2014 年 8 月 29 日

●商用化に歩むIBM「ワトソン」

【Wall Street Journal, 2014/08/28】
　IBM の質問応答システム「ワトソン」は３年前、クイズ番組「Jeopardy!」のチャンピオンと対戦して勝利。
　IBM は以後もその商用化を進めており、28 日にはワトソンをクラウド・サービスとして提供する「Watson Discovery Advisor」の企業、研究機関への提供を開始した。
　ワトソンの強みは、質問の文脈を理解した上で膨大な量のテキストから必要な情報を取捨選択し、推断を下すことにあり、その特長は、今週発表された IBM とベイラー医科大学、MD アンダーソンがんセンターとの共同研究報告でも示されている。
　この研究プロジェクトでは、がんの発生原因の１つと見られるタンパク質「p53」の変異を引き起こす別のタンパク質の選別をワトソンに行わせた。
　p53 の変異は、がん研究ではポピュラーなテーマで、これまで７万件以上の論文が発表されている。
　ワトソンは、これらの論文を全て自動的に読み込み、p53 の変異を引き起こすタンパク質を予測。その予測結果は実験で追試され、少なくともその一部は予備実験段階では正しいことが確認されたという。
　７万件の論文を人間が呼んだ場合、１日５件読んだとしても全て読み終わるには 38 年近くかかることから、ワトソンのテキストマイニング能力の有用性が分かる。
　しかし、この研究では、ワトソンに医学用語を学習させ、コンセプトや文脈を判別できるよう訓練する必要もあり、ワトソンを様々な分野で活用する上でのハードルも示している。
　ワトソンの実用化がどの段階まで来ており、IBM がそれからどれだけの利益を得るかはまだ不明な点も多い。
　IBM は、１月にワトソンを事業部門化。マンハッタン、イーストビレッジにオフィスを新設するとともに、今後 10 億ドルをワトソンに投入する方針も発表した。
　この中には、ワトソンを活用する新興企業を支援する１億ドルのベンチャー資金も含まれており、IBM は、ワトソンを一種の人口知能 OS にしたいとの考えを見せている。

（参考）本件報道記事
Questions for IBM's Watson

Steve Lohr
By STEVE LOHR

Sometimes, figuring out the right question is harder than finding the answer. Just ask Watson.

Watson's claim to fame rests on beating human champions in the question-and-answer game "Jeopardy!" In the three years since, IBM has been working to move Watson into the marketplace, step by step. The next step came on Thursday, when the company made a Watson technology, Discovery Advisor, available for companies and research organizations to use as a cloud service.

In fact, IBM announced Watson Discovery Advisor back in January. But now, John Gordon, vice president for strategy and product commercialization for Watson, said, "We're ready to open this up."

The new service builds on Watson's turbocharged text-mining and identification technology, which was so impressively on display in its "Jeopardy!" triumph. In its current version, Discovery Advisor is tuned for science, specifically the life sciences and medicine. Beyond mining text, the discovery tool not only finds connections among words but also links related concepts together to generate hypotheses. What might be the right place to look? What path of scientific inquiry is most likely to yield new knowledge?

  "Before, the answer was there, and the challenge for Watson was really just to find it," said W. Scott Spangler, a data scientist at IBM's Almaden Research Center in San Jose, Calif. "But this is about what's the right question for the scientist to ask."

A strong case for the power of the Watson technology was made in a research paper published this week and presented at the Association for Computing Machinery's annual conference that focuses on knowledge discovery and data-mining ? what we now call data science. Mr. Spangler is one of several researchers from IBM, the Baylor College of Medicine and the MD Anderson

Cancer Center in Houston who are co-authors on the paper, "Automated Hypothesis Generation Based on Mining Scientific Literature."

In the research project, biologists and data scientists used Watson to identify proteins that modify p53, a crucial protein that is sometimes called "the guardian of the genome." When p53 is mutated, it can set the stage for tumor growth of many kinds of cancer.

It is a most popular subject of research. More than 70,000 papers have been published on p53. Watson read them all in an automated effort to predict proteins that turn p53's activity on or off. Using Watson's analysis, the cancer researchers identified six potential proteins to target for new research. Watson went beyond digging for a known fact; it found previously unrecognized connections.

The predictions were then tested in biology laboratory experiments. "Some of the things that were predicted turned out to be true, at least in preliminary experiments," said Dr. Olivier Lichtarge, director of Baylor's Center of Computational and Integrative Biomedical Research. "We've shown that this technology can mine scientific literature and reason about it in molecular biology."

Dr. Lichtarge pointed to the efficiency of the automated system, and its potential to accelerate scientific discovery, by observing that a scientist might read five research papers a day at most. Even at that pace, he noted, it would take a human scientist nearly 38 years to read the more than 70,000 papers available today on p53. Hastening the pace of discovery should open the door to more effective drugs and other therapies.

IBM is by no means the only technology company applying the tools of artificial intelligence to try to create the equivalent of smart digital assistants. In the consumer market, there is Apple's Siri, Google's Now and Microsoft's Cortana. In the corporate market, there is less emphasis on having the software talk to users, but the underlying principle is similar, said David Schubmehl, an analyst at IDC. He singled out Palantir Technologies,Digital Reasoning and Saffron Technology among those pursuing the same market opportunity as IBM's Watson.

"The technology in this field is rapidly improving, but we're in the early days," Mr. Schubmehl said.

How early it is for Watson and how large the payoff might be for the company is uncertain. In January, IBM placed Watson in its own business unit, with new offices in the East Village in Manhattan and a pledge to spend $1 billion on Watson. The financial commitment includes $100 million for a venture fund to support start-ups and entrepreneurs making applications that run on the Watson artificial intelligence software. The goal is that Watson will be not just a product but a so-called platform on which other technology and businesses are built. Operating systems like Microsoft's Windows, Apple's iOS and Google's Android are classic technology platforms, and IBM aspires to make Watson an artificial intelligence operating system.

"The platform strategy for Watson is very smart," said Tom Austin, an analyst for Gartner. "But I'm waiting for evidence that there is action on the entrepreneurial side ? lots of businesses and developers making Watson-based products and trying to make money off it."

An IBM spokeswoman said the company had "hundreds of clients and partners with active projects" using the Watson technology.

The IBM research project with the Baylor College of Medicine shed some light on both the hurdles and the opportunity for Watson. Its artificial intelligence software may blitz through thousands of scientific papers in minutes, but only after it was trained to identify medical terms, discern concepts and make connections. Mr. Spangler said he spent two years working with the Baylor researchers. The software that brought success was an application based on Watson called the Baylor Knowledge Integration Toolkit or KnIT.

Proof-of-concept projects like the one with Baylor are always time-consuming. The goal, though, is that what is learned on early projects can be translated to code that is used again and again. Mr. Gordon, the IBM strategist, is confident that Watson can scale up in "co-creation projects with clients that can transform an industry."

Watson's quiz-show triumph gave many people the impression that the technology was a general-purpose artificial intelligence, which could be applied to any field. IBM's early marketing communication did nothing to dampen that enthusiasm. But while Watson 1.0 had general properties, it was designed to win at "Jeopardy!"

Could some version of the Watson technology ? now a cloud service instead of a hulking computer ? be adapted to technically daunting tasks with the potential for a significant payoff in insights and dollars?

  "That question pinpoints our key challenge," Mr. Spangler observed. And the progress shown in the protein-identification project, he added, suggests that the answer may be yes. "This is what makes me most excited," Mr. Spangler said.

Source :

以　上