Semantic Portals for Cultural Heritage

Eero Hyvönen

Semantic Computing Research Group
Helsinki University of Technology (TKK) and University of Helsinki
eero.hyvonen@tkk.fi, http://www.seco.tkk.fi/

1 Benefits of Cultural Semantic Portals

Cultural content on the web is available in various forms (documents, images, audio tracks, videos, collection items, learning objects etc.), concern various topics (art, history, handicraft, etc.), is written in different languages, is targeted to both laymen and experts, and is provided by different independent memory organizations (museums, archives, and libraries) and individuals. The difficulty of finding and relating information in this kind of heterogenous content provision and data format environment creates an obstacle for end-users of cultural contents, and a challenge to organizations and communities producing the contents.

Portals try to ease these problems by collecting content of various publishers into a single site [50]. Portal types include *service portals* collecting a large set of services together (e.g., Yahoo! and other "start pages"), *community portals* [53] acting as virtual meeting places of communities, and *information portals* [43] acting as hubs of data. Much of the semantic web content will be published using *semantic information portals* [38, 43]. Such portals are based on semantic web standards[1] and machine "understandable" content, i.e., metadata, ontologies, and rules, in order to improve structure, extensibility, customization, usability, and sustainability of traditional portal designs.

Cultural heritage is a promising application domain for semantic portals [6, 25, 48, 57, 4, 3]. They are useful from the end-users' view point in several ways:

- *Global view to heterogeneous, distributed contents*. The contents of different content providers can accessed through one service as a single, seamless, and homogenous repository [25]. Only a single user interface has to be learned.
- *Automatic content aggregation*. Satisfying an end-user's information need often requires *aggregation* of content from several information providers [50, 26], a task suitable for semantic web technologies. For example, when looking for data about an artist, relevant information may be provided by museum collections, libraries, archives, authority records, ontologies, and other sources.

[1] http://www.w3.org/2004/SW/

- *Semantic search*. In traditional portals, search is usually based on free text search (e.g., Google), database queries, and/or a stable classification hierarchy (e.g., Yahoo! and dmoz.org). Semantic content makes it possible to provide the end-user with more "intelligent" facilities based on ontological concepts and structures, such as *semantic search* [10], *semantic autocompletion* [24], and *faceted search* [42, 19, 27, 47, 21].
- *Semantic browsing and recommendations*. Semantic content also facilitates semantic browsing [17] (cf. Chapter 36) and recommendations [58] (cf. Chapter 34). Here semantic associations between search objects can be exposed to the end-user as recommendation links, possibly with explicit explanations.
- *Other intelligent services*. Also other kind of intelligent services can be created based on machine interpretable content, such as knowledge and association discovery [49], personalization [2, 4], and semantic visualizations based on e.g. historical and contemporary maps and time lines [36].

Semantic portals are very attractive from the content publishers viewpoint, too:

- *Distributed content creation.* Portal content is usually created in a centralized fashion by using a content management system (CMS). This approach is costly and not feasible if content is created in a distributed fashion by independent publishers, e.g, by different of museums and other memory organizations. Semantic technologies can be used for harvesting and aggregating distributed heterogenous content (semi-)automatically into global content portals [25].
- *Automated link maintenance.* The problems of maintaining links up-to-date is costly from the portal maintenance viewpoint. In semantic portals, links can be created and maintained automatically based on the metadata and ontologies.
- *Shared content publication channel.* In the cultural domain the publishers usually share the common goal of promoting cultural knowledge in public and among professionals. A semantic portal can provide the participating organizations with a shared, cost-effective publication channel [28].
- *Enriching each other's contents semantically.* Interlinking content between collaborating organizations enriches the contents of everybody "for free".
- *Reusing aggregated content.* The content aggregated into a semantic portal can be reused in different applications and cross-portal systems [59].

A cultural semantic information portal includes the several major components. First, we need a *content model* for representing cultural metadata, ontologies, and rules. Second, a content creation system is needed for creating and harvesting content. Third, the portal publishes semantic services for 1) human end-users as *intelligent user interfaces* and possibly for 2) other portals and applications as *web services*. In the following these components are explained in more detail.

2 Content Models for Semantic Cultural Portals

The semantic web "layer cake model" makes the distinction between a syntactic data level based on XML[2], and semantic levels above it:

- *Metadata level*. The RDF data model[3] (cf. Chapter 3) is used for representing metadata about cultural resources.
- *Ontology level*. The RDF Schema and web ontology language OWL[4] (cf. Chapter 4) are used for representing ontologies [14] (cf. Chapter 19) that describe vocabularies and concepts concerning the real world and our conception of it.
- *Logic level*. Logic rules (Cf. Chapter 5) can be used for deriving new facts and knowledge based on the metadata and ontologies.
- *Trust level*. At the highest conceptual level issues of e.g. trustworthiness of content, copyrights etc. are of concern.

In the following, metadata, ontology, and logic layers are considered from the viewpoint of semantic cultural portals. Issues related to trust on the semantic web in the cultural heritage domain have thus far not been discussed much in the literature.

2.1 Metadata Schemas

Cultural content in museum collections, libraries, and other content repositories is usually described using *metadata schemas* (also called *annotation schemas* or *annotation ontologies*). These templates specify a set of obligatory and optional elements, i.e. properties, by which the metadata for content items should be described. For example, the Dublin Core (DC) Metadata Element Set[5] lists 15 standardized[6] elements, such as *dc:title*, *dc:creator*, and *dc:subject*, with additional elements and element refinements. Encoding guidelines tell how to express the elements in RDF/XML and using HTML/XHTML meta and link elements. Qualifiers, such as encoding schemes, enumerated lists of values, and other processing clues are used to provide more detailed information about a resource. For example, "date" is a DC element that can further be specified as "date published" or "date last modified". The core elements can be extended in an interoperable way by using the "dumb-down" principle. It means that in any use of a qualified DC element, the qualifier may be dropped and the remaining value of the element should still be a term that is useful for discovery, although with less precision.

