Conference Name	58th Hawaii International Conference on System Sciences (HICSS)
Conference Location	Waikoloa Village, Big Island, Hawaii, USA
Conference Date	07. January 2025 - 10. January 2025
Publisher / Database	AIS Affiliated
DOI / Link	https://doi.org/10.24251/HICSS.2025.520
Status	Published
Ranking	VHB: B (2024 Rating) CORE: A (2018 Rating)[3] ERA: A (2010 Rating)
Comment	The paper was nominated for the best paper award.

[3] The HICSS conference was last ranked by CORE in 2018 (see https://portal.core.edu.au/conf-ranks/575/, accessed on Feb. 10, 2026). It is not listed in the latest ICORE 2026 rating.

Proceedings of the 58th Hawaii International Conference on System Sciences | 2025

Strengthening Data Sovereignty Through Digital Watermarking in Data Spaces

Malte Hellmeier
Fraunhofer ISST
malte.hellmeier@isst.fraunhofer.de

Haydar Qarawlus
Fraunhofer ISST
haydar.qarawlus@isst.fraunhofer.de

Abstract

Data spaces have emerged as a paradigm for maintaining data sovereignty and interoperability in data sharing among various stakeholders. There is an increasing interest and progress in research and practice whereby currently used implementations only enable data sovereignty within their trusted environments. Digital watermarking, a key concept in the research domain of information hiding, addresses similar principles of securing data ownership, while its integration within data spaces remains unexplored. This paper combines both domains by presenting the first integrated digital watermarking solution in a data space. Guided by design science research, we designed and developed two extensions for the connector of the Eclipse Dataspace Components in the Java programming language to validate the solution's practicality. The resulting artifact shows a robust data sovereignty enhancement, paving the way for more protection and control in future data spaces.

Keywords: Data Ecosystems, Design Science Research, Digital Watermarking, Steganography, Information Hiding

1. Introduction

The increasing interest in securely sharing data between participants, with its growing demand for solutions, has stimulated great interest in the concepts of data spaces (Gieß et al., 2023). Different initiatives and associations of many global companies and experts arose over the last decade, like the International Data Spaces Association (IDSA), Gaia-X, and the Data Spaces Support Centre (DSSC). They advance the ideas shown by various practical data spaces and use cases worldwide (Mertens & Kuster, 2024), with increasing political interest, as exemplary shown by the European strategy of data (European Commission, 2020).

One key driver is data sovereignty, the possibility of controlling data in a self-determined manner (Jarke et al., 2019). However, current research identified challenges in controlling data after it leaves its boundaries or authorized environments (Hellmeier et al., 2023). This results in a minimum level of trust that needs to be implemented (Gil et al., 2020). To further strengthen data sovereignty with a baseline of trust, the concepts of digital watermarking act as an enabler for increased protection mechanisms. Due to the standardized body of data spaces for interoperability, including protocols (IDSA, 2024) and reference architectures (Otto et al., 2019), the topic is challenging.

We close this gap by combining the research domains of digital watermarking and data spaces by considering the framework conditions, existing protocols, semantics, and standards. In doing so, this paper uses the Design Science Research (DSR) approach following Hevner et al. (2004) and Hevner (2007) to design and evaluate our IT artifact. The presented version consists of the first implementation of two watermark extensions for data space connectors to increase data sovereignty, even if it leaves the control boundaries while being interoperable with existing data spaces.

Our paper is structured as follows based on the publication schema introduced by Gregor and Hevner (2013): Relevant concepts, methods, and terminologies necessary for understanding this paper with an overview of related work are introduced in Section 2 and 3. The design process of the IT artifact for incorporating digital watermarking techniques in data spaces with

URI: https://hdl.handle.net/10125/109366
978-0-9981331-8-8

its implementation in a connector is presented and evaluated in Section 4. An overview of the theoretical and practical implications with its limitations and future research opportunities are discussed in Section 5 and concluded in Section 6. Based on the outline, our paper makes the following concrete contributions:

(i) Extends existing Information Systems (IS) research by providing an integrated view of digital watermarking in data spaces.

(ii) Designs and implements an IT artifact based on DSR.

(iii) Discusses concrete theoretical and practical implications based on the evaluation of the constructed artifact.

2. Background & Related Work

This chapter describes the terminologies of data sovereignty, data spaces, and the connector component, along with the concepts of steganography and digital watermarking. It helps to unify the needed understanding and present related work, considering its distinction from this work.

2.1. Data Sovereignty

The capability of an entity to control the access and usage of data to increase trust is designated as *data sovereignty* (Hellmeier & von Scherenberg, 2023). Such entities could range from individuals to groups and organizations to companies and large enterprises (Firdausy et al., 2022b; Jarke et al., 2019). While the term's first usage can be traced back to the early 2000s (Hallinan, 2022), its interpretation changed over time in various domains and application scenarios. Besides data sovereignty concepts for indigenous people (Taylor & Kukutai, 2016) or in specific legal contexts (Lian, 2021; Woods, 2018), this paper is based on definitions in the context of information systems, computer science, and software engineering. Here, the conceptualization of the term consists of a data provider and a data consumer negotiating specific rules and contracts under which conditions data can be used, supported by a data infrastructure that helps ensure trust and the management of the rules and contracts (von Scherenberg et al., 2024). Therefore, "[d]ata sovereignty refers to the self-determination of individuals and organizations with regard to the use of their data" (Jarke et al., 2019, p. 550).

Various associations and organizations like Gaia-X, the IDSA, and the DSSC use data sovereignty as a foundation to create architectures (Firdausy et al., 2022a; Otto et al., 2019), frameworks (Gaia-X European Association for Data and Cloud AISBL, 2022), and protocols (IDSA, 2024) to bring conceptual ideas into realization. One possibility are data spaces with their core component of a connector, introduced in the upcoming Section 2.2.

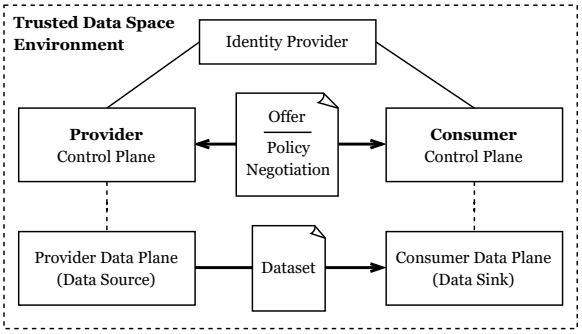


Figure 1. Architecture Overview of a Data Space Connector in its Trusted Environment.

2.2. Data Spaces & Connector

The concept of a data space (sometimes also written dataspace) is based on data sovereignty. It is a trustworthy decentralized environment for data sharing, where data providers and consumers can offer, sell, and exchange data (Nagel & Lycklama, 2021). Data spaces enable the technical implementation of data sovereignty by requiring several components to facilitate trust and interoperability among participants (Nagel & Lycklama, 2021). The main software artifact facilitating interoperability and connectivity within a data space is called a *connector*. Figure 1 illustrates the architecture of a system comprising two participants sharing data using connectors based on the standardized Dataspace Protocol (DSP) (IDSA, 2024). This Figure resembles the implementation of the connector in the Eclipse Dataspace Components (EDC) [1]. The EDC connector is a widespread open-source compatible implementation of a data space connector, which we additionally use as the basis for our developed artifact designed and described in this work.

Based on the DSP (IDSA, 2024), a transfer with a connector consists of two sub-modules facilitating controlled data sharing within a data space. The main component is the control plane, which is responsible for most of the management tasks related to data sharing. The list of functions includes cataloging, i.e., storing metadata about datasets intended to be offered within the data space. Additionally, the connector handles policy

[1] https://projects.eclipse.org/projects/technology.edc

management and their negotiation, i.e., the definition of access and usage policies defining the conditions attached to the use of data. Finally, the logical construct of a data plane facilitates the actual transmission of data between a provider and consumer.

As shown in Figure 1, each connector within a data space is connected to an identity provider, which is responsible for the issuance and validation of identities and certificates within a data space. It is essential to increase trust between different parties, like data providers and consumers. The connector uses the services of the identity provider whenever the verification of certificates from the participants is required (Nagel & Lycklama, 2021).

In order to offer data within a data space using the connector, the provider must create a representation of the data called a dataset (IDSA, 2024). The provider can further define policies connected to the dataset that determine their access and usage conditions. Combining these elements as a relationship is called an offer (IDSA, 2024). Once data sharing is needed, the potential consumer requests the catalog of datasets from the provider, which contains all the information required to start a contract negotiation process. During this process, a set of messages are exchanged among the two parties to reach an agreement. Once an agreement is established, the actual transfer process on the data plane level can take place.

To summarize existing concepts and delimit our work from related work, the IDSA forms a baseline for data spaces with the reference architecture, the DSP on the protocol level, and compatible implementations like the EDC on the implementation level for connectors. Nevertheless, ensuring data sovereignty is still challenging since the overall system only works within the trusted environment's scope. Even if control is ensured up to the data sink on the consumer side, data can still leave the trusted environment, leading to the overall system's limits and loss of control. In short, there is no data sovereignty. While related work mainly focuses on different access and usage control mechanisms inside a specific environment or architecture (Ferraiolo et al., 2011), maintaining control and data sovereignty is still unexplored after data leaves the data space. This also applies to related concepts and initiatives like GAIA-X (Otto, 2021), eXtensible Access Control Markup Language (XACML) applications (OASIS, 2013), and different solutions built on them (Appenzeller et al., 2020). To close this gap, we enhance data spaces with the help of information hiding. The two commonly known concepts of steganography and digital watermarking are introduced in the following.

2.3. Steganography & Digital Watermarking

The enrichment and expansion of data with secret information directly inside it belong to the research field of information hiding. Specialized techniques and algorithms arose over time and were classified as methods for steganography and watermarking. While *steganography* strives to extend or generate a cover medium with an invisible secret message inside it (Krishnan et al., 2017), watermarking or *digital watermarking* aims to protect a dataset by enriching it with additional details or copyright information, the so-called watermark (Podilchuk & Delp, 2001; Qadir & Ahmad, 2006). As a result, steganography methods often include encryption techniques with an unknown hiding method, as the focus is on securing the secret message that may include confidential information in a cover medium that is only a means to an end (Ahvanooey et al., 2022; Krishnan et al., 2017; Podilchuk & Delp, 2001). In contrast, digital watermarking focuses on protecting the intellectual property of the cover medium itself, where the embedding technique and the message itself do not need to be secret (Podilchuk & Delp, 2001; Rizzo et al., 2016). As a result, the robustness of watermarks to be permanently embedded is crucial (Ahvanooey et al., 2018; Jalil & Mirza, 2009), whereby a wide variety of techniques exist, focusing on different cover media types like image, text, video or audio (Podilchuk & Delp, 2001).

To our knowledge, there are currently no publications or usages of digital watermarking techniques inside data spaces available. Related work like Pan et al. (2010) only discusses the combination of Organization Based Access Control (OrBAC) with watermarking on an image level. Nevertheless, practitioners discussed watermarking techniques as a possible solution for strengthening data sovereignty (Hellmeier et al., 2023). Thus, we use a practical problem-solving method to combine both research areas and design a solution.

3. Method

This paper aims to mitigate the current practical problem of missing data sovereignty and usage control after data leaves a data space's trusted environment. We use input from research and practice to build and design an IT artifact that is evaluated toward this problem statement. We use the DSR methodology based on Hevner et al. (2004) and Hevner (2007) to derive design knowledge in a structured way, following the three-cycle approach. All three cycles with their application in this

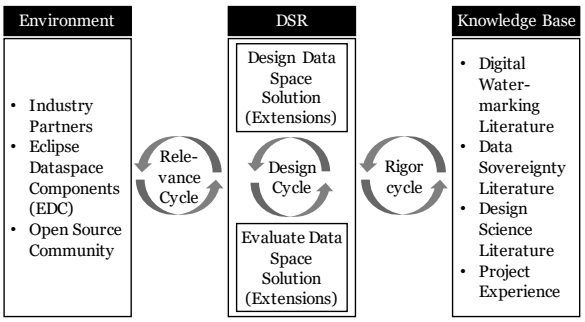


Figure 2. Applied DSR Cycles following Hevner (2007, p. 88).

work are described in the following with a summarized overview presented in Figure 2.

The *relevance cycle* initiates the DSR process and consists of external requirements and acceptance criteria (Hevner, 2007). In this study, the environment is based on the one hand on industry partners and their requirements. These requirements were carefully collected in a previous interview study of different sectors in different countries (Hellmeier et al., 2023). On the other hand, the open source community of the EDC are part of the environment since our artifact is based on an actual IT implementation.

The *rigor cycle* consists of existing sources, experiences, expertise, artifacts, and processes (Hevner, 2007). The present study's knowledge base comprises current literature findings from digital watermarking and data sovereignty research domains. It includes DSR literature for the methodological foundation and project experience from the development team.

The *design cycle* as the core of DSR aims to craft the artifact based on the input from the relevance and rigor cycle with a downstream evaluation (Hevner, 2007). We use the Java programming language to implement the IT artifact based on EDC connector extensions and test them afterward in an evaluation.

During the iterative execution of the design process, all seven DSR guidelines are carefully followed. Table 1 gives a comprehensive overview of our application from the guidelines introduced by Hevner et al. (2004).

4. Results

To combine the previously introduced domains of data spaces and digital watermarking, a concept of how to increase data sovereignty by dynamically watermarking data before it is shared is proposed.

Starting with gathered information from the knowledge base in the rigor cycle, watermarking techniques consist of a cover medium that needs to be protected and the watermark itself that gets included in the cover medium (Por et al., 2012; Qadir & Ahmad, 2006). In the context of a data space, we use the dataset shared between a data provider and a data consumer as a cover medium and the policy containing the conditions for data usage as a watermark. Depending on the use case of the environmental relevance cycle, a different semantic content of a watermark could be used. Through direct use of the policy as a watermark, it remains persistent within the dataset even if it leaves the data space's trusted environment. Therefore, we decided in the design cycle that the watermarking process only needs to be included on the provider's side, being interoperable between different connectors and data spaces due to the independence of the consumer side. To showcase the design functionality, we implemented it in Java based on the connector of the EDC, consisting of two extensions. One extension builds the watermark based on the policy, while the second extension integrates the watermark into the dataset. In the following subsections, the concept of each extension is described more generally, including details of our design.

Since most interactions like the policy negotiation or the data transfer remain unaffected by our solution, the watermarking starts directly after the requested transfer, as depicted in Figure 3. Therefore, the proposed solution only extends the currently used data sharing body without removing or changing existing interactions to preserve interoperability, shown by the *Watermarking* box in the sequence diagram of Figure 3.

4.1. Policy Information Extension

The designed Policy Information Extension (PIE) aims to extract the decision under which conditions a data transfer takes place to build the watermark. Based on the previously described architecture of the EDC as an environmental input from the relevance cycle, the contract terms are negotiated via the DSP between the owner of the dataset (data provider) and the external participant receiving it (data consumer) on the control plane. These contract terms are defined as a policy in the terminology of data spaces and their connectors (Otto et al., 2019), described in the Open Digital Rights Language (ODRL)[2]. Therefore, the PIE is built as an EDC control plane extension.

Our specific implementation of a possible PIE offers a REST API for external requests to return the watermark. Referring to Figure 3, the process starts after a consumer successfully negotiates a contract (*negotiateContract()*) and requests the start of the data

[2]https://www.w3.org/TR/odrl-model

Table 1. Applied DSR Guidelines following Hevner et al. (2004, p. 83).

No.	Guideline	Our Application
1	Design as an Artifact	Two implemented EDC connector extensions as IT artifact.
2	Problem Relevance	Identified lack of data sovereignty from industry partners after it leaves their control boundaries in heterogeneous system landscapes.
3	Design Evaluation	*Descriptive* evaluation based on the initial problem relevance and *experimental* simulation evaluation through technical testbed executions.
4	Research Contributions	Contributed IT artifact as a descriptive concept and implemented extensions.
5	Research Rigor	Knowledge based on existing data sovereignty and digital watermarking research.
6	Design as a Search Process	A development process for the artifact design.
7	Communication of Research	Published results in the form of this paper publication.

sharing (*startTransfer(destination)*). It accepts a transfer request ID as an input and uses it to locate and retrieve the corresponding contract with its policies that prompted the transfer request within the control plane's core services (*getPolicyData(transferRequestId)*). It uses the collected information to transform the ODRL policy into a serialized String, which forms the watermark. This watermark is returned as an HTTP response to fulfill the initial request.

4.2. Watermark Embedding Extension

The Watermark Embedding Extension (WEE) aims to inject the watermark inside the dataset. The actual data transfer occurs in the data plane of the EDC connector. Therefore, the WEE is integrated directly where the byte shoveling is done. Whenever a dataset designated for watermarking by its data provider is transferred via the EDC connector, the WEE intercepts the stream of the dataset, modifies it by injecting the watermark, and passes it on within the framework of the data plane from where it finally arrives at the data consumer side.

A specific watermarking algorithm injects the watermark into the cover medium of the dataset. Our specific WEE implementation is based on the text-based watermarking approach introduced in Hellmeier et al. (2025) that can include a byte-encoded watermark inside a cover text without increasing the content size or getting noticed by humans. Since different algorithms arose over time, specialized on different file types like text, image, or video (Su et al., 1998) with respective advantages and disadvantages, the usage of an appropriate algorithm for every data space is crucial and must be selected accordingly, depending on the use cases. If the application focuses on a high embedding capacity, alternative watermarking algorithms like UniSpaCh (Por et al., 2012) or zero-width character solutions can be used instead. Our implemented and utilized method fulfills the DSR guidelines four and seven for contributions and communication (Hevner et al., 2004) because the source code of the watermarking algorithm is made available publicly[3] in comparison to alternative solutions like UniSpaCh.

As described above, the negotiated contracts and their respective policies are stored within the control plane component of an EDC connector. Since the data plane only knows the transfer request ID, our implementation requests additional information from the control plane at the PIE. Whenever the data plane executes a data transfer, the WEE requests the watermark by calling the PIE to acquire the data representation of the policy incorporated in the contract. The dataset itself, acting as a cover medium, can be retrieved directly from the source location (*getAssetContentData()*). As soon as both the dataset and the watermark are available, the WEE injects the policy as a watermark into the dataset (*injectWatermark(policyData)* and forwards it within the data plane to the final location on the data consumer's side (*forwardData()*).

4.3. Evaluation

Since evaluation is crucial in DSR as stated in guideline three (see Table 1), we use an *experimental* simulation method based on a testbed setup and a *descriptive* evaluation method based on the information

[3] https://github.com/FraunhoferISST/TREND

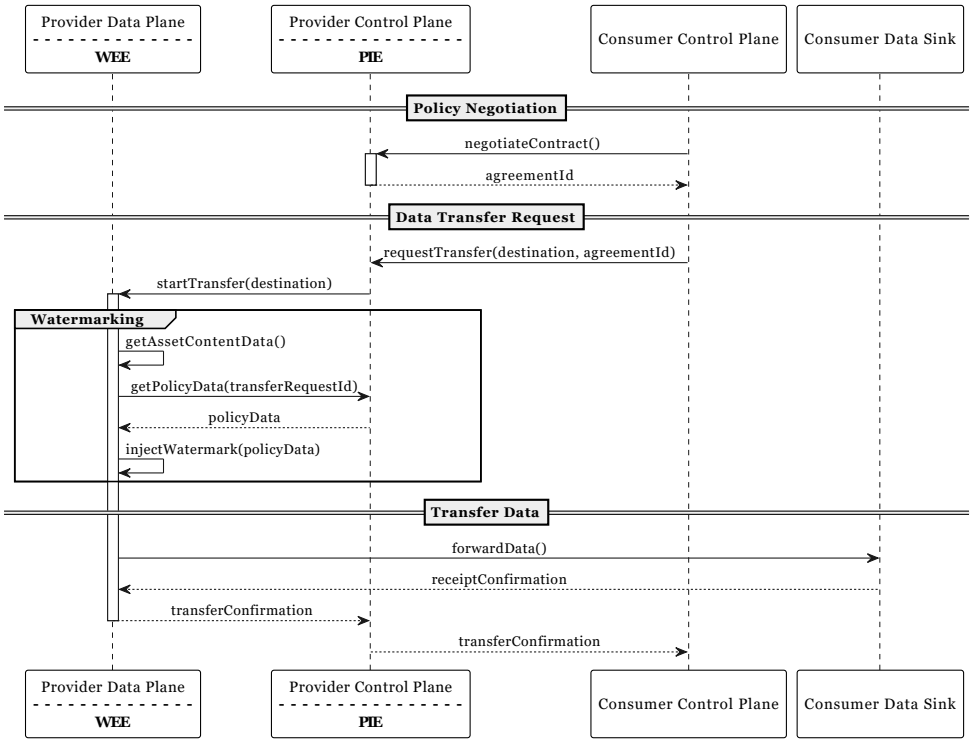


Figure 3. Unified Modeling Language (UML) Sequence Diagram of a Data Transfer with Watermarking.

from our knowledge base (Hevner et al., 2004). These two methods were deliberately chosen because others, like experimental usability tests, are not applicable since the resulting artifact is a backend solution without direct user interaction. Thus, frontend tests, interview-based field studies, or user interactions are not suitable. In the context of this work, one important aspect of the artifact evaluation is the validity criteria, analyzed and described with both methods in the following. "Validity means that the artifact works and does what it is meant to do; that it is dependable in operational terms in achieving its goals" (Gregor & Hevner, 2013, p. 351). While continuous evaluation during the design cycle and artifact development was conducted, this chapter focuses on the final testing. It aims to describe the setup of the testbed for transparency and reproducibility. To reduce the bias, the evaluation is executed by a group of researchers instead of a single person (Hevner & Chatterjee, 2010).

