

# 商务智能的理论和应用研究综述

## Review of Business Intelligence Theory and Practice

朱晓武 (北京大学光华管理学院 100871)

**摘要:**本文从学术界和企业界对商务智能的研究和应用遇到的问题出发,对 1990 年以来的商务智能的研究和应用进行总结,分别对商务智能的概念、组成部分进行讨论,得出一个综合的商务智能理论框架,明确了商务智能理论与应用的发展方向。

**关键词:**商务智能 数据仓库 数据挖掘

### 1 引言

企业无论其以前是通过信件、电话还是传真与客户联系,现在正努力进行数据化、信息化,因为它需要对客户更快的响应,提高客户的服务水平并降低成本。企业资源规划(ERP)、销售终端(POS)、市场调查、供应商、客户、网络、政府部门等都在不断地给企业添加数据,实际上平均每 18 个月企业的数据量就翻一番,图 1 是 Data Warehouse Institute 2004 年调查了 796 个使用数据库的企业用户,得到的关于用户更新其数据的速度分布图。

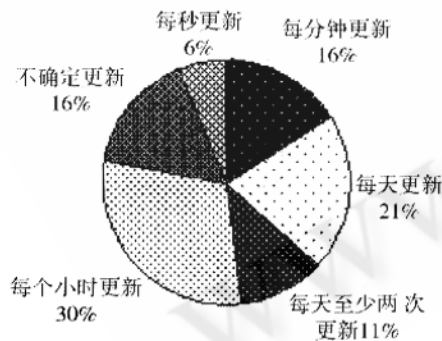


图 1

但是能被分析和利用的数据究竟有多少呢? 有些专家估计目前被利用的数据只有 5% ~ 10%。大量的数据是冗余、不相关、不完整或者质量很差的。怎样才能把大量的数据迅速转换成可靠的信息,发现数据背后隐藏的信息和知识? 因此,出现了商务智能(Business Intelligence, 以下简称 BI)。BI 的核心收益在于它

能够把信息转换为知识,然后把知识转变成利润。

### 2 商务智能理论与应用研究

BI 是当前 IT 领域和企业界关注的热点问题,IT 厂商不停地提出 BI 的解决方案,而企业用户对 BI 的认知度不足 20%,业务人员对 BI 不熟悉的超过了 90%,接近 60% 的企业,对 BI 所能带来的价值以及对企业信息决策系统的支持表示怀疑。BI 是否是 IT 厂商在数据仓库之后炒作的一个卖点,仅仅是一个概念的包装?

产生以上问题的原因主要有以下几个:

第一,BI 是一个新的研究领域,它的理论发展还不够完善,没有明确的界定和理论支持。企业用户迫切的想知道 BI,希望能够把 BI 应用到企业中提高绩效。很多 IT 厂商也对 BI 作了大量宣传,但是他们对 BI 的定义各不相同,有的说是商业绩效管理,有的说是数据挖掘,有的说是仪表盘等,这使企业用户产生了迷惑,到底什么是 BI? 企业应该采用什么样的 BI 解决方案?

Gartner Group 对理想的 BI 和当前的实际情况做了一个比较图,见图 2。

第二,从 BI 的供应商的角度来说,他们由于没有一个完备的理论支持,缺乏统一的标准。在一个企业中存在几个不同的 BI 软件,造成了 BI 的信息孤岛,其后果是在部署、维护和培训上花费多余的开支;最终用户因其商业问题不能被及时解决而感到灰心;不能利用企业 BI 而导致竞争优势的丧失;缺乏统一的标准使得难以整合数据。Karl Van den Bergh, Timo Elliott (2003) 指出,建立 BI 的标准框架是非常必要的。

第三, BI 在理论上的缺陷使得进行 BI 研究的研究者,常常赶到困惑,在国际好的学术刊物上到目前没有以 BI 为题目的论文,只是在一些应用型的产品介绍的刊物中出现 BI 的论文。国内外的研究者对 BI 进行研究大多是局限于对某一项技术如数据挖掘,数据仓库等等的研究,对 BI 的研究如同盲人摸象,缺乏一个整体的规划。

