International Journal of Semantic Computing
Vol. 6, No. 3 (2012) 315–327

SEMANTIC-BASED RETRIEVAL OF CULTURAL HERITAGE MULTIMEDIA OBJECTS

KAI STALMANN* and DENNIS WEGENER†

Fraunhofer IAIS, Schloss Birlinghoven
53754 Sankt Augustin, Germany
**kai@stalmann.org*
†*dennis.wegener@iais.fraunhofer.de*
www.iais.fraunhofer.de

MARTIN DOERR

Institute of Computer Science Foundation for Research and Technology - Hellas, N. Plastira 100, Vassilika Vouton
GR-700 13 Heraklion, Crete, Greece
‡*martin@ics.forth.gr*
www.ics.forth.gr

HERMANN JOSEF HILL

Ellingshohl 39, 56076 Koblenz, Germany
hill@hill4gis.de

NATALJA FRIESEN

Fraunhofer IAIS, Schloss Birlinghoven
53754 Sankt Augustin, Germany
natalja.friesen@iais.fraunhofer.de
www.iais.fraunhofer.de

Today's search interfaces typically offer keyword searches and facets for the retrieval of cultural heritage multimedia objects. Facets, however, are usually based on a static set of metadata fields. This set is often called an indexing profile. Graph-based repositories based on predicates about resources allow for more precise semantics. They offer stronger support for retrieval, and they can be adopted to almost any metadata format. Technically, those predicates may be serialized as RDF triples, but handling a huge amount of objects with numerous predicates puts an unpredictable load on the query engine. In this paper, we present an approach on analysing transition paths in the RDF triples at ingest time and using the results to create facets in the search index.

Keywords: Cultural heritage objects; digital library; semantics; ontologies; rdf; triples; linked data; repository.

1. Introduction

Cultural heritage multimedia objects may be digital born, but often have physical counterparts held in a museum somewhere on the globe, e.g., a book copy that is not

accessible locally or information possibly stored in an archive you do not even know it exists. Such multimedia objects of the cultural heritage domain may be text documents, sound files, video streams, 3D objects, but also datasets that need specialized software to make them intelligible for humans. The German digital library project "Deutsche Digitale Bibliothek" (DDB) [1] aims at making the full range of German cultural heritage objects available for end users and experts alike via a search interface. In order to do this, content from different cultural domains including museums, libraries, archives, scientific institutions, and others has to be integrated.

Business applications can usually rely on structured information and clearly modelled data, but it is widely understood that this is usually not the case with cultural heritage data. Metadata of cultural heritage objects differ in size (from less than 1 KB to over 100 MB for one item), format and semantic richness. Despite approaches towards standardizing metadata and metadata usage, the same metadata format can still be used in many different ways. Therefore, harmonizing heterogeneous data requires a model capable of covering all formats and schemas that are likely to be ingested into a repository. This model should preferably be semantically rich enough to preserve all information that might be relevant for users in machine-readable form — the search engine should be able to distinguish if, e.g., a book *is about* or *was created in* a certain epoch. Since each bit of information in the original metadata could be relevant for this, the target model must be at least as expressive as the sum of all original formats. The most universal and flexible approach for this is a graph-based representation of the semantics expressed in the original metadata.

If a repository built for holding cultural heritage multimedia objects makes use of more advanced technical approaches like modelling metadata as an ontology and storing knowledge in form of triples, the problem remains how this knowledge can be brought to and used by an end user. While Google-like search engines would not understand the 'semantic' question a user might want to pose, it is on the other hand the end user that has no command of SPARQL, which actually would allow for posing these questions. Data Repositories like DBPedia [2] could answer such questions that reach out for hidden treasures. But, and this raises a third kind of problems, resolving the semantically precise retrieval costs too much computation time when the query is applied to huge datasets.

Many commercial web applications have proven that search interfaces — at least slightly more sophisticated than what Google currently can offer — are feasible when using facets. In this paper we propose a method to accelerate the resolving of facet names and contents by pre-computing them at ingest time, i.e., when the information from a new metadata object is added to the already existing web of knowledge.

