From Biographies to Data Curation – the Making of www.deutsche-biographie.de

Matthias Reinert, Maximilian Schrott, Bernhard Ebneth, Team deutsche-biographie.de[1]
*Historische Kommission bei der Bayerischen Akademie der Wissenschaften, München,
†Bayerische Staatsbibliothek, München
E-mail: matthias.reinert@hk.badw.de, maximilian.schrott@hk.badw.de,
bernhard.ebneth@ndb.badw.de, info@deutsche-biographie.de

Abstract

The German Biography Portal "Deutsche Biographie" is a joint effort of the Historical Commission at the Bavarian Academy of Sciences and Humanities and the Bavarian State Library and supported by cultural heritage institutions to develop a historical and biographical information system for the German-speaking world. It includes digital full texts of more than 48.000 articles about persons and families of two biographical dictionaries and indices from associated institutions. We will describe our objectives: text-encoding, identifying individuals and places in authority files and aggregating further individual/biographical information from freely available, persistent, scientific and source-based websites and databases. The portal offers an entry point for historical biographical research. The potential of it lies in coordinated biographical data management and integration. The common database is gradually enlarged in a collaborative and modular manner together with partners in Germany and Europe. We will discuss on how the data base of scattered information can be managed.

Keywords: biographies, biographical dictionaries, digitisation, information extraction, authority files, biographical metadata aggregation, semantic web

[1] Prof. Dr. Malte Rehbein*, Klaus Kempf† as editors in chief for the website. Matthias Reinert*, Maximilian Schrott*, Sophia Stotz, Valentina Stuss as (part time) researchers employed in the project cf. http://www.deutsche-biographie.de/ueber.

1. Objectives

Biographical dictionaries constitute a unique source of historical knowledge. The articles not only inform about its subject, but they also give insight into the authors, their questions, methods, intentions, biases, and the zeitgeist. The factual knowledge expressed can be regarded as metadata. And, as we will show in this article, it can be successively derived from text, expressed using semantic web-inspired ontologies and hopefully reasoned upon. We achieved the best results by interlinking recognized entities (names, places). And we will try to pursue this approach further (works and similar subjects).

Our efforts relate to similar prosopographical projects. The biographical portal of the Netherlands successfully merged and enhanced several digitised historical biographical dictionaries as well as current online resources.[1] An API[2] allows automated request of database entries. A currently still ongoing project is described online and focuses on metadata and relations between individuals. The portal does not produce biographies on its own.

The Oxford University Press (OUP) set up an index combining individuals from the Oxford Dictionary of Biography with the National Portrait Gallery and National Archives content (images), as well as indexes of several monographs and dictionaries they published. The OUP is constantly extending their fee-based online database by about 200 biographies each year and published two supplement volumes, covering people who died 2001-08.[3] Available online is a search gateway "Oxford Index" based on identified names linking to 16 selected resources.[4] There are no known projects outside OUP which reuse these identifiers apart from Wikipedia/Wikidata.[5]

The Australian Dictionary of Biography[6] digitised their historical material as well. Their efforts concentrate on both the digital continuation of the dictionary and the enhancement of digitised articles.[7] The dictionary reconceived their publishing plan: new articles are still written, gaps are identified and subsequently filled. The articles now are provided with individuals mentioned, topics and nationality by birth using controlled vocabularies. Articles from a few other national relevant sources[8] are also integrated, obviously by manual redaction. There are only internal identifiers, no metadata is published separately. Users can aggregate several statistics on the database online.

[1] cf. http://www.biografischportaal.nl/about and the interchange format description. http://www.biografischportaal.nl/about/biodes.

[2] http://www.biografischportaal.nl/about/bioport-api-documentation.

[3] http://global.oup.com/oxforddnb/info/print/.

[4] http://oxfordindex.oup.com/. It extends the Oxford Biography Index containing about 55.000 individuals. See http://www.oxforddnb.com/index/42/101042206/ for Karl Friedrich Schinkel.

[5] https://www.wikidata.org/wiki/Property:P1415?uselang=de.

[6] http://adb.anu.edu.au.

[7] cf. http://adb.anu.edu.au/faqs/#determines and Paul Arthur in Amsterdam, 2015.

[8] F.i. Obituaries Australia, cf. http://adb.anu.edu.au/faqs/#editions.

