

T/QBDA

青岛市大数据发展促进会团体标准

T/QBDA 3301—2023

数据资产价值与收益分配评价模型

Data asset value and revenue distribution evaluation model

2023 - 04 - 27 发布

2023 - 05 - 01 实施

青岛市大数据发展促进会 发布

目 次

前 言	II
引 言	III
数据资产价值与收益分配评价	1
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	2
5 概述	2
5.1 数据资产特征	2
5.2 数据资产价值与收益分配评价模型	2
5.2.1 评价模型目的	2
5.2.2 评价对象范围和粒度	2
5.2.3 评价模型思路和框架	2
6 维度及指标	3
6.1 数据质量评价	3
6.2 数据应用评价	4
6.3 数据变现量评价	5
6.4 数据收益分配评价	5
7 评价过程	5
7.1 指标体系配置	5
7.2 评价实施	6

前 言

本文件按照GB/T 1.1-2020给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由青岛市大数据发展促进会提出并归口。

本文件起草单位：青岛华通智能科技研究院有限公司、中国信息通信研究院、雄安新区智能城市创新联合会（雄安区块链实验室）、中国联合网络通信集团有限公司、北京易华录信息技术股份有限公司。

本文件主要起草人：孙明铭、苏青林、赵传启、张云翰、王显坤、赵丽鹏、张峻齐、梁明君、李志民、刘贺贺、衣军成、王春萌、闫树、李雨霏、邱梦媛、马闻达、张斯睿、孟宏伟、王伟、张淼、柴立身、于良伟、刘锐、潘思宇、付钰、王爽、林镇阳、王运宏、王淑华、孟菊园、王雅萱、潘紫艺、张莉莉、乔方通、单亦栋、张树成、李雪。

版权声明

根据国家有关法律法规规定，团体标准享有版权，未经许可禁止复制和销售。

未经合法授权，严禁任何单位和个人对标准出版物及相关工作文件进行复制、销售、传播和翻译出版，严禁任何单位或个人将国际标准的任何部分通过电子信息网络或制作成标准数据库用于传播。标准化服务机构和中介机构的检验、鉴定、认证、咨询、评估和培训等活动必须使用正版标准。标准化研究机构馆藏的标准文本必须通过合法渠道购买。

引言

数据可作为一种产生价值的资产，这一观点逐渐得到认可。在认识到数据可以资产化的同时，如何评价数据资产的价值、如何分配由数据资产产生的收益、如何进一步评估数据资产价值，从而实现将数据资产纳入企业资产负债表中的“资产”，成为数据资产化的一项重要工作。

本文件将从数据运营过程中的数据资产价值评价与收益分配评价入手，根据数据在运营过程中的使用与收益情况，量化数据质量、数据应用情况、变现量和收益分配比例，进而对数据资产的价值与收益分配进行评价。该评价结果可作为数据资产评估的依据。

数据资产价值与收益分配评价模型

1 范围

本文件规定了数据资产价值与收益分配评价模型的框架、维度及指标和评价过程。

本文件适用于数据运营过程中，对数据资产的价值与收益分配进行量化计算、价值评价；可作为数据资产评估的依据。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 35274-2017 信息安全技术 大数据服务安全能力要求

GB/T 35295-2017 信息安全 大数据 术语

GB/T 36344-2018 信息技术 数据质量评价指标

GB/T 38667-2020 信息技术 大数据 数据分类指南

3 术语和定义

GB/T 35274-2017、GB/T 35295-2017、GB/T 36344-2018和GB/T 38667-2020界定的以及下列术语和定义适用于本文件。

3.1 数据资产 data asset

合法拥有或者控制的，以数据为载体和表现形式，能进行计量，可以产生经济和社会价值的数据资源和数据产品。

注：改写GB/T 37550-2019，定义2.4、GB/T 40685-2021，定义3.1。

3.2 数据运营 data operation

经政府有关部门授权符合条件的运营单位，在构建安全可控开发环境基础上，挖掘社会应用场景需求，围绕需求依法合规进行数据汇聚、治理、加工处理，提供数据资源、产品与服务的相关行为。