DC is used as a basis in more detailed cultural metadata schemas, such as the Visual Resource Association's (VRA) Core Categories[7]. Its element set provides a categorical organization for the description of works of visual culture as well as the

---

[2] http://www.w3.org/XML/
[3] http://www.w3.org/RDF/
[4] http://www.w3.org/2004/OWL/
[5] http://dublincore.org/documents/1998/09/dces/
[6] NISO Standard Z39.85-2001 and ISO Standard 15836-2003
[7] http://www.vraweb.org/

images that document them. Most VRA elements are defined as subproperties of corresponding DC elements. An example of an instance of VRA metadata in the CHIP portal [2, 4] is given below in RDF Turtle notation[8]. The schema has properties such as *vra:type* (the type of the art-work as a reference to the VRA vocabulary), *vra:title* (literal title of the art-work), *vra:creator*, *vra:subject*, *vra:culture*, and *vra:material*. Element values with a namespace are references to underlying ontologies.

```
rijks:artefactSK-C-K
  vra:type vra:Work ;
  vra:title "The Night Watch" ;
  vra:date "1642" ;
  vra:creator: 500011051 ;         # Rembrandt
  vra:subject iconclass:45F31 ;    # Call to arms
  vra:culture tgn:7006952 ;        # Amsterdam
  vra:material aat:30015050 .      # Oil paint
```

A metadata schema makes it possible to specify relevant aspects of the search objects, such as the "author", "title", and "subject" of a document, and focus search according to these. Sharing a metadata schema between different content providers facilitates, for example, multi- or metasearch[9]. Here the user types in a query in a metaportal. The query is then distributed to underlying systems and the results are aggregated for the end-user. Protocols such as Z39.52[10] and Search and Retrieve via URL (SRU)[11] of the Library of Congress can be used here. For example, the Australian Museums and Galleries Online[12] and Artefacts Canada[13] are multi-search engines over nation-wide distributed cultural collections. Another approach to creating metaportals is to first harvest the content into a global database, and search the global repository. Protocols such as Open Access Initiative Protocol for Metadata Harvesting (OAI-PMH)[14] can be used for distributed content publishing and harvesting.

Schema definitions tackle the problems of *syntactic* and *semantic interoperability* of content objects. Syntactic interoperability can be obtained by harmonizing encoding conventions (e.g., a date format) and other structural forms for representing data (e.g., an XML schema). Semantic interoperability is obtained by shared conventions for interpreting the syntactic representations, e.g., that the property *dc:subject* describes the subject matter of a document as a set of keywords taken from a thesaurus. Making different metadata schemas semantically interoperable includes two subtasks. First, semantic interoperability of element values has to be addressed using (shared) vocabularies and ontologies, and second, if multiple metadata schemas are involved, interoperability problems between different schema elements has to be solved. In below, these two issues are discussed in more detail.

---

[8] http://www.dajobe.org/2004/01/turtle/
[9] http://en.wikipedia.org/wiki/Metasearch_engine
[10] http://www.cni.org/pub/NISO/docs/Z39.50-brochure/
[11] http://www.loc.gov/standards/sru/
[12] http://www.amonline.net.au/
[13] http://www.chin.gc.ca/
[14] http://www.openarchives.org/

2.2 Vocabularies and Ontologies

Metadata schemas specify data formats but do not tell how to fill the element values in the formats. Additional standards and guidelines are necessary to guide the choice of terms or words (data values) as well as the selection, organization, and formatting of those words (data content). Data value standards have been traditionally specified by constructing controlled vocabularies and thesauri [15, 1]. Examples of cultural thesauri include the Thesaurus for Graphic Materials I (TGM I)[15] for indexing pictorial materials, ICONCLASS[16] for art, Art & Architecture Thesaurus (AAT)[17] for fine art, architecture, decorative arts, archival materials, and material culture, Union List of Artist Names (ULAN)[18], the Thesaurus of Geographic Names (TGN)[19], the Library of Congress Authority Files[20], and the terminologies and standards of the MDA (formerly Museum Documentation Association) [21]. An example of a data content standard is the Cataloging Cultural Objects (CCO) guidelines[22].

Many cultural thesauri have been transformed [56, 55] into SKOS format[23] to be used in cultural semantic portals [48, 57]. However, although a syntactic transformation into SKOS is useful, it is not always enough from a semantic viewpoint. The fundamental problem with traditional thesauri is that its semantic relations have been constructed mainly to help the indexer in finding indexing terms, and understanding the relations needs implicit human knowledge. Unless the meaning of the semantic relations of a thesaurus is made more explicit and accurate for the computer to interpret, the SKOS version is equally confusing to the computer as the original thesaurus, even if semantic web standards are used for representing it.

For example, there are many problems in utilizing the Broader Term (BT) relations of thesauri [30]: 1) BT relations do not necessarily structure the terms into a full-blown hierarchy that would be useful e.g. in faceted search, but into a forest of small subhierarchies. 2) The semantics of the BT relation is ambiguous: it may mean either subclass-of-relation, part-of relation (of different kinds, cf. [13]), or instance-of relation. As a result, the BT relation cannot e.g. be used for property inheritance. 3) The transitivity of the BT relation chains is not guaranteed from the instance-class-relation point of view. If x is an instance of class A whose broader term is B, then it is not necessarily the case that x is an instance of B, although this a basic assumption in RDFS and OWL. For example, assume that x is an instance of “make-up mirror”, whose broader term is “mirror”, and that its broader term is “furniture”. When searching with the concept “furniture” one would expect that instances of furniture are retrieved, but in this case the result would include confusingly make-up mirrors,

---

[15] http://www.loc.gov/rr/print/tgm1/
[16] http://www.iconclass.nl/
[17] http://www.getty.edu/research/conducting_research/vocabularies/aat/
[18] http://www.getty.edu/vow/ULANSearchPage.jsp
[19] http://www.getty.edu/research/conducting_research/vocabularies/tgn/
[20] http://authorities.loc.gov/
[21] http://www.mda.org.uk/stand.htm
[22] http://www.vraweb.org/ccoweb/cco/index.html
[23] http://www.w3.org/2004/02/skos/

too, if transitivity is assumed. A solution to these fundamental problems is to actually refine and reorganize the semantic structures of a thesaurus into a light-weight ontology e.g. along the lines proposed in [31].

Several domain ontologies are used in describing cultural metadata. This raises up the problem of making ontologies mutually interoperable. There are solution approaches for this, such as ontology mapping and alignment [18] (cf. Chapter 26), sharing common foundational logical principles DOLCE[24] (cf, Chapter 16), and using shared horizontal top ontologies, such as the IEEE SUMO[25]. It is likely, that in many cases several identifiers (URIs) will be in use for denoting a single concept even if this is not desirable in general. For example, registries of same geographical locations are maintained at different countries and by different service providers using their own identifiers. In such cases, dereferencing services will be needed to map resource identifiers denoting same concepts with each other.

2.3 Metadata Schema Interoperability

If a portal aggregates cultural contents described using different kind of schemas (e.g, for artifacts, music, maps, books, cultural sites etc.), the schema element structures have to be made interoperable in one way or another, including the element values. If the element structures in the schemas refine each other, then using subproperties and the dumb-down principle of DC applications may be applied. In other cases, the metadata schemas can be made interoperable by transforming them into a shared underlying form.

An approach to this is the CIDOC Conceptual Reference Model (CIDOC CRM) [11] (cf. Chapter 19), an annotation ontology standard[26] developed as an underlying schema into which other metadata schemas in the cultural domain can be transformed for interoperability. This model "provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation"[27]. The framework includes 81 classes, such as *crm:Man-Made Object*, *crm:Place*, and *crm:Time-Span*, and a large set of 132 properties relating the entities with each other, such as *crm:Has Time-Span* and *crm:Is Identified By*.

