

Word sense disambiguation for English Quranic IR system

Sabrina Tiun^a, Hamed Zakr^a, Masnizah Mohd^a, Norazlinda Zainal Abidin^b, Ahmad Irfan Ikmal Hisham^b

^a Faculty of Technology and Information Science, Universiti Kebangsaan Malaysia, Bangi, Malaysia

^b Centre for Modern Language and Human Sciences, Universiti Malaysia Pahang, Kuantan, Malaysia

ABSTRACT

Quranic text Information Retrieval (IR) is quite demanding yet very trivial due to that user will not always use the exact keywords to retrieve the relevant Quranic text (verse). Many have tried to overcome this problem by expanding or reformulating the query entered by users using semantic approaches with resources such as ontologies and thesauri. Word Sense Disambiguation (WSD) has been less interest to the IR research community due to the insignificant or very little significant impact on the IR performance. Recently, researchers pay interest on applying WSD to the IR problem due to the intuition that deep semantic analysis on the query process will give good impact on the IR performance. However, we have not seen any articles mentioning the use of WSD for Quranic IR, which we are assuming less or none research on WSD for Quranic IR have been carried out. Thus, this motivates us to explore WSD impact on Quranic IR performance. This paper will describe our on-going project on building an English Quranic WSD at the early stage, which is still at the proposal stage, where we layout what could be the best approach, resources and disambiguation algorithm for Quranic WSD for IR.

KEYWORDS

Unsupervised word sense disambiguation; Quranic translated text; Quranic IR

ACKNOWLEDGEMENTS

This research project is funded by Malaysian Government under research grant ERGS/I/2013/ICT07/UKM/03/1.