3.3 数据产品 data product

由开展经济活动的自然人、法人或非法人组织，使用运营单位提供的的数据资源与服务，开发的有明确应用场景的新的数据资源或数据服务，以及对数据资源投入实质性加工和创新性劳动形成的数据衍生品，产品形式包括但不限于数据集、数据分析报告、数据可视化产品、数据指数、API 数据、加密数据等。

4 缩略语

下列缩略语适用于本文件。
API: 应用程序编程接口 (Application Programming Interface)
ETL: 数据提取、转换和加载 (Extract-Transform-Load)
ID: 身份标识号码 (Identity Document)

5 概述

5.1 数据资产特征

数据资产应以数据为载体和表现形式，由特定主体合法拥有或者运营，能持续发挥作用并且能带来直接或者间接经济利益，其价值能够被计量。

5.2 数据资产价值与收益分配评价模型

5.2.1 评价模型目的

该模型适用于数据运营过程中，根据数据的使用与收益情况，对数据质量、数据应用情况、变现量和收益分配比例进行量化，进而对数据资产的价值与收益分配进行评价。该评价结果可作为数据资产评估的依据。

5.2.2 评价对象范围和粒度

评价对象包括所有可运营的数据资产。
评价对象的粒度可以参照数据资源的分类目录的层次结构来划分。可选择数据分类目录的某一级下的全部数据作为评价对象，也可选择其中数个并列的层级进行评价比较。

5.2.3 评价模型思路 and 框架

数据资产价值评价包括数据质量评价、数据应用评价、数据变现量评价。
数据收益分配评价用于建立数据要素按数据提供者、运营者和加工者三方的价值贡献参与利益分配机制。

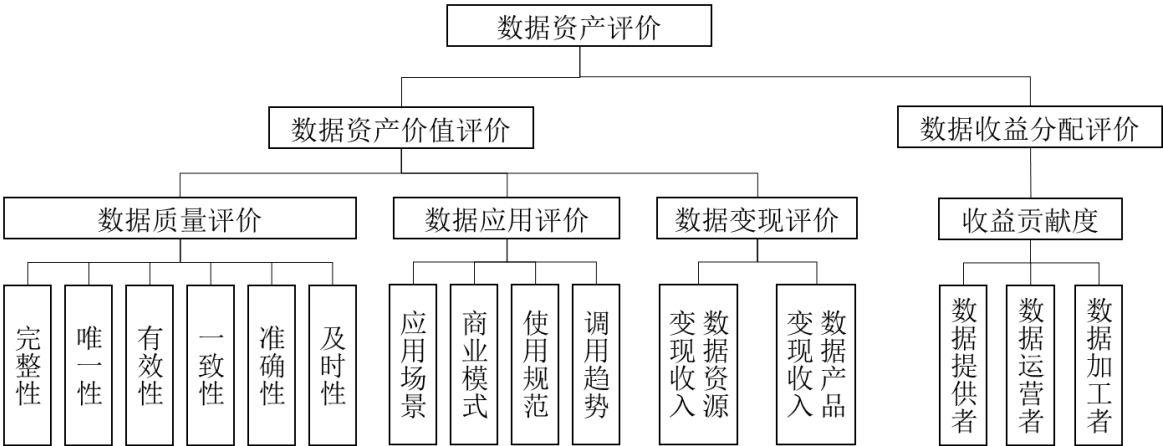


图 1 数据资产价值与收益分配评价模型框架

6 维度及指标

6.1 数据质量评价

数据质量是保证数据应用的基础，是获取数据价值的重要保障。根据目前业界对于数据质量的衡量标准，并结合数据治理经验，可基于数据完整性、唯一性、有效性、一致性、准确性和及时性 6 个维度来评估数据质量。依据以上指标，结合 ETL 等工具，针对不同的数据源做出定量的数据质量评估，也可根据实际情况，在评估执行中进行取舍。

——完整性：完整性指的是数据信息是否存在缺失的状况，数据缺失的情况可能是整个数据记录缺失，也可能是数据中某个字段信息的记录缺失。不完整的数据所能借鉴的价值就会大大降低，也是数据质量最为基础的一项评估标准。

