

Digitale Edition neu denken

Natascha Naumann, Matthias Wunsch, Martin Lemke

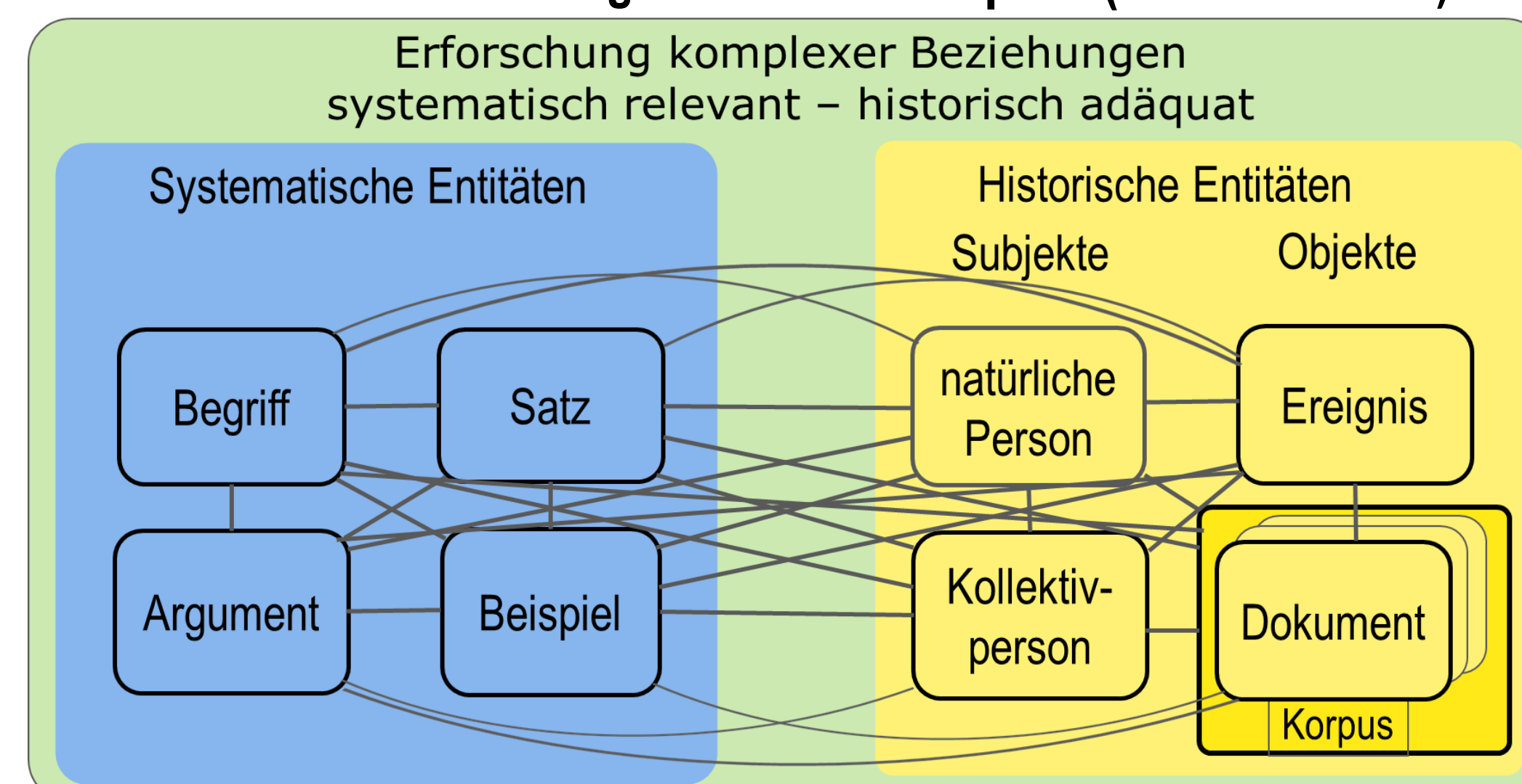
Digitale Sammelkorpus-Edition in Form einer Forschungsplattform für Konstellationsforschung zum Wiener Kreis

Projekt in Planungsphase

Wiener Kreis

- Was**
- bedeutende Gruppe von Philosoph:innen und Einzelwissenschaftler:innen
- Wann**
- Mitte 1920er bis Ende 1930er Jahre
- Wie**
- lehnte hierarchische Diskussionskultur in der Philosophie ab
 - übertrug die hochgradig arbeitsteilige Art naturwissenschaftlicher Forschung auf die Philosophie
 - legte Wert auf Themen- und Methodenvielfalt
 - praktizierte „geistige Gütergemeinschaft“
- Warum wichtig**
- beeinflusste viele Bereiche von Wissenschaft und Gesellschaft
 - ist Ausgangspunkt verschiedener heute noch aktiver Denkrichtungen der Philosophie

