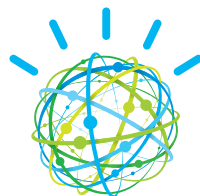


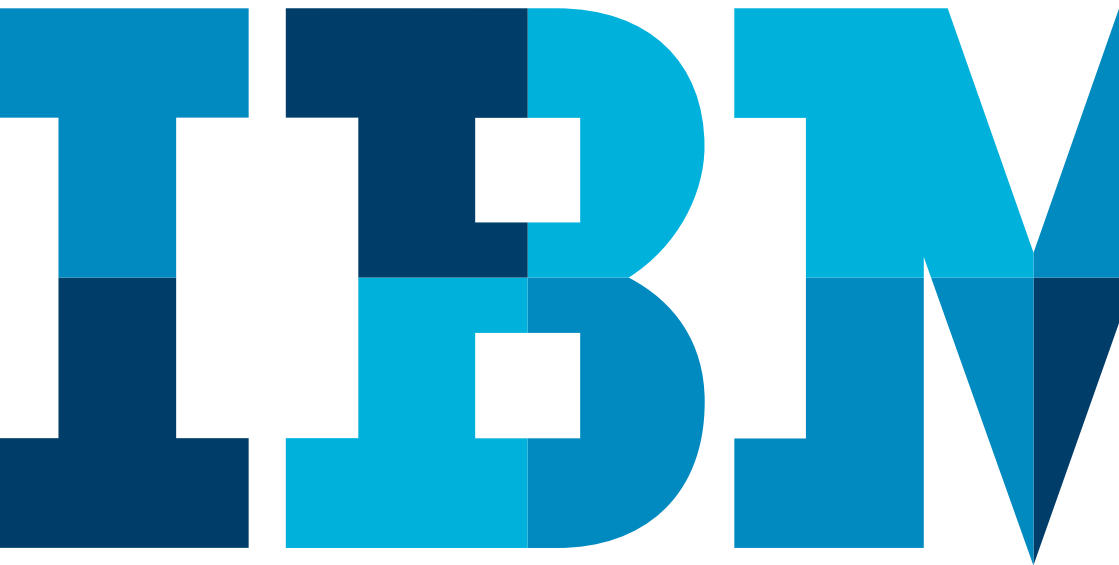
IBM Software

エンタープライズ・コンテンツ管理



# 「Jeopardy!」でのWatsonの 勝利の原動力となった IBM自然言語分析ソフトウェア

IBM Content Analytics (ICA) ソフトウェアをビジネスで活用する



---

「Watsonを支えるテクノロジーは、データ管理とデータ分析の未来の姿を示します。実社会では、このテクノロジーにより、物流から医療まで、幅広い分野に関する洞察を明らかにすることができるようになるでしょう」

— John Cohn, IBM フェロー、IBM Systems and Technology Group

## Watsonについて

IBMのWatsonは世界最先端の質問応答システムです。自然言語で書かれた質問について、その文脈に隠された意味を理解し、解答を導き出します。高度な自然言語処理(NLP)とDeepQAという自動質問応答テクノロジーを併せ持つWatsonは、分析、システム設計、コンテンツ管理、さらにはデータ管理の未来の姿を表しています。Watsonはコアとなる自然言語分析とその他多数の先進テクノロジーを活用して、非常に短い時間で1つの正しい答えを導き出します。このテクノロジーは、医療やカスタマー・サービスへの適用から官公庁の情報収集までなど、幅広い応用が期待されています。Watsonの詳細については、[ibmwatson.com](http://ibmwatson.com)をご覧ください。



## Watsonの仕組み

Watsonは、「Jeopardy!」の質問内容を分析します。まず、何を聞かれているのかを判断するために質問を理解しようとします。その際、Watsonは自然言語を分析しなければなりません。次に、多種多様なコンテンツ(大半が自然言語文書形式)を分析して、根拠を持った解答を複数見つけようとします。最後に、見つけ出した解答が正しいかどうかを示す相対的な確からしさを、点数付けの手法に基づいて判断します。

Stephen Baker氏はその著書「*Final Jeopardy: Man vs. Machine and the Quest to Know Everything*」の中で、この挑戦について以下のようなわかりやすい例を挙げて説明しています。「When 60 Minutes premiered, this man was U.S. President(米国のドキュメンタリー番組『60 Minutes』の初回放送時、この人物は米大統領でした)―従来、『premiere(初めて放映する)』がどういう意味を持ち、そしてそれが特定の日付に関連しているということを理解するのは、コンピューターには困難なことでした。コンピューターにとって、『premiere』は『premier(首相)』を意味する可能性もあるからです。この質問は人の肩書きに関するものなのか、それとも番組開始に関するものなのか。その答えが分かっても今度は、『60 Minutes』が最初に放映された日を見つけ出し、続いてその当時の米大統領が誰だったのかを見つけ出さなければなりません」。つまり、この質問に答えるには文脈(コンテキスト)に関する高い理解力が求められるということです。