First, we use the *experimental* evaluation method following Hevner et al. (2004) by building a real testbed. The testbed consists of two configured EDC versions[4], one provider EDC with our PIE and WEE extension, and one consumer EDC without our extensions to span a minimal data space. Both EDCs are connected to a minimal user interface for monitoring and controlling the tests. The entire evaluation is executed in a Docker setup on a Linux virtual machine running Ubuntu 22.04 LTS. We created two copies of a plain *.txt* file containing a dummy Lorem ipsum text block. The first file is transferred over the EDC from the provider into the data sink of the consumer with the basic setup and without enabling our extensions. Such a basic setup is mostly used in currently operated data spaces. The second file uses both extensions to enable our proposed watermarking strategy and test the designed artifact. After analyzing both files on the consumer side, the first one is unprotected and could easily be misused, strengthening the initially introduced problem statement and proving the concerns raised by the practitioners in the relevance cycle (Hellmeier et al., 2023). In contrast, the second file still contains the watermark and thus has lasting protection.

Second, we use the *descriptive* evaluation method based on the initially motivated problem for strengthening data sovereignty after it leaves its controlled boundaries. This scenario is used "around the artifact to demonstrate its utility" (Hevner et al., 2004, p. 86). Based on the results from the testbed of the

[4]A minimally configured variant based on EDC v0.5.1 with a custom file transfer extension, a mocked identity provider and without cloud extensions: https://github.com/eclipse-edc/Connector

experimental evaluation, the watermark stays persistent in the data sink of the consumer side. This ensures that the watermark remains persistent even after the data leaves the trusted environment of the data space.

5. Discussion

To our knowledge, this paper is the first combination of the research domains of data spaces and digital watermarking in the context of data ecosystems. Based on DSR with input from research and focus on a practical problem domain, we designed, implemented, and evaluated an IT artifact based on existing components and protocols and showed the practicable use by two EDC connector extensions.

Using watermarking techniques inside data spaces has clear implications for increasing data usage control and, thus, data sovereignty and trust in the overall network. This leads to specific implications for theory and practice.

On the one hand, regarding *theoretical implications*, this work acts as a starting point for watermarking in data spaces. Research needs to follow up by developing metrics on measuring data sovereignty and evaluating the amount digital watermarking has on it. Further, digital watermarking offers new perspectives on the application of data sovereignty. The integration into existing data spaces and data ecosystems can influence existing theoretical approaches, processes, models, and architectures. By improving technical enforcement mechanisms, the existing literature on trust frameworks must be checked to reflect the new opportunities and their downstream effects.

On the other hand, regarding *practical implications*, DSR is used as a problem-solving process for designing artifacts that can be used in reality (Hevner et al., 2004). With regard to our designed artifact, existing data spaces need to adopt and test digital watermarking with existing use cases in practice. Building on this work's proposed EDC connector extensions, it helps practitioners integrate watermarking inside different stages of the data sharing process. Nevertheless, the integration and continuous adoption of the EDC leads to several implementation challenges due to a lack of documentation in the upstream project. Additional experience is needed to check the feasibility of the solution within companies' existing software systems and architectures.

However, different questions are still open that are not discussed in this paper. These points of contact lead to future research opportunities guided by this work's limitations, which are discussed below.

5.1. Limitations & Future Research

Despite careful implementation, this study, with its presented IT artifact, has limitations and offers future research opportunities.

First, people familiar with the watermarking algorithm could selectively remove or change watermarked datasets. Future research needs to check whether concepts like checksums or encryption mechanisms can help to tackle this limitation. Further, reviewing and using alternative steganography methods focusing on counterfeit protection could lead to more robust solutions against attackers.

Second, the proposed solution has a high level of generalization as the variable structure allows the integration of any suitable watermarking algorithm, e.g., for specific file types like images (Begum & Uddin, 2020; Wan et al., 2022) or text (Ahvanooey et al., 2022; Jalil & Mirza, 2009). Nevertheless, if needed, this could lead to semantic interoperability issues without agreement on one watermarking method inside a data space.

Third, the practicability must be checked in a real-world environment in a concrete use case. Future research needs to conduct a field or case study to check its practicability and identify further improvement points. This can be done on a technical level with benchmarking and software quality analysis, as well as with expert interviews for qualitative research and an information systems point of view. It needs to include the involement of adjacent research domains and concepts beyond data spaces and data ecosystems, such as data usage control or digital rights management.

Fourth, watermarks in data spaces only help to a limited extent with policy enforcement. Even if the policies are directly integrated into a dataset that stays present after data leaves the trusted environment, this does not immediately ensure full enforcement. Future work is in progress to determine whether using mechanisms on lower Open Systems Interconnection (OSI) levels, like the network layer instead of the application layer (ISO/IEC, 1994), can help guarantee policy enforcement and increase data sovereignty further.

6. Conclusion

We have designed and implemented an IT artifact consisting of two data space connector extensions for watermarking datasets before sharing them with other participants in a data ecosystem, following DSR. The system described is lightweight and integrated on the provider side while being interoperable with existing

data spaces. The implemented version based on the EDC and our text-based watermarking technique proves the practicability of the theoretical solution.

Our results showcase how digital watermarking methods help increase data sovereignty in data spaces while being consistent with existing standards, protocols, semantics, and reference architectures. The evaluation in our testbed environment shows its valid functionality and identifies open points for future research. Further work is needed to evaluate the increase in policy enforcement with additional tests in other real-world data spaces.

Acknowledgments

This research was supported by the Cluster of Excellence Cognitive Internet Technologies CCIT and the Center of Excellence Logistics and IT, which are funded by the Fraunhofer-Gesellschaft. We also thank Ernst-Christoph Schrewe for his input and support during the extension development process.

References

Ahvanooey, M. T., Li, Q., Hou, J., Dana Mazraeh, H., & Zhang, J. (2018). Aitsteg: An innovative text steganography technique for hidden transmission of text message via social media. *IEEE Access*, *6*, 65981–65995. https://doi.org/10.1109/ACCESS.2018.2866063

Ahvanooey, M. T., Zhu, M. X., Mazurczyk, W., & Bendechache, M. (2022). Information hiding in digital textual contents: Techniques and current challenges. *Computer*, *55*(6), 56–65. https://doi.org/10.1109/MC.2021.3113922

Appenzeller, A., Rode, E., Krempel, E., & Beyerer, J. (2020). Enabling data sovereignty for patients through digital consent enforcement. In F. Makedon (Ed.), *Proceedings of the 13th acm international conference on pervasive technologies related to assistive environments* (pp. 1–4). ACM. https://doi.org/10.1145/3389189.3393745

Begum, M., & Uddin, M. S. (2020). Digital image watermarking techniques: A review. *Information*, *11*(2), 110. https://doi.org/10.3390/info11020110

European Commission. (2020). A european strategy for data: Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions. Retrieved September 16, 2024, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066

Ferraiolo, D., Atluri, V., & Gavrila, S. (2011). The policy machine: A novel architecture and framework for access control policy specification and enforcement. *Journal of Systems Architecture*, *57*(4), 412–424. https://doi.org/10.1016/j.sysarc.2010.04.005

Firdausy, D. R., de Alencar Silva, P., van Sinderen, M., & Iacob, M.-E. (2022a). Towards a reference enterprise architecture to enforce digital sovereignty in international data spaces. *2022 IEEE 24th Conference on Business Informatics (CBI)*, 117–125. https://doi.org/10.1109/CBI54897.2022.00020

Firdausy, D. R., de Alencar Silva, P., van Sinderen, M., & Iacob, M. E. (2022b). Semantic discovery and selection of data connectors in international data spaces. In M. Zelm, A. Boza, R.-D. León, & R. Rodriguez (Eds.), *Interoperability for enterprise systems and applications workshops, i-esa workshops 2022* (Vol. 3214).

Gaia-X European Assocication for Data and Cloud AISBL. (2022). Gaia-x trust framework: 22.10 release (Gaia-X European Assocication for Data and Cloud AISBL, Ed.). Retrieved September 16, 2024, from https://docs.gaia-x.eu/policy-rules-committee/trust-framework/22.10/

Gieß, A., Möller, F., Schoormann, T., & Otto, B. (2023). Design options for data spaces. *ECIS 2023 Research Papers*. Retrieved September 16, 2024, from https://aisel.aisnet.org/ecis2023_rp/287

Gil, G., Arnaiz, A., Diez, F. J., & Higuero, M. V. (2020). Evaluation methodology for distributed data usage control solutions. *2020 Global Internet of Things Summit (GIoTS)*, 1–6. https://doi.org/10.1109/GIOTS49054.2020.9119565

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, *37*(2), 337–355. https://doi.org/10.25300/MISQ/2013/37.2.01

Hallinan, D. (2022). *Data protection and privacy, volume 14: Enforcing rights in a changing world*. Bloomsbury Publishing Plc.

Hellmeier, M., Pampus, J., Qarawlus, H., & Howar, F. (2023). Implementing data sovereignty: Requirements & challenges from practice. *Proceedings of the 18th International Conference on Availability, Reliability*

and Security, 1–9. https://doi.org/10.1145/3600160.3604995

Hellmeier, M., Qarawlus, H., Norkowski, H., & Howar, F. (2025). A hidden digital text watermarking method using unicode whitespace replacement. *Proceedings of the 58th Hawaii International Conference on System Sciences*.

Hellmeier, M., & von Scherenberg, F. (2023). A delimitation of data sovereignty from digital and technological sovereignty. *ECIS 2023 Research Papers*. Retrieved September 16, 2024, from https://aisel.aisnet.org/ecis2023_rp/306

Hevner, A., & Chatterjee, S. (2010). *Design research in information systems* (Vol. 22). Springer US. https://doi.org/10.1007/978-1-4419-5653-8

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Jorunal of Information Systems*, *19*(2).

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75. https://doi.org/10.2307/25148625

IDSA. (2024). Dataspace protocol 2024-1 (International Data Spaces Association, Ed.). Retrieved September 16, 2024, from https://docs.internationaldataspaces.org/dataspace-protocol/

ISO/IEC. (1994). Open systems interconnection: Basic reference model: The basic model. Retrieved September 16, 2024, from https://www.iso.org/standard/20269.html

Jalil, Z., & Mirza, A. M. (2009). A review of digital watermarking techniques for text documents. *2009 International Conference on Information and Multimedia Technology*, 230–234. https://doi.org/10.1109/ICIMT.2009.11

Jarke, M., Otto, B., & Ram, S. (2019). Data sovereignty and data space ecosystems. *Business & Information Systems Engineering*, *61*(5), 549–550. https://doi.org/10.1007/s12599-019-00614-2

Krishnan, R. B., Thandra, P. K., & Baba, M. S. (2017). An overview of text steganography. *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, 1–6. https://doi.org/10.1109/ICSCN.2017.8085643

Lian, Y. (2021). *Data rights law 3.0*. Peter Lang UK.

Mertens, C., & Kuster, A. (2024). The data space radar (International Data Spaces Association, Ed.). Retrieved September 16, 2024, from https://internationaldataspaces.org/download/45467/?tmstv=1718188509

Nagel, L., & Lycklama, D. (2021). Design principles for data spaces - position paper. https://doi.org/10.5281/ZENODO.5105744

OASIS. (2013). Extensible access control markup language (xacml) version 3.0. https://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html

Otto, B. (2021). Gaia-x and ids (International Data Spaces Association, Ed.). https://doi.org/10.5281/ZENODO.5269077

Otto, B., Steinbuss, S., Teuscher, A., & Lohmann, S. (2019). Ids reference architecture model (International Data Spaces Association, Ed.; 3.0). https://doi.org/10.5281/ZENODO.5105529

Pan, W., Coatrieux, G., Cuppens-Boulahia, N., Cuppens, F., & Roux, C. (2010). Watermarking to enforce medical image access and usage control policy. *2010 Sixth International Conference on Signal-Image Technology and Internet Based Systems*, 251–260. https://doi.org/10.1109/SITIS.2010.50

Podilchuk, C. I., & Delp, E. J. (2001). Digital watermarking: Algorithms and applications. *IEEE Signal Processing Magazine*, *18*(4), 33–46. https://doi.org/10.1109/79.939835

Por, L. Y., Wong, K., & Chee, K. O. (2012). Unispach: A text-based data hiding method using unicode space characters. *Journal of Systems and Software*, *85*(5), 1075–1082. https://doi.org/10.1016/j.jss.2011.12.023

Qadir, M., & Ahmad, I. (2006). Digital text watermarking: Secure content delivery and data hiding in digital documents. *IEEE Aerospace and Electronic Systems Magazine*, *21*(11), 18–21. https://doi.org/10.1109/MAES.2006.284353

Rizzo, S. G., Bertini, F., & Montesi, D. (2016). Content-preserving text watermarking through unicode homoglyph substitution. In B. C. Desai, M. Toyama, J. Bernardino, & E. Desai (Eds.), *Proceedings of the 20th international database engineering & applications symposium on - ideas '16* (pp. 97–104). ACM Press. https://doi.org/10.1145/2938503.2938510

Su, J. K., Hartung, F., & Girod, B. (1998). Digital watermarking of text, image, and video documents. *Computers & Graphics*, *22*(6), 687–695. https://doi.org/10.1016/S0097-8493(98)00089-2

Taylor, J., & Kukutai, T. (Eds.). (2016). *Indigenous data sovereignty: Toward an agenda* (Vol. no. 38). Australian National University Press.

von Scherenberg, F., Hellmeier, M., & Otto, B. (2024). Data sovereignty in information systems. *Electronic Markets*, *34*(1). https://doi.org/10.1007/s12525-024-00693-4

Wan, W., Wang, J., Zhang, Y., Li, J., Yu, H., & Sun, J. (2022). A comprehensive survey on robust image watermarking. *Neurocomputing*, *488*, 226–247. https://doi.org/10.1016/j.neucom.2022.02.083

Woods, A. K. (2018). Litigating data sovereignty. *The Yale Law Journal*, *128*. Retrieved September 16, 2024, from http://hdl.handle.net/20.500.13051/10358

Paper VI

Table A.6 Metadata Overview of Paper VI

Title	Innamark: A Whitespace Replacement Information-Hiding Method
Authors	Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Hendrik Norkowski *Montsecure GmbH, Bochum, Germany* Ernst-Christoph Schrewe *Fraunhofer ISST, Dortmund, Germany* Haydar Qarawlus *Fraunhofer ISST, Dortmund, Germany* Falk Howar *TU Dortmund & Fraunhofer ISST, Dortmund, Germany*
Publication Year	2025
Publication Type	Journal
Journal Name	IEEE Access
Publisher / Database	IEEE Xplore
DOI / Link	https://doi.org/10.1109/ACCESS.2025.3583591
Status	Published
Ranking	VHB: - (2024 Rating) CORE: - (2020 Rating) ERA: - (2010 Rating) SJR: Q1 (2024 Rating)
Comment	Extended version of Paper IV [99] and previously published as preprint on arXiv in [93].

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 April 2025, accepted 22 June 2025, date of publication 26 June 2025, date of current version 18 July 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3583591

RESEARCH ARTICLE

Innamark: A Whitespace Replacement Information-Hiding Method

MALTE HELLMEIER[1], HENDRIK NORKOWSKI[2], ERNST-CHRISTOPH SCHREWE[1], HAYDAR QARAWLUS[1], AND FALK HOWAR[3,1]

[1]Fraunhofer ISST, 44147 Dortmund, Germany
[2]Montsecure GmbH, 44799 Bochum, Germany
[3]Department of Computer Science, Technical University Dortmund, 44227 Dortmund, Germany

Corresponding author: Malte Hellmeier (malte.hellmeier@isst.fraunhofer.de)

This work was supported by the Cluster of Excellence Cognitive Internet Technologies (CCIT) and the Center of Excellence Logistics and IT, which are funded by the Fraunhofer-Gesellschaft.

ABSTRACT Large language models (LLMs) have gained significant popularity in recent years. Differentiating between a text written by a human and one generated by an LLM has become almost impossible. Information-hiding techniques such as digital watermarking or steganography can help by embedding information inside text in a form that is unlikely to be noticed. However, existing techniques, such as linguistic-based or format-based methods, change the semantics or cannot be applied to pure, unformatted text. In this paper, we introduce a novel method for information hiding called Innamark, which can conceal any byte-encoded sequence within a sufficiently long cover text. This method is implemented as a multi-platform library using the Kotlin programming language, which is accompanied by a command-line tool and a web interface. By substituting conventional whitespace characters with visually similar Unicode whitespace characters, our proposed scheme preserves the semantics of the cover text without changing the number of characters. Furthermore, we propose a specified structure for secret messages that enables configurable compression, encryption, hashing, and error correction. An experimental benchmark comparison on a dataset of 1 000 000 Wikipedia articles compares ten algorithms. The results demonstrate the robustness of our proposed Innamark method in various applications and the imperceptibility of its watermarks to humans. We discuss the limits to the embedding capacity and robustness of the algorithm and how these could be addressed in future work.

INDEX TERMS Blind watermarking, copyright protection, data hiding, data sovereignty, digital text watermarking, information hiding, steganography, Unicode characters.

I. INTRODUCTION

Interest in large language models (LLMs) has grown rapidly in recent years, with a variety of promising applications emerging for individuals and businesses. The enduring process of digitization has led to numerous documents being stored directly in a machine-readable format that can be processed with LLMs. However, it is becoming increasingly difficult to protect this intellectual property, especially when data are shared. In addition, recent improvements in AI-generated text make it increasingly challenging to distinguish between text written by a human and text generated by an LLM [1]. This can make it more complicated to maintain control over data, often discussed under the umbrella term *data sovereignty* [2], [3], [4]. Information-hiding techniques such as digital watermarking or steganography can help to address these concerns by hiding information in a cover document [5], [6], [7], [8], [9].

Researchers and practitioners have developed a variety of methods for hiding data in different files, such as image, video, audio, or text files, of which the latter are the most challenging [7], [10]. Those techniques can be classified as *watermarking* or fingerprinting methods, which focus on robustly storing copyright information and securing

The associate editor coordinating the review of this manuscript and approving it for publication was Peter Langendoerfer.

intellectual property, and *steganography* methods, which focus on encoding secret information imperceptibly [5], [6], [7], [11]. To hide information in text, format-based methods shift words or lines or change fonts or text colors, whereas linguistic-based methods replace words with synonyms or generate new text [8], [12]. Recently, researchers have proposed and discussed the possibilities for watermarking and steganography in LLMs [13], [14], [15], [16], [17]. This is motivated by political regulations like the European Union's AI Act, which states that AI-generated texts must be detectable and mentions watermarks and fingerprints as possible solutions [18].

However, existing approaches show significant drawbacks in real-world use cases. For example, words cannot be replaced with synonyms in legal cover texts, as small semantic changes can have profound legal impacts. Format-based methods that employ line shifts or font colors cannot be applied to plain text files. Other methods using specific characters like Unicode confusables or spaces without width, called zero-width characters, are either visually recognizable by humans or not robust in different applications when using simple copy-and-paste tampering attacks [9].

To close this gap, we extend the method proposed in [19] and call it Innamark. This method can hide any byte-encoded string inside a cover text by replacing all classical whitespace characters with a curated set of similar-looking whitespace characters from the Unicode standard [20]. This leads to a text with a secret message hidden inconspicuously inside. This approach is applicable to plain text in different file formats due to its format independence, and the secret message remains when copied and pasted without alerting users. Our design can support optional functionalities like compression, encryption, hashing, or error-correcting codes. This makes it usable for both watermarking and steganography use cases. To sum up, the paper makes the following concrete contributions:

- We propose Innamark, an information-hiding technique that can hide any byte-encoded sequence inside a cover text.
- We provide a Kotlin library as a reference implementation.
- We introduce InnamarkTags, a specific secret message structure, to enable compression, encryption, hashing, and error-correcting codes to strengthen the security and robustness further.
- We evaluate the capacity, imperceptibility, and robustness of the proposed method.
- We conduct a comparison with related work in an implemented testbed of ten algorithms.

The remaining sections of this paper are structured as follows: The necessary background information and nomenclature are introduced in Section II. An overview of similar algorithms and concepts is presented in Section III. Our proposed embedding and extraction technique is detailed in Section IV, including information about our reference implementation. The results of a benchmark evaluation are presented in Section V and discussed in Section VI. The paper concludes in Section VII, with additional information in the Appendix.

TABLE 1. Digital watermarking definitions.

Definition	Reference
"A digital watermark can be described as a visible or an invisible, preferably the latter, identification code that permanently is embedded in the data."	[24, p. 230]
"In digital watermarking, relevant information is embedded in an imperceptive way into a digital document. The embedded information is called a watermark."	[25, p. 19]
"Digital watermarking technology is a typical information hiding method, which covers text, image and video."	[26, p. 1311]

II. BACKGROUND

This section introduces the terminology and notation needed to understand this work.

A. WATERMARKING AND STEGANOGRAPHY

Distinguishing between the different terminologies used in the domains of *cryptography* and *information hiding* is crucial and has been discussed comprehensively in the literature [5], [9], [11], [21], [22]. Starting with the broadest concept, "information hiding is the science of concealing a secret message or watermark inside a cover media (a host file/message) for providing various security purposes such as content authentication, integrity verification, covert communication, and so on" [9, p. 56]. Information hiding can further be divided into steganography and watermarking. It is distinct from cryptography because, at its root, it is not about transforming plain text into an encrypted cipher text [21]. A detailed survey and classification of information-hiding techniques has been published in [5].

