



Berner
Fachhochschule

Bachelor Thesis

Gender Gap Tracker für die Schweizer Medien

Studiengang

Bachelor of Science in Computer Science

Autor

Anina Aeschbacher und Felix Neidhart

Dozent

Prof. Dr. Mascha Kurpicz-Briki

Experte

Alain Joray

Version 1.0 vom 1. Juni 2023

Abstract

Die Bachelor Thesis „Gender Gap Tracker für die Schweizer Medien“ untersucht den *Gender Gap* in textbasierten Deutschschweizer Online-Medien anhand der Anzahl Zitate von Männern und Frauen. Die Datengrundlage bildet eine Datenbank mit über 350'000 Artikeln aus vier Nachrichtenportalen, die in einem separaten Vorprojekt [4] mittels Webcrawling aufgebaut wurde. Die Inspiration für dieses Projekt und das gewählte Vorgehen kam von der kanadischen Studie „The Gender Gap Tracker: *Using Natural Language Processing to measure gender bias in media*“ [5]. entweder alles gross oder klein...

Zum Ermitteln des Gender Gaps setzt diese Arbeit - in Anlehnung an die Vorbildsstudie - auf *Machine Learning* basierte Werkzeuge aus dem Bereich *Natural Language Processing*. Konkret verwendet die geschriebene Software ein *Spacy* Modell um mithilfe von *NER*, *POS*, *Coreference Resolution* und *Dependency Parsing* Zitate und Personen aus den Texten zu extrahieren. Zum Überprüfen der Qualität haben wir manuelle Test Sets erstellt, anhand derer wir die Performance der Software messen können.

Die Resultate zeigen einen *Gender Gap* von 50.34%, wobei das beste Nachrichtenportal bei 41.77% und das schlechteste bei 57.70% liegt. Das bedeutet, dass der Algorithmus im Schnitt 50.34% weniger Zitate von Frauen gefunden hat als von Männern.

Unter Berücksichtigung der Limitationen (Kapitel 6.3) bedeuten diese Ergebnisse, dass in der Deutschschweiz ein signifikanter *Gender Gap* in den textbasierten Online-Medien vorliegt und Frauen deutlich unterrepräsentiert sind, gemessen an der Anzahl der gefundenen Zitate. Der Unterschied ist damit grösser, als in den meisten anderen Studien, die für diese Arbeit zum Vergleich herbeigezogen wurden (Kapitel 2).

Inhaltsverzeichnis

Abstract	ii
1 Einleitung	1
2 Verwandte Arbeiten	4
2.1 Vorarbeit	4
2.2 Literaturrecherche	5
2.2.1 Methodik der Literaturanalyse	5
2.2.2 The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media	7
2.2.3 Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles	11
2.2.4 DIANES: A DEI Audit Toolkit for News Sources	13
2.2.5 The gender gap in commenting: Women are less likely than men to comment on (men's) published research	14
2.2.6 Using citation network analysis to enhance scholarship in psychological science: A case study of the human aggression literature	17
2.2.7 A Large-Scale Test of Gender Bias in the Media	18
2.2.8 Cancer in the news: Bias and quality in media reporting of cancer research	21
2.2.9 Academic medicine's glass ceiling: Author's gender in top three medical research journals impacts probability of future publication success	21
3 Methoden	24
3.1 Technologien	25
3.1.1 Python	25
3.1.2 Git	25
3.1.3 GitLab	25
3.1.4 MongoDB	26
3.1.5 Natural Language Processing (NLP)	27
3.1.6 Named Entity Recognition (NER)	27
3.1.7 Part Of Speech (POS) Tagging	27
3.1.8 Coreference Resolution	27
3.1.9 Dependency Parsing	27
3.1.10 Spacy	28
3.2 Projektplanung	28

4	Technische Umsetzung	30
4.1	Datensammlung bereinigen	32
4.2	Personen extrahieren	34
4.3	Geschlecht identifizieren	34
4.4	Extraktion der Zitate	35
4.4.1	Syntaktische Zitate	35
4.4.2	Parse Trees	35
4.4.3	Ablauf des Algorithmus	37
4.5	Qualitätssicherung	40
4.5.1	Recall	40
4.5.2	Qualitätsnote	42
4.5.3	Nicht erkannte Zitate	43
4.6	Datenbank	44
5	Resultate	47
6	Diskussion	51
6.1	Zusammenfassung der Ergebnisse	51
6.2	Interpretation der Resultate	51
6.3	Limitationen	53
6.4	Weiterführende Forschung	54
7	Schlussfolgerungen	55
7.1	Ziele evaluieren	55
7.1.1	Extraktion Zitate	55
7.1.2	Extraktion Personen	56
7.1.3	Evaluierung Geschlecht	56
7.1.4	Aggregation und Visualisierung	56
7.2	Persönliche Erkenntnisse	56
7.2.1	Literaturrecherche	56
7.2.2	Programmieren	57
7.2.3	LaTeX	57
7.2.4	MongoDB	57
7.2.5	Projektplanung	58
7.2.6	Zusammenarbeit in der Gruppe	59
	Literatur	62
	Abbildungsverzeichnis	68
	Tabellenverzeichnis	69
	Glossar	70
	Abkürzungen	73

1 Einleitung

Der siebte Februar 1971 war ein besonderer Tag für die Schweiz. Insbesondere aber für die Schweizerinnen, denn an diesem Tag nahmen die Schweizer - im zweiten Anlauf nach 1959 - das nationale Frauenstimmrecht mit 65% an bei einer Stimmbeteiligung von 57.72% ¹.

Ab diesem Zeitpunkt durften Frauen erstmals wählen, abstimmen und sich selber **zur Wahl stellen** lassen. Dieser Meilenstein beschleunigte die Anliegen der Frauen nach Gleichberechtigung. So wurden in den folgenden Jahren regelmässig Abstimmungen zur Annäherung der Rechte der Frau an den Mann verabschiedet.

Der nächste grosse Schritt erfolgte 10 Jahre später, als die rechtliche Gleichstellung von Frauen und Männern erstmals in die Verfassung aufgenommen wurde. Dieser Artikel wurde im Jahr 1996 verschärft mit der Ergänzung, dass der Bund nicht nur für die rechtliche, sondern auch für die tatsächliche Gleichstellung der Frau sorgen soll.

Ab 1988 stellte die Verfassungsreform zum Eherecht die Frau und den Mann in der Ehe gleich und verpflichtete beide Parteien **zur die** Pflege und Erziehung der Kinder und den Familienunterhalt.

Dann erst 1990 wurden auch die Frauen aus dem Appenzell Innerrhoden stimmberechtigt, als das Bundesgericht den Kanton durch eine Neuinterpretation seiner Verfassung gezwungen hatte, das Frauenstimmrecht einzuführen.

Doch trotz dieser gesetzlichen und gesellschaftlichen Fortschritte war die Schweiz im Jahre 1991 noch weit entfernt von rechtlicher und tatsächlicher Gleichstellung der Geschlechter. Deshalb riefen 1991 Frauenorganisationen zum ersten nationalen Frauenstreik der Schweiz auf.

Mehr als eine halbe Million Frauen gingen auf die Strasse im grössten Protest, den die Schweiz seit dem Landesstreik 1918 gesehen **hat**. Sie forderten die tatsächliche (nicht bloss rechtliche) Gleichstellung in der Gesellschaft, die Bekämpfung der horrenden Lohnungleichheit am Arbeitsplatz, gleiche Ausbildung für Frauen und ein Stopp der alltäglichen sexuellen Belästigung am Arbeitsplatz ².

Der Streik hatte bestimmt auch einen grossen Einfluss darauf, dass 1996 schliesslich die tatsächliche Gleichstellung der Frau gegenüber dem Mann in der Bundesverfassung verankert wurde ³.

Der nächste Meilenstein wurde 2002 mit der Legalisierung des Schwangerschaftsabbruchs (Abtreibung) bis zur 12. Schwangerschaftswoche gesetzt.

2004 wurde ein weiteres zentrales Anliegen der Frauenrechtsbewegung adressiert: Gewalt in der Ehe und in der Partnerschaft wurde als Offizialdelikt eingestuft. Dazu gehört Körperverletzung wie auch sexuelle Nötigung oder Vergewaltigung.

¹<https://www.bk.admin.ch/ch/d/pore/va/19710207/index.html>

²<https://hls-dhs-dss.ch/de/articles/058286/2019-06-12/#HDerStreiktagundseinEchoindenMedien>

³<https://www.ebg.admin.ch/ebg/de/home/themen/recht/gleichstellungsgesetz.html>

In den darauffolgenden Jahren wurden Gleichstellungsfragen breit in den öffentlichen und bald auch in den sozialen Medien diskutiert, dessen Höhepunkt sich wahrscheinlich in der #MeToo Debatte fand.

Diese Debatte feuerte 2019 auch den zweiten nationalen Frauenstreik an. Auch dieses Mal **mobilisieren** die Frauen über eine halbe Million Demonstrierende auf ihren Märschen durch die Schweizer Innenstädte. Sie prangerten den immer noch anhaltenden Sexismus in der Schweizerischen Gesellschaft und das Problem der sexuellen Gewalt an und forderten Anerkennung für Betreuungsarbeit und die Eliminierung der nach wie vor grossen Lohnungleichheit.

Denn zu diesem letzten Thema, der Lohnungleichheit oder *Pay Gap*, hatte das Statistik (BFS) nur einige Monate zuvor eine neue Studie veröffentlicht. Diese weist einen durchschnittlichen Lohnunterschied von 18.3% zwischen Männern und Frauen nach. Auch nach Eliminierung von Faktoren wie Ausbildung, Arbeitserfahrung oder Berufsgattung verbleiben 7.7% unerklärlich, die auf einen *Gender Bias* beim Lohn hindeuten.

Doch nicht nur beim Lohn gibt es noch immer bedeutende Unterschiede zwischen den Geschlechtern. Frauen sind in unserer Gesellschaft in wichtigen Positionen der Politik und Wirtschaft stark untervertreten. Nach den letzten Wahlen sassen im Nationalrat mit 96 Vertreterinnen gegenüber 150 Vertretern zwar so viele Frauen in der Grossen Kammer wie noch nie, doch von der Hälfte sind wir noch immer ein gutes Stück entfernt. Immerhin zeigt der Trend in die richtige Richtung mit einem Plus von 14% (+12) Frauen gegenüber dem Vorjahr ⁴. Doch der konservativere Ständerat und die Kantonsräte hinken mit 26% (12 / 46) respektive 30% deutlich hinterher ⁵.

Noch gravierender ist die Situation in der Wirtschaft bei den Führungspositionen und in den Unternehmensleitungen. So ist der Frauenanteil in diesen Positionen im Verlauf von 2012 bis 2022 nur um knapp 2% auf 36% angestiegen ⁶.

Neben den *Pay Gap* und dem sogenannten *Gender Bias in der Verteilung* in der Politik und Wirtschaft gibt es auch einen *Gender Gap* in den Medien, wie verschiedene Studien beweisen [5, 15, 50].

So messen beispielsweise Dacon und Liu in ihrer Studie „Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles“ [15] in zwei unterschiedlichen Datensets mit mit 160'000 bzw. 210'000 englischsprachigen Nachrichtenartikel einen signifikanten *Gender Bias in der Verteilung* von bis zu 53.9%.

Die Arbeit „A Large-Scale Test of Gender Bias in the Media“ von Shor, Rijt und Fotouhi [50] misst nicht nur den *Gender Bias in der Verteilung*, sondern teilt diesen zusätzlich auch auf in einen Bias, der der Struktur unserer Gesellschaft mit überproportional mehr Männern in Nachrichtenwerten Positionen geschuldet ist und in einen zweiten Bias, der die Medienschaffenden dazu veranlasst, mehr über Männer als Frauen zu schreiben.

Aber die wichtigste Studie für diese Arbeit wurde von Asr u. a. unter dem Titel „The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media“ [5]

⁴<https://www.parlament.ch/de/%C3%BCber-das-parlament/politfrauen/anteil-frauen-politischen-macht/praesenz-frauen-schweizer-parlament>

⁵<https://www.bfs.admin.ch/bfs/de/home/statistiken/nachhaltige-entwicklung/monet-2030/gleichstellung-geschlechter/anteil-frauen-nationalrat-kantonsparlamenten.html>

⁶<https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken.assetdetail.24065920.html>

veröffentlicht. Sie untersucht den *Gender Gap* in ausgewählten englisch- wie auch französischsprachigen kanadischen Medien anhand moderner *Natural Language Processing (NLP)* Werkzeuge und schaltet die neusten Erkenntnisse regelmässig online [6] auf, damit Interessierte den Fortschritt (oder Rückschritt) auf täglicher Basis mitverfolgen können. Zum Zeitpunkt der Veröffentlichung des Papers waren 71% aller Zitate Männern zuzuordnen.

An diese Studie wollen wir mit dieser Arbeit anknüpfen und nach dem gleichen Prinzip den *Gender Gap* in den Deutschschweizer Online-Medien messen. Zumindest in **denen** vier, von denen wir über die Daten verfügen. Die Arbeit soll zeigen, dass das **erkennen** von Zitaten mittels *NLP* Werkzeugen auch in deutschsprachigen Texten möglich ist. Anhand dessen soll sie herausfinden, wie viele der gefundenen Zitate von Männern sind und wie viele von Frauen. Diese Zahlen wollen wir im Anschluss untereinander, **vergleichen** und mithilfe der *Gender Gap* Formel (vgl. Abbildung 3.1) in Kontext mit den vergleichbaren Arbeiten setzen.

Dazu wollen wir zeitgemässe Werkzeuge des *NLP* einsetzen, wie *Named Entity Recognition (NER)*, *Part of Speech (POS)* Tagging oder *Dependency Parsing* um bestmögliche Resultate zu erzielen. Diese fungieren als Grundlage für unsere selbst geschriebenen Algorithmen, die je nach Aufgabe Personen herauslesen, Zitate extrahieren oder das Geschlecht einer Person bestimmen.

Die Arbeit ist in sieben Kapitel unterteilt, wobei Sie das erste nun schon fast hinter sich haben. Das nächste handelt von den verwandten Arbeiten und der durchgeführten Literaturrecherche zum Thema. Danach folgt ein Kapitel zu den Methoden, wo auch die verwendeten Technologien und theoretischen Konzepte genauer erläutert werden. Dort finden Sie auch einen Abschnitt zur Projektplanung inklusive Gantt Diagramm. Das Kapitel vier erörtert dann die technische Umsetzung und geht vertieft auf die Hauptaufgaben des *NLP* ein. Dieses technische Kapitel wird gefolgt von den Resultaten, wo sie mit den hübschen und aussagekräftigen Auswertungen dieser Arbeit beglückt werden. Im Anschluss werden diese Resultate im Rahmen der Diskussion in Kontext gesetzt und Limitationen und weiterführende Forschung erklärt. Und zum Schluss - wie könnte es anders sein - fasst das Kapitel Schlussfolgerungen die Ziele und Achievements der Arbeit zusammen und ergänzt diese mit einer persönlichen Reflexion. Wir wünschen viel Spass beim Lesen!

2 Verwandte Arbeiten

2.1 Vorarbeit

Im Projekt2 „Erstellung einer Datensammlung von Deutschschweizer Online Zeitungsartikeln“ [4] wurde eine Datensammlung von vier Deutschschweizer Online-Medien erstellt von Artikeln welche in den letzten 22 Jahren veröffentlicht wurden. Bestandteil dieser Arbeit war, einen spezifischen Webcrawler pro Nachrichtenportal zu bauen und automatisiert über einen GitLab Job Scheduler laufen zu lassen. Der gesammelte Datensatz umfasst mit über 300'000 Artikeln aus vier der fünf grössten frei zugänglichen Nachrichtenportalen [55] der Deutschschweiz einen bedeutenden Teil der online zugänglichen Zeitungsartikel, welche die hiesige Bevölkerung in den letzten Jahren gelesen hat. In der Abbildung 2.1 ist ersichtlich, wie viele Artikel wir am Ende von Projekt2 pro Nachrichtenportal gesammelt haben. Die Artikel von Blick, Watson, SRF und 20 Minuten sind in einer MongoDB Datenbank auf einer virtuellen Maschine der BFH gespeichert. In dieser Bachelor Thesis dient dieser Datensatz als Grundlage für die Analyse des Gender Gaps in den Deutschschweizer Online-Medien.

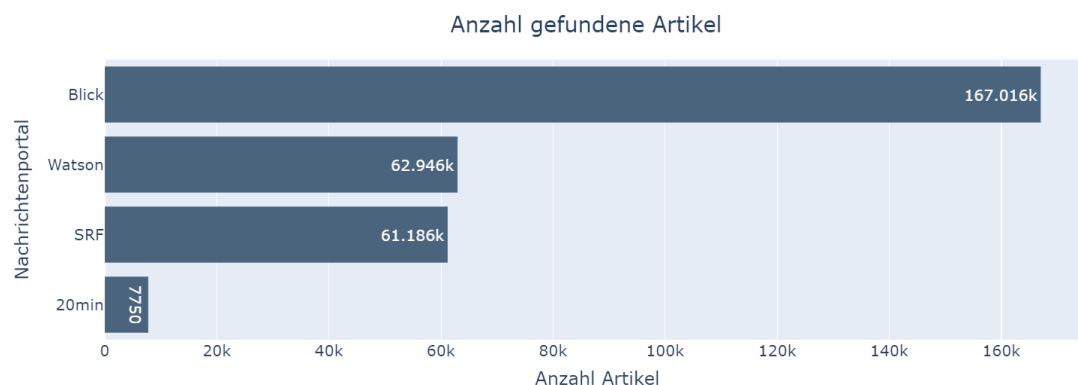


Abbildung 2.1: Auszug aus der DB Anzahl Artikel pro Nachrichtenportal, Stand 02.01.2023

Als Vorbereitung für diese Bachelor Thesis wurde eine grobe Methodik definiert für das Extrahieren der Zitate, Identifizieren der Personen und das Identifizieren des Geschlechts. Die Methodik wurde inspiriert vom kanadischen „The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media“ [5], welcher die Grundinspiration dieser Bachelor Thesis ist.

2.2 Literaturrecherche

2.2.1 Methodik der Literaturanalyse

Um die Resultate der Arbeit in Kontext setzen zu können und Inspiration von existierender Forschung zu erhalten, war es für uns - aber auch für den Erfolg der Arbeit - wichtig, den aktuellen Stand der Forschung aufzuarbeiten. Da die Menge an zur Verfügung stehenden Informationen sehr gross ist, war ein methodischer Ansatz gefragt. Mithilfe unserer Betreuerin, Prof. Dr. Mascha Kurpicz-Briki, haben wir die nachfolgende Herangehensweise besprochen, um die wichtigste Literatur abdecken zu können.

Die erarbeitete Methodik sah vor, dass als Erstes die relevantesten Portale von wissenschaftlichen Arbeiten identifiziert werden sollten. Als wir diese definiert hatten, suchten wir als Nächstes einen Filter, der die Arbeiten auf etwa 100 potenziell relevante Artikel eingrenzen sollte. Aus diesem Kandidatenset wollten wir im Anschluss die Abstracts durchlesen, um mindestens fünf und maximal 20 vergleichbare Arbeiten zu finden. Diese galt es dann als **ganzes** zu lesen und in den nachfolgenden Unterkapiteln der Literaturrecherche zusammenzufassen.

Auswahl der Portale

Für die Auswahl der Portale stützten wir uns auf die Empfehlungen unserer Betreuerin und ergänzten diese mit weiteren uns bekannten Plattformen. Schlussendlich durchsuchten wir die folgenden drei online Datenbanken mit wissenschaftlichen Artikeln.

1. ACM (dl.acm.org)
2. IEEE (ieeexplore.ieee.org)
3. PLOS (journals.plos.org)

ACM und IEEE sind bekannte Datenbanken für Informatik Themen und PLOS haben wir dazugenommen aufgrund der sozialwissenschaftlichen Perspektive.

Filterung der Portale

Der Filter sollte so eng gefasst sein, dass er möglichst viele irrelevante Arbeiten ausschliesst, gleichzeitig **auch offen genug sein**, damit wir eine genügend grosse Resultatliste erhalten. Um die geeignete Abfrage zu finden, mussten wir Variationen probieren, bis wir mit der Grösse des Resultats zufrieden waren.

Schlussendlich sollte unser Filter alle Artikel bezüglich *Gender Gap* oder Bias in den Medien ausgeben. Die genauen Kriterien sind **schwierig** in natürlicher Sprache auszuformulieren. Aus diesem Grund haben wir **Ihnen** nachfolgend beispielhaft das verwendete Query in der PLOS Syntax eingefügt.

Da die ursprüngliche Liste viele Artikel zu Social Media Themen oder der Covid Pandemie enthielt, haben wir diese explizit ausgeschlossen.

```
(abstract:gender) AND  
(  
  (abstract:media) OR  
  (abstract:news)  
) AND  
(  
  (abstract:gap) OR  
  (abstract:inequality) OR  
  (abstract:equality) OR  
  (abstract:bias)  
) AND NOT  
(  
  (abstract:"social media") OR  
  (abstract:pandemic) OR  
  (abstract:covid)  
)
```

Abbildung 2.2: Query zur Suche der relevanten Literatur

Manuelle Aussortierung anhand der Abstracts

Die Anwendung dieses Filters resultierte in einer Liste von 115 Arbeiten, **dessen** Abstracts wir im Anschluss durchlasen und dazu jeweils drei Fragen beantworteten.

1. Befasst sich der Artikel mit Geschlechter(un)gleichheit in den Medien?
2. Sind die untersuchten Medien textbasiert?
3. Wurde die Auswertung mithilfe von *NLP* durchgeführt?

Schlussendlich entschieden wir uns dazu, diejenigen Arbeiten zu verwenden, bei denen wir die ersten zwei Fragen mit „Ja“ beantworten konnten. Die dritte Frage hätte uns geholfen, die Liste weiter herunterzubrechen, wären mit den ersten zwei noch zu viele Arbeiten übrig geblieben.

Bei der Auswertung der Liste haben wir definiert, was wir unter textbasierten Medien alles einschliessen wollen. Wir entschieden uns dazu, Nachrichtenportale und wissenschaftliche Arbeiten als Medien zu definieren, Schulbücher und Kinderbücher jedoch aus dieser Definition auszuschliessen.

Nach dieser manuellen Suche nach den relevantesten Arbeiten enthielt unsere Liste noch acht Artikel aus zwei Portalen. So hat es kein Artikel von IEEE durch unsere Aussortierung geschafft.

Analyse der Arbeiten

Diese acht Arbeiten haben wir im Anschluss gelesen und in den nachfolgenden Kapiteln zusammengefasst. Der Fokus lag dabei auf den Themen der Geschlechter(un)gleichheit und den Analysemethoden, die der unseren ähneln. Für diese Arbeit irrelevante Themen, die ebenfalls in den Artikeln beschrieben wurden, haben wir weggelassen, um die Arbeit nicht unnötig zu verlängern.

Abweichungen von der Methode

In zwei Fällen sind wir von der oben beschriebenen Methodik abgewichen. **Nämlich konnten wir** aus der Vorarbeit noch einen weiteren Forschungsartikel, der uns relevant für unsere Arbeit schien. Dieser war in der Liste jedoch nicht enthalten, da er aus einem anderen Portal stammt. Es handelt sich dabei um die Arbeit namens „A Large-Scale Test of Gender Bias in the Media“ [50]. Die Zusammenfassung dieser Arbeit finden Sie zusätzlich zu den anderen in den nachfolgenden Seiten. Die zweite Abweichung ist der Ausschluss der Arbeit „Information Ecosystem Threats in Minoritized Communities: Challenges, Open Problems and Research Directions“ [17]. Es hat sich herausgestellt, dass dieser Artikel keine Untersuchung beschreibt, sondern die Zusammensetzung und Stossrichtung einer geplanten Forschungsgruppe. Diese wird unter anderem auch die Geschlechter(un)gleichheit in den Medien untersuchen. Im vorliegenden Paper ist jedoch noch kein Ergebnis vorhanden. Diese Tatsache war im Abstract jedoch nicht klar beschrieben.

2.2.2 The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media

Die wissenschaftliche Arbeit „The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media“ [5] von Asr u. a. aus dem Jahr 2021 ist das Vorbild für die vorliegende Thesis. Im Rahmen ihrer umfassenden Untersuchung der kanadischen, englischsprachigen Medien haben die Autor:innen nicht nur den Gender Gap anhand an der Anzahl Zitate nach Geschlecht und Nachrichtenportal untersucht, sondern auch neue Standards zur Messung und kontinuierlichen Überwachung der Repräsentation von Männern und Frauen in den Medien geschaffen.

Aufgrund der ausführlichen Beschreibung ihrer Vorgehensweise und der geringen Eintrittshürde in Bezug auf notwendiges vorprozessiertes Datenmaterial, kann diese Methodik auch gut in anderen **Kultur** und Sprachräumen nachgeahmt werden. Da dieses Projekt genau das vorhat, ist diese Arbeit in einem grossen Detailgrad beschrieben.

Ziele der Arbeit

Das Ziel der Arbeit war es **den** Unterschied in der Repräsentation der Geschlechter in den kanadischen Medien sichtbar zu machen. Als Mass für die Repräsentation verwenden Asr u. a. die Anzahl der Zitate in den Online-Nachrichtenportalen, die von Männern und Frauen publiziert werden.

