北京金融科技产业联盟
BEIJING FINTECH INDUSTRY ALLIANCE

公共数据在金融领域规范应用研究

北京金融科技产业联盟

2024 年 3 月

版权声明

本报告版权属于北京金融科技产业联盟，并受法律保护。转载、编摘或利用其他方式使用本报告文字或观点的，应注明来源。违反上述声明者，将被追究相关法律责任。

编制委员会

编委会成员：

何　军　　聂丽琴　　王立新　　常　征

编写组成员：

赵亚敏　　王　东　　李　聪　　薛佳梅　　郑平华　　王　雪

李武璐　　陆　燕　　徐宏杰　　关乃迁　　任　帅　　张浩瑀

汪翊琪　　张文韬　　张敬之　　余沛玥　　马福忠　　金银玉

单进勇　　黄翠婷　　陈　涛　　卞　阳　　杨天雅　　胡姣姣

李　博　　郑华祥　　艾　龙　　强　锋　　吴叶国　　马　利

编　审：

黄本涛　　郭　栋　　刘宝龙　　王　东

主编单位：

中国建设银行股份有限公司

参编单位：

建信金融科技有限责任公司

上海浦东发展银行股份有限公司

中国银行股份有限公司

浙商银行股份有限公司

北京数牍科技有限公司

同盾科技有限公司

上海富数科技有限公司

联易融数字科技集团有限公司

深圳市洞见智慧科技有限公司

天融信科技集团股份有限公司

深圳微言科技有限责任公司

目 录

一、 背景和意义 1

（一）研究背景 1
（二）研究意义 2

二、 公共数据市场情况 7

（一）概念和分类 7
（二）政策及法规 11
（三）公共数据应用模式 16
（四）公共数据产业发展情况 18

三、 问题与挑战 23

（一）合法合规 23
（二）数据安全 24
（三）质量保障 28

四、 公共数据在金融领域应用情况 31

（一）概况 31
（二）数据接入能力 32
（三）数据安全保障 35
（四）数据质量保障 46
（五）业务应用场景 53
（六）业务应用事例 58

五、 发展与建议 61

（一）法律政策趋势 61
（二）技术发展趋势 62
（三）发展建议 64

引言：党中央、国务院高度重视公共数据体系建设，不断推进数字政府建设，加强公共数据汇聚融合、共享开放和开发利用。2021 年，发布《国务院办公厅关于建立健全政务及公共数据共享协调机制 加快推进数据有序共享的意见》；2022 年，党中央、国务院先后印发《国务院关于加强数字政府建设的指导意见》、《全国一体化政务大数据体系建设指南》和《中共中央国务院关于构建数据基础制度更好发挥数据要素作用的意见》；2023 年组建国家数据局标志着公共数据开放共享进入新阶段。金融行业属于数据密集型领域，具有天然数据禀赋，积极响应国家对于数据要素市场建设的要求，强化数据的创新利用，以数据为关键要素，以价值释放为核心，加速数据要素市场的流通，加强内、外部数据的融合应用，如何在合法合规的基础上引入与应用公共数据，发挥公共数据要素价值已成为当前焦点。本文通过对我国公共数据的现状、特点进行梳理和总结，重点对政务及公共数据建设情况、金融机构的应用情况进行了深入调查和分析，并结合公共数据应用实践，提出公共数据接入、共享及应用的关键点和技术要求，供参考借鉴。

一、背景和意义

（一）研究背景

公共数据作为社会数据资源的重要组成部分，具有高权威性、高准确率、高可信度，在金融服务、医疗健康、城市治理等场景中具有极高的价值和市场需求，其开放利用已成为国际上备受瞩目的数据治理新方向。公共数据价值释放既是推动经济发展、完善社会治理、提升政府服务和监管能力的需要，也是数字经济全球竞争背景下增强国家竞争力的重要抓手。中国政府积极推进公共数据开放应用，着力提升公共数据治理水平，在政策制定、平台系统建设、数据共享开放、数据开发应用等方面开展大量实践并取得积极进展。2020 年，《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》提出“推进政府数据开放共享”，将数据开放作为经济社会快速发展的重要因素；2021 年，《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》将公共数据开放共享确定为“十四五”时期重点任务之一，明确提出“开展政府数据授权运营试点，鼓励第三方深化对公共数据的挖掘利用”。相关文件的出台，立足推动数据要素流通和市场化发展的价值目标，提出数据分级分类授权思路，强调要对各级党政机关、企事业单位依法履职或提供公共服务过程中产生的公共数据加强汇聚共享和开放开发，强化

统筹授权使用和管理，推进互联互通，明确公共数据的整体治理思路。

同时，各地各部门也都在加速公共数据开放进程，积极建设地方性公共数据开放平台，各地先后出台《上海市公共数据开放暂行办法》、《北京市公共数据管理办法》、《浙江省公共数据条例》、《广东省公共数据管理办法》等地方性法规和政府规章制度。据不完全统计，截至 2021 年，中国省级地方政府出台的公共数据治理相关政策已达 146 项，年投入资金超过 28 亿元，地市级及以上政府上线的公共数据开放平台达 193 个，为全面推动公共数据开放提供政策供给与制度安排。

从实践效果上看，政府依托数据采集打造公共数据资源库，数据存储方式向集约化存储深化推进，多层级数据流通框架初步构建，依托大数据平台围绕城市治理、环境保护、生态建设、交通运输、食品安全、金融服务、经济运行等应用场景开展数据分析应用，以服务场景为牵引推动数据融合，提供多样化共享服务。

（二）研究意义

目前，中国公共数据政策体系正在逐步搭建，在政务决策、治理、应用服务等各个环节发挥的作用日益凸显，但是从实际情况来看，公共数据开放度仍然较低，企业使用的公共数据占比不高，利用程度与市场需求之间仍存在较大缺口。目前存在的问题，主要包括以下方面：

1. 公共数据开放的基础设施及生态建设有待完善

第一，开放的公共数据可用性较低，影响开发用效果。数据可用性是数据开发利用的前提条件，直接关系到数据开放后的使用效果。因此，国际上在推进公共数据开放时均会设置相应的开放标准，如 2007 年总部位于英国的开放知识基金会下设的开放政府工作组（Open Government Working Group）首次提出了公共数据获取和开放的八项标准：完整性、原始性、及时性、可获得性、机器可读、非歧视、非财产性、免于许可。2015 年，《国际开放数据宪章》进一步提出了公共数据开放的六大准则：默认开放、及时全面、可获取和可使用、可比较及客户操作、改善政府治理及扩大公民参与、包容性发展与创新，以此来为公共数据开放提供可用性保障。但从实际执行情况来看，目前各国开放情况均不理想，达标的开放数据总量几乎仍然处于停滞状态。目前，中国各地数据开放平台的数据普遍存在碎片化、低容量、更新慢或不更新、未使用开放许可、机器不可读等问题，严重影响了公共数据的开发和利用。

第二，支持公共数据高效率调用、深层次开发与利用的基础能力和平台尚不完善。目前，中国部分地区建立了省级、地市级数据开放平台，归集了部分公共数据并开放其统计数据集以供查询和下载，但受资源预算和技术能力掣肘，海量的原始数据的归集、加工与开发应用尚未形成规模。例如，浙江、上海、深圳等地的平台虽然提供了 API 接口，但是普遍存在接口调用难度高，

可调取的数据容量小、更新频率低等问题；海南、北京等地虽然已经在积极探索打造公共数据开发应用平台，但是绝大多数平台尚不具备足够的数据处理能力，在数据接入过程中，经常出现系统间调用容量受限、服务稳定性得不到有效保障、需要反复重试等问题。当前，中国不同地区的公共数据运营模式虽然各不相同，但是大多由于缺乏明确的路径指引而导致各地在深层次发展上面临着诸多困境，因此，对基础性平台和深度开发能力的建设很难取得实质性的进展。

第三，数据产品和数据服务丰富程度和定制化程度仍有待加强。目前，中国公共数据产品和服务模式单一，地方的公共数据开放门户网站更多是服务于公众对统计数据集的查询和下载，细颗粒度的数据开放不足。此外，通过数据开放平台开放的动态数据占比较低，开放的公共数据商业化应用有限。结合当前中国公共数据开放的实践来看，中国公共数据运营被归纳为“行业主导模式”、“区域一体化模式”、“场景牵引模式”，其中以前两者为代表，反映出当前中国公共数据产品中存在以“供给导向”为主、与产业需求匹配度不够高、缺少紧密结合产业需求的有针对性的产品设计和开发等问题。以金融场景为例，由于不同金融机构在对待同类数据的对客方式、使用场景、风控机制上各有不同，所以目前数据接口标准化难以快速适应金融市场的需求。开放银行模式为丰富数据服务类型提供了宝贵的模板，然而金融系

统之外的部门，囿于价值回馈方式的不同，这种数据共享机制很难被直接复制和借鉴。

2. 公共数据开放面临数据安全挑战，掣肘数据开放进程

公共数据的开放在创造经济社会价值，提高政府治理能力的同时，还面临着数据安全风险与合规的挑战。2005 年，美国某医疗机构因其员工违规操作导致患者信息大批量外泄;2017 年，因美国健身软件发布用户定位的热力地图，导致美军军事基地及作战信息泄露;Facebook（现更名为 Meta）继被指控利用数据操纵美国大选后再次暴露出风险隐患，超过 5 亿用户信息被泄露甚至被非法销售等等。上述事件在推动美国等国进行数据安全立法的同时，也为公共数据的开放利用敲响了警钟。《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》三大支柱性立法的出台，为中国数据的安全流通提供了制度保障。但在发展层面，配套的开放制度却尚未建立，这也导致公共数据开放缺乏安全可行的规则指引，数据安全和个人信息保护成为悬在各地政府头上的达摩克利斯之剑，在客观上限制了公共数据的开放。

3. 公平有效的公共数据开放模式尚未形成影响利用效果

中国公共数据开放处于起步阶段，相关机构和企业对于公共数据开放持较为谨慎的态度，虽然我国基本构建公共数据授权开放的基础框架，各地政府也在不断探索开放利用机制，但是具体的实施措施仍然有待建立。目前，受限于数据安全要求、技术水

平等方面的因素，公共数据在利用上存在失衡的现象，部分地区对于公共数据的开放力度和影响有限，难以形成常态化、规模化、市场化的公共数据开放生态。此外，在公共数据开放过程中，由于欠缺规范的运营机制，数据资源提供方、数据开发服务商往往拥有优先获得开发利用的权利，且利用这一优势实现不当获利的情形也时有发生，破坏了数据的公共属性。另外，相关机构面对大量数据需求的服务能力仍有待提高。因此，公共数据开放的步伐较慢或者有选择性地开放数据，其中较明显的是更倾向向国资企业、大型机构开放，因此也造成公共数据开放利用过程中不同主体面临不同处境。这种数字鸿沟亦成为数据要素市场基础设施建设中的重要掣肘与重点关切。例如，电力等能源数据仅向大型银行开放，对中小微型金融机构准入设有限制，影响了实际利用效果，很难切实满足行业需求，中小微型金融机构难以实质性分享公共数据开放红利，不利于公共数据价值的充分释放。

本文重点针对上文所列的问题，并结合我国公共数据的现状、特点进行梳理和总结，重点对国家十余个部委及其各分支机构、各省市地方政府数据建设情况及各金融机构的应用进行深入调查和分析；并结合公共数据应用实践，提出公共数据成功接入、共享及应用的关键风险点和技术要求，供参考借鉴。

二、公共数据市场情况

（一）概念和分类

数据开放背景下的公共数据，最初指政府数据、政务数据，其内涵呈现出公共机构主体和行政职权行为属性。随着我国数据要素流通战略的深入推进，公共数据概念呈现出明显的扩张化态势，“公共数据”由原来所指的“公共的”数据，扩展为“具有公共价值的”数据，由此，企事业单位、社会团体的数据被纳入到公共数据开放制度的调整范围。

