CAICT 中国信通院

隐私保护计算与合规应用研究报告

（2021 年）

中国信息通信研究院安全研究所

2021 年 4 月

版权声明

参与单位

牵头单位：

中国信息通信研究院安全研究所

参编单位（排名不分先后）：

阿里巴巴集团安全部

北京数牍科技有限公司

腾讯研究院

智联出行研究院

中国工商银行

前　言

“纲举目张，执本末从”。站在“两个一百年”奋斗目标的历史交汇点上，中国共产党第十九届中央委员会第五次全体会议谋划长远，为中国擘画了一幅波澜壮阔的新图景[1]，提出**“发展数字经济，推进数字产业化和产业数字化，推动数字经济和实体经济深度融合，打造具有国际竞争力的数字产业集群”**的明确目标[2]。作为数字经济发展的核心内容，数据要素的**流通共享**与**协同应用**对于实现数据要素价值充分释放，解决数据市场的竞争与垄断问题，推动制造业高质量发展，提振实体经济具有重要意义。

“识变之智，应变之方”。综合当前实践来看，囿于数据的法律属性和产权规则在理论和立法层面长期未能清晰界定，规范有效的数据交易流通市场始终未能真正形成[3]，使得基于数据流转的货币化数据交易模式所产生的直接经济价值仍存在巨大释放空间。而依托于隐私保护计算技术的平台化数据协同应用服务模式所体现的间接经济价值更为明显。隐私保护计算将**数据持有**与**使用**分离，在保障数据持有者对数据控制的前提下，将数据加以利用，实现了数据“可用不可见”，进一步激发了数据要素的价值赋能。在近期发布的《金融业数据能力建设指引》（JR/T0218-2021）中，也将“数据可用不可见”作为金融业数据能力建设遵循的五大基本原则之一。

“备豫不虞，为国常道”。科学研判“时”与“势”，辩证

---

[1] 高远务实的时代擘画——党的十九届五中全会侧记 新华社

[2] 《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五远景目标的建议》

[3] 关于数据要素交易流通模式的新思考．张钦坤、朱开鑫．腾讯研究院

把握“危”与“机”，方能决胜于千里之外。鉴于数据在全球竞争中的经济价值及战略价值日益凸显，不当的数据利用所引发的公众个人信息和隐私安全恐慌也备受关注。2020 年，我国相继出台了《中华人民共和国民法典》（以下简称《民法典》），发布了《中华人民共和国数据安全法（草案）》（以下简称《数据安全法（草案）》）《中华人民共和国个人信息保护法（草案）》（以下简称《个人信息保护法（草案）》）等法律法规，丰富了我国在个人信息和隐私保护领域的立法内容，初步构建了我国个人信息和隐私保护的法律监管框架。

“天下将兴，其积必有源”。日趋严格的合规监管、日渐强化的政策引导、日益旺盛的市场需求以及日臻成熟的技术进步，共同助推了隐私保护计算的发展，为其提供了广阔的应用前景。但是出于对个人信息和隐私保护政策法规理解不足、隐私保护计算前沿技术了解不详、产业落地缺乏参考等诸多原因，使得行业对数据隐私保护计算技术应用和数据合规之间产生了一定的理解“鸿沟”，行业整体对于数据的流通共享及协同应用仍持有相对保守且谨慎的观望态度。据此，为数据要素流通市场之兴，需积累隐私保护计算技术合规应用之源。

“凡益之道，与时偕行”。本报告通过总结隐私保护计算关键技术、详细梳理我国个人信息和隐私保护法律法规框架，分析隐私保护计算技术相关的个人信息和隐私保护的关键合规点，深入讨论了隐私保护计算技术如何帮助企业更好地满足个人信息和隐私保护的要求。最后，围绕建立健全法律法规体系、加快标准体系建设、强化风

险防控、明确安全与发展并举以及人才培养进行了展望，为关注隐私保护计算技术及个人信息和隐私保护的社会各界提供有益参考。针对报告中的诸多不足，恳请各界同仁批评指正。

本报告的编制过程，得到了来自阿里巴巴集团安全部、北京数牍科技有限公司、腾讯研究院、智联出行研究院及中国工商银行的大力支持。特此，向支持、参与本报告编制工作的各位领导及专家表示感谢。

目 录

图 目 录

表 目 录

一、隐私保护计算技术概述

（一）隐私保护计算技术定义

随着移动互联网、5G、大数据、云计算等新一代信息技术的迅猛发展，数据应用与隐私保护的矛盾日益突出，隐私保护计算被认为是解决这对矛盾的有效技术手段而备受关注。2016 年李凤华等学者提出，**隐私计算是面向隐私信息全生命周期保护的计算理论和方法，是隐私信息的所有权、管理权和使用权分离时隐私度量、隐私泄漏代价、隐私保护与隐私分析复杂性的可计算模型与公理化系统**[4]。2019 年《UN Handbook on Privacy-Preserving Computation Techniques》中提到**隐私保护计算是在提供隐私保护的前提下，实现数据价值挖掘的技术体系**[5]。隐私保护计算并不是一种单一的技术，它是一套包含人工智能、密码学、数据科学等众多领域交叉融合的跨学科技术体系[6]，实现数据“可用不可见”。通过数据价值的流通，促进企业数据的合法合规应用，激发数据要素价值释放，进一步培育数据要素市场。

（二）隐私保护计算关键技术

隐私保护计算作为涉及多领域交叉融合的跨学科技术体系，重点提供了数据计算过程和数据计算结果的隐私安全保护能力，总体来说包含**联邦学习、安全多方计算、机密计算、差分隐私和同态加密**等。

[4] 李凤华，李晖，贾焰等. 隐私计算研究范畴及发展趋势[J]. 通信学报, 2016, 37(4): 1-11.

[5] UN Handbook on Privacy-Preserving Computation Techniques.

[6] 中国信息通信研究院、阿里巴巴集团安全部、北京数牍科技有限公司：《隐私保护计算技术研究报告 2020》

1. 联邦学习（Federated Learning, FL）

联邦学习涉及“联邦”和“学习”两部分，是密码学和人工智能相结合的分布式学习技术。根据数据集的分布特点，联邦学习分为横向联邦学习、纵向联邦学习和联邦迁移学习三类[7]。谷歌提出的（横向）联邦学习的技术原理是：客户终端从中央服务器下载现有预测模型，通过使用本地数据对模型进行训练，并将模型的更新内容加密上传至云端，再由中央服务器聚合产生新的模型。反复上述过程，原始数据始终存储在本地。联邦学习无需交换和传输原始数据，各数据提供方仅需传输部分中间结果，如梯度、模型参数等，得到的模型和中心化训练的模型相比，性能几乎无损。

2. 安全多方计算（Secure Multi-Party Computation, SMPC）

安全多方计算是密码学领域的重要分支之一，最早由图灵奖获得者、中国科学院院士姚期智教授于 1982 年通过“百万富翁问题”提出。安全多方计算研究两个或多个持有私有输入的参与者，在不泄漏各自私有输入数据的情况下联合计算一个函数，各自得到他们预定的输出。当前商用领域的大多数安全多方计算方案都是基于具有通信量优势的秘密共享技术，其基本原理是：将需要计算的函数转换成由加法和乘法“门”组成的算术电路，利用秘密共享实现加法和乘法两个基础运算，理论上就可以对整个算术电路进行计算。安全多方计算具

[7] 横向联邦学习针对的是不同数据集之间，特征重合较多而样本重合较少的情形；纵向联邦学习针对的是不同数据集之间，样本重合较多而特征重合较少的情形；联邦迁移学习针对的是不同数据集之间，样本和特征重合均较少的情形。

