Proc. 7th Int. Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Luxembourg, Dec. 2012.

Semantic personalisation in networked media: determining the background knowledge

Dorothea Tsatsou, Vasileios Mezaris, Ioannis Kompatsiaris
Information Technologies Institute (ITI)
Centre for Research and Technology Hellas
Thessaloniki, Greece
dorothea, bmezaris, ikom@iti.gr

Abstract — **This paper examines the requirements and argues about the features that an ontological knowledge base should encompass in order to represent semantic information pertinent to a user and suitable for context-aware preference elicitation and content recommendation. It explores existing formal ontologies suitable to represent user characteristics and the diverse information in digital media, as well as the opportunities and shortcomings that linked open data vocabularies bring into play. Based on these aspects, the paper concludes with the reasoning behind the composition of an appropriate reference ontology for the personalisation task of the LinkedTV EU project.**

Ontologies, linked open data, personalisation, context

I. INTRODUCTION

The convergence of TV and the Web has introduced a new perspective in personalised systems by offering the possibility to interlink multidisciplinary content for consumer usage, whether that might be audio, video, text or social media, while tailoring it to the user's needs and concrete situation. Thus a challenge arises in capturing user preferences and context: intelligent interpretation of the vastly heterogeneous digital media information and sophisticated representation of user-pertinent information under a homogeneous vocabulary, which would also be scalable and usable for subsequent matching of user preferences to the disparate media content.

The question arises how this information can/should be used for personalisation and contextualisation. In a linked media environment, the user is expected to interact with all kinds of content, spanning from the multimedia content itself (with information varying from visual features, the audio transcripts and media annotation), information from the user's social online activity (comments, likes, interrelations with similar users), as well as a wide variety of miscellaneous content relevant to the media (news articles, program summaries, bios, ads, encyclopedic information, photo tags etc).

The **Television Linked To The Web (LinkedTV)** [1] project aims to provide a novel practical approach to Future Networked Media. Within the scope of LinkedTV, triggerable video objects will be connected to additional Web content and information, thus augmenting the features of the audiovisual content with more comprehensive descriptions. Advanced personalisation in networked media platforms, such as LinkedTV is required to efficiently handle and take advantage of the information stemming from the users' digital traces by unobtrusively capturing, understanding and unifying the information originating from the users' transactional behaviour and interaction with peers. In addition, LinkedTV aims to also address the need for contextualising user preferences based on user situations such as time, location and on-platform actions (on-platform actions, dwell time etc), as well as on the physical state of the user (alone, with company, at work etc) and disposition towards the content (mood, attention etc).

Therefore, it is significant for the efficient elicitation of user preferences to have a holistic and dense, though somewhat lightweight vocabulary to classify this information under. To this end, ontologies provide the needed expressivity and conceptual basis. We may assume an ontology forming the vocabulary and reference knowledge base (KB) of the user model. This knowledge base can contain all relevant domain and/or user-specific concepts and their relationships and should provide uniform, compact conceptualisations for ambiguous, synonymous and multilingual knowledge. Such an ontology can be used as the backbone for predictive elicitation of user preferences, as well as for targeted content recommendation.

A core ontology aiming to adequately describe knowledge relevant to a user in a heterogeneous hypermedia environment is expected to be rather broad. It needs to cover anything from the high level topic conceptualizations and the vastness of the named entities encompassed across various domains, to dedicated entities and relations pertaining a user's context, emotional and physical situation. On the other hand, efficient handling of this immensity of information requires dense conceptualisations in a highly expressive, formal ontology for it to scale well and maintain the accuracy advantage of logical inference algorithms.

This paper conducts an overview of the opportunities and challenges of different ontologies and vocabularies that are germane at understanding and representing fundamental knowledge in digital media, while addressing methods for aligning such information under a rich and holistic knowledge base. Taking this information into account, it concludes with the envisioned approach for a comprehensive

[1] www.linkedtv.eu; FP7 IST IP, 2011-2015. Twitter @linkedtv

knowledge base suitable for intelligent personalisation in a networked media environment such as LinkedTV.

