

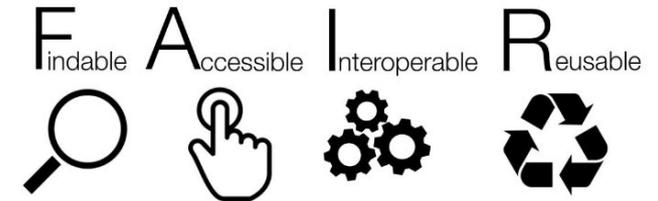
# I wie BEACON

3. FAIR February Meetup,  
5. bis 26. Februar 2025

NFDI Konsortium Text+, Task Area Editions  
Fernanda Alvares Freire (TU Darmstadt), Philipp Hegel (AdW Mainz),  
Harald Lordick (STI)

# Programm FAIR February 2025

- 5. Februar 2025 – **F** wie Qualitätssicherung
- 12. Februar 2025 – **A** wie User Experience
- **19. Februar 2025 – I wie BEACON**
- 26. Februar 2025 – **R** wie Dokumentation



Graphics: SangyaPundir - Own work, CC BY-SA 4.0,  
<https://commons.wikimedia.org/w/index.php?curid=53414062>

# Organisationsteam

- **Fernanda Alvares Freire (Technische Universität Darmstadt)**
- **Dr. Philipp Hegel (Akademie der Wissenschaften und der Literatur Mainz)**
- Kilian Hensen (Nordrhein-Westfälische Akademie der Wissenschaften und der Künste)
- Dr. Sandra König (Nationale Akademie der Wissenschaften Leopoldina, Halle (Saale))
- Karoline Lemke (Berlin-Brandenburgische Akademie der Wissenschaften)
- **Harald Lordick (Salomon Ludwig Steinheim-Institut für deutsch-jüdische Geschichte)**

# Ablauf

- Digitale Editionen, LOD, RDF und BEACON
- GND-BEACON-Hub
- Erzeugung von BEACON-Dateien mit Python

# Organisatorische Hinweise

- Der Workshop wird nicht aufgezeichnet.
- Der Chat steht jederzeit für Kommentare und Fragen offen.
- In der Diskussion können Sie sich mit dem Handzeichen-Button zu Wort melden.
- Handles: <https://nfdi.social/@Textplus> | [@Textplus\\_NFDI](https://nfdi.social/@Textplus_NFDI)
- Hashtag: [#FAIRfebruary](https://nfdi.social/#FAIRfebruary)

# Rückblick: I wie Interoperabilität

- (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- (Meta)data use vocabularies that follow the FAIR principles.
- (Meta)data include qualified references to other (meta)data.

# Rückblick: I wie Interoperabilität

- Renat Shigapov (Universitätsbibliothek Mannheim): *Knowledge Graphs for Interoperable NFDI: Digital Editions*
- Dieta Svoboda-Baas (Heidelberger Akademie der Wissenschaften): *Editions2LD: Modellierung von Texteditionen als Linked Data*
- Stefan Dumont (Berlin-Brandenburgische Akademie der Wissenschaften): *CMIF und CorrespSearch: Lessons learned*

# Rückblick: I wie BEACON

- DH quo vadis?, Passau 2024: *BEACONe unde venis et quo vadis?*
- Open Up DSE, Zürich 2024: *Pivot and Derivative Data Formats as Approaches to Interoperability in Digital Scholarly Editions*
- Edition als Transformation, Wuppertal 2024: *Die Edition sub specie aeternitatis. Infrastrukturelle Bedingungen nachhaltiger digitaler Editionen.* (Veröffentlicht in Editio)
- Pessoa Digital: Development and Sustainability Measures for the Digital Edition of Fernando Pessoa, Rostock 2023.

# Umfrage



# Lösung nach Berners-Lee



# Erster Stern: Offene Daten

*Im Netz unter freier Lizenz verfügbar.*

*Beispiel: Creative Commons*

- *BY = Nennung der Urheberin, des Urhebers*
- *ND = Unverändert*
- *NC = Nicht kommerzielle Nutzung*
- *SA = Weitergabe dieser, ggf. auch veränderter Daten unter gleicher Lizenz*



# Zweiter Stern: Maschinenlesbarkeit

In Ergänzungen zum ersten Kriterium maschinenlesbarer, „strukturierter“ Daten

*Beispiel:* Tabelle statt Bild der Tabelle



# Dritter Stern: Offene Formate

In Ergänzungen zu den ersten beiden Kriterien Verwendung nicht-proprietärer Formate

*Beispiel: CSV statt XLS*



# Vierter Stern: Offene W3C-Standards

In Ergänzungen zu den ersten drei Kriterien Verwendung von W3C-Standards

*Beispiel:* RDF und SPARQL



Graphik: W3C, Bill Schwappacher



# Fünfter Stern: Verknüpfung der eigenen Daten

In Ergänzungen zu den ersten vier Kriterien die Verknüpfung der eigenen Daten mit den Daten anderer

*Beispiel:* Gegenstand dieser Veranstaltung



# Ergänzungen

- URI
- Metadaten



# BEACON

```
#FORMAT: BEACON
#PREFIX: http://d-nb.info/gnd/
#TARGET: http://www.steinheim-institut.de/ebib-djg-nrw/query.html?database=NRW-Bibliografie&text1={ID}&kategorie1=gnd
#MESSAGE: Bibliografie deutsch-jüdische Geschichte Nordrhein-Westfalen
#DESCRIPTION: Literatur zur deutsch-jüdischen Geschichte in Nordrhein-Westfalen (Rheinland und Westfalen), Print und Online, ca. 7.500 Titel (806 Einträge in der BEACON-Datei)
#INSTITUTION: Steinheim-Institut
#CONTACT: Harald Lordick <+49(0)201-82162900>
#TIMESTAMP: 2022-04-05T18:24:25.648+02:00
#FEED: http://www.steinheim-institut.de/ebib-djg-nrw/ebib-djg-nrw-beacon.txt
100373550
10076410X
100880363
100969984
101140614
1011477629
```

BEACON spec.: <https://gbv.github.io/beaconspec/beacon.html>

# Verpflichtende Metadatenfelder

#FORMAT: BEACON

#PREFIX: <http://d-nb.info/gnd/>

#TARGET: <http://www.steinheim-institut.de/ebib-djg-nrw/query.html?database=NRW-Bibliografie&text1={ID}&kategorie1=gnd>

# Einige optionale Metadatenfelder

#MESSAGE: Bibliografie deutsch-jüdische Geschichte Nordrhein-Westfalen

#DESCRIPTION: Literatur zur deutsch-jüdischen Geschichte in Nordrhein-Westfalen (Rheinland und Westfalen), Print und Online, ca. 7.500 Titel (806 Einträge in der BEACON-Datei)

#INSTITUTION: Steinheim-Institut

#CONTACT: Harald Lordick <+49(0)201-82162900>

#TIMESTAMP: 2022-04-05T18:24:25.648+02:00

#FEED: <http://www.steinheim-institut.de/ebib-djg-nrw/ebib-djg-nrw-beacon.txt>

# Datenfelder

Normdateinummer (entsprechend #PREFIX)|Treffer|ID (gemäß #TARGET)

123456789|10|Vorname\_Nachname

entspricht zum Beispiel

[http://d-nb.info/gnd/123456789|10|http://de.wikipedia.org/wiki/Vorname\\_Nachname](http://d-nb.info/gnd/123456789|10|http://de.wikipedia.org/wiki/Vorname_Nachname)

# BEACON

```
#FORMAT: BEACON
#PREFIX: http://d-nb.info/gnd/
#TARGET: http://www.steinheim-institut.de/ebib-djg-nrw/query.html?database=NRW-Bibliografie&text1={ID}&kategorie1=gnd
#MESSAGE: Bibliografie deutsch-jüdische Geschichte Nordrhein-Westfalen
#DESCRIPTION: Literatur zur deutsch-jüdischen Geschichte in Nordrhein-Westfalen (Rheinland und Westfalen), Print und Online, ca. 7.500 Titel (806 Einträge in der BEACON-Datei)
#INSTITUTION: Steinheim-Institut
#CONTACT: Harald Lordick <+49(0)201-82162900>
#TIMESTAMP: 2022-04-05T18:24:25.648+02:00
#FEED: http://www.steinheim-institut.de/ebib-djg-nrw/ebib-djg-nrw-beacon.txt
100373550
10076410X
100880363
100969984
101140614
1011477629
```

In diesem Fall werden nur die Normdatennummern genannt, die den Target IDs entsprechen.

