Journal of Information Assurance and Security.
ISSN 1554-1010 Volume 11 (2016) pp. 322-330

Interlinking video with DBpedia using knowledge-based Word Sense Disambiguation algorithms

Olfa Ben Said[1], Ali Wali[1], Adel M. Alimi[1]
[1]REGIM: Research Groups on Intelligent Machines, University of Sfax,
National Engineering School of Sfax (ENIS),
BP 1173, Sfax, 3038, Tunisia
olfa.bensaid@ieee.org, ali.wali@ieee.org, adel.alimi@ieee.org

***Abstract*—In this paper we present a system of enhancement of videos with linked open data. We aim by this system to link videos with DBpedia dataset, based on semantic information, and so enrich them and make viewer in an intellectual experience when he is watching a video. We detail the proposed architecture, and we focus on the interlinking process, which comprises several steps. The adopted method relies on semantic technologies and search. We propose two different algorithms for the disambiguation, and we illustrated them by some examples and results.**

Keywords- Enhancing video; interlinking; Open Linked Data; Semantic Web; Dbpedia; disambiguation; VSM; TVSM

I. Introduction

Multimedia resources such as video, photo and music, need special requirements for processing and enrichment. The treatment of multimedia content can be based on different features, such as text-based, visual or audio features and others. The trend in annotation, tagging or any Natural Language Processing system (NLP) is the integration of Semantic Web technology. Indeed, the success of linking Textual or Multimedia data to Linked Data Web, either in annotation or interlinking is due to its clarity, unambiguity and the varied and linked information that we can find. For example when a program is tagged with places, it can be identified on a map, or films which can be linked with information about the actor, as well as with other films where the same actor are playing their role similarly an advertising can be enriched with information about the similar products. Another example that has been very useful recently is the annotation by semantic linked data. By using the Linked Data web, we can accommodate a wide range of unforeseen use-cases.

Our idea consists of enhancing videos with extra-contents by using Linked Data as our enrichment resource. We use specifically DBpedia to link it with videos, because of its richest contents. We base on a semantic description of the video, to ensure the interlinking process. In this context, we propose an architecture containing different components. In the next, we describe briefly each component and we focus on the interlinking process.

The article is structured as follows: section two describes related work in enhancing videos and in using linked data as an enrichment resource. In the next section we present the architecture of our system, the problems encountered and the interlinking process. In section four, we detail the disambiguation step using VSM and TVSM algorithms. In the next, we illustrate by some experiments and results. Finally, we finish by conclusions and future work.

II. Related Work

In this section we present the related works about the enhancement video systems and the Linked Open Data cloud, specifically DBpedia dataset.

A. Enhancing videos

In the age of Big Data we notice a huge increase in the amount of added videos. Hours of videos are uploaded for sharing each minute. This important increase requires specific treatment regarding these multimedia resources. Our idea consists of enhancing videos with extra content. The proposal of enhancing multimedia content takes many forms. In this context, an edutainment (education +entertainment) experience was presented in [1]. The proposed system combines TV programs and learning contents, in order to give attractive environment and to engage viewers in intellectual environment. The team has developed an architecture for T-learning interactive services called ATLAS. The system contains two fundamental parts: an intelligent tutoring system called T-MAESTRO; and an authoring tool called ATTOS, which allows teacher to create adaptive courses.

In addition, the evolutionary numbers of existing videos on the web make this last a good source of enrichment. Some works focus on enrichment by hyperlinking videos with other ones, for example the Hyperlinking task in TRECVID 2015 competition[1]. In this task, the candidates are asked to develop a system which can create links between semantically related sections of videos in a vast video database. This task is based on different features to create the link between different videos, noting text-based, audio or visual features or combination methods.

Other form of enhancing multimedia content is to interlink videos with extra links. Due to the large number of

[1] http://www-nlpir.nist.gov/projects/tv2015/

data sources within BBC, in [2] they use Semantic Web technology to integrate data and ensure a better interlinking of existing systems and connections across BBC domains. They work on two BBC services: BBC programs and BBC music, and they describe categorization system which aims to interlink data items of both services with an extra Linked Open Data, in order to make richer relationships between concepts. For the same database; BBC World Service; another work was conceived in [3]. The team conceives a system which aims to process the existing text and audio and automatically annotates programs within the archive with the Linked Open Data identifiers. The result of interlinks are used to improve search and navigation within this archive and expose it to users.

In this field we aim to conceive a system that enrich a video basing on its textual description with extra-content linked to Open Data Cloud. In the next section we give an abstract about the linked Open Data Cloud, and its different uses.

B. Linked Open Data project

The linked Open Data cloud (LOD) plays an increasingly important role in enhancement for several domains. On the other hand, collaborative websites like Wikipedia[2] behaved able to manage and create large public knowledge bases. It has grown into one of the central Knowledge sources and widely used encyclopedia. The marriage between both of LOD and collaborative websites, give us a new born called DBpedia. Its content is directly extracted from Wikipedia and provides this knowledge as Linked Open Data on the Web. Indeed, DBpedia [3] is considered among the largest dataset in the web of data which enables access to many data sources in the LOD cloud. It explores and extracts data from about 3,5 million resources from Wikipedia articles in 125 languages [7]. Otherwise DBpedia is the RDF version of Wikipedia. It is based on entities in the description of their resources. In 2014, DBpedia is described by more than 38.3 million entities in different languages where 4.58 described at English language. Entities are described by Resource Description Framework (RDF) triples. More than 3 billion RDF triples describes the DBpedia 2014 version. DBpedia presents abstract in different languages, links to images, links to external web pages and datsets, links to Wikipedia, Yago and Wordnet categories, and further other data. Moreover, DBpedia provides a rich resource of attributes and relations between the resources, for example, it connects products to their makers, or films to their actors and producer, etc. Entities in DBpedia knowledge base are classified according to DBpedia Ontology, including Persons, places, creative works, organizations, species and diseases. [17] present an overview about the DBpedia information extraction process as presented in Fig. 1.

