Qualifying Audiovisual Searching Results with Ontologies and Semantic Algorithms

Luiz Rolim, Ismael Ávila, Alexandre Osorio
Service Technologies Department
CPqD R&D Foundation
Campinas, São Paulo, Brazil
[lrolim, avila_an, aosorio]@cpqd.com.br

Abstract— **Multimedia capabilities in end-user terminals, improvements on audiovisual (AV) encoding technologies and the ease of handling AV contents in the Internet have all contributed to the growing use of this media on the Web. Nowadays, searching for videos has become as common as searching for documents, news, web pages or other types of media, being the amount of non-relevant results returned as response to user's queries a common problem posed by the majority of searching engines. Among the myriad of approaches under consideration for qualifying the results of the queries, the usage of semantic technologies is one of the most attracting techniques. In this work, we present how an OWL ontology of subjects, or themes, can improve the efficiency of searching engines through the adoption of semantic algorithms operating over selected contents metadata descriptors based on DCMI and MPEG-7 standards. The main goal is to develop an algorithm that explores the semantic relationships of the supporting ontology and allow searching engines to return results that more appropriately match the actual interest of the end-users.**

Keywords-ontology; metadata; audiovisual; content searching; semantic algorithm.

I. INTRODUCTION

The advent of the Internet, of the high speed networks and multimedia capabilities on personal devices has lead to a widespread production of audiovisual (AV) contents by companies, institutions and end-users. In addition to text and static images, videos are employed pervasively as a way of documenting facts and situations of the everyday life, as well as a tool to transmit messages, express points of view or for artistic purposes. Today's Web provides various options for video search or video sharing, such as YouTube, Mubi and Vimeo. A common aspect encountered in these services is that the searching options can take advantage of the metadata description which generally accompanies the AV contents. Since the metadata contain specific information regarding the production, context, protagonists, themes and other aspects of content, specialized searching engines can provide advanced searching and presentation options when compared to traditional searching engines based on generic text comparisons. However, even on the AV specialized searching tools mentioned above, the adopted cataloguing and searching models are generally based on plain text descriptions and semantic-less keywords or categories, returning results that are non-qualified from a semantic perspective. As a consequence, in large repositories, many of the returned items may not meet the actual interests of the user.

Concurrently with the on-going efforts on improving the efficiency of existing searching engines, semantic-based technologies could play an important role in video browsing and cataloguing as described in [1] and [2], studies based on structured sets of metadata descriptors, such as MPEG-7 [3] and Dublin Core Metadata Initiative - DCMI [4], and on a supporting ontology. The work presented here-in adopts a similar approach that aims at developing algorithms capable of exploring the semantics of video content metadata. In this paper, we present a proposal for a video searching semantic algorithm and the structure of the supporting ontology.

Due to its inherent simplicity and its widespread use as a resource description scheme, we selected DCMI as the metadata standard for the overall description of the contents. For the description of specific AV elements, such as video-segments, the choice was for MPEG-7, which provides a comprehensive descriptor set to represent specific AV content structures. Together, the two descriptor sets form an application profile similar to the one described in [5]. Semantic capabilities are provided through a supporting ontology containing structured subject terms which are made available to cataloguing and searching tools.

This research is an activity of the Experimental TV project, a part of the GIGA R&D (Research & Development) program, consisting of a high speed optical network and associated services, currently being developed by the CPqD Foundation (www.cpqd.com.br), a Brazilian R&D Center. The research comprises the elaboration of the supporting OWL (Web Ontology Language) [6] ontology, the semantic cataloguing and searching tools, and a field experiment with community TVs and independent video producers [7]. The goal is to evaluate how semantic enabled searching engines can provide more qualified results to end users and make the searching process more effective. At the same time, it will be observed the influence of semantic enabled cataloguing and searching tools in promoting the sharing of video contents and the participation of end-users in an established video description collaboration process.

The remainder of the paper is organized as follows. Section II discusses related works on the area of semantic video searching. In Section III, the cataloguing process for AV contents is presented together with the structure of the supporting ontology. Section IV describes the semantic algorithm proposed, exemplified with a use case in Section V. The paper ends with conclusions and a discussion on further work.

