

数据中台：让数据用起来

作者：付登坡

数据中台：让数据用起来

付登坡 等著

ISBN: 978-7-111-64240-4

本书纸版由机械工业出版社于2020年出版，电子版由华章分社（北京华章图文信息有限公司，北京奥维博世图书发行有限公司）全球范围内制作与发行。

版权所有，侵权必究

客服热线：+ 86-10-68995265

客服信箱：service@bbbvip.com 官方网址：www.hzmedia.com.cn 新浪微博 @华章数媒

目录

[赞誉](#)

[作者简介](#)

[前言](#)

[第1章 数据中台：信息化的下一站](#)

- [1.1 数据中台产生的大背景](#)
- [1.2 数据中台的3个核心认知](#)
- [1.3 数据中台的3个发展阶段](#)
- [1.4 开启信息化的下一站](#)

[第2章 什么是数据中台](#)

- [2.1 解码数据中台](#)
- [2.2 数据中台必备的4个核心能力](#)
- [2.3 数据中台需要厘清的2个概念](#)
- [2.4 数据中台VS现有信息架构](#)
- [2.5 数据中台的业务价值与技术价值](#)

[第3章 数据中台建设与架构](#)

- [3.1 持续让数据用起来的价值框架](#)
- [3.2 数据中台建设方法论](#)
- [3.3 数据中台架构](#)
- [3.4 中台手记（一）：我说服老板立项了](#)

[第4章 数据中台建设的评估与选择](#)

- [4.1 企业数据应用的成熟度评估](#)
- [4.2 企业数据中台建设的应用场景](#)
- [4.3 中台手记（二）：打仗前手里得有一张“粮草”清单](#)

[第5章 数据汇聚联通：打破企业数据孤岛](#)

- [5.1 数据采集、汇聚的方法和工具](#)
- [5.2 数据交换产品](#)
- [5.3 数据存储的选择](#)

[第6章 数据开发：数据价值提炼工厂](#)

- [6.1 数据计算能力的4种类型](#)
- [6.2 离线开发](#)
- [6.3 实时开发](#)
- [6.4 算法开发](#)
- [6.5 中台手记（三）：选一个适合自己的技术平台真的很重要](#)

[第7章 数据体系建设](#)

- [7.1 数据体系规划](#)
- [7.2 贴源数据层建设——全域数据统一存储](#)
- [7.3 统一数仓层建设——标准化的数据底座](#)
- [7.4 标签数据层建设——数据价值魅力所在](#)

[7.5 应用数据层建设——灵活支撑业务需求](#)

[7.6 中台手记（四）：即将开启的数据淘金之旅](#)

第8章 数据资产管理

[8.1 数据资产的定义和3个特征](#)

[8.2 数据资产管理现状和挑战](#)

[8.3 数据资产管理的4个目标](#)

[8.4 数据资产管理在数据中台架构中的位置](#)

[8.5 数据治理](#)

[8.6 数据资产管理与数据治理的关系](#)

[8.7 数据资产管理职能](#)

[8.8 数据资产管理效果评估](#)

[8.9 数据资产管理的7个成功要素](#)

[8.10 中台手记（五）：家里的这点家底可得管好了](#)

第9章 数据服务体系建设

[9.1 补全数据应用的最后“一公里”](#)

[9.2 4种常见的数据服务](#)

[9.3 3种常见的数据应用](#)

[9.4 数据服务背后的产品技术](#)

[9.5 中台手记（六）：解决“数据应用最后一公里”问题](#)

第10章 数据中台运营机制

[10.1 数据中台运营效果评估模型](#)

[10.2 数据中台运营的4个价值切入点](#)

[10.3 数据资产运营](#)

[10.4 数据成本运营](#)

[10.5 数据中台运营的实践经验](#)

[10.6 数据中台运营的要素与口诀](#)

[10.7 中台手记（七）：让数据用起来](#)

第11章 数据安全治理

[11.1 数据安全面临的挑战](#)

[11.2 贯穿数据全生命周期的数据安全治理体系](#)

[11.3 大数据平台安全管理技术手段](#)

[11.4 中台手记（八）：数据安全！数据安全！数据安全！](#)

附录 6大行业解决方案架构图

赞誉

这本书的大部分作者曾是阿里数据中台部门的实干者，“干过”是一种很宝贵的实力。今天，他们正在用这套数据中台方法论赋能各行各业的数字化转型，并取得了卓越的业务效果。撰写此书是他们践行“让数据用起来”这一使命的又一尝试，将为各行各业的数字化转型的探索者和参与者提供体系化的指导。

——谢世煌 阿里巴巴集团联合创始人/

湖畔山南资本创始人

未来的企业家都应该关注数据的应用，但不是哪个企业都适合建设数据中台。要了解数据中台就从这本书开始吧，因为作者们都是经验丰富的一线实践者。

——邱昌恒（卜鹰） 鲲鹏基金创始合伙人/

原阿里集团副总裁

大数据的本质是数据的融合，把原本各自孤立的数据互相关联、融合，通过抽象、加工构建数据资产标签类目体系，从而赋予数据更深层次的语义和价值，洞察事物的本质。大数据背景下的数据突出了数据的人本特性、数据的充分利用，不仅极大地发展了生产力，同时还将深刻地改变生产关系。

本书理论研究和实践应用并重，对于破解当前的信息化发展难题具有非常重要的现实意义。本书会为所有想要了解数据中台丰富内涵的读者，以及从事数据治理的技术和管理人员带来深刻的启发，指导他们的实践和创新。

——周傲英 华东师范大学副校长/数据学院教授/博导

数据中台是企业数字化转型的战略选择，是数字化时代对企业的组织重构、流程再造与技术升级。数据中台并非一个简单的平台，而是对海量数据进行采集、存储、计算、加工与融合的产物。数据中台旨在消除数据标准和口径不一致的问题，实现面向应用的数据共享。数据中台让数据与业务分离，实现面向客户需求的弹性化试错与快速迭代，为客户提供高效服务。

——金小刚 浙江大学计算机科学与技术学院研究员

借助“互联网+”和“智能+”，人类社会正以前所未有的速度向智能时代前进。在这个进程中，数据成为宝贵的资产，数据资产管理与数据服务能力将决定政府管理与社会服务的效率，也影响企业竞争的胜败。如何应对挑战、占尽先机，本书将带给你别具新意的解读。

——陈强 上海工程技术大学计算中心教授

当各行各业的数据累积到一定规模时，数据存储、管理、挖掘、应用等新技术就能帮助我们“把握现在，预知未来”。数据中台就是这样一套让数据持续用起来的机制，能够助力企业的数字化转型。

