

EADH 2018
Data In Digital Humanities
Galway, 2018

Building a domain specific Research Ontology from external Databases of Academic History

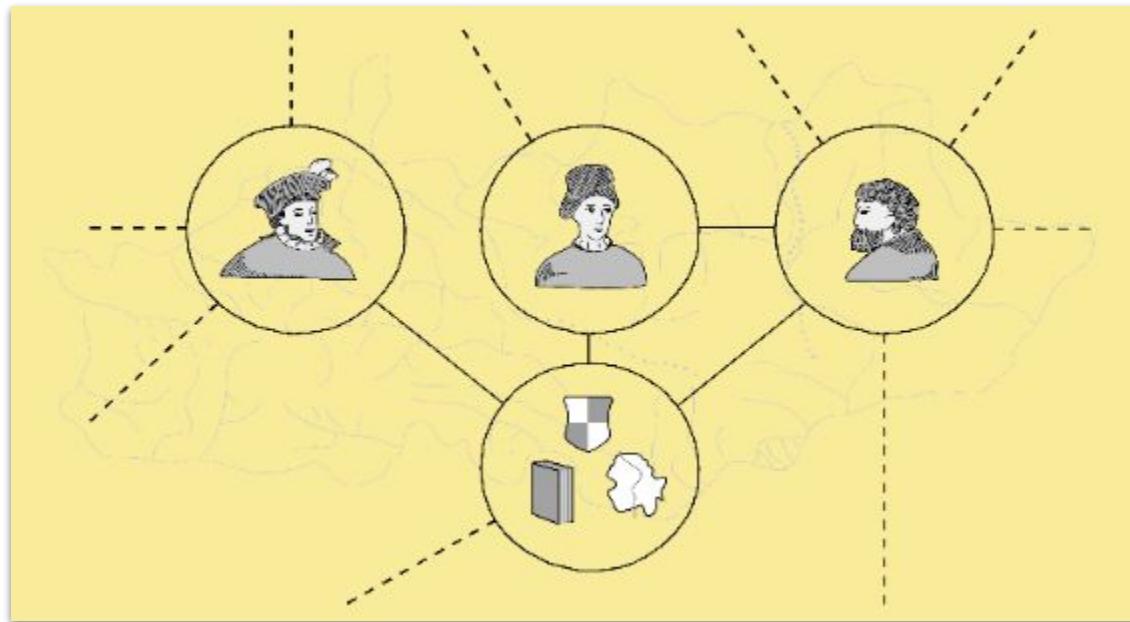
Thomas Riechert, Jennifer Blanke, **Edgard Marx**

Outline

- **Introduction**
- **Why?**
- **Data (Research) Life Cycle**

Professors' Career Patterns on the Web

Early modern university professors (German Protestant)



- Relevant criterias for early modern (German protestant) professors' recruitment



- Structured (You can access to its content through formal languages)

The screenshot shows a Google search results page for the query "nike inc". The results are categorized by type: Web, News, Maps, Images, Shopping, More, and Search tools. The first result is a local search for "Nike Portland" with a snippet about finding favorite Nike products and a link to the website. The second result is "Nike News - The official news website for NIKE, Inc." with a snippet about reading news and press releases. The third result is "NIKE, Inc.— Inspiration and Innovation for Every Athlete in..." with a snippet about experiencing sports, training, shopping, and corporate responsibility. A red arrow points from the third result to the right-hand sidebar. The sidebar contains detailed information about Nike, Inc., including its stock price (\$93.10), foundation date (January 25, 1964), customer service number (1 (800) 344-6453), CEO (Mark Parker), headquarters (Beaverton, OR), founders (Bill Bowerman, Phil Knight), and social media links for Facebook, Twitter, LinkedIn, Instagram, and Google+.

Google search results for "nike inc":

- Nike Portland**
Ad www.nike.com/NikePortland ⓘ
4.4 ★★★★★ rating for nike.com
Find Your Favorite Nike Products at Nike Portland.
Nike has 3,261,712 followers on Google+
Men's Gear Women's Gear
Jordan Shoes Custom Nike Shoes
- Nike News - The official news website for NIKE, Inc.**
news.nike.com/ ⓘ
Read the latest NIKE, Inc. news and press releases, search for jobs, find stock information and learn about corporate responsibility.
Running - Football (Soccer) - Basketball - Nike Sportswear
- NIKE, Inc.— Inspiration and Innovation for Every Athlete in...**
www.nike.com/ ⓘ Nike, Inc. ⓘ
Experience sports, training, shopping and everything else that's new at Nike.
Nike - Nike.com (ES) - Nike.com (UK) - Nike.com (PT)
- Careers at Nike | Corporate - Nike Jobs**
jobs.nike.com/article/Corporate ⓘ Nike, Inc. ⓘ
Nike corporate employees shatter limits daily at company headquarters and offices ... A corporate career at Nike opens countless opportunities for employees to ...

Nike, Inc.
Footwear Manufacturing company

Nike, Inc. is an American multinational corporation that is engaged in the design, development, manufacturing and worldwide marketing and selling of footwear, apparel, equipment, accessories and services. [Wikipedia](#)

