UNIVERSITÀ DEGLI STUDI DI TORINO

Alessio Antonini, Luca Vignaroli, Claudio Schifanella, Ruggero G. Pensa, Maria Luisa Sapino

MeSoOnTV: A Media and Social-driven Ontology-based TV Knowledge Management System

Editor: ACM
2013
ISBN: 9781450319676

in

Proceedings of the 24th ACM Conference on Hypertext and Social Media - HT '13
208 - 213
The 24th ACM Conference on Hypertext and Social Media - HT'13
Paris, France
1-3 May 2013

MeSoOnTV: A Media and Social-driven Ontology-based TV Knowledge Management System

Alessio Antonini
University of Torino
Torino, Italy
antonini@di.unito.it

Luca Vignaroli
RAI-CRIT
Torino, Italy
l.vignaroli@rai.it

Claudio Schifanella
University of Torino
Torino, Italy
schi@di.unito.it

Ruggero G. Pensa
University of Torino
Torino, Italy
pensa@di.unito.it

Maria Luisa Sapino
University of Torino
Torino, Italy
mlsapino@di.unito.it

ABSTRACT

Searching, browsing and analyzing web contents is today a challenging problem when compared to early Internet ages. This is due to the fact that web content is *multimedial*, *social* and *dynamic*. Moreover, concepts referred by videos, news, comments, posts, are implicitly linked by the fact that people on the Web talks about something, somewhere at some time and these connections may change as the perception of users on the Web changes over time. We define a model for the integration of the heterogeneous and dynamic data coming from different knowledge sources (broadcasters' archives, online newspapers, blogs, web encyclopedias, social media platforms, social networks, etc.). We use a knowledge graph to model all the heterogenous aspects of the information in an homogeneous way. Through a case study on social TV, we provide a non trivial cross-domain analysis scenario on real data gathered from YouTube and Twitter, and related to an Italian TV talk show on politics, broadcasted by RAI, the Italian public-service broadcasting organization.

Categories and Subject Descriptors

H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*information networks*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems

1. INTRODUCTION

Searching, browsing and analyzing web contents is today a challenging problem when compared to early Internet ages. This is due to multiple reasons. First, web content is *multimedial*: mere textual information has been replaced by combinations of text, pictures, sounds, videos, animations and interactive forms of content. Second, web content is *social*: Internet users interact with each other in social networking sites, blogs, forums; they add comments and tags to any kind of multimedia contents. Third, web content is *dynamic*: web objects evolve rapidly over time, as users enrich them with new descriptions, external references, opinions; the perception on items, events, persons changes consequently. To capture all these aspects and make the analysis more reliable and closer to the complex and dynamic nature of the reality, simply gathering web contents and analyzing them, even using enhanced data mining and analytics tools is not enough. In fact, concepts referred by videos, news, comments, posts, are implicitly linked by the fact that people on the Web talks about something, somewhere at some time. Moreover, concepts and web content are semantically linked to each other, and these connections may change as the perception of users on the Web changes over time.

To explain this position, we instantiate it in the context of our investigations: the integration of the cultures of TV and Web. In this paper we propose *MeSoOnTV*, our Media and Social-driven Ontology-based TV Knowledge Management System through which the social media (user provided) information and the TV content are integrated, and can be leveraged to improve users' experience as well as broadcasters returns.

More specifically, the integrated domain is modelled as a *knowledge graph*, in which nodes represent the concepts, while edges capture the relationships existing among them. This type of knowledge representation has been defined in the 70's, but, today, knowledge graphs are exploited by semantic analysis [14], sentiment analysis [15] and opinion mining [17] state of the art technologies. Furthermore, time is also a key question in knowledge representation and analysis [6]. As an example, Google uses a knowledge graph for its search engine[1]. In the recent literature, a number of interesting works integrating different knowledge-driven contexts have been presented, for instance in medical imaging and advanced knowledge technologies for cancer diagnosis [4], or personalized web browsing and search by the construction of user models based on a semantic representation of the user activity [16]. Other approaches propose conceptual frameworks supporting context-specific naming and representation of conceptual entities and related action executions [9]. Also in the field of web document model-