 ISBN: 978-1-61208-175-5

II. RELATED WORK

To correctly situate the solution proposed herein in the area of semantically enabled search [14] (SES) it is first necessary to review some current approaches in SES. An SES solution can be aimed to solve different types of search, depending on how narrow is the initial target defined by the user, or even depending on whether there is a clear target at all. In the so-called "navigational searches", the users know precisely what content they are looking for, and the process of finding it is navigating (browsing) to that particular document. In the "exploratory searches", on the other hand, the users have no precise idea of what will be the outcome of the search, probably because they are not familiar with the topic being searched, and their interests can change as they are presented to new search results. In between these two extremes one can distinguish "research searches" [13], where the users have some topics in mind, but no particular document.

According to [16], one can consider exploratory searches as a specialization of information exploration, and interface features such as dynamic queries can help users to see the immediate result of their decisions. To evaluate such systems it is necessary, for instance, to compare the time spent in finding and selecting the information. The solution for exploratory video searching described in [14] combines results from a specific video index with complementary data from DBpedia, which is an initiative to semantically structure information from Wikipedia and dispose the results on the Web. In order to determine, for the query string, a list of related entities, a set of heuristics are applied to the entities in DBpedia. The objective is to determine the relevance of one property based on the frequency it occurs on instances of a category or type in DBpedia. The resources suggested to the user are the ones connected to the highest frequent properties and that are available in the video index. Another approach to the problem is to conceive search engines totally based on the Semantic Web, such as the one described in [15].

In this work, our expectation is to contribute to the audiovisual searching area with the conception of semantic algorithms supported by an OWL ontology containing the knowledge to be applied to the searching process. Overall, we expect to explore functionalities that provide benefits in all the search categories described above.

III. METADATA AND ONTOLOGY

The infrastructure for the semantic AV content searching engine consists of a database containing the metadata descriptors and the supporting ontology. The AV contents files may be stored in one or more repositories and the access to the content is ruled by property rights defined by the owner(s). Searching results comprise an URL providing either direct access to the content or specific instructions for accessing it.

As indicated previously, the semantic capabilities will be implemented around the topics related to the AV content, treated here as a whole, complete entity. The corresponding field in the DCMI set is the *subject* descriptor, a multi-valued element that stores the relevant topics associated to the content. The role of the ontology will be to function as a controlled vocabulary for the terms that potentially can be assigned to the DCMI *subject* descriptor and capture the semantics relationships of all defined terms. The general architecture is depicted in the figure below:

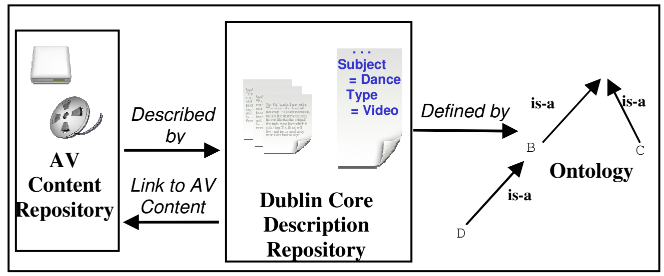

Figure 1. Semantic Searching Architecture

As shown in the figure, applicable values to the *subject* field are defined by the supporting ontology, which also provides the relationships to be explored by the semantic algorithms.

AV content descriptions are inserted into the repository via a cataloguing tool, not depicted in the figure. In general, the description process is initiated by the producer or author of the content. The process consists of textual annotations and requires that all project defined mandatory DCMI fields be filled out in order to allow the retrieval of the content and support the semantic based queries. Semantically enabled fields, such as *subject*, are manually annotated with the aid of specialized end-user interfaces driven by the ontology terms and relationships.

As users access the content descriptions, the catalogued information is improved by means of an established collaboration process. The ontology also evolves collaboratively by means of a cataloguing tool that allows users to suggest new terms and relationships for inclusion in the controlled vocabulary. This contribution will then be analyzed, refined and eventually incorporated into an updated ontology. As a result, the same users that participate in the collaborative content description process may also participate in the improvement of the supporting ontology.