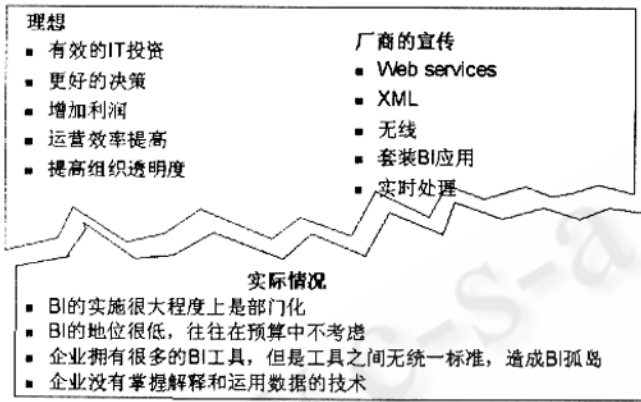


图 2 BI 的理想和实际情况比较

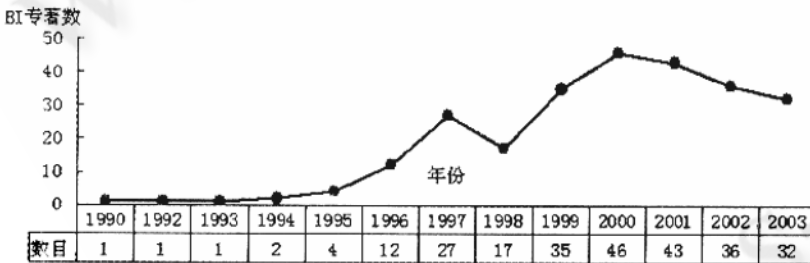


图 3 BI 专著数目趋势图

因此,有必要对 BI 的理论和实践进行研究。本文的主要目的是梳理 BI 的理论研究,消除对于 BI 的误解,避免盲目实施 BI 项目。通过对各个 BI 应用方案的比较,重新对 BI 进行界定,提出一套 BI 的衡量标准。

2.1 研究方法和数据来源

本文查阅了从 1989 年 Gartner 第一次提出 BI,一直到 2005 年的关于 BI 的研究论文和著作,以及 BI 主流厂商的 BI 解决方案的白皮书。

主要的文献来源:中国期刊网, JSTOR, 商业资源电子文献全文数据库 (Business Source Premier); 主要的书籍来源,王茁 (2004) 给出了一个从 1990 到 2003 年的关于 BI 的理论书目 257 本; 互联网搜索,一些关于 BI

的网站以及论坛;关于 BI 的厂商资料来源于厂商的网站,产品说明文档,以及与一些 BI 厂商的交流。

研究方法,主要对 1990 到 2003 年关于 BI 的书籍进行比较分析。一般来说刊物代表了研究的方向,而专著 (书籍) 则是这些研究理论的汇编。然后以英文单词 Business Intelligence 为关键词搜索,或者以中文商务智能 (或者商业智能) 为关键词进行搜索。把得到的商务智能厂商的应用方案和发表的论文进行综合归纳。

2.2 商务智能理论发展

商务智能 1989 年由 Gartner Group 的 Howard Dresner 首次提出,它描述了一系列的概念和方法,通过应用基于事实的支持系统,来辅助商业决策的制定。BI 技术提供帮助企业迅速分析数据的技术和方法,包括收集、管理和分析数据,将这些数据转化为有用的信息,然后分发到企业各处。

BI 的出现经历了一个渐进的、复杂的演进过程,而且仍处于发展之中,它包括事务处理系统 (TPS)、高级管理人员信息系统 (EIS), 管理信息系统 (MIS) 和决策支持系统 (DSS) 等阶段,最终演变成今天的企业 BI。

依据王茁 (2004) 给出的 BI 英文原版数目,得到从 1990 年到 2003 年 BI 的书籍数目的趋势见图 3。

Bernard Lioutaud 指出 BI 这个研究领域有时候会出现一个转折点:技术比理论先行 (通常是理论比技术先行)。1990 年到 1995 年对于 BI 研究主要是基于数据库,数据建模, Edward R. Tufte 于 1990 和 1991 年关于信息的著作,特别是信息定量可视化对 BI 奠定了基础。Inmon 于 1994 年提出数据仓库的概念,1996 年 Inmon 的《Managing the Data Warehouse》是数据仓库的里程碑。1997 年 Inmon 的《Corporate Information Factory》为数据仓库和数据挖掘建立了一个信息工厂的框架。

