

DATA

DESIGN

数据价值与产品化设计

Designing Data Products

李满海 辛向阳 著

 机械工业出版社
CHINA MACHINE PRESS

不要被这本书名吓住，这是一本有故事、有案例、有学术的好书，数据在这里不再是冰冷的，是有温度、有生命的！无论是对在校学生，还是职场相关行业的同仁，这本书值得一读和收藏。

——周晓农，益普索中国总裁&数据科学研究院首任院长

数据是21世纪的资源，算力是21世纪的生产力，本书带领产品化设计进入21世纪。

——范凌，特赞信息科技创始人，同济大学设计人工智能实验室主任&博导

作者针对数据资源和数据产品的各方面属性进行了剖析，提出了具有启发性的新概念、新观点、新问题。对于解决数据定价、数据产品设计等实践中的紧要问题，这本书是个很好的开始。

——陈为，浙江大学计算机学院CAD&CG国家重点实验室教授&博导

在ABC（AI，Bigdata，Clouding）时代，数据成为核心资源。李满海博士和辛向阳教授为数据产品研究做出了开创性和奠基性的贡献。强烈推荐给相关科研工作者和业内人士研读！

——徐毅，中国科学技术大学管理学院国际项目负责人、管理学博士

李满海博士以他独特的研发、产品、用户体验和大数据的复合背景，结合近20年的学术研究和商业实战经验，围绕如何将数据实现产品化价值创造为大家带来这本著作，值得一读。

——黄峰，唐硕体验咨询机构创始人&CEO，UXPA中国共同创始人



机工教育微信服务号

ISBN 978-7-111-64364-7

策划编辑◎冯春生 / 封面设计◎李满海



定价：32.00元

数据价值与产品化设计

李满海 辛向阳 著



机械工业出版社

本书从设计学角度系统地整理了数据用于决策的设计工具和方法，阐述了从数据作为商品视角去应用设计思维，提出了多级的数据产品理念以及产品化设计路径，同时列举了大量的案例，增强了本书的实用性。

本书内容主要包括数据经济时代、数据作为产品、数据加工模式、数据的价值感知、数据产品化要点、数据升级加工和数据产品化设计。

本书可供从事与数据相关的社会实践者，包括管理者、工程师、设计师，以及从事与数据相关的学术研究者，包括哲学、设计学、艺术学等人员阅读；同时也可供设计类专业的高校师生参考。

图书在版编目 (CIP) 数据

数据价值与产品化设计 / 李满海, 辛向阳著.

—北京: 机械工业出版社, 2020. 2

ISBN 978-7-111-64364-7

I. ①数… II. ①李… ②辛… III. ①数据处理—产品设计
IV. ①TP274

中国版本图书馆 CIP 数据核字 (2019) 第 290441 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 冯春生

责任编辑: 冯春生

责任校对: 宋道兰 杜雨霏

责任印制: 张博

三河市宏达印刷有限公司印刷

2020 年 5 月第 1 版 · 第 1 次印刷

184mm × 260mm · 9.75 印张 · 151 千字

标准书号: ISBN 978-7-111-64364-7

定价: 32.00 元

电话服务

客服电话: 010-88361066

010-88379833

010-68326294

封底无防伪标均为盗版

网络服务

机工官网: www.cmpbook.com

机工官博: weibo.com/cmp1952

金书网: www.golden-book.com

机工教育服务网: www.cmpedu.com

前 言

从2004年到2018年，我一直在中兴通讯从事产品研发工作，期间研究院的名字换了几回，但是工作内容基本没变，始终围绕着“云计算”和“大数据”去设计和开发几十种业务产品。2005年我申请了“利用智能网实现语音识别自动拨号”的技术发明专利，那年科大讯飞研究院才刚刚成立。2011年我所在的团队研发的语音信箱业务已经服务全球超过1亿人，那年微信才上线。

在通信行业，技术发展可谓日新月异。仔细研究发现：所有技术都在为“数据产品化”服务；所谓的数字化系统实质是数据的加工场所；诸如5G技术、人工智能技术、虚拟现实技术都只是手段，数据才是根本！我们需要搞清楚数据是什么以及数据产品化的内在逻辑，才能发现深层次的创新机会。