——魏凯 中国信通院云计算与大数据研究所副所长

因为工作的原因，我经常与国内外众多行业的管理者、数字化的实践者和同行们交流，其中数据中台这个话题谈论得非常多。很多人在深度地研究和实践，不仅产出了很多原创的思考和文字，而且有着越来越多的实践案例。作为数据中台领域的布道者，我相信数据中台将有助于推动中国的数字化转型，同时，我也相信这本书将会给数据中台领域的实践者带来有意义的启发和思考。

——史凯 ThoughtWorks中国区数据智能事业部总经理

数据是数字经济的“石油”，但数据的价值不仅在数据量，更在于如何让数据有效释放能量。数据分析能力是让“石油”精炼进而充分“燃烧”的技术，“数据中台”更是当前支持数据分析能力提升的技术热点。这本书系统、全面地讲述了“数据中台”的构建方法论，阐述了从IT到DT的发展方向，从技术实现到体系建设，可谓面面俱到，充分体现了“数据中台”的业务价值和技术价值，为企业开展“数据中台”建设提供了很好的规划蓝本。架构重在实践、贵在落地，相信本书能够为业内人士提供很好的指引。

——付晓岩 建信金融风险合规法律部副总经理/

《企业级业务架构设计》作者

随着DT时代的到来，如何让内部数据助力企业数字化建设已经成为各公司必须面对的问题，而最近大火的数据中台正好迎合了这一历史发展趋势。本书的作者们结合自己的中台建设经验，为读者讲述了数据中台的概念、架构、开发、建设、管理、服务、运营、安全等内容，深入浅出，娓娓道来。

——韩锋 宜信数据库开发与与管理主任工程师

作者简介

付登坡（花名：天湛）

资深大数据专家，数澜科技联合创始人&地产事业部总经理

有10余年大数据领域从业经验，擅长数据建模、海量数据产品架构设计与实现。原阿里巴巴集团大数据专家，曾在阿里巴巴集团负责消费者数据标签体系、DMP平台等大数据项目设计与实施。2015年以创始人身份组建阿里巴巴集团“11维数据创新工作室”，探索数据创新与数据商业化。

2016年6月离职，联合创办数澜科技，在数澜科技先后负责技术部、咨询服务部和地产事业部。

江敏（花名：江敏）

资深大数据专家，数澜科技联合创始人&CTO

有10年大数据平台规划、数据安全交换使用、数据应用场景建设方面的实践经验。曾任职于阿里数据平台事业部、阿里云数据事业部，负责阿里数据能力及平台的行业客户赋能，并打造行业的数据共享交换，是ID-Mapping体系能力构建及服务化的核心参与者、数据交易模式的早期探索者。

数澜科技联合创始人，负责管理公司产品技术团队，为客户输出构建和经营数据中台的能力。基于数据中台建设的实践经验，带领团队打造一站式数据应用基础设施数澜，并完成多家行业龙头客户基于数澜的数据中台建设。

任寅姿（花名：影姿）

资深数据产品专家，数澜科技创新事业部总经理

曾任阿里巴巴数据产品专家、数据创新梧桐工作室负责人等。对大数据资产设计、资产服务、资产应用在实践中形成了一套完整的数据标签类目体系方法论，擅长对各种复杂业务场景进行需求拆解、数据抽象和数据应用建模，关注采用大数据方法切实解决场景痛点，提升业务效率。

孙少忆（花名：守正）

资深数字化转型咨询专家，数澜科技战略副总裁

20年企业信息化工作经验，积累了丰富的信息化内部运营、解决方案销售及交付等方面的实践经验。拥有MBCI、CISSP-ISSMP、CGEIT、COBIT 5、ITIL Expert、P30等国际专业资质证书。曾任职华为ICT规划咨询部，面向企业、政府提供“以数据为核心，聚焦业务场景和价值”的流程信息化与数字化转型规划和落地咨询业务。

武凯（花名：行竹）

资深数据产品专家，数澜科技COO

有10余年数据产品经验，曾任阿里巴巴集团数据平台产品与运营部负责人，是营销、零售和医疗健康等领域数据应用的探索实践者，专注于企业大数据资产化及应用增值。

沈金（花名：铁平）

资深数据业务架构专家，数澜科技解决方案总监

10余年数据行业经验，擅长业务架构、数据架构、技术架构的规划和落地实施。曾在阿里巴巴担任DBA，后参与阿里数据中台建设，拥有用户识别、标签设计、动态数据组织等多个发明专利。

2017年加入数澜科技，负责带领解决方案团队，推动数据中台在零售、地产、金融等行业以及综合性集团的落地。

蔣珍波（花名：乐天）

大数据咨询专家，数澜科技高级咨询专家

15年信息化和大数据行业从业经验，具备广阔的知识面、丰富的咨询经验，擅长为客户提供创造性的解决方案，尤其擅长数据治理方面的咨询规划和产品设计，服务过数十家政府和大中型企业客户。

前言

“一切业务数据化，一切数据业务化”，回顾几十年的中国企业信息化发展历程，就是“业务数据化”的过程——企业持续在IT方面进行投入和建设，不断将发展过程中业务和经营管理端的各种能力以数据形态沉淀下来。而接下来的“数据业务化”则是将已经成为资产的数据作为生产资料融入业务价值的创造过程，使之持续产生价值。

但是随着DT时代的来临，一路高歌、突飞猛进的企业信息化建设却开始出现诸多发展瓶颈和痛点。

首先，随着信息化的深入，在传统烟囱式IT建设方式下，企业独立采购或者自建的各种企业信息系统，在内部形成诸多数据孤岛；而在互联网、移动互联网背景下，服务号、小程序、O2O平台等新模式下产生的外部数据与传统系统的内部数据无法互通，这进一步加剧了数据孤岛问题……系统多样性和多态性，增加了企业IT架构的复杂度。

其次，随着多云环境的出现，硬件基础设施从IT时代的服务器演变成DT时代的“云”，多数企业将选取多云策略，以避免被单一云厂商锁定，且多云的使用可以让企业IT架构更灵活、更符合自身情况。但这也让企业IT架构变得更为复杂，底层数据的互联互通成为困扰企业发展的痛点之一。

因此在传统企业底层IT架构下，新旧IT系统中沉淀的数据之间难以打通，而在多云环境中，企业内外部数据亦难以快速连接。分散各处难以融合的数据，无法很好地支撑企业经营决策，也无法很好地应对快速变化的前端业务，如何突破发展瓶颈，构建适应新时代的企业IT架构，以阿里、华为等为首的国内顶级公司开始提出“数据中台”的概念。

