

# More than the Sum of its Parts: CONTENTUS – A Semantic Multimodal Search User Interface

J. Waitelonis, J. Osterhoff, H. Sack  
Hasso-Plattner-Institute  
Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam,  
Germany  
{joerg.waitelonis |  
johannes.osterhoff | harald.sack}  
@hpi.uni-potsdam.de

## ABSTRACT

This paper presents the semantic search engine CONTENTUS [7] and its user interface, a multimedia retrieval system developed in the context of the German research project THESEUS<sup>1</sup>. This interface uses content-based suggestions and a multi-modal presentation of search results to support semantic search. In addition, the system deploys a combination of a faceted browsing and breadcrumb-based navigation; a time-line enables time based filtering of the search results and the system suggests related search results according to the user's preferences. Finally, CONTENTUS has become more than the sum of its parts. Its unique feature combination facilitates search to become a more efficient and overall more pleasant user experience.

## Author Keywords

Semantic Search, Multimedia Retrieval, Search Interfaces

## ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: Miscellaneous

## INTRODUCTION

According to ComScore the total U.S. Internet audience engaged in more than 5.2 billion video viewing sessions during September 2010<sup>2</sup> alone and the demand is still growing. The immense presence of multimedia in the WWW requires technologies for managing, organizing, and searching multimedia content in an efficient way. Within WWW search engines, the tremendous diversity of multimedia data affects

<sup>1</sup>CONTENTUS is an application scenario of THESEUS, supported by the Federal Ministry of Economics and Technology on the basis of a decision by the German Bundestag.

<sup>2</sup>[http://www.comscore.com/Press\\_Events/Press\\_Releases/2010/9/Video\\_Makes\\_it\\_Big\\_on\\_the\\_Small\\_Screen\\_in\\_Europe](http://www.comscore.com/Press_Events/Press_Releases/2010/9/Video_Makes_it_Big_on_the_Small_Screen_in_Europe)

the retrieval and search process in every part. New media analysis and retrieval methods produce metadata, which is not necessarily textual data anymore, as e.g. color features of an image. Image and video search is evolving into a more and more high level and fine granular multimedia retrieval. New metrics of similarity have to be applied to adapt the search result ranking. New user interfaces are needed to enable query formulation and intelligent visualization of search results. Today's search engines are not only focussed on finding keywords in documents anymore. Moreover their main goal has shifted to achieve a satisfying search experience for the user. E. g. personalization enables to adapt the search engine characteristics to the user's behavior and personal preferences. Semantic search technology promises more accurate and complete search results. Ontologies and semantic data repositories, as e.g., Linked Open Data [8] enable to incorporate external resources into the search process to satisfy the user's information needs and to enable a new more comprehensive search experience. Bringing together multimedia and semantic search applications bears new challenges in visualizing search results and facilitating navigation. In combination with semantic technologies, searching multimedia on-line in the WWW or in closed repositories, is subject of a paradigm shift – the route with the user's experience is more important than the original goal of fact finding.

While taking all these aspects into consideration, this paper presents a web based graphical user interface (GUI) for the multimodal semantic search engine CONTENTUS. The objective of the interface design was not to extend the research on a single aspect of current approaches in semantic multimedia retrieval, but to create an interface whose strength lays in exploiting synergies by combining state-of-the-art semantic and multimedia technologies with interface design approaches.

The paper is structured as follows: In section 2 we describe related work and introduce the relevant technologies and paradigms. Section 3 deals with the realization of the user interface including design aspects. Then, section 4 presents and discusses different show cases. Finally, section 5 concludes the paper with a short discussion of achieved results and an outlook on future work.