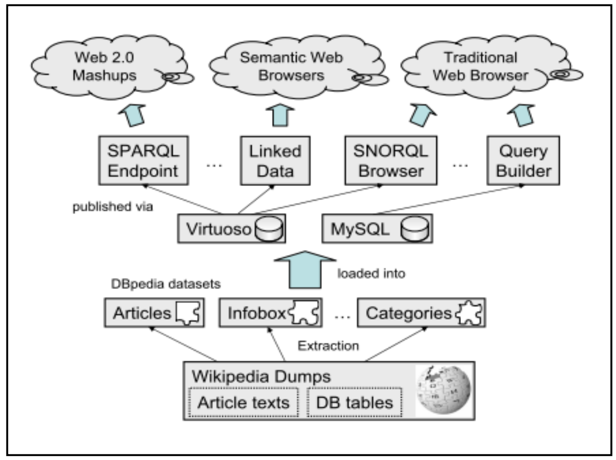

Figure 1. Overview of DBpedia components

Covering a vast subject ranges and being multilingual, DBpedia is useful within several application domains like browsing and knowledge exploration, search and querying, content annotation and tags. One of the main applications adopted by DBpedia Dataset is annotating and tagging texts and other content with semantic information. Many applications are performing the task of annotation using DBpedia entities. *DBpedia Spotlight* [4], is an automatic annotating system that aims to annotate text documents with DBpedia URI. *DBpedia Spotlight* allows users to configure the annotations to their specific needs through DBpedia Ontology and quality measures as topical pertinence, prominence, disambiguation confidence and contextual ambiguity. *Faviki*[4] is a social bookmarking that allows users to keep choose his own tags referring to Dbpedia concepts. In Faviki everybody uses the same names for tags from the world's largest collection of knowledge. Thereby that instead of having users haphazardly entering in tags to describe the links they save, Faviki will suggest tags to be used. *ImageSnippets* [5] is a system that annotates images. It refers to DBpedia for identifying unambiguously entities depicted within an image. In a similar context *xLime Semantic Search Engine* project [6] present an architecture to a search system within media items. An infrastructure is build that enables annotation, store, index and retrieve of media items such as television, news articles and social media posts. Architecture is composed from several components and uses several encyclopedic knowledge bases: DBpedia, Wikidata and Freebase.

DBpedia Mobile [11] offers to users the ability to explore background information about locations and navigate into interlinked dataset such as DBpedia, Revyu, Eurosat, GeoNames and Flickr. It is a browser based area map that gives the user's position and its proximity from DBpedia resources with appropriate labels and icons. Then the user can navigate from DBpedia into other datasets like

[2] http://www.wikipedia.org.
[3] http://wiki.dbpedia.org/about
[4] http://www.faviki.com/
[5] http://www.imagesnippets.com/
[6] http://xlime.eu/

GeoNames. DBpedia Mobile enables users to publish their location, pictures and reviews to the Semantic Web as Flicker.

As mentioned above, *BBC* [2] conceive a system to automatically annotate programs within the archive with Linked Data Web identifiers from DBpedia. This task is preceded by an identifying topics and speakers. Then users can validate and correct this task. In another work described in [3], *BBC* has employed DBpedia URI for tagging their programs including short clip and full episodes. SemanticLIFE [23] is a semantic desktop system which proposes a semantic based way for bridging the gap between semantic life and social life of the user where he share their personal resources. A mashup framework was developed to overcome the steadily increasing problem of information overload of Semantic Desktops and Social networking Site. Video Metadata represents a vital factor for effective representation, management and multimedia content's retrieval. Basing on this assumption, [24] conceives a new system called CAMO which aim to enrich multimedia metadata by integrating LOD. The integration of LOD sources was based on ontology matching and instance linkage techniques.

III. Design Of our System

In this paper we describe how we use linked open technologies and specifically DBpedia, to enhance videos and give it more interactivity with the viewer. In order to realize this system we conceive an architecture based on many sub-systems and resources. In this section we present the proposed architecture and describe the encountered problems.

A. Our proposal

A large amount of videos are published every day into the web taking different topics. Sometimes a video can take different subject. Our idea consists of enriching the main topic of the video with extra information. Information can be text, images or videos that take the form of tags. Fig.2 illustrates the idea of our application as described. We aim, by this work, to engage video viewers in a learning experience, therefore making video an entercation [7] (education/entertainment) environment. Also we aim to interlink videos with extra-content and enrich it and make videos branched with Open linked data. Starting from an input video, we must ensure an automatic tagging system towards information from an encyclopedic resource.

Figure 2. Enhancing video with Open Linked Data

Fig. 3 presents the proposed architecture. In our previous work [8], [9] and [10]we detail more this architecture. Our system consists of three modules. The first consists of taking information about video. Information can be extracted from different sources. It can be described by metadata (such as title, video descriptions, keyword, context, etc), and it can be also taken from audio/visual information and other data. In our case we take only metadata information represented as keywords. The second bloc is for the interlinking between the two resources, Video and extra one. This is the aim of our work. The objective of this task is to interlink video basing on its description with other data on the W3C Linking Open Data community project (in our case we adopt DBpedia). The final bloc is based on filtering selected information, basing on user's profile and preferences.