Another approach to semantic metadata schema interoperability has been developed in the CULTURESAMPO portal [26, 45]. The cultural content types in this system include a wide variety of cultural objects, such as artifacts, paintings, photographs, videos, music, biographies, epics, cultural sites, and historical events. The original metadata from the content providing memory organizations use several schemas, including DC, ULAN, and CIDOC CRM. Content integration is performed by transforming content into a light-weight knowledge base describing the domain world based on events and their thematic roles [52], such as agent, goal, and place. For example, the DC metadata of a painting tells that there has been a painting event

[24] http://www.loa-cnr.it/DOLCE.html
[25] http://suo.ieee.org/
[26] Since 2006 it has been an official ISO standard 21127:2006.
[27] http://cidoc.ics.forth.gr/

with the value of *dc:creator* in the agent role. This event instance can be used for enriching the painter's biography, that is also represented in terms of underlying events, such as the painter "being born" at a certain place—another event that can be derived from the relational embedded meaning of the relation *ulan:birthPlace* used in ULAN. In contrast to CIDOC CRM, the events and thematic role values are based on large shared domain ontologies of tens of thousands of concepts, and only few thematic and other relationships between them. The domain ontologies are used not only for explicating relational meaning of metadata schemas in an interoperable way, but also for making *element values* semantically interoperable, an issue not addressed by the CIDOC CRM standard. The homogenized event-based knowledge can be used, e.g., as a basis for semantic recommendations in the portal [46].

2.4 Logic Rules for Cultural Heritage

A collection of cultural metadata and related ontologies constitute a knowledge base. On the logical level, rules can be used for deriving new facts and knowledge based on the repository, i.e., for explicating the implicit content of the repository, and enriching the content semantically. Some examples illustrating different ways of using rules in semantic cultural portals and systems are given below.

- *Explicating content of metadata schemas*. Many metadata formats contain implicit knowledge embedded e.g. in the relational meaning of the element names. In [45] rule sets for three cultural metadata schemas are presented for explicating such knowledge in terms of events.
- *Enriching semantic content*. Common sense rules may be used for enriching annotations, thus extending the machine's understanding about culture. In [27], for example, family relation rules (and others) we used to explicate implicit family relations, such as "grand father of", between persons in order to link photographs of relatives together while browsing the repository.
- *Semantic recommendations with explanations*. In [25] some 300 rules and associations, such as "doctoral hats are related to academic ceremonies" or "distaffs are related to spinning events", were used to represent simple common sense knowledge and associations between ontological concepts. A semantic recommendation service was then established that, based on additional logical rules, could 1) dynamically find out chained semantic associations between cultural objects based on ontologies and the common sense relations, and 2) at the same time construct literal explanations of why the association would be of interest. In [32] semantic process descriptions of cultural processes, such as traditional farming and fishing, were used as basis for relating cultural objects with each in meaningful ways.
- *Projecting search facets*. In faceted search, rules can be used for constructing facet hierarchies based on ontological structures, such as the subclass-of and part-of-relations. Furthermore, rules can be used to solve problem of projecting search items to facet categories, which may be complicated [58, 29, 27]. From a software engineering viewpoint, using logic rules for projections separates facets from the

annotation ontologies and annotations, which makes it possible to apply the same faceted search engine to knowledge bases based on different kind of ontologies and annotation schemas [39].
- *Association discovery*. Association discovery can be based on rules trying to find paths between resources in a knowledge base [48, 49, 26].

3 Cultural Content Creation

Several kinds of content need to be created for a semantic portal, including ontologies, terminologies, and semantic annotations. Also creating rules for e.g. semantic recommendations can be seen as a form of content to be created. In below, ontology, terminology, and annotation creation are discussed in some more detail.

The core of a semantic cultural heritage portal is typically a set of domain ontologies that are used for annotating cultural contents. Many vocabularies and ontologies, such as AAT, are used for defining *universals*, i.e., general concepts, classes, or types of individuals, such as “chair” (artifact ontology), “wood” (material ontology), “painter” (actor types), or “city” (geographical concepts). In creating ontologies, it is advisable to try to re-use existing ontologies or transform existing thesauri into semantic web formats, as discussed earlier. The ontologies can also be created or enhanced manually using an ontology editor such as Protégé-2000[28].

Another basic type of ontologies, are instance-rich ontologies or registries of individuals. Such ontologies include, for example, geo-ontologies, such as TGN, and actor ontologies (persons and organizations), such as ULAN. This kind of ontologies of individuals are based on a (usually small) ontology of classes (universals), such as “city” or “person”, that is *populated* with individuals from e.g. a database. This kind of instance ontologies can be used for annotating content (e.g., used as *dc:creator* values), but the instances may, at the same time, be used as a content type of its own value in the portal (e.g, a biography).

The terminology used in a portal is typically defined by associating ontological resources with preferable and alternative labels (e.g., using properties *rdfs:label* or *skos:altLabel*). Resource identifiers (URIs) of concepts, used by the machine, refer to concepts that are in principle language independent. However, labels used by humans can be multi-lingual, based on XML markup (e.g., *xml:lang*). This is essential when creating multilingual portals.

The content providers often use different literal terms to refer to same resources when describing metadata in legacy systems. For example, literals “United States” and “US” may be used to refer to the same country. This problem of *synonymy* can be approached by using alternative labels. On the other hand, the same term may be used to refer different concepts, such as river “bank” and financial “bank”. In order to eliminate such *homonymy* in terminology, it is advisable that an ontology uses a unique labeling of terms for concepts. However, this does not solve the problem disambiguating meanings of terms occurring in natural language descriptions.

[28] http://protege.stanford.edu/

Content in memory organizations is usually available as relational legacy databases, whose annotations are literal terms and free text descriptions. Such annotations are often indented for human usage, use various syntactic conventions, are often semantically ambiguous, and may contain syntactic typing errors. When transforming legacy metadata into semantic web formats, a key problem is how map textual descriptions in the metadata with ontological concepts, e.g., how to determine that the string “bank” in a *dc:subject* description of a photograph refers to the concept “river bank” and not “financial bank”. In below, the task of transforming literal element values used in legacy systems into ontological references needed on the semantic web is discussed. The semantic portal MUSEUMFINLAND [25] and its content creation model [28] is used as a concrete example of the more general problem.

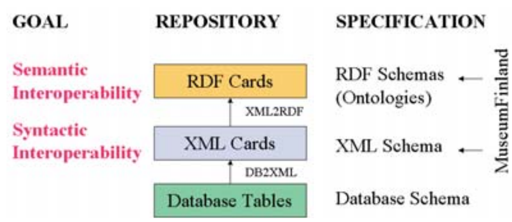

Fig. 1. Transforming legacy museum collection data from database tables into RDF.

