Title	Knowledge-based recommendations of media content case magazine articles
Author(s)	Vainikainen, Sari; Melin, Magnus; Södergård, Caj
Citation	Journal of Print and Media Technology Research vol. 3(2013):3, pp. 169 – 181
Date	2013
Rights	Copyright © [2013] IARIGAI. This article may be downloaded for personal use only

Journal of Print and Media Technology Research

3-2013

September 2013

Thematic issue
Content technologies

Guest Editor
Caj Södergård

J. Print Media Technol. Res. 2(2013)3, 125-214

Contents

A word from the Guest Editor

Caj Södergård
VTT - Technical Research Centre of Finland, Espoo

E-mail: Caj.Sodergard@vtt.fi

Content technologies provide tools for processing content to be delivered via any media to the target audience. These tools are applied in numerous ways in media production. Research into content technologies is very active and opens new possibilities to improve production efficiency as well as to enhance the user experience and thereby the business value of media products and services.

This thematic issue focuses on several applications of content technologies. All papers address the user, and the ability to objectively measure and predict the responses various content causes in users is a much needed tool for the media professional. An emerging application proposed in this issue helps journalists find interesting topics for articles from the excessive information available on the internet. Another class of applications dealt with here is recommending content to the users. Relevant recommendations motivate the user to visit and spend time on a web service. Recommenders are therefore important in designing attractive - and monetizable - digital services. As a consequence, this technology is found in many services recommending media items such as music, books, television programmes and news articles. The papers on recommenders in this issue cover the three main methods in the field - content-based, knowledge-based and collaborative - and they bring new perspectives to all three. One such novel perspective which has been evaluated in user studies is that of a portable personal profile.

Most of the included papers are outcomes of the Finnish *Next Media* research program (www.nextmedia.fi) of Digile Oy. Next Media has run from 2010 through 2013 with the participation of 57 companies and eight research organisations. The volume of the program has been substantial; annually around 80 person years with half of the work done by companies and half by research partners. The program has three foci: e-reading, personal media day, and hyperlocal. The papers in this issue represent only a small part of the results of Next Media. As an example, during 2012 the program produced 101 reports, most of which are available on the web.

Even if this thematic issue is centred on work done within the Finnish Next Media program, content technologies are of course studied in many other places around the world. The paper by NTNU in Norway presented here is just one example. Computer and information technology departments at universities and research institutes often pursue content related topics ranging from multimedia "big data" analysis to multimodal user interfaces and user experience. In the upcoming EU Horizon 2020 program, "Content technologies and information management" is a major topic covering eight challenges. This will keep the theme for this thematic issue in the forefront of European research during the years to come.

Caj Södergård, guest editor of this issue of JMTR, holds a doctoral degree in Information Technologies from the Helsinki University of Technology. After some years in industry, he has held positions at VTT as researcher, senior researcher, team manager and technology manager. His work has resulted in several patents and products used in the media field. Currently Caj Södergård is Permanent Research Professor in Digital Media Technologies at VTT.

JPMTR 024 | 1315
UDC 004.42:81'37

Research paper
Received: 2013-08-19
Accepted: 2013-11-01

Knowledge-based recommendations of media content - case magazine articles

Sari Vainikainen, Magnus Melin, Caj Södergård

VTT Technical Research Centre of Finland
Vuorimiehentie 3
P. O. Box 1000
FI-02044 VTT, Finland

E-mails: sari.vainikainen@vtt.fi
magnus.melin@vtt.fi
caj.sodergard@vtt.fi

Abstract

A successful media service must ensure that its content grabs the attention of the audience. Recommendations are a central way to gain attention. The drawback of current collaborative and content-based recommendation systems is their shallow understanding of the user and the content.

In this work, we propose recommenders with a deep semantic knowledge of both user and content. We express this knowledge with the tools of semantic web and linked data, making it possible to capture multilingual knowledge and to infer additional user interests and content meanings. In addition, linked data allows knowledge to be automatically derived from various sources with minimal user input. We apply our methods on magazine articles and show, in a user test with 119 participants, that semantic methods generate relevant recommendations. Semantic methods are especially strong when there is little initial information about the user and the content. We also show how user modelling can help avoiding the recommendation of unsuitable items.

Keywords: recommendation systems, personalization, semantics, semantic web, linked data, media services, metadata, user profiles, ontology

1. Introduction

1.1 Background and objectives

Media services compete for the attention of the users. As pointed out in the theory of attention economy (Simon, 1971; Davenport and Beck, 2002), human attention is a scarce commodity that determines what enters our awareness and what we decide to act on. To be successful, a media service must ensure that its content grabs the attention of the audience. One way to do this is to offer content that the user finds relevant in his/her current situation. Well-known examples from the web search domain are Google's PageRank algorithm (Brin and Page, 1998) that determines the presentation order of the search results and the AdSense program linking advertisements to search queries.

Recommendations are a central way to gain attention in digital media services. A recommender or recommendation system (RS) is a computer-aided tool that helps users find relevant items such as magazine articles, books, songs, movies, suppliers or people. An RS should assist users in avoiding poor decisions. The information that forms the basis for the recommendation system may be collected explicitly (typically from registration or profile data, users' ratings, social networks) and implicitly (typically by monitoring user behaviour, like web pages browsed) (Bobadilla et al., 2013). Recommendation can be viewed as *filtering* of data. Filtering methods are usually divided into three main categories (Meymandpour and Davis, 2013): *collaborative filtering* that employs user ratings and browsing history without further information on the items, *content-based filtering* that uses item descriptions and compares them to user profiles, and *knowledge-based filtering* that matches knowledge about user interests and preferences with knowledge about items. Knowledge is expressed as structures, such as ontologies (Middleton, Shadbolt and De Roure, 2004), restrictions and cases. In some taxonomies, the term *demographic filtering* is used to depict that common personal attributes (sex, age, profession, etc.) influence the recommendations. *Social methods* taking into consideration a user's social networks are becoming more and more important. *Hybrid approaches* integrate features from several methods. They have been found to produce better results than any single method alone, like in the Netflix Prize competition (Bell, Koren and Volinsky, 2008). One drawback in current collaborative and content-based approaches is the shallow knowledge of user interests and meanings in content items. This weakness is especially severe when new users enter a recommendation service - this is the so called cold start problem. By acquiring knowledge about the user and storing it into a user profile, recommendations can be generated right from the start. To be able to express deep knowledge about user and content, we, like some others in the field, use semantic web technologies and linked data. This makes it possible to capture multilingual knowledge and to acquire additional knowledge from external databases with minimal user input. Many semantic databases have made their knowledge available through open API's (Application Programming Interface) and use the linked data principles defined by Berners-Lee (2006). Based on these principles, Uniform Resource Identifiers (URIs) are used as pointers to concept definitions. The same concept (e.g., temperature) may be present with various naming (e.g., Fahrenheit, Celsius) in different databases but, as long as the databases include links to the same concept, the ambiguity can be resolved. The semantic web defines knowledge resources as subject-predicate-object triplets using the Resource Description Framework (RDF)[1] and the Web Ontology Language (OWL). These resources can be queried from RDF databases using SPARQL queries[2]. Large open knowledge bases include a Google-owned community-built database called Freebase[3], the Wikipedia based DBpedia[4], the geographical database Geonames[5], and the Finnish KOKO ontology of the Finnish Ontology Library Service ONKI[6]. These databases overlap to some extent, but each of them contains unique knowledge and, to maximise the covered concepts, several databases need to be used. KOKO is good for general concepts and gives information about related concepts, whereas Freebase and DBpedia contain a large amount of knowledge about persons, music, and movies. Our aim is to develop recommendations that rely on deep understanding of both the user and the content and in this way to improve the recommendation relevancy. The recommender must allow for almost automatic knowledge capture, minimizing the amount of user input necessary. In this paper, we report the results of applying these methods in recommending magazine articles. The recommendation quality was evaluated in a user test.

1.2 Related work

Meymandpour and Davis (2012) have developed linked data based similarity metrics for recommending closely related resources. Their metrics is based on shared concept features and information theory. In our work, we also match the semantic meaning of content and user profiles, but we use semantic reasoning instead of mathematical similarity measures.

Middeton, Shadbolt and De Roure (2004) take an ontological approach to user profiling for recommending on-line academic research papers and claim a performance improvement compared to non-semantic recommenders. Unlike our work, where knowledge about the user is captured from external sources, they rely entirely on monitoring the user behaviour in the actual service.

Safoury and Salah (2013) propose a recommender based on user demographics as a way to avoid the cold start problem of content-based and *collaborative filtering*. Applying the method on MovieLens[7] data, they found that the demographic data in the MovieLens dataset did not influence differentially on users' ratings. In our work, we use more knowledge of the user than just demographics. Fernandez-Tobias et al. (2011) have developed semantic-based cross-domain recommendations. Their aim is to recommend music that is relevant to a particular place. Linked data (DBpedia) is used for finding semantic relations between places and music and a weighted graph is generated to match items between the target and source domains.

In comparison, we use several linked data sources to get additional information of different domains, e.g. music, books and movies, to recommend content items based on the user interests on various levels (e.g., genre/artist name, movie/actor). The remainder of this paper is structured as follows. First, we lay out our framework with its central principle of semantic enrichment. We then present our knowledge-based recommendation methods. In the results part, we describe a user trial with magazine article recommendations and evaluate the performance of our methods.

2. Framework and methods

2.1 Semantic portable profile platform

The user profile is central in our knowledge oriented approach, as pointed out above. Our Semantic Portable Profile Platform (SP3) supports creating, managing and utilising semantic user profiles (Figure 1). The SP3 platform has been used in a multitude of applications and its methods have continuously been developed. The platform contains tools for linking interests, context and content to semantic metadata (see Section 2.2) as well as methods for generating recommendations (Section 2.3). In this article we concentrate on using semantic enrichment for recommending magazine articles. The profile portability aspects of the SP3 platform are not within the scope of this article.

SP3 lets users create and maintain their profiles by importing data from their social media accounts and by ma-

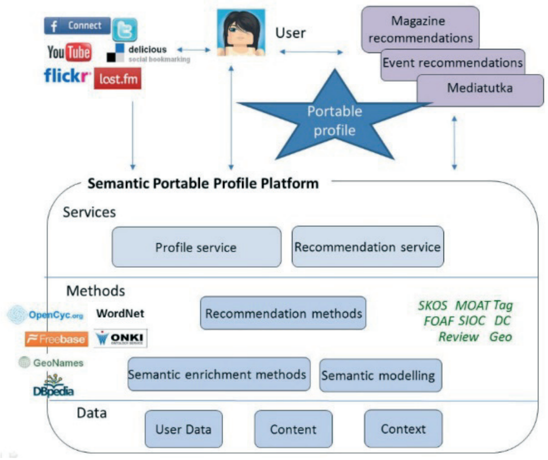

Figure 1: Semantic Portable Profile Platform (SP3). (Mediatutka refers to a mobile application which supports location based recommendations.)

nually inputting interests with the help of a semantic autocomplete widget (Figure 2). Life situation information such as family and employment situation can also be added.

The profiles are *portable,* which means that they can be used in several services. Third parties that wish to utilize the user's existing profile can do so with the user's acceptance using OAuth[8] - an open standard for authorization. REST (Representional State Transfer) APIs are offered to create, update, read and delete profiles from other services.

User and content metadata is modelled in RDF format, reusing concepts and properties from existing ontologies. The user data is coded with the help of following ontologies: FOAF[9] is used for describing demographics and online accounts, vcard[10] for home and work addresses, geo[11] for location co-ordinates, Review[12] for user ratings of interests with the values hate, like, and love, tags[13] for user interests, dcterms[14] for links to URIs of interests, and SCOT[15] for recording the account from where an interest is generated.

Content is modelled using Dublin Core dc[16] (e.g., title, description, subjects as keywords), dcterm (subjects as URIs, issue date) and prism[17] (name and number of publication, location).

