

# 采用认知系统转变组织的思考方式

IBM® Redbooks® 观点出版物，  
来自 IBM Academy of  
Technology



作者：**Rob High**（IBM 员工）  
和 **Bill Rapp**（IBM 杰出的工程师）

## 要点

认知系统可以转变组织思考、行动和运作的方式：

- ▶ 可以结合自然语言处理、假设生成和评估以及动态学习，提供强大、快速且准确的解决方案。
- ▶ 有助于理解非结构化数据的复杂性。
- ▶ 可以将高级分析应用于权衡和评估答案。
- ▶ 可帮助基于成果改善学习内容，使每一次迭代和交互都更加智慧。
- ▶ 可以使用深层自然语言处理（NLP）来评估各种主题中所使用的语言，然后作出明智的判断。

## 从大数据中发掘洞察力

数据量正在呈爆炸式增长。不仅是数据量，数据类型也在不断增加，我们需要加速对数据的理解。我们需要可以解读此类海量信息的计算机，以便从数据中发掘最大价值。此数据即**大数据**。

全球 **90%** 的数据是在过去两年中创造的。随着我们的世界变得更全面的互联互通以及更透彻的度量 and 感知，数据量还会继续增长。而世界上 **80%** 的信息是非结构化信息，其中包括以下文本类型：文学、报告、文章、研究论文、电子邮件、博客、推特、论坛、聊天内容和消息。

我们需要以更好的方式来理解所有这些数据，从中挖掘出可以应用于解决最紧迫且最重要问题的知识。我们需要能够对所阐述的内容加以理解，并使用此信息来解答问题、获取洞察力并促进制定更好、更明智的决策。但是，对于传统计算机系统而言，与人类相同的准确性来阅读此类信息并对其加以理解并非易事。我们需要新类型的计算系统，从而能够理解人类语言的精妙、特性、习语和用语的细微差别。这种计算方式即**认知型计算**。

认知系统模仿人类理解和处理信息的方式。不同于通过编程以进行快速计算并执行确定性任务的传统计算机，认知系统可分析信息，然后使用可能性分析从此分析结果中提取洞察力。系统在与数据的交互过程中学习改进，然后连续不断地有效地对自身进行重新编程。

IBM 正在领跑认知系统时代。认知系统正在转变我们使用计算机和技术来处理信息和自然语言的方式，从而创造机会来改变企业使用信息和制定决策的方式。

例如，医疗保健组织可以使用认知系统（例如，**IBM Watson™**）来分析所有可用数据（尤其是文本信息），以改善患者的治疗成果，同时使流程更高效。金融公司可以使用此技术来分析大量非结构化数据，以改善信用决策、投资分析和风险管理。认知系统可以转变将来组织思考、行动和运作的方式。



## 语言处理的复杂性

语言是表达思想以及沟通、理解、思考、推测及想象的媒介。它是我们理解周遭世界的途径以及我们传递和处理此理解内容的方式。

计算机难以理解人类语言的深度、复杂性和广度。我们可以打破差异、矛盾、不规则、分歧和缺乏条理性等限制，并且仍能准确理解彼此。计算机无法轻松以我们的方式来“理解万物”，在使用人类对事物的理解流程来分析大量数据时会出现问题，主要因为人类语言极为复杂。

这种精确性与准确性之间的差异至关重要。**精确性**是在文本段落上机械的或者科学的正确性。**准确性**是从一段话推断出的另一段话内容的真实程度。要准确回答一个问题，您必须考虑其可用环境。如果缺乏足够的证据信息，那么即使可以精确地回答部分问题，仍将难以准确解答问题。这种精确性对准确性的显著差别给正致力于理解人类语言的计算机又添加了另一层困难。

人类即使有时会困惑，但是仍可以轻松处理语言，因为他们可以通过比计算机更好的方式来解析信息的意义。这种能力是由于人类天生就有能力识别有关我们想要获取信息的使用环境。此概念是过去十余载人工智能群体想要实现的关键目标。虽然我们已提高了语言处理的精确度，但是，只有通过利用 **Watson** 才实现了计算机所需的精确度等级，使其在更广泛的自然语言现实世界中正常运作。