Dazu haben sie eine ausgeklügelte Applikation entwickelt, die auf regelmässiger Basis die neusten Artikel der Portale mittels Webcrawling herunterlädt und analysiert. Für die Untersuchung der Texte verwenden sie verschiedene Arten von *NLP*, über die folgenden Unterkapitel mehr Aufschluss geben sollen.

Schlussendlich visualisieren sie die Daten auf einer öffentlich zugänglichen Plattform in der Form einer interaktiven Webapplikation. Die Resultate aus der Arbeit stammen direkt von dieser Software.

Die stets aktuellen Daten des *Gender Gap Tracker (GGT)* werden bewusst öffentlich zur Verfügung gestellt, um den Medienschaffenden Anreize zu verschaffen, für mehr Diversität in ihrer Berichterstattung zu sorgen. In der Arbeit vergleichen sie den *GGT* mit einem Fitness Tracker, der den Menschen helfen soll für ihrer Fitness zu sorgen. Analog zu diesem Prinzip soll dieses Tool den Medienkonzernen mithilfe von Transparenz und Rückverfolgbarkeit helfen, ihren *Gender Bias* zu verringern.

Zitate extrahieren

Der Kern der Aufgabe des *GGT* ist das Extrahieren von Zitaten. Die Entwickler:innen haben sich dazu eines vortrainierten, syntaktischen Parsers von Spacy bedient, der Parsetrees aus Sätzen generieren kann. Die Software teilt die zu analysierenden Artikel also in die einzelnen Sätze auf und analysiert mithilfe des Parsers deren einzelne Wörter. Mithilfe dieser Wörter und der zugehörigen Bedeutung, die das Modell ihnen zuschreibt, kann die Applikation im Anschluss zwei unterschiedliche Arten von Zitaten erkennen:

1. *Syntaktisches Zitat*
2. *Schwimmendes Zitat*

Diese Art Zitate zu unterscheiden erscheint auf den ersten Blick nicht intuitiv. So werden Zitate ja entweder in direkter, also mit Anführungs- und Schlusszeichen, oder in indirekter Rede verfasst. Der naheliegende Ansatz wäre also die Unterscheidung nach diesen zwei Mustern. Spronck und Nikitina haben aber herausgefunden, dass ein syntaktischer Ansatz zum Parsen (mithilfe der Satzstruktur) bessere Resultate liefert, als semantisches Parsing (anhand von semantischen Merkmalen wie Satzzeichen oder bestimmten Wörtern) [51]. Aufgrund dieser Erkenntnis orientieren sie sich am syntaktischen Ansatz.

Was sind also syntaktische und schwimmende Zitate? Syntaktische Zitate sind diejenigen Satzstrukturen, die alle Informationen beinhalten, die zu einem Zitat gehören:

1. Ein Zitat in der direkten oder indirekten Rede
2. Ein Subjekt, die Person, die zitiert wurde
3. Ein Zitat-einleitendes Verb

Ein Beispiel dafür ist „Der Bundesrat Berset versicherte, dass genügend Masken vorhanden seien.“ oder „Der Bundesrat Berset versicherte: ‚Es sind genügend Masken vorhanden.‘“. In diesem Beispiel ist das Zitat „dass genügend Masken vorhanden seien“ beziehungsweise

„Es sind genügend Masken vorhanden“. Das Subjekt ist in beiden Fällen „Der Bundesrat Berset“. Auch das einleitende Verb ist in diesen Fällen gleich, nämlich „versicherte“.

Schwimmende Zitate finden sich besonders in Texten, in denen eine Person wiederholt zitiert wird und das Subjekt nicht jedes Mal wiederholt wird. Um beim vorherigen Beispiel zu bleiben, können wir uns eine Weiterführung des Textes so vorstellen: „Der Bundesrat Berset versicherte, dass genügend Masken vorhanden seien. „Es sind genug für alle da.““. Das Zitat „Es sind genug für alle da.“ *schwimmt* im Text ohne selbst von einem Verb eingeleitet oder im selben Satz von einem Subjekt begleitet zu werden. Für uns Menschen ist damit implizit klar, dass das Zitat zu der zuvor zitierten Person gehört. Für einen Algorithmus sind solche impliziten Muster schwieriger zu erkennen und zu verbinden, weshalb solche schwimmenden Zitate anders behandelt werden müssen.

Personen identifizieren und deren Geschlecht bestimmen

Für ihre Auswertungen benötigen die Forscher:innen eine Sammlung von Personen, die mit dem Text verbunden sind und deren Geschlechterzugehörigkeit. Insbesondere interessieren sie sich für

1. die **Autorinnen und Autoren** der der Texte,
2. die zitierten Personen (Quellen)

aber auch generell für alle anderen Menschen, die im Text erwähnt werden.

Um Personen in den Artikeln zu finden, bedienen sie sich eines *NERs (Named Entity Recognition)*. Dieser spezialisierte Parser kann aus Texten Entitäten wie Personen, Orte und Firmen extrahieren. In diesem Fall werden nur die Personen weiterverarbeitet.

Das Resultat dieses Verfahrens ist eine Liste mit allen Wörtern und Satzteilen, die Personen repräsentieren. Diese Liste ist unbereinigt und wird mithilfe eines *Coreference Resolutions (Coreference Resolution)* in zusammengehörige Gruppen aufgeteilt. So wird der *NER* in einem Text über den Bundesrat Berset viele unterschiedliche Repräsentationen dieses Mannes finden, wie zum Beispiel „Bundesrat Berset“, „er“, „Alain Berset“ etc. Da diese Nennungen alle **dergleichen** Person gelten, ist es für die Auswertung relevant, diese zu clustern.

Zur Identifikation des Geschlechts verwendet die Applikation eine Web API, die entweder nur mithilfe des Vornamens, oder mit Vornamen und Nachnamen, das Geschlecht bestimmen kann. Einzelne Fälle, die die API falsch zuordnet, führt die Applikation in einer dedizierten Liste.

Evaluation

Zur Qualitätssicherung haben die Forschenden kontinuierlich Messungen zur Performance ihrer Auswertungen durchgeführt. Das ermöglichte es ihnen, den Effekt von implementierten Verbesserungen direkt messen zu können. Als Grundlage dazu dienten 14 repräsentative, manuell ausgewertete Artikel aus allen sieben verwendeten Nachrichtenportalen mit

einer Mindestlänge von 3000 Zeichen. Mithilfe dieses Datensatzes konnten sie die *Precision*, den *Recall* und den *F1 Score* für die Extraktion von Zitaten, Personen und Verben bestimmen.

Resultate

Die Resultate der Arbeit können einen klaren Gender Bias in der Verteilung in den untersuchten Onlineartikel aus sieben kanadischen Medienhäusern feststellen. So sind die Zitate in ihrem Datensatz in einem Durchschnitt über zwei Jahre (01.10.2018 - 30.09.2020) in nur 29% der Fälle Frauen zuzuordnen. Die verbleibenden 71% sind Zitate von Männern.

Der Unterschied in der Verteilung der Zitate zwischen den Geschlechtern variiert zwischen den Medienhäusern, wobei CBC News mit 33% der Zitate am besten und The Globe and Mail mit 23% am schlechtesten abschneiden.

Aus ihren Daten ist ersichtlich, dass Frauen während der Coronapandemie häufiger zitiert wurden als zuvor. Die Autor:innen weisen auf die Tatsache hin, dass Frauen in den Gesundheitsämtern häufiger vertreten sind als in anderen Bereichen der Arbeitswelt und dass eine ausgeglichene Repräsentation der Geschlechter in wichtigen Ämtern den *Gender Gap* zu verringern scheint.

Die für diese Auswertung gesammelten Daten boten auch die Möglichkeit, Untersuchungen zu den am häufigsten zitierten Menschen zu machen. So haben die Forscher:innen einen Ausschnitt der Top 15 zitierten Personen aus dem Datensatz genauer untersucht. Sie haben diese Kategorien wie Politik, öffentliche Gesundheit und Private zugewiesen. Die meistzitierte Frau in dieser Liste würde mit 2'239 Zitierungen gerade mal den 8 Platz in der Rangliste der Männer einnehmen. Interessanterweise sind alle 15 Vertreter der meistzitierten Männer im Sektor der Politik tätig. Auch bei den Frauen sind die meisten Politikerinnen. Drei sind aus dem Bereich der öffentlichen Gesundheit und eine aus der Privatwirtschaft. Dass es drei Vertreterinnen aus dem Gesundheitssektor in die Top 15 meist zitierten Frauen im Datensatz schaffen, weist ebenfalls darauf hin, dass eine ausgeglichene Frauenquote in für die Öffentlichkeit interessanten Bereichen den Geschlechtergraben verkleinern könnte.

Eine weitere Analyse der Wissenschaftler:innen zeigt die Aufteilung der Zitate nach Geschlecht und Tätigkeitsgebiet der Quelle. Die tabellarische Auswertung zeigt, dass die meistzitierten Männer Politiker (73.8%), Sportler (7.7%) und Regierungsangestellte sind. Bei den Frauen finden sich die meisten Zitate in den Tätigkeitsgebieten Politik (52.6%), Gesundheit (16.7%) und Exekutive (8.3%) (Regierungsangestellte). Auch hier fällt wieder auf, wie massiv weniger Frauen zitiert werden. So werden Politiker 103'378 und Politikerinnen 29'007 Mal zitiert, Sportler 10'723 und Sportlerinnen 1'415 Mal und Regierungsangestellte Männer 9'175 und Frauen 4'583 Mal. Nur in den Bereichen öffentliche Gesundheit, NGO's und Wissenschaft werden Frauen leicht häufiger zitiert.

Eine interessante Nebenerkenntnis dieser Auswertung ist die Tatsache, dass insgesamt weniger Zitate von Frauen in Online-Nachrichtenportalen erscheinen als von Männern, jedoch absolut gesehen mehr individuelle Frauen zitiert werden. Das bedeutet, dass die Zitate pro Mann deutlich höher ist als die Zitate pro Frau.

Sie erkennen das Muster, dass Männern mehr Raum gegeben wird als Frauen und einige wenige Kategorien den grössten Teil der Berichterstattung dominieren. Die Forscher:innen spekulieren aufgrund dieser Erkenntnis, dass die Daten einer Pareto Verteilung unterliegen. Also dass das grösste Hindernis mehr von Frauen zu hören die Bevorzugung derer ist, die bereits sehr viel Raum einnehmen.

In einer weiteren Auswertung haben Asr u. a. den Zusammenhang zwischen dem Geschlecht des Autors oder der Autorin und der Anzahl und Verteilung der Zitate von Männern und Frauen untersucht. Sie wollten damit herausfinden, ob Frauen häufiger weibliche Quellen zitieren als Männer. In der Tat scheinen Journalistinnen häufiger Frauen zu zitieren als Journalisten. Während Männer im Schnitt nur in einem Viertel (25%) ihrer Zitate auf Frauen zurückgreifen, sind es bei den Frauen immerhin 34%. Daraus schliessen die Autor:innen, dass ein Teil der Lösung zum Gender Gap in der Erhöhung der Quote von Reporterinnen besteht. Nicht nur, weil sie in selbst geschriebenen Artikeln häufiger Frauen zitieren, sondern weil sie auch in gemischten Teams einen positiven Einfluss auf die Wahl der Quellen ausüben, wie die Auswertung zeigt.

Die Applikation, die auf täglicher Basis die für diese Auswertungen notwendigen Daten aggregiert, ist nach wie vor im Einsatz und wird - wenn auch sporadisch, doch immer wieder weiterentwickelt. Die aktuellsten Auswertungen sind in interaktiver Form auf GenderGapTracker.InformedOpinions.org zu finden. Der Code dazu ist öffentlich auf GitHub einsehbar.

Weiterentwicklung des Gender Gap Trackers

Nach der Fertigstellung des Papers im Jahr 2021, wurde die Applikation um die Möglichkeit, auch französische Artikel zu analysieren, erweitert. Damit analysiert der *GGT* zum aktuellen Stand sieben englisch- und sechs französischsprachige Nachrichtenportale.

Da die Release Notes auf GitHub erst seit September 2022 geführt werden, bleiben wohl viele neue Features, die in der Arbeit nicht beschrieben sind, verborgen. Die Entwickler:innen sind jedoch fleissig am **Refactor** und in voneinander unabhängige Teile aufzusplitten. Wir freuen uns auf zukünftige Innovationen des *GGT*.

2.2.3 Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles

In ihrer Arbeit „Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles“ [15] erforschten Dacon und Liu den impliziten und expliziten Bias in Abstracts von Nachrichtenartikeln aus zwei Datensets (*Microsoft News Dataset (MIND)*¹, *News Category Dataset[36] (NCD)* ²) auf Bias zwischen den Geschlechtern. Im **speziellen** untersuchten sie dabei den *Gender Bias in der Verteilung*, den *Gender Bias im Inhalt* und den *Gender Bias in der Wortwahl*.

Das *MIND* Dataset besteht aus ungefähr 160'000 englischen Nachrichtenartikeln mit Text, Titel, Abstract und Kategorien. Das Datenset enthält auch anonymisierte User Logs,

¹<https://msnews.github.io/>

²<https://www.kaggle.com/rmisra/news-category-dataset>

die für diese Auswertung jedoch irrelevant sind. *NCD* ist ein Kaggle Dataset bestehend aus etwa 210'000 englischsprachigen Schlagzeilen mit dazugehörigen Abstracts.

Im ersten Experiment untersuchten die Forschenden den *Gender Bias in der Verteilung*. Dazu teilten sie die Abstracts beider Datensets in drei Kategorien auf: Männlich (M), Weiblich (F) und Neutral (N) und verglichen die Anzahl Abstracts in den jeweiligen Kategorien anschliessend tabellarisch. Um die Aufteilung vorzunehmen, bedienten sie sich eines Datensets von 465 geschlechterbezogenen *Possessive Nouns* und 357 geschlechterspezifischen *Attribute Words*. Mit diesem Experiment konnten sie im *MIND* Datenset einen klaren *Gender Bias in der Verteilung* von 53.9% zugunsten der Männer nachweisen. Das andere Datenset schnitt mit einem *Gender Bias in der Verteilung* von 14.5% zugunsten der Männer nur deshalb besser ab in der absoluten Verteilung, weil Frauen in den Kategorien „Style & Beauty“, „Parenting“ und „Entertainment“ stark übervertreten sind, was auf einen Bias in der Art der Präsentation von Frauen hindeutet, wie auch das zweite Experiment zeigt.

Dieses soll den *Gender Bias im Inhalt* messen, in dem es die Häufigkeit vergleicht, mit der Frauen und Männer in den Kontexten Familie und Karriere vorkommen. Dazu setzen sie zwei weitere *Attribute Words* Sets ein, mit denen sie die geschlechterbezogenen Abstracts den gefragten Kontexten zuweisen können. Das Resultat zeigt, dass Frauen im *MIND* Datensatz 16.21% weniger häufig mit Karriere assoziiert werden, als Männer. Im *NCD* Datensatz ist der Unterschied mit 8.51% halb so gross. Im Gegenzug werden sie häufiger im Kontext von Familie beschrieben. Im *MIND* Datensatz 4.1% häufiger und im *NCD* Datensatz 3.19% häufiger. Diese Erkenntnis wird auch durch das dritte Experiment gestützt.

In diesem Experiment untersuchen die Forschenden basierend auf Sentiment Analysis und *Centering Resonance Analysis (CRA)* den *Gender Bias in der Wortwahl*. Mithilfe des Tools *VADER* extrahierten sie aus den nach Geschlecht aufgeschlüsselten Abstracts diejenigen mit positiver Benotung des Algorithmus. Anschliessend liessen sie den *CRA* die 20 wichtigsten Wörter in den jeweiligen Sammlungen als Graph darstellen. In diesem ist die Wichtigkeit des jeweiligen Wortes als Grösse der Knoten ersichtlich. Wörter, die in denselben Kontexten vorkommen sind durch Kanten verbunden. Das Resultat deutet auf einen starken *Gender Bias in der Wortwahl* hin. Während die wichtigsten 20 Wörter bezogen auf Männer „president“, „manager“ und „mayor“ beinhalten, beginnt die Top 20 Liste bei den Frauen mit „wife“, „mother“ und „beloved“. Dieses Experiment stützt also auch die Erkenntnis aus dem zweiten Experiment, dass Frauen in den Medien besonders im Kontext der Familie, ihres Ehemannes oder ihres Körpers porträtiert werden und nicht wie die Männer im Zusammenhang mit Politik, Wirtschaft oder Karriere.

Um die oben beschriebenen Experimente durchzuführen haben Dacon und Liu einiges an Datenaufbereitung betrieben. Damit künftige Forschende einen einfacheren Zugang zu solchen Daten haben, veröffentlichten **Sie** diese im Rahmen der Arbeit für die freie Verwendung in der Forschung³.

³https://github.com/daconjam/Detecting_Gender_Bias

2.2.4 DIANES: A DEI Audit Toolkit for News Sources

Die Autor:innen der Arbeit „DIANES: A DEI Audit Toolkit for News Sources“ [49] haben ein Audit Toolkit entwickelt, um Journalist:innen und Reporter:innen von Nachrichtenportalen dabei zu helfen, den Anteil der Quellenvielfalt (z. B. Geschlecht und Rasse) für Zitate in ihren Artikeln zu verbessern. Laut den Autor:innen der Arbeit wird in professionellen Nachrichtenmedien immer betont wie wichtig die Vielfalt der Perspektiven ist, aber in der Praxis wird dem nur wenig Beachtung geschenkt. Daher sind die Nachrichtenmedien unter ethische Kritik geraten im Rahmen der neuen Normen für Vielfalt, Gleichberechtigung und Integration, im Englischen *Diversity, Equity and Inclusion (DEI)*. Das Toolkit soll die Medienschaffenden dabei unterstützen, diesen Normen mehr Beachtung zu schenken. *Diversity Auditor for News Sources (DIANES)* ist ein *DEI* Toolkit, das eine NLP-Pipeline im Backend zur Extraktion von Zitaten, Sprechern, Titeln und Organisationen aus Nachrichtenartikeln verwendet und diese in Echtzeit verarbeitet.

Im Frontend unterstützt *DIANES* eine Vielzahl von Benutzeranfragen zum *DEI* Audit auf der Grundlage der im Backend extrahierten Informationen. Auf dem Artikel Dashboard können Medienschaffende ihren Artikeltext als Entwurf eingeben und dann die entsprechenden Informationen zur Quellenvielfalt in diesem Artikel abrufen. Das Dashboard der Quellenvielfalt ist dazu da, die *DEI* Daten für einen ausgewählten Zeitraum von allen Artikeln zu visualisieren.

Das Application Programming Interface (API) implementiert die Möglichkeit, alle Zitate, Namen, Geschlechter sowie Titel der Sprecher:innen zu extrahieren, die in **Artikel Texten** vorkommen, welche der API gesendet werden.

Das Backend Modul nimmt den Nachrichtenartikel als Input und extrahiert die Zitate **und** Sprecher:innen mit ihren Titeln und prognostiziert dann Geschlecht und Rasse/Ethnie dieser zitierten Person. Das Modul ist in folgende vier Stages unterteilt, von welchen die ersten zwei Stages die Stanford CoreNLP Bibliothek ⁴ nutzen.

Text Processing

In der „Text Processing“ Stage werden zunächst unnötige Informationen aus dem Text gefiltert wie *XML* Tags. Nach dem Filtern wird der Text tokenisiert und im nächsten Schritt wird *POS* Tagging angewendet um z. B. zu bestimmen, ob die einzelnen Wörter (Tokens) Verben oder Substantive sind. Danach wird eine *Lemmatisierung* durchgeführt, um jedes Wort auf seine Grundform zu reduzieren.

NLP Analysis

Die „NLP Analysis“ Stage enthält die folgenden spezifischen Komponenten. Die Satzspaltung teilt den Text in Sätze auf. Dies wird anhand des CoreNLP Regelwerks gemacht. Mit *NER* werden Personen, Titel und Organisationen im Text bestimmt. Dependency Parsing analysiert die grammatikalischen Beziehungen zwischen Wörtern in einem Satz und extrahiert textuelle Beziehungen auf Basis der Abhängigkeiten. Die *Coreference Resolution (Core-*

⁴<https://stanfordnlp.github.io/CoreNLP/>

ference Resolution) findet Erwähnungen der gleichen Entität in einem Text, z. B. wenn sich „Anne“ und „sie“ auf dieselbe Person beziehen. *Knowledge Base Population* wird verwendet um Titel und Organisationen einer Person zu finden, wenn diese im Text vorkommen.

Annotations

In dieser **Stage** werden Annotationen erstellt für Zitate, Sprecher:innen und ihre Titel und Organisationen auf der Grundlage der vorherigen *NLP* Analyse. Wenn bei einem Zitat ein Anführungszeichen oder Schlusszeichen fehlt, hat CoreNLP den Text bis zum nächsten Anführungszeichen/Schlusszeichen des nächsten Zitates als Zitat Inhalt deklariert, was die Genauigkeit der Zitat-Auflösung verringern kann. Um dieses Problem zu umgehen, arbeiteten die Entwickler:innen mit der Anzahl Wörter in vorgängigen Zitaten. Zitate mit über 100 Wörter oder unter fünf Wörter wurden für die Auswertung ausgeschlossen. Die Entwickler:innen machten die Erfahrung, dass CoreNLP Probleme hatte und oft die falsche Person zum Zitat zuwies.

Representation Recognition:

Diese Stage dient dazu Geschlecht und Rasse/Ethnie der zitierten Person zu definieren. Diese zwei Repräsentationsattribute von *DEI* sind die entscheidenden Informationen zum Verständnis der Repräsentativität der Nachrichtenbeschaffung von *DIANES*. Das Geschlecht wurde mit der Gender API ⁵ bestimmt. Es gibt viel weniger existierende Arbeiten zur Vorhersage der Rasse/Ethnizität einer Person mit einem Namen, und die Aufgabe ist anspruchsvoller als Geschlechtserkennung. Darum wurde zur Vorhersage der Rasse/Ethnizität einer Person ein eigenes *Machine Learning (ML)* Model implementiert und trainiert, welches als Input den Namen der Person braucht. Der Datensatz, der zum Trainieren des Rassendetektors verwendet wurde, stammt von dem United States Census Bureau aus den Jahren 2000 und 2010. Er enthält insgesamt 151'670 eindeutige Namen in sechs Kategorien (White, African, American, American Indian, Alaska Native, Asian, Native Hawaiian und Other Pacific Islander).

DIANES wurde zum Zeitpunkt der Veröffentlichung der Arbeit von mehreren Redaktionen in der Praxis getestet.

2.2.5 The gender gap in commenting: Women are less likely than men to comment on (men's) published research

Die Studie „The gender gap in commenting: Women are less likely than men to comment on (men's) published research“ [56] behandelt den Gender Gap im Kommentieren von veröffentlichten Forschungsartikeln. Die Autor:innen gehen der Frage nach, ob auch das Kommentieren von veröffentlichten Forschungsartikeln geschlechtsspezifisch ist. In der Studie haben die Autor:innen anhand von 1'350 Kommentaren zu 1'236 Forschungsartikeln, welche in den Fachzeitschriften - *Proceedings of the National Academy of Sciences (PNAS)*

⁵<https://gender-api.com/>

(PNAS) und Science - veröffentlicht wurden festgestellt, dass es eine geschlechtsspezifische Diskrepanz bei den **Verfassers**:innen von Kommentaren gibt.

Die Autor:innen haben sich gefragt, wieso Frauen seltener Kommentieren als Männer. Sie haben die Theorie aufgestellt, dass es beim Kommentieren um Herausforderung, Risiko und Belohnung geht.

Allgemeine Risikoaversion

Wenn eine Person sich entscheidet, einen für alle sichtbaren Kommentar unter einen veröffentlichten Forschungsartikel zu schreiben, geht sie auch das Risiko ein, den eigenen Ruf zu schädigen. Das Kommentieren setzt die Bereitschaft voraus, die Überlegenheit der eigenen wissenschaftlichen Erkenntnisse über die eines anderen zu behaupten. Die Autor:innen treffen die Annahme, dass Wissenschaftler:innen warten mit der Abgabe von Kommentaren, bis sich ihre Karriere besser etabliert hat und da Frauen meistens auf tieferen Karriere Stufen sind als Männer, weniger Artikel kommentieren. In diesem Fall wäre der Unterschied bei der Verteilung der Geschlechter im Kommentieren von Forschungsartikeln ein Pipeline-Effekt.

Geschlechtsspezifische Risikoaversion

Möglicherweise könnte das Geschlecht die relative Abwägung der Risiken und der Chancen durch einen Kommentar beeinflussen. Aus jahrzehntelanger Forschung ist bekannt, dass Männer Risikofreudiger sind als Frauen [9, 14, 45, 46, 23]. Daraus ist zu schliessen, dass sich Frauen mehr Sorgen über die negativen Folgen machen, wenn sie einen Forschungsartikel kommentieren wollen und dies darum seltener tun. Die Autor:innen erwarten auch, dass die geschlechtsspezifischen Unterschiede bei der Kommentierung schwächer ausfallen, wenn der Zielartikel von einer jüngeren Person verfasst wurde und somit das Risiko, sich gegen sie zu stellen, geringer ist.