The process of hiding a secret message inside a cover medium "in a way that one cannot detect it" [7, p. 6367] is termed *steganography*. Techniques used to perform steganography can be classified into character-level, bit-level, and hybrid methods [23]. Another definition is as follows: "Steganography embeds a secret message inside an innocent looking cover medium, stealthily, without creating any attention. The cover medium used can be a text, image, audio, video, network packets, etc." [23, p. 1].

In contrast, *digital watermarking* focuses on inserting "a visible or an invisible, preferably the latter, identification code that permanently is embedded in the data" [24, p. 230]. Watermarked data assets can range from images, audio, and video to the under-researched and most difficult cover medium of plain text [10], often referred to as digital watermarking, text watermarking, or digital text watermarking. An overview of the definitions that have arisen over time is presented in Table 1.

To bring everything together, Rizzo et al. summarize the concepts as follows: "While cryptography algorithms make unreadable the information by applying a kind of

permutation or substitution to the original content, the steganography algorithms provide techniques to hide new information into the carrier, that is a readable document. Whereas watermarking algorithms ensure the authentication and the copyright protection by applying a watermark to the digital content." [11, p. 97]

1) CLASSIFICATION

The existing methods can be further classified according to the following attributes based on [5], [22], [26]:

- *Cover Medium:* The type of data in which the secret message is embedded, such as image, audio, video, or text files [5].
- *Imperceptibility:* The secret message can be embedded visibly or invisibly inside the cover [22], [26].
- *Blindness:* A blind method allows the secret message to be extracted without knowledge of the original cover, while a non-blind method requires the original cover for extraction [22], [26].
- *Robustness:* A robust method can withstand attacks and intentional or unintentional modifications, whereas a fragile method can not [5], [22].

Our proposed method is an invisible and blind method that uses text as the cover medium, aiming for high robustness.

B. NOTATION AND UNICODE WHITESPACE CHARACTERS

In this subsection, we introduce the notation used in this paper to describe our proposed method, which follows [27]. All iterable elements in the algorithmic descriptions, like lists, are indexed starting at one. Since the information-hiding scheme is based on plain text, let u be a member of the set $\mathcal{U} := \{u : u$ is Unicode character$\}$ of the 154 998 characters in version 16.0.0 of the Unicode standard [20]. Let s be a Unicode whitespace character with positive width, and let $\mathcal{S} := \{s : s$ is space character $\wedge\, s \in U\}$ be the set of all 17 such characters in this version of the standard. A character like the zero-width space (U+200B) "although called a 'space' in its name, does not actually have any width or visible glyph in display (...) and is treated as a format control character, rather than as a space character" [20, p. 326] and, therefore, is not an element of $\mathcal{S}$. We define the classical and most commonly used space character U+0020 as $\delta \in \mathcal{S}$. We initially evaluated every s in [19] to define our own subset of whitespace characters as the alphabet $\mathcal{A}_+ := \{a : a \in S \wedge a \in U \wedge a$ meets criteria$\}$, where the criteria are *non-noticeability* for humans and *robustness* in different applications and file formats.

For the *non-noticeability* or visibility criteria, we compared the widths of the whitespace characters, taken from [28], with that of the standard space δ ($\approx 1/4$ em). If abnormalities are present that cause unusual space, we classify the character as having a different visibility from δ, depicted as "✗" in Table 2. When the difference is not noticeable, "✓" is shown ($\approx 1/3$ to $1/5$ em), whereas "(✗)" indicates that the difference could be noticeable to human eyes (e.g., $1/2$ or $1/6$ em).

TABLE 2. Whitespace evaluation based on [28].

Name	Code	Visibil.	.txt	.docx	.pdf	Mail	Teams
Space	U+0020	✓	✓	✓	✓	✓	✓
No-Break Space	U+00A0	✓	✓	✗	✗	✗	✗
Ogham Space Mark	U+1680	✗	✓	✓	✓	✓	✓
En Quad	U+2000	✗	✓	✓	(✓)	✓	✓
Em Quad	U+2001	✗	✓	✓	(✓)	✓	✓
En Space	U+2002	✗	✓	✓	✗	✓	✓
Em Space	U+2003	✗	✓	✓	✗	✓	✓
Three-per-Em Space	U+2004	✓	✓	✓	(✓)	✓	✓
Four-per-Em Space	U+2005	✓	✓	✗	✗	✗	✓
Six-per-Em Space	U+2006	(✗)	✓	✓	(✓)	✓	✓
Figure Space	U+2007	✗	✓	✓	(✓)	✓	✓
Punctuation Space	U+2008	✓	✓	✓	(✓)	✓	✓
Thin Space	U+2009	✓	✓	✓	(✓)	✓	✓
Hair Space	U+200A	(✗)	✓	✓	(✓)	✓	✓
Narrow No-Break Space	U+202F	✓	✓	✓	(✓)	✓	✓
Medium Mathematical Space	U+205F	✓	✓	✓	(✓)	✓	✓
Ideographic Space	U+3000	✗	✓	✓	✗	✓	✓

For the application and file type *robustness* criteria, we tested the different whitespace characters in text (.txt) files, Microsoft Word files (.docx), PDF files created using Word, emails, and Microsoft Teams Chat. All tests were executed on Windows, Linux (Ubuntu), and macOS for operating system independence. These file types and programs were selected because they are considered industry standards for office and collaboration tools in many fields, except phone calls and SMS [29].

Only the three-per-em space (U+2004), the punctuation space (U+2008), the thin space (U+2009), the narrow no-break space (U+202F), and the medium mathematical space (U+205F) are not noticed by humans and are robust in most of our tested applications and file formats. Therefore, these form the whitespace homoglyph alphabet $\mathcal{A}_+$; their names are bold in Table 2. Thus, we can see that $\mathcal{A}_+ \subset \mathcal{S} \subset \mathcal{U}$. Our proposed Innamark technique, introduced in Section IV, uses four of the five elements of $\mathcal{A}_+$ to encode the secret message because one element $\phi \in A_+$ is used as a separator character. We denote the alphabet without the separator character as $\mathcal{A}_-$, where $\mathcal{A}_+ = \mathcal{A}_- \cup \{\phi\}$ and $\phi \notin \mathcal{A}_-$. Depending on the final application, the modular design of our proposed technique allows the usage of other characters in $\mathcal{A}_+$ if a different set is needed due to framework restrictions or the requirements of specific use cases. A summary of the notation is presented in Table 3.

III. RELATED WORK

Over time, a variety of information-hiding algorithms and implementations for watermarking and steganography have been published. Due to the increasing diversity of methods, cover media, and applications, several literature reviews and surveys have been published to organize the cluttered research and application landscape. A more detailed discussion of selected methods is provided in section III-A.

TABLE 3. Overview of notation, following [27].

Symbol	Meaning
CT	Cover text
SM	Secret message text
SM_{bytes}	Byte representation of SM as a sequence of digits
SM_H	Hidden whitespace representation of SM_{bytes}
CT_{SM}	Cover text containing a hidden secret message
$\mathcal{U}$	Set of all Unicode characters
$\mathcal{S}$	Set of all 17 Unicode space characters
δ	Classical Unicode whitespace (U+0020)
ϕ	Separator whitespace character
$\mathcal{A}_+$	Whitespace alphabet with ϕ
$\mathcal{A}_-$	Whitespace alphabet without ϕ
θ	Option parameter for encryption, compression, etc.
$Emb(CT, SM_{bytes}, \theta)$	Embedding algorithm
$Ext(CT_{SM})$	Extraction algorithm

Bender et al. [30] provided one of the first comprehensive overviews of data-hiding methods for image, audio, and text files. Later on, Petitcolas et al. [5] published a survey on information-hiding techniques, focusing on steganography, watermarking, and fingerprinting, including information about possible attacks and a basic overall theoretical framework. More specialized overviews with solutions focusing on text steganography have been provided by Ahvanooey et al. [31], Krishnan et al. [23], and Majeed et al. [8]. Current challenges are discussed by Ahvanooey et al. [9] and Tyagi et al. [32], with the latter considering concrete application possibilities.

In addition to reviews, researchers have formally compared existing methods to identify their strengths and weaknesses. Ahvanooey et al. [33] compared watermarking and steganography methods by differentiating their embedding techniques and evaluating them according to the criteria of imperceptibility, embedding capacity, robustness, security, and computational cost [33]. One of the latest evaluations of text steganography methods was published by Knöchel and Karius [34], who compared their capacity, imperceptibility, robustness, and complexity with a specialized focus on malware.

A. RELATED METHODS

To situate our proposed Innamark method within the research landscape, we consider the most relevant text watermarking and steganography methods. In the LLM problem domain initially set out, Kirchenbauer et al. [13] and Christ et al. [14] presented token-based watermarking schemes, with the latter focusing on undetectability, completeness, and soundness. These ideas were integrated into SynthID, the watermarking engine used by Google's Gemini LLM [16]. Steinebach [15] generated a text based on sets of letters. These methods are classified as linguistic since they make use of LLM text generation. Such methods are problematic for cover texts whose semantics are essential. Thus, we focus on format-based methods that use insertion- or substitution-based embedding techniques [34]. Other types of format-based methods, as well as linguistic and random or statistical generation methods [8], [34], are not considered further since either they do not work on plain text documents due to the lack of formatting options or they change the semantics or structure of the cover text. We implemented all the methods presented here for our benchmark evaluation in Section V. A summary is provided in Table 4.

1) SNOW

One of the oldest whitespace steganography methods for ASCII texts is Steganographic Nature of Whitespace (SNOW) [35]. Although the first release goes back to the 20th century, the last update, with a change to the open-source Apache 2.0 license, was made in 2013. It has Java and Windows DOS versions and a C implementation last updated in 2016 [36]. The embedding process encodes the secret message into tab and space characters and appends it to the cover text, starting with a tab character under the consideration of a predefined line length [35]. Upstream compression and encryption can be enabled before the encoding process.

2) UNISPACH

A well-known algorithm in the field of information hiding for text documents is UniSpaCh, proposed by Por et al. [6]. It is an extended version of WhiteSteg, which replaces a single whitespace character between two words or paragraphs with either one or two characters to encode a zero or one [37]. UniSpaCh uses two different methods to embed the secret message in the text. For spaces between words and sentences, regular whitespace characters either remain as they are or are extended by adding a thin, six-per-em, or hair space to encode two bits per embedding location [6]. For end-of-line and inter-paragraph spacings, the remaining space is filled with a combination of hair, six-per-em, punctuation, and thin spaces to encode two bits per character [6].

3) AITSTEG

Ahvanooey et al. [21] proposed a text steganography technique for SMS or social media communication. The embedding method transforms the secret message into zero-width characters with a Gödel function and uses the sending/receiving time and the length of the secret message to insert it before the cover text.

4) SHIU ET AL.

The data hiding method proposed by Shiu et al. [38] focuses on communication over messengers of social media networks. Due to the small width of social media messaging windows, the method is based on a fixed line length and can hide three bits per line of the cover text [38]. After encoding a secret message into a bit stream based on the ASCII mapping, it embeds the first bit by adding a whitespace at the end of a line, changing the length of the line to embed the second bit, and adding a whitespace between two words to embed the third bit [38].

5) RIZZO ET AL.

A text-watermarking technique based on replacing Unicode characters with specific confusables, also known as homoglyphs, was initially proposed in [11] and extended to a fine-grain watermarking approach in [22]. The latter method generates a watermark by using a keyed hash function with a secret message as a watermark and a secret password [22]. Afterward, the watermark is embedded in the cover text by replacing specific characters with their confusables or leaving them unchanged to embed one bit in each and replacing spaces with a set of specific whitespace characters to embed three bits in each [22].

6) STEGCLOAK

The open-source implementation StegCloak published by [39], as described in [40], is a JavaScript steganography tool that is able to hide a secret message inside a cover text with optional password encryption and hash-based message authentication code (HMAC). In the embedding process, the secret message is compressed, optionally encrypted, and encoded in a set of zero-width characters to be inserted in one location after a classical whitespace of the cover text [40].

7) LOOKALIKES

Another implementation is the Unicode Lookalikes algorithm by [41] as part of the Python package pyUnicodeSteganography. Similar to [22], the method replaces specific characters with their confusables to encode a secret message inside the cover text [41].

8) COVERTSYS

Ahvanooey et al. [42] presented a multilingual steganography method focusing on short messages in social media networks. Like the previous approach in [21], four zero-width characters and a timestamp are used to encode the secret message. Further, a password-based approach with a one-time pad (OTP) and an XOR operation are used to transform the secret message into an encrypted bit stream that is appended to the cover text [42].

9) SHAZZAD-UR-RAHMEN ET AL.

The data-hiding approach of Shazzad-Ur-Rahmen et al. [43] can embed five bits per embeddable location, whereas their updated version [44] can embed six. The main idea of the latter procedure is to encrypt the secret message using AES and convert the resulting binary stream into blocks of six bits [44]. With the help of two lists, specific Unicode characters are replaced with their confusables, and whitespace characters are replaced with a particular combination of smaller whitespace characters to embed the secret message in the cover text [44].

IV. PROPOSED METHOD

In this section, we present Innamark, our invisible and blind information-hiding technique for plain text. Since existing methods either lack robustness in some applications, increase the number of characters, or are recognizable by humans, they are unsuitable for the LLM use case described in Section I. Therefore, this section presents the embedding and extraction method in Algorithm 1 and Algorithm 2 based on our nomenclature introduced in Section II-B. It further includes a concrete example based on a *Lorem ipsum* dummy text and information about our implemented prototypes.

Algorithm 1 Embedding $Emb(CT, SM_{bytes}, \theta)$

Input: Cover text ($CT := \{c_1, c_2, \ldots, c_n\}, \forall c \in \mathcal{U}$);
Secret message ($SM_{bytes} := \{q_1, q_2, \ldots, q_n\}$, $\forall q \in \{0, 1, \ldots, 255\}$);
Configuration parameter (θ)
Output: Cover text with hidden secret message (CT_{SM})

▷ *Insert tag*
$SM_{bytes}, SM_H \leftarrow applyTag(SM_{bytes}, \theta)$
▷ *Encode secret message*
$d \leftarrow \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil$
for each $q \in SM_{bytes}$ **do**
  for $i \leftarrow 1$ **to** d **do**
    $r \leftarrow q \mod |\mathcal{A}_-|$
    $q \leftarrow \left\lfloor \frac{q}{|\mathcal{A}_-|} \right\rfloor$
    $SM_H \leftarrow SM_H + a_{r+1}$ ▷ $a_{r+1} \in \mathcal{A}_-$
  end for
end for
▷ *Insert secret message*
$i \leftarrow 0$
for each $c \in CT$ **do**
  if $c = \delta$ **then**
    if $i = 0$ **then**
      $CT_{SM} \leftarrow CT_{SM} + \phi$
      $i \leftarrow i + 1$
    else if $0 < i \leq |SM_H|$ **then**
      $w_{H_i} \leftarrow SM_{H_i}$ ▷ $w_{H_i} \in \mathcal{A}_-$
      $CT_{SM} \leftarrow CT_{SM} + w_{H_i}$
      $i \leftarrow i + 1$
    else
      $i \leftarrow 0$
    end if
  else
    $CT_{SM} \leftarrow CT_{SM} + c$
  end if
end for
return CT_{SM}

A. EMBEDDING

The proposed embedding method, detailed in Algorithm 1, can hide any byte-encoded sequence in a Unicode-encoded cover text CT. The examples in this paper are based on a secret message SM in text form, in which every character is transformed into its UTF-8 byte representation SM_{bytes}.

TABLE 4. Overview of Related Methods.

Name	Release	Publication	Techniques
SNOW [35], [36]	Before 1998	Docum., Software	Appending additional tabs and whitespace characters at the end of the text.
UniSpaCh [6]	2012	Paper	Adds small whitespace characters between words and sentences and fills up lines and inter-paragraph spacings.
AITSteg [21]	2018	Paper	Hides data at the beginning of the cover text by using symmetric-key encoding and a transformation into zero-width characters.
Shiu et al. [38]	2018	Paper	Hides data line-wise by either changing the line length or adding whitespace between words or at the end of the line.
Rizzo et al. [11], [22]	2016/2019	Paper	Replaces whitespace and other Unicode characters with their confusables.
StegCloak [39], [40]	2020	Blog post, Software	Inserts the secret message at one location in the cover text using zero-width characters.
Lookalikes [41]	2021	Software	Replaces a specific set of Latin characters with their Unicode confusables.
CovertSYS [42]	2022	Paper	Adds zero-width characters at the end of the cover text by using the current date and time and an OTP.
Shazzad-Ur-Rahman et al. [43], [44]	2021/2023	Paper	Replaces confusables and changes whitespace to a specific combination of small whitespace characters.
Innamark [19], [45]	2025	Paper, Software	Replaces whitespaces with a specific set of robust, similar-looking whitespace characters.

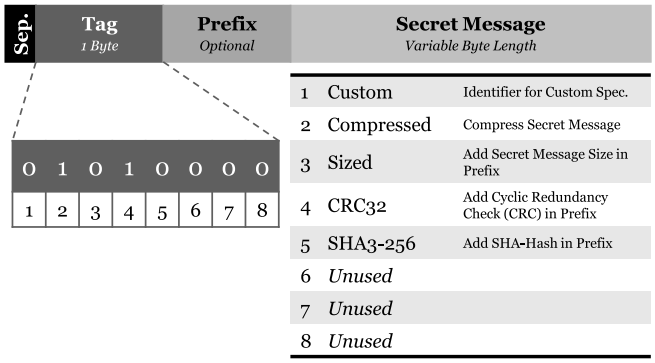


FIGURE 1. InnamarkTag structure.

The embedding function $Emb(CT, SM_{bytes}, \theta)$ begins by analyzing the configuration parameter θ. This specifies the type of the secret message, which we call a *InnamarkTag*. It offers optional functionalities that can be enabled in any combination by the end user, namely, encryption, compression, hashing, and error-correcting codes. Depending on the user's choice, θ defines the type used to calculate the *tag*. The tag has a fixed length of one byte and is returned by the *applyTag* method with the updated secret message, depending on the user's choice. Each bit in the tag indicates whether an option, such as compression, is enabled (1) or not (0). Thus, the tag describes the format of the secret message, similar to tags or headers in network packets [46]. Fig. 1 provides an overview of the InnamarkTag structure, the definition of each bit in the tag, and an example of a tag with enabled compression (second compression bit set to one) and error correction (fourth CRC32 bit set to one).

An InnamarkTag starts after the separator character ϕ, with the tag having a fixed length of one byte, followed by an optional prefix whose content depends on the tag, followed by the secret message itself.

Next, the encoding process starts transforming the secret message with the InnamarkTag and optional prefixes SM_{bytes} into a sequence of whitespace characters SM_H from our homoglyph alphabet $\mathcal{A}_{\_}$. Since $|\mathcal{A}_{\_}| = 4$, where $|\cdot|$ denotes the cardinality, each byte of the secret message is represented by four elements of $\mathcal{A}_{\_}$ because

$$\left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_{\_}|} \right\rceil = 4. \tag{1}$$

To fully include SM_H in a cover text CT, the number of standard space characters δ in CT must be at least the number of elements of the hidden secret message SM_H:

$$|\{x \in CT : x = \delta\}| \geq |SM_H|. \tag{2}$$

The final version of the cover text with the hidden secret message CT_{SM} is created by replacing all δ successively with the elements of SM_H. If, on the one hand, the input cover text CT does not have any normal whitespace δ, e.g., if the algorithm has already been applied to it, it is not possible to embed the secret message. If, on the other hand, CT has more δ characters than $|SM_H|$, the insertion process starts again until all δ are replaced to include the secret message multiple times. Thus, the resulting text does not contain standard space characters:

$$\forall x \in CT_{SM} : x \in \mathcal{U} \wedge x \neq \delta. \tag{3}$$

The use of multiple insertions improves the robustness of the information-hiding scheme, as changes do not necessarily destroy the secret message. The entire embedding algorithm $Emb(CT, SM_{bytes}, \theta) = CT_{SM}$ is presented in Algorithm 1.

An example of the proposed embedding algorithm is shown in Fig. 2 for the process and in Fig. 6 for input/output comparison. The secret message SM = "John" is to be hidden inside the cover text CT = "Lorem ipsum ..." with an empty configuration parameter θ to use the default InnamarkTag without compression, hashing, or error correction. Each character of SM is encoded into its byte representation SM_{bytes} and transformed into the whitespace alphabet. In this case, the first letter "J" of the secret

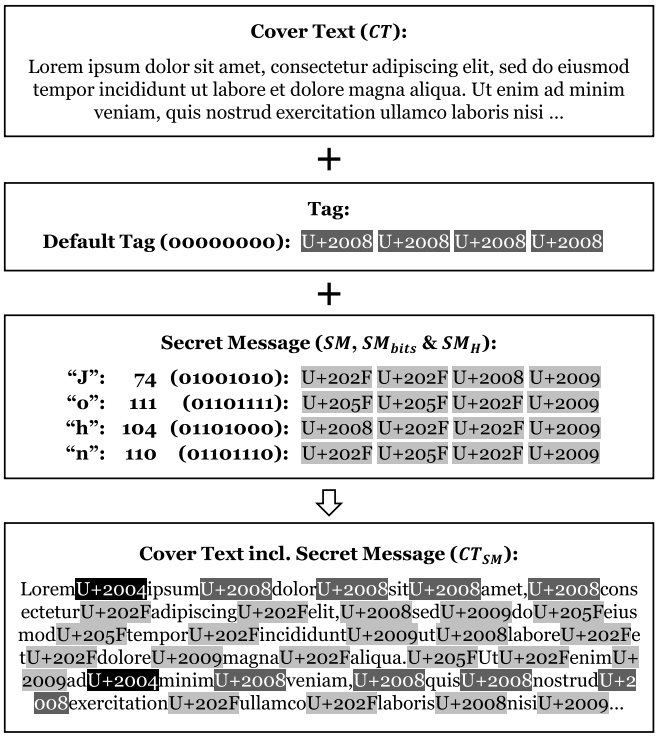


FIGURE 2. Example embedding with a default tag.

