



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Verortung geschichtswissenschaftlicher Quellen mittels
Geoparsing“

verfasst von / submitted by

André Novak, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2021 / Vienna 2021

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 066 856

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Kartographie und Geoinformation

Betreut von / Supervisor:

Ass.-Prof. Mag. Dr. Andreas Riedl

Inhaltsverzeichnis

Abbildungsverzeichnis.....	vi
Tabellenverzeichnis.....	vii
Kurzfassung/Abstract	viii
Vorwort	x
1. Einleitung.....	1
2. Theoretische Grundlagen.....	5
2.1. Digital Humanities	5
2.2. Spatial Turn in den Digital Humanities und der Geschichtswissenschaft	9
2.3. Spatial Humanities und Spatial History	12
3. Historical GIS	15
4. Qualitative GIS.....	21
4.1. Qualitative Humanities und Qualitative Historical GIS.....	25
5. Quellen in der Geschichtswissenschaft und Quantitative vs. Qualitative Daten in der Geschichtswissenschaft und in der Geoinformation	28
5.1. Quellen	28
5.2. Quantitative Forschung in der Geschichtswissenschaft.....	30
5.3. Qualitative Forschung in der Geschichtswissenschaft	31
5.4. Vergleich quantitative und qualitative Forschungsansätze in der Geschichtswissenschaft...	32
5.5. Quantitative und Qualitative Daten in GIS.....	33
6. Data Mining	35
6.1. Text Mining.....	37
6.1.1. Data/Text Mining in den Digital Humanities	40
6.1.2. Location Extraction via Text Mining	40
6.1.3. Geoparsing.....	42
6.1.3.1. Gazetteer-Lookup basiertes Geoparsing	44
6.1.3.2. Regelbasiertes Geoparsing	46

6.1.3.3.	Machinelearningbasiertes Geoparsing.....	46
6.1.3.4.	Toponym Recognition	47
6.1.3.5.	Toponym Resolution	48
7.	Verortung geschichtswissenschaftlicher Quellen mittels Geoparsing	49
7.1.	Wahl der Entwicklungsumgebung.....	51
7.2.	Quellenauswahl.....	54
7.2.1.	Quelle 1: Brief Weiß 1940	56
7.2.2.	Quelle 2: Neue Freie Presse, 24.09.1886	57
7.2.3.	Quelle 3: Ida Pfeiffer, Reise einer Wienerin in das Heilige Land	58
7.2.4.	Quelle 4: Brief WW1_Training.....	60
7.3.	Geoparsing im engeren Sinne	61
7.3.1.	Toponym Recognition	61
7.3.2.	Toponym Resolution	67
7.4.	Ergebnisse und deren Visualisierung.....	74
7.4.1.	Q1_Brief_Weiss_1940	75
7.4.2.	Q2_NFP_24091886.....	78
7.4.3.	Q3_Pfeiffer	81
7.4.4.	Weitere Visualisierungsmöglichkeiten	84
7.5.	Vergleich der Ergebnisse und manuelle Nachbearbeitung	85
7.6.	Zukünftige Optimierungsmöglichkeiten.....	89
8.	Diskussion.....	91
9.	Fazit	95
10.	Literaturverzeichnis	96
11.	Anhang.....	104
11.1.	Anhang I Quellen:	104
11.1.1.	Otto Weiß an seine Ehefrau am 31.5.1940 (3.2002.1202) - Transkript	104
11.1.2.	Neue Freie Presse, 24.09.1886, Seiten 1-2.....	104
11.1.3.	Ida Pfeiffer: Reise einer Wienerin in das Heilige Land, Kapitel 1	113

11.1.4.	Training: „drill before breakfast“ - Transkript	118
11.2.	Anhang II: Python Code Geoparsing.....	119
11.3.	Anhang III: Ergebnisse des Geoparsings, Tabellen	125
11.3.1.	Q1_Brief_Weiss_1940	125
11.3.2.	Q2_NFP_24091886.....	125
11.3.3.	Q3_Pfeiffer, Kapitel 1	126
12.	Eidesstattliche Erklärung	128

Abbildungsverzeichnis

Abbildung 1: Data-Mining-Prozess (Quelle: CLEVE und LÄMMEL 2020: 3)	37
Abbildung 2: Workflow Arbeit mit „geo-text data“ (Quelle: HU 2018: 11)	42
Abbildung 3: Geoparsing-Pipeline (verändert nach: GRITTA, et al. 2018: 604)	44
Abbildung 4: Workflow-Diagramm Verortung geschichtswissenschaftlicher Quellen via Geoparsing	50
Abbildung 5: Q1_Brief_Weiss_1940, erste Seite (Bildquelle: Museumsstiftung Post und Telekommunikation o.J.)	56
Abbildung 6: Beispiel OCR (Quelle: Neue Freie Presse 24.09.1886: 1)	63
Abbildung 7: spaCy-Workflow (Bildquelle: Explosion 2021)	62
Abbildung 8: spaCy Part-of-Speech-Tagging	63
Abbildung 9: spaCy Named Entity Recognition Q1_Brief_Weiss_1940	64
Abbildung 10: spaCy Named Entity Recognition Englisch	66
Abbildung 11: Ausschnitt GeoNames (Denis CARRIERE 2013)	67
Abbildung 12: Gazetteer-Lookup basiertes Geoparsing	72
Abbildung 13: Verortung mit GeoNames-Gazetteer	73
Abbildung 14: Verortung mit OSM-Gazetteer	73
Abbildung 15: Screenshot aus GeoNames (Bildquelle: geonames.org o.J.)	76
Abbildung 16: Karte Verortung Q1_Brief_Weiss_1940	77
Abbildung 17: Karte Verortung Q2_NFP_24091886	80
Abbildung 18: Karte Verortung Q3_Pfeiffer, Kapitel 1	81
Abbildung 19: Karte Verortung Q3_Pfeiffer	83
Abbildung 20: Heatmap Q3_Pfeiffer, bearbeitet	84
Abbildung 21: Karte Verortung Q3_Pfeiffer, manuell bearbeitet	87
Abbildung 22: Q3_Pfeiffer, richtig und falsch verortete Ortsnamen	88

Tabellenverzeichnis

Tabelle 1: Vergleich der Eigenschaften historischer Quellen und Daten und der Anforderungen eines GIS (vgl. GREGORY und ELL 2007: 17; BODENHAMER 2013: 6; BARTELME 2005).....	18
Tabelle 2: Einteilung textlicher Quellen	30
Tabelle 3: Vergleich Quantitative und Qualitative Forschung (basierend auf: HUG und POSCHESCHNIK 2015: 86; SCHUURMAN 2009: 148)	32
Tabelle 4: Quelleneigenschaften Q1_Brief_Weiss_1940	56
Tabelle 5: Quelleneigenschaften Q2_NFP_24091886.....	58
Tabelle 6: Quelleneigenschaften Q3_Pfeiffer	59
Tabelle 7: Quelleneigenschaften Q4	61
Tabelle 8: Confusion Matrix Q1_Brief_Weiss_1940 nach NER	75
Tabelle 9: Confusion Matrix Q1_Brief_Weiss_1940 nach Geocoding.....	76
Tabelle 10: Confusion Matrix Q2_NFP_24091886	78
Tabelle 11: Confusion Matrix Q3_Pfeiffer, Kapitel 1.....	82
Tabelle 12: Vergleich der Geoparsing-Ergebnisse.....	85

Kurzfassung/Abstract

„Historical GIS“, also die Anwendung von Methoden der Geoinformation in der Geschichtswissenschaft, entstand im Zuge des *spatial turns*. Verbunden mit der Verbreitung der „Digital Humanities“ entwickeln sich so neue Arbeitsweisen und die Möglichkeit, auch ganz neue Fragen zu stellen. Da in der Geschichtswissenschaft vor allem mit textlichen Quellen gearbeitet wird, sollte es daher das Ziel sein, eine Arbeit mit diesen auch im Bereich des Historical GIS zu ermöglichen. Daher wurden „Qualitative GIS“-Ansätze gewählt, um dies zu verwirklichen. „Text Mining“ kann dabei helfen, räumliche Informationen aus Texten zu extrahieren. Daher wurde ein „Geoparsing“-Workflow (Toponym Recognition + Resolution) entwickelt, welcher die Verortung geschichtswissenschaftlicher Quellen einfacher und automatisierbar macht. Dabei wurde das Online-Gazetteer „GeoNames“ für das Geocoding gewählt. Die Programmierung wurde mittels „Python“ und dem Natural-Language-Package „spaCy“ umgesetzt. Der Workflow wurde mit drei verschiedenen Quellentypen (Brief, Zeitung, Reisebericht) getestet. Auch wenn eine automatische Transkription einer Quelle mittels „OCR“ die Verortung beeinflusst, können solche Texte trotzdem zur Verortung herangezogen werden. Eine Evaluation der Ergebnisse zeigte, dass auch wenn der Workflow nur einen Teil der Ortsnamen richtig verorten konnte, trotzdem aussagekräftige Aussagen über den jeweiligen Text getroffen werden konnte, ohne diesen jemals genau gelesen haben zu müssen. Eine manuelle Nachbearbeitung kann dabei das Ergebnis deutlich verbessern. Trotzdem gibt es noch einige Optimierungsmöglichkeiten für zukünftige Arbeiten.

“Historical GIS”, that is the application of geographic information methods in historical studies, developed at the time of the *spatial turn*. In connection with “Digital Humanities”, new work methods and the possibility to answer completely new questions developed. Since in historical studies people mainly work with textual sources, the aim should be to work with them in the area of Historical GIS. For that reason, “Qualitative GIS”-approaches were chosen to realise this. “Text mining” can help to extract spatial information from texts. Therefore, a “geoparsing”-workflow (toponym recognition + resolution) was developed, which makes the localisation of historical sources easier and automatable. The online gazetteer “GeoNames” was chosen for geocoding. The programming was done in “Python” and the natural-language-processing package used was “spaCy”. The workflow was tested with three different types of

sources (letter, newspaper, travel report). Even if an automatic transcription of a source using “OCR” influences the localisation, such texts can still be used for the “geoparsing”-process. An evaluation of the results showed that, although the workflow could only correctly locate parts of the place names, meaningful conclusions could still be drawn about the respective text without reading it closely. Manual post-processing can significantly improve the result. Nevertheless, there are still some optimisation options for future work.

Vorwort

In dieser Arbeit wurde, soweit notwendig, geschlechtergerechte Sprache verwendet. Alle Bezeichnungen, bei denen dies übersehen wurde, sind als geschlechtsneutral anzusehen.

Bei der Erstellung dieser Arbeit waren einige Personen besonders unterstützend, bei diesen möchte ich mich nun folgend bedanken: Ass.-Prof. Mag. Dr. Andreas Riedl für die Betreuung der Arbeit und immer nützliches Feedback, Mag.^a Sarah Anja Novak, meiner Schwester, und ihrem Kollegen Michael Wildauer für die unermüdliche Korrektur der Arbeit und meinen Studienkollegen Stefan Rauscha und Volodymyr Ostash für hilfreiches Feedback im Anfangsstadium der Arbeit.

André Novak, Wien am 16.09.2021

1. Einleitung

Das Aufkommen der „Digital Humanities“ ist eine Entwicklung der Geistes- und Kulturwissenschaften seit dem Beginn des 21. Jahrhunderts. Diese sind ein Teilbereich, welcher sich mit der Anwendung informationstechnologischer Methoden auf geisteswissenschaftliche Themen beschäftigt. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 13) Dadurch ist es möglich, schon bisher bearbeitbare Fragestellungen zu beantworten aber auch neue Themenbereiche zu erschließen.

Zusätzlich kam es im Zuge des „Spatial Turns“ zur Herausbildung der „Spatial Humanities“, welche zusätzlich zu klassischen Fragestellungen der Geistes- und Kulturwissenschaften auch den Faktor „Raum“ in Überlegungen miteinbeziehen. Dies geschieht nicht nur durch Karten, sondern auch mittels Geographischer Informationssysteme (GIS). (vgl. DÖRING und THIELMANN 2009: 10ff; AYERS 2010: 1ff)

Während in GIS jedoch vor allem mit quantitativen Daten gearbeitet wird, liegt in den Geistes- und Kulturwissenschaften und in der Geschichtswissenschaft der Fokus vor allem auf der Bearbeitung und Verwendung qualitativer Daten und textlicher Quellen. GREGORY, DEBATS und LAFRENIERE (2018) bezeichnen daher den Einsatz qualitativer Daten aus den Geistes- und Kulturwissenschaften in GIS als wichtigstes potentiell Feld für das Gebiet „Historical GIS“. Dies bezeichnet dabei den Einsatz Geographischer Informationssysteme in der Geschichtswissenschaft.

Die Einbindung nicht-quantitativer Daten in GIS wird auch als „Qualitative GIS“ bezeichnet (ARPAGIAN und AITKEN 2018). Damit sind aber nicht nur die Darstellung, sondern auch die für Geoinformatiker*innen äußerst relevanten Analyse-Funktionen gemeint. Eine Kombination dieser beiden Bereiche, also Historical GIS und Qualitative GIS, sollte also das Ziel sein, um im Bereich der Geschichtswissenschaft durch GIS eine wirkliche Bereicherung zu erreichen.

Es gibt zwar einige wenige Beispiele, die dies im Bereich der Geschichtswissenschaft bereits umzusetzen versuchten, jedoch schöpften diese noch nicht das volle Potential dieses Ansatzes aus. (vgl. GREGORY und ELL 2007: 195ff) Eine räumliche Verortung textlicher Quellen, welche eine der Hauptquellen der Geschichtswissenschaft darstellen, scheint dabei besonders relevant. Denn dadurch ist eine visuelle Interpretation der Geschehnisse auf einer Karte und eine weitere Analyse dieser Ergebnisse möglich. Eine „Location Extraction“ von Texten fand jedoch

bisher zumeist nur für Nachrichten in Social-Media-Diensten statt (vgl. MIDDLETON, et al. 2018: 4ff). GREGORY und COOPER (2009), MONCLA, et al. (2017), ALEX, et al. (2019) oder das Projekt „Ein Literarischer Atlas Europas“ (vgl. Institut für Kartografie und Geoinformation, ETH Zürich 2019) arbeiten zwar mit älteren und längeren Texten, da es sich bei diesen jedoch um literarische Texte handelt, fehlt der Fokus auf die geschichtswissenschaftliche Relevanz der Texte. In der vorliegenden Arbeit wurde daher ein Geoparsing-Workflow entwickelt, um das Potential und Probleme der Verortung geschichtswissenschaftlicher Quellen aufzuzeigen.

Dieser Ansatz stellt auch für die Geschichtswissenschaft ein relevantes Entwicklungsfeld dar, da in diesen viele textliche Quellen vorliegen, welche räumliche Informationen enthalten. (vgl. YUAN 2010) Zusätzlich bietet eine interdisziplinäre Arbeitsweise für beide Forschungsbereiche (Geschichtswissenschaft und GIScience) die Möglichkeit, sich weiterzuentwickeln und neue Ansätze herauszubilden. Die dadurch geschaffene Verbindung zwischen „sciences“ und „humanities“ ermöglicht eine Nutzung von Synergien, welche bei einer streng getrennten Teilung in zwei Bereiche nicht möglich wäre. (vgl. DRUCKER 2021: 10; JANNIDIS, KOHLE und REHBEIN 2017: 14f)

Ziel dieser Arbeit ist es daher, aufzuzeigen, inwieweit es möglich ist, in GIS textliche Quellen aus dem Bereich der Geschichtswissenschaft zu verorten und mit diesen dann weiterarbeiten zu können. Die theoretische Grundlage ermöglicht es dabei, die Ergebnisse des praktischen Teils zu bewerten und zu evaluieren. Hierfür ist es von besonderer Wichtigkeit, eine gute Datengrundlage zu schaffen, da es in diesem Bereich nahezu keine Vorarbeiten gibt. Denn erst mit einer solchen kann das Potential der „Historical GIS“ anschaulich erläutert werden.

Um diese Problemstellung abdecken zu können, ergibt sich folgende Forschungsfrage: *Wie lassen sich textliche Quellen aus dem Bereich der Geschichtswissenschaft mittels Geoparsing in Geographischen Informationssystemen (GIS) verorten?*

Um diese Frage ausreichend zu beantworten, wird in der Arbeit auf folgende Teilfragen genauer eingegangen:

- Wie sehen Methoden und Ansätze im Bereich des „Qualitative GIS“ aus?
- Wie müssen geschichtswissenschaftliche Texte aufbereitet werden, um sie in GIS verwenden zu können?

- Welche geschichtswissenschaftlichen Quellen sind für eine Bearbeitung mittels GIS relevant?
- Wie können textliche Quellen aus dem Bereich der Geschichtswissenschaft mittels GIS verortet und präsentiert werden?
- Welche Einschränkungen gibt es bei der Bearbeitung textlicher Quellen aus dem Bereich der Geschichtswissenschaft mittels GIS?

Zunächst wurde eine Literaturrecherche zu den beiden grundlegenden Themen „Historical GIS“ und „Qualitative GIS“ durchgeführt, wodurch eine theoretische Grundlage für weitere Schritte gelegt wurde. Dabei wurden das Potential und mögliche Probleme dieser beiden Gebiete für die Geschichtswissenschaft klar gemacht. Anschließend wurde mittels praktischer Beispiele aufgezeigt, welche Möglichkeiten GIS bieten, um textliche Quellen aus dem Bereich der Geschichtswissenschaft zu verarbeiten, zu verorten und zu präsentieren. Diese werden dann evaluiert und mögliche Probleme zur Diskussion gestellt. Dadurch werden potentielle Vorgangsweisen und Optimierungsmöglichkeiten für künftige Projekte in diesem Bereich aufgezeigt.

In der vorliegenden Arbeit werden Methoden und Ansätze aus dem Bereich des „Qualitative GIS“ in das Gebiet „Historical GIS“ übertragen, damit die für Geschichtswissenschaftler*innen relevanten qualitativen Daten auch im Sinne des „Spatial Turns“ bearbeitet werden können. Dadurch wird es möglich, das Feld der Geschichtswissenschaft im Bereich der „Digital Humanities“ zu erweitern. Dieser aufstrebende Forschungsansatz ermöglicht es, bisher ungenutztes Potential zu erschließen und dadurch nicht nur neue Antworten zu finden, sondern auch ganz neue Fragen zu stellen.

Nachfolgend wird zuerst ein Überblick über die theoretischen Grundlagen gegeben. Der Ansatz der „Digital Humanities“ ist eine Voraussetzung für eine theoretisch und methodisch fundierte Arbeit mit Ansätzen aus der Informationstechnologie im Bereich der Geistes- und Kulturwissenschaften. Ein Anwendungsbereich aus diesen wird im Kapitel „Historical GIS“ vorgestellt. Hierbei geht es um eine Anwendung von Methoden und Tools aus dem Bereich der Geographischen Informationssysteme in der Geschichtswissenschaft. Bei der Beschreibung dieses Teilbereichs wird ein Fokus auf den *spatial turn* gelegt.

In Kapitel 4 wird ein weiterer Teilbereich der Geographischen Informationssysteme vorgestellt: „Qualitative GIS“. Bei diesen geht es die Verwendung nicht-quantitativer Daten in GIS. Kapitel 5 beleuchtet die Eigenschaften geschichtswissenschaftlicher Quellenarbeit und die quantitative und qualitative Forschung im Bereich der Geschichtswissenschaft und der Geographischen Informationssysteme. Dabei werden auch die Vor- und Nachteile unterschiedlicher Datentypen beschrieben.

Im 6. Kapitel, „Data Mining“, wird ein Fokus auf Location Extraction via Geoparsing gelegt. Dabei werden unterschiedliche Vorgangsweisen vorgestellt und der im praktischen Teil verwendete Workflow definiert.

Im 7. Kapitel wird der praktische Teil der Arbeit beschrieben. Zuerst werden die verwendeten Quellen genauer vorgestellt. Außerdem ist der Ablauf des Workflows Schritt für Schritt erläutert. Dabei werden auch im endgültigen Ablauf nicht verwendete Ansätze ausprobiert und vorgestellt (zum Beispiel gazetteer-basiertes Geoparsing), um die Vorteile des verwendeten Ansatzes gegenüber anderen zu veranschaulichen. Anschließend folgt eine Visualisierung der Ergebnisse geordnet nach den verschiedenen Quellen. Die Ergebnisse der Verortungen der Quellen werden anschließend miteinander verglichen und mittels Confusion Matrix evaluiert.

Danach folgt eine Diskussion des Ergebnisses mit Bezug zur in den vorherigen Kapiteln vorgestellten Theorie und ein Ausblick auf zukünftige Möglichkeiten in diesem Arbeitsbereich. Im Anhang anschließend an das Literaturverzeichnis können die verwendeten Quellen, der geschriebene Python-Code und das Ergebnis des Geoparsings als Tabelle genauer betrachtet werden.

2. Theoretische Grundlagen

Im folgenden Kapitel wird ein Überblick über Bereiche gegeben, die für die Entwicklung der „Historical GIS“ relevant sind.

2.1. Digital Humanities

Der Begriff „Digital Humanities“ (Digitale Geisteswissenschaften) bezeichnet die Anwendung von Informationstechnologie in den Geistes- und Kulturwissenschaften. Dies können zum Beispiel digitale Texteditionen von literarischen Werken oder Deep-Learning-Verfahren zur Unterscheidung von durch Grabdiebe geöffnete alte Gräber und ungestörte Gräber sein. Projekte, in denen digitale Methoden angewendet werden, unterliegen dabei immer technologischen Beschränkungen. Denn nur was die verwendete Technologie leisten kann, kann auch in Projekten im Bereich der Digital Humanities umgesetzt werden. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 13, DRUCKER 2012: 85)

Die Geistes- und Kulturwissenschaften sind dabei jene Wissenschaftsbereiche, welche sich mit den Menschen, ihrem Handeln, Erfahrungen und kulturellen Traditionen auseinandersetzen. Dabei werden viele verschiedene Teilbereiche dieser Themengebiete erfasst: Archäologie, Literatur- und Sprachwissenschaften, Musikwissenschaft, Geschichtswissenschaft oder Kunstgeschichte. Mit „Digital“ ist die Speicherung von Informationen in den jeweiligen Fachbereichen in binärer Form gemeint. (vgl. DRUCKER 2021: 9)

Informationstechnologische Methoden, mit welchem im Bereich der Digital Humanities gearbeitet wird, sind dabei unter anderem: Digitalisierung zur Erhaltung von Schriftstücken oder Kunstwerken, TEI (Text Encoding Initiative), GIS (Geographische Informationssysteme), Deep Learning/Machine Learning, Gaming oder Visual Representation. (vgl. ebd.)

Johanna DRUCKER (2021) unterteilt die Komponenten der Digital Humanities in drei Teile:

- Material: Unter Material sind Bilder, Texte, Karten, 3D-Modelle, etc. und Kombinationen dieser zu verstehen, welche zentral für ein Digital-Humanities-Projekt sind. Denn diese beinhalten die Informationen über die Menschen und ihre Kultur. Sie können entweder schon digital vorliegen oder müssen erst digitalisiert werden, um anschließend im Zuge des Projektes weiterverarbeitet werden zu können. Dies kann vor allem bei der Erfassung von nur analog vorliegendem Text, wie zum Beispiel archivierten

Zeitungen, zu einem erheblichen Aufwand führen. Diese sind entweder nur durch automatisierte Verfahren, welche dann auch fehlerhafte Ergebnisse bringen können, realisierbar, oder durch manuelle Bearbeitung, welche ebenso sehr zeit- und arbeitsaufwändig ist. Zusätzlich ist es auch notwendig, die digital vorliegenden Daten so vorzubereiten, dass sie effizient weiterverarbeitet werden können.

- Verarbeitung: In Digital-Humanities-Projekten werden rechnergestützte Methoden und Werkzeuge genutzt. Beispiele hierfür sind unter anderem Data Mining, statistische oder räumliche Analysen oder Deep Learning. Die informationstechnologische Verarbeitung ist für die meisten Geistes- und Kulturwissenschaftler*innen eine „black box“, da ihnen eine Ausbildung in jenem Bereich fehlt. Daher ist nicht immer klar wie die Methoden und Werkzeuge, mit welchen sie im Bereich der Digital Humanities arbeiten, überhaupt funktionieren. Das Verstehen dieser Prozesse ist jedoch wichtig für eine anschließende kritische Reflexion dieser.
- Präsentation: Die Ergebnisse der vorangegangenen Schritte werden meist via Online-Plattformen oder digitalen Visualisierungsmethoden präsentiert. (z.B.: Blogs, WordPress-Websites, Webmaps, etc.) Zusätzlich werden die Ergebnisse meist auch in wissenschaftlichen Zeitschriften publiziert.

(vgl. ebd.: 1ff)

Hier ist eine Ähnlichkeit zu dem in der Geoinformation üblichen „EVAP“-Modell zu sehen. Dieses besteht aus den Komponenten Erfassung, Verwaltung, Analyse, Präsentation. (vgl. LANGE 2013: 341). Während jedoch in den Digital Humanities vor allem die Verarbeitung und Präsentation des Materials im Vordergrund steht, ist bei Geographischen Informationssystemen auch die Verwaltung der Daten und eine Analyse dieser wichtig.

HUGHES, CONSTANTOPOULOS und DALLAS (2016) nennen zusätzlich Methoden und Tools als wichtigen Bestandteil der Digital Humanities. Tools, auch Werkzeuge genannt, ermöglichen dabei die Interpretation und Analyse des Materials, während Methoden dem*der Wissenschaftler*in einen bewährten theoretischen Rahmen vorgeben, in welchem Forschung praktiziert und Ergebnisse interpretiert werden können.

Dies zeigt, dass für die Arbeit im Bereich der Digital Humanities nicht nur geistes- oder kulturwissenschaftliche, sondern auch informationstechnologische Kenntnisse wichtig sind. Daher arbeiten in den meisten Digital-Humanities-Projekten Geistes- und

Kulturwissenschaftler*innen mit Informatiker*innen oder Computerwissenschaftler*innen zusammen. Trotzdem bleibt die Rolle der Geistes- und Kulturwissenschaftler*innen äußerst wichtig, da es ihre Aufgabe ist, die Ergebnisse des Arbeitsprozesses zu verstehen, um diese dann auch fachwissenschaftlich interpretieren zu können. (vgl. DRUCKER 2012: 85)

Schon seit den 1940er Jahren gibt es Entwicklungen hin zu den „Digital Humanities“ (DH). Der Jesuit und Sprachwissenschaftler Robert Busa nutzte 1940 Lochkartensortiermaschinen, um Karteikarten mit allen bei Thomas von Aquin vorkommenden Wörtern zu ordnen. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 3)

Seitdem passierte eine Menge. Im Zuge der Entwicklung von Computern und deren weiter Verbreitung vor allem ab den 1980er Jahren kam es auch zu einer verstärkten Nutzung dieser in den Geistes- und Kulturwissenschaften.

Seit dem Beginn des 21. Jahrhunderts nahm diese Entwicklung nochmals ordentlich Fahrt auf. Institute und Forschungsschwerpunkte mit Bezug zu den Digital Humanities wurden gegründet, Studienrichtungen etabliert und Forschungsprojekte initiiert. An der Universität Wien wurde 2014 ein interdisziplinärer Forschungsschwerpunkt mit Fokus auf die Digital Humanities eingerichtet. (vgl. Universität Wien o.J.) Ebenso gibt es an der Österreichischen Akademie der Wissenschaften (ÖAW) seit 2015 das „Austrian Center for Digital Humanities and Cultural Heritage“ (ACDH-CH). An diesem arbeiten unter anderem Sprachwissenschaftler*innen, Archäolog*innen, Musikwissenschaftler*innen (seit 2020 Teil des Instituts) und Historiker*innen, sowie Computerwissenschaftler*innen und Informatiker*innen an Projekten, welche digitale Methoden auf geistes- und kulturwissenschaftliche Themen anzuwenden versuchen. (vgl. Österreichische Akademie der Wissenschaften o.J.)

Dabei stellt sich die Frage, inwieweit Digital Humanities in der Wissenschaftscommunity verbreitet und akzeptiert ist. Claire WARWICK (2016) stellt einige Kritikpunkte vor, unter anderem, dass in Digital-Humanities-Projekten vor allem nach Mustern in Daten gesucht wird und es dabei vor allem um die Suche geht, während die Interpretation der Ergebnisse und eine Erklärung dieser dabei zu kurz kommt.

Auch Johanna DRUCKER (2012) stellt die Frage, ob Geistes- und Kulturwissenschaftler*innen, welche in diesem Bereich arbeiten, wirklich Neues schaffen oder nur die Vorteile neuer Technologien nutzen, um bisherige Arbeitsbereiche einfacher bearbeiten zu können. Diese

Diskussion ist ähnlich der der Geographischen Informationssysteme (GIS), welche immer wieder nur als Werkzeug bezeichnet werden, während andere Vorteile dieser unter den Tisch fallen. Durch Digital Humanities werden nicht nur neue Arbeitstechniken und Werkzeuge, sondern auch Fragestellungen und Methoden in den Geistes- und Kulturwissenschaften ermöglicht. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 13)

„Methoden“ bezeichnet dabei die Summe aller Techniken und Tools, welche dazu dienen, in einem bestimmten Wissenschaftsbereich neues Wissen zu generieren, indem digitale Daten oder Quellen geschaffen und analysiert werden. Informationstechnologische Methoden beziehen sich dabei immer auf Informations- und Kommunikationstechnologien. Diese ermöglichen es unter anderem große Datensätze oder -mengen schneller zu durchsuchen oder zu analysieren, als es durch manuelle Arbeit möglich wäre. Dadurch können bestehende Forschungsansätze effizienter und genauer durchgeführt werden. Zusätzlich können auch Themen und Fragestellungen Ziel von Forschungsansätzen werden, welche bisher aus verschiedenen Gründen nicht bearbeitbar waren. (vgl. HUGHES, CONSTANTOPOULOS und DALLAS 2016: 152f)

Des Weiteren kritisiert DRUCKER (2012), dass innerhalb der Digital Humanities Methoden und Tools aus anderen Wissenschaftsbereichen, wie zum Beispiel Data Mining oder räumliche Analysen, einfach übernommen worden seien, obwohl die Grundlagen dieser den Werten der Geistes- und Kulturwissenschaften diametral gegenüberstehen. Sie nennt diese sogar „hostile“ (DRUCKER 2012: 86), da sie „*positivistic, strictly quantitative, mechanistic, reductive and literal [...]*“ (ebd.: 86) seien. Dadurch würden eine kritische Reflexion des Themas und der Inhalte, sowie qualitative Ansätze in der Bearbeitung von Quellen ausgeschlossen, obwohl diese eigentlich den Geistes- und Kulturwissenschaften inhärent sein sollte. Außerdem merkt sie an, dass durch die informationstechnologische Arbeitsweise qualitative Ansätze nicht möglich seien. Daher plädiert sie für „*[...] computational protocols grounded in humanistic theory and methods.*“ (ebd.:86) Diese werden aber genau durch die Beschäftigung mit digitalen Ansätzen gefördert, da durch die Möglichkeit, neue Fragestellungen entwickeln zu können, diese immer auch neue theoretische Grundlagen brauchen. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 14) Im Gegensatz zu DRUCKER denke ich, dass diese Grundlagen jedoch nicht zwangsläufig nur geistes- und kulturwissenschaftlich sein sollten. Auch solche aus anderen Disziplinen, welche Kompetenzen in informationstechnologischen Arbeitsweisen besitzen (z.B.: Informatik, Data Science, Geoinformatik, Artificial Intelligence), sollten berücksichtigt werden. Diese

Kombination, auch Interdisziplinarität genannt, sollte daher unumgänglich Teil der Digital Humanities sein. Dadurch können neue Forschungsmethoden etabliert werden, welche Probleme, die vor allem im Bereich der Digital Humanities relevant sind, besser lösen könnten. (vgl. HUGHES, CONSTANTOPOULOS und DALLAS 2016: 153)

Johanna DRUCKER zeichnet hier eine Art Schreckgespenst an die Wand, nämlich dass die Geistes- und Kulturwissenschaften durch eine Beschäftigung mit informationstechnologischen Methoden ihren Charakter verlieren würden, da Daten und Quellen aus diesem Bereich sich nicht für die informationstechnologisch notwendige Parametrisierung eignen würden. (vgl. DRUCKER 2012: 89) Dass eine Interdisziplinarität, welche Vorteile aus unterschiedlichen Wissenschaftsbereichen in sich vereint, zu neuen Fragestellungen, Methoden und Erkenntnissen führen kann, ist jedoch ebenso relevant. (vgl. MCCARTHY 2016: 77ff) DRUCKER plädiert daher für theoriegebundene Wissenschaft: *„The question is not, Does digital humanities need theory? but rather, How will digital scholarship be humanistic without it?“* (DRUCKER 2012: 94)

2.2. Spatial Turn in den Digital Humanities und der Geschichtswissenschaft

Da die Geistes- und Kulturwissenschaften, und in ihnen vor allem die Geschichtswissenschaft, sich besonders mit dem Menschen beschäftigen und dieser nicht ohne Räumlichkeit existieren kann, ist es nur logisch, dass die Kategorie „Raum“ in diesem Wissenschaftsbereich besonders relevant sein sollte. Eine Erhöhung der Relevanz dieser ermöglichte der sogenannte *spatial turn*, welcher eine Hinwendung zu räumlichen Fragestellungen darstellt.

Bei der Arbeit mit räumlichen Fragestellungen bezogen auf die Vergangenheit sollten die verschiedenen Raumvorstellungen vergangener Jahrhunderte immer kritisch betrachtet werden, da zum Beispiel im 19. Jahrhundert Räume als homogene Einheiten verstanden wurden, wodurch eine differenzierte Auseinandersetzung mit der Kategorie „Raum“ nicht möglich war. Räume dienten damals als Abgrenzung territorial-kultureller Einheiten. Diese Vorstellung eines solchen „Raumes“ war die Grundlage für die nationalsozialistische Kriegs- und Expansionsbestrebungen, welche Lebensraum für arische Bürger*innen im Osten Europas schaffen sollte. (vgl. CSÁKY und LEITGEB 2009: 8f)

Dieser schreckliche Missbrauch des Begriffes „Raum“ durch die Nationalsozialisten führte „[...] nach 1945 [...] zu einer Diskreditierung und tendenziellen Dethematisierung des Raumbegriffs

[...]“ (LOSSAU 2009: 29) Trotzdem kam es auch im deutschsprachigen Raum, wenn auch später als im angloamerikanischen und französischen, zu einer Art Wiederentdeckung dieses Konzeptes vor circa 30 Jahren. Im Zuge dieser wurden raumbezogene Fragestellungen und Perspektiven aufgewertet und in den Fokus wissenschaftlicher Projekte gerückt. (vgl. ebd.: 29ff)

In den Geistes- und Kulturwissenschaften kam es immer wieder zu Veränderungen im Zugang zu Forschungsfragen. Beispiele hierfür sind: linguistic turn, postcolonial turn, translational turn, iconic/pictorial turn, material turn, social turn und digital turn (BACHMANN-MEDICK 2016: 1) Diese verändern den Fokus eines Wissenschaftsbereiches und ermöglichen es, mit anderen Augen auf ein Thema zu blicken beziehungsweise dieses zu bearbeiten. Sie entstanden nach und nach und können entweder einzeln oder auch kombiniert angewandt werden.

Diese Turns sind jedoch nicht im Sinne von Thomas Kuhns „Paradigmenwechsel“ zu verstehen. Kuhn nennt Paradigmen „[...] *universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners.*“ (KUHN und HACKING 2012: 30) Die Turns im Gegensatz dazu sind interdisziplinär und können sich ablösen, müssen es aber nicht. Daher sind diese auch nicht als Paradigmenwechsel zu verstehen. (vgl. BACHMANN-MEDICK 2016: 9f)

In der zweiten Hälfte des 20. Jahrhunderts und vor allem ab der Jahrtausendwende kam es zum *spatial turn* und dem Hinwenden zum Faktor „Raum“. Hierbei ist jedoch nicht nur die Analyse räumlicher Verhältnisse gemeint, sondern auch das kritische Hinterfragen dieser Räumlichkeiten. Als Beispiel dafür kann die Unterscheidung zwischen Nord-/Süd-Halbkugel, West-/Ost-Europa oder Stadt/Land gesehen werden. (vgl. ebd.: 213ff)

Der Raum stellte dabei lange eine Art „Container“ dar, welcher als physischer Behälter für die soziale Materie funktionierte. Somit setzt der Raum dem Sozialen Grenzen, welche unveränderbar sind. Später kam es jedoch zu anderen, sozialen Definitionen des Raums, welcher nun durch Beziehungen zwischen Objekten funktionierte. Der Raum entsteht dabei erst durch die Beziehungen und sozialen Distanzen. (vgl. NECKEL 2009: 46; LIPPUNER und LOSSAU 2004: 52) KINGSTON (2015) argumentiert hier: „*Human beings [...] interact through space, and only secondarily in space.*“ (ebd.: 114)

Auch wenn der Faktor „Raum“ und die räumliche Dimension gesellschaftlicher Prozesse im Zuge des *spatial turns* enorm an Bedeutung gewonnen haben und dadurch suggeriert wird,

dass Probleme vor allem räumlicher Art seien, darf trotzdem nicht auf die gesellschaftlichen Phänomene, welche sich im Raum abspielen, vergessen werden. (vgl. LIPPUNER und LOSSAU 2004: 47) Denn anschließend an KINGSTON (2015) gibt es den Raum ja nur durch die Menschen, welche in ihm agieren.

LOSSAU (2009) nennt die Entwicklung, welche durch den *spatial turn* angestoßen wurde, eine „Verräumlichung der Kultur“, welche in den Geistes- und Kulturwissenschaften Einzug gefunden hat. Der Raum als solcher wurde in bisher nicht räumlich denkenden Wissenschaftsbereichen relevant, indem „[...] nach dem Zusammenhang zwischen der physisch-materiellen Realität der geographischen Räume einerseits sowie den bedeutungsvollen sozial-kulturellen Wirklichkeiten und deren räumlicher Strukturiertheit andererseits“ (LOSSAU 2009: 31) gefragt wurde. Diese sozial-kulturellen Wirklichkeiten werden durch alltägliche Prozesse immer wieder neu hergestellt und sind dadurch nicht statisch. Diese Ansicht stellt eine große Differenz zum Raumbegriff des 19. und Anfang bis Mitte des 20. Jahrhunderts dar. Der Raum ist nunmehr nicht mehr nur physisch vorhanden, sondern wird auch durch gesellschaftliche Beziehungen hergestellt. Trotzdem sollte die physische Dimension nicht außen vor gelassen werden, da historische Geschehnisse und soziale Beziehungen im physischen Raum stattfindet und dadurch auch von ihm mitgeprägt werden. (vgl. ASSMANN 2009: 15)

Der Geschichtswissenschaft bietet der *spatial turn* insoweit Vorteile, als dass durch ihn die Forschungsarbeit neu gedacht werden kann: „*The spatial turn is given enhanced status as a new focus that seeks to represent the complexity of historical processes on the basis of the simultaneity and interwovenness of disparate spatial dimensions that were previously studied separately.*“ (BACHMANN-MEDICK 2016: 224) Diese Verbundenheit der räumlichen Situation mit den historischen Prozessen, welche im Raum stattfinden, ermöglicht es, einen neuen Blick auf historische Entwicklungen und Phänomene werfen zu können. Außerdem bricht diese Entwicklung die „[...] lange unangefochtenen Privilegierung der Zeit gegenüber dem Raum im Kategoriengefüge der Historiker [...]“ (vgl. DÖRING und THIELMANN 2009: 23) auf. Die Zeit wird zwar trotzdem von den meisten Geschichtswissenschaftler*innen als inhärente Kategorie der Geschichtswissenschaft gesehen, aber nun wird auch der Raum als einflussreiche Kategorie angesehen.

Auch in den Digital Humanities hatte der *spatial turn* seinen Einfluss, auch wenn es lange dauerte und dieser sich in jenem Bereich nur eingeschränkt durchsetzte. Vor allem die

Visualisierung räumlicher Daten und in kleinem Maße auch die Analyse dieser zeigt, dass räumliche Denk- und Arbeitsweisen Anklang gefunden haben und verwendet werden. Diese eingeschränkte Nutzung liegt unter anderem an der Unschärfe und dadurch nicht genauen Bestimmbarkeit von Daten aus dem Bereich der Geistes- und Kulturwissenschaften. (siehe Kapitel 5) Textliche, visuelle, auditive oder audiovisuelle Quellen lassen sich nur schwer so abstrahieren, dass sie gut für räumliche Analysen und Präsentationen verwendet werden können. (vgl. PRESNER und SHEPARD 2016: 205) Trotzdem wird zumeist mit quantitativen Daten gearbeitet, da diese einfacher in bestehende Methoden und Tools anwendbar sind. Johanna DRUCKER (2012) argumentiert daher, dass es neue methodische Überlegungen geben sollte, damit Arbeiten im Bereich der Digital Humanities nicht ohne Grundlagen aus dem Bereich der Geistes- und Kulturwissenschaften stattfindet. Im Laufe der letzten Jahre gab es hier jedoch weitere Fortschritte bzw. Entwicklungen. Beispiele hierfür sind unter anderem: „David Rumsey Map Collection“ (vgl. Cartography Associates 2021), „A Vision of Britian Through Time“ (vgl. University of Portsmouth et al. 2017) oder „The Valley of the Shadow“ (vgl. AYERS 2007).

Räumliche Informationen können dabei in unterschiedlichen Quellentypen vorhanden sein. Allerdings ist es wichtig zu beachten, dass der *„der kulturwissenschaftliche [Raum erst dort zugänglich gemacht wird], wo er oder etwas an ihm sich in Text verwandelt hat (oder in etwas Textanaloges), das lesbar ist wie eine Sprache“* [...]“ (DÖRING und THIELMANN 2009: 17) Daher sind auch im Sinne des *spatial turns* arbeitende Geistes- oder Kulturwissenschaftler*innen vor allem auf textliche Quellen angewiesen. Daher ist es von besonderer Relevanz, dass räumliche Informationen aus solchen Quellen extrahiert werden können. (siehe Kapitel 6.1.2)

2.3. Spatial Humanities und Spatial History

In den Geistes- und Kulturwissenschaften führte der *spatial turn* zu neuen Forschungsrichtungen, welche unter dem Begriff „Spatial Humanities“ zusammengefasst werden können. Diese bieten dabei Vorteile in der Untersuchung räumlicher Phänomene, weil *„in the humanities, GIS is used to show experience and cultural patterns that are encoded in geospatial information.“* (DRUCKER 2021: 141) Im Bereich der Geschichtswissenschaft kam es zum Entstehen der Spatial History, welche ein Teilbereich der Spatial Humanities ist. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 305)

Dabei geht es nicht nur um „[...] *simply ‘recreating’ historical environments and making historical data available to the wider public [...]*“ (PRESNER und SHEPARD 2016: 207), sondern vor allem auch darum, Geovisualisierung und räumliche Analysen durchzuführen und deren Funktionsweise und Ergebnisse zu verstehen, interpretieren und erklären zu können. Dadurch kann auch das Problem der statischen und genauen Präsentation und Analyse nicht vollständig bekannter oder darstellbarer Phänomene und Entwicklungen umgangen werden, da die Analysen auch als Argumente dienen können. (vgl. ebd.: 207)

Als methodische Grundlage und Tool im Sinne HUGHES, CONSTANTOPOULOS und DALLAS (2016) funktionieren Geographische Informationssysteme (GIS). Diese ermöglichen es, räumliche Datensätze unterschiedlicher Beschaffenheit, Qualität und Quantität zu erstellen, verarbeiten, analysieren und präsentieren. (vgl. LANGE 2013: 339ff) Außerdem ist es für Geschichtswissenschaftler*innen möglich, interaktive Visualisierungen zu erstellen, welche Veränderungen in Raum und Zeit darstellen können. (vgl. KINGSTON 2015: 111)

Einer der grundlegendsten Kritikpunkte an den Spatial Humanities ist die Verbindung einer „[...] *positivist technology with predominantly humanist traditions.*“ (HARRIS, CORRIGAN und BODENHAMER 2010: 168), da viele Wissenschaftler*innen meinten, dass es nicht möglich sei, diese zufriedenstellend zu kombinieren. GIS seien positivistisch und würden daher nicht mit den Ansichten und Arbeitsweisen von Geistes- und Kulturwissenschaftler*innen verknüpft werden können. Dem ist entgegenzuhalten, dass die Verbindung zwischen beiden Wissenschaftsbereichen immer mehr verschwimmt und daher genau hier angesetzt werden muss, um neue Ansätze zu generieren. Daher sollten Geistes- und Kulturwissenschaftler „[...] *GIS as science (GISci) [...]*“ (ebd.: 170) und nicht nur als Werkzeug bzw. Tools sehen. (vgl. BODENHAMER, CORRIGAN und HARRIS 2010: viii) Dies ermöglicht es, die Geistes- und Kulturwissenschaften neu zu denken und neue Einblicke in bekannte Themen zu bekommen, sowie neue zu entwickeln. Außerdem können dabei auch andere Arten der Visualisierung verwendet werden. Online-Karten, virtuelle Globen, 3D-Modelle, mobile Applikationen und Verortungstools sind Teil unseres täglichen Lebens geworden. Daher sollten die Spatial Humanities auf die Vorteile dieser Werkzeuge zugreifen, um ihre Forschungsarbeiten weiter zu verbreiten. (vgl. GREGORY, DEBATS und LAFRENIERE 2018: 2) Erst wenn dies erreicht, sind die Ziele des *spatial turns* wirklich umgesetzt.

GREGORY, DEBATS und LAFRENIERE (2018) benennen einige wichtige Punkte, in welchen sich der Bereich Spatial History noch verbessern sollte. Auch wenn durch GIS neue Erkenntnisse geschaffen werden, sollte es eigentlich vor allem darum gehend, diese auch zu erklären. Daher muss zukünftig darauf geachtet werden, zu erklären „[...] *how GIS-generated knowledge provides new answers to old questions and allows new historical insights.*“ (ebd.: 3) Zusätzlich ist es derzeit so, dass vor allem die Visualisierung räumlicher Daten aus dem Bereich der Geschichtswissenschaft im Vordergrund steht. GIS werden dabei hauptsächlich zur Visualisierung oder Verwaltung räumlicher Daten verwendet. (vgl. HARRIS, BERGERON und ROUSE 2011: 227) Jedoch liegt einer der Hauptvorteile von GIS und dadurch auch der Spatial History darin, räumliche Analysen durchzuführen und Attribute und deren Einflüsse erklären zu können. Die dritte von GREGORY, DEBATS und LAFRENIERE (2018) festgestellte Diskrepanz ist „[...] *the relationship between the quantitative objectives of GIS and the qualitative data central to humanities scholars.*“ (GREGORY, DEBATS und LAFRENIERE 2018: 3) Diese Dichotomie gilt es zu überwinden, um eine optimale Zusammenarbeit zwischen beiden Bereichen zu ermöglichen. In der vorliegenden Arbeit soll daher dieser Bereich besondere Beachtung finden, indem Ansätze aus dem Bereich „Qualitative GIS“ mit der Geschichtswissenschaft verknüpft werden. Dabei geht es um die Verortung geschichtswissenschaftlicher Quellen mittels Geoparsing (siehe Kapitel 6.1.3).

3. Historical GIS

Die zuvor beschriebene Verbindung zwischen der Geschichtswissenschaft und Geographischen Informationssystemen (GIS) wird auch als „Historical GIS“ bezeichnet. Diese Entwicklung begann vor circa 20 Jahren und findet seitdem immer breitere Anwendung. Historisch relevante räumliche Daten werden dabei in GIS mittels relationaler Datenbanken erfasst und gespeichert. Zusätzlich können diese Daten auch analysiert und visualisiert werden. (vgl. GREGORY, DEBATS und LAFRENIERE 2018: 1f) Dabei funktionieren GIS nicht nur als Werkzeug, sondern helfen auch dabei, über Disziplinengrenzen hinweg Erkenntnisse zu gewinnen und zu erklären, was diese für unser Verständnis der Vergangenheit bedeuten. (vgl. ebd.: 2): *„[...] Human activity is about time and space, and GIS provides a way to manage, relate, and query events, as well as to visualize them [...]“*. (BODENHAMER 2010: 22) Historical GIS bezeichnet dabei einerseits die Anwendung von Funktionen der GIS auf geschichtswissenschaftliche Themen und Fragestellungen und andererseits die Erschaffung von Datenbanken, welche große Mengen an räumlich verorteten geschichtswissenschaftlich relevanten Daten in digitalem Format enthalten. (vgl. GREGORY, DEBATS und LAFRENIERE 2018: 4; GREGORY 2010: 67) Diese werden im Zuge dieser Arbeit jedoch nicht behandelt und sollen daher nur kurz erwähnt werden, weil sie für die Entwicklung des Bereiches Historical GIS äußerst relevant waren.

Diese sogenannten „National Historical GIS“-Datenbanken dienen der Sammlung und Zurverfügungstellung von historisch relevanten räumlichen Daten, welche durch staatliche Förderungen unterstützt im Zuge von großen Projekten gesammelt und bereitgestellt werden. (vgl. GREGORY und ELL 2007: 186ff)

Historical GIS ermöglicht es Geschichtswissenschaftler*innen nicht nur, bereits bestehende Themen und Fragestellungen zu bearbeiten, sondern auch neue Fragen und neue Perspektiven zu ergründen. Dies wird auch dadurch ermöglicht, dass Daten und Quellen, welche zuvor mit klassischen historischen Methoden nicht bearbeitbar gewesen sind, nun in GIS integriert und analysiert werden können. Außerdem können auch schon bisher genutzte Quellen nun neu betrachtet werden. Um daher das gesamte Potential der Historical GIS auszuschöpfen, sollte es eigentlich das Ziel sein, neue Fragestellungen und Perspektiven zu finden. (vgl. RICHARDSON 2011: 211; BODENHAMER 2010: 23f)

Historical GIS stellen die konsequente Umsetzung des *spatial turns* (siehe Kapitel 2.2) in der Geschichtswissenschaft mittels Geographischer Informationssysteme (GIS) dar. Dabei werden

jedoch nicht nur Tools und Methoden aus der Geographie, Kartographie oder Geoinformation in den Bereich der Geistes- und Kulturwissenschaften transferiert. Auch diesen Wissenschaftsbereichen inhärente methodische Zugänge und das räumliche Denken müssen beachtet werden, um Erkenntnisse über räumliche Phänomene zu bekommen. (vgl. HARRIS, BERGERON und ROUSE 2011: 238; BODENHAMER 2010: 24)

Neben der Visualisierung von Informationen sind vor allem auch die Analysefunktionen relevant, welche es ermöglichen, räumliche Beziehungen und Verhältnisse aufzudecken und zu erklären, welche Beziehungen in Texten (siehe Kapitel 6.1) oder Tabellen enthalten sein können, da die räumliche Dimension für Geschichtswissenschaftler*innen immer wichtiger wird. Dies zeigt, dass geographische Methoden und Tools auf geschichtswissenschaftliche Fragestellungen angewandt werden sollten. Die Ergebnisse werden dann zumeist in Form von Karten präsentiert. (vgl. BODENHAMER, CORRIGAN und HARRIS 2010: vii; BODENHAMER 2010: 22; RICHARDSON 2011: 210; HARRIS, BERGERON und ROUSE 2011: 227)

GIS bieten somit unter anderem die Möglichkeit, den räumlichen Kontext bei der Bearbeitung von Büchern und anderen textlichen oder visuellen Quellen, welche einen räumlich-zeitlichen Verlauf abbilden, darzustellen. Dies kann wichtig sein, da Menschen und ihre Aktivitäten eigentlich immer im Raum stattfinden und sie diese dann auch schriftlich festhalten. (vgl. GREGORY und ELL 2007: 16)

BOONSTRA (2013) nennt einige limitierende Faktoren, welche den optimalen und weitreichenden Einsatz von GIS in der Geschichtswissenschaft verhindern: Daten sind zwar teilweise im Internet vorhanden, sind jedoch zumeist unorganisiert und daher schwer zu georeferenzieren. Des Weiteren bieten die gängigen Softwarepakete wenige oder keine Möglichkeiten, geschichtswissenschaftliche Fragestellungen abzubilden oder zu bearbeiten. Zeitliche Änderungen oder die Abbildung fehlender, inkonsistenter oder ungenauer Daten wird dadurch erschwert.