有严格的安全定义，包括**隐私性、正确性、公平性（可选）、结果传递保证（可选）等方面**。整个计算过程在保证计算结果正确的同时，不泄露任何参与方的原始秘密数据和计算过程中的明文中间信息，各数据提供方对其拥有的数据始终拥有控制能力。

3. 机密计算（Confidential Computing，CC）

机密计算是指在基于硬件的可信执行环境（Trusted Execution Environment, TEE）中执行代码来保护数据应用中的隐私安全的技术之一，其中 TEE 定义为在数据机密性、数据完整性和代码完整性三方面提供一定保护水平的环境[8]。机密计算的基本原理是：将需要保护的数据和代码存储在 TEE 中，对这些数据和代码的任何访问都必须通过基于硬件的访问控制，防止它们在使用中未经授权被访问或修改，从而提高机构管理敏感数据的安全水平。基于 TEE 的隐私保护方案不受算法和网络限制，其安全性完全依赖于 TEE 自身的安全性，实际使用中需要信任 TEE 硬件厂商或平台服务商。

4. 差分隐私（Differential Privacy, DP）

差分隐私是 Dwork 等人在 2006 年针对数据库隐私问题提出的一种严格的、可量化的隐私定义和技术[9]，被麻省理工科技评论为 2020 全球十大突破性技术之一[10]。以中心化差分隐私为例，其基本原理是：在计算结果中添加噪声（如适用于数值型输出的拉普拉斯噪声和适用于非数值型输出的指数噪声），使得修改数据集中单条记录不会对统

[8] Confidential Computing Deep Dive v1.0

[9] Dwork C . Calibrating noise to sensitivity in private data analysis[J]. Lecture Notes in Computer Science, 2012, 3876(8): 265-284.

[10] https://www.technologyreview.com/10-breakthrough-technologies/2020/

计结果造成显著的影响，从而保证攻击者在拥有背景知识的情况下也无法推断出该记录对应的敏感信息。差分隐私具有两个重要的优点：一是提出背景知识无关的隐私保护模型，实现攻击者背景知识最大化假设；二是为隐私保护水平提供严格定义和量化评估方法。

5. 同态加密（Homomorphic Encryption, HE）

同态加密是一种特殊的加密算法，它允许在加密之后的密文上直接进行计算，且解密后的计算结果与基于明文的计算结果一致。根据支持密文运算的程度，同态加密方案可以分为部分同态加密方案和全同态加密方案两类。部分同态加密方案只能支持有限的密文计算深度，如 Paillier[11]支持密文间的加法运算，但是不支持密文间的乘法运算；BGN[12]能够支持无限次密文间的加法运算，但是只能支持一次密文间的乘法运算。目前，部分同态加密在一些运算并不复杂的场景中得到了应用。全同态加密方案对密文上的计算深度没有限制，理论上可以支持任意的密文计算。FHE 的理论应用研究已经有很多，但是 FHE 的计算代价比较高，目前还未了解到在商业场景中有大规模应用。

（三）隐私保护计算技术解决的主要问题

诸多现实场景中，只有足够的数据量、丰富的特征维度才能得到真正有意义的结果，往往需要多个实体机构共同提供数据。传统的、集中式的数据使用方式存在数据安全、法律合规等诸多风险。隐私保

[11] Paillier P. Public-key cryptosystems based on composite degree residuosity classes[J]. 1999.

[12] Boneh D. Evaluating 2-DNF formulas on ciphertexts[J]. TCC'05, 2005.

护计算从技术上，实现原始数据不出库、数据“价值”和“知识”流通的目标，促进跨领域多维度数据的融合，构建“数据可用不可见”的合作新模式。目前的商用场景中，隐私保护计算技术主要用于解决模型训练（建模）、预测、匹配、联合统计分析等场景下的隐私安全问题。

关于建模。模型训练是隐私保护计算技术的重要应用场景。为了训练出效果更好的模型，需要联合多个数据实体来扩充样本数量或丰富特征维度。整个训练过程中需要保护原始数据、梯度、模型参数等信息，在实现保护企业自身利益的同时满足隐私合规要求。隐私保护计算技术，将敏感信息通过秘密碎片、加密等形式进行传递，保证参与方在整个计算过程中难以得到除计算结果之外的额外信息，也难以逆推原始输入数据和其他隐私信息。

关于预测。模型训练的最终目的是为了预测，在安全预测的场景中，各参与方只拥有模型或数据的一部分，所以预测过程既要保证各方数据安全也要保证各方模型安全，有时甚至需要隐藏预测的 ID 信息。比如借贷机构和征信机构联合预测某借款人的信用时，借贷机构并不想让征信机构知道该借款人有借款需求，所以借贷机构希望在预测借款人信用的同时不泄露借款人的 ID。

关于匹配。匹配常用于黑名单查询、多头借贷、广告推荐等场景。比如对于黑名单查询，作为借贷机构的一方想要从黑名单拥有方查询自身用户是否在黑名单中，以决定是否向该用户放款。在这个过程中，借贷机构不希望把用户信息告诉黑名单拥有方，同时数据拥有方也不

希望把整个黑名单告诉借贷机构。使用隐私集合求交集，可以让双方只获得交集结果，更加符合双方的安全需求。此外，还可以通过隐私信息检索技术来获得交集中每个元素对应的附加消息。

关于联合统计分析。隐私保护计算技术另一个重要的应用场景是联合统计分析。如典型的需求是安全 SQL 查询：涉及更灵活多样的 SQL 运算，既要保护数据的隐私安全，也要保护 SQL 语句的隐私安全。通用安全多方计算技术及机密计算技术都可以支持隐私保护的 SQL 数据库查询，能够支持自定义 SQL 运算，同时最大程度保护数据库和 SQL 语句的安全。

二、个人信息和隐私保护立法与监管

（一）个人信息和隐私保护立法

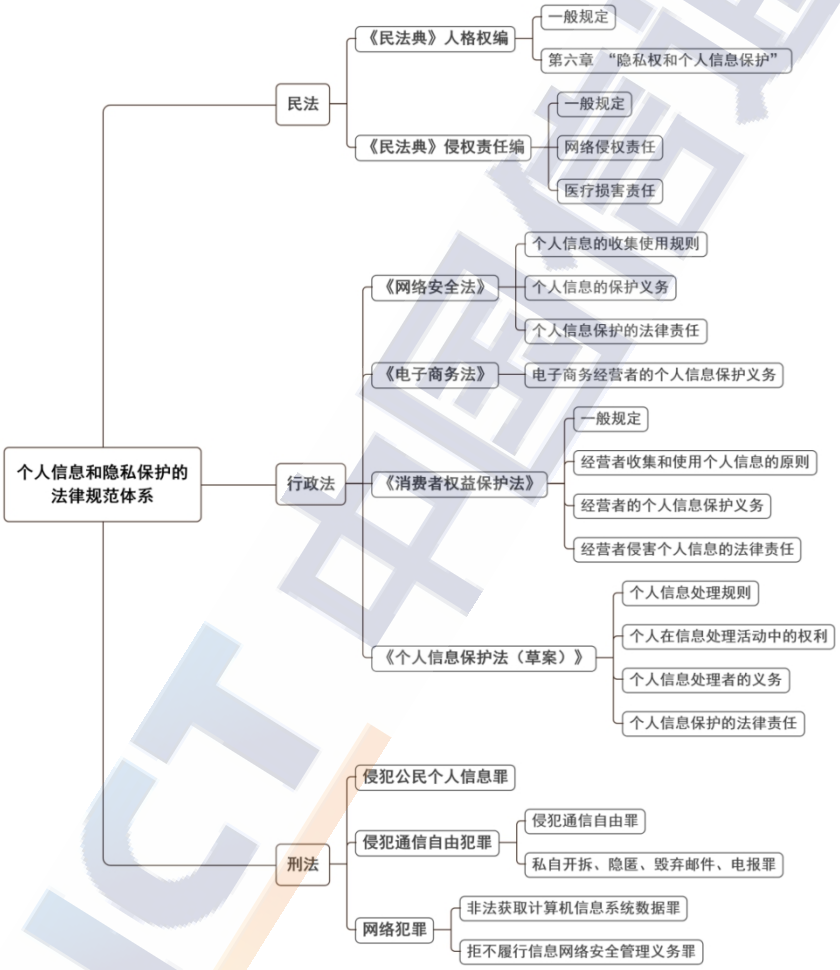
1. 个人信息和隐私保护立法概览

2020 年，随着《民法典》《数据安全法（草案）》《个人信息保护法（草案）》等涉及数据安全、个人信息和隐私保护的关键立法相继制定、出台或公布，个人信息和隐私保护的法制化程度进展明显。从法律的角度看，个人信息和隐私保护的规则，主要是通过三个层次的法律维度构建的（如图 1 所示）：