1998 年以后 BI 的研究开始多起来, Larry Kahaner 提出竞争智能,还有一些学者对数据仓库和数据挖掘技术,数据质量、机器学习等不同方向进行研究。1999 年,众多 IT 厂商纷纷给出相应的 OLAP 产品,如 Microsoft、Oracle、SYBASE、IBM/DB2 的 OLAP、多维数据集研究。2000 年随着知识发现和智能决策以及在不同行业的数据挖掘的深入,对 BI 的研究开始转向战略

性的数据库营销 (Hughes, 2000)。2001 年以后, BI 的研究的主题从技术层面上升到了企业的战略层面, 平衡计分卡、战略竞争分析等。进行 BI 研究的涵盖了企业的各个层面, 相应的研究的主题越来越广泛, 著作也越来越多。

从图 3 的趋势可以看出, 1994 - 1997 研究著作数目的增多是由于数据库技术的不断成熟。1997 年 Inmon 的 CIF (Cooperate Information Factory) 的模型, 整合了企业内部运营资料存储 (ODS), 为 BI 的研究带来了新的理论支持。1998 以后, 研究 BI 的著作数量大幅增多。2003 年有稍稍下降的趋势, 其原因就在于 BI 涵盖的面很广, 容易使得研究者产生困惑, 因此需要新的理论来支持 BI。

目前关于 BI 的定义主要可以分为企业界和学术界两类定义, 这里分别列举了几个比较全面的最新的定义: IBM (企业界) 认为 BI 是一种能力: 通过使用企业的数据库资产来制订更好的商务决策。企业的决策人员以数据仓库为基础, 经过各种查询分析工具、联机分析处理或者是数据挖掘加上决策人员的行业知识, 从数据仓库中获得有利的信息, 进而帮助企业提高利润, 增加生产力和竞争力。

Business Objects (企业界) 公司对 BI 的定义: BI 是一种基于大量信息基础上的提炼和重新整合的过程, 这个过程与知识共享和知识创造密切结合, 完成了从信息到知识的转变, 最终为商家提供网络时代的竞争优势和实实在在的利润。

Kamel Rouibah, Samia Ould - ali (2002, 学术界), BI 是一种系统地依据战略决策, 瞄准、跟踪、传达、转换企业的弱信号, 成为可行信息的一种战略方法。

Olszak (2002, 学术界) BI 是一系列的概念、方法和流程的集合, 其目标不仅仅是帮助决策, 而且支持企业的战略实施。它的主要任务是面向不同信息源的智能浏览、集中、综合以及多维分析。

Sal March (2004, 学术界) 对 BI 的研究按照时间的顺序进行了总结, 他指出传统的 BI 只是被当作一种智能处理后台数据的方式, 用以支持经理主管人员的商业决策。而今天的 BI 帮助企业管理者决策, 预测未来的趋势, 提高企业战略和流程的效率, 从而使得企业获得竞争优势。

以上这些定义从不同的角度对 BI 给出了解释, 但

这些定义都不够全面, 没有从根本上阐述 BI 的本质。

BI 和知识管理之间也容易产生混淆, 因此有必要把这二者之间的关系阐述清楚。Nonaka (1990) 和 Takeuchi (1995) 建立了显性知识和隐性知识的体系, 并提出知识经济时代最大的挑战就是如何有效实现知识的共享。O' Dell 和 Grayson (1998) 认为知识管理是一种有意识的战略, 它把恰当的知识在恰当的时间传递给恰当的人, 并帮助人们把这些知识共享, 把知识融入到提升组织的绩效的行动中。但是无论知识共享的技术如何发展, 实现知识的共享还是存在较大的问题——人本身的问题。事实上这是管理理论中的集体、个体一致性问题。一方面要求知识必须共享, 另一方面对于个人而言, 他们面临巨大的竞争, 所以缺乏知识共享的激励。对于知识共享中出现的两难的尴尬, 有很多解决方案。Hansen 和 Nohira (1999) 提出外部代理的方法, Osterloh 和 Frey (2000) 提出直觉和协作公平的理论。还有很多知识管理学者提出大量的理论方法, 但是却没有一个简单统一的理论让学者和企业界满意。因此需要跳出这个理论的框架, 重新来看这个问题。BI 采用与知识管理不同的另外一种方法来解决知识共享的问题。每一个决策的环节都因为得到了 BI 的支持, 实现双赢或者多赢的结果, 从而解决了知识管理中的问题。

