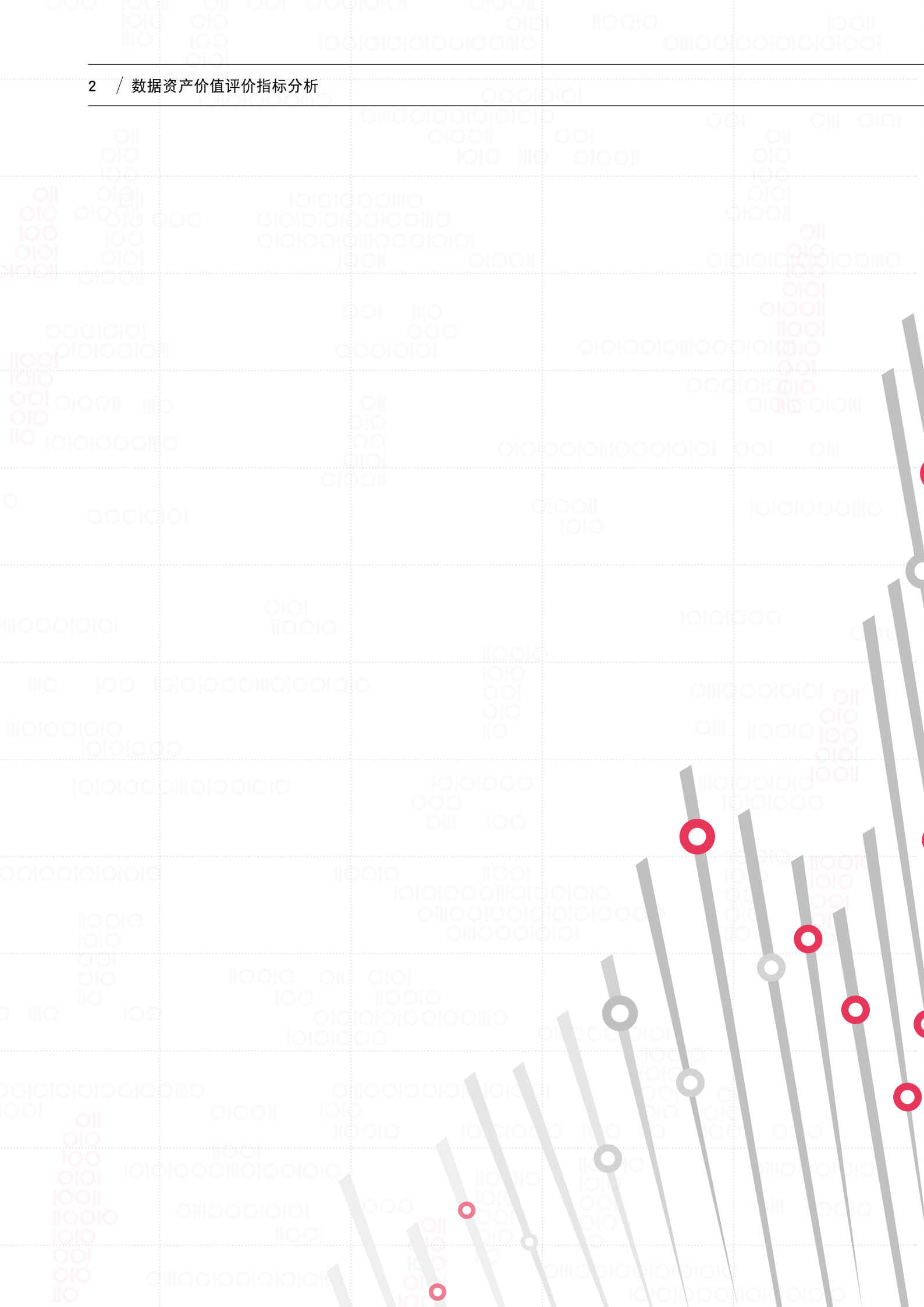
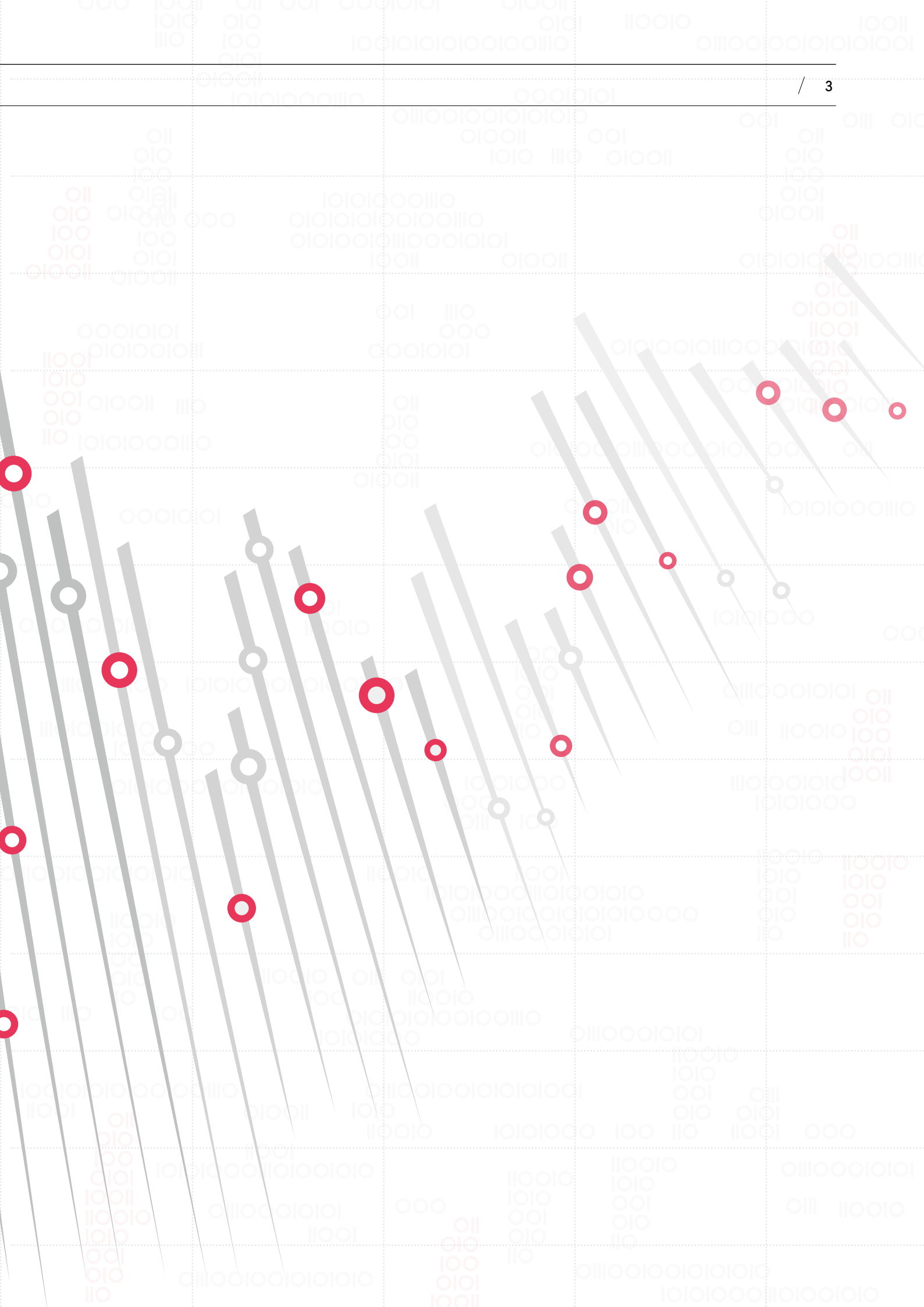