现阶段，相关地方政策和标准中已对公共数据进行初步的定义，本文挑选部分政策和标准进行示例，如表 1 所示。

表 1：现阶段公共数据相关定义归纳

政策/标准文件名	定义内容
《广东省公共数据管理办法》	是指公共管理和服务机构依法履行职责、提供公共服务过程中制作或者获取的，以电子或者非电子形式对信息的记录
《江苏省公共数据管理办法》	是指本省各级行政机关、法律法规授权的具有管理公共事务职能的组织、公共企事业单位（以下统称公共管理和服务机构）为履行法定职责、提供公共服务收集、产生的，以电子或者其他方式对具有公共使用价值的信息的记录。
《浙江省公共数据开放与安全管理暂行办法》	是指各级行政机关以及具有公共管理和服务职能的事业单位(以下统称公共管理和服务机构),在依法履行职责过程中获得的各类数据资源
《上海市公共数据开放暂行办法》	是指本市各级行政机关以及履行公共管理和服务职能的事业单位（以下统称公共管理和服务机构）在依法履职过程中，采集和产生的各类数据资源。
《宁波市公共数据安全管理暂行规定》	是指行政机关以及履行公共管理和服务职能的事业单位(以下统称公共管理和服务机构) 在依法履行职责过程中获得的各类数据资源。

《福州市公共数据开放管理暂行办法》	是指本市各级行政机关、事业单位、社会团体(以下简称“数据提供部门”)在依法履行公共管理和服务职能的过程中，产生、采集和获取的各类数据资源。
《重庆市公共数据开放管理暂行办法》	是指本市各级行政机关以及履行公共管理和服务职能的事业单位(以下简称公共管理和服务机构)，在依法履职过程中产生、采集和制作的，以一定形式记录、保存的各类数据资源。
《武汉市公共数据资源管理办法》	是指本市各级政务部门在履行职责和公共企事业单位在提供服务过程中产生或者获取的各类数据的总称。
浙江省地方标准 DB33/T 2349——2021 《数字化改革 公共数据目录编制规》	国家机关、法律法规规章授权的具有管理公共事务职能的组织（以下统称公共管理和服务机构）在依法履行职责和提供公共服务过程中获取的数据资源，以及法律、法规规定纳入公共数据管理的其他数据资源。
深圳市地方标准 DB4403/T 271—2022 《公共数据安全要求》	公共管理和服务机构及处理大量个人信息的服务平台在依法履行公共管理职责或者提供公共服务过程中产生、处理的数据。

综合分析当前各地政策和标准，结合我国数据要素流通政策趋势，本文初步定义，公共数据是指由国家机关、法律法规授权的具有管理公共事务职能或者提供公共服务的各级行政机关、企事业单位、社会团体等组织，在履行公共管理职责或者提供公共服务过程中，收集、产生的涉及公共利益的各类数据。其中涉及的主体包括国家机关、企事业单位、经依法授权具有管理公共事务职能的企事业单位、社会团体，以及供水、供电、供气、公共交通等提供公共服务的部门。

按照公共数据的共享开放程度进行分类，公共数据可以分为无限制使用的数据、授权使用的数据和禁止使用的数据。无限制使用的数据又称无条件共享类数据或普遍公开的数据，是指任何组织和个人都可以基于公共数据的使用目标不受限制地获取和

使用的公共数据。授权使用的公共数据又称有条件共享类数据或受限开放的数据，是指数据内容较为敏感，对开放主体、开放流程有要求或限制的数据类型。禁止使用的数据是指不予共享开放的数据。

从公共数据来源主体看，公共数据主要包含五种类型，一类是政务数据，即政务部门依法履职过程中采集、获取的数据；二类是具有公共职能的公共企事业单位，在提供公共服务和公共管理过程中产生、收集、掌握的各类数据资源，如教育医疗数据、水电煤气数据、交通通信数据、民航铁路数据等；三类是由政府资金资助的专业组织在公共利益领域内收集、获取的具有公共价值的数据，如基础科学研究的数据；四类是具有公共管理和服务性质的社会团体掌握的与重大公共利益关切的数据；五类是涉及公共服务领域的其他数据来源，如其他社会组织和个人利用公共资源或公共权力，在提供公共服务过程中收集、产生的涉及公共利益的数据。

根据公共数据的数据类型来划分，则分为：自然人类公共数据、法人类公共数据、信用类公共数据、自然地理类公共数据、感知类公共数据、统计类公共数据等。

自然人类公共数据：主要来源于公安、民政、人力社保等部门，规定了公共数据中的自然人基本信息、资产信息、社会活动、荣誉资质、涉事涉法等类数据的定义与属性，具体包括姓名、民族、证件号码、婚姻状况、文化程度、从业状况等公共数据元；

法人类公共数据：主要来源于市场监管、发改、经信等部门，规定了公共数据中的法人基本信息、资本与资产、许可、资质与荣誉、纳税、参保与缴费、生产经营、行政执法、司法信息、信用评价等类数据的定义与属性，具体包括统一社会信用代码、法人名称、住所、营业收入、许可编号、资质等级等公共数据元；

信用类公共数据：主要来源于发改、市场监管、财政等部门，规定了公共数据中的自然人信用信息、企业信用信息、社会组织信用信息、事业单位信用信息、政府机构信用信息、特征人群及领域信用信息等类数据的定义与属性，具体包括荣誉类型、评价等级、舆情内容、列入经营异常名录原因、行政处罚决定书文号、司法案件案号等公共数据元；

自然地理类公共数据：主要来源于自然资源、生态环境、气象等部门，规定了公共数据中的基础地理信息、地质信息、土地信息、覆被信息、海洋信息、生态环境信息、气象灾害信息等类数据的定义与属性，具体包括联系地址、经度、纬度、空间坐标系、地面分辨率、地理标识符、气温、场所用途等数据元；

感知类公共数据主要来源于建设、公安等部门，规定了公共数据中的感知采集信息、感知设备信息等类数据的定义与属性，具体包括感知对象、数据传送方式、数据摘要、设备参数、技术特征信息等数据元；

统计类公共数据：主要来源于统计、经信等部门，规定了公共数据中的统计指标信息、统计制度信息、统计目录信息、统计

报表信息等类数据的定义与属性，具体包括指标名称、指标编码、统计周期、统计时间、指标数据值、统计层级、统计模型等数据元。

（二）政策及法规

2022 年 12 月，中共中央国务院发布《关于构建数据基础制度更好发挥数据要素作用的意见》（“数据二十条”），描绘了数据基础制度的四梁八柱，提出以产权制度为基础、以流通制度为核心、以收益分配制度为导向、以安全制度为保障的数据基础制度顶层框架，对于充分激发数据要素价值具有全局性、奠基性、引领性重要作用。其中公共数据，作为数据要素中权威性、通用性、基础性、可控性、公益性较强的数据类型，在“数据二十条”中作出重要的系统部署。要求公共数据“加强汇聚共享和开放开发，强化统筹授权使用和管理，推进互联互通，打破“数据孤岛”。鼓励公共数据在保护个人隐私和确保公共安全的前提下，按照“原始数据不出域、数据可用不可见”的要求，以模型、核验等产品和服务等形式向社会提供，对不承载个人信息和不影响公共安全的公共数据，推动按用途加大供给使用范围。推动用于公共治理、公益事业的公共数据有条件无偿使用，探索用于产业发展、行业发展的公共数据有条件有偿使用。依法依规予以保密的公共数据不予开放，严格管控未依法依规公开的原始公共数据直接进入市场，保障公共数据供给使用的公共利益。”从顶层设计来推

动公共数据的高质量开发利用。公共数据相关政策及法规如表 2 所示。

表 2：公共数据相关政策及法规

时间	名称	主要内容
2015/08	《促进大数据发展行动纲要》	稳步推动公共数据资源开放。
2016/12	《大数据产业发展规划（2016-2020）年》	在全国建设若干国家大数据综合试验区，在大数据制度创新、公共数据开放共享等方面开展系统性探索试验。
2017/02	《关于推动公共信息资源开放的若干意见》	要求推进公共信息资源开放，要加强规划布局等。
2017/05	《政务信息系统整合共享实施方案》	加快公共数据开放网站建设。
2017/06	《政务信息资源目录编制指南（试行)》	对政务信息资源的目录编制、管理进行了规定。
2018/01	《公共信息资源开放试点工作方案》	确定北京市、上海市、浙江省、福建台、贵州台为试点地区，开展公共信息资源开放试点工作，并确定了重点开放领域。
2020/04	《关于构建更加完善的要素市场化配置体制机制的意见》	推进政府数据开放共享。优化经济治理基础数据库，加快推动各地区各部门间数据共享交换，制定出台新一批数据共享责任清单。研究建立促进企业登记、交通运输、气象等公共数据开放和数据资源有效流动的制度规范。
2021/06	《中华人民共和国数据安全法》	国家制定政务数据开放目录，构建统一规范、互联互通、安全可控的政务数据开放平台，推动政务数据开放利用。
2022/12	《中共中央、国务院关于构建数据基础制度更好发挥数据要素作用的意见》	从数据产权、流通交易、收益分配、安全治理等方面构建数据基础制度，提出 20 条政策举措。

此外地方层面公共数据相关的制度规范及政策文件也相继印发。目前已有多个省正式出台了公共数据管理条例或开放条例，以促进公共数据利用为基本定位，最大程度促进公共数据流通和

开发利用，促进数字经济和社会发展。地方公共数据相关政策及法规如表 3 所示。

表 3：地方公共数据相关政策及法规

时间	地区（省级）	名称
2017/03	浙江	《浙江省公共数据和电子政务管理办法》
2022/01		《浙江省公共数据条例》
2023/08		《浙江省公共数据授权运营管理办法（试行）》
2018/09	上海	《上海市公共数据和一网通办管理办法》
2023/03		《上海市公共数据共享实施办法(试行)》
2021/01	北京	《北京公共数据管理办法》
2023/07		《北京经济技术开发区公共数据管理办法（试行）》
2019/01	吉林	《吉林省公共数据和一网通办管理办法(试行)》
2020/07	天津	《天津市公共数据资源开放管理暂行办法》
2021/10	广东	《 广东省公共数据管理办法》
2021/12	江苏	《江苏省公共数据管理办法》
2022/01	江西	《江西省公共数据管理办法》
2022/02	山东	《山东省公共数据开放办法》

其中《浙江省公共数据条例》由浙江省第十三届人们代表大会第六次会议通过，自 2022 年 3 月 1 日起施行，为充分激发公共数据新型生产要素价值、推动治理能力现代化提供制度样本。在该条例中将“公共数据”定义为：国家机关、法律法规规章授权的具有管理公共事务职能的组织以及供水、供电、供气、公共交通等公共服务运营单位，在依法履行职责或者提供公共服务过程中收集、产生的数据，等等。对这类数据应当以共享为原则，不共享为例外。同时根据公共数据所承载的信息的保密程度，推动公共数据无条件开放、有条件开放和不开放。一是完善政务数据共享制度，完善公共数据标准、目录编制、质量管控等，提升

数据及时性、完整性、准确性，进一步探索公共数据资产凭证化管理。二是支持地方加快探索公共数据授权运营机制，明确授权标准、条件和具体程序要求，建立授权运营评价和退出机制。此外，浙江省还出台《浙江省公共数据授权运营管理暂行办法（征求意见稿）》，建立全省公共数据资源授权运营管理机制。建立数据产权制度，可在相关地方探索的基础上，将被证明行之有效的有益经验适时上升到制度层面。探索成立专门机构支撑公共数据主管部门对公共数据开展授权运营，推动实现对公共数据加工使用权和产品经营权的有序流转，加大公共数据供给使用范围，推动公共数据以产品形式向社会提供服务。

上海市在 2019 年 10 月施行《上海市公共数据开放暂行办法》是国内首部针对公共数据开放进行专门立法的政府规章，要求“明确公共数据开放的管理体制”“建立公共数据开放的长效机制”“优化公共数据开放平台建设”“确保公共数据利用合法正当”“打造公共数据多元开放系统”以及“强化公共数据开放的监督保障”。此后，在 2022 年 12 月上海进一步印发《上海市公共数据开放实施细则》对立法目的和依据、适用范围、概念意涵、工作原则以及职责分工进行规定，进一步规范上海市公共数据开放制度举措，为国家层面形成统一规则进一步积累地方经验。在 2023 年 2 月则发布《上海市公共数据共享实施办法(试行)》，以“共享为原则，不共享为例外，除法律法规另有规定外，公共数据应当全量上链、上云、充分共享”作为基本原则，鼓励公共