[1] http://www.google.com/insidesearch/features/search/knowledge.html

24th ACM Conference on Hypertext and Social Media
1–3 May 2013, Paris, France

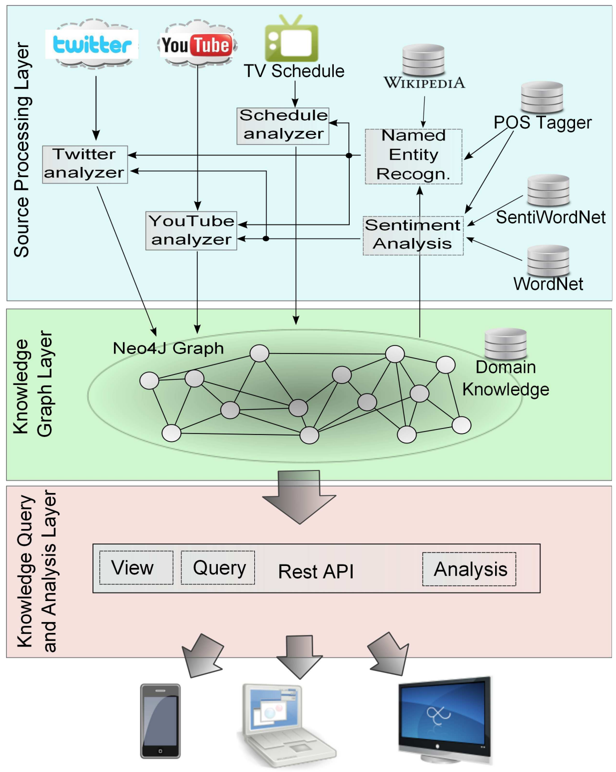

Figure 1: The *MeSoOnTV* integration framework and the related system architecture

ing [8], knowledge integration supports information search, focused web crawling and content adaptation. Similarly to these works, the key idea that we convey in our model is that the meaning of each entity and relationship within the knowledge graph depends on the context in which they are considered.

Through a case study, we show how our model captures multiple aspects of the considered domain, from the semantic characterization of the TV content, to the temporal dimension of the problem, to the social characterization and the social perception of a TV event. Last but not least, we provide a non trivial cross-domain analysis scenario on real data gathered from YouTube and Twitter, and related to an Italian TV talk show on politics, broadcasted by RAI.

2. A FRAMEWORK FOR SOCIAL MEDIA DATA INTEGRATION AND ANALYSIS

In this section we introduce the framework which enables the integration of various social and non social sources of information in a unique knowledge base. The knowledge base, modelled as a knowledge graph integrating domain and general purpose ontologies as well as social interactions among users and social media, can be queried and analyzed as a whole, enabling the discovery of new and interesting cross-domain patterns.

2.1 The integration framework

Figure 1 presents an overview of our integration framework. It consists of three main layers: a source processing layer, a knowledge graph layer and a knowledge query and analysis layer. The **source processing layer** has the role of collecting all the data which will be conveyed in the model. It accesses a number of predefined web/social/media sources (e.g., broadcasters official web sites, social networks, TV channels, etc) and processes them in order to extract those information units which will be represented as nodes in the knowledge graph, as well as the information that supports the existence of relationships (modelled as edges in the graph) among them.

The **knowledge graph layer** manages the knowledge graph, which is the core of our proposal. The graph contains essentially three types of nodes: social objects, subjects and concepts, and all social, representation and structural interactions among them.

The **knowledge query and analysis layer** consists in a set of components for querying, browsing and analyzing the knowledge graph. A query module extracts subgraphs from the knowledge graph based on user's requirements and constraints. Each extracted subgraph can be seen as a "view" over the complete knowledge graph, only containing nodes and edges potentially relevant to the user query. An analysis module, provides a set of analysis and data mining tool to obtain models and patterns from the knowledge graph. It can act directly on the knowledge graph, or it can handle the views extracted from the query module also in terms of matrices or tensors.

The core of our framework is the knowledge graph. In particular, we are interested in capturing the dynamic evolution in time of the graph by using temporal nodes associated to social objects and describing their lifecycle.