普华永道

数据资产价值

评价指标分析





■ 前言

随着计算机相关软硬件技术的迅猛发展，大数据的相关应用已经深入到生产生活的各个角落。随着数据应用领域的扩展，对于数据的需求也随之水涨船高，利用数据进行分析、将分析成果应用于运营决策也已成为企业发展的新动力。党的十九届四中全会将数据增列为一种生产要素，要求建立健全由市场评价贡献、按贡献决定报酬的机制。在此背景下，数据作为一种新型生产要素，针对其进行价值评价将是商业化交易流通过程中必不可少的一环。本文立足于数据价值与应用场景相结合的理论基础，认为在不同应用场景下，数据所贡献的经济价值有所不同，提出开放式数据资产价值评价指标分析框架，初步建立评价思路、评价维度与指标，由数据市场的主要参与者“买卖双方”协商选择标的的数据资产的评价指标，并确定各指标的权重，尝试对数据资产价值影响因素标准化并进行计算，为数据资产的价值评价研究提供探索经验。



目录

· 数据资产分类与应用价值	6
· 数据资产价值评价方法介绍	8
· 开放式数据资产价值评价指标分析框架	10
(一) 评价思路概述	10
(二) 评价指标介绍	12
成本维度评价指标	12
质量维度评价指标	13
应用维度评价指标	14
风险维度评价指标	14
(三) 评价结果处理	16
(四) 评价运用实例	17
· 数据资产价值评价应用	22
· 数据资产价值评价展望	23
(一) 数据权属合理拆分奠定数据流通基础	23
(二) 隐私计算确保数据流转安全	23
(三) 数据交易撮合平台助力数据流通	24
· 结束语	25
· 参考文献	26
· 联系人	27

数据资产分类与应用价值

未经清洗加工的原始数据，存在冗余、无序等方面的缺陷，导致其应用价值有限，而经过脱敏、分类、清洗、建模分析等数据加工处理后形成的可采、可见、标准、互通、可信的高质量数据，具备较高的应用价值。依照发展阶段，数据资产可以分为原始数据、粗加工后数据、精加工后数据、初探应用场景的数据、实现商业化的数据等。

数据的价值在于与应用场景的结合，不同应用场景下，数据所贡献的经济价值有所不同。数据资产按照应用领域不同，可划分为交通数据、医疗数据、金融数据、科研数据、社交数据、产业数据等。

表1数据资产分类示例

数据资产大类	数据资产子类	数据资产子类示例
交通数据	驾驶数据	踩刹车次数、驾驶稳定系数、里程数等
	位置数据	经度、纬度、车距等
	高速数据	高速车辆通行次数、高速运力指数、高速通行支付方式等
	车辆数据	车牌号、出厂日期、发动机型号等
	出行数据	出行时长、出行时间点、出行里程、出行起始点、行程停留情况等
医疗数据	疾病数据	诊断数据、病史、影像等
	健康数据	基因数据、体检数据、智能穿戴数据等
	物资数据	物资费用、物流费用、医药研发费用等
金融数据	基本面数据	监管文件数据、商业报告数据等
	市场数据	FIX数据、BWIC数据等
	另类数据	网络搜索、交易、天气等
教育数据	教学资源数据	教师个数、教师教育水平、教学设备费用等
	教育教学管理数据	网课网站个数、网课上线人数、寄宿学生个数等
	教育教学行为数据	学生出勤率、学生作业完成率、学生课数等
	教育教学评价数据	考试次数、学生成绩、教师考核成绩等
电商数据	总体运营数据	流量、订单、总体销售业绩等
	流量数据	流量规模、流量成本、会员数据等
	销售数据	购物车个数、下单个数、支付金额等
	商品类数据	商品总数、品牌存量、上架个数等
	市场营销活动数据	新增访问个数、新增注册个数、广告投资回报率等
	风控数据	评论卖家数、好评率、投诉率等
	市场竞争数据	市场占有率、交易额、用户份额等
房地产数据	土地数据	成交楼面价、土地出让金、土地供应面积等
	房企数据	拿地金额、销售面积、销售金额等
	住宅数据	成交面积、成交金额、成交套数等

数据资产的应用场景广泛，不仅可帮助企业进行精准营销、风险管控、商业决策等，也可应用于民生建设，促进社会发展和提高人民生活水平。例如，交通数据资产大类下的“出行数据”，可应用于移动出行的乘客风险监控。滴滴等出行服务商的数据分析师和算法开发人员，根据出行订单相关的“出行时长”、“出行时间点”、“出行里程”、“出行起始点”、“行程停留情况”等记录，判断乘客发生风险的可能性，以便随时监控风险和及时联系乘客确认乘车安全。

再如，医疗数据资产大类下的“疾病数据”，可应用于提高疾病治疗水平。医疗机构的科研人员，根据病患的“诊断数据”、“病史”、“影像”等记录，研究了解新冠肺炎的流行病学规律，优化治疗方案。同时，可基于医疗数据建立新冠肺炎流行病学模拟数理模型，预判疫情发展的速度和方向，帮助各级政府科学决策，让抗疫过程少走弯路。

表2数据资产应用示例

		数据资产	使用者	应用场景	应用价值
交通数据	驾驶数据	踩刹车次数、驾驶稳定系数、里程数等	保险公司的数据分析师、算法开发人员等	保费计算	分析司机驾驶数据，判断事故风险概率
	出行数据	出行时长、出行时间点、出行里程、出行起始点、行程停留情况等	出行服务商的数据分析师、算法开发人员等	风险监控	监测乘客出行数据，判断风险发生的可能性
医疗数据	疾病数据	诊断数据、病史、影像等	医疗机构的科研人员等	疾病研究	根据疾病数据进行实验，改进治疗水平
金融数据	分析数据	基本面的分析数据、市场行情的分析数据、另类数据的统计分析报告等	量化分析师	投资决策	根据各类分析数据得到研究报告，改善投资策略
教育数据	教育教学评价数据	考试次数、学生成绩、教师考核成绩等	学校教学团队	教学发展	通过教学评价数据，调整教学计划
电商数据	市场竞争数据	市场占有率、交易额、用户份额等	市场部门数据分析师、算法开发人员等	精准营销	通过市场数据分析可发现新的市场机会
房地产数据	住宅数据	成交面积、成交金额、成交套数等	房地产公司的数据分析师、算法开发人员等	客户挖掘	通过住宅数据分析识别和转化潜在客户群体

数据资产价值评价方法介绍

由上文我们了解到，数据资产具有应用价值，而不同应用场景下数据所贡献的价值不同，那么数据资产价值如何进行评价呢？比照无形资产价值评价方法，目前传统的数据资产价值评价方法主要为成本法、收益法和市场法三种基本方法。



成本法的原理是从产生数据资产所需花费的成本进行评价，在此基础上扣除各种贬值因素，并考虑数据资产的预期使用溢价，加入数据质量、数据基数、数据流通以及数据价值实现风险等数据资产价值影响因素进行修正，从而估算出标的的数据资产的价值。

计算公式

$$P=TC \times (1-L) \times (1+R) \times U$$

其中，P代表标的的数据资产价值，TC代表数据资产总成本，L代表数据资产贬值率，R代表数据资产成本投资回报率，U代表数据资产价值调整系数。

收益法的原理是对数据资产投入使用后的预期收益能力进行评价，考虑资金的时间价值，将未来各期收益进行加总，从而估算出标的的数据资产的价值。

计算公式

$$P=\sum_{t=1}^n F_t \frac{1}{(1+i)^t}$$

其中，P代表标的的数据资产价值， F_t 代表数据资产未来第t个收益期的收益额，n代表剩余经济寿命期，t代表未来第t年，i代表折现率。

市场法的原理是基于相同或相似数据资产的可比市场交易案例进行评价，对数据资产的价值密度、交易期日、容量等数据资产的性质等相关因素进行修正，从而估算出标的的数据资产的价值。

计算公式

$$P=P_o \times N$$

其中，P代表标的的数据资产价值， P_o 代表可比案例的数据资产交易价格，N代表数据资产性质修正系数。

此外, 业内学者在成本法、收益法和市场法三种基本方法的基础上, 衍生出不同的数据资产评估公式, 以及从不同的角度提出数据资产评估公式。

李永红等人基于市场法原理, 提出衍生的市场法评价数据资产价值。

计算公式

$$P = \frac{P_1 \times K_1 + P_2 \times K_2 + \dots + P_n \times K_n}{n}$$

其中, P 代表标的的数据资产价值, n 代表可比数据资产数目, P_i 代表第 i 个可比数据资产的市场价值, K_i 代表第 i 个可比数据资产根据各项指标确定的综合调整系数。

梁艳基于收益法原理, 提出使用多期超额收益法评价数据资产价值。

计算公式

$$P = \sum_{t=1}^n (E - E_w - E_f - E_i) \times (1+i)^{-t} \times K$$

其中, P 代表标的的数据资产价值, E 代表企业的自由现金流, E_w 、 E_f 、 E_i 分别代表流动资产贡献值、固定资产贡献值、除数据资产外的其他无形资产贡献值, K 代表数据资产价值调整系数, i 代表折现率, n 代表收益期限。

李希君认为数据的价值在于其减少不确定性的决策效用, 提出以“信息熵”进行数据资产价值评价。

计算公式

$$Pr(\cdot) = I(H(\cdot))$$

其中, $I(\cdot)$ 是一个非递减的联系函数, 它应该满足如下条件:

$$\forall x_1 \geq x_2, I(x_1) \geq I(x_2),$$

$$\forall x_1, x_2 \geq 0, I(x_1 + x_2) \leq I(x_1) + I(x_2).$$

张弛提出基于深度学习方法进行数据资产价值评价:

$$P = (1+G) \times (1+D) \times (1+A) \times (1+S) \times (1+R)$$

其中, P 代表标的的数据资产价值, G 代表颗粒度, D 代表多维度, A 代表活性度, S 代表规模度, R 代表关联度, G 、 D 、 A 、 S 、 $R \in [0, 10]$ 。

开放式数据资产价值评价指标分析框架

在数据资产交易逐渐成为社会各界关注重点的当下,设计和构建科学的数据资产价值评价体系对厘清数据资产属性、制定统一的定价机制、促进数据市场健康发展具有至关重要的意义。

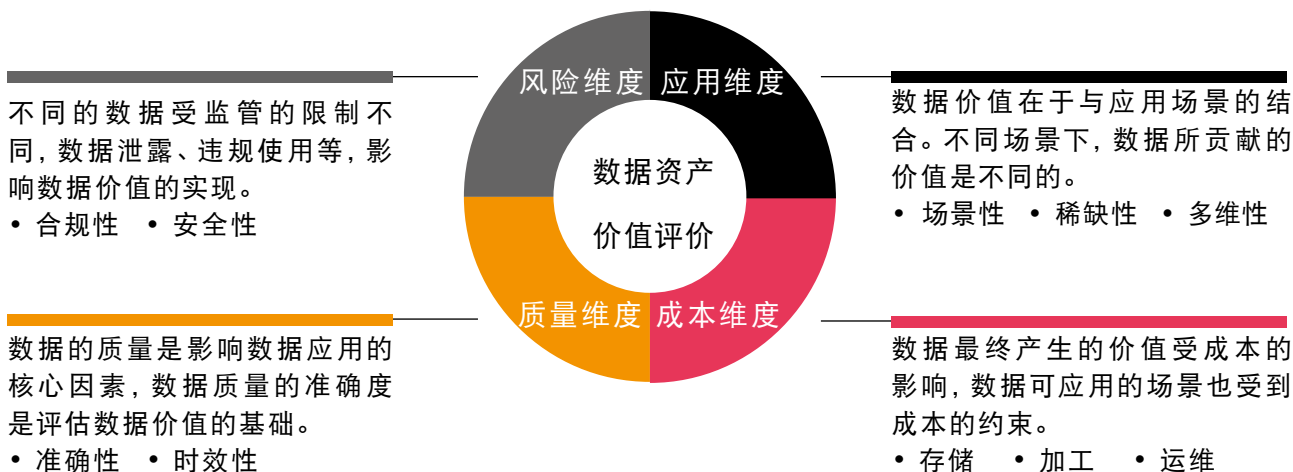
本文提出开放式数据资产价值评价指标分析框架,初步建立评价思路、评价维度与指标,由数据市场的主要参与者“买卖双方”协商选择标的的数据资产的评价指标,并确定各指标的权重,尝试对数据资产价值影响因素标准化并进行计算,为数据资产的价值评价研究提供探索经验。