管理和服务机构通过公共数据共享应用，创新管理和服务方式，提升管理和服务水平。并指出要进一步完善公共数据共享规则，压实共享责任，夯实基础底座，培育应用生态，加强跨层级联通，加快推动公共数据上链、上云，持续打造高效集约、互通共用的系统平台，确保数据安全可控，以高质量数据供给推动治理能力和服务能级不断提升。

2022 年 11 月，广东省发布《广东省公共数据开放暂行办法》，是以规范和促进公共数据开放和开发利用，释放公共数据价值，深化数字政府改革建设，提升政府治理能力和公共服务水平，推动数字经济、数字社会发展为目标。指出，公共数据开放工作应当遵循统筹管理、分类分级、便捷高效、安全可控的原则；保障网络和信息安全是防范公共数据开放风险的基础保障，明确各安全管理主体的职责和义务，建立预警、响应、处理和灾难恢复机制。2023 年 4 月广州市进一步发布了《广州市公共数据开放管理办法》，扩大开放范围，明确公共数据开放及管理行为的适用范围包括供水、供电、供气、通信、民航、铁路、道路客运、公共资源交易等提供公共服务的企事业单位，结合优化营商环境等要求，在合法有序前提下适度扩大公共数据开放的覆盖面。此外提出多元主体参与，鼓励企业、行业协会、科研机构、社会组织等依法主动开放自有数据，促进公共数据和非公共数据的多维度开放和融合应用，形成内容有别、各具特色、各有侧重的数据开放格局。

（三）公共数据应用模式

公共数据主要有数据开放、数据共享、数据交易等流通形式。公共数据开放是指公共数据持有机构通过数据接口、网站等形式，将其工作行为过程中获取的公共数据资料向公民、企业或社会组织合理合法合规地开放获取与使用权限的行为无偿地向社会公开，主要手段是建立公共数据开放平台，重在通过平台寻求数据价值创造生态的打造。数据共享是指参与方互为对方的供应方和需求方，可分为政府间的公共数据共享、政企间的公共数据共享，。重在解决需求方的某一类具体问题。数据交易 是指在数据要素市场中数据供应方和数据需求方以货币或与货币等值的其他产品或服务作为媒介进行的数据交换行为，重在通过数据交易所等数据交易平台，逐步构建完整、有效、规范的数据交易管理体系。

目前，我国政府正在积极推进公共数据的开放共享，建立了一些公共数据授权运营模式，并建设了一些公共数据运营平台。如国家数据交换与共享平台，中国可持续发展信息平台，公众数据平台。政府在公共数据方面具有权威性和专业性，能够更好地管理和利用公共数据资源。同时政府在公共数据方面加大了投入，形成了一定规模的数据资源库。

根据“数据二十条”，公共数据按照“原始数据不出域、数据可用不可见”的要求，以模型、核验等产品和服务等形式向社会提供，用于公共治理、公益事业的公共数据有条件无偿使用;

用于产业发展、行业发展的公共数据有条件有偿使用。围绕公共数据流通，当前主要任务为政务大数据平台建设与公共数据运营。

公共数据流通需要依托全国一体化大数据中心体系。根据国家发改委解读，实现“原始数据不出域、数据可用不可见”的新型开发利用模式，需要安全可控、集约高效的基础环境，并提供充足的算力、存力资源支撑。2022 年 9 月，国务院办公厅印发《全国一体化政务大数据体系建设指南》提出了全国一体化政务大数据体系建设目标、总体架构和重点任务。根据《指南》，全国一体化政务大数据体系包括三类平台，为“1+31+N”框架结构，即 1 个国家政务大数据平台，31 个省（自治区、直辖市）和新疆生产建设兵团统筹建设的省级政务数据平台，“N”个国务院有关部门的政务数据平台。除以省为单位统筹建设的政务云外，各省、市及各部委政务大数据平台建设需求或将得到释放。

公共数据授权运营是数据要素市场新方向，目前浙江、山东、湖南、北京等省市已相继出台了公共数据授权运营的实施办法。公共数据授权运营的模式主要分为特许开发模式、独家经营模式、市场供应链模式、主题牌照模式等。特许开发模式，是以数据资源为特许资源授权于特定的加工利用方，生产应用性的数据产品或服务，满足场景应用需求。独家经营模式，是全市构建统一平台，并为平台选择唯一运营方，该平台方独家负责平台运营，并负责开发利用和平台绑定的数据资源，以产出要素性数据产品或服务为主，服务于下游数据应用方，同时产生应用性数据产品或

服务为辅。市场供应链模式，是按职能分工构建多级授权，汇聚不同职能主体链式合作运营公共数据，同样赋能下游数据应用方，但采用类似数据交易所提出的“数商”生态一样，在“数据”到“应用”这一加工链条上，充分拆解智能，引入竞争，去培育以“数据加工主体”和“数据运营主体”为代表的“授权运营商”的多类型市场。主体牌照模式，是按照应用主题授予牌照，取得牌照方可在准许的应用主题范畴内，进一步按需获取共享数据，开展运营，其产出要素性数据产品或服务，或产出应用性数据产品或服务。

（四）公共数据产业发展情况

数据要素市场蓬勃发展，据预测，“十四五”期间我国的数据要素市场规模将达到 5 千亿—1 万亿元，考虑数据要素全链条投入产出，以及围绕数据资产的评估、质押、融资等衍生市场，整体规模可能会超过 60 万亿。数据要素的重要性日益显著，是信息时代的核心资源，也是国家竞争力的重要组成部分，目前美国、欧洲、英国等很多国家都发布了国家数据战略，国家间的数据竞争将会很激烈，我国同样要从数据大国发展为数据强国。而公共数据因为与公共生活紧密相连、与社会效益息息相关，具有覆盖面广，数量大，价值高的特点，同时具有权威性、通用性、公益性与可控性等特性，是数据要素产业发展的极为重要的组成部分，具有赋能经济发展的强大潜能。公共数据，一方面不仅可以提升政府部门之间的治理协作能力，还可以助力政府掌握社会

运行过程中的相关内在规则，制定相关应对策略，为解决社会复杂问题提供新的手段及方式；另一方面在社会经济方面，公共数据的开放有助于促进传统经济运行机制得以变革，促使产业结构实现优化及调整，推动经济增值创新。因此公共数据，不仅是数据要素流通的重要组成部分，也是数字经济发展的重要物质基础，具有重大战略意义。

目前，国际、国内都在积极探索公共数据开放模式。越来越多的国家认识到了公共数据的重大经济意义，促进数据开放共享也因而成为国际上的重要议题。英国、美国、澳大利亚、新加坡等国家为实现公共数据开放做了一系列的准备工作，具体措施包括建立规范与标准、出台政策、建立政府数据开放平台等，致力于促进政府数据的开放、获取、共享和利用，极大地推动了数据的汇聚共享和政府透明度的提升，为社会创新注入新的活力源泉。

目前国际上的公共数据开放分为三种主流模式：政府主导模式、公众参与模式、政企合作模式。

政府主导模式：即指在数据开放活动中 由政府占据主导地位、进行统筹规划。政府从构建顶层设计着手、并在之后不断推出并完善相关的法律法规，自上而下层层推进，从中央到地方、从概况到具体，逐渐部署更加具体和 底层的法律法规与基础设施，最终建立起健全的公共数据开放体系。由于具有来自政府的强大支持，政府主导模式相比而言具有强大的发展动力，可以在

较短时间内使公共数据开放步入正轨，并以更快速度建立起完备的底层支持基础。

公共参与模式：由民众同心协力共同探索监督，自下而上推动政府实施。与政府主导模式截然不同，公共参与型数据开放模式由公众为主要驱动力。作为数据开放共同体，社会公众对政 府和公共机构对开放数据的需求更强，也因而具有更强烈的监督意识。因此，形成了由社会公众推动政府建设和实施数据开放的模式。与其他两种模式相比，公共参与模式中用户具有更高的数据开放参与度，机构与社会公众的协同互动、数据的宣传与利用也同样是模式中的重中之重，致力于鼓舞更多的主体投身于数据的开发利用。

政企合作模式：实现政府与企业的有利互补。政企合作型数据开放模式中，政府向合作企业或第三方机构开放数据，借助企业丰富的人才储备和技术水平开发出高使用价值的数据产品与服务，通过提供给社会公众实现商业化、获得经济效益，实现 政府、企业与民众的三赢。与前两种模式相比，政企合作模式既拥有政府引导的强进推力，又具备市场竞争赋予企业的蓬勃活力，实现有利互补，不仅可以让公共数据发挥出更大的价值，还可以实现该数据与企业数据之间的互联互通、促进整个数据产业的良性发展。此外，政企合作还有效解决数据确权问题、有效保障数据安全。

在国内，国家高度重视公共数据的开放共享，相关政策持续落地，持续探索公共数据的开放共享路径。目前国内的公共数据开放，主要分为政府公共平台开放和公共数据授权运营两种模式。

政府公共平台开放模式：是政府搭建公共数据开放平台，通过无条件自由开放或通过协议有条件提供等方式，向社会无偿开放公共数据，这种方式大大丰富了数据开发主体，但由于具有高价值的公共数据往往其承载的信息也较为敏感，难以实现最大范围的公共数据开放，且这种方式是由政府单向向社会开放数据，因缺少收益反馈而难以得到快速发展。

公共数据授权运营模式：通过授权运营，政府将具有高价值、但由于其承载的信息不适宜向社会开放的公共数据授权给可信的企业进行开发利用，从而在保护个人信息与国家安全不受侵犯的前提下，实现最大程度与最大范围的公共数据开放。公共数据授权运营的重点任务在于建设起完备的授权机制，构建健康的公共数据开放生态。公共数据授权运营具有许多优点，不仅可以促进公共数据的共享和开放、从而促进数据的再度利用和开发，还可以通过指导政府部门做出针对性的决策，提升服务的质量和效率。

公共数据的开放，目前在国际、国内均获得显著的成效。在国际方面，一些国际组织和跨国企业在推进公共数据的开放共享，如截止到 2023 年 6 月，新加坡政府开放数据平台，提供 1944 个数据集、9 个数据目录；美国官网数据超市，提供 265911 个数据

集；美国特拉华州数据开放平台，通过网络门户开放该州发生自2009 以来到半年以前的交通事故数据，数据来源于该州社会健康与服务部门（DSHS，Department of Social Health and Services）公开发布的事故报告，包含 45 万行，37 个数据项，该数据具有追溯性好、颗粒度细、可读性高的特点，每月更新一次。这些平台提供了大量的公共数据资源，促进了数据在各个领域的应用。在基础设施建设方面，各个国家和地区都在积极推进相关的建设，例如建设数据中心、云计算平台等，以促进数据的存储和传输。而在国内，截至 2022 年 10 月，我国已有 208 个省级和城市的地方政府上线了政府数据开放平台,其中省级平台 21 个（含省和自治区，不包括直辖市和港澳台），城市平台 187 个（含直辖市、副省级与地级行政区），目前，我国 74.07%的省级（不含直辖市）和 55.49%的城市（包括直辖市、副省级与地级行政区）已上线了政府数据开放平台,公共数据开放水平逐步提升。

三、问题与挑战

（一）合法合规

一是公共数据的识别与判定不清晰。虽然现阶段公共数据尤其是政务数据在概念上得到初步定义和区别，但是在识别范围上仍然存在一定模糊、难以区分的情况。现阶段政务数据在对外开放或授权运营过程中缺少上位法依据，数据供方在共享开放政务数据时存在一定合规风险，数据需方在获取与使用数据时也会面临一定的困境。

二是数据确权难，安全责任模糊。在公共数据应用与流通的场景下，数据来源也变得越来越多样化和复杂化，数据所有权与使用权分离致使安全责任模糊。公共数据作为特殊资产，其所有权和控制权的分离会导致在数据共享和交换过程中面临无法满足安全共享和开发的难题，一旦数据离开组织通过共享、交换等途径，其跟踪和溯源将变得困难。目前，尚未针对数据权属问题颁布相关的法律法规，因此作为数据使用者的组织需要落实数据管理义务，这是一个难点所在。此外，作为数据所有者的组织在数据共享和交换过程中将面临技术防护的“盲区”问题，需要采取相应的技术手段弥补安全漏洞。