Stock price: NKE (NYSE) \$93.10 -0.61 (-0.65%)
Jan 15, 1:50 PM EST - Disclaimer

Founded: January 25, 1964

Customer service: 1 (800) 344-6453

CEO: Mark Parker

Headquarters: Beaverton, OR

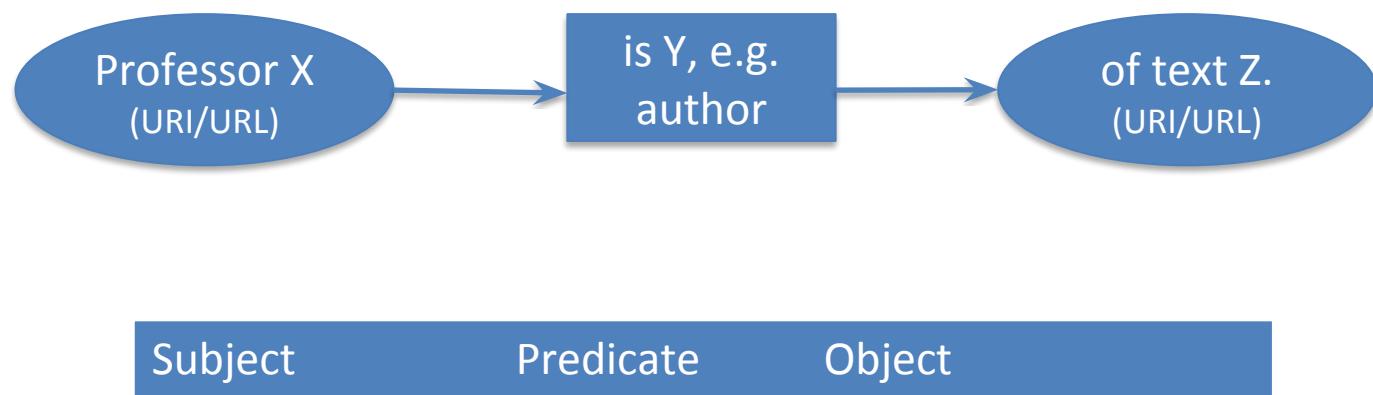
Founders: Bill Bowerman, Phil Knight

Profiles

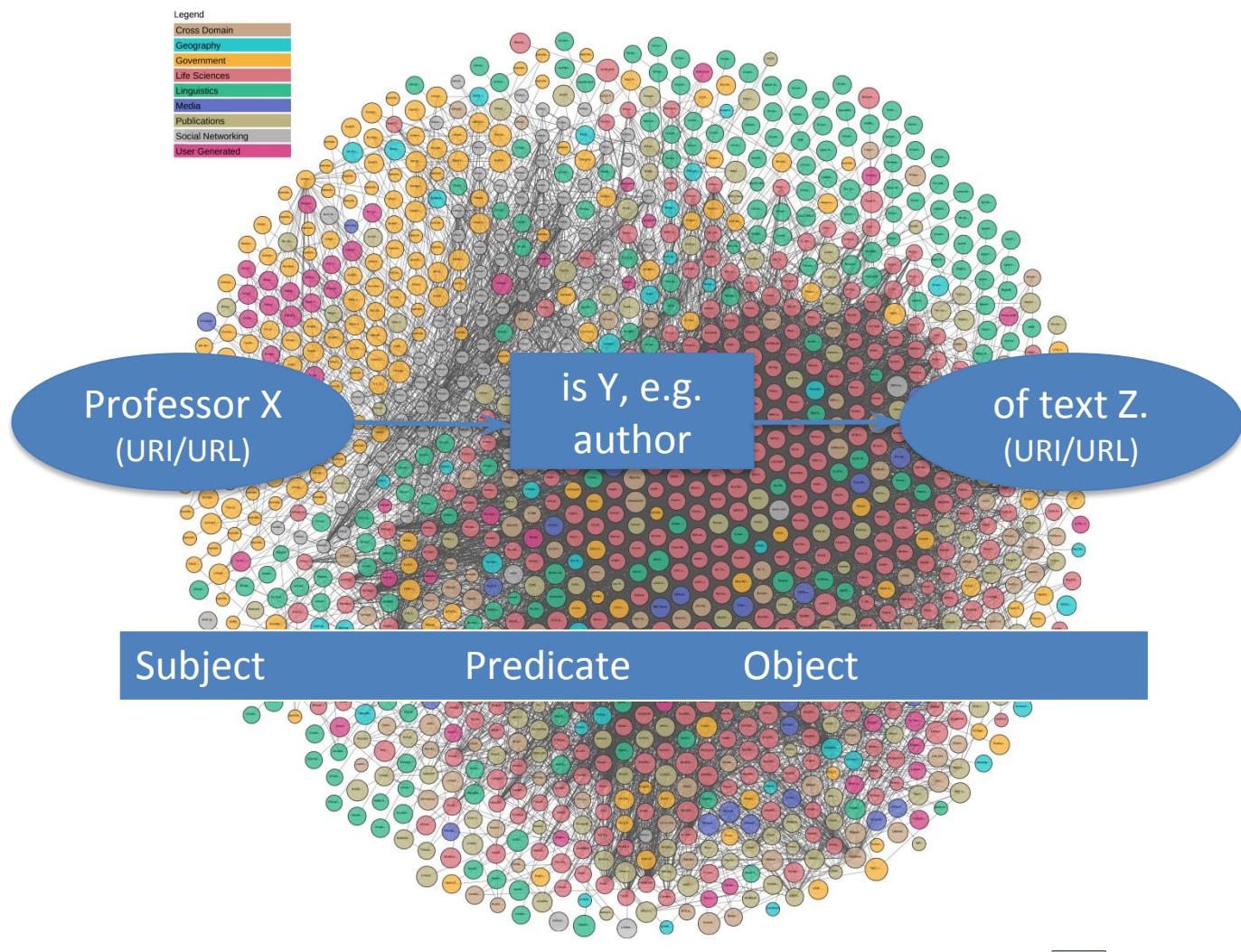
Facebook Twitter LinkedIn Instagram Google+

- Structured (You can access to its content through formal languages)
- Share/Describe data and metadata
- A standardized way to describe and name things
 - Reusable Schema
 - Unified Schema (vs multi-schema RDB, spreadsheets)
- Describe logic relations
 - $\text{father}(r) \leftrightarrow \text{hasChild}(r) \wedge \text{isMale}(r)$
- Technology and implementation independent
- **Interoperability**

How it works?

