



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



Maschinelles Lernen in den digitalen Literaturwissenschaften

8.12.2023

Hans Ole Hatzel

Über Mich

- Doktorand Informatik, Universität Hamburg, Language Technology Gruppe (Prof. Chris Biemann)
- Beschäftigt in Unterprojekten im Rahmen des SPP-CLS
 - Dissertation zur Modellierung von Geschichten/Erzählungen
- SPP-CLS: Schwerpunktprojekt der DFG zum Thema **Computational Literary Studies**

Agenda

- Was sind digitalen Literaturwissenschaften?
 - Beispiele
- Verwendete Methoden
- Forschungsprozess und Anforderungen an LLMs
- Ausblick

1

Was verstehen wir unter digitalen Literaturwissenschaften?

Digitale Literaturwissenschaft

- Eng. Computational Literary Studies
- Verwendung digitaler Methoden für die Analyse von (typischerweise) Belletristik
- Historisch eher statistische Analyse, heute immer mehr maschinelles Lernen

Beispiele für digitale Literatur Wissenschaft

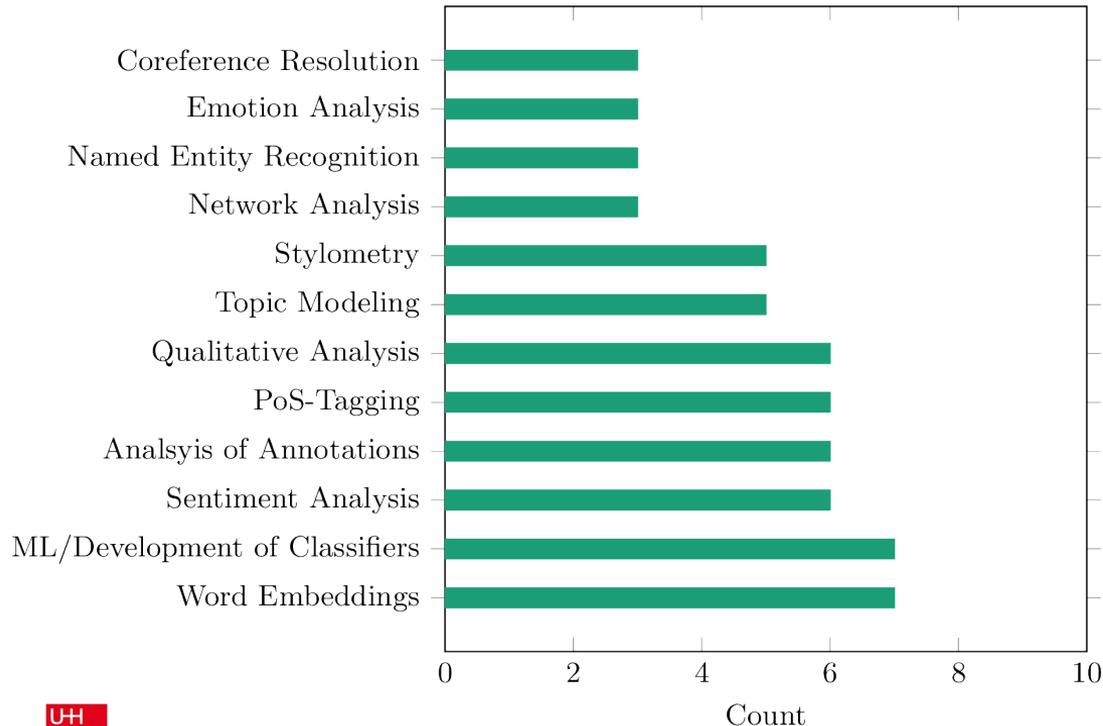
- Stylometrie
- Häufigste Wort-Vektoren Abstände
 - Most Frequent Word (MFW) Analyse
- Erfolgreiche Anwendungen z.B.
 - Etwa Identifikation von J.K. Rowling als Autorin von „The Cuckoo’s Calling“
 - Aber auch z.B. für literarische Qualität
- All das fällt eher unter Statistik

und 0.001
die 0.02
der ...
in
sie
ich
zu
den
er
das
dem
es
...

2

Welche Methoden kommen zum Einsatz?

