xLiMe

xLiMe – crossLingual crossMedia knowledge extraction

MISSION

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.

xLiMe proposes to extract 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 will provide a continuously updated and comprehensive view on knowledge diffusion across media, e.g., from European communities like Catalonia to worldwide content in English.
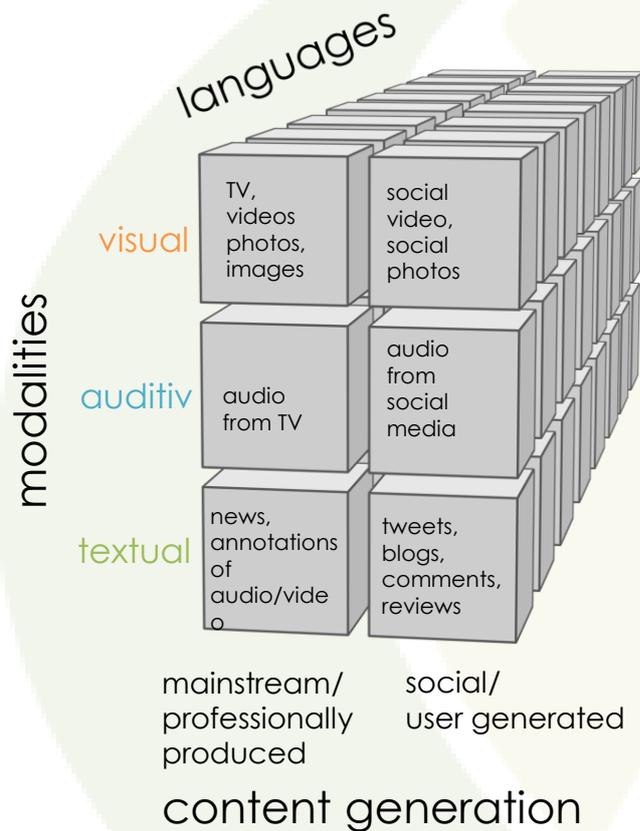
EXPECTED OUTCOME

Tools and methods developed in xLiMe will be applied in three complementary case studies and evaluated by several business clients and up to 10mio end users.

We will

1. augment more than 250 TV channels in different languages with up-to-date information from social media and news,
2. monitor brands and the diffusion of opinions across languages and media, and
3. analyse online shop performance with regard to external cross-lingual, cross-media factors, like campaigns for brands and the emergence of public opinions.

By combining speech recognition, natural language processing, machine learning and semantic technologies we will advance key open research problems, by

1. extracting machine-readable know-ledge (entities, sentiment, events, opinions) from multilingual, multi-media and social media content and integrate it with cross-lingual, cross-media knowledge bases,
2. searching this knowledge with structured and unstructured queries in near real-time,
3. monitoring its provenance, con-sumption and diffusion and
4. analysing the interdependency between media exposure and behavioural patterns.

www.xlime.eu

languages

modalities

visual

TV, videos photos, images

social video, social photos

auditiv

audio from TV

audio from social media

textual

news, annotations of audio/video

tweets, blogs, comments, reviews

mainstream/ professionally produced

social/ user generated

content generation

The goal of xLiMe is to develop technologies to extract and integrate knowledge from 3 different content dimensions, which in most approaches have been treated independently from each other: (i) Content generated by and for different channels and formats (e.g., traditional mainstream media vs. social media), (ii) different content types (audio, video, text) and (iii) different languages.

ADMINISTRATIVE DETAILS

Start	November 2013
End	October 2016
EC Contribution	2,987,000 €
Project Coordinator	achim.rettinger@kit.edu
Project Manager	anja.hess@kit.edu

PARTNERS

www.xlime.eu