Smart Media Navigator: Visualizing Recommendations based on Linked Data

Tabea Tietz, Jörg Waitelonis, Joscha Jäger, and Harald Sack

yovisto GmbH,
August-Bebel-Str. 26-53, 14482 Potsdam, Germany
{tabea,joerg,joscha,harald}@yovisto.com
http://www.yovisto.com

Abstract. The growing content in online media libraries results in an increasing challenge for editors to curate their information and make it visible to their users. As a result, the users are confronted with the great amount of content, unable to find the information they are really interested in. The *Smart Media Navigator* (SMN) aims to analyze a platform's entire content, enrich it with Linked Data content, and present it to the user in an innovative user interface with the help of semantic web technologies. The SMN enables the user to dynamically explore and navigate media library content.

Keywords: recommender systems, interface design, dbpedia, linked data, semantic web

1 Introduction

Content in most of today's online blogs and media libraries is arranged in chronological order, which means that the top entries receive the most attention and older yet relevant content remains hidden to the users. Linking current posts entries to older ones manually is one possibility, but contrariwise linking older content to current posts requires much time and effort and thus, cannot be achieved by most editors. In order to search and retrieve the content of multimedia platforms, manually edited tags are generally used, which lack completeness and often are ambiguous and heterogenous. A standardized vocabulary of tags may limit the editors in creativity and topics of interest that appear in future entries may not be considered. In case the user is interested in further information on a specific topic, which is actually not part of the platform, she usually has to leave the platform and may never return.

The solution to overcome these problem are *intelligent recommender systems* which are also able to help the authors to structure their content in a way that navigation and retrieval is aligned with the users needs as optimal as possible. But, many reommender systems rely on usage data and suffer from a general cold-start problem, when they are applied to a new content library without an established user community. The most approaches incorporate natural language

processing and therefore have to cope with ambiguity. By making use of formal knowledgebases, e.g. provided as *DBpedia*[1], semantic technologies help to improve the quality of recommender systems for language processing, recommendation generation, and content enrichment.

2 Smart Media Navigator

The **Smart Media Navigator (SMN)** is a navigation and recommendation framework, which uplifts a platform's content by enriching it with *Linked Open Data* (LOD). Therby, the content is complemented with additional information from the underlying LOD knowledgebase DBpedia. To support the user in navigating and exploring the content, semantic relations between different content items are visualized in a browsing interface.

As an online recommender system based on Linked Open Data, the main target groups for the SMN are broadcasting companies with online media libraries, archives with multimedia content, video-on-demand platforms, or blogs. The SMN aims to improve the user's and author's experience while curating and navigating the content. The first release will be implemented as a plugin for the blogging platform Wordpress[2]. Four main features (see Fig. 1) will be implemented in the first release:

1. Automated mapping: Fig 1 depicts the wordpress editing interface, new buttons were created in tinyMCE[3] in order to automatically annotate paragraphs with DBpedia entities. For this step, yovisto's named entity mapping was integrated through a RESTful[4] webservice. With the possibilities to automatically annotate the content and manually edit the annotations, editors and authors can decide, which entities they really want to have linked and enriched with the DBpedia data.

2. Embedded RDFa: The annotated information is directly embedded as RDFa in the articles HTML markup. When a webservice is triggered, this information is extracted out of the blog post with an *Apache Jena*[5] based RDFa extractor implementation. The extracted entities are stored in a local triple store[6].

3. Enrichment In order, to import additional information about the extracted entities, a federated SPARQL update[7] query is executed incorporating the public SPARQL endpoint of DBpedia[8]. It imports all triples the entity is involved as

[1] `http://dbpedia.org/`
[2] `https://wordpress.org/`
[3] `http://www.tinymce.com/`
[4] `https://jersey.java.net/`
[5] `https://jena.apache.org/`
[6] `http://jena.apache.org/documentation/tdb/`
[7] `http://www.w3.org/TR/sparql11-federated-query/`
[8] `http://dbpedia.org/sparql`

object or subject from DBpedia to the local triple store. This practice has the advantage, that for efficiency and availablity reasons there is no local copy of DBpedia neccessary. The update is only neccessary once when the article, or when the DBpedia changes. Furthermore, the local triple store only stores the triples, which are really used. It stays as small as possible and therefore efficient.

The imported additional information is then displayed on the articles web page. If the user hovers over an annotated entity in the articles text, additional information, such as abstracts and depictions, is appearing in a tooltip (see Fig.2).