Techniken - Laut Umfrage (bis 2022)

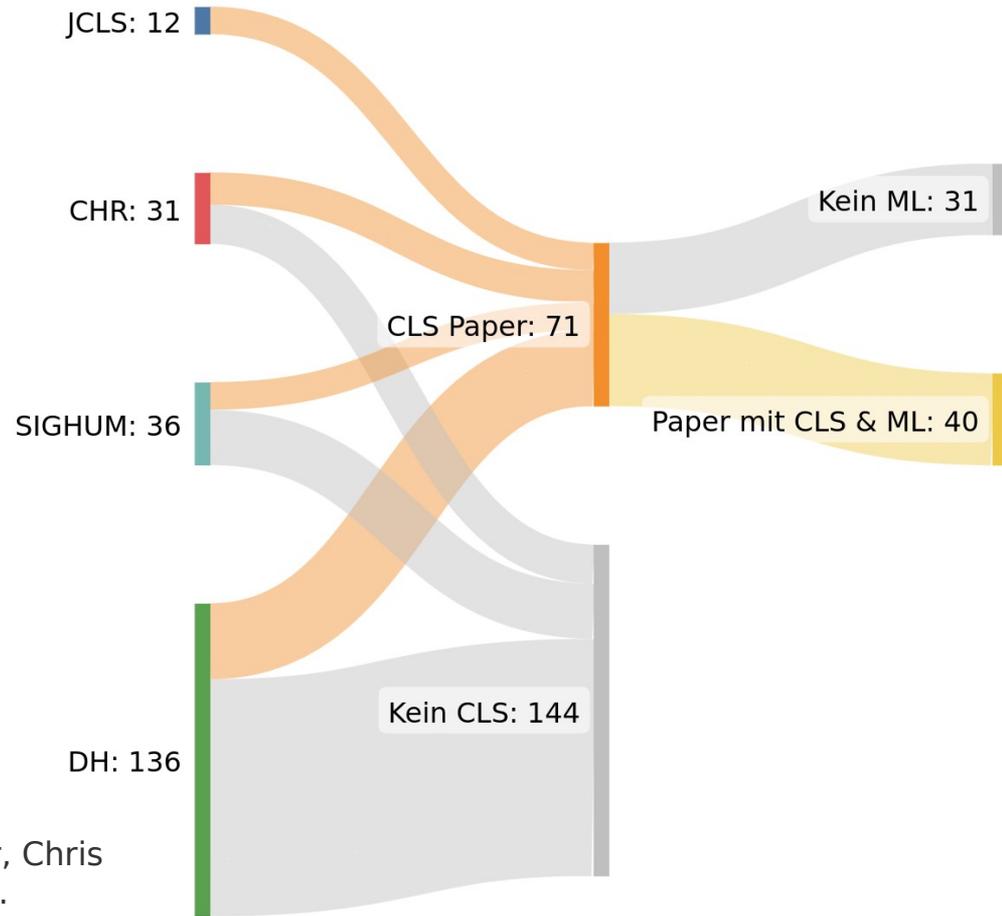


- Umfrage der Projekte im SPP-CLS
 - Insgesamt 11 Projekte
 - Helling et al., 2022
- Word Embeddings und Classifier sind sehr verbreitet
- Auch Methoden ohne maschinelles Lernen finden Anwendung
 - Z.B. Sentiment- und Annotations-Analyse
- Große Diversität bzgl. der Methoden

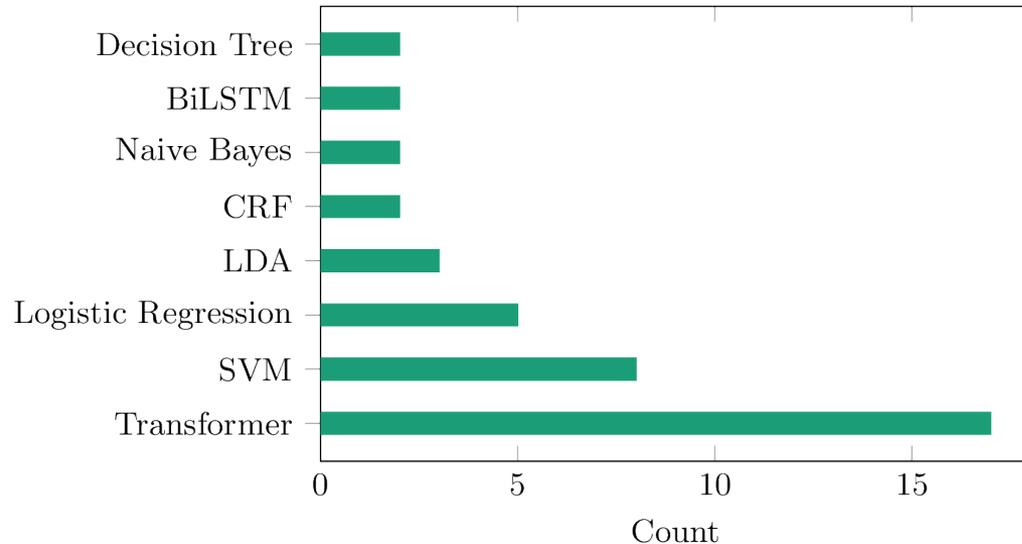
Survey Methode

- Erhebung aus Hatzel et al., 2023
- 4 beliebte Venues die für Inhalte aus der Digitale Literaturwissenschafts (CLS) bekannt sind
- Selektion nach tatsächlichen Inhalten zu CLS und ML (Machine Learning)
- Detailliertes Untersuchen von 40 Arbeiten