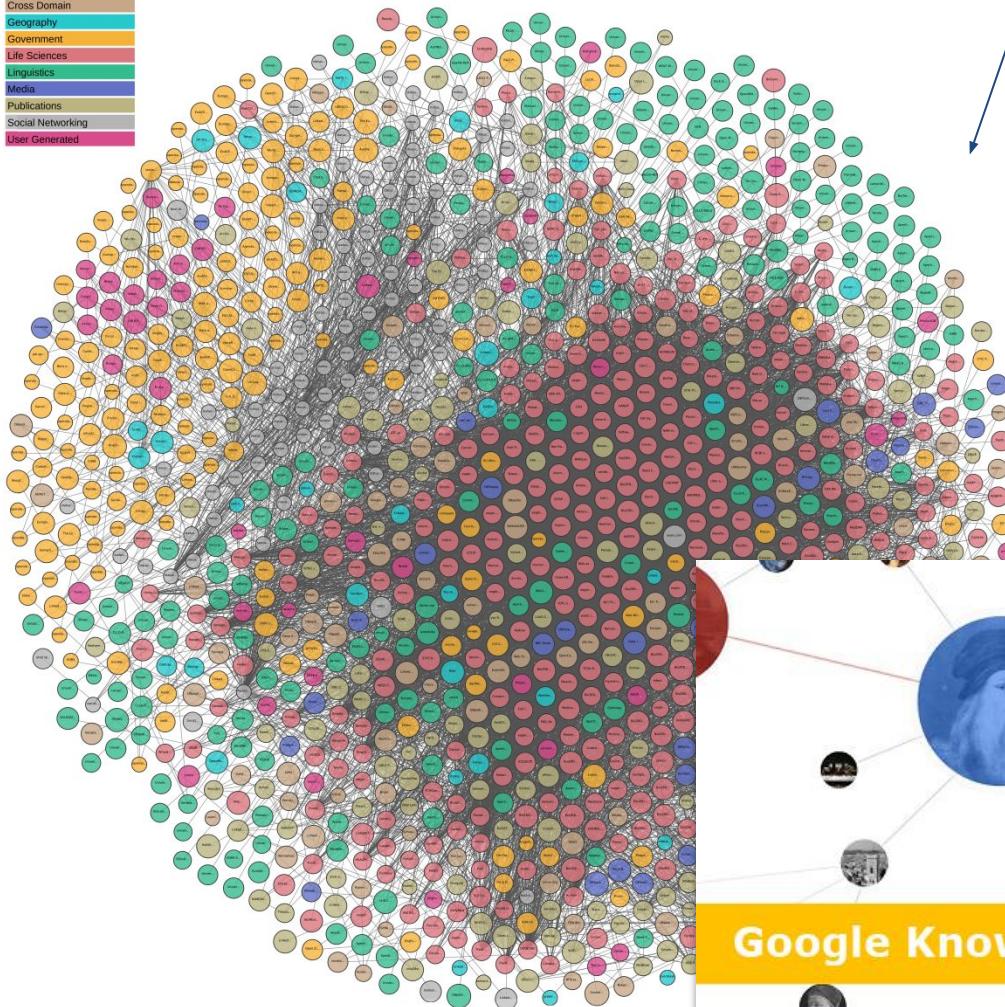


Semantic Web relations RDF -triple



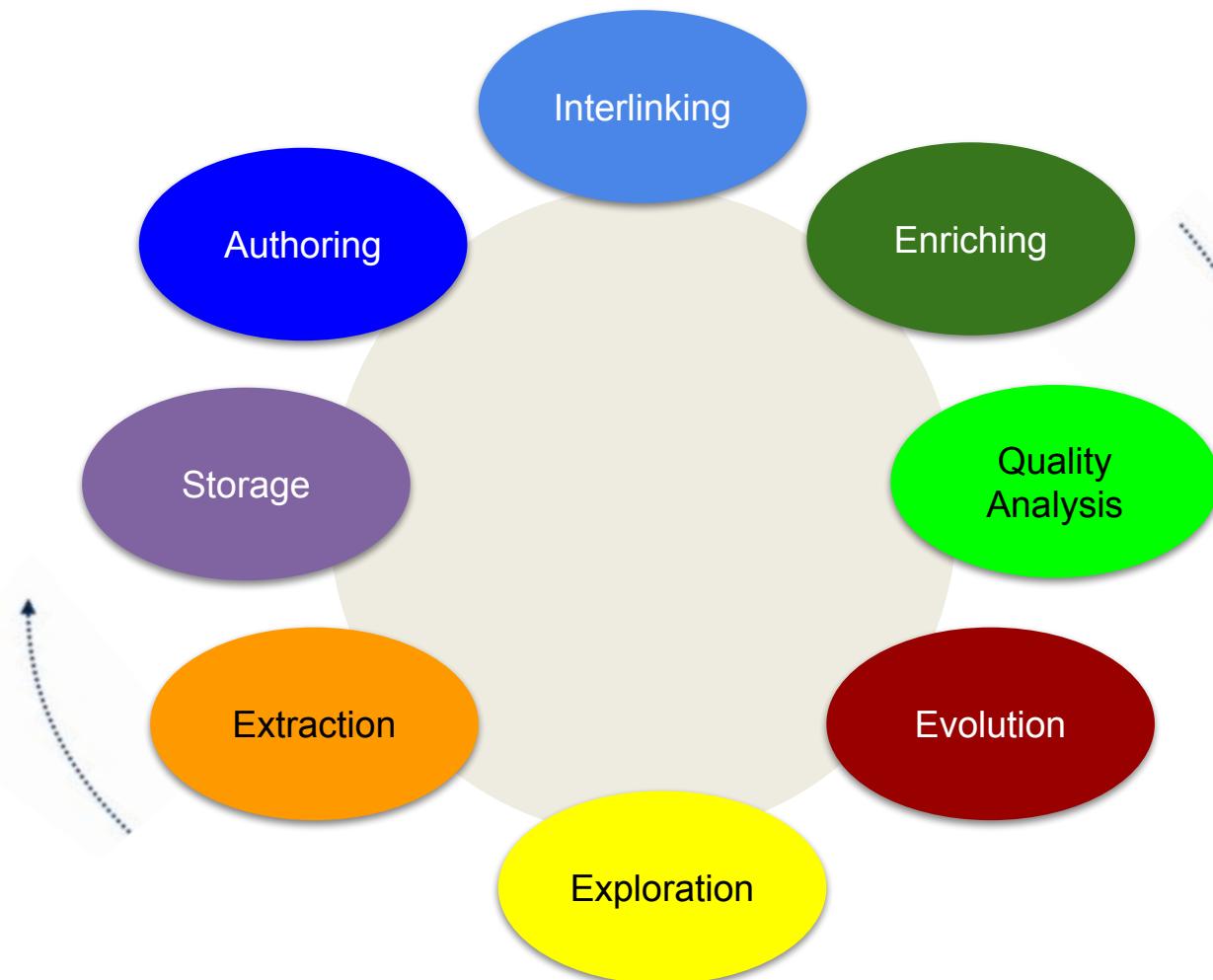
The Linked Open Data Cloud

Legend
Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated



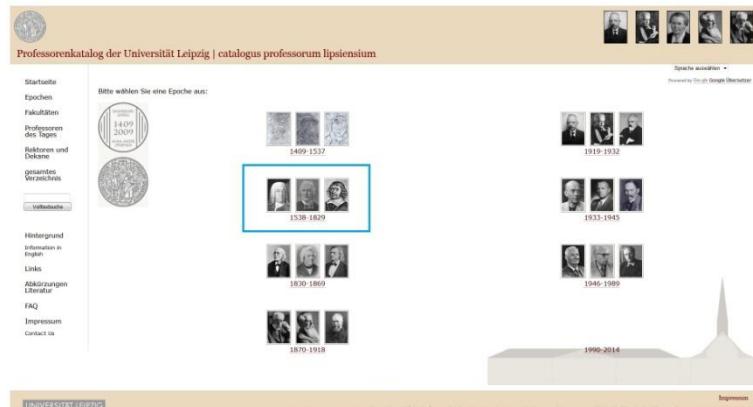
The Linked Open Data Cloud from lod-cloud.net

Linked Data (Research) Life Cycle



Professors' Catalogues - prosopographic datasets

Extraction



The screenshot shows a timeline of professor eras from 1409 to 1990-2014. The eras are represented by small portraits of professors and their names:

- 1409-1537
- 1538-1629 (highlighted with a blue box)
- 1630-1669
- 1670-1718
- 1719-1759
- 1760-1809
- 1810-1852
- 1853-1895
- 1896-1932
- 1933-1945
- 1946-1989
- 1990-2014

Below the timeline is a section titled "Impressum" which includes the university's name and a note about the page being part of the Professorenkatalog.

Epochen > 1538-1629
Fakultäten > Theologische Fakultät der Universität Leipzig
Daten: Ressourcen [RDF](#) | [PDF](#) | [Dokumente](#)

Prof. Dr. theol. Johann Benedikt Carpov I.

Lebensdaten

geb. 22.06.1607 in Rochlitz
gest. 22.10.1657 in Leipzig
Konfession: evangelisch-lutherisch
PND: 117674214

Lebenslauf

Studium

- 1618-1627 Studium in Wittenberg
- 1628-1634 Studium: Theologie in Leipzig

Qualifikation

- 1627 Promotion zum Mag. art. an der Universität Wittenberg
- 1634 Promotion zum Bacc. theol. an der Universität Leipzig
- 1651 Promotion zum Dr. theol. an der Universität Leipzig

Akademische Karriere

- 1643-1646 adProf für Theologie an der Theologischen Fakultät der Universität Leipzig
- 1646-1657 oProf (quartus) für Theologie an der Theologischen Fakultät der Universität Leipzig
