

Connecting people digitally - a semantic web based approach to linking heterogeneous data sets

Katalin Lejtovicz, Amelie Dorn

Austrian Centre for Digital Humanities, Vienna, Austria

{katalin.lejtovicz, amelie.dorn}@oeaw.ac.at

Abstract

In this paper we present a semantic enrichment approach for linking two distinct data sets: the ÖBL (Austrian Biographical Dictionary) and the dbo@ema (Database of Bavarian Dialects in Austria electronically mapped). Although the data sets are different in their content and in the structuring of data, they contain similar common “entities” such as names of persons. Here we describe the semantic enrichment process of how these data sets can be inter-linked through URIs (Uniform Resource Identifiers) taking person names as a concrete example. Moreover, we also point to societal benefits of applying such semantic enrichment methods in order to open and connect our resources to various services.

1 Introduction

In the Digital Humanities discourse, the establishment of data networks and creation of links between different resources has been a key aspect. The linking of resources not only aims at enrichment, but more importantly also at providing wider access to data resources in local but also global digital infrastructures. As a consequence data use and re-use is enabled.

One widely practised way of enabling semantic enrichment and linking is by means of open-source tools relying on semantic web technologies. For example DBpedia Spotlight (Mendes et al., 2011) provides the possibility to automatically annotate documents with mentions of DBpedia resources. The tool uses as resource types the classes of the DBpedia Ontology, thus enabling the user to annotate

documents with 272 different entity types. Furthermore, the user can choose the annotation domain by selecting the classes of the Ontology or by defining them via a SPARQL query. Although DBpedia Spotlight is a powerful tool, it limits entity linking to only one resource, and was developed for the English language. To apply it on documents written in other languages, the models used by Spotlight have to be adapted. Babelify (Moro et al., 2014) uses a graph-based approach to perform entity linking and word sense disambiguation, relying on BabelNet 1.1.1 - a semantic network of Wikipedia and WordNet¹ - in order to provide LOD² links to identified text fragments. Babelify's main asset is the use of a multilingual resource that incorporates encyclopedic knowledge as well, however it has the drawback, that the resources used for word sense disambiguation and entity linking cannot be defined or chosen by the user. For knowledge networks to be created across resources and applied to various data sets, there is a need for data to be processed by means of computational linguistic tools and matched preferably against domain specific authority resources.

In this paper we introduce and exemplify such a linking process developed and applied in the context of two connected Digital Humanities projects, APIS³ (Lejtovicz et al., 2015) and exploreAT!⁴ (Wandl-Vogt et al, 2015; Benito et al., 2016; Dorn et al, 2016). The diverse digital networks available to-date have been created around a variety of topics. Some

¹ <https://wordnet.princeton.edu/> [last accessed: 23.06.2017]

² <http://lod-cloud.net/> [last accessed: 23.06.2017]

³ <https://www.oeaw.ac.at/acdh/projects/apis/> [last accessed: 23.06.2017]

⁴ <https://www.oeaw.ac.at/acdh/projects/exploreat/> [last accessed: 23.06.2017]

evolve around networks of places (The Historical GIS Research Network⁵) or of art (e.g. EuropeanaArt⁶), etc. In our case, we apply semantic web tools to interlink person names. In the Digital Humanities project APIS, it is a main goal to unveil connections among people in biographical sources, which provides insightful information on the lives of well-known people. Applying entity-linking in connection with relation extraction - a task addressed in the project APIS - allows us to identify and visualize connections among entities mentioned in different data sources.

This study thus aims at linking existing resources partly containing the same information through the use of semantic web technologies. Through the additional enrichment with LOD, our study aims to show how these data sets can first be connected, and later opened to a wider user audience. This in turn adds to their prolonged re-use and sustainability by ensuring that additions and corrections to the data set only have to be added once to the reference resource, instead of updating all the distinct data resources. In addition, the results of our study also contribute to making information on people networks more widely available also to knowledge society.