Hatzel, Hans Ole, Haimo Stierner, Chris Biemann, und Evelyn Gius. 2023. „Machine Learning in Computational Literary Studies“. It - Information Technology, August. <https://doi.org/10.1515/itit-2022-0041>



Techniken - Literatur (2022)



- LLMs finden hier namentlich keine Erwähnung
- Transformer-basierte Sprachmodelle: entsprechen LLMs, allerdings nicht generative LLMs
 - Nichtmal die Hälfte benutzen Transformer
- Meist: Embedding-basierte Ansätze oder Klassifikationsaufgaben

LLM Generative Modelle

- Klassifikationsmodelle: Eine der vordefinierten Klassen vorhersagen
- Oft meint man mit LLMs generative Modelle wie ChatGPT oder auch GPT-3
- BERT allerdings hat z.B. keine generativen Fähigkeiten
 - Kann aber trotzdem etwa das Sentiment eines Textes bestimmen



Beispiel Prompt Sentiment Analyse

H You

Hat der Satz: "Die geschäftlichen Aufregungen sind viel größer, als im eigentlichen Geschäft zuhause, und außerdem ist mir noch diese Plage des Reisens auferlegt [...]" ein positives oder ein negatives Sentiment?

 ChatGPT

Der gegebene Satz hat eher ein negatives Sentiment. Der Sprecher beschreibt die geschäftlichen Aufregungen als größer als im eigentlichen Geschäft zuhause und bezeichnet das Reisen als eine "Plage". Diese Ausdrücke deuten darauf hin, dass die Erfahrungen im Geschäftsbereich und das Reisen als belastend oder unangenehm empfunden werden.