Notice that in our integration framework a fundamental role is played by a *semantic engine* in two places. First, it is adopted in the source processing layer to provide an interpretation to web/social/media elements taken by the heterogeneous sources. Within this layer, the semantic engine helps understand whether the considered entities should be modelled as a node or an edge in the graph, and helps provide a congruent set of features based on their characteristics. Second, it plays an important role in the graph query and analysis layer, where it is employed to assign a semantic role to each selected node/edge. In the following sections, we describe our framework in details.

3. MANAGING MULTIPLE KNOWLEDGE SOURCES

The core of our framework is the knowledge base that represents the result of public actions of users in social environments, combining different theories from cognitive science [2, 5], language philosophy [12] and social ontology [11]. In this domain we recognize three entities (corresponding to three types of nodes in the knowledge graph): *subjects*, users that act, *social objects*, the result of public acts, and *concepts*, physical and ideal objects referred by subjects via their public actions. Any act (or a set of acts) that can be identified by its trace, and have a recognized social value is a social object. However, we do not represent single subjects' actions but a unique social object for each group of similar actions.

We introduce relationships between subjects and social objects and between social objects and concepts in that way: a group of subjects that recognize a social value of an act *supports* the resulting social object (e.g. the contractors

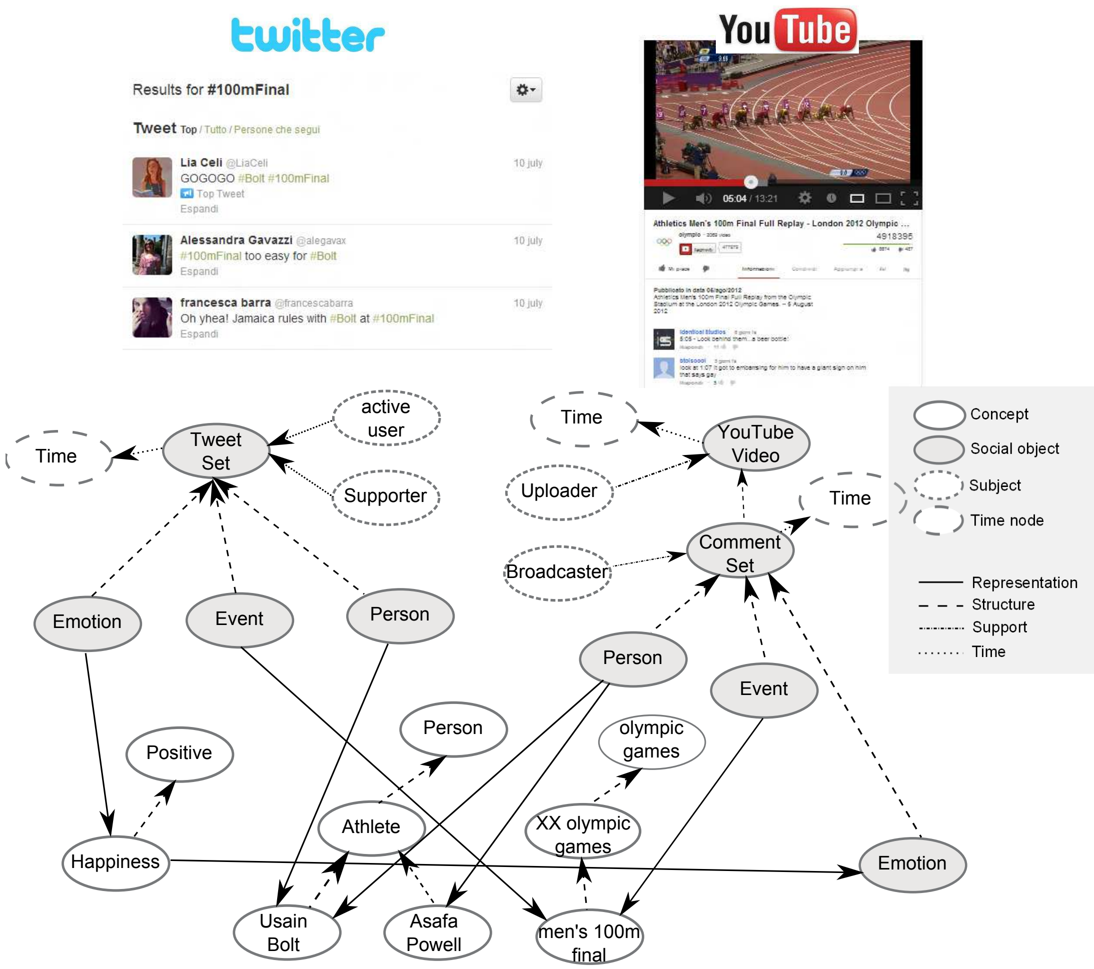

Figure 2: An instance of knowledge graph from different social sources

support the contract); a social object *represents* a social instance of some concepts on a precise context (e.g. a video may represent a volleyball match). Other relationships interest entities of the same type. We call these relationships *structural dependencies*. A social object o_1 is *structural* of another object o_2 if o_1 is part of o_2 (e.g. a comment is part of a video). A subject is *structural* of a group of subjects (e.g. a subscriber is part of playlist subscribers) that performed the same kind of actions on the same social object. A concept may be *structural* of a more general concept (e.g. hilarity is a specialization of joy).

Finally, social objects evolve with time. Hence, as a special case of representation relationship, we consider the *temporal representation* of a social object towards a special type of concept called *time concept* (e.g. a video has been posted in a specific time instant, and has been viewed during a specific time period).

3.1 Knowledge Graph

