

# Data Mining in der Geschichtswissenschaft – Historiker als die neuen Data Scientists



HENDRICKS,  
ROST & CIE.

Philipps



Universität  
Marburg



# Agenda

---

1. Wie arbeitet ein Historiker?
2. Gemeinsamkeiten mit anderen Disziplinen
3. Besonderheiten der Geschichtswissenschaft
4. Beispiele aus der Forschung
5. Historiker als Data Scientists

# Wie arbeitet ein Historiker?

---

1. „Alle“ potentielle Quellen ermitteln
2. Ausgewählte Quellen scannen
3. 90% der Quellen verwerfen
4. 10% genauer analysieren

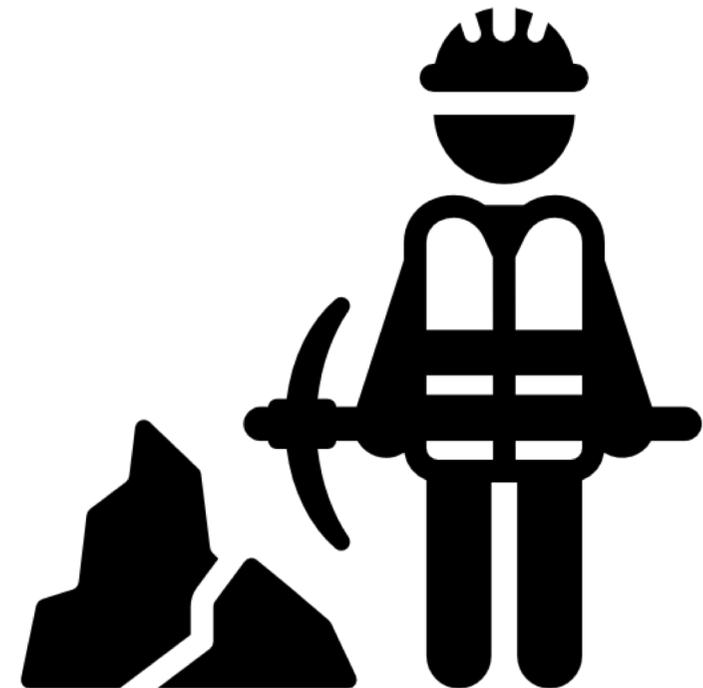


Icon made by Vecteezy from [www.flaticon.com](http://www.flaticon.com)

# Gemeinsamkeiten mit anderen Disziplinen

---

- „klassischer“ ETL-Prozess
  - Rohdaten erfassen
  - Bereinigen
  - Anreichern mit Metadaten
  - Auswerten
- Analyse
  - hypothesengetrieben
  - datengetrieben
- Identisches Tool Set
- Bedeutung der Datenqualität

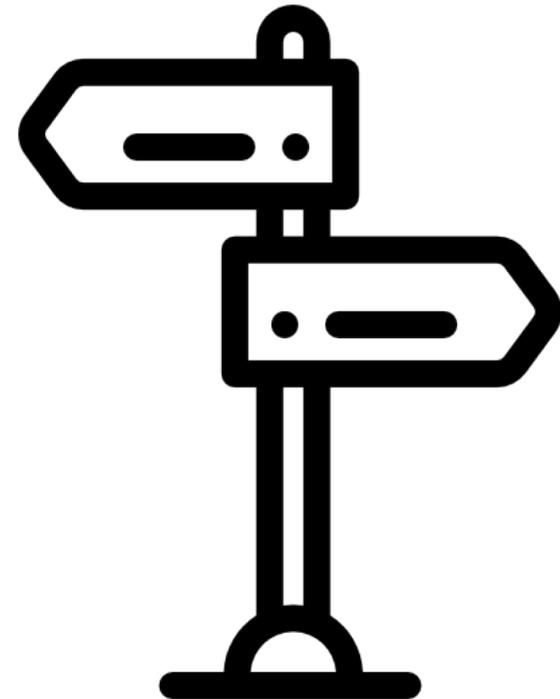


Icon made by freepik from [www.flaticon.com](http://www.flaticon.com)