(一) 评价思路概述

党的十九届四中全会将数据增列为一种生产要素,要求建立健全由市场评价贡献、按贡献决定报酬的机制,因此本白皮书从数据交易市场的主要参与者“买卖双方”的关注因素出发,来探讨数据价值评价思路。通常情况下,卖方关注数据资产的交易价格能否覆盖数据投入的成本并实现溢价,买方关注购买数据资

产后带来多少应用收益,而数据质量是数据应用的基础,同时在数据应用过程中存在数据泄露、违规使用等风险,影响数据应用价值的实现,因此,我们认为影响数据价值的最基本因素包括数据成本、数据应用、数据质量和数据风险。



资料来源：普华永道分析

基于上述4项影响数据价值的基本因素，我们提出”数据资产价值指数“计算公式：

计算公式

$$P=C \times Q \times U \times (1-R) \times (1-E) \times F$$

其中，P代表标的的数据资产价值指数，C代表数据成本系数，Q代表数据质量系数，U代表数据应用系数，R代表数据风险系数，E代表其他负向影响系数，F代表其他正向影响系数。其他影响指数E、F需根据数据市场发展进行调

整。对于成本、质量、应用、风险这4项数据价值影响因素，可设置具体指标和权重进行量化评价，经标准化处理后，汇总得到百分制的评价分数。以数据质量系数Q为例，其计算公式为：

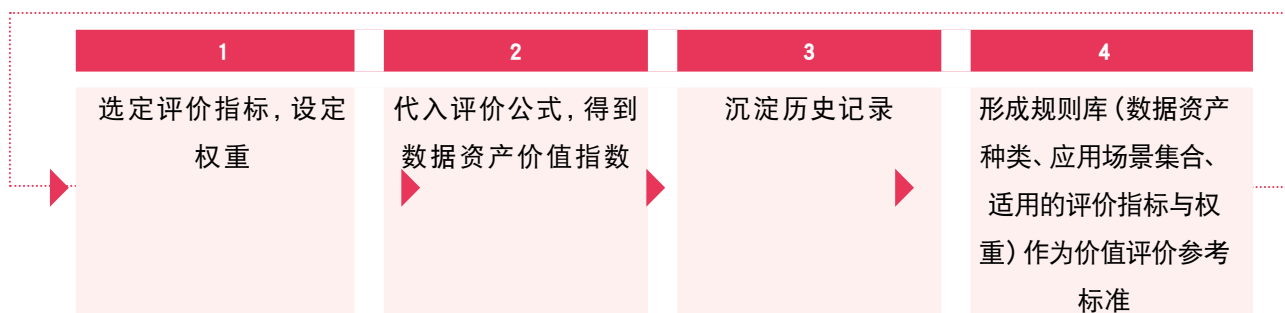
计算公式

$$Q = \sum_{j=1}^m Q'_j \times w_j \quad Q' = \sum_{i=1}^n p_i \times w_i$$

其中，Q'代表数据质量相关的一级指标评价价值， w_j 代表一级指标的权重， p_i 代表数据质量相关的二级指标评价价值， w_i 代表二级指标的权重。后文将在“评价指标介绍”章节，对成本、质量、应用、风险这4项数据价值影响因素的相关指标展开叙述。

遵循市场评价贡献、按贡献决定报酬的机制，对于不同应用场景下数据资产价值评价指标的选择与权重设定，由数据市场的主要参与者“买卖双方”协商确定。随着数据交易市场

的发展，大量的数据交易记录将沉淀出，各类数据资产在不同应用场景下的指标选择、权重设定记录。通过这些记录，我们可以获得各类数据资产的应用场景集合，并筛选出不同应用场景下针对不同数据资产最为适用的评价指标及对应权重，最终形成数据价值评价的规则库。未来，在积累了一定数据资产交易记录后，将历史交易信息与规则库结合，通过AI建模形成数据资产价值评价与定价模型，最终实现数据资产的自动化定价。



（二）评价指标介绍

1. 成本维度评价指标

数据资产的取得成本需要根据创建数据资产生命的流程特点，分阶段进行统计。我们将数据资产的取得成本，划分为数据规划、数据获取、数据处理三个阶段的成本。

数据规划阶段的成本，主要包括项目经费、市场调研费、设计评审费、咨询费和数据规划相关的人工工资等。数据获取阶段的成本，与数据的获取方式有关，包括内部生产运营过程获取的数据、间接调研方式获取的数据，以及外购方式获取的数据。通过内部生产运营过程获取的数据，其成本主要包括数据存储相关的设备折旧费、场地租金、水费、电费、空调费、网络费和下包商运维费等；通过间接调研方式获取的数据，其成本主要包括人工工资、设备折旧、场地租金、打印费、数据采集相关的网络费和下包商采集调研费等；通过外购方式获取的数据，其成本主要包括数据采买相

关的人工工资、数据购买价款、税费、注册费和手续费等。数据处理阶段的成本，与数据处理流程有关，包括数据核验、数据预处理和数据分析挖掘相关的成本。数据核验成本，主要包括核验相关的人工工资、设备折旧和下包商核验费等；数据预处理成本，主要包括数据预处理相关的人工工资、设备折旧和下包商数据处理费等；数据分析挖掘成本，主要包括数据打标相关的人工工资、设备折旧和下包商数据处理费，以及数据挖掘相关的人工工资、设备折旧、下包商数据研发费和可容试错费等。

以上各成本项目的发生额，可通过数据拥有者的财务记录、数据购买合同等相关文档获取。在获取各成本项目发生额的基础上，对各项成本进行加总，得到数据资产的成本总额。成本总额越大，数据资产的价值越大。

2. 质量维度评价指标

我们以全国信息技术标准化技术委员会提出的数据质量评价指标（GB/T36344-2018 ICS 35.24.01）为依据，从规范性、完整性、时效性、准确性、一致性、可访问性六个方面，对数据质量进行评价。

表3数据质量评价指标示例

一级指标	二级指标	释义	示例	
			指标约束规则	指标评价结果
规范性	长度规范率	数据长度符合规定格式要求的数量占该字段总数据量的比率	字段“营业执照号”的数据值长度规定格式为15位数字	数据表中“营业执照号”的数据值有80%符合长度规范要求，则长度规范率为80%