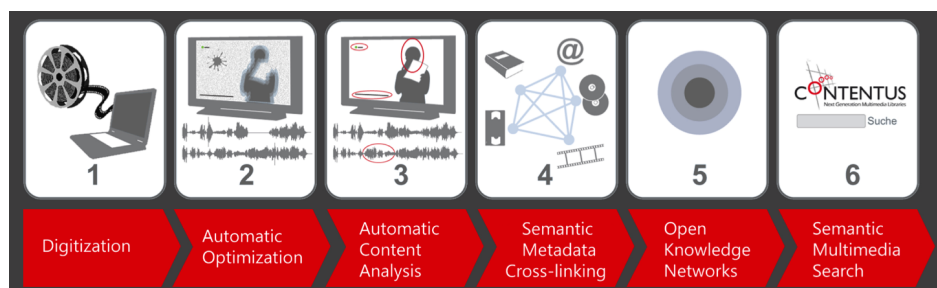


Figure 1. The CONTENTUS processing chain.

## DESIGN ASPECTS AND RELATED SYSTEMS

Germany's 30.000 libraries, museums and archives contain an incredible wealth of knowledge in the form of millions of books, images, tapes and films. Researchers involved in CONTENTUS are exploring how these cultural assets can be made available to as many people as possible and be preserved for future generations. The CONTENTUS search engine demonstrates how semantic technologies and interface design can be deployed in concert to facilitate a better and more livelier search experience within large virtual collections of cultural heritage domain assets. Fig. 1 illustrates the complex processing chain implemented by the CONTENTUS framework [10]. After digitizing all types of media assets such as books, newspaper, images, music/audio, and videos, complex workflows including media optimization, content analysis, and semantic metadata generation provide a highly heterogeneous collection of semantically annotated data. This comprises also automatically generated transcripts from speech, optical character recognition, or the extraction of semantic entities such as persons, places, and events. The main challenge of the user interface is to harness this heterogeneity of multimedia and semantic metadata in a consistent, clear, and meaningful way [6].

Another already existing system trying to meet these requirements is the the MultimediaN E-Culture Demonstrator<sup>3</sup>. It offers access to virtual cultural heritage collections based on semantic technologies[4]. MultimediaN maps painters and their oeuvre to a time-line to visualize historical references. Furthermore, search results are clustered based on the distance in the RDF graph between corresponding instances of the given search terms [3]. The 'Europeana' project<sup>4</sup> supports searching for text, images, video, and sound in a tremendous collection of paintings, music, films, and books from Europe's galleries and archives. A defaulted multimodal view shows text, images, videos and sounds within a single result page. With simple facet filtering the results can be refined, but a distinct selection on persons, locations, events, or other semantic entities is not possible. 'Culture-Sampo' is a system for creating a collective semantic memory of the cultural heritage of a nation. It provides various interfaces for faceted semantic recommendations, organizing places, people, and relations from a collaboratively generated ontology [2]. In contrast, the CONTENTUS ap-

proach is based on the semantic information, that was automatically extracted from historical documents and it provides one unified interface. The 'Parallax' interface<sup>5</sup> of the Freebase project uses a combination of faceted search, fact views, time-lines, and geo-maps guiding the users to explore a wide range of heterogeneous semantic data. But, in contrast to the CONTENTUS system, the Freebase interface is based on manually and collaboratively composed semantic datasets from the Freebase repository and does not include automatically extracted information from multimedia content. The CONTENTUS user interface is not explicitly designed to enable exploratory search, such as established in the video search engine yovisto [9]. In yovisto, exploratory search enables to broaden the search scope by expanding the search with related information.

Hence, the state-of-the-art systems provide either multimodal *or* semantic search. The objective of the CONTENTUS user interface is to melt together these paradigms by improving search results and navigation while preserving usability. Therefore, the interface provides not only information found in documents, but also information about documents and resources, its comprising entities, and their in-between relationships. For example, an instance of a person can arise as an author of a document, but also as a subject a document describes. This subtlety enables a more specific search and solidifies a new standard in semantic multimedia search.

The next section describes the CONTENTUS user interface in detail and points out its new features and their purpose.

## THE CONTENTUS USER INTERFACE

Considering Fig. 2, the layout of the CONTENTUS user interface is arranged in a *Search Area* (1) and a *Result Area*<sup>6</sup>. The *Search Area* spreads over the whole width of the layout. The *Result Area* below is divided into three columns. The left column (2) hosts the functions to organize and filter, the right column (3) the functions to refine and filter the search query. The main column (4) in the middle shows the actual search results. The *Search Area* contains the *Search Field* and the *Search Path*. The *Search Field* is the central element and is located in the upper left area. When entering a new search term, achieved results are listed in the main column.

