



破解数据化 运营困境

建立价值导向的企业级数据字典体系

主题 专家



张玉明

IBM 副合伙人，
IBM Consulting 金融核心锐变团队
数据转型与创新负责人
zhangyum@cn.ibm.com

张玉明先生是 IBM 副合伙人，IBM Consulting 金融核心锐变团队数据转型与创新负责人，主要关注于金融行业数据中台、大数据战略与分析咨询、数字化战略转型、数据架构规划、企业级数据治理及数据资产管理解决方案等方面。他领导过多个金融企业的大数据分析与人工智能技术应用、大数据应用架构、数据运营体系、业务咨询乃至系统实施项目。

韩玲玲

IBM Consulting 金融核心锐变团队
数据转型与创新咨询经理
hllhanbj@cn.ibm.com

韩玲玲女士是 IBM Consulting 金融核心锐变团队的咨询经理，她拥有超过 16 年丰富的数据咨询与实施项目经验，在数据战略规划、数据治理管控、数据资产管理等领域具备成熟的实践背景和专业知识。她服务于多个重要行业，如银行、保险和制造业等，为客户提供了一系列高质量的咨询服务和解决方案，涵盖企业级数据战略、数据治理、数据资产等领域的解决方案。

侯志伟

IBM Consulting 金融核心锐变团队
数据转型与创新咨询经理
hzwhou@cn.ibm.com

侯志伟先生是 IBM Consulting 金融核心锐变团队的咨询经理，他具备超过 18 年丰富的数据咨询与实施项目经验，他在数据治理、IT 战略规划、数据战略规划、数据中台建设以及数字化战略转型等领域均有着深厚的实践背景。凭借卓越的专业知识和经验，为金融业和制造业等大型机构提供企业级数字化转型策略、数据治理方案以及数据战略规划等全方位的解决方案。

张微

IBM Consulting 金融核心锐变团队
数据转型与创新咨询经理
Wei.zhw.Zhang@ibm.com

张微女士是 IBM Consulting 金融核心锐变团队的咨询经理。她拥有 15 年以上的数据咨询与实施项目经验。在数据质量、数据标准、数据架构、数据模型、数据资产盘点等方面均拥有成熟的实践经验，为银行、证券、电信等大型机构提供数据治理、数据集市等解决方案。

程倾

IBM Consulting 金融核心锐变团队
数据转型与创新高级咨询顾问
Qing.Cheng@ibm.com

程倾女士是 IBM Consulting 金融核心锐变团队的高级咨询顾问。她拥有 5 年以上的数据咨询与实施项目经验。在数据标准、主数据、数据安全等方面均拥有成熟的实践经验，为银行、券商、汽车等大型机构提供企业级的数字化转型、数据治理、AI 智能应用等解决方案。

王莉

IBM 商业价值研究院
高级咨询经理
gbswangl@cn.ibm.com

王莉女士是 IBM 商业价值研究院的高级咨询经理，担任 IBM 全球高管调研项目和对标分析项目的大中华区项目负责人。她拥有 17 年以上的管理咨询和管理研究经验。王莉女士目前所关注的研究领域包括数字化转型、无边界企业、人工智能等，旨在帮助各行业客户创造新的商业机会、发现和传递价值。



摘要

■ 难以挣脱的“数据旋涡”

数据管理挑战严峻，众多数据治理措施仍难以有效根除顽疾，陷入漩涡。

■ 数据字典建设之“营造法式”

数据字典建设的关键策略：顶层设计，精益求精，彰显价值，反哺业务。

■ 未来行动指南

数据字典成为企业数字化转型关键：全员参与、融合 AI 技术、获得高层支持。

前言

随着数字化转型的深入发展，数据作为新生产要素有望成为企业增长和创新的核​​心驱动力，其重要性被提到了空前高度。2023 年 8 月，财政部发布了《企业数据资源相关会计处理暂行规定》，¹ 为企业如何合理有效地管理和利用数据资源提供了指导和规范，强调了清晰界定和准确计量数据资源的重要性。2024 年 1 月，国家数据局启动了《“数据要素 ×”三年行动计划(2024—2026 年)》，² 旨在推动数据要素市场的发展，从国家层面对企业如何利用数据进行了深入规划和布局，强调了发挥各领域数据要素乘数效应，推动数据要素应用的广度和深度。在这样的大环境下，各领域各行业如何打造自身的数据要素基础，尤其是构筑规范化、标准化的数据要素底座将成为企业数据管理的首要任务，其重要性不言而喻。

