MixedEmotions: Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics Markets

ICT-15-2014 Big data and Open Data Innovation and take-up - Innovation Action

Industry Partners	Academic Partners
Expert System, Italy	Insight Centre for Data Analytics, National University of Ireland, Galway (coordinator)
Millward Brown, Czech Republic	
Paradigma Tecnológico, Spain	Universidad Politécnica de Madrid, Spain
Phonexia, Czech Republic	University of Passau, Germany
SindiceTech, Ireland	Brno University of Technology, Czech Republic
Deutsche Welle, Germany	

Project duration: April 2015 - March 2017

Summary

Emotion analysis is central to tracking customer and user behaviour and satisfaction, which can be observed from user interaction in the form of explicit feedback through email, call centre interaction, social media comments, etc., as well as implicit acknowledgment of approval or rejection through facial expressions, speech or other non-verbal feedback. In Europe specifically, but increasingly also globally, an added factor here is that user feedback can be in multiple languages, in text as well as in speech and audio-visual content. This implies different cultural backgrounds and thus different ways to produce and perceive emotions in everyday interactions, beyond the fact of having specific rules for encoding and decoding emotions in each language.

Making sense of accumulated user interaction from different ('mixed') data sources, modalities and languages is challenging and has not yet been explored in fullness in an industrial context. Commercial solutions exist but do not address the multilingual aspect in a robust and large-scale setting and do not scale up to huge data volumes that need to be processed, or the integration of emotion analysis observations across data sources and/or modalities on a meaningful level, i.e. keeping track of entities involved as well the connections between them (who said what? to whom? in the context of which event, product, service?)

The MixedEmotions project will implement an integrated Big Linked Data platform for emotion analysis across heterogeneous data sources, languages and modalities, building on existing state-of-the-art tools, services and approaches that will enable the tracking of emotional aspects of user interaction and feedback on an entity level. The platform will provide an integrated solution for:

- Large-scale emotion analysis and fusion on heterogeneous, multilingual, text, speech, video and social media data streams, leveraging open access and proprietary data sources, exploiting also social context by leveraging social network graphs
- Semantic-level emotion information aggregation and integration through robust extraction of social semantic knowledge graphs for emotion analysis along multidimensional clusters

The platform will be developed and evaluated in the context of three cross-domain pilot projects that are representative of a variety of data analytics markets: Social TV, Brand Reputation Management, Call Centre Operations.