

Die „Deutsche Biographie“ als vernetzte Infrastruktur für biographische Forschung



Zukunft des digitalen Erbes? Nachhaltige Lösungen für Gedächtnis- und
Wissenschaftseinrichtungen. 7. Digitale Bibliothek, Graz, 3.3.2017

Malte Rehbein, Bernhard Ebner, Matthias Reinert, Maximilian Schrott
Historische Kommission bei der Bayerischen Akademie der Wissenschaften

HISTORISCHE
KOMMISSION
BEI DER BAYERISCHEN
AKADEMIE DER
WISSENSCHAFTEN

IAK
MÜNCHEN

Überblick

- Biografische Basis
- Digitalisierungsprozess
- Datengewinnung und Aggregation
- „HBIS“
- Rolle der GND
- Schnittstellen
- Ausblick: Forschungslabor

**„... alle relevanten,
quellenbasierten,
zitierfähigen Angebote
zu einer Person o. Familie“
www.deutsche-biographie.de**

Biografische Basis

- lexikalische Arbeit seit der Allgemeinen Deutschen Biographie (ADB, 1875-1912), bis in die Gegenwart
- Redaktion der Neuen Deutschen Biographie, Herausgeber Hans-Christof Kraus im Auftrag der Historischen Kommission
- aktuell NDB 26, Tecklenburg – Vocke, 2016)
- Artikel zu verstorbenen Personen und Familienverbänden, mit bedeutender Leistung in Bezug auf die deutsche Kultur, Kunst, Literatur, Wirtschaft, Politik und Sport

Digitalisierungsprozess

- Kooperation mit der Bayerischen Staatsbibliothek / MDZ seit Ende der 1990er Jahre, gefördert durch die DFG
- Bildangebot der Bände, Registerdatenbank der erwähnten Personen, Volltexterfassung

Datengewinnung und Aggregation

- GND-Vollversorgung des Registers durch die BSB (2009)
- Textstrukturierung in XML
- Tiefenerschließung von Kopfzeilen
- Personenerkennung im Text

Datengewinnung und Aggregation II

- Gewinnung von Partnern zur gemeinsamen Erschließung der personenbezogenen Bestände mit Hilfe der GND
- Abgleich- und Redaktionsverfahren, zusammen mit der BSB
- computerlinguistische Personen- und Relationenerkennung (prototypisch Lehrer/Schüler)
- heuristische Identifikation von Personen

50.000 Artikel
738.000 Personen
19.000 Orte

„HBIS“ - Kartenzugang

Deutsche Biographie

Name

NDB/ADB Lexika Quellen Literatur Portraits Objekte

Filter: Ortstypen
... ändert berücksichtigte
Orte (Summe im Oval)

Suche nach Ortstypen
... "Klick" definiert einen
Suchbereich (Rechteck)

Ovale ... aggregieren
Orte nach Anzahl
d. Nennungen

„HBIS“ - facettierte Personensuche

Erweiterte Suche

Name

NDB/ADB Lexika

Quellen Literatur

Portraits Objekte

Beruf/Funktion

graz

Geburtsjahr Sterbejahr

GND

Berufe

Schriftsteller	162
Schauspieler	77
Historiker	69
Komponist	68
Jurist	63
...	50

Treffer 1 - 10 von 1632 Treffern

Relevanz « »

Goethe, Johann Wolfgang von (seit 1782)
1749 bis 1832, Dichter; weimarer Staatsmann; Naturforscher; Jurist; Zeichner; Theaterleiter