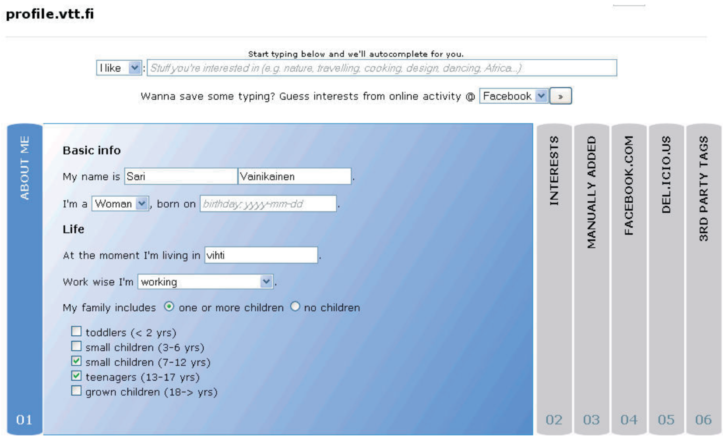

Figure 2: User profile creation in the profile service. Demographic and life situation information and user interests retrieved from different data sources can be viewed under the different tabs

2.2 Semantic enrichment

2.2.1 Enrichment methods

Our semantic enrichment methods use linked data for attaching metadata to users' interests or content metadata. Semantically enriched user profiles and content metadata are then used for personalized recommendations.

The workflow of our semantic enrichment methods (Figure 3) is described in the following sections.

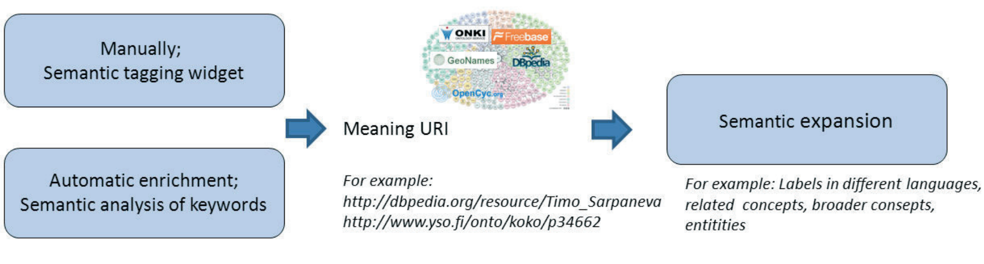

Figure 3: The workflow of semantic enrichment

2.2.2 Semantic tagging widget

The Semantic tagging widget (Vainikainen, Näkki and Bäck, 2012) lets the users manually add semantic tags for describing their interests. The same widget can also be used for content annotation. This autocomplete widget can be configured to use one or several linked data sources depending on the requirements of each particular case. When the user selects the suggested tag, its meaning is captured as a Linked Open Data URI.

In the profile service, automatic suggestions have been limited to general terms from the Finnish KOKO-ontology; music, movies, books, theatre and sports from Freebase; and places from Geonames. KOKO supports the management of general concepts in Finnish, Swedish and English. With Freebase, suggestions can be expanded to persons and their creations. When automatic suggestions are limited to fewer semantic databases and categories, automatic suggestions become clearer and suggesting the same concept from different semantic databases can be avoided.

2.2.3 Semantic analysis of keywords

Our semantic keyword analysis (Nummiaho, Vainikainen and Melin, 2010) uses publicly available knowledge bases - WordNet[18], KOKO, Geonames, DBpedia and Freebase - to analyse tags and keywords and turns them into semantic elements. We use it in the semantic analysis of short texts (such as TV synopses or tweets), and for analysing the keywords and tags coming from a user's social media accounts. Also service providers' vocabularies and categories are semantically enriched using this method.

The process of adding semantic meaning to keywords consists of several steps. First, we try to detect the language used. If the language is identified as English, Finnish or Swedish, we proceed to semantic annotation.

a) If the keyword is in English, we look it up in WordNet and determine its most likely meaning by finding similarities with the other keywords that it was used in conjunction with.
b) If the detected language is Finnish or Swedish, we first try to figure out its meaning using KOKO. Finnish words are POS (Part-Of-Speech)-tagged, which means that they are identified as nouns, verbs, adjectives, etc. If there is no direct match in KOKO, we try to find a match for the word's plural form, and if this is not successful, we check the spelling suggestions for the word and their plural forms. If there is no match in KOKO, we translate the word to English and look it up in WordNet.
c) For English and Swedish words without match in WordNet, we first translate the word into Finnish and look up meanings for it in KOKO, also trying the word's plural form if needed. In case we still do not find a match, we look up the word in DBPedia, Freebase, and Geonames and choose the one with highest confidence. Confidence is estimated using the Jaro-Winkler distance (Winkler, 1990).

If the language was not detected as English, Finnish or Swedish, we modify the word using spelling suggestions for English and Finnish in order to see if the word would, after all, be in Finnish or English, and do the same tests as described above.

We obtain the plural forms of Finnish words from Joukahainen[19] and Finnish spelling suggestions from Tmispell[20]. We use Suomi-Malaga[21] to POS-tag Finnish words. We get the English spelling suggestions from ASpell[22] and translations from Microsoft Translator[23]. We store all alternative meanings for the analysed concepts in RDF format using SKOS[24], MOAT[25], SCOT and Tags ontology. The primary meaning is linked to the analysed concept with the skos:closeMatch property.

2.2.4 Semantic expansion of concepts

Once the semantic meanings and their URIs have been defined, additional information relating to the concepts will be retrieved from the original linked datasets and stored in the SP3 platform databases. We combine data from the different linked data sets and load it into a common schema. The schema is based on the SKOS ontology and it defines the concepts and their relations. Additional properties will be used to define location data. The integration and simplification of heterogeneous data enables us to use it efficiently in generating the recommendations.

For every meaning (skos:Concept), we store its language versions (skos:prefLabel), its type (rdf:type) and links to similar concepts (owl:sameAs, skos:closeMatch). Relations between concepts are stored in skos:narrower, skos:broader, skos:related, skos:narrowerTransitive and skos:broaderTransitive properties. We retrieve additional information depending on the type of the concept. For example, if the concept is a music artist, information relating to music genre, bands, and other artists in the band will be retrieved; if the concept is a movie, information relating to its actors, directors, writers, genre and other movies in the same genre will be retrieved. For sports related concepts, additional information about the league, teams, players or athletes will be retrieved. For Geonames location concepts, geo coordinates and place hierarchy, such as continent, country, administrative divisions and nearby places, are retrieved. In addition to the SKOS ontology, Geonames ontology[26] properties such as featureClass, featureCode, inCountry, countryCode as well as the geo ontology properties geo:lat and geo:long are used to store location information into the profile database. We use an OpenLink Virtuoso RDF[27] database to store the semantic data.

