Transforming the Way Organizations Think with Cognitive Systems

An IBM® Redbooks®
Point-of-View publication by the
IBM Academy of Technology

By Rob High, IBM Fellow, and **Bill Rapp**,
IBM Distinguished Engineer

Highlights

Cognitive systems can transform how organizations think, act, and operate:

- They can combine natural language processing, hypothesis generation and evaluation, and dynamic learning for a powerful, fast, and accurate solution.
- They can help to understand the complexities of unstructured data.
- They can apply advanced analytics to weigh and evaluate responses.
- They can help to improve learning based on outcomes to get smarter with each iteration and interaction.
- They can use deep natural language processing (NLP) to assess and evaluate language over unlimited topics and to then make informed judgments.

Gaining insight from big data

Our data volume is exploding. But not just the volume of data, the type of data is taking on increasing forms, and the demand for speed at which we need to understand this data is accelerating. We need computers that can assess this flood of information so that we can mine the most value from our data. This data is *big data*.

Ninety percent of all the data in the world was produced in the last two years. Data volume is expected to grow as we interconnect and instrument more of our world. And 80 percent of all information in the world is unstructured information, which includes text such as literature, reports, articles, research papers, emails, blogs, tweets, forums, chats, and messages.

With all this data, we need better ways of understanding it to find knowledge that can be applied to solving our most pressing and important issues. We need to be able to make sense of what is being said and use that information to answer questions, gain insight, and drive better, more informed decisions. However, reading this information and understanding it with the same accuracy that people can is hard for traditional computing systems. We need a new class of computing that is capable of understanding the subtleties, idiosyncrasies, idioms, and nuance of human language. This computing is *cognitive computing*.

Cognitive systems mimic how humans reason and process information. Unlike traditional computers that are programmed to calculate rapidly and perform deterministic tasks, cognitive systems analyze information and then draw insights from that analysis by using probabilistic analytics. They learn from their interactions with data, in effect continuously reprogramming themselves.

IBM is leading the Cognitive Systems Era. It is transforming the way we use computers and technology to process information and natural language, creating the potential to change the way businesses use information and make decisions.

For example, healthcare organizations can use cognitive systems, such as IBM Watson™, to analyze all available data, especially textual information, to improve patient outcomes while making processes more efficient. Financial companies can use this technology to analyze vast amounts of unstructured data to improve credit decisions, investment analysis and risk management. Cognitive systems can transform how organizations think, act, and operate in the future.

Language processing complexity

Language is the expression of ideas and the medium by which we communicate and understand, think, speculate, and imagine. It is the way we understand the world around us and how we convey and manipulate that understanding.

The depth, complexity, and vastness of human language makes it difficult for computers to understand. We are able to see through inconsistencies, contradictions, irregularities, gaps, and lack of clarity, and we can still understand each other with a great deal of accuracy. Computers cannot easily "see" as we do and have problems when analyzing enormous amounts of data by using human reasoning processes, mainly because human language is incredibly complex.

This difference between precision and accuracy is important. *Precision* is the mechanical or scientific exactness in a passage of text. *Accuracy* is the degree to which a passage infers that another passage is considered to be true. To accurately answer a question, you have to consider its available context. Without enough evidential information, it is difficult to accurately respond to a question, even though you can precisely answer parts of it. This precision versus accuracy dichotomy poses another layer of difficulty for computers that are working to understand human language.

Humans can easily process language, even if they do get confused at times, because they can resolve the meaning of information better than computers can. This ability is due to our innate quality to provide a context for the information that we want to capture. This concept has been a key goal of the artificial intelligence community for decades. Though we have increased the precision of language processing, it is only with Watson that we have reached the level of accuracy necessary for computers to function well in the real world of broad natural language.

Effectively navigating through the vastness of unstructured information requires a new era of computing called cognitive systems.

Meet Watson

IBM Watson exemplifies a technological break through by IBM Research in its 100-year history of scientific discovery and continuous effort to create systems to make our world better and to live smarter.