在笔者看来，数据中台是企业数字化转型的必然产物。在企业IT架构日益复杂的今天，亟须通过一套机制，联通传统IT架构和各类数据，融合新老模式，整合孤岛数据，沉淀数据资产，快速形成数据服务能力，为企业经营决策、精细化运营提供支撑，这套机制就是数据中台。

DT时代数据中台的使命是“让数据持续用起来”，它的一个根本性创新就是把“数据资产”作为一个基础要素独立出来，将成为资产的数据作为生产资料融入业务价值创造过程，提供推动企业发展源源不断的生产力。

笔者们认为，数据中台作为整个企业各个业务所需数据服务的提供方，通过自身的平台能力和业务对数据的不断滋养（业务数据化），会形成一套高效可靠的数据资产体系和数据服务能力（数据资产化和资产服务化）。这样当出现新的市场变化，需要构建新的前台应用时，数据中台可以迅速提供数据服务（服务业务化），从而敏捷地响应企业的创新。业务产生数据，数据服务业务，业务在阳，数据在阴，阴阳互补，形成闭环。

值得一提的是，数据中台不仅仅是一种技术平台，倘若仅停留于此，就完全忽略了从IT到DT的本质变化是“围绕数据资产进行价值的持续积累和释放”。单纯增大技术投入和人才投入无法保障企业经营效能的持续提升，只有站在数据价值观和方法论的高度，才可能系统性解决企业经营发展中关于数据的诸多问题。谁能率先解决面向数字经济特征的全新数据价值观和方法论的问题，并在其指引下打造出平台级能力，谁就能真正意义上帮助企业把数据用起来。

因此，本书重点落笔于数据中台建设方法论体系的阐述，这也是笔者们多年大数据领域从业经验以及多个数据中台建设经验总结所得。希望这套数据中台建设方法论能为计划进行数字化转型，或已经在数字化转型之路上奋力前行的企业决策者、业务推动者和执行者，提供认知升级的有益借鉴，帮助企业结合自身特点，在战略规划牵引下，从组织、保障、准则、内容、步骤等五个层面全面考虑，建立起一套可持续运行的中台建设机制，以保障数据中台建设实施如期完成，从而加速企业的数字化转型进程。

2019年是数据中台爆发的元年，笔者们认为，数据中台必将依循从概念引爆到迭代试错，再到规模复制的认知升级路径，从行业头部企业普惠至更多中小企业，成为数据应用的“基础设施”。未来，每个企业都会像20年前上ERP、5年前上云那样，标配自己的“数据中台”，为企业数字化转型奠定坚实基础。而作为国内最早的一批数据服务创新者，笔者们希望能将过去十余年沉淀的数据认知和行业经验，与更为广泛的行业人士分享、碰撞、交流，并在此过程中实现认知升级，从而更好地增援企业未来，提升企业硬实力，这就是从2019年5月起，笔者们历时半年撰写本书的初心与梦想。

本书的写作历程是一次共创与感恩之旅。数澜三年，与百家行业头部客户共建数据中台，本书虽只有薄薄数百页，但却承载了数澜与客户合力共创的千余个日日夜夜的思考与探索。感恩在与中信集团、万科地产、兴业银行、百果园等诸多客户共创中沉淀的数据认知与实战经验，让我们得以不断思考“让数据用起来”的真谛与本质，而唯有将这些思想的火花具象成文，集结成册，分享给更多数字化转型之路上同行的人们，方不负大家对我们的期望和厚爱！

这还是一次自我超越之旅，非常感谢机械工业出版社华章公司的策划编辑杨福川老师、责任编辑罗天亮老师、美术编辑王建敏老师，他们的鼓励与指导点燃了我们心中的梦想，让我们这群拙于表达的技术人，竟产出了近20万字的中台理论知识与实践经验总结。半年前看似不可能完成的任务，我们坚持了下来，并沉淀了多年工作认知和经验思考，于我们自身来说是一份特别珍贵的精神财富。

这更是集体智慧的结晶，感谢赵东辉、谢辉、白松、蔡勇剑、黄飞、黄耐寒、贾松、罗玲、李舵、李远泽、秦文艺、孙昂、陶胜刚、张梦琴等多位数据专家，他们奋战在数据产品、开发、算法、咨询、治理一线，在异常繁忙的年中冲刺季，挤出时间为本书撰文，可谓字字珠玑；感谢徐锦、谭琼、邵雪、杨莎莎、朱清漪、程国强、李德亮、陶丽婷等组成的内部编委会，十年经验，四月烧脑总结写作，两周封闭萃取，杭京深三地联动，通宵达旦，激情燃烧，让专业的技术术语鲜活起来，让抽象的思维认知灵动起来，让本书逻辑更清晰明了，更具有可读性与吸引力！

数据中台已经掀起了幕布的一角，幕布后面的精彩世界需要政府、产业、行业、领先企业共同激荡演绎。让我们走进数据中台的世界，共同描绘属于我们的新纪元！

第1章 数据中台：信息化的下一站

经过几年的沉淀和酝酿，数据中台已经成为新的风口。大家好奇它的过去，希望一窥它的全貌，憧憬它的未来，“让数据用起来”是驱动数据中台发展的原动力和大家为之奋斗的目标。本书的重点在于揭示数据中台全貌并分享实战经验，本章则尝试从多个视角梳理数据中台产生的大背景，希望与读者一同感受气势磅礴的时代大潮交汇在数据中台这个点的始末，并抛砖引玉，提炼数据中台应该成为共识的几个认知，展望数据中台应该有的发展阶段，让读者自己形成对数据中台的构想，做好准备与我们共同开启数据中台之旅。

1.1 数据中台产生的大背景

1.1.1 新的时代浪潮

时光机回到公元2019年9月，中国IT界正在掀起一股新的汹涌大浪，“中台”这股技术之浪正在席卷IT界的每个角落，并经由IT工厂人员和各界媒体传导至各个行业。数据中台在DT时代的大背景下尤为引人注目，一些先知先觉的企业在讨论和探索数字化转型，谈论有关“数据中台”的概念，有人认为这是新一波的厂商向甲方企业收“智商税”的概念泡沫，有人认为这是给予CIO影响力的权柄，也有人认为这是企业应对“危”与“机”的快速创新利器。利用数据进行创新，看似机遇近在眼前，但各种问题又充斥在每个相关人士的心中。

“什么是数据中台？”

“数据中台有什么用？”

“什么样的企业适合建数据中台？”

“应该怎样保证数据中台的效果？”

“怎么才能说服我企业的董事长和CEO同意立项数据中台？”