2.3 Knowledge-based recommendation methods

2.3.1 Algorithms and workflow

Knowledge-based recommendation algorithms compare semantic content items with semantic user preferences and select the closest matches to recommendations.

We create the semantic representations using the above described methods. Mathematically, the user i gets an ordered list of content item recommendations $\{A'_{i,1},.....A'_{i,p}\}$ from the matching function M having as arguments the user profiles U_i and the content items A_j for positive integers I [Equation 1]. The matching function value is called *rank value r*.

$$\forall j \in \{1,..p-1\},\ M(U_i,A'_{i,j}) \geq M(U_i,A'_{i,j+1})\ \vee\ \forall s \in I, M(U_i,A'_{i,p}) \geq M(U_i,A'_{i,p+s}) \qquad [1]$$

We have developed two variants of knowledge-based recommendations representing increasing amounts of semantic enrichment and knowledge about the user: semantic and life situation specific.

As a baseline method, we use free text indexing without semantics in the content metadata creation.

The workflow of the recommendation methods is depicted in Figure 3 and described in the following sections. It is important to point out that we generate the recommendations based on the users' semantic profiles but we use additional criteria, such as publication date or location, to make the final decision on which items to show as top recommendations to the end users.

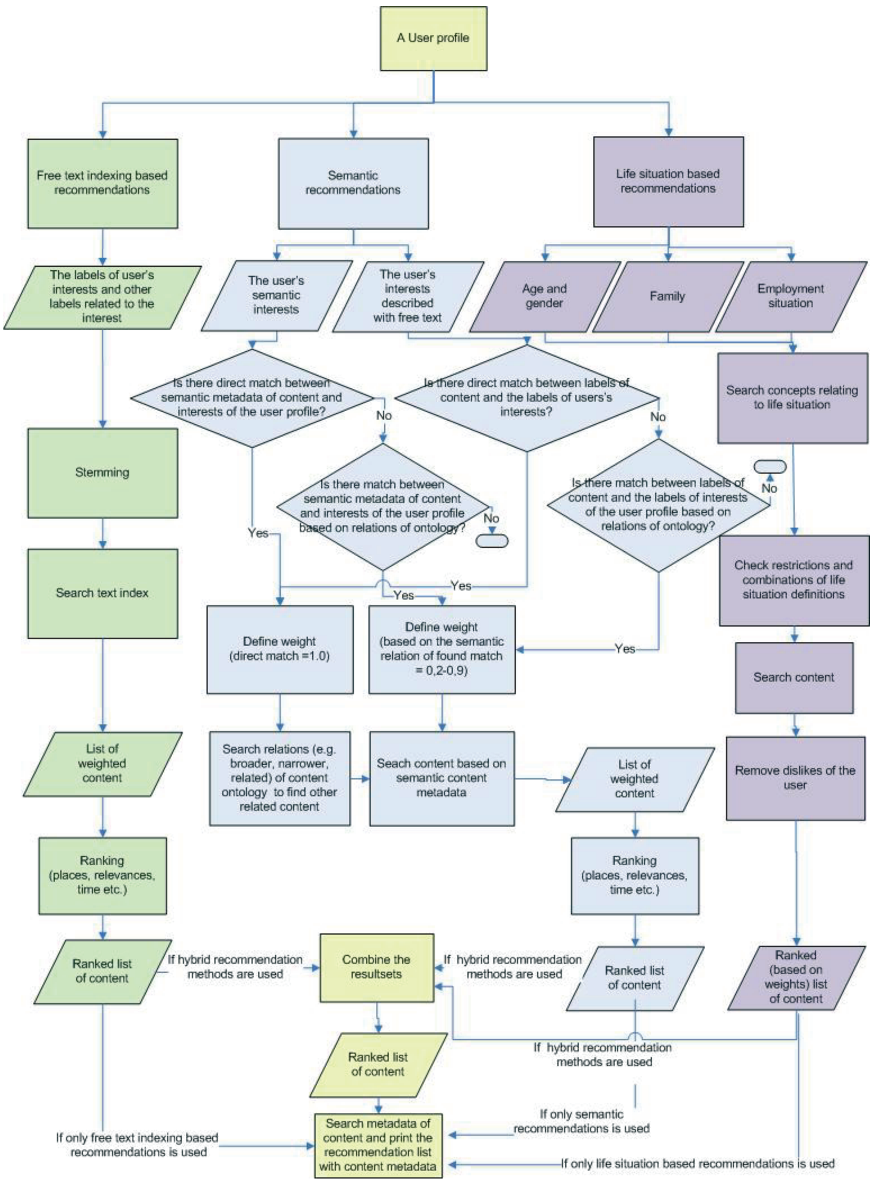

Figure 3: The workflow of the free text based, semantic and life situation based recommendation methods

2.3.2 Free text indexing based recommendations

Our free text indexing method matches semantically enriched user profiles with the content text index without any metadata enrichment. The semantic enrichment of user profiles gives support for multiple languages and enables extending searches to related subjects. We have used the Lucene[28] engine to index article texts and their existing metadata. Terms were stemmed in indexing.

The matching used the Solr[29] search engine which gives relevance values normalized between 0.0-1.0. These values were used for ranking recommendation results according to Equation [1].

2.3.3 Semantic recommendations

In this method, both users' interest profiles and content metadata were semantically enriched as described in section 2.2.

We used SPARQL queries to match semantic user profiles and content metadata. Ontological concepts may match directly or via closer or more distant concept relations. Weights were defined based on the distance of the match, so that a shorter distance gave a higher weight.

When an exact match was found between an article and a user interest, ontology relations were used for searching additional content to recommend. Weighted distances and the user-defined level of user interests (hate, like, love) were used for calculating the rank values. When several interests of the user profile match, the rank value increases.

2.3.4 Life situation based recommendations

In this method, additional knowledge was added to the user profile. We studied how to model the user's life situation and how to take it into consideration in recommendations. Life situation definitions were based on the user's age and gender, family situation such as the age of children and the employment situation, such as working, student or retired (Table 1).

Each life situation is described by an ontology with a certain concepts set. First, we looked up all concept URIs that matched each life situation that the user belonged to, meaning gender and age, employment situation, and family situation. We removed from this set of URIs all concepts that, while relevant, should not be recommended (e.g., "how to prevent burn-out at work" for the unemployed). The resulting set of URI's represents the user's stereotypical interests.