II. Vocabularies of Linked Data

The Linked Open Data [2] (LOD) initiative attempts to provide structure to the vast mass of information available online. Most current personalisation approaches for networked media environments have been directed towards employing such open linked vocabularies to efficiently describe and expand the diverse and continuously evolving information in digital media content and in extension reflect and address the variety in user preferences.

For instance, the NoTube [1] project's approach for personalised TV content recommendation involved building a user profile, in which interests were identified using categories from the DBPedia [2] concept space.

DBPedia stands out as the most prominent organization and concentration of knowledge in the current literature in LOD datasets. It is a shallow ontology that interlinks and semantically structures cross-domain information from Wikipedia[3] and constitutes 359 classes described by 1775 properties (as of 8/2012) and ~2.35M instances. It is released in a variety of languages, allowing for cross-language alignment. However the broadness of this information restrain DBPedia to relatively low expressivity, corresponding to the $\mathcal{ALF}$(D) complexity of Description Logics (DLs), namely entailing subsumption between classes and property restrictions [3].

Another important vocabulary prominent for interlinking and categorizing cross-domain entities is Freebase [4]. Freebase is a public collection of community-contributed interlinked data, or as the community itself describes it "an entity graph of people, places and things". The Freebase ontologies are user-generated and edited, consisting of semi-structured information in the form of folksonomies. It was recently employed by the Google Knowledge Graph [5] to expand Google search results about such entities with related information.

While the list of general-purpose semantic LOD vocabularies of more or less the same expressivity that can encompass knowledge useful for interpreting information about a user based on consumed content has not ended, we wrap up this overview with YAGO [6]. YAGO unifies WordNet [7] with Wikipedia, thus enhancing the semantic relations between entities and individuals of Wikipedia with more descriptive properties. It additionally offers correspondences between the entities and their lexical description (term) while taking into account synonymy and term ambiguity, thus allowing for advanced content classification. However, it is easily understandable that while such a vocabulary adds to the semantics of Wikipedia information, it also adds to the complexity of Wikipedia-based knowledge.

Refraining for enumerating them, we must also mention that the LOD cloud encompasses many interconnected datasets of domain specific knowledge, a collection of which can be found in [8], describing for instance detailed geographical information (Geonames), music relevant semantics (MusicBrainz), etc. These KBs may offer richer information or/and deeper semantics on many aspects important to represent a user's preferences and context, as well as mappings to more general upper knowledge bases and other ontologies in the cloud.

A. *Why use LOD vocabularies?*

First and foremost, the wealth of information is the most significant advantage of the LOD cloud. In a vastly heterogeneous and broad environment such as networked media, the LOD datasets offer structure over the magnitude of data. This structure and information abundance is additionally augmented though the interconnectivity between different datasets within the LOD cloud.

Moreover, the knowledge encompassed in LOD datasets is not static and does not require manual contribution from experts. Evolving knowledge is constantly updated, mostly through community contributed metadata. This process is further alleviated through the conformity of the knowledge bases to widely accepted and used standards (e.g. Dublin Core[4], SKOS[5], SIOC[6]).

B. *So why aren't LOD vocabularies enough for intelligent personalisation and contextualisation?*

At the outset, the expressivity of LOD datasets is rather low. They provide structure and semantics to a large amount of entities in a shallow structure [22]. For example DBPedia only deals with concepts, instances and generic relations connecting them. No complex information is available conveying the distinct axioms and specific relations between concepts that adequately describe the semantics prominent to a user or regarding the user's context across domains. Furthermore, the extremely high dimensionality of the reference knowledge raises serious management issues, especially with regard to safeguarding sensitive user information.

1) *Lack of user-related semantics:*

The majority, if not entirety, of LOD vocabularies and ontologies comprise of shallow ontologies of class-property-instance triples consisting of hierarchical and non-hierarchical relations (at best) or frequently of mere taxonomies or meronymies. Thus, their expressivity and granularity is very low, which is only natural since they describe a vast amount of generic conceptualisations about the world that might concern different applications of

---

[2] http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
[3] http://www.wikipedia.org/
[4] http://dublincore.org/
[5] http://www.w3.org/2004/02/skos/
[6] http://sioc-project.org/

different purposes. Consequently, there are no semantic descriptions in such ontologies for people-related facts that discern from general world perceptions. However a personalisation system should be able to comprehensibly understand how people (the users) generally perceive and interact with the world.