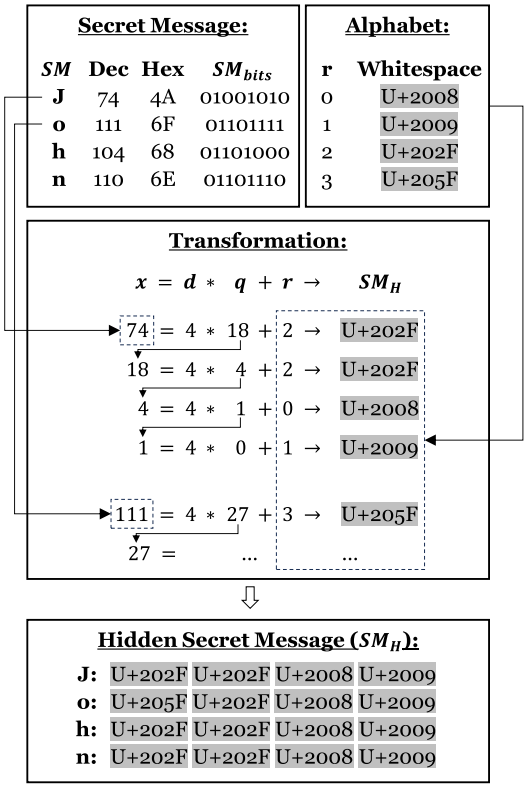


FIGURE 3. Example of transforming a secret message into the whitespace alphabet.

message is represented by the UTF-8 hexadecimal value 4A (U+004A), which equals the decimal value 74. Afterward, each value of SM_{bytes} is encoded into SM_H by transforming it into the alphabet $\mathcal{A}_-$ with a loop-based modulo operation, as described in Algorithm 1. This uses the length of the alphabet without separator characters as the divisor $d = |\mathcal{A}_-| = 4$, as illustrated in Fig. 3. The remainder of each division operation indicates the index of the whitespace in $\mathcal{A}_-$ needed to build the complete sequence of whitespace characters SM_H as a representation of the secret message SM. Next, each whitespace character δ of CT is replaced with the corresponding space in SM_H. Since CT has more whitespace characters than needed, the insertion process starts again with the separator character ϕ, represented by the black U+2004 in Fig. 2. Afterward, it inserts the default tag and the first character "J" of the secret message a second time and stops after all whitespace characters have been replaced.

B. EXTRACTION

The overall extraction method $Ext(CT_{SM}) = SM_{bytes}$ is split into three parts: extraction of the encoded message, tag analysis, and decoding. The first part starts by iterating over all characters of the input text, including the secret message CT_{SM}, until it finds the first occurrence of ϕ as the separator character. Through the filtering of $\mathcal{A}_+$, it extracts the full InnamarkTag, an optional prefix, and a hidden secret message SM_H, as specified in Fig. 1.

The second part analyzes and evaluates the entire InnamarkTag by calling *analyzeTag*(). If hashing is enabled, it checks and verifies the hash of SM_H and returns an error if problems occur. It can also decompress the message, check the size, or apply the CRC32 prefix.

The third part decodes the hidden secret message SM_H into its byte representation SM_{bytes}. The step size for the decoding part depends on the length of the alphabet without the separator character and is defined as

$$d := \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil \tag{4}$$

with $d = 4$ for our alphabet because each byte is represented by four whitespace characters from $\mathcal{A}_-$. The result of the cascading modulo operation from Algorithm 1 can be transformed back into its byte representation b. All b form the secret message representation SM_{bytes}, which in turn can be converted into the UTF-8 representation of the decoded secret message text SM. The overall extraction process is summarized in Algorithm 2.

C. IMPLEMENTATION

The watermark embedding and extraction methods were implemented as a generic library in the Kotlin programming language to test and validate them. Kotlin was chosen since it is interoperable with the widely used Java programming language while supporting multiplatform targets. Thus, our implementation can be used in applications supporting the Java virtual machine (JVM) and in JavaScript solutions due to the availability of both build targets.

Algorithm 2 Extraction $Ext(CT_{SM})$

Input: Cover text with a hidden secret message (CT_{SM})
Output: Extracted secret message bytes (SM_{bytes}) or error

▷ *Extract secret message*
for each $c \in CT_{SM}$ **do**
  if $c \in \mathcal{A}_+$ **then**
    if $c = \phi$ **and** $SM_H \neq \emptyset$ **then**
      break
    else if $c \in \mathcal{A}_-$ **then**
      $SM_H \leftarrow SM_H + c$
    end if
  end if
end for
▷ *Analyze tag and prefix*
$SM_H, error \leftarrow analyzeTag(SM_H)$
if error **then**
  return error
end if
▷ *Decode secret message*
$d \leftarrow \left\lceil \frac{\log_2 2^8}{\log_2 |\mathcal{A}_-|} \right\rceil$
for $i \leftarrow 0$ **to** $|SM_H|$ step d **do**
  for $y \leftarrow 0$ **to** $d - 1$ **do**
    $a_{r+1} \leftarrow SM_{H_{i+y+1}}$ ▷ $a_{r+1} \in \mathcal{A}_-$
    $b \leftarrow b + r \cdot d^y$ ▷ $r \in [0, \ldots, d-1]$
  end for
  $SM_{bytes} \leftarrow SM_{bytes} + b$
end for
return SM_{bytes}

To test the library, we developed a command-line interface (CLI) tool for the JVM that can embed and extract a byte-encoded string into another string or text-based document, such as a plain text file. Additionally, a web interface was implemented as a second usage example for the JavaScript build target. This front end acts as a graphical user interface and is likewise able to embed and extract watermarks in cover texts.

The source code of our implementations is made publicly available in [45] to ensure full transparency and applicability.

V. EVALUATION AND EXPERIMENTAL RESULTS

To analyze and evaluate our proposed Innamark technique, this section compares it with state-of-the-art methods for text watermarking and steganography.

Several empirical research methodologies are employed in software engineering, including simulations, benchmarks, case studies, and controlled experiments [47]. We use benchmarking in this evaluation since it is a "standard tool for the competitive evaluation and comparison of competing systems or components according to specific characteristics" [48, p. 333]. Numerous types of benchmarks exist, but we focus on *specification-based benchmarks* since they concentrate on a business problem and require development work before running the benchmark [48], [49].

We based our benchmark on criteria from existing comparisons and evaluations because benchmarks should be developed by the community instead of a single researcher [47], [50]. This leads to the following set of criteria, also used in related work [8], [33], [34]:

- **Capacity:** Describes the embedding amount as a relationship between the length of the secret message and the cover text.
- **Imperceptibility:** Also known as invisibility, refers to the visual and perceived differences between a text with a secret message and one without.
- **Robustness:** A broad term mainly focuses on how reliably the secret message stays inside the cover text in different environments or when attacks are carried out.

Runtime or execution speed is not benchmarked since the evaluation aims to differentiate the methods according to their core properties rather than efficiency. In the following, we introduce our experimental setup and dataset and present the results of an evaluation conducted for each criterion.

A. EXPERIMENTAL SETUP AND DATASET

We created a testbed and implemented all relevant existing methods presented in Section III-A and Table 4 in the Java programming language on the basis of the published descriptions, reference implementations, and examples. To ensure a uniform basis and comparability without unnecessary overhead, we consider the embedding and extraction methods only. Optional functionalities like encryption or compression are excluded since they can be applied upstream to all algorithms. Fig. 4 shows our implemented evaluation graphical user interface (GUI), consisting of the embedding functionality on the left side, a drop-down box to select the method in the middle with some additional options, and the extraction tool on the right.

For the benchmark evaluation, we used a large dataset of 1 000 000 random English Wikipedia articles as cover text. We tried to embed the largest possible secret message in each of the articles. For imperceptibility and robustness, we applied each algorithm in two execution runs on the 1 000 000 cover texts. The first run tried to hide a short four-character secret message inside the dataset, whereas the second run tried to hide a long 455-character secret message. The different lengths and forms of the Wikipedia cover texts, as well as the two different lengths of secret messages, ensure a comprehensive evaluation. More details about the data and evaluation process for transparency and reproducibility are provided in the Appendix.

B. CAPACITY

The embedding capacity analysis provides information about the relationship between the length of the cover text and that of the longest secret message that can be embedded within it. We distinguish the algorithms in Table 4 into two types:

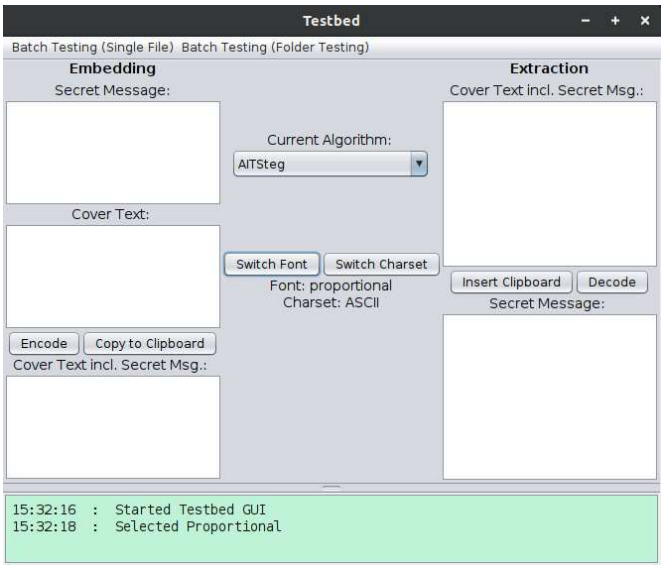


FIGURE 4. Evaluation GUI for embedding and extraction.

1) **Bounded-capacity algorithms:** These have a limited embedding capacity depending on the length, structure, and used characters in the cover text and secret message, primarily due to specific replacements or insertions.
   Algorithms: Shiu et al., Innamark, Rizzo et al., Lookalikes, Shazzad-Ur-Rahman et al., UniSpaCh.
2) **Unbounded-capacity algorithms:** These have no general limit on their embedding capacity, primarily due to the use of (zero-width) characters. Embedding capacity restrictions only apply if the text length is specified, like SMS or X (formerly Twitter) posts.
   Algorithms: AITSteg, CovertSYS, StegCloak, SNOW.

In our testbed, we analyzed all bounded-capacity algorithms by executing each on the dataset of 1 000 000 Wikipedia articles. This process started by hiding a one-byte secret message inside the cover and then repeatedly increasing the length of the secret message until an error occurred, yielding the maximum number of embeddable bytes. This process cannot be executed for the unbounded-capacity algorithms, which can theoretically embed a message of any length inside a given cover text.

We based the calculations on the approach described by Rizzo et al.: "The embedding capacity is computed as the average ratio between the number of embedded bits and the number of characters in each document" [22, p. 13]. Our dataset has an average cover text size of ∼2514 characters. The resulting capacity ratios are depicted in Fig. 5.

We assign all unbounded-capacity algorithms a value of 1.0 to illustrate the unprescribed limit. Comparing the others, we find that Shiu et al. has the lowest embedding capacity, with $26/2514 \approx 0.01$ bits/character, and that UniSpaCh has the highest of $1983/2514 \approx 0.79$ bits/character. The other four bounded-capacity algorithms are close to each other and have an embedding capacity of around and below

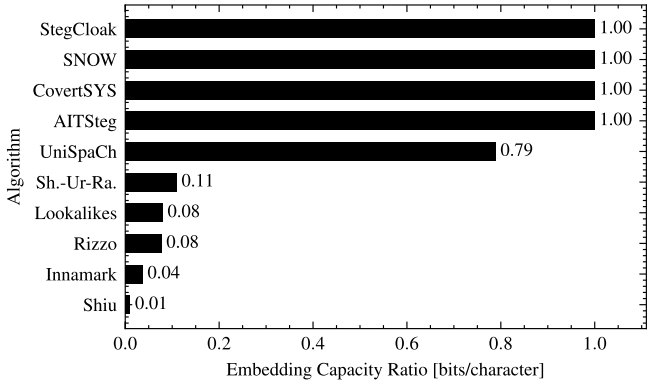


FIGURE 5. Maximum capacity evaluation results (higher values are better).

0.1 bits/character, whereas our proposed Innamark method has a capacity of $93/2514 \approx 0.04$ bits/character.

In a direct comparison, our results partly differ from previous studies because the benchmark evaluation depends strongly on the structure, format, and length of the input cover text and the secret message. In [22], a similar analysis with a different cover text dataset based on the New York Times Corpus resulted in an embedding capacity of 0.321 bits/character for UniSpaCh. UniSpaCh embeds the secret message between words, sentences, and paragraphs, with a significant amount hidden between paragraphs. Since [22] only used cover texts with a single paragraph as input, UniSpaCh's strength in embedding between paragraphs was not considered in that work.

C. IMPERCEPTIBILITY

Imperceptibility or invisibility is the ability for a secret message to be concealed in a cover document without causing any visible abnormalities [21]. Therefore, this benchmark is more relevant for steganography use cases than watermarking [6], [11]. It is essential to distinguish between imperceptibility to humans and to machines, which can detect differences, for example, by using statistical metrics. Due to the varying perceptions of humans, this benchmark "is the most subjective of all the metrics" [34, p. 121]. Fig. 6 shows an illustrative example of our Innamark algorithm as a direct comparison of a plain cover text and a cover text with the embedded secret message "John" in Microsoft Office Word version 2408 using the default font "Calibri (Body)" in font size 11. To further analyze and compare the imperceptibility of all algorithms, we use four different measurement metrics in the following, namely:

1) the Jaro–Winkler Similarity;
2) the number of characters;
3) the file size;
4) caret navigation.

1) JARO–WINKLER SIMILARITY

A standard numerical measurement used to compare the similarity between two character sequences is the

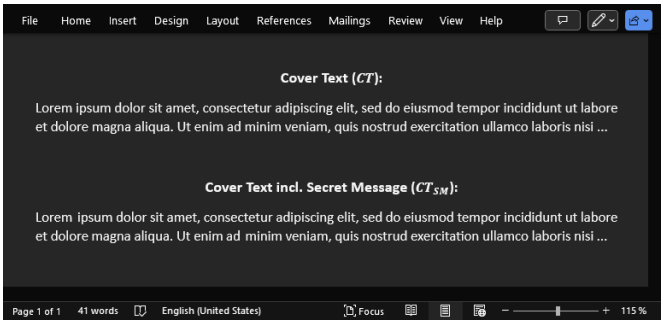


FIGURE 6. Innamark comparison example in Microsoft Word.

Jaro–Winkler similarity, also known by the misleading name "Jaro–Winkler distance" [51]. It is often used to evaluate information-hiding techniques (see [27], [31], [42], [52], [53], [54]). The benchmark is based on the Jaro string comparator Φ, for which a value of 1 indicates that two strings s_1 and s_2 are identical and 0 indicates that the strings have no common characters [55]:

$$\Phi(s_1, s_2) = \begin{cases} 1 & : s_1 = s_2 \\ \frac{1}{3}\left(\frac{c}{|s_1|} + \frac{c}{|s_2|} + \frac{c-\tau}{c}\right) & : m > 0 \\ 0 & : \text{otherwise.} \end{cases} \quad (5)$$

Here, $|s_1|$ and $|s_2|$ are the lengths of the strings, c is the number of matching characters, τ is the number of transpositions based on the characters and m for all matching characters [51], [55]. The newer Jaro–Winkler similarity Φ_n from [55] builds on the Jaro similarity:

$$\Phi_n(s_1, s_2) = \Phi(s_1, s_2) + i \cdot 0.1 \cdot (1 - \Phi(s_1, s_2)). \quad (6)$$

It adds the scaling factor 0.1 and a prefix length i to compare the first characters of the strings [55].

In our testbed, we used a Java implementation in version 1.12.0 of the Apache Commons Text package [56] to calculate the Jaro–Winkler similarity of a plain cover text and a text with an integrated secret message. Like in [34], Fig. 7 shows the average Jaro–Winkler similarity for each algorithm and for both short (four-character) and long (455-character) secret message execution runs. For short secret messages, UniSpaCh shows the best results with a Jaro–Winkler similarity of $\Phi_n = 0.996$, whereas Rizzo et al. has the worst result with $\Phi_n = 0.729$. Our proposed Innamark technique has the best similarity of $\Phi_n = 0.931$ for longer messages, whereas AITSteg has the lowest similarity value of $\Phi_n = 0.362$. It is noteworthy that the unbounded algorithms without capacity restrictions have a comparatively high difference in Φ_n between short and long secret messages.

2) NUMBER OF CHARACTERS

A change in the number of characters can reveal that a document has embedded hidden content, leading to poorer imperceptibility. This may be detected automatically by

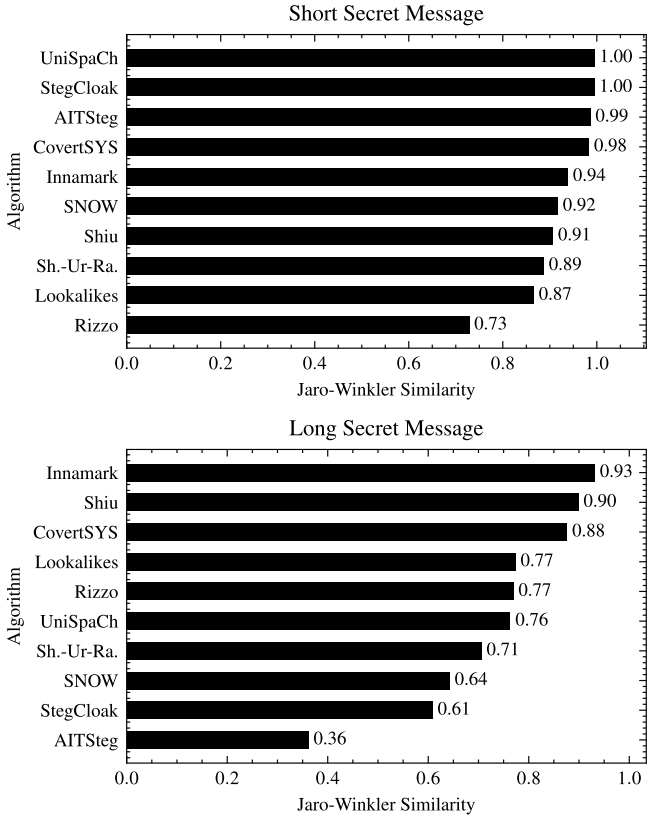


FIGURE 7. Jaro–Winkler similarity evaluation results (higher values are better).

software or by a human, for example, when submitting a text document to a publisher that imposes a character limit.

In our testbed, we compared the number of characters of the original cover text with a text that included an embedded secret message. The mean absolute differences Δ computed across the dataset are depicted for both long and short messages in Fig. 8.

Only the pure one-to-one replacement techniques Innamark, Lookalikes, and Rizzo et al. do not show a difference. The four unbounded algorithms and UniSpaCh show a significant increase in the number of characters due to the addition of zero-width or small whitespace characters, making them recognizable. The negative value of $\Delta = -8.957$ for Shiu et al. is caused by the design of the hiding algorithm. It replaces whitespace characters after a specific length with newlines and removes formatting characters like tabulators, decreasing the number of characters for short secret messages.

3) FILE SIZE

Humans and systems can detect a text file with a hidden secret message by its increased file size. In particular, suspicions may be raised if a file with a small amount of text has a large file size. Thus, Majeed et al. [8] argue for developing data-hiding methods that create results with minimal file sizes.

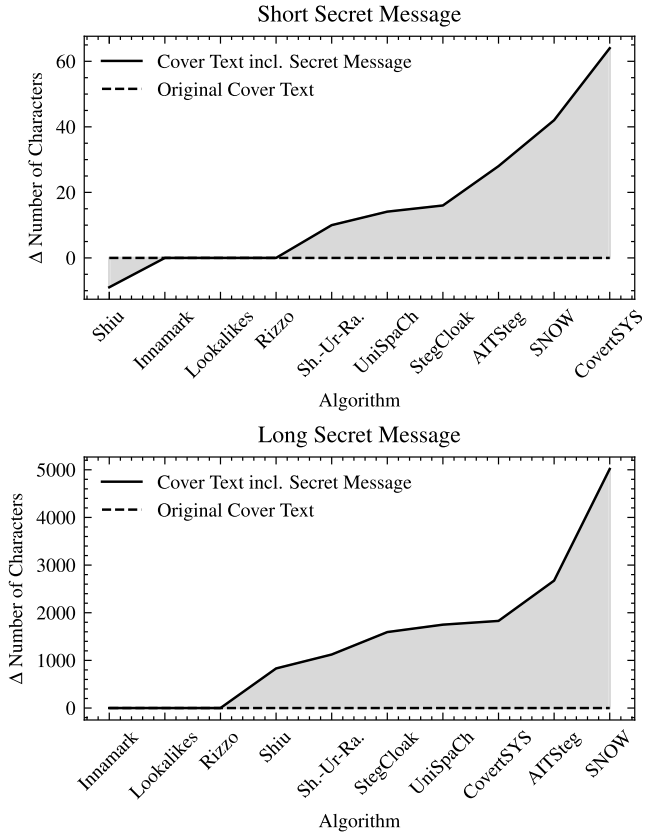


FIGURE 8. Character size evaluation results (smaller absolute values are better).

