

数据资产

数据质量

数据治理

数据清洗

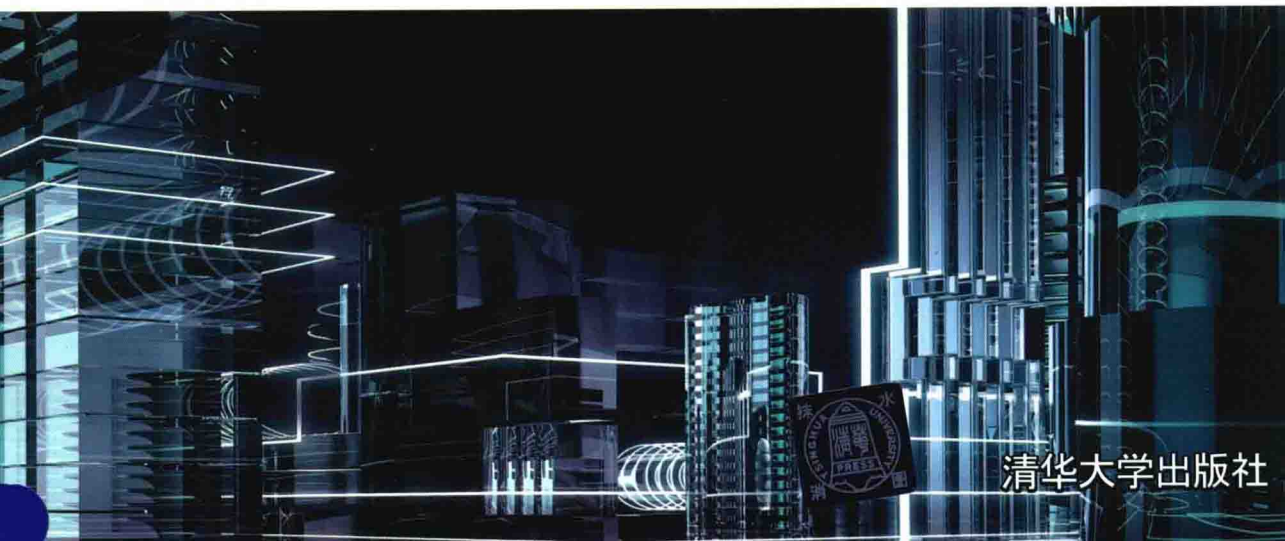
基于**全生命周期**的**主数据管理**

MDM

详解与实践

赵飞 编著

- 大数据应用之**核心推动力**
- 基于全生命周期的**主数据管理成熟度驱动模型**



清华大学出版社

基于**全生命周期**的**主数据管理**

MDM

详解与实践

赵飞 编著

清华大学出版社
北京

内 容 简 介

随着信息技术的发展和信息化的深入,数据逐步成为企业宝贵的战略性资产。主数据管理正是从企业杂乱的数据中捕捉那些具有高业务价值的、被企业内各个业务部门重复使用的关键数据进行管理,构建单一、准确、权威的数据来源,从而提高企业的整体数据质量,提升数据资产价值,推动业务创新,全面增强企业竞争力。本书编者将在近十年的主数据管理咨询中积累的经验 and 知识进行总结,通过对主数据管理的原理、技术、实施、产品、案例、发展趋势等内容的介绍,为读者揭开了主数据管理这一新兴概念的神秘面纱,为读者进行数据治理、主数据管理实践提供重要参考。

全书共分9章。第1章引入数据资产的概念以及数据治理的概念、内容和重点;第2章讨论主数据管理的基本概念,为读者揭示主数据体系的内涵和主数据管理的任务;第3章讨论主数据管理过程中的关键基础技术,系统地介绍数据标准、IT架构、信息系统应用等相关的技术知识;第4章讨论主数据全生命周期管理的概念、模型、方法和技术;第5章讨论为企业进行主数据体系架构设计的步骤和方法;第6章说明主数据管理系统实施阶段的主要任务和流程;第7章介绍目前国内主流的主数据管理解决方案和产品;第8章讨论国内主数据管理的先进案例,并分析国内外主数据管理的现状;第9章展望主数据与云服务、数据挖掘和大数据应用的发展趋势。

本书既可补充从事信息化建设的IT部门人员的专业知识,更能为组织管理者提供信息化知识储备和工作思路,助力组织优化IT架构,也可以作为IT咨询从业者、相关专业在校师生的参考读物。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