After this, we constructed SPARQL queries to find articles that match each of these stereotypical interests. Known user dislikes were filtered out of the results. We also searched for articles that would match skos:narrowerTransitive or skos:narrower of the concepts URIs of the defined life situation. The resulting articles were given weights depending on which hierarchic relation was used to obtain the match.

Table 1: Overview of the life situation definitions used in our work

Life situation		The ontology used for modelling	Examples
Age and gender	12 different groups based on gender and age limits	FOAF foaf:Group, foaf:gender, foaf: topic interest owl:Restriction for age limits	women_15to19; movies, sport, music, shopping, fashion, make-up and trends. women_50to64; food and drink, travelling, handicrafts, culture and nutrition.
Employment situation	entrepreneur, working, unemployed, student, a stay at home mom/dad, and retired	SKOS owl:NegativePropertyAssertion for defining which concepts should not be recommended	working; wages, work, travelling to work… unemployed; unemployment, job interview...
Family situation	Number of children (one or more children, no children) Ages of the children	FOAF foaf:Group, foaf:topic_interest owl:NegativePropertyAssertion for defining which concepts should not be recommended together	toddlers (<=2 years) small children (3-6 years) small children (7-12 years) teenagers (13-17 years) grown-up children (>=18) familyNoChild E.g. children, child diseases, parenthood, children's culture and children's clothes should not be recommended to a person with no children

2.3.5 Hybrid methods

We combined our semantic and text index based recommendations by calculating normalized rank values of the two recommendation result lists. The items that were found in only one of the lists were given the rank value of this list, and the items that were present in both lists were assigned a rank value that combined the two rank values. We used the following formula in combining the rank values: $r = max(r_1, r_2) + sqrt(min(r_1, r_2))$, where r_1 is the semantic recommendation rank value, r_2 the text based recommendation rank value, and r the final rank value of the recommendation. We also tested linear ($r = r_1 + r_2$) and squared ($r = max(r_1, r_2) + min(r_1, r_2)$^2) methods for combining the rank values, but the square root approach produced the best results.

3. User tests

We tested the quality of recommendations with actual users and 606 magazine articles from 12 women's magazines published by Sanoma Magazines Finland as the test set. Users were recruited by email. In total, 337 users created a profile and 119 out of them completed the entire user study.

The age distribution was between 18 and 64 with 58% of the participants being under 35 years of age. Most of the participants were women (116/119). Four recommendation methods - free text based indexing, basic and modified semantic as well as life situation semantic - were developed and tested (Table 2).

Table 2: Recommendation methods and their testing

Method	Content and logic	Profile	Testing
Basic semantic	Manually created metadata using the magazine vocabulary linked to semantic databases such as KOKO, Freebase and DBpedia to obtain semantic meanings of the concepts.	Semantically enriched concepts.	Online in two phases. 1. Users created profiles (interests, dreams, etc.) manually either by using free words or by choosing words that were linked to a semantic database. Linking to social networks (YouTube, etc.) was offered. 2. Users rated the recommendations
Free text indexing based	Text index of articles	Same as above	Same data as above
Modified semantic	Magazine vocabulary and recommendation logic were updated based on the experiences of previous tests.	Same as above	Same data as above
Life situation semantic	Magazine vocabulary ontology extended. Recommendation logic developed.	The profile page for inputting life situation data was developed	23 life situations modelled in software and recommendations subjectively tested.

Magazine articles on various topics and of various story types were chosen and a Sanoma Magazines representative added metadata manually using a magazine vocabulary[30] that was at the time under development in the company. The concepts of the magazine vocabulary were automatically analysed and links to semantic databases were added to give them the semantic meaning. The results of the automatic semantic analysis were checked manually. Of the 658 defined concepts in the magazine vocabulary, 6.2% needed correcting. In addition, names of places in the metadata of travel related articles were semantically enriched using the Geonames dataset.

The articles were available online on a web site. The users were given the tasks of creating profiles for themselves and then to rate the article recommendations that were offered to them based on their profile data. In the profile creation phase, users' interests, future plans, dreams, and current problems were asked for. Users could give their input either by entering words freely or by selecting among suggested tags using the semantic tagging widget. Users could read the recommended articles on a website that was developed for the test and they were asked to rate the recommended articles on a scale from -2 (not at all relevant) to 2 (highly relevant).

337 user profiles with a total of 4892 concepts or tags (14.8 tags/user) were created; the lowest quarter entered 9 tags at the most, half of the users gave at least 13 tags, and the top quarter entered at least 19 tags, the maximum being 45 tags. We obtained 8116 article recommendation ratings from the test users.

4. Results and discussion

4.1 Evaluation and recommendations

We calculated the precision P from the complete sets[31] of *relevant* and *recommended* articles according to Equation [2]:

$$P = \frac{|\{recommended\} \cap \{relevant\}|}{|\{recommended\}|} \qquad [2]$$

In accordance with, e.g., Bobadilla et al. (2013), we consider the recommended article to be relevant, if the user gives it a rating of 1 or 2 on the 5 point scale from -2 to 2.

The precision P_r (Equation 3) is the number of relevant ratings M divided by the total number of user ratings N

$$P_r = \frac{M}{N} \quad [3]$$

Alternatively, the precision per user is (Equation 4):

$$P_u = \frac{\sum_{i=1}^{I} P_i}{I} \quad [4]$$

where P_i is the precision for user i, and I is the number of users.

A measure of how well the recommendations are ordered is the Pearson correlation R between the rank value r and the user rating u (Equation 5).

$$R = \frac{\sum_{i=1}^{N}(r_i-\bar{r})(u_i-\bar{u})}{\sqrt{\sum_{i=1}^{N}(r_i-\bar{r})^2}\sqrt{\sum_{i=1}^{N}(u_i-\bar{u})^2}} \quad [5]$$

4.2 Semantic recommendations

The central idea of enriching the magazine vocabulary with linked data, especially with the concepts of the Finnish KOKO ontology, and utilising this information in recommendations, worked well. The magazine vocabulary is an high level ontology offering the main concepts that are important for the different magazines. With the help of linked data, these concepts could be extended into more detailed concepts. The mappings between the concepts of the magazine vocabulary and the concepts of linked data helped in generating recommendations for user interests that were not directly included in the magazine vocabulary. For example, a user had defined her interest in carpets with the Swedish word "mattor". Thanks to semantics, we obtained support for multiple languages and it did not matter in which language the interests had originally been defined. The magazine vocabulary did not contain the concept "carpets" but, based on the KOKO relations, the system knew that it was related to the concept of "furnishing fabrics" and was able to recommend articles relating to that subject.

