

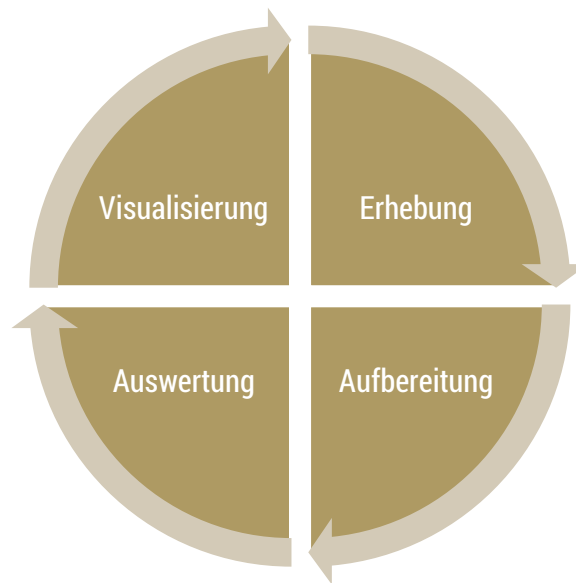


Digitale Tools zur Datenerhebung und Auswertung in den Geistes- und Sozialwissenschaften

Dr. Elisa Hofmann, IWK Uni Jena

Agenda

1. Willkommen und Einstieg
2. Was sind (digitale) Daten?
3. Vorstellung verschiedener Tools zur:
 - a. Datenerhebung
 - b. Datenaufbereitung
 - c. Datenauswertung
 - d. Datenvisualisierung
4. Übung
5. Workshop-Abschluss



1. Willkommen und Einstieg



Wer bin ich und warum gebe ich einen Workshop über digitale Tools?

- Seit Januar 2021: Wissenschaftliche Mitarbeiterin in der IWK
- 2016 - 2021: Promotion in Volkswirtschaftslehre mit Schnittstellen zur Sozialpsychologie und experimenteller Wirtschaftsforschung (Uni Jena und Max Planck Research School on “Adapting Behavior in a Fundamentally Uncertain World”)
- 2011 – 2015: Angestellt in einem Tageszeitungsverlag; Interne Unternehmens-Kommunikation, Marktforschung, Personal- und Organisationsentwicklung
- 2005 – 2011: Masterstudium der Medienwissenschaft, Philosophie und Psychologie (Uni Jena und Queensland University of Technology (Brisbane, Australia))
- Was mich sonst begeistert: Data Science, Nachhaltigkeit, Innovationen, Musik, Filme

Einstieg

Bitte kleben Sie ein buntes Post-It vor Ihre Kamera.

Nehmen Sie es ab, wenn Sie die gestellte Frage mit „Ja“ beantworten.

Kleben Sie es anschließend wieder drauf.



Ziele des Workshops

- Am Ende des Workshops haben Sie:
 - Erfahren, was (digitale) Daten sind
 - Einige digitale Tools zur Datenerhebung, -aufbereitung, -auswertung und –visualisierung kennengelernt
 - Eines oder mehrere digitale Tools selber ausprobiert
 - Einen Überblick erhalten, wo Sie weitere Informationen erhalten rund um die Themen (digitale) Daten



2. Was sind (digitale) Daten?



Was sind Daten? Was sind digitale Daten?

- Keine einheitliche Definition (teils disziplinspezifisch)
- Annäherung: Forschungsdaten sind "Dokumente in digitaler Form, mit Ausnahme von wissenschaftlichen Publikationen, die im Rahmen wissenschaftlicher Forschungstätigkeiten gesammelt oder erstellt werden" (EU DIRECTIVE 2019/1024)

Deutsche Forschungsgemeinschaft (DFG (2015) „DFG Guidelines on the Handling of Research Data“):

- Messwerte
- Laborergebnisse
- Umfragedaten
- Audiovisuelle Informationen
- Texte
- Objekte aus einer Sammlung/Proben



Was sind Ihre (digitalen) Daten?

- Mit welchen (digitalen) Daten arbeiten Sie aktuell bzw. haben Sie vor, zu arbeiten?
- Welche Forschungsfragen beantworten Sie damit?
- Welche digitalen Tools zur Datenerhebung und –auswertung kennen Sie schon?



3. Welche Tools gibt es?



Wie finde ich das richtige Tool?

1. Wie will ich welche (digitalen) Daten in welchem Format erheben?
2. In welchem Format liegen die Daten vor? Muss ich meine Daten vor der Auswertung noch aufbereiten?
3. Was kann ich mit meinen Daten für Fragen beantworten? Wie will ich diese Fragen beantworten?
4. Wie kann ich meine Ergebnisse zusammenzufassen und/oder visualisieren?

