

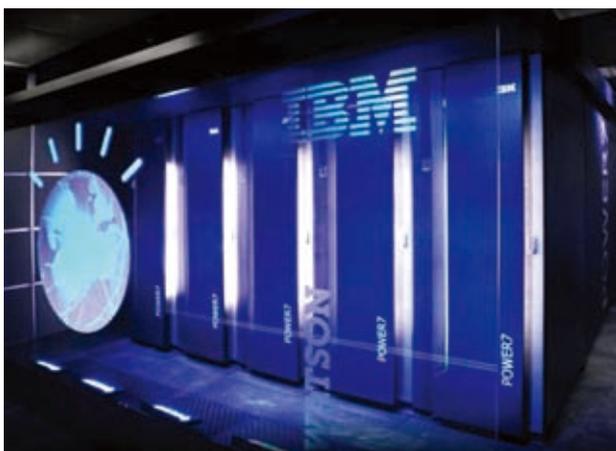
IBM Content Analytics软件助力 Watson在Jeopardy! 上获得成功

它能为您的企业做些什么?



“Watson背后的技术代表着数据管理和分析的未来。在真实世界中，此技术将帮助我们揭示从交通到医疗的每个方面的洞察。”

— John Cohn, IBM院士,
IBM Systems and Technology Group



什么是Watson?

IBM Watson DeepQA是全球最先进的问答机器，它通过理解自然语言问题上下文中隐含的含义来揭示答案。通过将先进的自然语言处理(NLP)与DeepQA自动问答技术相结合，Watson代表着分析、系统设计以及内容和数据管理的未来。Watson利用核心内容分析，以及众多其他先进技术，在非常短的时间内得到单一、准确的答案。此技术的业务应用是无限的，从临床医疗和客户关怀到政府情报等等。关于Watson的更多信息，请访问ibmwatson.com。

Watson的工作原理

Watson像人的大脑一样分析Jeopardy!问题的内容。首先，它尝试理解问题以确定所询问的内容。在这么做的过程中，它首先需要分析自然语言文本。接下来，它通过分析各种不同的内容(大部分都是自然语言文档的形式)来尝试找到合理的答案。最后，Watson基于一个置信度评分，评估并确定所找到的答案正确的几率。

Stephen Baker在图书《*Final Jeopardy: Man vs. Machine and the Quest to Know Everything*》中描述了该挑战的一个出色例子。“‘When 60 Minutes premiered, this man was U.S. President.’ 传统上，计算机很难理解‘premiered’的含义，并且这与一个时间相关联。对于计算机，‘premiere’可能也表示‘premier’。这是一个关于人的职位还是公司开业的问题？然后，当‘60 Minutes’出现时，它必须确定一个日期。然后找出谁是那时的‘U.S. President’。”简言之，它需要理解大量的上下文。

为什么这对您的业务意义深远?

Jeopardy!带来了一组与如今的业务领域中所发现的挑战类似的上下文信息挑战：

- 所存储的超过80%的信息存储是非结构化的(基于文本)。
- 理解这超过80%的部分可不简单。就像Jeopardy!一样，为了弄清楚含义和启示，微妙的语义、讽刺、谜语、同义词、缩写和其他复杂性必须被解决，它们带来了结构化数据中找不到的独特计算挑战。这就是NLP存在的意义。

在Watson中使用的核心NLP技术现在已经应用于提供业务价值。IBM Content Analytics软件使用了Watson中的相同功能来揭示您目前拥有的众多系统中多种格式的海量非结构化信息中蕴含的价值。理解这种非结构化信息的内容、上下文和价值为您的企业带来巨大商机。通过利用Watson中的技术，IBM软件解决方案将Watson的强大力量融入您的业务。

IBM Content Analytics软件将帮助您的企业迅速获得宝贵的洞察，这些洞察目前隐藏在非结构化的信息中：

- 从多个类别和外部来源和类型聚合和提取内容，包括企业内容管理(ECM)存储库、结构化数据、社交媒体、呼叫中心日志、研究报告、演讲脚本、电子邮件、安全报告和法律合同
- 使用NLP和其他分析来组织、分析和可视化企业内容(和数据)，理解含义并识别趋势、模式、关联、异常和业务上下文
- 交互式地搜索和探索，通过确认哪些内容值得怀疑或揭示新内容来迅速获得洞察，所有这些都无需构建模型或部署复杂的系统

案例研究

以下组织如今已从他们的内容中获得了宝贵的洞察。

医疗研究——类似于大部分医疗服务提供商，BJC Healthcare将大量的历史信息埋藏在非结构化的临床记录和诊断报告中，这些记录和报告包含对研究疾病进程、治疗效果和长期成果至关重要的信息。它现有的生物医疗信息资源是脱节的，而不是可互操作的，仅可用于少部分研究人员并常常闲置着，无法发现隐藏在非结构化临床记录、诊断报告中的广泛的研究信息。

使用IBM Content Analytics软件，BJC和大学研究人员现在能够分析非结构化信息来回答以前无法解决的关键问题。一些问题的示例包括：患者吸烟吗？如果吸烟，那么他吸烟的频率和烟龄是多少？如果已经戒烟，那么戒烟多长时间了？患者自己用过哪些药物？患者自己对病情是如何判断的？对患者执行了哪些治疗过程？BJC现在能够更深入地了解医疗信息，可揭示内容中存在的趋势和模式，向患者提供更有有效的医疗服务。