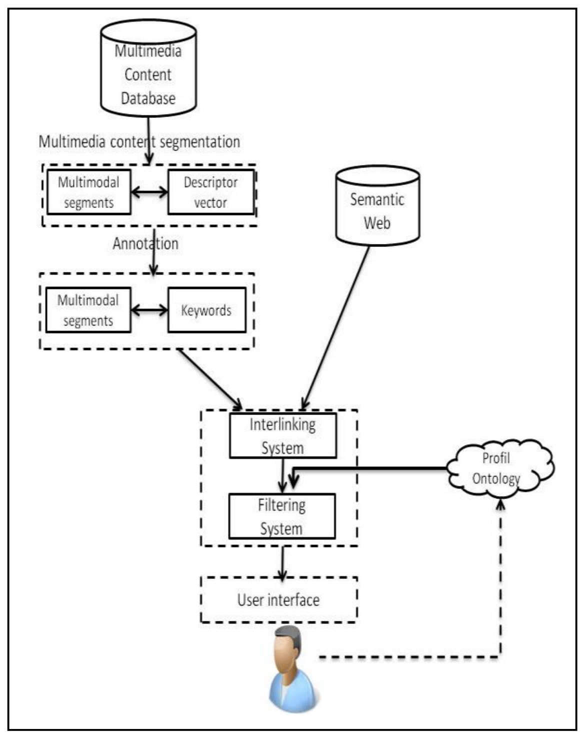

Figure 3. Video interlinking system architecture

The system must contain in addition to the treatment for the linking between the two resources (Videos and DBpedia), a series of steps for data preprocessing. The pretreatment steps (the left top side on Fig. 3) can take

[7] This term is used in [16] also in the web, having the mean of entertainment that educates

different forms depending on information needed for video description. Metadata is presented as a set of keywords describing the main idea of the video contents, thus we based on these metadata to make linking to the extra-resources. Indeed, the interlinking step consists of assigning, to each segment of video, an external resource that resembles it semantically. This step is illustrated in the middle side of the architecture in Fig. 3. The extra content will be extracted from DBpedia resource. This step requires semantic reasoning techniques, as well as techniques for searching the semantic correspondence between the two contents (video description and DBpedia one). We adopt DBpedia as the encyclopedic resource, since it is considered as the largest dataset in the Linked Open Data Cloud and it is extracted from Wikipedia, the biggest multilingual free encyclopedia.

B. *Problem definition*

One of the main problems of annotation, tagging, and any semantic treatment in Natural Language Processing (NLP) is ambiguity. However when a word is ambiguous, it may be used in different contexts to refer to different concepts. So, a word can have more than one sense. For example "*Apple"* can refer to "*a fruit tree*" or to the "*famous company*" or to the "*Apple Band: a British psychedelic rock band"*. The disambiguation process depends strongly to the notion of context. Different methods have been developed in this regard, to resolve this problem.

The goal of our work is to propose an approach to disambiguate metadata related to each video and make linking with the resource that provide a varied and rich information about its context. The largest DBpedia knowledge base is the best resource that can help to ensure these goals. It can be used to help select the candidate source from an ambiguous context and provide information to enrich our video metadata. Otherwise DBpedia will be used as a controlled vocabulary and a data provider.

C. *Interlinking process*

The main challenges in interlinking resources tasks, is ambiguity. Many different methods have been developed in this regard, to resolve this problem.

1) *Context in disambiguiation process*

In the search fields, many methods are proposed for disambiguation process, we distinct in the following the three main categories [14]. Knowledge-based approaches are based on different knowledge sources as dictionaries, thesaurus, lexical databases, etc. This approach includes Knowledge-Based WSD, Lesk algorithm, Semantic Similarity, Selectional Preferences and Heuristic Methods. As to supervised Word Sense Disambiguation (WSD) algorithms, they are based on the use of labeled training data to build a classifier which can determine the right sense for a given word in its context. This last exist in different methods: Decision List, Naïve Bayes, Neural Networks and others. The third approach is called, unsupervised WSD, which generally do not assign meaning to the words instead they discriminate the word meanings based on information, found in unannotated corpora. Main approaches of unsupervised WSD are: Context Clustering; Word Clustering, Co-Occurrence Graph and Spanning Tree based approach [14]. In the following, we focus on the first approach.

In FLOOR [12]- a mechanism for automatic folksonomy tags enrichment- the team propose a disambiguation process. They use Wordnet as dictionary resource to extract their hierarchy of senses and online ontologies and then they calculate the similarity with the senses of all tags. The meaning of the analyzed tag is represented by the most similar one. In the same context, *TagPlus* [13] uses WordNet for the disambiguation of senses of Flickr tags. It performs two step queries: First, the user looks for a tag, then the system returns all the possible WordNet senses that define the tag and the user selects the sense he meant. Finally the system searched the similar Flickr photos having the same tag and its synonyms. Some other work uses DBpedia, for example DBpedia Spotlight [14] is a tool for automatically annotating mentions of DBpedia resources in text. It provides a solution for linking unstructured information sources to the Linked Open Data cloud through DBpedia. The main advantage of this system is its comprehensiveness and flexibility, allowing one to configure it based on quality measures such as prominence, contextual ambiguity, topical pertinence and disambiguation confidence, as well as the DBpedia ontology. As mentioned above, DBpedia is also used in[2][3]. They use DBpedia URIs to tag programs, clips, and radio archives. They developed an algorithm containing two parts: DBpedia label lookup and context-based result disambiguation. Finally [6] proposes a context-based tag disambiguation algorithm that selects the meaning of a tag among a set of candidate DBpedia entries, using a common information Retrieval similarity measure. In a following work [23] authors present the technique used for the disambiguation step to automatically interlink a speech radio archives, noting the Enhanced Topic-based Vector Space Model (eTVSM).