### 2.3 商务智能的实践

目前 BI 产品大致可划分为大型数据库厂商 (如 Oracle、Sybase、IBM 等)、统计软件厂商 (如 SAS) 的衍生产品和独立的软件产品 (如 COGNOS, Hyperion) 两类。前者与数据库技术密切结合, 有强大的后台数据库支持。后者则有些仅有前端展示工具, 没有后端数据库的支持。从技术发展来看, 今后的 BI 产品将把数据仓库建模及数据挖掘等技术实质性地应用进来。同时 BI 技术也将和 CRM、ERP 等技术实现融合, 在企业决策分析中发挥更大的作用, 从而拓展出更大的市场发展空间。

查询、报告和 OLAP 工具占主流通常认为, 能够称得上 BI 产品的, 一定要包含查询、报告和 OLAP 这些基本功能, 否则不能称之为 BI 产品。但是目前的 BI 市场还没有形成统一的标准, 有些数据仓库的供应商也称自己的产品为 BI 产品, 因此市场上就会出现同被称为 BI 的两种产品却大相径庭的怪现象。在由数据仓库、查询、报告、OLAP、分析模型、数据挖掘这些不同的功能

构成的 BI 软件里,查询、报告和 OLAP 工具占据了绝对主流地位为 65%。而最能体现 BI 价值的分析模型和数据挖掘工具目前还只有少数的几家专门做 BI 软件的厂商,如 Business Objects、Hyperion、COGNOS 和 SAS 这样的专门做统计分析软件的厂商能够提供。其它像 Oracle、微软、IBM、Sybase 等这类数据库厂商对 BI 软件功能的侧重则各不相同,因此推出的 BI 产品各具特色,往往是把 BI 打包在其它软件中推广。由于 BI 的分析型应用将在未来占主导地位,而对于每个行业又都需要不同的行业知识,很难找到一个适合各种行业的分析模型来。所以在大的行业市场里,BI 将会与行业系统解决方案融合在一起。

据集市、运营数据存储和 ETL 几个部分,没有揭示在整个信息工厂中的价值链。M. Kathryn Brohman(2000)等人提出 BI 价值链,并对链上的活动进行了实证研究,但是没有从组织、战略等方面的角度来分析。

结合企业界和学术界,我们把对于 BI 的研究整合到一个框架下面。见图 4。

在战略层面,运用平衡计分卡,衡量财务指标和非财务的指标,把战略推向执行。在组织层通过企业绩效管理、企业活动监控、前后台信息部门分工,实现合理的结构分布,达到管理与监控的目的。在功能层,BI 系统必须包括查询、报告、和 OLAP 工具,同时必须设定良好的模型库、知识库、关键绩效指标。在技术层,BI 系