Metadata in this system originates from different DC like metadata schemas used in three museums, represented in different kind of database tables using different cataloging database systems. These tables are transformed into an RDF repository in two steps depicted in figure 1: First, the heterogenous relational tables in each museum are harmonized by transforming them into an XML metadata schema format that is shared by the co-operating content providers. This transformation ensures syntactic interoperability among all data sources, and partial semantic interoperability in terms the meaning of the metadata schema elements, since a single element set is used. Second, semantic interoperability between metadata sources is obtained by transforming the XML descriptions into the final RDF metadata schema format used by the portal. During this XML-to-RDF transformation the essential task is to move from term space into concept space by changing literal terms, used at the XML level as element values, into corresponding concept URIs referring to seven underlying domain ontologies (e.g., Artifacts, Places, etc.). The URIs created in this phase connect metadata RDF with domain ontology RDF, resulting into a single large semantic RDF triple store used for querying and as a basis for logical reasoning.

A major problem in the RDF transformation above is how to disambiguate the meanings of homonyms (e.g., “bank”) that may occur as keywords, free indexing terms, or in free text descriptions in different element values. Several methods can be applied here. For example, the type of the metadata element in which a homonymous expression is used, can often be used for semantic disambiguation effectively [28]. However, when dealing with the *dc:subject* element (or similar ones) that can have values taken from different vocabularies, such contextual disambiguating information is not available, and human decision help is more often needed.

Another practical problem is spelling errors in metadata, and the variance of synonyms and correct syntactic encoding practices used at different organizations at different times, in different languages, and even by different catalogers. For example, the name of Ivan Ayvazovsky (Russian painter, 1817–1900) has 13 different labels in ULAN (Ajvazovskij, Aivazovski, Aiwasoffski etc.), and the first, middle, and last names can be ordered and shortened in many different ways.

Still another problem of transforming literals into URIs is complicated free text descriptions that may be used as element values, such as the material description "cow leather with decorations". Free text descriptions in metadata are in general difficult to search for due their syntactic variance, and for the same reason, difficult to transform into URI references automatically. The problem can be approached by using in indexing controlled vocabularies or ontologies. However, even then the problem remains when dealing with *free indexing terms* . These terms are, by definition, legal keywords of a thesaurus that are not listed as entries. For example, plant and animal types as well as person and location names can be used as free indexing terms. When encountering such a term, it cannot usually be associated with the underlying ontologies without human help.

In a distributed content creation environment, free indexing concepts pose a challenge for ontology maintenance, too. In many cases new concepts should to populated into the ontologies and be shared, too. For example, when a painting of a new, formerly unknown artist is cataloged in a museum, the other catalogers and organizations should be made aware of her/him in order to prevent creation of multiple identifiers for the artist and later confusion of identities.

A solution approach to this is to connect annotation creation tools to centrally maintained *ontology library services* that provide the clients with up-to-date information about the vocabulary resources available, and facilitates creation and sharing of new resources collaboratively. An implementation of such a service is the ONKI Ontology Server[29] [59, 31] that can be used for creating mash-up annotation applications in a way analogous to creating Google Maps mash-ups.

Sharing unique URIs for concepts is preferable on the semantic web, but in practice there will be multiple URIs referring to a single resource. Creation of multiple identifiers for free indexing concepts cannot be eliminated totally in practice, and multiple identifiers will be created purposefully, too. For example, different countries are likely to use their own identifiers already in use for their geographical locations. Global dereferencing services will be needed in the future telling, e.g., that the concept of "London" in U.K. refers to the same thing as "Londres" in France.

After creating semantically interoperable RDF metadata, content harvesting and aggregation can be done either 1) off-line before starting the portal or 2) on-line dynamically when answering end-user queries. The on-line approach is more dynamic. However, from the viewpoint of creating intelligent end-user services, the off-line approach seems more promising: 1) By creating a global knowledge base first off-line,

[29] http://www.seco.tkk.fi/services/onki/

reasoning can be easily done at the global scale across local contents, which facilitates e.g. generation of recommendation links between the content of different content providers. 2) Knowledge can be compiled and critical reasoning tasks performed off-line beforehand for faster response times. For example, the *rdf:type* instance-class-relations can be explicated as RDF-triples based on the transitive closures of the subclass-of hierarchies. 3) The portal is independent of the content providers possibly unreliable web services when running the system.

4 Semantic Portal Services

The goal of semantic information portals for cultural heritage is to provide the end-user with intelligent services for finding and learning the right information based on her own preferences and the context of using the system. In the following, some possibilities of providing the end-users with intelligent services using semantically annotated metadata are shortly explored.

4.1 Semantic Search

In information retrieval [5] search is usually based on finding occurrences of words in documents. On the semantic web, search can be based on finding the concepts related to the documents at the metadata and ontology levels, in addition to the actual text or other features of the data. With concept-based methods document meanings and queries can be specified more accurately which usually leads to better recall and precision, especially if both the query and the underlying content descriptions are concept-based.

With non-textual cultural documents, such as paintings, photographs, and videos, metadata-based search techniques are a must in practice, although also content-based information retrieval methods [44] (CBIR) and multimedia information retrieval (MIR) [37] can be used as complementary techniques. Here the idea is to utilize actual document features (at the data level), such as color, texture, and shape in images, as a basis for information retrieval. For example, an image of Abraham Lincoln could be used as a query for finding other pictures of him, or a piece of music could be searched for by humming it. Bridging the "semantic gap" between low level image and multimedia features and semantic annotations is an important but challenging research theme [23]. Still another approach to do semantic search is to analyze and build search on the content using linguistic and/or statistical methods, without using annotated semantic metadata [8].

A key problem of semantic search is mapping the literal search words, used by humans, to underlying ontological concepts, used by the computer. Depending on the application, only queries expressed by terms that are relevant to the domain and content available are meaningful, other queries result in frustrating "no hits" answers. A way solve the problem is to provide the end-user with a vocabulary as a subject heading category tree, a facet , as in Yahoo! and dmoz.org. By selecting a category, related documents are retrieved. Faceted search [42, 19, 27, 47, 21] is a natural generalization of this, where the user can make several simultaneous selections from *multiple* orthogonal facets. They are exposed to the end-user in order to 1) provide her with the right query vocabulary, and 2) for presenting the repository contents and search results and the amounts of hits in facet categories. The result set can be presented to the end-user according to the facet hierarchies for better readability. This is in contrast with traditional search where results are typically presented as a list of decreasing relevance. The number of hits resulting from a category selection is always shown to the user before the selection. This eliminates queries leading to "no hits" dead-ends, and guides the user in making next constraining selections on the facets.

