

# 数据挖掘与知识管理的结合应用

北京信息职业技术学院 罗锦珠

[摘要] 21 世纪,知识已成为经济增长和社会发展及企业成长的关键性资源,搞好知识管理已成为全球企业,生存发展的主要支柱。知识管理是知识有效利用的手段。数据挖掘是知识管理的基础,是一类深层次的数据分析,是实现数据上升到知识的必然过程。

[关键词] 知识 知识管理 知识转化 数据挖掘 数据挖掘的实现

## 一、知识和知识管理

### 1. 知识

管理大师彼得·德鲁克(Peter. F. Druker)指出:“知识是一种能够改变某些人或某些事物的信息(information),这既包括使信息成为行动的基础的方式,也包括通过对信息的运用使单个个体(或机构)有能力进行改变或进行更为有效的行为的方式”。另一位学者车驰曼(Churchman)认为:“将知识设想或看作为一种对信息的集合的观点,事实上已经将知识这一概念从其全部生活之中剥离出去;知识只存在于其使用者身上。而不存在于信息的集合中。使用者对信息的集合的反应才是最为重要的”。从这两个定义可以归纳出,知识来源于信息,但信息必须结合人的自身知识及实际行为作用,才具备延伸为知识的基础。以信息表现出来的知识称为显性知识,存在于人头脑中的自身知识称为隐性知识,包括经验、教训、技能等。

### 2. 知识管理

知识管理作为获取、存储、学习、传播、应用、共享知识的一种管理方式,对于当今的企业是决定着命运的一种管理,是继工业时代泰勒的管理革命之后的又一次管理上的新革命。搞好知识管理已成为全球企业,生存发展的主要支柱。

Yogesh Malhotra 博士认为,知识管理是企业面对日益增长的非连续性的环境变化时,针对组织的适应性、组织的生存和竞争能力等重要方面的一种迎合性措施。本质上,它包含了组织的发展进程,并寻求将信息技术所提供的对数据和信息的处理能力以及人的发明创造能力这两方面进行有机的结合。

对知识管理的内涵的阐述,可以得到下面对知识管理的三点界定:

第一,知识管理是一个知识生产以及利用的过程。

第二,知识管理需要相关技术和企业内部结构的支撑,要建立一个知识交流的环境。

第三,知识管理需要充分与人相联系。知识管理只有和人及其工作相结合,才能真正发挥效用。

知识管理的重点是知识的识别、获取、开发、分解、存储和共享,并为其构建有效的途径和机制,以运用集体的智慧提高企业的应变和创新能力。

### 3. 企业知识管理

企业知识管理是一种新型的管理模式,它的基本要求如下:

第一,企业知识管理是以知识生产管理为前提。知识的生产不同于其它物质和服务产品,它是在对知识

## 三、结论

综上所述,从库恩的“科学范式”到“技术范式”、“技术经济范式”和“创新范式”,范式在其概念和内涵上发生了显著的变化。但从总体上看,“科学范式”、“技术范式”、“技术经济范式”和“创新范式”,有一个共同点,即都在某种程度上沿袭了范式作为范例和模式的本意。它们的显著区别就在于,“科学范式”更多是从抽象的思辨角度阐释科学的动态发展,而“技术范式”和“技术经济范式”则对技术系统的发展及其对经济系统的现实影响和作用予以关注。至于“创新范式”,它所考察和解释的不只是存在竞争和选择的动态发展过程,也包括从国家的宏观层面上考察社会系统、技术系统和经济系统之间的交互作用与演进。

### 参考文献

- [1] Thomas s. Kuhn The Structure of Scientific Revolutions [M]. Chicago & London: the University of Chicago Press 1962.
- [2] G. Dosi, Technological Paradigm and Technological Trajectories: A Suggested Interpretation of the Determinants and Directions of Technical Change [J]. Research Policy, 1982, (11):147 - 162.
- [3] Koenraas Debackere, Bart Clarysse, etc. Science and Industry: A Theory of Networks and Paradigms [J]. Technology Analysis & Strategic Management, 1994, 6 (1).
- [4] Fumio Kodama. Analyzing Japanese High Technologies: The Technological Paradigm Shift [M]. New York: Printer Publishers, 1991
- [5] Freeman, C., and C. Perez. Structural Crises of Adjustment, Business Cycles and Investment Behavior. In: G. Dosi., C. Freeman, R. Nelson, G. Silverberg. And L.L. G. Soete, eds. Technical Change and Economic Theory [M]. London: Frances Pinter, 1988. pp. 38 - 61

的识别、获取、开发、研究、分解、使用和共享的全过程中进行的,在这些过程中,知识既被使用,又在不断地产生、形成和发展,新的知识在产生,原有的知识在发展。

第二,知识管理是以信息管理为基础,是信息管理的延伸和发展。信息管理是将各种各样的信息以一定的方式汇总、组织起来,方便人们利用计算机来查询和检索,利用技术来进行信息采集、存贮和控制。知识管理是通过一组解决方案的集合寻找和识别与问题有关的关键性信息,并将这些信息进行提取,形成对某一问题的专门知识,作为决策的依据。有人将知识管理混同于信息管理,其实不然,知识与信息无论在内涵和外延,以及表现形式上都有所不同。近年来,由于经济发展的需要,知识管理开始从信息管理中孵化出来,逐步形成适应知识经济的新的管理模式。知识管理要求把信息与信息、信息与活动、信息与人联结起来,在人际交流的互动过程中实现知识的共享、运用和创新。从信息管理到知识管理的进化,是管理理论与实践“人本”管理主线的进一步体现和发展。正如库珀所言:“正是由于信息与人的知识能力相结合,才导致知识的产生。它是一个运用信息创造某种行为的过程。这就是知识管理的目标”。

第三,企业的知识管理是为企业实现知识的显性化和共享寻找新的途径。关于知识的分类的内容在 5 中详细论述,这里不再赘述。

#### 4. 知识链

知识管理的实质在于知识的创新与共享,它并不在乎拥有多少知识,而是利用现存的知识进行创新,创造出新的价值,让需要知识的人,很方便的利用知识。知识管理注重“做正确的事情”(结果),而不是“正确地做事情”(过程)。

知识管理从“信息价值链”转向“知识价值链”。所谓“信息价值链”,认为技术系统是指导组织活动过程中的关键要素,而人只是相关的被动“处理机”,其作用是保证正确执行信息库中得到的“最佳方案”;“信息价值链”是以信息技术为核心,强调的是信息的价值;而“知识价值链”,认为人力系统才是最关键的组成部分,只有人们不断地认识和评估技术系统提供的信息,才能够更好地实现制定的目标。

知识的生产力将会成为一个公司,一个行业,甚至整个国家竞争地位的决定性因素。没有哪一个国家,行业或公司“天生的”具有优势或者劣势。它所能拥有的唯一优势就是开发利用全世界都可以利用的知识的能力。

#### 5. 知识的分类

根据知识的获取方式,可分为显性知识和隐性知识。

显性知识是指以文字、图像、符号表达,以印刷或电子方式记载,可供人们交流的结构化知识,如事实、自然原理和科学的知识、标准化作业程序、系统化的文件等。显性知识比较容易获得、理解和交流,它具有公共性,可以存储在图书馆、局域网或数据库中,其传播和复制的成本较为低廉。这类知识在总的知识存量中最多占 20%,所谓“你知道的总比你说不出的多”。这类知识采用编码管理模式(Codification mode),编码模式强调投资信息技术,以便开发能迅速收集、传播知识的管理系统,然后通过这些知识的重复利用获得收益。这种模式中的知识可以极为迅速、无限多次地重复利用,它节省了工作时间,减少了信息交流成本,信息技术的飞速发展使得编码管理更加方便了。

隐性知识是指很难用语言、文字表述,即“只可意会、不可言传”的知识。它是一种高度个人化的知识,比较复杂,无法用文字描述的经验式知识,相对主观,依附于人的大脑或技能中,通常通过行动表现出来,以及必须通过人际互动才能产生共享的知识,这种知识与拥有知识的人息息相关,由于隐性知识具有非结构化和专有属性,其传播成本较高,范围较小。对这种知识采用人物化管理模式(Personalization mode);人物化模式强调投资人力资源,一般是引进经验丰富专家、学者,鼓励他们直接与其他人员进行交流,来传播他们的知识。隐性知识由认知、情感、信仰、经验和技能等 5 个要素共同组成,可细化为个人隐性知识,集体隐性知识,专业隐性知识。

#### 6. 知识转化的形式

根据知识的划分形成不同的知识转化形式,从内隐——内隐、内隐——外显、外显——外显、外显——内隐四种模式,如下表。

内隐知识外显知识内隐知识知识社会化 知识内化外显知识知识外部化 知识组合化 以日本松下公司发明自动面包机为例,说明知识转化的这四种模式。当年,松下电器公司开发一款新型家用烤面包机时,遇到的一个难题,就是怎样让面包机揉出合格的面团。他们做了许多的试验,但是,烤出的面包总是“外糊内生”。当时大阪国际饭店生产的面包远近有名。公司一名软件工程师叫田中郁子直接到大阪饭店拜师学艺。田中跟着面包师口传心授,学了一年,最后学会了面包师拉面团的独特技术,此后,松下公司利用了这种技术,成功地开发出了面包烘烤机。

内隐知识——内隐知识,这是一个知识社会化的过程。例如田中郁子获得了面包大师的技术。这种内隐知识社会化的过程是通过观察、模仿、实践而实现。

内隐知识——外显知识,这是知识的外化。尽管内隐知识已经转化成另一个人的内隐知识,但是其它人仍然不能共享。尽管田中习得了面包大师的技术,但是其它成员仍不知晓。只有田中郁子将学到的内隐知

识转化成设计方案、或程序(外显知识),别人才能共享。

**外显知识**——外显知识,这是知识的组合,知识的总量并没有增减。将现有的外显知识汇总到原有的知识构架中,例如将设计、开发烤面包机的知识汇总到操作手册或工作手册,并在产品设计中体现出来。

**外显知识**——内隐知识,这是知识的内化,这是一个通过做而学的过程(learning by doing),是指新的显形知识在整个组织中实现共享和运用,其它成员又将进一步延伸、扩散,内化为自身新的隐形成知识,这又构成了一个新的循环起点。

## 二、数据挖掘

近十几年来,信息化进程加快,人们利用信息技术生产和搜集数据的能力大幅度提高,千万个数据库被用于商业管理、政府办公、科学研究和工程开发等,并且这一发展趋势仍将持续下去。面对“人们被数据淹没,人们却饥饿于知识”的挑战,如何才能不被信息的汪洋大海所淹没,又能够从信息的海洋中及时发现有用的知识,提高信息利用率。数据挖掘技术应运而生,并得以蓬勃发展,越来越显示出其强大的生命力。

**数据挖掘(Data Mining)**就是从大量的、不完全的、有噪声的、模糊的、随机的数据中,提取隐含在其中的、事先未知而潜在有用的信息和知识的过程。提取的知识表示为概念、规则、规律、模式等形式。也可以说,数据挖掘是一类深层次的数据分析,是实现数据上升到知识的必然过程。

数据挖掘是知识管理的基础。数据挖掘与知识发现是等同的概念。

**知识发现(KDD, Knowledge Discovery in Database)**是 1995 年,在加拿大召开了第一届 KDD 和 DM 国际学术会议。会议对 KDD 做了确切的定义。知识发现是一个在数据中识别新颖有效模式的重复过程。这些模式具有潜在的可用性,并且最终可以被理解。

知识发现的过程可以分为三个主要阶段:数据准备、数据挖掘和结果表达和理解。

### (1) 数据准备

①**数据集成**。将多文件或多数据库运行环境中的数据进行合并处理,解决语义模糊性,处理数据中的遗漏和清洗脏数据等。

②**数据选择**。为知识发现的目标搜索和选择有关的数据,这包括不同模式数据的转换和数据的统一和汇总。数据选择的目的是辨别出需要分析的数据集合,缩小处理范围,提高数据挖掘的质量。

③**数据预处理**。对数据进行清理和充实等预处理工作。

④**数据转换**。对数据编码,数据库中字段的不同取值转换成数码形式将有利于搜索。

### (2) 数据挖掘

此阶段进行实际的挖掘操作,利用机器学习、统计分析等方法,从数据库中发现有用的模式或知识。

### (3) 结果表达与解释

根据最终用户的决策目的对提取的信息进行分析,把最有价值的信息区分出来,并且通过决策支持工具提交给决策者。这一步骤的任务不仅是把结果表达出来,还要对信息进行过滤处理。如果不能另决策者满意,需要重复以上数据挖掘的过程。

## 三、数据挖掘在银行中的应用

银行卡在今天已被各家商业银行视为可带来丰厚收益的重点业务。与国外银行信用卡为主的产品结构相比,国内银行卡是以借记卡为主的产品体系,银行付出的是存款利息成本,却不能得到相应的增值收益,如信用卡管理年费、信用卡透支利息等。从目前的条件看,我国银行向一般客户发放贷记卡又确实存在困难,原因是需要掌握该客户的消费习惯是否支持他有效利用贷记卡,能够利用贷记卡先消费后还款的功能,再有,需要了解该客户的信用情况,以避免出现恶意透支等使银行遭受损失。从银行角度可看,银行确实掌握着大量的客户信息,这些信息如果没有完整的数据仓库和挖掘工具,仅靠传统业务系统和银行掌握的借记卡客户信息,根本无法完成是否可向该客户发放贷记卡的业务决策,也无法对其用卡情况进行实时监督或开展相关金融产品营销。

如果银行可以利用数据挖掘工具,支持他们做客户信息分析,就可获得相当丰富的业务信息。比如,银行发现 A 客户信用卡消费类交易占到其过去一年来交易总额的 70% 以上,进一步的分析发现,该客户的卡日均存款余额总是保持在 2 万元左右,即该客户总是习惯于在消费后很快把钱存入卡中,从银行贡献度评价来看,该客户的平均贡献也进入了银行重点营销区——如果这些信息能准确地被业务人员掌握,那么他就可以向 A 客户重点推销“贷记卡”产品,辅加“周期转账”等金融产品,由银行在约定时间(如消费 45 天后或还款期前 3 天等),根据其贷记卡内透支余额,转入现金实现还款,让客户既享受先消费后还款的优惠,又可享受存款利息,满足客户消费和资产增值的组合需求。

从目前的技术发展水平来看,在银行管理客户生命周期的各个阶段都可以用到数据挖掘技术。数据挖掘能够帮助银行确定客户的特点,从而可以为客户提供有针对性的服务;通过数据挖掘,可以发现购买某类金融产品的客户特征,有的放矢的扩大业务;如果找到了流失客户的特征,就可以在具有相似特征的客户流失之前,采取针对性的措施。银行的客户获得、交叉销售(Cross-selling)、客户关怀与保持等方面,数据挖掘工具都将发挥重要作用。

# 信息技术在港口发展中的定位和应用

燕山大学 赫连志巍 秦皇岛港务集团公司 朱朝阳

[摘要] 信息化是港口可持续发展的重要力量,本文从国内港口信息技术发展的现状及存在的问题入手,对信息技术在港口发展中的重要作用进行了阐述,根据港口的特点,对信息技术的应用进行了详细的论述,对一些关键问题提出了解决方案。

[关键词] 信息技术 港口 应用

随着世界经济全球化及区域经济一体化时代的到来,我国经济持续迅猛发展,这为港口的发展带来了前所未有的契机。统计资料表明,2004 年中国港口货物吞吐总量达到 33 亿吨,集装箱总量接近 6000 万标准箱,遥遥领先于世界其它国家及地区。与此同时,我国上海港、青岛港、天津港、大连港等相继投入巨额资金用于码头及相关基础设施建设,这使我国港口的实际规模及竞争能力都得到了相当程度的提高。