三是如何规范公共数据合法合规利用问题。企业对公共数据进行加工、使用、发布信息时，存在因为公共数据的公共属性而放弃履行其保护义务的合规风险。过程中，数据运营商应当与相

关部门协商出协议机制，一方面对授权主体进行背景资质审查，确保授权主体的历史运营背景、数据安全资质、经营风险、外资风险等内容;另一方面应当通过协议表明该数据处理活动是基于当地政策安排和当地行政机构，约束授权主体身份、数据范围、数据处理目的等内容,明确各方在公共数据处理活动中的保护义务，对于重大负面敏感数据应履行更高的保护义务。

（二）数据安全

各参与主体应当以合法合规为基础,做好公共数据应用全生命周期场景下的安全保护,预防公共数据在应用与流转过程中因保护不当引发的数据泄露、破坏、窃取、滥用等安全事件。在公共数据安全保护方面当前面临的主要问题与挑战主要来自于如下几个方面。

1. 缺乏体系化顶层设计，各环节零散

随着国家加快培育数据要素市场的决策部署,数据资产作为重要的生产要素，蕴藏着巨大的商业价值，在数字经济时代帮助企业在经济活动中释放价值，通过数据的流通，实现业务间的贯通，以及多源的数据在新的场景中应用，使各行业在经营、服务和治理等环节的决策更智能与精准，伴随这个过程，各组织如何在安全前提下有序推动公共数据流转和应用的问题尚处于探索阶段，未形成统一的标准与共识。网络安全防护体系的管理对象流动性较弱且管控颗粒度可控,可通过各类管理技术手段实时掌控管理对象和管控现状，可以保证安全管理的覆盖度。数据安全

防护体系的管理对象具有极强的流动性且管控颗粒度极为细致，企业无法掌握全部数据资产，只能采用重点保护的方式明确管理场景，无法保证数据安全管理措施的覆盖率。公共数据的安全治理需要多主体多角色协作，区别于传统网络安全管理，数据安全可能会涉及法律、合规、业务、技术、人力等部门共同开展工作，因此组织需要建立可协调内外部多方资源的工作协调机制和管理体系。

公共数据流转和应用方式众多，数据接收方安全防护情况不可控，一旦数据脱离了组织环境，组织就不再能够通过技术手段保证自身数据安全，若发生安全问题，难以确定安全责任人，因此，建立一套可适用于不同数据共享开放需求、数据接收方的数据交换技术保护体系是各行业各组织的工作难点。

2. 保护过程繁琐，各环节实施细则待完善

首先，公共数据分类分级难以实施。有效的数据分类分级实施方案是确保公共数据使用范围合规、以及落实各类公共数据使用场景下的精细化安全保护的基础。虽然《网络安全法》、《数据安全法》和《个人信息保护法》在宏观层面上提出了数据分类分级的要求，但它们在具体规定方面比较简略。各地区目前也在探索公共数据分类分级的标准规范，但各参与主体需要自行根据相关法律规定，结合实际情况，制定具体的数据分类和分级方案，以确保公共数据保护的合规性。公共数据范围涉及信用服务、医疗卫生、社保就业、公共安全、城建住房、交通运输、教育文化、

科技创新、资源能源、生态环境、工业农业、商贸流通、财税金融、安全生产、市场监督、社会救助、法律服务、生活服务、气象服务、地理空间、机构团体、人民防空、应急管理等众多行业领域，现阶段部分行业已有数据分类分级规范，其相互之间如何衔接，如何避免冲突，如何体系化的构建构建全局公共数据分类分级规范机制，是需要持续研究和探索的重要任务。在实践落地层面，随着各企业或组织网络化、信息化和智能化水平的不断提高，业务系统的数据日渐庞大，一方面由于公共数据分类分级的行业和企业属性较强，目前国家和行业配套的标准体系仍在制定或征求意见的过程中，细则不完善，不足以支撑落地数据分类分级工作的全行业落地；另一方面，很多未发布行业数据分类分级标准的企业担心自发的数据分类分级建设不符合后期发布的标准和要求，需要重新调整，抑制了各单位落地数据分类分级工作的积极性。

其次，公共数据安全风险评估实践过程困境重重。公共数据的应用涉及多元化的各类主体角色和复杂的使用场景，其面临的数据安全风险也在不断加大。现阶段实施数据安全风险评估的困境主要来自于如下方面。一是数据安全风险评估标准尚不完善，评估实践缺乏可落地的方法论支撑。传统的信息安全风险评估主要是面向网络环境下的数据安全载体资产，基于某个标准作为基准来设置评估项，展开相对静态、固化的风险评估，而数据资产具备流动性，会在不同的载体和场景环境下流动，且其资产价值

也依据量级、周期性等因素动态发生变化，因此传统的信息安全风险评估无法顺应数据流动过程中不同环境、不同目标下的安全评估要求。二是，公共数据的应用模式和场景复杂。数据安全风险评估是一个定制化的过程，其核心是基于业务场景展开评估，传统的风评模板难以适用于不同行业和业务场景中的风险评估，因为其评估侧重点和评估流程都不同。在制定评估流程和检查项时，需要结合业务场景的具体特点，包括公共数据的敏感性、涉密性、完整性等方面进行考虑，同时还需要考虑到数据的生命周期管理、数据安全保障、应急响应机制等方面。

3. 保护技术需要加速迭代

数据要素市场背景下于对公共数据应用需求在不断地更新迭代，包括新技术的应用，新场景的建立，也都会带来新的威胁。数据安全技术起步较晚，碎片化严重，大部分技 术尚不成熟，距离大规模落地应用仍有差距。公共数据在共享开放过程中新场景、新领域、新模式 不断涌现，需要新型的数据安全技术产品提供支撑和保障。

4. 缺乏数据安全专业人才

随着数字经济的飞速发展，数据安全的人才培养与人才队伍建设也已经上升到了国家战略层面。现阶段，我国网络及数据安全人才的培育力度、人才的输出数量和质量与迅猛增长的人才市场需求还存在一定的差距，人才需求与供给的失衡，导致大多数组织缺乏数据安全专业人才，以及一线人员缺乏数据安全意识。

从决策层到技术层，从业务部门到 IT 部门，再到合规部门，从管理制度到技术支撑，数据安全人才需求贯穿企业各个部门，数据安全工作呈现出了高度的复合性，对数据安全领域下各从业人员提出了更高的要求。在专业知识方面，要求从业人员将业务知识、合规知识和安全知识融会贯通。在实践经验方面，新技术新场景带来的新威胁、数据安全产品与技术的不断迭代、数据安全保护与业务发展的紧密结合，使得工程和实践经验成为用人单位选才的重要标准。

（三）质量保障

1. 公共数据质量要求

公共数据质量是指公共领域中所使用的数据的规范度、准确性、完整性、一致性和时效性等方面的特征。涉及到数据的来源、收集、处理、存储和传输过程中的各种因素和环节，以确保数据能够有效地支持决策制定、政策实施、研究分析和公众信息需求的满足。衡量公共数据质量的维度包括:

规范性：公共数据应符合数据标准、数据模型、业务规则等程度。

准确性：公共数据应遵循数据标准及规范的要求，正确描述实际信息，数据记录的信息不存在异常或错误。

完整性：公共数据应全面、完整、无缺失，覆盖各项业务和管理数据。

一致性：公共数据与共享交换平台中的数据或不同公共数据提供单位所提供的同类数据之前无矛盾程度。

时效性：开放数据及时有效的程度。

2. 公共数据质量挑战

数据质量管理范围未界定。公共数据作为一种数据的特定类型，是由多种社会主体产生的，在数据终端上以多种数据格式进行储存，导致公共数据在其规范性方面存在原生性的不足。其管理范围在金融行业缺乏统一规范，没有明确定义管理和维护公共数据质量的界限。在其采集阶段易出现完整性方面的问题。部分公共服务机构在采集数据时的流程不规范，导致数据录入不完整或关键数据丢失等数据完整性问题，依靠不完整的公共数据无法分析得出全面的社会或经济发展信息，影响数据结果的全面性。其缺乏明确数据质量检测规则致使错误数据的决策，从而影响业务流程和战略计划。

数据质量控制不明确。公共数据来源复杂，来自不同部门、系统和渠道，采集、处理和使用时通过的数据流程过多，公共数据的处理和储存周期过长，且存在对同样数据的差异化解释，阻碍了公共数据的信息一致化进程，导致数据应用时的信息混乱，出现部分公共数据存在数据错误等数据准确性方面的问题，直接影响数据应用结果的准确性。实现数据生命周期各阶段的质量控制要求对于管控难度高，需要覆盖采集、传输、存储、处理、使用、退役等过程，形成闭环的数据质量控制。没有数据标准及规

范进行设计开发，数据在产品或系统中可能无法满足外部监管和业务管理的需要。

数据质量检测不统一。公共数据难以有效利用的关键是无法在多区域、多部门间建立统一监控机制。当前公共数据生态系统复杂，包含多系统、多平台、多来源，同时数据结构各异。难以形成全方位覆盖。考虑到数据项的重要程度和更新频率，部分应用场景需要实时质量检测，无法满足监测方式准确的同时一旦发现数据质量问题，迅速确定根本原因并采取解决措施。最后，不断变化的公共数据业务需求导致现有数据质量监控规则难以满足后期检测要求。规则的复杂性和变化性使得原有规则无法适应不同场景，需要及时调整和优化检核规则以适应多变的业务环境。

数据质量体系不健全。在组织保障方面，数据质量事宜需要获得足够的资源和支持。责任分工模糊不清，缺乏明确的数据质量负责人，导致问题处理时的拖延和混乱。此外，对数据质量重要性的培训和意识不足，影响了组织成员对数据质量的认知和积极参与。

在制度方面缺乏明确的数据质量标准和指南，导致不同部门在数据质量要求上存在差异，从而影响数据的一致性和可信度。其次，数据质量评估机制不完善，缺乏定期的检查和验证，使得潜在的质量问题难以及时发现和解决。

四、公共数据在金融领域应用情况

（一）概况

金融机构各分支机构积极与各地政府机构等合作，引入公共数据，赋能业务发展。公共数据具有“本源权威”的天然属性，且数据价值巨大，合法合规性也有保障，但由于政府机构信息化程度及数据集中程度等原因，很多数据只能通过“分对分”的方式从地方引入。截至目前，金融机构通过总部统筹对接的方式，引入工商、社保、知识产权等多项政务数据，各分支机构接入公积金、社保、税务、不动产、电力、燃气等多种公共数据，数据引入模式除传统的实时联机查询和批量传输之外，联合建模方式也越来越受政府部门推崇，尤其是国家重点业务领域，比如普惠金融、乡村振兴金融、双碳战略、科技金融等等。

为合法、合规、安全应用公共数据，各金融机构在公共数据计入和应用过程中，多样化的数据接入能力、隐私授权保护能力和数据整体分析能力。

金融行业已实现多样化数据接入能力，包括实时联机查询、批量数据传输、联合建模、隐私计算等方式接入外部数据。尤其是联合建模和隐私计算方式是解决当前政务数据共享应用的有效手段。

实时联机查询是指数据使用方通过实时调用数据提供方发布的 API 接口查询相关数据的方式，适用于外部数据嵌入到业务

流程或者模型中的外部数据，尤其是需要取得客户授权同意的个人和企业敏感数据；批量数据传输是指双方通过约定的格式、频度等数据提供方将数据批量传输给数据使用方的模式，对于公开的政务数据可以采用这种模式。

联合建模方式：以业务需求为导向，成立专门的联合建模团队，与政府合作，由合作双方评估可行性后，制定联合建模方案，共同提供建模数据，在相关建模环境中建立数据模型，形成最终模型报告，在联合建模环境中完成产品部署，客户授权等工作，避免了数据传递而产生的各项问题，“可用不可见”应用政务数据。

隐私计算方式：是借助联邦学习、隐私求交、匿踪查询等技术，在保证合作方均不泄露原始数据的前提下，支持多方数据的联合分析应用，实现了数据不出域的数据共享模式。

（二）数据接入能力

金融机构接入公共数据主要包括普通接口对接，联合建模，隐私计算等方式。

普通接口对接是指在不同的系统中，通过规定的接口来实现数据交换。一般常用的 API 接口可分为库/程序、框架和 Web 服务三类，在数据交互中数据提供方将原始数据加工成可输出的标签形式进而通过接口方式输出给数据需求方。接口数据对接的优点是可以更好地实现数据交换，降低系统之间的数据交互成本，也可以实现多个系统之间的数据高效共享。