Faceted search has been integrated with the idea of ontologies and the semantic web [27]. The facets can be constructed algorithmically from a set of underlying ontologies that are used as the basis for annotating search items. Furthermore, the mapping of search items onto search facets can be defined using logic rules. This facilitated more intelligent semantic search of indirectly related items. A method for ranking the search results in faceted search based on fuzzy logic is presented in [22], and [54] presents a card sorting approach for specifying and using end-user facets independently from the indexing ontologies.

The faceted search paradigm is based on *facet analysis* [41], a classification scheme introduced in information sciences by S. R. Ranganathan already in the 1930's. The idea of faceted search has been invented and developed independently by several research groups, and is also called view-based search [42] and dynamic taxonomies [47]. Several semantic cultural heritage portals make use of faceted search, such as [25, 21]. However, faceted search is not a panacea for all information retrieval tasks. Google-like keyword search interface is usually preferred if the user is capable of expressing her information need terms of accurate keywords [12].

4.2 Semantic Autocompletion

Keyword search can be integrated with semantic search by extending search to the labels of ontological resources or facet categories. For example, in [25] keyword search is integrated with faceted search in the following way: First, search keywords are matched against category names in the facets in addition to text fields in the metadata. The result set of hits is shown containing all objects in any of the categories matched in addition to all objects whose metadata directly contains the keyword. The hits are grouped by the categories found. Second, a new dynamic facet is created in the user interface for disambiguating the different possible ontological interpretations and roles of the keyword. This facet contains all facet categories whose name (or other property values) matches the keyword. They tell the end-user the different interpretations and roles of the keyword. For example, the keyword "Nokia" matches in the portal with the mobile phone company resource in the "Manufacturer" facet role, with the city of Nokia in the facet roles "Place of manufacturing" and "Place of usage", and with some other resources that have the string in their name. By selecting one of the interpretations, the user is able to disambiguate the meanings easily and constrain search further.

The idea of searching ontologies and facet categories for disambiguating intended meanings and roles has been generalized into the notion of *semantic autocompletion* [24]. The idea here is to generalize traditional text autocompletion by trying to guess, based on ontologies and reasoning, the search concept the user is trying to formulate after each input character in an input field. For example, the user may type in the query in French and the semantic autocompletion service finds the possible intended search concepts in English after each input character.

Autocompletion has become a popular way to find meaningful keywords in large search vocabularies after Google Suggest[30] was released. The idea is applied in several semantic cultural portals, such as [33, 26, 48, 57].

4.3 Semantic Browsing and Recommending

In addition to semantic search, semantic content facilitates *semantic browsing*. Faceted search is already a kind of combination of searching and browsing because search is based on selecting links on facets. However, in semantic browsing the general idea is not to constrain the results set but rather to expand it by trying find objects of potential interest outside of the hit list. The idea is to support browsing documents through associative links that are created based on the underlying metadata and ontologies, not on hardwired anchor links encoded by humans in HTML pages.

A simple form of a semantic browser are RDF browsers and tabulators [7]. Their underlying idea has been explicated as the "linked data"[31] principle proposing that when an RDF resource (URI) is rendered in a browser, the attached RDF links to related resources should be shown. When one of these links is selected, the corresponding a new resource is rendered, and so on.

A more developed related idea is *recommendation systems* [9]. Here the logic of selecting and recommending of related resources can be based on also other principles than the underlying RDF graph. For example, collaborative filtering [20] is based on browsing statistics of other users. Also logic rules on top of an RDF knowledge base can be used for creating semantic recommendation links [58] and, at the same time, *explanations* telling the end-user why the recommendation link was selected in this context. In [2, 4] explanations for recommended art works can be obtained based on a user profile of interest and features of the artworks. In [32] ontological models of narrative stories and processes in the society, such as fishing or slash farming, were used as a basis for creating recommendation links between cultural resources. Still another approach to create recommendation links with explanations is to use similarity measures of event-based annotations [46].

[30] http://www.google.com/webhp?complete=1&hl=en
[31] http://www.w3.org/DesignIssues/LinkedData.html

4.4 Relational Search

Semantic recommending is related to *relational search*, where the idea is to try to search and discover serendipitous semantic associations between different content items [49, 48, 26]. The idea is to make it possible for the end-user to formulate queries such as ”How is X related to Y” by selecting the end-point resources, and the search result is a set of semantic connection paths between X and Y. For example, in figure 2 the user has specified two historical persons, the Finnish artist Akseli Gallen-Kallela (1865–1931) and the French emperor Napoleon I (1769–1821) in the CULTURESAMPO portal [26]. The underlying knowledge base contains an ontologized version of the ULAN vocabulary in RDF with over 100,000 persons and organizations, and semantic autocompletion is used for finding the right query resources. The system has discovered an association chain between the persons based on “patronOf”, “teacherOf”, “knows”, and “studentOf” properties.

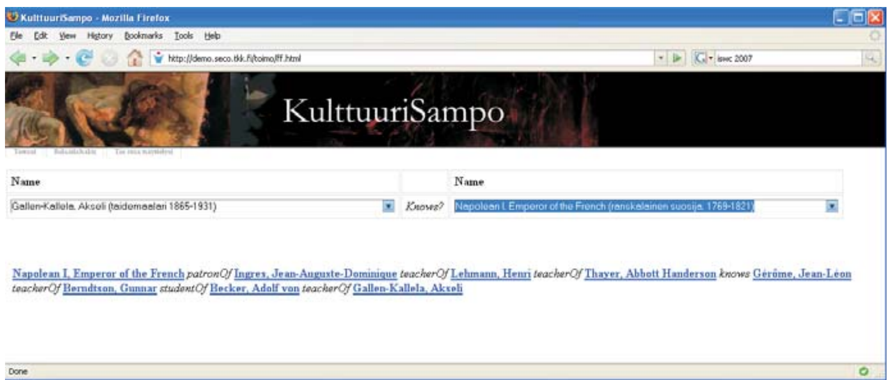

Fig. 2. An example of relational search in [26] using the ULAN vocabulary and database.

4.5 Personalization and Context Awareness

In many occasions the functioning of a semantic portal should not be static but adapt dynamically according to the 1) personal interests of the end-user and 2) the context of usage, such as time and location [51]. Visitors in semantic cultural portals, like in physical museums, are usually not interested in everything found in the underlying collections, and would like to get information at different levels of detail. An important aspect of a semantic cultural portal is then adaptation of the portal to different personal information needs and interests. An example of a personalized cultural semantic portal is [2, 4], where user profiling and personalization is based on metadata obtained by asking the users about her interests by rating pieces of artworks.