“如何才能保证企业数据不被滥用，又能被开发出价值？”

“在我的行业里，有没有可以参考的成功的数据中台案例？”

“需要多长时间和多少钱才能建起来？建完之后，如何运营才能创造价值呢？有没有持续优化成本的办法？”

风起于青萍之末，浪成于微澜之间。那么数据中台之浪，又成于哪一朵微澜之间呢？

本书的作者们都是在数据领域里摸爬滚打十余年的老数据人，我们将与各位读者一起穿透时光的迷雾，去追溯数据中台之源，回顾25年的IT到DT的演进史，畅谈“数据创新与数据中台”的实践思路，用不太出彩甚至有点干巴的技术文笔，梳理“让数据用起来”的数据中台机制，回答“数据中台百问”，为在数据创新之路上的读者们贡献一点力量。

1.1.2 从IT到DT，中国信息化演进之路

1996年8月，爱特信息技术有限公司成立；1998年2月，爱特信推出搜狐；2000年7月，搜狐公司正式在美国纳斯达克挂牌上市。

1997年6月，网易公司成立；2000年6月，网易公司正式在美国纳斯达克挂牌上市。

1998年12月，新浪公司成立；2000年4月，新浪公司成功在美国纳斯达克挂牌上市。

而在1998年至2000年间陆续成立的腾讯、阿里、百度在十年之后接过中国互联网的大旗，形成了以BAT为首的新一轮互联网化浪潮。

2015年3月5日，国务院总理李克强在政府工作报告^[1]中提出：“制定‘互联网+’行动计划，推动移动互联网、云计算、大数据、物联网等与现代制造业结合，促进电子商务、工业互联网和互联网金融健康发展，引导互联网企业拓展国际市场。”

2019年3月5日，国务院总理李克强在政府工作报告^[2]中提出：“推动传统产业改造提升。围绕推动制造业高质量发展，强化工业基础和技术创新能力，促进先进制造业和现代服务业融合发展，加快建设制造强国。打造工业互联网平台，拓展‘智能+’，为制造业转型升级赋能……促进新兴产业加快发展。深化大数据、人工智能等研发应用，培育新一代信息技术、高端装备、生物医药、新能源汽车、新材料等新兴产业集群，壮大数字经济。”

从1995年到2015年的20年间，互联网科技改变了众多面向个体用户端“2C”的生产关系，通过构建线上平台，方便了人们的衣食住行，丰富了人们的生活体验。而在同样的时间里，在企业内部有一群为了企业生产力而奋斗的技术人士，他们利用IT技术提升企业内部的生产力。2014年，马云先生正式提出“DT（Data Technology）”的概念，“人类正从IT时代走向DT时代”。他认为，IT时代是以自我控制、自我管理为主，而DT时代是以服务大众、激发生产力为主。这两者之间看起来似乎是技术的差异，但实际上是思想观念层面的差异。

同样也是在这一年年初，阿里内部的数据平台事业部正在大刀阔斧地建立整个集团的数据资产，笔者们也很有幸深度参与其中，构建了多笔数据资产，此为题外话。

2015年，“互联网+”行动计划的提出，让企业内部IT与企业外部互联网思维产生火花，云和SaaS形态的应用开始出现，从IT到DT正式有了广泛的落地实践。

1995年到2015年，互联网科技在中国从萌芽到提出“互联网+”行动计划，用了20年时间。从2015年“互联网+”提出到2019年“智能+”提出，用时仅4年。我们惊叹于这演进的速度之快，就像一个常常被提及的例子：“从整个地球史来看，人类科技进化速度的陡峭曲线，可以类比成一个猿人扔起一根骨头，等掉下来的时候骨头已经变成了火箭。”

1.1.3 外太空与黑土地，阿里与华为对中国数字化进程的贡献

1.外太空视角——阿里以数据为核心，推动数字产业化

回溯“数据中台”这个在中国被创新和实践落地的产物，就不得不去看它顶端的企业——阿里巴巴。

笔者认为阿里的贡献巨大，主要有两个原因。

第一，基于内部海量数据应用的数据中台实践经验，以及对新零售、新金融等互联网技术和思维为核心的数据赋能业务的创新尝试，唤醒行业全面跟进和尝试“中台”理念。

2014年阿里从芬兰Supercell公司接触到中台概念后，在集团内部积极践行，开创了“大中台、小前台”的组织机制和业务机制，通过高效、统一的后方系统来支持前端的机动部队，提高作战效率，减少冗余投入。2018年，中台概念开始逐渐深入互联网企业。

·2018年9月，腾讯宣布新成立云与智慧产业事业群（CSIG）和技术委员会，后者将负责打造技术中台。

·2018年11月，阿里云事业群升级为阿里云与智能事业群，并开始对外输出中台能力。

·2018年11月，美团被曝正在打通大众点评、摩拜等各业务间的数据，构建数据中台。

·2018年12月，百度调整组织架构，高级副总裁王海峰同时负责基础技术体系（TG）和AI技术平台体系（AIG）。此后，王海峰曾在公开场合表示，打造技术中台是百度调整组织架构的战略方向之一。

·2018年12月，京东进行了有史以来最大幅度的组织架构调整，增设中台部门。京东商城CEO徐雷还在企业年会上强调：要将中台提升为“永不停歇”的超级引擎。

第二，持续对政府、对社会拓宽基于数据的宏伟认知，并积极实践基于数据创新的城市大脑。

比如，持续提倡：“未来，数据将会是生产资料，计算是生产力，互联网是生产关系，智能时代是基于这些改变而随之发生的巨大的社会变革。未来30年，智能技术将深入到社会的方方面面，改变传统制造业，改变服务业，改变教育、医疗，所有的生活会因数据、计算而改变。IT让20%的人受益，而DT时代和AI时代的数据技术会让80%的人受益，这就是这个世界未来巨大的机会所在。并进一步落地为城市大脑、ET（Evolutionary Technology）系列。”

2016年10月，云栖大会上城市大脑首次亮相。

2017年3月，云栖大会深圳峰会推出ET工业大脑和ET医疗大脑。

2017年6月，云栖大会上海峰会推出ET环境大脑。

就像ET命名本身的科幻性隐喻，阿里所倡导的数据理念和实践犹如外太空的智慧传递，广泛而又深刻地影响着我们这个社会。

2.黑土地的尝试——华为扎根产业，引领产业数字化转型

华为是推动中国信息化进程的另一个时代巨人，华为在CT（通信技术）领域成长为国际知名企业，进而转入IT（信息技术）领域和移动终端领域，都取得了举世瞩目的成绩。华为自身独特的流程信息化实践已成为传统行业高效利用信息化、加速数字化转型的问计首选。