BODENHAMER (2010) stellt fest, dass die Methoden und Tools aus dem Bereich der Geoinformation schwer beziehungsweise langsam Eingang in die Arbeitsweisen von Historiker*innen finden. Auch wenn seit dieser Feststellung mehr als zehn Jahre vergangen sind und es Entwicklungen wie den Digital Humanities gelungen ist, sie vermehrt zu etablieren, gibt es unter manchen Geschichtswissenschaftler*innen noch immer eine Ablehnung der Technik. (vgl. GREGORY

und ELL 2007: 17) Auch wenn Geograph*innen und Geowissenschaftler*innen GIS immer wieder an die Wünsche und Bedürfnisse von Geschichtswissenschaftler*innen anzupassen und zu erweitern versuchen, werden diese zumeist immer noch alleinig als Werkzeug zur Erstellung von Karten oder zur Verwaltung räumlicher Daten gesehen, während die eigentlich besonders wichtigen Analyse-Funktionen nicht bekannt oder ungenutzt sind. (vgl. HARRIS, BERGERON und ROUSE 2011: 227)

Ein weiteres Hemmnis ist der Umstand, dass, während sich mittels GIS vor allem quantitative Daten bequem bearbeiten lassen, Geistes- und Kulturwissenschaftler*innen zumeist mit qualitativ relevanten Themen und Fragestellungen arbeiten, welche sich quantitativ nicht darstellen oder erklären lassen. (vgl. ebd.: 228; BODENHAMER 2010: 23) „*The emphasis in the humanities on qualitative data, linearity, narrative, storytelling, and multiple perspectives does not merge easily with the heavily positivist and scientific emphasis in GIS.*“ (HARRIS, BERGERON und ROUSE 2011: 229) Eine große Schwierigkeit ist dabei vor allem: „*Incorporating and accessing qualitative information such as photographs, paintings, oral history, moving images, text [...].*“ (ebd.: 228) Diese Quellentypen sind jedoch eigentlich genau jene, welche in der Geschichtswissenschaft besonders relevant sind. Zusätzlich sehen die meisten Geistes- und Kulturwissenschaftler*innen GIS nur als Werkzeug zur Bearbeitung, und da vor allem Visualisierung quantitativer Datensätze, da sie keine Ausbildung in diesem Bereich erfahren haben. Dies bedeutet, dass viele Funktionen und Möglichkeiten, welche GIS zur Arbeit mit historisch relevanten räumlichen Daten bieten, unbekannt sind und daher auch nicht zur Anwendung kommen. Gerade eine Ausbildung im Bereich Digital Humanities kann hier Abhilfe schaffen, da Geistes- und Kulturwissenschaftler*innen hier ebenjene Vorteile der Arbeit mit digitalen Methoden aufgezeigt bekommen. Außerdem wird in Projekten in diesem Bereich meist auch mit Spezialist*innen aus der jeweiligen Fachwissenschaft zusammengearbeitet, wodurch eine optimale Anwendung aller Möglichkeiten der digitalen Umsetzung gewährleistet werden kann. Außerdem sind die Ergebnisse somit nicht nur für GIS-Spezialist*innen interessant und verständlich, sondern können auch von den Fachwissenschaftler*innen, welche zumeist das Resultat und nicht den Weg dorthin als wichtig erachten, verstanden und interpretiert werden. (vgl. BODENHAMER 2010: 24; GREGORY und GEDDES 2014b: xii)

Tabelle 1: Vergleich der Eigenschaften historischer Quellen und Daten und der Anforderungen eines GIS (vgl. GREGORY und ELL 2007: 17; BODENHAMER 2013: 6; BARTELME 2005)

Eigenschaften historischer Quellen und Daten	Anforderungen eines GIS für eine Verortung
Unvollständigkeit	Vollständigkeit
Mehrdeutigkeit	Eindeutigkeit
inakkurat	hohe Genauigkeit
unstrukturiertheit	Strukturiertheit
meist alphanumerisch	alphanumerisch und numerisch möglich
meist qualitativ	meist quantitativ
Zeit wichtig	Raum wichtig

Ein zusätzliches Problem in der Arbeit mit historisch relevanten räumlichen Daten in GIS ist die Datenqualität (siehe Tabelle 1). Historische Quellen und Daten sind zumeist nicht vollständig, inakkurat, komplex und manchmal nicht eindeutig. Dies entspricht jedoch nicht den Anforderungen von GIS, da diese abhängig vom Datentyp eine besonders hohe Genauigkeit der Daten verlangen, welche dann in Koordinatensystemen oder mittels räumlicher Modelle dargestellt werden können. (vgl. LANGE 2013: 91) Auch die Genauigkeit der Nachbarschaftsverhältnisse sollte so eindeutig wie möglich sein. Diese Diskrepanz zwischen einzigartigen und zufällig auftretenden Phänomenen und der Generalisierung und Reduktionismus im Bereich der GIS ist dabei schwer zu überwinden. Während also GIS eher der Organisation und Analyse von Wissen in einem eher statischen System dienen, wird in den Geistes- und Kulturwissenschaften das Wissen als polyvalent, mehrdeutig und leicht veränderlich gesehen. (vgl. GREGORY und ELL 2007: 17; BODENHAMER 2013: 6; BODENHAMER 2010: 23f) Dadurch wird die Bearbeitung historisch relevanter räumlicher Daten in GIS eingeschränkt, da ein gemeinsamer Nenner gefunden werden muss, was überhaupt bearbeitet werden soll und wie es analysiert und dargestellt werden kann. Für diese Arbeit ist daher auch zumeist das Wissen von GIS-Expert*innen notwendig, da GIS sehr komplex sind und dadurch deren Funktionsweise schwierig zu verstehen ist. Außerdem sind sie meist umständlich zu bedienen und dadurch auch schwer zu erlernen. Dabei ist es auch für Geschichtswissenschaftler*innen wichtig zu verstehen, wie diese funktionieren, um aus diesen einen Mehrwert beziehen zu können. (vgl. HARRIS, BERGERON

und ROUSE 2011: 227f; GREGORY und GEDDES 2014b: xii) Wichtig dabei ist zu beachten, dass die Aussagekraft der Analysen und Darstellung immer von der Qualität der Daten abhängt. Ist dies bewusst, können trotzdem Aussagen über die Ergebnisse getroffen werden.

Um nun auch aus geschichtswissenschaftlicher Perspektive einen Mehrwert aus diesem Arbeitsbereich zu erhalten, sollten Geschichtswissenschaftler*innen GIS nicht nur als Werkzeug sehen („[...] *for those happy to embrace such changes and look at new ways of working with historical material, then GIS is worth consideration as a tool [sic!] for the historian.*“ (vgl. LILLEY und PORTER 2016: 140)), sondern als Methode beziehungsweise GIScience im Sinne von HUGHES, CONSTANTOPOULOS und DALLAS (2016), damit sie auch die zugrundeliegenden Methoden und Theorien in ihre Arbeit mit einfließen lassen können. Trotzdem sollte auch aus geoinformatischer Perspektive ein Umdenken stattfinden, damit GIS-Methoden und -Tools auch an die Bedürfnisse von Geschichtswissenschaftler*innen angepasst werden. Denn allein die Übernahme von Methoden und Theorien aus einem Fachbereich in den anderen kann zwar am Anfang zu neuen Erkenntnissen führen, hemmt jedoch eine innovative Weiterentwicklung des Bereiches Historical GIS. Erst ein umfangreiches Neudenken dieses Arbeitsbereiches mit einem Einfluss beider Teilbereiche kann zur Entwicklung neuer Methoden und dadurch zu neuen Erkenntnissen führen. Von diesen können dann auch wieder beide Teilbereiche profitieren und zugrundeliegende Methoden und Theorien weiterentwickeln. (vgl. HARRIS, BERGERON und ROUSE 2011: 228)

In einem ersten Schritt sollte es daher das Ziel sein, GIS „fit“ für die Ansprüche von Geschichtswissenschaftler*innen zu machen. Einige Geschichtswissenschaftler*innen erkennen dabei schon das Potential, welches GIS für die Geschichtswissenschaft haben kann:

„Certainly GIS is not for everyone, but as its software becomes more easy to use, more intuitive, and as our everyday lives (and those of our students) become ever more technologically orientated (most of us now turn to Google Maps without a thought, for example), spatial technologies and digital mapping are increasingly going to be part of our research and publication environments.“ (LILLEY und PORTER 2016: 140)

Wenn qualitative Daten und Quellen die Hauptarbeitsmittel in der Geschichtswissenschaft sind, sollten GIS-Spezialist*innen Methoden entwickeln bzw. aufzeigen, welche eine Arbeit mit diesen Quellentypen in GIS ermöglichen. Denn erst dann kann gesagt werden, dass eine

geschichtswissenschaftliche Arbeit mit GIS im Sinne des *spatial turns* und der Digital Humanities zufriedenstellend möglich ist. Es sollte daher ermöglicht werden, dass „[...] *multi-perspectival mapping, ambiguity and uncertainty, and differential approaches that decolonize maps [and] mapmaking [...]*“ (PRESNER und SHEPARD 2016: 210) in GIS Wirklichkeit werden.

Für Geschichtswissenschaftler*innen stellen GIS derzeit zumeist eine Datenbank dar, welche historische Daten mit einem räumlichen Bezug beinhalten. Dies können zum Beispiel georeferenzierte historische Karten sein. Des Weiteren werden GIS vielleicht noch für die Erstellung einfacher Karten mit historischem Bezug verwendet. Eine weiterführende Analyse von Daten, mit welcher dann räumliche oder zeitliche Muster und Veränderungen erfasst werden können, findet äußerst selten statt. (vgl. LILLEY und PORTER 2016: 129ff)

Ein nicht zu vernachlässigender Nachteil ist derzeit außerdem, dass Datensätze, mit welchen in diesem Bereich gearbeitet wird, zumeist selbst erstellt werden müssen, da bisher sehr wenig räumlich verortete historische Datensätze online verfügbar sind. Dadurch erhöht sich der Arbeitsaufwand und die Komplexität bei Projekten um ein Vielfaches, da zu Anfangs erst solche Datensätze generiert werden müssen, bevor diese visualisiert bzw. analysiert werden können.

4. Qualitative GIS

Geographische Informationssysteme (GIS) sind Systeme zur digitalen Erfassung, Verwaltung, Analyse und Präsentation räumlicher Daten (vgl. LANGE 2013: 337ff)

Einen Teilbereich davon stellt Qualitative GIS dar. „*Qualitative geographic information systems (qual-GIS) incorporates nonquantitative data into GIS, integrates qualitative data collection and analysis with quantitative spatial analysis facilitated by GIS [...].*“ (ARPAGIAN und AITKEN 2018) Dabei sind Daten und Quellen aus den Geistes- und Kulturwissenschaften, wie zum Beispiel Beobachtungen, Interviews oder schriftliche Dokumente, zumeist qualitativ und unstrukturiert und nicht numerischen Ursprungs. (vgl. MCHAFFIE, HWANG und FOLLETT 2018: 306) GIS wird dabei gesehen als

„[...] social and political practice, as discipline- and industry-inscribed ways of making knowledge, and as an assemblage of hardware, software, data structures, and procedures for working with digital spatial data. This conceptualization of GIS allows us to consider GIS-based knowledge production as something that is occurring explicitly and implicitly in several sites: spatial data, data structure, spatial analysis, cartographic representation, and the application or use of any of these phenomena in social and political practice.“ (ELWOOD und COPE 2009: 6)

PRESTON und WILSON (2014) sehen hierbei zwei große Arbeitsrichtungen, in welchen Qualitative-GIS-Ansätze derzeit verwendet werden: Einerseits geht es um die technische Integration qualitativer Quellen und Daten in GIS. Zweitens werden auch verschiedene Wissenschaftstheorien konzeptuell überarbeitet und angepasst, um den Bereich Qualitative GIS besser abbilden zu können.

Der Qualitative-GIS-Ansatz ermöglicht es, als „mixed-methods approach“ die dualistische Trennung in quantitative und qualitative Geographie beziehungsweise GIS-Expert*innen und Außenstehende, welche großes Wissen in geistes- kulturwissenschaftlichen Methoden und Arbeitsweisen haben aber GIS wegen der Konzentration auf quantitative Arbeitsweisen kritisch sahen, zu überwinden. (vgl. ARPAGIAN und AITKEN 2018; BITTNER und MICHEL 2018: 154) Dabei steht die Verwendung unterschiedlicher Daten und Quellen als Grundlage und der Einsatz verschiedener Analyse und Visualisierungstechniken im Vordergrund. Ziel ist es dabei, neue Praktiken der Wissensproduktion und neue Formen des Wissens zu ermöglichen.

Untersucht werden sollen dabei komplexe Beziehungen, Bedeutungszuschreibungen, nicht quantifizierbare Eigenschaften, unterprivilegierte soziale Gruppen, nicht lineare menschliche Erfahrungswelten und marginalisierte räumliche Perspektiven, um diese besser zu verstehen. (vgl. ELWOOD und COPE 2009: 4; COPE und ELWOOD 2009: 175; PAVLOVSKAYA 2009: 31; BITTNER und MICHEL 2018: 152)

Hierbei kann eine Verbindung zu zwei geschichtswissenschaftlichen Ansätzen, welche auf ähnlichen Konzepten basieren, hilfreich sein. Während die „Subaltern Studies“ sich mit untergeordneten Gesellschaftsschichten (den Subalternen) beschäftigen (vgl. LUDDEN 2002: 1ff), konzentriert sich die „History from Below“ auf die Sicht einfacher Leute auf historische Ereignisse und deren Alltag. (vgl. CREW 1989: 395ff)

Dabei entstand dieser Ansatz, weil GIS als „[...] *rooted in positivist epistemologies and most suited for quantitative techniques associated with spatial science*“ (ELWOOD und COPE 2009: 1) gesehen wurde. Dadurch war eine Arbeit mit GIS in einem Großteil der Forschungsbereiche nicht möglich, da die Methoden der GIS eine qualitative Arbeit nicht ermöglichten. Diskussionen aufgrund dieses Umstandes gibt es schon länger, da Wissenschaftler*innen nicht zufrieden waren mit den Möglichkeiten, welche GIS boten, um neue Arbeitsbereiche zu erschließen. (vgl. KWAN und KNIGGE 2006) Daher arbeiteten unterschiedliche Wissenschaftler*innen daran, GIS zu erweitern, indem „[...] *[they] incorporate multiple data and forms of knowledge, extend its representational capabilities to incorporate non-cartographic information, support quantitative and qualitative forms of analysis, and illustrate that multiple epistemologies may be part of GIS-based research.*“ (ELWOOD und COPE 2009: 1)

Die GIS-basierte Visualisierungen und Analysen lassen sich damit nun auch auf qualitative Daten und Quellen anwenden, welche die wichtigste Grundlage für das Arbeiten in den Geistes- und Kulturwissenschaften darstellen (siehe Kapitel 5). Mögliche bearbeitbare Quellen sind dabei: „*Photographs, sketches, mental maps, social behavior maps, audio and video records, field notes, interviews, and narratives [...]*.“ (BAGHERI 2014: 168) Zusätzlich zu diesen Quellentypen sind noch Texte in unterschiedlichen Ausführungen, zum Beispiel als Tagebücher, Briefe, Zeitungsartikel oder Reiseberichte, zu nennen. Trotzdem eignen sich bestehende GIS-Ansätze mittels Vektor- und Rasterdatenformate aufgrund einer schwer durchzuführenden Kompatibilität nur bedingt dafür, solche qualitativen Quellen und Daten optimal zu visualisieren und analysieren. Mögliche Ansätze, um dies zu erreichen, sind eine Umwandlung der qualitativen

Daten in GIS-kompatible Datenformate, eine Verknüpfung von Features mit Hyperlinks, welche dann die qualitativen Daten anzeigen oder die Einführung qualitativer Analysetechniken in GIS. (vgl. SCHÄFER, MUENCHOW und HARNISCH 2018: 171; MCHAFFIE, HWANG und FOLLETT 2018: 306)

Qualitativ meint im Bereich der Qualitative GIS jedoch nicht nur Daten oder andere Quellen, welche nicht numerisch und meist ambivalent und von Widersprüchen geprägt sind. Qualitative Daten und Quellen haben reichhaltige kontextbezogene Details über die soziale und materielle Situation. Außerdem enthalten solche Daten oder Quellen auch Interpretationen von Situationen oder Prozessen, welche sie zu beschreiben versuchen. Trotzdem heißt das nicht, dass sich qualitative Daten räumlichen Visualisierungen oder Analysen entziehen, da GIS nur für quantitative Ansätze zugänglich wäre. (vgl. ELWOOD und COPE 2009: 3; BITTNER und MICHEL 2018: 152)

Zusätzlich geht es nicht nur um die Visualisierung dieser qualitativen Daten, sondern auch um die Analyse, was natürlich für GIS-Projekte besonders wichtig ist. Analysen sind qualitativ, wenn sie „[...] are intended to draw out the situated interpretive detail of qualitative data.“ (ELWOOD und COPE 2009: 4) Dabei werden qualitative Daten und Quellen einerseits dazu verwendet, räumliche Features zu kategorisieren, erklären oder prognostizieren und räumliche Muster und Zusammenhänge zu erkennen, welche durch andere Zugänge nicht entdeckbar wären. Zusätzlich ist es auch möglich, Emotionen besser abzubilden. Dafür können entweder bestehende GIS-Funktionen so angepasst werden, dass mit ihnen qualitative Daten gut nutzbar sind, oder bestehende so verwendet werden, dass eine induktive interpretative Visualisierung möglich ist. Qualitative Daten werden dabei mittels GIS erst erstellt, oder bereits bestehende Daten analysiert, visualisiert oder mit anderen räumlich verorteten Daten verlinkt. (vgl. HARRIS 2008: 355; ELWOOD und COPE 2009: 1f; WILSON 2009: 163; BITTNER und MICHEL 2018: 156)

Ziel ist es dabei, die Bedeutung dieser herauszuarbeiten. Dies gelingt durch eine systematische Kategorisierung der Daten, um Regelmäßigkeiten aufzudecken, oder durch eine Triangulation von verschiedenen Quellen, um dadurch Gemeinsamkeiten, Widersprüche und alle Nuancen dazwischen herauszuarbeiten. Dadurch werden qualitative Techniken der Visualisierung, Interpretation, Reflexion und Kontextualisierung ermöglicht. Außerdem finden die Wissensproduktion und Verortung zumeist idiographisch statt. Dabei gilt es immer den Prozess der Wissensgenerierung, also das Design des Ablaufs, die zugrundeliegenden Daten, die

Analyseschritte und deren Ergebnisse, kritisch zu reflektieren, da diese, wenn sie einer bestimmten Konvention folgen, zu einseitig beeinflussten Ergebnissen führen können. Die erzeugten Ergebnisse sind grundsätzlich immer als offen und unabgeschlossen zu sehen. „*Qualitative GIS betonen die Perspektivität und Situiertheit räumlichen Wissens und führen [...] heterogene und kontextuell reichhaltige räumliche Daten zusammen.*“ (ebd.: 152) Dabei stehen immer die Produktion und Bedeutung des Raums im Mittelpunkt, wobei der Raum nie als statisch und absolut gesehen werden sollte. (vgl. ELWOOD und COPE 2009: 4; COPE und ELWOOD 2009: 175f; PAVLOVSKAYA 2009: 31; WILSON 2009: 163; BITTNER und MICHEL 2018: 154)

In den letzten Jahren wurde durch einige methodische Forschungsarbeiten eine Verbesserung in diesem Bereich erreicht, da es nun teilweise möglich ist, qualitative Daten in GIS einzubinden und auszuwerten. Multimediale und personenbezogene räumliche Daten sind nun Teil des GIS-Arbeitsprozesses. (vgl. ebd.: 156)

Trotzdem erkennen BITTNER und MICHEL (2018) noch immer eine gewisse Diskrepanz zwischen „[...] *situiertem und kontextuellen [sic!] Wissen einerseits und informatischen Formalisierungen andererseits [...]*“ (ebd.: 162). Da GIS aber vermehrt Funktionen für quantitative Daten bietet, welche genau verortbar sind, besteht immer die Gefahr, dass bei der Arbeit mit qualitativen Daten „[...] *die dem Untersuchungsgegenstand angemessensten Erhebungsverfahren zugunsten von GIS-Kompatibilität aus dem Blick [...]*“ (vgl. ebd.: 163) verloren werden können.

PRESTON und WILSON (2014) stellen daher fünf Anforderungen an eine bessere Arbeit mit Qualitative-GIS-Ansätzen:

- Breite: Die verwendeten Daten sollten in ihrem Format und Inhalt weit gefächert sein, dass unterschiedliche Perspektiven auf den Forschungsgegenstand gewonnen werden können. Dadurch können auch mögliche Datenschutzbedenken ausgeschlossen werden.
- Tiefe: Unterschiedliche Darstellungsformen und verschiedene Sichtweisen und -ebenen ermöglichen es, eine inhaltliche und kontextuelle Tiefe zu erreichen. Dies kann unter anderem durch Interviews oder partizipative Erhebungsmethoden erreicht werden.
- Iterativität und Flexibilität: „*Sowohl die Erhebung als auch die Auswertung des Materials sollten iterativ und flexibel gestaltet werden und eher einen explorativen Charakter*

haben, um auf dabei entstehende Fragen und Probleme reagieren zu können.“ (BITTNER und MICHEL 2018: 155) Diese Offenheit ermöglicht es, marginalisierte Erzählungen und Sichtweisen stärker zu betonen und dadurch einen ganzheitlicheren Eindruck des Forschungsgegenstandes zu erhalten.

- Visualisierbarkeit: Daten und Ergebnisse von Analysen, welche im Zuge eines Qualitative-GIS-Projektes verwendet und erzeugt wurden, sollten immer auch räumlich visualisiert werden, da diese Visualisierung (als Karte oder andere Darstellungsformen) ein einfaches Verständnis für den Forschungsgegenstand ermöglicht.
- Qualitative GIS als Prozess: Nicht nur das Produkt des Forschungsprozesses, sondern der Prozess an sich ist wichtig, da nicht unbedingt kohärente Antworten auf Qualitative-GIS-Fragestellungen gefunden, sondern bestehende Positionen und Annahmen hinterfragt werden können. Dadurch werden Beteiligte zur Reflexion des Forschungsgegenstandes und des Arbeitsprozesses angeregt.

(vgl. PRESTON und WILSON 2014: 520ff; BITTNER und MICHEL 2018: 155)

Diese Anforderungen können laut PRESTON und WILSON (2014) jedoch auch als Limitationen gesehen werden, da durch die Iterativität und die Ansicht von Qualitative GIS als Prozess nie wirklich ein Abschluss und dadurch ein feststehendes Ergebnis gefunden werden kann. Deshalb müsste die Untersuchung ad infinitum durchgeführt werden. Solange man sich dieses Umstandes jedoch bewusst ist, dass man keine endgültigen und unumstößlichen Aussagen tätigen kann und dies auch in der eigenen Arbeit festhält, kann die Arbeit auch abgeschlossen werden. Des Weiteren warnen PRESTON und WILSON (2014) vor der Unordnung eines solchen explorativen Forschungsansatzes. Daher muss während des ganzen Prozesses reflexiv gearbeitet werden, um mögliche Problemstellen rechtzeitig zu entdecken und gegebenenfalls den Forschungsprozess anpassen zu können.

4.1. Qualitative Humanities und Qualitative Historical GIS

Aufgrund der stark verbreiteten Verwendung von qualitativen Quellen und Daten in den Geistes- und Kulturwissenschaften, sowie vor allem auch in der Geschichtswissenschaft, soll nun nachfolgend auf Probleme und das Potential von Qualitative-GIS-Ansätzen in diesem Bereich (Humanities GIS und Historical GIS) eingegangen werden. Schon GREGORY und ELL (2007) und

auch John CORRIGAN (2010) sehen eine erhöhte Verwendung von qualitativen Daten in GIS und dadurch eine vielversprechende Möglichkeit, dass Geistes- und Kulturwissenschaftler*innen und vor allem Geschichtswissenschaftler*innen ihre Arbeitsweise erweitern, neu gestalten und mit neuen Ansätzen durchführen können, da nun auch eine informationstechnologisch bearbeitbare räumliche Komponente Teil ihrer Forschungsarbeit werden kann. Es wird dabei ermöglicht, GIS zu nutzen, „[...] *to locate people in time and space, and to explore changing spatial patterns through a variety of attributes.*“ (GREGORY und ELL 2007: 197) Also wird GIS nicht nur als Tool für quantitative Datenbearbeitung gesehen, sondern auch das Potential, um mit qualitativen Daten und Quellen neue Arbeitsmethoden zu entwickeln.

Auch GREGORY und GEDDES (2014a) sehen hier ein großes Potential, auch wenn sie anmerken, dass zur Erreichung derselben Selbstverständlichkeit wie in der Bearbeitung von quantitativen Daten noch einiges an Arbeit geleistet werden muss. Aber schon jetzt profitieren beide Seiten, also Geschichtswissenschaftler*innen und GIS-Expert*innen, von dieser interdisziplinären Zusammenarbeit, da für beide Forschungsbereiche neue Ansätze und methodische Vorgangsweisen entwickelt und etabliert werden.

Trotzdem ist das Feld der Historical GIS im Bereich der Geschichtswissenschaft noch ein relativ kleiner Bereich, da noch immer vor allem mit quantitativen Daten gearbeitet wird. Daher ist auch die Entwicklung neuer qualitativer Arbeitsansätze so wichtig, da qualitative Quellen ebenjene der Geschichtswissenschaft sind und daher nur mit quantitativen Ansätzen keine weitgreifendere Verbreitung der Historical GIS möglich sein wird. Außerdem können durch diesen Ansatz neue Perspektiven und neues Verständnis für räumliche Beziehungen und Verhältnisse im Bereich der Geschichtswissenschaft gewonnen werden. (vgl. BODENHAMER 2013: 7) Von dieser Entwicklung kann aber auch der Bereich GIScience profitieren, da neu entwickelte Forschungsmethoden im Bereich Historical GIS auch in anderen GIS-Bereichen eingesetzt werden können. (vgl. GREGORY und GEDDES 2014a: 174f)

Um nun eine Weiterentwicklung in beiden Fachbereichen erreichen zu können, müssen noch einige Problemfelder angegangen und überwunden werden. Einerseits geht es darum, räumliche Datenbanken zu entwickeln, in welchen qualitative Daten effektiv gespeichert werden können. Außerdem muss es ermöglicht werden, dass Attribute, welche nicht traditionellen Formaten entsprechen, zum Beispiel Bilder, Videos, Tonaufnahmen und textliche Quellen, in solchen Datenbanken zu Features hinzugefügt werden können. Dies liegt daran, dass diese

Quellentypen täglich Brot für Geschichtswissenschaftler*innen sind und daher auch in der Arbeit mit GIS eingebunden werden können müssen. Außerdem sollten diese Sammlungen auch georeferenzierbar sein. Dadurch kann die Bedeutung der Quellen und ihr räumlicher Horizont mittels Karten visualisiert und unterstrichen werden. Diese Verortung ist, aufgrund sich in der Zeit ändernder Grenzen, meist mittels Punktkoordinaten, welche Orte darstellen, einfacher. Eine einfache Verortung dieser ist mittels Gazetteer erreichbar (siehe Kapitel 6.1.3.1). Des Weiteren ist es wohl am wichtigsten, eine effiziente und umfangreiche Arbeit mit textlichen Quellen in GIS zu ermöglichen (siehe Kapitel 6.1.2). Solche Quellen sind die in der Geschichtswissenschaft am häufigsten genutzten Quellentypen und sollten daher auch in einer geschichtswissenschaftlichen Arbeit mit GIS ihren unverzichtbaren Platz finden. (vgl. GREGORY und HEALEY 2007: 647; GREGORY und GEDDES 2014a: 175f; BODENHAMER 2013: 7; PALMER 2013: 46) Diese können dann unter anderem mittels der Zusammenführung und gemeinsamen Umsetzung der Bereiche „[...] *knowledge representation, data mining, natural language processing, machine learning, information retrieval and visual analytics*“ (OWENS und WACHOWICZ 2013: 140) (siehe auch Kapitel 6.1) weiter bearbeitet und analysiert werden.

Außerdem nennen GREGORY und GEDDES (2014a) noch die Entwicklung von Online-Sammlungen von historischen Karten, welche auch georeferenziert abrufbar sind. Solche Sammlungen wurden bereits in den letzten Jahren entwickelt (vgl. Cartography Associates 2021; University of Portsmouth o.J.) und daher ist dieser Punkt schon als abgeschlossen zu betrachten.

5. Quellen in der Geschichtswissenschaft und Quantitative vs. Qualitative Daten in der Geschichtswissenschaft und in der Geoinformation

5.1. Quellen

In der Geschichtswissenschaft wird mit Quellen gearbeitet, welche als Grundlage für die Forschung dienen. *„Alle Quellen sind Medien, aber nicht alle Medien sind Quellen. [...] Medien sind alle Funktionen kultureller Praktiken und Objekte, die von Subjekten genutzt werden, um Informationen über Alterität zu gewinnen oder an sie zu richten.“* (TSCHIGGERL, WALACH und ZAHLMANN 2019: 81) Das bedeutet, dass Quellen immer im Kontext kultureller Praktiken gesehen werden müssen, wobei diese immer dazu dienen Informationen über Alterität, also alles, was man nicht selbst ist, zu gewinnen. Also nehmen Medien und dadurch auch Quellen eine Art Vermittlerfunktion zwischen dem*der Betrachter*in und etwas anderem ein. Jedoch muss darauf geachtet werden, dass *„ein Medium [...] nur dann ein Medium [ist], wenn es von jemandem in diesem Sinne genutzt wird.“* (ebd.: 81) Außerdem liegt die Bedeutung von Medien und dadurch auch von Quellen immer in ihrem Gebrauch, was bedeutet, dass eine Quelle erst durch die Auswahl und Verwendung als solche zu einer Quelle wird, da ihr dadurch Bedeutung zugewiesen wird. (vgl. ebd.: 86ff) Dabei sollte jedoch darauf geachtet werden, dass schon durch die Auswahl der Quellen die Forschungsergebnisse entscheidend beeinflusst werden. Und auch die Verfügbarkeit von Quellen hat einen entscheidenden Einfluss. Außerdem werden Quellen dabei immer von einer Gegenwartsperspektive aus betrachtet. (vgl. BUDDE 2010: 54f; WIERLING 1995: 47f) *„Wie bei jedem anderen Medium auch – und eine Quelle ist schließlich immer ein Medium – liegt die Bedeutung der Quelle in ihrem Gebrauch und ist ausschließlich von dem sie nutzenden Subjekt abhängig.“* (TSCHIGGERL, WALACH und ZAHLMANN 2019: 97) Dies zeigt, dass eine Quelle nicht als solche von selbst existiert, sondern durch den*die Wissenschaftler*in erst zu einer solchen gemacht wird.

Quellen ermöglichen es, Aussagen über die Vergangenheit zu tätigen. Sie geben Thesen ein Fundament, welches diese erst glaubwürdig macht und fundierte Darstellungen ermöglicht. Außerdem dienen sie dazu, fehlerhafte Thesen zu widerlegen. Da jedoch viele Quellen in ihrer Aussage nicht eindeutig sind, und daher auch nicht für sich sprechen können, müssen Fragen in diese gestellt werden, um eine Verbindung zur Vergangenheit herstellen zu können, ihre Wichtigkeit herauszustreichen und eine Validität festzustellen (vgl. BUDDE 2010: 54; WIERLING 1995: 47f):

- Wer hat die Quelle verfasst?
- Wann entstand die Quelle?
- Wo wurde die Quelle erstellt?
- Welche Art von Quelle ist es?
- Wen hat die Quelle als Adressaten im Visier?
- Wie ist die Quelle überliefert?
- Warum wurde sie erstellt?
- Wovon kündigt die Quelle, wovon schweigt sie?

(BUDDE 2010: 67)

In der Geschichtswissenschaft gibt es viele verschiedene Quellentypen, welche sich auch je nach Herkunftszeitraum unterscheiden können: Archäologische Funde, Inschriften, Münzen, Biographien, Fachliteratur zu beispielsweise: Rhetorik, Religion oder Hauswirtschaft, Annalen, Chroniken, Vitae, Hagiographien, Briefe, Predigten, Urkunden, Genealogien, Güterverzeichnisse, Inventare, Rechnungen, Testamente, Gerichtsprotokolle, Inquisitionsverfahren, Ego-Dokumente, Photographien, Tonaufnahmen, Filme, uvm. (vgl. ebd.: 57f)

Ego-Dokumente, welche eine der wichtigsten Quellentypen der Geschichtswissenschaft darstellen, sind dabei Texte, welche als Kriterium haben, dass

„[...] Aussagen oder Aussagenpartikel vorliegen, die – wenn auch in rudimentärer und verdeckter Form – über die freiwillige oder erzwungene Selbstwahrnehmung eines Menschen in seiner Familie, seiner Gemeinde, seinem Land oder seiner sozialen Schicht Auskunft geben oder sein Verhältnis zu diesen Systemen und deren Veränderungen reflektieren. Sie sollten individuell-menschliches Verhalten rechtfertigen, Ängste offenbaren, Wissensbestände darlegen, Wertvorstellungen beleuchten, Lebenserfahrungen und -erwartungen widerspiegeln.“ (SCHULZE 1996: 28)

Neben der Einteilung in unterschiedliche Quellentypen, ist es auch möglich, einzelne dieser nochmals genauer zu unterscheiden. Textliche Quellen lassen sich so durch verschiedene Eigenschaften charakterisieren:

Tabelle 2: Einteilung textlicher Quellen

handschriftlich	gedruckt
analog vorliegend	digital vorliegend
manuell transkribiert	automatisch transkribiert (OCR)
standardisiert	nicht standardisiert/erzählend
deutschsprachig	nicht deutschsprachig

Zusätzlich ist es noch möglich, textliche Quellen unter anderem nach dem Texttyp und dem Herkunftszeitraum einzuteilen.

Quellen sollten immer kritisch betrachtet werden, um einen Missbrauch dieser im Sinne einer Umschreibung oder -deutung der Geschichte zu verhindern. (vgl. BUDDE 2010: 68) Diese sogenannte Quellenkritik hilft dabei, den historisch wissenschaftlichen Wert einer Quelle zu bestimmen. Dabei ist die wichtigste Frage die nach der Echtheit, nämlich ob die Quelle wirklich aus der angegebenen Zeit stammt, aus welcher sie zu sein scheint. (vgl. JORDAN 2018: 45) Anlehnend an DROYSEN (1925) sollen dann auch noch die Kritik der Richtigkeit und die Kritik des Früheren und Späteren geleistet werden. Erstere bezeichnet eine Überprüfung des Materials, ob es als Beleg dafür gelten kann, was es darzustellen versucht. Zweitere hinterfragt die zeitliche Einordnung des Materials und ob es noch unverändert vorhanden ist.

5.2. Quantitative Forschung in der Geschichtswissenschaft

In der Geschichtswissenschaft arbeiten vor allem Wirtschaftshistoriker*innen mit quantitativen Daten, um empirische Sachverhalte zu zählen und als Zahlen darzustellen. Dies ist nur mit Sachverhalten möglich, welche sich numerisch erfassen oder in numerische Relationen umwandeln lassen. Damit wird versucht, die Wirklichkeit abzubilden, aber es werden nur einzelne Merkmale verwendet, während andere nicht beachtet werden. „*Quantitative Untersuchungen erheben Einstellungen, Orientierungsmuster, Handlungen und Strukturen auf standardisierte Weise, um sie dann in statistisch verarbeitbare Zahlen zu übersetzen.*“ (Schirmer und Blinkert 2009: 67) Sie verwenden unter anderem Bevölkerungsdaten oder Versorgungsdaten, um mittels der Berechnung meist einfacher statistischer Kennzahlen, wie zum Beispiel

Häufigkeit, Mittelwerte, Verteilungen, Prozentangaben, Wahrscheinlichkeiten oder Streuungsmaße, Erkenntnisse zu erlangen und unterschiedliche Variablen Zusammenhänge abzubilden. Diese Messungen beziehungsweise Berechnungen müssen folgenden Gütekriterien genügen: Objektivität, Reliabilität, Validität und Repräsentativität. (vgl. Budde und Freist 2010: 162ff; SCHIRMER und BLINKERT 2009: 66ff)

Außerdem findet diese Methode auch noch Anwendung in der historischen Mobilitätsforschung, wo durch standesamtliche Personenstandsregister auf Umzüge und Berufe geschlossen werden kann, und auch historische Wahlforschung bedient sich Methoden dieses Forschungsbereiches. Jedoch sollte dabei immer beachtet werden, dass durch solche Berechnungen Informationen verloren gehen, da sie sich nicht in Kennzahlen ausdrücken lassen. Dadurch wird die Aussagekraft immer reduziert sein, die Bearbeitungsfähigkeit wird aber erhöht. (vgl. Hug und Poscheschnik 2015: 86ff; Schirmer und Blinkert 2009: 66ff; Schuurman 2009: 149)

5.3. Qualitative Forschung in der Geschichtswissenschaft

Im Gegensatz zu quantitativen Forschungsansätzen, werden qualitative dazu verwendet, „[...] Ursachen bestimmter Verhaltensweisen herauszufinden und um Erfahrungen, Wahrnehmungen und Deutungen historischer Akteure zu analysieren.“ (BUDDE und FREIST 2010: 165) Hierfür werden der Mensch und seine Lebenssituation mittels einer holistischen Sichtweise betrachtet und auch nur in dieser Ganzheit verstanden. (vgl. SCHUURMAN 2009: 149) Ziel der Forschung sind dabei unter anderem „[...] soziale Regeln, kulturelle Orientierungen und individuelle Sinnstrukturen.“ (HUG und POSCHESCHNIK 2015: 89) Die Qualitätskriterien in der qualitativen Forschung sind: Offenheit gegenüber dem Forschungsgegenstand und in ihrem eigenen Handeln, Reflexivität in Bezug auf die Beforschten und des Forschenden sowie eine Iterativität des Forschungsprozesses. (vgl. SCHIRMER und BLINKERT 2009: 78ff) Dabei werden unterschiedliche Quellentypen, zum Beispiel bildliche Darstellungen, Ego-Dokumente oder Gerichtsakten, analysiert. Die Hauptmedien für die Arbeit in qualitativer Forschung sind Texte, welche Erkenntnisse ermöglichen, aber auch für Interpretationen offenstehen. Wichtig ist dabei zu beachten, dass die Inhalte der Texte nie wirklich sind, sondern die Wirklichkeit immer nur abbilden. Mögliche Methoden der qualitativen Arbeit mit Quellen sind die Diskursanalyse, qualitative Inhaltsanalyse, Bildinterpretation oder Oral History. (vgl. BUDDE und FREIST 2010: 165ff; SCHIRMER und BLINKERT 2009: 75f)

5.4. Vergleich quantitative und qualitative Forschungsansätze in der Geschichtswissenschaft

„Jede Methode – egal ob qualitativ oder quantitativ – ist für bestimmte Forschungsgegenstände und Forschungsfragen besser und für andere schlechter geeignet.“ (HUG und POSCHESCHNIK 2015: 86) Dieser wichtige Leitsatz ist beim Vergleich qualitativer und quantitativer Forschungsansätze immer zu beachten. Das bedeutet, dass weder qualitativ noch quantitativ zu arbeiten besser ist, sondern nur jeweils geeigneter für einen gewissen Forschungsgegenstand.

In der Geschichtswissenschaft finden vor allem qualitative Forschungsansätze ihren Einsatz. Jedoch werden in manchen Bereichen auch quantitative eingesetzt. Ein historischer Vergleich von verschiedenen Ländern wurde zu Beginn der Entwicklung dieses Forschungsbereiches zu meist quantitativ durchgeführt, da Massendaten zur Wirtschaft oder Bevölkerung vorhanden waren, auch wenn diese aufgrund unterschiedlicher Aufnahmeverfahren eigentlich nicht wirklich vergleichbar waren. Daher begann man auch wegen des *cultural turns* immer mehr dazu darüber hinaus Vergleiche anzustellen, welche sich nun auch qualitativer Methoden bedienen. (vgl. BUDDE und FREIST 2010: 174) Hier kann auch der Bereich Historical GIS Teil der Entwicklung sein, indem nicht nur mit quantitativen Methoden gearbeitet wird, sondern auch qualitative Ansätze (siehe Kapitel 4) entwickelt werden.

HUG und POSCHESCHNIK (2015) stellen die wichtigsten Eigenschaften und Unterschiede quantitativer und qualitativer Forschung gegenüber:

Tabelle 3: Vergleich Quantitative und Qualitative Forschung (basierend auf: HUG und POSCHESCHNIK 2015: 86; SCHUURMAN 2009: 148)

Quantitative Forschung	Qualitative Forschung
Quantifizierung von Sachverhalten	Verstehen von Sinn
große Stichproben, Repräsentativität	einzelne Fälle
linearer Ablauf	iterativer, zirkulärer Ablauf
standardisiertes Vorgehen	offenes, flexibles Vorgehen
interessiert an objektiven Fakten	interessiert an subjektiven Sichtweisen
Laborforschung	Feldforschung
eher hypothesentestend	eher hypothesengenerierend
materialistisch-deterministisches Welt- und Menschenbild	humanistisches Menschenbild
Betrachtung einzelner Merkmale	holistische Sichtweise
Analyse von Massendaten	Betonung der Analyse von Einzelfällen

Quantitative Forschung geht also eher standardisiert vor, quantifiziert dabei Sachverhalte und drückt diese mittels statistischer Kennzahlen aus. Sie will dadurch Objektivität erzeugen, welche eigentlich sowieso nie zu erreichen ist, da der*die Forscher*in selbst immer als Subjekt Teil des Forschungsprozesses ist. In der qualitativen Forschung wird hingegen versucht, Erfahrungen und Verhaltensweisen zu verstehen, indem subjektive Sichtweisen im Zentrum der Untersuchung stehen. Dabei werden meist Einzelfälle betrachtet. Die Unterscheidung von Laborforschung und Feldforschung ist dabei in der Geschichtswissenschaft und in der Geoinformation nicht wirklich anwendbar. Während in der Geschichtswissenschaft fast nie eine Laborforschung stattfindet, sondern zumeist mit Quelle qualitativ oder quantitativ gearbeitet wird, können in der Geoinformation beide Forschungstypen sowohl im „Labor“, also am Computer“, und im „Feld“, also vor Ort, stattfinden.

5.5. Quantitative und Qualitative Daten in GIS

Auch im Bereich der Geographischen Informationssysteme (GIS) wird mit quantitativen und qualitativen Daten gearbeitet. Während die Arbeit mit qualitativen Daten bisher zumeist stark eingeschränkt und untergeordnet stattfindet (siehe Kapitel 4), sind quantitative Daten die wichtigsten für geoinformatische Forschungsansätze.

Daten, welche in GIS verwendet werden, lassen sich gut nach dem Skalenniveau einteilen:

- Nominaldaten: Dies sind Daten, welche sich nur durch Gleichheit/Unterschiedlichkeit unterscheiden lassen bzw. keiner Rangfolge unterliegen. Eine Reihung dieser ist nicht möglich. Beispiel: Bodenbedeckung (Wasser, Wald, Ackerland)
- Ordinaldaten: Diese Daten lassen sich ordnen (größer/kleiner), sind jedoch keine numerischen Daten, das heißt, eine Anwendung statistischer Methoden oder mathematischer Operatoren ist nicht möglich. Beispiel: Erosionsgefahr (gering, mittel, hoch)
- Intervalldaten: Diese Daten bestehen aus relativen Zahlenwerten, welche sich auf einer linearen Skala befinden, bei welcher der Nullpunkt willkürlich festgelegt wurde. Mathematische Operationen und die Anwendung statistischer Berechnungen sind mit

diesen Daten möglich. Auch negative Werte sind als Ergebnis von Berechnungen möglich. Beispiel: Temperaturwerte in Grad Celsius (-10°C , 0°C , 25°C)

- Ratiodaten: Ratiodaten sind Daten, welche absolute Werte auf einer linearen Skala darstellen. Dies bedeutet, dass ein absoluter Nullpunkt vorhanden ist und dadurch die Berechnung von Proportionalitäten und relative Vergleiche möglich sind. Beispiel: Alter (Kein Alter unter 0 möglich, dass doppelte Alter jedoch schon)

(vgl. BARTELME 2005: 148; CLEVE und LÄMMEL 2020: 43f)

Nominal- und Ordinaldaten können dabei als qualitative Daten gesehen werden, welche sich aber noch immer in Attributen abbilden lassen, während Intervall- und Ratiodaten klar quantitativ zu verstehen sind.

6. Data Mining

Um also die räumliche Komponente aus dem in der Geschichtswissenschaft so wichtigen Quellentyp der Texte zu extrahieren, soll im nun folgenden Kapitel „Data Mining“ besprochen werden. Im Bereich „Data Mining“ gibt es unterschiedliche Vorgehensweisen, um räumliche Informationen aus Texten zu erhalten. Diese werden anschließend an einen Überblick über Data Mining an sich besprochen.

Daten bezeichnen im Kontext des Data Mining eine „[...] *Ansammlung von Zeichen mit der dazugehörigen Syntax.*“ (CLEVE und LÄMMEL 2020: 41) Diese können in drei Kategorien unterschieden werden:

- Unstrukturierte Daten: Dies sind zum Beispiel Bilder oder Texte. Mögliche Verarbeitungsschritte ist das Text Mining, welches, da in der Geschichtswissenschaft vor allem textliche Quellen verwendet werden, anschließend näher betrachtet werden soll (siehe Kapitel 6.1).
- Semistrukturierte Daten: Webseiten, welche zwar vor allem Text beinhalten, aber immer eine gewisse Struktur aufweisen, werden als semistrukturierte Daten bezeichnet. Diese können mittels Web-Mining bearbeitet werden.
- Strukturierte Daten: Daten, welche in relationalen Datenbanken oder Tabellen in einer festen Struktur gespeichert sind, werden als strukturierte Daten bezeichnet. Data Mining im engeren Sinn beschäftigt sich mit solchen Daten.

(vgl. ebd.: 41ff)

Data Mining bezeichnet also die „[...] *Extraktion von Wissen aus Daten.*“ (ebd.: 42) Dabei sollen neue Erkenntnisse geschaffen werden, welche für weitere Arbeiten in den jeweiligen Forschungsgebieten nützlich sein können. Hierfür können auch schon vorhandene Datenquellen neu verwendet werden, wenn dadurch neue Erkenntnisse gewonnen werden können. Ziel ist es außerdem, dass der Prozess so automatisiert wie möglich ablaufen sollte, auch wenn es dafür einiges an Vorarbeit benötigt. (vg. ebd.: 42; WITTEN, FRANK und HALL 2011: 4f)

Beim Data Mining wird nach unbekannten Mustern, Zusammenhängen, Abhängigkeiten oder anderen Informationen in Daten gesucht. Dabei kann es das Ziel sein, die Datensätze zu klassifizieren, zu clustern oder andere Informationen aus ihnen zu gewinnen. (vgl. CLEVE und

LÄMMEL 2020: 1) Data Mining bezeichnet zumeist nicht nur den eigentlichen Schritt des Data Minings, sondern auch vorangehende Schritte und die nachfolgende Evaluation (siehe Aufzählung unten).

Der Prozess des Data Mining sieht wie folgt aus:

1. Selektion: Am Anfang werden die für die nachfolgende Verarbeitung geeigneten Daten ausgewählt, welche nachfolgend weiter bearbeitet werden sollen. Dabei ist es auch möglich, eine „[...] repräsentative Teildatenmenge des Gesamtdatenbestands“ (ebd.: 8) auszuwählen. Hierbei sollte immer reflexiv vorgegangen werden: Sollten die gewählten Daten in einem weiteren Schritt nicht verwendbar sein, müssen die Kriterien der Datenauswahl angepasst werden.
2. Datenvorverarbeitung: In diesem Schritt wird zuerst die Qualität der vorhandenen Daten untersucht und, falls notwendig, fehlerhafte oder fehlende Daten bearbeitet. Sie werden bereinigt, also fehlende Werte ersetzt oder fehlerhafte korrigiert. Dies ist notwendig, um falsche oder verzerrende Ergebnisse zu verhindern.
3. Transformation: Die Daten werden in adäquate Datenformate umgewandelt, da nicht alle Datenformate mittels des folgenden Prozesses weiterverarbeitet werden können. Dabei ist es auch möglich, benötigte Attribute oder Datensätze neu zu generieren. Dies kann notwendig sein, falls numerische Daten vorhanden sind, aber nominale benötigt werden, oder umgekehrt. Außerdem sind Daten aus unterschiedlichen Quellen aufeinander abzustimmen.
4. Data Mining: Nun findet das eigentliche Data Mining statt, indem ein Modell zur Bearbeitung ausgewählt oder erstellt und dann angewandt wird. Dabei wird in den Daten nach Mustern gesucht, beziehungsweise neue Informationen werden extrahiert.
5. Interpretation und Evaluation: Die Ergebnisse des vorangegangenen Schrittes werden anschließend überprüft, bewertet und interpretiert, um mögliche neue Erkenntnisse zu bewerten und daraus Handlungsmaßnahmen ableiten zu können. Dabei müssen die folgenden Kriterien beachtet werden: *Validität* – liefert das Modell auch mit neuen Daten gültige Erkenntnisse, *Neuartigkeit* – wie steht die neue Erkenntnis zu bisherigem Wissen, *Nützlichkeit* – gibt es einen praktischen Nutzen für Anwender*innen, *Verständlichkeit* – ist die Erkenntnis vom* von der Anwender*in auch zu verstehen.

(vgl. AGGARWAL 2015: 3f; CLEVE und LÄMMELE 2020: 2ff; WITTEN, FRANK und HALL 2011: 52)

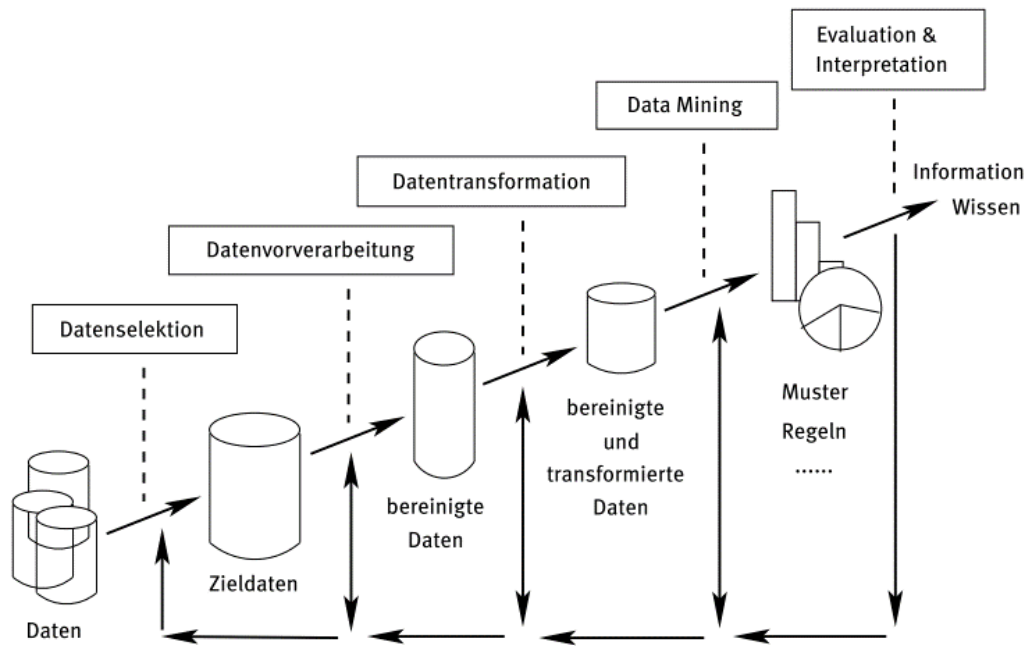


Abbildung 1: Data-Mining-Prozess (Quelle: CLEVE und LÄMMELE 2020: 3)

Dabei ist der Prozess nicht als Einbahn, sondern als reflexiver Kreislauf zu sehen. (siehe Abbildung 1) Im Laufe der einzelnen Schritte sind die vorangegangenen nicht als abgeschlossen zu betrachten, sie werden immer wieder hinterfragt und Erkenntnisse aus nachfolgenden in die vorangegangenen eingebaut. Dadurch wird eine optimale Umsetzung eines Projektes mittels Data Mining möglich.

6.1. Text Mining

„Text mining can be defined as a knowledge-intensive process in which a user interacts with a collection of documents by using analytic tools in order to identify and explore interesting patterns.“ (ŽIŽKA, DAŘENA und SVOBODA 2019: 1f) Es ist ein Subtyp des Data Mining. Dabei werden Texte beziehungsweise Textdokumente, welche als unstrukturiert bezeichnet werden (siehe oben), in die Data-Mining-Prozesse einbezogen. Dokumente der menschlichen Kommunikation liegen in einer unmessbaren Quantität, verschiedener Komplexitätsstufen und aus unterschiedlichsten Quellen vor, sodass eine umfassende, zeitnahe und zufriedenstellende Analyse durch einzelne Personen unmöglich ist. Daher werden diese mittels informationstechnologischer Methoden auf mögliche Muster und Regelmäßigkeiten oder unerwartete Unregelmäßigkeiten überprüft. Als Ergebnis sollen dabei nützliche Informationen extrahiert und neue

Erkenntnisse generiert werden. Ziel dabei ist es natürlich, eine Informationsdichte ähnlich der durch menschliche Analysen zu erreichen oder sogar Informationen darüber hinaus zu erlangen und das in einer arbeitsextensiveren und zeitsparenden Weise. Doch auch Anwendungen von geringerer Genauigkeit und Qualität können relevante Ergebnisse liefern. Mögliche Anwendungsbereiche sind zum Beispiel eine Klassifizierung von Dokumenten nach Themen, was durch eine Extraktion von relevanten Begriffen und Schlüsselwörtern erreicht wird. Texte können mittels Text Mining auch zusammengefasst oder Trends zwischen unterschiedlichen Dokumenten erkannt werden. Außerdem können auch neu digitalisierte Dokumente, welche zuvor nur analog zugänglich waren, nun auch informationstechnologisch analysiert werden. (vgl. CLEVE und LÄMMEL 2020: 69; FELDMAN und SANGER 2007: 1; ebd.: 94; KWARTLER 2017: 1ff; ŽIŽKA, DAŘENA und SVOBODA 2019: 1ff)

Relevante Dokumententeile, welche analysiert werden können, sind:

- Characters: Dies sind Zeichen wie Buchstaben, Ziffern, Leer-, Satz- oder andere Sonderzeichen. Eine Analyse dieser ist zwar möglich, jedoch meist schwer handzuhaben und zu interpretieren.
- Words: Bei diesen handelt es sich um „[...] *the basic level of semantic richness*.“ (FELDMAN und SANGER 2007: 5) Hierbei müssen bei Analysen jedoch einige Einschränkungen (siehe unten) beachtet werden, um ein aussagekräftiges Ergebnis zu erzielen.
- Terms: Terms sind einzelne Wörter und Phrasen, welche direkt aus dem Dokument mittels „term extraction“-Methoden und „speech tagging“ gefiltert wurden.
- Concepts: Dabei werden Features aus Texten extrahiert, welche bestimmten „concept identifiers“ genügen, also meist eine inhaltliche Konsistenz aufweisen. Diese können von einzelnen Wörtern bis zu ganzen Absätzen reichen.

