

## Charles Huot & Sonia Collada



**Bio. Charles Huot**, Expert System, worked ten years at IBM, before founding in 2000 the company Temis which publishes semantic enrichment software. About fifteen years later, he created People in the sun, a consulting firm in the field of big data and artificial intelligence. Representative of the college of Small and Micro Enterprises, president of the Cluster's "knowledge" thematic commission since 2011, is president of the board of directors of Cap Digital.

His areas of expertise are Big Data, data & text mining, Artificial Intelligence, linguistic processing and data analysis.



**Bio.** After studying at the Carlos III University of Madrid in Telecommunications Engineering, Systems and Networks, **Sonia Collada** worked for 4 years as a solutions consultant at Dedalus. She has now been working for 7 years at Expert System, initially as a solutions consultant and later as a project manager.

Her areas of expertise are Artificial Intelligence, Machine Learning, Big Data and Machine Learning.

# Detecting fake news in social media content

Charles Huot, Sonia Collada

Expert System

**Abstract.** SocialTruth is an European project designed and developed by an international consortium of 11 partners, whose slogan is “Embedding veracity for social media and web”. Creating a system that ranks news and certifies their reliability, targeted to professionals and ordinary users of social media and the web.

Expert System has conceived a system providing textual and semantic analysis; more concretely a meta-verification system on story classification and ranking. The global aim of the approach we have worked on is, based on a golden corpus, comparing an untargeted document to our base of “qualified documents”. The hypothesis applied lies in the fact that “true news” have a pattern that our system will highlight, being able to go from a large group of articles to automatically select a relevant group to compare to a possible “fake news”.

The system defined goes through the following milestones: The first milestone is the categorization, having the untargeted document; we analyze it using the categorization process and reduce the list to compare it to. The second milestone corresponds to narrowing the search by using the clustering, which means applying an unsupervised categorization among a whole corpus. The third milestone is the similarity analysis: how similar this document is it to other equivalent documents? The fourth milestone corresponds to sentiment analysis, whether there is a strong sentiment present or not in the text and the writing style. Finally, we end up by extracting the writeprint of the document, language level, tone and type of vocabulary.

The outcome of our work is the conception of a semantic analyzer that aims at providing information that will be the input of the expert meta-verification system. The expert meta-verification system will combine the verification results from the content verification services created from social, semantic and multimedia content, in order to compute a meta-score that accurately depicts the credibility of the digital content under consideration.

**Keywords.** Social media and social networks, Content verification, Semantic analysis, Machine learning, Expert system, Fake news.