3

Forschungsprozess und Anforderungen an LLMs

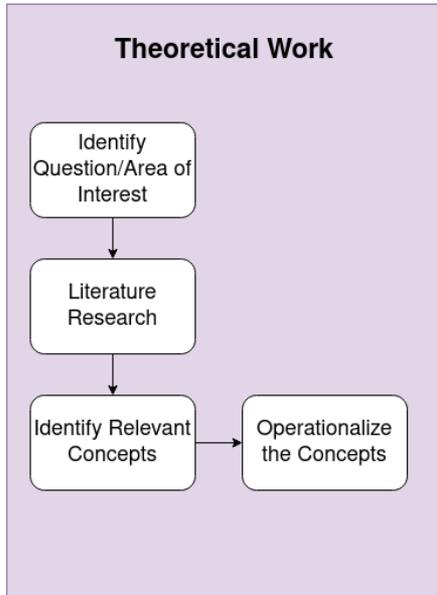
Typischer Forschungsprozess

Theoretical Work

Machine Learning
Methodology

- Theoretische Überlegungen werden in operationalisierbare Konzepte überführt

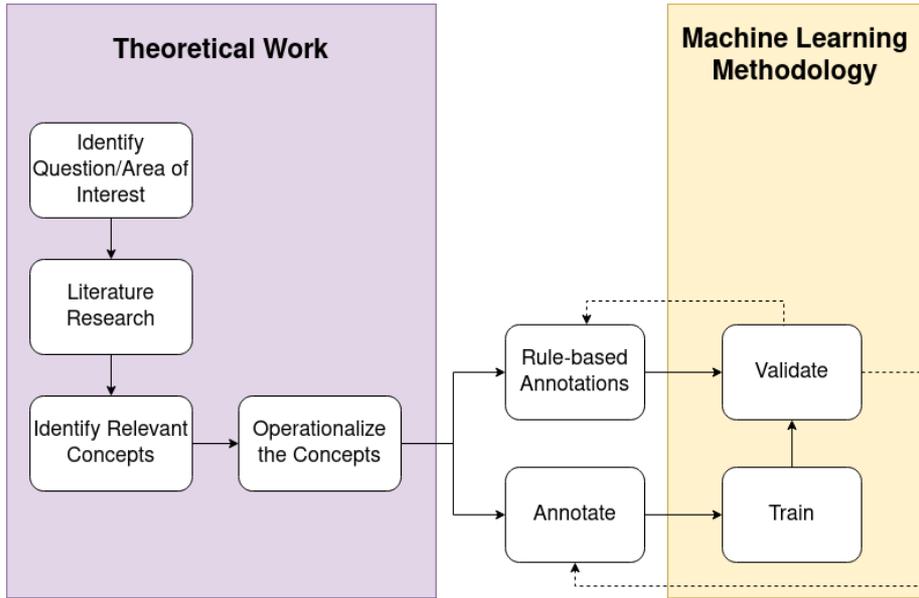
Typischer Forschungsprozess



Machine Learning Methodology

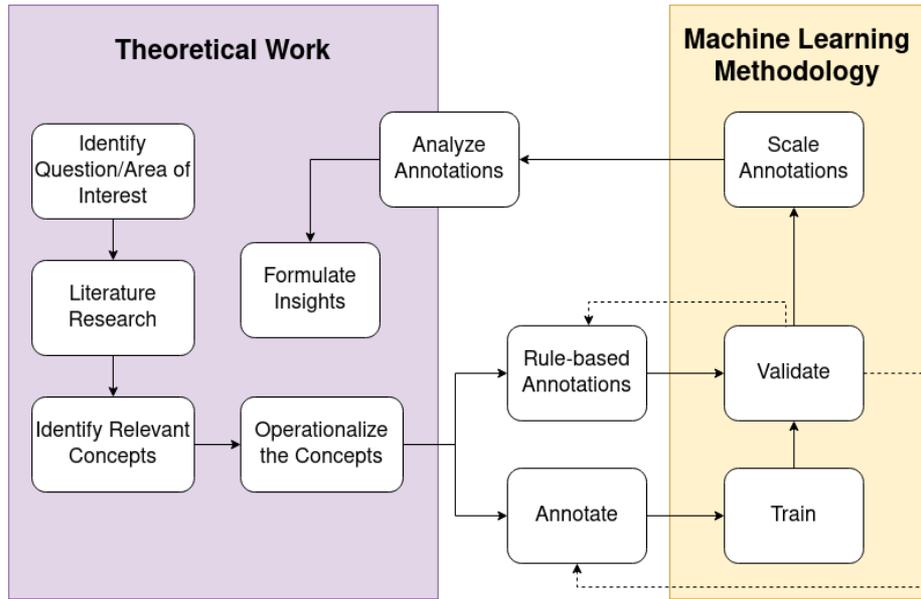
- Theoretische Überlegungen werden in operationalisierbare Konzepte überführt

Typischer Forschungsprozess



- Theoretische Überlegungen werden in operationalisierbare Konzepte überführt
- Diese werden entweder durch Training oder durch Regel-Systeme auf den untersuchten Text angewandt

Typischer Forschungsprozess



- Theoretische Überlegungen werden in operationalisierbare Konzepte überführt
- Diese werden entweder durch Training oder durch Regel-Systeme auf den untersuchten Text angewandt
- Auf Basis der automatisierten Annotationen kann wieder theoretische Arbeit vollzogen werden
- Rückweg in die theoretischen Konzepte!

Anforderungen an Modelle

- Eingabe von Konzepten: Definition der Konzepte für das LLM
 - Als Trainingsdaten oder als explizite Definition
 - Annotationen auf der Mikro-Ebene, etwa für einen einzelnen Satz
 - Aber auch Annotationen auf der Dokumentenebene
- Quantifizierbare Ausgaben
 - Die dennoch interpretierbar sein sollten
 - Kein generierter Text, der die Analyse vorwegnimmt

Modellintrospektion zum Erkenntnisgewinn

- Beispiel: Identifikation von „Translationese“ (Kunilovskaya et al. 2022)
 - Welche Eigenschaften unterscheiden russische Originaltexte von Übersetzungen ins Russische
 - Linear-Kernel SVM -> Feature Gewichte
 - Z.B. Feature Modal Verben:
 - „Er *könnte* die Antwort kennen“
 - **Statt des typischeren:** „*Wahrscheinlich* kennt er die Antwort.“
 - Viele Modal Verben -> Wahrscheinlich eine Übersetzung

Probleme bei der Anwendung von LLMs

- Modellinterpretation
- Ansatz: zero/few-shot Anwendung von Annotationsschemata
 - „Gegeben folgende Definitionen ... Markiere in diesem Text ...:“
- In unserer Survey gab es (fast) keine generativen-LLM Ansätze
 - 1. Zeitliche Gründe (ChatGPT Release Ende 2022)
 - 2. Konzepte aus den Literaturwissenschaften sind nicht immer Alltagssprachen kompatibel
 - Spezifische Trainingsdaten aus dem Kontext der DH in LLMs wären möglicherweise hilfreich