An example of location-based adaptability is the mobile phone user interface of [25]. By pushing a special button on the interface, collection artifacts either manufactured or used nearby can be retrieved based on a geolocation service proving the coordinate information of the phone. It can be envisioned that this kind of location-based and navigational services will be available in future cultural portals based on phones supporting GPS positioning and radio-frequency identification (RFID) tags.

Also time is an important parameter for contextualizing portal services. For example, recommending the end-user to visit a site in the nature during winter may not be wise due to snow, or direct her to a museum when it happens to be closed.

4.6 Visualization and Mash-ups

Visualization is an important aspect of the semantic web dealing with semantically complicated and interlinked contents [16]. In the cultural heritage domain, maps, time lines, and methods for visualizing complicated and large semantic networks are of special interest.

Maps are useful in both searching content and in visualizing the results. A widely used approach to using maps in portals is to use mash-up map services. For example, GeoMuseoSuomi[32] [36] is a mash-up combining Google Maps[33] and a semantic cultural portal [25]. The map interface is used for showing the places of the underlying location ontology on the map as interactive buttons (e.g., cities, villages, etc.). By selecting one of them, a query is executed by which all museum collection items manufactured or used in the selected place are retrieved. At the same time, additional search links to seven different traditional portals are shown. For example, by selecting the Wikipedia link, an article about the location (if available) is opened.

In the cultural heritage domain, historical maps are of interest of their own. For example, they depict old place names and borders not available anymore in contemporary maps. An approach to visualize historical changes is developed in the Temp-O-Map system [34, 36] that makes it possible to lay old maps semi-transparently on top of the contemporary maps and satellite images of Google Maps. To demonstrate the idea, the Karelia region of Finland was selected as a test case. This region was annexed to Soviet-Union as a result of the Second World War, after which most old Finnish place names in the region were changed into Russian ones making it difficult to the end-user to bridge the sematic gap between old and new names and locations. The system is connected into an ontology modeling over 1000 regional changes of Finnish municipalities in 1860–2007. Historical municipalities of different time periods are available as facets for finding historical places on the maps. By selecting a category, the tool focuses the map view to the center point of the region [35].

Another important dimension for visualizing cultural content is time. A standard approach for temporal visualization is to project search objects on a time line, as in [48, 26]. A generic mash-up tool for creating time lines is the Simile time line[34]. A time line can be used both for querying and for visualizing search results.

[32] http://users.tkk.fi/ tomik/geo/karttahaku.html.
[33] http://maps.google.com/
[34] http://simile.mit.edu/timeline/

4.7 Cross-portal Re-use of Content

Portal contents can be re-used in other web applications and portals due to semantic web standards. Re-using semantic content in this way is a kind of generalization of the idea of "multi-channel publication" of XML, where a single syntactic structure can be rendered in different ways. In a similar vein, semantic metadata can be re-used without modifying it through *multi-application publication*.

One possibility to facilitate cross-portal re-use is to merge triple stores, and provide services to end-users based on the extended knowledge base. For example, the learning object video portal [33] is able to provide recommendation links to the cultural museum collection portal [25] in this way. Another way of re-using content is to keep the portals separate and publish their functionalities as web services to be used by other semantic portals [59]. Both traditional web services or light-weight mash-ups based on the REST principle can be used. Here portal functionalities can be used in other portals on the HTML user interface level with just a pair of additional Javascript code added on the HTML level. This approach is related to the idea of using Google AdSense[35] advertisements, but generalized on a semantic level and used for publishing portal services. For example, there is a semantic widget for re-using the semantic search functionality and contents of the portal [25] in external web pages [40]. If a page, for instance, contains information about skating, then the widget can query and show dynamically, using AJAX, images and semantic links to skates and related objects in the museum collection portal.

5 Conclusions

Cultural heritage provides a semantically rich application domain in which useful vocabularies and collection contents are available, and where the organizations are eager to make their content publicly accessible. A major application type in the area has been semantic portals, often aggregating content from different collections, thus providing cultural organizations with a shared cost-effective publication channel and the possibility of enriching collaboratively the contents of each other's collections. For the end-user, new kinds of intelligent semantic services and ways of visualizing content can be provided. We envision that in the near future ever larger cultural semantic portals crossing geographical, cultural, and linguistic barriers of content providers at different countries will be developed. Also more systems for enriching the collections by end-user created content and tagging in the spirit of Web 2.0 will be seen, such as Steve Museum[36] and Powerhouse Museum[37].

A major practical hinder for publishing cultural content on the semantic web is that current legacy cataloging system do not support creation of ontology-based annotations. If semantic annotations cannot be created in memory organization when

[35] http://www.google.com/adsense/
[36] http://www.steve.com/
[37] http://www.powerhousemuseum.com/collection/database/

cataloging content, then costly manual work is needed when transforming and disambiguating literal legacy metadata into ontological references in semantic portals. A solution approach to this fundamental problem is to provide ontologies as publicly available ontology services, and to re-use them—as well as semantically annotated portal contents—as ready-to-use functionalities (widgets) in legacy systems using mash-up techniques [59, 31].

Acknowledgements

This work was supported by the national semantic web ontology project FinnONTO[38] 2003-2007 in Finland, funded by the National Funding Agency for Technology and Innovation (Tekes) and a consortium of 36 public organizations and companies.

References

1. J. Aitchison, A. Gilchrist, and D. Bawden. *Thesaurus construction and use: a practical manual*. Europa Publications, London, 2000.
2. L. Aroyo, R. Brussee, L. Rutledge, P. Gorgels, N. Stash, and Y. Wang. Personalized museum experience: The Rijksmuseum use case. In J. Trant and D. Bearman, editors, *Museums and the Web 2007: Proceedings*, pages 137–144, 2007.
3. L. Aroyo, E. Hyvönen, and J. van Ossenbruggen, editors. *Cultural Heritage on the Semantic Web. Workshop Proceedings. The 6th International Semantic Web Confereence and the 2nd Asian Semantic Web Conference, Busan, Korea.* 2007.
4. L. Aroyo, N. Stash, Y. Wang, P. Gorgels, and L. Rutledge. CHIP demonstrator: Semantics-driven recommendations and museum tour generation. In *Proceedings of ISWC 2007 + ASWC 2007, Busan, Korea*, pages 879–886. Springer–Verlag, 2007.
5. R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley, 1999.
6. V. R. Benjamins, J. Contreras, M. Blázquez, J.M. Dodero, A. Garcia, E. Navas, F. Hernandez, and C. Wert. Cultural heritage and the semantic web. In *The Semantic Web: Research and Applications*, pages 433–444. Springer–Verlag, 2004.
7. T. Berners-Lee, Y. Chen, L. Chilton, D. Connolly, J. Hollenbach R. Dhanaraj, A. Lerer, and D. Sheets. Tabulator: Exploring and analyzing linked data on the semantic web. In *The 3rd International Semantic Web User Interaction Workshop (SWUI 2006)*, Nov 2006.
8. W. Buntine, K. Valtonen, and M. Taylor. The ALVIS document model for a semantic search engine. In *The 2nd Annual European Semantic Web Conference, Demos and Posters, Heraklion, Crete*, May 2005.
9. R. Burke. Knowledge-based recommender systems. In A. Kent, editor, *Encyclopedia of Library and Information Systems*, volume 69. Marcel Dekker, New York, 2000.
10. S. Decker, M. Erdmann, D. Fensel, and R. Studer. Ontobroker: Ontology based access to distributed and semi-structured unformation. In R. Meersman et al., editor, *DS-8: Semantic Issues in Multimedia Systems*, pages 351–369. Kluwer, New York, 1999.
11. M. Doerr. The CIDOC CRM—an ontological approach to semantic interoperability of metadata. *AI Magazine*, 24(3):75–92, 2003.