For the development of semantic searching algorithms, it is necessary the combination of metadata descriptions with the supporting ontology [8][9]. Therefore it is important to point out what is required from the structure of the proposed ontology. For this purpose, we will use as example a Brazilian party named *Fandango* [18], a folkloric event in some coastal states, characterized by music and dances that honor sailors and fishermen.

If one catalogs an AV content about *Fandango*, a natural choice would be to set the *subject* field as *Fandango* in the repository. Thus, a semantic enabled searching engine could easily return this item as a response to a query with one of the keywords *Folklore* or *Dance*, as long as the appropriate relationships are present in the ontology. Besides that, if the keywords *Folklore*, *Dance* and *Sailor* were part of the query, the engine could highlight items catalogued as *Fandango* as the ones with the highest probability of matching the actual

 ISBN: 978-1-61208-175-5

interests of the user. Note that non-semantic engines would need to rely on the presence of these keywords in one or more description fields and perform partial match comparisons to get to similar conclusions, thus making the overall process less efficient and error prone.

The question that arises from this example is how the engine will get to such conclusion if the desired item is marked solely as *Fandango*. This is the point where the ontology makes its contribution by providing relationships that make semantic inferences possible. In summary, the ontology needs to be structured in a way that facilitates the categorization of subjects likely to be associated to AV contents [10] and define object properties that establish the semantic relationships between them. Another aspect to be taken into account is to base the ontology on an already established work in order to ease its acceptance by the users. These are key points for elaborating a stable structure of an ontology which can grow in terms of elements and relationships without requiring continuous updates to the deployed software engines. In this project, we selected the controlled vocabulary of the Brazilian Cinemateca [11], a repository of topics for cataloguing contents from independent producers, as the basis for our ontology. This vocabulary is composed of an extensive list of *subjects* which can have one of the following relationships with other *subjects* of the vocabulary:

Subject_A **isA** type of *Subject_B*
Subject_A **isRelatedTo** *Subject_C*

Mapping these relationships to an OWL ontology is straightforward. While **isA** can be directly mapped to the *class<–subclass* or *class<–individuals* OWL relationships, **isRelatedTo** is mapped to an object property whose domain and range are individuals of the generic class *Subject* or any of its subclasses. Note that for this specific application, **isRelatedTo** is not meant to capture the specific aspects that make two given subjects to be related to each other since this would require the definition of an extensive set of properties far beyond what is necessary to accomplish the goals of this work. In our case, the qualified results are obtained by exploring the generic **isRelatedTo** property that may exist between *subjects* defined in the ontology. By applying these relationships to the *Fandango* example mentioned above, we can draw the relationship diagram depicted in Figure 2. As shown in the figure, *subject* derived classes are represented as non-filled rectangles and correspond to groups or categories of topics. Ontology individuals correspond to specialized topics and are represented as solid-filled ellipses. The relationships **isA** and **isRelatedTo** are represented by the solid and dashed arrows respectively. According to the figure, a semantic searching engine could use the following relationships when processing the queries:

Fandango **isA** *Party*, a *Brazilian_Party* and a *Folklore*
Fandango **isRelatedTo** *Sailors* and *Fishermen*

Once the structure of the ontology is defined, we can turn our attention to the proposed semantic algorithms and an illustrative use-case.

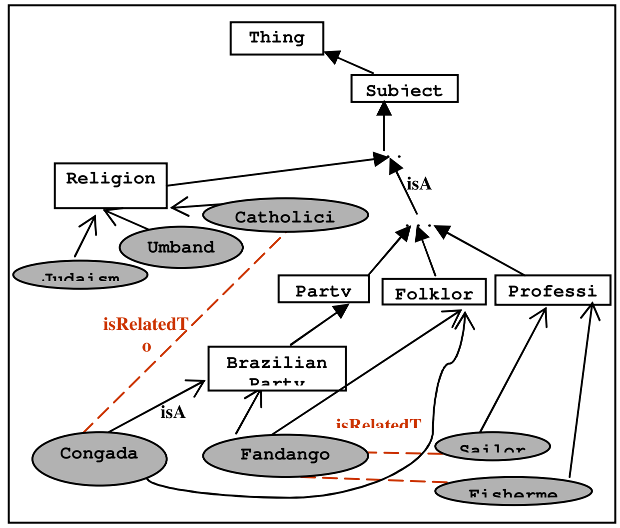

Figure 2. Ontology Structure

IV. SEMANTIC ALGORITHM

In this section we describe the semantic searching algorithm applied in the experiments. The input is the list of query parameters specified by the user for the *subject* descriptor.

