

数据安全治理白皮书 4.0

中关村网络安全与信息化产业联盟

数据安全治理专业委员会 编著

2022



前言

Preface

■ 数据安全新形势背景

数据被定义为新时代重要的生产要素，是国家基础性战略资源，数据价值的发挥是推动我国社会经济转型发展的新动力。

2021 年以来，我国陆续发布了《“十四五”国家信息化发展规划》、《“十四五”数字经济发展规划》等重要国家数据战略，强调建设数字中国，加快数据要素市场化流通，创新数据要素开发利用机制。为促进数据共享和数据交易，上海、深圳等十余个省市也密集发布了数据条例。伴随数据要素市场化进程的发展，数据的巨大价值和重要意义已得到强调和凸显，但数据的开发利用也是一把双刃剑，当数据创造价值的同时，也面临着数据被泄露、篡改、滥用等风险，造成对个人、组织、社会公共利益甚至国家利益的严重威胁和损害。为了规范数据处理活动，保障数据安全，促进数据开发利用，《中华人民共和国数据安全法》（以下简称数据安全法）、《中华人民共和国个人信息保护法》（以下简称个人信息保护法）正式发布，强调统筹数据发展和安全防护并重，在保障安全和隐私前提下推动数据依法合理有效利用。

■ 编写目标

在上述形势背景下，企业或组织在数据收集、存储、使用、加工、传输、提供、公开等数据处理活动过程中，针对数据安全威胁与监管合规要求，无可避免地需要面对越来越严峻和紧迫的数据安全挑战：如何在数据资产的开发利用、价值实现与安全保护、履行合规义务之间进行恰当平衡？如何在数据安全方面编制合理的管理制度和选取适宜的技术方案？

面向日益严峻的数据安全形势，为了帮助企业或组织应对以上众多数据安全方面的困惑和难题，实现数据利用与安全防护一体两翼、平衡发展的目标，本白皮书在《数据安全治理白皮书 3.0》的前序版本基础上，进一步研判当前数据安全形势与动态，解读密集颁布的法律法规监管要求，梳理数据安全治理面临的痛点和问题，完善数据安全治理框架体系及相关技术，归纳新形势下的数据安全治理实践案例及典型数据安全事件，并提出未来的展望与倡议，力图尽可能全面、系统地整理和总结当前与数据安全治理有关的各类资料和最新进展，探索“让数据使用自由而安全”的治理方案，在数据要素释放价值的同时，

坚守安全底线。为数据安全治理相关人员提供指引和参考，期望为进一步推广、普及和完善数据安全治理的理念、方法、体系与实践添砖加瓦、贡献力量。

■ 读者对象

本白皮书主要面向以下几类读者：

组织内部数据安全治理相关人员：面向企业或组织内开展数据安全治理工作相关的决策人员、方案规划和实施人员、安全管理人员、技术培训人员，期望能够帮助他们更加深入全面地了解在企业或组织的数字化转型过程中正在和将要面对的数据安全威胁、风险和合规性要求以及行业内场景化治理实践，从而更积极主动地筹划和开展系统化的数据安全治理，确保企业或组织能够有效应对新形势下的各种数据安全挑战。

数据安全治理相关产品 / 服务提供商：面向数据安全行业的方案及产品策划人员、安全治理咨询服务人员及项目实施人员，期望通过白皮书中对治理框架、技术应用与实践案例等内容的介绍为产品 / 服务提供商开展治理方案编制、产品研发和实施服务工作提供启迪与参考，以更好的服务用户。

数据安全治理相关的法务工作者：对组织合规部门、律师等法务相关工作者，期望通过对实际数据安全相关用例介绍和数据安全技术介绍，给予他们在进行数据安全合规要求和建议工作中提供一定的借鉴参考。

其他相关读者：此外，本白皮书对于数据安全相关政策法规和标准规范的编制人员、数据安全领域的研究人员，也具有一定参考价值。

■ 导读

整个白皮书共分为六个正文章节和一个附录章节。其中：

第一章：数据安全形势与挑战：自 2021 年来，数据安全形势发生很大变化，梳理、分析国内外数据安全战略、数据安全风险形势及国家与行业监管新形势，通过剖析新形势指导数据安全治理需求。

第二章：数据安全治理目前面临的主要痛点和难题：面向数据安全风险和监管合规要求严峻形势，梳理数据安全治理的主要痛点、问题和关键需求。

第三章：数据安全治理理念及框架：围绕数据安全的风险和合规驱动需求，梳理数据安全治理思路，提出治理愿景，构建治理理念，形成覆盖管理、技术、运营体系的治理框架，并给出治理规划建设实施路径。

第四章：数据安全相关法律法规解读：在数据安全治理过程中，满足监管合规要求是重要驱动力之一，体系化解读国内外数据安全相关法律及法规的监管合规要求。

第五章：数据安全治理落地实践案例集：以数据安全治理框架为指导，面向金融、政务、能源、教育、医疗等行业众多数据处理活动场景，给出丰富的实际场景化数据安全治理实践案例集，为相关方开展数据安全治理建设提供参考与借鉴。

第六章：未来展望与倡议：面对数据安全治理实践涉及的管理、技术与运营过程中短期内尚无法有效解决的问题，以展望与倡议的形式予以表述，供行业内人士进行探讨。

附录：对近似概念联系与区别尝试解读，对数据安全关键技术、国内外相关治理理论、数据安全相关标准进行介绍，对2020年以来重大数据安全事件与法律案件进行分析，为业务人士提供参考。

■ 数据安全治理白皮书历史沿革

北京安华金和科技有限公司是在中国倡导数据安全治理理念的发起者和先行者，也是多年来推动数据安全治理理念在我国各行各业应用和实施的实践引领者。2017年，在中国网络安全产业联盟的支持倡导下，正式成立国内首个“数据安全治理工作组”，安华金和成功当选数据安全治理工作组组长单位，并组织召开了首届中国数据安全治理高峰论坛。

2018年，在中关村网络安全与信息化产业联盟的支持和指导下，全国首家数据安全领域的专项委员会——数据安全治理专业委员会成立，安华金和成功当选委员会主任单位。积极促进数据安全治理理念在我国的推广和普及，为我国数据安全领域的学术研究、技术创新、产业发展和人才培养提供了资源丰富的交流和支撑平台。成立当年，就组织召开了第二届中国数据安全治理高峰论坛，并组织编撰《数据安全治理白皮书》，首次提出数据安全治理的概念定义、治理理念及治理框架，强调业务需求与数据安全的平衡，在论坛上向社会公开发布。

2019年，第三届中国数据安全治理高峰论坛召开，《数据安全治理白皮书2.0》和《数据安全治理建设指南》公开发布。在国家及行业高度重视数据安全及个人隐私安全的背景下，增加了数据安全相关法规和标准列表说明；补充了个人信息收集与隐私政策测评报告相关解读，扩展了各行业的数据安全治理实践，对数据库防勒索、透明加解密等数据安全相关热点技术进行了介绍，并更新了2019年重要的数据安全事件。

2021年，第四届中国数据安全治理高峰论坛隆重召开，《数据安全治理白皮书3.0》重磅发布。进一步诠释了“让数据使用自由而安全”的治理理念，分析了业内关注的数据安全与网络安全等概念的近似联系与区别，汇集了数据安全关键技术与前沿技术，解读了相关法律法规标准要求，整理了覆盖政务、金融、能源、教育、电信及医疗行业丰富的治理案例，成为数据安全行业较为全面且具有影响力的数据安全治理参考书。

主编方 / 出品方

中关村网络安全与信息化产业联盟数据安全治理专业委员会

指导单位

中国计算机学会计算机安全专业委员会

工业信息安全产业发展联盟

中关村网络安全与信息化产业联盟

北京工业互联网技术创新与产业发展联盟

中国电子商会自主创新与安全技术委员会

参编单位

公安部第一研究所、国家工业信息安全发展研究中心、中国软件评测中心、国家信息技术安全研究中心、公安部第三研究所网络安全法律研究中心、江苏省信息安全测评中心、北京市政务信息安全保障中心、中国船级社信息中心、陕西省信息化工程研究院、西安交通大学苏州信息安全法学所、华北电力大学能源电力大数据研究院、南开大学网络空间安全学院、电子科技大学网络空间安全研究院、北京世辉律师事务所、北京市中伦律师事务所、泰和泰律师事务所、北京博遵律师事务所、北京大成律师事务所、北京安华金和科技有限公司、光大科技有限公司、北京数字认证股份有限公司、北京安信天行科技有限公司、深信服科技股份有限公司、国网英大股份有限公司、国网智能电网研究院有限公司、中金金融认证中心有限公司、北京谷安天下科技有限公司、北京中测安华科技有限公司、南京壹进制信息科技有限公司、深圳市联软科技股份有限公司、杭州立思辰安科科技有限公司

致 谢

感谢以下人员为《数据安全治理白皮书 4.0》的编制付出的辛勤劳动。

指导专家（按姓氏笔画排序，排名不分先后）

于锐、马民虎、王克、王新锐、冯艳春、任卫红、向小佳、刘龙庚、刘海峰、刘哲理、刘晓韬、严明、李新社、李新友、李斌、李建彬、李俊、李洪伟、吴兰、吴沈括、何跃鹰、张伟、季亨卡、金涛、俞克群、姚相振、黄道丽

编审专家（按姓氏笔画排序，排名不分先后）

李安伦、李向锋、杨海峰、何晋昊、宋博韬、陈青民、录洋、郝文江、崔媛媛、蔚晨、魏力

参编人员（按姓氏笔画排序，排名不分先后）

马佑平、王玮、王晓民、尹晓东、白倩、朱剑楠、阮东辉、孙铮、李吉音、李松涛、李海鹏、杨帅锋、张兵、张建耀、张勇、张腾标、苗维杰、周娜、高阳春、郭丽丽、麻建、隆峰、雷嘉宾、廖怀学

■ 版权声明

白皮书版权属于中关村网络安全与信息化产业联盟数据安全治理专业委员会（简称数据安全治理专业委员会），并受法律保护。转载、摘编或利用其他方式使用本白皮书文字或观点的，应注明“来源：中关村网络安全与信息化产业联盟数据安全治理专业委员会”，违者将被追究法律责任。

目录

Catalog

第一章 数据安全形势	2
1.1. 重新认识数据的价值	2
1.1.1. 我国数据发展战略	2
1.1.2. 国外数据发展战略	3
1.2. 数据安全风险态势	5
1.2.1. 外部数据安全威胁持续升级	5
1.2.2. 内部数据安全风险日益严峻	6
1.2.3. 数据安全事件危害性愈发严重	8
1.3. 监管新举措	8
第二章 数据安全治理目前面临的主要痛点和需求	11
2.1. 管理层面	11
2.1.1. 多法并轨、多标准并行下缺乏一致性的合规治理手段	11
2.1.2. 合规治理下数据安全实施细则尚不完善	11
2.1.3. 传统的网络安全风险评估应对数据安全适应度不足	12
2.1.4. 数据安全防护和网络安全防护体系尚未形成有机融合	12
2.2. 技术层面	13
2.2.1. 面向海量数据的数据资产梳理与分类分级实施难度大	13
2.2.2. 数据所有权与控制权（使用权）分离导致数据共享和开放困境	13
2.2.3. 单点安全防护能力无法有效应对复杂数据流动风险	14
2.3. 运营层面	14
2.3.1. 业务动态变化下按需管控的运营机制尚不健全	14

第三章 数据安全治理理念及框架	15
3.1. 数据安全治理理念逐步深化	15
3.1.1. 治理“愿景”更加清晰	16
3.1.2. 治理目标愈发明确	16
3.1.3. 核心理念全面深化	16
3.1.4. 建设步骤指导实践	17
3.1.5. 能力支撑推动治理框架设计	18
3.2. 数据安全治理整体框架	19
3.3. 数据安全组织架构	21
3.3.1. 职能架构	21
3.3.2. 定岗定员	22
3.4. 数据安全管理体系	23
3.4.1. 管理体系制度架构	23
3.4.2. 管理体系建设思路	24
3.5. 数据安全技术体系	24
3.5.1. 数据安全技术演进趋势	24
3.5.2. 技术体系架构	25
3.5.3. 识别技术（I）	26
3.5.4. 防护技术（P）	30
3.5.5. 监测技术（D）	36
3.5.6. 响应技术（R）	38
3.5.7. 集中管控能力	38
3.6. 数据安全运营体系	39
3.6.1. 运营体系架构	39
3.6.2. 运营体系建设思路	40
3.6.3. 运营服务场景及流程	41
3.7. 数据安全治理规划建设	42
3.7.1. 数据安全治理整体建设思路	42
3.7.2. 数据安全治理迭代式建设思路	42

第四章 法律法规标准解读	44
4.1. 国内法律法规与行业标准解读	44
4.1.1. 法律	44
4.1.2. 行政法规	48
4.1.3. 部门规章及规范性文件	49
4.1.4. 地方性法规	51
4.1.5. 国家标准、行业标准及相关指南	51
4.2. 国际数据安全政策与法规概述	52
4.2.1. 欧洲的政策和法律	52
4.2.2. 美国的政策和法律	53
4.2.3. 其他国家的政策和法律	55
第五章 数据安全治理落地实践案例集	56
5.1. 政务行业数据安全治理	57
5.1.1. 某政务大数据中心安全运营实践	57
5.1.2. 基于隐私计算技术的政务数据安全共享开放实践	61
5.1.3. 政务大数据中心数据流转安全监测实践	64
5.2. 金融行业的数据安全治理	70
5.2.1. 某国有银行集中化安全审计应用实践	70
5.2.2. 某城商行大数据中心安全运维实践	73
5.2.3. 某国有银行数据防泄露实践	77
5.3. 能源行业的数据安全治理	80
5.3.1. 某电网公司数据安全治理实践	80
5.3.2. 某能源云平台数据安全密码应用实践	83
5.4. 教育行业的数据安全治理	87
5.4.1. 某部委数据安全治理实践	87
5.5. 电信运营商行业的数据安全治理	89
5.5.1. 某电信国际的数据安全治理实践	89
5.6. 医疗行业的数据安全治理	93

5.6.1. 某医院医疗数据灾备实践	93
5.7. 车联网行业的数据安全治理	95
5.7.1. 某车联网企业基于数据分类分级的防护体系建设实践	95
5.8. 烟草行业的数据安全治理	98
5.8.1. 某烟草公司基于零信任和隐私计算技术的防护体系实践	98
5.9. 互联网广告行业的数据安全治理	102
5.9.1. 某互联网公司个人信息保护法律实践	102
 第六章 未来展望与倡议	105
6.1. 未来展望	105
6.1.1. 国家法律法规和标准体系日趋完善	105
6.1.2. 数据安全执法力度逐步增强	105
6.1.3. 数据安全治理从合规驱动逐步走向自驱动	106
6.1.4. 建立持续迭代的数据安全治理体系	107
6.1.5. 提升数据安全评估 / 审查自动化水平	107
6.1.6. 新兴技术的逐步发展和应用，持续激发数据安全新需求	107
6.2. 持续加强数据安全治理能力的倡议	108
6.2.1. 加快复合型数据安全人才的培养	108
6.2.2. 持续加强数据安全技术研究	108
6.2.3. 发挥行业联盟推动数据安全治理产业发展的重要力量	109
6.2.4. 加速推进数据安全治理国际化	109
 附录	111
A. 词汇表	111
B. 近似概念的联系与区别	112
B.1 数据安全、信息安全、网络安全	112
B.2 数据安全治理与数据治理	112
B.3 数据安全治理与数据安全治理	114
B.4 数据安全治理与数据安全成熟度模型	114

C. 数据安全关键技术简介	115
C.1 新兴数据安全技术	115
C.2 成熟数据安全技术	147
D.2020 年以来重大数据安全事件分析	184
D.1 勒索软件持续流行，对不同国家不同行业造成的损失案件不计其数	184
D.2 网络攻击，热点事件背后看不见硝烟的战场	185
D.3 内部员工泄露和删除数据事件频发	186
D.4 非法数据交易造成数据泄露	187
D.5 商业间谍窃取竞争对手商业秘密	188
D.6 系统或权限配置不当造成数据泄露	188
D.7 第三方人员或软件漏洞造成数据泄露	189
D.8 疫情之下，远程办公导致数据泄露事件频发	190
D.9 人脸信息泄露，对个人生活造成危害	191
D.10 儿童数据泄露问题日趋严重	192
E.2020 年以来典型数据安全相关法律案件分析	193
E.1 滴滴因数据安全问题，被启动网络安全审查	193
E.2 科勒卫浴违规收集人脸信息，遭“3·15 晚会” 曝光并被罚款 50 万元	193
E.3 某科技公司利用爬虫窃取简历数据，被判 4000 万元罚款	194
E.4 淘宝 12 亿用户数据遭爬取，危及用户个人信息安全	194
E.5 圆通员工泄露用户数据，引发负面舆情	195
E.6 多家银行因网络安全及个人信息保护合规问题，被处以高额罚款	195
F. 主要数据安全相关标准汇总	196
F.1 国家标准汇总	196
F.2 行业标准汇总	197
G. 国内外相关理论与介绍	200
G.1 微软的整体化数据治理方法 DGPC	200
G.2 高德纳的数据安全治理框架 DSG	204
G.3 DSMM 能力成熟度评估模型	206
G.4 信通院 DSG 框架	207



《数据安全治理白皮书 4.0》版本修订说明

《数据安全治理白皮书 4.0》版本中，针对以下内容进行修订完善：

- 1、添加数据安全形势与挑战，对当前国内、国际数据战略、最新的数据安全风险态势、监管形势等进行剖析；
- 2、添加数据安全治理目前的痛点和难题分析，分析当前治理工作的关键需求；
- 3、深化数据安全治理理念，完善数据安全治理框架，扩展治理框架中的运营体系、技术体系及治理规划建设；
- 4、添加国内外数据安全相关的法律法规的体系解读，对数据安全治理建设提供合规落地指引；
- 5、基于当前数据安全发展形势，更新有代表性的场景化行业领域数据安全治理实践案例；
- 6、更新数据安全发展的未来展望与倡议；
- 7、完善数据安全关键技术简介，按新兴和成熟技术分类补充数据安全技术介绍；
- 8、更新 2020 年度以来国内外重大数据安全事件分析；
- 9、添加 2020 年度以来典型数据安全相关法律案件；
- 10、补充国内外相关理论和介绍；
- 11、汇总最新的重要国家与行业的数据安全相关标准列表；
- 12、全文内容文字和段落结构优化。



第一章 数据安全形势

1.1. 重新认识数据的价值

数据是数字经济时代的核心生产要素。数字经济是继农业经济、工业经济之后的主要经济形态，是以数据资源为关键要素，以现代信息网络为主要载体，以信息通信技术融合应用、全要素数字化转型为重要推动力，促进公平与效率更加统一的新经济形态。对数据要素掌控和利用能力，已成为衡量国家之间竞争力的核心要素。近年来，随着《网络安全法》、《数据安全法》、《个人信息保护法》等上位法的发布，数据要素掌控和利用过程中面临刚性的数据安全合规要求。加强数据安全保护，既是数据产业自身发展的主观需要，也是国家监管的客观要求。但我国数据产业正处于发展初期，数据利用与数据安全的平衡兼顾存在“对数据资产识别不完整”、“数据流动路径无法溯源”、“安全与业务难以兼顾”等现实问题，与国外大多数国家一样，数据利用与数据安全的协同发展在很长一段时间内将是我国数据产业发展的主要矛盾。

1.1.1. 我国数据发展战略

国家对数据资源十分重视，强调了国家数据战略对于中国经济社会发展的重要地位和时代意义，而确保数据安全则具有极为重要的战略意义。

2015年8月，国务院印发《促进大数据发展行动纲要》，明确提出数据已成为国家基础性战略资源，并要求完善法规制度和标准体系。

2016年11月，第十二届全国人民代表大会常务委员会通过了《网络安全法》，鼓励开发网络安全数据保护和利用技术，同年，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。

2020年4月，中共中央、国务院发布《关于构建更加完善的要素市场化配制体制机制的意见》，将数据与土地、劳动力、资本、技术并称为五种要素，提出加快培育数据要素市场。5月，《关于新时代加快完善社会主义市场经济体制的意见》中提出进一步加快培育发展数据要素市场。这意味着，数据已经不仅是一种产业或应用，而已成为经济发展赖以依托的基础性、战略性资源。同时，还要求制定数据隐私保护制度和审查制度，推动完善适用于大数据环境下的数据分类分级安全保护制度，加强对政务数据、企业商业秘密和个人数据的保护。

2020年10月，中国共产党第十九届中央委员会第五次全体会议通过《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》，对“十四五”时期我国经济社会发展作出了贯彻新发展理念、构建新发展格局的重要部署，并作出了我国进入新发展阶段的重要判断。新发展理念的核心是创新，创新是发展的第一动力。随着数字经济蓬勃发展，数据保存量逐年倍增，建设安全的数据基础设施将是支撑数据产业长期可持续发展的重要要素。

2021年10月，国资委、工业和信息化部签署《关于加快推进中央企业两化融合和数字化转型





战略合作协议》，共同推动中央企业加快信息化工业化融合和数字化转型，促进数字技术与实体经济深度融合，为建设制造强国、网络强国、数字中国和发展数字经济更好发挥引领支撑作用。

2021 年 11 月，工业和信息化部发布《“十四五”大数据产业发展规划》，《规划》围绕加快培育数据要素市场、发挥大数据特性优势、夯实产业发展基础、构建稳定高效产业链、打造繁荣有序产业生态、筑牢数据安全保障防线六个方面提出重点任务，设置数据治理能力提升、重点标准研制及应用推广、工业大数据价值提升、行业大数据开发利用、企业主体发展能级跃升、数据安全铸盾六个专项行动。

2021 年 12 月，中央网络安全和信息化委员会办公室（以下简称中央网信部门）印发《“十四五”国家信息化规划》（以下简称《规划》），《规划》指出“十四五”时期，信息化进入加快数字化发展、建设数字中国的新阶段，数据要素是参与社会生产经营活动，带来经济效益，以电子方式记录的数据资源，是国家发展的战略性基础性资源，也是驱动数字经济发展的强大动力。加强数据治理、深化数据开发、保障数据安全是释放数据资源价值的关键环节。《规划》提出建立数据要素资源体系，以数据治理为突破提升数据质量，以数据开发利用为抓手激活要素价值，以立法规范为重点保障数据安全，加快完善与我国发展实际相符合的数据要素资源体系，释放数据要素价值。

2022 年 1 月，国务院印发《“十四五”数字经济发展规划》指出要充分发挥数据要素作用，必须强化高质量数据要素供给、加快数据要素市场流通、创新数据要素开发利用机制。提出着力强化数字经济安全体系，需要增强网络安全防护能力、提升数据安全保障水平、切实有效防范各类风险。

2022 年 3 月，国务院总理李克强作 2022 年政府工作报告，报告中提到推进国家安全体系和能力建设，强化网络安全、数据安全和个人信息保护。这是自 2021 年政府工作报告以来，再次对数据安全和个人信息保护的强调。

1.1.2. 国外数据发展战略

1) 美国

美国在发展大数据方面全球率先启动，并将大数据上升为国家战略，并制定了较为完善的政策体系。美国的数据战略尤其在奥巴马和特朗普任期更为凸显。近年来美国推出了一系列大数据战略举措，并同时重视数据安全，强调数据安全。

早在 2009 年，美国政府推出 Data.gov，使数以千计的政府数据资源以机器可读形式发布到该网站，向公众免费开放。

2011 年美国公布《网络空间可信任身份的国家战略（The National Strategy for Trusted Identities in Cyberspace）》，为公共和私有部门合作，增强网上身份的隐私和信任提供路线图。

2014 年，奥巴马演讲中要求了解公共和私有部门面临的大数据挑战。

2018 年，签署并通过《澄清域外合法使用数据法案》（Clarifying Lawful Overseas Use of Data Act，以下简称 CLOUD 法案），确立了境外数据管辖控制者标准，使其合法访问境外数据，并对他国获取数据资源的资格进行限制，通过“长臂管辖”的方式不断扩大其跨境数据的执法权力，甚至抢夺他国的“治外法权”。与此同时，美国还大力推动 APEC（亚洲太平洋经济合作组织）的跨境隐私



保护准则（APEC Cross Border Privacy Rules，以下简称 CBPR），规定若不同国家的不同企业承诺遵守 APEC 隐私框架中的 9 项个人数据保护原则，则数据就可以在上述公司进行自由的、跨境地流动，从而实现数据向美国流动，掌握全球数据资源。

2019 年 12 月，美国白宫行政管理和预算办公室（Office of Management and Budget, OMB）发布了《联邦数据战略与 2020 年行动计划》（以下简称《联邦数据战略》），明确提出“将数据作为核心战略资源开发”。其内容包括 1 项使命宣言、10 项原则、40 项实践指导及年度行动计划四个组成部分。《联邦数据战略》是美国把科技创新和数字化转型提到国家战略核心层面部署的重要体现，描述了美国联邦政府在未来十年加速数据使用的愿景，并逐年确定行动计划，并突出了对数据认识的深化，坚持数据是最有价值的国家资产理念，并着重改进特定数据资产组合的管理和使用。

2020 年 10 月，美国国防部发布了《国防部数据战略》（《DoD Data Strategy》），该战略是继 2019 年美国国防部发布的《国防部云战略》、《国防部人工智能战略》和《国防部数字现代化战略》以后又一 IT（信息化）现代化领域指导性战略文件，其中提出国防部应加快向“以数据为中心”的过渡，强调通过数据融合实现军种联合、重视数据的安全性、强调数据全生命周期各个阶段的标准化处理能力。

2) 欧盟

欧盟的战略目标是确保其成为数据世界的榜样和领导者，为了实现这一目标，欧盟在数据保护、公民权利和数据安全等方面构建了完善的法律框架，并指导建立欧盟内部市场，同时加强环境保护、提升治理透明度，通过构建“泛欧数据市场”，以“内松外紧”的形式主导其数据战略。

近年来，欧盟委员会采取了一系列措施加强数据治理工作：

2016 年颁布了《通用数据保护条例》（General Data Protection Regulation, GDPR），经过两年缓冲期后于 2018 年 5 月份正式生效，其注重强调消费者的权利和提升效率，自其 2016 年出台以来就被人们普遍认为是当前世界范围内最为严格的数据保护法案《通用数据保护条例》。

2019 年，欧盟出台了《欧盟非个人数据自由流动框架条例》（Regulation on a Framework for the Free Flow of Non-Personal Data in the European Union），与《网络安全法案》和《开放数据指令》相辅相成，共同促进欧盟内部的数据流动。此外，欧盟还启动了数字外交计划，认可了 13 个为个人数据提供足够保障的国家。

2020 年 2 月，欧盟发布欧盟数据战略（A European Strategy for Data），指出“确保欧盟称为数据驱动型社会的领导者，以便公共和商业部门能利用数据更好地进行决策”。全文包括背景介绍、关键点、愿景、问题、战略、国际路径、结论以及附录（欧洲战略部门和公共利益领域公共数据空间创建计划）等八个部分，数据战略阐述了欧盟未来五年实现数字经济所需的政策措施和投资策略。该战略积极推进数字化转型工作，提出要打造一个共同的欧洲数据空间且面向世界开放，构建大规模数据池以支持欧盟内部成员之间的数据交流和使用，提升企业竞争力，从而使欧盟在数据发展浪潮中占得先机，同时最大程度地减少人为碳排放和环境破坏。

2022 年 2 月，欧盟发布数据法案（Data Act）草案，全文包括涉及数据共享、公共机构访问、国际数据传输、云转换和互操作性等方面规定，旨在确保数字环境的公平性，刺激竞争激烈的数据市场，为数据创新驱动提供机会。该法案主要的监管对象为互联网产品的制造商、数字服务提供商





和用户等。

1.2. 数据安全风险态势

1.2.1. 外部数据安全威胁持续升级

业务数字化、信息系统云化、安全边界模糊化等众多正在加速演进的时代发展趋势，使得企业或组织面临的外部数据安全威胁也在相应发生变化。

在数据时代，通过漏洞利用、防护绕过等手段侵入企业或组织的内部网络实现数据窃取或破坏的安全事件仍常有发生，但随着网络安全防护设施的普及和加强，明显增加了侵入内部网络的难度。伴生而来的新的攻击和外部数据安全威胁层出不穷，导致数据的窃取、篡改和非法使用等威胁。

1) 数据窃取

● 通过向软件中植入木马窃取数据

数据盗窃团伙定制开发大量已植入木马的盗版应用程序和手机 App，并在各种下载网站和应用商店广泛发布，诱骗和等候用户下载安装；更猖獗的方式是向电商、酒店等服务行业使用的客户端软件中植入木马，批量盗取电商和酒店客户的个人隐私数据。

● 盗用特权账户导致数据泄露

特权账户的防御往往是保护数据泄露的最后一道防线。据国际权威咨询机构 Forrester 最新报告显示，80% 的数据泄露事件都与特权账户失窃有关。美国电信巨头威瑞森从 2018 年发生的 5.3 万多起安全事件中的调查结果显示，81.1% 的安全事件通过窃取身份凭证实现。盗用凭证仍然是黑客最常用的攻击手段，是造成数据泄露最重要的因素之一。

● 数据委托处理时被合作方盗用数据

为更好地挖掘和变现数据中的价值，数据融合和委托分析正变得越来越普遍。在此过程中，某些数据分析服务提供商会因觊觎数据价值，超出委托协议范畴盗用或滥用数据。

● 关联利用已泄露数据发动撞库攻击

数据盗窃团伙会相互交换或到暗网购买已经泄露的用户名和密码数据用来撞库，撞库成功后，再爬取这个账户的数据和信息，并会将其用于开展进一步的数据黑灰产活动。

● APT 高级可持续威胁攻击窃取组织核心机密

APT 攻击（高级可持续威胁攻击），通过具有专业水平的攻击人员和丰富的攻击资源，有组织的对特定对象展开持续有效的定向威胁攻击，隐蔽地运用绕过传统安全方案（如防病毒软件、防火墙、IPS 等）、供应链攻击和社会工程学等一切手段实施持久且有效的威胁和攻击，在目标基础设施上建立并扩展立足点，并进行长期潜伏，逐步学习和适应安全防护措施，对目标进行高水平、低频次交互攻击，窃取组织核心机密。

2) 数据篡改

● 用勒索病毒加密数据敲诈赎金

勒索病毒是一种新型电脑病毒，主要以邮件、程序木马、网页挂马的形式传播，病毒会利用各



种加密算法加密感染者的数据文件，感染者若想拿到恢复数据的解密私钥，需要按照病毒的敲诈提示信息，支付高额赎金。

● 利用区块链漏洞攻击智能合约

随着区块链的应用越来越广泛，针对区块链的数据攻击主要体现为两类：一是利用智能合约的漏洞发起对智能合约的重入攻击、溢出攻击、所有权攻击和拥塞攻击，导致智能合约检查失效或发生节点仿冒及优先级提升；二是对未正确的进行匿名化操作进行隐私保护的用户，通过交易双方信息推断出地址关联性，造成用户隐私泄露。

3) 数据非法使用

● 利用网络爬虫抓取和分析隐私数据

数据掘金团体利用爬虫批量访问企业或组织向客户提供的正常服务接口大量抓取数据，然后，再通过与其他渠道获取的具有相关属性的批量数据进行关联，尝试从这些普通数据中分析出具有高风险的敏感数据。例如通过追踪用户的 ID 来推断用户注册时间、通过电商评论推断购买行为、通过统计抽样获悉竞争对手商业机密等。

数据价值和重要性的持续提升，刺激了贩卖数据牟利的作案动机，窃取数据成为大规模、有组织网络犯罪集团的首要目标。而且，随着大国间战略博弈态势的加剧，网络斗争与政治对抗的交融日趋紧密，针对国家高价值重要数据及相关基础设施的网络攻击也正在变得“高发、高危、扩散、伴生、难防”。再加上安全本身就存在“易攻难守”的特点，进一步导致了数据安全攻防双方当前力量对比的严重失衡，攻方已占据明显优势，外部数据安全威胁持续升级。

1.2.2. 内部数据安全风险日益严峻

相比持续升级的外部数据安全威胁，日益严峻的内部数据安全风险更是让企业或组织感到迫在眉睫。

1) 内部人员有意或无意行为引发的数据安全风险

比特币、暗网等增强网络匿名性技术的出现和发展、数据贩卖黑灰色产业链的产生和成熟，都给企业或组织内部人员利用数据非法牟取私利提供了更容易的变现途径。这种诱惑无疑会催生严重的内部数据安全风险。

● 数据泄露

➢ 越权访问攻击成为数据破坏、数据泄露的重要来源

由开发人员在数据增删改查时过于相信客户端请求的数据，从而遗漏的对权限的判定而产生的系统漏洞，可能造成批量数据破坏和泄露的严重数据安全事件，越权访问漏洞可分为水平越权和垂直越权两种。

➢ 数据跨境自由流动加剧数据安全风险

全球海量数据在网络空间中不断移动和流转，跨境数据中不可避免地涵盖了个人敏感信息、企业或组织运营数据和重要信息数据，并且随着数据规模不断地扩大、数据种类不断的丰富，数据安全风险也在日益加剧。





● 数据窃取

➢ 内部人员窃取商业秘密非法牟利

客服或销售人员可能会将一些重要客户的信息数据卖给企业的竞争对手；即将离职的员工可能窃取秘密数据后投奔竞争对手或者自立门户。

➢ 员工亲友冒用员工身份窃取数据

随着员工在家办公或使用自己设备办公的情况日渐普及，员工的亲友越来越容易使用员工的电脑、手机等终端设备以员工的身份访问和窃取企业或组织的数据。

➢ 研发人员向业务系统中放置后门

研发人员可能在编写代码时向业务系统中放置后门，等到系统被部署上线后，再悄悄利用后门盗取数据。

● 数据篡改

➢ 业务人员违规篡改数据牟利

能够接触考试成绩、交通违章记录等数据并对其拥有修改权限的业务人员，因收受贿赂、帮助亲友、线上非法交易牟利等目的，违规篡改数据。

➢ 研发人员向业务系统中放置后门

研发人员可能在编写代码时向业务系统中放置后门，在业务系统上线后，再利用后门篡改业务数据。

● 运维人员为泄愤恶意破坏数据

拥有数据和系统高级操作权限的运维人员，可能出于个人恩怨等报复动机，对企业或组织的重要数据采取篡改、删除等破坏行为。

● 数据意外丢失风险

造成数据失效的原因很多，包括：云平台故障、存储故障、服务器故障、软件故障、设备老化、火灾、机房漏水、机房结构性破坏、供电系统瘫痪、空调故障、人为破坏、人为误操作、病毒感染等，最主要问题是硬故障和人为因素。

2) 敏感个人信息非法利用严重侵害个人权益

● 敏感个人信息过度收集对个人权益造成巨大损害

组织未遵循法律法规要求的最小化、必要性原则，过度收集用户个人信息，虽然会给组织带来商业利益，但也潜藏着巨大的风险。

过度收集用户敏感信息且未经用户同意滥用用户数据，以及数据保存不当发生泄露等情况，将严重侵犯用户隐私，甚至给用户个人造成巨大的经济损失和精神伤害。

● 大数据杀熟严重侵害消费者权益

大数据杀熟是指通过收集用户个人信息，运用大数据和算法针对用户的消费能力建立画像，进而针对同一价格和服务对不同用户进行区别定价的情况。大数据“杀熟”主要表现为三种类型：一是根据用户使用的设备差别定价，如针对高端设备的用户，其商品定价更高；二是根据用户消费频率差别定价，针对消费频率高的老用户定价更高，针对低频消费用户定价较低；三是根据用户消费时所处的场所差别定价，如针对购买场所较远的用户定价更高。与不同消费场所商品或服务定价不





同的相比，“大数据杀熟”处于隐蔽状态，大数据杀熟背离了诚信和公平原则，导致了商业伦理的缺失，给组织造成了信任危机。

3) 业务频繁变化引起的数据误用、滥用

在企业或组织数字化转型过程中，难免要不断引入新的业务系统或者对旧业务系统进行升级改造，因培训不足、操作欠熟练、个人疏忽等原因导致员工非法利用数据，如误用、滥用等行为正在变多变频，这也是企业或组织在未来一段时间内需要积极准备应对的内部数据安全风险之一。

1.2.3. 数据安全事件危害性愈发严重

1) 危害程度日益加剧

根据风险基础安全（Risk Based Security）的数据显示，2020 年全球数据泄露达到 360 亿条。数据安全威胁更加多样化，集中在弱口令、特权账号访问、非法访问、越权访问等。

IBM《2021 年数据泄露成本报告》，更是指出每次数据泄露事件平均为公司带来 424 万美元的损失，为 17 年来之最。其中，大型数据泄露的平均成本达到 4.01 亿美元，泄露记录达到 5000 万到 6500 万条。

据 Verizon《2021 年数据泄露调查报告》指出，2021 年数据泄露的主要原因是 Web 应用程序攻击、网络钓鱼和勒索软件，其中 85% 涉及人为因素。该报告分析了 79635 起事件，其中 29207 起符合分析标准，5258 起确认是数据泄露，这些事件来自全球 88 个国家。

据中国软件评测中心持续对电信和互联网行业的威胁监测和远程检测结果分析，发现安全漏洞或问题 80% 和数据安全相关，主要包括 SQL 注入、非授权访问、数据泄露三大类。其中实现非授权访问的原因多样，包括弱口令、授权绕过、未进行身份验证等。数据泄露方面甚至存在源代码泄露。

2) 影响范围不断扩大

从受到影响的对象来看，数据安全事件的影响范围已经从最初的企业或组织和个人，逐步延伸到国家各个行业、整个社会，甚至可能会引发国际性的舆论关注。各国非常重视大数据在国家安全领域的运用，随着各国的数据战略部署，数据治理逐步上升到国家战略层面，某些数据安全事件极有可能发展为影响社会秩序、政治稳定、国家安全的非常事件。

从促进行业数字化转型和国家数字经济发展的角度看，数据安全问题会从多方面产生显著的负面影响：首先，会抑制行业数据合理开放共享的意愿和积极性；其次，会增加公众和运营企业对于个人信息和隐私数据的收集和使用以及对于包括重要数据在内的数据交易、共享和价值挖掘活动的顾虑，从而阻碍基于数据的新技术、新业务与实体经济的深度融合与创新；再者，数据安全问题给数据的跨境流通、驻留规管增添了新的难度与挑战。

1.3. 监管新举措

面对严峻数据安全风险态势，为推动数据依法合理有效利用，保障数据依法有序自由流动，近年来，根据国家数据战略布局，我国数据安全相关法律法规密集发布，中央网信部门统筹协调、行业及地方主管部门各司其职的数据安全监管模式正在逐步形成。





按国家、行业及地方政府各个层面列举说明数据安全相关法律、法规、规章如下：

序号	监管层面	法律法规	发布时间	备注
1	国家层面	《数据安全法》	2021 年 6 月	将个人、企业和公共机构的数据安全纳入保障体系，确立了对数据领域的全方位监管、治理和保护，既约束了数据的非法采集和滥用，又保护了数据提供方和民众的信息使用，推动以数据开放、数据保护、数据流动等为基础的数据规则进一步完善。作为我国数据安全领域内的基础性法律和我国国家安全领域内的重要法律，给企业数据经营合规、以及进一步的数据资产化治理与发展提供指引。
2		《个人信息保护法》	2021 年 8 月	规定个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息；个人信息处理包括个人信息的收集、存储、使用、加工、传输、提供、公开等活动。确立了个人信息处理应遵循的原则，强调处理个人信息应当采用合法、正当的方式，具有明确、合理的目的，限于实现处理目的的最小范围，公开处理规则，保证信息准确，采取安全保护措施等，并将上述原则贯穿于个人信息处理的全过程、各环节。
3		《网络数据安全管理条例（征求意见稿）》	2021 年 11 月	进一步规范了互联网产业的数据要求，包括爬虫的使用、第三方插件的个人信息收集、隐私协议的合规要求、互联互通义务等，并对重要数据进行了定义，明确了数据分类分级保护制度和跨境数据提供的条件，建立数据安全审计制度，促进数据依法合理利用。
4	行业层面	《银行业金融机构数据治理指引》	2020 年 5 月	引导银行业金融机构加强数据治理，提高数据质量，充分发挥数据价值，提升经营管理水平，全面向高质量发展转变而制定的法规。明确了监管机构的监管责任、监管方式和监管要求。对于数据治理不满足有关法律法规和监管规则要求的银行业金融机构，要求其制定整改方案，责令限期改正；或与公司治理评价、监管评级等挂钩；也可依法采取其他相应监管措施及实施行政处罚。
5		中国银保监会监管数据安全管理办法（试行）	2021 年 1 月	规范了银保监会监管数据安全管理工作，提高了监管数据安全保护能力，目的是建立健全监管数据安全协同管理体系，推动银保监会有关业务部门、各级派出机构、受托机构等共同参与监管数据安全保护工作，加强培训教育，形成共同维护监管数据安全的良好环境。
6		《工业和信息化领域数据安全管理办法（试行）（征求意见稿）》	2021 年 9 月	强调建立工业和信息化领域重要数据和核心数据备案管理制度，统筹建设备案管理平台鼓励数据开发利用和数据安全技术研究，支持推广数据安全产品和服务，培育数据安全企业、研究和服务机构，壮大数据安全保障能力，促进数据的创新应用。
7		《金融科技发展规划（2022-2025 年）》	2022 年 1 月	强调全面加强数据能力建设，在保障安全和隐私前提下推动数据有序共享与综合应用，充分激活数据要素潜能，有力提升金融服务质效。



序号	监管层面	法律法规	发布时间	备注
8	地方层面	《贵阳市大数据安全管理条例》	2018 年 6 月	强化数据安全管理制度；立大数据安全审计、应急处置、监测预警机制，实现对重要数据安全风险的全天候实时、动态监测；健全数据安全监管制度、大数据安全投诉举报制度。
9		《天津市促进大数据发展应用条例》	2018 年 12 月	明确数据全生命周期各环节保障数据的范围边界、主体责任、具体要求；采取关键信息基础设施安全防护措施，加强防攻击、防泄露、防窃取的技术和管理能力建设。
10		《深圳经济特区数据条例》	2021 年 7 月	对于数据安全的监管和义务做了详细规定，其中规定市网信部门负责统筹协调本市个人数据保护、网络安全、数据安全、跨境数据流通等相关监督管理工作。市发展改革、工业和信息化等部门依照有关法律、法规，在各自职责范围内履行数据监督管理相关职能。
11		《上海市数据条例》	2021 年 11 月	该条例分为总则、数据权益保障、公共数据、数据要素市场、数据资源开发和应用、浦东新区数据改革、长三角区域数据合作、数据安全、法律责任和附则，后续上海还将制定数据分级分类保护、重要数据目录管理、数据安全管理制度等配套措施，落实重要数据备案和数据安全评估制度。建立健全全流程数据安全管理制度，组织开展数据安全教育培训等。
12		《江苏省公共数据管理办法》	2021 年 12 月	对公共数据供给、共享、开放、利用与安全等提出了明确要求。立足规范公共数据管理的全流程、构建全域数字化发展的公共数据底座。
13		《浙江省公共数据条例》	2022 年 1 月	强调公共数据安全管理，对公共数据收集与归集、公共数据共享、公共数据开放与利用、公共数据安全、法律责任等方面作出明确规定。为了有效推进公共数据的安全管理，该条例设置公共数据安全专章，对加强公共数据安全管理、规范公共数据安全行为等都做出了体系化、实质性的制度设计。
14		《山东省公共数据开放办法》	2022 年 1 月	拓展了公共数据的适用范围，明确国家机关，法律法规授权的具有管理公共事务职能的组织，具有公共服务职能的企业事业单位，人民团体等在依法履行公共管理职责、提供公共服务过程中，收集和产生的各类数据均属于公共数据，应纳入公共数据开放办法管理，强调重点和优先开放就业教育等数据，并公共数据开放可“容错免责”。

法律法规详细解读参见第四章法律法规标准解读章节。





第二章 数据安全治理目前面临的主要痛点和需求

伴随数字化转型成为国家十四五的重点规划，数据要素市场化建设上升到国家战略高度，基于前序章节对数据安全形势与法律法规的分析、解读，企业或组织在数据收集、存储、使用、加工、传输、提供、公开等处理活动场景中，数据流转过程中暴露面风险急剧加大，面临严峻的数据安全风险防护和合规监管要求形势，在安全风险与监管合规双重驱动下，为促进数据发展与数据安全的平衡发展，各企业或组织纷纷开展数据安全治理工作，鉴于数据动态、频繁产生与流转带来的数据安全防护复杂性，如何在数据释放价值的同时，保障数据安全，推动平衡发展？企业或组织在治理实践过程中，面临着诸多的痛点和难题，本白皮书从数据安全治理、技术、运营层面对亟待解决的主要痛点和关键需求进行了梳理和归纳。

2.1. 管理层面

2.1.1. 多法并轨、多标准并行下缺乏一致性的合规治理手段

● 痛点

数据安全治理领域以《网络安全法》、《个人信息保护法》、《数据安全法》、《关键信息基础设施安全保护条例》、《网络数据安全条例（征求意见稿）》、《网络安全等级保护条例（征求意见稿）》构成“三法三条例”体系，成为了网络安全防护和数据安全治理的重要合规驱动力。在此前提下，企业或组织需要积极响应相关法律法规及相应的标准规范，落实不同维度、不同侧重的各类监管合规要求。如何应对众多且仍持续颁布的法律法规的合规要求，是摆在企业或组织面前的合规遵循难题。

● 需求

“上面千条线，下面一根针”，面向客观的数据处理活动场景中数据流动，应对合规要求，建立与各法律法规相适配的数据识别、监测与防护策略映射，实现一致性的合规治理。

2.1.2. 合规治理下数据安全实施细则尚不完善

● 痛点

随着有关数据安全和个人信息保护相关法律和法规要求发布，各单位也纷纷开展相关的安全自查和整改工作，逐步推进数据安全治理工作的落地。

面向数据的收集、存储、使用、加工、传输、提供、公开等处理活动及数据跨境流动，存在数据所有权与控制权（或使用权）分离的复杂场景，由于相关细化的数据安全监测与管控等实施细则建设尚不完善，不健全，导致数据流动控制与数据确权困难。同时，频繁的数据流动是数字经济时代的显著特征，如果管理不善，可能会侵犯个人隐私、泄露商业秘密，甚至会泄露国家秘密和威胁



国家安全。

- 需求

需要面向数据处理活动场景，遵循法律法规要求，制定完备的数据安全管理制度、实施细则与技术标准规范，并建立相应的数据安全治理组织，定岗定责，贯彻落实管理规定。

2.1.3. 传统的网络安全风险评估应对数据安全适应度不足

- 痛点

《数据安全法》、《个人信息保护法》、《关键信息基础设施安全保护条例》等数据安全相关法律法规中都要求落地数据安全风险评估，以及时发现数据安全风险，确保数据安全风险可控。数据资产具备流动性，会在不同的载体和场景环境下流动，且其资产价值也依据量级、周期性等因素动态发生变化，而传统的信息安全风险评估主要是面向网络环境下的数据安全载体资产，基于某个标准作为基准来设置评估项，展开相对静态、固化的风险评估，无法顺应数据流动过程中不同环境、不同目标下的安全评估要求。

- 需求

以网络安全风险评估的框架为基础，面向数据本身及其数据处理活动，围绕资产的重要程度、面临的安全威胁、脆弱性及安全措施等评估维度，在数据资产识别、法律法规遵从、数据处理活动、数据跨境流动、数据支撑环境等方面建设针对特定数据应用场景的安全风险评估机制。

2.1.4. 数据安全防护和网络安全防护体系尚未形成有机融合

- 痛点

企业或组织在网络安全管理上，通常已按照等保合规要求，在安全物理环境、安全区域边界、安全网络通信、安全计算环境、安全管理中心等方面建立了较为完备的网络安全防护体系，包括网络安全管理组织与制度流程、技术体系及运维保障体系，形成了面向终端、主机、网络、应用等数据载体的安全防护能力，在数据安全防护层面，也开展了针对数据本身的数据分类分级、访问控制、数据加密、数据审计、数据脱敏等单点数据安全防护能力建设，但是面向数据本身的数据安全防护与面向数据载体的网络安全防护体系尚未形成有机融合，缺乏围绕数据生命周期流转过程的整体化防护管理能力。

- 需求

针对数据安全体系的构建，需要以网络安全体系等保合规建设成果为基础，充分依托现有成熟的网络安全管理组织及人员、网络安全管理制度及流程、运维保障机制，扩展以数据为中心，围绕数据处理活动场景的数据安全管理体系，并将数据安全和网络安全进行充分融合，整体保障数据安全建设，促进企业或组织业务有序发展。



2.2. 技术层面

2.2.1. 面向海量数据的数据资产梳理与分类分级实施难度大

● 痛点

随着各企业或组织网络化、信息化和智能化水平的不断提高，来自互联网数据、移动数据、社交媒体数据和物联网等新型数据同原有业务系统的数据日渐庞大。但是由于数据分类分级的行业和企业属性较强，目前国家和行业配套的标准体系仍在制定或征求意见的过程中，细则尚不完善，不足以支撑落地数据分类分级工作的全行业落地，针对数据分类分级的行业级、精细化标准要求强烈。另一方面，很多未发布行业数据分类分级标准的企业担心自发的数据分类分级建设不符合后期发布的标准和要求，需要重新调整，抑制了各单位落地数据分类分级工作的积极性。在技术维度，现实中组织往往存在数据量大、结构复杂的问题，海量多源异构的原始数据量难以对每一条数据进行分类分级管理，在数据交换、共享等过程中，数据不可避免要脱离原有载体，由此引出需要再次进行数据分级分类并设置安全防护策略，如何快速摸清数据底账并保持数据分类分级的一致性以及防护策略的有效性，是企业或组织当前数据安全治理工作面临的难题。

● 需求

建立自动化的数据识别发现能力，并基于行业级、精细化的分类分级标准，实现智能化、快速的分类分级，并根据分类分级成果进一步开展数据安全评估与动态防护策略设置。

2.2.2. 数据所有权与控制权（使用权）分离导致数据共享和开放困境

● 痛点

只有数据流转起来、加快数据利用，才能最大化挖掘数据价值。伴随着种类繁多的数据处理活动，数据来源愈发多样，数据源信息种类愈发复杂。数据作为特殊资产，其所有权和控制权（使用权）存在分离的情况，导致在数据共享交换层面面临着无法满足其安全共享和开发的困境。在数据通过共享、交换等过程离开组织之后，数据的跟踪与溯源问题变得愈加困难。从法律层面上讲，当前数据权属相关的法律法规尚未颁布，在这样的前提下作为数据使用者的组织如何向外部组织或者个人的数据所有者落实数据管理义务，是企业或组织面临的难点；从实际落地层面上讲，作为数据所有者的组织，通过交换、共享、委托等处理活动对外提供的数据流出了自己可控的安全域，对于此类数据的防范属于自身技术防护的“盲区”。作为数据使用者的企业或组织，在使用数据中，如何落实外部组织或个人的数据所有者提出的数据安全义务，并提供合规性检测证明，以及数据安全事件发生后的安全取证、责任鉴定等，存在数据监管的挑战。在企业或组织内部，数据处理活动涉及多个部门，主体多元、利益多元，数据安全与发展、秩序与突破，诸多价值目标在这一问题上交织碰撞，如何满足企业或组织内部各方利益关系，实现数据开发利用和安全合规的平衡需求的满足度，进一步影响了数据的共享和开放。





● 需求

建立面向数据的全链路的数据流动监测与数据泄露风险分析机制。有效监控数据共享和开放过程的安全风险，并通过访问控制、数据脱敏等防护手段，保障数据共享和开放的安全性，针对数据所有权和控制权（使用权）分离情况，构建“可算不可见”的隐私计算技术体系，实现跨组织、跨部门的安全共享和开放。

2.2.3. 单点安全防护能力无法有效应对复杂数据流动风险

● 痛点

就数据安全产品而言，目前市面上的数据安全产品提供的技术防护能力，如数据分类分级、认证及访问控制、加密、审计、脱敏等都是基于单点能力，而鉴于数据的流动性与易传播性，会在不同的主机、网络及应用等载体间进行流转和移动，依据木桶原理的短板效应，单点的安全防护手段，缺乏之间的协调联动能力，导致安全策略一致性差、管控效率低、全面性弱等问题，出现“头痛医头，脚痛医脚”的现象，无法发挥围绕数据全生命周期的整体防护合力。

● 需求

以“平台化、体系化、可视化、实用化”为出发点，建立一站式数据安全集中管控能力，对包括数据资产梳理、数据安全防护、数据行为审计、数据脱敏、数据安全运维、数据加密等在内的各类数据安全产品进行“统一部署、统一监控、统一管理”，实现面向各数据处理活动的事前防护、事中监测、事后审计的整体、动态安全防护能力。

2.3. 运营层面

2.3.1. 业务动态变化下按需管控的运营机制尚不健全

● 痛点

数据驱动业务发展，是数字经济时代的显著特征，顺应市场的发展，企业或组织的业务在持续发生的变化或调整，承载业务的数据也伴随的业务的变化也在动态的发生着改变，同时，外部的数据安全监管合规要求也在不断完善、调整，单纯依靠管理和技术构建相对静态的防护体系，无法及时跟进业务和合规的变化，存在安全策略设置滞后风险，导致出现数据资产新增或调整识别不及时、安全风险监测不准确、安全防护不到位的防护难题。

● 需求

以外部合规监管要求和内部数据安全风险防范为驱动，建立“可实用、可持续”的安全运营体系，覆盖日常运营，安全评估、应急和重保等运营场景，对数据资产、安全策略、安全风险和安全事件进行整体运营管控，全面跟踪监测数据流转过程、风险发现和事件处置进度，实现数据识别，安全合规，事件稽核，风险感知。



第三章 数据安全治理理念及框架

针对当前企业或组织开展数据安全治理面临的主要痛点和关键需求分析，构建管理、技术、运营相融合的体系化、系统化的数据安全治理需求日益强烈。本白皮书面向企业或组织，从治理理念、治理框架、规划建设三个方面，持续探索和演进适合我国、易于落地的数据安全治理体系之路。

3.1. 数据安全治理理念逐步深化

自 2019 年《数据安全治理白皮书 2.0》推出数据安全治理理念框架，并在 2021 年《数据安全治理白皮书 3.0》进行进一步诠释，数据安全治理理念框架取得越来越多的业界共识，并在越来越多政府、机构、企事业单位得以尝试与实践，其涵义伴随治理的落地开展也越来越丰富。

数据安全治理是以“让数据使用自由而安全”为愿景，旨在安全有序推动数据流动，平衡数据发展与数据安全，在我国易于落地的数据安全建设的体系化方法论，核心内容包括：

（1）满足数据安全保护（Protection）、合规性（Compliance）、敏感数据管理（Sensitive management）三个需求目标；

（2）以数据为中心的分类分级安全治理理念包括：分类分级（Classifying）、角色授权（Privilege）、风险评估（Assessment）、场景化安全（Scene）；

（3）数据安全治理的建设步骤包括：组织构建、资产梳理、策略制定、过程控制、行为稽核和持续改善；

（4）核心安全框架为数据安全人员组织（Person）、数据安全使用的策略和流程（Policy&Process）、数据安全技术支持（Technology）大部分。



图 1 数据安全治理理念



3.1.1. 治理“愿景”更加清晰

数据安全法正式发布，对数据的安全和发展在国家层面给出明确指示，提出“国家统筹发展和安全，坚持以数据开发利用和产业发展促进数据安全，以数据安全保障数据开发利用和产业发展。”，进一步表明数据加工使用、开放共享与数据交易将更加推进数据价值的体现，数据只有在使用中创造价值，而数据价值越大越需要保护，将治理愿景由“让数据使用更安全”更加清晰地表述为“让数据使用自由而安全”，构筑数据安全保障底座，更好地促进数据的自由运用、助力数据价值的释放，有效推动数据开发利用与数据安全的一体两翼、平衡发展。

3.1.2. 治理目标愈发明确

1) 数据保护

伴随着政府、企事业单位、金融机构、能源集团、运营商等各行各业的数字化转型进程，数据资产在企业或组织发展中的重要性不断提升，数据泄露、篡改、破坏导致的影响日趋严重，企业或组织自身的风险驱动的数据保护、治理需求愈发主动、强烈。

2) 监管合规

伴随国家层面对数据安全的重视，国家及行业涉及数据安全相关法律法规标准持续推出与完善，安全合规要求持续增加，监管力度不断加大，监管内容不断细化，面对众多的法律法规标准条文，需要围绕管理和技术要求进行解析，寻找合规途径，落实合规措施。

3) 敏感数据管理

以往各行业对敏感数据的保护仅仅是自己摸索，缺乏对敏感数据的标准定义，随着数据分类分级相关标准的发布，金融、运营商、工业等行业对敏感数据有了更清晰的标准定义，当前对国家核心数据、重要数据、敏感个人信息等敏感数据的编目界定也在进一步明确。

综上所述，以数据保护为目的，以安全合规为驱动、以敏感数据管理为核心的数据安全治理需求愈发明确。

3.1.3. 核心理念全面深化

1) 数据分类分级

数据资产保护的核心在于数据分类分级。

通过对数据资产识别、梳理及数据的有效理解和分析，对数据进行不同类别和密级的划分；根据数据的类别和密级制定不同的管理和使用策略，尽可能对数据做到有差别和针对性的防护，实现在适当安全保护下的数据自由流动，避免敏感数据的防护不足，非敏感数据的过度防护。

2) 角色授权

数据安全访问控制核心在于数据访问主体的角色授权。

数据安全访问分为两条线，客体数据线与主体用户线，在数据线上明确数据资产及分类分级梳





理后，另外一条用户线，通过业务流、数据流分析，梳理用户角色对数据的主客体访问关系，明确数据的访问角色以及数据的使用方式，在不影响数据资源正常访问的前提下，围绕不同分类和级别的数据，针对各用户角色赋予相应的访问权限，实现按需、动态的数据访问和使用安全。

3) 安全评估

是指企业或组织在业务经营模式发生调整时，信息系统上线或发生升级时，系统运营环境发生变化时或发生重大安全事件时等情况下，企业或组织通过数据安全自评估工作，发现当前面临的数据安全问题，找到安全防护差距并及时落实整改，保障数据安全。

数据安全评估是数据安全治理的重要基础，针对数据收集、存储、使用、加工、传输、提供、公开等数据处理活动以及数据出境风险，根据相关法规标准要求，围绕敏感数据处理活动与数据跨境转移涉及的数据本体内容及数据载体环境，按照风险分析的方法，运用工具测试、配置检查、旁站验证、人员访谈等手段，从管理脆弱性、技术脆弱性、面临的暴露面威胁及已采用的防护措施等多个角度进行风险评估，分析、梳理数据处理相关活动存在的安全风险，为后续针对性的完善管理制度与流程及防护技术手段提供依据，确保有的放矢地开展数据安全治理。

4) 场景化安全

数据安全治理的核心在于场景化安全。

随着数据要素市场持续健全和深化，数据加工、共享、开放等应用场景不断丰富，数据交易、数据出境等数据运用新模式持续涌现，不同用户基于业务、访问途径、使用需求，会产生不同的使用场景。一刀切、固化的防护方式已无法适应不同场景下安全防护要求，在保证数据被正常使用的目标下，基于不同的使用场景特点制定相应的数据安全策略。实现场景化的数据安全治理，能及时发现不同场景下的数据风险暴露面，使数据安全治理更具针对性，从而实现数据使用自由而安全。

3.1.4. 建设步骤指导实践

从组织构建到持续改善，是一个系统化的建设步骤，可有效指导数据安全治理全面建设的落地。

1) 组织构建

组建专门的数据安全组织团队，是作为数据安全治理建设的首要任务，是保证数据安全治理工作能够持续执行的基础。

2) 资产梳理

资产梳理是数据资产安全管理的第一步，通过资产梳理能够掌握数据资产分布、数据责任确权、数据使用流向等，使数据资产安全管理更全面。

3) 策略制定

掌握了数据资产概况后，需要制定安全策略来作为数据资产管控的安全规则。

通过数据分类分级与重要数据识别，区分人员角色权限及场景，制定针对性的安全策略，能够实现对敏感数据进行分级管控，使数据在生命周期中安全流动。

4) 过程控制

策略制定后需要将安全规则落地，通过数据安全管理体系、技术体系、运营体系的有效配合，

能够帮助企业或组织进行数据资产安全管理的过程控制。

5) 行为稽核

要对数据的访问过程进行审计，要判断这些数据访问行为过程是否符合所制定的安全策略；要对数据的安全访问状况进行深度评估，看在当前的安全策略有效执行的情况下，是否还有潜在的安全风险。

6) 持续改善

数据安全需要动态跟踪，持续改善。可通过资产梳理，持续掌握数据资产动态；通过预警演练，提升应急响应能力；通过数据安全评估，了解数据安全管控现状，持续优化安全策略等。

依据上述指导性的建设步骤，数据运营者可根据企业或组织的业务关注点、风险容忍度等实际情况，首先考虑核心业务且易实施、见效的防护手段先行实践，然后在治理深度上逐步构建形成体系化、全周期的数据安全防护体系，再在治理广度上逐步覆盖到数据运营全业务，并在过程中持续优化。详细的治理规划建设参考见“3.7 数据安全治理规划建设”。

3.1.5. 能力支撑推动治理框架设计

数据安全治理绝非是平地起高楼，与传统网络安全和数据治理既有紧密的关联性，又有面向数据的独特性，整体框架需要多体系间融合展开治理。

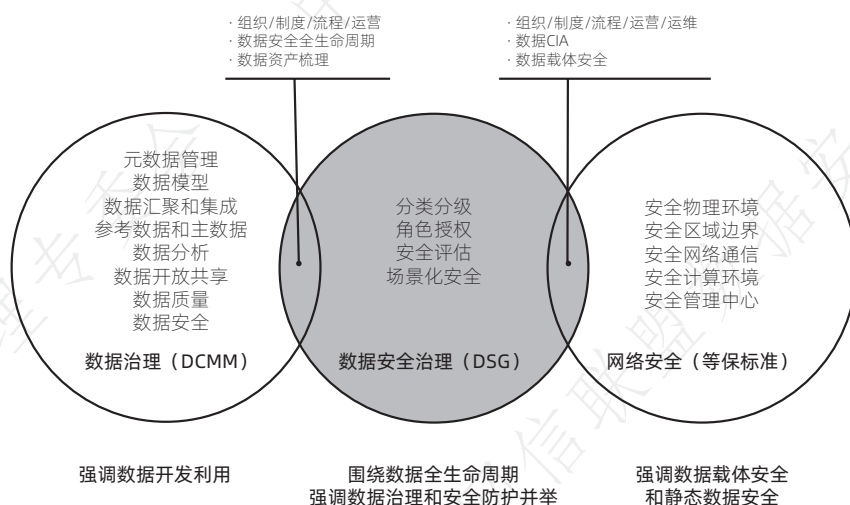


图2 数据安全治理与网络安全和数据治理关系图

在与传统网络安全关系方面：数据安全治理是以数据为中心，面向数据全生命周期的分类分级防护体系，其核心思想虽是面向业务数据流转的动态、按需防护，但在防护技术上，仍需依托网络安全等级保护中面向网络、设备、应用等数据载体的静态防护能力为基石，扩展面向数据本身的分类分级动态防护能力。安全管理上，仍需基于网络安全管理体系基础上，补充、完善面向数据安全管理的制度、策略和运营规范，形成围绕数据本身和数据载体的整体数据安全防护能力。

在与数据治理的关系方面：治理中数据分类分级的核心是数据资产的识别与梳理，可通过与数据治理中的元数据管理进行集成打通，直接获取元数据管理中的统一数据标准，作为数据分类分级



的资产识别输入，提高数据梳理准确性，避免重复工作，形成面向数据应用的数据治理体系与面向数据安全的安全治理体系的融合与统一。

围绕“让数据使用自由而安全”的愿景，结合实际的数据安全防护情况，数据安全治理能力划分为“组织人员、策略流程、技术支撑”三个核心能力领域，其中：

组织人员：建立数据安全治理团队，并明确团队中各成员的管理职责，团队组成依企业或组织具体情况，可以是实际组织也可以是各部门成员组成的虚拟组织，是数据安全治理工作开展的基础资源保障。

策略流程：设定相关的管理制度、标准规范、管理策略及流程，并围绕策略流程，构建运营管控机制，以运营思路开展数据安全治理工作，在制度、策略及流程时，充分考虑与传统网络安全运营管理的融合，实现“可持续化的数据安全治理能力”。

