

A Workflow for Cross Media Recommendations based on Linked Data Analysis

Thomas Köllmer¹, Emanuel Berndl², Thomas Weißgerber²,
Patrick Aichroth¹, and Harald Kosch²

¹ Fraunhofer Institute for Digital Media Technology IDMT,
Ehrenbergstraße 31, 98693 Ilmenau, Germany

`thomas.koellmer|patrick.aichroth@idmt.fraunhofer.de`

² University of Passau, Chair for Distributed Multimedia Systems,
Innstraße 43, 94032 Passau, Germany

`berndl|weissger|kosch@dimis.fim.uni-passau.de`

Abstract. The quality of content-based recommendation depends to a very high degree on the quality of the metadata available. We propose a workflow that combines novel cross media analysis platforms with linked data analysis to generate recommendations. The focus is set on an “editor user story” that combines live analysis of currently created content with a stored data backlog to select suitable content for article enrichment.

Keywords: Recommendation, Cross Media Analysis, Named Entity Recognition, Linked Data, Semantic Web

1 Introduction

Nowadays, a high percentage of (online) computer systems can be supported by recommender systems, assisting the decision taking process of customers. One of the biggest and well known domains of recommenders is the field of shopping. Customers are presented with masses of choices to buy, while facing the problem of finding their (preferably perfect) fit item. Recommenders work in the background in order to facilitate the process of finding relevant items for a given user. The recommender gathers information in order to generate a profile for every user (and possibly every item), which in terms can be used to determine which items to recommend.

Multimedia items are very similar to shopping items in a sense of supply and demand, so recommenders can also be used to facilitate various processes that involve the finding of multimedia. For example, knowledge discovery poses similar problems, as users are looking for content items that fit their current needs and fields of interest. A main issue of this domain is not only content discovery, but it is also more difficult to generate so called features of items – characteristics and properties of the item – which are needed in order to compare them amongst each other. Extracting those features manually is a cumbersome task and too expensive in terms of time, even more aggravated by the sheer amount of data. Automated processes can help to overcome this step, but need careful setup in order to generate useful results.

To overcome the aforementioned shortcomings of recommendation in knowledge discovery, the contributions of this paper are the design of a workflow that:

- is based on linked data analysis results, possibly allowing the reuse of data from other problem domains,
- is placed in an autonomous platform, enabling automated analysis of inserted content,
- can be altered and adjusted to given use cases, and
- is kept cross-modal, allowing recommendations being generated and consumed across different multimedia types.

In order to convey this, the paper is structured as follows. Section 2 discusses related work in the fields of recommendation and multimedia platforms. After positioning the work, section 3 describes identified use cases and their requirements, then section 4 specifies the recommendation workflow in detail. Section 5 will conclude this position paper depicting ongoing future work.

2 Related Work

The two main foundations of this work are the fields of *Recommender Systems* and *Media Analysis Platforms*. The next two sections show how the presented work fits inside the vast amount of proposed recommender systems and introduces two possible analysis platforms. The actual analysis components are considered part of the platform.

Recommender Systems

Ricci et al. describe recommender systems in [10] as systems that “are software tools and techniques providing suggestions for items to be of use to a user”. While this definition is very broad, it sets the focus to the user and its expectations. In a sense, recommender systems are more than a shopping guide, but useful in every setting that helps a human to select items from a wide set of items faster than without technical help.

The majority of recommender systems and the research on recommender systems is divided into two categories or two sources of data: The first approach is to observe user behaviour and draw conclusions from that, a technique called *Collaborative Filtering*. This assumes that within a certain user group, the user behaviour is similar. The second approach, *Content Based Recommendation*, relies on analysing the recommendation items, extracting certain features, and recommending similar items, based on that analysis.

The approach discussed here is clearly content based: Multimedia items will be processed by an analysis platform, and suggested according to their semantic similarity. However, collaborative filtering techniques can be used to reorder the list or prioritize found items. A recent survey of recommender systems is provided by Bobadilla et al. in [3]. An overview of collaborative filtering techniques was compiled by Su et al. ([12]), content based techniques are discussed by Azzani et al. in [8].

Cross Media Analysis Platforms

The work proposed in this paper was conducted within the scope of the MICO project³, a project with the aim to develop a multimodal analysis platform for various kinds of data. However, the proposed setting is not limited to MICO, but can be used by alternative frameworks as well.

The MICO project provides a platform that helps orchestrating different combinations of registered extractors. As every extractor might produce its results in different varying formats, it is desired to find one common denominator in order to make full use of the combined metadata and even enhance the degree of information by its recombination. One approach to achieve this is posed by the Semantic Web and its technologies to produce Linked Data (LD). Among those, the Resource Description Framework (RDF) [7] is the commonly known standard for producing metadata that allows semantic interlinking on the level of single resources and comprehensive querying with SPARQL[9]. The MICO Platform [11][5][2] utilised in our approach in combination with its MICO Metadata Model MMM [1] is an environment that allows to discover the hidden semantics of media in context by orchestrating sets of different components in a pipeline that jointly analyse content, each adding its bit of extra information to the final result.