——唯一性：针对某个数据项或某组数据，没有重复的数据值。值必须是唯一的如 ID 类数据。

——有效性：有效性指的是描述数据遵循预定的语法规则的程度，是否符合其定义，比如数据的类型、格式、取值范围等。

——一致性：一致性是指数据元素的类型和含义必须一致和清晰，如数据迁移或加工的前后数据表大小一致性对比。

——准确性：准确性是指数据记录的信息是否存在异常或错误。和一致性不一样，存在准确性问题的数据不仅仅只是规则上的不一致，更为常见的数据准确性错误就如乱码，其次异常的大或者小的数据也是不符合条件的数据。常见的准确性指标有：缺失值占比、错误值占比、异常值占比、抽样偏差、数据噪声。

——及时性：及时性是指数据从产生到可以查看的时间间隔，也叫数据的延时时长。对于数据更新频率的满足程度，针对用户对信息获取的时间及时性要求，确保数据及时更新。

表 1 数据质量评价指标

评估维度	满分	评估指标	指标说明	细分项满分
1.完整性	15	非空约束	数据记录内容中存在空值的情况。	5
		编码表是否缺失	对数据编码维表的描述是否为空的情况。	5
		字段是否缺失	重点字段内容是否为空值。	5
2.唯一性	15	主键是否唯一	主键数据 ID 必须为唯一的。	5
		编码表是否唯一	编码含义，编码规则必须唯一。	5
		数据是否重复	数据记录中是否存在重复内容的情况。	5
3.有效性	20	代码值域约束	是否符合数据值域约束，无超出情况。	4
		长度约束	符合数据长度要求。	4
		内容规范约束	符合数据类型约束要求。	4
		取值范围约束	符合数据范围的约束要求。	4
		标志位取值约束	重点标识数据必须满足要求。如：身份证号。	4
4.一致性	20	相同数据一致性	同一数据在不同位置存储或被不同应用或用户使用时，数据的一致情况。	7
		关联数据一致性	符合基础数据元组合的数据集的比例。	6

		逻辑一致性约束	跨表逻辑含义一致的数据定义及内容要一致。	7
5.准确性	15	取值准确约束	数据获取内容同来源要保障准确，无偏差。	5
		脏数据重复情况	数据中出现脏数据的占比情况。	5
		数据格式约束	数据是否符合标准格式校验要求的情况。	5
6.及时性	15	时段数据准确	日期范围的记录或者频率分布符合业务需求的情况。	5
		时点数据准确性	特定时点的记录数、频率分布或延迟时间符合业务需求的情况。	5
		数据时序准确性	数据时序的正确性。	5
合计	100	-	-	100

注1：细分项得分 = 细分项满分 × (A/B)。

注2：A=满足要求的数据个数，B=被评估的数据总数。

6.2 数据应用评价

数据应用评价是从数据资源的应用层面评价数据资产的价值。本评价模型从应用场景、商业模式、使用范围和数据调用趋势 4 个维度来评价数据的应用情况。

——应用场景：从数据目录和数据条目的被调用次数对数据的应用情况做综合评价。具体可按月、年计算；按模型或产品类别计算；按使用用途类别计算，使用用途包括数据资产涉及的行业、领域和区域等。

——商业模式：从产品或模型通过服务或场景产生的效益（包括社会效益和经济效益）的角度评价数据应用情况。该评价维度主要考虑数据服务的模式、赋能模式和金融模式等，例如金融科技产品给实体经济带来的融资或贷款额度。

——使用范围：从数据资源覆盖模型或产品的数量、广度的角度评价数据应用情况。具体可按数据目录被模型或产品调用的次数和频次计算。

——数据调用趋势：由于某一数据资源的使用在时间上可能存在集中性，为减弱在某一时段对数据应用评价的片面性，将数据目录和条目的历史调用情况纳入评价范畴。

表 2 数据应用评价指标

指标	指标说明	评价方法	满分	计算方法	示例说明
1.应用场景	数据目录和数据条目被调用次数，具体按月、年计算，被某个模型或产品调用次数	数据分析	25	按月、年计算；按模型或产品类别计算；按使用用途类别计算，使用用途包括数据资产涉及的行业、领域和区域等。	
2.商业模式	产品或模型通过服务或场景产生的效益，包括社会效益和经济效益	数据分析	25	主要考虑数据服务的模式、赋能模式和金融模式等。	比如金融科技产品给实体经济带来的融资或贷款额度。

