

“Vernetzte Kirche”: Building a Semantic Web

Sören Auer* and Bart Pieterse+

* University of Leipzig, 04109 Leipzig, Germany,
auer@informatik.uni-leipzig.de,

WWW home page: <http://www.informatik.uni-leipzig.de/~auer/>

+ Vernetzte Kirche, Marsstr. 19, 80335 München,
bartel.pieterse@elkb.de,

WWW home page: <http://www.vernetzte-kirche.de>

Abstract The only possibility for federally structured organizations and communities to enable consistent views on their meta-data and contents is to establish methods to gather the meta-data from the distributed peers, integrate it into a common conceptual model and finally make it accessible to humans and software systems. We present the approach taken to implement this strategy for the over 2000 affiliated organizations of the Bavarian Lutheran Church. It showcases how different Semantic Web standards, vocabularies, and methodologies can be coherently integrated into a consistent framework delivering added value to the large audience of participating peers.

1 Introduction

The Lutheran Church in the German federal state Bavaria has around 2000 affiliated clerical, cultural or charitable communities, organizations, and societies. All of them have their separate organizational structure with different spheres of activities, websites, contact information and various other information.

Unlike strictly hierarchical structured organizations such as business cooperations, for organizations which profit most from their social and cultural diversity, such as the Lutheran Church, the usage of a single unified data model or system (e.g. content or knowledge management system) to manage such information will be impossible.

Nevertheless it is highly desirable to give people an overview of lutheran organizations, activities, and content related to a geographical region or to a specific topic. Thus a web portal was requested interlinking all these resources and making them easily accessible for interested parties within and outside the Lutheran Church of Bavaria.

To reach this aim a Semantic Web application was envisioned, which collects information and meta data from distributed sources with diverse ownerships, integrates them into a common extensible conceptual model, which should be finally represented on the Web, exposing as many semantic relations as possible. An additional aim is to establish a meta-data initiative focused on religious

related content and to support individuals and organizations accordingly with its application.

The implementation of a prototype was realized by the department “Vernetzte Kirche” within the Bavarian Lutheran Church in cooperation with the University of Leipzig. The implementation is based on the Semantic Web application development framework Powl [1]. Powl itself makes use of the Web technologies with greatest deployed base (PHP and MySQL) and enables the rapid development of Semantic Web applications as well as appropriate ontologies.

The “Vernetzte Kirche” use case of Semantic Web technologies is presented here as follows: We first describe the information structure used to represent the various types of information. In Section 3 we give account how this information structure is populated with concrete data from the distributed, heterogeneous sources. We elaborate in Section 4 on ways to make this information accessible to humans and software systems. Finally we give some concluding remarks and an outlook on planned enhancements in Section 5.

2 Structuring Information

The first challenge was to find an integrated conceptual model for the information about relevant organizations and resources, which should be flexible enough to allow consistent integration of new information as it arises.

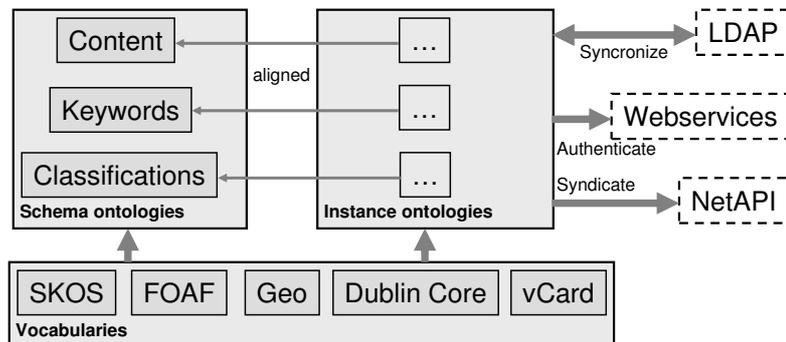


Figure 1. The ontological structure and interactions

Because of the heterogenic structure of this information an ontology seems to be more adequate to capture it than for example a fixed database schema. Such an ontology was developed in OWL [2] using Powl’s schema editor. Its overall ontological structure is depicted in Figure 1 and described in more detail in the remainder of this section.