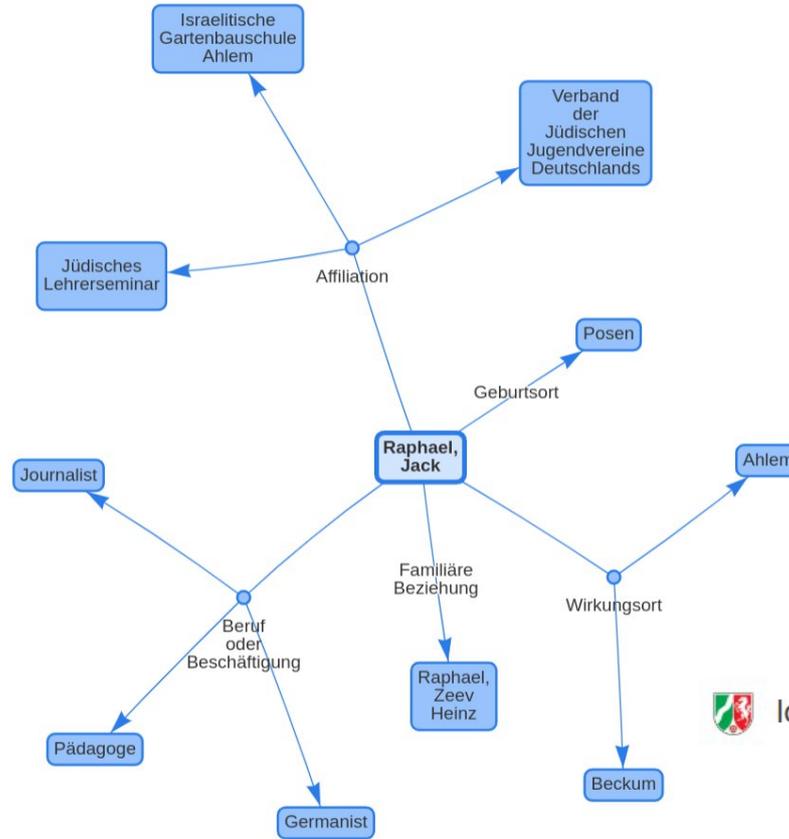
# See Also

- Eine Perspektive auf das Thema „BEACON“: Recherche-Netzwerke
- Entitäten wie etwa
  - Personen
  - Geografika
  - Körperschaften
  - Werke
  - Sachbegriffe / Themen / Phänomene

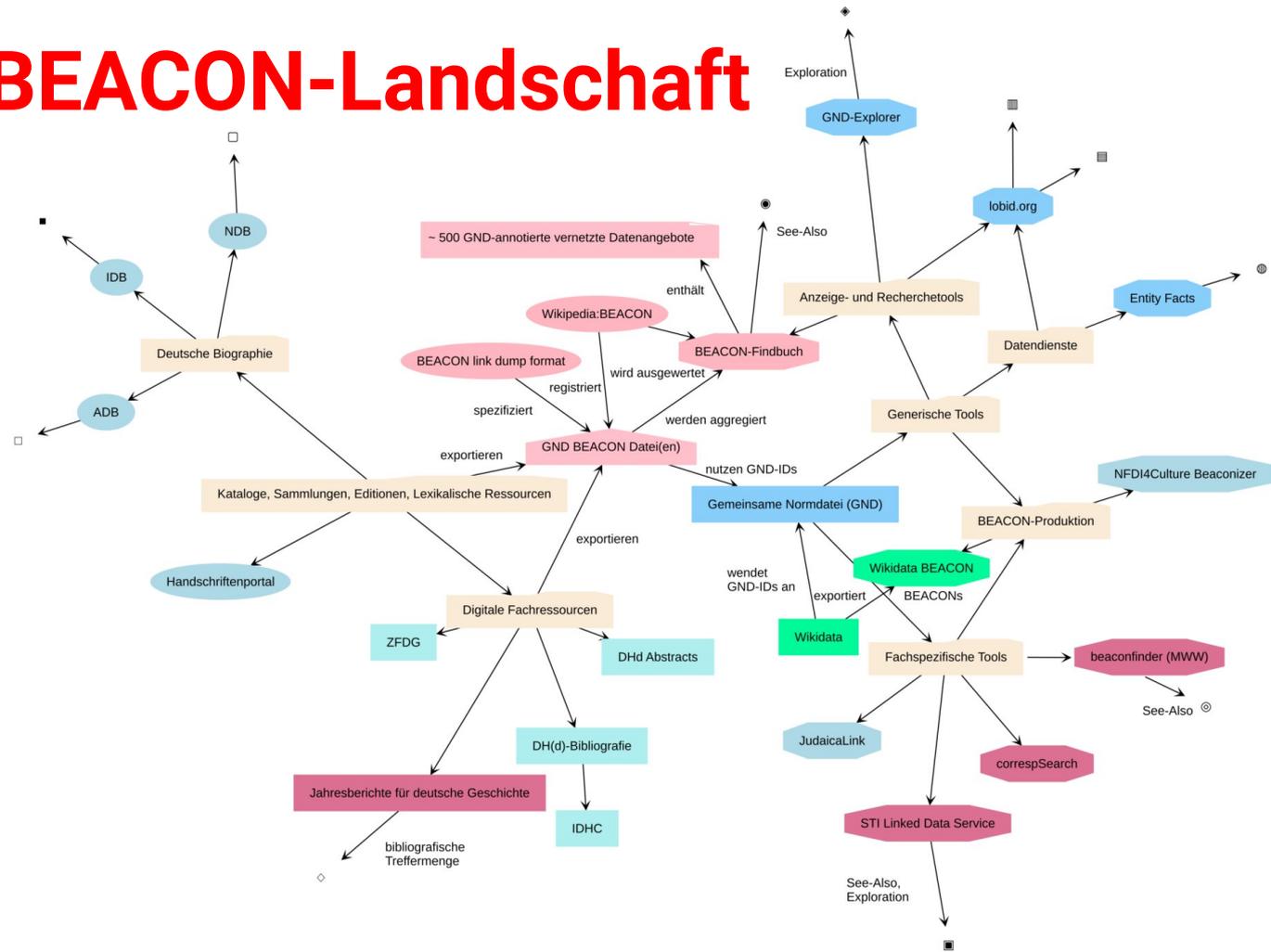
# See Also: Durch eindeutige Identifikation



# See Also: ... eindeutige Identifikation



# See Also: BEACON-Landschaft



# See Also

## 🔗 (STI Linked Data)<sup>α</sup>

ID eingeben und **Suchen** betätigen

ID eingeben

**Normdatei**

GND

**Suchen**

URLfragr {GND} URLfragr

---

Beispiel

Ludwig Philippson

---

Bookmarklet installieren

See-also STI

(Button in die Symbolleiste des Browsers ziehen)