Kepler, Johannes
1571 bis 1630, Astronom; Mathematiker; Naturphilosoph

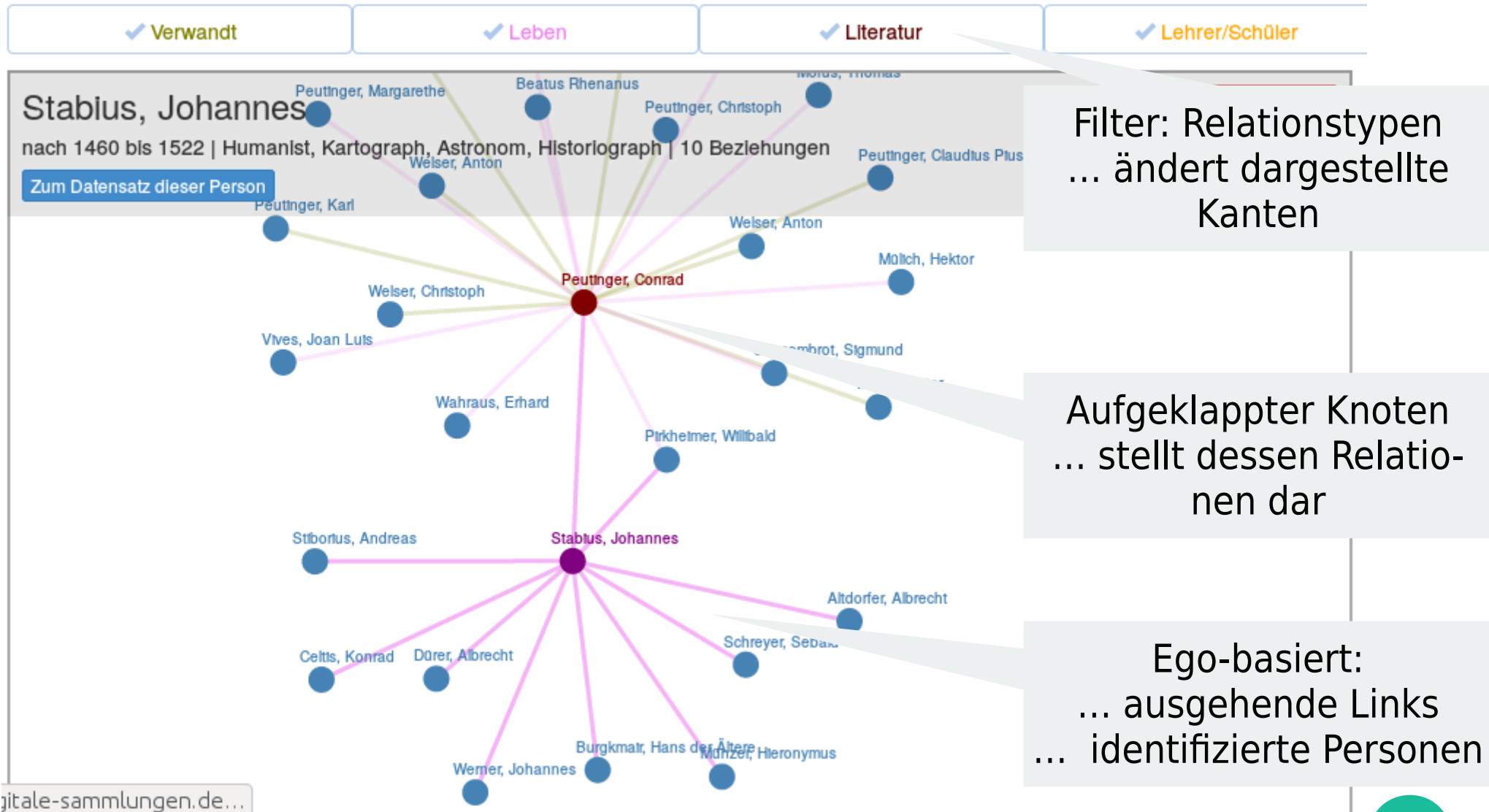
Leopold II.
1747 bis 1792, Kaiser; Großherzog von Toscana

Schubert, Franz
1797 bis 1828, Komponist

Johann
1782 bis 1859, Erzherzog von Österreich; österreichischer Feldmarschall; deutscher Reichsverweser



„HBIS“ - Graph-Darstellung



Rolle der GND

- **Personen**

- 99% der Personen identifiziert
- 226 Angebote via GND-Beacon verlinkt
- 101.000 „Beziehungen“ durch Textanalyse (2/3) und GND (1/3)

- **Orte**

- 8.000 Orte aus GND
- 11.000 eigene Orte in OpenStreetMap lokalisiert
- davon rd. 3.500 GND

- **Institutionen / Organisationen (Körperschaften): ./.**

- **Werktitel: ./.** Literaturnachweise: ./.

Schnittstellen

- Sparql: ~3 Mio RDF triples
- Solr: ~ 738.000 entries
- Neo4J: ~ 800.000 nodes, 1 Mio edges
- BeaconLinker: ~ 6 Mio entries

Deutsche Biographie - Services

This Webpage is currently under development. Please do not rely on in production environment.

The screenshot displays a grid of service cards for the Deutsche Biographie project. Each card has a title, a description, and a 'View details' button. The cards are:

- BeaconLinker**: relevant heritage, by authority, identifier. View details »
- access metadata**: mapped to RDF. View details »
- formulate queries**: against our self - index, using lucene's query language. View details »
- query**: node spanning relations, made easy with Neo4J. View details »
- link to**: sth

Ausblick: Forschungslabor

- **Abteilungsleiter: Malte Rehbein**
- **Accounts für Intensiv-Nutzende, Gruppenfunktion**
- **Suchhistorie, Kommentarfunktion**
- **zitierfähige Visualisierungen, Export**
- **Operationen und Auswertungen auf Ergebnismengen**

Ausblick: Forschungslabor & Ontologie

- **Entwicklung von**
 - GND-Ontologie,
 - über CIDOC-CRM (eCRM) Mapping
 - zu einem genuin biographischen, interoperablen Ontologie(-Schema)

Nachhaltigkeit

- Digitalisierung ist mit dem Digitalisat nicht beendet. Nur die wissenschaftliche Arbeit damit/daran kann Relevanz erzeugen.
- Volltextangebote müssen immer wieder und erneut strukturiert und aufbereitet werden, lese- und Recherchegewohnheiten wie -fragestellungen gerecht werden ohne den Bezug zum Original aufzugeben.

Nachhaltigkeit II

- Persistenz von Identifikatoren einzurichten und zu erhalten ist notwendig. Alle anderen Verlinkungen sind flüchtig.
- Datenaggregation - als Gesamtkonvolut wie als Prozess - ist ein Gegenstand wissenschaftlichen Interesses. Sie sollte so erfolgen, dass wissenschaftliches Bearbeiten möglich wird.
- Die Daten müssen frei nachnutzbar sein:
 - CC0 für Register/Metadaten
 - CC-BY-NC-SA - ADB Artikel
 - CC-BY-NC-ND - NDB Artikel