- 1657 oProf (tertius) für Theologie an der Theologischen Fakultät der Universität Leipzig

Ämter

- 1650-1651 Dekan der Theologischen Fakultät der Universität Leipzig

Publikationen (Auswahl)

- Hodegeticum brevissimum aphorismis olim pro collegio concionatorio conceptum et nunc revisum, Leipzig 1636.
- Vindictio Doctrinae Orthodoxae de Sacramentis in genere a Corruptis Tridentinorum, & Suppetitis, quas illis contra Chemicum ferre voluit Bellarminus lib. I. & II. de Sacramentis, Leipzig 1651.
- Systema theologiae, Leipzig 1653.
- Disputatio Theologica De Conversione Hominum Ad Salutem, Leipzig 1655.
- Isagoge in libros ecclesiarum Lutheranum symbolicos; opus posth. a J. Oleario continuatum, Leipzig 1665.

Weitere Informationen

- 1632 Pfarrer in Meuselwitz
- 1633 Subdiakonus, 1637 Diakonus, 1643 Archidiakonus St. Thomas Leipzig
- seit 1636 verheiratet mit Elisabeth Wrieppenring, 6 Kinder
- 1657 Domherr zu Zeitz
- Konsistorialassessor

Referenzen zu anderen Projekten

- Katalog der Deutschen Nationalbibliothek
- Allgemeine Deutsche Biographie (Wikisource)
- Eintrag im Biographisch-Bibliographischen Kirchenlexikon (BBKL)
- Titel im VDAB
- Biographien in der Allgemeinen Deutschen Biographie
- Biographische Informationen des Registers auf www.deutsche-biographie.de
- Deutschsprachige Wikipedia
- Forschungsstelle für Personalschriften, Germany
- Titel im Verbundkatalog des GBV
- VIAF

Referenzen und Quellen

HeinJunghans, Die Professoren und Dozenten der Theologischen Fakultät der Universität Leipzig von 1409 bis 2009, S. 178.; ADB, Bd. 4, S. 20 f.; Bautz, Bd. 1, Sp. 937.

Links

- http://de.wikipedia.org/wiki/Johann_Benedikt_Carpov_I.

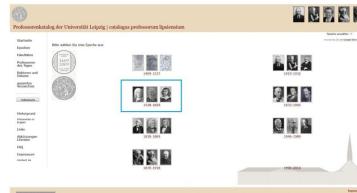
Professors' Catalogues - prosopographic datasets

