Topic-Oriented Portals

Alexander Sigel
University of Cologne, Germany

Khalil Ahmed
Networked Planet Limited, UK

INTRODUCTION: FROM PRESENTATION-LEVEL INTEGRATION TO CONCEPT-LEVEL INTEGRATION IN PORTALS

In general, portals are regarded as gateways to networked information and services, facilitating access to other related sites. Typically, portals provide transparent one-stop access to functionalities needed in a common context and make these appear as a single integrated application. Although these functionalities are implemented and made available by heterogeneous applications, they are integrated by presenting the output of such networked applications side-by-side, with only limited interaction between them. This superficial integration on the presentation level is quite useful, but it has considerable drawbacks compared to basing portals on the notion of subjects to achieve "seamless knowledge" and the "semantic superhighway" (Pepper, 2004, 2006; Pepper & Garshol, 2004). Consider these challenges:

- **Knowledge Organization:** How can a portal user be directed in a principled way from a given asset about a subject to other relevant assets about this subject, to other related subjects and the assets linked to them?
- **Portal Integration:** How can all assets about a given subject be made virtually accessible from one page in the portal, even if the content about the subject is distributed to several independently maintained portals?

Both challenges require an appropriate information architecture with an integration on the conceptual level. Content- and semantics-based portal approaches allow expressing a domain knowledge model (concepts and their interrelations) and connecting it to related resources.

After briefly sketching how topic-oriented portals (TOPs) are related to knowledge portals, ontology-based and semantic portals, we discuss Topic Maps-based portals (TMPs), a specific form of topic-oriented (or subject-oriented) portals with advantages in their creation, usage and maintenance. What are examples for such portals and their virtual integration? Which elements are appropriate for their information architecture? How could a subject-oriented information architecture be based on knowledge organization systems (KOS)?

For the remainder of this article, it is assumed that the reader is acquainted with basic Semantic Web concepts, including Topic Maps (Ahmed & Moore, 2005; Passin, 2004).

BACKGROUND: TOPIC-ORIENTED PORTALS

A *topic-oriented (or subject-oriented) portal*[1] is a Web-based information or knowledge portal application where informational content is presented in relationship to subjects and whose structure, or "information architecture" (Rosenfeld & Morville, 2002), makes use of Semantic Web technologies. A TOP is not necessarily a portal dedicated to a special narrow selection of topics or subjects. A TOP typically presents a page-per-subject view of a knowledge domain: Each subject has its own "homepage" that displays the information related to that subject and the relationships between the subject and other subjects in the knowledge domain. The pages about the related subjects may reside on the same portal, or on separate portals. A shallow ontology, specifying the subjects of interest and their interconnecting relationships, is the basis of the information architecture which structures the navigation, content integration and rendering, and the search facilities of content-rich Web sites. This knowledge structure on the ontology layer is by design separated from the resource layer. The resource layer describes which resources (assets such as documents) exist, and the connection between both layers specifies which resources are relevant to which subjects, and in which way. By this separation of layers, the knowledge structure becomes portable. This means that it can be superimposed upon different content, thus grouping assertions and relevant resources referring to the same subject together, and identifying and showing assertions and resources referring to related subjects in a principled way. A TOP can be implemented with different Semantic Web technologies, in particular with RDF/OWL or Topic Maps.

Related Work on Knowledge Portals, Ontology-Based and Semantic Portals

No reference to TOPs is made in recent work on knowledge, ontology-based and semantic portals (cf. Hädrich & Priebe, 2005a, b; Lara et al., 2004; Lausen et al., 2005), or (Hartmann & Sure, 2004) for the SEAL (SEmantic portAL) conceptual framework. However, the approaches are all closely related:

A *knowledge portal* is an information portal supporting knowledge workers in their tasks. It is a specific type of enterprise portal, comprising support for information content storage and retrieval, organizational communication and group collaboration (cf. Detlor, 2004, p. 13). An *ontology-based* portal (cf. Staab et al., 2000) is a portal employing ontologies as its semantic backbone, mainly for information integration, navigation and search. A *semantic portal* is a portal using Semantic Web technologies.

A TOP can be a knowledge portal, since the subject-centric integration supports knowledge management and organization, and knowledge workers may be supported in all three of Detlor's dimensions. Since the knowledge net has an ontologic layer, it is an ontology-based portal, and it is a semantic portal, because Semantic Web technologies are used for the representation and manipulation of the ontology and the metadata about the resources. The creation of semantically linked Web pages from Semantic Web content can result in TOPs (Hyvönen, Valo, Viljanen, & Holi, 2003). These authors acknowledge the similarity of their approach to Topic Maps, except they infer the linkage structure instead of specifying it.

TOPIC MAPS-BASED PORTALS

This article focuses on *Topic Maps-based (or Topic Maps-driven) portals* (TMPs), that is, TOPs realized with Topic Maps. TMPs are "the most common application of Topic Maps today," and "also by far the most visible" (Garshol, 2006c). Topic Maps are understood as standardized in the second edition of ISO/IEC 13250 Topic Maps (ISO 13250), with TMDM (Topic Maps Data Model) (ISO 13250-2) to become part of this standard. The interested reader is directed to the introductory book on XML Topic Maps (Park & Hunting, 2002) and to the discussion of research issues at the TMRA conferences (Maicher & Park, 2006).

Examples of Topic Maps-Based Portals

Almost any Topic Map rendered in a subject-centric way as interlinked Web pages, aggregating everything known about a particular subject, will lead to a TMP, for example, the so-called "Italian Opera Topic Map" by Steve Pepper.[2] Every topic has its own "homepage" which displays the semantic relation of this topic to other topics and the assets (resources) relevant to this topic, and all pages are interlinked. Several TMPs are in practical use, predominantly in Norway, for example Kulturnett,[3] a Norwegian public sector portal to cultural information. The IRS Topic Map[4] by Michel Biezunski is a prominent example in English language. Barta (2004) communicates his experiences in developing a Perl-based knowledge portal using Topic Maps; Pepper and Garshol (2002) based theirs on building a TMP of conference papers. For more practical examples of TMPs in the public sector and their virtual integration, see Pepper (2004) and Garshol (2006c).

Examples of the Virtual Integration of Topic Maps-based Portals

Consider you want to connect at least two portals such that topic pages about the same topic provided by both portals are virtually integrated (cf. the simple portal connecting scenario of Pepper, 2004). Portal A can ask portal B if it knows anything about this topic, update its own knowledge base with the answer on-the-fly, and present its changed pages to the user. For example, the three following Norwegian TMPs (of the Research Council Web site for young adults,[5] the public site of the Consumer Association,[6] and the biosecurity portal of the Department of Agriculture[7]) have topics about the common subject "genetically modified food" which they can mutually share, based on published subjects. Connecting portals is just one of four use cases for TMRAP, a remote access protocol for Topic Maps (Garshol, 2006a). Using published subjects and TMRAP as a vehicle to realize the idea of "seamless knowledge," portals can also share little Topic Map fragments, thus automatically syndicating, synchronizing and aggregating knowledge structures and the accompanying resources (Garshol, 2006a, c).

Elements of the Information Architecture of Topic Maps-Based Portals

Topic Maps and TOPs match well, because Topic Maps already exhibit all features for the implementation of TOPs. In particular, basic principles of Topic Maps and knowledge organization aid in the design of an appropriate information architecture:

- **Subject Centric View:** In contrast to the more resource-centric RDF, with subjects and topics, Topic Maps are by design subject-centric, making it easier to talk about subjects.
- **Semantic Interoperability with RDF:** All work on RDF-based portals can be reused, since Topic Maps are semantically interoperable with RDF (Garshol,