2 Data and resources

The data behind the inter-linking process of the projects APIS and exploreAT! are extracted from the resources ÖBL (Austrian Biographical Dictionary; Gruber and Feigl, 2009) and dbo@ema⁷ (Database of Bavarian Dialects in Austria electronically mapped) (cf. Wandl-Vogt et al., 2008). In the realization of both projects, the Austrian Centre for Digital Humanities (ACDH-ÖAW⁸) plays an important role. They rely on data from the respective resources (ÖBL, dbo@ema) which contain similar types of elements such as persons, locations, institutions and titles of written works. In ÖBL this concerns the names of important historical figures, names of cities and countries relevant to

the lives of the people in the biographies, as well as titles of books, journals, or publications mentioned in the biographies. In the dbo@ema, on the other hand, we are dealing with names of locations and regions, names of data collectors or authors and also titles of dictionaries, dissertations and literature. The benefit of linking the above mentioned data sets resides in the possibility to enrich the biographies with missing information contained in the entries of the dbo@ema and vice versa. Often for example the list of literature works is incomplete in either ÖBL or dbo@ema, by linking the two resources, the missing information can be added the other resource.

The *ÖBL* contains around 18.500 biographies and serves as the reference work for APIS, a project which aims to investigate whether a large scale lexicon can be used as the basis of quantitative data analysis and how biographical research can benefit from the digital transformation process realized in APIS. The lexicon contains biographies of important historical figures from the Austro-Hungarian Monarchy having lived in the time period of 1815-1950. The data is not only published in print, but it is also available in the machine readable XML format for the APIS project. An example of a typical ÖBL data entry in XML format is provided in Appendix. It is taken from the biography of *Johann Willibald Nagl*, an Austrian writer and germanist having lived and worked on the turn of the century. The entry contains some structured information in XML elements such as *Geburt* (containing place and date of birth), however the majority of the information (in this specific example referring to the studies and the career path of August Schreiber) is embedded in the unstructured XML element *Haupttext* (i.e. main text). The ÖBL data set contains not only the 18.500 persons the biographies were written about but also additional individuals mentioned in the main text. This set of names together with the persons in dbo@ema creates the basis for connecting the two projects APIS and exploreAT! via an automatic alignment process.

The dbo@ema, on the other hand is to-date a part of the Database of Bavarian dialects in Austria (DBÖ) which forms the basis of the project exploreAT!. The project explores this

⁵ <http://www.hgis.org.uk/> [accessed: 23.06.2017]

⁶ <http://www.europeana.eu/portal/de/collections/art> [accessed: 23.06.2017]

⁷ <https://wboe.oew.ac.at/projekt/beschreibung/> [accessed: 23.06.2017]

⁸ <https://www.oew.ac.at/acdh/acdh-home/> [accessed: 23.06.2017]

large heterogeneous collection of 20th century dialect data of the Bavarian dialects in Austria from perspectives of cultural lexicography, semantic technologies, visual analysis and citizen science. The `dbo@ema` is a MySQL database that comprises of a collection of dialect words of various fields of everyday life. Part of the database comprises of the digitised data originally collected by means of paper questionnaires as well as the digitized entries of the plants (~32.000 headwords) and mushrooms collections (~ 1.000 headwords), also include names of places and regions in the former Austro-Hungarian Empire, as well as names of data collectors or authors of dictionaries, dissertations or literature. Data concerning persons involved in the collection are for the bigger part derived from internal archival material of the institute. Initially, the available questionnaire data was manually entered in TUSTEP (TUEbingen System of TEXT processing Programs)⁹, a professional toolbox for scholarly processing of textual data. All in all, the DBÖ counts around 3.5 million records and an estimated 200,000 headwords.

3 Applying semantic web technologies to inter-link heterogeneous DH data sets

In many projects dealing with digital collections, digital content is generated from scanned books, dictionaries, maps, etc. This is, however, just the prerequisite for establishing a knowledge base which is usable and reusable within and across different disciplines. In order to make data more widely available in a network of relevant sources, the enrichment with Linked Open Data (LOD) is key. Enrichment is a process that has to be established in order to open up DH data sets (e.g. lexicons, encyclopedia, dictionaries, etc.) not only to the public, but also to the members of the research community and to industry.

The projects APIS and exploreAT! face the challenge that the valuable information they contain is embedded in different data models and data formats, and therefore they are not completely transparent and reusable for the

researchers, domain experts and interested citizens. It is also the case in many other Digital Humanities (DH) projects that they partially comprise of the same information embedded in different resources. APIS and exploreAT! have common entity types, among them being persons, locations, names of written works, which when being identified and aligned, can serve as the basis for inter-linking the two projects. This allows for adding missing information from the complementary data set, uncovering and visualizing networks of common entities, and expanding the search space by introducing new, joined data sets to the previously limited research environment.