Geschlechtsspezifisches Risiko

Denkbar ist auch, dass die Folgen von Kommentaren geschlechtsspezifisch sind. Frauen und Männer werden für dasselbe Verhalten unterschiedlich bestraft und belohnt. Von Frauen wird erwartet, dass sie gemeinschaftlich und nicht durchsetzungsfähig sind [8, 18]. Bei Männer hingegen hat die Gesellschaft die Erwartung, dass sie durchsetzungsfähig und wettbewerbsorientiert sind. Ein Verstoß gegen diese geschlechtsspezifischen Stereotypen führt in der Regel zu Missbilligung [45, 34, 37, 47]. Wenn Frauen als durchsetzungsfähig wahrgenommen werden und in männerdominierten Bereichen Erfolg haben, sinkt ihre „Beliebtheit“ in der Regel mit negativen sozialen und wirtschaftlichen Folgen [45, 34, 37, 47]. Während Kommentare im Allgemeinen für Frauen unverhältnismässig riskant sein könnten, dürfte dies noch ausgeprägter sein, wenn es sich bei der Zielperson um einen männlichen Wissenschaftler handelt, da dies die vermeintliche Überlegenheit von Männern infrage stellt. Dies bedeutet, dass Frauen im Vergleich zu männlichen Kommentatoren weniger wahrscheinlich Artikel angreifen, die von Männern verfasst wurden.

Geschlechtsspezifische Fürsorge unterdrückt Herausforderungen

Frauen gelten oft als das fürsorglichere Geschlecht [33]. Das Kommentieren ist nicht unbedingt eine angenehme Praxis, da es die Arbeit eines anderen infrage stellt. Es ist durchaus denkbar, dass potenzielle Verfasser:innen von Kommentaren befürchten, dass ihr Kommentar negative Folgen für die betroffenen **n** haben könnte. Dies könnte die Quote der Kommentare von Frauen senken.

Die Resultate der Untersuchung zeigen, dass unter den jüngeren, gefährdeten Wissenschaftler:innen der Anteil der Autorinnen sowohl bei Kommentaren als auch bei Artikeln höher ist, wahrscheinlich weil Frauen in den unteren akademischen Rängen stärker vertreten sind.

Wenn eine grössere Risikoaversion der Grund für die geringere Rate an Kommentaren von Frauen ist, erwarten die Forscher:innen auch eine geringere geschlechtsspezifische Diskrepanz bei der Verfasserschaft von Kommentaren, wenn der Erstautor relativ machtlos ist, da dies die Risiken beim Kommentieren im Allgemeinen weniger auffällig machen sollte. Umgekehrt, wenn Frauen insgesamt weniger kommentieren, weil sie sich mehr Sorgen machen, die Karriere anderer zu schädigen (geschlechtsspezifische Fürsorge drückt die Herausforderung), wird der geschlechtsspezifische Unterschied bei der Kommentierung zugunsten der Männer grösser sein, wenn die Zielperson stärker gefährdet ist. Tatsächlich sind die geschlechtsspezifischen Unterschiede in der Wahrscheinlichkeit, auf solche Artikel zu zielen, gering und statistisch nicht signifikant. Die Ergebnisse deuten also darauf hin, dass die unterschiedliche Risikoorientierung von Frauen an sich (sei es für sie selbst oder für die Zielpersonen) wahrscheinlich nicht der Grund für ihre geringere Rate an Kommentaren ist.

Während sowohl Männer als auch Frauen eher auf Artikel abzielen, die zuerst von Männern verfasst wurden, ist dies bei Männern deutlich ausgeprägter. Da Frauen unter den Autoren von *PNAS*- und Science-Artikeln unterrepräsentiert sind, würde dies an sich schon die Gesamtzahl der Kommentare von Frauen im Vergleich zu Männern senken. Es lässt sich nicht feststellen, ob diese Dynamik darauf zurückzuführen ist, dass Frauen andere Frauen als Zielpersonen bevorzugen (wie das Argument des geschlechtsspezifischen Risikos behauptet) oder dass Männer andere Männer überproportional bevorzugen (oder eine Kombination aus beidem). Es ist auch möglich, dass Frauen weniger als Männer auf Artikel abzielen, die von Männern geschrieben wurden, und zwar nicht aufgrund unterschiedlicher Konsequenzen, sondern weil ihr Fachwissen in Bezug auf die Arbeit von Frauen relevanter ist. Da Arbeiten **n** welche von Frauen verfasst wurden, auch ein wenig die Sichtweisen einer Frau repräsentieren.

Letztendlich konnte mit den Daten nicht direkt überprüft werden, ob Frauen mehr negative Konsequenzen erleiden, wenn sie den aktuellen Stand infrage stellen. Angesichts der aufgedeckten Muster und der mangelnden Unterstützung für die anderen Argumente bleibt dies jedoch eine plausible Erklärung für die niedrigere Rate an Kommentaren von Frauen.

Die Studie kommt zu dem Ergebnis, dass sich Frauen seltener an akademischen Kommentaren beteiligen, eine Diskrepanz, die die geschlechtsspezifischen Unterschiede bei der Veröffentlichung von Artikeln übersteigt. Die Diskrepanz kann nicht durch Schwan-

kungen im fachspezifischen Geschlechterverhältnis erklärt werden. Auch die Überrepräsentation von Frauen in den für Karriererisiken empfindlichsten Positionen oder eine größere allgemeine Sensibilität für Risiken für andere können die Diskrepanz nicht erklären. Frauen richten auch einen geringeren Anteil ihrer Kommentare auf die Forschung von Männern als Männer. Zusammengefasst stimmen die beiden Ergebnisse am ehesten mit dem Argument überein, dass die geringere Anzahl von Kommentaren von Frauen auf die geschlechtsspezifischen Kosten zurückzuführen ist, die mit der Infragestellung massgeblicher Forschung verbunden sind, insbesondere wenn es um die Infragestellung traditioneller Statushierarchien geht. Dieses Argument kann aber mit den Daten nur indirekt überprüft werden.

Diese geschlechtsspezifischen Muster bei akademischen Kommentaren könnten den wissenschaftlichen Austausch zwischen Männern und Frauen behindern und Frauen innerhalb der wissenschaftlichen Gemeinschaft weiter marginalisieren. Wenn Expertinnen ausgeschlossen werden, fehlt es der akademischen Gemeinschaft als Ganzes an frischen Ideen und vielfältigen Meinungen.

2.2.6 Using citation network analysis to enhance scholarship in psychological science: A case study of the human aggression literature

In dem Forschungsartikel „Using citation network analysis to enhance scholarship in psychological science: A case study of the human aggression literature“ untersuchten Iancarelli u. a. mithilfe einer graphbasierten Analyse die Zusammensetzung der Forschungsgemeinschaft zum Thema „Aggression“ [24]. Neben der Untersuchung zu den einflussreichsten Gemeinschaften auf diesem Feld wollten die Forscher:innen ausserdem herausfinden, ob ihre Netzwerkanalyse dazu beitragen kann, den *Gender Bias* beim Zitieren von wissenschaftlichen Arbeiten zu reduzieren.

Der Grund für diesen Bias, so mutmassen die Autor:innen des Artikels, liegt in der Art wie die Relevanz von wissenschaftlichen Arbeiten beurteilt wird. So werden aktuell diejenigen Papers als am relevantesten eingestuft, die am meisten zitiert wurden. Dieser Fokus auf der Anzahl Zitierungen begünstigt vermutlich einen Bias zum Vorteil von etablierten Forschungsgemeinschaften und führt zu einer Unterrepräsentation von anderen wichtigen Forschungsgruppen, so Iancarelli u. a. Dass dieser Ansatz im Falle von Publikationen im Bereich der Astronomie *Gender Bias* begünstigt, haben Caplar, Tacchella und Birrer bewiesen [10].

In ihrem Experiment vergleichen Iancarelli u. a. den Frauenanteil unter den Autor:innen aus den Top 75 Artikeln, ausgewählt nach drei Methodiken.

1. Anzahl Zitierungen
2. Kompositrangierung
3. Kompositrangierung mit Einbezug der Gemeinschaften

Die erste Methodik, „Anzahl Zitierungen“, entspricht dem Status Quo. Also, dass die wichtigsten Arbeiten anhand der absoluten Anzahl der Zitierungen identifiziert werden.

Die zweite Methodik nennen die Forschenden „composite score“. Auf Deutsch wäre das eine „Kompositrangierung“. Diese besteht aus sechs Teilbewertungen (Komposite) zur Messung der Zentralität einer Arbeit im gesamten Zitatnetzwerk. Der letzte Ansatz ist eine Erweiterung der Kompositrangierung und teilt allen Gemeinschaften im Graphen gleich viele Ranglistenplätze zu. Diese werden dann anhand der „Kompositrangierung“ sortiert.

Die Auswertung betrachtet die 75 bestplatzierten Arbeiten und vergleicht den Anteil der Autorinnen untereinander. Während dieser gemessen an der traditionellen Methode gerade mal 16% (12) beträgt, sind es bei der „Kompositrangierung“ bereits 20% (15) und unter Berücksichtigung der Gemeinschaften gleich 32% (24). Die „Kompositrangierung mit Einbezug der Gemeinschaften“ hat im verwendeten Datensatz also zu einer Verdoppelung der Quote von Autorinnen geführt. Iancarelli u. a. bemerken, dass diese Art der Relevanzbewertung nicht nur dabei helfen kann, unterrepräsentierte Gemeinschaften in der Forschung besser sichtbar zu machen, sondern auch den *Gender Gap* zu reduzieren.

Leider bietet die Arbeit keine Auskunft zum Anteil der Autorinnen über das gesamte Datensatz, was das Einordnen ihrer Zahlen erschwert. So ist es nicht möglich zu beurteilen, ob der Anteil der Arbeiten, die von Frauen verfasst wurden, insgesamt mehr oder weniger als 32% ausmacht und wie gross oder klein der *Gender Gap* damit noch immer ist.

2.2.7 A Large-Scale Test of Gender Bias in the Media

Die Autor:innen der Arbeit „A Large-Scale Test of Gender Bias in the Media“ [50] untersuchten den Gender Bias in einem Datensatz von Lydia news analysis system [31]. Dieser enthält Zeitgestempelte Nennungen von Personen aus mehr als 2000 gescannten oder digitalen Zeitungen, Magazinen und online Nachrichtenartikeln aus den Jahren 2004 bis 2009. Die Quellen stammen vorwiegend aus dem US **Amer**ikanischen Raum. Sie stellen fest, dass Frauen im Vergleich zu Männern in ähnlichen Positionen und aus denselben Berufsschichten ein grösseres öffentliches Interesse geniessen, aber weniger Medienberichterstattung erhalten.

In über 10'000 untersuchten Paaren von Männern und Frauen gleichen Alters, die im selben Sektor tätig sind und gleiches öffentliches Interesse geniessen, haben die Forschenden einen Unterschied von 46% in der Anzahl jährlicher Nennungen zugunsten der Männer gemessen. So wurden die Frauen aus den gebildeten Paaren 300 Mal erwähnt und die Männer 826 Mal.

Die Forscher:innen argumentieren mit den zwei folgenden Erklärungen wieso die Berichterstattung über Männer unverhältnismässig ist. Die erste Erklärung könnte die berufliche Ungleichheit zwischen Frauen und Männer sein. Studien haben gezeigt, dass Männer häufiger vertreten sind in Machtpositionen [19, 20]. Besonders in Bereichen, denen die Medien grössere Aufmerksamkeit schenken. Daraus folgt natürlich, dass die Mainstream-Medien mehr über Männer als über Frauen berichten.

Als zweite Erklärung erkennen die Forscher:innen, dass die Ungleichheit in der Berichterstattung ein unverhältnismässig grosses öffentliches Interesse an Männern widerspiegeln könnte. Dies kann darauf zurückzuführen sein, dass das Publikum Nachrichten über Männer bevorzugt oder dass Männer ein grösseres Interesse wecken, weil sie Verhaltensweisen an den Tag legen, die die Öffentlichkeit für berichtenswert hält.

Die Forscher:innen gehen davon aus, dass Ungleichheiten in der Medienberichterstattung über Männer und Frauen in erster Linie auf eine der folgenden drei Hauptkategorien von Erklärungen zurückzuführen sind:

- ▶ strukturelle Ungleichheiten in der Arbeitswelt und in hochrangigen Positionen
- ▶ Ungleichheiten in Bezug auf öffentliches Interesse und Nachrichtenwert bei gleichwertigen Positionen
- ▶ Medienspezifische Faktoren: Voreingenommenheit und Diskriminierung in der Medienberichterstattung

Strukturelle Ungleichheiten in der Arbeitswelt und in hochrangigen Positionen

Frauen haben oft einen eingeschränkten Zugang zu hochrangigen Berufen und Positionen und werden durch Stereotypen und Diskriminierung benachteiligt [21]. Frauen stoßen häufig auf eine "gläserne Decke", die ihr Fortkommen in der Arbeitswelt behindert [13, 21, 26]. Obwohl diese Ungleichheiten die Berichterstattung von Männern begünstigen, bleibt unklar, ob sie allein für den Unterschied in der Medienpräsenz von Männern und Frauen verantwortlich sind. Die Forscher:innen denken **an, Frauen und Männer** in ähnlichen Positionen oder mit ähnlichen Errungenschaften zu vergleichen, um zu bestimmen, ob strukturelle Ungleichheiten allein für die Unterschiede in der Berichterstattung verantwortlich sind.

Ungleichheiten in Bezug auf öffentliches Interesse und Nachrichtenwert bei gleichwertigen Positionen

Die Forscher:innen weisen darauf hin, dass selbst ein systematischer Vergleich von Männern und Frauen, die gleichwertige Positionen in verschiedenen sozialen und beruflichen Teilbereichen erreicht haben, nicht ausreicht. Denn auch das öffentliche Interesse muss berücksichtigt werden. Die Medien sind mit kommerziellen Interessen und Zwängen konfrontiert und müssen Leser:innen und Zuschauer:innen Quoten berücksichtigen. Es stellt sich die Frage, ob es systematische Unterschiede im öffentlichen Interesse zwischen Männern und Frauen in vergleichbaren Positionen gibt. Einerseits könnten Frauen, die hohe soziale und berufliche Positionen erreicht haben, im Vergleich zu ihren männlichen Kollegen mehr Interesse auf sich ziehen, da sie aufgrund von Barrieren und Diskriminierung besonders talentiert und erfolgreich sein müssen [21]. Andererseits deuten feministische und geschlechtsspezifische Theorien darauf hin, dass Frauen tatsächlich weniger öffentliches Interesse auf sich ziehen könnten, selbst wenn sie führende soziale und berufliche Positionen erreichen [30, 38, 40]. Die Gründe dafür liegen in fest verwurzelten kulturellen Überzeugungen über Geschlechterstatus und die Arten von Personen, die besonderen Respekt verdienen [38].

Medienspezifische Faktoren: Voreingenommenheit und Diskriminierung in der Medienberichterstattung

Ähnlich wie bei geschlechtsspezifischen Lohnunterschieden können Unterschiede, die nicht auf strukturelle Ungleichheiten zurückzuführen sind, als geschlechtsspezifische Verzerrungen betrachtet werden [11, 16]. Die Verzerrung kann Frauen sowohl begünstigen als auch benachteiligen. Sie kann durch Statusvorurteile entstehen, die die Bereitschaft beeinflussen, Frauen Aufmerksamkeit zu schenken und ihre Leistungen positiv zu bewerten [38, 39, 40]. Die täglichen Entscheidungen von Journalist:innen, Redakteur:innen und Verleger:innen, die in einem männlich geprägten kulturellen Umfeld arbeiten, tragen zur Verschärfung der Ungleichgewichte in der Berichterstattung bei und verzerren **n die sie** zugunsten von Männern. Die historische Dominanz von Männern in redaktionellen Positionen hat eine Machtstruktur geschaffen, die eine männlich geprägte Berichterstattung und redaktionelle Normen diktiert [41, 42].

Schlussfolgerung und Diskussion

Die Forscher:innen untersuchten mehr als 20'000 namhafte Männer und Frauen aus verschiedenen sozialen und beruflichen Bereichen und verglichen die Medienberichterstattung und das öffentliche Interesse. Die Studie ergab, dass Männer trotz des grösseren öffentlichen Interesses an Frauen mehr als doppelt so oft erwähnt wurde. Nach einem Abgleich von Männern und Frauen nach Erfahrung, struktureller Position und öffentlichem Interesse, haben die Zeitungen den Männern in fast allen Bereichen einen deutlich grösseren Umfang an Berichterstattung zugestanden. Auch wenn Frauen bedeutende Leistungen vollbringen und zweifellos im öffentlichen Fokus stehen, erhalten sie normalerweise weniger Medienaufmerksamkeit als Männer mit ähnlichem öffentlichem Interesse. Die Studie zeigt, dass theoretische Vorschläge, die entweder strukturelle Ungleichheiten oder individuelle Unterschiede in Motivation, Anstrengung, Talent oder Charisma hervorheben, unzureichend sind, wenn es darum geht, die geschlechtsspezifischen Unterschiede in der Berichterstattung zu erklären. Die Forschung zu geschlechtsspezifischen Ungleichheiten auf dem Arbeitsmarkt hat gezeigt, dass Frauen beim Berufseinstieg und beim Aufstieg innerhalb ihrer Berufe aufgrund von strukturellen Hindernissen benachteiligt sind und somit in der Öffentlichkeit weniger präsent sind [13, 21, 26]. Die vorliegende Studie signalisiert jedoch, dass diese strukturellen Ungleichheiten die Lücke nicht vollständig erklären können, die selbst beim Vergleich von Männern und Frauen, die ähnliche Positionen und Leistungen erreicht haben, erheblich bleibt. Es gibt Hinweise darauf, dass Frauen als weniger wertvoll für die öffentliche Aufmerksamkeit angesehen werden. Dies ist sowohl auf kulturelle Überzeugungen zurückzuführen, die Frauen und ihre Handlungen abwerten, als auch darauf, dass Frauen eine kürzere Karrieredauer in prominenten Ämtern haben, ihnen weniger wichtige Positionen zugewiesen werden und sie sich seltener selbst vermarkten [38, 40]. Die Studie leistet somit einen wichtigen Beitrag zur Massenkommunikationsliteratur über geschlechtsspezifische Unterschiede in der Medienberichterstattung und zur umfassenderen Literatur über kulturelle und berufliche Ungleichheiten zwischen den Geschlechtern.

2.2.8 Cancer in the news: Bias and quality in media reporting of cancer research

Die Studie „Cancer in the news: Bias and quality in media reporting of cancer research“ von Amberg und Saunders [3] untersucht die Medienberichterstattung von ausgewählten grossen Medienhäusern im Angelsächsischen Raum. Sie zeigt die Verteilung der Berichterstattung über die Art der Studien (Primär- oder Sekundärliteratur) auf, untersucht die Qualität der entsprechenden Artikel und analysiert den Bias in Bezug auf das Geschlecht und die Nationalität.

Als Datengrundlage verwenden die Forschenden je 20 Nachrichtenartikel aus den Online-Nachrichtenportalen von *The Guardian* (Edition aus dem Vereinigten Königreich), *New York Times* (NYT), *Sydney Morning Herald* (SMH) und *Australian Broadcast Corporation* (ABC). Die Artikel stammen aus dem Zeitraum vom März bis September 2017 und berichten über Studienergebnisse zur Krebsforschung. Dabei wurden nur Originalartikel berücksichtigt, welche die Ergebnisse jeweils einer Studie zusammenfassen. Für die Analyse zum Gender Bias haben die Forschenden die Anzahl Autorinnen und Autoren und die Anzahl männlicher und weiblicher Quellen verglichen.

Die Untersuchung über die 80 Nachrichtenartikel hat ergeben, dass Männer mit 60% (67/112) bei den „senior authors“ signifikant übervertreten sind. Dieser Trend ist konstant über alle untersuchten Medienhäuser. Die Ergebnisse zu den verwendeten Quellen zeigen zudem, dass der Bias bezüglich zitierten Fachpersonen mit 68% (100/148) Männeranteil noch stärker ausgeprägt ist. Auch diese Erkenntnis lässt sich über die untersuchten Nachrichtenportale generalisieren, mit der Ausnahme von ABC, bei der eine gleiche Anzahl Expertenmeinungen von Männern und Frauen gezählt wurde.

Amberg und Saunders haben einen erheblicher *Gender Bias* in der Berichterstattung über die Krebsforschung nachgewiesen. Sowohl bei den Hauptautor:innen wie auch bei den Quellen sind Männer deutlich übervertreten. Sie weisen darauf hin, dass der *Mathilda Effekt*, der die systematisch kleinere Anerkennung von Wissenschaftlerinnen beschreibt, auch bei der Kommunikation von wissenschaftlichen Erkenntnissen gilt [44, 27]. Zusätzlich ist die Unterrepräsentation von Frauen in den Medien im Allgemeinen gut erforscht [25, 43], so ihr Hinweis. Der *Gender Bias* in all seinen Facetten habe das Potenzial, qualitativ hochstehende Berichterstattung von Forschung durch die Limitierung von Meinungsdiversität zu kompromittieren, argumentieren sie. Zusätzlich scheint es wahrscheinlich, dass dieser die existierenden Vorurteile verstärkt und die Ungleichheit in der Sichtbarkeit und Anerkennung zwischen Wissenschaftlerinnen und Wissenschaftlern zementiert, so Amberg und Saunders. Dementsprechend betonen sie die Notwendigkeit, den Effort zur gleichen Repräsentation von Menschen in der Wissenschaft zu verstärken.

2.2.9 Academic medicine's glass ceiling: Author's gender in top three medical research journals impacts probability of future publication success

In der Arbeit „Academic medicine's glass ceiling: Author's gender in top three medical research journals impacts probability of future publication success“ [29] haben die Autor:innen die drei unten aufgelisteten Zeitschriften auf das Geschlecht der Autor:innen untersucht.

1. *New England Journal of Medicine (NEJM)*
2. *Journal of the American Medical Association (JAMA)*
3. Lancet

Diese Zeitschriften haben laut Krstacic u. a. den Verfasser:innen, den höchsten Einfluss in der medizinischen Forschung.

Medizinische Fachzeitschriften haben einen erheblichen Einfluss auf die klinische Praxis und die Empfehlungen von medizinischen Fachgesellschaften. In der akademischen Medizin gibt es jedoch nach wie vor geschlechtsspezifische Ungleichheiten, insbesondere in Bezug auf die Forschungsproduktivität und den beruflichen Aufstieg [32, 2, 12]. Frauen sind seltener in Spitzenzeitschriften vertreten und ihre Arbeiten werden seltener zitiert [7, 28, 35]. Die Zeitschrift Lancet hat sich zur Untersuchung geschlechtsspezifischer Ungleichheiten verpflichtet [48]. Diese Studie hat gezeigt, dass Frauen seltener als Hauptautor:innen, Zweit- oder Senior-Autor:innen in medizinischen Spitzenzeitschriften vertreten sind.

Materialien und Methoden

Um die primäre Variable der Studie, das Geschlecht, zu ermitteln, haben die Forschenden Informationen aus Biografien, Lebensläufen, Pronomen, Vornamen und Autor:innen-Fotos auf Websites extrahiert. Wenn das Geschlecht nicht eindeutig festgestellt werden konnte, wurde es auf der Grundlage einer Diskussion im Studienteam zugewiesen oder als „unbekannt“ bezeichnet. Zusätzlich zu Geschlecht haben die Forschenden auch andere Autor:innen- und Publikationsmerkmale, wie akademische Abschlüsse, Fachgebiete und Studienergebnisse, extrahiert. Die Zuverlässigkeit der webbasierten Datenextraktion, einschliesslich des Geschlechts, haben die Forschenden durch eine hohe „inter-rater reliability“ bestätigt.

Die Studie untersuchte geschlechtsspezifische Unterschiede in drei medizinischen Fachzeitschriften mittels Chi-Quadrat-Tests. Um den Stichprobenumfang zu berechnen, haben die Forschenden Effektgrößen mithilfe des R-Pakets „pwr“ geschätzt. Die Forschenden haben mindestens 1'033 Erstautor:innen-Publikationen verwendet, um die gewünschten geschlechtsspezifischen Unterschiede festzustellen. Die Studie konzentriert sich hauptsächlich auf Erstautor:innen und die Forschenden haben 20 Artikel pro Jahr und Zeitschrift nach dem Zufallsprinzip ausgewählt, was insgesamt 1'080 Veröffentlichungen ergibt. Geschlechtsunterschiede bei Zweit-, Letzt- und bedeutenden Autor:innen können aufgrund des Stichprobenumfangs höchstwahrscheinlich nicht festgestellt werden.

Folgende statistischen Methoden haben die Forschenden für die Analyse der Publikationsraten von Frauen und Männern in medizinischen Zeitschriften verwendet. Die Forschenden haben verallgemeinerte Schätzgleichungsmodelle (GEE) eingesetzt, um Publikationsraten von Frauen und Männern innerhalb und zwischen verschiedenen Zeitschriften zu vergleichen. Chi-Quadrat-Tests haben die Forschenden verwendet, um Merkmale auf Autor:innenebene zu vergleichen, wie zum Beispiel Fachgebiet, Abschluss, Führungsposition und geografische Lage. Zeitliche Vergleiche wurden durchgeführt, und die Forschenden

den haben exakte Tests nach Fisher verwendet, um die Mehrfachveröffentlichungsraten zwischen Frauen und Männern in verschiedenen Autor:innenrollen zu vergleichen. Die Ergebnisse haben die Forschenden mit einem p-Wert von ≤ 0.05 als statistisch signifikant betrachtet, und etwas höhere p-Werte bis zu $p \leq 0.15$ wurden als Trends betrachtet. Alle statistischen Analysen wurden mit SAS 9.4 durchgeführt.

