大数据安全标准化白皮书

（2018 版）

全国信息安全标准化技术委员会

大数据安全标准特别工作组

2018 年 4 月

《大数据安全标准化白皮书》（2018 版）

顾问指导组

赵泽良	高　林	胡　啸	杜　巍
顾建国	杜　虹	杨建军	陈吉学

全国信息安全标准化技术委员会大数据安全标准特别工作组

组　长：王建民

副组长：陈兴蜀

秘　书：金　涛

编写单位

清华大学
大唐电信科技产业集团
中国信息通信研究院
西北大学
阿里巴巴（北京）软件服务有限公司
深信服科技股份有限公司
北京中测安华科技有限公司
联想集团
北京天融信科技股份有限公司
成都秦川物联网科技股份有限公司
国际商业机器（中国）有限公司
中国信息安全测评中心
北京三未信安科技发展有限公司
中国信息安全认证中心
CSA 云安全联盟
湖南科创信息技术股份有限公司
普华永道
医渡云（北京）技术有限公司
腾讯云计算（北京）有限责任公司
贵州国卫信安科技有限公司
甲骨文（中国）软件系统有限公司
中国电子技术标准化研究院
启明星辰信息技术集团股份有限公司
北京奇安信科技有限公司（360 企业安全集团）
浙江蚂蚁小微金融服务集团股份有限公司
中电长城网际系统应用有限公司
公安部第三研究所
勤智数码科技股份有限公司
四川大学
陕西省网络与信息安全测评中心
天津南大通用数据技术股份有限公司
北京奇虎科技有限公司
中国移动通信集团有限公司
西安电子科技大学
国家信息中心
新华三技术有限公司
山西智杰软件工程有限公司
上海数据交易中心有限公司
海信集团有限公司
西安未来国际信息股份有限公司
深圳市腾讯计算机系统有限公司

编写人员

王建民　刘贤刚　金　涛　谢安明　汪　坤　韩晓露　郑新华　孙　骞　叶晓俊
王　昕　李克鹏　闵京华　叶润国　刘伯仲　唐　迪　陈世武　李　正　李小丁
赵　泰　阮树骅　徐雨晴　吴少华　杨　帆　张　磊　孙　卡　程海旭　鲍旭华
都　婧　陈　湉　江为强　任兰芳　鹿淑煜　詹　阳　吴　迪　吕　欣　叶思海
孙晓军　方　明　石在辉　叶荣伟　张丽萍　赵元勋　张　伟　桂　丽　张敏翀
梁文韬　胡　影　李　怡　苗光胜　王永霞　孙茵茵　陈先来　宋好好　钱晓斌
代　威　金　丹

前言

中共中央总书记习近平 2017 年 12 月 8 日下午在主持中共中央政治局就实施国家大数据战略学习时强调，大数据发展日新月异，我们应该审时度势、精心谋划、超前布局、力争主动，深入了解大数据发展现状和趋势及其对经济社会发展的影响，分析我国大数据发展取得的成绩和存在的问题，推动实施国家大数据战略，加快完善数字基础设施，推进数据资源整合和开放共享，保障数据安全，加快建设数字中国，更好服务我国经济社会发展和人民生活改善。

大数据已经上升为国家战略，数据被视为国家基础性战略资源，各行各业的大数据应用风起云涌，大数据在国民经济发展中发挥的作用越来越大。伴随着大数据的广泛应用，大数据安全问题也日益凸显，大数据安全标准作为大数据安全保障的重要抓手越来越被重视。随着大数据安全标准化工作的开展，由于缺乏顶层设计和统筹规划，大数据安全标准之间的交叉重复开始出现。为了更好的引导未来大数据安全标准化工作有序开展，全国信息安全标准化技术委员会大数据安全标准特别工作组集众多成员单位之合力，联合大数据系统软件国家工程实验室、大数据流通与交易技术国家工程实验室、大数据协同安全技术国家工程实验室、医疗大数据应用技术国家工程实验室，梳理了大数据应用中面临的安全风险和挑战，研究了国内外大数据安全相关的法律法规，分析了大数据安全标准化需求和目前已有的相关标准，建立了大数据安全标准体系，并给出了大数据安全标准化工作建议，最终形成本版白皮书。

全书组织如下：

第 1 章介绍了本书的背景、目的及意义。

第 2 章从保障大数据安全和利用大数据保障网络空间安全两个方面对本书中的大数据安全进行了范围界定，阐述了大数据安全的发展状况和重要意义。

第 3 章从大数据平台与技术、数据安全和个人信息保护、国家社会安全和法规标准三方面分析了大数据安全面临的挑战。

第 4 章介绍了国内外大数据安全相关的法律法规、相关的标准化组织及相应大数据安全标准化工作情况，并介绍了国内外大数据安全相关标准。

第 5 章首先汇总了大数据安全标准化的需求，给出了大数据安全标准的分类，基于分类制定了大数据安全标准图谱，并介绍了大数据安全标准特别工作组已经开展的大数据安全标准工作，指出了急需开展的标准化重点工作。

第 6 章给出了大数据安全标准化工作建议。

附录 A 由各参编成员单位介绍了相关领域大数据应用的特点、安全风险、安全需求以及大数据安全标准化需求。

附录 B 由各参编成员单位介绍了各自在大数据应用中的安全标准应用实践情况，包括使用了哪些标准，成效如何以及基于实践的大数据安全标准化需求。

附录 C 介绍了大数据安全相关的其它网络资源。

附录 D 摘录了大数据安全标准相关的术语定义。

附录 E 介绍了信安标委标准工作程序。

附录 F 给出了缩略语。

参考文献部分列出了本白皮书编写过程中参考的相关文献。

由于时间仓促，水平有限，错误疏漏在所难免，针对此版白皮书如有任何意见或建议，敬请联系 jintao16@mail.tsinghua.edu.cn。

目录

第 1 章 导论

1.1 背景

随着大数据时代的到来，数据已经成为与物质资产和人力资本同样重要的基础生产要素。2013 年 7 月，习近平总书记指出：“大数据是工业社会的‘自由’资源，谁掌握了数据，谁就掌握了主动权”。2014 年 2 月 27 日，习近平总书记在中央网络安全和信息化领导小组第一次会议又进一步强调：“网络信息是跨国界流动的，信息流引领技术流、资金流、人才流，信息资源日益成为重要生产要素和社会财富，信息掌握的多寡成为国家软实力和竞争力的重要标志”。2017 年 10 月 18 日，习近平同志代表第十八届中央委员会向党的十九大作报告指出：“加快建设制造强国，加快发展先进制造业，推动互联网、大数据、人工智能和实体经济深度融合，在中高端消费、创新引领、绿色低碳、共享经济、现代供应链、人力资本服务等领域培育新增长点、形成新动能。”国家拥有的数据规模及运用能力已逐步成为综合国力的重要组成部分，对数据的占有权和控制权将成为陆权、海权、空权之外的国家核心权力。大数据正在重塑世界新格局，被誉为是“21 世纪的钻石矿”，更是国家基础性战略资源，正逐步对国家治理能力、经济运行机制、社会生活方式产生深刻影响，国家竞争焦点也已经从资本、土地、人口、资源的争夺扩展到对大数据的竞争。

在大数据时代，机遇与挑战并存，大数据开辟了国家治理的新路径，国家社会管理现代化面临着由碎片型向整体型、由应急型向预防型、由管控型向参与型、由粗放型向精细型，以及由静态型向动态型转变的五位一体的全面变革。大数据可以通过对海量、动态、高增长、多元化、多样化数据的高速处理，快速获得有价值信息，提高公共决策能力，从而逐步改变国家治理架构和模式。2016 年 10 月 9 日，习近平同志主持中共中央政治局第三十六次集体学习指出，我们要深刻认识互联网在国家管理和社会治理中的作用，以推行电子政务、建设新型智慧城市等为抓手，以数据集中和共享为途径，建设全国一体化的国家大数据中心，推进技术融合、业务融合、数据融合，实现跨层级、跨地域、跨系统、跨部门、跨业务的协同管理和服务。要强化互联网思维，利用互联网扁平化、交互式、快捷性优势，推进政府决策科学化、社会治理精准化、公共服务高效化，用信息化手段更好感知社会态势、畅通沟通渠道、辅助决策施政。2016 年 12 月，国务院印发《“十三五”国家信息化规划》（以下称《“十三五”规划》）建议提出：“实施国家大数据战略，推进数据资源开放共享。”因此，必须要正视大数据安全对社会发展带来的挑战，在大数据应用推广过程中，坚持安全与发展并重的方针，既充分发挥大数据价值，又避免数据泄露和个人隐私暴露等带来的安全问题，从而构建大数据安全保障体系，完善大数据社会管理体制机制建设和大数据国家战略，促进大数据时代的社会发展。

大数据作为产业发展的创新要素，不仅在数据科学与技术层面，而且在商业模式、产业格局、生态价值与教育层面，均带来了新理念和新思维。大数据与现有产业深度融合，在人工智能、自动驾驶、金融商业服务、医疗健康管理、科学研究等领域展现出广阔的前景，使得生产更加绿色智能、生活更加便捷高

效。大数据已经逐渐成为企业升级转型发展的有力引擎，在提升产业竞争力和推动商业模式创新方面发挥越来越重要的作用。为坚持技术创新与应用创新协同共进，国家战略加快经济社会各领域的大数据开发与利用，催生出更多的新产业、新业态、新模式，推动国家、行业、企业在数据的应用需求和发展水平方面进入新的阶段。在内部技术条件成熟、外部政策因素推动的激励下，我国涌现出一批从传统业务扩展甚至转型到大数据业务的企业，尤其是大数据细分市场，新应用新模式层出不穷，大数据产业呈现出蓬勃发展的态势。在此背景下，在跟踪研究大数据，提升对大数据的认知和理解的同时，也要充分意识到大数据安全与大数据应用是一体之两翼、驱动之双轮，必须从国家网络空间安全战略的高度认真研究与应对当前大数据安全面临的复杂问题。

大数据安全标准是大数据安全保障体系的基础组成部分，对大数据安全保障体系的实施起到引领和指导作用，主要体现在如下方面：一是规范大数据系统所有者、建设者、运营者对大数据平台和应用的安全建设、运维和风险管理；二是指导数据控制者完善数据采集、传输、存储、处理、交换、销毁等大数据全生命周期的管理，防控来自组织内外部的安全风险；三是通过规范大数据服务组织的基础安全管理、数据安全管理、系统安全建设、安全运维，提升系统防范安全风险的能力；四是规范行业大数据安全体系。在构建大数据安全标准体系时，须统筹考虑数据在行业之间或组织之间的交换与共享问题，以指导各行业的大数据安全建设和运营，支撑行业大数据应用的快速发展。为此，亟待从技术和产业发展角度加快推进大数据安全标准化工作，为我国大数据产业的健康发展提供有效支撑。

党中央、国务院高度重视大数据安全及其标准化工作，将其作为国家发展战略予以推动。2015 年 9 月，国务院发布《促进大数据发展行动纲要》（以下称《纲要》），要求“完善法规制度和标准体系”并“推进大数据产业标准体系建设”。2016 年 8 月，中央网信办、国家质检总局、国家标准委联合印发了《关于加强国家网络安全标准化工作的若干意见》，其中对加强网络安全标准化工作做出部署，要求围绕国家战略需求，开展关键信息基础设施保护、网络安全审查、工业控制系统安全、大数据安全、个人信息保护、网络安全信息共享等领域标准研制工作，从而提升标准信息服务能力和标准符合性测试能力，并在政策文件制定、相关工作部署时积极采用国家标准，积极参与制定相关国际标准并发挥作用。2016 年 11 月，第十二届全国人民代表大会常务委员会通过了《中华人民共和国网络安全法》，鼓励开发网络数据安全保护和利用技术。2016 年 12 月，国家互联网信息办公室发布了《国家网络空间安全战略》，提出实施国家大数据战略、建立大数据安全管理制度、支持大数据信息技术创新和应用要求。全国人大常委会和工信部、公安部等部门为加快构建大数据安全保障体系，相继出台了《加强网络信息保护的决定》、《电信和互联网用户个人信息保护规定》等一系列法规和部门规章制度。与此同时，相关标准研制机构还发布了国家和行业的网络个人信息保护相关标准，开展以数据安全为重点的网络安全防护检查。

为了贯彻落实国家大数据安全标准化工作要求，全国信息安全标准化技术委员会（以下简称“全国信安标委”，委员会编号 TC260）下设了大数据安全标准特别工作组（SWG-BDS），并在已开展的大数据安全相关标准工作的基础上，启动了《大数据安全标准化白皮书》的编制工作。

1.2 目的及意义

本白皮书从法律法规、政策、标准及产业应用等角度，勾画出大数据安全的整体轮廓，从国家安全、社会公共利益，保护公民、法人和其他组织的合法利益的角度，综合分析大数据安全标准化需求，从而为我国后续的大数据安全标准化工作提供指导。

本白皮书从多维度阐述大数据安全的重要性，分析大数据面临的安全风险和挑战，梳理国内外的大数据安全法规政策和标准化工作现状，制定大数据安全标准体系框架，提出开展大数据安全标准化工作的建议。

本白皮书旨在全面、客观的反映国内外大数据安全标准化相关工作基础和进展，根据业界最佳实践、认知水平，分享大数据安全标准特别工作组在大数据安全标准化领域的研究成果和实践经验，呼吁社会各界共同关注大数据安全的法规政策、技术创新和标准建设，为大数据产业的健康、安全、有序发展奠定坚实基础。

第 2 章 大数据安全

本章从保障大数据安全和利用大数据保障网络空间安全两个方面介绍了大数据安全的含义，阐述了我国大数据安全发展状况和大数据安全的重要意义。

2.1 大数据安全含义

2.1.1 保障大数据安全

当今社会进入大数据时代，越来越多的数据共享开放，交叉使用。针对关键信息基础设施缺乏保护、敏感数据泄露严重、智能终端危险化、信息访问权限混乱、个人敏感信息滥用等问题，急需通过加强网络空间安全保障、做好关键信息基础设施保护、强化数据加密、加固智能终端、保护个人敏感信息等手段，保障大数据背景下的数据安全。

大数据应用涉及海量数据的分散获取、集中存储和分析处理，表现出数据容量大、数据变化快等特征。同时，大数据所面临的安全威胁和攻击种类多，且攻击行为具有一定的隐蔽性、攻击特征变化快，单纯依赖传统信息安全防护技术来防范大数据攻击存在一定局限性。大数据环境下，虽然很多传统安全技术手段和管理措施可以在大数据环境下提供一定安全保障能力。但与此同时，大数据环境下，数据量巨大、数据变化快等特征导致大数据分析及应用场景更为复杂，这就需要我们对传统信息安全技术优化改进基础之上进行创新，从而改善海量数据分析场景下的应用和数据安全问题。

大数据安全主要是保障数据不被窃取、破坏和滥用，以及确保大数据系统的安全可靠运行。需要构建包括系统层面、数据层面和服务层面的大数据安全框架，从技术保障、管理保障、过程保障和运行保障多维度保障大数据应用和数据安全。

从系统层面来看，保障大数据应用和数据安全需要构建立体纵深的安全防护体系，通过系统性、全局性地采取安全防护措施，保障大数据系统正确、安全可靠的运行，防止大数据被泄密、篡改或滥用。主流大数据系统是由通用的云计算、云存储、数据采集终端、应用软件、网络通信等部分组成，保障大数据应用和数据安全的前提是要保障大数据系统中各组成部分的安全，是大数据安全保障的重要内容。

从数据层面来看，大数据应用涉及到采集、传输、存储、处理、交换、销毁等各个环节，每个环节都面临不同的安全威胁，需要采取不同的安全防护措施，确保数据在各个环节的保密性、完整性、可用性，并且要采取分级分类、去标识化、脱敏等方法保护用户个人信息安全。

从服务层面来看，大数据应用在各行业得到了蓬勃发展，为用户提供数据驱动的信息技术服务，因此，需要在服务层面加强大数据的安全运营管理、风险管理，做好数据资产保护，确保大数据服务安全可靠运行，从而充分挖掘大数据的价值，提高生产效率，同时又防范针对大数据应用的各种安全隐患。

2.1.2 利用大数据保障网络空间安全

国家互联网信息办公室 2016 年发布的《国家网络空间安全战略》指出：网络空间安全事关人类共同利益，事关世界和平与发展，事关各国国家安全，并提出要实施国家大数据战略，建立大数据安全管理制度，支持大数据、云计算等新一代信息技术创新和应用，为保障国家网络安全夯实产业基础，大数据安全已成为国家网络空间安全的核心组成。

随着大数据应用的蓬勃发展，安全行业正发生重大转变，利用大数据来保障网络空间安全成为一种趋势。网络空间安全涉及到网络空间中电磁设备、信息通信系统、运行数据、系统应用所面临的安全威胁防护，既要防止包括互联网、电信网与通信系统、传播系统与广电网、计算机系统、工业控制网络系统及其所承载的数据免遭破坏，也要防止对这些网络基础设施和重要信息系统的攻击或滥用波及到政治安全、经济安全、文化安全、社会安全和国防安全。针对上述安全风险，需要采取法律、管理、技术等综合手段来进行积极应对，确保网络基础设施、重要信息系统及其所承载数据的保密性、完整性、可鉴别性、可用性、可靠性、可控性得到保障。

目前，大数据技术已经广泛应用到网络空间安全中的网络安全态势感知、高级持续威胁（APT）检测、伪基站发现与追踪、反钓鱼攻击、金融反欺诈等领域，并不断有新的应用场景出现。

大数据是实现网络空间安全保障的重要技术。综合考虑当前大数据应用的特点，利用大数据技术构建网络空间安全防护体系，建设以数据为核心的安全防护系统，集成态势感知、人工智能综合分析等功能，利用大数据技术工具，将传统的事中检测和事后响应防御体系转变为包括事前评估预防、事中检测和事后响应恢复的全面安全防护体系，为网络空间安全带来新的管理理念和技术创新，从而大幅提升网络空间安全治理能力。

2.2 我国大数据安全发展状况

为了保障大数据安全和网络空间安全，我国网络安全企业近年来发展迅速，网络安全初创企业不断涌现，各种先进的安全技术也被及时引入到国内。中国信息通信研究院于 2017 年 9 月 19 日发布的《2017 网络安全产业白皮书》表明，我国网络安全产业在近几年步入快速发展的新阶段；网络安全领域创新活跃；网络安全企业实力有较大提高，出现了一批具有整合能力的龙头企业。我国网络安全企业的业务类型基本覆盖了大数据安全涉及的各方面，包括基础设施安全、应用安全、数据安全、身份与访问管理、云安全、安全管理、安全服务等领域，这些企业是我国大数据安全市场的主力军。

据赛迪顾问统计，2016 年我国信息安全市场（包括大数据安全市场）整体规模达到 336.2 亿元，比 2015 年增长 21.5%；其中信息安全硬件仍为主力，占比达到 50.7%，信息安全软件与服务分别占 37.7%、11.6%。经统计，截止 2017 年 6 月，我国在主板上市的网络安全企业共有 12 家，总市值 1171.49 亿元，营业收入 148.18 亿元。在新三板上市的网络安全企业共有 36 家，营业收入共计 23.54 亿元。

作为企业发展的聚集区和孵化区，大数据安全产业园区建设也已逐步展开。

例如，2017 年 5 月 26 日，贵阳市被授牌成为全国首个大数据安全示范试点城市，《贵阳市大数据安全保障体系及产业规划》提出了“1+1+3+N”的大数据安全发展总体思路。其中，第一个“1”，“大数据安全示范试点城市”已实现落地，成为推动大数据安全发展的载体；而第二个“1”，1 个大数据安全靶场也正在着力建设中；“3”表示构建“城市安全态势感知中心”、“城市安全监管中心”、“大数据安全创新中心”3 个中心；“N”表示在不同领域、不同行业，围绕数据安全以及网络安全构建 N 个不同的平台。目前，已经启动建设占地一千多亩的大数据安全产业示范区，预计到“十三五”末期，贵阳大数据安全产业园将成为国内大数据安全产业的重要聚集区和大数据安全产业地标。

大数据安全市场蓬勃发展，市场预期良好，但问题也不断暴露。由于缺乏相应的监管措施、配套的安全标准以及相应的产品检测机制，一些不具备相关资质和能力的企业看到商机后趁机涌入，导致安全市场的从业企业鱼龙混杂、良莠不齐，呈现出“野蛮发展”的态势，市场乱象频出，亟待规范和引导。

随着国家对大数据安全的高度重视，一批大数据安全相关的国家标准将陆续出台，将对规范市场秩序、扶持优质企业起到重要作用。

2.3 大数据安全的重要意义

大数据已经逐步应用于产业发展、政府治理、民生改善等领域，大幅度提高了人们的生产效率和生活水平。适应、把握、引领大数据，将成为时代潮流。在大数据时代，数据是重要的战略资源，但数据资源的价值只有在流通和应用过程中才能够充分体现出来。这就要求打破传统垂直应用中所形成的数据孤岛，形成适应大数据时代的数据湖，并需要数据在不同应用之间流动，这难免会出现数据泄露和滥用问题。在发展大数据的同时，也容易出现政府重要数据、法人和其他组织商业机密、个人敏感数据泄露，给国家安全、社会秩序、公共利益以及个人安全造成威胁。没有安全，发展就是空谈。大数据安全是发展大数据的前提，必须将它摆在更加重要的位置。

大数据系统自身安全防护具有重要意义。大数据的数据量大且相互关联，黑客一次成功的攻击就能够获得大量的数据，可以从大数据中快速捕捉到有价值的信息，尤其是个人敏感信息。因此，蕴含着海量数据和潜在价值的大数据成为网络攻击的显著目标。另一方面，传统网络安全防御技术以及现有网络安全行政监管手段与大数据安全保护的需求之间还存在较大差距：Hadoop 对数据的聚合增加了数据泄露的风险；NoSQL 技术在维护数据安全方面缺乏严格的访问控制和隐私管理；复杂多样的数据存储在一起，在数据管理和使用环节也容易形成安全隐患；安全防护手段的更新升级速度无法跟上数据量指数级增长的步伐等。因此，需要各层面、各环节保障大数据的安全。从数据的层面来看，大数据自身安全涉及到采集、传输、存储、处理、交换、销毁等各个环节，每个环节都面临不同的威胁，需要采取不同的安全保障措施，这些工作都是保障大数据安全的重要内容。从系统的层面来看，保障大数据自身安全需要从大数据系统的各部分采取措施，建立坚固、缜密、健壮的防护体系，保障大数据系统正确、安全、可靠的运行，防止大数据系统被破坏、被渗透或被非法使用。从服务的层面来看，规范大数据安全服务内容，提高对大数据安全的风险识别能力，建立健全的大数据安全保障体系，降低大数据安全隐患和安全事件发生

频率。

大数据在保障网络安全方面也具有重要作用。当前，各种网络攻击频发，攻击过程越来越复杂，网络攻击手段变得越来越隐蔽，传统的入侵检测、防御等网络安全产品往往难以奏效，采用大数据技术来检测高级网络攻击成为一种趋势。当前，为了利用大数据来加强企业信息安全能力，包括采用大数据技术来实现网络安全威胁信息分析，采用基于大数据的深度学习方法来替代传统入侵检测方法中的攻击特征模式提取，采用大数据技术来实现网络安全态势感知，以及对多步复杂网络攻击的检测、溯源和场景重现，都已开始应用。可以说，大数据技术将重塑未来的网络安全技术和产业发展趋势。

未来，在大数据应用的飞速发展过程中，大数据安全问题将始终伴随左右。针对大数据安全问题和安全风险，必须加大大数据安全技术的研究力度，必须以现有安全技术为依托，深入研究新型的大数据安全技术，比如同态加密技术等。确保大数据在存储、处理、传输等过程的安全性，在充分挖掘数据价值的同时保护用户隐私，从而避免因大数据安全问题而给用户的利益造成损失。需要进一步完善大数据安全相关法律体系建设，对数据权属界定、数据流动管理、个人信息保护等各种问题，给出明确规定。需要创新研制和推广大数据安全保护的产品和服务，基于大数据研制网络安全产品和服务，推动大数据安全市场发展，保障大数据时代的信息安全。

第 3 章 大数据安全挑战

大数据安全风险伴随大数据应用而生。我们在享受大数据福祉的同时，也面临着前所未有的安全挑战。随着互联网、大数据应用的爆发，系统遭受攻击、数据丢失和个人信息泄露的事件时有发生，而地下数据交易黑灰产也导致了大量的数据滥用和网络诈骗事件。这些安全事件，有的造成个人的财产损失，有的引发恶性社会事件，有的甚至危及国家安全。可以说当前环境下，大数据平台与技术、大数据环境下的数据和个人信息、大数据应用等方面都面临着极大的安全挑战，这些挑战不仅对个人有着重大影响，更直接威胁到社会的繁荣稳定和国家的安全利益。

3.1 大数据技术和平台安全挑战

伴随着大数据的飞速发展，各种大数据技术层出不穷，新的技术架构、支撑平台和大数据软件不断涌现，大数据安全技术和平台发展也面临着新的挑战。

3.1.1 传统安全措施难以适配

大数据的一个显著特点是数量巨大，即“Volume”，指的是要采集、存储和处理体量非常大的数据。同时，大数据还有另外一个特点是类型多，即“Variety”，指的是数据种类和来源非常多，类型上包括结构化、半结构化和非结构化数据，来源上包括生产、财务等业务数据，也包括文本、音频、视频、图片、地理位置信息等。这些海量、多源、异构等大数据特征导致其与传统封闭环境下的数据应用安全环境有很大区别。

大数据技术架构复杂，大数据应用一般采用底层复杂、开放的分布式计算和存储架构为其提供海量数据分布式存储和高效计算服务，这些新的技术和架构使得大数据应用的系统边界变得模糊，传统基于边界的安全保护措施将变得不再有效。如在大数据系统中，数据一般都是分布式存储的，数据可能动态分散在很多个不同的存储设备、甚至不同的物理地点存储，这样导致难以准确划定传统意义上的每个数据集的“边界”，传统的基于网关模式的防护手段也就失去了安全防护效果。

同时，大数据系统表现为系统的系统（System of System），其分布式计算安全问题也将显得更加突出。在分布式计算环境下，计算涉及的软件和硬件较多，任何一点遭受故障或攻击，都可能导致整体安全出现问题。攻击者也可以从防护能力最弱的节点着手进行突破，通过破坏计算节点、篡改传输数据和渗透攻击，最终达到破坏或控制整个分布式系统的目的。传统基于单点的认证鉴别、访问控制和安全审计的手段将面临巨大的挑战。

此外，传统的安全检测技术能够将大量的日志数据集中到一起，进行整体性的安全分析，试图从中发现安全事件。然而，这些安全检测技术往往存在误报过多的问题，随着大数据系统建设，日志数据规模增大，数据的种类将更加丰富。过多的误判会造成安全检测系统失效，降低安全检测能力。因此，在大数据环境下，大数据安全审计检测方面也面临着巨大的挑战。随着大数据技术的应用，为了保证大数据安全，需要进一步提高安全检测技术能力，提升安全

检测技术在大数据时代的适用性。

3.1.2 平台安全机制严重不足

现有大数据应用中多采用开源的大数据管理平台和技术，如基于 Hadoop 生态架构的 HBase/Hive、Cassandra/Spark、MongoDB 等。这些平台和技术在设计之初，大部分考虑是在可信的内部网络中使用，对大数据应用用户的身份鉴别、授权访问以及安全审计等安全功能需求考虑较少。近年来，随着更新发展，这些软件通过调用外部安全组件、修补安全补丁的方式逐步增加了一些安全措施，如调用外部 Kerberos 身份鉴别组件、扩展访问控制管理能力、允许使用存储加密以及增加安全审计功能等。即便如此，大部分大数据软件仍然是围绕大容量、高速率的数据处理功能开发，而缺乏原生的安全特性，在整体安全规划方面考虑不足，甚至没有良好的安全实现。

同时，大数据系统建设过程中，现有的基础软件和应用多采用第三方开源组件。这些开源系统本身功能复杂、模块众多、复杂性很高，因此对使用人员的技术要求较高，稍有不慎，可能导致系统崩溃或数据丢失。在开源软件开发和维护过程中，由于软件管理松散、开发人员混杂，软件在发布前几乎都没有经过权威和严格的安全测试，使得这些软件大都缺乏有效的漏洞管理和恶意后门防范能力。如 2017 年 6 月，Hadoop 的发行版本被发现存在安全漏洞，由于该软件没有对输入进行严格的验证，导致攻击者可以利用该漏洞攻击系统，并获得最高管理员权限。

物联网技术的快速发展，使得当前设备连接和数据规模都达到了前所未有的程度，不仅手机、电脑、电视机等传统信息化设备已连入网络，汽车、家用电器和工厂设备、基础设施等也将逐步成为互联网的终端。而在这些新终端的安全防护上，现有的安全防护体系尚不成熟，有效的安全手段还不多，急需研发和应用更好的安全保护机制。

3.1.3 应用访问控制愈加困难

大数据应用的特点之一是数据类型复杂、应用范围广泛，它通常要为来自不同组织或部门、不同身份与目的的用户提供服务。因而随着大数据应用的发展，其在应用访问控制方面也面临着巨大的挑战。

首先是用户身份鉴别。大数据只有经过开放和流动，才能创造出更大的价值。目前，政府部门、央企及其它重要单位的数据正在逐步开放，或开放给组织内部不同部门使用，或开放给不同政府部门和上级监管部门，或者开放给定向企业和社会公众使用。数据的开放共享意味着会有更多的用户可以访问数据。大量的用户以及复杂的共享应用环境，导致大数据系统需要更准确地识别和鉴别用户身份，传统基于集中数据存储的用户身份鉴别难以满足安全需求。

其次是用户访问控制。目前常见的用户访问控制是基于用户身份或角色进行的。而在大数据应用场景中，由于存在大量未知的用户和数据，预先设置角色及权限十分困难。即使可以事先对用户权限分类，但由于用户角色众多，难以精细化和细粒度地控制每个角色的实际权限，从而导致无法准确为每个用户指定其可以访问的数据范围。

再次是用户数据安全审计和追踪溯源。针对大数据量时的细粒度数据审计

能力不足，用户访问控制策略需要创新。当前常见的操作系统审计、网络审计、日志审计等软件在审计粒度上较粗，不能完全满足复杂大数据应用场景下审计多种数据源日志的需求，尚难以达到良好的溯源效果。

3.1.4 基础密码技术亟待突破

随着大数据的发展，数据的处理环境、相关角色和传统的数据处理有了很大的不同，如在大数据应用中，常常使用云计算、分布式等环境来处理数据，相关的角色包括数据所有者、应用服务提供者等。在这种情况下，数据可能被云服务提供商或其他非数据所有者访问和处理，他们甚至能够删除和篡改数据，这对数据的保密性和完整性保护方面带来了极大的安全风险。

密码技术作为信息安全技术的基石，也是实现大数据安全保护与共享的基础。面对日益发展的云计算和大数据应用，现有密码算法在适用场景、计算效率以及密钥管理等方面存在明显不足。为此，针对数据权益保护、多方计算、访问控制、可追溯性等多方面的安全需求，近年来提出了大量的用于大数据安全保护的密码技术，包括同态加密算法、完整性校验、密文搜索和密文数据去重等，以及相关算法和机制的高效实现技术。为更好地保护大数据，这些基础密码技术亟待突破。

如在上世纪七十年代提出的同态加密思想，由于这种加密算法可以直接对加密数据进行各种运算，运算后数据再解密的结果和对原始未加密数据进行同样运算的结果是一致的，因此同态加密非常适合于云计算环境中，可以从根本上解决将数据及其操作委托给第三方时的保密问题。尽管近几年来，同态加密技术已经得到了较大的发展，但是离大规模实用还有一定距离。考虑到应用需求和诱人的前景，同态加密算法亟待得到突破性创新发展。

3.2 数据安全和个人信息保护挑战

大数据中包含了大量的数据，而其中又蕴含着巨大的价值。数据安全和个人信息保护是大数据应用和发展中必须面临的重大挑战。

3.2.1 数据安全保护难度加大

大数据拥有大量的数据，使得其更容易成为网络攻击的目标。在开放的网络化社会，蕴含着海量数据和潜在价值的大数据更受黑客青睐，近年来也频繁爆发邮箱账号、社保信息、银行卡号等数据大量被窃的安全事件。分布式的系统部署、开放的网络环境、复杂的数据应用和众多的用户访问，都使得大数据在保密性、完整性、可用性等方面面临更大的挑战。

历史上发生过多起大数据平台数据泄露的安全事件。如 2016 年年底，因系统漏洞和配置问题，全球范围内数以万计的 MongoDB 系统遭到攻击，数百 TB 的数据被攻击者下载，涉及包括医疗、金融、旅游在内的诸多行业。一部分攻击者甚至在入侵 MongoDB 数据库后，将数据清除并向受害者索取赎金。又如在 2017 年 6 月，因 HDFS 服务器配置不当，导致全球近 4500 台服务器遭受攻击，泄露数据量高达 5120 TB。

针对数据的安全防护，应当围绕数据的采集、传输、存储、处理、交换、

销毁等生命周期阶段进行。针对不同阶段的不同特点，应当采取适合该阶段的安全技术进行保护。如在数据存储阶段，大数据应用中的数据类型包括结构化、半结构化和非结构化数据，且半结构化和非结构化数据占据相当大的比例。因此在存储大数据时，不仅仅要正确使用关系型数据库已有的安全机制，还应当为半结构化和非结构化数据存储设计安全的存储保护机制。

3.2.2 个人信息泄露风险加剧

由于大数据系统中普遍存在大量的个人信息，在发生数据滥用、内部偷窃、网络攻击等安全事件时，常常伴随着个人信息泄露。另一方面，随着数据挖掘、机器学习、人工智能等技术的研究和应用，使得大数据分析的能力越来越强大，由于海量数据本身就蕴藏着价值，在对大数据中多源数据进行综合分析时，分析人员更容易通过关联分析挖掘出更多的个人信息，从而进一步加剧了个人信息泄露的风险。在大数据时代，要对数据进行安全保护，既要注意防止因数据丢失而直接导致的个人信息泄露，也要注意防止因挖掘分析而间接导致的个人信息泄露，这种综合保护需求带来的安全挑战是巨大的。

在大数据时代，不能禁止外部人员挖掘公开、半公开信息，即使想限制数据共享对象、合作伙伴挖掘共享的信息也很难做到。目前，各社交网站均不同程度地开放其所产生的实时数据，其中既可能包括商务、业务数据，也可能包括个人信息。市场上已经出现了许多监测数据的数据分析机构。这些机构通过对数据的挖掘分析，以及和历史数据对比分析、和其他手段得到的公开、私有数据进行综合挖掘分析，可能得到非常多的新信息，如分析某个地区经济趋势、某种流行病的医学分析，甚至直接分析出某个人的具体个人信息来。

个人信息泄露产生的后果将远比一般数据泄露严重，2016 年 8 月，犯罪团伙利用非法获取得到的数万条高考考生信息实施诈骗，山东女孩徐某因学费被骗出现心脏骤停，最终不幸逝世。近几年来，个人信息泄露的事件时有发生，如在 2015 年 5 月，美国国税局宣布其系统遭受攻击，约 71 万人的纳税记录被泄露，同时约 39 万个纳税人账户被冒名访问； 2016 年 12 月，雅虎公司宣布其超过 10 亿的用户账号被黑客窃取，相关信息包括姓名、邮箱口令、生日、邮箱密保问题及答案等内容。

需要注意的是，如经过“清洗”、“脱敏”后的数据也不能说肯定是安全的。如 2006 年，为了学术研究，美国在线（AOL）将 65 万条用户数据匿名处理后，公开发布。而《纽约时报》通过综合推断，竟然分析出了数据集中某个匿名用户的真实姓名和地址等个人信息。因此，在大数据环境下，对个人信息的保护将面临极大的挑战。

3.2.3 数据真实性保障更困难

大数据的特点中，类型多（Variety），是指数据种类和来源非常多。实际上，在当前的万物互联时代，数据的来源非常广泛，各种非结构化数据、半结构化数据与结构化数据混杂在一起。数据采集者将不得不接受的现实是：要收集的信息太多，甚至很多数据不是来自第一手收集，而是经过多次转手之后收集到的。

从来源上看，大数据系统中的数据来源可能来源于各种传感器、主动上传

者以及公开网站。除了可信的数据来源外，也存在大量不可信的数据来源。甚至有些攻击者会故意伪造数据，企图误导数据分析结果。因此，对数据的真实性确认、来源验证等需求非常迫切，数据真实性保障面临的挑战更加严峻。

事实上，由于采集终端性能限制、鉴别技术不足、信息量有限、来源种类繁杂等原因，对所有数据进行真实性验证存在很大的困难。收集者无法验证到手的数据是否是原始数据，甚至无法确认数据是否被篡改、伪造。那么产生的一个问题是，依赖于大数据进行的应用，很可能得到错误的结果。

如在 2008 年，Google 发布一款名为“谷歌流感趋势”（Google Flu Trends，GFT）的产品。该产品的基本思路是：搜索流感相关主题的人数与实际患有流感症状的人数之间存在着密切的关系，用大数据分析网络上用户的搜索词有助于了解流感疫情。该产品在 2008 年大获成功，基于用户的搜索数据，比美国疾病预防控制中心（Centers for Disease Control and Prevention）提前两个星期预测到了流感的爆发。但是，消息公布后，众多的网民都对这个预测很感兴趣，于是网络中出现了大量的类似搜索记录，从而导致了很多“虚假”的数据记录到搜索数据中。所以后来该产品的预测结果就不准确了，尤其是到了 2012 年，偏差最大甚至高出了标准值一倍多。因此，在大数据环境下，对数据真实性保障面临巨大的挑战。

3.2.4 数据所有者权益难保障

数据脱离数据所有者控制将损害数据所有者的权益。大数据应用过程中，数据的生命周期包括采集、传输、存储、处理、交换、销毁等各个阶段，在每个阶段中可能会被不同角色的用户所接触，会从一个控制者流向另一个控制者。因此，在大数据应用流通过程中，会出现数据拥有者与管理者不同、数据所有权和使用权分离的情况，即数据会脱离数据所有者的控制而存在。从而，数据的实际控制者可以不受数据所有者的约束而自由地使用、分享、交换、转移、删除这些数据，也就是在大数据应用中容易存在数据滥用、权属不明确、安全监管责任不清晰等安全风险，而这将严重损害数据所有者的权益。

数据产权归属分歧严重。数据的开放、流通和共享是大数据产业发展的关键，而数据的产权清晰是大数据共享交换、交易流通的基础。但是，当前的大数据应用场景中，存在数据产权不清晰的情况。如大数据挖掘分析者经过对原始数据集处理后，会分析出新的数据，这些数据的所有权到底属于原始数据所有方，还是挖掘分析者，目前在很多应用场景中还是各执一词，没有明确的说法。又如在一些提供交通出行、位置服务的应用中，服务提供商在为客户提供导航、交通工具等服务时，同时记录了客户端运动轨迹信息，对于此类运动轨迹信息的权属到底属于谁，以及是否属于客户端个人信息，到目前为止，分歧仍然比较大。对此类数据权属不清的数据，首要解决的是数据归谁所有、谁能授权等问题，才能明确数据能用来干什么、不能用来干什么，以及采用什么安全保护措施，尤其是当数据中含有重要数据或个人信息的时候。

3.3 国家社会安全和法规标准挑战

大数据正日益对全球经济运行机制、社会生活方式和国家治理能力产生重要影响。全球范围内，运用大数据推动经济发展、完善社会治理、提升政府服

务和监管能力正成为趋势。与此同时，随着大数据的应用和发展，数据量越来越大、内容越来越丰富、交流领域越来越广、应用越来越重要，大数据的安全问题引发了世界各国的普遍担忧。可以说，大数据时代的到来在给我们带来机遇的同时，也给国家安全、社会治理以及法规标准制定等带来了巨大的挑战。

3.3.1 国家安全深受大数据影响

国家安全是伴随着国家的出现而产生的，它是一个国家生存和发展的前提。随着时代发展，当前国家安全的内容已发展的十分丰富，包含了政治安全、国土安全、军事安全、经济安全、文化安全、社会安全、科技安全、信息安全、生态安全、资源安全、核安全等内容。这些内容相互联系、相互作用，影响着整个国家安全。

大数据不仅仅带来了技术和产业的变更，更是改变了我们的工作方式、生活方式乃至思维模式。大数据是信息化发展的新阶段，运用大数据可以提升国家治理现代化水平，通过建立健全大数据辅助科学决策和社会治理的机制，有助于推进政府管理和社会治理模式创新。

信息技术与经济社会的交汇融合引发了数据迅猛增长，数据已成为国家基础性战略资源。而同时，大数据的应用范围越来越广泛，国家的政治、经济、军事、文化等各个领域都离不开数据和数字基础设施。各类大数据平台承载着海量的数据资源，其中不乏大量敏感资源和重要数据，必然会成为包括黑客在内的各类敌对势力对一个国家进行网络攻击的重要目标。实际上，各类数据已经成为一些不法分子和敌对势力用来策划、实施、推动各种违法犯罪活动的理想工具，对国家安全和社会稳定造成了极大的破坏。上升到国家战略层面，涉及国计民生的关键信息基础设施的大数据资源一旦受到破坏，将使得国家在政治、经济、军事等各领域受到巨大的损失。

面对汹涌的数据洪流，站在国家安全的角度来思考和研究大数据安全，已经成为一个紧迫而现实的挑战。大数据全球化、开放化的特点，使国家的“信息边疆”不断拓展和延伸。大数据安全和国家安全息息相关，没有大数据安全，就没有真正意义上的国家安全。

3.3.2 社会治理面临大数据挑战

大数据应用能够揭示传统技术方式难以展现的关联关系，推动政府数据开放共享，促进社会事业数据融合和资源整合，将极大提升政府整体数据分析能力，为有效处理复杂社会问题提供新的手段。建立“用数据说话、用数据决策、用数据管理、用数据创新”的管理机制，实现基于数据的科学决策。但是，从我国信息化发展的现实情况看，“不敢共享开放”、“不会共享开放”的情况依然较为普遍。相关人员担心数据共享开放会引起信息安全问题，担心数据泄密和失控。尤其是掌握大量数据的各级政务部门，因大数据安全措施不到位，导致他们对数据不敢共享开放，也不会实施安全地共享开放。因此，加强大数据环境下的网络安全问题研究和基于大数据的网络安全技术研究，建立健全大数据安全保障体系，切实保障数据安全，才能确保大数据“敢共享开放”和“会共享开放”，才能真正促进社会发展。

此外，创新社会治理，是我国应对社会转型、化解社会矛盾、协调利益关

系所面临的一项重大战略任务。针对目前社会治理领域普遍存在的一些问题，大数据技术通过对海量数据的快速收集与挖掘、及时研判与共享，成为支持社会治理科学决策和准确预判的有力手段，为转型期的社会治理带来了新机遇。而现实问题是，在大数据时代，可以说每个人都是数据的制造者、传递者和消费者，大量现实问题在虚拟的网络环境中讨论和传播，其中不乏存在大量的误导、篡改及谣传的信息。一方面，这些虚假、错误的信息进入到社会治理的数据集中后，将会误导基于大数据的科学决策，影响社会治理重点和效果；另一方面，虚假、错误的信息不被及时发现和处理，极有可能带来恶劣的负面效果，甚至导致爆发社会群体性事件。因此，如何甄别大数据中虚假和错误信息对社会治理带来了巨大挑战。

3.3.3 大数据安全法规标准尚需完善

大数据应用的场景越来越多，越来越重要，因此，要科学规范利用大数据并切实保障数据安全，在完善法规制度和标准体系方面也将面临着不小的挑战。

一方面，大数据的发展推动了经济发展，但也给监管和法律带来了新的挑战。法律带来的是稳定的预期和权利义务关系的平衡。大数据以及它给政治、经济、社会带来的深刻变革，终将需要法律规范的保障。《促进大数据发展行动纲要》指出，推进大数据健康发展，要加强政策、监管、法律的统筹协调，加快法规制度建设。要制定数据资源确权、开放、流通、交易相关法规，完善数据产权保护法规。通过积极研究数据开放、保护等方面的法规，有利于实现对数据资源的采集、传输、存储、处理、交换、销毁的规范管理，可以促进数据在风险可控原则下最大程度开放，明确市场主体大数据的权限及范围，界定数据资源的所有权及使用权，加强对数据滥用、侵犯个人信息安全等行为的管理和惩戒。如通过制定个人信息方面的法规制度细则，可以界定哪些数据属于个人信息，如非法使用则将受到相应的惩戒；又如通过制定跨境数据流动方面的法规制度细则，可以加速形成跨境数据安全流动框架，明确相应的部门职责、数据分类管理要求以及数据主体的权利和义务等。

另一方面，大数据的发展也给标准规范配套带来了新的挑战。标准是法规制度的支撑，肩负着规范市场客体质量和技术要求的重要职能。因此，除了在立法层面要明确数据保护方面的法规外，还应制定相应的数据采集、储存、处理、推送和应用的标准规范。通过制定符合实际的大数据应用和安全标准，能有效促进大数据安全应用，从而既能引导、规范、促进大数据的发展，又确保了数据开放共享、个人信息保护需求和安全保障需求之间的平衡。如制定了个人信息分类、责任原则、保护要求和安全评估方面的标准内容，有利于更好地规范实施个人信息的安全采集、存储和处理过程，防止个人信息被误用和滥用；又如制定了数据确权、访问接口、服务安全要求等标准内容，有利于建立安全的大数据市场交易体系，促进大数据交易流通的发展。

第 4 章 大数据安全法规政策和标准化现状

为积极应对大数据安全风险和挑战，确保大数据产业的健康发展，各国政府历来都非常重视大数据相关法规政策和标准的建设，以便对大数据安全进行规范。法律法规作为约束大数据用户行为的规范化文件，是确保大数据平台及大数据应用安全可控，防范大数据服务安全风险，维护国家安全和公共利益的重要手段。本章介绍国内外大数据安全相关的法规政策和标准化现状。

4.1 大数据安全法规政策现状

大数据安全相关的法规、政策环境是大数据行业发展的基础和保障，是大数据安全标准制定的重要依据，我国及世界各国充分重视大数据相关法律法规的建设与制定，为大数据发展营造了健康的发展环境，本节将介绍国内外大数据相关安全法规的发展及政策环境。

4.1.1 国外数据安全法律法规和政策

数据保护是大数据安全的重要基础和组成部分。美国、欧盟、俄罗斯、新加坡等网络安全产业发展强国先后颁布了众多的数据保护法律法规。尽管这些国家制定的相关法律法规思路和策略不同，但涉及的要素是基本一致的。本节梳理了国外数据保护法律法规核心要素，通过分析比较这些国外的数据安全法律法规和政策，为我国后续制定和出台数据保护相关法律法规和行政规章以及标准提供参考。

4.1.1.1 各国现有数据保护法律法规情况

表 4.1 梳理了美国、欧盟、澳大利亚、俄罗斯、新加坡等已制定或发布的数据保护相关法律法规，总结起来这些国家的数据保护法律法规分两类：

1. 制定专门的数据保护法律法规，并明确相应的数据安全管理部门，如欧盟、俄罗斯、新加坡等。其中，俄罗斯进行数据保护的主要法律是《个人数据保护法案》，涉及到的主要监管部门是俄罗斯电信/信息技术和大众传媒联邦监管局（Roskomnadzor）。新加坡进行数据保护的主要法律是《个人数据保护法令》（PDPA）。同时，为了执行 PDPA，新加坡专门成立了个人数据保护委员会（PDPC）来承担 PDPA 的制定和实施工作。
2. 数据保护的相关要求分散地体现在本国各项法律法规及部门规章的相关条款中，但尚未颁布数据保护的专门法律法规，也未设置相应的管理部门，如美国、澳大利亚、日本等。

表 4.1 主要国家的数据保护法律法规

序号	法律法规和部门规章	发布/生效时间	备注
一、美国			
1	《隐私盾协议》（替代《安全港协议》）	2016 年发布	通用法律
2	《加州在线隐私保护法案》	2014 年生效	州法律
3	《联邦隐私法案》	2014 年发布	通用法律
4	《数字问责和透明法案》（FFATA）	2014 年发布	部门规章

序号	法律法规和部门规章	发布/生效时间	备注
5	《数字政府战略》	2012 年发布	通用法律
6	《开放政府指令》	2009 年发布	通用法律
7	《加州安全违约告知法律》	2002 年生效	州法律
8	《金融服务现代化法案》（GLBA）	1999 年发布	部门规章
9	《健康保险携带和责任法案》（HIPAA）	1996 年发布	部门规章
10	《联邦贸易委员会法案》（FTCAct）	1914 年发布	部门规章
二、欧盟			
1	《通用数据保护规则》（GDPR）	2016 年发布	通用法律
2	《欧盟数据留存指令》	2006 年发布	通用法律
3	《隐私与电子通讯指令》	2002 年发布	通用法律
4	《欧盟数据保护指令》	1995 年发布	通用法律
三、澳大利亚			
1	《电信法案》	1997 年发布	部门规章
2	《联邦隐私法案》	1988 年发布	通用法律
四、俄罗斯			
1	俄罗斯联邦法律第 152-FZ 条中 2006 年个人数据相关内容（PersonalDataProtectionAct，个人数据保护法案）	2015 年发布	通用法律
2	俄罗斯联邦法律第 149－FZ 条 2006 年信息、信息技术和数据保护相关内容（DataProtectionAct，数据保护法案）	2006 年发布	通用法律
3	《斯特拉斯堡公约》	2005 年发布	通用法律
五、新加坡			
1	《个人数据保护法令》(PDPA)	2012 年发布	通用法律

从一定意义上说，各国数据保护法律法规的宗旨就是围绕数据提供者、数据基础设施提供者、数据服务提供者、数据消费者、数据监管者等参与方，力图将数据保护范围、各参与方对应的权利和义务、相关行为准则等要点界定清晰。

4.1.1.2 数据保护范围

在法律法规层面上，数据保护是有范围的，要针对可监管的辖区范围、需保护的数据对象、需监管的数据应用场景，以及需监管的数据处理行为等明确数据保护范围。数据保护范围一般在数据保护相关法律法规中都明确界定，并通过各种配套标准加以细化，以支撑法律法规的落地。