# Besonderheiten der Geschichtswissenschaft

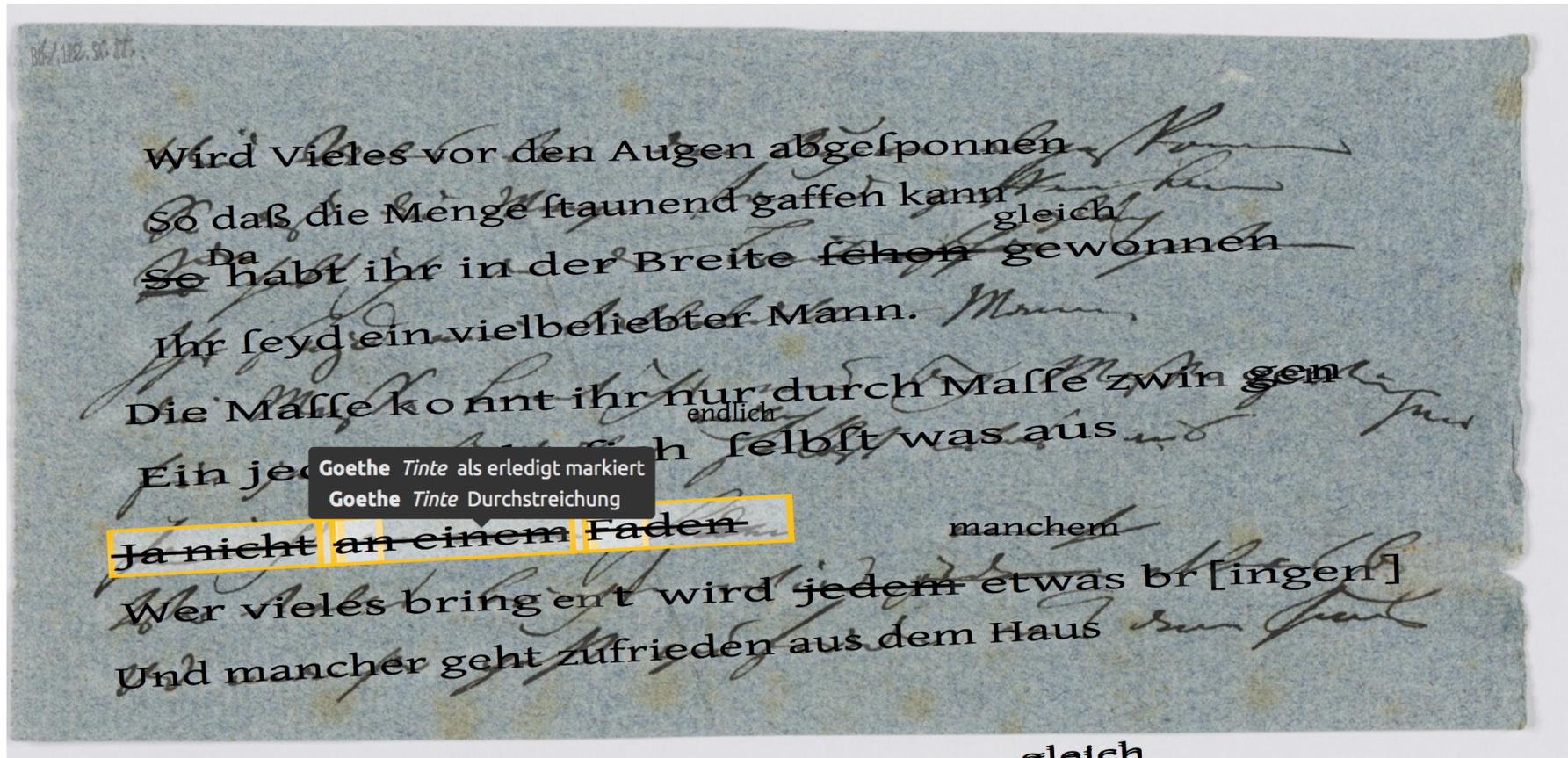
---

- Großteil der Daten nicht digital
  - Einschränkung durch Verfügbarkeit
- Komplexität der Metadaten
  - Schreibform, Drucktiefe, Anmerkungen, Aufbau
- Höhere Notwendigkeit für Interpretation
  - Zeitlicher Bedeutungswandel von Begriffen
  - Interpretation von Zahlen und Daten



Icon made by freepik from [www.flaticon.com](http://www.flaticon.com)