笔者认为，华为对于产业数字化转型的巨大贡献在于其基于一套创造性的“管理架构、流程与IT支撑”的管理体系，在研发、生产制造、供应交付、销售、财经等领域进行持续不断地数据管理和数据应用实践，沉淀了完整且成熟的数据管理相关流程、管控机制、组织建设、人才保障、平台建设、使能服务等体系和机制。在多业务、超大规模和全球化分布式管理的环境下做到了“以数据使能业务”，帮助华为公司在人员不显著增加的情况下，收入、利润、现金流持续有效增长。

2016年华为提出了数字化转型战略，当时希望通过数字化变革重塑华为的商业流程。公司由此提出“+互联网”的概念，希望利用先进的数字、数据技术，改造华为业务流程，致力于率先实现ROADS（实时、按需订阅、在线、自助、社交）体验并成为行业的标杆。

2018年华为建立了数字化转型实践中心（DIPC）并正式投入运营，与运营商一起开展数字化转型的实践探索和能力构建。提出统一的数字化平台必须具备以下特征：**充分协同并融入业务流程，统一数据模型并可平滑交换数据，云原生和能力开放，以及智能化。**基于该数字化平台，与内外部生态协作创新，可以快速提供最佳解决方案以满足客户需求，实现商业敏捷。

对于数据中台，虽然华为公司内部流程的IT架构中没有明确的提法，但是华为在对外提供的解决方案中已不断出现“数据中台”字眼，致力于打造信息化、自动化、智能化“黑土地”的华为势必融入数据中台的洪流之中。通过以下几则有代表的信息，我们可以从侧面了解华为的数据中台。

华为云官网：“（华为智慧）园区中台包含数据中台和业务中台；数据中台对园区数据进行标准化建模，园区各单体子系统数据上传到云端后经过数据治理并存储在数据中台的主题库中……”

江苏移动：“江苏公司秉承集团2019年工作重点‘用户满意度领先’和‘降本增效’两大方针，在提升用户满意度工作方面，携手华为整合SEQ平台、MR地理化平台以及性能平台等多个大数据平台，初步实现江苏移动满意度提升数字化中台战略，通过大数据研究、分析及建模，江苏移动与华为联合研发，全国首次将全网用户业务感知地理化，抢

占‘用户满意度领先’的网络高地。”^[3]

国网天津信通公司：“本次验证工作中，天津公司作为国网公司试点单位，负责开展对华为公司数据中台和云平台相关产品的技术验证测试，从而客观论证华为产品全面支撑天津公司泛在电力物联网建设的可行性。测试工作基于华为产品构建的数据中台和云平台，重点针对功能、性能、开放性、扩展性、兼容性，以及对业务应用的支撑能力开展验证工作。”^[4]

1.1.4 中国政府的支持与引导，为数据中台生长提供阳光雨露

1.两会政府工作报告

自2015年起，“互联网+”在政府工作报告中经常被提及。政务在国家层面指引和推动中国经济模式的升级和迭代。既扶植了数字产业化，又推动传统行业和产业数字化转型。

2015年制定“互联网+”行动计划。推动移动互联网、云计算、大数据、物联网等与现代制造业结合，促进电子商务、工业互联网和互联网金融健康发展，引导互联网企业拓展国际市场。

2016年落实“互联网+”行动计划。增强经济发展新动力，大力推行“互联网+政务服务”，实现部门间数据共享，发挥大众创业、万众创新和“互联网+”集众智、汇众力的乘数效应。

2017年扩充“互联网+”模式。深入推进“互联网+”行动计划和国家大数据战略，推动“互联网+”深入发展、促进数字经济加快成长，让企业广泛受益、群众普遍受惠。

2018年完善“互联网+”模式。“互联网+”广泛融入各行各业，加强新一代人工智能研发应用，在医疗、养老、教育、文化、体育等多领域推进“互联网+”，深入推进“互联网+农业”，多渠道增加农民收入，促进农村第一、第二、第三产业融合发展。

2019年全面推进“互联网+”。运用新技术和新模式改造传统产业，推行信用监管和“互联网+监管”改革，优化环保、消防、税务、市场监管等执法方式，加快在各行各业、各领域推进“互联网+”。

2.数博会

中国国际大数据产业博览会（简称“数博会”），2015年在贵阳创办，2017年正式升级为国家级展会活动。作为全球首个大数据主题博览会，凭借国际化、专业化、市场化的领先优势，数博会成为全球大数据发展的风向标和业界最具国际性和权威性的成果交流平台。

历届数博会均受到国家领导人的关怀和指示。中共中央总书记、国家主席、中央军委主席习近平向首届和近两届数博会连续发来贺信。国务院总理李克强、副总理马凯、全国人大常委会副委员长王晨先后出席前五届数博会开幕式并致辞。其中，2019年数博会共吸引448家国内外企业参展、举办专业论坛53场，专业赛事6场、各类活动近百场，来自50多个国家和地区的政要、知名企业家、专家学者、协会组织、科研机构及媒体相继参加各类活动，共话大数据前沿热点，共绘大数据发展蓝图，共享大数据时代发展新机遇。

3.数字中国建设峰会

“数字中国建设峰会”2018年在福州举办第一届，中共中央总书记、国家主席、中央军委主席习近平发来贺信，向峰会的召开表示衷心的祝贺，向出席会议的各界人士表示热烈的欢迎。

2019年第二届峰会在福州举行，为期3天的峰会共对接数字经济项目587项，总投资额4569亿元，其中签约项目308项，总投资额2520亿元。

本届峰会的主题是“以信息化培育新动能用新动能推动新发展以新发展创造新辉煌”。峰会定位为中国信息化发展政策发布平台、电子政务和数字经济发展成果展示平台、数字中国建设理论经验和实践交流平台、汇聚全球力量助推数字中国建设的合作平台。

4.智博会

中国国际智能产业博览会（Smart China Expo，简称“智博会”），是经党中央、国务院正式批准，由科学技术部、工业和信息化部、中国科学院、中国工程院、中国科学技术协会和重庆市人民政府共同主办的展会。

2018年5月，经党中央、国务院同意将中国重庆国际汽车工业展与中国（重庆）国际云计算博览会合并，并更为现名；决定智博会从2018年起，每年在重庆市举办一届。

2019年4月，中共中央总书记、国家主席、中央军委主席习近平视察重庆工作时作出“继续高标准办好智博会，深度参与数字经济国际合作”的重要指示，遵照这一重要指示，2019年智博会坚持“智能化：为经济赋能，为生活添彩”的主题，智汇八方、博采众长，重点围绕“会”“展”“赛”及“论”，集中展示全球智能产业的新产品、新技术、新业态和新模式等。

数字经济的发展已呈现出越来越清晰的特征：

数据信息资源逐步成为新的关键要素资源；

数字技术创新是数字经济持续发展的原动力；

平台化是数字经济的主要产业组织形态；

产业融合是数字经济的主要表现形式；

多元共治是数字经济时代必然的治理要求；

网络空间成为驱动实体世界变革的关键力量。

以构建数据资产体系、释放数据资产价值为核心的数据中台被推到了广阔的舞台中央。

[1] http://www.gov.cn/guowuyuan/2015-03/16/content_2835101.htm [2] http://www.gov.cn/premier/2019-03/16/content_5374314.htm [3] <http://www.c114.com.cn/wireless/2935/a1093040.html> [4] http://www.tj.xinhuanet.com/xhvision/2019-07/23/c_1124787988.htm

1.2 数据中台的3个核心认知

数据中台能否从自发的单点状态、媒体热点，变成数字经济的基础、普惠性的数据服务机制，有赖于以下几个认知能否被业界广泛接受，并为之共同努力。

1.数据中台需要提升到企业下一代基础设施的高度，进行规模化投入

数据中台的目标是提供普惠数据服务，在“互联网+”行动计划和“智能+”的推动下，数字产业化和产业数字化成为数字经济的两大基础。数字产业化（互联网）从C端市场起步进而走向B端市场（互联网+），产业数字化天然在B端市场（+互联网）。数据中台只有在B端市场被企业提升到下一代基础设施的高度，才能帮助企业从根本上解决数字化转型过程中遇到的瓶颈和痛点，例如数据孤岛林立（其实是底层计算和存储架构的复杂性和异构造成的）、数据资产化程度低、数据服务提供效率与业务诉求严重不匹配等。相比于信息化部门把数据中台中的某些功能和特性作为新技术来局部验证和引入，数据中台更需要企业从战略高度进行顶层设计、确定规模化投入政策、设置更合理的组织架构，才能够确保数据中台作为数据应用的基础设施并落地建设，承担起企业数据资产全生命周期的管理。

2.数据中台需要全新的数据价值观和方法论，并在其指引下形成平台级能力

数据中台所包含的数据技术创新可以在成熟的平台型企业内部孕育，技术的创新和融合应用于很多贴近业务的创新应用场景。但数据中台不仅仅是技术平台，倘若停留于此，就完全忽略了IT到DT的本质变化是围绕数据资产，企业面临的主要矛盾是无法解决业务端的灵活性和经营管理稳定性之间的冲突，单纯地增大技术投入和人才投入都无法保障企业经营效能的持续提升。只有秉持数据价值观和方法论，才可能系统性地解决企业经营发展围绕数据的诸多问题，谁能率先解决面向数字经济特征的全新数据价值观和方法论的问题，并在其指引下打造出平台级能力，谁就能真正意义上帮助企业把数据用起来。

3.数据中台围绕业务、数据、分析会衍生出全新人才素养要求，需要尽快启动人才储备

人才永远是瓶颈，并且人才的具体定义在动态变化，需要为人才准备成长的土壤。信息化历程中从简单的搭建网站、单功能系统开发，到复杂系统开发、建设、运营，再到新技术引入等都曾经是人才具体定义的重要关注点。在社会范围内，信息化人才天然趋向两类企业：成熟稳定的平台型企业或有成熟平台潜力的企业。企业只有围绕数据中台明确了人才在企业的定位和职业通道，才可能吸引到或培养出拥有业务、数据、分析等综合素养的新型信息化人才，企业在数据中台人才储备上需要尽快做起来。

1.3 数据中台的3个发展阶段

“让数据用起来”，既是终极目标，也是数据中台要为处于不同数据认知成熟度阶段的企业实现的一个个具体目标。业务不会停滞，信息化不断追求自身的价值，数据部门力图与业务部门具有同等组织地位和话语权，业务部门不断提出新的挑战，政府在加速拉动数字经济建设……在这些因素的共同作用下，结合普惠数据服务按需取用、业务自流程化^[1]、数据自我治理的特点，在笔者看来，数据中台未来会经过以下几个发展阶段。

1.3.1 第一阶段：数据中台探索

这个阶段是个过渡阶段。一方面，传统的数据应用过往都是从外往内的（利用外部的技术、数据和资源来服务内部需求）。例如，零售行业要做精准营销，在广告上砸钱，做用户画像分析，利用外部的技术、数据、资源来服务内部需求，但是做完了会发现企业自身没有沉淀，又回到了原点。另外一方面，还是要借助一个个具体的场景化数据应用来推动企业对数据中台的认知，积累各行业（特别是头部客户）的业务和服务经验快速迭代和打造数据中台。

这个阶段会将数据生命周期各个阶段的技术与现有业务场景或创新业务场景结合，迅速形成可见、可展示的业务成果。特点是项目短小精悍，容易见效果，缺点是由于缺乏数据中台整体规划及让数据用起来完整流程设计，无法对众多单个数据应用沉淀的数据形成通用数据资产，每个项目都需要从头到尾走一遍，当应用需求爆发式增长时，底层数据支撑的效率会大幅度下降，甚至影响最终的业务效果。

1.3.2 第二阶段：数据中台整合数据应用提升效率

这一阶段的特点是构建数据中台的技术、理念、方法论是可复制的，市场上已有成熟的支撑数据中台高效运转的平台级产品。企业通过规划、建设、实施数据中台能够具备三方面的基础能力：

·数据多样性、多态性、多云连接能力（汇聚/交换能力）。交换的能力用来解决企业有哪些数据、数据在哪里等问题。

·数据资产化的能力是数据中台建设的关键，包括清洗、加工、治理、安全、质量等工具模块及实施方法论。（说明：能直接作用于业务领域，业务能阅读、能理解的数据才叫数据资产。）

·数据服务化的能力，用数据技术来使用数据的方法。

有了这三个能力，就能将上一阶段构建起来的场景级数据应用，甚至是历史建成的系统都整合成企业级数据应用平台，既能满足原有系统对数据的需求，又能快速满足新业务场景对数据的需求，将数据作为资产上架，成为共享的生产要素。