一、可监管的辖区范围

可监管的辖区范围是指法律法规里规定的所能管辖的数据涉及的领土范围，尤其是设立在境外的数据中心是否受到本国法律法规的监管，这也是目前业界关注的重点之一。不同国家和地区对此规定有一定的差异性。如美国、澳大利亚目前的管辖范围是本国领土，也就是说，外国企业以及本国企业设在境外的数据中心，均不受本国法律法规约束（但某些特殊情况下，美国有可能会运用长臂管辖权等特殊原则，以国家安全的名义，在认为必要的时候实施强制管辖）。但欧盟、俄罗斯、新加坡等国家和地区则相对监管较严，如俄罗斯规定其数据保护法律法规不受领土管辖权的限制，适用于任何在俄罗斯发生的所有数据处理过程，包括所有对俄罗斯公民数据的收集和使用，而无论数据中心是否建立或位于俄罗斯境内。对于跨境的数据流，如果俄罗斯公民是对应的数据传输协定中的一方，那么俄罗斯的数据保护法律法规也可在一定程度上应用。

二、需保护的数据对象

目前，美国、欧盟、俄罗斯等国的数据保护主要针对个人信息，一般说来可划分为两类：个人识别信息（PII，Personal Identity Information）和个人隐私/敏感数据。其中，PII 是指能直接根据该信息识别和定位到个人的信息，如姓名、身份证号码、银行卡号、家庭住址等；个人隐私/敏感数据是指虽不能直接识别和定位到个人，但通过关联和综合分析，有可能定位到个人的信息，如健康信息、教育经历、征信记录等。各国对个人隐私/敏感数据的定义不同，其保护的数据范围也就各不相同，如美国在一些部门规章（如 HIPAA）中划定了个人隐私保护的具体范围，而俄罗斯、新加坡等国则规定凡是和个人相关的信息，均被认为是个人隐私/敏感数据，均在保护范围内。

此外，在这两类需要监管的数据中，也有因例外豁免条款成为不需监管的数据，如新加坡规定商务联系信息、已存在了 100 年的个人资料以及已经死去超过十年的个人数据等均不在保护范围内。

三、需监管的数据应用场景

一般情况下，所有涉及数据收集、存储、处理、利用的数据控制者都是被监管的对象，但各国也根据自己国情划定了可免除监管的例外条例，如新加坡规定了公民个人行为、员工就业过程中的必要行为、政府/新闻/科研等公共机构的部分行为、某些获取了明确证明或书面合同的数据中介机构等，可免于数据保护法律法规的监管。

四、需监管的数据处理行为

目前，美国、欧盟、俄罗斯、新加坡等国均提出应对数据的全生命周期进行监管，包括收集、记录、组织、积累、存储、变更(更新、修改)、检索、恢复、使用、转让(传播，提供接入等)、脱敏、删除、销毁等行为，但各国也根据自己国情划定了可免除监管的例外条例，如俄罗斯规定了专为个人和家庭需求处理个人数据(前提是不侵犯数据对象的权利)、处理国家保密数据、依照有关法律由主管当局向俄罗斯法院提供相关数据等情况，则属于相应的例外豁免情形。

4.1.1.3 监管部门及其权力

为保证数据安全法律法规的落实，监管部门需设立相应的机构和人员，并赋予相应的权力，如执法权、处罚权等。

一、美国

美国联邦贸易委员会（FTC）是美国国家隐私法律的主要执行者。虽然其他机构（如银行机构）也被授权执行各种隐私法，但 FTC 采取的措施相对更加强势。例如，FTC 可以发起调查、停止令，甚至在法庭上提出申诉。此外，FTC 还向国会报告隐私问题，并制定隐私立法所需的建议。相比而言，《金融服务现代化法案 FSMC》则由 FTC、联邦银行监管机构和国家保险机构共同执行（在执行过程中，FTC 比其他两家机构更为积极）。HIPAA 则由健康与人类服务部（HHS）公民权利办公室（The Office of Civil Rights）执行。该办公室可对下属机构的信息处理实践活动发起调查，确认其是否符合 HIPAA 隐私规则，并允许个人对侵犯隐私的行为进行投诉。在加州，《加州安全违约告知法律》和《加州在线隐私保护法案》则由加州总检察长和地方检察官执行。

二、欧盟

2016 年 4 月 14 日，欧洲议会投票通过了商讨四年的《一般数据保护法案》

（General Data Protection Regulation，GDPR），该法案将于 2018 年 5 月 25 日正式生效。GDPR 的通过意味着欧盟对个人信息保护及其监管达到了前所未有的高度，堪称史上最严格的数据保护法案。GDPR 对于业务范围涉及欧盟成员国领土及其公民的企业都具有约束力，通过设立欧盟数据保护理事会（European Data Protection Board），赋予其欧盟数据监管的最高机构的地位，并保证其独立性。理事会可以单独行动，直接对欧盟委员会负责。

三、澳大利亚

澳大利亚成立了专门机构——澳大利亚信息专员办公室（Office of the Australian Information Commissioner，OAIC），并设立了信息专员作为执行《联邦隐私法案》的关键角色。OAIC 有权接受并处理个人对于隐私的相关控诉。如果相关控诉属实，则 OAIC 可做以下处理：1）通知被告不能重复或继续侵害隐私；2）做出相应的决定，如申诉人有权指定赔偿方式和数额；3）OAIC 和申诉人有权向法院提起诉讼决定，违反隐私法案的个人和公司，可能分别面临高达 34 万澳元和 170 万澳元的处罚。

信息专员的权力范围包括调查违规、促进合法行为等，如审计实体的合法性、接受书面保证（并推进执行）、注册有约束力的业务法规、决定是否开展自愿调查。同时，针对数据主体的投诉，信息专员可通过调解去解决双方争端，或根据投诉做出决定。此外，信息专员还可开展下列流程：1）执行承诺和决策；2）寻求强制救济；3）申请民事处罚。对于严重的或反复违反《隐私法案》的行为，法庭可进行罚款。

此外，除了 OAIC 及其信息专员，澳大利亚通信及媒体管理局 ACMA（《垃圾邮件法案》的监管者）也有独立的执法权。

四、俄罗斯

俄罗斯数据保护最主要的监管部门是俄罗斯电信/信息技术和大众传媒联邦监管局（Roskomnadzor，相当于美国 FCC）。此外俄罗斯政府、俄罗斯联邦技术和出口服务局（FSTEC）以及俄罗斯联邦安全局（FSS）等主管监管部门也制定了一些对数据保护的特定条款。

五、新加坡

新加坡数据保护最主要的法律依据是《个人数据保护法令》（PDPA），同时为了执行 PDPA，新加坡专门成立个人数据保护委员会（PDPC）来承担 PDPA 的制定和实施工作。与俄罗斯的 Roskomnadzor 类似，新加坡的 PDPC 也具有一定的执法权。

4.1.1.4 数据提供者的权利

数据提供者的权利主要指用户在使用信息服务过程中被收集相关信息，对自身信息所拥有的权利，包括知情权、授权处理权、访问/查询/更正权、停止收集/删除权、投诉权等，具体见表 4.2。

表 4.2 各国对数据提供者权利的规定

权利内容	美国	欧盟	澳大利亚	俄罗斯	新加坡
数据收集/处理前被告知的权利	允许[1]	允许[1]	允许[1]	允许[1]	允许[1]
授权个人数据收集/处理的权利	允许	允许	允许	允许	允许
访问/查询个人信息的权利	指定情况下允许[2]	允许	允许	允许	允许
更正个人信息的权利	指定情况	允许	允许	允许	允许

	下允许[2]				
停止收集个人信息的权利	指定情况下允许[3]	允许	未明确规定	允许	未明确规定
删除个人信息的权利	指定情况下允许[4]	允许	不允许	指定情况下允许[4]	不允许[4]
投诉的权利	未明确规定	允许	允许	未明确规定	允许
其他权利	未明确规定	允许[5]	未明确规定	允许[5]	未明确规定
注 1：美国、欧盟、澳大利亚、俄罗斯、新加坡对收集数据前征求数据提供者同意的形式、提供的内容、例外情况等规定各不相同。 注 2：除了 HIPAA、加州法律等，FTC、GLBA 等大多数美国隐私法一般不支持为用户提供相关访问权限,但《儿童在线隐私保护法案》允许父母查看网站所收集的孩子个人信息。此外，HIPAA、《儿童在线隐私保护法案》还支持用户的删除/更正信息的要求。 注 3：美国 GLBA、HIPAA、加州法律都支持公司提供渠道，允许用户退出其提供的信息服务。 注 4：美国《儿童在线隐私保护法案》允许父母删除信息的要求。俄罗斯个人数据保护法案规定，当个人数据不完整、过期、不准确、非法取得、数据处理声明的目的不是必须的等情况下，数据提供者可以请求删除个人数据。新加坡 PDPA 不给个人请求删除自己个人信息的权利，但保留有限制的责任。 注 5：欧盟 GDPR 在第 17 章中明确定义数据遗忘权和删除权。俄罗斯个人数据保护法案规定的数据主体其他权利还包括获取数据使用相关信息、反对直接营销等权利。					

可以看到，各国数据保护法律法规对数据提供者权利基本都进行了规定，但规定的粒度又各不相同，如针对知情权，俄罗斯则规定，针对数据主体全名/地址/身份证明 ID(如护照)/身份证明的发行时间与发证机关/签名等信息、数据控制者的全名/地址/数据处理目的等关键信息，需要以书面形式给出（包括电子签名的方式）。新加坡则规定，数据控制者在收集个人数据之前需经数据主体的同意，但不指定通知形式。

4.1.1.5 数据使用者的义务

美国、欧盟、澳大利亚、俄罗斯、新加坡等国均规定数据使用者除了有义务在数据收集、存储、处理等全生命周期中配合数据主体实现其权利外，还有确保数据安全、对数据监管者进行数据收集和利用情况报备、发生异常事件时的通报、数据境外流转/存储前向数据监管者申请等义务，如表 4.3 所示。

表 4.3 各国对数据使用者义务的规定

义务内容	美国	欧盟	澳大利亚	俄罗斯	新加坡
保证数据中心在境内的义务	指定情况下不要求	指定情况下不要求	指定情况下不要求	要求数据中心在境内	未明确规定
收集数据前征求政府部门同意的义务	未明确规定	未明确规定	明确规定[1]	明确规定[2]	明确规定[3]
发生异常时向指定政府部门及数据主体等报告的义务	明确规定	明确规定	明确规定	明确规定	未明确规定
数据境外流转/存储前向数据监管者申请的义务	见表 4.4	见表 4.4	见表 4.4	见表 4.4	见表 4.4
注 1：澳大利亚没有要求必须通知 OAIC 和信息专员。 注 2：俄罗斯规定数据控制者在操作处理个人数据之前必须通知 Roskomnadzor（通知可以是纸质版或电子版），Roskomnadzor 会在收到通知后 30 天内对数据操作者进行登记。 注 3：新加坡规定有义务通知政府主管部门收集、使用或披露个人信息的目的（包括书面和口头两种形式）。					

4.1.1.6 数据境外流转/存储和转移协议要求

数据跨境主要包括数据跨境流动、数据跨境存储以及所涉及的跨境协议等方面内容，各国对数据境外流转/存储和转移协议的规定如下表 4.4 所示。

表 4.4 各国数据境外流转/存储和转移协议要求

数据境外流转/存储和转移协议要求	美国	欧盟	澳大利亚	俄罗斯	新加坡
数据是否可存储在境外	允许	允许	允许	不允许	允许
数据是否可流转到境外	在指定条件下流转	在指定条件下流转	在指定条件下流转	在指定条件下流转	在指定条件下流转
数据流转到境外的转移协议	未明确规定	未明确规定	未明确规定	未明确规定	未明确规定
发生数据境外存储/流转时是否告知数据监管者	未明确规定	未明确规定	未明确规定	必须告知主管部门	未明确规定

一、美国

美国对个人数据跨境传输限制较少，只有部分州颁布相关法律限制国外组织或机构开展数据服务,但通常仅限于为政府机构提供服务或产品的企业。

FTC 和其他监管机构的立场是，美国法律法规适用于跨境传输的美国数据，监管企业如下方面:

1. 数据出口到美国国外；
2. 海外分包商处理的数据；
3. 分包商使用相同的保护措施(如通过使用安全保障协议,审核和合同规定)监管跨境后的数据。

二、欧盟

GDPR 规定：数据接收国的法律、监管机构能够有效地保护欧盟数据主体的权利，并且有充分的司法救济权力；数据接收国必须是欧盟认可的数据保护充分的国家。

三、澳大利亚

根据《联邦隐私法案》，组织或机构将持有的个人信息披露给位于澳大利亚以外的第三方之前，必须采取合理的步骤以确保海外的数据接收者不会违反法案中规定的原则（除了特殊情况，组织或机构将持有的个人信息披露给位于澳大利亚以外的第三方之前，必须采取合理的步骤以确保海外的数据接收者不会违反《联邦隐私法案》的要求。在一定程度上，甚至认为组织或机构须对任何境外接收者违反《联邦隐私法案》的行为负责）。但在下列情况下不适用：

1. 组织或机构已按照《联邦隐私法案》规定的方式获得了相关数据主体的同意；
2. 境外接收者是由和《联邦隐私法案》类似，且可被《联邦隐私法案》强制执行的境外法律所约束；
3. 其他例外情况（如个人数据披露是由澳大利亚相关法律要求的）。

通常获得知情同意很困难，因此大多数情况下境外接收者不受类似的由相关数据主体强制执行的境外法律的约束。因此，在大多数情况下实体必须采用“合理步骤”来保证境外接收者不会违反《联邦隐私法案》的优先级高低出境数据存在以下情况之一的将信息披露给境外接收者。组织或机构通常寻求获得境外收件人的合同承诺即按照澳大利亚隐私法处理个人数据来遵守此规定（通常组织或机构通过和境外接收者签署合同来符合相关要求，这些合同可以保证

在《联邦隐私法案》的框架下处理所有用户数据。

在某些情况下，将会被追究法律责任，如组织或机构披露个人数据到境外而该境外接收者又违反了《联邦隐私法案》。

四、俄罗斯

俄罗斯个人数据保护相关法律法规指出：在个人数据跨境流动的情况下，所有的数据控制者必须确保（做出转让之前）其各自的数据主体的权利和利益在相应的其他国家能以适当的方式得到充分的保障。所有签署斯特拉斯堡公约的国家都被认为是能为数据主体提供权利和利益“充分保护”的司法管辖区。只要征得数据主体同意，数据跨境流动到具有对等保护级别的司法管辖区是不受任何限制的。同时，Roskomnadzor 已列出能够对个人数据跨境流动提供足够级别保护的国家正式名单，包括澳大利亚、阿根廷、加拿大、以色列、墨西哥和新西兰。只有在特定例外情况下才允许个人资料跨境流动到无足够保护水平的国家。

通常情况下，作为数据控制者的公司会在向境外传输个人数据前检查对方是否具有足够的数据保护水平。此外，公司将获得数据主体各自给出的书面同意，或执行遵循自身目的的跨境数据传输协议。按照这些步骤，公司将按照自己公司内部的规则或政策开展数据跨境传输。

2014 年 7 月 21 日，俄罗斯联邦总统签署第 242-FZ 号联邦法律，将修订在信息和电信网络方面对于个人数据处理的某些立法行为，成为新数据保护法。新数据保护法的修订已于 2014 年 12 月 31 日在第 526-FZ 号联邦法律通过，并于 2015 年 9 月 1 日生效。新的数据保护法主要修正了个人数据保护法的两个问题：

1. 介绍了数据控制者关于俄罗斯公民个人数据的收集、存储和处理方面的新义务；
2. 引入了 Roskomnadzor 阻止非法加工俄罗斯公民个人数据的网站和在线资源的新机制。

具体而言，新数据保护法对所有数据控制者都引入了一项义务，要求其“在收集个人相关数据时，包括通过互联网的过程中，确保能够通过位于俄罗斯联邦境内的数据中心记录、系统化、累积、存储、变更以及提取俄罗斯公民个人数据。”这意味着，任何由数据控制者收集的俄罗斯公民个人数据都需要存储在俄罗斯境内的服务器、IT 系统或数据中心上。

虽然新的数据保护法律没有明确规定这一点，但这些要求很可能被解释为禁止将俄罗斯公民个人数据存储在俄罗斯境外。因此，根据新数据保护法的文字描述和解释，为满足数据保护立法的所有其他通用要求，本地和外国公司（数据控制者）都需要在俄罗斯境内处理或组织处理俄罗斯公民个人数据。

此外，新的数据保护法律并没有禁止从境外访问位于俄罗斯境内的服务器、IT 系统和数据中心，也没有对包括境外数据传输和数据复制在内的数据转移进行特别的限制。

俄罗斯法律法规并没有规定数据跨境流动所采用的传输协议，虽然这些跨境传输协议应用广泛，但 Roskomnadzor 没有给出其标准形式，所用到的每一个协议都是视情况而定的。

只要收到数据主体的同意，或数据主体的同意在协议里有表达，则个人数据的跨境流转就是合法的。此外，数据控制者在向 Roskomnadzor 登记时，必须告知 Roskomnadzor 其数据跨境流动的权利。

五、新加坡

原则上组织不能把任何个人资料转移到新加坡以外的国家或地区，除非其符合 PDPA 第 26 节规定的要求，这是为了保证组织能够提供与 PDPA 保护框架相对应的保护水平。只有当组织在采取合理步骤，满足特定条件时，才可将个人数据传输到境外。

4.1.2 国内数据安全法律法规和政策

我国在推进大数据产业发展的过程中，越来越重视数据安全问题，不断完善数据开放共享、数据跨境流动和用户个人信息保护等方面的法律法规和政策，为大数据产业健康发展保驾护航。

4.1.2.1 数据开放共享相关法律法规政策

近年来，中央和地方政府高度重视数据开放共享工作，相继出台数据开放共享相关政策法规。

一、国家层面，制定数据开放共享相关的纲要规划，加强数据开放共享的顶层设计。

2015 年 8 月，国务院印发《促进大数据发展行动纲要》，提出加快政府数据开放共享，推动资源整合，提升治理能力。要大力推动政府部门数据共享，明确各部门数据共享的范围边界和使用方式，厘清数据管理及共享的义务和权利，依托政府数据统一共享交换平台，大力推进国家基础数据资源共享；稳步推动公共数据资源开放，建立公共机构数据资源清单，建设国家政府数据统一开放平台，推进公共机构数据资源统一汇聚和集中向社会开放，提升政府数据开放共享标准化程度；建立政府和社会互动的大数据采集形成机制，制定政府数据共享开放目录，通过政务数据公开共享，引导企业、行业协会、科研机构、公共组织等主动采集并开放数据。

2016 年 12 月，工信部印发《大数据产业发展规划（2016－2020 年）》，提出加强资源共享和沟通协作，协调制定政策措施和行动计划，解决大数据产业发展过程中的重大问题，建立大数据发展部省协调机制，加强地方与中央大数据产业相关规划、法律、政策等的衔接配套，通过联合开展产业规划等措施促进区域间大数据政策协调；推动制定公共信息资源保护和开放的制度性文件，以及政府信息资源管理办法，逐步扩大开放数据的范围，提高开放数据质量。

2018 年 3 月，国务院办公厅发布《科学数据管理办法》，深刻把握大数据时代科学数据发展趋势，充分借鉴国内外先进经验和成熟做法，针对目前我国科学数据管理中存在的薄弱环节，进行了系统的部署和安排，围绕科学数据的全生命周期，加强和规范科学数据的采集生产、加工整理、开放共享等各个环节的工作。把确保数据安全放在首要位置，建立数据共享和对外交流的安全审查机制；按照“开放为常态、不开放为例外”的共享理念，明确为公益事业无偿服务的政策导向，充分发挥科学数据的重要作用。

二、地方层面，地方政府也积极配合推动大数据产业发展应用，加快制定数据开放共享方面的法规政策。

2014 年，贵州省发布《关于加快大数据产业发展应用若干政策的意见》（以下简称《意见》）和《贵州省大数据产业发展应用规划纲要（2014-2020）》（以下简称《纲要》），提出数据开放共享方面需加快建立相关标准规范，实现

有效整合数据资源的目标；制定大数据采集、管理、共享、交易等标准规范，明确收集数据的范围和格式、数据管理的权限和程序以及开放数据的内容、格式和访问方式等。《意见》还指出建立政府和社会互动的大数据采集形成机制，通过政务数据公开共享，引导企业、行业协会、科研机构、公共组织等主动采集并开放数据，形成数据的大开放。《纲要》也指出制定出台数据资源开放指导办法和数据资源安全开放标准规范，按照“开放优先、安全例外、分类分级”的原则，对大数据中心的数据资源进行梳理和开放风险评估，制定数据开放目录并及时更新。

2016 年，浙江省发布《浙江省促进大数据发展实施计划》（以下简称《浙江计划》），提出建立数据开放共享相关标准规范体系。《浙江计划》提出打造数据共享、交换和开放统一平台，要推进政府数据资源共享交换，建设统一的政府信息资源管理服务系统，厘清数据管理及共享的义务和权利，明确共享的范围边界和使用方式；深化公共数据统一开放平台建设，建立公共数据资源开放目录，制定数据开放标准，落实数据开放和维护责任，推动相关领域的政府数据向社会开放。《浙江计划》还提出研究制定数据采集标准及分级分类标准、政府数据共享标准、数据交换标准、政府及公共数据开放标准、统计标准，对共享开放的方式、内容、对象、条件等进行规范；积极参与大数据关键共性技术国际标准、国家标准、行业标准的制修定，推进大数据产业标准体系建设，探索建立大数据市场交易标准体系。

2016 年，广东省发布《广东省促进大数据发展行动计划（2016-2020 年）》（以下简称《广东计划》），对数据开放共享做出相关规定。《广东计划》提出将推动政府数据资源整合、共享和开放作为重点行动，完善大数据采集机制，整合公共数据资源，推动政府数据共享，推动公共数据资源开放。完善政务信息资源共享平台，健全政府信息资源共享机制，建设政府数据统一开放平台，提供面向公众的政府数据服务，推动公共数据资源向社会开放。《广东计划》还提出加快建立公共机构的数据标准和统计标准体系，推进大数据采集、管理、共享、交易等标准规范的制定和实施。统一政务数据编码、格式标准、交换接口规范，研究制定一批基础共性、重点应用和关键技术标准。

2017 年 5 月，《贵阳市政府数据共享开放条例》正式实施，这标志着贵阳填补了在大数据方面的地方性法规空白，是贵阳市以大数据为引领加快打造创新型中心城市的重大举措，对于推进数据资源开放共享有法可依，提供了理论依据和法律支撑，为贵阳大数据产业发展法治构建打下了坚实的基础。

4.1.2.2 数据跨境相关法律法规和政策

随着全球数字经济快速发展，数据跨境流动日趋频繁，数据跨境传输引发的国家重要数据资源安全风险也与日俱增。为了加强数据跨境安全保护，我国在《网络安全法》中首次明确了关键信息基础设施有关个人信息和重要数据本地存储和向境外提供的规定。此外，国家网信部门正在依据《网络安全法》加紧制定《个人信息和重要数据出境安全评估办法》（现已形成征求意见稿），提出网络运营者在我国境内运营中收集和产生的个人信息和重要数据，因业务需要向境外提供的，应进行安全评估。

《个人信息和重要数据出境安全评估办法（征求意见稿）》（以下简称《评估办法》）明确规定，出境数据存在以下情况之一的，要经过安全评估：含有或累计含有 50 万人以上的个人信息；包含核设施、化学生物、国防军工、人口健

康等领域数据，大型工程活动、海洋环境以及敏感地理信息数据等；包含关键信息基础设施的系统漏洞、安全防护等网络安全信息；关键信息基础设施运营者向境外提供个人信息和重要数据等。评估重点包括：数据出境的必要性；目的地是否有能力及网络安全保护水平能否确保数据安全；可能对国家安全、社会公共利益、个人合法利益带来的风险等。国家网信部门统筹协调数据出境安全评估工作，指导行业主管或监管部门组织开展数据出境安全评估。《评估办法》中还规定，个人信息出境，应向个人信息主体说明，并经其同意。未成年人个人信息出境须经其监护人同意。可能影响国家安全、损害社会公共利益的以及其他经国家网信部门、公安部门、安全部门等认定不能出境的数据不得出境。

在《网络安全法》发布前，我国已经在金融、卫生医疗、交通、地理、电子商务、征信等行业制定了有关数据跨境的法律法规和政策要求。已有的数据跨境相关政策要求集中于数据本地化存储，按照管理方法可以分为两类，一类是限制出境，一类是禁止出境。

一、限制出境

我国在征信、云计算、电子商务等行业采取限制出境的管理方式。

2012 年 12 月，国务院第 228 次常务会议通过《征信业管理条例》（以下简称《条例》）。《条例》明确要求征信机构在中国境内采集的信息的整理、保存和加工，需在中国境内进行。若征信机构确因业务需要向境外组织或者个人提供信息，应当遵守法律、行政法规和国务院征信业监管部门的有关规定。

2016 年 11 月，工业和信息化部发布了《关于规范云服务市场经营行为的通知（公开征求意见稿）》（以下简称《云服务通知》），对数据出境做出有关规定。《云服务通知》明确指出，面向境内用户提供服务，应将服务设施和网络数据存放于境内，跨境实施运维及数据流动应符合国家有关规定。

2016 年 12 月，十二届全国人大常委会第二十五次会议初次审议了《中华人民共和国电子商务法（草案）》，为电子商务数据出境提供法律依据。该法案明确指出电子商务经营主体从事跨境电子商务活动，应当依法保护交易中获得的个人信息和商业数据。国家建立跨境电子商务交易数据的存储、交换和保护机制，努力做好数据出境安全保障。

二、禁止出境

我国在金融、保险、卫生医疗、交通、气象、新闻出版等行业采用禁止出境的管理方式。

2011 年 1 月，中国人民银行印发《关于银行业金融机构做好个人金融信息保护工作的通知》（以下简称《金融通知》），将保护个人金融信息定义为一项法定义务，要求做好金融领域个人信息保护工作。《金融通知》规定，在中华人民共和国境内收集的个人金融信息的存储、处理和分析应当在境内进行。除法律法规及中国人民银行另有规定外，银行业金融机构不得向境外提供境内个人金融信息。

2011 年 3 月，中国保险监督管理委员会印发《保险公司开业验收指引》，要求原则上业务数据、财务数据等重要数据应在中国境内存储，且具有独立的数据存储设备及相应的安全防护和异地备份措施。2015 年发布《保险机构信息化监管规定（征求意见稿）》，规定数据来源于中华人民共和国境内的，数据中心的物理位置应当位于境内。外资保险机构信息系统所载数据移至中华人民共和国境外的，应当符合我国有关法律法规。

2014 年 5 月，国家卫生计生委印发《人口健康信息管理办法（试行）》（以

下简称《健康办法》)。《健康办法》规定不得将人口健康信息在境外的服务器中存储，不得托管、租赁在境外的服务器，明确禁止了有关我国人口健康信息的境外存储。

2015 年 11 月，国务院第 111 次常务会议通过《地图管理条例》，要求互联网地图服务单位应当将存放地图数据的服务器设在中华人民共和国境内，并要求其制定互联网地图数据安全管理制度和保障措施。

2016 年 7 月，交通运输部、工业与信息化部等七部委联合发布《网络预约出租汽车经营服务管理暂行办法》（以下简称《网约车办法》），严格规范网约车平台的经营行为。《网约车办法》明确要求平台在网络安全与信息安全方面遵守国家有关规定，在提供服务的过程中采集的个人信息和生成的业务数据，应当在中国内地存储和使用，保存期限不少于 2 年；除法律法规另有规定外，个人信息与业务数据不得外流。

2008 年 6 月，中国气象局发布的《气象资料共享管理办法》（第 4 号令）规定，用户不得直接将其从各级气象主管机构获得的气象资料，用作向外分发或供外部使用的数据库、产品和服务的一部分，也不得间接用作生成它们的基础。

2016 年 2 月，国家新闻出版广电总局、工业和信息化部公布《网络出版服务管理规定》（第 5 号令），明确要求图书、音像、电子、报纸、期刊出版单位必须将从事网络出版服务所需的必要的技术设备、相关服务器和存储设备存放在中华人民共和国境内。

4.1.2.3 个人信息保护相关法律法规与政策

目前，我国尚未出台专门的个人信息保护法，个人信息保护相关规定散见于多个法律法规和规章制度之中。

4.1.2.3.1 法律层面

我国正逐步加强公民个人信息保护方面的顶层立法工作，陆续在《网络安全法》、《刑法》和《民法》等基本法中加入个人信息保护的内容，不断完善个人信息保护法律体系。

一、《网络安全法》关于个人信息保护

《网络安全法》第四十条至第四十五条，对个人信息保护做出有关规定，明确了我国个人信息保护的基本原则和框架。第四十条是对网络运营者保护用户信息义务的原则规定，要求网络运营者对其收集的用户信息严格保密，建立健全用户信息保护制度。第四十一条对网络运营者收集、使用个人信息应遵守的规则进行了规定，这些规定与国际通行规则是一致的。第四十二条是关于个人信息安全原则、个人信息匿名化处理和个人信息泄露报告义务的规定，首次明确提出建立数据泄露通知报告机制。第四十三条是关于个人信息删除权和更正权的规定，信息主体在具备法定理由的情形下，拥有请求删除其个人信息的权利；在个人信息不完整或不准确时，拥有要求及时改正、补充的权利。第四十四条是关于禁止非法获取、非法出售、非法提供个人信息的规定。第四十五条是关于负有网络安全监督管理职责的部门及其工作人员的保密义务的规定。

二、《刑法》关于个人信息保护

我国正逐步加大威胁个人信息安全行为的刑事罪责，从法律的强制性上加强个人信息保护。

2009 年 2 月 28 日，十一届全国人民代表大会常务委员会第七次会议通过《中华人民共和国刑法修正案（七）》，在刑法第二百五十三条后增加一条，作

为第二百五十三条之一：“国家机关或者金融、电信、交通、教育、医疗等单位的工作人员，违反国家规定，将本单位在履行职责或者提供服务过程中获得的公民个人信息，出售或者非法提供给他人，情节严重的，处三年以下有期徒刑或者拘役，并处或者单处罚金。《刑法修正案（七）》的出台具有重大意义，我国第一次将个人信息保护写入刑法，规定了国家机关与金融、电信等领域工作人员出售或非法提供个人信息的法律后果。但《刑法修正案（七）》有关规定也存在不足，没能有效覆盖犯罪主体，在实际操作中存在犯罪行为认定困难等问题。

2015 年 8 月 29 日十二届全国人民代表大会常务委员会第十六次会议通过《刑法修正案（九）》，对二百五十三之一做出修改，将“出售、非法提供公民个人信息罪”和“非法获取公民个人信息罪”整合为“侵犯公民个人信息罪”。《刑法修正案（九）》与《刑法修正案（七）》相比，从三个方面完善了对公民个人信息保护的规定：一是扩大犯罪主体的范围，规定任何单位、组织、个人违反有关规定，出售或向他人提供公民个人信息，情节严重的，都将构成犯罪；二是严打“内部人”犯罪，明确规定任何单位、组织、个人违反国家有关规定，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，从重处罚；三是提高量刑标准，规定侵犯公民信息罪情节特别严重的，处三年以上七年以下有期徒刑，并处罚金。《刑法修正案（九）》施行以来，公检法机关办理的侵犯公民个人信息案件显著增加，一定程度上遏制了侵犯公民个人信息行为快速增长的趋势。

《刑法修正案（九）》施行一段时间后，相关人员在司法实践中发现侵犯公民个人信息罪的定罪量刑标准较为原则，不易把握；另有一些法律适用问题存在认识分歧，影响案件办理。为保障法律正确、统一适用，2017 年 5 月 9 日，最高人民法院会同最高人民检察院联合发布《最高人民法院、最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释》（《刑九解释》）。《刑九解释》在《刑法修正案（九）》基础上列出了十三条具体的司法解释，明确了“公民个人信息的范围”包括身份识别信息和活动情况信息，细化了非法获取、提供公民个人信息的认定标准，对侵犯公民个人信息犯罪的定罪量刑标准和有关法律适用问题作了全面、系统的规定，为司法实践中开展公民个人信息保护提供了强有力的支撑。

三、《民法》关于个人信息保护

为了进一步保障公民的个人信息安全，我国将个人信息保护的内容纳入民法中，强化了公民个人信息民事权益保护。

2017 年 3 月 15 日，十二届全国人大五次会议表决通过了《中华人民共和国民法总则》，并于 2017 年 10 月 1 日起施行。《民法总则》规定自然人的个人信息受法律保护，任何组织和个人需要获取他人信息的，应当依法取得并确保信息安全，不得非法收集、使用、加工、传输他人个人信息，不得非法买卖、提供或者公开他人个人信息。

个人信息权被纳入《民法总则》具有重大意义，表明个人信息权利拥有了基本民事权利的地位。今后除了严重侵犯公民人身权利、财产权利的重大违法犯罪行为应当依照《刑法》承担刑事责任（可以附带提起民事诉讼）外，对于一般的侵害个人信息权的侵权行为，任何自然人或组织均可以从侵权法的角度进行维权，以个人信息权被侵犯为由提起民事诉讼。

4.1.2.3.2 具体行业和领域的法律法规和部门规章

除国家层面加快个人信息保护相关立法进程外，我国各行业和领域也开始高度重视个人信息保护工作，出台专门的个人信息保护行业法律法规或部门规章，或是将个人信息保护有关内容写入相关法律法规中，进一步完善了个人信息保护法律体系。

一、电信和互联网行业

为了进一步完善电信和互联网行业个人信息保护法律体系，保护电信和互联网用户的合法权益，维护网络与信息安全，工业和信息化部于 2013 年出台了《电信和互联网用户个人信息保护规定》(《电信规定》)。《电信规定》明确了电信业务经营者和互联网信息服务提供者在提供服务的过程中收集的个人信息范畴，即能够单独或者与其他信息结合识别用户的信息。进一步明确了电信业务经营者、互联网信息服务提供者收集、使用用户个人信息的规则和信息安全保障措施要求。同时对电信管理机构实施监督检查和违反个人信息保护的行为应当承担的法律责任进行规定和说明。

二、消费者权益保护

《消费权益保护法》修订案增加了个人信息保护相关内容，中国消费者的个人信息受保护权益正式被确认。2013 年 10 月 25 日，第十二届全国人民代表大会常务委员会第五次会议修正通过了新版《消费者权益保护法》，并于 2014 年 3 月 15 日开始正式实施。第二十九条对个人信息保护作了明确规定：“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。”

2016 年 11 月，国家工商总局公布《消费者权益保护实施条例（征求意见稿)》。第二十二条规定：“经营者收集、使用消费者个人信息应当遵循合法、必要、正当的原则，明示收集、使用信息的目的、方式和范围并征得消费者同意，经营者不得收集与经营业务无关的信息或者采取不正当方式收集信息。消费者明确要求经营者删除、修改其个人信息的，除法律法规另有规定外，经营者应当按照消费者的要求予以删除、修改。”《条例》对实施消费者个人信息保护做出明确规定，成为《消费者权益保护法》的护航者，对保障个人信息安全起到重要作用。

4.1.2.4 存在的问题

在大数据应用快速发展的推动下，数据保护越来越成为各国关注的焦点，但数据保护的范围和力度则在一定程度上取决于各国信息产业的发展现状以及未来发展方向。我国在制定数据安全相关法律法规过程中，应结合国家大数据产业发展纲要，与各主管监管部门、研究机构、研发应用企业等产业链各环节充分互动，建立既能维护国家安全和公众利益、又能有效促进产业发展的数据安全法律法规和标准体系，从而推动我国大数据产业健康发展。

目前我国法律法规的核心是监管和规范企业或个人对于数据的相关行为，以防止滥用数据，以及监管跨境数据转移等内容，重点保障国家、企业、个人等利益。我国虽然已经出台了以《网络安全法》为代表的相关法律法规，但我国在数据领域的建规立法还存在以下不足：

一、我国数据安全法律法规的制定仍需进一步完善

以个人数据保护方面为例，虽然对于个人数据保护及相关基本原则进行了

规定，尤其是对于使用的正当和必要程序等方面进行了规定，但对于不同类别的个人数据还应该对其保护水平和要求进行细分，还需要进一步细化使用规范。参照国外相关法律内容，涉及到基因数据、医疗数据等方面需要提供更为严苛的保护规定，对于从事此类数据的企业的数据保护能力及水平提出更高的要求。联合国《人类基因数据国际宣言》提出基因数据应当以人权高度进行特殊保护。欧盟规定对个人重要数据需要采取更多的技术防护措施。而目前我国《网络安全法》还是以网络、系统及数据为重点，今后还需要对更大范围的数据保护立法进行补充和优化。

二、我国缺乏数据安全相关的长期执法实践

虽然有了相关的法律条款作为数据安全执法依据，但数据保护的执法落地需要进一步强化，一方面亟待配套建立相关的数据安全标准及指南等，规范当前大数据应用的新业态健康发展。例如，许多城市都已开展了健康医疗大数据、消费大数据等方面的研究和产业发展，其中个人信息如何通过具体执法得到切实保护，仍然存在着执法成本、执法效果及执法环境等方面问题，需要进一步研究；另一方面，需要不断借鉴国外法律实践的经验，提高企业、个人等数据所有者的主动保护意识，鼓励自愿或无偿共享网络安全威胁信息，出台相应的鼓励措施，推动全社会对于数据安全保护的法律意识进一步提高。

三、国内立法与国外法律的对接尚需实践

我国立法是从国内当前的情况及法律需求角度出发，但涉及国内外数据跨境交换时，就需要考虑国内法律与现有国外相关法律条款是否能够对接，国外现行的法律不一定符合我国的国情。因此，我国应进一步研究在跨国法律实践中，能够与国外法律有效对话，以切实维护我国利益。

4.1.3 国内数据安全标准化相关政策

目前国内已经出台关于数据安全标准化的法律政策文件，为推进数据安全标准化工作提供了法律保障和意见指导。

4.1.3.1 法律层面

全国人大针对《中华人民共和国标准化法》（《标准化法》）进行了多次修订，对加强和推进国家标准化工作起到了重要作用。《标准化法》对标准的制定、实施和标准化工作的监督做出明确规定和要求，同时对违反该法和相关法律规定的行为依法追究法律责任。数据安全标准化工作的开展应当和必须依据《标准化法》，在其指导下进行数据安全相关标准的制定、实施和监督，确保数据安全标准化工作顺利进行。

4.1.3.2 顶层设计

2015 年 8 月，国务院印发《促进大数据发展行动纲要》，在指导思想中明确提出“完善法规制度和标准体系，科学规范利用大数据，切实保障数据安全”：在政策机制中明确提出建立标准规范体系；推进大数据产业标准体系建设，加快建立政府部门、事业单位等公共机构的数据标准和统计标准体系，推进数据采集、政府数据开放、指标口径、分类目录、交换接口、访问接口、数据质量、数据交易、技术产品、安全保密等关键共性标准的制定和实施；加快建立大数据市场交易标准体系；开展标准验证和应用试点示范，建立标准符合性评估体系；充分发挥标准在培育服务市场、提升服务能力、支撑行业管理等方面的作

用；积极参与相关国际标准制定工作。

2015 年 12 月，国务院办公厅印发《国家标准化体系建设发展规划（2016-2020）》（《标准化规划》），部署推动实施标准化战略，加快完善标准化体系，全面提升我国标准化水平。《标准化规划》明确了标准化建设的指导思想、基本原则和发展目标，提出了六项主要任务，确定了五个重点领域，规划了十个重大标准化工程，争取实现关键领域的标准化工作突破，还提出推进标准化工作的保障措施。《规划》在重大标准化工程部分，提出发展新一代信息技术标准化工程，编制新一代信息技术标准体系规划，建立面向未来、服务产业、重点突出、统筹兼顾的标准体系，支撑信息产业创新发展，推动各行业信息化水平全面提升，保障网络安全和信息安全自主可控。指出围绕数据安全等领域研究制定关键技术和共性基础标准，搭建标准化公共服务平台，建立国家网络安全审查技术标准体系并试点应用，开展标准化创新服务机制研究，助力企业实现创新发展。

2016 年 12 月，国务院印发《“十三五”国家信息化规划》，提出建设统一开放的大数据体系，强化数据资源管理：建立健全国家数据资源管理体制机制，建立数据开放、产权保护、隐私保护相关政策法规和标准体系；制定政府数据资源管理办法，推动数据资源分类分级管理，建立数据采集、管理、交换、体系架构、评估认证等标准制度；加强数据资源目录管理、整合管理、质量管理、安全管理，提高数据准确性、可用性、可靠性；完善数据资产登记、定价、交易和知识产权保护等制度，探索培育数据交易市场。此外，在国家互联网大数据平台建设工程中还特别提出“制定国家或行业大数据平台技术标准，形成统一的数据采集、分析处理、安全访问等机制”。

4.1.3.3 政策文件

2016 年 8 月 24 日，中央网信办、国家质检总局、国家标准委联合印发《关于加强国家网络安全标准化工作的若干意见》（以下简称《标准化意见》），对加强网络安全标准化工作做出部署。《标准化意见》指出，建立统一的国家标准工作机制、网络安全行业标准联络员机制和会商机制、重大工程和重大科技项目标准信息共享机制、军民网络安全标准协调机制和联络员机制，以建立统筹协调、分工协作的工作机制；提出科学构建标准体系，优化完善各级标准，推进急需重点标准制定，以加强标准体系建设；要提高标准适用性、先进性、规范性和基础能力建设。《标准化意见》在加强标准体系建设中指出，推进急需重点标准制定，加快开展网络安全审查、大数据安全、个人信息保护等领域的标准研究和制定工作，为加强数据安全标准化工作提供方向和指导。

4.2 主要标准化组织

目前，世界范围内有多个标准化组织正在开展大数据和大数据安全相关标准化工作，主要有国际标准化组织/国际电工委员会的信息技术联合委员会（ISO/IEC JTC1）下属的大数据工作组（WG9）和安全技术分委员会（SC27）、国际电信联盟电信标准化部门（ITU-T）、美国国家标准与技术研究院（NIST）等。国内正在开展大数据和大数据安全相关标准化工作的标准化组织，主要有国家标准化管理委员会（SAC）下属的全国信息技术标准化委员会（以下简称“信标委”，委员会编号为 TC28）和全国信息安全标准化技术委员会（以下简

称“信安标委”，委员会编号为 TC260）等。

4.2.1 ISO/IEC JTC1

4.2.1.1 ISO/IEC JTC1 WG9

ISO/IEC JTC1 WG9 是 ISO/IEC JTC1 于 2014 年 11 月成立的大数据工作组，其工作范围包括聚焦和支持 JTC1 的大数据标准计划，编制大数据基础标准（包括参考架构和术语标准），识别大数据标准化中的差距，与涉及大数据相关工作的其他标准组织建立和维护联络关系等。

目前正在编制 ISO/IEC 20546《信息技术 大数据 概述和词汇》和 ISO/IEC 20547《信息技术 大数据参考架构》两项国际标准。ISO/IEC 20547 为多部分标准，包括 ISO/IEC TR 20547-1《第 1 部分：框架和应用过程》、ISO/IEC TR 20547-2《第 2 部分：用例和衍生需求》、ISO/IEC 20547-3《第 3 部分：参考架构》、ISO/IEC 20547-4《第 4 部分：安全与隐私保护》、ISO/IEC TR 20547-5《第 5 部分：标准路线图》。

其中，ISO/IEC 20547-4《信息技术 大数据参考架构 第 4 部分：安全与隐私保护》标准编制项目根据 ISO/IEC JTC1 JAG（JTC1 咨询小组）2016 年 3 月巴黎会议决定被转交给 ISO/IEC JTC1 SC27，现由 SC27 下属的安全控制与服务工作组（WG4）负责、身份管理与隐私保护技术工作组（WG5）配合，其编辑工作由中国专家担任，目前进入第三版工作草案阶段。

4.2.1.2 ISO/IEC JTC1 SC27

ISO/IEC JTC1 SC27 是 ISO/IEC JTC1 下属安全技术分委员会，成立于 1990 年，其工作范围涵盖信息和信息与通信技术（ICT）保护的标准开发，包括安全与隐私保护方面的方法、技术和指南。目前下设五个工作组，分别为信息安全管理体系工作组（WG1）、密码技术与安全机制工作组（WG2）、安全评价、测试和规范工作组（WG3）、安全控制与服务工作组（WG4）和身份管理与隐私保护技术工作组（WG5）。各工作组负责各自工作范围内的多项标准开发，并根据需要设立相应的研究项目。

WG4 负责信息安全控制与服务方面的标准研制和维护。WG4 负责制定的标准中直接与大数据安全相关的标准有正在编制中的 ISO/IEC 20547-4《信息技术 大数据参考架构 第 4 部分：安全与隐私保护》；涉及大数据运行平台云计算安全的标准有正在编制中的 ISO/IEC 19086-4《云计算 服务水平协议（SLA）框架 第 4 部分：安全与隐私保护》；涉及数据存储安全的标准有 ISO/IEC 27040:2015《信息技术 安全技术 存储安全》；涉及信息安全事件管理及调查取证的标准有 ISO/IEC 27035《信息技术 安全技术 信息安全事件管理》（包括三个部分）、ISO/IEC 27037:2012《信息技术 安全技术 数字证据的识别、收集、获得和保全指南》、ISO/IEC 27038:2014《信息技术 安全技术 数字脱敏规范》、ISO/IEC 27041:2015《信息技术 安全技术 确保事件调查方法适宜性和充足性的指南》、ISO/IEC 27042:2015《信息技术 安全技术 数字证据分析和解释指南》、ISO/IEC 27043:2015《信息技术 安全技术 事件调查原则和过程》和 ISO/IEC 27050《信息技术 安全技术 电子发现》（包括四个部分）。

WG5 负责身份管理和隐私保护方面的标准研制和维护。WG5 负责制定的标准中涉及个人隐私保护方面的标准有 ISO/IEC 29100:2011《信息技术 安全技术

隐私保护框架》、ISO/IEC 29101:2013《信息技术　安全技术　隐私保护体系结构框架》、ISO/IEC 29134《信息技术　安全技术　隐私影响评估指南》、ISO/IEC 29151《信息技术　安全技术　可识别个人信息（PII）保护实践指南》、ISO/IEC 29184《信息技术　安全技术　在线隐私通知和准许指南》、ISO/IEC 29190:2015《信息技术　安全技术　隐私保护能力评估模型》、ISO/IEC 29191:2012《信息技术　安全技术　部分匿名、部分不可链接鉴别要求》、ISO/IEC 27018:2014《信息技术　安全技术　可识别个人信息（PII）处理者在公有云中保护 PII 的实践指南》、ISO/IEC 27550《信息技术　安全技术　隐私保护工程》和 ISO/IEC 27551《信息技术　安全技术　对 ISO/IEC 27001 在隐私保护管理方面的增强要求》。

我国专家积极参与 SC27 大数据安全标准化工作，目前 SC27 大数据安全直接相关的标准化工作 3 项：ISO/IEC 20547-4《信息技术　大数据参考架构　第 4 部分：安全与隐私保护》以及两项研究项目《大数据安全能力成熟度模型》、《大数据安全实施指南》，均由我国专家主导。

4.2.2 NIST

美国国家标准与技术研究院（NIST）于 2012 年 6 月启动了大数据相关基本概念、技术和标准需求的研究，2013 年 5 月成立了 NIST 大数据公共工作组（NBG-PWG），对所有感兴趣的相关方开放，无会员费，旨在通过结合行业、学术和政府等各方力量加速对大数据这一新兴产业的采纳，其成果由 NIST 评审和发布。

2015 年 9 月编写形成并发布了 NIST SP 1500《NIST 大数据互操作框架》系列标准（第一版），包括 7 个分册，即 NIST SP 1500-1《第 1 册　定义》、NIST SP 1500-2《第 2 册　大数据分类法》、NIST SP 1500-3《第 3 册　用例和一般要求》、NIST SP 1500-4《第 4 册　安全和隐私保护》、NIST SP 1500-5《第 5 册　架构调研白皮书》、NIST SP 1500-6《第 6 册　参考架构》和 NIST SP 1500-7《第 7 册　标准路线图》。目前正在进行第二版的编制工作，并增加了 2 个分册，即 NIST SP 1500-8《第 8 册　大数据参考架构接口》和 NIST SP 1500-9《第 9 册　大数据采用与现代化》。

其中，NIST SP 1500-4《NIST 大数据互操作框架：第 4 册　安全与隐私保护》由 NIST NBD-PWG 的安全与隐私保护小组负责编写。

4.2.3 ITU-T

ITU-T 在 2013 年 11 月发布了《大数据：今天巨大，明天平常》报告，并在其下属相关研究组开展了多项大数据和大数据安全相关的标准化工作。

ITU-T SG13（聚焦于 IMT-2020、云计算和可信网络基础设施的未来网络研究组）负责制定的大数据相关标准有 ITU-T Y.3600《大数据　基于云计算的要求和能力》、《大数据　元数据框架和概念模型》、《大数据　数据溯源要求》、《大数据交换框架和要求》、《数据存储联合要求》、《大数据即服务的功能架构》、《大数据　数据保全概述和要求》、《大数据驱动联网要求》、《基于 DPI 的大数据驱动联网框架》、《大数据驱动联网的用例和应用场景》、《大数据驱动的移动网络流量管理与规划》和《应用于网络大数据语境下的深度包检测机制》。

ITU-T SG17（安全研究组）负责制定的大数据安全相关标准有《移动互联网服务中的大数据分析安全要求和框架》、《大数据即服务的安全指南》、《电子商务业务数据生命周期管理安全参考架构》和《电信大数据生存周期安全指南》。

4.2.4 SAC TC28

为规范和推动我国大数据产业的快速发展，培育大数据产业链，并与国际标准接轨，全国信标委于 2014 年 12 月成立了大数据标准工作组，工作组主要负责制定和完善我国大数据领域标准体系，组织开展大数据相关技术和标准的研究，推动国际标准化活动，对口 ISO/IEC JTC1 WG9 大数据工作组。

目前，大数据标准工作组已经发布《信息技术 大数据 术语》、《信息技术 大数据 技术参考模型》、《信息技术 科学数据引用》三项标准，正在制定的国家标准有 16 项，包括《数据能力成熟度评价模型》、《信息技术 通用数据导入接口规范》、《信息技术 数据质量评价指标》、《信息技术 数据溯源描述模型》、《信息技术 数据交易服务平台 通用功能要求》、《信息技术 数据交易服务平台 交易数据描述》、《多媒体数据语义描述要求》、《信息技术 大数据 面向应用的基础计算平台基本性能要求》、《信息技术 大数据 开放共享》（第 1 部分：总则；第 2 部分：政府数据开放共享基本要求；第 3 部分：开放程度评价）、《信息技术 大数据 分类指南》、《信息技术 大数据 分析系统功能测试规范》、《信息技术 大数据 基于参考架构下的接口框架》、《信息技术 大数据 存储与处理系统功能测试规范》和《信息技术 大数据 系统通用规范》。