The Historical Dictionary of Switzerland (HLS) started in 1989 and published both online and in print from the start (cf. Jorio 1990 and Jorio 2000). The articles link to the authority file Gemeinsame Normdatei (GND) and the Virtual International Authority File (VIAF) where appropriate, missing identifiers are added to the GND resource.[9]

Finally the tripartite complex Wikipedia, DBpedia and Wikidata must be mentioned. The DBpedia arose out the Wikipedia-dictionary as a means to conceptualize facts and assertions in a structured way. In addition the practical requirements to manage obvious facts scattered in different language versions gave birth to the Wikidata effort. Wikidata manages merely simple assertions on articles and struggle with the problem that facts and ontological structures are not always translatable.[10]

The digitisation of the NDB/ADB fits into the field. The latest volumes (V-Z) are not finished, all volumes in print are digitised and published online after an embargo period of 2 ½ years. The demand for data in science is met by creating and publishing structured metadata and offering link protocols.

2. The "Neue Deutsche Biographie" – in print

The precursor of the "Neue Deutsche Biographie" (NDB) originates in the 19th century. The "Allgemeine Deutsche Biographie" (ADB) was compiled and printed in 55 vol., between 1875 and 1912. This dictionary was embedded into a programme to shape the national identity on the basis of a German culture and language. The concept was to reinforce a kind of federated cultural nation on the basis of protestant liberal middle-class milieus including opponent and minority positions (Hockerts 2008, 238).

The NDB was outlined towards the end of World War II in order to renew the former dictionary ADB and in the preface dedicated to the German people as a whole ("das dem ganzen deutschen Volke gehören soll" Vorwort, NDB I, 1953, p. VI). It was conceived to cover distinguished individuals of German language and culture and was not confined by national borders. The culturally grounded concept included members of the German *Volk* (in Austria and shortly before detached territories) and foreigners with cultural influence in Germany (Vorwort, NDB I, 1953, p.VII f., in detail Hockerts 2008, 244-254).

As of 2015 25 volumes in alphabetical order up to "Tecklenborg" have been published. Three volumes („Tecklenburg – Zyrl") are planed to be released until 2020.

The NDB gives concise, thoroughly prepared biographies of deceased persons who have had a significant impact on developments in scholarship, literature, arts, politics, economics, social life, and technology (Hockerts 2008). The NDB is regarded as an authoritative biographical dictionary for all regions in which German is spoken ("deutscher Sprachraum") and German culture is prevalent (Hockerts 2008, Kraus 2010).

The NDB covers the period from the early middle ages down to the present and is arranged alphabetically. The 25 published volumes, contain about 22.000 articles, roughly 19.000 are biographies on individual persons and 3.000 cover families. The articles include detailed information on genealogy (cf. Ebneth 2012), selective lists of works as well as secondary literature, and references to portraits. In total about 8.000 different authors, often distinguished experts in their field, have contributed articles.

The NDB is edited by the Historical Commission at the Bavarian Academy of Sciences and Humanities, Munich, under the guidance of the editors in chief, who are members of this commission.[11]

The publishing house Duncker & Humblot, Berlin, printed and distributed all volumes, each containing about 800 articles on 830 pages and an index of names. The NDB is mostly used in public and scientific libraries in Europe and Overseas.

Each volume is prepared by an editorial staff of five historians each of whom is responsible for a particular subject area. The staff selects the people to be included in the NDB, appoints qualified authors, and edits the articles for printing. The preparation process builds upon an internal database which is continuously extended by systematic examination of online resources, monographs, periodicals, newspapers, obituaries, bibliographies, editions and exhibitions.

Each volume covers a selection of 8.000 names in the internal database, the editor responsible for a specific field (humanities, sciences, literature, arts, politics, business, medicine) creates a preselection which is discussed with selected relevant experts. The author is chosen by the editor and the article undergoes an informal peer review process in the editorial office and a revision in order to fit the criteria of style, content and completeness.

A strict structure is imposed on all articles (cf. Redaktion der NDB 2009):

1. Full name, occupation, date and place of birth, date and place of death, tomb, religious denomination
2. Family (genealogy)
3. Career, achievements, critical evaluation
4. List of selected works
5. List of sources and secondary literature
6. References to portraits (but no illustrations)
7. Name of the author.

3. Principles of Digitisation

The "making of" www.deutsche-biographie.de rests on three pillars: (1) the digitisation of ADB and NDB with

[9] A GND-Beacon is provided: http://www.hls-dhs-dss.ch/pnd/BEACON-PND-HLS.txt.