完整性	空值率	字段取值为空的数据量占总数据量的比率	客户信息表中的客户号不能为空	数据表中“客户号”有10%取值为空，则空值率为10%

时效性	更新频率	实际信息发生变化后，间隔多久可正确同步至数据表	门店新增会员人数更新至客户信息表的时间间隔	门店新增会员人数T+1日更新至客户信息表，则更新频率为T+1

准确性	精确度	满足精度要求的数据量占该字段总数据量的比率	利率数据需精确到小数点后4位	数据表“利率”有85%符合精确度要求，则精确度为85%

一致性	逻辑一致率	字段间的数据值满足逻辑约束关系的数据量占相关字段总数据量的比率	“投保开始时间”应符合小于等于“投保结束时间”的逻辑约束关系	相关字段的数据值有95%符合逻辑约束要求，则逻辑一致率为95%

可访问性	访问成功率	数据产品在约定时间长度内的可获取性	访问者在约定时间长度内通过数据接口对目标数据表发起访问请求且成功获取数据	访问者在1小时内通过数据接口对目标数据表发起10次访问请求，有9次请求成功获取所需数据，则访问成功率为90%

3.应用维度评价指标

数据资产在应用场景中的价值,可从数据资源的稀缺程度、数据覆盖范围的多样性,以及在该场景中的应用深度等方面进行评价。商业竞争的本质,部分来自于对于稀缺资源的竞争。在制造差异化趋平的情况下,稀缺数据资源背后潜在的商业信息更加凸显价值。数据资源的稀缺程度,是数据资产拥有者对数据独占程度的体现,可通过数据资产所拥有的数据量占该类型数据总量的比例来量化评价。数据覆盖范围的多样性,可通过数据维度(字段)丰富度进行评价,数据维度(字段)越多,数据表的信息覆盖范围越广,数据应用价值实现程度越高。在数据维度(字段)丰富度的基础上,可进一步通过数据维度(字段)适用率进行评价。例如,某张客户信息表中有8个关于描述个人信息的字段,包括住址、薪资、性别、年龄、

身份证号、手机号、职业、驾驶习惯,适用于精准营销场景的是住址、薪资、性别、年龄、职业这5个字段,那么该数据表在精准营销场景下的维度适用率为 $5/8$,数据维度(字段)适用率越高,数据表的价值密度越大,数据应用价值实现程度也越高。

数据资产在某个场景中的应用深度,反映的是数据资产在应用时的可挖掘价值大小,可通过数据访问记录、接口调用频次等指标进行评价。数据访问记录或接口调用频次低,说明数据使用者在数次使用后,因其应用价值有限或可挖掘价值较少,无需再进行使用,业务场景的应用深度低,数据价值实现程度低。数据访问记录或接口调用频次高,说明需高频次使用或深度挖掘,业务场景的应用深度高,数据价值实现程度高。

4.风险维度评价指标

数据资产在转让过程中,通常面临数据泄露、违规使用等风险,可能造成数据资产的大幅减值甚至价值完全损失。由于数据泄露、违规使用等风险发生的概率与数据控制者的数据安全能

力直接相关,我们引入“数据安全能力成熟度模型”(DSMM)作为数据控制者的数据安全能力评价标准,并将评价结果与数据风险发生概率进行关联,如下表:

表4数据安全能力与风险发生概率对照表

数据安全能力成熟度	数据风险发生概率
级别1：非正式执行	100%
级别2：计划跟踪	10%
级别3：充分定义	1%
级别4：量化控制	0.1%
级别5：持续优化	0.01%

由此可得，当标的的数据资产在经过n次转让后，数据风险发生概率r的计算公式为：

计算公式

$$r=1-(1-r_0)\prod_1^n(1-r_i)$$

其中， r_0 为初始数据控制者的数据安全能力成熟度所对应的数据风险发生概率， r_i 为数据第i次转让至新的数据控制者时，该数据控制者的数据安全能力成熟度所对应的数据风险发生概率。

数据资产作为一种新型生产要素，当前国内外对其风险度量方面的研究较为欠缺，尤其是定量计算。上表所列示的“数据风险发生概率”应在未来相关研究、实践的成果上进行调整和修正，实现对风险维度的准确量化。

（三）评价结果处理

由于数据资产各项评价指标的正负向性质不同，对不同性质指标直接加总不能正确反映不同作用力的综合结果，因此需要改变逆指标的数据性质，确保所有指标对数据资产价值评价结果的作用力同趋化。同时，数据资产各项指标的评价结果，存在百分数、金额、时长等单位不统一的问题，因此需要将不同量纲的特征转化为同一数值量级，确保数据的可比性。因此，在明确数据资产评价指标的基础上，需对各指标的评价结果进行数据标准化处理，统一转化为0~1之间的评价值。

本文借鉴优劣解距离法和同趋化处理，对各项指标的评价结果进行数据标准化。优劣解距离法的基本逻辑为，数据资产的各项指标评价结果，与该项指标的最优值越接近（最优值由第三方评价机构基于行业调研和实践经验提供），则评分越高。同时，通过同趋化处理各

评价指标的性质不同的问题。根据指标性质分为正向指标和负向指标，正向指标数值越大，评分越高，负向指标数值越小，评分越高，这将决定最优值为较小值还是较大值。计算思路为，分别计算各项指标评价结果与该项指标的最优值间的距离M、与最劣值间的距离N。计算公式为：假设某项指标的评价结果为x，最优值为y，最劣值为z，则与最优值距离为 $M=|x-y|$ ，与最劣值距离为 $N=|x-z|$ ，指标项评价值 $p=N/(M+N)$ 。

以“维度数量”这项指标为例，假设该项指标的最值分别为5个和100个，标的数据资产的实际维度数量为80个。经判断，指标性质为正向，因此最优值为100个，最劣值为5个，即 $M=|80-100|=20$ ， $N=|80-5|=75$ ，“维度数量”这项指标的评价值 $p=N/(M+N)=75/(20+75)=0.79$ 。

表5 指标评价结果数据标准化处理（注：指标评价结果x与指标最值y,z均为假定数值）

评价维度	一级指标	二级指标	指标估结果x	指标最值y,z	指标性质	数据标准化处理 $p=[N/(M+N)]$ $M= x-y $; $N= x-z $	指标评价值p
数据应用	多维性	维度数量	80个	100个， 5个	正向	$p=[80-5 /(80-100 + 80-5)]$	0.79
		维度适用率	70%	100%， 30%	正向	$p=[70\%-30\% /(70\%-100\% + 70\%-30\%)]$	0.57
数据质量	时效性	更新频率	T+1日	T+0日， T+30日	负向	$p=[1-30 /(1-0 + 1-30)]$	0.97
		更新延迟率	5%	0%， 30%	负向	$p=[5\%-30\% /(5\%-0\% + 5\%-30\%)]$	0.83
		时间跨度	4年	5年， 1年	正向	$p=[4-1 /(4-5 + 4-1)]$	0.75