»Philippson, Ludwig«

28. Dezember 1811 Dessau – 29. Dezember 1889 Bonn

Dt. Rabbiner, Schriftsteller und Publizist

GND-ID [116176288](#) (Person)



2022-04-11 | 14:47:10

**Ressourcen im Steinheim-Institut zur GND-ID [116176288](#)**

- Posener jüdische Publizistik | Bibliografie
- Kalonymos. Beiträge zur deutsch-jüdischen Geschichte aus dem Salomon Ludwig Steinheim-Institut an der Universität Duisburg-Essen
- Bibliografie deutsch-jüdische Geschichte Nordrhein-Westfalen
- Gidal-Bildarchiv
- Digitale Edition der Universal-Kirchenzeitung
- Deutsch-jüdische Geschichte digital (ISSN 2363-6343)
- Posener Heimat deutscher Juden (ISSN 2367-2285)
- Epigraphische Datenbank epidat
- Biographisches Handbuch der Rabbiner

**Weitere Fachangebote ...**

- GND-Einträge in der Judaica-Sammlung der UB Frankfurt/M.

**Externe Datenangebote zur GND-ID [116176288](#)**

-  [Ibid](#) (Rechercheoberfläche für die GND)
- Deutsche Biographie: Informationen zu Personen im Index von NDB, ADB und Register
- [BEACON-Findbuch](#)

STI-Domäne (GND)

Sachgruppe(n)

Personen zu allgemeiner und vergleichender Religionswissenschaft, Personen zu nichtchristlichen Religionen

- 1: Wolf, Albert
- 2: Bamberger, Simon
- 3: Cohn, Emil
- 4: Philippson, Ludwig
- 5: Prinz, Joachim
- 6: Brann, Marcus
- 7: ...

**➤ Berufe / Funktionen**

- Rabbiner
- Zeitungsverleger
- Autor

**➤ Personen**

- Geschwister
  - Philippson, Phöbus
- Kinder
  - Philippson, Martin
  - Philippson, Alfred

**➤ Wirkungsort(e)**

- Bonn
- Magdeburg

**➤ Werk(e)**

- Fragmente. - 1830
- Aristotelis doctrina de sensibus. - 1831
- Hylē antrōpinē. - 1831
- Theophrastus: Fragmenta de sensu, phantasia et intellectu. - 1831
- Kleines israelitisches Gesangbuch. - 1855
- Jakob Tirado. - 1867
- Die Enthronen. - 1869
- Die Hochschule für die Wissenschaft des Judentums. - 1872

**➤ Sachgruppe(n)**

- Personen zu allgemeiner und vergleichender Religionswissenschaft, Personen zu nichtchristlichen Religionen
- Personen zu Presse

**➤ andere Namen**

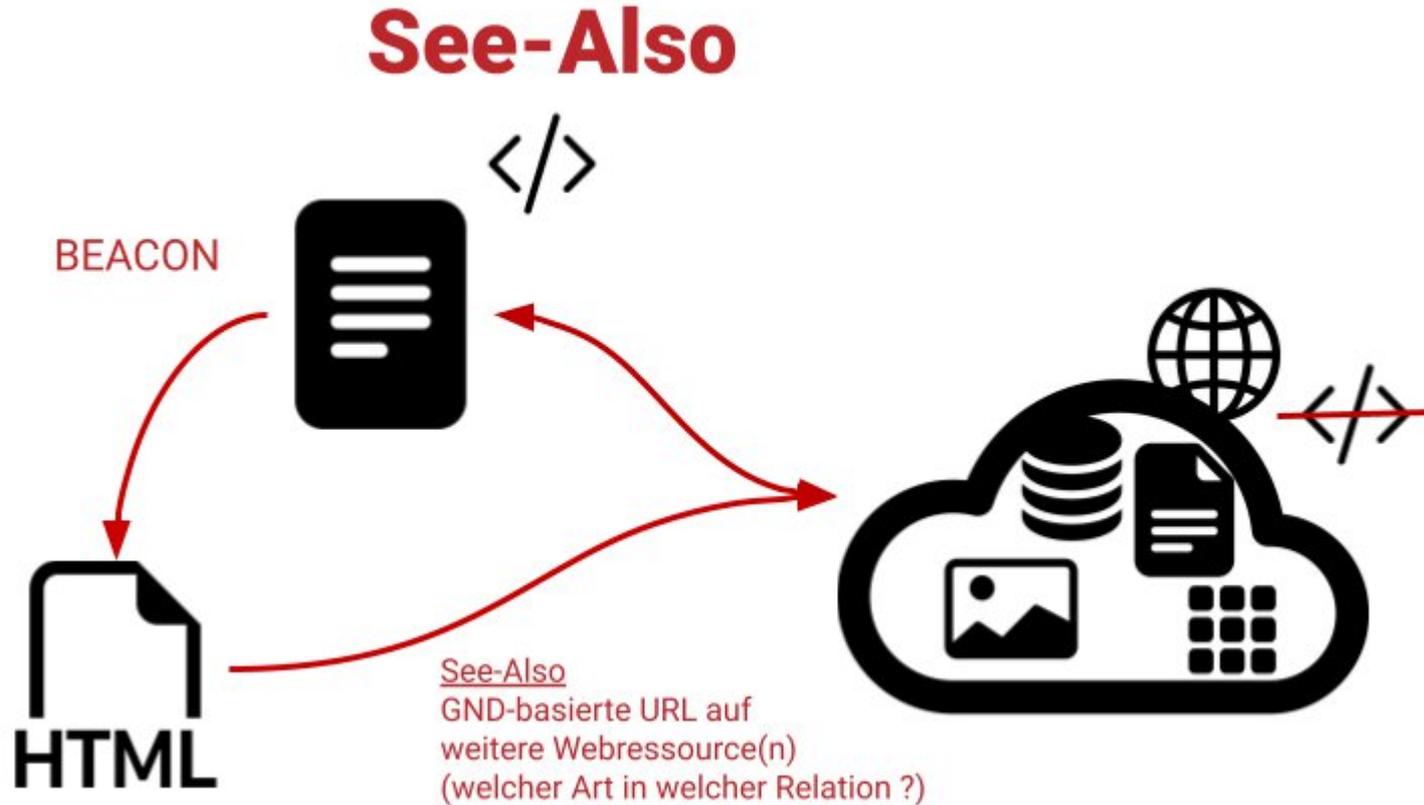
- Filipson, Yehudah Ludvig ben Mosheh
- Yehudah Ludvig ben Mosheh Filipson
- Philipzohn, Ludvig
- Ludvig Philipzohn
- Filipson, L.
- L. Filipson
- Philipzohn, L.
- L. Philinzohn

- Filipson, Yehudah Ludvig ben Mosheh
- Yehudah Ludvig ben Mosheh Filipson
- Philipzohn, Ludvig
- Ludvig Philipzohn
- Filipson, L.
- L. Filipson
- Philipzohn, L.
- L. Philipzohn
- Philippson, Ludw.
- Ludw. Philippson
- Yehûda Ben-Mosheh
- Yehûda Ben-Mosheh
- Schragge, Ludwig
- Ludvig Schragge
- Philippson, Ludwig M.
- Ludwig M. Philippson
- Philippson, Ludovicus
- Ludovicus Philippson
- Philippson, Jehûdâ
- Jehûdâ Philippson
- Filipson, Lúdvig
- Lúdvig Filipson
- פִּילִיפּוֹן, יהודה לודוויג בן משה