2) *Adopted Interlinking process*

The system follows three steps: Candidate Lookup service, context-based disambiguation and Entity Filtering Service. In the Lookup step, the system builds his space candidate resources according to the searched keyword. In the next one, the system chooses the most relevant candidate according to the context of the keyword. In the Filtering -service, we aim to filter only the main proprieties for the selected resource, according to its appurtenance. Otherwise, this service aims to summarize selected entities and present main information in a concise form. The interlinking process is described in Fig.4. In the next, we focus on the two first steps.

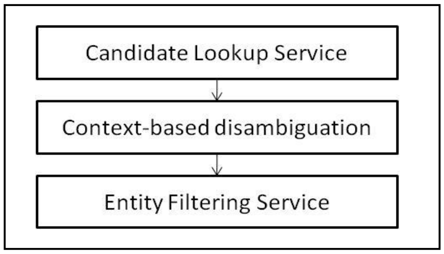

Figure 4. Interlinking system

a) Candidate Lookup Service:

In the first step, we describe the possible candidates of ambiguous word by querying DBpedia with a label matching the keyword as describing in Fig.5. The outputs of this query describe our DBpedia resource candidates, which represent possible senses of the keyword.

b) Context Based disambigution:

In this step, we have to select the proper sense among all candidates according to the context of the keyword. In order to disambiguate the possible matches, the system ranks all candidate resources by using a similarity score for each candidate. To do so, we compute the similarity between the keywords context and the resource description. We leverage the correspondence of DBpedia resources with Wikipedia articles to obtain textual descriptions of each resource. Then we extract the most frequent terms and their frequency, from the Wikipedia pages. We obtain by that, Vectors represent each DBpedia sense and one other representing the context, and then we can calculate the similarity among those vectors using as measure the cosine of the angle they form. Fig.6 presents a general idea about the disambiguation process.

The objective of disambiguation process is to choose among all senses associated to a keyword, the most appropriate one, according to the context of the searched keyword. In this paper we adopt two different algorithms, Vector Space Model (VSM) and Topic-based Vector Space Model (TVSM) and then we evaluate and compare between both. In the next we present a background of the adopted methods.

```
PREFIX   rdfs:   <http://www.w3.org/2000/01/rdf-
schema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dbo: <http://dbpedia.org/ontology/>

SELECT DISTINCT ?label ?dbp ?wiki
 WHERE {
 ?dbp rdfs:label ?label.
 ?dbp foaf:isPrimaryTopicOf ?wiki.
 ?label <bif:contains> "Apple".

FILTER ( langMatches( lang(?label), "EN" )).
MINUS {
?dbp
dbo:wikiPageRedirects|dbo:wikiPageDisambiguates
 ?dis
        }}
```

Figure 5. SPARQL Query over DBpedia

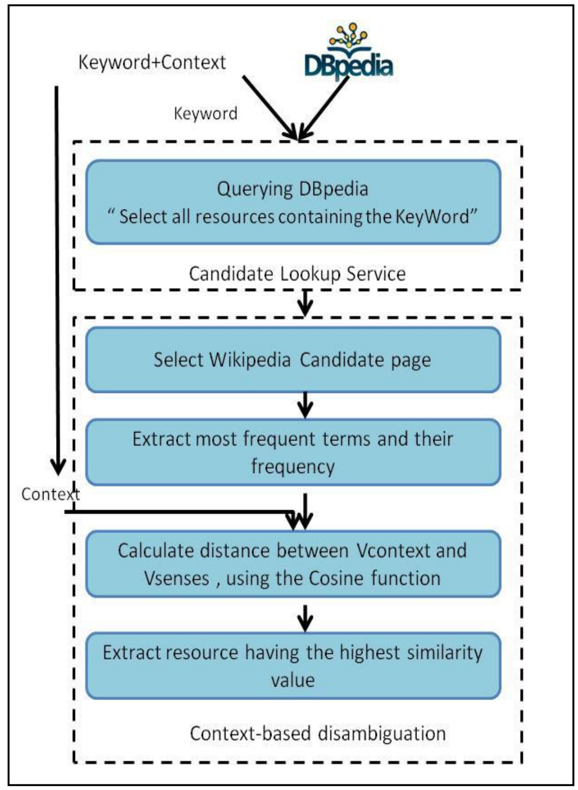

Figure 6. Interlinking Module Diagram

IV. Background

This section introduces common terms and definitions for the adopted information retrieval model.

A. Term Weighting

Each vector document d, is described by a term weighting, which represent the term t_i and its weight $w_{d,ti}$. different factors have been adopted in literature. The simplest term weighting schema, that assumes that term weight is equal to the term occurrences $w_{d,ti}=Occ_{d,ti}$. The classic VSM proposed by [20] has adopted both local and global parameters as the term weight, known as Tf_Idf (Term frequency_Index Term Frequency).

$$W_{d,ti} = tf_{d,ti} * \log\left(\frac{D}{df_{ti}}\right) \qquad (1)$$

With :

- $tf_{d,ti}$ is the number of term t_i in a the document d,
- df_{ti} is the number of documents containing the term t_i,

- D is the number of documents in the candidate space.