联合建模一般是指数据提供方提供安全空间，原始数据部直接输出，各数据需求方将自身标签样本数据以脱敏等安全措施携带出域，而后在安全空间中进行数据交互，匹配特征数据来联合开发模型或查询运算。是指原始数据不直接输出，数据需求方提供标签样本数据，去匹配特征数据来联合开发模型，开发出来的模型则部署在数据提供方，只以模型预测结果的方式来输出数据。联合建模的优点是可以通过线下数据融合的方式更深入挖掘数据特征。

隐私计算则是指基于联邦学习、多方安全计算、可信执行环境等主流隐私计算技术，在保护数据本身不对外泄露的前提下实现数据的联合建模、联合统计分析等，达到对数据“可用不可见”的目的，在充分保护数据和隐私安全的前提下，实现数据价值的转化和释放。基于隐私计算技术，可实现基于联邦学习的联合建模、基于多方安全计算的联合统计分析、基于隐私信息检索技术的匿踪查询、基于隐私集合求交的安全求交等能力。隐私计算的优点是基于底层密码学技术实现线上数据交互，可有效解决跨机构、跨地域数据的安全融合问题。具体技术领域可分为联邦学习、多方安全计算和可信执行环境。

联邦学习(Federated machine learning/Federated Learning)，又名联邦机器学习，联合学习，联盟学习，被认为是解决数据安全与人工智能两难问题的一个重要技术。联邦学习是一个机器学习框架，能有效帮助多个机构在满足用户隐私保护、

数据安全和政府法规的要求下，进行数据使用和机器学习建模。联邦学习作为分布式的机器学习范式,可以有效解决数据孤岛问题,让参与方在不交换原始数据的基础上联合建模,能从技术上打破数据孤岛,实现AI协作。

多方安全计算（Secure Multiparty Computation）是解决多源数据隐私保护问题的另一个重要方法。多方安全计算是指在无可信第三方情况下，通过多方共同参与，安全地完成某种协同计算。即在一个分布式的网络中，每个参与者都各自持有秘密输入，希望共同完成对某个函数的计算，但要求每个参与者除计算结果外均不能得到其他参与实体的任何输入信息。也就是参与者各自完成运算的一部份,最后的计算结果由部分参与者掌握或公开共享。也就是说多方安全计算技术可以获取数据使用价值，却不泄露原始数据内容，保护隐私。因此在大数据发展并立法保护的今天，多方安全计算技术从多年的学术研究，变得具有巨大的商业价值。

可信执行环境（Trusted Execution Environment）是计算平台上由软硬件方法构建的一个安全区域,可保证在安全区域内部加载的代码和数据在机密性和完整性方面得到保护。其目标是确保一个任务按照预期执行,保证初始状态和运行时状态的机密性、完整性。相较于其他基于密码学的隐私计算技术，TEE的优势在于性能强大，能够解决复杂场景的业务诉求，同时开发成本低，能够迅速做出满足业务诉求的产品。

（三）数据安全保障

金融数据是金融体系的重要组成部分，应对金融数据安全风险挑战，防范化解金融数据安全风险，有利于更好发挥金融数据价值，实现金融业的数字化转型。为保障数据安全工作得到有效落实，需要建立一套数据生命周期安全框架。数据生命周期安全框架遵循数据安全原则，以数据安全分级为基础。数据生命周期安全防护要求是数据生命周期安全框架的核心，针对不同安全级别的数据，明确在采集、传输、存储、使用、删除以及销毁等数据生命周期各个环节的安全防护要求，这是金融业机构开展数据安全防护工作的基本依据。安全组织保障是确保数据安全相关工作的组织和落实奠定基础。在金融业机构日常运营过程中，信息系统运维过程的数据安全防控工作也不容忽视。数据安全组织保障、信息系统运维保障共同构成了数据生命周期安全防护机制能够有效落实和严格执行的基石。通过建立覆盖金融数据生命周期安全框架是确保金融数据安全的必经之路。

1. 实施数据资产安全分级管理

在日常金融活动中，会产生大量金融数据，什么样的数据在什么环节需要什么样的保护。首先要梳理需保护的数据资产目录，在此基础上根据其价值及敏感程度建立安全等级清单，为后续落实数据保护措施夯实基础。考虑到整体经营活动中面临着各式各样的数据，各部门在梳理时难免口径不一，需要建立统一的数据

资产安全等级标准,在各部门的协同配合下形成涵盖业务、运营、财务、科技、人事、合规等各类型数据资产安全等级清单。

（1）明确数据安全分级原则

为保证数据的可用性，保障数据使用的便捷和管理安全，对数据安全进行分级，采用精细的差异化管理，应依据中国人民银行发布的《JRT 0197-2020 金融数据安全数据安全分级指南》中的建议遵循以下原则:

1） 合法合规性原则: 满足国家法律法规及行业主管部门有关规定。

2） 可执行性原则: 数据定级规则避免过于复杂，以确保数据定级工作的可行性。

3） 时效性原则: 数据安全级别具有一定的有效期限，金融业机构宜按照级别变更策略对数据级别进行及时调整。

4） 自主性原则: 结合金融业机构自身数据管理需要（如战略需要、业务需要、风险接受程度等），在本标准的框架下自主确定数据安全级别。

5） 差异性原则:根据本机构数据的类型、敏感程度等差异，划分不同的数据安全层级，并将数据分散至不同的级别中，不宜将所有数据集中划分到其中若干个级别中。

6） 客观性原则: 数据定级规则是客观且可校验的，即通过数据自身的属性和定级规则即可判定其级别，并且数据的定级是可复核和检查的。

（2）明确数据安全分级范围

依据中国人民银行发布的《JRT 0197-2020 金融数据安全数据安全分级指南》金融数据安全定级过程中安全定级工作所涉及的金融数据包括：提供金融产品或服务过程中直接（或间接）采集的数据，包括通过柜面以纸质协议签署或收集，并经信息处理后在计算机系统中流转或保存的数据，以及通过信息系统签约或收集的电子信息。金融业机构信息系统内生成和存储的数据，包括业务数据、经营管理数据等金融业机构内部办公网络与办公设备（终端）中产生、交换、归档的电子数据，如机构内部日常事务处理信息、政策法规与部门规章、业务终端临时存储的业务或经营管理数据、电子邮件信息等。金融业机构原纸质文件经过扫描或其他电子化手段形成的电子数据。

2. 建立数据安全管理组织架构

全国人大通过的《数据安全法》以及中国人民银行发布的《JRT 0223-2021 金融数据安全数据生命周期安全规范》，均提出保障数据安全需建立层次清晰、职责明确的完整组织架构。搭建起健全的组织架构可以更好地明确数据安全管理过程中各部门的协作机制和分工，为数据安全管理提供组织保障。

（1） 明确数据安全管理组织岗位设置

建立自上而下的覆盖决策、管理、执行、监督四个层面的数据安全管理体系，明确组织架构和岗位设置，保障数据生命周期

安全防护要求的有效落实。按照中国人民银行发布的《JRT 0223-2021 金融数据安全数据生命周期安全规范》要求:

应设立由金融业机构高级管理层组成的领导小组，由机构高级管理层构成数据安全管理的决策层。科技、业务、法律合规、风险管理等部门主要负责人构成数据安全管理的管理层。科技、业务、法律合规、风险管理等部门具体数据安全岗位相关人员构成数据安全管理的执行层。审计、稽核等部门相关人员构成数据安全管理的监督层。

（2） 明确数据安全管理组织工作职责

建立统一的金融数据安全管理架构后，还需明确各层级部门与相关数据安全工作职责，规范工作流程。据安全管理的决策层总体负责数据安全工作的统筹组织、指导推进和协调落实，明确数据安全管理部门，协调机构内部数据安全管理资源调配。数据安全管理的管理层负责数据安全相关工作的实施、相关政策和制度的制定评审工作，保障数据安全管理工作所需资源，并设立数据安全管理专职岗位，负责日常数据安全管理工作。数据安全管理的执行层负责根据数据安全相关策略和规程，落实本部门数据安全防护措施。对本部门数据脱敏、对外提供等关键活动的数据安全控制有效性进行确认。数据安全管理的监督层负责根据本机构数据相关业务实际情况，确定相应审计策略及规范包括但不限于审计周期、审计方式、审计形式等内容。

（3） 提升数据安全参与人员能力

《数据安全法》指出组织开展数据安全教育培训，采取相应的技术措施和其他必要措施，保障数据安全。通过教育培训的方式，提升专业人员技术能力。在数据安全教育培训方面，金融机构应积极落实全员数据安全相关法律法规、监管要求、数据安全专项的宣导工作，积极开展针对专业人员的技术宣导和培训，包括国密算法、APP 隐私保护、敏感信息探测等内容。

（4） 实施数据安全绩效考核机制

数据安全需要所有人的参与和落实，为了考量不同岗位在数据安全工作上的参与度和落实情况，需要设计一套针对不同数据安全职责的绩效考核机制，在员工整体考核中占一定比例，以鼓励员工主动参与和配合相关工作。考核内容包括但不限于：考核数据安全整体实施成效、数据安全合规程度以及资源保障情况；数据安全分级、数据安全管理控制手段及技术防护体系建设情况；数据安全要求实际落实执行情况，包括数据安全等级完整性、数据合规使用情况等。

3. 建设统一数据安全运维保障

数据安全有管理要求，也需要有运维保障规范来保障相关要求可以遵照落实，考虑到数据全生命周期涉及的链路冗长，其中涉及的安全技术保障环节也众多，金融机构在落实过程中应注重统一标准，避免安全防护的篱笆此长彼短，产生木桶短板风险。通过建设一整套的数据安全运维保障规范，实现统一安全服务与管控，保障安全措施有效落地。

（1） 加强边界管控

数据安全中的边界管控是一种管理和控制数据流动的方法，旨在确保只有经过授权的用户、系统或实体可以访问、传输或处理特定的数据。这种控制是通过设置边界来实现的，边界可以是物理的、逻辑的或虚拟的，用于隔离不同的网络、系统或环境，以及限制数据的传输和访问。按照《JRT 0223-2021 金融数据安全数据生命周期安全规范》如下要求：应在内网边界按照“最小权限”原则严格控制外部机构的访问权限。互联网区和外联接入区为不可控区域，应在内部可控区域与不可控区域之间进行隔离，并根据应用需求和数据传输需要逐一开通访问关系，默认为禁止访问。应避免将重要网段部署在网络边界处且直接连接外部信息系统，重要网段与其他网段之间应采取技术手段进行隔离。应明确生产网络接入和数据传输接口开通相关审批流程。通过边界管控有助于减少数据泄露、未经授权访问以及其他安全风险。

（2） 强化访问控制

数据安全的访问控制是一种控制和管理谁可以访问特定数据的方法。它旨在确保只有经过授权的用户或实体可以获取、查看或处理特定数据，以防止未经授权的访问、数据泄露和滥用。访问控制策略包括物理环境访问控制、信息系统与介质访问控制、数据存储系统的访问控制，访问控制策略应建立面向数据应用的安全控制机制，包括访问控制时效的管理和验证，以及应用接入

数据存储的合法性和安全性取证机制，宜建立基于用户行为或设备行为的数据存储安全监测与控制机制。

（3） 加大检测力度

金融业机构宜具备数据溯源能力，对数据生命周期过程中数据的采集、查询、修改、删除、共享等相关操作进行跟踪，通过留存金融数据流动记录等方式，确保金融数据相关操作行为可追溯。并且应建立日常数据泄露、数据篡改、数据窃取、数据非法使用的风险监控机制，主动预防、发现和终止数据泄露异常行为，有效防范和化解风险。

（4） 建立审计机制

数据安全中的安全审计是指对系统、应用程序、网络和数据访问活动进行系统性检查和记录的过程。它旨在监测和追踪谁、何时、以何种方式访问了敏感数据，以及对系统的其他安全事件进行记录和分析。金融业机构应记录数据操作行为日志，并针对日志进行审计分析，识别并告警可疑行为，审计方式。安全审计对于维护数据的机密性、完整性和可用性，以及检测和预防潜在的安全问题非常重要。