For example, let's assume the case where a user's preferences include football and a particular football team, let's say Manchester United - "and" here denoting the last item of a list, not a constructor indicating conjunction between items. It is perceivable that an inferencing engine that relies on LOD knowledge bases to match available content to a user's profile would not derive any inconsistencies in recommending a video about Chelsea, in which Manchester United does not appear, given a KB that relates the "Chelsea" instance to football. However, it is a general truth that if a user is a fan of a particular team and not casually interested in the sport (that being determined by a preference learning mechanism) he would not be interested in recommendations about other teams-opponents, only because chances are that he would frequently view content containing the opposite team in relation to matches or events that his team was involved in. Conversely, an axiom in the reference ontology conveying such a general user-relevant condition (e.g. disjointness between all teams in the same national league) would permit a reasoner to detect the arising conflict.

2) Privacy and scalability:

While the plethora of information in linked open data ontologies and vocabularies is most valuable for efficiently understanding content, they are still hampered by the immense volume of data, as argued in [9]. The problem for personalised services in particular is twofold: a) Recommendation services are server-bound even in resource-rich devices. The background knowledge required for making inference between the user profiles and available content is too large to be handled outside of a server at any instance. This renders constant client-server communication obligatory, thus giving rise to user privacy compromise problems. b) The volume of data and complexity in the KBs itself and the additional information stemming from mappings across other vocabularies in the Linked Data cloud can prove to be unmanageable for intelligent inferencing services, such as reasoners, to handle.

III. FORMAL ONTOLOGIES FOR THE DIGITAL MEDIA DOMAIN

General upper ontologies such as SUMO[7], DOLCE[8], PROTON[9] and BFO[10] offer rich expressivity and can be used in support of middle or domain ontologies for describing the fundamental liaisons between the various cross-domain data encompassed in digital media. This paper will not delve into details about such ontologies as they are well-known in the semantic web community and are too generic (consisting of highly abstract conceptualisations) and voluminous to be used per se for meaningful description of user models or for supporting meaningful inferencing. They can however serve as the pillar for defining the semantics for a structured, formal ontology. Most importantly they can provide the means to align different middle or domain-specific ontologies and other vocabularies under a knowledge base of an expressivity appropriate for modelling and handling user preferences.

In the field of broadcasting in particular, the RDF[11]-based BBC Programmes Ontology [10] provides a descriptive vocabulary for TV programmes, describing concepts such as broadcasting events, brands, episodes etc. It is a lightweight ontology recording the broad spectrum of rather abstract broadcasting-relevant aspects. It notably provides semantics for media-related temporal concepts and objects, thus rendering it a considerable basis for an upper level vocabulary in personalised TV environments. On top of that, the ontology is based on the FOAF vocabulary (described further on), thus providing associated mappings that allow for handling information in a user's social media activity.

For the purpose of personalised TV-content recommendations, an expressive OWL[12] ontology was developed within the AVATAR system [11]. This ontology consisted of a full hierarchy based on three levels of granularity of program-related categories and subcategories used to classify TV programs. It also comprised the properties that interrelate them and different important entities within the context (actors, directors, places, scriptwriters, etc.) [12]. Although the ontology appears suitable for re-using as a core for a TV-related personalisation platform based on its design principles and conceptual description, the ontology itself appears to be no longer available for properly analysing its semantics and value of use, and most importantly for reuse as such.

In the interim we must point out that an overview of ontologies depicting multimedia (visual) attributes is out of the scope of this paper as they are deemed irrelevant to the requirement of semantically describing user preferences and behaviour but rather facilitate the annotation of content.

IV. USER-SPECIFIC ONTOLOGIES

In addition to ontologies and vocabularies expressing knowledge over general world domains, conceptualisations and semantics relevant to users of personalised platforms have also been proposed (e.g. skills, contextual situations, mood etc). The Cognitive Characteristics Ontology [13] for instance provides a vocabulary for describing cognitive patterns for users within certain contexts, their temporal dynamics and their origins, on/for the Semantic Web.