基于全生命周期的主数据管理:MDM详解与实践 / 赵飞 编著. --北京:清华大学出版社,2015

ISBN 978-7-302-38563-9

I. ①基… II. ①赵… III. ①企业管理—数据管理 IV. ①F270.7

中国版本图书馆CIP数据核字(2014)第273618号

责任编辑:栾大成

封面设计:杨玉芳

责任编辑:徐俊伟

责任印制:杨 艳

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:北京嘉实印刷有限公司

经 销:全国新华书店

开 本:188mm×260mm 印 张:22 字 数:543千字

版 次:2015年2月第1版 印 次:2015年2月第1次印刷

印 数:1~5000

定 价:69.00元

产品编号:060681-01

编委会

主任委员： 赵 飞

副主任委员：（按姓氏笔画排序）

王 乐

邓忠军

杨富春

咸奎桐

符 刚

曹国钧

委员：

马 玥

王玉荣

王 旭

王兆君

李 明

张明月

曹朝晖

序

“大数据”的概念火了，一时间，无论是社交媒体、电商平台等互联网企业，还是银行、医疗以及工业制造等传统行业，随处可见“大数据”的影子。大数据时代的到来意味着数据已经成为重要的生产要素，将掀起新一波提高生产率的浪潮。

在这样的潮流中，许多企业都希望从海量数据中挖掘新的契机，然而企业要做大数据，首先要了解自身的业务和核心竞争力，以此为基准定位和规范核心数据，然后再进一步延伸到外围相关数据（例如行业数据、社交数据等等），像滚雪球一样建立自己的大数据，从中挖掘出企业发展所需的重要信息。因此，在大数据时代，企业应该更加重视挖掘核心数据资源的价值。

然而现如今，国内许多企业面临的问题恰恰来自于企业内部的核心数据资源。与海量的企业外部数据相比，主数据描述了企业的核心业务实体，承载了企业的核心业务流程，其重要性不言而喻。随着企业 IT 架构的不断复杂化和企业数据的迅速积累，以及公司治理费用、运营风险的日益增加，企业急需更有效和根本的方式管理和维护分散保存于多个异构或同构的系统中的数据。而主数据管理正是从源头入手，为企业提供统一的主数据访问接口，建立全局的、规范的数据视图，从而为业务流程、经营决策以及大数据的集成和管理提供一个可靠的支撑载体。

对于国内许多的企业管理者来说，主数据管理还是一个新鲜的名词。其实，主数据管理在很早之前就已经存在，其致力解决的核心问题也一直伴随着企业信息化的发展，即数据标准化问题。但数据标准化经常被误解为一个单纯的技术问题，而实际上在企业内部贯彻实施统一的数据标准化，需要文化、制度、业务、流程等全方位的配合和改造，需要业务部门之间的沟通和协调，需要企业高层、行业专家、领域专家以及业务人员的认同和参与。也就是说，主数据管理实质上是个管理问题。也正因如此，企业管理者需要积极学习主数据管理的理念、方法和经验，从而指导自身的管理实践。

本书凝聚了编者在数据标准化和数据治理领域多年的从业经验，通过从原理到架构、从实施到案例，深入浅出地介绍了主数据管理这一新兴领域，不仅可以为正在从事企业数据治理工作的 IT 人员提供参考，更能为奋斗在数据时代的企业管理者开启新的思路。企业必须首先提升数据管理能力，确保核心数据的质量和真实性，才能更好地洞悉业务数据和外部数据中所隐藏的价值，真正把大数据转化为大机遇。

傅少川 教授、博士生导师

北京交通大学经济管理学院 工程硕士中心主任

2015年1月18日于北京

前言

随着信息技术的蓬勃发展和企业信息化的深入应用，数据一直都在以惊人的速度增长。美国互联网数据中心的研究表明，互联网上的数据每年将增长 50%，每两年便翻一番，而依据 IDC 的预测，到 2020 年全世界的数据存储总量将达到 40ZB，相当于 40 万亿 GB。

与此同时，大数据时代的来临也为我们带来了更多的烦恼。企业积累的数据越来越多，数据类型变得多样且散落在不同的系统中，零散的且不一致的数据逐渐成为瓶颈。虽然数据已经成为企业宝贵的战略性资产，但如何从海量数据中准确地挖掘出有价值的信息，已经成为企业数据管理和使用的全新挑战。