[38] http://www.seco.tkk.fi/projects/finnonto/

12. J. English, M. Hearst, R. Sinha, K. Swearingen, and K.-P. Lee. Flexible search and navigation using faceted metadata. Technical report, University of Berkeley, School of Information Management and Systems, 2003.
13. C. Fellbaum, editor. *WordNet. An electronic lexical database.* The MIT Press, Cambridge, Massachusetts, 2001.
14. D. Fensel. *Ontologies: Silver bullet for knowledge management and electronic commerce (2nd Edition)*. Springer-Verlag, 2004.
15. D. J. Foskett. Thesaurus. In *Encyclopaedia of Library and Information Science*, volume 30, pages 416–462. Marcel Dekker, New York, 1980.
16. V. Geroimenko and C. Chen, editors. *Visualizing the Semantic Web: XML-based Internet and Information Visualization*. Springer–Verlag, 2002.
17. C. Goble, S. Bechhofer, L. Carr, D. De Roure, and W. Hall. Conceptual open hypermedia = the semantic web? In *Proceedings of the WWW2001, Semantic Web Workshop, Hongkong*, 2001.
18. A. Hameed, A. Preese, and D. Sleeman. Ontology reconciliation. In S. Staab and R. Studer, editors, *Handbook on ontologies*, pages 231–250. Springer–Verlag, 2004.
19. M. Hearst, A. Elliott, J. English, R. Sinha, K. Swearingen, and K.-P. Lee. Finding the flow in web site search. *CACM*, 45(9):42–49, 2002.
20. J. H. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *Computer Supported Cooperative Work*, pages 241–250. ACM, 2000.
21. M. Hildebrand, J. van Ossenbruggen, and L. Hardman. /facet: A browser for heterogeneous semantic web repositories. In *Proceedings of the 5th International Semantic Web Conference (ISWC 2006)*. Springer–Verlag, Nov 2006.
22. M. Holi and E. Hyvönen. Fuzzy view-based semantic search. In *Proceedings of the 1st Asian Semantic Web Conference (ASWC2006), Beijing, China*. Springer–Verlag, 2006.
23. L. Hollink. *Semantic annotation for retrieval of visual resources*. PhD thesis, Free Univerisity of Amsterdam, 2006. SIKS Dissertation Series, No. 2006-24.
24. E. Hyvönen and E. Mäkelä. Semantic autocompletion. In *Proceedings of the first Asia Semantic Web Conference (ASWC 2006), Beijing*. Springer–Verlag, 2006.
25. E. Hyvönen, E. Mäkela, M. Salminen, A. Valo, K. Viljanen, S. Saarela, M. Junnila, and S. Kettula. MuseumFinland—Finnish museums on the semantic web. *Journal of Web Semantics*, 3(2):224–241, 2005.
26. E. Hyvönen, T. Ruotsalo, T. Häggström, M. Salminen, M. Junnila, M. Virkkil, M. Haaramo, T. Kauppinen, E. Mäkelä, and K. Viljanen. CultureSampo—Finnish culture on the semantic web. The vision and first results. In *Semantic Web at Work—Proceedings of STeP 2006*. Finnish AI Society, Espoo, Finland, 2006. To appear in: Klaus Robering (Ed.), Information Technology for the Virtual Museum. LIT Verlag, 2008.
27. E. Hyvönen, S. Saarela, and K. Viljanen. Application of ontology techniques to view-based semantic search and browsing. In *The Semantic Web: Research and Applications. Proceedings of the First European Semantic Web Symposium (ESWS 2004)*. Springer–Verlag, 2004.
28. E. Hyvönen, M. Salminen, S. Kettula, and M. Junnila. A content creation process for the Semantic Web, May 2004. Proceeding of OntoLex 2004: Ontologies and Lexical Resources in Distributed Environments, Lisbon, Portugal.
29. E. Hyvönen, A. Valo, K. Viljanen, and M. Holi. A logic-based semantic web HTML generator—a poor man's publishing approach. In *Proceedings of WWW2004, New York, Alternate Track Papers and Posters*, May 2004.
30. E. Hyvönen, K. Viljanen, E. Mäkelä, T. Kauppinen, T. Ruotsalo, O. Valkeapää, K. Seppälä, O. Suominen, O. Alm, R. Lindroos, T. Känsälä, R. Henriksson, M. Frosterus,

J. Tuominen, R. Sinkkilä, and J. Kurki. Elements of a national semantic web infrastructure—case study Finland on the semantic web (invited paper). In *Proceedings of the First International Semantic Computing Conference (IEEE ICSC 2007), Irvine, California*, Sept 2007. IEEE Press.

