

Shaping-Up Multimedia Analytics: Needs and Expectations of Media Professionals

Guillaume Gravier, Martin Ragot, Laurent Amsaleg, Rémi Bois, Grégoire Jadi, Eric Jamet, Laura Monceaux, Pascale Sébillot

▶ **To cite this version:**

Guillaume Gravier, Martin Ragot, Laurent Amsaleg, Rémi Bois, Grégoire Jadi, et al.. Shaping-Up Multimedia Analytics: Needs and Expectations of Media Professionals. The 22nd International Conference on Multimedia Modelling, Special Session Perspectives on Multimedia Analytics, Jan 2016, Miami, United States. hal-01214829

HAL Id: hal-01214829

https://inria.hal.science/hal-01214829v1

Submitted on 13 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Shaping-Up Multimedia Analytics: Needs and Expectations of Media Professionals

Guillaume Gravier[1], Martin Ragot[3], Laurent Amsaleg[1], Rémi Bois[1], Grégoire Jadi[4], Éric Jamet[3], Laura Monceaux[4], Pascale Sébillot[2]

(1) CNRS, IRISA & Inria Rennes, (2) INSA Rennes, IRISA & Inria Rennes
(3) Univ. Rennes 2, CRPCC, (4) Univ. Nantes, LINA

Abstract. This paper is intended to help clarifying what multimedia analytics encompasses by studying users expectations. As a showcase, we focus on the very specific family of applications doing search and navigation of broadcast and social news content. This paper is first describing what professional practitioners working with news currently do. Thanks to extensive conversations with media professionals, mockup interfaces and a human-centered design methodology, we analyze the perceived usefulness of a number of functionalities leveraging existing or upcoming technologies. This analysis helps (i) determining research directions for the technology underpinning the very recent field of (multi)media analytics and (ii) understanding how multimedia analytics should be defined. In particular, dependency to the domain is discussed: are multimedia analytics tasks domain-specific or can we find general definitions?

1 Introduction

From the early days of DARPA and NIST evaluations, news search and navigation have been widely studied. The fundamental technologies needed to search and navigate news have been designed and evaluated in a number of different, though closely related, contexts such as broadcast news transcription, topic detection and tracking, automatic content extraction, etc. Several systems and prototypes originated from there. Early technology-driven systems such as Broadcast News Navigator [13], Informedia [8] and Físchclár news [12] primarily targeted indexing for search based on transcription, topic segmentation and entity extraction. Departing from the index and search philosophy, pioneering work on topic threading aimed at organizing news collections with explicit links to offer event-oriented navigation capabilities, e.g., [9, 21]. Exploratory news visualization and exploration interfaces were also designed in lab settings [7].

However, no commercial news navigation product is widely operational today for media professionals. Practitioners mainly rely on general public search systems, synthesizing and organizing search results by themselves. For example, press agents in state offices are often asked to make a brief on a given topic, e.g., Charlie's killings in Paris, mentioning the main dates and the chronology of the event, the causes and consequences as well as the main characters. To do so, they search media databases to complement their personal knowledge.

Obviously, tools to organize, navigate, synthesize and extract information from news collections would be beneficial. This is even more true now that information sources have been significantly multiplied, in particular with social networks and user-generated news and comments.

Multimedia analytics exactly match this description: organize data collections and provide tools to extract knowledge by interacting with the data. Practitioners have at hand a large amount of multimedia material that they need to understand and explore in order to gain insight. Taking advantage of this insight, they will, e.g., create some new multimedia material for the general public to have a better understanding of the society they live in, and/or provide other practitioners with another source of knowledge useful to get insight, etc. While multimedia analytics is a recent field, there are already several technological solutions available to facilitate working and understanding multimedia collections. In particular, the sheer number of content description and search tools published in the literature clearly shows that multimedia analysis has reached maturity. So why haven't we real-world, commercial systems, facilitating multimedia analytics over news data?

