

TDM Studio als forschungsnaher Bibliotheksservice

Strategie und Anwendung an der Universität Mannheim



Worum geht's?

„**TDM Studio** is a *text and data mining solution* for research, teaching and learning.“

- **Text-and-Data-Mining (TDM):** “die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.“
(§ 44b UrhG)
 - **Solution:** Online-Plattform mit Tools
 - für Retrieval und Analyse von lizenzierten Proquest-Inhalten
 - zur Visualisierung von Analyse-Ergebnissen
- **Wofür ist das gut und warum haben wir es lizenziert?**

Ausgangslage

Universität Mannheim

- Fokus auf empirischen Wirtschafts- und Sozialwissenschaften
 - Regelmäßig Projekte und Anfragen zu TDM mit Nachrichtenquellen in Nexis Uni, ABI-Inform und Factiva
- Probleme aufgrund von lizenzrechtlichen und technischen Beschränkungen

Bibliothek

- Strategischer Schwerpunkt im Bereich Forschungsdaten
- Entwicklung eines forschungsorientierten Serviceportfolios



Besserer Service möglich?

TDM: rechtlicher Hintergrund

- § 60d UrhG: gesetzliche Erlaubnis für Text- und Data-Mining zu Forschungszwecken

Aber:

- Anbieter können systematischen Download zur Wahrung von „Sicherheit und Integrität ihrer Netze und Datenbanken“ einschränken (und tun das auch)
- Offene Datenpublikation ist nicht erlaubt

→ Unsicherheiten und Einschränkungen bleiben bestehen

Warum TDM Studio?

- Klarer lizenzrechtlicher Rahmen (auch für die Grauzonen des UrhG und für uniexterne Kooperationspartner)
 - Gut etablierte Tools für die Datenanalyse (Jupyter Notebooks, R, Python)
 - Nachnutzbare Analysemethoden (sentiment analysis, topic models...)
 - Speicher- und Rechenkapazitäten
- als
Bibliotheksservice!
- Bibliothek als Ansprechpartnerin mit Beratungsservices (auch zu alternativen Angeboten und Tools)