为什么写这本书

网上书店可检索出近八万本与数据相关的书籍，大致可以分为两大类：一类是介绍与数据相关的专业技术，如大数据挖掘、分布式存储和数据可视化等；另一类是阐述数据在各行各业的实践应用。这些书籍在不断强调数据拥有的巨大价值，但是以下这些不可回避的现实问题却没有被回答：

问题一：从理论上说，拥有稀缺数据源的企业将拥有持久的竞争优势，然而实际上，许多企业在数据投资上耗费了大量的人力财力，也取得了广泛的关注与专业性的口碑，但是却没有获得预期的回报。那么是否因为数据存在价值层次而导致数据的潜在价值没有被充分地挖掘出来？如果是，那么数据价值层次之间是怎样的递进关系？

问题二：市场上存在各种各样的数据应用平台，向用户提供多样化和个性化的

数据服务或产品。同样的数据经过不同方式的经营，产生的价值差异很大，为什么有的数据价值能够实现指数级增长？而为什么有的数据价值表现一般？

问题三：同样的数据加工品带给人们感受到的实际价值会因人而异，也会因使用场景而产生差异。在数据产品化前期进行价值机会点分析时，设计师需要考虑哪些影响因素？在数据产品化过程中通过哪些举措可以提升用户对数据价值的感知？

为了深究这些商业现象背后的问题根源，我决定攻读博士。在学习期间，我查阅了近500篇国内外学术论文，其中90%是国际刊物，包括最近十几年《自然》和《科学》等顶级杂志关于数据相关的研究进展。当然也少不了到一线调研，先后专程到美国硅谷和英国伦敦参访了脸书（Facebook）、谷歌（Google）、爱彼迎（Airbnb）、英国广播公司（BBC）等企业，走访了国内北上广深等9个城市，访谈了阿里巴巴、金蝶软件、今日头条等36位与数据产品相关的代表人物，有的是公司的首席运营官，有的是该领域的首席科学家，有的是部门的设计主管。这些调研材料经过哲学抽象后，诞生了我的博士论文《基于价值维度的数据产品化设计研究》。这是一篇从设计学角度论述数据产品化的理念和方法的学术研究成果，希望这些成果不仅能够解答企业实践中出现的现象问题，而且能够揭晓现象背后的规律，进而指导企业实践。

在博士导师辛向阳教授的鼓励和支持下，我们决定将博士研究过程积累的几十个案例打散重构，从读者方便理解的角度将核心成果表达出来，希望学术成果能够指导企业实践，产生出商业价值。

本书的主要内容

本书定位是一本偏学术的专著，通过理论探讨和案例验证，力争让每个章节都能有自己的观点，依次如下：

第1章，数字经济时代。本书从数据正在经历“离线”到“在线”、“静止”到“流动”、“封闭”到“开放”的过渡过程开始讲起，阐述在数据作为资源的大趋势下，数据产品化的市场前景值得期待，然后对本书所要描述的“数据”范畴做界定。

第2章，数据作为产品。这是本书的核心章节，提出数据产品化的相关观点，

不仅对“数据产品”概念和特性进行定义，而且将数据产品划分为零级、一级和二级三个级别，同时整理出不同级别数据产品的基本特征，并列举案例进行说明。

第3章，数据加工模式。数据产生价值的过程可以抽象成单环加工和双环加工模式；数据加工可以分为“点、线、面”三个层次；人们可以通过从“自动化、实时化、模板化”三个方式来提升数据产生价值的效能，实现从手工生产到智能生产范式的加工升级。

第4章，数据的价值感知。数据与人们经历之间的关系可以分为三种紧密度，不同紧密度会影响人们对数据价值的判定；人们对数据价值的认知方式有两种：被动感知和主动体验；数据的意义比功能更能影响价值感知；数据隐私和安全问题也是一个影响因素。

第5章，数据产品化要点。本章总结数据加工过程的共同注意要点和潜在问题。作为数据产品的生产者，有四个方面的加工要点需要关注，包括数据产品的实效性、数据产品的利益点、数据产品价值共创，以及数据与体验的关系。

第6章，数据升级加工。本章以数据的价值层次、价值效能和价值感知作为三个维度构建出数据产品化价值矩阵，提出移向右上角可以实现数据价值最大化；提出数据价值的三个侧重、数据加工品所对应的三个形态特征，以及不同级别之间的递进逻辑。