The remainder of the paper is structured as follows: In Sec. 2 we introduce the context of our work and discuss related work. Section 3 presents our approach to the problem of semantic-based retrieval in general and gives details on indexing, search and navigation in the object graph. Section 4 shows first benchmark results of our system. In Sec. 5 we draw our conclusions and present future work.

2. Background on the DDB and Related Work

The DDB project [1], which motivated the subjects discussed here, is in line with many other digital library projects which have been undertaken in recent years throughout the world. In contrast to most of these projects, the DDB is designed to integrate content from different cultural domains including museums, libraries, archives, scientific institutions, and others. However, even if only cross-domain cultural heritage projects are taken into account, a multitude of approaches for integrating heterogeneous metadata can be found:

- The lowest-level approach can be found in the CatchUp system [3] from the Netherlandic "Haags Gemeentemuseum": It solely relies on a full-text index of all metadata fields; distinctions like the already mentioned "book *is about* or *was created in* a certain epoch" are not expressible at all with this approach.
- Other cross domain cultural heritage portals like The Internet Archive [4] also provide access to various types of content but map the incoming metadata to a flat indexing profile similar to DC [5] or ESE [6]. This makes things easy at first glance, but comes at a price: the unspecific semantics of DC like profiles achieve high recall rates at the expense of precision.
- The project most similar to the DDB is Europeana [7], launched in 2008 "with the goal of making Europe's cultural and scientific heritage accessible to the public". But in contrast to Europeana the DDB may store binary multimedia files of any kind.

The question of how to support the user in exploratory search over a large number of items has been addressed by Urruty *et al.* [8]. They introduced an interface for video retrieval that automatically presents suggestions by extracting textual and visual features of relevant shots. The authors proposed clustering of the retrieved results based on low-level features to create groups of similar content. The description of the clusters is then used to find different aspects of search results. The authors demonstrated good retrieval results. However, the heterogeneous multimedia objects in the DDB can not be represented by a fixed set of features.

The principal idea of faceted search for exploring data is used by various systems. Schenk *et al.* present SemaPlorer [9], an application that enables search and visualisation of heterogeneous semantic data in real-time. The application explores data sources that are semantically well annotated: triples from DBpedia, GeoNames, and WordNet. The search operates with four predefined facets with fixed context view. However, the DDB search engine enables the generation of facets dynamically according to the search results.

The DDB approach shares some concepts with the work of Jankowski *et al.* [10], which is building a system inspired by the global network of knowledge [11], CIDOC CRM [12] and linked data [13]. We also map original metadata onto a CIDOC CRM graph. However, our approach on querying objects is not directly based on the graph, but on pre-compiled transitions derived from the graph because this needs much less computation resources at query time.

In [14], the authors present different types of handling multimedia objects in the cultural heritage domain. As examples, they describe a presentation-style visualization that allows conveying the semantic relation among objects and the semantic annotation of objects based on an underlying ontology. In our approach, different representations of an object can be accessed: a title (text), a preview (an HTML-snippet with a small amount of metadata and, if available, preview-images), and a detailed view (an HTML-snippet with lots of metadata and, if available, binary representations of the objects accessible via viewers). These views are generated from the metadata of the objects when the objects are ingested, and can contain the pre-compiled transitions derived from the CIDOC CRM graph. In addition, the pre-compiled transitions allow for presenting the semantic relation between objects and for navigating from one object to another object via these transitions.

3. Semantic-Based Ingest and Retrieval: The DDB Approach

A challenge for projects like DDB is to ingest a multitude of heterogeneous objects and to connect the objects correctly — thus overcoming the metaphor of metadata described as a filing box or catalogue. A more adequate representation would preserve a deeper knowledge using models like the CIDOC CRM (Conceptual Reference Model), which "provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation" [12]. Alternatively, the new data model of Europeana, EDM [15], would be usable. These models were designed with a vision of a network of knowledge in mind [11], but mapping to these often requires making implicit knowledge explicit, especially when mapping from a metadata format without precise semantics like DC (see [16]).

In this section, we present our approach on the ingest and retrieval of semantic-based representations of metadata, starting with a brief overview of the relevant aspects of the DDB architecture and its current implementation. After that, we describe our attempt to solving two important problems in the context of our platform: the dynamic generation of facets from transition paths and the identification of unique entities from literal attributes.

3.1. *Architecture and current implementation of the DDB*

The DDB consists of three main components: an ETL [17] like tool for data integration (the Augmented SIP Creator or ASC), the DDB core platform (Cortex) [18], and the presentation layer. Its architecture was developed in dependence on two reference models for archives resp. libraries: the Reference Model for an Open Archival Information System (OAIS) [19] and the DELOS Digital Library Reference Model [20]. Figure 1 visualizes our approach on the DDB architecture. It is designed to allow distributing services on different machines.

The ASC transforms arbitrary metadata describing cultural heritage objects of any kind into the internal data model which is based on CIDOC CRM. Its current

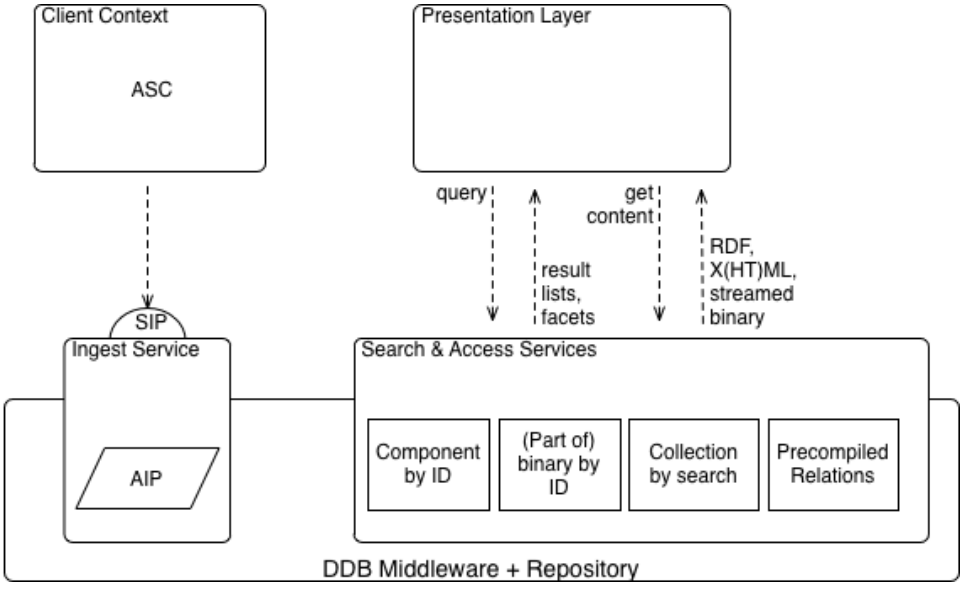

Fig. 1. High level view on the DDB.

implementation in the DDB project expects the incoming metadata to be in one of these XML-based metadata formats: DC [5], DIF [21], EAD [22], ESE [6], LIDO [23], MARC 21 [24] or METS[25]/MODS[26]. For formats that may contain more than one object in a single input file, the input file is split into single object representations. After that the input is processed with transformers written in XSLT to generate a SIP (Submission Information Package). If there are binaries available for the multimedia object, they are attached to the SIP and sent to the core platform.

When the core platform receives a SIP, it is being ingested: For every object a unique ID is created and the new object and its relations to other objects are persisted in the repository. In the current implementation, the repository consists of a file system (locally mounted, remote accessible via REST [27] or in an OpenStack cloud) and a Solr server that holds the following different indexes:

- The *search index*, which contains the full text index and the pre-computed facets for every digital object (described in detail in Sec. 3.2).
- The node store, which contains the "nodes" with their attributes and links to other nodes. A node is created for every entity and every event described in the RDF triples (described in detail in Sec. 3.3).
- The graph store, which contains the RDF triples in their unaltered form (not relevant for the topics discussed here).

After its ingestion, each object is accessible using an REST-structured URL hierarchy containing the object ID. Different URLs in this hierarchy allow for retrieving

specific parts of the object data, an RDFa description, or a binary stream. Depending on the configuration of the platform, every request may be checked against an "intellectual property rights" rule set protecting resources or parts of them. To this end, the platform offers an authorization and authentication service. Furthermore, the platform can handle user profiles which may be used to store the "journeys" the user experiences while navigating through the content.

Data managed by the core system of the DDB can be made accessible on the web through web browsers and mobile devices via the presentation layer. The most important client web application is called Object Discovery, and it provides functionality for searching, browsing and displaying objects.

3.2. *The search index: Keyword-based and faceted search*

The primary search facilities of the DDB are based on the well-known concept of faceted search. This concept has been successfully used in the context of metadata search in general [28] as well as in the context of searching in semantic web repositories [29]. The current implementation of the faceted search relies on the standard functionality of the underlying Solr index called the "search index". All information held in this index is public by definition.

While ingesting an object, the ingest service analyses the representational model of the object. The model contains transitions reflecting semantic relations. During ingest, we extract relevant transitions and reduce the transitional path to form a new synthetic property that directly combines the beginning node and the termination of the transition. The terminator may be a resource. In this case we also have to resolve the resource and extract one or multiple literals.

These synthetic properties are then being used to build the facets dynamically. All mapped values are indexed and can be queried. Figure 2 shows the concept of deducing an indexing profile from a transitional model based on an example.

Descriptive literals derived from original metadata reside in the index and refer to the items. While the model has been constructed to unveil semantic relations, facets are derived in a way that preserves the semantics of the modelled transitions, but also provides the users with easy to use filters. Facets can be computed in an acceptable time using standard technologies like Solr. In addition, they can be used for additional optimizations like topic based clustering of similar objects within result lists. With keyword search and facets, users have flexible instruments for expanding and reducing the result list. While these features depend on the model and the index configuration, additional flexibility can be achieved by expanding queries with vocabularies. Most of the metadata in the DDB is available in German only. Translations of the metadata are not in sight. However, translating the query terms is easier to achieve and could by computed during query time.

One of the main problems in this area is the huge amount of values for certain facets, e.g., time or place related facets, which can cause confusion when selecting the values for these facets via a user interface. We propose to address this issue by the use

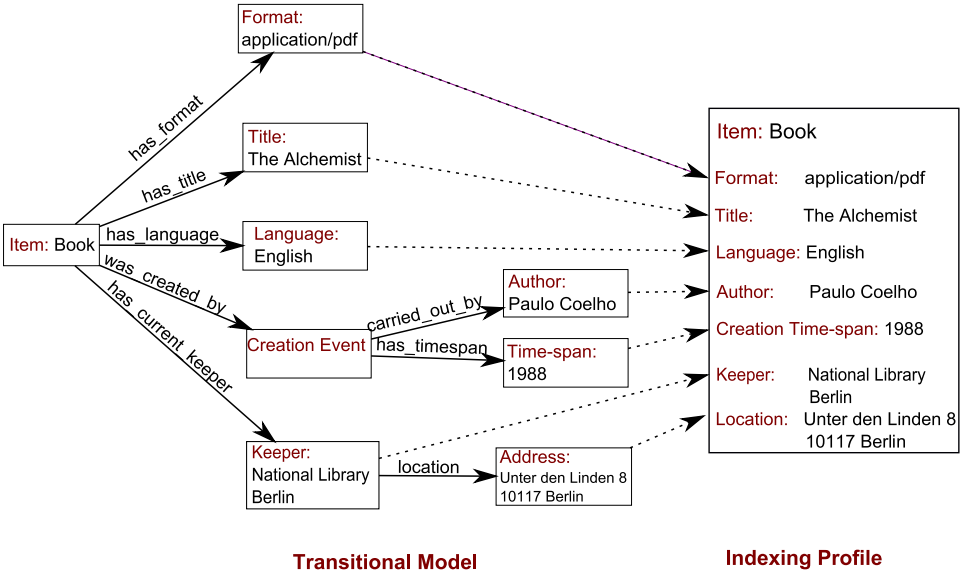

Fig. 2. Example for the deduction of an indexing profile from a transitional model.

of vocabularies to aggregate the facet values. For grouping date values in periods (50 year periods in the current implementation), we use a vocabulary that contains Latin numbers and a set of epoch names, which allows for offering time based facets that combine time values in a user friendly way instead of confronting the user with an endless number of single date values.