客户满意度——对于大部分企业来说，识别客户关于产品、服务和员工的满意度的趋势都很重要。Hertz Corporation and Mindshare Technologies是一家著名的企业反馈解决方案提供商，它使用了IBM Content Analytics软件来检查客户的调查数据，包括文本消息，以更好地识别汽车和设备租赁业绩水平，确定客户的满意水平并执行必要的调整来改善它们。通过使用IBM Content Analytics软件，Hertz推出了新的营销活动，调整了其产品和服务来满足其客户的需求。

“Hertz 每天会收集大量的客户洞察，包括来自 Web 调查的数千条评论、电子邮件和文本消息。我们希望利用这些洞察在战略上和局部上同时促进运营改进。”

— Joe Eckroth, 首席信息官, The Hertz Corporation
(ibm.com/press/us/en/pressrelease/32859.wss)

研究分析——对于北卡罗来纳州立大学来说，大学的本质不仅仅是教育培训，而是以多种形式推进和传播知识。北卡罗来纳州立大学面临的一个主要问题是处理可供它使用的大量数据来源。该大学想要一种解决方案来有效挖掘和分析庞大的数据量，以更好地识别可将其科研成果投入实际生产的公司。它的目标是这样一个解决方案：能够分析来自非结构化信息源的数千项内容，执行数据和文本分析，以及提供一组针对性的有用结果。

使用IBM Content Analytics软件，该大学能够将查找目标公司所需的时间从几个月缩短到几天。结果是识别了新的商业化机会，测试表明报考该学校的学生数量增长了大约300%。通过从北卡罗来纳州立大学广泛的内容来源中获取洞察，它的技术转让办公室 (Office of Technology Transfer) 能够以更经济有效的方式将本大学研制的技术授权给企业。



“该解决方案之所以如此强大，是因为它超越了传统的在线搜索方法，能够在结果中考虑上下文。”

— Billy Houghteling, 技术转让办公室执行主任,

北卡罗来纳州立大学(ibm.com/software/success/cssdb.nsf/CS/SSAO-8DFLBX?OpenDocument&Site=software&cty=en_us)

关于IBM企业内容管理

IBM企业内容管理软件使全球的著名公司能够更快制定更精明的决策。通过获取对非结构化信息的控制权，公司可以全新的方式开展协作并获取和影响业务决策，使内容成为洞察的最重要来源。借助针对具体行业的IBM ECM解决方案，公司能够捕获、管理和共享内容的整个生命周期，帮助确保合规性，减少成本和最大化生产力。IBM ECM产品组合包含广泛的功能，这些功能与现有系统相整合，帮助组织最大化信息的价值，这些功能包括文档捕获和成像、社交内容管理、高级案例管理、信息生命周期治理和内容分析。13,000多家全球公司、组织和政府机构依靠IBM ECM来通过创新改善绩效和保持竞争力。

通过利用为Watson提供强力支持的IBM Content Analytics功能，让未来的技术为您当前的业务服务。要了解所有利用了Watson技术的IBM ECM产品的更多信息，请访问以下网站：

IBM Content Analytics软件：

ibm.com/software/data/content-management/analytics

IBM Classification Module软件：

ibm.com/software/data/content-management/classification

IBM eDiscovery Analyzer软件：

ibm.com/software/data/content-management/products/ediscovery-analyzer

IBM OmniFind® Enterprise Edition软件：

ibm.com/software/data/enterprise-search/omnifind-enterprise

© 版权所有IBM Corporation 2011

IBM Corporation

Software Group

3565 Harbor Boulevard

Costa Mesa, CA 92626-1420

U.S.A.

在中国印刷

2011年8月

保留所有权利

IBM、IBM徽标和ibm.com是国际商业机器公司在全球许多司法管辖区注册的商标。其他产品和服务名称可能是IBM或其他公司的商标。可在网络上获得最新的IBM商标列表，请访问ibm.com/legal/copytrade.shtml上的“Copyright and trademark information”部分。

JEOPARDY! (c) 2011 Jeopardy Productions, Inc.。JEOPARDY!是Jeopardy Productions, Inc. 的注册商标。保留所有权利。

本出版物中对IBM产品和服务的引用不代表它们可用于所有IBM运营的国家。客户成功案例可从ibm.com/software/success/cssdb.nsf获得

本文中包含的信息仅供参考。虽然在检查本文信息时尽量保证其完整性和准确性，但它是“按原样”提供的，没有任何隐含或者明确的担保。此外，本文包含的信息根据IBM当前产品计划和策略提供，如有变更，恕不通知。IBM不承担因为使用本文内容和相关内容而造成损害的责任。本文中不包含的内容不打算、也不应该作为IBM或其供应商或其许可证销售商的担保或表示，或者修改适用于IBM软件的许可证协议的条款和条件。

每个IBM客户应负责确保遵守法律要求。对于可能影响客户业务的任何相关法律和规定要求的标识和解释，以及为符合这些法律读者可能必须采取的行动，客户自己负责获得合适的法律咨询。



请回收利用

ZZB03003USEN00