一是以《民法典》为代表的个人信息和隐私保护的民事法律路径；

二是以《中华人民共和国网络安全法》（以下简称《网络安全法》）《数据安全法（草案）》《个人信息保护法（草案）》等为代表的行政监管路径；

三是以《中华人民共和国刑法》（以下简称《刑法》）为代表的补充性规制路径。



资料来源：根据公开资料整理

图 1 个人信息和隐私保护立法框架

2. 个人信息和隐私保护立法具体内容

(1) 民法

2021 年 1 月 1 日起施行的《民法典》，在人格权编中专章规定了隐私权和个人信息保护问题。具体体现在以下方面：

一是用概括加列举的方式明确了“个人信息”的定义和范围。《民法典》规定“个人信息是以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人的各种信息，包括自然人的姓名、出生日期、身份证件号码、生物识别信息、住址、电话号码、电子邮箱、健康信息、行踪信息等。”

二是确立了个人信息处理的基本原则。个人信息的处理包括个人信息的收集、存储、使用、加工、传输、提供、公开等。《民法典》第一千零三十五条以及一千零三十六条分别规定了处理个人信息的原则以及责任豁免情形。

三是确立了自然人对其个人信息的查阅、复制、删除等权利。《民法典》第一千零三十七条规定，自然人可以依法向信息处理者查阅或者复制其个人信息；自然人发现其个人信息有误的有权要求更正，发现信息处理者违法违约处理其个人信息的，有权要求删除。

四是规定了信息处理者的个人信息保护义务。《民法典》规定，信息处理者不得泄露或者篡改其收集、存储的个人信息；未经自然人同意，不得向他人非法提供其个人信息，但是经过加工无法识别特定个人且不能复原的除外。同时，《民法典》也要求信息处理者采取技

术措施和其他必要措施，确保其收集、存储的个人信息安全，防止信息泄露、篡改、丢失；发生或者可能发生个人信息泄露、篡改、丢失的，应当及时采取补救措施，按照规定告知自然人并向有关主管部门报告。

此外，《民法典》侵权责任编中的部分条款也可适用于个人信息保护，主要包括侵权责任的一般条款、网络侵权责任以及医疗损害责任等方面的条款。2021 年 1 月 1 日修改的《最高人民法院关于审理利用信息网络侵害人身权益民事纠纷案件适用法律若干问题的规定》也强调了对网络用户或者网络服务提供者侵害他人人身权益，造成财产损失或者严重精神损害的行为，被侵权人有权依据《民法典》相关规定，请求其承担赔偿责任。

(2) 行政法

2016 年通过的《网络安全法》第四章“网络信息安全”，也被称为“个人信息保护专章”。《网络安全法》确立了个人信息收集使用的基本原则，即**合法正当原则**、**知情同意原则**、**目的限制原则**、**安全保密原则**以及**删除改正原则等**。同时，《网络安全法》规定了相关主体的**个人信息保护义务**。此外，《网络安全法》第六十四条、第七十四条还规定了违反个人信息保护的法律责任。

2018 年发布的**《中华人民共和国电子商务法》**（以下简称《**电子商务法**》）中规定了电子商务经营者有保护用户个人信息的义务，并在用户对其个人信息的主体权利方面做出规定，要求电子商务经营

者明示用户信息查询、更正、删除以及用户注销的方式、程序，不得对用户信息查询、更正、删除以及用户注销设置不合理条件。同时要求电子商务经营者在核实身份后及时响应用户查询或者更正、删除的请求，并在用户注销后立即删除该用户的信息（依照法律、行政法规的规定或者双方约定保存的，依照其规定）。

《中华人民共和国消费者权益保护法》（以下简称《消费者权益保护法》）中规定了消费者“享有个人信息依法得到保护的权利”，此外，还规定了经营者收集、使用消费者个人信息的原则以及对个人信息的保护义务。在法律责任方面，经营者侵害消费者个人信息依法得到保护的权利，应当承担停止侵害、恢复名誉、消除影响、赔礼道歉并赔偿损失的民事责任，以及相应行政责任。

《数据安全法（草案）》于 2020 年 7 月 3 日发布并向社会公开征求意见。《数据安全法（草案）》没有专门区分数据与个人数据，其在第三条规定：本法所称数据，是指任何以电子或者非电子形式对信息的记录。同时，草案中规定了数据安全的十大制度，包括**数据分级分类、数据安全审查、数据出口管制、数据安全风险评估监测预警、全流程数据安全管理制度、数据收集制度、数据交易制度**等，都涉及对数据安全的规定。

《个人信息保护法（草案）》于 2020 年 10 月 21 日发布，其既需和现有法律规范相协调，又需为现有法律保护框架带来完善和创新。具体内容包括：**一是明确了“个人信息”定义**，采用“识别+关联”的标准将个人信息定义为“与已识别或可识别的自然人有关”；

二是明确了个人信息处理规则，细化了前述提及的《民法典》中关于“告知同意”的规定，并以“告知同意”作为个人信息处理中的合法基础之一；**三是确立了个人信息保护权利制度。**明确了个人在信息处理活动中的权利，包括但不限于知情权、决定权、查阅复制权、补充更正权、删除权以及限制或拒绝处理等。**四是明确了个人信息处理者的保护义务。**包括但不限于：指定个人信息保护负责人；对个人信息实施分级分类管理；采取相应的加密、去标识化等安全技术措施；合理确定个人信息处理的操作权限，并定期对从业人员进行安全教育和培训；设立事前风险评估制度、建立事后补救措施与通知义务等。**五是明确了个人信息保护的法律责任。**在民事赔偿方面，对赔偿数额的计算标准、规则原则以及起诉主体进行了规定。在行政处罚方面，对违反《个人信息保护法（草案）》规定处理个人信息的行为，惩罚措施包括并处五千万以下或者上一年度营业额百分之五以下的罚款，可以并处责令暂停相关业务、停业整顿、吊销相关业务许可证或营业执照；对直接主管人员处十万元以上一百万元以下的罚款；记入信用档案，并予以公示等。

(3) 刑法

《刑法》第二百五十三条之一规定了侵犯公民个人信息罪，对违反国家有关规定，向他人出售或者提供公民个人信息，窃取或者以其他方法非法获取公民个人信息，情节严重的，追究刑事责任。该罪的犯罪主体包括自然人和单位，而“违反国家有关规定，将在履行职责

或者提供服务过程中获得的公民个人信息，出售或者提供给他人的”，则会从重处罚。

此外，《刑法》中还有一些罪名虽非专门为保护个人信息而设，但也可用于规制某些侵犯公民个人信息的行为。如：侵犯通信自由罪；私自开拆、隐匿、毁弃邮件、电报罪；非法获取计算机信息系统数据罪；拒不履行信息网络安全管理义务罪等。