A similar approach is taken to manage places. Here, one of the most intriguing issues is to recognize historic names of places and to assign particular places to broader units using an appropriate hierarchy. For Germany, such a hierarchy must at least cover cities, Länder (states like Bayern or Nordrhein-Westfalen) and regions like Rheinland or Friesland. Facets containing places therefore must reflect the hierarchy.

What also come in handy for managing the multitude of facet values is that indexes like Solr can group facet values by relevance (which actually is the default behaviour).

The screenshot in Fig. 3 shows an early prototype of the faceted search, the result list, and the detail view. The detail view renders static content and currently integrates three widgets as a proof of concept: a Google map, a Wikipedia timeline, and a graph visualizer. Please note that the screenshot shows a functional prototype but not the layout of the DDB which has been implemented already but may not be published before the official launch of the website.

Another type of interactive data exploring is a map and timeline based browser that gives users the opportunity of accessing data in an explorative manner. Such a widget can be synchronized with the facetted search for an integrated way of

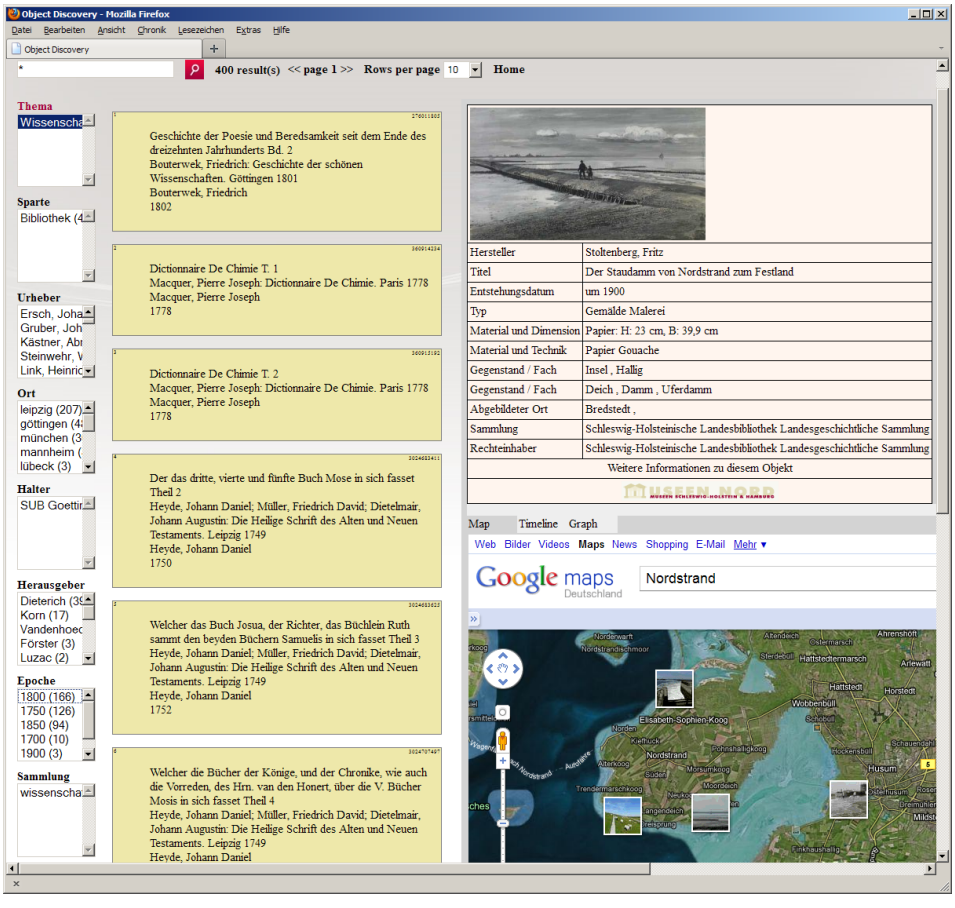

Fig. 3. Screenshot of the prototype. *Note*: This is a functional protoype, not the layout of the DDB.

inspecting search results or collections of objects. For example, users are enabled to select facets to reduce the number of objects shown in the map or timeline, or reducing the set of objects by selecting a certain area in the map. Widgets like the map and timeline browser profit from the semantics the facets reflect.