4.2.5 SAC TC260

为了加快推动我国大数据安全标准化工作，全国信安标委于 2016 年 4 月成立大数据安全标准特别工作组，主要负责制定和完善我国大数据安全领域标准体系，组织开展大数据安全相关技术和标准研究。

目前，大数据安全标准特别工作组已经发布《信息安全技术 个人信息安全规范》、《信息安全技术 大数据服务安全能力要求》两项标准，正在制定的大数据和个人信息安全相关国家标准有 6 项，包括《信息安全技术 大数据安全管理指南》、《信息安全技术 数据安全能力成熟度模型》、《信息安全技术 数据交易服务安全要求》、《信息安全技术 数据出境安全评估指南》、《信息安全技术 个人信息安全影响评估指南》和《信息安全技术 个人信息去标识化指南》。

4.3 大数据安全相关标准现状

大数据安全问题不仅有外部因素影响，从内部安全管理角度分析，监督管理体系健全与否，也关乎大数据的生存境况。为应对复杂的信息安全环境，满足大数据时代的信息安全管理需求，全球众多政府机构、科研组织陆续出台有关大数据安全管理的标准或规范；特别是在重要行业和关键领域，比如电子政务、通信、金融等早已出台相应的安全标准和技术规范。

下面将从数据安全相关标准、当前被广泛关注的个人信息安全标准以及大

数据安全标准三个方面展开，进行详细分析。

本节重点关注全国信安标委大数据安全标准特别工作组之外开展的相关安全标准工作，大数据安全标准特别工作组开展的大数据安全标准工作将在第 5 章详细介绍。

4.3.1 数据安全相关标准

按照数据采集、传输、存储、处理、交换、销毁的全生命周期，梳理出目前数据安全相关标准如表 4.5 所示。可以看出，数据安全标准化工作主要还是重点研究数据收集、存储、传输等环节，**数据使用、共享、销毁等环节的工作还需进一步加强**。

表 4.5 数据安全相关标准

序号	标准类型	标准编号	标准名称
基础标准			
1	国际标准	ISO/IEC 9579:2000	信息技术　具有安全增强的 SQL 远程数据库访问
2	国家标准	GB/T 18391—2009	信息技术 元数据注册系统(MDR)
3	国家标准	GB/T 17859—1999	计算机信息系统 安全保护等级划分准则
4	国家标准	GB/T 22239—2008	信息安全技术 信息系统安全等级保护基本要求
5	国家标准	GB/T 31503—2015	信息安全技术 电子文档加密与签名消息语法
6	国家标准	GB/T 32918—2016	信息安全技术 SM2 椭圆曲线公钥密码算法
7	国家标准	GB/T 32905—2016	信息安全技术 SM3 密码杂凑算法
8	国家标准	GB/T 32907—2016	信息安全技术 SM4 分组密码算法
数据收集阶段			
9	国家标准	GB/T 14258—2003	信息技术 自动识别与数据采集技术条码符号印制质量的检验
10	国家标准	GB/T 28788—2012	公路地理信息数据采集与质量控制
11	国家标准	GB/T 26237—2010	信息技术　生物特征识别数据交换格式
12	国家标准	GB/T 27912—2011	金融服务　生物特征识别　安全框架
数据存储阶段			
13	国家标准	GB/T 20273—2006	信息安全技术　数据库管理系统安全技术要求
14	国家标准	GB/T 20009—2005	信息安全技术　数据库管理系统安全评估准则
15	通信行标	YD/T 2390—2011	通信存储介质（SSD）加密安全技术要求
16	通信行标	YD/T 2665—2013	通信存储介质（SSD）加密安全测试方法
数据传输阶段			
17	国家标准	GB/T 17963—2000	信息技术　开放系统互连　网络层安全协议
18	国家标准	GB/T 28456—2012	IPSec 协议应用测试规范
19	国家标准	GB/T 28457—2012	SSL 协议应用测试规范
20	电子行标	SJ 20951—2005	通用数据加密模块接口要求
21	通信行标	YD/T 1466—2006	IP 安全协议（IPSec）技术要求
22	通信行标	YD/T 1467—2006	IP 安全协议（IPSec）测试方法
23	通信行标	YD/T 1468—2006	IP 安全协议（IPSec）穿越网络地址翻译（NAT）技术要求
24	通信行标	YD/T 2908—2015	基于域名系统（DNS）的 IP 安全协议（IPSec）认证密钥存储技术要求
25	国家标准/国际标准	GB/T 18794—2002 (ISO 10181:1996, IDT)	信息技术　开放系统互连　开放系统安全框架

序号	标准类型	标准编号	标准名称
数据使用阶段			
26	国家标准	GB/T 32908—2016	非结构化数据访问接口规范
27	国家标准	GB/T 25000.12—2017	系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第 12 部分：数据质量模型
28	国家标准	GB/T 31594—2015	社会保险核心业务数据质量规范
29	国家标准	GB/T18784—2002	CAD-CAM 数据质量
30	国家标准	GB/T 28441—2012	车载导航电子地图数据质量规范
数据共享阶段			
31	国家标准	GB/T 7408—2005	数据元和交换格式 信息交换 日期和时间表示法
32	国家标准	GB/T 21062—2007	政务信息资源交换体系
数据销毁阶段			
33	公安行标	GA/T 1143—2014	信息安全技术　数据销毁软件产品安全技术要求

1. ISO/IEC 9579:2000《信息技术　具有安全增强的 SQL 远程数据库访问》

该标准与 ISO/IEC 10032:2003《信息技术　数据管理参考模型》（Information technology — Reference Model of Data Management）和 ISO/IEC 9075-1:2016《信息技术　数据库语言　SQL　第一部分：框架（SQL/框架）》（Information technology — Database languages — SQL — Part 1: Framework (SQL/Framework)）配套使用，其目标是促进数据库系统在不同架构、不同管理策略、不同复杂性水平、采用不同技术的情况下的互相访问与内在连接。

2. GB/T 18391《信息技术 元数据注册系统(MDR)》

该系列标准描述了数据的语义、数据的表示以及这些数据描述的注册。通过这些描述，可以找到语义的确切理解及数据的有用描述。本标准的目的在于促进：数据的标准描述、组织内以及组织间对数据的一致理解、跨越时间、空间和应用对数据的重用和标准化、组织内和组织间数据的协同及标准化、数据成分的管理、数据成分的重用。第 1 部分：框架。标准和基本概念；第 2 部分：分类。元数据注册系统中分类方案的管理；第 3 部分：注册系统元模型与基本属性。元数据注册系统基本概念模型，包括基本属性和关系；第 4 部分：数据定义的行程。给出了构成高质量数据元及其成分定义的规则与指南；第 5 部分：命名和标识原则。描述了如何为数据元及其成分建立命名约定。第 6 部分：注册。规定了符合 GB/T 18391 的元数据注册系统注册过程的角色和要求。

3. GB/T 17859—1999《计算机信息系统 安全保护等级划分准则》

该标准主要有三个目的：一、为计算机信息系统安全法规的制定和执法部门的监督检查提供依据；二、为安全产品的研制提供技术支持；三、为安全系统的建设和管理提供技术指导。该标准规定了计算机信息系统安全保护能力的五个等级，适用于计算机信息系统安全保护技术能力等级的划分。计算机信息系统安全保护能力随着安全保护等级的增高，逐渐增强。

4. GB/T 22239—2008《信息安全技术 信息系统安全等级保护基本要求》

该标准规定了不同安全保护等级信息系统的基本保护要求，包括基本技术要求和基本管理要求。适用于指导分等级的信息系统的安全建设和监督管理。

5. GB/T 31503—2015《信息安全技术 电子文档加密与签名消息语法》

该标准规定了电子文档加密与签名消息语法,此语法可用于对任意消息内容进行数字签名、摘要、鉴别或加密。适用于电子商务和电子政务中电子文档加密与签名消息的产生、处理以及验证。

6. GB/T 32918—2016《信息安全技术 SM2 椭圆曲线公钥密码算法》

该标准规定了 SM2 椭圆曲线公钥密码算法，适用于基域为素域和二元扩域的椭圆曲线公钥密码算法的设计、开发、使用。

7. GB/T 32905—2016《信息安全技术 SM3 密码杂凑算法》

该标准规定了 SM3 密码杂凑算法的计算方法和计算步骤，并给出了运算示例，适用于商用密码应用中的数字签名和验证、消息认证码的生成与验证以及随机数的生成。

8. GB/T 32907—2016《信息安全技术 SM4 分组密码算法》

该标准规定了 SM4 分组密码算法的算法结构和算法描述，并给出了运算示例，适用于商用密码产品中分组密码算法的实现、检测和应用。

9. GB/T 14258—2003《信息技术 自动识别与数据采集技术条码符号印制质量的检验》

该标准规定了一维条码符号印制质量的检验方法。适用于印制的一维条码符号的质量检验，具体应用领域有专用条码符号检验国家标准时，应按其检验标准进行检验。

10. GB/T 28788—2012《公路地理信息数据采集与质量控制》

该标准规定了公路地理信息采集的技术指标、采集内容和方法、数据属性结构、质量控制和成果验收等要求，同时也规定了使用移动道路测量技术采集公路地理信息的作业流程。该标准适用于建立各级公路地理信息数据库及相关地理信息系统时，对公路工程建设的基础地理信息的采集、处理与交换；适用于 1:10000、1:50000、1:100000、1:250000 国道、省道和县道的数据采集与更新。乡道及专用公路数据采集、1:5000 公路数据采集与更新也可参照执行。

11. GB/T 26237《信息技术　生物特征识别数据交换格式》系列标准

该系列标准规定了生物特征识别数据结构的用法概述；生物特征识别数据结构的类型；生物特征识别数据结构的命名思想；格式类型的编码方案，并从指纹细节点、指纹型谱、指纹图像、人脸图像、虹膜图像、签名/签字时间序列、指纹型骨架、血管图像、手型轮廓等方面提出生物特征识别涉及的数据交换格式。

12. GB/T 27912—2011《金融服务　生物特征识别　安全框架》

该标准规定了金融业使用生物特征识别机制鉴别人员身份的安全框架，介绍了生物特征识别技术的类型，阐述了有关应用问题。该标准也描述了实现架构，详细规定了有效管理的最小安全要求，也为专业人员提供了控制目标和使用建议。该标准包括：1）使用生物特征识别技术，通过验证其声称的身份或识别其个体身份，对参与金融服务的人员和雇员身份进行鉴别；2）根据风险管理的要求，对用户登记时提交的凭证进行确认，以支持身份鉴别；3）在整个生命周期内，包括登记、传输、存储、身份确认、身份识别以及终止等过程，对生物特征信息进行管理；4）生物特征识别信息在其生命周期内的安全性，包括数据完整性、源鉴别和机密性；5）生物特征识别机制在逻辑和物理访问控制中的应用；6）保护金融机构及其客户的监控措施；7）在整个生物特征识别信息生命周期中所使用的物理硬件的安全性。该标准不包括：1）个体生物特征识别信息的隐私权和所有权；2）有关数据采集、信号处理与生物特征数据匹配、以及

生物特征匹配决策流程等方面的具体技术；3）生物特征识别技术在非鉴别方面的便利性应用，如语音识别、用户交互和匿名访问控制等方面的使用。该标准适用于由于数据机密性或其他原因而对生物特征信息进行加密的强制方式。

13. GB/T 20273—2006《信息安全技术　数据库管理系统安全技术要求》

该标准依据 GB 17859—1999 的五个安全保护等级的划分，根据数据库管理系统在信息系统中的作用，规定了各个安全等级的数据库管理系统所需要的安全技术要求。该标准适用于按等级化要求进行的安全数据库管理系统的设计和实现，对按等级化要求进行的数据库管理系统安全的测试和管理可参照使用。

14. GB/T 20009—2005《信息安全技术　数据库管理系统安全评估准则》

该标准从信息技术方面规定了按照 GB 17859—1999 的五个安全保护等级对数据库管理系统安全保护等级划分所需要的评估内容。该标准适用于数据库管理系统的安全保护等级的评估，对于数据库管理系统安全功能的研制、开发和测试亦可参照使用。

15. YD/T 2390—2011《通信存储介质（SSD）加密安全技术要求》

该标准规定了固态硬盘（SSD，Solid State Disk）加解密系统架构、加解密流程、身份认证模块要求、加解密算法模块要求和密钥管理模块要求等内容。该标准适用于固态硬盘产品。

16. YD/T 2665—2013《通信存储介质（SSD）加密安全测试方法》

该标准规定了通信存储介质（SSD）加密安全测试方法，包括设备缺省设置、启动流程、身份认证、加/解密算法、密钥管理等相关测试内容。该标准中出现的所有未指明的受测设备、加密硬盘等均特指用于通信领域的 SSD 类存储设备。该标准适用于通用的通信存储介质（SSD）产品的加密安全特性的测试。不支持用户数据加密功能的 SSD 不适用该标准。

17. GB/T 17963—2000《信息技术　开放系统互连　网络层安全协议》

该标准规定的协议将由端系统和中间系统使用，以在网络层提供安全服务，而网络层由 GB/T 15126 和 GB/T 15274 定义。该标准中定义的协议称为网络层安全协议（NLSP）。

该标准规定：A)支持 GB/T 9387.2 中定义的下列安全服务：1）对等实体鉴别；2）数据原发鉴别；3）访问控制；4）连接保密性;5）无连接保密性;6）通信流量保密性；7）无恢复的连接完整性（包括数据单元完整性，其中连接上的各个 SDU 具有完整性保护）；8）无连接完整性。B)声称与该标准一致的实现的功能要求。

该协议的规程根据下列定义：1）可用于该协议实例的加密技术的要求；2）用于通信实例安全联系中携带信息的要求。

尽管一些安全机制提供的保护程序取决于一些特定加密技术，而该协议的正确操作并不取决于某种特定的加密或解密算法的选择。这是通信系统的本地事情。此外，特定的安全策略的选择和实现都不在该标准的范围之内。特定的安全策略的选择以及因此将达到的保护程度，留作使用安全通信的单个实例的系统之间的本地事情。该标准不要求涉及同一开发系统的多个安全通信的实例必须采用相同的协议。 附录 D 按照 ISO/IEC 9646-2 中给出的相关指导为网络层协议提供了 PICS 形式表。

18. GB/T 28456—2012《IPSec 协议应用测试规范》

该标准对IPSec协议应用的测试内容及测试步骤进行了规范。该标准适用于IPSec 协议应用的开发单位、第三方授权测试认证机构、用户等对 IPSec 协议应用测试时参考使用。

19. GB/T 28457—2012《SSL 协议应用测试规范》

该标准规定了 SSL 协议应用的测试内容和基本测试步骤。该标准适用于SSL 协议应用的开发单位、第三方授权测试认证机构、用户等对 SSL 协议应用的测试。

20. SJ 20951—2005《通用数据加密模块接口要求》

该标准规定了通用数据加密模块（以下简称数据加密模块）与主机的接口要求，包括物理连接、数据交互协议以及对主机的基本要求等。该标准适用于数据加密模块和主机的接口设计、研制、检验和验收。

21. YD/T 1466—2006《IP 安全协议（IPSec）技术要求》

该标准规定了 IP 安全协议(IPSec)的技术要求，包括 IPSec 体系结构、IPSec 的安全联盟、AH 协议和 ESP 协议等。该标准适用于支持 IPSec 的数据设备。

22. YD/T 1467—2006《IP 安全协议（IPSec）测试方法》

该标准规定了 IPSec 的测试方法，包括 AH 和 ESP 功能测试和性能测试等。该标准适用于支持 IPSec 协议的数据设备。

23. YD/T 1468—2006《IP 安全协议（IPSec）穿越网络地址翻译（NAT）技术要求》

该标准规定了 IPSec 穿越 NAT 的技术要求，包括 IPSec 穿越 NAT 存在的兼容性问题、兼容性要求、解决方法以及穿越 NAT 对 IPSec 的影响等。该标准适用于支持 IPSec 穿越 NAT 的数据设备。

24. YD/T 2908—2015《基于域名系统(DNS)的 IP 安全协议(IPSec)认证密钥存储技术要求》

该标准规定了一种基于 DNS 的 IPSec 认证密钥及加密点信息存储方法，该方法可用于从 DNS 权威服务器获取 IPSec 目标系统的密钥信息和加密点信息。该标准规定了该资源记录的数据格式及其使用方法。该标准适用于部署有安全的 DNS 服务（通过 DNSSEC 或类似技术实现）的系统。

25. GB/T 18794—2002《信息技术　开放系统互连　开放系统安全框架》（ISO 10181，IDT）

安全框架涉及在开放系统环境中安全服务的应用，其中术语“开放系统”系指包括诸如数据库、分布式应用、ODP 和 OSI 一类的领域。安全框架主要是用来提供在系统内和系统间交互时对系统和客体的保护方法。安全框架不考虑用于构造系统或者机制的方法学。

安全框架涉及用于获取具体安全服务所使用的数据元素和操作序列（但不是协议元素）。这些安全服务可适用于系统的通信实体，也可以用于系统间交互的数据和由系统管理的数据。

安全框架提供了进一步标准化的基础，对特定安全需求的通用抽象服务接口提供了一致性的术语和定义。安全框架还对能够用于实现这些需求的机制进行了分类。

一些安全服务经常依赖于其他的安全服务，使得安全的一部分与其他的部分进行隔离很困难。安全框架描述了特定的安全服务，描述能够用于提供这些安全服务的机制范围，并标识这些服务和机制间的相互关系。这些机制的描述可能涉及对不同安全服务的依赖关系，安全框架用此方式描述一个安全服务对

另一个安全服务的依赖关系。

该系列标准包括：第 1 部分：框架，第 2 部分：鉴别框架，第 3 部分：访问控制框架，第 4 部分：抗抵赖框架，第 5 部分：机密性框架，第 6 部分：完整性框架，第 7 部分：安全审计和报警框架。

26. GB/T 32908—2016《非结构化数据访问接口规范》

该标准规定非结构化数据管理系统的访问接口要求，包括查询语言访问接口、应用程序访问接口和 Web 服务访问接口。该标准适用于非结构化数据管理系统产品的研制、开发和测试。

27. GB/T 25000.12—2017《系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第 12 部分：数据质量模型》

该标准针对计算机系统中以某种结构化形式保存的数据，定义了通用的数据质量模型。该标准关注于作为计算机系统一个组成部分的数据的质量，并定义由人和系统使用的目标数据的质量特性。

28. GB/T 31594—2015《社会保险核心业务数据质量规范》

该标准涉及养老、医疗、失业、工伤、生育等社会保险的核心业务数据，是社会保险国家标准的一个重要组成部分。该标准结合我国实际状况，对社会保险核心业务数据的指标、数据质量两方面做出了规范性要求。

29. GB/T18784—2002《CAD-CAM 数据质量》

该标准主要涉及到不同 CAD/CAM 系统之间共享信息、交换 CAD 模型数据的问题。同时还要求公司的各部门之间、不同公司之间，在交换 CAD 模型数据时对模型信息的组织和质量水平达成共同的理解和协定。该标准的使用者包括企业中 CAD/CAM 数据质量的负责人员和软件开发人员，也包括 CAD/CAM 数据和软件系统的最终用户。该标准的数据质量是指产品数据精度和可用性满足数据用户要求的程度，好的数据质量是指在正确的时间将正确的数据传送给正确的人。

30. GB/T 28441—2012《车载导航电子地图数据质量规范》

该标准规定了车载导航电子地图数据质量描述原则、评价内容、评价指标、评价方法和评价流程。该标准适用于车载导航电子地图数据的检查验收和质量评定。

31. GB/T 7408—2005 《数据元和交换格式 信息交换 日期和时间表示法》

该标准规定了公历日期和时间以及时间间隔的表示法，适用于在信息交换中涉及的日期和时间表示。

32. GB/T 21062.1～GB/T 21062.4《政务信息资源交换体系》

该系列标准提出了政务信息资源交换体系的总体技术架构，规定了政务信息资源交换体系技术支撑环境的组成；规范了政务信息资源交换体系技术支撑环境功能组成及要求，规定了信息交换系统间互联互通的技术要求；规定了信息交换时封装业务数据采用的数据接口规范，提出了交换指标项；规定了政务信息资源交换体系的技术管理总体架构、管理角色的职责、交换体系各环节的技术管理要求。

33. GA/T 1143—2014《信息安全技术 数据销毁软件产品安全技术要求》

针对磁性介质的存储原理和数据读写方法，普通的数据销毁，如低级格式化、数据删除等方法都无法彻底清除数据、操作系统和磁盘的隐形操作产生的

残留数据。根据《中共中央保密委员会办公室、国家保密局关于国家秘密载体保密的规定》销毁秘密载体时必须确保秘密信息无法还原。因此，根据以上特点和要求，该标准对在计算机信息系统中使用的、针对磁性存储介质的数据销毁软件产品的设计、开发和检测提出了技术要求。标准对数据销毁软件的定义、安全环境、安全目的、安全功能、安全保障进行了定义和要求。同时，描述了技术要求的基本原理，对产品的技术要求进行了等级划分。

4.3.2 个人信息安全标准

个人信息安全相关国际标准如表 4.6 所示。

表 4.6 个人信息安全相关国际标准

序号	标准类型	标准编号	标准名称
1.	国际标准	ISO/IEC 29100:2011	信息技术　安全技术　隐私保护框架
2.	国际标准	ISO/IEC 29101:2013	信息技术　安全技术　隐私保护体系结构框架
3.	国际标准	ISO/IEC 29190:2015	信息技术　安全技术　隐私保护能力评估模型
4.	国际标准	ISO/IEC 29191:2012	信息技术　安全技术　部分匿名、部分不可链接鉴别要求
5.	国际标准	ISO/IEC 27018:2014	信息技术　安全技术　可识别个人信息（PII）处理者在公有云中保护 PII 的实践指南
6.	国际标准	ISO/IEC 29134	信息技术　安全技术　隐私影响评估指南
7.	国际标准	ISO/IEC 29151	信息技术　安全技术　可识别个人信息（PII）保护实践指南
8.	国际标准	ISO/IEC 27550	信息技术　安全技术　隐私保护工程
9.	国际标准	ISO/IEC 27551	对 ISO/IEC 27001 在隐私保护管理方面的增强要求
10.	国际标准	ISO/IEC 29184	信息技术　安全技术　在线隐私通知和准许指南
11.	英国标准	BS 10012:2009	数据保护　个人信息管理系统规范

1　　**ISO/IEC 29100:2011《信息技术　安全技术　隐私保护框架》**

该标准为信息与通信技术（ICT）系统内可识别个人信息（PII）的保护提供了一个高层次隐私保护框架。该隐私保护框架规范了通用的隐私保护术语；定义了处理 PII 中的参与者及其角色；描述了隐私保护的考虑事项；为实现由许多国际组织开发的 11 个隐私保护原则提供指导。11 个隐私保护原则包括同意和选择、意图合法性和规约、收集限制、数据最小化、使用/保留/披露限制、准确和质量、开放/透明/告知、个体参与和访问、可核查性、信息安全、隐私保护合规。该标准适用于涉及规范、获取、构建、设计、开发、测试、维护、管理和运行需要隐私保护控制措施来处理 PII 的 ICT 系统或服务的任何自然人和组织。

2　　**ISO/IEC 29101:2013《信息技术　安全技术　隐私保护体系结构框架》**

该标准定义了一个隐私参考体系结构框架，该框架明确提出了处理 PII 的 ICT 系统的关心点，列出了实现这种系统的组件，并提供了将这些组件语境化的体系结构视图。该标准适用于涉及规划、获取、构建、设计、测试、维护、

管理和运行处理 PII 的 ICT 系统的实体。

3 **ISO/IEC 29190:2015《信息技术　安全技术　隐私保护能力评估模型》**

该标准为组织评估其管理隐私保护相关过程的能力提供高层指南，规范了确定隐私保护能力的评估过程和评估级别，为评估隐私保护能力的关键过程域及其实现，以及如何将隐私保护能力评估继承到组织运行中提供了指南。

4 **ISO/IEC 29191:2012《信息技术　安全技术　部分匿名、部分不可链接鉴别要求》**

该标准为部分匿名和部分不可链接鉴别要求提供基本框架并建立相应要求。

5 **ISO/IEC 27018:2014《信息技术　安全技术　可识别个人信息（PII）处理者在公有云中保护 PII 的实践指南》**

该标准依据 ISO/IEC 29100 给出的隐私保护原则，为在公有云计算环境中保护可识别个人信息（PII），建立了普遍接受的控制目标、控制措施和测量实现指南。特别是，该标准考虑到在公有云提供者的信息安全风险环境下适用的 PII 保护法规要求，基于 ISO/IEC 27002 给出指南。该标准适用于作为 PII 处理者通过云计算提供信息处理服务的所有类型和规模的组织。

6 **ISO/IEC 29134《信息技术　安全技术　隐私影响评估指南》**

该标准为隐私影响评估（PIA）过程以及 PIA 报告的结构和内容给出指南。该标准适用于所有类型和规模组织。

7 **ISO/IEC 29151《信息技术　安全技术　可识别个人信息（PII）保护实践指南》**

该标准为满足通过可识别个人信息（PII）保护相关的风险和影响评估而识别的要求，建立了控制目标和控制措施，并提供了控制措施实现指南。该标准考虑到在组织信息安全风险环境下适用的 PII 处理要求，基于 ISO/IEC 27002 给出指南。该标准适用于作为 PII 控制者的所有类型和规模的组织。

8 **ISO/IEC 27550《信息技术　安全技术　隐私保护工程》**

该标准针对企业如何将隐私保护工程与自身工程实践相结合，给出了相应的体系架构和指南，具体包括以下内容：1）隐私工程和其他工程（如系统工程、安全工程、风险管理）之间的关系；2）与知识管理、风险管理、需求分析、架构设计等关键工程过程相关的隐私工程活动；3）隐私工程附录，如与隐私工程相关的实体，考虑到域管理、供应链、软件开发方法等因素后的隐私工程实践，用于指导隐私工程活动的目录，以及隐私风险分析的案例。

该标准的目标人员是需在系统开发、应用或操作维护过程中考虑隐私工程的专业人员和工程师，也适用于企业中负责隐私工程、系统开发、产品管理、市场和运维的相关人员。

9 **ISO/IEC 27551《信息技术　安全技术　对 ISO/IEC 27001 在隐私保护管理方面的增强要求》**

目前实体认证都要求被认证实体提供可识别的身份信息，但在很多交易中，实体更倾向于维持匿名化或非链接性，这就使得完成两笔交易时，很难区分交易是由一个用户还是两个不同的用户完成的。该标准正是针对基于属性的非链接实体认证提出了架构并建立相应要求。

10 **ISO/IEC 29184《信息技术　安全技术　在线隐私通知和准许指南》**

宽带网络等通信基础设施的快速普及、智能手机和可穿戴设备等可收集用户详细信息的终端的广泛应用、信息处理能力的大幅度提升，使得大范围信息

收集和分析成为可能。在技术升级给用户带来使用便利性和有吸引力的服务并催生新商机的同时，用户也变得对“隐私”越来越敏感，对在线服务中的 PII（个人识别数据）收集和使用产生的影响越来越存疑。这种质疑通常是由于未对如何使用、处理、存储个人 PII 数据进行明确的解释造成的。

该标准为企业提供了一个基本架构，可向被收集 PII 数据的用户提供明晰、易于理解的基本信息，解释企业将如何处理这些 PII 数据。同时，该标准为如何落实 ISO/IEC 29100 中的两个隐私原则（原则 1：同意和授权；原则 7：开放、透明和通知）提供了详细指南。

11 **BS 10012:2009《数据保护　个人信息管理系统规范》**

该标准由英国标准协会（BSI）于 2009 年 6 月发布，主要是针对个人信息保护提出了“个人信息保护标准”，其中参考了经济开发合作组织（OECD）的个人隐私权保护的八大原则，用于支撑欧盟隐私保护条例和英国的数据保护法案，强调要建立一个管理体系，并且就个人信息跨境管理的情况给出了建议。

该标准规范了个人信息管理体系（PIMS）要求，提供了一个框架用于维护和改进数据保护的合规性和最佳实践。该标准适用于任何规模和行业的组织，主要为在其内部启动、实施和维护 PIMS 的组织所用。该标准旨在提供个人信息管理的共同基础，以便增强个人信息管理的信心，并使得内部和外部评估者能够有效地评估数据保护的合规性和最佳实践。

4.3.3 其它大数据安全标准

4.3.3.1 国际标准

大数据安全相关国际标准如表 4.7 所示。

表 4.7 大数据安全相关国际标准

序号	标准类型	标准编号	标准名称
1.	国际标准	ISO/IEC 20547-4	信息技术　大数据参考架构　第 4 部分：安全与隐私保护
2.	国际标准	ISO/IEC 19086-4	云计算　服务水平协议（SLA）框架　第 4 部分：安全与隐私保护
3.	国际标准	ITU-T Y.3600	大数据　基于云计算的要求和能力

1 **ISO/IEC 20547-4《信息技术　大数据参考架构　第 4 部分：安全与隐私保护》**

该标准由我国专家担任编辑，分析了大数据面临的安全与隐私保护问题和相关风险，在 ISO/IEC 20547-3《信息技术　大数据参考架构　第 3 部分：参考架构》给出的大数据参考架构（BDRA）基础上，提出了大数据安全与隐私保护参考架构（BDRA-S&P）。BDRA-S&P 包括用户视角的大数据安全与隐私保护角色和活动，以及功能视角的支持大数据安全与隐私保护活动的功能组件。该标准还汇集了信息安全领域中已有的安全控制措施和隐私保护控制措施，作为大数据安全与隐私保护功能组件的选项。

2 **ISO/IEC 19086-4《云计算　服务水平协议（SLA）框架　第 4 部分：安全与隐私保护》**

该标准定义了云计算服务水平协议（SLAs）框架，为考虑迁移到云的组织和云服务提供商提供指导，该框架提供了一种结构用于在选用云时，确定相应的性能、服务、数据管理和治理目标和要求。其中第四部分将识别云 SLAs 的安

全和隐私要求。

3　　ITU-T Y.3600《大数据　基于云计算的要求和能力》

该标准明确指出了大数据定义、大数据生态系统的特征等基本问题，描述了大数据与云计算之间的关系，并从数据采集、数据预处理、数据存储、数据分析、数据可视化、数据管理、数据安全与保护等 7 个方面给出了基于云计算的大数据要求、能力和用例。

4.3.3.2 国外标准

大数据安全相关国外标准如表 4.8 所示。

表 4.8 大数据安全相关国外标准

序号	标准类型	标准编号	标准名称
1.	美国标准	NIST 1500-4	NIST 大数据互操作框架：第 4 册 安全与隐私

1　　NIST 1500-4《NIST 大数据互操作框架：第 4 册 安全与隐私》

该标准聚焦于提出、分析和解决大数据特有的安全与隐私保护问题。在理解和执行安全与隐私保护要求上，大数据触发了需求模式的根本转变，从而满足大数据的体量大、种类多、速度快和易变化的特点。基础架构的安全解决方案目标也发生了变化，例如，分布式计算系统和非关系型数据存储的安全。大数据环境下新的安全问题需要解决，其中包括平衡隐私与实用性，对加密数据开展分析和治理，以及核查认证用户和匿名用户。该标准分析了特定应用场景（包括医疗、政府、零售、航空等）下的大数据安全与隐私保护问题，提出了大数据安全与隐私保护的主要概念和角色，开发了一个大数据安全与隐私保护参考架构来补充 NIST 大数据参考架构（NBDRA），并对行业应用案例和 NBDRA 之间的映射进行了相关探索。

第 5 章 大数据安全标准体系

本章首先汇总分析大数据安全标准化需求，划分出大数据安全标准的类别，然后给出大数据安全标准图谱，并介绍已有大数据安全标准研制工作情况，最后指出需要优先开展的重点方向，为大数据安全标准化的后续工作提供基础和指导。

5.1 大数据安全标准化需求

大数据安全标准是应对大数据安全需求的重要抓手。为贯彻《国家标准化体系建设发展规划（2016-2020 年）》中“需求引领、系统布局”的基本原则，本节基于对大数据安全挑战的综合分析，结合当前大数据技术和应用的发展现状，以及当前我国对大数据安全合规方面的要求，归纳出六个方面的大数据安全标准化需求。

一、规范大数据安全相关术语和框架

当前，大数据技术和应用在快速变化之中，人们对大数据安全概念和术语的认知水平不同，包括大数据安全定义、大数据安全角色、大数据生命周期等，所有这些都将影响大数据行业的快速和健康发展。然而，目前缺乏一个通用的参考框架，能够清晰描述大数据生态中各安全角色之间关系以及各角色安全活动，用以指导后续大数据安全标准的制定。因此，应优先制定包括大数据安全概念、角色、模型和框架等基础标准，为其它标准的制定打好坚实基础。

二、推动大数据平台安全建设及相关技术应用

大数据平台和应用是支撑数据采集、传输、存储、处理、交换、销毁等数据活动的分布式信息系统，它包括底层的基础平台和上层的大数据应用。大数据平台和应用的安全建设和安全运维对整个大数据系统的安全产生重要影响。当前我国缺乏针对大数据基础平台和上层大数据应用的安全规范和指南来覆盖管理、工程、技术、平台系统和应用服务等各个方面，以指导大数据系统所有者、建设者、运营者对大数据平台和应用的安全规划、建设、安全运维和安全风险管理。因此，需要制定包括大数据平台安全运维、系统安全以及相关安全技术机制等标准，为大数据平台安全建设提供标准支撑。

三、保障数据生命周期的安全管理

数据安全是大数据安全的核心之一。数据是大数据系统中的重要资源，其安全性至关重要。当前我国缺乏针对大数据环境下的数据安全管理规范，尤其在个人信息安全、重要数据安全以及数据跨境安全等方面，需要制定规范大数据系统中的数据安全管理活动、流程和方法的安全标准，以指导数据生命周期管理活动，包括数据收集、传输、存储、共享、处理、销毁等安全活动，减少来自组织内部和外部的各种大数据安全风险。因此，应针对个人信息、重要数据及跨境数据等优先制定相关的大数据安全标准，为数据在生命周期各个环节提供安全保障。

四、支撑大数据服务的安全管理

大数据服务可以为大数据生态中的数据提供者和数据消费者提供数据分析处理、数据交易等服务。在提供大数据服务的过程中，大数据服务组织的安全

能力至关重要，直接影响到数据的安全。当前，我国缺乏指导建立大数据服务安全能力的规范，以及对大数据服务组织的安全能力成熟度进行评级的标准规范。在大数据交换和共享方面，我国大数据交易服务安全也面临没有标准规范的局面。因此，亟需制定大数据服务安全能力、交换共享安全、数据安全治理等相关标准，以规范大数据服务组织的基础安全能力、数据安全管理能力，对组织安全能力成熟度进行有效评价，支撑《网络安全法》在大数据交易等领域的落地实施。

五、促进大数据应用的安全和持续发展

不同行业和领域的大数据应用具有不同特点，所涉及的数据敏感度因政策环境、行业环境不同存在差异。因此，需要制定相应的行业大数据安全标准，以解决数据在行业之间或组织之间的交换与共享问题，支撑行业大数据应用的快速发展；同时，需要对电子政务、电子商务、电信、健康医疗等重点行业大数据应用适时出台相应的大数据安全指南类标准，指导各行业的大数据安全建设和运营。

六、促进安全应用大数据的健康发展

安全大数据是大数据的一个重要应用。随着《网络安全法》的正式实施，利用大数据解决安全问题成为了网络安全行业的一个热门课题。目前，安全行业中的大数据分析应用缺乏明确的技术规范，传统的数据分析方法往往被包装成安全大数据应用；缺乏针对性的安全标准，大多数安全应用系统本身就存在严重的脆弱性和漏洞；缺乏相关的评估标准，基于大数据的安全分析产品效果参差不齐。因此，需要针对安全大数据应用特点，制定包括技术要求、实施指南、检测评估等大数据安全标准，为安全大数据应用的健康发展提供保障。

5.2 大数据安全标准分类

大数据安全标准可基于标准主题、标准类型等进行初步的划分。具体描述如下：

5.2.1 标准主题分类

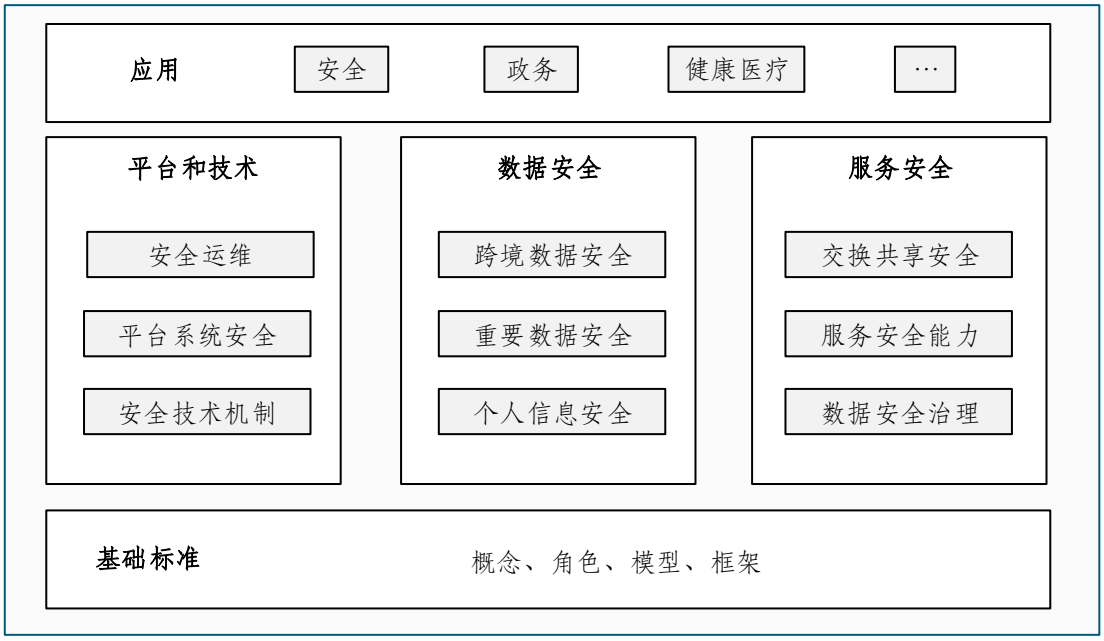

图 5.1 大数据安全标准主题分类图

图 5.1 展示了大数据安全标准主题分类图，核心是数据安全，围绕数据安全，需要技术、系统、平台方面的安全标准以及业务、服务、管理方面的安全标准支撑，分别纳入了平台和技术类标准以及服务安全类标准。

5.2.1.1 基础标准类

大数据基础类安全标准为整个大数据安全标准体系提供包括概念、角色、模型、框架等基础标准，明确大数据生态中各类安全角色及相关的安全活动或功能定义，为其它类别标准的制定奠定基础。

一、概念

大数据安全技术术语相关标准是在大数据安全方面进行技术交流的基础语言，规范术语定义和术语之间的关系，有助于准确理解和表达技术内容，方便技术交流和研究。

二、角色

大数据安全相关角色定义了大数据安全涉及到的各类人员以及相应的活动，便于明确各类人员在大数据安全中承担的责任，以及针对各类活动进行安全方面的考虑。

三、模型

大数据安全模型用于理解大数据安全，表达大数据安全相关的概念以及概念之间的关系。

四、框架

大数据安全参考架构相关标准是对大数据安全内在的要求、设计结构和运行建立的一个开放的大数据安全技术模型，规范大数据安全体系架构有助于准确理解大数据安全保障包含的结构层次、功能要素及其关系，是大数据安全其他标准制定参考的基础。大数据安全参考架构通过借鉴国际上现有的研究成果，针对大数据安全的需求，给出大数据安全参考模型，并作为对大数据参考模型

的重要补充，给出大数据安全参考架构的结构层次和功能要素，以及各结构层次和功能要素之间的关系。

5.2.1.2 平台和技术类

本类标准主要针对大数据服务所依托的大数据基础平台、业务应用平台及其安全防护技术、平台安全运行维护技术展开，具体包括安全技术与机制、系统平台安全和安全运维三部分。

一、安全技术机制

本类标准主要涉及大数据安全相关的技术、机制方面的标准，包括分布式安全计算、安全存储、数据溯源、密钥服务、细粒度审计等技术和机制。通过这些技术、机制的标准化工作，有利于经过实践检验的技术、机制的推广应用，从而整体提升大数据安全水平。

二、平台系统安全

本类标准主要涉及大数据平台系统建设和交付相关的安全标准，为大数据安全运行提供基础保障。主要包括基础设施、网络系统、数据采集、数据存储、数据处理等多层次的安全技术防护。

三、安全运维

本类标准主要涉及大数据安全运行相关的安全标准，针对大数据运行过程中可能发生的各种事件和风险做好事前、事中、事后的安全保障。包括大数据系统运行维护过程中的风险管理、系统测评等技术标准等。

5.2.1.3 数据安全类

本类标准主要包括个人信息、重要数据、数据跨境安全等安全管理与技术标准，覆盖数据生命周期的数据安全，包括分类分级、去标识化、数据跨境、风险评估等内容。

一、个人信息安全

本类标准主要涉及针对个人信息处理活动应遵循的原则和安全要求、个人信息安全影响评估等标准内容，用以健全个人信息安全标准体系，指导组织内部建立个人信息保护策略，指导产品、服务、内部信息系统的设计、开发和实现，并指导个人信息保护实践，为《网络安全法》的实践落地提供技术支撑，切实保护个人信息。

二、重要数据安全

本类标准主要围绕重要数据的生命周期，从重要数据治理、管理、技术、基础保障、安全评价等全方位、细粒度的制定对应的重要数据安全标准，用以指导重要数据的管理和保护，并为《网络安全法》的实践落地提供技术支撑。

三、跨境数据安全

本类标准旨在规范指导跨境数据处理。包括为国家开展数据出境安全评估提供技术标准支撑，为企业开展数据出境安全风险自评估提供规范指南。通过制定相关标准，使企业可以按照规定的安全评估流程、评估要点、评估方法等内容，合理有效地开展数据出境安全评估，同时为行业主管或监管部门对本行业（领域）数据出境安全评估指导、监督等工作提供依据。

5.2.1.4 服务安全类

本类标准主要是针对开展大数据服务过程中的活动、角色与职责、系统和应用服务等要素提出相应的服务安全类标准；针对数据交易、开放共享等应用

场景，提出交易服务安全类标准，包括大数据交易服务安全要求、实施指南及评估方法等。

一、数据安全治理

为了保护数据应用过程中所涉及的相关数据的安全，在数据安全治理方面需要开展各项工作，包括但不限于识别数据的敏感性并进行分类分级管理等，适用于企业等在数据安全分类分级管理的技术指导等，帮助解决数据违规收集、数据开放与隐私保护相矛盾及粗放式“一刀切”的问题，实现大数据应用、权益及安全的有效平衡。

二、服务安全能力

本类标准针对大数据服务过程及支撑系统安全提出规范要求，为大数据服务提供者的组织能力建设、数据业务服务安全管理、大数据平台安全建设和大数据安全运营等规范提出安全能力要求。一方面可以为大数据服务提供者提升大数据服务安全能力提供指导，一方面则为第三方机构对大数据服务安全测评提供依据。

三、交换共享安全

本类标准用以规范数据交换共享过程的安全性和规范性，保护个人信息安全不受侵犯、企业利益不受损害等，保证数据交易服务产业的健康规范发展，促进政府、企业、社会资源的融合运用，支撑行业应用和服务创新，提升经济社会运行效率等。

5.2.1.5 应用安全类

本类标准主要是针对重要行业和领域大数据应用，对涉及国家安全、国计民生、公共利益的大数据应用的安全防护，形成面向重要行业和领域的大数据安全指南，指导相关的大数据安全规划、建设和运营工作。本类标准主要包括以下两类：

一、安全应用

主要围绕利用大数据保障网络空间安全开展相应的标准化工作，引导相关安全技术、产品及安全产业的健康发展。

二、领域应用安全

针对不同的应用领域，围绕领域大数据应用的特点，对特殊性细化或适配相应的通用大数据安全标准，统筹考虑数据在行业之间或者组织之间的交换、共享等问题，支撑领域大数据的快速发展，指导领域大数据安全建设和运营等。

5.2.2 标准类型分类

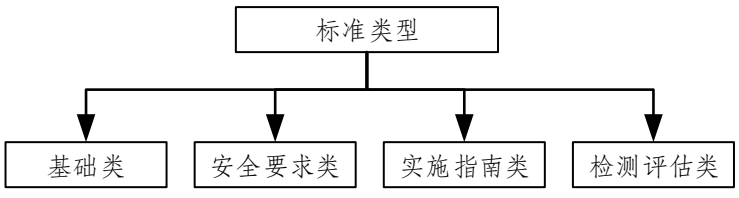

图 5.2 大数据安全标准类型分类图

大数据安全标准按类型可划分为：基础类、安全要求类、实施指南类和检测评估类。其中基础类标准旨在提供基础性的符号、术语、模型、框架等；安全要求类标准主要衔接上位法律法规，围绕大数据安全提出更具体明确的要求；

实施指南类标准主要围绕安全要求的落实，基于最佳实践，给出具体的实施指导；检测评估类标准主要围绕具体的实施是否满足安全要求展开。

5.2.3 其它分类

除了上文提到的标准主题和类型两个分类维度外，还存在诸多其它分类维度，比如表 5.1 列出了一些常见的其它分类。这些分类还需要在后续工作中不断丰富完善使用。

表 5.1 其它分类

其它分类维度	分类内容
数据状态	静态、流动、在用
数据权利主体	数据所有者、数据控制者、数据处理者、数据消费者
数据生命周期	采集、传输、存储、处理（包括计算、分析、可视化等）、交换、销毁
大数据系统角色	数据提供者、大数据框架提供者、大数据应用提供者、系统协调者、数据消费者
使用服务生命周期	准备、选择、部署、使用、变更、终止
应用领域	通用、分行业
……	……

5.3 大数据安全标准图谱

基于上述标准分类维度，可构建大数据安全标准分类体系。图 5.3 是基于大数据安全标准主题分类和大数据安全标准类型分类这两个维度构建的一种图谱，其中每个单元格可基于其它分类维度进一步细分。图 5.3 标示了大数据安全标准特别工作组截止到目前为止开展的所有标准研制项目。在今后的大数据安全标准化工作中，可依据图 5.3 基于大数据安全标准对象和标准类型进行拟研制大数据安全标准定位，明晰与单元格内已有标准的关系，明确与其它标准的区别联系之后再行研制。

需要特别说明的是，各单元格内已有标准并不仅限于大数据安全标准特别工作组范围内的标准，需要对拟研制标准主题相关的其它已有标准进行系统分析梳理，尽可能复用，针对差异点或者新的标准化点开展标准研制工作。可以看出，大数据安全标准需要尽可能的发挥框架性作用，将已有网络空间安全相关标准有机组织使用，保障大数据安全。

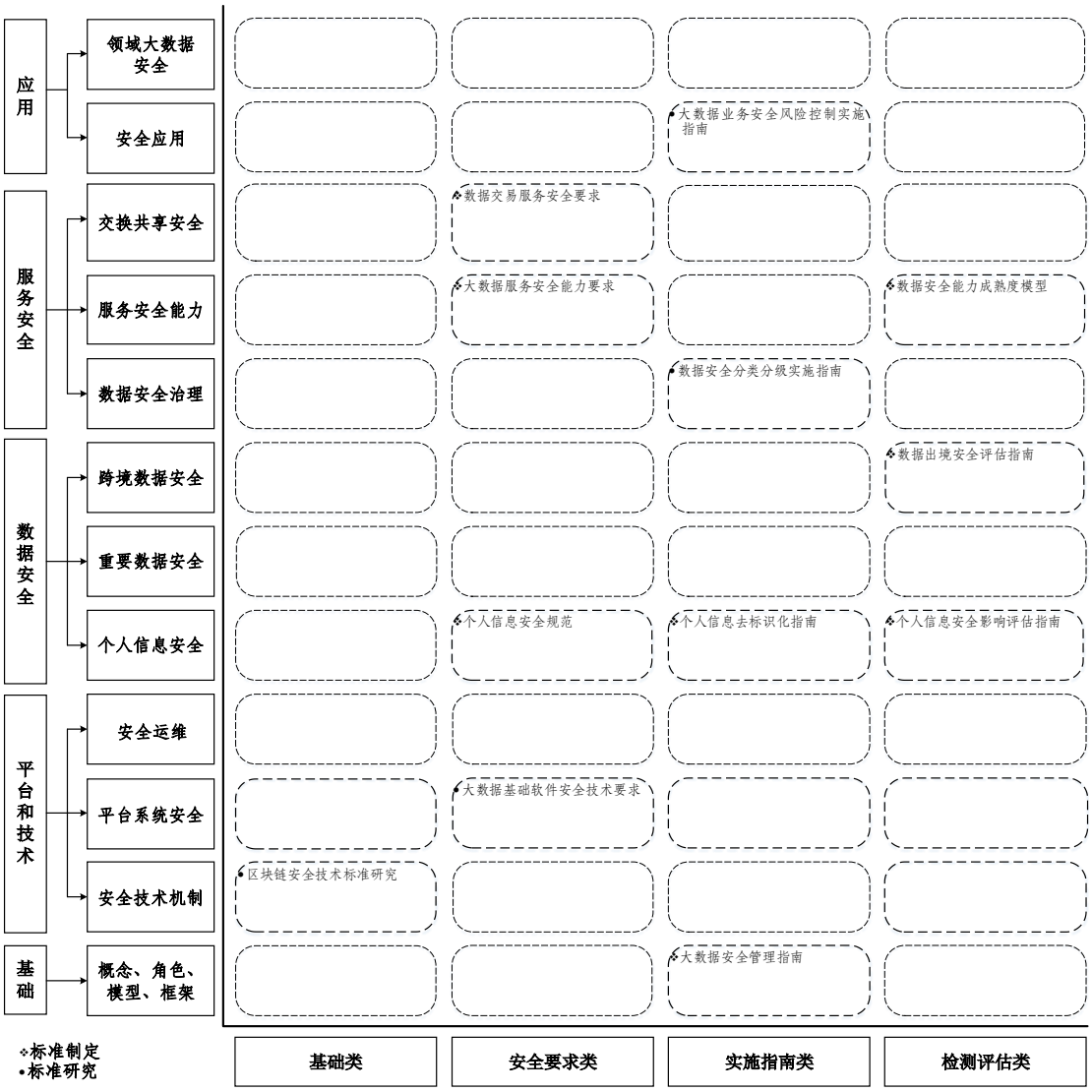

图 5.3 大数据安全标准图谱

5.4 大数据安全标准特别工作组标准工作

自 2016 年 4 月大数据安全标准特别工作组成立，已经启动了 8 项大数据安全标准制定项目和 6 项大数据安全标准研究项目。这些标准研制工作状态及内容范围分别介绍如下。