To better understand why practitioners have not adopted the latest technology and to foster new research directions in media analytics targeting users' needs, this paper studies current practices and expectations in the news media business. In contrast to many other papers, the study here is not technologically oriented but rather user-oriented. It does not expose users to the vast bestiary of existing tools, asking them to imagine what they could be doing with these tools. The study in this paper goes the other way around and adopts a general human-centered design approach for interactive systems. We use here an ergonomic analysis based on interviews of practitioners, pursuing multiple goals via media analytics. First, we want to gain a better understanding of potential users and of their professional activity to identify relevant functionalities to develop. Second, we want to measure the perceived usefulness of a number of functionalities to prioritize research and development. The idea is to measure how users perceive the utility of functionalities that can be developed based on existing and upcoming technology, before actually developing a prototype system. Finally, we also want to gather suggestions and new ideas from discussions with future users, so as to shape the future of multimedia analytics.

2 Methodology

2.1 The UCD approach

The approach taken to better understand how media professionals work and how media analytics can assist them borrows from human-centered design (UCD). The ISO 9241-210 standard identifies six basic principles governing a human-centered design approach, leading to four main activities in application design: (i) understanding and characterizing the context of use: skills or habits of potential users, tasks to perform with the system, analysis of the environment in

which the tasks will be performed; (*ii*) identifying users' demands; (*iii*) producing design solutions (scenarios, mock-ups or prototypes) based on technical knowledge; (*iv*) assessing these solutions in relation to the demands. With the goal of defining major trends for (multi)media analytics and shaping the related research directions, we consider here the first three activities applied to the design of an interface to search and explore a collection of documents related to the news domain so as to gain insight and extract information.

The scenario considers large-scale collections of newswires and online newspapers, radio podcasts and videos along with comments and reactions from the web, e.g., via blogs, or on social networks. We explicitly envision two distinct uses of the application. The first case is a standard *information extraction* usage where the interface is used to search for a precise piece of information, be it a simple fact such as a date, a location, or the reaction of public persons to an event. The second case considers *information exploration* (or analytics) where the interface is typically provided to apprehend a subject or an event, gain insight and get a global understanding, e.g., to make a synthesis of an event. This is typically what press agents do when they're asked a press kit or a brief on a subject. Contrary to the search scenario, we believe that advancing the state of the art in Multimedia Analytics is needed to address this task and develop better analytics tools.

The target population for such a navigation interface is typically that of media professionals—journalists, press agents, news-related community managers and website designers/editors, politicians and their press assistant, etc.—rather than the general public whose needs are typically very heterogeneous.

2.2 Interview Stage #1: Knowing the Crowd

Over a period of one month, we interviewed 13 media professionals representing three different professions: journalists, press agents, and community managers. All analyze the press on a daily basis though with different purposes. The professions were chosen to provide a variety of practices for which we believe technology can help. Each interview was divided in four main parts. A questionnaire was first used to gain a better understanding of the interviewee's profile, in particular regarding his relation to the press and his degree of acquaintance with Internet and multimedia technology. The interview itself, led by an ergonomist, started with an analysis of the current activity and practices. Each interview lasted from 1 hour to 1:30. Thirteen professionals were selected for these user tests. This sample seems adequate compared to studies on this subject. Indeed, a user test conducted using a sample between 5 and 10 persons is enough to reveal the majority of usability problems [15].

The 13 interviewed persons were mostly male (9/13): 4 press agents, 7 journalists and 2 community managers. Average age is 32 with an average professional experience of 5 years. We measured several aspects of their technology profile. Personal innovativeness, i.e., willingness to test innovative technology [1], gets a score around 5 on a Likert scale from 0 (innovation reluctant) to 10 (innovation victim), except for community managers having an average score of

8.5. Skills in Internet browsing and information retrieval range from average for press agents (6 on a scale from 0 to 10) to high for community managers (10) with journalists in between. All practitioners make heavy use of social networks, mostly Facebook (100 %) and Twitter (91.7 %), with 58 % of the persons spending more than 5 hours daily on social networks. Search engines on the Internet are frequently used for all groups.

We also measured how traditional media sources are used. Surprisingly, TV news are not frequently used by any of the professional categories. Newspapers are regular sources for press agents and journalists who however mostly rely on the radio news, with many of them listening to radio news several times a day. News aggregation sites are not usual sources of information except maybe for some journalists where 3 out of 7 visit such sites at least once a day.