3. 个人信息和隐私保护立法趋势

一是个人信息定义逐渐明晰。《网络安全法》将个人信息定义为“电子或者其他方式记录的能够单独或者与其他信息结合识别自然人个人身份的各种信息”。《民法典》基本上沿用了“识别”路径。《个人信息保护法（草案）》将个人信息定义为“以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息”，在“识别”基础上，兼顾了“关联”路径。

二是个人信息处理规则更加具体且有针对性。《个人信息保护法（草案）》将“告知同意”作为一项大多数国家在进行个人信息保护立法时都普遍认可的核心处理原则，体现了信息主体对信息的自主决定权。《个人信息保护法（草案）》明确规定：告知需要告知充分、明晰；同意（包含单独同意、自愿明确同意、书面同意、重新取得同意等）需要真实有效，除此之外还有其他规则同样构成了个人信息处理的合法性基础，比如：为订立或者履行个人作为一方当事人的合同所必需；为履行法定职责或者法定义务所必需；为应对突发公共卫生事件，或者紧急情况下为保护自然人的生命健康和财产安全所必需；

为公共利益实施新闻报道、舆论监督等行为在合理的范围内处理个人信息等，更好地回应了规范和实践中不同信息处理场景的个人信息保护需求，也为处理过程中必需的利益权衡提供了更有针对性的依据，更贴切地落实了个人信息保护制度的初衷。

三是个人信息保护的责任划分和追责机制更加完整。民事赔偿和行政处罚制度的细化使得对个人信息保护法律责任的界定更符合比例原则，在一定程度上消解了此前刑法前置带来的问题。责任划分方面，《个人信息保护法（草案）》对政府和企业的责任进行了界定，规定了从事共同处理、跨境处理、向第三方提供等行为时所要承担的责任。追责方面，从不同的部门法角度解读，个人信息处理者可能会具备侵权人、服务提供者、经营者等多重身份，可以基于具体的行为人找到和责任相适应的处罚措施，既要保证处罚的效用，也要防止追责手段的滥用。

（二）个人信息和隐私保护监管与执法

随着数据价值的进一步凸显，侵犯个人信息和隐私的情形不断增多。2020 年，监管机关针对个人信息和隐私监管与执法都呈现了趋严趋紧的特点，不断加大对个人信息和隐私保护的力度。具体而言，表现为以下几方面：

一是多主体的监管结构逐步优化。个人信息和隐私保护监管由国家网信部门负责统筹协调，公安机关、工信部门、市场监督管理部门以及其他行业主管部门在各职责范围内行使监督管理职责。在《网络

安全法》第八条、《数据安全法（草案）》第七条及《个人信息保护法（草案）》第五十六条均有体现。需要注意的是，在《数据安全法（草案）》第六条增加了中央国家安全领导机构负责数据安全工作的决策和统筹协调的职责，同时将数据区分为网络数据和非网络数据，其中对于网络数据而言，明确：国家网信部门依照本法和有关法律、行政法规的规定，负责统筹协调网络数据安全和相关监管工作（表 1）。

表 1 监管结构梳理

《网络安全法》	**第八条** 国家网信部门负责统筹协调网络安全工作和相关监督管理工作。国务院电信主管部门、公安部门和其他有关机关依照本法和有关法律、行政法规的规定，在各自职责范围内负责网络安全保护和监督管理工作。 县级以上地方人民政府有关部门的网络安全保护和监督管理职责，按照国家有关规定确定。
《数据安全法（草案）》	**第六条 中央国家安全领导机构**负责数据安全工作的决策和统筹协调，研究制定、指导实施国家数据安全战略和有关重大方针政策。
	第七条 各地区、各部门对本地区、本部门工作中产生、汇总、加工的数据及数据安全负主体责任。 工业、电信、自然资源、卫生健康、教育、国防科技工业、金融业等行业主管部门承担本行业、本领域数据安全监管职责。 公安机关、国家安全机关等依照本法和有关法律、行政法规的规定，在各自职责范围内承担数据安全监管职责。 国家网信部门依照本法和有关法律、行政法规的规定，负责统筹协调网络数据安全和相关监管工作。
《个人信息保护法（草案）》	**第五十六条** 国家网信部门负责统筹协调个人信息保护工作和相关监督管理工作。国务院有关部门依照本法和有关法律、行政法规的规定，在各自职责范围内负责个人信息保护和监督管理工作。

资料来源：根据公开资料整理

二是监管执法强度持续提高。2020 年，由网信部门综合协调，公安机关、工信部门、市场监督管理部门以及其他行业主管部门，在其各自职责范围内，围绕 APP 监管和个人信息违法犯罪活动开展了

诸多个人信息、隐私保护执法专项和涉及行政处罚的案件。如工信部门聚焦互联网领域的 APP 治理，根据《网络安全法》《中华人民共和国电信条例》《电信和互联网用户个人信息保护规定》等法律法规和《GB/T 35273-2020 信息安全技术 个人信息安全规范》等标准规范，对 APP 主体违法违规收集、处理个人信息的行为进行持续的专项整治；截至 2020 年 12 月，已经对 52 万款 APP 进行了技术检测工作，责令 1571 款违规 APP 进行整改，公开通报了 500 款 APP，下架 120 款整改不到位及拒不整改的 APP[13]。同时，工信部目前也正在起草《移动互联网应用程序个人信息保护管理暂行规定》。2021 年工信部在监管上也将继续开展 APP 治理，进一步强化消费者个人信息保护。

此外，市场监督管理部门聚焦消费者权益保护领域，依据《消费者权益保护法》等法律法规，对未经消费者同意收集其个人信息以及发送商业性信息的行为进行处罚。2020 年，公安机关深入推进“净网 2020”专项行动，对侵犯公民个人信息犯罪，开展集中打击行动，侦办侵犯公民个人信息类案件 6524 起，抓获犯罪嫌疑人 1.3 万名[14]。2021 年，全国公安机关将继续开展“净网”专项行动，持续打击侵犯公民个人信息的行为。

三是监管措施日趋多样化。从检索案例以及实践经验来看，当前的监管措施形式表现更加多样化。网信部门以联合多部门进行专项执法并约谈相关违法违规主体为主；工信部门依职权对相关 APP 进行检测调查，并采取通报、约谈、责令整改、下架 APP 等为主；市场

[13] 国务院新闻办公室 2020 年 12 月 24 日 新闻发布会发布要点：APP 用户个人信息保护工作取得明显成效

[14] 公安部新闻发布会 2021 年 3 月 8 日［人民公安报］2020 年全国公安机关 打击整治网络犯罪成效明显

监管部门和公安机关往往直接予以警告、罚款等行政处罚，对于情节严重的给予了吊销营业执照和行政拘留的处罚。此外，2020 年 10 月，工信部发布了《信息通信行业信用记分实施方案（试行）（征求意见稿）》，拟采用信用记分的方式对企业进行监管。2020 年 12 月 18 日，国务院办公厅发布了《关于进一步完善失信约束制度构建诚信建设长效机制的指导意见》，同样关注到了信用监管这一新监管措施。**对于这种基于信用的新监管方式对个人信息和隐私保护带来的监管变化，行业需要予以关注。**

（三）金融领域个人信息和隐私保护立法与监管

为发挥更大的金融数据价值，金融行业作为数字化水平最高的行业之一，围绕数据协同应用积极探索并开展了一系列的数字化经营和实践工作，如交易欺诈识别、反洗钱、多方黑名单查询等。但是，在金融数据协作过程中，也出现了多种针对个人信息和隐私信息的违法行为。例如未经审批查询个人金融信息、违规保存客户身份资料和交易记录、侵害消费者个人信息等。围绕金融数据安全，国家和监管部门陆续出台一系列法律法规，个人金融信息保护被提高到前所未有的程度。