Firstly, the algorithm will expand the list of parameters with terms from the ontology having an **isA** relationship with each of the terms entered by the user. This operation expands the list with all sub-classes and individuals members of each parameter entered by the user, thus making the searching process semantically comprehensive. By doing this, any AV content eventually catalogued with one of the specialized terms will also be considered a potential result of the query. The steps of this portion of the algorithm are presented as follows:

(1) ***CTL*** = *Original List of Controlled Terms* – user´s input.
Nc = *number of terms in CTL.*

(2) Let $\boldsymbol{i}$ be an integer varying from 1 to ***Nc.***
For each term $\boldsymbol{T_i}$ in ***CTL***, create the set $\boldsymbol{S_i}$ defined as:
$\boldsymbol{S_i}$ = {***T,*** (all $\boldsymbol{T_i}$ sub-classes),
(all $\boldsymbol{T_i}$ type individuals)}.

(3) Search the AV contents description repository and select as a result to the query items whose subject descriptor contains values present in all $\boldsymbol{S_i}$ sets.

Note that steps (1) through (3) perform a semantically extended logic AND over the input parameters and furnish a comprehensive list of categories to organize the presentation of the results to the user. For simplicity, we omitted the handling of terms not present in the ontology, which would be treated like any other non-semantic field.

These steps of the algorithm allow a searching engine to return an item marked only as *Fandango* or *Congada* [17] as a possible result for a generic query on the keyword *Brazilian-Party*. Conversely, if the query parameter is the keyword *Fandango*, the searching tool could be enhanced to inform the user that he/she might be also interested in AV contents eventually marked as *Brazilian-Party*. All these

 ISBN: 978-1-61208-175-5

inferences derive from the **isA** semantic relationships defined in the ontology, which were captured in each of the S_i sets created at the step 2. Additionally, for presentation purposes the returned items can be grouped by the elements of each S_i set$_i$, offering to the user a friendlier and more organized interface to navigate into the results of the query.

The benefits of the algorithm are even stronger if we consider that the ontology and the description itself can be evolved as part of a collaborative program. As described in [12], new terms and semantic relationships can be added to the ontology over time and become available to the searching engine by loading the updated ontology into the reasoner. As an example, we can take the term *Fandango*, which according to the ontology, is a type of *Brazilian Party*. However, *Fandango* can also be regarded as a *Brazilian Dance*. As a result of the collaboration process, an updated version of the ontology can specify that *Fandango* is also an individual of *Brazilian-Dance*, a sub-class of *Dance*. As the new ontology is reloaded into the reasoner, *Fandango* annotated items will also be returned for a query on the keyword *Brazilian-Dance*. Note that the results of the query are improved without requiring any updates to the metadata description repository.

The next part of the algorithm aims at obtaining results as close as possible to the actual interests of the user by employing the **isRelatedTo** relationship, as described in the following steps:

(4) Let i be an integer varying from 1 to Nc.
For each term T_i, create the set R_i defined as:
$R_i = \{S_i$, (all individuals which are related to T_i)$\}$.

(5) Let ***PITL*** be the list of terms present in all R_i sets
PITL is defined as: $\{R_1 \cap R_2 \cap R_3 \cap \ldots \cap R_n\}$.

(6) Include as a qualified result to the query any AV contents whose subject descriptor contains at least one of the terms present in ***PITL***.

The short list of qualified results, ***PITL***, is represented by the intersection of all the R_i sets, as shown in step (5). If not null, ***PITL*** contains one or more items common to all elements of the semantically extended lists of terms, the R_i sets, built by exploring the **isA** and **isRelatedTo** relationships over the input parameters. The algorithm infers that ***PITL*** contains the terms with the highest probabilities of representing the actual intent of the user when submitted the query.

