



Ein *Tag* sagt mehr als tausend Worte?

Kreatives Potenzial und Neotags in Tagging-Systemen

Christof Niemann



Themen

- Wissensorganisation im Internet
- Intelligenz und Kreativität im Web 2.0
- Tagging-Systeme als Forschungsgegenstand
- Vermessungen zur Kreativität der Sprache
- Intelligenz im Chaos? Erste Ergebnisse.
- Ausblick

Wissensorganisation im Internet

Der Information Overload

- Problem: Sinnvolle Annotation und Strukturierung von Informationsmassen
- Lösung: Kombination der Erschließungstypen



Intelligenz und Kreativität im Web 2.0

Emergierende Ordnung = Schwarmintelligenz?

- Vorbild Natur: Selbstorganisation
- Relativ unintelligente Einzelelemente bilden übergreifende Ordnung
- Flexibel, robust, dezentral, fehlertolerant
- Zielgerichtete Aufgabenerfüllung, Lösung von Optimierungsproblemen
- Mensch: Schwarmverhalten lediglich erste Stufe der kollektiven Intelligenz



Intelligenz und Kreativität im Web 2.0

Kollektive Intelligenz

- Netzwerke: nicht hierarchische Kommunikation
- Intelligente Einzelelemente
- Unkontrollierte (unkontrollierbare) Dynamik
- Anything goes: Massenproduktion von Anregungen und Denkbarem



Intelligenz und Kreativität im Web 2.0

Kollektive Intelligenz II: Die Dynamik zähmen

- Problem: Chaos in den Daten, Intelligenz „erst danach“
- Lösung: klare Rahmenbedingungen, Distanz zu den Einzelobjekten, Strukturierung der Anregungen
- *Hierarchisches Handeln zur Beurteilung und Umsetzung der Ideen notwendig*





Tagging-Systeme als Forschungsgegenstand

Forschungsstand

Defizite:

- Geisteswissenschaftliche Zugänge?
- Tags als „seelenlose“ Zeichenketten?
- Lediglich funktionaler Beleg der These
emergierenden Sinns

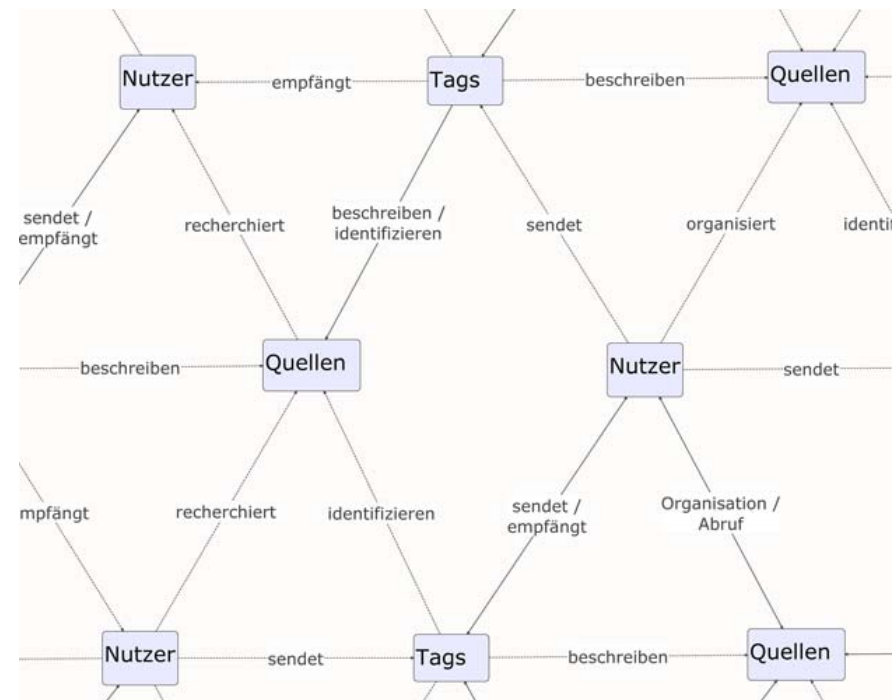
Vermessungen zur Kreativität der Sprache

Netzwerkkommunikation in Tagging-Systemen

- Medium: Sprachliche Symbole
- Ziel: Organisation und inhaltliche Erschließung von Ressourcen, Information Retrieval