One interesting outcome of the study of practices is that media aggregation interfaces are not used (and not trusted) by media professionals. We see that as a sign that search engines and content aggregation hardly make analytics! With this in mind, we explore functionalities that could be useful for media analytics.

2.3 Interview Stage #2: Assessing Design Decisions

The second stage of the interview addressed the assessment of the design solution, focusing on the acceptability of various functionalities. Acceptability refers to an individual's perception of the value of a system or a technology. According to the technology acceptance model [3], the two most influential theoretical constructs in terms of acceptability are the perceived usefulness of a technology (i.e., the degree to which a person believes that using a given system will improve performance) and its perceived ease of use (i.e., the degree to which a person believes that using a given system will require little or no effort). In this study, we focus on perceived usefulness to evaluate more specifically each of the proposed functionalities. Finally, ranking the functionalities was proposed to conclude the interview and global suggestions were collected.

From a practical point of view, evaluation of the functionalities was done using a psychometric Likert scale from 0 (strongly disagree with the proposition) to 10 (strongly agree with the proposition). In addition, we analyzed verbatim transcripts of the interviews for a better understanding of the judgments. Note that ranking the functionalities was done interactively with the interviewees who were allowed to ask for details. We observed that this interaction was needed for persons to get a good understanding of non-conventional functionalities, thus fully justifying the use of interviews as opposed to a online questionnaire. In preliminary experiments, the latter was deemed not adequate for complex functionalities and usage significantly departing from standard practices. We shall note here that this situation might arise on a regular basis in the assessment of what multimedia analytics can do and how it can help, with foreseen practices departing from the standard search philosophy.

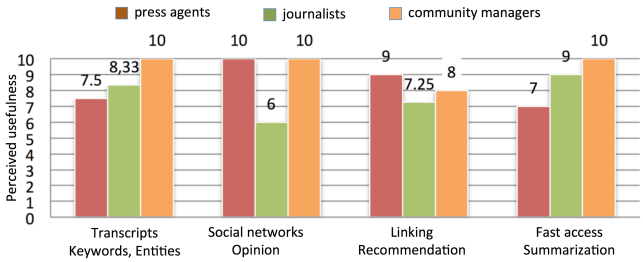


Fig. 1. Perceived usefulness of each group of functionalities.

3 Perceived Usefulness of Analytics Functionalities

During the interviews, the navigation interface was globally described to a person as a web-based system to explore the various types of media data used in his/her professional activity, offering a number of functionalities to search information, have a synthetic view of documents, see links between documents in the collection, explain those links, etc. It is important to stress again that no actual system was used at this stage of the design process. The understanding of the functionalities is thus based on explanations where the interviewee imagines how he would use the functionality rather than based on experience. Following the UCD approach, we built mockups screens such as the one illustrated in Fig. 2 to help describing a wide range of functionalities. Some of the functionalities obviously build on existing technology while others are clearly innovative if not futuristic, e.g., generating a short text that explains the links between two documents. Functionalities were grouped in four broad categories, respectively related to text and keywords, social networks and opinions, links and recommendation, and content abstraction and fast access (details are given below).

A synthetic view of the trends is given in Fig. 1, where perceived usefulness is reported for each professional profile and for each group of functionalities. Globally, for press agents, social network analysis and linking/recommendation are the two most useful functionalities. For journalists, fast access to information ranks first along with transcripts, keywords and entities. Community managers find everything useful, with a weaker interest for links and recommendation.

3.1 Transcripts, Keywords and Entities

The first group gathers fairly classical tools highlighting salient information in a document or in the collection based on language data, mostly from automatic speech transcription for radio and video content. Showing the transcript appears as the most straightforward possibility. Keywords as well as entity extraction and characterization can further be used to present key information to users.

Six persons judged displaying transcripts useful and 3 mentioned this functionality would help them save time. Displaying word clouds instead of the whole transcript was globally perceived negatively, 3 persons mentioning that there is

no interest in doing so, 2 mentioning that word clouds are not practical. Highlighting proper names was judged useful or interesting by 5 persons, but only one found interest in having direct access to the corresponding biography (if not from Wikipedia). Highlighting locations was surprisingly judged of no use by 3 while 3 mentions that using a map would be more appropriate. In the free suggestion phase, 5 persons for whom classifying keywords according to relevance and frequency would be nice, better than displaying word clouds.