(vgl. ebd.: 5ff)

Dokumente und textliche Quellen bilden menschliche Aktivitäten ab, da sie Teil der zwischenmenschlichen Kommunikation sind. Um diese nun weiterverwenden oder analysieren zu können, werden sie getaggt, also mit bestimmten Informationen gekennzeichnet, um wichtige Entitäten und deren Beziehungen (bezeichnet als „facts“ und „events“) extrahieren zu können. Diese sollen dabei sinnvoll und inhaltlich relevant sein. (vgl. ŽIŽKA, DAŘENA und SVOBODA 2019: 1; FELDMAN und SANGER 2007: 94) Die dabei wichtigsten Informationen sind:

- Entities: Dies sind grundlegende Bauteile von Texten. Beispiele hierfür sind Personen, Unternehmen oder Ortsangaben. Die Ortsangaben sind für räumliche Analysen besonders relevant. (siehe Kapitel 6.1.2)
- Attributes: Attribute sind Teile der extrahierten Entitäten. Dies können zum Beispiel der Titel oder das Alter von Personen oder Unternehmenstypen sein.
- Facts: Dies bezeichnet Beziehungen zwischen einzelnen Entitäten. Ein Beispiel hierfür ist eine Arbeitsbeziehung zwischen einer Person und einem Unternehmen an einem bestimmten Ort.
- Events: Events sind Aktivitäten oder Ereignisse von Interesse, an welchem Entitäten teilhaben. Dies können zum Beispiel ein Geburtstag einer Person oder eine Zusammenschließung von Unternehmen sein.

(vgl. ebd.: 96)

Um die wichtigsten Informationen aus Texten zu extrahieren, ist vor allem eine Durchführung der sogenannten „Named Entity Recognition“ relevant. Dabei werden alle relevanten Informationen wie zum Beispiel Namen, geographische Ortsangaben, Organisationen, Daten, Zeitangaben oder bestimmte Zahlenwerte wie Geldmengen oder Prozentangaben bestimmt und getaggt. (vgl. ebd.: 96)

Besonders wichtig bei der Durchführung von solchen Analysen ist dabei die Beachtung von Einschränkungen, welche sich durch die sprachlichen Eigenheiten einzelner Sprachen ergeben. Auch die zunehmende Komplexität durch eine Kombination Tausenden von Wörtern in unterschiedlichen Kombinationen ist ein Problem für eine zufriedenstellende und allumfassende Analyse. Einerseits müssen irrelevante Wörter („stop words“) ausgeschlossen und Abkürzungen und Synonyme als solche erkannt werden. Zusätzlich beeinträchtigen Schreib- oder grammatikalische Fehler eine genaue Analyse. Außerdem müssen, vor allem in der deutschen Sprache, Wortvarianten aufgrund von grammatikalischen Umformungen als solche erkannt und auf den Wortstamm zurückgeführt werden („Stemming“). Mit der nun reduzierten Wortmenge („bag of words“) wird anschließend die eigentliche Analyse durchgeführt. (vgl. CLEVE und LÄMMEL 2020: 69f; ŽIŽKA, DAŘENA und SVOBODA 2019: 4)

6.1.1. Data/Text Mining in den Digital Humanities

Vor allem für die Geistes- und Kulturwissenschaften ergibt sich durch diese Arbeitsweise ein großer Vorteil, nämlich „[...] *the analysis of cultural materials in their native format — as texts, images, and media. But much processing has to occur before text or image analysis can proceed.*“ (DRUCKER 2021: 110) Dies liegt daran, dass unstrukturierte Daten beziehungsweise Texte in unterschiedlichen Formaten vorliegen und der Inhalt historischer Texte besonders beachtet werden muss, da dieser im Vergleich zu heutigen Texten historische Veränderungen abbildet und diesen auch unterliegt. Rechtschreibung und Grammatik waren vor einigen Jahrhunderten noch ganz anders als heutzutage und auch Bezeichnungen und Bedeutungen von zum Beispiel Wörtern, Berufen oder Orten hat sich stark verändert. Natürlich können diese Veränderungen auch Ziel der Analyse sein, zumeist jedoch muss daher eine gewisse Standardisierung durchgeführt oder auf diese Änderungen Acht gegeben werden. (JOCKERS und UNDERWOOD 2016: 299)

Johanna DRUCKER (2012) sieht im Data Mining ein großes Potential für die Digital Humanities, auch wenn die Verbindung der Geistes- und Kulturwissenschaften dem Prinzip des Data Mining diametral gegenüberzustehen scheint. Vor allem der Bereich Text Mining (siehe Kapitel 6.1) kann dabei aber eine große Rolle spielen, da nun textliche Quellen informationstechnologisch bearbeitet werden können. Dies bringt Vorteile für Wissenschaftsbereiche wie zum Beispiel Literaturwissenschaft oder Geschichtswissenschaft, welche sich vor allem auf Texte als Quellen beziehen. Aber auch andere Bereiche wie zum Beispiel die Musikwissenschaft oder Kunstwissenschaft können vom Data Mining profitieren, indem Tonaufnahmen oder Bilder analysiert werden können. Dabei sollte meist eine Kooperation mit Data- oder Text-Mining-Expert*innen angestrebt werden, da diese die Expertise für diesen Bereich haben und mögliche Anwendungsbereiche aufzeigen und umsetzen können. (vgl. DRUCKER 2021: 110; JOCKERS und UNDERWOOD 2016: 292ff)

6.1.2. Location Extraction via Text Mining

Um räumliche Informationen aus den geschichtswissenschaftlichen Quellen zu bekommen, ist es nun besonders relevant, mittels Text Mining ebenjene zu extrahieren. Dabei gibt es unterschiedliche Vorgangsweisen, welche nun nachfolgend vorgestellt werden sollen. (Kapitel 6.1.3.1-6.1.3.5) Die extrahierten Daten sollen dann einerseits visualisiert werden, um eine

visuelle Betrachtung und Interpretation dieser zu ermöglichen. Dieser Prozess kann als „spatialization“ bezeichnet werden: „*The process of spatialization transforms non-geographic data to spatial forms for visual analysis. [...] Fundamentally, spatialization reduces data dimensionality by compressing multidimensional variables into two dimensional displays.*“ (YUAN 2010: 111) Andererseits können dadurch auch räumliche Analysen durchgeführt werden, welche Muster und Eigenheiten offenbaren können, welche auf den ersten Blick vielleicht nicht offensichtlich sichtbar sind.

Um nun Texte, welche bei flüchtiger Inaugenscheinnahme hin nicht unbedingt räumlich sind, in GIS zu verorten, gibt es mehrere Möglichkeiten: Einerseits ist es möglich, einen Text vollständig einem Ort zuzuweisen, oder man verknüpft nur einen bestimmten Teil des Textes mit einem Ort. Dies ist nur dann möglich, wenn die räumlichen Informationen aus den Texten bekannt sind. Hier kommt das Text Mining ins Spiel, da durch dieses eine schnellere und automatisierbare Möglichkeit besteht, Orte aus Texten zu extrahieren. (vgl. BITTNER und MICHEL 2018: 156; OWENS und WACHOWICZ 2013: 127f) Schon GREGORY und GEDDES (2014) stellten fest, dass in diesem Bereich großes Potential nicht nur für GIS, aber auch für die Geistes- und Kulturwissenschaften liegt: „*Combining these techniques to extract place-names with the abilities of gazetteers to georeference place-names quickly and easily has the potential to allow unstructured texts to be converted into GIS databases.*“ (GREGORY und GEDDES 2014:ebd.: 176) und auch BODENHAMER (2013) sieht hier großes Potential:

„[...] the ability to transform unstructured text into structured maps through computational text and data mining, semantic synthesis, geoparsing, place-name matching using natural language processing and digital gazetteers, and georeferencing methods promises to be a major contribution to the armory of tools available to the spatial humanities.“ (BODENHAMER 2013: 9)

Wobei es auch wichtig ist, dass neben der reinen Kartierung der extrahierten Informationen auch weitergedacht wird, indem die Informationen dazu genutzt werden, auch räumliche Analysen durchzuführen. (vgl. ebd.: 9; GREGORY 2010: 70)

6.1.3. Geoparsing

Eine Möglichkeit, räumliche Informationen aus unstrukturierten textlichen Quellen zu bekommen, ist „Geoparsing“:

„Geoparsing is the process of identifying geographic references in text and linking geospatial locations to these references so that the text can be accessed through spatial retrieval methods and suitable for spatial analysis. Geoparsing is used to add geospatial locations to written text, oral discourse, and legacy scientific data where referencing to location was done with placename references only.“ (HILL 2008b: 200)

Dabei werden Texte mittels unterschiedlicher Methoden (siehe unten) linguistisch analysiert, wodurch Teile dieser getaggt werden, welche räumliche Informationen darstellen. *„Text geoparsing allows us to link geographic properties of places to place names and enrich our understanding of discourses communicated in the text.“* (YUAN 2010: 115) Durch die Verlinkung von Ortsnamen in unstrukturierten Texten mit geographischen Koordinaten können diese dann räumlich verortet werden. (vgl. HILL 2008b: 200; HILL 2008a: 161)

Anschließend an Hu (2018) ist Geoparsing Teil eines Workflows (siehe Abbildung 2), in welchem mit „geo-text data“ gearbeitet wird. „Geo-text data“ bezeichnet dabei Texte, in welchen Verbindungen zwischen Orten und Texten bestehen. Dies kann entweder durch eine Angabe von Koordinaten, geotagged social media posts oder durch Erwähnungen von Ortsnamen oder anderer Toponyme geschehen.

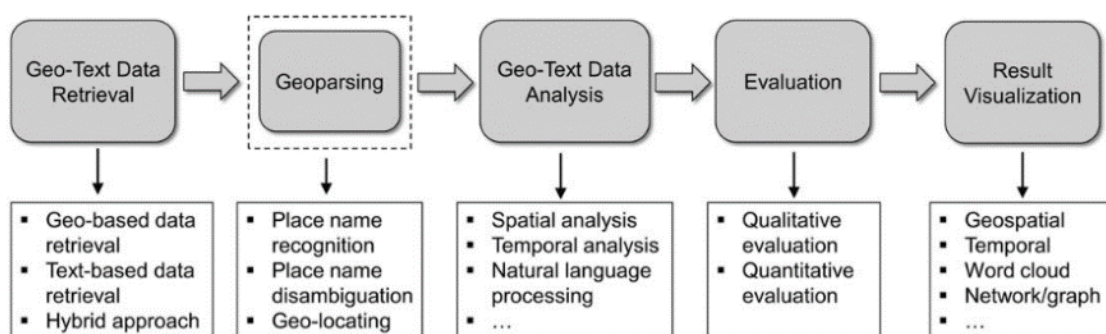


Abbildung 2: Workflow Arbeit mit „geo-text data“ (Quelle: Hu 2018: 11)

Der Workflow sieht dabei wie folgend aus:

1. Geo-Text Data Retrieval: In einem ersten Schritt werden relevante Texte (geo-texts) ausgewählt. Diese können entweder räumlich oder inhaltlich zusammenhängend sein und dadurch gewählt werden. Auch eine Kombination beider Ansätze ist möglich.
2. Geoparsing: Beim Geoparsing werden Wörter im Text, welche Ortsnamen repräsentieren, identifiziert und mit Koordinaten versehen. Dieser Schritt ist bei schon vorhandener Verortung von Texten (zum Beispiel bei geotagged social media posts) nicht notwendig. Da historische Quellen aber noch nicht verortet sind, ist dies im Zuge der vorliegenden Arbeit ein äußerst relevanter Schritt. Die genauere Ausgestaltung dieses Prozesses wird nachfolgend noch genauer beschrieben.
3. Geo-Text Data Analysis: Anschließend an die Verortung folgt nun die Analyse, eine der wichtigsten Funktionen von GIS. Diese kann entweder mit sprachwissenschaftlichen Methoden (z.B.: sentiment analysis) oder mit räumlichen Methoden (z.B.: clustering, kernel density estimation) durchgeführt werden. Wenn Zeitangaben vorhanden sind, kann auch eine zeitliche Analyse durchgeführt werden.
4. Evaluation: Die Evaluation dient der Sicherstellung eines validen und nützlichen Ergebnisses. Dabei können quantitative oder qualitative Ansätze gewählt werden. Eine Kombination beider ermöglicht es, eine robuste Belastbarkeit (quantitativ) und die Möglichkeit eines intuitiven Verständnisses (qualitativ) zu schaffen.
5. Result Visualization: Da es in diesem Workflow um die Extraktion räumlicher Informationen geht, sollten diese nach Abschluss dieser auch visualisiert werden. Dies ermöglicht eine visuelle Interpretation der zuvor nur schriftlich vorliegenden Informationen. Potentielle Visualisierungsmöglichkeiten sind unter anderem Karten, space-time-cubes, word clouds oder Netzwerke.

(vgl. Hu 2018: 11ff)

Der Bereich des Geoparsing besteht aus zwei Schritten. (siehe Abbildung 3) Einerseits müssen Ortsnamen aus dem Text extrahiert werden (Toponym Recognition) und anschließend werden diesen räumliche Koordinaten zugewiesen (Toponym Resolution). (siehe Kapitel 6.1.3.4 und 6.1.3.5)

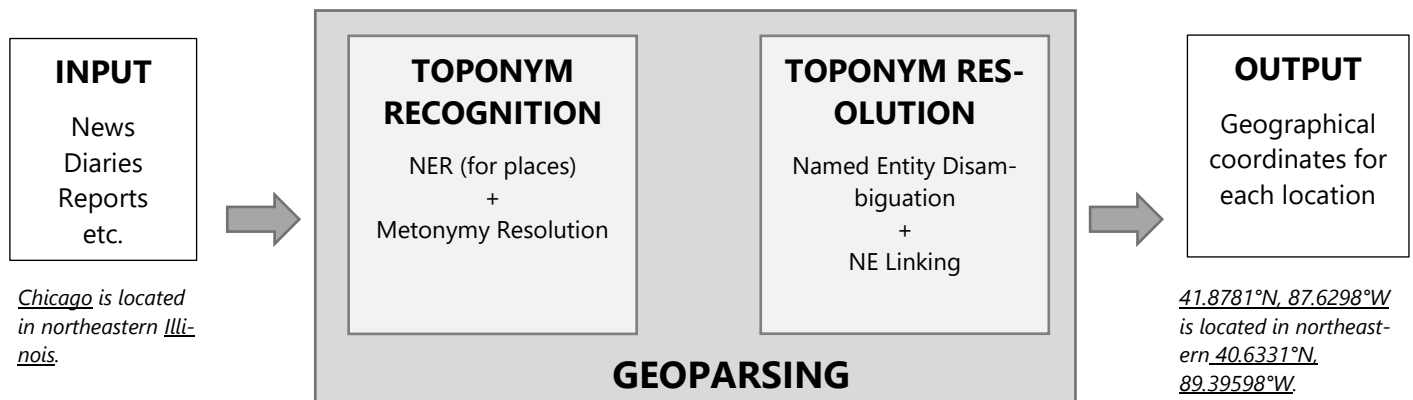


Abbildung 3: Geoparsing-Pipeline (verändert nach: GRITTA, et al. 2018: 604)

Im Bereich des Geoparsing gibt es drei verschiedene Methoden für die Identifizierung räumlicher Informationen in Texten:

6.1.3.1. Gazetteer-Lookup basiertes Geoparsing

Bei dieser Methode wird der Text Wort für Wort durchsucht und mit einem vordefinierten Set von Wörtern, welche im sogenannten Gazetteer aufgelistet sind, abgeglichen. Diese Methode funktioniert äußerst schnell und ist durch das Gegenchecken von Wörtern auch sprachunabhängig. Zusätzlich ist es mittels Vorauswahl möglicher relevanter Wörter (mittels Named Entity Recognition) möglich, nur einen gewissen Teil der vorhandenen Wörter zu verwenden. Solche Gazetteers sind eine Art Wörterbuch von Ortsnamen, welche heutzutage digital und in Form von Datenbanken vorliegen, in welchen Ortsnamen mit Koordinaten verbunden sind. „[...] Each entry represents a named geographic feature and contains at a minimum the essential elements of names, types, and locations.” (HILL 2008a: 159)

Natürlich beeinflusst der Umfang und die Qualität eines solchen Gazetteers auch das Ergebnis des Geoparsings, da nur Ortsnamen gefunden werden können, welche auch Teil des verwendeten Gazetteers sind. Dabei können Matches entweder nur dann gelten, wenn die Schreibweise genau die gleiche ist (full match) oder auch dann, wenn das Wort im Text und Ortsnamen nur teilweise übereinstimmen (partial match). Hier gilt es genau abzuwägen, welche Methode für das eigene Projekt die bessere ist. Die Beziehungen zwischen solchen Einträgen, also

zum Beispiel die Stadt Wien liegt im Land Österreich, können durch solche Gazetteers ebenso abgebildet werden. Außerdem ist es durch diese Datenbanken möglich, Ortsnamen, welche in Texten vorkommen, zu verorten. Dadurch kann die Brücke zwischen der räumlichen Verortung mittels Koordinaten und der Benennung von diesen Orten geschlagen werden. Außerdem ermöglichen Gazetteers durch die räumliche Verortung von Texten bzw. Textteilen auch eine räumliche Analyse dieser. (vgl. LEIDNER und LIEBERMAN 2011: 7; MONTEIRO, DAVIS und FONSECA 2016: 29f; PURVES, et al. 2018: 35; HILL 2008a: 159f; HU 2017: 5ff; vgl. MOSTERN und JOHNSON 2008: 1091) *„From a perspective of geospatial semantics, digital gazetteers help machines understand the geographic meaning [...] of a textual place name as well as the relations among places.“* (HU 2017: 5)

Problematisch dabei kann es sein, dass einige Orte durch unterschiedliche Namen gekennzeichnet sind, während manche Namen unterschiedliche Orte benennen. Ein Beispiel hierfür ist das Vorkommen des Städtenamens Springfield, welches mit 64 verschiedenen Nennungen eine der häufigsten in den USA und daher nur bei der Nennung des Städtenamens nicht zuzuordnen ist. (Wikipedia 2021) Andererseits wird zum Beispiel die Marktgemeinde Perchtoldsdorf südlich von Wien umgangssprachlich auch Petersdorf genannt. Zusätzlich gibt es noch die Probleme, dass unterschiedliche Bezeichnungen von Städten oder Orten in mehreren Sprachen bestehen (Bratislava – Pressburg als Endonym bzw. Exonym) oder sich Ortsnamen mit der Zeit verändert haben (Byzanz – Konstantinopel – Istanbul). Ein weiteres Problem stellt die Verwendung von Wörtern dar, welche zwar Ortsnamen sind, aber nicht in diesem Kontext verwendet werden. „Washington“ zum Beispiel bezeichnet einerseits eine Stadt, ist aber auch der Name des ersten US-Präsidenten. Außerdem gibt es keine vollständigen Gazetteers, vor allem nicht im Bereich von historischen Ortsnamen. Dies muss bei der Umsetzung des Geoparsings mittels Gazetteer immer mitbetrachtet werden. Zusätzlich sollten auch oftmals verwendete Abkürzungen (UK statt United Kingdom) in einem Gazetteer verzeichnet sein. Aus diesen Gründen ist das Arbeiten mit einem Gazetteer immer fehleranfällig, wobei es das Ziel sein sollte, diese Problemfälle so gut wie möglich zu minimieren. (vgl. GREGORY 2010: 70; MCHAFFIE, HWANG und FOLLETT 2018: 300f; PURVES, et al. 2018: 36)

Im Bereich des Historical GIS haben Gazetteers unterschiedliche Anforderungen und Merkmale. *„In recent years, scholars have called for gazetteers to record temporal change by date-stamping place names, spatial footprints, and feature types, and to record alternative place*

names and source citations that reflect both temporal change and multiple nomenclatures.” (MOSTERN und JOHNSON 2008: 1091) Einerseits gibt es historische Gazetteers, das heißt, sie bilden Ortsnamen ab, welche in einem gewissen Zeitraum in der Vergangenheit gültig waren. Daher ist immer fraglich, ob auch wirklich allen historischen Veränderungen im Gazetteer Rechnung getragen wird. Außerdem sollten immer die Quellen, aus welchen die Informationen zum jeweiligen Namen stammen, seien es historische Karten oder textliche Quellen, angegeben sein, damit eine Überprüfung der Informationen stattfinden kann. Außerdem wäre es optimal, wenn die Ortsnamen in mehreren Sprachen vorlägen, um eine Unterscheidung zwischen Endonymen und Exonymen möglich machen zu können. Dies ist vor allem bei historischen Ortsnamen relevant, da Verwaltungs- und Volkssprache in vielen Gebieten lange Zeit nicht die gleichen waren. Zusätzlich sollten unterschiedliche Schreibweisen von Ortsnamen auch Eingang in solche Gazetteers finden.(vgl. BOL 2011: 304f; GREGORY und GEDDES 2014a: 175f)

6.1.3.2. Regelbasiertes Geoparsing

Ein Regelsatz bezeichnet einen Entscheidungsprozess, welcher es ermöglicht, zwischen Ortsnamen und anderen Wörtern zu unterscheiden. Hierbei wird der Kontext der Wörter miteinbezogen, wodurch dieser Ansatz als sprachabhängig zu bezeichnen ist. Bei diesem Ansatz können neben internen auch externe Komponenten miteinbezogen werden. Erstere bezeichnet wortinterne Strukturen wie zum Beispiel die Großschreibung von Wörtern. In der englischen Sprache werden Ortsnamen im Gegensatz zu anderen Nomen immer großgeschrieben. Zweiteres bezeichnet die Beachtung des Kontextes im Text. Eine Angabe von Himmelsrichtungen (z.B.: nördlich von Washington) deutet auf einen räumlichen Kontext hin, während die Konstruktion Vorname + Ort (George Washington) eher auf die Verwendung eines Ortsnamens als Familienname hindeutet. (vgl. LEIDNER und LIEBERMAN 2011: 7f; vgl. MONTEIRO, DAVIS und FONSECA 2016: 29f; PURVES, et al. 2018: 35f)

6.1.3.3. Machinelearningbasiertes Geoparsing

Bei diesem Ansatz werden die Methoden des Machinelearnings angewandt. Das heißt, es wird mittels Trainingsdaten ein Algorithmus trainiert, welcher dann mittels Testdaten validiert wird, um seine Einsatzfähigkeit abzubilden. Diese Methode ist daher eher bei größeren Textmengen anzuwenden. Mittels dieser Methode werden die zugrundeliegenden Regeln, welche

auch schon beim regelbasierten Ansatz Teil der Analyse sind, automatisch aus dem Text heraus berechnet. Der Trainingsdatensatz ist so vorzubereiten, dass alle Wörter entsprechend ihrer Wortart getaggt sind und weitere Informationen über die einzelnen Wörter, wie zum Beispiel Wortlänge, Wortart, Groß- und Kleinschreibung und andere, welche kontextuelle Informationen wie voranstehende und nachfolgende Wörter abbilden, festgehalten sind. Die enthaltenen Ortsnamen sollten sich dabei nicht zu sehr von jenen unterscheiden, welche dann auch im Testdatensatz und bei weiterer Verwendung zum Einsatz kommen. (vgl. LEIDNER und LIEBERMAN 2011: 8; vgl. MONTEIRO, DAVIS und FONSECA 2016: 29f; PURVES, et al. 2018: 35ff)

In der nun folgenden Aufstellung der Vorgangsweise beim Geoparsing findet eine Kombination zweier Methoden statt (siehe auch Abbildung 3). Einerseits werden „Named Entity-Recognition“-Ansätze (NER) verwendet, um den Text auf mögliche relevante Wörter vorzufiltern und mittels Part-of-Speech-Methoden (POS) zu taggen. In einem zweiten Schritt werden diese dann mittels Gazetteer abgeprüft und georeferenziert. Diese Kombination kann auch bezeichnet werden als: „[...] *gazetteer-based entity matching used to geocode an entity recognition result.*“ (MIDDLETON, et al. 2018: 5)

6.1.3.4. Toponym Recognition

Toponym Recognition (Toponym=Ortsname) ist Teil der Named Entity Recognition (NER), welche dazu dient, Wörter in Texten definierten Klassen zuzuweisen. Während bei der NER alle Wörter unterschiedlichen Klassen zugewiesen werden (z.B.: Namen, Daten, Organisationen, Zahlen), wird bei der Toponym Recognition nur Wörter und Phrasen der Klasse Ortsname zugewiesen. Dies geschieht unter anderem mittels Part-of-Speech-Tagging, also der Zuweisung der Wörter zu Wortarten. Daher kann diese auch als „Extraction of Location Names“ bezeichnet werden. (vgl. GRITTA, et al. 2018: 604; HU 2018: 5; GREGORY und HARDIE 2011: 301f; GREGORY, et al. 2015: 3f; PURVES, et al. 2018: 34ff) Named Entity Recognition wurde bisher vor allem bei Social-Media-Nachrichten getestet, wobei diese einen besonderen Fall darstellen, da sie sehr kurze Texte beinhalten und zumeist auf Grammatik und Rechtschreibung nicht so viel Acht gegeben wird. (vgl. MIDDLETON, et al. 2018: 4ff) Natürlich wäre es auch möglich, alle Wörter im Text sofort mittels Gazetteer zu überprüfen, jedoch gibt es dabei einige Probleme, welche eine optimale Auswertung behindern würden. Dabei ist vor allem der Ansatz der Metonymie relevant, bei welchem auch Ortsnamen erkannt werden sollen, ohne schon zu wissen, dass es sich

bei dem Wort um einen solchen handelt. Außerdem gibt es auch die umgekehrte Variante, nämlich dass Ortsnamen zwar vorhanden, diese aber nicht als solche zu verstehen sind. (z.B.: Stockholm-Syndrom, Paris als Personennamen) (vgl. GRITTA, et al. 2018: 620)

Ein Problem dabei ist, dass zumeist nur der individuelle Ortsname in diesen Prozess mit einbezogen wird, während der ganze Terminus, welcher diesen enthält, nicht beachtet wird (zum Beispiel: nördlich von Wien, nahe Salzburgs, der Platz vor dem Wiener Rathaus). Dadurch können wichtige räumliche Informationen verloren gehen. Dieser Punkt sollte in zukünftigen Arbeiten vermehrt betrachtet werden. (PURVES, et al. 2018: 8)

6.1.3.5. Toponym Resolution

Toponym Resolution, auch Geocoding (bei strukturierten Texten) genannt, ist anschließend an die Toponym Recognition die Zuweisung der räumlichen Lage zu den getaggten Ortsnamen mittels Koordinaten. Dabei wird versucht, Probleme zu lösen, die auftreten können, wenn ein Ortsname mehrere reale Orte bezeichnen kann. (Location Disambiguation) Dies kann entweder mittels des vorhandenen Textes oder über eine räumliche Korrelation der anderen im Text vorhandenen Ortsnamen durchgeführt werden. Dabei wird also immer der Kontext miteinbezogen und das dann wahrscheinlichste Ergebnis ausgewählt. (vgl.; HU 2018: 5; PURVES, et al. 2018: 37ff; MIDDLETON, et al. 2018: 2ff) „*Given a list of candidate coordinates for each location and the surrounding context, the goal is to select the correct coordinate i.e. disambiguate*“ (GRITTA, et al. 2018: 604) Anschließend werden den nun eindeutigen im Text vorhandenen Ortsnamen Koordinaten zugewiesen. Dies geschieht mithilfe eines Gazetteers, in welchen sowohl Ortsnamen als auch Koordinaten verzeichnet sind. (vgl. GREGORY und HARDIE 2011: 301f; GREGORY, et al. 2015: 3f) Diese nun extrahierten räumlichen Informationen sind nun noch in einem GIS zu verorten, um diese anschließend auch visualisieren und analysieren zu können.

7. Verortung geschichtswissenschaftlicher Quellen mittels Geoparsing

In diesem Abschnitt geht es darum, die Erkenntnisse aus den vorangegangenen Kapiteln in einem praktischen Teil umzusetzen. Dabei ist es das Ziel, räumliche Informationen aus historischen Quellen zu extrahieren, was mit Hilfe des in Kapitel 6.1.3 vorgestellten Geoparsings geschehen soll.

GREGORY und COOPER (2009) kritisieren bei solchen Natural-Language-Processing-Arbeiten (Natural Language Processing ist eine Analyse natürlicher Sprache mittels Computeralgorithmen (vgl. DALE 2010: 4f)) im geschichtswissenschaftlichen Bereich, dass

„although most literary texts are set in particular geographical locations at specific historical moments, there has been a tendency for literary criticism to privilege the representation of temporality over issues of spatiality. In other words, geographical setting is seen to be of secondary importance to the ‘unfolding’ of individual characters over time.“ (GREGORY und COOPER 2009: 62)

Dies zeigt, dass bei der Analyse literarischer Texte die Zeitkomponente über die räumliche Komponente gestellt wird. Dies würde ich auch für geschichtswissenschaftliche Quellen so sehen, da diese meist zur Beobachtung zeitlicher Phänomene herangezogen werden, während eine räumliche Dimension oft keine Betrachtung findet.

Geoparsing soll nun dabei helfen, räumliche Informationen aus Texten zu extrahieren, um diese in GIS weiter nutzen zu können. Im Gegensatz zu oftmals genutzten Ansätzen, welche sich auf das manuelle Parsing von Texten konzentrieren (vgl. LAFRENIERE und GILLILAND 2015) soll in dieser Arbeit ein automatisches Parsing mittels Gazetteer eingesetzt werden. Dies bringt den Vorteil, dass die Texte schneller und arbeitsextensiver räumlich nutzbar gemacht werden können.

Wie in Abbildung 4 zu sehen ist, besteht ein solcher Arbeitsablauf nicht nur aus dem Geoparsing an sich, sondern aus weit mehr Arbeitsschritten. Daher soll nachfolgend auch kurz auf diese eingegangen werden. Untenstehende Abbildung zeichnet den Arbeitsablauf, welcher in dieser Arbeit angewandt wurde, nach. Im Vorlauf an das Geoparsing ist es wichtig, zu überlegen, mit welchen Quellen gearbeitet werden soll. Einerseits ist daher ein Thema zu wählen und passend dazu dann die Quellen auszuwählen. Diese sind dann entsprechend vorzubereiten, damit sie im nächsten Schritt weiterverwendet werden können.

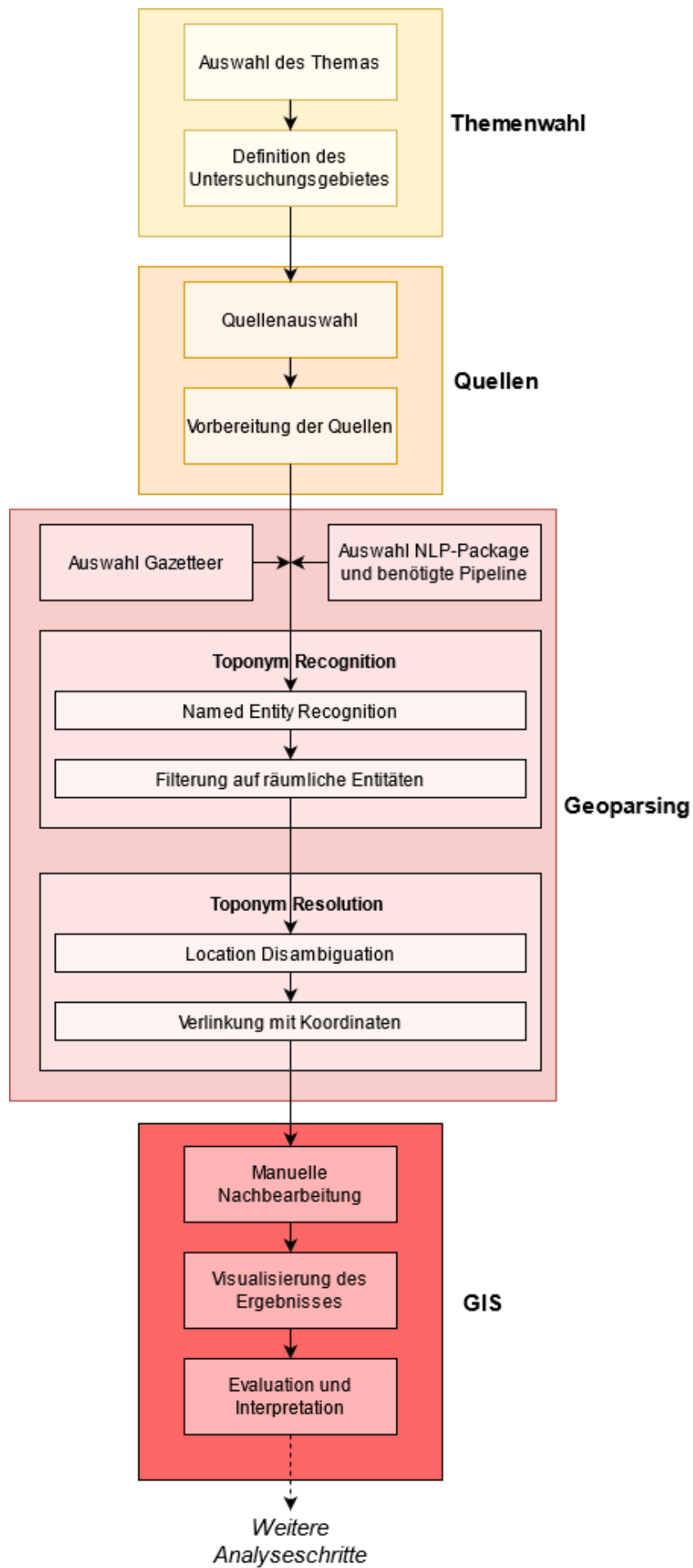


Abbildung 4: Workflow-Diagramm Verortung geschichtswissenschaftlicher Quellen via Geoparsing

Das Geoparsing an sich besteht einerseits in der Auswahl des gewünschten Gazetteers und des Natural-Language-Package, mit welchen gearbeitet werden soll. Anschließend findet das eigentliche Geoparsing, also Toponym Recognition und Toponym Resolution statt. Im GIS kann dann eine manuelle Nachbearbeitung stattfinden, um das Ergebnis noch zu verbessern. Dieses kann auch visualisiert werden, was dann bei der anschließenden Evaluation des Workflows und der Interpretation des Ergebnisses helfen kann. Nachfolgend wäre es noch möglich, die erstellten Daten weiteren Analyseschritten zu unterziehen. Dies fand im Zuge dieser Arbeit jedoch nicht statt.

Nachfolgend werden nun die einzelnen Schritte vorgestellt, welche für eine erfolgreiche Verortung geschichtswissenschaftlicher Quellen notwendig ist. Der Hauptfokus soll dabei jedoch auf der programmiertechnischen Umsetzung des Geoparsings liegen. Anschließend werden die Ergebnisse visualisiert, evaluiert und miteinander verglichen. Von diesen Erkenntnissen aus werden daraufhin Optimierungsmöglichkeiten für zukünftige Arbeiten in diesem Bereich geschlossen.

Eine Visualisierung der Ergebnisse ist wichtig, da durch diese ein schneller Überblick über das Ergebnis gewonnen werden kann und diese anschaulich kommuniziert werden. Solche können von Leser*innen einfacher erfasst und auch verstanden werden als eine Präsentation der Ergebnisse in Tabellenform. (vgl. DRUCKER 2021: 86ff) Vor allem bei räumlichen Daten, welche ja das Ergebnis einer Verortung mittels Geoparsing sind, lassen sich Zusammenhänge durch eine Visualisierung mittels Karten besser und schneller erkennen. Räumliche Cluster oder der Einfluss von anderen räumlichen Elementen wie Bergen oder Flüssen können zusätzliche Erkenntnisse über Texte liefern, welche bisher nicht erkennbar waren. (vgl. JANNIDIS, KOHLE und REHBEIN 2017: 331f) Außerdem kann durch eine Visualisierung der Handlungshorizont eines Textes skizziert werden, ohne dass der Text im Detail gelesen werden musste.

7.1. Wahl der Entwicklungsumgebung

Für die Umsetzung des Workflows in einen funktionierenden Code mussten folgende Bestandteile ausgewählt werden, wobei die nachfolgenden Kriterien bei der Auswahl dieser wichtig waren:

- Gazetteer: Das verwendete Gazetteer soll ohne Kosten frei verfügbar sein, damit eine einfache und unkomplizierte Nutzung möglich ist. Außerdem soll dieses möglichst viele Daten zu Ortsnamen enthalten. Wichtig dabei ist, dass die Daten so valide wie möglich sein sollen.
- Programmiersprache und Packages: Für mich bietet sich hier nur die Nutzung zweier Programmiersprachen an, welche ich schon in Grundzügen beherrschte: Python oder R. In der ausgewählten Programmiersprache sollten die wichtigsten Natural-Language-Processing-Packages vorhanden sein, um diese dann in der Programmierung auch nutzen zu können. Außerdem muss ein Package vorhanden sein, welches ermöglicht, das gewählte Gazetteer in der Programmiersprache verwenden zu können. Diese Packages sollen online gut dokumentiert sein, da es sich hierbei um meine erste Arbeit mit Natural Language Processing handelt und dadurch ein schnelles und einfaches Verständnis dieser von hoher Relevanz ist. Außerdem ist es wichtig, ein „Virtual Environment“ einrichten zu können, welches es ermöglicht, die Programmierumgebung auf einen einzelnen Workflow genau anzupassen. Dadurch ist es nämlich möglich, Packages in unterschiedlichen Versionen installiert zu haben. (vgl. Python Software Foundation 2021)
- Entwicklungsumgebung: Die Entwicklungsumgebung soll eine einfache Verwaltung der verwendeten Packages ermöglichen. Außerdem ist es wichtig, dass die Verbindung zum genutzten Geographischen Informationssystem besteht, damit Zwischenergebnisse schnell und einfach überprüft werden können.
- Geographisches Informationssystem: Dieses soll in erster Linie zur manuellen Nachbearbeitung und Visualisierung dienen. Außerdem ist es wichtig, dass es die oben genannten Bestandteile effizient miteinander verbindet, damit das Arbeiten so einfach wie möglich gemacht wird.

Als Gazetteer wurde „GeoNames“ ausgewählt, da dieses über elf Millionen Angaben zu Ortsnamen enthält. Diese Daten sind außerdem sehr vertrauenswürdig, da sie in hohem Maße von öffentlichen Stellen stammen. Ein Nachteil bei der Nutzung von GeoNames ist das Vorhandensein von in meinem Fall unnötigen Featureklassen wie „farm“ oder „undersea“. Diese können jedoch beim Geocoding-Prozess gefiltert werden. Ein weiterer nicht zu vernachlässigender

Nachteil ist eine limitierte Nutzung aufgrund einer notwendigen Anmeldung und eines stündlichen und täglichen Limits an freien Abfragen. Gerade bei längeren Texten könnte dies zu Problemen führen. (vgl. [geonames.org](https://www.geonames.org) o.J.)

Als Programmiersprache wurde „Python“ gewählt, da sie die wichtigsten NLP-Packages enthält. „SpaCy“, „cltk“ oder „StanfordNLP“ sind in Python enthalten. Dies ermöglichte es mir, verschiedene NLP-Packages auszuprobieren und das beste für meinen Anwendungsfall auszuwählen. Auch wenn „R“ ein eigenes Package namens „geonames“ enthält (vgl. ROWLINGSON 2019), wurde Python aufgrund der leichten Verfügbarkeit der wichtigsten NLP-Packages gewählt. In Python lässt sich GeoNames mittels „Geocoder“ (vgl. Denis Carriere 2013) einbinden. Dieses Package ermöglicht es auch, andere Geocoding-Services wie HERE, Bing oder Google zu verwenden.

Als Entwicklungsumgebung wurde ein „ArcGIS Notebook“ gewählt. Dies ist die Einbindung eines „jupyter notebooks“ in „ArcGIS Pro“. Dadurch ist es möglich, schon bei der Entwicklung des Codes in jenem Programm zu arbeiten, in welchem anschließend auch eine Überprüfung und eine Visualisierung der Ergebnisse stattfindet. Diese Verbindung ermöglicht es auch einfach und schnell, Zwischenergebnisse anzusehen und daraufhin Änderungen im Code vorzunehmen. Durch die Einrichtung eines „Virtual Environments“ bleibt die voreingestellte Entwicklungsumgebung unverändert. Dabei ist es jedoch wichtig zu beachten, kein Update von ArcGIS Pro vorzunehmen, da sonst die gesamte Entwicklungsumgebung nicht mehr nutzbar gemacht wird und dann neu aufgesetzt werden muss. Ansonsten bietet ArcGIS Pro eine einfache Verwaltung von Packages und eine gleichzeitige Nutzung von Programmierumgebung, Tabellenansicht zur Überprüfung und manuellen Nachbearbeitung der Ergebnisse und einer Karte zur Darstellung dieser. (vgl. GIMMLER und KALISKY 2020) Hierbei wäre es noch vorteilhaft, wenn sich auch das benutzte Gazetteer in ArcGIS Pro einbinden ließe, da so eine vollständige Integration aller Bestandteile eines Geoparsing-Workflows in einer Umgebung realisiert wäre. So ist es immer wieder notwendig, GeoNames via Browser aufzurufen.

7.2. Quellenauswahl

Anschließend an Kapitel 5.1 wurden verschiedene textliche Quellen ausgewählt, um verschiedene Sprachen, Entstehungszeiträume, Transkriptionsmethoden und Textarten auf ihre Bearbeitbarkeit mit dem oben angeführten Workflow (siehe Abbildung 4) zu überprüfen.

Quellen in der Geschichtswissenschaft werden heutzutage auf unterschiedlichem Wege nutzbar gemacht. Neben der klassischen analogen Nutzung in Archiven, Museen und Bibliotheken bietet die Nutzung digitaler Quellen ganz neue Möglichkeiten in der Arbeit mit Quellen. Viele von ihnen sind transkribiert vorhanden, das heißt, dass die Texte ohne großen Aufwand durchsucht und weiterverarbeitet werden können. Dadurch ist es möglich, eine bisher nicht bearbeitbare Menge an Quellen in die eigene Forschungsarbeit einzubeziehen. Dadurch ist es auch möglich, ganz neue Fragestellungen an die Texte zu richten. (vgl. TSCHIGGERL, WALACH und ZAHLMANN 2019: 90)

Eine Transkription von Texten aus Bilddateien, also zumeist von Fotos analoger Quellen wie Bücher, passiert zumeist automatisch mittels OCR-Methoden. „OCR“ steht für „optical character recognition“ und beschreibt das automatische Erkennen von Schrift (handschriftlich oder maschinelle), Buchstaben und Zeichen in Bildern und eine Transkription dieser. (vgl. CHERIET, et al. 2007: 1ff) Dadurch können aus Fotos Texte automatisch extrahiert werden.

Auch wenn eine Nutzung digitaler Quellen nach einer Vereinfachung in der geschichtswissenschaftlichen Arbeit klingt, sind zwei wichtige Punkte dabei zu beachten: Standardisierung und Obsoleszenz. Wenn Quellen nämlich nicht in standardisierten Formaten vorhanden sind, ist eine einfache und universelle Arbeit mit diesen nicht so leicht möglich. Digitale Speichermedien haben auch nur eine gewisse Lebensdauer und auch Cloud-Speicher versprechen keine andauernde ungehinderte Nutzbarkeit, da Quellen dadurch auch immer leicht vom Netz genommen werden können und dadurch keine Zugänglichkeit mehr besteht. (vgl. TSCHIGGERL, WALACH und ZAHLMANN 2019: 91)

Trotzdem habe ich mich dafür entschieden, online verfügbare Quellen zu nutzen, da diese schon in transkribierter Form vorliegen und dadurch einfacher und schneller nutzbar sind. Eine eigenständige Digitalisierung von verschiedenen Texten nähme zu viel Zeit in Anspruch und außerdem sind online verschiedene Arten an Quellen in transkribierter Form vorhanden.

Aus diesem Grund wählte ich vier verschiedene Quellen aus, wobei drei von diesen (1-3) anschließend vollständig weiterverarbeitet wurden:

- Quelle 1: handschriftlicher deutschsprachiger Brief, manuell transkribiert, Zweiter Weltkrieg
- Quelle 2: gedruckte deutschsprachige Zeitung, automatisch mittels OCR-Methoden transkribiert, 19. Jahrhundert
- Quelle 3: gedrucktes deutschsprachiges Buch, automatisch mittels OCR-Methoden transkribiert und manuell nachbearbeitet, 19. Jahrhundert
- Quelle 4: handschriftlicher englischsprachiger Brief, manuell transkribiert, Erster Weltkrieg

Während also die Quellen 1, 3 und 4 folglich keine Fehler enthalten sollten, enthält die Quelle 2 einige Fehler, welche bei der automatischen Texterkennung fälschlicherweise eingebaut wurden. ALEX, et al. (2019) bestätigen die Nutzung unterschiedlich transkribierter Texte als besonders relevant, da dadurch eine Vergleichbarkeit zwischen den einzelnen Methoden im Hinblick auf den Workflow ermöglicht wird.

Die oben genannten Quellen sind nur eine kleine Auswahl der schon in Kapitel 5.1 beschriebenen Typen. Hierbei musste ich eine Auswahl treffen, da eine umfassende Bearbeitung aller Quellentypen im Rahmen dieser Arbeit nicht möglich ist. Bei der Auswahl der Quellen wurde ein Augenmerk darauf gelegt, dass diese räumliche Informationen enthalten. Außerdem sollten diese aus einem ähnlichen Zeitraum kommen und dadurch eine sprachliche Vergleichbarkeit möglich ist, damit die Ergebnisse nicht durch diese Faktoren beeinflusst werden.

7.2.1. Quelle 1: Brief Weiß 1940

Für eine erste Entwicklung des Python-Codes wurde ein relativ kurzer Text (550 Wörter) gewählt, um diesen dann schnell und einfach umsetzen zu können. Bei diesem Text (nachfolgend bezeichnet als „Q1_Brief_Weiss_1940“, Transkript siehe Anhang 11.3.1) handelt es sich um einen Brief, welchen Otto Weiß 1940 im Zweiten Weltkrieg aus Frankreich an seine Ehefrau schrieb. Der Brief ist Teil eines Konvolutes, welches von Otto Weiß zwischen 1940 und 1943 verfasst wurde. Es wird vermutet, dass der Verfasser im Krieg gefallen ist. Der betreffende Brief wurde im Zuge des Frankreich-

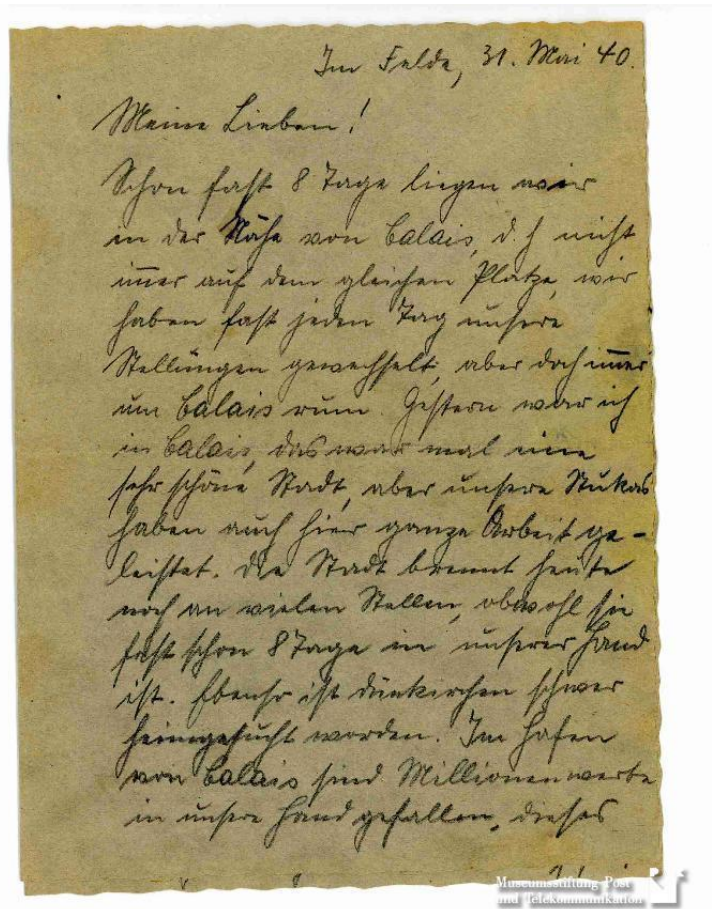


Abbildung 5: Q1_Brief_Weiss_1940, erste Seite (Bildquelle: Museumsstiftung Post und Telekommunikation o.J.)

Feldzuges in der Nähe von Calais verfasst und berichtet von der Zerstörung der Stadt und der näheren Umgebung. Des Weiteren berichtet der Verfasser noch über die Situation von Flüchtlingen vor Ort und das Leben der Soldaten im besetzten Frankreich. Der Brief liegt in der Briefsammlung der „Museumsstiftung Post und Telekommunikation“ in Frankfurt am Main. (vgl. Museumsstiftung Post und Telekommunikation o.J.)

Tabelle 4: Quelleneigenschaften Q1_Brief_Weiss_1940

handschriftlich	gedruckt
analog vorliegend	digital vorliegend
manuell transkribiert	automatisch transkribiert (OCR)
standardisiert	nicht standardisiert/erzählend
deutschsprachig	nicht deutschsprachig

Dieser Brief wurde als erste Quelle gewählt, da er online transkribiert vorhanden und dadurch keine aufwändige Quellenedition notwendig war (siehe Tabelle 4). Außerdem ist er in deutscher Sprache verfasst, wodurch ein einfaches und schnelles Verständnis möglich ist. Dadurch konnte ich mich bei der Entwicklung des Python-Codes komplett auf ebenjene Tätigkeit konzentrieren. Die Kürze des Textes ermöglichte dann auch eine schnelle Überprüfung des Ergebnisses, ohne dabei in vielen Absätzen nach Ortsnamen zu suchen.

Briefe sind dabei als Ego-Dokumente im Sinne SCHULZES (1996) zu verstehen. Als geschichtswissenschaftliche Quelle bieten sie insofern einen Vorteil, da sie einerseits das individuelle menschliche Handeln und die Ängste und Gefühle des*der Verfasser*in darstellen und dies so auch reflektiert betrachtet werden kann. Außerdem bieten sie so eine Sicht der*des Einzelnen auf größere Zusammenhänge wie zum Beispiel den Zweiten Weltkrieg. Dabei muss dem*der Leser*in immer bewusst sein, dass es sich hierbei eben um eine singuläre Sicht auf größere Zusammenhänge handelt und dies auch in möglichen Interpretationen beachtet werden muss.

Da ich mich in der vorliegenden Arbeit aber nicht mit der Interpretation und Einbindung solcher Quellen in größere Zusammenhänge beschäftige, sondern mit der Verortung dieser Quelle, behalte ich diese Einschränkungen zwar im Hinterkopf, diese behinderten mich aber nicht bei der weiteren Arbeit.

7.2.2. Quelle 2: Neue Freie Presse, 24.09.1886

Die „Neue Freie Presse“ war eine österreichische Zeitung (1864-1939) und eine Vorgängerin der aktuellen „Die Presse“. Sie erschien mehrmals täglich. (vgl. Österreichische Nationalbibliothek 2011) Die ersten zwei Seiten der ausgewählte Ausgaben vom 24.09.1886 (Nr. 7931, Morgenblatt) enthalten neben den üblichen Informationen zum Abonnement einen Artikel, welcher die Enthüllung eines Denkmals für Admiral Wilhelm von Tegetthoff am nächsten Tag beschreibt. Außerdem gibt es einen Artikel über Tegetthoff selbst, welcher sich über die ersten zwei Seiten zieht. Seite zwei enthält außerdem noch Artikel über Diskussionen um das Schwurgericht und weitere innenpolitische Debatten. Diese Seiten wurden stellvertretend für eine andere Art an wichtigen Quellen in der Geschichtswissenschaft gewählt: Zeitungen. (siehe Tabelle 5) Außerdem ist die vorliegende Ausgabe für weitere Schritte besonders

relevant, da sie als Zeitung einerseits einen andern Schreibstil als die anderen verwendeten Texte aufweist und andererseits mittels OCR-Methoden automatisch transkribiert wurde. Ein automatischer Übertrag aus der damals in Druckwerken üblichen Frakturschrift (siehe Abbildung 6) kann dabei zu einigen Textfehlern führen. Daher ist es interessant zu prüfen, inwieweit sich solche Quellen für einen Geoparsing-Workflow eignen. (vgl. o.V. 24.09.1886) Diese Quelle wird nachfolgend als „Q2_NFP_24091886“ bezeichnet. Die Transkription der ersten zwei Seiten ist im Anhang 11.1.2 zu sehen.