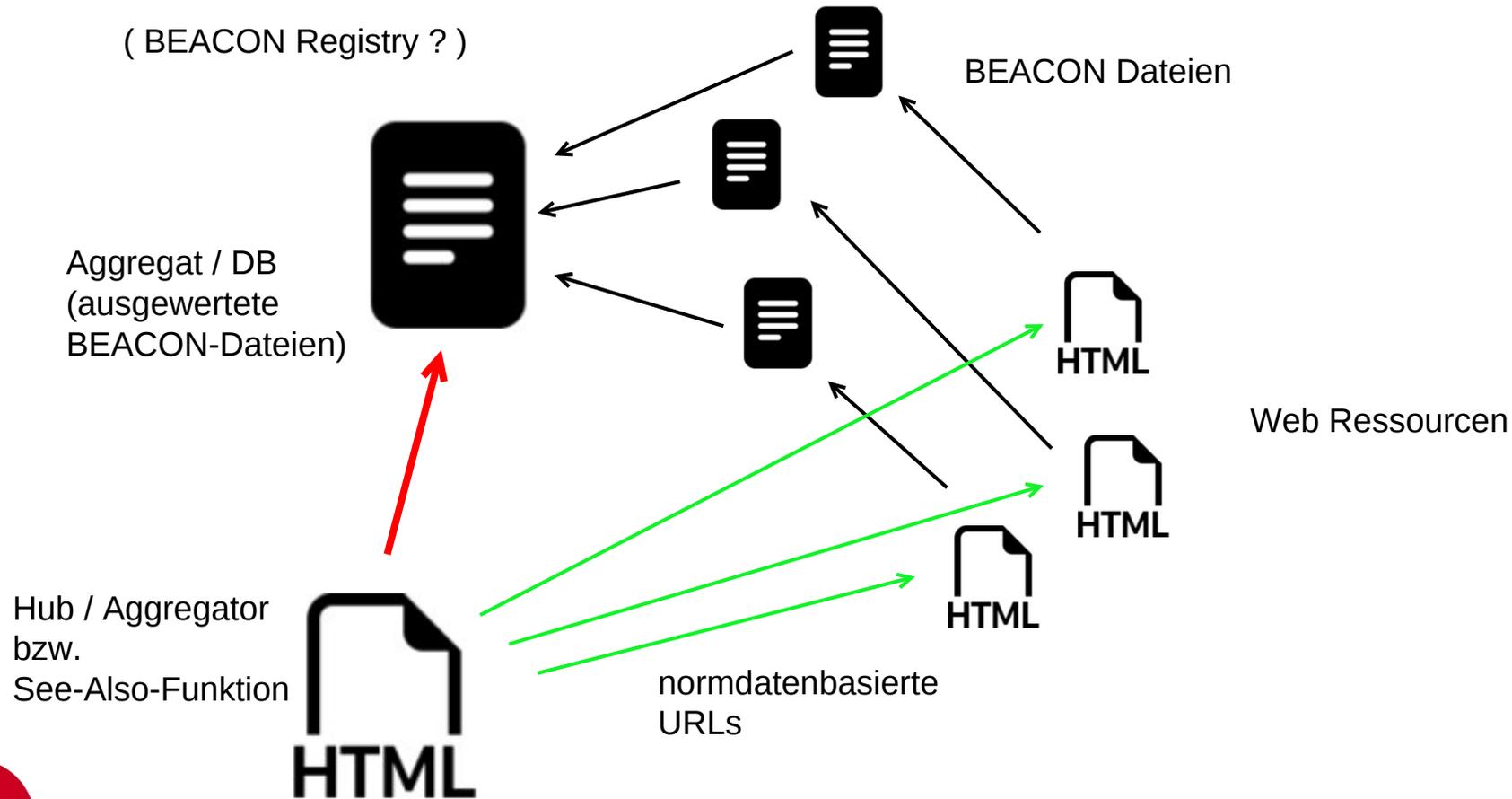
# See Also: »Leuchtfeuer«



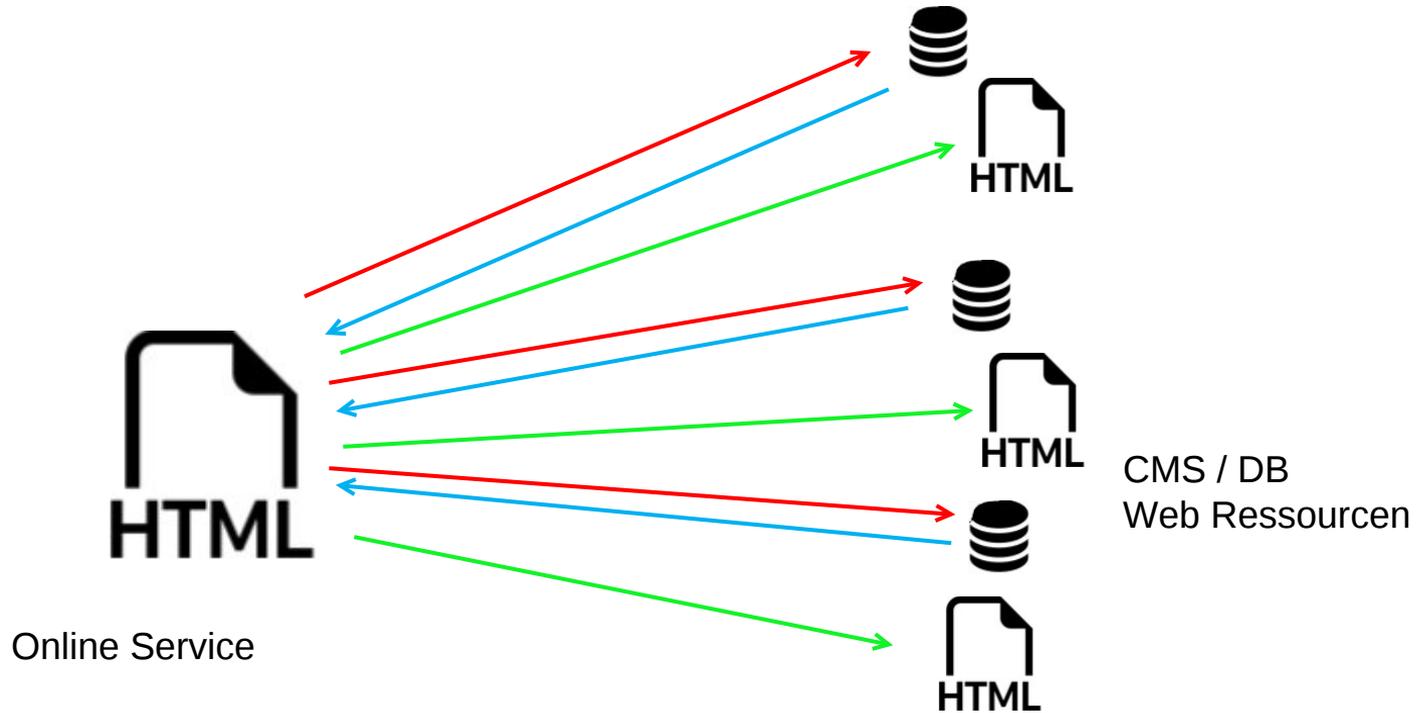
# See Also: Normdatenbasiert recherchieren



# See Also: BEACON-Dateien / Aggregatoren



# See Also: Alternative dynam. Schnittstellen ?



# See Also: BEACON vs. Live-Requests

- Dezentralität
- Resilienz
- Green DH / Green Computing / Green IT /
- FDM-Perspektive: nachhaltige Speicherbarkeit
- Ressourcenaufwand

# See Also: ... das macht doch alles KI !

Eine BEACON-Datei ist eine Datei, die zu Entitäten, die in einem Text vorkommen, die IDs der Gemeinsamen Normdatei (GND) auflistet. Sie enthält zudem einige weitere Metadatenfelder wie z.B. #TARGET. Könntest Du bitte zu folgendem Text die BEACON-Datei erstellen?