However, the method also produced false recommendations. We analysed the false recommendations and made the following modifications to the basic method:

Limiting semantic extensions

An example of a false recommendation is when an article tagged with "MS disease" leads to recommending an article about "children's diseases".

The reason for this was that, in KOKO, the concept "MS disease" links to "disease" which in the recommendation process was narrowed to "children's diseases". In this case, the recommendations should not have been extended to other diseases.

When the link to the magazine vocabulary has been found with the help of extended links in the KOKO hierarchy, it means that the match to the concept of the magazine vocabulary is not an exact match. In such cases, the recommendations should not be extended to more specified concepts since this takes recommendations further away from the user's actual interest.

Refining class hierarchy, metadata and concept relations

Some concepts in the magazine vocabulary connected two different topics into one class (e.g., travelling and nature) and this caused problems in recommendations.

If a user had indicated that she was interested in "travelling", the recommendations were extended based on the hierarchy of the vocabulary and the user also received recommendations relating to nature. Several users who had more negative than positive ratings had used these concepts for expressing their interests.

Person related roles (babies, children, adolescents, women, men, siblings and relatives) had been defined on the same level of the hierarchy. It was difficult to recommend family related articles because a simple tag "family" or "children" does not reveal what in particular the user is interested in. If a user inserted the tag "family" into her profile, the recommendations included articles on topics from childbirth and babies to teenagers.

The articles should be annotated more in detail and users should be encouraged to define more precisely what aspects of "family" they are interested in.

More relations between concepts of the magazine vocabulary would be useful, for example to indicate a link between "weight control" and "diets".

Refining the mappings between KOKO and the magazine vocabulary

We found that the hierarchical relationships between terms were sometimes problematic, especially in situations where the mappings had been made between very general concepts such as "events" and "phenomena".

Mapping to the concepts of "organized events" and "cultural phenomena" would have been more appropriate and prevented some false recommendations.

The quantitative results from the user tests of the basic and modified semantic methods are presented in Table 3.

The precision of the basic method was 0.58 and, if rating 0 was seen as relevant, 0.74. The modified method was slightly more precise (0.63 and 0.77, respectively). The precision per user was 0.68 for the basic method and 0.70 for the modified method. Using the basic method, 12% of the users gave a negative average rating, i.e., their recommendation list contained more unsuccessful than successful recommendations. Using the modified method, the number of unsuccessful recommendations could be decreased. The average ratings of three users changed from negative to positive and the average ratings of 6 users changed to highly relevant.

The number of recommendations with negative ratings was smaller using the modified method. On the other hand, the added restrictions on ontology relations also reduced the number of relevant recommendations. In addition, the Pearson correlation between rank value and rating decreased from 0.23 using the basic to 0.15 using the modified method. This implies that the rank value algorithm needs more tuning.

Table 3: The relevance of the recommendations made using the semantic methods and evaluated with the following metrics: 1. Precision and correlation. 2. The percentage of the different rating values from the total amount of user given ratings for the generated recommendation list. The rating scale is from -2 (not at all relevant) to +2 (highly relevant). N is the total number of rated articles on test users recommendation lists. 3. The percentage of users having a rating average equal to one, positive but less than one, or negative

	Basic semantic (Number of rated recommendations N=8116)	**Modified semantic** (N=4809)
Precision P_r	0.58	0.63
Precision P_u	0.68	0.70
Pearson correlation R	0.23	0.15
Rating	**Percentage of all ratings**	
2	29	32
1	29	30
0	16	15
-1	12	11
-2	14	12
Average rating per user	**User %** (Number of users=119)	
≥1	40	45
0 ≤ average rating < 1	47	45
<0	12	10

To understand the correlation between rank values and user ratings better than just by looking at the Pearson coefficient we assume that, if the rank value is above a threshold (=1), the item exactly matches the user's interests. If the value is below the threshold and the match has been found based on the relations of the semantic concepts, it is assumed to match less well with the user's interests (Table 4).

Table 4: The correlation between calculated rank values and user given ratings of the recommended articles. The recommendations were generated using the basic semantic recommendation method

Calculated rank value of recommended article	**User rating**	**Percentage of all ratings** (N=8116)
High rank value (r ≥1)	Positive rating (0, 1 or 2)	48
High rank value (r ≥1)	Negative rating (-1 or -2)	11
Low rank value (r < 1)	Positive rating (0, 1 or 2)	25
Low rank value (r < 1)	Negative rating (-1 or -2)	17

Table 4 shows a good overall quality of our recommendations. However, 11% of the rated items had a high rank value but received a negative rating from the user. This may be because the subject area (e.g., family) was so wide and diverse that not all recommended articles were of interest to the user.

The numbers indicate that, while semantic relations have the potential of finding additional interesting articles to recommend, it is important to consider carefully how these deep semantic relations should be used in order to avoid going too far away from the user interests. We also analysed which tags caused a conflict between the rank values and user ratings. Family, children, nature, vacation and travelling are examples of such tags. This supports the observation mentioned above that too general terms produce false recommendations.

4.3 Comparing semantic recommendations to traditional free text indexing

To acquire further insights, we compared the results of our basic semantic method to results obtained when the content metadata was produced using free text indexing, and also to results gained by combining semantic and free text indexing (Table 5).

Table 5: Number of recommendations using different recommendation methods

	Semantic (basic)	Free text indexing	Semantic and free text indexing
Total number of recommendations for all users	33 519	35 516	49 620
Total number of recommendations for all users with user ratings	8 116	5 077	8 116
Average number of recommendations per user	252	272	375
Number of different articles that were recommended (Total 606 articles)	603	569	603
Number of different recommended articles that had at least one rating	541	456	541

The free text indexing method generated 6% more recommendations than the semantic recommendation method. A combination of both methods produced recommendations that could not be found using single methods; thus it generated the largest number of recommendations.