（5） 优化评估方法

数据安全的检查评估是一种系统性的方法，用于评估组织的数据安全状况，识别潜在的风险和漏洞，并提出改进建议。这种评估可以帮助组织了解其数据安全的薄弱环节，从而采取适当的措施来提升数据的保护水平。金融业机构应建立数据安全检查评

估机制，定期制定数据安全检查评估计划。在国家及行业主管部门的相关要求发生变化时，或在业务模式、信息系统、运行环境发生重大变更时，或发生重大数据安全事件时，应进行数据安全评估。数据安全检查时宜采取多种形式，如自查和外部检查等，执行管理和技术并重的检查原则，并通过技术工具对相关管理检查内容进行验证和确认。

（6） 应急响应与事件处置

在数据安全中的应急响应与事件处理至关重要，它可以帮助组织迅速应对和缓解安全事件，从而减少潜在的损害和影响。金融业机构制定应急响应与事件处置规范，建立完善的应急响应与事件处置和问责机制，做好应急预案，组织应急演练，确保在紧急情况下重要信息资源的可用性。发生金融数据泄露事件时，金融业机构应及时采取补救措施，及时启动应急响应机制。事件处置结束后，应分析和总结原因和存在的问题，形成调查记录和事件清单，调整数据安全策略，避免事件再次发生，并形成总结报告。

（7） 驱动技术保障

金融机构在完善合规流程基础之上，还需要配合引入相关前沿科技，是提升流程效率、打破数据保护与价值流转内在取舍、打通业务创新瓶颈的核心对策。包括但不限于通过以区块链、隐私计算为代表的前沿技术，来驱动数据安全体系的升级。借助分布式进行可信分级授权治理、数据最小化披露和全密态数据计算

等，以此有效减少事后追责成本，显著提升数据安全流程可行性和实施效率，促进数据生产要素价值的发掘创新，进一步的将技术手段来保障数据全生命周期涉及的全链路。

4. 加强数据安全生命周期管理

在数据收集、存储、使用、加工、传输、提供、公开的各流程环节中都存在数据安全问题，相应地在业务流程的每一个环节都应落实数据安全管理责任。《数据安全法》也明确提出要建立健全全流程数据安全生命周期管理。为保证数据全生命周期的安全性，在数据的整个生命周期中，需要对数据的采集、传输、存储、使用、共享、销毁等过程进行安全管理，同时保障数据的机密性、完整性、可用性和合规性。安全管理需要涵盖数据的所有环节，防范数据泄露、损毁或其他安全事件的发生。数据全生命周期安全管理的关键是要制定相应的政策和措施，并落实执行。银行金融机构可以参照《JRT 0223-2021 金融数据安全数据生命周期安全规范》，构建并不断完善自身的数据安全生命周期控制矩阵，对于数据从采集、传输、存储、使用、共享、销毁等全生命周期节点进行详细控制。

（1） 实施数据采集的安全管控措施

在数据采集过程中应针对具体关键业务场景制定数据分类分级的实施原则。在数据采集安全管理方面，为防止采集过程中个人信息和商业数据的滥用，需要信息主体授权，并应当依照法律法规和与用户约定来处理其相关数据，另外还应在数据应用和

数据安全保护间寻找适度的平衡。数据鉴别及记录方面，为防止采集到非法数据源产生的数据或失真的数据，需要对采集的数据需要进行数据来源的标识，以便在必要时对数据源进行追踪和溯源。在数据质量管理方面，需要保证对数据采集过程中收集和产生的准确性、一致性、完整性。

（2） 实施数据传输的安全管控措施

在数据传输加密过程中，为防止数据在传输过程中发生数据被窃取、伪造和篡改等安全风险，需要建立相关加密措施来保障数据在传输过程中的机密性、完整性和可信任性。在网络可用性管理方面，为避免一旦发生网络故障或者瘫痪，数据传输也会受到影响甚至中断。因此，需要建设高可用性的网络，从而保证数据传输过程的稳定性。

（3） 实施数据存储的安全管控措施

在存储介质安全方面，为避免数据存储介质损坏和不当使用造成安全风险，应对数据存在的物理实体介质，和虚拟存储介质等进行安全管理。在逻辑存储安全方面，针对存储容器和存储架构的安全要求，比如认证鉴权、访问控制、日志管理、通信举证、文件防病毒等建立必要的安全配置和安全配置策略，以保证数据存储安全。数据备份和恢复过程中，为提高信息系统的高可用性和灾难可恢复性，在数据库系统崩溃的时候，通过对数据的建立备份和恢复策略和部署相关工具，确保数据的可用性。

（4） 实施数据处理的安全管控措施

在数据脱敏方面，可以通过将敏感数据进行数据的变形，为户提供非真实数据，这样可以在开发、测试和其它非生产环境以及外包环境中安全地使用脱敏后的真实数据集，既保护了组织的敏感信息不泄露，又达到了挖掘数据价值的目标。在数据分析安全方面，可以通过在数据分析过程采取适当的安全控制措施，防止数据挖掘、分析过程中有价值信息和个人隐私产生泄露的安全风险。为保证数据正当使用，应当建立数据使用过程中的相关责任和管控机制，以保证组织内部数据的正当使用，防止内部合法人员被数据的价值吸引而滥用有价值的敏感数据。

（5） 实施数据交换的安全管控措施

由于数据导入导出广泛存在于数据交换过程中，而且一般数据导入导出的数据量都很大，因此相关安全风险和安全危害也会被成倍放大。因此，需要采用有效的制度和工具控制数据导入导出的安全风险。为保证数据共享安全，防止数据在共享后产生丢失、篡改、假冒和泄露的风险，需要采取安全保护措施保障数据共享后数据的完整性、保密性和可用性。在数据发布时，应建立安全保障措施，确保数据发布内容(企业宣传、网站内容发布、社交媒体发布、开放数据、对外宣讲材料等)的真实性、正确性、实效性和准确性。

（6） 实施数据销毁的安全管控措施

在进行数据销毁处置时，组织要根据不同要求和需要采取不同的数据销毁策略和技术手段，已实现对数据的有效销毁，防止

因对存储介质中的数据内容进行恶意恢复而导致的数据泄漏风险。在对介质销毁处置时，应对存储介质进行彻底的物理销毁，保证数据无法复原，以免造成信息泄露。

（四）数据质量保障

金融领域数据保障的意义在于帮助金融机构降低风险、提高决策的准确度。直接关系到金融机构的决策制定、风险管理和客户信任，保障所使用的市场数据、经济数据、财务数据等准确、一致、完整且可靠。为有效保障数据质量，可以通过构建适合组织管理的治理框架、一套数据质量保障体系、制定制度流程、优化信息系统以达到公共数据质量的规范要求。在保障数据质量的基础上，进一步提升客户服务质量，加大金融业务经营管理支撑，实现金融数据资产的利用最大化。

1. 明确数据质量要求

（1） 明确数据质量管理范围

数据质量管理范围包括数据质量源头管控、数据质量检核规则管理、数据变更管理、数据质量问题管理、数据质量监控与检查、数据质量评价、数据质量管理工具等内容，覆盖数据生命周期全流程，使得公共数据在其整个生命周期里满足金融领域，包括商业银行、证券等内外部管理的需要。在数据质量管理范围内涉及的数据项，包含但不限于基准数据、指标数据、针对外部监管要求，整理公开数据中的必填数据项、业务决策和分析的重要数据项。

在需求分析阶段，深入分析公共数据需求，明确数据质量管理范围，与需求提出部门确认业务数据项清单，与 IT 实施人员确认技术数据项清单。识别关键业务流程和数据来源，包括决策制定、风险管理和客户服务等方面。

（2） 明确数据质量检核规则

要求制定数据质量衡量指标，包含但不限于数据规范性、完整性、唯一性、准确性、时效性、一致性等指标定义。数据质量管理范围涉及的数据项，要求与需求提出方，明确数据质量检核规则，包含但不限于对应的数据质量指标、数据质量控制方式和规则、监测方式和规则等。依据数据标准及其他业务需求，参考数据分类和认责体系梳理业务数据项，并根据质量的特性分类，制定形成数据质量业务规则，进一步将业务规则技术化，进而生成技术规则。

此外，金融机构在制定数据质量检核规则时需要考虑法律法规和监督要求，满足数据隐私和保护法规，以 ISO 8000 标准作为参考，进一步实现指标定量和定性，如数据准确性百分比、数据缺陷百分比和数据源可靠性、数据更新的及时性。

2. 落实数据质量控制

（1） 定义数据质量控制方式

根据制定完成的金融公共数据质量检核规则，细化定义数据质量校验或控制，包含数据质量控制方式和监测方式。防止“数据越权、数据欺诈、数据作假”等违规现象。

（2） 加强数据源头质量管控

在公共数据采集录入阶段，加强数据源头质量管控，从系统录入角度、指标数据血缘分析角度、定期数据质量检核、数据安全等角度，落实数据质量要求。分析数据质量规则和数据质量控制方式，理解业务对数据质量控制的相关需求。根据数据质量规则表中的控制方式，在系统中开发相应的功能以实现对数据质量的过程控制。

（3） 落实数据生命周期质量要求

从数据采集、流转、整合、应用等数据生命周期阶段落实数据质量管控要求，明确数据创建、加工、流转和存储各环节的数据质量控制方式，形成数据质量控制闭环。

1） 数据采集质量

建立数据创建阶段的控制措施，综合考虑数据创建的业务场景、数据质量规则、取值范围等，明确录入界面控制等控制方式。

满足数据系统化建设、数字化管理，数据目录设立、数据分类分级等方面的质量要求。

2） 数据传输质量

建立数据流转的质量验证和控制措施，要求明确接口文件控制和文件传输控制方式，保证数据文件及内容的质量，避免错误数据加载，保障数据传输质量。

满足数据安全稳定传输、数据异常检测、数据流程续传等过程控制方式，以及数据规律更新等方面的质量保障要求。

3） 数据存储质量

建立统一规范编制数据资源目录，根据数据资源共享方式分层，分级存储到相应的共享交换平台，支持结构化与非结构化数据存储以及批量、交互查询、实时流、内存计算等大数据场景。

满足存储备份机制、数据库系统、数据存储备份要素，以及存储备份功能可靠性、安全性、时效性等方面的质量要求。

4） 数据处理质量

建立数据处理的质量验证和控制措施，明确加工、整合数据过程的控制方式，保证不丢数、不失真、不变坏，应在数据质量问题产生点对数据进行修改、删除操作，遵循但不限于满足数据清洗规则、数据加工范围和程序等方面的相关质量要求。

5） 数据使用质量

要求在满足数据服务安全、共享的前提下开展数据应用，满足可收集、可开发、可使用、可交互等方面质量要求。

6） 数据退役质量

建立数据退役的质量控制措施，要求保障数据退役前后系统的业务延续性、异常情况下的数据可恢复性。

执行数据备份，数据清理，应急恢复等情景，遵循确保生产系统稳定、高效运行及对外服务影响最小化的原则，保证不破坏业务数据的完整性和一致性。符合数据清理规则的，具有关联关系的业务数据应一并清理，避免出现孤岛数据。

（4） 强化项目实施数据质量管控

系统验收测试阶段，应将数据标准要求、数据质量要求等作为测试案例的重要组成部分，验证数据质量源头管控的有效性。及时执行测试，验证质量检核有效性。要求应用系统能自动提示数据项异常变动及错误情况，发现并提升其数据质量问题。

遵循数据标准及规范。在产品或系统的设计开发中，数据除满足外部监管要求及业务管理需要外，还应符合金融业企业级数据字典等标准及规范的要求。进行源头数据质量控制、数据加工质量控制和监督控制。

3. 数据质量监控与检查

（1） 建立数据质量监测机制

通过技术方式，建立数据质量监测机制，形成系统化、线上化、可量化的数据质量管控。包含但不限于以下要求：

数据质量监测设计要求：根据数据项重要程度、更新频率等，确定对数据项的监测方式。根据数据质量规则和数据质量控制方式，确定数据质量监测方式。确定数据质量监测结果的展现方式，包括报表形式、报表内容等；同时，确定数据质量监测的实现方式，包括应用系统内实现数据质量监测功能、实现数据质量检查和控制、执行数据质量规则运行和监测等。

