



FEMtech FTI



Sprachliche Konstruktion von Geschlecht. Ein Thema für die Computerlinguistik?

Brigitte Krenn (OFAI) & Karin Wetschanow (Uni Wien & INSET)



INSET[✿]
RESEARCH & ADVISORY

Spezifizierung

- Kommunikation in Social Media
- Automatische Genderklassifizierung (=gender guessing, identification, classification, ...)

Ist ein automatisches „Gender Guessing“
aus Sicht konstruktivistischer
Gendertheorien **sinnvoll** und mit
konstruktivistischen Gendertheorien
vereinbar?

- ▣ Gendered Technology
- ▣ Gendered Software
- ▣ Gendered Structures
- ▣ Gendered Communication

Dominante Genderstrukturen sind auch im Internet und damit in der Kommunikation im Internet Realität.

Gendered CMC

strukturell

- ▣ Insistieren auf Geschlechterbinarität in Anmeldeformularen
- ▣ Löschattacken bzw. -aufforderungen von Seiten mit genderquerdenkenden Inhalten

individuell

- ▣ Tendenz, in Social Media Foren immer weniger anonym zu agieren (Herring&Stöger 2013)
- ▣ "Genderswapping" eher die Ausnahme (Dorer, 2000; Lübke, 2005)
- ▣ Tendenz zu stereotyper Weiblichkeit und Männlichkeit
- ▣ „comeback of insisting on performing a ‚real‘ gender identity“ (Carstensen, 2009).

Schlüsselfragen (Speer, 2011)

Erforschung von „Geschlechterdifferenzen“

- ▣ Sprechen und interagieren Frauen und Männer unterschiedlich? Welche Eigenschaften charakterisieren ihr kommunikatives Handeln?
- ▣ Interpretation von Unterschieden?
- ▣ Andere Variablen?
- ▣ Wie wird Sexismus realisiert?

Erforschung von Genderkonstruktion

- ▣ Wie stellen Menschen Geschlechtsidentitäten interaktiv her und wie nutzen sie diese? Wie vollführen „Frauen“ Femininität und wie „Männer“ Maskulinität?
- ▣ Gender/Sexualität/Diskurs
- ▣ Micro-ebene innerhalb dominanter Makro-Ebene (Diskurs)?
- ▣ Wie gestalten sie communities of practice und wie werden vergeschlechtlichte Institutionen verhandelt?

Hypothesen

- Die diskursive Konstruktion von „Geschlechtsidentität“ tritt in virtueller Kommunikation in den Vordergrund.
- Diskurs ist hier der Ort der Konstruktion von „Gender“ und wird einer Untersuchung direkt zugänglich.
- Automatische Genderklassifizierung kann nützlich sein, um Prozesse der Genderkonstruktion zu erkennen.

Behauptete methodologische Kluft (vgl. MAASS et al. 2007)

Computer Technik/ Computer Wissenschaft

- ▣ Braucht klare distinkte Kategorien

de(konstruktivistische) gender studies

- ▣ Hinterfragen von Kategorien
- ▣ Dekonstruktion von Kategorien

Klassifizierung=Differenz?

- Klassifizierung=Differenz= Reifizierung bestehender Dichotomien?
- Ist eine Klassifizierung notwendigerweise eine Erforschung der „Differenz“?
- Kann Sie auch der Erforschung der „Differenzierung“ (Gildemeister 2008) dienen?

Was tun automatische Verfahren?

Automatische Verfahren

- ▣ schreiben nicht automatisch eine Differenz in die Daten ein.
- ▣ klassifizieren und differenzieren nicht in einem sozialen Wertesystem.
- ▣ sind Werkzeuge.

Was tut automatische Textverarbeitung?

- ▣ **Identifiziert sprachliche Einheiten**

z.B.: Wörter, Part-of-Speech, Phrasen, Satzstrukturen, Satzgrenzen, Emoticons, ...

- ▣ **Weist sprachlichen Einheiten Bedeutung zu**

z.B.: Wortbedeutungen: LIWC-Kategorien, Eigennamen, Ortsnamen, Ober- u. Unterkategorien aus Thesauri , ...

- ▣ **Identifiziert Wortkontexte**

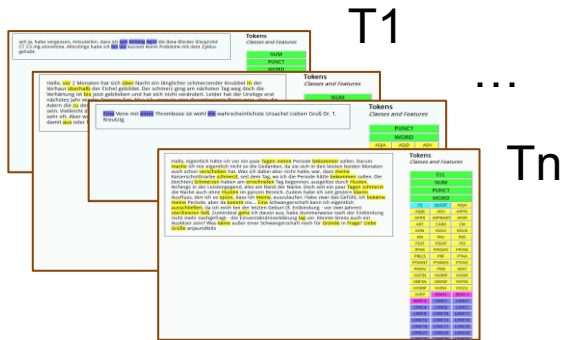
z.B.: n-Gramme $w_{i-n}, \dots, w_{i-1}, w_i, w_{i+1}, \dots, w_{i+n}$

Alles mit einer gewissen Fehleranfälligkeit (~1-40%) aufgrund der hohen Ambiguität von Sprache.