近年来，越来越多的大型机构开始意识到企业级数据字典的重要性，并相继开展企业级数据字典的建设和落地工作，这也成为数据要素建设的基础性工作。IBM 认为，企业级数据字典是“企业范围内数据项的集合（去重），覆盖企业内各业务和技术活动中涉及的所有数据信息，通过对数据信息的表达、格式及定义的一致约定，实现对数据标准化、结构化的描述，为企业的业务、技术人员提供统一的数据项规范”。相比于传统 IT 开发过程中使用的数据字典，企业级数据字典将数据标准化工作范围拓广到全领域的数据，并且更关注跨业务领域、跨应用系统间的数据共享，以及业务与技术间的对数据理解和实现的一致性。

在数据治理领域，数据字典曾被视为一种强化管理的工具。然而，随着大模型技术的日臻成熟，数据字典的角色已悄然转变。企业级数据字典对企业内部数据元素的标准化定义和描述，构建起一套统一的数据语言和知识体系，几乎涵盖了企业内所有业务相关的知识信息。它不再仅仅是管控数据的技术手段，而是成为企业构建专业领域知识库的关键能力，能够助力企业在智能化的大模型应用中脱颖而出。

挑战：难以挣脱的“数据旋涡”

在数据驱动的商业环境中，企业面临着日益严峻的数据管理挑战。数据的处理和利用效率已成为衡量企业竞争力的新标杆。然而，数据生命周期中的采集、存储、分析和应用环节常常问题重重，让企业应接不暇。如果企业仅仅依赖临时性的应对措施，例如“搞运动”式的数据治理专项整改，虽能暂时缓解眼前之困，却无法从根本上解决问题，这将严重制约企业和组织的发展。

数据碎片化是企业普遍面临的问题。尽管企业每日生成海量数据，但这些宝贵的信息资源往往散落在各个部门和系统之中，从而形成信息孤岛。企业内部信息壁垒难以打破，难以实现数据价值的最大化。

此外，数据质量问题亦不容忽视。在许多情况下，由于数据采集和处理的非标准化，大量噪音和错误数据随之产生，严重影响了数据的准确性和可靠性。尽管企业坐拥庞大的数据资源，如何从中提炼出具有价值的信息，并将这些信息转化为切实可行的洞察和行动，仍然是一项极具挑战性的任务。IBM 商业价值研究院的调研表明，有 40% 的全球 CEO 和 38% 的大中华区 CEO 认同，他们的企业在很大程度上面临的数据挑战是难以识别有意义的洞察。³



策略：数据字典之“营造法式”

美国管理学家弗雷德里克·泰勒的科学管理思想体系的核心就是标准化。当代诸如国际标准化组织（ISO）、良好操作规范（GMP）等标准化管理体系，沿用的基本也是泰勒的标准化思想和工作方法。⁴ 虽然科学管理的标准化思想诞生于 100 年前，但仍然对今天的企业推动数字化转型具有重要的借鉴意义。

企业级数据字典是标准化思想在企业数字化转型过程中的一种科学实践。数据字典通过提供统一的数据定义、格式和标准，确保整个组织内部数据的一致性，这对于跨部门协作和系统集成至关重要。此外，数据字典作为一个数据资产全域目录，使得数据更容易被发现和重用，促进了数据驱动的创新和业务增长。

但在具体实践中，企业级数据字典的构建与应用常常陷入认识误区。首先，数据字典的重要性尚未得到充分认识，往往被视为技术部门的专属责任，缺乏高层支持。如果没有业务部门的积极参与，数据字典的建设将变得异常艰难。其次，数据字典的落地实施也面临挑战。它要求改变现有的产品研发流程，而这往往容易遭到业务用户和科技人员的抵触，他们对于遵循标准化要求常常持保留态度。

基于对行业领先实践和项目实施经验的深入分析，IBM 系统地提炼了一套高效的企业级数据字典工作方法及其应用策略。通过综合考量宏观和微观两个层面的因素，我们提出了三项关键策略：

提高站位, 收益导向

企业级数据字典超越了单一系统的局限，与传统数据标准有着明显的区别。它融合了企业内多个业务领域和信息系统的数据库资源，其应用范围横跨整个集团或企业分支机构，而不仅限于个别部门或局部系统的使用。这种数据字典能够支持跨部门、跨流程、跨系统对数据的使用，实现了数据的全方位整合和共享。相较于数据标准，企业级数据字典的覆盖范围更广泛、更全面，它是对数据标准的深化和延展，提供了一个更为丰富和具体的语义描述和管理框架。

业务收益的绩效归因

企业级数据字典的建设不仅是技术问题，更是涉及企业战略、管理、文化和流程的全面变革。因此，企业高层领导的重视和参与至关重要。而企业高管层的重视往往建立在数据字典工作对企业的业务价值提升的基础之上。为此，IBM 根据多年数据治理的实践经验，总结出数据字典工作提升业务价值的绩效归因指标，示例如下：

表 1

建设数据字典所带来的业务价值示例

数据字典赋能的典型业务领域	业务收益指标	归因分析	预估收益
客户营销	1. 客户满意度 2. 客户流失率 3. 销售收入	数据字典支持企业构建更全面的客户视图，通过整合和分析客户数据，企业可以更深入地了解客户需求，提供个性化的产品和服务，增加精准营销度，提高客户满意度和忠诚度，从而增加重复购买和口碑推荐，直接或间接帮助企业提高营业收入。	例如：帮助企业提高销售收入 2%~5%
财务管理	1. 财务对账时间	通过数据字典统一科目体系和各类数据的标准化，使得财务人员在核对账目时能够快速追踪数据来源，验证数据的准确性和可靠性，从而减少了对账过程中的混淆和错误。 数据字典有助于企业实施统一的数据整合和转换流程，减少了在财务对账过程中对数据进行手工调整和转换的需要，从而降低了时间成本。帮助企业缩短对账周期，提高财务管理的整体效率和准确性，实现业财一体化。	例如：财务对账时间从每月 N 天缩短到每月 M 小时 例如：年报发布时间提前 N 天 ⁵
产品研发	1. 产品从需求到交付的周期 2. 项目成本超支的概率	通过数据字典的使用，业务需求能够在早期阶段变得更为明确和详细，这有助于减少在开发过程中因需求不明确导致的变更次数。减少返工和修正的次数，可以缩短需求分析和确认的时间。通过减少需求变更和返工，可以降低项目成本超支的风险。从而缩短整个产品从需求到交付的周期。	例如：银行产品系统的平均上线周期从原来的 6 个月缩短为 5 个月
风险与合规	1.(集团) 风险报告及时率和准确率 2. 企业因违反监管要求而受到的罚款金额	数据字典在(集团) 风险管控方面的作用主要体现在提高信息透明度、确保数据一致性、加强风险识别和监控等方面，有助于降低集团风险管理的成本，包括人力成本、时间成本。同时降低因数据错误或不一致导致的合规风险和潜在合规处罚。	例如：在一年内，企业因违反监管规定而收到的罚款减少 XX 万元
数据管理	1. 数据资产利用率 2. 分析决策及时率和准确率	数据字典促进了企业内部数据资源的更广泛利用，提高了数据的整体价值。有助于提高数据分析的准确性，减少因数据误解或误用导致的错误。提高数据分析师为业务决策提供支持的效率。	例如：数据资产利用率可提升 20%~40%
	3. 数据质量问题数量 4. 数据治理活动所耗费成本 (时间 × 人力)	数据字典降低了部门间责任推诿的事件数量，提高了数据溯源的效率，缩短了溯源所需的时间，减少了数据转换和整合的工作量，加快了数据治理问题的解决速度，有助于降低数据治理的成本，包括人力成本、时间成本等。提高了企业内部对数据治理工作的满意度，包括员工满意度和管理层满意度。	例如：数据治理活动成本(人月)降低 30%

以上绩效归因指标证明了数据字典工作对业务收益增加所带来的贡献和价值，也论证了数据字典工作的有效性和成本持续投入的合理性。有了高层领导和相关利益干系人的支持（包括资金支持、资源支持、政策支持等），数据字典工作的开展才会更加顺利。

人人参与，人人有责

企业级数据字典的编制是跨部门的协作成果，尤其需要业务部门的深度参与。现实中，技术或数据管理部门牵头、业务部门辅助的模式常因缺乏业务部门的充分投入而效果不佳。建议由公司“一把手”直接督导，以业务部门为核心，使其提供专业知识输入，确保数据字典的准确性和实用性。同时，还需要明确各部门的职责划分，确定每个数据项的责任主体，为数据的归属提供坚实的依据，确保数据字典的实施有据可依。通过这种协同作业，不仅能够提升数据字典的质量，还能够促进跨部门的沟通与协作，为企业信息系统的稳健发展奠定坚实基础。