[10] See f.i. Mark Grahams early remarks in http://www.theatlantic.com/technology/archive/2012/04/the-problem-with-wikidata/255564/.

[11] http://www.historischekommission-muenchen.de; The Historical Commission, elected as editors in chief during the digitisation campaign Hans Günter Hockerts (1998-2012), Maximilian Lanzinner (2013) and Hans-Christoph Kraus (since 2014).

the ongoing edition of biographies, (2) a joint index of ADB and NDB and (3) the interlinking and integration of distributed sources.

3.1. Ease access to the dictionary

A dictionary of notable individuals and families is generally understood as a principal source of biographical information and should provide a low threshold of accessibility. At first the pages of the older series ADB and subsequently the NDB had been scanned and put online. The index of names was already compiled for the ADB in an additional volume in print. It was and is enlarged by all the names from the indices of each new volume of NDB. The index of names, professions, dates and pages of occurrence accompanied the digital images. A relational databases back-end served as a comfortable means to find the specific pages.

In a second step the full text was digitised, encoded in XML (partially TEI[12]) and the structure reintroduced. In parallel each name was aligned to the bibliographical authority file Personennamendatei (PND, in 2012 extended to GND, cf. Pfeifer 2015). By exploiting structural tagging and indexed information the cross linking of articles and index entries could be realized.

The third stage introduced faceted filters and map based search abilities.

3.2. Collecting biographical information on outstanding personalities

A constant interest of the editorial board is to extend the knowledge base, to search for reliable biographical information on historical personalities and families. In 2009 a decade after initial thoughts and plans an international cooperation started (Menges/Ebneth 1998, Ebneth 2009, Ebneth 2010). The Biographical Portal brought together the index entries of several national biographical dictionaries and made them searchable in a unified way (cf. Gruber/Feigl 2011).

In addition the NDB gained substantial support through cooperation. Relevant cultural heritage institutions delivered identifiers and the meta data could be drawn from the authority file GND.

A basic form of cooperation with over 100 websites is realised by cross-linking on the basis of authority file identifiers and a suitable interface (BEACON).[13] By publishing a list of GND identifiers and a concordance to subpages on a website they become interlinkable through aggregators.[14] These concordances could as well be created by a third party researcher.

3.3. Adopting authority files

The first database online already offered selections along groups of profession. An early classification with a two level hierarchy of occupations distinguishes between humanities, sciences, arts, administration & church, and business & technology. It is still in use and could in the future support grouping queries and information extraction alike.

Instead of creating an own authority file the NDB sought the cooperation with the Munich Digitization Center (Münchener Digitalisierungszentrum MDZ) at the Bavarian State Library in order to supply each individual in the NDB/ADB with an identifier in the bibliographical authority file GND (Busch/Jordan 2011, Ebneth/Busch 2012).[15]

By using GND the Deutsche Digitale Bibliothek (DDB[16]) and the Consortium of European Research Libraries (CERL[17]) Thesaurus are frequently linking to Deutsche Biographie.

The latest efforts in this direction had been the locating of birth and death places of the NDB in geographical databases.

3.4. Link to subsidiary resources on the web

To ease access to all relevant scientific sources it is required to improve and systematize linking. The NDB besides others adopted and promoted the use of authority files and identifiers in historiographical research projects (cf. Akademienunion 2009).

Following prominent proponents like BSB and Wikipedia the NDB promoted simple protocols (BEACON) to offer and share lists of identifiers and concordances.[18] With BEACON everyone can aggregate lists of links on an individual basis automatically.[19]

3.5. Linked Open Data

In late 2011 the Historical Commission applied succesful for consulting in an EU-funded project Linked Open Data (LOD2, Riechert 2011).[20] Together with experts from the Leipzig AKSW[21] the metadata were translated into a semantic web ontology expressed in RDF-XML (Brümmer 2011). This prototype[22] was based upon the ontology-schemata developed for the authority file GND and comprised roughly 2.7 million statements.

---

12 http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html.

13 Beacon https://de.wikipedia.org/wiki/Wikipedia:BEACON; https://meta.wikimedia.org/wiki/Dynamic_links_to_external_resources.

14 http://beacon.findbuch.de/seealso/pnd-aks maintained by Thomas Berger, Bonn.

15 The GND is curated by the German National Library and widely used by german, austrian, and swiss libraries to collect and identify personal and organisational names, geographic entities and subject headings. http://www.dnb.de/gnd.