Automatische Textverarbeitung

- lernt Muster basierend auf großen annotierten Datenmengen (Tausende bis hin zu vielen Millionen Daten).
- leitet aus vorklassifizierten und annotierten Daten statistische Modelle ab.
- wendet diese Modelle auf neue Daten angewendet werde an.
- Annotierungen erfolgen manuell und automatisch.

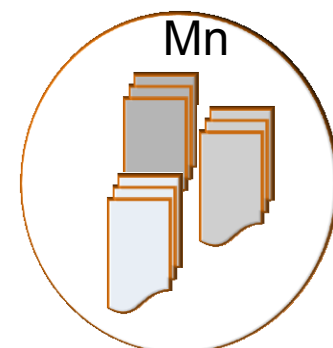
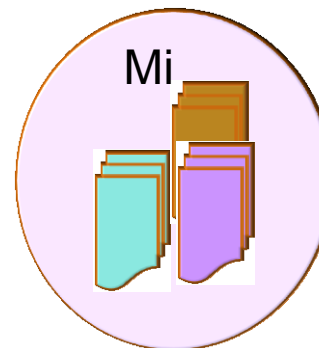
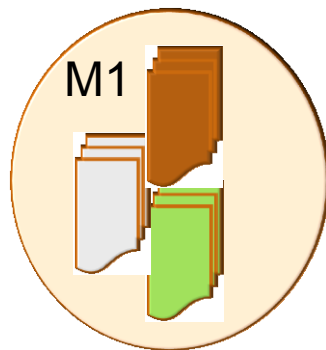
Annotierte Trainingstexte



Maschinelles Lernen

Modellbildung abhängig von

- Was annotiert wird
- Welche statistischen Modelle angewandt werden



rangeing ... | Mi Es reicht ... | FemSM... | google - ... | anekdotis... | erich gstr... | Hotel Gol... | +

Femsma.ofai.at/internal/corpus_workbench.html | winterpalais prin |

Most Visited | An alle Frauen. Penisl... | Getting Started | Latest Headlines | ORF 2 - tv.ORF.at

FemSMA Corpus Workbench

brigitte ?

Select Annotations

Annotator: Karin Wetschanow

Label Group: Solidarity

Haarausfall Forum Frauen / Superunglücklich

Haarausfall Forum Frauen / Superunglücklich

Type:	Forum
Topic:	Haarausfall
URL:	http://www.alopezie.de/foren/frauen/index.php/t/3833/
Number of messages:	122
Average message length:	1042
Annotation counts per label group:	
	1 Gender indexicity
	539 I
	12 Opposition
	64 Solidarity

HarleyQuinn + Assign Gender

2012-11-30 09:43:00 1547

Es muss auf jedenfall ein anständiger Bluttest her. Und zwar muss der zwischen ZT3 und 5 erfolgen. Für Progesteron sollte man um ZT21 testen. Wenn es heißt die Schilddrüse sei ok, muss das nicht stimmen. Da bräuchten wir mal die Werte. Man darf sich auf keinen Fall von Ärzten einfach abspeisen lassen. Mach am besten direkt nen Termin beim Endokrinologen. Der ist Spezialist für Hormone. Das dauert allerdings, bis man da nen Termin kriegt. Ferritin, Vitamin D und B12 sollten auch getestet werden. Wenn da ein Mangel vorhanden ist, ist das für die Haare schädlich. Ich finde, es ist eine Unverschämtheit einer 12jährigen die Pille zu geben. Zu hohes Testosteron kann in dieser Lebensphase normal sein, weil der Hormonhaushalt sich erst einpendeln muss. Nicht umsonst soll man nach der Pille 3 bis 6 Monate warten, bis sich der Hormonhaushalt reguliert hat. Ansonsten sind Tests verfälscht. Die Pille kann sehr viel Schaden anrichten. Ich habe das selber feststellen müssen. Ich hatte