3.3. *The node store: Linking objects and identifying entities*

In addition to keyword search and facets, we wish to integrate widgets that allow navigating on the data interactively. A linked data browser would allow for traversing the properties that crosslink entities. According to the CIDOC CRM format, data is represented as a graph. The nodes of this graph represent entities related to the digital objects that are stored in the DDB as well as helper-entities such as events (e.g., a book, its creation, the person who has written it, the date of its creation, the

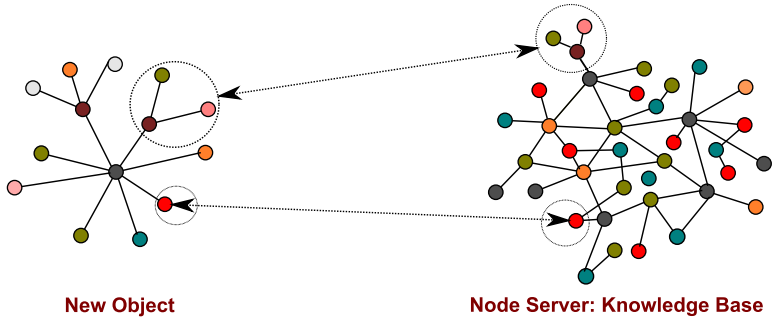

Fig. 4. Connecting multimedia objects through references to identical entities

library where it is kept, etc.). The edges represent the relations between these entities (e.g., carried_out_by, has_timespan, has_current_keeper, etc.).

Each node is stored in a Solr index called the "node store" as an independent resource with a unique identifier that refers to the original entity. Therefore, multimedia objects that share a connection to the same entity (e.g., two books that were written by the same author) are implicitly connected (see Fig. 4).

The main challenge is to ensure that no more than one single node in the repository represents the same real-world entity. Thus, the ingest service has to identify whether an entity from an incoming SIP is already present in the repository or if it has to be added. This would not be a problem if every incoming entity contained a globally unique identifier like an URI or URN reference to an authority file like PND (Personennormdatei), VIAF, or others, and if these repositories would not itself overlap. Since the name or title of the entity is needed as a literal for inclusion in the search index, it would be best if the entity would contain both the reference to the authority file and the literal. If the entity contains only the reference, it would have to be resolved to get the corresponding literal. This might be a performance problem, but that could be mitigated by pre-filling the node store with the contents of the desired authority files.

However, most of the incoming entities are not linked to an authority file but are described solely with simple literals, or are helper-entities (e.g., the creation of a book) that have no globally unique identifier. We address this problem by using the distance metric learning method presented in [30]. The key idea of this approach is to learn a function which measures the similarity of a pair of items. Since the computing of a pairwise similarity over a large amount of data is a time and resource consuming task, we propose an iterative approach involving two steps. The first step aims at quickly discovering the potentially similar objects from the repository, while the second one intends to ensure the accurate prediction of the similarity.

Automatically calculated similarities can be used for constructing relations and annotating objects with certain likelihoods. While hard and reliable links between resources and authority file entries are not always achievable, machine generated links can bridge the gap. In addition, machine generated links could be evaluated by

a crowd of competent users. Both, the semantic annotation of information and its organization into a well structured knowledge base, open up great opportunities for expressing and querying advanced knowledge about entities and relations.

However, this concept is not always understandable and usable for end users, as by design there are no direct relations between digital objects. Digital objects can only be connected via another intermediate object (e.g., the author who wrote different books), but this enables us to present this intermediate object to the end user as the source of the connection. Additional information, a detailled description of the functionality, and the concepts behind the implementation can found in [18].

