

## New research platform provides support to journalists

RESEARCH NEWS

07 | 2015 || Topic 4

An important summit meeting about the debt crisis in Greece has just come to an end, and any minute now the participants will emerge to face the press. Has Athens presented a list of reform proposals, and if so what are its central points? Have there been any statements by leading representatives of the German government and business community? How are the financial markets responding? And what's the mood among the general populace? It isn't easy to keep tabs on a situation like this and not miss out on any information. Every reporter worth their salt will take the investigation seriously, consulting endless different sources – from press conferences, news tickers and live video streams to commentaries posted on the social networks and the latest tweets. This mass of information is so huge that it is barely manageable without recourse to smart data analysis technologies. While solutions do exist in the form of analysis tools that search through news databases on the basis of semantic matching algorithms – and undoubtedly simplify the journalist's work – these are nothing more than island solutions according to Dr. Daniel Stein of the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS. He explains, "Because they are tailored to specific formats, they provide only a partial view of the complete news stream."

### BMBF-funded project "News-Stream 3.0"

As part of the "News-Stream 3.0" project funded by the German Federal Ministry of Education and Research (BMBF), the Fraunhofer Institute in St. Augustin has joined forces with Neofonie GmbH, Deutsche Welle and the German Press Agency dpa to develop a novel analysis tool for journalists. As Dr. Stein explains, "Our aim is to create a research platform capable of processing heterogeneous, multimedia data streams rapidly and efficiently, performing a semantic analysis, pooling the results, and presenting a condensed overview in a uniform format." The new analysis process can scan information from different sources, including text messages, videos and audio files, in a matter of seconds. "It is an amalgam of existing approaches and product-ready solutions, blended together to create new tools," says Stein. With the help of the new technology, journalists and editors will soon be able to gather together the information they need to report on a specific subject and document the results of their research in what are known as "news cases". The content is continuously updated to ensure that users always stay abreast of the latest information.

One of the most important areas of expertise that the Fraunhofer scientists have brought to the project is that of audio analysis. The audio mining technology developed by Fraunhofer IAIS enables the spoken word in audio and video recordings to be analyzed in the same way as the content of text files. Their solution also takes voice recognition a step further by automatically identifying the person speaking. This technology is based on the i-vector paradigm that assigns an acoustic signature to each sound bite. The storage capacity of each i-vector is such that they are capable of analyzing even

large, heterogeneous, high-bitrate-streaming audio files in a minimum of time. Among other things, the speaker recognition function implemented in the News Stream platform is designed to ensure that editors don't miss a single statement by important personalities. For example, whenever the German chancellor appears before the camera to speak about an issue of topical interest, a reference to this broadcast is shown on the platform user's screen. This audio-based search function allows spoken-word content in the news streams to be linked with the name of the person speaking and filtered on this basis.

Deutsche Welle and dpa employees have recently started testing an early prototype in their daily work. "This user feedback is a crucial element of the project because it enables us to make further improvements to the platform while it is still at the development stage and thus ensure it meets the requirements of future users," says Stein.



Simultaneous speaker recognition and tweet analysis (example taken from the 2015 UK general election). (© Fraunhofer IAIS) | Picture in color and printing quality: [www.fraunhofer.de/press](http://www.fraunhofer.de/press)