Extraction

The screenshot shows a web interface for a professor catalog. At the top, there is a navigation bar with a red vertical bar and a grey bar. On the right side of the header, there are logos for the Herzog August Bibliothek Wolfenbüttel (red barcode logo) and the Niedersächsisches Ministerium für Wissenschaft und Kultur (red horse logo). Below the header, the text "Wissensproduktion an der UNIVERSITÄT HELMSTEDT" is displayed. The main content area has a dark grey header with the text "PROFESSORENKATALOG" and a portrait of two men. Below this, there is a search form with fields for "Name" (set to "Suchbegriff") and "Suchen" (Search) button. A dropdown menu shows "Name" selected. There is also a "GND" link. The search results are listed under the heading "Treffer 288, sortiert nach Nachname". The first result is for "Adamius, Theodor", with details: born *1566 in Berka bei Osterode, died †1613, Phil./Jur. Fakultät, Phil. Prof. d. Eloquenz in Helmstedt (1596-1605); Jur. Prof. d. Pandekten in Helmstedt (1605-1613), GND-Ansetzung: [104106379](#). The second result is for "Adolph, Johann Traugott", with details: born *1728 in Hirschberg/ Schlesien, died †1771, Med. Fakultät, Med. Prof. d. Anatomie und Chirurgie in Helmstedt (1760-1768); Med. Prof. der Anatomie, Chirurgie und Physiologie in Altdorf (1768-1771), GND-Ansetzung: [10016191X](#). To the right of the search results, there is a decorative illustration of a building labeled "NOVUM". At the bottom of the page, there is a footer with the text "© HAB Wolfenbüttel | Impressum".

Professors Career Patterns on the Web

Extraction



Leipzig Catalog

Extraction

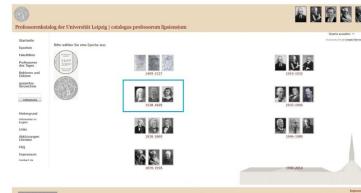


Extraction

Helmstedt Catalog

Professors Career Patterns on the Web

Interlinking



Leipzig Catalog

Extraction



Helmstedt Catalog

Extraction

PCP-On-Web
Dataset

Vocabulary Alignment

Interlinking

	Lipsius Predicates (Properties)	Helmstedt Predicates (Properties)	Merged Properties
1			<http://ns.aksw.org/spatialHierarchy/priorTo>
2			<http://ns.aksw.org/spatialHierarchy/subsequentTo>
3			
4		<http://uni-helmstedt.hab.de/cph/model/alternativeWritingOfName> <http://uni-helmstedt.hab.de/cph/model/author>	<http://purl.org/pcp-on-web#alternativeWritingOfName> <http://purl.org/pcp-on-web#author>
5			
6	<http://catalogus-professorum.org/cpd/academicalTitle>		<http://purl.org/pcp-on-web#academicalTitle>
7	<http://catalogus-professorum.org/cpd/additionToSurname>		<http://purl.org/pcp-on-web#additionToSurname>
8	<http://catalogus-professorum.org/cpd/alternativeWritingOfSurname>		<http://purl.org/pcp-on-web#alternativeWritingOfSurname>
9	<http://catalogus-professorum.org/cpd/archiveCode>		<http://purl.org/pcp-on-web#archiveCode>
10	<http://catalogus-professorum.org/cpd/bodyPlace>		<http://purl.org/pcp-on-web#bodyPlace>
11	<http://catalogus-professorum.org/cpd/cognateBody>		<http://purl.org/pcp-on-web#cognateBody>
12	<http://catalogus-professorum.org/cpd/consistsOf>	<http://uni-helmstedt.hab.de/cph/model/consistsOf>	<http://purl.org/pcp-on-web#consistsOf>
13	<http://catalogus-professorum.org/cpd/date>	<http://uni-helmstedt.hab.de/cph/model/date>	<http://purl.org/pcp-on-web#date>
14		<http://uni-helmstedt.hab.de/cph/model/dedicatee>	<http://purl.org/pcp-on-web#dedicatedTo>
15		<http://uni-helmstedt.hab.de/cph/model/digitalisat>	<http://purl.org/pcp-on-web#digitalisat>
16		<http://uni-helmstedt.hab.de/cph/model/documentBody>	<http://purl.org/pcp-on-web#faculty>
17		<http://uni-helmstedt.hab.de/cph/model/endDate>	<http://purl.org/pcp-on-web#endDate>

Task Description

<https://github.com/pcp-on-web/helmstedt-leipzig-dataset/wiki/Data-sets-fusion:-task-description>

Ontology

<http://purl.org/pcp-on-web/ontology>

Interlinking

- Task performed with Limes Link Discovery Framework using Wombat Simple algorithm with two specs
 - Highest acceptance rate (0.8)
 - <leipzig:heinrichmatthiasheinrichs> <helmstedt:13084>
0.8164965809277261
 - <helmstedt:13084> “Andreas Heinrich Matthias”
 - <leipzig:heinrichmatthiasheinrichs> “Heinrich Matthias Heinrich”
 - Lowest acceptance rate (0.5) had 1683 matching instances

Task Description

<https://github.com/pcp-on-web/helmstedt-leipzig-dataset/wiki/Instance-Matching:-Link-Discovery>

Quality
Analysis

- **Labeling problems**
 - **Properties**
 - **use of multi languages (e.g. leipzig:hasMatrikel)**
 - **use of different patterns (e.g. leipzig:surname_lat)**
 - **wrong label (e.g. leipzig:lecture -> pcp:lecturer)**
 - **Classes**
 - **wrong concepts (e.g. leipzig:course -> pcp:lecture)**

Task Description

<https://github.com/pcp-on-web/helmstedt-leipzig-dataset/wiki/Data-sets-fusion:-task-description>