V. USE CASE

The benefits of the semantic algorithm can be better visualized through a practical example in which a description repository contains a couple of instances referencing the term *Fandango* in their *subjects* descriptor fields. Then, let's consider that the user submits a query with the following terms: *Party* and *Folklore*. The execution of steps 1 and 2 will lead to the following:

(1) ***CTL** = {Party, Folklore}*
***Nc** =2*

(2) According to CTL in step 1:
T_1 = *Party*
T_2 = *Folklore*
Now the S_i sets are calculated:
S_1 = *{Party, Brazilian Party, Congada, Fandango}*
S_2 = *{Folklore, Congada, Fandango}*

At this point, step (3) will return all AV contents whose subject descriptor are marked as *Fandango* since this term is present in both S_1 and S_2 sets. Note that items marked as *Congada* would also be selected as a result to the query, similarly to items marked as *{Brazilian Party*, *Folklore}*.

For illustrating the second part of the algorithm, a slightly different example will be used. The query parameters are now *Party*, *Folklore* and *Sailor* and the goal of semantic query is to obtain AV contents whose associated topics are some how related to all these three terms. The execution of steps (1) thru (2) would lead to the following S_i sets:

S_1 = *{Party, Brazilian Party, Congada, Fandango}*
S_2 = *{Folklore, Congada, Fandango}*
S_3 = *{Sailor}*

Following with the execution of the algorithm, steps (4) thru (6) would lead to:

(4) R_1 = *{Party, Brazilian Party, Congada, **Fandango**}*
R_2 = *{Folklore, Congada, **Fandango**}*
R_3 = *{Sailor, **Fandango**}*

(5) The intersection of all all R_i sets will lead to:
PITL = {Fandango}

(6) Now, the engine will search the repository and select AV items marked as *Fandango* as results to the query.

The intersection operation over the R_i sets in step (5) leads to a short list of topics with good probability of representing the real interest of the user. Of course, this depends on the accuracy of the relationships defined in the ontology. In this example, *Fandango* is a subject related to *Sailor* and also a type of *Party* and *Folklore*. Consequently, the algorithm infers that AV contents with *subject* descriptors marked as *Fandango* are the ones with the best chances of meeting the expectation of the user.

It is important to note this conclusion is obtained entirely through inferences made over the supporting ontology. Another benefit of this approach is that as new relationships are added to the ontology, the inference power of the engine increases without requiring updates to the searching engine software.

VI. CONCLUSION AND FOLLOW-ON WORK

The expectation of this work is that the conceived algorithms and ontology structures can effectively contribute for improving the efficiency of searching engines and become a valid mechanism for identifying results more likely to represent the actual interests of the user. At the same time, it is also expected that the algorithms become building blocks for the execution of more complex logical operations involving the entire set of AV contents descriptors fields. At the same time, the availability of semantic enabled searching and cataloguing tools can act as a way to promote the sharing of AV contents in a distributive and collaborative

 ISBN: 978-1-61208-175-5

environment in which both the description and the ontology are continuously improved by the users.

However, validating all these ideas in a real environment is a must. So, the next activities of the project comprise a field test with Brazilian community TV stations and independent AV producers, connected to the cataloguing and searching tools through the GIGA high speed network and the Internet. The diversity of end-users and richness of subjects that can be assigned to AV contents form the ideal combination for establishment of a *de facto* collaboration process where the AV content description and the ontology are gradually refined by the participants.

During the evaluations, we will attempt to test how engines enhanced with semantic capabilities can provide higher levels of effectiveness, accuracy and ease of use when compared to traditional, non-semantic, searching tools. One way to evaluate the proposed algorithm is to define a set of search tasks to be executed by the users, in which some videos must be found. One group of users would then execute searches supported by the proposed semantic algorithm and another group would perform searches in a traditional way. A similar approach, with both quantitative and qualitative results, is presented in [14].

It will be also an opportunity to evolve the structure of the proposed ontology, conceive new algorithms and procedures for the cataloguing and searching process. At the same time, implementation and deployment aspects, such as space-time complexity analysis, scalability issues and performance of the ontology queries will deserve special attention from our research team in order to make sure that the benefits observed in the field experience can be replicated in other environments.