Konstellationsforschung in der Philosophie (Henrich 1991)



Methode um philosophische Konstellation wieder „zum Leben“ zu erwecken

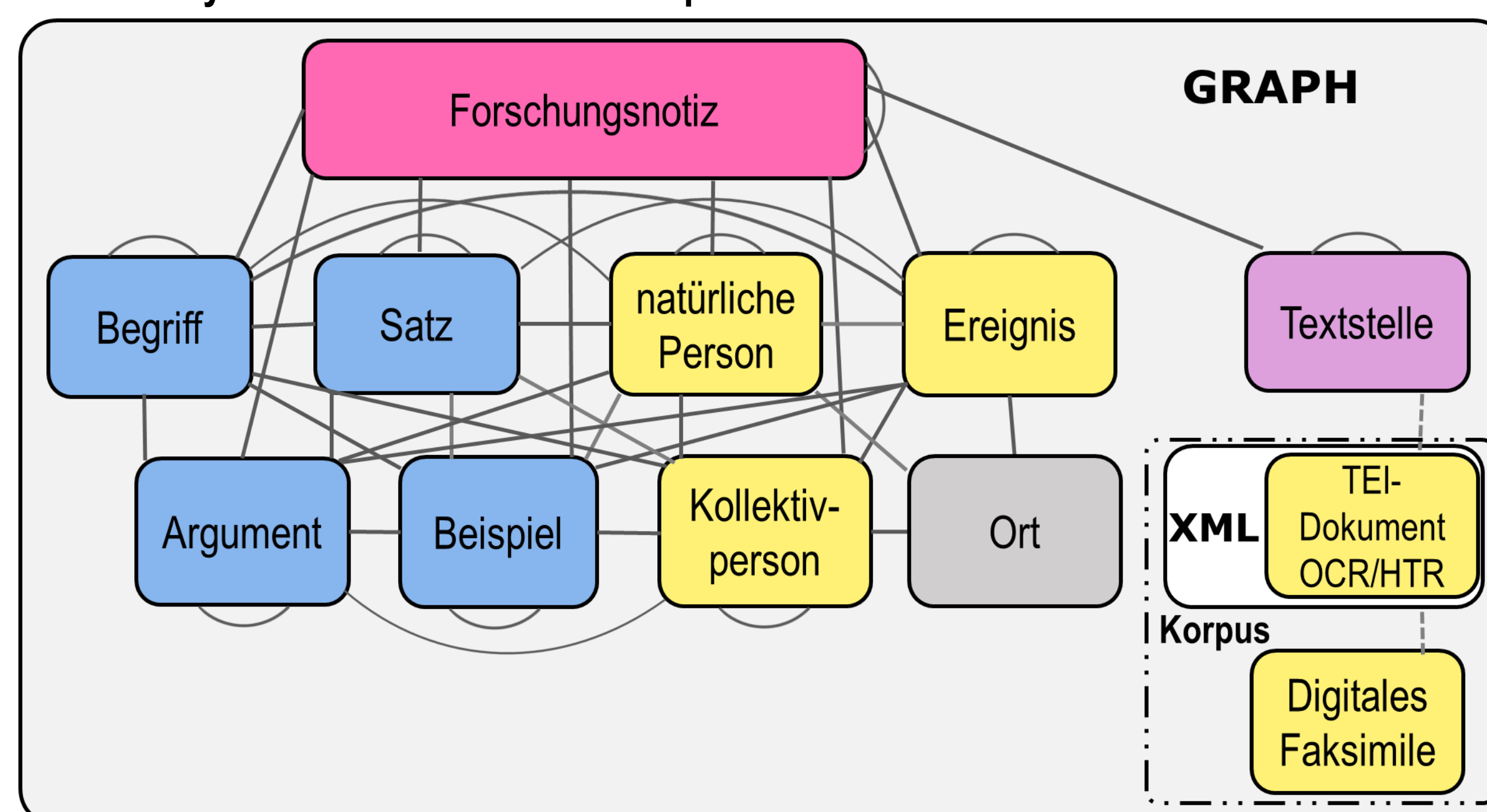
Probleme für Konstellationsforschung

- mehr Material, als ein Mensch alleine lesen könnte
- Werke vieler Mitglieder einer Konstellation sind nicht ediert
- vorhandene Editionen erlauben keine übergreifenden Suchen nach relevanter Information
- Korpus von Konstellation ist zu Beginn nicht komplett konturierbar
- forschungsrelevante Information ist nicht im Vorhinein festlegbar

Lösung ausgehend von Methode

- komputational**
- statt detaillierter Auszeichnung von Teilmengen: maschinengestützte Verarbeitung und Auswertung großer Bestände von Dokumenten
- dynamisch**
- Forschungsergebnisse in wachsender Datenstruktur standardisiert ablegbar
 - Forschungssituation ändert sich: Dokumente zur Datenbasis hinzufügbare

