

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.

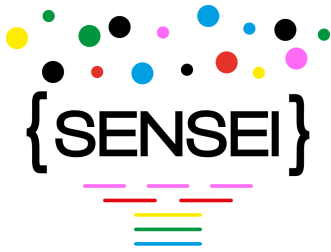
Target User Groups

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.



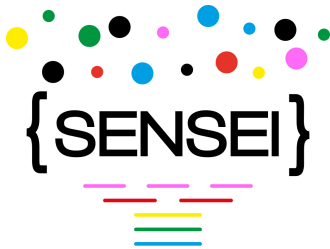
Work performed and the main results

The main three results in the second period of the project are the following:

First Release of Summarization Prototype. One of the most important achievements in Period 2 has been the development of the first version of the summarization prototype. The development of the summarization prototype is based on the first release and deployment of the technology pipeline starting at Work package 3 (semantic and para-semantic), Work package 4 (discourse) and ending in Work package 5 (summarization algorithms). The summarization prototype has been released for both spoken and social media conversations. In the speech use case, we have the prototype released for Italian and French. In the social media use case we have summarization prototype versions for English, French and Italian. In this period the prototype framework has been developed in such way that system developed can assist an end-user in understanding the conversation by using various NLP and data mining algorithm categories. The algorithms that have been implemented can be categorized into: a) Sentiment Analysis Algorithms, b) Summarization Algorithms (abstractive and extractive summaries) c) Conversation Filtering Algorithms d) Topic Detection Algorithms, Clustering Algorithms e) Argument Structure Algorithms.

First Extrinsic Evaluation of the Conversation Summaries. The second most important achievement of P2 has been the user-evaluation of some of the prototype versions. The evaluation experiments have been carried out following the principle of engaging users in realistic task settings. The goal has been to gather feedback on the present state of development of the prototype and providing insights for improving it. For the speech use cases the users were professionals of a call centre company, for the social media use case they were graduates with experience of using online news and reader comments. The prototype versions we have been evaluating include different types of summaries, including short summaries of call centre calls (synopses), summaries of reader comments, (including text summaries, linked to clusters of comments and graphical summaries, where clusters of comments are represented in a pie-chart), and filled questionnaires used to summarise some aspects of call centre agents' communication behaviour when they interact with their customers. In general, the results of the evaluation have highlighted that the evaluation protocol and the tasks are realistic, potentially accepted by the users, and feasible on a larger scale.

Sharing and Dissemination of SENSEI Results. In Period 2 there have been quite a few scientific and business initiatives to share and disseminate SENSEI's vision and preliminary results and resources. In particular we have co-organized and run two shared tasks at the SIGDIAL-Multiling 2015 multilingual summarization evaluation campaign: the Online Forum Summarization (OnForumS) task and the Call-Centre Conversation Summarization (CCCS) task. The SIGDIAL event was an opportunity to meet with the summarization community which gave good feedback and willingness to continue organizing summarization evaluation campaigns, in particular reiterate both of the SENSEI tasks, and develop the community at the European level for instance with a COST. For the SENSEI project, the organization of the shared task has been an opportunity to reach out to the community and create a well-defined evaluation benchmark on tasks that matter for the project, and for the community it has been and opening towards new research directions for which the SENSEI project lowered the barrier of entry.



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	3,560,044 EUR (EC Contribution 2,650,000 EUR)
Contact	University of Trento Giuseppe Riccardi Department of Information Engineering and Computer Science +390461282087 giuseppe.riccardi@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/