Kann ich meine Fragen mit den Daten und mithilfe der Tools beantworten?



Tools zur Datenerhebung

- (Digitale) Daten können auf vielfältige Weise erhoben werden – je nach Erhebungsart benötigt man unterschiedliche Geräte und Tools
 - z.B. Fotos in analogem (Kamera und Film) und digitalem Format (Digitalkamera, Einscannen)
 - z.B. Texte in analogem (Buch, Zeitungen, Zeitschriften, Briefe, Karteikarten, etc.) und digitalem Format (E-Book, E-Paper, Einscannen, Transkripte von Interviews, digitale Archive)
 - z.B. Audio in analogem (z.B. Schallplatte, Kassette, Tonband) und digitalem Format (z.B. analoge Tonaufnahmen digitalisieren, MP3, digitale Aufnahmegeräte)
 - z.B. Video auf VHS-Kassetten oder Mini-DV und digitalem Format (z.B. Videoaufnahmen auf SD-Karte, Aufzeichnungen von Online-Meetings)
 - z.B. Pencil-und-paper Fragebögen und Online-Umfragen (z.B. mit LimeSurvey, SoSci Survey, UniPark, Qualtrics, LamaPoll, Mentimeter) bzw. in digitale Formate eingetippte Fragebögen
 - z.B. Inhalte aus dem WWW: Web-Scraping (z.B. mit Python, R, ParseHub)
- Wenn man die Daten nicht selber erheben will / kann, kann man auch auf Archive und frühere Forschung zurückgreifen

Tools zur Datenerhebung - Digitale Archive

- Digitales Archiv nichtstaatlicher Archive Thüringens: https://dana.thulb.uni-jena.de/templates/master/template_dana/index.xml
- Digitalisierte Adressbücher, Kalendarien, Zeitungen und Zeitschriften (historisch); teilweise maschinenlesbar: <https://zs.thulb.uni-jena.de/content/main/journalList.xml#A>
- Digitalisierte „Kultur- und Wissensschätze Thüringens“ (z.B. Bauhaus-Bibliothek, kirchliches und schulisches Zeitschriftenwesen) <https://kuwi-thueringen.de/>
- Digitale Bestände THULB: https://collections.thulb.uni-jena.de/templates/master/template_collections/index.xml
- Digitale Quellen zur Gesetzgebung in Thüringen: <https://www.thulex.de/>
- Digitales Archiv Thüringer Flurnamen: <http://projekte.thulb.uni-jena.de/flurnamen/projekt.html>



Tools zur Datenerhebung - Korpora

Korpora-Archive:

- Korpora am Archiv für Gesprochenes Deutsch: http://agd.ids-mannheim.de/korpus_index.shtml
- Deutsches Textarchiv (Fokus auf deutschsprachige Texte 1600-1900): <https://www.deutschestextarchiv.de/>
- Clarin Data Archive (Fokus auf digital language data): <https://www.clarin.eu/content/data>
- Dariah (Datenbanken und Links zu Tools): <https://de.dariah.eu/web/guest/daten>
- HU Berlin (Korpusdatenbank): https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/links/korpora_links



Tools zur Datenaufbereitung

- **QualiAnon:** Textdaten pseudonymisieren und anonymisieren



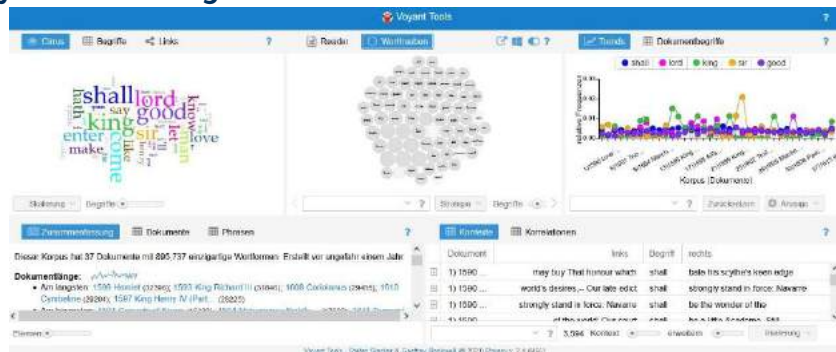
- **Open refine:** Aufbereiten von Daten(sätzen), z.B. Vereinheitlichen von Abkürzungen, Groß-/Kleinschreibung, Erkennen und Löschen von Duplikaten

