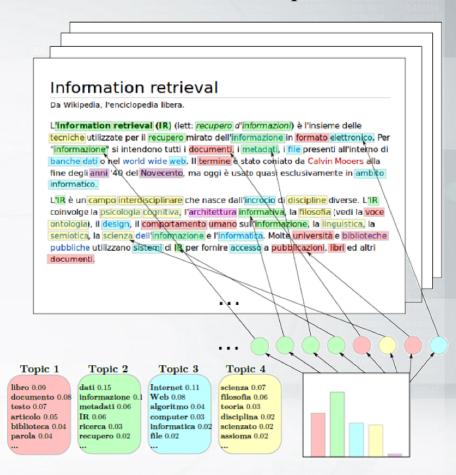
## Multilingual topic models

## English corpus

## Information retrieval From Wikipedia, the free encyclopedia This article is about information retrieval in general. For the fictional government department, see Brazil (film). Information retrieval (IR) is the area of study concerned with searching for documents, for information within documents, and for metadata about documents, as well as that of searching structured storage, relational databases, and the World Wide Web. There is overlap in the usage of the terms data retrieval, document etrieval, information retrieval, and text retrieval, but each also has its own body of literature, theory, praxis, and technologies. IR is interdisciplinary, based on omputer science, mathematics, library science, information science, information architecture, cognitive psychology, linguistics, statistics and law Automated information retrieval systems are used to reduce what has been called "information overload". Many universities and public libraries use IR systems to provide access to books, journals and other documents. Web search engines are the most visible IR applications. Topic 1 Topic 2 Topic 3 Topic 4 document 0.12 data 0.17 computer 0.13 science 0.08 text 0.09 information 0.1 Internet 0.09 theory 0.06 book 0.08 search 0.08 Web 0.07 theorem 0.04 library 0.08 retrieval 0.06 informatics 0.05 math 0.02 word 0.07 metadata 0.04 system 0.04 statistics 0.02 term 0.06 IR 0.03 application 0.02 praxis 0.01

## Italian corpus



I. Vulic, W. De Smet, J. Tang, M.-F. Moens. Probabilistic topic modeling in multilingual settings: a short overview of its methodology with applications, NIPS-2012.