Ergebnisse

Die Studie stellt fest, dass Frauen insgesamt weniger häufig als Erst-, Zweit- oder Letztautorinnen auftraten als Männer. Insgesamt und für jede der drei führenden medizinischen Forschungszeitschriften wurden geschlechtsspezifische Unterschiede bei den Erst-, Zweit- und Letztautorinnen Rollen festgestellt. Der Anteil der Frauen als Erstautorinnen lag insgesamt bei 26.82 %. Dieses Geschlechtergefälle blieb stabil, im ganzen Zeitraum den untersuchten 17 Jahren. Die Autorinnen hatten zudem tendenziell weniger Zitierungen, waren weniger häufig Koautorin von veröffentlichten Artikeln, besaßen weniger häufig einen Dokortitel und hatten seltener Führungspositionen inne. Zudem veröffentlichten Frauen weniger klinische Studien als Beobachtungsstudien ($p < 0.001$), und ihre Projekte konzentrierten sich häufiger auf Infektionskrankheiten als die von Männern, deren Projekte sich am häufigsten auf kardiovaskuläre Themen konzentrierten ($p < 0.001$). Frauen waren auch seltener als Zweit- und Letztautorinnen vertreten. Bei 34.89 % der Veröffentlichungen in medizinischen Forschungsjournalen waren Frauen Zweitautorinnen. Ein Geschlechtergefälle wurde sowohl bei den Zeitschriften insgesamt als auch innerhalb der einzelnen Zeitschriften festgestellt 27.32 % im *NEJM*, 34.12 % im *Lancet* und 43.81 % im *JAMA*, (alle $p < 0.01$). Die Gesamtrate der weiblichen Letztautorinnen lag bei 18.60 %. Diese variierte zwischen 15.08 % im *NEJM*, 19.83 % im *Lancet* und 20.96 % im *JAMA* (alle $p < 0.001$). Die Unterschiede bei den Veröffentlichungsquoten von Frauen waren bei der Rolle der letzten Autorin am dramatischsten.

Diskussion

Die Anzahl von Autorinnen, welche in medizinischen Spitzenforschungsjournalen publizieren, ist nach wie vor unterrepräsentiert, trotz Bemühungen, die Sensibilität dafür zu erhöhen. Diese Studie zeigt, dass Frauen nicht nur als Erst-, Zweit- und Letztautorinnen in hochrangigen medizinischen Fachzeitschriften unterrepräsentiert sind, sondern auch erhebliche Unterschiede in der Vertretung von Wissenschaftlerinnen in diesen Zeitschriften bestehen. Die Veröffentlichung in diesen wichtigen Zeitschriften gilt als Indikator für den beruflichen Erfolg und das künftige akademische Potenzial der **Autorinnen**, was für Beförderungen, Fördermittel und Führungspositionen ausschlaggebend sein kann. Die Unterschiede zwischen den Publikationsergebnissen von Frauen und Männern in medizinischen Zeitschriften tragen zur Aufrechterhaltung der Geschlechterungleichheit in der gesamten Medizin bei. Die Studie zeigt dramatische Unterschiede in der Autorenschaft von Frauen im Vergleich zu Männern, was Anlass zu ernsthaften Bedenken hinsichtlich der Geschlechtergerechtigkeit gibt.

3 Methoden

Das grundsätzliche Vorgehen (Methodik) wurde bereits im Vorprojekt [4] definiert. In diesem wurden folgende technische Teilaufgaben identifiziert, die untereinander keine oder wenige Abhängigkeiten aufweisen:

- ▶ Zitate extrahieren
- ▶ Personen extrahieren
- ▶ Geschlechter von Personen identifizieren
- ▶ Personen den Zitaten zuordnen

Das Extrahieren der Zitate ist von allen Teilaufgaben am unklarsten. Teil dieser Aufgabe ist das Erkennen von eindeutigen Mustern in Zitaten und deren Ausnutzung zur Extraktion. Der Kanadische *GGT* kann dabei Inspiration liefern. Weil das deutsche Sprachmodell fürs *Dependency Parsing* aber andere Labels verwendet als das französische und englische, welche die **Autorinnen und Autoren** des Kanadischen *GGT* verwendet haben, muss dieser Teil neu erfunden werden. Die technische Umsetzung dazu findet sich im Kapitel 4.4.

Die Funktion zum Extrahieren der Personen (und dazugehörigen Beschreibungen) wird die Software mithilfe der *POS Tagging* und *NER* Funktionen des *Spacy* Parsers umsetzen. Die grösste Herausforderung bei dieser Aufgabe ist das *Coreference Resolution*. Denn für das Bestimmen des Geschlechts einer Person ist es hilfreich, möglichst viele Aliase, Artikel und Pronomen zu kennen. Die schlussendliche Realisation ist im Kapitel 4.2 zu finden.

Zum Bestimmen des Geschlechts einer Person können alle zugehörigen Informationen verwendet werden. Am **sprechendsten** sind Artikel und Pronomen, wobei die Zuverlässigkeit ihrer Extraktion nicht sehr hoch ist. Ein weiteres Indiz für das Geschlecht gibt der Vorname. Dazu wollen wir das Namensregister vom Bund zum Abfragen verwenden [52, 53]. Als Backup sollen APIs aus dem Internet dienen. Diese haben ein Rate-Limit und können nicht zu häufig abgefragt werden. Dieser Teil ist im Kapitel 4.3 beschrieben.

Als letzte Teilaufgabe gilt es die Zitate den extrahierten Personen zuzuordnen. Dazu wird die Position im Text und Textähnlichkeit des Namens im Zitat und aller Eigenschaften der extrahierten Person verwendet.

Die schlussendliche Aggregation führt für alle Artikel in der Datenbank die bereits beschriebenen Funktionen aus und speichert das Resultat wie im Kapitel 4.6 beschrieben wieder zurück in die Datenbank.

Nachdem die Software alle rund 350'000 Artikel analysiert hat, gilt es diese auszuwerten. Das Ziel dieser Arbeit ist das Bestimmen des *Gender Gaps*, dem Unterschied im Raum, den Männer und Frauen in den Medien einnehmen. Um diese klar mit einer Zahl benennen zu können, definieren wir den *Gender Gap* wie folgt:

$$\text{Gender Gap} = \frac{RF}{RF + RM} - \frac{RM}{RF + RM} \quad (3.1)$$

Wobei RF der Raum ist, der Frauen gegeben wird und RM derjenige, der den Männern gegeben wird.

Abbildung 3.1: Formel für den *Gender Gap*

Im Fall von dieser Arbeit ist mit „Raum“ die Anzahl Zitate gemeint. Die Abstraktion „Raum“ ist deshalb nützlich, weil sich damit die *Gender Gaps* unterschiedlicher Arbeiten vergleichen lassen. So können wir beispielsweise den *Gender Gap* des „Body Counts“ der Ringier EqualVoice Initiative [1] mit unserem *Gender Gap* anhand der Anzahl Zitate vergleichen. Wir machen uns dies im Kapitel 6.2 zunutze, um die Ergebnisse mit den relevantesten Arbeiten aus dem Kapitel 2 zu vergleichen.

3.1 Technologien

3.1.1 Python

Die Applikation nutzt Python Version 3.10, eine schwach typisierte Skriptsprache, die aufgrund ihrer Flexibilität eine schnelle Entwicklung auch für grössere Anwendungen ermöglicht und dank ihres reichen Ökosystems viele Verwendungsmöglichkeiten bietet. Diese Flexibilität bringt jedoch auch gewisse Herausforderungen mit sich, wie z.B. die Abwesenheit von Typen, Klammern und expliziten Regeln zur Code-Verwendung, die die Zusammenarbeit im Team erschweren können. Aus diesem Grund wurde in diesem Projekt für den produktiven Code streng auf PEP8 mit MyPy-Typisierung gesetzt und dies mithilfe von Linting in den relevanten Code Teilen durchgesetzt.

3.1.2 Git

GIT ist ein verteiltes Versionskontrollsystem, das bei der Verwaltung von Quellcode und dessen Änderungen hilft. Entwickler:innen können mit GIT Änderungen an Code vornehmen, ihre Arbeit mit anderen teilen und zusammenarbeiten. GIT bietet auch Funktionen wie Branching und Merging, um komplexe Entwicklungsaufgaben zu unterstützen.

3.1.3 GitLab

GitLab ist eine Webanwendung, die auf GIT aufbaut und ein umfangreiches Set von Tools für die Zusammenarbeit an Softwareprojekten bereitstellt. GitLab bietet Funktionen wie Projektmanagement, Issue-Tracking, Continuous Integration/Continuous Deployment (CI/CD) und vieles mehr, um die Entwicklung von Software zu erleichtern.

Die BFH hostet eine eigene Instanz von GitLab, auf der wir unser Repo ¹ veröffentlicht haben.

¹<https://gitlab.ti.bfh.ch/aesca4/bachelor-thesis-2023-gender-gap-tracker-schweizer-medien>

In unserem Projekt haben wir GitLab Pipelines eingesetzt um die Codequalität mittels Linting sicherzustellen und das LaTeX Dokument zu builden, sodass stets eine gültige Version verfügbar ist.

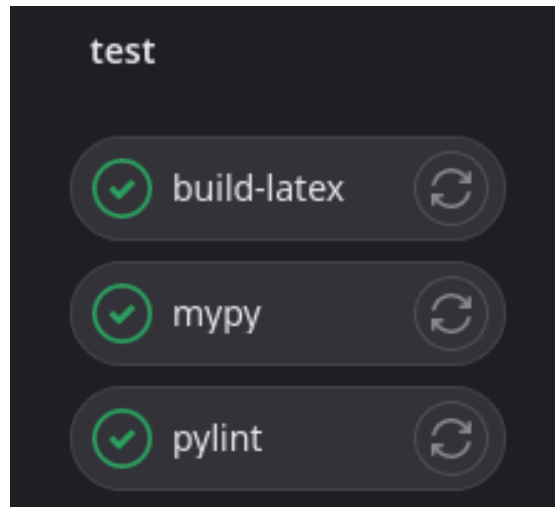


Abbildung 3.2: GitLab Pipeline Web Crawler

Die Abbildung 3.2 zeigt die drei Jobs, die wir dafür verwenden.

3.1.4 MongoDB

MongoDB ist eine dokumentenorientierte NoSQL-Datenbank, die auf flexible Datenmodellierung und Skalierbarkeit ausgelegt ist. Im Gegensatz zu relationalen Datenbanken verwendet MongoDB keine Tabellen, sondern speichert Daten in Dokumenten, die in Collections organisiert sind. Dies ermöglicht eine einfache Handhabung von unstrukturierten Daten und eine schnelle Abfrage von Informationen. Diese Eigenschaften ermöglichten uns eine schnelle und flexible Entwicklung des Programmcodes.

MongoDB bietet auch ein umfangreiches Set von Funktionen wie Aggregation, Indexierung und Volltextsuche, um Entwickler:innen die Arbeit zu erleichtern.

Ausführlichere Informationen zu MongoDB sind auf der offiziellen Webseite ² zu finden. Zum Einbinden des Treibers und der Verwendung mittels Python haben wir öffentlich zugängliche Tutorials der offiziellen Seite von MongoDB³ und der Seite GeeksForGeeks ⁴ verwendet.

²<https://www.mongodb.com/>

³<https://pymongo.readthedocs.io/en/stable/index.html>

⁴<https://www.geeksforgeeks.org/python-mongodb-tutorial/>

3.1.5 Natural Language Processing (NLP)

Natural Language Processing (NLP) beschäftigt sich mit der Verarbeitung von natürlicher Sprache durch Computer und ermöglicht es, diese in eine für Maschinen verständliche Form zu bringen. Dabei kommen Algorithmen und Techniken zum Einsatz, die es ermöglichen, Texte zu verstehen, zu analysieren und zu generieren. Dazu werden häufig Libraries wie *Spacy* (3.1.10) oder *NLTK* ⁵ verwendet.

3.1.6 Named Entity Recognition (NER)

NER ist eine *NLP* Technologie und ermöglicht das automatische Identifizieren von Entitäten mit Namen wie Personen, Orten, Organisationen und Produkten in Texten. Die in diesem Projekt geschriebene Software verwendet *NER* von *Spacy*

3.1.7 Part Of Speech (POS) Tagging

POS Tagging ist eine *NLP* Funktionalität, die dazu dient die Wortarten (Nomen, Verb, Adjektiv...) aller Wörter in einem Satz zu bestimmen. Im Allgemeinen werden für *POS ML* gestützte Programme verwendet, um die Zuordnung der Wörter zu den jeweiligen Kategorien zu bestimmen.

3.1.8 Coreference Resolution

Coreference Resolution ist eine Technik der Sprachanalyse, die darauf abzielt, Pronomen im Text auf ihre Referenz im selben Text zu beziehen. Sie identifiziert das Substantiv, auf das sich ein Pronomen bezieht, um eine eindeutige Bedeutung des Satzes zu ermöglichen. Coreference Resolution ist ein wichtiger Schritt bei der automatisierten Textanalyse und ermöglicht es, die Bedeutung von Texten besser zu verstehen. In dem Fall dieses Projekts ermöglicht die Zuordnung der Pronomen zu den Personen eine effizientere Bestimmung des Geschlechts der Person. Da diese zum Teil Hinweise zum Geschlecht einer Person enthalten (er, sie, sein, ihr...).

3.1.9 Dependency Parsing

Dependency Parsing ist eine Technik des *NLPs*, die Beziehungen zwischen Wörtern in einem Satz untersucht und hierarchisch darstellt. Dabei wird jeder Satz in **einen** Baumstruktur umgewandelt, in dem jedes Wort **eine** Knoten darstellt und die Beziehungen zwischen den Wörtern durch Kanten dargestellt werden. Die Beziehungen sind abhängig von der Bedeutung des Satzes und geben an, welches Wort mit welchem anderen Wort zusammenhängt.

⁵<https://www.nltk.org/>

3.1.10 Spacy

Spacy ist eine Open-Source-Bibliothek für *NLP*, die in Python geschrieben wurde. Sie bietet Entwicklern eine Vielzahl von Funktionen zur Verarbeitung von Texten, einschliesslich Tokenisierung, POS-Tagging, Named Entity Recognition und Dependency Parsing. Spacy wurde für hohe Leistung und Geschwindigkeit entwickelt und ist auf die Verarbeitung grosser Textmengen ausgelegt. Die Library verfügt über Modelle, die in mehreren Sprachen verfügbar sind und es ermöglichen, Texte in verschiedenen Sprachen zu verarbeiten.

In unserem Projekt haben wir das Spacy Modell „de_core_news_lg“⁶ verwendet, um die Funktionen NER, POS Tagging und Dependency Parsing auf deutschsprachige Texte anzuwenden.

3.2 Projektplanung

Der Projektplan orientiert sich am Rahmen, der die BFH für die Bachelor Arbeit vorgibt und ist in vier zum Teil überlappende Projektphasen aufgeteilt:

1. Literaturrecherche
2. Praxisteil
3. Dokumentationsteil
4. Präsentationen

Die Phasen dauern alle in etwa gleich lange, doch der Fokus liegt klar auf der schriftlichen Arbeit, die am meisten Zeit in Anspruch nimmt.

Die Abbildung 3.3 bildet unseren Zeitplan in der Form eines Gantt-Diagrammes ab. In der ersten Spalte von links sind die Arbeitspakete aufgelistet. In der zweiten die geplanten und effektiv verwendeten Zeiten. Die weiteren Spalten repräsentieren dann die zeitliche Dimension im Projekt, dargestellt durch die nummerierten Arbeitswochen sowie dem Datum vom Montag und dem Sonntag der jeweiligen Woche.

Die fettgeschriebenen, einzeiligen Einträge in der ersten Spalte stellen die Projektphasen dar. Deren geplante Dauer wird in dunkelblauer Farbe dargestellt.

Jedes Arbeitspaket nimmt zwei Zeilen ein. Die erste Zeile bildet die Planung ab, die zweite die dokumentierte Realität. Die Zahlen in der Spalte rechts des Paketnamens enthält die geplante und die effektive Zeit in Personentagen, wobei ein Personentag 8 Stunden repräsentiert. Die **Farbigen** Balken stellen den Zeitraum dar, in dem die Arbeitspakete geplant respektive durchgeführt wurden. Blau steht für die Planung und Orange für den effektiven benötigten Zeitraum.

Die Wochen sind viergeteilt, wobei in jedem Viertel ein Personentag geleistet werden kann. Parallel abgearbeitete Arbeitspakete teilen sich diese Zeit. Die Hälfte der Woche stellt in unserem Arbeitsplan die Spanne von Montag Morgen bis Donnerstag Abend dar. Die **Zweite** Hälfte den Zeitraum bis zum Ende des Wochenendes.

⁶https://spacy.io/models/de/#de_core_news_lg

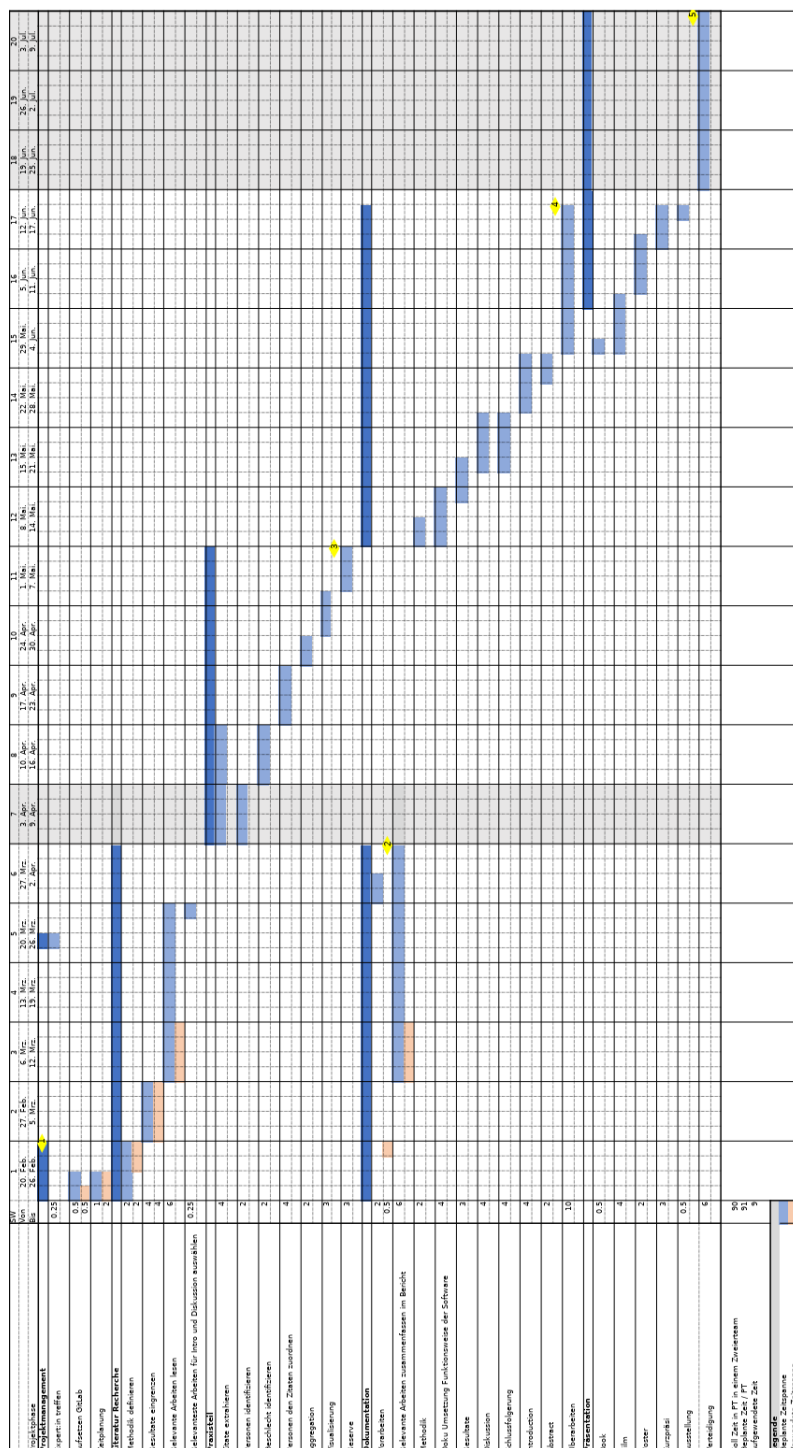


Abbildung 3.3: Gantt Diagramm der Projektplanung

4 Technische Umsetzung

Die technische Umsetzung der durchgeführten Analyse ist von der Pipeline Architektur inspiriert. Die geschriebenen Skripts profitieren von der isolierten Natur der Aufgabe und wenden wo immer möglich funktionale Prinzipien wie Immutability, Higher order- und pure Functions an.

Die untenstehende Abbildung 4.1 zeigt den grundsätzlichen Ablauf der Analyse der Daten. Dabei implementiert die Software sowohl Aspekte des Batch- wie auch des Stream Processings: Einerseits läuft die Auswertung von Anfang bis zum Schluss auf dem gesamten Datensatz durch, andererseits nutzt sie einen MongoDB Cursor, der die Artikel aus der DB Stück für Stück herunterlädt. Dieser Cursor ermöglicht es dem Programm, bereits heruntergeladene Artikel zu analysieren und nach Abschluss wieder aus dem Speicher zu löschen, um Ressourcen zu schonen. Des Weiteren verwendet das Skript einen Processor Pool, der die gleichmässige Verteilung der Last auf alle CPU Kerne garantiert und damit die zur Verfügung stehenden Ressourcen ganz ausnutzt.

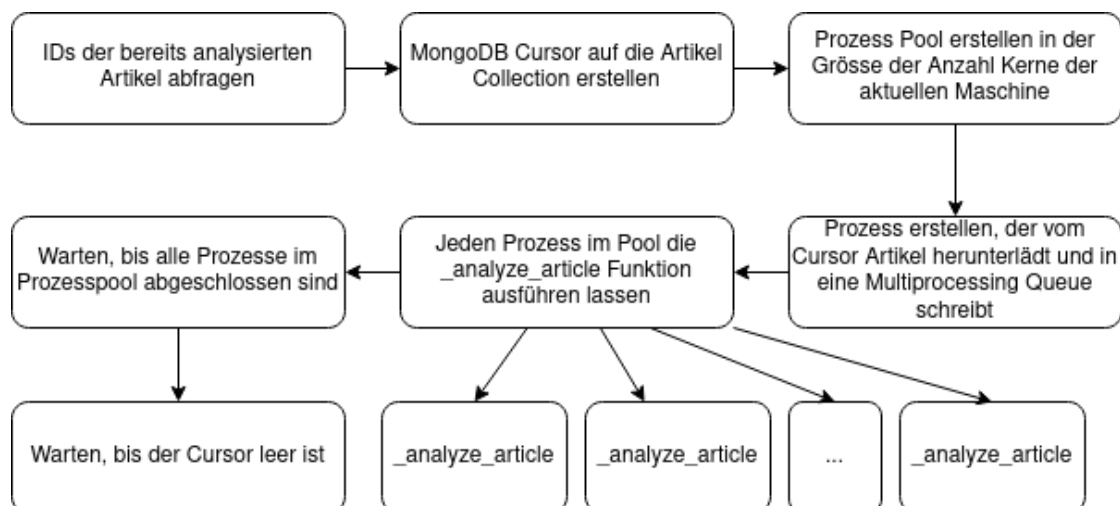


Abbildung 4.1: Ablaufdiagramm der Auswertung

Die Abbildung 4.2 zeigt den Python Code, der über den Prozessor Pool auf die einzelnen CPU Kerne verteilt wird. Diese Funktion konsumiert Artikel aus der „articles.queue“, einer Multiprocessing-sicheren Datenstruktur, die nach und nach vom MongoDB Cursor mit Artikeln aus der DB gefüllt wird.

Die Funktion läuft so lange, bis sie für mindestens zehn Minuten keine neuen Artikel in der Queue mehr findet. Der Grund für dieses Timeout **ehrer hohe sind** mögliche Netzwerkis-

sues oder andere unvorhergesehene Unterbrüche. Da das Skript bei der grossen Datenmenge sowieso mehrere Tage läuft, spielen die 10min warten am Schluss keine entscheidende Rolle auf die Performanz.

Als Erstes nach dem Erhalt eines neuen Artikels muss der Algorithmus, bestimmen, ob der aktuelle Artikel bereits analysiert wurde. Falls ja, verwirft er ihn wieder und holt sich einen neuen.

Danach beginnt die effektive Analyse. Einzelne Funktionen daraus sind in den nachfolgenden Kapiteln näher beschrieben.

Zuerst bestimmt die Funktion das Geschlecht des Autors, danach durchsucht sie den Text nach Personen und Zitaten. Zum Schluss weist sie die gefundenen Zitate den gefundenen Zitaten zu und speichert das Resultat in der Datenbank.