Herausforderungen für den Tagger

- Kompression und Distinktion
- Mehrdeutigkeiten





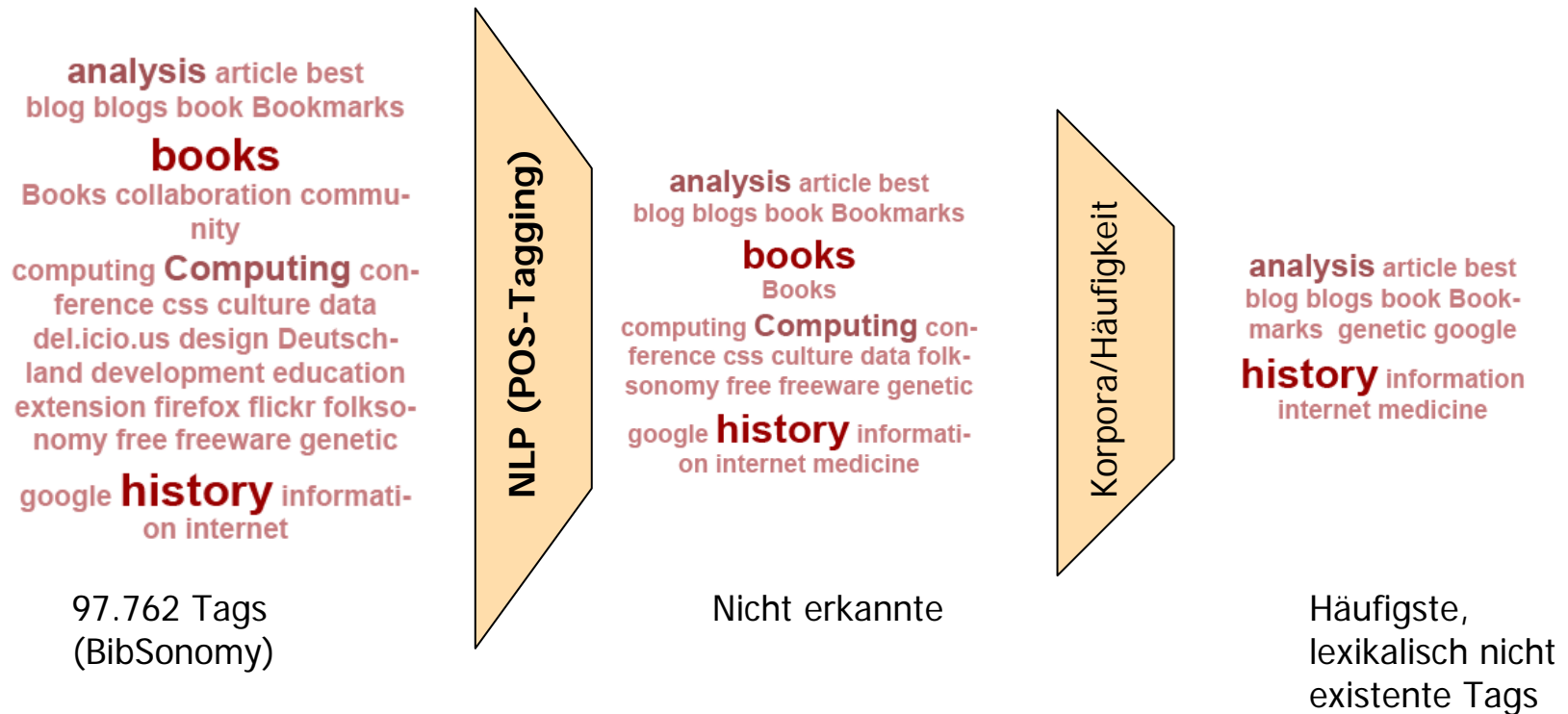
Vermessungen zur Kreativität der Sprache

Strategien der Tag-Vergabe

- Mehrheitlich: Übernahme / Verkettung bestehender Informationsfragmente (Zeichenfolgen aus Titel oder Volltext, Verwendung anderswo lexikalisch existenter Begriffe)
- Wortneuschöpfungen: Kreatives Potenzial (Neotags)
- **These:** Anhand der Neotags lässt sich die Kreativität eines Tagging-Systems messbar machen.
- Probleme: Identifikation von Neotags für intellektuelle Analyse, Messung der Kreativität