4. Benchmark Results

The method of deriving facets from semantic transitions presented here aims to combine a flexible representation of cultural heritage metadata with the computational efficiency of indexing profiles. Although the DDB website is not yet launched officially, an earlier version of the DDB platform (Cortex) was already benchmarked against the open-source Fedora platform [31] for use in another context.

The test consisted of ingesting 100,000 metadata objects with a small JPEG picture (less than 200 KB) attached. It was carried out on a virtualized Linux server with the following technical data:

- 4 cores of an AMD Opteron processor 6172
- 3.86 GB RAM
- Local harddrive (79 GB)
- SUSE Linux Enterprise Server Version 11 Patchlevel 1 (64 Bit GNU/Linux Kernel 2.6.32.36-0.5-default)

The results are shown in Table 1. Note that in spite of having to translate the incoming semantic transitions to facets, the ingest process in Cortex is slightly faster while using less CPU time. The query times while still ingesting new data are even significantly faster than those of Fedora.

As of January 2012, the Cortex system has successfully ingested a mixed set of about 6 million items from all relevant domains, with search and item access performing at a reasonable speed.

Table 1. Benchmark results.

Platform	Ingest time (100,000 objects)	Query time[a] (average)	Query time[a] (std. deviation)	CPU usage
Cortex	1 h 55 min. 05 sec.	24.462 ms	68.851 ms	5–8%
Fedora[b]	2 h 07 min. 49 sec.	68.851 ms	1105.200 ms	13–19%

[a]Query times were measured while the ingest was still running.
[b]Version 3.4.2 with mySQL database 5.0.67-13.26.1 x86_64

5. Conclusions and Future Work

This paper presented an overview over some of the challenges experienced while setting up a system that aims at making cultural multimedia heritage data accessible to the public. We first described our technical approach in general, hoping this may be found helpful for similar projects. For two important issues we proposed solutions based on a rich semantic model. First, we showed how semantic transitions can be offered to the user in a user friendly and efficient way. Second, we discussed how a user may be navigating between objects.

In general, all mappings and processing steps handled during pre-ingest apply to metadata. For certain objects, more specific metadata would be desirable. For example, for books, page level metadata would allow searching for terms matching particular pages. For videos, certain frames could be addressed if the content (persons, places, dates) could be recognized and if this information was assigned to the video frames where matches occur. Similar techniques could open audio files to more sophisticated searches. Technologies that could help synthesizing metadata have been developed for the CONTENTUS use case of the Theseus project [32]. CONTENTUS offers a wide range of enrichment services working on textual content like entity recognition and keyword extraction, but also a number of functions applicable to binary formats like voice and speech recognition.

Acknowledgments

We would like to thank Timm Kißels and Jobst Löffler for providing us with their benchmark results.

References

[1] S. Becker, M. Borowski, M. Gnasa, K. Stalmann and S. Wrobel, Humanities: Intelligent analysis and information system for humanities and culture, *GI Jahrestagung* (**2**) (2010) 552−556.
[2] DBPEdia, http://www.dbpedia.org/.
[3] M. Koolen, A. Arampatzis, J. Kamps, V. D. Keijzer and N. Nussbaum, Unified access to heterogeneous data in cultural heritage, in *Proc. RIAO*, 2007.
[4] E. Jaffe and S. Kirkpatrick, Architecture of the internet archive, in *Proceedings of SYSTOR 2009: The Israeli Experimental Systems Conference (SYSTOR '09)*, ACM, New York, NY, USA, 2009.
[5] Dublin Core Metatdata Element Set, Version 1.1. http://dublincore.org/documents/dces/.
[6] Europeana Semantic Elements specification v3.3, http://version1.europeana.eu/c/document_library/get_file?uuid=84c10c80-84b8-4771-acd8-f6f0c1879a85&groupId=10602.
[7] C. Concordia, S. Gradmann and S. Siebinga, Not (just) a Repository, nor (just) a Digital Library, nor (just) a Portal: A Portrait of Europeana as an API, IFLA, 2009.

[8] T. Urruty, F. Hopfgartner, D. Hannah, D. Elliott and J. M. Jose, Supporting aspect-based video browsing — analysis of a user study, in *Proceedings of the ACM International Conference on Image and Video Retrieval 2009*, New York, NY, USA, 2009.
[9] S. Schenk, C. Saathoff, S. Staab and A. Scherp, SemaPlorer-Interactive semantic exploration of data and media based on a federated cloud infrastructure, *Web Semant.* **7**(4) (2009) 298−304.
[10] J. Jankowski, Y. Cobos, M. Hausenblas and S. Decker, Accessing cultural heritage using the web of data, in *10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2009)*, 2009.
[11] M. Doerr and D. Iorizzo, The dream of a global knowledge network — A new approach, *J. Comput. Cult. Herit.* **1**(1) (2008) 1−23.
[12] CIDOC Conceptual Reference Model, http://www.cidoc-crm.org/.
[13] T. Berners-Lee, Linked data, design issues, http://www.w3.org/DesignIssues/LinkedData.html.
[14] A. W. M. Smeulders, L. Hardman, G. Schreiber and J. M. Geusebroek, An integrated multimedia approach to cultural heritage E-Documents, *ACM International Conference on Multimedia Information Retrieval*, 2002.
[15] M. Doerr, S. Gradmann, S. Hennicke, A. Isaac, C. Meghini and H. van de Sompel, The Europeana Data Model (EDM), in *World Library and Information Congress: 76th IFLA General Conference and Assembly*, 2010.
[16] K. Kakali, M. Doerr, C. Papatheodorou and T. Stasinopoulou, DC type mapping to CIDOC/CRM (Department of Archives and Library Science, Ionian University, 2007).
[17] P. Vassiliadis and A. Simitsis, Extraction, Transformation, and Loading, in *Encyclopedia of Database Systems* (Springer, 2009).
[18] K. Stalmann, Rationale zur IAIS CORTEX Konzeption, den Datenmodellen und Mappings, 2011, http://www.iais.fraunhofer.de/fileadmin/user_upload/Abteilungen/NM/pdfs/DDB_IAIS-CORTEX_Rationale20111221.pdf.
[19] Reference Model for an Open Archival Information System (OAIS), CCSDS Secretariat, Washington, DC, 2002.
[20] L. Candela, D. Castelli, N. Ferro, Y. Ioannidis, G. Koutrika, C. Meghini, P. Pagano, S. Ross, D. Soergel, M. Agosti, M. Dobreva, V. Katifori and H. Schuldt, The DELOS Digital Library Reference Model — Foundations for Digital Libraries, Version 0.98, 2008.
[21] Directory Interchange Format (DIF) Writer's Guide, http://gcmd.nasa.gov/User/difguide/.
[22] EAD — Encoded Archival Description, Library of Congress, http://www.loc.gov/ead/.
[23] LIDO. Lightweight Information Describing Objects, http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf.
[24] Marc Standards, Library of Congress, http://www.loc.gov/marc/.
[25] METS, Metadata Encoding and Transmission Standard, Library of Congress, http://www.loc.gov/standards/mets/.
[26] MODS, Metadata Object Description Schema, Library of Congress, http://www.loc.gov/standards/mods/.
[27] R. T. Fielding, Architectural styles and the design of network-based software architectures, PhD Thesis, University of California, Irvine, 2000.
[28] K.-P. Yee, K. Swearingen, K. Li and M. Hearst, Faceted metadata for image search and browsing, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*, New York, NY, USA, 2003, pp. 401−408.
[29] M. Hildebrand, J. van Ossenbruggen and L. Hardman, /facet: A Browser for Heterogenous Semantic Web Repositories, in *ISWC*, 2006.

[30] E. P. Xing, A. Y. Ng, M. I. Jordan and S. Russell, Distance metric learning, with application to clustering with side-information, *Advances in Neural Information Processing Systems* **15** (2002) 505−512.

[31] C. Lagoze, S. Payette, E. Shin and C. Wilper, Fedora: An architecture for complex objects and their relationships, *Int. J. Digit. Libr.* **6**(2) (2006) 124−138.

[32] G. Paaß, S. Eickeler and S. Wrobel, Text mining and multimedia search in a large content repository, in *Proceedings of the Sabre Conference on Text Mining Services (TMS 2009)*, Leipzig, 2009.