Wenn die Abfrage der Queue in ein Timeout läuft (zehn Minuten), dann wirft die Funktion eine „Empty“ Exception, die vom „except“ Block abgefangen wird und dazu führt, dass die Funktion und damit der Prozess beendet wird.

```
def _analyze_article(processor_nr: int) -> None:
    log.info("Starting processor %d", processor_nr)
    while True:
        try:
            article_dict = articles_queue.get(
                timeout=600
            ) # 10min, in case of network issues

            if article_dict["_id"] in analyzed_articles:
                continue

            builder = ArticleBuilder()
            article = builder.from_db_result(article_dict).build()

            genderized_author = get_genderized_person(
                get_person_from_string(article["author"])
            )

            article_text = article["text"]
            people = get_people_from_string(article_text)
            genderized_people = map(get_genderized_person, people)

            citations = get_syntactic_quotes(article_text)
            citations_with_person = assign_people_to_citations(
                genderized_people, citations
            )

            analyzed_article = get_analyzed_article(
                article, genderized_author, citations_with_person
            )
            insert_analyzed_article(analyzed_article)

        except Empty:
            log.info("Queue is empty! Terminating processor %d!", processor_nr)
            return
```

Abbildung 4.2: Funktion: _analyze_article()

Die nachfolgenden Kapitel enthalten genauere Beschreibungen zu den Funktionen, die Personen und Zitate finden, sowie das Geschlecht von Personen bestimmen.

4.1 Datensammlung bereinigen

Die Vorarbeit hat eine Datensammlung von etwa 370'000 Nachrichtenartikeln ergeben, die jeweils über eine eindeutige URL verfügen. Wir haben durch die Verwendung von Unique Constraints sichergestellt, dass jede URL einzigartig ist. Allerdings vermuteten wir,

dass die Datensammlung Duplikate von Texten enthält, die unter verschiedenen URLs veröffentlicht wurden. Dies könnte durch Weiterleitungen auf andere Seiten oder veränderte URLs verursacht worden sein. Um eine genaue Aussage über den Gender Gap in diesen Nachrichtenportalen machen zu können, mussten wir alle Duplikate pro Nachrichtenportal eliminieren.

Nach einigen Testversuchen mit 1000 Artikeln entschieden wir uns dafür, die Cosine-Similarity zu verwenden, um zwei Artikeltexte zu vergleichen. Dieser Vergleich war bei den Tests performanter als die Levenshtein-Distanz, die ebenfalls zur Auswahl stand. Der Rechenaufwand für den Vergleich zweier Zeichenketten der Länge m bzw. n ist bei der Levenshtein-Distanz $O(m*n)^1$, während er bei der Cosine-Similarity nur $O(m+n)$ beträgt.

Die Cosine-Similarity ist eine Metrik zur Messung der Ähnlichkeit zwischen zwei Vektoren, bei der der Kosinus des Winkels zwischen den Vektoren verwendet wird. In unserem Fall sind die Vektoren die Artikeltexte. Um die Ähnlichkeit zwischen einem Vektor und einer Menge von anderen Vektoren zu berechnen, erstellen wir eine Ähnlichkeitsmatrix. Wenn der Kosinus eins ist, sind die Texte gleich. Als Schwellwert für die Ähnlichkeit haben wir 0.9 festgelegt, da dieser bei unseren Tests am besten funktionierte. Wenn wir einen höheren Schwellwert genommen hätten, hätten wir nicht alle Duplikate gefunden und bei einem zu niedrigen Schwellwert hätten wir zu viele Texte als Duplikate interpretiert, die gar keine sind.

Da jeder Artikeltext mit jedem anderen verglichen werden muss, hätte der Algorithmus eine Matrix in der Grösse von 371'653 x 371'653 erstellen müssen und 42,9 GB Arbeitsspeicher benötigt. Das war für unsere Arbeitsmaschinen zu viel. Daher haben wir das „divide and conquer“ Verfahren angewendet und aus einem grossen Problem kleine Teilprobleme gemacht. Für jeden Artikel, den wir verglichen haben, haben wir eine neue Matrix erstellt, sodass sie nur noch 1x371'653 gross war und dann nacheinander 371'653 Mal verglichen wurde. Unten sind Auszüge aus der Monog-DB der Anzahl Artikel vor und nach der Entfernung der Duplikate zu sehen.

```
{ "_id" : "20min", "count" : 20838 }
{ "_id" : "Watson", "count" : 77276 }
{ "_id" : "blick", "count" : 197311 }
{ "_id" : "srf", "count" : 76228 }
```

Abbildung 4.3: Anzahl Artikel pro Nachrichtenportal vor der Bereinigung der Duplikate

```
{ "_id" : "20min", "count" : 20487 }
{ "_id" : "Watson", "count" : 62381 }
{ "_id" : "blick", "count" : 194881 }
{ "_id" : "srf", "count" : 73272 }
```

Abbildung 4.4: Anzahl Artikel pro Nachrichtenportal ohne Duplikate

4.2 Personen extrahieren

Das Extrahieren von Personen aus einem Text umfasst mehrere Schritte. Zunächst wird mithilfe von *NER* die Entitäten mit Personen Tag aus dem Text herausgelesen. Als Nächstes wird *Coreference Resolution* verwendet, um Personencluster zu bilden. Dies bedeutet, dass der Text auf Referenzen für Personen überprüft wird, z.B. wenn eine Person im Text als „sie“ oder „er“ bezeichnet wird. *Coreference Resolution* hilft dabei, alle Referenzen für dieselbe Person zu identifizieren und in Gruppen zu organisieren. Um eindeutige Personen zu erhalten, müssen dann die NER-Entitäten und die Coreference-Entitäten zusammengeführt werden. Dazu werden die Schnittmengen der Wörter verwendet, die den NER- und Coreference-Entitäten zugeordnet sind. Auf diese Weise können Personen und zugehörige Informationen eindeutig identifiziert werden. Zusätzlich zu den Vornamen und Nachnamen der Person werden auch Pronomen und Artikel, die mit der Person in Verbindung stehen aus der Coreference-Entität, mithilfe von *POS* extrahiert und der Person zugewiesen. Diese zusätzlichen Informationen zur Person sind im nächsten Schritt hilfreich, um das Geschlecht der Person zu bestimmen. Das Resultat vom Personen extrahieren ist am Ende eine Liste von Personen in der Form wie unten aus der zugehörigen Python Klasse ersichtlich.

```
class Person:
    first_name: str
    last_name: str
    pronouns_and_articles: list[str] # er, sie, ihr, ihre, sein, seine der,
    die, das
    substitute_nouns: list[str] # Informatikerin, Studentin, Schwester,
    Tochter, Prinz, Experte
    positions_in_text: list[int]
```

Abbildung 4.5: Person Klasse

4.3 Geschlecht identifizieren

Um das Geschlecht einer Person herauszufinden, werden als Erstes die offiziellen Listen¹ von Vornamen aus dem Jahr 2021 des Bundesamts für Statistik konsultiert. Davon gibt es zwei, eine mit männlichen und eine mit weiblichen Vornamen. Dann werden noch die Pronomen der Person, falls vorhanden, verglichen, ob sie wahrscheinlicher männlich oder weiblich sind. Wenn die Wahrscheinlichkeit aus den Pronomen und der Liste vom Bundesamts für Statistik unter 66% liegt, dass es zum einen Geschlecht gehört, wird noch die API von Genderize.io mit dem Vornamen abgefragt. Bei dieser API gibt es jedoch eine Begrenzung von 1000 Requests pro Tag. Darum wird, wenn dieses Limit erreicht ist, trotzdem das wahrscheinlichere Geschlecht aus dem Resultat der Liste vom Bundesamts für Statistik und den Pronomen gewählt. Das Resultat welches Geschlecht die Person wird dann in

¹<https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/geburten-todesfaelle/namen-schweiz.html>

folgender Form zurückgegeben mithilfe des Gender Enums:

```
class GenderizedPerson(Person):  
    gender: Optional[Gender] # Unknown => None  
    probability: float  
  
class Gender(Enum):  
    FEMALE = 1  
    MALE = 2  
    NEUTRAL = 3
```

Abbildung 4.6: GenderizedPerson Klasse und Gender Enum

4.4 Extraktion der Zitate

Die vorliegende Auswertung bezüglich der Anzahl Zitate von Männern und Frauen pro Nachrichtenportal fokussiert sich auf das Erkennen und Extrahieren der sogenannten *Syntaktischen Zitate*. Der Grund dafür ist, dass diese von der Form her die klarste Struktur aufweisen und auch die Basis für das Erkennen der anderen Arten von Zitaten darstellen.

Weitere Arten von Zitaten, die in dieser Arbeit aufgrund des zeitlichen Rahmens nicht berücksichtigt werden konnten sind

1. *Schwimmende Zitate*,
2. *Heuristische Zitate* und
3. *Syntaktische Zitate*, die ein zu unspezifisches oder gar kein Verb enthalten.

Nachfolgend ist deshalb ausschliesslich die Extraktion der *Syntaktischen Zitate* mit spezifischen Verben beschrieben. Auf die weiteren Arten und deren mögliche Erkennung wird im Kapitel 6.4 eingegangen.

4.4.1 Syntaktische Zitate

Die *Syntaktischen Zitate* zeichnen sich in der Textanalyse mit *Spacy Parser* dadurch aus, dass sie im resultierenden Parse Tree in einer fast eindeutigen Struktur auftauchen und deshalb mit einem klar definierten Regelwerk extrahiert werden können. Beispiele dazu finden sich in den nachfolgenden Unterkapiteln.

4.4.2 Parse Trees

Die Analyse eines Texts (Strings) mit dem *Spacy Parser* resultiert in einem Parse Tree, der die Abhängigkeit der Wörter untereinander enthält. Die Aufgabe des Programms besteht also darin, Subtrees in einer gewissen Struktur zu erkennen. Die Abbildung 4.7 ist eine Abstraktion eines solchen Subtrees.

In den eckigen Klammern ist die Art des Worts beschrieben. Diese wird gefolgt von einem Bodenstrich (_) und der *Spacy*-Abkürzung für die Kategorie des Worts.

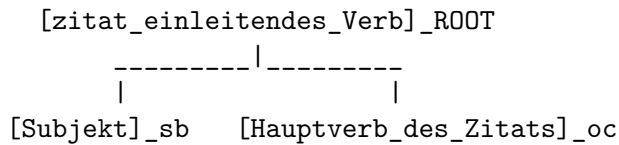


Abbildung 4.7: Subtree Repräsentation eines Syntaktischen Zitats

Die möglichen Kategorien sind in der Tabelle 4.1 beschrieben. Diese Kategorien widerspiegeln nicht konsequent Wortarten oder Begriffe aus der Satzlehre, wie sie im Deutschunterricht gelehrt werden. Deshalb ist auch mit der Erklärung oft nicht klar, wofür sie stehen. Deren Auftreten ist jedoch konsistent genug, sodass sie zur Erkennung der gewünschten Satzstrukturen genutzt werden können.

Spacy Parser Label	Spacy Erklärung
ROOT	root
da	dative
mnr	postnominal modifier
mo	modifier
nk	noun kernel element
oc	clausal object
punct	punctuation
sb	subject

Tabelle 4.1: Legende zu den wichtigsten Spacy Parser Labels

Die nachfolgenden Abbildungen 4.8 und 4.9 sind konkrete Beispiele solcher Trees. Der *Spacy Parser* überrascht mit seiner Fähigkeit, Zitate in der direkten und indirekten Rede gleich strukturieren zu können.

Die Abbildung 4.8 repräsentiert den Satz „Macron sagte zu Xi, »Die Aggression hat der Stabilität einen Schlag versetzt.«“ mit einem Zitat in der direkten Rede.

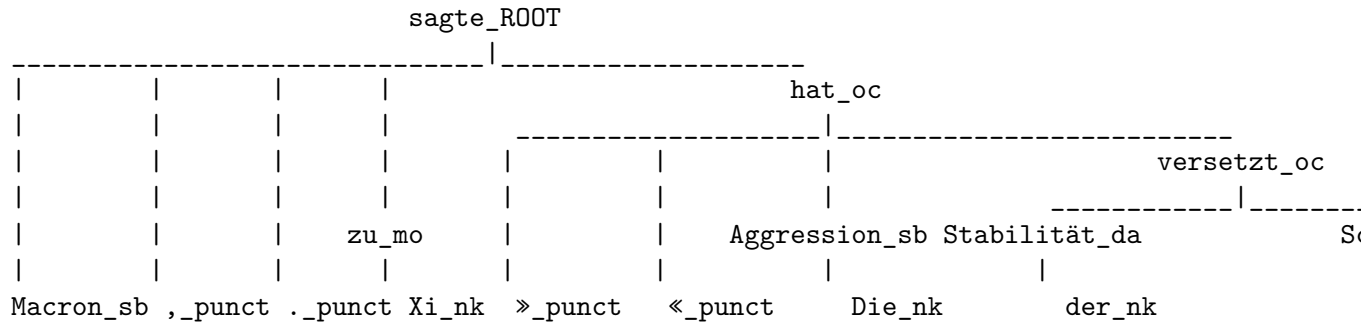


Abbildung 4.8: Parse Tree eines Satzes mit einem Syntaktischem Zitat in der direkten Rede

Die Abbildung 4.9 repräsentiert den Satz „Macron sagte zu Xi dass die Aggression der Stabilität einen Schlag versetzt habe.“ mit einem Zitat in der indirekten Rede.

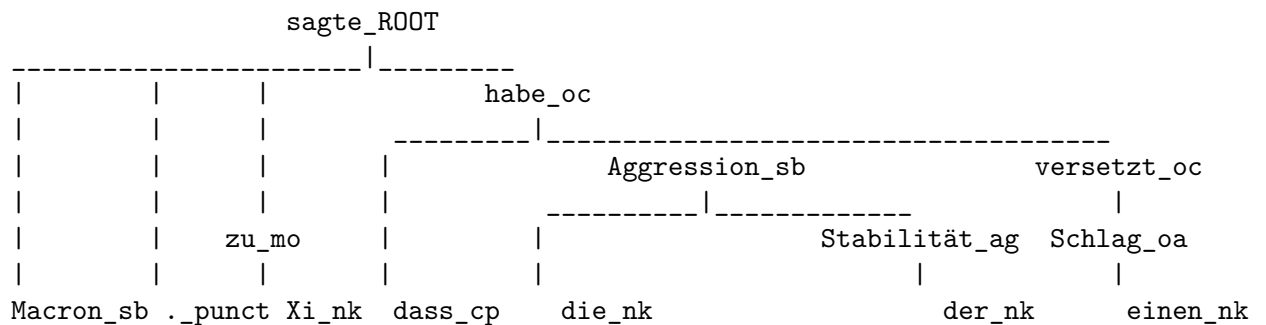


Abbildung 4.9: Parse Tree eines Satzes mit einem Syntaktischem Zitat in der indirekten Rede.

4.4.3 Ablauf des Algorithmus

Der nachfolgende Abschnitt soll konzeptuell und mit einigen Code-Beispielen erklären, wie der Algorithmus vorgeht, um die Zitate in der oben beschriebenen Struktur zu Erkennen und zu extrahieren.

Hinweis: Im nachfolgenden Text werden die Terme „Node“, „Knoten“ und „Token“ als Synonyme verwendet. Der Grund dafür ist, dass „Token“ die Namensgebung von *Spacy* ist und „Node“ und „Knoten“ die passenden Begriffe dazu aus der Graphtheorie sind.

Der nachfolgende Code 4.10 zeigt die Funktion, die das Programm verwendet, um den „Zitat Node“ aus dem Parse Tree zu extrahieren. Gemeint ist damit derjenige Knoten, der als Kinder alle Wörter des Zitats enthält. Grundsätzlich muss die Funktion dabei unterscheiden, ob der vorliegende Satz in einer Zeitform mit Hilfsverb vorliegt (Perfekt, Plusquamperfekt oder Futur) oder in einer Zeitform ohne (Präsens oder Präteritum).

Die Funktion bedient sich einer weiteren Funktion „`_get_nearest_token_by_condition`“ (4.11), um in einer Breitensuche nach dem entsprechenden Node zu suchen.

```
def __get_quote_node(root: Token) -> Token:
    result = []

    # Perfekt, Plusquamperfekt, Futur
    if root.lemma_ in __get_hilfsverben():
        condition = (
            lambda n: n.head.lemma_ in __get_quotation_verbs() and n.dep_ == "oc"
        )
        result = __get_flattened_list(
            [__get_nearest_tokens_by_condition(c, condition) for c in
             root.children]
        )

    # Präsens, Präteritum
    if root.lemma_ in __get_quotation_verbs():
        result = [x for x in root.children if x.dep_ == "oc"]

    if len(result) < 1:
        raise _NotFoundError()

    return result[0]
```

Abbildung 4.10: Function __get_quote_node

Die Funktion „__get_nearest_token_by_condition“ (4.11) wird vom Algorithmus verwendet, um die Knoten des Subjekts, des Zitat-einleitenden Verbes und des Zitats zu finden. Die Funktion durchsucht den gegebenen Subtree mit der Strategie der Breitensuche, um den nächsten Knoten im Baum zu finden, der eine vorgegebene Bedingung erfüllt. Die Bedingung ist dabei jeweils spezifisch für den gesuchten Knotentyp.

Das Resultat ist derjenige Knoten, der den kürzesten Pfad zum gegebenen Ursprungsknoten aufweist und die Bedingung erfüllt.

```
# Breadth First Search
def __get_nearest_tokens_by_condition(
    node: Token, condition: Callable[[Token], bool]
) -> List[Token]:
    def get_nearest_by(
        node: Token, condition: Callable[[Token], bool], depth: int
    ) -> List[Tuple[Token, int]]:
        # check current node
        if condition(node):
            return [(node, depth)]

        # Recursion step (flatten result)
        return __get_flattened_list(
            [get_nearest_by(n, condition, depth + 1) for n in node.children]
        )

    results = get_nearest_by(node, condition, 0)
    if len(results) < 1:
        return []
    min_depth = min(results, key=lambda t: t[1])[1]

    return list(map(lambda t: t[0], filter(lambda t: t[1] == min_depth,
        results)))
```

Abbildung 4.11: Function `__get_nearest_token_by_condition`

Da die identifizierte Struktur 4.7 von **Syntaktischen** Zitaten nicht eindeutig ist, müssen die Zitat-Kandidaten mithilfe einer Liste von „Zitat-einleitenden Verben“ gefiltert werden, um False-Positives zu vermeiden.

Die nachfolgende Liste 4.12 wird verwendet, um die Zitate zu identifizieren.

ankündigen, argumentieren, aufrufen, begrüßen, behaupten, berichten,
bestätigen, betonen, bezeichnen, bringen, dementieren, empfehlen,
erklären, erwidern, erzählen, fassen, feststellen, fragen, kontern,
kündigen, meinen, mitteilen, nennen, rechnen, rufen, sagen, schreiben,
stellen, teilen, twittern, verraten, versichern, verweisen, werfen,
zitieren, zusammenfassen

Abbildung 4.12: Quotation verbs

Dieser Ansatz ist Fehleranfällig, da die Liste umfassend genug sein muss, um alle möglichen Verben zu beinhalten, mit denen Zitate eingeleitet werden können. Gleichzeitig, müssen die Verben sehr spezifisch sein, weil sonst Satzstrukturen als Zitate erkannt werden, die gar keine sind. Weil das Resultat des Parsers nicht eindeutig ist (ambiguous), ist die Liste nötig. Hier verliert das Programm wahrscheinlich einiges an Präzision. Mehr dazu im folgenden Kapitel 4.5

4.5 Qualitätssicherung

Weil die Extraktion der Zitate keiner wasserdichten Logik folgen kann, war es für die Entwicklung des Algorithmus, aber auch um eine Aussage zu der Qualität der Resultate treffen zu können, wichtig, die Qualität und die Verbesserung des Algorithmus zu messen ~~zu~~.

Ein eigens dafür entwickelter Bewertungsalgorithmus soll einen möglichst guten Überblick über die Anzahl der gefundenen Zitate und deren Qualität geben.

Dazu haben wir in manueller Arbeit pro Nachrichtenportal fünf Artikel ausgewertet und als JSON abgelegt. Insgesamt ergeben sich daraus 20 Test-Artikel. Der Bewertungsalgorithmus ruft die Funktionalität zum Extrahieren der Zitate mit dem rohen Text der Artikel auf und vergleicht das Resultat im Anschluss mit den manuell erstellten Lösungen. Als wichtigste Metrik dient dabei der „Recall“ (Abbildung 4.13).

4.5.1 Recall

Diese Metrik wird häufig auch im *ML* verwendet, um die Genauigkeit des trainierten Modells zu messen. Der Recall sagt in unserem Fall aus, wie viele der relevanten Zitate der Algorithmus finden konnte, als Prozentzahl der totalen Anzahl Zitate, die er hätte finden können.

$$Recall = \frac{Anzahl\ gefundene\ Zitate}{Totale\ Anzahl\ Zitate} \quad (4.1)$$

Abbildung 4.13: Formel Recall

Als **gefundene Zitate** zählt der Bewertungsalgorithmus all **diejenigen** Zitate, die eine Qualität von mindestens 60% aufweisen. Mehr dazu **im** Kapitel 4.5.2

Die nachfolgende Abbildung 4.14 visualisiert den Recall von 61.2% über alle Nachrichtenportale.

Dieser Wert sagt aus, dass das Programm von allen vorhandenen Zitaten in den Lösungen 61.2% gefunden hat.

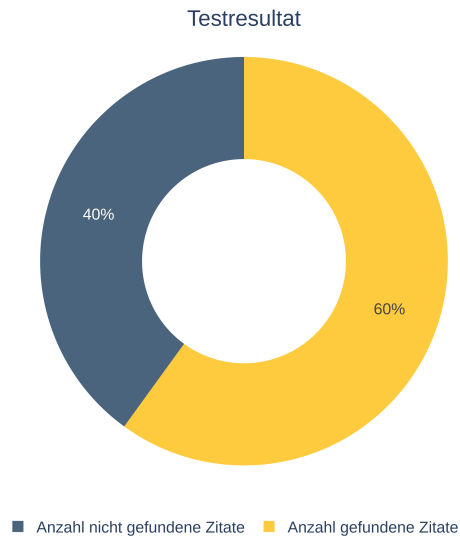


Abbildung 4.14: Recall der Zitate auf dem Testset

Das untenstehende Balkendiagramm 4.15 bietet einen Überblick über die Performance aufgeschlüsselt nach Nachrichtenportal. Es fällt auf, dass die Daten von Blick mit 70.6% deutlich besser analysiert werden als diejenigen von 20min mit 50.0%.

Wahrscheinlich ist diese Diskrepanz auf die geringe Anzahl Testfälle zurückzuführen. Mit mehr Testfällen würde sich der Wert wahrscheinlich bei allen Nachrichtenportalen bei 60% einpendeln.

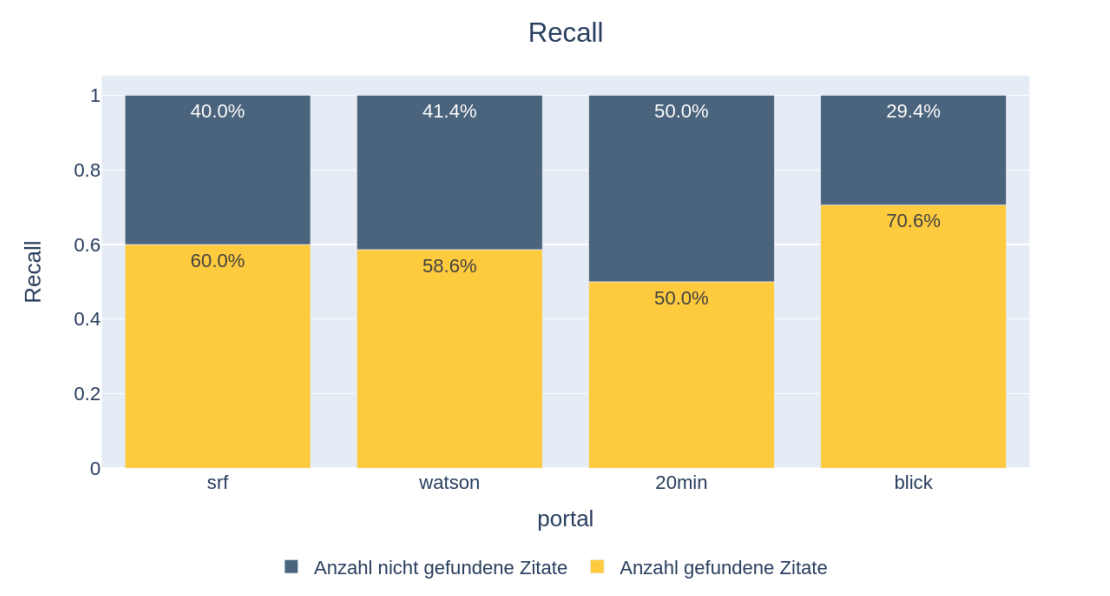


Abbildung 4.15: Recall der Zitate pro Nachrichtenportal

4.5.2 Qualitätsnote

Die Qualitätsnote setzt sich aus dem Durchschnitt der Stringähnlichkeiten der extrahierten Zitate mit der Lösung zusammen. Zum Bestimmen der Stringähnlichkeit verwendet der Algorithmus den „SequenceMatcher“² von Python. Dieser findet den längsten übereinstimmenden Substring aus zwei Strings relativ zur Länge der beiden zu vergleichenden Strings.