1.3.3 第三阶段：数据中台重构数据空间和业务空间

到了这一阶段，数据中台已经成为企业数据资产的核心能力和基础，通过快速构建数据资产体系，帮助企业真正实现对其全量数据的有效管理。业务和业务流程本身都可以通过适当的颗粒度进行数字化解耦和标准化，企业能够以自我为中心构建更加宏大的产业、行业价值链范围的数据空间和业务空间，以数据编排的方式响应业务需求，彻底颠覆传统的软件工程方式，业务实现自流程化，数据实现自我管理。

这里需要引入业务空间和数据空间的基本概念。

·企业业务空间：企业任何一个业务条线从初始设立到日益精细化，一般都遵循一个共性的演进过程：清晰定义该业务条线内专项业务的“毛细血管”功能体系、建设或升级相应技术支撑系统、生成专项业务数据。当所有业务条线都遵循这个发展规律，纵横交错的业务条线构成了企业实际运营的多维业务空间。企业的业务空间是产生和形成全量数据的根本依据和前提。

·企业数据空间：在数字化时代，任何一家企业都是市场生态中的一个节点，从数据交换的宏观视角来看，任何一家企业的数据全集只是整个市场数据生态空间中的一个子集。从企业自身视角来看，依据数据的生成和交互方式，企业全量数据的数据空间大致由三个维度构成：自主生产和消费的数据、外部数据（含单向外部获取数据和单向对外提供数据）、内外部交互数据。

[1] 此前没有人正式提过这一说法，在与政府、金融客户沟通时经常会提到：当业务能够实现对象数字化、规则数字化、结果数据化时，业务自身的流程也就可以按照规则自由、自行组建和优化了。

1.4 开启信息化的下一站

1.4.1 在多建设模式并存中做好准备

数字化需要探寻、挖掘数据的多维度业务价值，数据能力要在组织中孕育继而成为业务能力的一部分，需要摆脱“需求→立项→建设实施→日常运维”的简单模式，深入思考数据的业务本质。

回顾信息化建设的历程，笔者们发现因为企业自身组织变革和建设能力的限制，在相当长时间会有多种建设模式并存，以数据中台为驱动的建设模式将有可能成为融汇归一的新型建设模式，如图1-1所示。

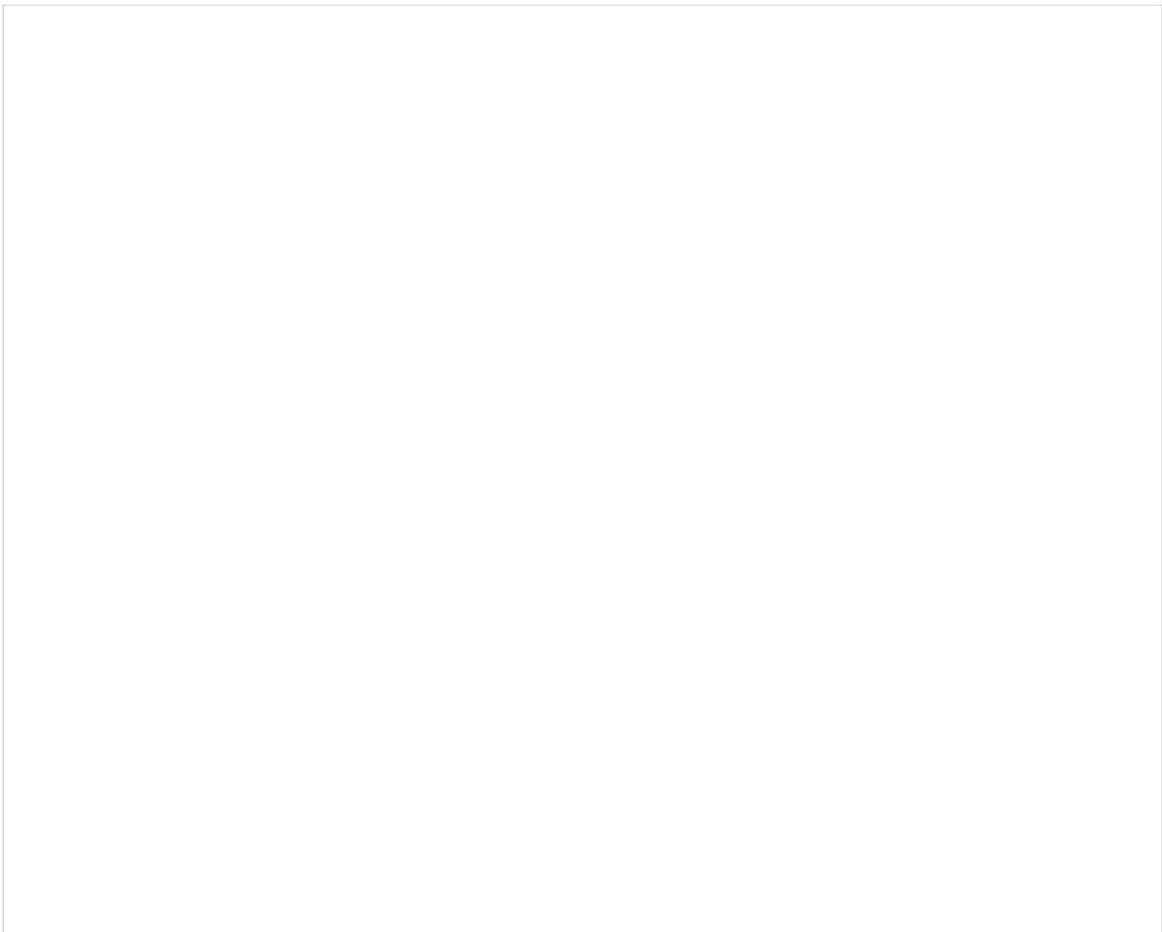


图1-1 数据中台建设模式融合

以下为4种主要的信息化建设模式：

·软件功能驱动模式：该模式对组织变革和建设能力要求最低，通常以采购和实施成熟产品为主，目标是业务部门直接能用。

·数据治理驱动模式：该模式的目标是针对同一数据不同问题或不同数据同一问题进行分类治理，通常是业务上遇到了难题，立个专项来解决。

·业务能力驱动模式：该模式对组织变革和建设能力要求最高，目标基于企业架构（EA）自上而下开展规划建设，覆盖组织从战略到执行全业务过程，从业务设计到IT实现。该建设模式实施难度极高，通常会形成顶层规划和一系列实施项目。

·业务服务化驱动模式：该模式专注于新技术的引入，通常是面向用户提升体验、面向业务拉通资源调度。

上述4种模式在建设数据中台时都可以有效融合，让组织变革和建设能力能够充分支持业务数据化、数据资产化、资产服务化、服务业务化的数据中台建设特点。

1.4.2 迎接数据中台新时代

2019年是数据中台爆发的元年，除了数澜科技与Forrester共同发布了首部行业白皮书《拥抱数据中台，加速数字化转型》，围绕数据中台的各种展会、发布会、产品也纷至沓来。

