Publishable summary

About

Conversational interaction is the most natural and persistent paradigm for business relations with customers. In contact centres millions of calls are handled daily. On social media platforms millions of blog posts are exchanged amongst users.

Can we make sense of such conversations and help create assets and value for private and public organizations' decision makers? And indeed for anyone interested in conversational content?

The overall goals of the SENSEI project are twofold. First, SENSEI will develop summarization/analytics technology to help users make sense of human conversation streams from diverse media channels. Second, SENSEI will design and evaluate its summarization technology in real-world environments, aiming to improve task performance and productivity of end-users.

Target User Groups

SENSEI's end-user groups are from the contact center and social and news media application domains. In the contact centre domain, the end-users of summarization analytics will be data analysts, quality control professionals and managers. In the media domain, the end-users of summarization analytics will be news comment readers, news comment authors, journalists and editors/media analysts.

Objectives and Outcomes

SENSEI's scientific and technological objectives are to develop new technologies that will empower users to make sense of conversations through the following advances:

- Parse human conversations for both content, affect and other behavioural traits.
- Create adaptive technology to address the diversity and velocity of the media sources.
- Automatically generate human-readable multimedia, graphical and tabular summaries of dialogues and/or multiparty conversations.
- Evaluate technology where it is being used and not only in the lab. We will engage end-users ranging from language data analysts to quality assurance professionals and news media analysts in real task settings.

Work performed and the main results

The main three results in first year of the project are the following:

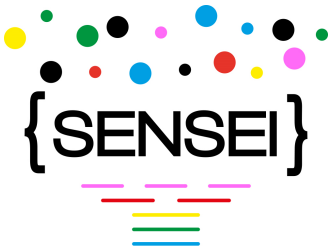
Analytics Summarization use case design and validation. A fundamental and unique achievement has been the empirical investigation and participatory development of the summarization end-user use cases. We believe this first study on the summarization technology requirements will boost the interest of industries and applicability of researchers' results. In the speech technology there is a operational need for such technology and we have identified professional end-user categories and needs. The outcome of the empirical investigation has generated a set of requirements for the summaries to be operational within companies' processes. In the case of social media conversations, the need for summarization technology is latent and the impact of summarization technology maybe a game-changer for the social media industry. Both in the speech and social media domain, we have designed a roadmap for the development of summarization technology driven by end-user participation. We have focused on identifying communities of potential users, collecting their requirements, and selecting multiple use cases both for speech and social media domains. At the same time we have (1) developed the evaluation model that will be applied throughout project, (2) identified the evaluation scenarios based on the use cases and on the evaluation method, and (3) set up the methodology for acquiring baseline parameters for the evaluation.

Unified Data Schema for Social Media Conversations. As part of our goal to summarize the conversations over social media, we first need to represent it. We have studied and observed conversation sources from diverse news media channels, different languages and geographical and cultural settings, and across domain topics. The result of such consortium-wide investigation and deliberation has provided a general representation of social media conversations. Such representation will be published as part of our deliverable to the general and technical public. Inside SENSEI, such data schema has been tested and is used as part of our process of data acquisition, processing and repository management.

Conversation Summarization prototype for Spoken Conversations. We have developed the first summarization prototype that processes call center spoken conversations. The spoken conversation are real-life recordings of customer calling into the contact center of a public service company. The prototype provides a web user interface to retrieve, view the customer-operator transcript and listen to the audio of the call-center conversations. In addition, it features an automatic summarization component which generates textual synopses of what happened during a conversation. The component showcases the backbone of the components that will support the development of SENSEI prototype versions: the general purpose conversation repository that stores all the data and metadata, the conversation viewer, and multiple baselines for generating conversation synopses.

Impact

We expect SENSEI to advance the state-of-the-art in conversation understanding towards the next-generation of analytics technology. SENSEI's is committed to develop methodologies for professional conversation data analysts and create innovative analytics services from large scale



data streams. Given the diverse target user groups, SENSEI will impact diverse industry sectors, such as contact centres, news and social media.

Project Information

Making Sense of Human-Human Conversation Data – SENSEI FP7-ICT-610916

Duration	1 November 2013 – 30 October 2016
Budget	3,500,000 EUR (EC Contribution 2,650,000 EUR)
Contact	University of Trento Giuseppe Riccardi Department of Information Engineering and Computer Science +390461283175 riccardi@disi.unitn.it
Partners	University of Trento**,** Italy Université d'Aix Marseille, France University of Sheffield, United Kingdom University of Essex, United Kingdom Teleperformance, Italy Websays, Spain
Website	http://www.sensei-conversation.eu/