<sup>5</sup><http://www.freebase.com/labs/parallax/>

<sup>6</sup>Search results have been modified to English language for better readability. Actually, the search engine content is in German.

<sup>3</sup><http://e-culture.multimedien.nl/demo/search>

<sup>4</sup><http://www.europeana.eu>

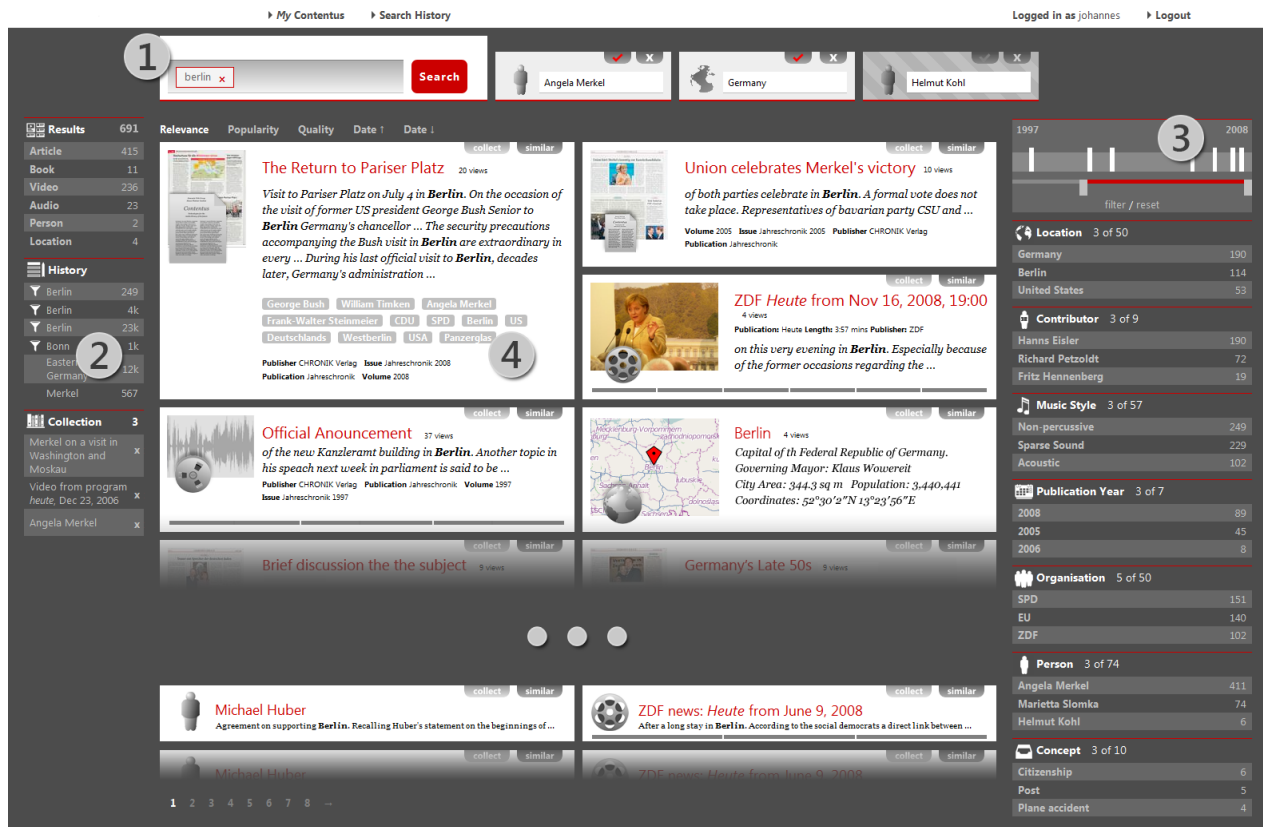


Figure 2. The CONTENTUS user interface.