## ビジネスに与える影響

Watsonのクイズ番組への挑戦は、今日のビジネスの世界で見られるようなコンテキストの情報に基づく難題を提起しています。

- 保存されている情報の80%以上は構造化されていない(テキストベースの)情報です。
- この80%以上を理解するのは容易なことではありません。Jeopardy!のクイズ形式では、問題に含まれる微妙な意味、風刺、謎掛けなどの複雑な要素の分析を、人間よりもそうした処理が苦手なコンピューターが行わなくてはならず、コンピューター・システムに究極の課題を突きつけます。ここで、登場するのが自然言語処理技術です。

ワトソンは、質問応答技術の更なる向上を目的に、自然言語処理技術をさらに進化させることを目的に設計されました。自然言語処理技術は現在、ビジネス価値の実現に活用することができます。Watsonで実証された技術の開発実績を持つIBMは、お客様の課題解決に役立つさまざまな製品を提供しています。

IBM Content Analyticsソフトウェアは、非構造化情報の中に埋もれている貴重な洞察を引き出し、さまざまな分野において、専門家や企業が迅速に重要な意思決定を行う支援をします。

- エンタープライズ・コンテンツ管理 (ECM) リポジトリ、構造化データ、ソーシャル・メディア、コールセンター・ログ、研究レポート、トランスクリプト、電子メール、安全レポート、契約といった社内外の複数のソースや異なるタイプの情報からコンテンツを集約および抽出する
- 意味を理解する、また傾向、パターン、相関関係、偏差、ビジネス・コンテキストを識別するために、NLPやその他の分析手法を使用してエンタープライズ・コンテンツ(およびデータ)を整理、分析、視覚化する

---

「Hertzは毎日、お客様に関する大量の洞察を収集しています。

その中にはWeb調査、電子メール、テキスト・メッセージから届く数千ものコメントも含まれます。私たちは、この洞察を戦略レベルとローカル・レベルの両方で活用し、業務の向上につなげたいと考えました」

— Joe Eckroth, CIO, The Hertz Corporation

- モデルの構築や複雑なシステムの導入を行うことなく、疑問点を確認したり、新たな事柄を明らかにしたりするなどして迅速に洞察を引き出すために、対話式の検索および調査を行う

## ケース・スタディー

以下に、自社コンテンツから貴重な洞察を得ているお客様をご紹介します。

**医療研究** — ほとんどの医療機関と同様、BJC Healthcareでは、病気の進行、治療の有効性、および長期転帰の研究に不可欠な情報などの貴重な履歴情報を、構造化されていない臨床記録や診断レポートに埋もれさせていました。既存の生物医学情報リソースはまとまりがなく、相互運用がなされておらず、ごく一部の研究者しか利用できない状態で、冗長な情報が多く、また構造化されていない臨床記録や診断レポートなどに埋もれている大量の研究情報をうまく活用する機能もありませんでした。

ところが、IBM自然言語分析ソフトウェアにより、BJCと大学の研究者は現在、非構造化情報を分析して、これまで不可能であった重要な質問に答えることができました。質問とは、例えば、「患者はたばこを吸うか」「吸う場合、1日に何本吸い、吸い始めてどのくらいたつか」「吸わない場合、禁煙期間はどのくらいか」「患者は自宅でどのような薬を服用しているか」「患者の診断結果はどのようなものか」「これまでどのような治療を受けてきたか」などです。BJCは現在、医療情報に関してより深い洞察を入手し、コンテンツ内の傾向とパターンを発見できるようになった結果、より良い医療を患者に提供できるようになりました。

---

「このソリューションを強力なものにしているのは、従来のオンライン検索による方法を超越る、コンテキストを分解して結果に組み込む機能です」

— Billy Houghteling、エグゼクティブ・ディレクター、技術移転室、North Carolina State University ([ibm.com/software/success/cssdb.nsf/CS/SSAO-8DFLBX?OpenDocument&Site=software&cty=en\\_us](http://ibm.com/software/success/cssdb.nsf/CS/SSAO-8DFLBX?OpenDocument&Site=software&cty=en_us))