技术支撑：落实策略流程贯彻所涉及的识别、防护、监测、处置、溯源等数据安全防护技术建设，并融合数据治理和网络安全相关技术，形成完整的技术支撑体系。

通过专业的数据安全治理团队、明确的数据安全治理策略和流程、全面的数据安全运营机制、覆盖数据全生命周期的技术手段为支撑，围绕数据使用的业务场景活动，分析安全需求，在数据安全管理及技术等方面综合指导数据安全框架设计，提升数据资产的体系化保障能力。

3.2. 数据安全治理整体框架

在《数据安全治理白皮书 3.0》基础上，依据数据安全治理理念，围绕常态化安全运营服务需求与安全防护技术体系化、平台化演进趋势，数据安全治理框架重点扩展了可持续化的数据安全运营体系，完善了闭环管控的数据安全技术体系，深化了以数据安全管控策略为核心，以管理体系为指导，以运营体系为纽带，以技术体系为支撑的治理框架构建思路，构筑形成管理、技术、运营三位一体的数据安全治理框架，数据安全策略通过管理体系制定，通过安全运营体系发布，通过技术体系落地。依据形成的三大体系框架，动态、持续保障数据处理活动的有序、安全开展。

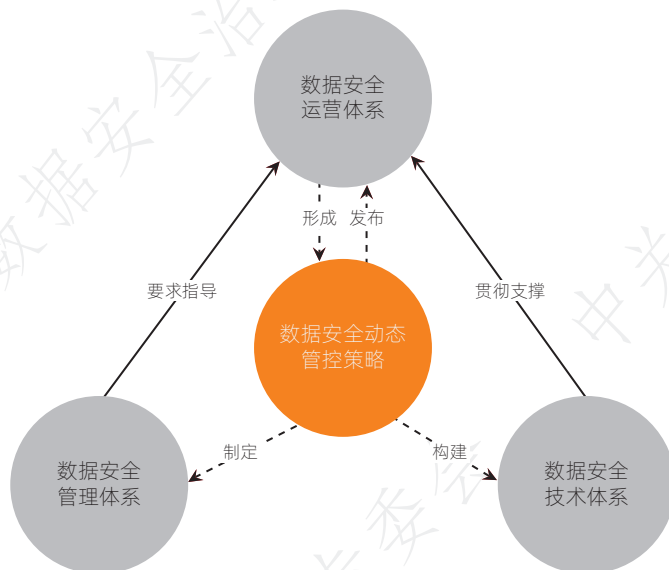


图3 数据安全治理体系关系图



作为帮助企业或组织应对数据安全挑战的一种理念和一套贯穿整个组织架构的建设数据安全能力的方法论，数据安全治理主要包括组织、制度、技术和运营四个方面的方案规划与实施。

在组织方面，需要从企业或组织内选派合适人员构成一支由决策层（高级管理人员）、管理层（安全部门）、支撑层（业务部门）、监督层（审计部门）构成的负责推动开展数据安全治理工作的团队。

在管理方面，深入调研和解读国家及本行业在数据安全方面的合规要求，建立数据安全管理制度四级文件，最终形成一级方针文件，二级制度规范文件，三级细则指引文件，四级表单、模板、记录文件。

在技术方面，需要基于组织人员架构和拟定的制度规范，选择和实施适宜的数据安全产品、服务等技术手段。

在运营方面，以组织业务安全需求和满足法律法规及行业标准为驱动力，建立一套日常化、集中化、规范化、流程化的数据安全运营工作方法。

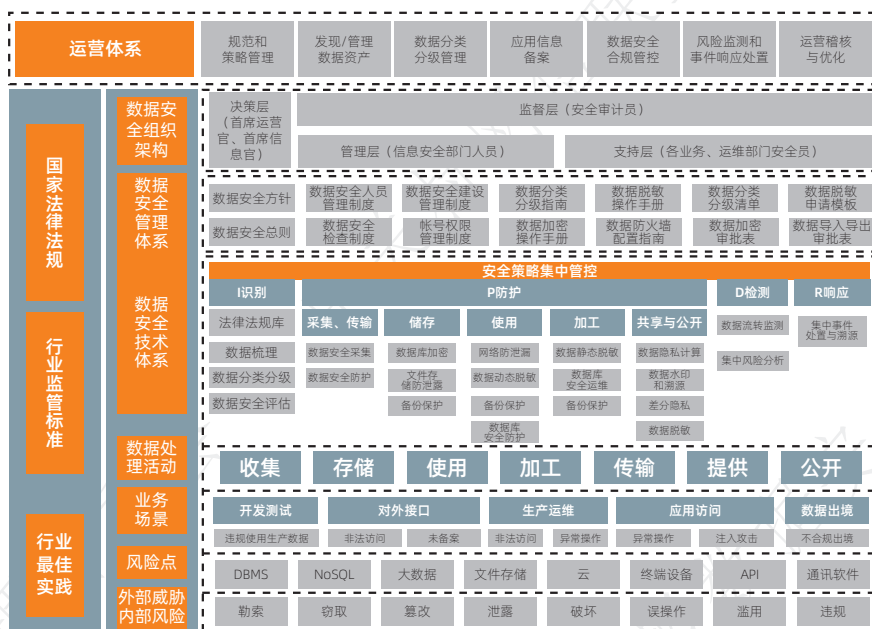


图 4 数据安全治理的总体框架

以数据为中心保护数据安全是数据安全治理的重要理念和发展趋势，因此，在数据安全治理过程中应充分考虑到数据处理活动的所有环节，包括数据的收集、存储、使用、加工、传输、提供、公开等。

数据安全治理必须要能够实现数据安全的全部核心目标，即保护数据的完整性（防篡改）、保密性（防泄露、防滥用）、可用性（防勒索、防破坏）。在某些具有特殊或更高安全需求的情况下，还经常会包括一些额外目标，例如可追溯性（不可否认性）、可靠性、可控性、真实性等等。

最后，可利用目前数据价值实现涉及的各种典型应用场景（如业务访问、运维管理、开发测试、数据智能分析、数据共享等）和数据环境（关系型数据库、非关系型数据库、大数据环境、文件存储系统、云、终端设备、数据访问 API、通信软件等）中存在的外部威胁（如勒索、窃取、篡改、泄露、破坏等）和内部风险（如窃取、篡改、泄露、破坏、误操作、滥用、违规等），对数据安全治理结果的全面



性和有效性进行考查、评估和深入分析。

3.3. 数据安全组织架构

3.3.1. 职能架构

企业或组织内部负责数据安全治理例行事宜的通常是一个常设的虚拟团队，一般称为数据安全治理委员会或数据安全治理小组。团队的职责是制定对数据进行分类、分级、保护、使用和管理的原则、策略和过程。团队的成员应包括企业或组织内的数据安全专家，以及所有与数据安全有关部门（如IT支持、人资、法律、财务、业务和市场、运营和维护、知识产权、保密等）的人员代表；在一些大型的企业或组织中，因为数据安全正日益变成企业或组织生死攸关的重要问题，甚至会包括主管副总裁、董事会成员等高级管理人员。

数据安全治理团队的成员同时也是企业或组织数据安全制度的受众。他们是数据安全策略、规范和流程的执行者和被管理者，同时也是数据的使用者、管理者、维护者、分发者。只有将这些角色的人员代表纳入到团队中，才能使得在数据安全治理中制定的安全原则、安全措施和安全规范能够在具体执行中得到有效贯彻落实。

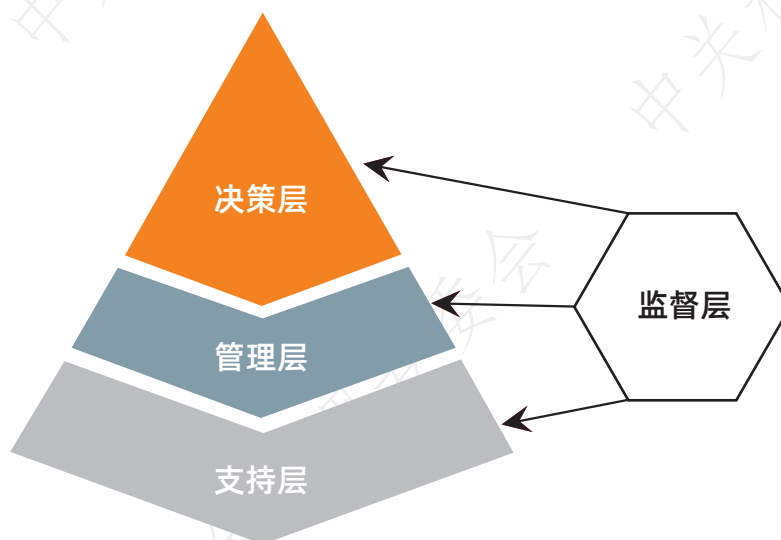


图5 数据安全治理团队的职能架构

数据安全治理团队常用的职能架构如上图所示，自顶而下依次为决策层、管理层、支持层，外加一个贯穿数据安全治理全程并负责对上述三层进行监督审计的监督层。各层的职能分工和成员建议如下：

1) 决策层

负责对企业或组织开展和实施数据安全治理的体系目标、范围、策略等进行决策。成员包括企业或组织内主管数据价值实现的最高负责人（如首席运营官、首席战略官等）和信息安全方面的最高负责人（如首席信息官、首席信息安全官等），甚至可以考虑由负责推动企业或组织数字化转型的高级副总裁或者负责战略新兴业务拓展的高级副总裁来出任决策层的组长。



2) 管理层

管理层一般由来自信息安全部门或专门的数据安全管理部门人员组成，负责数据安全治理体系的建设、培训和运营维护工作。在数据安全治理启动建设的早期，管理层需要牵头对企业或组织现有的数据资产进行梳理，向业务运营、数据分析等数据使用部门充分了解与数据安全有关的业务需求，分析评估安全威胁和潜在风险，并详尽调研各政策、法律、标准、规范中的数据安全合规要求，然后根据企业或组织的风险承受能力和财务预算，规划和起草适宜自身的数据安全操作规程等制度文档。在制度规范得到决策层认可后，管理层要组织相关培训，以推动制度在企业或组织内的推广和落地实施。管理层还要承担起维护数据安全制度持续运转的保障工作，并及时做出更新、调整和优化，以更好适应和支撑企业或组织的业务发展。

3) 支持层

支持层一般由来自业务部门和运维部门的人员组成。这些人员是数据的使用者、管理者、维护者、分发者，同时也是数据安全策略、规范和流程的重要执行者和管理对象。在数据安全治理启动建设的早期，支持层负责协助管理层详细了解并深入理解企业或组织在业务开展过程中的各种数据安全风险；对管理层提出的数据安全操作规程等制度和方案的可行性和易用性，进行细致的分析和评估，并将结果反馈给决策层，以支撑后者做出明智、正确的决策。数据安全制度正式发布实施后，支持层要在其日常例行工作中严格遵守数据安全操作规程，并积极发现和报告制度规范中的漏洞和潜在风险，促进管理层及时响应，尽快对数据安全的制度和措施做出更新、调整和优化。

4) 监督层

监督层一般由企业或组织内的审计部门承担，负责定期对数据安全方面的制度、策略、规范等的贯彻落实和执行遵守情况进行考查与审核，并将结果汇报给决策层。监督层的关键是要具有独立性，确保其审计核查工作不会受到来自其他三层，特别是管理层和支持层的相关利益或动机的影响和干扰，从而保证企业或组织能够及时发觉其数据安全制度在落地执行层面的问题和风险。

提供重要互联网平台服务、用户数量巨大、业务类型复杂的个人信息处理者，还需成立主要由外部成员组成的独立机构对个人信息保护情况进行监督。

3.3.2. 定岗定员

数据安全治理团队的职能架构确定后，如何制定出高质可行的操作规程和管理制度并实现这些制度规范的高效运作和部门职责的有效达成，就成为企业或组织要面对和解决的首要问题。定岗定员、专业化分工是解决问题、实现目标的基本方法。

定岗是指对数据安全职能架构中各个层次内的职责进行更细致的分工，将一系列相互间关联紧密的工作任务集合设定为一个岗位。定员则是指明确固定某一个或一组人承担某个岗位的任务和职责。定岗的本质是分工，定员的核心是明确职责，二者被共同用于在分工基础上为企业或组织实现降低成本、提高效率的终极目标。

在数据安全治理实践中，定岗定员应特别注意遵循以下原则：



1) 不应出现人员跨层交叉

从上节对数据安全治理团队职能架构的介绍不难发现，决策层、管理层、支持层、监督层之间一方面存在着密切的配合与合作，另一方面也同时存在着相互制约与平衡。这些不同层间的相互制约与平衡是必不可少的，只有如此，才能有效降低人的主观因素中潜藏的负面影响。一旦有人在不同层内同时扮演多个角色，难免会出现因考虑个人相关利益而降低标准、通融迁就的情况。例如，如果一个人既在管理层内负责起草数据安全的技术实施方案，又在决策层内负责对方案的可行性进行审批，则难免会出现方案敷衍或滥用数据安全建设预算的问题；而如果有人同时出现在监督层和其他层，则监督层对后者的审计核查作用必然会大打折扣，甚至形同虚设。

2) 仅授予相关角色或人员完成岗位职责所需的最小权限

3) 对于重要岗位或操作规程，应权限分离或安排多人互相监督，并对重要岗位人员开展背景调查

对重要数据的修改应安排不同人员分别进行操作、审批和复核，以防范内部人员对数据的违规篡改；例如，加密数据库的密码不应告知操作系统管理员，以防范后者用拖库的方法窃取数据等。

4) 信任要基于岗位职能和人员角色，而非人员身份

要培训员工树立和加强安全意识：在数据安全治理的职能框架下，默认不信任企业或组织内外部的任何人或设备，除非按照定岗定员的分工结果被授予了相应权限，否则即使是企业或组织内的上级领导，也不能违反数据安全制度规定的操作规程，越权访问或使用敏感数据。

3.4. 数据安全管理体系

3.4.1. 管理体系制度架构

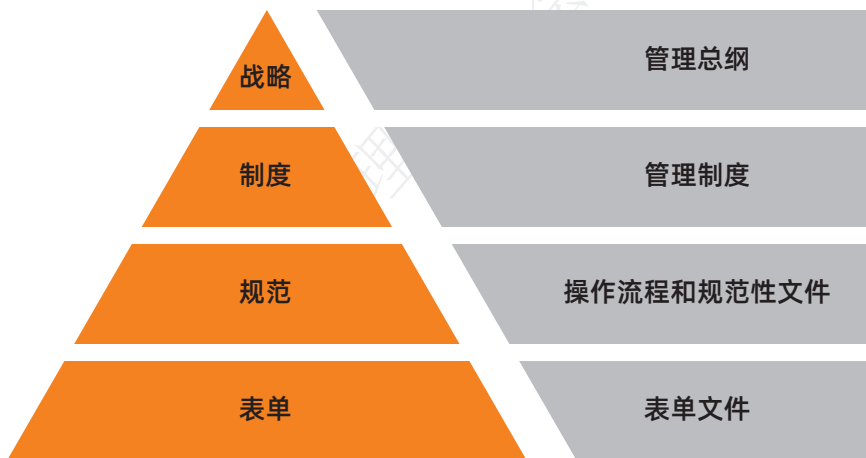


图6 管理体系架构图

数据安全管理体系可分为四层架构，每一层作为上一层的支撑。

第一层是管理总纲，是组织数据安全治理的战略导向。应明确组织数据安全治理的目标重点，如“以分类分级为基准，以权限控制为措施，管理与技术并重”的数据安全治理方针。

第二层是管理制度，是数据安全治理体系建设导向。应建立数据安全管理制度、组织人员与岗位职责、应急响应、监测预警、合规评估、检查评价、教育培训等制度。



第三层是操作流程和规范性文件，是组织安全规范导向。作为制度要求下指导数据安全策略落地的指南。应建立数据分类分级操作指南、技术防护操作规范、数据安全审计规范等指导性文件。

第四层是表单文件，是组织安全执行导向。作为数据安全落地运营过程中产生的执行文件。应建立数据资产管理台账清单、数据使用申请审批表、安全审计记录表、账号权限配置记录表等。

3.4.2. 管理体系建设思路

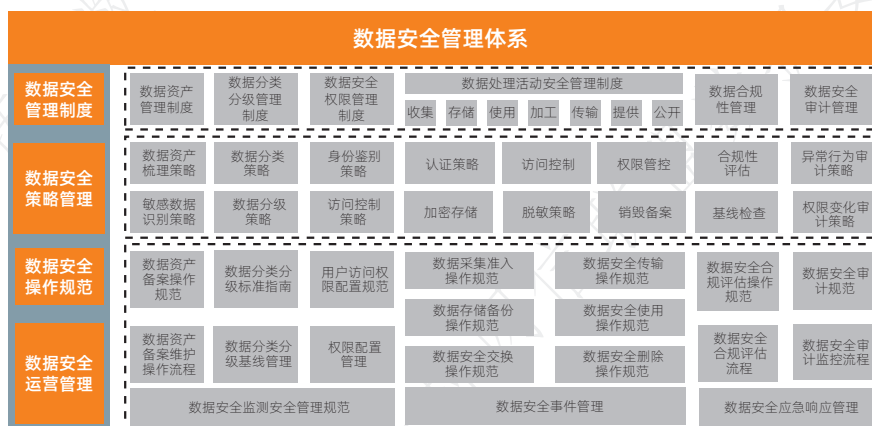


图 7 数据安全管理体系建设思路

数据安全管理体系并不是摒弃企业或组织内部建设以网络为中心的安全管理规范，而是在此基础上融合数据安全要求，形成全面的数据安全管理规范，是数据安全技术体系的指引和基础。

3.5. 数据安全技术体系

3.5.1. 数据安全技术演进趋势

面向数据安全治理涉及的复杂的监管与合规需求，数据安全防护技术也需要进行体系化的应对，数据防护技术在传统的注入或漏洞攻击防护、认证与访问控制、加密、审计、去标识化技术的基础上，也在进行着持续的技术演进，同时，这些技术也在快速实现国产化落地，拥抱国产化趋势。对演进趋势总结归纳为以下几点：

1) 场景化主动防控手段纳入数据安全技术体系

鉴于数据资产与业务的强相关性，需要根据业务的流程和特点，面向数据资产价值及防护诉求不同，针对数据流转的风险暴露面，确定合适的数据安全防护技术手段，突破传统的基于网元的一刀切式防护手段，构建动态、按需的体系化技术防护框架。

2) 使用密码技术保护数据安全使用成为热点

在数据处理活动中，数据收集、存储、使用、加工、传输、提供、公开等每一个环节，其目标



不外乎“使用”这些数据，或“加工”这些数据，创造新的价值，从而释放数据红利，为社会发展提供强大的推动力。从数据安全的角度来看，保护数据完整生命周期的安全，例如数据安全收集、安全传输、安全存储，其根本目的是确保数据在“使用”时是真实可信的，对这些数据进行“加工”所获得的价值是有效的。当前，在保证数据使用安全方面，机密计算、隐私计算等基于密码技术保护数据交易、共享安全的落地化实践场景已经展开。

3) 数据安全防护与 AI 技术融合愈加成熟商用

面向数据资产识别、数据分类分级、数据流转的风险分析等安全防护手段，单纯依赖传统的基于人工设定规则/策略的防护技术，在处理效率、准确性、全面性等方面已经无法满足日趋复杂的数据安全防护需求，通过引入自然语言处理（NLP）、用户异常行为分析（UEBA）、知识图谱（KG）等人工智能技术，与传统的安全技术进行有效融合并实用化，已较好的推进了智能化数据分类分级、智能化风险分析的进展，极大的提升了数据安全防护的水平。

4) 单点技术向平台化融合技术发展趋势愈加强烈

围绕数据全生命周期，在周期各阶段节点的单点防护技术手段已日益健全，但基于安全木桶效应，数据流转需要体系化防护，建立集中化、联动化的安全防护平台，将这些单点技术进行有效串联，实现面向数据安全风险的动态、纵深防御已成业内共识，零信任平台，隐私计算平台、安全监测平台、安全运营平台等平台化技术得到蓬勃发展，在数据收集、存储、使用、加工、传输、提供、公开等应用场景中开展了广泛实践应用，有效的支撑了数据安全治理实践。

5) 信息技术应用创新进程加速

伴随《网络安全法》、《数据安全法》等法律法规的出台和实施，数据安全技术发展又迎来了一个新的里程碑，翻开了基于信息技术应用创新的时代新篇章，推动数据安全领域的技术和产品创新发展，通过数据安全技术和芯片、操作系统、云平台、AI 算法及国产密码算法等自主可控技术充分融合，形成安全可信的防护体系，筑牢数据安全基石，助力我国数字化转型。

3.5.2. 技术体系架构

数据安全技术体系并非单一产品或平台的构建，而是覆盖数据全生命周期，结合企业或组织自身使用场景的体系建设。要依照企业或组织数据安全建设的方针总则，围绕数据处理活动各场景的安全要求，借鉴 I（识别）P（防护）D（监测）R（响应）模型，建立与制度流程相配套的技术和工具，并将这些技术和工具形成平台化应用，发挥技术合力作用。通过持续对数据生命周期内各使用场景进行风险监测，评估企业或组织现有数据安全控制措施的有效性及其薄弱环节，对有风险的风险场景及时进行数据安全整改，优化数据安全相关制度流程，进而持续的提升企业或组织的数据安全防护能力。技术体系整体技术架构如下图所示：

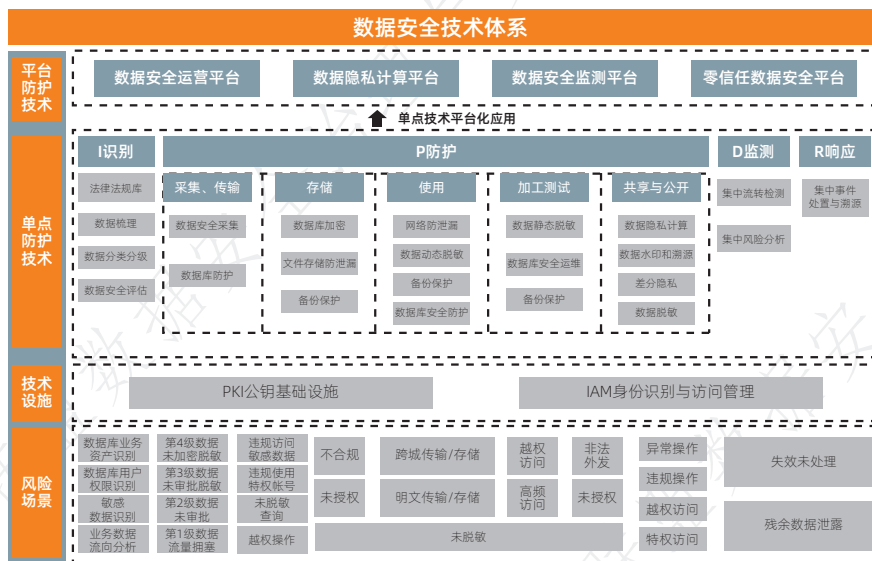


图 8 数据安全技术体系

3.5.3. 识别技术（I）

3.5.3.1. 法律法规知识库

法律法规是指引组织数据安全建设的重要依据，应对解读的安全标准和规范的合规项进行落实，形成合规库，收录并拆解各类与数据安全相关的规范、标准，为数据安全治理提供参考及评估标准，并根据合规库中的合规项制定各类安全策略规则。

传统的法律法规库多为文档形式，难以融入到日常的安全运营工作中去，可通过技术手段建立法律法规库系统，系统提供规范标准的录入、修订、废止、检索功能，通过系统可快速检索组织应执行的合规项，及根据各合规项执行的安全策略和具体实施的安全措施。

3.5.3.2. 数据资产梳理

数据资产梳理是指对目标环境中的数据资产进行全面清查、摸排，通过了解数据资产类型、数据资产分布、数据资产权限、数据资产使用等信息，构建数据资产目录的过程。在数据安全治理实践中，尤其关注针对敏感数据资产的梳理，这是数据安全体系建设及数据资产管理中的一项基础性工作。对数据资产进行及时准确的梳理以掌握其中敏感资产的分布、数量、权限及使用状况，是进行后续数据安全治理工作的基础与先导。

对数据资产进行梳理，主要包括两种技术：一种是静态梳理技术，通常采用对 IP 地址段和端口范围进行扫描的主动嗅探方式，发现数据库资产。然后再应用数据库字段识别技术、数据样本特征识别等技术和元数据接口对接、数据字典导入等方式，形成数据资产清单；另一种是动态梳理技术，通常采用对网络流量进行协议分析的被动监测方式，用于形成数据访问关系清单。二者共同为后续的数据安全治理建设提供必要的基础信息。

有关数据资产梳理的更多技术详情，可参考本白皮书附录 C.2.2。



3.5.3.3. 数据分类分级

为了更好地在保持数据使用的便捷性和实现数据保护的安全性二者之间进行平衡，促进数据安全能力建设降本增效，通常会对企业或组织的数据（资产）进行分类分级，以便对数据采用精细化的安全管控手段，避免一刀切。

为应对企业或组织海量数据的分类分级，可采用谓词切分与语义识别技术、知识库与匹配技术和机器建模与匹配等技术，辅助人工快速形成数据分类分级清单。

有关数据分类分级的更多技术详情，可参考本白皮书附录 C.1.1。

1) 数据分类方法

对数据的分类因为要密切贴合企业或组织的自身业务特点，一般没有特别通用的规则和方法，通常会根据数据的用途、内容、业务领域等因素进行，而且可能需要随着业务的变化而动态变化。

2) 数据分级方法

数据分级一般会根据数据的价值、敏感程度、被泄露后的影响和后果等因素进行定义，一旦确定就很少再发生变化。

3) 数据分类分级落地

在通用规范层面，全国信息安全标准化技术委员会秘书处发布了《网络安全标准实践指南——网络数据分类分级指引》（TC260-PG-2-20212A），对全领域的网络数据分类分级给出了指引，将数据类别分为个人信息、公共数据、传播数据、行业领域数据等，将数据级别由低到高分为一般数据、重要数据、核心数据。针对个人信息，明确将个人信息细划为 16 个子类，清晰定义了每个子类的标识信息项，并给出了针对敏感个人信息的定级标准。在行业领域规范层面，金融、电信等重点行业分别颁布了数据分类分级相关的行业标准规范（《金融数据安全 数据安全分级指南》（JR/T0197—2020）、《基础电信企业数据分类分级方法》（YD/T3813—2020）等），定义了行业领域内数据分类分级方法与实际规则参考或示例，为行业内企业或组织界定核心数据、重要数据、一般数据开展分类分级安全治理落地工作提供了有效的指引和依据。

针对通用规范和行业规范中有标准可参照的数据，通过建立元数据（或元数据集）与标准规范中分类分级标识信息项的映射关系，结合企业或组织实际情况，可以较为快速的落实数据分类分级，并基于分类分级结果清单开展针对不同类别与级别数据的安全防护措施实践；对于尚未颁布行业数据分类分级标准规范的企业或组织，在对数据进行分类分级的过程中，由于类别和级数增多会导致管理和维护成本上升，因此，建议企业或组织在满足国家和行业监管合规要求的前提下，选用尽量简单的分类分级规则，以保障以分类分级为基础的数据安全治理的可实施性，后续再伴随治理工作机制的完善和成熟，结合行业监管要求的规范化，逐步进行深入和细化安全管控。

3.5.3.4. 数据安全评估

数据安全风险评估的结果通常包括五个基础维度：

● 资产识别

通过数据处理活动要素识别，对数据资产进行识别和业务分类，根据该类数据泄露或篡改后影响的范围定义数据泄露或篡改后影响程度，再根据业务分类和数据泄露或篡改后的影响程度进行数



据资产定级，明确数据资产价值。

● 脆弱性识别

脆弱性一方面源于数据本身及数据载体可能被威胁利用的薄弱环节，另一方面源于法律法规对数据资产利用和防护的要求。数据安全风险脆弱性识别包括脆弱性主体识别、脆弱性类型识别和脆弱性问题描述。

● 威胁识别

威胁识别是对数据资产在数据处理活动中，可能发生的对保密性、完整性、可用性造成的危害识别。通过威胁识别掌握脆弱性被利用的可能性及对数据资产的影响程度，判断数据安全风险的大小。

● 已有安全措施识别

已有安全措施的识别是针对企业或组织已采取安全措施的有效性进行分析确认，已有安全措施可作用于数据本身，如对数据加密后传输，也可以作用于数据载体，如对数据传输链路进行加密。

有效的安全措施实施，可降低企业或组织在数据安全管理和技术两个方面的脆弱性，缓解数据资产面对的威胁，整体降低数据安全风险评级。

● 残余风险分析

结合脆弱性、威胁及已有安全措施，针对实施安全措施后满足脆弱性需求和增强抵御威胁的能力进行分析。如针对某脆弱性问题，通过安全设备实施了相应的防护策略，则数据安全整体风险降低。

● 安全风险评估报告

基于以上五个基础维度的评估结果，分析数据处理活动过程中的脆弱性问题，及该脆弱性面临的威胁，并相应的检查已有安全措施，判断残余风险，并根据数据资产价值和残余风险判定数据安全风险值，最终形成数据安全风险评估报告。

数据安全评估主要包括数据处理活动评估和数据出境安全评估两个场景。

1) 数据处理活动评估

数据安全风险评估是指借鉴信息安全风险评估的基本原理和步骤，基于组织的数据战略，从数据全生命周期安全出发，通过对组织目标环境中数据资产的重要程度、数据载体的安全状况、敏感数据的访问状况、数据安全相关管理制度和基础设施的安全性等多方面的信息进行收集、统计和分析，从不同维度评估风险状况，并最终计算得出综合评估结果的过程。

数据安全风险评估主要包含两个步骤：首先，要明确待评估的数据资产范围；然后，再基于数据安全相关的法律法规和标准规范设定具体的评估指标。评估范畴应覆盖数据的收集、存储、使用、加工、传输、提供、公开等数据处理活动的各个环节：

● 数据收集

评估外部数据源的合法性；评估数据收集的数据源、频度、渠道、方式、数据范围和类型；针对个人信息要评估是否存在过度收集等问题。

● 数据存储

评估企业或组织的数据存储管理安全规范、配置规则和访问控制手段；评估是否对个人信息和重要数据采取了加密存储机制；评估数据备份和恢复的管理制度、技术建设及其执行情况。

● 数据使用和加工



评估企业或组织的数据使用及加工制度及相关安全技术措施的建设情况；评估针对不同应用场景的脱敏处理流程及其涉及的部门和人员的职责分工、数据导入导出（特别是大量或敏感数据导出）的授权审批规范，评估加工后的数据是否引入了新的安全风险等。

● 数据传输

评估是否对数据尤其是敏感数据的传输进行了加密，是否对认证和票据相关数据使用了数字签名等。

● 数据提供和公开

评估数据共享接口的安全策略和保护机制；评估合作协议中关于数据的使用目的、供应方式、保密约定、数据安全责任等事宜是否明确；评估数据公开的内容、适用范围、发布者与使用者的权利和义务等事宜是否明确。

2) 数据出境安全评估

《数据出境安全评估办法（征求意见稿）》针对数据跨境转移行为提出了明确要求。数据出境安全评估重点关注合法性、数据识别、数据处理、数据环境等全链条评估。

数据出境安全评估的内容包括：

● 出境数据梳理及分类分级识别

企业或组织需识别出境的数据种类和级别，并针对不同类型的数据在出境前采取差异化保护措施，并根据最小化原则，确认出境的数据的充分性和相关性，确保出境数据的内容仅限于向数据接收方转移和处理的必要数据。

● 识别内外部合规要点

企业或组织需从数据接收方国家层面的整体法律环境，数据接收方的组织措施及技术措施三个方面开展评估。主要评估要点如下：

- (1) 评估数据接收方的法律政策环境；
- (2) 评估数据接收方的主体资格及行业信誉；
- (3) 评估是否和数据接收方明确约定数据出境管理的权利、责任、义务；
- (4) 评估是否和数据接收方明确了出境数据的范围、使用期限以及再转移限制；
- (5) 评估在数据出境后对出境数据是否具有有效的监管控制能力；
- (6) 评估数据接收方是否提供相应的数据安全管理和技术保障措施；
- (7) 评估数据接收方是否提供方法和渠道及时响应和配合数据主体的诉求。

● 补充安全措施

基于数据出境风险评估结论，有针对性的采取补充安全措施，并在需要时与数据接收方合作，确保数据出境后可以得到企业或组织所规定的的数据保护要求。补充的安全措施包括合同措施、组织措施和技术措施。

● 审批评估结果

设立评审组织对评估结果进行审批，评估结果将决定企业数据是否出境。

● 持续监测数据接收方的保护水平

对数据接收方的数据安全组织措施和技术措施进行持续监测，确保数据接收方能够保持对出境



数据的安全保护水平。

3.5.4. 防护技术（P）

3.5.4.1. 身份安全基础设施

身份安全基础设施可以为收集、存储、使用、加工、传输、提供、公开等数据处理活动的访问控制提供基础的数据来源，以便对用户身份、各类设备和应用系统进行身份管理和权限管理，基础的身份安全基础设施包括 IAM 身份识别与访问管理和 PKI 基础公钥设施。

1) IAM 身份识别与访问管理

IAM 是一套全面建立和维护数字身份，提供有效安全的 IT 资源访问的业务流程和管理手段，实现组织信息资产统一的身份认证、授权和身份数据集中管理与审计。身份和访问管理是一套业务处理流程，也是一个用于创建、维护和使用数字身份的支持基础结构。

在数据安全治理所有过程中都需要安全的身份标识，在数据治理的交互过程中识别对端的身份，初始自身身份，IAM 是数据安全治理不可或缺的一个中心环节。

IAM 能够起到的作用主要包括：

- 在数据的收集阶段，保证数据来源的真实性
- 在数据传输阶段保证数据传递路径中接收者和传递者身份的可靠
- 在数据存储过程中保证对数据进行存取的实体身份是可靠的
- 在数据权属转移、数据共享、数据交易中保证数据业务层安全
- 在存在多个异构的数据应用域的情况下，IAM 保证的访问者身份的一致性。

有关 IAM 身份识别与访问管理的更多技术详情，可参考本白皮书附录 C.2.3。

2) PKI 基础公钥设施

公钥基础设施 PKI (Public Key Infrastructure)，是一种遵循既定标准的密钥管理平台，它能够为用户提供生命周期防护涉及的身份认证、数据传输、数据存储、数据访问、数据应用提供加密和数字签名等密码服务及所必需的密钥和证书管理体系。简单来说，PKI 就是利用公钥理论和技术建立的提供安全服务的基础设施。PKI 技术是信息安全技术的核心，也是电子商务的关键和基础技术。

PKI 公钥基础设施包括密钥管理服务、数字证书服务、密码计算服务、时间戳服务、密码设备管理和硬件安全模块服务。

● 密钥管理服务

密钥是指控制加解密、签名验签等密码变换运算的符号序列。密钥的分类：按照密码算法类型分为对称密钥和非对称密钥；按照使用场景分为根密钥、系统密钥、用户密钥、会话密钥、密钥加密密钥和数据加密密钥等多种类型。对不同类型的密钥按照密钥生命周期进行管理的过程。具体功能包括：密钥生命周期管理、密钥体系分层、密钥传输安全和密钥服务功能。

● 数字证书服务

为各类实体提供数字证书，并进行数字证书的生命周期管理，在认证或签名过程中确保实体身份的可信。具体功能包括：证书生命周期管理、注册审核服务、证书签发存储、证书撤销列表维护、



证书在线查询服务和公钥证书下载服务。

● 密码计算服务

为密码管理内部模块和外部应用提供的基础密码计算服务。具体功能包括：密码计算服务接口、加密服务、签名服务。

● 时间戳服务

基于国家权威时间源和密码等技术，为信息系统提供精确可信的时间戳，为业务处理的抗抵赖性和可审计性提供有效支持。具体功能包括：时间同步和校准、时间戳签发、时间戳验证。

● 硬件安全模块服务

利用计算机硬件设备，保护和管理强认证系统所用的密钥，并提供相关密码操作。具体功能包括：关键密钥存储、密码管理、防篡改。

3.5.4.2. 数据收集、传输安全防护

数据收集应满足数据收集相关的合法性要求，满足数据收集的必要性原则和最小化原则，确保用户数据在收集前已征得用户同意，需要单独征得用户同意的业务确已征得用户同意，应当重新取得用户同意的情况确已重新取得用户同意，并按规定向用户公示了其个人信息收集的目的，防止出现超范围收集用户数据、超权限使用用户数据、违法违规交易用户数据及差异化用户决策等违法行为。同时，数据收集源头的安全是数据价值利用的先决条件，在数据收集、传输时，应对数据提供方的身份进行有效验证，并通过安全防护技术防止数据泄露及数据篡改等数据安全事件的发生。在数据安全收集方面，应对数据收集设备进行持续的身份认证和数据传输安全防护。在数据中心防护方面，应在网络安全防护的基础上，对数据库操作加强安全防护，构建基于数据安全的纵深防御体系，针对业务数据进行细粒度访问控制，防止数据的越权/非授权访问和利用数据库漏洞发起的攻击行为。

1) 数据安全采集

在数据收集处理活动中，可采用业务数据关联分析、流量分析等手段，对数据收集的合法性进行监测。梳理系统业务说明中用户数据收集的告知内容，识别系统数据流量中的敏感数据内容，分析是否存在超范围收集用户数据的情况，并结合实际业务需要，评估数据收集是否满足最小、必要原则；分析业务系统的数据结构，通过数据关联分析，识别用户未标记同意但却产生相关用户数据的情况，评估是否存在未征得用户同意收集用户数据的情况；分析业务系统的数据结构，通过数据关联分析，识别用户未标记同意但系统却使用用户数据进行自动化决策的情况。

数据安全采集是集密码技术与网络隔离技术于一身的物联网应用层边界防护技术，可对数据采集设备的通信数据进行内外网物理隔离、专用安全协议包解析和过滤、通道加密、接入认证等，实现由端至云覆盖应用、网络、感知各层的全面安全防护。

数据安全采集对数据采集设备进行发现和识别，对前端设备进行准入与访问控制，匹配信令白名单列表，阻断未登记在信令白名单列表中的控制信令传输到数据采集设备，持续检测数据采集设备运行状态，计算周期内设备使用上下行流量的平均值，发现设备上下行流量波动超过阈值的情况，并进行策略处置。

有关数据安全采集的更多技术详情，可参考本白皮书附录 C.2.4。





2) 数据库安全防护

对于攻击者通过 WEB 应用，使用“SQL 注入”攻击的方法从后台数据库服务器尝试进行“刷库”等攻击行为，应用数据库安全防护技术，针对应用系统访问数据库根据数据分类分级规则进行访问控制和防止漏洞利用等安全防护，在数据库系统账户及权限管理的基础上提供二次防护，一方面规范数据库特权用户的操作行为，解决数据库用户权限过度分配、权限未及时收回等问题。另一方面通过 SQL 注入防护可以从根本上帮助管理员防止 SQL 注入的发生。此外，数据库防火墙提供的虚拟补丁功能可实现数据库不打补丁也能防止补丁攻击的需求。

有关数据库安全防护的更多技术详情，可参考本白皮书附录 C.2.8。

3.5.4.3. 数据存储安全防护

存储作为 IT 数据基础设施堆栈的底座，对保障数据安全可靠尤为重要。应用数据库加密技术保障结构化数据存储安全、数据 DLP 技术保障非结构化数据安全、以及数据保护技术保障数据的可用性，建立一套完整的数据存储安全机制，对数据的存储安全具有重要意义。

1) 数据库加密

数据加密技术是指将一个蕴含信息的明文数据经过加密密钥（加密密钥）及加密函数转换，变成无意义的密文数据；以后需要获取数据内蕴含的信息时，要先将该密文数据经过解密函数、解密密钥（解密密钥）处理，恢复成原来的明文数据，然后才能对数据及其内的信息加以使用。加密密钥和解密密钥都可泛称为密钥，而且在对称加密算法中，加密密钥和解密密钥完全相同。关于密钥的生成、保存、传递和使用等机制是数据加密技术必须具备的关键组成部分，一般统称为密钥管理。数据加密技术是保护数据机密性和完整性的重要手段，是网络和数据安全技术领域的基石。

让数据使用者参与甚至负责实现数据的加密、解密或者密钥管理的缺点是显而易见的，不仅会显著增加数据使用者负担，严重损害数据存储系统的易用性、稳定性，而且容易因数据使用者对数据加密缺乏经验而引入漏洞，对数据造成损害，也可能因密钥被遗失或盗取而造成数据无法被解密或者泄露。为消除这些缺点，数据加密领域发明了透明加密技术。

所谓透明加密，是指数据的加密、解密和密钥管理等过程对数据使用者是完全透明的。为此，透明加密技术需要在数据使用者和存储系统之间置入中间件。一方面，透明加密中间件向数据使用者提供与存储系统完全相同的操作接口，并预先生成好密钥备用。另一方面，当数据使用者写入数据时，中间件要先将接收到的明文数据用密钥加密，再将密文数据保存到存储系统中；当数据使用者读取数据时，中间件要先从存储系统中读出密文数据，用正确的密钥解密后，再将明文数据提供给数据使用者。

有关数据加密的更多技术详情，可参考本白皮书附录 C.2.6。

2) 文件存储防泄露

存储 DLP 对驻留在电脑终端、服务器、文件共享和数据库上的敏感信息进行识别，对文件共享、邮件、Web、应用程序等传输的数据进行监控，在数据进行操作之前对其进行管控，如：打开、另存、复制、打印、拷贝、刻录、在线传输（如：QQ、邮件客户端、蓝牙）、文件共享等行为进行管控，





并根据安全策略产生相关的动作（如阻止、审计、提示、加密等），同时生成预警日志和审计日志，使企业或组织能够最大化地加强对关键数据的管控。

有关文件存储防泄露的更多技术详情，可参考本白皮书附录 C.2.10。

3) 备份保护

数据备份保护是指为防止因物理故障、操作失误或遭受攻击破坏导致数据不可用，而对生产位置的全部或部分数据集合制作一或多份副本并以某种方式保存到其他位置，以备在需要时重新加以恢复和利用的一个过程。根据数据的分类分级，可以采用不同的数据备份策略，确保数据的可用性和完整性。当前，数据备份保护热点包括：大数据的备份恢复、混合云的备份恢复、副本数据利用、灾备服务化、应用的容灾、多种保护能力的平台化等。

数据备份是容灾的基础，即当灾难发生时，能够保证尽量少丢失数据情况下尽快恢复信息系统运行，保证业务的连续服务能力。

有关数据备份保护的更多技术详情，可参考本白皮书附录 C.2.11。

3.5.4.4. 数据使用安全防护

数据在使用中发挥价值，但数据在使用过程中的流动性特征，极易导致数据泄露事件的发生。因此，在数据使用阶段，应从数据内容识别和数据权限细粒度管控两个方面实施数据安全防护措施，一方面通过网络 DLP，在网络传输过程中精确识别出各类敏感数据，及时切断数据传输通道，避免数据外泄。另一方面，在传统的用户身份认证技术的基础上，通过数据动态脱敏技术，根据用户身份，对用户查询的数据进行动态脱敏，在业务应用无需配合改造的情况下，实现了基于“身份-数据”的访问权限控制。同时，也可应用数据库安全防护能力对应用访问数据库的权限进行细粒度管控，加强数据库访问的安全防护。

1) 网络防泄露

网络 DLP 提供网络流量的可见性并可以对流量进行控制，具备全流量检测、审计、简单加密算法解密、流量内容检查、自定义协议功能。检查网络传输的流量，解析包括电子邮件（SMTP）、web（HTTP）、文件传输（FTP）、文件共享（SMB）和其他自定义的 TCP 会话流量协议，识别传输的敏感数据，并执行强制数据访问控制策略，阻断非授权的敏感信息传输，并及时告警。

有关网络存储防泄露的更多技术详情，可参考本白皮书附录 C.2.10。

2) 数据动态脱敏

动态数据脱敏一般用在生产环境中。在数据访问者读取数据的过程中，部署于访问通道上的动态脱敏系统会监测和拦截数据访问请求，并根据请求中数据使用者的角色、权限、待访问数据的类别级别等信息，按照脱敏策略和规则实时对数据中的敏感信息进行脱敏处理，然后将脱敏后的数据提供给数据访问者。

动态脱敏与数据访问控制，作为两种保障数据安全的机制和技术手段，通常会在数据安全的产品和系统实现中被联合使用。

有关静态脱敏的更多技术详情，可参考本白皮书附录 C.2.5。



3.5.4.5. 数据加工安全防护

在数据加工阶段，运维人员需要协助业务部门将大量的生产数据频繁的导出到加工、测试环节，采用手工导出的方式除效率较低外，还难以确保数据脱敏的有效性。同时，传统的如堡垒机等安全运维管控技术无法对运维人员的特权账户进行细化到字段级别的访问控制，这都给数据加工、测试阶段的数据安全造成了巨大的数据泄露风险。

1) 数据静态脱敏

静态脱敏一般用于数据在非生产环境（如开发、测试、培训、外包和数据分析等）被使用时，对数据中含有的敏感信息进行保护。静态脱敏通常会涉及对较大数量的数据进行批量化的处理：静态脱敏系统首先从数据的原始存储环境（通常为生产环境）读入含有敏感信息的数据，然后在非持久化存储条件（系统内存）下按照脱敏策略、规则和算法对数据进行变形等脱敏处理，再将经过处理后的脱敏数据存储到新的目标存储环境中。

存储到新的目标环境中的脱敏数据固然已被消除信息敏感性，可以安全地在开发、测试、外包和数据分析等场景中放心使用，但为了能够得到具有业务应用价值的最终结果，脱敏数据应尽可能在统计意义上保留与生产环境中原始数据一致的数据特征。为此，在静态脱敏场景下，通常需要关注脱敏数据的有效性、真实性、高效性、可重现性、关联性与可配置性等六大基本原则。

有关静态脱敏的更多技术详情，可参考本白皮书附录 C.2.5。

2) 数据库安全运维

数据访问控制通常被用在数据生命周期中的数据使用环节，主要包括两种应用场景：一种是内部业务人员通过业务系统访问数据，另一种是运维或系统管理人员通过运维客户端访问数据存储系统。

内部业务人员通过业务系统访问数据存在的数据安全风险主要有：

- 业务系统可能因存在漏洞或后门而被攻击者控制和恶意利用；
- 业务人员可能出于自身牟利目的而盗用、滥用业务系统访问数据；
- 业务人员（尤其是有高权限者）可能因疏忽而执行错误操作。

运维或系统管理人员通过运维客户端访问数据存储系统存在的数据安全风险有：

- 运维工作中的操作、流程等违反外部合规要求或内部制度规范；
- 管理人员因疏忽错误执行对数据具有破坏性的高位操作；
- 出于个人牟利或发泄报复等原因而对数据进行恶意删除或篡改。

为防范和应对以上数据安全风险，数据访问控制系统通常具备以下主要功能：

- 对用户的身份、权限和访问控制策略的配置管理功能
- 对访问请求进行拦截和解析，认证发起请求的用户身份，识别请求的操作目标；
- 根据身份认证和操作目标的解析结果，判断数据访问请求是否访问控制策略；
- 按照访问控制策略，对非法的数据访问请求进行阻断和告警。

下图以数据库安全运维系统为例，介绍了数据访问控制的工作原理和基本流程。

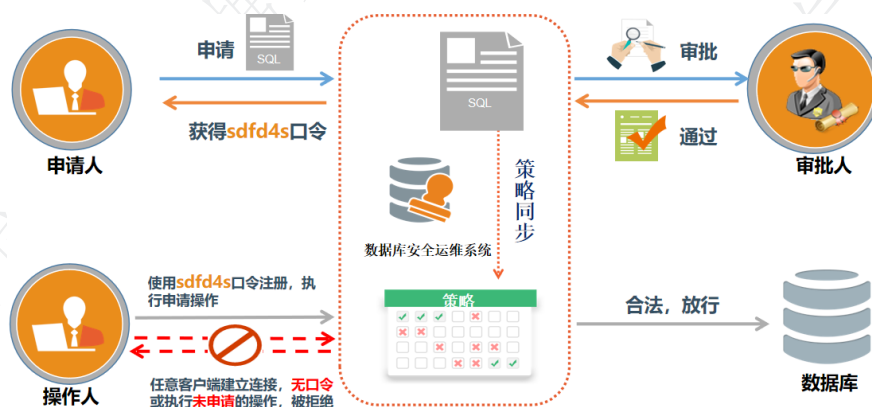


图9 数据访问控制的工作原理示意图

有关数据访问控制的更多技术详情，可参考本白皮书附录 C.2.7。

3.5.4.6. 数据提供与公开安全防护

在对外提供数据共享与公开时，为保障数据安全，可综合运用静态脱敏和动态脱敏能力提供脱敏后的数据给数据使用方，但数据的二次传播特性决定了数据一旦共享或公开给其他使用方或处理方，数据所有者可能全面失去数据的管理权和监督权，数据水印技术能够在一定程度上对二次传播的数据进行溯源，却还不能从根本上阻断数据二次传播。当前更为安全的数据共享与公开方式是通过隐私计算技术，按照数据使用方的数据使用需求，直接向使用方提供数据结果。隐私计算技术可在数据不出本地的情况下，和其他组织共同挖掘数据的价值，既保证了组织数据的安全性，又实现了数据价值的最大化。同时，在数据对外共享与公开、数据出境时，可应用网络 DLP 实时监测非授权敏感数据外发的风险，防止数据泄露。

1) 数据隐私计算

数据有着不同于其他生产要素的独特特征：数据的复制成本极低，数据可以被无限制地被复制；数据不会因被使用而耗损或灭失；它可以被重复使用、被多方同时使用；使用数据的过程通常会产生新的数据，因此数据是取之不尽，用之不竭的，只会越用越多。数据的特征使得其有巨大的价值潜力，但同时也带来极大的使用隐患，主要在于：（1）数据一旦被“看见”就会泄露具体信息，即可被复制，复制成本极低并可以被无限地复制；（2）数据一旦被泄露或复制，就无法限制其用途和用量，被滥用的风险极高。正是这些使用隐患影响了数据成为生产要素并实现大规模的流通交易。

要实现数据要素化的核心思想，在于建立“新数据观”：将明文数据分割为“具体信息”和“计算价值”（使用价值），对于数据的“具体信息”部分，由于其涉及敏感信息需要做好保护，而对于数据的“计算价值”部分则需有效利用才能充分释放数据价值。

数据安全计算技术是指让数据在充分保护隐私的前提下参与计算，这样的数据计算技术将有助于在提高数据处理安全性的同时，更好地平衡数据融合应用和安全合规的需求。其中，安全多方计算、联邦学习作为数据计算的核心技术，能在多个数据归属方在融合使用数据的过程中保证各自数据不被泄露，让数据有效地融合利用，助力多个数据归属方的数据有效实现使用价值。

在新数据观下，实现数据要素化的关键在于两点：一是通过多方安全计算等新数据技术实现数



据“可用不可见”，保护数据敏感信息同时也能充分利用数据的计算价值；二是通过多方安全计算等新数据技术规定数据具体的用途和用量。基于上述两点，流通于市场的不再是明文数据，而是数据的特定使用权。市场可对数据的特定使用权进行定性、定量以形成可控的供给，数据才能成为生产要素并实现大规模流通，同时也有效规避数据被滥用的风险，数据的责、权、利得到厘清。通过数据可用不可见和规定用途、用量的技术手段，可以保障数据在使用时不泄露其具体信息，从而能针对数据的特定使用权和受益权进行确权，为数据所有权立法和数据资产化提供有益的探索和实践。

有关数据隐私计算的更多技术详情，可参考本白皮书附录 C.1.4。

2) 数据水印和溯源

数据水印和溯源，是指通过事先向数据中嵌入类似于水印的特征而实现事后能够对数据的违规使用行为进行追溯和核查。这项技术主要用于在数据安全治理中的事后追责活动：通过对数据安全损害已经发生的事件中的特定行为成因进行探究，从导致结果的成因中找出权重最大或最初阶段的行为主体，再以规章为依据追究其相关责任。

相关数据安全事件既可以是数据泄露，也可以是数据被滥用或者盗版、贩卖等情况。数据溯源工作的最终效果能实现对滥用数据、泄露数据的行为形成威慑，对盗用数据者给予法律惩罚，还可以进一步达到对数据的安全保护责任形成定向委托的效果（即对接收数据者要保护好自己接收到的信息或数据形成约束）。因此，数据溯源常被视作面向数据使用、共享的一项基本安全治理工作。

有关数据水印和数据使用行为溯源的更多技术详情，可参考本白皮书附录 C.1.6。

3) 差分隐私

差分隐私是一种基于严格数学证明的隐私定义，可用于数据收集、发布、分析等场景中，为用户数据提供合理的隐私保护。作为一种被学界广泛认同的隐私黄金标准，差分隐私要求输出数据集上的统计结果对于任意单个记录的变化是不敏感的。其背后的原始直觉在于，如果统计结果是隐私的，改变一条记录应当不会对输出产生太大的影响。

作为时下最流行的隐私定义，差分隐私被广泛应用于数据库查询，数据挖掘，网络踪迹分析，智能计量系统等多个领域。

有关差分隐私的更多技术详情，可参考本白皮书附录 C.1.5。

3.5.5. 监测技术（D）

3.5.5.1. 集中流转监测

数据安全体系建设从来不会是一蹴而就的，构建针对业务访问的全链路全面审计体系，支撑数据安全策略持续优化。在各个安全防护系统运行过程中，需要结合数据分类分级结果，对业务应用系统和数据库系统的操作进行全面的审计，对安全规则、敏感数据、用户行为进行不断的策略调整，以便于更加贴合各项业务系统的使用情况。

构建全链路数据安全审计能力体系：全链路数据安全审计综合采用数据库审计技术和应用及 API 审计技术，贯穿数据全生命周期各个阶段，覆盖用户访问、系统运维、开发测试、数据对外提供、





数据出境等全部业务场景，通过可视化方式按数据分类分级全面分析、展示数据的分布、流动、命令执行成功/失败、风险事件等综合态势。

持续监测应用访问全生命周期风险：构建智能化应用及 API 审计能力，管理应用资产全生命周期，对用户访问敏感数据的行为进行建模，持续监测敏感数据访问行为，发现敏感数据泄露风险，为在系统上发生的风险行为提供追溯能力。

以数据分类分级为基础加强数据库风险分析能力：构建数据库审计与分析能力，在传统数据库审计能力的基础上，将审计结果与数据分类分级结果进行关联分析，避免关键性的高风险审计信息淹没在大量的一般性访问记录中，为实现数据分类分级管理提供实时、精准、明确的管理依据。

完善的审计机制是基于敏感数据、策略、数据流转基线等多个维度的集合体，对数据的生产流转，数据操作进行监控、审计、分析，及时发现异常数据流向、异常数据操作行为，并进行告警，输出报告。其为数据安全带来的价值主要体现在两个方面：

1) 事中告警

一旦发现可能导致数据外泄、受损的恶意行为，审计机制可以第一时间发出威胁告警，通知管理人员。管理人员在及时掌握情况后，可以针对性地阻止该威胁，从而降低或避免损失。为此，审计机制应具备告警能力（如通过邮件、短信等方式发出告警通知），并能够有效识别系统漏洞（例如针对 CVE 等漏洞库的漏洞特征进行检测）、注入攻击（如对 SQL 注入攻击进行监控）、口令破解、高危操作（如变更数据库的表结构等）等威胁和风险。

2) 事后溯源

发生数据安全事件后，可以通过审计机制记录的日志信息对该事件进行追踪溯源，确定事件的源头，还原事件的发生过程，分析事件造成的损失，进而对违规人员实现定责和追究，为调整防御策略提供非常必要的参考。为此，审计机制就需要具备丰富的检索能力，可以将全要素作为检索条件来检索其记录的日志信息。

有关集中流转监测的更多技术详情，可参考本白皮书附录 C.1.8。

3.5.5.2. 集中风险分析

在传统依赖于规则检查、专家经验和阈值设定的基础上，补充用户和实体行为分析技术（UEBA），从单维度检测到多维度检测，从单点检测到长周期分析，从基于规则分析到关联分析、行为建模、异常分析来发现更多更隐蔽的安全风险。

● 用户和实体行为分析技术（UEBA）

UEBA 是用户和实体行为分析技术（user and entity behavior analytics），UEBA 更多的是关注人异常行为，行为主体通常是企业内部的员工。UEBA 技术基于海量数据对内部用户的异常行为或内部威胁进行预测，直接以“人”的视角给出判定，抓住“坏人”、主动出击，在数据泄露之前进行阻止，并为安全分析人员提供可靠的依据。

有关集中风险分析的更多技术详情，可参考本白皮书附录 C.2.9。



3.5.6. 响应技术（R）

3.5.6.1. 集中事件处置及溯源

当前建立的基于 SIEM/SOC 的安全事件以人工防守为主的响应和处置方式，涉及到防守方案制定及变更，多角色、岗位人员协作，人员知识传递，多厂商、多类型安全设备联动，关键操作审批等工作，存在人员能力要求高、工作流程复杂、事件响应时间长、无法在短时间内阻断攻击等不足。

● 安全编排和自动化和响应（SOAR）

安全编排和自动化和响应（SOAR）通过联动多个系统和设备来调度不同的安全能力，将人、技术与制度相融合，建立手动与自动相结合的协同响应与处置机制，简化事件处置流程，加快事件响应速度，减少事件响应时间。

SOAR 可收集不同来源、不同类型的安全威胁数据，运用数据分析的方法进行安全事件分析与分类，根据定义的事件响应与处置流程制定自动化事件响应与处置流程。SOAR 为安全运营团队提供定制化的自动编排程序框架，加速安全事件的响应与处置，安全运营团队的日常事务性工作也可借助 SOAR 自动化完成。同时，还可以帮助运营团队在单一平台上响应和缓解威胁，节约事件响应和处置的宝贵时间。

有关集中事件处置及溯源的更多技术详情，可参考本白皮书附录 C.1.7。

3.5.7. 集中管控能力

为充分发挥各单点技术能力的优势，形成安全能力合力，可通过单点技术的平台化应用，构建满足数据安全集中运营的数据安全运营平台，支撑监管单位统一监测和预警的数据安全监测平台、以及持续评估访问过程可信的零信任数据安全平台。

1) 数据安全运营平台

以“可实用、可持续”为设计初衷建设一站式、体系化的数据安全运营平台，集成全生命周期防护与监测的安全能力于一身，实现数据资产运营、标准策略运营、安全事件运营和安全风险运营。其中，数据资产运营包括数据资产分布视图、敏感数据分布视图以及资产安全评估指标稽核、资产认领指标稽核和备案指标稽核等；标准策略运营包括合规指标稽核、策略落实指标稽核、策略下发指标稽核和策略生效指标稽核等；安全事件运营包括事件处置稽核指标和数据流动视图等；安全风险运营包括风险监测视图和风险分析视图。通过可视化的信息呈现与工作引导，真正实现“统一部署、统一监控、统一管理、统一运营”的数据安全日常化、可持续的运营管控目标。

有关数据安全运营平台的更多技术详情，可参考本白皮书附录 C.1.2。

2) 数据安全监测平台

面向数据安全监管机构和大型企业集团数据安全总体责任部门，依托国家数据安全法规、行业数据安全监管政策和企业数据安全总体策略，提供面向物理和网络分散的数据使用组织和信息系统的分布式数据安全检测平台；以分布式的数据库流量探针、应用流量探针、运维流量探针、数据安全指标收集前置机、数据安全扫描评估工具为基础，实现对数据全生命周期的在线和离线合规监测；





通过平台化的数据安全备案、数据访问行为分析、数据安全合规汇总能力，实现集中式的数据安全统一监测和预警。

有关数据安全监测平台的更多技术详情，可参考本白皮书附录 C.1.8。

3) 零信任数据安全接入平台

传统场景中，网络安全都是在以防火墙为边界的，其前提是内网完全可信。但是，随着技术的更新，远程办公变得普遍，诸如 VPN 之类的技术在网络应用中非常普遍。这就使得企业的安全边界必须进行扩大以允许来自全球各地各种类型终端的访问。随着以 5G、工业互联网为代表的新基建的不断推进，还会进一步加速“无边界”的进化过程。

零信任代表了新一代的网络安全防护理念，它的关键在于打破默认的“信任”，用一句通俗的话来概括，就是“持续验证，永不信任”。零信任数据安全平台默认不信任企业网络内外的任何人、设备和系统，而是以身份认证和授权重新构建访问控制的信任基础，任何在授权前的人/设备/应用程序都视为不被信任的、在访问数据前都应取得信任，细粒度数据访问权限，采用最小权限原则，将网络防御的对象从边界缩小到了单个或者更小的资源组，从而确保身份可信、设备可信、应用可信和链路可信。基于零信任原则，可以保障办公系统的三个“安全”：终端安全、链路安全和访问控制安全。

有关零信任数据安全接入平台的更多技术详情，可参考本白皮书附录 C.1.3。

3.6. 数据安全运营体系

3.6.1. 运营体系架构

在《数据安全治理白皮书 3.0》版本数据安全治理体系的基础上，参照经典的 PDCA 模型，强化数据安全运营体系在数据安全治理中的轴心作用，有效上承管理体系，下接技术体系，落实数据流动性带来的持续、动态、闭环管理。

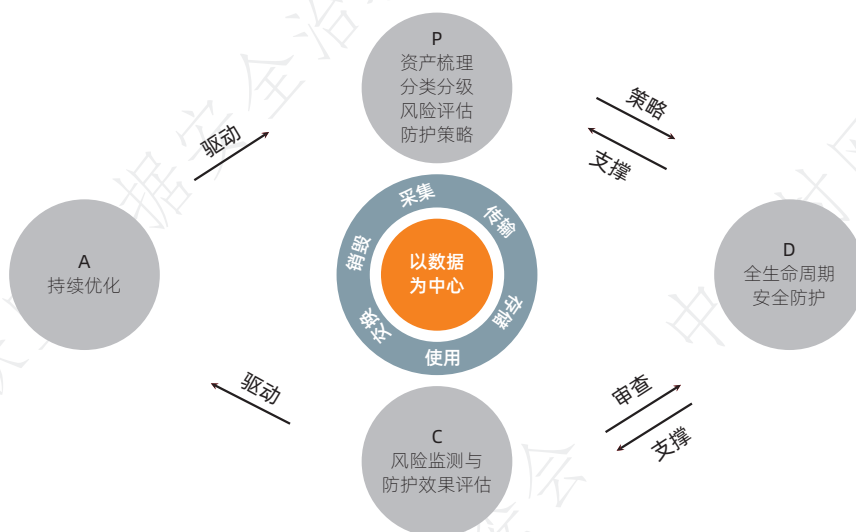


图 10 运营体系建设思路



首先制定数据安全规划（P），通过对组织应遵循的法律法规和行业标准进行解读，结合组织的业务情况，输出并持续更新数据资产分类分级知识库和组织数据安全合规库，并进行数据资产梳理及分类分级，摸清数据资产家底，评估风险暴露面缺陷，再依照合规性要求，针对风险点设定动态分级防护策略，然后落实全生命周期安全防护（D），依据规划制定的安全策略，面向不同级别的敏感数据对象，构建覆盖数据全生命周期节点的按需、动态防御技术能力体系，其次，展开风险监测与防护效果评估（C），实时监测数据安全运行风险，对安全事件进行响应处置，并对安全防护效果进行合规性综合评价，最后，根据风险监测和防护效果评估结果，结合业务变革，进行持续改善、优化（A），迭代驱动下一个安全规划（P）。

3.6.2. 运营体系建设思路

管理和技术体系需要落地，离不开数据安全运营，因此需要从“数据资产、安全策略合规、安全事件、安全风险”四大维度来建设运营手段，量化每个维度的数据安全管控建设指标，明确哪里做的好、好到什么程度，又有哪些做的不足、哪里需要改进和优化等等，不断丰富和提升数据安全建设的完整性和成熟度。

通过运营平台支撑数据安全合规管理和运营，实现数据安全运营流程化、规范化，持续保护数据安全，整体安全运营体系包括如下几个关键节点：

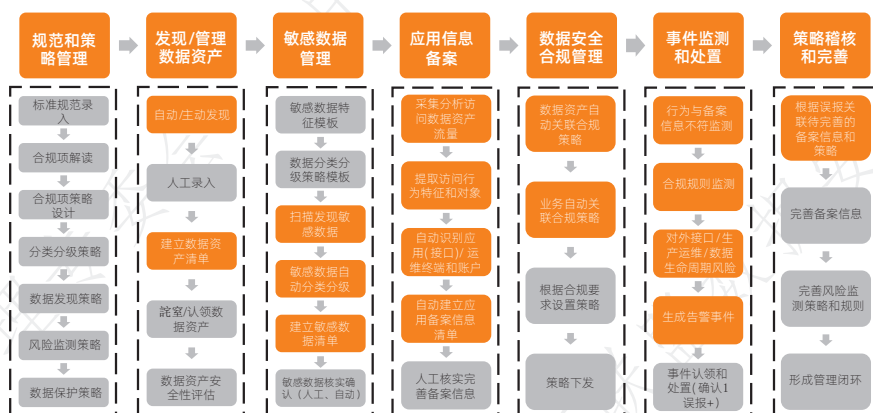


图 11 运营体系框架图

1) 规范和策略管理

调研与组织相关的国内外数据安全相关监管要求、业界数据安全治理相关方法论和架构体系，作为组织数据安全治理优化工作的方法论和框架体系依据，形成数据分类分级策略、风险监测策略和数据保护策略。

2) 发现/管理数据资产

运用数据安全治理平台的自动发现工具，结合手工调研梳理，对数据资产进行持续维护，发现敏感数据库和敏感数据的位置和分布，统计重要数据，并对数据资产的归属部门、责任人、使用方等信息进行备案登记，形成数据资产列表。并对数据库资产进行安全评估，识别数据库资产的脆弱点。





3) 数据分类分级管理

基于数据分类分级策略模板，借助敏感数据发现技术和数据自动分类分级工具，对新发现的数据资产进行分类分级识别与打标，快速建立数据分类分级清单，并进行人工核实确认。

4) 应用信息备案

收集分析数据资产访问流量，提取访问行为特征和对象，识别应用账号和运维账号，建立应用信息备案清单，并由人工进行备案信息核实完善。

5) 数据安全合规管控

梳理评估各级别数据资产在整个数据生命周期流动中的安全风险，借助数据安全平台的策略管理工具，针对风险设定各安全防护节点的防护策略，对数据安全进行合规管控。

6) 风险监测和事件响应处置

借助数据安全平台的风险监测与处置能力，持续监测数据流转风险，针对产生的安全事件告警，落实应急响应与事件处置，并进行事件溯源取证。

7) 运营稽核与优化

建立面向安全策略、安全风险、安全事件的 KPI 指标分析，对数据安全运营情况进行整体稽核，根据稽核考核结果进行持续完善和优化。

3.6.3. 运营服务场景及流程

3.6.3.1. 日常运营服务

日常数据安全运营服务参照运营体系中的内容从数据安全摸底、数据安全策略的制定与升级、数据安全风险管理，以及数据安全策略优化等方面对数据安全开展全方位的工作。

数据安全运营是一个持续化运营的过程，在日常运营服务中需安排安全运营人员持续进行数据安全风险监测。新的法律法规、行业标准的实施，和对已实施法律法规及行业标准的重新认识，都会触发运营人员进行重新解读，转化为新的安全管控策略；业务系统的变更需对涉及的数据资产进行重新分类分级，更新安全策略；安全事件的发生也会促进安全运营人员优化数据安全措施，不断完善数据安全治理体系。

3.6.3.2. 数据安全评估服务

数据安全评估服务首先通过人工调研及自动化工具等手段识别数据资产、收集业务应用场景并进行汇总；然后结合行业监管标准规范的安全防护要求，汇总出基于管理制度规范、操作规范流程和技术防护工具三个维度的应遵循的数据安全防护要求，形成数据安全风险评估策略；其次基于场景的风险暴露面（即自身脆弱性）、安全威胁主体、安全威胁等多个方面进行识别分析；最终通过从制度保障、数据分类分级、权限管理、安全审计、合作伙伴、应急响应等多个维度的配置核查，发现可能存在业务应用场景的高风险问题和威胁行为，形成数据安全评估报告。



3.6.3.3. 应急保障

应急保障针对每年国内重大、重要事件，以及全网突发的重大安全事件提供应急支持服务。主要工作内容包含但不限于：制定数据安全应急方案、工作要求及相关制度；在事前为应急响应做好预备性的工作，做好数据备份；在安全事件发生后，按要求及时对异常的系统、网络进行分析，确定安全事件的各项技术细节，保留相关证据并制定进一步的响应策略；及时采取行动限制安全事件扩散和影响的范围，限制潜在的损失与破坏，保障系统正常运行，恢复受到毁损的数据；事后通过对有关安全事件或异常行为的分析结果，找出根源，明确相应的补救措施并协助完成彻底清除；协助恢复安全事件所涉及的系统，并还原到正常状态，使业务能够正常运行。

3.7. 数据安全治理规划建设

3.7.1. 数据安全治理整体建设思路



图 12 数据安全治理整体建设思路

数据安全治理规划建设，围绕收集、存储、使用、加工、传输、提供、公开等数据处理活动，分为三个阶段建设：

基础阶段－基础防护建设，主要以咨询服务为主，包括资产管理、数据分类分级、安全评估、管理制度建设、安全策略建设、方案规划。

优化阶段－策略优化及能力提升建设，主要以产品为主，依据管理流程和安全策略部署自动化防护工具，解决业务风险。

运营阶段－数据安全管控平台建设，主要基于管理流程、安全策略、业务场景和防护积累的经验，形成行为特征库、安全事件知识库、结合业务场景的积累，形成风险分析模型，建立数据安全管控平台，有监测、有预警、有管理、有控制、有审计、有运维、可感知。

3.7.2. 数据安全治理迭代式建设思路

在数据安全治理实践中，一次性建立完整的数据安全体系存在方案复杂、投入高、周期长，见效慢、



效果不可控等问题,且各组织数据安全建设基础情况不一,宜循序渐进地开展数据安全治理建设工作。

用户需坚持以业务数据资产为核心,基于企业或组织当前的数据安全治理现状和数据安全治理目标,将数据安全治理体系的建设划分为多个阶段,并将管理+技术+运营的建设思想贯穿在每一个阶段的建设任务中,快速达成阶段性目标,再由前一个阶段的建设成果指导下一个阶段的建设目标,一步步证明路线选择的正确性,通过“小步快跑,不断迭代”的方式,横向形成应用全面纳管,纵向持续优化防护策略,逐步建成数据安全闭环管控体系。

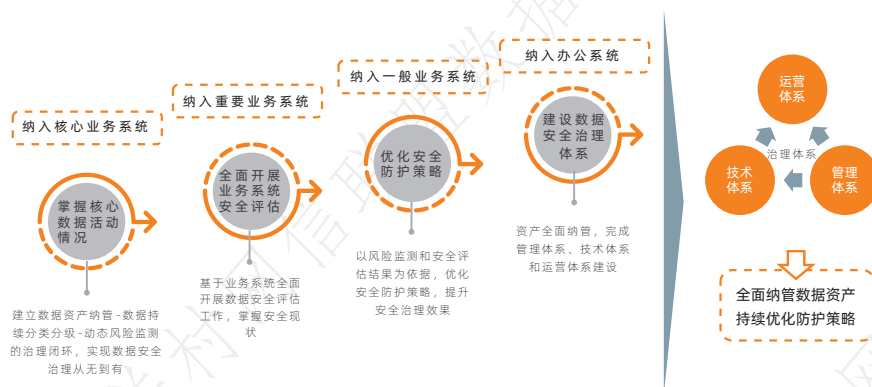


图 13 数据安全治理迭代建设思路



第四章 法律法规标准解读

数据的开发利用必须要依法依规开展，满足合规要求是数据安全底线。通过对国内、国际数据安全相关的法律法规解读，了解监管层面对有序推动数据价值发挥的安全保障要求，为数据安全治理的开展奠定合规依据。

4.1. 国内法律法规与行业标准解读

我国在数据安全治理已经形成以法律、行政法规、部门规章及规范性文件、地方性法规、以及相关行业标准、指南等相结合的综合性治理体系，不断加强数据治理的规范建设。其中，《网络安全法》从法律层面保障了广大人民群众在网络空间的利益，有效维护了国家网络空间主权和安全，是国家基本法律；《数据安全法》、《个人信息保护法》从法律层面提供了数据安全保障和个人信息保护，是相关领域的基础性法律。《网络安全法》、《数据安全法》以及《个人信息保护法》共同构成了我国数据保护的基础体系，整体的法律法规及规范脉络关系如下图所示：

层级	名称	
上位法	基本法律	基础性法律
	《网络安全法》	《数据安全法》 《个人信息保护法》
下位法	行政法规	
	《关键信息基础设施安全保护条例》	《网络数据安全管理条例（征求意见稿）》
部门规章及规范性文件	《网络安全审查办法》	《数据出境安全评估办法（征求意见稿）》
	《汽车数据安全管理若干规定（试行）》	《工业和信息化领域数据安全管理办法（试行）（征求意见稿）》
地方性法规	《银行业金融机构数据治理指引》	《国家健康医疗大数据标准、安全和服务管理办法（试行）》
		
标准、指南等	《深圳经济特区数据条例》	《上海市数据条例》
	《信息安全技术 个人信息安全规范》（GB/T 35273-2020）	《网络安全标准实践指南——网络数据分类分级指引》
		

图 14 数据安全相关法律法规及规范关系脉络图

4.1.1. 法律

在数据安全治理体系上，我国已形成以《网络安全法》、《数据安全法》、《个人信息保护法》三部基础法律为核心的法律架构，并分别从数据所承载系统的安全保障、一般数据的安全保护以及个人信息的安全保护的层面构建较为完整的数据安全保护体系。

1) 《网络安全法》保障网络与信息安全

我国于 2016 年颁布并于 2017 年施行的《网络安全法》是我国网络安全管理方面的第一部基础性立法，旨在应对我国网络安全领域的严峻形势，以制度建设加强网络空间治理，规范网络信息传





播秩序，惩治网络违法犯罪。《网络安全法》全面地规定网络与信息安全治理的基本规则，以网络运营者及关键信息基础设施运营者为主要规制对象，明确网络运行安全、网络信息安全、监测预警与应急处置等方面的义务。

近年来，各机构因网络安全防护能力薄弱进而导致的网络攻击、网络安全事件层出不穷、数据泄露及毁损危机凸显，由于网络安全的一大重要内容即是保障网络数据的完整性、保密性、可用性的能力，因此《网络安全法》对于构建数据安全保障体系有着重要的意义。具体而言，《网络安全法》通过规定下述措施，以强化数据安全保障。

● 保障网络运行安全

《网络安全法》第 21 条构建了网络安全等级保护制度，在保障网络系统安全的组织架构及管理體系上，《网络安全法》要求制定内部安全管理制度和操作规程，确定网络安全负责人，落实网络安全保护责任。

在保障网络系统安全的技术体系上，《网络安全法》要求采取包括防范计算机病毒和网络攻击、网络侵入等危害网络安全行为的技术措施；采取监测、记录网络运行状态、网络安全事件的技术措施，并按照规定留存相关的网络日志不少于六个月；采取数据分类、重要数据备份和加密等措施。

针对关键信息基础设施的运营者，《网络安全法》在第三章第二节中施加了更多的安全保护义务。而且，《网络安全法》第 37 条强化对关键信息基础设施运营者数据跨境传输的监管，提出数据本地化存储及跨境传输安全评估的要求，明确“关键信息基础设施的运营者在中华人民共和国境内运营中收集和产生的个人信息和重要数据应当在境内存储。因业务需要，确需向境外提供的，应当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；法律、行政法规另有规定的，依照其规定”。

● 保障网络用户信息安全

在保障网络用户信息安全的管理体系上，《网络安全法》第 40 条要求建立健全用户信息保护制度。

在保障网络用户信息安全的技术措施上，《网络安全法》第 40、42 条重点要求对其收集的用户信息严格保密；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。

同时，《网络安全法》第 47 条加强对其用户发布的信息的管理，发现法律、行政法规禁止发布或者传输的信息的，应当立即停止传输该信息，采取消除等处置措施，防止信息扩散，保存有关记录，并向有关主管部门报告。

● 保障网络数据安全

在保障网络数据安全上，《网络安全法》第 10 条提出保障网络数据完整性、保密性和可用性的要求；《网络安全法》第 37 条提出对关键信息基础设施运营者数据本地化存储及跨境传输安全评估的要求。

在保障网络数据安全与促进利用上，《网络安全法》第 18 条强调了数据利用与数据安全之间的平衡，鼓励开发网络数据安全保护和利用技术。

2) 《数据安全法》构建数据安全治理框架

由于数据安全已成为事关国家安全与经济社会发展的重大问题，为了有效应对数据这一非传统



领域的国家安全风险与挑战，切实加强数据安全保护，维护公民、组织的合法权益，并发挥数据的基础资源作用和创新引擎作用，推进政务数据资源开放和开发利用，2021 年我国颁布并实施《数据安全法》。《数据安全法》作为数据领域的基础性法律，重点确立数据安全保护管理各项基本制度，构建起数据安全治理的框架，强调数据开发利用与保障数据安全并重的思路，将个人、企业和公共机构的数据安全纳入保障体系，确立了对数据领域的全方位监管。

● 数据开发利用与数据安全并重

《数据安全法》在第 1 条立法目的中就将“保障数据安全，促进数据开发利用”纳入，强调数据开发利用与保障数据安全并重的思路。而且，《数据安全法》在第二章也对数据安全与发展进行专章规定，提出一系列促进数据开发利用的思路及措施，包括：国家统筹发展和安全，坚持以数据开发利用和产业发展促进数据安全，以数据安全保障数据开发利用和产业发展；国家实施大数据战略，推进数据基础设施建设，鼓励和支持数据在各行业、各领域的创新应用；国家推进数据开发利用技术和数据安全标准体系建设。国务院标准化行政主管部门和国务院有关部门根据各自的职责，组织制定并适时修订有关数据开发利用技术、产品和数据安全相关标准。国家支持企业、社会团体和教育、科研机构等参与标准制定；国家建立健全数据交易管理制度，规范数据交易行为，培育数据交易市场等。

● 构建数据安全制度体系

由于不同维度的数据的价值不一，而且对于国家利益、社会利益、个人利益有着不同程度的影响，数据安全治理首先需要实施数据的分类分级保护，以避免因重要数据泄露、损毁带来影响国家安全、社会安全的严重后果。

鉴于此，《数据安全法》第 21 条确立了以数据分类分级为核心的安全制度，根据数据在经济社会发展中的重要程度，以及一旦遭到篡改、破坏、泄露或者非法获取、非法利用，对国家安全、公共利益或者个人、组织合法权益造成的危害程度，对数据实行分类分级保护。同时，《数据安全法》第 22 条要求建立相应的数据安全风险评估、报告、信息共享、监测预警机制及数据安全应急处置机制等。对于开展数据处理活动的主体，可以数据分类分级为基础，形成组织、管理、技术体系相融合的数据安全治理体系。

● 实施全生命周期的数据安全保护

《数据安全法》第四章紧盯数据泄露、数据漏洞以及非法使用数据的风险，从数据处理的全生命周期提出合规要求，包括开展数据处理活动应当建立健全全流程数据安全管理制度；采取相应的技术措施和其他必要措施，保障数据安全；加强风险监测，发现数据安全缺陷、漏洞等风险时，应当立即采取补救措施；不得窃取或者以其他非法方式获取数据等。

为了实现数据生命周期安全保护的要求，开展数据处理活动可采取集中策略管控能力与单点防护能力结合的防控措施，统一部署防护策略，在数据收集、存储、使用、加工、传输、提供、公开等数据处理活动中采取相适应的防护技术和能力。

● 推进政务数据的安全与开放

《数据安全法》第五章推进政务数据的安全与开放，强调提升运用数据服务经济社会发展的能力。在保障政务数据安全方面，要求国家机关应当依照法律、行政法规的规定，建立健全数据安全





管理制度，落实数据安全保护责任，保障政务数据安全；国家机关委托他人建设、维护电子政务系统，存储、加工政务数据，应当经过严格的批准程序，并应当监督受托方履行相应的数据安全保护义务。受托方应当依照法律、法规的规定和合同约定履行数据安全保护义务，不得擅自留存、使用、泄露或者向他人提供政务数据。

在推进政务数据开放方面，要求国家机关应当遵循公正、公平、便民的原则，按照规定及时、准确地公开政务数据；国家制定政务数据开放目录，构建统一规范、互联互通、安全可控的政务数据开放平台，推动政务数据开放利用。

3) 《个人信息保护法》保障个人信息全生命周期处理的安全

相较于一般数据，个人信息因对个人权益的影响需要进行专门的保护。为了解决一些企业、机构甚至个人，从商业利益等出发，随意收集、违法获取、过度使用、非法买卖个人信息，利用个人信息侵扰人民群众生活安宁、危害人民群众生命健康和财产安全等问题，在保障个人信息权益的基础上，促进信息数据依法合理有效利用，2021 年颁布并施行的《个人信息保护法》以数据中的“个人信息”为主要规范对象，划定个人信息全生命周期处理的安全保护规则，以保护个人信息权益、促进个人信息合理利用。具体如下：

● 个人信息全生命周期处理的防护

根据《个人信息保护法》，个人信息的处理包括个人信息的收集、存储、使用、加工、传输、提供、公开、删除等的全生命周期，相应个人信息处理的风险也贯穿于个人信息处理的始终。例如，在个人信息采集传输阶段，因网络端口及传输通道的安全性问题而导致个人信息的毁损和丢失；在个人信息存储阶段，因未采取加密、脱敏存储而导致敏感信息的泄露；在个人信息使用阶段，因超权限的访问或者未经授权的使用而导致个人信息对外泄露、个人信息滥用；在个人信息销毁阶段，因未及时、有效销毁存储介质上的个人信息而导致的个人信息泄露等等。

鉴于此，《个人信息保护法》以第一章总则及第二章的一般规定划定了个人信息全生命周期处理的规则，要求个人信息处理应具备包括同意在内的合法基础、告知个人完整的个人信息处理事项。同时，《个人信息保护法》也对需要重点保护的敏感个人信息、风险程度较高的个人信息跨境提供予以特别规定。在《个人信息保护法》第二章第二节项下，对于处理敏感个人信息的，需具备具有特定的目的和充分的必要性及采取严格保护措施下方可处理，并且还要求取得个人的单独同意等；在《个人信息保护法》第三章项下，对于个人信息跨境提供，划定应当具备的条件，根据第 38 条，满足下列条件之一才可向境外提供个人信息，包括：（1）通过国家网信部门组织的安全评估；（2）经专业机构进行个人信息保护认证；（3）按照国家网信部门制定的标准合同与境外接收方订立合同等条件；（4）法律、行政法规或者国家网信部门规定的其他条件。

● 明确个人信息处理者所应采取的组织措施及技术措施

在《个人信息保护法》第五章规定了个人信息保护的组织架构及管理体系，个人信息处理者应当承担个人信息保护的主要责任；在作为个人信息处理者的组织机构内部，要求制定内部管理制度和操作规程；合理确定个人信息处理的操作权限；指定个人信息保护负责人，负责对个人信息处理活动以及采取的保护措施等进行监督等。

在个人信息保护的技术措施上，要求采取措施防止未经授权的访问以及个人信息泄露、篡改、





丢失，包括对个人信息实行分类管理；采取相应的加密、去标识化等安全技术措施。

- 赋予个人充分的个人信息权利，保障个人信息权益

根据《个人信息保护法》第四章，我国赋予个人充分的个人信息权利，包括个人对其个人信息的处理享有知情权、决定权，有权限制或者拒绝他人对其个人信息进行处理；个人有权向个人信息处理者查阅、复制其个人信息；个人有权请求个人信息处理者更正、补充不准确或不完整的个人信息；个人在特定情形下有权请求个人信息处理者删除个人信息；个人有权要求个人信息处理者对其个人信息处理规则进行解释说明等。

而且，《个人信息保护法》第 50 条要求个人信息处理者应当建立便捷的个人行使权利的申请受理和处理机制，并赋予个人可以在个人信息处理者拒绝个人行使权利的请求下向法院起诉的权利。上述做法充分保障了个人在其个人信息处理活动中的权益。

4.1.2. 行政法规

1) 《关键信息基础设施安全保护条例》实施关键信息基础设施重点保护

- 重点防范关键信息基础设施风险

我国秉承“抓重点、保关键”的立法思路，2017 年 6 月 1 日生效的《网络安全法》第 31 条以列举的方式引入了关键信息基础设施的概念，列举了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 7 个重要行业和领域，并规定关键信息基础设施的具体范围和安全保护办法由国务院制定。

为配合《网络安全法》第 31 条的实施，2021 年 9 月 1 日生效的《关键信息基础设施安全保护条例》第二章专门规定了关键信息基础设施的认定程序，规定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、国防科技工业等重要行业和领域的主管部门、监督管理部门是负责关键信息基础设施安全保护工作的部门。保护工作部门制定关键信息基础设施认定规则，并根据该认定规则组织认定本行业、本领域的关键信息基础设施，及时将认定结果通知运营者，并通报国务院公安部门。通过制定《关键信息基础设施安全保护条例》，我国实施对关键信息基础设施的重要保护。

- 重点强调对数据安全的保护

《网络安全法》、《数据安全法》、《关键信息基础设施安全保护条例》均对涉及关键信息基础设施的数据进行特别保护。《网络安全法》第 37 条强调关键信息基础设施运营者在中国境内收集个人信息和重要数据本地化存储及跨境数据传输安全评估的义务。《数据安全法》第 31 条也提出关键信息基础设施的运营者在中国境内运营中收集和产生的重要数据的出境安全管理适用于《网络安全法》的规定。

《关键信息基础设施安全保护条例》进一步强调对数据安全的保障，包括关键信息基础设施运营者应维护数据的完整性、保密性和可用性；履行个人信息和数据安全保护责任，建立健全个人信息和数据安全保护制度；发生重要数据泄露、较大规模个人信息泄露等情形时，保护工作部门及时向国家网信部门、国务院公安部门报告的义务等。

同时，为了进一步发现安全风险，《关键信息基础设施安全保护条例》第 17 条也要求运营者应





当自行或者委托网络安全服务机构对关键信息基础设施每年至少进行一次网络安全检测和风险评估，对发现的安全问题及时整改，并按照保护工作部门要求报送情况。

2) 《网络数据安全条例（征求意见稿）》细化数据安全治理规则

2021 年国家互联网信息办公室代国务院颁布《网络数据安全条例（征求意见稿）》，对《网络安全法》、《数据安全法》、《个人信息保护法》的规则予以进一步的细化，推动上述法律的进一步落地。

《网络数据安全条例（征求意见稿）》规定了数据安全的一般规则、个人信息保护、重要数据安全、数据跨境安全管理、互联网平台运营者义务等的重要内容，具体而言：（1）在数据安全部分，要求数据处理者应当采取备份、加密、访问控制等必要措施，保障数据免遭泄露、窃取、篡改、毁损、丢失、非法使用，应对数据安全事件，防范针对和利用数据的违法犯罪活动，维护数据的完整性、保密性、可用性；按照网络安全等级保护的要求，加强数据处理系统、数据传输网络、数据存储环境等安全防护等；（2）在个人信息保护部分，细化取得个人同意的合规要求、个人信息处理规则的内容等；（3）在重要数据安全部分，进一步明确数据安全机构的具体职责，并要求数据处理者履行制定数据安全培训计划、优先采购安全可信的网络产品和服务等的义务；（4）在数据跨境安全管理部分，要求建立健全数据跨境安全相关技术和管理措施等；（5）在互联网平台运营者义务部分，要求互联网平台运营者对接入其平台的第三方产品和服务承担数据安全治理责任，利用人工智能、虚拟现实、深度合成等新技术开展数据处理活动按照国家有关规定进行安全评估等。

4.1.3. 部门规章及规范性文件

1) 各行业各部门数据安全的协同治理

数据安全对于各行各业都十分重要，我国的数据治理也呈现出各部门协同治理的局面。为了应对各自领域的数据安全治理的风险，网信部门、工信部门、公安部门以及金融、卫生等行业主管部门均在各自职责范围内颁布数据安全治理的规范文件。例如，国家互联网信息办公室、工信部等五部门发布《汽车数据安全管理若干规定（试行）》，国家互联网信息办公室、国家发改委等十三部门联合发布《网络安全审查办法》，国家互联网信息办公室发布《数据出境安全评估办法（征求意见稿）》、工信部发布《工业和信息化领域数据安全管理办法（试行）（征求意见稿）》、中国银行保险监督管理委员会发布《银行业金融机构数据治理指引》、国家卫生健康委员会发布的《国家健康医疗大数据标准、安全和服务管理办法（试行）》等，上述主管部门均提出各自领域的数据安全治理规范。

2) 数据安全治理领域的重要规范

● 整体要求

➤2021 年 12 月 28 日，国家互联网信息办公室、国家发改委等十三部门联合发布《网络安全审查办法》（“《办法》”），适用于关键信息基础设施运营者采购网络产品和服务，网络平台运营者开展数据处理活动，影响或者可能影响国家安全的情形。《办法》规定了网络安全审查应考虑的国家安全风险因素，明确网络安全审查的具体流程。其中，在网络安全审查应考虑的国家安全风险



因素中，特别考虑数据安全的因素，将“核心数据、重要数据或者大量个人信息被窃取、泄露、毁损以及非法利用、非法出境的风险”、“上市存在关键信息基础设施、核心数据、重要数据或者大量个人信息被外国政府影响、控制、恶意利用的风险，以及网络信息安全风险”纳入。

►2021年10月29日，国家互联网信息办公室发布《数据出境安全评估办法（征求意见稿）》，旨在规范数据出境活动，保护个人信息权益，维护国家安全和社会公共利益，促进数据跨境安全、自由流动。该办法明确需要申报数据出境安全评估的情形、数据出境风险自评估及数据出境安全评估所考虑的事项等。

● 金融领域

►2020年5月21日，中国银行保险监督管理委员会发布了《银行业金融机构数据治理指引》（“《指引》”），《指引》是为引导银行业金融机构加强数据治理，提高数据质量，充分发挥数据价值，提升经营管理水平，全面向高质量发展转变而制定的法规。《指引》共七章五十五条，其中明确了数据治理架构，要求确保数据治理资源充足配置，明确董事会、监事会和高管层等的职责分工，提出可结合实际情况设立首席数据官。《指引》还明确了监管机构的监管责任、监管方式和监管要求。对于数据治理不满足有关法律法规和监管规则要求的银行业金融机构，要求其制定整改方案，责令限期改正；或与公司治理评价、监管评级等挂钩；也可依法采取其他相应监管措施及实施行政处罚。《指引》实施将进一步提升银行业金融机构的数据治理能力，同时也对数据合规提出了更高的要求。

►2021年1月，银保监会发布《中国银保监会监管数据安全管理办法（试行）》，该办法规范了银保监会监管数据安全管理工作，提高了监管数据安全保护能力，目的是建立健全监管数据安全协同管理体系，推动银保监会有关业务部门、各级派出机构、受托机构等共同参与监管数据安全保护工作，加强培训教育，形成共同维护监管数据安全的良好环境。

►2022年1月，中国人民银行印发《金融科技发展规划（2022—2025年）》（“《规划》”），强调全面加强数据能力建设，在保障安全和隐私前提下推动数据有序共享与综合应用，充分激活数据要素潜能，有力提升金融服务质效。

● 工业和互联网

►2019年11月，国家互联网信息办公室、工业和信息化部、公安部、市场监管总局联合制定了《App违法违规收集使用个人信息行为认定方法》，该办法落实《网络安全法》等法律法规，明确App违法违规收集使用个人信息的主要情形，可以为监督执法提供参考，也可以为企业履责合规指明方向，为网民社会监督提供依据。

►2021年3月，国家互联网信息办公室、工业和信息化部、公安部、国家市场监督管理总局联合制定了《常见类型移动互联网应用程序必要个人信息范围规定》（“《个人信息范围规定》”），该规定明确了39类常见App的基本功能服务和必要个人信息范围，这些App覆盖了大众衣食住行、学习工作等日常生活主要方面。《个人信息范围规定》明确要求App运营者不得因用户不同意收集非必要个人信息，而拒绝用户使用其基本功能服务，在保障App正常运行的同时，保障了用户对App基本功能服务的使用权，以及对收集使用非必要个人信息的知情权和决定权。

►2021年9月，为提升工业、电信行业数据安全保护能力，防范数据安全风险，工业和信息化部发布了《工业和信息化领域数据安全管理办法（试行）（征求意见稿）》（“《办法（征求意见稿）》”），





该办法充分贯彻《数据安全法》的法治精神，强调建立工业和信息化领域重要数据和核心数据备案管理制度，统筹建设备案管理平台鼓励数据开发利用和数据安全技术研究，支持推广数据安全产品和服务，培育数据安全企业、研究和服务机构，壮大数据安全产业，提升数据安全保障能力，促进数据的创新应用。《办法（征求意见稿）》第二章明确数据分类分级的方法，将工业和电信数据分为一般数据、重要数据和核心数据，并说明可以纳入上述分类数据的具体条件。

4.1.4. 地方性法规

我国各地方同样在不断探索数据安全治理的规则及模式，已出台的《深圳经济特区数据条例》、《上海市数据条例》、《浙江省公共数据条例》、《贵州省大数据安全保障条例》等均在不断深化我国的数据安全治理的模式，提出了一些数据安全治理的新措施。

1) 创新数据安全治理新模式

《深圳经济特区数据条例》涵盖个人数据、公共数据、数据要素市场、数据安全等方面，可以认为是国内数据领域首部基础性、综合性立法。在个人数据部分，《深圳经济特区数据条例》明确规定自然人对其个人数据享有人格权益，细化个人信息处理、告知与同意、个人信息权利等的具体规则；在数据安全部分，《深圳经济特区数据条例》明确数据处理全流程记录、数据存储进行分域分级管理、重要系统和核心数据的容灾备份、建立数据销毁规程等的要求。

《上海市数据条例》涉及数据权益保障、公共数据、数据要素市场、数据资源开发和应用、浦东新区数据改革、长三角区域数据合作等重点内容，强调数据资源的利用与开发，以及数据资源的合作，并通过明确数据安全责任制、开展数据处理活动所应履行的义务、健全数据分类分级保护制度等保障数据安全。

2) 提供公共数据治理的模式借鉴

上述地方性法规也突出对公共数据的处理的保护，明确公共数据共享、利用的规则。这也为相关政务大数据共享、政务数据安全检测及敏感数据防泄露等提供有益的指导。

《深圳经济特区数据条例》明确对公共数据进行分类管理，实行公共数据目录管理制度，推动公共数据的共享，以共享为原则，以不共享为例外，并构建公共数据共享、公共数据开放、公共数据利用的治理体系；《上海市数据条例》同样建立公共数据共享和开放的机制，也特别规定了公共数据授权运营的模式，提高公共数据社会化开发利用水平；《浙江省公共数据条例》全面规定公共数据平台、公共数据收集与归集、公共数据共享、公共数据开放与利用、公共数据安全等的内容，促进公共数据应用创新，保护自然人、法人和非法人组织合法权益，保障数字化改革。

4.1.5. 国家标准、行业标准及相关指南

我国出台了一系列的国家标准、行业标准及相关指南，对法律法规等规范性文件中的较为原则性的规定给予了具体的指导。从防止数据泄露、保护个人权益、提升数据安全治理能力等多方面多角度，提出细化的保护要求及保护措施，为我国数据安全治理提供有力的支撑。具体请见附录F主





要数据安全技术标准汇总。

4.2. 国际数据安全政策与法规概述

在数字经济时代，数据安全是各行各业的运行关键，数据安全所带来的风险及损失也不可估量。例如，在 2021 年，美国管道运输公司一条输油干线遭遇黑客攻击，导致停运数天，对美国的经济活动、民众生活、国家安全造成威胁；而在个人信息安全领域，美国知名征信机构 Equifax 也曾遭受数据泄露事件，黑客窃取大量个人信息，超过 1.4 亿美国人受到影响。

各国均高度重视数据安全治理，严厉惩处造成数据安全后果的违规事件。例如，在英国 ICO 曾对英国航空作出的 1.83 亿英镑的罚款，原因在于 2018 年 9 月英国航空网站发生的数据泄露事件，最终导致大约 50 万人的客户信息遭到泄露。同时，全球各主要国家纷纷出台相关法律法规，通过各种组织措施和技术措施强化数据安全，推动数据开发利用、价值挖掘与安全防护的并举发展。

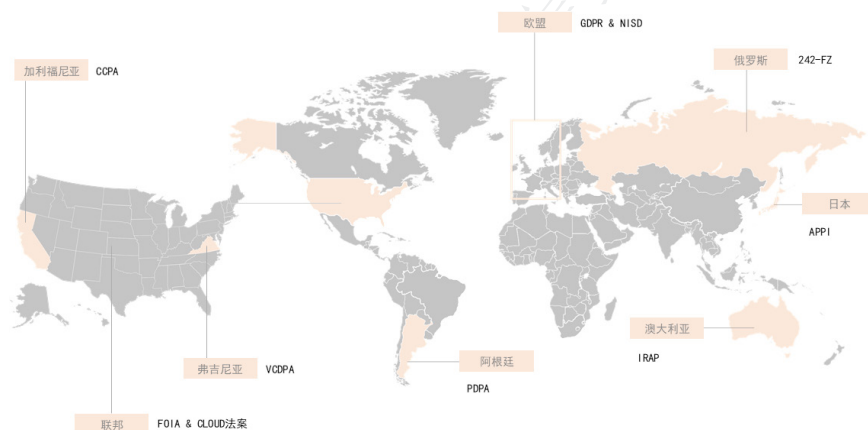


图 15 主要国家的数据安全政策与法律

4.2.1. 欧洲的政策和法律

1) 欧盟

2018 年 5 月 25 日施行的《通用数据保护条例》（GDPR）适用于向欧盟民众提供商品和服务或收集并分析欧盟居民相关数据的组织。GDPR 授予人们管理组织收集的个人数据的权限，可通过数据主体请求（DSR）来行使这些权限，同时，聚焦个人数据的全流程处理，提出个人数据假名化、匿名化设计的和默认的数据保护手段等，加强个人数据的保护。在 GDPR 框架下，组织需要及时提供有关 DSR 和数据泄露的信息，并执行数据保护影响评估（DPIA）。

《网络信息系统安全指令》（NISD）旨在加强基础服务运营商、数字服务提供商的网络和信息系统安全，并要求成员国及时转化为国内立法。NISD 和 GDPR 分别从安全防护和基本权利保障两个角度推行欧盟的网络和数据安全战略。

2) 法国

1978 年 1 月 6 日，法国通过了关于数据处理、数据文件和个人自由的第 78-17 号法案。自 1978 年以来，法国没有效仿其他许多欧洲司法管辖区，进行新的数据保护独立立法，而是选择多次



修改该法案，在历次修正案引入了被遗忘权，罚金刑和数据可迁移性规则。为了衔接 GDPR，该法案于 2018 年 6 月进行了修订，并对 GDPR 的多个规定进行了调整。例如，尽管不要求强制性的事先通知，但该法案仍然要求主管部门（CNIL）对某些特定处理活动进行事先授权或征求意见，例如出于研究或公共利益目的处理健康数据。此外，该法案允许数据主体指定其个人数据在死亡后的处理策略。

3) 德国

德国在 2017 年 6 月 30 日通过了《联邦数据安全法》以落实对 GDPR 国内立法的承诺。德国数据安全执法采取了联邦—地方的两级架构，联邦数据保护与信息自由委员（BfDI）在联邦一级实施数据保护，而 16 个地区数据保护机构则在各自州的公共和私营部门执行数据保护法。所有监管机构会定期在德国数据保护会议（DSK）上举行会谈，发布详细的指导方针，进一步完善了德国的隐私法律框架。其中重要指南包括有关 Cookie 同意的指南、基于 GDPR 的罚金刑评估标准以及标准数据保护模型。此外，德国是首批实现卫生系统数字化的欧洲国家之一，《患者数据保护法》（PDSG）于 2020 年生效。

4) 英国

虽然英国已经退出欧盟，但是作为重要的一项政治遗产，GDPR 已在早前立法中写入英国法律，并量身定制为“英国 GDPR”（UK GDPR）。为此，欧盟委员会为英国通过了两项专门决定，一项是根据 GDPR 的委员会决定，另一项则是适用于英国的补充指令：有关执法行动的保护指令 EU-2016/680。UK GDPR 规定，信息专员办公室（ICO）是英国的监管机构，并详细说明了其职能，包括执法权限，以及制定行为准则的义务。近年来，ICO 表现活跃，发布了涉及广泛主题的指导文件，开展了一系列执法行动。2022 年 3 月 8 日，ICO 发布决定，对英国皇家邮政集团违反隐私条例，在未持有效同意的情况下向个人发送营销邮件的行为处以罚款。

5) 俄罗斯

242-FZ(12.31.2014) 全称“关于阐明信息和电信网络中个人数据处理过程的俄罗斯联邦一揽子法律法案的修订”。该法案确认，自 2015 年 9 月 1 日起，被视为个人数据操作员的组织必须确保在收集个人数据时，俄罗斯公民的个人数据记录、系统化、保存、存储、修改（更新、更改）和提取操作是通过位于俄罗斯的数据库执行的。

4.2.2. 美国的政策和法律

1) 联邦

美国在联邦层面，还未出台统一的联邦隐私法规。但是有多个部门法条款，和一些约束专门领域或事项的法令在联邦层面发挥作用。包括：规制联邦机构信息公开的美国法典第 552 条以及 1986《信息自由法》（FOIA），1996 年的《健康保险流通与责任法案》（HIPAA），该法案规范了健康信息的隐私和安全；1999 年的《格雷姆—里奇—比利雷法案》（GLBA），该法案要求金融机构向客户解释其信息共享做法，并保护敏感数据；1998 年的《儿童在线隐私保护法》（COPPA），对面向 13 岁以下儿童的网站或在线服务的运营商施加了额外的要求。缺乏联邦隐私法或监管机构的现实使联邦贸易委员会（FTC）成为事实上的监管机构，FTC 就侵犯消费者隐私权或未能维护敏感消费者信息的安全性达成了一系列有约束力的判例法和和解。





此外美国还通过了多个法案和法律文件，来确保其对数据的控制和在数字领域的领先地位。2018年，美国通过《澄清域外合法使用数据法案》（Clarifying Lawful Overseas Use of Data Act，以下简称 CLOUD 法案），确立了境外数据管辖控制者标准，使其合法访问境外数据，并对他国获取数据资源的资格进行限制，通过“长臂管辖”的方式不断扩大其跨境数据的执法权力，甚至抢夺他国的“治外法权”。与此同时，美国还大力推动 APEC（亚洲太平洋经济合作组织）的跨境隐私保护准则（APEC Cross Border Privacy Rules，以下简称 CBPR），规定若不同国家的不同企业承诺遵守 APEC 隐私框架中的 9 项个人数据保护原则，则数据就可以在上述公司进行自由的、跨境地流动，从而实现数据向美国流动，掌握全球数据资源。

2019 年 12 月，美国白宫行政管理和预算办公室（Office of Management and Budget, OMB）发布了《联邦数据战略与 2020 年行动计划》（以下简称《联邦数据战略》），其内容包括 1 项使命宣言、10 项原则、40 项实践指导及年度行动计划四个组成部分。《联邦数据战略》是美国把科技创新和数字化转型提到国家战略核心层面部署的重要体现，描述了美国联邦政府在未来十年加速数据使用的愿景，并逐年确定行动计划，并突出了对数据认识的深化，坚持数据是最有价值的国家资产理念，并着重改进特定数据资产组合的管理和使用。

2020 年 10 月，美国国防部发布了《国防部数据战略》（《DoD Data Strategy》），该战略是继 2019 年美国国防部发布的《国防部云战略》、《国防部人工智能战略》和《国防部数字现代化战略》以后又一 IT（信息化）现代化领域指导性战略文件，其中提出国防部应加快向“以数据为中心”的过渡，强调通过数据融合实现军种联合、重视数据的安全性、强调数据全生命周期各个阶段的标准化处理能力。

2021 年 12 月，根据《外国公司问责法案》（HFCAA）规定的监管机构如何界定将被摘牌的公司以及相应的退市程序，美国证券交易委员会（SEC）要求外国公司提供审计底稿供美国检查，不能提供审查的，可能在三年内被纽约证交所和纳斯达克摘牌。受到该法规影响的中概股超过 200 支。美国自 2002 年以来一直要求提供审计底稿，但中国大陆和中国香港一直拒绝检查。2022 年 3 月，SEC 再次公布新版规定征求意见稿，进一步加强要求上市公司关于网络安全风险管理、战略、治理和事件报告的披露，并使之标准化，其中明确了上市公司信息安全事件应在 4 天内披露。2022 年 3 月 -4 月，SEC 分五批将搜狐、百度、微博、富途控股等中概股列入预退市名单。

2) 州

弗吉尼亚州消费者数据保护法（VCDPA）是一项综合隐私法，即将于 2023 年 1 月 1 日开始由弗吉尼亚州总检察长（AG）执行。总检察长可以申请高达 7,500 美元的单次违规赔偿。VCDPA 为弗吉尼亚州消费者提供各种隐私权保护。受 VCDPA 监管的企业对这些消费者负有多项关键义务，包括提供披露，以类似方式响应通用数据保护条例（GDPR）消费者数据主体请求（DSR），以及遵守某些数据处理义务（例如，数据最小化原则，负责任的数据保全措施）。

加州消费者隐私法案（CCPA）美国第一个全面的隐私法。CCPA 赋予加州消费者各种隐私权利，包括知晓其个人信息是否被出售或者披露及其（加州）流向，拒绝个人信息的出售，访问其个人信息，享有非歧视的服务与价格。为了保障这些权利，CCPA 监管下的企业有责任履行多项义务，包括披露、一般数据保护条例（GDPR）等消费者数据主体权利（DSR）、某些数据传输的“选择退出”和未成年人



的“选择加入”要求。

4.2.3. 其他国家的政策和法律

1) 澳大利亚

信息安全注册评估（IRAP）由澳大利亚网络安全中心（ACSC）管理，为针对澳大利亚政府政策和指南的独立评估系统安全提供了一个综合流程。IRAP 的目标是通过专注于存储、处理和传达信息的信息技术基础结构，最大限度地提高澳大利亚联邦、州以及当地政府数据的安全性。IRAP 提供了一个框架，用于认可来自私有和公共部门的个人，以便向澳大利亚政府提供网络安全评估服务，搭建了根据澳大利亚政府政策和指南对系统安全进行独立评估的综合流程。IRAP 的目标是通过专注于存储、处理和传达信息的信息技术基础结构，最大限度地提高澳大利亚联邦、州以及当地政府数据的安全性。

2) 日本

日本《个人信息保护法》（APPI,《令和2年6月12日法律第44号》）旨在就个人信息的正当处理，对其基本理念、政府制定的基本方针以及其他个人信息保护措施的基本事项作出规定，对国家及地方公共团体的职责等予以明确，同时，对个人信息处理业者应遵守的义务等作出规定，以期在确保个人信息有效利用的同时，对个人的权利和利益加以保护。2021年5月12日，日本国会通过了《个人信息保护法》的新修正案。

3) 阿根廷

阿根廷《个人数据保护法》（PDPA）（Ley de Protección de los Datos Personales）已于2000年执行，用以帮助保护个人数据隐私，并为用户提供对存储在公用和专用数据库和注册表中的任何信息的访问权限。下属于最高检察院的阿根廷公共信息获取机构（Agencia de Acceso a la Información Pública, AAIP）负责强制实施这条法律。PDPA 与保护数据隐私的欧洲立法模式一致，使阿根廷成为拉丁美洲第一个获得从欧盟进行数据传输资格的国家。



第五章 数据安全治理落地实践案例集

数据安全治理框架中的管理、技术、运营体系根据业务场景不同具有差异化的落地方式，为方便读者阅读，下表针对不同行业业务场景下的数据安全治理体系落地进行导读。

数据安全治理落地实践案例导读表

章节号	案例名称	案例简介	数据安全治理		
			管理	技术	运营
5.1.1	某政务大数据中心安全运营实践	针对政务大数据中心建立一套覆盖管理、技术和运营三大体系的完整数据安全治理体系	数据安全管理制度；关键数据安全管理制度；	集中管控； 分类分级； 加密；脱敏；	建立统一运营平台
5.1.2	基于隐私计算技术的政务数据安全共享开放实践	针对政务数据实现“可算不可见”的安全数据开放与共享，促进数据在不同组织间安全交换，充分发挥数据价值	/	联邦学习；混淆电路；同态加密；秘密共享；不经意传输；国密算法；	/
5.1.3	政务大数据中心数据流转安全监测实践	实时监测政务大数据中心的数据流转情况，形成数据资产视图和数据流视图，识别数据安全风险，形成数据安全态势感知	/	资产探查；数据流转监测；安全风险监测；泄密审计追溯；UEBA用户行为分析技术；可疑第三方检测模型（STP）算法；	/
5.2.1	某国有银行集中化安全审计应用实践	全面监测大型银行的业务数据处理活动，根据银行业务模型识别数据安全风险	/	集中管控；数据库审计；应用审计；	/
5.2.2	某城商行大数据中心安全运维实践	在运维场景针对数据实现运维账户的细粒度权限控制，并运用脱敏技术实现面向特权运维账户的敏感数据动态遮蔽，和面向加工测试环境的数据批量脱敏	/	集中管控；数据分类分级；数据库运维管控；数据脱敏；数据库安全审计；	/
5.2.3	某国有银行数据防泄露实践	基于数据分类分级结果清单，从终端、网络、邮件等多个维度全面监测敏感数据的泄露风险	/	数据分类分级；终端数据防泄露系统；网络数据防泄露；邮件防泄露；	/
5.3.1	某电网公司数据安全治理实践	全面开展数据安全治理工作，建立数据安全管理体系，落地数据安全技术保障体系	建立组织保障体系；建设管理制度体系；	数据库安全评估；数据库审计；数据库安全运维；数据安全防护；	/
5.3.2	某能源云平台数据安全密码应用实践	运用密码学算法在云平台上实现密码基础设施服务，支撑身份认证、传输加密等安全应用	/	密码前置服务；密码服务平台；安全认证网关；SSL VPN；	/





章节号	案例名称	案例简介	数据安全治理		
			管理	技术	运营
5.4.1	某部委数据安全治理实践	加强教育数据的安全管理，基于数据分类分级结果，确立数据安全治理技术，并落地数据安全治理稽核	敏感数据规范标准；制定教育行业的数据安全管理规范；	/	数据治理稽核；
5.5.1	某电信国际的数据安全治理实践	通过数据安全治理咨询，制定方针、制度、规范、表单四级文件，建立数据安全运营稽核体系	制定数据安全方针、组织架构、制度要求、流程规范、记录表单；	/	实现数据资产管理、数据安全策略管理、数据安全事件监测和数据安全风险运营能力。
5.6.1	某医院医疗数据灾备实践	根据医疗数据分类分级结果，实现数据分级灾备管理	/	数据分类分级；统一灾备管理；	/
5.7.1	某车联网企业基于数据分类分级的防护体系建设实践	根据车联网数据分类分级结果，进行合规性检查，实现数据安全防护	建立数据分类分级台账；	基于国密算法的签名验签技术；数据流转监测；数据加密传输；	/
5.8.1	某烟草公司基于零信任和隐私计算的防护体系实践	按数据生命周期实施烟草大数据安全隔离措施和零信任安全防护，实现工业数据安全防护的动态化、主动化和纵深加强。	/	数据分类分级；零信任安全访问接入；隐私计算技术；	/
5.9.1	某互联网公司个人信息保护法律实践	在广告中避免虚假宣传、保障消费者权益、保护用户个人信息，营造良好的合规舆论及监管环境，提升企业社会声誉及评价	设立数据隐私部门；数据合规相关法律法规规范、监管政策和司法案例解读；报告研究；员工培训。	/	/

5.1. 政务行业数据安全治理

5.1.1. 某政务大数据中心安全运营实践

5.1.1.1. 背景

为主动适应数字信息技术的快速发展，解决部门信息“孤岛”和数据“烟囱”问题，加快推进“互联网+电子政务”，建设“数字政务”，在省政府办公厅大数据和电子政务等管理职责的基础上，组建某省政务大数据中心。

数据正从各个委办局通过一体化大数据平台逐步汇聚，汇聚的数据通过治理、加工后实现数据价值的最大化。政务数据中心划分：从直属单位前置区采用ETL的方式将数据抽取到数据中心前置库，通过中心前置库汇集为原始库，基于原始库的数据进行数据清洗，数据转换等数据治理工作，形成各类主题库，专题库等标准库。结合公共服务开放平台，以及各直属单位所需的数据类型对外提供相应的接口服务或者数据共享服务。



在这个过程中能够接触到数据的有多个团队，数据收集团队、数据治理团队、数据开放团队等等。但是数据安全的建设并没有一个明确的团队，而是由业务团队同时建设，如何保证数据安全能力能够合法合规，如何保证安全能力能持续发挥价值，“确保数据处于有效保护和合法利用的状态，以及具备保障持续安全状态的能力”，是用户亟待解决的问题。

某省大数据局对于电子政务数据的处理，是委托第三方进行的，依据法律要求，“国家机关委托他人建设、维护电子政务系统，存储、加工政务数据，应当经过严格的批准程序，并应当监督受托方履行相应的数据安全保护义务”。鉴于数据接触者较多，数据流转使用场景复杂，由何种角色对海量的政务数据进行安全监测、安全防护是用户面临的第一个问题。

第一，海量数据缺乏完整的政务数据的分类分级的标准，无法区分重要数据，数据安全建设缺少抓手。

第二，涉及的网络现状复杂，缺乏从数据使用角度对访问行为监测，管理，阻断拦截技术手段。

第三，数据的存量，流转，策略管理，缺乏统一的管控手段，缺乏可视化的综合管理平台，对于数据管理者的数据安全风险工作带来不便和障碍。

第四，数据安全建设缺乏常态化、持久化的运营来应对业务系统变化、数据安全需求变化、技术能力更迭带来的新型挑战。

5.1.1.2. 安全对策

依据 IPDR 数据安全治理的思路，建立一套从资产梳理、数据分类分级、安全策略管理、数据流转监测与防护、持续风险分析到响应处置的动态、闭环管控体系。

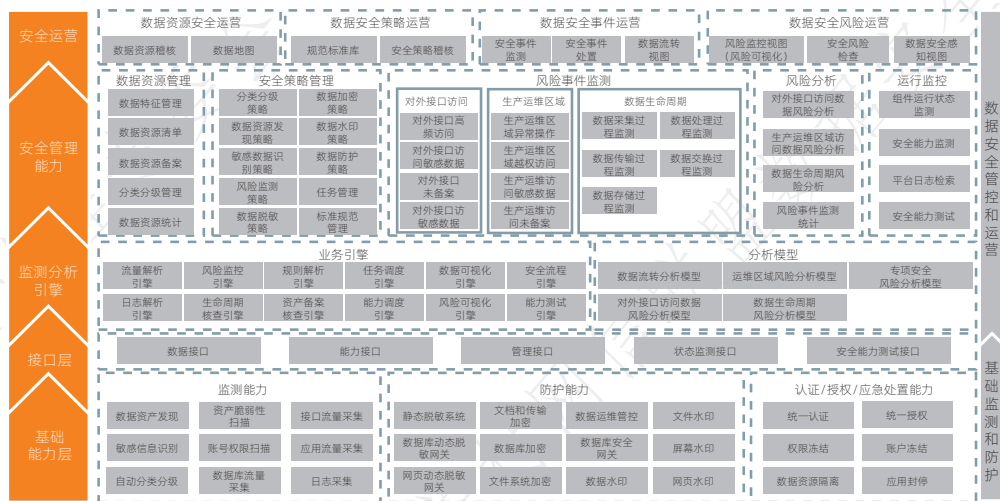


图 16 数据安全运营平台功能架构图

1) 数据分类分级

根据数据分类分级要求，对于一体化大数据平台中的新增数据、变更数据的数据库资产以及数据库账户、权限情况进行梳理清查，同时针对重要数据库进行敏感数据分布梳理、敏感数据访问操作汇总，通过对数据资产的摸底，形成数据资产清单，全局掌握数据资产情况，清晰了解重要数据库的访问情况及风险情况。结合政务数据分类分级指南，进行数据分类分级的打标或变更，持续跟踪数据分类分级的情况，依据数据分类分级结果动态调整安全监测和防护策略。

2) 数据安全监测与防护

通过建设数据安全监测能力、安全防护能力、安全运维能力、安全加密能力、安全溯源能力、安全脱敏能力，形成完整的数据安全防护能力框架。整个数据安全的防护从数据汇聚开始，到数据的共享开放为止。其中涵盖了数据收集、数据加工、数据运维、数据测试、数据开放等多个环节。

要进行数据安全事件、安全风险的判定，先决条件是能监测到所有的数据流向和数据使用情况。通过数据监测节点，对数据的交互，数据流向，访问来源等元素进行监测，对于涉及的数据交互过程中的访问动作与安全策略进行匹配，判定是否存在非法行为。

对于监测的非法操作通过安全防护探针对于动作进行细粒度过滤；对于来自于数据运维人员的操作，通过安全运维探针进行权限管控和高风险动作的流程审批，对于数据分类分级后的重要数据通过加密探针进行存储层加密保护；通过数据安全评估探针持续的对数据进行分类分级打标，通过数据安全分发探针，实现数据分发安全。

3) 统一运营

通过建设数据安全管控平台，形成“数据资产管理、规范和策略管理、风险事件监测、风险态势分析”四大核心能力，实现数据安全“统一运营、集中管控”，提升数据安全技术支撑保障能力。

5.1.1.3. 具体实践

1) 系统部署示意图

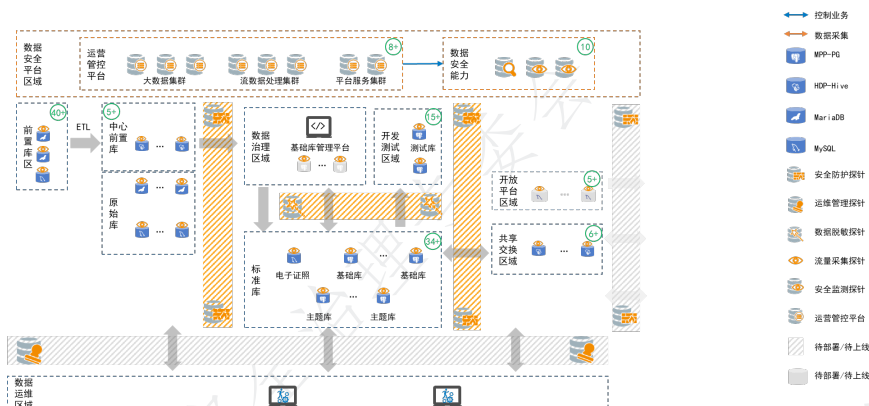


图 17 数据安全运营平台部署示意图

数据安全防护节点部署在数据前置机到中心数据汇聚前置机，中心数据汇聚前置机到数据治理区域，数据治理区域到数据标准库，数据标准库到数据共享开放区域，通过防护节点，保证数据的流入和流出的安全性。

对于整个平台的运维人员接触数据，需要经过数据运维安全节点经过身份认证后，方可介入数据的流转过程中。

对于开发测试环境的数据，经过数据脱敏节点处理完成之后，方才允许流入测试环境。

而对于数据监测节点，则贯穿了整个数据的流传和使用过程，所有的数据的操作都在安全监测节点的监控之下进行。



同时对于核心数据的安全，根据数据分类分级结果，四级数据进行了加密存储的保障。

治理后的数据在进行与地市同步分发时，进行了水印处理，保障数据外发后的可溯源。

在整个数据安全基础监测及防护能力之上，则部署了基于大数据组件的安全运营平台，做到安全能力集中管控、安全事件集中处置、数据资产集中管理、安全防护协同工作。

2) 建设步骤

方案实施过程遵循以下五个建设步骤：梳理评估阶段、制度建设、数据分类分级及技术管控、策略优化。

第一步、梳理评估阶段

通过数据安全运营平台对数据资产现状进行了统一的资产梳理和数据安全风险评估。通过人工的方式对数据应用、数据安全现状、现有数据管理制度进行了调研和摸底。结合政务资源共享目录，形成了初步的数据分类分级规范，经与某省大数据局负责人反复碰撞，形成了最终数据分类分级标准，将标准落地成相应的数据分类分级知识库导入安全运营平台。

第二步、制度建设

基于前期调研的数据应用与数据安全管理制度，结合国家和省政府针对数据安全颁布的法律法规，安全指南，初步制定了一系列的数据安全管理制度、关键数据安全流程等相关规范文件，并结合数据分类分级规范，将制定的流程，制度规范转化为相关的安全策略，导入数据安全运营平台中，形成数据安全策略。

第三步、数据分类分级

结合上述数据分类分级策略，数据安全策略，通过数据安全运营平台对一体化大数据平台涉及的所有数据资产进行了初步的数据分类分级，并基于数据的分类分级策略和数据安全管控策略实施了监测和管控，将数据安全基础监测及防护能力融入数据生命周期过程中，对数据的使用全程留痕，风险事件及时告警，违规动作有效阻断。

第四步、技术管控

以业务场景为出发点，基于数据资产备案及数据分类分级的结果，将数据安全能力嵌入数据业务过程中。伴随着技术管控落地的还有基础的防护策略，来保障数据使用的合规、安全。

第五步、策略优化

通过数据安全管控平台对一体化大数据平台的数据资产展开运营工作，监控数据流转过程，定期输出分析报告及整改措施，跟进整改进度。同时对数据资产进行安全备案，根据备案情况形成数据安全基线。平台运行过程中，适时调整安全策略，跟踪安全管理制度落地情况，持续修订制度文件。使数据安全能力随着持续运营呈稳固上升的趋势。

5.1.1.4. 总结

建设“数据安全集中监测和管控、统一运营”的“数据安全运营管控平台”，实现数据安全技术+数据安全治理+数据安全运营的能力，形成闭环可持续的数据安全管理、服务、运营能力。

1) 数据安全监测和防护能力

通过建设数据安全基础监测能力和数据安全基础防护能力，形成数据安全的基础能力。





2) 数据安全集中管控

通过实现集中的数据资产管理、数据安全策略管理、数据安全事件监测与处置、数据风险分析能力，形成数据安全闭环管理。

3) 数据安全运营

通过实现数据资产安全运营、数据安全策略运营、数据安全事件运营、数据安全风险运营等能力的建设，为数据安全提供信息化支撑手段。

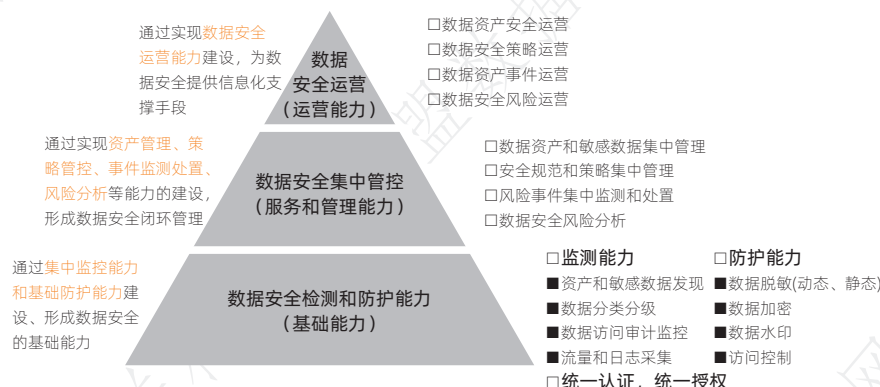


图 18 数据安全运营能力

通过上述能力建设实现了以下数据安全运营管控的效果：

1) 以数据使用场景驱动的数据安全能力建设

首先根据数据使用情况抽象出多个数据业务场景，再根据业务场景的数据交互方式，进一步抽象出对应业务场景需要何种数据安全措施来支撑和保障业务的安全，并分析出相关安全措施之间的先后依赖关系；继而梳理出对应业务场景的数据安全监测与防护工作业务流，并最终形成完整的数据安全场景化模块。

2) 管控与运营并举，以运营提高安全管控能效

从“资产、合规、事件、风险”四大视角施展运营手段，通过量化每个维度的数据安全管控建设指标，以明确哪里做的好、好到什么程度，又有哪些做的有所欠缺、需要改进和优化等等，从而不断丰富和提升数据安全建设的完整性与成熟度。

3) 以工作台引导工作，更轻松地完成日常运营

数据安全的管理、运营与执行是一系列庞杂的工作过程，数据安全运营管控平台以日常运营工作台、待办事项工作台等可视化的方式，将日常运营工作规范化、流程化、指标化——以清晰的业务流程和数据统计，引导监管与执行人员“要做什么、先做什么、后做什么、怎么去做”等等，从而解决相关工作落地难的问题，在提高工作效率的同时，大大降低运营成本投入。

5.1.2. 基于隐私计算技术的政务数据安全共享开放实践

5.1.2.1. 背景

2021 年国家出台了关于数据安全的法律法规：《数据安全法》和《个人信息保护法》。外部监



管的要求和力度逐渐加大，而实际政务数据的安全规划和应用却不能满足现有的市场需求，不能做到真正意义上的数据共享、数据开放和数据交易。

数据安全保护的大环境条件逐渐完善，数据联合使用的商业化需求也在提升。以业务协同为主的政务数据共享，涉及到跨区域、跨级别、跨部门、跨组织之间的配合，但由于配合、协同体制机制不够完善，供需双方权责不够明确，共享范围和边界不够清晰，比如各业务部门的应用系统基于各自业务需求而建，数据存储结构、标准不统一，数据流动蕴藏巨大风险。尤其数据要素区别于传统要素资源的特点是虚拟化的数据要素更易泄露，而数据在流动过程中，环境复杂，并且大部分属于敏感数据的政府数据要素中，隐私数据的保护和安全成为制约数据共享的主要原因。

所以亟待一种能数据安全协同生态平台，融合政府和第三方数据进行公共数据密文开发共享、联合建模，使数据在各业务线之间流通，打通跨域数据的应用价值链。

数据属于可以复制并可以反复利用的资产。但是数据只有通过流通才能产生价值，不能交融的数据信息只能固封在各领域企业的资产库中形成数据孤岛，不能展现最终的生产价值。

而对数据的合法保护，却产生了一个矛盾：政府部门拥有大量数据信息，融合了民生和社会运营数据，涉及多种行业 and 多个业务部门，政务业务数据必然涉及到跨部门协同。可实际能对外共享的数据多为宏观数据，应用价值有限，导致政企、企企数据共享并没有大规模实际应用，大量的有效数据因为需要依法保护而无法被联合在一起使用、计算。

随着数据安全问题的备受重视，数据的融通协作也将迎来更大的发展突破。身处大数据时代，数据协同合作过程中隐私信息的保护将成为企业数据商业竞争的重要砝码，随着公众隐私信息的保护意识增强，隐私计算技术将受到大数据融合应用和隐私保护的双重需求驱动，也是目前国内外政策法规的必然趋势。

5.1.2.2. 安全对策

建设数据安全协同生态平台，首要树立数据公有化的理念。

一是要树立政务数据“公有”的概念。在大数据时代，政务数据不再是部门的“专属品”，所以要破除部门管理、部门专属的旧观念，大力强调数据的公共属性。二是要树立数据“流通共享才能增值”的理念。

政府掌握着大量公共数据资源，只有充分共享和利用，才能不会造成资源浪费。而通过数据共享，充分开发应用政务数据安全隐患计算平台，才能让沉睡的政务数据流动起来，并且产生新的价值。现在一些省市地区已经建立起区域性大数据交易中心，所以构建一体化、整合式的政务数据协同共享平台，无论从国家层面还是地域层面，都是推进公共数据开放和基础数据资源跨部门、跨区域共享。

传统模式通过可信第三方可以实现多信息系统之间的隐私信息交换，但交换后的隐私信息的授权并不能完好控制，而采用去第三方的端到端数据流，数据不出本地库，安全性上也更加容易证明，让用户也有更好的安全感。





隐私计算技术中的多方安全计算和联邦学习是基于密码学等多种技术实现的隐私计算，其核心在于解决如何在无可信第三方的情况下，安全的计算一个约定的函数问题，同时要求每个参与的主体，除了计算结果之外，不能得到其他参与者任何的输入信息。多方参与一起联合运算、联合建模、联合预测，分布式部署的模式使得多方能够相互协同。

隐私计算技术整体过程中主要关注协议安全、密码安全、数据安全、模型安全等多个层面的安全问题。协议安全，主要包括联邦学习、秘密共享、不经意传输等在内隐私计算安全协议，应能够根据场景灵活的配应不同的标准的协议安全假设；密码安全，所使用的隐私计算安全协议所涉及密码算法满足国家标准规范的安全强度要求；数据安全，是指对数据从包括存储、使用、回收在内的全生命周期进行的安全管理；模型安全，主要指在进行联邦学习模型训练的时候，对模型参数的保护，部署时模型访问权限的管控，下架时的历史记录审计等。

5.1.2.3. 具体实践

政务数据融合了民生和社会运营数据，涉及多种行业 and 多个业务部门，政务业务数据必然涉及到跨部门协同。通过数据安全隐私计算平台，融合政府和第三方机构数据进行公共数据密文开发共享、联合计算、联合建模，使数据在各业务线之间流通，打通跨域数据的应用价值链。

政务数据安全隐私计算是以汇集数据处理、数据共享、数据分析、密态运算、数据联合建模、可视化展现于一体的综合性数据共享生态平台。

数据处理：数据安全隐私计算平台独有专利技术在数据处理前期，保障上传数据统一性，并可以根据需求自动将数据转化为用于多方安全计划和联邦学习的数据信息，在后期进行加密传输时无缝衔接多种隐私计算技术的使用。

数据共享：数据平台概念，数据集市理念，去中心化部署方式，保障用户数据通过密码学相关隐私协议数据不出本地可以相互进行加密协同，安全共享。

数据分析：根据用户实际需求场景，对双方或多方数据进行融合并进行协同任务开展，整体过程全流程由数据发起方管控，不仅可以进行相应数据运算分析结果输出，还帮助用户未来对数据使用情况进行梳理和分析。

密态运算：通过同态加密、秘密共享、不经意传输、混淆电路等多种密码学和计算机科学组合的加密运算方式，对双方或多方数据进行加密协同运算，数据安全不出域并且可以获得正确计算结果。整体流程安全可靠，审批流程控制结果可见性，保证数据共享的过程和结果都安全安心。

联合建模：在单方机器学习基础上的联邦学习模型构建，可以让政务数据联合多机构数据构建联邦学习模型进行预测，也可以帮助企业安全引入了政务部门多方数据，丰富用户特征维度，服务企业挖掘目标客户，减少企业推广成本，提升企业纳税营收。

可视化：可视化界面 + 无代码或低代码操作，用户通过拖拽即可在短时间内完成多维度、全方位的联合建模，并以多元可视化方式展示预测结果。

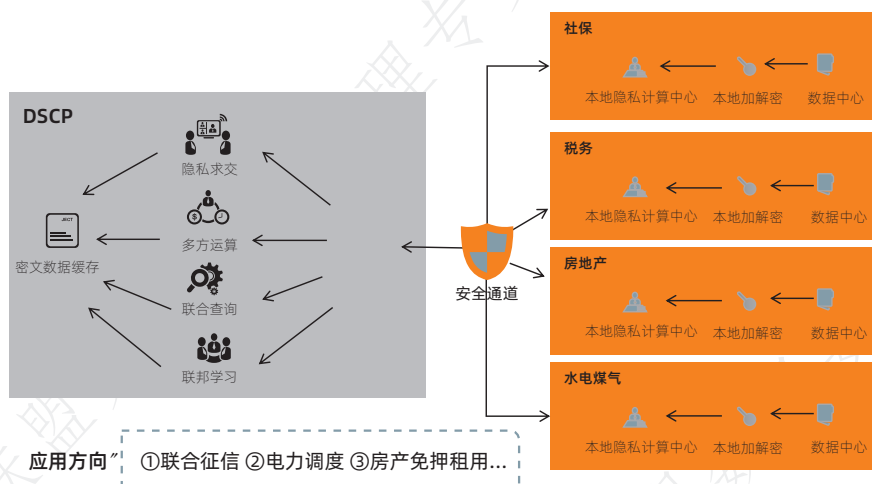


图 19 隐私计算技术应用场景

政务数据安全隐私计算平台部署在政务中心。国有企事业单位数据可以选择由政务中心数据部门统一管理、审批。明文数据不出所在单位，全程加密传输、缓存和运算，保护数据的安全和隐私。通过政务数据协同共享平台，既可以连接各行各业数据信息，综合使用政务服务，提升社会治理水平，又可以通过多方安全计算和联邦学习等隐私计算技术，搭建统一规范、互联互通、安全可控的政务数据开放服务共享平台，促进数据交易市场培育，促使数据流通价值最大化，安全合规的推动政务数据开放与共享。

5.1.2.4. 总结

大数据时代，云计算和大数据等技术为数字政府建设赋能。隐私计算时代的到来，政务数据安全建设也应借助隐私计算技术和数据安全专业化安全能力，实现各类公共信息和各机构数据的端到端联通共享和敏感信息保护，保障政务业务安全和合规的运行，为政务数据安全流通创造价值保驾护航。

因此隐私计算平台和实际政务业务的深度结合才能更好地实现政务数据的公开共享。在政务隐私计算平台上，我们可以通过各方数据的汇总、融合、分析、运算、建模，提高政务业务协同能力，真正落地实现政务数据互联互通，通过科学分析和决策优化工作方式、提高工作效率，提升人民的满意度。

5.1.3. 政务大数据中心数据流转安全监测实践

5.1.3.1. 背景

随着数字化转型加速，“数字政府”改革已进入深水区，国务院办公厅《关于印发政务信息系统整合共享实施方案的通知》（国办发〔2017〕39号文）要求，加快推进政务信息系统整合共享，加强政务信息资源收集、共享、使用的安全保障工作，加强统一数据共享交换平台安全防护，切实保障政务信息资源共享交换的数据安全。某市成为“数字政府”综合改革试点城市以来，就将政务网络和数据安全作为关乎数字政府改革成败的关键，针对集约化建设和一体化技术支撑平台对安全



防护提出的更高要求，从制度先行、制定标准、夯实防护技术保障等方面着手，筑牢安全防护基础，有效保障政务数据的安全。

政务数据作为国家基础性战略资源，事关国家政治经济运行、国防和社会稳定。随着国家数据共享交换平台的建设，政务数据流转日益频繁，呈现出新的特点：一是数据从有限流动到广泛流动变化；二是数据从特定安全域到跨不同安全域使用变化；三是数据从分散存储到汇聚集中存储变化；四是数据业务流程变长，接触人员变多。如何保障数据在共享与流动过程中的安全成为行业关注的重点。

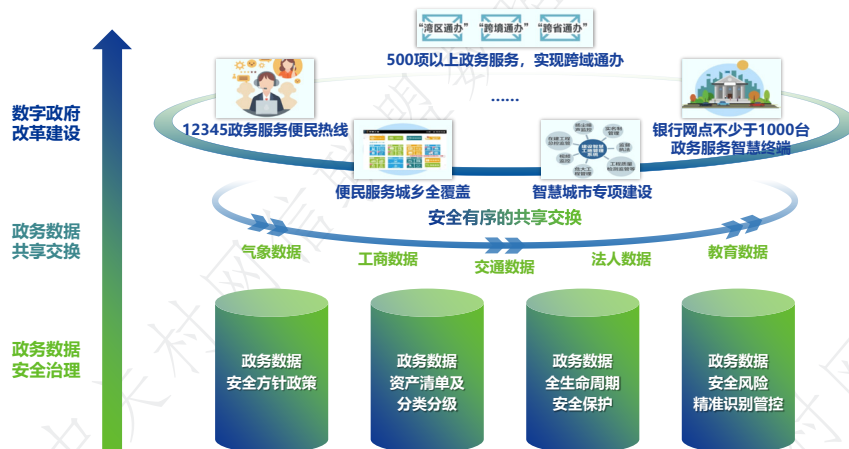


图 20 政务数据流转场景示意图

政务数据利用开发的高速发展、错综复杂的数据安全威胁、日益严峻的数据安全形势，对政务数据的安全治理提出了更高的要求。如何解决政务数据共享过程中数据状况难看清、数据风险难监测、数据泄密难溯源成为政府数据安全建设的重点关注内容。

5.1.3.2. 安全对策

以大数据和 AI 为核心技术，通过数据资产探查、数据流转监测、安全风险监测、泄密审计追溯等能力，实现“现状可知、安全可管、风险可视、泄露可溯”，让数据共享流转更安全，数据价值释放更充分。

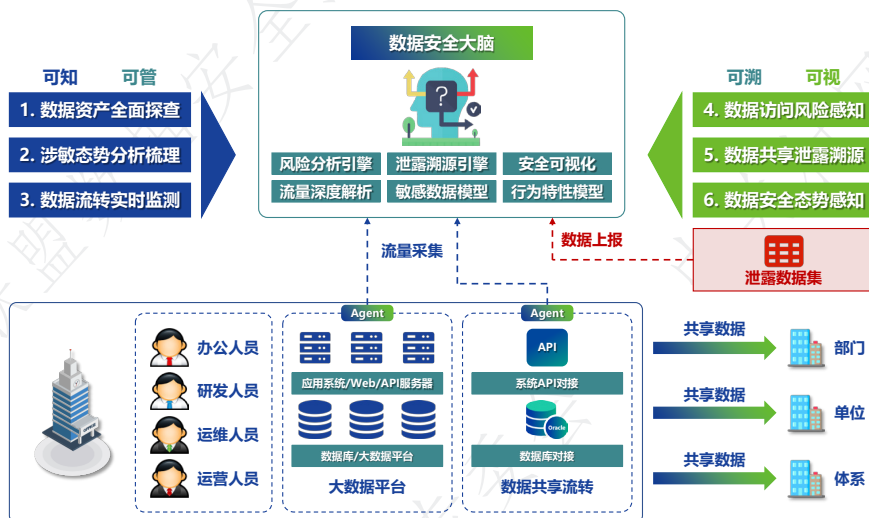


图 21 政务数据流转功能架构图



1) 数据资产探查与敏感数据梳理，数据全景心中有数

基于网络扫描发现和识别数据资产，采用定时任务自动发现分布在 Oracle、SQL Server、MySQL、Hive 等数据库服务中的数据资产。通过定义数据源，连接大数据平台中的数据仓库/数据湖，对数据库的字段进行全量深度扫描，遍历整个数据库，自动梳理数据资产清单，包括数据库、实例、表、字段，结合人工审核校验，最终确定数据资产清单，并定期监测数据库字段变化，自动更新数据资产清单。

采用内置的算法读取字段内容、自动提取数据特征，利用基于语义语法的上下文识别引擎、机器学习的聚类算法模型和基于关键词的多维带权匹配识别算法，建立数据字段与数据分类分级之间的映射，实现智能数据分类分级。通过 AI 算法学习已标识数据特征模型，自动为字段关联数据标签，出现自动关联出错时可人工进行干预，从而提高自动关联的准确率。

平台提供对外的 Rest API 接口，数据共享平台以及数据安全产品均可以通过 API 接口引用或者调用数据分类分级的结果，并且通过全量同步结构化的数据，可以保证新同步的全量数据也能自动打标，满足业务部门和安全部门在不同场景下对数据分类分级的时效性要求，解决了人工分类错误率较高、数据量增长带来人力与时间成本线性增加的难题，为精细化的数据安全管控提供支撑。

根据国家、行业和地方标准及业务需求，利用内置数据识别规则及模型，智能识别敏感数据。

通过对数据、应用和 API 等进行智能数据收集和自动化梳理，形成数据资产分布、敏感数据资产清单、敏感数据流转视图和数据权责清单，同时实时采集业务动态访问过程中的数据流量，对流量进行智能分析，生成涉敏数据集、涉敏应用集、数据流转视图等，做到对数据全景与流转“心中有数”。

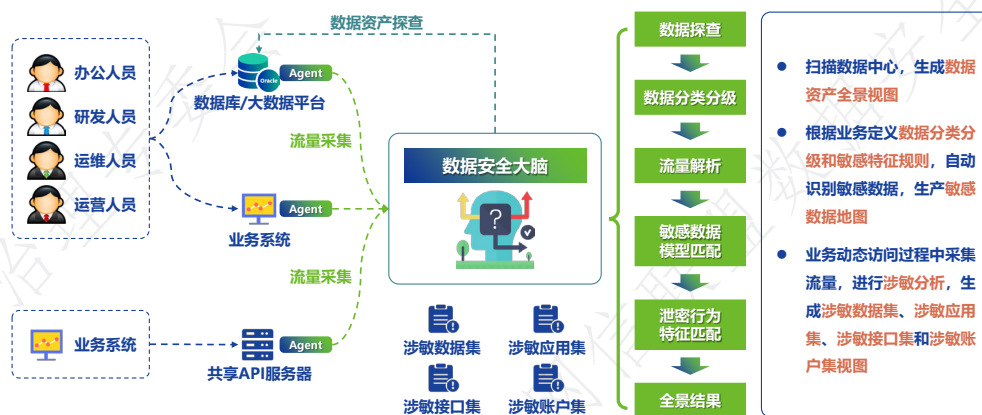


图 22 数据资产探查

2) 数据流转监测、风险实时感知

通过大数据与 UEBA 用户行为分析技术，对用户数据访问流量进行建模，自动生成安全基线，基于安全基线以及异常行为特征模型对数据访问行为进行研判，实时感知风险并告警，如：数据越权使用、API 异常调用、运维人员批量读取敏感数据等。

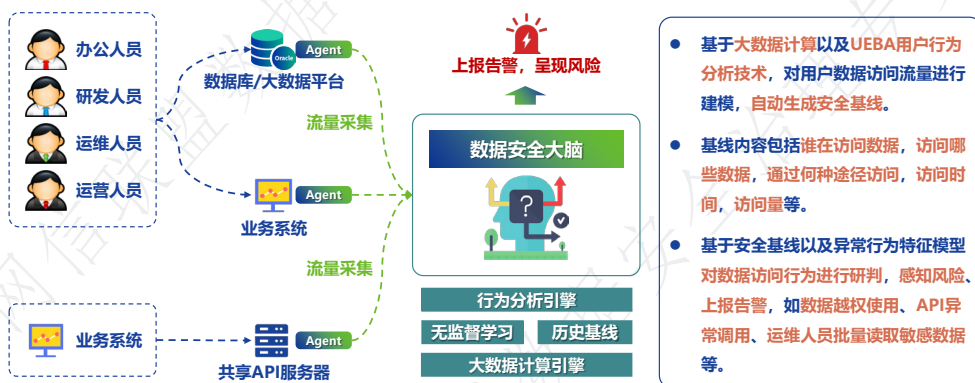


图 23 数据流转监测、风险感知

3) 数据共享交换泄露溯源, 责任界定有理有据

以AI技术为核心的可疑第三方检测模型(STP)算法,对共享交换过程中可能泄露的数据进行追溯,利用事件同源分析、样本同源分析、多种关联分析等自动化分析技术,快速定位可能的泄露源头,提升追踪溯源的效率和能力。

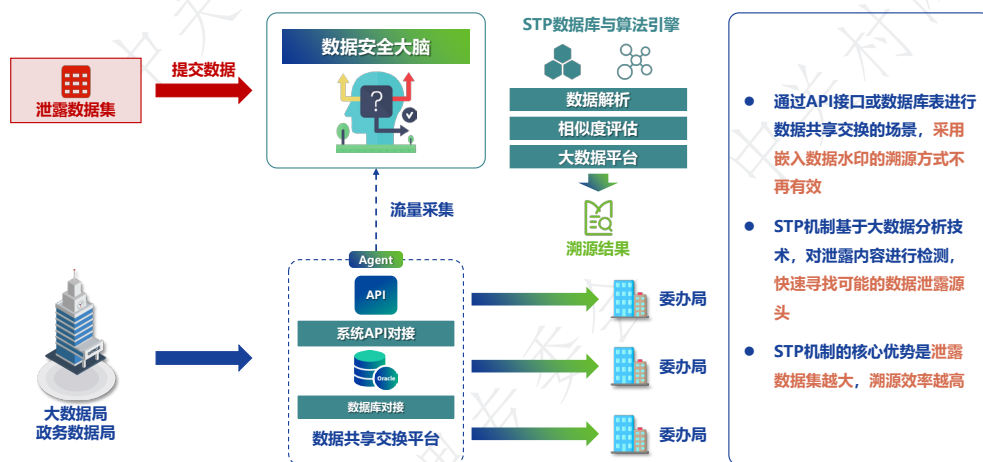


图 24 数据共享交换溯源

采集政务数据共享服务接口的流量,实现对于数据共享状态的实时监测和历史记录。当发生数据泄露之后,管理员通过将泄露数据导入到STP监测模型,实现对于泄露数据路径和时间的快速特征匹配和展示,辅助管理员实现泄露数据人员和机构的追溯和定责。

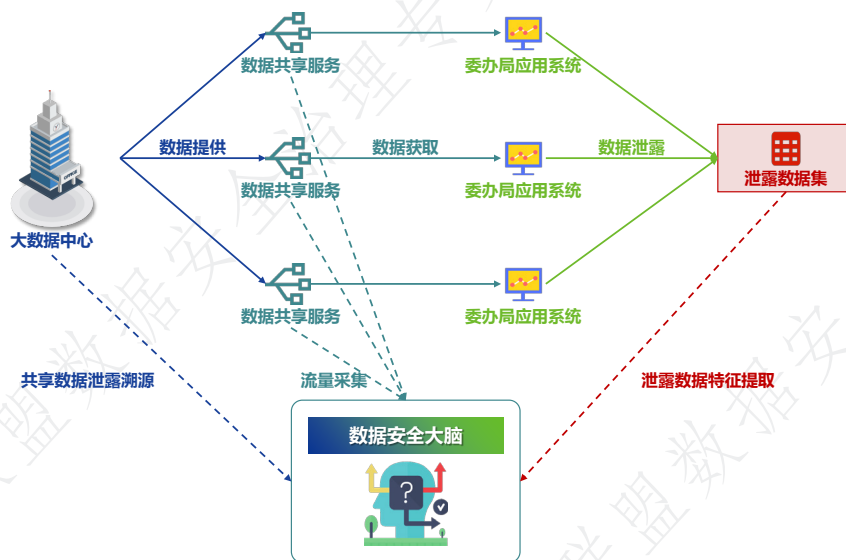


图 25 STP 检测模型示意图

STP 检测模型的工作方式：

(1) 流量采集代理通过在数据共享服务服务器上安装流量采集 agent 实现对于数据共享流量的采集，采集之后将流量发送到流量解析探针。(2) 流量解析探针将流量采集代理发送过来的流量进行解析，将原始的数据访问流量根据数据访问应用的类型解析成流量访问日志，并将流量访问日志发送到流量日志存储组件进行存储。(3) 流量日志存储基于自主研发的大数据存储平台，将全量的数据访问日志进行安全、高效的存储，能够提供稳定的日志存储和高效的日志索引。(4) 泄露溯源分析模块基于 STP 检测模型（即可疑第三方检测模型），将泄露的数据进行特征提取，并将提取的特征在存储的全量数据访问日志中通过高效的算法进行检索，获取到可疑的数据泄露路径和泄露时间，快速定位数据泄露的机构和人员。

4) 数据安全态势，全方位掌控

从数据资产、数据访问两种视角出发，对异常状态、异常行为进行监测预警，全面展示数据资产分布、数据流转可视、安全风险监测、异常行为分析、数据泄露溯源等，从而全方位掌控数据安全态势。



图 26 数据安全态势感知



5.1.3.3. 具体实践



图 27 实施路径图

政务数据安全治理规范数据处理活动，监控数据安全风险，推动数字政府改革建设工作新局面，整理包括四个阶段：

第一阶段、构筑安全合规基石

梳理法律政策要求，建立政务数据安全合规库，规范政务数据处理活动；编制《某市政务数据安全管理办法》，为政务数据安全定责立规。

第二阶段、厘清数据安全现状

梳理政务数据流转全景，厘清政务数据安全治理界面；定位数据流转重要节点、关键系统，形成数据资产清单，明确数据安全风险。

第三阶段、定义分类分级标准

基于数据资产清单和标准规范，定义政务数据分类分级标准；立足重要系统节点，定位重要业务数据和敏感个人信息，圈定保护重点。

第四阶段、管好交换流转风险

可视化展现重要系统节点数据交换和访问情况，发现数据安全隐患；溯源安全风险，数据安全事件前期有效处置。

5.1.3.4. 总结

某市通过政务数据安全治理体系的规划与实施，显著提升了政务信息资源共享交换的数据安全保障能力，已基本实现以数据安全法律法规为基础，以政务数据为中心，融合技术、管理和运营的数据安全治理模式，确保政务数据，这一支撑数字政府有效运转的关键要素，持续处于有效保护、合法利用状态，保障政务数据共享开放释放价值。

可知：落实数据分类分级，对数据进行自动化、智能化、动静结合的探查与梳理，生成敏感数据分布地图以及多维度涉敏信息集合，大大提升用户对重要数据的全局掌控力度。

可视：基于大数据、机器学习、UEBA 等技术构建数据安全访问模型，直观呈现数据流转过程的





全景路径，研判潜在的风险隐患，及时进行通报预警。

可控：以数据为中心，结合数据分类分级结果，实现精细化的身份管理、权限管控、数据库管控、数据防泄露等系列技术栈，构筑立体化数据安全防护体系，“进不来、拿不走、看不懂、抓得住”，保障数据共享交换生命周期的安全可控。

可溯：基于 AI 和大数据分析技术为核心的 STP 可疑第三方检测模型，不在原始数据中嵌入任何标记，不破坏原始数据，快速有效定位数据泄露源头，数据安全事件定责有据。

政务数据安全与国家网络安全、国家安全紧密相关。政务数据释放价值的同时面临业务风险和责任，以数据安全监测为核心的政务数据安全治理实践，在建立数据分类分级保护制度的基础上，智能追踪数据使用情况及流向，让数据用途明晰化、安全可视化，及时发现数据安全隐患，避免数据安全事件发生，确保政务数据持续处于有效保护、合法利用的状态，最大程度为社会治理赋能增效。该项目助力地方政府持续深化政务服务改革，为提高社会的治理效率、推进国家治理体系的建设、提升政府治理能力的现代化发展保驾护航。

5.2. 金融行业的数据安全治理

5.2.1. 某国有银行集中化安全审计应用实践

5.2.1.1. 背景

随着金融业务市场的规模不断扩大，数字化转型和升级提速，数据已成为金融行业基础设施，成为业务发展重要原动力，数据价值在的进一步凸显，也让数据安全面临更加严峻的挑战。近几年，国家和监管机构对数据安全的重视程度不断加强，国家层面《网络安全法》、《数据安全法》、《个人信息保护法》的数据安全合规监管要求密集颁布，行业层面《中国银保监会监管数据安全管理办法（试行）》、《金融科技发展规划（2022-2025）》陆续出台，这些国家和监管机构的法律法规密集颁布，对于数据存储、使用、运营提出了明确要求，如何更好的对数据进行有效防护，保障数据全生命周期的安全性，如何以事前发现、事中管控、事后审计、持续加固的方式，提供更好的服务成为当前环境下数据安全的重点工作。

行内数字化、国产化建设的工作快速推进，行内的数据库主机环境具有虚拟化程度高、网络环境复杂、数据流量大、数据库主机系统种类多、数量大等情况，如何在如此复杂的环境下建设数据库审计系统，安全部门面临着巨大的挑战：

- 缺乏敏感数据分类分级，在面对海量的审计结果时，无法精确识别不同等级敏感数据的访问行为。
- 传统的镜像流量的数据库审计方式，无法承载超大流量的实时解析工作，并且对于数据库本地运维操作无法进行审计。
- 对于一些数据库高危操作并不能及时进行告警，只能在安全事件发生后进行事后追溯，无法进行事中监测和告警。



5.2.1.2. 安全对策

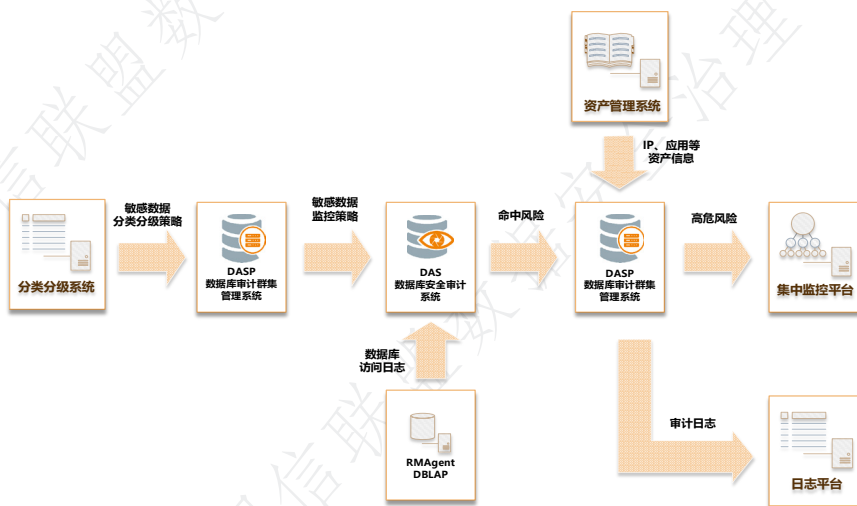


图 28 集中化安全审计处理流程图

为解决行内在数据审计管理中遇到的痛点，提升数据库的数据安全管理水平。行内构建了一套覆盖总行、分行和租户环境的数据库审计系统，构建事中监测、事后审计的数据安全防护体系。

1) 数据分类分级是安全防护的基石，结合规范，对行内所有的业务系统进行梳理，识别敏感数据，完成数据分类分级工作，为数据库审计提供分类分级策略支撑，并且将分类分级策略下发给数据库审计群集管理系统（DASP）。

2) 数据库审计群集管理系统（DASP）统一纳管所有的数据库安全审计系统，依据分类分级策略，统一制定数据库安全审计策略，并将策略统一下发给所有的数据库安全审计系统（DAS）。

3) 旁路部署在各数据库主机上的数据库探针（RMAgent）和数据库插件（DBLAP）将审计到的数据库网络访问和本地操作日志上传给自己对应的数据库安全审计系统（DAS）。

4) 数据库安全审计系统（DAS）根据敏感数据访问监控策略进行风险监控，并将命中的风险上报数据库审计群集管理系统（DASP）。

5) 数据库审计群集管理系统（DASP）与资产管理系统对接，确认涉敏应用。将审计日志上报日志平台备查，同时将高位风险上报集中监控平台，由平台统一管理和处置。

5.2.1.3. 具体实践

1) 部署架构

针对行内数据库访问流量大，网络环境复杂等现状，审计系统采用分布式轻量插件方式实现旁路部署。通过在宿主机上部署网络探针（RMAgent），对于网络访问进行全面审计，通过在容器中部署数据插件（DBLAP），实现运维人员本地数据库操作审计，从而全方位的实现数据库审计。并且 Agent 和 DBLap 通过智能动态防御机制，在数据库主机处于高耗能情况下，释放资源保障数据库业务的正常运行。



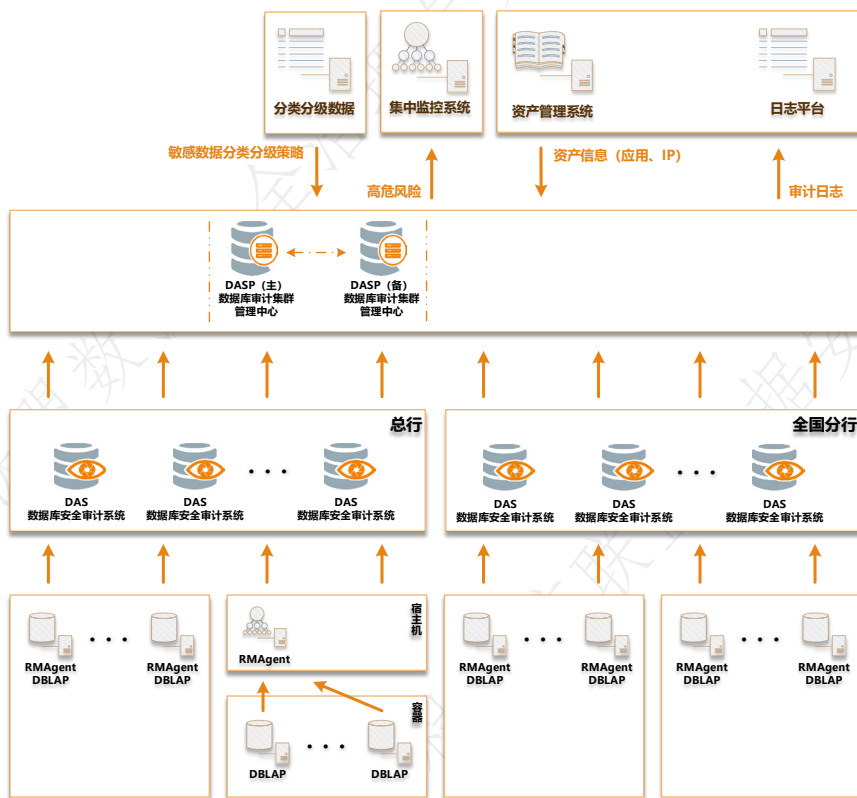


图 29 集中化安全审计部署示意图

2) 治理实践

基于分类分级系统产生的数据分类分级结果，行内形成了一套敏感数据资源库。根据敏感数据资源库中不同级别的敏感数据，数据库安全审计系统制定了不同的安全防护策略，通过这些策略，完成针对不同用户、不同应用访问的监测，及时发现涉敏应用或用户。



图 30 数据安全集中化审计可视化

5.2.1.4. 总结

通过系统建设，目前已经完成第一阶段和第二阶段的工作，覆盖总行上万数据库系统，并且与





行内数十万敏感数据对接，完成从数据资产梳理、分类分级策略下发、数据库风险识别和处置的全周期闭环监控，实现对数据库操作、访问用户和访问应用的全量审计：

- 实现数据库审计统筹管理，以往开发平台日志不全、生产故障无法分析以及日常运维审计的问题彻底改变。

- 成功解决容器环境本地运维日志不记录、堡垒机操作无法成功溯源等业界难题。

- 实现数据从分类分级、策略制定、访问监测、集中监控管理的全周期闭环管理。

接下来，将按照计划完成分行和租户环境的部署，全面覆盖行内所有数据库系统，实现全面防护。

5.2.2. 某城商行大数据中心安全运维实践

5.2.2.1. 背景

某农商行的业务数据具有高敏感性特点，数据一旦被泄露、滥用带来的风险和损失难以估量，数据安全尤为重要。随着数字化转型升级，前端生产数据集中整合到大数据平台，并通过下游大数据创新应用进一步加工处理和挖掘分析，数据的流动过程涉及到的场景复杂、角色繁多，给数据保护带来了新的挑战。因此，某农商行早期即对数据安全建设十分重视，并在行内已采取一定技术措施保护数据安全。

首先、现场已对数据库用户权限进行管控，核心生产库的运维账户权限默认只授予查询权限，但这些运维人员可以查询到数据库中的真实数据，存在数据泄露的风险；其次、现有的堡垒机作为运维行为管控设备，无法对具体 SQL 操作进行限制，若出现误操作或恶意操作等情况，无法及时拦截阻断，可能造成不可估量的损失。

因此，针对大数据系统安全运维场景，还需进一步提升数据资产管理水平、强化数据安全防护细粒度、增强数据安全集中管控能力。解决某农商行数据中心在数据开发、运维过程中遇到的数据安全痛点及难点，构建围绕敏感数据的分类分级动态运维安全防护体系，预防敏感数据在开发、运维场景下因外部攻击和内部攻击带来的盗取、泄露、滥用带来的内部安全风险。

5.2.2.2. 安全对策

本项目针对研发侧、运维侧不同的数据使用场景提供事前审批、事中管控、事后审计以及敏感数据脱敏的一整套安全防护方案。

传统依靠孤立的产品技术手段，难以对敏感数据开发工作和大数据系统运维工作的安全开展形成系统化的支撑。某农商银行大数据系统安全运维管控平台的建设，以数据为中心，参照 I（风险识别）P（安全防御）D（安全检测）R（安全响应）模型，实现对数据库安全评估系统、数据库运维管控系统、数据库脱敏系统和数据库安全审计等单点安全产品的集中管控，汇聚各安全产品的安全信息进行关联分析，统一下发安全管控策略，构建一套从数据资产识别、安全防护、风险监测、响应和处置事件、到策略优化的大数据运维闭环管控系统，覆盖大数据系统运维场景的各个环节，增强行内数据安全集中管控能力。平台结合数据安全相关管理制度，制定规范化的运维操作审批流程，结合数据分类分级结果，将防护策略细粒度的施加到库表及字段级别，做到在同一平台上将安全策略、



审批、管控和追责有效结合，形成了覆盖事前预防、事中监测和事后审计的整体大数据系统安全运维防护体系。

针对某农商银行的敏感数据安全开发和大数据系统运维安全运维需求，可通过某农商银行大数据系统安全运维管控平台，建立以下对策，以应对在开发和运维工作中面临的运维接入终端安全风险、越权访问风险、高危操作指令误操作风险、敏感数据伪脱敏风险，以及从内部直接针对大数据系统发起的攻击风险。

1) 运维账户统一接入

对运维账户接入的 IP、连接工具、账户进行认证，规范运维人员的运维接入行为。

2) 数据库操作安全防护

根据内置的安全规则库，有效阻断 SQL 注入等攻击行为，防止数据泄露。

3) 建立细粒度的运维账户权限访问控制策略

通过对大数据系统中的数据资产分类分级，及对数据的流转情况进行梳理，建立可细化到字段级的主客体访问控制策略。

4) 基于运维账户身份的高敏感数据查询动态脱敏

基于数据库账户的运维访问权限只能控制到表级别，一旦授予账户运维权限，运维账户对于表中存储的所有高级别数据均具有访问权限，平台可设置动态脱敏规则，针对不同的运维账户身份，动态遮蔽查询返回的高级别数据内容，防止某农商行核心数据通过水滴式收集和屏幕拍摄的方式大量泄露。

5) 定义高危操作指令审批流程

对执行 Drop、Update 等高危操作指令的操作，先进入审批流程，得到审批后，指令下发执行。

6) 实现生产环境数据到加工、测试环境的高效脱敏

根据业务需求，自定义数据脱敏规则，批量导出数据到加工、测试环境。

7) 集中风险监测与处置

对大数据系统安全运维操作进行全面审计，有效识别暴力登录、拖库、越权访问、数据篡改、网络攻击等风险事件，并进行事件处置情况跟踪。

5.2.2.3. 具体实践

1) 设备部署方案

本次项目部署的集中管理平台、数据库安全审计系统、数据库运维管控系统和数据脱敏系统均采用旁路部署方式，通过代理访问的方式进行安全运维访问控制。

其中，集中管理平台部署在生产环境，对本方案其他安全系统进行集中管理。数据库安全审计系统和数据库运维管控系统部署在运维区，管理和审计运维行为。数据脱敏系统部署在开发环境，从生产环境采集管理员授权脱敏的业务数据，并对数据脱敏后导出到加工、测试系统。系统部署示意如下图所示：



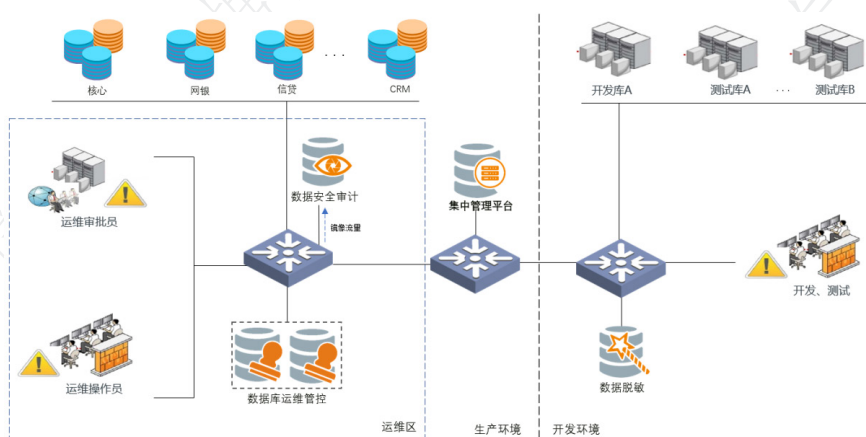


图 31 大数据系统安全运维部署图

2) 项目实施

项目实施共分为五个步骤：

第一步、摸清数据资产家底，快速形成数据分类分级清单

对基金系统、理财系统、信贷风险预警等系统进行数据分类分级，快速形成“数据资产清单”，3 级数据将作为核心业务数据资产成为运维管控和数据脱敏管控策略中的重点管控对象。数据分类分级步骤如下：（1）依据《金融数据安全数据安全分级指南》，结合某农商行实际业务，制定数据分类分级知识库；（2）系统负责人梳理字段元数据描述信息；（3）系统对知识库和描述信息进行智能匹配，自动分类分级；（4）通过拆词、分布分析等智能手段对未形成唯一分类的数据进行智能分类；（5）人工分类仍未形成唯一分类的数据，并全面核查数据分类分级结果。

第二步、全链路业务数据流转监测，形成业务资源备案，识别数据流转中的安全风险

通过数据库审计日志分析，智能梳理基于字段级的数据访问主客体关系，形成“IP- 运维工具 - 运维账户 - 数据库（表）- 字段”的运维业务资源备案，为实施数据运维账户到数据库字段的主客体访问控制策略，遵循必要、最小原则，提供基础依据。

第三步、实施运维过程中针对数据的安全防护策略

基于分类分级结果和业务资源备案信息，对基金系统、理财系统、网银、柜面 F 等 40 多套系统实施运维防护策略。安全策略包括：

- 细粒度访问控制规则：根据业务资源备案清单中的主客体访问关系制定细粒度的用户访问控制策略；
- 规范账户登录行为：针对运维账户设置登录频率、IP、登录时间等行为控制策略；
- 敏感数据动态遮蔽：针对运维账户身份设置 3 级数据返回动态遮蔽策略；
- 高危操作审批：对所有高危操作（如：Drop、Update 语句）进行上级审批流程定义；
- 数据脱敏：实施高敏感数据批量导出到加工、测试场景的安全策略，策略包括数据源、导出目的、脱敏规则等，可实现自动对策略覆盖的增量数据进行脱敏，并导出到加工、测试环境；
- 加载风险规则库：加载高危风险识别规则库，识别 SQL 注入、脚本攻击等风险行为。

第四步、形成运维场景数据安全态势感知，发现运维场景中的攻击行为。

大数据系统运维系统实时收集各个安全组件上报的安全信息和事件信息，对这些信息进行关联和分析，形成面向大数据系统运维场景的整体数据安全态势感知平台，全面掌握敏感数据分布情况、业务资源备案情况、监测数据流转情况，并集中展示安全风险事件。

通过整体数据安全态势感知，可识别运维账户违规登录、禁用账户成功登录、多人共用账号、未通过堡垒机访问数据库，批量查询数据、越权查询数据、越权篡改数据、非变更时段执行操作指令，以及口令爆破、拖库、SQL 注入语句、超长语句、嵌套型语句等外部攻击行为。

第五步、集中响应和事件处置

通过大数据运维系统各个安全组件的集中管理，统一下发安全策略，快速实现对安全事件的响应和处置。管理人员可以通过系统持续跟踪事件处置情况和事件处置趋势，并通过事件稽核 KPI 指标督促运营人员提高事件处置效率。

3) 实施效果

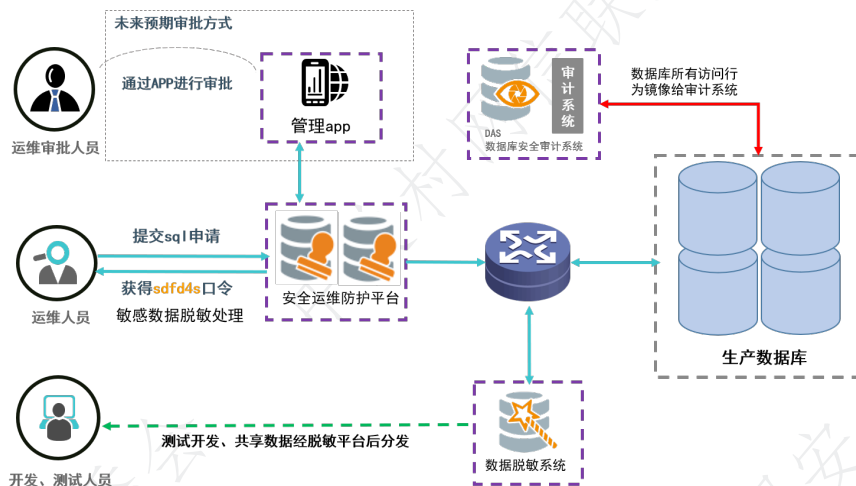


图 32 大数据系统安全运维效果图

● 数据库运维管控

➤ 运维人员把需要访问的生产库的 SQL 操作提交给运维审批人员；运维审批人员通过管理 app 对运维人员的操作申请进行审批，审批通过后，对应运维人员方可获取相应的 code 码；

➤ 在整个运维过程中均基于运维账户实现敏感数据动态脱敏，确保敏感数据不会被运维人员越权访问；

➤ 在运维操作结束后运维权限随时终止。运维人员所有的 SQL 操作都会被运维管控设备记录，方便后期追溯；

➤ 在周、月、季度等时间节点生成特定的运维报表，方便运维用户统计分析。

● 数据脱敏

➤ 自动发现敏感数据并进行核实；

➤ 添加脱敏任务，为核实后的敏感数据匹配对应的脱敏算法进行脱敏处理，之后将处理好的数据提供给研发测试使用；

➤ 添加定时任务以免去用户的重复操作；



- 定期统计并生成任务报表。

- 数据库安全审计

- 开启自学习建模功能，并在学习期结束后确认学习结果并形成数据库模型；

- 审计系统开启数据库自动发现功能，自动发现镜像流量中的数据资源并自动添加审计，省去用户手工添加数据的人力、时间成本与可能存在的误操作等风险；

- 审计系统中针对不同数据库配置相应的告警策略，发现风险行为及时告知用户；

- 开启虚拟补丁功能，系统内置已知数据库漏洞和风险规则库，一旦数据库访问行为匹配到这些漏洞将及时告警；

- 定期生成专项报表，满足合规要求。

5.2.2.4. 总结

通过项目的实施，我行建立了数据分类分级长效机制，为大数据系统安全运维场景实现了运维账户集中接入、高危操作行为阻断、异常用户行为阻断、高危操作指令审批、3级数据查询返回动态遮蔽等安全防护能力，在生产数据导入加工、测试场景实现了在保留数据原始分析价值的前提下，对数据进行高效、高仿真的批量脱敏，解决了通过手工脱敏数据存在的效率低下及数据伪脱敏问题。

大数据系统安全运维系统的应用，有效降低了某农商行数据资产管理成本和数据安全事件运营成本，提高了数据安全事件响应和处置效率，最大限度降低了数据安全事件造成的经济损失。同时，有效预防数据泄露事件和恶意数据破坏事件的发生，避免给社会和企业造成不良的社会影响。

5.2.3. 某国有银行数据防泄露实践

5.2.3.1. 背景

大数据时代，数据已经成为企业的核心资产。在金融行业中，除了自身的商业隐密数据以外，还有保护客户隐私信息等需求，保护数据安全已经成为安全建设的重要挑战。然而现代企业中一方面存在海量文档，如何精准的定位关键数据，并加强敏感数据在创建、存储、使用、传输和销毁全生命周期的安全管理；另一方面，如何更好地监控和保障已上线及新上线的信息系统运行，及时识别和防范安全风险，同时满足国家网络安全法和监管要求，保证信息安全管理工作的依法合规；此外，如何对这些文档信息数据进行安全防护，确保不被外部入侵违法获得，或内部违规访问外泄，也是数据安全过程中面临的重要挑战。

5.2.3.2. 安全对策

为防范和控制银行商业秘密和国家秘密通过终端以及网络泄露的风险，XX银行发起数据防泄露系统建设项目。通过数据防泄露项目的建设，则可有效地解决用户的“有意”、“无意”数据泄露行为。

数据防泄露项目首先要解决关键数据的定位问题，所以需要以数据内容为核心，进行识别和分类分级；通过数据分类分级形成的标准与策略，及时对外发网络流量和邮件进行管控，根据敏感数据类别和级别的不同，提供审计、提醒、审批等不同级别的防护；并为管理员提供审计和统计报表，



监测到的事件可形成多维度的可视化报表，形成数据安全态势分析图表，及时有效发现针对数据的攻击和威胁。

网络防泄露系统是通过关键字、自然语言识别等技术对网络外发的数据进行分类分级的管控，响应策略包括监控、通知（短信和邮件）、阻断等方式。同时，与银行内审批系统结合，实现重要数据的审批外发。

终端数据防泄露系统是通过文档信息分类分级方法，扫描磁盘上存储的文档，对文档进行定级，采用加密和权限管理的方法，对文档实施有重点的保护，控制重要文档的使用权限和范围。

最后，通过文档追溯技术，结合系统中的日志，对文档的创建、存储、使用、交换、销毁等生命周期进行追溯，形成一套较为全面、系统的数据安全管理方法，以支持银行高效、安全的管理好内部重要文档信息。

5.2.3.3. 具体实践

1) 终端数据防泄露系统建设内容

终端数据防泄露系统建设充分考虑采用先进而成熟的技术、完善的体系结构，既保证实用成熟，又能够适应未来的业务发展和技术的更新要求。

● 敏感文件扫描

通过设置灵活的扫描策略，按照指定的关键字及敏感信息指纹进行终端文件扫描；

● 文件加解密功能

通过设置灵活的策略实现对办公终端或移动存储内文件进行自动加密、主动加解密、主动标注密级等；

● 打印文件加标水印功能

对于指定用户或用户组终端中标注为指定密级的文件，在打印时自动加标水印。水印可以通过普通水印或二维码等方式实现，支持水印的打印位置、透明度等参数的灵活设置，同时也支持记录、追踪指定密级文件的打印行为。

● 策略离线支持

指定用户或用户组办公终端在未接入银行内网络情况下，指定策略仍生效可用。

● 日志审计

终端数据防泄露系统客户端和管理员操作所产生的日志均可以支持自动保存在服务器上，提供日志查询和统计报表管理。

2) 网络数据防泄露系统建设内容

网络数据防泄露监测设备实现对行内邮箱外发行为的监测功能，设备支持负载均衡和热备，以保证监测的连续性。

● 内容识别

网络数据防泄露监测设备支持关键字内容识别、词典模糊内容识别、正则表达式内容识别，通过过后处理脚本对发现的敏感数据进行进一步校验，提高准确率；对文件可以指纹内容识别、自然语

言特征匹配、图片内容文字识别，可识别图片格式的文字内容。

● 文件识别

网络数据防泄露监测设备支持识别常规文件格式、多层压缩文件、多层嵌套文件（如 Word 里嵌套 Excel）；可根据文件特征识别文件类型和文件内容；同时也可识别被密码保护的常规文件（如 Office 2003–2013、RAR、ZIP、PDF 文档等）；支持识别单个大容量文件（如文件大小为 500M 以上）；支持 Unicode、GB18030、GBK 中文编码格式。

● 邮件及网络协议识别

对邮件标题、正文、附件进行内容识别；对 126、163、Sohu、Sogou、Sina、Tom、21cn 等常见邮件（传输协议非加密）的信息识别；对 HTTP、SMTP、FTP、POP3 协议发送敏感数据发现；可以通过网络全流量抓取数据审计，可抓取 IP 范围内的所有上行数据。

3) 综合管理平台建设内容

● 综合管理系统为终端数据防泄露系统、网络数据防泄露系统和终端安全管理系统提供统一用户服务，统一用户服务主要包括用户认证和用户同步。

● 综合管理系统通过同步银行统一用户源完成全量和增量用户同步。

● 综合管理系统为各级管理员提供相应身份的组织机构管理、用户管理、审计员设置、管理员设置、统一网管和管理员日志相关接口服务。

4) 系统部署架构

充分考虑总行、分行、辖属行、部门的管理模式，支持策略分级管理、报表分级管理、角色分级管理。管理角色分系统管理员、策略管理员、安全审计员，不同管理员功能各不相同、身份互斥。终端安装配合已部署的终端安全管理系统推送安全客户端，实现全面、快速的客户端部署。总行管理员可以对全行制定全局的安全策略，各分行管理员制定和管理分行范围内用户的策略，实现数据防泄露系统的统一认证、统一管理和统一运维。

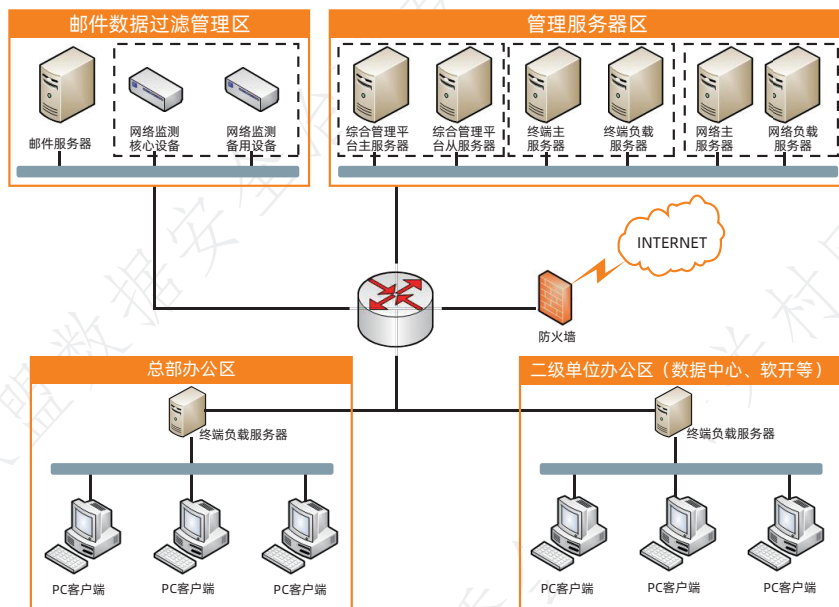


图 33 XX 银行一期部署示意图



为实现集中统一化的管理，终端数据防泄露系统分别部署一台主服务器、一台热备服务器及多台负载服务器；网络数据防泄露系统和综合管理平台分别部署一台主服务器、一台热备服务器，组成数据防泄露系统的服务器区域，确保冗余和高可用性，满足本期项目需要及后续扩容；由综合管理平台统一提供多级管理模式，实现对各分行的分级管理（部署方式可调整，可以针对各分行用户数、应用和需求等情况，进行灵活调整）。

5.2.3.4. 总结

“源头管控、边界防护、审计溯源”的理念紧密贴合用户需求，兼顾了安全性和应用便利性，对用户组织业务流程基本无影响，用户接受程度高。终端数据防泄露、邮件数据防泄露和综合管理平台整体规划，成体系建设，应用效果更好，全面满足银行数据安全主动管控需要；利用集中管控平台，实现各功能模块统一管控，方便行内运维；先进的密钥管理体系、实用的透明加解密技术、一文一密防暴力破解等特性，确保系统的安全性更强、对用户影响程度更低。

5.3. 能源行业的数据安全治理

5.3.1. 某电网公司数据安全治理实践

5.3.1.1. 背景

某电网公司提出要全力打造“三型两网，世界一流”的企业，随着泛在电力物联网的建设，某电网公司信息化建设必然打破传统的数据孤岛，转而走向共享、开放。因此，某电网公司大数据将呈现日益活跃的“流动”趋势，在“流动”中发挥价值，必然驱动数据的有序流动、合理利用和安全分享。

但是，在数据“流动”的过程中，存在诸多安全隐患问题：对个体而言，特别关心数据的隐私泄露问题；对主管部门来说，则关心数据是否“健康”，也即数据是否真实、完整、可信，关心数据使用过程中是否会被批量泄露和篡改。数据安全治理，已经成为了某电网公司业务部门与安全部门共同的核心关切点。

5.3.1.2. 目标和挑战

随着某电网公司智能电网技术的深化应用和“三集五大”管理体系的全面建设，信息系统已全面融入公司生产、经营、管理、业务的各个方面，积累了大量的结构化数据、非结构化数据、海量历史等实时数据和地理信息数据。



图 34 某电网数据安全挑战

数据的安全性已经迫在眉睫，需要应对和解决的挑战主要有：

1) 数据繁多复杂

随着某电网公司“业务贯通，数据共享”的发展思路逐渐深入落实，国网公司各单位的数据体量越来越大，并且数据的种类越来越多，数据管理部门很难清晰、系统地认识当前的数据现状，无法为数据的安全管理提出有效且可行的保障措施。因此，无论在安全层面还是在管理层面都给国网公司提出了巨大挑战。

2) 数据集中存储风险

当前某电网公司的数据呈现集中化存储趋势，而在数据的存储层当前缺乏有效的保护措施，在数据库中重要敏感的数据都是以明文的方式存储。一旦黑客通过攻击手段形成“脱库”，或者后台运维人员通过权限滥用将数据文件拷走，都将直接形成数据的全部泄露。

3) 第三方服务风险

某电网公司的业务系统繁多，系统的运维压力巨大，因此国网公司众多业务系统的运行维护大多交由第三方服务商。而第三方服务商本身在安全管理方面普遍偏弱，但第三方服务商却掌管着系统中重要数据的访问权限，在利益的驱使下，存在数据被第三方服务商误用、滥用甚至泄露的风险。

4) 共享与分发风险

随着“业务贯通，数据共享”的发展思路逐渐深入落实，某电网公司会将大量的业务数据、个人隐私数据进行部门间共享，及面向外部机构的数据分发。如何平衡数据共享、分发与数据安全之间的利与弊，已成为某电网公司进行数据安全治理工作的核心环节。

5) 新技术风险

随着某电网公司全业务统一数据中心及大数据中心的建立，更多的大数据技术已经得到应用，而 Hadoop 体系中的数据存储组件，如 HBase、Hive、Mongodb 这些 NoSQL 数据库，其自身的安全性相对薄弱，因此针对 NoSQL 数据库这类新技术的安全管理也是国网公司面临的挑战之一。

6) 政策性要求

某电网公司作为国内重要的能源企业，自身也要遵循国家层面、行业层面、监管层面、公司层面各项关于数据安全的政策法规。

5.3.1.3. 治理实践

1) 建立组织保障体系

某电网公司的数据安全治理从组织保障层面开始入手，形成“三级垂直组织管理保障”体系（如下图所示）：

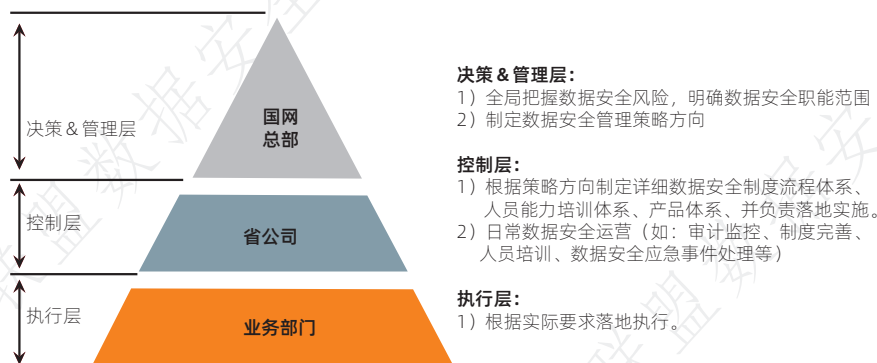


图 35 组织保障体系

第一级，某电网公司总部安全部门牵头，负责制定数据安全总体方针、重大政策和重大事项，并提出了“谁主管谁负责，谁运行谁负责，谁使用谁负责，管业务必须管安全”的总体原则。

第二级，由省公司牵头，以总部的数据安全方针为指引，结合本省公司的数据安全现状，制定数据安全治理可落地的技术与管理办法，建立具体的数据安全治理运营体系。

第三级，省公司业务部门牵头，负责具体技术与管理措施的落地。

2) 建设管理制度体系

完成组织层面的保障体系建设后，数据安全治理工作继续在管理制度层面持续推进，制定出一系列数据安全相关的管理制度：

- 《关于加强对外提供数据规范管理的指导意见》

重点提出了在面向某电网公司外部组织提供数据时的流程规范，以及提供数据应采取具体技术措施的细则。

- 《某电网公司客服服务中心数据导出实施细则》

重点提出了客服中心导出敏感数据并提供给外部组织时的业务流程以及应遵循的管理要求。

- 《某电网公司关于进一步加强数据安全工作的通知》某电网公司信通【2017】65号

重点提出了某电网公司在全生命周期过程中应采取的技术措施以及管理办法。

- 《国网营销部关于加强营销专业网络与信息安全管理的工作意见》营销综【2017】4号

重点提出了国网营销系统应提供加密、脱敏、访问控制等管控措施以及具体要求。

- 《营销专业客户敏感信息脱敏规范》营销综【2017】65号

重点提出了国网营销系统的脱敏要求，具体到营销系统客户敏感信息6大类23子项数据类型的划分依据，以及具体的脱敏方法及规则要求。

3) 落实技术手段保障体系

- 定期进行数据安全督查，发掘数据安全隐患

某电网公司各省公司安全督查相关部门定期对公司内部各信息系统、数据库系统进行安全督查，



通过专业的检查工具对数据库系统的安全漏洞以及安全配置基线进行安全检查，并将检查结果通报被检单位。被检单位将根据检查结果进行安全漏洞的修复以及安全配置调整，并接受复查。

- 构建主动数据防护机制，抵御非法窃取数据

根据业务部门以及安全部门需求调研，明确其业务开展必需的数据访问权限，形成《数据访问权限矩阵》，通过对重要数据建立增强级访问控制，针对业务访问、运维访问环节实施精准防护，针对不同角色用户在访问数据对象时进行不同力度的安全防护，并细化到数据表、字段级。

- 业务数据访问管控，敏感数据细粒度防护

针对某电网公司营销系统，国网营销部发文（营销综【2017】65号）要求针对前台业务人员访问以及运维人员访问营销系统客户敏感数据进行脱敏防护。通过对业务人员数据以及客户敏感数据梳理，将业务人员分为不同等级以及不同的角色，按照“数据访问最小化原则”，在不影响业务正常开展的前提下，最大化实施客户敏感数据脱敏。

- 强化数据操作行为审计，落实数据访问稽核

某电网公司全面推广实施数据库审计，在针对核心业务系统的数据库操作行为进行精准记录的同时，进行了数据库审计深化应用，推广并实施了数据库审计典设，针对敏感数据操作、攻击行为检测、违规操作检测、异常登录检测等典型场景进行了细化风险分析，达到了实时洞悉数据库安全风险的目标。

5.3.1.4. 实践总结

近年来，某电网公司持续开展数据安全治理工作，不断加强并深化在“组织保障、制度保障、技术保障”三方面的工作，并在实践中发现不足、弥补不足。加强智能电网的深入应用以及泛在电力物联网的建设，对某电网公司而言是机遇亦是挑战。但可以肯定的是，数据的安全问题将愈加凸显，而贯彻数据安全治理工作思路势必会将某电网公司的整体数据安全水平提升到一个新的台阶。

5.3.2. 某能源云平台数据安全密码应用实践

5.3.2.1. 背景

国家密码主管部门对密码应用，特别是国产密码算法的推广应用多次下发文件要求：2015年中办、国办发布《关于加强重要领域密码应用的指导意见》（厅字2015.4号），要求各行业部门梳理密码应用需求并制定《行业规划》；2016年中央办公厅下发《44号文》，要求为《行业规划》制定任务分工落实相关工作；2017年能源局印发《石油石化行业国产密码应用规划》，明确中石油、中石化、中海油需建设统一密码服务体系，推广国产密码应用等，进一步推动石油石化行业对国产密码应用落实；2020年1月1日《密码法》开始施行，要求使用商用密码进行保护的关键信息基础设施，其运营者应当使用商用密码进行保护。

在能源行业，目前广泛采用云计算技术、云架构，通过计算资源虚拟化、存储资源虚拟化、网络虚拟化等技术，形成虚拟资源池，实现物理和虚拟化资源管理和调度，建设能源云平台，为各种能源行业业务系统提供计算服务、存储服务等基础设施、网络、平台服务。能源云平台应用目前存



在密码应用高风险问题，具体表现在：在机房内部，云平台运维管理员使用内部网络访问安全管理运维区，使用用户名+口令登录运维堡垒机；系统用户使用内部网络访问安全运维管理区，使用用户名+口令的方式登录云平台的管理系统；在机房外部访问云平台时，云平台系统运维人员使用HTTP连接进入互联网和公共网的安全管理运维区。

5.3.2.2. 安全对策

根据密码应用工作要求，现需要对能源云平台密码应用进行改造，助推云平台建设符合信息化主管部门要求，保障能源云平台业务开展的安全性、合法性、合规性。

按照《信息安全技术信息系统密码应用基本要求》（GB/T39786-2021）的要求和云平台及云上租户系统的实际密码应用需求，采取符合国家密码管理政策和标准规范的密码进行保护，做到“同步规划、同步建设、同步运行、定期评估”，通过部署一系列符合国家密码管理局要求的商用密码产品搭建密码服务平台，并合理有效完成商用密码应用集成建设，实现身份鉴别、数据完整性、数据机密性等安全应用功能，来实现云平台上重要数据全生命周期安全。涉及到的重要数据和安全需求如下表所示：

序号	数据	所属应用	存储位置	安全需求
1	系统用户鉴别数据（如口令）	XX 市政务云平台	平台数据库	传输和存储机密性
2	快照、镜像数据	XX 市政务云平台	平台虚拟机	存储完整性
3	日志审计数据	XX 市政务云平台	平台日志数据库	存储完整性
4	关键配置数据	XX 市政务云平台	平台数据库	存储完整性

1) 密码应用技术框架

根据《信息安全技术信息系统密码应用基本要求》（GB/T39786-2021）、GM/T0115《信息系统密码应用测评要求》，结合能源领域基于云架构而设计这一实际情况，通过采用符合国家相关技术标准的专用密码产品和服务，搭建密码服务平台，为云平台、云上租户系统、以及云平台的下三层提供密码支撑能力，保障系统用户身份真实性、数据机密性、数据完整性、不可否认性。系统的密码应用总体技术框架如下图所示：

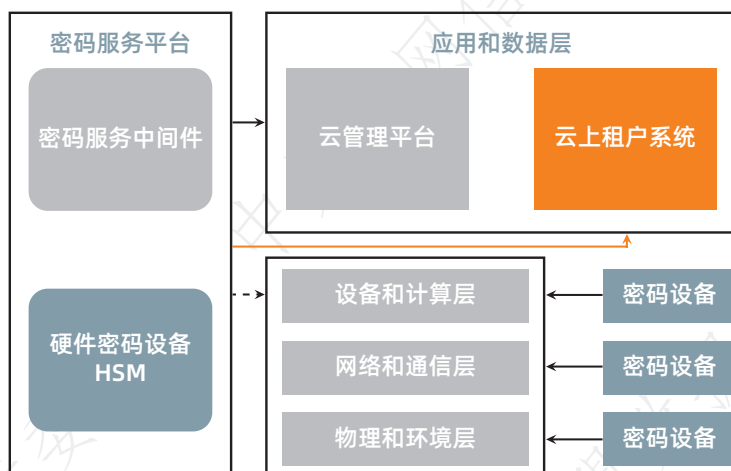


图 36 密码应用技术框架图



- 通过密码服务平台，给云管理平台应用层提供密码支撑能力；
- 通过密码服务平台，给云上租户系统应用层提供密码支撑能力；
- 通过部署密码设备和产品，给云平台下三层提供密码支撑能力；
- （可选）通过密码服务平台，给云平台设备层（包括设备层的身份鉴别、日志完整性等）提供密码支撑能力。

➤ 身份鉴别

云管系统用户通过使用基于国密算法且具备商密认证证书的智能密码钥匙，使用公共网络资源区部署的密码服务平台提供的身份鉴别功能，该功能采用符合 GB/T15843.3 的非对称密码的身份鉴别，基于 SM2 算法的挑战-响应机制实现云上租户系统用户的身份鉴别，云平台管理员的身份鉴别方式参考设备层远程运维人员。

➤ 访问控制信息完整性

系统管理员分配用户权限时，在权限数据入库存储时在服务端调用密码服务平台云服务器密码机的电子签名功能，基于 SM2 算法的电子签名技术实现用户访问控制信息的完整性保护。在用户登录环节以及其他关键业务操作环节，系统均会通过与数据库里的用户权限信息进行权限验证，使用密码服务平台云服务器密码机的签名验签功能校验电子签名值。

➤ 重要数据传输机密性、完整性

重要数据在云平台云管系统应用层传输时，需使用密码技术实现数据传输的机密性、完整性保护。目前仅通过网络层措施来缓解应用层高风险，主要通过互联网区部署的已有的 SSL VPN、公共网络区部署的国密 SSL VPN、安全运维管理区部署的国密安全认证网关，在网络层实现对登录过程中身份鉴别信息（用户名口令）传输的机密性保护，以及对迁移镜像文件传输的完整性保护。

➤ 数据存储的机密性

使用密码服务平台云服务器密码机的数据加密功能，基于 SM4 的对称密码技术，对政务云平台系统用户身份鉴别信息（用户名口令）进行加密保护，确保数据库里的系统用户身份信息敏感信息内容不被非法泄露，实现重要数据的存储机密性保护。