Die Vermutung, dass die geringe Anzahl Tests Schwankungen in der Qualität verursacht wird durch den Überblick in den folgenden Diagrammen (Abbildung 4.16) gestützt. 20min hat deutlich weniger Zitate in den Lösungen als die anderen Nachrichtenportale.

Die Histogramme legen ausserdem nahe, dass die Rate der False-Positives gering ist. Denn diese machen sich durch eine tiefe Note bemerkbar, da sie sehr unähnlich zu den Zitaten aus den Lösungen sind. Dass davon wenige zu erkennen sind, suggeriert, dass False Positives selten sind.

²<https://docs.python.org/3/library/difflib.html#difflib.SequenceMatcher>

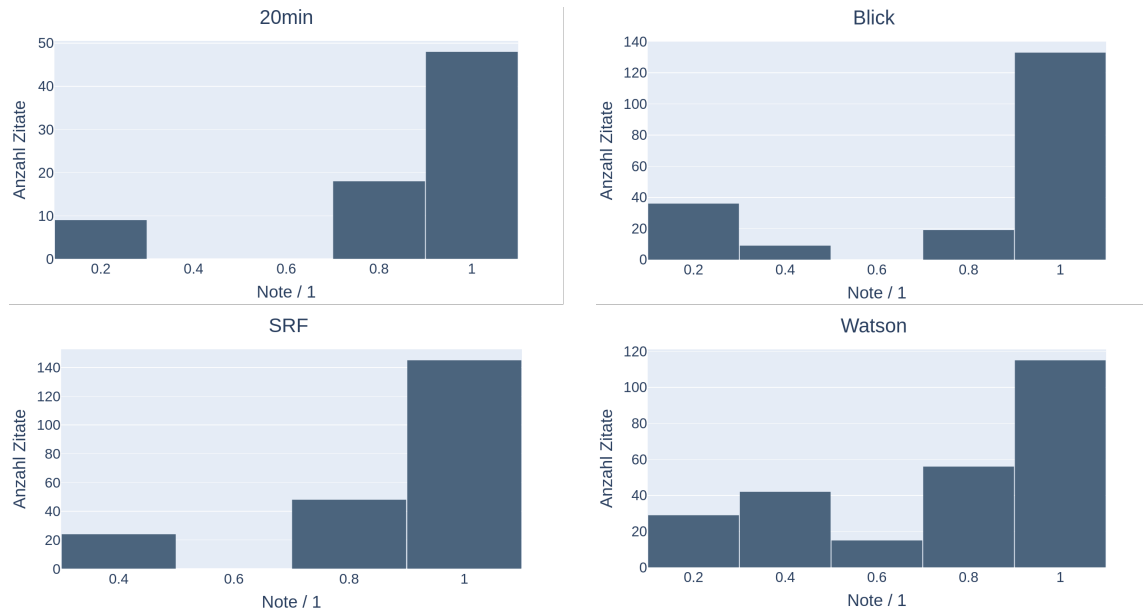


Abbildung 4.16: Histogramme zur Verteilung der Qualität der gefundenen Zitate

4.5.3 Nicht erkannte Zitate

Das nachfolgende Unterkapitel soll einen Überblick über die nicht erkannten Zitate liefern und erklären, weshalb der Algorithmus aktuell noch nicht in der Lage ist, diese zu erkennen.

Grundsätzlich gibt es viele mögliche Erklärungen, weshalb das Programm gewisse Zitate nicht erkennen kann. Aufgrund der analysierten Fälle scheinen die meisten fehlenden Zitate entweder neuen, nicht implementierten Satzstrukturen oder Fehlern des Parsers geschuldet zu sein.

So hat sich gezeigt, dass der Parser in langen Sätzen mit mehr als zwei Satzteilen die Abhängigkeiten anders aufbaut. So wird der Subtree mit dem Zitat einleitenden Verb anders aufgebaut und ~~wird~~ vom Algorithmus ignoriert.

So wird das Zitat einleitende Verb „schreibt“ im folgenden Beispielsatz (Abbildung 4.17) als „mnr“ (postnominal modifier) gekennzeichnet. Wann und wie die gesuchten Verben mit diesem Label auftreten, konnten wir nicht mehr herausfinden. „Dort habe er in einer Wohnung auch einen Grossteil seiner Sachen, unter anderem Familienerbstücke, wie das Inventar aus dem Restaurant Rossberg, wie der «Landbote» schreibt.“

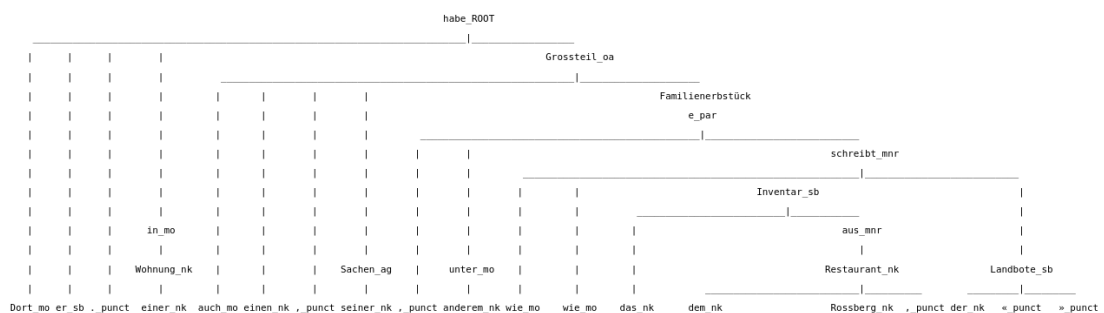


Abbildung 4.17: Zitat mit anderer Baumstruktur

Bei diesem Satz (Abbildung 4.18) wird das signifikante Verb „mitteilte“ als „mo“ (modifizier) gelabelt. In diesem Fall scheint der Parser das Label falsch vergeben zu haben, denn „modifizier“s werden sonst eher den Adverbien und Präpositionen zugeordnet. Möglicherweise hat der Wortteil „mit“ in diesem Fall den Ausschlag gegeben. „Das Departement für Verteidigung, Bevölkerungsschutz und Sport (VBS) wird bis Ende Jahr die rechtlichen Grundlagen erarbeiten, wie der Bundesrat am Mittwoch mitteilte.“

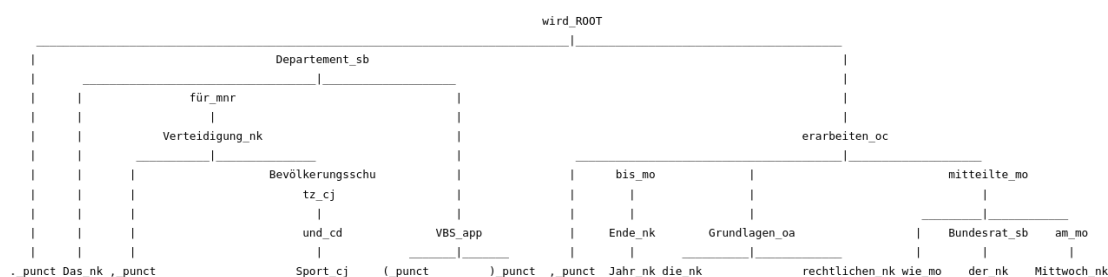


Abbildung 4.18: Zitat mit anderer Baumstruktur

4.6 Datenbank

Wir haben im Vorprojekt von der BFH eine virtuelle Maschine gemietet, auf die wir über *SSH* zugreifen können. Auf dieser haben wir eine MongoDB, um dort die Artikel Daten zu speichern. Die Collection „analyzed_articles“ enthält die analysierten Artikel. Diese enthalten die gefundenen Zitate mit Personen und deren Geschlecht. Unten ein Beispiel Eintrag aus dieser Collection.

analyze_articles

- ▶ **title:** Titel des Artikels
- ▶ **lead:** Lead Text des Artikels

- ▶ **url:** *URL* zum Artikel, eindeutiger Schlüssel
- ▶ **author:** Autor:in des Artikels
- ▶ **source:** Das Nachrichtenportal
- ▶ **published:** Timestamp der Veröffentlichung
- ▶ **updated:** Timestamp der letzten Aktualisierung
- ▶ **text:** Text des Artikels
- ▶ **quotes:** Zitate in Form von Liste
 - **designation:** Person von Zitat
 - **gender:** Geschlecht der Person
 - **quotation verb:** Verb welches zum Zitat einleitet
 - **quote:** Zitat Text
 - **start of quote in text:** Start Position von Zitat im Text

```
{
  "_id": "794aceeeea0911eda58075d660e6a249",
  "article": {
    "_id": "cf49cff695fa1ed8a280242ac110002",
    "title": "Nimm dich nicht zu wichtig!",
    "lead": "Bescheidenheit ist eine Zier, doch weiter kommt man  
ohne ihr, besagt ein Sprichwort. Doch zahlt sich  
Bescheidenheit wirklich weniger aus als Selbstdarstellung?  
Eine Bestandesaufnahme.",
    "author": {
      "designation": "Raphael Zehnder",
      "gender": "male"
    },
    "source": "srf",
    "url":
      "www.srf.ch/kultur/gesellschaft-religion/wochenende-gesellschaft/
      bonus-bescheidenheit-nimm-dich-nicht-zu-wichtig",
    "published": 1577005620,
    "updated": 1577005620,
    "text": "Egozentrische Pfauen tummeln sich allenthalben:  
Wirtschaftskapitäne blasen sich auf, Politikerinnen und  
Politiker lobpreisen ..."
  },
  "quotes": [
    {
      "subject": {
        "designation": "Walter Slezak",
        "gender": "male"
      },
      "quotation_verb": "brachte",
      "quote": "Wir kaufen Dinge, die wir nicht brauchen, um  
Menschen zu beeindrucken, die wir nicht mögen",
      "start_of_quote_in_text": 555
    }
  ]
}
```

Abbildung 4.19: analyzed_articles Collection Eintrag Beispiel

5 Resultate

Wir haben aus der `analyzed_articles` Collection verschiedene Informationen zu der Anzahl Artikel pro Geschlecht extrahiert. Diese Informationen und sonstige wichtige Resultate wie die Datenbereinigung, welche wir zur Untersuchung des Gender Gaps anhand von Zitaten brauchen, sind in den folgenden Diagrammen dargestellt.

Die Abbildung 5.1 zeigt die Anzahl einzigartiger Artikel pro Nachrichtenportal nach der Datenbereinigung. Wir fanden bei Blick 2'430 Duplikate, bei SRF 2'956, bei Watson 14'895 und bei 20 Minuten 351.

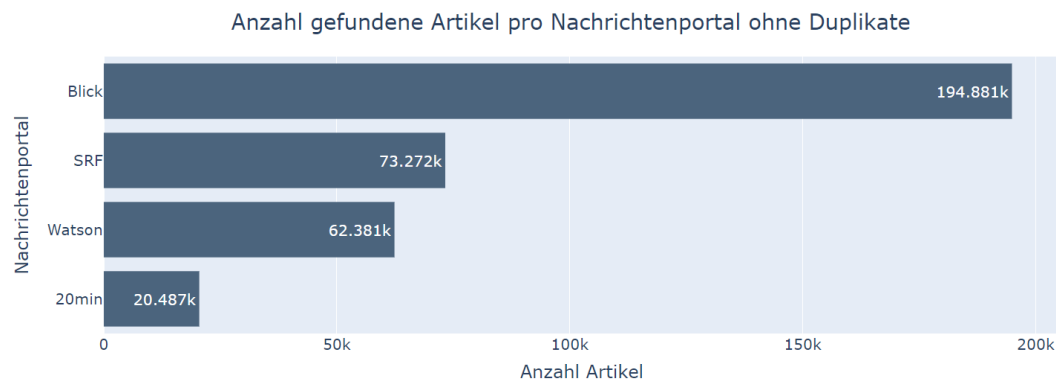


Abbildung 5.1: Anzahl gefundene Artikel, Stand 22.04.2023

Wie in Abbildung 5.2 zu sehen ist, haben wir die meisten Zitate bei Blick gefunden. Dies liegt wahrscheinlich daran, dass wir von diesem Portal auch die meisten Artikel in unserer Datensammlung haben, wie in Abbildung 5.1 ersichtlich. Im Vergleich zu Watson haben wir bei SRF mehr Zitate gefunden, obwohl die Anzahl der Artikel geringer ist. Eine Aufschlüsselung der Anzahl der Zitate pro Artikel ist in Abbildung 5.4 zu finden.

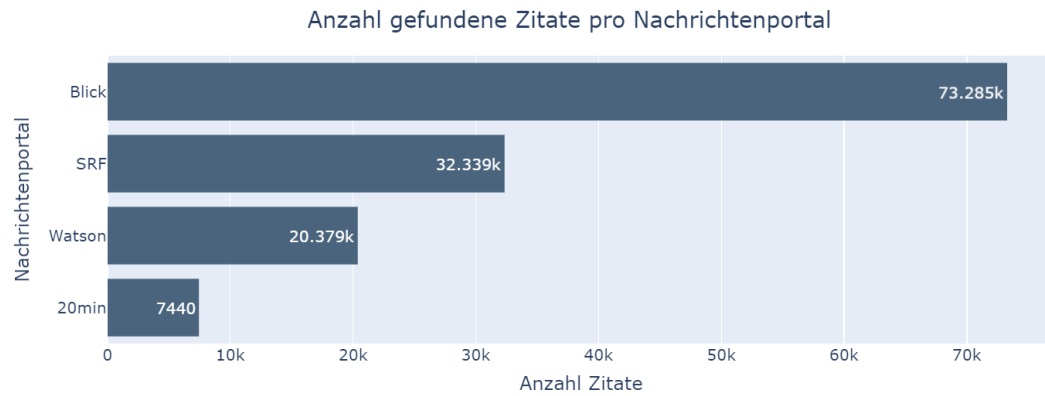


Abbildung 5.2: Anzahl gefundene Zitate

Die wichtigste Erkenntnis aus unserer Untersuchung ist in folgendem Diagramm (siehe Abbildung 5.3) abgebildet, das die prozentuale Verteilung von Zitaten zwischen Frauen und Männern zeigt. Zitate von Personen, bei denen wir das Geschlecht nicht eindeutig zuordnen konnten, sind unter der Kategorie „unbekannt“ aufgeführt. Die Analyse aller untersuchten Nachrichtenportale ergab, dass etwa zwei Drittel aller Zitate von männlichen Personen stammen. Mithilfe automatisierter Auswertung wurden die Zitate in Texten erkannt, entsprechend dem Geschlecht gruppiert und gezählt. Basierend auf unseren Qualitätssicherungstests schätzen wir, dass wir etwa 60% aller Syntaktischen Zitate erfasst haben.

Die Farben im Diagramm (siehe Abbildung 5.3) repräsentieren die Farben der BFH, wobei Männer mit Gelb und Frauen mit Blau dargestellt werden. Wir haben uns bewusst für diese Farbwahl entschieden, weil Blau normalerweise als Männerfarbe gilt, auch für Kleinkinder. Diese Geschlechterklischees werden leider schon frühzeitig aufgedrückt und dem wollten wir hier etwas entgegen wirken.

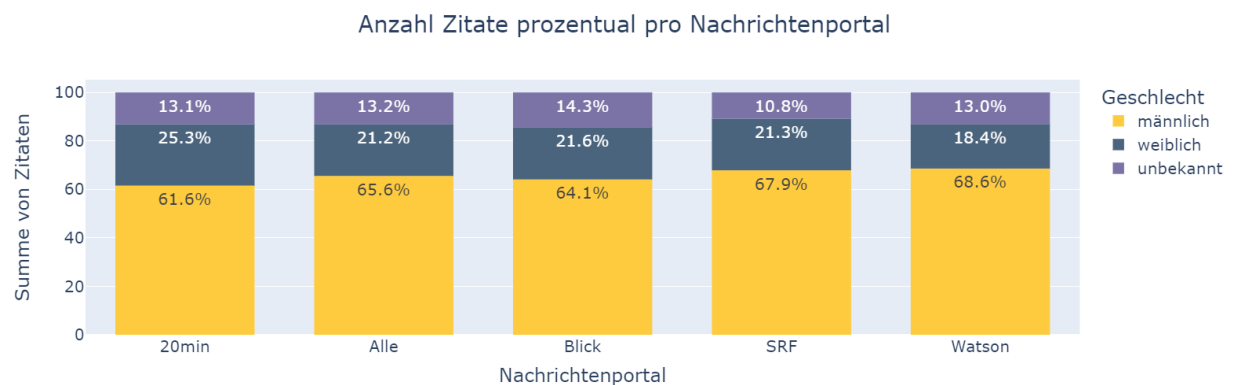


Abbildung 5.3: Verteilung der Anzahl gefundener Zitate pro Geschlecht und Nachrichtenportal

Die Abbildung 5.4 veranschaulicht das Verhältnis zwischen der Anzahl der gefundenen Zitate und der Anzahl Artikel pro Nachrichtenportal. Dabei wird deutlich, dass wir bei SRF die höchste Anzahl an Zitaten pro Artikel extrahieren konnten. Dieses Ergebnis weist auf interessante Aspekte hin. Es könnte bedeuten, dass SRF einen höheren Anteil an Artikeln mit zitierfähigem Inhalt hat oder dass sie eine spezifische Art der Berichterstattung bevorzugen, die häufiger zu zitierbarem Material führt.

Es ist wichtig anzumerken, dass wir nicht in jedem Artikel Zitate gefunden haben, weshalb die Anzahl der Zitate pro Artikel unter eins liegt.

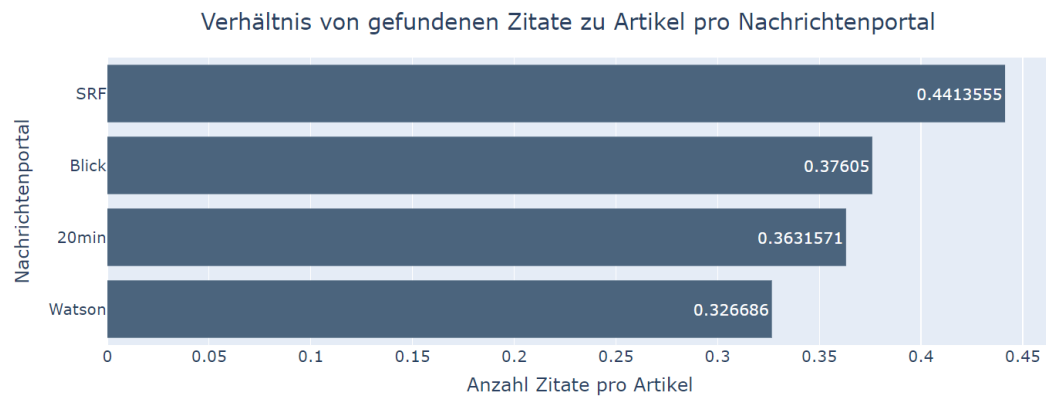


Abbildung 5.4: Anzahl gefundene Zitate pro Artikel pro Nachrichtenportal

Zum Schluss dieses Kapitel unten in Abbildung 5.5 und Tabelle 5.1 der jeweilige Gender Gap pro Portal.

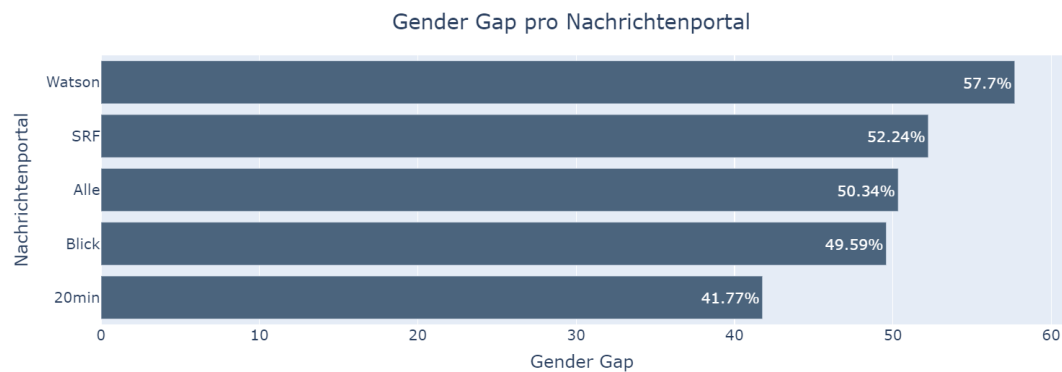


Abbildung 5.5: Gender Gap pro Nachrichtenportal

Nachrichtenportal	Frau	Mann	Gender Gap
20min	1882	4580	41.77%
Blick	15805	46981	49.59%
SRF	6888	21951	52.24%
Watson	3745	13979	57.70%
	28320	87491	50.34%

Tabelle 5.1: Gender Gap pro Nachrichtenportal

6 Diskussion

6.1 Zusammenfassung der Ergebnisse

Das Resultat der durchgeführten Untersuchung zeigt einen klaren und erstaunlich gleichmässigen *Gender Gap* in allen analysierten Nachrichtenportalen von ungefähr 41.77% bis 57.70%. Die Ergebnisse erfüllen die gestellte Aufgabenstellung, indem sie die Grösse ebenjener Gaps in den grössten gratis Schweizer Online-Nachrichtenportalen benennen. Die grosse Anzahl untersuchten Artikel erlaubt ein vorsichtiges Vertrauen in die Resultate. Dieses wird gestützt durch die Plausibilität der Resultate, wenn man sie mit Erkenntnissen aus anderen Studien vergleicht.

6.2 Interpretation der Resultate

Die resultierenden Ergebnisse der Untersuchung ergänzen das Bild, das von der aktuell Forschung gezeichnet wird (vgl. Kapitel 2). Der Gender Bias wurde bereits auf unterschiedlichste Arten gemessen: anhand der Anzahl Nennungen von Frauen und Männern [15], anhand von Zitaten [5], Quellenangaben [24] oder absolut vergleichbaren Paaren [50]. Stets scheint der gemessene Bias, der im Endeffekt den Unterschied im Raum, der den Geschlechtern in den Medien eingeräumt wird, misst, zwischen 20% und 60% zu liegen, je nach Studie.

Im Vergleich zum kanadischen *GGT* [5], der die Inspiration für diese Arbeit geliefert hat, sind die Ergebnisse ähnlich aber im Schnitt doch etwas schlechter. So lag der Gender Gap über alle Nachrichtenportale in der Arbeit „The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media“ bei 42%, bei unseren Resultaten jedoch bei 50.34%. Dies widerlegt die Vermutung, die wir in der Vorarbeit [4] auf Basis des *Global Gender Gap Report 2022* [22] getroffen haben. Wir sind davon ausgegangen, dass der Gender Gap in der Deutschschweiz etwas geringer ausfällt, als in Kanada, weil die Schweiz im internationalen Vergleich eine deutlich bessere Geschlechtergerechtigkeit kennt, als Kanada [22].

Die Studie „Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles“ [15] kommt in Bezug auf den *Gender Gap* je nach Datensatz auf sehr unterschiedliche Ergebnisse. So weisen die Artikel aus *MIND* einen Gap von 53.9% auf, gemessen an der Anzahl Nennungen von Männern und Frauen in Abstracts von Nachrichtenartikeln. *NCD* schneidet mit 14.5% unter Betrachtung dieser Metrik sehr gut ab. Doch zu beachten ist bei diesem Resultat die Verteilung der Artikel über die Kategorien. Denn in den Kategorien „Style & Beauty“, „Parenting“ und „Entertainment“ sind Frauen deutlich übervertreten.

Auch in der Studie „A Large-Scale Test of Gender Bias in the Media“ [50] wird ein starker Bias in der Anzahl Nennungen von Personen der unterschiedlichen Geschlechter gemessen. So liegt der *Gender Gap* im dritten Experiment, das Paare von Frauen und **Männer** im gleichen Alter, in derselben beruflichen Kategorie von gleichem öffentlichen Interesse vergleicht bei 46.71%. Eine grosse Zahl, jedoch immer noch deutlich kleiner als derjenige, den wir gemessen haben. So scheint der *Gender Gap* in der Deutschschweiz insgesamt grösser zu sein als im **Amerikanischen und Kanadischen** Raum.

Recherchen zu ähnlichen Projekten in der Schweiz ergeben jedoch ein hoffnungsvolles Bild. So misst der Ringier Verlag ~~seit~~ in seinem Projekt *EqualVoice* [1] die Repräsentation von Männern und Frauen in seinen Medienhäusern anhand des „Body Counts“, der Anzahl Nennungen von Männern und Frauen und hat im Jahr 2022 einen Unterschied von noch 34.95%. Wobei auch hier ein grosser Unterschied bei den unterschiedlichen Zeitungen feststellbar ist. So schneidet die Print Ausgabe von Blick mit 42% auch in dieser Auswertung ähnlich schlecht ab, wie in den vorliegenden Resultaten mit 49.59%. Es ist jedoch zu beachten, dass das Resultat von *EqualVoice* nur das Jahr 2022 betrifft und unsere Auswertung Artikel bis ins Jahr 2004 zurückreichen. Leider ist auf der Webseite auch zu erkennen, dass sich der *Gender Gap* seit dem Ende der Coronapandemie mit dem zusätzlichen Fokus auf den Gesundheitsbereich wieder weitet.

