

Abstract

Introduction: The final goal for all Information retrieval systems is to retrieve the most relevant documents in the minimum time and in information retrieval system relevance usually determined according to similarity of words. Farsi language like other Languages has many problems in information retrieval. The differences in spelling and semantic problems are some problems searching in through searching tools. In this study, we investigated 15 Farsi Language morphological variants that make problem in retrieving Farsi documents. Research community is consisting of Four database: magiran, iranmedex, markaz-e-mantaghei shiraz, medlib and two search engines: Parsik and Parsijoo.

Method: This study is a comparative survey. We selected morphological problems from the literature (15 problems) and for every morphological problem we selected one keyword with two variants. Farsi medical thesaurus. Then, we searched these words in selected search tools and recorded Time and frequency in tables. The result were analysed by Chi-square test, Cramer coefficient and Phi coefficient using spss (v. 16)

Results: IranMedex retrieved more documents than other databases, totally and parsijoo found more results than parsik. When searching the databases, the longest search period belonged to Markaz-e- mantaghe -ei-Shiraz and shortest time dedicated to Medlib. For search engines, searching parsik took shortest time

Conclusion: Although all search tools retrieved the selected form of the keywords (based on farsi medical thesaurus) but also not-selected keywords have been restored. all the Farsi Language morphological problems, just “tashdid” was solved for search engines in retrieving Farsi medical documents but other Farsi language morphological problems were not solved for search Tools that searching different keywords in search Tools make different results in retrieving documents.

Keywords: morphology, Farsi language, information retrieval, Farsi search engines, Farsi web, medical documents, medical information, Farsi medical thesaurus, information searching, medical key words.