确保数据字典的权威性

企业级数据字典不应仅存于纸上，而应深入融合到企业的信息系统与业务流程中，充分发挥其应用潜力。为此，数据字典的发布质量与权威性显得尤为重要。通常情况下，企业级数据字典的规模可达数万级别，涵盖众多参与者及复杂的定义规范，牵涉到大量的细节问题。若完全依赖人力投入，不仅耗时耗力，而且难以保证质量。因此，有必要借助自动化、智能化工具来提升数据字典的精确度、实用性与时效性。例如，依据数据项的名称自动生成数据的分类、业务定义等属性，同时实施规范化检测以提高数据字典编制的质量，并对不符合规范的内容提供相应的提示。

此外，数据字典的发布应经过广泛的意见征集，并严格遵循审批流程，确保其全面性和适用性。建立清晰的数据字典版本管理机制，是维护其稳定性和可追溯性的关键。同时，针对现有和新建系统，采取差异化的实施策略，确保数据字典能够在不同环境中有效落地。

案例

某股份制银行在开展企业级数据字典编制工作时，通过公司董事长领导挂帅、业务部门牵头、科技部门配合、数据管理部门复核的协同方式，积极调动全行超过 20 个业务部门参与数据字典的编制工作。其中，董事长直接过问数据字典建设进程和落地进展；业务部门负责对数据中文名称、业务定义、业务规则等进行初稿编制；科技部门负责对数据类型、长度等内容进行定义；数据管理部门负责对数据字典整体内容进行复核、完善和确认。各部门充分发挥各自的专业领域知识和经验，共同努力、高效配合，对全行关键系统进行多轮滚动整合和萃取，最终形成了覆盖对公、零售、同业等业务领域的万余项企业级数据字典。⁶

案例

纵横平台智能匹配工具（Smart Data Mapping, 以下简称 SDM）是 IBM 以通用数据字典为核心，利用 NLP 自然语言处理算法、深度学习模型等 AI 技术，通过内置的 IBM 行业数据字典规范与预训练的 AI 算法模型，进行数据匹配和自动映射的工具。IBM 曾使用 SDM 的智能映射匹配、词根分析等功能，为某股份制银行完成数万项数据字典的编制工作。其中包括：1) 自动化检核字典项属性值的内容，完成对共计超过 30 万个字典项属性值的内容进行质量检核，相较于纯人工质检节省了 80% 时间；2) 自动生成字典项英文名称与简称，大幅减少人工逐项编写的时间成本，同时保障词根英文翻译的一致性。

精益求精,做到“统、实、严”

在编制数据字典的过程中,确立一套统一的标准和规范是至关重要的。这意味着对数据字典中的每个元素,包括业务属性、技术属性和管理属性,都需要有一套明确的编制规范。业务属性包括数据字典项的中英文名称、业务含义、业务规则和引用出处等,技术属性涉及数据类型、格式、长度和约束等,而管理属性则包括数据字典项的安全级别、责任归属和生命周期等。对于这些属性的每个元素,都应制定清晰的定义和描述规则,以确保数据字典的一致性、准确性和易用性。

