

第3章



主数据驱动的数据治理

企业数据资产可划分成为主数据、业务数据、分析数据三个主要部分。企业运营的主体活动便是围绕这三种数据资产展开。主数据和业务数据支撑起企业的业务流程,而主数据和分析数据则是企业商务智能的基础。其中,主数据会出现在所有重要的业务流程和分析任务中,是企业数据资产中的黄金部分。

为确保企业可进行跨业务领域、跨职能部门、跨信息系统的业务协作和整体分析,需对主数据、业务数据、分析数据进行数据治理,保证其一致性,提升数据质量和数据安全水平。因此,主数据、业务数据和分析数据构成了企业数据治理中三个核心的治理域。源于主数据的基础作用,主数据治理是业务数据治理和分析数据治理的前提,为业务系统和分析系统提供基础性的数据服务。

主数据驱动的数据治理以主数据作为数据治理的具体切入点而展开,为企业数据治理提供了可操作的治理框架和治理过程。

- 数据治理框架是指为了实现数据治理的总体战略和目标,将数据治理领域所蕴含的基本概念组织起来的一种逻辑结构,可为企业的数据治理实践提供理论指导。
- 主数据驱动的数据治理框架抓住主数据在企业数据资产中的核心位置,综合考虑数据治理在战略、管理、过程、技术等方面的任务和要求,明确了企业数据治理中的关键要素和核心过程。

3.1 数据治理框架

数据治理是围绕数据资产展开的系列工作,以服务组织各层决策为目标,是数据管理的技术、过程、标准和政策的集合。数据治理是一个复杂的系统工程,需要决策者、管理者、系统开发人员、系统使用人员、系统维护人员多方协作才能进行,因此构建科学的数据治理框架是开展数据治理工作的首要任务。

数据治理框架是指为了实现数据治理的总体战略和目标,将数据治理领域所蕴含的基本概念(如原则、组织架构、过程和规则等),利用概念间关系组织起来的一种逻辑结构。它

用于描述数据治理领域的基本组件(概念)以及组件间的逻辑关系。引入数据治理框架的目的是为组织的数据治理具体实践提供理论指导,确保数据治理付出的努力获得应有的价值回报。

数据治理框架通常需要明确如下几个方面:

- 每个组件的职能以及组件间的逻辑关系。
- 数据治理的工作范围和重点。
- 数据治理的工作任务和目标。
- 建立清晰的组织架构和职责分工。
- 建立数据治理成效的评估标准。

为了指导组织有效开展数据治理工作,国际研究机构在各自研究成果和实践经验的基础上,提出了一些通用的数据治理框架,这些框架为各机构的数据治理工作提供了不同的价值视角和关注维度,下面将对其中最具影响力的机构及其数据治理框架进行介绍。

3.1.1 国际标准化组织

国际标准化组织(ISO/IEC^①)所提出的数据治理框架建立在IT治理的基础上。2015年,国际标准化组织IT服务管理与IT治理分技术委员会制定了ISO/IEC 38500系列标准,提出了IT治理的通用模型和方法论,并认为该模型同样适用于数据治理领域。

在数据治理规范相关的ISO/IEC 38505标准中,阐述了基于原则驱动的数据治理方法论,提出通过评估现在和将来的数据利用情况,“指导数据治理准备及实施”,并监督数据治理实施的符合性等。

ISO/IEC 38505为组织的治理主体提供数据治理指南,组织的数据治理主体可以应用上述基于原则的方法来开展数据治理活动,在减少数据风险的同时提升数据的价值。如图3-1所示,该标准主要关注治理主体评估、指导和监督数据利用的过程,而不关注数据存储结构、恢复等数据管理活动。该标准强调数据治理的责任主体在治理层,治理层在开展数据治理的过程中主要通过制定数据战略来指导数据管理活动,而管理层需要通过管理活动来实现战略目标。同时,治理主体需要通过建立数据策略来保障数据管理活动符合数据战略的需要,进而满足组织的战略目标。

该标准实际上是对IT治理方法论的进一步扩展,并未对数据治理的实施和落地提供有效的手段。在实践中,数据治理虽根植于IT治理,但两者之间又有明显的区别,IT治理的对象是IT系统、设备和相关基础设施,而数据治理的对象是可记录的数据。因此IT治理过程中过于强调IT投资和系统实施,忽视了商业价值增长中的数据创建、处理、消耗和交换方式。

^① ISO/IEC: International Organization for Standardization/International Electrotechnical Commission,国际标准化组织/国际电工技术委员会。

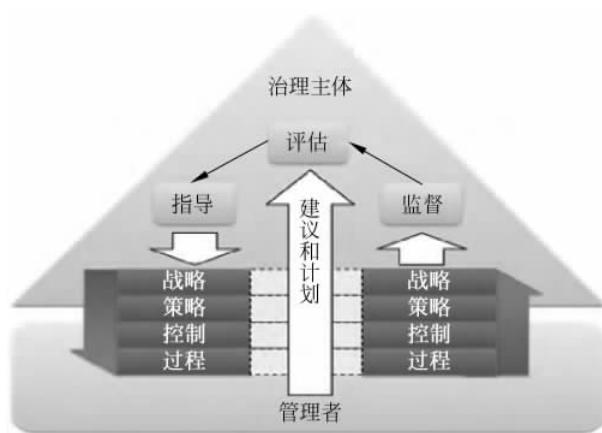


图 3-1 ISO/IEC 38505 数据治理框架

3.1.2 国际数据管理协会

国际数据管理协会(DAMA^①)提出的 DAMA-DMBOK 框架以数据管理为中心,认为数据治理是数据管理的组成部分,是数据管理的核心功能。DAMA 框架包括两个子框架:功能子框架和环境要素子框架(图 3-2 和图 3-3)。功能子框架总结了数据管理的 10 个功能,并将数据治理置于核心位置。环境要素子框架提出了数据管理的 7 个环境要素,并最终建立起 10 个功能和 7 个要素之间的对应关系。DAMA 框架中数据治理的核心工作就是解决数据管理的 10 个功能与 7 个要素之间的匹配问题。