3.使用范围	覆盖模型或产品的数量、广度	数据分析	15	按数据目录被模型或产品调用的次数和频次计算。	
4.数据调用趋势	数据目录和条目被调用趋势	历史数据分析	35	按数据目录和条目的历史调用量和年均调用量。	
合计			100	-	

6.3 数据变现量评价

数据资源或数据产品能够变现产生的收入能够直接反映数据资产的价值。在数据资产价值评价模型里，使用资源或产品在某一时间段内，按照合同或协议产生的变现收入和累计变现收入对该指标评分。

表 3 数据变现量评价指标

指标	指标说明	评价方法	满分	计算方法	示例说明
数据资源或者数据产品变现收入	资源或产品变现收入	合同或协议	100	合同或协议收入	
合计			100	-	

6.4 数据收益分配评价

数据收益分配评价用于判定数据流通过程中各参与方的收益分配比例。该评价可依据数据流通过程中，各参与方产生的收益贡献度进行收益分配。例如，根据流通过程中数据提供者（所有权和使用权）、数据运营者（运营权）、数据加工者（使用权）三方贡献度进行收益分配。贡献度可依照各参与方投入的成本作为评判依据。

表 4 数据收益分配评价指标

指标	指标说明	评价方法	满分	计算方法	示例说明
数据流通过程中产生的收益按贡献度分配	获得收益按照数据提供者（所有权和使用权）、数据运营者（运营权）、数据加工者（使用权）三方分配	按贡献度分配	100	协议	
合计			100	-	

7 评价过程

7.1 指标体系配置

根据评价目的和对象的具体情况，配置评价指标、确定指标计算方法、配置权重。

a) 配置评价指标

选择合适的评价指标及评价内容，可根据实际需要进行动态调整、扩展。

b) 确定指标计算方法

指标计算方法是指依照评价内容、用于检查和分析数据应用效果的技术方法，包括工具计算评价和人工计算评价两种方法。在条件允许的情况下，数据应用效果评价指标计算应采用工具计算评价的方法。

c) 配置指标权重

配置一维指标和二维指标的权重。指标的权重设置可以根据自身业务特性及上一评价周期的得分情况进行调整。

7.2 评价实施

在本规范的数据资产价值评价模型中包含多个评价维度和指标层级，首先需要确定数据质量、数据应用和数据变现量三个一维指标的权重，进而计算各维指标得分和总得分。评价所得的各一维指标得分和总得分，均可作为数据资产评估的标准依据。

a) 确定一维指标权重

数据价值评价的3个一维指标满分均为100分。在计算数据价值评价指标得分时，首先需要确定各一维指标的权重。3个一维指标的初始权重分别为：数据质量得分 X_1 占比30%，数据应用得分 X_2 占比30%，数据变现量得分 X_3 占比40%。由此可得数据资产价值评价得分 N 的计算公式为：

$$N = X_1 * 30\% + X_2 * 30\% + X_3 * 40\%$$

b) 计算一维和二维指标得分

根据数据资源情况对二维指标进行打分，或按既定规则计算其得分。将各一维指标下的二维指标得分累加，得到一维指标得分。对于包含 m 个二维指标的第 n 个一维指标，其得分为：

$$X_n = \sum_{i=1}^m A_i$$

以第2个一维指标“数据应用”为例，该指标包含4个二维指标。假定这4个二维指标得分依次为18、22、20、15，则该一维指标得分为75（ $X_2 = 18 + 22 + 20 + 15 = 75$ ）。

c) 计算指标总得分

将各一维指标得分，乘以相应权重后求和，所得数值即为数据价值评价指标最终总得分。

假定一项数据资源的三个一维指标得分为 $X_1 = 80$ ， $X_2 = 75$ ， $X_3 = 60$ ，则数据资产价值评价得分 $N = 80 * 30\% + 75 * 30\% + 60 * 40\% = 70.5$ 。