Gerne! Ich helfe dir dabei, eine BEACON-Datei zu erstellen. Dafür benötige ich jedoch den Text, den du verwenden möchtest. Kannst du mir diesen zur Verfügung stellen?



# See Also: ... das macht doch alles KI !

Ich habe nun den Text und werde die BEACON-Datei mit den relevanten GND-IDs für die genannten Entitäten erstellen. Damit ich dir präzise helfen kann, werde ich zunächst die Entitäten im Text identifizieren und anschließend die GND-IDs eintragen. Einige Entitäten, wie Paul Frank, Dr. Hugo Neumann oder der Central-Verein Deutscher Staatsbürger jüdischen Glaubens (CV), sollten leicht zuzuordnen sein. Ich werde dir den nächsten Schritt der BEACON-Datei generieren.

Hier ist ein Entwurf der BEACON-Datei, die du für die genannten Entitäten verwenden kannst:

# See Also: ... das macht doch alles KI !

## BEACON-Datei (Beispiel)

plaintext

Kopieren

# BEACON-Datei: Juden in der sozialen Hilfsarbeit von Paul Frank

#TARGET: 1924

#DATE: 1924-04-03

#SOURCE: Central-Verein-Zeitung 3 (1924), 14 (3.4.1924), s. 162-163

[Entitäten]

1. Paul Frank: GND-ID: 118750554
2. Dr. Hugo Neumann: GND-ID: 118720307
3. Central-Verein Deutscher Staatsbürger jüdischen Glaubens (CV): GND-ID: 118664268
4. Lina Morgenstern: GND-ID: 118528362
5. Henriette Goldschmidt: GND-ID: 118569520
6. Alice Salomon: GND-ID: 118602320
7. Jeanette Schwerin: GND-ID: 118596229
8. Albert Levy: GND-ID: 118570380
9. Josephine Levy-Rathenau: GND-ID: 118570682
10. Bertha Pappenheim: GND-ID: 118588406
11. Helene Simon: GND-ID: 118562409
12. Emil Münsterberg: GND-ID: 118613316
13. Dr. Paul Koehne: GND-ID: 118657462
14. Dr. E. Friedeberg: GND-ID: 118656474
15. Rosenstock (Königsberg): GND-ID: 118742589

# See Also: ... das macht doch alles KI !

```
    "description": "Sozialreformer und Mitbegründer der Zentralen f 📄 Kopieren  
  },  
  {  
    "entity": "Hugo Neumann",  
    "gnd_id": "GND101112233",  
    "type": "Person",  
    "description": "Kinderarzt und Gründer des Muster-Kinderhauses."  
  }  
],  
"metadata": {  
  "target": "Juden in der sozialen Hilfsarbeit, Paul Frank, 1924",  
  "source": "Central-Verein-Zeitung 3 (1924), 14 (3.4.1924), s. 162-163",  
  "version": "1.0"  
}  
}  
}
```

# See Also: ... das macht doch alles KI !

## Erläuterungen:

1. **Entities:** Hier wird jede relevante Entität im Text (z.B. Personen, Begriffe, Organisationen) aufgelistet.
2. **GND\_ID:** Diese ID verweist auf den jeweiligen Eintrag in der GND-Datenbank (die Platzhalter-IDs müssen durch die korrekten GND-IDs ersetzt werden).
3. **Description:** Eine kurze Beschreibung, die den Kontext der Entität im Text erläutert.
4. **Metadata:** Zusätzliche Informationen, wie die Quelle des Textes und die Zielsetzung (TARGET).

Damit die BEACON-Datei korrekt funktioniert, müssen die GND-IDs der jeweiligen Entitäten aus der Gemeinsamen Normdatei (GND) besorgt und in den Platzhalterstellen ersetzt werden. Diese Daten können meist über Bibliotheken oder Archivdatenbanken bezogen werden.

# See Also: ... das macht doch alles KI !

Hilf mir, eine BEACON-Datei zu erstellen.

Eine BEACON-Datei wird häufig im Zusammenhang mit der Bioinformatik verwendet, speziell im Bereich der Genomforschung und der Datenanalyse. Es handelt sich um ein Format, das in der Regel zur Darstellung von genetischen Varianten genutzt wird, etwa bei der Arbeit mit VCF-Dateien (Variant Call Format).

Könntest du ein bisschen mehr Kontext geben? Möchtest du eine BEACON-Datei für genetische Daten erstellen, oder geht es um eine andere Art von BEACON-Datei? Wenn es um genetische Daten geht, kann ich dir zeigen, wie man eine BEACON-Datei strukturiert, welche Felder enthalten sein sollten und wie man sie erstellt.

Bitte gib mir mehr Details, damit ich dir gezielt helfen kann!

# See Also: ... das macht doch alles KI !

Gib mir bitte ein Biogramm der Sozialarbeiterin Bertha Pappenheim. Füge noch Links auf relevante Online-Angebote zu seiner Person hinzu. Die Links sollen auf der ID von Bertha Pappenheim in der Gemeinsamen Normdatei (GND) basieren.



Bertha Pappenheim (27. Februar 1859 – 28. Mai 1936) war eine österreichisch-jüdische Sozialarbeiterin, Frauenrechtlerin und Pionierin der sozialen Arbeit. Sie wurde als "Anna O." bekannt, einer der ersten dokumentierten Fälle in der Psychoanalyse, behandelt von Josef Breuer. EN.WIKIPEDIA.ORG

# See Also: ... das macht doch alles KI ?

Für weiterführende Informationen über Bertha Pappenheim können folgende Online-Ressourcen konsultiert werden:

- [Bertha Pappenheim – Leo Baeck Institute](#)
- [Bertha Pappenheim – Digitales Deutsches Frauenarchiv](#)
- [Bertha Pappenheim – Jewish Women's Archive](#)
- [Bertha Pappenheim – Wikipedia](#)
- [Bertha Pappenheim's Gender Revolution, Religion, and Modernity – Leo Baeck Institute London](#)

Diese Links basieren auf der Gemeinsamen Normdatei (GND) und bieten vertiefte Einblicke in ihr Leben und Werk.

# Beispieldaten

- [Fernando Pessoa – Digitale Edition Projekte und Publikationen](#)
- Kooperationsprojekt (2023) – Ulrike Henny-Krahmer (Universität Rostock):

ergänzende Anreicherung der Registerdaten mit Normdaten; Auslieferung der Registerdaten im BEACON- und RDF-Format

- Register Typen:

## Namen

Namen der in den Texten der Edition erwähnten Personen

## Titel

Titel von FPs veröffentlichten Texten

## Zeitschriften

Zeitschriften, die FPs Texten veröffentlicht haben

The screenshot shows a website interface with a dark red navigation bar containing the following items: Autoren, Dokumente, Publikationen, Genres, Chronologie, Register, and Projekt, followed by a search icon. Below this is a secondary menu with the items: Bibliographie, Namen, Titel, Zeitschriften, and LOD. The main content area is titled "Linked Open Data (LOD)" and contains the following text:

Auf dieser Seite finden Sie Links zu Dateien, die Ressourcen aus der Digitalen Edition von Fernando Pessoa in Linked Open Data-Formaten anbieten. Die folgenden Arten von Ressourcen sind in den LOD-Dateien enthalten:

- Namen von Personen und Figuren, die in den Dokumenten und Publikationen der digitalen Edition vorkommen
- Namen von Periodika, die in den Dokumenten und Publikationen der digitalen Edition vorkommen
- Literarische Werke, die in der digitalen Edition in Form von Form von Publikationen vorkommen

Zurzeit wird das Vorkommen dieser Entitäten in einer allgemeinen Weise angegeben, ohne Angabe der jeweiligen einzelnen Quelle, in der sie vorkommen.