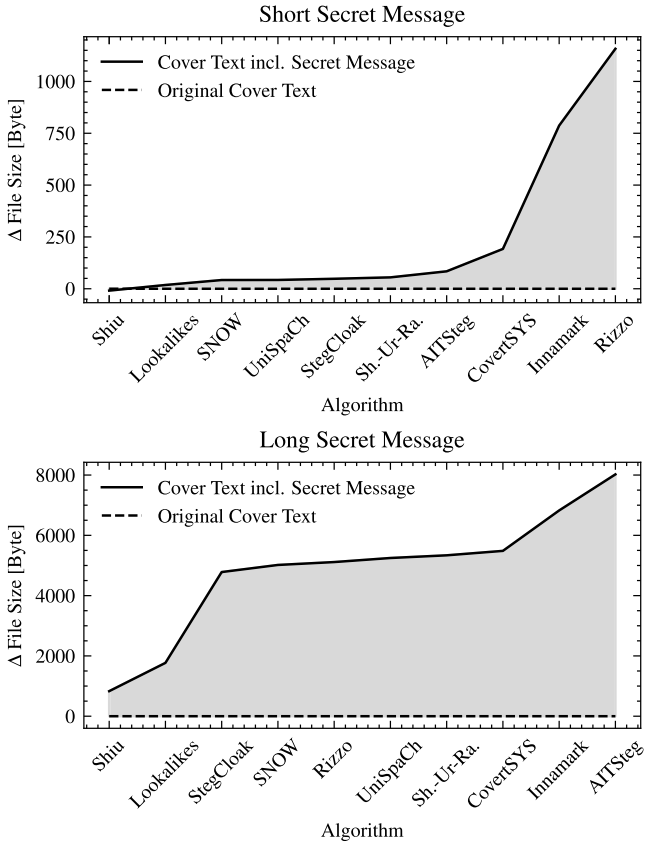


FIGURE 9. File size evaluation results (smaller absolute values are better).

In our testbed, we compared the differences in file size between the original cover text and the text including a secret message. Although these benchmark criteria are closely related to the previous analysis of the number of characters, the results differ due to varying character storage requirements. For example, Shazzad-Ur-Rahman et al.'s [43], [44] method replaces a classical small Latin letter "g" (U+0067) with the visually similar-looking mathematical alphanumeric symbol U+1D5C0 as part of the letterlike symbols in the Unicode standard [20]. Whereas the first needs one byte of storage space in the UTF-8 encoding, the second needs four bytes. Fig. 9 illustrates the mean absolute differences based on the two runs on our 1 000 000-article dataset. The negative value for short secret messages from the Shiu et al. algorithm indicates a decreased file size. This can be attributed to replacing whitespace characters with newlines, as seen in the results for the number of characters. Our proposed Innamark technique is just in the lower third in direct comparison because it is one of the few algorithms that embeds the secret message multiple times in the cover text for increased robustness.

4) CARET NAVIGATION

The last imperceptibility metric relates to suspicious behavior when navigating through a digital text document with the arrow keys on a computer keyboard. The *caret* is the blinking pipe in a text field that indicates the cursor position. If, for example, a document contains zero-width characters and the caret arrives at such a character when arrow keys are used for navigation, the caret moves one zero-width character further but visually stays at the same position, raising suspicion. Similar abnormalities occur when multiple small whitespace characters are used.

In our testbed, we manually checked various texts with secret messages. We moved the caret from the start to the end to identify any unusual behavior. In cases of hanging caret positions or identifying unexpected characters, we classified the algorithm as visible because the caret attack revealed the secret message's position. Table 5 shows an overview of the results with a justification for each algorithm as to why it is noticeable by users or not. It was found that only three algorithms (Rizzo et al., Lookalikes, and Innamark) are invisible through caret navigations.

D. ROBUSTNESS

The last primary benchmark criterion relates to the persistence of a secret message inside the text. We focus on the following two types of robustness checks:

1) modification robustness (insertion, replacement, deletion);
2) usage robustness (retyping, formatting, file type and application change via copy and paste).

TABLE 5. Caret navigation attack evaluation.

Name	Invisible	Reason
SNOW	✗	Adds multiple whitespace characters at the end of the text.
UniSpaCh	✗	Adds small spaces between words, sentences, and paragraphs.
AITSteg	✗	Adds zero-width characters at the beginning of the text.
Shiu et al.	✗	Adds whitespace characters at specific positions to encode 0 or 1.
Rizzo et al.	✓	Only replaces characters without adding additional ones.
StegCloak	✗	Adds zero-width characters in one position in the text.
Lookalikes	✓	Only replaces letters without adding additional characters.
CovertSYS	✗	Adds zero-width characters at the end of the text.
Shazzad-Ur-Rahman et al.	✗	Replaces letters and single whitespace characters with combinations of multiple whitespace characters.
Innamark	✓	Only replaces whitespace characters without adding additional characters.

1) MODIFICATION ROBUSTNESS

Modifications to a file that may compromise the secret message, such as text insertions or deletions, are also known as tampering attacks in the information-hiding literature [33]. We have analyzed replacements, each consisting of a deletion followed by an insertion, for all implemented algorithms in batch runs for both the short and long secret messages on the 1 000 000-article dataset. For each article, we embedded the secret message and then replaced a block of 10% of the length of the original cover text with just one letter. The starting position of the replacement was randomly chosen using seeding for reproducibility. A success rate was calculated, indicating the percentage of articles for which a secret message was successfully extracted after modification. Next, the process was repeated with replacements of 20% up to 90% in steps of 10 percentage points. This resulted in a total of 180 000 000 processed articles for the test (10 algorithms × 9 modification percentages × 2 execution runs for short and long secret message × 1 000 000 articles). The results are presented in Fig. 10.

For short secret messages, our proposed Innamark technique proved to be the most robust method in our testbed, even at high modification levels. Lookalikes and Shiu et al. had a particularly low success rate, but the other algorithms show a similar linear robustness trend. For long secret messages, all tested techniques are roughly similar, with StegCloak showing the best results because the algorithm embeds the secret message at a single position.

2) USAGE ROBUSTNESS

The usage robustness considered several aspects of how users work with text. We tested various usage scenarios classified as attacks on the information-hiding schemes on the full testbed of all implemented algorithms. The results are summarized in Table 6 and explained in the following.

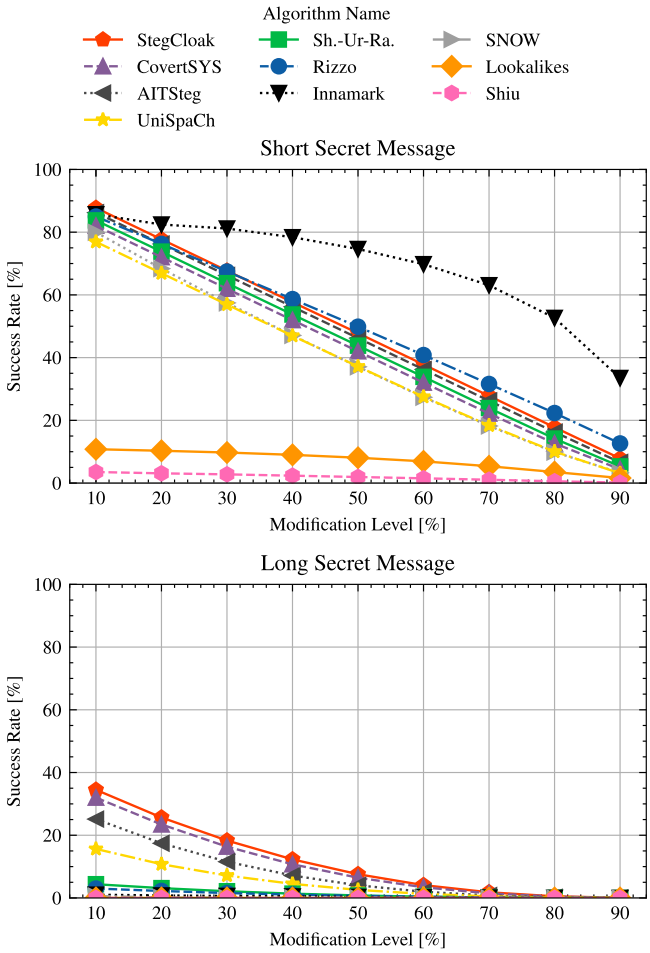


FIGURE 10. Modification robustness (higher values are better).

First, we analyzed a text reproduction or retyping attack [24], [33], in which users manually type the text with an embedded secret message in a new document. None of the algorithms could withstand this attack since they all use specific Unicode symbols that are lost on retyping.

Second, we applied different formatting changes and checked whether the secret message could be extracted afterward [9], [33]. We embedded a secret message inside a cover, made the text bold, and changed the font, color, and size. The secret message could be successfully extracted in all cases because all algorithms work with Unicode symbols that are not affected by such styling attacks. Nevertheless, changing the font can affect the perceptibility of the secret messages if the algorithm uses specific characters unsupported by the font family.

Third, and in accordance with the whitespace evaluation mentioned in Section II-B, we analyzed the robustness of a cover text with an embedded secret message when used in different applications and file formats. Simple copy and paste operations form "one of the most common attacks in that the malicious users copy the whole of text and paste into their own files" [33, p. 7]. Following [43], we extended the set of business-related targets to social media applications, namely, WhatsApp, Facebook Messenger, and X (formerly Twitter).

TABLE 6. Usage robustness.

Algorithm	Retyping	Formatting	.txt	.docx	.pdf	Mail	Teams	WhatsApp	Facebook Msg.	X/Twitter
SNOW	✗	✓	✓	✓	✗	✗	✗	✗	✗	✗
UniSpaCh	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓
AITSteg	✗	✓	✓	✗	✗	✗	✓	✓	✓	✗
Shiu et al.	✗	✓	✓	✓	✗	✓	✓	✓	✓	(✓)
Rizzo et al.	✗	✓	✓	✗	✗	✗	(✓)	✓	✓	✗
StegCloak	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓
Lookalikes	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓
CovertSYS	✗	✓	✓	(✓)	✗	✗	✓	✓	✗	✗
Shazzad-Ur-Rahman et al.	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓
Innamark	✗	✓	✓	✓	(✓)	✓	✓	✓	✓	✓

In our testbed, we applied each algorithm to a *Lorem ipsum* dummy text to hide a secret message in the cover text. We copied each result into the relevant application and checked whether the secret message could still be extracted after copying it back to our testbed GUI. In Table 6, "✓" indicates that the secret message could fully be extracted, whereas "✗" indicates a corrupted output. Edge cases are depicted as "(✓)," like CovertSYS in a .docx document, in which additional characters were shown in the extracted result but the original secret message could still be recognized. Only our proposed Innamark technique worked partially in the PDF format, depending on the PDF viewer used. For example, some whitespace characters were replaced with a standard U+0020 space when copying the content from Adobe Acrobat Reader, whereas they remained the same with PDF24 Reader. In such cases, "(✓)" is shown in Table 6 because the respective whitespace characters remain in the original PDF file but the robustness depends on the PDF viewer used.

The results show that some applications remove specific characters like the four-per-em space (U+2005), which do not work in file types like .docx and PDF or in emails (see Table 2) but are used by algorithms like Rizzo et al. [22]. Further, the tested messenger software and social media networks often remove trailing whitespace characters from messages, which is why SNOW and CovertSYS encountered problems. Only our proposed Innamark technique was robust in all tested applications and file types.

VI. DISCUSSION

This section discusses our proposed Innamark scheme on the basis of the experimental evaluation results, considering its limitations and future research directions. We evaluated and compared it against a testbed of ten algorithms from the literature, conducting tests on a dataset of 1 000 000 articles. To our knowledge, the reported Innamark algorithm is the first method that can hide a secret message inside a cover text without increasing the number of characters or being noticed by humans while ensuring robustness to replacing portions of text and copying it between applications.

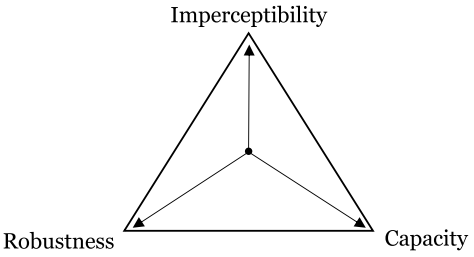


FIGURE 11. Requirement trade-off based on [57].

In general, it should be emphasized that there is no one-fits-all information-hiding scheme with high robustness, embedding capacity, and imperceptibility. Information-hiding techniques often strive for high embedding capacities, which often conflict with imperceptibility criteria since imperceptibility decreases if the embedding capacity increases [58]. This can be illustrated with a *trade-off triangle* [57], shown in Fig. 11. Therefore, researchers and practitioners need to select an appropriate algorithm for their use case under the consideration of boundary conditions and application scenarios. Following [34], we provide an overview in Fig. 12 based on the evaluation results to support the decision-making process.

An unbounded algorithm like AITSteg, CovertSYS, StegCloak, or SNOW should be selected if a high embedding capacity is essential. Our Innamark algorithm is a favorable choice if data are often transferred between different applications and both robustness and imperceptibility are important.

A. LIMITATIONS AND FUTURE WORK

Nevertheless, our approach has limitations that need to be discussed in future work.

First, Innamark's small embedding capacity is a weakness. The structure of an InnamarkTag, as shown in Fig. 1, is designed to enable optional compression that can help to increase the capacity. Work is in progress to increase the capacity further by using compression libraries or developing hybrid approaches that use non-printable characters. Furthermore, the impact of using different InnamarkTag options needs to be analyzed to determine to what extent the error correction improves robustness or how hashing affects the embedding capacity.

Second, the proposed solution is based on and tested on the Unicode standard and the UTF-8 scheme. Future research should consider the influence of other encoding schemes like UTF-16 and ISO-standardized Latin-1 (ISO 8859-1), and the potential impacts on the process of recoding to a small scheme like ASCII.

Third, the algorithm is robust to the actions of typical users who do not recognize documents with hidden secret messages. However, people familiar with the strategy can use smart attacks to apply targeted destruction of the secret message, for example, the random replacement of $\mathcal{A}_+$. This can be eliminated by using a different random subset of $\mathcal{A}_+$

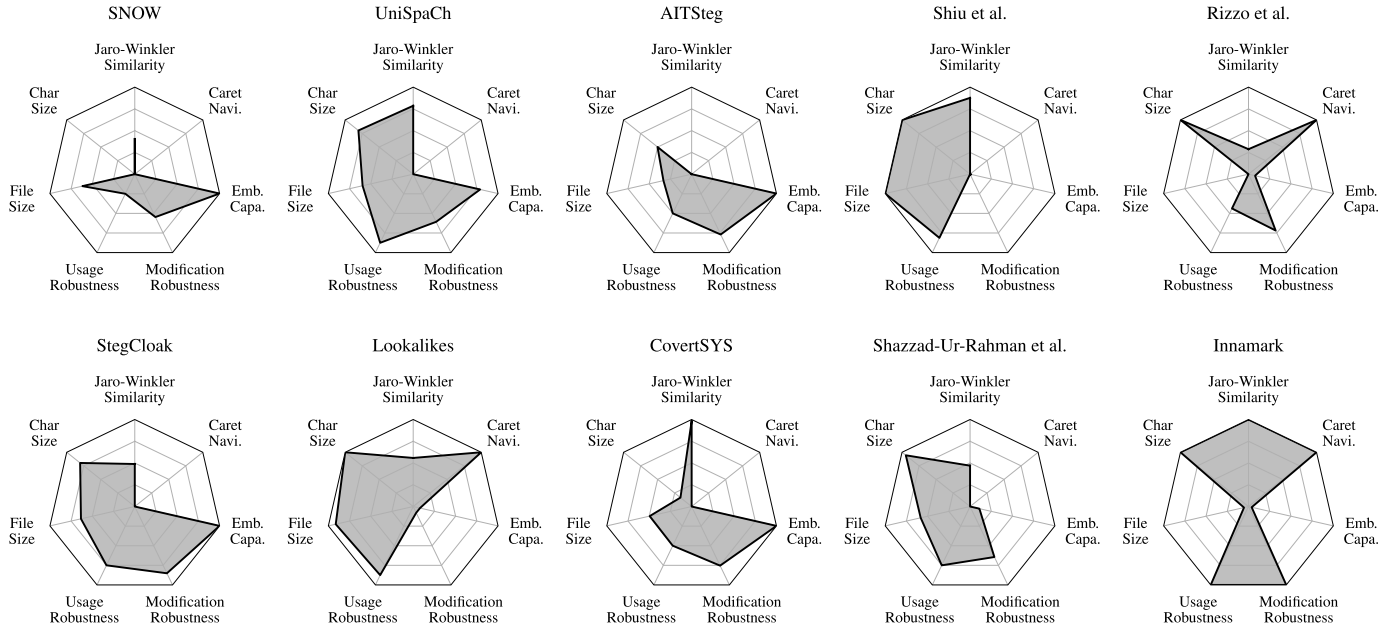


FIGURE 12. Summarizing evaluation comparison inspired by [34].

for every embedding operation. With the help of the modular InnamarkTag structure, future research can use smart analysis in encoding to increase the noise tolerance and allow the extraction and restoration of broken secret messages.

Fourth, printing and rescanning attacks using optical character recognition (OCR) may destroy secret messages. However, most spaces have slightly different widths that are not recognizable by human eyes, so machines may restore the original whitespace characters from a scan if configured correctly. This strongly depends on the font used, the applied OCR technique, and physical conditions like the scan quality.

Fifth, accessibility concerns may arise if a text with a secret message is read by screen readers or automatically translated into another language. Future research is needed to analyze how those accessibility tools handle specific Unicode characters. This emphasizes the need for a deeper analysis of related text processing tools, such as machine learning-based natural language processing (NLP) units, minifiers, compressors, or deobfuscation tools.

VII. CONCLUSION

We have designed and implemented Innamark, a blind and invisible information-hiding technique that can embed byte-encoded sequences inside a cover text. Although several solutions for digital text watermarking and steganography have been published in recent years, existing approaches change the semantics or style of the cover text, increase the number of characters, or lack robustness against the output being copied into different applications. By encoding and mapping a secret message into our embedding alphabet of five Unicode whitespace characters, we can embed the information in the cover text by substituting all whitespace characters. The specified structure of our InnamarkTag has been designed to enable additional functionalities like compression, encryption, hashing, and error correction. The experimental evaluation shows strengths in imperceptibility and robustness, with limitations in embedding capacity based on a direct benchmark comparison with ten algorithms. Our method can help LLM operators fulfill regulations like the European Union's AI Act [18] and can assist businesses in securing sensitive data before it is shared with external parties, especially if they are concerned about robustness to copying between applications. Work is in progress to increase the algorithm's embedding capacity and enable the restoration of secret messages broken by text alterations.

APPENDIX
EVALUATION DATA

To evaluate, compare, and benchmark our proposed Innamark method against existing solutions, we performed two batch runs with each algorithm. The first used the English example name "John" as a short secret message. The second used the following 455-character *Lorem ipsum* dummy text as a long secret message: "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum."

The set of cover texts consists of 1 000 000 random English Wikipedia articles, chosen because they are publicly available texts that have different lengths and structures and cover a wide range of domains. We used the cleaned article version from Hugging Face based on the dump from the Wikimedia

Foundation [59]. To reproduce our randomized selection, the articles' IDs are available from the corresponding author upon request and can be mapped back to the original texts and URLs.

REFERENCES

[1] M. Christ and S. Gunn, "Pseudorandom error-correcting codes," in *Proc. Adv. Cryptol. (CRYPTO)*, Jan. 2024, pp. 325–347, doi: 10.1007/978-3-031-68391-6_10.

[2] M. Jarke, B. Otto, and S. Ram, "Data sovereignty and data space ecosystems," *Bus. Inf. Syst. Eng.*, vol. 61, no. 5, pp. 549–550, Oct. 2019, doi: 10.1007/s12599-019-00614-2.

[3] M. Hellmeier and F. von Scherenberg, "A delimitation of data sovereignty from digital and technological sovereignty," in *Proc. ECIS Res. Papers*, 2023, pp. 1–19. [Online]. Available: https://aisel.aisnet.org/ecis2023_rp/306

[4] F. von Scherenberg, M. Hellmeier, and B. Otto, "Data sovereignty in information systems," *Electron. Markets*, vol. 34, no. 1, Dec. 2024, Art. no. 15, doi: 10.1007/s12525-024-00693-4.

[5] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding—A survey," *Proc. IEEE*, vol. 87, no. 7, pp. 1062–1078, Jul. 1999, doi: 10.1109/5.771065.

[6] L. Y. Por, K. Wong, and K. O. Chee, "UniSpaCh: A text-based data hiding method using unicode space characters," *J. Syst. Softw.*, vol. 85, no. 5, pp. 1075–1082, May 2012, doi: 10.1016/j.jss.2011.12.023.

[7] M. H. Alkawaz, G. Sulong, T. Saba, A. S. Almazyad, and A. Rehman, "Concise analysis of current text automation and watermarking approaches," *Secur. Commun. Netw.*, vol. 9, no. 18, pp. 6365–6378, Dec. 2016, doi: 10.1002/sec.1738.

[8] M. A. Majeed, R. Sulaiman, Z. Shukur, and M. K. Hasan, "A review on text steganography techniques," *Mathematics*, vol. 9, no. 21, p. 2829, Nov. 2021, doi: 10.3390/math9212829.

[9] M. T. Ahvanooey, M. X. Zhu, W. Mazurczyk, and M. Bendechache, "Information hiding in digital textual contents: Techniques and current challenges," *Computer*, vol. 55, no. 6, pp. 56–65, Jun. 2022, doi: 10.1109/MC.2021.3113922.

[10] F. Bertini, S. G. Rizzo, and D. Montesi, "Can information hiding in social media posts represent a threat?" *Computer*, vol. 52, no. 10, pp. 52–60, Oct. 2019, doi: 10.1109/MC.2019.2917199.