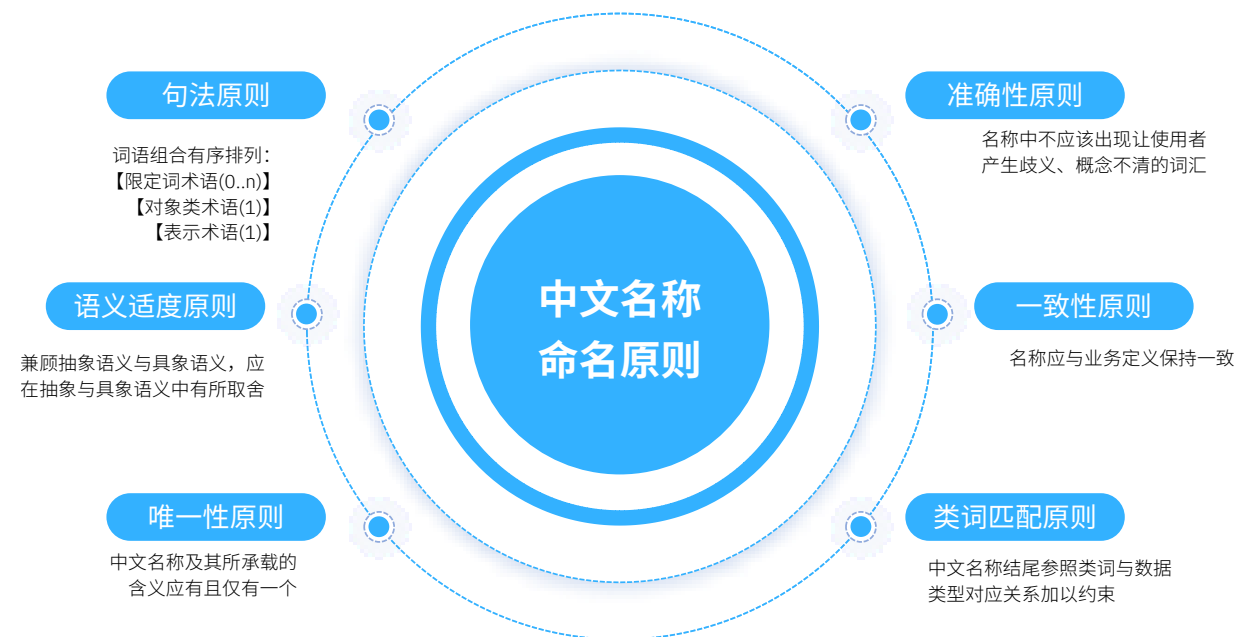
字典项命名规范

字典项的命名包括中文命名和英文命名,需要遵循一系列的命名原则,确保命名的合理性与一致性。IBM总结了字典项中文名称的六大命名原则(见图1):

- 句法原则,即规范中文命名的基本结构及词语的排列顺序。
- 准确性原则,即确保中文表达的信息准确,避免含混、歧义等情况的发生。
- 语义适度原则,即兼顾抽象语义与具象语义的平衡,避免名称表义过于宽泛或者过于细节。
- 一致性原则,确保名称与业务定义间的内涵一致性。
- 唯一性原则,中文名称的全局唯一,所承载的业务含义信息有且仅有一个。
- 类词匹配原则,中文名称结尾参照类词语数据类型对应关系加以约束,避免随意造词。

图1

数据字典项中文命名原则



以下是具体实践中参考国家标准⁷使用的字典项中文名称句法原则示例（见图 2）。

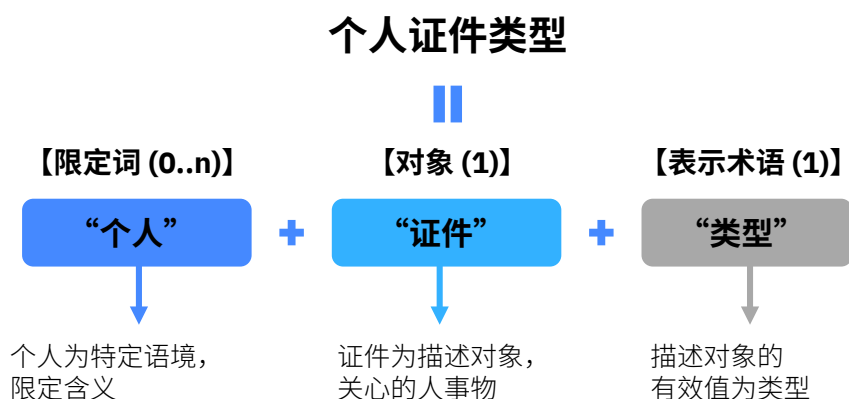
中文名称的句法原则要求如下：**【限定词术语 (0..n)】**、**【对象类术语 (1)】**、**【表示术语 (1)】**。

- 句法原则 1：当没有限定词时，名称中应当有且仅有一个对象类术语且置于名称中的最左侧 / 首位。
- 句法原则 2：名称中应有且仅有一个表示术语且应置于最右侧 / 末尾。
- 句法原则 3：限定词术语应置于所限定部分的前面，不应使用限定词的顺序区别名称。限定词术语是可选的。

值得注意的是英文名称和简称关乎到未来数据字典项的落地实现，既要确保全局唯一，又要确保词义和语义的准确，单纯依靠人工编制、复核是难以实现效率与质量并重的目标。所以有必要利用词根分析工具，提升同词同译的规范化效率，制定英文缩写重复情况的处理方法与步骤，以确保英文简称的结构化和英文缩写的范式化。

图 2

中文名称的句法示例



字典定义，既是重点也是难点

业务定义和业务规则为数据字典项提供了必要的业务上下文，使得数据不仅仅是数字和文字的组合，而是具有明确含义和用途的业务资产。它们是数据理解和正确应用的基础，确保了所有用户对数据项的理解一致，减少了误解和混淆。定义业务定义和业务规则时，语言的使用需简洁明了且精准，既要符合描述规范，又要易于理解，以准确反映实际业务含义。需明确指出业务定义中业务对象和适用业务范围，并应与字典项名称所指范围一致，业务场景的定义应清晰明确，特殊词汇应给予适当的解释，业务用途的描述必须符合业务实际，规则的描述亦需做到准确严谨，口径统一。

