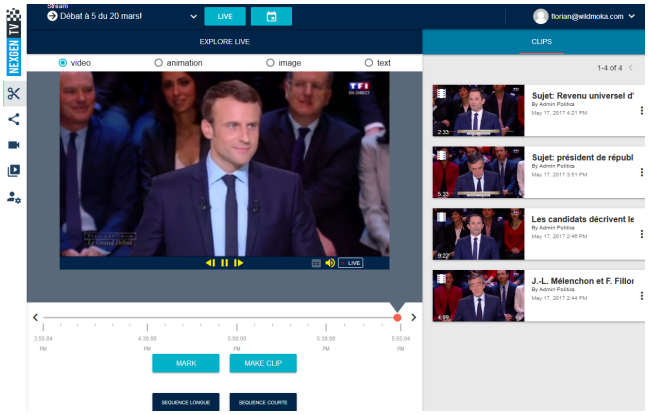

Figure 1: Screenshot of the back-office interface.

noun adj) are extracted and normalized to account for small variations. Each of the resulting candidates is evaluated based on its frequency and discriminativeness, as measured with the Okapi-BM25 weighting scheme [8]. Normalized compound sequences are finally ranked from the most relevant to the least.

Content hyperlinking. One distinguishing feature of second screen applications is the ability to provide links from a clip to other related clips, a task that we refer to as hyperlinking [1, 4, 5]. In NexGenTV, hyperlinking exploits subtitles. State-of-the-art document embedding [7], trained on French newspapers, is used to represent each clip by a vector. Using a database of such clip vectors, approximate nearest neighbor search techniques are used to efficiently retrieve on the fly a small number of related entries in the database for a given clip, the latter acting as a query.

Tweet collection and analysis. Social media is an important source of insight for broadcasted political events. A large number of persons share their views on the debate in quasi real-time, making it possible to analyze the TV broadcast in light of social reactions. Messages related to a program are collected on Twitter based on predefined hashtags and keywords. Relevant named entities are then extracted combining regular expressions and conditional random fields. Sentiment analysis [6] is also performed on each tweet, allowing to detect the polarity (neutral, positive, negative or mixed) relying on a recurrent neural network. This approach has been shown to obtain state-of-the-art results while being robust enough to handle grammatical variability.

3 INTERFACES

The NexGenTV platform has two interfaces: the back-office desktop interface allows TV channels to easily select and enrich content, partially automating the selection and enrichment of key clips to push; the front-end consists of a mobile application receiving selected content and enrichment from the back-office.

TV streams are ingested within the back-office where the main screen provides single-click clip selection, exploiting the result from speech and speaker turn detection: options are provided to select long or short clips and to easily adjust the boundaries of the clip. A list of clips already selected also appears in the main screen, as illustrated in Fig. 1. After selecting a clip, the edition interface enables adding information to the clip before pushing the enriched clip to the front-end application. An initial description of the clip is automatically generated thanks to term extraction from subtitles, entity linking leveraging from face recognition and topic characterization [3] exploiting an ontological description of political life during the 2017 presidential election. Combining these elements, links can also be made to a description of the political program of the candidate for the clip's topic. Content linking also provides potential links to related clips or to additional sources previously ingested, e.g. previous debates or news shows. The back-office editing screen enables to re-arrange all these elements and the selection of relevant links before publication to the front-end.

In the front-end interface, new clips appear on the timeline as they are published from the back-office. They can be viewed, along with their description and with links to entries in our knowledge base (e.g., the bio of a politician, description of events of interest, program of a candidate on a given topic) and to related content. The front-end application also enables to view real-time statistics on the debate: time of speech or visual presence per candidate, amount of reactions on Twitter with clips at key instants, popularity for each candidate from opinion mining.

The demonstration operates on a collection of about 192 h of videos broadcasted by major French TV channels, totaling around 100 h of political debates and 92 h of TV news for enrichment. The average length of a debate is approx. 3 h, varying from less than 30 m to almost 4 h. The database for enrichment of a clip with hyperlinks consists of clips that were extracted from the debates and from the news shows. The ontology of the 2017 presidential election was created from scratch and describes the bio of all politicians involved, a list of topics discussed in the campaign (relation with EU, wealth tax, etc.) and synthetic records of the political program of each candidate on each topic.

REFERENCES

[1] Evlampios Apostolidis, Vasileios Mezaris, Mathilde Sahuguet, Benoit Huet, Barbora Červenková, Daniel Stein, Stefan Eickeler, José Luis Redondo Garcia, Raphaël Troncy, and Lukás Pikora. 2014. Automatic fine-grained hyperlinking of videos within a closed collection using scene segmentation. In *ACM Multimedia 2014*. ACM, 1033–1036.

[2] Mathieu Ben, Frédéric Bimbot, and Guillaume Gravier. 2005. A model space framework for efficient speaker detection. In *European Conference on Speech Communication and Technology*.

[3] Vincent Claveau and Sébastien Lefèvre. 2015. Topic segmentation of TV-streams by watershed transform and vectorization. *Computer Speech and Language* 29, 1 (2015), 63–80. https://doi.org/10.1016/j.csl.2014.04.006

[4] Maria Eskevich, Gareth J.F. Jones, Shu Chen, Robin Aly, Roeland Ordelman, Danish Nadeem, Camille Guinaudeau, Guillaume Gravier, Pascale Sébillot, Tom De Nies, Pedro Debevere, Rik Van de Walle, Petra Galuščáková, Pavel Pecina, and Martha Larson. 2013. Multimedia Information Seeking through Search And Hyperlinking. In *Intl. Conf. on Multimedia Retrieval*.

[5] Maria Eskevich, Huynh Nguyen, Mathilde Sahuguet, and Benoit Huet. 2015. Hyper Video Browser: Search and Hyperlinking in Broadcast Media.. In *ACM Multimedia 2015*. ACM, 817–818.

[6] Grégoire Jadi, Vincent Claveau, Béatrice Daille, and Laura Monceaux-Cachard. 2016. Evaluating Lexical Similarity to build Sentiment Similarity. In *Language and Resource Conference, LREC*. portoroz, Slovenia.

[7] Quoc V. Le and Tomas Mikolov. 2014. Distributed Representations of Sentences and Documents. In *International Conference on Machine Learning*.

[8] Stephen Robertson and Hugo Zaragoza. 2009. The Probabilistic Relevance Framework: BM25 and Beyond. *Found. Trends Inf. Retr.* 3, 4 (April 2009), 333–389. https://doi.org/10.1561/1500000019

[9] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 815–823.

[10] Shujun Yang, Lei Pang, Chong-Wah Ngo, and Benoit Huet. 2016. Serendipity-driven Celebrity Video Hyperlinking. In *ACM ICMR'16*. ACM, New York, NY, USA, 413–416. https://doi.org/10.1145/2911996.2912029