主数据管理正是从杂乱的数据中捕捉那些具有高业务价值的、被企业内各个业务部门重复使用的关键数据进行管理，通过构建单一、准确、权威的数据来源来建立数据的应用基础，从而提高企业的整体数据质量，提升数据资产价值，推动业务创新，全面增强企业竞争实力。主数据管理并不仅仅是搭建一个信息系统，而是要建立包括主数据标准、主数据组织、主数据质量体系和主数据安全体系在内的完整、持续运行的主数据管控体系，这样才能构建企业数据的核心治理能力，合理利用企业数据来寻求竞争优势，形成数据资产分析和利用的长效机制。

本书编者从事主数据管理咨询和数据治理工作近十年，亲身体会了主数据和数据治理的理念在中国企业信息化浪潮中的兴起、演进和实践。目前，包括能源、金融等多个行业已经将主数据管理和数据治理工作列入到信息技术的“十二五”规划中，企业亟需科学有效的管理模式、技术手段对企业数据进行管理。然而，国内主数据管理的应用仍处于初级阶段，许多企业管理者对主数据管理理念的认识不足。而且国内关于主数据管理相关资料仍十分有限，对主数据体系建设工作的系统介绍更是凤毛麟角。编者将在主数据领域的从业经验和知识积累进行总结并编著出版，希望弥补这一空白，与大家分享和探讨。

本书内容

本书坚持“理论结合实践”的思路，不仅介绍主数据管理的产生背景、概念、模型和技术等理论知识，同时涵盖主数据管理项目的实施方法和过程、主数据管理的产品和应用案例，使读者对主数据管理项目从底层技术知识到上层应用实践都能有系统的理解。同时，本书有针对性地对行业主流厂家的主数据管理产品进行了全面介绍，让读者能够更加深入地了解行业主流产品与趋势。书中案例都是近几年国内相关行业的领先企业的优秀实践，对其他企业



的主数据管理和数据治理工作具有很高的参考价值。另外，对数据挖掘、云计算、大数据等新兴技术与主数据管理的结合应用也进行了探讨及趋势分析。

全书通过对主数据管理的背景、概念、模型、技术、实施、产品、案例、发展等内容的全面介绍，为读者揭开主数据管理这一新兴概念的神秘面纱，为读者进行数据治理、主数据管理实践提供重要参考。

全书共分 9 章。第 1 章引入数据资产的概念以及数据治理的概念、内容和重点；第 2 章讨论主数据管理的基本概念，为读者揭示主数据体系的内涵和主数据管理的任务；第 3 章讨论主数据管理过程中的关键基础技术，系统地介绍了数据标准、IT 架构、信息系统应用等相关的技术知识；第 4 章讨论主数据全生命周期管理的概念、模型、方法和技术；第 5 章讨论为企业进行主数据体系架构设计的步骤和方法；第 6 章说明主数据管理系统实施阶段的主要任务和流程；第 7 章介绍目前国内主流的主数据管理解决方案和产品；第 8 章讨论国内主数据管理的先进案例，并分析了国内外主数据管理的现状；第 9 章展望主数据与云服务、数据挖掘和大数据应用的发展趋势。

读者群

本书既可补充从事信息化建设的 IT 部门人员的专业知识，更能为组织管理者提供信息化知识储备和工作思路，助力组织优化 IT 架构。

本书也面向咨询公司的顾问和实施人员，不仅针对主数据管理项目，对处理各类信息系统项目中可能出现的数据问题都具有一定参考价值。

同时，还可以作为企业管理软件开发人员的自学和参考书，以及作为相关专业在校师生开拓视野、理论联系实践的参考书。

在本书的编写过程中参考和引用了国内外很多书籍和网站的相关内容，部分图片素材和个别实例的初始原型也来源于网络，部分互联网相关资源无法一一列举出处，在此一并予以感谢。最后特别感谢清华大学出版社的栾大成编辑在本书出版过程所给予的大力帮助。