Tabelle 5: Quelleneigenschaften Q2_NFP_24091886

handschriftlich	gedruckt
analog vorliegend	digital vorliegend
manuell transkribiert	automatisch transkribiert (OCR)
standardisiert	nicht standardisiert/erzählend
deutschsprachig	nicht deutschsprachig

Im Gegensatz zu den anderen drei Quellen, welche den Ego-Dokumenten zuzuordnen sind, sind Zeitungen Medien, welche neben der einfachen Übertragung von Informationen auch Diskurse der damaligen Zeit abbilden und diese auch selbst generieren. Dadurch kann ein anderer Blick auf vergangene Ereignisse gewonnen werden. Des Weiteren sollte auch die Zeitung selbst, also die Blattlinie, die politische Ausrichtung oder das Nahestehen zu Parteien, Eingang in eine Analyse finden. (vgl. TSCHIGGERL, WALACH und ZAHLMANN 2019: 129)

Vor allem die Digitalisierung von Zeitungen wie im Zuge des „ANNO“-Projektes der Österreichischen Nationalbibliothek (vgl. Österreichische Nationalbibliothek 2021) führt zu einem einfacheren Zugang zu geschichtswissenschaftlich relevanten Quellen. Dadurch ist es möglich, die Zeitungen auch schnell und einfach in eigenen Analysen zu verwenden.

7.2.3. Quelle 3: Ida Pfeiffer, Reise einer Wienerin in das Heilige Land

Ida Pfeiffers „Reise einer Wienerin in das Heilige Land“ (Q3_Pfeiffer) wurde gewählt, da es sich bei diesem Werk um einen Reisebericht handelt. Dieser ist für eine Verortung mittels Geoparsing besonders geeignet, da in solchen viele verschiedene Ortsnamen vorkommen und so ein

gutes Bild der Reiseroute gezeichnet werden kann. Eine Transkription des ersten Kapitels ist im Anhang 11.1.3 nachzulesen.

Ida Pfeiffer (1797-1858) gilt als eine der bedeutendsten Forschungsreisenden und Reise-schriftstellerin des 19. Jahrhunderts. Sie war die erste Frau, welche die Welt umrundete. Ab dem Jahr 1842 verbrachte sie ihr Leben fast ausschließlich mit Reisen und sammelte dabei auch Objekte, welche noch heute im Naturhistorischen Museum Wien zu sehen sind. Insgesamt unternahm sie fünf Fernreisen, welche sie in Reiseberichten detailliert beschrieb. Diese wurden mehrfach aufgelegt und in mehrere Sprachen übersetzt. Die erste ihrer Reisen, welche sie im Alleingang durchführte, brachte sie dabei ins Heilige Land. Dabei passierte sie Konstantinopel, Beirut, Libanon, Syrien und Ägypten. Der Bericht dieser Reise (vollständiger Titel: *„Reise einer Wienerin in das Heilige Land – Konstantinopel, Palästina, Ägypten nämlich: von Wien nach Konstantinopel, Brussa, Beirut, Jaffa, Jerusalem, dem Jordan und todten Meere, nach Nazareth, Damaskus, Libanon, Alexandrien, Kairo, durch die Wüste an das rothe Meer u.s.w.“*) soll nachfolgend als Quelle weiterverwendet werden. In diesem beschreibt Pfeiffer den Verlauf ihrer Reise sowie auf dem Weg erfahrenen Erlebnisse. Diese Quelle wurde mittels OCR-Methoden transkribiert und anschließend wurde diese Transkription durch Freiwillige noch manuell verbessert. (vgl. Projekt Gutenberg-DE o.J.) Tabelle 6 zeigt die weiteren Eigenschaften der Quelle. Auch diese ist eine gedruckte und digital vorliegende erzählende Quelle.

Tabelle 6: Quelleneigenschaften Q3_Pfeiffer

handschriftlich	gedruckt
analog vorliegend	digital vorliegend
manuell transkribiert	automatisch transkribiert (OCR)
standardisiert	nicht standardisiert/erzählend
deutschsprachig	nicht deutschsprachig

Ihre nächste Reise führten sie nach Skandinavien und Island (1845), außerdem machte sie zwischen 1846 und 1848 eine Weltreise, bei welcher sie zum Beispiel Brasilien, China, Indien, Persien und Russland besuchte. Eine zweite Weltreise (1851-1855) führte sie unter anderem nach Singapur und Indonesien sowie Nord- und Südamerika. Ihre letzte Reise führte sie dann

nach Madagaskar und Indonesien (1856-1858). Die dabei gesammelten Informationen wurden in der historischen und ethnographischen Forschung verwendet. (vgl. o.V. 2016: 2522 ff)

An der Universität Wien gibt es ihr zu Ehren die „Ida-Pfeiffer-Professur“. Diese ermöglicht es semesterweise Forschungs- und Lehrprojekte an der Fakultät für Geowissenschaften, Geographie und Astronomie durchzuführen. (vgl. Fakultät für Geowissenschaften, Geographie und Astronomie 2021)

Reiseberichte eignen sich dabei besonders für das Geoparsing, weil sie, durch die Beschaffenheit solcher Berichte, meist eine Menge räumlicher Information enthalten.

Zusätzlich enthalten sie auch Informationen für verschiedene Fachdisziplinen (u.a. Geschichtswissenschaft, Geographie, Ethnologie, Anthropologie oder Politikwissenschaft), wobei der Fokus dabei natürlich immer auch auf räumlichen Gegebenheiten und Zusammenhängen liegt. Daher sollte, im Sinne des *spatial turns*, auf solche räumlich relevanten Quellen vermehrt eingegangen werden. Denn diese bieten Informationen, welche vor Ort gesammelt worden sind. Auch wenn dabei immer der Blick des*der Autor*in hinterfragt werden muss, sind diese hilfreiche Quellen, um Informationen über gewisse Gebiete vor dem 20. Jahrhundert zu erhalten. (vgl. CAMPBELL 2002: 261ff)

7.2.4. Quelle 4: Brief WW1_Training

Als vierte Quelle, nachzulesen im Anhang 11.1.4, wurde der Brief „Training: ‘drill before breakfast’“ (SMITH 1915) gewählt. Dieser ist über die Website des britischen Nationalarchivs abrufbar. Er wurde gewählt, da er in englischer Sprache verfasst wurde und so einen Vergleich zur Methode des Geoparsings deutschsprachiger Quellen ermöglicht. Der Brief wurde am 04. April 1915 von Arthur Smith verfasst und stammt aus Frankreich im ersten Weltkrieg. Dies ermöglicht einen Vergleich zu Quelle 1, welche ebenso ein Brief von der Front aus dem zweiten Weltkrieg ist. Der Verfasser berichtet im Brief von seinem militärischen Training im Longmoor Camp und der Überfahrt nach und der Landung in Frankreich. (vgl. ebd.)

Bei der Auswahl dieser Quelle war es von Relevanz eine Vergleichbarkeit mit der ersten Quelle zu schaffen, daher wurde eine Quelle ausgewählt, welche bis auf die Sprache identische Merkmale aufweist. (siehe Tabelle 7) Die Quelle wurde als Vergleich zwischen der deutschen und

englischen Named Entity Recognitionin mittels spaCy verwendet (siehe Kapitel 7.3.1) und wurde daher auch nicht nachfolgend in weitere Vergleiche einbezogen oder visualisiert.

Tabelle 7: Quelleneigenschaften Q4

handschriftlich	gedruckt
analog vorliegend	digital vorliegend
manuell transkribiert	automatisch transkribiert (OCR)
standardisiert	nicht standardisiert/erzählend
deutschsprachig	nicht deutschsprachig

7.3. Geoparsing im engeren Sinne

7.3.1. Toponym Recognition

Um nun die vorliegenden Quellen mittels Geoparsing zu verorten, wurden einige zusätzliche Python Packages benötigt:

- pprint
- pandas
- spacy
- geocoder

Die Packages „pprint“ und „pandas“ wurden zur Abspeicherung der erfassten Entitäten als Liste beziehungsweise Dataframe genutzt, damit diese anschaulich und effizient weiterverarbeitet werden konnten. Die anderen beiden Packages, „spacy“ und „geocoder“, waren für die Umsetzung des Geoparsings essenziell, daher sollen diese nun am Beginn des jeweiligen Unterkapitels genauer vorgestellt werden.

„spaCy“ ist ein Open-Source-Natural-Language-Processing-Package für Python, entwickelt von der Firma „Explosion“. Dieses ermöglicht es, in Python Texte zu analysieren. Dabei können unterschiedliche Funktionen wie Tokenization, Part-of-speech-Tagging, Lemmatization, Named Entity Recognition, Text Classification oder Rule-based Matching durchgeführt werden. Diese Funktionen sind für über 60 verschiedene Sprachen (wie Englisch, Französisch oder

Deutsch, aber auch zum Beispiel Vietnamesisch, Slowenisch oder Persisch) verfügbar. Dies ist ungewöhnlich, da jede Sprache ihre grammatikalischen und lexikalischen Eigenheiten besitzt und die Funktionen dabei immer auf diese angepasst werden müssen. Für einige dieser Sprachen sind auch vortrainierte Pipelines vorhanden, sodass zur Durchführung eines Natural-Language-Processing-Projektes das Trainieren eigener Daten nicht unbedingt notwendig ist. Dies wäre mit einem erheblichen Aufwand verbunden, da bei diesem eine ausreichende Menge an Entitäten zuerst selbst getaggt werden müsste, um diese dann in ein Modell einfließen zu lassen. Auch das Modifizieren bereits vorhandener Pipelines ist möglich. Die vortrainierten Modelle, welche in unterschiedlichem Umfang vorhanden sind, haben jedoch auch Nachteile, da sie mit bestimmten Textsorten (Englisch: Blogs, Nachrichten, Kommentare aus dem Web; Deutsch: Nachrichten und Medien) trainiert wurden. (vgl. Explosion 2021) Wenn nun mit geschichtswissenschaftlichen Quellen Named-Entity-Recognition-Aufgaben durchgeführt werden, können manche vortrainierte Modelle gut für die jeweilige Quelle funktionieren, während bei anderen dies nicht der Fall ist. Dies kann von unterschiedlichen Faktoren abhängen: Einerseits hat die Textsorte einen gewissen Einfluss, da bei manchen eine formellere Sprache genutzt wird als bei anderen. Außerdem hat auch der Herkunftszeitraum der Quelle einen Einfluss, da Sprachen sich im Laufe der Zeit verändern und dadurch unterschiedliche Ausdrucksweisen, Grammatik und Rechtschreibung nicht unveränderlich sind. Außerdem ist die Genauigkeit der Transkription von vor allem älteren Texten besonders relevant. Zwar liefern automatische Texterkennungsmethoden wie OCR (Optical Character Recognition) schnell und umfangreich Ergebnisse, die Richtigkeit dieser ist aber ziemlich unterschiedlich. Dies hängt zumeist von der Qualität des gescannten Dokuments und der verwendeten Schriftart ab. Vor allem die Frakturschrift, welche im Deutschen lange Zeit benutzt wurde, sorgt hier oft für Probleme, wie im Beispiel unten aus der „Neuen Freien Presse vom 24.09.1886 zu sehen ist:

Feuilleton.

Aus Tegetthoff's Mappe.

Von den Männern, die im letzten Vierteljahrhundert sich einen hervorragenden Namen gemacht haben, wird Wilhelm v. Tegetthoff wol am längsten in dem Andenken der Mit- und Nachwelt fortleben. Und gewiß würde sein Bild an Lebendigkeit gewinnen, wenn über sein Privatleben und über seinen Charakter mehr bekannt würde, als dies thatsächlich der Fall ist. Das knappe Lebensbild, welches ich vor einigen Jahren entwarf, beruhte zum überwiegenden Theile auf Familienpapieren. Mit dem Tode des Vaters brechen diese vertraulichen Mittheilungen leider ab. Wenn auch Tegetthoff mit inniger Liebe an seiner Mutter hing, die ihn bekanntlich überlebte, so hat er sich ihr gegenüber doch nicht über seine amtliche Stellung und seine Thätigkeit ausgesprochen, während er seinem Vater, der an allen öffentlichen Angelegenheiten den regsten Antheil nahm und mit seltener Klarheit Menschen und Verhältnisse beurtheilte, die kleinsten Vorgänge mittheilte und sein volles Herz ausschüttete. Von den vielen Privatbriefen ist bisher wenigstens nichts in die Oeffentlichkeit gedrungen, und doch wären wir an der Hand derselben allein in der Lage, mancherlei Einzelheiten aufzuhellen, um Tegetthoff und die ganze Lebenswürdigkeit seines Charakters voll kennen zu lernen. Die Briefe an seinen Vorgesetzten Fautz enthielten über die Vorgänge bei Helgoland interessante Angaben. Wohin sind diese gerathen? Auch der jetzige Leiter des Marine-Ministeriums würde sich ein Verdienst erwerben, wenn er bei seinen Kameraden Nachfrage hielte und vielleicht aus seinem Schatzkästlein Einiges mittheilen würde.

Abbildung 6: Beispiel OCR (Quelle: Neue Freie Presse 24.09.1886: 1)

Feuilleton.

J Aus Tegetthoff's Mappe.

Von den Männern, die im letzten Vierteljahrhundert

sich einen hervorragenden Namen gemacht haben, wird Wilhelm v. Tegetthoff wol am längsten in dem Andenken der

Mit- und Nachwelt fortleben. Und gewiß würde sein Bild an Lebendigkeit gewinnen, wenn über sein Privatleben und

über seinen Charakter mehr bekannt würde, als dies thatsächlich der Fall ist. Das knappe Lebensbild, welches ich vor

einigen Jahren entwarf, beruhte zum überwiegenden Theile auf Familienpapieren. Mit dem Tode des Vaters brechen

diese vertraulichen Mittheilungen leider ab. Wenn auch Tegetthoff mit inniger Liebe an seiner Mutter hing, die ihn

bekanntlich überlebte, so hat er sich ihr gegenüber doch nicht über seine amtliche Stellung und seine Thätigkeit ausgespro-

chen, während er seinem Vater, der an allen öffentlichen Angelegenheiten den regsten Antheil nahm und mit seltener

Klarheit Menschen und Verhältnisse beurtheilte, die kleinsten Vorgänge mittheilte und sein volles Herz ausschüttete. Von

den vielen Privatbriefen ist bisher wenigstens nichts in die Oeffentlichkeit gedrungen, und doch wären wir an der Hand

derselben allein in der Lage, mancherlei Einzelheiten aufzuhellen, um Tegetthoff und die ganze Lebenswürdigkeit

seines Charakters voll kennen zu lernen. Die Briefe an seinen Vorgesetzten Fautz enthielten über die Vorgänge bei

Helgoland interessante Angaben. Wohin sind diese gerathen? Auch der jetzige Leiter des Marine-Ministeriums würde sich

ein Verdienst erwerben, wenn er bei seinen Kameraden Nachfrage hielte und vielleicht aus seinem Schatzkästlein Einiges

mittheilen würde.

(Neue Freie Presse 24.09.1886: 1)

Das richtige Erkennen einzelner Buchstaben wird durch die Frakturschrift eingeschränkt. Dies liegt daran, dass diese vor 1900 nicht normiert war und dadurch verschiedene sich stark ähnelnde Schreibweisen üblich waren. Vor allem das lange s („ſ“) führt bei OCR-transkribierten Frakturschrift-Texten zu Problemen, da dieses entweder nicht als solches erkannt wird oder dieses dann nicht weiterverarbeitet werden kann, da viele Natural-Language-Processing-Pipelines dieses nicht kennen. (vgl. KLEIN, DÄUMER und HÜTIG 2016: 334) Daher muss in der Arbeit mit solchen Texten immer auch dieses Problem mitgedacht werden. Im oben angegebenen Text wurde das lange s zu einem „normalen“ s umgewandelt, um etwaige Verständnisprobleme zu vermeiden.

Trotz all dieser Einschränkungen entschied ich mich, mit vortrainierten Pipelines zu arbeiten, da das Trainieren eigener ein zu großer Aufwand gewesen wäre, vor allem deswegen, da ich

mit verschiedenen Texten unterschiedlicher Sprache, verschiedener Textsorten und Herkunftszeiträume arbeitete.

Explosion gibt auf der Website eine Genauigkeit bei der Named Entity Recognition von ca. 85% an (geprüft wurde mit dem OntoNotes 5.0 Corpus). Die vortrainierten Pipelines sind damit etwa ebenso genau wie andere vergleichbare NLP-Tools, aber dafür besonders schnell. Die Gründe, warum ich dieses Package für die Durchführung des NER auswählte, waren unter anderem der leicht gemachte Einstieg durch eine ausführliche Dokumentation auf der Website sowie die Verfügbarkeit von Videotutorials, welche das Kennenlernen von spaCy erleichterten.

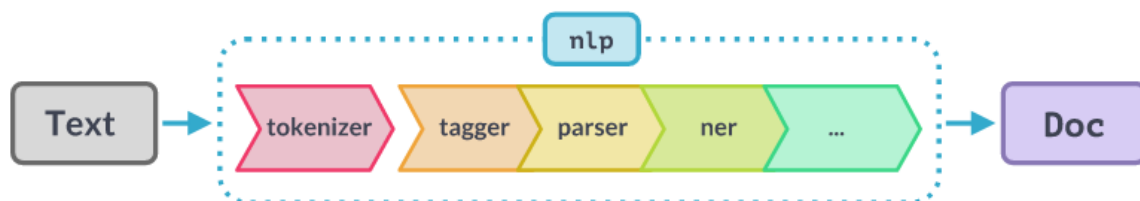


Abbildung 7: spaCy-Workflow (Bildquelle: Explosion 2021)

SpaCy funktioniert so, dass der Funktion ein Text übergeben wird (siehe Abbildung 7) und dieser dann mittels Natural Language Processing weiter analysiert wird. Dabei werden Funktionen wie Tokenization, Part-of-speech-Tagging oder Named Entity Recognition durchgeführt. Wenn dies abgeschlossen ist, wird der Text mit allen erfassten Informationen als Doc, also einer Sequenz von Token-Objekten, ausgegeben. Diese Tokens sind dabei zum Beispiel einzelne Wörter, Leer- oder Satzzeichen. Die erfassten Attribute sind in der Doc-Datei den einzelnen Tokens beigelegt. (vgl. Explosion 2021)

Abbildung 8 zeigt eine Visualisierung des Part-of-Speech-Taggings mittels spaCy. Dabei sind unter dem Satz „Schon fast 8 Tage liegen wir in der Nähe von Calais.“ die jeweils zugeordneten Wortarten (z.B.: Verb, Adverb, Nomen, Präpositionen, Artikel) abgebildet. Die Pfeile oberhalb der Wörter zeigen dann die Abhängigkeit der einzelnen Wörter von anderen an. Zum Beispiel gehören die Wörter „in“ und „der“ zum Wort „Nähe“. Die Wörter „liegen“ und „wir“ sind über eine Verb-Subjekt-Beziehung miteinander verbunden. Dies zeigt, dass spaCy viele weitere Funktionen bietet, um Texte sprachlich zu analysieren.

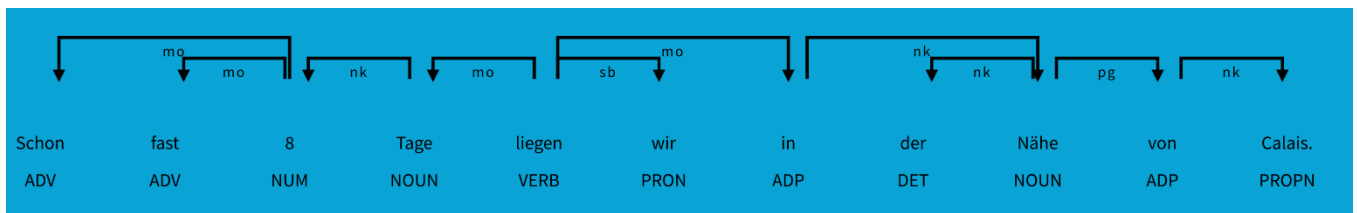


Abbildung 8: spaCy Part-of-Speech-Tagging

Die Programmierung wurde in „ArcGIS Pro“ in einem jupyter-Notebook vorgenommen, damit die Ergebnisse sogleich auch via Kartenansicht überprüft werden konnten. Der zu bearbeitende Text wurde mittels folgender Funktionen eingelesen (gesamter Python-Code siehe Anhang 11.2):

```
## Einstellung der Workspace-Direction und des Filenamen
dir_q = "C:/Users/andre/OneDrive/Dokumente/UNI/Master/S4_SS2021/SE
Konversatorium zur Kartographie und Geoinformation (gilt als Priva-
tissimum zur Masterarbeit aus KGI)/Masterarbeit/praxis/texts/Q1_Brief_Weiss_1940/"
filename = "Q1_Brief_Weiss_1940"

## Laden der Datei, welche geoparsed werden soll
file = open(dir_q+filename+".txt", "r", encoding='utf8')

## Speichern unter dem Namen q1
q1 = file.read()
file.close()
print(q1)
```

Somit ist der Text „Q1_Brief_Weiss_1940“ (siehe Anhang 11.1) nun in der Variable „q1“ gespeichert und kann weiter verarbeitet werden. Anschließend wurde das Natural Language Processing angewandt und dadurch eine Doc-Datei erzeugt:

```
## Anwendung der Named Entity Recognition (mittels spacy) auf den
Text

#Laden des trainierten deutschen Models
#nlp = de_core_news_sm.load()
nlp = de_core_news_lg.load()

##Anwendung der spacy-Funktion nlp zur Extraktion von Entitäten auf
die Quelle
doc = nlp(q1)
```

Hierfür wurden unterschiedliche vortrainierte Pipelines ausprobiert. Da die Pipeline „de_core_news_lg“ (lg steht für Large) die besten Werte in Bezug auf die Named Entity Recognition lieferte (beste Precision- und Recall-Werte, vom Entwickler angegeben, vgl. Explosion 2021), wurde dieses für eine weitere Verarbeitung ausgewählt.

Die nun erfassten Entitäten der Named-Entity-Recognition-Funktion wurden mittels der „Displacy“-Funktion (ein Subpackage von spaCy) „.render“ visualisiert:

```
##Darstellung des Textes mit getaggten Entitäten
displacy.render(doc, style="ent", jupyter=True)
```

Vielleicht kommt mal plötzlich eine Ansichtskarte von England LOC bei Euch ORG an? Dort drüben würden wir bestimmt noch anders aufräumen als in Frankreich LOC . – In dieser Gegend sind alle Einwohner zu Hause LOC geblieben, wir liegen z.Zt. in einem kleinen Städtchen Guines LOC (Fiennes PER) da sind viele Flüchtlinge von Calais LOC und auch viele Belgier LOC von der Gegend von Antwerpen LOC .

Abbildung 9: spaCy Named Entity Recognition Q1_Brief_Weiss_1940

Abbildung 9 zeigt das Ergebnis dieser Visualisierung. Die erfassten Entitäten wurden farblich markiert und jeder ein Entity-Label zugewiesen. Auch wenn die meisten Orte richtigerweise mit „LOC“ (Erklärung siehe unten) getaggt wurden, sind schon auf den ersten Blick einige Fehlklassifizierungen zu erkennen: Das Wort „Euch“ wird fälschlicherweise als „ORG“ klassifiziert. Dies ist in diesem Fall nicht so relevant, da es sich hierbei ja um keine räumliche Entität handelt. Problematischer ist das bei den Entitäten „Fiennes“ und „Belgier“. Während Fiennes fälschlicherweise als „PER“ getaggt wurde (es gibt zwar einen Schauspieler mit jenem Namen, dieser ist hier aber eindeutig nicht gemeint), wurde das Wort „Belgier“ als „LOC“ klassifiziert, obwohl es sich hier um eine Bezeichnung für die Nationalität einer Gruppe handelt.

Schon dieser kleine Textausschnitt zeigt, dass die Named-Entity-Recognition mittels spaCy zwar gut funktioniert, jedoch auch einige Fehlklassifikationen liefern kann.

Die Entitätenklassen in spaCy unterscheiden sich je nach Sprache und vortrainierter Pipeline. Bei eigens erstellten Pipelines müssten diese selbst erstellt werden. In den in dieser Arbeit relevanten Sprachen (Deutsch und Englisch) drücken diese hierbei Folgendes aus:

Deutsch:

- LOC – Non-GPE locations, mountain ranges, bodies of water
- MISC – Miscellaneous entities, e.g. events, nationalities, products or works of art
- ORG – Companies, agencies, institutions, etc.
- PER – Named person or family

Englisch:

- CARDINAL – Numerals that do not fall under another type
- DATE – Absolute or relative dates or periods
- EVENT – Named hurricanes, battles, wars, sports events, etc.
- FAC – Buildings, airports, highways, bridges, etc.
- GPE – Countries, cities, states
- LANGUAGE – Any named language
- LAW – Named documents made into laws
- LOC – Non-GPE locations, mountain ranges, bodies of water
- MONEY – Monetary values, including unit
- NORP – Nationalities or religious or political groups
- ORDINAL – "first", "second", etc.
- ORG – Companies, agencies, institutions, etc.
- PERCENT – Percentage, including "%"
- PERSON – People, including fictional
- PRODUCT – Objects, vehicles, foods, etc. (not services)
- QUANTITY – Measurements, as of weight or distance
- TIME – Times smaller than a day
- WORK_OF_ART – Titles of books, songs, etc.

(vgl. Explosion 2021)

Dies zeigt schon, wie sehr sich verschiedene Pipelines unterscheiden können. Während die deutschsprachigen ein Mindestmaß an unterscheidbaren Entitäten enthalten, sind die englischsprachigen vielfältiger untergliedert. Dies kann dann auch zu Problemen führen, da im Text (siehe Abbildung 9) das Wort „Belgier“ der Kategorie LOC zugewiesen wird, da es keine andere passende Kategorie gibt. Im Englischen wäre hierfür die Kategorie „NORP“ vorhanden. Die Named-Entity-Recognition eines anderen Briefes zeigt dies anschaulich (siehe Abbildung 10). Dieser englischsprachige Brief (Volltext siehe Anhang 11.1.4) stammt aus dem Ersten Weltkrieg.

anyone from coming but I warn him to bring a knife and fork with him. After the first day
DATE we were issued with kits and we treasure our knives and forks now. We came
across on the mail boat Hantonia GPE from Southampton GPE to Havre GPE
... We had to sleep in huts there on the docks but the sleeping accommodation was
better. We Audit fellows were able to keep together till last Saturday night TIME ,
when Pond and I were sent off to different railheads. He is at 3rd Corps Railhead ORG
. We had to travel in horseboxes and during night time. We travelled in this manner
Saturday, Sunday and DATE Monday DATE nights TIME . French NORP
shunting is severe and we were shaken up. Saturday DATE and Sunday DATE we
were about 30 CARDINAL in the trucks, Monday 17 DATE .

Abbildung 10: spaCy Named Entity Recognition Englisch

Hier ist eine größere Anzahl an unterschiedlichen Kategorien für die einzelnen Entitäten ausgewiesen. Zum Beispiel werden hier Zeitangaben, welche ebenso für eine zeitliche Einordnung von geschichtswissenschaftlichen Quellen relevant ist, ausgewiesen. Außerdem wird das Wort „French“ nun richtigerweise der Kategorie „NORP“ zugewiesen und kann somit nicht fälschlicherweise als Bezeichnung eines Landes klassifiziert werden. Falls man sich mit bestimmten Entitäten beschäftigen möchte, welche nicht Teil der vorhandenen Pipelines sind, können vortrainierte Pipelines durch weiteres Trainieren noch um weiteren Entitätenklassen erweitert werden. Dies ist jedoch sehr zeitaufwändig und benötigt einiges an Erfahrung.

Somit werden für die weiteren Schritte die ausgewiesenen Entitäten auf die Kategorien „GPE“ und „LOC“ eingeschränkt, da diese räumliche Entitäten darstellen sollten.

Hierfür wurden alle erfassten Entitäten als Dataframe gespeichert. Dieser wurde dann auf die Entitätsklassen „GPE“ und „LOC“ gefiltert und dann in eine Liste „places_list“ ausgegeben. Diese Liste wird für die weitere Verarbeitung mittels geocoder benötigt.

```
## Abspeichern der Liste list_test als Dataframe df
df=pd.DataFrame(list_test, columns=['ents', 'types'])
df

## Extraktion der Räumlichen Entitäten (LOC und GPE) und Abspeichern
als Dataframe df_places
df_places=df.loc[df['types'].isin(['LOC', 'GPE'])]

## Extraktion der Entitäten (Namen der Orte) aus der Liste, Wegfall
des Typs, Abspeicherung als places_list
places_list = df_places['ents'].tolist()
places_list
```

7.3.2. Toponym Resolution

Bei der Toponym Resolution geht es nun darum, die via Toponym Extraction extrahierten Entitäten, welche als solche mit Raumbezug klassifiziert wurden, mittels Geocoding Koordinaten zur Verortung zuzuweisen. Hierfür wurde das Python-Package „Geocoder“ gewählt. Mit diesem ist es möglich, Geocoding mit unterschiedlichen online verfügbaren Gazetteers durchzuführen. Beispiele hierfür sind: ArcGIS, BingMaps, Google oder OpenStreetMap. (vgl. Denis CARRIERE 2013) Für die vorliegende Geocoding-Aufgabe wurde jedoch der Online-Gazetteer „GeoNames“ (siehe Abbildung 11) gewählt.

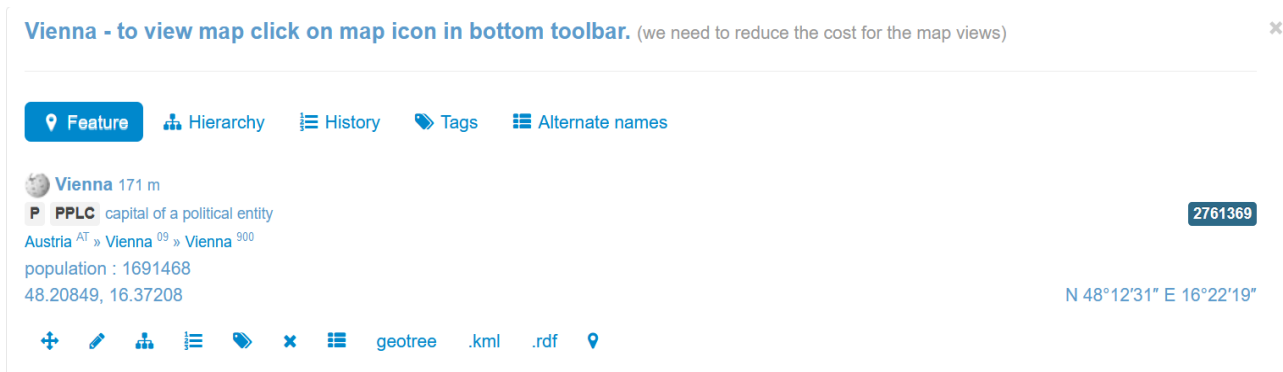


Abbildung 11: Ausschnitt GeoNames (Denis CARRIERE 2013)

Dieser ist unter der Lizenz „CC BY 4.0“ frei verfügbar und bietet selbst Schnittstellen, um die vorhandenen Daten für Geocoding weiterverwenden zu können. Die Daten können auch frei heruntergeladen werden. Dieser Gazetteer bietet über elf Millionen Ortsnamen und zusätzlich

auch Informationen zur Hierarchie dieser Toponyme, ebenso wie Informationen zu Exonymen, Koordinaten, Bevölkerungsinformationen, Höhenangaben oder sogar zum Wetter. (vgl. geo-names.org o.J.) Die vorhandenen Features sind dabei in unterschiedliche Klassen eingeteilt:

- A – country, state, region, ...
- H – stream, lake, ...
- L – parks, area, ...
- P – city, village, ...
- R – road, railroad, ...
- S – spot, building, farm, ...
- T – mountain, hill, rock, ...
- U – undersea
- V – forest, heath, ...

(vgl. Denis CARRIERE 2013)

Diese werden noch weiter unterteilt, dies ist hier aber nicht relevant und wird daher ausgelassen.

Ein großer Vorteil bei der Nutzung von GeoNames sind die validen Datenquellen. Für Österreich zum Beispiel stammen die Daten vom Bundesamt für Eich- und Vermessungswesen, den einzelnen Ländern und der Statistik Austria. Aber auch bei allen anderen Daten sind die Quellen zumeist offizielle Stellen. Dies bedeutet, dass die Daten von GeoNames sehr vertrauenswürdig sind. Um nun die Webservices der Website nutzen zu können, ist es notwendig, ein Nutzerkonto zu erstellen und sich zur Nutzung der frei verfügbaren Webservices freizuschalten. Dies ist ohne viel Aufwand möglich und wird mittels Argument „key“ der geocoder-Funktion übergeben. Ein Nachteil bei der Arbeit mit GeoNames ist ein Limit der täglichen Abfragen auf 20.000 und der stündlichen auf 1.000, solange kostenlose Webservices genutzt werden. Somit ist ein einfaches Gazetteer-Lookup basiertes Geoparsing (siehe Kapitel 6.1.3.1) auf diesem Wege nicht möglich. Hierzu müssten erst zusätzliche Abfragen zugekauft werden. (vgl. ebd.)

Eine Abfrage mittels Geocoder über GeoNames sieht nun wie folgt aus:


```

## Überprüfung einzelner Ergebnisse mittels Geocoding um zu sehen, ob
dies auch funktioniert

## Filterung auf Kontinent Europa und FeatureClasses A, H, L, P, R,
T, V

g = geocoder.geonames(places_list[0],continentCode=["EU"], fea-
tureClass=["A","H","L","P","R","T","V"], key='XXX')

## Abfrage einzelner Attribute des codierten Ergebnisses
g.address

## Überprüfung der Koordinaten (Länge und Breite in Grad) für die
Durchführung der Verortung
print("Breite/Latitude = "+g.latlng[0], "\nLänge/Longitude = "+g.lat-
lng[1])

```

Einerseits wird hier der erste Wert der Entitäten-Liste („Calais“) der Funktion übergeben. Mittels Filterung auf den Kontinent Europa ist sichergestellt, dass nur Ergebnisse in jenem Gebiet in den Geocoding-Prozess einfließen. Dies ist notwendig, da viele Ortsnamen in Europa und Nordamerika gleich sind und es dadurch sonst zu Verwechslungen kommen könnte. Außerdem wird noch auf einige der oben genannten Featureklassen gefiltert, um so falsche Verortungen mit nicht benötigten Klassen (S und U) auszuschließen. Das Ergebnis davon ist dann eine Variable, welche unterschiedliche Attribute enthält. Diese können zum Beispiel der erfasste Name (g.address), eine ID-Nummer, die Bevölkerungsanzahl, das zugehörige Land oder die zugehörigen Koordinaten sein. Diese sind mit dem Attribut „latlng“ abrufbar. Die Koordinaten liegen im Koordinatensystem „World Geodetic System 1984“ vor.

Um nun eine Geokodierung der gesamten Entitäten zu ermöglichen, wurde eine Schleife programmiert, welche oben beschriebene Schritte für die gesamte Liste „places_list“ durchführt. Die geokodierten Ergebnisse werden dann einem zuvor erstellten Dataframe „df_geocoded“ hinzugefügt, in welchem dann die relevanten Attribute „name“, „lon“ und „lat“ beigefügt werden. Dieser enthält zusätzlich Originalwerte der jeweils mittels Named Entity Recognition erfassten räumlichen Entitäten, um diese später auch nochmals überprüfen zu können.

```

for x in range(len(places_list)): ## Limitierung der Schleife auf die
Anzahl der Entitäten in places_list

    ## Geocoding der einzelnen Entitäten (x) in places_list

    g = geocoder.geonames(places_list[x],continentCode=["EU"], fea-
tureClass=["A","H","L","P","R","T","V"], key='XXX')

```

```

    print(g.address, g.feature_class) ## Ausgabe des gematchten Ortes
    und des Typs zur Überprüfung
    if g.address is None:
        print(places_list[x]+": Entity not geocoded!")
        ## Schreiben der Zeile (new_row, welche dann dem Dataframe
        hinzugefügt wird (Werte: Name, Länge, Breite)
    else:
        new_row = {"name": g.address, "lat": g.latlng[0], "lon":
g.latlng[1], "original": places_list[x]}
        #continue
        ## Hinzufügen der erstellten Reihe zum Datframe df_geocoded
        df_geocoded = df_geocoded.append(new_row, ignore_index=True)
        print("New Row has been added-----")
--")

## Ausgabe des fertiggestellten Dataframes
df_geocoded

## Änderung der Spalten lon und lat in Zahlenwerte (float)
df_geocoded["lat"]=df_geocoded["lat"].astype(float)
df_geocoded["lon"]=df_geocoded["lon"].astype(float)

## Ausgabe der Datentypen um zu überprüfen, ob Länge und Breite nun
Zahlenwerte sind
df_geocoded.info()

```

Entitäten, bei welchen das Geocoding zu keinem Ergebnis führte, werden nicht in den Dataframe übergeben. Dadurch werden zuvor durch spaCy fälschlicherweise getaggte Entitäten großteils noch ausgeschieden. Zum Beispiel wurde mittels spaCy das Wort „Millionenwerte“ mit dem Typ „LOC“ getaggt. Dieser fand beim Geocoding jedoch kein Gegenstück und wurde daher nicht in den Dataframe übergeben. Trotzdem gab es bei der Quelle „Q1_Brief_Weiss_1940“ am Ende noch eine Fehlkodierung, da das Wort „Hause“, stammend von den Begriffen „zu/nach Hause“, fälschlicherweise bei der Toponym Extraction (hier noch als undefinierte Ortsbezeichnung des Zuhauses) und der Toponym Resolution als Ortsname erfasst wurde. Dies ist ein Berg in Norwegen und hat inhaltlich mit dem Brief natürlich nichts zu tun. Daher muss dieser noch aus dem Ergebnis manuell gelöscht werden.

Anschließend müssen die Spalten „lat“ und „lon“ noch in Zahlenwerte (float) umgewandelt werden, um diese dann weiterverarbeiten zu können. Das Ergebnis wird dann in einen Geodataframe umgewandelt und als Shapefile exportiert, um es dann auch im MapViewer anzeigen und nutzen zu können.

```
## Umwandlung des Dataframes in einen Geodataframe mittels geopandas
gdf = geopandas.GeoDataFrame(df_geocoded, geometry=geopandas.points_from_xy(df_geocoded.lon, df_geocoded.lat),
                             crs="epsg:4326")

## Überprüfung des Geodataframes
print(gdf.head())

## Export des Geodataframes als shapefile
gdf.to_file(dir_q+filename+".shp", driver="ESRI Shapefile")
```

Ein Versuch, nun im Sinne des Gazetteer-Lookup basierten Geoparsings (siehe Kapitel 6.1.3.1) vorzugehen, indem statt dem GeoNames-Gazetteer jener von OSM (über Geocoder) gewählt wurde, zeigt das Problem jener Vorgangsweise, welche im Folgenden beschrieben wird. Dafür galt es nun den gesamten Text in eine Liste zu schreiben, wodurch dann jedes Wort abgefragt werden konnte.

```
# Python code to convert text to list
q1 = q1.replace("\n", " ")

## Quelle: https://www.geeksforgeeks.org/python-program-convert-string-list/
def Convert(string):
    li = list(string.split(" "))
    return li

# Driver code
q1_list = Convert(q1)
print(q1_list)
```

Die Abfrage fand dabei ebenfalls mittels schon oben beschriebener Schleife statt, nur das nun der OSM-Geocoder verwendet wurde:

```
g = geocoder.osm(q1_list[x])
```

Das Ergebnis dieses Prozesses (siehe Abbildung 12) zeigt schon die schwere Bearbeitbarkeit auf diesem Wege. Schon die ersten zwei Wörter im Text „Im Felde“ (die Originalwörter sind in der vierten Spalte zu sehen) führten zu zwei Fehlkodierungen, da „Im“ mit „Indien“ und „Felde“ mit der gleichnamigen Gemeinde in Deutschland kodiert wurde. Dieser Vorgang ist daher nicht alltagstauglich, vor allem auch da durch eine Überprüfung jedes einzelnen Wortes das Geokodieren einen erheblich größeren Zeitaufwand benötigt. Daher ist eine Vorprozessierung mittels Named Entity Recognition zu empfehlen, da dadurch das Ergebnis schneller verfügbar und inhaltlich richtiger ist.

	name	lon	lat	original
0	India	78.667743	22.351115	Im
1	Felde, Achterwehr, Rendsburg-Eckernförde, Schl...	9.933327	54.300000	Felde,
2	Haute-Garonne, Occitanie, France métropolitain...	0.971679	43.305455	31.
3	Maillé, Fontenay-le-Comte, Vendée, Pays de la ...	-0.789006	46.342367	Mai
4	Landes, Nouvelle-Aquitaine, France métropolita...	-0.643387	44.009969	40.
5	Meine, Samtgemeinde Papenteich, Gifhorn, Niede...	10.539227	52.384917	Meine
6	Liben, ሰማሌ ክልል / Somali, ኢትዮጵያ	40.711791	4.544255	Lieben!
7	Schön, Archshofen, Creglingen, Main-Tauber-Kre...	10.082689	49.456998	Schon
8	五百米口径球面射电望远镜, 克度镇, 罗甸县, 黔南布依族苗族自治州, 贵州省, 558...	106.856567	25.652853	fast
9	パイオニア株式会社, 8, 不忍通り, 千石三丁目, 文京区, 東京都, 113-0021, 日本	139.747396	35.729798	8
10	Tage, Saint-Avit-Rivière, Bergerac, Dordogne, ...	0.897046	44.734108	Tage
11	Üçlerkayası, İhsaniye, Afyonkarahisar, Ege Böl...	30.417016	39.086606	liegen

Abbildung 12: Gazetteer-Lookup basiertes Geoparsing

Anschließend wurde außerdem ein Vergleich beider Gazetteers (GeoNames und OpenStreet-Map, Abbildungen 13 und 14) durchgeführt. Hierfür wurde das Ergebnis der Named Entity Recognition mit beiden Gazetteers verortet. Die Verortung mittels GeoNames ist, wie in den Abbildungen 13 und 14 zu sehen ist, weit weniger gestreut als das Ergebnis mittels OSM-Gazetteer und an den Fahrtweg Pfeiffers entlang der Donau. Auch wenn die OSM-Daten viel umfangreicher sind (21 Mio. zu 11. Mio.) und auch nicht offizielle Namen enthalten können, ist dem GeoNames-Gazetteer vertrauenswürdiger, da dieser Daten offizieller Stellen enthält. Daher wurde dieser Gazetteer für weitere Verortungsschritte genutzt.

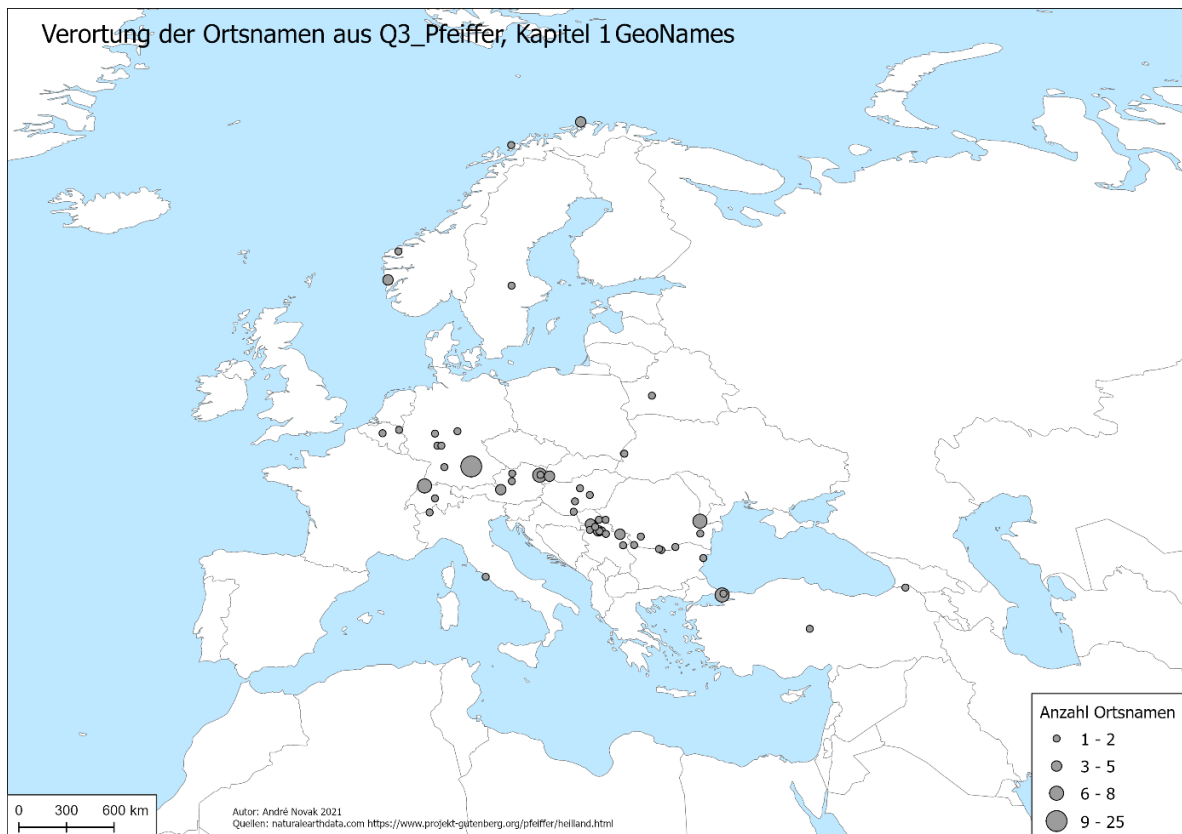


Abbildung 13: Verortung mit GeoNames-Gazetteer

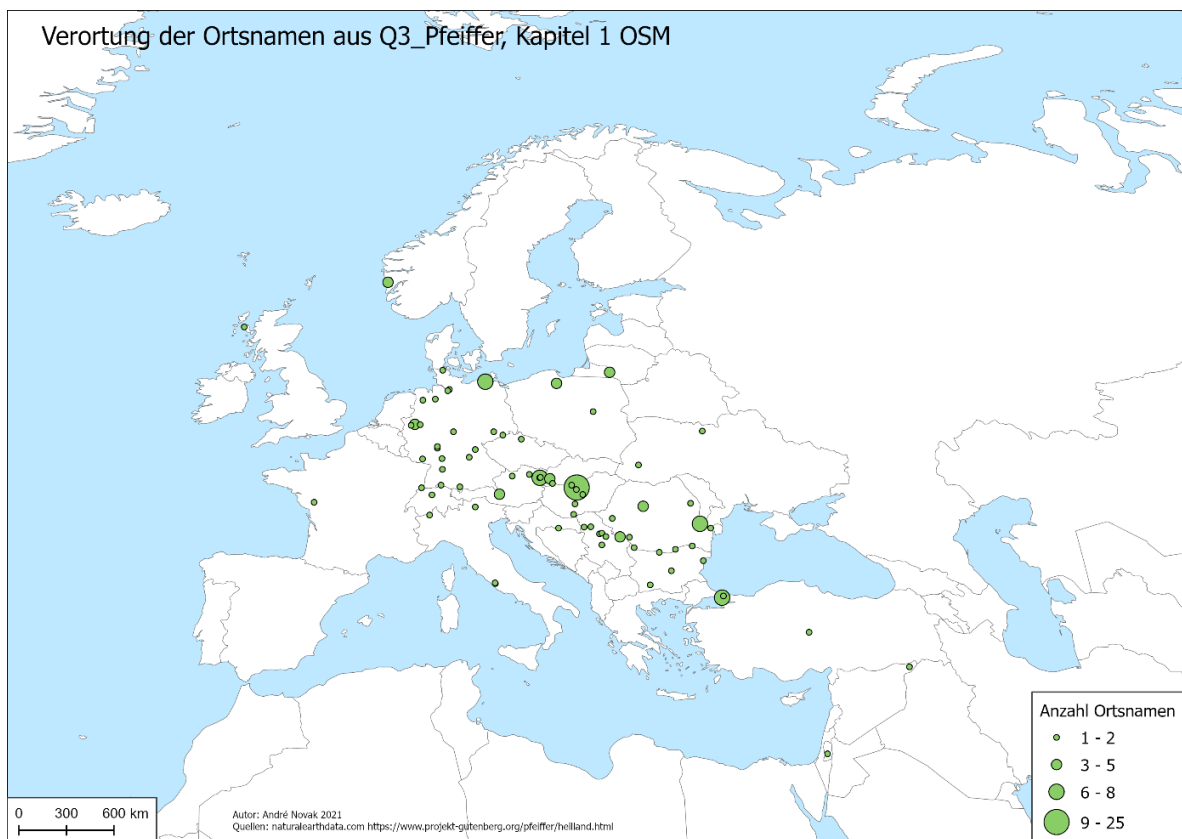


Abbildung 14: Verortung mit OSM-Gazetteer

7.4. Ergebnisse und deren Visualisierung

Das Ergebnis des Geoparsings, also die verorteten Ortsnamen, werden in diesem Kapitel mittels Karten visualisiert. Dabei werden gleiche Ortsnamen in einem Symbol zusammengefasst, welches bei steigender Anzahl größer dargestellt wird. Als Projektion für alle Karten wurde dabei jene von Robinson (epsg = 54030) ausgewählt.

Außerdem wurden statistische Kennzahlen für die Klassifikation berechnet und in mehreren Confusion-Matrix-Tabellen dargestellt. Die Berechnung der einzelnen Werte sieht dabei wie folgt aus:

- True Positive (TP): Das Wort wurde als räumliche Entität klassifiziert und ist tatsächlich eine räumliche Entität.
- True Negative (TN): Das Wort wurde nicht als räumliche Entität klassifiziert und ist tatsächlich keine räumliche Entität.
- False Positive (FP): Das Wort wurde als räumliche Entität klassifiziert, ist aber keine räumliche Entität.
- False Negative (FN): Das Wort wurde nicht als räumliche Entität klassifiziert, aber ist eigentlich eine räumliche Entität.
- Precision: Dieser Wert gibt den Anteil der korrekt als räumliche Entitäten klassifizierten Werte an all jenen, welche als räumliche Entitäten klassifiziert wurden, wieder. Berechnung: $\frac{TP}{TP+FP}$
- Accuracy: Dieser Wert gibt die richtige Klassifizierung in nicht-räumliche und räumliche Entitäten wieder. Berechnung: $\frac{TP+TN}{\text{Gesamtzahl}}$
- Recall: Dieser Wert gibt den Anteil der klassifizierten räumlichen Entitäten an all jenen wieder, welche tatsächlich räumliche Entitäten sind. Berechnung: $\frac{TP}{TP+FN}$

(vgl. HOFMANN und CHISHOLM 2015: 202f)

7.4.1. Q1_Brief_Weiss_1940

Aus dem Text Q1_Brief_Weiss_1940 wurden mittels spaCy insgesamt 20 räumliche Entitäten extrahiert: Calais, Calais, Dünkirchen, Calais, **Millionenwerte**, England, Calais, England, England, England, Frankreich, **Hause**, Guines, Calais, **Belgier**, Antwerpen, **Belgier**, **Engländer**, **Engländer**, **Hause**. Tabelle 8 zeigt die Confusion Matrix nach der Named Entity Recognition. Es wurden alle räumlichen Entitäten (13) richtig getaggt. Zusätzlich zu diesen wurden weitere sieben fälschlicherweise als räumliche Entitäten getaggt. Diese sind oben in rot markiert. Dies zeigt, dass bei diesem Beispiel ca. 1/3 der Entitäten fälschlicherweise als räumliche Entität interpretiert wurde. Wobei bei der Entität „Hause“ im Sinne von „nach/zu Hause“ eigentlich schon von einer räumlichen gesprochen werden kann. Diese ist jedoch zu unbestimmt, um sie mittels Geocoding genauer verorten zu können.

Der Accuracy-Wert, welcher die richtige Klassifizierung in nicht-räumliche Entitäten und räumliche Entitäten angibt, liegt bei 0,9872. Im Gegensatz dazu liegt die Precision der räumlichen Entitäten bei 0,65. Dieser Wert beschreibt den Anteil der korrekt als räumliche Entität klassifizierten Werte an all jenen, welche als räumliche Entität klassifiziert wurden. Dies bedeutet also, dass nach der Named Entity Recognition noch gut ein Drittel der als räumliche Entität klassifizierten Werte eigentlich keine sind. Bei diesem Beispiel ergab der Recall 1, das heißt, dass alle räumlichen Entitäten im Text durch die NER mittels spaCy erfasst wurden.

Tabelle 8: Confusion Matrix Q1_Brief_Weiss_1940 nach NER

Räumliche Entitäten	Predicted: NO	Predicted: YES	
Actual: NO	530	7	537
Actual: YES	0	13	13
n=550	530	20	

Precision = 0,65

Accuracy = 0,9872

Recall = 1

Diese Entitäten wurden nun mittels geocoder und GeoNames geokodiert. Dabei kamen folgende Ergebnisse heraus: Calais, Calais, Dunkirk, Calais, London, Calais, London, London, London, France, **Hause**, Guînes, Calais, Antwerp, **Hause**. (Ergebnis: siehe Anhang 11.3.1) Dies

zeigt, dass durch eine Kombination einer Toponym Recognition und einer Toponym Resolution mittels Gazetteer ein großer Teil der zuvor falsch klassifizierten Entitäten getilgt werden konnten.

