



xLiMe – *crossLingual crossMedia knowledge extraction*

Title of the presentation: Cross-Lingual Cross-Media Knowledge Extraction from Live Media Streams¹

Type of the presentation: Research contribution (with system demonstration)

Summary of the presentation: We will demonstrate our recent achievements in semantic analysis of large-scale real-time media streams. On the one hand we focus on bridging the varieties of data: colloquial (social media) vs. well-formed (news articles) text, different languages and different modalities (video, audio, text). On the other hand we scale to high data velocities: hundreds of Live-TV channels, hundred thousands of news articles per day and millions of social media documents per day. Our prototypes demonstrate its use in real-world applications, using semantic search across media channels, modalities and languages or content-based recommendations to viewers of Live-TV.

Contributor names and short CVs

Achim Rettinger: Dr. Achim Rettinger is a junior research group leader at AIFB. He did his PhD studies in computer science at the Technische Universität München, Germany, and at the Siemens AG in Munich, Germany. Among the research projects he is and was involved in are xLiMe (coordinator), XLike, SFB CGS (DFG), Theseus and LarkC. Before he worked on research projects at the Osaka University (Japan), University of Bath (UK), University of Alberta (Canada) and University of Georgia (USA). His research interests and publications are in combining machine learning, knowledge discovery and human computer systems with semantic technologies. He organized workshops on cross-lingual technologies (xLiTe) and published on text mining, topic modelling, recommendation engines and statistical analysis of structured data.

Marko Grobelnik: Marko Grobelnik (m) is an expert in the areas of analysis and knowledge discovery in large complex databases. Marko collaborates with major European and US academic institutions and consults industries such as British Telecom, Microsoft Research, Nature, New York Times, Bloomberg, and Accenture. Marko is author of several books in the area of machine learning, data mining, text mining and semantic technologies and authors of many scientific papers. He is also W3C AC representative for IJS, CEO of the company Quintelligence and co-founder of the company Cycorp Europe. Marko served also as a program chair for European Machine Learning conference (ECMLPKDD 2009)

¹ The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 611346. (<http://www.xlime.eu>)

and for European Semantic Web Conference (ESWC 2011). In terms of the past project experience, Marko has been technical coordinator for projects FP6 IST-World and FP7 VIDI and coordinator of FP7 project X-LIKE (Cross-Lingual Knowledge-Extraction); he was a member of project management board in several FP6 and FP7 Projects (SEKT, NEON, ACTIVE, COIN).

Nicu Sebe: Nicu Sebe is a Professor of Computer Science and he is leading the research in the areas of multimedia analysis and human-computer interaction in computer vision applications. He has been involved in the organization of the major conferences and workshops addressing the computer vision and human-centered aspects of multimedia information retrieval, among which as a General ACM Multimedia 2013 and ACM International Conference on Image and Video Retrieval (CIVR) 2017 and 2010. He was the program chair of ACM Multimedia 2011, a track chair of WWW 2009 and ICPR 2010 and will be a program chair of ECCV 2016 and ICCV 2017. He has participated in several EC projects.

Extended abstract of the presentation

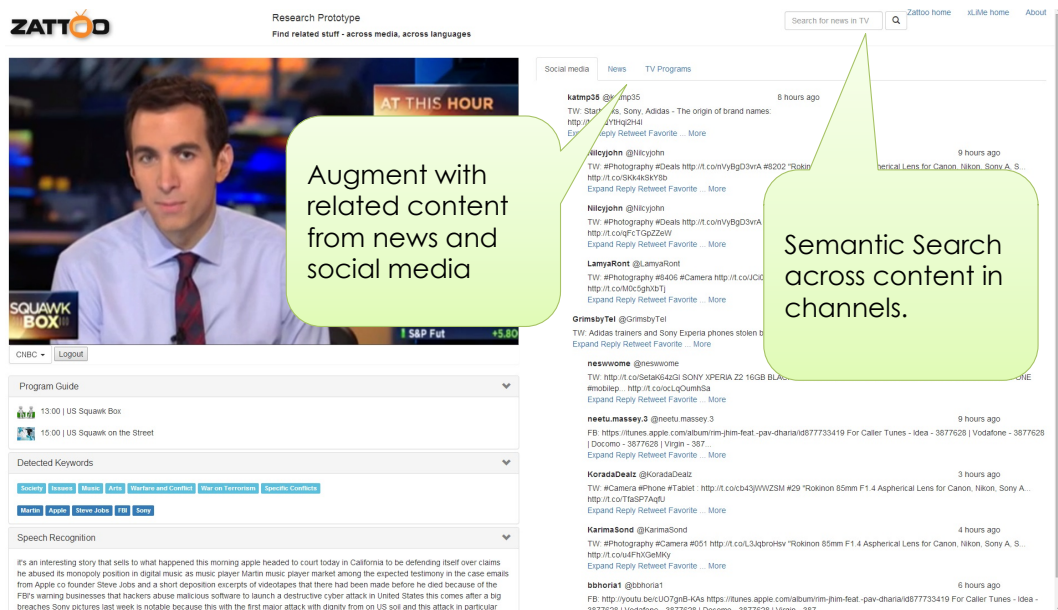
Europe is different from other large media markets such as the US or China in that information is being generated in different languages and distributed via diverse streams of localized media channels. Automatic analysis is complicated further by different content types (audio, video, text) and different channels (mainstream, social media). Thus, information can only be analysed independently for each dimension. This restricts the extractable knowledge and keeps it fragmented, which ultimately constrains the exchange of information.