图 3-2 功能子框架

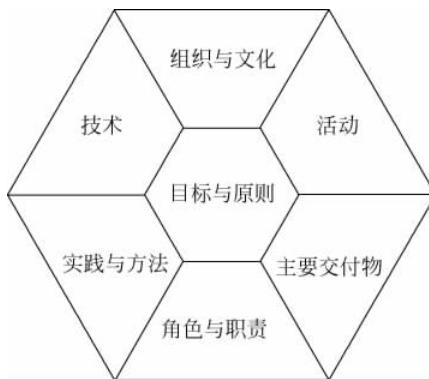


图 3-3 环境要素子框架

^① DAMA: Data Management Association International, 国际数据管理协会。

DAMA-DMBOK 框架对数据治理和数据管理的界定扩大了数据管理的范畴。一般情况下,我们更倾向于数据治理是为了确保有效管理而做的决策,强调决策制定的责任路径,而数据管理仅仅涉及决策的执行。同时,DAMA-DMBOK 框架更强调数据管理的各项职能以及关键活动,而对于实施数据治理的过程、评估的准则等未能给予清晰而系统的指导。

3.1.3 国际数据治理研究所

国际数据治理研究所(DGI^①)从组织、规则、过程三个层面,提炼出数据治理的 10 个基本组件,并在此基础上提出了 DGI 数据治理框架(简称 DGI 框架),如图 3-4 所示。该框架既包含从管理角度提出的促成因素(如目标、数据利益相关者和组织结构等),也包含项目管理的相关内容(如数据治理生命周期)。

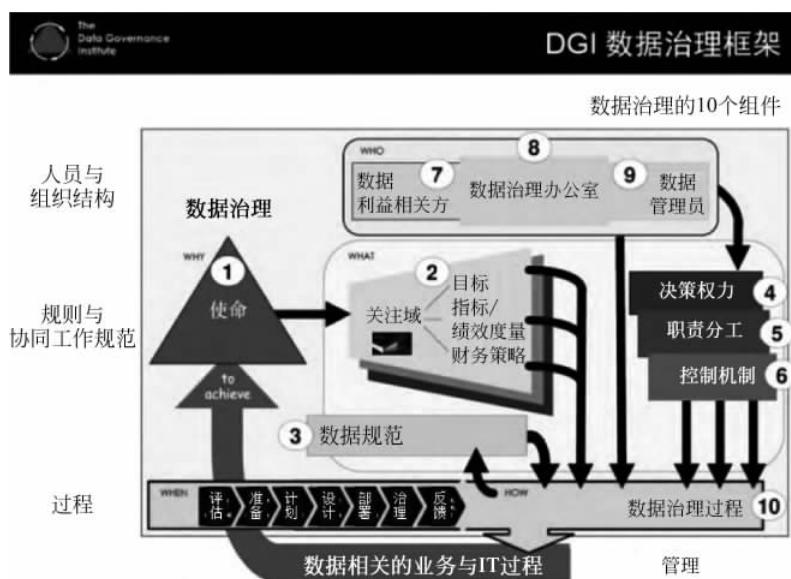


图 3-4 DGI 框架中数据治理基本组件

DGI 框架将 10 个基本组件按职能划分为三组:

- 规则与协同工作规范: 包括使命、目标、数据规范、决策权力、职责分工、控制机制。
- 人员与组织结构: 包括数据利益相关方、数据治理办公室、数据管理员。
- 过程: 包括评估状态、准备路线图、规划、设计/开发程序、数据治理、监测/度量/反馈等阶段。

DGI 框架以访问路径的形式,非常直观地展示了 10 个基本组件之间的逻辑关系,形成了一个从方法到实施的自成一体的完整系统。

DGI 与 DAMA 不同,它认为治理和管理是完全不同的活动,治理是有关管理活动的指

① DGI: Data Governance Institute,数据治理研究所。

导、监督和评估,而管理则是根据治理制定的决策来执行具体的计划、建设和运营。因此,数据治理独立于数据管理,前者负责决策,后者负责执行和反馈,前者对后者负有领导职能。因此,相比 DAMA 框架,DGI 框架的设计完全从数据治理角度出发,是一个更加独立、完整和系统的数据治理框架。

3.1.4 IBM 数据治理委员会

IBM 数据治理委员会(IBM Data Governance Council)通过结合数据特性和实践经验,有针对性地提出了数据治理的成熟度模型,将数据治理分为五级,即初始阶段、基本管理、主动管理、量化管理和持续优化。同时在构建数据治理统一框架方面,提出了数据治理的要素模型,将数据治理要素划分为支持规程、核心规程、支持条件和成果 4 个层级,如图 3-5 所示。



图 3-5 IBM 数据治理要素模型

IBM 数据治理委员会重点关注治理过程的可操作性,认为业务目标或成果是数据治理的最关键命题,在支持规程、核心规程、支持条件作用下,组织最终可以获得业务目标或成果,实现数据价值。

3.1.5 中国电子工业标准化技术协会信息技术服务分会

在积极参与并辅助国际标准化组织推进 ISO/IEC 38505 系列标准的同时,中国电子工业标准化技术协会信息技术服务分会(ITSS^①)服务管控组带领国内近百家机构开展了《信息技术服务 治理 第 5 部分: 数据治理规范》国家标准的制定。

ITSS 服务管控工作组是国内信息技术服务领域的信息技术治理和数据治理的标准制

① ITSS: Information Technology Service Standards, 信息技术服务标准。

定和研究机构。ITSS 结合国际数据治理标准的研制思路,遵循“理论性和实践性相结合、国内与国际同步推进、通用性与开放性相结合、前瞻性和适用性相结合”的原则,明确了数据治理规范实施的方法和过程,旨在评估组织数据管理能力的成熟度,指导组织建立数据治理体系,并监督数据管理体系的建设和完善。