To provide a meaningful representation of the knowledge base, we employ a *Knowledge Graph* that enables to model all relationships between social objects, subjects and concepts. At this purpose, let us first introduce the required notation. Let $\mathcal{O}$, $\mathcal{S}$ and $\mathcal{C}$ be, respectively, the sets of all social objects, subjects and concepts. Let $\mathcal{T} \subseteq \mathcal{C}$ be the set of time concepts. Our *Knowledge Graph* is defined as follows:

Definition 1 (Knowledge Graph). *Given $\mathcal{O}$, $\mathcal{S}$, $\mathcal{C}$, the corresponding* Knowledge Graph *is a directed weighted graph $G^K(V, E, W)$, where, $V = V^O \cup V^S \cup V^C$ is the set of vertices built on $\mathcal{O}$, $\mathcal{S}$ and $\mathcal{C}$, $E = E^{sup} \cup E^{rep} \cup E^{str}$, with $E^{sup} = \{(v_i^S, v_j^O)\ s.t.\ v_i^S \in V^S,\ v_j^O \in V^O\}$ the set of* support *edges, $E^{rep} = \{(v_i^O, v_j^C)\ s.t.\ v_i^O \in V^O,\ v_j^C \in V^C\}$ the set of* representation *edges, $E^{str} = \{(v_i, v_j)\ s.t.\ v_i, v_j \in V^S \vee v_i, v_j \in V^O \vee v_i, v_j \in V^C,\ i \neq j\}$ the set of* structural *edges, and $W : E \rightarrow [0, 1]$ is the function that associates a weight w_{ij} to each edges $(v_i, v_j) \in E$, $v_i, v_j \in V$, $i \neq j$. Moreover, given the set of time concepts $\mathcal{T} \subseteq \mathcal{C}$, $E^{tmp} \subseteq E^{rep}$ is the set of edges (v_i^O, v_j^T), where $v_i^O \in V^O$, and $v_j^T \in V^T$, $V^T \subseteq V^C$ being the set of vertices representing the time concepts.*

A special subgraph of G^K is the one consisting of all concepts of $\mathcal{C}$, i.e., the forest of the ontology of concepts, or the users' shared conventional knowledge.

Definition 2 (Ontology Graph G^{KB}). *The* Ontology Graph *$G^{KB}(V^{KB}, E^{KB}, W^{KB})$ is the subgraph of G^K induced by $V^{KB} = V^C$.*

To populate the knowledge graph, our framework may interact with different and heterogenous sources of information. Each source is first analyzed, then relevant items and relationships are extracted and added to the graph. In the following, we explain how to analyze the sources of interest.

3.2 Source analysis

Our knowledge graph can be feeded from any source of information. However we distinguish between two kinds of sources: *social sources* and *non social sources*. The first ones consist essentially of social networking platforms, social media platforms and blogs. The second group of sources consists of general purpose or domain ontologies, online newspapers, news feeds, broadcasting websites that are needed to provide a human view on the results of social interactions. In our framework external sources are analyzed in order to extract resources that can be added to the knowledge graph following a set of specific rules. In a nutshell, a resource is an independent information unit that can be mapped into vertices and edges in the knowledge graph.