³ http://cordis.europa.eu/project/rcn/111088_en.html

We will extend this context in order to design a recommendation engine as part of a MICO workflow pipeline. Consequently, the recommendation process and especially its results can be stored as metadata, making it possible to generate cross-media recommendations based on metadata results of various extractor combinations.

A similar semantic web-based approach is proposed by De Meester et al. [4] and Verbogh et al. [13]. They also tackle the problem of the always increasing masses of multimedia data and the accompanying task of costly or tedious multimedia retrieval. By supporting a framework that allows automated analysis of multimedia data, they create a rich metadata background for every multimedia item, which in terms allows users to find their desired content more quickly and efficiently. Next to this, the process of annotating multimedia is automated and hence it is not another burden to the user or data-publisher itself.

Their proposed framework meets state-of-the-art standards and analysis techniques and, in order to operate with an input base that is as broad and diverse as possible, is kept domain-agnostic by including multimedia analysis procedures as web services. The output follows Semantic-Web standards, enabling it to be understood and interpreted at various different locations. When “injecting” a multimedia item, the platform makes use of the (possibly) already present metadata background. Then, it induces a three step algorithm that can be re-iterated as often as needed, until no more change in metadata and an associated increase of quality is achieved. At first, the algorithm combines the results of different analysis processes that are registered at the platform with the already present metadata. Then it can determine which results need improvement before deciding a new analysis plan for the given multimedia item. This results in a rich and robust metadata background.

3 Cross media recommendation

As discussed in the previous section, content based recommendation depends to a large amount on the quality of the content analysis. Cross media analysis, i.e., combined analysis on different media types on the input side (e.g., video and corresponding user responses) is one promising approach to obtain the needed metadata quality.

The main user story we focus on in this report is the enrichment of journalistic articles with fitting media items, the so called *editor support usecase*. Within MICO, there is more: for example, a showcase partner, Zooniverse⁴, has a related use case in which crowd-sourced discussions about a given item can generate recommendations or pointers to other discussions and consecutively other items. In this case, both the items and user discussions are analysed.

The proposed recommendation workflow is designed to deal with both problems, the main difference are the used analysis components inside the platform.

Editor Support Usecase

In today’s web, it is vital practice, to add (linked) content to an article for reasons of giving the reader further information and commitment to the site, but also for search engine optimization. Independent of the motivation, an editor profits highly from a recommendation system, that suggests fitting items *while the new article is written*. On a high level, the related user story is:

As an editor, while I create or edit articles using WordPress, I want to automatically get related articles and videos that I might link to the article.

⁴ <https://www.zooniverse.org/>

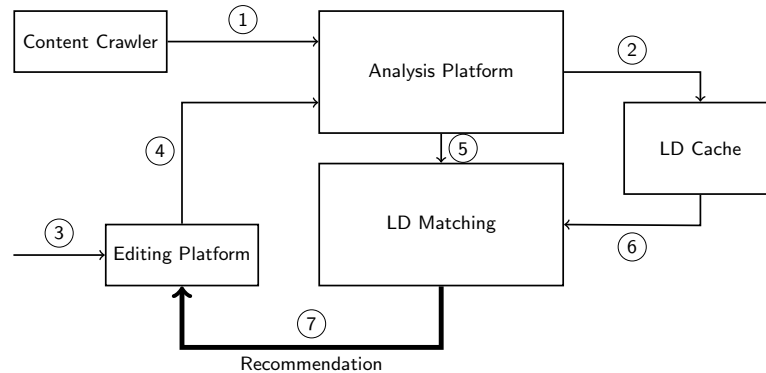


Fig. 1. Recommendation workflow combining stored results and live analysis to create recommendations. The crawled content will be forwarded to the analysis platform at any time to create background knowledge, while the current item is parsed immediately *using the same analysis platform instance*.

As a first step for generating recommendations, we implemented a pipeline designed to recommend content videos or text-comments to a given user writing her or his own comment. We will base this on the general “meaning” or topic of both, which is done by a named entity recognition (NER) component. The recommendation exposes a REST API, that can be used by a WordPress plugin to acquire the recommendation data.