[11] S. G. Rizzo, F. Bertini, and D. Montesi, "Content-preserving text watermarking through unicode homoglyph substitution," in *Proc. 20th Int. Database Eng. Appl. Symp. (IDEAS)*, 2016, pp. 97–104, doi: 10.1145/2938503.2938510.

[12] Y. Zhang, C. Huang, S. Liu, L. Huang, T. Yang, X. Zhang, and H. Wu, "Screen-shooting resistant robust document watermarking in the discrete Fourier transform domain," *Int. J. Netw. Manage.*, vol. 35, no. 1, Jan. 2025, Art. no. e2278, doi: 10.1002/nem.2278.

[13] J. Kirchenbauer, J. Geiping, Y. Wen, J. Katz, I. Miers, and T. Goldstein, "A watermark for large language models," in *Proc. 40th Int. Conf. Mach. Learn.*, Jan. 2023, pp. 17061–17084. [Online]. Available: https://proceedings.mlr.press/v202/kirchenbauer23a/kirchenbauer23a.pdf

[14] M. Christ, S. Gunn, and O. Zamir, "Undetectable watermarks for language models," in *Proc. 37th Conf. Learn. Theory*, Jan. 2023, pp. 1125–1139. [Online]. Available: https://proceedings.mlr.press/v247/christ24a/christ24a.pdf

[15] M. Steinebach, "Natural language steganography by ChatGPT," in *Proc. 19th Int. Conf. Availability, Rel. Secur.*, Jul. 2024, pp. 1–9, doi: 10.1145/3664476.3670930.

[16] S. Dathathri et al., "Scalable watermarking for identifying large language model outputs," *Nature*, vol. 634, no. 8035, pp. 818–823, Oct. 2024, doi: 10.1038/s41586-024-08025-4.

[17] Z. Xu, R. Xu, and V. S. Sheng, "Beyond binary classification: Customizable text watermark on large language models," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jun. 2024, pp. 1–8, doi: 10.1109/ijcnn60899.2024.10650062.

[18] Eur. Commission, Brussels, Belgium. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13, Jun. 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC), no. 300/2008, (EU), no. 167/2013, (EU), no. 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)*. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2024/1689

[19] M. Hellmeier, H. Qarawlus, H. Norkowski, and F. Howar, "A hidden digital text watermarking method using unicode whitespace replacement," in *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, 2025, pp. 7411–7420. [Online]. Available: https://hdl.handle.net/10125/109738

[20] Unicode Consortium, San Francisco, CA, USA. (2024). *The Unicode Standard, Version 16.0.0*. [Online]. Available: https://www.unicode.org/versions/Unicode16.0.0/

[21] M. T. Ahvanooey, Q. Li, J. Hou, H. D. Mazraeh, and J. Zhang, "AITSteg: An innovative text steganography technique for hidden transmission of text message via social media," *IEEE Access*, vol. 6, pp. 65981–65995, 2018, doi: 10.1109/ACCESS.2018.2866063.

[22] S. G. Rizzo, F. Bertini, and D. Montesi, "Fine-grain watermarking for intellectual property protection," *EURASIP J. Inf. Secur.*, vol. 2019, no. 1, Dec. 2019, Art. no. 10, doi: 10.1186/s13635-019-0094-2.

[23] R. B. Krishnan, P. K. Thandra, and M. S. Baba, "An overview of text steganography," in *Proc. 4th Int. Conf. Signal Process., Commun. Netw. (ICSCN)*, Mar. 2017, pp. 1–6, doi: 10.1109/ICSCN.2017.8085643.

[24] Z. Jalil and A. M. Mirza, "A review of digital watermarking techniques for text documents," in *Proc. Int. Conf. Inf. Multimedia Technol.*, Dec. 2009, pp. 230–234, doi: 10.1109/ICIMT.2009.11.

[25] M. Qadir and I. Ahmad, "Digital text watermarking: Secure content delivery and data hiding in digital documents," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 21, no. 11, pp. 18–21, Nov. 2006, doi: 10.1109/MAES.2006.284353.

[26] Q. Wang, B. Yue, C. Wangdu, P. Xinghao, C. Zhipeng, S. Wang, Y. Wang, and C. Wang, "An overview on digital content watermarking," in *Proc. 8th Int. Conf. Signal Inf. Process. Netw. Comput.*, Jul. 2022, pp. 1311–1318, doi: 10.1007/978-981-19-3387-5_157.

[27] M. Taleby Ahvanooey, Q. Li, X. Zhu, M. Alazab, and J. Zhang, "ANiTW: A novel intelligent text watermarking technique for forensic identification of spurious information on social media," *Comput. Secur.*, vol. 90, Mar. 2020, Art. no. 101702, doi: 10.1016/j.cose.2019.101702.

[28] J. Korpela. (2002). *Unicode Spaces*. Accessed: Apr. 17, 2025. [Online]. Available: https://www.jkorpela.fi/chars/spaces.html

[29] DataReportal, GWI, and Meltwater. (2023). *Share of Professionals Worldwide Using Selected Communication Channels and Digital Tools for Work As of 3rd Quarter 2022, By Frequency*. [Online]. Available: https://www.statista.com/statistics/1306580/usage-communication-tools-for-work-worldwide-by-frequency/

[30] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3.4, pp. 313–336, 1996, doi: 10.1147/sj.353.0313.

[31] M. T. Ahvanooey, Q. Li, J. Hou, A. R. Rajput, and Y. Chen, "Modern text hiding, text steganalysis, and applications: A comparative analysis," *Entropy*, vol. 21, no. 4, Apr. 2019, Art. no. 355, doi: 10.3390/e21040355.

[32] L. K. Tyagi, A. Gupta, and A. Mohamed, "Unveiling the invisible an in-depth analysis of text steganography techniques, challenges, and advancement," in *Proc. 3rd Int. Conf. Technol. Advancements Comput. Sci. (ICTACS)*, Nov. 2023, pp. 177–183, doi: 10.1109/ictacs59847.2023.10390024.

[33] M. Taleby Ahvanooey, Q. Li, H. J. Shim, and Y. Huang, "A comparative analysis of information hiding techniques for copyright protection of text documents," *Secur. Commun. Netw.*, vol. 2018, pp. 1–22, Sep. 2018, doi: 10.1155/2018/5325040.

[34] M. Knöchel and S. Karius, "Text steganography methods and their influence in malware: A comprehensive overview and evaluation," in *Proc. ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2024, pp. 113–124, doi: 10.1145/3658664.3659637.

[35] M. Kwan. (2013). *How SNOW Works*. Accessed: Apr. 17, 2025. [Online]. Available: https://darkside.com.au/snow/description.html

[36] M. Kwan. (2016). *SNOW*. Accessed: Apr. 17, 2025. [Online]. Available: https://github.com/mattkwan-zz/snow

[37] L. Y. Por, T. F. Ang, and B. Delina, "WhiteSteg: A new scheme in information hiding using text steganography," *WSEAS Trans. Comput.*, vol. 7, no. 6, pp. 735–745, Jun. 2008. [Online]. Available: https://dl.acm.org/doi/10.5555/1458369.1458384

[38] H.-J. Shiu, B.-S. Lin, B. Lin, P.-Y. Huang, C. Huang, and C.-L. Lei, "Data hiding on social media communications using text steganography," in *Proc. 12th Int. Conf. Risks Secur. Internet Syst.*, Jan. 2018, pp. 217–224, doi: 10.1007/978-3-319-76687-4_15.

[39] StegCloak. (2020). *StegCloak*. Accessed: Apr. 17, 2025. [Online]. Available: https://github.com/KuroLabs/stegcloak

[40] Mohanasundar. (2020). *How to Hide Secrets in Strings—Modern Text Hiding in JavaScript*. [Online]. Available: https://blog.bitsrc.io/how-to-hide-secrets-in-strings-modern-text-hiding-in-javascript-613a9faa5787

[41] G. Thompson. (2021). *PyUnicodeSteganography Lookalikes*. Accessed: Apr. 17, 2025. [Online]. Available: https://github.com/bunnylab/pyUnicodeSteganography/blob/main/pyUnicodeSteganography/lookalikes.py

[42] M. T. Ahvanooey, M. X. Zhu, W. Mazurczyk, Q. Li, M. Kilger, K.-K.-R. Choo, and M. Conti, "CovertSYS: A systematic covert communication approach for providing secure end-to-end conversation via social networks," *J. Inf. Secur. Appl.*, vol. 71, Dec. 2022, Art. no. 103368, doi: 10.1016/j.jisa.2022.103368.

[43] M. Shazzad-Ur-Rahman, M. M. H. Ornob, A. Singha, M. S. Kaiser, and N. I. Akhter, "An effective text steganographic scheme based on multilingual approach for secure data communication," in *Proc. Joint 10th Int. Conf. Informat., Electron. Vis. (ICIEV), 5th Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR)*, Aug. 2021, pp. 1–8, doi: 10.1109/ICIEVicIVPR52578.2021.9564231.

[44] M. Shazzad-Ur-Rahman, M. S. Kaiser, M. B. Alam, and S. N. Nova, "A data hiding technique combining steganography and cryptography for secured communication," in *Proc. Int. Conf. Inf. Commun. Technol. Sustain. Develop. (ICICT4SD)*, Sep. 2023, pp. 432–437, doi: 10.1109/icict4sd59951.2023.10303563.

[45] Fraunhofer ISST. (2025). *Innamark*. Accessed: Apr. 17, 2025. [Online]. Available: https://github.com/FraunhoferISST/Innamark

[46] *Internet Protocol, Version 6 (IPv6)*, document RFC 8200, Internet Eng. Task Force, 2017.

[47] W. Hasselbring, "Benchmarking as empirical standard in software engineering research," in *Proc. Eval. Assessment Softw. Eng.*, Jun. 2021, pp. 365–372, doi: 10.1145/3463274.3463361.

[48] J. V. Kistowski, J. A. Arnold, K. Huppler, K.-D. Lange, J. L. Henning, and P. Cao, "How to build a benchmark," in *Proc. 6th ACM/SPEC Int. Conf. Perform. Eng.*, Jan. 2015, pp. 333–336, doi: 10.1145/2668930.2688819.

[49] S. Kounev, K.-D. Lange, and J. von Kistowski, *Systems Benchmarking*. Cham, Switzerland: Springer, 2020, doi: 10.1007/978-3-030-41705-5.

[50] S. E. Sim, S. Easterbrook, and R. C. Holt, "Using benchmarking to advance research: A challenge to software engineering," in *Proc. 25th Int. Conf. Softw. Eng.*, 2003, pp. 74–83, doi: 10.1109/icse.2003.1201189.

[51] J. M. Keil, "Efficient bounded jaro-winkler similarity based search," in *Proc. Datenbanksysteme Bus. Technol. Web (BTW)*, Jan. 2019, pp. 205–214, doi: 10.18420/btw2019-13.

[52] M. A. Majeed, R. Sulaiman, and Z. Shukur, "New text steganography technique based on multilayer encoding with format-preserving encryption and Huffman coding," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 12, pp. 163–172, 2022, doi: 10.14569/ijacsa.2022.0131222.

[53] M. A. Majeed, R. Sulaiman, and Z. Shukur, "New text steganography technique based on part-of-speech tagging and format-preserving encryption," *KSII Trans. Internet Inf. Syst.*, vol. 18, no. 1, pp. 170–191, 2024, doi: 10.3837/tiis.2024.01.010.

[54] S. S. Baawi, M. R. Mokhtar, and R. Sulaiman, "Enhancement of text steganography technique using Lempel-Ziv-Welch algorithm and two-letter word technique," in *Proc. 3rd Int. Conf. Rel. Inf. Commun. Technol. (IRICT)*, Sep. 2018, pp. 525–537, doi: 10.1007/978-3-319-99007-1_49.

[55] W. E. Winkler, "String comparator metrics and enhanced decision rules in the fellegi-sunter model of record linkage," *Proc. Sect. Surv. Res. Methods*, vol. 1990, pp. 354–359, Jan. 1990. [Online]. Available: https://files.eric.ed.gov/fulltext/ED325505.pdf

[56] Apache Softw. Found. *Apache Commons Text: Class JaroWinklerSimilarity*. Accessed: Oct. 31, 2024. [Online]. Available: https://commons.apache.org/proper/commons-text/apidocs/org/apache/commons/text/similarity/JaroWinklerSimilarity.html

[57] Y. Li, H. Wang, and M. Barni, "A survey of deep neural network watermarking techniques," *Neurocomputing*, vol. 461, pp. 171–193, Oct. 2021, doi: 10.1016/j.neucom.2021.07.051.

[58] Z.-L. Yang, X.-Q. Guo, Z.-M. Chen, Y.-F. Huang, and Y.-J. Zhang, "RNN-stega: Linguistic steganography based on recurrent neural networks," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1280–1295, May 2019, doi: 10.1109/TIFS.2018.2871746.

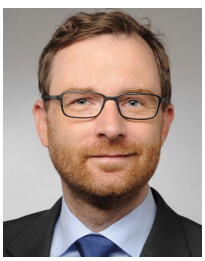
[59] Wikimedia. (2023). *Wikimedia Downloads*. [Online]. Available: https://huggingface.co/datasets/wikimedia/wikipedia

MALTE HELLMEIER received the B.Sc. degree in computer science/business information systems from Clausthal University of Technology, Clausthal-Zellerfeld, Germany, in 2018, and the M.Sc. degree in business information systems from Georg August University, Göttingen, Germany, in 2021. He is currently pursuing the Ph.D. degree with Technical University Dortmund, Germany. He is a Research Associate with the Applied Research Institute Fraunhofer ISST, Dortmund, Germany. His research interests include data sovereignty and information hiding, such as watermarking and steganography, with a strong focus on software engineering and software architecture.

HENDRIK NORKOWSKI received the B.Sc. degree in computer science from Technical University Dortmund, Dortmund, Germany, in 2020, and the M.Sc. degree in IT security/networks and systems from Ruhr University Bochum, Bochum, Germany, in 2024. He is currently a Software Developer with Montsecure, Bochum.

ERNST-CHRISTOPH SCHREWE received the B.Sc. degree in computer science from FernUniversität in Hagen, Hagen, Germany, in 2024, where he is currently pursuing the M.Sc. degree in practical computer science. He has been working full-time as a Technical Employee, focusing on software development at Fraunhofer ISST, Dortmund, Germany, since 2024.

HAYDAR QARAWLUS received the B.Sc. degree in information technology (IT) from American University of Iraq, Sulaymaniyah, in 2016, and the M.Sc. degree in computer science from Paderborn University, Paderborn, Germany, in 2020. He is currently pursuing the Ph.D. degree with Technical University Dortmund, Germany. He is currently a Research Associate with the Applied Research Institute Fraunhofer ISST, Dortmund, Germany. His research focuses on data sovereignty, data usage control, and steganography.



FALK HOWAR is currently a Professor of rigorous software engineering with the Department of Computer Science, Technical University Dortmund, Dortmund, Germany. He is also the Coordinator of software engineering research at Fraunhofer ISST, Dortmund. His research focuses on safe and trustworthy intelligent software systems with automated analysis, testing, and verification.

Paper VII

Table A.7 Metadata Overview of Paper VII

Title	A Fragile Watermarking Technique for Integrity Authentication of CSV-Files Using Invisible Line-Ending Control Characters
Authors	Florian Zimmer *Fraunhofer ISST, Dortmund, Germany* Malte Hellmeier *Fraunhofer ISST, Dortmund, Germany* Motoki Nakamura *Fujitsu Limited, Kanagawa, Japan* Tobias Urbanek *Fraunhofer ISST, Dortmund, Germany*
Publication Year	2025
Publication Type	Conference
Conference Name	22nd International Conference on Security and Cryptography (SECRYPT)
Conference Location	Bilbao, Spain
Conference Date	11. June 2025 - 13. June 2025
Publisher / Database	SCITEPRESS Digital Library
DOI / Link	https://doi.org/10.5220/0013559600003979
Status	Published
Ranking	VHB: - (2024 Rating) ICORE: C (2026 Rating) ERA: B (2010 Rating)
Comment	The paper was certified as a best paper candidate.

A Fragile Watermarking Technique for Integrity Authentication of CSV-Files Using Invisible Line-Ending Control Characters

Florian Zimmer[1][a], Malte Hellmeier[1][b], Motoki Nakamura[2][c] and Tobias Urbanek[1][d]
[1]*Fraunhofer Institute for Software and Systems Engineering ISST, Speicherstr. 6, 44147 Dortmund, Germany*
[2]*Data & Security Research Laboratory, Fujitsu Limited, Kanagawa, Japan*
{*florian.zimmer, malte.hellmeier, tobias.urbanek*}*@isst.fraunhofer.de, nakamura-motoki@fujitsu.com*

Keywords: Digital Watermarking, Fragile, CSV, Integrity, Authentication, Line-Ending, Unicode, CRLF, LF.

Abstract: Every day, a growing amount of data, including audio, video, images, and plain text, is published and shared online. Facilitating its interoperable exchange, a range of standards and formats has emerged, establishing common ground. Among plain text formats, CSV prevails as one of the most used text formats. However, being a simplistic, plain text format, it lacks built-in security measures. Consequently, data users cannot authenticate the integrity of CSV texts they receive. A recognised method in research for ensuring text integrity is fragile watermarking. Accordingly, numerous watermarking techniques are available for tamper detection. However, many of these methods are either incompatible with the CSV format or visible to the human eye. To address these shortcomings, we propose a novel fragile watermarking technique for CSV files. Using invisible line-ending control characters, we are able to embed any byte-encodable information into a CSV cover text, making it truly imperceptible. We evaluated our technique by conducting three experiments to benchmark robustness, capacity and imperceptibility and comparing it with existing solutions. We found that our technique successfully achieves complete imperceptibility in all cases. However, a limited capacity and line-ending normalisation sensitivity must be considered when applying it.

1 INTRODUCTION

The proliferation of an increasingly interconnected world has led to an ever-growing amount of data, with a projected growth to more than 394 zettabytes within the next five years (Taylor, 2024). Driven by the digital transformation, more and more digital assets are created, published, and shared over the internet every day, such as audio, video, images, or simply plain text (Rizzo et al., 2019). Moreover, active research in inter-organisational data sharing suggests that to fully utilise the value of data, it needs to be shared (Otto, 2022).

One of the most used plain text data formats besides HTML and PDF is the Comma-Separated Values (CSV) format. According to Vitagliano et al., CSV makes up to 31% of available formats on governmental portals (Vitagliano et al., 2023). Being simple in nature, the CSV format provides an easy way of storing, processing, and transferring data (Abba and Hassan, 2018). Especially for information exchange between heterogeneous systems and processing of raw data, CSV prevails to be a common choice due to its broad compatibility and lightweight processing capabilities (Ito, 2024).

However, as CSV is a plain text format focused on simplicity, it was not designed with security in mind and thus fails to provide any security features out of the box (Ito, 2024). This leaves data users of third-party CSV files exposed to various risks when using them. Following recent work, data integrity attacks are considered one of the most fundamental ones (Tian and Nogales, 2023), potentially resulting in financial losses up to human harm (Jaigirdar et al., 2019; Hisham et al., 2013). Therefore, as CSV does not provide any protection schemes itself, other solutions are needed.

Besides well-known security approaches and cryptography techniques, *watermarking* – especially *fragile watermarking* – has been identified as a potential solution aiming at mitigating data integrity risks. Therefore, various watermarking schemes have been proposed that aim to enable integrity

[a] https://orcid.org/0009-0002-8060-7162
[b] https://orcid.org/0000-0002-2095-662X
[c] https://orcid.org/0009-0004-3894-5023
[d] https://orcid.org/0009-0007-3121-0245

authentication by making the watermark susceptible to any changes made to the cover medium. For example, (He et al., 2020) proposed a novel watermarking technique for semi-structured text data such as JSON, XML, or CSV by embedding error-correction codes into the least significant bit (LSB) of numeric values. Other text watermarking schemes, on the other hand, often focus on homoglyph substitutions, as demonstrated in (Rizzo et al., 2016).

However, most of the existing watermarking approaches for integrity authentication are either not applicable for the CSV format or alter the values themselves. Consequently, recent work proposed a novel data hiding scheme which embeds a digital signature into a CSV cover text using alternating double quotation marks (Ito, 2024). The proposed scheme does not change the values themselves but exploits the syntactical definition of the CSV format. Yet, a significant drawback of this approach is that it is visible to its user, affecting the text's fidelity.

In this study, we aim to address the shortcomings of existing work by proposing a novel CSV fragile watermarking technique. Using invisible, non-printable line-ending control characters to embed a digital signature in a CSV cover text, our approach is imperceptible in nature. We demonstrate how our technique manages to incorporate any byte-encodable information into plain CSV text and how it can be utilised alongside digital signatures to authenticate the text's integrity. Furthermore, we evaluate our approach in an experimental setup and compare it with related work. Our main contributions include:

(i) A novel fragile watermarking technique for CSV text, outlining the embedding and extraction procedure.

(ii) An experimental setup used to evaluate and compare our approach with relevant related work.

The remainder is structured as follows: In Section 2, we present relevant background information and related work. Section 3 introduces our CSV watermarking approach. Section 4 describes the experimental setup and results. In Section 5, we discuss results and limitations. Section 6 concludes the study with a summary.

2 BACKGROUND

2.1 Watermarking

The idea of hiding data inside multimedia content goes back to the 20th century, with a substantial increase in academic publications since the 90s (Petitcolas et al., 1999). Since then, data hiding has mainly focused on proofing copyright and assuring the content integrity of digital media (Bender et al., 1996). Existing methods aim to hide a secret message (like a watermark or signature) inside a cover medium. Those cover mediums can range from images, text, audio, and video (Rizzo et al., 2019) to more specialised types like Word documents (Liu and Tsai, 2007), CSV files (Ito, 2024), or databases (Rani and Halder, 2022). An alternative delimitation is a classification into the categories of cryptography, steganography, and watermarking (Taleby Ahvanooey et al., 2018; Podilchuk and Delp, 2001; Rizzo et al., 2019). An overview of the interrelationship between the categories is shown in Figure 1.