For each source, we must set an extractor agent that should map each resource into a valid set of social objects, subjects, concepts and relationships among them. To identify correctly each entity, the extractor employs the set of ontologies G^{KB}. To map each identified entity into a congruent set of vertices and edges in the graph G^K, the extractor leverages a set of rules whose complexity depends on the specific source to be analyzed. In particular, as we mentioned earlier, we use two basic types of extractors: one for social sources, and one for non social sources.

Once the extractor agent has analyzed the source, it provides a set of concepts, subjects and social objects that should now be translated into new or updated vertices and edges in the graph. In Figure 2, we show a small example of knowledge graph obtained after the processing of the two resources extracted from Twitter and YouTube. Notice how the common concepts are the nodes that connect communities coming from different social media sites (YouTube and Twitter). Thanks to this structure, it becomes possible to extract new cross-domain patterns, as we will see in Section 5.

In real applications, the graph will not be instantiated with all possible resources extracted from any social or non social source. The reasons are essentially twofold: on the one hand, the huge amount of information could be untractable in practice; on the other hand, many social sources set a limit to the number of resources that can be retrieved in a time slice. For this reason, the way the knowledge graph is populated is somehow constrained by the specific application.

In the following section, we provide the details of the system architecture implementing our framework.

4. THE SYSTEM ARCHITECTURE

Figure 1 depicts the *MeSoOnTV* architecture. It can be noticed that, coherently with our framework, each source (both social and non social) is associated to an analyzer module (the *Twitter analyzer*, *YouTube analyzer* and *Schedule analyzer* boxes), whose task is to collect the data from the sources and extract concepts, subjects, social objects and their relations through the combined use of different shared modules (the *Name Entity Recognition* and *Sentiment Analysis* boxes). The knowledge base extracted from each analyzer will be used to properly update the graph G^K. More in detail, for each TV program that a Schedule Analyzer inserted in the knowledge graph, the Twitter module collects in real-time all related tweets, grouping them into time dependent slices, where each slice contains the tweets published from time t to $t + \Delta$. Each tweet set is then processed in order to detect the named entities (people, places and events) trough the use of a NER (Name Entity Recognition) module, while a Sentiment Analysis module allows to extract the opinions contained in a tweet set. Similarly, at each time slice, the YouTube analyzer looks for new videos or new user comments that belong to previously analyzed media and performs the same type of analysis described for Twitter.

Within the NER module, we can distinguish two different phases: entity detection and entity disambiguation [7]. Entity detection is performed by a combined use of the Freeling POS Tagger [10] and Wikipedia articles[2] as reference knowledge base. In particular, through the use of the Wikipedia search API, the NER module is able to detect the presence of entities starting from hashtags: for example, the hashtag *#barackobama* will be recognized by Wikipedia as the string "Barack Obama". Nevertheless, the most challenging task in Named Entity Recognition is represented by the entity disambiguation (or resolution) [7]. Since our scenario is characterized by the presence of short and sparse texts (both for Twitter and YouTube comments), many of the existing approaches based on the Bag of Words model will fail: for this reason our NER module tries to leverage additional information provided by the context defined by the TV program in which the resolution process is involved, in order to establish which entity is the best among the set of the candidate real-world entities. In details, the context of a TV program is defined by using the Wikipedia categories it belongs to and the set of all entities contained in the knowledge graph previously associated with the program. In this manner, for each detected entity, the NER module tries to establish an order among all real-world candidates extracted from Wikipedia. For example, if the text "Michael Jordan" is contained in a tweet set related to a TV sports program, it is very likely that the tweeter is referring to the famous basketball player rather than the Berkeley's professor, and this is computed by a comparison between the Wikipedia categories of the candidates and the corresponding categories of the TV program. Moreover, if, for example, Michael Jordan is present within the knowledge graph as a real-world entity recognized and associated with the considered TV program (i.e. because he is the presenter or a frequent guest), the NER module will choose it among all the possible real-world entity candidates. Finally, our module supports the integration of external knowledge generated by a supervised scenario and it allows for user feedback, using an active learning process. In our application, we filter out infrequent recognized entities with the energy cutoff method.

The Sentiment Analysis module is used to extract polarity values and emotions from tweet sets. Concerning the former, a first phase of lemmatization is performed by the Freeling POS tagger, while SentiwordNet [1] is used to extract the polarity values: hence, an aggregation function allows us to enrich each tweet set in the knowledge graph with a degree of positivity, negativity and neutrality. With the same approach, WordNet-Affect [13] is used to extract emotions. Where necessary, MultiwordNet[3] is used for cross-language

[2] `http://www.wikipedia.org`
[3] `http://multiwordnet.fbk.eu`

(a)

(b)

Figure 3: Two examples of hashtag clusters.

purposes.

The knowledge graph is realized and stored in Neo4j[4], the well known NoSQL graph database: it offers a comprehensive REST interface, an object-oriented API, and it scales up to billions of nodes and relationships with properties.

The last component of the *MeSoOnTV* architecture is the module dedicated to the data analysis and publication of the results to the end users of the system. This module is able to extract both simple views of the graph and more complex query and analysis algorithm, in order to expose the corresponding results to devices like smartphones, computers and TVs by defining a standard REST API.

5. AN EXAMPLE ON ITALIAN POLITICS

In this section, we describe a real use-case of *MeSoOnTV* on an Italian TV show (Ballarò) dealing with politics and broadcasted by RAI. We focused our analysis on the episodes scheduled from October 2, 2012 to November 27, 2012 (nine episodes). This period is interestingly full of political events for many reasons: the past or future elections in many big Italian regions (Sicily, Lazio and Lombardy); the upcoming Italian general elections; the recession; the rise of the populist extra-parliamentarian group M5S (Movimento 5 Stelle) that many polling institutes consider as one of the favorite parties for the 2012 political elections in Italy.

We considered two social sources: Twitter and YouTube. For each episode, we collected all tweets containing *#Ballarò* (the official program hashtag) or *@RaiBallaro* (the official program username). YouTube videos were extracted at once by including in the search fields the keyword related to the TV program title ("Ballarò") and the date each episode was broadcasted (e.g., "2-10-2012" or "2 ottobre 2012").

5.1 An Example of Cross-Source Analysis

As an example of the potential analysis scenarios that our framework may enable, we consider non trivial associations betweens YouTube videos and Twitter hashtags. These two objects are not immediately linked: users' communities and social platforms are different. However, they may have in common several entities (persons, nouns, events, emotions). Thanks to our framework, it is quite simple to compute the entities that connect videos and hashtags. We then construct a hashtags × videos matrix (called M) in the following way. For a given video v and a given hashtag h, we call $TS(h) = \{ts_i^h\}$ the set of all tweet sets which h is associated to in G^K. Then, for each element $ts_i^h \in TS(h)$, we compute the number of concepts associated to both v and ts_i^h. We call this number c_{iv}^h. Then, the value m_{hv} of matrix M, is given by $m_{hv} = \sum_i c_{iv}^h$. We repeat this computation for each pair (h, v) of hashtags h and videos v. We ignore all concept nodes related to emotions in this case. As a result, the association of all videos and tweets related to the monitored period, bring to a matrix M of 258 hashtags, 249 videos and a 12167 non-zero values.

It is now interesting to obtain associations between groups of hashtags and groups of videos. As an example of application, we may imagine cross-domain recommendation of interesting hashtags to people watching YouTube videos, or interesting videos to people using some hashtags in Twitter. To compute relevant cross-associations, we use the well-known *information theoretic co-clustering* algorithm [3]. It identifies a clustering of rows and an associated clustering of columns by optimizing the loss in mutual information objective function. We apply this algorithm to compute two co-clustering results: the first with a grid of 10×10 co-clusters; the second with 5×5 co-clusters. We then associate to each cluster R of rows (videos) the cluster C of columns (hashtags) such that $\frac{1}{|R| \cdot |C|} \sum_{h \in R} \sum_{v \in C} m_{hv}$ is max. As an example of results, we consider two co-clusters, one for each result set. The first one, extracted from the 10×10 grid, associates a group of 27 videos mostly related to the 2012 elections in Sicily to the list of 17 hashtags summarized by the tagcloud in Figure 3(a). This results makes sense since this electoral competition was won by the coalition headed by the party of Pierluigi Bersani, but the M5S (Beppe Grillo's political movement) reported the highest number of preferences. The second co-cluster comes from the 5×5 grid. It consists in 60 videos mainly related to a satirist (Maurizio Crozza), that leads a 10 minutes' intervention during each episode of Ballarò TV programs. As such, it usually performs imitations of politicians (like Pierluigi Bersani and Matteo Renzi), and it is often cited or posted by audiences watching other political talk shows (such as Piazzapulita and Servizio Pubblico). This is clear from Figure 3(b), that depicts the tagcloud of the 66 hashtags associated to the described cluster of videos.