5.4.1 标准制定项目

大数据安全标准特别工作组已启动的大数据安全国家标准制定项目如表 5.2 所示。

表 5.2 大数据安全国家标准制定项目

序号	标准名称	启动时间	标准状态
1.	信息安全技术　个人信息安全规范	2016 年	已发布
2.	信息安全技术　大数据服务安全能力要求	2016 年	已发布
3.	信息安全技术　大数据安全管理指南	2016 年	报批稿

序号	标准名称	启动时间	标准状态
4.	信息安全技术　个人信息安全影响评估指南	2017 年	草案
5.	信息安全技术　个人信息去标识化指南	2017 年	报批稿
6.	信息安全技术　数据安全能力成熟度模型	2017 年	报批稿
7.	信息安全技术　数据交易服务安全要求	2017 年	报批稿
8.	信息安全技术　数据出境安全评估指南	2017 年	征求意见稿

1.　《信息安全技术　个人信息安全规范》

随着信息技术的快速发展和互联网应用的普及，越来越多的组织大量收集、使用个人信息，给人们生活带来便利的同时，也出现了对个人信息的非法收集、滥用、泄露等问题，个人信息安全面临严重威胁。

该标准针对个人信息面临的安全问题，规范个人信息控制者在收集、保存、使用、共享、转让、公开披露等信息处理环节中的相关行为，旨在遏制个人信息非法收集、滥用、泄露等乱象，最大程度地保障个人的合法权益和社会公共利益。

该标准规范了开展收集、保存、使用、共享、转让、公开披露等个人信息处理活动应遵循的原则和安全要求。适用于规范各类组织个人信息处理活动，也适用于主管监管部门、第三方评估机构等组织对个人信息处理活动进行监督、管理和评估。对标准中的具体事项，法律法规另有规定的，应遵照其规定执行。

按照国家标准化管理委员会 2017 年第 32 号中国国家标准公告，该标准已于 2017 年 12 月 29 日正式发布，标准号为 GB/T 35273-2017。

2.　《信息安全技术　大数据服务安全能力要求》

数据服务是针对数量巨大、种类多样、流动速度快、特征多变等特性的数据集，通过底层可伸缩的大数据平台和上层多种大数据应用，提供覆盖数据生命周期相关数据活动的一种网络信息服务。大数据服务提供者要确保大数据平台与应用安全可靠地运行，满足保密性、完整性、可用性等大数据服务安全目标。该标准规定了大数据服务提供者应具有的组织相关的基础安全要求和数据生命周期相关的数据服务安全要求。

该标准将大数据服务安全能力分为一般要求和增强要求两个级别。一般要求是大数据服务提供者在开展大数据服务时应具备的安全能力，能够抵御或应对常见的威胁，能将大数据服务受到破坏后的损失控制在有限的范围和程度内，具备基本的事件追溯能力。增强要求是在大数据服务涉及国家安全，或对经济发展和社会公共利益有较大影响时，大数据服务提供者应具备的安全能力，即具备一定的主动识别并防范潜在攻击的能力，能高效应对安全事件并将其损失控制在较小范围内，能保证安全事件追溯的有效性、大数据服务的可靠性、可扩展性和可伸缩性。根据所承载数据的重要性和大数据服务不能正常提供服务或遭受到破坏时可能造成的影响范围和严重程度，大数据服务提供者应具备的安全能力也各不相同。

该标准规定了大数据服务提供者应具有的组织相关基础安全能力和数据生命周期相关的数据服务安全能力。可为政府部门、企事业单位等组织机构的大数据服务安全能力建设提供参考，也适用于第三方机构对大数据服务提供者的大数据服务安全能力进行审查和评估。

按照国家标准化管理委员会 2017 年第 32 号中国国家标准公告，该标准已于 2017 年 12 月 29 日正式发布，标准号为 GB/T 35274-2017。

3.　《信息安全技术　大数据安全管理指南》

大数据技术的发展和应用影响着国家的治理模式、企业的决策架构、商业

的业务模式以及个人的生活方式。我国大数据仍处于起步发展阶段，各地发展大数据积极性高，行业应用得到快速推广，市场规模迅速扩大。在面向大量用户的应用和服务中，数据采集者希望能获得更多的信息，以提供更加丰富、高效的个性化服务。随着数据的聚集和应用，数据价值不断提升。而伴随大量数据集中，新技术不断涌现和应用，使数据面临新的安全风险，大数据安全受到高度重视。

目前拥有大量数据的组织的管理和技术水平参差不齐，有不少组织缺乏技术、运维等方面的专业安全人员，容易因数据平台和计算平台的脆弱性遭受网络攻击，导致数据泄露。在大数据的生命周期中，将有不同的组织对数据做出不同的操作，关键是要加强掌握数据的组织的技术和管理能力的建设，加强数据采集、存储、处理、分发等环节的技术和管理措施，使组织从管理和技术上有效保护数据，使数据的安全风险可控。

该标准为组织的大数据安全管理提供指导，提出了大数据安全管理基本原则，从大数据安全需求、数据分类分级、大数据活动的安全要求、评估大数据安全风险等方面，指导组织针对大数据的特点开展数据保护的管理工作。该标准适用于所有的组织，包括企业、事业单位、政府部门等，也适用于第三方机构对组织的数据安全管理能力进行评估。

4. 《信息安全技术　个人信息安全影响评估指南》

该标准是《信息安全技术　个人信息安全规范》的配套标准，将借鉴美、欧等国家和地区在个人信息安全风险评估（国际上习惯称为隐私影响评估（PIA））方面最新的法律规定、制度设计、实践做法，以国内现有立法、行政法规、标准要求为出发点，提出科学有效、符合信息化发展需要、具有明确实施指导意义的个人信息安全风险评估指南。指南将针对机构、企业提出个人信息安全风险评估的基本框架、方法和流程，供其自评估使用，同时为国家主管部门、第三方测评机构等开展个人信息安全监管、检查、风险评估等工作提供的指导和依据。该标准规定了个人信息安全风险评估的基本概念、框架、方法和流程。

5. 《信息安全技术　个人信息去标识化指南》

在大数据、云计算、万物互联的时代，基于数据的应用日益广泛，同时也带来了巨大的个人信息安全问题。为了保护个人信息安全，同时促进数据的共享使用，特制定个人信息去标识化指南标准。

该标准旨在借鉴国内外个人信息去标识化的最新研究理论，提炼业内当前通行的最佳实践，研究个人信息去标识化的目标、原则、技术、模型、过程和组织措施，提出能科学有效地抵御安全风险、符合信息化发展需要的个人信息去标识化指南。

该标准关注的待去标识化的数据集是微数据（以记录集合表示的数据集，逻辑上可通过表格形式表示）。去标识化不仅仅是对数据集中的直接标识符、准标识符进行删除或变换，而且应当结合后期应用场景考虑数据集被重标识的风险，进而选择恰当的去标识化模型和技术措施，并实施合适的效果评估。

对于不是微数据的数据集，可以转化为微数据进行处理，也可以参照该标准的目标、原则和方法进行处理。比如针对表格数据，如果关于同一个人的记录有多条，则可将多条记录拼接成一条，从而形成微数据，其中同一个人的记录只有一条。

该标准描述了个人信息去标识化的目标和原则，提出了去标识化过程和管

理措施。针对微数据提供具体的个人信息去标识化指导，适用于个人信息处理相关方，也适用于网络安全相关主管部门、第三方评估机构等组织开展个人信息安全监督管理、评估等工作。

6. 《信息安全技术 数据安全能力成熟度模型》

随着互联网、物联网、云计算等技术的快速发展，以及智能终端、网络社会、数字地球等信息体的普及和建设，全球数据量出现爆炸式增长，形成了大数据环境。伴随着大数据技术的发展和普及，组织机构在业务发展、企业运营等关键环节利用大数据技术对业务进行优化以发掘出更多的数据价值。在组织的内部管理运营过程中，组织机构利用大数据技术使能业务的发展和组织的运营，极大程度改变了其传统工作模式和业务发展方向，同时，对组织机构的数据安全管理带来了新的挑战。数据的高速流通性让组织机构内部信息系统、网络区域之间的边界越发模糊；而在大数据技术的广泛应用中，大数据的特性如大容量、多种类和可变性都对组织机构的数据管理能力提出了更高的要求。

组织机构除了关注自身业务中产生的数据之外，也开始采集外部第三方组织或人员的数据来丰富自己的数据资源，数据在不同组织机构间的流通和处理成为不可避免的趋势。各组织机构在大数据产业中提供或获取各种数据服务，成为数据源提供者、数据计算平台提供者、数据服务或应用提供者等大数据产业相关的角色。同时数据作为组织机构的重要资产，一方面面临着传统环境中数据安全的相关风险，另一方面也面临着大数据环境下所特有的数据安全风险。数据安全成为了当前产业环境下各类组织机构共同关注的安全命题。

数据安全的管理需要基于以数据为中心的管理思路，从组织机构业务范围内的数据生命周期的角度出发，结合组织机构各类数据业务发展后所体现出来的安全需求，开展数据安全保障。数据安全能力成熟度模型（以下简称“模型”）关注于组织机构开展数据安全工作时应具备的数据安全能力，定义数据安全保障的模型框架和方法论，提出对组织机构的数据安全能力成熟度的分级评估方法，来衡量组织机构的数据安全能力，促进组织机构了解并提升自身的数据安全水平，促进数据在组织机构之间的交换与共享，发挥数据的价值。

该标准基于大数据环境下电子化数据在组织机构业务场景中的数据生命周期，从组织建设、制度流程、技术工具以及人员能力四个方面构建了数据安全过程的规范性数据安全保障能力的成熟度分级模型及其评估方法。适用于组织机构数据安全能力的自身评估，也适用于第三方机构对组织机构的数据安全保障能力进行评估。

7. 《信息安全技术 数据交易服务安全要求》

数据正日益对全球生产、流通、分配、消费活动以及经济运行机制、社会生活方式和国家治理能力产生重要影响。数据交易可以促进数据资源流通，破除数据孤岛，有效支撑数据应用的快速发展，发挥数据资源的经济价值。然而，数据交易面临诸多安全问题和挑战，影响了数据应用的进一步健康发展。

为规范数据资源交易行为，建立良好的数据交易秩序，促进数据交易服务参与者安全保障能力提升，该标准将对数据交易服务进行安全规范，增强对数据交易服务的安全管控能力，在确保数据安全的前提下，促进数据资源自由流通，从而带动整个数据产业的安全、健康、快速发展。

该标准规定了通过数据交易服务机构进行的数据交易服务所涉及的交易参与方、交易对象和交易过程的安全要求。适用于提供数据交易服务的机构进行安全自评估，也适用于第三方机构对数据交易服务机构进行安全评估。

8. 《信息安全技术　数据出境安全评估指南》

随着互联网的蓬勃发展，数据流动无处不在。数据流动所产生的经济价值和社会价值凸显，这一过程中的安全风险也随之增加，国家安全、经济发展、社会公共利益、个人信息安全受到严重威胁，防范数据泄露和滥用所产生的风险日益紧迫。

该标准规定了数据出境安全评估流程、评估要点、评估方法等内容，国家网信部门、行业主管部门以及网络运营者按照本指南对其向境外提供的个人信息和重要数据进行主管部门评估和安全自评估，发现存在的安全问题和风险，及时采取措施，确保个人信息和重要数据合法流动的同时，避免其对国家安全、经济发展、社会公共利益和个人信息主体权益造成不利影响。

5.4.2 标准研究项目

大数据安全标准特别工作组已启动的大数据安全国家标准研究项目如表 5.3 所示。

表 5.3 大数据安全国家标准研究项目

序号	项目名称	启动时间	状态
1.	大数据交易服务平台安全要求	2016 年	已完成
2.	大数据安全能力成熟度评估模型	2016 年	已完成
3.	大数据基础软件安全技术要求	2017 年	已完成
4.	大数据业务安全风险控制实施指南	2017 年	已完成
5.	数据安全分类分级实施指南	2017 年	已完成
6.	区块链安全技术标准研究	2017 年	研究中

1. 《大数据交易服务平台安全要求》

该研究项目顺利结题，并在 2017 年通过申请成为标准制定项目，其详细内容见 5.4.1 节《数据交易服务安全要求》内容。

2. 《大数据安全能力成熟度评估模型》

该研究项目顺利结题，并在 2017 年通过申请成为标准制定项目，其详细内容见 5.4.1 节《数据安全能力成熟度模型》内容。

3. 《大数据基础软件安全技术要求》

国家“十三五”规划明确指出，实施国家大数据战略。经过几年发展，大数据在政务、金融、交通、互联网等诸多领域发展迅速，为社会带来巨大的价值。与此同时，安全问题已成为制约大数据平台建设部署及业务发展的重要阻碍。包括：（1）大数据平台使用开源软件，这些软件设计初衷是为了高效数据处理，但是在安全功能方面缺乏整体规划，安全防护能力较差，存在安全风险；（2）大数据平台需要汇集多源数据，包括用户敏感数据，大规模数据的集中管理也带来风险的积聚效应；（3）数据开放是大数据业务发展的重要方向，在数据开放、共享的过程中必然存在用户隐私泄露等关键问题。这些问题对大数据安全管理和技术提出了更高的要求，尤其是大数据基础软件作为构建大数据平台、承载大数据业务的重要基础，关系到整个大数据应用的安全。大数据基础软件是指包括负责完成大数据平台中数据的传输交换、存储管理软件、计算框架以及一系列通用软件的组合。因此，需要对业界主流的大数据基础软件开展研究，识别大数据基础软件特有安全风险，并提出共性的安全技术要求。

该项目参照国际、国内大数据安全相关法规、政策、标准、技术和管理等研究成果，明确大数据基础软件的范畴，调研业界主流的大数据基础软件，分

析识别大数据基础软件所面临的安全风险，研究设计大数据基础软件安全技术总体架构，提出具体的安全防护技术要求，形成完整的研究报告和标准草案，并通过本项目研究，为大数据基础软件安全技术要求标准制定做好技术研究和前期准备。

4. 《大数据业务安全风险控制实施指南》

互联网行业不仅在业务过程中产生大量的数据，并且在利用大数据开展多种业务，互联网环境下的大数据业务安全风险控制越来越广泛。企业在利用大数据开展各种业务过程中，需要进行恶意用户识别、恶意行为监控等。但是目前没有标准参照，无法达到规范一致的水平和效果。

互联网行业典型的业务风险，包括但不限于：网络金融领域贷款环节面临着巨大的骗贷风险，电商、O2O 领域羊毛党盗刷，社交领域和直播领域传播色情、政治反动、暴恐等违法内容，内容载体可以是文字、图片、URL、视频，这些会对国家带来很大的安全威胁，并且也会影响到企业的正常运营。

然而不少企业还停留在人工+简单规则+简单计算系统的对抗方式上，人力消耗大，效果不理想，亟需进行能力升级。例如，一般的网络金融企业会采取让用户多上传个人资料，进行电话、实地考察的方式来控制风险，然而这会极大地降低企业的放贷效率，并且面对灵活的黑产团伙，一些传统的验证方式会快速失效。

2017 年 6 月开始正式实施的《网络安全法》第十条：建设、运营网络或者通过网络提供服务，应当依照法律、行政法规的规定和国家标准的强制性要求，采取技术措施和其他必要措施，保障网络安全、稳定运行，有效应对网络安全事件，防范网络违法犯罪活动。

随着大数据技术的应用逐渐成熟，能对上述互联网业务层面的安全风险起到良好的控制作用。该标准研究项目通过广泛的调查研究，提出《大数据业务安全风险控制实施指南》研究报告，建立大数据风控模型，识别恶意的文本、视频、音频、图片、URL、用户行为等；探讨应用于直播、金融、电商、O2O、社交等互联网业务领域；为制定标准做好准备。

5. 《数据安全分类分级实施指南》

随着国家大数据发展战略的实施，“互联网+”行动的深入推进，大数据资源价值不断提升，电信、互联网、金融、政务、交通等各领域相关的大数据应用也在蓬勃发展。这些大数据应用涉及的数据量大、种类多，同时又包含有很多用户相关重要数据。而大数据应用在不断发展创新的同时，由于数据违规收集、数据开放与隐私保护相矛盾以及粗放式“一刀切”管理方式等给大数据应用的发展带来严峻的安全挑战。大数据资源的过度保护不利于大数据应用的健康发展，数据分类分级的安全管控方式能够避免“一刀切”带来的问题，实现大数据应用与个人权益的有效平衡。

该研究是对目前已有标准的进一步细化和实施参考。对现有标准《大数据服务安全能力要求》中关于数据分类分级要求的进一步落地实施。也是对现有标准《大数据安全管理指南》中关于个人数据分类分级安全管理的要求进一步落地实施。

6. 《区块链安全技术标准研究》

随着互联网和信息技术的发展，电子商务和网络金融产品不断创新，出现了各种基于区块链技术的应用，比如虚拟电子货币。虚拟电子货币在电子商务环境下正发挥着日益重要的作用，虚拟电子货币的流通对传统货币应用方式产

生了巨大的影响，尽管带来了非常好的用户体验和便捷性，但对于那些受众范围广的虚拟电子货币应用，时刻面临着安全风险。对于这种基于新技术和新交易模式的虚拟电子货币，特别是区块链技术采用，目前还缺少相关技术应用的安全规范。区块链技术目前已得到了广泛应用，在虚拟电子货币这类特殊应用上，安全不当则可能造成资金损失，甚至严重冲击稳定的社会金融秩序，有待加强监管。

该项目主要基于狭义非法定电子货币（也称虚拟电子货币），分析虚拟电子货币的发行与交易模式，重点研究区块链技术的应用。通过分析这类电子货币系统的安全风险和威胁，针对性地提出区块链技术应用的安全风险和威胁，提出安全技术要求。其意义是通过规范虚拟电子货币这类区块链技术的安全应用，在可评估验证的条件下，确保采用区块链技术实现的系统安全可靠，从而促进社会金融和网络安全环境的健康发展。

5.5 近期重点工作方向

根据国家标准委 2018 年全国标准化工作要点和全国信息安全标准化技术委员会 2018 年工作要点，对照大数据安全标准化需求和已有大数据安全标准研制工作情况，近期需要优先开展下述相关工作。

5.5.1 开展大数据安全参考框架研制

加强大数据环境下的网络安全问题研究和基于大数据的网络安全技术研究，明确数据采集、传输、存储、处理、交换、销毁等各环节保障网络安全的范围边界、责任主体和具体要求。基于我国大数据技术框架相关标准，明确大数据安全相关要素以及各要素之间的关系，包括大数据角色、角色安全职责、安全功能组件以及它们之间的关系，形成大数据安全参考框架。

5.5.2 完善个人信息安全相关标准研制

近年来，随着互联网与各种行业的广泛融合，越来越多的信息系统运营应用会采集个人信息，并对个人信息进行存储、处理，甚至交易。个人信息的非法收集、泄露、滥用等已成为社会关注的焦点问题，个人权益侵害情况也屡见不鲜。为有效支撑《网络安全法》的落地实施，配合已有个人信息安全标准工作，制定配套的安全保护技术等支撑标准，在保障用户合法权益和维护社会公共利益的同时，最大程度的促进数据自由流通，挖掘数据价值。

5.5.3 推进数据交换共享相关安全标准研制

数据作为一种战略性基础资源，在交换共享过程中将会产生更大价值。但数据在交换共享过程中由于缺乏必要的安全技术或管理能力，会暴露更多安全问题，比如地下数据交易黑灰产业激增，凸显侵犯用户个人信息安全、数据滥用等问题，这些问题已经成为数据自由流通的严重障碍。建立健全数据交换共享相关安全管理办法，加快数据交易安全相关标准的制定，规范数据交易市场，从数据交易主体、交易对象、交易过程等方面规范数据交易服务，加强对大数

据交易服务提供商的监管；有效解决数据共享中的各种安全问题，为数据流通过程提供有效的安全支撑环境，保障数据供应链相关方的合法权益，促进大数据产业的安全和健康发展。

5.5.4 加快数据出境安全相关标准研制

数据流动产生价值，但无序的数据跨境流动则可能危害到个人安全、社会稳定甚至国家安全。为保障国家安全、公共利益和公民权益，《网络安全法》明确规定，关键信息基础设施运营者因业务需要，向境外提供在中国境内运营中收集和产生的个人信息和重要数据的，需按照国家网信部门会同国务院有关部门制定的办法进行安全评估。为有效落实实施《网络安全法》有关规定，支撑数据出境安全评估工作实施，亟需制定数据出境的相关安全标准。明确研究数据出境安全评估的主要风险指标、数据属性特征指标，判断出境数据的重要性，设计数据出境活动评估指标，综合评判出境活动的风险性。研制数据出境安全相关标准，能够为国家开展数据出境安全评估工作机制和有关制度落地提供标准支撑，为企业开展数据跨境安全风险自评估提供指导。

5.5.5 推动大数据安全检测评估相关标准研制

大数据安全关系到国家安全、社会稳定和个人权益，依法对大数据应用进行安全审查和检测评估是保障我国国家安全、公共利益和公民权益的一个重要手段。大数据产业发展需要大数据安全检测评估相关标准支撑。针对当前缺乏大数据安全检测评估标准的问题，需要通过对大数据安全现状、安全需求和攻防技术的深入调研，全面分析大数据可能存在的安全隐患，确定大数据安全审查和检测评估的方向和重点，梳理和细化安全审查和检测评估内容，编制可量化、可操作的大数据安全审查和检测评估技术要求和测试评价方法相关标准。同时，推动网络安全等级保护制度在大数据技术和应用领域的贯彻落实，指导数据生命周期的数据安全保护，以及涉及的等级保护对象的安全建设和监督管理，促进大数据应用的健康发展。

5.5.6 启动重点领域大数据安全标准研制

各领域大数据应用由于其独特的业务特点，在通用大数据安全标准基础上还需要有针对性的开展标准化工作。目前，政务大数据和健康医疗大数据的发展急需相应的安全标准保驾护航。

政务大数据方面：《促进大数据发展行动纲要》（以下简称《纲要》）部署了三方面主要任务，其中政府数据的开放共享被排在首位。《纲要》强调要大力推动政府部门数据共享，稳步推动公共数据资源开放，统筹规划大数据基础设施建设，支持宏观调控科学化，推动政府治理精准化，推进商事服务便捷化，促进安全保障高效化，加快民生服务普惠化。中共中央政治局2016年10月9日下午就实施网络强国战略进行第三十六次集体学习。中共中央总书记习近平在主持学习时指出要深刻认识互联网在国家管理和社会治理中的作用，以推行电子政务、建设新型智慧城市等为抓手，以数据集中和共享为途径，建设全国一体化的国家大数据中心，推进技术融合、业务融合、数据融合，实现跨层级、跨地域、跨系统、跨部门、跨业务的协同管理和服务。

健康医疗大数据方面：《国务院办公厅关于促进和规范健康医疗大数据应用发展的指导意见 国办发〔2016〕47 号》指出要推动健康医疗大数据资源共享开放，全面深化健康医疗大数据应用。习近平总书记在全国卫生与健康大会上强调把人民健康放在优先发展战略地位，努力全方位全周期保障人民健康，指出要推进健康医疗大数据应用。中共中央国务院印发的《“健康中国 2030”规划纲要》指出要推进健康医疗大数据应用，加强健康医疗大数据应用体系建设，推进基于区域人口健康信息平台的医疗健康大数据开放共享、深度挖掘和广泛应用。

第 6 章 大数据安全标准化工作建议

为应对大数据安全风险和挑战，促进大数据应用健康有序发展，结合目前大数据安全标准化工作基础，对我国今后大数据安全标准化工作建议如下。

6.1 健全大数据安全法律法规体系

加快推动数据资源开放共享和开发应用的同时，必须建立大数据安全保障体系，构筑适应大数据发展的法规制度，健全大数据时代信息安全新秩序。从政策上关注大数据战略性和基础性重点领域，加快相关法律法规的出台步伐，依法保护公民和国家的大数据安全。平衡释放数据经济活力、规范商业利用与数据资源安全和个人信息保护之间的关系，重点针对数据的收集和使用环节建立规则，明确大数据生态中不同主体的责任，促进网络基础设施的发展，开放数据资源，加强网络安全与个人信息保护。

6.2 加强大数据安全核心技术研发

大数据安全核心技术研发是我国大数据产业自主发展的关键驱动，同时也是我国大数据安全标准的制定和落地实施的重要支撑。建议加强我国大数据安全核心技术研究，包括分布式环境下的数据加密、数据完整性验证、数据标签、区块链、细粒度访问控制、密文透明运算、数据溯源、数据脱敏与安全审计等技术；同时，建议加强大数据技术在网络安全防护方面的研究，包括入侵检测、安全态势感知、网络攻击取证、威胁情报分析等，以利用大数据技术来抵御针对大数据的网络攻击威胁。

6.3 大力推广大数据安全标准示范应用

推进大数据安全标准应用，覆盖标准研制、验证和推广等标准化活动。要加快大数据安全标准在产业中应用，切实发挥大数据安全标准对产业技术发展的保障支撑作用，维持产业秩序；促进产业大数据安全标准研制与科技研发的衔接，打通最佳实践与科技成果转化渠道；建设一批典型标准试点示范工程，提升标准孵化与研制质量，增强标准与技术环境的适应能力，保障重点领域大数据安全标准实施应用效果。

6.4 建立大数据安全标准体系研究长效机制

为提升产业大数据安全保障能力，维护网络安全秩序，考虑到大数据安全问题的泛在性、复杂性、专业性，以及有关术语和标准匮乏的现状，需要建立大数据安全标准体系研究长效机制。建议立足我国国家安全管理要求和大数据有关产业发展现状，借鉴国际国外标准化工作模式和经验，逐步建立我国大数据安全标准体系研究长效机制，持续规划大数据安全标准有关术语规范、标准

化体系等方面研究，以充分发挥大数据安全标准国家质量技术基础的支撑作用，有效引导大数据安全标准化工作科学推进。

6.5 加强大数据安全标准化人才培养

建议建立健全多层次、多类型、国际化的大数据安全人才培养体系，鼓励高校、企业、测评机构等单位加强合作，在高等院校网络空间安全学科下设立大数据安全课程，在培养大数据安全标准人才的同时，加强大数据安全标准制定、宣贯、检测、评估类专业人才的培养。通过健全大数据安全标准化人才的联合培养模式，依托社会化教育资源，开展大数据安全知识普及和教育培训，提高社会整体大数据安全标准的认知水平，加快大数据安全的标准化和大数据安全标准的国际化进程。

6.6 深度参与大数据安全国际标准化工作

目前，我国在大数据产业基础和应用探索实践方面积累了一定经验，需要抓住机遇，积极参与国际标准化有关活动来提升我国的国际影响力。建议密切跟踪国际国外大数据安全标准化发展趋势和工作动态，加强大数据安全国际标准提案研究，深度参与大数据安全国际标准编制工作。加大对我国企事业单位和学术专家在大数据安全国际标准项目中担任编辑并主导编制的工作支持力度，充分发挥我国现有国际标准化交流与合作机制的优势，举办大数据安全标准化国际交流合作活动。推动大数据安全领域国际标准提案，将国内成熟的大数据安全标准转化为国际标准，贡献中国智慧，提升我国在大数据安全国际标准制定方面的国际话语权和影响力。

附录 A 典型领域大数据安全标准需求

不同领域大数据应用特点各异，所呈现的安全风险及需求也不尽相同，不免会带来新的标准化需求。本附录列出了典型领域大数据应用的特点、大数据安全风险和挑战以及大数据安全标准化需求。一方面，有助于归纳总结通用的大数据安全标准化需求；另一方面，有助于指导通用的大数据安全标准在具体大数据应用领域的应用。

A.1 安全应用大数据

A.1.1 安全应用大数据特点

随着移动网络、云和虚拟化、物联网、工控系统等技术领域的快速发展，网络空间面临的安全威胁越来越多样化，复杂的攻击路径和高级的攻击手段使得针对特定对象的网络安全防护工作正日益受到挑战。攻击威胁来源从早期的个人黑客变为犯罪团伙、政治势力、网络部队等更严密的组织。为了应对不断升级的网络安全威胁，网络安全防护能力也需要不断提升，“安全应用大数据”技术应运而生。安全应用大数据包括以下行业特点：

一、以大数据技术为基础，大数据资源为驱动。需要有持续不断的、海量异构的安全数据，才能推动产业运作和发展。

二、依赖协同机制发挥作用，包括不同数据来源的协同汇聚，海量异构数据的协同分析，云端数据平台和本地安全产品的协同工作。孤立节点很难形成足够的数据规模，也难以发挥作用。

三、大数据技术在安全生命周期中最有效的部分是在预警和检测环节，而形成的情报可以对防护和响应做出指导。

四、安全应用大数据技术需要与人工分析紧密结合，安全行业应对的主要威胁来自于人，而不是地震海啸这样的自然现象，所以任何固定模式都无法长期有效，必须由人工参与才能发挥最大作用。

A.1.2 安全应用大数据应用领域

目前安全应用大数据对于保障网络安全具有重要作用，主要应用在以下几个领域。

一、基于大数据的威胁发现

这是当前大数据在网络安全领域的应用的两个主要方面，包括宏观层面的网络安全态势感知和微观层面的高级持续威胁（APT）检测。

1. 网络安全态势感知

近年来，网络安全事件层出不穷，传统网络安全防御措施很难及时、有效的发现安全威胁。把互联网的海量安全数据作为安全要素，通过大数据技术对这些安全要素信息进行分析，可全面、精准的掌握网络安全状态，形成安全监控的闭环，才能改变当前“黑客主动攻击、企业被动防御”的局面。态势感知在网络安全领域是指广泛采集和收集区域网络中的安全状态和事件信息，并加以处理、分析和展现，从而明确当前网络的总体安全态势，为大范围的预警和响应提供决策支持的技术，具体包括海量异构数据分析、深度学习、网络综合度量指标、网络测绘、资产建模、威胁情报、知识图谱、安全可视化等。

2. 高级持续威胁（APT）检测

高级持续性威胁具有精心伪装、定点攻击、长期潜伏、持续渗透等特点，已经成为网络犯罪和间谍活动的首选攻击方式。过去要发现针对特定网络 APT 的攻击有两个难点：一是未知威胁分析过程缺少历史数据的支持，难以进行回溯关联，遗漏了很多关键信息；二是缺少外部威胁情报，只依赖于自有的黑域名/黑 IP 库，检测的精度和效率都难以满足需求。通过全面收集重要终端和服务器上的日志信息以及采集网络设备上的原始流量，利用大数据技术进行分析和挖掘，检测并还原整个 APT 攻击场景，从攻击源头进行精准定位，最终达到对高级持续威胁（APT）的检测与溯源。

3. 反恐维稳

国际恐怖组织早已开始通过网络阵地宣传自己的主张，并且通过互联网招募人员、筹集资金。此外，网络还给“志同道合者”提供了交流接触的机会。面对新媒体环境下国内恐怖活动猖獗的严峻挑战，急需大数据技术提高网络反恐斗争的效率。例如利用大数据技术结合音素分析、语图匹配等技术对网络传播的音视频进行精确阻断，利用大数据结合上下文关联分析技术整治网络谣言传播泛滥化问题。

二、基于大数据的业务风险管理

由于大数据能够从多个维度反映业务的真实状况，可以从不同侧面反映企业经营状态、经营质量、经营能力的数据，因此，进行业务风险管理（包括风险识别、风险分析与风险评价等）成为大数据的一个重要应用领域。其中一个典型应用就是金融反欺诈。

欺诈风险是消费金融业务发展中重要的风险，信息不对称是导致欺诈风险的主要原因。可以借助大数据大幅度提高欺诈检测中数据和信息收集的完整性和准确性，找到欺诈者留下的线索，以防止欺诈发生或在欺诈发生后进行调查取证。主要包括以下几个步骤：一、获取海量数据作为分析的基础，即获取基础数据，包括 UGC 样本、设备画像库、手机号、IP、每日新增数据等等，也包括关系图谱即数据之间的关联；二、数据预处理，即需要利用文本挖掘、图片挖掘等数据分析工具，对语音、文本、照片等数据的内容进行特征抽取、文本或图片的分类、聚类等预处理操作，统一数据格式，为深入分析做准备；三、基于数据分析构建动态的欺诈风险模型，包括抽取关键风险场景要素，多维关联关系分析，大数据风险聚合分析等，构建欺诈风险模型。

三、基于大数据的身份认证

虚假身份是网络应用的重要安全威胁，可以应用大数据从多个维度对身份信息进行认证，从而有效破解虚假身份难题。这方面的典型应用包括伪基站发现与追踪、反钓鱼攻击。

1. 伪基站发现与追踪

伪基站是一种和运营商的真实基站功能类似的小型或微型信号收发装置，能够获取周围的手机与基站的设备信息，通过模拟真实基站通信机制，迫使周围的手机连接到该仿冒的基站上，向普通用户发送垃圾短信，甚至冒用号码、群发诈骗信息。采用传统检测技术很难发现和追踪到伪基站。然而采用大数据技术，则可以极大提高发现伪基站的效率，并可以及时阻断诈骗短信中的钓鱼链接，打破诈骗链条。基于大数据的伪基站检测方法具体包括以下步骤：第一，通过手机用户举报垃圾短信，或者通过手机防护软件主动拦截并上报垃圾短信，大量收集伪基站短信中包含的时间、地点、内容、仿冒的基站号等各种信息；第二，在大数据处理平台运用自然语言处理与机器学习方法，去掉大量的噪声点，从海量的垃圾短信中以较高的精度提取出伪基站短信；第三，将伪基站发出的短信与经纬度信息结合，就可以发现并定位伪基站；结合伪基站的历史数据，可以进一步找到伪基站的活动规律，并以此对其运动轨迹进行预判；第四，与地理信息系统联动，展现伪基站位置、伪基站的行为、历史运行路径、数量分布等信息，从而帮助执法部门的抓捕行动。

2. 反钓鱼攻击

钓鱼攻击是一种利用社会工程学手段，伪装在线金融或交易平台网站，针对客户个人

身份数据和金融账号进行盗窃的犯罪行为。采用传统技术很难发现钓鱼网站，但可以利用大数据技术来实现对钓鱼网站的发现，比如，一方面可以利用搜索引擎扫描相关互联网址，并通过大数据建模过滤掉可信页面与重复页面，筛选出有嫌疑的钓鱼网址页面，将这些页面输入到分析引擎中；另一方面利用用户举报数据，将钓鱼网址上报到分析引擎的数据库中；最后，分析引擎通过规则模型综合研判、机器学习等方式检测出钓鱼网址和页面，并将发现的钓鱼网站和网页汇集成为网址信誉库，据此提示或阻止用户的访问行为。

四、基于大数据的真实性分析

大数据可以应用于多种场合的真实性分析，消除相关的安全威胁。典型的应用包括关系验证、反洗钱等。

1. 关系验证

在金融信贷业务中，为了防范欺诈，需要在用户授权情况下，验证申请人和所填联系人之间是否存在真实的关系，以验证个人身份的真实性和可靠性。在营销推荐、案件团伙识别等应用场合，也也需要验证相关人员之间的关系及其强弱。利用各种业务平台和信息监控平台，对相关人员的活动记录、通信记录、行为记录等大数据进行分析和关联，可以验证特定人员之间的关系是否属实以及关系的强弱，从而为相关业务提供强有力的决策支撑。

2. 反洗钱

互联网金融行业中，第三方支付减少了交易成本和风险，但也存在隐蔽转移资金、便利任意性套现、潜在跨境支付风险、隐匿现金形式的“黑钱”处置及多功能资金等洗钱风险。利用大数据技术可以实现虚假交易识别、账户控制人和开立人关系判定等，能够大大提高业务平台方反洗钱能力，降低互联网金融洗钱风险。

A.1.3 安全应用大数据标准需求

安全应用大数据标准需求主要集中在以下三个方面：

一、安全应用大数据表达类标准

为了应对 APT 攻击、零日漏洞等网络安全威胁，需要共享和协同分析威胁情报，进行主动防御和协同防御，需要制定适用不同场景的威胁情报的表达、传输和协同分析等标准。

二、安全应用大数据协同类标准

安全应用大数据需要大量协同，不同安全产品之间的接口需要标准支持才能互联，例如态势感知类系统标准、威胁情报交换标准等。

三、安全应用大数据应用类标准

将网络安全应用大数据用于身份认证、业务风险管理、真实性分析等领域时，需要制定相关应用指南标准，规范产品的研制和业务的应用工作。

A.2 政务大数据

A.2.1 政务大数据特点

政务大数据是指政府在推动大数据应用发展的过程中或大数据在公共服务领域的应用实践中产生的大数据。政务大数据是建设新型智慧城市的基础，具体应用场景包括：为政府提供智能办公、智能监管、智能服务、智能决策等大数据服务；帮助政府更好的治理城市，提高政府办公、监管、服务、决策的智能化水平等等。

为形成政务大数据，需要将各级政府部门、各单位管辖的数据资源汇集起来，实现政府数据的互联互通，并对大量的多源异构数据融合进行大数据综合分析、挖掘，从而帮助政府将现有的数据资源进行转化并创造出价值，有效提升政府管理和决策能力。因此，政务大数据具有如下特点：涉及的行业范围广泛、数据结构多样、关联关系复杂，并涉及大量个人隐私数据、国家敏感数据等。

A.2.2 政务大数据安全风险和需求

政务大数据的安全风险和挑战主要包括：

一、平台安全

大数据平台是政府使用数据资源的基础平台，确保该平台的安全是各级政府安全可靠地使用数据资源的基础。大数据平台除了面临诸如恶意代码、攻击软件套件、物理损坏与丢失等传统安全威胁外，还面临大数据平台自身的安全问题。

二、服务安全

以政务大数据服务于新型智慧城市为例，政府可通过互联网向百姓提供基于政务大数据的便民服务。但是由于该服务基于互联网，可能会面临基于 Web 的攻击、Web 应用程序攻击/注入攻击、拒绝服务攻击、网络钓鱼、用户身份盗窃等一系列威胁，导致基于政务大数据的便民服务存在一定的安全隐患，可能产生信息泄露、网络瘫痪、服务中断等安全问题。

三、数据安全

数据是核心资产，因此要充分重视数据安全。电子政务多年的运行为政府积累了海量的数据资源，在对这些数据资源进行开发利用的同时，必须重视数据自身的安全，否则将导致数据泄露等安全问题。因此，在政府部门进行数据公开、各级部门间及各部门内部数据的平台化共享过程中，数据安全是迫切需要解决的问题，也是政府大数据资源得以共享开放、规范化、平台化、相关“掘金”应用得以发展的关键。

四、数据确权

由于数据的所有权、使用权、管理权可能涉及多个部门，政府在进行大数据应用过程中，需要做到权责分明，通过厘清数据的权属关系，防止数据流通过程中的非法使用，保障数据安全流通。

五、用户安全

在政务大数据使用中，建立全网统一的用户注册中心，用以统一管理全网用户，统一注册登记、信息管理系统等；汇聚各政府部门原有平台的用户注册信息，所有用户信息自动与用户大数据库打通，便于精准推送。在这个过程中，需要建立各种商业性应用程序、用户信息存储、支付系统、用户及资金安全保障体系等。

A.2.3 政务大数据安全标准需求

在推进政务大数据共享的过程中，存在“不愿共享”、“不敢共享”、“不能共享”三个难题，表现在：1. 有些政府部门不愿意与其他部门共享本部门或本系统的管理数据，其原因主要包括三个方面：（1）出于权力本位；（2）缺乏法律约束和考核机制；（3）政府部门自身惰性和“路径依赖”；2. 有些政府部门基于风险的考虑，不敢与其他部门共享管理数据；3. 因为“信息壁垒”的存在，有些政府部门不能将管理数据及时与其他部门共享。究其原因，一方面是由于标准不统一阻滞共享，原先的政务信息系统建设中，由于缺乏标准体系的支撑，各部门采集的数据格式不统一、标准不一致，采取的处理技术和应用平台各异，数据库接口也不互通；另一方面是体制问题拖延共享，由于管理边界不清晰、责任区分不

明确，导致政府对数据资源的归属、采集、开发等相关管理规则不够明确。

因此，政务大数据领域的标准需求主要集中在以下几个方面：

一、政务大数据安全交换与共享标准

由于缺乏政务大数据交换共享标准体系的支撑，政府各部门采集的数据格式不统一、标准不一致，采取的数据处理技术和搭建的应用平台具有一定的差异性，数据库接口也不互通。因此，信息管理平台难以整合，导致数据采集、数据获取、交互交换中发生迟滞、偏差，信息资源的共享存在困难。政务大数据交换共享安全标准能有效提升各部门数据采集、获取及交换的效率。

二、政务大数据敏感信息保护标准

由于担心共享后泄露敏感信息会带来负面影响和不利后果，或者共享同时暴露出本部门原有数据不真实、不精确而引发问责，或者认为数据安全与保密比共享更重要、采取封闭行为更妥当，导致有些政府部门不愿意将管理数据与其他部门共享。在此背景下，规范政务大数据敏感信息保护标准，可以在政务大数据开放共享服务于民的同时，有效保障政府敏感信息不被泄露。

A.3 健康医疗大数据

A.3.1 健康医疗大数据特点

健康医疗大数据是指与健康医疗相关，满足大数据基本特征的数据集合，是国家重要的基础性战略资源，正快速发展成为新一代信息技术和新型健康医疗服务业态。健康医疗大数据覆盖全员人口和全生命周期，涉及国家战略安全和人民生命安全。通过对健康医疗大数据进行标准化治理、综合行业数据、深度挖掘分析，可以达到优化医疗卫生资源配置、改善医疗服务、提升管理水平、推动经济发展的效果。

当前，健康医疗大数据主要应用于公共卫生和医疗服务，包括疾病监测预防、精细化运营管理、临床科研分析、患者问诊服务、辅助新药研发、提升疑难病症诊治能力等场景，并逐步应用于临床路径优化、智能辅助诊疗、个性化精准医疗等方面。不久的将来，健康医疗大数据将在个体疾病预防、群体疾病预测、种族差异分析、全民健康生态产业等方面发挥巨大作用，为惠民、惠政、惠业、惠医做出贡献。

目前，健康医疗大数据主要应用于如下几个领域：

一、医疗机构服务

（一）比较结果研究

在海量病例中，对不同病例提取出特征值（如年龄、性别、各项生理指标、疾病严重程度等），利用大数据技术对不同疗法的疗效进行建模，对各种疗法进行比较研究，从而为病人精确地提供最有效的疗法。

（二）临床决策支持

通过全面分析既往患者的临床特征、检查检验、医嘱、疗效等数据，挖掘医学文献数据，建立医疗专家大数据，辅助医生进行临床决策。

（三）基于基因测序的个性化医疗服务

利用健康医疗大数据形成个性化医疗服务和治疗，即基于基因科学的医疗模式、个体特征和身体情况，通过对居民健康影响因素进行分析，对患者健康信息进行整合，为疾病的诊断和治疗提供更好的数据证据，进行居民健康知识库的积累，从而改进居民健康。

（四）慢性病检测和预警

通过连续性的医疗监测形成的大数据，发现常识中得不到的信息并捕捉到与慢性疾病有关的紊乱的体征波动信息。可以对这些慢性疾病进行预警及为用户提供护理建议。

（五）个人健康管理

利用大数据技术，对个人健康进行全生命周期管理。健康分析人员能够有效地对个人健康状况进行分析，以便在身体处于非健康状态时得到及时的干预。

二、医疗保险管理

（一）医疗保险决策支持

通过大数据技术，重构医保对医疗费用审核监管的全新模式，从而达到遏制“过度诊疗”行为，控制医疗费用不合理上涨，规范诊疗行为的目的，为将来医保谈判购买性价比高的医疗服务奠定技术基础和提供数据支撑。

（二）医疗保险有效支付和治理

通过对积累起来的大数据进行挖掘，支持医疗保险政策调整和医保支付制度改革。

（三）商业医疗保险管理

通过健康医疗大数据分析，为客户分析、医疗控费、理赔风险管理等业务活动提供数据支撑；通过剖析客户参保人群的费用驱动因素及健康情况，为优化保障设计与精算定价提供有力支持，为客户量身定制相关增值服务；通过健康医疗大数据分析，为医疗保险找出影响费用的关键驱动因素，并以此作为战略决策的依据，使决策者有针对性地制定措施，解决问题关键，并在保障医疗质量的前提下有效控制医疗费用；通过健康医疗大数据分析，帮助找出一些典型的理赔费用风险问题等。

三、医药研发生产经营

（一）新药研发

在新药物的研发阶段，通过大数据建模和分析确定最有效率的投入产出比，从而配备最佳资源组合；基于药物临床试验阶段之前的数据集及早期临床阶段的数据集，尽可能及时地预测临床结果，同时可利用大数据技术提高临床试验的统计工具和算法，提升分析效率。

（二）医药生产

医药生产企业利用医药大数据可以精准了解终端市场，实现生产和销售的匹配，同时也将为整个社会服务——全面提升医疗环境，助力药品安全。

（三）药品定价

药监局可通过大数据技术将生产、销售、使用等各个环节的数据汇聚分析，基于卫生经济学和疗效研究对药品进行定价，从而使药物价格处在合理的范围内。

四、公共卫生服务

通过覆盖全国的患者电子病历数据库，快速检测传染病，进行全面的疫情监测，并通过集成疾病监测和响应程序，快速进行响应，从而实现传染病监控。

A.3.2 健康医疗大数据安全风险和需求

随着健康医疗大数据的应用范围不断扩大，尤其是基因、转录、蛋白质、代谢等组学数据的深入研究和应用，健康医疗大数据的价值越来越高，而其敏感数据较多，广泛涉及个人隐私、群体生活、种族安全、全面健康。因此，除了传统的数据安全，大数据技术发展的多样性，以及健康医疗大数据的多源性和高隐私性，也将带来新的安全风险。鉴于此，建立健康医疗大数据安全体系和安全管理责任制度将成为健康医疗大数据应用的核心基础和有力保障。

对于健康医疗大数据来说，尤其是在数据安全和管理安全方面，有其独特的行业安全

特性：

一、数据本身的安全风险

数据本身的安全风险可分为静态数据的安全风险和动态数据的安全风险两方面。静态数据的安全方面，要设置严谨的访问权限控制和安全风险的分级分类管理策略。健康医疗大数据涉及到隐私的数据共享和存储要实现分级隔离、数据加密等安全技术手段，单个信息脱敏后可识别性不强，但仍需考虑多源碰撞后敏感信息易还原的安全风险。动态数据的安全方面，主要是加密和动态审计能力，要对重要敏感数据，比如涉及个人隐私的电子病历、电子健康档案、人口健康数据库产生的大数据进行分级、标识，实现跨平台（端点、移动设备、网络和存储系统）的统一管理。

二、数据使用过程中的安全风险

健康医疗大数据开放对于医疗行业的应用是必须的，但数据不能无条件向公众或者第三方开放而不考虑使用过程的安全风险，因此只能做点对点的共享，或者基于某种特殊约束的多边交易，例如共享健康档案、电子病历、患者用药信息、医疗影像等大数据信息。另外，在数据开放过程中，需要对数据进行脱敏处理，对隐私进行保护。在形成初步疾病诊疗干预措施或者重疾筛查的应用场景下，医疗工作人员不一定拥有整个数据集，但可以借助平台基于特定的目的租赁相关的数据进行计算。大数据租赁和使用的过程中，要保证数据的权利，可以借助同态加密、支持 SQL 的加密数据库，基于加密协议的多方安全计算，基于可信计算环境的多方安全计算，基于隐私保护的机器学习算法等手段实现数据可使用，但不可见。

三、数据处理过程中的安全风险

在健康医疗大数据使用过程中，即使隐去个人信息，在更加深度和广度的搜索下，仍能还原个人信息，从而造成隐私权侵犯。因此，在数据被授权其他方处理后，最重要的问题是处理过程中是否产生滥用和恶意还原敏感数据，是否符合法律法规，是否符合双方或各方同意的隐私条款。在多方计算中，即使加密协议或可信执行环境能够保障数据传输和计算过程中数据处理者不能直接看到数据，但分析程序是可以“看见”和记录数据的。

A.3.3 健康医疗大数据安全标准需求

我国对于健康医疗领域相关的数据安全和隐私保护的立法依然比较滞后。《“健康中国2030”规划纲要》关于推进健康医疗大数据应用中提出：“加强健康医疗大数据相关法规和标准体系建设，强化国家、区域人口健康信息工程技术能力，制定分级分类分域的数据应用政策规范，推进网络可信体系建设，注重内容安全、数据安全和技术安全，加强健康医疗数据安全保障和患者隐私保护。加强互联网健康服务监管”。

尽管国内相关法律法规和指导性文件对保护个人隐私提供了相应的规定，但仍需要针对健康医疗大数据的安全风险和安全需求进行标准体系建设和专项立法，从而保障健康医疗行业全面的信息安全与隐私权保护。

健康医疗大数据安全标准需求包括但不限于：

一、建立健全健康医疗大数据安全体系，形成个人隐私脱敏行业规范，对于涉及敏感数据制定标识赋码、科学分类、风险分级、安全审查规则。注重内容安全、数据安全和技术安全，加强全民健康领域国产密码应用，确保关键信息基础设施和核心系统自主可控稳定安全。

二、建立健全健康医疗大数据网络可信体系，包括强化健康医疗数字身份管理，建设全国统一标识的医疗卫生人员和医疗卫生机构可信医学数字身份、电子实名认证等。

三、建立健全健康医疗大数据安全风险评估机制，加强健康医疗数据安全保障，开展

健康医疗大数据平台及服务商的可靠性、可控性和安全性评测以及应用的安全性评测和风险评估，建立安全防护、系统互联共享、公民隐私保护等软件评价和安全审查制度。

四、建立健全健康医疗大数据安全监测和风险应对机制，包括机构内和机构间的安全信息通报和应急处置联动机制，“互联网+健康医疗”服务安全工作机制，风险隐患化解和应对工作措施等。