The main concepts (or classes) were quickly identified: **Organizations** and **Content**. The class **Content** is categorized into the subclasses **Images**, **Books**,

News, Events and Journals; additional classes can be added as needed. The classification of organizations though is more complicated, since every organization can be classified along different taxonomies:

- Regional,
 - By German administrative units (federal state; district; region; community or city),
 - By established regional church structure (regional church; deanery; clerical district; religious community),
- Sphere of activity (e.g. education, arts, and culture; administration; souls care),
- Type of organization (e.g. choir; community; publishing house),
- Topics and keywords.

For the spheres of activity, types of organizations, topics and keywords separate classes containing taxonomies of instances are introduced. The taxonomic structure is encoded using a functional object property `subTypeOf` having the respective class as domain and range. An elaborated keyword taxonomy was created using the SKOS [9] methodology and much inspired by the IPTC subject reference system [11]. This allows for organizations and content to (such as images provided by the photo service “FotoBayern”¹ or news provided by RSS feeds) to be annotated in a consistent manner.

Address information for a specific organization is used in conjunction with the data of OpenGeoDB² to calculate its geographical coordinates and relate the instance to the regional classifications. The geographical vocabulary (with the namespace `http://www.w3.org/2003/01/geo/wgs84_pos#`) is used here to capture this information within the ontology.

For initial acquisition of instance data Powl’s feature to import rows from an Excel sheet as instances was very beneficial, since most of the information was already available within conventional Office documents. As Powl enables the customized generation of reports in spreadsheet compatible formats the traditional storing of information within Excel documents became needless.

The complete ontology schema is available from the “Vernetzte Kirche” web site³ and intended to be reusable by other established regional churches or similar federally structured associations. Currently, it contains around 70 classes and 40 properties.

3 Continuous Acquisition of Information

To populate the information structure presented in the last section, instance data has to be continuously acquired, synchronized, or updated. For content related meta-data, specialized acquisition methods have been developed:

¹ <http://www.fotoBayern.de>

² <http://opengeodb.de/>

³ <http://www.vernetzte-kirche.de>

- *Images* - meta data is envisioned to be integrated into web-accessible image files according to the proposed “Vernetzte Kirche” classifications. Common image formats, such as JPEG or PNG offer the ability to save meta-data within the file in custom text formats. Images and meta-data thus can be gathered from specified web sources, such as the photo service “FotoBayern”.
- *News* - such as from publications as the “Sonntagsblatt”⁴ or the lutheran press service EPD⁵ are syndicated by accessing respective RSS feeds. News feeds are encouraged to use the “Vernetzte Kirche” or IPTC classifications to tag their news.
- *Events* - for events as concerts, religious services or courses a specialized database and administration portal “Evangelische Termine”⁶ existed. Information is exchanged by establishing a synchronization between the database schema and the ontology in the spirit of D2RQ [3].

To continuously acquire instance data about organizations a distributed ownership strategy was selected:

1. Organizations integrate meta information about them self inside the HTML header of their website.
2. A crawler regularly gathers this meta data and updates existing instances accordingly.

To ease the task of integrating meta information into the organization websites a meta-tag generator was developed. A publisher thus has to complete a simple web form to generate an HTML header snippet for integration into the organizations website. The integration of meta data and its extraction is implemented according to the GRDDL mechanism [6]. The following code shows an example meta tag snippet for integration into an HTML header:

```
<meta name="DC.title" content="Vernetzte Kirche" />
<meta name="DC.creator" content="Bart Pieterse" />
<meta name="DC.subject" content="Semantisch Vernetzte Kirche" />
<meta name="DC.description" content="Informationen zur AG
  Vernetzte Kirche der Evangelisch-Lutherischen Kirche in
  Bayern und semantische Suche nach ELKB Ressourcen" />
<meta name="DC.publisher" content="Vernetzte Kirche" />
<meta name="DC.date" content="2005-05-03" />
<meta name="DC.language" content="de" />
<meta name="DC.coverage" content="Bayern" />
<meta name="DC.rights" content="GFDL" />
<meta name="geo.lat" content="48.1333" />
<meta name="geo.long" content="-11.5833" />
<meta name="geo.country" content="BRD" />
<meta name="dmoz.id" content="" />
<meta name="keywords" content="Vernetzte Kirche, Semantic Web,
  ELKB, semantische Suche" />
<meta name="robots" content="follow" />
<meta name="revisit-after" content="20 days" />
```

⁴ <http://www.sonntagsblatt.de>

⁵ <http://epd.de>

⁶ <http://www.evangelische-termine.de/>

By using attributes of commonly used vocabularies (such as Dublin Core [8], vCard [7], FOAF [5], Geo [4]) this meta information can be easily interpreted by other (search) agents as well.

