



Upgrade

Open in app



Published in Lingvo Masino · Follow



Ranjan Satapathy · Follow

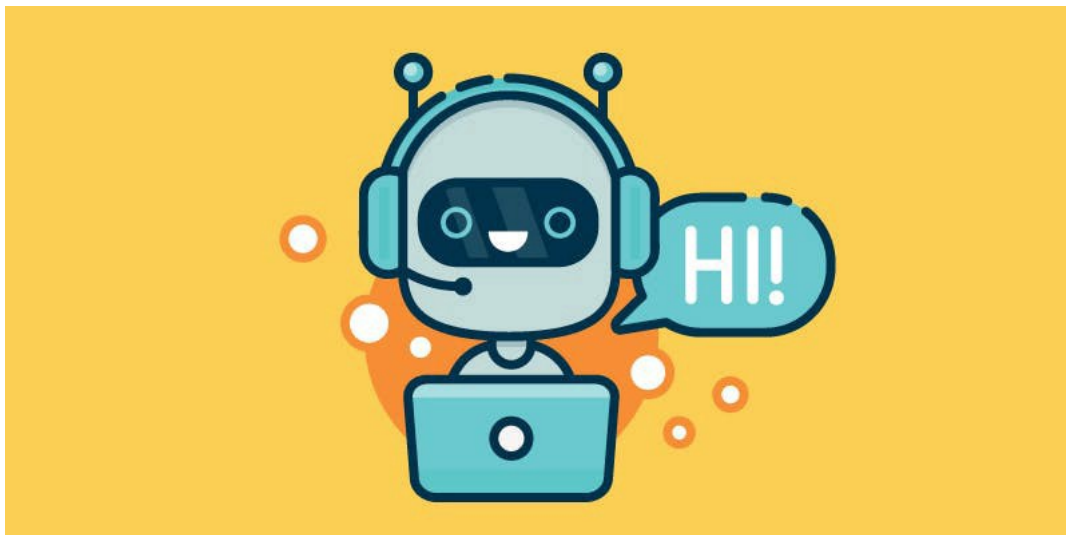
Aug 11, 2018 · 5 min read · Listen



# Question Answering in Natural Language Processing [Part-I]

## Introduction

Question Answering is a computer science discipline within the fields of information retrieval and natural language processing, which focuses on building systems that automatically answer questions posed by humans in a natural language. A computer understanding of natural language consists of the capability of a program system to translate sentences into an internal representation so that this system generates valid answers to questions asked by an user [1]. Valid answers mean answers relevant to the questions posed by the user. As the internal representation of natural language, sentences must adequately map semantics of this statement, the most natural approach is in the simulation of facts contained in the sentences using a description of real objects as well as actions and events connected with these objects. To form an answer it is necessary, in the first place, to execute the syntax and semantic analysis of a question. This article covers the introduction to Question Answering, types and challenges posed by the systems in real world.

[Source](#)

## Open Datasets available for Question Answering

1. **Stanford Question Answering Dataset (SQuAD)** [2] is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage, or the question might be unanswerable. There is an awesome article on this [here](#).
2. **WikiQA** dataset [3], is a publicly available set of question and answer pairs, collected and annotated for research on open-domain question answering. It is constructed using a more natural process and is more than an order of magnitude larger than the previous dataset. In addition, the WikiQA dataset also includes questions for which there are no correct sentences, enabling researchers to work on answer triggering, a critical component in any QA system.
3. The **TREC-QA** dataset contains questions and answer patterns, as well as a pool of documents returned by participating teams.
4. **NewsQA** dataset [4] is to help the research community build algorithms that are capable of answering questions requiring human-level comprehension and reasoning skills. Leveraging CNN articles from the [DeepMind Q&A Dataset](#), authors prepared a crowd-sourced machine reading comprehension dataset of 120K Q&A pairs.

## Types of Question Answering

There are three major modern paradigms of question answering:





Upgrade

Open in app

The answer type specifies the kind of entity the answer consists of (person, location, time, etc.). The query specifies the keywords that should be used for the IR system to use in searching for documents.

b) **Knowledge-based question answering** is the idea of answering a natural language question by mapping it to a query over a structured database. The logical form of the question is thus either in the form of a query or can easily be converted into one. The database can be a full relational database, or simpler structured databases like sets of RDF triples. Systems for mapping from a text string to any logical form are called semantic parsers. Semantic parsers for question answering usually map either to some version of predicate calculus or a query language like SQL or SPARQL.

c) **Using multiple information sources: IBM's Watson [5,6]** system from IBM that won the Jeopardy! challenge in 2011 is an example of a system that relies on a wide variety of resources to answer questions. The first stage is question processing. The DeepQA system runs parsing, named entity tagging, and relation extraction on the question. Then, like the text-based systems, the DeepQA system extracts the focus, the answer type (also called the lexical answer type or LAT), and performs question classification and question sectioning. Next DeepQA extracts the question focus. Finally the question is classified by type as definition question, multiple-choice, puzzle or fill-in-the-blank. Next is the candidate answer generation stage according to the question type, where the processed question is combined with external documents and other knowledge sources to suggest many candidate answers. These candidate answers can either be extracted from text documents or from structured knowledge bases. Then it is passed through the candidate answer scoring stage, which uses many sources of evidence to score the candidates. One of the most important is the lexical answer type. In the final answer merging and scoring step, it first merges the candidate answers that are equivalent. The merging and ranking is actually run iteratively; first the candidates are ranked by the classifier, giving a rough first value for each candidate answer, then that value is used to decide which of the variants of a name to select as the merged answer, then the merged answers are re-ranked.

## Challenges in Question Answering

The main challenges [7] posed by a Question Answering System are described below:

1. **Lexical Gap** : In a natural language, the same meaning can be expressed in different ways. Because a question can usually only be answered if every referred concept is identified, bridging this gap significantly increases the proportion of questions that can be answered by a system.
2. **Ambiguity** : It is the phenomenon of the same phrase having different meanings; this can be structural and syntactic (like “flying planes”) or lexical and semantic (like “bank”). The same string accidentally refers to different concepts (as in money bank vs. river bank) and polysemy, where the same string refers to different but related concepts (as in bank as a company vs. bank as a building).
3. **Multilingualism** : Knowledge on the Web is expressed in various languages. While RDF resources can be described in multiple languages at once using language tags, there is not a single language that is always used in Web documents. Additionally, users have different native languages. A QA system is expected to recognize a language and get the results on the go!

## References

[1] <https://arxiv.org/pdf/1111.4343.pdf>

[2] <https://rajpurkar.github.io/SQuAD-explorer/>

[3] Yang, Y., Yih, W.T. and Meek, C., 2015. Wikiqa: A challenge dataset for open-domain question answering. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 2013–2018).

[4] Trischler, A., Wang, T., Yuan, X., Harris, J., Sordoni, A., Bachman, P. and Suleman, K., 2016. Newsqa: A machine comprehension dataset. *arXiv preprint arXiv:1611.09830*.

[5] Kalyanpur, A., Patwardhan, S., Boguraev, B.K., Lally, A. and Chu-Carroll, J., 2012. Fact-based question decomposition in DeepQA. *IBM Journal of Research and Development*, 56(3.4), pp.13–1.

[6] <https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=6177717>