➤ 数据存储的完整性

在云平台管理系统中的重要数据，包括镜像迁移数据、日志审计数据、关键配置数据、系统用户权限数据等，通过使用密码服务平台云服务器密码机的电子签名功能实现对重要数据的电子签名，实现重要数据存储的完整性保护。在取出已进行完整性保护的数据时，使用密码服务平台云服务器密码机的电子签名验签功校验存储数据的电子签名。

2) 密码应用部署

根据系统密码应用的需求和建设目标，通过本次密码应用建设对上述系统和系统用户提供安全可靠服务，其服务核心是实现“可信身份、可信行为、可信数据、可信通道、可信时间”等密码应用，保障系统网络身份和实体身份的真实对应，建设重要数据防篡改、重要数据防窃取等服务的重要支撑基础设施。

方案设计的部署图如图所示：

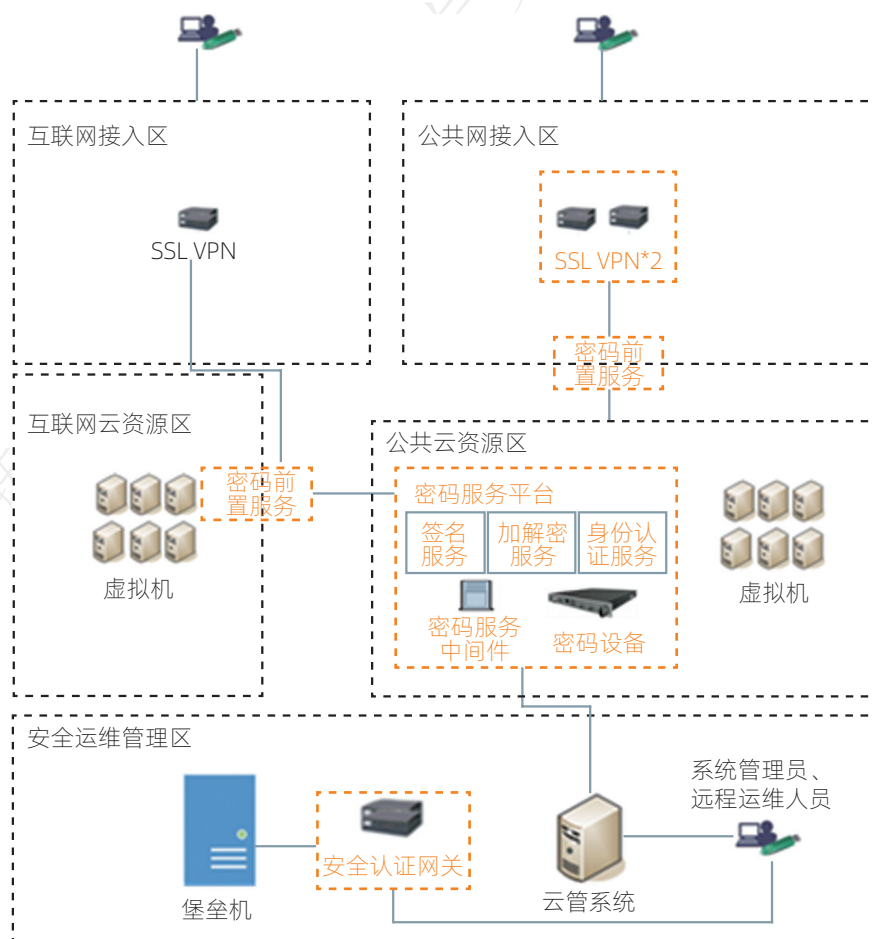


图 37 云平台密码应用部署示意图

5.3.2.3. 具体实践

能源云平台密码改造总体分为以下几个步骤：

第一步：明确密码改造流程节点

对于已经开发完成的部分功能模块，应用厂商（含云服务提供商等）明确需要加入密码改造的模块，对流程进行重新整理。由模块负责人进行梳理，整理出文档，与项目经理进行确认，确认后方可进行改造。

第二步：密码服务对接、联调

根据密码厂商提供应用改造对接接口，应用厂商结合自身的密码应用场景进行密码应用改造。密码厂商配合应用厂商进行联调测试。

第三步：试运行

应用厂商试运行时间需按照项目实际进度安排调整。试运行期间，需要严密监控应用系统和密码服务的使用情况。一旦出现异常情况，需要及时处理反馈。



阶段	XXX 局	应用厂商	密码厂商
明确密码改造流程节点	统筹组织开展系统建设、体制机制、标准规范制定工作，并组织项目的评价验收	明确需要加入密码改造的模块，对流程进行重新整理。	/
密码服务对接、联调		结合自身的密码应用场景进行密码应用改造；主导联调测试工作。	提供应用改造对接接口；配合联调测试工作。
生成环境部署、上线		部署到生产环境中	/
试运行		试运行	/

5.3.2.4. 总结

信息系统上云已成为主流趋势，云平台作为信息系统的重要载体，其安全直接关系到系统的安全，更关系到各业务系统数据的安全，数据安全目前已经从“防漏洞、补漏洞”的应对式防护，转化到“为数据访问重建安全规则”的主动式防护，将密码技术等安全手段直接作用于数据，是最直接有效的主动防护措施，密码技术可以直接保障数据的机密性和完整性，同时也可以保障信息传递中的真实性和不可否认性。

为了支撑能源行业中信息系统和应用的密码技术安全需求，也为了密码技术能与各信息系统深度融合应用，充分发挥保障信息系统和重要数据的安全性等方面的重要作用，本案例建设了统一的密码基础设施，一方面，使用密码技术保证云平台本身的安全，另一方面，使用密码技术给云上租户系统提供基础密码服务支撑，实现身份真实性、数据完整性、机密性、操作不可否认性等。确保云平台密码应用满足商用密码应用安全评估要求，助推云平台建设符合信息化主管部门要求，保障云平台业务开展的安全性、合法性、合规性。同时形成应用接入密码服务平台标准规范、制度，指导各信息系统快速实现密码服务的集成；在管理角度实现了统一密码资源管理、统一密钥管理、统一业务监控分析考核等监管功能，通过数据来全面体现国产密码推进和管理工作效率，提升能源行业各集团管理的精细度。

5.4. 教育行业的数据安全治理

5.4.1. 某部委数据安全治理实践

5.4.1.1. 背景

教育行业是关乎国家发展和社会进步的基础行业之一，同时也是网络安全法定义的关键基础设施行业之一。在互联网时代，随着教育行业信息化建设的高速发展，诸多信息安全问题纷纷暴露出来。为深入贯彻中央关于网络信息安全工作的总体部署，落实教育信息化中长期发展规划（2021-2035年）、“十四五”规划与信息安全等级保护制度的有关要求，亟须加快建立健全教育行业网络与信息安全保障体系，提高防护能力和水平，从而保障教育事业健康有序发展。

某部委业务种类众多，同时伴随着教育行业数据敏感、业务特殊、系统繁杂的特点，并且需要



数百个应用系统和数据库做支撑。对于存储着海量数据资产的某部委来说，如何发现有价值的数据，在做到全量的盘点和梳理的同时，对其中存在的敏感数据使用情况进行管理和管控，就变得尤为困难。这些不足直接导致现阶段某部委信息化建设无法正确决策后续数据安全建设的方向和防控策略。

目前，某部委数据安全主要存在如下不足：

- 某部委在数据安全防护过程中，对数据的使用、流动、敏感数据范围和分布都缺少全面的掌控，这就造成某部委以往对数据安全手段是采取“完全封闭”的态度和措施。

- 某部委对数据保护的主要防护手段，是依靠传统的网络防火墙、入侵防御等网络安全产品进行防护，而这些传统的安全防护体系，对数据库这个层面的防护力度不足。

5.4.1.2. 安全对策

针对数据泄露及各种安全事件存在的原因，为了提升数据安全保护的效果，切实保护教育行业数据的安全性，提出如下对策和建议：

建立以数据为核心的泄露防护体系，通过对敏感数据的梳理、加密等方式来保障数据的安全可控。

制定教育行业的敏感数据规范标准，在数据全生命周期过程中，依据数据信息的敏感程度对数据进行分类，依据敏感数据的标准对数据进行分级管理。用数据标签对创建数据、应用数据、存储数据、传输数据和销毁数据提供技术上和操作上的规范要求。对于关系到教育行业发展以及系统用户个人信息的重要数据，需严格控制其存储位置、传输和使用方式。对于完全可以对外公开的数据，不需要采取特殊的保护方式。提升行业信息系统用户安全意识，在教育行业中的数据泄露问题处理方面，需要增强用户的网络安全意识。

5.4.1.3. 治理实践

某部委将数据安全治理工作合理划分为四个步骤：（1）全面梳理某部委各系统的敏感数据资产；（2）制定教育行业的数据安全管理规范；（3）选择教育行业的数据安全治理技术；（4）开展教育行业的数据安全落地稽核。

1) 梳理敏感数据的资产

某部委拥有着庞大且错综复杂的业务系统，进行资产梳理的目的是：扫描检测数据资产，梳理数据库分布情况，通过扫描发现敏感数据，并对资产的访问热度做出分析，协助某部委完成数据等级分类，合理规范使用教育数据资产。

此阶段数据梳理的主要工作内容包括：

- 数据资产梳理（主要包括数据库、表、字段等）；
- 针对敏感数据的分布情况梳理、访问情况梳理、授权情况梳理；
- 当前环境下数据安全措施、产品、工具的梳理；
- 评估当前环境下数据安全风险等。

2) 确立符合行业需求的数据治理技术

基于数据梳理的成果，将学籍数据、教师、校舍、学生资助数据作为海量数据的重中之重，作为某部委第一阶段数据治理重点关注的对象。





在治理过程中，将数据表中涉及的“姓名”、“身份证号”、“地址”作为条件，分级别建立数据治理方案：

- 数据表同时含有三个因素，级别★★★，采用数据加密保护手段；
- 数据表同时含有其中两个因素，级别★★，建议采用数据脱敏保护手段；
- 数据表只含有其中一个因素，级别★，建议采用数据运维管控保护手段；
- 其中针对数据脱敏功能，选择了两种数据脱敏保护手段：

➤ 静态脱敏：对于某部委生产重要数据，在开发、测试环境的使用进行数据演练和测试时采用静态仿真脱敏的方式。

➤ 动态脱敏：某部委领导定期检查系统报表信息，领导对系统查看要求实时性检查，对人员敏感信息进行遮蔽处理，防止数据泄露。

3) 数据安全落地稽核

数据治理稽核目的是监督数据安全落实情况，针对数据在使用访问过程中的潜在风险进行预警，利用数据库审计和数据态势感知技术做好数据的安全稽核与风险预警。将监控的所有节点实时上传至监控统一平台，平台系统监测节点状态，业务流向；对分析发现的可疑威胁源与高风险事件，进行持续跟踪分析，并结合历史告警事件进行关联分析，从而挖掘出隐藏的真正威胁。

5.4.1.4. 实践总结

数据安全治理是一项复杂的系统工程、某部委通过数据安全治理实践成果，有效满足了自身所面临的数据安全管理需求：使数据安全可视、可控、合规。除带来上述主要价值外，数据安全解决方案还带给客户如下价值：

- 简化业务治理流程，提高数据安全能力；
- 完善纵深防御体系，提升整体安全防护能力；
- 减少核心数据资产被侵犯，保障业务连续性；
- 提升了信息化综合管理水平。

同时，上述实践工作对推进教育领域的政务信息资源整合和数据共享，提升政府信息资源再配置、再利用也有着深远意义。

5.5. 电信运营商行业的数据安全治理

5.5.1. 某电信国际的数据安全治理实践

5.5.1.1. 背景

某电信国际有限公司为拓展海外业务而组建的国际业务公司。伴随全球经济经历数字化转型，数据逐渐成为继土地、劳动力、资本之外的第四大关键生产要素。与此同时，作为企业三大资产——固定资产、财务资产和数字资产，数据安全已经成为网络安全一个非常重要的核心。国际公司伴随业务的快速拓展，数据随之增多，也开始关注数据安全，并着手准备针对数据防护的安全体系建设工作。

5.5.1.2. 目标和挑战

1) 数据安全威胁对现有传统防范思路的挑战

传统的安全防护大多数将网络设备、服务器、主机、操作系统或数据库及相关应用程序本身视为保护对象，宏观来讲属于静态的防护；但数据本身是动态的，随时跨部门、跨区域、跨应用传输和运转，而且有些业务或个人数据是具备高价值的，一旦泄露，将造成巨大损失。因此，传统的防护思路是否可以应对现阶段的数据安全防护需求成为国际公司面临的第一个巨大挑战。

2) 来自外部法律法规和监管合规的压力

从法律法规的角度，我国已经建立多部法律法规，包括《网络安全法》、《刑法》、《个人信息保护法》（草案）《数据安全法》（草案），且愈发有针对性地提出对数据安全的保护。

国际公司作为运营商企业，除了遵从法律法规的制约，也要符合行业监管机构的合规管理，包含《2020 省级基础电信企业网络与信息安全工作考核要点与评价标准》、《IT 安全能力建设规范—5. 数据安全能力分册》以及《电信网和互联网大数据平台安全防护要求》。

5.5.1.3. 国际公司数据安全现状

此前，国际公司还未能对数据安全防护有清晰的思路，对数据资产的识别与数据分类分级缺乏专业性的支撑与解读，对数据资产重要性的了解也不够清晰；数据安全的全生命周期管理制度要求与实际业务场景结合落地困难，管理制度未能根据不同业务、不同专业实践落实；缺乏数据安全业务保护评估，未能有效识别数据安全关键薄弱环节；数据安全防护手段缺失，且对数据安全防护能力的补足也缺乏依据，存在采用无针对性的安全策略进行防护建设等问题。因此，亟需一套完整、体系化、全覆盖的数据安全治理体系思路和方法，辅助国际公司落实数据安全防范的相关具体措施。

5.5.1.4. 安全对策

1) 设计思路



图 38 数据安全治理体系框架图

针对国际公司的现状以及具体的需求，基于“数据安全管理体系 + 数据安全技术体系 + 数据安全运营体系”的组合，形成落实国际公司数据安全治理体系的方法论：

- 数据安全管理体系，确定数据安全治理的对象、管理者并建立管理制度，做到“有制可依”。

- 数据安全技术体系，确定制度落实的操作流程，明确管理者的手段和工具，做到“有规可行”。
- 数据安全运营体系，基于自动化数据安全运营管控平台，结合数据安全治理服务保障团队的支撑，做到“风险必查”，“行为可溯”。

2) 体系规划



图 39 数据安全治理体系规划路线图

国际公司的数据安全治理服务规划路线分为三个阶段：

第一阶段是数据安全基础防护阶段，对于数据安全治理体系基本的治理组织机构、治理对象、治理基本要求和策略进行总体调研和把控，然后进行整体性的数据安全风险评估，发现业务/个人数据在业务流转中会出现的风险场景，最后提供符合国际公司现状的数据安全治理规划方案。

第二阶段是策略优化及能力提升建设阶段，建立符合国际公司数据安全需求的管理制度，形成操作规范，梳理完善数据安全防护落实策略流程，补足数据安全防护工具。

第三阶段是形成数据安全运营管控阶段，配备自动化数据安全管控平台，结合一线专业数据安全工程师，二线数据安全咨询顾问，实现对国际公司数据安全日常运营、监管与高效的数据安全服务保障。

5.5.1.5. 治理实践

1) 路线概览



图 40 数据安全治理体系规划路线 - 第一阶段



基于国际公司对数据安全防护思路不清晰，缺乏体系化的防御手段，本方案先从基础的数据防护建设服务开启。

- 现状调研——确定数据安全治理的业务范围；
- 数据资产梳理——确定安全治理对象、对象分布范围及承载数据的数据库权限情况；
- 数据分类分级——确定数据安全治理对象的类别和等级，识别重要敏感数据，区别防控策略的基础；
- 数据安全评估——发现数据安全治理业务范围和对象的风险场景，从而建设规划的基础；
- 规划方案蓝图——针对国际公司现状、数据安全治理对象及业务范围，以及在流转中所产生的数据安全风险，基于管理、技术、运营三方面进行方案蓝图设计。

2) 业务数据资源调研摸底

客户的困难在于不知道数据资产具体分布，不清楚数据资产家底状况。

数据安全咨询顾问采用“工具+人工”的方式，利用数据资产梳理工具进行了数据资产识别，梳理了数据库账号、权限、表和字段等，形成了数据资产清单，明确了数据资产的分布情况。

3) 数据安全风险评估

数据安全评估工作将围绕与敏感信息相关的业务流程进行评估，厘清各流程中数据的流转和使用情况，重点评估那些在信息流转和使用的过程中能够直接或间接接触到这些数据信息的情况。

结合对敏感信息保护的实际情况，针对项目范围内的数据库进行漏洞扫描，包括弱口令、缺省配置、配置缺陷、漏洞等。

同时，依据国家和工信部、香港和国际重要数据安全监管要求，设计符合客户的评估矩阵，采用人员访谈、文档查阅、系统查验、技术测评等方法，帮助客户查找文档和策略配置，评估存在的潜在风险，提出整改建议，形成评估报告。

4) 安全保障体系规划

依据评估报告中存在的风险，结合客户现状，规划未来客户数据安全建设蓝图，包含数据安全管理体系、技术体系和运营体系建设。

数据安全管理体系，通过建立四级文档的形式，明确数据安全方针、组织架构、制度要求、流程规范、记录表单等。

数据安全技术体系，通过规划数据资产识别梳理、数据处理活动安全防护能力（数据管控、数据审计等）实现事前诊断、事中控制、事后分析溯源的数据安全技术能力。

数据安全运营体系，通过规划数据安全监管平台，实现数据资产管理、数据安全策略管理、数据安全事件监测和数据安全风险运营能力。

5.5.1.6. 实践总结

本项目中的数据安全治理为客户提供的价值主要有：

- 加强了国际公司对于数据的安全管理；
- 建立了符合国际公司的数据安全治理体系；





- 通过建立数据安全治理体系提升了国际公司的数据安全保障能力；
- 满足国家相关法律法规及行业监管要求。

5.6. 医疗行业的数据安全治理

5.6.1. 某医院医疗数据灾备实践

5.6.1.1. 背景

为解决不断增长的医疗需求，提升医疗服务的质量和效率，国家加速推动医院信息化转型升级，实施“健康中国 2030 战略”。但医疗信息化的发展快速，使得医疗业务对信息系统、基础设施的依赖逐渐增强，在各种勒索病毒、网络攻击等安全事件频发的环境下，任何人为或其他因素导致的应用中断和数据丢失，都会给医院造成巨大的社会影响和经济损失。因此，对医疗数据的备份及应用容灾的需求逐渐凸显，同时，《信息安全技术网络安全等级保护基本要求》、《数据安全法》以及医疗行业相关政策的制定也推动着医疗数据安全建设，对医疗数据全生命周期管理以及数据安全治理提出了更高的要求。

各级医院为加快信息化转型，大部分已建成涵盖医院管理和临床应用的主要信息系统，包括 HIS、PACS、EMR、LIS、自助挂号查询等信息系统，同时，部分地区已建设医疗云平台，实现医疗业务上云，随着医院智慧医疗水平的提升，医院运营对医疗数据的依赖越来越强。

与其他行业相比，医院信息化系统呈现专业性强、非结构化数据多、业务连续性要求高等特点。目前，多数医院已实现本地数据库备份保护，对于医疗数据异地备份、应用容灾及非结构化 PACS 数据备份手段欠缺。同时，卫健委在数据安全方面制定了多个管理规范 and 指引，3 级电子病历要求医院对全部系统有软件 and 数据的备份，4 级电子病历要求医院具备异地灾备机房 and 异地数据备份能力。同时根据法律法规，医院信息系统要满足相应合规性要求，严格按照等保 2.0 定级标准落实相应数据安全措施。

5.6.1.2. 安全对策

在遵循卫健委相关安全标准与规范的基础上，根据网络安全等级保护要求，为医院规划建设本异地灾备体系，保证医疗数据的安全性。由于医疗应用分散部署于云环境及物理环境中，具有业务主机繁多、数据量大等特性，因此对数据与业务保护提出更高的挑战：

1) 数据安全需求

各类医院 HIS、LIS、EMR 等关键业务已实现上云，数据库多为 Oracle、Mysql 和 SQLserver，其中 PACS 系统具有海量非结构化影像数据，为确保不同场景下的医疗数据安全，针对医疗数据需提供本异地分级备份手段满足不同 RPO 要求。

2) 医疗业务连续性需求

当灾难发生时，需能够保障医疗应用的连续性，尤其是面向关键业务，一旦中断，会导致医疗

服务停摆，造成严重的社会影响，因此需提供面向关键业务的容灾措施，保障医疗业务永续。

3) 合规性需求

根据等保 2.0 要求，等保定级为二级和三级的信息系统需要具有本异地备份能力，三级系统要满足业务高可用；同时根据《医疗机构病历管理规定》门诊病例要留存至少 15 年，住院病历留存至少 30 年，要求对医疗病例进行长期归档保存，满足合规性要求。

4) 统一灾备运维

为实现高效灾备管理与统一运维，需要满足灾备资源的集中监控和调度，降低运维复杂度，提高运维效率。

5.6.1.3. 具体实践

为解决各级医院的数据和业务安全问题，综合医院投入及业务特点，建设本异地灾备体系，在本异地部署统一灾备系统，实现结构化及非结构化数据的本异地分级数据保护、业务容灾及数据归档，同时通过异地灾备平台实现本异地灾备系统的统一管理和运维。方案拓扑如下图所示：

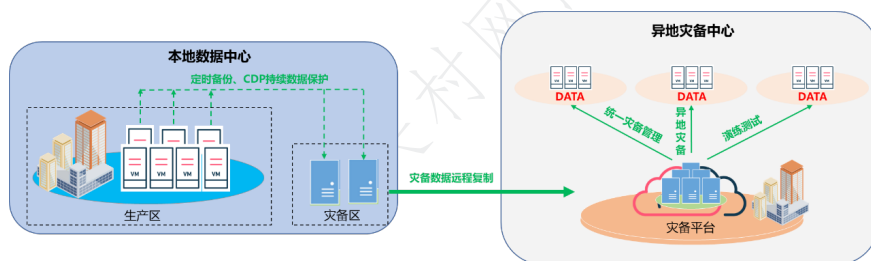


图 41 容灾备份拓扑图

1) 本异地数据分级备份

通过在本地数据中心、异地灾备中心部署灾备系统，针对关键及一般数据提供分级数据保护机制，采用小时级、分钟级和秒级分层数据捕获技术，结合远程数据复制技术，实现医疗数据本异地定时备份和实时备份，满足不同 RPO 及合规性要求；同时，结合数据可用性校验技术，确保备份数据的可用性。

2) 海量非结构化数据备份及归档

对于海量的非结构化 PACS 数据，采用灾备系统分布式可扩展架构，灵活的在线扩展灾备容量，实现海量备份任务的并发、数据的负载均衡，提升海量数据备份效率。同时，通过在各本地数据中心部署光盘库，通过在线归档、离线保存等方式，实现医疗病历等数据的长期归档保存。

3) 本异地容灾接管

本异地灾备系统提供面向关键业务的容灾接管，采用实时保护技术、自动/一键式应急容灾接管技术、自动/一键式业务演练技术等，当业务系统发生故障或因恶意软件、人为因素导致系统停机时，可提供任意历史时间点状态的异地应用级容灾接管，确保关键业务 7×24 小时不间断运行。

4) 统一灾备管理平台

通过异地灾备系统构建灾难恢复演练区、可视化监控与集中管理区，实现医院定期的演练恢复



和数据管理，并对相关设备资源、灾备资源、策略管理等进行统一管控，提供全域、综合化、规范化的运营支撑。

5.6.1.4. 总结

通过构建本异地灾备体系，对医疗数据实现分级分类保护和全生命周期数据管理，完善了针对不同等级数据的分层保护机制；建立了异地容灾保障体系，为医疗核心业务系统提供异地应急保障措施，满足合规性要求；同时实现灾备统一管理运维，从而提升网络数据安全治理效率，优化了医疗服务水平、提升了工作效率和增加人民群众满意度。

5.7. 车联网行业的数据安全治理

5.7.1. 某车联网企业基于数据分类分级的防护体系建设实践

5.7.1.1. 背景

随着企业数字化程度加深，数据成为企业重要的信息资产，是推动业务发展的重要原生动力。国家层面强调要切实保障国家数据安全，加大关键信息基础设施保护力度，强化国家关键数据资源保护能力。以工业－车联网行业为例，大型企业在网络安全方面已经采用多种安全控制措施，但在数据安全治理层面还面对诸多挑战。现以工业－车联网行业道路交通管理车联网数据安全治理项目为例，为数据安全治理落地实践案例集提供参考。

本案例涉及数据主要有红绿灯展示数据、交通绿波协调控制信息、道路可变限速信息、车辆位置信息、车辆行驶速度信息、车辆传感器信息、总线数据等多项重要数据。以上数据如果得不到有效监管，将对我国交通关键信息基础设施造成严重威胁，对我国重要数据及公民个人隐私安全造成潜在影响。

5.7.1.2. 安全对策

车联网应用需要收集车外数据、车内数据、车辆运行数据、车辆位置轨迹数据等，依据车联网网络安全和数据安全相关标准要求，确保数据处于有效保护和合法利用状态，保障车联网安全稳定运行。因此，需要建立覆盖车联网数据全生命周期的数据安全治理方案，从技术与管理两方面共同对抗车联网数据安全风险。

5.7.1.3. 具体实践

本方案以政策为导向，技术要求为支撑，按照以下工作事项进行建设。

一是数据分级分类。按照“谁主管、谁负责，谁运营、谁负责”的原则，建立数据管理台账，实施数据分类分级管理，加强个人信息与重要数据保护。定期开展数据安全风险评估，强化隐患排查整改。





二是数据安全技术保障。采取合法、正当的方式收集数据，针对数据全生命周期采取有效技术保护措施。强化数据安全监测预警和应急处置能力建设，及时处置数据安全事件。

三是规范数据开发利用和共享使用。合理开发利用数据资源，明确数据共享和开发利用的安全管理和责任要求，对数据合作方数据安全保护能力进行审核评估，对数据共享使用情况进行监督管理。

四是强化数据出境安全管理。如需向境外提供在中华人民共和国境内收集和产生的重要数据的，保证依法依规进行数据出境安全评估并向各级主管部门报备。

本项目通过四个阶段完成：

第一阶段：针对重要系统——智能网联汽车、路侧智能设施身份管理及认证系统，对相关数字资产落实资产梳理，进行数据分类分级，形成数据清单，进行合规性检查，制定数据防护方案与策略。

在实施数据资产梳理过程中，首先依据组织业务系统的数据安全需求以及相关法律法规的规定，按照数据安全管理的目标和原则，梳理组织核心数据处理活动有关的数据情况，形成了初步的数据清单。其次，根据数据的业务属性以及数据在处理活动中的重要程度，对数据进行分类，形成数据目录。最后，按照已分类数据在经济社会发展中的重要程度，数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用等造成的损害，以及对国家安全、公共利益或者个人、组织合法权益造成的影响和危害程度进行分级，形成最后的数据分级分类目录。

数据目录（节选）如下：

数据名称	应用场景
红绿灯展示数据	车与路方向－红绿灯数据车内展示
交通绿波协调控制信息	车与路方向－交通绿波协调控制信息提示
交叉口可变车道功能状态信息	车与路方向－交叉口可变车道功能状态显示
道路可变限速信息	车与路方向－道路可变限速信息提示
车辆位置信息	车与车方向、车与路方向
车辆行驶速度信息	车与车方向、车与路方向
车辆传感器信息	车与车方向、车与路方向
总线数据	车与车方向、车与路方向

第二阶段：依据工作方案，根据数据分类分级防护策略，部署与整合数据安全组件能力，形成数据使用、共享等关键环节的防护能力。

本案例中车联网身份认证及认证管理平台进行了证书管理平台硬件基础设施、网络环境的建设以及管理软件功能模块的部署。针对车路直连通信安全，基于国密技术，开展信号机数字身份安全模块的研制、测试工作。信号机按照 GA/T1743《信号机信息发布接口规范》，推送经过数字签名的信息。

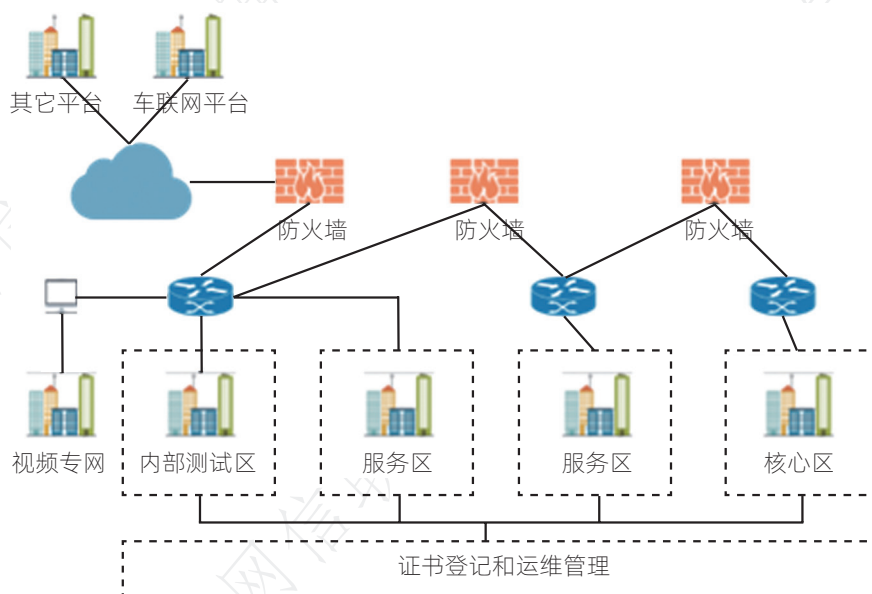


图 42 身份认证平台架构图

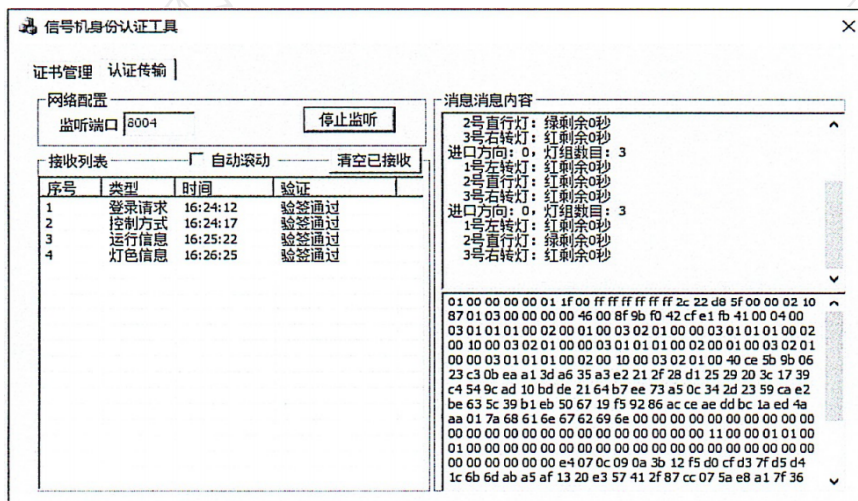


图 43 推送经数字签名的红绿灯灯色信息

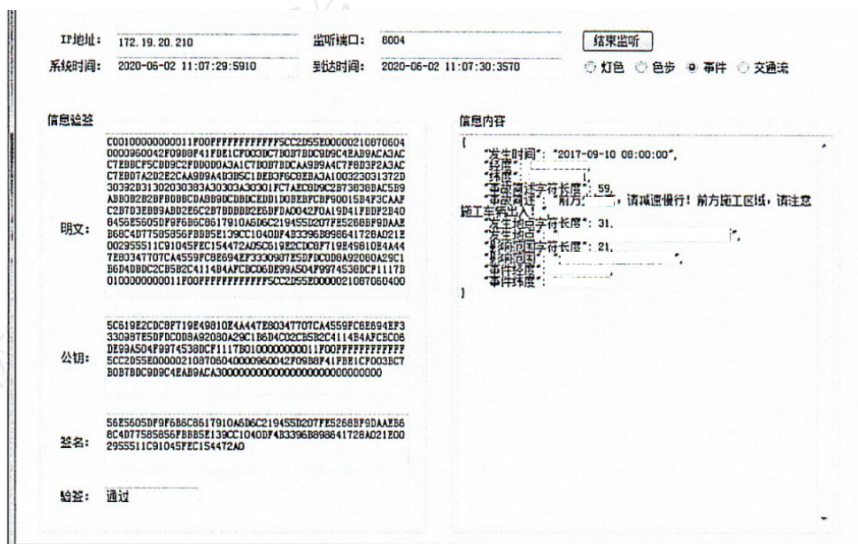


图 44 推送经数字签名的事件信息



针对车车直连通信安全，基于国密技术，应用内嵌智能网联汽车数字身份芯片的车载 OBU 对车辆发送消息进行签名，将消息和证书广播发送，接收端车载 OBU 验证接收到的证书，通过后使用接收到的证书验证消息真实性和完整性。

在车路安全通信中，分别使用 SM2、SM3 算法对信号机推送的信号灯灯色信息、事件信息等进行数据的完整性验证，使用多种安全芯片，设计信号机数字身份模块；在车车安全通信中，分别使用 SM2、SM3 算法对车辆广播的预警、提醒等消息进行数据的合法性和完整性验证；对于关键敏感数据，使用 SM4 算法对车辆广播的消息进行机密性保护。

第三阶段：构建数据安全防护体系，及时发现数据安全风险并进行事件响应控制，形成持续运维能力。

在平台建设方面，建立网络安全防护基础设施，设置服务区、管理区、核心区、密码管理区等安全区域，提出严格的网络安全管理要求；在路侧交通管控设施方面，进行设备接入认证、信息签名等安全通信措施；在智能网联汽车方面，依托网络侧安全防护基础设施，结合业务侧车辆注册登记、定期安全检验、注销等业务，对数字证书全生命周期进行安全管理，有效保障数据的传输安全。

第四阶段：对数据安全治理体系进行完善优化，建立数据安全大脑，实时监测数据的分布及流转情况，并面向全业务场景进行治理。

持续完善车联网身份管理及认证系统的设计，适时升级测试道路数量、扩大测试范围，稳步推进车联网车车、车路应用场景试点工作，同步验证证书体系功能完备性与适用性，以及路侧车联网通信设施、交通管控设施的性能效率。

5.7.1.4. 总结

本案例完成了车联网身份认证系统建设，实现了车联网安全通信，支持跨信任域的互信互认互通，为数据提供了基于全生命周期的安全保护，降低了车联网中包括红绿灯灯色信息、交通事件信息等在内的交通重要信息与包括车辆速度信息、车辆位置信息等在内的车辆重要信息泄露与破坏的风险，为车联网数据安全提供了保障。

本案例探索了车联网数据安全治理的实现路径与形式，为全方面推广车联网数据安全治理提供了可复制可推广的经验。将车联网数据安全治理方案全面推广还需要进一步畅通反馈机制，优化方案设计，提高应用实效性，及时总结提炼、完善规范、争取更大范围的推广。

5.8. 烟草行业的数据安全治理

5.8.1. 某烟草公司基于零信任和隐私计算技术的防护体系实践

5.8.1.1. 背景

随着各类数字化技术大量应用，依托数采工作和样本终端建设，烟草行业系统近年来积累了庞





大的数据资源。烟草行业大数据中心旨在通过汇聚行业数据，实现烟草行业数据价值的转化和释放，增强烟草系统灵活性、开放性、交互性、经济性和共享性，在这一过程中由于数据安全问题导致数据流转困难，也不可避免地带来数据集中后的安全问题。根据烟草大数据中心的业务特点和网络架构，其安全治理过程中的安全风险主要来自以下几个方面。

1) 数据隔离安全

随着各卷烟厂与大数据中心的协同交互不断加强，远程办公规模的不断扩大，内外网边界越来越模糊，传统安全技术构成的安全边界无法阻挡向内部渗透、外部攻击的威胁。

2) 数据流转安全

目前烟草大数据中心网络内部服务器/主机与数控系统网络连接存在安全风险；数据流转和交换很可能被外部攻击者利用，侵入内部系统，执行越权提交作业、修改运行状态、篡改数据等恶意行为，影响数据安全；数据流量明文传输，存在被窃取风险；数据传输缺少正确性校验，数据包存在被篡改风险。

3) 隐私保护需要

目前在数采工作中尤其是数据收集和分析过程中未能对个人隐私数据提供有效的去标识化措施，数据存储未能加密，缺少明确的数据分级分类规定，需要满足数据安全法、个人信息安全保护法的合规要求。

为保障烟草大数据中心的落地，根据网络安全法、数据安全法等国家有关要求，在全面了解行业网络安全建设现状的基础上，在数据处理活动（收集、存储、使用、加工、传输、提供、公开等）防护的基础上，开展安全监控和审计，以实现对数据安全风险的监控，建立相关的措施对非法收集、未授权访问、数据滥用、数据泄露进行监控和审计，通过数据分析来支撑有效的安全、合规决策，从而降低数据安全风险，推动网络安全整体防护水平取得长足进步，指导各部门及各卷烟厂的数据安全管理工作。

- 需实现各卷烟厂与大数据中心的安全接入，以及业务用户远程办公的安全接入，确保内部网络不被渗透攻击，防止出现非授权的访问，造成数据被篡改或泄露等安全风险；

- 需建立数据安全平台，基于数据分类分级应用隐私计算技术实现对数据的安全管理和保护，在确保数据不出安全域的情况下实现数据的分析计算，释放数据价值的增长。

5.8.1.2. 安全对策

1) 总体架构

采用“分层解耦”的设计理念，将大数据平台分为安全接入层、数据资源层和数据应用层，并配套相应的数据安全防护体系，各层之间采用集中数据总线进行数据传输和交换，以此降低各类数据应用对底层数据存储之间的强依赖性，各层之间独立工作，方便后期的安全业务扩展，并保障各层之间稳定运行。数据安全管控总体框架如下图所示。



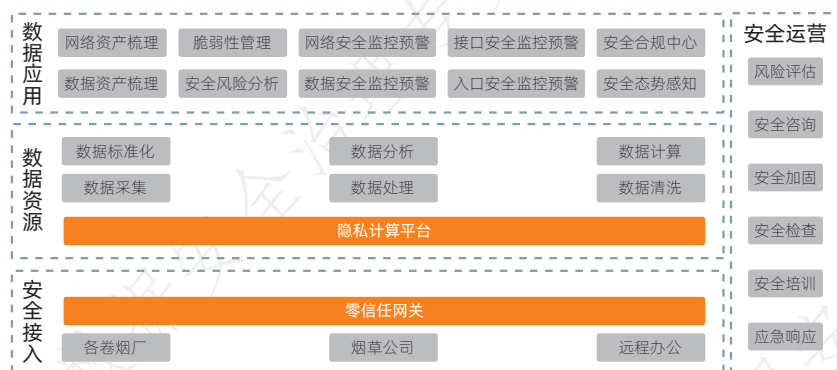


图 45 数据安全管控总体框架

2) 部署方式

数据安全管控支持包括云环境部署、分布式部署、旁路部署、集群部署等方式，无需改变现有网络结构，也不需要安装客户端软件，数据安全管控通过WEB方式登录，执行数据安全管控操作，重要边界部署安全网关。

3) 管控措施

数据安全管控通过对各类数据日志、流量采集进行数据安全审计、异常行为分析、漏洞管理、威胁管理安全数据分析、数据安全预警等，实现大数据安全运营管理。通过漏洞扫描、敏感数据发现、数据脱敏、水印溯源、数据加密、敏感数据分级分类等，实现数据安全流转及监测安全管理，提高大数据中心敏感及隐私数据安全能力。通过身份安全与访问控制基础设施系统为大数据中心各应用系统提供统一账号管理、统一认证管理、统一授权管理和统一访问控制，为推动零信任体系的部署提供支撑。

4) 数据分类分级

开展数据安全治理首先要定义敏感数据，基于业务特点进行数据的识别、数据分类、数据分级。数据分类分级的准确清晰，是后续数据保护的基础。由于数据类型不同，根据《网络安全法》要求对个人信息和重要数据分开进行评估与定级，再按照“就高不就低”的原则对数据条目进行整体定级，进而根据不同需要对关键数据进行重点监管。记录对这些数据的访问行为，并通过各种策略模型发现异常的访问，比如规定敏感数据只能由指定的用户从规定的IP地址访问，当某个用户的IP发生改变时，及时报警，因为有可能该用户的账号发生盗用。

5) 隐私增强

严格按照《网络安全法》收集、存储和使用个人信息，对存储、处理和使用个人信息的信息系统实行备案管理。定期应用系统隐私政策设置、信息系统数据和公民个人信息的约定条款进行评估和问题整改；对所有信息系统进行逐一核查，重点排查违规“读取、存储和使用行业数据”的情况及风险隐患，隐私增强主要是保障隐私数据在上传、融合计算等过程中的安全性。主要提供如下三种安全机制。

● 数据全加密传输

所有的隐私数据采用数据提供方私有的密钥进行加密上传，保障这类数据只有数据提供方可解



密使用；

● 隐私数据脱敏

通过脱敏技术实现对数值和文本类型的数据脱敏，支持多种脱敏方式，包括不可逆加密、区间随机、掩码替换等。脱敏技术需要能自动扫描发现敏感信息，实现高效、方便、准确的信息脱敏；

● 密文计算

在隐私数据的融合计算过程中，采用密文计算技术，保障数据无需解密，即可完成数据碰撞、查询、交集等运算。

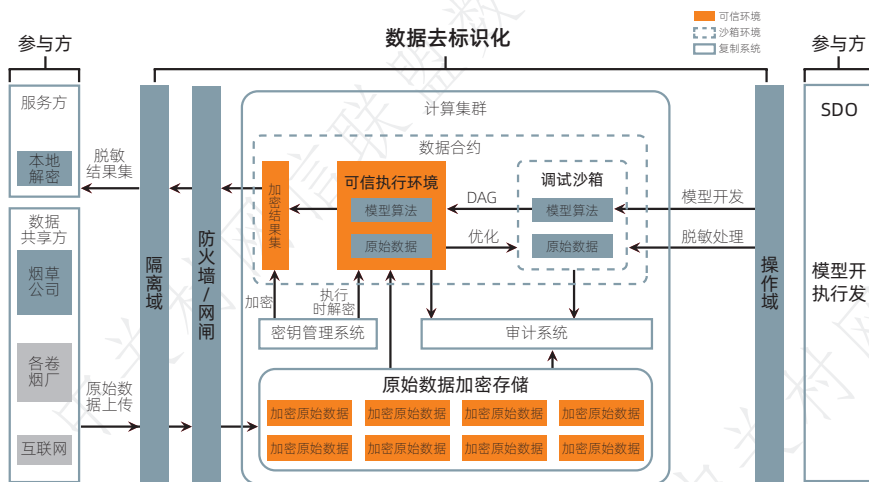


图 46 数据安全管控场景图

6) 隔离域

落实行业安全要求，积极开展人、终端、主机、应用之间的白名单访问控制和零信任服务，打破传统边界安全理念，提高安全事件的发现和响应能力，重点开展远程办公和各卷烟厂远程访问的零信任安全试点。通过零信任系统提供统一的业务安全访问通道，取消终端直连内部业务系统的网络策略，避免内部服务完全暴露在办公网络中的情况。各卷烟厂远程访问内部系统在零信任网关保护之后，只有通过零信任的认证和授权后才可以被访问使用。通过将访问流量统一接入零信任网关、零信任访问控制与保护引擎，实现在各卷烟厂网络环境下的内部资源访问。

5.8.1.3. 具体实践

本方案的数据安全管控平台是一个专注于保障数据安全流通，致力于解决数据共享过程中的安全、信任和隐私保护问题的隐私计算平台，综合应用安全计算沙箱，联邦学习，MPC 等多种前沿隐私处理技术，基于数据分类分级和多层次的数据安全隔离措施，实现共享数据的所有权和使用权分离，确保原始数据的“可用不可见”、“可用不可取”，保障数据全生命周期的可靠、可控和可溯。另外还考虑了零信任安全的部署，基础统一的安全认证和安全网关，实现各卷烟厂、烟草公司访问内部资源以及远程办公，降低了数据访问的安全风险，提高了管控的力度。

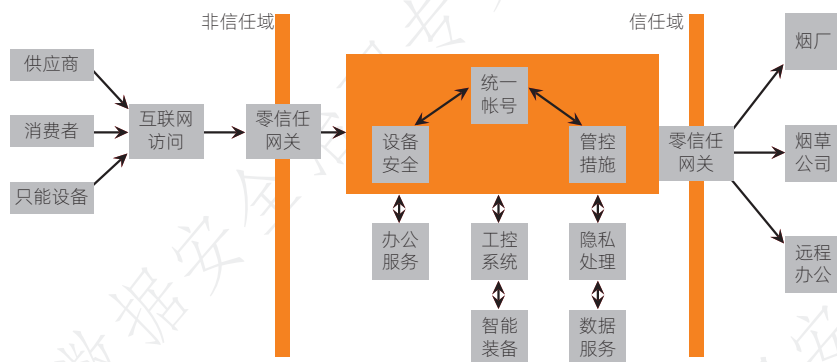


图 47 数据安全管控系统部署示意图

5.8.1.4. 总结

烟草行业大数据中心的安全治理是巩固现有数据安全建设成效，聚焦合规建设和基础性建设需求，综合考虑可操作性和国家政策需要，优先实施一批基础性、通用性和关联性的数据安全保障和显著提升基础防护能力的项目，重点推动数据安全治理的体系化，深化数据全生命周期的安全隔离措施、零信任安全防护、工业数据安全，实现防护的动态化、主动化和纵深加强。全面提升整体防御能力，打造数据安全管控的全面可视化、体系化，对数据服务技术与产品的风险、隐患、漏洞、潜在问题做到“心中有数、手中有招、控制有道”。各部门和卷烟厂的安全风险全面可控，积极应用零信任理念实现数据安全闭环管理，打破传统边界安全理念，提高安全事件的发现和响应能力，提升数据安全响应和管理水平，助力数据中心安全稳定运行。

5.9. 互联网广告行业的数据安全治理

5.9.1. 某互联网公司个人信息保护法律实践

5.9.1.1. 背景

2022 年 1 月，中关村互动营销实验室发布《2021 中国互联网广告数据报告》。据《报告》统计，2021 年，我国互联网广告市场稳中向好。后疫情时代，受益于内生需求的增长，互联网广告总体收入首次突破 5000 亿元，达到 5435 亿人民币（不含港澳台地区），同比增长 9.32%。互联网广告市场规模也在持续高速增长，中国已成为仅次于美国的全球第二大互联网广告市场。与传统媒体广告相比，互联网广告有着成本低廉、传播范围更广、投放目标更精准等特点，越来越多的经营者青睐于通过互联网广告开展宣传。但与此同时，互联网广告中的虚假宣传、侵害消费者权益、用户个人信息保护等问题也日益突出，引发各监管部门更多关注。自 2019 年起，互联网广告整治工作连续多年被列为网络市场监管专项行动的工作重点。

互联网广告依赖于大数据的应用，其中不可避免涉及到对个人信息的合规使用。利用用户数据实现广告主需求、用户需求和媒体广告资源的高效匹配，是互联网广告的基本运行逻辑，数据在行业各服务主体之间进行高并发的交互，是互联网广告业的基础需求。

2021 年 11 月 1 日，《个人信息保护法》实施，规定个人信息处理者向其他个人信息处理者提



供其处理的个人信息的，应取得个人的单独同意，同时要求个人信息处理者通过自动化决策方式向个人进行信息推送、商业营销，应当同时提供不针对其个人特征的选项，或者向个人提供便捷的拒绝方式，以此充分保障消费者的知情权和选择权。上述法律要求，对互联网广告运营的基本模式提出了较高标准。

2021 年 11 月，国家网信办发布《网络数据安全条例（征求意见稿）》（以下简称“《数安条例》”）进一步将取得个人“单独同意”设为个性化推荐的前置条件，并将“单独同意”限定为，不包括一次性针对多项个人信息、多种处理活动的同意。工业和信息化部自 2019 年 11 月启动了 App 侵害用户权益专项整治工作。2021 年 11 月，工业和信息化部发布《关于开展信息通信服务感知提升行动的通知》。同时还将重点整治应用软件启动弹窗欺骗误导用户、强制提供个性化服务等问题，包括开屏弹窗未提供关闭按钮、弹窗整屏为跳转链接、定向推送时提供虚假关闭按钮等场景。综合相关立法与监管动态已充分表明，我国对互联网广告行业个人信息保护的要求日趋严格。

互联网广告治理应与个人信息保护协同进行，互联网广告企业应严格遵守法律、法规以及行业标准的要求，针对互联网广告精准投放所使用的主要技术，其个人信息风险的不同以及用户对它们的接受度也不同，对广告场景下的个人信息保护施行不同的侧重点，但核心应遵守相关法律法规的要求，充分保障用户的合法权益。重点列举互联网广告企业应遵循的法律法规及行业标准如下：

- 法律层级

《民法典》、《广告法》、《网络安全法》、《数据安全法》、《个人信息保护法》等；

- 行政法规以及部门规章层级

《互联网广告监督管理暂行办法》、《电信和互联网用户个人信息保护规定》、《网络安全审查办法》、《网络安全等级保护条例》、《网络信息内容生态治理规定》、《数据安全管理办法》等规定要求；

- 国家及行业标准层级

《信息安全技术个人信息安全规范》（2020 年）、《个人信息去标识化指南》、《个人信息安全影响评估指南》、《APP 违法违规收集使用个人信息行为认定方法》。

5.9.1.2. 安全对策

互联网广告合规应该综合处理好数据安全、个人信息保护与发挥数据要素促进个性化广告发展之间的关系，三者兼顾，不可偏废。互联网广告形式多样，涉及的主体类型复杂，互联网广告合规需要针对不同广告类型，以及不同主体需求提供合规服务。广告数据合规要求企业在广告数据处理的全生命周期内，严格按照法律法规及政策标准履行数据安全管理和个人信息保护义务。既包括协助企业开展具体产品、服务的数据合规，也包括协助企业搭建广告数据合规体系。

总体而言，在个人信息收集、使用方面，应当遵循合法、正当、必要的原则。网络服务提供者应当在应用显著位置披露隐私政策，明确告知用户收集、使用个人信息的目的、方式、范围，同时配备相关专门人员和部门对收集、使用的用户个人信息的安全负责。个性化广告的精准投放，通常是根据无法直接识别个人的标识符进行的，主要依赖用户画像技术将具有共同特质的用户进行标签化，这一过程中不会定位到消费者个人，且个性化广告的精准服务一般是通程序化购买实现的。因此，在个人信息共享方面，企业需进一步完善自己的隐私政策，告知个人信息的收集会被用于用



户画像和个性化展示，并征得用户的同意，充分保障用户知情权。同时，应充分尊重用户选择权，为用户提供关闭个性化广告的便捷通道，加强广告场景下个人信息的技术保护，允许用户通过删除 Cookie 的方式，实现“无痕浏览”的权利，推动用户隐私保护合规与互联网公司利益的平衡，促进信息数据依法合理有效利用。此外，还应加强行业自律自治，在主管部门的指导和帮助下，行业协会制定标准和规范细则，为产业具体实践提供指引，支撑相关监管工作。在注重合规之外，还需要重视通过各种渠道，以生动易懂的方式增进用户对互联网广告精准投放的实质和所采用的相应技术的了解，以消除用户误解，打消用户顾虑。

5.9.1.3. 具体实践

企业应主动提升合规能力，配备相关专门人员和部门负责个人信息保护工作、主要包括：设立数据隐私部门，主动跟进并解读数据合规相关法律法规、监管政策和司法案例；研究国内外广告数据领域前沿法律问题并提供专业研究报告；定制并开展数据与个人信息保护相关主题的员工培训。同时，配置专业广告合规团队，负责起草广告数据合规相关法律文件，包括但不限于合同文本、个人信息保护指引等；提供数据合规问题咨询并出具法律意见书；协助排查产品数据与个人信息保护相关的法律风险并提供法律意见；协助应对数据合规相关监管约谈、通报和社会舆论事件，并提供应对策略；提供互联网广告相关法律咨询并出具法律意见书，包括三品一械、医疗、化妆品、教育培训、未成年人广告等。

5.9.1.4. 总结

近两年，个人信息保护相关法律法规逐步完善，为响应法律、落实行业标准，企业合规成本必然增加，部分业务需进行较大调整，由此产生的效率下降、业务重构、管理系统的构建以及运维成本等，将成为企业无可回避的支出。但《个人信息保护法》等相关法律法规的初衷是为了促进数字经济健康发展，做好个人信息保护合规才能更好的推动企业长久持续发展。跟进最新监管要求，及时调整服务协议、隐私政策，可有效避免因违法处理个人信息而受到处罚，避免更大经济损失。同时，可通过对外宣传合规文化，展示企业合规形象，为企业发展营造良好的合规舆论及监管环境，提升企业社会声誉及评价。





第六章 未来展望与倡议

为有效促进安全有序的开展数据开发利用，数据安全治理伴随业务和数据的发展是一个持续迭代演进的过程，需要国家、行业、企业或组织的多方力量共同推进。本白皮书一方面提出展望，对数据安全相关监管要求、管理和技术发展等方面进行趋势分析。另一方面提出倡议，对数据安全涉及的人才培养、产学研技术研究合作，行业联盟指导及治理国际化等方面向国家或行业组织予以建议。

6.1. 未来展望

6.1.1. 国家法律法规和标准体系日趋完善

近年来，以“三法三条例”为基础的数据安全相关的法律法规、政策、标准等频繁出台。这些已颁布、修订的法律，发布、实施的政策、标准等，一方面为组织开展数据安全治理工作，提升对重要数据和个人信息的保护能力提供了良好的指导参考信息和材料；另一方面，也是对组织提出了强制力的硬性要求。数据安全治理是个系统性工程，需要国家立法部门、政策制定部门、监管部门引导和组织数据处理相关单位相互协作，合力应对数字经济时代的数据安全风险和挑战，尽快为我国尽快建立和发展数据安全治理的国家法律法规和标准体系，保障数字经济健康有序发展。

参考目前金融、医疗、工业等不同领域以及多省市已经发布的数据安全相关的条例和制度，通过目前各个行业主管部门发布的十四五规划，也基本都明确了进一步细化数据分级分类等相关制度和联动机制，对相关制度进行统筹管理和执行。在安全标准层面，国家目前已将发布了通用的数据安全治理要求、个人信息保护相关标准，但在一些数据安全关键治理环节、关键业务场景、特定行业要求方面还有待进一步细化，明确数据安全治理各个活动指南以及数据安全评估方法和指标。通过 2022 年 3 月 6 日全国信息安全标准化技术委员会《关于发布 2022 年度网络安全国家标准需求的通知》（信安秘字〔2022〕47 号）中显示，2022 年计划制修订的 34 个标准中，数据安全和个人信息保护相关标准超过三分之一，也充分体现了国家完善数据安全治理相关标准体系的态度。标准和指南作为法律法规和政策落地实践的重要桥梁，可有效解决前述法律、法规的基本概念厘清、配套制度建立以及相关举措细则确定等问题。

综上所述，可以预期未来在法律法规和标准体系建设会日趋完善，“数据”“个人信息”“网络数据”等基本概念将会日渐清晰。各部法律所涵盖的范围中出现重叠的部分，会在具体落地的管理办法中逐步明确各种场景下进行数据处理活动时的具体落地方案，数据所有权和数据加工新成果的责权界定等问题会逐步解决。

6.1.2. 数据安全执法力度逐步增强

2021 年，多部数据安全相关法律法规面世，总体呈现以下三个特点：一是法律法规覆盖内容广泛，





数据安全法律框架逐步构建。随着《数据安全法》、《个人信息保护法》、《关键信息基础设施安全保护条例》的颁布和实施，有“小数安法”之称的《网络数据安全管理条例》也正在征求意见阶段，数据安全保护、个人信息保护、关键信息基础设施安全保护等均有法可依，数据安全法律框架逐步完善。二是由法律到地方性法规，地方结合实际做好数据安全治理工作。2021年，安徽、深圳、上海先后出台地方数据条例，结合地方的优势明确数据安全责任、促进数据资源高效利用。三是监管文件密集出台，数据安全落实到各行各业，如在移动互联网应用领域，2021年监管部门通报下架APP的数量高达数百款，“下架”成为数据安全年度热词。随着监管部门检测能力不断提升、执法覆盖面不断扩大，执法趋严是明确的监管态势，APP运营者贯彻落实《个人信息保护法》、《网络安全法》等相关法律，认真开展数据安全合规工作势在必行。分析APP被通报的理由，主要集中在违规收集个人信息、APP强制、频繁、过度索取权限、违规使用个人信息、超范围收集个人信息、欺骗误导用户下载五大类型，违法违规行为主要集中在数据收集和使用两个环节。APP运营者在数据安全合规中应当重视上述风险高发点，规范隐私设计，及时根据执法案例及立法更新情况调整隐私政策及相关文案内容。

通过《网络安全法》的落地执行过程可以看出，从法律出台到后期执法，中间为各个企业深入解读法律并进行整改、执法部门对评定准则和处罚尺度均需要进行适应。目前，国内已经出现数起针对数据安全法和个人信息保护法的处罚案例。随着我国数据安全法律体系的进一步完善，为后续立法、执法、司法相关实践提供了重要法律依据，可以预见，为数字经济的安全健康发展提供有力支撑，保护公民隐私和财产安全的呼声越来越高，围绕个人信息保护和数据安全领域的执法力度会越来越大，执法频率越来越高，且这一趋势还将延续。

6.1.3. 数据安全治理从合规驱动逐步走向自驱动

目前在大多数组织内，数据安全治理工作都是以合规为驱动，由信息安全技术部门或信息技术支持部门负责牵头推进和落实，尚未能得到高级管理层的充分重视。此种模式，无法建立起包括数据安全专家以及IT支持、人资、法律、财务、业务和市场、运营和维护、知识产权、保密等所有与数据安全有关部门的人员代表的数据安全治理团队，导致数据安全治理价值得不到体现。

随着数据规模的不断提升，以及组织生产依赖数据治理效果进行提升，数据安全治理工作应会逐步引起高层管理人员的充分重视。因此，除了强调满足数据安全合规要求对于组织维持当前业务的重要性外，将数据安全治理的意义与数据治理、数据资产的价值挖掘与实现、整个组织的数字化转型升级联系起来，能有效消除高层管理人员将数据安全治理视作组织的成本开销而非潜在资产增值机会的习惯性误解。因此，随着组织开展数据资产盘点，厘清自身拥有数据资源的价值。根据合规要求和标准规范对数据资产进行分类分级，结合自身的业务需求和人员组织架构，分析和评估现有和潜在的数据安全威胁和风险，以及自身业务的风险承受能力和容忍度，再根据评估的结果，科学地设计和规划适合自身情况的数据安全保护方案，真正让数据安全治理活动为业务有效开展保驾护航。





6.1.4. 建立持续迭代的数据安全治理体系

数据安全与企业经营活动联系紧密，对组织而言，一成不变的数据安全保护策略显然不足以应对高速发展中的新态势。数据安全治理体系包括数据、业务、安全、技术、管理等多个方面，并且应满足组织的战略一致、风险可控、运营合规等各种要求，其本身就属于动态变化中的活动。数据安全治理体系化建设，应该遵循循序渐进的原则，充分了解合规及行业要求，建设满足合规性要求同时，需要考虑灵活性、可扩展性及各阶段衔接性，使数据安全治理更加合理规范。从业界实践而言，数据安全治理体系的建设，组织早期经常会选择传统数据防泄露、数据加解密、数据库审计等简单数据安全的保护产品。后期随着数据治理方面的发展，对于数据的分级分类、流转溯源监测等需求逐步增加。从数据相关技术产品演进角度看，数据应用正在经历数据仓库、大数据平台、数据湖的不断演进，以及多种场景并存的实际需求。

因此数据安全治理体系作为一个新生事物，其发展演进也应该遵循持续迭代、不断完善的过程。各组织应在数据安全治理目标、成本控制和技术发展的多重约束下，对数据资产的运行情况、风险分布、数据流转需求等进行持续监管和分析，不断完善技术防护措施和管理体系。

6.1.5. 提升数据安全评估 / 审查自动化水平

《数据安全法》、《个人信息保护法》、《关键基础设施保护条例》等法律法规都明确提出对重要数据和个人隐私的数据处理活动定期开展数据安全风险评估及审查工作，并在安全评估与审查的指导下建立完善数据安全管理制度与管控体系。但目前组织独立开展相关数据安全评估工作确实存在困难，需借助于第三方服务机构对数据安全提供咨询、评估和运维服务。服务机构的评估与审查更多的是依赖调研、访谈、旁站、代码核查等人工手段开展，审查机关投入的人力和时间成本很大，且评估的全面性、客观性以及时效性得不到有效保证，面向数据及隐私的安全风险以及涉及个人知情权、同意权等合法权益的评估与审查，亟待提升自动化技术水平，依据自动化工具检测数据安全面临的各类脆弱性与安全威胁，以支持审查机关高效、准确的落实数据安全评估及审查工作，并可组织开展自评工作提供有效手段。

6.1.6. 新兴技术的逐步发展和应用，持续激发数据安全新需求

随着人工智能技术的飞速发展，2021 年“深度伪造”和“元宇宙”的概念进入大众的视野。从抖音到微博，多位大咖齐唱“蚂蚁呀嘿”的画面不断刷屏，这其实便是深度伪造技术的应用。深度合成的图像、视频、音频、文本等内容，例如热门影视剧片段、话题明星的换脸视频等等，具有极强的娱乐性与传播性，伴随着技术的发展成熟，越来越多的创作者在互联网上发布和分享深度合成内容，数量逐年高速增长。如果此技术被不法分子用于伪造音频、视频，实施诬陷、诽谤、诈骗、勒索等违法行为，甚至捏造国家政要言论扰乱社会与政治秩序。在未来“亲眼所见未必为实，亲耳所听可能为虚”，将对社会信任、媒体信任、政治信任产生巨大的影响。因此 AI 生成或修改各类数



据内容相关技术引发一系列数据安全风险的应对与解决方案已经提上了日程，如何完成该类数据的识别发现、取证固证、响应处置等，传统基于生物特征的鉴别方式越来越难以发挥作用，以 AI 对抗 AI 成为此类数据安全治理领域的发展方向。

2021 年，Facebook 宣布该公司的名字将更改为“Meta”，体现出其全力投入“元宇宙”的决心，作为一个通过数字技术构建的与现实世界平行的虚拟世界，虚拟人、数字人即是深度合成的主要应用，也是“元宇宙”的重要组成部分。任何一类元宇宙产品，从实体看，无论具体依靠何种技术手段和方法，对个人信息进行深度挖掘，以实现定制化的设计和运转。而元宇宙产品自诞生起，就同区块链等去中心化技术深度绑定，其应用范围跨越多个主权国家之间，因此带来了国家地理边界范围内对用户个体隐私数据的系统保障问题，跨越地理边界的敏感数据传输实现有效规制问题，以及在不同尺度的数据管制体系之间构建共有知识基础上的协调与管辖机制问题。如何在新时期通过构建生态化治理体系方式，有效保障国家数据安全免受新型威胁，成为未来数据安全治理的新领域。

6.2. 持续加强数据安全治理能力的倡议

6.2.1. 加快复合型数据安全人才的培养

数据安全治理活动涉及法律合规解读、业务知识理解、安全技术及管理等多维度，但法务人员、业务人员和安全技术人员各自有擅长的领域，都存在知识短板。企业中满足法律、业务、技术、管理四大核心基础知识领域的人才十分稀少，只能通过长期业务实践自行培养相关人才，培养周期长、难度大。承载数据安全行业相关人才培养任务的主体主要在于高等院校、高等院校、职业院校、专业认证培训机构。高校和科研院所培养后的人才需要实践的打磨，通过丰富的实际业务场景经历才能见效，难于满足当下人才紧迫的需要。

数据安全治理的竞争，归根结底是人才竞争。缓解数据安全人才缺口问题的长期解决之道是做好顶层设计、政策供给和机制建设，开展全民数据安全意识教育和培训，建立有层次、标准统一的人才发展体系。国家、行业应鼓励院校开展数据安全人才培养计划，立足于结合创新型人才培养要素的宽口径基础教育，培养网络空间安全与法学的复合型人才。专业培训机构或网络安全公司的培训部门，结合实际和技术趋势，在培训计划中加入数据安全相关法律知识，有助于培养出一批具有扎实的数据安全领域技术与理论知识，又懂得法律与管理的复合型人才。企业自身可建立师徒制，选拔动手能力强、基础较好、符合实际需要的数据安全人才开展传帮带，师傅带徒弟不仅是一种高效的人才培养与孵化方式，更是一种可以彰显团结协作精神的企业文化。