<https://openrefine.org/>



Tools zur Datenauswertung - Texte

- Webbasierte Lese- und Analyseumgebung für digitale Texte
 - **Voyanttools:** <https://voyant-tools.org/>



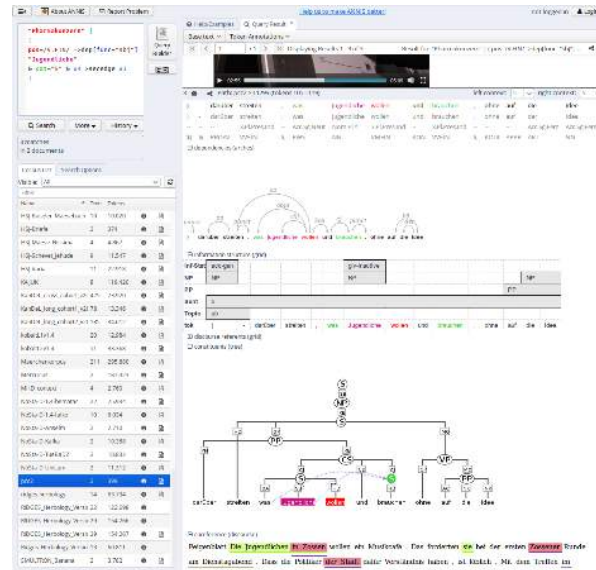
- „Klassisches“ Annotieren, Codieren und Auswerten z.B. Inhaltsanalyse:

- € • MAXQDA: <https://www.maxqda.de/>
- € • Atlas.ti: <https://atlasti.com/de/>
- QCAmap: <https://www.qcamap.org/ui/home>
- R und Python

Page 10 of 10

Korpora-Tools:

- Web-basierte Korpus-Annotation: <https://corpus-tools.org/annis/>
- Korpus-Annotation: <http://www.corpustool.com/>
- Korpus-Analyse-Tool mit Visualisierungsmöglichkeiten:
<http://corpora.lancs.ac.uk/lancsbox/help.php#tabs2>
- Korpus-Analyse-Tool:
<https://www.laurenceanthony.net/software/antconc/>



Tools zur Datenauswertung - Videos

- € • Mangold InterAct: <https://www.mangold-international.com/de/produkte/software/interact-videographie-software>
- € • Noldus The Observer XT: <https://www.noldus.com/observer-xt>
- The Communication Analysis Tool (CAT): <https://cat.ctwd.com.au/login.html>
- ANVIL: <https://www.anvil-software.org/>
- BORIS: <http://www.boris.unito.it/>



Tools zur Datenauswertung - Umfragedaten

- Manche Anbieter haben Auswertungen im Tool integriert
- Falls man die Daten separat hat (z.B. als Excel-Datei, .csv-Datei, SPSS-Datei etc.):
 - Auswertung in Excel
 - € • Auswertung in einer Statistiksoftware: STATA, SPSS
 - Auswertung mit Hilfe einer Programmiersprache und Oberfläche: R und R Studio, Python und pycharm



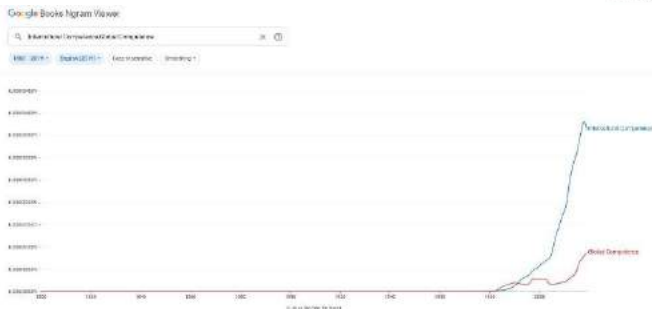
Tools zur Datenauswertung – Bibliometrische Daten

Digitales Wörterbuch der Deutschen Sprache (DWDS):