The motivation to semantically enrich the ÖBL data collection - a historically and culturally rich heritage data - is a main goal in the APIS project. We designed a workflow that is also applicable for the semantic annotation of other DH collections as well. This workflow is set up by first identifying candidates for the linking process, in the second step linking them automatically to LOD resources and finally approving and curating the results. In our study, we link entities to GeoNames and GND, and plan to further extend the pool of used LOD resources with VIAF¹⁰. We use the linked LOD resources to enrich our data with missing information (e.g. to add name variants, latitude, longitude, if available URI of corresponding Wikipedia article, etc. to our data sets), to detect possible errors in our data sets by comparing the information in ÖBL/`dbo@ema` with the information contained in GeoNames/GND, and to make it machine readable and searchable through publishing it eventually in the LOD cloud. However linking to significant vocabularies such as GeoNames and GND do not only provide valuable information, but also challenge computational linguistic systems. Some of the problems are caused by the incompleteness of authority files, not all person/place/institution names are contained in LOD vocabularies. However this problem can be addressed by adding further resources to the system, for this reason we are planning to index VIAF in addition to GeoNames and GND. If an entity is present in a vocabulary, information in

⁹ www.tustep.uni-tuebingen.de/ [last accessed: 23.06.2017]

¹⁰ <https://viaf.org/> [last accessed: 23.06.2017]

a biography might still not be enough to automatically identify the connection. Often the only information about spouses, siblings, tutors, etc. mentioned in the biography are their name and their relationship (father of, spouse, tutor of, etc.) to the person the biography was written about. In this case relation extraction can help to correctly identifying the matching entity. Relational information collected from the biographies can be compared with information in the dictionaries, and in case of matching values, the link between the entities can be proposed by the system. In APIS we implemented a rule based approach using the JAPE¹¹ grammar to detect relations. Further difficulties arise from names, where more than one match is possible with vocabulary entries. Choosing the correct match is called disambiguation, the heuristics we apply for automatic disambiguation consist of fine-tuning the Solr indexes of place names and person names, and adapting them to the characteristics of the input data. We apply heuristics such as indexing only person names from geographical areas relevant to the data sets ÖBL and dbo@ema. Thus we can decrease false matches caused by name-collisions between individuals having born, lived and died in areas other than ÖBL/ dbo@ema related ones.

For the realization of the entity linking, Apache Stanbol¹² has been chosen as an open-source, customizable and extendible implementation framework to work with. The benefit of using Apache Stanbol is, on the one hand its ability to create Referenced Sites (i.e. a local Apache Solr¹³ index of a knowledge base) from any (publicly available) RDF/XML resource and to perform Entity Linking against the compiled site. Furthermore, Stanbol allows the user to take advantage of the integrated Natural Language Processing (NLP) frameworks such as OpenNLP¹⁴ in a free, open source environment. In APIS we have set up a procedure to convert unstructured, full text biographies into structured, semantically enriched and machine-readable documents. This

procedure currently consists of two steps. First, we resolve the abbreviations including the shortened forms of person names, institution names, academic titles, location names, frequent verbs, etc. with a regular expression based Java program to substitute them with their corresponding resolution taken from an ÖBL-intern abbreviations list. Second, we configure and run Stanbol's Entityhub Indexing Tool to create Solr indexes from the resources GeoNames¹⁵ and GND¹⁶ After initializing the index an Enhancement Chain is set up. The Enhancement Chain is on the one hand responsible for running NLP tasks on the biographies (language detection, sentence splitting, tokenization, part-of-speech tagging and chunking) and on the other hand for matching the entities identified by the NLP processor with the Solr index. In our project, the NLP pipeline runs the OpenNLP software with the German model files.

Although correction methods can reduce the error rate of automatic Entity Linking, some manual correction is still required, hence we foresee a manual data curation process to complement and correct the shortcomings of the automatic process.