数据质量监测执行要求：对预先定义的数据质量规则进行分析，根据不同环境的监测对象，按需生成数据质量验证报告，发现数据质量问题。

（2） 妥善处置数据质量问题

应用系统应自动提示数据项异常变动及错误情况，发现并提升其数据质量问题。找出发生数据质量问题的根本原因并采取相关的策略进行解决。

确定根本原因：确定引起数据质量问题的相关因素，并区分它们的优先次序，以及为解决这些问题形成具体的建议。

制定和实施改进方案：最终确定关于行动的具体建议和措施，基于这些建议制定并且执行提高方案，预防未来数据质量问题的发生。

（3） 定期优化检核规则

在开发新应用系统时，技术人员应根据系统的业务层面的数据质量检核规则，针对技术部分补充检核要求，并落实数据质量检核控制内容。发现数据质量问题进行分析解决时，若定位是数据检核规则不完善的，技术人员应对数据质量检核规则进行优化。在发生应用系统数据结构变化、业务规则变化等情形时，需求提出方应对业务数据项及时提供新的校验规则，技术人员应及时根据新的校验规则更新数据质量检核规则。

4. 厘清数据质量保障管理范围

（1） 确定数据质量管理组织职责

确定数据质量保障最高决策组织，设立并明确相关部门（或组织）及其职责，包括但不限于：

数据质量部门、组织及其负责人，主要负责统筹、规划数据质量保障工作。

数据质量部门下辖数据质量团队（或组织）及其负责人，主要负责数据质量相关工作的组织、协调、管理、审核、评审等工作。

信息科技部门及其负责人在数据质量保障工作中的角色，主要负责制定数据质量战略和目标，数据质量保障与机构的整体战略一致，将数据质量纳入企业文化中。

业务部门及其负责人在数据质量保障工作中的角色，主要负责落实数据质量有关要求，推动数据质量培训，提高数据质量意识和技能。

其他相关部门在数据质量保障工作中的角色、职责及负责人。实施数据质量监控体系，实时跟踪数据质量情况，及早发现和解决问题。

（2） 明晰数据质量保障制度要求

建立数据质量管理制度，明确数据质量的标准、流程、责任并落实相关工作要求，确保数据质量管理的系统性和规范性，包括但不限于:

数据质量保障的目标和原则。

数据质量保障工作涉及的角色、部门及相关职责。

数据质量保障的方法和具体要求。

数据质量保障的日常管理流程和操作规程，以及数据质量保障结果的确定、评审、批准、发布和变更机制。

数据质量管理相关绩效考核和评价机制。

数据质量保障结果的发布、备案和管理的相关规定。

（五）业务应用场景

在中国经济走向新常态的转型中，在中国金融业向支持实体经济、创新驱动的转换中，公共数据能够充分利用金融市场形成的海量数据来更好地服务金融业的发展，释放出强大的生命力。公共数据在金融领域有广泛的应用场景，主要包括以下几个方面：

1. 精准获客

当前商业银行在营销获客的过程中采集的客户信息并不全面，基于银行自身拥有的数据有时候难以得出理想的结果甚至可能得出错误的结论。公共数据则拥有综合政务办理数据、税务数据、社保公积金缴费数据等，可以有效补充金融机构所不掌握的客户行为特征。因此金融机构不仅仅要应用自身业务所采集到的数据，还应考虑整合更多维度的公共数据，以丰富客户画像，从而达到精准营销的目的。通过接入的政府公共数据以及海量第三方数据可以丰富客户的精准画像，从而对客户实现精准分层，制定更加精准和有针对性的营销策略，更有针对性地提供符合客户

需求的金融产品和服务，为获客准入的初筛提供参考，提升金融机构整体的盈利水平和运行效率，如图 1 所示。

图 1 公共数据在精准获客场景应用示意图

2. 流程优化

近年来线上化金融服务成为了大趋势，金融机构借助公共数据实现流程优化和自动化，通过分析业务流程，对整个运营模式进行重构和自动化处理，进一步提高金融机构的运营效率和服务质量。以对公客户融资贷款流程优化为例子，工商数据包含对公客户开户时所需的企业要素信息，在企业申请融资贷款流程中，数据录入方式由人工转为自动化，改变耗时长易出错的情况，实现客户信息各栏位的自动输入，提高业务办理效率，改善客户、员工体验，提高客户的满意度，同时规范信贷管理，提升了风险管理水平。除支持柜面开户外，还支持客户线上渠道预约办理。同时，还与国家市场监管总局对接，接入了全国范围的电子营业

执照数据，实现了在线查验电子营业执照辅助业务办理的功能，借助公共数据真正实现金融业务流程优化。如图 2 所示。

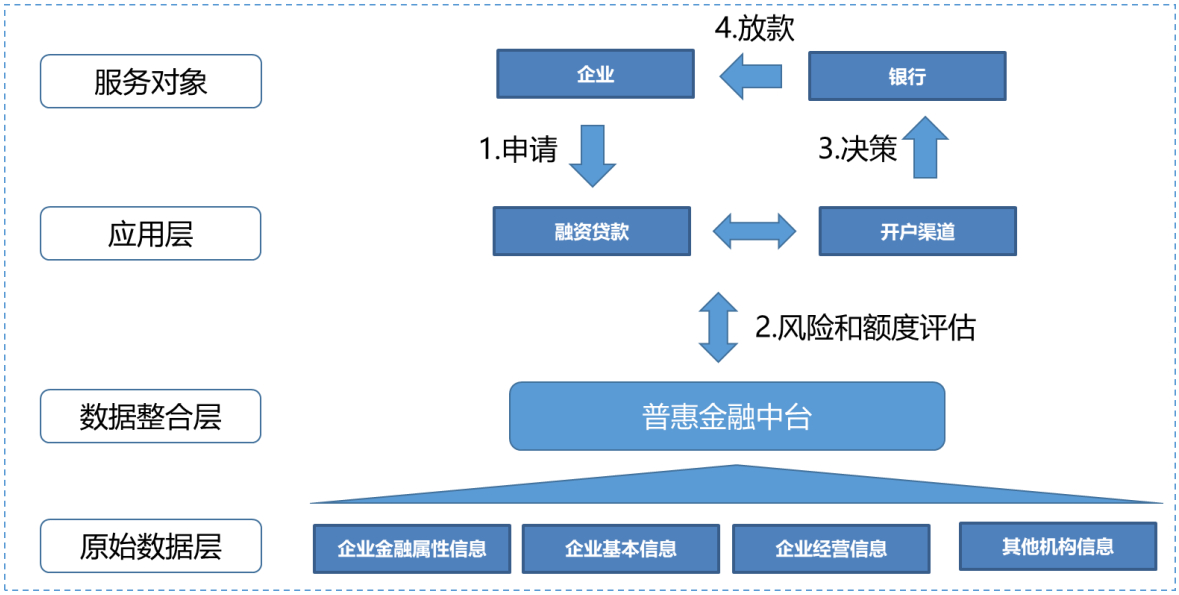

图 2 公共数据在金融业务流程优化场景应用示意图

3. 欺诈监测

随着近年来纯信用贷款产品的兴起，随之而来的信贷欺诈也越来越猖獗，恶意骗贷、仿冒他人骗贷、团伙欺诈等欺诈行为对金融机构造成了严重的损失。目前对于欺诈监测的痛点主要体现在以下两个方面：一方面，从传统欺诈监测工作来讲，而金融机构基于自身掌握的历史还款行为、关联关系以及外部渠道获取的征信数据来做贷前反欺诈，但对于当前骗贷层出不穷的手段来说普遍存在数据维度缺乏、数据量较少等情况，无法准确开展欺诈监测和识别工作；另一方面，在融合多方数据联合构建反欺诈模型的场景中，出于不同机构间对于黑灰名单等敏感信息保护等多

方面原因，数据融合壁垒较高，间接提升了机构之间欺诈相关数据交互的工作难度。综合以上两方面痛点，公共数据在金融领域实现欺诈团伙和欺诈分子精准识别的需求则尤为强烈。一方面，金融机构掌握其交易行为轨迹及欺诈洗钱黑灰名单等信息，政府各委办局依靠各自掌握的行政、司法、欠费等犯罪分子或失信名单及相关联黑灰名单等公共数据，与金融机构之间实时融合可以实现欺诈联防联控；另一方面各机构在数据交互方面尝试使用密码学算法，对电诈、洗钱、骗贷等行为的黑灰名单用户进行匿踪识别，在保障各方隐私的条件下实现特征信息核验和比对，确认其安全可信背景，真正实现企业公共数据与金融机构数据的价值融合，阻止潜在欺诈操作，为客户提供聚合安全的服务，同时树立金融机构联合政府机构为民服务的良好形象。如图 3 所示。

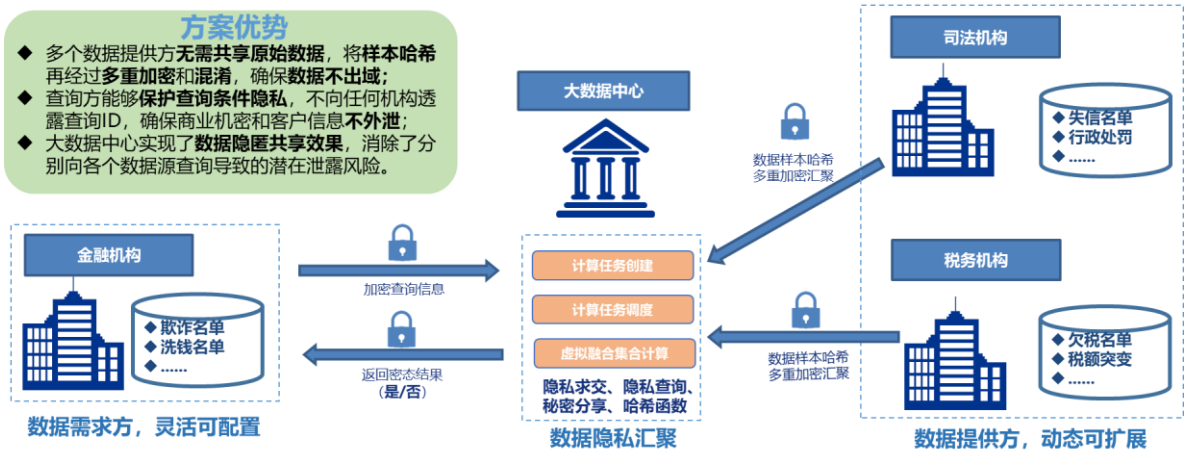

图 3 公共数据在欺诈监测场景应用示意图

4. 智能风控

风险控制管理是金融行业主要任务之一，是使金融业正常运行的保障。无论在传统的银行、证券、保险业，还是新兴的互联网金融业，风险控制都是极为重要的。目前政府和金融机构对于风控的痛点主要体现在以下两个方面：一方面，金融机构传统风控一般构建风险评分卡模型评价客户的风险等级，但对于一些下沉客户，普遍存在数据维度缺乏、数据量较少等情况，而随着实际金融产品及渠道的不断丰富，客户在金融机构内留存的金融业务数据难以满足新的风控需求；另一方面，在联合外部数据优化当前风控模型的场景中，出于不同机构间数据分散及保护等多方面原因，金融机构之间以及金融机构与其他行业机构之间的数据融合壁垒较高，“数据孤岛”现象非常严重，间接提升了机构之间数据交互的工作难度。综合以上两方面痛点，公共数据在金融信贷风控应用探索工作则意义重大，一方面公共数据包括社会经济的方方面面而拥有更多维的数据信息，可以帮助金融机构实现跨机构间数据特征价值的联合挖掘，解决单个金融机构样本量及维度有限的问题；另一方面各机构在数据交互方面也在不断探索尝试科技创新的手段保障原始数据不出域，数据可用不可见，从而在一定程度上解决“数据孤岛”的问题。在实际场景应用中，金融机构自身拥有企业基本信息、关联信息、金融资产信息、交易信息、征信信息等，通过引入工商、环保、司法、海关、公安、税务及舆情等公共数据，积极应用于贷前、贷中、贷后的风险评