众所周知，一本书难免出现不足和疏漏之处，恳请广大读者将意见和建议反馈给我们，以便在后续版本中不断改进和完善。

目 录

第1章 主数据管理的背景	1
1.1 信息时代的企业发展	2
1.2 数据资产	3
1.2.1 数据资产的概念	3
1.2.2 数据资产的重要性	5
1.2.3 数据资产现状	6
1.3 数据治理	10
1.3.1 数据治理的概念	10
1.3.2 数据治理的意义	11
1.3.3 数据治理的内容	11
1.3.4 数据治理的重点	14
第2章 主数据管理的内涵	17
2.1 主数据的概念	18
2.1.1 主数据的定义	18
2.1.2 主数据的特征	19
2.1.3 主数据的范围	19
2.2 主数据管理的概念	24
2.2.1 主数据管理的定义	24
2.2.2 主数据体系架构	25
2.2.3 主数据管理系统	29
2.3 主数据管理的意义	38
2.3.1 主数据管理的必要性	38
2.3.2 主数据管理的意义	40
第3章 主数据管理的关键技术	43
3.1 数据标准	44
3.1.1 数据标准的分类	44
3.1.2 物资类数据标准	51
3.1.3 产品数据标准	55
3.1.4 财务数据标准	57
3.2 信息分类编码技术	59
3.2.1 信息分类编码的概念与功能	60
3.2.2 信息分类编码原则	62
3.2.3 信息分类编码方法	64
3.2.4 典型编码结构及其标准	67



3.3 企业应用集成	69
3.3.1 企业应用集成的概念	69
3.3.2 企业应用集成的分类	72
3.3.3 企业应用集成的方法	74
3.3.4 企业服务总线	78
3.3.5 主流企业应用集成平台简介	79
3.4 面向服务架构	82
3.4.1 面向服务架构的概念	82
3.4.2 面向服务架构的基本特征	84
3.4.3 服务导向的架构和 Web 服务协议	87
3.4.4 基于面向服务架构的企业应用集成	90
3.5 数据仓库	91
3.5.1 数据仓库的定义及特征	91
3.5.2 数据仓库的发展	94
3.5.3 数据仓库的体系结构	95
3.5.4 数据仓库的数据模型	97
3.5.5 主流数据仓库软件简介	99
第 4 章 主数据全生命周期管理	103
4.1 主数据全生命周期管理的概念、架构和模型	104
4.2 主数据模型管理	106
4.2.1 元数据	106
4.2.2 数据编码管理	108
4.2.3 校验规则	109
4.2.4 workflow 引擎	110
4.2.5 显示规则	111
4.2.6 分发策略	113
4.3 主数据业务管理	113
4.3.1 数据申请	114
4.3.2 数据审核	115
4.3.3 数据变更	117
4.3.4 数据集成和数据分发	117
4.3.5 数据查询	121
4.3.6 数据归档	122
4.4 数据清洗管理	123
4.4.1 数据清洗的内容	124
4.4.2 数据清洗的一般过程	127
4.4.3 数据清洗的工具	128
4.5 主数据质量管理	129

4.5.1	主数据质量管理的目标和任务	129
4.5.2	主数据质量管理评估维度	130
4.5.3	影响数据质量的因素	131
4.6	主数据安全	132
4.6.1	系统应用安全	133
4.6.2	数据内容安全	136
4.6.3	系统日志管理	136
4.6.4	系统版本管理	138
第5章	主数据体系规划方法	139
5.1	主数据体系规划的任务和步骤	140
5.2	主数据体系评估方法论	141
5.2.1	主数据管理成熟度模型	142
5.2.2	主数据管理成熟度模型的评价指标	146
5.2.3	主数据管理成熟度评估方法	148
5.3	现状调研与需求分析	151
5.3.1	现状调研	151
5.3.2	现状评估与差距分析	156
5.3.3	需求分析	157
5.4	主数据识别分析方法	164
5.4.1	多因素分析方法	164
5.4.2	主数据类型识别分析	165
5.4.3	主数据元属性识别分析	166
5.5	主数据体系规划设计	167
5.6	主数据体系架构设计	169
5.6.1	主数据管控体系	171
5.6.2	主数据标准体系	175
5.6.3	主数据质量体系	179
5.6.4	主数据安全体系	180
5.7	主数据管理实施规划	180
第6章	主数据管理体系实施方法	185
6.1	系统实施阶段的任务	186
6.2	数据准备	188
6.2.1	数据准备方案制定	188
6.2.2	数据采集	188
6.2.3	数据清洗	189
6.2.4	数据导入	190
6.3	人员培训	192