Es ist jedoch höchst erfreulich, dass sich die Verlage zunehmend Gedanken zu diesem Thema machen und Lösungen anstreben. So hat uns SRF bei unserer Anfrage nach einer Stellungnahme darauf hingewiesen, dass sie beispielsweise in der beliebten Schweizer Radiosendung „Echo der Zeit“ ¹ im Februar erstmals mehr Gesprächspartnerinnen als Gesprächspartner in den Echo-Gesprächen hatten ².

Es scheint als wären sich zumindest SRF und Ringier (Blick Mutterkonzern) der Problematik bewusst zu sein und Aufwand zur Verringerung des *Gender Gap* zu betreiben. Trotzdem scheint die Deutschschweiz auf diesem Gebiet im Vergleich mit den **Amerikanischen** und **Kanadischen** Nachrichtenportalen hinterherzuhinken und noch viel Arbeit betreiben zu müssen, um den Unterschied weiter zu verkleinern.

Unter dieser Betrachtung stellen sich auch die Fragen, was die Ursachen für dieses Problem sind. So zeigt beispielsweise die Arbeit „A Large-Scale Test of Gender Bias in the Media“ [50] schön auf, dass dieser Bias nicht alleine durch eine Präferenz der Medienschaffenden zur Porträtierung von Männern zustande kommt, sondern auch der Struktur unserer Gesellschaft geschuldet ist, in der ein Grossteil der Nachrichtenwerten Positionen von Männern belegt wird. Dies ist nicht zuletzt dem Phänomen des *Glass Ceiling* geschuldet, wie verschiedene Studien zeigen [13, 21].

Die Bereinigung dieser Ungerechtigkeit kann also nicht von den Medienschaffenden alleine gelöst werden, denn ihre Aufgabe in unserer Gesellschaft besteht in **Erster** Linie in der Informationsverbreitung und der Schaffung von Transparenz in den für die Öffentlichkeit interessanten Bereichen. Trotzdem können sie durch bewusste und reflektierte Berichterstattungen und Recherchen ihren Anteil am Problem verringern und so die Gesamtsituation verbessern. Die eigentliche Herkulesaufgabe, die Transformation unserer

¹<https://www.srf.ch/audio/echo-der-zeit>

²Dieser Beitrag konnte öffentlich nicht gefunden werden.

Gesellschaft zu einer ausgeglicheneren und faireren Zivilisation wird jedoch noch viel Effort von der Politik, Wirtschaft und jedem und jeder einzelnen von uns erfordern.

6.3 Limitationen

Obwohl die grosse Datenmenge den Resultaten eine gewisse Zuverlässigkeit garantiert, ist bei deren Betrachtung immer noch Vorsicht geboten. Der Grund dafür liegt in der grundsätzlichen Schwierigkeit, natürlichen Text zuverlässig automatisiert auswerten zu können. Die verwendeten Werkzeuge dazu erkennen Strukturen und Elemente stets mit einer gewissen Unsicherheit. Auf Basis dieser **Im**perfekten Daten wendet die Auswertung eigene Abläufe an, die ihrerseits nicht mit absoluter Sicherheit ein korrektes Resultat liefern müssen. Auch wenn diese Unsicherheiten in den einzelnen Schritten jeweils klein wären - oder zum Teil auch sind - so können sie sich trotzdem gegenseitig aufsummieren und das Ergebnis als Ganzes ungenauer machen als vermutet. So gibt *Spacy* für das von uns verwendete Model „Accuracy Scores“ von > 90% an³. Für das POS Tagging wird eine Präzision von 98% versprochen und für das gelabelte Dependency Parsing eine von 91%. Diese Unsicherheiten sollten nicht vergessen werden, doch die grösste Unsicherheit entsteht wohl durch die Extraktion der Zitate.

Aufgrund des zeitlichen Rahmens konnte in dieser Arbeit nur die Extraktion der gängigsten Formen von **Syntaktischen** Zitaten implementiert werden. Die **Schwimmenden** Zitate konnten hingegen gar nicht berücksichtigt werden. Auch komplexere Formen **Syntaktischen** Zitate erfasst der Algorithmus aufgrund anderer Strukturen im resultierenden Parse Tree nicht. Das Programm erfasst laut unseren Tests etwa 60% der **Syntaktischen** Zitate. Ohne genauere Angaben zu der Verteilung der Anzahl Zitate auf deren unterschiedlichen Kategorien zu haben, schätzen wir den Anteil der **Syntaktischen** Zitate auf etwa 50%. Denn schwimmende Zitate kommen unbegleitet von **Syntaktischen** Zitaten eigentlich nicht vor, doch wenn sie vorkommen, treten sie meist in grösseren Mengen auf (bsp. längeres Interview). Unter der Annahme, dass die verwendeten Tests ein repräsentatives Bild der Performance des Algorithmus abgeben und dass die oben beschriebenen ~~en~~ Schätzung korrekt ist, hat das Programm also 60% von 50%, also insgesamt 30%, aller Zitate erkannt.

Auch bei der Erkennung der Personen herrscht eine gewisse Unsicherheit ~~gegeben~~ durch die probabilistische Natur des *Spacy* Models. So verkündet die offizielle Webseite zwar eine Accuracy von 98%, doch nach unserer Erfahrung ist zumindest die Anzahl von False-Positives deutlich höher als diese Zahl würde vermuten lassen. Wir konnten beobachten, dass die Entitäten zwar richtig erkannt wurden, jedoch unter einem falschen Label. Meist waren dies Organisationen, die *Spacy* als Personen identifiziert hatte. Leider konnten wir nicht herausfinden, welche Metrik *Spacy* für die Messung der Accuracy verwendet. Aufgrund der subjektiven Erfahrung gehen wir aber davon aus, dass es sich wohl um eine Metrik handeln muss, die False-Positives nicht oder nur schwach berücksichtigt. So könnte es sich bspw. um den *Recall* handeln.

³https://spacy.io/models/de#de_core_news_lg-accuracy

6.4 Weiterführende Forschung

Die Resultate belegen, dass in den untersuchten Nachrichtenportalen ein grosser Gender Gap herrscht. Der zeitliche Rahmen hat jedoch weiterführende Untersuchungen verhindert. Aus dieser Erkenntnis ergeben sich Fragen zur Ursache und Beschaffenheit dieses Gaps.

So wäre es spannend zu wissen, in welchen Kategorien von Nachrichtenartikeln dieser Unterschied wie fest ausgeprägt ist. Diese Erkenntnisse könnte man in Vergleich setzen mit der Arbeit von Shor, Rijt und Fotouhi „A Large-Scale Test of Gender Bias in the Media“ [50] oder dem Paper von Dacon und Liu „Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles“. Des Weiteren wäre von Interesse, welcher Anteil der strukturelle Bias in unserer Gesellschaft an diesem Gap hat und welcher Anteil dem Bias in der Berichterstattung von Medienschaffenden geschuldet ist. Dazu könnten interessierte Forschende eine ähnliche Methodik wie Shor, Rijt und Fotouhi in „A Large-Scale Test of Gender Bias in the Media“ [50] anwenden.

Eher moralischer Natur ist die Forschungsfrage, ob dieser Gap überhaupt etwas Schlimmes ist. Die Antwort zu dieser Frage liegt wohl in der Meinung der von dieser Tatsache diskriminierten Gruppe, der Frauen. Hierzu könnte eine Meinungsumfrage Aufschluss bieten.

Interessant wäre auch zu wissen, welche Auswirkungen sich aus diesem Gap ergeben. Es ist unbestritten, dass Medien die Meinung einer Gesellschaft prägen. So kann wohl auch angenommen werden, dass die Unterrepräsentation der Frauen in den Medien Nachteile für diese mit sich bringt. Wie diese Nachteile aussehen, die sich aufgrund dieser Unterrepräsentation ergeben, lässt sich aufgrund des aktuellen Forschungsstands nicht sagen. Obwohl die strukturellen Benachteiligungen der Frauen gegenüber den Männern klar benennbar sind, ist es schwierig diese auf einzelne Ursachen zurückzuführen. Weiterführende Forschung könnte hier Klarheit schaffen.

7 Schlussfolgerungen

7.1 Ziele evaluieren

In diesem Teil der Arbeit gehen wir auf die Ziele ein, welche wir am Anfang des Projekts definiert haben.

7.1.1 Extraktion Zitate

Ein Hauptziel dieser Arbeit bestand darin, Zitate aus Texten zu extrahieren. Die vorliegende Auswertung konzentriert sich auf die Anzahl der Zitate von Männern und Frauen pro Nachrichtenportal. Dabei lag der Fokus auf der Erkennung und Extraktion sogenannter **Syntaktischer Zitate**. Der Grund dafür ist, dass sie die klarste strukturelle Form aufweisen und auch als Grundlage für die Erkennung anderer Arten von Zitaten dienen.

Es ist wichtig zu beachten, dass aufgrund des zeitlichen Rahmens dieser Arbeit weitere Arten von Zitaten nicht berücksichtigt werden konnten. Dazu gehören **Schwimmende Zitate**, **Heuristische Zitate** und **Syntaktische Zitate**, die ein zu unspezifisches oder gar kein Verb enthalten.

Die Fokussierung auf syntaktische Zitate ermöglichte es uns, eine solide Grundlage für die Untersuchung des Gender Gaps in Bezug auf Zitate zu schaffen. Indem wir uns auf diese klar definierte Art von Zitaten konzentrierten, konnten wir eine präzise Analyse durchführen und spezifische Erkenntnisse über die Verteilung von Zitaten zwischen Männern und Frauen gewinnen.

Es ist anzumerken, dass die Nicht-Berücksichtigung anderer Zitatarten nicht bedeutet, dass sie in Zukunft keine Rolle spielen könnten. In weiteren Untersuchungen könnten diese zusätzlichen Zitatarten genauer analysiert werden, um ein umfassenderes Verständnis der Zitatlandschaft und möglicher geschlechtsspezifischer Unterschiede zu erlangen.

Wir denken, dass wir 60% der **Syntaktischen Zitate** extrahieren konnten aufgrund der Tests, welche wir gemacht haben. Wir gehen davon aus, dass diese Art etwa die Hälfte aller Zitate ausmacht. Insgesamt rechnen wir also damit, dass wir 30% der Zitate analysiert haben. Mehr zu diesen Tests im Kapitel 4.5. Diese Recall-Rate von 60% beim Herauslesen von Zitaten ist enttäuschend tief. Es war nicht immer möglich, alle Zitate zuverlässig zu identifizieren und zuzuordnen. Dies erfordert weitere Optimierungen und Verfeinerungen des Algorithmus, um die Genauigkeit zu verbessern. Leider war dies aber zeitbedingt nicht möglich.

7.1.2 Extraktion Personen

Die Kombination von *NER* und *Coreference Resolution* ermöglichte eine weitgehende Extraktion und Identifizierung der im Text genannten Personen. Durch den Einsatz dieser beiden Verfahren konnten die Personen grundsätzlich erfolgreich aus dem Text herausgefiltert werden.

Dennoch ist es wichtig zu betonen, dass die Extraktion von Personen aus Texten keine absolute Gewährleistung für eine fehlerfreie Identifizierung bietet. Es besteht immer die Möglichkeit von Fehlern, insbesondere wenn die Namen mehrdeutig sind oder wenn im Text Unklarheiten auftreten oder Verweise auf nicht eindeutige Personen gemacht werden.

7.1.3 Evaluierung Geschlecht

Das Geschlecht konnte bei 87,2% aller gefundenen Personen, denen ein Zitat zugewiesen wurde, erfolgreich bestimmt werden. Die restlichen 12,8% der Personen, bei denen das Geschlecht nicht ermittelt werden konnte, sind entweder vermutlich keine tatsächlichen Personen, da sie fälschlicherweise vom *NER* als solche erkannt wurden, oder es handelt sich um Namen, denen kein eindeutiges männliches oder weibliches Geschlecht zugeordnet werden können.

Wir hatten auch die Idee, das Geschlecht der Autor:innen der Artikel zu bestimmen. Leider war die Datenqualität der vorliegenden Datengrundlage aus dem Vorgängerprojekt unzureichend. Oftmals waren im Feld „author“ der Datenbank keine Namen, sondern andere Informationen vermerkt, mit denen wir nichts anfangen konnten. Wir vermuten, dass Teile des Artikels als Autor:in erkannt wurden, aufgrund der Tatsache, dass die Autor:innen Informationen häufig am Ende des Texts in einer nicht eindeutigen Form aufgeführt wurden.

7.1.4 Aggregation und Visualisierung

Die Artikel konnten mit unserem Algorithmus erfolgreich analysiert und aggregiert werden. Die visualisierten Ergebnisse der Aggregation sind im Kapitel 5 ersichtlich.

7.2 Persönliche Erkenntnisse

7.2.1 Literaturrecherche

Zu Beginn der ersten Phase unserer Arbeit, der Literaturrecherche, hatten wir das Glück, dass Frau Kurpicz uns bei der Definition unserer Forschungsmethode unterstützte. Sie half uns eine geeignete Methode für die Literaturrecherche zu entwickeln. Ihre Anleitung und Erfahrung waren sehr hilfreich und führten uns in eine richtige Richtung.

Während der Literaturrecherche hatten wir mit einer beträchtlichen Menge an wissenschaftlicher Literatur zu tun. Das Lesen und Zusammenfassen von zahlreichen Papers wurde mit der Zeit anstrengend und ermüdend. Es war eine Herausforderung, die Konzentration

on aufrechtzuerhalten und gleichzeitig den Überblick über die wichtigsten Informationen zu behalten.

Das Analysieren der gefundenen Literatur stellte eine neue und herausfordernde Aufgabe dar. Wir waren ungeübt in diesem Bereich und mussten uns mit den grundlegenden Techniken des Literatur-Analysierens vertraut machen. Es erforderte Zeit und Mühe, die verschiedenen Ansätze zur Bewertung und Zusammenfassung von Forschungsergebnissen zu erlernen. Es gab Momente, in denen es frustrierend war. Mit der Zeit und Übung konnten wir unsere Fähigkeiten verbessern. Wir haben gelernt, dass Ausdauer und Organisation essenziell sind, um mit der Fülle an Informationen umzugehen.

7.2.2 Programmieren

Nach der Literaturrecherche haben wir festgestellt, dass es für uns wichtig ist, Spass am Programmieren zu haben und nicht ausschliesslich mit dem Lesen und Dokumentieren von Papers beschäftigt zu sein. Umso mehr hatten wir Freude aktiv zu programmieren und Lösungen zu entwickeln, die einen Mehrwert bieten für unsere Arbeit. So konnten wir unsere Motivation und Kreativität aufrechterhalten, weil in der letzten Phase der Fokus auf dem Dokumentieren liegt.

Des Weiteren haben wir festgestellt, dass die Bearbeitung grosser Datenmengen mehr Zeit in Anspruch nehmen kann als ursprünglich geplant. Insbesondere die Datenbereinigung erforderte viele Stunden bis alle Daten verarbeitet waren. Oftmals zwischen 20 und 30 Stunden pro Nachrichtenportal. Teilweise sogar mehrmals, da wir Fehler im Script bemerkten und diese nochmals neu starten mussten. Bei der Aggregation waren auch 40 Stunden erforderlich, um alle Daten korrekt zu verarbeiten. Für zukünftige Projekte würden wir dies bei der Zeitplanung mehr berücksichtigen und ausreichend Zeit für die Datenbereinigung und Aggregation einplanen.

Insgesamt betrachtet war das **s programmieren** mit Python für uns der beste Teil der Arbeit, da wir immer wieder neue Herausforderungen und Lernmöglichkeiten entdeckten.

7.2.3 LaTeX

Unsere Vorkenntnisse in LaTeX aus Project 2 waren äusserst hilfreich. Sie haben uns ermöglicht, effizient mit unserem LaTeX-Dokument umzugehen und ansprechende Ergebnisse zu erzielen.

Dank unserer vorhandenen LaTeX-Kenntnisse konnten wir schnell und präzise neue Kapitel erstellen, Formatierungen anpassen und Inhalte strukturieren. Wir waren mit den grundlegenden Befehlen und Konzepten vertraut, was uns geholfen hat, uns auf den Inhalt unseres Dokuments zu konzentrieren, anstatt viel Zeit mit dem Erlernen der grundlegenden Syntax zu verbringen wie im Vorprojekt.

7.2.4 MongoDB

Auch mit MongoDB hatten wir hilfreiche Erfahrungen aus dem Vorprojekt. Die integrierte Abfragesprache „MongoDB Query Language“ war uns auch schon bekannt und so konnten

wir viel wirksamer Abfragen machen und auch die Komplexität der Abfragen steigern.

7.2.5 Projektplanung

Im Vergleich zum Vorprojekt war es herausfordernd, die Planung im Voraus so detailliert durchzuführen. Im Vorprojekt hatten wir hauptsächlich auf Wochenbasis geplant, während der Bachelor Thesis ist ein umfassender Projektplan Teil der Bewertung. Das Festlegen von konkreten Aufgaben und Meilensteinen auf langfristiger Basis erforderte eine sorgfältige Abwägung und eine genaue Einschätzung der Aufwände.

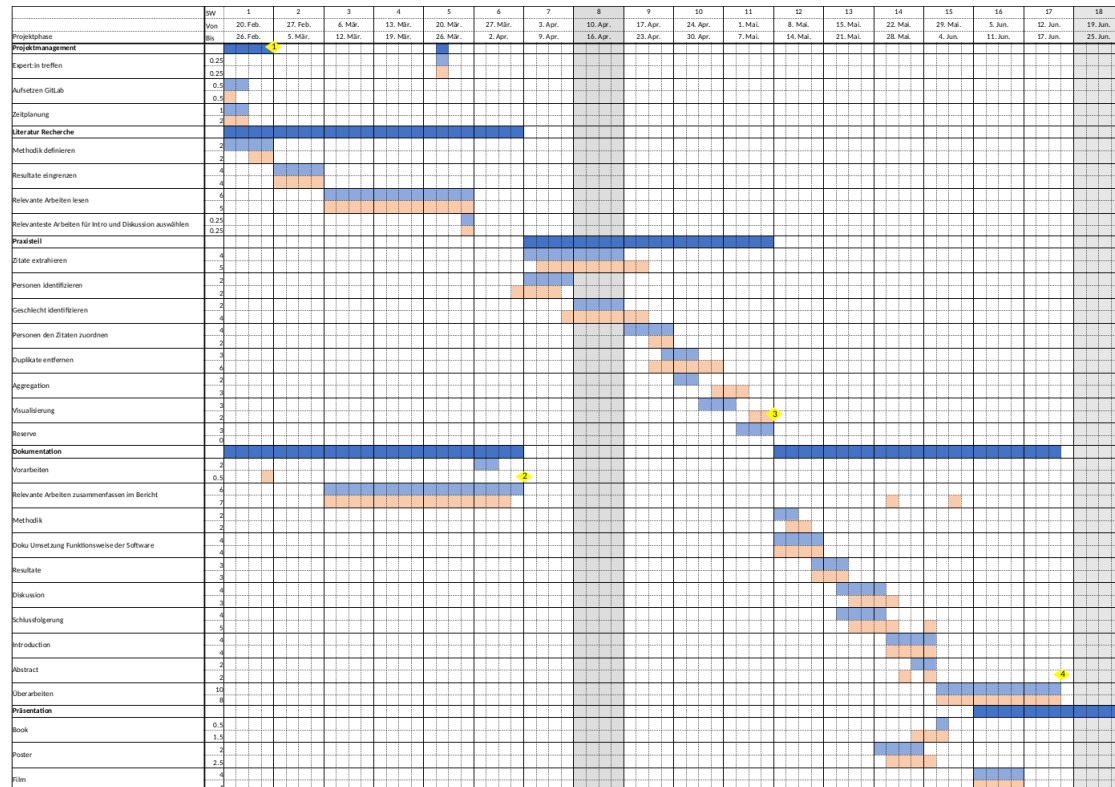


Abbildung 7.1: Zeitplan ausgefüllt

Trotz der Komplexität der Projektplanung konnten wir den Zeitplan im Grossen und Ganzen einhalten. Dies war ein Erfolg, da es uns ermöglichte, die Fortschritte des Projekts zu steuern und sicherzustellen, dass wir die gesetzten Ziele erreichen können. Die Disziplin und Organisation aller Teammitglieder trugen dazu bei, dass wir den Überblick über die Arbeitsauslastung behielten und unsere Deadlines einhielten. Die Verwendung eines Kanban Boards für jede Projektphase erwies sich als äusserst hilfreich. Es ermöglichte uns den Fortschritt der Aufgaben zu überblicken. Das Kanban Board ermöglichte uns, Aufgaben zu visualisieren, den Status jedes Tasks zu verfolgen und Engpässe oder Verzögerungen frühzeitig zu erkennen. Dadurch konnten wir flexibel auf Änderungen reagieren und

die Prioritäten entsprechend anpassen.

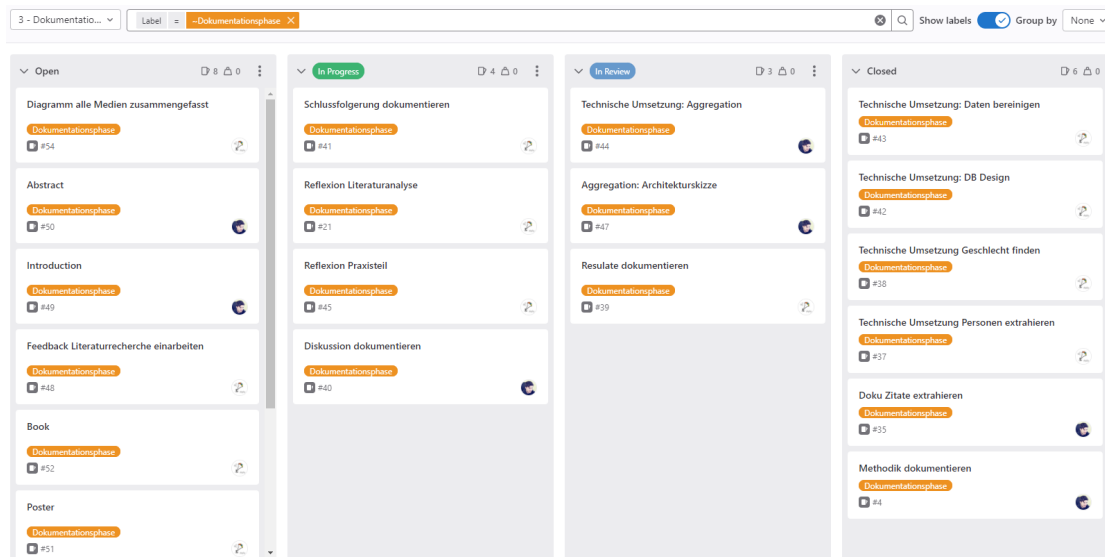


Abbildung 7.2: Kanban Board der Dokumentationsphase

In unserer Projektplanung war die Datenbereinigung nicht als einzelner Task vorgesehen. Es stellte sich jedoch heraus, dass dieser Task erforderlich war, um mit den gesammelten Daten arbeiten zu können. Das Hinzufügen dieses zusätzlichen Tasks brachte eine gewisse Herausforderung mit sich, da wir unsere Zeitplanung und Ressourcen neu anpassen mussten, um diese Aufgabe bewältigen zu können.

Das Schreiben von Besprechungsprotokollen erwies sich als wertvoll für die Nachvollziehbarkeit von Entscheidungen und die Umsetzung von Tasks. Die Protokolle von den Meetings mit Frau Kurpicz sowie von den Besprechungen mit uns beiden halfen uns, wichtige Informationen und Diskussionsergebnisse festzuhalten. Ausserdem war es hilfreich sicherzustellen, dass alle Teammitglieder über den Stand der Dinge informiert waren und die vereinbarten Aufgaben richtig ausgeführt wurden.

7.2.6 Zusammenarbeit in der Gruppe

Bei unserer Selbstreflexion zur Zusammenarbeit in der Gruppe haben wir folgende wichtige Erkenntnisse gewonnen:

- ▶ **Eingespieltes Teamwork durch vorherige Projekte:** Dank unserer vorherigen Zusammenarbeit in anderen Projekten konnten wir als Team gut zusammenarbeiten. Wir kannten bereits unsere Stärken und Schwächen und wussten, wie wir effektiv kommunizieren und unsere Aufgaben koordinieren können.
- ▶ **Schnelles Feedback von Frau Kurpicz und gute Kommunikation:** Das rasche Feedback von Frau Kurpicz auf unsere Mails hat uns sehr geholfen. Die Kommunikation

mit ihr war reibungslos und wir konnten jederzeit Fragen stellen und Unterstützung erhalten und haben diese Möglichkeit auch wo nötig in Anspruch genommen.