字典编制的“实用性”理念

数据字典编制按照各类数据的表现形式，对数据类型进行统一规范，定义数据更细粒度的规范参考，包括数据长度、精度、代码取值等，保证数据格式的统一。例如，针对文本类、编码类数据长度，涉及中文字符时，需要结合各系统字符集规则，从逻辑上规定可包含字母、数字、汉字字符的数量，以确保数据字典贯标的可操作性。

为了确保数据字典的科学性和实用性，其制定不仅需满足企业内部的业务和系统使用需求，还应充分借鉴外部相关的法律法规、行业标准、监管要求和最佳实践。这样的做法旨在避免数据字典内容出现歧义或误导，使其成为可靠的信息资源。

案例

某股份制银行在推进企业级数据字典的实施过程中，采取了“明确主体认责、实施分层授权、确保源头落地、灵活分类施策、保持优化迭代”的落地策略，确保数据字典能够高效、准确地应用。为实现这一目标，该银行构建了三层数据字典落地防御体系，强化了总行与各分行的标杆责任，使每个部门都能够积极履行数据字典的主体责任，形成合力共同推动其落地实施。同时，根据数据的覆盖范围和应用场景，将数据字典划分为通用类和专有类，并据此进行分层授权管理。

在新系统建设中，坚持“应落尽落”原则，确保系统能够严格按照标准采集和生成数据。对于已有的存量系统，该银行则结合项目建设的契机，有计划、有步骤地推进贯标工作。此外，各部门还制定了现有系统的贯标总体规划，确保贯标工作的有序开展。整个贯标过程从源系统开始，带动相关联的系统参与贯标。同时，总行及各分行制定了自我检查方案，并由数据管理部门负责检查和评估，确保数据字典的落地效果符合预期。审计部门则作为独立的第三方，负责执行日常的或专门的审计工作，确保整个过程的合规性和准确性。

字典落地，“严”字当头

数据字典的落地执行与监督过程需保持高度的责任感和纪律性。对违反数据字典的行为要进行及时的纠正和处理，确保数据字典的权威性和有效性。

明确责任主体、夯实责任。从业务、技术和管理维度，明确数据字典的责任主体类型，为有效的数据管理提供支持和保障。根据数据属性信息，结合部门职责建立数据责任认定原则，为数据的归属提供有力依据。

制定数据字典贯标总体策略。结合流程制度、应用软件开发工艺，明确谁（Who）来落标、落什么（What）内容、何时（When）落标、在哪（Where）落标、如何（How）落标的总体策略，推动数据字典在企业中的有效落地。

赋能应用，彰显字典价值

在企业环境中，企业级数据字典的规模往往达到万量级，这庞大的知识库如何有效地应用于企业内部的各个业务环节，以提升业务效率并降低成本，成为了一个关键问题。为了充分发挥数据字典的价值，我们建议针对企业内不同类型用户的核心需求，在以下典型场景中加强数据字典与业务需求的融合。

嵌入业务场景的智能助手

数据字典是企业内部最有价值业务信息的集合。在生成式人工智能技术的助力下，数据字典可以更加轻松和便捷地融入到员工日常工作的工具和平台中，例如 ERP 系统、客户关系管理系统（CRM）、产品管理系统（PMS）、管理驾驶舱、风险看板、业务研发一体化平台等。



这样的集成旨在提升数据字典的普及度和员工接触面，确保数据字典中蕴含的高价值知识能够辅助员工高效地完成各项任务。此外，数据字典的分类体系还能够为营销人员提供设计精准客户画像的参考，辅助产品负责人构建更加全面和灵活的产品配置条件及属性设置。

通过这种方式，数据字典不仅增强了企业数据资产的管理效率，还提升了员工的工作效率和业务决策的质量，为企业的数字化转型和持续创新提供了坚实的支撑。

业务需求标准化

业务需求的标准化是许多组织在产品快速设计和研发过程中面临的核心挑战之一。一方面，业务需求人员往往难以提出与 IT 开发实现相符合的需求输入；另一方面，业务需求对场景和用例的设计往往过于笼统，常常需要在后续的迭代过程中逐步完善，这大大降低了产品设计和研发的效率。

如果在业务需求编制过程中引入数据字典的使用，则可以显著提升业务需求的标准化程度。数据字典能够提供业务流程和场景所依赖的详细数据信息，使用户能够更深入地理解业务需求及其上下文。这有助于业务需求方在业务需求中建立与实际业务流程相一致的场景和用例。同时，数据字典为业务术语和定义提供了统一的参考，有助于所有用户对业务需求的理解保持一致。这支持用户在业务需求分析和设计中建立一致和可靠的标准，从而提高业务需求的可理解性、可维护性和可执行性。

