



**HAL**  
open science

## Exploration de corpus scientifiques et techniques.

Fidelia Ibekwe-Sanjuan

► **To cite this version:**

Fidelia Ibekwe-Sanjuan. Exploration de corpus scientifiques et techniques.. domain\_other. Université Jean Moulin - Lyon III, 2010. tel-00635294

**HAL Id: tel-00635294**

**<https://theses.hal.science/tel-00635294>**

Submitted on 26 Oct 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Jean Moulin, Lyon 3

Mémoire en vue de l'Habilitation à Diriger des Recherches  
Présenté et soutenu publiquement par :

**Fidelia Ibekwe-SanJuan**

Le 30 juin 2010

**Exploration de corpus scientifiques et techniques**

**Devant le jury composé de :**

**Bruno Bachimont**, Habilité à Diriger des Recherches, Université Technologique de Compiègne (Rapporteur)

**Anne Condamines**, Directrice de Recherches CNRS, Université Toulouse Le Mirail

**Viviane Couzinet**, Professeur à l'Université Paul Sabatier, Toulouse (Rapporteur)

**Geneviève Lallich-Boidin**, Professeur, Université Claude Bernard, Lyon

**Ahmed Silem**, Professeur, Université Jean Moulin, Lyon.

**Khaldoun Zreik**, Professeur, Université de Paris 8 (Rapporteur)

## Résumé.

Ce mémoire d'HDR met en perspective les recherches que nous avons conduites en Sciences de l'information et de la communication (SIC) au cours des douze dernières années. Partie d'une problématique pratique d'accès à l'information spécialisée dans le cadre de la veille scientifique et technologique, nous avons été amenée peu à peu à interroger les fondements mêmes de la discipline, ses concepts fondamentaux et les paradigmes qui la traversent. Les quatre premiers chapitres reviennent sur le chemin parcouru par les SIC depuis leurs origines (1972) jusqu'à présent.

Les cinq chapitres suivants mettent en perspective les axes de recherche que nous avons développés dont l'exploration de corpus scientifiques et techniques forme le noyau. Cet axe de travail a conduit à la mise au point d'un dispositif technique d'analyse des publications permettant de générer une cartographie thématique des principales tendances observées dans le corpus.

Un deuxième axe, conçu dans le prolongement du premier, porte sur l'analyse discursive des textes scientifiques pour la caractérisation des marques d'énonciation des auteurs. Cet axe de recherche trouve ses racines dans la théorie des facettes développée dans le milieu documentaire par Ranganathan en 1933.

Le troisième axe porte sur une analyse du paradigme orienté-système qui domine actuellement les recherches en recherche d'information. Nous revenons sur les modalités d'évaluation des systèmes de recherche d'information et sur leur réelle utilité pour de *vrais* usagers.

Un quatrième axe portant sur les usages des TIC (Technologies de l'Information et de la Communication) étudie les ressorts du concept de « bibliothèques participatives » (*participatory libraries*) outre-Atlantique. Les thèmes abordés dans les chapitres précédents permettent de dégager les perspectives d'avenir.

**Mots-clés :** épistémologie de l'information-communication, théories de l'information, exploration de corpus, cartographie thématique, analyse terminologique, analyse discursive, recherche d'information, annotation de corpus, bibliothèques participatives, usage des TIC, web social.

## Abstract.

This *mémoire* retraces our research trajectory since our doctoral dissertation in the late nineties to the present day. Originally of a more pragmatic strain, our research gradually led us to investigate the theoretical foundations of Information & Communication Sciences (ICS). The first four chapters recall attempts to define the object of the discipline, to understand the nature of its fundamental concepts – information and communication, and the different epistemological tendencies and paradigms that drive research in the discipline.

The five following chapters analyze our own research poles of which the exploration of scientific corpora forms the center. This led us to design a technical tool for mapping research trends from corpora. A second pole on discourse analysis analyzes scientific texts structure using rhetorical markers and lexico-syntactic cues in order to characterize sentences according to author's intent. Next, the validity of the system-oriented paradigm currently dominating research in information retrieval comes under scrutiny. We show that there is room for more user-oriented approaches and a symbolic representation of queries and documents. Turning to more user-oriented studies, we analyze the impact of technological evolution, in particular web 2.0 on information professionals (librarians) with regard to the emergence of participatory librarianship in North America. Finally, perspectives for future research are outlined.

**Keywords:** epistemology of information – communication, information theory, corpus exploration, scientific and technical information, topic mapping, terminology analysis, discourse analysis, information retrieval, corpus annotation, participatory libraries, social web.