Text Annotations
Solidarity

agreement

good_wishes

we

encouragement

thanks

sympathie

Save Annotations

Delete Selected

Delete All

Find: | Previous | Next | Highlight all | Match case

Manuelle Inhaltsanalyse und Annotierung

Hallo, vielen Dank für die ersten Infos. Ich versuche mal auf alles einzugehen... QUOTATION Vermännlichung, oh Gott, ja... kann ich denn gar nichts dagegen tun??? smileysad QUOTATION Ja, so sieht's leider aus... Ich wünschte das hätte ich nicht genommen, aber kann man ja jetzt nicht mehr ändern. Mich regt es aber tierisch auf dass man nie über die Nebenwirkungen der Pille informiert wird!! Als ich vor kurzem gelesen habe was für Probleme die Pille alles verursachen kann... diese ganzen Nährstoffmängel die davon auftreten, das erklärt so viele Probleme die ich hatte und nie hat mir irgendein Arzt gesagt dass es von der Pille kommt!! Find ich unverschämt!! QUOTATION Ich fühl mich gerade etwas erschlagen von den ganzen Abkürzungen und Fachwörtern in diesem Forum, bin ja ganz neu dabei. ZT steht für Zyklastag?? Demnach würde ich diesen Bluttest beim Frauenarzt machen lassen? Könnt ihr mir sagen ob das was kostet? QUOTATION Dafür nochmal ein extra Bluttest? Bei welchem Arzt kann ich den machen? Beim Endokrinologen war ich schon, der hat mir vor ca. einem Jahr die Schilddrüse untersucht. Werte habe ich nie gesagt bekommen. In der Praxis wird das so gehandhabt dass man zu einer bestimmten Sprechzeit bei dem Arzt anruft um die Ergebnisse der Untersuchung gesagt zu bekommen, und der meinte dann am Telefon nur zu mir dass alles OK wäre... Mehr weiß ich dazu nicht, ich weiß ehrlich gesagt nicht mal genau was eigentlich untersucht wurde. Meint ihr ich kann mir das an der Rezeption der Praxis ausdrucken lassen? Und bringt ein erneuter Test überhaupt was nach ca. einem Jahr? Immerhin habe ich ja die Pille jetzt abgesetzt, da müsste sich doch was geändert haben... QUOTATION Was, ich könnte Krebs kriegen wegen dem Haarausfall??? smileyeek

Tokens

Classes and Features

ELL		
EXCLAM		
IMGEMO		
PUNCT		
QUOTE		
WORD		
CAP	ITJ	SWEAR
QUEST	ADJA	ADJD
ADV	APPO	APPR
APPRART	APZR	ART
CM	KOKOM	KON
KOUI	KOUS	NE
NN	PAV	PDAT
PDS	PIAT	PIDAT
PIS	PPER	PPOSAT
PRELS	PRF	PTKA
PTKANT	PTKNEG	PTKVZ
PTKZU	PWAT	PWS
SENT	VAFIN	VAIMP
VAINF	VMFIN	VVFIN
VVIMP	VVINP	VVIZU
VVPP	SENT1	SENT2
SENT-3	SENT-4	LIWC1
LIWC2	LIWC4	LIWC5
LIWC6	LIWC7	LIWC8
LIWC9	LIWC10	LIWC12
LIWC13	LIWC15	LIWC16
LIWC18	LIWC19	LIWC20
LIWC21	LIWC22	LIWC23
LIWC25	LIWC26	LIWC31
LIWC32	LIWC33	LIWC35
LIWC36	LIWC37	LIWC38
LIWC39	LIWC40	LIWC41
LIWC42	LIWC44	LIWC45
LIWC46	LIWC47	LIWC48
LIWC49	LIWC50	LIWC51
LIWC53	LIWC56	LIWC57
LIWC58	LIWC60	LIWC61
LIWC65		

Automat. Inhaltsanalyse und Annotierung

Was muss ein automatisches „Gender Guessing“, das von einem konstruktivistischen Verständnis von Gender ausgeht, **beachten**?

Geschlecht als

- ▣ kommunizierte Geschlechtsidentität,
- ▣ semiotische Gestalt,
- ▣ performativ inszenierte Bedeutung (Butler 1991:61) verstehen.

Eine konstruktivistische automatische Klassifizierung muss

- ▣ den **Genderkategorien der User*innen** Rechnung tragen.
- ▣ **explizite Genderklassifizierungen** qualitativ erheben
- ▣ ihre **Klassen** aus der qualitativen Analyse ableiten.
- ▣ Relevante Kontextfaktoren erheben und
- ▣ **kontextsensitive** Modellierungen vornehmen.
- ▣ Sprachliche „Merkmalscluster“ müssen in einem **Kontinuum** angeordnet werden.

Beispiel explizite Genderindexikalisierung

- ▣ <http://www.butch-femme.de/forum/viewtopic.php?f=8&t=1939&start=90>
- ▣ **Thread: genderlose Erziehung?**
- ▣ 97 Beiträge
- ▣ 24 User*innen
- ▣ 14 verschiedene Genderidentitäten

Beispiel explizite Genderindexikalisierung

- ▣ Femme (4)
- ▣ Butch (3)
- ▣ Queere femme (2)
- ▣ Weiblich (2)
- ▣ Tg gentle Butch
- ▣ TG Stone Butch
- ▣ trans*butch
- ▣ TM
- ▣ High Femme
- ▣ queer femme freak
- ▣ ?
- ▣ Queer
- ▣ Androgyn/pansexuell
- ▣ Gender fluid

Welche **Theorien** erscheinen
nützlich, um ein sprachliches
doing/performing/indexing gender
sinnvoll zu operationalisieren?

Indexing Gender (Kotthoff, 2012; Ochs 1992)

- Wenn wir Gender als „semiotische Gestalt“ konzipieren, dann können wir diese Gestalt klassifizieren, ohne bestehende Dichotomien zu reifizieren.
- Indexing Gender ist eine nützliche Theorie
 - Explizite Genderindexikalisierung
 - Implizite Genderindexikalisierung

Indexing Gender

- Explizite (direkte) Genderindexikalisierung kann auf Identitäten jenseits bestehender Dichotomien verweisen/ sie erzeugen
- Implizite (indirekte) GI
 - stilistische Realisierungen von Sprechaktivitäten
 - innerhalb einer Gesellschaft,
 - die mit historisch entstandenen Assoziationen verbunden
 - d.h. Gender-Folien also Indexe (Verweise) auf in einer Community als weiblich/männlich assoziiertes Verhalten
- Implizite GI kann nur funktionieren
 - Innerhalb einer community of practice,
 - die dieses Doing identifizieren und interpretieren kann,
 - was aber zwangsläufig auch zu einer Klassifizierung führt.

Vorgehen

Qualitative DA von

Expliziter
Genderindexikalisierung
Impliziter GI =
Stilisierungsmittel
Relevanter Kontextfaktoren

Automatische

Annotierung
Musterbildung/Typenbildung
(Kontinuum)
Modellbildung

Antwort

- ▣ Klassifizierung im Sinne einer automatischen Sprachverarbeitung heißt nicht mehr als **datengetriebene Kategorisierung**.
- ▣ Was hinter der Interpretation der jeweiligen Kategorien steht, ist **theorie- bzw. hypothesengetrieben** und eine Frage der Geistes- und Sozialwissenschaften
- ▣ Die automat. Textverarbeitung und das maschinelle Lernen stellen **Werkzeuge** bereit, um das, was theoriegetrieben modelliert wurde mit automat. Mitteln in Texten wiederzufinden.
- ▣ Manche Mittel sind besser geeignet manche schlechter; das gilt auch für die Theorien.

Antwort

- ▣ Ja, wenn Genderexpert_innen im Team eingebunden sind
- ▣ Wenn Kategorien vor einem gendertheoretischen Hintergrund erstellt werden.
- ▣ Wenn Ergebnisse vor einem gendertheoretischen Hintergrund reflektiert werden.
- ▣ Wenn Ergebnisse innerhalb einer gendertheoretisch interessierten Community diskutiert werden.

Genderklassifizierung als Emanzipatorisches Tool?

- Automatische Genderklassifizierung von Texten
- Kann dazu beitragen komplexe Vergeschlechtlichungsmechanismen im Internet zu erforschen.

Butler, Judith (1991): *Das Unbehagen der Geschlechter*. Frankfurt a.M.: Suhrkamp.

Carstensen, T. (2009, March 5-7, 2009). *Gender Trouble in Web 2.0. Gender perspectives on social network sites, wikis and weblogs*. Paper presented at the 5th European Symposium on Gender & ICT. Digital Cultures: Participation - Empowerment - Diversity, University of Bremen.

Dorer, J. (2000). Geschlechterkonstruktionen in der Aneignung und Anwendung des Internet. Ergebnisse einer qualitativen Studie. *Medien & Zeit*, 2.

Herring, S. C., & Störger, S. (in press). Gender and (A)nonymity in Computer-Mediated Communication. In J. Holmes, M. Meyerhoff & S. Ehrlich (Eds.), *Handbook of Language and Gender* (2nd edition ed.). Hoboken, N.Y.: Wiley-Blackwell Publishing.

Kotthoff, H. (2012). "Indexing gender" unter weiblichen Jugendlichen in privaten Telefongesprächen. In Günthner, S., Hüpper, D. & Spieß, C. (Hrsg.), *Genderlinguistik. Sprachliche Konstruktionen von Geschlechtsidentitäten* (S. 251-285). Berlin: de Gruyter.

Lübke, V. (2005). *CyberGender : Geschlecht und Körper im Internet*. Königstein/Taunus: Helmer.

Ochs, E. (1992). Indexing Gender. In Duranti, A. & Goodwin, C. (Hrsg.), *Rethinking Context* (S. 335-358). Cambridge: Cambridge University Press.

Speer, S., & Stokoe, E. H. (2011). An introduction to conversation and gender. In S. Speer & E. H. Stokoe (Eds.), *Conversation and gender* (pp. 1-29). Cambridge: Cambridge Univ. Press.