当前,港口之间的竞争已从单纯的自然条件的竞争、技术装备的竞争、服务软环境的竞争上升到综合竞争能力的竞争。港口业务也从单纯的码头装卸、仓储等物流操作发展到涵盖贸易、物流、监管、金融支付等各个环节。要提高港口综合竞争能力,把大量、复杂的信息资源组织得有条不紊,保证生产高效运行,这就对物流链的信息网络化水平,特别是作为物流中心节点的港口的信息化程度提出了较高要求。为此,港口企业已经把信息化作为推动和实现企业体制创新、技术创新、管理创新,增强企业核心竞争力的重要手段和必由之路。

## 一、目前国内港口信息技术现状

考察当前国内港口信息化状况,硬件配置均达到了一定程度,但内部处理系统的应用水平、信息集成度高低不一,多以满足自身需要为主要目标,系统的开放程度有一定局限性。对外界而言,各类信息的收发方式暂时还无法摆脱传统的手工录入、纸面传递及一对一的现象,既复杂又浪费资源,同时也无法及时得到回应。现有的信息传递方式整体上造成港口业务的动态信息分布在相对封闭、不同的系统内,同一信息在不同的环节被重复录入并表现为不同的数据格式。相关部门和单位的信息水平发展不平衡,相互间处于割裂状态,业务流程难以互动,无法实现横向的电子数据交换和信息共享。港口信息化建设总体规划工作力度稍显不够,各单位独立建设和发展自身的信应用与管理系统,信息代码、格式、标准不统一,难以与相关单位进行信息交换及资源共享,港口信息分布在各个独立的系统中,增值信息未被充分开发。

## 二、信息技术在港口发展中的定位

信息化是实现企业战略的有力工具,信息化的核心是实现资源共享,目的是为企业赢得客户和市场。通过网络与信息技术,既可以改善现有业务流程,使信息流、资金流、物流等有机结合,创造出新的生产力;同时

一般来看,银行数据挖掘的任务,可以划分成四个层次:数据分析、知识发现、决策支持和金融智能<sup>[2]</sup>。

## 四、数据挖掘的实现

数据挖掘的实现主要有以下 4 个步骤:

首先,要踏踏实实做好基础数据库的建设。一个企业实现数据挖掘的前提和基础是拥有大量、真实的数据积累。基础数据是企业宝贵的财富,没有数据积累,数据挖掘将无用武之地。一般用于挖掘的数据有两种来源:数据仓库或数据库。基于数据仓库的数据挖掘,它有三个显著优势:一是数据挖掘必须要对数据进行抽取、清洗、转换和装载,这个过程比较耗时。二是数据仓库的数据是按主题组织的,这为数据挖掘选择合适的数据源提供了方便。三是数据库不能存放历史数据,因此直接在数据库中挖掘,许多知识无法挖掘出来,如预测型应用;相反,数据仓库能够做到。

其次,企业要有比较明确的挖掘目标。盲目的数据挖掘是很难成功的。

再次,数据挖掘必须由来自不同领域的人员共同参与,包括行业专家、数据管理员、数据分析人员、业务分析人员、数据挖掘专家等。大家需要通力合作,寻找一套适合自己企业的开发方法,并逐步建立起挖掘的模型库。

最后,构建数据挖掘系统。挖掘结果是供决策层决策使用的,因此必须得到最高决策管理层的支持、认可和参与。<sup>[5]</sup>

## 结论

社会信息化的发展,使得各类社会组织中的数据大量存在,如何对数据进行有效的利用是面临的尖锐问题。没有数据时认为,有了数据才能有效的决策,但是数据量膨胀时又出现无从下手、不知所措的局面。数据挖掘是我们从数据到知识,再到决策的正确道路。

## 参考文献

[1]“基于决策支持的银行数据挖掘应用”王广宇

[2]“利用数据挖掘实现电信行业客户流失分析”李军