16 https://www.deutsche-digitale-bibliothek.de.

17 http://thesaurus.cerl.org.

18 http://www.historische-kommission-muenchen-editionen.de/pnd.html.

19 http://www.deutsche-biographie.de/vernetzte_angebote, see also http://beacon.findbuch.de/seealso/pnd-aks, equally possible for institutions http://beacon.findbuch.de/seemore/gnd-aks.

20 http://lod2.eu; http://blog.aksw.org/german-biographies-as-part-of-the-linked-open-data-cloud.

21 http://aksw.org.

22 http://data.deutsche-biographie.de:8888/bigdata/.

The concept of linked open data is nourished by prospects of machine reasoning and automated knowledge management. Interlinking of inhomogeneous web content by use of identifiers, decidable ontology-schemata and expression of facts in simple structured phrases (RDF) seems to allow the cross-checking of information, the recomposition of knowledge and distillation of new knowledge. Digital key player adopted it (BBC, Google). We estimated that it is to laborious to check each difference in factual assertions, especially as often a severe scientific controversy lies behind a seemingly simple difference.

In order to interoperate on linked open data and infer statements a significant amount of ontology-mapping and value-combing is necessary. The GND contained a small number of error, e.g. impossible dates. Our own linked data set contains normalised values and in some cases information is lost. For example, the year of birth "um 1225"[23] or place of birth "wohl in Aquitanien"[24] would require more specific vocabulary which would make inference more complicated.

4. The Deutsche Biographie online – Ongoing Project

The recent project, that aims to establish a historical and biographical information system, runs from 2012 to 2015 (Hockerts 2012; Jordan 2012; Hagn/Schrott 2015; Ebneth 2015). It deepens and extends the objectives mentioned in ch. 3.

The Deutsche Biographie aggregates data from 15 selected renowned partners in Germany who provide content related to individuals of importance. In a first stage (2012-14) we worked with Deutsches Literaturarchiv Marbach, Bundesarchiv Koblenz/Berlin, Germanisches Nationalmuseum Nürnberg, Foto Marburg, Deutsches Rundfunkarchiv Frankfurt/Main and Deutsches Museum, München to get their individual records in selected databases linkable.[25] The second stage in 2015 includes Staatsbibliothek zu Berlin as coordinating body of "Kalliope", Deutsche Fotothek Dresden, Deutsches Filminstitut Frankfurt/Main, Landesarchiv Baden-Württemberg (Landeskundliches Informationssystem Baden-Württemberg LEO-BW[26]), as well as selected projects of four academies of sciences and humanities (Berlin-Brandenburg, North Rhine-Westphalia, Mayence and Bavaria). Partners in the second stage usually offer to align their biographical information with an authority file (GND) on their own and agree that their names should be integrated into the Deutsche Biographie.

4.1. Extending the Knowledge Base

Biographical information from our cultural heritage partners increased the Deutsche Biographie by 134.000 additional records.

As of late 2014 the Deutsche Biographie offers more than 260.000 records on personalities with further biographical information and individual linking to up to 150 web resources. By the end of 2015 the number of persons in the Deutsche Biographie will rise above 500.000.

4.2. Information extraction on biographical articles

Cross linking articles by recognizing and identifying persons mentioned in the index files was a first approach to a more general extraction of information. It started with regular expressions and outstanding features like names and dates in headlines as well as page numbers. With the 152.000 references in the printed index as a basis, nearly all article headings (48.000) were identified automatically and about 55.000 occurrences of names could be automatically located in the corpus (cf. Reinert 2010).

The headlines of articles regularly state the place of birth and of death. These places and the places of burial (if documented) were extracted and linked with geographic data bases. The API provided by Nominatim, a joint service of Mapquest and OpenStreetMap was used.[27] One third of the 12.000 different place names could be automatically matched and provided with coordinates. Unfortunately Nominatim and OpenStreetMap (OSM) did not provide persistent identifiers. To some extent our model and the OSM approach mismatched, because we do not differentiate yet between geographical entities like OSM and treated everything as a timeless point. But at least the coordinates persist and we use them to offer search result on a map.

After that our efforts went to extracting information in interpersonal relationships as found in the articles of the NDB. The prospect was content enhancement and better search functions by adding context to search results.