### Search Results

The search interface combines different media and entity types within one result page. The types comprise books, newspaper-articles, videos, images, audio, locations, and persons. Thereby, locations and persons are entities and not documents. All types can be visually distinguished by an individual layout and icon set. For example, search results for locations reveal a geo-map, and an image preview is presented in video image results. The relevance ranking of the search result is supported by different sizes of the single result items. The more relevant an item is, the more space is provided to show related information. This enables to expose the most relevant results in more detail and thus provide more information to the user. Every search result is featured with a title and specific information depending on its media or entity type. For example, text results and results based on transcripts (e.g. video, audio) are presented with highlighted text snippets and are underpinned with semantic entities extracted from the documents content (e.g. as shown in the first result item).

### Navigation

The design of the search process outweighs the design of the single page and the former page based navigation on the web changes by the means of Ajax to a more coherent user experience. The interface makes use of this technologies with the result that elements add to and change on the screen with the aim to make the separation between entering a search and

examining results on different pages disappear.

### Timeline Filter

To filter the search results, a horizontal slider has been placed in the upper area of the right column. The handles of the slider select the starting and ending date of search results. An intuitive use of a time-line based navigation is difficult to realize because usually there are too many different dates available for different contexts. In the case of a person this can be the date of birth, the date of death, or any other event related to that person. In the case of a document it is not clear if the date of production or dates related to the documents' content should be depicted on the time-line. The time-line in the CONTENTUS user interface exclusively makes use of the publication dates. In comparison to MultimediaN the CONTENTUS user interface does not use plenty of room for its time-line. MultimediaN has more space to show dates and therefore lets its user render more complex time-based operations. The time-line of CONTENTUS plays a significant role in the interplay of the interface elements, yet it does not play the lead. It has to share the space with other equally important elements.

### Facet Filter

The list of facets is located below the time-line. The facets used by CONTENTUS are extracted entities grounded to Wikipedia articles. The location facet includes various places, sites, cities and countries. The person facet lists persons that

are mentioned in the various media items. As opposed to this, the Contributor facet lists people that have been engaged in the media document creation process, as e.g., authors of articles and books, directors of videos, musicians or composers. Facets for music style is derived from the meta data attached to audio files. Publication years are listed in the publication year facet. Various accumulations, political parties and companies are listed in the organization facet. The concept facet collects subjects and notions not fitting into the other facets.

#### *Breadcrumbs*

Selected facets are collected as boxes to the right end of the search path (1). As a result the user is always aware of her current refinements by viewing an ordered list of activated filters. After several iterations the search path usually contains a number of filter boxes. Since the users should be able to refine their search, the filter boxes in the search path contain buttons to deactivate and to remove items from the search path and thus to change the results quickly.

#### *History*

The search history enables the user to return to any possible step of the preceding search. Within a list the user can access recent actions sorted by date in ascending order. Icons indicate, if a new search term was entered, a filter was set or unset, or if a detail has been accessed.

#### *Collections*

The user can store and collect relevant search result items in his own collection simply by clicking the item's collect button. A preview of this collection is always shown on the lower end of the left column. This presence collection of media items is used to generate a user profile. The user can adapt the profile according to her preferences. The user profile can be utilized to personalize the search results.

#### *Media and Entity Details*

By clicking on an item within the search results, the media or entity detail page displays all relevant information about this item. An example for a newspaper article is shown in Fig. 3. The article viewer (left) displays the original scanned document page. All detected semantic entities within the page are highlighted. The color key encodes the entities class (e.g. red for locations, purple for persons, etc.). On the right, all pages of the underlying newspaper can be selected. Below, the article text extracted by OCR is provided and all related entities extracted from the text are registered. By clicking on an entity, its detail page appears (not depicted) and provides the further information. On the entity detail page, relationships to other resources and entities are exposed in a relationship graph and allow further navigation. The relationships are derived from the PND [5] provided by DNB. Furthermore, entities are mapped to DBpedia [1] to expose related information. The depicted article detail page (cf. Fig. 3) is just an example of visualizing media types in CONTENTUS. For video and audio items a media player, a transcript, and a temporal segmentation are displayed on the detail page to enable quick navigation within the media.