There is no big difference in the average number of recommendations per user. However, some individual users had a large variation in the number of their recommendations. User profiles influence the number of recommended articles to a great extent. General profile terms such as family, home, children and food lead to a large number of text index based recommendations because these topics are very common in women's magazines.

Although the text index based method as a whole produced more recommendations than the semantic method, 32% of users received more recommendations using the semantic method. These users had terms such as travelling, health, nature, decoration and illness in their profiles. These terms had been semantically expanded to several other terms. There were 85 articles that had not been included in any free text indexing based recommendation but had been recommended based on semantic relations. These articles were related to science, domestic appliances, food, fashion and welfare.

Relevance

We subjectively compared the relevance of recommendations produced using the two methods. We studied users with irrelevant semantic recommendations (negative average user rating) and users whose semantic recommendations were good (average user rating higher than or equal to 1.5).

Even if free text indexing increased the number of recommended articles, it also produced irrelevant recommendations and some relevant articles that had been found using the semantic method were not found at all using free text indexing. Many of the irrelevant articles at the top of the recommendation list produced by the text method were interview articles and articles relating to divorces. They ended up at the top because they often covered several topics, such as living, friends, children, family terms also and economy - popular in the users' interests. Another problem with free text indexing is the ambiguity of words: the term "renovation" (Finnish: "remontti") was used in a divorce article in the meaning of "renovation of life", whereas the user's intended meaning was "renovation of houses".

The influence of the user profile

The number of interests and the actual words used in the user profile had large influence on the results of the different recommendation methods. When there were only few interests in the user profile, the free text indexing based method performed worse that the semantic method. This was as expected, as the latter is able to infer more terms and enrich the profile. For example, 'fashion' can be extended to different accessories and 'food' to different types of courses. Very general terms in the user's interest profile caused irrelevant recommendations using both recommendation methods. It is important to guide users to describe their interests as specifically as possible.

Strengths and weaknesses of the methods

The semantic recommendation method provided the best results when both the user profile and the content metadata were semantically enriched. The drawback is that semantic annotation, even when performed automatically, sets additional requirements on the production systems. The benefit of free text indexing is that it is easy and cost effective to create.

The best recommendation results were achieved by combining the two recommendation methods. The combination creates benefits, such as:

- Finding relevant articles that cannot be found using one method only.
- In the semantic recommendation method, free tags are matched to the concepts of the magazine vocabulary but, since the number of the defined concepts is fairly low, matches cannot be always found. Text indexing produces many descriptive terms for each article.
- When a user interest is not included in the magazine ontology, semantic relations can be used to find related subjects. For example 'golf' is not part of the magazine vocabulary but if the user has 'golf' in her interests, other sports related articles can be recommended using semantics. Using text indexing, articles containing the word 'golf' can be found as well.

4.4 Life situation semantic recommendations

When analysing the false recommendations of the basic method, we found a good opportunity for improving recommendations by modelling life situations. For instance, the terms "abortion", "childlessness" and "miscarriage" should not lead to recommendation of family related articles. The problems with these terms were caused by both the magazine vocabulary and KOKO. In the magazine vocabulary, the concept "childlessness" was linked to family-related terms. The term "miscarriage" was more complicated: in KOKO, it was related to "pregnancy" and this created relations to the "pregnancy", "family" and even "breastfeeding" concepts of the magazine vocabulary. This produced inappropriate article recommendations and gave us the idea of creating a solution for managing negation.

Using the life situation semantic method, we produced and subjectively evaluated recommendations based on 23 life situations. In defining the situations, we used the demographic classification of their readers that Sanoma Magazines uses. An example: a woman, born in 1981, is connected to the life situation 'women_30to39' and she will get recommendations about subjects such as living, nature, food and drink, children, family, and sport. If she adds to her profile that she has children under 2 years, she will get recommendations relating to babies, and the recommendations will not any more include articles about older children. If she tells that she works outside of home, she will additionally get recommendations of articles relating to work. When she adds some interests, such as dance, she will also get recommendations relating to these topic.

When we subjectively evaluated the results of the life situation based recommendations, we concluded that:

- *The life situation is especially important when there is only little information about the user.*
- *Recommendations based on incompatible life situations should be excluded*

  When analysing the recommendations for retired persons, we noticed that there were irrelevant recommendations relating to the concept "sleep" ("uni", in Finnish). The concept "sleep" was connected to the stereotypical definitions of retired persons, but a problem emerged because the result set included recommendations to articles dealing with "sleep and children".

  In order to avoid this kind of recommendations, we updated the stereotypical definition of retired persons by adding definitions that described which concepts, such as children, should not be recommended together with retirement. owl:NegativePropertyAssertion definitions were used for this purpose.

- *Combining information from different aspects of life situation is a sensitive matter*

  The employment situation definition, such as "stay _at_home_mom_dad", includes mappings to children but the family definition can contain more detailed information about the age of the user's children. In this case, recommendations should prioritise articles relating to the correct age of the children.

  The same problem occurs with life situations that are defined based on age and gender, because they include mappings to the concepts of children and family. If there is information about the user's real family situation, it should be used in recommendations.

- *Life situation definitions and user interests may cause conflicts*

  An example of possible conflicts between user's interests and the life situation is that a user had defined that childlessness is a sorrow in her life. In this case, there should not be recommendations of articles relating to children, although the user may match a life situation that includes such a definition. One opportunity for future work is to make the life situation based profiles visible to users so that they can modify and control them.

5. Conclusion

We have developed recommendation methods for media content based on knowledge about the user and the content. The knowledge is expressed using the tools of the semantic web and linked data, making it possible to capture multilingual knowledge and to infer user's interests and content metadata. Linked data also allows us to automatically retrieve knowledge from various sources with minimal user effort. In addition, *a priori* knowledge of the user solves the so called "cold start" problem that *collaborative filtering* and content-based methods encounter, because with new users these methods lack enough user interaction data to make reliable recommendations.

We have applied our methods to recommending articles from women's magazines. The user tests with 119 participants verified our hypothesis that semantic methods generate relevant recommendations. Semantic methods are especially strong in cases where there is little knowledge about the user and the content. This is because semantic ontologies make it possible to infer additional user interests and content metadata. Therefore, an exact match between user interest and content metadata is not required as is the case when using non-semantic methods.