Hybride Datenarchitektur Graphdatenbank und TEI-XML-Datenbank

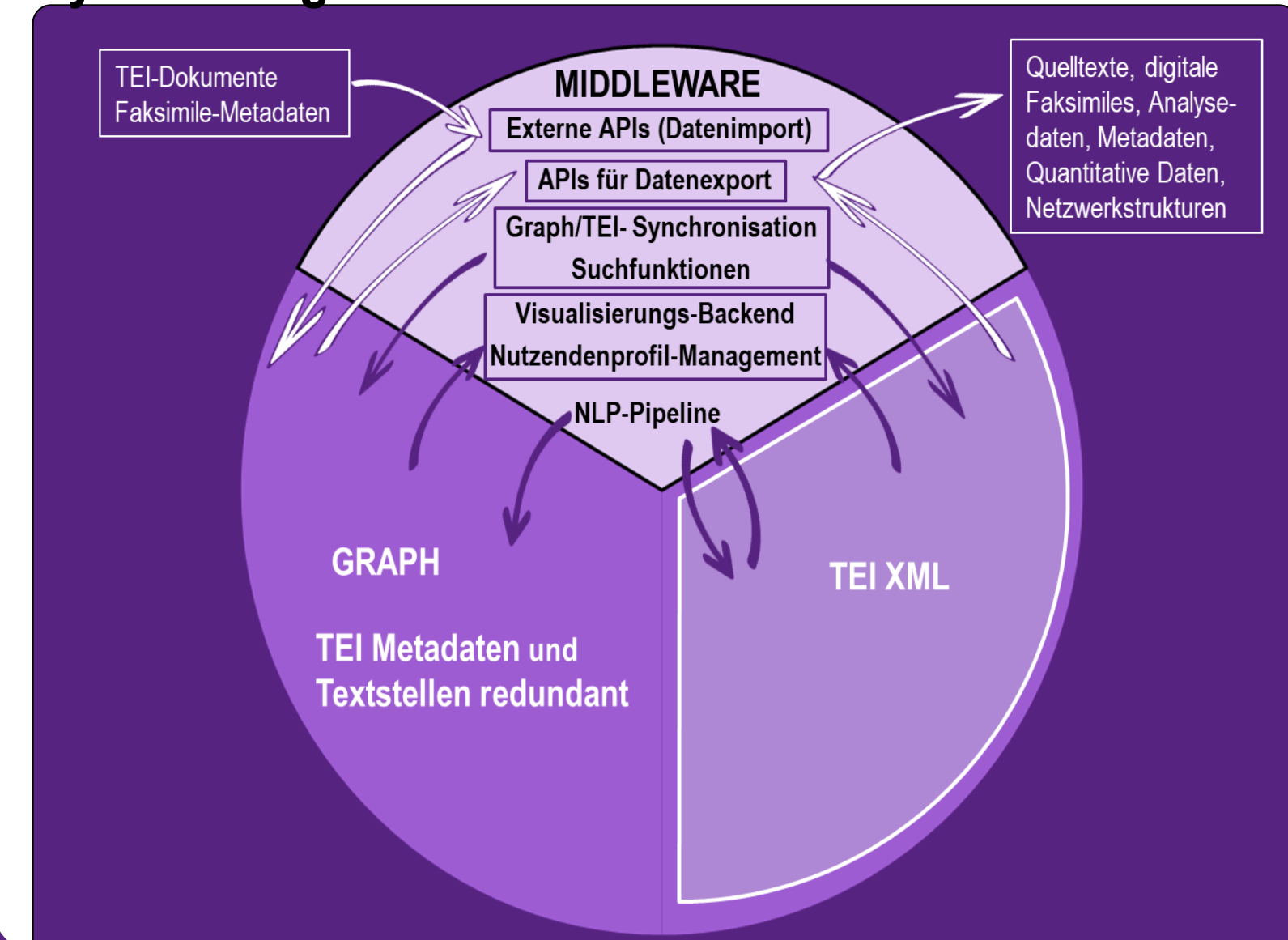


Entitäten					Beziehungen	
Systematische Entität	Historische Entität	Forschungsnotiz	Textstelle	Ort	Forschungsbeziehung (interpretativ)	Herkunftsbeziehung (dokumentativ)
Spezifische Attribute, Validierungsattribute, Entstehungsnachweis						

Zusätzliche Entitäten im Graphen

- Forschungsnotiz**
- modelliert mehrstellige Beziehungen
 - speichert textuelle Information zu Forschungsergebnissen
- Textstelle**
- speichert redundant relevante Textauszüge aus TEI Dokument für Graphperformanz
- Ort**
- ermöglicht räumliche Analysen