例如，一些领先的金融机构早已经在业务需求模板中嵌入对数据字典的强制引用，其目的是为了加强标准化数据的贯标，增强业务需求的一致性和准确性。看似增加了管控节点，但实质上提升了产品设计和研发的效率，使得组织能够更快地推出满足市场需求的产品。

支持快而准的数据分析和决策

当下许多先进组织在企业内部及其分支机构设立了数据分析岗位，致力于深入的数据分析和为决策提供支持。然而，在日常工作实践中，数据分析师常常面临诸如“我要的数据在哪里？”、“多个系统都有这个数据，哪个才是准确的？”等挑战。数据字典的引入能够有效减轻数据分析师的工作负担。

首先，数据字典作为一个关键的数据资产目录，允许数据分析师迅速定位到所需的数据，并指导他们如何访问这些数据，从而大幅提升数据检索的效率。其次，数据字典详尽地记录了数据的业务规则和逻辑，包括计算公式、取值范围和合法性检查等。这些信息帮助数据分析师深入理解数据的生成和处理机制，确保他们所依据的业务逻辑是准确无误的。第三，数据字典揭示了数据之间的复杂关系和依赖，如引用和被引用关系。这有助于数据分析师识别并分析数据之间的相互作用，从而做出更为全面和精准的决策。

例如，一些领先的组织正在建立数据字典标签视图体系，旨在实现针对各类数据应用场景的数据视图，从而实现从“人找数据”到“数据找人”的转变。这种体系不仅提高了数据分析师的工作效率，还增强了数据分析的深度和广度，为企业的数据驱动决策提供了强有力的支持。

降低数据治理活动的复杂度

数据治理活动已经成为每一个先进数字化企业的共同认知和基本配置。然而，在实际推动数据治理的过程中，由于责任体系不明确和系统架构复杂多变，数据治理的有效性往往受到影响。数据字典的引入能够显著降低数据治理的复杂性。

首先，数据字典明确了数据元素的所有者和责任部门，使用户能够在数据治理活动中建立清晰的数据管理和维护责任，从而提高执行效率，并解决部门间责任推诿和整改落实困难的问题。

其次，数据字典记录了数据的系统来源、转换过程和使用情况，使用户能够追踪数据在整个数据生态系统中的流动。这有助于用户深入理解数据的来龙去脉，简化数据溯源和影响分析的工作流程。

第三，数据字典定义了数据元素的标准格式、度量精度和取值范围，使用户能够在数据治理活动中实施标准化策略，减少数据转换和整合的工作量，提高数据治理的效率和质量。

通过这些方式，数据字典不仅提升了数据治理的效率，还增强了数据治理的执行力和效果，为企业的数据资产管理 and 数字化转型提供了坚实的支持。

缓释数据合规风险

在近年来的数据合规领域，监管合规报送、数据安全、数据出境以及个人信息保护等方面已成为数据合规风险的关键焦点。⁸ 企业面临着政策密集发布、合规要求日益增多、监管手段不断升级等挑战。在这种背景下，数据字典的引入成为合规人员有效识别和管理合规风险的重要工具。

数据字典为合规人员提供了一个详尽的数据资产清单，其中包括个人数据和敏感数据的分类。借助数据字典，合规人员能够确保数据收集和处理活动符合各种数据保护法规，如《个人信息保护法》(PIPL)、《通用数据保护条例》(GDPR) 等。数据字典详细记录了数据的安全分类和合规要求，使得合规人员在业务需求提出和开发过程中，能够确保数据的使用严格遵守相关法规和政策，从而保护数据隐私和安全。

此外，数据字典还规定了每一项数据的质量标准和规则，包括完整性、准确性和一致性要求等。这些质量检核规则可以部署到数据质量检核平台中，帮助监管报送岗位人员在 EAST5.0 报送以及各地试点的一表通报送中，检查数据的准确性和可靠性，降低因数据错误或不一致导致的合规风险和潜在合规处罚。



数据字典建设行动指南

一些领先的组织已经将企业级数据字典视为数据驱动的核心工具。尽管在行业内对数据字典的理解和应用尚存在差异，但普遍认为，数据字典很可能成为推动企业数字化转型、数据要素市场化和创新发展的关键力量。

首先，构建全面的企业级数据字典需要企业内所有业务和科技部门的共同参与。鉴于所涉及的范围广泛、资源投入巨大，多数领先组织选择分阶段推进，逐步完善数据字典的规模、落地机制和配套的管控流程。