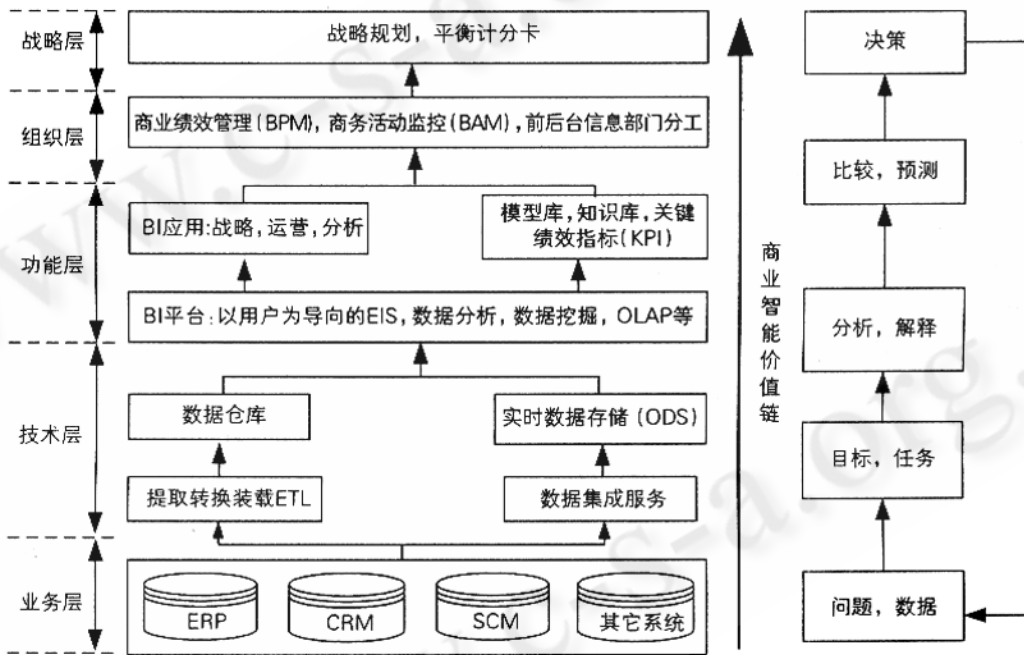


图 4 商务智能理论框架图

### 3 商务智能的理论框架

综合 90 年代到现在的研究者对 BI 的研究,我们可以发现,BI 是一系列的概念、方法和过程的综合,通过这些概念、方法和过程来获取和分析数据,提取有用信息,帮助更好的决策,特别是战略决策。一个真正意义上的 BI 必须跟战略、组织、技术等密切配合,监控企业的关键绩效指标(KPI),包括企业外部环境,顾客,供应商,竞争者等,及时给各层决策者提供智能支持,帮助企业构建更好的盈利模式。

Inmon(1996)提出信息工厂,但是仅仅是限于数

据必须包括 ETL 工具、数据仓库、数据挖掘、数据集成服务、实施数据存储。在业务层面,BI 必须有较好的结构和标准,符合企业各个部门的需求,确保 ERP、CRM、SCM 等系统收集到真实的数据,防止信息的孤岛。

BI 涉及到企业战略、组织、功能、技术、业务五个层面的整体解决方案,通过这五个层面,把企业整合成一个信息工厂。在 BI 的价值链中实现数据到信息、知识、智能、利润的价值增值,从而使企业取得竞争优势。同时在问题和决策之间,有信息的反馈,保证战略决策和执行对环境变化的适应性。

(下转第 54 页)

## 4 结论以及展望

本文对 BI 的理论与应用进行总结分析,整合了企业界和学术界对于 BI 的研究成果,得出了一个 BI 的理论框架,并且对 BI 进行了明确的界定,把战略、组织、功能、技术和业务结合起来,使得学术界和企业界能够在统一的框架下讨论 BI。

但是 BI 面临不断变化的环境,数据仓库和数据挖掘,知识发现一直是 BI 研究的热点问题。数据集成、协同商务、 workflow 管理、BI 战略和移动商务等问题是 BI 最新的研究方向。

### 参考文献

- 1 王茁、顾洁,三位一体的商务智能管理、技术与应用[M],北京:电子工业出版社,2004.
- 2 商务智能(BI)白皮书, <http://www-900.ibm.com/>

- 3 Liataud, Hammond. E - Business Intelligence: Turning Information into Knowledge into Profit. New York: McGraw - Hill, 2002.
- 4 Turban. E., Aronson, and J. E.. Decision Support Systems and Intelligent Systems. Prentice Hall, 1998.
- 5 Karl Van den Bergh, Timo Elliott. Why Standardizing Business Intelligence Is Critical. Business Objects Whitepaper, 2003.
- 6 Steven M. Shaker, Mark P. Gembicki. The War-room Guide To Competitive Intelligence. New York: McGraw - Hill, 1999.
- 7 Kamel Rouibah, Samia Ould - ali. Puzzle: a concept and prototype for linking business intelligence to business strategy. Journal of Strategic Information System 2002, 11: 133 - 152.