B. Vector Space Model (VSM)

In VSM[18][19]we consider a vector per document d where each dimension of the document corresponds to a term i with a weighting $w_{d,ti}$.

$$d = (w_{d,t1}, w_{d,t2}, w_{d,t3}, \dots, w_{d,tn}) \quad (2)$$

$$q = (w_{q,t1}, w_{q,t2}, w_{q,t3}, \dots, w_{q,tn})$$

Similarity between two documents can be performing by cosine similarity measure. $W_{d,ti}$ represents the weight of term t_i in document d. Different form of term weighting can be taken, as stated previously. In our case we consider two the occurrence weight as the term weighting.

$$Sim\ (q,d) = \frac{\sum_j W_{q,j} \quad W_{d,j}}{\sqrt{\sum W_{q,j}^2}\sqrt{\sum W_{d,j}^2}} \quad (3)$$

In this scenario we adopted an approach presented in [6] that disambiguate polysemous tags automatically, after some changes to adapt it with our system. By using a common information retrieval similarity measure, our approach selects the meaning of a keyword from a set of candidate DBpedia entries. The approach depends on DBpedia and Wikipedia information.

C. Topic-based Vector Space Model (TVSM)

Topic-based Vector Space Model proposes another form of vectors [21]. It considers documents as vectors in a vector space, which represents the fundamental topics of terms, whereas the VSM represents terms.

A Term i ∈{1,…,n} is represented by a *term vector* t_i and has a *term weight* between 0 and 1. Term Weight is defined by:

$$\vec{t_i} = (t_{i1}, t_{i2}, \dots, t_{id}) \in \mathrm{R}$$

$$|\vec{t_i}| = \sqrt{t_{i1}^2 + t_{i2}^2 + \cdots + t_{in}^2} \in [0; 1] \quad (4)$$

A document K∈{1,..n} is represented by a document vector $\vec{d_k}$, where vector length is standardized to 1.

$$\vec{d_k} = \frac{1}{\vec{\delta}_k} \rightarrow |\vec{d_k}| = 1$$

$$\vec{\delta}_k = \sum_{i=1}^{n} e_{ki}\vec{t_i} \quad (5)$$

Where e_{ki} =number of term i in document k

Now we can define the similarity between two documents l and k as the scalar product from vector documents, basing on their document representation.

$$\vec{d_k}\vec{d_l} = |\vec{d_k}|.|\vec{d_l}|.Cos(w_{kl}) = Cos(w_{kl})$$

Where $|\vec{d_k}| = |\vec{d_l}| = 1$

$$|\vec{d_k}|.|\vec{d_l}| = \frac{1}{|\vec{\delta}_k|}\vec{\delta}_k \frac{1}{|\vec{\delta}_l|}\vec{\delta}_{l} = \frac{1}{|\vec{\delta}_k||\vec{\delta}_l|}\sum_{i=1}^{n}\sum_{j=1}^{n} e_{ki}e_{lj}\vec{t_i}\vec{t_j} \quad (6)$$

The length of the unnormed document-vector $\vec{\delta}_k$ can be defined as:

$$|\vec{\delta}_k| = \sum_{i=1}^{n} e_{ki}\vec{t_i} = \sqrt{\left(\sum_{i=1}^{n} e_{ki}\vec{t_i}\right)^2} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} e_{ki}e_{kj}\vec{t_i}\vec{t_j}} \quad (7)$$

V. Experiments and Result

First, we have to describe the possible candidates of ambiguous word, then select from the meaning of the ambiguous word candidates, one that describe the sense of this word depending on its context. Inspired by information retrieval techniques [5], the sense and the tag context information are represented by means of vectors, which then can be compared by measuring the angle between them.

Our Model can be defined as *E={ <T,R,Y,S,W,X,Z>}*

Where:

- T is the set of terms describing the video,
- *R* is the set of resources
- *S* is the set of senses,
- *W* is the set of terms related to those sense
- *X ⊆ S×W* is the relation between senses and terms,
- *Z ⊆ S×T* is the relation between tags and senses.
- *Y ⊆ U×T×R* is the relation denoting a tagging activity.

We describe the following data sets related to a metadata.

- *Senses(t ∈ T) = {s_j: (t,s_j) ∈ Z }*. The set of the senses associated to a tag.
- *Terms(s ∈ S) = {w_k: (s,w_k) ∈ W }*. The set of terms associated to a sense.
- *Voc(t ∈ T) = ∪ Terms(s_j): s_j ∈ Senses(t).* The set of terms of all the senses associated to a tag. Voc stands for vocabulary.
- *Context (t ∈ T, r ∈ R) = {t_l ∈ T : (t_l, r) ∈ Y}*. The set of all co-occurring tags to annotate the resource r.

Now we can define a tag, its context and the senses associated to that tag using vectors. Those vectors are in $\Re$ |Voc(t)| and each position in the vector corresponds to a term in Voc(t).

- The vector $V_{context} = (v_i)$ represent a tag and its context, where $1 \leq i \leq |Voc(t)|$ and $v_i = 1$ if the corresponding term w_i in Voc(t) appears in Context(t,r), otherwise $v_i = 0$.
- The vector $V_{sense} = (v_i)$ represent the sense associated to a tag, where $1 \leq i \leq |Voc(t)|$ and v_i is the frequency of the corresponding term $w_i \in Voc(t)$ in the sense .

By this way when we want to analyze a keyword, we have to create a $V_{context}$ for the specific keyword and then a V_{sense} for each sense related to that keyword. Then, we use a similarity measure to compare $V_{context}$ with each V_{sense} . Cosine function represents the most known similarity measure of vectors, which represents the angle between two vectors where his value is between 0 and 1.