阿里、腾讯、华为做的是云计算基础设施，客户要做云计算的时候，他们就会给出解决方案。同样，数据应用也需要基础设施，当企业需要数据应用时，数据中台就会给出整体解决方案，真正“让企业的数据用起来”。

数据中台的需求不是来源于外部，而是来自内部，来自企业对自身未来发展的担忧。数据中台是增援未来，是以发展的观点解决企业面临的问题，面对不确定的未来，企业无法确认今天的数据未来会怎么用，会产生什么样的价值，所以才需要数据中台。现在把数据源源不断地接进来，源源不断地进行资产化、服务化，未来当企业看清楚业务场景，把对数据的需求输入数据中台时，才知道原来数据可以这样使用，才知道怎么去适配。数据中台是对未来场景的能力支撑，是增援未来的能力。

数据中台已经掀起了幕布的一角，幕布后面的精彩世界需要政府、产业、行业、领先企业共同激荡演绎。欢迎走进数据中台的世界。

第2章 什么是数据中台

伴随着云计算、大数据、人工智能等技术的迅速发展，以及这些技术与传统行业的快速融合，企业数字化、智能化转型的步伐逐渐加快。IDC预测，到2021年，全球至少50%的GDP将被数字化，而每个行业的增长都会受到数字产品与服务、数据化运营的驱动。

数字化转型成功的企业，其内部和外部的交互均以数据为基础。业务的变化快速反馈在数据上，企业能够迅速感知并做出反应，而其决策与考核基于客观数据。同时，数据是活的，是流动的，越用越多，越用越有价值。随着数据与业务场景的不断交融，业务场景将逐步实现通过数据自动运转和自动优化，进而推动企业进入数字化和智能化的阶段。

传统IT建设方式下，企业的各种信息系统大多是独立采购或者独立建设的，无法做到信息的互联互通，导致企业内部形成多个数据孤岛。互联网、移动互联网的发展带来很多新的业务模式，很多企业尝试通过服务号、小程序、O2O平台等新模式触达客户、服务客户，新模式是通过新的平台支撑的，产生的数据与传统模式下的数据也无法互通，这进一步加剧了数据孤岛问题。分散在各个孤岛的数据无法很好地支撑企业的经营决策，也无法很好地应对快速变化的前端业务。因此需要一套机制，通过这套机制融合新老模式，整合分散在各个孤岛上的数据，快速形成数据服务能力，为企业经营决策、精细化运营提供支撑，这套机制就是数据中台，如图2-1所示。



图2-1 数据中台定位

本章主要阐述数据中台的定义和核心能力，并澄清几个与数据中台相关的概念，最后总结数据中台建设能为客户带来的业务价值和技术价值。

2.1 解码数据中台

与许多新概念诞生之初的境遇一样，数据中台目前正处于“定义混乱期”。

有人认为数据中台是云平台的一部分，同时包括业务中台和技术中台；有人认为数据中台是数据的共享、整合和深度分析；还有人认为数据中台是“计算平台+算法模型+智能硬件”，不仅有云端，还需要智能设备帮企业在终端收集线下数据……从服务方到客户方，对数据中台的理解并不相同，如同一千个观众心中就有一千个哈姆雷特。

笔者有幸见证了数据中台在中国从0到1的全过程，并在其中实践多年，对于数据中台的定义，笔者认为：**数据中台是一套可持续“让企业的数据用起来”的机制，是一种战略选择和组织形式，是依据企业特有的业务模式和组织架构，通过有形的产品和实施方法论支撑，构建的一套持续不断把数据变成资产并服务于业务的机制。数据来自于业务，并反哺业务，不断循环迭代，实现数据可见、可用、可运营，如图2-2所示。**

通过数据中台把数据变为一种服务能力，既能提升管理、决策水平，又能直接支撑企业业务。数据中台不仅仅是技术，也不仅仅是产品，而是一套完整的让数据用起来的机制。既然是“机制”，就需要从企业战略、组织、人才等方面来全方位地规划和配合，而不能仅仅停留在工具和产品层面。

以中国某大型央企集团的数据中台为例，该集团旗下拥有横跨金融、地产、零售的多条业务线。要做数字化转型，不仅是技术问题，更是组织与业务运转模式改变的问题，需要顶层战略规划和组织架构上的改变。这也是为什么各大互联网公司在宣布中台战略时，会伴随着组织架构调整。



图2-2 数据中台是一套“让企业的数据用起来”的机制

每家企业的业务与数据状况各不相同，业务对数据服务的诉求不同，数据中台的建设将呈现出不同的特点，没有任何两家企业的数据中台是完全相同的。数据中台的实施不仅需要一整套技术产品，更需要针对不同业务、数据、应用场景的体系化的实施方法和经验，过程中涉及企业战略、组织、技术、人才等全面的保障和配合。

2.2 数据中台必备的4个核心能力

早在2015年，数字化领域的领先者已经开始从顶层战略设计入手，调整组织架构，协调内外部利益，更新方法论和认知体系，着手构建数据中台体系。从2018年下半年开始，以数据中台战略为核心的变革潮流席卷互联网行业，然而多数企业对数据中台内涵的认识仍不够全面，导致业务落地和商业创新还是困难重重。

数据中台需要具备数据汇聚整合、数据提纯加工、数据服务可视化、数据价值变现4个核心能力，让企业员工、客户、伙伴能够方便地应用数据。^[1]

1. 汇聚整合

随着业务的多元化发展，企业内部往往有多个信息部门和数据中心，大量系统、功能和应用重复建设，存在巨大的数据资源、计算资源和人力资源的浪费，同时组织壁垒也导致数据孤岛的出现，使得内外部数据难以全局规划。

数据中台需要对数据进行整合和完善，提供适用、适配、成熟、完善的一站式大数据平台工具，在简便有效的基础上，实现数据采集、交换等任务配置以及监控管理。

数据中台必须具备数据集成与运营方面的能力，能够接入、转换、写入或缓存企业内外部多种来源的数据，协助不同部门和团队的数据使用者更好地定位数据、理解数据。同时数据安全、灵活可用也是绝大多数企业看重的，他们期望数据中台能协助企业提升数据可用性和易用性，且在系统部署上能支持多种模式（见图2-3）。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.cn>)

文档名称：《数据中台：让数据用起来》付登坡 著.epub

请登录 <https://shgis.cn/post/1007.html> 下载完整文档。

手机端请扫码查看：