Since some, user and organization related information, is held in an LDAP directory, for extranet authentication purposes, a synchronization between distinct instance and property values and LDAP nodes was implemented.

The overall instance data related to the Bavarian Lutheran Church contains currently around 5,000 instances, accounting for 60,000 triples.

4 Sharing the Information

Projekt Vernetzte Kirche

Nach Regierungskreisen:
 Mittelfranken
 Niederbayern
 Oberbayern
 Oberfranken
 Oberpfalz
 Schwaben
 Unterfranken

Nach Dekanaten:
 Kirchenkreis Ansbach - Würzburg
 Dekanat Ansbach
 Dekanat Aschaffenburg
 Dekanat Bad Neustadt
 Dekanat Bad Windsheim
 Dekanat Castell
 Dekanat Dinkelsbühl
 Dekanat Feuchtswangen
 Dekanat Gunzenhausen
 Dekanat Heidenheim
 Dekanat Kitzingen
 Dekanat Leutershausen
 Dekanat Lohr am Main
 Dekanat Markt Einersheim
 Dekanat Rothenburg o.T
 Dekanat Schweinfurt
 Dekanat Uffenheim
 Dekanat Wassertrüdingen
 Dekanat Windsbach
 Dekanat Würzburg

Kirchenkreis Ansbach-Würzburg
 Kirchenkreis Bayreuth
 Kirchenkreis Nürnberg
 Kirchenkreis Regensburg
 Kirchenkreis Augsburg
 Kirchenkreis München

Handlungsfelder:
 Spiritualität, Gottesdienst,
 Verkündigung und
 Kirchenmusik
 Gemeindeaufbau und
 Gemeindeentwicklung
 Seelsorge und Beratung
 Themen- und
 zielgruppenbezogene
 gesellschaftliche Dienste
 Ökumene, Mission,
 Entwicklungsdienst und
 Partnerschaft
 Diakonisches Handeln
 Presse, Öffentlichkeit, Medien
 Aus-, Fort- und Weiterbildung

Organisationsart:
 Amt
 Beratungsstelle
 Bildungszentrum
 Chor
 Diakonisches Werk
 Evangelische Jugend
 Hochschule
 Institut
 Kirchengemeinde
 Materialstelle
 Medienstelle
 Sonstige Organisation
 Stadtmission
 Tagungsstätte
 Verlag
 Wohnheim

Suche:

Organisationen im Web
 Die Suche ergab **954** Ergebnisse.[c]
 |< | 1 | 2 | 3 | 4 | ... | >|

Amt Amt für Gemeindedienst Nürnberg
[Details](#) [Website](#)

Diakonisches Werk Innere Mission München
[Details](#) [Website](#)

Beratungsstelle AIDS-Beratung Mittelfranken der Stadtmission Nürnberg
[Details](#) [Website](#)

Kirchenkreis Augsburg
 Kirchenkreis Bayreuth
 Kirchenkreis München
 Kirchenkreis Nürnberg
 Kirchenkreis Regensburg

Figure 2. The “Vernetzte Kirche” Web Portal

After filling the meta-data repository, this information should be shared with “human” and artificial agents. For machine consumption access is provided by Web-Services or the NetAPI [10], which allows to share RDF ontology (fragments) by using the HTTP protocol. For humans a bunch of different, domain or topic specific views on the data is planned. Initially Powl’s application programming interface is used to present the information in a portal style way on the web site <http://www.vernetzte-kirche.de>. Users are enabled to efficiently search and filter the instance data in the ontology.

