

大数据时代知识管理的新风向

吴若溪

(黑龙江大学 信息管理学院 哈尔滨 150080)

摘 要 大数据是一个新兴的、包容性非常强的概念,对各个行业和学科都有着很大的影响。知识管理作为一个与信息息息相关的数据学科,对大数据的兴起很敏感。如何在大数据时代中重新认知和定位知识管理,对于提升知识管理的学科能力和学科地位都有着很重要的作用。

关键词 大数据;知识管理;知识管理系统

中图分类号 G20

文献标志码 A

文章编号 1002-2589(2015)19-0161-02

引言

大数据的兴起是计算机科学发展产物,计算机科学的发展使人们能够实现对事物的全面描画,而由此产生的大量数据催生了大数据这样一个概念,以及围绕大数据产生的各种数据技术。科研数据、互联网个人数据、企业数据以及社会和政府数据都在迅猛增长。大数据环境下的知识管理,有着很多不同以往的新特点,这些新特点既是知识管理本身更新换代所必需的品质,也是嫁接大数据技术所引发的知识管理革命。

一、大数据的概念与特点

大数据,顾名思义,即巨量的数据。但是相对于传统的小数据来说,大数据并非简单的量的增长,而是借由量变达到质变,大数据已经呈现出许多新的特质,所以要求人们用新的思维去看待和利用大数据。一般来说,大数据指的是所涉及的数据量规模巨大到无法通过目前主流软件工具,在合理时间内达到抽取、管理、处理、并整理的信息^[1]。大数据的特点一般被概括为 4V,即 Volume(大量)、Velocity(高速)、Variety(多样)、Value(价值)^[2]。

大数据所应对和处理的数据量非常庞大,是传统的数据分析和处理难以想象的。面对如此庞杂的数据,大数据还要求非常高的处理速度,甚至达到实时分析。大数据一般来说价值密度特别低,但是经过数据挖掘之后,所产生的数据价值就会特别大,所产生的外部效应也是小数据时代所不能相比的。

二、大数据时代中知识管理的新特点

大数据是互联网以及计算机科学发展必然结果,几乎所有的领域和行业都面临着大数据时代的转型任务。知识管理作为一门面向知识与信息的学科,必然要与大数据短兵相接。特别是对于大型企业以及社会层面的知识管理来说,大数据是一个必不可少的研究课题。

(一)知识管理系统所处理的数据量与信息量激增

知识管理是建立在对信息和数据搜集整理的基础之

上,任何一个知识管理系统都是由最基本的信息和数据建立而成,而后才是进行知识抽取和转化的过程。知识管理最初出现的时候,所面临的信息量和数据量并不多。第一代知识管理阶段(20 世纪 70 年代至 80 年代初中期),企业知识管理是以邮件及文档数据库为核心的办公业务自动化管理,重点在文档管理。在当时,纵使对于一家巨型企业来说,其所拥有的文档数量在 21 世纪的今天也算不上巨量。一方面是当时的数据生产和存储能力有限,使企业生产、搜集和存储数据和信息的工作受到限制。当时的生产并未实现完全的电子化,更没有如今琳琅满目的各种传感器时刻监控着企业各方面的运行,所以并未累计出可观的数据量。另外,当时的数据存储技术并不发达,数据存储的代价很高,也限制了企业存储大量的数据。另一方面,当时的数据处理技术也没有今天这么发达,企业存储的大量数据既无法及时也无法有效地得到处理和分析,使企业也没有主观意愿去存储大量的数据。

进入 21 世纪特别是近几年,不论是数据的生产还是存储以及处理技术,都得到了质的提升,计算机以及互联网科学以前所未有的势头迅猛发展,使企业完全有能力处理以前无法想象的海量任务。IDC 于 2013 年发布的《2013 年德国存储情况调查:通过高效的存储技术优化计算中心》报告中,调查了 219 家德国企业,并预测未来企业数据量增长将达到 40%左右^[3]。

企业数据量的增长,并非企业知识管理系统的简单扩容就能解决。量变引起质变,企业数据量的增长已经为企业数据管理带来了一些新的特点,用以前企业知识管理的理念和技术已经无法应对如此海量的数据。企业必须以全新的知识管理理念和技术来革新企业知识管理系统,以应对大数据带来的新挑战。

(二)知识管理系统的运行速度更快

随着技术的高速发展,知识管理系统的更新速度与处理速度都得到了质的飞升。但是,大数据时代所要求的知

收稿日期 2015-04-24

作者简介 吴若溪(1989-),女,山东莱州人,硕士研究生,从事信息生态、知识管理、图书馆等方面研究。

识管理系统的更新速度与处理速度,已经超过了以往知识管理系统传统的提速方式的逻辑负荷,必须以新兴的大数据技术来重塑整个业务流程。

首先,知识管理所存储的数据与信息总量飞速增长,那么知识管理系统的运行速度也势必要同步增长,才能促进知识管理系统的及时更新与流畅处理。其次,知识管理在企业运行中所扮演的角色也更加要求及时性。知识管理既有积累隐性知识的慢节奏任务,也有要求快速产出显性知识的快节奏任务,特别是现代企业竞争十分激烈,要求企业有极强的知识更新速度,相应地也要求企业的知识管理系统具备很高的运行速度。最后,计算机和互联网技术,特别是大数据处理技术的发展,为企业知识管理系统提升运行速度提供了技术基础,使企业得以开展大数据战略。现代企业的电子化程度越来越高,企业运作的整个流程,不论粗细,都能被企业的知识管理系统收录,从而产生和积累了大批量的数据。在这个过程中,数据和信息的积累速度超过历史上的任何时期,企业的任何一个动作都会被迅速地纳入到数据库中,不仅总量巨大,而且流量也大。

