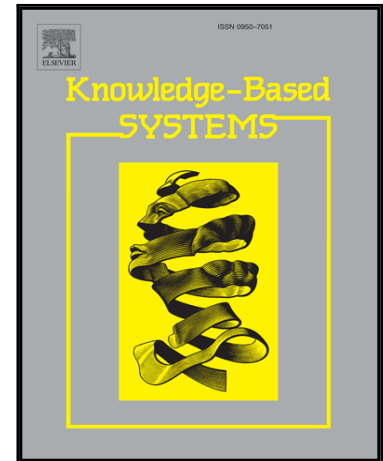


# Accepted Manuscript

A Study about the Future Evaluation of Question-Answering Systems

Alvaro Rodrigo, Anselmo Peñas

PII: S0950-7051(17)30416-1  
DOI: [10.1016/j.knosys.2017.09.015](https://doi.org/10.1016/j.knosys.2017.09.015)  
Reference: KNOSYS 4036



To appear in: *Knowledge-Based Systems*

Received date: 22 December 2016  
Revised date: 29 July 2017  
Accepted date: 8 September 2017

Please cite this article as: Alvaro Rodrigo, Anselmo Peñas, A Study about the Future Evaluation of Question-Answering Systems, *Knowledge-Based Systems* (2017), doi: [10.1016/j.knosys.2017.09.015](https://doi.org/10.1016/j.knosys.2017.09.015)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## A Study about the Future Evaluation of Question-Answering Systems

Alvaro Rodrigo\*, Anselmo Peñas  
{alvarory,anselmo}@lsi.uned.es  
NLP & IR Group at UNED  
Juan del Rosal 16, Madrid, Spain  
Tel.: +34 91 398 96 93  
Fax: +34 91 398 93 82  
\*(Correspondence Author)

---

### Abstract

Evaluation campaigns of Question Answering (QA) systems have contributed to the development of such technologies. These campaigns have promoted some changes oriented to overcome results. However, at this period we see how systems have reached an upper bound, as well as systems are still far way from answering complex questions. In this paper, we overview the main QA evaluations over free text, paying special attention to the changes encouraged at such campaigns. We observe that systems still return a high proportion of incorrect answers and that the changes are almost not included in traditional approaches. Moreover, we analyze QA collections in order to obtain better insights about the main challenges for current QA systems. We detect that QA systems find very difficult to deal with different rewordings in questions and documents, as well as to infer information that is not explicitly mentioned in texts. Based on those observations, we recommend a set of directions for future evaluations, suggesting the application of textual inference and knowledge bases as a way for improving results.

*Keywords:* Question Answering, Evaluation Campaigns, Validation, Textual Inference

---

Download English Version:

<https://daneshyari.com/en/article/6862133>

Download Persian Version:

<https://daneshyari.com/article/6862133>

[Daneshyari.com](https://daneshyari.com)