Although the research started with very limited resources we could rely on advice and help from experts at the Centrum für Informations- und Sprachverarbeitung (CIS, LMU München), namely Franz Guenthner and Michaela Geierhos (now University of Paderborn, cf. Geierhos 2010). The concept of Local Grammar (Maurice Gross) was adopted and grammars drafted to be applied with Unitex[28], a versatile open source corpus processor. Through cooperation we were allowed to use the CISLEX dictionaries for German (Guenthner & Maier 1994; Langer et al. 1996).

Due to limited resources we started with scientific teachers and students, and slightly extended linguistic work to cover friends and circles. The focus lay on the most frequently recurring phenomena (Stotz/Reinert, 2013; Stotz/Stuß 2014).

4.3. Enhanced search and Information visualisation

In addition to our newly developed faceted search we

23 http://data.deutsche-biographie.de/rest/sfz17523.rdf.

24 http://data.deutsche-biographie.de/rest/sfz70566.rdf.

25 List of Partners cf. http://www.deutsche-biographie.de/partner.

26 http://www.leo-bw.de/web/guest/about.

27 http://nominatim.openstreetmap.org.

28 http://www-igm.univ-mlv.fr/~unitex.

were able to offer maps with identified places of birth, death and burial as starting point. Relations between identified individuals are now additionally displayed as an ego-centered graph, utilizing the JavaScript library D3.js. The RDF-store prototype worked with a generic tool to visualise relationships (Relfinder, cf. Heim et al. 2009), but it did not offer simple graph based queries.

4.4. Metadata to RDF

The accompanying RDF-store will be set up anew in OpenLink Virtuoso because of problems of speed, and scalability. The prototype developed with AKSW Leipzig (Brümmer 2011) exhibited metadata and first genealogical relations.

Current work consists of extending the ontology schema to express interpersonal relations in RDF/OWL and proceeding to express these relations: In addition to the schema of GND[29] we look at Agrelon[30] as a reference, but more refinement on predicates would still need to be done. Another influence was the Catalogus Professorum Model (CPM) which is expressive for statements and constraining on argument values, but it is specified on careers of university teachers and seems to be discontinued (Riechert et al. 2010a, 2010b). Further we considered the Conceptual Reference Model (CIDOC-CRM) expressed in "Erlangen OWL". It follows an "agents and events" based meta model which requires a certain overhead in order to express an interpersonal relation.

5. Results and discussion

The digitisation of NDB and ADB formed a first step into "world of data". Normalisation, homogenisation, and categorisation was forced upon historical biographical accounts. Authority files offered of big help in data management and interlinking opportunities.

For persons even our internal identifiers subsist because GND identifiers were not always initially available. The ontology schemata did not fit in every situation, but it can and will be technically specialised.

But there are even more types of identifiers provided by authority files: works of art, not only literature, professions, organisations, fields of study and subject headings (cf. Reinert 2011b).

Future work should also clarify the uncertainty of individual statements which have been normalized and respect the historical complexity of denominations (places, organisations, individuals) in an appropriate way.

6. Acknowledgements

The NDB received funding from the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) in 2001-03, 2008-09, 2010/11 and 2012-15.[31]

[29] http://d-nb.info/standards/elementset/gnd.

[30] http://www.contentus-projekt.de.

[31] http://gepris.dfg.de/gepris/projekt/53346764; http://gepris.dfg.de/gepris/projekt/165972532; http://gepris.dfg.de/gepris/projekt/213818920.

7. References

Akademienunion (ed.) (2009) Personendateien – *Workshop der Arbeitsgruppe Elektronisches Publizieren der Union der deutschen Akademien der Wissenschaften in Zusammenarbeit mit der Sächsischen Akademie der Wissenschaften und der Deutschen Nationalbibliothek*, 21.-23.9.2009, Leipzig, http://www.akademienunion.de/fileadmin/redaktion/user_upload/-Personendateien.pdf.

Brümmer, Martin (2011). Realisierung eines RDF-Interfaces für die Neue Deutsche Biographie SKIL 2011. In *Studentenkonferenz Informatik Leipzig 2011, Leipziger Beiträge zur Informatik*, vol. 27, Leipziger Informatik-Verbund (LIV), http://skil.informatik.uni-leipzig.de/blog/wp-content/uploads/proceedings/2011/Brummer2011.31.pdf.

Busch, Thomas, Jordan, Stefan (2011). Vernetzte Lebensläufe. Der Einsatz von Normdatenbanken zur Verlinkung biographischer und bibliographischer Angebote im Internet. *Geschichte in Wissenschaft und Unterricht*, 11/12, pp. 684-691.