五、加强对涉及国家利益、公共安全、患者隐私、商业秘密等重要的健康医疗大数据信息的保护，加强医学院校、科研机构对于健康医疗大数据的安全防范。

六、推进亮照行医中的电子证照（执照）应用，为多点行医和互联网行医提供安全的技术保障。

七、明确各级卫生医疗行政机构和服务机构的数据安全职责和问责制度，包括建立有效的内审机制，必要时进行外审以验证安全措施的有效性，对恶意类数据安全事件进行严厉处罚等。

八、明确健康医疗数据的权属关系及相关法律义务，包括许可权、占有权、隐私权、审批权、收益权、患者知情权、民众选择权等。

A.4 教育大数据

A.4.1 教育大数据特点

教育大数据包括日常教育活动中参与者的全部数据。基于大数据的精确学情诊断、个性化学习分析和智能决策支持、学生精准帮扶、心理健康干预、贫困学生精准资助、学生舆情综合分析、图书馆资源应用分析、科研数据分析等一系列创新应用，对促进教育公平、提高教育质量具有重要作用，已逐步成为各级各类教育机构实现教育现代化、提升教育治理能力的最新实践，将会给我国教育行业带来极其深远的影响。教育大数据的发展对教育行业的影响体现在以下四个方面：

一、个性化、差异化的教学模式创新

利用教育大数据分析学习者个人学习相关数据，精细刻画学习者的学习特征，掌握学生特点、发现学习需求、引导学习过程、诊断学习结果，为学习者提供个性化的学习支持。与此同时，通过对学习者学习过程数据、教学结果数据采集，结合教育理论科学，对教学过程和教学效果等进行科学的综合评价，根据评价结果优化调整教学策略和教学过程，真正达到因材施教，开展规模化的差异化教学。

二、数据驱动的科学研究新范式

大数据对科学研究的影响是革命性的，继实验归纳、模型推演、仿真模拟之后，以数据作为核心驱动力的数据密集型科学发现已成为科学研究的第四范式，即“科学大数据”。科学研究过程中直接通过对相关数据的挖掘、集成、分析建模，揭示数据背后的规律和趋势，并做出新的科学发现。一批新兴交叉学科以海量数据的分析利用为基础，蓬勃发展，跨越了时空的限制，实现了我国某些科研领域的跨越式发展。

三、教育治理能力现代化建设

教育治理能力现代化提升是推动我国教育改革的重要内容，汇聚社会各方面数据逐步形成教育大数据，已经突破了传统教育部门的数据局限性，实现了精确观察和分析，推动了教育管理模式转变，从经验型、粗放型、封闭型转向精细化、智能化、可视化的教育管理新模式，形成用数据说话、用数据决策、用数据管理的教育创新，教育大数据对于加快教育治理能力的现代化建设具有重要意义。

四、智能化的精准教育服务

通过融合分析管理者、家长、教师、学生等多源数据，能够更好地理解师生需求，改变传统被动服务，主动为学习者、教师、家长等提供更智能化的教育服务，例如根据学生社交及性格特点智能分配宿舍、依据学生的学习兴趣主动推荐阅读书籍、依据测试及考试分析为学生提供课程优化方案、结合学生体质及运动情况主动推荐运动方案等。大数据应用服务已成为学校提升服务能力的先进举措，为广大师生营造一个智能化、个性化的良好学习空间。

A.4.2 教育大数据安全风险和需求

教育是国家之本，教育大数据是当前教育现代化建设的科学实践，教育大数据在促进个性化学习、推动教学创新、实现教育公平、引领教育变革等方面发挥了积极作用，但也面临着大数据时代的诸多安全风险，教育参与者个人隐私保护、数据安全规范管理、数据平台访问控制安全等问题亟待解决。教育大数据面临的安全风险和需求主要包括：

一、学生隐私保护亟待加强

教育大数据涉及受教育者与教育者群体，特别是对于我国人数规模较大的未成年学生而言，在教育过程中会采集学生个人学习、疾病、家庭、生活等方方面面的数据，个人隐私保护就显得至关重要。各类教育教学软硬件系统、互联网教育资源等在招生、就业、测评、辅导等过程中都要求学生提供相关个人信息，但缺乏相应的隐私保护措施，存在着较大的隐私安全风险。

二、学校数据安全管理有待规范

目前教育领域的数据管理还存在着诸多现实问题与安全盲区，包括：各类教育信息系统尚未完全覆盖学校所有业务，存部分数据缺失，数据无法共享和利用；已建管理信息系统未遵循现有的数据标准，数据难以整合和共享；数据使用普遍缺乏安全审计，使用者与监管者模糊不清；对于学校科研及资产、财务数据的分级分类管理粗放。因此，当前数据资源安全管理机制尚未成熟，亟需结合学校需求研究出台相应的教育行业数据安全管理标准规范，提升整体教育行业大数据安全能力。

三、教育大数据云计算安全风险

目前，很多教育大数据在云环境下被广泛应用，云环境使得教育数据面临的安全威胁更为复杂和多样，云服务商、黑客、恶意租户等都可能成为需要防范的对象。云服务虚拟机滥用、租户隔离失效、数据被泄露、篡改或丢失等问题，应用程序接口安全以及代码级安全与测试等程序级安全问题都成为了主要风险。在云服务模式下，还存在安全责任不清晰、业务权限不透明等监管漏洞，使得云端的安全功能并不完全可控。

四、教育大数据安全标准严重缺失

当前教育大数据的规范使用和安全管理缺乏相应标准的指导，教育大数据基础设施安全防护的标准规范要求和教育用户与业务数据的安全标准亟待完善，尤其是教育大数据采集处理、共享利用、开放交易、安全能力等方面普遍缺乏有效监管，也没有相应的教育行业安全标准规范，这已经严重影响了教育大数据的健康发展。

A.4.3 教育大数据安全标准需求

目前亟需加快推进我国教育大数据相关安全标准规范的研究制定工作，切实保障教育大数据采集、交易、使用的安全规范，保障我国教育信息化稳健发展。根据大数据安全通用标准体系框架，结合教育大数据在个人隐私保护、数据分级分类及共享利用、数据平台安全等方面的实际需求，制定教育大数据安全标准规划，为教育机构建设大数据应用提供

标准依据。教育大数据安全标准需求主要包括：

一、教育大数据个人信息分级分类保护标准

教育信息化各类应用系统及平台产生和存储了大量的师生信息，还包括科研项目、成果等重要数据，目前缺乏对于个人信息管理的相关标准，对师生个人信息的共享利用仍存在着非法采集、泄露、滥用等安全风险，亟需制定师生个人信息保护相关的安全标准，对于个人信息建立分类分级管理的标准，规范个人信息收集、存储、处理、共享等数据全生**命周期的相关行为，确保数据来源清晰、责权明确、应用有度，保障广大师生的合法权益。**

二、教育大数据共享开发相关安全标准

教育数据共享利用过程中往往由于安全技术保障及管理不到位出现安全问题，需要建立教育数据共享利用开发相关的安全标准，从应用平台、交易对象、交易共享过程等方面规范数据共享利用的服务安全，确保数据开发利用过程中得到充分保护，有效解决教育数据共享利用中的安全问题，完善数据使用的安全审计工作。

三、教育大数据基础设施安全实施指南

教育大数据基础设施包括云计算、高性能计算、校园网络等重要设施，结合教育行业安全稳定的特殊要求，以及基础设施安全防护需求，围绕基础网络、业务系统、配置安全和防护框架等方面，制定教育大数据基础设施安全防护实施指南。

四、教育大数据脱敏实施指南

教育大数据既有个人隐私数据也有重要科技成果信息等重要数据，这些重要数据在成果发布、项目评审、信息共享、数据分析等教育应用时需要进行数据脱敏，需要结合当前个人信息去标识化指南以及大数据脱敏相关标准，制定教育大数据脱敏实施指南。

A.5 金融大数据

A.5.1 金融大数据特点

金融大数据是指现代化的金融机构广泛收集的各渠道海量结构化数据、半结构化数据和非结构化数据。利用大数据技术对这些数据进行实时分析与整合，从而获取客户全方位信息，对客户体验做深度挖掘，掌握客户真实需求，增加客户粘性，并提供有效、合理的安全保障。通过分析和挖掘客户的交易和消费信息，掌握客户的消费习惯，准确预测客户行为，使金融机构和金融服务平台在营销和风控方面有的放矢。

大数据技术的应用在金融行业中发挥的作用有以下特点：

一、实现精准快的营销

大数据时代，金融机构迫切的需要掌握更多用户信息。应用大数据技术，金融机构可以实现在海量信息中快速提取有用信息，并进行分析整合，继而构建用户 360 度立体画像，从而可对细分的客户进行精准营销、实时营销等个性化智慧营销。

二、支持精细化管理，促进经营模式的升级

技术环境的革新使得传统金融模式面临着巨大变革，特别是大数据的快速发展，传统经营模式面临革新。通过大数据分析方法进行数据分析后，金融大数据可用于改善经营决策，为管理层提供可靠的数据支撑，使经营决策更加高效、敏捷，精确性更高。例如，银行通过收集和分析大量中小微企业用户日常交易行为的数据，可以快速判断其业务范畴、经营状况、用户定位、资金需求和行业发展趋势，从而为管理层提供精确的营销策略。

三、实现金融服务创新和产品创新

通过大数据技术，金融机构可监控各种市场推广运作情况，将客户行为转化为咨询流，

从中分析出客户的个性特征、风险偏好，进一步分析及预测客户潜在的需求，将精准营销扩展至服务的创新与优化；同时，通过高端数据分析和综合化数据分享，可有效对接银行、保险、信托、基金等各类金融产品，使金融机构能够从其他领域借鉴并创造出新的金融产品。

四、加强风险的可审计性和管理力度，推动风险管理模式的创新

金融机构可以应用大数据技术，统一管理内部多源异构数据与外部征信数据，解决传统金融风险管理中的信息不对称难题，提升风险判断和风险预警能力，实现风险管理的精确化和前瞻性；同时，也可以借助大数据技术，挖掘数据之间的多维关系，建立更加准确的决策模型。例如，银行可以通过大数据技术打破信息孤岛，全面整合客户的多渠道交易数据以及个人金融、消费、行为等信息，降低信贷风险。

五、带来新的用户体验

大数据时代的到来使得金融机构为客户带来更多新的用户体验。例如，花旗银行通过社交网络、公共网页上得到的客户记录来细分客户，按照客户行为进行分类，为客户提供质量一致的客户体验。

A.5.2 金融大数据安全风险和需求

随着“互联网+”和“大数据”时代的来临，传统银行如今不仅面临着前所未有的来自其他领域的跨界挑战，也不得不面对“大数据+互联网”模式所带来的关于数据、个人隐私等安全风险和隐患。近年来，金融界数据安全事件频频发生，不仅给客户带来直接经济损失，也给金融业的声誉带来负面影响。金融行业需要建立并完善一套大数据平台安全保障体系，最大范围地保护数据安全、行为安全和环境安全，保证大数据的完整性、一致性、准确性、保密性、可用性和可追溯性。

金融大数据的安全风险和需求可概括为：

一、金融大数据的高度集中对金融机构基础设施带来威胁

金融行业作为世界上数据最为密集的行业之一，大数据时代的金融数据更是以几何级数增长。金融大数据的高度集中带来诸多安全风险。第一，高度集中、高敏感度的数据造成数据暴露的风险加大，使其更容易吸引潜在的攻击者将其作为攻击目标。第二，数据的高度集中使得攻击者一旦得手将一次性获得更多的有价值数据。

金融大数据的高度集中对金融机构提出了更高的数据安全要求和技术挑战。目前，金融信息安全保障体系并不完善，有些网络技术、通信设备和应用系统严重依赖国外技术，不能做到对核心技术的自主可控。例如，很多金融机构，尤其是中小银行，在信息化建设过程中普遍存在核心系统落后、管理与决策信息化薄弱等问题。在此背景下，各大银行及其他金融行业的核心系统一旦出现安全问题，将带来难以估量的经济和信誉损失。

二、智能终端的普及对个人金融数据信息泄密的威胁

随着国内智能移动终端市场的不断扩大，第三方支付渠道得到普及，从电话到电视，从 POS 机到 PAD，从网络到手机各种支付手段层出不穷。而智能数据终端中存储了大量的个人数据信息，这些信息包括客户各种网上银行的账号、密码，以及第三方支付终端在交易过程中所存储的大量的金融信息，一旦被攻击将会导致个人金融信息泄露。同时，金融信息的网络化，必然促使金融信息系统通过互联网与终端智能设备相连接，并参与到金融信息系统的数据采集、储存、传输和处理中来，导致信息量也会越来越多。在与外部终端设备的数据交换中，本来就封闭的网络对外开放，无疑会增加了被入侵和攻击的几率。因此，智能终端的数据采集、存储、传输、处理都会增加金融信息受到攻击的威胁。

三、数据虚拟化技术的发展和应用对金融大数据泄密的威胁

数据虚拟化技术是最近比较流行的一种用户访问权限及管理和优化异构基础架构的技术。虚拟化的应用使不同程度的信息混存于同一物理介质上，造成信息访问权限的混乱以及数据泄密等问题。大数据配合虚拟化技术虽然实现了实时交易和迅速拓展业务，但这也反映出风险会实时快速的扩散。因此如何保管虚拟化后的不同密级的信息，避免越权访问或数据泄密就成为了关键。一方面，采用虚拟化管理程序的大多数企业所使用的虚拟化技术软件可能存在安全漏洞，很大可能会导致黑客的攻击；另一方面；随着虚拟化技术的发展和金融电子渠道的不断拓展以及网上业务的普及，数据处理的复杂度越来越高，各种金融卡号的失窃、电子欺骗等犯罪活动也逐年增多，来自互联网的虚拟数据的入侵和攻击也成为这种金融信息安全受到威胁的原因之一。

四、监管政策的缺失对金融大数据安全的威胁

随着我国金融事业的快速发展，而相对应的相关的监管政策并没有与之快速的匹配起来。首先，法律监管政策缺失。由于中国互联网金融发展时间较短，现有的法律法规并不完善。同时，由于金融大数据涉及的方面十分广泛，与现有的法律监管体质并不能完全重合，从而导致法律监管不到位。其次，金融机构内部监管政策的落后和缺失。很多金融机构缺少客户金融信息保护的应急保障机制，当客户信息泄露和资金发生损失时，不能在第一时间采取措施。再次，金融消费者自我监管意识不到位。金融消费者并不能切实认识到对自身重要信息的保护，尤其涉及财产和自身安全的金融数据。

A.5.3 金融大数据安全标准需求

根据金融行业特性，中国人民银行、银监会、证监会、保监会、国家标准委联合发布了《金融业标准化体系建设发展规划（2016-2020 年）》（银发〔2017〕115 号）（以下简称《规划》），其中明确提出了“十三五”金融业标准化工作的指导思想、基本原则、发展目标、主要任务、重点工程和保障措施。当前，金融行业的安全标准主要围绕《金融行业信息系统信息安全等级保护实施指引》和《金融行业信息系统信息安全等级保护测评指南》两个行业标准展开。2015 年至今，已立项的金融行业安全标准主要是针对移动支付、网上银行、密码应用等细分领域的。

基于对金融大数据安全风险和挑战的综合分析、当前大数据技术和应用发展现状的研究，以及当前我国对大数据整体安全合规方面的要求，提出如下金融大数据安全标准化需求：

一、金融大数据安全相关术语和框架

当前，金融大数据技术和应用在快速变化之中，然而，由于金融机构与用户对金融大数据定义、金融大数据安全角色、金融大数据生命周期等金融大数据安全相关术语和框架的认知水平不同，严重影响了其对金融大数据安全的建设需求与建设目的的理解。因此，应优先制定金融大数据安全相关术语与安全框架等基础标准，从而为其它标准的制定打下坚实基础。

二、金融大数据生命周期管理标准

数据是金融大数据系统中的重要资源，其安全性至关重要。当前，我国缺乏针对金融大数据的数据安全管理规范，因此需要制定金融大数据生命周期管理安全标准，以指导对金融大数据全生命周期的安全管理。金融大数据生命周期管理标准主要包括如下内容：1. 在数据创建、数据保护、数据访问、数据迁移、数据归档和数据回收/销毁等金融大数据生命周期管理中，要减少来自组织内部和外部的各种数据安全风险；2. 通过规范金融大数据生命周期管理中的安全管控、安全建设、安全运维等，提高金融大数据创建者和运维者的风险防范意识；3. 考虑到金融大数据之间的互联互通、交互与共享问题，建议建设金融行

业大数据平台，从而实现对数据全生命周期立体化、全方位的安全管控，为我国金融行业大数据发展奠定坚实的基础。

三、金融大数据个人信息保护标准

用户个人信息的潜在价值刺激着人们不断收集和使用金融大数据的欲望，巨大的经济利益催生地下产业链非法牟利，严重威胁用户个人信息安全。因此，需要制定金融大数据个人信息保护的标准，遏制大数据时代个人信息安全的系统性风险。金融大数据个人信息保护标准首先应引导重点金融机构开展个人信息安全相关国际标准和国内标准的实施认证，其次要规范个人信息处理的全流程活动，规定个人敏感信息在收集和利用之前须获得个人信息主体明确授权。最后，政府应进一步研究出台标准采用的激励措施。

四、金融大数据脱敏实施指南

结合金融行业业务的实际需求，保护金融大数据在查询、迁移、共享等应用场景中的安全使用，制定金融大数据脱敏实施指南。该指南的重点内容应包括金融大数据脱敏方法、金融大数据脱敏安全要求、金融大数据脱敏方法安全评估标准、金融大数据脱敏实施存在的风险、金融大数据脱敏存在的问题、金融大数据脱敏实施步骤等。

A.6 互联网金融大数据

A.6.1 互联网金融大数据特点

互联网金融大数据是互联网金融在运行中所依赖、产生、收集、分析、挖掘、使用的数据，是互联网金融的核心和基石。在互联网技术下，大数据驱动金融发展的特点更加显著。借助大数据技术、互联网技术、移动通信技术、密码技术等新技术和手段，互联网金融具备速度更快、参与度更高、中间成本更低、操作上更便捷等一系列特征。

数据是互联网的基础，互联网为金融插上翅膀。互联网金融大数据具备以下特点：

一、影响大

由于互联网加快了数据的传播，而金融大数据又属于个人核心隐私材料。在我国互联网金融发展现状下，信用体系尚不完善，互联网金融的相关法律还有待配套。互联网金融单位的违约成本较低，容易引发多种金融风险问题，造成群体性事件。

二、数量多

互联网金融大数据是获取的个人的金融行为数据，而这是属于个人数据中非常高频使用的部分。国内互联网金融服务企业获取的互金大数据已经达到数百 PB，而且还在不断高速增长中。

三、速度快

互联网金融业务主要信息由系统处理，操作流程完全标准化，业务处理速度更快。在用户画像和信用数据库等金融大数据的支持下，经过数据挖掘和分析，引入风险分析和资信调查模型，一笔业务从申请到完成只需要几秒钟。

四、覆盖广

根据最新的相关统计报告，中国互联网用户的人口渗透率已经高达 50%，互联网金融已经覆盖我们身边的很多人。互联网金融大数据使得用户能够突破时空的限制，在互联网上寻找符合自己需要的金融资源，而金融服务提供者可以更直接、更精确地提供金融服务。

五、价值高

金融数据是用户个人数据的核心数据和高频数据，用户无时不刻在进行着购物、交通、餐饮、住宿、租赁等金融活动，通过大数据平台对用户的收入水平、消费偏好、行动位置、

消费特征、品牌倾向等维度实时勾勒出用户的“真实画像”，从而深入介入和了解用户的生活。因此，互联网金融大数据具备极高的价值。只要拥有更大规模、更全维度、更多用户的互联网金融大数据，就能在互联网金融大潮中取得优势地位。

六、安全弱

近年来，依赖于大数据和电子商务的发展，互联网金融得到了快速增长，但总体来说仍然存在巨大的安全风险。从全国范围来看，至今还没有统一的互联网金融大数据方面的标准，也缺乏统一的互联网金融监管系统，更不存在全国范围的互联网金融信息共享机制，这使得互联网金融大数据信息不能得到充分发挥和利用。而互联网金融大数据的“速度快”的特性导致整个系统行走在不稳定和不安全的钢索上。

A.6.2 互联网金融大数据安全风险和需求

互联网金融大数据在给用户带来便捷和效率的同时，也隐含着很大的安全风险。总体来说，可以从标准、管理、技术这三个层面对互联网金融大数据所面临的安全风险进行分析。

一、从业机构管理水平和风险控制能力不足

互联网金融作为新生事物，国家层面的互联网金融相关的监管标准和指导意见在逐步完善中。互联网金融相关从业机构和单位对金融安全并没有像传统金融机构那么重视，其内部管理水平和风险控制能力与其服务的大量互联网金融用户和数据并不匹配，在日常业务活动中存在不当的操作、内部控制程序存在缺陷、信息系统也需要不断完善。

二、个人信息、数据、隐私泄露

互联网金融的个人信息和数据是重要的敏感数据，但是一些从业机构和单位片面追求利润和业绩，并没有在个人信息、数据、隐私“传输、存储、使用、销毁”等环节建立保护敏感信息的完整机制，这大大加剧了个人信息、数据、隐私的泄露风险。

三、互联网金融大数据业务系统自身的安全漏洞

传统的银行金融网络属于相对封闭的行业网络，对外提供的网络服务也通过 UKEY、动态密码等安全等级较高的手段进行防护。但互联网金融就没有这一层保护，其业务完全在互联网中运行，而网络中的病毒、蠕虫、钓鱼、僵尸网络、木马、间谍软件、DDoS 等威胁层出不穷，互联网金融业务系统本身的安全漏洞直接放大了上述威胁。举例来说，2017 年 5 月全球爆发的“WannaCry 蠕虫式勒索病毒”，就给全球范围内的多个机构、单位和个人造成了严重的经济损失。

四、泛滥的外部威胁和海量的外部攻击

金融安全涉及到用户的财产，病毒、木马程序、密码嗅探、钓鱼等通过盗取客户资料，直接威胁互联网金融的安全。虽然互联网金融机构重视防护这些外部威胁，但由于互联网技术的飞速发展，使得病毒、木马等威胁手段也不断翻新升级，同时大量的病毒、木马、钓鱼的传播途径均来自海外的网络空间，加大了安全监管的难度。而互联网金融由于兼备了“网络，金融，长尾，新生业态”等特性，更是黑客和不法分子进攻的“首选”。互联网金融面临着传统 SYN 攻击、DNS 泛洪攻击、DNS 放大攻击，以及针对应用层和内容更加难以防御的应用层 DDoS 攻击，APT 攻击等多种网络恶意攻击。

为了规范互联网金融的发展，需要在统一的标准要求下，充分考虑金融行业信息系统现有标准和互联网信息技术发展情况，制定符合我国互联网金融领域发展需要的相关标准及规范，指导互联网金融从业单位进行相应的信息安全建设和安全运维管理，提高互联网金融从业单位的安全准入门槛。达到规范互联网金融市场秩序、推动互联网金融行业技术进步、促进互联网金融业的跨行业发展、提高互联网金融行业管理水平的目的。

安全需求方面，互联网金融大数据的安全需求主要体现在管理和技术两个方面。

一、管理方面

（一）外部金融监管技术的提升

互联网金融监管部门也需要提升监控水平，加强技术力量，不断完善网络安全技术，如防火墙技术、加密技术。同时应大力发展我国先进的信息技术，提高计算机系统的关键技术水平，在硬件设备方面缩小与发达国家之间的差距，提高关键设备的安全防御能力。

（二）内部管理制度的健全和完善

互联网金融机构应该成立专门的风险监控部门以及大数据业务部门，并分别向机构的最高领导和最高技术领导进行汇报。风险监控部门应该根据各机构自身的业务特点和风险特征，充分利用大数据技术，建立起独立的风险评估、分析和审核机制，以及完整的工作流程体系，并制定详细的操作方案，确保每个风险点都被严格监控。

（三）专业技术人才队伍的培养和考核

互联网金融领域是人才密集型领域，必须紧紧围绕我国金融改革发展，保证专业技术人才的知识更新，进一步提高互联网金融行业创新能力和竞争力。在产品开发、风险管理、财务管理、信息技术、法律等专业技术领域，以适应互联网金融业知识和技能等要素密集的特点，以提高专业水平和创新能力为核心，培养造就一支数量充足、素质优良、门类齐全、梯次匹配的专业技术人才队伍。

二、技术方面

（一）平台及系统的审计与风险控制

互联网金融企业是互联网和金融企业的有机结合。由于互联网企业具有的虚拟化、交易对象以及交易流程难以确定等特点，其所面对的风险比传统的金融机构要高很多，因此对互联网金融行业内部审计的风险管理极其重要。对高速运转的互联网金融大数据业务系统，必须部署全面的运维审计与风险控制系统，通过账号管理、身份认证、自动改密、资源授权、实时阻断、同步监控、审计回放、自动化运维、流程管理等手段提升安全性。

（二）国内自主知识产权产品的进一步推广使用

金融系统是国民经济的支柱，安全问题关乎国家发展。我国金融行业发展迅速，对信息化手段支持的突发性高要求，导致我们为尽快实现业务稳定运行而直接引用国外成熟技术，很多关键的应用系统大多是建立在非自主产品上，形成了一定程度的技术依赖。随着国产品牌的不断发展，我们需要持续开展安全审计、强制访问控制、系统结构化、多级系统安全互联访问控制、产品符合性检验等相关技术。研发用于保护互联网金融业务系统的安全计算环境、安全区域边界、安全通信网络和安全管理中心的核心技术产品。研发自主可控的计算环境、操作系统、中间件、数据库等基础产品，实现对国外软硬件的替代，降低对国外技术的依赖。

A.6.3 互联网金融大数据安全标准需求

根据前面对互联网金融行业大数据安全风险的分析，进一步明晰了互联网金融行业在大数据安全方面的需求，现初步提出互联网金融行业大数据安全标准体系规划：

一、互联网金融大数据信息安全指南

互联网飞速发展的特性和金融行业对安全的需求，使得互联网金融机构迫切需要制定符合行业发展需求的大数据信息安全指南。该指南应包括互联网机构实施大数据的相关策略讨论，机构和方案的结构化法律法规组件，探讨互联网金融机构在选择和实施大数据安全控制措施方面需考虑的内容，以及在互联网金融服务机构中大数据信息安全风险的要素，并给出了基于互联网金融机构大数据业务环境、实践和规程方面应考虑的建议。

二、互联网金融大数据信息安全通用规范

目前，互联网金融大数据信息安全缺乏针对性的安全要求，需要从技术、管理、业务等多方面统一规范，提升互联网金融大数据安全防范能力，同时为金融行业主管部门、专业检测机构进行检查、检测、认证提供依据。

三、互联网金融大数据服务安全框架

为了保证互联网金融大数据在安全应用方面遵循统一的标准，需要针对互联网金融大数据关键安全问题进行汇总和解答，并提供互联网金融机构在应用大数据服务方面的标准参考。

A.7 电信大数据

A.7.1 电信大数据特点

电信大数据是指电信运营商在业务经营过程中积累的大量数据资源，包括用户属性数据、通话数据、位置数据、终端数据、上网行为数据、消费数据等。电信大数据具有数据规模大、信息种类多、覆盖用户广、应用价值高等特点。近年来，在保障数据安全的前提下，电信行业聚合生产运营、网络承载和企业管理等三大来源的数据，利用大数据深度挖掘分析技术，向客户提供数据服务能力。

可拓展的大数据应用服务主要包括内部支撑、社会服务、商业运营三大类。

一、内部经营支撑

通过对电信行业的运营管理数据和业务数据进行大数据分析，提高业务经营效率。例如，通过大数据分析掌握电信基础网络中网络流量变化、网络行令数据以及设备运行情况，可以及时调整资源配置，进行全网络优化，提升网络质量和网络利用率；通过分析不同用户群流量使用特征以及存量用户流量趋势，可以调整流量服务产品设计，实现流量经营。

二、社会服务支撑

通过对其拥有的各种数据进行深度挖掘，结合不同行业客户的业务特点和行业数据，改善公共社会服务能力。如提供城市规划与交通线网规划、路网状态实时监测与公共交通调度、信息验真服务、公共区域安全监测等不同的社会服务支撑信息服务。例如，基于移动蜂窝网络产生的位置信息，可为政府公共管理、城市规划、交通规划提供数据，并为零售商提供人群分布、流向、热点等信息。

三、商业运营支撑

通过对用户终端信息、用户基础数据、订购产品及行为数据进行大数据分析，刻画用户画像，为商业行为提供用户标签支撑。例如，可以生成用户基础数据特征与兴趣特征模型，进而预测客户行为，进行程序化广告投放，实现精准营销；可以进行产品与推广的规划，以个性化、精准型的业务内容不断增强客户黏性。

A.7.2 电信大数据安全风险和需求

大数据给电信行业带来了新的发展机遇，电信运营商借助大数据积累了竞争优势，并不断发展各类大数据应用。但同时也应注意到，对数据的集中管理、数据对外开放等新技术特点和业务新形态的应用，也使电信行业大数据面临新的安全风险和挑战，主要包括：

一、敏感数据泄露风险

由于缺乏敏感数据衡量标准，对敏感数据的识别与分类不规范，缺乏敏感数据的标识

与处置工具，造成对敏感数据的保护措施选择不恰当，在开发测试、数据分析、数据开放过程中无法有效界定数据共享和开放的边界，容易导致敏感数据泄露风险。

二、供应链管理风险

电信行业的移动网络设备由多家供应商提供，这些设备在运行过程中不断采集并传输用户数据。同时，有些大数据平台或系统由第三方代建设、代维护，这将导致在特定阶段，部分设备的操作权掌握在供应商手中，这意味着供应链提供的设备和平台可能存在安全管理风险。

三、合作方留存数据风险

合作方在使用业务系统过程中可能会违规收集敏感数据，或者利用某些漏洞还原敏感数据，或通过数据沉淀获取全量数据，或在业务使用场景外非法拷贝并留存敏感数据，这将带来敏感数据泄露风险。

四、数据共享安全风险

在电信运营商内部，信息系统建设相对分散，敏感数据跨部门、跨系统共享留存比较常见，一旦其中某个环节存在安全防护措施不当问题，则可能发生数据泄露事件。另外，由于数据去敏感化处理缺乏安全开放衡量标准，难以保证去敏感化强度，在重要数据去敏感化后再对外开放的过程中，第三方可能会采用大数据关联分析、聚合分析等技术非法获取部分敏感数据。

A.7.3 电信大数据安全标准需求

为了促进电信行业大数据业务健康有序发展，保障电信企业的数据资产安全，保障用户的合法权益，亟需在大数据安全通用标准的基础上，结合电信大数据行业特征来制定电信行业大数据安全标准，为电信行业建设大数据安全保障体系提供依据和指导意见。

行业安全标准需求主要包括：

一、电信大数据分类分级实施指南

由于目前缺乏适用于电信行业数据分类分级的实施标准，无法落地大数据安全分类分级差异化管控措施，造成实际应用中数据过度共享或开放不足的问题。为解决这个问题，可基于通用大数据分类分级标准，结合电信行业大数据特征制定电信大数据分类分级实施指南，充分识别、标记、分类、分级运营商采集、生产的各类结构化、半结构化、非结构化类型数据。

二、电信大数据去标识化实施指南

为保护电信大数据在开放、测试及共享等应用场景中的安全，基于目前已有的大数据去标识化通用标准、个人信息去标识化指南和上述电信大数据分类分级指南，结合电信行业实际业务需要，制定电信大数据去标识化实施指南，其内容可包括电信大数据去标识化方法、数据去标识化安全要求、数据去标识化方法安全评估标准等。

三、电信大数据应用业务安全技术要求

随着电信大数据内外部应用业务的拓展，需要制定适用于运营商对外合作过程中各类大数据应用业务的安全保障框架和机制，内容可包括大数据应用业务流程与安全管控框架和大数据采集、大数据存储、大数据挖掘、大数据输出、大数据传输、大数据运营等方面的安全技术要求。

四、电信互联网大数据开放平台安全管理要求

基于通用的大数据平台安全框架，结合电信行业大数据业务开放应用多样性特征和安全防护需求，重点围绕行业数据互联互通、数据安全和个人隐私保护等目标，制定电信大数据开放平台安全管理要求。

A.8 能源大数据

A.8.1 能源大数据特点

“互联网+”智慧能源（以下简称能源互联网）是一种互联网与能源生产、传输、存储、消费以及能源市场深度融合的能源产业发展新形态，具有设备智能、多能协同、信息对称、供需分散、系统扁平、交易开放等主要特征。在全球新一轮科技革命和产业变革中，互联网理念、先进信息技术与能源产业深度融合，正在推动能源互联网新技术、新模式和新业态的兴起。能源互联网是推动我国能源革命的重要战略支撑，对提高可再生能源比重，促进化石能源清洁高效利用，提升能源综合效率，推动能源市场开放和产业升级，形成新的经济增长点，对提升能源国际合作水平具有重要意义。

一、实现能源大数据的集成和安全共享

实施能源领域的国家大数据战略，积极拓展能源大数据的采集范围，逐步覆盖电、煤、油、气等能源领域及气象、经济、交通等其他领域；实现多领域能源大数据的集成融合；建设国家能源大数据中心，逐渐实现与相关市场主体的数据集成和共享；在安全、公平的基础上，以有效监管为前提，打通政府部门、企事业单位之间的数据壁垒，促进各类数据资源整合，提升能源统计、分析、预测等业务的时效性和准确度。

二、创新能源大数据的业务服务体系

促进基于能源大数据的创新创业，开展面向能源生产、流通、消费等环节的新业务应用与增值服务；鼓励能源生产、服务企业和第三方企业投资建设面向风电、光伏等能源大数据运营平台，为能源资源评估、选址优化等业务提供专业化服务；鼓励发展基于能源大数据的信息挖掘与智能预测业务，对能源设备的运行管理进行精准调度、故障诊断和状态检修；鼓励发展基于能源大数据的温室气体排放相关专业化服务；鼓励开展面向能源终端用户的用能大数据信息服务，对用能行为进行实时感知与动态分析，实现远程、友好、互动的智能用能控制。

三、建立基于能源大数据的行业管理与监管体系

探索建立基于能源大数据技术，精确需求导向的能源规划新模式，推动多能协同的综合规划模式，提升政府对能源重大基础设施规划的科学决策水平，推进简政放权和能源体制机制持续创新；推动基于能源互联网的能源监管模式创新，发挥能源大数据技术在能源监管中的基础性作用，建立覆盖能源生产、流通、消费全链条，透明高效的现代能源监督管理网络体系，提升能源监管的效率和效益。建设基于互联网、分级分层的能源统计、分析与预测预警平台，指导监督能源消费总量控制。

A.8.2 能源大数据安全风险和需求

能源大数据在促进能源互联网创新发展的同时，面临的安全挑战主要表现在：

一、敏感数据保护

大数据平台中包含多个能源行业产生的大量原始数据。随着大数据分析业务的深入开展，将有更多的分析和应用开发团队使用。数据的融合和共享将带来巨大的商业机会，然而数据在发布、存储、处理、使用等环节存在敏感数据泄露的风险。既要实现不同数据分析团队和应用开发团队间的高质量数据共享与发布，也要保护好数据的安全与隐私，避免敏感数据的非法访问与泄露，这是亟待解决的问题。

二、数据脱敏与保护装备缺乏

与传统的关系型数据库相比，大数据面临的场景往往是：数据的发布是动态的，且数据来源众多，总量巨大。如何在数据发布阶段，在保证数据可用的前提下，高效、可靠地去除敏感的数据内容，持续评估和降低系统脱敏和监控技术应用对大数据平台并发处理能力、访问用户数及数据量影响，这是亟待解决的问题。

三、大数据体系安全验证机理缺失

大数据技术体系包括了数据采集、预处理、存储管理、分析挖掘、安全、敏感数据处理、可视化等，技术复杂且发展迅速。平台、组件和算法的选型、应用成果验证往往缺乏有效的方法和技术手段，大数据模型和应用的研发效果也缺乏评价方法和能力。

A.8.3 能源大数据安全标准需求

能源大数据理念是将电力、石油、燃气等能源领域数据及人口、地理、气象等其他领域数据进行综合采集、处理、分析与应用的相关技术与思想。能源大数据不仅是大数据技术在能源领域的深入应用，也是能源生产、消费及相关技术革命与大数据理念的深度融合，将加速推进能源产业发展及商业新模式。面对能源大数据的新模式，为指导和促进大数据在能源行业的应用。同时，根据实际需求及规划，能源大数据应从大数据体系架构、关键技术、集成开发、应用管理、数据管理、数据服务和运行维护等方面制定技术和应用标准规范，形成能源大数据标准体系，指导和规范能源大数据平台和应用建设及其安全运维工作。

能源行业大数据的行业标准需求如下：

一、能源大数据的数据安全类标准

能源行业涉及多个领域，比如电力、石油、燃气等，并在各个领域都会将大量的原始数据汇集到大数据平台中。面对海量的数据，数据的分类分级是数据的合理保护的前提和基础。随着新业务模式的发展，数据安全成为一个挑战，能源大数据的分类分级、数据脱敏及数据防护标准是数据融合、数据共享的前提。

二、能源大数据开放平台类安全标准

随着云计算及大数据的快速发展，基于云架构的企业级大数据平台也成为能源行业各单位建设的重点，同时，大数据也成为推动能源行业发展的关键核心技术，在多个领域得到广泛应用。在这种形势下，大数据平台的开放性也为平台安全带来了挑战。因此，有必要建立大数据开放平台安全标准，并使之成为大数据开放平台建设的重要依据。

A.9 交通大数据

A.9.1 交通大数据特点

交通行业是国民经济正常运转和持续发展的基石，涉及各级政府监管部门、各运输企业以及每个出行者，涵盖交通网络基础设施、载运工具及装备、交通组织与管理等多个方面。随着信息技术的不断深入发展，交通行业积累了大量数据资源。交通大数据的主要来源包括：(1)内部自身产生的数据；(2)相关行业的导入数据；(3)公众交互的数据，等。

乘着大数据发展的东风，交通行业也开始了大数据时代的转型升级，致力于提高资源利用效率和管理精细化水平，全面提升交通行业服务品质和科学治理能力。交通行业中的大数据可以为管理者、出行者和环境保护提供有力的技术支持，典型的应用主要包括：

一、对管理者的支持

大数据有助于提高管理者对交通体系的规划、建设、管理、营运和养护水平，体现在：

（一）帮助优化交通网络的规划和建设。通过大数据对网络客货流的周转时间、服务水平的分析，可以找出交通网络中的待优化点，为交通网络的规划、建设提供优化依据；

（二）提高交通体系的信息化、智能决策水平。基于条码、射频、全球定位、手机、视频、身份识别等汇聚的大数据，发掘各种交通方式之间关联信息，为智能决策提供参考；

（三）有效提升交通服务水平。通过各种运输服务之间的关联分析，促进各种运输服务之间的无缝衔接与合作，提高客货运输服务效率，降低社会物流成本。

（四）提高运输安全水平。基于对运营车辆、铁路、船舶、飞机等定位、联网联控系统及交通动态监控，利用历史积累的数据分析安全隐患点及区域、事故易发环境，起到运营安全预测，为安全监管提供决策支持，提高应急反应速度和救援成功率，从信息技术上保障综合交通体系安全。

（五）提高养护效果，降低养护成本。汇聚交通基础设施健康监测数据、动态交通运行数据，通过大数据分析，预警列车、道路、桥梁、隧道的全寿命周期健康状态，及时更换零部件并保养，以确保最大化利用和最小化投入。

二、对出行者的支持

大数据有助于提高出行者的出行服务质量和效率。大数据通过各种运输方式信息系统的互联互通，为公众提供全方位、立体化的出行信息服务，例如飞机票价预测、航班延误预测、出发地目的地联运信息及预订等。

三、对环境保护的支持

大数据的应用，有助于降低交通对环境的影响。通过汇聚交通车辆行驶轨迹、道路港口航空等区域的大气监测数据，分析交通工具对环境的影响，正确衡量不同交通方式对环境的负贡献情况，为管理者提供面向生态交通的规划、建设和管理方案，从而最大限度地降低污染物和二氧化碳排放水平，有效控制噪声污染，建设绿色低碳交通。

A.9.2 交通大数据安全风险和需求

大数据的发展为交通行业带来了新的发展机遇，但是，大数据本身固有的新技术特点和业务新形态应用，使得来自多个部门的多源异构数据的汇聚整合、开放分享和深度挖掘成为大势所趋，这也导致交通行业大数据面临新的安全风险和挑战，主要包括：

一、数据权属不清

交通行业大数据涉及到政府管理部门、第三方平台、承运机构、乘客等多个主体，交通过程中产生的数据如何正确划分其所有权和使用权，是交通大数据健康发展必须要考虑的问题。在权属不明确的情况下，数据的使用不能名正言顺，容易引发法律纠纷，数据的安全问题也难以明确责任方，不利于数据安全保护工作的开展。为此，需要国家和行业尽快制定相关的法律和法规进行必要的规范。

二、数据泄露的追踪和定责困难

目前已有的各类交通数据平台，多为行业内自行采集、自行管理、自行使用，但是在大数据业务应用发展的驱动下，交通行业的数据由原来的各机构分散存储转变为大数据平台集中存储、多家共享的模式，这使得大数据资源的安全风险更加集中。同时，由于缺乏有效的数据追踪手段，一旦发生数据泄露事件，追查泄露责任方的难度很大，定责困难。

三、数据隐私保护面临巨大挑战

近年来，在积累互联网订票、定位导航的传统数据基础上，网约车、航班网络值机、互联网快递下单与跟踪、交通旅游信息综合服务等蓬勃发展，个人数据广泛分布于铁路公路售票、民航票务、快递物流跟踪、网约车服务平台以及交通旅游一体化服务平台中。这

些数据包含公民身份信息、旅客行程与物品运输信息，已成为国内外网络黑市交易的“黄金数据”，诱使非法个人或组织进行数据贩卖以谋取暴利，直接危害广大人民群众的经济利益与人身安全，严重阻碍交通大数据产业的健康发展，迫切需要研发有效的隐私保护方案。

A.9.3 交通大数据安全标准需求

如何在交通大数据利用价值最大化的同时确保信息安全风险最小化，这已经成为业界必须直面和深思的问题。由于各部门和机构各自不同的职能定位、发展需求和利益诉求，他们之间的数据交换、交易和共享问题日益突出，已经成为阻碍交通大数据产业发展的瓶颈，迫切需要国家和行业出台一系列权威标准进行指导和规范，以便加快建立安全、健康、有序的交通大数据生态体系，促进大交通数据产业的可持续发展。

交通大数据行业安全标准需求主要包括：

一、交通大数据确权定价模型和交易指南

目前，由于缺乏交通大数据的确权定价模型，数据的权属划分和价格确定都面临巨大挑战。同时，在数据交易方面也没有成熟的指导意见，交通大数据确权定价模型和交易指南应当对交通大数据资源的权责归属进行清晰划分，构建有效的价格生成机制，明确交易规范，实现交通大数据资源的快速价值发现和高效流通。

二、交通大数据隐私保护要求和测评方法

随着交通大数据应用的不断深入，隐私保护需求日益凸显。这一问题如不解决，将严重影响交通大数据产业的深入发展。同时，对隐私保护状况也需要有一个比较完善的测评方案，确定当前交通大数据所处的隐私保护安全等级，交通大数据隐私保护要求和测评方法希望可以对涉及到用户隐私的信息明确保护要求，提供实施建议指南，并对保护等级提供测评方案。

三、交通大数据安全事件定责与取证规范

在交通大数据的应用中，数据泄露等安全事件在所难免。如果可以对此类事件进行有效取证和定责，可以起到惩前毖后、防微杜渐的效果。交通大数据安全事件定责与取证规范负责对发生交通大数据安全事件后涉事各方的责任划定进行说明，并提供客观、可靠、可信的取证方法及其规范。

A.10 电商大数据

A.10.1 电商大数据特点

电商大数据是指电商行业中累积的大量商家数据、买家数据、商品数据，以及在买卖交易过程中产生的订单数据、交易数据和用户行为数据等，以及利用大数据技术分析、挖掘数据价值的过程中产生的数据。

电商大数据对电商行业的影响体现在以下三个方面：

一、精细化运营及管理

通过对电商数据分析和利用，改进电商业务运营模式，实现业务精细化的运营及管理，包括业务分析、业务智能化、精细化营销、风险管理和运营效率等方面。例如在营销方面，通过对以往的营销数据分析，能够最大化的利用数据资源建立适用的营销方案，并通过营销反馈数据的收集和分析，及时反作用于营销方案的改进工作；在风险管理方面，通过建立实时响应的风险监控系统，对电商业务流中的数据与风险大数据进行关联分析，更好地

识别和控制业务风险。

二、提升业务效率

通过对电商行业中的业务数据进行大数据分析，作用于电商企业的产品设计、绩效管理、配送效率、库存管理和客户关系管理等关键环节，可以提升业务能力和效率。例如，通过对历史销售数据进行大数据分析，结合后期的产业环境，商家可以预测销售数据，进而优化物流和仓储；通过客户大数据分析，物流企业能够更合理的选择派送方式、优选路径，并提供差异化服务，提高物流服务质量等。

三、改进消费体验

电商企业通过对消费者数据的分析可以产生如消费者“画像”等衍生数据。企业基于这些衍生数据，可以为消费者提供个性化的服务，如个性化商品推荐、个性化搜索以及智能机器人客服等服务，提升消费体验。

四、保障生态圈良性发展

电商行业的生态圈涉及电商业务平台、商家、消费者、为商家提供服务的独立软件提供商，以及相关服务机构等众多合法参与者，但也存在着诸如诈骗组织和炒信团伙之类的非法谋求利益的黑灰产组织。电商行业可利用大数据技术精准识别风险，打击炒信、欺诈和侵权等恶意行为，促进电商行业生态的良性发展。

A.10.2 电商大数据安全风险和需求

电商大数据在促进业务发展的同时，相应的安全风险与挑战也随之浮现，主要表现在：

一、数据权属不清

电商业务的开展主要涉及电商平台、商家和消费者三方。电商业务产生的数据如何划分其所有权、控制权和使用权，这是在电商业务开展中合理使用数据的前提。当前，在对电商业务大数据的应用中，通常利用电商平台对数据进行分析，同时也存在商家或商家授权独立软件提供商使用商家数据进行分析的情况，在权利归属不明确的情况下，责任的归属也难以界定，相关数据安全难以保障。

二、大数据聚合分析风险

电商业务中的大数据应用涉及对消费者相关数据的分析，虽然可以通过隐私保护政策、用户授权协议的形式获取相关数据的使用合法授权，而且在对电商业务分析的过程中采用匿名化处理的方式，从而保证用户个人信息安全。但是，在对大数据进行分析的过程中，如何保障不会因为大数据的聚合分析而实现“去匿名化”，依然是亟待解决的难题。

三、数据版权保护

电商行业生态圈内的数据流动和共享较为普遍，目前主要通过法律协议方式约束对数据的使用。但由于缺乏有效的数据版权保护技术手段及措施，难以甄别是否存在超出范围的数据扩散或使用问题。

四、数据跨境安全

目前，国家大力支持跨境电商业务，而跨境电商业务必然涉及数据的跨境问题。不同国家和地区的数据保护法规对数据跨境流动的要求存在差异性，比如俄罗斯明确提出俄罗斯公民的数据应在俄罗斯境内更新后方可传到海外进行处理；欧盟则扩大了数据保护法律适用的管辖范围。这些法规将给跨境电商企业带来高昂的合规成本，制约了跨境电子商务的发展。如何处理数据跨境安全合规与跨境电商战略发展之间的矛盾是亟待解决的难题。

A.10.3 电商大数据安全标准需求

电商业务涉及电商平台、物流厂商、消费者、支付机构、第三方软件提供商等多个角

色，产业链较长。因此，对于大数据安全，需要整体考虑整个生态链中各个环节中大数据安全标准的协同性，既要保障大数据安全，又要允许数据在整个电商行业生态链中的交换与共享。具体来说，电商行业的安全标准需求体现在：

一、电商行业数据开放安全标准

为促进电商业务的开展，电商平台需要对第三方软件提供商或物流厂商进行数据开放。但由于缺乏统一的安全标准，各方的安全水准不一致，容易在数据开放时造成信息泄露。因此，应制定电子商务数据开放的标准，对数据提供者和数据访问者双方的安全性进行规范化，从而在促进业务开展的同时，保障电商数据的安全性。

二、电商行业参与方安全能力评估标准

在电子商务生态环境中，涉及到电子商务平台、第三方软件提供商、支付平台、物流厂商等众多角色。由于缺乏安全能力评估标准，无法衡量和识别整个生态链中的短板，容易造成信息泄露。因此，应制定电商生态参与者的安全能力评估标准，对各个参与者的能力进行评估，提升电商生态中各个参与者的安全能力。

三、电商行业数据脱敏标准

在电商行业中，存在大量的敏感数据，比如个人信息数据、交易数据等。在数据共享和开放的过程中，需要进行脱敏处理。但目前缺乏统一的数据脱敏标准，导致脱敏方式不一致，脱敏效果难以衡量等，影响数据的共享和开放。因此，急需制定电商行业数据脱敏标准，其中定义数据脱敏安全方法和评估标准，规范数据共享、交易、开放等过程中数据的安全管理要求，保护电商业务中的业务数据、个人信息等敏感数据。

四、电商数据跨境安全标准

在电商行业中，存在多种跨境业务模式，比如国外客户从国内购买商品，国内客户购买国外商品等。在这些业务中，都涉及到数据的跨境传输。由于缺乏数据跨境安全标准，容易造成重要数据、个人信息泄露。因此，应制定电商业务中数据跨境安全标准，提升跨境数据传输的安全性，促进跨境电商业务开展。

附录 B 大数据安全标准应用实践

本附录列出了相关企事业单位在大数据实践中应用大数据相关安全标准的情况，包括选用标准、应用场景、应用成效，以及发展展望。这些实践为大数据安全标准化工作的开展奠定了坚实的基础。

B.1 360 企业安全集团大数据安全标准应用实践

为了帮助企业快速实现基于大数据的应用，360 企业安全集团研制了 360 网神安全大数据平台，支持 PB 至 EB 级别大数据的应用。为了保障该平台的大数据安全，公司积极参照各类标准规范开展工作。

一、360 网神安全大数据平台简介

360 网神安全大数据平台是一个集成了多个开源系统的开放平台，包括数据接入、数据存储、数据计算、数据分析、数据应用、数据运维管理、安全防护等功能模块（见图 B.1），能够帮助企业快速构建海量数据分析应用的专业化产品，挖掘大数据的价值。