标准包括正文和附录两大部分。正文面向组织的决策层,提出了数据治理的目标、任务、框架、顶层设计、环境等,明确了决策层的作用和责任,为决策层规划、监督数据治理提供指引。附录面向组织的管理层,对数据治理涉及的核心治理域提出了明确的管理要求,为管理层实施数据治理提供指引,为决策层监督数据治理成效提供参考。正文和附录的结合,使治理可以通过管理有效落地,打通了从治理到实施的路径,解决了国际治理标准不易应用落地的问题。

该标准适用于组织数据治理现状自我评估及数据治理体系的建立,数据治理域和过程的明确,数据治理实施落地的指导。数据治理相关的软件或解决方案的研发、选择和评价,数据治理能力和绩效的第三方评价。

标准把实施数据治理的目标总结为运营合规、风险可控、价值实现三个层面,机构可根据自身业务需求进行选择。其中,运营合规是基础目标,应保证数据及其应用的合规;在合规的基础上,建立数据风险管控机制,确保数据及其应用满足风险偏好和风险容忍度;以合规、可控的数据应用为基础,构建数据价值实现体系,促进数据资产化和数据价值实现。

标准澄清了数据治理的任务,认为组织决策层作为数据治理的主体,在数据治理过程中的主要任务包括评估、指导、监督。决策层首先要明确数据治理的目标,引导管理层开展数据治理现状及需求、数据治理环境、数据资源管理和数据资产运营能力的评估,找出现状和目标之间的差距,指导管理层开展数据治理体系的构建、数据治理域的选择、数据治理的实施落地,通过一系列数据治理工作提升机构的数据应用能力,缩小现状和目标的差距。决策层还应制定合理的评价体系与审计规范,监督数据治理实施的合规性和有效性,对于有偏差或不符合的内容进行指导,最终引导机构实现数据治理目标。

数据治理框架包含顶层设计、数据治理环境、数据治理域、数据治理过程 4 部分,如图 3-6 所示。

顶层设计包含数据相关战略规划、组织构建和架构设计,是数据治理实施的基础。数据战略规划应保持与业务规划、信息技术规划的一致,并明确战略规划实施的策略;组织构建应聚焦责任主体及责权利,通过完善组织机制,获得利益相关方的理解和支持,制定数据管理的流程和制度,以支撑数据治理的实施;架构设计应关注技术架构、应用架构和管理体系架构等,通过持续的评估、改进和优化,支撑数据的应用和服务。

数据治理环境包含内外部环境及促成因素,是数据治理实施的保障。治理机构应分析业务、市场、利益相关方的需求,适应内外部环境变化。同时,还应关注决策层对治理工作的支持程度、相关人员的职业技能、内部治理文化等,以支撑数据治理的实施。

数据治理域包含数据管理体系和数据价值体系,描述了数据治理实施的对象。其中数据管理体系包括数据标准、数据质量、数据安全、元数据管理、数据生存周期 5 个治理域。

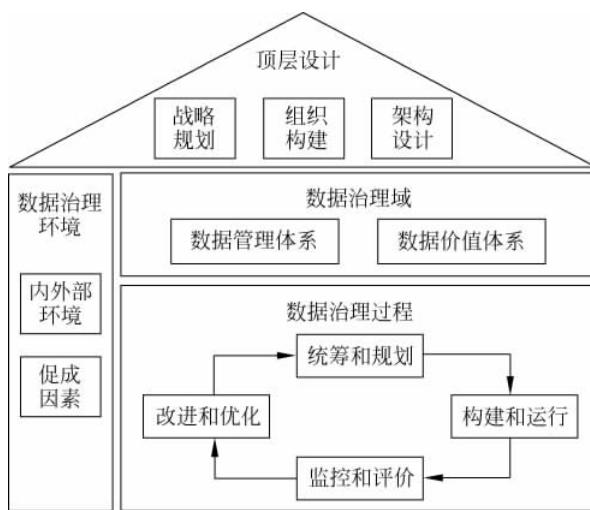


图 3-6 ITSS 的数据治理框架

数据治理过程包含统筹和规划、构建和运行、监控和评价、改进和优化 4 个步骤，描述了数据治理实施的方法。在统筹和规划阶段，应明确数据治理目标和任务，营造必要的治理环境，做好数据治理实施的准备；在构建和运行阶段，应构建数据治理实施的机制和路径，确保数据治理实施的有序运行；在监控和评价阶段，应监控数据治理的过程，评价数据治理的绩效、风险及合规，保障数据治理目标的实现；在改进和优化阶段，应改进数据治理方案，优化数据治理实施策略、方法和流程，促进数据治理体系的完善。

3.1.6 现有数据治理框架的局限

数据治理是综合性很强的领域，既涉及企业战略，又涉及具体的管理制度，同时还和数据管理的技术架构紧密相关。成功的数据治理，需要完善的方法论指导，合理的过程和制度保障，以及系统的平台和工具的支持。上述治理框架大多只涉及方法论、过程、制度和工具中的一部分，并且更侧重宏观的战略和管理问题，对于执行层面的工具方法关注不够，普遍存在落地困难的问题。同时，数据治理的范畴定义宽泛且模糊，导致数据治理的任务太过复杂庞大，数据治理从何处切入才能快速见效？如何将复杂的治理任务分解成为可操作的阶段性目标？这些企业特别关心的具体问题在上述一般性的框架中都没有针对性的说明。

同时，中国企业的数据治理在体制层面、管理对象层面和技术平台层面都存在显著的特色，这些特色在上述治理框架中关注不够。首先，国内强调数据标准建设，已经形成大量行业性数据标准（如金融、电子政务、公安、税务等）。这些企业、行业数据标准，主要覆盖元数据、主数据等静态数据内容，但也涉及交易数据等部分动态数据；同时，在技术平台层面，国内习惯将主数据、数据标准、数据质量、元数据等几部分功能统一形成数据资源管理平台并作定制化开发，力图通过一站式数据管理提升用户体验。因此，在数据治理的概念被广泛接受之前，已经存在大量数据标准化和数据管理的平台、工具和方法，但跨业务、跨部门、跨系