1. 个人金融信息

当前针对个人金融信息的立法及规范制定的情况主要如表 2 所示：

表 2 金融领域法律法规梳理

法律	《中华人民共和国反洗钱法》	
	《中华人民共和国商业银行法》	
行政法规	《个人存款账户实名制规定》	
	《征信业管理条例》	
部门规章	中国人民银行	《中国人民银行金融消费者权益保护实施办法》
	银保监会	《商业银行信用卡业务监督管理办法》
		《电子银行业务管理办法》
	证监会	《货币市场基金监督管理办法》
部门规范性文件	中国人民银行	《全国银行间债券市场柜台业务管理办法》
		《中国人民银行关于银行业金融机构做好个人金融信息保护工作的通知》
		《中国人民银行关于金融机构进一步做好客户个人金融信息保护工作的通知》
		《移动金融客户端应用软件安全管理规范》
		《网上银行系统信息安全通用规范(2020)》
		《中国人民银行关于进一步加强银行卡风险管理的通知》
	银保监会	《银行业金融机构外包风险管理指引》
		《银行业金融机构数据治理指引》
		《中国银保监会关于银行保险机构加强消费者权益保护工作体制机制建设的指导意见》
	银监会、财政部等五部委联合印发	《金融资产管理公司监管办法》
	金标委	《个人金融信息保护技术规范》

资料来源：根据公开资料整理

个人金融信息的定义。现阶段对个人金融信息有明确定义的是规章层级的《中国人民银行金融消费者权益保护实施办法》（以下简称《办法》）和属于规范性文件的《个人金融信息保护技术规范》（以下简称《规范》），《办法》所称的消费者金融信息是“银行、支付机构通过开展业务或者其他合法渠道处理的消费者信息”，《规范》所称的个人金融信息是“金融业机构通过提供金融产品和服务或者其他渠道获取、加工和保存的个人信息”，虽然规制对象不完全等同，但覆盖范围较为一致，主要包括个人身份信息、财产信息、账户信息、

信用信息、金融交易信息及其他与特定消费者购买、使用金融产品或服务相关的信息。

目前对个人金融信息立法相对比较分散，主要是结合征信、支付、存款等具体业务场景和相关部门的职能范围进行专门规定，但从立法以及各推荐性标准的制定趋势上看，规范的制定方向也在逐渐回应个人金融信息处理场景有所增设和处理范围不断扩大的现实情况，个人金融信息处理者所包含的对象也从传统的银行扩大到了金融机构，对行为的规定也从一般的原则性规定逐渐具体到对金融数据的收集、存储、提供、披露等行为，按照事前预防和事后规制的思路结合金融领域的特殊情况进行规定，比如《个人金融信息保护技术规范》就根据信息遭到未经授权的查看或未经授权的变更后所产生的影响和危害，将个人金融信息按敏感程度从高到低分为了 C3、C2、C1 三类。

据不完全统计，在金融领域，中国人民银行总行及各地分支行针对包括未经审批查询个人金融信息、未按照规定保存客户身份资料和交易记录、侵害消费者个人信息依法得到保护的权利等在内的“个人金融信息”违法行为，开出了近 200 张罚单，涉及罚单金额近 2 亿元[15]。

[15] 《中国个人金融信息保护执法白皮书 2020》

三、合规要点与隐私保护计算技术

（一）重要合规要点

个人信息保护的具体合规要求包含数据最小化、数据分级分类、数据匿名化、知情同意、规范操作、应急补偿等多项要求。其中有三个与隐私保护计算相关且相对重要的合规要点分别是数据最小化、数据分级分类和数据匿名化。一般情况下，如果能做到其中的一点或多点，即可显著减少数据安全风险。其中数据最小化强调个人数据收集要尽量克制，与收集和使用目的保持一致；数据分级分类强调对不同数据的管理与保护有所区别，敏感、重要的数据加强保护；数据匿名化强调对于不需要识别自然人身份的情况，通过匿名化改变数据的颗粒度，减少被识别后给自然人带来的风险。

表 3 个人信息和隐私保护重要合规要点梳理

合规要求	法律规定 （义务来源）	具体条款
数据最小化	《民法典 人格权编》	第一千零三十五条 第一款 处理个人信息的，应当遵循合法、正当、必要原则，不得过度处理，并符合下列条件：（一）征得该自然人或者其监护人同意，但是法律、行政法规另有规定的除外；（二）公开处理信息的规则；（三）明示处理信息的目的、方式和范围；（四）不违反法律、行政法规的规定和双方的约定。
	《网络安全法》	第四十一条 网络运营者收集、使用个人信息，应当遵循合法、正当、必要的原则，公开收集、使用规则，明示收集、使用信息的目的、方式和范围，并经被收集者同意。 网络运营者不得收集与其提供的服务无关的个人信息，不得违反法律、行政法规的规定和双方的约定收集、使用个人信息，并应当依照法律、行政法规的规定和与用户的约定，处理其保存的个人信息。
	《数据安全法（草案）》	第二十九条 任何组织、个人收集数据，必须采取合法、正当的方式，不得窃取或者以其他非法方式获取数据。 法律、行政法规对收集、使用数据的目的、范围有规定的，应当在法律、行政法规规定的目的和范围内收集、使用数据，不得超过必要的限度。
	《个人信息保护法（草案）》	第六条 处理个人信息应当具有明确、合理的目的，并应当限于实现处理目的最小范围，不得进行与处理目的无关的个人信息处理。
	《消费者权益保护法》	第二十九条 第一款 经营者收集、使用消费者个人信息，应当遵循合法、正当、必要的原则，明示收集、使用信息的目的、方式和范围，并经消费者同意。经营者收集、使用消费者个人信息，应当公开其收集、使用规则，不得违反法律、法规的规定和双方的约定收集、使用信息。
数据分级分类	《数据安全法（草案）》	第十九条 国家根据数据在经济社会发展中的重要程度，以及一旦遭到篡改、破坏、泄露或者非法获取、非法利用，对国家安全、公共利益或者公民、组织合法权益造成的危害程度，对数据实行分级分类保护。
	《个人信息保护法（草案）》	第五十条 个人信息处理者应当根据个人信息的处理目的、处理方式、个人信息的种类以及对个人的影响、可能存在的安全风险等，采取必要措施确保个人信息处理活动符合法律、行政法规的规定，并防止未经授权的访问以及个人信息泄露或者被窃取、篡改、删除：(一)制定内部管理制度和操作规程；(二)对个人信息实行

		分级分类管理；(三)采取相应的加密、去标识化等安全技术措施；(四)合理确定个人信息处理的操作权限,并定期对从业人员进行安全教育和培训；(五)制定并组织实施个人信息安全事件应急预案；(六)法律、行政法规规定的其他措施。
数据匿名化	《民法典 人格权编》	第一千零三十八条 第一款 信息处理者不得泄露或者篡改其收集、存储的个人信息；未经自然人同意，不得向他人非法提供其个人信息，但是经过加工无法识别特定个人且不能复原的除外。
	《网络安全法》	第四十二条 第一款 网络运营者不得泄露、篡改、毁损其收集的个人信息；未经被收集者同意，不得向他人提供个人信息。但是，经过处理无法识别特定个人且不能复原的除外。
	《个人信息保护法（草案）》	第二十四条 个人信息处理者向第三方提供匿名化信息的，第三方不得利用技术等手段重新识别个人身份。