Vocabulary Enhancement

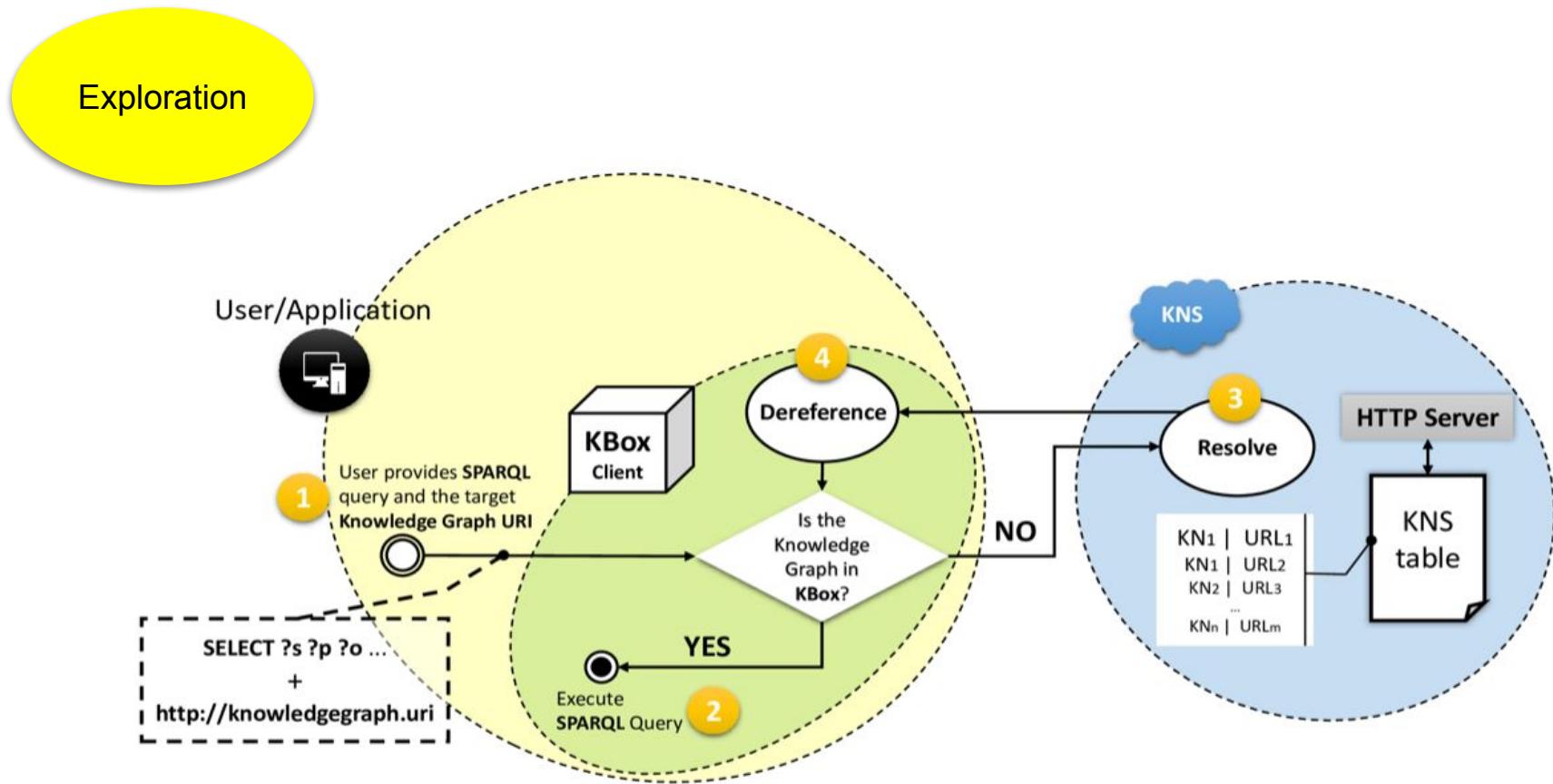
Evolution

Properties	
Original	Renamed
http://uni-helmstedt.h##ab.de/cph/model/dedicatee	http://purl.org/pcp-on-web/ontology#dedicatedTo
http://uni-helmstedt.hab.de/cph/model/documentBody	http://purl.org/pcp-on-web/ontology#faculty
http://uni-helmstedt.hab.de/cph/model/gratulant	http://purl.org/pcp-on-web/ontology#well-wisher
http://uni-helmstedt.hab.de/cph/model/lecture	http://purl.org/pcp-on-web/ontology#lecturer
http://uni-helmstedt.hab.de/cph/model/page	http://purl.org/pcp-on-web/ontology#folio
http://uni-helmstedt.hab.de/cph/model/forename_lat	http://purl.org/pcp-on-web/ontology#latinForename
http://uni-helmstedt.hab.de/cph/model/surname_lat	http://purl.org/pcp-on-web/ontology#latinSurname
http://uni-helmstedt.hab.de/cph/model/hasMatrikel	http://purl.org/pcp-on-web/ontology#hasEnrollment

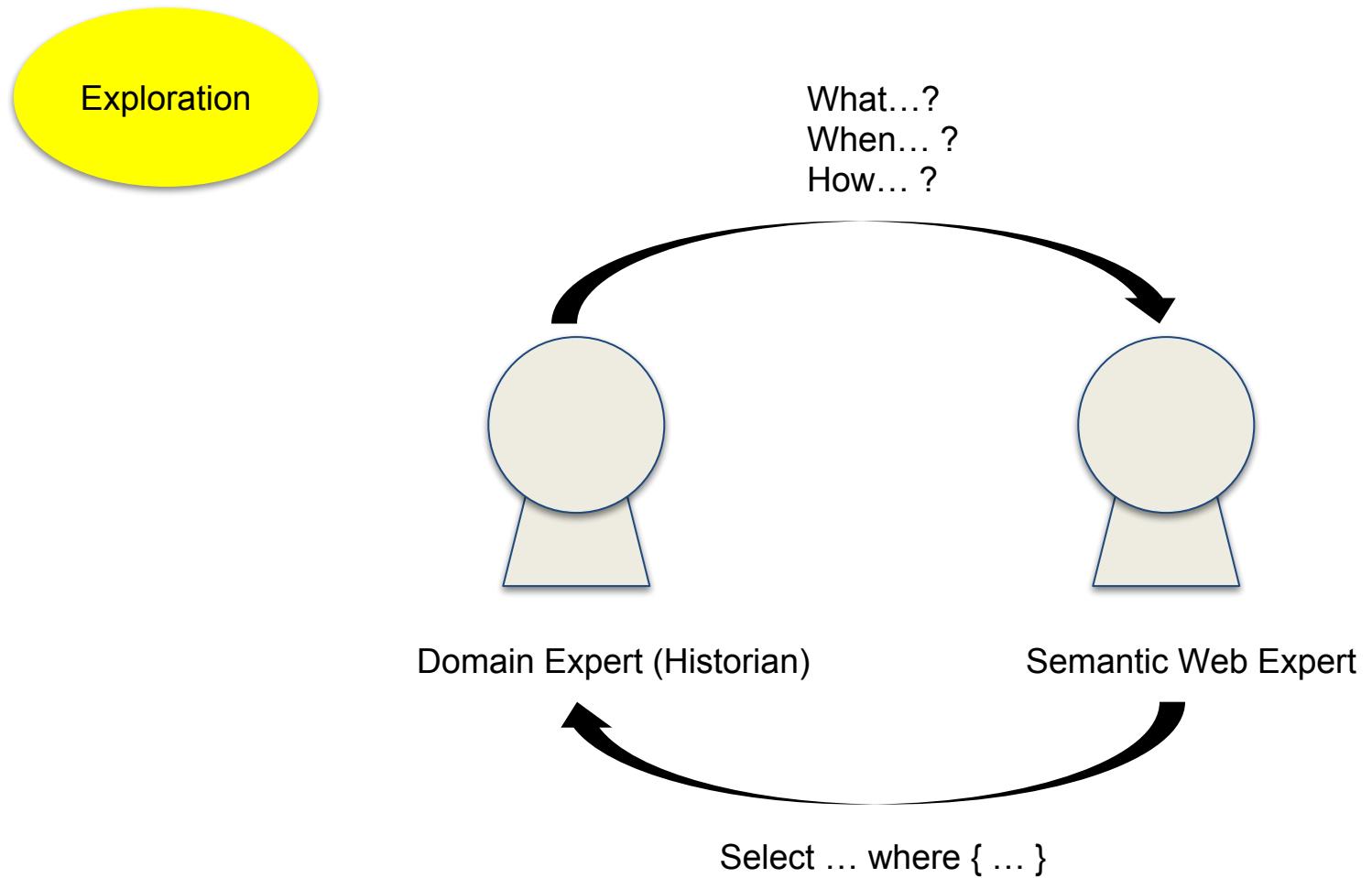
Classes	
Original	Renamed
http://uni-helmstedt.hab.de/cph/model/Course	http://purl.org/pcp-on-web/ontology#Lecture