- ▶ **Hit The Ground Running:** Wir waren von Anfang an äusserst engagiert und haben sofort mit der Arbeit begonnen. Dadurch konnten wir effizient voranschreiten und die Projektziele frühzeitig erreichen.
- ▶ **Gute Aufteilung der Arbeit in Arbeitspakete:** Durch die sorgfältige Aufteilung der Arbeit in Arbeitspakete konnten wir parallel an verschiedenen Aufgaben arbeiten. Dies erhöhte unser Arbeitstempo und ermöglichte uns, schneller voranzukommen.
- ▶ **Nützliche gegenseitige Reviews:** Das gegenseitige Review unserer Arbeit war sehr sinnvoll. Durch konstruktives Feedback konnten wir unsere Ergebnisse verbessern und sicherstellen, dass sie unseren Anforderungen entsprachen.
- ▶ **Umgang mit Schwierigkeiten im Team:** Wir haben Schwierigkeiten im Team offen angesprochen und gemeinsam Lösungen gefunden. Ein Beispiel dafür war die Entscheidung, nur syntaktische Zitate zu extrahieren. Durch solche Diskussionen konnten wir unsere Arbeitsweise optimieren und bessere Ergebnisse erzielen.
- ▶ **Veränderung klassischer Geschlechterrollen:** Wir haben klassische Geschlechterrollen in der Projektzusammenarbeit identifiziert und aktiv verändert. Wir haben uns abgewechselt bei Präsentationen und der Leitung von Meetings, um sicherzustellen, dass alle Mitglieder des Teams gleiche Chancen haben, ihre Fähigkeiten zu entwickeln und zu zeigen.
- ▶ **Effiziente und zeitlich gut geplante Meetings mit Frau Kurpicz:** Die Meetings mit Frau Kurpicz verliefen immer recht speditiv und innerhalb des festgelegten Zeitrahmens. Wir hatten regelmässige Austausche, bei denen wir den Fortschritt präsentierten, die weiteren Schritte besprachen und Fragen diskutierten.
- ▶ **Lösung für Terminfindung:** Die spontane Terminfindung hat nicht immer reibungslos funktioniert in unserem 2er Team. Als Lösung haben wir uns auf wöchentliche Termine ab der Praxisteil-Phase geeinigt, um sicherzustellen, dass wir uns regelmässig treffen. Um den Fortschritt unserer Arbeiten kontinuierlich besprechen und uns über Probleme austauschen zu können.

Insgesamt war die Zusammenarbeit in der Gruppe sehr positiv. Unsere vorherige Erfahrung als Team, effektive Kommunikation, gegenseitiges Feedback und die Fähigkeit, Schwierigkeiten anzusprechen und Lösungen zu finden, haben dazu beigetragen, dass wir so gut zusammen arbeiten konnten. Die identifizierten Verbesserungen in der Geschlechterrolle beim Meeting führen und die regelmässige Terminfindung haben ebenfalls zu einer positiven Teamdynamik beigetragen.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die hier vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Sämtliche Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, habe ich als solche kenntlich gemacht.

Hiermit stimme ich zu, dass die vorliegende Arbeit in elektronischer Form mit entsprechender Software überprüft wird.

1. Juni 2023



A. Aeschbacher



F. Neidhart

Literatur

- [1] Ringier AG. *EqualVoice-Factor Halbjahresbilanz 2022: Neu algorithmische Auswertung auch bei Print-Artikeln möglich*. URL: <https://www.ringier.com/de/equalvoice-factor-halbjahresbilanz-2022-neu-algorithmische-auswertung-auch-bei-print-artikeln-moeglich/>. (accessed: 03.01.2023).
- [2] Adolfo Alonso-Arroyo u. a. „Gender inequalities on editorial boards of indexed pediatrics journals“. In: *Pediatric Research* 90.2 (2021), S. 300–314.
- [3] Amanda Amberg und Darren N. Saunders. „Cancer in the news: Bias and quality in media reporting of cancer research“. In: *PLOS ONE* 15.11 (Nov. 2020), S. 1–13. DOI: 10.1371/journal.pone.0242133. URL: <https://doi.org/10.1371/journal.pone.0242133>.
- [4] Felix Neidhart Anina Aeschbacher. „Erstellung einer Datensammlung von Deutschschweizer Online Zeitungsartikeln“. 2023.
- [5] Fatemeh Torabi Asr u. a. „The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media“. In: *PLOS ONE* 16.1 (Jan. 2021), S. 1–28. DOI: 10.1371/journal.pone.0245533. URL: <https://doi.org/10.1371/journal.pone.0245533>.
- [6] Fatemeh Torabi Asr u. a. *The gender gap tracker: Using natural language processing to measure gender bias in media*. 2021. URL: <https://gendergaptracker.informedopinions.org/>. (accessed: 26.09.2022).
- [7] Stan Benjamens u. a. „Gender disparities in transplantation research“. In: *Transplantation* 104.S3 (2020), S25.
- [8] Diana Burgess und Eugene Borgida. „Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination.“ In: *Psychology, public policy, and law* 5.3 (1999), S. 665. URL: <https://doi.org/10.1037/1076-8971.5.3.665>.
- [9] James P Byrnes, David C Miller und William D Schafer. „Gender differences in risk taking: A meta-analysis.“ In: *Psychological bulletin* 125.3 (1999), S. 367. URL: <https://doi.org/10.1037/0033-2909.125.3.367>.
- [10] Neven Caplar, Sandro Tacchella und Simon Birrer. „Quantitative evaluation of gender bias in astronomical publications from citation counts“. In: *Nature Astronomy* 26 (Mai 2017). URL: <https://doi.org/10.1038/s41550-017-0141>.

- [11] Carolina Castagnetti und Luisa Rosti. „Unfair Tournaments: Gender Stereotyping and Wage Discrimination among Italian Graduates“. In: *Gender & Society* 27.5 (2013), S. 630–658. DOI: 10.1177/0891243213490231. URL: <https://doi.org/10.1177/0891243213490231>.
- [12] Andrew J Chadwick und Rosaleen Baruah. „Gender disparity and implicit gender bias amongst doctors in intensive care medicine: a ”disease” we need to recognise and treat“. In: *Journal of the Intensive Care Society* 21.1 (2020), S. 12–17.
- [13] David A. Cotter u. a. „The Glass Ceiling Effect“. In: *Social Forces* 80.2 (Dez. 2001), S. 655–681. ISSN: 0037-7732. DOI: 10.1353/sof.2001.0091. eprint: <https://academic.oup.com/sf/article-pdf/80/2/655/6519837/80-2-655.pdf>. URL: <https://doi.org/10.1353/sof.2001.0091>.
- [14] Rachel Croson und Nancy Buchan. „Gender and Culture: International Experimental Evidence from Trust Games“. In: *American Economic Review* 89.2 (Mai 1999), S. 386–391. DOI: 10.1257/aer.89.2.386. URL: <https://www.aeaweb.org/articles?id=10.1257/aer.89.2.386>.
- [15] Jamell Dacon und Haochen Liu. „Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles“. In: *Companion Proceedings of the Web Conference 2021*. WWW ’21. Ljubljana, Slovenia, 2021, S. 385–392. ISBN: 9781450383134. DOI: 10.1145/3442442.3452325. URL: <https://doi.org/10.1145/3442442.3452325>.
- [16] Ronit Dinovitzer, Nancy Reichman und Joyce Sterling. „The Differential Valuation of Women’s Work: A New Look at the Gender Gap in Lawyers’ Incomes“. In: *Social Forces* 88.2 (Dez. 2009), S. 819–864. ISSN: 0037-7732. DOI: 10.1353/sof.0.0260. URL: <https://doi.org/10.1353/sof.0.0260>.
- [17] Shiri Dori-Hacohen und Scott A. Hale. „Information Ecosystem Threats in Minoritized Communities: Challenges, Open Problems and Research Directions“. In: *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR ’22. Madrid, Spain: Association for Computing Machinery, 2022, S. 3384–3385. ISBN: 9781450387323. DOI: 10.1145/3477495.3536327. URL: <https://doi.org/10.1145/3477495.3536327>.
- [18] Alice H Eagly und Steven J Karau. „Role congruity theory of prejudice toward female leaders.“ In: *Psychological review* 109.3 (2002), S. 573. URL: <https://doi.org/10.1037/0033-295X.109.3.573>.
- [19] Paula England. „Gender Inequality in Labor Markets: The Role of Motherhood and Segregation“. In: *Social Politics: International Studies in Gender, State & Society* 12.2 (Juli 2005), S. 264–288. ISSN: 1072-4745. DOI: 10.1093/sp/jxi014. eprint: <https://academic.oup.com/sp/article-pdf/12/2/264/4721557/jxi014.pdf>. URL: <https://doi.org/10.1093/sp/jxi014>.

- [20] Paula England. „The Gender Revolution: Uneven and Stalled“. In: *Gender & Society* 24.2 (2010), S. 149–166. DOI: 10.1177/0891243210361475. eprint: <https://doi.org/10.1177/0891243210361475>. URL: <https://doi.org/10.1177/0891243210361475>.
- [21] Olle Folke und Johanna Rickne. „The Glass Ceiling in Politics: Formalization and Empirical Tests“. In: *Comparative Political Studies* 49.5 (2016), S. 567–599. DOI: 10.1177/0010414015621073. eprint: <https://doi.org/10.1177/0010414015621073>. URL: <https://doi.org/10.1177/0010414015621073>.
- [22] World Economic Forum. *Global Gender Gap Report 2022*. URL: https://www3.weforum.org/docs/WEF_GGGR_2022.pdf. (accessed: 31.12.2022).
- [23] Christine R Harris und Michael Jenkins. „Gender differences in risk assessment: why do women take fewer risks than men?“ In: *Judgment and Decision making* 1.1 (2006), S. 48–63. URL: <https://doi.org/10.1037/0022-3514.74.3.629>.
- [24] Alessia Iancarelli u. a. „Using citation network analysis to enhance scholarship in psychological science: A case study of the human aggression literature“. In: *PLOS ONE* 17.4 (Apr. 2022), S. 1–21. DOI: 10.1371/journal.pone.0266513. URL: <https://doi.org/10.1371/journal.pone.0266513>.
- [25] Sen Jia u. a. „Women Are Seen More than Heard in Online Newspapers“. In: *PLOS ONE* 11.2 (Feb. 2016), S. 1–11. DOI: 10.1371/journal.pone.0148434. URL: <https://doi.org/10.1371/journal.pone.0148434>.
- [26] Fiona M. Kay und John Hagan. „The Persistent Glass Ceiling: Gendered Inequalities in the Earnings of Lawyers“. In: *The British Journal of Sociology* 46.2 (1995), S. 279–310. ISSN: 00071315, 14684446. URL: <http://www.jstor.org/stable/591789> (besucht am 16. 03. 2023).
- [27] Silvia Knobloch-Westerwick, Carroll J Glynn und Michael Huge. „The Matilda effect in science communication: an experiment on gender bias in publication quality perceptions and collaboration interest“. In: *Science communication* 35.5 (2013), S. 603–625.
- [28] Philipp W Kramer u. a. „Sex disparities in ophthalmic research: a descriptive bibliometric study on scientific authorships“. In: *JAMA ophthalmology* 137.11 (2019), S. 1223–1231.
- [29] John E. Krstacic u. a. „Academic medicine’s glass ceiling: Author’s gender in top three medical research journals impacts probability of future publication success“. In: *PLOS ONE* 17.4 (Apr. 2022), S. 1–21. DOI: 10.1371/journal.pone.0261209. URL: <https://doi.org/10.1371/journal.pone.0261209>.
- [30] Michèle Lamont. „Toward a Comparative Sociology of Valuation and Evaluation“. In: *Annual Review of Sociology* 38.1 (2012), S. 201–221. DOI: 10.1146/annurev-soc-070308-120022. URL: <https://doi.org/10.1146/annurev-soc-070308-120022>.

- [31] Levon Lloyd, Dimitrios Kechagias und Steven Skiena. „Lydia: A System for Large-Scale News Analysis“. In: *Proceedings of the 12th International Conference on String Processing and Information Retrieval*. SPIRE’05. Buenos Aires, Argentina: Springer-Verlag, 2005, S. 161–166. ISBN: 3540297405. DOI: 10.1007/11575832_18. URL: https://doi.org/10.1007/11575832_18.
- [32] Leah M Marcotte, Vineet M Arora und Ishani Ganguli. „Toward gender equity in academic promotions“. In: *JAMA Internal Medicine* 181.9 (2021), S. 1155–1156.
- [33] Abigail A Marsh. „The caring continuum: Evolved hormonal and proximal mechanisms explain prosocial and antisocial extremes“. In: *Annual Review of Psychology* 70 (2019), S. 347–371. URL: <https://doi.org/10.1146/annurev-psych-010418-103010>.
- [34] Heather McLaughlin, Christopher Uggen und Amy Blackstone. „Sexual harassment, workplace authority, and the paradox of power“. In: *American sociological review* 77.4 (2012), S. 625–647. URL: <https://doi.org/10.1177/0003122412451728>.
- [35] Leoni C Menzel u. a. „Gender disparities in authorships of Alzheimer’s disease and dementia research articles“. In: *Journal of Alzheimer’s Disease* 70.4 (2019), S. 1143–1152.
- [36] Rishabh Misra. „News Category Dataset“. In: *arXiv preprint arXiv:2209.11429* (2022).
- [37] Natasha Quadlin. „The Mark of a Woman’s Record: Gender and Academic Performance in Hiring“. In: *American Sociological Review* 83.2 (2018), S. 331–360. DOI: 10.1177/0003122418762291. URL: <https://doi.org/10.1177/0003122418762291>.
- [38] Cecilia L. Ridgeway. „Why Status Matters for Inequality“. In: *American Sociological Review* 79.1 (2014), S. 1–16. DOI: 10.1177/0003122413515997. URL: <https://doi.org/10.1177/0003122413515997>.
- [39] Cecilia L. Ridgeway und Shelley J. Correll. „Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations“. In: *Gender & Society* 18.4 (2004), S. 510–531. DOI: 10.1177/0891243204265269. URL: <https://doi.org/10.1177/0891243204265269>.
- [40] Cecilia L. Ridgeway u. a. „How Easily Does a Social Difference Become a Status Distinction? Gender Matters“. In: *American Sociological Review* 74.1 (2009), S. 44–62. DOI: 10.1177/000312240907400103. URL: <https://doi.org/10.1177/000312240907400103>.
- [41] Shelly Rodgers und Esther Thorson. „A Socialization Perspective on Male and Female Reporting“. In: *Journal of Communication* 53.4 (Jan. 2003), S. 658–675. ISSN: 0021-9916. DOI: 10.1111/j.1460-2466.2003.tb02916.x. URL: <https://doi.org/10.1111/j.1460-2466.2003.tb02916.x>.
- [42] Karen Ross und Cynthia Carter. „Women and news: A long and winding road“. In: *Media, Culture & Society* 33.8 (2011), S. 1148–1165. DOI: 10.1177/0163443711418272. URL: <https://doi.org/10.1177/0163443711418272>.

- [43] Karen Ross und Cynthia Carter. „Women and news: A long and winding road“. In: *Media, Culture & Society* 33.8 (2011), S. 1148–1165.
- [44] Margaret W Rossiter. „The Matthew Matilda effect in science“. In: *Social studies of science* 23.2 (1993), S. 325–341.
- [45] Laurie A Rudman. „Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management.“ In: *Journal of personality and social psychology* 74.3 (1998), S. 629. URL: <https://doi.org/10.1037/0022-3514.74.3.629>.
- [46] Laurie A Rudman und Stephanie A Goodwin. „Gender differences in automatic in-group bias: Why do women like women more than men like men?“ In: *Journal of personality and social psychology* 87.4 (2004), S. 494. URL: <https://doi.org/10.1037/0022-3514.87.4.494>.
- [47] Laurie A Rudman und Julie E Phelan. „Backlash effects for disconfirming gender stereotypes in organizations“. In: *Research in organizational behavior* 28 (2008), S. 61–79. URL: <https://doi.org/10.1016/j.riob.2008.04.003>.
- [48] Nina Schwalbe und Jennifer Fearon. „Time’s up for journal gender bias“. In: *The Lancet* 391.10140 (2018), S. 2601–2602.
- [49] Xiaoxiao Shang u. a. „DIANES: A DEI Audit Toolkit for News Sources“. In: *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR ’22. Madrid, Spain: Association for Computing Machinery, 2022, S. 3312–3317. ISBN: 9781450387323. DOI: 10.1145/3477495.3531660. URL: <https://doi.org/10.1145/3477495.3531660>.
- [50] Eran Shor, Arnout van de Rijt und Babak Fotouhi. „A Large-Scale Test of Gender Bias in the Media“. In: *Sociological Science* 6.20 (2019), S. 526–550. ISSN: 2330-6696. DOI: 10.15195/v6.a20. URL: <http://dx.doi.org/10.15195/v6.a20>.
- [51] Stef Spronck und Tatiana Nikitina. In: *Linguistic Typology* 23.1 (2019), S. 119–159. DOI: doi:10.1515/lingty-2019-0005. URL: <https://doi.org/10.1515/lingty-2019-0005>.
- [52] Bundesamt für Statistik. *Männliche Vornamen der Bevölkerung nach Jahrgang, Schweiz und Sprachgebiete, 2021*. URL: <https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/geburten-todesfaelle/namen-schweiz.assetdetail.23045213.html>. (accessed: 29.12.2022).
- [53] Bundesamt für Statistik. *Weibliche Vornamen der Bevölkerung nach Jahrgang, Schweiz und Sprachgebiete, 2021*. URL: <https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/geburten-todesfaelle/namen-schweiz.assetdetail.23045206.html>. (accessed: 29.12.2022).
- [54] Bundesamt für Statistik (BFS). *Analyse der Lohnunterschiede zwischen Frauen und Männern anhand der Schweizerischen Lohnstrukturerhebung (LSE) 2016*. DE. 8266033. Neuchâtel, Mai 2019, S. 212. ISBN: 978-3-303-03292-3. URL: <https://dam-api.bfs.admin.ch/hub/api/dam/assets/8266033/master>.

- [55] Patrick Wagner. *Top 12 der deutschsprachigen Nachrichten der Schweizer*. URL: <https://de.statista.com/infografik/14295/beliebteste-deutschsprachige-medien-in-der-schweiz/>. (accessed: 02.01.2023).
- [56] Cary Wu u. a. „The gender gap in commenting: Women are less likely than men to comment on (men’s) published research“. In: *PLOS ONE* 15.4 (Apr. 2020), S. 1–15. DOI: 10.1371/journal.pone.0230043. URL: <https://doi.org/10.1371/journal.pone.0230043>.

Abbildungsverzeichnis

2.1	Auszug aus der DB Anzahl Artikel pro Nachrichtenportal, Stand 02.01.2023	4
2.2	Query zur Suche der relevanten Literatur	6
3.1	Formel für den <i>Gender Gapp</i>	25
3.2	GitLab Pipeline Web Crawler	26
3.3	Gantt Diagramm der Projektplanung	29
4.1	Ablaufdiagramm der Auswertung	30
4.2	Funktion: <code>_analyze_article()</code>	32
4.3	Anzahl Artikel pro Nachrichtenportal vor der Bereinigung der Duplikate . .	33
4.4	Anzahl Artikel pro Nachrichtenportal ohne Duplikate	33
4.5	Person Klasse	34
4.6	GenderizedPerson Klasse und Gender Enum	35
4.7	Subtree Repräsentation eines Syntaktischen Zitats	36
4.8	Parse Tree eines Satzes mit einem Syntaktischem Zitat in der direkten Rede	37
4.9	Parse Tree eines Satzes mit einem Syntaktischem Zitat in der indirekten Rede.	37
4.10	Function <code>_get_quote_node</code>	38
4.11	Function <code>_get_nearest_token_by_condition</code>	39
4.12	Quotation verbs	39
4.13	Formel Recall	40
4.14	Recall der Zitate auf dem Testset	41
4.15	Recall der Zitate pro Nachrichtenportal	42
4.16	Histogramme zur Verteilung der Qualität der gefundenen Zitate	43
4.17	Zitat mit anderer Baumstruktur	44
4.18	Zitat mit anderer Baumstruktur	44
4.19	<code>analyzed_articles</code> Collection Eintrag Beispiel	46
5.1	Anzahl gefundene Artikel, Stand 22.04.2023	47
5.2	Anzahl gefundene Zitate	48
5.3	Verteilung der Anzahl gefundener Zitate pro Geschlecht und Nachrichtenportal	48
5.4	Anzahl gefundene Zitate pro Artikel pro Nachrichtenportal	49
5.5	Gender Gap pro Nachrichtenportal	49
7.1	Zeitplan ausgefüllt	58
7.2	Kanban Board der Dokumentationsphase	59

Tabellenverzeichnis

4.1	Legende zu den wichtigsten Spacy Parser Labels	36
5.1	Gender Gap pro Nachrichtenportal	50

Glossar

Attribute Word

Englisch für ein Wort, dass einem Subjekt eine oder mehrere Eigenschaften zuschreibt.
Beispiel: „Politikerin“. 12

Dependency Parsing

Eine *NLP* Technik, die die Abhängigkeiten von Wörtern innerhalb von Sätzen bestimmen kann.. ii, 3, 24, 72

EqualVoice

Eine Initiative vom Ringier Verlang um die Präsenz von Frauen auf ihren Plattformen zu messen [1].. 52

F1 Score

Das Harmonische Mittel aus *Precision* und *Recall*. 10

Gender Bias

Geschlechterbias. 2, 8, 17, 21

Gender Bias im Inhalt

Bias in der Bedeutung der Beschreibungen unterschiedlicher Geschlechter.. 11, 12

Gender Bias in der Verteilung

Der Unterschied in der Verteilung der Geschlechter (in einem Datenset). 2, 10–12

Gender Bias in der Wortwahl

Bias in der Wortwahl mit der Personen unterschiedlicher Geschlechter beschrieben werden.. 11, 12

Gender Gap

Unterschied zwischen Geschlechtern. ii, 2, 3, 5, 10, 14, 18, 24, 25, 51, 52, 68

Geschlechterbias

Wahrnehmungsverzerrungseffekte in Bezug auf Gender, auch Genderbias genannten.. 70

Glass Ceiling

Der Begriff "gläserne Decke" bezieht sich auf eine unsichtbare Barriere oder Einschränkung, die Frauen daran hindert, in höhere Positionen am Arbeitsplatz aufzusteigen, insbesondere in Macht- und Führungspositionen. Der Begriff wird in der Geschlechterforschung und der Soziologie verwendet, um die systemischen und oft subtilen Formen der Diskriminierung und Ungleichheit hervorzuheben, mit denen Frauen im beruflichen Umfeld konfrontiert sind. Die gläserne Decke steht für die tief verwurzelten geschlechtsspezifischen Vorurteile und strukturellen Hindernisse, die den beruflichen Aufstieg von Frauen behindern und die Ungleichheiten zwischen den Geschlechtern in der Arbeitswelt verstärken.. 52

Knowledge Base Population

Knowledge Base Population (KBP) bezieht sich auf die automatisierte Extraktion von strukturierten Informationen aus unstrukturierten Texten, um eine Wissensdatenbank aufzubauen oder zu erweitern. . 14

Lemmatisierung

Lemmatisierung ist die Reduktion eines Wortes auf die Grundform (das Lemma). 13

Mathilda Effekt

Beschreibt die systematisch kleinere Anerkennung von Wissenschaftlerinnen verglichen mit Wissenschaftlern. 21

Pay Gap

Lohnunterschied (zwischen den Geschlechtern). 2

Possessive Noun

Englisch für ein Nomen, dass Besitz oder eine Richtung beschreibt. Beispiel: „It's Maria's dog“. „Maria's“ is a possessive noun. 12

Precision

Der Bruchteil der gefundenen relevanten Daten über alle relevanten Daten im Datensatz. 10, 70

Recall

Der Bruchteil der gefundenen relevanten Daten über alle als relevant eingestuft Daten. 10, 53, 70

Ringier

Ein grosser Medienkonzern der Schweiz, zu dem auch Blick gehört.. 52

Schwimmendes Zitat

Ein Zitat ohne Subjekt. Im *GGT* „Floating Quote“ genannt. Meist in der direkten Rede und in der Nähe eines *Syntaktisches Zitat*. „Das Wetter ist schön“ hat Mark laut gesagt. „Obwohl es etwas Wolken hat“. 8

Spacy

Eine *NLP Library* ¹, die *NER*, *POS* und *Dependency Parsing* beinhaltet.. ii, 24, 27, 28, 36, 37, 53, 72

Spacy Parser

Ein Teil der *Spacy* Komponente.. 35, 36

Syntaktisches Zitat

Ein Zitat, das über Subjekt, einleitendes Verb und Aussage verfügt. 8, 72

VADER

Ein tool für Sentiment Analysis. 12

¹<https://spacy.io/>

Abkürzungen

ABC

Australian Broadcast Corporation. 21

API

Application Programming Interface. 13

Coreference Resolution

Coreference Resolution. ii, 9, 13, 24, 27, 34, 56

CRA

Centering Resonance Analysis. 12

DEI

Diversity, Equity and Inclusion. 13, 14

DIANES

Diversity Auditor for News Sources. 13, 14

GGT

Gender Gap Tracker. 8, 11, 24, 51, 72

JAMA

Journal of the American Medical Association. 22, 23

KBP

Knowledge Base Population. 71

MIND

Microsoft News Dataset. 11, 12, 51

ML

Machine Learning. ii, 14, 27, 40

NCD

News Category Dataset[36]. 11, 12, 51

NEJM

New England Journal of Medicine. 22, 23

NER

Named Entity Recognition. ii, 3, 9, 13, 24, 27, 34, 56, 72

NLP

Natural Language Processing. ii, 3, 6, 8, 14, 27, 28, 70, 72

NYT

New York Times. 21

PNAS

Proceedings of the National Academy of Sciences (PNAS). 14, 16

POS

Part of Speech. ii, 3, 13, 24, 27, 34, 53, 72

SMH

Sydney Morning Herald. 21

SSH

Secure Shell. 44

URL

Uniform Resource Locator. 45

XML

Extensible Markup Language. 13