Tabelle 9: Confusion Matrix Q1_Brief_Weiss_1940 nach Geocoding

Räumliche Entitäten	Predicted: NO	Predicted: YES	
Actual: NO	535	2	537
Actual: YES	0	13	13
n=550	535	15	

Precision = 0,86666

Accuracy = 0,9963

Recall = 1

Tabelle 9 zeigt die Confusion Matrix nach der Geokodierung mittels GeoNames. Während der Recall-Wert gleich blieb und der Accuracy-Wert leicht gestiegen ist (auf 0,9963), stieg der Precision-Wert jedoch deutlich auf 0,867. Dies bedeutet, dass ein nach der Named Entity Recognition durchgeführtes Geocoding nochmals zu einer Qualitätssteigerung führte, da somit Entitäten, welche fälschlicherweise als räumlich klassifiziert wurden (z.B.: Millionenwerte), nun richtigerweise ausgeschieden wurden.

Name	Country	Feature class	Latitude	Longitude
1 New England Novaja Anglija,Новая Англия	United States	region population 15,000,000	N 44° 0' 0"	W 71° 0' 0"
2 East of England	United Kingdom , England	region population 6,130,542	N 52° 14' 26"	E 0° 30' 14"
3 New England Region - 2-digit Hydrologic Unit Code - 01	United States	area	N 44° 12' 20"	W 70° 18' 23"
4 London London, LON, Lakana, Landan, Landen, Ljondan, Llundain, Lodonj, Londain, London, Londer, Londe, Londen, Londin, L...	United Kingdom , England Greater London	capital of a political entity population 7,556,900	N 51° 30' 30"	W 0° 7' 32"
5 England 'Ingland, AEnglound, Anghiltera, Anghiltèra, Anglaterra, Anglatèrra, Anglete, Angletterre, Angletè, Angli, Ang...	United Kingdom , England	first-order administrative division population 56,286,961	N 52° 9' 37"	W 0° 42' 11"
6 Great Britain Bretainia Haundia, Britain, Britio, Gran Bretagna, Gran Bretana, Gran Bretaña, Grand Bretagne, Great Britai...	United Kingdom , England	island population 64,163,534	N 54° 0' 0"	W 2° 0' 0"
7 Birmingham BHX, Birmin'gkam, Birmingham, Birmingamas, Birmingem, Birmingema, Birmingham, Birminghamia, Birminghamo, Birmi...	United Kingdom , England Birmingham	seat of a second-order administrative division population 984,333	N 52° 28' 53"	W 1° 53' 59"
8 Old England	Jamaica , Manchester Old England	second-order administrative division population 4,788	N 17° 59' 6"	W 77° 27' 36"

Abbildung 15: Screenshot aus GeoNames (Bildquelle: geonames.org o.J.)

Trotzdem gibt es weiterhin zwei Problemfälle. Einerseits wurde das Wort „Hause“ nun mit einem Berg in Norwegen mit dem gleichen Namen geokodiert. Dies zeigt das Problem, wenn Toponyme genauso heißen wie Wörter, die etwas ganz anderes bezeichnen. Auch wenn durch eine Named Entity Recognition dadurch viele Problemfälle eliminiert wurden, ist bei der Quelle Q1_Brief_Weiss_1940 noch immer eine falsch geokodierte Entität vorhanden. Der zweite (kleinere) Problemfall ist die Geokodierung des Wortes „England“ mit dem Toponym

„London“. Warum GeoNames hier nicht auf den Landesteil Großbritanniens sondern auf dessen Hauptstadt verlinkt, ist nicht ganz klar, da ja auch „Frankreich“ mit „France“ richtig erkannt wurde. Dieser Fall ist jedoch nicht ganz so ausschlaggebend, da ja trotzdem eine Verlinkung zum richtigen Gebiet besteht. Ein Blick auf die GeoNames-Website (siehe Abbildung 15) zeigt die Suchergebnisse bei der Suche nach England. Hier ist „New England“ als erstes Ergebnis aufgelistet. Dieses wurde jedoch mittels Eingrenzung auf bestimmte Kontinente im Python-Code ausgeschieden. London wurde wohl durch die Geokodierung gewählt, da es in der Ergebnisliste vor England liegt. Warum also die Hauptstadt bei der Suche nach einem Landesteil vor diesem angezeigt wird, ist fraglich und sollte daher in GeoNames nochmals überdacht werden.

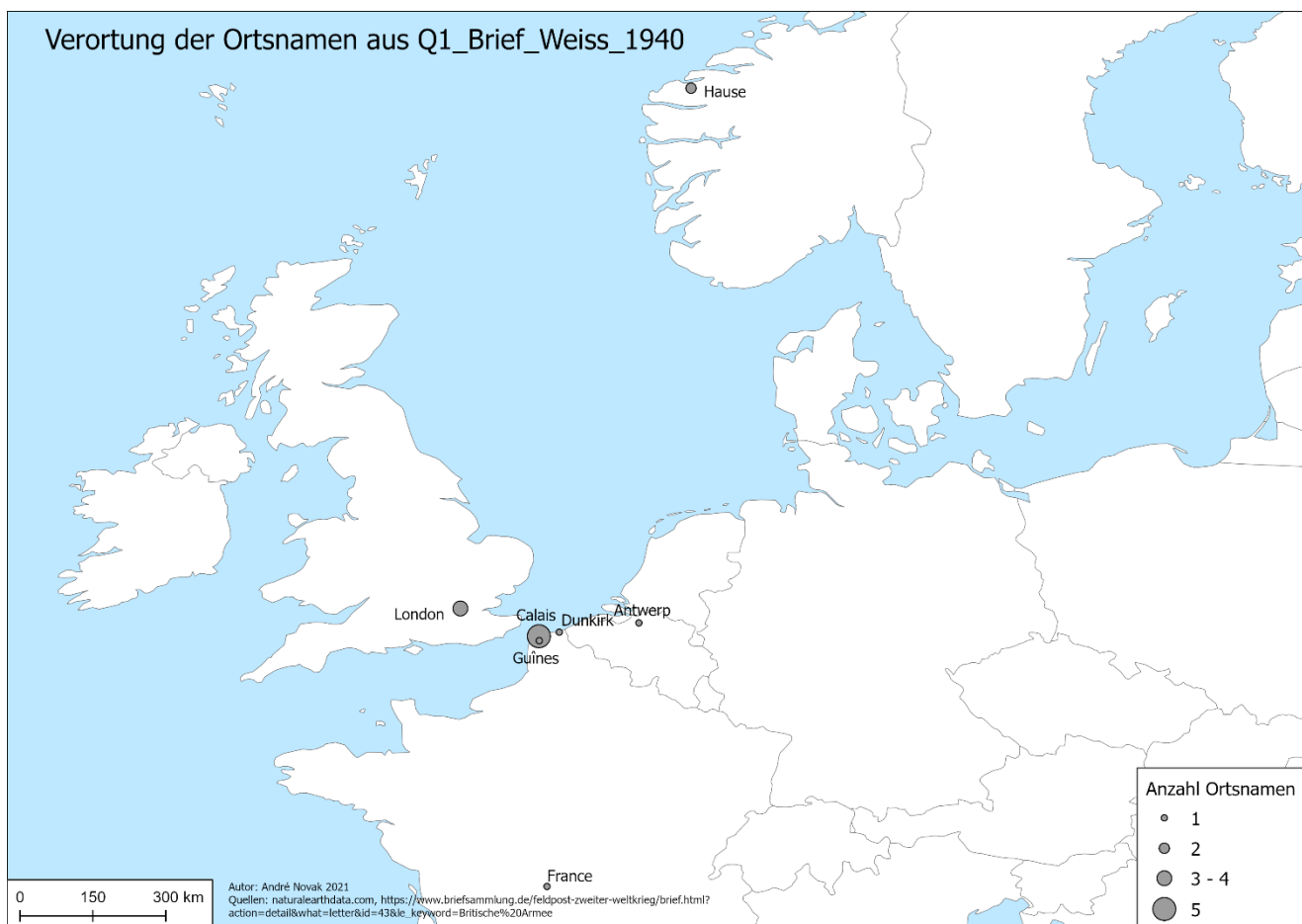


Abbildung 16: Karte Verortung Q1_Brief_Weiss_1940

Abbildung 16 zeigt das Ergebnis des Geoparsings mittels in Kapitel 7 beschriebenen Workflows. Zu sehen ist eine Konzentration des Ergebnisses im Bereich Nordost-Frankreich. Wenn das falsch erkannte Feature händisch eliminiert wird, ist das Ergebnis aussagekräftig für eine Verortung der Quelle „Q1_Brief_Weiss_1940“. Bei einer Interpretation des Ergebnisses muss also immer darauf geachtet werden, dass etwaige Fehler des Geoparsings mitbetrachtet

werden. Diese können entweder durch weitere Schritte (siehe Kapitel 7.5) noch eliminiert werden oder müssen bei einer Analyse des Ergebnisses als Fehler deklariert werden. Außerdem ist wichtig anzumerken, dass dieses Ergebnis nicht gleichzusetzen ist mit dem tatsächlichen Bewegungsradius des Verfassers, da dieser im Text ja auch auf England Bezug nimmt, obwohl er bis zum Verfassen dieses Briefes nicht dort war. Daher ist das Ergebnis des Geoparsing eher zu sehen als eine Verortung aller im Text vorkommenden Toponyme als eine Verortung des Verfassers im zweiten Weltkrieg durch seinen Brief. Somit könnte ein Radius/Horizont gezogen werden, in welchem sich der Verfasser gedanklich befunden hat. Um nun eine umfangreichere Analyse über Otto Weiß und seinen Bewegungs- und Gedankenhorizont im Zweiten Weltkrieg durchführen zu können, müssten dazu weitere Briefe von ihm analysiert werden. Eine vollständige Auflistung der einzelnen Geoparsing-Ergebnisse als Tabelle sind in Anhang 11.3 nachzulesen.

7.4.2. Q2_NFP_24091886

Das Geoparsing der ersten zwei Seiten der „Neuen Freien Presse“ vom 24.09.1886 führte zur Extraktion von 96 räumlichen Entitäten. (Ergebnis: siehe Anhang 11.3.2) Die verwendete Textgrundlage wurde unverändert aus dem ANNO-System der Österreichischen Nationalbibliothek (vgl. Österreichische Nationalbibliothek 2021) übernommen, um so eine Textgrundlage, erschaffen durch OCR-Methoden, verwenden zu können.

Eine Analyse der wichtigsten Kennwerte der Confusion-Matrix (siehe Tabelle 10) zeigt, dass der Accuracy-Wert bei 0,9872 liegt. Der Precision-Wert ist mit 0,6875 gut, da durch diese Methode mehr als zwei Drittel der als räumliche Entitäten gekennzeichneten Wörter auch tatsächlich solche sind. Und auch die True-Positive-Rate, also der Recall, weist mit 0,7765 einen ansprechenden Wert auf. Mehr als drei Viertel aller räumlichen Entitäten wurden erkannt.

Tabelle 10: Confusion Matrix Q2_NFP_24091886

Räumliche Entitäten	Predicted: NO	Predicted: YES	
Actual: NO	2671	30	2701
Actual: YES	19	66	85
n=2786	2690	96	

Precision = 0,6875

Accuracy = 0,9872

Recall = 0,7765

In der Arbeit mit OCR-Methoden transkribierten Texten ergeben sich jedoch einige Probleme, welche die optimale Umsetzung des Geoparsings behindern können. Einerseits sorgen diese für eine zeichengetreue Transkription des Originals. Das heißt, dass vor allem bei Zeitungen, aber natürlich auch in anderen Druckwerken das Layout und die Zeichensetzung (im Sinne der Positionierung aller verwendeten Schriftzeichen am Blatt), großen Einfluss auf das Ergebnis des OCR haben. So erstrecken sich zum Beispiel Wörter durch Trennzeichen über mehrere Zeilen. Dies kann dann beim Geoparsing jedoch zu Fehlinterpretationen führen. So wurde im Text zum Beispiel der Name „Wilhelm von Tegetthoff“ wie folgt geteilt: „Will-helm von Tegetthoff“. (o.V. 24.09.1886: 1) Dies führte jedoch zu einem Tagging als Ortsname, da „Wil“ auch eine Stadt in der Schweiz ist. Ebenso gab es zum Beispiel eine Worttrennung bei „Bra-silien“. Dies führte jedoch zu keinen Problemen, da auch „Bra“ mit dem Land in Südamerika geokodiert werden konnte.

Des Weiteren führt die automatische Transkription auch zur Teilung von eigentlich zusammengehörigen Texten. Der Text „Aus Tegetthoff's Mappe (vgl. ebd.: 1f) erstreckt sich am Blatte über zwei Seiten, wird jedoch aufgrund der OCR-Methode geteilt, da hier jede Seite für sich transkribiert wurde. Dadurch kann es zu Verwirrung und Fehlern bei einer automatischen Analyse einzelner Artikel kommen. Probleme, welche sich durch eine automatische Übertragung von Frakturschrifttexten ergeben können, sind schon in Kapitel 7.3.1 aufgeführt worden. Um also das Ergebnis weiter zu verbessern, sollte angedacht werden, den Text manuell nochmals auf das Geoparsing vorzubereiten.

Beispielhaft soll hier noch ein weiteres textspezifisches Problem genannt werden. „Lissa“ bezeichnet im vorliegenden Text eine Seeschlacht aus dem Jahr 1866 („Seeschlacht von Lissa“), durch welche Admiral Tegetthoff unter anderem seine Berühmtheit erlangt hat. (vgl. ebd.: 1f) Diese Entität wird bei der Named Entity Recognition teilweise richtig als räumliche Entität getaggt und teilweise als Personen-Entität. Dies liegt wohl an der Ähnlichkeit zum Namen „Lisa“, welcher hier zu einer Verwechslung führte.

Trotzdem wurde der Text unbearbeitet analysiert, um ebenjene Problemstellen aufzeigen zu können.

Abbildung 17 zeigt das Ergebnis der Verortung der ersten zwei Seiten der Quelle „Q2_NFP_24091886“. Neben einiger weniger Fehlklassifikationen im Bereich Nordamerika und Ostafrika zeigt die Karte gut, in welchen Bereichen Admiral Tegetthoff, welcher in einem Großteil des analysierten Textes aufgrund der geplanten Enthüllung eines Denkmals Ziel der Berichterstattung ist, seinen Handlungsbereich hatte. Neben Brasilien und Mexiko war er vor allem in Europa tätig. Dort ist, neben Wien, wo das Denkmal enthüllt werden sollte und die Zeitung ihren Hauptsitz hat, vor allem im Bereich der Insel „Vis“, wo die „Seeschlacht von Lissa“ stattfand und Helgoland, wo eine weitere wichtige Schlacht in Tegetthoffs Karriere stattfand, relevant. Zusätzlich sind Ortsnamen in Italien und Deutschland sowie die Namen dieser Länder wichtig, da dies die Gegner in den jeweiligen Schlachten waren.

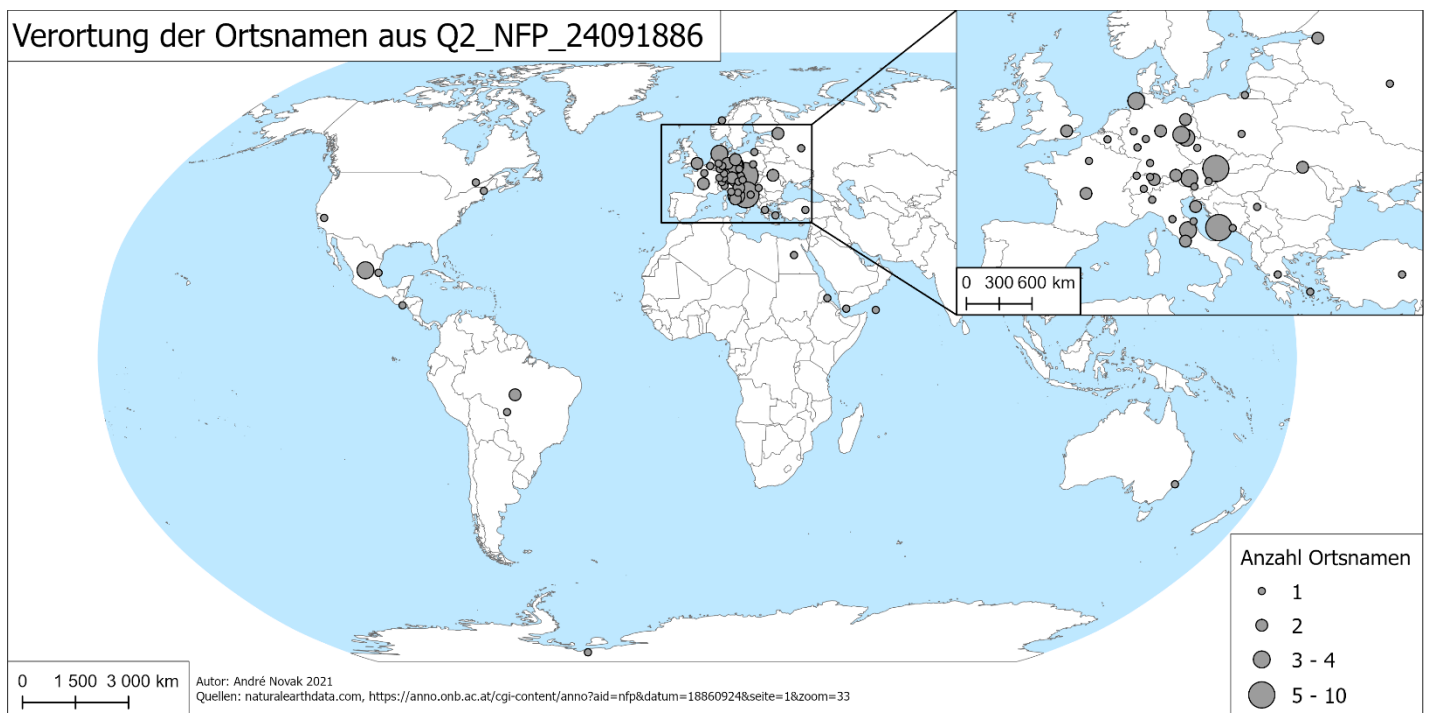


Abbildung 17: Karte Verortung Q2_NFP_24091886

7.4.3. Q3_Pfeiffer

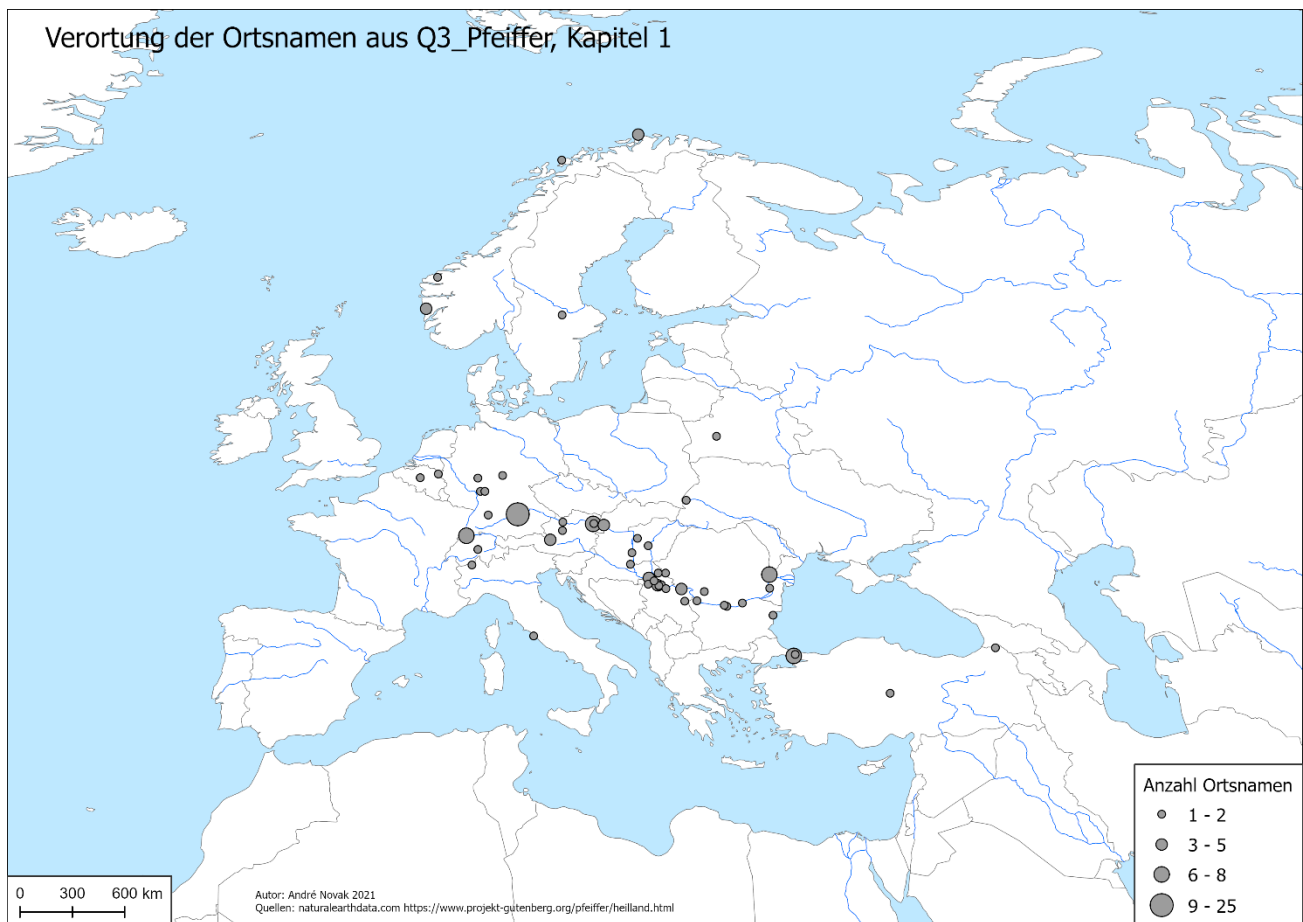


Abbildung 18: Karte Verortung Q3_Pfeiffer, Kapitel 1

Die Verortung des ersten Kapitels der Quelle Q3_Pfeiffer „Reise von Wien nach Konstantinopel“ ist in Abbildung 18 zu sehen. (Ergebnis als Tabelle: siehe Anhang 11.3.3) Auch wenn auch hier einige Fehlklassifikationen durchgeführt wurden (z.B.: „Stadt“ mit „Basel“ oder der Fluss „Donau“ mit „Neuburg an der Donau“), zeigt sich hier klar ein Verlauf entlang der Donau von Wien nach Konstantinopel. Dies bedeutet wohl, dass man davon ausgehen kann, dass Ida Pfeiffer auf einem Schiff auf der Donau nach Konstantinopel fuhr. Ein Blick in die Quelle bestätigt dies. In dieser berichtet Pfeiffer tageweise von ihrer Reise per Schiff entlang der Donau. Auch wenn sie immer wieder Halt in einigen Städten machte, fand der Hauptweg zu Wasser statt. Somit lässt sich über Pfeiffers Reise einiges aussagen, ohne den Text genau gelesen zu haben.

Da Pfeiffer im Kapitel eins der Quelle wie in einem Tagebuch ihren Weg nach Konstantinopel beschreibt, eignet sich diese Quelle besonders gut für das Geoparsing.

Tabelle 11: Confusion Matrix Q3_Pfeiffer, Kapitel 1

Räumliche Entitäten	Predicted: NO	Predicted: YES	
Actual: NO	6255	63	6318
Actual: YES	17	77	94
n=6412	6272	140	

Precision = 0,55

Accuracy = 0,9875

Recall = 0,8191

Ein Blick auf die Confusion-Matrix (Tabelle 11) zeigt, dass sich das Ergebnis der Verortung des ersten Kapitels von Ida Pfeiffers „Reise einer Wienerin in das Heilige Land“ (6412 Wörter) sehen lassen kann. Auch wenn der Precision-Wert mit 0,55 nicht besonders hoch ist, zeigt die Abbildung 18 doch ein aussagekräftiges Ergebnis. Dieser Wert ist so niedrig, da einige Wörter öfters im Text vorkommen und immer wieder zu den gleichen Problemen geführt haben. So führte das Wort „Stadt“ immer wieder zu einer Geokodierung als „Basel“, da „Basel-Stadt“ ein Kanton in der Schweiz ist. Auch die Entität „Pest“, die Krankheit und damals noch eine eigene Stadt in Ungarn, führte immer wieder zu einer Fehlklassifizierung. Die Accuracy liegt bei einem hohen Wert (0,9875) und der Recall bei 0,8191. Probleme gab es hierbei vor allem bei Formulierungen, welche damals üblich waren, aber heutzutage wohl nicht mehr genutzt werden („Um vier Uhr gelangten wir zwar ohne Unglück, aber ganz erstarrt zu Drencova an [...]“ (PFEIFFER 1886) GeoNames deckt einen Großteil der damals üblichen deutschsprachigen Namen für Städte und Gebiete im Bereich der ehemaligen Donaumonarchie überraschend gut ab, wodurch das Geocodieren vieler Entitäten möglich war.

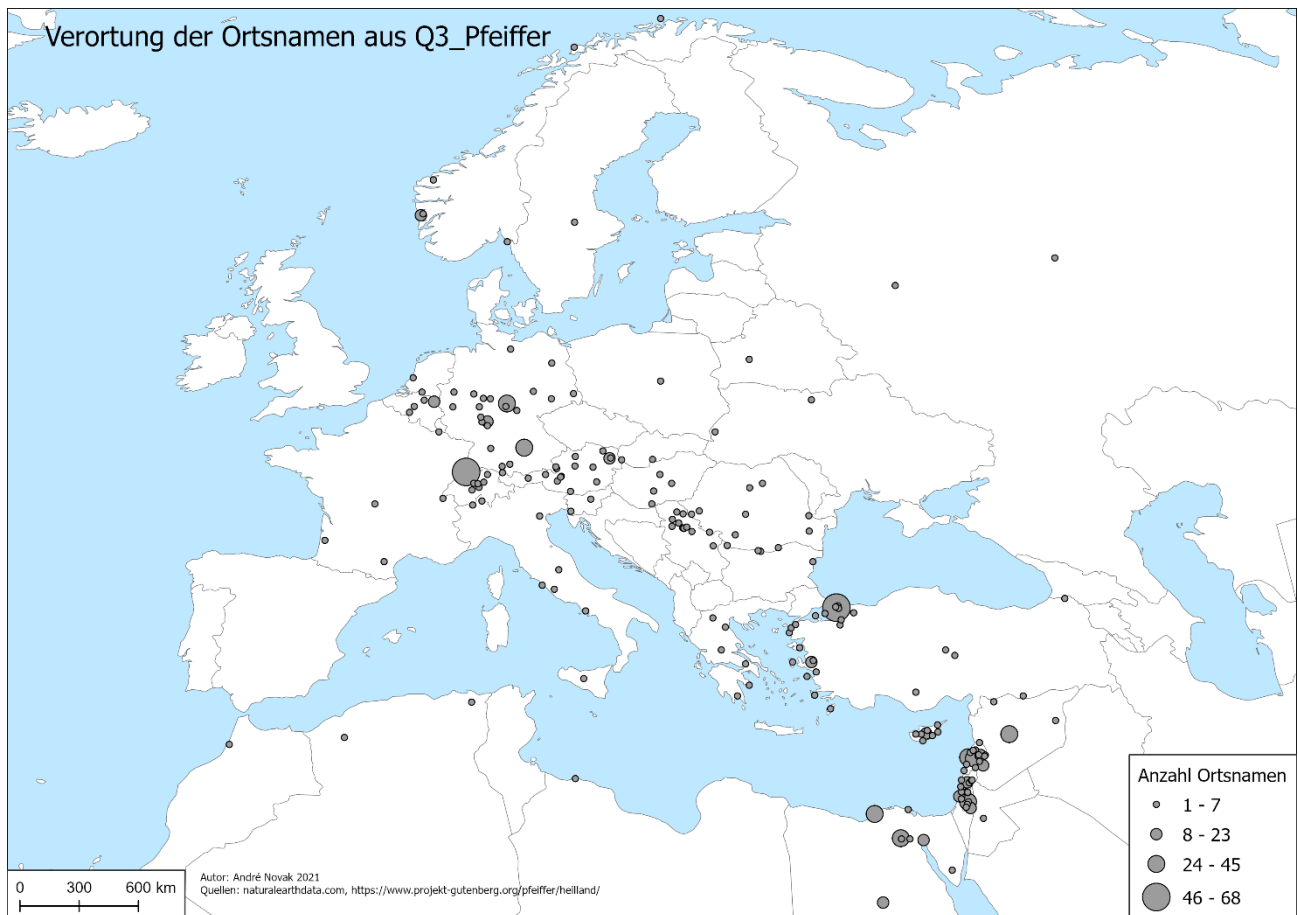


Abbildung 19: Karte Verortung Q3_Pfeiffer

Abbildung 19 zeigt eine Verortung der Ortsnamen des gesamten Textes. Es kam zwar auch zu Verortungen außerhalb des dargestellten Bereiches, da der Text jedoch Geschehnisse in Europa, Nordafrika und im Nahen Osten beschreibt, ist dieser Bereich ausreichend zur Darstellung. In diesem liegen 86,80% der getaggten Entitäten. Neben der Häufung der Werte im Bereich der Donau (siehe Beschreibung des ersten Kapitels oben), sind auch im östlichen Ägäischen Meer, in Zypern und im Bereich des heutigen Israel und Ägypten viele räumliche Entitäten getaggt worden. Daher ist wohl anzunehmen, dass Ida Pfeiffer via Schiff diese Route für ihre Reise ins Heilige Land wählte.

7.4.4. Weitere Visualisierungsmöglichkeiten

Die erfassten Punktdaten lassen sich auch auf verschiedene andere Arten visualisieren. Einerseits ist es möglich, die Verortungen mittels Visualisierungsmöglichkeiten von Punktdaten darzustellen. Hierzu zählt zum Beispiel die „Heatmap“ (siehe Abbildung 20). In dieser werden Punktdaten mittels Farbskala von geringem bis dichtem Vorkommen abgebildet. Dadurch ist eine schnellere visuelle Erfassung wichtiger Bereiche möglich. In der untenstehenden Abbildung ist zu erkennen, dass das Hauptaugenmerk im Werk Pfeiffers auf Konstantinopel und im Nahen Osten gelegen haben muss.

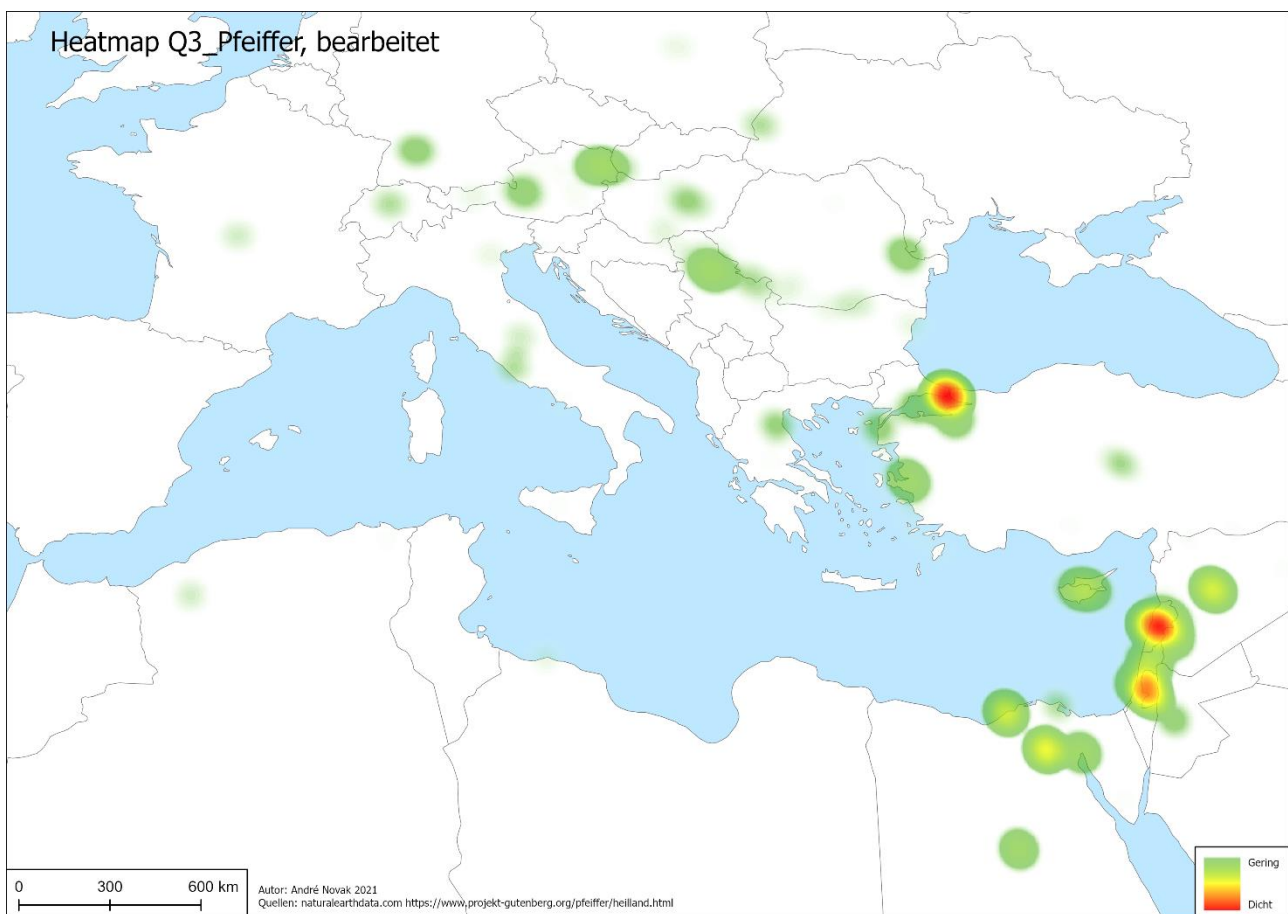


Abbildung 20: Heatmap Q3_Pfeiffer, bearbeitet

Weitere statische Darstellungsmöglichkeiten von Punkten sind unter anderem Kernel-Density-Estimation-Karten, welche auf Basis eines zugrundeliegenden Rasters Dichtewerte (points per unit) darstellen. Dadurch lassen sich Hotspots gut erkennen. Auch ein Clustern von Punktdaten ist möglich, um wichtige Bereiche zu detektieren.

Da es sich bei der Reise Ida Pfeiffers jedoch um kein statisches Ereignis handelte, ist eine Visualisierung mittels Animation eine Möglichkeit, ihren Weg dynamisch darstellen zu können.

Hierfür gäbe es zwei Vorgangsweisen: Einerseits könnten die Orte abhängig von ihrem Vorkommen im Reisebericht eingeblendet beziehungsweise animiert und so der Weg nachvollziehbar gemacht werden. Eine zweite Möglichkeit wäre es, die Orte je nach Datum in die Animation einzubeziehen. Da Pfeiffers Bericht tagebuchähnlich verfasst wurde, böte sich dieser für eine solche Animation an. Außerdem können in einer solchen Animation abhängig vom jeweiligen Ort oder Datum noch zusätzliche Informationen eingeblendet oder auf wichtige Orte gezoomt werden. Dadurch wäre vor allem im Bereich des Nahen Ostens eine übersichtlichere Darstellung der relevanten Orte möglich.

Eine dritte Möglichkeit, den verorteten Text zu visualisieren, wäre eine Verlinkung des Textes mit einer Karte. Wenn der Text gelesen wird, kann so durch gleichzeitiges Betrachten der jeweiligen Orte auf der Karte, welche durch Klicken auf diese im Text auf der Karte markiert würden, ein Verständnis dafür erlangt werden, wo sich Pfeiffer auf ihrer Reise gerade befindet. (vgl. O'SULLIVAN und UNWIN 2010: 63ff)

7.5. Vergleich der Ergebnisse und manuelle Nachbearbeitung

Die Ergebnisse des Geoparsing-Workflows wurden nun durch die Auflistung der Faktoren Precision, Accuracy und Recall evaluiert. Dafür wurden die Werte True Positive, True Negative, False Positive und False Negative in die Berechnung miteinbezogen.

Tabelle 12: Vergleich der Geoparsing-Ergebnisse

	Gesamtzahl	True Positive	True Negative	False Positive	False Negative	Precision	Accuracy	Recall
Q1_Brief_Weiss_1940	550	13	535	2	0	0,8666	0,9963	1
Q2_NFP_24091886	2786	66	2671	30	19	0,6875	0,9872	0,7765
Q3_Pfeiffer_Kapitel1	6412	77	6255	63	17	0,55	0,9875	0,8191

Tabelle 12 zeigt die Ergebnisse der Verortung der unterschiedlichen Quellen mittels Geoparsing. (Einfärbung nach Grenzwerten >0,9 und <0,75) Ein Vergleich zwischen den einzelnen Quellenarten zeigt, dass der Accuracy-Wert überall gleich hoch ist. Dies liegt wohl daran, dass ein Großteil der Wörter immer richtig als nicht-räumliche Entitäten klassifiziert wurden und

dadurch der True-Negative-Anteil sehr hoch ist. Spannender deshalb sind die Werte Precision und Recall. Ersterer beschreibt den True-Positive-Anteil an allen klassifizierten räumlichen Entitäten. Dieser sinkt mit Komplexität und Länge des Textes. Während der Brief von Otto Weiß mit 0,87 noch einen relativ guten Wert hat, ist beim ersten Kapitel aus Ida Pfeiffers „Reise einer Wienerin in das Heilige Land“ dieser Wert nur mehr bei 0,55. Dies könnte unter anderem auch daran liegen, dass viele Wörter Pfeiffers mit Ortsnamen gleichgesetzt wurden, obwohl sie in einem anderen Sinne verwendet worden sind (siehe dazu auch Kapitel 7.4.3) Der Recall-Wert liegt beim Brief bei 1, das heißt, dass hier alle räumlichen Entitäten im Text richtigerweise als solche klassifiziert wurden. Dieser Wert ist beim Zeitungsausschnitt, welcher mittels OCR-Methode transkribiert wurde, am niedrigsten (0,7765). Und auch jener des Pfeiffer-Kapitels ist mit 0,8191 nur unwesentlich höher. Trotzdem zeigt dieser Wert gut, dass bei allen drei Quellen mehr als drei Viertel der in den Texten vorhandenen räumlichen Entitäten auch als solche klassifiziert wurden. Dadurch geben die einzelnen Ergebnisse schon einen guten Überblick über die räumliche Erstreckung der einzelnen Texte.

Ein Vergleich der Verortungen der einzelnen Quellen zeigt, dass sich unterschiedliche Quellentypen unterschiedlich gut für den oben beschriebenen Workflow eignen. Der erste Quellentyp, also Briefe von Soldaten in Kriegen, kann durch die Verortung einen guten Überblick darüber geben, in welchem räumlichen Horizont sich die Gedanken des*der jeweiligen Soldat*in befindet.

Bei allen Quellen gilt es zu beachten, dass die verorteten Toponyme nicht nur die Orte abbilden, an welchen sich der*die Autor*in wirklich aufgehalten hat, sondern auch Orte, welche nur erwähnt worden sind. Dies kann gut anhand der zweiten Quelle, der Zeitung, gezeigt werden, da die Redakteur*innen ja nicht immer am Schauplatz des jeweiligen Handlungsortes des Artikels aufhältig sind. Ebenso verweist Ida Pfeiffer zum Beispiel in ihrem Text mehrmals auf die Stadt Rom, ohne im Zuge der beschriebenen Reise dort gewesen zu sein:

„Der Platz, wo Christus den Weltweisen gezeigt wurde, ist nur einige Schritte davon entfernt. Gegenüber dieser Stelle erhebt sich ein Altar an dem Ort, wo einst die Krippe stand, vor welcher die Hirten Christus anbeteten. [...] Die Krippe selbst befindet sich in **Rom** [eigene Hervorhebung] in der Basilika Santa Maria Maggiore. (PFEIFFER 1886)“

Doch werden auch diese Orte mittels des Geoparsing-Workflows herausgefiltert. Dieser Umstand ist daher in weiterführenden Analysen immer mitzubedenken beziehungsweise manuell herauszufiltern.

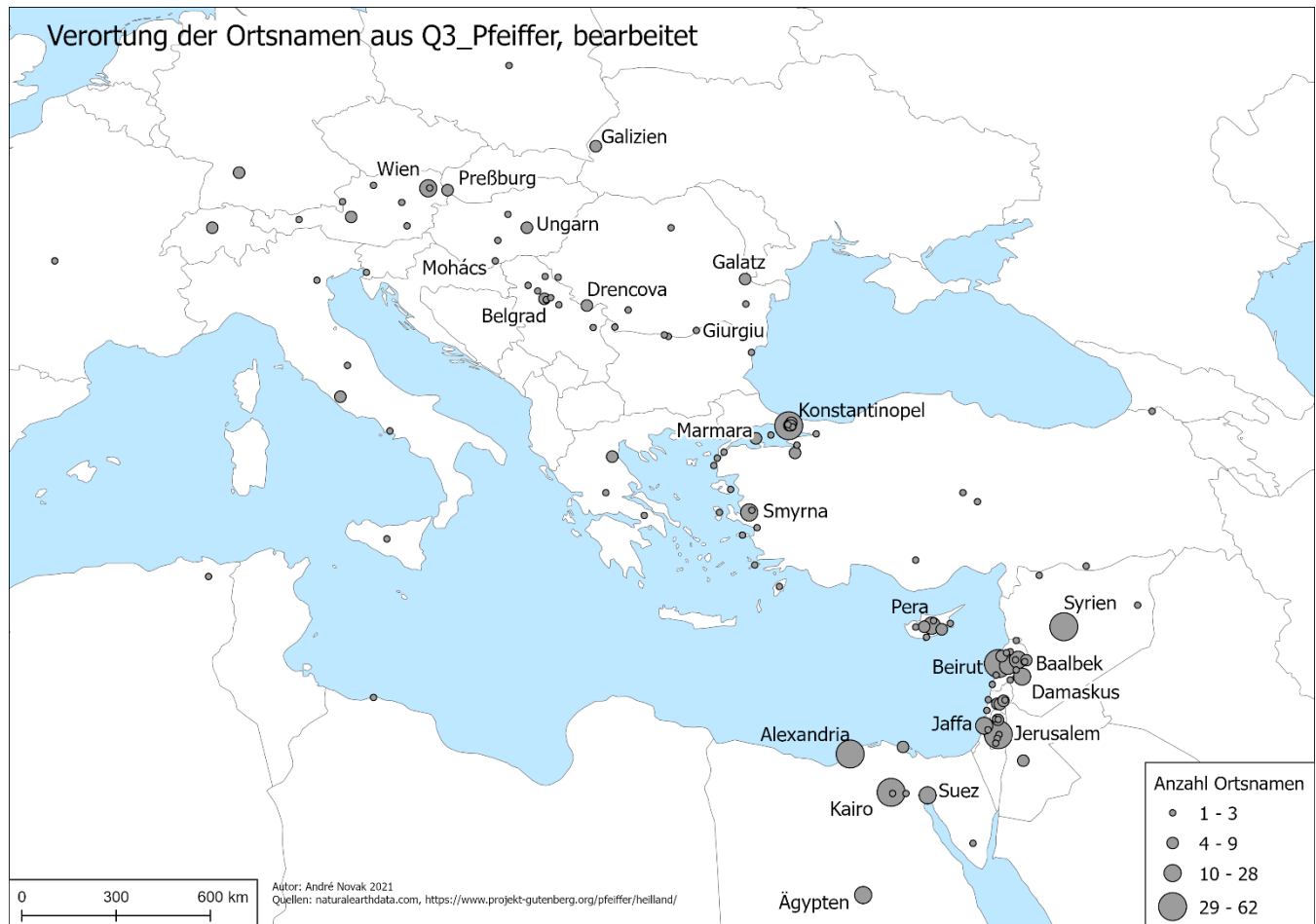


Abbildung 21: Karte Verortung Q3_Pfeiffer, manuell bearbeitet

In Abbildung 21 wurden beispielhaft die fehlerhaften Verortungen manuell herausgelöscht, um einen guten Überblick darüber zu erhalten, inwieweit sich das Ergebnis visuell von dem unbearbeiteten unterscheidet. Ein Vergleich des unbearbeiteten (Abbildung 19) und des bearbeiteten Ergebnisses (Abbildung 21) zeigt einige Unterschiede. Durch die manuelle Bearbeitung ist der Weg, welchen Ida Pfeiffer nach Jerusalem und Ägypten genommen hat, noch deutlicher zu erkennen. Eine gewisse Streuung durch falsch verortete Toponyme oder Wörter, welche fälschlicherweise als Ortsnamen erkannt wurden, fielen dadurch natürlich weg. Es wurden 449 Punktfeatures entfernt. Dies klingt viel, jedoch handelt es sich dabei oftmals um mehrfach vorkommende Features wie zum Beispiel das Wort „Stadt“, welches 68 Mal mit der „Stadt Basel“ verortet wurde. Durch die manuelle Bearbeitung fielen auch viele falsch verortete Ortsnamen im deutschsprachigen Bereich weg. Dies liegt wohl daran, dass es im Deutschen viele

Wörter gibt, welche auch Ortsnamen beziehungsweise Teil dieser sind. Zum Beispiel wurde das Wort „Kreise“ mit dem „Kreiseberg“ in Deutschland verortet. Im Bereich des Nahen Ostens funktionierte die Verortung wiederum sehr gut. In diesem Bereich mussten nahezu keine Features entfernt werden. Dies liegt wohl an der Unterschiedlichkeit der Ortsnamen dort zu Wörtern der deutschen Sprache. Abbildung 22 zeigt die richtig und falsch verorteten Entitäten des Textes. Falsch verortete Ortsnamen sind über die ganze Welt verstreut, jedoch liegt ein Großteil dieser im deutschen Sprachraum. Dies liegt wohl daran, dass Ortsnamen in diesem Gebiet teilweise gleich Wörtern der deutschen Sprache sind. Da klar ist, dass Pfeiffers Reise sie von Europa aus in den Nahen Osten und nach Ägypten führte, können alle Verortungen außerhalb dieses Gebietes getrost entfernt werden. In Europa ist hier mit genauerem Blick zu arbeiten.

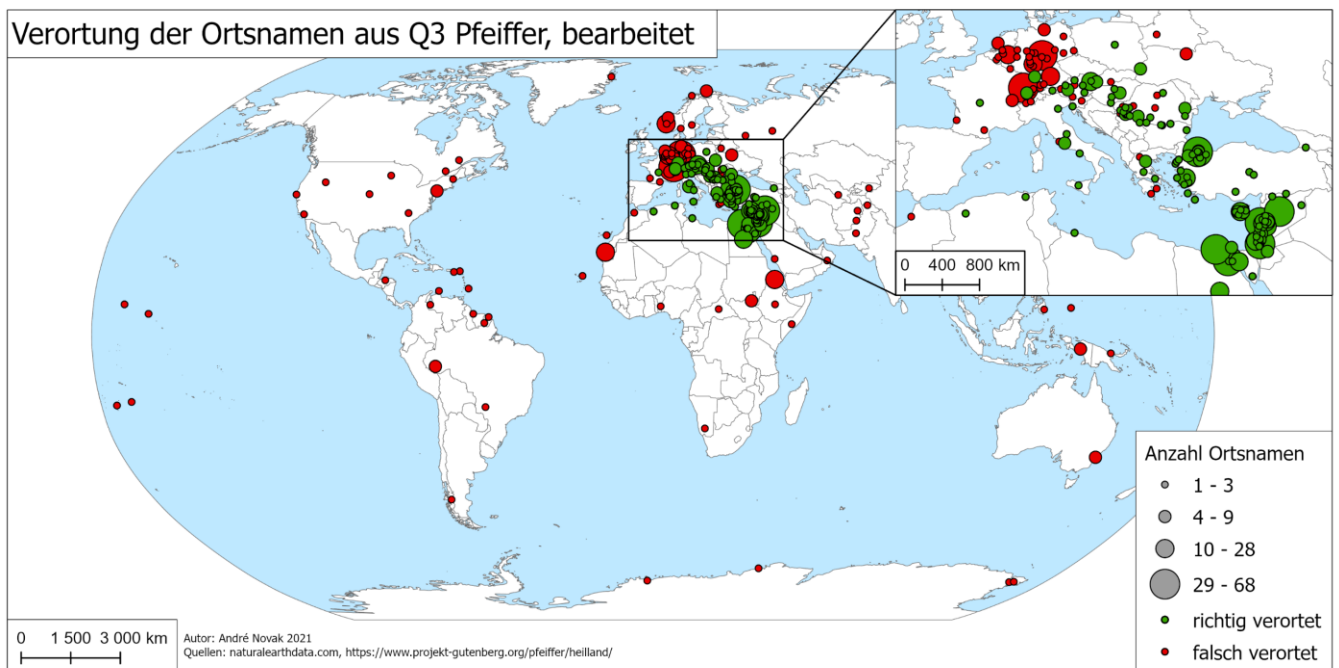


Abbildung 22: Q3_Pfeiffer, richtig und falsch verortete Ortsnamen

Auch wenn schon das unbearbeitete Ergebnis ein interpretierbares Bild lieferte, sorgt erst die manuelle Nachbearbeitung dafür, dass das Bild klarer und lesbarer wurde. Der Weg Ida Pfeiffers ist nun Schritt für Schritt nachvollziehbar.

7.6. Zukünftige Optimierungsmöglichkeiten

Auch wenn die oben angeführten Beispiele schon zu visuell interpretierbaren Ergebnissen geführt haben, sollte dieser Ansatz zukünftig weiterentwickelt werden, um zu noch besseren Ergebnissen zu kommen.

Einerseits sollte bei größeren Projekten eine eigene Pipeline (zum Beispiel in spaCy) trainiert werden, um schon bei der Named Entity Recognition zu einem besseren Ergebnis zu kommen. Dadurch ist es nämlich möglich, text- und sprachspezifische Eigenheiten in den Algorithmus miteinfließen zu lassen. Außerdem bieten die bereits trainierten Algorithmen nur eine eingeschränkte Anzahl an Entitäten. Vor allem im Deutschen ist diese stark limitiert. Bei einer selbsttrainierten oder erweiterten Pipeline können jedoch alle benötigten Entitäten hinzugefügt werden. Dies können zum Beispiel Zeitangaben, welche im geschichtswissenschaftlichen Kontext besonders relevant sind, oder Ländernamen, im Gegensatz zu Ortsnamen, sein.

Auch wenn Ortsnamen richtig getaggt wurden, muss das Ergebnis nicht immer mit dem beschriebenen Ort übereinstimmen. Ungenaue Ortsbeschreibungen wie „nahe Wiens“ oder „südlich von Wien“ lassen sich nicht genauer bestimmen. Aber auch genau definierte Ortsbeschreibungen, welche Mittels Distanzangaben genauer definiert werden (z.B.: 5km südlich von Wien), könnten als solche erfasst werden. Dabei geht es nicht darum den erwähnten Ort zu erfassen, sondern dabei auch eine Entfernungsrechnung miteinzubeziehen.

Alle getaggtten räumlichen Entitäten werden mittels Punktfeature markiert. Dies ist für Ortsnamen oder Berge zutreffend, jedoch für linienhafte oder polygonale Features nur einschränkend zutreffend. Länder wurden durch Punktfeatures im Zentroid der Fläche markiert. Diese nun zu flächenhaften Features zu geocoden ist ebenso wichtig wie auch die Geokodierung linienhafter Features wie Flüsse. Dies wäre unter anderem bei der Quelle „Q3_Pfeiffer“ relevant, da sich die Autorin hier lange auf der Donau flussabwärts fortbewegt.

Bei geschichtswissenschaftlichen Quellen sind Zeitangaben von besonderer Relevanz. Diese ermöglichen es, das Beschriebene zeitlich genauer einordnen zu können. Daher ist es relevant, Geschehnisse und Ortsangaben mit der zeitlichen Komponente verknüpfen zu können. Vor allem wenn mehrere unterschiedliche Quellen genutzt werden, kann eine zeitliche Einordnung dieser wichtig sein. So könnten zum Beispiel verschiedene Briefe von Soldaten an der Front herangezogen werden, um einen räumlichen und zeitlichen Verlauf dieser feststellen zu

können. Auch die Quelle „Q3_Pfeiffer“ bietet Zeitangaben, da das Werk im Tagebuchstil verfasst wurde. Durch eine zeitliche Einordnung der einzelnen Stationen der Reise könnte diese so nicht nur räumlich, sondern auch zeitlich oder raum-zeitlich visualisiert werden.

Die Ergebnisse des Geoparsings können natürlich nicht nur für eine Visualisierung herangezogen werden. Auch Analysen sind mit dem erstellten Datensatz möglich. So ergäben Punktmusteranalysen wie Hot-Spot-Analysen oder Kerndichteschätzer weiteren Aufschluss über wichtige Orte des Werkes.