图 B.1 360 网神安全大数据平台体系架构

二、安全标准应用情况

由于缺乏直接相关的大数据安全标准和大数据系统安全标准的指导（第一个大数据安全标准 GB/T 35274-2017《信息安全技术 大数据服务安全能力要求》在 2017 底才发布），360 网神安全大数据平台从系统、网络、数据、应用等各方面采取安全防护措施来综合保障大数据安全，包括行为审计、数据安全、认证授权、操作系统安全、网络安全、技术设施安全等方面。

1. 行为审计

所有实体在平台上的行为都会被详尽地记录，作为审计记录，并以加密形式进行保存。对审计数据进行分析能够实现实体行为挖掘、异常行为发现等安全功能。这些工作参考了 GB/T 18794.7-2003 、GB/T 17143.8-1997 等标准。

2. 数据安全

所有保存至本平台的数据均采用加密保存，根据数据类型不同采用不同的加密方式。

加密的数据在本平台内部使用时，能够实现透明加解密，不影响数据分析和数据挖掘等工作。这方面的工作遵守国家有关部门的规定和技术要求，如保密管理局颁发的有关标准。

3. 认证授权

认证授权贯穿于本平台每一个环节，包括但不限于数据的获取、数据存储、数据计算、用户操作等方面。认证授权采用国际通用的 RBAC 模型，实现用户组、用户、角色、权限的细粒度管控。采用 Kerberos 进行账户信息安全认证，提供单点登录功能，能够实现用户的统一管理以及认证。

4. 操作系统安全

这主要包括操作内核安全加固、操作系统补丁更新、操作系统权限控制、操作系统端口管理、操作系统运行程序检测等，参考了 GB/T 20272-2006《信息安全技术 操作系统安全技术要求》。

5. 网络安全

在部署之前，根据客户实际情况进行网络规划，使网络具有隔离性、保密性、稳定性等特点。平台的网络划分成 2 个层面，包括业务层面与管理层面，两个层面之间采用物理隔离。在运营过程中，对网络安全事件及时进行处置，参考了 GB/T 28517-2012、GB/T 32924-2016 和 GB/T 25068 系列标准等等。

6. 基础设施安全

综合采用防病毒、边界防护、入侵防护、态势感知、威胁情报等手段，保障平台各类设备的安全。这些工作分别参考了 GB/T 20281-2015、GB/T 31505-2015、GB/T 20275-2013 等标准。

360 企业安全集团也积极参与了大数据安全有关标准的制修订工作，是 GB/T 35274-2017《信息安全技术 大数据服务安全能力要求》和多个即将发布的多个大数据安全标准的编制单位，平台研制人员及时将这些标准的有关要求落实到产品的研制和运营管理工作中。

三、对大数据安全标准的需求

目前，“数据驱动安全”已经成为业界共识，基于网络安全大数据可以及时识别网络安全威胁，有效支撑网络安全的管理和控制，在新时期的网络安全保障工作中发挥了越来越大的作用。网络安全态势感知、威胁情报、APT 发现以及防范电信诈骗、防范网络欺诈等各种基于网络安全大数据的新应用层出不穷，并发挥越来越大的作用，但是，这些工作目前都还没有正式的标准进行指导和规范。希望国家有关方面加快这方面的工作力度，将正在编制过程中的标准尽早正式发布，并支持立项更多的安全大数据应用有关的标准规范。

B.2 IBM 大数据安全标准应用实践

企业核心信息的 80%是以结构化信息，即数据形式存在的。因此，企业需对数据进行自身内在及所依存的基础环境等全方位的保护，来逐步完善企业数据安全。

一、IBM 大数据安全标准及合规实践

大数据环境允许组织聚合越来越多的数据，这些数据包括金融、个人隐私、知识产权、财产或其他敏感数据。这些敏感数据大多有合规标准，例如萨班斯-奥克斯利法案(SOX)、健康保险的可携性和责任法 (HIPAA)、支付卡行业数据安全标准(PCI-DSS)、美国的联邦信息安全管理法 (FISMA) 以及欧盟的通用数据保护法规(GDPR)等。这些合规标准的一些数据审计要求例子如表 B.1 所列：

表 B.1 大数据安全保障体系框架

审计要求	COBIT (SOX)	PCI-DSS	ISO 27002	GDPR	FISMA
1. 敏感数据访问（成功 / 失败的查询）		✓	✓	✓	✓
2. Schema 变更(DDL)（创建/删除/修改数据库表等）	✓	✓	✓	✓	✓
3. 数据修改（DML)（插入，更新，删除）	✓		✓		
4. 安全异常（登录失败，SQL 错误等）	✓	✓	✓	✓	✓
5. 账号、角色、权（DCL)（授予，收回）	✓	✓	✓	✓	✓

IBM 所有安全产品均充分参考并遵从 ISO/IEC 27017:2015《基于 ISO/IEC 27002 的云服务应用的信息安全控制措施》、ISO/IEC 29101:2013《隐私参考体系架构》等标准体系。

IBM Security Guardium 是一个完整的数据安全平台， 对大数据提供了全面的安全保护，同时也能帮助企业快速完成表 B-1 所列的大数据的合规要求。

IBM Security Guardium 支持 Hadoop、Impala、Mongodb 等各种大数据平台，图 B.2 是 IBM Security Guardium 与大数据应用, 安全政策，及标准合规流程和报告集成的视图：

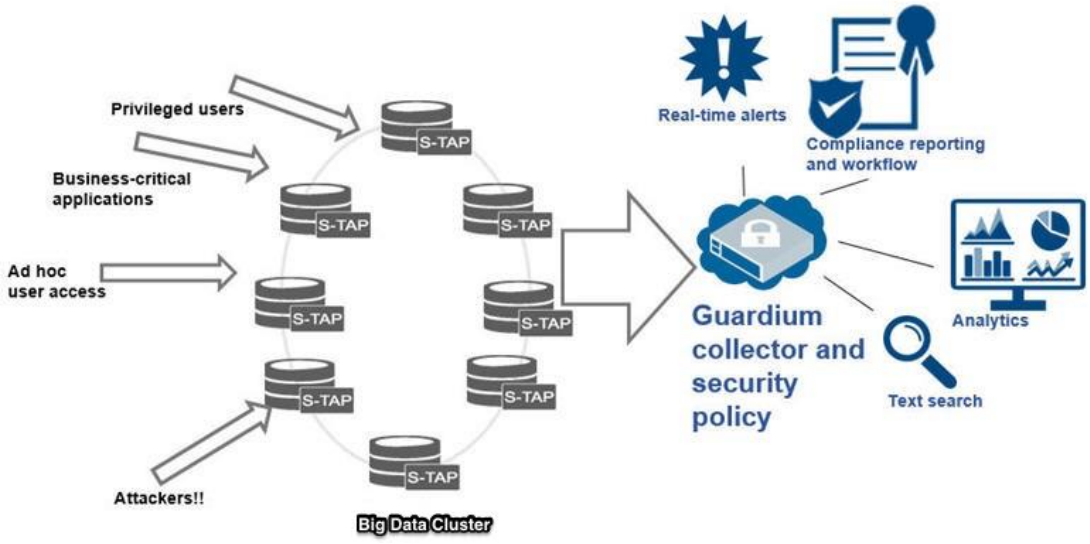

图 B.2 IBM Security Guardium 大数据应用，安全政策及标准合规整合视图

IBM Security Guardium 总结并实施了各种数据安全标准和最佳实践、提供了数据安全保护体系：

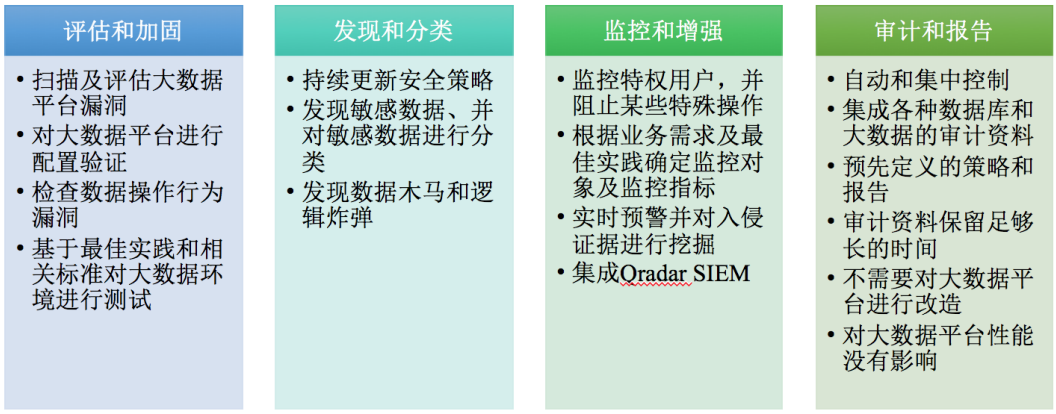

图 B.3 数据安全保护体系

IBM Security Guardium 对大数据提供保护，具体特色体现在：

- 对来自各种应用和用户的 Hadoop 和 NoSQL 的数据访问进行全面实时监控。在应用或用户违反安全策略时进行实时预警并在 SIEM 仪表盘中显示警告信息。
- 通过审计和报告方式来满足合规要求并能提供法律所要求的合规证据。
- 对高流量、高速度、多类型的大数据进行全面的变更管理。
- 对企业的全部数据（数据库、应用、文件、大数据）等进行集中式、自动化的管控。
- 通过加密、屏蔽、掩码等方式保护敏感数据。
- 评估和解决大数据环境中的漏洞，保证大数据系统自身是安全的。

IBM Security Guardium 内置了标准合规加速器，这些加速器包含了标准及合规所需的策略、规则、报告模版。用户按照加速器引导，经过配置后就可以完成合规要求。这里以 Sox 为例看如何使用 IBM Security Guardium 来快速实现 SOX 合规。

在部署 IBM Security Guardium 平台后，并配置好数据源后，通过这 5 个步骤可以快速实现 Sox 合规：

表 B.2 快速实现 SOX 合规 5 步骤

步骤	任务	说明
1	建立 SOX 用户组	SOX 用户组是报告和策略的基础
2	识别报告（在 SOX 加速器中）	确认已经通过配置记录了所有需要的数据
3	定义策略	策略帮助审计和保护生成财务报告所需的数据
4	设定审计流程	对数据处理过程和内部控制方式进行记录和验证
5	对违反审计策略的事故进行复盘	证明对问题调查和披露进行了恰当的控制

IBM 还做了 IBM 解决方案对美国网际安全框架的合规和对标。表 B.3 是数据安全部分。

表 B.3 IBM 解决方案对美国网际安全框架的合规和对标中的数据安全

类别	子类别	IBM 产品和服务
数据安全	PR.DS-1: 静止的数据是受保护的。	IBM Tivoli Storage Manager, IBM Security® Optim™，IBM Security Guardium
	PR.DS-2: 运动中的数据是安全的。	IBM Security Key Lifecycle Manager，IBM Security Guardium
	PR.DS-3: 资产在整个迁移、转让和处理的过程中被正式管理的	QRadar SIEM, Tivoli Storage Manager, Global Technology Services – Data security strategy and assessment

	PR.DS-4: 足够的能力，以确保可用性是可持续的.	Global Technology Services, Global Business Services
	PR.DS-5: 有防止数据泄漏的保护。	QRadar SIEM, （EOS soon）IBM Power Systems™ GX adapters, IBM Security Guardium®, Global Technology Services – Endpoint & network data loss prevention
	PR.DS-6: 知识产权受到保护。	QRadar SIEM, Tivoli Storage Manager, IBM Security Key Lifecycle Manager, IBM Security Guardium, Global Technology Services – Endpoint and network data loss prevention
	PR.DS-7: 取消不必要的资产。	Tivoli, QRadar SIEM, BigFix
	PR.DS-8: 系统开发有单独的测试环境。	Global Technology Services – Data security strategy and assessment, Global Business Services，IBM Security AppScan
	PR.DS-9: 个人隐私和个人可识别信息 (PII) 受到保护。	IBM Security Optim, IBM Security Guardium, Power Systems GX adapters, QRadar SIEM, IBM Security Access Manager，IBM Security Identity Manager

二、对数据安全标准的思考及建议

IBM 在具体实施数据安全项目过程中，从分析、保护、适应三方面，建议项目遵循以下总体性要求，依照法律法规，兼顾国际标准、国家标准、行业规范及标准和企业组织自身需求特点，打造全面的数据安全体系。

- 构建纵深的防御体系
- 采取互补的安全措施
- 保证一致的安全强度
- 规划统一的支撑平台

在制定我国数据安全标准时，应该贯彻落实《国家网络空间安全战略》和《网络空间国际合作战略》，《网络安全法》及《关于扩大对外开放积极利用外资若干措施的通知》（国发[2017]5 号）等国家政策，提升我国标准国际化水平，对已有的国际标准，应尽量等同采用以避免重复，且把有限的资源用到必要的标准创新上。已有的标准对数据分类覆盖缺失，可能导致对重要数据保护不足并对需要开放的数据过分保护。安全威胁来自全球范围，安全防护应该有全球一体化的观念，对全球范围内的安全情报研究应该加强。

B.3 阿里巴巴大数据安全标准应用实践

一、阿里巴巴大数据安全体系架构

为了保障整个数据业务链路的合规与安全，阿里巴巴提供面向电商行业提供的大数据平台，从业务、数据和生态三个层面来保障和应对其数据在消费者隐私保护、商业秘密保护等方面的安全风险与挑战，具体如图 B.4 所示：

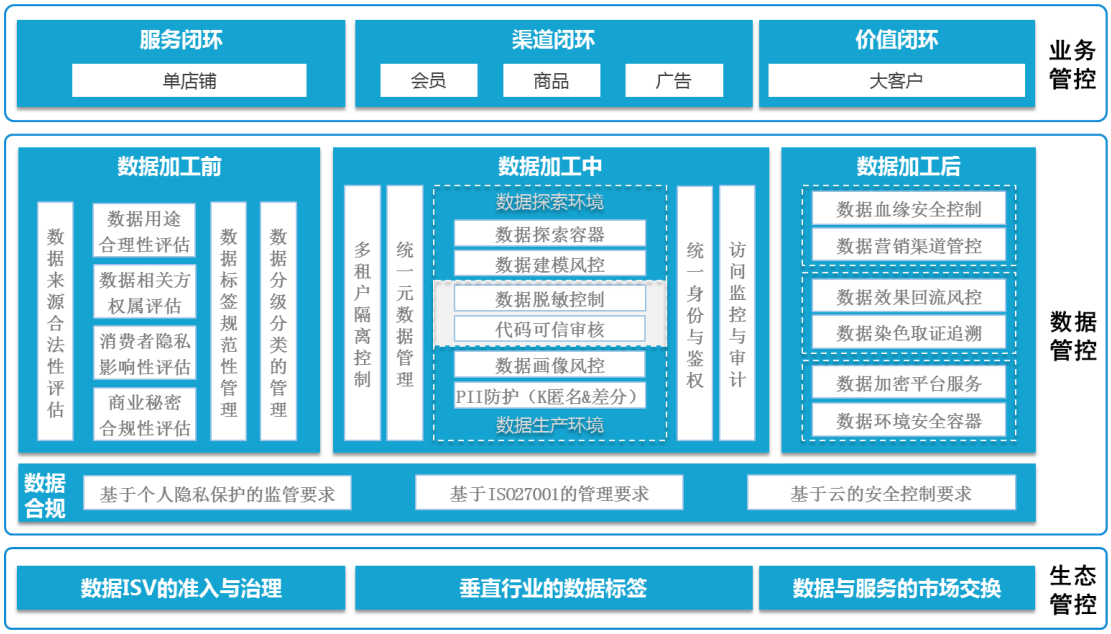

图 B.4 阿里巴巴大数据安全体系架构

首先，在业务模式设计上，大数据安全平台依据电商自身的业务特性和其数据权属关系的边界，建立了以私域数据为基础的店铺内服务闭环、以公域数据为基础的平台内渠道闭环和价值闭环，从而确保了业务整体对数据的授权边界是合理清晰的、对数据的处理逻辑是基于可用不可见的安全原则以及数据的应用产出是基于数据价值而不是裸数据输出的。

其次，此大数据安全平台基于数据业务链路构建了全面的数据管控体系，包括：数据加工前、数据加工中、数据加工后、数据合规等方面的数据安全管控。在数据合规层，参考了《GB/T 35273-2017 个人信息安全规范》、《GB/T 35274-2017 信息安全技术 大数据服务安全能力要求》、《GB/T 31168-2014 信息安全技术　云计算服务安全能力要求》，以及 ISO 27001 系列标准进行实施。通过遵循这些标准，实现了对个人隐私信息的保护、保障了云服务的安全控制，保障了大数据服务的安全性，同时符合了国家的监管要求。

最后，通过对数据 ISV 的准入准出、基于垂直化行业的标签体系建立以及数据生态的市场管理机制建立，确保业务和安全间找到有效的平衡点。

二、阿里巴巴大数据安全实践

阿里巴巴建立了一套标准的大数据采集、计算存储、服务和应用的架构方案，伴随着大数据体系架构的建设，阿里巴巴同步逐渐形成了以数据为中心的大数据安全管理理念，大数据安全工作的开展如图 B.5 所示。

图 B.5 阿里巴巴的大数据安全实践

此安全实践基于《信息安全技术 数据安全能力成熟度模型》来进行，以数据为中心，围绕数据生命周期，对组织机构的数据进行安全保障，通过遵照标准，有效地控制了数据安全风险，提升了公司自身及生态伙伴的数据安全能力，促进了生态内的数据资源的流通与共享，更大地发挥了数据的价值。

三、大数据安全能力成熟度模型

基于长期在内部业务形成的及在外部组织机构学习了解的大数据安全实践经验，阿里巴巴提炼形成了大数据安全能力成熟度模型，如图 B.6 所示。该模型可用于对当前大数据安全能力的有效评估，并为阿里巴巴后期的大数据安全工作开展提供了明确的提升路线。

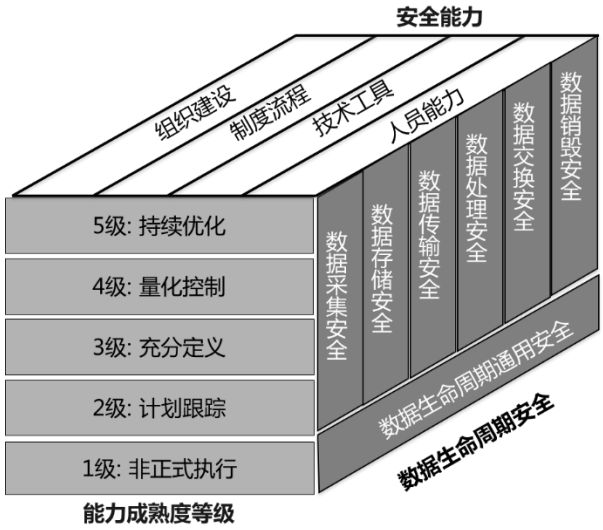

图 B.6 阿里巴巴大数据安全能力成熟度模型

目前，全国信息安全标准化技术委员会大数据安全特别工作组正在制定《数据安全能

力成熟度模型》的国家标准。目前阿里巴巴的大数据安全实践，遵循了正在制定的《数据安全能力成熟度模型》的标准。阿里巴巴已完成在内部业务、生态圈、行业领域多家组织机构的落地推广，实现在各行各业 30 多家企业的试点落地，很好地帮助这些企业提升自身的数据安全保护能力。

阿里巴巴将这些大数据安全实践的经验与外部机构进行针对性的交流和共享，并开展针对生态圈及外部各行各业的组织机构的大数据安全能力的评估和交流，在持续的与外界的互动交流中学习提升对大数据安全实践的重点和方法的认识。

四、建议的标准化需求和说明

基于这些产业实践，对大数据安全标准化工作的建议如下：

（一）建议《数据安全能力成熟度模型》的国家标准在研制过程中，从这些行业实践经验中进行总结和归纳，提炼出通用的安全过程域和安全要求，完善和补充到国标中，对标准草案进行完善。

（二）建议启动制定《数据安全能力成熟度评测方法》、《数据安全能力成熟度提升指南》等配套标准，促进大数据环境下各组织的数据安全能力的评测以及能力提升。

（三）建议启动制定《数据安全人员能力要求》的标准，为组织机构的数据安全人员的人员选取、技能培训、技能规划和能力分级评定等活动提供框架参考和依据，并促进数据安全从业人员的能力提升。

（四）建议同步推进《大数据安全能力成熟度模型》到 ISO、ITU-T 等国际标准中，推动中国标准走向世界，获得更高的认知度和认可度。

B.4 海信交通大数据安全标准应用实践

随着智慧交通行业的快速发展，交通数据出现了爆发性的增长，作为智慧交通产业的引领者，海信将数据视作公司最重要的核心资产加以严密保护，并将这一理念贯彻到交通大数据的采集、汇聚、清洗、挖掘和应用等各个环节。

一、平台安全架构

海信交通大数据平台涵盖了数据采集层、传输层、处理层、应用层等，提供从数据采集、加工、存储、数据分析、机器学习到最后数据应用的全链路技术和服务。针对各层面不同的安全需求，海信设计了以数据安全为中心的系统性安全框架，如图 B.7 所示：

图 B.7 海信交通大数据安全平台架构

二、现有标准应用

海信交通大数据平台安全框架的构建过程中参考了多项信息安全国家标准，如在数据传输层的接入安全模块，参考 GB/T 32213-2015《信息安全技术 公钥基础设施 远程口令鉴别与密钥建立规范》，构建了双向接入认证平台，避免恶意节点混入系统；在数据处理层的数据安全模块，参考 GB/T 20281-2015《信息安全技术 防火墙安全技术要求和测试评价方法》，构建了数据存储防火墙，有效阻止非法访问和入侵行为；在数据处理层的云安全模块，参考 GB/T 31167-2014《信息安全技术 云计算服务安全指南》，构建了安全的云服务平台，作为大数据系统的支撑基础；在数据处理层的入侵检测模块，参考 GB/T 28454-2012《信息技术 安全技术 入侵检测系统的选择、部署和操作》，部署了入侵检测系统，对常见网络入侵行为进行有效监测、发现和处置。

此外，还在采集层部署了适应采集设备和应用场景的弹性安全技术，对采集的数据进行加密保护，并通过安全路由向上层提交。传输层利用高强度加密技术保证数据的保密性，通过完整性校验保证数据的完整性，并借助流量清洗技术防范常见的 DDOS 攻击。处理层采用严格的访问控制策略，对数据的访问权限进行严格管理，并严密监测各类违规行为，保障平台和数据的安全性。应用层对容易遭受攻击的 Web 接口进行加固，并按照用户的信用情况进行分级管理，重点保障大数据的隐私保护需求。

此外，在交通大数据平台中，海信还部署了全面的安全态势感知系统，对平台的安全运行情况进行实时感知，全面监控硬件、软件服务状态，发现异常时及时报警。

三、未来标准需求

随着交通大数据行业的深入发展，涉及到的部门越来越多，同时大数据本身固有的新技术特点和业务新形态应用，使得来自多个部门的多源异构数据的大汇聚大整合成为大势所趋，例如交通大数据直接涉及的部门就涵盖交警支队、交通委员会、市政管理局、公交公司、出租车公司、网约车公司、共享车公司、民航管理局、航空公司、客运公司、铁路局、机场、港口、长途车站、火车站等，因为各部门和机构各自不同的职能定位、发展需求和利益诉求，他们之间的数据交换、交易和共享问题日益突出，尤其是数据资源的权责归属划定、定价机制、智能交易、隐私保护和安全定责问题，已经成为阻碍交通大数据产业发展的瓶颈，迫切需要国家和行业出台一系列权威标准进行指导和规范，以便加快建立安全、健康、有序的交通大数据生态体系，促进大交通数据产业的可持续发展。

具体来说，建议制定《大数据确权定价模型和交易指南》，对大数据资源的权责归属进行清晰划分，形成有效的价格生成机制，明确交易规范，实现大数据资源的价值发现和有效流通；制定《大数据隐私保护要求和测评方法》，对涉及到用户隐私的信息明确保护要求，提供实施建议指南，并对保护等级提供高效测评方法，构建《大数据安全事件定责与取证规范》，对发生大数据安全事件后涉事各方的责任划定进行说明，并提供客观、可靠、可信的取证方法及其规范等。

B.5 联想大数据安全标准应用实践

联想通过 6 年时间，实现了全球 2 亿台设备及个人数据的实时采集、超过 200 余个信息系统数据的同步、混合架构下 9 个大数据中心及 2000 个节点的安全运维，在大数据及安全领域有了丰富的实践与经验积累，为联想自身及企业客户研发并搭建了成熟、安全、高可用的大数据分析与计算平台。

联想在大数据产品的设计、开发、实施、咨询及服务等环节中参考了《大数据服务安全能力要求》等标准。基于可信链技术，实现硬件、系统、软件、数据、应用多层次的安全防护，并对大数据采集、处理、存储、应用、销毁等全生命周期的安全管控。

基于《大数据服务安全能力要求》，联想在大数据产品研发上，首先从基础框架出发，不断对平台所包含的开源的底层组件，如 Hadoop、Spark 等的危险程序、弱安全配置、系统缺陷及渗透注入等多全方面漏洞进行检测与安全加固；其次，按照数据生命周期应用要求，实现采集、传输、存储、处理、访问、交换直至销毁的全周期安全保障。对核心敏感的数据，会对其进行发现、跟踪、加密、脱敏，在使用过程中遵循法律法规要求，并严格的管控数据的使用用途与方式。对于非核心敏感数据，也提供了基于 RBAC 授权管理模型，支持数据行列级权限控制。

根据标准对于大数据安全分级要求，联想大数据在解决方案上提供了基础设施安全、存储安全、系统安全、应用安全、平台网络服务安全、认证授权、审计与日志跟踪、安全事件监控等全方位的解安全保护。同时参考指南要求，帮助企业或机构建立大数据安全制度、原则、策略、管理方案以及实施细则，一方面使数据在资源整合、共享、发布、交换过程中满足相关合规性要求，也在个人数据、重要数据、数据跨境传输、密码及密钥管理提供合规性支持。所做工作如图 B.8 所示：

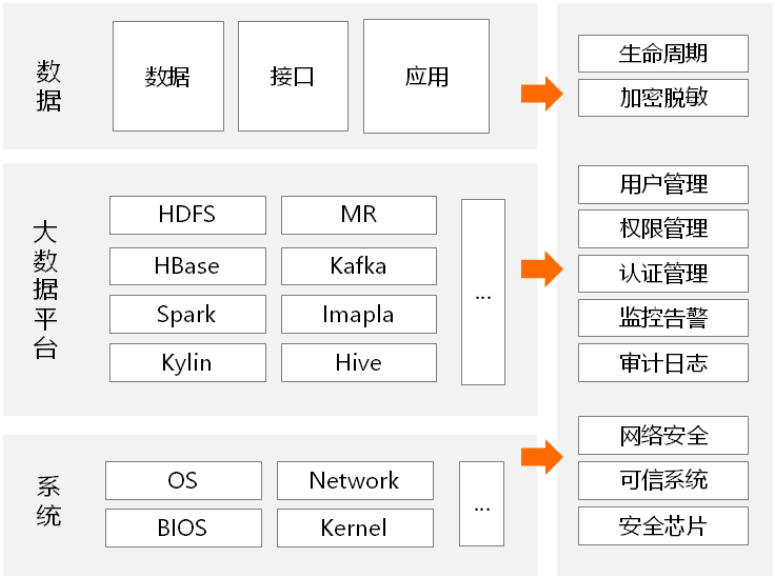

图 B.8 联想大数据平台安全方案

联想通过对标准的实践形成基于芯片级的大数据安全解决方案，基于可信度量理论，可建立完整的信任链。提供统一的身份认证与访问控制，基于 Kerberos 提供主机、服务、用户间的访问安全控制框架，并可以通过可信计算实现硬件级票据的生成与认证增强。实现全服务、全组件的高可用策略，不仅限于数据存储、计算的组件，包括运维、管理等组件也实现高可用，可满足 7x24 小时不间断运行的高可用需求。

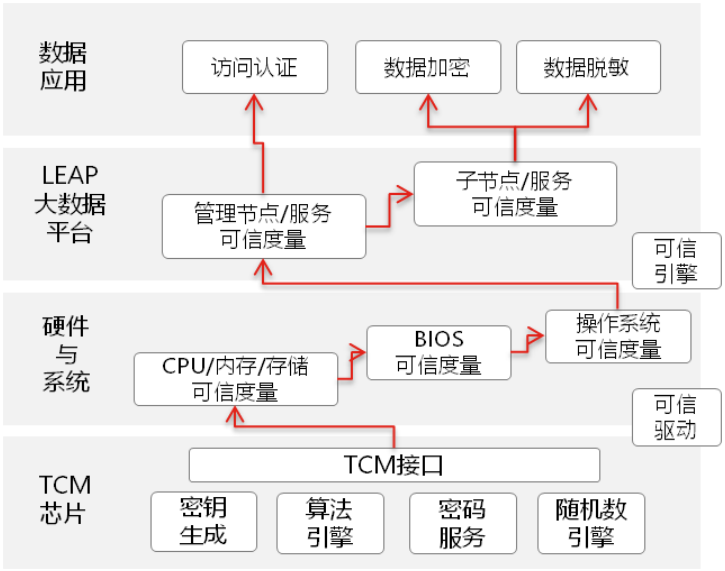

图 B.9 联想大数据可信计算方案

对标准的需求

建立起大数据安全标准以及一系列测评体系是实现大数据安全的重要支柱，大数据安全相关标准是企业用来度量大数据产品安全目标和大数据安全服务能力的尺度，也为大数据服务提供企业提供重要的参考。

大数据安全标准应提供用户数据安全保护目标、所属资产安全保护的范围和程度、用户安全管理需求；大数据安全标准应提供相应大数据服务安全能力的评估方法，并支持大数据服务的安全水平等级化，便于企业和用户理解和选择；大数据安全标准应规定大数据服务安全目标的验证方法，以便企业能够提供正确有效的证据证明其安全性。

从目前联想的大数据实践过程看，缺乏指导行业大数据安全的标准依据，比如在工业、医疗以及一些安全要求级别高的行业，这些行业对大数据的安全标准应该是有区别的，所以希望后续大数据标准制定考虑行业的应用属性，针对不同行业属性和需求，进行不同的安全级别要求，这样能更好的指导不同行业的大数据发展。

B.6 蚂蚁金服大数据安全标准应用实践

浙江蚂蚁小微金融服务集团有限公司(以下简称“蚂蚁金服”)深刻理解国家对大数据发展的战略部署，以大数据为核心能力，将数据安全视作公司发展的生命线，在从事金融业务的过程中，坚持依法合规、安全可控的数据安全策略，遵循《大数据安全管理指南》中职责明确、意图合规、质量保障、数据最小化、最小授权、数据保护等原则，不断提升消费者权益保护能力，持续完善数据安全治理体系。在此基础上，坚持将大数据作为创新发展的动力，一方面建立了基于大数据技术能力的数据安全体系，提升自身数据安全保障能力；另一方面，蚂蚁金服依托互联网技术及数据能力构建生态数据安全赋能产品，联合生态伙伴，共同提升生态数据安全能力。

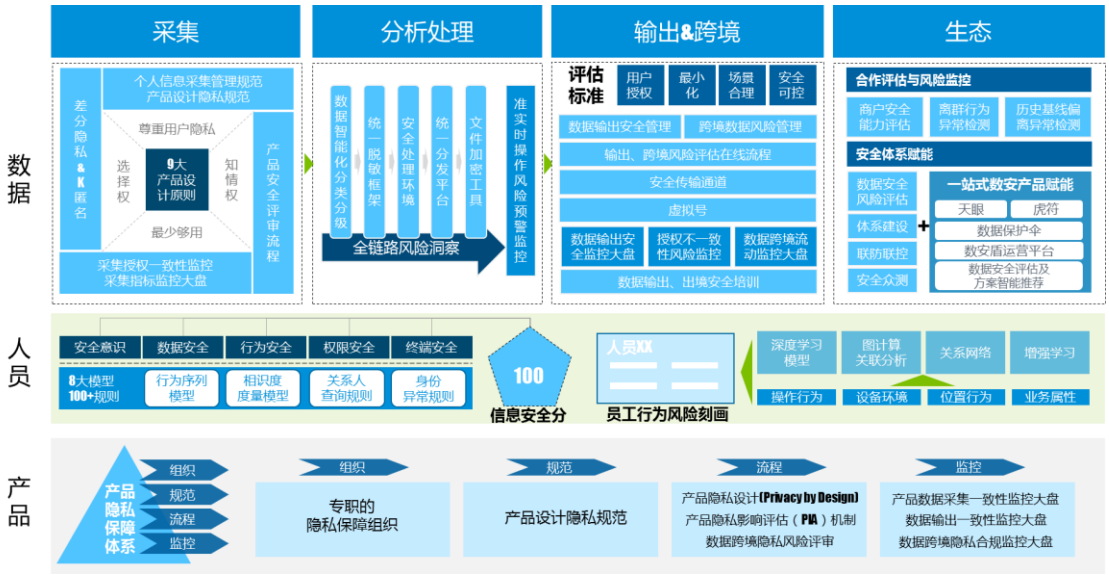

图 B.10 蚂蚁金服基于大数据技术的数据安全防控体系

一、坚持依法合规、安全可控的数据安全策略，持续完善数据安全治理能力

1）持续完善数据安全顶层治理。蚂蚁金服高度重视自身的数据安全体系建设工作，一直遵循业务沉淀、用户授权、合法正当必要的数据采集原则；坚持依法合规、安全可控的数据使用原则；在数据安全策略层面，成立了以高级管理层为核心的风险委员会，下设数据工作小组，确定数据安全策略，统筹集团层面的数据安全工作；在数据安全管理层面，设立了数据安全中心以及各业务部门数据安全责任人职责，全面负责公司内部数据安全风险管理能力建设，规划和推动各项管控措施落地，开展数据安全风险运营管理，确保 360 度、7*24 小时的全方位数据安全运营保障和风险监控；在数据安全执行层面，要求公司全员参与，持续加强员工安全意识教育，全面执行公司数据安全策略，深入落实本岗位的数据安全责任要求。

2）建立健全数据安全制度流程。蚂蚁金服基于 ISO27001 国际信息安全标准建立了信息安全管理体系，从方针策略、安全标准、安全控制指引、安全流程记录四个层面对数据安全管理的安全策略、规程、管控措施进行了明确说明和体系化定义，并针对技术领域的安全控制措施建立了一系列的安全基线。蚂蚁金服制定了《数据安全管理规范总则》和《数据分级规范》等制度，建立了严格的数据分级标准，以风险为本的指导原则，对所有

存量数据按风险等级（机密、保密、内部、公开）进行管理。在整个制度建设方面，现已发布 30 多个数据安全相关的制度流程，与集团业务运营紧密结合，确保在业务运营过程中的数据安全风险可控，数据使用有章可循。

3）建立数据安全内控体系和审计监督机制。通过内部制度明确公司整体的数据安全管理岗位及对应的职责，规范所有员工的数据使用行为，同时，以《个人信息安全规范》、《数据出境安全评估指南》标准为指导，通过统一身份管理、统一鉴权、统一日志等方式建立体系化的审计监督机制，利用大数据风险分析技术，建立了数据使用异常分析控制，及时识别业务运营过程中的数据使用风险，把对用户信息保护的责任落实到公司的日常业务运营过程中。

4）建立以数据为中心的风险管理体系。蚂蚁金服从数据、人员、产品三个方面重点进行风险管理体系建设。在数据方面，覆盖采集、分析处理、输出、生态等多个大数据管理重点；在人员方面，基于大数据安全管理需要，利用大数据、人工智能等技术，建立了信息安全评分及员工行为风险量化机制，准确识别和管控员工使用、处理数据过程中的各维度风险。另外蚂蚁金服从产品角度对用户隐私进行全方位保护，构建了评估产品隐私设计的组织、规范、流程、监控等一系列体系化控制机制。

二是构建生态数据安全赋能产品，联合生态伙伴，共同提升生态数据安全能力

生态方面，在加强自身数据安全能力建设的同时，蚂蚁金服也积极开展数据安全对生态的赋能。通过提升生态的数据安全能力，带动行业个人信息保护水平的整体提升。

1）生态数据安全风险监控

蚂蚁金服与合作伙伴合作的过程中，建立了一套完整的合作伙伴数据安全风险识别机制，通过敏感数据检测、调用历史基线偏离、离群行为等大数据异常检测技术，实现对生态合作伙伴端的敏感信息泄露等风险的监控。同时，也通过差分隐私和 K 匿名等技术措施提升个人隐私和数据安全保障能力。

2）生态数据安全风险评估和体系建设咨询服务

首先，蚂蚁金服建立了生态伙伴数据安全风险评估体系，定期进行安全排查并推动落实高风险整改。线上化、产品化的数据安全评估及方案智能推荐方案正在研发中。当前正在建设线上化的风险评估能力，为生态伙伴的数据安全评估提供更高效、更智能化的支撑平台。

其次，蚂蚁金服协助生态企业从组织职责、流程规范、平台技术等多个角度，推动其健全数据安全管理组织与管控体系，并通过日常风险运营支撑其安全体系落地。已协助多家核心生态伙伴完成信息安全管理体系规划、信息安全团队组建及技术能力提升。

最后，蚂蚁金服重视日常数据安全风险的运营工作，与核心生态伙伴建立了日常风险沟通、安全教育、应急演练等协同防控机制。

3）生态数据安全产品赋能

蚂蚁金服正在推动将自身数据安全能力产品化并覆盖蚂蚁金融云、阿里云等平台上的生态企业，已形成包括数据安全（数据保护伞）、权限安全（虎符）、人员风险（天眼）、风险运营（数安宝）等四大类产品。

随着《网络安全法》的落地实施，网络运营者须“采取技术措施和其他必要措施，确保其收集的个人信息安全，防止信息泄露、毁损、丢失”。个人信息安全的防护成为企业的重中之重，“个人信息去标识化”是保证用户个人信息在企业内部甚至企业间安全流转与共享的基础，蚂蚁金服已经在个人信息去标识化方面进行了模块化、自动化的实践，但同时也迫切希望能有相关标准依据，包括去标识化配套的检测方法，指导和提升企业用户个人信息安全防护工作的体系化和标准化程度，进一步提升互联网行业用户个人信息的安全保障能力。

B.7 南大通用大数据安全标准应用实践

南大通用参考大数据安全相关标准中的《大数据服务安全能力要求》的相关要求，综合考虑该标准的存储架构安全、数据交换安全、存储访问控制、数据归档管理和数据时效性管理等方面，设计了南大通用大数据分析平台。南大通用大数据平台融合了事务型数据库、分析数据库和 Hadoop 大数据平台产品，兼顾大规模分布式并行数据库集群系统、稳定高效的事务数据库，以及 Hadoop 生态系统的多种大规模结构化与非结构化数据处理技术，平台中所有组件都能提供基于 Kerberos 的认证功能，能够帮助企业快速构建安全的大数据平台产品。

一、存储架构安全。如图 B.11 所示，南大通用大数据分析平台采用开放式架构，融合交易型数据库、分析数据库和 Hadoop 生态产品，便于横向扩展系统规模，能够满足由于数据量递增带来的动态扩展容量需求。数据交换层在确保数据安全的情况下，提供高效的数据交换服务。在安全访问控制层使用多种信息安全技术，守护用户数据的安全。

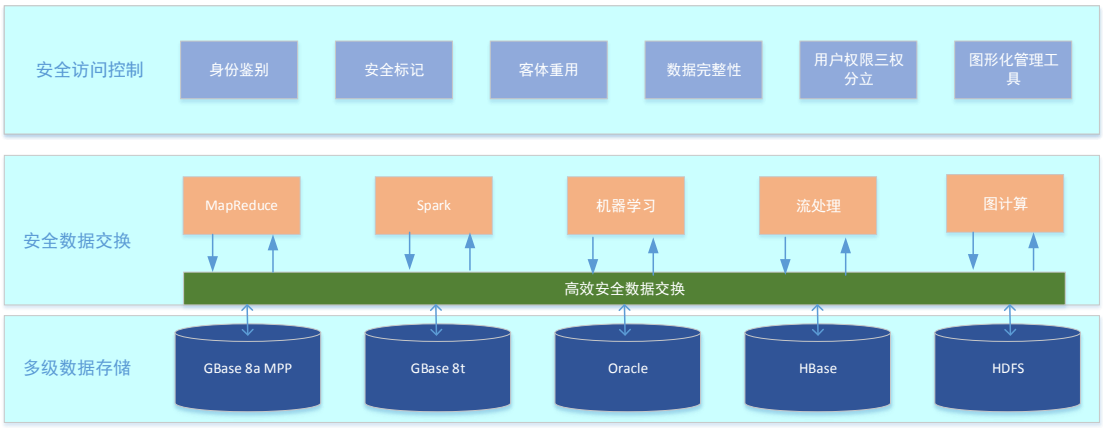

图 B.11 南大通用大数据分析平台架构

二、数据交换安全。在数据交换节点之间采用 SSL 安全通道，确保端到端的真正安全。计算节点对数据的每一次操作都需要经过安全的身份验证和加密，确保数据交换的安全可靠。在存储引擎之间交换数据采用加密 HDFS 作为中间媒介，先将数据导出到 HDFS 上，然后从 HDFS 加载到目标存储引擎或目标计算引擎。在 HDFS 上的中间文件采用内置或用户指定加密方式存储，以防止数据泄密。在需要流式处理的数据交换情景，采用安全 kafka 作为中间媒介,启用 Kerberos 认证模式，确保 kafka 集群的安全性。

三、存储访问控制。南大通用大数据分析平台采用用户密码以及数字证书双重认证的鉴别机制，使用杂凑算法保证密码自身安全性。采用自主访问控制机制，根据客体和用户生成访问控制表，当一个主体访问某个客体时，根据访问控制表确定该主体是否有该客体的访问权限。并提供强制访问控制机制，客体在主体之间共享的控制，将平台中的信息分类和分密级进行管理，以保证每个用户只能访问到那些被标明可以由他访问的信息。实施三权分立的互相制约的机制，防止滥用数据库超级用户特权的安全漏洞。

四、数据归档管理。如图 B.12 所示，南大通用大数据分析平台采用双活集群方案架构，主集群和备集群之间的数据同步是利用同步工具，在特定的时间点完成数据同步工作。主集群主要承担每日的数据跑批作业、准实时数据的复杂查询分析；备集群存储主集群的归档数据，一旦主集群故障立即代替主集群提供服务。双活集群系统根据上层应用的特征分区隔离，为上层应用提供安全可靠的数据服务。

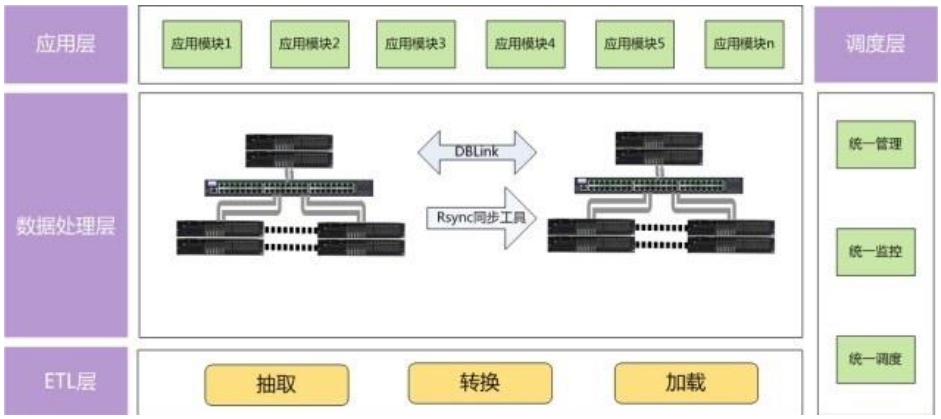

图 B.12 南大通用大数据分析平台双活集群方案

五、数据时效性管理。南大通用大数据分析平台为不同时效性的数据建立分层的数据存储方法，按照时效性自动迁移数据到不同的存储引擎，确保大数据用户能高效地获得有效数据。建立过期存储数据的安全保护机制，将超出有效期的存储数据自动备份到 HDFS 存储引擎，如若再次需要使用超出有效期的存储数据，需要再次获取数据提供者授权。

六、安全审计。南大通用大数据分析平台具有独立的审计系统，能定义有关的审计事件，记录用户的有关操作，记录用户标识、身份鉴别、自主访问控制中的有关审计数据，进行相关的审计分析并自动报警，并对审计数据进行查阅。另外需要保证审计记录在数据库系统运行时的访问控制安全，即只有合法的管理员才能够访问到符合其自身权限范围之内的审计记录，但审计记录的安全性还可能受到其他安全威胁，主要包括：用户冒充、物理窃取等。冒充的安全隐患通过采用 SSL 安全套接字协议可以大大降低，通过强化的用户身份鉴别可有效的预防用户冒充的威胁。物理窃取的安全问题仍然采用数据加密的方式来解决，即对独立的审计库进行单独的加密处理，从而保证物理存储的审计记录是密文态，即使发生物理文件失窃也能保证审计记录的安全性。

审计数据的加密机制与普通用户数据的加密机制有所区别，普通用户数据的加密设置由用户自行指定，如指定哪些表或哪些字段进行加密，而审计记录的加脱密处理则由数据库内核自行处理，鉴于审计记录的数据量通常会很大，因此系统缺省采用按表加密的方式，一表一密，使用对称加密算法进行加解密处理，管理员仅需自行设定密钥变更周期等加密策略。通过如上安全策略，实现了对于审计记录的加密处理，以较小的代价增强了审计记录的安全性。

七、数据安全保护。需要保证数据库内数据的安全，防止数据被窃取，需要支持数据存储加密功能，数据存储加密功能采用库内加密的方式，在数据库管理系统的内核存储引擎进行数据加解密处理，从而对于合法用户来讲是完全透明的，因此称为透明加密，用户创建表的语法支持指定加密关键字 ENCRYPT，ENCRYPT 关键字可以出现在表属性中，也可以出现在列属性中，以支持表级及列级加密；在后端修改数据存储流程，在数据实际写入存储文件之前，将数据以数据块为单位调用加密函数进行加密，然后再将加密后的数据写入存储文件中，此后在读取数据文件中的数据时，先将数据解密，再提供给数据库上层使用。

八、数据通讯安全。南大通用大数据分析平台通过实现对 SSL（Secure Sockets Layer）安全套接字协议的支持来保证数据通讯的安全性，通过 SSL 协议提供了数据库服务器与客户端之间的可信通信路径，实现了用户与南大通用大数据分析平台服务器间的安全数据交换。SSL 安全套接层协议为 TCP/IP 连接提供数据加密、服务器认证、消息完整性以及可选的客户机认证，SSL 把对称加密技术和非对称加密技术相结合，可以实现保密性、完整性、认证性三个通信目标。

九、资源限制。可以对用户按最大限额的要求，进行资源的管理和分配，确保用户和主体不会独占某种受控资源（内存、硬盘），另外提供资源使用预警功能，提供配置参数对剩余磁盘空间大小进行设置功能。提供限制数据库某项资源的配额，使数据库无法使用超过该配额资源的功能。提供配置某项资源的阈值，当数据库资源使用超过该阈值具备报警功能。提供限制系统并发会话最大数量的功能。提供会话超时功能，能够限制用户会话活动的超时终止时间。

十、数据完整性。支持数据防篡改特性，对存储数据块中的数据进行校验，能够检测数据的完整性。支持数据传输过程中的完整性,利用密码算法和散列(HASH)函数，通过对传输信息特征值的提取来保证信息的完整性，确保要传输的信息全部正确完整的到达目的地，可以避免服务器和客户机之间的信息受到破坏。

十一、数据可用性。数据的可用性包括提供数据备份和灾难恢复两个方面。南大通用大数据分析平台的数据冗余可以通过副本来实现。另外可以进行数据的全备和增备来进行数据的备份和恢复。

从目前南大通用的大数据安全实践过程看，大数据安全首要考虑的是基础设施安全，关键取决于搭建大数据平台的技术及数据库的安全技术。南大通用大数据分析平台从访问控制、安全传输、存储加密等方面增强了数据安全能力，在基础设施方面消除大数据安全隐患。从实践过程发现，当前缺乏一套能够对大数据基础设施安全验证与测试的权威、全面、面向商用的测试标准和方法，企业面临着无标准测试依据的境况，建议加快制定大数据安全测试标准方面的工作，同时能够细化对基础设施安全的测试项。

B.8 启明星辰能源大数据安全标准应用实践

能源行业信息网络是国家关键信息基础设施，关系到能源行业安全稳定生产，是国家信息安全检查和保卫的重要行业。能源行业大数据具有明显的三个特征：海量（Volume）、高速（Variety）、多样（Velocity）。目前能源信息网络中的网络行为数据呈现出了大数据化，并且还具有网络行为数据专有的特征。网络行为大数据的处理更需要结合网络行为大数据化的特征进行相应的处理。因此，为了提升能源行业网络与信息安全防护水平，确保国家关键信息基础设施安全，面对海量事件管理带来的挑战，启明星辰集团率先利用国内领先的高性能日志采集范式化技术、大数据分布式存储与索引技术和流式分析技术等大数据技术，应用在某重要能源企业信息安全领域的实践获得成功，探索出一款能够支撑能源大数据安全数据处理的，具有能源行业特性的安全管控系统。具体如图 B.13 所示：