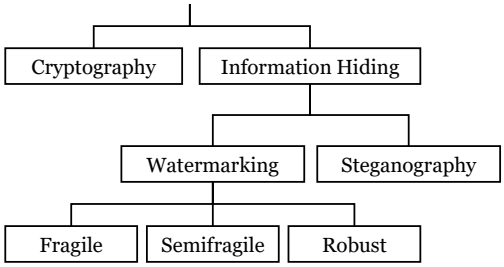


Figure 1: Information Hiding Classification (Podilchuk and Delp, 2001; Taleby Ahvanooey et al., 2018; Rizzo et al., 2019; Hellmeier et al., 2025).

Cryptography uses encryption and decryption techniques by working with cipher to focus on data hiding and data protection (Alkawaz et al., 2016). In contrast, *steganography* focuses on secure communication by hiding data invisibly to prevent third parties from detecting it (Hartung and Kutter, 1999). *Watermarking* aims to embed copyright information visibly or invisibly inside the cover (Jalil and Mirza, 2009; Kamaruddin et al., 2018). The latter can further be divided into *robust* techniques aiming for security and copyright protection, *fragile* techniques aiming for tamper detection, and *semi-fragile* techniques for something in between (Podilchuk and Delp, 2001; Alkawaz et al., 2016).

This work presents a fragile watermarking technique for CSV cover files, introduced in the following.

2.2 CSV Text Format

The CSV format is a plain text format for sharing, processing and storing tabular data (Abba and Hassan, 2018). CSV has been used for decades, especially

in the domain of databases. As CSV was developed out of need for an exchange format for heterogeneous systems, CSV was long lacking a sound definition or standard (Mitlöhner et al., 2016). This has led to various dialects and formats that persist to this day, resisting standardisation efforts.

In 2005, the IEFT published the RFC4180 specification, aiming at defining a common CSV format based on how CSV was predominantly used at the time. According to this, the most basic structure of CSV text is records of the form `aaa,bbb,ccc` delimited by a CRLF line-break. On top of that, they define a more sophisticated grammar on how to format different contents. Nevertheless, the resulting specification still acknowledges the variety of different dialects that exist by recommending that "Implementors should be conservative in what (they) do (and) be liberal in what (they) accept from others" (Shafranovich, 2005, p. 5) when adhering to this specification.

Roughly ten years later, in 2016, the W3C formalised a non-normative document, aiming at increasing the interoperability of CSV on the web by defining a data model for CSV, as well as by enhancing it with an additional metadata model. This was to enrich the overall capabilities of CSV, increasing the interoperability and overall accessibility of CSV documents by formally describing them (Brickely et al., 2016).

Nevertheless, the interoperability challenge associated with utilising the CSV format has persisted to this day. Therefore, recent work emphasised the need for either a more consistent usage of the CSV format or more robust processing and parsing tools (Mitlöhner et al., 2016; van den Burg et al., 2019; Vitagliano et al., 2023).

As the proposed watermarking scheme in this work makes use of the fact that different dialects exist, we investigate how major CSV tools handle our watermarked content more closely in Section 4.

2.3 Line-Ending Control Characters

Line-Ending control characters are non-printable control characters often found in text encoding standards such as Unicode[1] or ASCII[2]. Often referred to as *newline*, their function is to indicate the end of a line or the beginning of a new line, respectively (Allen, 2007). Being utilised for text formatting purposes only, they are invisible to regular users as they are non-printable.

There is a variety of different line-ending control characters for different text encoding standards. Unicode, e.g., defines eight different ones, with the most prominent ones being carriage return (CR), line feed (LF), and the combination of both carriage return and line feed (CRLF) (Allen, 2007). Moreover, different operating systems use different line-ending control characters by default. An overview can be seen in Table 1. The variety of different line-ending control characters originated from a time when typewriters required a combination of carriage return to move the cursor back to the beginning of the page and line feed to move the page up in order to continue writing in a new line. With the shift to the digital world, some operating systems, such as Windows, kept the combination of CRLF and others, such as UNIX-based operating systems, chose LF only (Saltzer and Ossanna, 1970; IEEE, 1986).

Table 1: Newline per Operating System (Allen, 2007).

Operating System	Newline
MacOS 9.x and earlier	CR
MacOS X	LF
Unix	LF
Windows	CRLF

Nowadays, many operating systems and tools are capable of handling both line-endings. This becomes evident, as many messaging protocols such as HTTP[3], MTP[4] or FTP[5] stipulate the usage of CRLF. Yet, most server operating systems which use those messaging formats for information exchange are UNIX-based (Fortune Business Insights, 2024). As a result, platform users are usually free to use both types of line-endings. However, most operating systems or tools usually normalise line-endings to the default platform-specific line-ending. E.g., Git[6], a major version control tool, provides an option which automatically converts CRLF into LF on any push.

In this work, we take advantage of the variety of accepted line-ending control characters and intentionally use a mixture of both CRLF and LF. We analyse any implications this might have in Section 4.

2.4 Related Work

According to (Liu et al., 2025), text watermarking approaches can roughly be divided into four categories: format-, lexical-, syntactic- or generation-

[1] https://www.ietf.org/rfc/rfc3629.txt [25.02.25]

[2] https://www.ietf.org/rfc/rfc20.txt [25.02.25]

[3] https://www.ietf.org/rfc/rfc2616.txt [25.02.25]

[4] https://www.ietf.org/rfc/rfc780.txt [25.02.25]

[5] https://www.ietf.org/rfc/rfc354.txt [25.02.25]

[6] https://www.git-scm.com/book/ms/v2/Customizing-Git-Git-Configuration.html [25.02.25]

based approaches. Our line-ending-based approach mainly falls into the category of format-based watermarking or, more specifically, into the subcategory of Unicode-based substitution.

Among the Unicode-based substitution techniques, one of the most notable approaches to mention is UniSpaCh (Por et al., 2012) and the proposed watermarking technique in (Rizzo et al., 2019). Both of them are representative of many more techniques which embed a watermark into a cover text by either replacing whitespaces or other confusables with similar-looking whitespaces or characters or by adding additional zero-width characters. Although this class of techniques could be adapted for use with CSV text, the applicability might be limited as these kinds of approaches strongly depend on text values. On top of that, these approaches would alter the text values themselves and are, therefore, not suitable for use cases where accuracy is an important requirement.

Therefore, in (He et al., 2020), a data protection scheme for semi-structured text is proposed, which allows not only for content integrity authentication but also for data recovery if the data has been tampered with. They achieve this by embedding error correction codes into the least-significant bits of whitelisted numeric values. In their work, they mainly focus on JSON text, yet they highlight the applicability to other semi-structured text data like CSV. However, even if it is negligible for some use cases, altering the least significant bit of numeric values might not be appropriate in others.

Consequently, in recent work, (Ito, 2024) proposed a novel embedding scheme in order to integrate digital signatures in CSV text, allowing for integrity authentication. They do this by exploiting the vague definition of the CSV format, which makes double quotes for values optional in most cases. Therefore, using alternating double quotes, they are able to embed a byte-encoded digital signature into CSV text.

A similar approach is followed in (Wen and Wang, 2013), which uses alternating double quotes and single apostrophes to enclose text values. Although their approach is demonstrated for XML text, it could also be adapted to CSV as well, given the variation of different dialects.

In contrast to other approaches, both of the latter succeed at leaving the values untouched. This way, they enable users to authenticate the content's integrity by verifying the digital signature included in the watermark. However, a major shortcoming of both approaches is that the changes introduced to the text are visible to the human eye, affecting its fidelity.

3 PROPOSED SOLUTION

In the following, we present our novel fragile CSV watermarking technique and describe the embedding and extraction procedures in detail. Our approach addresses the shortcomings of existing approaches, as it is imperceptible by using invisible, non-printable line-ending control characters. More specifically, using a combination of alternating CRLF and LF control characters, we are able to embed any byte-encodable information in a CSV cover text. We do this by mapping 0 or 1 to either control characters, respectively. As discussed in Section 2, most major platforms, as well as text encoding standards, are capable of handling both representations. Thus, most CSV editors and tools are able to display and parse a mixed set of line-endings, as we demonstrate in Section 4.

Furthermore, the following properties, acknowledged in literature, characterise our watermarking technique (Rizzo et al., 2016):

- *Fragile* - the fragility of a watermark is given if it is susceptible to any changes made to the cover text. In our case, the fragility is grounded on two facts: First, as mentioned in Section 2, mixed line-endings are prone to normalisation. Therefore, different tools tend to wipe the watermark if any changes are made, as we discuss in Section 4. Second, we integrate a digital signature into the watermark's content to make sure that the remaining modifications, which might not be detected by line-ending normalisation, are covered as well. This way, the recipient is able to securely authenticate the text's integrity. According to (Cox et al., 2000), doing this is a feasible approach to enable integrity authentication capabilities.
- *Invisible* - a watermark is invisible if it is hidden in the carrier text and does not appear to the user. As line-ending control characters are non-printable, this holds true for our approach.
- *Distortion-free* - a distortion-based watermarking technique introduces slight modifications to the data itself, whereas a distortion-free technique leaves the data itself untouched (He et al., 2020). As our approach keeps the values intact and rather alters the syntax within acceptable boundaries, our approach is considered distortion-free.
- *Blind* - a watermark is blind if the extraction procedure does not require the original cover text. As our approach is able to extract the content given solely the watermarked CSV text, it can be considered blind.
- *Secure* - a watermark technique is secure if it

adheres to Kerckhoffs' Law (Kerckhoffs, 1883). Specifically, if a malicious actor is aware of the embedding and extraction procedures, they still cannot read or alter the watermark's content without access to a private shared secret (Petitcolas et al., 1999). Since we aim to employ digital signatures in practice, this principle applies to our case.

In the following, both the embedding and extraction procedures are detailed. To establish them by general means, we consider a random bitstring as watermark content in our notions. However, the same can be applied to digital signatures or any other kind of information that can be encoded as bitstring.

3.1 Embedding Procedure

Using alternating CRLF and LF line-endings, our approach is capable of embedding any byte-encodable information into a CSV text. More precisely, we are able to embed a bitstring B of length m into a CSV cover text, where $B := \{b_0, b_1, \ldots, b_m\}$ and $b \in \{0,1\}$. The CSV text can be represented by a set of n rows R, where $R := \{r_0, r_1, \ldots, r_n\}$. Furthermore, each row r_i has a trailing line-ending ℓ_j with $i = 1, \ldots, n$ and $j = 1, \ldots, \tilde{n}$. Since a line-ending for the last row, r_n is often times optional, the following applies: $\tilde{n} \leq n$. Accordingly, the total watermark capacity C_{max} can be described with $C_{max} = \tilde{n}$, limiting the size of B with $m \leq C_{max}$, meaning only as many bits fit in the CSV text as there are line-endings. However, for the following notions, we assume $\tilde{n} = n$ for the sake of simplicity.

In order to embed B into a set of rows R, resulting in a watermarked CSV text denoted as CSV_{wm}, we define the following watermark embedding function $W : (R, B) \rightarrow CSV_{wm}$ as follows:

$$W(R,B) = \bigoplus_{i=0}^{n} \begin{cases} r_i f(b_i), & \text{if } i \leq m, \\ r_i \ell_i, & \text{if } i > m \end{cases} \tag{1}$$

where $\oplus$ denotes a concatenation operator, putting all rows back together, each with a trailing line-ending. However, when choosing what line-ending to place, we distinguish between the following two cases: In cases where $i \leq m$ the bitstring B is not fully embedded into the CSV text yet. Thus, we apply a mapping function f, which determines what line-ending to append to each r_i in order to embed bit b_i. In all other cases where $i > m$ the bitstring is fully embedded within the cover text. Accordingly, all remaining rows r_i simply keep their original line-ending ℓ_i. As a result, we receive the watermarked CSV text CSV_{wm} for which $W(R,B) = CSV_{wm}$ applies.

Mapping function $f : \{0,1\} \rightarrow \{\text{CRLF}, \text{LF}\}$, determining what line-ending to append in order to embed a single bit b of bitstring B, is defined in the following way:

$$f(b) = \begin{cases} \text{CRLF}, & \text{if } b = 0, \\ \text{LF}, & \text{if } b = 1 \end{cases} \tag{2}$$

It is worth noting that the mapping function was arbitrarily chosen and may also be switched.

The embedding procedure can be implemented as described in Algorithm 1. Accordingly, all rows need to be iterated, and for each row r_i either a new line-ending is appended according to the mapping function or the original line-ending ℓ_i is maintained as soon as all bits b_i are embedded.

Data: $R \leftarrow$ CSV rows, with
$R := \{r_0, r_1, \ldots, r_n\}$, with line-ending
ℓ_i for each r_i
Data: $B \leftarrow$ Watermark bitstring, with
$B := \{b_0, b_1, \ldots, b_m\}$ and
$b_i \in \{0,1\}$ and $m \leq n$
Result: $CSV_{wm} \leftarrow$ watermarked CSV
Initialise $CSV_{wm}, lineEnding \leftarrow$ as empty
for $i = 1$ **to** n **do**
  if $i > m$ **then**
    $lineEnding \leftarrow \ell_i$
  else if $b_i = 0$ **then**
    $lineEnding \leftarrow$ CRLF
  else
    $lineEnding \leftarrow$ LF
  end
  $CSV_{wm} \leftarrow CSV_{wm} + r_i + lineEnding$
end
return CSV_{wm}

Algorithm 1: CSV Watermark Embedding.

3.2 Extraction Procedure

In order to extract the watermark's content, i.e. bitstring B, each row of the watermarked CSV text CSV_{wm} needs to be iterated by applying an inverse mapping function until all m bits are extracted. Accordingly, the extraction function $W^{-1} : CSV_{wm} \rightarrow \{0,1\}^m$ is denoted as:

$$W^{-1}(CSW_{wm}) = \bigoplus_{i=0}^{m} \left(f^{-1}(\ell_i) \right) \tag{3}$$

Thus $B = W^{-1}(CSW_{wm})$. Moreover, the inverse mapping function $f^{-1} : \{\text{CRLF}, \text{LF}\} \rightarrow \{0,1\}$ is defined as follows:

$$f^{-1}(\ell) = \begin{cases} 0, & \text{if } \ell = \text{CRLF}, \\ 1, & \text{if } \ell = \text{LF} \end{cases} \tag{4}$$

Consequently, if there is no tampering with the watermarked CSV text CSV_{wm}, we expect the following equation to hold true:

$$W^{-1}(W(R,B)) = B \tag{5}$$

Based on the prior, the extraction procedure can algorithmically be described as detailed in Algorithm 2. Therefore, each line-ending ℓ_i must be checked to extract each bit b_i respectively. The entire procedure is carried out until all m bits are extracted, resulting in B.

Data: $R \leftarrow$ CSV rows of CSV_{wm}, with $R := \{r_0, r_1, \ldots, r_n\}$, with line-ending ℓ_i for each r_i
Result: $B \leftarrow$ bitstring
Initialise $B \leftarrow$ as empty
for $i = 1$ **to** m **do**
  if $\ell_i = CRLF$ **then**
    $B \leftarrow B + 0$
  else if $\ell_i = LF$ **then**
    $B \leftarrow B + 1$
  end
end
return B

Algorithm 2: CSV Watermark Extraction.

It is worth noting that determining the size m of bitstring B, i.e., the number of line-endings required to read in order to extract the embedded information, might not be straightforward. In our case of embedding digital signatures, the resulting bitstring sizes are fixed length, dependent on the signature used. Therefore, extracting the watermark is no issue as long as the recipient knows when to stop reading. As knowing what signature is embedded is a precondition to be able to validate the signature at all, we assume this as given. However, in other cases with dynamic content size, the extraction procedure may require adjustments to be able to identify the last bit included. This could be done, e.g., by embedding special delimiter bits or bytes that clearly signal the end of contents.

4 EXPERIMENTAL EVALUATION

In order to evaluate the proposed fragile CSV watermarking technique, we conducted three experiments described in detail in the upcoming section. We base our evaluation criteria on related work by analysing the robustness, capacity, and imperceptibility (Knöchel and Karius, 2024; Li et al., 2021). It is essential to note that despite "the differences between watermarking techniques [...], the requirements that any watermarking system must satisfy can be summarised by the so-called watermarking tradeoff triangle" (Li et al., 2021, p. 172). This visual representation of the criteria as a triangle illustrates their interdependence and conflicts with one another (Li et al., 2021).

4.1 Experimental Setup

Several experiments were carried out to assess our CSV watermarking technique, focusing on its robustness, capacity, and imperceptibility. Each experiment used a set of RFC4180 conform CSV files. More specifically, two distinct datasets were compiled for the experiments. The datasets are based on prior work of (Vitagliano et al., 2023). In their work, the authors aimed to compile a representative real-world set of CSV files by scraping various data sources to evaluate their CSV dialects. In total, they collected 3712 files, which are accessible on GitHub[7], along with accompanying annotation JSON files denoting the dialect characteristics of each CSV. Furthermore, they designed an additional CSV file, that is intended to represent an average CSV file, both in dialect and content.

Consequently, we assembled two datasets: *DS1* and *DS2*. A full overview of both datasets' characteristics can be seen in Table 2. DS1 is a subset of (Vitagliano et al., 2023) initial sample set, excluding all files that were not RFC4180-conforming, not Unicode-encoded or Unicode-compatible, and exceeded a total file size of 1 Mb. This was to ensure a consistent and manageable set of files. Doing this resulted in 380 distinct CSV files. The median amount of rows and columns are 64 and 8, respectively. The median file size is 7.33 KB.

Table 2: Dataset Overview as Median Values.

Dataset	Files	Rows	Columns	Size (KB)
DS1	380	64	8	7.33
DS2	1	84	9	21.4

DS2, on the other hand, comprises solely the presented average CSV file. However, as this file used LF line-endings instead of the RFC4180 stipulated CRLF line-endings, we converted them accordingly.

[7]https://github.com/HPI-Information-Systems/Pollock [24.02.25]

DS2 has 84 rows, 9 columns and a file size of 21.4 KB. The content is a broad mixture of numeric values of different formats, as well as various text values encompassing dates, short text, long multi-line descriptions, and special characters.

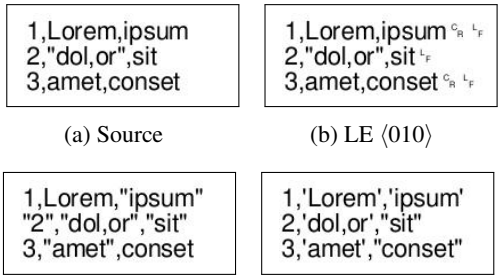
Using the datasets mentioned above, the following experiments were conducted to analyse our watermarking technique:

Experiment A: This experiment aimed to analyse both the capacity and imperceptibility of the technique. Therefore, we implemented a testbed in Python, which allowed us to embed a random bitstring of maximum length C_{max} into all CSV files of DS1. In a second step, we computed the following metrics: Char and file size difference of original and watermarked CSV, embedded bits per character, and the Structural Similarity Index Measure (SSIM). According to (Setiadi, 2021), SSIM is particularly well suited to measure visual similarity as it closely matches human perception. To apply SSIM, we used the Python package Pillow to render a visual representation of CSV files, to be then able to calculate the SSIM between original and watermarked files using scikit-image's SSIM implementation. This and all following experiments were run on a Desktop Computer, running Windows 11 Pro 64-Bit 24H2, equipped with an AMD Ryzen 7 3800XT 8-Core processor, running with a base clock speed of 3.9 GHz, as well as 32GB of DDR4 3200MHz C16 memory. Moreover, Python 3.10 was used as an interpreter.

Experiment B: As changing the line-endings of CSV text modifies its syntax, this experiment investigated whether the alterations fall within acceptable boundaries. This is in line with (Vitagliano et al., 2023) as they found that both CRLF and LF are used widely for CSV files. Therefore, it's important to validate whether CSV tools can handle a mixture of them. Otherwise, CSV tools would prompt any syntax errors directly to its users, affecting the watermark's imperceptibility. Therefore, we chose different CSV linters to check whether they are able to validate a watermarked CSV file successfully. To do this, we first watermarked the average CSV file of DS2 to then manually conduct the experiment on five online available CSV linters. The CSV linters used are listed in Table 6. The linters were chosen based on the fact that they offer direct file uploads. This was an important consideration, as most other linters would otherwise normalise line-endings if printed to a text field before validating it.

Experiment C: This experiment aimed at analysing the robustness of the watermarking technique by investigating how different CSV tools and text editors affect the embedded watermark. More specifically, this experiment examined the normalisation of line-endings. To do so, we first watermarked the average CSV file of DS2. Next, we used a set of candidate tools and manually opened and saved the file without making any changes. This was to trigger a potential normalisation or reformatting of the CSV file. The candidate tools used are displayed in Table 3. The set comprises major CSV tools and text editors commonly used by regular users to edit and view CSV text or files.

Lastly, in our effort to address the shortcomings of related work, we decided to carry out all of the three experiments for the Double Quote approach (DQ) mentioned in (Ito, 2024) and for the Double Single Quotes Code approach (DSQC) described in (Wen and Wang, 2013) as well. A brief description of their embedding technique is outlined in Section 2.4. Doing this allowed for a comprehensive comparison of our approach and existing work. Hence, we implemented both approaches in our testbed, adhering to the explanations provided by the authors in their work. A comparison of all three methods based on a lorem ipsum CSV text is displayed in Figure 2. It is important to highlight that we needed to make the usually invisible line-endings visible to observe the difference in our line-ending-based approach (LE).