统的横向协同机制不顺畅,治理效果欠佳;同时,在数据治理体系的中层和基层缺乏可操作的数据治理方法和标准。如何在数据治理实施中妥善处理这些特色和问题,也是我国企业实施数据治理过程中关注的重点。

考虑数据治理的可操作性、中国企业数据治理的特色和普遍问题,本书综合多种数据治理框架,提出主数据驱动的数据治理框架。

3.2 主数据驱动的数据治理框架

3.2.1 治理思路和治理目标

数据治理的核心是加强对数据资产的管控,通过深化数据服务,持续创造价值。数据治理是在数据资产价值创造的过程中,治理团队对数据资产管理的评价、指导、控制,如图 3-7 所示。

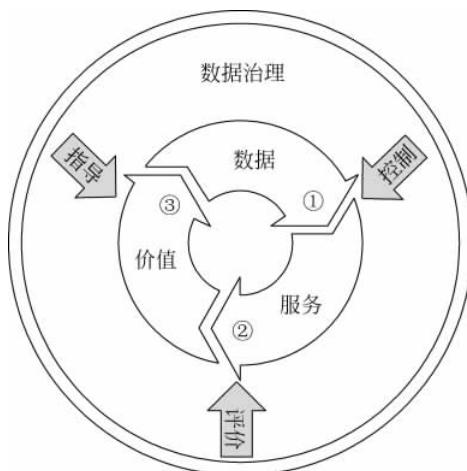


图 3-7 数据治理核心工作

根据企业中数据的特征、作用以及管理需求的不同,可将数据资产划分成为主数据、业务数据、分析数据三个主要部分。主数据包含元数据、引用数据、企业结构数据、业务结构数据等内容,这些数据相对慢变,但对企业具有全局的重要作用;业务数据是在交易和企业活动过程中动态产生的,通常具有实时性的要求;分析数据是对业务数据梳理和加工的产物,相对业务数据而言,实时性要求较低,通常按照分析的主题进行组织和管理。

企业运营的主体活动便是围绕这三种数据资产展开。主数据和业务数据支撑起企业的业务流程,而主数据和分析数据则是企业商务智能的基础。其中,主数据会出现在所有重要的业务流程和分析任务中,是企业数据资产中的黄金部分,如图 3-8 所示。

为确保企业可进行跨业务领域、跨职能部门、跨信息系统的业务协作和整体分析,需对主数据、业务数据、分析数据进行数据治理,保证其一致性,提升数据质量和数据安全水平。

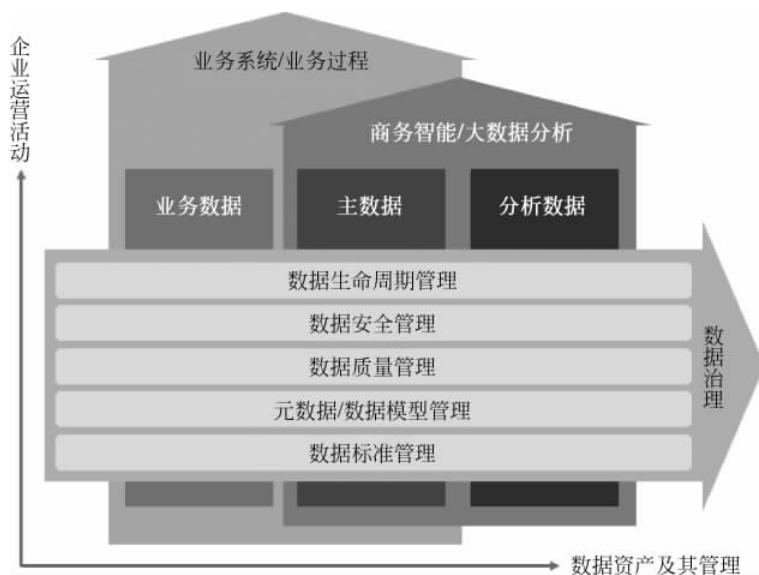


图 3-8 企业数据资产构成

因此,主数据、业务数据和分析数据构成了企业数据治理中三个核心的治理域。每个治理域都包含数据标准管理、元数据/数据模型管理、数据质量管理、数据安全管理、数据生命周期管理 5 个基本的管理组件。源于主数据的基础作用,主数据治理是业务数据治理和分析数据治理的前提,为业务系统和分析系统提供基础性的数据服务。而后续业务数据治理更关注改善数据流通,分析数据治理更关注改善数据洞察。

有别于传统的企业资产,数据资产来源丰富,可拷贝、可重用,这将导致数据搜集、存储、使用都具有特殊性。同时,数据还涉及个人隐私、运行安全等问题。当跨业务、跨部门、跨系统进行协作时,更需要数据的一致性。这些都是数据治理要解决的关键问题。因此,数据治理不但与数据标准、数据产生过程的业务规范相关,也涉及企业战略、管理决策架构等因素,是战略问题、管理问题、技术问题的综合。

企业必须使数据治理项目更贴近企业整体业务目标,需要真正将数据视为一种战略资产,构建统一的数据架构和管控体系以满足企业信息化的整体要求,并制定路线图,在实现短期目标的同时为企业未来目标做好准备。因此,构建数据治理框架可以从 4 个方面入手,分别是战略、信息基础架构、路线图、治理任务界定。如图 3-9 所示,首先需要确定信息能够影响的主要业务目标(战略愿景),其次标识那些能够满足信息需求的技术组件和功能,并充分利用现存信息资产获得治理的速度和灵活性(信息基础架构),再为企业建立长期和短期的规划(路线图),最后需要界定具体的治理任务(治理任务界定),以保证信息的一致性。数据治理通过协调人、管理流程和技术体系,提高企业信息质量、可用性、完整性和安全性,达到掌控信息的目的。