(三)知识管理需要应对的数据和信息更加多样化

传统的知识管理系统所面对的信息和数据格式,多以结构化数据为主。在企业数据中,非结构化数据和半结构化数据占据了绝大部分,而结构化数据只占了一小部分。这一小部分的结构化数据虽然价值较大,但从价值总量上来讲,非结构化数据和半结构化数据的价值总量仍不可忽视。

随着大数据技术的发展,企业的数据库与知识管理系统越来越能处理更多样化更多格式的数据和信息,不仅包括传统的文档,更包括迅猛增长的图像与视频等丰富格式的数据和信息。

(四)知识管理系统的价值密度降低和价值总量增加

与传统的知识管理系统很大不同的是,大数据时代的知识管理系统,其所包含的价值密度会降低。这一方面是因为知识管理系统中数据和信息总量的迅猛增长,另一方面也是因为在大数据环境下的知识管理系统中大部分数据和信息都是机器自动搜集和组织的,这与传统中人工进行搜集和组织相比,价值密度必然会降低。但价值密度的降低,并不代表大数据的出现降低了知识管理系统的总体价值。在知识管理系统中价值密度降低的同时,其所包含的数据和信息量正呈几何级数增长,海量数据的价值总量非常巨大,远远超出价值密度低所带来的负面影响。

(五)知识管理系统的智能化

传统的知识管理系统,不论是知识库的建设还是扩展上,很大程度上依靠人的智力与智慧。而大数据时代的知识管理系统,不仅能够替代人们自动地建设知识库,还能像人类一样具有智力与智慧去思考和推理,并生产和组织知识。

谷歌公司于2014年8月末公布了在建的“知识库”(Knowledge Vault),这是谷歌继2012年开始的“知识图谱”(Knowledge Graph)之后又一大作。“知识库”通过算法自动搜集网上信息,通过机器学习把数据变成可用知识。目前,Knowledge Vault已经收集了16亿件事实,其中2.71亿件是“可信的事实”^[4]。据专家估计,大型企业特别是大型的科技企业,都在建设类似的知识库。当机器系统掌握了全人类的知识,它的智力将远远超过现有的智能助手。

知识管理系统的智能化发展,将把人类从知识管理流程的低端阶段中解放出来,从而更专注于高级知识的生产

和应用。这将是大数据时代中的知识管理系统对传统的知识管理系统的一大革命。

三、大数据时代中的知识管理关键技术

(一)云计算

云计算是一个与大数据相伴相生的概念,大数据是云计算的计算对象,云计算是大数据兑现的工具。如果说大数据是一个无尽的宝藏,那么云计算就是开发宝藏的利器。首先,云计算为企业知识管理系统中的大数据提供了可以弹性扩展、成本较低的存储空间和计算能力,以使中小型企业也可以像大型企业一样利用大数据的技术来完成知识管理系统的构建和运行。其次,云计算的计算模式丰富、计算资源庞大,可以为异构系统较多的各类企业提供经济适用的数据处理方式。

(二)数据挖掘

面对庞杂的各类数据和信息,传统的知识管理系统只能依靠人力智能去完成知识的发现。而在大数据时代中,数据挖掘技术的发展使人们可以应用技术轻松地从事数量巨大的数据和信息资源中挖掘出各类有价值的知识,从而更有效地建设知识库。

(三)可视化

数据和信息的可视化由来已久,但一直久盛不衰。特别是近几年,数据可视化和信息可视化技术进入了新阶段的蓬勃发展期。可视化技术是知识管理系统不可忽视的一项重大技术基础,因为可视化是促进知识传播和转化的极重要工具。首先,可视化技术将知识的表现形式丰富起来,使人们更乐于去接受知识;其次,可视化技术降低了知识的理解吸收难度,促进知识在组织内更迅速和容易地传播开来;最后,可视化是一个知识管理系统中所有知识的总结性表达,更有利于对知识库中的知识进行全面掌控,从而促进知识的更新迭代。

(四)机器学习和人工智能

未来的知识管理系统,是一个相对独立的大脑,能代替人类进行一般的知识思考和推理,从而将人类从低端的知识管理流程中解放出来。机器学习和人工智能就是进行这一任务的主要技术,它们将机器构造成一个具有学习能力和理解能力的大脑,从而实现知识管理的智能化。

四、结语

大数据是知识管理绕不开的一个重要概念,知识管理唯有积极拥抱大数据、利用大数据革新自我,才能在大数据时代找到立足之处。一方面要重新定位和认识知识管理的业务流程和学科原理,另一方面也要积极应用大数据时代中的各类实用技术,使知识管理不仅从内涵上,更从实践上成为大数据时代中的佼佼者。

参考文献:

- [1]孟小峰,慈祥.大数据管理:概念、技术与挑战[J].计算机研究与发展,2013(1).
- [2]大数据[EB/OL].[2014-8-20].<http://baike.baidu.com/subview/6954399/13647476.htm>.
- [3]IDC:未来八年地球产生数据量将超40泽字节[EB/OL].(2014-8-22)[2015-03-23].<http://network.51ctacom/art/201212/372335.htm>.
- [4]张林.谷歌KnowledgeVault知识库准确性90%[EB/OL].(2014-8-22)[2015-03-23].<http://soft.zol.com.cn/474/4747352.html>.

(责任编辑 张 璐)