(a) Source (b) LE ⟨010⟩

(c) DQ ⟨00111010⟩ (d) DSQC ⟨111010⟩

Figure 2: Rendered CSV Texts With Watermark Contents.

4.2 Robustness

The robustness of a watermark is connected to its persistence, which refers to the ability to withstand both intentional and unintentional modifications or attacks (Swanson et al., 1998). Therefore, it is typically assessed by simulating various attack scenarios such as insertion, deletion, or replacement attacks, thereby demonstrating its persistence across

Table 3: Experiment C: Normalisation Test.

Tool	LE	DQ	DSQC	DS2
Excel (LTSC Pro+ 2021 v2108)	✓	✓	(✓)	✓
LibreOffice Calc (v25.2.0)	✓	✓	(✓)	✓
Google Sheets	✓	✓	(✓)	✓
Windows Editor (v11.2410.21.0)	✓	✗	✗	✗
VSCode (v1.97.2)	✓	✗	✗	✗
Sublime Text (v4192)	✓	✗	✗	✗
Notepad++ (v8.6.9)	✗	✗	✗	✗
Atom (v1.60.0)	✗	✗	✗	✗
Vim (v9.1.0)	✗	✗	✗	✗

different cases (Rizzo et al., 2019). However, in the case of fragile watermarking, the opposite is true. A fragile watermark should be capable of detecting any changes made to the cover medium and enable the user to authenticate its integrity.

In our case, two factors must be considered when analysing its fragility: line-ending normalisation and digital signatures. As digital signatures are a proven method of data integrity authentication (NIST, 2023), we mainly focused on the implications of line-ending normalisation, investigating whether the embedded digital signature would easily be wiped or not. Therefore, we conducted Experiment C, opening and saving a sample watermarked CSV file in various commonly used CSV and text editors. The results are presented in Table 3. A check mark '✓' indicates that a normalisation occurred, whereas a '✗' indicates that the file remained unchanged. On the other hand, a '(✓)' indicates that although the file was normalised, its structure was compromised during the process and, as a result, not accurately parsed.

Interestingly, normalisation occurred for all three types of watermarked files as well as the original unwatermarked source file. This is based on the fact that two styles of using quotation exist: a minimal one, enclosing only cells which require to be escaped due to special characters like commas or line-endings within the cell itself, or the holistic one, which encloses all cells regardless of their content (Vitagliano et al., 2023). Therefore, especially CSV tools like Excel, LibreOffice Calc or Google Sheets normalised both line-endings and quotations, resulting in a consistent minimal CSV text. Moreover, no CSV tool was able to parse the DSQC watermarked CSV file, as they did not manage to handle a mixed use of single and double quotation marks. As a result, they broke the structure of the CSV format when trying to normalise it.

Furthermore, it becomes evident that solely CSV tools performed CSV specific normalisation. All other general text editors did not normalise quotations. However, some of them do normalise line-endings. Considering the variety of text editors available, it appears that more sophisticated text editors such as Notepad++, Atom, or Vim do not perform any modifications, like normalisation. In contrast, more user-friendly tools like Windows Editor tend to normalise line-endings.

Based on this, the robustness of all three approaches is prone to normalisation to some degree. Yet, our approach seems to be affected by it in more cases, as line-ending normalisation is format-agnostic. However, in all three cases, a normalisation can technically be seen as a modification to the original cover text, as it is indeed a modification of the CSV text as a whole. Even if the watermark would stay persistent, verifying the extracted digital signature would fail regardless.

4.3 Capacity

A watermark's capacity is often defined as the number of bits the watermark achieves to embed into the cover medium (Li et al., 2021). Besides robustness and imperceptibility, it is also an important evaluation criterion for watermarks, as watermarking techniques usually attempt to achieve a high embedding capacity. However, according to (Liu et al., 2025), the greater the watermark's size, the more it negatively impacts the cover text's fidelity. As a result, one typically has to choose between the two.

We conducted Experiment A to assess the capacity of our approach alongside the two other methods. Using 380 representative RFC4180 conform CSV files of varying length and contents, we embedded a random bitstring of maximum length into each file. The results can be seen in Table 4. The embedding capacity, that is, the embedded bits per character, was the highest for the DQ approach with 0.07 bits/char and a median of 475 embedded bits, followed by DSQC with 0.05 bits/char and a median of 312 embedded bits. Our approach accomplished

0.01 bits/char with a median of 64 embedded bits.

Table 4: Experiment A: Results as Median Values.

	LE	DQ	DSQC
Total Embedded Bits	64	475	312
Emb. Bits/Char	0.01	0.07	0.05
Char Count Difference	-32	407	485
SSIM Score	1.0	0.62	0.56

The significant capacity difference among the three approaches arises from the embedding technique. Whereas the DQ and DSQC approaches both embed their bits per cell, our line-ending-based approach is row-based. Thus, it can only include as many bits as there are rows. As the sample set's median of rows was 64, so was the embedding capacity of our approach. The slight difference in DQ and DSQC is based on the fact that DQ considers numeric values as well, whereas DSQC solely enquotes text values.

Since all three methods depend on embedding a digital signature into the CSV text, we must also consider the size of common signatures. A brief overview is pictured in Table 5. It becomes evident that given our sample set, DQ is the only approach which can reliably fit a digital signature in most cases. DSQC, on the other hand, would be able to utilise smaller signatures such as DSA. In contrast, our approach can only embed a digital signature into CSV files, with a minimum amount of 320 rows. Therefore, all three approaches depend highly on the size and content of a CSV text and are thus only applicable for larger files. Yet, the size affects our line-ending-based approach to a higher degree.

Table 5: Common Digital Signature Sizes Based On (NIST, 2013; NIST, 2023).

Digital Signature	Signature Size
DSA	320 - 512 Bits
RSA	1024 - 4096 Bits
ECDSA	512 - 1024 Bits
EdDSA	512 - 896 Bits

4.4 Imperceptibility

The imperceptibility of watermarks is highly connected to human perception. According to (Swanson et al., 1998), a truly imperceptible embedding procedure is given in case humans cannot differentiate between original and watermarked content. However, as capacity and imperceptibility are conflicting goals, watermarking approaches usually aim for either one of them.

Accordingly, our goal in conducting Experiment A was, besides analysing the capacity, to evaluate the visual similarity of the three approaches. To do so, for each CSV file, we created a rendered image of the plain CSV text both for the original and watermarked contents. Based on this, we calculated the SSIM score. The median values for each approach are displayed in Table 4.

Following this, our approach has a median SSIM score of 1.0, denoting full similarity between all original and watermarked CSV texts and is therefore indistinguishable. DQ on the other hand has a median similarity of 0.62, slightly better than DSQC with a similarity score of 0.56. The difference arises because, in certain cases, DQ does not add any additional double quotes to the row's first or proceeding cells. Doing this shifts the entire line to the right, resulting in greater dissimilarity.

Additionally, we examined the char size difference between the source and the watermarked file. As an increase in char and, thus, in file size affects both practicality and imperceptibility, we chose to include this metric in our experiment. The resulting differences are also displayed in Table 4. Therefore, our approach managed to decrease the char size with a median of -32 chars, whereas both DQ and DSQC increased the char size by 407 and 485, respectively. The decrease in char size is because our approach replaces some of the CRLF line-ending control characters with a single LF character. It is worth noting that a file having LF line-endings per default would lead to an increase in char size. However, as we used RFC4180 conform samples, the default line-ending was CRLF.

Furthermore, by conducting Experiment B, we evaluated whether the modifications lay within acceptable boundaries. A CSV tool or linter prompting any warnings or format errors to a potential user would significantly decrease the imperceptibility. We, therefore, manually used five online CSV linters. The results can be seen in Table 6.

Table 6: Experiment B: CSV Linter Validity.

Linter	LE	DQ	DSQC	DS2
CSVLint.io	✗	✓	(✗)	✓
ToolkitBay.com	✓	✓	(✗)	✓
CSVLint.com	✓	✓	✓	✓
Zazuko.com	✓	✓	(✗)	✓
ExtendsClass.com	✗	✗	(✗)	✗

The DQ approach was successfully validated by four out of five linters. In contrast, two out of five linters could not validate our approach. One mentioned inconsistent line-endings, whereas the

other had problems parsing the file. Similarly, only one linter successfully validated the DSQC approach. All others were unable to parse it correctly and, consequently, could not validate it at all. It is worth noting that the ExtendsClass CSV validator could not validate any of the files, including the original unwatermarked source file. The reason for this was the presence of an empty last row in the file. However, according to RFC4180, an empty last row is permissible.

Following the previous results, our method exhibits superior imperceptibility compared to the other two approaches. This advantage arises as our approach utilises invisible control characters, which are not detectable by the human eye. In contrast, both DQ and DSQC utilise visible quotations, degrading the visual appearance.

5 DISCUSSION

In this study, we present a novel CSV fragile watermarking technique. Our aim is to overcome the shortcomings of existing approaches by addressing the imperceptibility. Consequently, we investigated the application of invisible line-ending control characters and compared our technique with relevant existing approaches by conducting three experiments using two distinct representative datasets.

Our experiments show that robustness, characterised by fragility in our case, is affected by normalisation across all three methods. However, normalisation is notably more influential on our technique because of the usage of line-endings. Nevertheless, a wiped watermark presents no issue in most cases, as modifying a file would lead to an invalid digital signature anyway. Therefore, a failed validation might be as good as having no signature at all for many users, as they would not be able to estimate what changes have been made and whether they affect the accuracy of the data or the format only after all.

Furthermore, the experiments highlight the difference in the overall watermarks' capacities. Our findings show that our approach is inferior to the other approaches, as it embeds bits row-based rather than cell-based. Considering the size of commonly used digital signatures, it is evident that our approach is applicable to files that are at least 320 rows in size only. However, as we excluded all files greater than 1 Mb from our dataset, the median value might be higher in reality. For example, in relevant related work, researchers analysed 104.826 CSV files from various sources (Mitlöhner et al., 2016). They found that their sample set had a mean value of 379 rows, with a min and max value of 1 and 8684, respectively. This emphasises the great variation in row sizes.

Lastly, we demonstrate the complete imperceptibility of our approach and its superiority over the two other candidates. Consequently, a watermarked CSV is indistinguishable by the human eye using our technique. Only by comparing the difference in file size, a user is able to detect the modification. In contrast, when using DQ and DSQC, a user can clearly identify the changes made both visually and by size. However, it is worth noting that a typical user who solely views the watermarked CSV text without a side-by-side comparison might not anticipate a watermark embedding scheme behind it.

Our findings validate the already known trade-off watermark techniques must make regarding the three criteria. We suggest that our contribution introduces a novel fragile watermarking technique for CSV text. To the best of our knowledge, this is the first fragile watermarking technique for CSV, which is genuinely imperceptible. It is, therefore, particularly well-suited for use cases where imperceptibility is the most important goal. However, if a use case requires higher capacity, then either DQ or DSQC may be more appropriate. In terms of robustness, the difference between the approaches is, in reality, negligible. Therefore, the choice between the approaches largely depends on the use case and the objectives one aims to achieve.

5.1 Limitations & Future Work

In this section, we explore the limitations of our work and outline future touch points. Firstly, although we utilised the wide range of dialects available and relied on the robustness of popular parsers and tools to handle such inconsistencies, mixing line-endings adds to the already challenging landscape of inconsistent CSV files. Our results support the conclusions drawn by (Vitagliano et al., 2023), indicating that the issue of inconsistent CSV dialects remains a significant challenge for various tools and parsers. Consequently, the effectiveness of our method is highly dependent on the context. Future research could improve this by exploring which environments might benefit from our approach specifically, where normalisation and inconsistencies are not an issue.

Second, commonly used normalisation wipes our watermark. While this may not undermine the goal of solely utilising text whose integrity has been verified, some users may prefer standard formatting and normalisation optimisations, given that the contents

remain unchanged. Consequently, future work could address this by developing a semi-fragile approach for CSV text, which permits simple formatting-related modifications. To do this, error correction codes might be suitable for localising any changes.

Third, our method uses a row-based embedding scheme, which limits its capacity. Since a digital signature is necessary to verify the text's integrity, our solution is impractical for small CSV files. Hence, future research could investigate other more compact tamper-detection methods, balancing size and security. For example, in (Rizzo et al., 2016) *SipHash*, a key-based hash is used, which is 64 bits in size only. Additionally, the feasibility of expanding the set of line-ending control characters could be explored to increase overall capacity.

6 CONCLUSION

In this study, we proposed a novel fragile watermarking technique for CSV text. We aimed to address the shortcomings of existing techniques, focusing on imperceptibility specifically. Using a combination of different invisible line-ending control characters, we are able to embed any byte-encodable information into a CSV cover text. Moreover, we conducted three experiments with representative datasets in order to evaluate and compare our approach with relevant existing work.

We found that while our approach has limited capacity compared to existing techniques, it excels in imperceptibility. Therefore, our approach is most suitable in situations where imperceptibility is the primary goal. However, due to the line-ending-based embedding scheme, our approach is more vulnerable to normalisation, making the watermark sensitive to formatting procedures. Consequently, careful consideration is required when choosing to implement our embedding scheme. Future work should address this issue by developing a semi-fragile watermarking technique which allows for format optimisations.

ACKNOWLEDGEMENTS

CRediT Author Statement

Florian Zimmer: Conceptualisation, Methodology, Software, Data Curation, Investigation, Writing - Original Draft, Visualisation. **Malte Hellmeier:** Conceptualisation, Methodology, Investigation, Writing - Original Draft, Visualisation. **Motoki Nakamura:** Conceptualisation, Investigation. **Tobias Urbanek:** Software.

REFERENCES

Abba, A. H. and Hassan, M. (2018). Design and implementation of a csv validation system. In Bogach, N., Pyshkin, E., and Klyuev, V., editors, *Proceedings of the 3rd International Conference on Applications in Information Technology*, pages 111–116, New York, NY, USA. ACM.

Alkawaz, M. H., Sulong, G., Saba, T., Almazyad, A. S., and Rehman, A. (2016). Concise analysis of current text automation and watermarking approaches. *Security and Communication Networks*, 9(18):6365–6378.

Allen, J. D. (2007). *The Unicode standard 5.0*. Addison-Wesley, Upper Saddle River NJ, new ed. edition.

Bender, W., Gruhl, D., Morimoto, N., and Lu, A. (1996). Techniques for data hiding. *IBM Systems Journal*, 35(3.4):313–336.

Brickely, D., Tennison, J., and Herman, I. (2016). Csv on the web working group. Accessed February 2025 at: https://www.w3.org/2013/csvw/wiki/Main_Page.html.

Cox, I. J., Miller, M. L., and Bloom, J. A. (2000). Watermarking applications and their properties. In *Proceedings International Conference on Information Technology: Coding and Computing (Cat. No.PR00540)*, pages 6–10. IEEE Comput. Soc.

Fortune Business Insights (2024). Server operating system market volume. Accessed February 2025 at: https://www.fortunebusinessinsights.com/server-operating-system-market-106601.

Hartung, F. and Kutter, M. (1999). Multimedia watermarking techniques. *Proceedings of the IEEE*, 87(7):1079–1107.

He, J., Ying, Q., Qian, Z., Feng, G., and Zhang, X. (2020). Semi-structured data protection scheme based on robust watermarking. *EURASIP Journal on Image and Video Processing*, 2020(1).

Hellmeier, M., Qarawlus, H., Norkowski, H., and Howar, F. (2025). A hidden digital text watermarking method using unicode whitespace replacement. In Bui, T. X., editor, *Proceedings of the 58th Hawaii International Conference on System Sciences*, pages 7411–7420. Hawaii International Conference on System Sciences.

Hisham, S. I., Muhammad, A. N., Zain, J. M., Badshah, G., and Arshad, N. W. (2013). Digital watermarking for recovering attack areas of medical images using spiral numbering. In *2013 International Conference on Electronics, Computer and Computation (ICECCO)*, pages 285–288. IEEE.

IEEE (1986). Portable operating system interface for computer environments (POSIX).

Ito, A. (2024). Embedding digital signature into csv files using data hiding.

Jaigirdar, F. T., Rudolph, C., and Bain, C. (2019). Can i trust the data i see? In *Proceedings of the Australasian Computer Science Week Multiconference*, pages 1–10, New York, NY, USA. ACM.

Jalil, Z. and Mirza, A. M. (2009). A review of digital watermarking techniques for text documents. In *2009 International Conference on Information and Multimedia Technology*, pages 230–234. IEEE.

Kamaruddin, N. S., Kamsin, A., Por, L. Y., and Rahman, H. (2018). A review of text watermarking: Theory, methods, and applications. *IEEE Access*, 6:8011–8028.

Kerckhoffs, A. (1883). La cryptographie militaire: pp. 5-38. *J. Sciences Militaires*.

Knöchel, M. and Karius, S. (2024). Text steganography methods and their influence in malware: A comprehensive overview and evaluation. In Pérez-González, F., Comesaña-Alfaro, P., Krätzer, C., and Vicky Zhao, H., editors, *Proceedings of the 2024 ACM Workshop on Information Hiding and Multimedia Security*, pages 113–124, New York, NY, USA. ACM.

Li, Y., Wang, H., and Barni, M. (2021). A survey of deep neural network watermarking techniques. *Neurocomputing*, 461:171–193.

Liu, A., Pan, L., Lu, Y., Li, J., Hu, X., Zhang, X., Wen, L., King, I., Xiong, H., and Yu, P. (2025). A survey of text watermarking in the era of large language models. *ACM Computing Surveys*, 57(2):1–36.

Liu, T.-Y. and Tsai, W.-H. (2007). A new steganographic method for data hiding in microsoft word documents by a change tracking technique. *IEEE Transactions on Information Forensics and Security*, 2(1):24–30.

Mitlöhner, J., Neumaier, S., Umbrich, J., and Polleres, A. (2016). Characteristics of open data csv files. In *2016 2nd International Conference on Open and Big Data (OBD)*, pages 72–79. IEEE.

NIST (2013). Digital signature standard (DSS) - FIPS 186-4.

NIST (2023). Digital signature standard (DSS) - FIPS 186-5.

Otto, B. (2022). The evolution of data spaces. In Otto, B., ten Hompel, M., and Wrobel, S., editors, *Designing Data Spaces*, pages 3–15. Springer International Publishing, Cham.

Petitcolas, F., Anderson, R. J., and Kuhn, M. G. (1999). Information hiding-a survey. *Proceedings of the IEEE*, 87(7):1062–1078.

Podilchuk, C. I. and Delp, E. J. (2001). Digital watermarking: algorithms and applications. *IEEE Signal Processing Magazine*, 18(4):33–46.

Por, L. Y., Wong, K., and Chee, K. O. (2012). Unispach: A text-based data hiding method using unicode space characters. *Journal of Systems and Software*, 85(5):1075–1082.

Rani, S. and Halder, R. (2022). Comparative analysis of relational database watermarking techniques: An empirical study. *IEEE Access*, 10:27970–27989.

Rizzo, S. G., Bertini, F., and Montesi, D. (2016). Content-preserving text watermarking through unicode homoglyph substitution. In Desai, B. C., Toyama, M., Bernardino, J., and Desai, E., editors, *Proceedings of the 20th International Database Engineering & Applications Symposium on - IDEAS '16*, pages 97–104, New York, New York, USA. ACM Press.

Rizzo, S. G., Bertini, F., and Montesi, D. (2019). Fine-grain watermarking for intellectual property protection. *EURASIP Journal on Information Security*, 2019(1).

Saltzer, J. H. and Ossanna, J. F. (1970). Remote terminal character stream processing in multics. In Cooke, H. L., editor, *Proceedings of the May 5-7, 1970, spring joint computer conference on - AFIPS '70 (Spring)*, page 621, New York, New York, USA. ACM Press.

Setiadi, D. R. I. M. (2021). Psnr vs ssim: imperceptibility quality assessment for image steganography. *Multimedia Tools and Applications*, 80(6):8423–8444.

Shafranovich, Y. (2005). Rfc 4180. Accessed February 2025 at: https://www.rfc-editor.org/rfc/rfc4180.html.

Swanson, M. D., Kobayashi, M., and Tewfik, A. H. (1998). Multimedia data-embedding and watermarking technologies. *Proceedings of the IEEE*, 86(6):1064–1087.

Taleby Ahvanooey, M., Li, Q., Hou, J., Dana Mazraeh, H., and Zhang, J. (2018). Aitsteg: An innovative text steganography technique for hidden transmission of text message via social media. *IEEE Access*, 6:65981–65995.

Taylor, P. (2024). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2023, with forecasts from 2024 to 2028. Accessed February 2025 at: https://www.statista.com/statistics/871513/worldwide-data-created/.

Tian, Y. and Nogales, A. F. R. (2023). A survey on data integrity attacks and ddos attacks in cloud computing. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0788–0794. IEEE.

van den Burg, G. J. J., Nazábal, A., and Sutton, C. (2019). Wrangling messy csv files by detecting row and type patterns. *Data Mining and Knowledge Discovery*, 33(6):1799–1820.

Vitagliano, G., Hameed, M., Jiang, L., Reisener, L., Wu, E., and Naumann, F. (2023). Pollock: A data loading benchmark. *Proceedings of the VLDB Endowment*, 16(8):1870–1882.

Wen, Q. and Wang, Y. (2013). An efficient fragile web pages watermarking for integrity protection of xml documents. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Shi, Y. Q., Kim, H.-J., and Pérez-González, F., editors, *Digital Forensics and Watermaking*, volume 7809 of *Lecture Notes in Computer Science*, pages 135–144. Springer Berlin Heidelberg, Berlin, Heidelberg.