From the technology point of view, this set of functionalities builds on existing tools (automatic speech recognition, named entity detection, term and keyword extraction, etc.) and common practices in the field of interfaces (word cloud, Google map localization, etc.). Comments during the interviews however highlighted the importance of the accuracy of those techniques: accuracy and robustness of the underlying content extraction tools involved should still be improved to make them fully acceptable in media analytics. Automatic processing quality assessment, i.e., saying whether or not the outcome of content analysis can be trusted, is needed. Comments on Wikipedia also highlight that external knowledge sources must be trustworthy to be of interest for media analytics.

3.2 Social Networks and Opinions

A number of functionalities related to opinion mining in social networks were also considered, ranging from standard and existing opinion mining techniques targeting valence (positive, negative, neutral) to upcoming fine-grain characterization techniques. In particular, recent work in opinion mining addresses the identification of the emotions (anger, surprise, fear, etc.) expressed in addition to valence [6] and it is expected that these techniques will be mature in the near future. The identification of the aspects on which people are reacting also appears as a promising functionality. Typically, on a news item such as the Strauss-Kahn Sofitel scandal, people reacted on several aspects: the main offender, the director of the International Monetary Fund, the candidate to the French presidential election, etc. After identifying the different aspects targeted in social networks, opinions and feelings can be analyzed, displayed and synthesized for each aspect individually to have a better synthetic view of social reactions.

The group of functionalities was globally judged interesting and useful, with 5 persons mentioning usefulness and 3 time gain. Perceived usefulness is particularly high for press agents, while journalists are more doubtful. The ability to analyze the evolution of opinions over a period of time was well perceived in terms of usefulness, with 5 persons mentioning this functionality. Extracting opinions and feelings, whether globally on a subject or more precisely for an event, was also found interesting. Looking at feelings appears more interesting at a global level while opinions seem more interesting for a precise event. Unsurprisingly, opinion analysis in traditional media (as opposed to social media) is judged of limited interest. The lack of confidence in the outcome of automated opinion analysis was again mentioned on several occasions. Finally 5 persons suggested a functionality implementing filters to sort comments (by social network, by keywords, by number of retweets, likes, followers, etc.).

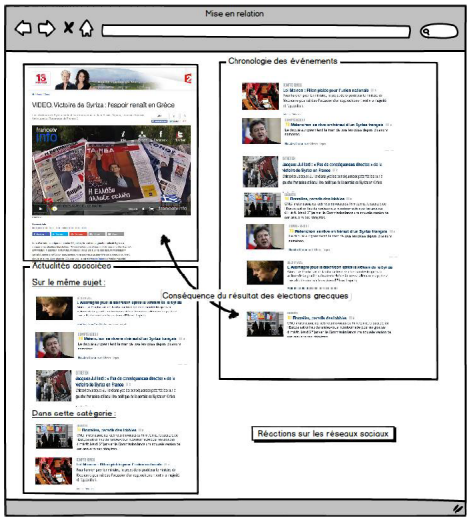


Fig. 2. Mockup user interface illustrating the functionalities related to links and recommendation. The current item appears in the top-left corner. Below are links to items on the same subject; The right panel gives a timeline related to the current item. Explanation of a link is illustrated with the text 'Consequence of the Greek elections'.

There clearly is a high demand for social network analysis to extend the classical sources of information of media analytics. There is apparently a clear interest to go further into the characterization of opinions and feelings, beyond valence. This is an ongoing research topic and confidence in the result indicates that progress are still required for this technology to be considered as mature for social media analytics: better ressources (lexicons, tree-banking of social texts, etc.), fine grain lexical and syntactic analysis on degraded language are among the elements that will shape the future. As for the first group of functionalities, being able to measure the confidence in the outcome of processing modules is vital. Explaining the decision, which is something out of reach of the current technology, appears like a good option to help with this issue.