The authors of [14] use their OCUM (Ontological Cognitive User Model) as an upper ontology that "encompasses the core human factors' elements for

7 http://www.ontologyportal.org/
8 http://www.loa.istc.cnr.it/DOLCE.html
9 http://proton.semanticweb.org/
10 http://www.ifomis.org/bfo
11 http://www.w3.org/RDF/
12 http://www.w3.org/TR/owl-features/

hypertext computer-mediated systems" and can be reused to provide enhanced user-centric mappings for personalisation systems.

The NAZOU project [15] also specifies a dedicated user model ontology. In this approach, an ontology-based user model defines concepts representing user characteristics and identifies relationships between individual characteristics connected to a domain ontology. The user ontology in NAZOU is composed of two standalone (OWL) ontologies, which separate domain-dependent and general characteristics: the Generic user ontology which defines general user characteristics and the Job offer user ontology which defines characteristics bound to the domain of job offers represented by a domain ontology.

The GUMO (General user model ontology) [16] and its descendants/hybrids records general upper concepts in combination with characteristic attributes of a user. The ontology is very general and broad, consisting of hierarchical/categorical relationships only. However, its user-related subsets efficiently depict top-level, user-relevant concepts such as user state and actions, e.g. personality (agreeable, timid etc), facial expression (happy, scared etc), motion (sitting, standing, walking), knowledge (e.g. computer skills), location (coordinates, spatial location, virtual location), social environment (friends, family etc). Such concepts are especially useful to semantically describe contextual semantics of sensor extracted information, i.e. the reactional behaviour of a user to the TV content.

The LOD initiative has not refrained from addressing the needs for a vocabulary pertaining user information. With an interest in representing peer-related information in social media environments, the well-known FOAF[13] vocabulary was introduced. FOAF encompasses the most prominent semantics for representing relationships between people. The Weighted Interests (WI[14]) vocabulary introduced in NoTube and derived from FOAF is used to express user interests and disinterests within context along with their temporal dynamics. The vocabulary contains a small set of concepts and properties that define relevant semantics. Both of these vocabularies consist of simple RDF-based triples. WI includes support for annotation representation in RDFa.

V. ALIGNING ONTOLOGIES AND LOD DATASETS

Although the domain of networked media is rather broad, relevant knowledge can still be limited under a consistent subset of information that mean something to the user. To achieve this task, two comprehensive requirements arise: a) unifying relevant schemata under a core ontology of expressive formal semantics appropriate for comprehensive inferencing (these schemata encompassing knowledge about user characteristics and about the domains pertinent to digital media at an appropriate granularity) and b) unifying individuals across different vocabularies under a single conceptualisation and mapping them to the uniform schema.

The latter can be achieved through tools focused on aligning instances and concepts within LOD ontologies. The NERD ontology [17] for instance provides a frame for mapping named entities (NEs) described across several multi-discipline vocabularies (Alchemy, DBpedia Spotlight, Extractiv, OpenCalais, Zemanta) on top of the NER [17] named entity extraction, classification and disambiguation tool. This ontology can be used for extracting NEs within textual manifestations of digital media (audio transcripts, articles etc) and supports semantic annotation of media content with coherent and interlinked instances belonging to popular LOD schemata, thus substantially enhancing semantic interpretation of diverse user-consumed content.

The former involves the non-exhaustive subject of ontology alignment and is mostly dependent on discovering relations from the lexical manifestations of ontological concepts. Some recent approaches include AROMA [19] and S-Match [20]. AROMA is an association rule mining-based ontology matcher. S-Match discovers semantic correspondences between concepts by computing the semantic information implicitly or explicitly codified in class labels.

BLOOMS [18] is a bootstrapping ontology matching tool for aligning information between datasets/ontologies in the LOD cloud at the schema level. The system generates links between class hierarchies (taxonomies), with subclass and equivalent relations, using Wikipedia (articles and categories) and DBPedia. The advantage of this tool is that it not only maps LOD datasets but can also use and map to upper level ontologies such as SUMO and DOLCE.