在实际操作中,随着行业的变迁,可能出现标的的数据资产的某一指标数值超过第三方评价机构认定的最值范围。当超过范围的数据样本累计到一定程度,考虑进一步更新最值范围。同时,当区间内数值分布不均匀时,优劣解距离法不适用,可考虑采用同类排名、正态分布等方法,处理单位不统一的问题,将不同量纲的特征转化为同一数值量级。

经过上述标准化处理后,各项指标的评价价值分别乘以各自权重,得到上级指标的评价

(四) 评价运用实例

数据的价值在于与应用场景的结合,不同应用场景下,数据所贡献的经济价值有所不同,以下我们以金融信贷里的风控场景为例,来估算标的的数据资产的价值。

金融信贷生命周期管理大致可分为贷前、贷中、贷后三个阶段。贷前风控,主要进行准入审核、额度授信、支用审批等,这就需要收集用户在信用贷款之前的个人财务数据和其他基本数据。例如:年龄、家庭成员等。贷中风控,主要进行额度管理、流失预测、营销响应等,这就需要收集用户在贷款后的还款情况和各金融产品的使用状况信息。通过贷款人在贷款之后的各种还款情况,可以预测该贷款人或同类型贷款人在未来时间的还款能力和违约风险。贷后风控,主要进行催收、还款率预测、账龄滚动、失联预测,收集那些已经违约用户的

值,经层层分级加权汇总,得到数据资产的成本系数、质量系数、应用系数和风险系数,代入数据资产价值指数评价公式,即可得到标的的数据资产的价值指数。由于各指标评价结果经数据标准化处理后,均转化为0~1之间的评价价值,且各级指标的权重之和均为100%,因此最终得到的数据资产成本系数、质量系数、应用系数和风险系数取值也在0~1范围内。因此,数据资产价值指数P的取值范围为0~1。

各种金融数据。通过逾期客户的金融数据预测未来该用户无法偿还的概率,推测该笔贷款成为坏账的可能性。

我们选取贷中风控场景,该应用场景下的标的的数据资产为“客户贷款后的还款情况和各金融产品的使用状况”相关的8张数据表,包括描述贷款关联静态信息的信贷合同表及字段数据、对公客户表及字段数据,描述贷款动态变化信息的对公信贷业务借据表及字段数据、贷款展期表及字段数据,描述贷款主体分账户资金流水信息的对公信贷分户账表及字段数据、对公信贷分户账明细记录表及字段数据,描述贷款主体其他账户总体资金流水信息的对公活期存款分户账表及字段数据、对公活期存款分户账明细记录表及字段数据。鉴于篇幅限制,我们选取“贷款展期表”中的部分数据进行价值评价。

贷款关联静态信息	贷款动态变化态信息	贷款主体分账户 资金流水信息	贷款主体其他账户 总体资金流水信息
<ul style="list-style-type: none">• 信贷合同表• 对公客户	<ul style="list-style-type: none">• 对公信贷业务借据• 贷款展期	<ul style="list-style-type: none">• 对公信贷分户账• 对公信贷分户账明细记录	<ul style="list-style-type: none">• 对公活期存款分户账• 对公活期存款分户账明细记录

表6标的资产数据基本信息

数据表名称	数据元素（字段）	数据表获取方式	数据表转让次数	初始数据控制者的数据安全管理成熟度
贷款展期表	银行机构代码、展期金额、原利率、信贷借据号、展期利率、原借据号等	通过内部生产运营过程获取	0	级别3：充分定义

以下详细展开对标的资产数据资产的价值评价过程：

第一步：选取标的资产适用的评价指标，计算得到各项指标的评价结果。

首先，选取标的资产的成本维度适用的评价指标，计算得到成本总额。计算内容如表8所示。

表7成本维度评价结果表

总指标	一级指标	二级指标	取值	成本总额
数据成本	数据规划成本	项目经费	7000元	46220元
		市场调研费	600元	
		设计评审费	1000元	
	数据获取成本	存储设备折旧	500元	
		场地租金	820元	
		数据核验人工工资	2600元	
	数据处理成本	数据打标人工工资	3700元	
		下包商数据研发费	30000元	

其次，选取标的的数据资产的质量维度适用的评价指标，计算得到指标评价结果。计算内容如表9所示。

例如，在计算表中字段“银行机构代码”的长度规范率时，根据指标约束规则“12位人行行号数字”，我们统计出该字段取值为12位数字的数量是1266804，以及该字段取值不为12位数字的数量是12796。长度规范率表示数据长度符合规定格式要求的数量占该字段总数据量的比率，则字段“银行机构代码”的长度规范率=（数据长度符合规定格式要求的数量/总数据量*100%）=[1266804/(1266804+12796)]*100%=99%。

表8质量维度评价结果表

数据表	数据元素（字段）	一级指标	二级指标	指标约束规则	指标评价结果
贷款展期表	银行机构代码	规范性	长度规范率	12位人行行号数字	99%
	原利率	准确性	精确度	精确到小数点后4位	92%
	展期金额、原利率		值域符合率	>0	87%
	银行机构代码、展期金额、信贷借据号、展期利率、原借据号	完整性	空值率	必填项空（为null、空串、空格）	10%
	银行机构代码、展期金额、信贷借据号、展期利率、原借据号、原利率	时效性	更新频率	/	T+1日

再次，选取标的的数据资产的应用维度适用的评价指标，计算得到指标评价结果。计算内容如表10所示。

表9应用维度评价结果表

总指标	一级指标	二级指标	指标评价结果
数据应用	多维性	维度数量	6个
		维度适用率	70%
	应用深度	数据访问记录	100次

最后，评价标的数据资产的风险发生概率。标的的数据资产未经过转让，初始数据控制者的安全能力成熟度为级别3“充分定义”，根据上文的“数据安全能力与风险发生概率对照表”可知，初始数据控制者的数据安全能力成熟度所对应的数据风险发生概率 r_0 为1%，代入数据风险发生概率 r 的计算公式：