ACKNOWLEDGMENT

This work has been sponsored by the Brazilian Ministry of Communications through the FUNTTEL program and funded by the FINEP innovation agency.

REFERENCES

[1] J. Hunter, "Adding Multimedia to the Semantic Web – Building an MPEG-7 Ontology", Proc. of the International Semantic Web Working Symposium (SWWS), July 2001, pp. 261–283.

[2] F. Nack and L. Hardman, "Towards a Syntax for Multimedia Semantics", Technical Report CWI, Amsterdam, The Netherlands, 2002.

[3] J. M. Martínez, R. Koenen, and F. Pereira, "MPEG-7: The Generic Multimedia Content Description Standard", Multimedia, IEEE, vol. 9(2), 2002, pp. 78-87.

[4] Dublin Core Metadata Initiative, "DCMI Metadata Terms", http://dublincore.org/documents/dcmi-terms, <retreived: Sept., 2011>.

[5] J. Hunter, "An Application Profile which combines Dublin Core and MPEG-7 Metadata Terms for Simple Video Description", http://metadata.net/harmony/video_appln_profile.html, <retreived: Sept., 2011>. DSTC- Australia, 2002.

[6] W3C Recommendation, "OWL Web Ontology Language Overview", http://www.w3.org/TR/owl-features/, <retrieved: Sept., 2011>, February, 2004.

[7] L. Rolim, A. Osorio, and I. Ávila, "Collaborative System for Semantic Annotation of Audiovisual Contents - Applications in the Context of Brazilian Independent Culture" (in portuguese), Proc. of SBSC – Brazillian Symposium of Collaborative Systems, IEEE, 2011, pp. 1-4.

[8] A. Th. Schreiber, B. Dubbeldam, J. Wielemaker, and B. Wielinga, "Ontology-Based Photo Annotation", University of Amsterdam, IEEE Intelligent Systems, 2001, pp. 66-74.

[9] L. Hollink, S. Little, and J. Hunter, "Evaluating the Application of Semantic Inferencing Rules to Image Annotation", Proc. of the K-CAP´05 – 3rd International Conference on Knowledge Capture. Alberta, Canada, 2005, pp. 91-98.

[10] A. Isaac and R. Troncy, "Designing and Using an Audio-Visual Description Core Ontology", Institute National de l´Audiovisuel (INA), France, 2004.

[11] Brazilian Cinematheque, "Controlled Vocabulary", http://www.cinemateca.com.br/, <retrieved: Sept., 2011>. Audiovisual Department – Brazilian Ministry of Culture, 2011.

[12] H. Klotz and E. P. Wach, "Collaborative Ontology Building", Seminar Applied Ontology Engineering, STI-Innsbruck – December 2010.

[13] R. Guha, R. McCool, and E. Miller. "Semantic search." In WWW '03: Proc. of the 12th Int. Conf. on World Wide Web, New York, NY, USA, 2003, pp. 700-709.

[14] .J. Waitelonis, H. Sack, J. Hercher, and Z. Kramer. "Semantically enabled exploratory video search," Proceedings of the 3rd International Semantic Search Workshop (SEMSEARCH '10), ACM, New York, USA, 2010, pp. 1-8.

[15] E. Oren, R. Delbru, M. Catasta, R. Cyganiak, H. Stenzhorn, and G. Tummarello, "Sindice.com: a document-oriented lookup index for open linked data," Int. J. Metadata, Semantics and Ontologies. IJMSO, vol. 3(1), 2008, pp. 37-52.

[16] R. W. White, G. Marchionini, and G. Muresan, "Evaluating exploratory search systems: Introduction to special topic issue of information processing and management," Information Processing & Management, vol. 44 (2), 2008, pp. 433-436.

[17] L. Glaura, "The Sounds of Rosário" (in portuguese), UFMG, Belo Horizonte, 2002.

[18] Caburé Cultural Society, "Live Museum of Fandango" (in portuguese), http://www.museuvivodofandango.com.br/, <retrieved: Sept., 2011>.

 ISBN: 978-1-61208-175-5