6.4	程序设计	193
6.4.1	程序设计的基本要求	193
6.4.2	程序设计方法	194
6.4.3	产品定制开发	195
6.5	系统测试	196
6.6	系统试运行及上线	201
6.6.1	系统试运行	201
6.6.2	系统切换	202
6.7	系统评价	204
6.8	项目管理	207
第7章	典型主数据管理系统平台简介	215
7.1	主数据管理系统模式的分类	216
7.2	主流主数据管理平台产品介绍	218
7.2.1	北京三维天地科技有限公司的主数据全生命周期管理平台	218
7.2.2	SAP 的 MDM 解决方案	250
7.2.3	IBM 的 MDM 解决方案	258
7.2.4	Oracle 的 MDM 解决方案	264
7.2.5	Informatica MDM	268
7.3	主数据管理系统选型	272
第8章	典型主数据管理应用案例	275
8.1	国际先进企业的主数据管理现状	276
8.1.1	法国 VINCI 集团通过主数据集中管理有力促进全球项目一体化运作	276
8.1.2	美洲银行公司通过构建主数据体系实现了对数据的高效管控	277
8.2	国内先进企业的主数据管理现状	278
8.3	石油石化行业应用举例——某特大型石油总公司	279
8.4	煤炭行业应用举例——两个大型能源集团	283
8.4.1	某特大型能源集团公司	283
8.4.2	某大型能源集团	286
8.5	电力行业应用举例——某特大型电力投资集团公司	288
8.6	机械制造行业应用举例——某大型饲料机械集团	291
8.7	水泥行业应用举例——某水泥控股有限公司	294
8.8	建筑行业应用举例——某特大型建筑股份有限公司	298
第9章	主数据管理的应用前景展望	303
9.1	主数据管理应用市场发展趋势	304
9.2	基于云服务的主数据管理	306
9.2.1	云服务的定义和发展现状	306
9.2.2	主数据管理的云服务模式	309

9.2.3	主数据管理云服务平台的技术基础	311
9.2.4	云服务对企业主数据管理的影响	312
9.3	主数据管理在数据挖掘中的应用	314
9.3.1	数据挖掘的定义及企业应用	314
9.3.2	数据挖掘的价值与挑战	318
9.3.3	主数据管理与数据挖掘的关系	320
9.3.4	主数据管理支持的数据挖掘体系构建	322
9.4	大数据时代的主数据管理	324
9.4.1	大数据的定义及特征	324
9.4.2	大数据时代企业管理的新模式	327
9.4.3	主数据管理在大数据分析中的作用	329
9.4.4	大数据对主数据管理的挑战	331
9.5	主数据管理——企业发展的坚实根基	332

第1章

主数据管理的背景



导读

随着农业时代和工业时代的衰落，以信息技术为代表的技术革命从根本上改变着我们的社会经济生活。人类社会形态已由工业社会发展为信息社会，传统的农业经济、工业经济正在被知识经济逐渐取代。在经济全球化的浪潮中，企业面临的竞争环境也发生了巨大变化，企业竞争环境的变化促使企业尽快进行发展模式、管理、商业模式的升级转型。知识经济的逐渐形成，使得信息资源日益成为不可忽视的生产要素和无形资产，在创造价值和社会财富的过程中发挥着越来越重要的作用。为了使庞大的企业数据发挥更大的价值，企业必须着眼于数据治理和综合利用，从信息化方面引领企业的转型升级。而主数据管理作为数据治理的核心环节，是企业充分利用数据资产，在市场竞争中立足的必经途径。

- 信息时代就是信息产生价值的时代。信息化是当今时代发展的大趋势，代表着先进生产力。
- 大数据权威专家维克托·迈尔-舍恩伯格提出“虽然数据还没有被列入企业的资产负债表，但这只是一个时间问题。”
- 数据作为信息时代的重要战略资源，如同一座有待开采、矿藏丰富的矿山，而数据治理则是具体的开采方法和手段。



1.1 信息时代的企业发展

第二次世界大战以来,尤其是近30年来,以信息技术为核心的科技革命推动了世界经济两大趋势的出现,即经济全球化和知识经济。

随着信息技术的迅速发展,经济全球化的浪潮使得企业面临的竞争环境发生了巨大的变化。不同国家和地区之间的商品、资本、劳动力和技术流动日益自由化。由于信息高速公路以及国际互联网的日益普及,世界经济、贸易、金融也日益网络化、全球化,传统的时空界限被打破。物流、信息流、资金流实现了全球的流通,时空在压缩,我们生活的地球已经成为小小的村落。以产品生产为例,许多产品已不再局限在一个国家生产,而依赖各个国家之间的紧密合作,一种产品可能同时分布在十几个、几十个国家和地区生产,这些国家各自发挥其技术、劳动力成本方面的优势,最终生产出的将是国际性的产品。信息化与经济全球化的迅猛发展,推动着全球产业结构不断调整升级,产业分工日益深化,企业面临的竞争环境日趋激烈。