计算公式

$$r=1-(1-r_0)\prod_1^n(1-r_i)=1-(1-1\%)\prod_1^n(1-0)=1\%$$

第二步：对评价结果进行标准化处理，统一转化为0~1之间的评价值。具体计算内容如表11所示。

表10 指标评价值计算表

总指标	一级指标	二级指标	指标评价结果 x	指标最值y,z	指标性质	数据标准化处理 p=[N/(M+N)] M= x-y ; N= x-z	指标评价值p _i
数据成本	成本总额	/	46220元	100000元， 0元	正向	p=[46220-0]/((46220-100000)+46220-0))	0.46
	规范性	长度规范率	99%	100%， 50%	正向	p=[99%-50%]/((99%-100%)+99%-50%))	0.98
	准确性	精确度	92%	100%， 50%	正向	p=[92%-50%]/((92%-100%)+92%-50%))	0.84
数据质量		值域符合率	87%	100%， 50%	正向	p=[87%-50%]/((87%-100%)+87%-50%))	0.74
	完整性	空值率	10%	0%， 50%	负向	p=[10%-50%]/((10%-0%)+10%-50%))	0.8
	时效性	更新频率	T+1日	T+0日， T+30日	负向	p=[1-30]/((1-0)+1-30))	0.97
	多维性	维度数量	6个	100个, 5个	正向	p=[6-5]/((6-100)+6-5))	0.01
数据应用		维度适用率	70%	100%， 30%	正向	p=[70%-30%]/((70%-100%)+70%-30%))	0.57
	应用深度	数据访问记录	100次	300次， 0次	正向	p=[100-0]/((100-300)+100-0))	0.33
数据风险	风险发生概率	/	1%	100%， 0.01%	正向	p=[1%-0.01%]/ (1%-100%+1%-0.01%))	0.01

第三步：对各指标项的评价值进行加权汇总，得到标的的数据资产价值指数。计算内容如表12所示。

表11数据资产价值指数计算表

总指标	一级指标	一级指标权重 w_i	二级指标	二级指标权重 w_i	指标评价 p_i
数据成本	成本总额	100%	/	/	0.46
数据质量	规范性	15%	长度规范率	100%	0.98
	准确性	40%	精确度	40%	0.84
			值域符合率	60%	0.74
			空值率	100%	0.8
	时效性	15%	更新频率	100%	0.97
数据应用	多维性	30%	维度数量	50%	0.01
			维度适用率	50%	0.57
	应用深度	70%	数据访问记录	100%	0.33
数据风险	风险发生概率	100%	/	/	0.01

综上，标的的数据资产的成本系数、质量系数、应用系数、风险系数分别为：

$$C=0.46 \times 100\%=0.46;$$

$$Q=0.98 \times 100\% \times 15\% + (0.84 \times 40\% + 0.74 \times 60\%) \times 40\% + 0.8 \times 100\% \times 30\% + 0.97 \times 100\% \times 15\% = 0.8445$$

$$U=(0.01 \times 50\% + 0.57 \times 50\%) \times 30\% + 0.33 \times 100\% \times 70\% = 0.318$$

$$R=0.01 \times 100\% = 0.01$$

代入数据资产价值指数评价公式，得到标的的数据资产的价值指数为：

$$P=C \times Q \times U \times (1-R) = 0.46 \times 0.8445 \times 0.318 \times (1-0.01) = 0.1223$$



数据资产价值评价应用

对社会而言，数据资产的价值评价研究为数据资产交易统一定价模式提供了指导性框架体系及参考，促进构建流通的数据要素市场，提升社会各界对数据资产的认知，引导蓄势待发的数据交易市场正向发展。

对企业而言，推进数据资产的价值评价工作能够有效提升企业各业务线的运营效率和管理能力，定期的价值评价工作可以有效地协助管理层分析数据资产价值与企业价值的高度相关性，从而发掘高价值密度的数据，制定或修正业务发展目标及战略。此外，具有准确量化价值的数据资产可有效解决企业各部门在绩效理解、数字化运营等方面的沟通障碍，减少企业沟通成本，提高运营效率，推动企业整体朝着健康的方向发展。



无论是社会还是企业，可根据实际情况，结合成本法、收益法、市场法，将“数据资产价值指数P”应用到数据资产化计量，得到数据资产估价。“数据资产价值指数” $P=C \times Q \times U \times (1-R) \times (1-E) \times F$ （其中，P代表数据资产价值指数，C代表数据成本系数，Q代表数据质量系数，U代表数据应用系数，R代表数据风险系数，E代表其他负向影响系数，F代表其他正向影响系数。）

成本法是基于产生数据资产所需花费的成本进行评价，未考虑数据应用、数据风险、数据质量等因素对数据价值的影响，因此调整数据资产价值指数P为 P_c ， $P_c = Q \times U \times (1-R) \times$

$(1-E) \times F$ ，数据资产估价 $V = \text{成本} \times (1 + \text{期望利润率}) \times P_c$ ；

收益法是基于数据资产投入使用后的预期收益能力进行评价，未考虑数据成本等因素对数据价值的影响，因此调整数据资产价值指数P为 P_u ， $P_u = C \times (1-E) \times F$ ，数据资产估价 $V = P_u \times \sum_{t=1}^n F_t \frac{1}{(1+i)^t}$

市场法是基于相同或相似数据资产的可比市场交易案例进行评价，可将标的数据资产的价值指数 P_m 与可比市场交易数据资产的价值指数 P_s 进行比对，数据资产估价 $V = P_0 \times (P_m / P_s)$ 。

数据资产价值评价展望

（一）数据权属合理拆分奠定数据流通基础

在倡导充分发挥数据价值、利用数据要素助力企业智能化发展的现今，数据确权是横亘在数据自由流通运用道路上的一块巨石。数据应该是属于谁的？是属于创造数据的用户还是收集数据的商业机构？这一问题自从大数据应用于生产生活的各个领域开始就一直未得到解决。目前，海量的用户相关数据作为国内互联网巨头拥有的核心资产，若强硬地进行数据所有权的限制与转移无疑会造成国内数据行业的巨大震动。在此情境之下，拆分数据权属，合理分配数据相关的不同权利是助力数据流通的必经之路。

根据复旦大学特聘教授、重庆市原市长黄奇帆在第三届外滩金融峰会上的发言，数据涉

及五项基本权利：管辖权、交易权、所有权、使用权、分配权，各类主体基于自身在数据交易环节中的位置来行使相应权利。

目前，主流大数据多是由个人在社会生活中的各个数据平台上的行为记录所构成。从这个角度出发，平台与个人应共享这些数据的所有权，因为数据的产生离不开个人与平台的共同存在。数据的分配权则并不唯一，应该由数据生成、收集、处理、销售等一系列环节中所有的重要参与方共同协商分享，这也符合现代商业规则。数据的使用权则是考虑由对原始数据进行处理，生成数据产品的技术厂家或公司所有。