---

*有效处理大量非结构化信息迫切需要一种新计算时代的到来，我们将它称之为认知系统。*

---

## Watson 简介

**IBM Research** 百年来致力于科学发现和持续不懈地创造系统，使我们的世界更美好并以更智慧的方式生活，而 **IBM Watson** 正是其所实现的技术突破的佐证。

**Watson** 代表着技术系统识别模式、获取关键洞察力和增强决策制定能力的重大飞跃，尽管其复杂性仍令人敬畏。尤其是，**Watson** 是 **Deep Question-Answering (DeepQA)** 技术发展的重大进步。**Deep QA** 可促进对涉及一些主题所有内容的自然人类语言进行理解，然后对这些主题作出明智的判断。**Watson** 所配备的知识可以显著提升人机沟通能力。还可以扩展高级分析的能力，以理解大量结构化数据和非结构化数据。例如，**Watson** 可以使用 **DeepQA** 来提供关键且及时的信息，以帮助医疗工作者诊断和治疗患者。

## Watson 和语言

许多自然语言系统均在已形成的特定规则约束范围内强调精确性。这些系统会寻找一组特定词汇和同义词，而不对其使用环境进行评估。然后，记录找到这些词的次数。此概念通常被称为**浅层自然语言处理**或**浅层 NLP**，在狭义范围内可实现精确性，但是由于不考虑使用环境，因此常欠缺准确性。随着使用环境的变更，精确性也会降低。

当对准确性的需求高于精确性时，我们使用**深层自然语言处理**或**深层 NLP** 来对正在评估的问题中的使用环境进行分析。**Watson** 是深层 NLP 系统，它会尽可能对从即时信息、广泛可用的信息、知识库（称为**资料汇编**）以及从源数据库中提取的所有使用环境进行评估。

作为认知系统，**Watson** 可以将与人类相似的特征应用于传递和处理思想意见。它可分解语言以挖掘文本段落的推论信息，同时确切回答问题。当与传统计算系统相结合时，认知系统可以更快速、更准确且更弹性化地大规模解决大量信息所带来的问题。

但是 **Watson** 无需理解语言中的文字，而是理解人类使用的语言的特征。通过这些特征，它可以确定某一文本段落（称为问题）是否可推断出另一文本段落（称为答案），并且在不断变化的环境中仍能够保持高度准确度。

环境、时间和空间约束的概念都是不可或缺的，这些概念对于支持认知系统以类人的特征来运作方面至关重要。

---

*环境和推理在构建语言处理的基础中至关重要。*

---

## Watson 如何响应问题

**Watson** 是在上一代具备健全功能的计算系统上构建的，这些功能都不是特有功能，但是，通过结合这些功能，可以交付一个强大的解决方案：

- ▶ **自然语言处理**有助于理解非结构化数据的复杂性。
- ▶ **假设生成和评估**可应用高级分析，仅根据相关的证据来权衡并评估一系列响应。
- ▶ **动态学习**基于成果改善学习内容，使每一次迭代和交互都更加智慧。