Des Weiteren gibt es nicht nur deutschsprachige geschichtswissenschaftliche Quellen. Auch anderssprachige Quellen können durch Natural Language Processing in Geoparsing-Workflows miteinbezogen werden. So bietet zum Beispiel das „The Classical Language Toolkit“ (vgl. JOHNSON 2021) die Möglichkeit, mit alten Sprachen wie zum Beispiel Mittelhochdeutsch, Altgriechisch, Latein oder Altnordisch zu arbeiten. Auch eine Arbeit mit Gazetteers, welche besonders auf historische Namen fokussiert sind, ist möglich. Beispiele hierfür sind: „World Historical Gazetteer“ (vgl. World Historical Gazetteer o.J.) oder „pleiades“ (vgl. BAGNALL o.J.). GeoNames bietet zwar sehr viele unterschiedliche Ortsnamen für moderne Orte, sollten jedoch nicht mehr existierende Orte oder in GeoNames nicht vorhandene Ortsnamen Ziel der Verortung sein, sind diese Gazetteers eine gute Grundlage für die weitere Arbeit.

8. Diskussion

Auch wenn das Ergebnis des Geoparsings nicht perfekt ist und einige Fehlverortungen Teil dieses waren, zeigt es dennoch das Potential eines solchen Ansatzes auf. Wenn zukünftig ein Ausgleich zwischen den Eigenschaften geschichtswissenschaftlicher Daten und Quellen und den Anforderungen Geographischer Informationssysteme (siehe Tabelle 1) geschaffen werden kann, eröffnet sich ein ganz neues Feld für beide Fachbereiche.

Geschichtswissenschaftliche Quellen lassen sich mittels Geoparsing automatisch und ohne viel Aufwand als Karte darstellen. Dadurch ist es möglich, einerseits einen Überblick über den räumlichen Horizont einer Quelle zu erhalten. Ich schreibe hier absichtlich nicht, dass es sich hierbei um eine Verortung des Inhaltes der Quelle und bei Reiseberichten oder Tagebüchern somit des*der Autor*in handelt, da in einem Text ja auch Orte vorkommen, welche unabhängig des Inhaltes zu sehen sind. Johanna DRUCKER (2012) verweist hierbei auf einen Faktor, welcher auch bei der Betrachtung und visuellen Interpretation des Ergebnisses der Verortung immer mit zu beachten ist:

„But if the premises on which quantitative information might be abstracted from texts or corpora raise one set of issues, the use of graphical techniques from social and natural sciences raise others. Graphs and charts reify statistical information. They give it a look of certainty. Only a naive viewer, unskilled and untrained in matters of statistics or critical thought, would accept an information visualization at face value.” (DRUCKER 2012:89)

Ein*e Autor*in könnte zum Beispiel auf eine Erinnerung seiner*ihrer Kindheit Bezug nehmen, welche an einem anderen Ort stattfand als der eigentliche Inhalt der Quelle. Diesem muss sich der*die Interpret*in immer bewusst sein, damit es nicht zu Fehlschlüssen kommt. Außerdem sind Quellen der Geschichtswissenschaft als solche und dadurch auch die Ergebnisse einer Arbeit mit diesen immer kritisch zu hinterfragen, da Quellen immer auch missinterpretiert werden können. (vgl. BUDDE 2010: 68) Dies kann ohne, aber auch mit Vorsatz geschehen. Daher ist es ebenso wichtig, dass die Geschichtswissenschaftler*innen die Funktionsweise der angewendeten Methode auch verstehen und diese nicht nur als „black box“ betrachten. Und auch die Eigenschaften des jeweiligen Quellentyps sowie die Intention des*der Autor*in sind immer zu beachten.

Trotzdem zeigt die Verortung der Quelle „Q3_Pfeiffer“ die Vorteile einer solchen. Auf der Karte ist deutlich der Verlauf der Reise von Ida Pfeiffer in das Heilige Land zu sehen. Auch Personen, welche keinen Blick in das Werk geworfen haben, können somit einfach die Stationen der Reise der Autorin nachvollziehen. Auch wenn es beim Geocoding immer wieder zu Fehlverortungen kommt, lässt sich trotzdem eine fundierte Aussage über die Quelle treffen. Eine manuelle Nachbearbeitung verbessert das Ergebnis stark. Daher stellen LAFRENIERE und GILLILAND (2015) fest: „*The georeferencing of narrative sources is a semi-automated procedure.*“ (LAFRENIERE und GILLILAND 2015: 237)

Die Arbeit mit OCR-transkribierten Texten zeigt, dass auch automatisch erfasste Texte sich gut zur Verortung mittels Geoparsing eignen, auch wenn dadurch einige Fehler in die Texte eingebracht werden. Ein weiterer Vorteil einer Nutzung solcher Quellen ist ein niedriger Arbeitsaufwand in der Beschaffung der Grundlagen für die Verortung. Sobald Quellen also digital vorliegen, kann mit diesen auch weitergearbeitet werden. Geschichtswissenschaftliche Quellen weisen in der Arbeit mit Geoparsing besonders zu beachtende Eigenheiten auf. Im Gegensatz zu Ansätzen, welche die Verortung von Social-Media-Nachrichten verfolgen, sind die Texte viel länger. Außerdem bedeutet das Alter der Texte je nach Text auch Unterschiede in Grammatik, Rechtschreibung, Satzbau und Ausdruck. Daher kann eine Arbeit mit vortrainierten Natural-Language-Processing-Pipelines immer auch zu einer Beeinträchtigung des Ergebnisses führen.

Trotzdem ergibt sich durch eine solche Arbeit auch die Möglichkeit, mit größeren Quellenbeständen effizient arbeiten zu können. So wäre es zum Beispiel möglich, verschiedene Briefe von der Front aus einem Weltkrieg zu verorten und die Ergebnisse dann in einem weiteren Schritt zu analysieren. Dabei könnten unter anderem folgende Fragen gestellt werden: Von wo wurden Briefe in die Heimat geschickt? Welche Orte waren Inhalte dieser Briefe und welche nicht? In einer solchen Arbeit muss es dabei dann auch von Relevanz sein, die in Kapitel 4 erwähnten Anforderungen bei der Arbeit im Bereich Qualitative GIS vollständig umzusetzen. Neben einer Verwendung einer breiteren Quellenbasis, welche dann auch mit unterschiedlichen Methoden analysiert und auf verschiedene Arten und Weisen präsentiert werden können, sollte diese als iterativer Prozess mit unterschiedlichen Fokussen durchgeführt werden. Dadurch ist es möglich, sich ganzheitlicher mit einem Thema zu befassen.

Eine Verortung geschichtswissenschaftlicher Quellen mittels Geoparsing stellt eine interdisziplinäre Arbeit über verschiedene Fachbereiche hinweg dar. Einerseits bringt diese Arbeit

Vorteile für Geschichtswissenschaftler*innen, da ihnen durch diesen Ansatz ganz neue Werkzeuge und Methoden zur Verfügung stehen, um mit geschichtswissenschaftlichen Quellen zu arbeiten. Dadurch ist es für sie möglich, neue Antworten auf bestehende Fragen zu finden, aber auch ganz neue Fragen zu stellen und neue Perspektiven einnehmen zu können. Demgemäß können nun neue Aussagen über die Vergangenheit getätigt werden. Der räumliche Fokus dieser ist nun schneller und einfacher greifbar als mittels herkömmlicher Umsetzung solcher Arbeiten.

Der vorgestellte Ansatz ermöglicht es, auch mit unstrukturierten Texten zu arbeiten und dadurch neue Datentypen in Geographische Informationssysteme einzubinden. Zusätzlich wird dadurch eine Arbeit mit qualitativen Daten in GIS gestärkt. Somit wird zumindest der erste der beiden von PRESTON und WILSON (2014) benannten Arbeitsbereiche der Qualitative GIS umgesetzt: Qualitative Daten, in diesem Fall textliche Quellen der Geschichtswissenschaft, werden in Geographische Informationssysteme integriert. Außerdem ermöglicht es dieser Ansatz, dem schon von BOONSTRA (2013) benannten Problem der mangelnden Verfügbarkeit an räumlichen historischen Daten, welche in GIS nutzbar sind, ein bisschen entgegenzutreten. Eine solches fachübergreifendes Zusammenwirken ist ganz im Sinne der Digital Humanities.

In vorgestellten Geoparsing-Workflow wurden Methoden und Tools aus dem Bereich der Geoinformatik in die Geschichtswissenschaft übertragen, um hier einen Mehrwert zu erreichen. Dabei geht es jedoch nicht darum, einem anderen Wissenschaftsbereich die eigene Arbeitsweise vorzuschreiben, sondern darum, Synergien zu nutzen und so zu Vorteilen und Erkenntnissen in beiden Bereichen zu kommen. Einer dadurch stattfindenden Weiterverbreitung des *spatial turns* in den Geistes- und Kulturwissenschaften kann somit außerdem Rechnung getragen werden.

Dabei wurde im Sinne Johanna DRUCKERS (2012) versucht, durch die Arbeit mit unterschiedlichen Quellentypen die Vorteile beider Fachrichtungen auszunützen und so eine bereits bestehende Methode an die Bedürfnisse der Fachwissenschaft anzupassen und nicht nur die Technologie zu übertragen.

Eine Arbeit mit Texten (als geschichtswissenschaftliche Quellen) in GIS kann so zu einer weitreichenderen Verbreitung von räumlichen Methoden in der Geschichtswissenschaft und den Digital Humanities führen. Es sollte das Ziel jedes*r Geoinformatiker*in sein, die eigene

Fachrichtung mit anderen Wissenschaftsbereichen in Kontakt zu bringen und so zur Nutzung von Synergien und zur Entwicklung neuer Arbeitsweisen beizutragen. Wichtig dabei ist anzumerken, dass der vorgestellte Workflow dank Entwicklung eines Python-Codes die Verortung automatisch vornimmt. Dadurch wäre es in einem weiteren Schritt möglich, diesen zum Beispiel als Plugin für gängige GIS-Softwarepakete umzusetzen und dadurch auch Nichtprogrammierer*innen die Möglichkeit zu geben, diesen einfach nutzen zu können. Dies ist besonders wichtig, da viele Geschichtswissenschaftler*innen wenig oder keine Kenntnisse im Programmieren besitzen und dadurch ihnen sonst die Möglichkeit verloren geht, ebenso eine Verortung geschichtswissenschaftlicher Quellen vorzunehmen.

Im oben gewählten Ansatz wurden die erzeugten Daten vor allem zur Visualisierung und visuellen Interpretation herangezogen. Eine wichtige Funktion Geographischer Informationssysteme ist aber auch die Analyse der erzeugten Daten. Diese soll aus Zeit- und Platzgründen in dieser Arbeit nicht stattfinden, muss jedoch in einem weiteren Schritt Ziel solcher Geoparsing-Arbeiten sein und ist daher auch im Workflow (siehe Abbildung 4) angeführt. Denn erst durch eine solche kann das ganze Potential der Verortung geschichtswissenschaftlicher Quellen mittels Geoparsing ausgeschöpft werden.

9. Fazit

Historical GIS liegt im Zentrum der Entwicklungen *spatial turn* und Digital Humanities im Bereich der Geschichtswissenschaft. Dabei ergeben sich für Geschichtswissenschaftler*innen ganz neue Möglichkeiten bisherige Themenbereiche zu bearbeiten, aber auch ganz neue zu entdecken. Um nun auch aus geoinformatischer Sicht einen wirklichen Mehrwert für diese zu ermöglichen, sollten geschichtswissenschaftliche Quellen Ziel einer räumlichen Arbeit sein. Texte, welche einen Großteil in der Geschichtswissenschaft verwendeten Quellen darstellen, zu verorten, kann daher zu einer weiteren Verbreitung und Akzeptanz von Historical-GIS-Methoden in der Geschichtswissenschaft führen. Es war deshalb mein Ziel, eine automatische Verortung solcher Quellen zu ermöglichen und dabei besonders auf die Eigenheiten textlicher Quellen der Geschichtswissenschaft einzugehen. Im Internet sind dabei unglaublich viele Quellen vorhanden, daher wurden auch nur solche genutzt.

Um nun den Ansätzen des Text Minings zu folgen, entwickelte ich einen Geoparsing-Workflow für die Verortung unterschiedlicher Quellen via Geoparsing. Dadurch wurde es möglich, ohne die Quelle jemals gelesen zu haben, den räumlichen Horizont dieser auf einer Karte präsentiert zu sehen. Dies zeigt schon das Potential jenes Ansatzes. Denn durch eine räumliche Präsentation ergibt sich ein ganz neuer Blick auf Texte. In einem ersten Schritt kann daher eine visuelle Analyse schon zu neuen Erkenntnissen führen. Hier sollte jedoch nicht gestoppt werden, sondern in einem weiteren Schritt auch eine räumliche Analyse der erstellten Daten durchgeführt werden, um das ganze Potential der Geographischen Informationssysteme auch der Geschichtswissenschaft zugänglich zu machen. Auch wenn der entwickelte Workflow noch einiges an Weiterentwicklungspotential für zukünftige Arbeiten aufweist, lassen sich die visualisierten Ergebnisse einer Arbeit mit diesem trotzdem aussagekräftig analysieren.

10. Literaturverzeichnis

- AGGARWAL, C. C. (2015): Data Mining. The Textbook.- Cham, Heidelberg, New York, Dordrecht, London.
- ALEX, B., GROVER, C., TOBIN, R. und OBERLANDER, J. (2019): Geoparsing historical and contemporary literary text set in the City of Edinburgh. - In: Language Resources and Evaluation 4, 651–675. - doi: 10.1007/s10579-019-09443-x.
- ARPAGIAN, J. und AITKEN, S. (2018): Qualitative GIS. - <https://www.oxfordbibliographies.com/view/document/obo-9780199874002/obo-9780199874002-0192.xml> (11.03.2021).
- ASSMANN, A. (2009): Geschichte findet Stadt. - In: CSÁKY und LEITGEB (Hrsg.): Kommunikation - Gedächtnis - Raum. - Bielefeld, 13–27.
- AYERS, E. L. (2007): The Valley of The Shadow. - <https://valley.lib.virginia.edu/> (25.05.2021).
- AYERS, E. L. (2010): Turning toward Place, Space, and Time. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, 1–13.
- BACHMANN-MEDICK, D. (2016): Cultural turns. New orientations in the study of culture.- Berlin, Boston.
- BAGHERI, N. (2014): What qualitative GIS maps tell and don't tell. Insights from mapping women in Tehran's public spaces. - In: Journal of Cultural Geography Vol. 31, No. 2, 166–178. - doi: 10.1080/08873631.2014.906848.
- BAGNALL, R. (o.J.): Pleiades. - <https://pleiades.stoa.org/> (10.08.2021).
- BARTELME, N. (2005): Geoinformatik. Modelle , Strukturen , Funktionen.- Berlin, Heidelberg.
- BITTNER, C. und MICHEL, B. (2018): Qualitative Geographische Informationssysteme. - In: WINTZER (Hrsg.): Sozialraum erforschen. - Berlin, Heidelberg, 151–166.
- BODENHAMER, D. J. (2010): The Potential of Spatial Humanities. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, 14–30.
- BODENHAMER, D. J. (2013): Beyond GIS. Geospatial Technologies and the Future of History. - In: VON LÜNEN und TRAVIS (Hrsg.): History and GIS. - Dordrecht, 1–13.
- BODENHAMER, D. J., CORRIGAN, J. und HARRIS, T. M. (2010): Introduction. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, vii–xv.
- BOL, P. K. (2011): What do humanists want? What do humanists need? What might humanists get? - In: DEAR; KETCHUM; LURIA und RICHARDSON (Hrsg.): Geohumanities. - London, New York, 296–308.
- BOONSTRA, O. (2013): The Dawn of a Golden Age? Historical GIS and the History of Choropleth Mapping in the Netherlands. - In: VON LÜNEN und TRAVIS (Hrsg.): History and GIS. - Dordrecht, 27–38.

- BUDDE, G. (2010): Quellen, Quellen, Quellen... - In: BUDDE; FREIST und GÜNTHER-ARNDT (Hrsg.): Geschichte. - Berlin, 52–69.
- BUDDE, G. und FREIST, D. (2010): Verfahren, Methoden, Praktiken. - In: BUDDE; FREIST und GÜNTHER-ARNDT (Hrsg.): Geschichte. - Berlin, 158–177.
- CAMPBELL, M. B. (2002): Travel writing and its theory. - In: HULME und YOUNGS (Hrsg.): The Cambridge companion to travel writing. - Cambridge, 261–278.
- Cartography Associates (Hrsg.) (2021): David Rumsey Map Collection. - www.davidrumsey.com (20.05.2021).
- CHERIET, M., KHARMA, N., LIU, C.-L. und SUEN, C. Y. (2007): Character recognition systems. A guide for students and practioners.- Hoboken, N.J.
- CLEVE, J. und LÄMMEL, U. (2020): Data Mining. - München, Wien.
- COPE, M. und ELWOOD, S. (2009): Conclusion. For Qualitative GIS. - In: COPE und ELWOOD (Hrsg.): Qualitative GIS. - Los Angeles, London, 171–177.
- CORRIGAN, J. (2010): Qualitative GIS and Emergent Semantics. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, 76–88.
- CREW, D. F. (1989): Alltagsgeschichte: A New Social History "From Below"? - In: Central European History 22, 394–407.
- CSÁKY, M. und LEITGEB, C. (2009): Kommunikation – Gedächtnis – Raum: Orientierungen im spatial turn der Kulturwissenschaften. - In: CSÁKY und LEITGEB (Hrsg.): Kommunikation - Gedächtnis - Raum. - Bielefeld, 7–10.
- DALE, R. (2010): Classical Approaches to Natural Language Processing Robert Dale. - In: INDURKHYA und DAMERAU (Hrsg.): Handbook of Natural Language Processing. - New York, 3–7.
- Denis Carriere (Hrsg.) (2013): Geocoder: Simple, Consistent. - <https://geocoder.readthedocs.io/> (22.07.2021).
- DÖRING, J. und THIELMANN, T. (2009): Einleitung: Was lesen wir im Raume? Der Spatial Turn und das geheime Wissen der Geographen. - In: DÖRING und THIELMANN (Hrsg.): Spatial Turn. - Bielefeld, 7–45.
- DROYSEN, J. G. (1925): Grundriss der Historik. - Halle/Saale.
- DRUCKER, J. (2012): Humanistic Theory and Digital Scholarship. - In: GOLD (Hrsg.): Debates in the digital humanities. - Minneapolis, MN, 85–95.

DRUCKER, J. (2021): The digital humanities coursebook. An introduction to digital methods for research and scholarship.- Abingdon, Oxon, New York, NY.

ELWOOD, S. und COPE, M. (2009): Introduction. Qualitative GIS: Forging mixed methods through representations, analytical innovations, and conceptual engagements. - In: COPE und ELWOOD (Hrsg.): Qualitative GIS. - Los Angeles, London, 1–12.

Explosion (Hrsg.) (2021): spaCy. - <https://spacy.io/> (21.07.2021).

Fakultät für Geowissenschaften, Geographie und Astronomie (Hrsg.) (2021): Ida Pfeiffer Professur. - <https://fgga.univie.ac.at/ueber-uns/ida-pfeiffer-professur/> (26.07.2021).

FELDMAN, R. und SANGER, J. (2007): The text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data.- Cambridge, New York.

geonames.org (Hrsg.) (o.J.): GeoNames. - <https://www.geonames.org/> (22.07.2021).

GIMMLER, L. und KALISKY, S. (2020): Introducing ArcGIS Notebooks in ArcGIS Pro. - <https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/introducing-arcgis-notebooks-in-arcgis-pro/> (13.08.2021).

GREGORY, I. (2010): Exploiting Time and Space: A Challenge for GIS in the Digital Humanities. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, 58–75.

GREGORY, I. und COOPER, D. (2009): Thomas Gray, Samuel Taylor Coleridge and Geographical Information Systems: A Literary GIS of Two Lake District Tours. - In: International Journal of Humanities and Arts Computing 3, 61–84. - doi: 10.3366/E1753854810000534.

GREGORY, I., DEBATS, D. A. und LAFRENIERE, D. (2018): Introduction. Spatial history, history, and GIS. - In: GREGORY; DEBATS und LAFRENIERE (Hrsg.): The Routledge companion to Spatial History. - London, New York, 1-6.

GREGORY, I., DONALDSON, C., MURRIETA-FLORES, P. und RAYSON, P. (2015): Geoparsing, GIS, and Textual Analysis. Current Developments in Spatial Humanities Research. - In: International Journal of Humanities and Arts Computing 1, 1–14. - doi: 10.3366/ijhac.2015.0135.

GREGORY, I. und ELL, P. S. (2007): Historical GIS. Technologies, methodologies, and scholarship.- Cambridge.

GREGORY, I. und GEDDES, A. (2014a): Conclusions. From Historical GIS to Spatial Humanities: Challenges and Opportunities. - In: GREGORY und GEDDES (Hrsg.): Toward spatial humanities. - Bloomington, 172–185.

GREGORY, I. und GEDDES, A. (2014b): Introduction. From Historical GIS to Spatial Humanities: Deepening Scholarship and Broadening Technology. - In: GREGORY und GEDDES (Hrsg.): Toward spatial humanities. - Bloomington, ix–xix.

GREGORY, I. und HARDIE, A. (2011): Visual GISTing. Bringing together corpus linguistics and Geographical Information Systems. - In: Literary and Linguistic Computing 26 (3), 297–314. - doi: 10.1093/llc/fqr022.

GREGORY, I. und HEALEY, R. G. (2007): Historical GIS. Structuring, mapping and analysing geographies of the past. - In: Progress in Human Geography 5, 638–653. - doi: 10.1177/0309132507081495.

GRITTA, M., PILEHVAR, M. T., LIMSOPATHAM, N. und COLLIER, N. (2018): What's missing in geographical parsing? - In: Language resources and evaluation 52 (2), 603–623. - doi: 10.1007/s10579-017-9385-8.

HARRIS, R. (2008): Qualitative Analysis. - In: Encyclopedia of geographic information science, 355–357.

HARRIS, T. M., BERGERON, S. und ROUSE, L. J. (2011): Humanities GIS. Place, spatial storytelling, and immersive visualization in the humanities. - In: DEAR; KETCHUM; LURIA und RICHARDSON (Hrsg.): Geohumanities. - London, New York, 226–240.

HARRIS, T. M., CORRIGAN, J. und BODENHAMER, D. J. (2010): Challenges for the Spatial Humanities: Toward a Research Agenda. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, 167–176.

HILL, L. L. (2008a): Gazetteers. - In: KEMP (Hrsg.): Encyclopedia of Geographic Information Science. - Los Angeles, 159–161.

HILL, L. L. (2008b): Geoparsing. - In: KEMP (Hrsg.): Encyclopedia of Geographic Information Science. - Los Angeles, 200–201.

HOFMANN, M. und CHISHOLM, A. (2015): Text Mining and Visualization. Case Studies Using Open-Source Tools.- Boca Raton.

HU, Y. (2017): Geospatial Semantics. - In: HUANG; COVA und TSOU (Hrsg.): Comprehensive Geographic Information Systems. - Oxford.

HU, Y. (2018): Geo-text data and data-driven geospatial semantics. - In: Geography Compass 12 (11). - doi: 10.1111/gec3.12404.

HUG, T. und POSCHESCHNIK, G. (2015): Empirisch forschen. Die Planung und Umsetzung von Projekten im Studium.- Konstanz, München.

HUGHES, L., CONSTANTOPOULOS, P. und DALLAS, C. (2016): Digital Methods in the Humanities: Understanding and Describing their Use across the Disciplines. - In: UNSWORTH; SIEMENS und SCHREIBMAN (Hrsg.): A new companion to digital humanities. - Chichester, West Sussex, UK, Malden, MA, USA, 150–170.

Institut für Kartografie und Geoinformation, ETH Zürich (Hrsg.) (2019): Ein Literarischer Atlas Europas. - <http://www.literaturatlas.eu/index.html> (10.08.2021).

JANNIDIS, F., KOHLE, H. und REHBEIN, M. (2017): Digital Humanities. Eine Einführung.- Stuttgart.

JOCKERS, M. L. und UNDERWOOD, T. (2016): Text-Mining the Humanities. - In: UNSWORTH; SIEMENS und SCHREIBMAN (Hrsg.): A new companion to digital humanities. - Chichester, West Sussex, UK, Malden, MA, USA, 291–306.

JOHNSON, K. P. (2021): The Classical Language Toolkit. - <http://cltk.org/> (10.08.2021).

JORDAN, S. (2018): Theorien und Methoden der Geschichtswissenschaft. - Paderborn, Stuttgart.

KINGSTON, R. (2015): Mind Over Matter? - In: Cultural and Social History 1, 111–121. - doi: 10.2752/147800410X477368.

KLEIN, E., DÄUMER, M. und HÜTIG, A. (2016): Gute Ergebnisse aus „schlechten“ Textvorlagen. Einflüsse der OCR-Qualität auf kulturwissenschaftsgeschichtliche Forschungsergebnisse am Beispiel von Drucken des 18. Jahrhunderts. - In: Information. Wissenschaft & Praxis 67 (5-6), 331–338.

KUHN, T. S. und HACKING, I. (2012): The structure of scientific revolutions. - Chicago, London.

KWAN, M.-P. und KNIGGE, L. (2006): Doing Qualitative Research Using GIS: An Oxymoronic Endeavor? - In: Environment and Planning A 38, 1999–2002.

KWARTLER, T. (2017): Text Mining in Practice with R. - Hoboken, NJ.

LAFRENIERE, D. und GILLILAND, J. (2015): “All the World's a Stage”. A GIS Framework for Recreating Personal Time-Space from Qualitative and Quantitative Sources. - In: Transactions in GIS 2, 225–246. - doi: 10.1111/tgis.12089.

LANGE, N. (2013): Geoinformatik. In Theorie und Praxis.- Berlin, Heidelberg.

LEIDNER, J. L. und LIEBERMAN, M. D. (2011): Detecting Geographical References in the Form of Place Names and Associated Spatial Natural Language. - In: SIGSPATIAL Special 3(2), 5–11.

LILLEY, K. und PORTER, C. (2016): GIS, Spatial Technologies and Digital Mapping. - In: GUNN und FAIRE (Hrsg.): Research methods for history. - Edinburgh, 125–146.

LIPPUNER, R. und LOSSAU, J. (2004): In der Raumfalle. Eine Kritik des spatial turn in den Sozialwissenschaften. - In: MEIN und RIEGER-LADICH (Hrsg.): Soziale Räume und kulturelle Praktiken. - Bielefeld, 47–63. - doi: 10.14361/9783839402160-003.

LOSSAU, J. (2009): Räume von Bedeutung. Spatial turn, cultural turn und Kulturgeographie. - In: CSÁKY und LEITGEB (Hrsg.): Kommunikation - Gedächtnis - Raum. - Bielefeld, 29–43.

LUDDEN, D. (2002): A Brief History of Subalternity. - In: LUDDEN (Hrsg.): Reading subaltern studies. - London, 1–42.

MCCARTHY, W. (2016): Becoming Interdisciplinary. - In: UNSWORTH; SIEMENS und SCHREIBMAN (Hrsg.): A new companion to digital humanities. - Chichester, West Sussex, UK, Malden, MA, USA, 69–83.

MCHAFFIE, P., HWANG, S. und FOLLETT, C. (2018): GIS. An Introduction to Mapping Technologies.- Milton.

MIDDLETON, S. E., KORDOPATIS-ZILOS, G., PAPADOPOULOS, S. und KOMPATSIARIS, Y. (2018): Location Extraction from Social Media. - In: ACM Transactions on Information Systems 4, 1–27. - doi: 10.1145/3202662.

MONCLA, L., GAIO, M., JOLIVEAU, T. und LE LAY, Y.-F. (2017): Automated Geoparsing of Paris Street Names in 19th Century Novels. - In: 1st ACM SIGSPATIAL Workshop on Geospatial Humanities.

MONTEIRO, B. R., DAVIS, C. A. und FONSECA, F. (2016): A survey on the geographic scope of textual documents. - In: Computers & Geosciences 96, 23–34. - doi: 10.1016/j.cageo.2016.07.017.

MOSTERN, R. und JOHNSON, I. (2008): From named place to naming event. Creating gazetteers for history. - In: International Journal of Geographical Information Science 22 (10), 1091–1108. - doi: 10.1080/13658810701851438.

Museumsstiftung Post und Telekommunikation (Hrsg.) (o.J.): Otto Weiß an seine Ehefrau am 31.5.1940 (3.2002.1202). Brief (Transkript).- https://www.briefsammlung.de/feldpost-zweiter-weltkrieg/brief.html?action=detail&what=letter&id=43&le_keyword=Britische%20Armee (21.07.2021).

NECKEL, S. (2009): Felder, Relationen, Ortseffekte: Sozialer und physischer Raum. - In: CSÁKY und LEITGEB (Hrsg.): Kommunikation - Gedächtnis - Raum. - Bielefeld, 45–55.

Neue Freie Presse (24.09.1886): Aus Tegetthoff's Mappe. – Wien. - <https://anno.onb.ac.at/cgi-content/anno?aid=nfp&datum=18860924&seite=1&zoom=33> (21.07.2021)

o.V. (Hrsg.) (24.09.1886): Aus Tegetthoff's Mappe. - In: Neue Freie Presse, 1–2.

o.V. (Hrsg.) (2016): Pfeiffer Ida. - In: KOROTIN (Hrsg.): biografiA. - Wien, Köln, Weimar, 2522–2527.

Österreichische Akademie der Wissenschaften (Hrsg.) (o.J.): Austrian Center for Digital Humanities and Cultural Heritage. - <https://www.oeaw.ac.at/acdh/> (14.05.2021).

Österreichische Nationalbibliothek (Hrsg.) (2011): Neue Freie Presse. - https://anno.onb.ac.at/info/nfp_info.htm (27.07.2021).

Österreichische Nationalbibliothek (Hrsg.) (2021): ANNO Historische Zeitungen und Zeitschriften. - <https://anno.onb.ac.at/> (27.07.2021).

O'SULLIVAN, D. UND UNWIN, D. J. (2010): Geographic Information Analysis. - Hoboken, New Jersey.

- OWENS, J. B. und WACHOWICZ, M. (2013): The Role of Knowledge Spaces in Geographically-Oriented History. - In: VON LÜNEN und TRAVIS (Hrsg.): History and GIS. - Dordrecht, 127–144.
- PALMER, M. H. (2013): (In)Digitizing Cáuigú Historical Geographies. Technoscience as a Postcolonial Discourse. - In: VON LÜNEN und TRAVIS (Hrsg.): History and GIS. - Dordrecht, 39–58.
- PAVLOVSKAYA, M. (2009): Non-quantitative GIS. - In: COPE und ELWOOD (Hrsg.): Qualitative GIS. - Los Angeles, London, 13–38.
- PFEIFFER, I. (1886): Reise einer Wienerin in das Heilige Land. - <https://www.projekt-gutenberg.org/pfeiffer/heilland/> (26.07.2021).
- PRESNER, T. und SHEPARD, D. (2016): Mapping the Geospatial Turn. - In: UNSWORTH; SIEMENS und SCHREIBMAN (Hrsg.): A new companion to digital humanities. - Chichester, West Sussex, UK, Malden, MA, USA, 201–212.
- PRESTON, B. und WILSON, M. W. (2014): Practicing GIS as Mixed Method. Affordances and Limitations in an Urban Gardening Study. - In: Annals of the Association of American Geographers 104 (3), 510–529. - doi: 10.1080/00045608.2014.892325.
- Projekt Gutenberg-DE (Hrsg.) (o.J.): Information. - <https://www.projekt-gutenberg.org/info/texte/info.html> (16.08.2021).
- PURVES, R. S., CLOUGH, P., JONES, C. B., HALL, M. H. und MURDOCK, V. (2018): Geographic information retrieval. Progress and challenges in spatial search of text.- Boston.
- Python Software Foundation (Hrsg.) (2021): Virtual Environments and Packages. - <https://docs.python.org/3/tutorial/venv.html> (13.08.2021).
- RICHARDSON, D. (2011): Geohistories. - In: DEAR; KETCHUM; LURIA und RICHARDSON (Hrsg.): Geohumanities. - London, New York, 209–214.
- ROWLINGSON, B. (2019): geonames: Interface to the "Geonames" Spatial Query Web Service. - <https://cran.r-project.org/web/packages/geonames/index.html> (13.08.2021).
- SCHÄFER, S., MUENCHOW, J. und HARNISCH, B. (2018): Qualitative Forschung und Geographische Informationssysteme. - In: WINTZER (Hrsg.): Sozialraum erforschen. - Berlin, Heidelberg, 167–180.
- SCHIRMER, D. und BLINKERT, B. (2009): Empirische Methoden der Sozialforschung. Grundlagen und Techniken.- Paderborn.
- SCHULZE, W. (1996): Ego-Dokumente: Annäherung an den Menschen in der Geschichte? Vorüberlegungen für die Tagung „Ego-Dokumente“. - In: SCHULZE (Hrsg.): Ego-Dokumente. - Berlin, 11–30.

- SCHUURMAN, N. (2009): Metadata as a site for imbuing GIS with qualitative information. - In: COPE und ELWOOD (Hrsg.): Qualitative GIS. - Los Angeles, London, 41–56.
- SMITH, A. (1915): Training: 'drill before breakfast'. - <https://www.nationalarchives.gov.uk/education/resources/letters-first-world-war-1915/training-drill-breakfast/> (21.07.2021).
- TSCHIGGERL, M., WALACH, T. und ZAHLMANN, S. (2019): Geschichtstheorie. - Wiesbaden.
- Universität Wien (Hrsg.) (o.J.): Forschungsschwerpunkt Digital Humanities. - <https://fsp-digital-humanities.univie.ac.at/> (14.05.2021).
- University of Portsmouth (Hrsg.) (o.J.): Old Maps Online. - <http://www.oldmapsonline.org/> (20.05.2021).
- University of Portsmouth et al. (Hrsg.) (2017): A Vision of Britain Through Time. - <https://www.vision-of-britain.org.uk/> (25.05.2021).
- WARWICK, C. (2016): Building Theories or Theories of Building? A Tension at the Heart of Digital Humanities. - In: UNSWORTH; SIEMENS und SCHREIBMAN (Hrsg.): A new companion to digital humanities. - Chichester, West Sussex, UK, Malden, MA, USA, 538–552.
- WIERLING, D. (1995): Geschichte. - In: FLICK; KARDORFF und KEUPP (Hrsg.): Handbuch qualitative Sozialforschung. - Weinheim, 47–52.
- Wikipedia (Hrsg.) (2021): Springfield (Ortsname). - [https://de.wikipedia.org/wiki/Springfield_\(Ortsname\)](https://de.wikipedia.org/wiki/Springfield_(Ortsname)) (14.06.2021).
- WILSON, M. W. (2009): Towards a geanology of qualitative GIS. - In: COPE und ELWOOD (Hrsg.): Qualitative GIS. - Los Angeles, London, 156–170.
- WITTEN, I. H., FRANK, E. und HALL, M. A. (2011): Data mining. Practical machine learning tools and techniques.- Amsterdam.
- World Historical Gazetteer (Hrsg.) (o.J.): World Historical Gazetteer. - <https://whgazetteer.org/> (10.08.2021).
- YUAN, M. (2010): Mapping Text. - In: BODENHAMER; CORRIGAN und HARRIS (Hrsg.): The spatial humanities. - Bloomington, 109–123.
- ŽIŽKA, J., DAŘENA, F. und SVOBODA, A. (2019): Text mining with machine learning. Principles and techniques.- Boca Raton.

11. Anhang

11.1. Anhang I Quellen:

11.1.1. Otto Weiß an seine Ehefrau am 31.5.1940 (3.2002.1202) - Transkript

Im Felde, 31. Mai 40.

Meine Lieben!

Schon fast 8 Tage liegen wir in der Nähe von Calais, d.h. nicht immer auf dem gleichen Platze, wir haben fast jeden Tag unsere Stellungen gewechselt, aber doch immer um Calais, das war mal eine sehr schöne Stadt, aber unsere Stukas haben auch hier ganze Arbeit geleistet.

Die Stadt brennt heute noch an vielen Stellen, obwohl sie fast schon 8 Tage in unserer Hand ist. Ebenso ist Dünkirchen schwer heimgesucht worden. Im Hafen von Calais sind Millionenwerte in unsere Hand gefallen, dieses Briefpapier ist auch davon. Es stammt aus England. Die Tommis haben fluchtartig Calais verlassen und wollten sich in England in Sicherheit bringen. Aber unterwegs hat sich ja ihr Schicksal erfüllt, davon sind nicht mehr viel nach England zurückgekehrt, weil unsere Flieger eine kleine Unterredung auf hoher See mit ihnen hatten. Ich habe zugesehen wie unsere Bomber 3 englische Schiffe bombardiert haben. Das war herrlich anzusehen. Augenblicklich haben wir 3 Ruhetage, da ist heute der zweite davon. Es ist vermutlich eine Ruhe vor dem Sturm. Welche Richtung wir dann einschlagen ist noch nicht bekannt

Vielleicht kommt mal plötzlich eine Ansichtskarte von England bei Euch an? Dort drüben würden wir bestimmt noch anders aufräumen als in Frankreich. – In dieser Gegend sind alle Einwohner zu Hause geblieben, wir liegen z.Zt. in einem kleinen Städtchen Guines (Fiennes) da sind viele Flüchtlinge von Calais und auch viele Belgier von der Gegend von Antwerpen. Die sind geflüchtet weil sie Angst hatten, aber wir waren mal wieder schneller wie die Flüchtlinge. Es sind meist Flamen wenn sie langsam sprechen verstehen wir fast jedes Wort Auch meine Schulkenntnisse in französisch kann ich jetzt gut gebrauchen, ich spiele manchmal den Dolmetscher. Ich verstehe noch ziemlich viel davon. Die Belgier sind fast alle sehr freundlich auch die franz. Bevölkerung bisher, nur in dieser Gegend sind sie nicht sehr entgegenkommend, aber sie werden sich noch daran gewöhnen müssen Die Engländer sind überall gehaßt. Das konnte man bei den Gefangenen zügen feststellen, die Engländer haben von den Franzosen manche Rippenstöße bekommen. – In dem Städtchen in dem wir liegen haben wir's uns sehr gemütlich gemacht. Es fehlt uns gar nichts wir leben wie die Fürsten. Schokolade in rauen Mengen Kaffeebohnen zentnerweise. Wein und Likör zum Schwimmen. Jeden Tag ein frisches Hemd u. Unterhose. u.s.w. Also mit anderen Worten. Es geht uns glänzend. Zigaretten haben wir auch tausendweise, in der Hauptsache englische. Die französischen sind uns zu stark. Wir sind sehr verwöhnt und anspruchsvoll. Eure zwei Päckchen mit Photo u Schlafanzug u.s.w. habe ich vorgestern erhalten. Da haben wir den richtigen Zeitpunkt erwischt. Wenn ichs früher mal mitgenommen hätte dann hätte ich es wieder nach Hause schicken können, So muß ich es eben immer mir herumschleppen. Ich hatte schon gehofft daß Ihr's nicht weggeschickt habt weil es so lange gedauert hat. Paketpost können wir leider keine zurückschicken. Ist ja auch ganz gut sonst würden wir nur den ganzen Tag Pakete zurechtmachen. –

Jetzt muß ich allmählich Schluß machen habe noch viel zu schreiben. Ein andermal mehr, natürlich kann ich nicht immer 6 Seiten schreiben, weil nicht alle Tage Ruhetage sind. Also Päckchen braucht Ihr keines zu schicken. Es geht uns besser wie Euch. Lebt alle wohl und seid recht herzlich begrüßt von

Eurem Otto

Bald wird der Krieg zu Ende sein. Es geht bestimmt keine 4 Jahre

(Quelle: Museumsstiftung Post und Telekommunikation o.J.)

11.1.2. Neue Freie Presse, 24.09.1886, Seiten 1-2

[Neue Freie Presse - 18860924 - Seite 1]

ßsdactlen, Administration L Druakercl:

Kolowrat-Ring, Fichtegasse Nr. 11.

ÜWMkirte Briefe werden nicht anrennoimaest WM&

Mannscripte nicht xuröekg, e»e 7 idet v,

AnkttiKH^nn^e'BoireAii:

Mfflit» Wollzeile Nr. 20 . InsertLonsprei» nach Tarife

luerte übernehmen im Auslande: Agenca Hat#®

S'Ari» und SrtfKftCl; Rudolf Messe in Revt»

lin.Franlisurt a. 9 f M Üliinchn.

W. Ktttrlcll» G, L. Daube & Cie. in Igrainltfeyg

S - M« Berlin, < öln, lieipxix und Purina

aaseustem & Vogler in Mmuburs'. Berlin«

*r»nlisuvt n. Bl» u. IKhscl; Annoneen-ISxpe»

Atlton „Invaiddcn<iank“,l>s'es«l«*»; Deutschen Zel»

«Bfigsbur., „In validendank“ in Rerlhi; Adolsleinejr
in llambnr&; Eduard Schlotte in Bärucusucns
Orall Fiisali & Cie. in Lsrrifh .- Kajehton L Frend*
Jer, Gener, *Annonc. *Bur. in U nv»cliHM. 33 jiti!>ikau
■ U Z.O«Iz; Fried. Patrick, Centml-Anroncen-Buroan
■ t Sl.Petersburg; E.E.QbHoe-ht in Mailimd,
alovenx and Rom; Aug. Siegle in Loadott*
Abonnement für Wien:
9m HaupWerlage, Wollzeile 20 : Ganzjähr. fl. S1.0GL
ÄOWtL fl* 1.80, Mit Zustellung ins Haast Viertel),
fl. 6.80, monatl. fl. 2.10,
Sfualfls Morgenblatt s kr.» Abendblatt 8 kn
Bst 7031.
Morgenblatt.
Abonnement ffir das Inland»
Kit tltg-l. einmal. Po'stversendung: Ganzj. fl.28, halb)
fl. 14* viertel), fl.7. Mittägl. zweimal.Postvere, : GansÄ
fl. SS, halbj. fl. 16, vtertelj. fl. S,
Abonnement für das Ausland»
r Vierteljährig:
Be* «ns flir Dcmttchland lofl., ftrr allettbrig«*
fliaaten von Barop» u. Nordenmerihn 11 fl.%
ausserdem bei den Postämtern oder Zeitung*-Expedl«
«lonen in Dentschlund is M. 5 Pf., Kchtvcls
90 Fr. So er., Bumtlnien, Serbien und Bai«
Karlen IS Fr. 40 Ct., Etassland, St. Petersburg
end Moskau 6 stob. 28 Kop., in den anderen Städten
E Bnb, is Kop., Türkei 5 fl. 97 kr., Griechen«
l»n«l 7 fl., Egypten 6 fl. 60 kr. (Alexandrien
fl fl. 97 kr.); ferner in Italien : bei E. E. Obliegh*
h» Mailand» Florenz n. Rom 28 Fr. 50 Ct. ; Frank«
Nicii: bei der Agence Havas in Paris, 34 Kai
Motre Dame des Victoires, 29 Fr. 50Ct.; Englandl
hal A. Siegle, SO Lirae>Street E. C., London.
14 Sh. \$ daselbst für die eugsl. Kolonien mit
Zuschlag das englischen Portos; Nordamerlies*t
hei £, Steiger* 85 Park-Pln.ee in Kewyork» S holl
40 Cts. Gold.
Vit dl« «ft Agenten, Austräger oder VerschlelSMS

hecahlten Beträge leisten wir keine Garantie*

Wien, Freitag, den 24. September

1S8G.

Wien, 23. September.

Morgen wird in Wien ein Denkmal zu Ehren Wilhelm
v. Tegetthosf's mit Feierlichkeit enthüllt. In Marburg,
Tegetthosf's Geüurtsstadt, ehrt bereits ein Standbild den
steierischen Helden, und in: Angesichte des Meeres, das
Tcgetthoff's Element und der Schauplatz seiner Thaten ge -
wesen, erhebt sich bereits seit neun Jahren, aus unserem
Kriegshafen hinausblickend in die weite See, ein erhebender
Abschiedsgruß dem Ausfahrenden, ein freudiger Willkomm
dem Heimkehrenden, die Gestalt des Sehelden. Nun endlich
wird in der Reichshauptstadt, zwanzig Jahre nach dem See -
siede bei Lissa, dem Sieger von 1866, eine Ehrensäule er -
richtet, ihm zur Ehre, der Stadt, in welcher er die letzten
Jahre seines vielbewcgten Lebens zugebracht, zum dauernden
Schmucke, in Zukunft ein Wahrzeichen Wiens, wie die Nelson-
Säule auf dem Trafalgar-Square für London, ein sichtbarer
Ausdruck der Dankbarkeit des Volkes für den glänzenden
Vertreter von Oesterreichs Waffenruhm zur See, den Vor -
kämpfer der Rchichseinheit, den unerschütterlichen Vertreter
von des Kaiserstaates ruhmvoller deutscher Tradition, die in
Wien ihre lebendige Darstellung findet.
' _ Tcgetthoff war uns ein Geschenk des Himmels; plötz -
lich stieg dieser Sohn der steierischen Berge als Held auf
der tückischen See empor, wie der Ruhm unserer Marine
selbst, die bis dahin das Stiefkind der österreichischen Heeres -
verwaltung gewesen, keine Geschichte und also keine Geltung
hatte. Der Name Tegetthoff, an fremden Gestaden, in fernen
Welttheilen bekannter als in der Heimat, ging erst vom
9. Mai 1864 von Mund zu Mund. Seine Kraft, Umsicht,
Energie und Genialität waren allensalls dem Ober-Comman -
danten Erzherzog Ferdinand Max bekannt, der ihn vielfach
verwendete, an der Sulina-Mündung, um Ordnung zu hal -
ten während des Krimkrieges, ihn dann wieder einmal die
Küsten des Rothen Meeres bereisen ließ, um eine Kohlen -
station aufzusuchen, während des italienischen Krieges ihn
im Golfe von Venedig kreuzen läßt, dann ihn an seine Seite
ruft, um in seiner Gesellschaft Brasilien zu bereisen, wie uns

der unglückliche Ferdinand Max in. dem Buche „Ans meinem Leben" erzählt, und ihm nach der Abreise König Otto's eine Flottenabtheilung zum Schutze der in Griechenland lebenden Oesterreicher anvertraut. Da kommt der Krieg mit Dänemark sink Schleswig-Holsteins willen, und Tegetthoff erhält die Aufgabe, die deutsche Küste zu schützen. Das Seegefecht vom 9. Mai 1864 gegen den überlegenen dänischen Feind offenbart mit Einem Schlage dem genialen Admiral, dessen glückliche Manöver, Kühnheit, Ausdauer, von Kennern viel bewundert, nicht den Sieg zu erringen, dennoch den Gegner zur Aufhebung der Küstenblockade zu zwingen vermochten, während er selbst wegen der brennenden Schraubenfregatte „Schwarzenberg" Helgoland zu gewinnen suchen mußte. Jetzt ward sein Name weltbekannt. Die Heimat erkannte doch noch immer nicht seine Bedeutung; die Bürgerschaft Hamburgs feierte ihn durch kostbare Ehrengeschenke, auf den nordfriesischen Inseln wurde sein Name mit ehrfurchtsvoller Scheu genannt. Immerhin, Tegetthoff hatte den seit Jahrhunderten entschlafenen Ruhm der deutschen Seetüchtigkeit dauernd geweckt; er hatte Oesterreich auf ein neues Element seiner Kraft aufmerksam gemacht, er schuf der Marine, deren Officiere und Führer sich nicht täglich in der Gunst der einflußreichsten Hofkreise sonnen konnten, in deren Mitte sich nicht die Officierspatente mit dem Adelsbriefe vererbten, die auf fernen Meeren ihren Muth und ihre Kunst erprobten, Geltung und bereitet sie zu großen Thaten vor.

Mit welchem Siegesbewußtsein zog unsere Nordarmee gegen Preußen! Und in sieben Tagen vollzogen sich die traurigen Geschehnisse Oesterreichs, erlag eine schlecht bewaffnete, schlecht geführte, an Zahl schwache, schlecht gepflegte, aller Fortschritte spottende Armee in beispiellosem, opfervollem, heroischem Kampfe, und Oesterreich verlor darüber seine vielhundertjährige Stellung in Deutschland, die geschichtlichen Wurzeln seines Bestandes, die Grundlage seiner innern Entwicklung mit dem in der deutschen Bundesverfassung schon allein begründeten Uebergewichte des deutschen Volksstammes.

Was war von der Südarkmee, was nun gar erst von der Flotte zu erwarten? Doch die Südarkmee siegte glänzend bei Custvizza; allein dieser Sieg, wie sehr er zu preisen sein mag, er war nur eine rühmliche Fortsetzung der Großthaten des Kaiserlichen in Ober-Italien. Der große Seesieg Tegetthoff's bei Lissa hingegen, er steht einzig da in unserer Geschichte.

Mit gerechtem Stolze gedenken wir allzeit der See - schlacht vom 20. Juli 1866! Sie gilt aller Welt als eine der wichtigsten, sie war epochemachend für die seemannische Kriegsgeschichte der Neuzeit. Sie war der glänzendste Erfolg einer viel schwächeren über eine stärkere Flotte. In Zahlen spricht sich die Lage der Kämpfer aus: 12, beziehungsweise 24 Riesenpanzerschiffe der Italiener standen gegen 7 der Oesterreicher, 248 italienische Geschütze feuerten gegen 173 österreichische! Oesterreich besaß keine der neuen schweren Geschütze gegen Italiens gepanzerte Rammcolosse mit ihren Armstrongs! Aber Oesterreich hatte einen Tegetthoff, der am Vorabend des Kampfes seine Holzschiffe mit Kettenpanzern bekleidete, der die Lehren des amerikanischen Krieges mit Scharfsinn und Fleiß sich angeeignet hatte, der rasch seine Taktik zu ändern verstand und mit einer Kühnheit den ungleichen Kampf aufnahm. Das ist eben Tegetthoff, der einstige Genosse Theodor Heuglin's, der ans offene arabische Boot (vor dreißig Jahren) über Kossair, Suakim, Massauah die Korallen-Inseln, Sokotora — heute geläufige Namen! — durchforschte, als Gefangener der Beduinen seine Ruhe, seinen Trotz nicht verlor. Und diesen selben Tegetthoff finden wir auf der Reconoscirungs-Fahrt nach Ancona, den Persano nicht anzugreifen wagt, als Leiter der Schlacht, mitten in der Gefahr, wie vorher in rascher Ausbildung und Vorbereitung seiner Truppen. Persano erschien endlich vor Lissa und beschoß die Befestigungen der Insel. Nach zweitägigem Kampfe war er dabei, die Landungsmannschaft auszuschießen, am 20. Juli, als der „Esploratore" meldete, die österreichische Flotte nahe heran. Tegetthoff hatte seine Flotte in drei Treffen formirt, ließ den Feind anlaufen und dann in concentrischen Lagen feuern. Wie er durch den erspähten Zwischenraum der Gegner Aufstellung durchbrach, die Mittelgruppe angriff, sich mit vier Panzern auf den „Re d'Italia" warf, das feindliche Hauptschiff von allen Seiten umstellte, plötzlich mit seinem „Max" gegen dessen Breitseite anfuhr und ihm einen furchtbaren Riß beibrachte, so daß die eindringenden Wogen das stolze Fahrzeug in die Tiefe rissen — das ist, wie die Manöver der anderen Schiffe, unzählige Male zu beiden Seiten des Weltmeeres erzählt worden. Der Untergang des „Re d'Italia" war die Einleitung und der in die Luft fliegende „Palestro" der Schluß der Seeschlacht, die Tegetthoff's Namen unsterblich gemacht. Es fand sich

in Italien kein zweiter Admiral, der mit dem Sieger den Kampf wagen wollte.

Sein Sieg war für die damalige Lage unfruchtbar, aber er war es nicht für Oesterreich, seinen Ruhm, seine Tradition, seine Ehre. Inmitten der Katastrophe von 1866, nach den Niederlagen in Böhmen, dem Bankrott einer dem Rechte treubruchigen Staatsverwaltung, den Zuckungen im Innern, dem entfesselten Nationalitätenkampfe, einer Trübsal ohnegleichen, weckte Tegetthof's Sieg den Glauben an Oesterreich, sein Können, seine Volkskraft. An seiner That konnten wir uns ausrichten, hat sich Oesterreich aufgerichtet. Er war unser Stolz, und wir sahen ihn gefeiert in Amerika 1867, in Frankreich und England. Sein Ruhm half ihm bei seiner Mission, die Leiche seines theuren Herrn und Freundes, des Kaisers von Mexico, nach Wien zu holen. Dann ließ er sich in unserer Mitte nieder, als Leiter der Marine-Angelegenheiten, wir sahen ihn schaffen für die Zukunft unserer Seemacht; er begleitete seinen Kaiser zur Einweihung des Suez-Canals, den er in einer unübertroffenen Abhandlung und mit einer freimüthigen Kritik der österreichischen handelsmaritimen Verhältnisse einst schriftstellerisch gewürdigt hatte, und wir sahen ihn auch dort hochgeehrt. Heimgekehrt, nahm er in unserm Herrenhause regen Antheil an der inneren Politik und, wie nicht anders denkbar, im Sinne des Reichsgedankens. Dort saß er auf der äußersten Linken, treu zur Verfassungspartei stehend, ein Liberaler ohne Frage, auch in allen damals vielbewegten konfessionellen Reformen, allen Hohenwart'schen Versnuckmungen mit kräftigem „Nein“ widerstehend; denn sein Conspass wies nur auf Einheit, Ehre, Macht des Reiches. Ein wissenschaftlich und politisch hochgebildeter Mann — die „Neue Freie Presse“ hatte Gelegenheit, dies zu zeigen — wußte er die Feder trefflich zu führen, stand er mitten in seiner Zeit, auf der Höhe der Zeit-Ideen. Das Buch, welches dem das Bewußtsein Verlierenden aus der Hand fiel, war Darwin's „Ursprung des Menschen“. Und als ein grausames Geschick ihn uns am 7. April 1871 im vierundvierzigsten (!) Jahre seines Lebens in einem Augenblicke entriß, wo der Kampf gegen Hohenwart's staatszerreißende Pläne die Gemüther so sehr erregte und man in ihm einen Zukunftslenker feilschen konnte.