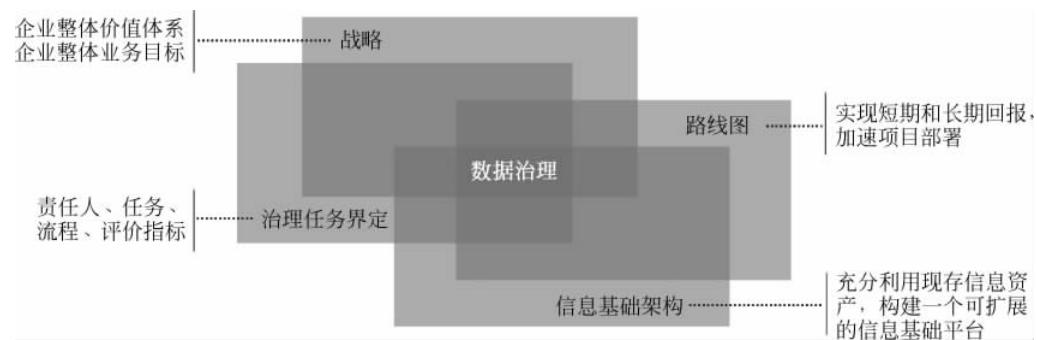


图 3-9 数据治理治理框架的 4 个方面

3.2.2 治理框架

基于对企业数据资产和经营活动的分析，综合考虑数据治理在战略、管理、过程、技术等方面的任务和要求，提出主数据驱动的数据治理框架，如图 3-10 所示。

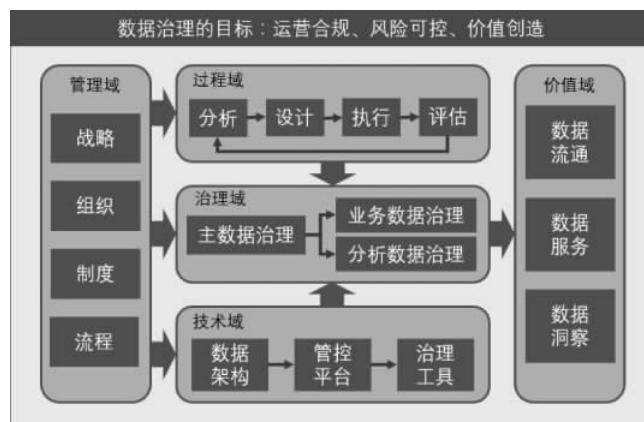


图 3-10 企业数据治理框架

企业开展数据治理之前，应首先明确数据治理的目标。参照 ITSS 提出的数据治理规范，本框架把实施数据治理的目标总结为运营合规、风险可控、价值创造三个层面，企业可根据自身业务需求进行选择。其中，运营合规是基础目标；在合规的基础上，建立数据风险管理机制，确保数据及其应用满足风险偏好和风险容忍度；以合规、可控的数据应用为基础，构建数据价值实现体系，促进数据资产化和数据价值实现。

将数据治理的各项任务和要素划分在 5 个不同的域内。

(1) 管理域。管理域是数据治理的主要驱动力量，负责确定数据治理的战略、组织、制度和流程。数据战略规划应保持与业务规划、信息技术规划的一致，并明确战略规划实施的策略。组织架构设计明确责任主体及责权利，通过完善组织机制，获得利益相关方的理解和

支持,制定数据管理的流程和制度,以支撑数据治理的实施。

(2) 治理域。治理域是数据治理的主体,明确数据治理的具体目标和责任。依据对数据资产构成的分析,将治理域分为主数据治理、业务数据治理、分析数据治理三部分,其中主数据治理是业务数据治理和分析数据治理的前提,为业务系统和分析系统提供基础性的数据服务。因数据特征和管理需求不同,三部分的治理任务有所区别,但都应包含以下基本的数据治理组件。

- 数据标准管理:规范了数据治理活动的内容、程序和方法,是相关管理人员和治理活动的行为准则,一般包括数据管理规范、数据应用规范、数据集成服务规范等内容。
- 数据模型管理:数据模型管理实现对元属性、数据约束条件、校验规则、编码规则等方面的规定与管理,以及对数据模型的创建申请、审批和变更申请、审批过程等流程的规定和管控。
- 数据质量管理:数据质量管理对数据从计划、获取、存储、共享、维护、应用和消亡过程中每个阶段可能引发的各类数据质量问题进行识别、度量、监控和预警,并通过改善和提高组织的管理水平,使数据质量获得进一步提高。
- 数据安全管理:数据安全管理通过数据分级、用户级别及权限的规定来防止无意、故意甚至恶意对数据进行非授权的访问、浏览、修改或删除。通过数据安全体系的建设,增强信息安全风险防范能力,有效地防范和化解风险,保证业务持续开展,满足内控和外部法律、法规的要求。
- 数据生命周期管理:数据生命周期管理用于管理信息系统的数据在整个生命周期内的流动,从创建和初始存储,到它过时被删除。通过数据生命周期管理可实现数据申请/转入、数据清洗、数据校验、数据审核、数据发布、数据维护等功能,最大限度地体现数据的价值。

(3) 技术域。技术域是数据治理的支撑条件,提供治理所需的数据架构、管控平台和治理工具,在IT整体规划的基础上,通过持续的评估、改进和优化,支撑数据治理的应用和服务。

(4) 过程域。过程域是数据治理实施的具体方法。数据治理过程包含分析、设计、执行、评估4个步骤。在分析阶段,应评价数据治理的成熟度、风险及合规性,发现问题;在设计阶段,应明确数据治理目标和任务,设计数据标准、数据模型、数据架构,做好数据治理实施的准备;在执行阶段,应构建数据治理实施的机制和路径,确保数据治理实施的有序运行;在评估阶段,应监控数据治理的过程,改进数据治理方案,优化数据治理实施策略、方法和流程,促进数据治理体系的完善。

(5) 价值域。数据治理的目标是通过对数据资产的有效管控持续创造价值,价值域通过对治理结果的有效整理,通过构建具体化的数据产品,实现上述的价值创造。数据治理的价值体系具体包括三个方面。

- 数据服务:通过数据的采集、清洗、导入,提升数据质量,确保数据的一致性。这部

分体现着主数据治理的关键价值。

- 数据流通：通过实现信息整合和分发机制，支持跨业务、跨部门、跨系统的信息流转和协同。这部分体现着业务数据治理的关键价值。
- 数据洞察：通过消除数据内在的质量缺陷，明确数据之间的关联关系，帮助数据分析人员更好地理解数据，实现数据洞察。这部分体现着分析数据治理的关键价值。

3.2.3 技术架构

为有效支持数据治理的开展，需要高效、灵活的技术架构和信息管控的工具。在主数据驱动的数据治理框架下，数据治理的技术架构如图 3-11 所示。其中基础数据平台、业务支持平台、数据分析平台分别承载着和主数据、业务数据、分析数据有关的企业运营、管理活动；数据交换平台将上述三个平台连接在一起，完成彼此之间的数据交换，是平台之间的数据通道；数据管控平台完成上述平台之间的协调，是平台之间的控制通道。

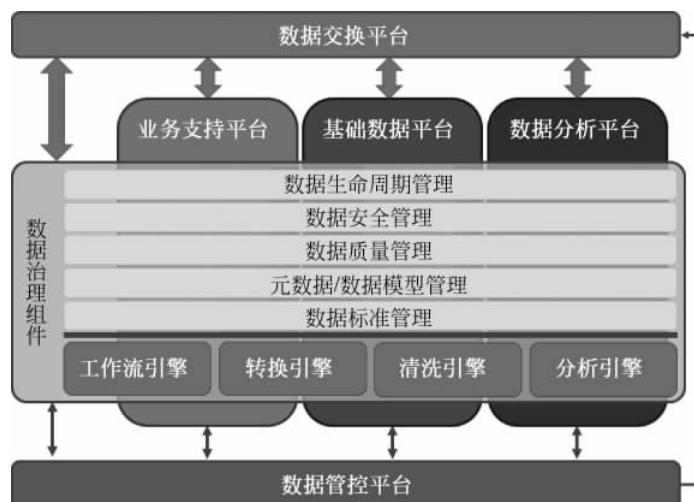


图 3-11 数据治理技术架构

数据治理组件辅助完成对主数据、业务数据和分析数据的治理。数据治理组件可看作是实施数据治理的 IT 工具包，其中包括数据标准管理、元数据/数据模型管理、数据质量管理、数据安全管理、数据生命周期管理 5 个基本的管理组件，实现治理框架中治理域的基本功能。同时，以工作引擎的方式提供一系列通用工具，包括但不限于：

- 工作流引擎：管理申请、校验、审核、发布等数据治理的工作流程。
- 清洗引擎：完成采集数据的清洗工作。
- 转换引擎：完成符合特定条件数据的批量修改和转换。
- 分析引擎：实现数据质量、数据生命周期管理中的分析任务。

数据治理组件也通过数据交换平台实现数据导入导出，通过数据管控平台完成和其他平台的协调。数据交换平台和数据管控平台以总线化的方式提供了可扩展的数据通道和控

制通道,而工作引擎则实现了数据治理组件的能力扩展。

3.3 主数据驱动的数据治理过程

3.3.1 过程框架

如图 3-12 所示,数据治理的任务可分解成为 5 个阶段,即架构阶段、主数据治理阶段、业务数据治理阶段、分析数据治理阶段、优化治理阶段。每个阶段都包含分析、设计、执行、评估 4 个基本环节,循环迭代,推动阶段任务的达成。

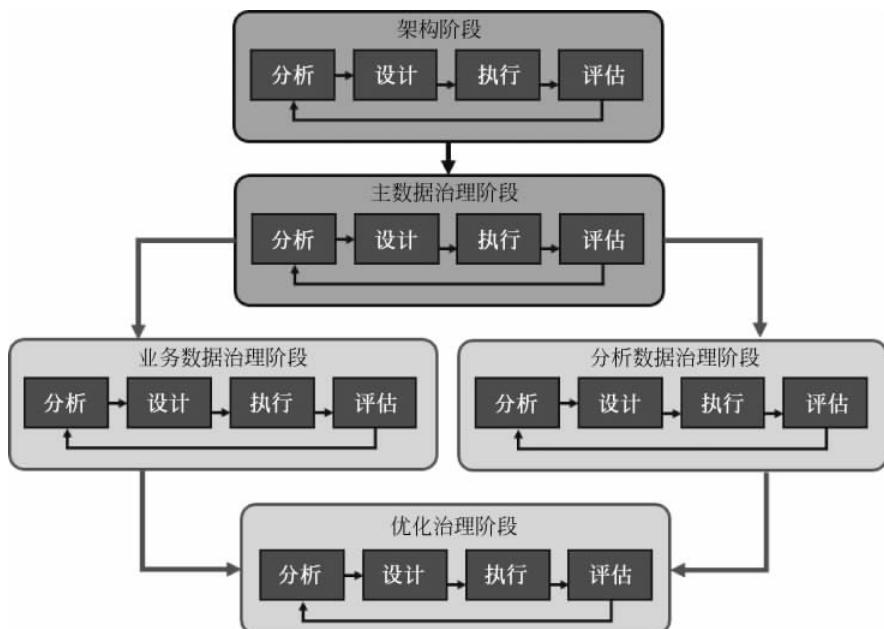


图 3-12 数据治理过程架构

上述过程框架提出的阶段性任务划分,将复杂庞大的数据治理任务分解成为较小的独立子集,增强了数据治理的可实施性,有利于企业管理者进行整体规划和安排,并对企业整体的信息化建设有一定的参考价值。

同时,上述过程框架可根据企业的整体战略和治理目标进行灵活剪裁,以满足不同企业数据治理的不同要求。其中架构阶段和主数据治理阶段是数据治理的基础,一般企业实施数据治理项目均需从架构和主数据治理开始,这也是本书第二篇重点介绍的内容。而业务数据治理阶段、分析数据治理阶段、优化治理阶段则可根据需要灵活选择。

3.3.2 架构阶段

架构阶段是数据治理的准备阶段。架构阶段通过现状调研和需求分析,识别业务问题

和实施风险,完成数据治理的整体规划和体系设计。同时,获得高层支持,创建数据治理管理组织,并完成管理成熟度的评估。架构阶段的主要任务及其要点如图 3-13 所示。

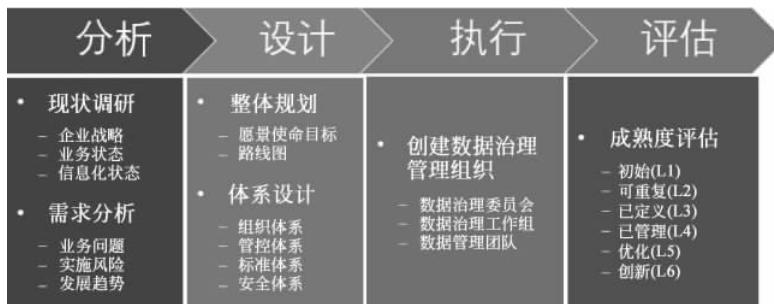


图 3-13 架构阶段的主要任务及其要点

3.3.3 治理阶段

主数据治理、业务数据治理、分析数据治理、优化治理 4 个阶段只是治理的对象和工作范畴存在差别,可参照相同的过程框架,主要任务及其要点如图 3-14 所示。



图 3-14 治理阶段的主要任务及其要点

考虑到主数据治理的核心作用,本书的第二篇将针对主数据治理的具体实施过程进行详细介绍,该过程也可作为业务数据和分析数据治理的参考。

分析环节完成业务过程分析和业务问题聚焦,并从数据质量、数据安全、数据生命周期三个方面分析、梳理数据缺陷,形成治理的阶段性目标和具体的工作计划;设计环节完成数据标准和数据模型的设计、开发,为数据治理的实际执行做好准备;执行环节依次完成数据采集、数据清洗、数据导入、应用集成、系统测试和上线切换,最终提交数据服务和数据产品,完成数据资产的价值提升;最后,在评估环节从运行情况、数据质量、数据安全、系统性能、管理水平、经济效益等方面对数据治理的效果进行评估,为进一步的优化提供条件。

3.3.4 任务、角色、分工、职责

数据治理是一项复杂的任务,仅仅依靠企业IT部门难以完成,往往需要专业的外部咨询服务和开发人员的辅助。因此,需要在过程框架中澄清各方的分工和职责,以便配合。表3-1以主数据治理为例,对数据治理过程中的任务、角色、分工、职责进行梳理,供企业的管理者参考。

表3-1 数据治理过程各方分工和职责

阶段	参考时长/周	任务描述	职责(RASIC)				
			R—负责,A—批准,S—支持 I—知情,C—咨询				
架构阶段	分析	现状调研	I	R	S	C	S
		需求分析	I	A	I	R	I
		差异对比	I	I	I	R	I
	设计	整体规划	A	I	I	R	S
		体系设计	A	R	I	R	S
	执行	建立数据治理委员会	A	R	I	C	S
		建立数据治理工作组	A	R	I	C	S
		建立数据管理团队	I	A	S	C	R
	评估	成熟度评估	I	R	I	C	R
治理阶段	分析	业务分析	I	A	S	RC	S
		数据分析	I	A	R	C	RS
		目标和计划	A	R	S	C	S
	设计	数据标准制定	A	R	R	C	R
		数据模型建立	I	A	I	C	R
	执行	数据采集	I	A	R	C	S
		数据清洗	I	A	R	C	S
		数据导入	I	A	R	C	S
		应用集成	I	I	I	SC	R
		系统测试	I	I	R	SC	R
		上线切换	A	I	I	C	R
	评估	运行情况评估	I	A	I	SC	R
		数据质量评估	I	A	R	R	S
		数据安全评估	I	A	I	R	R
		系统性能评估	I	I	I	C	R
		管理水平评估	A	R	I	C	S
		经济效益评估	A	R	I	C	S

3.4 数据治理工具和系统选型

数据治理的实施需要 IT 技术的有效支持。根据权威数据管理研究机构 TDWI 对数百家国际企业的调研结果,数据治理所涉及的 IT 技术主题及其重要性排序如图 3-15 所示。其中,元数据管理、主数据管理、数据质量的重要性在全部技术主题中位列前三。

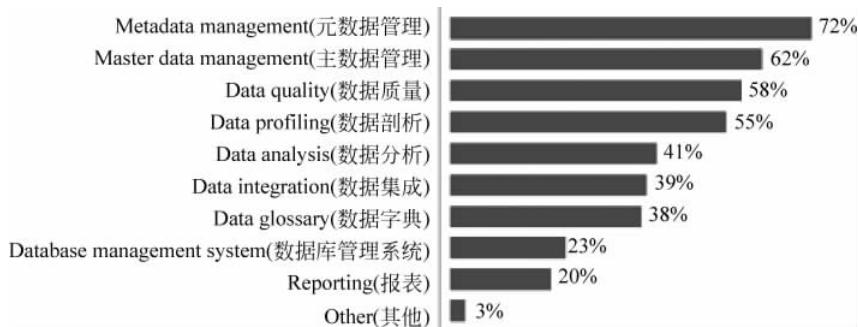


图 3-15 数据治理中的关键 IT 技术

在国内,习惯将元数据、主数据、数据质量以及数据标准等几部分功能统一形成主数据管理平台并作定制化开发,通过一站式数据管理提升用户体验。因此,主数据管理平台(MDM)是数据治理最重要的工具,能够为数据治理建立一套基础数据资料,存储治理范围内的数据定义、负责人、来源、转换关系、目标、质量等级、依赖关系、安全权限等,这些基础数据对于商业整合、数据质量、可审计性等数据治理目标的实现至关重要。同时,主数据管理平台通过定制化的流程对上述基础数据的创建、清洗、存储、更新、分发、删除的全生命周期进行管控,确保基础数据的完整、准确、一致。