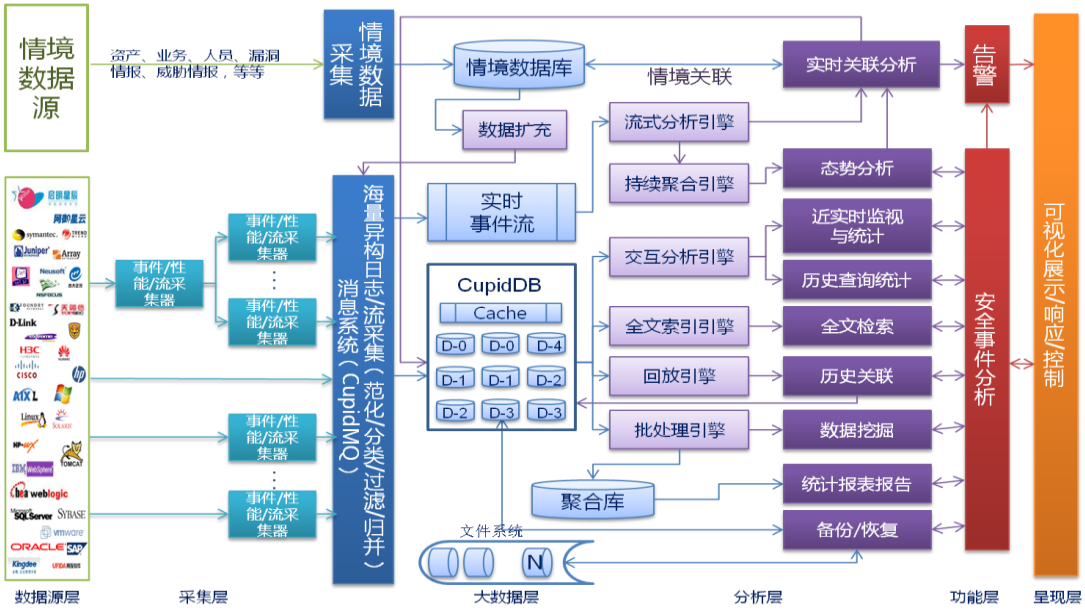

图 B.13 启明星辰大数据分析安全框架

系统的安全分析技术架构从总体上划分为五个部分，分别是：信息采集（Collection）、大数据存储（Big Data）、信息分析（Analysis）、功能层（Function）、呈现层（Presentation）。

一、信息采集，主要包括事件/流/性能采集和情境数据采集。

二、大数据存储。针对大数据安全事件（日志），启明采用了具有自主知识产权的分布式非关系型数据库 CupidDB，从根本上解决了采用传统关系型数据库的安全管理平台的性能瓶颈。

三、信息分析。针对采集上来的各类安全要素信息，系统实现了性能与可用性分析、配置符合性分析、大数据安全事件分析、流行为安全与合规分析、宏观态势等分析。信息分析的方法包括实时流式分析、交互式分析、历史数据批量分析和数据回放等多种先进技术。

四、功能层。实现日常安全管理的功能，对发现的安全问题进行处置，包括例行处置和应急处置。

五、呈现层，系统为不同层级、不同角色的用户提供了层次化的用户视图，从多个维度进行展示，帮助安全分析师快速获取安全数据，进行威胁发现；帮助管理层了解网络的整体安全态势，及时掌握安全态势，以求做出清晰、有效的决策等。

从功能方面，系统主要从六个方面开展安全防护：

一、身份认证：对访问来源的合法性进行验证，确保所有来源身份合法；

二、访问控制：控制访问的条件，只允许在合法的条件下访问大数据系统；

三、授权鉴权：配置用户（角色）可以访问的数据资源以及对数据资源执行的操作；

四、数据脱敏：配置数据资源需要脱敏，实现事中脱敏；

五、应用审计：记录用户访问的应用，便于事后审计；

六、数据审计：记录用户访问的数据，便于事后审计。

系统的框架如图 B.14 所示:

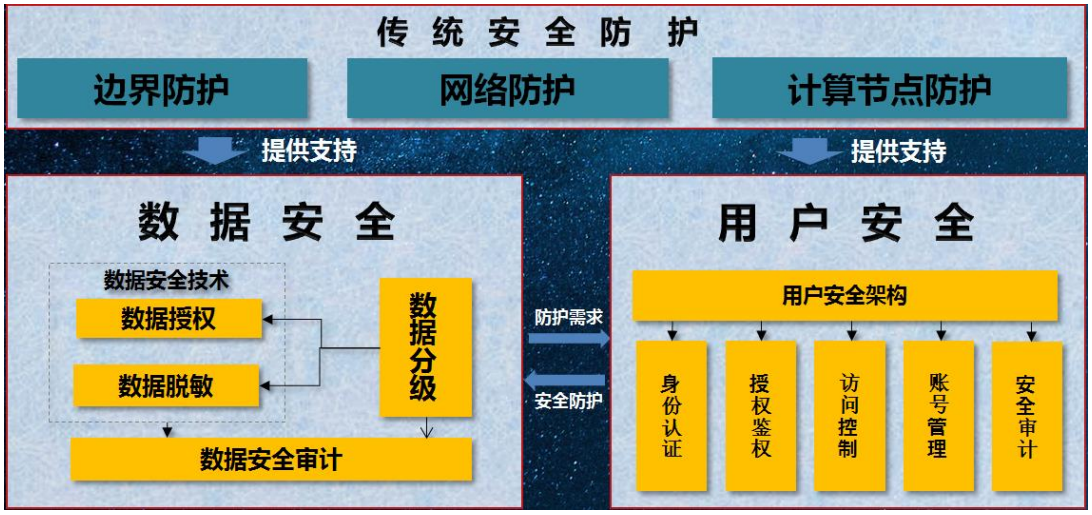

图 B.14 启明星辰大数据分析安全框架

面对日益严峻的网络安全环境，系统可以实现：

1） 代理式访问：对平台无干扰，无需修改平台配置，可以拥有更灵活的访问管理，且能充当大数据系统的一层防火墙；

2） 实时脱敏：与应用层无关，对上层应用开发没有任何影响；

3） 审计全面：支持对大数据平台安全管控系统的应用事件和日志进行管理，收集日志，并对日志进行标准化、过滤，实现上发、检索、生成统计报表等功能；

4）组件支持全：支持常见大数据平台 SQL 类组件，支持 HBase、HDFS 的代理访问。

在产品研发等过程中，如在大数据安全管控系统的研发、数据分析等过程，主要参考 NIST SP 1500 系列标准，如 Big Data Definitions 、Big Data Taxonomies、Big Data Use Cases and Requirements、Big Data Security and Privacy 、 Big Data Reference Architecture 、Big Data Standards Roadmap 等。

在产品应用方面，公司大数据安全管控系统目前已经成功运用到能源、政府等多个行业中，在 2016 年的某公司大数据模糊化及数据安全加固项目中，主要依据《信息安全技术云计算服务安全能力要求》、网络安全等级保护等相关规范和标准进行项目实施：（一）建立大数据平台安全管控规范。根据大数据平台的业务要求及特性，结合两部委、等保、公司安全要求从安全策略体系、安全管理体系、安全运营体系、安全技术体系、安全评测体系进行整体设计及大数据平台安全规范管理，建立大数据平台安全总体架构，逐步建设和完善企业大数据安全管理和使用规范；（二）对大数据平台 IT 资产进行全方位安全加固。对大数据平台的网络设备、安全设备、中间件系统、操作系统、Web 应用系统进行安全基线配置、检测及加固，以达到安全检测要求；（三）实现大数据平台安全审计。通过部署大数据安全审计平台，采集用户日常业务操作日志数据进行建模分析，从而可以有效审计对于敏感数据的操作行为，加强针对内部及外部用户的审计、取证与事后追责；（四）建立大数据平台敏感数据防泄露及模糊化的综合防护能力。对大数据平台上的敏感数据进行识别、分级、分类；通过配置不同的脱敏策略实现 hive 和 mpp 平台（基于 vertica 数据库）访问数据的实时脱敏等四方面的建设总体实现从大数据从采集、加工、存储、开发、使用，优化全流程的安全管控，实现完整的数据生命周期管理。通过安全技术加管理的方式实现大数据平台全面的安全防护。

在实施过程中，对于大数据安全标准的主要需求如下：

一、大数据安全服务能力要求（目前报批阶段）。在提供安全服务过程中，尚未有统一的衡量标尺和参考依据，大数据安全服务能力要求类标准能为安全服务能力提供实施和评测依据。

二、大数据安全体系架构类标准。在系统建设过程中，尚未有对大数据安全体系架构的规范要求，以至于产品的体系架构多种多样，难以统一，给研发人员带来挑战的同时对用户的统一管理造成了一定的负担。

三、大数据安全类标准。数据安全对于大数据安全管控系统来说至关重要，在数据共享、数据融合、数据保护过程中，如何合理的采取保护措施，是数据安全的一大挑战。在实际实施过程中，数据脱敏、数据安全分类分级等标准能给数据的分析处理、数据存储等过程提供依据，为能用有限的资源为更多的数据提供保护，实现利益最大化提供保障。

B.9 勤智数码互联网金融大数据安全标准应用实践

互联网金融安全关系到国计民生，勤智 DeepOneTM BDP 互联网金融大数据监管平台通过对互联网金融主管机关业务数据、政务数据共享交换平台、互联网金融企业的经营管理数据、互联网数据和第三方数据的采集与融合。采用可信接入保护机制，对上述互金数据在采集、传输、存储、交换、处理、销毁等过程中加以保护，确保数据的完整性、保密性和可用性，并提供标准化访问接口。监管者通过平台对区域内的互联网金融产业实现业务监管、风险预警以及风险处置等功能，提高互联网金融监管机关业务监管的有效性，防范行业风险和区域风险，促进区域金融行业的健康规范发展。

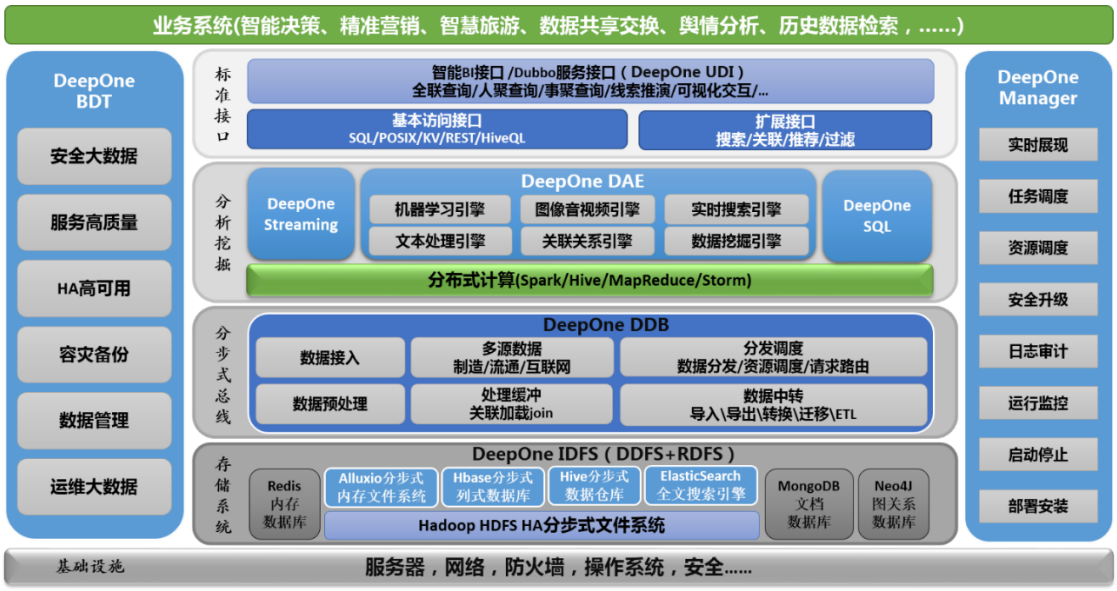

图 B.15 DeepOneTM BD 勤智互联网金融大数据监管平台产品架构

勤智 DeepOneTM BDP 互联网金融大数据监管平台从几个方面设计和架构了安全

（1）系统架构安全，DeepOneTM BDP 以安全管理中心为核心，构建对应的安全计算环境、安全区域边界和安全通信网络，确保系统能够在安全管理中心的统一管控下运行。设计过程中参考了国标《GB22239 信息安全技术 信息系统安全等级保护基本要求》、《GB25070 信息安全技术 信息系统安全等级保护设计技术要求》等标准。

（2）访问认证安全，平台基于用户和角色的认证统一体系，遵从帐户/角色 RBAC（Role-Based Access Control）模型，实现通过角色进行权限管理，对用户进行批量授权管理。

（3）数据安全，提供数据冗余、数据备份、在线副本等多种方式，确保用户数据的安全性。同时采用国产密码算法对数据进行加密，由于涉及到互联网金融机构的业务数据，信息安全显得更为重要。平台遵循了《GB T 35273-2017 信息安全技术 个人信息安全规范》、

以及《网络安全法》的相关要求，从而起到保护金融信息安全的作用。

（4）不间断运行，平台服务节点实现 HA（High Availability），将故障对业务的影响降低到最小程度。HDFS、Hbase、Spark 等服务实现高可用可视化配置，解决了大数据处理中的各种不稳定的问题，保证业务 7×24 小时不间断运行。

（5）存储安全，实现集群异地灾备，当业务需要灾难迁移时，容灾中心快速扩容至可满足日常业务需求的 IT 容量，并启动接替业务中心工作，降低企业 IT 的容灾成本。

（6）审计安全，审计记录时间由系统内唯一确定的时钟产生，审计记录包括事件的日期和时间、用户、事件类型、主体标识、客体标识和事件是否成功及其他与审计相关的信息。

为了规范互联网金融大数据的发展，推动互联网金融大数据的技术进步、提高互联网金融行业管理水平。勤智数码结合自身在互联网金融大数据的安全实践经验，建议在国家统一的标准规范下，制定在《互联网金融大数据服务安全框架》、《互联网金融大数据信息安全指南》等方面符合我国互联网金融大数据发展所需的相关标准，指导互联网金融从业单位和监管机构进行相应的安全建设和安全运维管理。

B.10 三未信安大数据安全标准应用实践

大数据安全隶属于数据安全的范畴，具备传统数据安全保护的特征，例如：大数据的数据保密性，大数据的可信性，大数据的隐私保护，大数据的访问控制都是密码技术重点解决的问题。

密码技术是一种基础的安全支撑技术，很多安全场景的实现都依赖于密码技术的合理有效应用。大数据安全面临全新的安全挑战的同时，基于密码学的全新的解决手段也在更好的解决大数据安全带来的新问题：数据发布匿名保护技术，数据水印技术，数据溯源技术，角色挖掘技术，自适应访问控制等。可以说，大数据安全是绕不开密码技术的应用的。但是在上述技术的发展中，不同的大数据系统采用的密码技术没有统一的标准可以遵循，对于数据的保护能力也不一而同，往往会对大数据的安全保护带来截然不同的保护效果。同时，大数据安全标准在密码技术的应用方式和应用场景上还是空白，未能很好的做到规范和指引作用。

北京三未信安科技发展有限公司经过初步研究和技术实践，认为大数据安全与密码技术需要有机的结合，通过合适的密码技术采用和现有密码技术标准的改良，来有效的规范大数据安全技术，提高大数据安全的风险控制能力。因此，通过大数据中的密码技术应用标准，规范大数据安全领域的密码技术使用方式方法，是大数据安全标准中对安全技术方面及其必要的补充，是对大数据安全标准及其有价值的工作。

三未信安致力于通过密码技术为大数据领域提供数据安全保障。为完善大数据平台安全能力，确保大数据平台对用户数据的管理和控制权，三未信安与国内某公有云在大数据平台建设的安全实践中，首先参考依照大数据安全相关标准现有的《大数据安全管理指南》要求，从该标准的合规原则和安全存储要求两方面入手，提出保护大数据平台内数据安全的可行方案。同时，将现有的商用密码技术标准进行改进，应用于大数据安全领域，数据使用权通过认证及访问控制等密码技术手段控制在数据所有者手中，从而符合标准要求，解决大数据平台的数据安全存储和安全访问问题。

在诸多 Hadoop 大数据安全实践中，普遍采用 Kerberos 作为服务的安全认证方式。对敏感数据通过在大数据系统中建立加密区并启用密钥管理服务进行加密存储，加密区的密钥则使用密钥文件存储于系统目录中。上述安全实践通过身份认证、数据加密等手段提高了 Hadoop 等大数据系统的安全性，但在管理、安全合规及实践应用中还有很多不足。首

先，在大规模部署分布式系统时将会对密钥分发中心带来管理和控制上的挑战；其次，对加密密钥的保护安全性考虑不足，且无法对密钥进行生命周期及策略上的控制；再有，加密区中的数据加解密会对系统带来性能上的极大损耗，拖累数据分析响应时间；另外，这套基于密码技术的安全体系来源于 Hadoop 等各类平台本身的实现，在合规性上也必然存在不符合国家密码监管部门规范要求的问题。

针对这些问题，三未信安结合密码技术标准，在以 Hadoop 为原型的大数据安全体系进行了一系列的探索和功能补充，所做工作如图 B.16 右半部分所示：

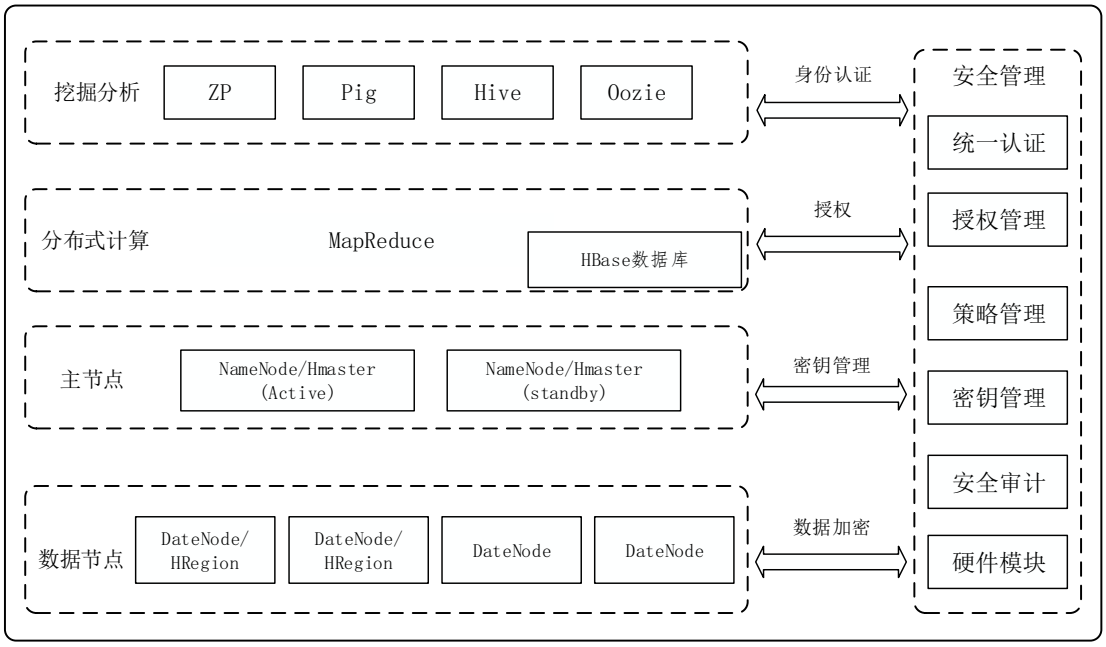

图 B.16 三未信安大数据平台安全架构

一、大数据平台认证授权标准的完善

有效的身份认证与访问控制是确保用户数据不被非授权访问的关键。三未信安结合商用密码标准 GM/T 0014-2012《数字证书认证系统密码协议规范》，通过分布式的身份认证技术的引入，可对 Pig、Hive、MapReduce、DataNode 等服务节点进行统一认证管理。该认证体系通过密码技术，与原有的 Hadoop 认证体系完全独立，确保节点不会被冒充，保证集群中的服务的身份可靠性。其中认证服务模块支持密码硬件生成认证票据，并实现了国产算法在大数据领域的应用，符合国家密码相关政策标准的要求。

二、大数据平台适用的密钥管理标准的应用探索

三未信安深入研究了大数据平台中的加密密钥管理需求，结合商用密码标准 GM/T 0051-2016《对称密钥管理规范》的要求，将传统对称密钥管理与大数据中的密钥分布式使用，统一授权的的使用硬件安全模块保护大数据主密钥安全的高可用、可扩展的密钥管理服务。同时，三未信安引入国际标准的密钥管理互操作协议（KMIP），将大数据系统中密钥管理操作与国际标准的密钥管理服务对接，可完成对密钥生成、获取、注销、归档、更新等全生命周期管理，并加入密钥访问时间、访问用户等策略配置，提高密钥安全性。

三、大数据平台的数据加密标准实践

结合 GM/T 0009-2012 《SM2 密码算法使用规范》，GM/T 0010-2012 《SM2 密码算法加密签名消息语法规范》的要求，三未信安将支持国密算法的密码服务模块加入 Hadoop 大数据平台的密码服务中，完成与国密算法的无缝调用。对认证完成后的用户通过调用三未信安自主实现的高速硬件密码模块，完成对数据的加解密操作,大幅提高加密区内数据存取性

能，从而保证标准所采用算法的技术可行性。

通过上述实践，三未信安对现有的大数据安全标准进行了如下几方面的完善和补充：

1. 对现有大数据安全标准中的密码技术部分进行了实践，总结归纳最适宜的密码技术使用方式，可形成具备实践基础的大数据标准密码技术使用指南。
2. 对现有商用密码技术标准在大数据安全领域的应用进行实践，依据实践结果对传统标准提出大数据安全实际使用中的改进意见。
3. 针对国产密码算法和大数据等应用等特点，研究解决并规范大数据安全实现过程中如何更好的使用国产密码算法、满足用户的安全需求、增强系统间的互通性等问题。通过数据积累为算法在性能及可用性方面的可行性提供依据。

目前的实践工作仍有进一步演化和改进的空间，这主要表现在两个方面：

1. 密码技术领域新技术的应用，可以进行更多的研究实践，以保证技术标准的与时俱进；
2. 更细粒度的安全工作，如用户访问控制、数据权限管理及数据脱敏、数据溯源等安全工作仍需要完善，后续可以根据大数据安全中的《数据溯源描述模型》《数据脱敏指南》等相关标准对功能进一步加强。

三未信安通过大数据安全产品的完善及标准实践工作，进一步总结提炼了大数据安全与密码技术的契合点，为大数据安全中的密码应用标准的后续补充提供了宝贵的实践积累和应用指南。随着实践的积累，可以在大数据安全技术标准范畴中，适时的推出“大数据安全密码技术应用标准”，以更好的补充大数据安全中技术环节的标准体系，打通密码技术与大数据安全技术的标准体系联系，规范指导大数据安全的密码技术应用模式。

B.11 腾讯云大数据安全标准应用实践

腾讯云大数据安全方面的标准应用积累了丰富的实践，凭借在安全领域的丰富经验，腾讯云搭建了多层次全方位的纵深安全防御系统，为客户提供一个绝对值得信赖的云平台。腾讯云将数据安全的理念融入每一个产品的需求设计和开发过程中，并贯穿产品运营的每一个环节。同时，也积极应用多种国内外安全标准提升和展示安全能力。腾讯云的安全保护和控制流程，例如数据分类标准、访问控制策略、虚拟化安全策略、运维安全控制等安全内控标准，均已经通过多个权威第三方独立安全评估的验证。

自 2014 年国内首家云服务商获得 ISO27001：2014 认证以来，陆续运用和实施了多项安全标准，2016 年通过贯彻以 CSA CCM（云安全联盟 云安全控制矩阵）为基础的 STAR 云安全体系并获得金牌评估结果，提升了安全标准的运用成熟度。2017 年 5 月公有云和金融云分别通过了网络安全等级保护三级和四级评测，验证了腾讯云内部的安全管理体系的效果和效率。为了将腾讯云打造成为客户值得信赖的云服务供应商，腾讯云在客户云端数据安全保护的实践中稳步前行，内部流程持续优化，控制技术不断更新。腾讯云的数据安全实践以强大的安全研究团队为核心，通过专业的安全运维团队提供 7*24 小时的服务支持，并建立了独立的安全合规团队，紧跟不同行业、领域、国家的合规性要求。

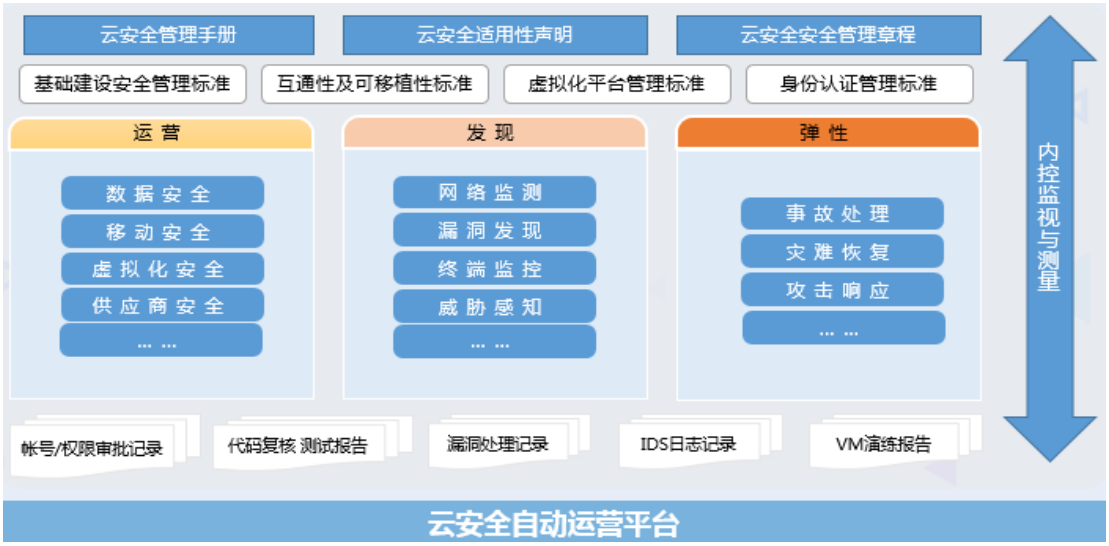

图 B.17 腾讯云安全标准体系示意图

图 B.17 所示安全标准体系的应用，为腾讯大数据安全提供了保障，同时也为腾讯利用大数据技术实现安全控制目标提供了基础。

互联网企业的业务安全现状

互联网业务在近些年高速地发展，互联网已和多个领域都发生了深度的结合，对社会的进步产生了巨大的推动作用。然而在一片繁荣的发展背后，国家和企业都面临着黑色产业链带来的巨大风险。比如说在社交产品中出现大量诸如色情、政治反动、欺诈钓鱼等违法内容，开始极大地危害着国家的网络安全。然而随着国家和企业对互联网投入的加大，黑产有越演愈烈的趋势。如何做好业务的风控，成为了国家、企业需要重点解决好的问题。

用户在互联网产品中产生了海量的行为，每一次行为都蕴含了大量的数据信息，并且每一笔行为涉及到主体（比如说账号、IP、设备等）也拥有非常多的属性信息。就腾讯的业务来说，目前每天用户会产生数万亿条即时通讯消息、数十亿条社区消息、数亿张相片上传。如何应对这种量级数据的接入、存储、管理，并通过数据的有效挖掘，生成业务风控模型以保障业务的安全至关重要。

腾讯多年互联网业务安全运营的经验证明，大数据风控技术充分运用大数据的挖掘，完全可以应对以下难题：

1）恶意的有效刻画

使用画像大数据对用户行为进行全方位的刻画，为风控模型提供坚实的高质量、规范化数据基础，而不只仅仅是基于单笔的用户操作。

2）恶意变化的态势感知

通过人工智能的方式生成业务风控模型，而不是采用人工拍脑袋、简单数据分析的方式；通过人工智能的方式自动感知风险变化的趋势，而不是采用人工反馈的方式。

3）降低人工在恶意对抗中的介入

使用大数据处理技术有效应对每天万亿级别的消息量，而不是采取降级处理的方式。

腾讯云基于腾讯大数据安全实践，实现云端大数据业务安全

腾讯云以大数据为基础的业务安全风险控制领域，已经推出了八大功能，分别为：活动防刷、注册保护、登录保护、注册码、消息过滤、图片鉴黄、URL 安全扫描、金融反欺诈；实际使用的企业客户已经达到了上百家。

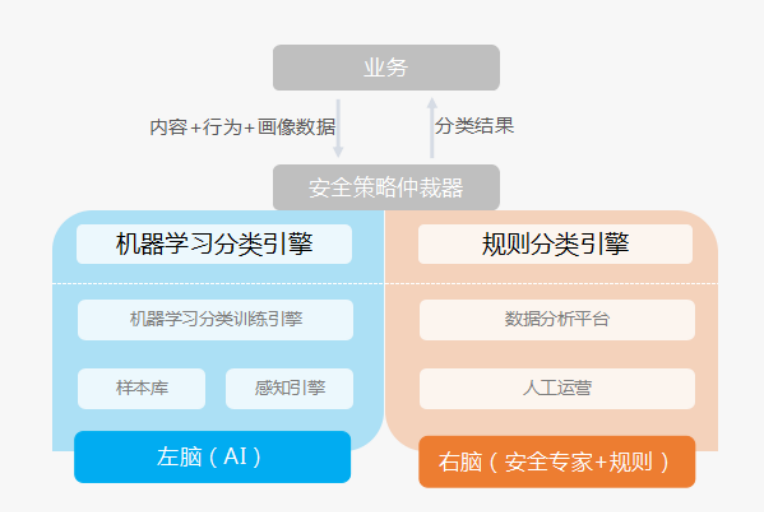

图 B.18 腾讯云基于大数据业务安全风险控制模型

业务安全控制能力已经从原始的状态进化成了以大数据为基础的业务安全风险控制，主要的特点体现在以下几个方面：

1）业务风控能力

企业的风控能力从刚开始基于 iptables、配置文件的黑名单风控策略，演进到了基于机器学习+人工智能的风控策略，并且开始通过画像系统对主体（比如说账号、手机号等）的历史信息进行信息沉淀，以更好地刻画恶意；企业的恶意发现能力已经从基于用户的投诉反馈，演进到了通过态势感知自动发现恶意。

2）数据处理能力

数据处理能力已经从原始的单机器处理 MB 级别的数据，演进到了能够通过多机器分布式并发处理 PB 级别数据的地步；每日处理的实时消息量，已从初期的每日几千条达到了目前每天万亿级别的量级。

3）大数据技术

互联网上每天都需要应对超过万亿级别的用户行为，可称为“行为大数据”。每一笔行为的背后，还蕴含着海量的属性信息。

为应对风险，企业需使用“大数据技术”来对“行为大数据”进行有效的处理，以建立“画像大数据”（也即是高质量、标准化的行为描述信息）；使用“大数据技术”中的人工智能方式生成风控模型，对风险进行高精度的刻画和对抗，并实时感知风险的动向。

腾讯云基于大数据业务安全风险控制框架

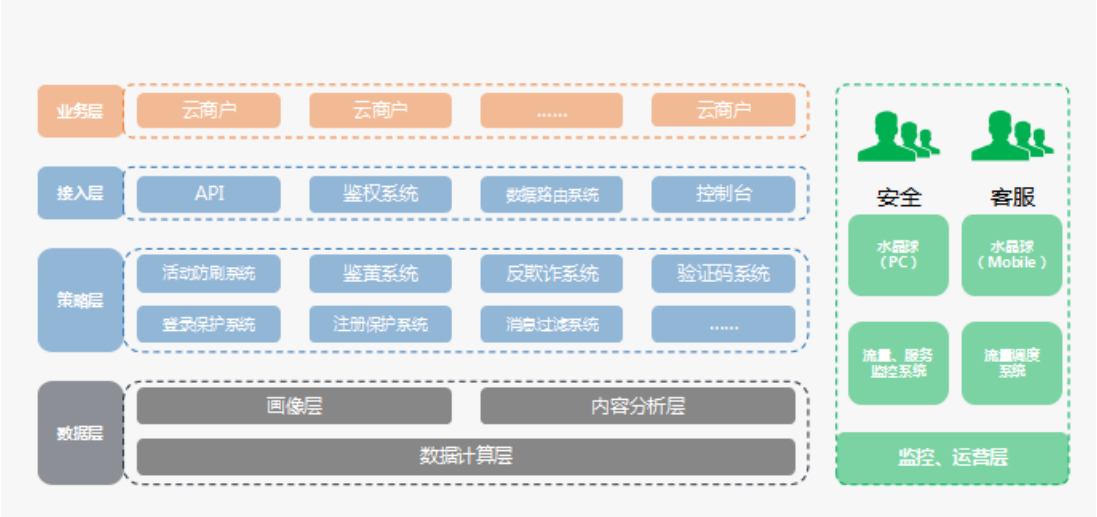

图 B.19 腾讯大数据业务安全风险控制框架

1）数据层：构建画像系统，特别是用户的画像信息，为区分恶意构建标准的底层数据；构建大数据处理系统，以应对海量用户行为数据的处理；

2）策略层：结合画像数据，对用户的操作进行综合判定；对新型的恶意变种进行感知和告警；对用户行为的重要要素进行恶意分析和判定，这些基本要素包含但不仅局限于：文本、图片、视频、URL、行为；

3）接入层：如何让企业方便地使用大数据风控，并保证数据的安全；

4）安全专家团队：对安全系统、策略、样本库进行运营。

腾讯云 2017 年牵头标准研究项目《大数据业务安全风险控制实践指南》

利用大数据技术保障业务安全的领域，目前标准化还是空白；国际上，包括美国 NIST 在内，也缺少对这一领域的标准化研究成果。实际上，利用大数据开展业务安全风险控制，互联网行业已经有充分的良好实践。腾讯云的天御，产品化产业化应用已经非常成熟，在视频直播、金融与互联网金融、游戏、电商等行业应用带来了显著的社会效益。基于这方面的实践，经过抽象分析和理论总结，腾讯云 2017 年牵头开展国标研究项目《大数据业务安全风险控制实践指南》，这个标准研究的成果，不仅适用于互联网行业，也将广泛适用于传统行业面向互联网转型时，遇到的层出不穷的业务安全风险，向各种类型的组织，提供解决业务安全风险的标准规范的思路。

B.12 医渡云大数据安全标准应用实践

随着信息技术的飞速发展,全球正在进入 DT 时代，医疗大数据已经成为我国的重要战略资源。与此同时,大数据在安全方面面临着诸多挑战，安全问题也日益凸显。医渡云（北京）技术有限公司（以下简称“医渡云”）为保障医渡云医疗大数据平台能够在安全稳定的环境下为医院提供数据服务，结合大数据基础平台安全要求，参照大数据基础平台安全框架和各项设施安全要求，围绕医渡云大数据平台完成安全实践与建设。

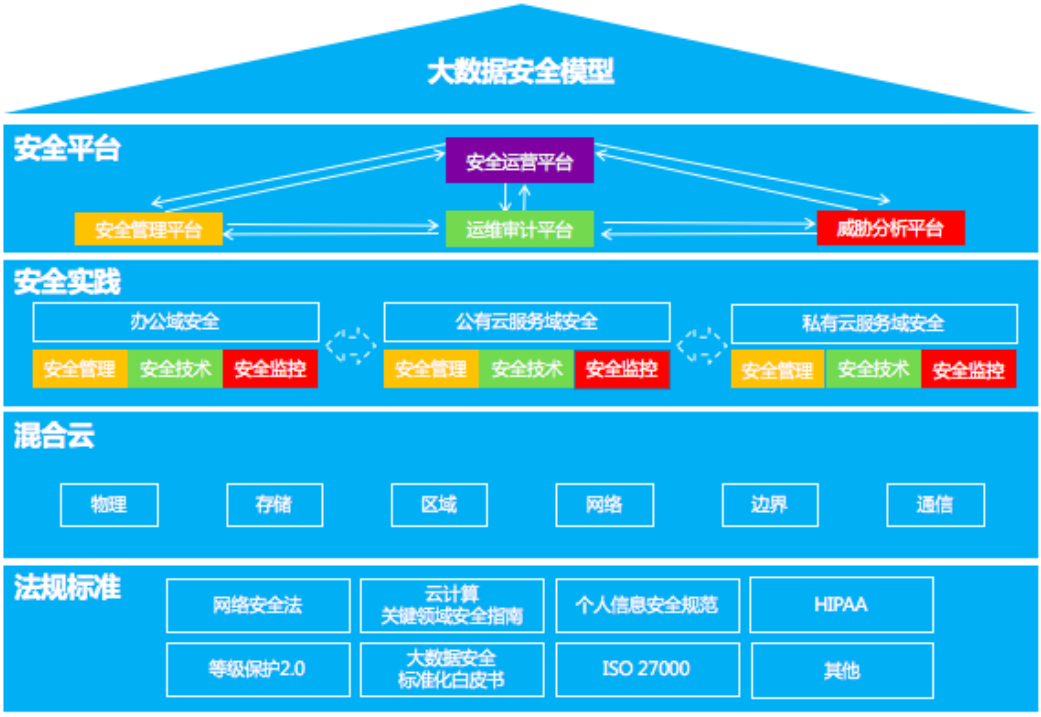

图 B.20 医渡云大数据安全模型

医渡云大数据平台利用先进的机器学习和人工智能技术，对医疗数据（HIS、CIS、EMR 等依赖的传统关系型数据）进行集成、挖掘、利用，辅助开展新型临床、科研、医院管理等服务，助力医疗机构和医生，为患者提供更好的医疗服务。综合考虑法规要求、业务需求、安全需求和客户实际需求，医渡云大数据平台以私有云独立隔离部署的方式部署在医院内部，医渡云作为大数据平台服务提供商，负责大数据平台开发安全等重要工作内容。

一、大数据平台和技术的实践

（一）系统平台安全的实践：

物理安全。大数据平台依赖的机房环境进行安全保障，如：防火、防火、防雷击、安全系统、7*24 小时 CCTV 监控、温湿度监控等。

网络安全。 医渡云用户对大数据平台的日常管理维护、开发测试必须通过“VPN+堡垒机”的方式接入；大数据平台独立隔离部署，网络边界部署防火墙，严格访问控制；在大数据平台网络边界出口部署流量控制系统，对异常流量实现实时阻断。

主机安全。采用 LDAP 统一认证身份认证和主机权限控制；部署 HIDS 安全监控组件对主机实现 7*24 小时安全防护，及时发现并修复安全漏洞；部署堡垒机对登录主机用户进行操作审计，主机端的 HIDS 可以有效的解决安全边界模糊的问题。

账号权限管理。大数据平台运维人员、开发人员进行统一身份认证、权限管理如：VPN、堡垒机、服务器等。大数据服务平台进行 LDAP 统一身份认证、权限管理，如：HDFS、Hive、Hue、Kylin、Spark 等。

应用安全。统一应用发布出口，对应用会话进行日志审计，提供审计依据；组织安全厂商定期对应用系统进行渗透测试，及时发现并修复应用层安全漏洞。为保证大数据平台应用接口安全，除采用 HTTPS 通信以外，应用接口实现身份认证、权限控制、审计等安全需求。

大数据组件安全。不定期对大数据平台依赖的 HDFS、MarReduce、Spark、Hive 等相关组件漏洞进行跟踪修复，防止组件安全隐患造成数据泄密。

（二）平台安全运维的实践

医渡云大数据平台安全运维分为以下几个部分：

安全监控：安全监控包括对资产细粒度的梳理和监控，包括主机、安全设备、网络设备的监控。

安全检查：包括每天通过安全产品自动化的对现有资产进行风险的评估。

安全预警：当随着 1day 漏洞的爆发，我们会采取相应的预警策略。

安全加固：在相关重点业务中，参照 CIS 国际标准对所有大数据业务进行相关的基线检查并后续进行加固。

应急响应：在预警之后，我们通过相关产品进行风险定位，并进行漏洞的修复流程。

（三）平台安全相关技术的实践

平台相关技术主要在密钥管理和细粒度审计进行了实践。对于非对称和对称的密码算法的所有密钥进行了统一的管理和分配，建立了 KMS 系统。在审计上有流量层面的审计甚至包括服务器终端命令的审计。

二、数据安全的实践

对于数据领域，重点参照了 NIST SP 1500-4 中关于大数据本身的特点 5V：多样性、大量性、高速性、真实性、不稳定性进行思考，结合用例的实际特点，对大数据安全和隐私提出相应的解决思路。同时针对 NCHHSTP 给出的实现数据收集、存储、共享和使用过程安全和私密性的 10 大指导原则，来对数据进行处理。NIST IR 7497 讲述了医疗信息交换（HIE）需要注意的方面和分类。结合上述标准医渡云做了以下实践。

（一）个人信息安全

个人信息收集。秉承收集个人信息最小化的原则，跟医院签署授权协议和隐私协议，将相关个人病历数据进行保密性传输到大数据系统。

个人信息存储。对个人信息存储遵守最小时间原则，原则上存储个人病历时间在大数据系统中不超过初次使用和索引时间，使用之后即对个人信息进行匿名化处理。

个人信息使用。对个人信息使用保证责任分离和最小授权。对于特定业务进行拆分，不同业务之间能够查看的数据是不同的和分离的。

（二）重要数据的安全实践

敏感文件管理。大数据平台日常开发、测试涉及到的过程文档、电子文件、电子病历、纸质病历等敏感文件进行审批管理。

数据脱敏管理办法。参照 HIPAA、个人信息安全规范、以及医疗行业特有的数据安全需求（如：“统方”数据脱敏等），对敏感字段进行脱敏处理，在原始数据进入大数据平台之前，完成数据脱敏操作。对现有业务系统的字段脱敏，从患者、新生儿、统方三个角度出发，覆盖 16 个分类，共计 28 个字段进行了去标识化处理。如图 B.21 所示。

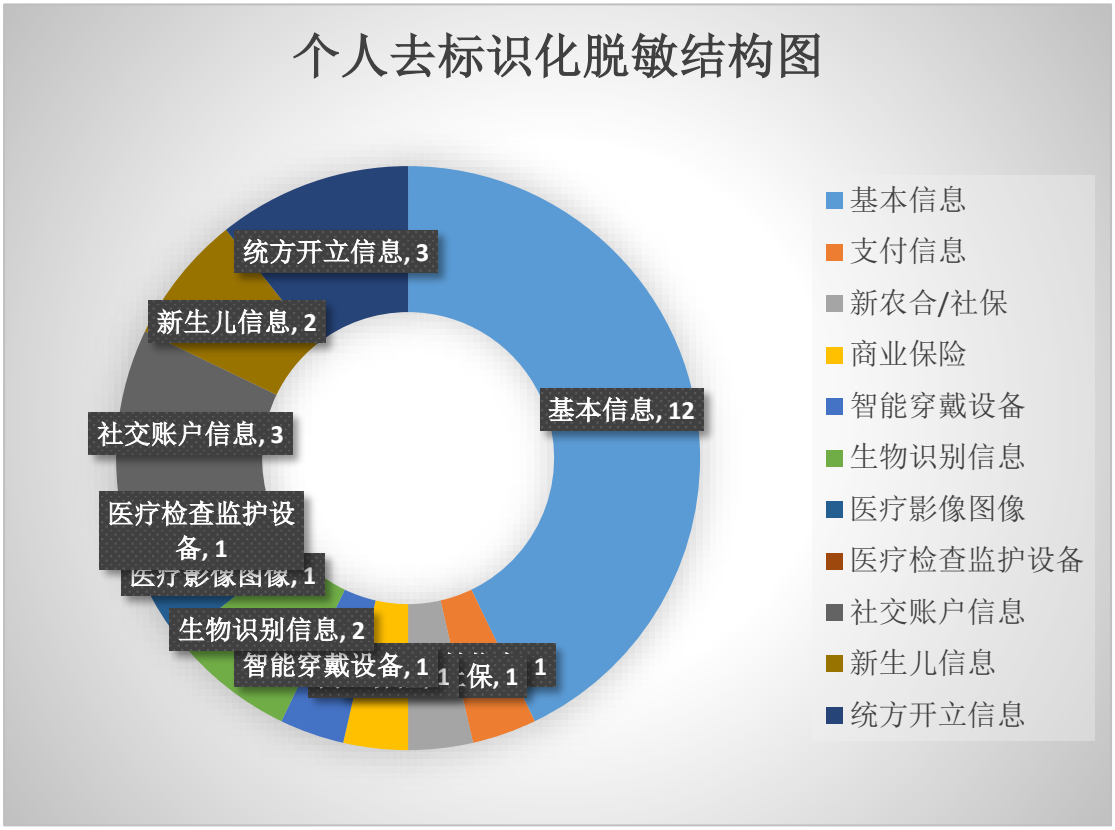

图 B.21 医渡云个人去标识化脱敏结构图

数据加密管理办法。首先对大数据平台传输、存储的敏感数据进行识别，然后对识别的敏感数据进行加密处理，加密算法选择强度较高的加密算法，如 AES 等国际通用算法或 SCB2 等国有商密算法。同时对数据加密密钥进行集中化与分发管理机制，实现对数据加密密钥的安全管理。

数据分级存储管理办法。由于医院信息系统的局限性，目前大数据平台主要获取的数据为离线数据，非实时数据。对离线数据加工前后采用不同的存储管理方式，加工完成后，对识别的重要数据进行分库、分表存储。

数据备份安全。对大数据平台依赖的加工前后的数据进行数据备份，未脱敏数据在加工处理前单独备份，备份介质单独进行管理。

三、数据安全标准的思考和建议

医渡云在大数据平台安全建设过程中，基础设施安全方面，主要根据 GB/T 22239-2008《信息安全技术 信息系统安全等级保护基本要求》进行基础设施安全建设，但为了保障大数据平台的先进性和适应性，还参考等级保护 2.0 中的安全通用要求和云计算扩展要求内容进行安全实践。在实践过程中，对于涉及的标准不适用项进行研讨，对以不同形式交付的大数据平台，建议应着重强调在大数据平台安全建设中要明确对大数据平台服务厂商和大数据平台使用者的管理职责和义务，不能仅靠其中一方来对大数据平台技术设施安全和数据安全进行负责。

由于大数据平台的 5V 特点，以往的信息系统设施无法满足大数据平台计算需求。大数据平台若要完成一定的业务目标必须依赖更先进的信息系统技术实现，比如：云计算、物联网、容器、分布式存储等内容。但是先进技术的应用对大数据平台的安全提出了巨大挑战，因此必须参考和借鉴《云计算关键领域安全指南》等类似的安全标准建设大数据平台的安全技术内容，在建设过程中由于大数据本身的特点再结合复杂的业务需求，致使大数据平台在安全方面建设成本和开销增加，是否可以对当前大数据涉及的相关技术要点进行

归纳总结，形成统一的标准规范，对重要技术点的要求和参考内容进行说明，以便于企业更有效的进行大数据平台安全建设，这样既可以减少大数据平台安全建设成本，又对后期大数据平台的安全服务能力评估和测评提供了保障。

在数据安全方面，主要参照最新发布的《个人信息安全规范》和国际《HIPPA》法案。以《个人信息安全规范》为指导，参考《HIPAA》的具体安全要求进行数据安全实践。在实践过程中发现，重要数据的管理和维护在明确职责基础之上，应对具体的数据内容进行明确安全要求，比如：姓名、身份证、联系方式等内容应以通用要求和扩展要求的方式对具体数据字段进行安全说明，并且对数据的交付方式和交付安全应进行明确要求。通用要求有利于各企业梳理现有大数据平台数据内容，扩展要求有利于个企业结合业务特点和行业特色形成具体的实践指南，安全的交付方式有利于企业提高数据交付质量和内容。

B.13 中电长城网际大数据安全标准应用实践

一、应用场景

目前大数据协同共享越来越频繁，医疗健康、信息安全、公共安全、能源电力、金融等领域和行业都有数据协同需求，例如在安全领域，需要通过引入不同维度的数据进行协同分析，从线索中分析得到安全攻击事件，并从事件分析中还原出整个安全攻击的全貌，才能进行有效的应急响应处置。但是，由于部门壁垒、行业分隔，数据归属权，隐私保护，数据协同共享后的控制权等各种因素，导致了大量数据无法真正协同共享利用。如何破解这些矛盾，在保证数据安全的前提下，让数据协同利用起来，发挥大数据作用。长城网际提出了“前置分析、本地计算、多级融合”的大数据平台设计架构，利用大数据相关技术标准、管理规范，突破多项核心技术，解决多源异构数据协同的隐私保护，跨安全域的数据管控，数据协同汇聚融合等问题，有利于推动智慧城市、电子政务等大数据跨行业、跨政府部门、跨区域的协同共享利用。

二、系统方案

本系统是为满足政府管理、公安、医院、金融等领域的决策及研究的需要，解决这些部门对于大量数据的采集、汇聚、存储、分析、融合等问题而研发的，平台架构如图 B.22 所示：

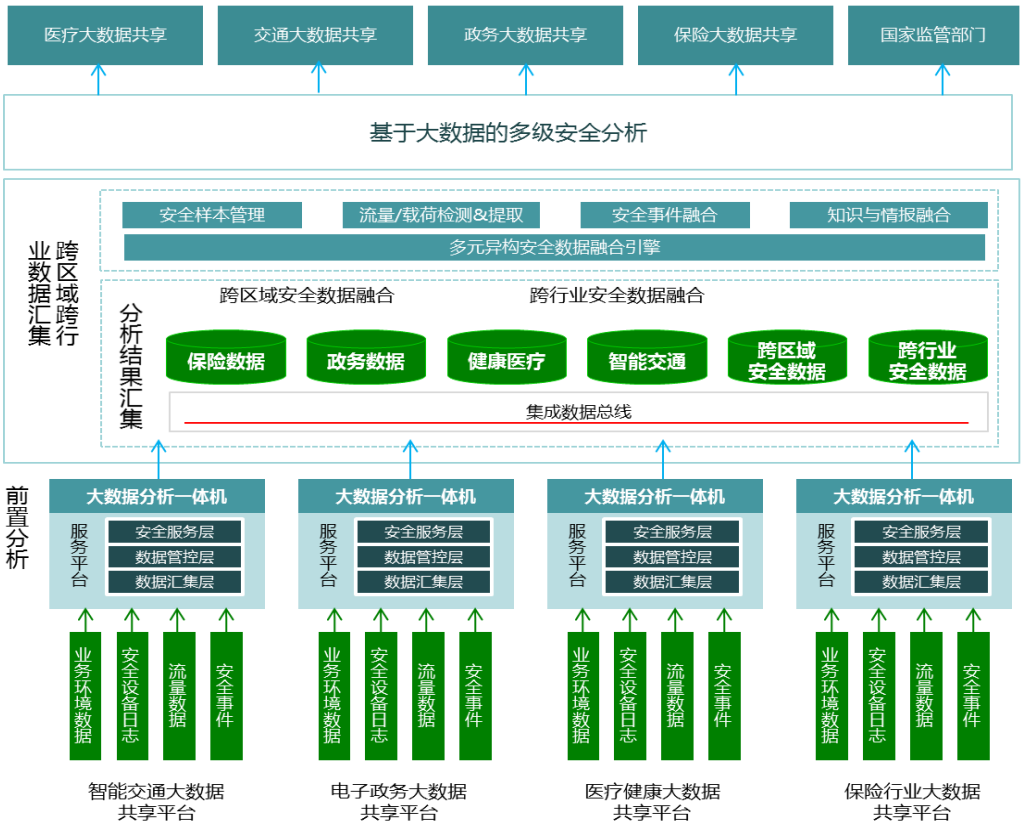

图 B.22 长城网际多源异构分布式大数据协同安全

总体架构采用“前置分析，本地计算，多级融合”的模式，通过在各个数据中心部署大数据分析一体机（称之为“数据安全屋”），实现了原始数据不搬家，物理上分布式存储，逻辑上集中管理。“数据安全屋”是一个分布在各个数据汇聚节点的大数据分析系统，自身由安全操作系统＋大数据分析平台＋增强级安全模块等组成，遵从与之对接的信息系统的安全要求。