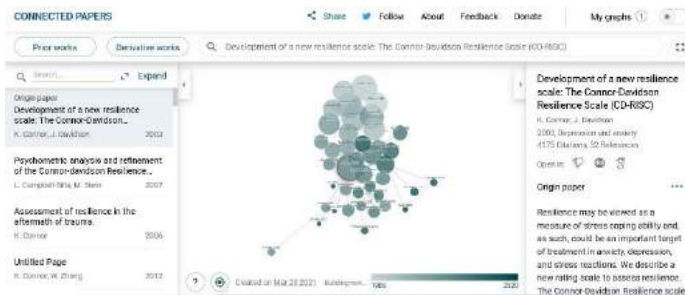
<https://www.dwds.de/>



Google Ngram Viewer: <https://books.google.com/ngrams>



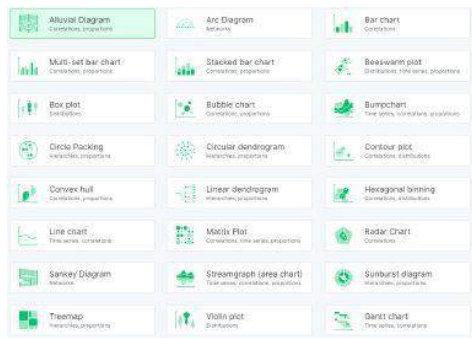
Connectedpapers: www.connectedpapers.com



Tools zur Datenvisualisierung 1

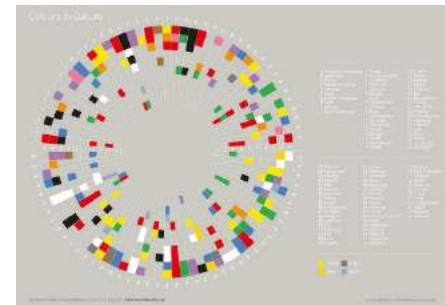
Grafiken mit eigenen Daten erstellen:

- Tipp: Datensatz vorher „aufräumen“ und nur die Daten hochladen, die man in der Grafik haben will
- <https://rawgraphs.io/>
- <https://www.datawrapper.de/>



Inspirationen:

- <https://datavizcatalogue.com/>
- **45 Ways to Communicate Two Quantities:**
 - <https://rockcontent.com/blog/45-ways-to-communicate-two-quantities/?view=undefined>
- <https://www.informationisbeautiful.net/visualizations/colours-in-cultures/>



Tools zur Datenvisualisierung 2

Zeitverlauf

- Timeline Knightlab:
<https://timeline.knightlab.com/>



Ort

- Dariah GeoBrowser:
<https://geobrowser.de.dariah.eu/index.html>



4. Übung



Übung

- Nehmen Sie sich 10 Minuten Zeit und probieren Sie eines oder mehrere Tools aus
- Beantworten Sie dabei folgende Fragen:
 - Welche Vorteile hat das Tool?
 - Welche Schwächen hat das Tool?
 - Wie kann ich es für meine Forschung verwenden?

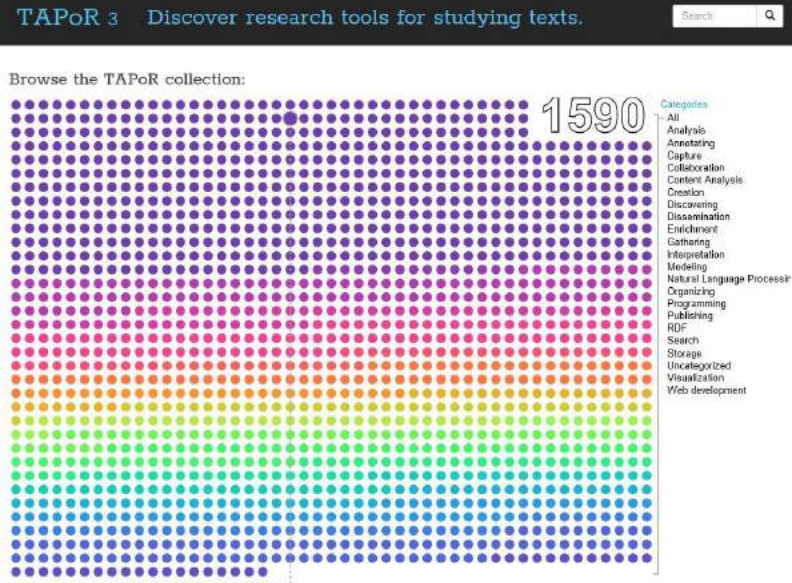


5. Workshop- Abschluss



Noch nicht das richtige Tool dabei gewesen?

- Tool-Finder TAPoR3: <http://tapor.ca/home>



TAPoR 3 Discover research tools for studying texts.