Watson represents a major leap in the capacity of technology systems to identify patterns, to gain critical insight, and to enhance decision making despite daunting complexity. Specifically, Watson is a significant advance in Deep Question-Answering (DeepQA) technology. Deep QA drives the ability to understand natural human speech about a limitless range of topics and to then make informed judgments about those topics. The science behind Watson can elevate human and computer communication to new levels. It can also expand the power of advanced analytics to understand vast amounts of structured and unstructured data. For example, Watson can use DeepQA to provide critical and timely information to help medical staff diagnose and treat patients.

Watson and language

Many natural language systems emphasize precision within the confines of specific well-formed rules. They look for sets of specific words and synonyms without assessing their context. They then tally the number of times those words were found. This concept is generally referred to as *shallow natural language processing*, or *shallow NLP*, which can be precise within narrow confines, but it is often not accurate because it does not consider context. It can also become less precise as the context changes.

When accuracy is needed over precision, we use *deep natural language processing*, or *deep NLP*, that analyzes context in evaluating a question. Watson is a deep NLP system that assesses as much context as possible that it derives from immediate information, from more broadly available information, from the knowledge base (called a *corpus*), and from source databases.

As a cognitive system, Watson can apply human-like characteristics to conveying and manipulating ideas. It can break down language to identify inferences between text passages while formulating the correct answer to a question. When combined with traditional computing, cognitive systems can solve problems with more speed, accuracy, and resilience, and on a massive scale over broad bodies of information.

But even Watson does not necessarily understand words in the language, rather it understands language features that are used by humans. From those features, it can determine whether one text passage (question) infers another text passage (answer) with great accuracy under changing circumstances.

The concepts of context, temporal, and spatial constraints all matter and are critical in enabling a cognitive system that can behave with human-like characteristics.

Context and reasoning are critical in building a basis for processing language.

How Watson responds to questions

Watson builds on previous generations of computing with robust capabilities, none of which are unique but which, in combination, deliver a powerful solution:

- *Natural language processing* helps to understand the complexities of unstructured data.
- *Hypothesis generation and evaluation* applies advanced analytics to weigh and evaluate a panel of responses based on only relevant evidence.
- *Dynamic learning* improves learning based on outcomes to get smarter with each iteration and interaction.

A critical element of the way that Watson functions is the corpus, which consists of unstructured knowledge, such as text books, guidelines, how-to manuals, FAQs, benefit plans, and news. Watson ingests the entire corpus to modify and curate content into a workable form. It focuses on whether the corpus content is appropriate, and sifts out information that is out of date, irrelevant, or derived from potentially unreliable sources.

Watson breaks down the question, ferrets out potential responses in the corpus, and then examines the response and the statement context in hundreds of ways. It uses the results to gain a degree of confidence in its interpretation of the question and potential answers. Figure 1 outlines the process that Watson uses to respond to a question.

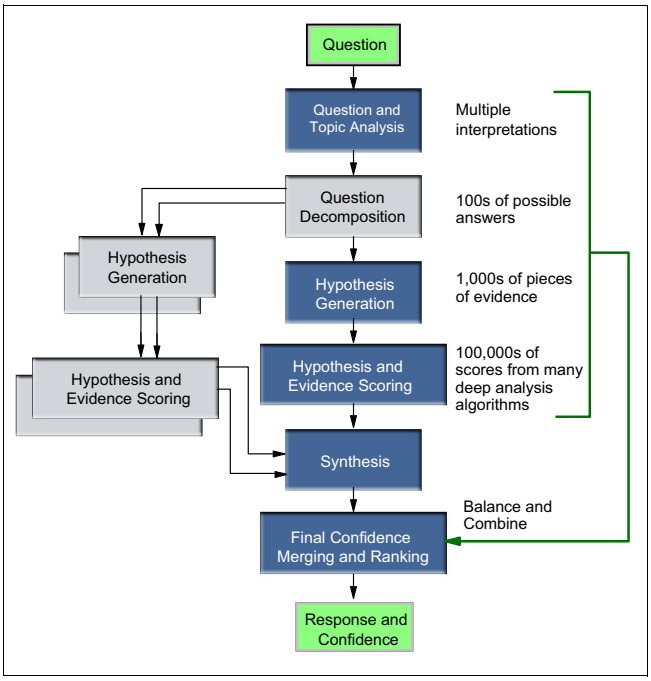

Figure 1 How Watson responds to a question

The combination of the capabilities that Watson has creates a unique solution that achieves the following goals:

- Move beyond the constraints of programmatic computing.
- Move from reliance on structured, local data to unlock the world of global, unstructured data.
- Move from decision tree-driven, deterministic applications to probabilistic systems that evolve with users.
- Move from a keyword-based search that provides a list of locations where an answer might, or might not, be located, to an intuitive, conversational set of confidence ranked responses.

Putting Watson to work

Many organizations are leveraging Watson with traditional computing to solve business problems by using statistical analytics, rules and business processing, collaboration, and reporting. For example, you can combine statistical analysis with answers from Watson about potential events to signal investment risk.

Healthcare

We all want our doctor to know the right, best answer to our health problems. But, with the volume of biomedical data doubling every seven years, how can any medical professionals stay on top of everything they need to know to improve our health?

Watson can help the healthcare industry to process, store, retrieve, and analyze the immense field of biomedical information to improve patient health.

Watson will never replace a nuanced medical diagnosis. But, it can analyze a patient's complex conditions and, with incredible speed and accuracy, present the medical professional with treatment possibilities that are unique to that patient to narrow the choices to a correct medical decision. For example, in a 45-year-old patient with chest pain, Watson can analyze and combine all the patient's conditions and history to help create treatments that are specific to this patient, resulting in better and faster care.

Business and finance

Cognitive systems will radically change how businesses interact with both their customers and employees. Watson can help businesses improve risk and valuation processes. If you had a highly reliable, incredibly fast system to ask the right question and select the right answer with confidence, how would you change your business? If you were able to know your customers' buying patterns and habits, how would you change your product line and retail process?

The volume of data in the financial industry is growing 70 percent yearly. With such depth of data, how do you extract real knowledge and key insight quickly and reliably? By simultaneously seeing all available financial and business news, economic trends and market information, you can make valuable logical connections between disparate pieces of information. For example, you can analyze such information to seek companies that might soon become acquisition targets.

Customer service

The first level of applying Watson to the marketplace is in the area of call centers, where its capabilities are expected to change the call center business model.

Call centers are often a large cost for industries. Many call centers receive millions of calls that create long, frustrating wait times for callers. The goal of Watson is to increase the quality of service and to provide correct answers quickly and more accurately. Watson is the next evolutionary step in cognitive systems that will move the call center model to far greater efficiency and enable users to get data on their terms.

What's next: How IBM can help

IBM is investing in substantial improvements to Watson that will lead to further breakthroughs in healthcare, finance, call centers, government, and energy industries that will help build a more intelligent planet.

IBM will continue to develop Watson to be a smarter, leaner and faster expert advisor to human experts in many fields. To learn more about IBM Watson, see:

http://www.ibm.com/innovation/us/watson

Resources for more information

Use the following resources for more information about the concepts highlighted in this paper:

- *The Era of Cognitive Systems: An Inside Look at IBM Watson and How it Works*, REDP-4955

  http://www.redbooks.ibm.com/abstracts/redp4955.html?Open
- IBM Watson: Ushering in a new era of computing

  http://www.ibm.com/innovation/us/watson
- IBM Watson: Smarter answers for a smarter planet

  http://www-03.ibm.com/innovation/us/watson/watson-for-a-smarter-planet/smarter-answers-for-a-smarter-planet.html

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to: *IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document, REDP-4961-00, was created or updated on December 13, 2012.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (or), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at `http://www.ibm.com/legal/copytrade.shtml`

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

IBM®
IBM Watson™
Redbooks®
Redbooks (logo) 

The following terms are trademarks of other companies:

Other company, product, or service names may be trademarks or service marks of others.