Diese Daten sind mit drei Normdatenverzeichnissen verknüpft: die Gemeinsame Normdatei der Deutschen Nationalbibliothek (GND), die Virtuelle Internationale Authority File (VIAF), und Wikidata. Die verlinkten Daten werden in zwei Formaten angeboten, RDF und BEACON, und können hier in ihrer Gesamtheit eingesehen und heruntergeladen werden:

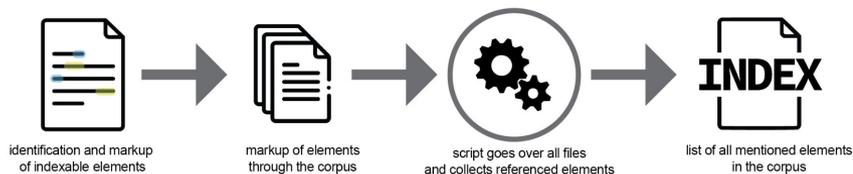
RDF  
BEACON

# Beispieldaten

[BNP 731r](#): eine Liste der übersetzten Texten (von Pessoa)

```
<rs type="name" key="P87">Edgar Poe</rs>
```

```
<item>  
→ <rs type="title">DESCIDA AO MAELSTROM</rs>, de <rs type="name" key="P87">Edgar  
Poe</rs>.</item>  
<item>  
  <rs type="title">A CAVALGADA ATÉ AO MAR</rs>, de <rs type="name" key="P270">  
  <choice>  
    <abbr>J<am>.</am> M<am>.</am>  
    </abbr>  
    <expan>J<ex>ohn</ex> M<ex>illington</ex>  
    </expan>  
  </choice> Syngde</rs>.</item>  
<item>  
  <rs type="title">A TERRA DOS CEGOS</rs>, de <rs type="name" key="P271">  
  <choice>  
    <abbr>H<am>.</am> G<am>.</am>  
    </abbr>  
    <expan>H<ex>erbert</ex> G<ex>eorge</ex>  
    </expan>  
  </choice> Wells</rs>.</item>  
<item>  
  <rs type="title">A PORTA NO MURO</rs>, de <rs type="name" key="P271">  
  <choice>  
    <abbr>H<am>.</am> G<am>.</am>  
    </abbr>  
    <expan>H<ex>erbert</ex> G<ex>eorge</ex>  
    </expan>  
  </choice> Wells</rs>.</item>  
</item>
```



# Beispieldaten: Register

- `indices.xml` Datei verfügbar in Github Repository:

<https://github.com/cceh/pessoa-corpus>

- Titel, Zeitschriften und Namen (Personen) mit Normendaten *enriched*:
  - GND – Gemeinsame Normdatei der Deutsche Nationalbibliothek
  - Wikidata
  - Vial – *Virtual International Authority File*

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Digital Edition of Fernando Pessoa - Project Knowledge</title>
      </titleStmt>
      <editionStmt>
        <p>Corpus version C2.1.1</p>
      </editionStmt>
      <publicationStmt>
        <p>Publication Information</p>
      </publicationStmt>
      <sourceDesc>
        <p>Born digital.</p>
      </sourceDesc>
    </fileDesc>
    <revisionDesc> [7 lines]
  </teiHeader>
  <text>
    <body>
      <div xml:id="Indices">
        <!-- ##### COMPLETE WORKS ##### -->
        <!-- mentioned works: mentioned titles, collections, translations, works -->
        <list type="titles"> [1346 lines]
        <!-- collections -->
        <list type="collections"> [52 lines]
        <!-- translations -->
        <list type="translations"> [5 lines]
        <!-- abstract works: includes published, unpublished or envisioned works by Pessoa -->
        <list type="works"> [1612 lines]
        <!-- publications: published works by Pessoa -->
        <list type="publications"> [668 lines]
        <!-- ##### JOURNALS ##### -->
        <list type="periodical"> [393 lines]
        <!-- ##### PERSONS ##### -->
        <listPerson type="all"> [3957 lines]
      </div>
    </body>
  </text>
```

# Beispieldaten: Titel

```
<item xml:id="Caeiro_De_O_Guardador_de_Rebanhos">
  <title>De "O Guardador de Rebanhos"</title>
  <idno type="wikidata">Q10339214</idno>
  <idno type="viaf">182801150</idno>
</item>
```

Viaf:

Colecção O guardador de rebanhos   
O guardador de rebanhos   
VIAF ID: 182801150 (Work)  
Permalink: <http://viaf.org/viaf/182801150> 

---

Preferred Forms

 130 0  Colecção O guardador de rebanhos  
 130 2  O guardador de rebanhos

Wikidata:

**The Keeper of Sheep** (Q10339214)

poetry collection by Fernando Pessoa 

[In more languages](#)  
[Configure](#)

Language	Label	Description	Also known as
default for all languages	No label defined	–	
English	The Keeper of Sheep	poetry collection by Fernando Pessoa	
German	No label defined	No description defined	
French	No label defined	No description defined	
Bavarian	No label defined	No description defined	

# Beispieldaten: Zeitschriften

```
<item xml:id="J14">
  <title>Século</title>
  <idno type="viaf">180234301</idno>
  <idno type="gnd">7567219-4</idno>
  <idno type="wikidata">Q3347813</idno>
</item>
```

Viaf:

O século (Sextante) 

VIAF ID: 180234301 (Work)

Permalink: <http://viaf.org/viaf/180234301> 

Wikidata:

O Século (Q3347813)

Portuguese daily newspaper 

O Seculo

[+ In more languages](#)

Configure

Language	Label	Description	Also known as
default for all languages	No label defined	–	
English	O Século	Portuguese daily newspaper	O Seculo
German	No label defined	No description defined	
French	O Século	No description defined	
Bavarian	No label defined	No description defined	

GND:

Werk (wit)

## O @século (Tageszeitung, Lissabon, 1881-1977)

1881-1977

---

**Titel** 

Bevorzugter Titel: Titel des Werks O @século Zusatz Tageszeitung, Lissabon, 1881-1977  
Früherer bevorzugter Titel: swd t Alte Ansetzungsform O @século <Tageszeitung, Lissabon, 1881-1977> Alte Normnummer 7567219-4

---

**Beschreibende Angaben** 

Sprachencode: por  (Portugiesisch)   
Entitätentyp: wit  (Werk) 

---

**Zeit** 

Gründungs- und Auflösungsdatum: Beginn 1881 Ende 1977 (Relation datb) 

---

**Geografischer Bezug** 

Land: [XA-PT](#)  (Portugal)   
Geografikum allgemein:  Portugal (Relation geoa) 

---

**Identifikatoren** 

GND: 7567219-4  
GND-URI <http://d-nb.info/gnd/7567219-4>   
Andere Identifikatoren: PPN 983883548  
Alte Identifikatoren: swd Alte Normnummer 7567219-4 (Bemerkungen zg)

---

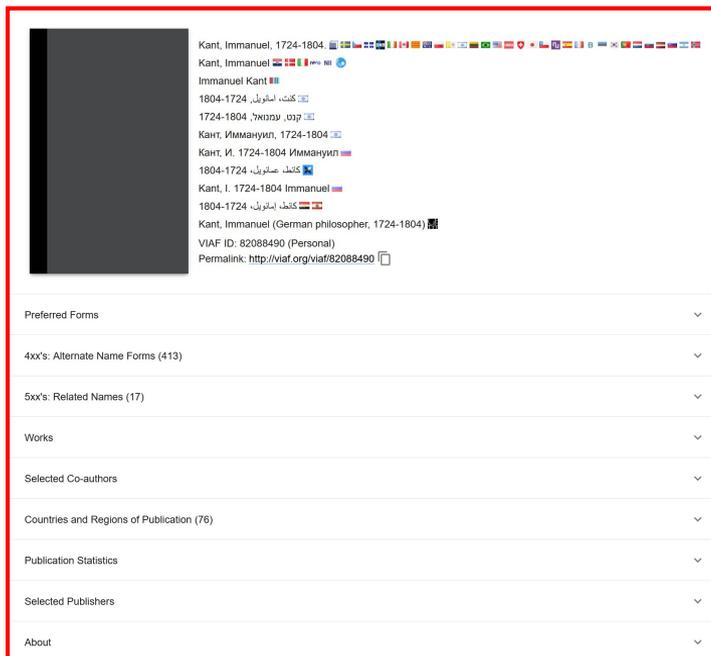
**Oberbegriffe** 

Beispiel für: [Zeitung](#) (Relation obin) 

# Beispieldaten: Namen