Vermessungen zur Kreativität der Sprache

Identifikation von Neotags I

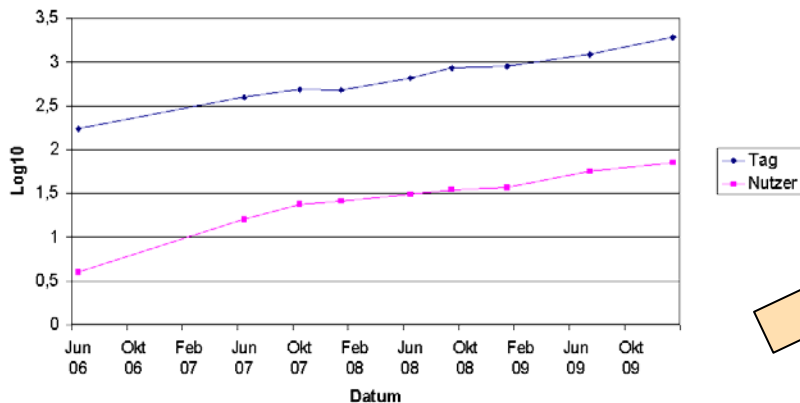


Vermessungen zur Kreativität der Sprache

Identifikation von Neotags II

1	N/Tag	Tag	Wortart	Zusatz1
2	4234	JRR	AK	
3	4234	CCP	AK	
4	1890	wismasys0809	NE	CD
5	1536	nlp	AK	
6	1452	file-import-09-02-13	XY	CD
7	1200	devo	XY	
8	1199	mjown	PPOSS	ADJ
9	1182	pfc	AK	
10	1167	folk.sonomy	NN	
11	1088	imtm-Bibliothek	AK	NN
12	1000	www.wiki	AK	

Intellektuelle
Recherche



Häufigkeit und Nutzeranzahl im
Zeitverlauf (2006-2009)

Rang (Nachnutzung)	Tag	Steigung
1	statecharts	3,66
2	mythesis	2,84
3	folkrank	2,80
4	socialnetworking	2,78
5	to-read	2,77
6	OpenSource	2,74
7	tagora	2,73
8	bibsonomy	2,73
9	openaccess	2,53
10	Varusschlacht	2,51
11	visualcommunica tion	2,49
12	eval	2,45

Neotags

Vermessungen zur Kreativität der Sprache

Messung der Kreativität

- Ranking nach linguistischen Kriterien
- Maß: kognitive Investition, Komplexität der Zeichenfolgenbildung

Zeichenbildung	Rang
Tag aus Titel	10
Tag aus Abstract/Text	9
Existent, aber nicht in Quelle	8
Komposition und Derivation (Aneinanderreihung/Affixe/Suffixe)	7
Konversion (unerwartete Wortklassenwechsel)	6
Kurzwortbildungen (Verkürzung)	5
Kontamination (Verschmelzung)	4
(Multiple) Mischformen	3
Analogiebildung	2
Neulexem	1



Intelligenz im Chaos? Erste Ergebnisse.

Struktur der Zeichenfolgen (500 häufigste Tags)

- Ausdruckslänge: sehr häufig 3-4 Zeichen
- Nutzeranzahl pro Tag: 1-2 (60%), 3-5 (16%), mind. 6 (22%)
- Wortarten: Nominative (50%), Akronyme (20%), Eigennamen (10%), Adjektive (10%)

Intelligenz im Chaos? Erste Ergebnisse.

Kreativität der Zeichenfolgen (500 häufigste Tags)

- Häufigste Klassen: Komposition zweier Wortformen, direkte Übernahmen aus Titel bzw. Text, Verkürzungen (Abkürzungen, Akronyme)
- Höhere kognitive Leistung eher selten
- Beispiele: socialnetworking (Komposition, Konversion), wwwkap2 (Abkürzung), semweb (Abkürzung und Komposition), open (-access, -source etc.) (Analogiebildung)

Fazit und Ausblick

- Tags sind mehrheitlich nicht besonders kreativ (keine „echten“ Neologismen)
- Tags mit hohen Steigerungsgraden (spreading) sind tendenziell „kreativer“
- Gut verständliche Neotags (z.B. folkrank, socialnetworking, eval, e-Learning etc.) können sich schnell durchsetzen
- Erhebung weniger häufiger Tags erforderlich (These: weniger verständliche Tags könnten kreativer sein)
- Verfahren und Modell aussichtsreich

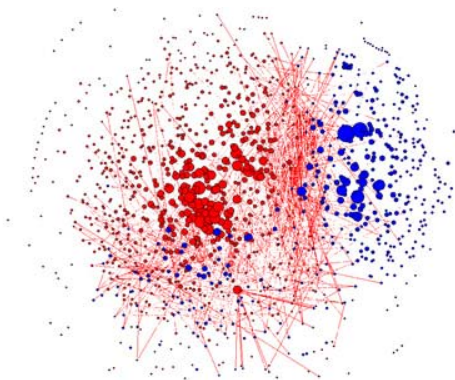


Vielen Dank für Ihre Aufmerksamkeit!

Kontakt:

christof.niemann@mail.uni-mannheim.de

www.ub.uni-mannheim.de/485.html



Gefördert durch