资料来源：根据公开资料整理

（二）隐私保护计算技术合规应用探讨

1. 数据最小化

数据最小化原则（data minimization principle）几乎是世界各国个人信息保护立法中共通的原则。**数据最小化原则可被理解为要求企业和公共机构收集和使用个人信息时，以实现产品和服务目的为标准，在功能可实现的前提下在最小范围内收集收据。**无论是欧盟《通用数据保护条例》（General Data Protection Regulation,GDPR）、美国《加州消费者隐私法案》（California Consumer Privacy Act of 2018，CCPA），还是印度[16]、巴西[17]等国家正在推动的个人信息保护立法，都把个人数据最小化原则作为基本的原则之一。如上表 3 所示，我国《民法典》《网络安全法》《数据安全法（草案）》《个人信息保护法（草案）》《电子商务法》《消费者权益保护法》等其中的相关规定均体现了个人数据最小化原则。**数据最小化原则之所以成为大多数国家个人信息保护立法的基本原则之一，其主要原因在于个人信息处理者存在扩大收集个人信息的倾向，在大数据时代尤为突出，需从法律上加以限制，使其保持克制。使数据收集行为受到数据处理目的限制。**

此外，根据英国信息专员办公室（Information Commissioner's Office, ICO）发布的《数据最小化指南》[18]，个人数据应该是充分的、

[16] 印度《个人数据保护法》（Personal Data Protection Act ）

[17] 巴西《通用数据保护法》（Lei Geral de Proteção de Dado, LGPD）

[18] 《数据最小化指南》（Guidance on the data minimisation principle Principle (c): Data minimisation ）

相关的及只限于处理这些数据的目的所需，既数据最小化。并根据**数据处理目的**围绕**充分性、相关性和限制性**给出评估是否满足数据最小化的方法：**首先**评估是否仅收集了实现该特定数据处理目的的数据；**其次**是有足够的数据实现该特定数据处理目的；**最后**是周期性检查所持有的数据，并删除不需要的数据。

满足数据最小化原则，是法律要求的一部分。在数字经济高速发展的大背景下，能够实现数据最小化原则的技术具有巨大的应用价值，值得鼓励，尤其从降低风险事件出现的可能性和现实损失这一角度来看，如英国信息专员办公室在其发布的《人工智能和数据保护指南》[19]中提及了隐私保护计算技术的多种类型，并鼓励通过使用如联邦学习、差分隐私等隐私保护计算技术来减少在机器学习模型训练阶段及预测阶段处理个人信息。并且说明了虽然有些技术不需要为遵守数据最小化要求而作出任何妥协，但更多技术仍然需要在数据最小化与其他合规或效用目标之间取得平衡。

(1) 联邦学习

联邦学习的原始数据始终存储在本地，不直接互相传输，直观上体现了数据最小化思想。以逻辑回归纵向联邦学习为例，训练模型需要双方共同计算梯度。最初的方法是将梯度直接暴露给对方或第三方，保留原始数据不出本地。但是梯度的本质是基于原始输入数据的函数，暴露梯度可能存在原始数据泄露的风险[20,21]。为进一步实现个

[19] 《人工智能和数据保护指南》（Guidance on AI and data protection）

[20] Li Z, Huang Z, Chen C, et al. Quantification of the leakage in federated learning[J]. arXiv preprint

人信息和隐私保护，一种方法是增加每次迭代训练的样本量，使得平均每个样本泄露的数据尽可能少；一种方法是使用差分隐私等方法，增强对中间梯度结果的保护。两种方法殊途同归，都以进一步减少中间计算过程的明文数据传输为原则，同样体现了数据最小化思想。此外在预测过程中，无限制地使用模型预测可能造成模型参数或样本数据的泄露。控制预测过程用法用量或使用更安全的密码学手段（如安全多方计算技术），可以降低预测过程的数据披露。总之联邦学习是以各个阶段尽可能降低数据共享来实现数据隐私保护的最大化。

(2) 安全多方计算

安全多方计算技术的特点就是在保证各参与方在结果计算的过程中，不会产生任何额外的信息泄露。以“百万富翁问题”为例，想知道两个富翁谁更富有，传统方案是由其中一个富翁张三通过直接采集另一个富翁李四的财富值以完成计算（或反之），但李四的财富值被张三获知，无法控制其财富信息的进一步传播蔓延。而安全多方计算的方案中，张三和李四运行一个密码学协议，双方仅计算得到“张三和李四谁更富”的结果，而不了解对方的具体财富值，也就不用担心对方会将数据用于其他处理目的。由此可见，传统方案中，我们一般需要通过法规制度来限制数据处理方不能“将数据用于与处理目的无关的场景”，而使用了安全多方计算技术之后，数据处理方是在技

arXiv:1910.05467, 2019.

[21] Zhu L, Liu Z, Han S. Deep leakage from gradients[C]//Advances in Neural Information Processing Systems. 2019: 14774-14784.

术角度就无法做到“将数据用于与处理目的无关的场景”，实现了目的受限的原则，进而满足了数据最小化要求（即是以实现产品和服务目的为标准，在功能可实现的前提下在最小范围内收集数据）。**需要注意的是，计算结果符合处理目的，这是方案满足数据最小化原则的重要前提。**以“百万富翁问题”为例，如果“张三和李四谁更富有”这个计算结果本身就与数据处理者提供的服务无关，不符合数据采集时的隐私政策，那么无论使用何种技术实现这一过程，都不符合数据最小化的原则。

(3) 机密计算

机密计算通过基于硬件的可信执行环境集中存储各参与方数据，并对这些数据进行操作。除事先被授权允许运行的代码之外，任何其他未授权的、恶意的代码都无法在可信环境中运行。**因此只要被授权运行的代码是按照数据最小化思想设计并实现的，那么整个机密计算的方案也就符合数据最小化思想。**在实际使用机密计算场景中，对代码目的进行确认（如利用远程认证机制进行检验）也是一个必不可少的关键步骤。

(4) 差分隐私

差分隐私是在保留统计学特征的前提下去除个体特征以保护用户隐私。当敌手试图从数据库中查询个体信息时将其混淆，使得敌手

无法从查询结果中辨别个体级别的敏感性[22]。具体而言，是通过结合统计计算及噪声添加实现。目前，差分隐私在隐私保护计算技术应用中，根据噪声添加及统计计算的先后执行顺序，可分为本地差分隐私和中心化差分隐私。例如谷歌 Chrome 浏览器和苹果 iPhone 手机，使用本地差分隐私对消费者点击广告或用户输入的单条记录（Record）先添加噪声，然后对添加噪声后的单条数据进行汇总，最后进行统计计算（如图 2 所示）。但是，由于本地差分隐私引入的噪声更大，仅能实现有限的统计计算，难以用于机器学习建模等复杂场景。

资料来源：根据公开资料整理

图 2 本地差分隐私

关于中心化差分隐私，主要分为基于多方计算的场景，以及基于单方计算的场景。基于多方计算的场景，数据提供方会对诸如联邦学习联合建模中的梯度或安全多方计算的中间值等内容添加噪声，防止通过梯度或其他中间值来逆推出关于数据提供方的训练样本、以及属性的特征分布等。基于单方计算的场景，首先是将原始数据库进行统计计算转变为不直接包含任何个人数据的统计信息，然后在统计信息之上添加噪声，进一步实现隐私保护（如图 3 所示）。与其他技术相

[22]杨强、刘洋、程勇、康焱、陈天健、于涵，《联邦学习》，电子工业出版社。

比，差分隐私技术在数据最小化中扮演的角色更特殊一些，主要用途有两类：一类是我们可以使用差分隐私技术，达到比“数据最小化”更高的要求；例如我们可以给一个本身已经满足了“数据最小化”的方案加入差分隐私，进一步降低其数据泄露风险；第二类是如果一个方案由于成本等种种原因，不得不传输或采集超出最小化目标之外的信息，可以使用差分隐私技术对这些信息增加干扰，对于数据下游的接收方能够获得的额外信息量更少，更符合数据最小化的思想。

资料来源：根据公开资料整理

图 3 基于单方计算的中心化差分隐私

(5) 同态加密

（全）同态加密技术支持在密文上进行计算操作，可以避免数据处理者接触数据明文，这与数据最小化原则的思想是相通的，同样能够减少数据泄露的现实风险。需要注意的是，这一最小化效果仅当数据处理者不了解数据密钥时才成立（请参考本节“数据分级分类”所述）。在目前业界的隐私保护计算应用中，同态加密一般不独立构成一个方案，而是作为安全多方计算、联邦学习等方案的一个组成部分，所以是否符合数据最小化原则，需要结合整体方案的处理过程和处理目的来判定。一般来说，将一个技术方案中的明文计算替换为同等功

能的同态计算，可以起到降低数据泄露风险的作用，更符合数据最小化思想。

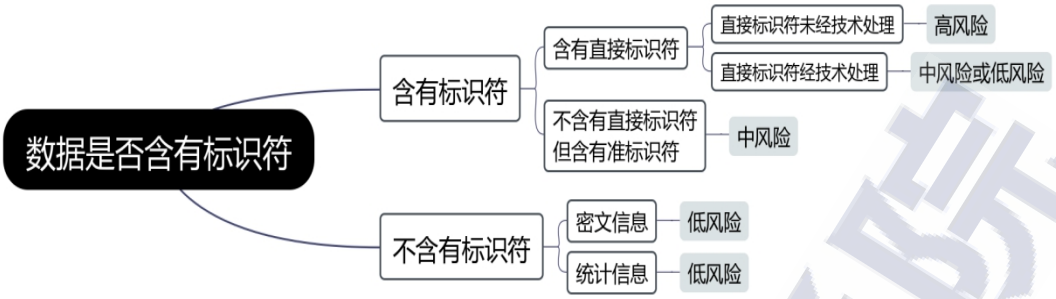
2. 数据分级分类

数据分级分类是开展数据安全治理的起始点，也是数据精细化管控的重要手段，在数据立法中也被反复强调。引入数据分类分级这一基础性数据安全管理方法，综合考虑数据属性、特点、数量、质量、格式、重要性、敏感程度等因素，对数据资源进行分类分级，梳理出非敏感、低风险等级、权属相对明确的数据资源，以要素形式优先进入数据交易市场，同时明确在市场交易过程中应配备的安全保护措施，可以在最大限度释放数据价值的同时，又兼顾数据安全和对个人隐私的保护[23]。分级分类大多从数据角度静态地进行评定，而本节尝试从整体解决方案流程角度进行评估探讨。

隐私保护计算技术可以减少不必要的数据披露及传播，一定程度上体现了数据最小化的思想。但是在一个应用中使用隐私保护计算技术，并不直接表明该应用满足了相关法律法规的要求，还需要视具体应用场景进行针对性分析。**本小节尝试从个人信息分级分类保护角度，根据隐私保护计算过程中数据是否含有标识符[24]信息来评估个人信息泄露风险程度的高低，对隐私保护计算技术的应用场景和数据进行分级分类，提供一种开放性的思路（如图 4 所示）。其他领域或其他类别的数据可根据其具体安全保护目标研提不同的分级分类方法。**

[23] 中国信通院《强化数据分类分级安全管理，推进完善数据要素市场化配置》

[24] 标识符 identifier：微数据中的一个或多个属性，可以实现对个人信息主体的唯一识别。来源 GB/T 37964—2019，定义 3.6



资料来源：根据公开资料整理

图 4 基于风险的隐私保护计算技术应用场景分级方法

1）高风险：含有未经技术处理的直接标识符[25]。该类数据能够**直接关联（或是在简单的外部条件帮助下可以关联）到唯一自然人，**一旦泄露可以直接暴露用户的身份或者用户的更多信息，因此将其判定为高风险数据，相应场景为高风险场景。如包含未经技术处理的身份证号、护照号、驾驶证号、手机号、银行卡号等直接标识符的数据。

2）中风险：**一类**是不含有直接标识符，但含有准标识符[26]的数据。该类数据由于其不含有直接标识符所以通常不再直接关联到唯一自然人，但在简单的外部帮助下，可以间接关联到自然人（如披露的住址、IP 地址、生日、年龄等）；**另外一类**是含有直接标识符且对直接标识符进行技术处理的数据（如采用泛化技术[27]、假名化[28]等技术对

[25] 直接标识符：direct identifier：微数据中的属性，在特定环境下可以单独识别个人信息主体。来源 GB/T 37964—2019，定义 3.7

[26] 准标识符：quasi-identifier：微数据中的属性，结合其它属性可唯一识别个人信息主体。来源 GB/T 37964—2019，定义 3.8

[27] 泛化技术是指一种降低数据集中所选属性粒度的去标识化技术，对数据进行更概括、抽象的描述。来源 GB/T 37964—2019 ，附录 A.5

[28] 假名化技术是一种使用假名替换直接标识符（或其他准标识符）的去标识化技术。来源 GB/T 37964—2019 ，附录 A.4

直接标识符进行处理），不再直接关联到自然人，但是仍具有一定概率的关联到自然人（如证件号、手机号中，将较少位数字标为*号）。

3）低风险：含有经过技术处理后的直接标识符数据，但重识别概率较低的数据（如证件号、手机号中，将较多位数字标为*号）。此外，统计信息和加密信息可以归为低风险。统计信息是指对一个包含足够多（例如 20 人以上）自然人群体信息的统计结果，如平均值，方差等[29]；加密信息是指在当前经典计算机的处理能力条件下无法解密的加密数据，且信息接收方不掌握其解密密钥。对于某些类型的中低风险类信息（如统计信息或脱敏信息），若数据规模特别大时，可以考虑提高其风险级别。

一般而言，在应用的某一环中使用隐私保护计算技术可以有效的降低这一环的数据泄露风险，但是整体的风险是否降低仍有待进一步确定。整个应用的风险应遵循“木桶原则”，即其数据泄露风险程度取决于其最弱一环。对于一个隐私保护计算应用，首先要分析其各个环节，包括数据传输、计算过程和计算结果是否会产生上述类型的风险。若有，则需要进一步分析其风险程度。对于高、中风险，需要获取用户明示同意，或者考虑进一步提高隐私保护计算技术的强度以降低这一环的风险；而对于低风险，需对其是否包含个人信息进行研判。

以下给出一例风险评估案例。《UN Handbook on Privacy-Preserving Computation Techniques》中有一个典型案例，该案例是对爱沙尼亚 1000 多万条纳税记录和 60 多万条自然人的教育信息进行联合分

[29] 《信息安全技术-个人信息去标识化效果分级与评定（草案）》

析，其目的是统计各类学生毕业后的工作和收入情况。主要的难点在于纳税数据和教育数据由不同的机构持有，而且相关法律法规规定，禁止数据流出本机构；为此技术人员设计了一个基于安全多方计算的密码学方案，保证数据安全的同时，完成了统计目的。

对这一案例而言，我们可以使用上述拟定的风险评估方案进行评估：**首先**，假设若不采取隐私保护计算技术，则为了完成这一统计，其中一个机构不可避免的需要将大量数据交付给另一机构；如此大量的教育信息和纳税信息传播无疑属于高风险，仅在涉及的每个自然人都明示同意时才能这么做，一定程度上增加了企业的成本，而获取这么多人的明示同意是不现实的，存在制约数据要素流通的可能；**然后，**本案例中的技术人员使用了基于秘密共享的安全多方计算技术，各机构在整个过程中所见的信息仅包括两类内容：秘密共享的随机分量（在各机构不共谋的前提下，属于低风险），以及最后的统计结果（低风险）。爱沙尼亚相关安全机构（The Data Protection Agency）对该案例进行了详细的评估，认为整个处理过程中没有触及个人信息。此外，欧盟支持的PRACTICE[30]项目也认为，使用最先进的（state of the art）密码学技术是可以达到GDPR要求的数据匿名化效果的，但是也要将私钥泄露的可能性，以及共谋风险考虑在内。