However, the semantic methods also produced false recommendations. In analysing these, we found out that there is a need to limit the usage and influence of broader concepts. If too distant concepts are used, there is a risk that the corresponding recommendations do not anymore interest the user. Another finding was that the life situation of the user is a good addition to the user model - especially when there is only little information about a user. For women's magazines, especially information about the family situation helps in identifying relevant articles to recommend. When we added the life situation model to our methods we were able to define which concepts should not occur together (e.g. childlessness - family related articles), thus avoiding inappropriate recommendations. The life situation models developed need to be tested in more extensive user tests in the future.

The tests also show that the best recommendation results were achieved by combining the semantic and the traditional text indexing methods. This helped finding relevant articles that would not have been found using one method only. This is in congruence with results from other studies that stress the merits of hybrid methods.

Creating domain knowledge is a challenge in knowledge-based recommendation systems. We can benefit from linked data in creating such knowledge and to model it so that the same concepts will be understood in the same way in different systems. This is an important aspect also in relation to the concept of portable profiles. The use of linked data sources creates new opportunities but also challenges when developing new recommendation methods. There are many different databases, APIs and data types. The amount and accuracy of data varies among concepts and databases. The future availability of open knowledge bases is an open question but many such data bases are currently available for downloading and can be installed locally.

In our future work, we intend to extend the user model to take into account additional aspects, such as values, roles and intentions of the user. The profiles need to be made adaptive to changes in user preferences. We also plan to include context information, such as the location of the user. Context information helps in recommending relevant content considering the current situation of the user. This requires more accurate user modelling including multiple sets of interests related to different contexts and roles. Similarly, the social networks of the user are important to consider in making recommendations.

Once the user profile is available in semantic form, it can be used for personalizing a wide range of services, such as search and content adoption. We envision that portable profiles under user control will become a central element in future digital services. Privacy issues and cross-service profile visibility must be considered. To advance this vision, we have created the SP3 platform that we have used in this work.

Acknowledgements

This work was supported by TEKES, the Finnish Funding Agency for Technology and Innovation, as a part of the TIVIT Flexible Services and Next Media research programs. Special thanks go to Sanoma Magazines for co-operation in this research.

References

Bell, R., Koren, Y., Volinsky, C., 2008. *The BellKor 2008 Solution to the Netflix Prize.* [Online] Available at http://www.netflixprize.com/assets/ProgressPrize2008_BellKor.pdf> [Accessed 10 October 2013]

Berners-Lee, T., 2006. *Linked Data.* [Online] Available at <http://www.w3.org/DesignIssues/LinkedData.html>. [Accessed 10 October 2013]

Bobadilla, J., Ortega, F., Hernando, A. and Gutiérrez, A., 2013. Recommender systems survey. *Knowledge-Based Systems,* 46, pp.109-132

Brin, S. and Page, L., 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems,* 30, pp. 107-117

Davenport, T. H. and Beck, J. C., 2002. *The Attention Economy: Understanding the New Currency of Business*. Boston: Harvard Business Press

Fernández-Tobías, I., Cantador, I., Kaminskas, M. and Ricci, F., 2011. A generic semantic-based framework for cross-domain recommendation. *ACM Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems,* pp. 25-32

Meymandpour, R. and Davis, J. G., 2012. Recommendations Using Linked Data. *ACM Proceedings of the 5th Ph.D. workshop on Information and knowledge,* pp. 75-82

Middeton, S. E., Shadbolt, N. R. and De Roure, D. C., 2004. Ontological user profiling in recommender systems. *ACM Transactions on Information Systems*, 22(1), pp. 54-88

Nummiaho, A., Vainikainen S. and Melin, M., 2010. Utilizing Linked Open Data Sources for Automatic Generation of Semantic Metadata. In: *Metadata and Semantic Research - Communications in Computer and Information Science*, Vol. 108, pp. 78-83. Berlin: Springer

Simon, H., 1971. Designing Organizations for an Information-Rich World. In: M. Greenberger, 1971. *Computers, Communication, and the Public Interest.* Baltimore, MD: The Johns Hopkins Press

Safoury L. and Salah A., 2013. Exploiting User Demographic Attributes for Solving Cold-Start Problem in Recommender System. *Lecture Notes on Software Engineering*, 1(3), pp. 303-307

Vainikainen, S., Näkki P. and Bäck A., 2012. Exploring Semantic Tagging with Tilkut. In: A. Lugmayr, H. Franssila, P. Näränen, O. Sotamaa, J. Vanhala and Z. Yu, eds. 2012. *Media in the Ubiquitous Era: Ambient, Social and Gaming Media.* Hersheu, PA: IGI Global. pp. 130-148

Winkler, W. E., 1990. String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage. *Proceedings of the Section on Survey Research Methods* (American Statistical Association), pp. 354-359

---

[1] http://www.w3.org/TR/PR-rdf-syntax/

[2] http://www.w3.org/TR/rdf-sparql-query/

[3] http://www.freebase.com/

[4] http://dbpedia.org/About

[5] http://www.geonames.org/

[6] http://onki.fi/

[7] http://movielens.org/

[8] http://oauth.net/

[9] http://xmlns.com/foaf/0.1/

[10] http://www.w3.org/2006/vcard/ns#

[11] http://www.w3.org/2003/01/geo/wgs84_pos#

[12] http://purl.org/stuff/rev#

[13] http://www.holygoat.co.uk/owl/redwood/0.1/tags/

[14] http://purl.org/dc/terms/

[15] http://scot-project.org/scot/ns#

[16] http://purl.org/dc/elements/1.1/

[17] http://prismstandard.org/namespaces/basic/2.1/

[18] http://wordnet.princeton.edu/

[19] http://joukahainen.puimula.org/

[20] https://github.com/voikko/tmispell

[21] https://github.com/voikko/corevoikko

[22] http://aspell.net/

[23] http://msdn.microsoft.com/en-us/library/ff512423.aspx

[24] http://www.w3.org/2004/02/skos/core#

[25] http://moat-project.org/ns#

[26] http://www.geonames.org/ontology#

[27] http://virtuoso.openlinksw.com/

[28] http://lucene.apache.org/core/

[29] http://lucene.apache.org/solr/

[30] http://onki.fi/en/browser/overview/magazine

[31] More precisely, we consider the recommendation set to be the recommendation list subset that the user has rated. Thus, the size of the set varies from user to user, the average being 48 ratings.