第7章，数据产品化设计。本章提出数据产品的机会分析可以从社会、技术、经济和意义四个方面去考量；同时提出数据产品化设计流程可以分为三个大步骤，然后通过一个完整数据产品的设计案例来系统阐述数据产品化设计的全过程。

本书的读者对象

数据在自然科学和人文社会科学研究中发挥着越来越重要的作用。在数据经济时代的语境下，阿里芝麻信用、联通智慧足迹，以及新零售等数据驱动的新兴商业生态逐步成熟时，数据价值所服务的产业环境、产品属性、设计流程都和传统产品大相径庭。各位读者从本书提供的案例中也可以体会到数据价值带来的社会变革、设计进步和行业转型等价值。

本书尝试从理念上突破该认知局限，启发读者跳出“数据是辅助工具”或“数

据是决策依据”等传统认知，尝试以“数据是设计对象”来全新看待已有的数据。

尤其推荐下面两类读者阅读本书：

1. 从事与数据相关的社会实践者，包括具有数据采集和处理能力的各类传感器厂商的工程师，拥有数据处理和应用能力的无人驾驶汽车、虚拟现实、人工智能等新技术领域的工程师和分析员，以及拥有稀有数据而不知如何产生更大商业价值的企业主管。

2. 从事与数据相关的学术研究者，包括工业产品设计专业和交互体验设计专业的学者，计算机与软件应用专业的学者，以及大数据技术与应用专业的学者。

李满海

2019年9月

前 言

第 1 章 数据经济时代

Chapter
One

1.1 数据发展趋势 / 002

1.1.1 数据在线 / 002

1.1.2 数据流动 / 004

1.1.3 数据开放 / 006

1.2 数据作为资源 / 007

1.2.1 变革行业应用 / 007

1.2.2 市场规模庞大 / 010

1.3 数据范畴界定 / 013

1.3.1 数据历史由来概述 / 013

1.3.2 本书给数据下定义 / 014

参考文献 / 016

第 2 章 数据作为产品

Chapter
Two

2.1 数据产品的定义 / 020

2.1.1 什么是数据产品 / 020

2.1.2 数据产品的特性 / 021

2.1.3 数据产品的载体 / 023

2.1.4 价值产生的原理 / 024

2.1.5 数据产品的级别 / 024

2.2 零级数据产品 / 026

2.2.1 数据本身就是产品 / 026

2.2.2 数据的可视化处理 / 028

2.2.3 数据的艺术化加工 / 033

2.3 一级数据产品 / 035

2.3.1 作为行为决策工具 / 035

2.3.2 作为信息服务范式 / 039

2.4 二级数据产品 / 042

2.4.1 普适性数据加工品 / 042

2.4.2 重新定义行业规则 / 044

参考文献 / 046

第 3 章 数据加工模式

Chapter
Three

3.1 数据加工的通用模式 / 050

3.1.1 单环加工模式 / 050

3.1.2 双环加工模式 / 051

3.2 数据加工的三个层次 / 052

3.2.1 点层次数据加工 / 054

3.2.2 线层次数据加工 / 055

3.2.3 面层次数据加工 / 056

3.3 提升加工效能的方式 / 058

3.3.1 自动化：端到端数据流管道 / 058

3.3.2 实时化：随时随地在线处理 / 059

3.3.3 模板化：从需求到解决方案 / 060

3.4 数据产品的生产范式 / 062

3.4.1 手工生产范式 / 063

3.4.2 智能生产范式 / 064

参考文献 / 067

第 4 章 数据的价值感知

Chapter
Four

4.1 数据与人的经历 / 070

4.1.1 两者没有关联 / 071

4.1.2 相互交叉影响 / 071

4.1.3 构成经历核心 / 072

4.2 用户被动地感知 / 073

4.2.1 数据的客观质量 / 073

4.2.2 数据的主观质量 / 074

4.2.3 数据解读成信息 / 076

4.3 用户主动去体验 / 076

4.3.1 意义层面的价值 / 077

4.3.2 意义比功能重要 / 078

4.3.3 主动体验的模型 / 079

4.4 数据隐私和安全 / 080

4.4.1 个人的隐私数据 / 080

4.4.2 敏感数据的安全 / 081

4.4.3 数据的四个类别 / 081

参考文献 / 084

第 5 章 数据产品化要点

