

UNSHELM, Nina; SILLER, Hans-Stefan  
Würzburg

## **Untersuchung von Medienaussagen mit großen Datensätzen – Auswahl von Daten als Komponente von Data Literacy**

Daten werden zunehmend in der Kommunikation von Informationen, z. B. durch Medienberichte, eingesetzt – etwa als Hintergrundinformationen oder Argumentationsmittel (Schüller et al., 2019). Damit einhergehend gewinnen Fähigkeiten des Umgangs mit Daten an Bedeutung. Diese werden als Data Literacy zusammengefasst und z. B. in den Methoden der Data Science verwendet (Schüller et al., 2019). Zusätzlich stellen der Einfluss von Fehl- und Desinformationen auf die öffentliche Meinung (Lewandowsky et al., 2017) sowie die komplexen Themen der Nachhaltigkeit, z. B. Sustainable Development Goals (SDGs; UN General Assembly, 2015), Herausforderungen dar. Insbesondere sind bei den SDGs, zur Darstellung von Problemen oder Lösungsansätzen, Daten erforderlich, weshalb im Umgang damit Data Literacy unumgänglich ist. Zur Befähigung der Bevölkerung zur aktiven Teilnahme an der Gesellschaft ist es deshalb notwendig, die Untersuchung von Medienaussagen und Data Literacy zu fördern. Dafür bietet sich im schulischen Kontext der Mathematikunterricht an.

### **Theoretischer Hintergrund**

Der Begriff Data Literacy umfasst Kenntnisse und praktische Fähigkeiten, die den verständnisorientierten Umgang mit Daten ermöglichen und zur Nutzung, Analyse und Interpretation von Daten befähigen (Schüller et al., 2019). Dies ist bei der alltäglichen Konfrontation mit vielen Daten unumgänglich, weshalb Data Literacy eine Schlüsselkompetenz des 21. Jahrhunderts ist (Schüller et al., 2019). In diesem Beitrag wird das *Untersuchen von Medienaussagen mit Mathematik und Daten* fokussiert. Dieses umfasst laut unserer Charakterisierung u. a. die Identifikation und Überprüfung von Informationen aus Medienberichten anhand von Daten und ist aufgrund der Adressierung einzelner Aspekte der Data Literacy ein Teilbereich davon (Unshelm et al., in Begutachtung). Weiterhin legen wir hier den Fokus auf die Auswahl von Daten als eine Komponente von Data Literacy, da diese im Rahmen der Untersuchung von Medienaussagen unerlässlich ist. Unter der Auswahl von Daten verstehen wir die Identifikation von zu einem Medienbericht passenden Daten aus (großen) Datensätzen. Passende Daten aus den Datensätzen sind solche, die dieselben Bezugspunkte (z. B. Daten für ein bestimmtes Land oder Jahr) sowie Datentypen (z. B. Daten, die einen Anteil beschreiben) wie die Daten aus den Medienberichten besitzen.

Die Förderung von Data Literacy im schulischen Kontext wird vermehrt gefordert (Schüller et al., 2021). Jedoch zeigen sich bei (schulischen) Ansätzen, die die Themenfelder Medien, Mathematik und Daten verbinden, Lücken: Studien, die Evidenz im Rahmen der Untersuchung von Medienaussagen zu Nachhaltigkeitsthemen mit Hilfe der Modellierung mit Daten ansprechen, sind selten (Unshelm et al., eingereicht) und aktuelle bayerische Mathematik-Schulbücher der 9. und 10. Jahrgangsstufe beinhalten nur eingeschränkt Möglichkeiten zur Einschätzung der Glaubwürdigkeit von Medienaussagen unter Verwendung von großen Datensätzen (Siller et al., 2025). Diese Probleme wurden von uns aufgegriffen, indem Aufgabenstellungen entwickelt wurden, die die Untersuchung von Medienaussagen im Nachhaltigkeitskontext mit Hilfe von Mathematik und großen Datensätzen fokussieren. Im Rahmen der Aufgabenstellungen sollen Lernende (widersprüchliche) Aussagen in zwei realen Medienberichten zu Themen der Nachhaltigkeit (globale Kohlenstoffdioxid-Emissionen (C) und globales Ausmaß der Armut (A)) identifizieren und diese anschließend mit Daten aus großen Datensätzen untersuchen. Dazu müssen aus vier vorgegebenen, nur teilweise passenden, großen, realen Datensätzen, zu den Inhalten der Medienaussagen passende Daten selbstständig ausgewählt und diese entsprechend verarbeitet werden (Materialien siehe [www.we-stem.it](http://www.we-stem.it)). In diesem Beitrag steht die Frage im Fokus, wie die Lernenden bei diesen Aufgaben (passende) Daten zur Untersuchung von Medienaussagen aus vorgegebenen, großen Datensätzen auswählen. Daran anknüpfend stellt sich die Frage, inwiefern die entwickelten Aufgabenstellungen Komponenten der Data Literacy ansprechen.