Systemintegration



Ergänzende Komponenten

URI-referenziert

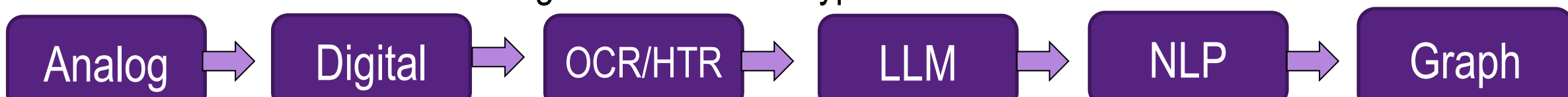
Datenbank für forschungs-relevante Ergebnisse automatisierter Analysen

Normdaten VIAF Geonames Wikidata

Externe Forschungsquellen Literatur

Erste Textbearbeitung

- Transkription**
- Transkribus + LLM-Nachbearbeitung für höhere Qualität
- NLP- Pipeline, erste Befüllung des Graphen**
- automatische Auszeichnung in TEI Dokumenten von Personen, Orten, Zeiträumen, philosophischen Begriffen,
 - Klassifizierung nach Dokumenttypen



Weitere Texterschließung

- korpusweite Ähnlichkeitssuche**
- lexikalische und semantische Verfahren
 - Textduplikate automatisiert als Instanz der Entität Textstelle erstellt

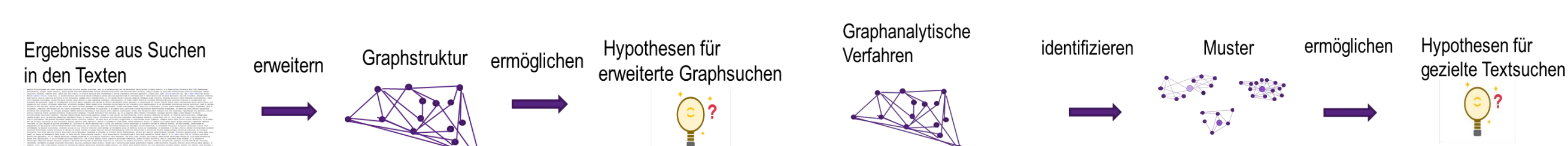
Mit wachsender Datendichte und zunehmender Validierung entwickelt sich die Plattform zu einem leistungsstarken Tool für Musterberechnung und Datenabfrage



- Gezielte Datenabfragen**
- Query-basierte Ähnlichkeitssuche zu Textstelle
- Filterbasierte Suche im Graphen
- Traversierungsbasierte Suche im Graphen

- Explorative und gezielte Zeitverlaufsanalysen**

- Bidirektionaler Forschungsansatz**



FAIR Prinzipien

- Findability**
- F1: URI, F2: TEI P5, PROV-O, Normdaten, F3: URI, TEI-Id, XPath, F4: OAI-PMH, DARIAH-DE Repository (geplant)
- Accessibility**
- A1.1: REST-APIs, IIF-Manifeste (geplant), A1.2: Authentifizierung noch zu spezifizieren
- A2: Metadaten-Persistenz in Entwicklung
- Interoperability**
- I1: TEI P5, RDF/JSON-LD, XPath, I2: PROV-O, SKOS, Normdaten, I3: URI, TEI Id, XPath
- Reusability**
- R1.1: Datennutzungslizenzen noch offen
- R1.2: PROV-O, R1.3: TEI P5, SKOS, Normdaten

Qualitätssicherung

Alle Einträge werden durch strukturierten Begutachtungsworkflow validiert

Editionsarbeit

Erschließung historischer Quellen nach dokumentierten intersubjektiv reproduzierbaren Prinzipien

Annotation in den TEI Dokumenten

- Auszeichnung von Entitäten in den TEI-Dokumenten
- Markierung relevanter Textstellen in TEI-Dokumenten
- Sukzessive Ergänzung der TEI-Header-Metadaten

Forschungsdiskurs im Graphen

- Erstellung neuer Knoten verschiedener Entitätstypen
- systematische Erstellung von Kanten zwischen Knoten
- wissenschaftliche Interpretationen werden sowohl in Entitäts- und Kantenattributen als auch in Textform in der Entität Forschungsnotiz dokumentiert

Schaffung einer Grundlage für weiterführende Forschung

- Forschungsdiskurs einseh- und exportierbar
- Korpus auch für andere Forschung zugänglich
- Forschung Grundlage für Entdeckung wichtiger Dokumente
- Metagraph für andere Gebiete der Konstellationsforschung nutzbar

Henrich, Dieter 1991: *Konstellationen: Probleme und Debatten am Ursprung der idealistischen Philosophie (1789–1795)*. Stuttgart.