A. Example with VSM algorithm

In this example, we need to disambiguate the keyword *Apple* which resource is R.

- Context(Apple, R)={ Apple, Steve, Jobs, iPhone }
- Senses(Apple)={*dbpedia: Apple, dbpedia: Aple_Inc*}, knowing that different senses of the keyword are obtained after querying the RDF DBpedia data, using SPARQL language.

For simplification reasons, we have limited our query to only the tow first's results, because of the big number of result that can contain the query.

- Terms(*dbpedia:Apple*)={Apple, malus, tree, cultivars} having the frequencies (125,74,26,23)
- Terms(*dbpedia: Apple_Inc*)={Apple, inc, jobs, iphone} having the frequencies (787,109,85,75)

We collect all this information to create Voc()={*Apple, malus, tree, cultivars, inc, Jobs, iphone* }

Now we create the vectors to represent the keyword context and the senses.

- $V_{context}$=(1,0,0,0,0,1,1)
- V_{Apple}=(125,74,26,23,0 ,0,0)
- V_{Apple_Inc}=(787,0,0,0, 109,85,75)

With the first Algorithm we apply the cosine function directly on Vector of senses and Vector of context as described in(1). Then we get the results below:

- Sim(Vcontext,VApple)=0,483
- Sim(Vcontext, VApple_Inc)=0,705

Results show that *dbpedia: Apple_Inc* is the most probable meaning of the keyword Apple according to its context.

B. Example with TVSM algorithm

In TVSM, we add the term vector $\vec{t_i}$. As we have described in the previous section, t_i represent the weight of the term i in each document. In our case, document are represented by *Senses(t ∈ T) U Context(t ∈ T).* Then we get this description:

- $\vec{t_i} = (t_{i1}, t_{i2}, \ldots, t_{id})$ where $t_{i,1}$ is the weight of the term i in the document d_i ;
- the weight represent the occurrence of the term i in document d_i / sum of the term i in all candidate resources.

We treat the same example of the last section and we compute all vector terms $t_{i,\ where}\ t_i \in$ Voc:

- $\vec{t}_{Apple} = (0.136;\ 0.86\ ;\ 0.001)$
- $\vec{t}_{malus} = (1;\ 0;\ 0)$
- $\vec{t}_{tree} = (1;\ 0;\ 0)$
- $\vec{t}_{cultivars} = (1;\ 0;\ 0)$
- $\vec{t}_{inc} = (0;\ 1;\ 0)$
- $\vec{t}_{jobs} = (0;\ 0.98;\ 0.011)$
- $\vec{t}_{iphone} = (0\ ;\ 0.98;\ 0.013)$

Now we can compute $\vec{\delta}_k = \sum_{i=1}^{n} e_{ki}\vec{t}_i$ and then apply the similarity between the two vectors δ_{sense} and $\delta_{context}$ by applying (6). We get as result:

- Sim(δ_{Apple}, $\delta_{context}$)=0,647
- Sim(δ_{Apple_Inc}, $\delta_{context}$)=0,998

This approach gives us a correct selection of the meaning of the keyword *Apple*.

C. Real Examples

In this section we give some real examples, we have used to test our system. The input of the system is a description of a video taken from TRECVID[8] database, composed of several terms that represent the keyword and the context describing the video.

The first example is a video about a presentation which defines the Bingo club; the presentation is illustrated by storyboards. The video is described by the following terms: {*Storyboard; illustration; animation; bingo; online bingo; editor; jackie*}. A request was sent to SPARQL DBpedia endpoint to extract different resources containing the word *Storyboard.* The result represents the candidate resources

[8] The main goal of the TREC Video Retrieval Evaluation (TRECVID) is to promote progress in content-based analysis of and etrieval from digital video via open, metrics-based evaluation. TRECVID is a laboratory-style evaluation that attempts to model real world situations or significant component tasks involved in such situations.

where the disambiguation algorithm will take them as inputs, and calculate the similarity measures between the context and each candidate.

In Table.I we present all the candidate resources obtained and its similarities. As it is represented, the resource *http://dbpedia.org/resource/Storyboard* was chosen by the algorithm as the most similar resource to the keyword and its context.

Storyboard	
http://dbpedia.org/resource/Storyboard	0.677
http://dbpedia.org/resource/Storyboard_Artist	0.461
http://dbpedia.org/resource/Bryan_Andrews_(storyboard_artist)	0.380
http://dbpedia.org/resource/StoryBoard_Quick	0.358
http://dbpedia.org/resource/Crank_Storyboard_Suite	0.339
http://dbpedia.org/resource/Quality_storyboard	0.0
http://dbpedia.org/resource/IBM_Storyboard_Plus	0.0
http://dbpedia.org/resource/Storyboard_(TV_series)	0.0

TABLE I. Disambiguiation results for “Storyboard” tag

The second example, presented in Table.II is a video about a man who presents a *Powerpoint Karaoke*. The video has been tagged by the following keywords: {*karaoke; powerpoint; talks; slideshow; Hippy; Awareness; Technology; PCI; Concrete; Paving* }. For simplification reasons, we present only the first most preferment results. The disambiguation algorithm chooses the resource *http://dbpedia.org/resource/Powerpoint-Karaok*.