Addressing the need identified in [9] to align the central LOD vocabularies under a uniform upper ontology, the authors of [22] present an approach to facilitate access of LOD data through a single ontology (namely PROTON). Their approach includes using "ontology expressions" to make matching rules for concepts and relations between PROTON and selected LOD datasets and adding new instances to PROTON through inferencing and concludes with an extended version of PROTON, with added classes and properties useful for uniformly accessing the LOD datasets.

This approach instigated a more sophisticated extension of BLOOMS, namely BLOOMS+ [23] that maps LOD datasets to PROTON, which takes into consideration contextual information to support the decision of if and which concepts/relations are going to be aligned.

VI. THE CHOICE OF AN ONTOLOGY FOR LINKEDTV

The requirements on deciding over the most suitable semantic knowledge for the users of LinkedTV includes determining the level of granularity, the semantic precision, and the expressivity of the ontology with regard to appropriate inferential services, such as logical reasoning. Another important issue to be elucidated is the content of the ontology. Every concept, relation and rule that may have

[13] http://www.foaf-project.org/

[14] http://xmlns.notu.be/wi/

meaning to a user in the scope of LinkedTV should be represented.

In order to keep user models and their usage in the information system lightweight and manageable we identify the need to build and maintain an ontological knowledge base that a) can support meaningful representation of world semantics under *a single uniform vocabulary*, b) will encompass the minimum possible concept space among the ample information in the networked media domain with regards to addressing user needs and c) will be able to sustain abstract user-specific conceptualisations such as user status, skill and situation. Evidently, various structured information (different sources, formats) should be integrated, mediated and exchanged within this ontology.

To this end, we will consider composing a core upper ontology based on existing broadcasting-related and user-related ontologies by re-using and extending existing upper/middle ontologies (such as the BBC programmes ontology and GUMO subsets). This ontology should bring together context-related, sensor data-related, user situation-related and interest-related semantic information. In extension, we will consider the appropriate alignment tool(s) for bringing together the relevant ontologies under a uniform, compact knowledge base and to produce mappings between the core ontology and different LOD vocabularies.

It should be noted that the multimedia content in LinkedTV will be annotated based on several LOD vocabularies and other general purpose schemata, including the use of DBPedia and the BBC Programmes ontology. Annotations will consist of concepts and individuals stemming from the videos' metadata, visual and audio analysis and will be extended with additional information from the Web. Mappings retrieval can be further supported by capitalizing on LinkedTV's envisioned Linked Media Layer. The Linked Media Layer will be interlinking fragments of multimedia content and annotation schemata and will comprise a core tool for expanded multimedia annotation in LinkedTV.

In the interest of further conveying more expressive knowledge pertinent to the user, we will further consider building upon the core ontology more granular domain-specific ontologies that will correspond to the LinkedTV scenarios. In order to retrieve all the comprehensive semantic information to be incorporated in these ontologies we can automatically analyse the domains by identifying relevant DBPedia subsets and their interrelations and using them to retrieve the conceptual representations of the domains in existing ontologies.

A. Minimizing a multidimensional, multilingual concept space

The alignment and mapping detection process will also aim at reducing the reference concept space into a more dense and meaningful vocabulary by integrating semantically equivalent information from different vocabularies under a single conceptualisation with multilingual support.

While media annotations can be expressed in a common, language-agnostic vocabulary, the same doesn't hold for text (subtitles, articles, social networks activity) and audio content. Target users of the LinkedTV platform do not share the same natural language since the project's envisioned scenarios involve content in several different languages, namely English, German, Dutch and possibly also French. In effect, the choice of an ontology for LinkedTV may arguably be based on a common natural language (e.g. English) with multilingual support across LinkedTV scenarios in order to facilitate re-use and exchange of knowledge.

To this end, we propose to enrich an English-based core ontology with adequate information to align cross-vocabulary and multilingual information under a single, unified conceptualisation. This can be achieved by indexing corresponding semantic information about the concepts and properties, derived by aligning the ontology with different LOD vocabularies and schemata, as well as with the German, French and Dutch DBPedia, under a single representation. In essence, all conceptual information within these vocabularies, semantically describing the same concept of the reference ontology in all the languages involved in the LinkedTV scenarios, will comprise descriptive features characterizing each concept and property. These features may be encompassed in the description or labels of each concept in the core ontology.