**顧客満足** — 多くのビジネスにとって、製品、サービス、担当者に関する顧客満足の傾向を明らかにすることは非常に大切です。Hertz CorporationとMindshare Technologies(エンタープライズ・フィードバック・ソリューションの主要プロバイダー)は、IBM自然言語分析ソフトウェアを使用してテキスト・メッセージを含む顧客調査データを検証し、自動車や機器のレンタル状況をより詳しく把握して、顧客満足度を正確に特定し、その向上に必要な調整を行っています。IBM自然言語分析ソフトウェアを使うことにより、Hertzのような企業は、新たなマーケティング・キャンペーンを展開したり、製品やサービスを変更したりすることで、顧客のニーズに応えています。

**研究分析** — North Carolina State University(ノースカロライナ州立大学)にとって、大学の本質は教育に留まらず、あらゆる形態で知識を発展させ、広めていくことにあります。ノースカロライナ州立大学が直面している主要な問題の1つは、大学で利用可能な膨大な数のデータソースの取り扱い方でした。そこで同大学は、大量のデータを効率よくマイニング/分析して、大学の研究内容の商品化を可能にする企業を特定するためのソリューションを探すことにしました。目標としたソリューションは、数千もの非構造化情報源のコンテンツを解析し、データとテキストの分析を実行し、有益な結果をまとめて提示できるものでした。

IBM自然言語分析ソフトウェアを導入することにより、同大学ではターゲット企業を見つけるための期間を数カ月から数日に短縮できました。その結果、新たな商用化の機会が見つかり、候補企業数が約300%増えました。また、この大学の広範な情報の本質を知ることにより、同大学の技術移転機関では大学の研究を通して作り出されたテクノロジーに関するライセンスをこれまでよりも効果的に供与できるようになりました。

## IBM エンタープライズ・コンテンツ管理について

IBMエンタープライズ・コンテンツ管理ソフトウェアの導入により、世界のトップ企業では、これまで以上に迅速かつ的確に意思決定をしています。非構造化情報を駆使することにより、新しい方法でビジネスの意思決定に臨み、新しい方法で連携して意思決定に取り組み、さらに新しい方法で意思決定に影響を与えることが可能になります。この結果、コンテンツは最高の洞察源となります。業種別のIBM ECMソリューションは、ライフサイクル全体を通してコンテンツを取り込み、管理および共有して、コンプライアンスの順守、コストの削減、生産性の最大化を達成するのを支援します。IBM ECMポートフォリオには、既存のシステムと統合して情報の価値を最大限に高める、文書の取り込みとイメージ処理、ソーシャル・コンテンツ管理、高度なケース管理、情報ライフサイクル・ガバナンス、自然言語分析といった多種多様な機能が含まれています。世界の13,000を超える企業および官公庁のお客様が既にIBM ECMソリューションを活用して、イノベーションによって業績を改善し、競争優位を維持しています。

Watsonの原動力となったIBM自然言語の分析機能を導入すると、明日のテクノロジーを今すぐご利用いただけます。Watsonのテクノロジーを活用するすべてのIBM自然言語の分析ソフトウェア製品の詳細については、以下の各サイトをご覧ください。

The Hertz Corporation事例:

<http://ibm.com/software/jp/data/casestudies/search/hertz.html>

IBM Content Analytics(ICA)ソフトウェア:

<http://ibm.com/software/jp/data/search/textmining.html>

IBM Classification Module(ICM)ソフトウェア:

<http://ibm.com/software/jp/data/search/classify.html>

IBM eDiscovery Analyzer(eDA)ソフトウェア:

<http://ibm.com/software/jp/data/archive/ediscovery.html>

IBM OmniFind Enterprise Edition(OEE)ソフトウェア:

<http://ibm.com/software/jp/data/search/search.html>



---

## 日本アイ・ビー・エム株式会社

〒103-8510 東京都中央区日本橋箱崎町19番21号

© Copyright IBM Corporation 2011  
All Rights Reserved

06-11 Printed in Japan

このカタログに掲載されているサービスは2011年6月のもので事前の予告なしに変更することがあります。

IBM、IBMロゴ、ibm.com、OmniFind、Watsonは、世界の多くの国で登録されたInternational Business Machines Corp.の商標です。

他の製品名およびサービス名等は、それぞれIBMまたは各社の商標である場合があります。

現時点でのIBMの商標リストについては、[www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)をご覧ください。

掲載されている製品・サービスはIBMがビジネスを行っているすべての国・地域でご提供可能なわけではありません。

---

詳細情報のお問い合わせは、

日本アイ・ビー・エム株式会社 ソフトウェア事業 ECM営業部  
[CMoffice@jp.ibm.com](mailto:CMoffice@jp.ibm.com)

---