Karaoke	
http://dbpedia.org/resource/Powerpoint-Karaok	0.999
http://dbpedia.org/resource/The_Karaoke_King	0.707
http://dbpedia.org/resource/Karaoke	0.639
http://dbpedia.org/resource/Karaoke_Studio	0.575
http://dbpedia.org/resource/Is_for_Karaoke	0.560
http://dbpedia.org/resource/Karaoke_clubs_in_Sri_Lanka	0.525
http://dbpedia.org/resource/Karaoke_Plays	0.452
http://dbpedia.org/resource/Karaoke_Joysound	0.438
http://dbpedia.org/resource/Sing!_Karaoke	0.426
http://dbpedia.org/resource/Fun_Bar_Karaoke	0.420
http://dbpedia.org/resource/Karaoke_(TV_series)	0.376
	0.375
	0.319

TABLE II. DIsambiguiation results for “Karaoke” tag

In the last examples, the system chooses the correct meaning but sometimes the algorithm fails to give the right meaning of the tag. This is due, first to the weakness of the context accompanying the tag, then to the choice of similarity measures and many other criteria. We aim in our next work to pay attention to these issues, to improve the results of the disambiguation algorithm. Furthermore we aim to test this approach with a larger set of data.

D. *Results and evaluation*

Many Evaluation systems have been present in literature. The most used in retrieval systems are presented in [22], where the author defines measures for evaluation of information retrieval systems; Recall ratio and Precision ratio. Recall is the proportion of retrieved documents that are relevant; Precision is the proportion of relevant documents that were retrieved.

$$Precision = \frac{|relevant\ and\ retrieved|}{|retrieved|}$$

$$Recall = \frac{|non-relevant\ and\ retrieved|}{|non-relevant|} \qquad (8)$$

F -measure is the weighted harmonic mean of precision and recall. It can be computed as:

$$F-measure = \frac{2*Recall*Precision}{Recall+Precision} \qquad (9)$$

Our enhancement system, need a disambiguation step. The evaluation of this later, give us the results in Table.III. Experiments show that TVSM algorithm outperform VSM one. This is due to the fact that TVSM does not assume independence between terms by being flexible in definition of term similarities, contrasting to VSM. The main difference of TVSM comparing to VSM is the operational vector space. TVSM represents documents as vectors in a d dimensional space R which has only positive axis intercepts [21].

	Recall	Precision	F-Measure
VSM	0.6	0.3	0.4
TVSM	**0.68**	**0.32**	**0.44**

TABLE III. EVALUATION OF VSM AND TVSM

VI. Conclusions and future work

As DBpedia reflects the largest Dataset in the Linked Open Data cloud, extracted from Wikipedia, we use it for linking with videos in regard of enhancing videos with structured information coming from the biggest encyclopedic database. To do this, we propose an architecture where the main steps are to disambiguate the description of video and linked it to structured information in LOD cloud. For disambiguation and ranking step of candidate resources, we use tow different algorithms: Vector Space Model and Topic-based Vector Space Model. We evaluated the system using TRECVID database where video is present with textual description.

Although the result was improved between the two algorithms, but there is still work to do. The failure of the algorithm is due to several criteria. The first problem is due to the video description that must to undergo a pretreatment step. Also the weakness of the context can provide problems and other criteria concerning disambiguation algorithm.

In the future work we aim to implement a term normalization service before disambiguation step. We aim also to add a feedback between the user and the system to enrich the context for better results. Then, we have to adopt more disambiguation algorithms and compare between the two studied in this paper. Finally we aim to implement the entity filtering service to have a complete and coordinated system.

Acknowledgment

The authors would acknowledge the financial support of this work by grants from General Direction of Scientific Research (DRGST), Tunisia, under the ARUB program

References

[1] M. Rey-López., R. P. D. Redondo, A. F. Vilas, J. J. P. Arias, M. L. Nores, J. G. Duque, A. Gil-Solla and M. R. Cabrer, "T-MAESTRO and its authoring tool: using adaptation to integrate entertainment into personalized t-learning.", *Multimedia Tools Appl*. pp. 409-451, 2008 .

[2] G.Kobilarov, T.Scott, Y. Raimond, S. Oliver, C. Sizemore, M. Smethurst, C. Bizer and R. Lee, "Media Meets Semantic Web- How the BBC Uses DBpedia and Linked Data to Make Connections", 6thEuropean *Semantic Web Conference (ESWC)*, Heraklin, Crete, Greece, pp. 723-737, 2009 .

[3] Y. Raimond, T. Ferne, M. Smethurst, G. Adams, "The BBC World Service Archive Prototype*", Web Semantics: Science, Services and Agents on the World Wide Web*, North America, V.27-28, August-October 2014.

[4] P. N.Mendes, M.Jakob, A.García-Silva, C.Bizer. "DBpedia spotlight: shedding light on the web of documents", the 7th *International Conference on Semantic Systems*, ACM, Graz, Australia, pp. 1—8, 2011.

[5] R.A.Baeza-Yates, B. Ribeiro-Neto, "Modern Information Retrieval", *Addison-Wesley Longman Publishing Co., Inc.*, Boston, MA,USA, 1999.

[6] Garcia, Andres; Szomszor, Martin; Alani, Harith and Corcho, Oscar. "Preliminary results in tag disambiguation using DBpedia". In: *The Fifth International Conference on Knowledge Capture (K-Cap'09)*, 1 Sept 2009, Redondo Beach, California, USA.

[7] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiack, S. Hellmann, "DBpedia—a crystallization point for the Web of data", *Journal of Web Semantics*, pp.154-165, V7, 2009.

[8] O. Ben Said, H. Karray, A.M Alimi and A. Haqiq. "T-Learning and Interactive Television for Edutainment", *International Conference on Engineering Education and Research 2013 (iceer 2013).* July 2013.