Ebneth, Bernhard (2009). Vom digitalen Namenregister zum europäischen Biographie-Portal im Internet. In Martina Schattkowsky and Frank Metasch (eds.), *Biografische Lexika im Internet, Bausteine aus dem Institut für Sächsische Geschichte u. Volkskund*e 15, pp. 13-44.

Ebneth, Bernhard (2010). Das europäische Biographie-Portal mit Allgemeiner Deutscher Biographie und Neuer Deutscher Biographie Online. In Ulf Morgenstern and Thomas Riechert (eds.), *Catalogus Professorum Lipsiensis. Konzeption, technische Umsetzung und Anwendungen für Professorenkataloge im Semantic Web*, Leipzig, pp. 159-168.

Ebneth, Bernhard (2012). Aktueller Stand der Genealogien in der Neuen Deutschen Biographie – Arbeit mit der Online-Version, http://www.ndb.badw.de/Genealogentag-NDB-2012.pdf.

Ebneth, Bernhard, Busch, Thomas (2013). Die Deutsche Biographie – Vernetzung mittels der Gemeinsamen Normdatei (GND), *Workshop: Mehr Personen – Mehr Daten – Mehr Repositorien, Berlin-Brandenburgische Akademie der Wissenschaften*, 4.-6.3.2013, http://pdr.bbaw.de/veranstaltungen/pdr-workshop-2013/die-deutsche-biographie-vernetzung-mittels-der-gemeinsamen-normdatei-gnd, http://www.ndb.badw.de/Deutsche_Biographie_Vernetzung_mit_GND.pdf.

Ebneth, Bernhard (2015). Auf dem Weg zu einem Historisch-biographischen Informationssystem. Daten integration und Einsatz von Normdaten am Beispiel der Deutschen Biographie und des Biographie-Portals. *Jahrbuch für Universitätsgeschichte* 16, pp. 261-290.

Geierhos, Michaela (2010). *BiographIE - Klassifikation und Extraktion karrierespezifischer Informationen. Linguistic Resources for Natural Language Processing 05*. München: Lincom.

Gruber, Christine, Feigl, Roland (2011). Das ‹Biographie-

Portal› – work in progress. *Germanistik in der Schweiz (GiS),* 8/2011, pp. 211-220, http://www.germanistik.ch/scripts/download.php?id=Das_Biographie-Portal.

Guenthner, Franz, Maier, Petra (eds.) (1994). *Das CISLEX Wörterbuchsystem*. München, http://www.cis.uni-muenchen.de/download/cis-berichte/94-076.pdf.

Hagn, Roxane, Schrott, Maximilian (2015). Workshop Historisch-biographisches Informationssystem, Tagungsbericht, http://www.hsozkult.de/conferencereport/id/tagungsberichte-5853.

Heim, Philipp, Hellmann, Sebastian, Lehmann, Jens, Lohmann, Steffen and Stegemann, Timo (2009). RelFinder - Revealing Relationships in RDF Knowledge Bases. In *Proceedings of the 4th International Conference on Semantic and Digital Media Technologies* (SAMT 2009), Berlin/Heidelberg: Springer, pp. 182-187.

Historische Kommission bei der Bayerischen Akademie der Wissenschaften (2010). Jahresbericht 2009, pp. 37-56, http://www.historischekommission-muenchen.de/fileadmin/user_upload/pdf/jahresberichte/jahresbericht 2009.pdf.

Hockerts, Hans Günter (2008). Vom nationalen Denkmal zum biographischen Portal. Die Geschichte von ADB und NDB 1858-2008. In Lothar Gall (ed.) *"... für deutsche Geschichts- und Quellenforschung". 150 Jahre Historische Kommission bei der Bayerischen Akademie der Wissenschaften*, München: Oldenbourg, pp. 229-269.

Hockerts, Hans Günter (2012). Zertifiziertes biographisches Wissen im Netz. Forschungsnahe Informationsinfrastruktur. Die „Deutsche Biographie“ auf dem Weg zum zentralen historisch-biographischen Informationssystem für den deutschsprachigen Raum. *Akademie Aktuell,* 4, pp. 34 sq., http://www.badw.de/de/publikationen/akademieAktuell/2012/43/0412_12_hockerts.pdf.

Jordan, Stefan (2012). Entwicklung eines zentralen Historisch-biographischen Informationssystems für den deutschsprachigen Raum. Tagungsbericht, http://www.hsozkult.de/conferencereport/id/tagungsberichte-4372.

Jorio, Marco (1990). Historisches Lexikon der Schweiz. *Geschichte und Informatik* 1(1990), pp. 24-27, http://dx.doi.org/10.5169/seals-1255.

Jorio, Marco (2000). Das Historische Lexikon der Schweiz im Jahre 2000. *Schweizerische Zeitschrift für Geschichte* vol. 50, no. 5, pp. 198-203, http://dx.doi.org/10.5169/seals-81277.

Kraus, Hans-Christof (2010). Neue Deutsche Biographie, 24. Bd.: Schwarz - Stader (Rezension). *Jahrbuch für die Geschichte Mittel- und Ostdeutschlands* 58, 2012, pp. 180-182, http://www.degruyter.com/view/supplement/s21919909_Inhaltsverzeichnis.pdf.

Kraus, Hans-Christof, Jorio, Marco, Schattkowsky, Martina, Ebneth, Bernhard, Reinert, Matthias, Declerck, Thierry, Gruber, Christine, Wandl-Vogt, Eva (2014). Sektion: Vernetzung von historisch-biographischen Lexika und Fachportalen im Linked (Open) Data Framework. *1. Jahrestagung der Digital Humanities im deutschsprachigen Raum (DHd 2014), Universität Passau, 25.-28.3.2014*. http://www.ndb.badw.de#DHd_2014.

Langer, Stefan, Maier, Petra, Oesterle, Jürgen (eds.) (1996). CISLEX – An Electronic Dictionary for German: Its Structure and a Lexicographic Application. In *CIS-Bericht-96-97, Proceedings of COMPLEX '96*, pp. 155-164.

Menges, Franz, Ebneth, Bernhard (1998). Die Neue Deutsche Biographie als Projekt und Aufgabe der Historischen Kommission bei der Bayerischen Akademie der Wissenschaften. In Peter Csendes et al (eds.) *Traditionelle und zukunftsorientierte Ansätze biographischer Forschung und Lexikographie, Symposium des Instituts Österreichisches Biographisches Lexikon und biographische Dokumentation*, Österreichisches Biographisches Lexikon, Schriften 4, pp. 9-15.

Neue Deutsche Biographie (1953 sqq.). Historische Kommission bei der Bayerischen Akademie der Wissenschaften (ed.), 25 vols., Berlin: Duncker & Humblot, reprint 1971, since 2008 also http://www.deutsche-biographie.de.

Pfeifer, Barbara (2015). Über Zweck und Nutzen der Gemeinsamen Normdatei (GND). *Jahrbuch für Universitätsgeschichte* 16, pp. 251-259.

Redaktion der NDB (Ed.) (2009). Richtlinien für die Erstellung von Beiträgen für die NDB (notes of guidance for authors), last revision May, http://www.ndb.badw.de/ndb_richtlinien.htm.

Reinert, Matthias (2010). Biographisches Wissen auf einen Klick. Die "Deutsche Biographie" führt das Angebot von Neuer Deutscher Biographie (NDB) und Allgemeiner Deutscher Biographie (ADB) im Internet zusammen (…) *Akademie Aktuell*, 4, pp. 44-46. http://www.badw.de/de/publikationen/akademie Aktuell/2010/35/16_reinert.pdf.

Reinert, Matthias (2011a). Die „Neue Deutsche Biographie“ auf dem Weg zum biographischen Informationssystem, *Vernetzungstage, Osnabrück, 4.3.2011*, http://www.dini.de/fileadmin/workshops/ver netzungstage_2011/reinert_neuedeutschebiografie.pdf.

Reinert, Matthias (2011b). Normdatenbasierte Vernetzung (in) der Neuen Deutschen Biographie, *.hist2011, Berlin 14.-15.9.2011*, http://www.hsozkult.de/event/id/termine-16859.

Riechert, Thomas, Morgenstern, Ulf, Auer, Sören, Tramp, Sebastian and Martin, Michael (2010). The Catalogus Professorum Lipsiensis – Semantics-based Collaboration and Exploration for Historians. In *Proceedings of the 9th International Semantic Web Conference (ISWC2010)*, ceur-ws.org/Vol-658/paper532.pdf.