**Watson** 正常运作的关键元素之一是资料汇编，其中包含各种非结构化知识，例如，文书、指南、操作手册、常见问题及解答、福利计划书和新闻。**Watson** 对整个资料汇编加以理解，以对内容进行修改并编辑成为一种

可运作的方式。它关注于资料汇编内容是否适当，剔除过时的、不相关的或者潜在不可靠来源的信息。

**Watson** 在资料汇编中将问题加以梳理、搜索出潜在的答案，然后以数百种方法来对答案及其使用环境进行验证。然后，使用生成的结果来确定所解释的问题和潜在答案的置信度。图 1 概括了 **Watson** 响应问题的过程。

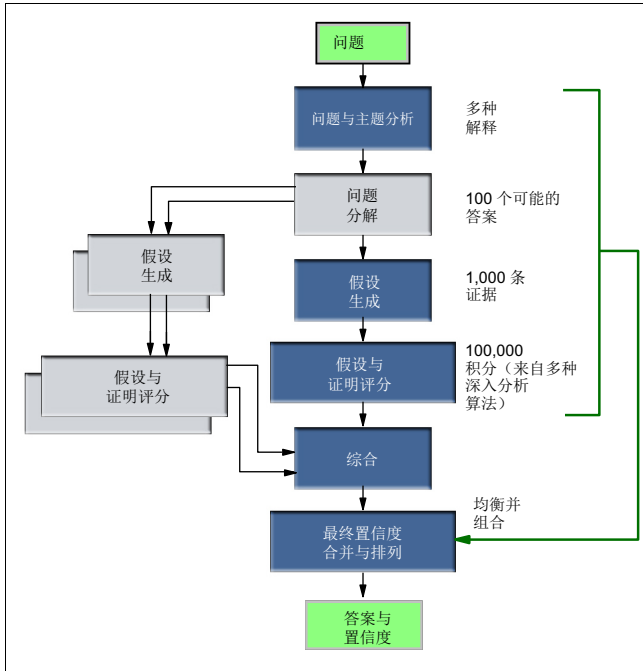


图 1 Watson 如何响应问题

将 **Watson** 所具有的功能加以组合即创造出了独一无二的解决方案，可以实现以下目标：

- ▶ 打破程式化计算的制约。
- ▶ 从对结构化的本地数据的依赖转移至全球化的非结构化数据世界。
- ▶ 从由基于决策树的确定性应用程序转变为与用户共同演进的概率性系统。
- ▶ 从基于关键字的搜索（结果涵盖包含或不包含答案的位置列表）转向一组按置信度排列的、直观的对话式答案。

## 使 Watson 投入工作

许多组织正在利用 **Watson** 与其他传统的计算形式一起，通过使用统计分析、规则和业务处理、协作和报告等方式，来解决各种业务问题。例如，您可以将统计分析与来自 **Watson** 的有关潜在事件的答案相结合，以警示投资风险。

## 医疗保健

我们都希望自己的医生能够知道有关我们健康问题的正确的、最佳答案。但是，由于生物医学数据量每七年翻一番的增长率，没有任何医学专家可以掌握改善我们的健康所需了解的一切信息。

**Watson** 可以帮助医疗保健行业处理、存储、检索和分析庞大的生物医学信息领域，以改善患者健康。

**Watson** 将永远无法取代细致入微的医学诊断。但是，它可以分析患者病情的复杂状况，并快速准确地为医学专家提供针对该患者的可行治疗方案，以缩小正确医疗决策的选择范围。例如，对于一名胸部疼痛的 45 岁患者，**Watson** 可以分析并结合患者的所有情况和病历，以帮助创建特定于该患者的治疗方案，从而提供更快更好的护理。

## 商业与金融

认知系统将大幅度改变企业与客户和员工的交互方式。**Watson** 有助于企业改善风险和估价流程。如果您拥有十分可靠且快速的系统，可以自信地提出适当的问题并选择适当的答案，那么您会如何更改自己的业务？如果您能够了解客户的购买模式和习惯，那么您会如何更改自己的产品线 and 零售流程？

金融行业中的数据量每年增长 70%。如何利用如此大量的数据，快速可靠地从中提取出真实的知识和关键洞察力？通过同时查看所有可用的金融和商业新闻、经济趋势和市场信息，您可以在分散的信息之间建立有价值的逻辑连接。例如，您可以分析此类信息，以寻找可能很快成为收购目标的公司。