4 Processing Workflow

Figure 1 depicts the general workflow for the proposed recommendation system. The involved components are described in the remainder of this section. Before discussing them, the data flow (as indicated by the numbered arrows) looks like this:

- 1 (Continuous process) A crawler feeds the domain specific media items (news, videos, ...) to the MICO analysis platform
- 2 (Continuous process) Analysis results are stored as RDF data, e.g., inside MICO’s Marmotta⁵, a Linked Data platform
- 3 An item gets into focus, e.g., someone is writing an article on a specific topic or a new user post is published somewhere
- 4 The editing platform feeds the item to the MICO platform
- 5 The analysis results are preprocessed by the LD Matching component
- 6 The LD Matching component queries the stored annotation results
- 7 The matching component calculates a similarity score and feeds relevant items back to the editing platform as a recommendation to the editing platform

Editing Platform The term *Editing Platform* stands for a component that produces content that has to be analysed and matched to the cached analysis results. A good example for this is a content management system like WordPress or TYPO3. All major systems allow to integrate plugins, therefore the communication with the new service is assumed to be a plugin that communicates with the REST endpoint of the recommender service.

⁵ <http://marmotta.apache.org/>

Content Crawler This workflow assumes that every use case has a defined subset of content that is supposed to be recommended. This might be an internal archive of media files, e.g., a in-house video collection or public databases, e.g. YouTube or Wikimedia Commons. Licensing issues are out of scope of this paper, however it should be noted that the task of storing and evaluating license information can be accomplished within the linked-data model as well. For the prototype implementation, the *Copyright Aware Crawler*⁶, developed within the CUBRIK⁷ project is used. All the crawled content is forwarded to the analysis platform for semantic indexing and further use in the recommendation process.

Analysis Platform & LD Cache As described in section 2, a central part of the proposed workflow is an analysis platform that is able to extract the desired features needed for a *content based recommendation* and is able to output its result in a linked data format, to profit from the additional semantics inside the following *LD Matching* step. The *LD Cache* component emphasizes the need of the Matching component to access precomputed analysis results. Depending on the used architecture it can be integrated into the analysis platform, as it is the case for MICO.

LD Matching As described in section 2, current recommender systems apply metrics to describe the similarity of concepts, behaviours, users, items or related item data. We adopt this kind of content-based approach, motivating the similarity of multimedia items on the NER linked data analysis produced by the MICO platform. Envisioned is an implementation of a component capable of computing a similarity score of stored videos towards written comments. By looking into the “meaning” of given analysis results the LD Matching component will try to identify the major topics of the associated resource. These will be used by a matching logic to gather relevant resources stored in the backing RDF store. To attain this goal it will try to group named entities into categories weighted by importance. Finding fitting categorisations of the initial input can be interpreted as a mapping of the given named entities into a semantic type hierarchy. The co-occurrence of relations of instances of these types as well as shared instances and similar graph structure properties form features for the similarity computation of types and their contained instances.

By interpreting this hierarchy a semantic distance metric will be derived as foundation for the recommendation system. This method will try to make use of non-textual characteristics implied by the RDF graph. Hereby it can be considered to apply information of newly analysed resources to expand the knowledge graph.

To clarify the idea of the matching process, consider following example in figure 2, showing an exemplary RDF subclass hierarchy:

Such an hierarchy will be used in our approach to match the results given by the NER analysis. After that, the semantic similarity of the given classes can be calculated to ultimately compute the similarity of two given items. For example, considering three items i_{flower} , i_{panda} , and i_{lion} with the extracted meanings of “flower”, “panda”, and “lion” respectively, one would assume that i_{panda} is much more similar to i_{lion} than i_{flower} because they are much “closer” in the hierarchy. As a result, a user writing a text about *flowers* can receive recommendations about items related to *animals*, in case there are no fitting videos about *plants* or *trees* that would be more similar.

4.1 Demo and Work in Progress

The proposed workflow was showcased as a demo inside the MICO project. During the remainder of the project it will be fully implemented for two show cases: An integration into WordPress

⁶ http://www.idmt.fraunhofer.de/en/projects/expired_publicly_financed_research_projects/cubrik.html#tabpanel-2

⁷ http://cordis.europa.eu/project/rcn/100872_en.html

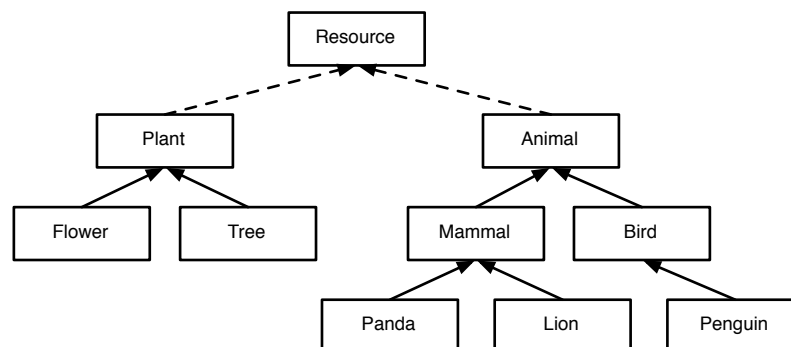


Fig. 2. Exemplary RDF class hierarchy. Solid arrows depict a direct subclass relationship, while dashed arrows symbolise a path of (possibly multiple) subclass relationships towards the top class. In RDF, every class is subclass of `rdf:Resource`, however, transitive edges derived through inference are not considered, as this would break the assumption of having a semantic distance between two classes, as every class reaches `Resource` with one “hop” to the `Resource` class.

that suggests related videos to the editor, based on speech-to-text results on the video, and NER on the draft post. The second user story is about analysing user discussions and linking it with content available for the named entities. A snapshot on the related activities for recommendation inside the MICO project can be found in [6]. The source code repository is hosted on the project’s Bitbucket repository⁸.