4. Visualization The imported entity information is used to find associations (direct connections and connections of path length two) between entities of different articles, and therfore enables to recommend blog posts related to a given entity. These associations are used to create the relation browser, as depicted in Fig. 3. A main entity from a blogpost is automatically chosen (e.g. Albert Einstein) and based on this entity, further related entities are visualized in the relation browser above. They are sorted by persons, places, other 'things' and possibly also events. Hovering one entity (e.g. Zurich) reveals its relations to further entities and the main entity. Clicking on the entity Zurich in the relation browser turns it into the new main entity and the relation browser re-arranges according to Zurich. Next to the navigation feature, recommendations are shown based on the entities, the user may be interested in.

3 Conclusion

In conclusion, editors can take advantage of an integrated annotation tool, which helps to map entities with DBpedia as they write and thus, helps to reduce their workload. The additional information retrieved from DBpedia in combination with the relation browser will not substitute but improve traditionally arranged online multimedia libraries without altering their actual content. It increases the existing content's value and motivates the platform users to stay in the site.

The first showcase using the plugin will be the yovisto blog[9], which currently contains more than 800 daily articles consisting of texts, images, and videos. A demonstration of SMN will be given at the conference. The SMN is a project supported by the MIZ Potsdam-Babelsberg[10] and will be finished by March 2015.

[9] `http://blog.yovisto.com/`
[10] `http://www.miz-babelsberg.de/`

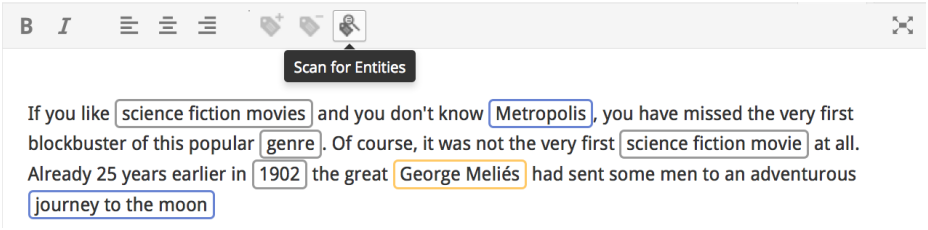

Fig. 1. Automated Annotation

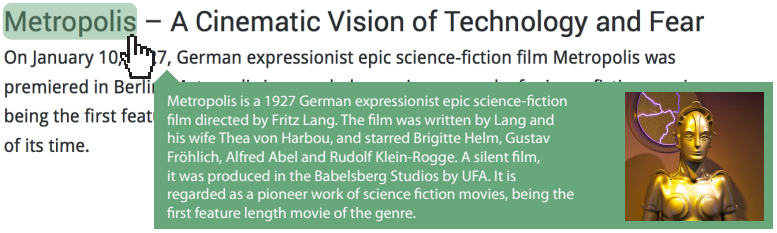

Fig. 2. Enrichment

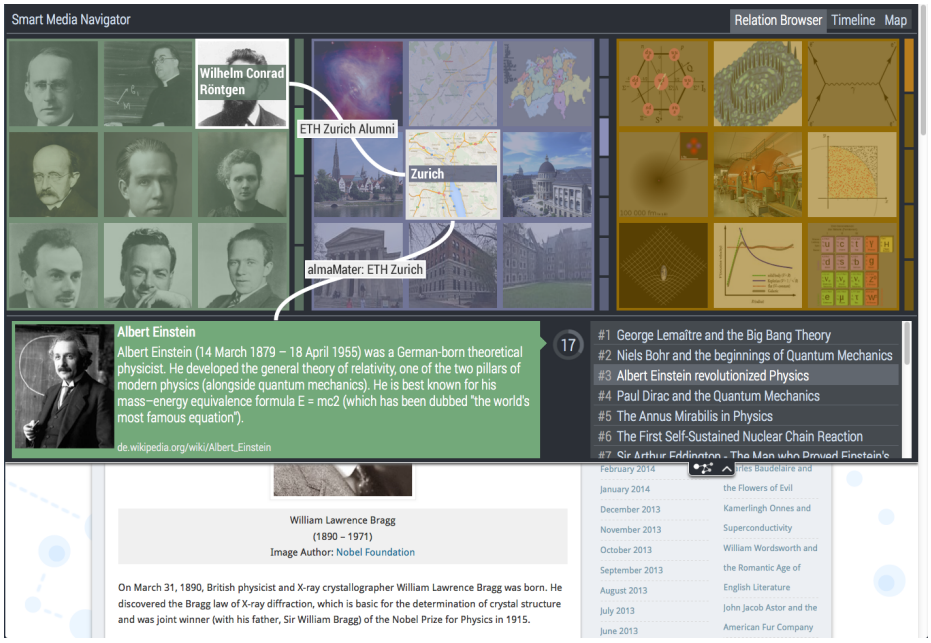

Fig. 3. Relation Browser and Recommender