## 客户服务

将 **Watson** 应用于市场的首要选择是在呼叫中心领域中，在此领域中，其功能有望更改呼叫中心的业务模型。

呼叫中心通常是各行各业投入大量成本的领域。许多呼叫中心接收到数百万次呼叫，呼叫者不得不忍受冗长、令人沮丧的长时间等待。**Watson** 的目标是更快速且更准确地提高服务质量，并提供正确的答案。**Watson** 是认知系统的下一个发展阶段，它将显著提升呼叫中心模式的效率，并支持用户按自己的条件来获取数据。

## 接下来：IBM 如何提供帮助

IBM 正致力于对 Watson 进行根本改进，这将进一步在医疗保健、金融、呼叫中心、政府和能源行业实现突破，从而帮助构建更智慧的地球。

IBM 将继续开发 Watson，使其成为更智慧、更精简且更快速的专家顾问程序，为各种领域中的人类专家提供帮助。要了解有关 IBM Watson 的更多信息，请参阅：

[ibm.com/innovation/us/watson](http://ibm.com/innovation/us/watson)

### 可提供更多信息的资源

使用以下资源以获取有关本文中强调的概念的更多信息：

- ▶ *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  
[ibm.com/innovation/us/watson](http://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>

# 声明

本信息是针对在美国提供的产品和服务编写的。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。但是，评估和验证任何非 IBM 产品、程序或服务，则由用户自行负责。

IBM 公司可能已拥有或正在申请与本文档内容有关的所有专利。提供本文档并未授予用户使用这些专利的任何许可。您可以用书面方式将许可查询寄往：  
*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*

## 本条款不适用英国或任何这样的条款与当地法律不一致的国家或地区：

INTERNATIONAL BUSINESS MACHINES CORPORATION “按现状”提供本出版物，不附有任何种类的（无论是明示的还是暗示的）保证，包括但不限于暗示的有关非侵权、适销和适用于某种特定用途的保证。某些国家或地区在某些交易中不允许免除明示或暗示的保证。因此本条款可能不适用于您。

本信息中可能包含技术方面不够准确的地方或印刷错误。此处的信息将定期更改；这些更改将编入本资料的新版本中。IBM 可以随时对本资料中描述的产品和 / 或程序进行改进和 / 或更改，而不另行通知。

本信息中对非 IBM Web 站点的任何引用都只是为了方便起见才提供的，以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

本信息包含在日常业务操作中使用的数据和报告的示例。为了尽可能完整地说明这些示例，示例中可能会包括个人、公司、品牌和产品的名称。所有这些名字都是虚构的，若现实生活中实际业务企业使用的名字和地址与此相似，纯属巧合。

此处包含的任何性能数据都是在受控环境中测得的。因此，在其他操作环境中获得的数据可能会有明显的不同。有些测量可能是在开发级的系统上进行的，因此不保证与一般可用系统上进行的测量结果相同。此外，有些测量是通过推算而估计的，实际结果可能会有所不同。本文档的用户应当验证其特定环境的适用数据。

版权许可：

本信息包括源语言形式的样本应用程序，这些样本说明不同操作平台上的编程方法。如果是为按照在编写样本程序的操作平台上的应用程序编程接口（API）进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例并未在所有条件下作全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。

本文档“REDP-4961-00”创建或更新于 2013 年 4 月 3 日。

IBM®



# 商标

IBM、IBM 徽标和 [ibm.com](http://ibm.com) 是 International Business Machines Corporation 在美国和 / 或其他国家或地区的商标或注册商标。这些术语和其他 IBM 已注册商标的术语在本信息中首次出现时都使用适当的符号（® 或 ™）加以标记，以表示在本信息发布时由 IBM 在美国注册或拥有的普通法商标。些商标也可能是其他国家或地区的注册商标或普通法商标。在 Web 地址 [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml) 中包含了 IBM 商标的最新列表。

以下术语是 International Business Machines Corporation 在美国和 / 或其他国家或地区的商标：

IBM®  
IBM Watson™  
Redbooks®  
Redbooks (徽标) 

以下术语是其他公司的商标：

其他公司、产品或服务名称可能是其他公司的商标或服务标记。