

关于 Watson 你必须要了解的几点

孔鹿

代表着 IBM 在认知计算领域最核心的技术 Watson(沃森)的一战成名,是在 2011 年 2 月的美国问答节目《Jeopardy!》上。在这次节目中,Watson 战胜了这一节目的两位冠军选手,这被和 1996 年同样来自 IBM 的“深蓝”战胜国际象棋大师卡斯帕罗夫相提并论,被认为是人工智能历史上的一个里程碑。

不过,在 IBM 董事长、总裁兼 CEO 罗睿兰(Ginni Rometty)宣布 IBM 已经转型为一家认知解决方案和云平台公司的当下,Watson 已经不再是会答题那么简单。

认知时代的商机

从技术角度来看,2011 年参加“Jeopardy!”电视问答挑战赛时,Watson 做了一件事——用自然语言进行深度问答。但问答只是 Watson 具备的众多能力之一,截至 2015 年 10 月,Watson 已经拥有包括问答在内的 28 项能力。

据 IBM 全球认知商业行业解决方案总经理 Neil Isford 透露,Watson 的这些能力如今均已被转变成数字服务或 API,在 IBM 内部将其视为认知能力的积木。包括关系抽取、性格分析、情绪分析、概念扩展及权衡分析等在内的 IBM Watson API 数量将在 2016 年达到 50 项。

而这背后,IBM 并没有打算让以 Watson 为代表的人工智能系统远离公众生活,这家“百年老店”已经充分表露了对 Watson 商业化的信心。

不可忽略的市场大环境是,随着物联网的普及,每个设备都在产生数据。然而 80% 的数据无法被计算机辨识,这 80% 的数据中既包括使用人类语言编写的书面材料,从教科书和公式到文学作品和对话;还包括系统从听、说和肢体动作中,捕获到的各类数据。

有调查数据预测,这些“非结构化”的数据增长速度飞快,到 2020 年数据总量将超过 44ZB (Zettabyte,十万亿亿字节),将占到全球数据总量中的一大部分。

如果从具体行业来看,在未来两年,医疗数据将增长 99%,其中 88% 的医疗数据都将是非结构化数据,包括电子病历、化验结果、医学影像、视频以及病患传感器(如可穿戴医疗设备);政府和教育数据也将增长 94%,其中 84% 是非结构化数据,这些数据来自各类传感器、建筑物、道路、车队等;传媒业的数据将增长 97%,其中 82% 是非结构化数据,这些数据包括书籍、期刊、报纸和其他出版物,以及视频、电影、录音及在线游戏等。

而 Neil Isford 的观点是,计算技术正在进入一个全新的时代,我们叫作认知的时代。数据呈现爆炸性增长的趋势之下,这些数据所蕴含的尚未被挖掘价值,在认知时代这意味着无限商机。“进入到认知的时代,这是第一次我们可以把这些无形的数据更好地运用起来,不管是我们的教科书,或一些信息,或一些文件,来自于传感器的数据,在运动当中产生的数据,我们都能够把数据结合起来进行挖掘,以前是没有发现它的价值,现在可以从中获得一些商业洞察。”

IBM 大中华区董事长陈黎明说:“认知时代已经开启。在不久的将来,我们将能看到整个商业模式由于认知技术的推动而发生巨大变化——小到每个人获得的服务和产品、创业者所能拥有的商业创新优势,大到传统企业行业的转型、甚至经济和整个社会治理效率的跨越式提升。我们坚信,认知商业是大势所趋。”

认知计算与人工智能

IBM 将认知系统的三项重要特质定义为理解、推理、学习。

所谓理解,就是通过感知和互动快速理解结构和非结构化数据,能够依据文本资料和感知与用

户进行交互，并理解和回答用户的问题。认知系统可以在理解的基础上，通过假设，推理并揭示洞察、发现模式和关系，实现以多种方式认知和产出多种结果而不仅仅是一种结果的传统方式，以帮助人们做出更好的决策。

此外，通过以证据为基础的学习能力，认知系统能够从所有文档中快速提取关键信息，使其能够像人类一样进行持续的学习。通过追踪用户对自身提出的解决方案和问题解答的范库和评价，以及专家训练，能够不断进步，提升解决方案和解答的能力。

尽管 IBM 提出 Watson 和认知计算多年，但在科技领域对于认知计算和人工智能两者的关系依然很模糊。

这一次，IBM 终于愿意出面将认知计算与人工智能说清楚。

来自 IBM 的观点是，“火了”二十多年的人工智能概念从历史和研究角度来讲主要目的是为了机器表现得更像人，我们称之为 IntelligentBehavior。

对此，IBM 也承认其认知计算从技术角度上来讲和 AI 有很多共性的地方，比如机器学习 (MachineLearning) 深度学习(DeepLearning)等方面都很类似。

但是 IBM 的认知计算目的并不是为了取代人，或者说 IntelligentBehavior 只是认知计算的一个维度，在讲认知计算的时候除了要能够表现人和计算机的交互更加自然之外，还会更多强调推理的部分，自学习的部分以及怎样把这样的能力结合具体的商业应用，解决商业的问题。后两个维度不是传统做人工智能的人关心的维度，他们更关心的只是怎样表现得更像人。

换句话说，认知计算不是制造为人们思考的机器，而是与增加人类智慧有关——认知计算系统通过与人的自然语言交流及不断学习帮助人们做到更多，使专家可以更好地从海量复杂的数据中获得更多洞察，从而做出更为精准的决策。

认知计算广义上讲是通过获取海量的不同类型的数据，根据信息进行推论，从自身与数据、与人们的交互中学习，并以对人类而言更加自然的方式与人类交互。它最重要的目的是如何整合这些能力，并结合具体的商业应用场景，来解决商业上的问题，帮助企业实现商业变革。现在，企业正面临着大数据带来巨大的挑战。传统计算方式会错过世界上 80% 的信息(非结构化数据)，而认知技术能够支持组织去发现数据中隐藏的模式，挖掘出令人惊喜的新的商机，也能够加速发现新药、发现新的登月方式，甚至发现未知领域。