3. 数据匿名化

个人信息概念的界定是个人信息保护立法的核心问题和逻辑起点[31]，直接关系到法律保护对象的边界。而一旦个人信息进行了匿名

[30] PRACTICE D31.3 Evaluation and integration and final report on legal aspects of data protection

[31] 何波：《试论个人信息概念之界定》，载《信息通信技术与政策》，2018 年第 6 期

化处理，就不再具有个人信息属性。《民法典》强调了个人信息是“识别”特定自然人的各种信息，而《个人信息保护法（草案）》中规定的个人信息为“与已识别或可识别的自然人有关”。过窄的“个人信息”定义无法实现对个人信息的充分保护，过宽的定义则一定程度上阻碍了数据要素的流通，无法促进数据要素市场培育。“能够识别”个人和与“自然人有关”的不同判断标准是否会对具体信息类型的定性产生影响还有待学界进一步探讨以及司法实践的检验[32]。数字经济高速发展的当下，数据流通共享与协同应用并非必须识别特定个人。

根据《个人信息保护法（草案）》，个人信息不包括匿名化处理后的信息。匿名化是指**个人信息经过处理无法识别特定自然人且不能复原的过程**，而相较于“去标识化”的定义，“匿名化”定义没有了“不借助额外信息”的条件。即是匿名化的数据即使允许借助额外信息，也不能复原。但实际应用中，如果给予足够多的资源和额外信息，没有任何技术能够实现绝对的“无法复原”。例如，即使是经过加密处理的数据，只要数据接收者能够拿到密钥这一额外信息，该数据也可复原。因此，应当考虑获取额外信息、复原数据所需花费的资源成本。例如对于像全同态加密、安全多方计算等以密码学为基础的隐私保护计算技术来说，其安全性的基础来自于加密算法的强度、加密密钥的长度和密钥管理的安全性。**因此，当启用了最先进的加密技术，且执行了严格的密钥管理，数据接收者无法获取解密密钥，在当前经典计算机的算力下，在一定的时间期限范围内（例如 PRACTICE 项**

[32] 中国信息通信研究院《互联网法律白皮书 2020》

目中的假设 10 年）想要破解几乎又是不可能的，则可认为构成匿名化。因此，实现现实中的解密只有在“不合理的努力”下（如违法攻击对方机构的密钥管理服务器等）才会发生。这类场景的安全性判定对技术细节的了解程度要求较高。鉴于此，一方面，鼓励具有相应技术能力的机构进行试点检测，另一方面，借鉴前文所述的欧盟 PRACTICE 项目，成立由业内专家组成的委员会对具体案例进行评估。

四、发展展望

在个人信息和隐私保护合规监管日渐趋严的大背景下，如何通过隐私保护计算技术进一步合规地助力数据流通共享和协同应用，我们提出如下展望。

（一）坚持良法善治，完善法律法规体系

立足我国国情，充分借鉴国内外个人信息和隐私保护经验，建议国家相关监管部门加快推动《网络安全法》配套立法、《数据安全法（草案）》《个人信息保护法（草案）》等法律法规立法进程。对于涉及数据流通与协同应用中的安全合规问题开展深入研究，在法律层面明确不同利益主体在数据收集、使用和协作等环节中所要遵循的原则（例如合法、公开、透明、必要等），细化不同主体的法律义务及责任，加强相关法律法规的可落地性，切实做好对个人信息、隐私保护与数字经济发展的平衡。如在个人信息收集环节已明确获得用户同

意（或存在其他合法性事由），且在隐私保护计算过程中并未获得相关个人信息，是否需再次告知用户并获得同意等。如需再次获取用户同意可能会增加企业成本，进而制约了数据要素的流通共享与应用，仍值得深入关注和探讨。

（二）强化标准引领，加快标准体系建设

聚焦联邦学习、安全多方计算、同态加密、差分隐私、机密计算等隐私保护关键技术，鼓励企业实质性参与数据流通共享与协同应用相关技术、产品与安全等标准的制修订工作，研制可操作性强、量化指标明确、与实际技术发展水平衔接紧密，更具参考价值的标准。例如加快关于“匿名化”“去标识化”等相关技术操作指引以及金融等重要场景隐私保护计算技术应用安全要求等相关标准的制定。此外，建议国家及相关监管部门积极开展数据领域国际交流与合作，参与数据安全、个人信息和隐私保护相关国际规则和标准的制修订工作，提升我国网络安全国际竞争力及标准话语权。充分借鉴国外在数据协同安全应用等方面的先进经验，促进我国个人信息和隐私保护标准体系不断完善。

（三）立足风险评估，强化全流程风险防控

隐私保护计算技术的应用涉及数据全生命周期的多项关键环节，建议从隐私保护计算技术应用的全流程出发，围绕数据源的合法性、隐私保护计算需求建立的合理性、隐私保护计算过程中的个人信息和

隐私泄露风险、以及隐私保护计算目标（隐私保护计算结果）的个人信息和隐私泄露风险，进行隐私保护计算全流程的风险评估（如本报告所述关于分级分类的风险评估），进而建立闭环式的数据流通共享与协同应用风险管控机制，以支持更精细化的数据管理和使用。例如在数据源合法性评估环节，若某数据参与方的数据为未经用户授权获得的，或虽经用户授权收集但未将后续使用方式告知并征得用户同意，则对其数据的使用行为本身即违规，并不因使用隐私保护计算技术而合法化。

（四）深化双轮驱动，明晰安全与发展并举

统筹发展和安全是对历史上大国兴衰经验的深刻总结，是对发展和安全辩证统一关系的深刻认识和把握。建议**一是**推动隐私保护计算应用落地，在实践中解决安全问题。基于当前技术发展水平，不切实际的隐私期待极易使得数据价值挖掘陷入“囚徒困境”。建议立足我国国情，建立健全数据流通与协同应用机制，鼓励不同市场主体开展面向实际产业需求的数据协同应用工程试点，在实践中探索基于隐私保护技术实现个人信息和隐私安全保护的新方法。**二是**提升隐私保护计算技术成熟度和产业化能力，以安全促发展。加大隐私保护计算等技术能力研究投入，围绕应用性能瓶颈、安全性证明、数据质量规范性等规模化应用难点，夯实关于安全多方计算、联邦学习、差分隐私等研究基础，不断提升个人信息和隐私保护能力，为数据流通共享和协同应用提供安全保障。

（五）着力固本培元，造就高水平人才队伍

隐私保护计算技术是一门多学科跨领域的综合技术体系，对人工智能、密码学、数据科学、甚至是法律合规的复合型人才有着迫切的需求。数据要素的流通共享与协同应用是一个厚积薄发的过程，隐私保护计算人才培养是“固本培元”的工作。当前，隐私保护计算技术及合规方面的人才队伍规模仍需扩大、专业素养有待进一步提升，我们要围绕学科设置、培养平台建设等方面建立健全涵盖密码学、人工智能、法律合规等重要学科的复合型人才的培养计划，打造体系化、高层次的、真正意义上的高精尖“专业”人才队伍。

中国信息通信研究院 安全研究所

地址：北京市海淀区花园北路 52 号

邮政编码：100191

联系电话：010-62300264

传真：010-62300264

网址：www.caict.ac.cn