Ansatz Prompting

- Problem: Wir hätten gerne eine quantifizierbare Antwort auf eine Frage wie „Welches Sentiment hat dieser Text?“
 - Wir erstellen einen Prompt/eine Anfrage
 - Verwenden dann Constrained Decoding oder ähnliche Techniken => Generatives Modell als Klassifikationsmodell
 - Ein Problem dabei: es ist quasi das Gegenteil von chain-of-thought Prompting

Handhaben Langer Texte

- Oft reicht es, die Texte zu zerlegen
- Recursive Summerization, z.B. Wu et al., 2021
- Große Eingabefenster
 - 128k Tokens in GPT-4-Turbo -> reicht oft für einen ganzen Roman
 - Ähnliches auch in frei verfügbaren Modellen
 - Qualität noch etwas unklar
 - Prinzipiell quadratischer Speicheraufwand durch den Attention-Mechanismus

Fazit

- LLMs erlauben uns jetzt schon, viele lokale Annotationsaufgaben zu lösen
 - Abstraktere Aufgaben, etwa auf der Dokumentenebene bleiben noch schwierig
 - => Die Erschließung ganzer Bücher ist aktuell noch nicht wirklich in Reichweite
- Das Verarbeiten ganzer Dokumente bzgl. komplexer Konzepte bleibt schwierig, LLMs haben hier noch Defizite
- Interpretierbarkeit bleibt ein Problem

Fragen und Diskussion



Literaturverzeichnis

- Cranenburgh, Andreas van, und Erik Ketzan. 2021. „Stylometric Literariness Classification: the Case of Stephen King“. In Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, 189–97. Punta Cana, Dominican Republic (online): Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.latechclfl-1.21>.
- Hatzel, Hans Ole, Haimo Stiemer, Chris Biemann, und Evelyn Gius. 2023. „Machine Learning in Computational Literary Studies“. It - Information Technology, August. <https://doi.org/10.1515/itit-2023-0041>.
- Helling, Patrick, Kerstin Jung, und Steffen Pielström. 2022. „Pragmatisches Forschungsdatenmanagement - qualitative und quantitative Analyse der Bedarfslandschaft in den Computational Literary Studies“. In , herausgegeben von Michaela Geierhos, Peer Trilcke, Ingo Börner, Sabine Seifert, Anna Busch, und Patrick Helling. <https://doi.org/10.5281/zenodo.6328021>.
- Kunilovskaya, Maria, Ekaterina Lapshinova-Koltunski, und Ruslan Mitkov. 2021. „Translationese in Russian Literary Texts“. In Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, 101–12. Punta Cana, Dominican Republic (online): Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.latechclfl-1.12>.
- Wu, Jeff, Long Ouyang, Daniel M. Ziegler, Nisan Stiennon, Ryan Lowe, Jan Leike, und Paul Christiano. 2021. „Recursively Summarizing Books with Human Feedback“. arXiv. <https://doi.org/10.48550/arXiv.2109.10862>.
- Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, und Diyi Yang. 2023. „Can Large Language Models Transform Computational Social Science?“ arXiv. <https://doi.org/10.48550/arXiv.2305.03514>.



Kontakt



Hans Ole Hatzel

hans.ole.hatzel@uni-hamburg.de

[@HansHatzel.bsky.social](https://bsky.app/profile/HansHatzel)

[@HansHatzel](https://www.instagram.com/HansHatzel)

Prompting Beispiel

„Als Gregor Samsa eines Morgens *<anno>*aus unruhigen Träumen erwachte*</anno>* [...]"

Um welchen der folgenden 4 Event Typen handelt es sich beim markierten Textabschnitt: Non-Event, Stative Event, Process-Event oder Change-of-State Event.

Antwort: Change-of-State

Prompting: Beispiel 2

H You

User

Hat der Satz: "Die geschäftlichen Aufregungen sind viel größer, als im eigentlichen Geschäft zuhause, und außerdem ist mir noch diese Plage des Reisens auferlegt [...]" ein positives oder ein negatives Sentiment? Gebe die Antwort auf einer Skala von 1-10 an, wobei 10 sehr positiv und 1 sehr negativ ist.

 ChatGPT

Der gegebene Satz hat ein eher negatives Sentiment. Auf einer Skala von 1-10, wobei 10 sehr positiv und 1 sehr negativ ist, würde ich diesem Satz eine Bewertung von etwa 3-4 geben, da er darauf hinweist, dass die geschäftlichen Aufregungen als belastend empfunden werden und das Reisen als eine Art Plage betrachtet wird.

