

Content Based Searching for INIS

V. Jain^{1,2}, R. K. Jain²

¹ICLE's Motilal Jhunjhunwala College of Science, Arts and Commerce, Navi Mumbai, India

²La Vastu labh Pvt. Ltd., India

²Bhabha Atomic Research Centre (BARC), Mumbai, India

Corresponding Author: V. Jain, vidhijain75@gmail.com

Whatever a user wants is available on the internet, but to retrieve the information efficiently, a multilingual and most-relevant document search engine is a must. Most current search engines are word based or pattern based. They do not consider the meaning of the query posed to them; purely based on the keywords of the query; no support of multilingual query and and dismissal of nonrelevant results. Current information-retrieval techniques either rely on an encoding process, using a certain perspective or classification scheme, to describe a given item, or perform a full-text analysis, searching for user-specified words. Neither case guarantees content matching because an encoded description might reflect only part of the content and the mere occurrence of a word does not necessarily reflect the document's content. For general documents, there doesn't yet seem to be a much better option than lazy full-text analysis, by manually going through those endless results pages. In contrast to this, new search engine should extract the meaning of the query and then perform the search based on this extracted meaning. New search engine should also employ Interlingua based machine translation technology to present information in the language of choice of the user.