#### *Demonstration Screencast*

Because of copyright protection on its content, a public live demo of CONTENTUS<sup>7</sup> can only be provided at the conference. Anyway, a screencast demonstrates the user interface in all details (c.f. <http://www.yovisto.com/labs/vissw2011>). The following section describes some of the scenarios shown in the screencast.

### **SEARCH SCENARIOS AND DEMONSTRATION**

This section demonstrates the capabilities of the CONTENTUS user interface by means of various search scenarios.

#### *Disambiguation*

The example starts with querying CONTENTUS for the current German chancellor. While typing the first name 'Angela' a drop down box with a disambiguation of the term comes up. In this case the user selects the suggested term 'Angela Merkel' and several results of various media types dealing with the German Bundeskanzler are displayed. To make sure that the interface only lists documents related to 'Angela Merkel' as a person, the user now can select the person facet. After the selection only Angela Merkel and ministers appointed to her cabinet are shown in the results.

#### *Disambiguating Roles of Persons*

A query for 'Hanns Eisler', a famous German composer and companion of playwright Bertolt Brecht, lists documents that contain the search terms 'Hanns Eisler', but lets the user distinguish between Hanns Eisler as detected entity, as author, and as topic in the media.

#### *Persons and Exploratory Search*

When the user toggles the facet box named 'Hanns Eisler' among the bread-crumbs, again more search results are shown and Bertolt Brecht is listed as a related person to Hanns Eisler in the person facet on the right hand's side. The user now clicks on 'Bertolt Brecht' to get documents related to the German author. The user now starts a new search and enters 'Bertolt Brecht', confirms the suggestion and finds Brecht's entity page listed as first search result. She clicks on it and finds relations derived from the underlying ontology. She selects Ernst Busch from the relationship *Has Professional Contact* and is directed to Busch's entity page. There, she finds a graph depicting the relationships of Ernst Busch. As e.g., she confirms the relation to Brecht and learns that Busch was not only a musician, and singer, but also a playwright and director.

#### *Video Search Example*

This example assumes that a researcher wants to find additional information for a report on the rescue of the Chilean miners that took place in Oct. this year. Therefore she requires footage of good quality showing similar events. She has a vague idea that a similar rescuing took place in Germany in the 1960s. She queries for 'Bergleute' (German for 'miners'), limits the results by the means of the Timeline on the right-hand side to the 1960s and now sees two videos, one from November 7, 1963 and one from November 10,