3.3 Links and Recommendation

Creating explicit links between documents in a collection provides a set of functionalities to group similar items, recommend content related in some way to a document or provide a chronological thread of an event. We also consider explicit linking as a potential prerequisite for analytics, exploiting generic graph analytics techniques and knowledge propagation.

Links can be established on a number of grounds—e.g., same content, same event, same topic, same person involved—and organized in various ways: chronological order to show the evolution of an event, reactions to an event, causes and

consequences. To help better apprehend why links are established, link characterization can further be used to inform users of the interface of the meaning of the link. The mockup screen in Fig. 2 illustrates these functionalities.

Providing links and recommendations with documents do not appear as a priority. We asked people to rank the types of links according to perceived usefulness and obtained the following ranking: by topic, keywords, date, location, etc. Linking by dates and/or locations was however mentioned several times as useful during the interviews. Regarding link explanation, journalists found fine-grain explanation of why the link is provided more useful than a simple coarse grain explanation such as a type (average rating of 8.5 vs. 6.3). Limiting redundancy by grouping highly similar documents wasn't judged as necessary. However, press agents rank the functionalities 'grouping similar documents' and 'highlight key/central documents' with a perceived utility of 10. Interestingly, these two features can easily be implemented using graph analytics should the data be organized as a graph. A suggestion, made by two persons, is to emphasize links to raw information sources, i.e., not processed by (other) journalists.

While the technology to compare content exists for all modalities, little has been done to create links on a large scale in collections. Evaluations on video hyperlinking over the past few years, e.g., [5], are typically first step in this direction. Clearly, analytics cannot satisfy itself of documents taken regardless of the collection. The question of knowing what usage for what types of links and structure (k-nn graphs, threads, etc.) thus appears as a crucial one for analytics. Fine grain semantic links are still missing, along with the ability for the machine to explain why a link was established, to facilitate navigation.

3.4 Fast Access to Information

The last group of functionalities gathers a number of features whose goal is to provide a more efficient access to the information and to ease a global view of the collection (or of part of the collection). Apart from the classical table of content and search engine, several rapid access features were considered: showing a timeline of the collection or of a selected event, providing a clickable word cloud depicting the whole collection, grouping similar documents along with a summary, or highlighting key documents.

While fast access to information was globally judged as a useful set of functionalities, we were able to gain limited insight on the detailed functionalities. Topic segmentation was globally perceived as useful and time saving. Links from topic segments to the corresponding (transcription of the) content were suggested when topic segmentation was mentioned. Several persons also found interesting displaying a chronological record of an event. In contrast, summarizing documents or groups of similar documents was judged of no interest: one reason given is that machine interpretation of content cannot be trusted by professionals.

Most of these functionalities rely on existing technology and interface design. Summarization, in particular in the context of multiple documents from multiple modalities, still needs improvement in order to be accepted as part of an interface for media analytics. Yet, the perceived usefulness of these functionalities remain

limited though we believe they are all relevant for access and fast selection of relevant information, an important feature of multimedia analytics. Progress in interface design to accommodate these functionalities might be considered for users to find utility in them.

4 Lessons for Multimedia Analytics

Multimedia analytics could be defined as the *process of organizing multimedia data collections and providing tools to extract knowledge, gain insight and help make decisions by interacting with the data*. This definition clearly goes beyond traditional multimedia analysis whose goal is to describe content, usually for indexing purposes, and requires organizing collections of data. Studying users practices and expectations in media analytics sheds light on the design of such applications, on technology requirements and on research directions to pursue.

Social Networks. A first important result that was highlighted by the interviews is the importance of social networks, at least in the news domain. While this does not come as a surprise, for news as for any other domain [22], this state of fact calls for tools to characterize opinions and feelings in social networks at a fine grain: better resources are required for fine grain semantic characterization; improved word representations (e.g., derived from word embeddings [20, 14]) along with the corresponding learning machinery; etc. Apart from opinons and sentiments, on which we focused in this paper, a huge amount of information can be extracted from social signals, e.g., for event detection (see, e.g., [19], to cite a famous example), fact checking, etc.