Chapter
Five

5.1 数据产品的实效性 / 086

5.1.1 用户关心的是问题 / 086

5.1.2 数据不是越多越好 / 087

5.1.3 数据清洗是体力活 / 087

5.2 数据产品的利益点 / 088

5.2.1 数据利益相关者 / 088

5.2.2 数据产品的收费 / 089

5.2.3 数据加工的伦理 / 090

5.2.4 数据隐私的应对 / 091

5.3 数据产品价值共创 / 092

5.3.1 用户生成内容 / 092

5.3.2 多方共同创造 / 093

5.3.3 使能数据平台 / 094

5.3.4 价值共创案例 / 096

5.4 数据与体验的关系 / 098

5.4.1 数据提升用户体验 / 098

5.4.2 数据助力体验设计 / 099

5.4.3 体验作为设计准则 / 100

5.4.4 体验作为设计对象 / 101

参考文献 / 103

第 6 章 数据升级加工

Chapter
Six

6.1 数据价值移向右上角 / 106

6.2 数据价值的三个侧重 / 107

6.2.1 还原和表达问题 / 108

6.2.2 发现和解决问题 / 110

6.2.3 重新再定义问题 / 111

6.3 数据产品的三个形态 / 112

6.3.1 初原自然的形态 / 113

6.3.2 具象实效的形态 / 115

6.3.3 通用普适的形态 / 116

6.4 三个级别的递进升级 / 116

6.4.1 高低包含的关系 / 116

6.4.2 逐级进化的关系 / 118

参考文献 / 121

第 7 章 数据产品化设计

Chapter
Seven

7.1 数据产品的机会分析 / 124

7.1.1 社会因素 / 125

7.1.2 技术因素 / 126

7.1.3 经济因素 / 127

7.1.4 意义因素 / 127

7.2 数据产品化设计流程 / 128

7.2.1 识别数据产品化机会 / 129

7.2.2 选择实现的价值层次 / 131

7.2.3 决定价值的效能程度 / 135

参考文献 / 140

鸣谢 / 141

Chapter One

第1章

数据经济时代

世界经济论坛 (World Economic Forum) 宣称, 数据作为新的经济资产类别, 正在成为一种与能源和材料相提并论的资源, 已经关系到制造企业未来的生存和发展^[1]。《人民日报》在官方网站上也在宣传和推广: 数据将成为继土地、能源之后的又一战略性资源, 将成为最重要的生产资料^[2]。当数据被看作像石油一样的生产原料时, 可以说我们身处于“数据 (Data) 经济时代”。

1.1 数据发展趋势

数据正在经历从“离线”到“在线”，从“静止”到“流动”，从“封闭”到“开放”的发展过程。随着人工智能、虚拟现实、物联网等技术的不断进步，该发展演进速度将加快，影响广度也将扩大。

1.1.1 数据在线

阿里巴巴技术委员会主席王坚在他的《在线》书中提到：数据在线，是信息社会和数据社会的本质区别。我们已经从过去企业内部的封闭信息化，走向了互联网上的在线开放。在线，让数据的价值体现变得容易，可以快速影响社会。数据价值变得越来越大，主要原因在于数据在线使得数据流动的时间和空间变大了^[3]。

1. 从数字化到数据化

计算机技术的早期应用是将物理实体转换成二进制码的数字世界，该过程称为数字化过程。例如，1986年王选教授发明的方正汉字激光照排系统解决了中文纸质书的数字化问题，使延续上百年的中国传统出版印刷行业得到了彻底改变。该系统

不仅可以将打印内容数字化，而且通过数字化接口可以操控印刷设备。

数字化不等于数据化，数字化是数据化的基础。例如，1971年就有人提出古登堡计划（Project Gutenberg），倡议将纸质书籍通过计算机技术制成电子书，方便更多人阅读，由于当时限于技术和资金没有做起来。2004年谷歌（Google）开始大规模对所有版权条例允许的纸质书内容进行数字化，每一页都被扫描存成高分辨率的数字图像。在完成纸质图书的数字化后，谷歌继续使用数字图像识别软件提取出页面的每个字、词、段落，将文字从“数字化”图像转化成“数据化”内容，方便人们按关键词检索和分析。其中很重要的一个应用就是提升了谷歌翻译软件的精准度，因为翻译软件需要对单个字或词进行数据处理，而不是对整个数字图像进行处理。

