

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

Nitish Aggarwal and Paul Buitelaar

Insight Centre for Data Analytics, National University of Ireland, Galway, Ireland  
firstname.lastname@insight-center.org

## 1 Summary

Emotion analysis is central to tracking customer behavior and satisfaction, which can be observed from their interaction in the form of explicit feedback through email, call center 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 also globally, an added factor here is that user feedback can be in multiple languages, in text as well as in 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 an industrial context. Commercial solutions exist but do not address the multilingual aspect in a robust and large-scale setting. MixedEmotions project aims provide a big data platform that integrates emotion analysis across different data sources and 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?).

In MixedEmotions<sup>1</sup> we are implementing an integrated Big Linked Data platform that will provide:

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

---

<sup>1</sup> <http://mixedemotions-project.eu/>

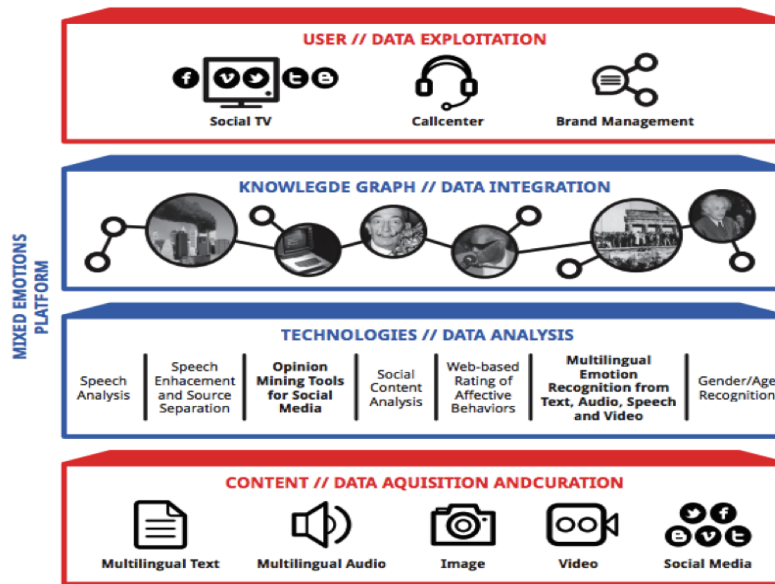


Fig. 1. Basic Architecture of the MixedEmotions Platform

Figure 1 illustrates the MixedEmotions platform that extracts data from heterogeneous data sources and perform several data analytics to build an integrated social knowledge graph. The platform will be evaluated in the context of three Pilot projects that represents a variety of data analytics markets: Social TV, Brand Reputation Management, and Call Centre Operations. Each of the companies involved in the pilots have specific innovation objectives as follows:

**Pilot I (Social TV):** “Expert System” aims at extending its text mining API “COGITO”<sup>2</sup> with added features for Social Multimedia Analytics. MixedEmotions will enable the inclusion of entity-level emotion and sentiment analysis, and Social TV (social networks activity related to TV programs).

**Pilot II (Brand Reputation Management):** Paradigma Tecnológico seeks to enhance their Online Brand Reputation Tool in three main areas: enhance real-time capabilities, multilingual data analysis and report generation to increase internationalization of the product. As all of these are central to MixedEmotions, its outcomes will enable Paradigma to provide better reputation management services to its customers.

**Pilot III (Call Centre Operations):** Phonexia offers solutions for call centre operations which will benefit from the MixedEmotions platform in the following ways: enable an advanced extraction of emotions from speech, and enhanced multilingual processing.

We will demonstrate an initial version of the platform that includes different data analytics over extracted heterogeneous data for its three pilot projects.

<sup>2</sup> <http://www.intelligenceapi.com>