Feature vectors describing the meaning of a concept have been introduced before in [21]. In essence, these descriptive features can be used to classify the content based on its annotation or information stemming from text mining under a uniform ontological vocabulary of reduced dimensionality. This approach can enable classification of synonymous and multilingual content information under a single conceptualization in order to avoid redundant steps in the inferencing process, such as Football ≡ Fußball or Church ≡ Temple.

B. Knowledge pulling and ontology learning

Evidently, the volume of background knowledge in a multidiscipline domain such as digital media still hampers real-time inferencing through reasoning mechanisms due to the unnecessarily high complexity it introduces, i.e. a large terminological box that has no relation to the domain or the situation of the user in a given session, however light and compact the reference knowledge might be.

Therefore, we will further explore and foster manifold user knowledge acquisition and adaption techniques, such as pulling subsets of (domain) knowledge based on user context to reduce the dimensionality of background knowledge, in order to facilitate intelligent and privacy-preserving concept and content recommendation. The approach aims to understand the extended context during a concrete viewing/browsing situation of the user based on content annotation and other contextual factors (e.g. geotemporal information) in order to pull the appropriate sub-ontology that is needed for inferencing at a given condition without having to transmit back to the server specific information about the user.

Moreover, we cannot assume that domain knowledge remains static as the world changes and new trends emerge and evolve. To this end, in extension to freely available DBPedia and other LOD datasets updates, we will consider dynamically adapting the user-pertinent ontology to new data and evolving it over time by considering strategies to analyse aggregated user generated information to discover new knowledge about the domains. In addition, we will investigate methods to expand the platform's capabilities by extending reference knowledge through automatically learning group-specific knowledge based on aggregated user information (e.g. through user clustering) and adapting new information to the initial knowledge, e.g. via learning persistent association rules within clusters of similar users.

VII. Conclusions

This paper argues on the opportunities and challenges in selecting an efficient knowledge base as the background for semantic, context-aware personalisation in a broad and multidiscipline environment such as networked media. The argumentation is supported by exemplifying trends and existing solutions in current knowledge bases and a brisk overview of methodologies aiming to bring together the rich semi-structured linked open data world and the deep but restricted world of formal ontologies. Finally it concludes with an envisioned solution for an expressive KB for the personalisation task in the LinkedTV EU project.

Acknowledgment

This work was supported by the European Commission under Contract FP7-287911 LinkedTV.

References

[1] http://www.notube.tv. Last accessed on 15th March 2012.

[2] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. "DBPedia: a nucleus for a web of open data". Proc. 6th international The semantic web (ISWC'07/ASWC'07), 2007, Springer-Verlag, Berlin, Heidelberg, pp. 722-735..

[3] J. Völker and M. Niepert. "Statistical schema induction". Proc. 8th extended semantic web conference on The semantic web: research and applications - Volume Part I (ESWC'11), 2011, Vol. Part I. Springer-Verlag, Berlin, Heidelberg, pp. 124-138.

[4] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor. "Freebase: a collaboratively created graph database for structuring human knowledge". Proc. 2008 ACM SIGMOD international conference on Management of data (SIGMOD '08), 2008. ACM, New York, NY, USA, pp. 1247-1250.

[5] "Introducing the Knowledge Graph: things, not strings". http://googleblog.blogspot.gr/2012/05/introducing-knowledge-graph-things-not.html. Last accessed on 14th June 2012.

[6] F. M. Suchanek, G. Kasneci, and G. Weikum. "Yago: a core of semantic knowledge". Proc. 16th international conference on World Wide Web (WWW '07) 2007, ACM, New York, NY, USA, pp. 697-706.

[7] G. A. Miller. "WordNet: A Lexical Database for English". Communications of the Association for Computing Machinery, 38(11), 1995, pp. 39-41.

[8] http://www4.wiwiss.fu-berlin.de/lodcloud/ckan/validator/. Last accessed on 10th June 2012.