6. CONCLUSIONS

In this paper we have proposed a model for the integration of the heterogeneous data coming from many different knowledge sources, including broadcasters archives, EPGs, ontologies and social networks. The model highlights the tight interactions between the Web world and the TV world. We have also provided a concrete example of the potential applications of our framework on real data.

Acknowledgments

We are grateful to Roberto Del Pero and Fulvio Negro for their constructive discussions during the formalization of the integration framework, to Paolo Pasteris for his technical support, and to Raffaele Teraoni Prioletti for his help in the implementation of the modules for YouTube data collection.

[4] http://www.neo4j.org

7. REFERENCES

[1] S. Baccianella, A. Esuli, and F. Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proc. LREC 2010*, Valletta.

[2] B. Bara. *Cognitive Pragmatics : The Mental Processes of Communication*. MIT Press, 2010.

[3] I. S. Dhillon, S. Mallela, and D. S. Modha. Information-theoretic co-clustering. In *Proc. of ACM SIGKDD 2003*, pages 89–98, Washington, DC.

[4] D. Dupplaw, S. Dasmahapatra, B. Hu, P. Lewis, and N. Shadbolt. A distributed, service-based framework for knowledge applications with multimedia. *ACM Trans. Inf. Syst.*, 27(4):22:1–22:29, 2009.

[5] P. Johnson-Laird. *Mental Models*. Harvard Univ Press, 1983.

[6] A. C. Kaluarachchi, D. Roychoudhury, A. S. Varde, and G. Weikum. Sitac: discovering *semantically identical temporally altering concepts* in text archives. In *Proc. of EDBT 2011*, pages 566–569, Uppsala.

[7] H. Kopcke and E. Rahm. Frameworks for entity matching: A comparison. *Data Knowl. Eng.*, 69(2):197 – 210, 2010.

[8] A. Micarelli, F. Sciarrone, and M. Marinilli. Web document modeling. In *The Adaptive Web*, volume 4321 of *LNCS*, pages 155–192. 2007.

[9] R. Motschnig-Pitrik. A generic framework for the modeling of contexts and its applications. *Data Knowl. Eng.*, 32:145–180, 2000.

[10] L. Padro and E. Stanilovsky. Freeling 3.0: Towards wider multilinguality. In *Proc. LREC 2012*, Istambul.

[11] J. Searle. *The construction of social reality*. Free Press, New York, 1995.

[12] J. R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1970.

[13] R. Valitutti and C. Strapparava. WordNet-Affect: an Affective Extension of WordNet. In *In Proc. LREC 2004*, pages 1083–1086, Lisbon.

[14] P. Yan and W. Jin. Improving cross-document knowledge discovery using explicit semantic analysis. In *Proc. of DaWaK 2012*, pages 378–389, Vienna.

[15] Y. Yoshida, T. Hirao, T. Iwata, M. Nagata, and Y. Matsumoto. Transfer learning for multiple-domain sentiment analysis - identifying domain dependent/independent word polarity. In *Proc. of AAAI 2011*, San Francisco, CA.

[16] H. Zhang, Y. Song, and H.-T. Song. Construction of ontology-based user model for web personalization. In *Proc. of the 11th Intl. Conf. on User Modeling, UM 2007*, pages 67–76, Corfu.

[17] Q. Zhang, Y. Wu, Y. Wu, and X. Huang. Opinion mining with sentiment graph. In *Proceedings of IEEE/WIC/ACM Web Intelligence 2011*, pages 249–252, Lyon.