Filters: Narrow the Selection of Tools

Type of analysis:

Type of license:

Background Processing:

YWC Usability:

Ease of Use:

Warning:

Usage:

Tool Family:

Historic Tool (developed before 2005):

Compute Canada:

Link to Recipe:

TaDIRAH Methods:

Tag:

YadIRAH Goals:

Type:

First Previous 1 of 128

Sort by Name Rating Date

☆☆☆☆☆

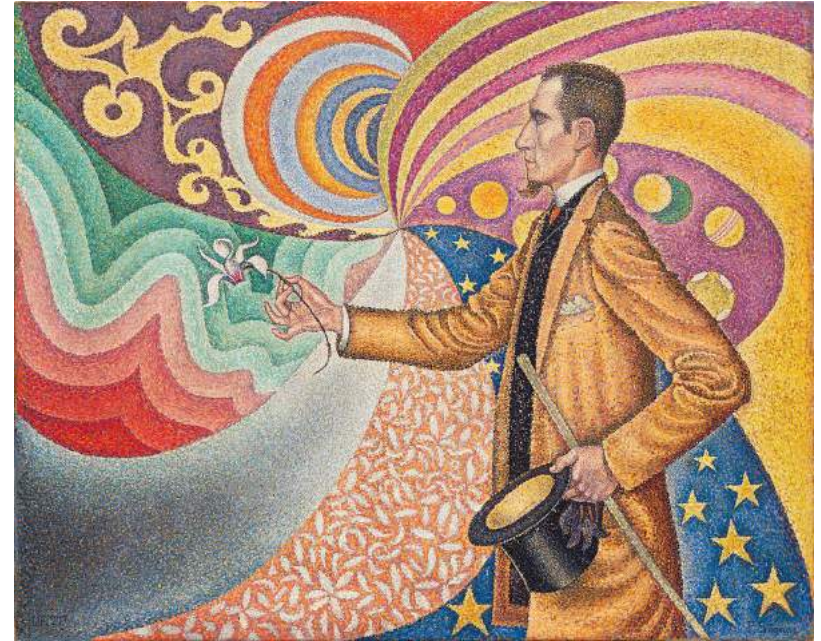
Noch nicht der richtige Datensatz dabei gewesen? Weitere Fragen?

- Die ThULB hat viele versteckte „Schätze“ und Nachfragen lohnt sich!
- Auch eine Suche bei Dataset Search von Google kann helfen: <https://datasetsearch.research.google.com/>
- Ansonsten gibt es noch disziplinen-spezifische Forschungsdatenbanken, die man durchsuchen kann
- Auch die Kontaktstelle Forschungsdatenmanagement der Uni Jena kann Ihnen behilflich sein, v.A. rund um das Thema Datenmanagement: <https://www.researchdata.uni-jena.de/>
- Weitere Workshops und Angebote rund um das Thema Digital Humanities werden vom Projekt Data Literacy Jena angeboten: <https://www.dataliteracy.uni-jena.de/>



3-Satz-Feedback

1. Was haben Sie gesucht und gefunden?
2. Was haben Sie gesucht, aber nicht gefunden?
3. Was haben Sie nicht gesucht, aber gefunden?



Signac, Paul - Portrait de Félix Fénéon

Vielen Dank für Ihre Aufmerksamkeit!

Ich wünsche Ihnen viel Erfolg mit Ihren Datenprojekten! Bei weiteren Fragen zu digitalen Tools:

Kontaktieren Sie mich gerne:

Dr. Elisa Hofmann, *Wissenschaftliche Mitarbeiterin*

Friedrich-Schiller-Universität Jena

Institut für Deutsch als Fremd- und Zweitsprache und Interkulturelle Studien

Bereich Interkulturelle Wirtschaftskommunikation (IWK)

Interkulturelle Personal- und Organisationsentwicklung

Ernst-Abbe-Platz 8, 07743 Jena

E-Mail: elisa.hofmann@uni-jena.de

Telefon: +49 (0)3641-944375

Viel Spaß noch bei der ISA 2021!

Quellen

- Workshop von Dr. Barbara Aehnlich zum Thema Digital Humanities & Data Literacy
 - Vortrag von Swantje Dogunke, Thüringer Universitäts- und Landesbibliothek (ThULB) Jena, Fachreferentin Geschichte, Alte Geschichte, Volks- und Völkerkunde
 - Vortrag von Petra Kunze, Thüringer Universitäts- und Landesbibliothek (ThULB) Jena, Projekterschließung Abteilung Medienerwerbung und -erschließung
 - Kontaktstelle Forschungsdatenmanagement (2021). Nachhaltig forschen -Zum Umgang mit Forschungsdaten in den Geistes- und Sozialwissenschaften. Präsentation von Volker Schwartze, 27.05.2021
- Bildnachweise: Pixabay, Websites der jeweiligen Tools, Uni Jena und privat