4 Data set analysis

Analyzing the person names in the data sets ÖBL and dbo@ema the following figures emerged: in the ÖBL (counting the biographies written until the beginning of the project) life stories of 18219 persons comprise the data set of the APIS project, whereas the dbo@ema data resource contains 8841 person names. When aligning the two data sets, results showed that 402 person names are identical, given the criteria that the first name and the last name of the corresponding dbo@ema and ÖBL entries have to match exactly. Due to the fact, that the two data sets differ in how they model personal data (e.g. the ÖBL *second name* contains all the name variants of a person in a comma separated format, whereas the dbo@ema contains a comma in the *second name* before noble titles) the number of matches between the two

¹¹ <https://gate.ac.uk/sale/tao/splitch8.html> [last accessed: 23.06.2017]

¹² <https://stanbol.apache.org/> [last accessed: 23.06.2017]

¹³ <http://lucene.apache.org/solr/> [last accessed: 23.06.2017]

¹⁴ <https://opennlp.apache.org/> [last accessed: 23.06.2017]

¹⁵ <http://www.geonames.org/about.html> [last accessed: 23.06.2017]

¹⁶ http://www.dnb.de/EN/Standardisierung/GND/gnd_nod_e.html [last accessed: 23.06.2017]

resources could be higher after reconciliation. Our analysis thus shows a first rough estimation about how many persons are potentially overlapping in the two collections. Further manual curation is necessary considering that information for the correct identification of a person is often missing in the database. The `dbo@ema` often lacks the information about date and place of birth. In this case additional knowledge, such as the publications or names of relatives can be used to identify and correctly find the person from the `dbo@ema` in the Austrian Biographical Dictionary. When narrowing down the criteria to exactly match on the first name, last name and year of birth, there are only 35 entries found that occur in both resources. The small number of matches can also be attributed to the fact, that in many cases basic information is missing for the exact identification of a person. To overcome this problem, a system has been developed in the frame of the APIS project, where manual curation of entities such as persons, locations, institutions, works and events is possible. We foresee that a manual review process will be carried out after the automatic linking of the `dbo@ema` and `ÖBL` person data sets, in order to approve correctly established links, revise erroneous connections and add missing information to both data sources.

The following example illustrates how the knowledge sources `ÖBL` and `dbo@ema` are connected to each other via the GND URI assigned to *Johann Willibald Nagl*, an Austrian writer and Germanist appearing in both data sets. Nagl's `ÖBL` biography has been published online, and his personal data (name, date and place of birth, date and place of death) is also recorded in the `dbo@ema` database (see the two entries of Nagl in the Appendix). The link between the two instances has been established by means of the Stanbol Entity Linking Module, which identifies *Johann Willibald Nagl* as a candidate for entity matching and looks it up in the Solr index created from GND person names. Below we show an excerpt of the semantic annotation created by Stanbol. The URI <http://d-nb.info/gnd/116880414> links the two occurrences of *Johann Willibald Nagl* and thus the two resources `ÖBL` and `dbo@ema`.

```
{
```

```
  "@id": "urn:enhancement-41adec0e-9ebc-8d19-7644-b799288d563b",
  "@type": [
    "Enhancement",
    "EntityAnnotation"
  ],
  "confidence": 1.0,
  "created": "2017-06-22T16:25:27.384Z",
  "creator": "org.apache.stanbol.enhancer.engine.s.entitylinking.engine.EntityLinkingEngine",
  "entity-label": "Nagl, Johann Willibald",
  "entity-reference": "http://d-nb.info/gnd/116880414",
  "entity-type": "http://d-nb.info/standards/elementset/gnd#DifferentiatedPerson",
  "extracted-from": "urn:content-item-sha1-3dee9b203b74c12fec298348e74a1a0f16ee7da2",
  "relation": "urn:enhancement-e1a4dcdd-e9fc-d9fc-42d4-b4e7cabb4685",
  "site": "gndPersons"
}
```

With the help of a web application being developed in APIS we are planning to evaluate the quality of the linking process. The application is designed to support automatic and manual annotation within one system, thus allowing automatic evaluation of annotation tasks.