6.2.2. 持续加强数据安全技术研究

技术是数据安全治理的核心要素和驱动力量，要充分利用各种先进技术手段和工具来保护数据安全。在大数据、云计算多种新技术出现的背景下，企业数字化转型升级加快，业务持续发展扩大。但就目前来说，业界普遍应用和研究的隐私计算技术正处于快速迭代的发展阶段，隐私计算产品的





算法协议差异化较大，难以形成统一的算法安全基础。隐私计算产品所使用的算法协议多种多样，各自的协议安全基础也不相同：多方安全计算、同态加密等密码学算法基于数学与密码学基础，联邦学习等隐私机器学习安全基于机器学习理论、差分隐私和相关密码学协议，可信执行环境则更多依赖于硬件厂商的安全技术。在数据追踪溯源方面，数字水印技术对处理资源和存储资源的高占用、高依赖，无法大规模应用于云计算、大数据等大量数据汇聚的场景，数据血缘技术产业化应用尚不成熟，技术能力仍待发展。

建议在产学研技术合作研究方面，积极推进多方安全计算、联邦学习、可信执行环境等隐私计算技术的发展和落地，并加快数据安全与人工智能、区块链的融合研究，尤其需要与深度学习算法模型等技术结合形成智能化、动态化安全防护闭环。同时在权限界定、价值挖掘和创新应用等核心技术上进一步拓展，形成更加丰富的解决方案，弥补现在技术防护供应能力不足的问题，更加有效的促进数据开发利用与安全防护的平衡发展。

6.2.3. 发挥行业联盟推动数据安全治理产业发展的重要力量

单一企业往往聚焦在某一技术领域开展工作，无法同行业主管部门建立直接连接，对于国家政策、技术发展的认知也相对浅薄。在“1+1 > 2”的原理被日益验证的时代，数据安全相关产业联盟可通过收集行业内数据安全问题并展开研讨宣贯，推动行业相关标准发布，保证行业内对数据安全相关问题解读的一致性，建立“标准/定制”、“现场/远程”的立体化数据安全服务体系。同时，通过联盟组织不同类型机构互相交流、协作，加强各服务机构服务方案的整合，推动产学研一体化的闭环的数据安全服务链，有力推进国内数据安全治理体系化的发展。

6.2.4. 加速推进数据安全治理国际化

数据安全治理是数字时代全球面临的共同问题，网络攻击从数字空间延伸到物理空间，继而对网络安全提出了严峻挑战，必须筑牢网络安全防线。中国这样一个网络大国、数据大国，数据安全治理应结合国情，提出中国特色的数据安全治理方案，推动各国共识下的数据安全国际规则发展，对数字经济合作全球化具有重要促进意义。2020年9月，在“抓住数字机遇，共谋合作发展”国际研讨会上，我国提出了《全球数据安全倡议》，提出“秉持发展和安全并重的原则，平衡处理技术进步、经济发展与保护国家和社会公共利益的关系”，并呼吁各国应尊重他国主权、司法管辖权和对数据的安全管理权。

目前世界各国加强数据安全立法，保护涉及本国国家安全、公共安全、经济安全和社会稳定的重要数据及个人信息安全。为了维护开放、公正、非歧视性的营商环境，有效解决数据出境不好管和不同国家法律长臂管辖问题，推动实现互利共赢、共同发展，无论企业还是个人，都应该严格遵守所在国法律法规，尊重他国主权、司法管辖权和对数据的安全管理权，确保中国企业顺利出海。同时，2021年10月，为了规范数据出境活动，保护个人信息权益，维护国家和社会公共利益，促进数据跨境安全、自由流动，中央网信办出台了《数据出境安全评估办法（征求意见稿）》，也



对数据在国际间安全流转进行了约束。

因此，随着国际间贸易来往和数据流转越来越频繁，数据安全治理不仅仅是国内的事情，更要做到国际化，既方便走进来，也便捷走出去，促进数字经济更加高速的发展。





附录

A. 词汇表

数据生命周期：是指在企业或组织开展业务和进行经营管理的过程中，某个集合的数据从产生或获取到销毁的全过程。参照数据参考《信息安全技术 数据安全能力成熟度模型》（简称 DSMM），将数据生命周期划分为采集、传输、存储、处理、交换、销毁的六个阶段。特定

数据经历的生命周期由实际的业务场景所决定，并非所有的数据都会完整的经历上述六个阶段。

数据处理活动：根据《数据安全法》第三条规定，数据是指任何以电子或者其他方式对信息的记录。数据处理，包括数据的收集、存储、使用、加工、传输、提供、公开等处理活动场景。

合规：对数据安全所适用的相关法律法规的符合程度。

DCAP: Data Centric Audit and Protection, 以数据为中心的审计与安全防护。一系列能够集中监控和管理用户与特定数据相关行为的数据安全技术的统称。

DGPC: Data Governance for Privacy, Confidentiality and Compliance, 以隐私、保密和合规为目标的整体化数据治理。由微软公司的可信计算研究部门于 2010 提出的一套数据治理方法，帮助企业或组织能够以统一、跨学科的方式来实现隐私、保密和合规的三个目标，避免让组织内不同部门各自独立解决三个割裂的问题。

DLP: Data Leakage Prevention/Protection, 数据泄密防护, 又称为数据防泄露。采用内容分析引擎, 利用关键字、正则表达式、文件指纹、自然语言处理等数据识别技术, 对外发数据进行解析与扫描, 实时识别、监控、保护企业或组织的敏感数据; 对即将发生、正在发生的泄露敏感数据行为, 按照预置策略及时阻断并告警, 防止敏感数据传输到外部, 有效避免数据泄露带来的安全风险, 实现对外发敏感数据的可知、可见、可控的一体化解决方案。

DSG: Data Security Governance, 数据安全治理。由全球信息技术领域的权威咨询公司提出的帮助企业或组织应对数据安全挑战的一套数据安全能力建设理念和方法论。

MPC: Multi-Party Computation, 多方计算。使多个非互信参与方之间可以在数据相互保密的前提下进行高效数据整合和计算, 做到既使用多源数据进行指定的计算任务, 又保证使用过程中数据不被泄露, 包括数据使用方和计算任务执行平台都无法接触到原数据的明文, 同时又能保证最终的计算结果与基于原数据明文的计算结果完全一致, 真正实现数据的可用而不可见。

UEBA: User and Entity Behavior Analytics, 用户和实体行为分析。以用户为中心分析用户行为, 关联端点、网络和应用行为, 并将这些数据与威胁情报相关联, 通过自动学习正常行为来发现异常行为。

撞库：攻击者利用大量用户会在不同的网络服务平台使用相同用户名和密码的特点, 收集并利用已泄露的批量用户和密码信息, 逐一尝试登录各网络服务平台或系统, 把碰撞成功的有效身份认证数据保存下来以用于后续攻击。





B. 近似概念的联系与区别

B.1 数据安全、信息安全、网络安全

国标《信息安全技术术语标准》（GB/T25069-2010）将信息安全明确定义为“保护、维持信息的保密性、完整性和可用性，也可包括真实性、可核查性、抗抵赖性、可靠性等性质。”如果仅是从字面上将定义中的保护对象由“信息”变更为“网络”和“数据”以分别引申出“网络安全”和“数据安全”的概念，对于理解“数据安全”与“信息安全”和“网络安全”的关联和差异并无太多帮助，为此，下面将从每个概念所涉及内容和范畴的角度区分三者。

在大数据的概念出现以前，广义的信息安全范围最广，能够涵盖网络安全和数据安全。狭义的信息安全则在不同语境中有多种不同的含义。例如，当信息安全特指为保障内容健康合规而封禁黄赌毒等有害内容的发布和传播时，它与网络安全和数据安全的概念几乎完全独立；而当信息安全特指企业或组织机构防范内部秘密信息泄露时，它就成了网络安全和数据安全两个研究范畴的子集。

早期的网络安全（Network Security）概念通常泛指保护由计算机网络（特别是接入互联网的网路）连接的软硬件信息基础设施能够持续正常运行，防范外部攻击者对信息系统及其中存储的数据进行破坏和篡改或者对秘密数据进行窃取。同时期的数据安全一般被理解为对信息系统中重要数据进行保护，因此，可视为网络安全范畴的子集。

近年来，随着移动互联网、物联网、人工智能等技术的飞速发展，大数据的应用和影响范围越来越广，人类正在全面转型跨入数字化社会。在这种时代背景下，网络的概念开始由传统的计算机网络不断向云、边缘、终端等新的衍生概念延伸，早期的网络安全也相应演变为与陆、海、空、天等国家安全概念相提并论的网络空间安全（Cyber Security）的简称。

数据安全概念的演进和发展速度比网络安全有过之而无不及，甚至仍会伴随数字化转型的浪潮加速向前奔涌，不断突破人类现有的认知和想象。当下一般将数据安全解读为以数据为中心，保护数据在其生命周期（包括数据的采集或生成、传输、存储、处理或使用、交换、销毁等众多流转环节）内的安全性，一个形象比喻就是“为数据配备贴身保镖”——既要不限数据自由流动，又要确保数据走到哪，安全措施就覆盖到哪。至此，数据安全的概念范畴尽管仍与网络安全拥有大量重合，但已不再完全是网络（空间）安全的子集。

总体而言，网络安全更侧重于计算资源与网络环境的安全防护范畴，而数据安全则被更多赋予了促进与保障数据开发利用、数据价值释放的重要时代使命，与数据经济社会下的业务创新关联更加紧密，追求数据发展与安全保障的最佳平衡。

B.2 数据安全治理与数据治理

数据治理是指从使用零散数据变为使用统一数据、从具有很少或没有组织和流程到企业范围内的综合数据流转、从尝试处理数据混乱状况到数据井井有条的一个过程。数据治理的作用就是确保企业的数据资产得到正确有效的管理，从组织架构、原则、过程和规则等方面确保数据管理的各项



职能得到正确的履行。数据治理其实是一个关注于信息系统执行层面的体系，这一体系的目的是整合信息技术支持（IT）部门与业务部门的知识和意见，通过一个类似于监督委员会或项目小组的虚拟组织对企业的信息化建设进行全方位的监管，这一组织的基础是企业高层的授权和业务部门与IT部门的建设性合作。从范围来讲，数据治理涵盖了从前端事务处理系统、后端业务数据库到终端的数据分析，从源头到终端再回到源头形成一个闭环负反馈系统（控制理论中趋稳的系统）。从目的来讲，数据治理就是要对数据的获取、处理、使用进行监管（监管就是我们在执行层面对信息系统的负反馈），而监管的职能主要通过发现、监督、控制、沟通、整合这五个方面的执行力来予以保证。

根据国标《信息技术服务治理第5部分：数据治理规范》（GB/T34960.5-2018），数据治理被明确定义为：“数据资源及其应用过程中相关管控活动、绩效和风险管理的集合。”这项标准还进一步给出了数据治理的目标、任务和框架等：数据治理是“源于组织的外部监管、内部数据管理及应用的需求”；数据治理的目标是“保障数据及其应用过程中的运营合规、风险可控和价值实现”；“组织应通过评估、指导和监督的方法，按照统筹和规划、构建和运行、监控和评价以及改进和优化的过程，实施数据治理的任务”。

在数据治理的三点目标中，运营合规和风险可控是约束条件，价值实现是追求的结果，合规和风控既是使数据能最终实现价值的前提，同时又是必要的保障。为此，数据治理框架包含了数据管理和数据价值两套体系，前者统摄合规与风控，后者则支撑价值实现。在数据治理的框架下，数据安全是数据管理体系的重要组成部分，企业或组织在实施数据治理的过程中，应“制定数据安全的管理目标、方针和策略，建立数据安全体系，实施数据安全管控，持续改进数据安全能力”。此处的数据安全概念相对狭义，基本仅限于为数据及其价值实现过程提供安全保障机制的范畴。

如果基于上述狭义的数据安全概念，将数据安全治理相应解读为以保护数据及其价值实现为目的而采取的风险评估和安全管控活动，那么，数据安全治理就可视作数据治理概念下专注于安全方面的子集。但从实际操作上来看，两者之间又有很大的不同：

- 从发起部门来看，数据治理主要是由IT部门在驱动，而数据安全治理主要是由安全合规部门在驱动。当然两者的成功都要涉及到业务、运维和管理部门甚至公司最高管理决策层。

- 从目标上看，数据治理的目标是数据驱动商业发展，提升企业数据资产价值。数据安全治理的目标是让数据的使用更安全，保障数据的安全使用和共享，实质也是保障数据资产价值。

- 从工作内容产出上看，数据治理工作产出的一个核心成果就是数据质量提升，通过对数据进行清洗和规范的过程，获得高质量的数据。数据安全治理的重要产出包括完成对企业数据资产的分类分级，制定合规的安全访问策略并规划适宜的管控措施。

- 从数据资产梳理上看，数据治理中资产梳理的主要产出物是元数据，即赋予数据上下文和含义的参考框架。数据安全治理中的资产梳理，则要明确数据分类分级的标准，弄清敏感数据资产的分布、授权和访问状况。

尽管当前在数据治理中也在不断加强对数据安全方面的要求，但相对数据安全治理而言，数据治理中的安全实践还是常被置于从属角色，如同信息安全在信息化建设中的角色一样，不够系统和深入。近年来，随着包括大数据在内的数据在数字化社会中的重要性不断提升，国家已将数据上升到新型生产资料和创新要素的高度，数据对安全形成影响的广度和深度已超越单纯为数据掌握者保



护数据自身价值的范畴，例如对个人（数据主体）隐私保护的问题就牵涉到了社会伦理以及关于数据权益的法律认定，而数据出境等问题更是需要在国家安全层面全盘统筹考虑。

本白皮书讨论的数据安全治理概念，将主要关注以保护数据及其价值实现过程安全为目标的风险管控活动，但在合规性方面，则将以全面覆盖和满足国家关于数据安全的所有法律、法规、政策、规范和标准的要求为目标。数据安全治理工作既可在数据治理的框架下作为一个环节或子过程来开展，也可直接依照本白皮书提出的数据安全治理框架独立实施。

B.3 数据安全治理与数据安全管理的区别

“治理”对应的英文单词是“governance”（源自拉丁文或古希腊语“steering”一词，原意是“引领导航”），是指遵照具有共识的指导原则，通过协调和配合共同追求一致目标的过程。“管理”对应的英文单词则是“management”，是指为达到组织的既定目标而以人为中心对组织拥有的资源进行有效的决策、计划、组织、领导和控制的过程。

治理的概念强调协调和合作，一般难以表述成一套严格的规则条例或正式制度，而通常是以一种方法论的形式呈现。管理的概念则强调控制和执行，通常会以形成制度和条例规范的形式加以表述和落实。

在数据安全范畴内，使用数据安全治理的概念比使用数据安全管理的概念，能够更好地适应当前的现实情况：一方面，人类正在全面跨入数据时代，为积极顺应这一势不可挡的历史潮流，国家和各层级的组织机构都迫切需要加强数据保护的意识，尽早启动和落实关于数据安全的准备和建设；另一方面，人类也仅是刚刚看到数据时代和数字化社会形态特征的端倪，数据无论是作为资产、生产资料，还是创新要素，都是方兴未艾的新事物，必然要经过一个认知逐渐深入和全面的复杂过程。在这个过程中，使用治理而非管理的思路来指引数据安全建设，无疑拥有更好的灵活性、丰富性和包容性，因而也更有利于平衡对数据进行安全保护和在数据使用和价值挖掘方面激发创新。

B.4 数据安全治理与数据安全成熟度模型

国标《信息安全技术数据安全能力成熟度模型》（GB/T37988-2019）给出了一个全方位评估企业或组织机构的数据安全能力成熟度的三维立体模型——DSMM。该模型将数据安全的能力成熟度分成从1到5共五个等级（由弱到强依次为非正式执行、计划跟踪、充分定义、量化控制、持续优化），分别从组织建设、制度流程、技术工具、人员能力四个方面对通用安全和构成数据生命周期（数据安全过程）的六个阶段（采集、传输、存储、处理、交换、销毁）及其进一步细分的三十个过程域内各自达到的数据安全能力成熟度等级进行评估。

DSMM适合用来作为评估组织数据安全能力的方法和标准，亦可在组织开展数据安全能力建设的过程中被用作参考目标和依据。然而，DSMM并未给出指导数据安全能力建设的具体方法和操作步骤，而这恰好是数据安全治理的核心内容和价值所在。

数据安全治理和DSMM的区别主要体现在以下方面：





● 数据安全治理是以数据的分类分级为核心，进行安全策略的设定；DSMM 是以数据的生命周期为核心，寻求安全策略的覆盖。

● 数据安全治理体系化地建立数据使用或服务的人员角色，根据角色对数据使用的主要场景，提供建议的安全措施以及所要使用的安全工具；而 DSMM 更多的是提供一种评估方法，看数据使用的过程中，企业是否定义了明确的控制措施，并无具体化的方法指导。

● 相比 DSMM 完全围绕数据生命周期展开，数据安全治理通常反对在数据生命周期中不区分化地采用安全措施，而是更强调根据数据使用场景和数据使用需求，针对性地选用适宜的安全措施。比如在开发测试环境，需要采用脱敏的技术获得高仿真数据，在这种场景下，数据安全治理就会弱化推荐审计、管控和加密等措施。

● 数据安全治理将安全体系的构建，明确归纳为安全政策的制定，技术支撑平台的建设，对安全政策执行有效性的监督，安全政策的改善这一持续循环过程；而 DSMM 对安全成熟度进行了等级化分级，将持续改善定义为了最高级别，不像在数据安全治理理念中将持续改善视作标准组成部分之一。

● 数据安全治理本质上是一种方法论，帮助企业如何迅速构建一套行之有效的数据安全体系；而 DSMM 是一种评估方法，像一个考试一样，帮助企业或监管机构评价和了解企业或组织的当前安全能力状况。

从某种意义上讲，数据安全治理和 DSMM 之间是相辅相成、互助互补的关系。一方面，DSMM 能够为数据安全治理在事前提供开展工作的目标和方向，在事后提供对治理效果的评估方法；另一方面，数据安全治理又为追求 DSMM 描述的高级数据安全能力目标和弥补 DSMM 发现的数据安全短板提供了一套切实、可操作且行之有效的方法论。

C. 数据安全关键技术简介

C.1 新兴数据安全技术

C.1.1 自动化数据分类分级技术

C.1.1.1 技术概述

数据安全治理的首要步骤是数据资产识别与数据分类分级，需要对数据的分布、位置、类型、级别进行摸底梳理，明确保护对象后，在此基础上通过区分不同的访问者身份，制定不同的数据安全防护策略，采用相对应的安全防护技术，做到安全防护有的放矢，确保数据的“用”，“护”结合。

近年来，国家密集出台了多项数据安全法律法规，包括《网络安全法》、《数据安全法》、《个人信息保护法》等，明确要求建立数据分类分级保护制度，对数据实行分类分级保护，并基于数据分类分级确定的国家核心数据和重要数据目录，进行重点保护。

但鉴于信息系统众多，敏感数据本身数据量大，存在的范围广且分布零散，质量良莠不齐，命名规范不统一等诸多问题，依赖人工手段进行数据分类分级，导致了数据资产梳理难、数据分级分

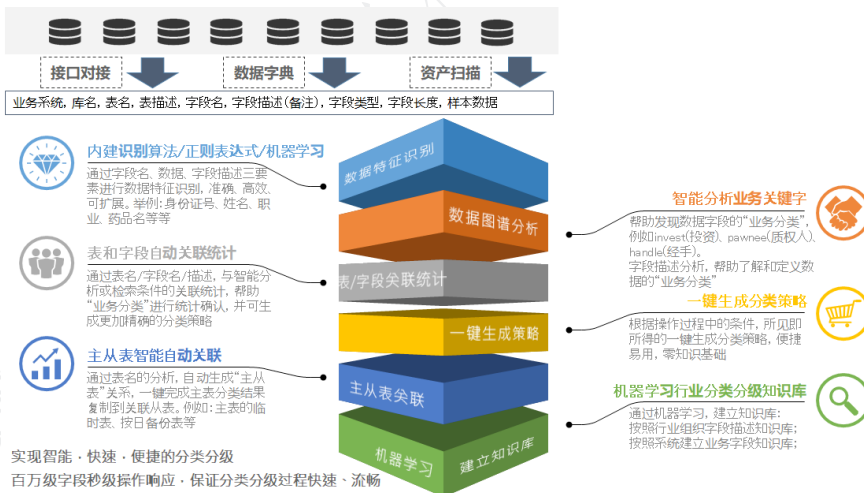
类耗时费力、敏感数据识别准确率低，且面向数据不断的产生、加工转换、流转过程，持续数据分类分级处理难，亟需通过自动化的技术手段梳理企业核心业务系统的数据资产，提高数据分类分级效率与质量，降低数据分类分级强度和难度，全面盘点数据资产并开展敏感数据识别和数据分级管理，为后续建立数据生命周期中的安全管理和技术防护提供依据和参考。

数据库作为数据资产的重要载体，针对存储在数据库中的结构化数据分类分级是重中之重，面向分布在数据库中的海量数据资产，如何快速、自动化、高质量的落实数据分类分级？是后续基于数据分类分级成果，面向数据在结构化和非结构化载体上的流转，形成分类分级防护策略，实现数据分类分级防护的核心基础。

自动化数据分类分级的核心是如何准确识别数据表/字段的业务含义并与行业数据分类分级标准模板智能建立绑定关系？从而通过绑定关系建立数据资产的分类分级标签。

C.1.1.2 功能架构

整体数据分类分级技术由数据资产发现、数据特征识别、数据图谱分析、数据分类分级策略集生成、主从表自动关联与智能化行业数据分类分级知识库建立等核心能力构成，实现智能、快速、便捷的数据分类分级。



1) 数据资产发现

数据分类分级，首先要发现数据资产，通过资产扫描、元数据接口对接、数据字典导入等方式，收集、发现数据资产，其中资产扫描方式是以任务形式对指定网段内的各类数据库进行嗅探，自动发现数据库资产并进行备案，再通过对指定数据库进行扫描，识别库中的表名、表描述、字段名、字段描述、字段类型、字段长度与样本数据；数据字典导入和接口方式可通过文件导入或与元数据系统的接口对接方式获取数据资产。

2) 数据特征识别

对发现的数据资产，面向字段名、字段描述、数据样本三要素，通过敏感数据识别引擎进行自动化的敏感数据特征识别，与数据分类分级模板中明细分类建立映射关系，快速生成数据分类分级



策略，完成数据资产分类分级的预处理。

3) 数据图谱分析

针对敏感字段和字段描述，通过运用谓词切分与语义分析技术，帮助了解和定义数据的“业务分类”，进一步通过表和字段自动关联分析，对业务分类关键字进行钻取、分析、统计，帮助对“业务分类”进行统计确认，并与数据分类分级模板中的明细分类建立映射关系，生成更加精确的数据分类分级策略。

4) 数据分类分级策略集生成

围绕数据特征识别与数据图谱分析中对“业务分类”还原操作形成的数据分类分级策略，进行梳理确认，所见即所得的一键形成数据分类分级策略集，并依据策略集中定义的映射规则，对未分类的数据资产进行策略匹配，对符合策略集的数据资产批量、自动化生成数据分类分级标签。

5) 主从表自动关联

通过表名分析，自动形成主表与临时表、备份表等从表的关系，将主表数据分类分级结果自动转移同步到从表相同字段的标签上。

6) 数据分类分级知识库

基于已分类分级的数据资产结果集，将结果中的表名、字段名、字段描述、数据特征等作为向量，进行机器学习建模，并根据知识库模型智能预测大规模数据的分类分级打标，实现数据自动分类分级，输出数据分类分级报告。并通过面向行业或系统的持续模型优化，形成数据分类分级模型知识库，向全自动数据分类分级迈进。

C.1.1.3 关键技术

1) 数据特征自动识别引擎

通过主动发起数据探测行为，对数据源或结构化文件中的数据进行分析，运用字段名、字段描述、内容关键字、正则表达式、自然语言识别等规则匹配，进行数据特征的智能识别，通过该技术，可以判断输入数据的特征是属于个人姓名、身份证号、手机/座机号、职业、地址、纳税人识别号等类型。

通过数据资产自动发现与识别技术，可以对需要分类分级的数据做预处理，加速人工判别过程。

2) 谓词切分与语义识别技术

数据分类分级所依托的重要因素，即为数据的业务属性，而数据的业务属性往往存在于字段的命名方式、字段描述或字段备注中，通过对这类数据基于文本上下文进行合理、准确的谓词切分，可以辅助人工进行判断，也为自动匹配数据分类分级规则提供了重要依据。例如“CARD_NUM”或“借款合同到期日期”，切分结果应为 CARD_NUM->CARD+NUM，借款合同到期日期->借款合同+到期+日期，通过切分结果，获得 CARD（银行卡信息）、借款合同等业务属性。

3) 知识库与匹配技术

对于某个特定行业，其“专业术语”是可被枚举、收集的，在结合谓词拆分的基础上，可以将该行业所涉及的谓词进行充分的归纳记录整理，并按照法规与行业标准，对这些谓词与谓词组合“归属于”哪个数据分类分级结果进行收集统计，将收集统计结果形成行业知识库。





一般来说，知识库分为行业知识库与客户自身的知识库，知识库越细化，对于知识库的自动匹配结果越准确。

当形成知识库后，对于相同命名规范和谓词切分的元数据成果，可以通过知识库自动匹配的机制，直接自动化生成初步数据分类分级结果，在此基础上再结合人工梳理，大幅度提升工作效率。

4) 机器建模与匹配技术

通过引入机器学习，数学建模等技术，采用监督学习模型，可以更进一步加速数据分类分级效率与自动化程度。对于遵循同样逻辑构造的大量数据，可先对其一小部分采用半自动数据分类分级技术先行处理，经过人工审核后，将这部分数据提交至机器学习引擎进行建模，反复训练达到较高匹配率后，该模型即为此场景下数据分类分级最佳匹配模型，并利用该模型完成后续数据的智能分类分级匹配。

C.1.1.4 应用场景

面向《数据安全法》的数据分类分级保护要求，金融、电信、工业、政务等各行业纷纷出台数据安全分类分级规范，展开基于分类分级的数据安全防护工作。

以金融行业应用场景为例，针对金融行业的数据中心，汇聚了海量的数据资源，为数据的处理、加工、使用、共享和开放场景提供了数据基础，涉及众多 RDMS、NoSQL 等各种异构的数据库存储资源，如何动态、按需保护这些数据库中的敏感数据资产？需要引入数据分类分级防护思路，依据金融行业数据分类分级标准《金融数据安全数据安全分级指南》（JR/T0197—2020）及《个人金融信息保护技术规范》（JR/T0171—2020），运用数据资产梳理及数据分类分级工具，通过智能数据分类分级技术，快速、准确、高效对海量数据资产进行分类分级标签化处理，再依据数据分类分级结果，驱动形成数据脱敏、数据加密、数据风险审计等策略，与脱敏、加密、审计等安全防护系统的联动，实现面向金融行业数据中心的众多应用访问及开发、运维人员的数据库访问场景的动态安全管控。

C.1.2 数据安全运营技术

C.1.2.1 技术概述

伴随移动网络、数字经济、电子政务等领域的持续高速发展，数据的价值和作用日益彰显，成为世界各国新兴、重要的生产资料，攸关国家安全与社会稳定。然而，面对因黑客攻击、内鬼泄密、操作失误、违规使用等导致的数据安全事件及各类隐患、风险、问题，过去市场上常见的数据安全方案大都通过提供单一化的功能解决客户在某个方面的问题和需求，这让不同安全产品间的数据“孤岛”现象凸显，继而影响到数据安全建设在落地执行过程中的效率、质量乃至成本投入等诸多层面。

在当前数据的价值利用与安全防护持续并行的需求下，将分散的安全产品能力进行整合，以“平台化、体系化、可视化、实用化”为出发点，提供具备的一站式管理能力，对包括数据资产梳理、数据库防火墙、数据库审计、数据脱敏、数据库运维管理、数据库加密等在内的各类数据安全产品进行“统一部署、统一监控、统一管理、统一运营”，将为数据安全运营工作的开展提供有力的支撑。



C.1.2.2 功能架构

数据安全运营技术为了满足项目需求建设，分为基础能力层、接口层、监测分析引擎层、安全管理和安全运营五个层次，如下图所示：

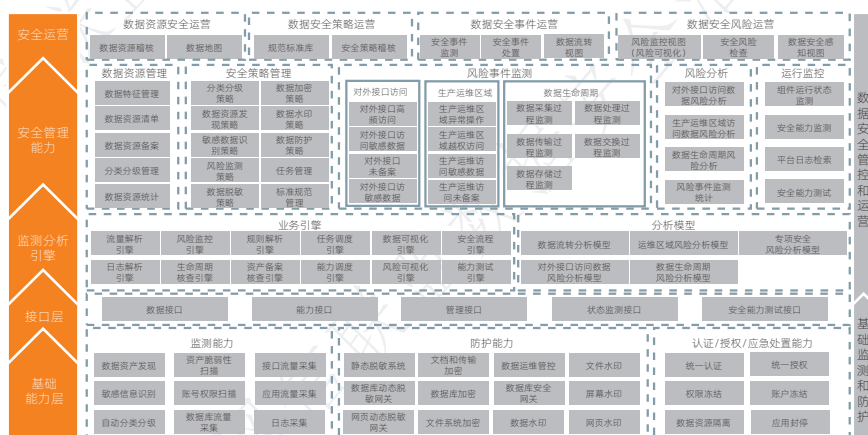


图 49 数据安全运营管控技术架构图

在用户可视化层横向分为数据安全运营和数据安全管理两层能力，纵向分为数据资源、数据安全策略、数据安全事件、数据安全风险共四个维度，另包含各组件的运行监控能力。

检测分析引擎层负责控制各类引擎来完成上层应用的各种功能。例如采集状态监控引擎，就是在收集日志后利用引擎做各种业务处理和相关数据分析，获得每个节点上对数据访问行为的状态监控。

接口层主要包含不同业务场景使用的五种接口。例如状态监测接口用于各个采集设备的工作情况，实时监控采集设备是否正常运行。策略配置接口是用于把规则向每台采集设备下放的，用于配置每台设备上对数据监控的策略。

最底层是基础能力层，主要为数据安全监测设备和探针、数据安全防护设备和探针，提供原始数据的输入。

数据安全运营技术实现以下核心功能：

1) 数据资源安全运营

针对数据库资产及敏感数据的分布及流转状态进行实时呈现。

● 数据资产视图

数据资产视图主要为对已注册的数据资产进行统计展现的结果，并进行相应的展现，展现过程体现为数据资产总数，并根据资产的重要度不同进行分级展现。并可根据数据资产类型分布主要物理位置、逻辑地址作为展示内容，为用户做整体展现。

● 敏感数据视图

敏感数据分布视图主要是基于敏感数据资产检测结果统计分析得出，采用多种维度、多种图表方式（比例图、趋势图、热度图等）进行展示，其中展示的内容包含敏感数据类型分布、敏感数据量展示、敏感数据属性展示、敏感数据类型分布主要物理位置、逻辑地址展示等。可针对非结构化数据、结构化数据进行独立展现。



● 数据资产访问热度视图

数据资产热度视图主要是对于使用比较频繁的数据资产进行视图展示。可根据数据资产活跃度进行相应的展示，并提供高热资产视图、低热资产视图以及高热/低热数据资产的详细信息，能够快速了解到当前数据资产的访问频度。

● 数据资产流向视图

数据资产流向视图用于对于数据资产与账号的访问关系视图。可通过账号访问的时间以及路径对数据资产的调用进行展示，并可根据时间范围进行选择。并建立全局数据资产流向展示，能够对某个数据资产或某个账号使用过的数据路径进行高亮展示，从而快速定位到账号在使用数据的业务流程，当出现异常访问时可给予相应的安全告警。

● 数据资产账号应用视图

提供数据资产账号管理功能。可针对操作数据资产的账号进行全局管理，包括提供资产账号详表，内容包含账号、准入状态、资产名称、资产类型、来源方式、活跃度等。建立账号行为热图，可针对日常账号访问的频次、时间进行矩阵式排序，从而直观看到账号的使用情况。根据账号使用的基线分析，对于异常账号访问、异常操作访问进行管理，发现数据资产风险，并给予通报。

● 应用行为流向视图

数据应用行为视图用于对于数据资产与应用的访问关系视图。可通过应用访问的时间以及路径对数据资产的调用进行展示，并可根据时间范围进行选择。并建立全局数据资产流向展示，能够对某个数据资产或某个应用使用过的数据路径进行高亮展示，从而快速定位到应用在使用数据的业务流程，当出现异常访问时可给予相应的安全告警。

2) 数据安全策略管理中心

建设数据安全策略管理中心，实现统一化数据防护策略管理。

● 安全策略集中设置

➢ 统一策略配置

集中展示各个数据安全组件状态，对各个数据安全组件涉及的识别类及防护类安全策略进行统一配置，形成安全策略集中管理。

➢ 统一策略发布

针对完成配置的各安全组件的策略，提供策略发布能力，将安全策略发布到相应的安全组件，以实现策略的落地执行。安全组件包括数据资产梳理、数据库防火墙、数据库审计、数据脱敏、数据库运维管理、数据库加密等。

➢ 策略状态统计

能够根据单独的数据安全组件进行已添加策略、已执行策略、已失效策略等多个维度状态展示，并能够根据策略执行的成功率给出业务风险值参考等。

● 安全策略核查

制定核查策略模板，基于模板，检验、核查数据安全策略设定的完备性及执行效果。

➢ 核查策略模板管理

提供内置检查模板，依据数据全生命周期安全场景不同提供一个或多个安全策略组，形成一套





策略评价 KPI 指标模型库，并能够自动匹配关联对象策略；模板可进行创建、变更、停用、同步，并提供模板的查询检索功能，包括不限于模板编号、模板名称、模板类别等。

➤ 数据安全策略核查

提供安全策略检查功能，依据选定的核查策略模板，展开安全策略核查工作，管理员执行过程中对此策略具备管控权限，结果自动上报及汇总，并提供分析报告及评价报告。

● 数据安全风险监测类视图

实时监测绘制数据安全风险相关视图，直观可视化呈现安全事件及风险趋势。

➤ 数据安全管控全景图

数据安全管控全景图用于展现数据安全管控节点以及防护手段的逻辑架构图，对于数据安全组件部署的位置、已执行的数据安全组件策略、已完成的策略结果等进行展现，并能够根据组件的类别进行单独展现。

➤ 数据安全事件视图

数据安全事件视图根据用户环境实际的事件内容直观展现当前环境中出现的数据安全事件的统计结果，包含数据资产事件维度、应用事件维度、账号事件维度，并可根据不同的时间期限进行展示，让客户快速、便捷的看到数据安全事件。

➤ 数据安全风险趋势视图

数据安全风险趋势视图适用于数据安全事件进行分类统计，并以曲线图展示，可根据不同的时间维度进行展示，也可根据不同的事件种类进行选择展示，并根据事件发生区间进行颜色区分，从而快速判断数据安全事件风险。

● 统一数据安全事件管理

➤ 事件响应处置

对各安全组件上报的安全事件按事件等级进行集中呈现，可钻取查看事件详情，并提供与工单、企业即时消息等系统的联动，进行统一告警通知，驱动针对安全事件的响应处置工作。当完成响应处置后，对安全事件记录处理人、处理结果，关闭安全事件，并依据事件典型性，定义处置措施标签集，记入事件处置知识库。

➤ 事件溯源

针对指定的安全事件，提供事件关联功能，从数据资产流动路径分析角度出发，对安全事件涉及数据资产流转过程进行追踪溯源。

➤ 事件处置知识库

对各类安全事件的处置措施按照事件分类进行归集，形成事件处置知识库，提供按分类、标签、关键字等丰富检索条件对处置知识库进行查询检索，以辅助支持安全事件处置决策。

C.1.2.3 关键技术

对各个设备侧的数据安全设备 / 系统产生及上报的流量日志、事件日志及安全事件日志等各类日志进行集中化大数据采集、汇聚、存储与归一化解析处理，形成标准化的统一数据格式，以作为支撑数据资产及数据分类分级纳管中心、数据安全策略管理中心、数据分布流转及风险监测中心的应用、





分析、与可视化呈现的数据基础。

对统一安全策略管理中心发布的安全防护策略进行解析和适配，形成各安全系统/设备可识别的安全策略，以实现安全防护策略的落地执行。

C.1.2.4 应用场景

数据安全运营技术针对政务、金融、能源、制造业、教育、医疗等行业的数据中心的数据安全运营工作全面的承载，将参与数据安全运营各角色的具体工作通过任务流转的方式全部纳入集中管理，形成从法律法规解读、数据库资产梳理、敏感数据发现、数据分类分级、防护策略统一下发、事件监测与处置、到策略稽核与完善的日常化、集中化、规范化、流程化的数据安全运营体系，推进组织数据安全运营各项工作的落地。

通过安全运营技术的应用，实现数据安全的集中化管理，形成了数据风险趋势的全面感知，对数据安全事件整体态势做到心中有数，形成了强有力的管理抓手，全面提升组织数据安全能力和安全防护能力，保障数据安全稳定的使用和运行，为数据价值的充分挖掘和利用奠定了坚实的基础。

以某省政务大数据中心为例，该中心正在逐步将各委办局的数据汇聚，汇聚的数据通过加工、共享后实现数据价值的最大化。在数据开发利用的过程中，数据收集团队、数据治理团队和数据开放团队都可以直接接触到海量的核心业务数据，数据安全风险陡增，安全管控难度极大。通过数据安全运营技术的应用，某省政务大数据中心建立了数据安全治理相关的制度、流程、表单、模板，在省级大数据系统建立了数据分类分级、数据脱敏、数据加密、数据访问控制、数据安全运维等安全能力，并实现了各安全组件的策略统一管理和安全信息集中汇聚与分析，交付了以安全能力进行赋能、安全数据辅助决策的可持续数据安全运营体系。

C.1.3 零信任接入安全技术

C.1.3.1 技术概述

零信任系统是企业内部的一个集成的安全平台，它能够根据身份数据、安全基础设施、风险分析数据等信息进行判断，进而来触发企业安全策略的动态执行。零信任将传统的基于边界的低效的安全模型，提升为高效的以资源和身份为中心的企业安全模型。这样做的好处就是，企业可以在一个动态变化的环境中进行合理的访问控制，最终提升安全性、降低风险、简化运营操作、增加业务敏捷性。

在谈到“零信任”前不得不提到“基于网络边界建立信任”的理念，传统网络安全防护大都通过在网络边界部署防火墙、IDS、入侵监测系统等防护设备，对企业网络层层防护，从外向内地将攻击拦截在外从而保障企业网络的安全。然而现在的网络架构和使用模式正在对这种基于边界的防护策略发起挑战：缺乏内部流量的检查、主机部署缺乏物理和逻辑上的灵活性、存在单点故障。比如：位于网络防护边界之外的远程用户和移动设备的接入，云上资产的访问与防护等问题，一旦某个单点被突破，就给了攻击者横行移动、进一步入侵的可能，传统防护策略逐渐显得力不从心。如果把企业网络比作一座城的话，那么传统边界防护就是给城修城墙、设栅栏，而在零信任网络下



就是给一个城市居民配置一名守卫兵。传统基于边界访问的“城墙”不会消失——它还能抵御大部分入侵，仅靠配置“守卫兵”的方式又会占用太多的资源和带宽，两者还需在实用性和安全性上作出平衡。

在2010年，Forrester的一位分析师在一份白皮书《No More Chewy Centers: Introducing The Zero Trust Model Of Information Security》中提出了零信任的概念。随着后续的更新，最终形成了目前知名的 Zero Trust eXtended 框架，简称 ZTX。

几乎同时，Google 开始了在内部构建 BeyondCorp。BeyondCorp 是一个零信任解决方案并且实现了基础的零信任组件，这使得 Google 能够移除自身企业内部的网络边界。Google 在2014年通过一系列的文章阐述了他们的实现理念，这很大的程度上影响了业界。

在2017年，Gartner的一位分析师重新定义了 Continuous Adaptive Risk and Trust Assessment(简称 CARTA) 的概念，使得 CARTA 和零信任有很多相似的原则。CARTA 不仅要关注身份和数据相关的因素，还要关注发生访问时身份和设备所可能带来的风险。

使得零信任被进一步关注的是，美国 NIST 在2020年发布的《NIST Special Publication——Zero Trust Architecture》以及一个关于零信任的项目 US National Cybersecurity Center of Excellence。发展至今，业界已经普遍认为，零信任所倡导的改变是必要的。因为通过零信任，可以防止恶意用户在企业边界内部访问私有资源、防止数据泄露以及恶意操作。

C.1.3.2 功能架构

融合“管理平台”+“安全组件”+“安全服务”的零信任逻辑架构如下图所示：

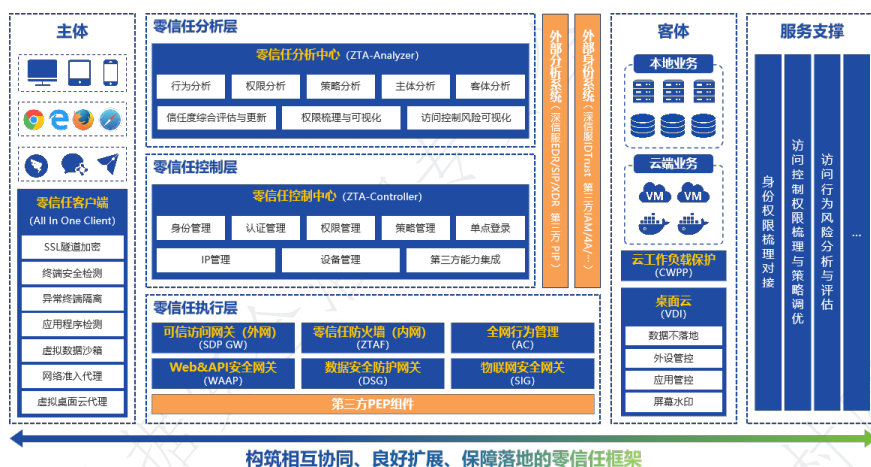


图 50 数据安全零信任架构图

1) 通过统一管理平台，打通端到端身份信息，实现全局风险研判与信任评估

管理平台部分可以进一步细分为“零信任分析中心”和“零信任控制中心”，零信任分析中心的主要职责是通过多维度获取的信息进行信任的评估，除了自身的分析模型与算法组件，还可以通过开放 API 与第三方分析系统进行对接，实现对访问行为和风险的持续评估；零信任控制中心主要职责是身份管理，以及根据分析中心的评估结果，生成动态的访问控制策略，向分散各处的、各种类型的安全组件进行下发，统筹全局的访问控制。



2) 通过丰富的安全组件，匹配用户端到端数据访问控制的场景需求

安全组件部分的核心职责是接收和执行管理平台的访问控制策略，其组成、形态、部署位置、功能选择等，可以随业务管控实际需求而定：1) 零信任一体化安全客户端，是零信任架构的重要组成部分，整合了安全基线核查、病毒与漏洞防护、网络安全接入、数据泄露防护等多种安全能力；2) 匹配不同业务场景的各类安全网关，包括外网访问、内网访问、数据中心服务间访问、物理网访问等等；3) 通过开放 API 接口实现异构厂商安全组件的按需灵活对接。

3) 通过融合专业安全服务，支撑零信任架构落地，保障安全效果，兑现安全承诺

零信任不是“交钥匙工程”，系统和设备的堆砌无法达成既定目标，应该结合安全专家知识与云端分析能力，提供线上或线下多样化的安全服务（如账户与访问权限梳理、业务访问路径梳理、访问控制策略优化、风险分析与处置等），实现安全威胁的快速识别与处置。

4) 通过“平台”+“组件”+“服务”的体系，最终构筑贯穿用户、终端、应用、连接、权限、数据，构筑端到端的信任体系

完善的零信任系统，应该尽可能覆盖端到端业务运行流程，考虑如何在用户、设备、应用、连接、权限和数据等多个环节，落地“什么值得信任”这个关键的问题。核心的功能包括：（1）通过多维度的终端环境检查，确保访问数据的终端设备安全合规；（2）通过跨场景全局的统一身份管理与多因素认证方式，保证访问数据的用户其身份合法；（3）通过精细化权限控制，对数据的访问进行细粒度授权，防止越权、串权、滥权的访问；（4）根据数据重要程度和密级程度的不同，通过云桌面、终端数据沙箱、网络隔离、数据访问行为审计等技术手段，实现差异化数据泄密防护，保障数据安全；（5）通过持续信任评估，发现数据访问过程中行为的异常，进而生成动态的访问控制策略。

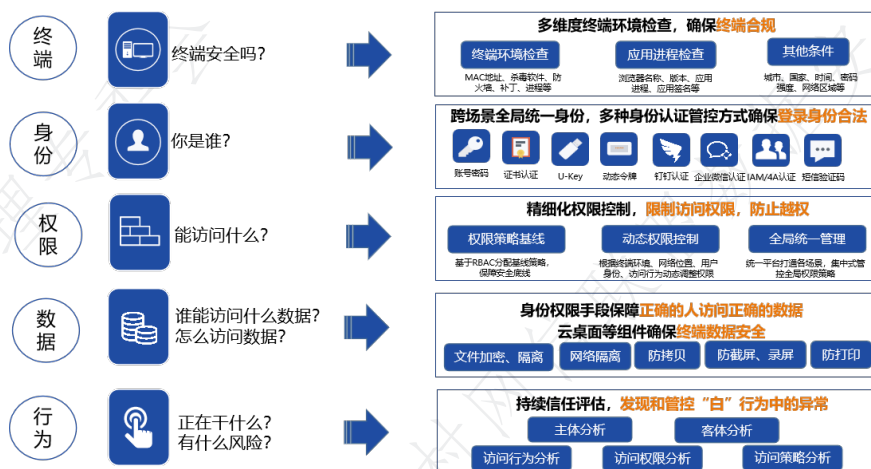


图 51 数据安全零信任防护效果图

根据不同用户的实际情况、需求和目标，实施路径可能有所区别，但大体遵循如下的过程：

1) 梳理数据资产，确定保护目标

需要与客户对待保护的资产目标进行明确，零信任涉及用户管理、认证、授权、收缩暴露面、调整权限、制定保护策略等一系列动作，保护目标不同，这些具体工作也会有所差异，对齐目标非常重要



2) 梳理访问路径，确定数据暴露面

在明确保护目标的基础上，通过业务访问流的梳理，搞清楚哪些业务可能通过互联网被访问，可能被什么人，什么设备，通过什么应用，访问什么域名 IP 端口等等，这些就是业务与外界发生关系的方式，就是业务的暴露面，对暴露面的剖析要尽可能清晰、完善

3) 梳理身份与权限，确定数据安全需求

明确访问保护目标的角色构成，权限授予以及实际运行过程中的访问情况，评估授权是否合理，有没有过紧过松的情况，尽量调优，寻找平衡

4) 梳理方案建设思路，确定落地实施计划

经过前面几个步骤的梳理，基本能够形成安全管控的需求，进而进行合理的方案设计、部件选择和部署实施。

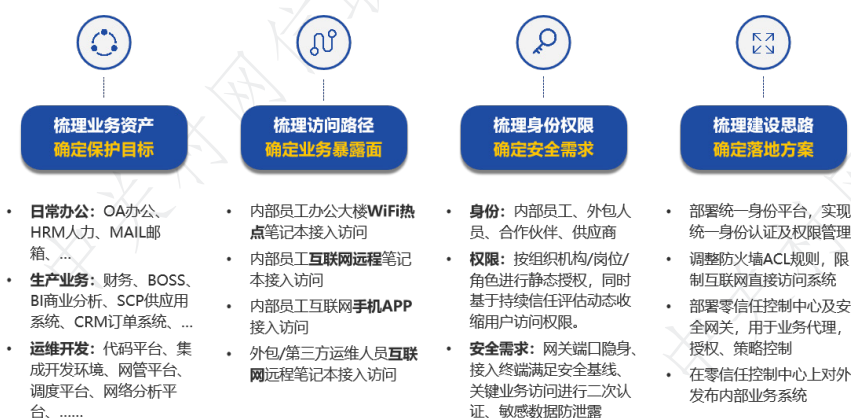


图 52 数据安全零信任实施路径

C.1.3.3 应用场景

面向不同的应用环境、业务场景，零信任架构有多种灵活的实现方式和部署模式。面向远程办公、大数据中心、物联网等典型场景，按照访问主体和资源之间的关系，数据平面的访问代理重点考虑采用和被保护资源相结合的部署模式，搭建安全访问通道，对访问请求进行分流。控制平面的访问控制引擎负责指挥，按照“先认证后连接”原则，建立、维持有效连接，实施对资源的安全访问控制。在此过程中，持续开展安全监控评估，对应用场景中出现的安全威胁及时响应，消减风险。

1) 远程办公

在现代企业的信息化建设环境中，远程办公涵盖的应用场景必然越来越复杂：接入人员和设备的多样性增加；企业资源暴露程度大幅度增加；数据泄露和滥用风险大幅增加；近年来外部攻击的规模、手段、目标等都在演化，利用远程办公找到漏洞，突破企业边界后进行横向移动访问，成为最常见和最有效的攻击手段之一。

常见的远程接入方式有两种，一种是通过端口映射将业务系统直接在公网上开放；另一种是使用 VPN 打通远程网络通道。各组织都在对自己的安全边界进行“加固”，尽量使用 VPN 远程计入而非字节开发业务端口。然而，这些手段基本上可以视作是传统的边界安全方案，难以系统性解决远程办公带来的安全威胁。



零信任架构针对远程办公应用场景，不再采用持续强化边界的思维，不区分内外网，针对核心业务和数据资产，梳理访问这些资产的各种访问路径和场景，在人员、设备和业务之间构建一张虚拟的、基于身份的逻辑边界，针对各种场景构建一体化的零信任动态访问控制体系，构建更安全的远程办公网络。

2) 大数据中心

大数据中心业务上要求数据集中和共享，一方面实现了多部门、多平台、多业务的数据融合；另一方面在数据内部打破了业务之间、部门之间的网络边界，实现互通互访。这将集中更多的风险，从而使其更容易成为攻击的目标。

目前大数据中心访问中东西向流量大幅度增加，而传统的安全产品基本都是在南北向业务模型的基础上进行研发设计的，无法满足大数据中心的安全需求。零信任架构通过微隔离技术，实现环境隔离、域间隔离、端到端隔离，根据环境变化自动调整策略，实现精细化隔离的网络安全策略及以身份为基础的逻辑边界，同时安全策略可以自适应调整。

3) 物联网

5G、大数据、人工智能等新技术为物联网带来了创新活力，催生了智能家居、智慧车联等新兴应用领域。在物联网越来越普及的同时，网络安全问题也甚嚣尘上。物联网中海量多样化的物联网设备，其入网方式和自身脆弱性都带来无法忽视的系统风险。

随着边缘计算技术的不断完善，边缘计算在本地执行技术和分析的思想也越来越被接受，云边协同成为新的基础架构。在物联网实施零信任架构，可借助边缘计算技术，解决终端的身份认证和访问控制价，允许身份可信、经过动态授权的物联设备入网，并动态监测，及时发现并处置假冒、伪造的非法连接。其核心思想是对物联设备建立身份指纹，由主体属性、环境属性和客体属性构成，通过持续主动扫描、被动监听检测、安全接入控制区等方式，对物联终端进行持续信任评估、访问控制，解决物联终端的身份仿冒和恶意访问。

C.1.4 隐私计算技术

C.1.4.1 技术概述

数字经济时代，数据已成为政府、企业、组织实现持续、稳定、健康发展的重要核心资产。国内社会对于用户隐私的法律保护逐步完善，数据安全建设已经步入“有法可依、有法必依”的时代。一方面，数据只有通过流通才能产生价值，不能交融的数据信息只能固封在各领域的资产库中形成数据孤岛，不能展现最终的生产价值。而另一方面，对数据的保护，却产生了一个矛盾：大量的数据因为需要各自依法保护而无法被联合在一起运算。

目前不管是多方安全计算还是联邦机器学习平台，都只能解决单一的隐私计算问题。多个设备分别部署，各自本地数据无法同时共享，只能通过多次操作才能反复利用。如何使数据在综合开发利用，创造数据价值的同时，能够得到充分的保护和隔离，是当前亟待解决的问题。

基于多方安全计算、联邦学习、隐私求交和匿踪查询等隐私计算技术已经成为主流的数据安全共享技术，不仅实现了数据的安全共享，还可以支持去隐私化数据开放、数据不出域计算、密文计





算等多种安全共享路径，保障数据主权的同时，保护软件、数据、信息、物理环境等其他基础环境不受偶然或恶意的原因而遭到破坏、更改、泄露，保障系统连续可靠正常地运行，信息服务不中断，最终实现数据安全流通的同时，助力多方业务融通的连续性。

C.1.4.2 功能架构

数据的安全共享，需要隐私计算技术手段和数据安全协同理念相结合，才能真正构建具备安全、合规及生态能力的数据共享、研究、利用及交易的平台。基于数据安全共享技术的协同生态圈，应该涵盖了从基础资源管理到计算引擎再到数据服务和应用服务的多层架构，具有“高安全、高性能、高兼容”三大优势特性：一是无第三方参与联邦学习（FL）技术，解决多方联合建模中的第三方可信风险问题；二是隐私增强 MPP 引擎，相对于高并发大数据量性能速度优化提升，解决联合建模、多方安全计算中的通信性能问题；三是通过“MPC+FL”的双计算引擎，适配不同计算场景，并在一定标准内支持异构计算框架之间的互联互通。

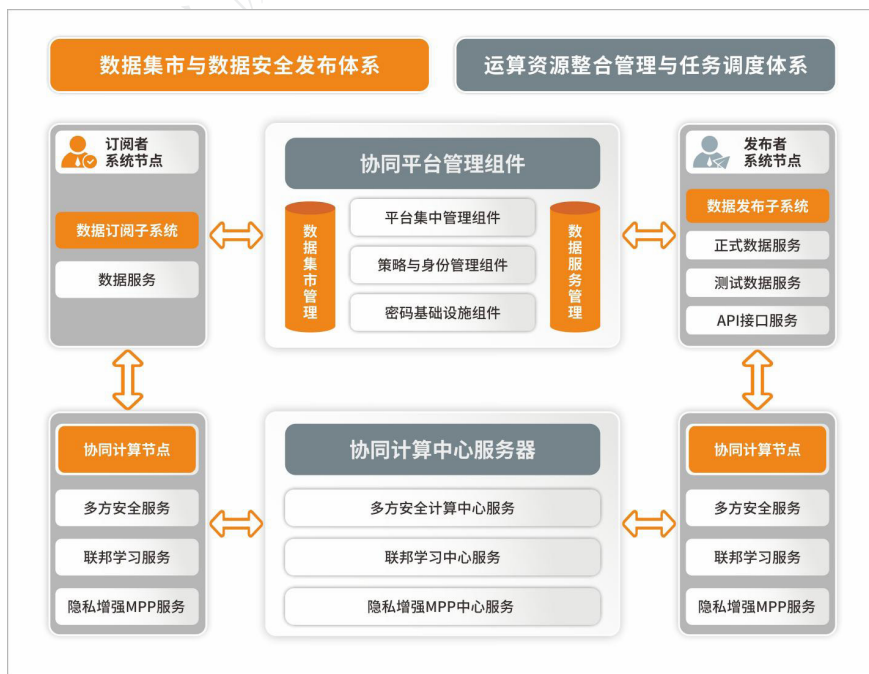


图 53 隐私计算技术架构图

数据安全协同生态圈的主要架构由数据集市和数据安全发布体系，以及运算资源整合管理与任务调度体系组成。

1) 数据集市和数据安全发布体系

数据集市和数据安全发布体系，可以由数据订阅者系统节点和数据发布者系统节点组成，各节点均部署自己的协同计算节点，用于进行多方安全计算、联邦学习计算和隐私增强计算等服务。

其中发布者系统节点为该生态系统的数据库持有方，负责管理本地用户的登录、数据主题的发布、发布数据和发布服务的维护、计算资源的管理、订阅请求的批准、数据协同任务的批准、数据协同任务的参与、协同结果的存储、本地数据处理过程日志的存储等。

订阅者系统节点为该生态系统的数据库需求方，负责管理软件负责管理订阅者成员、计算资源管理、



数据订阅、数据协同任务申请、数据协同任务的执行监控等。

2) 协同计算节点

数据发布子系统和数据订阅子系统的协同计算节点，依托密码学基础算法库和人工智能算法库组成的计算引擎层，形成多方安全服务、联邦学习服务两类隐私计算服务，并通过隐私增强 MPP 服务形成大规模并行处理能力。

● 多方安全计算服务 (MPC)

多方计算的核心思想是使多个非互信参与方之间可以在数据相互保密的前提下进行高效的数据整合和计算，做到既使用多源数据实现给定的计算任务，又保证数据在使用过程中不被外泄。主要用于两方或者多方的协同计算任务，在密文状态下进行基础四则运算，包括 (+ - * / ())；比较操作，包括 (>、<、=、<>、>=、<=、between、not between；函数运算，包括 max、min 等；聚合函数操作，包括聚合函数：sum、count、max、min 等，算子包含 select、join、union 等多种算子及其查询结果。

● 联邦学习服务 (FL)

联邦学习用于解决如何在有效利用用户终端上的数据进行 AI 模型训练的同时不侵犯用户数据隐私的问题。联邦学习不需要收集任何一个参与方的原始数据，而是由每个参与方在本地进行一些计算，得到完成模型训练任务所需的中间结果，然后再收集或者交换这些中间结果（如梯度或模型参数）等。联邦学习服务 (FL) 既可以用于本地进行单方数据机器学习建模，也可以用于联邦两方或多方数据进行协同任务进行联邦学习模型构建。联邦学习建模应用的算法库包括特征分箱、特征缩放、本地逻辑回归、横向逻辑回归、纵向逻辑回归、横向 XGBoost、纵向 XGBoost、纵向线性回归、纵向 K-means 等。

● 隐私计算引擎

隐私计算引擎由密码学基础算法库和人工智能算法库组成，包含秘密分享、姚式混淆电路、不经意传输、同态加密、多重签名、零知识证明等安全传输协议和数据密码分享方式；人工智能算法库包括逻辑回归、线性回归、决策树、XGBoost、神经网络等分析模型及人工智能算法。多种算法的集成适配，组成了联邦学习 FL 引擎、多方安全计算 MPC 引擎、隐私增强 MPP 引擎、去隐私化服务引擎等。

使用 MPC 以及同态加密技术构建底层安全计算协议，将不同的特征在加密状态下加以聚合，并以此支持不同种类的机器学习的安全计算。包括逻辑回归、基于决策树的算法、深度学习和迁移学习等，结合分布式系统的结构有监督的训练和增强机器学习模型。

● 隐私增强 MPP 服务

大规模并行处理的隐私增强引擎，可用来替代目前的计算节点中对于某些需要实时计算，并且计算任务量庞大、计算任务较复杂的应用场景，如多方安全计算无法解决的运算效率问题。它将任务并行的分散到多个服务器和节点上，在每个节点上完成计算，大大提高运算效能，突出增强隐私计算优势。



3) 运算资源整合管理与任务调度体系

运算资源整合管理与任务调度体系为平台整体调度管理中心节点，由隐私计算平台管理组件（平台集中管理组件、策略与身份管理组件、密码基础设施组件）和协同计算中心服务器（多方安全计算中心服务、联邦学习中心服务、隐私增强中心服务）组成，负责管理整个平台的成员身份管理，计算资源信息协调和集中分配，日志追踪记录，各节点审批流程和规则制定和下发等。

成员身份管理分为用户管理和角色管理。其中用户管理含用户分组、用户权限设置的具体划分，可以从平台方对成员进行用户管理和角色划分，审批成员注册，对注册成员进行管理，还可以对用户进行分组权限设置。

计算资源信息协调和集中分配，可以根据当前任务队列及任务量进行网络、内存等计算资源统筹和动态调度，分配最小耗能方案，优化计算资源，提高任务执行效率。

日志追踪记录主要记录数据发布信息、发布订阅信息、数据协同任务信息、以及平台成员间的数据使用和交互过程中产生的审批结果、操作日志、告警情况和流量数据等信息，同时提供日志查询以及报表输出功能。

各节点审批流程和规则制定和下发，平台方可制定任务审批流程和结算流程，制定多方安全计算和联邦学习的数据转换规则，制定数据分类分级规则、数据字典、数据行业标签等，并将以上可以由平台方制定的规则和流程进行对各子节点的下发。各子节点也可以单独制定节点规则，由平台方进行审批。

C.1.4.3 关键技术

1) 多方安全计算技术

自 MPC 理论创立以来，相关技术分支得到了长足发展。MPC 是一个多方协议，其采用的基本技术包括混淆电路（Garbled Circuit, 简称：GC），同态加密（Homomorphic Encryption, 简称：HE）、秘密共享（Secret Sharing, 简称：SS）等。

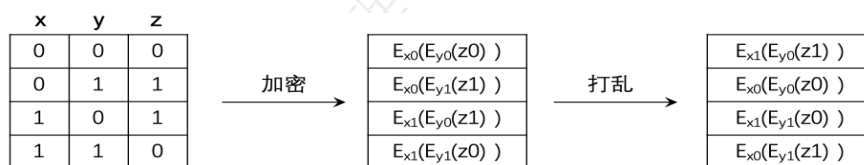
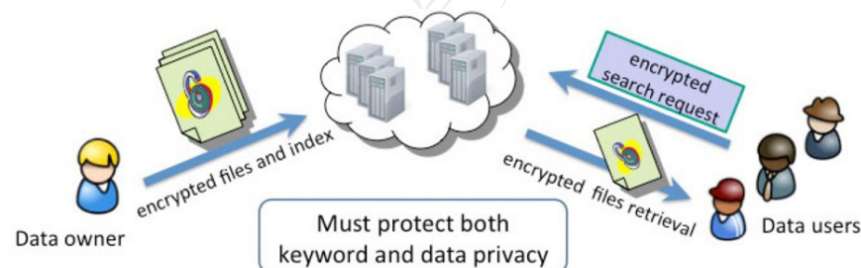


Figure A.1 加密并扰乱 XOR 门的过程。

图 54 混淆电路示意图

混淆电路的基本思想是将计算电路的每个门都加密并打乱，保证计算过程中不会泄露原始输入和中间结果。混淆电路的特点是具有通用性，但计算过程中通信开销较大，比较适合运算复杂度不高的场景。混淆电路研究者也已提出了很多优化方法，如 Free-xor、half-gates 可以节省加密计算量和密文的传输量。



$$\text{Enc}(x_1) \boxplus \text{Enc}(x_2) = \text{Enc}(x_1 + x_2)$$

$$\text{Enc}(x_1) \boxtimes \text{Enc}(x_2) = \text{Enc}(x_1 \times x_2)$$

图 55 同态加密示意图

同态加密的思想比较直观，即数据持有方将自己的输入加密后发给某计算服务器，服务器直接在密文上进行计算，计算后将得到的结果的密文发送给指定结果方，结果方再将结果的密文解密，即可得到最终的计算结果。如此一来，计算服务器一直在密文上操作，无法看到任何有效信息，而参与者也只拿到最后的结果，看不到中间结果。同态加密原理如图 55 所示。当前同态加密技术主要基于 LWE (the learning with errors)、RLWE (the ring learning with errors) 等问题，通常首先实现低阶运算上的全同态运算，然后拓展到高阶运算。上述过程计算性能低，导致目前的全同态加密库如 HElib、SEAL、libScarab 等在运行效率上离实用性尚有一段距离。同态加密可适用于对性能要求不高或者小数据量的情况。