谷歌通过“数字化”和“数据化”两个步骤将纸质图书的内容变成了可以被灵活加工的数据原料。快速发展的互联网（Internet）和物联网（IoT）技术正在将越来越多的内容变成类似的数据原料。这些数据原料构成了数字经济的基础，恰如从1859年美国第一次成功钻井采油开始，那些原油就构成了现代石油工业的基础。

2. 在线成为社会常态

根据中国互联网络信息中心发布的第43次《中国互联网络发展状况统计报告》来看，2018年中国的网民规模数量已发展到8.29亿人，普及率达59.6%，其中手机网民规模约为8.17亿人，网民通过手机接入互联网的比例高达98.6%。随着互联网覆盖范围进一步扩大，居民入网门槛进一步降低，信息交流效率得到提升。特别是今后5G技术应用的普及，上网的概念会完全被淡化，联网才是常态，各种终端设备永远保持连接，产生各种各样的数据。

互联网技术是第三次工业革命^[4]的基础，如今快速发展的物联网技术加快了数据从“离线”到“在线”的跨越。根据阿里研究院（AliResearch）2017年发布的《数字经济2.0》报告，预测未来十年超过一万亿个传感器将接入网络，超过10%的衣服、鞋和眼镜等都能接入物联网，首款植入人体式手机也即将上市。

数据在线，不仅可以让数据产生最大化的经济价值，而且可以给企业带来持久的竞争优势。例如，谷歌地球（Google Earth）是一款谷歌公司开发的虚拟地球软件，它把卫星照片、航空照相和地理位置数据集成显示在一个虚拟地球的三维模型上，

用户只需下载客户端软件，就可以在线浏览世界各地的高清晰度卫星图片和最新的地图。谷歌在 2005 年向全球推出该产品之后，用户量迅速增大，快速赢得市场，彻底颠覆了传统的离线电子地图的盈利模式。

如今，在线已成为了社会的常态。根据腾讯网公布的 2014 全年财务报告，QQ 月活跃账户数达 8.15 亿，微信月活跃账户数达 5 亿，手机淘宝日活跃用户已超 8000 万，新浪微博日活跃用户为 7600 万。

1.1.2 数据流动

百度前总裁、现任 YC 中国创始人陆奇在 2019 年一次演讲中提到：数据一定要是活的，活的数据才有价值^[5]。脸书（Facebook）、谷歌、腾讯、阿里巴巴和京东等互联网企业通过即时消息、社交网络、电子商务等在线服务方式加速了数据流动。

1. 数据交往行为

早期的通信技术主要是实现人与人之间的远程联络，如今的通信技术实现了人与人之间的在线协作，包括远程视频会议、实时文档协作、在线网络游戏等。例如，淘宝（Taobao）和京东将不同商家的商品信息产生关联，为人们提供在线交易服务；优步（Uber）和滴滴出行将用户信息、车辆位置和实时交通等数据在线共享，为人们提供便捷的出行服务；优兔（YouTube）为人们提供实时上传和分享视频的服务。社交网络犹如人类的神经网络，牵一发而动全身。数据像红细胞一样，是基础的媒介和载体，流到哪个行业，就把信息传递到哪里。

人们通过互联网开展的数据交往变得越来越频繁。2012 年脸书平台每天更新照片的数量超过 1000 万张，每天产生 30 亿次的单击“喜欢”（Like）按钮或者评论。推特（Twitter）每天新增 4 亿条微博数，每个月至少有 1.4 亿活跃用户的阅读和微博转发^[6]。这些庞大用户量的背后是密集的数据流动。任何人的数据都可能在极短时间内像地震波一样，从一个圈子传播到下一个圈子，扩散到世界各地。

数据交往行为改变了人们的生活方式，倒逼着企业的变革。以金融行业的数字化转型为例，目前中国国内开通手机银行的用户数已超过 15 亿，网上转账、在线购