其次，虽然数据字典的建立相对容易，但保持其长期的有效性和更新却更具挑战性。因此，在建设初期就应着手构建动态保鲜机制，尤其是加强元数据的基础能力建设，以确保数据字典的时效性和准确性。

第三，企业应积极探索将 AI 技术与数据字典相结合，以提升数据字典的自动化和智能化水平，从而在数据字典的编制、贯标和业务赋能过程中提高效率。

最后，企业级数据字典不仅是数据资产的重要组成部分，甚至可以说是数据资产的代名词。在未来的数据资源入表实现路径中，数据字典将扮演至关重要的角色。例如，未经有效治理且无法带来实际经济收益的数据将无法被认定为资产。正如前文所指出的，数据字典在降低数据治理活动复杂度方面起着重要作用。因此，企业的首席执行官（CEO）或首席数据官（CDO），不应将数据字典的建设视为可有可无的辅助工作，而应从数据要素的视角，为这项工作提供必要的资源保障和投入。

关于专家洞察

专家洞察代表了思想领袖对具有新闻价值的业务和相关技术主题的观点和看法。这些洞察是根据与全球主要的主题专家的对话总结得出。要了解更多信息，请联系 IBM 商业价值研究院：ibvchina@cn.ibm.com

选对合作伙伴, 驾驭多变的世界

在 IBM，我们积极与客户协作，运用业务洞察和先进的研究方法与技术，帮助他们在瞬息万变的商业环境中保持独特的竞争优势。

IBM 商业价值研究院

IBM 商业价值研究院 (IBV) 站在技术与商业的交汇点，将行业智库、主要学者和主题专家的专业知识与全球研究和绩效数据相结合，针对公共与私营领域的关键议题提供可信的战略洞察。

访问 IBM 商业价值研究院中国网站，免费下载研究报告：<https://www.ibm.com/ibv/cn>

备注和参考资料

- 1 “财政部印发《企业数据资源相关会计处理暂行规定》”。20230822. https://www.gov.cn/lianbo/bumen/202308/content_6899425.htm
- 2 “国家数据局等部门关于印发《“数据要素×”三年行动计划（2024—2026年）》的通知”。20240104. https://www.cac.gov.cn/2024-01/05/c_1706119078060945.htm#
- 3 “AI时代的CEO决策力：运筹帷幄，决胜未来（大中华区洞察）”。IBM商业价值研究院。202308. <https://www.ibm.com/cn-zh/services/insights/c-suite-ceo-2023>
- 4 《企业精细化成本管理：核算、分析与管控》。侯立新、崔刚。2022年7月。
- 5 “详解马明哲全面数据化经营术，为何他家年报能提速22天？”。20200222. <https://news.hexun.com/2020-02-22/200375586.html>
- 6 “编制数据字典，深化数据治理——吕家进董事长为《兴业银行企业级数据字典》作序”。20230912. https://mp.weixin.qq.com/s/JU7n4Fk2HHHTuf__1fylkA
- 7 《GBT+18391.5-2009+ 信息技术 + 元数据注册系统（MDR）+ 第5部分命名和标识原则》。20091201. <https://std.samr.gov.cn/>
- 8 “共罚没超23亿！透视2023年银行业罚单：信贷、内控仍是重灾区”。20231213. <https://www.ifnews.com/news.html?aid=562427>

© Copyright IBM Corporation 2024

国际商业机器（中国）有限公司
北京市朝阳区金和东路20号院3号楼
正大中心南塔12层
邮编：100020

美国出品 | 2024年4月

IBM、IBM 徽标和 ibm.com 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。以下 Web 站点上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表：
ibm.com/legal/copytrade.shtml。

本档为自最初公布日期起的最新版本，IBM 可能随时对其进行更改。IBM 并不一定在开展业务的所有国家或地区提供所有产品或服务。

本档内的信息“按现状”提供，不附有任何种类的（无论是明示的还是默示的）保证，包括不附有关于适销性、适用于某种特定用途的任何保证以及非侵权的任何保证或条件。IBM 产品根据其提供时所依据的协议条款和条件获得保证。

本报告的目的仅为提供通用指南。它并不旨在代替详尽的研究或专业判断依据。由于使用本出版物对任何企业或个人所造成的损失，IBM 概不负责。

本档中使用的数据可能源自第三方，IBM 并未对其进行独立核实、验证或审查。此类数据的使用结果均为“按现状”提供，IBM 不作出任何明示或默示的声明或保证。

扫码关注 IBM 商业价值研究院



官网



微博



微信公众号



微信小程序