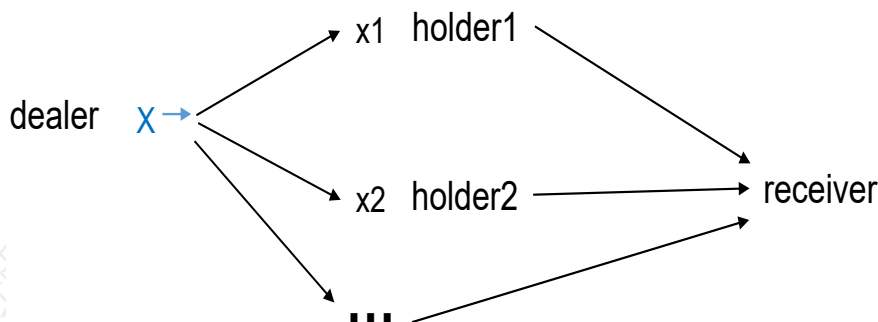


图 56 秘密分享基本原理

秘密分享的基本思想是将隐私数据拆分成若干伪随机的部分（称为分片），每个计算节点持有其中一份，协作完成计算任务（比如加法、乘法运算），因为看不到原消息的全量信息，从而实现数据隐私保护。基本原理如图 56 所示，一个被称为数据持有方的 dealer 将数据拆分为若干份后分别分发给若干个 holder，holder 可以基于分片进行计算，但不能获得数据原文 x ，最后 holder 将数据汇总给 receiver 进行数据还原。基于秘密共享理论基础，研究人员构造了 Shamir、SPDZ 等多方安全计算技术。秘密分享对于加法运算比较直观，但对于乘法运算需要提前预置生成大量的数据，并进行多次交互实现。PrivPy 使用了一种新的计算协议，能够对秘密信息进行二次分片，避免了数据提前预置。

2) 联邦学习技术

联邦学习技术作为一种多方联合建模的计算框架最早是由谷歌在 2017 年公开发布。在三年多的



时间里，国内外诸多科技巨头都已经开展联邦学习的深入研究与应用。联邦学习最初提出是如何有效利用用户终端上的数据进行AI模型训练，而又不侵犯用户的数据隐私。在从C端应用向B端产业应用演进中，联邦学习形成了更为通用的解决方案：

- 横向联邦学习

即样本中用户不同，特征相同，谷歌的方法即是这种。

- 纵向联邦学习

即样本中用户相同，特征不同，适用于拥有同一批用户的不同企业。

- 迁移联邦学习

即样本中用户、特征都可能不同，这样可以通过数据的升维或降维，从而进行另外数据子空间的迁移学习。

目前国内多家企业也都纷纷推出了可以进行产业落地应用的开源框架。微众银行基于联邦学习理论研究进行相关开源软件研发，正式开源全球首个工业级联邦学习框架FATE（Federated Learning Enabler）；百度在PaddlePaddle 2.0开放平台中，也增加了PaddleFL联合学习框架；平安科技在其提出的联邦智能体系的基础上自主研发了蜂巢联邦学习平台；华控清交在其多方安全计算平台上设计开发了基于明密文混合运算的联邦学习系统。

联邦学习的场景中，非互信的多个实体希望协同解决一个机器学习问题，如模型训练等。在这个过程中需要用到这些实体间的融合数据，但是每个参与方都不希望暴露自己的原始数据。因此联邦学习的过程中，不收集任何一个参与方的原始数据，而是由每个参与方在本地进行一些计算，得到完成模型训练等任务所需的中间结果，然后收集或者交换这些中间结果，如梯度或模型参数等。由于这些中间结果也包含原始数据的很多信息，因此收集或者交换中间结果的过程可以在密文上进行。通过明密文混合运算的联邦学习系统将能有效解决机器模型训练问题，发挥数据融合价值。

3) 同态加密技术

同态加密技术（Homomorphic Encryption，简称HE）是基于数学难题的计算复杂性理论的密码学技术。对经过同态加密的数据进行处理得到一个输出，将这一输出进行解密，其结果与用同一方法处理未加密的原始数据得到的输出结果是一样的。

一般的加密方案关注的都是数据存储安全：即要给其他人发个加密的东西，或者要在计算机或者其他服务器上存一个东西，要对数据进行加密后在发送或者存储。没有密钥的用户，不可能从加密结果中得到有关原始数据的任何信息。只有拥有密钥的用户才能够正确解密，得到原始的内容。这个过程中用户是不能对加密结果做任何操作的，只能进行存储、传输。对加密结果做任何操作，都将会导致错误的解密，甚至解密失败。

而同态加密方案最有趣的地方在于，其关注的是数据处理安全。同态加密提供了一种对加密数据进行处理的功能。也就是说，其他人可以对加密数据进行处理，但是处理过程不会泄露任何原始内容。同时，拥有密钥的用户对处理过的数据进行解密后，得到的正好是处理后的结果。

4) 隐私求交技术

隐私求交技术（Private Set Intersection，简称PSI）是多方安全计算领域的一个子问题，目的是在保护通信双方的数据隐私性前提下完成数据集的交集计算。隐私求交协议允许持有各自集合的

两方来共同计算两个集合的交集运算。在协议交互的最后，一方或是两方应该得到正确的交集，而且不会得到交集以外另一方集合中的任何信息。

保护集合的隐私性是在很多场景下是自然甚至是必要的需求，比如当集合是某用户的通讯录或是某基因诊断服务用户的基因组，这样的输入就一定要通过密码学的手段进行保护。

协议允许持有各自数据集合的两方执行双方集合的交集运算。PSI 协议结束之后，一方或两方能够得到交集结果，但是双方都无法获知交集以外的对方集合数据的任何信息。

5) 不经意传输技术

不经意传输 (Oblivious Transfer, 简称 OT) 是一个密码学协议，在这个协议中，消息发送者从一些待发送的消息中发送一条给接收者，但事后对发送了哪一条消息仍然 oblivious (不知道)，这个协议也叫茫然传输协议。

不经意传输是设计其他密码协议的基础，也可以构造更为复杂的协议，如不经意电路计算等。

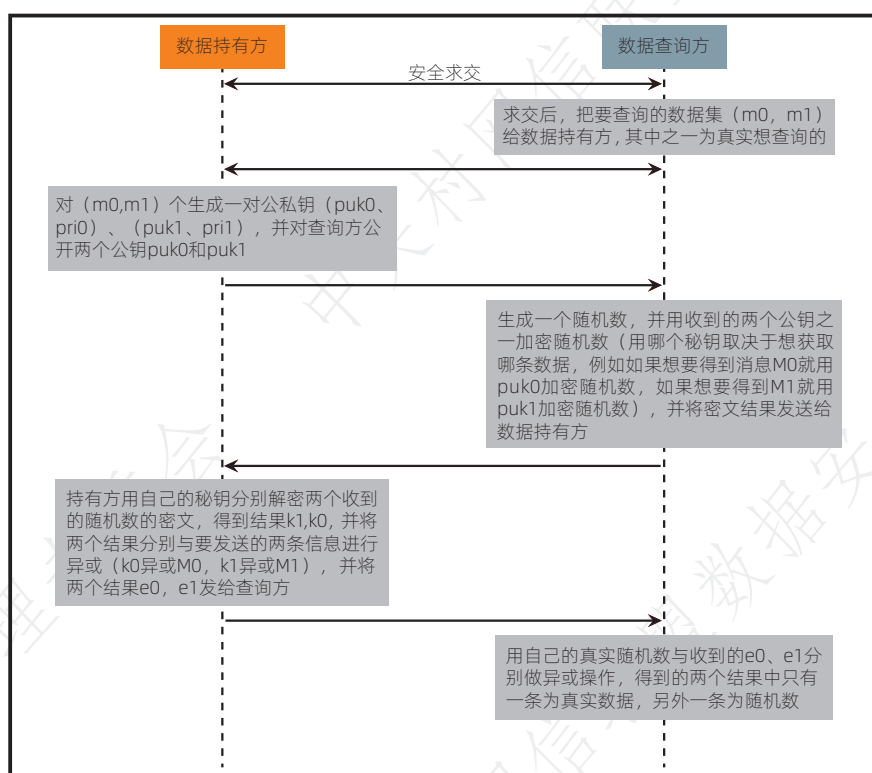


图 57 不经意传输协议示例过程

第一步、数据查询方把要查询的数据集（包含真实要查询的数据）发送给持有方

第二步、数据持有方生成两对 RSA 公私钥，并将两个公钥 puk_0 、 puk_1 发送给数据查询方。

第三步、查询方生成一个随机数，并用收到的两个公钥之一加密随机数（用哪个密钥取决于想获取哪条数据，例如如果想要得到消息 M_0 就用 puk_0 加密随机数，如果想要得到 M_1 就用 puk_1 加密随机数），并将密文结果发送给持有方。

第四步、持有方用自己的两个私钥分别解密收到随机数密文，并得到两个解密结果 k_0, k_1 ，

并将两个结果分别与要发送的两条信息进行异或（ k_0 异或 M_0 ， k_1 异或 M_1 ），并将两个结果 e_0, e_1 发给查询方。



第五步、查询方用自己的真实随机数与收到的 e0、e1 分别做异或操作，得到的两个结果中只有一条为真实数据，另外一条为随机数。

6) 接口服务能力

可手写表达式自定义函数接口信息，针对接口安全采用公私钥 + 签名 + 证书管理，结合匿踪查询技术，参与方只需传递加密的参数，隐匿查询对象关键信息。可通过 OpenAPI 方式直接自定义算法对接第三方系统。且平台在隐私计算技术中引入分级分类概念，对上传数据自动标识等级、分类，对开放数据进行标准化、精细化管理，保证用户使用的数据统一标准、分级分类、安全可控，提升平台数据质量。

C.1.4.4 应用场景

基于数据安全共享技术的隐私计算平台无论在数据价值分享的过程中，还是数据协同过程中，都应为各组织机构数据流通提供安全体验，推进政务、金融、医疗、互联网等机构在联合风控、联合营销、联合分析等场景的数据安全和隐私保护需求，实现数据价值极大释放。

以政务数据共享为例，通过隐私计算平台打通政企之间业务屏障，实现数据的安全共享。

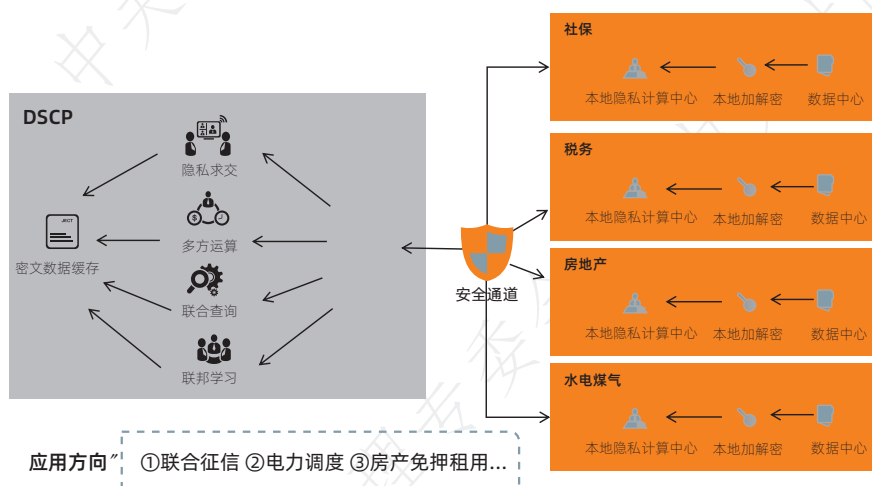


图 58 隐私计算应用场景图

整体数据安全共享技术中心服务可部署在政务中心，实现以下管控效果：

- 国有企事业单位数据可以选择由政务中心数据部门统一管理审批
- 明文数据不出所在单位，全程加密传输、缓存和运算，保护数据的安全和隐私
- 各行各业综合使用政务服务，提升社会治理水平
- 其他机构，如金融机构使用数据，可以注册节点统一使用数据，简化线下申请和审批环节

利用数据安全共享技术，既可以为政务中心构建全闭环安全体系，也可以形成覆盖事前、事中、事后的闭环安全隐私数据管控的生态系统。

事前防护可根据开放数据的敏感等级灵活设置安全共享策略：平台率先引入分级分类概念，对上传数据自动标识等级、分类，对开放数据进行标准化、精细化管理，保证用户使用的数据统一标准、分类分级、安全可控，提升平台的使用数据质量。

事中监控可实时监测数据开发利用过程中的状态和风险：根据用户实际生产环境需求，多渠道多路径消息通知，对风险进行预警告知和处理，使用户可以根据告警内容准确、快速地做出响应，定位异常行为提升问题解决效率。

事后审计对整个数据流通过程进行全程审计和取证溯源：对用户行为和数据操作行为进行记录、分析和汇总，用来帮助事后生成数据报告、事故追根溯源，结合区块链技术算法上链、内容存证、日志存证等，全面保障和提高数据资产安全。

C.1.5 差分隐私

差分隐私通过严格的数学保证数据集中任意个体的变化对最终发布的统计结果几乎没有影响。它要求对只有一个记录不同的两个数据集，进行相同的查询，随机算法输出相同结果的概率是相近的。这种隐私定义要求，任何个体是否选择加入或退出数据集对于查询结果都不会有太大的影响。这种定义隐含的思想是向用户解释，用户信息的泄露与他们的数据是否在数据集中基本是无关的，从而推动用户更广泛的分享数据。

差分隐私定义的稳定性在于，它提供的隐私保护程度与攻击者的背景知识是无关的。这是因为，差分隐私本质上是对随机算法的随机程度的度量，因此只要证明了随机算法的随机性，就能保证在对应场景下，使用该随机算法发布数据能提供一定程度上的隐私保护。

C.1.5.1 数据收集模型

差分隐私可以通过是否存在可信的数据管理者，分为中心化差分隐私模型与本地化差分隐私模型两种类型。

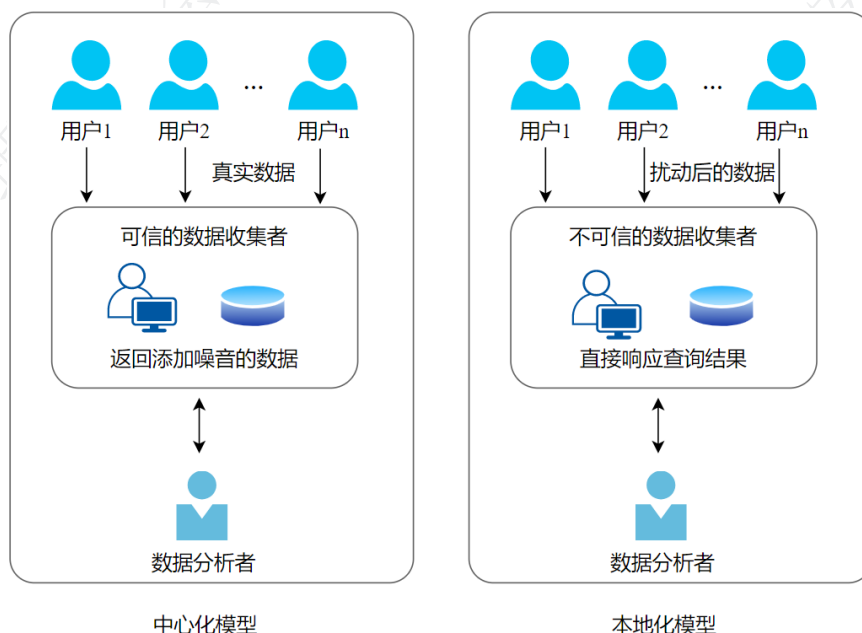


图 59 数据收集模型

1) 中心化差分隐私模型

中心化差分隐私模型如上图左所示。中心化模型假设存在一个可信的数据管理员，它可以收集



用户的真实信息。数据管理员在掌握了真实数据的基础上，统一对发布信息添加噪音，或者根据真实数据的具体分布情况，设计更加具有针对性的随机算法。由于数据管理员可以在真实的统计数据上统一添加噪音，因此中心化差分隐私模型往往具有较高的精确性。

中心化差分隐私模型通常用于数据库查询，广告投放等场景。Google 已公布了提供差分隐私保护的数据库开源代码，Uber 也同样发布了该公司对于差分隐私数据库的研究。

2) 本地化差分隐私模型

本地化差分隐私模型如上图右所示。本地化模型考虑不存在可信的数据管理员的场景。在这个场景下，用户不信任数据收集者，因此用户在上传数据时，通过一个本地的随机发生器，以一定概率上传真实结果，而以一定概率上传一个随机值。由于数据收集者收集到的数据集已经得到了隐私保证，因此可以直接对该数据集后续分析，而不需要额外的隐私保护。但由于每个用户都会以一定的概率上传随机值，因此本地化差分隐私模型的精确性往往会低于中心化差分隐私模型。

由于本地化差分隐私模型不需要一个可信的数据收集者，因此在数据收集阶段，本地化模型通常更受到用户的青睐。Google 已经在浏览器中通过本地化差分隐私机制 RAPPOR 收集用户的浏览器设置，Apple 也在其 iOS, Mac OS 系统中使用本地化差分隐私机制收集用户表情、健康等多种信息。Microsoft 同样发布了其关于本地差分隐私的研究，利用本地化机制收集 Windows 中的系统信息。

C.1.5.2 隐私保护数据发布模型

基于差分隐私保护的数据发布是差分隐私研究中的核心内容，其目的是在不披露任何个人记录的情况下向公众输出汇总信息。根据对于多次查询的响应方法不同，分为交互式和非交互式。

1) 交互式数据发布

在交互式设置中，给定数据集和查询集，需通过一种数据发布机制，使其能够在满足差分隐私保护的条件下，逐个回答查询集中的查询，直到耗尽全部隐私预算。即在发布前一个查询的答案之前不能发出新的查询。发布机制的性能通常由精确度来衡量。交互式数据发布即是要在满足一定精确度的条件下，以给定的隐私保护预算回答尽可能多的查询。交互式设置的方法主要考虑事务数据库、直方图、流数据和图数据发布等。

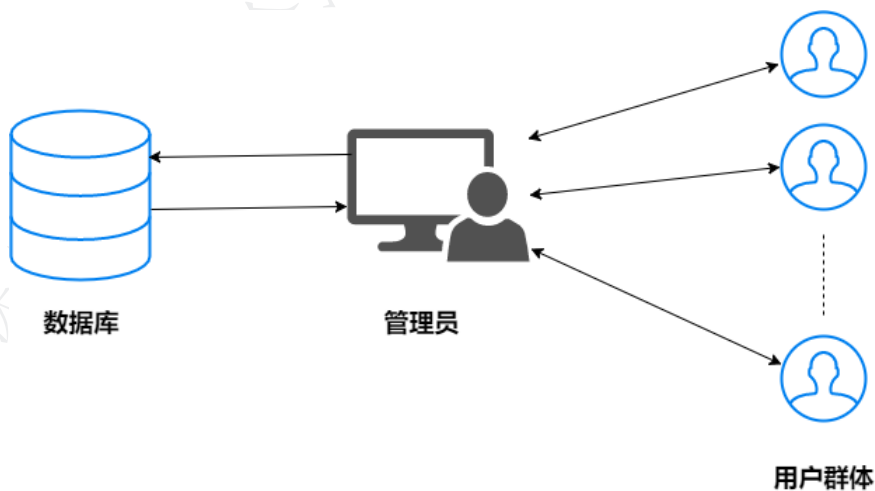


图 60 交互式数据发布模型



2) 非交互式数据发布

在非交互设置中，给定数据集和查询集，需通过一种数据发布机制，使其能够在满足差分隐私保护的条件下，一次性回答中的所有查询。数据管理者针对所有可能的查询，在满足差分隐私的条件下一次性发布所有查询的结果，或者发布一个原始数据集的净化版本，即一个不精确的数据集，用户可对该版本的数据集自行进行所需的查询操作。非交互式数据发布方法主要集中在批查询、列联表发布、基于分组的发布方法以及净化数据集发布。

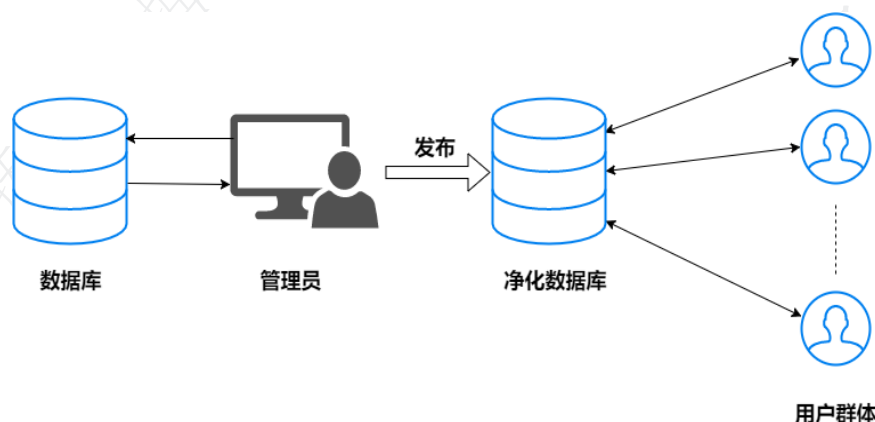


图 61 非交互式数据发布模型

C.1.6 数据水印和数据使用行为溯源

C.1.6.1 技术概述

在开展数据安全防护工作过程中，越来越多的数据通过网络进行存储和发布，这些数据往往包含各类敏感信息，即使经过脱敏或去标识化处理后，仍然有巨大的社会价值与经济价值，成为宝贵的资源，如果不采取有效的安全控制和版权保护措施，常常会给攻击者以可乘之机。

同时，如果缺乏数据完整性验证的有效措施，一旦出现数据泄露，后果也是无法想象的。因此，需要使用数据水印技术在数据交换及数据使用中的分发共享、委托处理等环节标明数据的所有者、数据分发对象、分发时间、分发途径及使用范围等信息。

当出现数据泄露事件时，可以从植入的数据水印信息中还原上述信息，从而达到追溯泄露途径，追责泄露人员的目的，并且可以促进安全建设，避免相同类型的数据泄露事件发生。

数据水印是指从原始环境向目标环境进行敏感数据交换时，通过一定的方法向数据中植入水印标记，从而使数据具有可识别分发者、分发对象、分发时间、分发目的等因素，同时保留目标环境业务所需的数据特性或内容的数据处理过程。数据水印应具有隐蔽性、可追溯性、确定性等特点。

自从 2015 年起，数据水印技术基于伪行，伪列与仿真水印三种原理进行发展，并逐渐产生小规模应用。

C.1.6.2 功能架构

1) 数据初始化

数据库表包含有多个元组，每个元组的数据模式可以用 $R = (Pk, F1, F2 \dots Fn, Fk)$ ，其中 R 代表元组，



Pk 代表主键，Fk 为外键，F1, F2...Fn 为属性；在属性中会包含有部分敏感信息，定义 C 为数据库中包含敏感属性的表的集合，对 C 进行数据抽取形成数据子集 S，S 中含有一个或多个表，S 中表的属性含有敏感信息，S 中表之间具有一定的关系，比如主外键表；

● 定义伪行伪列字段规则

字段规则根据敏感数据属性含义和数据类型进行定义；敏感数据属性定义会根据数据属性含义和数据类型结合，敏感数据属性往往具有明显的特征和数据类型；首先将分离出敏感数据属性，根据敏感数据属性的类型、值域范围及限定规则，抽象出敏感数据属性的规则，数据值型数据将通过正则表达式方式抽取数据属性规则，字符型属性可通过正则表达式方式或进行值域范围分解抽取数据属性规则；

● 对数据进行字段规则配置

选择数据子集 S 中一个或多个表的属性，进行伪行字段或者伪列字段规则配置；首先选择指定的表，根据表中的已有字段属性分布情况，选择采用伪行或伪列方式嵌入水印信息；选择伪行时，对表中已有字段属性进行规则定义，选择字段属性配置对应的伪行字段规则，由于表中字段属性会很多，限制至少两个字段属性需要配置，其余字段属性按默认值，配置生成的数据行比例因子；选择伪列时，直接配置对应的伪列字段规则，生成的数据行比例因子为 1；

● 水印数据的生成

根据选择的字段规则，进行水印数据的生成；伪行或伪列水印处理根据字段规则配置生成数据行或数据列，生成的数据中对应规则配置的数值型数据按正则表达式构建数据值，并通过随机数值组合生成水印标记信息并嵌入到数据值中；字符型数据按值域范围获取数据字典数据构建数据值，并通过值域数据字典对应的数值编码生成水印标记信息并嵌入到数据值中；

● 水印数据的嵌入

水印数据的插入处理；伪行水印处理时，按照配置的数据行比例因子，在原始数据中按照一定的间隔比例将水印数据分散的插入到原始数据中；伪列水印处理时，对原始数据增加新的字段属性并根据配置生成字段名称，然后将伪列数据插入到对应的字段属性中。

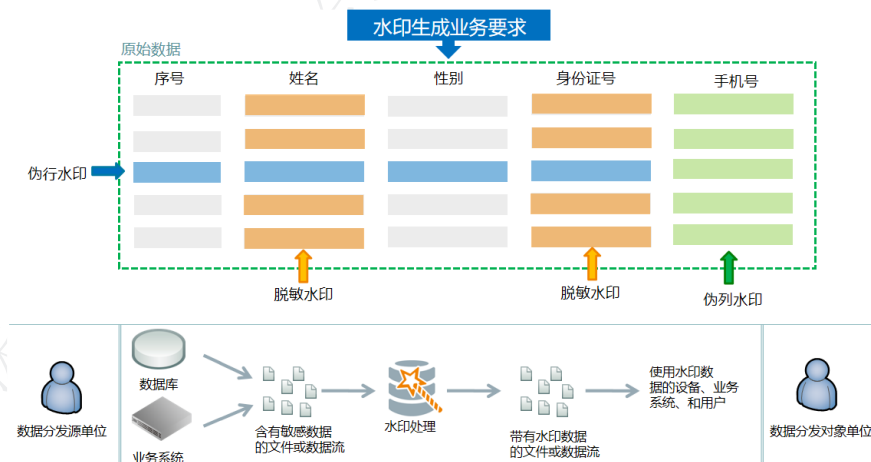


图 62 数据库水印处理示意图

2) 水印溯源

水印溯源技术是指通过泄露数据中包含的水印数据识别其数据特征，并最终解析出水印信息的技术。一般包括识别水印类型、解析水印信息、回溯分发内容等三个步骤。

● 识别水印类型

指通过对泄露数据分析，判断其中哪些字段为植入的水印信息的字段。

● 解析水印信息

通过对植入水印字段的解析，从中获取到水印代码内容。

● 回溯分发内容

指通过解析出的水印代码，查询映射码表，并获取分发者、被分发者、分发时间、用途的关键信息的步骤。



图 63 数据库水印溯源

C.1.6.3 关键技术

1) 伪行水印

伪行水印指在对某些外发数据添加水印时，通过添加人为生成的若干整行信息，并从中挑选某些字段植入水印信息的技术。这些筛选的植入水印的字段一般常见于身份证、电话号、银行卡号、交易金额等字段。

伪行水印主要处理步骤为：基于对数据表结构分析和数据类型分析，识别主键表和外键表；确定主键值的处理方式，自动生成或者手动生成；识别敏感数据字段，在伪行水印库中选取添加的数据类型与敏感字段匹配；对其他数据字段进行同类型数据的生成；选择生成数据的分组策略，分组策略主要解决伪行数据生成的行数及伪行数据在原始数据中如何分布；根据设置生成伪行数据，并在伪行数据中自动嵌入水印标记。

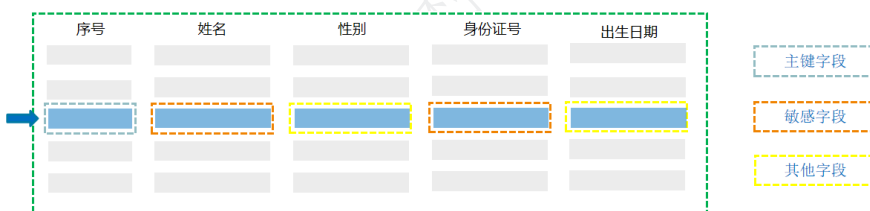


图 64 仿行水印

2) 伪列水印

伪列水印指通过对分发数据中，人为构造增加一列，并在其中植入水印的机制。



伪列水印主要处理步骤为：基于对数据表结构分析和数据类型分析，在伪列水印库中选取添加的数据类型，然后根据水印的类型，系统生成与原始数据量一致的伪列数据，并在伪列的数据中自动插入水印标记。



图 65 伪列水印

3) 脱敏仿真水印

脱敏仿真水印指在增加行和列的情况下，通过对原始数据中某些字段按照一定的水印植入规则进行脱敏变形，产生的脱敏后数据植入了水印信息的技术。

脱敏水印主要处理步骤为：基于对数据表结构分析和敏感数据发现分析，识别敏感字段。根据识别的敏感字段，选择相应的数据变形算法，对数据进行变形处理，在数据过程中添加水印标识到变形后的数据中。在数据置换的数据变形方式下，对数据置换的可逆编码记录并结合水印标记信息生成新的变形后数据，生成带有水印标记信息的数据。



图 66 脱敏水印

C.1.6.4 应用场景

1) 静态水印应用场景

静态水印指类似通过静态脱敏的机制，在一次性的完成大批量数据的水印添加处理。静态水印通常会在将生产环境中的数据交付至开发、测试或者外发环境时使用，在外发数据的同时，植入水印信息。

静态水印技术可以和静态脱敏技术结合使用。一方面降低外发数据的敏感性并且保证数据可用性，另一方面植入水印信息，确保数据一旦泄露，可以进行追溯。

● 开发测试应用场景

业务系统中存在大量的客户敏感信息，如姓名、年龄、手机号、银行卡号、地址、工作信息等，在系统建设前期，往往需要使用上述信息进行开发测试，即使使用了数据脱敏技术去除了敏感信息，





仍然需要使用水印技术向外发的数据中植入水印，标明分发者、分发目标、数据用途等信息。

在此场景中适合使用静态水印中的伪行水印技术，既能保证添加水印，又不会对数据结构有所影响，从而干扰开发测试正常进行。

● 数据分享应用场景

数据分享应用场景在一些特定需求下，部分隐私数据需要提供给其他机构或企业，需要使用水印技术向外发的数据中植入水印，标明分发者、分发目标、数据用途等信息。

在此场景中适合使用静态水印中的伪列水印技术。

● 数据科学研究应用场景

数据科学研究应用场景，其主要目的是通过数据进行研究，因此需要保留数据本身的一些特征。此场景中适合使用静态水印中的伪列水印技术，增加的部分列信息不会包含研究应用所用到的真实数据列，不会干扰研究结果。

2) 动态水印应用场景

动态水印指在数据的访问过程中，动态的在数据的查询访问请求返回的结果集中植入水印的方式。动态水印通常会在数据访问、数据实时交换、数据动态分发过程中使用，在实时交付访问者所需数据的同时，植入水印信息。

● 生产应用场景

生产场景主要指各类业务场景，当涉及访问敏感数据时，需要对获取到的数据植入水印信息，这种场景适合动态水印机制。

● 数据交换应用场景

数据交换场景主要是通过 API 接口方式向特定平台提供数据，这类场景适用动态水印机制。

C.1.7 响应与处置自动化 (SOAR)

C.1.7.1 技术概述

响应与处置自动化 (SOAR) 技术可以实现安全事件响应与处置动作的自动化，避免安全运营团队在面临复杂、繁忙、多变的工作时出现工作事项遗漏的情况，SOAR 技术能够帮助安全运营团队更智慧地开展安全运营工作。

C.1.7.2 功能架构

SOAR 的核心能力包括编排、自动化、安全响应和威胁情报，SOAR 具备对第三方设备 / 系统接口对接能力、剧本编排能力以及任务管理能力，通过预定义的剧本形成标准化流程，对不同类型的安全事件实现自动化响应，将企业的安全运营流程数字化管理，有效提高安全事件响应效率。通过 SOAR，企业可以大幅度缩短响应时间、充分发挥安全分析师能力价值、标准化安全运营流程、减少安全运营成本。

1) 安全编排与自动化

安全编排与安全自动化是两个不同的概念。其中，安全编排 (Orchestration) 是指将客户不同的



系统或者一个系统内部不同组件的安全能力通过可编程接口 (API) 和人工检查点, 按照一定的逻辑关系组合到一起, 用以完成某个特定安全操作的过程。譬如用户针对一封收到的可疑邮件进行深入检测与响应 (操作) 的过程可以分解为根据拆解出来的发件人、URL 链接和 IP 等信息查询威胁情报系统, 将附件送入沙箱系统进行分析, 并根据情报系统和沙箱系统返回的信息进一步决定是否要通知邮件系统删除该邮件或者附件, 是否要通过 EDR 获取收件人终端上的进一步信息做分析, 等等。上述这个可疑邮件分析的过程就是一个将邮件系统、威胁情报系统、沙箱系统、EDR 等等系统通过一定的逻辑编排到一起的实例。

安全自动化 (Automation) 在这里特指自动化的编排过程, 也就是一种特殊的编排。如果编排的过程完全都是依赖各个相关系统的 API 实现的, 那么它就是可以自动化执行的。与自动化编排对应的, 还有人工编排和部分自动化 (混合) 编排。不论是自动化的编排, 还是人工的编排, 都可以通过剧本 (playbook) 来进行表述。而支撑剧本执行的引擎通常是工作流引擎。为了方便管理人员维护剧本, SOAR 通常还提供一套可视化的剧本编辑器。

剧本是面向编排管理员的, 让其聚焦于编排安全操作的逻辑本身, 而隐藏了具体连接各个系统的编程接口及其指令实现。SOAR 通常通过应用 (App) 和动作 (Action) 机制来实现可编排指令与实际系统的对接。应用和动作的实现是面向编排指令开发者的。

2) 安全事件响应平台

安全事件 (Incident) 响应平台在 SOAR 出现之前就一直存在, 顾名思义就是一个针对 Incident 进行响应和处置的平台。但 SOAR 出现后, 安全事件响应与安全编排与自动化的结合使得响应的能力获得了极大地提升。通常, 安全事件响应包括告警管理、工单管理、案件 (Case) 管理等功能。

告警管理的核心不仅是对告警安全事件的收集、展示和响应, 更强调告警分诊和告警调查。只有通过告警分诊和告警调查才能提升告警的质量, 减少告警的数量。

工单管理适用于中大型的安全运维团队协同化、流程化地进行告警处置与响应, 并且确保响应过程可记录、可度量、可考核。

案件管理是现代安全事件响应管理的核心能力。案件管理帮助用户对一组相关的告警进行流程化、持续化的调查分析与响应处置, 并不断积累该案件相关的痕迹物证 (IOC) 和攻击者的战技过程指标信息 (TTP)。多个案件并行执行, 从而持续化地对一系列安全事件进行追踪处置。

3) 威胁情报平台

威胁情报平台 (TIP) 是 Gartner 在 2014 年定义的一个细分市场, 通多对多源威胁情报的收集、关联、分类、共享和集成, 以及与其他系统的整合, 协助用户实现攻击的阻断、检测和响应。威胁情报主要是以服务而非平台的形式存在。目前 TIP 市场规模不大, 厂商不多, 有的是独立存在, 有的依附于威胁情报服务, 还有的跟安全响应结合, 融合到 SOAR 里面。

通过上面的分析, 我们可以发现, SOAR 作为安全运维的综合响应平台, 具有极强的支撑作用。通过平均检测时间 (MTTD)、平均响应时间 (MTTR)、确认时间 (TTQ) 和调查时间 (TTI) 等明确且可跟踪的指标, SOAR 功能还能帮助分析师了解流程有效性, 快速识别可改善的地方, 进一步提升安全运营团队的效率。这些性能指标还可使安全主管量化团队的整体业务价值, 若有需要, 可作为相关监管机构的重要证据呈上。





另外，SOAR 有助于减少书面工作，改善报告环节。很多安全运营团队有大量基于管理的工作要做，比如撰写报告或记录安全过程等。

尽管 SOAR 的自动化能快速产生回报，值得指出的是，该项功能确实需要前期投资，因而与 IT 部门合作和买入就成了关键。SOAR 自动化整个 IT 部门的响应，所以必须将安全运营团队外部的 IT 团队纳入共事。

C.1.7.3 应用场景

通过在 SOC 中实现 SOAR，不仅可以完善 SOC 的安全响应的能力，尤其是编排和自动化能力，以及响应管理能力，并且能在整体上提升 SOC 的效能，包括安全事件调查分析（含 MTTD）的速度、安全响应（MTTR）的速度、将分散的安全系统整合的能力，以及单个安全运维人员的生产率。最终，SOAR 帮助业务与安全运营团队优化其快速检测与响应威胁的能力，量化 MTTD 和 MTTR 等关键性能指标，通过改善情报与报告、弥合自动化响应动作的流程与操作来减轻日常工作负荷。毫无疑问，自动化正快速成为 IT 人员手中最重要的工具。今天的威胁态势日趋纷繁复杂，安全运营团队越来越难以跟上威胁进化发展的脚步。SOAR 去除了大量手动事务性操作，令安全运营团队能够专注其他更重要的任务。

C.1.8 数据安全监测技术

C.1.8.1 技术概述

在数字经济时代，经济发展依靠的关键生产要素是数据，行业的数据安全关系是企业生存命脉，任何恶意的数据泄密行为都可能使企业遭受严重的损失，特别是如金融、电信、能源、政府、交通等重点行业的数据安全事件的发生，会给国家安全带来严峻挑战。因此，行业的数据安全直接关系到国家安全，数据安全将成为行业安全监管的核心之一，落实企业信息管理主体责任，全面提高网络治理能力，是数据安全监管的核心目标。

当前数据安全监管领域存在目标不统一、水平不一致、资源未整合等问题，而如何摸清各单位的安全，如何响应国家对关键信息基础设施安全监管的相关要求，深化网络安全监管机制创新，统筹网络空间安全监管防护资源，有效保障关键信息基础设施正常运行和不受侵害，是当前亟待解决的现实问题。

数据安全监管体系可采用“纵向贯通，横向融通”的建设思想，分地域、分行，纵向贯通各级单位业务数据安全，横向融通各同级单位安全信息，实现数据安全综合监管一体化，加强日常监管，从监管角度判断企业数据安全合规情况，从行业角度判断数据安全风险趋势，从地区角度分析区域内数据安全总体态势，从集团公司角度俯视各不同分子公司数据安全态势，从全局考虑和解决问题。

通过建立“纵向贯通、横向融通”的数据安全监测体系，组织指导有关单位开展数据安全信息汇总、分析研判工作，可实时掌握被监管单位当前存在的数据安全问题和安全管理水平，宏观把握数据安全态势，整体提升数据安全防护水平，保障政治、经济、文化和社会安全稳定。



C.1.8.2 功能架构

数据安全监测技术从多个维度抽取数据，并进行关联、分析，形成风险判断和违规告警，并对风险趋势进行预测，向被监管单位发布数据安全风险预警。

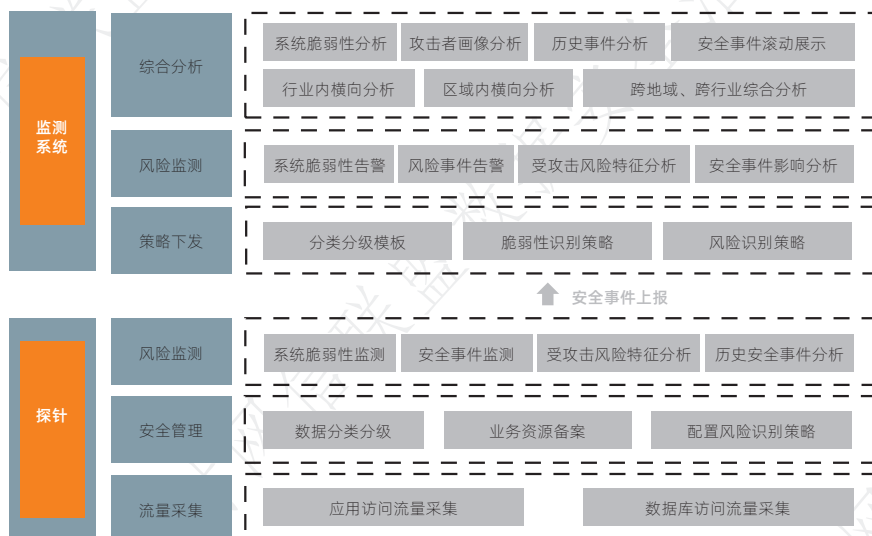


图 67 数据安全监测技术架构图

数据安全监测技术采用分布式系统架构设计，包括监测系统和探针。

监测系统向探针下发数据安全策略，及汇总探针上报的安全事件并进行分析，实时掌握被监管单位的数据安全现状，跟踪被监管单位数据安全事件的处置进度，并综合各监管单位的数据安全整体情况，按区域、行业、被监管单位、受攻击数据资产类型、安全事件数量、安全事件类型等条件进行综合分析，预测某一范围内的数据安全风险趋势。探针支持被监管单位按标准、规范进行数据分类分级、数据流转监测、制定数据风险监测策略及安全风险事件集中监测与响应处置。

监测系统的主要功能包括：

1) 策略下发

向探针下发监管策略，包括数据分类分级标准模板、系统脆弱性识别策略和风险事件识别策略。

数据分类分级标准模板是基于行业属性向各被监管单位下发数据分类分级模板，驱动各单位按照数据分类分级标准，落实数据分类分级工作。

系统脆弱性识别策略是针对数据库、应用 API 接口的脆弱性识别策略。脆弱性识别策略包括脆弱性识别特征库，及按严重、高危、中危、低危定义的脆弱性等级。监测系统仅收集探针上报的脆弱性等级和数量，不收集脆弱性详细信息。

风险事件识别策略是监管单位根据各被监管单位的合规要求下发的风险事件识别策略，并根据事件的严重程度对策略按严重、高危、中危和低危进行等级划分。

2) 系统脆弱性识别

对各监管单位实时上报的数据库脆弱性和应用 API 接口脆弱性的数量和严重程度进行实时统计、告警和趋势分析。



3) 风险事件告警

对被监管单位实时上报的数据安全风险事件实时告警。

4) 受攻击风险特征分析

分析受攻击单位、受攻击数据资产类型、漏洞利用手段、攻击手法（HTTPFlood、SQL 注入等）、每日攻击次数、攻击源所属区域，攻击源活跃度等信息。

5) 分析安全事件对被监管单位产生的影响

综合各被监管单位上报的攻击信息，分析攻击事件对被监管单位所在行业及被监管单位产生的影响。

6) 攻击者身份画像分析

综合各被监管单位上报的攻击信息，对攻击者进行身份画像分析，包括：攻击地址、攻击时间、攻击频率、攻击手段、攻击工具等。

7) 区域攻击事件分析

可在地图上，对被监管单位的所在地、安全事件等信息进行汇总分析，地图上的坐标位置可自动对应到单位所在行政区县，管理员可按照行政区县对被监管单位的整体数据安全态势进行统计分析。

● 行业内横向分析

对行业内各被监管单位上报的安全事件进行分类、关联分析，掌握整个行业当前的数据安全风险现状，预测风险趋势。

● 区域内横向分析

按区域对各被监管单位上报的安全事件进行分类、关联统计，形成区域数据安全风险态势分析。

● 跨地域、跨行业综合分析

跨地域、跨行业对相关被监管单位上报的安全事件进行分类和关联统计，为开展综合数据安全治理提供决策依据。

8) 历史事件分析

系统可对历史数据进行分析，展示历史攻击态势，按事件类别、事件级别、时间等条件进行事件可视化分析。

9) 安全事件滚动展示

针对所有被监管单位，通过滚动表的形式进行告警排名展示，直观了解单位安全状态，便于在重保期间或重大事件发生进行指挥决策。

探针的主要功能包括：

1) 数据分类分级

被监管单位参照数据分类分级模板对业务数据进行分类分级，明确资产价值，有针对性地实施数据安全监管策略。

2) 形成资产地图和数据流转地图

结合数据分类分级结果，对探针采集的流量进行分析，动态梳理数据资产所存储的位置和数据流转的情况，形成数据资产地图和数据流转地图。





3) 应用信息备案

全链路监测业务数据的流转情况，基于字段级的数据访问主客体关系，形成“用户终端 IP- 业务系统账户 - 应用 - API 接口”、“应用 API 接口 - 应用 API 接口”、“业务系统 IP- 业务系统连接数据库账户 - 数据库（表）- 字段”和“运维终端 IP- 运维工具 - 运维账户 - 数据库（表）- 字段”的运维业务资源备案，为实施数据访问控制策略，遵循必要、最小原则，提供基础依据。

4) 策略管理

加载监测系统下发的系统脆弱性识别策略和风险事件识别策略，更新漏洞规则库规则和攻击特征库规则。

5) 系统脆弱性监测

探针数据库和应用 API 接口进行动态脆弱性发现。探针向被监管单位展示脆弱性详情，辅助脆弱性修复，并将数据库脆弱性和 API 脆弱性的严重程度和数量上报到监测系统。

6) 安全事件监测

集中监测数据安全风险事件，跟踪事件处置进度，并上报到监测系统。

7) 受攻击风险特征分析

分析受攻击单位、受攻击数据资产类型、漏洞利用手段、攻击手法（HTTP Flood、SQL 注入等）、每日攻击次数、攻击源所属区域，攻击源活跃度等信息。

8) 历史安全事件分析

● 攻击事件统计

按事件类型、发生时间、攻击源等条件进行攻击事件统计。

● 安全事件级别比例分析

通过可视化方式对安全事件级别的比例进行分析展示。

● 攻击源分析

针对攻击源，可对攻击源分布情况动态展示，包括攻击源所属区域，攻击源活跃度等。

C.1.8.3 关键技术

1) 数据分析技术

应用模型运行引擎对安全大数据进行收集、存储、计算、挖掘和管理，并通过深度学习技术和数据建模技术，赋能数据安全。

● 模型适配

模型引擎是模型执行的基础组件，符合本系统可视化建模标准的模型，均可以被加载到模型引擎，实现模型的各类操作。模型引擎高效识别模型算子和数据资源，将算子调度到引擎组件服务中，并通过模型描述文件中描述的属性参数，调用本地数据资源和接口数据资源来运行生成模型结果。

● DAG 执行

DAG (Database Availability Group) 即数据库可用性组件，是服务器高可用性和站点恢复框架的基本组件。DAG 通过仲裁模式确保服务可用性。仲裁对于确保一致性，充当用于避免分区的关系断开裁判，以及确保群集响应能力而言非常重要。引擎加载模型的 DAG 执行计划，并进行在线解析，





完成解析后，按标准格式，提交到大数据平台上运行。

2) 用户行为分析

采用用户行为分析技术从网络流量中识别出数据安全风险，包括涉敏访问、业务数据未脱敏加密监测、超量访问、频次异常、非工作时间违规操作、密码泄露、批量导出等。

3) 漏洞规则库

系统内置丰富的漏洞攻击库，匹配流量中的攻击行为特征，识别数据安全攻击事件。包括：数据库函数、SQL 命令特征、跨站攻击等。

4) 可视化分析技术

通过可视化分析技术，以安全大数据为基础，辅助数据安全决策，从全局视角提升对安全风险的发现识别和响应处置能力。

C.1.8.4 应用场景

数据安全监测技术可应用于以下监管场景：

- 监管单位对所管辖的单位进行数据安全监管
- 集团公司对各不同分、子公司进行数据安全监管

被监管单位管理在对数据分类分级后，针对核心和业务数据进行持续风险监测，并通过探针将安全事件实时上报到监管单位。

通过探针可实时查看本单位的数据安全整体态势，包括当前正在处置的数据安全事件、历史数据安全事件统计等。如下图所示：

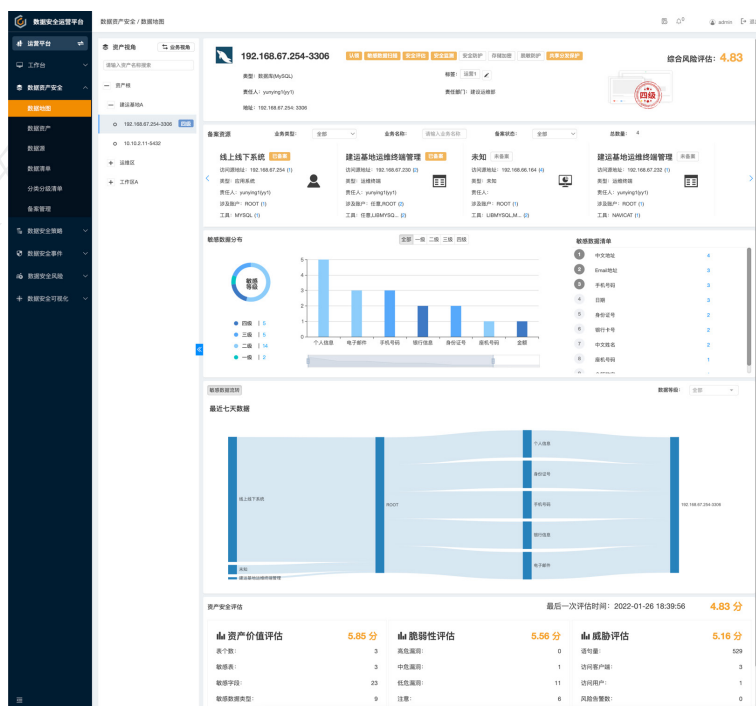


图 68 数据安全事件监测

监管单位对被监管单位下达数据安全监管策略，并实时汇聚、分析各被监管单位上报的数据安全事件，全面掌握数据安全整体态势。

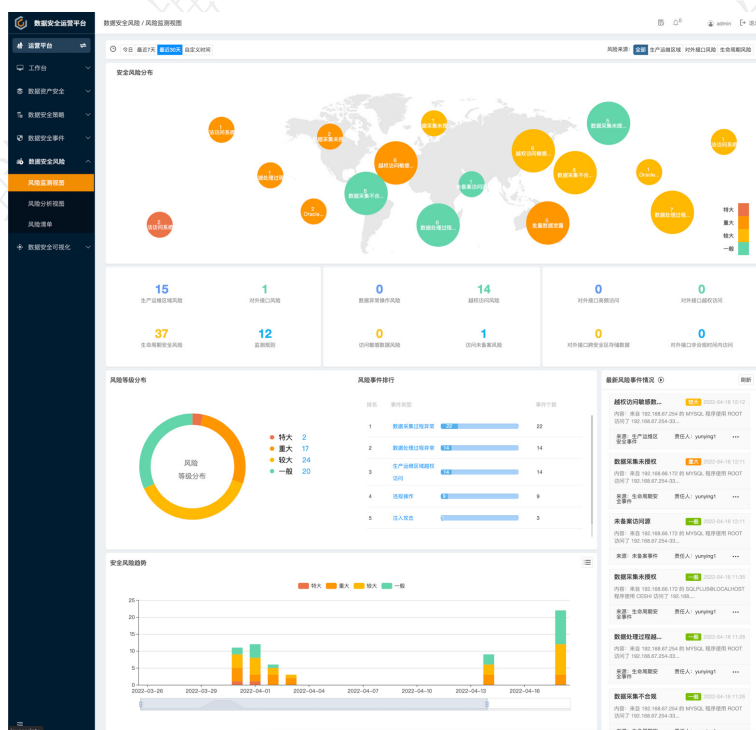


图 69 数据安全态势分析图

C.2 成熟数据安全技术

C.2.1 以数据为中心的审计和防护（DCAP）

以数据为中心的审计与安全防护（DCAP，Data Centric Audit and Protection）是一系列能够集中监控和管理用户与特定数据相关行为的数据安全技术的统称。一些安全企业正在开发机器学习或行为分析能力，通过行为监控和智能分析提供更高层次的洞察力。这种技术基于数据安全治理的原则，将数据安全策略中的控制要求在非结构化，半结构化和结构化数据库或数据孤岛进行实现。其核心功能由数据分类和发现、数据安全策略管理、监控用户权限和数据访问行为、行为分析、警报和阻断、数据加密、令牌化、数据脱敏等数据保护手段构成。其核心技术落地又细分为：数据库审计和保护（DAP）、数据访问治理（DAG）、云访问安全代理（CASB）和数据保护（DP）四个产品方向，不同的进化轨迹，意味着不同的产品在其产品路线图中具有不同的目标和基本功能。虽然没有一款产品完全符合 DCAP 的要求，但这些产品在每种类别中都在完成跨数据处理对象的兼容能力：

1）数据库审计和保护（DAP）

用于实施数据安全策略，数据分类和发现，特权访问管理，数据活动监控或行为分析，审计和数据保护。专注于 RDBMS 和数据仓库，某些产品开始兼容 Hadoop 和非结构化文件共享以及 DBaaS（数据库即服务）。

2）数据访问治理（DAG）

这是有时被称为以文件为中心的审计和保护（FCAP）。通常，这些产品专注于实施文件数据的安全访问策略，数据分类和发现，以及文件存储库和目录服务的活动监视和审计。这些产品与身份和访问管理（IAM）密切相关。一些产品也开始包含云 SaaS 应用的功能。



3) 云访问安全代理 (CASB)

在 SaaS 应用程序或云存储环境中保护数据的能力正在通过多种产品快速增长。这些产品具备跨越 DCAP, 数据丢失防护 (DLP) 和用户实体行为分析 (UEBA) 等能力的数据安全控制。CASB 正在不断发展数据分类和发现, 访问控制, 活动监控, 审计和阻断, 改写, 加密, 标记化和隔离等方面的不同组合。这些产品通常是独立的, 一些 CASB 可以从企业 DLP 产品导入策略, 但它们的管理并不与内部部署 DLP 集成。

4) 数据保护 (DP)

这些产品传统上侧重于通过多个数据仓库 RDBMS, 数据仓库, 非结构化大数据和一些基于云的企业文件同步和共享的加密, 标记化或遮蔽来保护数据。但是, 通过添加实时警报, 活动监控和审计功能, 几项产品已经创新发展。DAP 和 DAG 产品可以提供对文件或数据库中所有数据的访问的监视和审计, 但这些 DP 产品通常只关注敏感数据类型。产品可以使用各种技术通过应用层和 / 或数据层进行连接。通过需要应用层代理, 接口或与特权管理工具的集成, 穿透隐藏了身份标识的连接池, 从而对来自应用层的个人访问进行控制。

C.2.2 数据资产梳理

数据资产梳理是指对目标环境中数据资产进行全面清查、摸排, 通过了解数据资产类型、数据资产分布、数据资产权限、数据资产使用等信息, 构建数据资产目录的过程。在数据安全治理概念中, 数据资产梳理特指敏感数据资产梳理, 这是数据安全体系建设及数据资产管理中的一项基础性工作。

在组织机构内部, 存在着大量的数据资产载体, 例如数据库、存储数据的主机或系统节点, 这些载体中存在大量敏感数据资产, 包括个人信息如身份证号、姓名、联系方式; 金融信息如账户信息、鉴别信息、交易信息、财产信息; 其他信息如企业组织机构信息、纳税人识别号、日期时间等。对这些数据资产进行及时准确的梳理, 掌握其中敏感资产的分布、数量、权限及使用状况, 是进行后续数据安全治理工作的基础与先导。

对数据资产进行梳理, 主要包括两种技术: 一种是静态梳理技术, 通常采用对 IP 地址段和端口范围进行扫描的主动嗅探方式; 另一种是动态梳理技术, 通常采用对网络流量进行协议分析的被动监测方式。静态梳理技术被用于形成敏感数据字典, 动态梳理技术被用于形成敏感数据访问状况清单, 二者共同为后续数据安全治理建设提供必要的基础信息。

C.2.2.1 静态梳理技术

静态梳理技术是指对目标环境内数据资产载体进行探测发现、并对数据资产载体 (如数据库) 内部的敏感数据的类型、分布、数量进行扫描匹配, 对数据载体如数据库内部的账号权限情况进行梳理, 形成静态目录的过程。

在静态梳理的过程中, 根据梳理的对象和阶段不同, 主要包括三种技术手段。

1) 数据载体的自动探查与发现

数据载体的自动探查与发现, 是指通过对目标网段内的 IP、端口进行基于数据库协议的某种探测过程, 判断其主机上是否存在对外提供服务的数据库。在数据存储环境中, 可根据基于 IP 和端口的主动探测, 发现网段内开放了数据服务的数据资产载体 (数据库), 通过构造针对性的探测语句





或协议，可进一步获取数据库的类型、版本号和其他相关信息。通过数据库的自动探查与发现，能够形成目标环境内的数据资产载体情况清单，在形成清单后，可为后续进行安全建设做出指导。

2) 敏感数据发现与梳理

敏感数据发现与梳理，是指通过对数据载体中存储的数据进行抽样提取，并与数据特征模板进行匹配，进而判断其是否属于某种敏感数据，然后，再针对数据环境中的敏感数据，采用抽样扫描匹配加人工核实筛查的方式，形成敏感数据分布清单。进行敏感数据梳理的步骤通常分为如下四步：

- 对敏感数据进行采样，与特征库进行匹配，能够对数据特征明显且符合标准的数据进行类型精确判断，如身份证、银行卡号、手机号等；

- 对于只能根据数据自身属性判断其数据特征，但是无法判断数据类型的数据，则需要在自动特征库匹配的基础上，结合业务环境进行人工核实确认。例如 2020-10-12 12:00:03，可以根据数据特征自动判断为“日期时间”，但是这类数据归属于“出生日期”或是“开户日期”则需要根据业务环境人工梳理判断；

- 对于无法根据数据自身属性判断其数据特征，但是该类型数据取值有范围，可以通过枚举或遍历的方式进行判别的，可通过构建数据字典，并通过字典匹配的方式判断其数据类型。例如“济南、上海、天津、杭州”可通过“中国城市字典”判别为是城市名称；

- 对于上述三类都无法判别其数据类型与数据特征的资产，则需要人工结合实际环境，手动进行梳理。

3) 账号权限梳理

账号权限梳理，是指通过对数据载体的系统表或配置信息进行扫描、查询数据库中系统表和系统对象，判断其提供访问的账号情况，梳理出目标数据环境中存在哪些可登录可访问数据的账号，以及账号对数据资产尤其是敏感数据的访问权限情况。

C.2.2.2 动态梳理技术

动态梳理技术是一种基于协议及流量分析，识别敏感数据的访问状况的技术。动态梳理技术通过获取对数据载体的访问流量并进行解析，从中判别出敏感资产的访问状况包括访问源身份与位置、访问时间、访问频次、访问返回结果集内容与数量。

动态资产梳理主要依托于数据库协议精确解析技术，通过端口镜像或者在数据库所在主机部署探针引流的方式，将涉及数据资产访问的流量获取后进行协议分析。从流量中可以获取到以下三类信息：

- 访问源信息

主要包括访问资产的主机信息、用户信息等。

- 访问对象信息

主要包括被访问的数据资产载体、数据资产内容、访问的应答、结果集等。

- 其他访问信息

主要包括访问语句、访问工具、访问时间等信息。

通过对上述信息的汇总分析，可以得出数据资产的访问源统计，访问热度，访问频次，以此形成动态资产梳理清单。



动态资产梳理结果对于掌握数据资产的使用状况，从而进行数据安全综合评估，数据安全策略设定具有指导意义。

C.2.3 身份认证和访问管理 (IAM)

C.2.3.1 技术概述

IAM 是一套全面地建立和维护数字身份，并提供有效地、安全地进行 IT 资源访问的业务流程和管理手段，从而实现组织信息资产统一的身份认证、授权、访问控制和身份数据集中管理与审计。

IAM 体系建设的核心意义在于帮助组织进行应用系统的统一管理，从而提高组织数据资产的可管理性，同时，便于组织实施进一步的安全保护措施。IAM 体系对数据安全的意义主要体现在三个方面：

首先，实现应用系统的单点登录 SSO，用户通过 IAM 访问门户首次认证登录成功之后，在系统设定的票据有效周期中，可多次访问组织内各业务应用，用户身份票据自动保持、传递，不再需要进行多次身份认证。单点登录为组织内应用系统提供了统一的入口，整合了组织资产，降低了由用户账号被窃导致的数据资产遗失等安全隐患。

其次，IAM 体系可设计完善的权限管理机制，进行严格的访问控制，避免了数据资源的过度访问导致的数据泄露，配合完善的安全审计规则，可实现数据访问事前事后的全面管理。

另外，IAM 体系通过构建完整可信的用户资源信息库，整合各应用系统的用户库，集中实现用户信息统一管理，并通过数据同步功能向应用系统同步，大大提高了组织数据资产的可管理性，也便于系统的安全审计。

C.2.3.2 技术框架

IAM 体系的建设包括四个子域：身份数据、身份管理、访问执行及访问管理，通常分为身份管理、访问管理及门户三部分，集身份管理、身份认证、授权管理、应用资源访问控制及其安全审计于一体，构建多信息资源的应用整合和安全防护的安全基础服务平台。IAM 的体系框架如图所示：

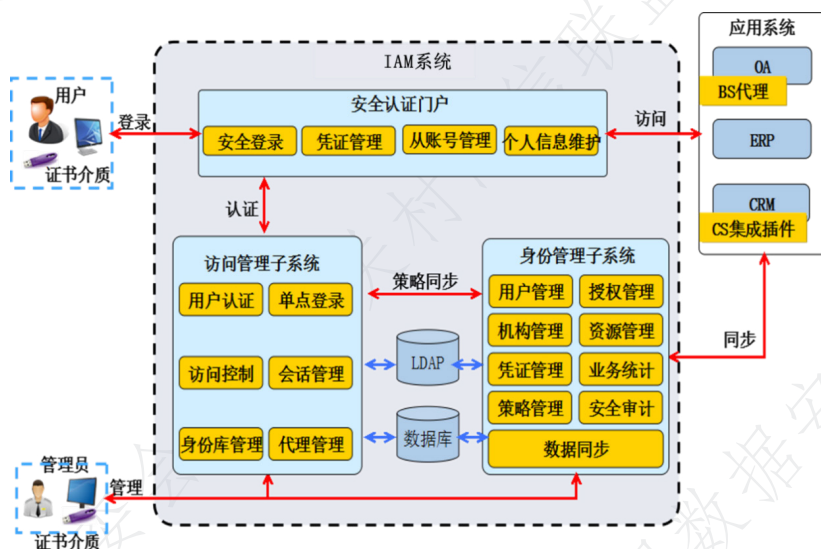


图 70 IAM 的技术体系框架



安全认证门户：安全认证门户为用户集中展现可访问的应用列表，为用户提供统一的认证入口。认证门户能对用户所提供的数字身份凭证进行在线认证，识别用户身份的真实性，并通知各个应用系统，同时在门户中提供单点登录功能，实现门户与各应用系统之间进行一次性认证与登录。

身份管理子系统：身份管理子系统提供用户身份完整的生命周期管理、应用系统管理，身份管理子系统可以通过接口与业务系统进行用户信息同步，同时也可以为访问管理子系统提供用户信息，可以通过本系统进行行为审计查询。

访问管理子系统：访问管理子系统主要为用户提供集中的身份认证机制，采用数字证书技术进行身份鉴别，提供面对业务系统的访问控制机制，用户全行为流程进行跟踪审计。

C.2.3.3 访问流程

IAM体系的用户访问业务系统的流程如图所示。IAM体系建设后，每一用户拥有唯一的账户信息，用户访问应用系统时，将请求重定向到IAM服务器，进行统一的身份认证和权限审核，认证通过后，即跳转进行业务访问，当用户需访问其权限范围内的其他应用系统时，无需再次进行身份认证。同时，通过基于机构、岗位、用户组和角色等不同维度的授权管理，灵活地实现对用户的权限控制，通过全面的用户行为数据收集，实现统一的安全审计，对用户的行为日志进行审计和综合分析。

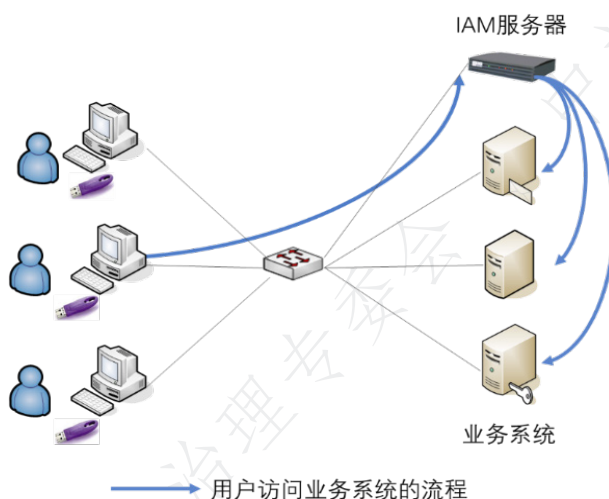


图 71 访问流程图

C.2.4 数据安全采集技术

C.2.4.1 技术概述

数据安全收集技术能够对工业现场的各类传感器、设备、智能产品/装备等设备进行接入和数据收集，并通过 5G/4G、NB-IoT、GPS、WiFi、有线等各类有线和无线通信技术，转发收集的数据或下发操作指令。

C.2.4.2 功能架构

数据安全采集技术包括数据采集设备的发现和识别、设备指纹管理、准入与访问控制、网络接



入异常监控、数据采集合规分析、流量行为监控、信令白名单检查、和探查分析。

1) 数据采集设备发现和识别

对数据采集设备进行资产管理，管理设备的特征指纹、设备类型、用户、操作系统、品牌、厂商、IP/MAC 信息，形成“一机一档”管理体系。有新增前端设备，准入系统通过被动监测方式发现并更新设备特征指纹信息。