<sup>7</sup><http://mediaglobe.yovisto.com:8080/contentus/>

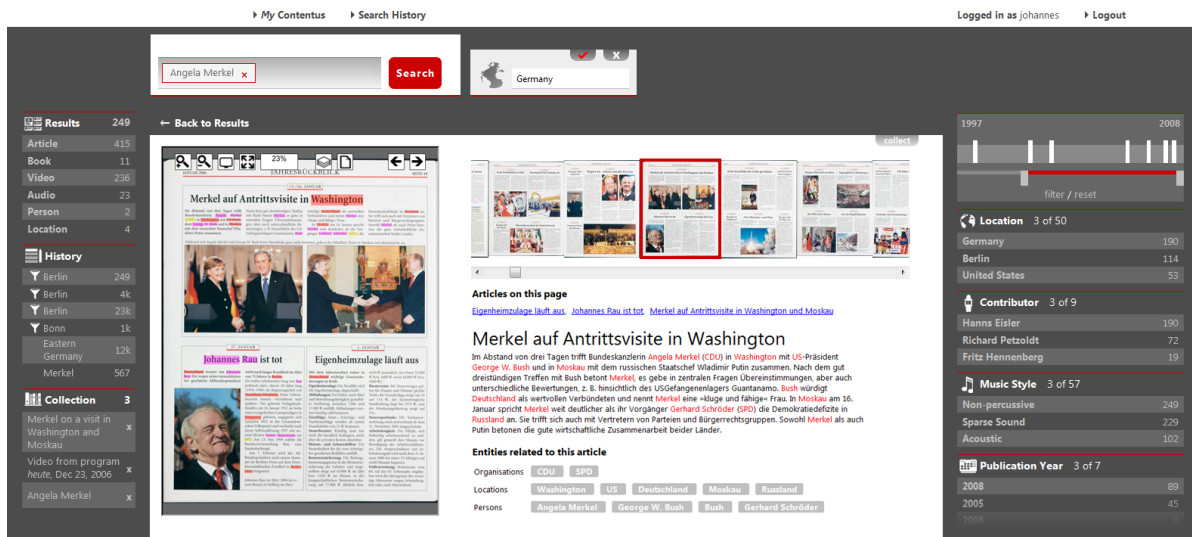


Figure 3. Details page of newspaper article with highlighted semantic entities.

1963. The first video deals with the accident of the miners, the second with their rescue.

#### Pre-selections and Final Decisions

In an exploratory search many users first pick up some candidates, continue their search and then make a final decision among these pre-selections. The interface makes it easy to select the findings in the *Media Shelf* collection. A click on the icon on top of a result tile or on the detail page adds the item to this collection. A quick preview supports the user's decision-making procedure.

#### CONCLUSION AND FUTURE WORK

When showing various media types on one result page, special care has to be devoted to the feedback that lets the user know why exactly the listed items have been presented. To find an appropriate balance between the demand for details on the one hand and ease of use on the other hand is subject to further research on CONTENTEUS. Currently CONTENTEUS is a system for individual users. In future versions we plan to expand the possibility to connect with friends and to share results with them. Then, it will be possible to make suggestions to the user not only by her personal profile but also by the profiles of her friends.

#### REFERENCES

1. S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. DBpedia: A Nucleus for a Web of Open Data. In *Proc. of 6th Int. Semantic Web Conf., 2nd Asian Semantic Web Conf.*, pages 722–735, November 2008.
2. E. Hyvönen et al. CultureSampo – A National Publication System of Cultural Heritage on the Semantic Web 2.0. In *Proc. of the 6th European Semantic Web Conference, Heraklion, Greece, May 2009*. Springer-Verlag.
3. G. Schreiber et al. . Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator. *Web Semant.*, 6:243–249, November 2008.
4. G. Schreiber et al. MultimediaN E-Culture Demonstrator. In I. Cruz, S. Decker et al., editor, *The Semantic Web*, volume 4273 of *Lecture Notes in Computer Science*, pages 951–958. Springer Berlin / Heidelberg, 2006.
5. German National Library. Liste der fachlichen Nachschlagewerke zu den Normdateien (GKD, PND, SWD). Leipzig ; Frankfurt, M. ; Berlin – ISSN 1438–1133., April 2009.
6. A. Heß. Technologies for next-generation multi-media libraries: the contentus project. In *Proceedings of the 3rd international workshop on Automated information extraction in media production*, AIEMPro '10, pages 1–2, New York, NY, USA, 2010. ACM.
7. J. Nandzik, A. Heß, J. Hannemann, N. Flores-Herr, and K. Bossert. Contentus - towards semantic multimedia libraries. In *World Library and Information Congress: 76th IFLA General Conference and Assembly*, Gothenburg, 2010.
8. Semantic Web Education and Outreach Interest Group. <http://esw.w3.org/topic/sweoig/taskforces/communityprojects/linkingopendata/>, 2009.
9. J. Waitelonis and H. Sack. Towards Exploratory Video Search Using Linked Data. In *ISM '09: Proc. of the 2009 11th IEEE Int. Symp. on Multimedia*, pages 540–545, Washington, DC, USA, 2009. IEEE Computer Society.
10. S. Weitbruch, I. Doser, and R. Zwing. Automatic structured content archiving towards improved multimedia search. In *NAB (National Association of Broadcasters) Conference*, Las Vegas, 2009. NAB.