（二）隐私计算确保数据流转安全

随着数据资源重要性的日益凸显，其监管要求也逐渐严格，社会各界开始将目光投向一种以保护数据不外泄为前提，同时实现数据分析和计算的新兴数据应用技术“隐私计算”。隐私计算综合多方安全计算、联邦学习、可信计算等前沿科技的技术，在无需改变数据存储位置的情况下支持数据查询、数据建模等

多方数据协同利用的场景，进而实现对于数据价值的挖掘。隐私计算和区块链等技术结合形成的整体解决方案，能够对数据真实性、准确性进行记录。当数据安全遭受威胁，例如数据被非法篡改时、可通过相应安全机制查明，数据在哪个环节遭到泄露、是哪个主体泄露了数据。

在隐私计算框架下,各参与方的明文数据在无需出库的情况下即可进行数据价值的共享、挖掘,实现数据的“可用不可见”。隐私计

算在保护数据安全的同时,实现多方数据跨域合作,破解了数据保护与融合应用的诸多难题,将成为数据流通交易的重要支撑技术。

(三) 数据交易撮合平台助力数据流通

在过去数据要素市场发展,数据交易平台建设环节存在几个主要难点:(1) 数据安全保障难;(2) 交易信任构建难;(3) 服务模式持续难;(4) 数据权属确认难。为解决这些直接影响数据交易安全、数据拥有者交易参与意愿、数据市场持续运营等方面的难点,需建立可支持多层次数据要素流通、保障数据交易行为合规安全、提升各方参与积极性及有助于形成数据流通交易生态的新型数据交易平台。对于数据交易所,平台需具备商品管理、市场调控、交易行为监管、安全环境监控的能力,保障

数据交易全流程的安全可控。对于数据供方,平台需具备数据产品包装、数据交易撮合、数据经营指引的能力,在保障供方合法权益的同时,能协助供方开发数据产品并寻找适合的交易方。对于数据需方,平台需具备数据产品选择、产品购买推荐、产品售后服务的能力,在提供丰富数据产品的同时,能针对需方的痛点进行精准撮合,提升交易效率。能满足数据交易各参与方需求,且提升其参与交易意愿的交易撮合平台,未来将成为构建数据流通交易生态的重要一环。

结束语

随着科学技术与数据应用的飞速发展，数据资产呈现细分化、差异化、专业化等特征。在复杂商业环境中，如何对数据资产公允计量，并形成具有广泛公信力的数据资产价值评价体系，已成为推进数据资产交易的核心课题之一。基于对数据资产价值评价领域的探索与研究，我们提出开放式数据资产价值评价指标分析框架。在本文中初步建立数据价值评价思路、评价维度与指标，尝试对数据资产价值的影响因素标准化并进行计算，为数据资产的价值评价研究提供探索经验。

不同应用场景下数据所贡献的价值不同，因此我们在评价数据资产价值时，需结合其应用场景方有意义。数据交易流通领域内的研究经验揭示，数据在转让过程中存在安全风险，越转让风险越高。本文中，我们将数据风险纳入数据资产价值的评价维度之一，并借鉴数据安全能力成熟度模型，初探数据资产风险维度的衡量标准。

“数据资产价值研究”方法论目前仍处于探索阶段，需进一步的深入研究与论证。普华永道希望通过本文诚邀学术界、各科研单位以及业内人士开展深度交流探讨，就数据资产的价值评价与定价、权属划分以及数据交易流通相关的隐私计算、交易撮合等各领域进行进一步的研究，共同助力我国数字经济产业发展，为数据要素市场化尽绵薄之力。

参考文献

1. 资产评估专家指引第9号——数据资产评估
《企业会计准则》，财务部会计司，<http://kjs.mof.gov.cn/zt/kjzzss/kuaijizhunzeshishi/>
2. 《数据生产要素对中国经济增长的贡献研究》2021年
3. 《关于技术商品交易定价的思考》，载《经济师》2003第7期
4. 《关于技术交易定价的新思考》2021年
5. 《基于熵权-TOPSIS法的农村电商物流服务质量评价研究》，载《中国储运》2022年第1期
6. 《数据隐私问题的维度扩展与议题转换：法律经济学视角》，载《交大法学》2019年第1期
7. 《基于博弈论和拍卖的数据定价综述》，载《大数据》2021第7期
8. 《数据定价研究综述》，载《计算机科学与探索》2021第15期
9. 《数据资产价值评估与定价：研究综述和展望》，载《大数据》2021第15期
10. 《案例解析自主知识产权资产评估》，载《中国会计报》2021第7期
11. 《数据资产价值分析模型与交易体系研究》2018年
12. 《无形资产价值评估方法分析》，载《居舍》2021第8期
13. 《从原材料到资产——数据资产化的挑战和思考》，载《中国科学院院刊》2018年第8期
14. 《基于信息熵的数据交易定价研究》2018年
15. 《关于数据资产会计核算的研究》2017年
16. 《关于我国技术商品交易定价的思考》载2006第4期
17. 《数据价值化与数据要素市场发展报告（2021年）》
18. 《国内外大数据交易平台调研分析》2019年
19. 《互联网企业数据资产价值评估》2020年
20. 《数据资产化框架初探》，载《大数据期刊》2020年3期
21. 《数据交易所必须由国家管理》2021年
22. 《知识交易及其定价研究》2006年
23. 《专利资产评估中收益分成率的估算方法》，载《国有资产管理》2021年第7期
24. 《专利价值预评估分析方法研究》，载《中国科技信息》2021年第2期
25. 《企业数据资产的确认，计量与报告研究》2020年

联系人

张立钧

区域经济及金融业主管合伙人

james.chang@cn.pwc.com

+86 (755) 8261 8882

王斌

广东省市场主管合伙人

joanne.wang@cn.pwc.com

+86 (20) 3819 2100

傅毓敏

风险管理和内部控制部华南区主管合伙人

charlie.fu@cn.pwc.com

+86 (755) 8261 8892

翁泽鸿

数据安全及隐私保护咨询服务合伙人

danny.weng@cn.pwc.com

+86 186 883 99918

张锐

大湾区企业并购咨询部合伙人

roger.zhang@cn.pwc.com

+86 (20) 3819 2274

研究团队成员:

姚立, 普华永道中国数据资产管理及交易定价服务总监

陈雪夫, 普华永道中国数据安全及隐私保护咨询服务经理

黄耀驹, 普华永道中国数据资产管理及交易定价服务经理

谭韵欢, 普华永道中国数据资产管理及交易定价服务经理

张萌, 普华永道中国数据资产管理及交易定价服务高级顾问

盛厚钦, 普华永道中国数据资产管理及交易定价服务顾问

冉汶鑫, 普华永道中国数据资产管理及交易定价服务顾问