Task Description

<https://github.com/pcp-on-web/helmstedt-leipzig-dataset/wiki/Data-sets-fusion:-task-description>



<https://github.com/AKSW/KBox>



Example

Exploration

Give me the amount of qualification documents with professors as praeses, arranged by year and faculty.

```

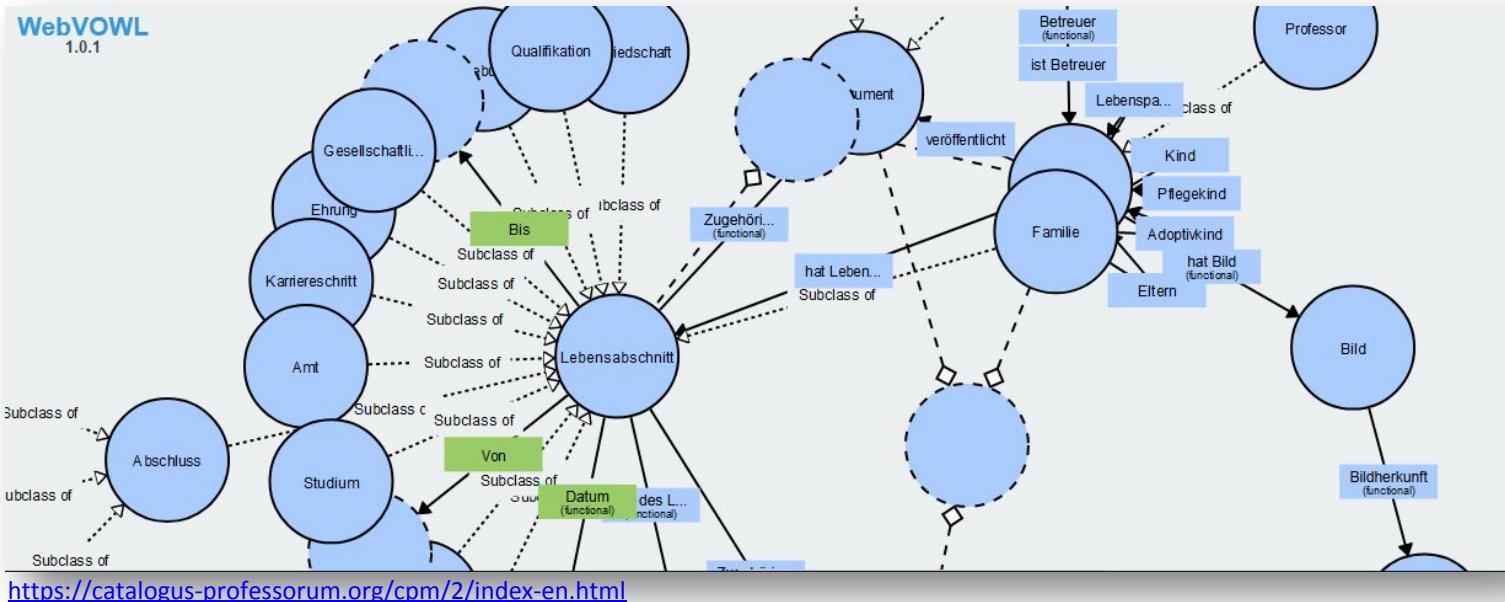
prefix pcp:<http://purl.org/pcp-on-web/ontology#>
prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix rdfs:<http://www.w3.org/2000/01/rdf-schema#>

select (count(?doc) as ?docN) ?faculty ?year
where {
    ?doc pcp:praeses ?professor.
    ?doc a pcp:QualificationDocument.
    ?professor a pcp:Professor .
    ?doc pcp:faculty ?faculty.
    ?doc pcp:date ?docDate.
    bind (year(?docDate) as ?year).
} group by ?faculty ?year order by asc(?year)
asc(?faculty)