5 Discussion and Conclusion

In this paper we discussed the linking of person names in two data sets, the `ÖBL` and `dbo@ema`. Our applied method has shown, that through the automatic entity linking process, the same persons occurring in different resources can be detected and connected. Through the established links and by applying the relation extraction method implemented in the APIS project, a link across the data sets `ÖBL` and `dbo@ema` can be revealed, giving valuable information of relations among persons mentioned. Our method is only in its developing stages and this paper is a first introduction. By generating person networks including additional information existent in the `ÖBL` or `dbo@ema`, our “social network” could provide a valuable

source of information also for non-specialists. As persons mentioned in the two resources are also connected to a variety of personal information (profession, birth place, etc.), opening up and connecting our data sets to other services for societal benefits is another main goal. Services that could potentially benefit from our generated knowledge include Europeana collections or Museums. Connecting the information from our ÖBL and dbo@ema resources to current collections would offer a fruitful collaboration for giving citizens access to otherwise hidden information.

References

- Benito, A., Losada, A. G., Therón, R., Dorn, A., Seltmann, M., Wandl-Vogt, E. (2016): A spatio-temporal visual analysis tool for historical dictionaries. TEEM 2016. Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality: pp. 985-990
- Gruber, C., Feigl, R. (2009) Von der Karteikarte zum biografischen Informationsmanagementsystem. Neue Wege am Institut Österreichisches Biographisches Lexikon und biographische Dokumentation, in: Martina Schattkowsky / Frank Metasch (eds.), Biografische Lexika im Internet. Internationale Tagung der „Sächsischen Biografie“ in Dresden (30. und 31. Mai 2008) (= Bausteine aus dem Institut für Sächsische Geschichte und Volkskunde 14), Dresden: Thelem Universitätsverlag, 2009, pp. 55–75
- Dorn, A., Wandl-Vogt, E., Bowers, J., Piringner, B., Seltmann, M. (2016) exploreAT! – perspectives of exploring a dialect language resource in a framework of European digital infrastructures. 1st International Congress on Sociolinguistics (ICS-1), Budapest, Hungary.
- Lejtovicz, K., Durco, M., Schlögl, M., Wandl-Vogt, E. (2015) APIS New Austrian Prosopographical Information System. Mapping Historical Networks. 2nd DHA Conference. Vienna, Austria. DOI: [10.15169/sci-gaia:1473321487.86](https://doi.org/10.15169/sci-gaia:1473321487.86)
- Mendes, Pablo N., Jakob, Max, García-Silva, Andrés, and Bizer, Christian, “DBpedia spotlight: shedding light on the web of documents”. In: Proceedings of the 7th International Conference on Semantic Systems, page 1-8. New York, NY, USA, ACM, (2011)
- Moro, A., Raganato, A., Navigli, R. (2014) Entity Linking meets Word Sense Disambiguation: a Unified Approach. Transactions of the Association for Computational Linguistics (TACL), 2, pp. 231-244.
- Schopper, D., Bowers J., Wandl-Vogt, E (2015) dbo@TEI: remodelling a database of dialects into a rich LOD resource. Proceedings of TEI conference 2015.
- Wandl-Vogt, E., Bartelme, N., Fliedl, G., Hassler, M., Kop, C., Mayr, H., Nickel, J., Scholz, J., Vöhringer, J. (2008): dbo@ema. A system for archiving, handling and mapping heterogeneous dialect data for dialect dictionaries. In: Bernal, Elisenda / De Cesaris, Janet (Hrsg.): Proceedings of the XIII Euralex International Congress, Barcelona, Universitat Pompeu Fabra, 15.-19. Juli 2008 (= Sèrie activitats 20). Barcelona (Documenta Universitaria). S. 1467-1472 (CD-ROM).
- Wandl-Vogt, E., Kieslinger, B., O’Connor, A., Theron, R. (2015). „exploreAT! Perspektiven einer Transformation am Beispiel eines lexikographischen Jahrhundertprojekts“, in: DHD-Tagung 2015. Graz. Austria. Accessed at: <http://dhd2015.uni-graz.at> [23.06.2017]

Appendix

ÖBL entry of Johann Willibald Nagl:

```
<?xml version="1.0"
encoding="utf-8"?>
<Eintrag
xmlns="http://www.biographien.ac.at"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.biographien.ac.at
https://aspix2.lgbs.at/GIDEON_NG_OEBL/userdefined/Biografien/XML/XSD/OEBL-Bio-V1.xsd"
Nummer="Nagl_Johann-Willibald_1856_1918.xml"
Version="01" pnd="116880414"
eobl_id="1410752">
<Lexikonartikel>
<Schlagwort>

<Hauptbezeichnung>Nagl</Hauptbezeichnung>
<Nebenbezeichnung
Type="Vorname">Johann
Willibald</Nebenbezeichnung>
</Schlagwort>

<Sortierung_Nachname>Nagl</Sortierung_Nachname>
```