```
<person xml:id="P6">
  <persName type="main">Immanuel Kant</persName>
  <idno type="viaf">82088490</idno>
  <idno type="gnd">118559796</idno>
  <idno type="wikidata">Q9312</idno>
  <note type="ontological-status">real</note>
  <sex>male</sex>
</person>
```

Viaf:



Kant, Immanuel, 1724-1804. 

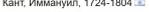
Kant, Immanuel 

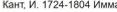
Immanuel Kant 

1804-1724. 

1724-1804. 

Кант, Иммануил, 1724-1804 

Кант, И. 1724-1804 Иммануил 

1804-1724 

Kant, I. 1724-1804 Immanuel 

1804-1724 

Kant, Immanuel (German philosopher, 1724-1804) 

VIAF ID: 82088490 (Personal)

Permalink: <http://viaf.org/viaf/82088490> 

Preferred Forms 

4xx's: Alternate Name Forms (413) 

5xx's: Related Names (17) 

Works 

Selected Co-authors 

Countries and Regions of Publication (76) 

Publication Statistics 

Selected Publishers 

About 



# Beispieldaten: Namen

```
<person xml:id="P6">
  <persName type="main">Immanuel Kant</persName>
  <idno type="viaf">82088490</idno>
  <idno type="gnd">118559796</idno>
  <idno type="wikidata">Q9312</idno>
  <note type="ontological-status">real</note>
  <sex>male</sex>
</person>
```

GND:

 Person (piz)

## Kant, Immanuel

1724-1804; Philosoph; Naturwissenschaftler; Hochschullehrer; Autor; Privatlehrer; Bibliothekar

---

**Namen** 

**Beschreibende Angaben** 

Geschlecht: [m \(männlich\)](#) 

Sprachencode: [ger](#)  (Deutsch) 

Charakteristischer Beruf: [Philosoph](#)  (Relation [beru](#)) 

Beruf: [Naturwissenschaftler](#)  (Relation [beru](#))  [Hochschullehrer](#)  (Relation [beru](#))  [Autor](#)  (Relation [beru](#))  [Privatlehrer](#)  (Relation [beru](#))  [Bibliothekar](#)  (Relation [beru](#)) 

Entitätentyp: [piz](#)  (Person) 

---

**Zeit** 

Lebensdaten: [Beginn 1724 Ende 1804](#)  (Relation [dati](#)) 

Exakte Lebensdaten: [Beginn 22.04.1724 Ende 12.02.1804](#)  (Relation [datx](#)) 

---

**Beziehungen zu Personen** 

Familiäre Beziehung: [Kant, Johann Georg](#)  (Relation [bezf](#) Bemerkungen [Vater](#))  [Kant, Johann Heinrich](#)  (Relation [bezf](#) Bemerkungen [Bruder](#)) 

Bekannschaft: [Feyerabend, Gottfried](#)  (Relation [beza](#) Bemerkungen [Student](#)) 

---

**Beziehungen zu Organisationen** 

**Geografischer Bezug** 

Land: [XA-DE](#)  (Deutschland) 

Geburtsort: [Königsberg](#)  (Relation [ortg](#)) 

Sterbeort: [Königsberg](#)  (Relation [orts](#)) 

Wirkungsort: [Königsberg](#)  (Relation [ortw](#)) 



# Registerdaten

- jeder Registertyp wird in einem `<list>` Element gesammelt; der Typ wird in einem `type`-Attribut definiert  
für Namen ist das `<listPerson>` statt `<list>`
- jede einzelne Einheit wird in einem `<item>` Element beschrieben und mit einer lokalen ID identifiziert, die im `xml:id`-Attribut definiert ist.  
für Namen ist das `<person>` statt `<item>`
- die Normdaten-IDs werden als Textinhalt in `<idno>` Elemente gegeben. Die Normdatenquellen werden in einem `type`-Attribut definiert.

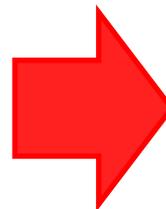
```
<!-- publications: published works by Pessoa -->  
<list type="publications"> [668 lines]  
<!-- ##### JOURNALS ##### -->  
<list type="periodical"> [393 lines]  
<!-- ##### PERSONS ##### -->  
<listPerson type="all"> [3956 lines]
```

```
<item xml:id="J14">  
  <title>Século</title>  
  <idno type="viaf">180234301</idno>  
  <idno type="gnd">7567219-4</idno>  
  <idno type="wikidata">Q3347813</idno>  
</item>
```

```
<person xml:id="P6">  
  <persName type="main">Immanuel Kant</persName>  
  <idno type="viaf">82088490</idno>  
  <idno type="gnd">118559796</idno>  
  <idno type="wikidata">Q9312</idno>  
  <note type="ontological-status">real</note>  
  <sex>male</sex>  
</person>
```

# Registerdaten: *in a nutshell*

```
<list type="[Namen|Titel|Zeitschrift]">
  <item xml:id="[lokalle ID]">
    <idno type="[Viaf|Wikidata|GND]"> [Normdaten] </idno>
    (...)
  </item>
  <item xml:id="[lokalle ID]">
    <idno type="[Viaf|Wikidata|GND]"> [Normdaten] </idno>
    (...)
  </item>
  (...)
</list>
```



```
[Namen|Titel|Zeitschrift]
[lokalle ID] - [Normdaten][Viaf|Wikidata|GND]
[lokalle ID] - [Normdaten][Viaf|Wikidata|GND]
[lokalle ID] - [Normdaten][Viaf|Wikidata|GND]
(...)
Namen:
P6 - 82088490 (Viaf)
P6 - 118559796 (GND)
P6 - Q9312 (Wikidata)
BEACON-Format:
http://viaf.org/viaf/82088490 | P6
https://d-nb.info/gnd/118559796 | P6
https://www.wikidata.org/wiki/Q9312 | P6
```

# Prefix/Target

## GND

Immanuel Kant

<https://d-nb.info/gnd/118559796>

Aristoteles

<http://d-nb.info/gnd/118650130>

Jean-Jacques Rousseau

<http://d-nb.info/gnd/118603426>

#PREFIX: <http://d-nb.info/gnd/>

## Pessoa Digital

Immanuel Kant

<http://www.pessoadigital.pt/index/names#P6>

Aristoteles

<http://www.pessoadigital.pt/index/names#P401>

Jean-Jacques Rousseau

<http://www.pessoadigital.pt/index/names#P434>

#TARGET: <http://www.pessoadigital.pt/index/names#>

# XML2LinkedData

- Verfügbar auf:

<https://github.com/fernandaalvaf/XML2LinkedData>

## wichtige Dateien:

- Beispieldatei: `indices.xml`
- Konfiguration-datei: `config.yaml`
- BEACON Skript: `BEACONgen.py/BEACONgenNotebook.ipynb`

- Direkt nutzen auf:

[https://colab.research.google.com/drive/1VVkEtIQNRo1uTa\\_WBUE1YZ3TyznVyrZs?usp=sharing](https://colab.research.google.com/drive/1VVkEtIQNRo1uTa_WBUE1YZ3TyznVyrZs?usp=sharing)

# config.yaml

## file\_location:

Pfad, in dem sich die XML-Datei befindet.

Wenn die Datei in den *Root-Ordner* des Projekts kopiert wird, reicht der Name der Datei aus