另一方面,知识经济已逐步形成,经济增长方式正在发生着重大转变。知识经济建立在信息技术、应用知识和智能的基础上。知识经济具有两大特征,即网络化的生产关系和数字化的生产力。知识是知识经济的生产支柱和主要产品。新技术的发展造就了一大批新的高科技产业群,旧的工业体系得到改变,从而使经济过程和经济产出中知识和技术的含量不断提高,其贡献率也不断提高。创新成为崭新动力,知识成为未来驱动经济发展的主要力量。著名管理学家彼得·德鲁克(Peter F. Drucker)^①说“在新经济时代,知识不仅仅是与土地、人力及资本这些传统资源并存的另一类资源,而且是当今唯一有意义的资源。”

与此同时,伴随着互联网技术的高速发展,数据正在呈现爆发式的增长,“大数据”的产生正在引发信息技术和商业模式的又一次变革。以互联网流量为例,2001年全世界一年的互联网流量是1EB^②,足以刻满1.68亿张DVD;到2004年,用掉这些流量仅需要一个月;到了2007年只需要一周便可以达到;而到2012年全世界互联网一天流量就是1EB。全世界每年产生的数据量以千倍为单位迅猛增长。从2012年至2014年全世界产生的数据量会超过人类有历史以来到2012年数据量的总和,两年的产出相当于一万年的积累^③。知识更新速度也在不断地加快,根据联合国世界科学情报系统^④统计,20世纪80年代以来,科学知识每年的增长率达到12.5%,同时知识跟进的周期不断缩短,19世纪的知识更新为20年,20世纪缩短为15年,计算机的处理能力每18个月翻一番。

在这个经济全球化和知识经济的时代,企业市场竞争的内容已经由质量的竞争转向了速度的竞争、信息的竞争和知识的竞争,市场竞争的范围也由地区扩展到全球范围。2012年,

① 彼得·德鲁克(Peter F. Drucker, 1909.11.19~2005.11.11), 现代管理学之父, 其著作影响了数代追求创新以及最佳管理实践的学者和企业家们, 各类商业管理课程也都深受彼得·德鲁克思想的影响。

② EB(ExaByte), 艾字节, 计算机存储容量单位。1艾字节=1,152,921,504,606,846,976字节(B)。1字节表示8个二进制位。

③ 参见中国工程院副院长邬贺铨, 2013年7月21日第二届中国大数据应用论坛, 大数据时代的发展趋势演讲。

④ 世界科学情报系统(World Science Information System, UNISIST)是联合国教育科学文化组织的一个所属机构。该机构于1972年在联合国教育科学文化组织的第17届大会上批准建立, 旨在促进科学技术情报在世界范围内的广泛传播。

全球技术研究和咨询公司 Gartner 公司^①发布的调查报告^②中显示,在新环境下,增加企业增长、吸引并留住新客户、降低企业成本、创造新的产品和服务、交付运营结果、提高效率等业务问题,和商业智能、移动技术、云计算、协同技术等技术问题得到了企业管理者的广泛关注。可见,面对知识经济大潮的冲击,企业在不断变化的环境中必须依靠先进的 IT 技术,深度挖掘信息资源,占领信息化的制高点,即时调整自身发展战略,才能保持在新的局面下持续发展的动力。

从企业角度而言,由于在新的经济环境下企业需要在战略与发展方面参与全球化和国内跨区域的各种竞争,促使企业在管理与运作方面不断推出和适应新的模式与理念。企业战略逐步向全球化、精益化、服务化、智能化和协同化的趋势发展。精细化的运营掌控战略和全球化的战略管控大大依赖于企业开发和利用信息资源的能力。企业信息化与新的商业运作模式相互融合,推动企业的管理变革,是企业面对新环境挑战的必要对策。面对信息时代日趋激烈的市场竞争环境,企业如果不重视信息化战略,不加强对企业信息技术和数据应用的投资,必将丧失从海量数据包围的社会竞争中突围的历史性机遇。

企业信息化是指企业利用现代信息技术,通过对信息资源的深度开发和广泛利用,不断提高生产、经营、管理、决策的效率和水平,从而提高企业经济效益和提升企业核心竞争力的过程。企业信息化是一个宽泛的概念,是对企业的集成化、系统化、规范化、创新性管理思想的体现。企业信息化是一个持续的过程,在复杂多变的内外部环境,企业必须针对商业模式演变、技术创新等机遇和挑战,对信息化建设工作的方向和目标不断提出新的、全面的要求。

