

M. Biltawi, S. Tedmori and A. Awajan, 2021, "Arabic Question Answering Systems: Gap Analysis," in IEEE Access, vol. 9, pp. 63876-63904, 2021, doi: 10.1109/ACCESS.2021.3074950. (ISI, Scopus).

Abstract

Question-answering (QA) systems aim to provide answers for given questions. The answers can be extracted or generated from either unstructured or structured text. Therefore, QA is considered an important field that can be used to evaluate machine text understanding. Arabic is a challenging language for many reasons; although it is spoken by more than 330 million native speakers, research on this language is limited. A few QA systems created for Arabic text are available. They were created to experiment on small datasets, some of which are unavailable. The research on QA systems can be expanded into different components of QA systems, such as question analysis, information retrieval, and answer extraction. The objective of this research is to analyze the QA systems created for Arabic text by reviewing, categorizing, and analyzing the gaps by providing advice to those who would like to work in this field. Six benchmark datasets are available for testing and evaluating Arabic QA systems, and 26 selected Arabic QA systems are analyzed and discussed in this research.