In this presentation we show our recent achievements in extracting knowledge from different media channels and languages and relate it to cross-lingual cross-media knowledge bases. By doing this in near real-time we provide a continuously updated and comprehensive view on knowledge across media in Europe.

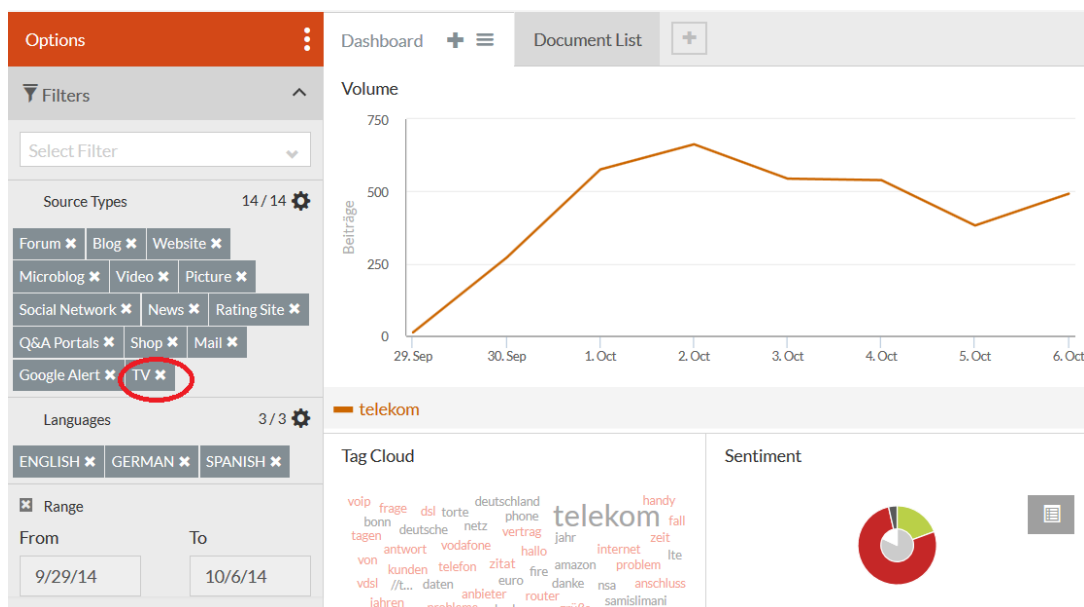
We achieve this by extending approaches from research on Written and Spoken Language Processing, Computer Vision and Knowledge Graphs in several directions. More than 15 scientific publications resulted from this research, and provide details on how this has been achieved with respect to various research challenges. All sub-systems have been integrated into one with the following main components:

- We developed a data model suited to describe and integrate all data produced by the different sources. This includes TV data, news and social media, plus accompanying information like provenance.
- We implemented a software architecture capable of processing and distributing this data from the data providers to the consumers in near real-time.
- We provided services to extract text from the various modalities and sources, including speech-to-text and OCR.
- Also initial tools for object and brand recognition from visual content are provided.
- All of this content is then being processed and integrated on two different levels. Firstly on a textual basis and secondly on a semantic level by representing the content of media items as graphs. In addition, early research has been conducted to learn a cross-modal latent representation, which bridges the gap between the textual and semantic level.
- The final technical achievement allows for querying the content by semantic and keyword search, including additional background knowledge taken from the DBpedia knowledge graph.

The core real-world tasks that have been addressed with this system is real-time content-based recommendations across modalities and languages and semantic search across all media channels



In our presentation we will demonstrate how this system can augment a Live-TV stream with related content from social media, online news and other TV channels in real-time. The figure above depicts the interface for TV viewers with related social media posts.



In addition we will showcase how the system allows complex semantic queries over the content of various media channels and languages, like appearances of cars of a certain manufacturers in the European media stream. This allows media monitoring and trend-mining applications as depicted in the figure above.

This complex real-time big data processing pipeline opens up opportunities in terms of solving challenging research questions and building innovations with societal and economic impact. Already now, the benefits for media monitoring and delivering more relevant content to TV users are becoming obvious.