## **Methode**

Die entwickelten Aufgabenstellungen wurden 2023 (C) und 2024 (A) mit einer 9. bzw. 10. Klasse (Gymnasium) erprobt. In diesem Beitrag werden punktuell Ergebnisse einer Analyse einer zufälligen Auswahl von jeweils der Hälfte der dabei entstandenen transkribierten Audio- und Bildschirmaufnahmen (insgesamt sieben Transkripte – vier zu C und drei zu A) vorgestellt. Die Analyse erfolgte durch eine qualitative Inhaltsanalyse nach Mayring (2022). Dazu wurden zunächst induktiv aus den Medienartikeln, relevante Charakteristika der Daten, für die Auswahl passender Daten, identifiziert und in drei übergreifende inhaltliche Auswahlkriterien klassifiziert – Datentyp, geographischer Bezugspunkt und zeitlicher Bezugspunkt. Anschließend wurden die Transkripte hinsichtlich dieser drei Kriterien qualitativ ausgewertet sowie die drei Kriterien induktiv erweitert. Die Reliabilität der Analyse wurde durch eine Zweitkodierung von über 50 % der Transkripte mit anschließender Konsensbildung sichergestellt.

## Ergebnisse und Interpretationen

Zur Untersuchung der identifizierten Medienaussagen wählen die Lernenden teils Daten mit Datentypen aus, die in den Medienaussagen nicht vorkommen: In der C-Aufgabe wählen die Lernenden z. B. häufig Daten aus Datensatz 2, Tabellenblatt "GHG Emission from Energy", obwohl diese Daten alle Treibhausgase und nicht nur Kohlenstoffdioxid-Emissionen beinhalten. Folglich wird der Datentyp, also die Art der Daten, teilweise nicht beachtet und dadurch unpassende Daten zur Untersuchung der Aussagen ausgewählt.

Hinsichtlich der geographischen Bezugspunkte der Daten, also der Länder oder Bereiche, auf die sich die Daten beziehen, gibt es einerseits Gruppen, die in Übereinstimmung mit den Aussagen der Medienberichte, globale Daten zur Untersuchung der Aussagen suchen:

[Lernende suchen World in Datensatz 3]

S15: Gibt es zufällig bei W world oder so? Hier. Oh ich bin so schlau. W steht natürlich für World. Also.

S14: Als ob es wirklich ne Spalte mit world gibt. [00:32:10 - 00:32:18]

Andererseits wählen viele Gruppen (fälschlicherweise) die Daten einzelner Länder, insbesondere häufig Deutschland, aus:

S8: Bei welchem Land sollen wir denn jetzt schauen?

S7: Boah mach mal irgendwas.

S8: Ist Deutschland mit? Hä. Leute. Ich find kein. Soll ich bei Germany schauen? [Lernender sucht Zeile von Deutschland in dem Datensatz]

S7: Ja, gucke mal bei Germany. Germany oder so was ähnliches. Dann nimm ein anderes Land, ist egal. [Lernende finden Germany nicht] [02:01:28 - 02:02:16]

Viele Gruppen vergleichen die zeitlichen Bezugspunkte der Daten, also die Jahre, auf die sich die Daten beziehen, in Medienberichten und Datensätzen und nutzen folglich dieses Auswahlkriterium zielführend:

S8: Also der zweite Artikel ging ja bis 21, ne?