Aus Tegetthof's Mappe.

Von den Männern, die im letzten Vierteljahrhundert sich einen hervorragenden Namen gemacht haben, wird Wilhelm v. Tegetthof wohl am längsten in dem Andenken der Mit- und Nachwelt fortleben. Und gewiß würde sein Bild an Lebendigkeit gewinnen, wenn über sein Privatleben und über seinen Charakter mehr bekannt würde, als dies thatsächlich der Fall ist. Das knappe Lebensbild, welches ich vor einigen Jahren entwarf, beruhte zum überwiegenden Theile auf Familienpapieren. Mit dem Tode des Vaters brechen diese vertraulichen Mittheilungen leider ab. Wenn auch Tegetthof mit inniger Liebe an seiner Mutter hing, die ihn bekanntlich überlebte, so hat er sich ihr gegenüber doch nicht über seine amtliche Stellung und seine Thätigkeit ausgesprochen, während er seinem Vater, der an allen öffentlichen Angelegenheiten den regsten Antheil nahm und mit seltener Klarheit Menschen und Verhältnisse beurtheilte, die kleinsten Vorgänge mittheilte und sein volles Herz ausschüttete. Von den vielen Privatbriefen ist bisher wenigstens nichts in die Öffentlichkeit gedrungen, und doch wären wir an der Hand derselben allein in der Lage, mancherlei Einzelheiten aufzuzuhellen, um Tegetthof und die ganze Liebenswürdigkeit seines Charakters voll kennen zu lernen. Die Briefe an seinen Vorgesetzten Fautz enthielten über die Vorgänge bei Helgoland interessante Angaben. Wohin sind diese gerathen? Auch der jetzige Leiter des Marine-Ministeriums würde sich ein Verdienst erwerben, wenn er bei seinen Kameraden Nachfrage hielte und vielleicht aus seinem Schatzkästlein Einiges mittheilen würde.

Tegetthof hat sich seine Stellung durch eigene Kraft errungen. Langsam stieg er empor. Die Legende war allerdings geschäftig, um ihm von seinem ersten Eintritt in den activen Marinedienst einen Beschützer zuzuweisen, dessen Scharfblick die ganze Bedeutung des Mannes erkannt und seine rasche Beförderung sich habe angelegen sein lassen. Daran ist kein wahres Wort. Als Tegetthof die Marine-Akademie zu verlassen im Begriffe stand, hoffte der zärtliche Vater allerdings, daß auf die Zukunft seines Sohnes eine fördernde Land Einflüsse haben dürfte. Ein Schreiben des Erzherzogs Friedrich an den Oberstlieutenant Tegetthof vom 1. Juli 1847 erweckte diese Hoffnung. „Es ist mit wahrem Vergnügen,“ heißt es daselbst, „daß ich im Laufe dieses Jahres, wo das Marine-Collegium unter meinen unmittel-

baren Befehlen steht, die guten Fortschritte und die vortreffliche Aufführung Ihres Sohnes Wilhelm bemerkte, und kann daher nicht unterlassen, ihm hiemit meine ganze Zufriedenheit zu spenden. Es wird mir eine sehr angenehme Pflicht sein, dafür zu sorgen, daß er bei seinem Austritte jene Be Stimmung erhalte, welche zur praktischen Ausbildung eines jungen angehenden Seemannes die beste ist."

Bekanntlich starb Erzherzog Friedrich bereits am 5. October 1847. Die ihm zugedachte Stellung in der Leitung der Marine fiel später Erzherzog Ferdinand Maximilian anheim. Es ist ganz unrichtig, wenn behauptet wird, daß dieser Tegetthoff besondere Gunst zugewendet habe; so lange der Erzherzog an der Spitze der Marine stand, ist Tegetthoff nicht rascher aufgestiegen, als irgend ein anderes Mitglied der Marine. Erst die Tage von Helgoland und Lissa verschafften ihm die höheren Chargen, nachdem der Erzherzog längst Miramar mit dem Palaste Chapultcpcc vertauscht hatte. So viel ist richtig, daß der Erzherzog Tegetthoff, der auch anfangs an die erzherzogliche Wirksamkeit große Hoffnungen knüpfte und für die Marine den Anbruch besserer Tage erwartete, in seine Nähe rief, um den brauchbaren Mann zu verwenden. Allein so oft Beide in nähere Berührung kamen, trat die entschiedene Unverträglichkeit der Charaktere beider Männer in nicht selten schroffer Weise hervor. Der Erzherzog hatte von den Menschen überhaupt die schlechteste Meinung. Frühzeitig von Schmeichlern umgeben, mit scharfem Auge für die Schwächen der Menschen, widerte der nicht immer würdige Wettstreit seiner Umgebung nm seine Gunst, das Rennen nach Macht und Einfluß ihn an und festigte bei ihm die Ueberzeugung, daß die Führung und Beherrschung der Menschen ein Kinderspiel sei.

Tegetthoff erhielt in Aden die Nachricht, daß er zum Vorstande der ersten Section beim Marine-Commando beim Erzherzog ernannt worden sei. Ein Anderer würde sich ganz und voll über diese Auszeichnung gefreut haben: nicht so Tegetthoff. Für den Bureaudienst wählte er nicht die erforderliche Eignung zu besitzen; aus dem Schiffe fühlte er sich heimisch: sein Lieblingsgedanke war damals, an die Spitze

einer Expedition gestellt zu werden, fremde Länder kennen zu lernen. Auch mochte er sich der Empfindung nicht entschlagen, daß große Kämpfe in dem ihm nun zugewiesenen Wirkungskreise seiner harren. Sein offener Charakter verschmähte die

vielen Mittekchn, um sich in der Gunst der Mächtigen zu erhalten. Keine Kunst war im Stande, ihn seiner Ueberzeugung und Pflicht abspenstig zu machen. Es ist heute noch nicht möglich, die Einzelheiten des Verhältnisses zwischen Tegetthoff und seinem Vorgesetzten bloßzulgeii; die Bemerkung möge genügen, daß Tegetthoff in den meisten Fragen, um die es sich damals handelte, durch Geradheit und Offenheit als Sieger hervorging. Allein die Kämpfe verleiteten ihm seine Stellung, und der spätere Sieger bei Lissa trug sich im Jahre 1862 ernstlich mit dem Gedanken, den Dienst ganz zu verlassen. Während der Reise des Erzherzogs nach Brasilien, wohin ihn Tegetthoff begleitete, war der Gegensatz der Charaktere ain herbsten zum Durchbruche gekommen, und seitdem ging Tegetthoff jedem Anwürfe, an der Seite des kaiserlichen Prinzen zu wirken, ans dem Wege.

Von den Vorfällen erfuhr Niemand ein Wort; über die intimen Vorgänge schwieg Tcgetthoff beharrlich. Nie rühmte er sich der Dienste, die er dem Erzherzoge geleistet.¹ Einem künftigen Geschichtsschreiber der Marine wird es vielleicht möglich sein, die großen Verdienste, die sich Tegetthoff um die Marine in seiner Stellung als Referent erworben hat, ins Helle Licht zu setzen. So viel läßt sich schon jetzt behaupten, daß manche Maßregel, für die Andere das vollste Lob einheimsten, von Tegetthoff angeregt und vorbereitet wurde. So sehr wahrte er aber auch später das Geheimniß, daß selbst seine vertrautesten Freunde über seine Leistungen nie ein Wort erfuhren. Ans der Reise nach Brasilien sendete der kaiserliche Prinz regelmäßig Berichte ein, die geradezu meisterhaft genannt werden können; herrliche plastische Schilderungen über Land und Leute, scharfe Charakteristikeil hervorragender Persönlichkeiten, mit denen die Reisenden in Berührung kanien, durchsichtige Darstellungen der politischen Verhältnisse, große Blicke über den Welthandel und die Aufgaben Oesterreichs im Wettbewerbe mit den anderen handeltreibenden Nationen füllen diese Blätter. Dazu¹ gesellt sich eine nicht selten bewunderungswürdige Beherrschung der Sprache, die einem jeden Schriftsteller Ehre machen würde. Zweifellos, daß diese Schriftstücke in den Kreisen, für

[Neue Freie Presse - 18860924 - Seite 2]

24 . septemner issv

Ufr. ' 0)31

Seite 2 Wien, Freitag Neue Freie Ir*resse.

des österreichischen^ Aatatschisfes zu erblicken währte, was seine außerordentliche Bescheidenheit niemals ernst nehmen wollte, da konnte einer seiner Genossen im Herrenhause in einem in diesem Blatte erschienenen Nachrufe sagen: „Eine der glänzendsten Zierden des Herrenhauses, stand er wie dieses selbst fest zur Verfassung, unerschütterlich.“ — Wenn heute sein Geist herniederstiege!

Er war ein Held, eilt Admiral, ein Soldat, aber auch ein freisinniger Denker, ein von Leidenschaft und Selbstsucht freies Herz, das mit der Wärme eines biedereren, echt deutschen Wesens an Oesterreich hing, an dem alten Oesterreich. Solch ein volksthümlicherHeld muß im Denkmale fortleben! Heute erneuert sich der Schmerz über seinen Verlust, denn „Er war unser!“ im vollen Sinne des Wortes. Unvergessen ist sein Verdienst, sein Name schwebt auf allen Lippen, wenn wir an Krieg denken; er dünkt uns ein Muster, wenn wir uns den Soldaten im Sinne der modernen Zeit darstellen. Möge sein Denkmal den Weg der Ehre, des Ruhmes den späteren Geschlechtern zeigen und er würdige Erben finden! Das Schwurgericht vor dem deutschen Juristentage.

Wien, 23. September.

Schon seit dem Bestände der gegenwärtigen deutschen Gerichtsverfassung herrscht unter den Juristen im deutschen Reiche ein mehr oder minder heftig geführter Streit über die Frage : ob Schöffen-, ob Schwurgericht? Die Anhänger des erstern wollen die Schöffen in größerer Anzahl auch bei den Gerichten höherer Ordnung in Verwendung bringen und damit die Schwurgerichte beseitigen. Von diesem Gedanken getragen wab der Antrag Olshausen's beim letzten deutschen Juristentage. Allein es fanden sich Männer genug, welche gegen dessen Ausspruch, als verdienten die Schwurgerichte'nicht das ihnen theilweise (?) „geschenkteVertrauen“, Entschieden Verwahrung einlegten, so daß schließlich in Betreff der Schwurgerichte der Beschluß nur dahin ging, daß die dermalige Einrichtung des schwurgerichtlichen Verfahrens der Reform dringend bedürftig sei.

Daß unter dieser Phrase verschiedene, vielleicht sogar

entgegengesetzte Wünsche' verborgen liegen, dürfte kaum bestritten werden. So können' die Vertreter der Schöffengerichte die erwähnte Reform leicht dahin verstehen, daß das Laien-Element im Strafprocesse seines selbstständigen Charakters, wie ihn das Schwurgerichts-Verfahren wahrte, entkleidet und dem gelehrten Richter-Collegium beigeordnet werde, womit die gewünschten Schöffengerichte höherer Ordnung von selbst gegeben sind, während dadurch die Schwurgerichte gleichzeitig außer Wirksamkeit treten. Allerdings wäre dies nicht mehr eine Reform, sondern eine Vernichtung der Schwurgerichte. Allein der gesunde, von juristischem Dünkel freie Sinn des Volkes wird dafür Sorge tragen, daß die für die allgemeine Wohlfahrt so wichtige Einrichtung der Geschwornengerichte unverkümmert erhalten werde.

— Als seinerzeit nach' dem Sturze des ersten Napoleon die Rheinlande wieder zu Deutschland kamen, baten deren Einwohner nur um die Belastung zweier unter der französischen Herrschaft erlangter Institutionen, nämlich: der Civil-Ehe und des Geschwornengerichtes. Sollten die Deutschen von heute den Werth des zuletzt genannten Instituts weniger erkennen und zu erhalten trachten? Man täusche sich doch nicht. Auch das besteingerichtete Schöffengericht kann die Jury nie und nimmer ersetzen. Es hält gar keinen Vergleich aus mit der Bedeutung des Geschwornengerichtes, weil ihm gerade dasjenige fehlt, was dem im Strafprocesse verwendeten Laien-Elemente Werth und Bedeutung gibt. Der Meister des deutschen Strafproceßes, unser zu früh verstorbener Glaser, würde sich nicht wenig wundern, daß man heute mit solchem Eifer eine Einrichtung bevorzugen will, von deren Werthlosigkeit er durchdrungen war, wofür die von ihm geschaffene Strafproceß-Ordnung Zeugniß gibt.

Man horche doch hin, wo sich die Feinde des Schwurgerichtes regen. Gewiß nicht in den Reihen derjenigen, welchen der gesetzmäßige Fortschritt der Freiheit am Herzen liegt. Wenn die modernen Gesetzgebungen, worunter selbst Rußland, die Heranziehung des Laien-Elementes zum Strafprocesse als eine im eigensten Interesse der Rechtspflege gebotene Maßregel anerkennen, so kann es nur in jener Form geschehen, daß die aus dem Volke beigezogenen Richter, wie es im Schwurgerichts-Verfahren der Fall ist, selbstständig und unabhängig über den ihnen zur Entscheidung vorgelegten Gegenstand berathen und

beschließen. Jede andere Form der Mitwirkung entbehrt der Bürgschaft für dasjenige, was mit der Heranziehung des Volkes Clementes zur Rechtsfindung geleistet werden soll. Auch aus die Geschwornen kann das bekannte Wort Anwendung finden: Link, nt sunt, aut non sint. Auch in Oesterreich gab es seinerzeit eine Art Schöffengerichte. Man hat sie fallen lassen, als in der Mitte dieses Jahrhunderts der Anklageproceß mit den Garantien der Oeffentlichkeit und Mündlichkeit sowie des Geschwornengerichtes in unsere Lande einzog. Allerdings stehen heute dem Einzelrichter bei uns keine Schöffen zur Seite; dafür schützt den Angeklagten einerseits die Oeffentlichkeit und Unmittelbarkeit des Verfahrens, andererseits können Schuld und Strafe durch ein Gericht höherer Ordnung überprüft werden.

Es befremdet fast, wenn man fort und fort die Geschwornen ob einzelner falscher Urtheile tadeln hört. Als ob die Urtheile gelehrter Richter von jedem Irrthum frei wären! Die unbegründeten Freisprechungen auf jener Seite werden sattem ausgewogen durch unrichtige Verurtheilungen auf dieser Seite. Und was das Schlimmere von Beiden ist, braucht nicht erst gesagt zu werden. Der in Sachen der Schwurgerichte Erfahrene wird es auch bestätigen können, daß, wenn unrichtige Wahrsprüche gefällt werden, hieran nur zu häufig nicht die Geschwornen, sondern ein oder der andere der mitwirkenden rechtskundigen Factoren die Schuld trägt. Es wurde in dieser Beziehung auch von verschiedenen Seiten auf die mangelhafte Fragestellung und Belehrung hingewiesen. Diese Mängel aber, welche überdies nicht die einzigen sind, haben doch nicht die Geschwornen zu verantworten.

Man lasse doch die Geschwornengerichte ruhig sich fortentwickeln! Die Mängel, welche dieser wie jeder andern menschlichen Einrichtung anhaften, werden hundertfach ausgewogen durch die Segnungen, welche sie bietet. Selbst in Frankreich, dessen Strafproceß lange nicht mehr aus der Höhe der heutigen Wissenschaft steht, ist es, trotz allem Wechsel der Regierungsform, noch Niemandem beigefallen, an dem Bestände der Jury zu rütteln. Uebrigens wären gerade diejenigen, welche die Geschwornengerichte am heftigsten tadeln, berufen, diesen Tadel durch die richtige Handhabung des Gesetzes verstummen zu machen. Wenn Gerichtshof und Geschworne, Ankläger und Vertheidiger im Sinne des Gesetzes zusammenwirken, dann, aber -nur dann gibt es einen guten Klang. Nimmer aber hängt es von

Einem Factor allein ab. Der Weg zum Bessern liegt daher auch nicht in der Umwandlung der Geschwornen in Schöffen, sondern in der Fortbildung der Ersteren. Wer da die Schöffen statt der Geschwornen wist, der wirft eine feste Säule um, um dafür einen schwanken Stab einzutauschen; der rüttelt an einem Bollwerke, welches gerade in den schwierigsten Zeiten berufen ist, dem Rechte eine sichere Stütze zu bieten.

Inland.

Wien, 23. September. (Die Session der Vertretungskörper.) Ueber die Dauer der Arbeitszeit, welche den Vertretungskörpern der beiden Reichshälften in diesem Jahre gegönnt sein wird, liegt heute eine officiële Aeußerung des ungarischen Minister-Präsidenten vor. Danach werden der Reichsrath und der ungarische Reichstag bis Ende October versammelt bleiben und dann den Delegationen Platz machen. Nach der Delegations-Session, welche diesmal mindestens vier Wochen beanspruchen dürfte, wird sich der ungarische Reichstag wieder versammeln. Ob in Oesterreich zu gleicher Zeit der Reichsrath wieder zusammentreten oder ob nach den Delegationen die Landtage einberufen werden, scheint noch nicht festgestellt zu sein.

Wie», 23. September. (Die Affaire des Freiburger Gymnasiums.) Gegen den Unterrichtsminister Gautsch hat sich wegen der von ihm angeordneten Schließung des widerrechtlich eröffneten Gymnasiums in Freiberg ein förmlicher Entrüstungssturm in der Brunner und Prager czechischen Presse erhoben. Ueber die Umstände, welche zur Schließung dieser Privatschule führten, sendet uns unser Brünner Correspondent die folgende Darstellung: „Das ehemals in Freiberg bestandene deutsche Staats-Untergymnasium wurde vor Jahresfrist auf Vorschlag des Landesschulrathes wegen geringer Frequenz und angesichts der örtlichen Verhältnisse in Freiberg, welche dem Gedeihen eines Gymnasiums ungünstig seien, vom Unterrichtsministerium aufgehoben. Die Gemeinde ersuchte um den Fortbestand des deutschen Gymnasiums oder eventuell um Umwandlung desselben in ein czechisches Gymnasium; der Landesschulrath und das Ministerium wiesen jedoch übereinstimmend dieses Ansuchen ab, da auch bei einem czechischen Gymnasium dieselben ungünstigen örtlichen Verhältnisse fortbestehen und überdies durch die Errichtung eines czechischen Gymnasiums in Freiberg den nahegelegenen czechischen Gymnasien in Walachisch-Meseritsch und Prerau Schüler entzogen

würden. In jüngster Zeit schritt nun die Freiburger Gemeinde um Bewilligung zur Errichtung eines czechischen Privat-Gymnasiums ein. Die czechische Majorität des Landesschulrathes be- fürwortete dieses Gesuch vor beiläufig acht Tagen beim Unterrichtsministerium. Einstweilen stellte die Freiburger Gemeinde Professoren an, es wurde von der czechischen Partei durch umfassende Agitation eine Anzahl Schüler beschafft, und die Gemeinde eröffnete, ohne die hiezu gesetzlich erforderliche Bewilligung abzuwarten, die erste Classe. Vor drei Tagen wurde vom Unterrichtsminister Dr. v. Gautsch telegraphisch die sofortige Schließung dieser widergesetzlich errichteten Privat-Lehranstalt angeordnet. Ueber das Herabblenden einer meritorischen Entscheidung des Ministers bezüglich des Gesuches der Gemeinde ist bis nun nichts bekannt. Auch in Angelegenheit des Freiburger Gymnasiums zeigt sich die in der letzten Zeit mehrfach zu Tage getretene Divergenz zwischen den Anschauungen des Unterrichtsministeriums und jenen der czechischen Mehrheit des Landesschulrathes." Die „Politik" macht den Versuch, die Schließung des Gymnasiums auf einen bloßen Formfehler zurückzuführen, den die Gemeinde Freiberg begangen haben soll; sie verschweigt aber, daß es sich um ein Privat-Gymnasium handelt, welches zu seiner Eröffnung einer ausdrücklichen Bewilligung bedarf. Es ist doch eine gar zu plumpe Bauernfängerei, wenn man behauptet, das Nichteintreffen des Verbotes sei gleichbedeutend mit der Bewilligung der Schließung.

Wie», 23. September. (Vom Verwaltungs-Gerichtshof. Eine gewerberechtliche Frage.) Julie Widmann, die Besitzerin einer Leihbibliothek und eines Antiquariats in Czernowitz, führte beim Verwaltungs-Gerichtshof Beschwerde gegen die ihr seitens der Gewerbe-Behörde verweigerte Ausfolgung eines Gewerbescheines für die Buchbinderei. Sie hielt nämlich seit mehreren Jahren auch eine Buchbinderwerkstätte, hauptsächlich für die Zwecke ihrer Antiquar-Buchhandlung und ihrer Leihbibliothek, und als sie nach dem Inlebensreten der Gewerbe-Novelle vom Jahre 1883 wegen Ausübung der Buchbinderei Anstände hatte, schritt sie im vorigen Jahre beim Magistrate in Czernowitz um einen Gewerbeschein für das Buchbinder-Gewerbe ein, dessen Ausübung sie auf die Zwecke ihres Antiquariats beschränken zu wollen erklärte. Sie wurde mit ihrem Ansuchen abgewiesen, weil die Buchbinderei ein handwerksmäßiges Gewerbe ist, zu dessen Betrieb der Nachweis der Befähigung erforderlich ist und ein Gewerbeschein für die beschränkte Ausübung des Gewerbes nicht ausgestellt werden könne. Die Landesregierung in Czernowitz und das Ministerium des Innern bestätigten diese

Entscheidung, Letzteres mit dem Bemerken, daß der Beschwerdeführerin gestattet sei, zur Erhaltung ihrer Leihbibliothek eine Buchbinderwerkstätte zu halten. Der Verwaltungs-Gerichtshof wies nach der heute durchgeführten Verhandlung die Beschwerde als unbegründet zurück, weil, wie in den Motiven ausgesprochen wurde, die Beschwerdeführerin zur Zeit, als noch das erste und zweite Hauptstück der Gewerbe-Ordnung vom Jahre 1859 in Wirksamkeit waren, den Betrieb des Buchbinder-Gewerbes nicht angemeldet hatte, daher zur selbstständigen Ausübung dieses Gewerbes nicht berechtigt war, sonach finde hier die Gewerbe-Novelle vom Jahre 1883 Anwendung, derzufolge der Betrieb der Buchbinderei als eines handwerksmäßigen Gewerbes an den Befähigungs-Nachweis gebunden ist, den die Beschwerdeführerin nicht erbracht hat und für dessen Beibringung der Umfang des Gewerbebetriebes nicht entscheidend sei.

die sie bestimmt waren, Aufsehen machten, aber aus Tegetthoff's Munde erfuhr Niemand, daß er die Feder geführt hatte.

Erst nach dem Gefechte von Helgoland und dem Siege bei Lissa mochte Kaiser Max die ganze Größe des Mannes erkannt haben. In einem Schreiben an Tegetthoff aus Mexico vom 6. Juli 1.864 sprach er seine Bewunderung über die Leistung der Marine, über die Umsicht und Tapferkeit des Führers aus, und zwei Jahre später sendet er von Chapultepec „dem Admiral und Freunde" seine tiefgefühlten Glückwünsche. Und in noch vollerm Tone spricht der Kaiser von Mexico in. einem Schreiben an den Vice-Admiral v. Dahlerup über die Leistungen der Marine. „Tegetthoff," schreibt er, „hat nicht nur für seine Flagge bei Lissa, er hat gleichzeitig für unsere österreichische Ehre gesiegt." Und mit besonderer Genugthuung gedenkt er der Reformen, die unter seiner eigenen Leitung angebahnt und durchgeführt worden sind. Auch die anderen Vorgesetzten Tegetthoff's können nicht das Verdienst für sich in Anspruch nehmen, seine unübertrefflichen Talente früh erkannt und zu seiner Förderung beigetragen zu haben. Mit Bourguignon, der an der Spitze der Marine-Behörde stand, hatte Tegetthoff nach seiner Ernennung zum Referenten oft in Verbindung zu treten, allein die guten Beziehungen wurden nicht selten, wenn es sich um ernste Geschäfte handelte, getrübt. Der „Homo degli affari di somma importanza", wie Bourguignon in der Marine genannt wurde, stand in den meisten Fragen in Hellem Gegensatz zu Tegetthoff. An althergebrachten Einrichtungen

mit Fähigkeit festhaltend, war Bourguignon für eine einschneidende reformatatorische Thätigkeit ganz ungeeignet, und der energische feurige Tegetthoff, der mit scharfem Blicke für die Nothwendigkeit, Umgestaltungen in der Marine vorzunehmen, schon als junger Officier eintrat, bereitete ihm manch bittere Stunde. Auch später, nachdem Tegetthoff mit der Leitung der Marine-Angelegenheiten betraut war und Bourguignon als Hafen-Admiral in Pola fungirte, das heißt auf einem Posten sich befand, dessen Jnhecker mehr als jeder Andere in der Marine die Fähigkeit besitzen muß, den Anforderungen der Neuzeit gerecht zu werden, fehlte es nicht an harten Kämpfen. Und hiebei bekundete Tegetthoff eine Güte und Weichheit des Herzens, die man bei ihm nicht geluht hätte. Der energische, uaciszuldiue Mann zögerte gegen

seine sonstige Gewohnheit, radical einzugreifen, einmal weil es sich um seinen ehemaligen Vorgesetzten handelte, sodann aber weil Bourguignon sich an seinen „edlen, ritterlichen und tapfern Freund“ mit der Bitte wendete, „ihn noch lange im activen Dienste zu belassen“, und Tegetthoff, ein seelen-guter Mensch, konnte es nicht über sich gewinnen, dem älteren Manne eine abschlägige Antwort zu ertheilen.

Mit dem Vice-Admiral Fautz unterhielt Tegetthoff gute Beziehungen. Schon in der vormärzlichen Marine erfreute sich Fautz eines trefflichen Rufes als wackerer, charakterfester Officier. Ein braver Soldat, ein tüchtiger Seemann, stellte er jederzeit seinen ganzen Mann. Aber ohne wissenschaftliche Bildung, ohne Befähigung für die Behandlung großer Fragen, ließ er, an die Spitze der Marine-Verwaltung gestellt, die Dinge gehen, ohne eingreifen zu wollen oder zu können. Ließen die Verhältnisse bei der Marine in der zweiten Epoche der erzherzoglichen Wirksamkeit Manches zu wünschen übrig, so traten unter Fautz die Gebrechen und Mängel noch schärfer hervor. Ohne Initiative, ohne schöpferischen Gedanken, blieb die Erledigung der Acten das Um und Aus der ganzen Verwaltung. Tegetthoff schätzte an dem Manne die Integrität des Charakters, seine Thätigkeit als Soldat und Seemann. Während der Expedition nach Helgoland erhielt Fautz ausführliche Briese, die in militärischen Kreisen großes Aufsehen machten. Seit dem Frühjahr 1866 trübte sich das bisher gute Verhältniß.

Tegetthoff wurde nach Wien berufen, um an die Spitze einer ostasiatischen Expedition gestellt zu werden. In der Re-

sidenz angelangt, gewann es immer mehr den Anschein, daß Oesterreich sich am Vorabende eines Krieges mit Preußen und wahrscheinlich auch mit Italien befinde. Tegetthoff hielt es für seine Pflicht, die maßgebenden Kreise darauf aufmerksam zu machen, welche große Aufgabe der Marine zu fallen würde, wenn nur rechtzeitig die nöthigen Mittel bewilligt werden. Er klopfte an alle Thüren, vergebens; man speiste ihn mit freundlichen Worten ab oder kehrte auch dem lästigen Manne den Rücken. Fast verzweiflungsvoll klingen die Aufzeichnungen aus jenen Tagen: „Es sei alle Aussicht vorhanden, eine gewaltige San aufzuheben“, „man ignore die Marine, besitze durchaus nicht das Verständniß, die vorhandene Seemacht zu verwerthen; das Kriegsministerium betrachte factisch die Marine nicht als in das eigene Ressort

gehörig, es treffe in einem Theile kriegerische Anstalten und erfreue sich im andern Theile eines friedlichen Schlummers“.

Daß nun Personen, welche der Marine fernstanden, seinen Darlegungen kein Gehör schenkten, darüber setzte sich Tegetthoff mit leichtem Spotte hinweg. Aber daß auch der Mann an der Spitze der Marine kein Verständniß für die Sache bekundete und dieselbe fast so behandelte, wie etwa die Erledigung eines Actenstückes, verdammt Tegetthoff scharf und eindringlich. Und wenn er auch die Zustände bei der Marine genügend kannte, einen vollen Einblick in die damals herrschende Mißwirthschaft gewann er doch erst, nachdem er nach mehrwöchentlichem Zögern und Zaudern endlich den Auftrag erhalten hatte, an die Ausrüstung einiger Schiffe zu gehen. Tegetthoff legte sich seitdem keinen Zwang auf und machte seinem Unmuth über Fautz vielleicht manchmal in allzu harten Ausdrücken Luft.

Am meisten fühlte sich Tegetthoff zu Wüllerstorff hingezogen. Die allgemeine Bildung, welche der Handelsminister unter Belcredi unstreitig besaß, seine feinen, gewinnenden Umgangsformen wirkten auf Tegetthoff faëcinirend — ohne daß er deßhalb den Werth des Mannes höher angeschlagen hätte, als dieser es in der That verdiente. Nichts schätzte der Sieger von Lissa höher, als Kenntnisse und Bildung, und Wüllerstorff überragte seine Marinegenossen in dieser Richtung um Kopfeslänge. Gerne hätte es Tegetthoff gesehen, wenn Wüllerstorff unter ihm eine dienstliche Verwendung angenommen hätte. Er trug ihm die Leitung der Expedition im Jahre 1868 an, hatte die Absicht, ihm das Hafen-Admiralat

in Pola anzuvertrauen, aber Wüllerstorff war aus seiner Ruhe in Graz nicht herauszulocken. Eine der schönsten Charakter-Eigenschaften Tegetthoff's war es unstreitig, fremdes Verdienst willig und neidlos anzuerkennen. Jede tüchtige Leistung seiner Kameraden fand an ihm den eifrigsten Lobspender. Sein Bericht über die Schlacht bei Lissa wird immer ein Denkmal echter Größe bleiben. Mit

welcher Objectivität und Bescheidenheit rechtfertigt er sein Unternehmen, wie rein und hell klingt das Lob seiner Officiere und der Mannschaft! Bei jedem Schiffe, welches an der Action Antheil nahm, gedenkt er der Verdienste der Commandanten. Nicht selten wortkarg im Umgänge, war er erprobte» Freunden gegenüber von einer seltenen Offenherzigkeit, dank.

11.1.3. Ida Pfeiffer: Reise einer Wienerin in das Heilige Land, Kapitel 1

Reise von Wien nach Konstantinopel

Seit Jahren lebte der Wunsch in mir, eine Reise in das Heilige Land zu machen. Jahre gehören dazu, um mit dem Gedanken eines so gewagten Unternehmens vertraut zu werden. Als daher meine häuslichen Verhältnisse sich so gestaltet hatten, daß ich mich wenigstens auf ein Jahr entfernen konnte, hatte ich nichts eifriger zu tun, als mich auf diese Reise vorzubereiten. Ich las manche Werke darüber und war auch so glücklich, mit einem Herrn bekannt zu werden, der einige Jahre früher jene Länder bereist hatte. Ich konnte mündlich manche Belehrung und manchen Rat über das Fortkommen und Verhalten auf dieser gefährvollen Wanderung erhalten.

Vergebens suchten meine Verwandten und Freunde, mich von diesem Vorsatz abzubringen. Höchst lebhaft stellte man mir all die Gefahren und Beschwerden vor, die den Reisenden dort erwarten. Männer hätten Ursache zu bedenken, ob ihr Körper die Mühen aushalten könne und ob ihr Geist den Mut habe, dem Klima, der Pest, den Plagen der Insekten, der schlechten Nahrung usw. kühn die Stirn zu bieten. Und dann erst eine Frau! So ganz allein, ohne alle Stütze hinauszuwandern in die weite Welt, über Berg und Tal und Meer, ach, das wäre unmöglich. Dies war die Meinung meiner Freunde.

Ich konnte nichts als meinen festen unabänderlichen Willen entgegensetzen. Mein inneres Vertrauen auf Gott gab mir Ruhe und Kraft, meine irdischen Angelegenheiten mit voller Besonnenheit zu ordnen. Ich machte mein Testament, bestellte alles derart, daß im Fall des Todes, worauf ich mehr gefaßt sein mußte als auf eine glückliche Rückkehr, die Meinigen alles in bester Ordnung fänden.

Und somit trat ich am 22. März 1842 meine Wanderung von Wien aus an.

Ich fuhr um ein Uhr mittags zu den Kaisermühlen, dem Platz, von welchem die Dampfschiffe nach Pest usw. abgehen. Freudig überraschte mich am Ufer die Anwesenheit einiger Verwandter und Freunde, die mir nochmals Lebewohl sagen wollten. Die Trennung war freilich recht hart, denn unwillkürlich erfaßte uns der Gedanke, ob wir uns in dieser Welt wohl noch einmal sehen würden.

Ein lebhafter Streit an Bord des Schiffes zerstreute ein bißchen unsern trüben Sinn. Ein Reisender mußte auf Ansuchen eines Herrn, statt mit Sack und Pack nach Ungarn zu flüchten, mit der Polizei in die Stadt zurückkehren. Ersterer schuldete letzterem tausenddreihundert Gulden, und glücklicherweise wurde er noch vor der Abfahrt des Schiffes eingeholt. Kaum war dies geordnet, so gab die Glocke das Zeichen der Abfahrt, die Räder begannen ihre Bewegung und entzogen mich für diesmal meinen Lieben nur zu schnell.

Reisende gab es noch wenige. Die Witterung war zwar schön und mild, aber die Jahreszeit noch zu früh, um andere Reisende als Geschäftsleute oder solche mit so umfassenden Plänen, wie ich sie im Kopfe hatte, in die Welt zu führen. Die meisten gingen nach Preßburg oder höchstens nach Pest. Bald hörte man vom Schiffskapitän, daß eine Frau auf dem Schiff sei, die bis Konstantinopel zu reisen gedenke, und nun betrachtete man mich von allen Seiten. Einer der Herren, der dieselbe Reise machte, sprach mich an und bot mir seine Dienste an, wenn ich deren benötigen sollte, und wirklich stand er mir überall schützend zur Seite.

Die schöne milde Witterung wechselte bald mit Wind und Kälte, als wir hinaus in die große Donau kamen. Ich schlug mich in meinen Mantel ein und blieb auf dem Verdeck, um die Umgebung zu sehen, die von Wien bis Preßburg wohl recht lieblich sein mag, wenn sie im Frühlingsschmuck prangt, jetzt aber nur kahle Bäume, nackten Boden darbot, ein unfreundliches Bild des Winters.

Hainburg mit dem alten Schloß auf dem Bergrücken und noch weiter hinab die bedeutende königliche Freistadt Preßburg nehmen sich recht artig aus.

In drei Stunden erreichten wir letztere und landeten in der Nähe des Krönungsberges, einer künstlichen Erhöhung am Ufer der Donau, wohin der König nach der Krönung im feierlichen Ornat mit dem Schwert in der Hand reiten und dasselbe gegen Osten, Westen, Süden und Norden schwingen muß, zum Zeichen, daß er das Königreich gegen alle Feinde, woher sie immer kommen mögen, verteidigen wolle. Unweit von diesem Hügel ist der schöne Gasthof »Zu den drei grünen Bäumen«, wo es so teuer, ja noch teurer wie in Wien ist. Stromabwärts darf man bis unter Pest nicht auf dem Schiff übernachten.

23. März 1842

Heute fuhren wir um sechs Uhr morgens ab. Gleich unterhalb Preßburg teilt sich die Donau in zwei Arme und bildet die sehr fruchtbare Insel Schütt, welche zehn Meilen lang und sechs Meilen breit ist. Die Gegend bis Gran ist ziemlich einförmig, dann aber wird sie hübscher. Schöne Hügel und mehrere Berge umschließen dieselbe und bringen Abwechslung ins Gemälde.

Gegen sieben Uhr abends kamen wir in Pest an. Schade, daß es schon ganz finster war. Die wunderschönen Häuser, man könnte sagen Paläste, welche das linke Ufer der Donau zieren, sowie gegenüber die altertümliche und berühmte Festung und Stadt Ofen gewähren einen herrlichen Anblick und verdienen einen längeren Aufenthalt. Ich blieb nur über Nacht, weil ich schon einige Jahre früher mehrere Tage in Pest zugebracht hatte.

Da hier das Dampfschiff gewechselt wird, muß man beim Landen vorzüglich auf jenes Gepäck achthaben, welches man zu Wien nicht im Bureau übergeben hat.

Ich stieg im Gasthof »Zum Jägerhorn« ab. Es ist ein höchst eleganter Ort, aber unverschämt teuer. Ein kleines Stübchen im Hof kostete über Nacht vierundfünfzig Kreuzer.

Schon diesen ganzen Tag war mir sehr unwohl. Heftige Kopfschmerzen, Fieberschauer und wiederholtes Erbrechen ließen mich eine Krankheit und Unterbrechung meiner Reise befürchten. Wahrscheinlich waren diese Übelkeiten eine Folge des schmerzlichen Abschieds von geliebten Freunden und der Veränderung

der Luft. Ich konnte nur mühsam mein bescheidenes Kämmerchen erreichen und legte mich gleich zu Bett. Doch glücklich besiegte meine gute Natur all diese Feinde der Gesundheit, und ziemlich erholt begab ich mich des folgenden Tages am

24. März 1842

auf unser neues Dampfschiff, die »Galatha« von sechzig Pferdekräften, welche mir aber nicht so nett und niedlich vorkam wie die »Marianne«, die uns von Wien nach Pest geführt hatte. Unsere Reise ging schnell vonstatten, denn schon um zehn Uhr morgens waren wir bei Földvár, welches sich von fern groß und schön ausnimmt, in der Nähe aber gleich einer Seifenblase in nichts zerfließt. Um zwei Uhr kamen wir nach Paks. Hier und an allen wichtigen Orten wird ein Viertelstündchen haltgemacht. Ein Kahn rudert vom Land her, bringt und holt Menschen mit einer solchen bewundernswürdigen Schnelligkeit, daß man kaum die an den Nachbarn gerichtete Rede vollenden kann. Man hat nicht Zeit, sich Lebewohl zu sagen.

Um acht Uhr abends erreichten wir den durch zwei Schlachten berühmten Marktflecken Mohács. Die Festung daselbst wird als Gefängnis für Verbrecher benützt. Wir sahen weder Festung noch Ortschaft. Es war finstere Nacht, als wir ankamen, und um zwei Uhr morgens, den

25. März 1842

lichteten wir schon die Anker. Man versicherte mir, daß ich dadurch nichts verloren hätte.

Nach einigen Stunden bekam unser Schiff plötzlich einen so gewaltigen Stoß, daß alles auf das Verdeck eilte, um nach der Ursache zu sehen. Unser Steuermann hatte vermutlich mehr Schlaf als Sehkraft im Auge, gab dem Schiff eine ungeschickte Wendung, und ein Rad blieb nach dem Verlust mehrerer Schaufeln an vorstehenden, über das Wasser ragenden Baumpflöcken hängen. Schnell eilten die Matrosen in die Boote, das Schiff wurde rückwärts geleitet, und so gelang es mit vieler Mühe, uns wieder flottzumachen.

Wir hielten auf Augenblicke zu Dalina und kamen gegen zwei Uhr an der herrlichen und großartigen Ruine des Stammschlusses der Grafen Palffy vorüber. Noch schöner nimmt sich das dem Fürsten Odescalchi gehörige Schloß Illok aus, welches auf einem Berge liegt.

Um vier Uhr landeten wir bei Neusatz, der berühmten Festung Peterwardein gegenüber, deren bedeutende Werke auf einer weit in die Donau reichenden Felszunge liegen. Von dem Freistädtchen Neusatz ist nicht viel zu sehen, indem vorspringende, den Strom selbst beengende Hügel es dem Blick entziehen. Die Donau ist hier ziemlich zusammengeengt. Eine Schiffsbrücke verbindet beide Ufer. Hier fängt die Militärgrenze von Österreich an. Die Gegend ist sehr hübsch, besonders gut nimmt sich das Städtchen Karlowitz aus, das in einer kleinen Entfernung vom Ufer an artigen Hügeln, umpflanzt von Reben, liegt. Von diesem Punkt an wird aber die Gegend einförmiger bis Semlin. Die Donau breitet sich schon recht stattlich aus und gleicht oft mehr einem See als einem Fluß.

Um neun Uhr nachts erreichten wir die Stadt Semlin, an deren Ufer haltgemacht wurde. Semlin ist befestigt, liegt an der Einmündung der Save in die Donau, hat dreizehntausend Einwohner und ist die letzte österreichische Stadt am rechten Donauufer.

Als wir uns dieser Stadt näherten, wurden einige Böller auf unserem Schiff gelöst. Man berichtete dem Kellner zu spät davon, er hatte nicht mehr Zeit, die Fenster zu öffnen, und leider zersprang eines davon; ein unersetzlicher Schaden für uns, da die Gegend überall in Schnee gehüllt und die Temperatur auf Null gesunken war. Den Ofen hatte man schon in Wien von seinem Plätzchen verbannt, da die Sonne durch einige Tage ihre milden Strahlen ausgebreitet hatte und man vermessen auf ihre Beständigkeit rechnete. Überhaupt würde ich keinem Reisenden raten, den zweiten Platz auf einem Dampfschiff der Wiener Gesellschaft zu nehmen. Eine größere Unordnung wie da kann es nicht leicht geben. Wer nicht imstande ist, den ersten Platz zu zahlen, der bleibe nur gleich auf dem dritten, nämlich auf dem Verdeck, besonders wenn die Reise nicht weiter als bis Mohács geht. Ist das Wetter schön, so bleibt man ohnehin lieber im Freien, um das Panorama der Donau an sich vorübergleiten zu sehen. Ist das Wetter unfreundlich, so kann man ohne Anstand in die Kajüte des zweiten Platzes gehen, denn niemand ist aufmerksam auf die Reisenden des zweiten und dritten Platzes. Sie können sich wenigstens bei Tag sowohl auf dem Verdeck als unter demselben aufhalten. Von Pest abwärts sind die Frauen gezwungen, mit den Männern in einer Kajüte die Nacht zuzubringen. Dies ist unangenehm und auch ungeschicklich. Ich lernte später die Schiffe des österreichischen Lloyd und auch französische und italienische kennen und mit ihnen eine gute zweckmäßige Einteilung, Absonderung beider Geschlechter und keine Vernachlässigung des zweiten Platzes.

Die Kälte war so unendlich, daß man gern jede Öffnung geschlossen hätte, aber des vielen Tabakrauchens und der Ausdünstung all der armen Leute wegen, die einen großen Teil der Fracht im Ungarland ausmachen und bei der geringsten schlechten Witterung von ihrem gezahlten dritten Platz auf den zweiten eilen, hätte man gerne Tür und Fenster aufgerissen. Es ist gar nicht zu beschreiben, was man alles auf diesen Schiffen auszustehen hat. Ungepolsterte Bänke gehören bei Tag zum Sitzen, bei Nacht zum Schlafen. Von einem Waschbecken des Morgens ist keine Spur zu entdecken; und so ging es fort bis zum dritten Dampfschiff, dem »Zriny«, welches wir unterhalb der Donaufälle bestiegen; da fanden wir wenigstens bequeme gepolsterte Bänke. Allein auf keinem Schiff, selbst nicht auf dem »Ferdinand«, mit welchem man schon in das Schwarze Meer kommt und der fatalen Seekrankheit anheimfällt, ist eine Absonderung von Männern und Frauen.

Ich sollte doch glauben, daß man für die hohen Preise dieser Fahrt auf etwas Besseres Anspruch machen könnte. Der erste Platz bis Konstantinopel kostet ohne Zehrung und mit Ausnahme der Nachtlager in Preßburg und Pest hundertzwanzig Gulden, der zweite Platz fünfundachtzig Gulden.

26. März 1842

Die verfllossene Nacht war für uns Reisende keine Nacht der Ruhe, sondern des Lärmens. Niemand konnte die Augen schließen.

Semlin ist ein bedeutender Ladeplatz; es wurden über hundertachtzig Zentner Waren ab- und dagegen Steinkohlen, Holz und wieder Waren aufgeladen. Die zerbrochene Maschine wurde zurechtgemacht, und dies alles geschah mit einem solchen Lärm und Gepolter, daß man glaubte, das ganze Gebäude müsse über uns zusammenbrechen. Dazu Kälte und Wind, die durch die gebrochene Scheibe ihren beständigen Eingang hielten und uns die Nacht zu einer wahren Höllenqual machten. Um sechs Uhr morgens wurden wir endlich flott. Diesem zufälligen Aufenthalt hatten wir das Glück zu danken, Belgrad, welches der Stadt Semlin gegenüber liegt, die erste türkische Festung und Stadt in Serbien mit 29 000 Einwohnern, sehr gut zu sehen.

Die Lage von Belgrad ist sehr schön. Die Festungswerke ziehen sich vom Ufer der Donau längs eines Berges stufenweise hinauf. Die Stadt mit ihren schlanken Minaretten liegt eine Viertelstunde rückwärts. Hier sah ich die ersten Moscheen und Minarette. Die Moscheen, welche ich von Bord aus sehen konnte, haben ungefähr die Form eines runden, nicht sehr hohen Gebäudes und sind mit einer Kuppel gedeckt, an welche sich ein bis zwei schlanke Minarette, eine Art hoher, runder Säulen schließen. Nun wird die Fahrt teilweise sehr interessant und reich an Abwechslung und an schönen, wie durch einen Zauber malerisch vorübergleitenden Bildern. Die Gebirge beengen den Strom, bis er sich frei und fessellos in der Nähe von Pancsova wieder zu einer Breite von achthundert Klaffern ausdehnt.

Von dem Innern der Städte und der meisten Orte, die man berührt, ist wenig zu sehen, weil nur auf Augenblicke angehalten wird. Da läuft und drängt sich alles durcheinander, die Glocke läutet plötzlich, die Bretter werden aufgezogen, und wer von den Fremden sich um einige Augenblicke verspätet hat, muß bis zur nächsten Station auf dem Schiff bleiben.

In Neusatz geschah dies einem Bedienten, der die Effekten seiner Herrschaft nicht gleich auf das Verdeck warf, sondern sie erst in die Kajüte trug. Der Ärmste mußte bis Semlin mitfahren, um dann eineinhalb Tage zu Fuß nach Hause zu wandern.

Von Pancsova aus erreichten wir, nach einer zweistündigen, äußerst angenehmen Fahrt, die türkische Festung Semendria, die eine wahrhaft schöne Lage hat. Besonders verleihen ihr die vielen Spitzen und Zacken ihrer im maurischen Stil gebauten Wälle und Türme einen eigentümlichen Reiz. Überhaupt zeichnen sich die türkischen Festungen durch ihre schöne Lage aus.

Die Dörfer aber, ganz vorzüglich jene am rechten serbischen Ufer, gleichen an Ärmlichkeit jenen, deren ich leider so viele in Galizien sah: elende Hütten von Lehm, mit Stroh gedeckt, und weit und breit kein Baum und Strauch, der sowohl für das Auge des Reisenden als auch für den Bewohner selbst sehr wünschenswert wäre. Der arme Landmann könnte im Schatten seinem müden Körper einige Erholung gönnen, und dem Reisenden bliebe die Nacktheit und Armut solcher Wohnplätze, die doch jedes fühlende Herz mit Wehmut erfüllt, ein bißchen verborgen.

An dem linken Ufer, welches zu Ungarn gehört und das Banat heißt, ist es wohl nicht gar so arg, es bleibt aber auch da noch gar manches zu wünschen übrig, und man muß sich um so mehr über diese Armut wundern, da dieser Landstrich, überreich an Naturprodukten, die Getreidekammer Ungarns genannt wird.

Auf der österreichischen Seite der Donau sind, von zweihundert zu zweihundert Schritten, Grenzwachen aufgestellt, welche Einrichtung auch von den andern Regierungen an dem linken Ufer bis an die Mündung dieses Stromes in das Schwarze Meer beibehalten wird.

Man würde sich aber sehr irren, wenn man dächte, daß diese Soldaten in Uniform auf ihren Posten stünden. Sie beziehen in ihren erbärmlichen und zerrissenen Kleidern, oft mit nackten Füßen, ihre Station auf acht Tage. Ihre Hütten gleichen einem Stall. Ich trat in einige, um die innere Einrichtung zu sehen, die unmöglich einfacher sein könnte. In der einen Ecke befindet sich eine Feuerstelle, in der andern ein seinsollender Ofen von Lehm zusammengestoppelt. Eine unförmliche Öffnung in der Wand, anstelle des Glases mit Papier überklebt, bildet das Fenster, eine hölzerne Bank die Einrichtung. Was der Bewohner während dieser Zeit zum Unterhalt des Lebens bedarf, muß er sich mitbringen. Dafür erhält er von der Regierung Grund und Boden.

Auf dem russischen Gebiet haben die Soldaten wenigstens Uniformen an.

Immer schöner und reizender wird nun die Reise. Der große mächtige Strom eilt oft brausend und schäumend an hohen Bergen dahin, die ihm kaum einen Ausweg zu gestatten scheinen. Bald bespült er wieder freundlich und ruhig die ihn umgebenden Ufer. Jede Wendung zeigt neue Schönheiten; man weiß nicht, auf welche Seite man das begierige Auge wenden soll. Und stolz und majestätisch beherrscht ihn das Schiff, das sicher und schnell durch die wildromantischen Gegenden dahineilt.

Gegen ein Uhr mittags kamen wir nach Pasiest. Hier ist weiter nichts als ein großer Vorrat von Steinkohlen für die Dampfschiffe, nebst einigen Hütten. Vom Städtchen selbst ist nichts zu sehen.

Eine Stunde unterhalb Pasiest gewährt der mitten aus den Fluten sich erhebende, einzeln stehende Fels Babakai einen imposanten Anblick. Mit ihm vereint sich die am serbischen Ufer lehrende Ruvia Golubac zu einer wunderschönen Landschaft.

27. März 1842

Daß doch nie alles vereint sein kann. Jetzt befinden wir uns in den schönsten Gegenden und hofften hier Entschädigung zu finden für die vielen Unannehmlichkeiten, mit denen wir bisher zu kämpfen hatten, allein der Himmel begünstigt uns nicht. Die Witterung ist häßlich, und Schneegestöber treibt uns in die Kajüte. Der Sturm bringt die Donau dermaßen in Aufruhr, daß sie, dem Meere gleich, Wellen wirft. Wir leiden ungemein von Kälte und können uns nirgends erwärmen und betrachten wehmütig die Stelle, wo einst – ein Ofen stand.

Um vier Uhr gelangten wir zwar ohne Unglück, aber ganz erstarrt zu Drencova an und eilten in das von der Dampfschiffahrtsgesellschaft errichtete Gasthaus, wo wir einen gutgeheizten Saal, vortreffliche Kost und ein ziemlich bequemes Lager fanden. Dies war seit Pest der erste Ort, an welchem man sich erholen und erwärmen konnte.

In Drencova ist nichts als dieses Gasthaus und unweit davon ein Wachtposten. Am Ufer zeigte man uns die Barke, welche im Jahre 1839 mit den Reisenden verunglückte, als sie aufwärts der Donaufälle fuhren. Acht Personen, die in der Kajüte saßen, verloren dabei ihr Leben, und nur jene, die außen waren, wurden gerettet.

28. März 1842

Frühmorgens bestiegen wir die mit einer Kajüte versehene Barke »Tünthe«.

Die Donau wird von Felsen und Bergen immer mehr zusammengedrängt, so daß ihre Breite zwischen Drencova und Fetislav an manchen Stellen nicht über achtzig Klafter beträgt und sie mit doppelter Eilfertigkeit ihrem nahen Ziele, dem Pontus Euxinus, zuströmt.

Der Donaufälle wegen, die zwischen Drencova und Fetislav zu passieren sind, muß man das Dampfschiff mit einer Barke vertauschen. Die Hinabfahrt des Schiffes würde ohne Gefahr stattfinden, allein die Rückkehr desselben wäre mit großen Schwierigkeiten verbunden. Darum bleiben die Dampfschiffe in Drencova zurück, und man befördert die Reisenden stromabwärts in Barken und seit dem Unglücksfall vom Jahre 1839 stromaufwärts in bequemen guten Wägen.

Die Kälte war heute so unendlich wie gestern, und wenn einer der Reisenden nicht so gefällig gewesen wäre, mir seine Bunda (großer ungarischer Pelz) zu leihen, hätte ich in der kleinen Kajüte sitzen bleiben müssen und würde die interessantesten Stellen der Donau nicht gesehen haben. So aber hüllte ich mich vom Kopf bis zum Fuß tüchtig in den Pelz ein, setzte mich außerhalb der Kajüte auf eine Bank und konnte mit voller Muße die herrlich abgeschlossenen Bilder des Stromes, einer Kette von Seen ähnlich, in mein Gedächtnis aufnehmen. Beinahe bis Alt-Orsova blieben diese Ansichten gleich pittoresk.