Heterogeneity of Information Sources. Regardless of the scientific challenges for NLP and machine learning, the importance of social media in the news domain (and in many other domains) points out to the variety of information sources that need to be ingested and connected within multimedia analytics applications. In turn, this raises a number of scientific challenges regarding heterogeneous and distributed information integration. There are no clear model to do so as of today and mostly ad-hoc solutions are adopted, often with probabilistic modeling. Mathematical tools for the joint analysis of modalities and/or sources are still poor as of today. We will come back to this point later in the discussion with the notion of *knowledge*.

Reliablity of Sources. With the variety of sources of information comes the notion of reliability of the information, which is clearly a crucial point. This is obviously true for media professionals who prefer unedited sources, who need to know who published and who read, etc. We believe that this is also true in most application domains and that analytics interfaces must implement a trust mechanism of some sort. From a multimedia modeling point of view, we see a number of implications. First, information automatically extracted need be reliable. At the very least, an automatic decision should come with a confidence. Confidence measures have been used for a long time in some domains, e.g., speech recognition [10], but need be generalized to all content-based analysis tools. Another possible option is the ability for a classification system to explain on what ground

the decision was made. Early attempts in this direction, e.g., recounting [4] or evidence [16], are encouraging and, again, should be generalized. But the current trends in big data machine learning goes in a different direction with neural models acting as black boxes. Second, content-based analysis can be used to validate information. This is clear from current trends in data journalism and fact checking [17]. These are recent research areas, mostly unexplored, implementing analytics on a limited scale. Here again, trust and confidence are crucial.

Contextualization. Another point that appears meaningful is the need for contextualization in general. Showing and analyzing a document must take into account the metadata context (authors, number of views, etc.) and the context of the collection: tweets often make sense in the context of more global threads, video fragments can hardly be understood outside the chronology of events they illustrate. Links between parts of documents form a nice basis to keep track of the context. From the technology point of view, organizing collections according to links materializing the relations between items within the collection as discussed in Sec. 3.3 is convenient for context-aware analysis, navigation and visualization. While the links themselves are not judged relevant by professionals for recommendation, they could easily be used for contextual display, a feature that was judged as highly relevant.

Dynamic Management of Knowledge. With the goal of analytics being to gain insight and knowledge, there is a need for research on knowledge representation in tight cooperation with content analysis and interpretation. We believe that multimedia modeling and collection organization should evolve as more knowledge is gained. To do so, we need knowledge representation mechanisms that can serve content-based analysis and, conversely, content-based analysis able to handle evolving knowledge. While knowledge representation and content-based modeling have mostly been two separate fields in the past, we can anticipate that they will converge, to some extent. Recent work on image processing [2] and natural language processing [18] hint in this direction. Multimedia analytics certainly goes with a trend towards using knowledge available as linked open data. Note that information trust remains an issue. Finally, as knowledge evolves with the process of multimedia analytics, a constant back and forth between data description and organization, on the one hand, and knowledge extraction via analytics on the other hand, is required. This in turn requires multimedia models and content-based description algorithms to adapt and evolve, e.g., relying on active learning techniques.

Conclusion. To conclude this discussion, let us highlight a few points that were left out of the study we performed (and thus out of the discussion) but that we judge very relevant. First of all, the dynamicity (aka velocity) of the data has been totally ignored. Many collections evolve at a very fast pace, in particular on social media. How do applications adapt to the constant stream of data and, more generally, to increasing knowledge? This question remains largely open and few multimedia modeling tools can today handle dynamic collections and knowledge evolutions. Second, user interfaces have also been disregarded to focus on functionalities regardless of their inclusion in an actual interface.

All interviewees in our study insisted on the ease of use of media analytics functionalities, calling for research on interfaces and navigation in multimedia databases, e.g., to handle knowledge-aware multi-faceted view of heterogeneous and multimodal data. Recent work such as [11] on photo browsing goes in this direction and require generalization to multimodal and heterogeneous data, a challenge for the database world.

As we can see from this discussion, the future of media analytics goes beyond the sole multimedia community, standing at the intersection of multiple domains including databases, knowledge discovery, or human-computer interaction. Targeted users should also be actively involved in the process of shaping-up multimedia analytics. In the near future, efforts should be made to put all these actors together.

