

Building Watson: An Overview of the DeepQA Project

David Ferrucci

IBM T. J. Watson Research Center

Eric Brown

IBM T. J. Watson Research Center

Jennifer Chu-Carroll

IBM T. J. Watson Research Center

James Fan

IBM T. J. Watson Research Center

David Gondek

IBM T. J. Watson Research Center

Aditya A. Kalyanpur

IBM T. J. Watson Research Center

Adam Lally

IBM T. J. Watson Research Center

J. William Murdock

IBM T. J. Watson Research Center

Eric Nyberg

Carnegie Mellon University

John Prager

IBM T. J. Watson Research Center

Nico Schlaefer

Carnegie Mellon University

Chris Welty

IBM T. J. Watson Research Center

DOI: <https://doi.org/10.1609/aimag.v31i3.2303>

Abstract

IBM Research undertook a challenge to build a computer system that could compete at the human champion level in real time on the American TV Quiz show, Jeopardy! The extent of the challenge includes fielding a real-time automatic contestant on the show, not merely a laboratory exercise. The Jeopardy! Challenge helped us address requirements that led to the design of the DeepQA architecture and the implementation of Watson. After 3 years of intense research and development by a core team of about 20 researchers, Watson is performing at human expert-levels in terms of precision, confidence and speed at the Jeopardy! Quiz show. Our results strongly suggest that DeepQA is an effective and extensible architecture that may be used as a foundation for combining, deploying, evaluating and advancing a wide range of algorithmic techniques to rapidly advance the field of QA.

Author Biographies

Eric Nyberg, Carnegie Mellon University

Professor

Language Technologies Institute, School of Computer Science



Published

2010-07-28

Issue

[Vol 31 No 3: Fall 2010](#)

Section

Articles

Authors who publish with this journal agree to the following terms:

1. Author(s) agree to transfer their copyrights in their article/paper to the Association for the Advancement of Artificial Intelligence (AAAI), in order to deal with future requests for reprints, translations, anthologies, reproductions, excerpts, and other publications. This grant will include, without limitation, the entire copyright in the article/paper in all countries of the world, including all renewals, extensions, and reversions thereof, whether such rights current exist or hereafter come into effect, and also the exclusive right to create electronic versions of the article/paper, to the extent that such right is not subsumed under copyright.
2. The author(s) warrants that they are the sole author and owner of the copyright in the above article/paper, except for those portions shown to be in quotations; that the article/paper is original throughout; and that the undersigned right to make the grants set forth above is complete and unencumbered.
3. The author(s) agree that if anyone brings any claim or action alleging facts that, if true, constitute a breach of any of the foregoing warranties, the author(s) will hold harmless and indemnify AAAI, their grantees, their licensees, and their distributors against any liability, whether under judgment, decree, or compromise, and any legal fees and expenses arising out of that claim or actions, and the undersigned will cooperate fully in any defense AAAI may make to such claim or action. Moreover, the undersigned agrees to cooperate in any claim or other action seeking to protect or enforce any right the undersigned has granted to AAAI in the article/paper. If any such claim or action fails because of facts that constitute a breach of any of the foregoing warranties, the undersigned agrees to reimburse whomever brings such claim or action for expenses and attorneys' fees incurred therein.
4. Author(s) retain all proprietary rights other than copyright (such as patent rights).
5. Author(s) may make personal reuse of all or portions of the above article/paper in other works of their own authorship.
6. Author(s) may reproduce, or have reproduced, their article/paper for the author's personal use, or for company use provided that AAAI copyright and the source are indicated, and that the copies are not used in a way that implies AAAI endorsement of a product or service of an employer, and that the copies per se are not offered for sale. The foregoing right shall not permit the posting of the article/paper in electronic or digital form on any computer network, except by the author or the author's employer, and then only on the author's or the employer's own web page or ftp site. Such web page or ftp site, in addition to the aforementioned requirements of this Paragraph, must provide an electronic reference or link back to the AAAI electronic server, and shall not post other AAAI copyrighted materials not of the author's or the employer's creation (including tables of contents with links to other papers) without AAAI's written permission.
7. Author(s) may make limited distribution of all or portions of their article/paper prior to publication.
8. In the case of work performed under U.S. Government contract, AAAI grants the U.S. Government royalty-free permission to reproduce all or portions of the above article/paper, and to authorize others to do so, for U.S. Government purposes.
9. In the event the above article/paper is not accepted and published by AAAI, or is withdrawn by the author(s) before acceptance by AAAI, this agreement becomes null and void.

[Open Journal Systems](#)

Information

[For Readers](#)

[For Authors](#)

The DeepQA Approach. Early on in the project, attempts to adapt PIQUANT (Chu-Carroll et al. 2003) failed to produce promising results. We devoted many months of effort to encoding algorithms from the literature. Figure 6 illustrates the DeepQA architecture at a very high level. The remaining parts of this section provide a bit more detail about the various architectural roles. Figure 6. DeepQA High-Level Architecture. Content Acquisition. The first step in any application of DeepQA to solve a QA problem is content acquisition, or identifying and gathering the content to use for the answer and evidence sources shown in figure 6. Content acquisition is a combination of manual and automatic steps. The DeepQA project is aimed at illustrating how the advancement and integration of Natural Language Processing (NLP), Information Retrieval (IR), Machine Learning (ML), massively parallel computation and Knowledge Representation and Reasoning (KR&R) can greatly advance open-domain automatic Question Answering. The goal of Watson is to achieve human-level question answering performance. Eric is involved in the architecture and scale out of DeepQA, as well as special question processing for Watson. Keynote Address II on Tuesday, September 14, 8:30 am. Towards a Science of Parallel Programming.