

TECHNOLOGY

AI in the ER: What Is Watson and What Does It Mean for Medical Librarianship?

Jason Bengtson

KEYWORDS: Watson, IBM, Artificial Intelligence, AI, Medical Librarianship, Semantic, Natural Language

Jason Bengtson is the Emerging Technologies/R&D Librarian for the University of New Mexico's Health Sciences Library and Informatics Center in Albuquerque, NM.(E-mail: jbengtson@salud.unm.edu).

Comments and suggestions should be sent to the Column Editors: Hope Leman (hleman@samhealth.org), Nichole Mitchell (nmitchell79@gmail.com) and Jason Bengtson (jbengtson@salud.unm.edu).

INTRODUCTION

It runs on ten server racks, has natural language analysis capability and recently won a Jeopardy! Tournament where it was pitted against two champions. If you answered, "What is Watson", you would be correct, but you probably didn't answer in time to beat the new champ. Watson, the latest incarnation of IBM's Deep Blue research project, recently astonished the Jeopardy! audience with its performance. This was a watershed moment in artificial intelligence

research because it required not only analytical reasoning skills (of the same sort that enabled IBM's Deep Blue to become a chess champion), but also natural language processing skills. Simply gathering data and organizing it is nothing new for a computer system, but being able to apply a solid level of semantic reasoning to complex natural language questions is a much more challenging goal. By coupling sophisticated natural language recognition with the brute power of modern digital computing, Watson was able to handily defeat human opponents.

Watson itself runs off of 90 IBM Power 750 servers on 10 racks (1). This constitutes a significant amount of digital real estate, yet Watson still can't replicate all of the natural language processing capabilities of one human brain (2). What Watson can do is process what information it has using its own proprietary algorithms faster than its human competition. For purposes of the Jeopardy! bout Watson didn't use the information resources of the web (which probably would have been too slow, anyway), relying instead on 500 Gigabytes worth of natural language documents stored on its local disks (1).

Watson works well when analyzing certain types of question, especially those that can be narrowed down by applying rules of English grammar (limiting by synonyms that rhyme for example), or those where Watson can isolate a central theme and ignore extraneous elements (2). What Watson cannot do well is hold a real dialog with a human, as physician have to do with their patients (2).

Given the promise shown by this advance, IBM has already publicly speculated about future directions for Watson, including a medical version of the application. Two hospitals have already signed on as partners with IBM for this development effort (2). There has even been some speculation on the web that Watson could make human physicians obsolete.

The idea that Watson, as impressive as it is, sounds the death toll for the medical profession seems rather absurd, but it could potentially be a powerful tool for making recommendations for diagnosis and treatment. Less a search engine than a synthesis engine, Watson could do something similar to Cochrane Systematic Reviews, albeit less skillfully (2). Essentially, Watson could offer a trade-off: services like Cochrane give higher quality, human generated systematic-reviews for a limited range of topics, whereas Watson would provide much more rapid, lower quality systematic reviews over almost any topic imaginable.

This leaves our profession with a number of important questions. What does Watson mean for medical librarians? What do we need to see from Watson as a tool before it is suitable for use in a medical environment? And, assuming a medical version of Watson lives up to the enthusiasm of its creators, how much will it cost?

WHAT FORM WILL MEDICAL WATSON TAKE?

Given Watson's hardware requirements, it seems like an obvious conclusion that Watson will function as a cloud application, in the same way that the databases most librarians rely on now are remotely hosted. This format would, of course, leave the system vulnerable to power and network outages, as well as server failures at IBM. It would also potentially make Watson vulnerable to outages affecting the databases and other sources the application uses to function (assuming databases form the back-end of a future medical Watson). Of course, even a locally hosted system can be affected by these issues, but cloud systems are, by their nature, highly dependent on environmental conditions over multiple nodes and networks. IBM would be wise

to use server arrangements that minimize any downtime for an application such as Watson that might need to be used clinically, without warning and at a moment's notice.

The interface itself is a question mark at this point. The version of Watson that appeared on Jeopardy! relied on input in the form of digitally transmitted ASCII files, but IBM could incorporate any number of input options within the final product (2). Will it simply have one large basic search box where natural language queries can be typed, will it have a more sophisticated interface, or will both approaches be incorporated? These are questions that IBM software engineers will have to give serious consideration to before any practical version of Watson can become a reality.

WILL WATSON MAKE MEDICAL LIBRARIANS OBSOLETE?

No more than databases or systematic review services have. Watson's ability to synthesize knowledge is potentially very impressive, but it still has severe limitations. The instincts and serendipity employed by professional researchers would be lost on Watson. Where its source materials end, Watson's queries will end, leaving alternative databases, uncataloged print materials, direct inquiries or any other tools for follow-up completely unexplored. Most importantly, however, Watson does not have human-level problem analysis capabilities. Watson looks for objective facts, but cannot make a judgment call; a fact that means it won't be replacing physicians anytime soon, either (3). Watson can synthesize facts from multiple sources into one solution, but to do so it relies on clues that can be easily misinterpreted. At this time it also has no ability to take into account the indirect verbal or visual cues that can add depth and disambiguation to information from a patient or patron.