[9] O. Ben Said, A. Wali, A.M. Alimi, "A new System for TV program contents improvement using a semantic matching technique", *13th International Conference on Hybrid intelligent System (HIS)*, 2013, pp.246-250.

[10] O. Ben Said, A. Wali, A.M. Alimi, "Interlinking video programs with Linked Open Data", *15th International Conference on Intelligent Systems Design and Applications (ISDA)*, pp. 462-467, December 2015.

[11] C. Becker, C. Bizer, "Exploring the Geospatial Semantic Web with DBpedia Mobile", *Web Semantics: Science, Services and Agents on the World Wide Web*, pp.278-286, V 7(4), December 2009.

[12] S. Angeletou, M.Sabou, E. Motta, "Semantically enriching folksonomies with FLOR", *1st International Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008), 5th Annual European Semantic Web Conference (ESWC),* Jun 2008.

[13] S. Lee and H. Yong, "Tagplus: A retrieval system using synonym tag in folksonomy", *International Conference on Multimedia and Ubiquitous Engineering*, pp. 294–298, Seoul, Korea, 2007.

[14] R. Navigli, "Wsd: a survey", *ACM Computing Surveys*, V.41(2), P.1–69, 2009.

[15] P.N. Mendes, M. Jakob, A. Garcia-Silva and C. Bizer, "DBpedia Spotlight: Shedding light on the web of documents", *Proceedings of the 7th International Conference on Semantic Systems (I-Semantics)*,pp.1-8, Graz, Austria, 2011.

[16] M. Rey-López, R.P. Díaz-Redondo, A. Fernández-Vilas and J.J. Pazos-Arias, "Entercation experiences: Engaging viewers in education through TV programs", *In Fourth European conference on interactive television (EuroITV)*, Athens, Greece pp. 310–319, 2006.

[17] Sören Auer, Chris Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives, "DBpedia: A Nucleus for a Web of Open Data,", *the 6th International Semantic Web Conference (ISWC*), vol. 4825,pp. 722--735. Springer, 2008.

[18] G.Salton, M. Lesk, "Computer evaluation of indexing and text processing", *Journal of the ACM*, Vol.15 (1), pp.8-36, 1968.

[19] R. Beaza-Yates, B. Ribeiro-Neito, "Modern Information Retrival*". ACM Press*, New York, 1999.

[20] G. Salton, M.J. McGill, "Introduction to Modern Information Retrieval, *McGraw-Hill Computer Science Series*, September 1986.

[21] J.Becker, D.Kuropka, "Topic-based Vector Space Model", *the 6th International Conference on Business Information Systems*, Colorado Springs, p. 7–12, July 2003.

[22] C.W. Cleverdon, "The Cranfield tests on index language devices", in *ASLIB Proceedings*, San Francisco, CA, USA, pp. 47–59, 1997.

[23] Yves Raimond, Chris Lowis, "Automated interlinking of speech radio archives", *Linked interlinking of speech radio archives, Word Wide, Wibe Conference (LDOW2012)*, Lyon, France, April 2012.

[24] Vo. Sao-Khue, A. Anjomoshoaa, A. Min. Tjoa, "Mashups Semantic-Aware pour Ressources personnelles en SemanticLIFE et SocialLIFE", Conference internationale de la Croix-Domain, Fribourg, Suisse, pp.138-154, V.7808, September, 2014.

[25] Wei Hu, Cunxin Jia, Lei Wan, II Liang, L. Zou, Y. Qu, "CAMO: Intégration des données liées Open for Metadata Multimedia Enrichissement ", *Conference Internationale Semantic Web*, pp.1-16, V.8796, Octobre 2014.

Author Biographies

Olfa Ben Said. was born in Béja in 1985. She received the four year university degree in Informatics-Mathematics from the Faculty of Sciences of Sfax (FSS), Tunisia, in 2008. Then she obtained the master degree in New Technologies of the Dedicated Computer Systems (NTSID) from the National School of Engineers (ENIS), University of Sfax, Tunisia. She is now preparing her ph.D in the Research Group in Intellegent Machines (REGIM-Labs), in ENIS at university of Sfax. She is interested in Information Retrieval systems, Data web and Video Enhancement.

Ali Wali. is an Assistant Professor on Computer Sciences at ISIM, University of Sfax. He got his Ph.D. in Engineering Computer Systems at National school of Engineers of Sfax, in 2013. He is member of the REsearch Groups on Intelligent Machines (REGIM). His research interests include Computer Vision and Image and video analysis. These research activities are centered on Video Events Detection and Pattern Recognition. He is a Graduate member of IEEE. He was the member of the organization committee of the International Conference on Machine Intelligence ACIDCA-ICMI2005, Third IEEE International Conference on Next Generation Networks and Services NGNS2011 and 4th International Conference on Logistics LOGISTIQUA2011, International Conference on Advanced Logistics and Transport (ICALT'2013, ICALT'2014,ICALT'2015), 13th International Conference on Document Analysis and Recognition (ICDAR 2015), International Conference on Hybrid Intelligent Systems (HIS'2013), International Conference on Information Assurance and Security (IAS'2013).

Adel M.Alimi. is an IEEE Student Member'91, Member'96 and Senior Member'00. He graduated in Electrical Engineering in 1990. He obtained a PhD and then an HDR both in Electrical & Computer Engineering in 1995 and 2000 respectively. He is full Professor in Electrical Engineering at the University of Sfax, ENIS, since 2006. He is founder and director of the research REGIM-Lab. on intelligent Machines.