31. E. Hyvönen, K. Viljanen, J. Tuominen, and K. Seppälä. Building a national semantic web ontology and ontology service infrastructure—the FinnONTO approach. In *Proceeedings of the 5th European Semantic Web Conference (ESWC 2008)*. Springer–Verlag, 2008.
32. M. Junnila, E. Hyvönen, and M. Salminen. Describing and linking cultural semantic content by using situations and actions. In *Semantic Web at Work—Proceedings of STeP 2006*. Finnish AI Society, Espoo, Finland, Nov 2006. To appear in: Klaus Robering (Ed.), Information Technology for the Virtual Museum. LIT Verlag, 2008.
33. T. Känsälä and E. Hyvönen. A semantic view-based portal utilizing Learning Object Metadata, August 2006. 1st Asian Semantic Web Conference (ASWC2006), Semantic Web Applications and Tools Workshop.
34. T. Kauppinen, C. Deichstetter, and E. Hyvönen. Temp-O-Map: Ontology-based search and visualization of spatio-temporal maps. In *Demo track at the European Semantic Web Conference ESWC 2007, Innsbruck, Austria*, June 4–5 2007. http://www.eswc2007.org/demonstrations.cfm.
35. T. Kauppinen, R. Henriksson, R. Sinkkilä, R. Lindroos, J. Väätäinen, and E. Hyvönen. *Ontology-based Annotation and Search with Spatiotemporal Locations*. Helsinki University of Technology, 2008. Submitted paper, http://www.seco.tkk.fi/publications/.
36. T. Kauppinen, R. Henriksson, J. Väätäinen, C. Deichstetter, and E. Hyvönen. Ontology-based modeling and visualization of cultural spatio-temporal knowledge. In *Semantic Web at Work—Proceedings of STeP 2006*. Finnish AI Society, Espoo, Finland, Nov 2006.
37. Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. Content-based multimedia information retrieval: state of the art and challenges. *ACM Transactions on Multimedia computing, communications, and applications*, pages 1–19, Feb 2006.
38. A. Maedche, S. Staab, N. Stojanovic, R. Struder, and Y. Sure. Semantic portal—the SEAL approach. Technical report, Institute AIFB, University of Karlsruhe, Germany, 2001.
39. E. Mäkelä, E. Hyvönen, and S. Saarela. Ontogator—a semantic view-based search engine service for web applications. In *Proceedings of the 5th International Semantic Web Conference (ISWC 2006)*, Nov 2006.
40. E. Mäkelä, K. Viljanen, O. Alm, J. Tuominen, O. Valkeapää, T. Kauppinen, J. Kurki, R. Sinkkilä, T. Känsälä, R. Lindroos, O. Suominen, T. Ruotsalo, and E. Hyvönen. In L. Nixon, R. Cuel, and C. Bergamini, editors, *First Industrial Results of Semantic Technologies, proceedings, co-located with ISWC 2007 + ASWC 2007, Busan, Korea*, 2007. CEUR Workshop Proceedings, Vol 293.
41. A. Maple. Faceted access: A review of the literature. Technical report, Working Group on Faceted Access to Music, Music Library Association, 1995.
42. A. S. Pollitt. The key role of classification and indexing in view-based searching. Technical report, University of Huddersfield, UK, 1998. http://www.ifla.org/IV/ifla63/63polst.pdf.
43. D. Reynolds, P. Shabajee, and S. Cayzer. Semantic Information Portals. In *Proceedings of the 13th International World Wide Web Conference on Alternate track papers & posters*, New York, NY, USA, May 2004. ACM Press.
44. Y. Rui, T. Huang, and S. Chang. Image retrieval: current techniques, promising directions and open issues. *Journal of Visual Communication and Image Representation*, 10(4):39–62, April 1999.

45. T. Ruotsalo and E. Hyvönen. An event-based method for making heterogeneous metadata schemas and annotations semantically interoperable. In *Proceedings of ISWC 2007 + ASWC 2007, Busan, Korea*, pages 409–422. Springer–Verlag, 2007.
46. T. Ruotsalo and E. Hyvönen. A method for determining ontology-based semantic relevance. In *Proceedings of the International Conference on Database and Expert Systems Applications DEXA 2007, Regensburg, Germany*. Springer–Verlag, 2007.
47. G. M. Sacco. Dynamic taxonomies: guided interactive diagnostic assistance. In N. Wickramasinghe, editor, *Encyclopedia of Healthcare Information Systems*. Idea Group, 2005.
48. G. Schreiber, A. Amin, M. van Assem, V. de Boer, L. Hardman, M. Hildebrand, L. Hollink, Z. Huang, J. van Kersen, M. de Niet, B. Omelayenko, J. van Ossenbruggen, R. Siebes, J. Taekema, J. Wielemaker, and B. Wielinga. MultimediaN E-Culture demonstrator. In *Proceedings of the Fifth International Semantic Web Conference (ISWC 2006)*, pages 951–958. Springer–Verlag, November 2006.
49. A. Sheth, B. Aleman-Meza, I. B. Arpinar, C. Bertram, Y. Warke, C. Ramakrishnan, C. Halaschek, K. Anyanwu, D. Avant, F. S. Arpinar, and K. Kochut. Semantic association identification and knowledge discovery for national security applications. *Journal of Database Management on Database Technology*, 16(1):33–53, Jan–March 2005.
50. T. Sidoroff and E. Hyvönen. Semantic e-goverment portals—a case study. In *Proceedings of the ISWC-2005 Workshop Semantic Web Case Studies and Best Practices for eBusiness SWCASE05*, Nov 2005.
51. S. Sirmakessis, editor. *Adaptive and Personalized Semantic Web*. Springer–Verlag, 2006.
52. J. Sowa. *Knowledge Representation. Logical, Philosophical, and Computational Foundations*. Brooks/Cole, 2000.
53. S. Staab, J. Angele, S. Decker, M. Erdmann, A. Hotho, A. Maedche, H.-P. Schnurr, R. Studer, and Y. Sure. Semantic Community Web Portals. In *Proceedings of the 9th International World Wide Web Conference*, Amsterdam, The Netherlands, May 2000. Elsevier.
54. O. Suominen, K. Viljanen, and E. Hyvönen. User-centric faceted search for semantic portals. In *Proceedings of the 4th European Semantic Web Conference (ESWC 2007), Innsbruck, Austria*, pages 356–370. Springer–Verlag, 2007.
55. M. van Assem, V. Malaise, A. Miles, and G. Schreiber. A method to convert thesauri to SKOS. In *Proceedings of the Third European Semantic Web Conference (ESWC 2006)*. Springer–Verlag, 2006.
56. M. van Assem, M. R. Menken, G. Schreiber, J. Wielemaker, and B. Wielinga. A method for converting thesauri to RDF/OWL. In *Proceedings of 3rd International Semantic Web Conference (ISWC 2004), Hiroshima, Japan*. Springer–Verlag, 2004.
57. J. van Ossenbruggen, A. Amin, L. Hardman, M. Hildebrand, M. van Assem, B. Omelayenko, G. Schreiber, A. Tordai, V. de Boer, B. Wielinga, J. Wielemaker, M. de Niet, J. Taekema, M.-F. van Orsouw, and A. Teesing. Searching and Annotating Virtual Heritage Collections with Semantic-Web Techniques. In *Proceedings of Museums and the Web 2007*, San Francisco, California, March 2007. Archives & Museum Informatics.
58. K. Viljanen, T. Känsälä, E. Hyvönen, and E. Mäkelä. ONTODELLA—a projection and linking service for semantic web applications. In *Proceedings of the 17th International Conference on Database and Expert Systems Applications (DEXA 2006), Krakow, Poland*. IEEE, September 4–8 2006.
59. K. Viljanen, J. Tuominen, T. Känsälä, and E. Hyvönen. Distributed semantic content creation and publication for cultural heritage legacy systems. In *Proceedings of the 2008 IEEE International Conference on Distibuted Human-Machine Systems, Athens, Greece*. IEEE Press, 2008.

Index

Author Index