<Sortierung_Vorname>Johann Willibald</Sortierung_Vorname>

<Schlagwort_Nachname>Nagl</Schlagwort_Nachname>

<Schlagwort_Vorname>Johann Willibald</Schlagwort_Vorname>

<Vita>

<Geburt Metadatum="1856" TT="11" MM="5">(1856-
<Geographischer_Begriff
OrtAlt="Natschbach b. Neunkirchen"
OrtNeu="?" LandAlt="NÖ"
LandNeu="Österreich/NÖ">Natschbach
b. Neunkirchen (?,
NÖ)</Geographischer_Begriff></Geburt
>

<Tod Metadatum="1918" TT="23" MM="7">1918)<Geographischer_Begriff
OrtAlt="Diepolz b. Neunkirchen"
OrtNeu="?" LandAlt="NÖ"
LandNeu="Österreich/NÖ">Diepolz b.
Neunkirchen (?,
NÖ)</Geographischer_Begriff></Tod>

<Beruf
Berufsgruppe="Geisteswissenschaft">
Germanist und
Schriftsteller</Beruf>

<Beruf
Berufsgruppe="Literatur, Buch- und
Zeitungswesen" />

</Vita>

<Geschlecht Type="m" />

<Kurzdefinition>Nagl Johann
Willibald, Germanist und
Schriftsteller. * Natschbach b.
Neunkirchen (NÖ), 11. 5. 1856; †
Diepolz b. Neunkirchen (NÖ), 23. 7.
1918.</Kurzdefinition>

<Haupttext>Stud. nach einem
bald wieder abgebrochenen Theol.-
Stud. Phil. und Germanistik an der
Univ. Wien, 1886 Dr. phil. Neben
seiner Lehrtätigkeit an
verschiedenen Schulen war N. ab
1890 als Priv. Doz. für
Mundartforschung an der Univ. Wien
tätig. Er darf neben Seemüller zu
den Initiatoren der Wr.
mundartkundlichen Schule (z. B. als
Hrsg. der Z. „Deutsche Mundarten“)
gezählt werden, wenn auch manche
von ihm angeschnittene Probleme
später anderen Lösungen zugeführt
wurden. Schon als Schottenkleriker
hatte N. begonnen, die alte
Tierfabel von Reineke Fuchs in

seiner niederösterr. Heimatmundart
darzustellen. Als Vorlage für das
Dialektepos „Der Fuchs Roáner, á
lehrreichs und kürzweiligs
Gleichnus aus derselbigen Zeit, wo
d'Viecher noh hab'n red'n künná.
Aus uralten, vierhundert- bis
sechshundertjährigen Büchern neu in
die Welt gestellt für die
österreichischen Landsleute“
dienten Goethes „Reineke Fuchs“,
aber auch die alten Texte des
Reinaert und des Reinke de vos. N.
gelang es dabei nicht nur, den
niederösterr. Bauerndialekt,
sondern auch die gesamte bäuerliche
Anschauungswelt des Neunkirchner
Raumes lebendig darzustellen. Gem.
mit Zeidler begründete N. außerdem
die vierbändige „Deutsch-
österreichische
Literaturgeschichte“, die später
von Castle fortgesetzt wurde.
Überdies befaßte sich N. mit Stud.
über den niederösterr. Bauernstand,
von denen er einige im Selbstverlag
veröff.

</Haupttext>

<Werke>W.: Da Roanad.
Grammatik des niederösterr.
Dialekts, 1886; Der Fuchs Roáner .
. ., 1889, 2. Aufl. 1909;
Vokalismus der bayr.-österr.
Mundart, 1895; Geograph. Namenkde.,
in: Die Erdkde. 18, 1903; Dt.
Sprachlehre . . ., 1905, 2. Aufl.
1906; etc. Hrsg.: Dt. Mundarten,
1896 ff.; Dt.-österr.
Literaturgeschichte, 4 Bde., gem.
mit J. Zeidler und E. Castle, 1899-
1937.