作为实施数据治理的核心 IT 技术,优秀的主数据管理平台将为数据质量、数据集成等技术的实施,以及数据治理目标的最终实现奠定坚实的基础。

不同的主数据管理平台和解决方案虽然都能够实现主数据的整合、标准化、存储和分发等基本功能,但在产品特性、产品实施复杂度、系统兼容性、可扩展性方面仍存在许多差异。企业需根据严密的流程来对 MDM 的产品进行选型。在选择 MDM 系统时,企业一定要充分做好自身的调研工作,例如:企业高管层需要的是什么?数据治理的难点是什么?具体的业务难点是什么?当前的 IT 架构和未来的 IT 规划是什么?否则,盲目按照所谓的同行业可借鉴的经验进行评估与选型,只会让企业更加无所适从。

大部分企业都比较接受“成熟套装软件+专业的实施团队+适当的个性化定制开发”的解决方案。在这种情况下,根据企业管理特点和建设目标,个性化的开发是必要的,但成熟软件与开发的比例越大,项目风险越小。对于某些重点行业,如机械、电子、冶金、汽配、化工等行业,一些实施经验丰富的厂商已经形成了一定的行业解决方案。这些行业解决方案不仅满足了企业标准应用,也对行业化应用进行了提取和升华,基本能够满足某一行业的个性

化应用要求,大大降低了实施的风险。例如,北京三维天地科技有限公司所提供的能源行业主数据解决方案和机械行业主数据解决方案,已经帮助该行业中的领头羊企业大大提升了主数据管理能力,成为数据治理方面的优秀范例。

在主数据管理系统的选型中,应该兼顾以下原则。

- **信息集成原则:** MDM 作为数据集成的平台,兼容性是产品选型时需要考虑的关键因素。主数据产品不但需要与企业现有应用软件实现流畅的数据共享,更需要支持企业未来的 IT 规划。大多数时候,MDM 产品购买者容易受到他们倾向的软件厂商的影响,或者受到他们正在使用的数据管理或企业应用软件厂商的制约,但随着企业应用的不断丰富,依赖某一技术框架下的主数据产品反而会影响数据集成的效率。
- **最终衡量标准原则:** 衡量 MDM 选型工作质量好坏的最终标准是看 MDM 在企业的推广应用中能否取得成功。成功是各项工作因素的综合表现,包括主数据体系的方方面面。因此,凡是与选型相关的因素都要尽可能考虑进去。
- **功能与技术并重原则:** 功能完善、技术先进的软件是不多的,技术过于先进的软件,不一定适合企业。不成熟的技术可能会被慢慢淘汰,造成未来系统升级、扩展的困难,要准确地把握二者的结合点。
- **服务至上原则:** MDM 软件的管理对象是面向整个 IT 架构及其应用,不同的 IT 架构对 MDM 产品的要求差异很大,因此,再好的产品也需要好的管理咨询、定制开发、日常维护等服务才能更好地满足用户差异化的需求。信息系统选型不单是选产品,同时也是为公司未来发展选择一个具有实力的合作伙伴,以促进企业的成长和壮大。

大多数时候,MDM 产品购买者容易受到软件厂商故意降低初始报价的销售策略的影响,或者受到他们正在使用的数据管理或企业应用软件厂商的制约,这些都会造成选型过程的草率,为以后的系统应用埋下后患。信息系统选型是一个系统化的工程,应从以下几个方面进行综合考虑。

3.4.1 软件公司的行业实践

软件公司的行业实践是一个成熟软件公司最大的价值体现。对于一个成熟的主数据管理产品,应该是经过很多企业使用,且有一定发展史的。这样产品的软件公司才会在长期的客户使用中,不断将优秀企业的管理思想、业务流程等与软件产品的进行结合,持续地优化更新产品。

软件产品并不是单纯的技术体系,更重要的是内置先进的管理思想和管理框架。因此,只有行业实践丰富的软件公司,才有可能在项目实施过程中依托其丰富的管理经验,帮助企业提升管理水平。

对于行业实践少的产品与软件公司,由于经验不足,许多遇到的问题需要不断摸索解决,为此企业将担负巨大的风险。

3.4.2 产品特性

从目前的软件行业来看,同规模同行业的软件公司的产品相对是同质化的,但非同规模同行业的产品相差就会很大。这不是体现在表面所看到的功能上,而是体现在系统内在的框架和逻辑关系,以及管理思想和应用上。

但是企业往往不能在有限的时间内深入了解产品,准确判断产品和企业自身的匹配程度。保险起见,就要优先选择那些大企业都在使用的软件产品。

此外,虽然很多领先的信息管理技术都源于国外,但是对于主数据管理这一新概念,国内的普及非常迅速,在产品上已经迅速跟进国际水准,甚至在一些特色应用上超过国际水平。国外基于成熟的产品体系开发的产品普遍架构庞大而复杂,所以对新技术的采用普遍采取比较保守的策略,实施过程烦琐。因此,在选择产品时不能一味地追求国际声誉,还需要结合成功实践,考察企业在主数据管理方面的真正实力。

3.4.3 软件公司的实力

软件公司的实力包括公司的规模、成立时间、团队构成等多个因素。软件公司的规模大,成立时间长,说明这个公司经历了中国企业发展期,其产品是得到业界认可的,能够满足企业的管理需要。

3.4.4 软件公司的实施

信息系统的实施不仅是技术实施,更重要的是带来的组织架构、业务流程的变革。系统实施的成功与否,需要站在管理的角度来评判。邀请具有丰富管理经验的团队来实施,才能利用实施团队的经验来提升企业的管理能力和水平。

3.4.5 软件的价格

通过以上的分析可以看出,IT项目最根本的目的是为了实现企业的战略目标,与企业管理思想相匹配,选择的是适合企业发展的信息系统,而不能一味地追求低价中标。

信息系统的总价中包括非常多的的因素,包括有形的和无形的,如新增组织费用、个性化需求开发、与其他系统接口、后续开发成本和解决方案的价值等。因此,不能仅用有形成本作为唯一的考虑因素和谈判依据,而是要靠竞争和多角度的谈判策略来控制总体成本。