Eine Stunde unterhalb Drencova, bei Islaz, riefen uns die Schiffer plötzlich zu: »Der erste Fall!« Gespannt vor Erwartung blickte ich vor. Das Wasser warf kleine Wellen, strömte etwas heftiger und verursachte ein leises Brausen. Wenn man mir es nicht gesagt hätte, daß die Donau hier einen Fall bildet, würde ich es nicht geahnt haben. Ich fand die Klippen und die Gewalt des Stromes zwischen Linz und Krems nicht viel unbedeutender. Freilich hatten wir großen Wasserstand, und da ist die Gefahr nicht halb so groß und das Ganze nicht so schaudererregend anzusehen. Die vielen Felsenzacken, die bei niederem Wasserstand überall drohend hervorblicken und durch die sich der Schiffer mit großer Kunstfertigkeit durchwinden muß, waren unseren Augen verborgen. Wir glitten unbeschädigt darüber weg, und nach ungefähr zwanzig Minuten hatten wir den ersten Fall im Rücken. Die beiden folgenden Fälle sind noch unbedeutender.

Auf der österreichisch-walachischen Seite zieht sich längs des Ufers eine sechs bis acht Stunden lange Straße, die oft von Mauerwerk unterstützt, oft auch den Felsen abgerungen ist. In der Mitte dieses Weges sieht man hoch oben in einer Felsenwand die berühmte Veteranihöhle, eine der unbezwingbarsten Stellen an der Donau, die es gibt. Sie ist mit etwas Schanzwerk umgeben und ganz geeignet, die Durchfahrt auf dem Strom zu sperren. Sie soll so geräumig sein, daß fünfhundert Mann Platz darin haben. Schon zu Zeiten der Römer wurde sie als Verteidigungspunkt der Donau benützt. Dritthalb Stunden abwärts von dieser Höhle sieht man die Trajanstafel, welche in den vorspringenden Felsen eingehauen ist.

Auf der türkisch-serbischen Seite erstrecken sich die Felsenmassen so nahe und tief in den Strom, daß an den meisten Stellen nicht Raum für einen Fußweg ist. Hier war die berühmte Trajansstraße. Man sieht weiter nichts mehr davon als längs des Stromes auf einer Strecke von vier bis fünf Meilen hin und wieder Löcher in die Felsen gehauen, worin starke Stämme eingelassen waren, auf denen einst Bretter lagen, welche die Straße gebildet haben sollen.

Um elf Uhr vormittags kamen wir zu Alt-Orsova an, der letzten Stadt Österreichs im Banater oder Walachischen Militärgrenzbezirk. Hier mußten wir den übrigen halben Tag bleiben.

Die Stadt nimmt sich ziemlich gut aus, sie hat hübsche, meistens neue Häuser. Besonders groß und schön ist jenes der Dampfschiffahrtsgesellschaft. Es gehört aber nicht zur Unterbringung der Reisenden wie zu Drencova. Hier, wie in Preßburg und Pest, muß der Reisende abermals die Kosten des Nachtquartiers zahlen; eine Einrichtung, die ich etwas sonderbar finde, da auf diese Art der Reisende doppelt zahlt, nämlich für seinen Platz auf dem Schiff und für jenen im Gasthof.

Es war gerade Sonntag, als wir ankamen, und ich sah viele Leute in die Kirche gehen. Die Bauern sind ziemlich gut und nett gekleidet. Männer und Weiber tragen lange, blautuchene Röcke. Die Weiber haben um die Köpfe große, weiße, leinene Tücher geschlagen, die hinten lang hinabhängen, und an den Füßen derbe Stiefel; die Männer runde Filzhüte und Sandalen von Baumrinde.

29. März 1842

Nachdem wir uns in dem guten Gasthof »Zum goldenen Hirschen« vollkommen erholt hatten, bestiegen wir heute morgen abermals eine neue Barke, den »Saturnus«, der nur oben gedeckt und von allen Seiten offen ist.

Sobald man diese Barke betritt, wird man schon für unrein, d. h. für halb verpestet, angesehen und darf nicht mehr an Land, ohne Quarantäne zu halten; auch begleitete uns ein Guardian bis Gallatz.

Gleich unterhalb Alt-Orsova verläßt man Österreichs Grund und Boden gänzlich.

Nach einer halben Stunde kommt man an der Festung Neu-Orsova vorüber, welche auf einer Insel liegt und eher den Namen einer Ruine verdient. Gegenüber liegt das Fort Elisabetha, bestehend aus einem Turm und mehrerem Mauerwerk, das sich längs des Berges hinanzieht. Dieses Fort liegt an einem der herrlichsten Punkte der Donau.

Nun nähert man sich immer mehr und mehr der gefährlichsten Stelle dieses Stromes, dem Eisernen Tor, von den Türken Demir-Kapu genannt. Wohl eine halbe Stunde vorher verkündet schon das Rauschen des Wassers den gefürchteten Ort. Viele Felsenriffe durchziehen den Strom und bilden eine Menge Wirbel. Diese gefährliche Strecke legten wir in fünfzehn Minuten zurück. Der große Wasserstand half uns ebenso glücklich über das Eiserne Tor wie vorher über die Fälle.

Ich fand diese Fälle tief unter meiner Erwartung und beinahe alles lang nicht den oft so poetisch schönen Beschreibungen entsprechend. Ich schildere alles, wie ich es finde, wie es meinen Augen erschien, ungeschmückt, aber wahr.

Am Ende des Eisernen Tores kommt man an einem Dorf vorüber, in dessen Nähe bei niederem Wasserstand einige Pfeiler der Trajansbrücke zu sehen sind.

Die Gegend fängt an flacher zu werden, besonders am linken Ufer, wo sich die ungeheure Ebene der Walachei ausbreitet und dem Auge nirgends ein Anhaltspunkt geboten wird. Rechts ziehen sich mehrere Staffagen von Hügeln und Bergen hin, deren Hintergrund die feingebildeten Linien des Balkans, der durch den Übergang der Russen im Jahr 1829 berühmt ist, schließen. Die Dörfer, deren man höchst selten einige an den Ufern sieht, werden immer erbärmlicher. Sie gleichen mehr Ställen für Vieh als menschlichen Wohnungen. Das Vieh kumpelt im Freien, obwohl das Klima nicht viel milder sein mag wie bei uns in Österreich, denn heute, so nahe an April, hatten wir einen Grad Kälte und gestern fünf Grad Wärme nach Réaumur.

Merkwürdig ist es auch, auf welcher schnellen und einfachen Art das Vieh in diesen Gegenden frei von der Pest erklärt wird. Wenn die Tiere von einem unreinen Ort zu Schiff in die Nähe eines gesunden kommen, so muß das Schiff ungefähr vierzig bis fünfzig Schritte vom Ufer entfernt halten, dann wird jedes Stück in das Wasser geworfen und an das Ufer getrieben, wo schon Leute harren, sie zu empfangen. Nach dieser einfachen Operation sind sie vom Peststoff befreit.

Der Viehbestand scheint in diesen Gegenden sehr bedeutend zu sein. Überall sieht man große Herden Hornvieh, darunter sehr viel Büffel. Ebenso sieht man ganze Scharen Ziegen und Schafe.

Auf dem »Saturnus« fuhren wir höchstens zwei Stunden und bestiegen sodann gegenüber der Festung Fetislav das Dampfschiff »Zriny«.

Um fünf Uhr abends kamen wir an der Festung Vidin vorüber und hielten gegenüber in der Nähe des Ortes Calafat. Hier sollten nur Waren abgeladen und sollte gleich wieder weitergefahren werden, allein der Agent war nirgends zu finden, und so mußten wir Reisende das Opfer dieser Fahrlässigkeit sein und hier über Nacht vor Anker bleiben.

30. März 1842

Noch immer war der Agent nicht zum Vorschein gekommen, und es blieb dem Kapitän nichts anders übrig, als den Oberkellner als Wache bei den Waren zurückzulassen. Um halb sieben Uhr früh wurde die Maschine endlich in Bewegung gesetzt, und nach einer schönen angenehmen Fahrt von sechs Stunden erreichten wir Nicopol.

Alle türkischen Festungen liegen am rechten Ufer meistens in schönen Gegenden. Die größeren Städte und Ortschaften sind umgeben von Gärten und Bäumen, welche ihnen ein gar freundliches Ansehen gewähren. Das Innere derselben soll freilich dem Äußern nicht entsprechen. Schmutzige, wirklich enge Gassen, baufällige Häuser und verbotene sollen dem Fremdling überall störend entgegenreten. Wir landeten bei keiner der Festungen und Städte, für uns war das rechtseitige Ufer das verbotene Paradies, und somit blieb uns das Schöne schön, die Enttäuschung ward uns nicht zuteil.

Ziemlich spät warfen wir Anker in der Nähe eines unbedeutenden Ortes.

31. März 1842

Früh morgens wurde abgefahren, und so kamen wir um acht Uhr schon nach Giurgiu.

Diese Stadt liegt am linken Ufer, der Festung Rustschuk gegenüber. Sie zählt sechzehntausend Einwohner und ist ein Hauptstapelplatz der Walachei. Wir mußten bis vier Uhr nachmittags hier verweilen, denn es wurden über sechshundert Zentner Waren nebst acht Wagen abgeladen und Steinkohlen dagegen eingenommen, und hatten daher Muße, das Innere dieser walachischen Stadt in Augenschein zu nehmen.

Doch wie wurden meine Reisegefährten von der Häßlichkeit dieser von außen so vielversprechenden Stadt unangenehm überrascht! Auf mich machte sie nicht halb diesen Eindruck, weil ich dergleichen noch von Galizien her im Gedächtnis hatte. Die Gassen und Plätze sind voll Gruben und Löcher, die Häuser ohne den geringsten Geschmack, ohne Symmetrie ausgeführt; das eine stand in die halbe Gasse hinein, das andere wieder ganz zurück usw. An einigen Orten zogen sich an beiden Seiten hölzerne Buden mit den gemeinsten Lebensmitteln oder sonstigen Bedürfnissen versehen hin, und dies nannte man den Bazar. Die Neugierde zog uns in ein Wein- und in ein Kaffeehaus. In beiden fanden wir nichts als hölzerne Tische und Bänke, beinahe keine Gäste, und diese wenigen der ärmsten Klasse angehörig. Gläser und Tassen werden den Gästen gereicht, ohne sie vorher auszuspülen.

Wir kauften Eier und Butter und gingen in ein Bürgerhaus, um uns ein Gericht nach deutscher Art zuzurichten. Bei dieser Gelegenheit sah ich auch die innere Beschaffenheit eines solchen Hauses. Der Boden des Zimmers war nicht gedeilt, die Fenster nur zur Hälfte mit Glas, der andere Teil entweder mit Papier oder

feiner Blase überklebt. Übrigens war alles nett und einfach eingerichtet. Ja sogar ein recht bequemer guter Diwan fehlte nicht. Um vier Uhr verließen wir diese Stadt.

Nun wird die Donau nur auf kurze Strecken breit. Sie ist mit Inseln wie besät und deshalb immer mehr getrennt als vereint.

In den Ortschaften sieht man schon griechische und türkische Trachten, jedoch sind die Frauen und Mädchen noch unverschleiert.

An der Festung Silistra kamen wir leider sehr spät vorüber und konnten sie nicht mehr sehen. Unweit davon blieben wir am linken Ufer über Nacht.

Den 1. April 1842

kamen wir zeitig an Hârsova vorüber, und um zwei Uhr hielten wir bei Braila, einer Festung, welche die Russen seit dem Jahre 1828 im Besitz haben. Hier wollte man die Reisenden nicht an Land steigen lassen, weil man sie für verpestet hielt, allein unser Guardian trat hervor und gab Zeugnis, daß weder am rechtsseitigen Ufer gelandet noch von dort jemand aufgenommen worden sei; darauf durften die Ankömmlinge das feste Land betreten.

Um vier Uhr lagen wir vor Galatz, einer der bedeutendsten Handelsstädte mit achttausend Einwohnern und dem einzigen Hafen der Russen an der Donau. Hier sahen wir die ersten Kauffahrer, Segelschiffe und Barken aller Art, die aus dem Schwarzen Meer kommen. Auch Möwen, die Verkündiger des nahen Meeres, schwirrten über unsern Köpfen.

Es geht hier schon äußerst bunt und lebhaft zu, denn Galatz ist der Sammelplatz von Kaufleuten und Reisenden aus zwei Weltteilen, aus Europa und Asien; es ist der Vereinigungspunkt von drei der größten Monarchien: Österreich, Rußland und der Türkei.

Nachdem der Guardian auch hier dieselben Versicherungen wiederholt hatte wie zu Braila, durften wir das Schiff verlassen. Ich hatte einen Empfehlungsbrief an den österreichischen Konsul, welcher mich nach Übergabe meines Briefes sehr freundlich empfing und für meine Unterkunft auf das gefälligste sorgte.

Die Stadt verspricht viel, aber man sieht ein ebenso schmutziges, erbärmliches Nest wie Giurgiu. Die meisten Häuser sind aus Holz oder Lehm und mit Stroh gedeckt; nur jene der Konsuln und reichen Kaufleute sind aus Stein. Die schönsten Gebäude sind die christliche Kirche und der moldauische Gasthof.

Obwohl Galatz an der Donau liegt, kommt den Einwohnern das Trinkwasser dennoch sehr teuer. Es gibt weder Brunnen in den Häusern noch auf den Plätzen. Die Leute müssen sich alles Wasser von der Donau zutragen und -führen lassen, was eine bedeutende Beschwerde für die Armen und eine ziemliche Ausgabe für die Wohlhabenden ist, da im Winter an den entfernteren Gegenden der Stadt für ein Faßchen Wasser von zwei Eimern zehn bis zwölf Kreuzer bezahlt werden müssen. Man begegnet beständig an allen Orten und Ecken nichts als Wasserträgern und Wägelchen mit Wasserfässern. Schon öfters hatte man Versuche gemacht, nach diesem unentbehrlichen Element zu graben; es kam zwar zum Vorschein, aber leider ungenießbar, da es salzig schmeckte.

In Galatz wird vierundzwanzig Stunden haltgemacht; ein Aufenthalt, der eben nicht zu den angenehmsten gehört, da weder Stadt noch Umgebung etwas Sehenswertes bieten. Und dennoch werde ich immer mit Vergnügen und Dankbarkeit an diesen Tag denken. Der Herr Konsul H. ist ein gebildeter und gefälliger Mann, der mir, da er selbst viel gereist ist, manchen Rat und manche Verhaltensregel mit auf die Reise gab. Die Ruhe, Ordnung und Bequemlichkeit, welche ich in seinem Hause fand, war nach einer Reihe so vieler Tage der Entbehrungen eben auch nicht zu verwerfen, und so fand ich hier Erholung für Geist und Körper.

2. April 1842

Die Gegend um die Stadt ist so wenig einladend, daß ich gar keine Lust bekam, einen Spaziergang dahin zu machen. Ich blieb also in der Stadt und ging in den holprigen Gassen bergauf und bergab. Kaffeehäuser gibt es hier schon eine Menge, wenn aber die Menschen nicht vor denselben saßen, Kaffee trinkend und Tabak rauchend, so würde man diesen schmutzigen Stuben schwerlich die Ehre antun, sie für solche zu halten.

Auf dem Markt und an den Plätzen sieht man bedeutend weniger Frauen als Männer: letztere tummeln sich überall umher und besorgen zum Teil gleich den Italienern auch die Geschäfte des anderen Geschlechtes. Man sieht ein Gemisch der verschiedenartigsten Nationen.

Der Bazar ist überhäuft mit Südfrüchten aller Art. Orangen und Zitronen sind in solcher Menge vorhanden wie bei uns das gemeinste Obst. Natürlich ist auch der Preis dafür sehr gering. Ganz besonders schön ist der Blumenkohl, der aus Kleinasien gebracht wird. Man findet viele Stücke darunter von der Größe eines Mannskopfes.

Abends mußte ich mich wieder nach dem Hafen begeben, um mich einzuschiffen.

Von dem Wirrwarr, der hier herrscht, kann man sich keinen Begriff machen. Ein hölzernes Geländer ist die Scheidewand zwischen den Gesunden und jenen, welche aus einem Land der Pest kommen oder in dasselbe gehen. Wer diese Grenze überschreitet, darf nicht mehr zurück. Soldaten, Offiziere, Beamte und Aufseher, letztere mit Stöcken und Zangen bewaffnet, stehen am Eingang, um jene, die sich mit Worten nicht abfertigen lassen, mit Gewalt zurückzutreiben. Die Lebensmittel oder sonstigen Effekten werden zum Teil hinübergeworfen oder an die Grenze gestellt, dürfen aber erst berührt werden, wenn sich die Überbringer davon entfernt haben. Ein Herr auf der verpesteten Seite wollte jemandem auf der anderen einen Brief geben; augenblicklich riß man ihm denselben aus der Hand und reichte ihn mittels einer Zange hinüber. Und dabei ist beständig ein solcher Lärm und ein solches Geschrei, daß man sein eigenes Wort kaum hört. Der eine ruft: »Langen Sie mir mein Gepäck herüber«, der andere: »Ach, kommen Sie mir nicht in die Nähe! Rühren Sie mich ja nicht an!« Dazwischen schreien wieder die Aufseher: »Zurück! Zurück!« usw.

Im ganzen kommt mir diese ängstliche Vorsicht doch gar zu übertrieben vor, besonders zu einer Zeit, wo in der Türkei weder die Pest noch sonst eine ansteckende Krankheit herrscht. Einer unserer Reisegefährten wurde schon den vorhergehenden Tag auf unser künftiges Schiff verbannt, weil er das Unglück hatte, an einen Guardian zu streifen, als er nach seinen Effekten sehen wollte.

Um sieben Uhr ertönt der Zapfenstreich, das Gitter wird geschlossen, und die Komödie hat ein Ende. Wir begaben uns nun auf das vierte und letzte Dampfschiff, auf den »Ferdinand«. Im ganzen werden von Wien bis Konstantinopel die Fahrzeuge sechsmal gewechselt, viermal die Dampfschiffe und zweimal die Barken, was eben nicht zu den Annehmlichkeiten der Donaureise gehört.

Der »Ferdinand« ist kein großes, aber ein starkes und bequem gebautes Schiff. Sogar die Kajüte des zweiten Platzes ist nett, und ein niedliches Öfchen verbreitete, da wir selten mehr als sechs bis acht Grad über Null hatten, eine sehr wohlthuende Wärme. Eine besondere Abteilung für Frauen ist leider auf dem zweiten Platz nicht vorhanden, jedoch wird wenigstens darauf gesehen, daß von dem dritten Platz niemand auf den zweiten darf. An den Wänden laufen ringsherum zwölf Schlafstellen, und vor denselben befinden sich gut gepolsterte breite Bänke.

3. April 1842

Um fünf Uhr morgens fuhren wir aus dem Hafen von Galatz. Etwas später wurden uns Waschbecken und Handtücher gereicht, eine Sache, die man auf den früheren Schiffen gar nicht kannte. Für die Verpflegung, welche ziemlich gut ist, zahlt man des Tages einen Gulden dreißig.

Gegen zehn Uhr gelangten wir an ein bessarabisches, sehr erbärmlich aussehendes Nest, Tchussu, wo eine Viertelstunde angelegt wurde, dann ging es unausgesetzt dem Meer zu.

Ich freute mich schon lang auf das Einlaufen in das Schwarze Meer und dachte mir die Donau in der Nähe dieser Stelle selbst einem Meer gleich. Wie es aber im Leben gewöhnlich geht – »große Erwartungen, kleine Erfolge« –, so war es auch hier. Bei Galatz ist die Donau sehr breit, aber eine geraume Strecke vor dem Ausfluß teilt sie sich in so viele Arme, daß eigentlich keiner majestätisch zu nennen ist.

Gegen drei Uhr nachmittags liefen wir endlich ins Schwarze Meer ein.

Da stürmen nun von allen Seiten die Arme der Donau heran und drängen mit Ungestüm das Meer so weit zurück, daß man nur in großer Ferne einen grünen Streifen desselben entdeckt. Über eine Stunde fährt man noch auf dem gelben, lehmigen, stark bewegten Süßwasser, bis man endlich die Grenze überschreitet und von den salzigen Meeresfluten getragen wird. Äquinoktialstürme und Unwetter trieben zu unserem Unglück ihr Unwesen noch so arg, daß ganze Ladungen des salzigen Elementes unser Verdeck überschütteten. Wir konnten uns kaum mehr auf dem Verdeck halten und gelangten nur mit Hilfe einiger Matrosen in die Kajüte, wohin uns ohnedies der Schall der Speiseglocke rief.

Einige der Reisenden, worunter auch ich gehörte, machten diesmal dem Koch wenig Ehre. Wir hatten kaum einige Löffel Suppe genossen, als uns das Seeübel so derb ergriff, daß wir nicht schnell genug vom Tisch eilen konnten. Ich legte mich nieder und war an diesem Tag nicht mehr imstande, mich zu bewegen und mich auf das Verdeck hinaufzuschleppen, um dies herrliche Schauspiel der Natur bewundern zu können. Die Wellen gingen oft so hoch, daß sie über der Heizröhre zusammenschlugen und uns von Zeit zu Zeit durch diese Öffnung ganze Ladungen Wasser in die Kajüte sandten.

4. April 1842

Der Sturm nahm von gestern auf heute bedeutend zu, so daß man sich in den Betten festhalten mußte, um nicht herausgeworfen zu werden. Einem der Reisenden geschah dieser Unfall, da er durch die großen Übelkeiten außerstande war, sich fest anzuklammern.

Da ich mich schon etwas besser fühlte, versuchte ich aufzustehen, wurde aber in demselben Augenblick mit solcher Gewalt an den gegenüberstehenden Tisch geschleudert, daß ich die Lust, einen abermaligen Versuch zu machen, auf lange verlor. An Schlaf war in der Nacht gar nicht zu denken. Das schreckliche Geheul des Windes in den Masten und Tauwerken, das furchtbare Gekrache des Schiffes, das aus seinen Fugen zu gehen schien, das ewige Hinundherlegen desselben, das Rollen der schweren Ankerketten über uns, das Rufen, Befehlen und Schreien des Kapitäns und der Matrosen, dieser vereinte, unaufhörliche Lärm gönnte uns keinen Augenblick Ruhe. Des Morgens schleppte ich mich noch halb krank mit Hilfe des Dieners hinauf auf das Verdeck in die Nähe des Steuermanns, um die wundervollste Szene der Natur, einen Seesturm, betrachten zu können.

Ich klammerte mich fest an und trotzte kühn den Wellen, die hoch über dem Schiff zusammenschlugen und mich von allen Seiten benetzten, als wollten sie die Hitze meiner Krankheit kühlen. Dafür bekam ich aber auch den klaren, deutlichen Begriff eines Sturmes auf dem Meer; ich sah die Wogen schäumend daherstürmen, sah das Schiff bald in den Abgrund tauchen, bald wieder mit Blitzesschnelle auf den höchsten Wellengipfeln sich erheben. Es war ein grauses, fürchterliches Bild, dessen Anblick mich so ergriff und beschäftigte, daß ich gänzlich mein Übelbefinden vergaß.

Erst spät in der Nacht ließ der Sturm etwas nach, so daß wir nun einlaufen und Anker werfen konnten im Hafen von Varna, den wir schon zehn bis zwölf Stunden früher hätten erreichen sollen.

5. April 1842

Heute morgen konnte ich diese schöne Festung und Stadt, die die Russen im Jahre 1828 belagert und eingenommen haben, mit Muße betrachten. Wir blieben daselbst mehrere Stunden. Der obere Raum des Schiffes wurde hier dermaßen mit Geflügel aller Art beladen, daß der Raum für uns Reisende höchst beschränkt war. Dieser Artikel scheint von Türken und Franken in Konstantinopel sehr gesucht zu sein, denn der Schiffskapitän versicherte mir, daß sie bei jeder Abfahrt von Varna mit dieser Ware vollgeladen nach Stambul führen.

6. April 1842

Der schönste Anblick der Welt, auf welchen ich mich schon bei meiner Abreise freute, die Fahrt durch den Bosphorus, wurde mir durch die Nacht entzogen. Erst einige Tage später machte ich diesen Ausflug auf einer Kaik, einem äußerst leicht und schmal gebauten Kahn, und genoß da in vollen Zügen Ansichten und Bilder, die ich nicht vermögend bin zu schildern.

(Quelle: Pfeiffer 1886)

11.1.4. Training: „drill before breakfast“ - Transkript

Dear Sir,

I have now been out in France a fortnight. I am getting on quite well in the army. Food is also good and plentiful. Of course, I find the life much rougher from that which I have been used to, but it is a fine experience all the same.

We were at Longmoor camp for a week where we were equipped and put through elementary drills. We slept in a big hut there and had a sack of straw and three blankets each. The huts, of which ours was one, were called Apple Pie and they are about a mile from the camp, so we had to be up about six every morning in order to march down for parade at the camp. The parade was at 7.15am. Some mornings we had an hour's drill before breakfast. It was about 11pm I should think, when we arrived at Longmoor on 31st March, and consequently we went straight to our hut. No light was provided so we had to fish about for our beds in the dark. Fortunately it was moonlight, and after some time a lantern was brought. We were only too glad when getting up time came the next morning. It was before six. We were marched down to the camp and were inspected by the Sergeant Major and after waiting some time went into breakfast. Roast sausage was the chief item on the menu but to our amazement no knives and forks were supplied for us...Well, the only thing was to convey the sausages to our mouths with fingers.

The nearest thing to a knife and fork that some had was a pair of nail scissors so these were used. Personally I did not know whether to feel shocked or to laugh at the situation; it ended in the latter. We Audit fellows got to one table and we had to have a good laugh before we could eat much. It was a hopeless situation to be in. Don't let this discourage anyone from coming but I warn him to bring a knife and fork with him. After the first day we were issued with kits and we

treasure our knives and forks now. We came across on the mail boat Hantonia from Southampton to Havre. ... We had to sleep in huts there on the docks but the sleeping accommodation was better. We Audit fellows were able to keep together till last Saturday night, when Pond and I were sent off to different railheads. He is at 3rd Corps Railhead. We had to travel in horseboxes and during night time. We travelled in this manner Saturday, Sunday and Monday nights. French shunting is severe and we were shaken up. Saturday and Sunday we were about 30 in the trucks, Monday 17.

We were fortunate enough to get out into every town at which we stopped during the day. On Tuesday night another young fellow and I were stranded at a station somewhere in France, and we were instructed by the military authorities to sleep in the goods shed. We had to sleep on the deck of the shed and shunting was going on during the night. This was about the limit, I thought.

I arrived finally at my destination about midday last Wednesday. Now I live in a truck at a small French station. There are two other Sappers and a Lieutenant here. We deal chiefly with gifts and comforts sent to the troops...

We are out in the country. I think this railhead is about seven miles from the firing line. Can hear the guns and see the star shells at night. I must ask you kindly to excuse me for writing in pencil because I have no pen and ink here. Would you mind letting this letter be treated as to the whole of the statistical department, please sir. With kindest regards to all. I am yours obediently,

Arthur Smith

(Quelle: SMITH 1915)

11.2. Anhang II: Python Code Geoparsing

Sofern nicht anders angegeben, stellt dieser Code die Umsetzung des Geoparsings für die Quelle „Q1_Brief_Weiss_1940“ dar. Änderungen, welche andere Quellen betreffen, sind rot markiert.

```
## Umsetzung mit Quelle Q1_Brief_Weiss_1940

## Laden der benötigten Packages

## Spacy für das NLP, zur Darstellung der getaggten Wörter und die benötigten Pipelines
import spacy #Package: https://spacy.io/ (11.08.2021)
from spacy import displacy
#import de_core_news_sm
import de_core_news_lg

#Quelle Q4_Brief_WW1_Training
import en_core_web_lg
#

##Importieren des Packages pprint für das Abspeichern als Liste
from pprint import pprint #Package: https://docs.python.org/3/library/pprint.html (11.08.2021)

## Geocoder um einen Abgleich mit dem Geonames Gazetteer durchzuführen und den Orten auch Koordinaten zuweisen zu können
import geocoder #Package: https://geocoder.readthedocs.io/ (11.08.2021)
```

```

## Geopandas wird benötigt, um einen Dataframe in einen Geodataframe umzu-
wandeln

import geopandas #Package: https://geopandas.org/ (11.08.2021)

## Einstellung der Workspace-Direction und des Filenamen

dir_q="C:/Users/andre/OneDrive/Dokumente/UNI/Master/S4_SS2021/SE Konversa-
torium zur Kartographie und Geoinformation (gilt als Privatissimum zur Mas-
terarbeit aus KGI)/Masterarbeit/praxis/texts/Q1_Brief_Weiss_1940/"

filename="Q1_Brief_Weiss_1940"

#Quelle Q2_NFP_24091886
dir_q="C:/Users/andre/OneDrive/Dokumente/UNI/Master/S4_SS2021/SE Konversa-
torium zur Kartographie und Geoinformation (gilt als Privatissimum zur Mas-
terarbeit aus KGI)/Masterarbeit/praxis/texts/Q8_OCR/"
filename="Zeitung_S2"
#

#Quelle Q3_Pfeiffer
dir_q6 = "C:/Users/andre/OneDrive/Dokumente/UNI/Master/S4_SS2021/SE Konver-
satorium zur Kartographie und Geoinformation (gilt als Privatissimum zur
Masterarbeit aus KGI)/Masterarbeit/praxis/texts/Q6_Pfeiffer/"
filename6 = "Q6_Pfeiffer_kap"

dir_q=dir_q6
filename=filename6
#

#Quelle Q4_Brief_WW1_Training
dir_q="C:/Users/andre/OneDrive/Dokumente/UNI/Master/S4_SS2021/SE Konversa-
torium zur Kartographie und Geoinformation (gilt als Privatissimum zur Mas-
terarbeit aus KGI)/Masterarbeit/praxis/texts/Q4_WW1_Training/"
filename="Q4_WW1_Training"
#

## Laden der Datei, welche geoparsed werden soll

file = open(dir_q+filename+".txt", "r", encoding='utf8')

## Speichern unter dem Namen q1

q1=file.read()

file.close()

print(q1)

#Quelle Q1_Brief_Weiss Gazetteer-basiertes Geoparsing

# Python code to convert text to list
q1 = q1.replace("\n"," ")
# Quelle: https://www.geeksforgeeks.org/python-program-convert-string-list/
(11.08.2021)
def Convert(string):
    li = list(string.split(" "))
    return li

```

```

# Driver code
q1_list = Convert(q1)
print(q1_list)
#

## Anwendung der Named Entity Recognition (mittels spacy) auf den Text

#Laden des trainierten deutschen Modells
#nlp = de_core_news_sm.load()
nlp = de_core_news_lg.load()

#Quelle Q4_Brief_WW1_Training
nlp = en_core_web_lg.load()
#

##Anwendung der spacy-Funktion nlp zur Extraktion von Entitäten auf die
Quelle
doc=nlp(q1)

##Darstellung des Textes mit getaggten Entitäten
displacy.render(doc, style="ent", jupyter=True)

## Testausgabe der erfassten Entitäten
#test=pprint([(X.text, X.label_) for X in doc.ents])

## Abspeichern der erfassten Entitäten in der Liste list_test
list_test=[]
for X in doc.ents:
    list_test.append([X.text, X.label_])

## Ausgabe der Liste list_test
list_test

## Abspeichern der Liste list_test als Dataframe df
df=pd.DataFrame(list_test, columns=['ents', 'types'])
df

## Extraktion der Räumlichen Entitäten (LOC und GPE) und Abspeichern als
Dataframe df_places
df_places=df.loc[df['types'].isin(['LOC', 'GPE'])]

## Extraktion der Entitäten (Namen der Orte) aus der Liste, Wegfall des
Typs, Abspeicherung als places_list
places_list = df_places['ents'].tolist()

```

```

places_list

## Ausdrucken als Liste mit Index, um die Entitäten einzeln Abfragen zu
können

a=0
for x in range(len(places_list)):
    print(a, places_list[x])
    a+=1

## Überprüfung einzelner Ergebnisse mittels Geocoding um zu sehen, ob dies
auch funktioniert

## Filterung auf Kontinent Europa und FeatureClasses A, H, L, P, R, T, V

g = geocoder.geonames(places_list[0],continentCode=["EU"], fea-
tureClass=["A","H","L","P","R","T","V"], key='XXX')

## Abfrage einzelner Attribute des codierten Ergebnisses
g.address
#g.geonames_id
#g.description
#g.population
#g.description
#g.country
#g.bbox
##Proximity to other placenames einfügen???

## Überprüfung der Koordinaten (Länge und Breite in Grad) für die Durchfüh-
rung der Verortung
print("Breite/Latitude = "+g.latlng[0], "\nLänge/Longitude = "+g.latlng[1])

##Erstellung eines Dataframes (df_geocoded) für Output, Spalten Name, Länge
und Breite
df_geocoded=pd.DataFrame(columns=['name', 'lon', "lat","original"])

##Hinzufügen der getaggten Entitäten (places_list) mit Schleife zu df_ge-
ocoded

for x in range(len(places_list)): ## Limitierung der Schleife auf die An-
zahl der Entiditäten in places_list
    ## Geocoding der einzelnen Entitäten (x) in places_list
    g = geocoder.geonames(places_list[x],continentCode=["EU"],fea-
tureClass=["A","H","L","P","R","T","V"], key='XXX')
    #Quelle Q1_Brief_Weiss Gazetteer-basiertes Geoparsing
    g = geocoder.osm(q1_list[x])

```

```

#

#Quellen Q2_NFP_24091886 und Q3_Pfeiffer
g = geocoder.geonames(places_list[x], featureClass=["A", "H", "L", "P", "R", "T", "V"], key='XXX')
#

print(g.address, g.feature_class) ## Ausgabe des gematchten Ortes und
des Typs zur Überprüfung
if g.address is None:
    print(places_list[x]+": Entity not geocoded!")
    ## Schreiben der Zeile (new_row, welche dann dem Dataframe hinzuge-
    fügt wird (Werte: Name, Länge, Breite)
else:
    new_row = {"name": g.address, "lat": g.latlng[0], "lon":
g.latlng[1], "original": places_list[x]}
    #continue
    ## Hinzufügen der erstellten Reihe zum Datframe df_geocoded
    df_geocoded = df_geocoded.append(new_row, ignore_index=True)
    print("New Row has been added-----")

## Ausgabe des fertiggestellten Dataframes
df_geocoded

## Änderung der Spalten lon und lat in Zahlenwerte (float)
df_geocoded["lat"]=df_geocoded["lat"].astype(float)
df_geocoded["lon"]=df_geocoded["lon"].astype(float)

## Ausgabe der Datentypen um zu überprüfen, ob Länge und Breite nun Zahlen-
werte sind
df_geocoded.info()

## Abspeicherung des Dataframes als csv-File
#df_geocoded.to_csv(dir_q+filename+"_table.csv")

## Umwandlung des Dataframes in einen Geodataframe mittels geopandas
gdf = geopandas.GeoDataFrame(df_geocoded, geometry=geopan-
das.points_from_xy(df_geocoded.lon, df_geocoded.lat), crs="epsg:4326")

## Überprüfung des Geodataframes
print(gdf.head())

##Export des Geodataframes als shapefile

```

```
gdf.to_file(dir_q+filename+".shp",driver="ESRI Shapefile")
```

```
## MIT License
```

```
Copyright (c) 2021 André Novak
```

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

(<https://choosealicense.com/licenses/mit/>, 23.08.2021)

11.3. Anhang III: Ergebnisse des Geoparsings, Tabellen

11.3.1. Q1_Brief_Weiss_1940

	name	lon	lat	original
0	Calais	1,85635	50,95194	Calais
1	Calais	1,85635	50,95194	Calais
2	Dunkirk	2,37681	51,0344	Dünkirchen
3	Calais	1,85635	50,95194	Calais
4	London	-0,12574	51,50853	England
5	Calais	1,85635	50,95194	Calais
6	London	-0,12574	51,50853	England
7	London	-0,12574	51,50853	England
8	London	-0,12574	51,50853	England
9	France	2	46	Frankreich
10	Hause	6,29332	62,24055	Hause
11	Guines	1,87025	50,86708	Guines
12	Calais	1,85635	50,95194	Calais
13	Antwerp	4,40026	51,22047	Antwerpen
14	Hause	6,29332	62,24055	Hause

11.3.2. Q2_NFP_24091886

	name	lon	lat	original
0	Berlin	13,41053	52,52437	Berlin
1	Berlin	13,41053	52,52437	Berlin
2	Belgium	4,5	50,75	Bl
3	St Peters-burg	30,31413	59,93863	Petersburg
4	Rome	12,51133	41,89193	Rom
5	Vienna	16,37208	48,20849	Wien
6	Barop	7,43197	51,48173	Barop
7	Serbia	20,45998	44,81892	Serbien
8	Tỉnh Yên Bái	104,58333	21,75	Bai
9	St Peters-burg	30,31413	59,93863	St. Petersburg
10	Moscow	37,61556	55,75222	Moskau
11	Kop'ung-gun	125,91778	40,56444	Kop
12	Turkey	35	39	Türkei
13	Egypt	30	27	Egypten
14	Italy	12,83333	42,83333	Italien
15	Milan	9,18951	45,46427	Mailand
16	Florence	11,24626	43,77925	Florenz
17	Rome	12,51133	41,89193	Rom
18	Paris	2,3488	48,85341	Paris
19	London	-0,12574	51,50853	London
20	Ái Giannis o Pórtos	25,21709	37,53635	Portos
21	Vienna	16,37208	48,20849	Wien
22	Vienna	16,37208	48,20849	Wien
23	Vienna	16,37208	48,20849	Wien
24	Marburg	8,77069	50,80904	Marburg
25	Massawa	39,47455	15,60811	Meeres
26	Basel	7,57327	47,55839	Stadt
27	Wiens Peak	-56,4	-83,983	Wiens
28	Trafalgar Square	151,30119	-33,6485	Trafalgar-Square
29	London	-0,12574	51,50853	London
30	Vienna	16,37208	48,20849	Wien
31	Greece	22	39	Griechenland
32	Schinder	11,86667	47,6	Oesterreicher
33	Germany	10,5	51,5	Deutschland
34	Ober Hohsand-joch	8,31152	46,40149	Ober-Italien
35	Vis	16,15167	43,045	Lissa
36	Italiener Straße	13,84133	46,60614	Italiener
37	Schinder	11,86667	47,6	Oesterreicher
38	Rapides des Italiens	-74,81286	46,36199	Italiens
39	Socotra	53,91678	12,49826	Sokotora
40	Ancona	13,50337	43,5942	Ancona
41	Vis	16,15167	43,045	Lissa

42	Italy	12,83333	42,83333	Italien
43	Austria	13,33333	47,33333	Oesterreich
44	Bohemia	14,5	50	Böhmen
45	Austria	13,33333	47,33333	Oesterreich
46	South Ame-rica	-57,65625	-	Amerika
			14,60485	
47	France	2	46	Frankreich
48	New Eng-land	-71	44	England
49	Mexico	-102	23	Mexico
50	Vienna	16,37208	48,20849	Wien
51	Wil	9,04552	47,46152	Wil
52	Heligoland	7,88458	54,18213	Helgoland
53	Heligoland	7,88458	54,18213	Helgoland
54	Vis	16,15167	43,045	Lissa
55	Miramar	-97,89997	22,36094	Miramar
56	Aden	45,03667	12,77944	Aden
57	Brazil	-55	-10	Brasilien
58	Brazil	-55	-10	Bra
59	Vienna	16,37208	48,20849	Wien
60	Vienna	16,37208	48,20849	Wien
61	Europe	9,14062	48,69096	Be
62	Germany	10,5	51,5	Deutschland
63	Großer Lor-cher Werth	7,81169	50,0366	Werth
64	Beiden Field	-	36,81328	Beiden
		119,75514		
65	France	2	46	Frankreich
66	Vienna	16,37208	48,20849	Wien
67	Austria	13,33333	47,33333	Oesterreich
68	Freiberg	13,33881	50,91089	Freiberg
69	Mlinište	17,61583	42,9925	Ministe
70	Freiberg	13,33881	50,91089	Freiberg
71	Freiberger Mulde	12,79856	51,16004	Freiberger
72	Freiberger Mulde	12,79856	51,16004	Freiberger
73	Gamprin	9,50995	47,21986	"Gemeinde
74	Freiberger Mulde	12,79856	51,16004	Freiberger
75	Gamprin	9,50995	47,21986	Gemeinde
76	Freiberg	13,33881	50,91089	Freiberg
77	Czerno	19,57045	51,2507	"Czerno-
78	Chernivtsi	25,93241	48,29045	Czernowitz
79	Chernivtsi	25,93241	48,29045	Czernowitz
80	Vis	16,15167	43,045	Lissa
81	Mexico	-102	23	Mexico
82	Mexico	-102	23	Mexico
83	Vis	16,15167	43,045	Lissa
84	Hellem	9,61667	63,8	Hellem
85	Pula	13,84806	44,86833	Pola

86	Heligoland	7,88458	54,18213	Helgoland
87	Vienna	16,37208	48,20849	Wien
88	Austria	13,33333	47,33333	Oesterreich
89	Kaliningrad	20,51095	54,70649	Preußen
90	Italy	12,83333	42,83333	Italien

91	San Salvador	-89,18718	13,68935	San
92	Vis	16,15167	43,045	Lissa
93	Pula	13,84806	44,86833	Pola
94	Graz	15,45	47,06667	Graz
95	Vis	16,15167	43,045	Lissa

11.3.3. Q3_Pfeiffer, Kapitel 1

	name	lon	lat	original
0	Europe	9,14062	48,69096	Europa
1	Prenzlauer Berg	13,42443	52,53878	Berg
2	Vienna	16,37208	48,20849	Wien
3	Kaisermühlen	16,425	48,22611	Kaisermühlen
4	Hungary	20	47	Ungarn
5	Berlin	13,41053	52,52437	Stadt
6	Bratislava	17,10674	48,14816	Preßburg
7	Bratislava	17,10674	48,14816	Preßburg
8	Neuburg an der Donau	11,18709	48,73218	Donau
9	Vienna	16,37208	48,20849	Wien
10	Bratislava	17,10674	48,14816	Preßburg
11	Bratislava	17,10674	48,14816	Preßburg
12	Bratislava	17,10674	48,14816	Preßburg
13	Neuburg an der Donau	11,18709	48,73218	Donau
14	Belgium	4,5	50,75	Königreich
15	Vienna	16,37208	48,20849	Wien
16	Bratislava	17,10674	48,14816	Preßburg
17	Neuburg an der Donau	11,18709	48,73218	Donau
18	Neuburg an der Donau	11,18709	48,73218	Donau
19	Las Palmas de Gran Canaria	- 15,41343	28,09973	Gran
20	Las Palmas de Gran Canaria	- 15,41343	28,09973	Gran
21	Pest megye	19,33333	47,41667	Pest
22	Neuburg an der Donau	11,18709	48,73218	Donau
23	Ofen	8,31544	46,8073	Ofen
24	Vienna	16,37208	48,20849	Wien
25	Jägerhorn	7,87723	45,95241	Jägerhorn
26	Vienna	16,37208	48,20849	Wien
27	Pest megye	19,33333	47,41667	Pest
28	Paks	18,85569	46,6221	Paks
29	Mohács	18,68621	45,9902	Mohács
30	Urochishche Glubokaya Dolina	25,77506	53,06641	Dalina
31	Urochishche Glubokaya Dolina	25,77506	53,06641	Dalina
32	Borgo Odescalchi	11,80296	42,08443	Odescalchi
33	Borgo Odescalchi	11,80296	42,08443	Odescalchi
34	Novi Sad	19,83694	45,25167	Neusatz
35	Novi Sad	19,83694	45,25167	Neusatz
36	Neuburg an der Donau	11,18709	48,73218	Donau
37	Neuburg an der Donau	11,18709	48,73218	Donau
38	Novi Sad	19,83694	45,25167	Neusatz
39	Neuburg an der Donau	11,18709	48,73218	Donau
40	Neuburg an der Donau	11,18709	48,73218	Donau
41	Austria	13,33333	47,33333	Österreich
42	Novi Karlovci	20,17948	45,07636	Karlowitz
43	Zemun	20,40116	44,8458	Semlin
44	Neuburg an der Donau	11,18709	48,73218	Donau
45	Zemun	20,40116	44,8458	Semlin

46	Zemun	20,40116	44,8458	Semlin
47	Montaigut-sur-Save	1,23133	43,68947	Save
48	Neuburg an der Donau	11,18709	48,73218	Donau
49	Neuburg an der Donau	11,18709	48,73218	Donau
50	Berlin	13,41053	52,52437	Stadt
51	Schnee Berge	8,54329	50,72938	Schnee
52	Vienna	16,37208	48,20849	Wien
53	Vienna	16,37208	48,20849	Wien
54	Vienna	16,37208	48,20849	Wien
55	Mohács	18,68621	45,9902	Mohács
56	Neuburg an der Donau	11,18709	48,73218	Donau
57	Neuburg an der Donau	11,18709	48,73218	Donau
58	Neuburg an der Donau	11,18709	48,73218	Donau
59	Neuburg an der Donau	11,18709	48,73218	Donau
60	Neuburg an der Donau	11,18709	48,73218	Donau
61	Bratislava	17,10674	48,14816	Preßburg
62	Belgrade	20,46513	44,80401	Belgrad
63	Zemun	20,40116	44,8458	Semlin
64	Serbia	20,45998	44,81892	Serbien
65	Belgrade	20,46513	44,80401	Belgrad
66	Neuburg an der Donau	11,18709	48,73218	Donau
67	Pančevo	20,64167	44,87177	Pancsova
68	Novi Sad	19,83694	45,25167	Neusatz
69	Zemun	20,40116	44,8458	Semlin
70	Hause	6,29332	62,24055	Hause
71	Pančevo	20,64167	44,87177	Pancsova
72	Smederevo	20,92763	44,66436	Semendria
73	Galicja	23	49,5	Galizien
74	Hungary	20	47	Ungarn
75	Banat of Temesvar	21	45,5	Banat
76	Neuburg an der Donau	11,18709	48,73218	Donau
77	Neuburg an der Donau	11,18709	48,73218	Donau
78	Bergen	5,32415	60,39299	Bergen
79	Bergen	5,32415	60,39299	Bergen
80	Bergen	5,32415	60,39299	Bergen
81	Neuburg an der Donau	11,18709	48,73218	Donau
82	Drencova	21,97464	44,63752	Drencova
83	Neuburg an der Donau	11,18709	48,73218	Donau
84	Bergen	5,32415	60,39299	Bergen
85	Drencova	21,97464	44,63752	Drencova
86	Drencova	21,97464	44,63752	Drencova
87	Drencova	21,97464	44,63752	Drencova
88	Drencova	21,97464	44,63752	Drencova
89	Södra Barken	15,65	60,01667	Barken
90	Neuburg an der Donau	11,18709	48,73218	Donau
91	Neuburg an der Donau	11,18709	48,73218	Donau
92	Drencova	21,97464	44,63752	Drencova
93	Comuna Islaz	24,74381	43,75179	Islaz
94	Neuburg an der Donau	11,18709	48,73218	Donau

95	Linz	14,28611	48,30639	Linz
96	Politischer Bezirk Kirchdorf an der Krems	14,21346	47,84346	Krems
97	Politischer Bezirk Kirchdorf an der Krems	14,21346	47,84346	Krems
98	Politischer Bezirk Kirchdorf an der Krems	14,21346	47,84346	Krems
99	Neuburg an der Donau	11,18709	48,73218	Donau
100	Jarčina	19,75835	44,91062	Römer
101	Neuburg an der Donau	11,18709	48,73218	Donau
102	Neuburg an der Donau	11,18709	48,73218	Donau
103	Neuburg an der Donau	11,18709	48,73218	Donau
104	Neuburg an der Donau	11,18709	48,73218	Donau
105	Neuburg an der Donau	11,18709	48,73218	Donau
106	Neuburg an der Donau	11,18709	48,73218	Donau
107	Berlin	13,41053	52,52437	Stadt
108	Berlin	13,41053	52,52437	Stadt
109	Banatski Dvor	20,51146	45,51866	Banater
110	Banatski Dvor	20,51146	45,51866	Banater
111	Bratislava	17,10674	48,14816	Preßburg
112	Bratislava	17,10674	48,14816	Preßburg
113	Bratislava	17,10674	48,14816	Preßburg
114	Bratislava	17,10674	48,14816	Preßburg
115	Bratislava	17,10674	48,14816	Preßburg
116	Bratislava	17,10674	48,14816	Preßburg
117	Bratislava	17,10674	48,14816	Preßburg
118	Neuburg an der Donau	11,18709	48,73218	Donau
119	Neuburg an der Donau	11,18709	48,73218	Donau
120	Türkenkopf	10,28333	50,88333	Türken
121	Gora Demir-Kapu	34,20823	44,58552	Demir-Kapu
122	Gora Demir-Kapu	34,20823	44,58552	Demir-Kapu
123	Gora Demir-Kapu	34,20823	44,58552	Demir-Kapu
124	Oltenia	23,5	44,5	Walachei
125	Bergen	5,32415	60,39299	Bergen
126	Bucharest	26,10626	44,43225	Balkans
127	Russen	24,123	71,08629	Russen
128	Austria	13,33333	47,33333	Österreich
129	Réaumur	-0,80305	46,72006	Réaumur
130	Réaumur	-0,80305	46,72006	Réaumur
131	Calafat	22,93328	43,99069	Calafat
132	Nikopol	24,89521	43,70528	Nicopol
133	Giurgiu	25,9627	43,88664	Giurgiu
134	Giurgiu	25,9627	43,88664	Giurgiu
135	Oltenia	23,5	44,5	Walachei
136	Berlin	13,41053	52,52437	Stadt
137	Galicia	23	49,5	Galizien
138	Novi Pazar	20,51222	43,13667	Bazar
139	Neuburg an der Donau	11,18709	48,73218	Donau
140	Neuburg an der Donau	11,18709	48,73218	Donau

141	Oraş Hârşova	27,94785	44,68721	Hârsova
142	Russen	24,123	71,08629	Russen
143	Galati	28,05028	45,43687	Galatz
144	Russen	24,123	71,08629	Russen
145	Neuburg an der Donau	11,18709	48,73218	Donau
146	Neuburg an der Donau	11,18709	48,73218	Donau
147	Neuburg an der Donau	11,18709	48,73218	Donau
148	Galati	28,05028	45,43687	Galatz
149	Galati	28,05028	45,43687	Galatz
150	Europe	9,14062	48,69096	Europa
151	Europe	9,14062	48,69096	Europa
152	Austria	13,33333	47,33333	Österreich
153	European Russia	40	55	Rußland
154	Big Turk	20,79806	41,99194	Türkei
155	Giurgiu	25,9627	43,88664	Giurgiu
156	Gemeente Stein	5,76126	50,97358	Stein
157	Galati	28,05028	45,43687	Galatz
158	Neuburg an der Donau	11,18709	48,73218	Donau
159	Neuburg an der Donau	11,18709	48,73218	Donau
160	Berlin	13,41053	52,52437	Stadt
161	Galati	28,05028	45,43687	Galatz
162	Berlin	13,41053	52,52437	Stadt
163	Berlin	13,41053	52,52437	Stadt
164	Berlin	13,41053	52,52437	Stadt
165	Berlin	13,41053	52,52437	Stadt
166	Berlin	13,41053	52,52437	Stadt
167	Novi Pazar	20,51222	43,13667	Bazar
168	Novi Pazar	20,51222	43,13667	Bazar
169	Langen	8,66852	49,98955	Langen
170	Big Turk	20,79806	41,99194	Türkei
171	Vienna	16,37208	48,20849	Wien
172	Vienna	16,37208	48,20849	Wien
173	Vienna	16,37208	48,20849	Wien
174	Vienna	16,37208	48,20849	Wien
175	Galati	28,05028	45,43687	Galatz
176	Galati	28,05028	45,43687	Galatz
177	Neuburg an der Donau	11,18709	48,73218	Donau
178	Galati	28,05028	45,43687	Galatz
179	Neuburg an der Donau	11,18709	48,73218	Donau
180	Neuburg an der Donau	11,18709	48,73218	Donau
181	Neuburg an der Donau	11,18709	48,73218	Donau
182	Neuburg an der Donau	11,18709	48,73218	Donau
183	Neuburg an der Donau	11,18709	48,73218	Donau
184	Neuburg an der Donau	11,18709	48,73218	Donau
185	Varna	27,91667	43,21667	Varna
186	Russen	24,123	71,08629	Russen
187	Russen	24,123	71,08629	Russen
188	Türkenkopf	10,28333	50,88333	Türken
189	Franken	9	50	Franken
190	Franken	9	50	Franken
191	Varna	27,91667	43,21667	Varna
192	Gura Stam- bulu Vechi	29,66667	45,3	Stambul
193	Gura Stam- bulu Vechi	29,66667	45,3	Stambul

12. Eidesstattliche Erklärung

Hiermit versichere ich,

- dass ich die vorliegende Masterarbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubter Hilfe bedient habe,
- dass ich dieses Masterarbeitsthema bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt habe
- und dass diese Arbeit mit der vom Begutachter beurteilten Arbeit vollständig übereinstimmt.

Wien, am 16.09.2021