</Werke>

<Literatur>L.: RP vom 2. und
11. 5. 1916, 27. 7. und 15. 8.
1918; Wr. Ztg. und N. Fr. Pr. vom
26. 7. 1918; Z. für österr.
Volkskde., Jg. 3, 1897, S. 319, Jg.
4, 1898, S. 52; Monatsbl. des Ver.
für Landeskd. von NÖ, Jg. 17,
1918, S. 190 ff.; Petermanns Mitt.,
1918, S. 228; Unsere Heimat, NF,
Bd. 11, 1938, S. 200 ff.; I. M.
Swift Peacock, Der grammat. Anhang
J. W. N.s „Fuchs Roánad“ im
Vergleich mit dem heute lebendigen
Wortschatz in der Mundart der
Gemeinde Hafning, Bez. Neunkirchen,
NÖ, phil. Diss. Wien, 1969;

Giebisch-Gugitz; Kosel; Rollett, Neue Beitr., Tl. 10, 1898, S. 80; Kosch, Das kath. Deutschland; Wer ist's? 1905-14.

</Literatur>

<Autor>(M. Hornung)

</Autor>

<PubInfo Reihe="ÖBL 1815-1950" Band="7" Lieferung="31" Seite="21" Jahr="1976" Monat="" Tag="">ÖBL 1815-1950, Bd. 7 (Lfg. 31, 1976), S. 21</PubInfo>

</Lexikonartikel>

</Eintrag>

Excerpt of the dbo@ema entry of Johann Willibald Nagl:

```
<database name="dboe_1">
  <table name="person">
    <column
name="id">12102</column>
    <column name="vorname">Johann
Willibald</column>
    <column
name="nachname">Nagl</column>
    <column
name="gebTag">11</column>
    <column
name="gebMonat">5</column>
    <column
name="gebJahr">1856</column>
    <column
name="gebOrt">Natschbach          b.
Neunkirchen, NÖ</column>
    <column
name="gebOrt_id">7082</column>
    <column
name="todTag">23</column>
    <column
name="todMonat">7</column>
    <column
name="todJahr">1918</column>
    <column name="todOrt">Diepolz
b. Neunkirchen, NÖ</column>
    <column
name="todOrt_id">NULL</column>
    <column
name="geschlecht">2</column>
    <column
name="adresse"></column>
    <column name="plz">-1</column>
    <column name="ort"></column>
    <column name="email"></column>
    <column name="tel1"></column>
    <column name="tel2"></column>
    <column name="tel3"></column>
```

```
<column
name="adressverlauf"></column>
  <column
name="verwandschaft">Mutter: ---
Geburtsdatum: --- Todesdatum: ---
Anm.: --- (bereits in Datenbank:
ja/nein) Vater: --- Geburtsdatum:--
- Todesdatum: --- Anm.: ---
(bereits in Datenbank: ja/nein)
Gattin/Gatte: --- Geburtsdatum: ---
Todesdatum: --- Anm.: --- (bereits
in Datenbank: ja/nein) Weitere
Verwandte: --- Anm./Verweise: ---
  </column>
  <column
name="kontaktperson"></column>
  <column
name="ausbildung">Regierungsrat
Dr.phil. Schule: Universität ---
Ort: --- von: --- bis: --- Anm:
Theologie; abgebrochen --- Schule:
Universität --- Ort: Wien --- von:
--- bis: 1886 --- Anm: Phil. und
Germanistik; 1886 Dr.phil. ---
Schule: --- Ort: --- von: --- bis:
--- Anm: --- Beruf: Lehrer --- Ort:
--- von: --- bis: --- Anm: an
verschiedenen Schulen --- Beruf:
Priv. Dozent für Mundartforschung -
-- Ort: Universität Wien --- von: -
-- bis: --- Anm: --- Beruf:
Schriftsteller --- Ort: --- von: --
- bis: --- Anm: --- Beruf:
Herausgeber der Zeitschrift
„Deutsche Mundarten“ --- Ort: ---
von: --- bis: --- Anm: --- Beruf: -
-- Ort: --- von: --- bis: --- Anm:
--- Ehrenamtl.
Tätigkeiten:</column>
  </table>
</database>
```