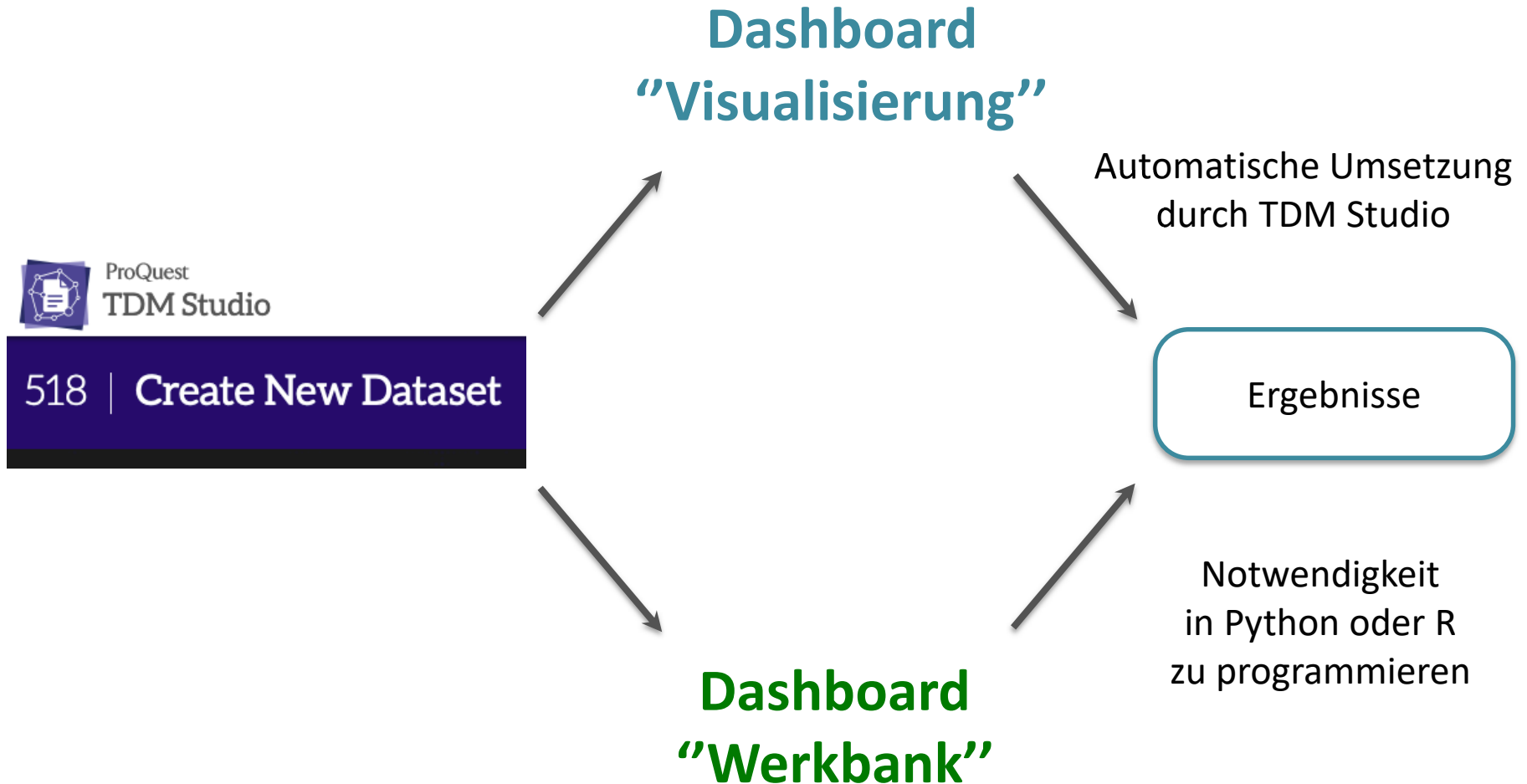
TDM Studio an der UB Mannheim



- Lizenz seit dem 01.12.2021, zunächst für 1 Jahr
- Aktuell Einsatz in Forschungsprojekten der BWL-Fakultät und in Forschungsdatenprojekten der UB
- Schulung von FachreferentInnen Ende April 2022
→ ggf. Infokampagne auch an anderen Fakultäten
- Evaluierung an der Fakultät BWL und an der UB geplant
- Hindernisse:
 - Funktionalität und konkreter Nutzen sind nicht ganz leicht zu erklären, v.a. für weniger technikaffine Zielgruppen
 - Graphische Oberfläche bietet nur eingeschränkte Möglichkeiten

TDM Studio: Anwendungsbeispiel

Workflow

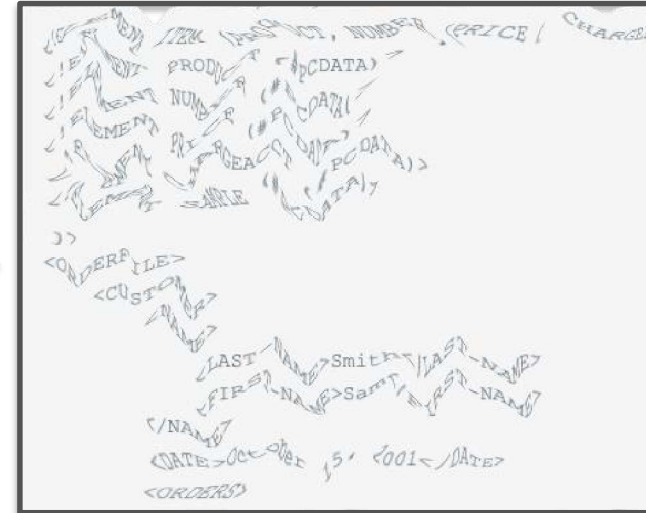
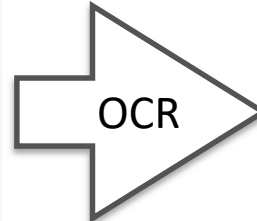


Fallbeispiel: Bearbeitung von historischen Zeitungen

PDF-Scan der NYT (02.06.1922)

XML

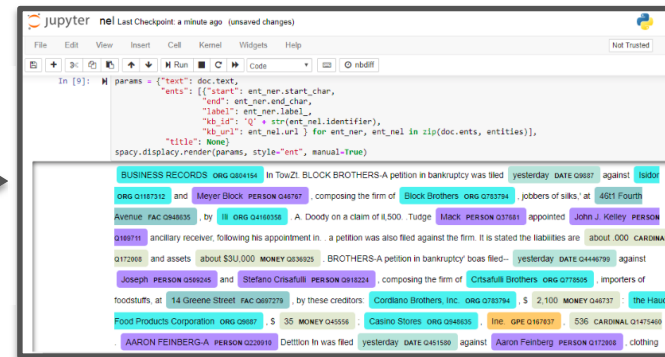
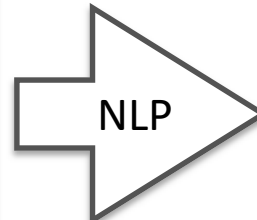
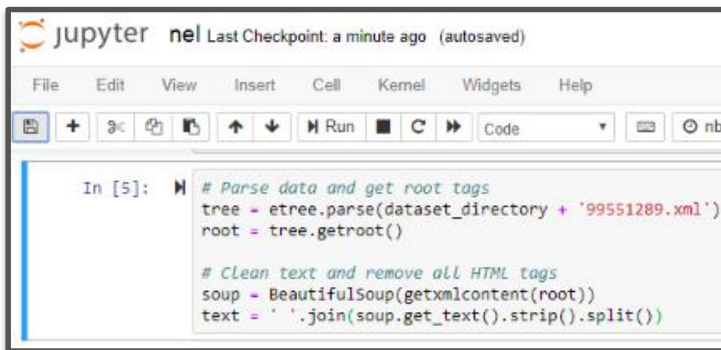
TDM Studio



Bearbeitung von xml-file in Jupyter Notebook

NER & NEL

Anwender



Links

- Login: <https://tdmstudio.proquest.com/home>
- Fallbeispiel:
<https://www.proquest.com/docview/99551289/fulltextPDF/B4F28CA78F674E99PQ/1?accountid=14570>

Full Text | Historical Newspaper

BUSINESS RECORDS

New York Times (1857-1922); New York, N.Y. [New York, N.Y]. 02 June 1922: 34.