信息化的实质是通过先进业务流程、管理理念与信息技术的融合,改变企业收集、处理、利用信息的方式,从而导致业务流程和组织形式的巨大变革。信息化的关键在于如何获得更优质的信息资源和如何从信息资源中挖掘更大的商业价值。数据作为信息资源的基本表现形式和重要组成部分,对数据的投资能更大化地实现企业的业务价值。所以,体现业务价值的企业信息化蓝图,必然需要以企业数据为核心。如果将企业的信息系统比喻成一辆承载企业发展战略和业务模式的列车,那么数据就是驱动列车前进的燃料,燃料的质量直接决定着列车的运行速度。因此,数据的治理应该是企业信息化的一项核心基础性工作,其目标是保证企业数据的有效性、可访问性、高质量、一致性、可审计性和安全性。

1.2 数据资产

1.2.1 数据资产的概念

数据是一种未经加工的原始资料,是对客观事物的逻辑归纳,用符号、字母等方式对客

^① Gartner 公司,高德纳公司,又译顾能公司,1979年创立于美国,是全球最具权威的IT研究与顾问咨询公司。

^② <http://www.gartner.com/newsroom/id/1897514>。



观事物进行直观描述。数据是进行各种统计、计算、科学研究或技术设计等所依据的数值（是反映客观事物属性的数值），是表达知识的字符的集合。信息是数据内涵的意义，是数据的内容和解释。数据经过解释并赋予一定的意义之后，便成为信息。

企业所应用的信息系统和数据仓库中存储了大量生产经营活动中的基础数据和业务数据，例如客户数据、BOM 清单^①数据、订单数据等等，通过 ERP^②、CRM^③等管理信息系统的加工处理。这些数据将转化为信息，用以支持企业各层级的管理决策，提高现有生产资源的利用率，充分发挥资源的整合效应，促进提高管理效率、降低经营费用、提高劳动生产率、提高客户满意度等，从而间接为企业带来经济效益。

信息资产是由企业拥有或者控制的能够为企业带来未来经济利益的信息资源。其本质是将信息作为一种经济资源参与企业的经济活动，减少和消除了企业经济活动中的风险，为企业的管理控制和科学决策提供合理依据，并预期给企业带来经济利益。国际标准 ISO/IEC17799《信息安全管理标准》^④中明确指出“信息是一种资产，像其他重要的业务资产一样，对组织具有价值，因此需要妥善保护”。根据 BS7799 以及国标 GB/T20984—2007《信息安全风险评估规范》^⑤，信息资产可分为数据、软件、硬件、服务、人员和其他类别等。由于数据是信息在企业中的主要表达形式，数据资产也成为信息资产中最重要的组成部分。

数据的直接效用是将企业的各项生产经营活动客观形象地记录下来，实现可计量、可存储、可复用的管理目标。数据的间接效用体现在：一是通过参与市场竞争，提高服务水平和营销能力来增加收入；二是通过改进业务流程或提高分析效率来降低运营、人力等各项成本；三是以真实完整的信息助力企业管理者的科学管理与决策。有效的管理和使用数据可以减少或消除企业经济活动中的风险，为企业管理控制和科学决策提供合理依据，给企业带来相关的经济效益。

《大数据时代》的作者维克托·迈尔-舍恩伯格（Viktor Mayer-Schönberger）^⑥指出：“在亚当·斯密论述 18 世纪劳动分工时所引用的著名的大头针制造案例中，监督员需要时刻看管所有工人、进行测量并用羽毛笔在厚纸上记下产出数据，而且测量时间在当时也较难把握，因为可靠的时钟尚未普及。技术环境的限制使古典经济学家在经济构成的认识上像是戴了一副墨镜，而他们却没有意识到这一点，就像鱼不知道自己是湿的一样。因此，当他们在考虑生产要素（土地、劳动力和资本）时，信息的作用严重地缺失了。”由此可知，当时数据获