At the crux of the matter remains the fact that reference tools are only ever as good as the questions we ask of them. Professional researchers understand the implications of this simple, but inescapable statement. The best literature searches are still conducted by humans, with the use of tools like databases, or, perhaps Watson, and that is not likely to change.

WHERE WILL THE DATA COME FROM?

This question goes to the heart of our profession. What sources will be used by a medical version of Watson? Will it have access to all of the major medical databases? How many journal subscriptions will Watson have available to it? Will Watson search through full text articles, or merely search through secondary, synthesized sources, like Up To Date, that may be of limited overall value? If Watson is really to function as an aid in matters of diagnosis and treatment, its base of information needs to be as broad as possible, and it should be firmly grounded in high authority, primary source material.

Currently, Watson is set up to use unstructured documents rather than the databases that house so much medical knowledge at present (2). John Kelly, the head of IBM's research lab, has expressed a vision of hospitals entering all of the new medical papers into Watson as they come out. . . a notion that is as optimistic as it is ignorant of the financial dimension of academic intellectual property (3). It remains to be seen whether or not IBM will change that vision or if their engineers intend to use an information bank comprised of raw papers, studies and other materials in a similar fashion to the 500 GB of natural language documents Watson used to win at Jeopardy! (1) Either way, access to that much full text will be expensive, and Watson's source material will need to be updated regularly as new studies are conducted and new breakthroughs

made. Even when divided among many customers, the cumulative cost of the various useful biomedical materials will be staggering. If Watson is given full access to what it needs to be searching, Watson customers can expect a hefty price tag.

HOW WILL THE RESULTS BE DELIVERED?

Providing answers in a trivia contest is a very different thing from providing advice to medical professionals. In a trivia contest, there is only one right answer, and how it was obtained is unimportant. In the medical world there is rarely one “right” answer, and where the information came from is at least as important as the information itself. Was it obtained from case studies, or double blind trials? Was it reported in a reputable medical journal, or found orphaned on the web? A medical version of Watson that simply spits out a best answer of “stomach cancer” would be of little real value. Medical professionals will need evidence from reliable sources, presented in a clear manner that ranks levels of evidence appropriately.

This being the case, what can the medical community expect from Watson? If Watson generates suggestions ranked by levels of confidence it will not be doing enough. Suggestions will also need to be ranked according to the level of evidence associated with the sources that support those suggestions. There will need to be a way for medical professionals to see the citations, and, in a best case scenario, the full text of the references that Watson drew on when generating its recommendations. Once again, the specter of price appears. As all librarians know, full text access does not come cheaply.

It is also worth considering the danger of overwhelming medical practitioners with useless data (4). Doctors and other health professionals often work within severe time

constraints, making it unrealistic for them to sift through large amounts of information on the fly. Here the synthesis possibilities of a tool like Watson could potentially make all the difference, if they are applied judiciously. The supporting information to a recommendation must be readily accessible, but it shouldn't clutter up the initial display of suggestions from Watson. Instead, that interface, while clearly displaying some indication of levels of evidence, should be relatively spartan, only supplying an option to call up supporting references as needed.

The logic used by Watson might also be relevant to medical professionals using the tool. One display option could be a logic tree showing Watson's interpretation of the query and the semantic steps it followed when gathering data and deciding on recommendations. This may lead medical practitioners to new and useful lines of reasoning in a difficult case, making Watson's approach to a question as least as valuable as its answer.

CONCLUSION

The allure of Watson as a tool is undeniable. As information accumulates in our society, the tools to manage that information become all the more indispensable just as our role as information professionals becomes more central to the sciences. Watson represents a fundamental shift away from traditional search engine solutions, offering a potential means to digitally process natural language, gather information, and synthesize a response in a meaningful way. Nonetheless, it remains to be seen if Watson will live up to its promise. If IBM is serious about turning Watson into a practical medical tool it must address issues of source authority, comprehensiveness, timeliness of data, mode of delivery, cost, methodology, and availability of

references. If they do not, they risk any future version of Watson becoming little more than an expensive curiosity or a better spoken variation on Wikipedia.

REFERENCES

1. IBM - what powers Watson? IBM. <http://www-03.ibm.com/innovation/us/watson/watson-for-a-smarter-planet/watson-schematic.html> (12 Mar. 2011).
2. Angelica A. D. (31 Jan. 2011) How watson works: a conversation with Eric Brown, IBM Research Manager. *KurzweilAI*. <http://www.kurzweilai.net/how-watson-works-a-conversation-with-eric-brown-ibm-research-manager> (10 Mar. 2011)
3. Thompson C. (16 Jun. 2010) What is I.B.M.'s Watson?. *New York Times Magazine*. <http://www.nytimes.com/2010/06/20/magazine/20Computer-t.html> (10 Mar. 2011)
4. Husain, I. (21 Feb. 2011). Why IBM's artificial intelligence "Watson" could not replace a physician *iMedicalApps*. <http://www.imedicalapps.com/2011/02/ibm-watson-replace-physician-artificial-intelligence/> (10 Mar. 2011).