# Digitale Faust-Edition

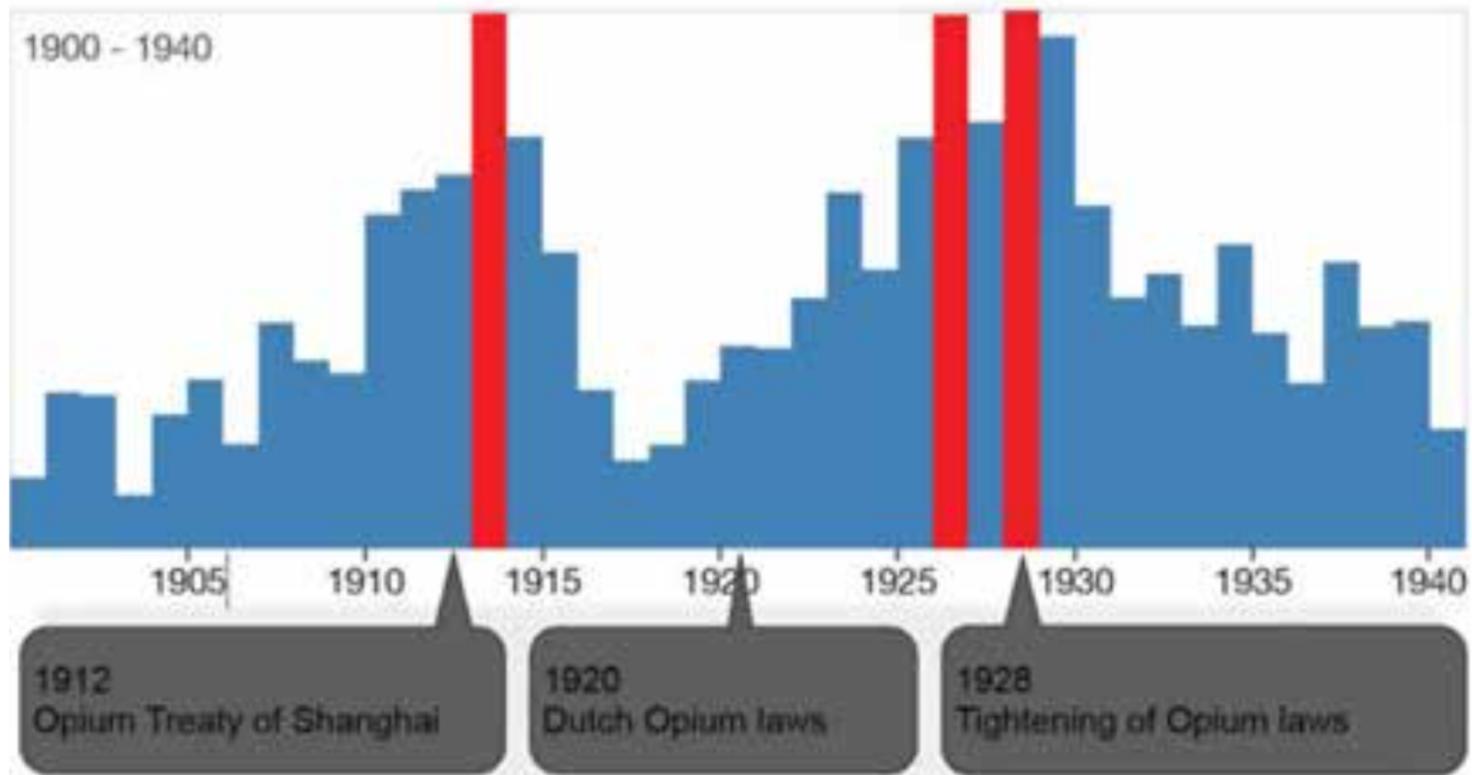


# Kostümmuster-Analyse

		Farbe			
		Rot	Gelb	Blau	
Kleidung	Kleid	4	3	2	Aggreg.lvl 2
	Abendkleid	3	2	1	Aggreg.lvl 1
	Schulterfreies Kleid	1	2	0	Detaillevel
	One-Shoulder Kleid	2	0	1	
	Sommerkleid	1	1	1	

Abbildung 1: Aggregation über Kategorieebenen von Merkmalstaxonomien

# Opium-Debatte im Zeitverlauf



# Historiker als Data Scientists

---