S9: Der zweite Artikel? Ähm ja, bis 2021, bis November 2021 mit späteren Änderungen.

S8: Aber zwei Datensätze gehen doch auf jeden Fall bis 21. Dann nehmen wir schonmal den erste äh den zweiten Datensatz, wo das so weit geht. [02:00:34 - 02:00:55]

Die Lernenden erkennen außerdem weitere Bezugspunkte der Daten in den Datensätzen und verwenden diese in ihrer Datenauswahl. Die Lernenden identifizieren bspw., dass sich die Daten in Datensatz 2 der A-Aufgabe auf unterschiedliche Geschlechter (Male, Female, Total) beziehen und entscheiden sich aufgrund des Medienberichts dann für die Daten zu "Total":

S10: Ja. Aber wir haben da kein. Ich würd das ganz ehrlich nur total machen.

S8: Ja würd ich auch machen, alle Menschen sind doch gleich.

S10: Ja, aber ganz ehrlich beim anderen gibt's auch keine Unterscheidung. Also in dem Text gibt's auch keine Unterscheidung.

S9: Dann müssen wir das auch nicht unterscheiden. [02:18:00 - 02:18:13]

Neben den vorgestellten inhaltlichen Auswahlkriterien ist auch das Aussehen und die Verständlichkeit der Datensätze ein Auswahlkriterium:

S10: Also ich würd den [Datensatz 2] ganz ehrlich auch verwenden, weil der sieht ziemlich gut aus.

S9: Übersichtlich, ja. [02:09:44 - 02:09:48]

## Fazit und Ausblick

Die exemplarischen Ergebnisse der Auswertung zeigen, dass Lernende verschiedene, mehr oder weniger geeignete Kriterien im Rahmen des Auswahlprozesses von Daten zur Untersuchung von Medienaussagen anwenden. Der zeitliche Bezugspunkt der Daten wird meist zielführend zur Datenauswahl verwendet. Die Kriterien Datentyp und geographischer Bezugspunkt werden jedoch teils ungenügend oder fehlerhaft eingesetzt, wodurch unpassende Daten ausgewählt werden. Zusätzlich zeigen sich weitere Bezugspunkte, die beachtet werden und solche Auswahlkriterien, die auf das Aussehen und die Verständlichkeit der Datensätze abzielen. Da insgesamt viele Prozesse der Datenauswahl gezeigt werden, sprechen die Aufgaben somit mindestens diese Komponente der Data Literacy an, auch wenn die Auswahl aus vorgegebenen Datensätzen eine Limitation darstellt. Damit sind sie ein Indiz für einen unterrichtlichen Ansatzpunkt von Data Literacy.

## Literatur

- Lewandowsky, S., Ecker, U. K. H. & Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369. <https://doi.org/10.1016/j.jarmac.2017.07.008>
- Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13. Neuauflage). Julius Beltz GmbH & Co. KG.
- Schüller, K., Busch, P. & Hindinger, C. (2019). *Future Skills: Ein Framework für Data Literacy*. <https://doi.org/10.5281/ZENODO.3349864>
- Schüller, K., Koch, H. & Rampelt, F. (2021). *Data-Literacy-Charta: Version 1.2*. <https://www.stifterverband.org/charta-data-literacy>
- Siller, H.-S., Unshelm, N., Geiger, V. & Digan, S. (2025). Implementation of Media and Data Literacy in Student Tasks: An Analysis of different Tasks. In A. Coles & K. Le Roux (Hrsg.), *ICMI Study 27*, Quezon City, Philippinen.
- UN General Assembly (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*.
- Unshelm, N., Digan, S., Geiger, V. & Siller, H.-S. (eingereicht). Empowering Students to Evaluate Media Claims with Large Data. In *Proceedings of ICTMA21*. Springer.
- Unshelm, N., Siller, H.-S., Digan, S. & Geiger, V. (in Begutachtung). Umgang mit großen Datensätzen zur Untersuchung von Medienaussagen: Vorstellung und erste empirische Erprobung eines Modells notwendiger Praktiken. *mathematica didactica*.