```
! config.yaml
1 file_location: "indices.xml"
2
3 register_types: {'names': {
4     'element': 'listPerson',
5     'attribute_type': 'type',
6     'attribute_value': 'all',
7     'target': 'http://www.pessoadigital.pt/index/names#'},
8     'periodicals': {
9     'element': 'list',
10    'attribute_type': 'type',
11    'attribute_value': 'periodical',
12    'target': 'http://www.pessoadigital.pt/index/periodicals#'},
13    'publications': {
14    'element': 'list',
15    'attribute_type': 'type',
16    'attribute_value': 'publications',
17    'target': 'http://www.pessoadigital.pt/pub/'
18    }
19 }
20 header_data: {'name' : 'Digital Edition of Fernando Pessoa',
21    'contact' : 'Ulrike Henny-Krahmer <email>',
22    'message' : 'Mentions in the digital edition of Fernando Pessoa',
23    }
24
25 authority_files: {'wikidata': 'https://www.wikidata.org/wiki/',
26    'viaf': 'http://viaf.org/viaf/',
27    'gnd': 'http://d-nb.info/gnd/'
28    }
29
30 namespaces: {"tei": "http://www.tei-c.org/ns/1.0",
31    "xml": "http://www.w3.org/XML/1998/namespace"
32    }
33
```

# config.yaml

## register\_types:

Allgemeine strukturelle Informationen über das Projektregister.

Es kann mehr als eine Art von Register angegeben werden, wenn alle sich in einer Datei befinden.

**element:** in welchem TEI-Element werden die Einträge gesammelt?

**attribute\_type:** gibt es in diesem Element ein identifizierendes Attribut? Wenn ja, das Attribut angeben, wenn nicht `None`

**attribute\_value:** was ist der Wert dieses Attributs? Wenn kein Attribut vorhanden ist, einfach `None` angeben

```
      element  attribute_type
      <listPerson type="all">
                attribute_value
```

**target:** Was ist der unveränderliche Teil der URL des Projekts, der zusammen mit den IDs die URI des Eintrags bildet?

```
! config.yaml
1  file_location: "indices.xml"
2
3  register_types: {'names': {
4                    'element': 'listPerson',
5                    'attribute_type': 'type',
6                    'attribute_value': 'all',
7                    'target': 'http://www.pessoadigital.pt/index/names#'},
8                    'periodicals': {
9                    'element': 'list',
10                   'attribute_type': 'type',
11                   'attribute_value': 'periodical',
12                   'target': 'http://www.pessoadigital.pt/index/periodicals#'},
13                   'publications': {
14                   'element': 'list',
15                   'attribute_type': 'type',
16                   'attribute_value': 'publications',
17                   'target': 'http://www.pessoadigital.pt/pub/'}
18                   }
19
20  header_data: {'name': 'Digital Edition of Fernando Pessoa',
21               'contact': 'Ulrike Henny-Krahmer <email>',
22               'message': 'Mentions in the digital edition of Fernando Pessoa',
23               }
24
25  authority_files: {'wikidata': 'https://www.wikidata.org/wiki/',
26                  'viaf': 'http://viaf.org/viaf/',
27                  'gnd': 'http://d-nb.info/gnd/'
28                  }
29
30  namespaces: {"tei": "http://www.tei-c.org/ns/1.0",
31              "xml": "http://www.w3.org/XML/1998/namespace"
32              }
33
```







# BEACONgen.py

1. Einlesen der Konfigurationsdatei (`config.yaml`)
2. Validierung der XML Datei
3. Extraktion der Normdaten per Quelle
4. Verarbeitung der Normdaten
5. Protokollierung von fehlender IDs
6. Erstellung der BEACON-Dateien

# Übung 1

```
#FORMAT: BEACON
#NAME: Digital Edition of Fernando Pessoa
#PREFIX: http://viaf.org/viaf/
#TARGET: http://www.pessoadigital.pt/index/names#
#CONTACT: Ulrike Henny-Krahmer <email>
#MESSAGE: Mentions in the digital edition of Fernando Pessoa
#RELATION: http://www.w3.org/2002/07/owl#sameAs
#TIMESTAMP: 2025-02-19T11:08:31
```

<https://www.w3.org/2000/01/rdf-schema#seeAlso>

```
rdfs:seeAlso a rdf:Property ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "seeAlso" ;
  rdfs:comment "Further information about the subject resource." ;
  rdfs:range rdfs:Resource ;
  rdfs:domain rdfs:Resource .
```

<https://www.w3.org/2002/07/owl#sameAs>

```
owl:someValuesFrom a rdf:Property ;
  rdfs:label "someValuesFrom" ;
  rdfs:comment "The property that determines the class that an existential property restriction refers to." ;
  rdfs:domain owl:Restriction ;
  rdfs:isDefinedBy <http://www.w3.org/2002/07/owl#> ;
  rdfs:range rdfs:Class .
```

**Wie können wir den Wert von RELATION, der im Header angezeigt wird, von rdfs:seeAlso in owl:sameAs ändern?**

# Übung 2

Jetzt versuchen wir, das Skript auf einen anderen Datensatz anzuwenden.

- Neue Beispieldaten: `Data/Orte.xml`
- Quelle:

„Lebenswelten, Erfahrungsräume und politische Horizonte der ostpreußischen Adelsfamilie Lehndorff vom 18. bis in das 20. Jahrhundert. Bearbeitet von Gaby Huch. Berlin-Brandenburgische Akademie der Wissenschaften 2019. URL: <https://lebenswelten-lehndorff.bbaw.de>“

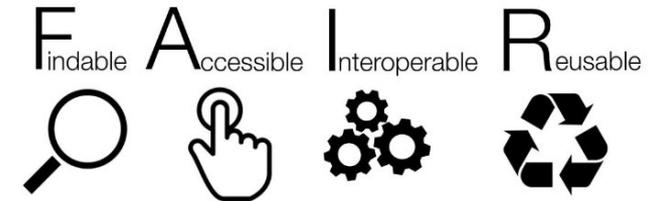
**Wie sollten wir die `config.yaml` bearbeiten, damit sie zu den neuen Daten passt?**

# Diskussion

- Saubere Daten vs. saubere Skripte?
- Verwendung verschiedener Normdateien?
- Hilft ChatGPT?

# Programm FAIR February 2025

- 5. Februar 2025 – **F** wie Qualitätssicherung
- 12. Februar 2025 – **A** wie User Experience
- 19. Februar 2025 – **I** wie BEACON
- **26. Februar 2025 – R wie Dokumentation**



Graphics: SangyaPundir - Own work, CC BY-SA 4.0,  
<https://commons.wikimedia.org/w/index.php?curid=53414062>