Acknowledgments

This work was funded via the CominLabs excellence laboratory financed by the National Research Agency under reference ANR-10-LABX-07-01.

References

1. R. Agarwal and A. Prasad. A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, 9(2):204–215, 1998.
2. J. Atif, C. Hudelot, and I. Bloch. Explanatory reasoning for image understanding using formal concept analysis and description logics. *IEEE Transactions on Systems, Man and Cybernetics: Systems*, 44(5):552–570, May 2014.
3. F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8):982–1003, 1989.
4. D. Ding, F. Metze, S. Rawat, P. F. Schulam, S. Burger, E. Younessian, L. Bao, M. G. Christel, and A. Hauptmann. Beyond audio and video retrieval: Towards multimedia summarization. In *Proc. ACM International Conference on Multimedia Retrieval*, 2012.
5. M. Eskevich, G. J. F. Jones, R. Aly, and et al. Multimedia information seeking through search and hyperlinking. In *ACM Intl. Conf. on Multimedia Retrieval*, 2013.
6. A. Fraisse and P. Paroubek. Toward a unifying model for opinion, sentiment and emotion information extraction. In *Intl. Conf. on Language Resources and Evaluation*, 2014.
7. M. Ghoniem, D. Luo, J. Yang, and W. Ribarsky. NewsLab: Exploratory broadcast news video analysis. In *IEEE Symposium on Visual Analytics Science And Technology*, pages 123–130, 2007.
8. A. G. Hauptmann and M. J. Witbrock. Informedia: News-on-demand multimedia information acquisition and retrieval. *Intelligent multimedia information retrieval*, pages 215–239, 1997.
9. I. Ide, H. Mo, N. Katayama, and S. Satoh. Topic threading for structuring a large-scale news video archive. In *Intl. Conf. on Image and Video Retrieval*, pages 123–131, 2004.

10. H. Jiang. Confidence measures for speech recognition: A survey. *Speech communication*, 45(4):455–470, 2005.
11. B. T. Jónsson, Á. Eiríksdóttir, Ó. Waage, G. Tómasson, H. Sigurthórsson, and L. Amsaleg. M3+ p3+ o3= multi-d photo browsing. In *MultiMedia Modeling*, pages 378–381, 2014.
12. H. Lee, A. F. Smeaton, N. E. O'Connor, and B. Smyth. User evaluation of Físchlár-news: An automatic broadcast news delivery system. *ACM Trans. on Information Systems*, 24(2):145–189, 2006.
13. A. Merlino, D. Morey, and M. Maybury. Broadcast news navigation using story segmentation. In *ACM Intl. Conf. on Multimedia*, pages 381–391, 1997.
14. T. Mikolov, W.-t. Yih, and G. Zweig. Linguistic regularities in continuous space word representations. In *HLT-NAACL*, pages 746–751, 2013.
15. J. Nielsen and T. K. Landauer. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pages 206–213, New York, NY, USA, 1993. ACM.
16. J. Poignant, H. Bredin, and C. Barras. Multimodal person discovery in broadcast tv at mediaeval 2015. In *Working Notes Proc. of MediaEval 2015 Workshop*, 2015.
17. J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer. Detecting and tracking political abuse in social media. In *ICWSM*, 2011.
18. G. Rizzo and R. Troncy. NERD: A framework for unifying named entity recognition and disambiguation web extraction tools. In *Conf. of the European Chapter of the Association for Computational Linguistics*, 2012.
19. T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, pages 851–860. ACM, 2010.
20. J. Turian, L. Ratinov, and Y. Bengio. Word representations: a simple and general method for semi-supervised learning. In *Proceedings of the 48th annual meeting of the association for computational linguistics*, pages 384–394, 2010.
21. X. Wu, C.-W. Ngo, and Q. Li. Threading and autodocumenting news videos: a promising solution to rapidly browse news topics. *IEEE Signal Processing Magazine*, 23(2):59–68, 2006.
22. D. Zeng, H. Chen, R. Lusch, and S.-H. Li. Social media analytics and intelligence. *IEEE Journal on Intelligent Systems*, 25(6):13–16, 2010.