各地的中心汇集自己服务对象的各类数据，形成区域、行业数据汇聚节点等，总中心将数据分析挖掘的任务分发到各个数据中心的“数据安全屋”。这样，通过“前置分析，本地计算，多级融合”模式，由“数据安全屋”从本地抽取数据，分析挖掘，将分析后的数据汇集到总中心，在总中心形成跨区域、跨行业的数据融合，对各级汇聚数据再进行协同分析。各个“数据安全屋”构成一个安全协同的闭环，以确保数据共享与协同分析过程的安全。在保障数据拥有者具有数据的控制权前提下，安全受控地进行数据共享与协同分析，实现可管理可控制，解决数据协同中最基本的信任问题。

三、标准应用

本方案一方面应用 GB/T 35274—2017《信息安全技术　大数据服务安全能力要求》、GB/T 31167-2014《信息安全技术　云计算服务安全指南》、GB/T 31168-2014《信息安全技术　云计算服务安全能力要求》和 GB/T 32923—2016《信息技术　安全技术　信息安全治理》等国家标准，实现大数据分析与协同服务安全；另一方面应用 GB/T 36073—2018《数据管理能力成熟度评估模型》，参考《信息技术　大数据　开放共享》（在研）和《信息技术　数据交易服务平台》（在研）等国家标准，提供大数据分析与协同安全服务。这些标准在本系统研究、开发和实施过程中发挥了重要作用，即从高起点做起，按高质量实现。同时，也发现现有标准的缺口，特别是在多源异构分布式大数据协同安全方面。比如，在医疗健康、公共安全等实际应用场景中，如果将数据脱敏后再共享给其他应用使用，所得到的数据质量大打折扣，难以支撑大数据应用分析。本方案提出的“前置分析，本地计算”

模式，将大数据分析前置到数据源头，避免了数据交易后发生数据失控、数据脱敏后质量下降等诸多问题。由此可见，在多源异构数据的共享与协同分析方面，急需相应的安全与隐私保护标准。

四、实施效果

中电长城网际在其参与的国家高技术研究发展计划（863 计划）重大专项“心血管疾病大数据平台的构建和应用研究”课题中采用了本系统方案，针对心血管疾病的大数据分析需求，集成电子病历、医学影像、检查检验等多类型数据，自主研发心血管疾病大数据集成、存储、分析、应用及安全技术，建立大数据管理和分析平台，构建心血管疾病预警预测模。目前已实现了安贞医院、唐山区域医院医疗机构、四川乐山区域医院医疗机构的 900 万人、200 多家医院的住院、门诊数据达 50T 的结构化数据，可作为医疗大数据应用示范。

本系统还在中电长城网际牵头的克拉玛依下一代互联网城市建设项目中实施，实现了多源异构数据安全交换共享，整合了社会数据、政府数据、互联网数据等十余种数据源。数据可以安全地被共享协同，但不会被滥用，解决了跨行业、跨政府部门、跨区域的数据共享协同及其安全问题，在推动智慧城市、电子政务等大数据跨行业、跨政府部门、跨区域的共享利用进行了有益探索。

B.14 中国移动大数据安全标准应用实践

中国移动在长期业务过程中沉淀了大量客户信息、生产数据和管理数据，这些数据具备规模大、类型多，精准度高等特点。同时，因数据的集中管理、数据对外开放、设备虚拟化等新技术新业态特点，给大数据业务发展带来了新的安全挑战。

中国移动自 2015 年以来逐步加强大数据安全保障体系建设步伐，体系框架涵盖安全策略、安全管理、安全运营、安全技术、合规评测、服务支撑等六大体系，涉及大数据安全采集、传输、存储、使用、共享、销毁六大过程。同时，通过推进“大数据安全防护”手段建设，积极开展“大数据安全应用”试点研究，全方位保护大数据的保密性、完整性、可用性和可追溯性，保障大数据环境安全可管、可控、可信。

通过近两年多的运行，大数据安全保障体系为实现公司数据开放共享、数据有效利用，同时又保障数据安全方面发挥了不可替代的作用。相关成果已推到成为国际、国家、行业标准，力争成为国内和国际最佳实践。

大数据安全保障体系框架如图 B.23 所示：

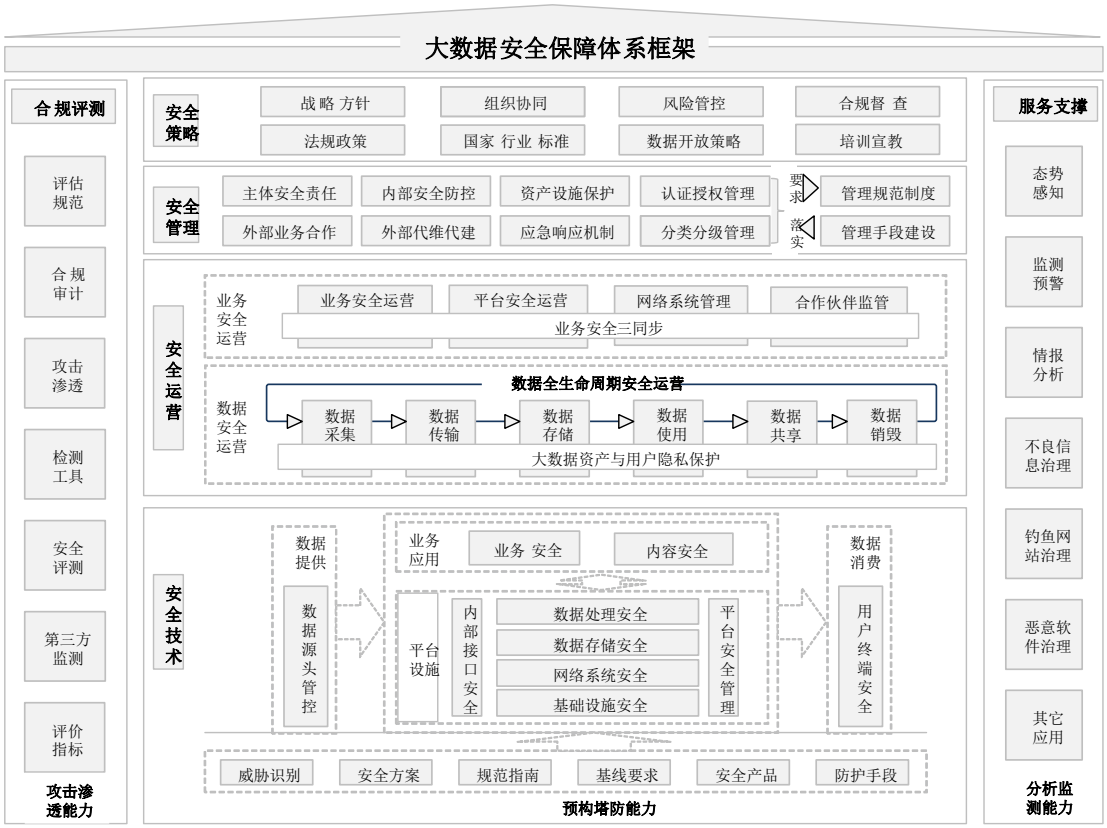

图 B.23 大数据安全保障体系框架

其中：

（一）安全策略体系从顶层设计层面明确了安全保障工作总体要求及方向指南，全面把握大数据安全工作的对象、提出安全管控的基本原则及核心策略、建立健全大数据安全内部管理流程、持续强化大数据对外合作管理力度。相应的管理制度有：《大数据安全风险防控工作指引》、《大数据安全保障总体策略》等。

（二）安全管理体系是通过管理制度建设，明确运营方安全主体责任，落实安全管理措施，相关制度包括第三方合作管理、内部安全管理、数据分类分级管理、应急响应机制、资产设施保护和认证授权管理等安全管理规范要求。相应的管理制度有：《大数据安全管理要求》、《大数据安全管控分类分级实施指南》等，主要针对市场、业务、支撑等部门对数据与平台系统进行安全管理。

（三）安全运营体系是通过定义运营角色，明确运营机构安全职责，实现对大数据业务及数据的全流程、全周期安全管理。体系设计涵盖业务安全运营、数据安全运营、安全应急管控等。相应的管理制度有：《大数据安全运营要求》等，主要针对业务部门、系统建设和支撑部门等进行规范执行。

（四）安全技术体系是公司开展大数据安全防护建设相关要求和实施方法。体系设计涵盖数量流转各环节数据安全防护通用技术要求、大数据平台各类基础设施及应用组件安全基线配置能力要求、基于数据防护通用要求和平台基线要求设计实现的大数据平台安全防护体系架构、敏感数据安全脱敏实施技术指南等。相应的企业标准有：《大数据安全防护通用技术要求》、《大数据平台安全基线要求》、《大数据安全防护技术实施指南》、《大数据安全脱敏实施指南》等，主要针对系统建设、运维部门等在系统规划、建设、运营、验收等场景规范应用。

（五）安全合规评测体系包括安全运营管理合规评测和安全技术合规评测方法、评测手段和评测流程。其建设目标是持续优化安全评估能力，实现对大数据业务各环节风险点

的全面评估，保障安全管理制度及技术要求的有效落实。相应的企业标准有：《大数据安全技术合规测评方法》等，主要针对安全管理部门、系统建设部门在安全验收、检查自查等场景应用。

（六）大数据服务支撑体系是基于大数据资源为信息安全保障提供支撑服务，开展大数据在安全领域的研究及推广应用，为公司信息安全治理提供新型技术手段，并支撑对外安全服务，实现数据增值。

基于中国移动的企业安全标准实践经验，国家标准在安全技术方面，特别是在大数据相关产品安全技术要求等方面比较欠缺，比如，大数据基础软件产品是大数据平台建设的核心部分，属于“大数据操作系统”，在其开发、选型、采购、建设等环节，亟需统一的安全要求，建议推进研制大数据基础软件安全功能要求相关国家标准。同时，在大数据安全评测、大数据安全应用等方面，也有必要研制配套的国家标准。

附录 C 其它相关资源介绍

C.1 大数据安全报告

- CSA 大数据安全性与隐私的十大挑战, 2012
  https://cloudsecurityalliance.org/download/top-ten-big-data-security-and-privacy-challenges/
- CSA 基于大数据的安全情报分析”, 2013
  https://cloudsecurityalliance.org/download/big-data-analytics-for-security-intelligence/
- CSA 大数据，大顾虑 -- 白宫任务清单, 2014
  https://cloudsecurityalliance.org/download/big-data-big-concerns-and-what-the-white-house-wants-to-do-about-it/
- CSA 大数据安全性与隐私白皮书：100 个最佳实践, 2016
  https://cloudsecurityalliance.org/download/big-data-security-and-privacy-handbook/
- IBM 企业信息保护 - 大数据的影响
  https://www-01.ibm.com/software/os/systemz/pdf/Info_Security_and_Big_Data_on_Z_White_Paper_Final.pdf
- 保护大数据：Hadoop 和 NoSQL 环境的安全建议
  https://securosis.com/assets/library/reports/SecuringBigData_FINAL.pdf

C.2 安全管理及框架

- NIST 网络安全框架, 2014 年 2 月 12 日，
  https://www.nist.gov/sites/default/files/documents/cyberframework/cybersecurity-framework-021214.pdf
- 保护性安全策略框架（PSPF），
  https://www.protectivesecurity.gov.au/Pages/default.aspx
- 开放数据中心联盟：数据安全框架版本 1.0，
  https://www.opendatacenteralliance.org//docs/Data_Security_Framework_Rev1.0.pdf
- ISO/IEC SC27 WG1 信息安全管理体系，标准样本例子：
  ISO/IEC 27014:2013: 信息安全治理，
  ISO/IEC 27017:2015 基于 ISO / IEC 27002 的云计算服务使用信息安全控制指南，
  通过下面链接可以找到所有 ISO/IEC SC27 WG1 标准，
  https://www.iso.org/committee/45306.html
- CSCC 用于大数据和分析的云客户体系架构
  http://www.cloud-council.org/deliverables/CSCC-Cloud-Customer-Architecture-for-Big-Data-and-Analytics.pdf
- CSCC 云客户体系架构，用于保护云服务的工作负载
  http://www.cloud-council.org/deliverables/CSCC-Cloud-Customer-Architecture-for-Securing-Workloads-on-Cloud-Services.pdf
- 确保成功云计算安全 10 个步骤
  http://www.cloud-council.org/deliverables/security-for-cloud-computing-10-steps-to-ensure-

success.htm

- CSCC 云安全标准的预期和谈判内容
  http://www.cloud-council.org/deliverables/CSCC-Cloud-Security-Standards-What-to-Expect-and-What-to-Negotiate.pdf

C.3 数据分类

- 联邦信息和信息系统安全分类标准
  http://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf
- 企业信息安全标准：数据分类
  http://www.mass.gov/anf/research-and-tech/cyber-security/security-for-state-employees/security-policies-and-standards/enterprise-information-security-standards.html
- ISO27001 信息分类政策
  http://www.iso27001security.com/html/27001.html
  https://www.itgovernance.co.uk/data-classification-software
  http://iso27001security.com/ISO27k_Model_policy_on_information_classification.pdf
- CSA 大数据分类, 2014
  https://cloudsecurityalliance.org/download/big-data-taxonomy/

C.4 个人信息保护

- ISO: 隐私保护框架，技术架构,标准路线图和一系列标准
  https://www.iso.org/obp/ui/#iso:std:iso-iec:29100:ed-1:v1:en
- 美国-欧盟数据隐私：从避风港至隐私盾
  https://fas.org/sgp/crs/misc/R44257.pdf
- 隐私盾
  https://www.privacyshield.gov/welcome
- 欧盟通用数据保护条例（GDPR）
  http://www.eugdpr.org/
- 澳大利亚隐私法案
  https://www.oaic.gov.au/privacy-law/privacy-act/
- 澳大利亚隐私法和实践
  http://www.alrc.gov.au/publications/report-108
  https://en.wikipedia.org/wiki/Privacy_in_Australian_law
- 美国隐私法案
  https://www.epa.gov/laws-regulations/summary-privacy-act
- NIST 特刊 800-53R4“联邦信息系统和组织的安全和隐私控制”
  http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf
- “网络空间信任身份的国家战略：提高在线选择，效率，安全和隐私”
  https://www.whitehouse.gov/sites/default/files/rss_viewer/NSTICstrategy_041511.pdf
- NISTIR 8062 联邦信息系统隐私风险管理
  http://csrc.nist.gov/publications/drafts/nistir-8062/nistir_8062_draft.pdf
- 隐私威胁评估框架

https://people.cs.kuleuven.be/~kim.wuyts/LINDDUN/LINDDUN.pdf

- 隐私工程框架
  http://www.mitre.org/publications/technical-papers/privacy-engineeringframework
- NIST SP800-188 政府数据集去标识化
  http://csrc.nist.gov/publications/drafts/800-188/sp800_188_draft2.pdf
- NISTIR 8053 个人信息去标识化
  http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf
- 亚太经合组织隐私框架
  https://www.apec.org/Publications/2005/12/APEC-Privacy-Framework

C.5 数据驻留和跨境流动

- CSCC 数据驻留的挑战，2017 年 5 月
  http://www.cloud-council.org/deliverables/CSCC-Data-Residency-Challenges.pdf
- GDPR 第 13 章 跨境数据流动，2016 年 7 月
  https://www.whitecase.com/publications/article/chapter-13-cross-border-data-transfers-unlocking-eu-general-data-protection
- 澳大利亚隐私法和实践（ALRC Report 108）第 31 章，跨境数据流动，2008 年 12 月。
  http://www.alrc.gov.au/publications/31.%20Cross-border%20Data%20Flows%20/introduction
- 欧盟隐私委员会，跨境数据流动 - 可能性
  https://www.privacycommission.be/en/cross-border-transfers
- 亚太经合组织跨境隐私规则系统
  http://www.cbprs.org/

C.6 行业数据安全

- 支付卡行业数据安全标准
  https://en.wikipedia.org/wiki/Payment_Card_Industry_Data_Security_Standard
- HIPAA / HITECH - 美国健康保险可移植性和责任法案（HIPAA）和健康信息技术经济与临床健康（HITECH） - 由美国联邦政府创建的法案包括保护患者私人信息的规定
  https://www.hhs.gov/hipaa/index.html
  https://www.hhs.gov/hipaa/for-professionals/special-topics/HITECH-act-enforcement-interim-final-rule/index.html
  https://en.wikipedia.org/wiki/Health_Insurance_Portability_and_Accountability_Act

C.7 标准文本

- 已发布的信息安全国家标准
  https://www.tc260.org.cn/advice/list.html

- 信息安全国家标准意见征求稿
  https://www.tc260.org.cn/front/bzzqyjList.html?start=0&length=10
- 国家标准全文公开系统
  http://www.gb688.cn/bzgk/gb/index
- ISO/IEC JTC 1/SC 27 主页
  https://www.iso.org/committee/45306.html
- NIST Big Data 主页
  https://bigdatawg.nist.gov/home.php

附录 D 大数据安全标准术语摘录

D.1 《信息安全技术　个人信息安全规范》术语

- **个人信息　personal information**

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身份或者反映特定自然人活动情况的各种信息。

注 1：个人信息包括姓名、出生日期、身份证件号码、个人生物识别信息、住址、通信通讯联系方式、通信记录和内容、账号密码、财产信息、征信信息、行踪轨迹、住宿信息、健康生理信息、交易信息等。

注 2：关于个人信息的范围和类型可参见《信息安全技术　个人信息安全规范》附录 A。

- **个人敏感信息　personal sensitive information**

一旦泄露、非法提供或滥用可能危害人身和财产安全，极易导致个人名誉、身心健康受到损害或歧视性待遇等的个人信息。

注 1：个人敏感信息包括身份证件号码、个人生物识别信息、银行账号、通信记录和内容、财产信息、征信信息、行踪轨迹、住宿信息、健康生理信息、交易信息、14 岁以下（含）儿童的个人信息等。

注 2：关于个人敏感信息的范围和类型可参见《信息安全技术　个人信息安全规范》附录 B。

- **个人信息主体　personal data subject**

个人信息所标识的自然人。

- **个人信息控制者　personal data controller**

有权决定个人信息处理目的、方式等的组织或个人。

- **收集　collect**

获得对个人信息的控制权的行为，包括由个人信息主体主动提供、通过与个人信息主体交互或记录个人信息主体行为等自动采集，以及通过共享、转让、搜集公开信息间接获取等方式。

注：如果产品或服务的提供者提供工具供个人信息主体使用，提供者不对个人信息进行访问的，则不属于本标准所称的收集行为。例如，离线导航软件在终端获取用户位置信息后，如不回传至软件提供者，则不属于个人信息收集行为。

- **明示同意　explicit consent**

个人信息主体通过书面声明或主动做出肯定性动作，对其个人信息进行特定处理做出明确授权的行为。

注：肯定性动作包括个人信息主体主动作出声明（电子或纸质形式）、主动勾选、主动点击“同意”、“注册”、“发送”、“拨打”等。

- **用户画像　user profiling**

通过收集、汇聚、分析个人信息，对某特定自然人个人特征，如其职业、经济、健康、教育、个人喜好、信用、行为等方面做出分析或预测，形成其个人特征模型的过程。

注：直接使用特定自然人的个人信息，形成该自然人的特征模型，称为直接用户画像。使用来源于特定自然人以外的个人信息，如其所在群体的数据，形成该自然人的特征模型，称为间接用户

画像。

- **个人信息安全影响评估** personal information security impact assessment

针对个人信息处理活动，检验其合法合规程度，判断其对个人信息主体合法权益造成损害的各种风险，以及评估用于保护个人信息主体的各项措施有效性的过程。

- **删除** delete

在实现日常业务功能所涉及的系统中去除个人信息的行为，使其保持不可被检索、访问的状态。

- **公开披露** public disclosure

向社会或不特定人群发布信息的行为。

- **转让** transfer of control

将个人信息控制权由一个控制者向另一个控制者转移的过程。

- **共享** sharing

个人信息控制者向其他控制者提供个人信息，且双方分别对个人信息拥有独立控制权的过程。

- **匿名化** anonymization

通过对个人信息的技术处理，使得个人信息主体无法被识别，且处理后的信息不能被复原的过程。

注：个人信息经匿名化处理后所得的信息不属于个人信息。

- **去标识化** de-identification

通过对个人信息的技术处理，使其在不借助额外信息的情况下，无法识别个人信息主体的过程。

注：去标识化建立在个体基础之上，保留了个体颗粒度，采用假名、加密、哈希函数等技术手段替代对个人信息的标识。

D.2 《信息安全技术　大数据服务安全能力要求》术语

- **大数据** big data

具有数量巨大、种类多样、流动速度快、特征多变等特性，并且难以用传统数据体系结构和数据处理技术进行有效组织、存储、计算、分析和管理的数据集。

- **数据生命周期** data lifecycle

数据从产生，经过数据采集、数据传输、数据存储、数据处理（包括计算、分析、可视化等）、数据交换，直至数据销毁等各种生存形态的演变过程。

- **数据服务** data service

提供数据采集、数据传输、数据存储、数据处理（包括计算、分析、可视化等）、数据交换、数据销毁等数据生存形态演变的一种网络信息服务。

- **大数据服务** big data service

支撑机构或个人对大数据采集、存储、使用和数据价值发现等数据生命周期相关的各种数据服务和系统服务。

大数据服务一般面对的是海量、异构、快速变化的结构化、半结构化和非结构化数据服务，且通过

底层可伸缩的大数据平台和上层各种大数据应用的系统服务提供。

- **大数据应用** big data application

执行数据生命周期相关的数据采集、数据传输、数据存储、数据处理（如计算、分析、可视化等）、数据交换、数据销毁等数据活动，运行在大数据平台，并提供大数据服务的各种应用系统。

- **大数据平台** big data platform

采用分布式存储和计算技术，提供大数据的访问和处理，支持大数据应用安全高效运行的软硬件集合，包括监视大数据的存储、输入/输出、操作控制等大数据服务软硬件基础设施。

- **大数据服务提供者** big data service provider

通过大数据平台和应用，提供大数据服务的机构。

- **大数据使用者** big data consumer

使用大数据平台或应用的末端用户、其它信息技术系统或智能感知设备。

- **大数据系统** big data system

包括大数据使用者、大数据服务提供者、大数据应用和大数据平台的信息系统。

- **数据供应链** data supply chain

对大数据服务提供者的数据采集、数据预处理、数据聚合、数据交换、数据访问等相关数据活动进行计划、协调、操作、控制和优化所需的可用数据资源形成的链状结构。

注：数据供应链目标是将大数据服务所需的各种数据和系统资产，通过计划、协调、操作、控制、优化等数据活动，确保大数据服务提供者能在正确的时间，按照正确的数据服务协议送给正确的大数据使用者。

- **数据交换** data interchange

为满足不同平台或应用间数据资源的传送和处理需要，依据一定的原则，采取相应的技术，实现不同平台和应用间数据资源的流动过程。

- **数据共享** data sharing

让不同大数据用户能够访问大数据服务整合的各种数据资源，并通过大数据服务或数据交换技术对这些数据资源进行相关的计算、分析、可视化等处理。

- **重要数据** important data

我国机构和个人在境内收集、产生的不涉及国家秘密，但与国家安全、经济发展以及公共利益密切相关的数据。

注：重要数据通常指公共通信和信息服务、能源、交通、水利、金融、公共服务、电子政务等重要行业和领域的各类机构在开展业务活动中收集和产生的，不涉及国家秘密，但一旦泄露、篡改或滥用将会对国家安全、经济发展和社会公共利益造成不利影响的数据（包括原始数据和衍生数据）。

D.3 《信息安全技术　个人信息去标识化指南》术语

- **个人信息** personal information

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身份或者反映特定自然人活动情况的各种信息。

[GB/T 35273-2017，定义3.1]

- **个人信息主体 personal data subject**

  个人信息所标识的自然人。

  [GB/T 35273-2017，定义3.3]

- **去标识化 de-identification**

  通过对个人信息的技术处理，使其在不借助额外信息的情况下，无法识别个人信息主体的过程。

  [GB/T 35273-2017，定义3.14]

  注：去除标识符与个人信息主体之间关联性。

- **微数据 microdata**

  一个结构化数据集，其中每条（行）记录对应一个个人信息主体，记录中的每个字段（列）对应一个属性。

- **聚合数据 aggregate data**

  表征一组个人信息主体的数据，比如各种统计值的集合。

- **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。

  注：标识符分为直接标识符和准标识符。

- **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。例如：姓名、身份证号、护照号、驾照号、地址、电子邮件地址、电话号码、传真号码、银行卡号码、车牌号码、车辆识别号码、社会保险号码、健康卡号码、病历号码、设备标识符、生物识别码、互联网协议（IP）地址号和网络通用资源定位符（URL）等。

  注：特定环境指个人信息使用的具体场景。比如，在一个具体的学校，通过学号可以直接识别出一个具体的学生。

- **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。比如：性别、出生日期或年龄、事件日期（例如入院、手术、出院、访问）、地点（例如邮政编码、建筑名称、地区）、族裔血统、出生国、语言、原住民身份、可见的少数民族地位、职业、婚姻状况、受教育水平、上学年限、犯罪历史、总收入和宗教信仰等。

- **重标识 re-identification**

  把去标识化的数据集重新关联到原始个人信息主体或一组个人信息主体的过程。

- **敏感属性 sensitive attribute**

  数据集中需要保护的属性，该属性值的泄露、修改、破坏或丢失会对个人产生损害。

  注：在潜在的重标识攻击期间需要防止其值与任何一个个人信息主体相关联。

- **有用性 usefulness**

数据对于应用有着具体含义、具有使用意义的特性。去标识化数据应用广泛，每种应用将要求去标识化数据具有某些特性以达到应用目的，因此在去标识化后，需要保证对这些特性的保留。

- **完全公开共享　completely public sharing**

  数据一旦发布，很难召回，一般通过互联网直接公开发布。

  注：同英文术语The Release and Forget Model。

- **受控公开共享　controlled public sharing**

  通过数据使用协议对数据的使用进行约束，数据使用协议规定内容应包含但不限于：

  a)　禁止信息接收方发起对数据集中个体的重标识攻击；

  b)　禁止信息接收方关联到外部数据集或信息；

  c)　禁止信息接收方未经许可共享数据集。

  比如，针对合格的研究者，可基于数据使用协议共享数据。

  注：同英文术语The Data Use Agreement Model。

- **领地公开共享　enclave public sharing**

  在物理或者虚拟的领地范围内共享，数据不能流出到领地范围外。

  注：同英文术语The Enclave Model。

- **去标识化技术　de-identification technique**

  降低数据集中信息和个人信息主体关联程度的技术。

  注1：降低信息的区分度，使得信息不能对应到特定个人，更低的区分度是不能判定不同的信息是否对应到同一个个人，实践中往往要求一条信息可能对应到的人数超过一定阈值。

  注2：断开和个人信息主体的关联，即将个人其它信息和标识信息分离。

- **去标识化模型　de-identification model**

  应用去标识化技术并能计算重标识风险的方法。

D.4 《信息安全技术　大数据安全管理指南》术语

- **大数据　big data**

  具有数量巨大、种类多样、流动速度快、特征多变等特性，并且难以用传统数据体系结构和数据处理技术进行有效组织、存储、计算、分析和管理的数据集。

- **重要数据　important data**

  我国机构和个人在境内收集、产生的不涉及国家秘密，但与国家安全、经济发展以及公共利益密切相关的数据。

  注：重要数据通常指公共通信和信息服务、能源、 交通、水利、金融、公共服务、电子政务等重要行业和领域的各类机构在开展业务活动中收集和产生的，不涉及国家秘密，但一旦泄露、篡改或滥用将会对国家安全、经济发展和社会公共利益造成不利影响的数据（包括原始数据和衍生数据）。

- **组织　organization**

  由作用不同的个体为实施共同的业务目标而建立的结构。组织可以是一个企业、事业单位、政府部门等。

- **大数据平台　big data platform**

采用分布式存储和计算技术，提供大数据的访问和处理，支持大数据应用安全高效运行的软硬件集合，包括监视大数据的存储、输入/输出、操作控制等大数据服务软硬件基础设施。

- **大数据环境** big data environment
  开展大数据活动所涉及的数据、平台、规程及人员等的要素集合。

- **大数据活动** big data activity
  组织针对大数据开展的一组特定任务的集合，大数据活动主要包括采集、存储、处理、分发、删除等。

D.5 《信息安全技术　数据安全能力成熟度模型》术语

- **数据安全** data security
  保护数据的可用性、完整性和机密性。

- **数据安全能力** data security capability
  组织机构在组织建设、制度流程、技术工具以及人员能力等方面对数据的安全保障能力。

- **成熟度** maturity
  对一个组织的有条理的持续改进能力的度量，对实现特定过程的连续性、可持续性、有效性和可信度的度量。

- **成熟度模型** maturity model
  对一个组织机构的成熟度进行度量的模型，包括一系列的代表能力和进展的特征、属性、指示或是模式。模型的内容通常是最佳实践的举例说明。成熟度模型提供一个组织机构衡量其当前的实践、流程、方法的能力水平的基准，并设置提升的目标和优先级。当一个模型被广泛应用于某个特定的行业，这个行业可以基于模型，来评估本行业的组织机构的成熟度等级。

- **组织机构** organization
  安排责任、权利和关系的一组人员和设施。

- **安全过程域** security process area
  实现同一安全目标的一系列数据安全相关活动、过程的集合。

- **基本实践** base practices
  是实现某一安全目标的数据安全相关的活动和过程，一个过程域由若干个基本实践组成。

- **通用实践** generic practices
  在评估中用于确定任何过程实施能力的评定准则。

- **数据脱敏** data desensitization
  通过模糊化等方法对原始数据的处理，达到屏蔽敏感信息的一种数据保护方法。

- **数据产品** data product

直接或间接使用数据的产品，包括但不限于能访问原始数据，提供数据计算、数据存储、数据交换、数据分析、数据挖掘、数据展示等应用的软件产品。

- **数据处理** data processing

  对原始数据进行抽取、转换、加载的过程；包括开发数据产品或数据分析。

- **数据供应链** data supply chain

  指为满足数据供应关系，通过资源和过程将需方、供方相互连接的网链结构，可用于供方将数据及其产品与服务提供给需方。

- **合规** compliance

对数据安全所适用的法律法规的遵循。

D.6 《信息安全技术 数据交易服务安全要求》术语

- **数据交易** data transaction

  数据供方和需方之间以数据商品作为交易对象，进行的以货币或货币等价物交换数据商品的行为。

  注 1：数据商品包括用于交易的原始数据或加工处理后的数据衍生产品。

  注 2：数据交易包括以大数据或其衍生品作为数据商品的数据交易，也包括以传统数据或其衍生品作为数据商品的数据交易。

- **数据供方** data supplier

  数据交易中提供数据的组织机构。

- **数据需方** data demander

  数据交易中购买和使用数据的组织机构。

- **数据交易服务** data transaction service

  帮助数据供方和需方完成数据交易的活动。

- **数据交易服务机构** data transaction service provider

  为数据供需双方提供数据交易服务的组织机构。

- **数据交易服务平台** data transaction service platform

  为数据交易提供各项服务的信息化平台。

- **在线数据交付** online data delivery

  数据供方通过网络向数据需方交付数据的模式。

- **离线数据交付** offline data delivery

  数据供需双方在达成数据交易协议后，由数据供方通过离线方式将数据从供方提供给需方的交付模式。

- **托管数据交易** custodian data delivery

  数据供需双方在达成数据交易协议后，由供方将数据拷贝到数据交易服务机构指定的数据托管服务平台，需方在数据托管服务平台内使用数据，数据不发生转移的交付模式。

- **数据交易过程** data exchanging process

数据供需双方依托数据交易服务平台针对具体的数据交易对象，进行的一次完整和具体的数据交易行为。

注：数据交易过程一般分为交易申请、交易磋商、交易实施和交易结束等环节。

- **重要数据** important data

  指与国家安全、经济发展和社会公共利益密切相关的数据。

  注：重要数据通常指公共通信和信息服务、能源、 交通、水利、金融、公共服务、电子政务等重要行业和领域的各类组织在开展业务活动中收集和产生的，不涉及国家秘密，但一旦泄露、篡改或滥用将会对国家安全、经济发展和社会公共利益造成不利影响的数据。

附录 E 信安标委标准工作程序

为帮助更多人了解大数据安全标准工作程序，更好的参与到大数据安全国家标准研制工作中来，简单介绍《信息安全国家标准项目管理办法》内容，主要围绕标准项目申请立项程序和标准项目制修订程序两方面进行简要介绍，更多细节请关注全国信息安全标准化技术委员会（以下简称“信安标委”）官网（http://www.tc260.org.cn）相关内容。

E.1 标准项目申请立项程序

信安标委标准项目立项程序如图 6.1 所示。基本流程为：申请单位提出标准项目立项申请，秘书处进行形式审查，工作组技术审查，WG1 专家组审查，信安标委全体委员投票表决，主任办公会审议，全部通过后，标准项目立项成立，后续工作按照标准项目制修订程序进行。

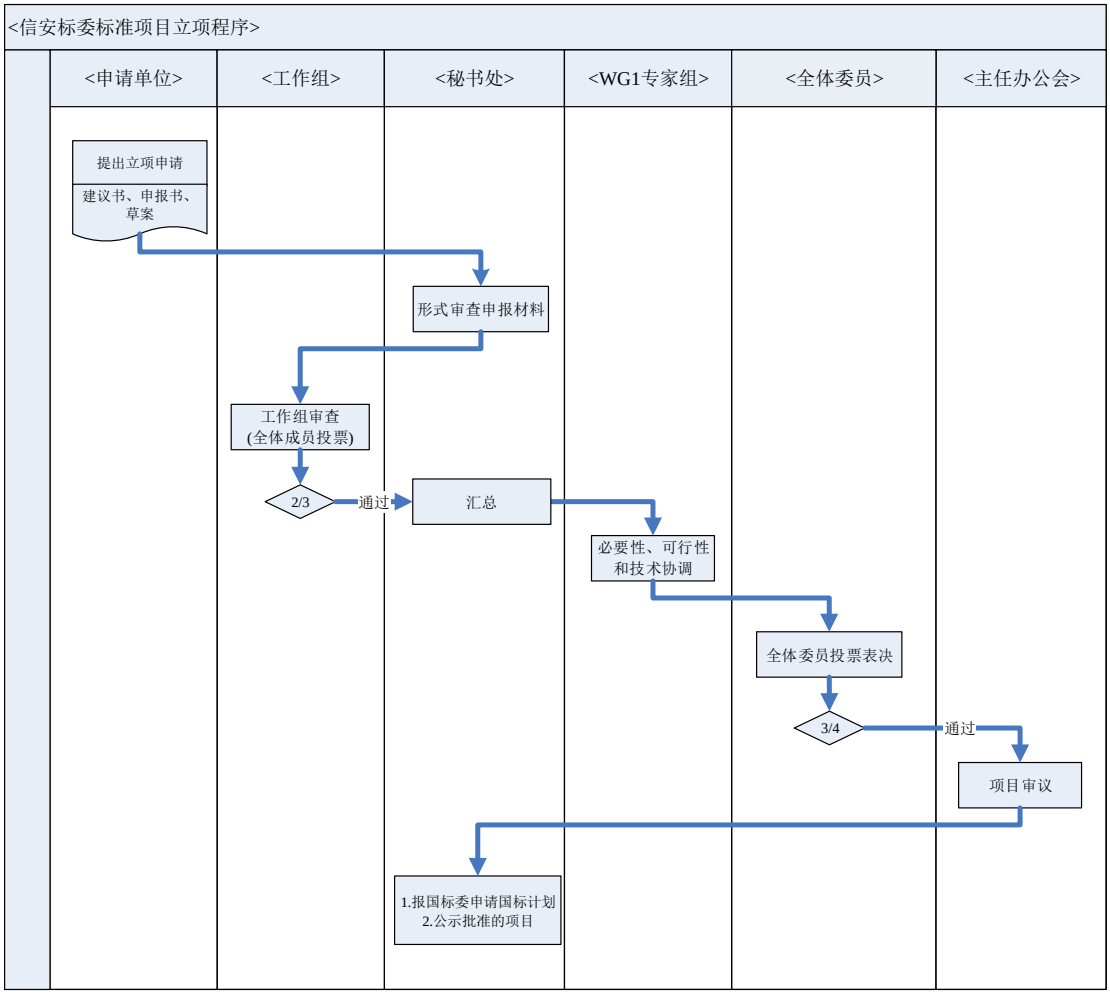

图 E.1 信安标委标准项目立项程序

E.2 标准项目制修订程序

信安标委标准项目制修订程序如图 6.2 所示。基本流程为：标准项目组起草完善标准草

案，工作组讨论审查，通过后转为征求意见稿，广泛征求各方意见，包括各大部门意见；标准项目组针对意见修改完善后转为送审稿，由标准专家进行讨论审查；标准项目组针对专家意见修改完善后转为报批稿，信安标委全体委员投票通过后进入主任办公会审查，审查通过后即上报国标委，进入报批发布程序。

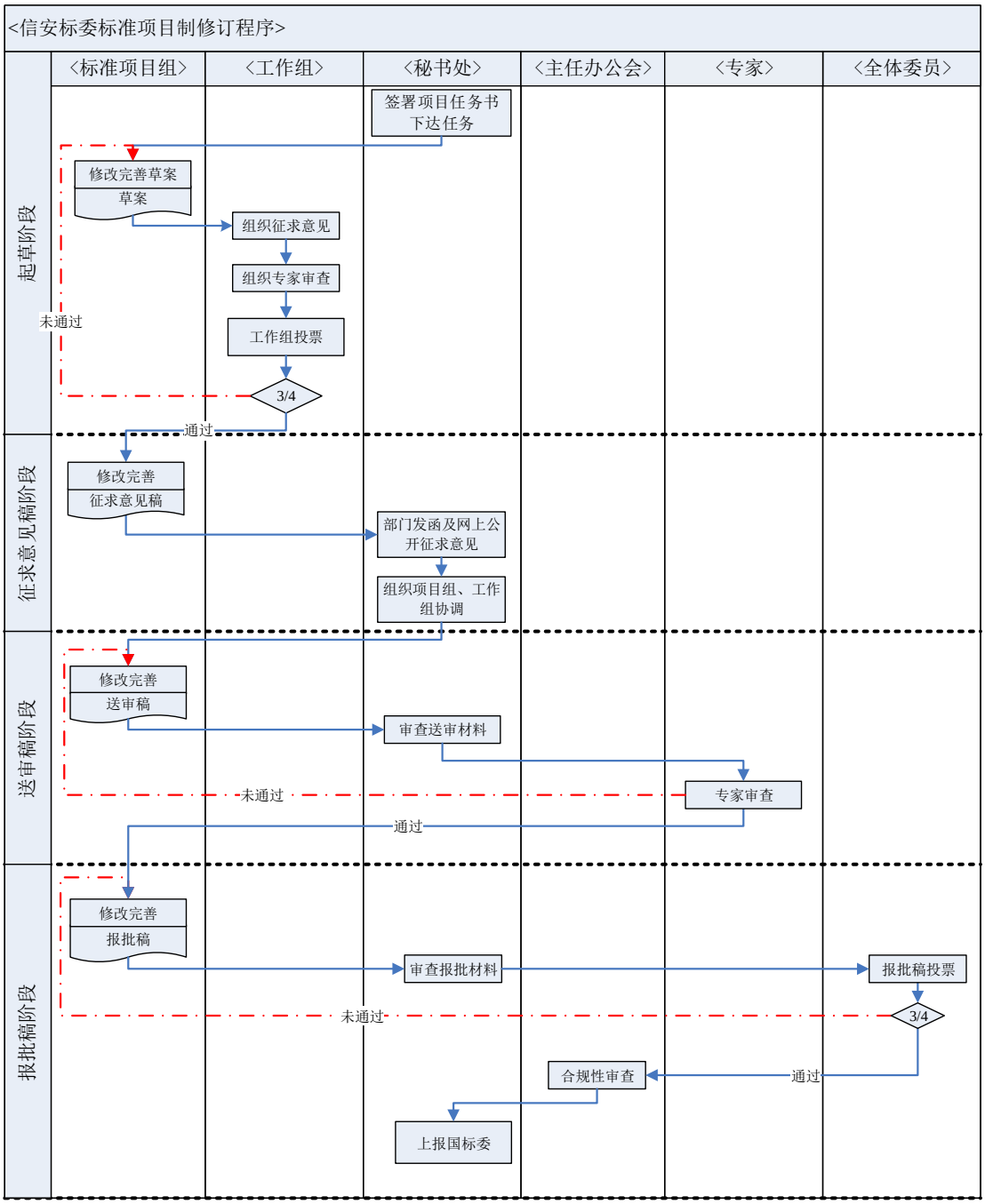

图 E.2 信安标委标准项目制修订程序

附录 F 缩略语

APEC	亚太经济合作组织（Asia-Pacific Economic Cooperation）
APT	高级持续威胁（Advanced Persistent Threat）
BDRA	大数据参考架构（Big Data Reference Architecture）
BDRA-S&P	大数据参考架构（Big Data Reference Architecture — Security and Privacy）
BDWG	大数据标准工作组（Big Data Working Group）
BSI	英国标准协会（British Standard Institute）
CCM	云安全控制矩阵（Cloud security Control Matrix）
CSA	云安全联盟（Cloud Security Alliance）
CSCC	云标准用户协会（Cloud Standards Customer Council）
DDoS	分布式拒绝服务（Distributed Denial-of-Service）
DPI	深度包检测（Deep Packet Inspection）
EDPB	欧盟数据保护理事会（European Data Protection Board）
ePHI	电子受保护健康信息（electronic Protected Health Information）
ETI	欧洲透明度倡议（European Transparency Initiative）
FERPA	[美国]家庭教育权和隐私权法案（Family Educational Rights and Privacy Act of 1974）
FFATA	[美国]数字问责和透明法案（Federal Funding Accountability and Transparency Act）
FIPS	[美国]联邦信息处理标准（Federal Information Processing Standard）
FISMA	[美国]联邦信息安全管理法（The Federal Information Security Management Act）
FTCAct	[美国]联邦贸易委员会法案（Federal Trade Commission Act）
GDPR	[欧盟]通用数据保护条例（General Data Protection Regulation）
GLBA	[美国]金融服务现代化法案（Gramm-Leach-Bliley Act）
HDFS	分布式文件系统（Hadoop Distributed File System）
HIDS	主机入侵侦测系统（Host Intrusion Detection System）

HIPAA	[美国]健康保险携带和责任法案（Health Insurance Portability and Accountability Act）
HTTPS	安全超文本传输协议(Hypertext Transfer Protocol Secure)
ICT	信息与通信技术（Information and Communications Technology）
IEC	国际电工委员会（International Electrotechnical Commission）
ISO	国际标准化组织（International Organization for Standardization）
ITU-T	国际电信联盟电信标准化组（International Telecommunication Union - Telecommunication Standardization Sector）
JTC1	ISO/IEC 联合技术委员会 1：信息技术（ISO/IEC Joint Technical Committee 1: Information technology）
NBDRA	NIST 大数据参考架构（NIST Big Data Reference Architecture）
NBD-PWG	NIST 大数据公开工作组（NIST Big Data Public Working Group）
NCHHSTP	[美国]艾滋病、肝炎、性传播疾病与结核病预防中心（National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention）
NIST	[美国]国家标准与技术研究院（National Institute of Standards and Technology）
NoSQL	非关系型的数据库（Not Only Structured Query Language）
OAIC	澳大利亚信息专员办公室（Office of the Australian Information Commissioner）
PCI DSS	支付卡行业数据安全标准（Payment Card Industry Data Security Standard）
PDPA	[新加坡]个人数据保护法令（Personal Data Protection Act）
PDPC	[新加坡]个人数据保护委员会（Personal Data Protection Commission）
PIA	隐私影响评估（Privacy Impact Assessment）
PII	可识别个人信息（Personally Identifiable Information）
PIPC	[日本]个人信息保护委员会（Personal Information Protection Committee）
PHI	受保护健康信息（Protected Health Information）
RBAC	基于角色的访问控制（Role Based Access Control）

SAC	国家标准化管理委员会（Standardization Administration of the People's Republic of China）
SC27	国际信息安全技术委员会（ISO/IEC JTC 1/SC 27）
SIEM	安全信息和事件管理（Security Information And Event Management）
SSL	安全套接字协议（Secure Sockets Layer）
SWG-BDS	大数据安全标准特别工作组（Special Working Group - Big Data Security）
TC28	全国信息技术标准化委员会（简称：信标委）
TC260	全国信息安全标准化技术委员会（简称：信安标委）
VPN	虚拟私人网络（Virtual Private Network）

参考文献

[1] 国务院. 促进大数据发展行动纲要. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm, 2015.

[2] 全国人民代表大会. 中华人民共和国网络安全法. http://www.npc.gov.cn/npc/xinwen/2016-11/07/content_2001605.htm, 2016.

[3] 中央网络安全和信息化领导小组办公室. 关于加强国家网络安全标准化工作的若干意见. http://www.cac.gov.cn/2016-08/22/c_1119430337.htm, 2016.

[4] 国家互联网信息办公室. 国家网络空间安全战略. http://www.xinhuanet.com/politics/2016-12/27/c_1120196479.htm, 2016.

[5] 国务院. “十三五”国家信息化规划. http://www.gov.cn/zhengce/content/2016-12/27/content_5153411.htm, 2016.

[6] 中国信息通信研究院. 网络安全产业白皮书. www.caict.ac.cn/kxyj/qwfb/bps/201709/P020170919308653198647.pdf, 2017.

[7] NIST Big Data Public Working Group, Security and Privacy Subgroup. DRAFT: NIST Big Data Interoperability Framework: Volume 4, Security and Privacy, DRAFT Version 2, https://bigdatawg.nist.gov/V2_output_docs.php, August 7, 2017.

[8] 全国人民代表大会. 全国人大常委会关于加强网络信息保护的决定. http://www.gov.cn/jrzg/2012-12/28/content_2301231.htm, 2012.

[9] 工业和信息化部. 电信和互联网用户个人信息保护规定. http://www.miit.gov.cn/n1146295/n1146557/n1146619/c4700556/content.html, 2016.

[10] 工业和信息化部. 大数据产业发展规划（2016-2020 年）. http://www.miit.gov.cn/n1146295/n1652858/n1652930/n3757016/c5464999/content.html, 2017.

[11] 国家互联网信息办公室. 个人信息和重要数据出境安全评估办法（征求意见稿）. http://www.cac.gov.cn/2017-04/11/c_1120785691.htm, 2017.

[12] 国务院. 征信业管理条例. http://www.gov.cn/zwgk/2013-01/29/content_2322231.htm, 2013.

[13] 工业和信息化部. 关于规范云服务市场经营行为的通知（公开征求意见稿）. http://www.miit.gov.cn/n1146295/n1652858/n1653100/n3767755/c5381367/content.html, 2016.

[14] 人民银行. 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. http://www.gov.cn/gongbao/content/2011/content_1918924.htm, 2011.

[15] 中国保险监督管理委员会. 保险公司开业验收指引. http://www.circ.gov.cn/web/site0/tab5225/info163158.htm, 2011.

[16] 中国保监会. 保险机构信息化监管规定(征求意见稿). http://www.circ.gov.cn/web/site0/tab5174/info3975814.htm, 2015.

[17] 国家卫生计生委. 人口健康信息管理办法（试行）. http://www.cac.gov.cn/2014-08/20/c_1112064075.htm, 2014.

[18] 国务院. 地图管理条例. http://www.gov.cn/zhengce/content/2015-12/14/content_10403.htm, 2015.

[19] 交通运输部, 工业和信息化部, 公安部, 商务部等. 网络预约出租汽车经营服务管理暂行办法. http://www.miit.gov.cn/n1146295/n1146557/n1146624/c5218603/content.html, 2016.

[20] 中国气象局. 气象资料共享管理办法. http://www.cma.gov.cn/2011zwxx/2011zflfg/2011zbmgz/201110/t20111027_135170.html, 2001.

[21] 国家新闻出版广电总局, 工业和信息化部. 网络出版服务管理规定. http://www.miit.gov.cn/n1146290/n4388791/c4638978/content.html, 2016.

[22] 全国人民代表大会. 中华人民共和国刑法修正案（七）. http://www.gov.cn/flfg/2009-02/28/content_1246438.htm, 2009.

[23] 全国人民代表大会. 中华人民共和国刑法修正案（九）. http://www.npc.gov.cn/npc/xinwen/2015-08/31/content_1945587.htm, 2015.

[24] 最高人民法院,最高人民检察院. 最高人民法院 最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释. http://www.spp.gov.cn/xwfbh/wsfbt/201705/t20170509_190088.shtml, 2017.

[25] 全国人民代表大会. 中华人民共和国民法总则. http://www.npc.gov.cn/npc/xinwen/2017-03/15/content_2018907.htm, 2017.

[26] 全国人民代表大会. 中华人民共和国消费者权益保护法. http://www.npc.gov.cn/npc/xinwen/2013-10/26/content_1811773.htm, 2013.

[27] 国家工商行政管理总局. 消费者权益保护法实施条例（征求意见稿）. http://www.gov.cn/xinwen/2016-08/05/content_5097833.htm, 2016.

[28] 全国人民代表大会. 中华人民共和国标准化法. http://www.npc.gov.cn/npc/xinwen/2017-11/04/content_2031446.htm, 2017.

[29] 国务院办公厅. 国家标准化体系建设发展规划（2016—2020 年）. http://www.gov.cn/zhengce/content/2015-12/30/content_10523.htm, 2015.

[30] 国务院办公厅. 国务院办公厅关于促进和规范健康医疗大数据应用发展的指导意见. http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm, 2016.

[31] 国务院. “健康中国 2030”规划纲要. http://www.gov.cn/zhengce/2016-10/25/content_5124174.htm, 2016.

[32] 贵阳市人民代表大会. 贵阳市政府数据共享开放条例. http://www.gyfg.gov.cn/article-21-12644.aspx, 2017.

[33] 国务院办公厅. 科学数据管理办法. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm, 2018.