Riechert, Thomas (2011). Deutsche Biographie becomes part of the LOD cloud. Workshop on June 27th at the Historical College in Munich. The AKSW group and the group of the New German Biography at the Bavarian Academy of Sciences and Humanities presented their results of the LOD2-supported PUBLINK project, http://blog.aksw.org/german-biographies-as-part-of-the-

linked-open-data-cloud.
Stotz, Sophia, Reinert, Matthias (2013). Detecting and Encoding Interpersonal Relations with Unitex/Local Grammars, *2nd Unitex/GramLab Workshop*, Université Paris Est-Marne-la-Vallée.
Stotz, Sophia, Stuß, Valentina (2014). Relationsextraktion mit lokalen Grammatiken am Beispiel der Relation „emigrieren“, *1. Jahrestagung der Digital Humanities im deutschsprachigen Raum (DHd 2014), Universität Passau, 25.-28. März 2014,* https://wiwi.uni-paderborn.de/uploads/media/Poster2.pdf.

8. Appendix: Chronology of Digitization

since 1996	Cooperation between NDB and Leibniz Supercomputing Centre (Leibniz Rechenzentrum, LRZ[32]), MS-Access-Databases
1997, nov.	International conference at Vienna (ÖAW): Traditionelle und zukunfts-orientierte Ansätze biographischer Forschung und Lexikographie[33]: Biographische Lexika im digitalen Zeitalter
since 1999	Cooperation between NDB and Munich Digitization Center (Münchener Digitalisierungszentrum MDZ[34]) at Bavarian State Library (Bayerische Staatsbibliothek[35])
2001-2003	Cumulated Index to ADB & NDB and Image files of ADB online, funded by Deutsche Forschungsgemeinschaft (DFG)
2003	CD-ROM with Cumulated Index to ADB & NDB[36], published by Duncker & Humblot, Berlin[37]
2008:	Image files of NDB online[38]
since 2008	New Database NDBIO (FAUST) in cooperation with LRZ for biographical documentation and management of articles, correspondance with authors and publisher, bibliography, and index of NDB
since 2009	Biographical Portal[39] (Biographie-Portal), cooperation between NDB and Austrian Academy of Sciences (Österreichisches Biographisches Lexikon, ÖBL), Foundation Historical Dictionary of Switzerland (Historisches Lexikon der Schweiz/Dictionnaire historiqe de la Suisse/Dizionario storico della Svizzera, HLS/DHS/DSS), and Bavarian State Library
2008-2009:	Full-Text Digitization of ADB & NDB and application of persistent identifiers of Gemeinsame Normdatei (GND), funded by DFG[40]
2010	Expansion of Index of Deutsche Biographie by data from NDBIO with usage of GND[41]
2012	Biographical Database of Rhineland-Palatinate (Rheinland-Pfälzische Personendatenbank, RPPD) and Saxon Biography (Sächsische Biografie, SäBi) take part in Biographical Portal[42]
since 2012	Deutsche Biographie will be enlarged by development of a central historical and biographical information system by networking with other institutions of cultural heritage in Germany and in close cooperation with GND and semantic analysis[43]
2014, dec.	Relaunch www.deutsche-biographie.de, funded by DFG, and workshop Historisch-biographisches Informationssystem at Munich[44]
2015, march	Slovenska biografija takes part in Biographical Portal[45]

32 http://www.lrz.de.
33 http://verlag.oeaw.ac.at/index.phtml?act=ps&aref=1819.
34 http://www.digitale-sammlungen.de/index.html?c=digitalisierung&l=en.
35 https://www.bsb-muenchen.de/en.
36 http://rzblx10.uni-regensburg.de/dbinfo/detail.php?titel_id=3697.
37 http://www.duncker-humblot.de/index.php/neue-deutsche-biographie-58.html?___store=gb.
38 https://idw-online.de/en/news226995.
39 http://www.biographie-portal.eu/en; http://www.germanistik.ch/scripts/download.php?id=Das_Biographie-Portal.
40 http://gepris.dfg.de/gepris/projekt/53346764.
41 http://gepris.dfg.de/gepris/projekt/165972532.
42 http://web.isgv.de/index.php?page=743.
43 http://gepris.dfg.de/gepris/projekt/213818920.
44 http://www.hsozkult.de/conferencereport/id/tagungsberich-te-5853; http://www.ndb.badw.de/index_e.htm#German_Biography_Portal; http://www.bonnerblogs.de/tag/deutsche-biographie.
45 http://www.slovenska-biografija.si.