

4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Contextual Exploration (EC3): A strategy for the detection, extraction and visualization of target data

Christophe Jouis

LIP6 (CNRS -UPMC Sorbonne Université), France

EC3 is intended to extract relevant information from large heterogeneous and multilingual text data, in particular in WEB 3.0. The project is based on an original method: contextual exploration. EC3 does not need syntactic analysis, statistical analysis or a "general" ontology. EC3 uses only small ontologies called "linguistic ontologies" that express the linguistic knowledge of a user who must concentrate on the relevant information from one point of view. This is why EC3 works very quickly on large corpus, whose components can be both whole books as well as "short texts": SMS to books. At the output, EC3 offers a visual representation of information using an original approach: the "Memory Islands". EC3 is implanted in the ACASA / LIP6 team. EC3 is tested on the large digitized corpus provided by the Labex OBVIL «Observatoire de la Vie Littéraire», in partnership with the Bibliothèque Nationale de France (<http://obvil.paris-sorbonne.fr/>). OBVIL intends to develop all the resources offered by digitization and computer applications to examine French literature from the sixteenth to the twentieth century, English and American literature, Italian literature, Spanish literature, in its most traditional formats and media Or the most innovative.

Biography

Christophe Jouis is assistant professor at the University Paris Sorbonne Nouvelle, France. He received a Ph.D. in Applied Mathematics at the "Ecole des Hautes Etudes en Sciences Sociales" (EHESS); and CAMS ("Centre d'Analyse et de Mathématiques Sociales"), OPTION: Science, Logic, Linguistics. From 2000 to 2004 he was associate professor in the Department of Computer Science at the University of Quebec at Trois-Rivieres (Canada), under the direction of Professor Ismail Biskri. In 2005, he joined the LIP6 ("Laboratoire d'Informatique de Paris 6), affiliated with the University Pierre et Marie Curie (UMPC) and the CNRS (France). Within the LIP6, he is currently a member of the research team ACASA ("Cognitive Agents and Automated Symbolic Learning"), led by Professor Jean-Gabriel Ganascia. His research interests are in natural language processing (NLP), cognitive sciences, ontology, typicality, data mining and information retrieval.

christophe.jouis@lip6.fr

Notes: