

大数据与决策研究

(政策与技术跟踪专题)

2021年第17期(总第60期)

广西壮族自治区信息中心

广西壮族自治区大数据研究院

2021年5月27日

编者按：随着数据量的爆发式增长、数据处理技术的进步，以及数据中台逐渐实现商业化，2019年常称被为数据中台元年。当前我国数据中台行业处于从萌芽转向高速发展的过渡期，由于企业数字化转型驱动市场需求不断增加，行业增长势头明显，市场规模快速扩张。本期主要介绍数据中台相关情况。

本期要目

- ◆ 数据中台的概念、架构与核心价值
- ◆ 数据中台相关进展及发展趋势

数据中台的概念、架构与核心价值

一、数据中台的概念

数据中台不是简单的一套软件系统或者一个标准化产品，更多的是一种强调资源整合、集中配置、能力沉淀、分步执行的运作机制，是一系列数据组件或模块的集合。目前在业界对于数据中台尚无明确的定义，一般认为数据中台居于前台和后台之间，是企业级的数据共享、能力复用平台，是数字化转型的基础和中枢系统。将企业全域海量、多源、异构的数据整合资产化，为业务前台提供数据资源和能力的支撑，以实现数据驱动的精细化运营。

从广义上讲，数据中台是一种企业组织管理模式和理念，集公司战略决心、组织架构、技术架构于一体，企业从战略上构建统一的协同基座即中台化组织，以协调和支持各业务部门。

二、数据中台的架构

从业内较为通用的架构来看，数据中台一般可分为四层：大数据技术平台、数据资产管理平台、数据分析挖掘平台、面向应用的主题式数据开放服务平台。总体架构图如图 1 所示。

(1) 大数据技术平台：为数据资产管理平台提供技术支撑，基于 Hadoop 生态体系构建，包含多个数据存储、计算框架，解决多源异构的海量数据采集、存储、计算等问题。

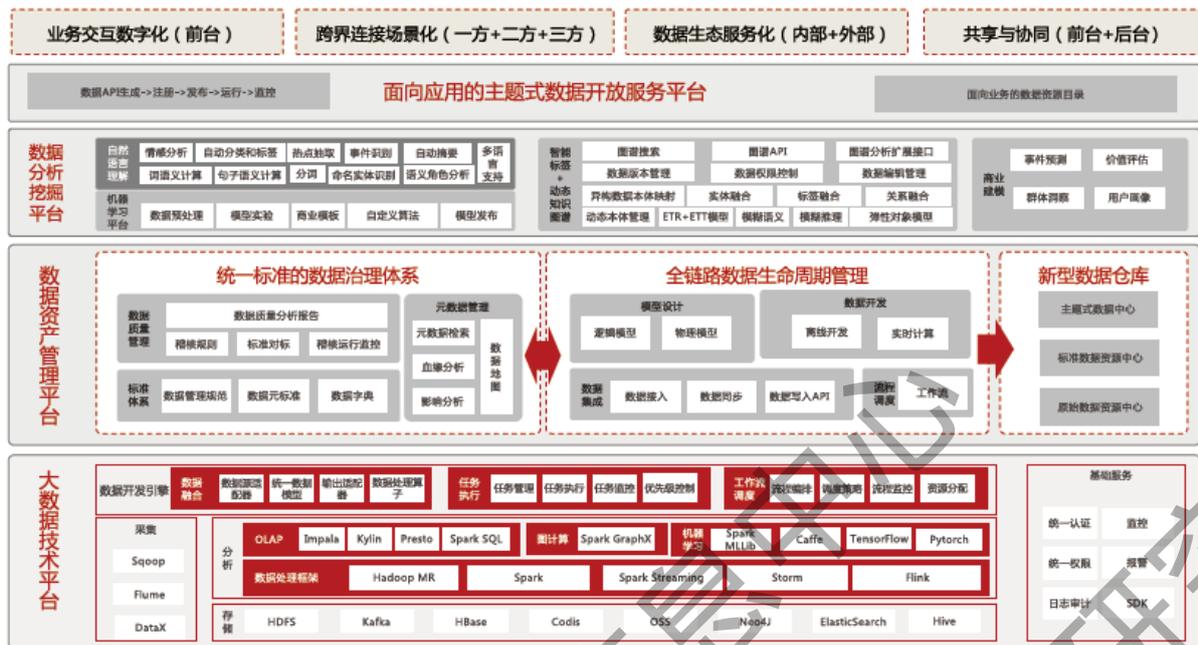


图 1 数据中台总体架构图

(2) 数据资产管理平台：基于大数据技术平台之上的数据管理中间件，用于盘点数据家底、构建统一的数据标准体系、构建行业化主题式数据仓库，以实现数据资产化为主要目的。数据资产管理平台通过数据开发引擎与底层大数据技术平台进行数据交互。

(3) 数据分析挖掘平台：架构在数据资产管理平台之上，为数据科学家、数据分析师提供稳定、高质量的跨主题数据资源。同时支持自然语言处理、机器学习建模平台、智能标签+动态知识图谱等多个易用的数据挖掘工具集。

(4) 面向应用的主题式数据开放服务平台：提供统一的、面向应用的、主题式的数据服务，将数据资产管理平台、数据分析挖掘平台的数据处理和分析结果以数据服务形式对外提供，同时生成以业务为导向的服务资源目录，让前台

应用更清晰的使用数据中台里的各类数据，实现以数据驱动业务，促进前台业务。

三、数据中台和数据仓库、数据平台的关键区别

数据中台的实质是构建全域数据共享的中心，提供数据采集、数据萃取、数据服务等全链路一体化的服务，提供面向企业业务应用的数据智能平台；数据平台是以存储、运算、显示为目的的平台，它是以处理海量数据存储、计算及不间断流数据实时计算等场景为主的一套基础设施，主要是直接为业务提供数据集服务；数据仓库则是信息的集中存储库，为企业提供所有类型数据支持的战略集合，创建的目的主要是出于分析和决策支持，它以前端查询和分析作为基础，存在的问题是冗余较大，需要的存储空间较大。通过对比可以看出，数据中台更加贴近具体业务，可以为业务提供更快捷的服务，企业可以在已有的数据平台和数据仓库之上构建数据中台，也可以把数据中台看作为企业从具体数据到业务价值实现过程的中间层。

四、数据中台的核心价值

（一）降低数据建设成本，提高数据治理效率。

数据中台的建设可以帮助企业打通数据孤岛，并建设统一的数据标准，包括数据建设规范和数据消费规范。数据中台基于原有的数据关系及 SOA（即面向服务的架构）等企业数据管理的经验，能解决企业信息管理中“数据烟囱”的问题，从全生命周期的角度管理数据。随着数据中台的建设，

数据二义性逐渐消除，透明度和利用率大大提高，有效发挥数据及分析技术对前台业务的复用价值，降低数据计算与数据存储成本，减少因数据体系建设不一致或重复建设导致的人力成本浪费等。

（二）激活数据商业价值，赋能企业运营与决策。

数据中台与过去的数据工具相比，通过将数据资产化，将不同系统、不同类型的数据纳入一个可对比、可计算的范围，使其更易于企业日常经营活动中进行搜索、过滤和管理，充分激活数据的商业价值。数据中台在实现数据接口标准化和在线交互实时化的基础上，集成可快速复用的数据生产力工具或模块，使数据具备敏捷地对外服务的能力，智能服务全流程的部门及人员，使每个层级的员工都能快速制定适合自己的数据决策服务，有效赋能业务决策。

（三）改造企业业务流程，升级企业组织架构。

传统的业务流程通常呈现“流水线”的特点，数据仅仅是用于监测业务进展和洞察规律的副产物，最终的决策由业务人员进行，因此决策不确定性较强，整个业务流程的迭代速度较慢，很难与当前快速变化的前端应用匹配。而随着数据中台在整个业务链条中的部署和应用，大数据进入决策阶段，企业的业务流程也逐渐快速、扁平化，由原先依赖业务人员经验的流程驱动逐步转向数据驱动。

此外，传统企业数据孤岛、业务割裂、资源分配等问题，其根源往往来自于组织架构的分割，尤其当业务需要涉及跨

部门协同时，“部门墙”的现象十分严重。数据中台的部署应用既是打通了数据的壁垒，更是打通了部门、事业群间的壁垒，使企业组织灵敏性得到提升。

五、数据中台建设的关键点

建设数据中台的关键点简要概括为两方面。

一方面，数据中台一定要与业务价值对齐。梳理清楚业务对于数据的诉求是构建数据中台的第一步，在业务场景还没有明确、优先级还不清晰、价值度量体系尚未建立起来的时候，不盲目建立大而全的数据平台。大而全的数据平台大量功能看上去很有用，但是缺乏应用场景；当真的有了场景，会发现不能开箱即用，还需更多的定制化。

另一方面，数据中台应该从小数据、小场景做起。开始的时候进行顶层设计，面向业务愿景制定中台的整体规划，全面梳理数据创新全景蓝图，通过业务愿景驱动出所有的业务场景，从而推导出数据中台的全景架构、技术支撑。但是在实施的时候，要从具体的业务场景出发。从高价值数据集场景做起，然后顺着这个场景竖切，找到数据全景图中的一个或多个数据集合，从小数据场景落地，快速验证价值。从大处思考，全局拉通，避免后续的数据孤岛，从小数据集、可实现性高的场景切入，然后将一个个的场景做起来，业务价值和中台能力也就可以同步地建立起来。（《2021年中国数据中台行业白皮书》《数据中台技术在业务系统中的应用研究》等）

数据中台相关进展及发展趋势

一、行业应用进展

（一）数据中台在互联网落地情况

阿里巴巴于 2015 年开始实施数据中台战略后，据 2015 财年统计，阿里集团批量数据计算总时长减少 50%，数据存储节约上百 PB 空间，直接降低成本数亿元。通过中台面向应用提供的数据服务方面，更是对营业额的大幅增长起到了直接的促进作用。如今阿里正在推广云数据中台，融通多元化、复杂化、丰富化的业务数据，以达到技术降本、应用提效、业务赋能的目标。

美团数据中台紧紧围绕业务开展建设，其中核心的一项业务场景即是建立集团统一的会员服务。美团获客渠道广，用户数量庞大，通过数据中台将不同应用中的用户数据标准统一、打通融合后，可精准刻画用户画像，为用户提供更好的生活体验。对美团来说，用会员打通各业务，使得各高频业务能够作为拉新手段让会员关注到并尝试其它低频业务，可从会员和业务双通道获取直接的价值收益。

2018 年腾讯、京东都对各自企业组织架构进行了调整，其中相同的一点是都对企业中台业务的战略发展做出了重要部署，强化数据中台的建设。如今，更多的互联网企业正抓紧建设符合自身业务发展的数据中台，通过数据中台实现

一切数据业务化、一切业务数据化，驱动企业业务精细化管理，带动产业数字化转型。

（二）数据中台在传统行业落地情况

金融行业是较早开始构建数据中台的行业。商业银行、证券公司和保险公司等，在互联网和大数据的技术推进下，业务不断转向线上，急需将线上产生的海量消费者行为数据与线下网点的数据结合起来，对客户进行综合的分析。多家金融企业发现，传统的 Oracle、SAP 数据库已经不能支撑大数据环境下的内部分析、营销和服务的需求，纷纷转向搭建基于大数据的数据中台，从而更好地支持业务需要的海量数据挖掘、搜索等服务。中信集团、招商证券、光大集团等金融企业在构建数据中台和业务中台方面已经走在前列。

在零售行业，百货公司、商超、品牌商也在布局全渠道战略和数字化转型，而搭建底层数据中台和业务中台是他们的首要选择。王府井集团搭建了集团大数据资产管理平台，完成集团线上线下用户、商品、交易层面的数据资产化。在数据中台支撑下，集团上线的智能营销系统覆盖了全国 30 多家门店，智能营销系统内嵌 11 个模型，完成潜客营销、大促营销、RFV 精准营销、广告投放四大场景设计，促进营销更加智能精准，实现营销活动数字化运营，有效提升活动业绩。

碧桂园、龙湖、新城等地产公司纷纷构建集团的数据中台，将住宅地产的业主数据、商业地产的会员数据和物业数

据等进行拉通，从而形成全面的用户画像和动态的标签体系，并进行跨业态的数据分析和精细化运营，均取得不错的成绩。

（三）数据中台在政府落地情况

各地政府借助数据中台的方法论和相关成熟配套产品实施政务数据中台，梳理数据标准体系，融合公安、民政、教育、工商、税务等众多部分数据，建立人口库、法人库、地理库、电子证照库等基础数据库，依托基础数据库结合各业务应用建立征信、城管、交通、教育、医疗等一系列专题数据库。

浙江省应用大数据、云计算技术，推进“互联网+政务”建设，人口、法人、信用库归集超 60 亿条数据，在全国率先出台首个省级公共数据和电子政务管理办法，让老百姓到政府办事只跑一次成为常态。杭州市不动产登记服务中心开发微信导航功能，群众可实时查询市区内 10 个办事网点的排队情况，选择空闲网点前往办理；杭州市通过对交通数据实时采集、融合、挖掘分析，创新应用 AI 智能技术，打造智慧化交通，可实现红绿灯信号灯自动配时，根据各路口交通流量和拥堵情况，灵活调整红绿灯时长，有效地提高了道路通行效率，降低道路拥堵程度。

上海市推进电子证照库的建设和应用，综合运用基于互联网+技术应用、数据资源共享、人工智能辅助等手段，实现了包括营业执照、身份证、居住证、出生证明、结婚证、

离婚证、食品经营许可证等 20 余种证照 221 个事项的场景应用。企业和市民在办理相关业务时，不再需要提交多样繁琐的纸质证件信息，系统可从电子证照库调取材料，节约办事时间，提高工作效率，为人民群众提供便利。

二、行业应用面临的挑战

（一）企业内外的认知和推广仍面临挑战。

从企业内部来看，管理者对数据治理一知半解，如果在没有深入梳理企业业务现状及需求的情况下盲目建设数据中台、追求“大而全”的概念，可能导致数据中台落地效果不佳。如果数据中台不能发挥降本提效的作用，反而使得实际使用者在适应时耗费更多的时间和精力，那么数据中台在企业内部的推广必然受到阻碍。

从市场环境来看，数据中台外在表现为偏定制化的解决方案，它们涉及不同的应用领域、覆盖不同的生产环节，很难有一套面对全行业、全领域通用的中台产品，因此企业对数据中台的认知非常依赖市场内正确信息的传递和头部企业的成功案例示范。整体而言，认知和推广是数据中台应用的挑战，这需要供应商和企业长期共同协作来解决。

（二）数据标准化困难、耗时费力。

多数企业在过去的信息化建设过程中，有一些遗留的数据问题待解决：一是多源异构的数据需要标准化：企业从不同的角度切入，多点建设了很多单链的业务系统，造成不同系统中架构、标准不一致，同一个项目中生产、运营、营销

的数据编码定义不一致。二是各种信息系统积累的数据需要集成打通：数据在保持及时性、准确性和完整性的同时进行 PaaS 层和 IaaS 层、各系统之间的集成打通。三是业务数据量大，多源异构数据处理的技术水平要求高。四是各部门数据共享意愿低：部分数据的敏感性、重要性较高，相关部门共享资源的意愿较低。这些问题在客观上阻碍了数据共享、复用的过程。

（三）技术与业务的融合存在壁垒。

尽管数据中台服务商逐渐形成了优势行业分布竞争的格局，但是数字化转型过程中的企业场景多样，同一行业的业务场景也会因为企业文化、经营模式的差异而有所不同，技术难以融合业务是市场上普遍存在的实际问题：供给方拥有技术能力，缺少业务场景的实际落地经验；需求方服务于场景，需要融合技术以带来切实的效益增长。深入理解业务并将技术成功与场景结合以产生收益是供给端与需求端共同面对的最大挑战。

（四）投资回报率等收益指标难以量化。

一方面，数据中台的成本投入有不确定性。数据中台的建设周期长，投入成本高，随着业务的快速增长，可能造成初期对数据增长的预估错误导致整个建设过程中成本的极大波动。另一方面，数据中台的回报难以量化。首先，数据中台输出的是无形的数据能力，这就要求其应对业务有充分的理解，否则企业将难以发挥效用。其次，不管是使用者还

是使用场景都多元而复杂，难以界定数据中台的直接受益。最后，企业业务日趋敏捷，数据中台对这种敏态提供相应的动态调整需要长时间检验其实际效果。因此，对于企业而言，存在数据中台选型困难，实施风险大，且投资收益难以量化的问题。

三、数据中台的发展趋势

（一）深入下沉市场，产品更加标准化。数据中台的核心在于共享和沉淀能力。在深度上，数据中台厂商承载细分行业的各类定制化业务，不断沉淀业务能力。在广度上，随着不同业务场景的持续输入，数据中台厂商产品的能力越来越丰富，覆盖的领域也越来越广泛。完善数据中台的深度和广度，提炼和整合数据中台的服务，尤其是对于对数据中台能力要求相对简单的中小企业，为客户提供标准化的整体解决方案将成为数据中台服务商的产品方向。

（二）深耕细分领域，场景愈加精细化。首先，数据中台所提供的底层技术支撑能力，需要供应商在软件架构、云技术、容器编排、DevOps 等多方面有充足的技术储备，还需要具备资本和技术实力的双重积累。我国数据中台行业大致形成了以阿里、腾讯等技术雄厚的头部企业侧重提供底层架构技术，其他中小供应商侧重提供行业化服务和产品的竞争格局。其次，没有一家供应商可以覆盖企业庞大的、所有的需求，尤其是多组织、多板块、跨业务的大型企业，所以在一个领域内已经完成实践和形成规模的供应商会优先深

耕本领域，提供更加细分的场景切入口。最后，企业也会根据业务需求面向不同领域的数据中台产品进行选择，不会局限于一家中台服务商。随着创业公司不断成长，细小赛道逐渐被填充，愈加激烈的市场竞争会使差异化成为供应商采取的产品战略。

（三）趋近软件形态，技术的 SaaS 化。从内部来看，数据中台不断沉淀跨行业、跨企业复用的组件、模块，存在朝 SaaS 和本地部署混合模式发展的趋势。从外部来看，随着云计算的普及，部分系统 SaaS 化趋势较强。因此，作为前台和后台的连接，数据中台与 SaaS 应用融合对接的实践越来越多，市场将逐渐形成一套成熟的中台+SaaS 系统融合闭环方案。随着数据中台市场渗透率的提高，允许通过零代码或少量代码就可以快速创建应用，对企业运维团队的要求降低，将充分提升数据中台的应用性。

（四）融合新兴技术，释放更多能量。现阶段数据中台更多的是描述基于云计算、大数据、人工智能等新一代技术打造的持续演进的企业级数据共享服务平台，是技术和方法论的结合。首先，人工智能使得数据采、存、通、治的加工链条大幅缩短。在应用侧，数据智能将更多地辅助业务决策、提高生产效率，赋能更多应用领域。其次，大数据以丰富的数据计算和存储技术为数据中台提供了强大的数据处理能力。最后，云端是如今企业数字化、智能化的基础，企业上云是大势所趋。云计算的按需付费、弹性扩展等特性，使其

使用和迁移成本更低，为更多企业运用数据中台等新技术提供了条件。随着数据源越来越丰富，数据使用场景越来越多元，数据中台将会融合更多新兴技术，释放出更多能量。

（五）人才需求从单一性向复合型转变。数据中台所包含的技术、组织、方法论的革新会衍生出全新人才需求。企业需要人才不仅会用数据，更要对业务有洞察、理解并能应用到数字化转型过程中，创造出更多的商业场景和机会。这种综合能力必须经过系统培养，企业原先的组织架构、人才培养计划不能适应快速发展的信息化要求，人才观念的转型成为数字化转型中不可忽略的重要部分。（数据与计算发展前沿、艾瑞咨询）

编辑部地址：南宁市体强路 18 号广西信息中心 1412 号房

联系电话：0771-6113592

电子邮箱：dsjyjs@gxi.gov.cn

网 址：<http://gxxxzx.gxzf.gov.cn/>



扫描二维码获取
更多决策参考信息