5 Ongoing and Future Work

This paper describes a workflow which is used in order to generate recommendations in a workflow driven environment. The process makes use of linked metadata that is produced inside the MICO platform, generating the recommendations based on a named entity recognition extractor that excerpts the meaning or trend of a written user comment towards a given video. The platform then uses that information in order to find (already categorised) fitting similar videos to recommend. As the workflow as well as its underlying utilised analysis process is kept cross-modal, other use cases, dealing with other various multimedia formats that can be analysed by NER, can easily be supported.

Next steps will also include ways of increasing the quality or reliability of the generated recommendations. Using the design of the platform, a feedback loop is envisioned, in which users can rate the received recommendations. This information will then be used in the analysis process as well as the recommendation generation process.

The matching algorithm gives room for extensibility as well. Especially in the text analysis, for example textual feature-based approaches that use `Word2Vec`⁹, a framework designed by Google to compute vector representations of words, can be used in order to further interpret the similarities of extracted named entities. This promises further improvement of the overall results, as written posts can be analysed more explicitly.

⁸ <https://bitbucket.org/mico-project/>

Please write an email to the authors to get full access to the recommendation and platform repository.

⁹ <https://code.google.com/archive/p/word2vec/>

Acknowledgements

This work has been partially funded by the European Commission 7th Framework Program, under grant agreement no. 610480.

References

1. Aichroth, P., Weigel, C., Kurz, T., Stadler, H., Drewes, F., Bjorklund, J., Schlegel, K., Berndl, E., Perez, A., Bowyer, A., Volpini, A.: Mico - media in context. In: Multimedia Expo Workshops (ICMEW), 2015 IEEE International Conference on. pp. 1–4 (June 2015)
2. Aichroth, P., Weigel, C., Kurz, T., Stadler, H., Drewes, F., Bjorklund, J., Schlegel, K., Berndl, E., Perez, A., Bowyer, A., et al.: Mico-media in context. In: Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on. pp. 1–4. IEEE (2015)
3. Bobadilla, J., Ortega, F., Hernando, A., Gutiérrez, A.: Recommender systems survey. *Knowledge-Based Systems* 46, 109–132 (2013)
4. De Meester, B., Verborgh, R., Pauwels, P., De Neve, W., Mannens, E., Van de Walle, R.: Towards robust and reliable multimedia analysis through semantic integration of services. *Multimedia Tools and Applications* pp. 1–20 (2015), <http://dx.doi.org/10.1007/s11042-014-2445-9>
5. Fernández, S., Schaffert, S., Kurz, T.: Mico. *Proceedings of the 24th International Conference on World Wide Web - WWW '15 Companion* (2015)
6. Köllmer, T.: Utilizing the mico platform for cross media recommendation (2016), <http://www.mico-project.eu/cross-media-recommendation/>
7. Manola, F., Miller, E.: RDF primer. W3C recommendation, W3C (Feb 2004), <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>
8. Pazzani, M.J., Billsus, D.: Content-based recommendation systems. In: *The adaptive web*, pp. 325–341. Springer (2007)
9. Prud'Hommeaux, E., Seaborne, A., et al.: Sparql query language for rdf. W3C recommendation 15 (2008), <https://www.w3.org/TR/rdf-sparql-query/>
10. Ricci, F., Rokach, L., Shapira, B., Kantor, P.B.: *Recommender systems handbook*. Springer (2011)
11. Schlegel, K., Berndl, E., Granitzer, M., Kosch, H., Kurz, T.: A platform for contextual multimedia data: Towards a unified metadata model and querying. In: *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*. pp. 1:1–1:8. i-KNOW '15, ACM, New York, NY, USA (2015), <http://doi.acm.org/10.1145/2809563.2809586>
12. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. *Advances in Artificial Intelligence* 2009(12), 1–19 (2009)
13. Verborgh, R., Deursen, D., Mannens, E., Poppe, C., Walle, R.: Enabling context-aware multimedia annotation by a novel generic semantic problem-solving platform. *Multimedia Tools and Applications* 61(1), 105–129 (2011), <http://dx.doi.org/10.1007/s11042-010-0709-6>