估系统模型中进行训练，得到融入公共数据后的综合风险评估结果，实现一键体检进行贷前还款能力和还款意愿的风险评估，同时跟踪企业的贷中贷后经营与资产的变动情况，实时监测感知舆情，提升贷后风险预警水平。如图 4 所示。

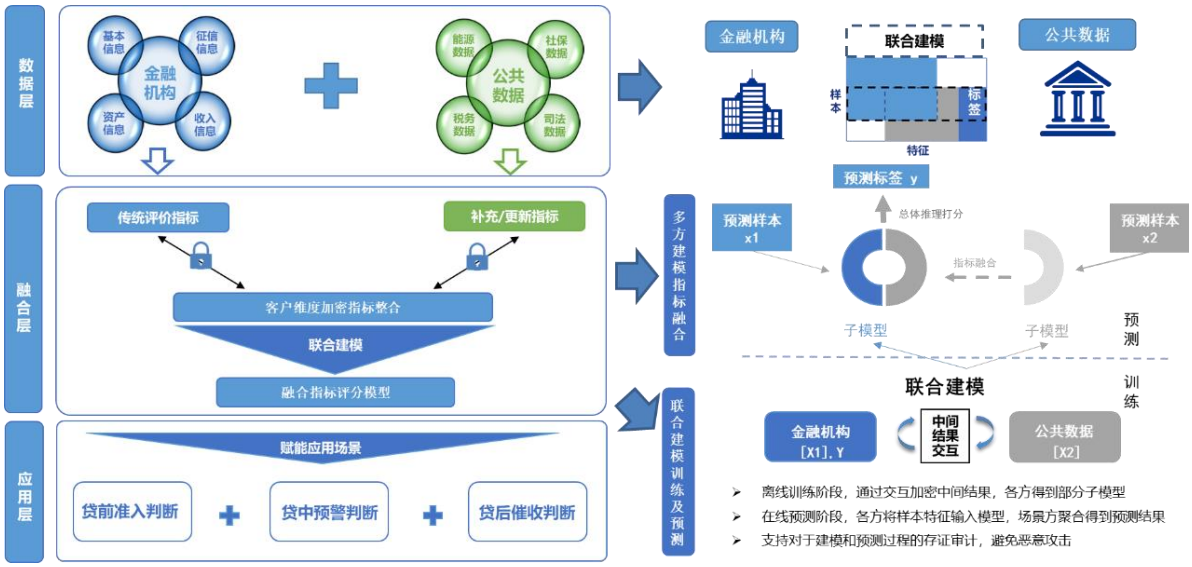

图 4 公共数据在智能风控场景应用示意图

（六）业务应用事例

1. 人社局农民工工资代发监管与保证金监管合作场景

为贯彻落实国务院《保障农民工工资支付条例》，进一步加强工程建设项目农民工劳动用工管理服务，切实提高农民工工资支付监管水平，维护农民工劳动报酬权益，部分地方人社局通过建设监管平台开展农民工工资是否按时发放的监管工作。银行可以通过内部信息系统对接地方人社局监管平台的方式来实现农民工工资代发，协助人社厅监管农民工工资代发专户，保障农民

工工资按时发放，解决农名工工资拖欠、讨薪顽疾。此类公共数据合作场景可促进银行与人社局深度合作，践行金融服务社会责任的同时，带来高质量的对公客户，以及农民工工资代发业务机会，形成大额存款沉淀。

2. 不动产登记中心零售押品信息数据治理合作场景

银行对零售押品数据的质量要求较高，而各地不动产中心的数据资源较银行更为及时、准确、全面，银行可通过与不动产登记中心线上对接，使用数字化手段打造零售押品智能核对管理系统，从而实现对数据进行有效分析和治理，提升数据的完整性和准确性，提高押品核对管理效率。在此生态合作场景中，银行通过与不动产登记机关的信息匹配，对现有押品信息进行核对补充，并将核对无误的完整押品信息回传行内零售信贷相关信息系统，从而提升行内零售押品数据质量的完整性和准确性，有利于后续的定期贷后预警、二次抵押查封涉诉预警，助力数据治理工作。

3. 公安局银行开户异常侦测联合建模合作场景

近年来，由于欺诈、买卖银行卡行为的日益猖獗以及国家“断卡行动”的指导方针要求，银行需担当起守护人民群众财产安全的责任，做好风险识别和拦截。但在面对新客开户时，由于对客户信息的掌握有限，银行无法准确判断欺诈客户。针对此类问题，银行可以借助公安局丰富精准的反欺诈数据资源，开展双方基于隐私计算平台的合作，在保证客户隐私且互不泄露客户明文信息的前提下，利用历史涉嫌欺诈、买卖银行卡等风险的有效

实名黑样本，建立集成模型以及规则组，开发开户异常侦测模型，给出每个账号是否属于开户异常的评分，用于开户场景的异常监测。此类公共数据合作场景落地，可以提升涉案账户管控的准确性和及时性，并从源头上降低银行欺诈风险。

4. 大数据中心个人证照、健康码亮证便民服务合作场景

银行业数字化转型是数字中国建设中的重要一环，金融科技与业务发展的加速融合，形成了数字生态银行的趋势。银行为提升客户体验，可与各省市大数据中心合作，通过大数据中心提供的个人居民身份证查询等数据接口接入证件，电话等信息数据，支持应用于柜面等网点线下渠道的排队取号，综合查询，小额存取转等个人基础金融业务服务场景，使市民无需出示实体证件，即可办理部分银行业务。此类公共数据合作场景，可以优化金融服务流程，提升金融服务效率，拓展金融服务场景，为市民提供便捷的金融服务体验，契合“一网通办”的社台气会化应用步伐。

五、发展与建议

（一）法律政策趋势

公共数据的法律法规政策趋势主要体现在以下几个方面:

1. 加强立法保障

政府重视数据立法工作，制定一系列法律法规，明确公共数据的定义、范围、权利和义务，为公共数据的开放、共享和利用提供法律依据。今年以来，数据要素市场获得地方性政策的密集支持。北京、上海、广东、江西等地陆续出台政策文件，进一步深化数据要素市场改革和创新。

2. 完善数据治理体系

政府通过制定政策和法规，规范数据收集、存储、处理和使用，保障数据安全和个人隐私，构建完善的数据治理体系。

3. 推动数据交易市场规范发展

政府鼓励创新数据交易模式，培育数据交易平台，同时加强对数据交易市场的监管，确保数据交易规范、公平、透明。政府鼓励创新数据交易模式，培育数据交易平台，以促进数据要素在各个行业的广泛应用和价值发挥。

4. 支持数据产业发展

政府制定一系列政策，支持数据产业创新发展，鼓励企业利用公共数据开发新产品和服务，促进数字经济的繁荣。各地政府

在推进数据要素市场化建设中，将公共数据开放共享利用视为率先发力方向，以促进数据要素的价值转化和数字经济发展。

5. 强化国际合作与交流

政府重视国际合作与交流，积极参与国际数据治理和交易规则制定，推动数据跨境流动和合作共赢。

总之，公共数据的法律法规政策趋势是朝着更加规范、透明和有利于创新发展的方向发展，以推动公共数据在各个领域的广泛应用和价值发挥，同时保障数据安全和个人隐私。

（二）技术发展趋势

公共数据的技术发展趋势可以概括为以下几个方面：

1. 云计算

云计算技术在公共数据领域得到了广泛应用。通过将公共数据迁移到云计算平台上，云计算技术可以提供更加高效、安全、灵活的存储和技术能力，可以实现弹性、灵活和可扩展的数据处理和管理，降低数据处理和数据分析的成本，提高数据的公共性和可用性。

2. 大数据技术

随着公共数据量的增长，大数据技术在公共数据领域变得越来越重要。大数据技术可以提供高效的数据处理能力，可以更快地实现数据的收集、存储、分析和可视化，帮助政府和企业更好地理解和利用公共数据，提高公共数据的公共性和价值。

3. 人工智能技术

人工智能(AI)技术在公共数据领域中的应用也越来越广泛。AI 技术可以用于公共数据的分类、预测、推荐等方面，例如机器学习、深度学习等技术可以对数据进行智能分析发现数据中的潜在规律和价值，帮助政府和企业更好地理解和利用公共数据以及提供更有针对性的决策支持，提高公共数据的价值和应用。

4. 区块链技术

区块链技术在公共数据领域中的应用也越来越广泛，区块链技术为公共数据提供一种去中心化、安全、可靠的数据存储和传输方式，可以有效地保证公共数据的完整性、真实性和可追溯性，同时也可以用于公共数据的共享和授权，促进公共数据的共享和利用。

5. 数据安全和隐私保护技术

随着公共数据的广泛应用，数据安全和隐私保护成为越来越关注的问题，政府和企业需要采用加密、脱敏、访问控制等技术手段，确保公共数据的安全和个人隐私保护。关于公共数据的隐私保护技术的前沿算法愈发受到重视，如联邦学习算法，致力于打破数据孤岛，在各建模或数据参与方不泄露本方客户数据和不违反国家数据保护法律法规的前提下，进行联合的机器学习模型训练，同时模型效果与将所有数据集中在一起建模的效果无异，在保障各方客户隐私和数据安全的前提下，带动跨领域的企业级数据合作，催生联合建模的新业态和模式。

总之，公共数据的技术发展趋势是朝着更加高效、智能、安全和可扩展的方向发展，以满足政府和企业在数据处理、分析和应用方面的不断增长的业务需求。同时，新技术的应用也需要关注数据安全和个人隐私包含问题，确保公共数据的合理合规使用。

（三）发展建议

当前，公共数据价值加速凸显，社会各界大力推动公共数据共享开发和开发利用，将加快释放公共数据的业务价值；政策法规体系加速健全，国家政策文件加快细化落地、地方制度创新持续迭代、不断优化；数据运营模式加速形成，建设运营模式更加多元、运营授权制度更加规范、运营生态图谱更加丰富；价值共创机制加速明晰，场景牵引价值释放成为共识，公共数据和社会数据融合应用，市场在数据产品交易定价中的决定性作用凸显；安全可信环境加速构建，安全管理制度和安全可信平台加速构建，数据安全和隐私保护技术得到广泛应用；合规监管机制加速规范，参与主体权责体系不断健全。

关于公共数据在金融领域的价值，有如下几点建议：

1. 完善内部管理体系

在大数据相关技术已经在金融机构当中得到普遍应用后，金融机构必须通过更为规范的管理避免保证运营的顺畅。如机构管理的所有措施均应当以大数据为平台开展，包括金融机构的管理人员以及其他工作人员在内，均应当严格按照工作流程在大数据平台完成操作。另外，金融市场的管理以及金融机构的运营均应

当通过分析大数据，确定是否存在管理的问题，以及管理改革过程当中采用的措施是否能够真正有效解决问题，达到预期目标。

2. 挖掘数据分析能力

大数据产业中，数据是大数据能够得到利用的基础，也是时代最大的资源。从海量的数据当中，金融市场存在的所有问题均能够出现，参与金融市场活动的所有市场参与者存在的问题以及发展的方向同样能够得到准确的判断。因此在大数据时代当中，如何充分而有效地利用大数据的资源，对市场和市场的参与者均具有重大的意义。对金融市场以及市场的参与者而言，通过专业的人才完成数据的分析以及利用，已经成为大数据金融市场的未来需要。

3. 规范数据定价体系

金融产品的价格变化较快，且与实体商品不同，金融产品的价格变化区间较大，因此对金融市场而言，如何根据市场的变化情况确定当前产品的价格，是金融市场需要研究的重要课题之一。大数据技术的应用，对确定金融产品的价格能够起到显著的促进作用，在大数据技术应用情况下，金融市场的变化不仅能够得到更为全面且及时的监测，对金融市场能够形成影响的政策以及其他因素同样能够更为快速纳入价格模型体系当中，使得金融产品的价格能够根据市场确定并及时变更。

4. 提升风险管理水平

在金融模式下大部分金融相关的信息以及数据均存在透明度较低的问题，透明度低使得金融交易双方经常处于不平等的状态，而大数据技术得到广泛应用后，金融相关的数据透明度将出现大幅的提高，金融交易双方，交易前以及交易后的信息获取不会存在较大的差异，该种情况下由于信息掌握不均匀造成的很多风险能够得到有效的规避。因此在大数据技术应用情况下，金融市场应该利用其完成风险的整体性规避，提高金融市场的安全性。