[9] P. Jain, P. Hitzler, P. Z. Yeh, K. Verma, and A. P. Sheth. "Linked Data is Merely More Data". In AAAI Spring Symposium 'Linked Data Meets Artificial Intelligence', 2010.

[10] Y. Raimond, P. Sinclair, N. J. Humfrey, and M. Smethurst. "BBC Programmes Ontology". 2009 http://www.bbc.co.uk/ontologies/programmes/2009-09-07.shtml. . Last accessed on 14th June 2012.

[11] Y. Blanco, J. J. Pazos, A. Gil, M. Ramos, A. Fernández, R. P. Díaz, M. López, and B. Barragáns. "AVATAR: an approach based on semantic reasoning to recommend personalized TV programs". Proc. Special interest tracks and posters of the 14th international conference on World Wide Web (WWW '05), 2005. ACM, New York, NY, USA, pp. 1078-1079.

[12] Y. Blanco, J. J. Pazos, A. Gil, M. Ramos, A. Fernández, R. P. Díaz, M. López, and J. Garcia. "Exploiting synergies between semantic reasoning and personalization strategies in intelligent recommender systems: A case study". Journal of Systems and Software, 81, 12, December 2008, pp. 2371-2385.

[13] http://smiy.sourceforge.net/cco/spec/cognitivecharacteristics.html. . Last accessed on 15th June 2012.

[14] M. Belk, P. Germanakos, N. Tsianos, Z. Lekkas, C. Mourlas, and G. Samaras. "Adapting Generic Web Structures with Semantic Web Technologies: A Cognitive Approach". Proc. 4th International Workshop on Personalised Access, Profile Management, and Context Awareness in Databases (PersDB 2010), 2010.

[15] http://nazou.fiit.stuba.sk/. . Last accessed on 23rd April 2012.

[16] D. Heckmann, B. Brandherm, M. Schmitz, T. Schwartz, B. M. von Wilamowitz-Moellendorf, B.M.. "Gumo - the general user model ontology". In User Modeling, Edinburgh, Scotland, 2005, pp. 428-432.

[17] G. Rizzo and R. Troncy. "Nerd: evaluating named entity recognition tools in the web of data". In ISWC'11, Workshop on Web Scale Knowledge Extraction (WEKEX'11), October 23-27, 2011, Bonn, Germany.

[18] P. Jain, P. Hitzler, A. P. Sheth, K. Verma, and P. Z. Yeh. "Ontology alignment for linked open data". Proc. 9th international semantic web conference on The semantic web - Volume Part I (ISWC'10), 2010, Vol. Part I. Springer-Verlag, Berlin, Heidelberg, pp. 402-417.

[19] J. David, F. Guillet, and H. Briand. "Matching directories and OWL ontologies with AROMA". Proc. 15th ACM International Conference on Information and Knowledge Management (CIKM '06), New York, NY, USA, ACM, 2006, pp. 830–831.

[20] F. Giunchiglia, P. Shvaiko, M. Yatskevich. "S-Match: an algorithm and an implementation of semantic matching". Semantic Interoperability and Integration. Number 04391 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2005.

[21] D. Tsatsou, F. Menemenis, I. Kompatsiaris, P. C. Davis. "A semantic framework for personalised ad recommendation based on advanced textual analysis". ACM International conference in recommender systems (RecSys09), New York, USA, 2009, pp. 217-220.

[22] M. Damova, A. Kiryakov, K. Simov, S. Petrov. "Mapping the Central LOD Ontologies to PROTON Upper-Level Ontology". Proc. 5th International Workshop on Ontology Matching at ISWC 2010 Shanghai, China, November 2010.

[23] P. Jain, P. Z. Yeh, K. Verma, R. G. Vasquez, M. Damova, P. Hitzler, and A. P. Sheth. "Contextual ontology alignment of LOD with an upper ontology: a case study with proton". Proc. 8th extended semantic web conference on The semantic web, Volume Part I (ESWC'11), 2011, Vol. Part I. Springer-Verlag, Berlin, Heidelberg, pp. 80-92.