The user interface of the web portal consists of two different types of pages: overview pages with lists of instances and detail pages, exposing the attributes of a distinct instance with all the specifics. The overview pages for organizations follow a three column layout (as depicted in Figure 2):

- the left part contains the regional filtering options (by German administrative units or established regional church structure),
- in the middle part the list of organizations in the selected region (and according to optionally applied filters) is presented,
- on the right the results can be filtered according to available taxonomical categorizations available for organizations in the selected region.

By choosing the details link for an organization in the list the user reaches a web page combining all information about that organization. The calculated geo coordinates for a specific organization are used to render a map with the location highlighted and to get a list of nearby organizations.

A full-text search is provided, to query the data for literal descriptions containing the search string. Results are grouped by the instances and ordered by the occurrence frequency of the search string. Search results may be filtered to contain only resources which are instances of a distinct class (e.g. restrict the search to organizations or books) or which are described by the literal only in conjunction with a distinct property (e.g. `dc:author`). This semantic search has significant advantages compared to conventional full-text searches, since it provides important feedback to the searching user on how to successfully refine the search.

The portal framework further allows to export or extract complete meta-data sets or information fragments by NetAPI [10] or Web-Service interfaces. This strategy is in particular used to syndicate content and authenticate participating organizations and individuals at partnering web sites. The photo service “FotoBayern” for example gives discounts to individuals and organizations collaborating within the “Vernetzte Kirche” project.

5 Conclusion and Outlook

Semantic Web knowledge representation standards and especially the generic Semantic Web application development framework Powl enabled the project “Vernetzte Kirche” to rapidly develop an initial version of the ontology and a prototype of the web portal with minimized resource and cost demands.

Though already publicly accessible the application is still in Beta stage. First experiences with real users showed the practicability of the approach. Currently, it is worked on the transition into an easy to use, productive, scalable industry strength system. To achieve this goal usability and documentation as well as the stability of the crawling solution has to be improved further. It is additionally planned to implement enhanced client-side meta-data annotation and extraction methods with the focus for usage within the “Vernetzte Kirche” project.

Summarizing we can state that the project showcases how different Semantic Web standards, vocabularies, and methodologies can be coherently integrated into an consistent framework delivering added value to a large audience of participating peers.

References

1. Sören Auer. Powl: A web based platform for collaborative semantic web development. In Sören Auer, Chris Bizer, and Libby Miller, editors, *Proceedings of the Workshop Scripting for the Semantic Web*, number 135 in CEUR Workshop Proceedings, Heraklion, Greece, 05 2005.
2. Sean Bechhofer, Frank van Harmelen, Jim Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein. Owl web ontology language reference. W3C Recommendation (<http://www.w3.org/TR/owl-ref/>), 2004.
3. Christian Bizer and Andy Seaborne. D2rq -treating non-rdf databases as virtual rdf graphs (poster). Poster at Third International Semantic Web Conference (ISWC2004), Hiroshima, Japan, November 2004.
4. Dan Brickley. Rdf vocabulary for representing lat(itude), long(itude) and other information about spatially-located things (technical report), 2001.
5. Dan Brickley and Libby Miller. Foaf vocabulary specification. <http://xmlns.com/foaf/0.1/>, 2005.
6. Dominique Hazaël-Massieux and Dan Connolly. Gleaning resource descriptions from dialects of languages (grddl). <http://www.w3.org/TR/grddl/>, 13 April 2004.
7. Renato Iannella. Representing vcard objects in rdf/xml (technical report), 2001.
8. Dublin Core Metadata Initiative. Dcmi metadata terms. <http://dublincore.org/documents/dcmi-terms/>, 2005.
9. Alistair Miles and Dan Brickley. Skos core guide. <http://www.w3.org/TR/swbp-skos-core-guide/>, 2005.
10. Andy Seaborne. An RDF netAPI. In I. Horrocks and J. Hendler, editors, *Proceedings of the First International Semantic Web Conference (ISWC2002)*, volume 1201 of *Lecture Notes in Computer Science*. Springer-Verlag GmbH, February 11 2002.
11. Misha Wolf. Metadata framework technical specification (iptc standards draft). http://iptc.org/pdl.php?fn=DRAFT-NAR_1.0-spec-NMDF-TechSpec_6.pdf, 2005.