-
- ① BOM（Bill of Material），物料清单，指产品所需要的零部件的清单及组成结构，即生产一件产品所需的子零件及其产品中零件数量的完全组合。
 - ② ERP（Enterprise Resource Planning），企业资源计划，是指建立在信息技术基础上，以系统化的管理思想，为企业决策层及员工提供决策运行手段的管理平台。
 - ③ CRM（Customer Relationship Management），客户关系管理。
 - ④ ISO/IEC17799《信息安全管理标准》，由国际标准化组织 ISO 于 2000 年 12 月正式发布，是参照英国国家标准 BS7799 而来的，该标准为信息安全管理提供建议，旨在为一个机构提供用来制定安全标准、实施有效的安全管理时的通用要素，并使得跨机构的交易实现互信。
 - ⑤ 《信息安全风险评估规范》GB/T 20984—2007，由国际标准信息技术安全性评估通用准则（CC）改进而来。标准提出了风险评估的基本概念、要素关系、分析原理、实施流程和评估方法，以及风险评估在信息系统生命周期不同阶段的实施要点和工作形式。
 - ⑥ 维克托·迈尔-舍恩伯格（Viktor Mayer-Schönberger），牛津大学教授、大数据权威专家，被誉为“大数据商业应用第一人”，其著作《大数据时代：生活、工作、思维的大变革》（Big Data: A Revolution That Will Transform How We Live, Work, and Think）是国外大数据研究的先河之作。

取、存储和应用的成本过高是信息资产被忽视的重要原因。但随着技术的不断发展,数据的存储密度大大提高,大量企业内部信息通过信息系统被便捷地记录和使用。管理者意识到数据所创造的价值已远远大于其成本。

在企业信息化水平不断提高、业务数字化程度日益加深的今天,几乎所有机构都卷入到数据及其处理(数据收集、存储、检索、传输、分析和表示)的浪潮中,数据已成为重要生产要素和无形资产。2012年年初的达沃斯世界经济论坛^①上,一份题为《大数据,大影响》(Big Data, Big Impact)的报告宣称,数据已经成为一种新的经济资产类别,就像货币或黄金一样。

数据资产作为信息资产的主要组成部分,具有以下基本特征。

- 共享性: 一个个体对数据的使用不会影响另一个体对其使用,数据的传播并不是数据的转移,而是数据的复制,因此,数据资产具有共享性。但由于企业数据作为一种重要资产,其共享性只体现于企业内部部门、合作伙伴或员工之间。
- 增值性: 众所周知,物理资产会在使用的过程中发生损耗,造成价值的降低,但是,随着时间的推移,数据量的积累反而会使得数据资产的价值增加,这就是数据资产的增值性。数据资产作为信息的载体,其传播和使用的过程也正是其扩张和创新的过程,结合先进的管理思想进行有效的组合、分析和挖掘,将产生更多有价值的信息。
- 时效性: 数据资产的特征来源于数据和信息的属性。信息具有时效性,同样的,数据资产也具有时效性。对于一些流动性强的信息而言,如果不能及时开发利用,其价值就会大打折扣,例如市场类信息。
- 低安全性: 一般而言,数据资产经常处于公共的介质或者处于流动状态,数据资产的复制成本较低,从而导致企业拥有和控制数据资产的安全性很差,这正是导致信息资产风险的一个重要因素。

1.2.2 数据资产的重要性

早在春秋战国时期,《鬼谷子》^②符言第十二中便提出:“目贵明,耳贵聪,心贵智。以天下之目视者,则无不见;以天下之耳听者,则无不闻;以天下之心思虑者,则无不知”。其中便反映了在进行了管理决策时,需要收集更多的数据以辅助决策的管理思想,但是在信息封闭的古代,这仅仅是一个梦想。在信息时代,数据就像农耕时代的土地、工业时代的石油,是其他价值创造过程的源泉。数据资产不仅具有一般物质资产的特征,而且兼有无形资产

^① 世界经济论坛(World Economic Forum, WEF)是一个非官方的国际组织,总部设在瑞士日内瓦。其前身是现任论坛主席、日内瓦商学院教授克劳斯·施瓦布1971年创建的“欧洲管理论世界经济论坛非洲峰会”。1987年,“欧洲管理论坛”更名为“世界经济论坛”。论坛因每年年会都在达沃斯召开,故也被称为“达沃斯论坛”。每年的世界经济论坛年会均有来自数十个国家的千余位政界、企业界和新闻机构的领袖人物参加。世界经济论坛已经成为世界政要、企业界人士以及民间和社会团体领导人研讨世界经济问题最重要的非官方聚会和进行私人会晤、商务谈判的场所之一。

^② 《鬼谷子》,又名《捭阖策》,是战国纵横家鼻祖鬼谷子王诩的著作。据传是由学者根据鬼谷子先生言论整理而成。该书侧重于权谋策略及言谈辩论技巧,是纵横家的代表著作。