```

Result

docN	Faculty	year
1	helmstedt:faculty/juristische-fakultaet	
1	helmstedt:faculty/theologische-fakultaet	1574
1	helmstedt:faculty/juristische-fakultaet	1574
2	helmstedt:faculty/theologische-fakultaet	1574
2	helmstedt:faculty/juristische-fakultaet	1575
2	helmstedt:faculty/theologische-fakultaet	1575
1	helmstedt:faculty/juristische-fakultaet	1576
1	helmstedt:faculty/theologische-fakultaet	1576
1	helmstedt:faculty/juristische-fakultaet	1577
1	helmstedt:faculty/theologische-fakultaet	1577
1	helmstedt:faculty/juristische-fakultaet	1578
1	helmstedt:faculty/theologische-fakultaet	1578
6	helmstedt:faculty/juristische-fakultaet	1579
6	helmstedt:faculty/theologische-fakultaet	1579
4	helmstedt:faculty/juristische-fakultaet	1580
4	helmstedt:faculty/theologische-fakultaet	1580
6	helmstedt:faculty/juristische-fakultaet	1581
1	helmstedt:faculty/philosophische-fakultaet	1581
6	helmstedt:faculty/theologische-fakultaet	1581
2	helmstedt:faculty/juristische-fakultaet	1582
2	helmstedt:faculty/theologische-fakultaet	1582
2	helmstedt:faculty/juristische-fakultaet	1583
1	helmstedt:faculty/philosophische-fakultaet	1583
5	helmstedt:faculty/theologische-fakultaet	1583
5	helmstedt:faculty/juristische-fakultaet	1585
1	helmstedt:faculty/medizinische-fakultaet	1585
1	helmstedt:faculty/philosophische-fakultaet	1585
1	helmstedt:faculty/theologische-fakultaet	1585
30	helmstedt:faculty/juristische-fakultaet	1586
30	helmstedt:faculty/theologische-fakultaet	1586
6	helmstedt:faculty/juristische-fakultaet	1587
2	helmstedt:faculty/medizinische-fakultaet	1587
1	helmstedt:faculty/philosophische-fakultaet	1587
6	helmstedt:faculty/theologische-fakultaet	1587
3	helmstedt:faculty/juristische-fakultaet	1588



- Shaped the relational database „Helmstedt Wissensportal“ by Semantic Web standards (RDF)
- Created a Mutual Ontology: <http://purl.org/pcp-on-web/ontology>
- Created the basis for a research ontology reflecting the complex research question of early modern scholarly career patterns (**in Germany**)
- Merged of the Helmstedt and Leipzig data:
<https://github.com/pcp-on-web/helmstedt-leipzig-dataset/wiki/Data-sets-fusion:-task-description>

```
java -jar kbox-v0.0.1-alpha3-RC16.jar -sparql "select ?p where { ?s ?p ?o} limit 10" -kb  
"http://purl.org/pcp-on-web/dataset,http://purl.org/pcp-on-web/ontology"
```

```
...  
| p |  
=====|  
| <http://www.w3.org/2000/01/rdf-schema#domain> |  
| <http://www.w3.org/2000/01/rdf-schema#range> |  
| <http://www.w3.org/2000/01/rdf-schema#subClassOf> |  
| <http://www.w3.org/2000/01/rdf-schema#subClassOf> |  
| <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> |  
| <http://www.w3.org/2000/01/rdf-schema#subPropertyOf> |  
| <http://www.w3.org/2000/01/rdf-schema#subClassOf> |  
| <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> |  
| <http://www.w3.org/2000/01/rdf-schema#subClassOf> |  
| <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> |
```

```
java -jar kbox-v0.0.1-alpha3-RC16.jar -list  
Knowledge base table list  
http://purl.org/pcp-on-web/ontology  
http://purl.org/pcp-on-web/dataset  
http://uni-helmstedt.hab.de/cph  
http://catalogus-professorum.org/cpd  
http://catalogus-professorum.org/lipsiensium  
http://dbpedia.org/3.9/en/full  
http://dbpedia.org/3.9/en  
...
```

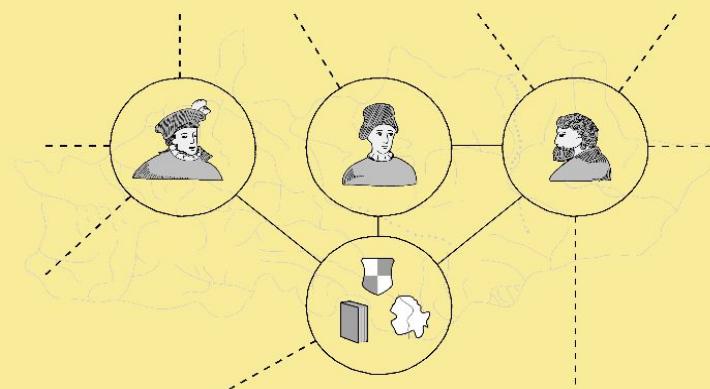
Upcoming Tasks

- **Interlink other relevant datasets**
 - German National Library
 - Gelehrte Journale (Scholarly Journals in the Age of Enlightenment)
- **Development and evolution vocabulary and dataset**
- **Research on new ways of data exploration and visualization**

PCP-on-Web - Early modern university professors as digital entities for research

FORSCHUNGSPROJEKT (PCP-ON-WEB)

PARTNER PROJEKTZIELE PORTFOLIO BLOG PUBLIKATIONEN KONTAKT DE - EN

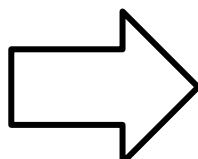


Professorale Karrieremuster der Frühen Neuzeit

Entwicklung einer wissenschaftlichen Methode zur Forschung auf online verfügbaren und verteilten Forschungsdatenbanken der Universitätsgeschichte.

DFG Projekt: 317044652

<https://pcp-on-web.htwk-leipzig.de/project/>



Thanks for your attention!

**Ontology engineering in the field of Academic History -
Research in the domain of database interoperability**

Jennifer Blanke <blanke@hab.de>

Thomas Riechert <thomas.riechert@htwk-leipzig.de>

Edgard Marx <edgard.marx@htwk-leipzig.de>

Research project: Early modern professorial career patterns
<https://pcp-on-web.htwk-leipzig.de>

Funded by Deutsche Forschungsgemeinschaft (DFG) - Projectnumber:
317044652

Save the Date!

**Doktorandensymposium (PhD Workshop)
about research with online databases on academic history –
Digital tools, methods and standards**

Date: 04-05/04/2019

Deadline: 15/02/2019

Acceptance information: 01/03/2019