2) 设备指纹管理

准入引擎内置指纹库，覆盖绝大多数终端类型，可通过指纹管理功能手动将该设备指纹提取进指纹库。

3) 准入与访问控制

数据安全采集技术对接入的数据采集设备进行指纹认证，如发现接入的数据采集设备指纹信息与登记的设备指纹信息不符，则断开采集设备的网络连接，并通知管理员进行处置。

4) 网络接入异常监控

数据安全采集技术具有仿冒检测能力，持续检测数据采集设备运行状态，当发现设备存在网络环境变更（如网络中断重新接入），或网络行为出现变更时触发仿冒确认过程，当发现设备硬件信息（MAC 地址）、指纹信息发生变化时，确认设备仿冒事件发生。

当确认设备发生仿冒，准入系统需要通过连接重置的方式对设备网络流量进行限制或联动网络设备的方式直接切断设备网络连接。同时记录完整的发现及处置日志，满足合规审计要求。

5) 数据采集合规分析

调研业务系统的目标和功能，梳理业务系统数据结构，通过数据采集的流量和日志信息，自动关联分析用户在系统上授权同意的相关操作行为与数据采集入库、数据在业务中使用、数据自动化决策等系统操作行为，生成《数据采集用户同意报告书》，向监管人员提供数据合法合规采集的相关监管报告，为监管人员评估数据采集、使用是否遵循用户同意原则提供数据支撑。

整理系统要收集的敏感数据特征，通过数据识别技术，审计业务系统采集的各类敏感数据，发现未列入数据采集范围的敏感数据采集行为，生成《数据超范围采集报告书》，向监管人员提供数据合法合规采集的相关监测报告，为监管人员评估数据采集是否遵循必要、最小原则提供数据支撑。

6) 流量行为监控

通过监测设备出入站流量，持续计算周期内设备使用上下行流量的平均值，当发现设备上下行流量波动超过阈值时及时告警并处置。

7) 信令白名单检查

检查发向数据采集设备的控制信令内容，匹配信令白名单列表，只放行白名单列表中允许放行的控制信令，阻断未登记在信令白名单列表中的控制信令传输到数据采集设备。

8) 探查分析

数据安全采集技术根据运维人员策略配置，能够对发现的数据采集设备进行全方位、多维度的深入识别。通过设备指纹特征，资产库、自动学习等技术，发现其 IP/MAC、操作系统、设备类型、厂商、型号、序列号、开放端口、接入位置等信息。

资产发现后可以对设备进行自动分类分组，并基于组织结构进行统计分析、基于类型和厂商进



行统计分析、基于接入位置进行统计分析，生成图形化报表，支持快速信息筛选。

C.2.4.3 关键技术

数据安全采集技术支持 RS485、RS232、CAN、ADC 等常见的工业接口接入，并对 Modbus、CAN、OPC 等协议解析与转换，实现数据格式转换和统一。

支持 SM 国密加密通信，实现数据加密传输，防止数据泄露、被侦听或篡改，保障数据采集和传输过程中的安全。

C.2.4.4 应用场景

数据安全采集技术可应用在各数据采集领域。如在互联网个人数据采集方面，可对数据采集的移动设备进行终端身份认证，数据加密传输，并收集数据采集流量，分析互联网用户个人数据采集是否满足合法、合规要求；在数控机床监测方面，运用数据安全采集技术，可以安全的接入数控机床设备，接入机床设备数据，远程监测并控制机床设备状态；在桥梁结构安全监测方面，通过数据安全采集技术，采集桥梁结构点的变形、受力、环境等关键信息，实现对桥梁结构的实时安全监测、远程管控、风险告警和统计分析。

C.2.5 数据脱敏

数据脱敏，是指对数据中包含的秘密或隐私信息（如个人身份识别信息、商业机密数据等）进行数据变形处理，使得恶意攻击者无法从经过脱敏处理的数据中直接获取敏感信息，从而实现对机密及隐私信息的保护。数据脱敏技术能够消除数据的信息敏感性，有效保证重要数据在分析、监管协作、开发测试等过程中的安全性。

C.2.5.1 脱敏算法

脱敏算法是指对数据进行脱敏变形中采用的具体方法，按照技术和效果不同，可分为泛化、抑制、扰乱三大类，在每一类中，又会有多种不同的具体实现方式。

1) 泛化类脱敏算法

泛化是指在保留原始数据局部特征的前提下使用一般值替代原始数据，泛化后的数据具有不可逆性，具体的技术方法包括但不限于：

● 截断

舍弃部分不需要的数据信息。数据截断后的结果往往无法较好地保持原有业务属性，因此在对数据截断时，应根据数据特点酌情选择截断位数。例如把身份证号 120104199804081123 截断为 12010419980408。

● 取整

按照一定区间与步长对数据进行向上或向下取整，此类急加速可在保证数据一定分布特征的情况下隐藏数据原始属性。取整的方法主要通过舍弃一定的精度来保证原始数据的安全性，可以一定程度上保持数据业务特性上的分布密度，适用于粗略统计分析的场景。例如将金额 6023.62 元按照百



位粒度向下取证得到 6000 元。

2) 抑制类脱敏算法

抑制是指通过遮蔽数据中部分信息的方式来对原始数据的值进行替换，又称为隐藏技术。该技术只保留部分信息，对敏感数据的部分内容用通用字符（如“#、*”等）进行统一替换，从而使得敏感数据保持部分内容公开，例如将手机号码 13500010001 经过掩码得到 135***0001。

3) 扰乱类脱敏算法

扰乱是指通过加入噪声的方式对原始数据进行干扰，以实现原始数据的扭曲、改变，扰乱后的数据仍保留着原始数据的分布特征，具体的技术方法包括但不限于：

● 重排

将原始数据按照特定的规则进行重新排列，对于跨行数据，采用随机互换来打破其与本行其他数据的关联关系，从而实现脱敏。例如：字母 abcde 重排为 adecb，按照一定的顺序进行打乱。重排可以在相当大范围内使脱敏后数据看起来跟原始数据更一致，与此同时也牺牲了一定的安全性。一般重排方法用于大数据集合且需要保留待脱敏数据特定特征的场景。

● 加密

对脱敏数据进行对称加密算法、非对称加密算法等常规加密算法处理，使外部用户只能看到无意义的加密后的数据，同时在特定场景下，可以提供解密能力。加密其安全程度取决于采用哪种加密算法，一般根据实际情况而定。

● 替换

按照特定规则对原始数据进行替换。常见的替换方式包括常数替换、查表替换、参数化替换：

➢ 常数替换（置常数）：所有敏感数据都替换为唯一的常数值，具有不可逆性。例如，将 500 替换为 0，将 635 替换为 0；

➢ 查表替换：从中间表中随机或按照特定算法选择数据进行替代；

➢ 参数化替换：以敏感数据作为输入，通过特定函数形成新的替换数据。

● 散列

对原始数据取散列值，使用散列值来代替原始数据；例如使 MD 散列函数，将客户的密码“123456”MD 哈希算法运算后得到“e10adc3949ba59abbe56e057f20f883e”；为了保证散列的安全性，建议避免采用弱安全性散列函数（如 MD5、SHA-1）。对于原文空间有限的散列，实际的应用场景中还建议加入随机因子。散列函数常用于密码等敏感信息存储的场景。

● 重写

参考原数据的特征，重新生成数据。重写与整体替换较为类似，但替换后的数据与原始数据通常存在特定规则的映射关系，而重写生成的数据与原始数据则一般不具有映射关系。例如对手机号码，可在一定范围内按照规则随机生成的方式构造数据，对身份证号，可在保证身份证特征“前六位是归属地，中间 8 位是出生年月，倒数第二位是性别”等特征生成数据。

● 固定偏移

将数据值增加 n 个固定的偏移量，隐藏数值部分特征；

● 局部混淆



保持前面 n 位不变，混淆其余部分，例如：保持座机号码区号不变的情况下，对座机号码 0571-123456 进行混淆脱敏后得到 0571-328192；

- 唯一值映射

将数据映射成一个唯一值，允许根据映射值找回原始值，支持正确的聚合或者连接操作；

- 均化

针对数值性的敏感数据，在保证脱敏后数据集总值或平均值与原数据集相同的情况下，改变数值的原始值，这种方法通常用于成本表、工资表等场合。

C.2.5.2 应用场景

脱敏应用技术指脱敏技术在具体应用中，根据实用性与场景不同的不同表现，分为静态脱敏技术与动态脱敏技术两种，静态数据脱敏一般用在非生产环境，将敏感数据从生产环境抽取并脱敏后用于培训、分析、测试、开发等非生产环境。动态数据脱敏一般用在生产环境，将敏感数据实时进行脱敏后用于应用访问等生产环境。

1) 静态脱敏

静态数据脱敏旨在通过类似 ETL 技术的处理方式，按照脱敏规则一次性完成大批量数据的变形转换处理。静态脱敏通常会在将生产环境中的敏感数据交付至开发、测试或者外发环境时使用，在降低数据敏感程度的同时，能够最大程度上保留原始数据集所具备的数据内在关联性等可挖掘价值。

静态数据脱敏主要特点：

- 适应性，即可为任意格式的敏感数据脱敏；
- 一致性，即数据脱敏后保留原始数据字段格式和属性；
- 复用性，即可重复使用数据脱敏规则 and 标准，通过定制数据隐私策略满足不同业务需求

2) 动态脱敏

动态数据脱敏旨在通过类似网络代理的中间件技术，按照脱敏规则对于外部申请访问的数据进行及时处理并返回脱敏后结果。动态脱敏通常会在数据对外提供查询服务的场景中使用，在降低数据敏感程度的同时，最大程度上降低了需求方获取脱敏后数据的延迟，请求实时产生的数据也能即时得到脱敏后结果。

动态数据脱敏主要特点：

- 实时性，即能够实时地对用户访问的敏感数据进行动态脱敏、加密和提醒；
- 多平台，即通过定义好的数据脱敏策略实现跨平台间、不同应用程序或者应用环境间的访问限制；

- 可用性，即能够保证脱敏数据的完整，满足业务系统的数据需要。

3) 动态双因子可逆脱敏

对于数据脱敏的可恢复数据能力提供更高更安全的算法能力极其重要，在非常多的业务场景中，都需要脱敏后数据具备可恢复能力，而脱敏后数据又需要满足一定的使用要求，格式要求，存储要求，性能要求等，这就对于算法安全性要求提出了新的挑战，通过多因子，动态双因子算法，从根本上去除脱敏后数据映射规律问题，让映射关系始终处于动态变化中，从而提供更为高效安全的算法能力。





同时我们加入格式保留技术，格式保留脱敏具有脱敏后数据格式和数据长度不变的特点，不会破坏数据格式约束，从而降低改造数据格式的成本。

传统现有敏感信息格式保留脱敏方案，均基于对称脱敏体制，存在密钥传输安全性低和密钥管理成本较高等问题，在此前提下，我们提出了多因子技术能力，基于格式保留的敏感信息脱敏方案，与现有的格式保留脱敏方案相比，脱敏和还原双方不需要传递多个密钥，通过密钥派生函数来生成脱敏密钥和还原密钥，利用混合脱敏的方式提高了敏感信息传输的安全性。

格式保留可逆脱敏基于数据不扩充，数据类型不改变，数据必须具备确定的保密性的原则。

该方式参考分组密码的思路，密钥共享的思路，通过动态多因子的方式实现密钥分组思路，同时通过格式保持技术，保证了数据在脱敏后依然具有其原有格式要求，从而满足在数据脱敏的数据使用。

C.2.6 数据加密

C.2.6.1 数据机密性保护

数据机密性是指保证数据不被泄露给非授权的个人、计算机等实体的性质，是网络与信息安全的重要属性之一。采用密码技术中的加密保护技术，可以实现数据的机密性保护。利用实现加密的一段计算机程序，对数据进行加密，可产生形如乱码的密文。攻击者即使截获了密文，由于加密算法具有足够的强度，也无法从密文中获取有价值的信息。而拥有密钥的一方，利用实现解密的一段计算机程序，可以从乱码中恢复原来的数据。

机密性保护技术包括基于对称密码的机密性保护和基于非对称密码的机密性保护。不同机密性保护技术应用用于传输、存储场景时的功能需求不同。数据传输过程中机密性保护的主要对象包括：传输的重要数据、敏感信息数据或整个报文、存储的重要数据和敏感信息数据、身份鉴别信息、密钥数据等。数据传输过程中的机密性保护技术包括对称密码技术和非对称密码技术，其中 SM4、ZUC 可作为对称密码的传输机密性保护，SM2、SM9 可作为非对称密码的传输机密性保护。对称密码保护的数据对象更多是实际应用数据，非对称密码保护的数据对称往往是密钥数据。数据存储机密性保护的数据既涉及存储于物理存储介质的数据，又涉及存储于逻辑数据结构中的数据。数据存储的机密性保护一般采用对称密码技术，SM4、ZUC 可作为对称密码的存储机密性保护。

1) 公钥加密技术

公钥加密，又称为非对称加密，在数据加密和解密的过程中使用不同的密钥。其中加密的密钥被公开，称为公钥，解密的密钥被保密，称为私钥。公钥和私钥是密切关联的，从私钥可以推导出公钥，但从公钥不能推导出私钥。

在应用公钥加密时，每个用户都有自己的一对公钥和私钥，需要跟谁秘密通信，只需用谁的公钥对消息进行加密，具有相应的私钥的人可以解密得到明文，而没有相应的私钥的人无法进行解密。公钥加密的速度一般比对称加密慢，主要用于运算数据的加密，例如，用公钥加密算法建立共享密钥。

SM2 算法是我国颁布的商用密码标准算法中的公钥密码算法。SM2 算法基于椭圆曲线上离散对数计算困难问题，密钥长度为 256 比特，具有密钥长度短、安全性高的特点。



2) 对称加密技术

对称加密的加密过程与解密过程使用相同或容易互相推导得出的密钥，即加密和解密两方的密钥是“对称”的。这如同往一个上了锁的箱子里放物品，放入时需要用到钥匙打开，想要取出物品时需要用同样的钥匙开锁。针对不同的数据类型和应用场景，对称密码算法有两种主要形式：一是序列密码，二是分组密码。

● 序列密码算法

序列密码适用于实时性要求较高的场景，如电话、视频等。序列密码的加密过程中将密钥和明文数据的字母逐个进行“相加”运算，解密时再用相同的密钥对密文进行逐个字母的“相减”运算。序列密钥需要快速产生一个足够长的密钥，因为有多长的明文，就要有多长的密钥。为此，序列密码的主要任务是从一个短的初始密钥，快速产生一个足够长的“密钥流”。因此，序列密码也叫做“流密码”。

ZUC 算法是我国颁布的商用密码标准算法中的序列密码算法，也成为祖冲之算法。ZUC 算法采用 128 比特的密钥和 128 比特的初始向量作为输入参数，算法体现了序列密码设计上的发展优势，具有较大的安全余量，算法速度快，软硬件实现性能都比较好，ZUC 算法已作为国际第四代移动通信加密标准使用。

● 分组密码算法

分组密码是讲明文数据分成多个等长的块，这样的数据块叫做分组。每个分组以同样的处理过程、同样的密钥进行加密和解密。它的加密过程一般采用混淆和扩散部件的多次迭代方式。SM4 是我国颁布的商用密码标准算法中的分组密码算法。SM4 算法的分组长度为 128 比特，密钥长度为 128 比特，解密算法与加密算法相同。它具有算法速度快、实现效率高、安全性好等优点，主要用于保护数据的机密性。

C.2.6.2 数据完整性保护

数据的完整性是指数据没有受到非授权的篡改或破坏的性质，它是网络与信息安全的一个重要属性。

数据完整性保护技术包括基于杂凑的完整性保护、基于消息鉴别码的完整性保护和基于数字签名的完整性保护。传输的完整性保护包括：基于消息鉴别码的传输完整性保护、基于数字签名的传输完整性保护。存储的完整性保护包括：基于消息鉴别码的存储完整性保护、基于数字签名的存储完整性保护、基于杂凑算法的存储完整性保护、基于可信计算的存储完整性保护。

1) 杂凑算法

杂凑算法是将任意长的消息压缩成固定长度的摘要值的函数，具有抗碰撞性和单向性。抗碰撞性是指寻找两个不同的消息，使二者的摘要值相同，在计算上是不可行的。单向性是指从摘要值不能反向求出输入消息。

杂凑算法的作用很多，它可以用于数字签名：先对消息进行压缩后，再对摘要值进行签名，因为很长的文件不能一页一页的简明。它可以检测消息的完整性，使用杂凑算法计算出数据的摘要，将这个新的摘要与之前保存的摘要进行比对，如果两个摘要一样，就说明这个文件没有被改动过。





它可以用于口令的安全存储，利用单向性的特点，即使暴露了摘要值也不会暴露口令。

SM3 算法是我国商用密码标准中的杂凑算法。SM3 算法摘要值长度为 256 比特，算法速度快，适用性强，可用于数字签名、口令安全存储等场景。

2) 数字签名

在应用数字签名时，每个签名者都有一对公钥和私钥，公钥公开，私钥保密。与公钥加密过程使用公钥和私钥的顺序不同，数字签名使用私钥对消息进行签名，这一过程成为签名过程，使用公钥对签名进行验证，这一过程成为验证过程。

数字签名技术可以用于确认数据的完整性、签名者身份的真实性和签名行为的不可否认性。因为数字签名是用私钥产生的，所以数字签名是不可伪造的。因为签名者和公钥一一对应的关系，可以用公钥对签名进行验证，因此合法的签名是可以公开验证的。因为只能用签名者的公钥进行验证，所以签名者不能否认自己签过的签名。

C.2.6.3 数据库加密

通常情况下，数据库中的数据是以明文形式进行存储和使用的，一旦数据文件或备份磁带丢失，可能引发严重的数据泄露问题；而在拖库攻击中，很多成熟的数据库文件解析软件，均可对明文存储的数据文件进行直接分析，并输出清晰的、结构化的数据，从而导致泄密。数据库加密技术可对数据库中存储的数据在存储层进行加密，即使有人想对此类数据文件进行反向解析，所得到的也不过是没有任何可读性的“乱码”，有效避免了因数据文件被拖库而造成数据泄露的问题，从根本上保证数据的安全。

对于数据库加密技术，从加解密位置区分，可以分为应用层加密、数据库中间件加密、数据库系统自身加密、数据库系统所依赖的文件系统加密以及存储硬件加密等不同技术路线。

1) 应用层加密

在应用系统层的源代码中对敏感数据进行加密，加密后将密文存储到数据库中。可以直接在应用系统的源代码中以独立的函数或模块形式完成加密；也可以通过源代码的方式封装出应用系统相关业务专用的加密组件或定制的加密 API 来完成加密。

通常情况下，当业务系统仅对有限的敏感数据存在加密需求时，可以考虑使用应用系统加密技术。这里的“有限”包含两方面含义：一方面，是需要加密处理的敏感数据对应的表或字段相对较少；另一方面，是需要加密处理的敏感数据在整个业务系统中的使用相对不多。

2) 前置代理加密

类似于应用系统加密技术，前置代理加密技术也是在数据保存到数据库之前对敏感数据进行加密，并将密文存储到数据库中；而不同于前者的是，前置代理加密技术通常是以“前置代理加密网关”这种独立组件产品的形式实现的。

通常情况下，当业务系统仅对有限的敏感数据存在加密需求，且用户自身无能力或不愿意进行加解密的相关研发工作时，可以考虑使用前置代理加密技术，即采用第三方厂商的前置代理加密网关系统对敏感数据进行加密保护。





3) 后置代理加密

后置代理加密技术是基于数据库自身能力的一种加密技术，可充分利用数据库自身提供的定制扩展能力实现数据的存储加密、加密后数据检索和应用透明等目标。其中，能够完美实现后置代理加密的典型代表是 Oracle 数据库——通过“视图 + 触发器 + 扩展索引 + 外部方法调用”的方式实现数据加密，同时保证应用的完全透明。

后置代理加密技术的价值主要体现在“应用透明和独立权控”两大能力上。因此，在不希望对应用系统在加密时进行改造，或需要对数据库超级用户的数据访问权限进行控制的场景下，如果查询涉及的加密列不多且查询结果集中包含的数据记录也相对不多时，可以考虑使用后置代理加密技术对数据库进行加密。

4) 数据库透明加密

透明数据加密（TDE, Transparent Data Encryption），是一种对应用系统完全透明的数据库端存储加密技术，通常由数据库厂商在数据库引擎中实现：在数据库引擎的存储管理层增加一个数据处理过程，当数据由数据库共享内存写入到数据文件时对其进行加密；当数据由数据文件读取到数据库共享内存时对其进行解密。也就是说，数据在数据库共享内存中是以明文形态存在的，而在数据文件中则以密文形态存在。同时，由于该技术的透明性，任何合法且有权限的数据库用户都可以访问和处理加密表中的数据。

数据库透明加密能够在无需数据库使用者改变访问接口甚至无需知情（透明）的情况下，使得数据库内存储的数据为经过加密的密文数据。这种技术的核心优点在于其对数据库使用者的高度友好性，几乎无需对使用数据库的已有应用（业务）系统进行任何改造和升级，就能实现对结构化数据的机密性进行有效保护。

透明数据加密技术凭借其自身的上述优势，几乎适用于全部有数据库加密需求的应用场景，尤其是在对数据加密透明化有要求，或需要对数据库超级用户进行数据访问权限控制，以及对数据加密后数据库性能有较高要求的场景中。

● 功能需求

➤ 数据库透明加密技术的关键就是对应用透明、使具有权限用户的数据库访问对加密“无感”。

➤ 应兼容主流加密算法，支持国家指定的 SM4 加密算法。

➤ 权控体系支持“三权分立”，在传统的数据库超级管理员（DBA）权控基础上，增设数据安全管理员（DSA），数据库安全审计员（DAA）。DBA 和 DSA 相互独立，在不影响数据库本身权限的同时，增强了密文的权限控制，分别从数据库用户，客户端 IP，应用系统等不同层面对密文的访问权限进行了控制，全面防止越权访问，防止数据泄露。而 DAA 可以对 DSA 的配置情况作出有效审计。

➤ 数据加密后，对数据库的性能影响应尽量小，特别对加密数据的查询、分组、排序、聚合等操作。

● 技术路线

对数据进行透明加密可在不同层次、不同位置进行，下图展示了若干典型多层数据加密方案在透明性、性能影响、安全强度等方面的不同效果和表现。



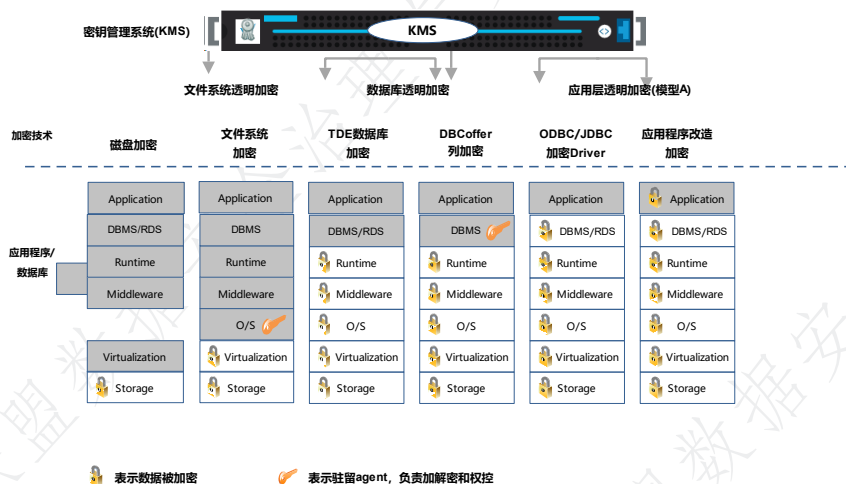


图 72 多层数据加密方案效果对比

➤ 列加密

数据库列透明加密是按列对数据进行加密的，针对指定的列，采用指定的加密算法和密钥、盐值等进行加密处理。加密后的数据以密文的形式存储在 DBMS 的表空间中。

只有经过授权的用户才能看到明文数据，并且授权也是按列进行的，这种方式具备很好的灵活性和安全性。非授权用户，将无法读取（查询）加密列和更改加密列的数据。

权限管理上，数据库列透明加密采用了分权的机制，实现了三权分立，有效制约了数据库管理员（DBA）这样的特权用户对数据的访问。同时这种保护又是透明的，不会对管理员的日常工作造成不便。

数据库列透明加密具有应用透明的特性，应用系统和外围维护工具无需改造，涵盖 SQL 语句透明、存储程序透明、开发接口透明、数据库对象透明、管理工具透明。

下图说明了列加密的数据处理机制：当明文数据到达数据库时，被列加密组件自动进行加密处理，之后数据以密文形式保存在数据文件中；在读取数据时，数据被列加密组件根据权限进行自动的解密处理。这些操作对应用都是透明的。

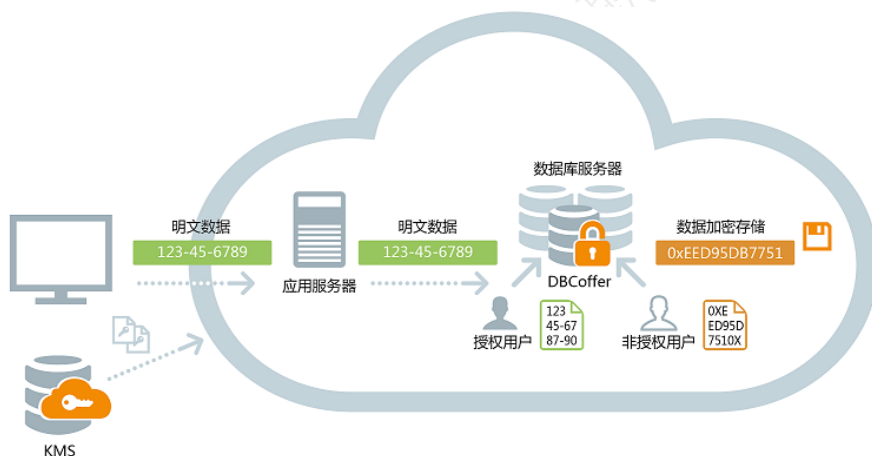


图 73 数据库列透明加密



➤ 表空间加密

下图以 Oracle 数据库为例，说明了表空间透明加解密的关键机制：TDE 插件作为 Oracle 的插件，负责对数据存储时对数据进行加密处理；在数据从表空间读取时，负责对数据进行解密处理；并且根据读取数据的数据库账户，进行相应的权限控制。

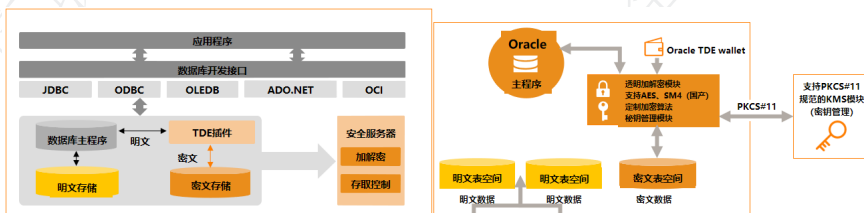


图 74 表空间 (OracleTDE) 加密技术

5) 文件系统加密

文件系统加密技术是在操作系统的文件管理子系统层面对文件进行加密，大多是通过与文件管理子系统相关的操作系统内核驱动程序进行改造实现的。不同于文件加密只对单个文件设置访问口令，或对单个文件的内容进行加密转换，文件系统加密提供了一种加密文件系统格式（类似于 ext4、xfs 等文件系统格式），通过把磁盘存储卷或其上的目录设置为该文件加密系统格式，达到对存储于卷或卷上目录中文件进行加密的目的。文件系统加密技术本质上并不是数据库加密技术，但可以用于对数据库的数据文件进行存储层面的加密。

文件系统加密技术几乎可以适用于任何基于文件系统的数据库存储加密需求，尤其是原生不支持透明数据加密的数据库系统和大数据数据库系统。但是，由于文件系统加密技术无法提供针对数据库用户的增强权限控制，对于需要防范内部数据库超级用户的场景并不适用。

文件系统级透明加密防护是通过在主机操作系统上部署专门的加解密 agent，实现对专门的数据文件或卷 (Volume) 的加解密。

对操作系统账户具有权限控制能力，通过指定只有专门的 DBMS 系统账户才有对文件或卷进行加解密的权限，没有权限的程序无法对数据文件进行读写操作。

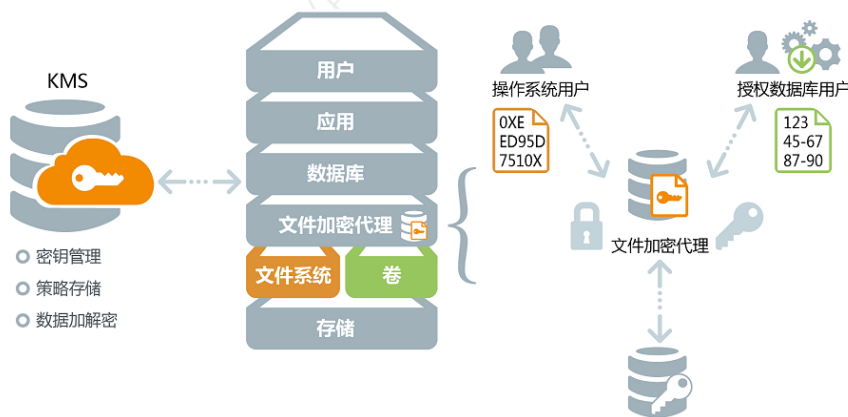


图 75 文件系统级透明加密

文件系统透明加密技术路线的主要优势有：（1）对数据库系统（DBMS）和应用透明；（2）同



时支持结构化（关系型数据库）和非结构化文件（NoSQL 数据库）的加密；（3）能有效控制操作系统用户的访问，满足通常的控制需求，例如数据库的执行账户（oracle、mysql、root 等）。

文件系统透明加密技术路线的主要局限有：（1）无法提供细粒度的数据访问审计能力；（2）需要针对不同的操作系统平台提供专门的 agent；（3）无法控制数据库账户对敏感数据的访问，例如 DBA 账户。

6) 磁盘加密

磁盘加密技术通过对磁盘进行加密以保障其内部数据的安全性，从实现上有软硬两种方式：软件方式的磁盘加密技术，大多是通过专用的磁盘加密软件对磁盘内容进行加密，典型代表如 Windows 操作系统自带的 BitLocker，同类型的商业软件在国内也有很多，但这类软件由于加密原理和使用方式等因素，基本上无法满足数据库系统的数据加密需求；而硬件方式的磁盘加密技术，在实现上则有两个思路：一种是针对单块硬盘的磁盘加密，一种是针对磁盘阵列或 SAN 存储设备的磁盘加密。

磁盘加密技术可以满足几乎所有的存储加密需求，但由于不能提供针对数据库用户的增强权限控制，无法满足对内部数据库超级用户泄露敏感数据的风险防范需求。

7) 不同加密技术路线比较

数据透明加密是否能够有效的关键，企业需要在关键数据的安全性、应用系统的功能可用性，和系统可维护性方面综合考虑，来确定适合企业需要的加密保护的技术方案。

下面的两张图表对前面提到的加密防护的关键技术能够提供的防护效果，安全性和部署复杂性进行了对比，供读者参考：

各种数据加密技术比较

风险	磁盘加密	文件系统加密	数据库透明加密		应用层加密
			TDE 表空间加密	列加密	
防止磁盘丢失引起数据泄露	✓	✓	✓	✓	✓
防止操作系统 root 账户和管理员账户访问数据	X	✓	✓	✓	✓
控制数据库管理员 DBA 访问数据	X	X	✓	✓	✓
抵抗 APT 持续攻击造成的数据泄露	X	✓	✓	✓	✓
提供细粒度的访问记录，法规遵从报告和风险分析	X	X	X	✓	✓
确保备份数据和数据快照加密存储	X	✓	✓	✓	✓
非结构化数据和文件的保护	✓	✓	X	X	✓
防止硬件和数据库厂商“偷窥”	X	X	✓	✓	✓

性能对比

文件系统的加解密是在文件系统层进行的，并且能够很好的利用文件缓存优化数据读写，所以一般情况下，性能要比数据库 TDE 加解密的性能好。

● 安全性对比

文件系统加解密的访问控制能力作用在文件系统层，能够控制对文件的访问权限；数据库 TDE



加解密的权限控制作用在数据库内部，能够控制数据库账户对数据为访问，所以在安全性上数据库 TDE 加密要比文件系统加密更加安全。

● 可靠性对比

数据加密的可靠性在一定程度上是与数据库系统的复杂性相关的；特别的由于 Oracle 数据库的复杂性高，其复杂的外围工具和程序，以及复杂的文件处理（ASM）机制，造成文件系统层的加解密处理要适配这些特性，这在一定程度上造成了文件系统加密机制对 Oracle 数据库的可靠性不足。因此从技术路线的选择上，对于复杂性不是很高的数据库系统加密，文件系统加密和数据库 TDE 加密的可靠性都可以满足，但对于像 Oracle 这种复杂性非常高的数据库系统，文件系统加密对数据库自身可靠性会带来一定的风险。以国外的某主流文件透明加密系统为例，在官方网站上可以看到该系统对 Oracle 数据库带来以下的影响：1）稳定性方面，可能造成系统宕机；2）兼容性方面，文件系统加密无法支持诸如 ASM 等典型的 Oracle 数据库高端特性；3）还会造成 Oracle 数据库性能下降。

Kernel panic at gsch_ftl_aio_write() function.

🔗 SOLUTION UNVERIFIED · Updated January 24 2017 at 3:24 PM · English ▾

Issue

- Kernel panic occurred during NULL pointer dereference inside gsch_ftl_aio_write() function.

```
[6428(bash)]: gsch_scan(525204,1,0) - interrupted & wait(1000)
[6428(bash)]: gsch_scan(525204,1,0) - interrupted & wait: done
BUG: unable to handle kernel NULL pointer dereference at 0000000000000018
IP: [ffffffffffa036a7b1>] gsch_ftl_aio_write+0x31/0x120 [gsch]
PGD 1d6215067 PUD 1d6216067 PID 0
Oops: 0000 [1] SMP
last sysfs file: /sys/devices/system/cpu/online
CPU 0
Modules linked in: autofs4 nfs lockd fscache auth_rpcgss nfs_acl sunrpc secvm2(P)(U) secfs2(P)(U) gsch(U) redirfs(U) nf_conntrack_ipv4
nf_defrag_ipv4 iptable_filter ip_tables vsock(U) dsa_filter(P)(U) ip6t_REJECT nf_conntrack_ipv6 nf_defrag_ipv6 xt_state nf_conntrack
ip6table_filter ip6_tables ipv6 ulpnet ppsdev microcode vmware_balloon parport_pc parport sg vmci(U) i2c_piix4 i2c_core shpchp ext4 jbd2
mbcache id_mod crc_t10dif sr_mod cdrom vmxnet3 vmw_pvscsi pata_acpi ata_generic ata_piix dm_mirror dm_region_hash dm_log dm_mod [last
unloaded: speedstep_lib]
```

* 3910356 (Tracking ID: 3908785)

SYMPTOM:

System panic observed because of null page address in writeback structure in case of kswapd process.

DESCRIPTION:

Secfs2/Encryptfs layers had used write VOP as a hook when Kswapd is triggered to free page. Ideally kswapd should call writepage() routine where writeback structure are correctly filled. When write VOP is called because of hook in secfs2/encryptfs, writeback structures are cleared, resulting in null page address.

RESOLUTION:

Code changes has been done to call VxFS kswapd routine only if valid page address is present.

图 76 文件系统加密导致 Oracle 数据库系统宕机异常

SYMPTOMS

ASM disks are changed from Character to Block devices using "vormetric" encryption and after that not able see ASM disks to add back

CHANGES

Changed Character to Block devices

crw-rw---- 1 oragrid asmadmin 18, 8 Sep 13 02:17 /dev/asmdevices/asm001

crw-rw---- 1 oragrid asmadmin 18, 12 Sep 13 02:17 /dev/asmdevices/asm003

crw-rw---- 1 oragrid asmadmin 18, 14 Sep 13 02:17 /dev/asmdevices/asm004

crw-rw---- 1 oragrid asmadmin 18, 23 Sep 13 02:17 /dev/asmdevices/asm005

to

brw-rw---- 1 oragrid asmadmin 18, 8 Sep 13 02:18 /dev/asmdevices/asmve001>>>>>>

brw-rw---- 1 oragrid asmadmin 18, 12 Sep 13 02:18 /dev/asmdevices/asmve003>>>>>>

brw-rw---- 1 oragrid asmadmin 18, 14 Sep 13 02:18 /dev/asmdevices/asmve004>>>>>>

brw-rw---- 1 oragrid asmadmin 18, 23 Sep 13 02:18 /dev/asmdevices/asmve005>>>>>>



☆ Block devices not supported for ASM (Doc ID 2310512.1)

In this Document

[Symptoms](#)

[Changes](#)

[Cause](#)

[Solution](#)

[References](#)

APPLIES TO:

Oracle Database - Enterprise Edition - Version 11.2.0.3 and later
Oracle Database Cloud Schema Service - Version N/A and later
Oracle Database Exadata Express Cloud Service - Version N/A and later
Oracle Database Exadata Cloud Machine - Version N/A and later
Oracle Cloud Infrastructure - Database Service - Version N/A and later
IBM AIX on POWER Systems (64-bit)
IBM AIX on POWER Systems (32-bit)

图 77 文件系统加密无法支持 Oracle 数据库的 ASM 高端特性

Copyright (c) 2019, Oracle. All rights reserved. Oracle Confidential.

☆ Performance issue when using Vormetric agent for encryption of database files (Doc ID 2098275.1)

In this Document

[Symptoms](#)

[Cause](#)

[Solution](#)

[References](#)

APPLIES TO:

Oracle Database - Enterprise Edition - Version 11.2.0.1 and later
Information in this document applies to any platform.
Specifically when using Vormetric agent for encryption of database files.

CAUSE

Introduction of a third-party encryption module that mediates interaction between the database instance and the database files. In this particular case, the customer configured Vormetric Encryption Expert Agent to protect Oracle Database files, including redo and archive logs. I/O speeds were impacted slightly for all write operations. Under normal database load, the impact did not introduce problems, but during heavy write activity, the increase in write time impacted the archiver's ability to copy archive logs to the encrypted file system enough that the archiver lagged the lgwr processes. This caused a cascade of wait events that resulted in significant performance issues.

SOLUTION

Modify Archive Log destination to a volume that does not use the Vormetric encryption agent.

Note:

Oracle does not support any 3rd party kernel/user modules, libraries or filters that inspect or change the behavior of the Oracle Database, including agents that modify the way in which the database interacts with the file system. When a database with these 3rd party modules, libraries or filters encounters performance, stability or corruption issues, Oracle will request that you reproduce the problem on systems that do not include such intrusive software.

REFERENCES

NOTE:1970146.1 Database Instability Caused by 3rd Party Kernel Modules

图 78 文件系统加密造成 Oracle 数据库性能下降

C.2.6.4 统一密码基础设施

1) 需求背景

基于密码技术的安全保护机制是保护数据安全、主动防御的有效措施。基于密码技术，可以采



从数据在各个应用、各个环节中流通、利用的角度看，在使用密码技术保障数据安全时，密码能力的互通、密码应用策略的一致至关重要，尤其是大规模、大范围、包含多种实体、跨越多个应用场景、跨越多个数据应用域进行数据流通、共享、开发和利用的场景中，如果基于密码的数据保护能力不互通、策略不一致，则可能造成数据流通和分享、交易过程中的数据不可用、数据无法保护、无法利用数据产生价值的情况。因此，应建设统一的密码基础设施，提供在数据的流通和利用中跨越不同数据应用域、不同时间段互通的密码能力、一致的密码策略，以支撑数据的全过程安全治理。

Figure 1 illustrates the architecture of the Data Security Service System, organized into three main layers:

- 数据安全服务层 (Data Security Service Layer):** This top layer includes services such as 数据安全存储 (Data Secure Storage), 数据安全流通 (Data Secure Circulation), 数据安全共享 (Data Secure Sharing), 数据安全交易 (Data Secure Transaction), and others (.....).
- 密码服务层 (Password Service Layer):** This middle layer includes 运营管理 (Operation Management), 运维支撑 (Maintenance Support), 安全审计 (Security Audit), and a central 服务 (Service) block. The service block contains 密码功能 (Password Function), 密钥管理 (Password Management), and 策略管理 (Policy Management).
- 基础设施层 (Infrastructure Layer):** The bottom layer includes 硬件基础设施 (Hardware Infrastructure), 软件基础设施 (Software Infrastructure), 网络基础设施 (Network Infrastructure), and 虚拟化基础设施 (Virtualization Infrastructure).

The system is supported by external frameworks: **国家法律法规 (National Laws and Regulations)** on the left and the **国家标准体系 (National Standard System)** on the right. Arrows indicate the flow of data and services between the layers and the external frameworks.

统一的密码基础设施的体系架构如上图所示。在符合国家法律法规、标准体系的框架下，通过统一建设的密码基础设施，提供了在数据流通、交易、共享和使用的全过程中，各个环节互通的密码基础支撑能力，通过密码服务层的组合、串接以及封装，以服务方式提供给各种数据运营，在数据安全服务层，借助这些密码能力，为数据安全存储、安全流通、安全共享、安全交易以及其它的数据管理、数据应用提供保障能力。

基础设施层通常包括硬件基础设施、软件基础设施、网络基础设施、虚拟化基础设施等。其中,



通常还包括以硬件模式出现的各种密码设备，以软件或硬件模式出现的各种密码系统例如数字签名系统等，还可能包括外部的各种密码服务例如第三方电子认证等等。

密码服务层通常包括以服务方式提供的密码功能、密钥管理、密码策略以及相配套的运营管理、运维支撑、安全审计等。密码功能可根据数据安全具体场景设置，可以动态增加或减少。通常包括：用以保证数据存储机密性的各种数据加密服务；用以保证数据传输安全的安全通信服务；用以保证数据完整性的加密算法服务；用以保证数据存取安全的身份鉴别、访问控制服务；用以确保数据在各个流通域的密码策略一致性；其它各种密码功能等。

数据安全层使用密码服务层提供的各种基本密码功能，按照各个数据应用域的安全要求，组织和实现数据在域内的安全以及跨域之间数据交换的安全保护能力，通常可包括：使用各种加密算法、完整性算法、身份鉴别机制保证数据存储安全；利用数据传输安全协议、数据加密算法保证数据的传输安全；利用电子认证、电子签名等机制提供数据确权能力、数据权属转移、数据安全转移能力。

3) 作用和价值

首先，统一密码基础设施能够保证在数据安全治理过程中密码技术应用一致性与有效性。

一致性指数据在广泛的网络空间中流通、共享、交易，需要进行各种确权、权属转移、数据开发利用的过程中，均可理解和接受上一环节的数据保护机制，确认其数据来源的真实性、验证数据的完整性，遵照数据来源一致的数据保护方式解密获得数据。一致性还包括数据在跨越较长时段的存储、流通、交易等环节中，保证数据保护机制是可用、安全的。

有效性指统一密码基础设施遵循国家法律法规、严格按照国家相关标准进行建设、运维、管理，提供可靠的服务，因此，能够保证用于数据保护的密码算法、密钥管理机制以及密码应用策略是有效的。

其次，统一密码机制能够为数据安全治理提供敏捷易用的安全支撑。

安全支撑的敏捷特性指密码应用中的部署和维护敏捷、应用敏捷以及扩展敏捷。部署和维护敏捷使用统一密码基础设施，数据应用域无须进行密码设施建设，而是可以直接使用基础设施提供的能力；应用敏捷指统一密码服务基础设施按照数据安全存储、安全流通、安全共享、安全交易等环节的需要，对基础密码算法、身份鉴别、密钥管理的技术细节进行组合串接，使各个数据应用域无缝接入密码锁提供的强大安全保护能力；扩展敏捷指当数据应用域需要密码功能特性和能力特性需要扩展时，包括添加新的密码算法、提升服务容量时，能够以最小代价完成。

C.2.7 数据安全运维

在数据安全治理的过程中，不但要预防外部攻击造成的敏感信息泄露，同时也要关注内部环境中的数据使用安全，尤其是数据在运维过程中的安全。

数据安全运维技术是通过人员、过程、技术和资源的整合为运维服务提供有效保障的一组技术的统称，以期解决运维过程中存在的特权身份冒用和滥用、访问权限管理混乱、运维人员权限划分粗粒度等问题。

数据安全运维技术的核心功能主要包括：特权访问管理（PAM, Privileged Access Management）、密码代填（密码桥）和运维访问防绕过等功能。





C.2.7.1 特权访问管理

特权访问安全是在整个企业中（不管是在本地、云端、终端设备上还是整个 DevOps 流程内）保护数据、基础架构和资产安全性的一个关键领域。一旦获得了特权账户和安全凭证，任何人都可以控制关键资源，访问大量敏感数据使安全保护系统形同虚设。

特权访问管理是指由人员，流程和技术组成的综合安全策略，用于控制，监控，保护和审计整个企业 IT 环境中所有人类和非人类特权身份和活动，以防止凭据盗用和特权滥用所造成的威胁。

特权访问管理的基础是最小权限原则，用户仅获得执行其工作职能所需的最小访问级别。最小权限原则被广泛认为是网络安全的最佳实践，并且是保护特权访问高价值数据和资产的基础。通过执行最小权限原则，组织可以减少攻击面并减轻恶意内部人员或外部网络攻击的风险，这些风险可能导致代价高昂的数据泄露。

1) 特权访问治理

提供管理特权分配、定期检查和认证特权访问、基于策略确保职责分离的能力。

2) 账号发现和启用

提供发现，识别和启用特权账号的功能，包括支持周期性，临时或连续扫描发现的功能。还包括自动发现目标服务和系统（包括虚拟机）以进一步发现其中包含的特权账号的能力。

3) 特权凭证管理

提供对组织和系统定义的特权账号凭证或机密（包括 SSH 密钥）进行管理和保护的核心能力。用户可通过生成，存储，轮换和检索等方式对凭证进行交互访问。还包括目标系统上服务和软件账号（即嵌入式账号）的轮换。这些功能至少需要能够通过 Web 控制台或 API 访问 PAM 工具。

4) 特权会话管理

为特权访问会话提供会话建立，管理，录制和回放，实时监视，基于协议的命令过滤以及会话分离。包括使用 PAM 工具管理从凭证签出到签入期间的交互式会话的功能，即使在正常情况下，凭证不会向用户公开。此功能还可能涉及一些限制，例如在登录目标系统时允许/拒绝某些类型的命令和功能。

5) 机密管理

此功能提供了针对非人类用例（例如机器，应用程序，服务，脚本，进程和 DevSecOps 流水线）的凭据（例如密码，OAuth 令牌，SSH 密钥等）进行访问管理的能力。它具有生成，存储，转换和向非人类实体提供凭据的能力（例如，通过 API）。还具有在不同的非人类实体之间交换机密和管理授权的信任代理能力。这些功能组合在一起支持动态环境的机密管理，并提供对 RPA 平台的支持。

6) 日志和报告

此功能提供了将所有单个事件（包括更改和操作）记录为 PAM 操作的功能。为监视和确定风险事件的根本原因，并识别未经授权的访问，将单个事件基于用户，时间，日期和位置，并通过逻辑顺序关联与其他事件一起处理。提供事件数据库的审计和报告的功能，支持预建报告和临时报告。事件数据还必须包含来自特权会话的信息。利用特权账号活动的分析（基线、风险评分和告警），实现检测和异常标记。目的是为了地更好地识别领先和滞后指标，从而识别特权访问异常，以触发对告警的自动应对策略。

7) 特权任务自动化



此功能提供针对多个系统上执行的特权操作，将与其相关的多步骤重复任务进行自动化的能力。使用常见 IT 系统和设备的可扩展的预配置特权操作库，在不同活动之间来回协调，并根据需要请求更多信息，同时通过对照策略和设置检查输入来提供防护。

8) 特权提升和委派

此功能基于主机强制执行应用程序的允许/拒绝/隔离等控制策略，并允许授权的命令或应用程序在提升的特权下运行。管理员将使用非特权账号登录并根据需要提升特权。任何需要额外特权的命令都必须通过这些工具，从而有效地防止管理员执行不安全的活动。这些功能必须在实际的操作系统（内核或进程级别）上执行。支持级别可能因平台（例如 Windows, UNIX / Linux 和 Mac）而异。特权提升和委派工具还可以提供文件完整性监视功能。

9) 相邻系统集成

此功能支持与相邻的安全和服务管理能力集成和交互，包括身份治理和管理（Identity Governance and Administration, IGA）、SSO、MFA 和企业目录，对灵活的连接器和集成框架的支持，常规 API 访问，与 IT 服务管理（IT Service Management, ITSM）系统的集成，安全信息和事件管理（Security Information and Event Management, SIEM）系统，以及漏洞管理。

10) 即时 PAM 方法

此功能提供按需特权访问，而无需共享拥有长期特权的账号。通常，这涉及到在有时间限制的基础上授予非特权账号适当的特权。实现此目的的常用方法可以是：使用特权提升和委派方法，使用临时和按需组成员身份、使用临时账号或安全令牌。此功能重点在于遵守最小权限原则，为 PAM 访问实现零站立特权（Zero Standing Privileges, ZSP）。

C.2.7.2 密码代填

运维工作中通常会存在多名运维人员共用一个数据库账号的情况，共用的人越多，风险也就越大。

密码代填（密码桥）技术运行机制为，预先建立运维账号和数据库账号的映射关系表，当运维账号通过数据库连接工具（如 SQL*Plus、PL/SQL Developer 等）访问数据库时，密码代填技术会在数据库连接过程中按照映射关系表对账号、密码进行替换，实现运维账号直接访问数据库的效果。同时数据库账号不对运维人员暴露，增强了账号的安全性，解决了账号共用问题。

C.2.7.3 运维访问防绕过

某些 C/S 架构的业务系统，存在大量客户端设备和动态 IP 的情况，无法通过客户端 IP 地址定位到客户端主机来实施控制。安装插件的方式工作量较大，极易出现遗漏，且存在绕过安全设备访问数据库的漏洞。安装的插件无法监测到数据库服务器本地的访问操作行为，同样存在管控的漏洞。

运维访问防绕过技术，通过数据库的 logon 触发器进行访问白名单控制。访问数据库的客户端设备，须先将其关键信息（IP 地址、用户、客户端主机名等）进行备案注册，按照注册内容对触发器进行调整，实现客户端与数据库建立连接时的有效控制。试图绕过安全设备的访问行为，也会被认定为非授信的异常访问。只有在授信终端上使用授信客户端运维工具、运维账号、并通过合规的访问通道，才可对数据库进行运维操作，真正实现全方位的访问控制，保障安全无死角。



C.2.7.4 相关技术

随着 PAM 的发展, Gartner 将 PAM 市场划分为特权账号和会话管理 (Privileged Account and Session Management, PASM)、特权提升和委派管理 (Privilege Elevation and Delegation Management, PEDM) 及机密管理三个细分领域, 以凸显 PAM 解决方案的不同。

1) 特权账号和会话管理 (PASM)

特权账号通过存储其凭据得到保护。然后, 将这些账号的访问权代理给各类用户, 服务和应用程序。特权会话管理 (PSM) 功能可建立具有可能的凭证注入和完整会话记录的会话。特权账号的密码和其他凭据得到有效管理, 例如以可定义的间隔或在发生特定事件时进行更改。PASM 解决方案还可以选择为 IT 人员和不需要 VPN 的第三方提供应用程序到应用程序密码管理 (AAPM) 和 / 或零安装远程特权访问功能。

PASM 通常通过密码库授予临时管理员访问权限。保管库授予用户访问所需服务器的权限, 然后在管理会话期间记录其所有活动, 以进行监视。会话完成后, 管理员访问权限将被撤销。

2) 特权提升和委派管理 (PEDM)

基于主机的代理在受管系统上授予登录用户的特定特权。PEDM 工具提供基于主机的命令控制 (过滤), 应用程序允许 / 拒绝 / 隔离控件和 / 或特权提升, 后者的形式允许以更高的特权级别运行特定命令。PEDM 工具必须在实际的操作系统 (内核或进程级别) 上执行。通过协议过滤进行的命令控制已明确排除在此定义之外, 因为控制点的可靠性较差。PEDM 工具还可以选择提供文件完整性监视功能。

PEDM 通常旨在提供细粒度的访问权限控制, 消除管理员账号, 只允许系统管理员使用常规用户账号进行操作。系统管理员仅被授予对他们需要管理的各个应用程序, 脚本或任务的管理员特权。因此, 组织可以更轻松地减少或消除其网络中具有任何管理员访问权限的账号数量, 从而减少攻击面并减少外部威胁或人为错误的风险。

3) 机密管理

软件和机器的凭据 (例如密码, OAuth 令牌, SSH 密钥等) 和机密通过 API 和 SDK 进行编程管理, 存储和检索。建立和代理信任关系是为了在不同的非人类实体 (例如机器, 容器, 应用程序, 服务, 脚本, 进程和 DevSecOps 流水线) 之间交换机密以及管理授权和相关功能。机密管理通常用于动态和敏捷的环境中, 例如 IaaS, PaaS 和容器管理平台, 确保跨工具栈, 平台和云环境的资源只能由经过身份验证和授权的实体访问。

4) 多因子身份认证

多因素身份认证 (Multi-factor authentication, MFA) 是一种访问控制的方法。用户需要通过两种或两种以上的认证机制之后, 才能得到授权, 访问资源。这种认证方式可以降低网络攻击的风险, 提高安全性。

MFA 的目标是创建分层防御, 并使未经授权的人员更难以访问目标, 例如物理位置, 计算设备, 网络或数据库等。如果一个因素被破坏, 攻击者在成功闯入目标之前至少还有一个障碍需要突破。过去, MFA 系统通常依赖于双因素身份验证 (Two-factor authentication, 2FA)。现今人们越来越多地使用 “多因素” 来描述需要多个身份凭证的身份验证方案。



身份认证因素是用于身份验证的一类凭据。对于 MFA，每个附加因素都旨在提高实际访问资源的实体与其所声明的对象匹配的准确度。认证因素包括知识因素（用户已知的东西），拥有因素（用户拥有的东西）和固有因素（用户自己的东西）、位置因素和时间因素。

- 知识因素

这种基于知识的身份验证（KBA）通常要求用户提供一个秘密问题的答案。

- 拥有因素

用户必须拥有特定的物品才能登录，例如安全令牌，密钥卡或电话的 SIM 卡。对于移动身份验证，智能手机通常会结合 OTP 应用程序提供占有因子。

- 固有因素

用户具有的可用于登录验证的任何生物学特征。生物特征认证方法的范围包括：视网膜扫描、虹膜扫描、指纹扫描、手部几何学识别、面部识别、耳垂几何学识别、语音识别。

- 位置因素

通常将用户的当前位置建议为进行身份验证的第四因素。例如通过智能手机的定位功能对登录位置进行合理的确认。

- 时间因素

当前时间有时也被视为身份认证的因素之一。根据工作时间表验证员工 ID，防止某些类型的用户账户劫持攻击。

C.2.8 数据库安全防护

C.2.8.1 技术概述

主动防御型的数据库安全防护模型以风险策略 (Strategy) 为核心，以风险识别 (Identification)、风险评估 (Assessment)、风险处置 (Response)、风险分析 (Analysis) 为手段。帮助用户快速构建主动式安全防护体系，减少因风险事件造成的经济损失，节省漏洞的修复本，降低投入到合规工作中人力成本。

C.2.8.2 功能架构

数据库安全防护系统是一款基于数据库协议分析与控制技术的增强级数据库安全防护系统，可主动、实时、全方位地保障数据库安全，使数据免受数据库漏洞、高危恶意操作及敏感数据泄露的威胁。面对来自外部的入侵行为，提供防 SQL 注入和数据库虚拟补丁技术。通过虚拟补丁，使数据库系统不依赖升级、打补丁，即可完成对主要漏洞的防护。结合针对数据库的安全访问控制规则，帮助用户应对来自内外部的数据安全威胁。



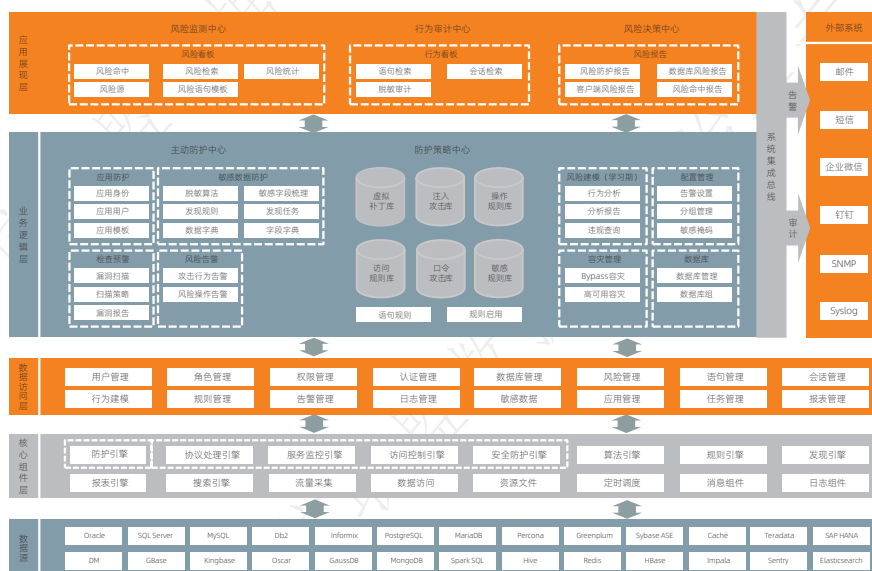


图 80 数据库安全防护架构图

数据库安全防护的主要功能如下：

1) 虚拟补丁，主动防护已知漏洞

虚拟补丁技术是系统通过协议解析技术，实时监控数据库漏洞攻击行为，防止利用已知漏洞对数据库发起的攻击，保障数据库安全与应用稳定。在数据库外围创建安全层，无需根据数据库厂商提供的补丁做数据库修复，也不需要停止服务和回归测试，使漏洞防御的实施更为轻量化和及时。

2) 防护规则，全面感知防护风险

系统支持多种类型的防护规则，针对不同风险提供相应的防护规则。

- 注入攻击：提供 SQL 注入防护和 XSS 攻击防护规则，防止利用应用程序漏洞的注入攻击。
- 操作规则：防止高危语句操作、批量数据篡改和大规模数据泄露等风险行为。
- 访问规则：针对应用端信息、客户端信息、时间等条件限定非法账户登录行为。
- 口令攻击：限定不同访问源和数据库账户的失败登录频次。
- 语句规则：针对 SQL 语句模板设置访问控制规则。
- 例外规则：可在每条风险防控规则中设置合规的例外访问操作，不拦截不审计。
- 信任规则：针对信任操作行为及语句建立白名单不拦截，避免影响正常的业务行为。
- 脱敏规则：配置脱敏规则及脱敏算法防止敏感数据泄露。

3) 敏感数据，防止敏感信息外泄

关联 schema、表、字段等对象形成敏感数据组，基于“与或”关系并结合其他防护项建立敏感数据防护规则。限定敏感数据的访问时间、来源 IP 地址和访问账户，针对高危操作进行语句拦截或会话阻断。

4) 应用关联，准确定位访问信息

基于应用会话捕获应用账户、应用登录 IP 及 URL 等关键防护信息，添加到防护规则进行风险行为管控，建立并完善应用关联防护体系。将业务信息和数据库操作行为进行关联，有效追溯应用系统的原始访问者和请求信息，对于应用端的数据库访问、操作行为，实施精确安全管控和防护。

5) 防漏扫，抵御漏扫通过测评



对数据库漏洞扫描进行防御，防止漏扫工具扫描出漏洞，协助用户通过测评。支持多种漏洞扫描工具及 Oracle、Db2、MySQL 等数据库类型。

6) 安全审计，关联分析追责溯源

从风险、语句和会话多个角度对风险行为进行审计。用户可基于风险模型进行风险行为的关联查询和深度挖掘，实现风险分析和问题追溯。

7) 风险告警，实时告警及时处置

系统提供数据库风险告警能力，基于灵活的防护规则，对外部发起的数据库漏洞攻击、恶意的 SQL 注入行为、非法的业务登录、高危的 SQL 操作和过量的数据下载，提供实时的风险告警。

C.2.8.3 关键技术

1) 立体防护，阻断拦截审计告警

会话阻断：基于数据库安全防护规则，对高危会话进行强制阻断，禁止该会话的所有操作行为。

语句拦截：比会话阻断更为细致的限制，只对风险语句进行拦截，不影响会话的整体操作。

2) 精确防护，协议解析语法分析

准确的协议解析能力，是数据库安全防护的基础。由于串联防护的特殊性，任何误报、漏报都可能造成不可估量的损失。传统解析手段采用正则匹配和较为简单的协议解析技术来分析 SQL 语句，解析结果不准确，数据库通讯协议解析内容需要覆盖 SQL 语句、参数化语句句柄、SQL 参数、应答结果信息、结果集结构信息及结果集数据。

3) 双向防护，黑白名单周密布防

通过学习模式及 SQL 语法分析构建动态模型，形成 SQL 黑白名单，放行 SQL 白名单语句，阻断 SQL 黑名单语句。

4) 高性能防护，业务性能零影响

支持每秒数万个 SQL 的请求速率，响应速度快，无感串接，可实现访问速度无明显变化，不影响业务。

5) 高可用防护，保障业务连续性

Bypass 容灾实现异常断电或系统宕机时自动进行网桥的强制连接。确保数据库通讯网络的畅通及应用系统的正常运行。

双机 HA 容灾支持双机主备部署，实现网络高可用，通过网络协议和 HA 跳线进行主备间探测与切换，采用会话同步、策略同步机制，保障双机间的一致性。

系统后台预置三通机制，当解析进程出现异常时，立即触发三通机制保障用户与服务器间的正常通信。并重新创建解析进程恢复对后续访问操作的安全防护。

C.2.9 用户和实体行为分析 (UEBA)

C.2.9.1 技术概述

用户和实体行为分析 (UEBA, User and Entity Behavior Analytics) 是指以用户为中心分析用户

行为，关联端点、网络和应用行为，并将这些数据与威胁情报相关联，通过自动学习正常行为来发现异常行为。UEBA 依赖高级安全分析方法实现免配置优化、自动发现异常行为，将大大提高告警准确性。UEBA 常用来检测恶意的内部人员和外部攻击者对企业信息系统的渗透和数据窃取。

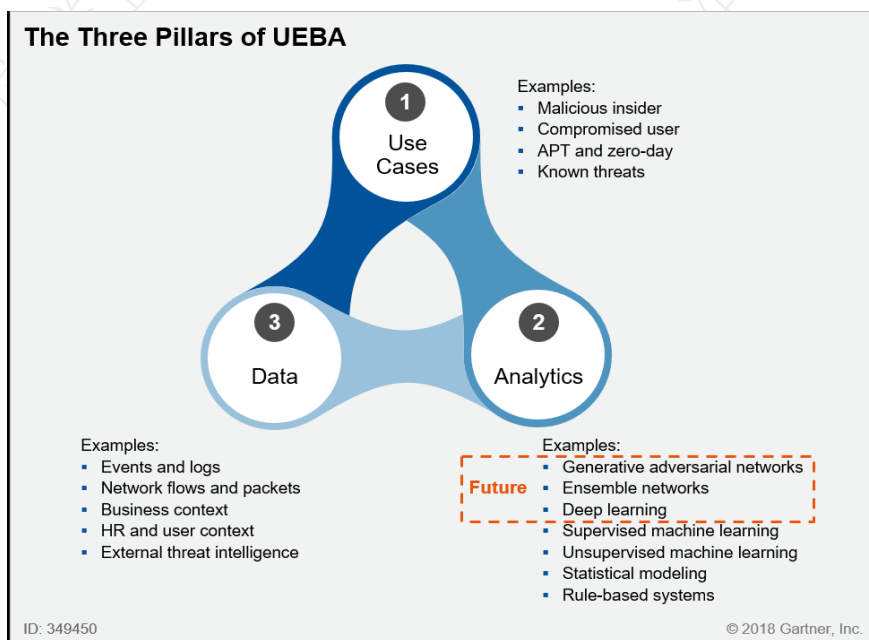


图 81 用户和实体行为分析

如上图所示，Gartner 从用例（应用场景）、数据、分析三个维度上定义了 UEBA 方案。

C.2.9.2 关键技术

1) 数据收集

收集的数据包括但不限于，安全设备产生的事件和日志、网络流量包、外部威胁情报、企业的基础信息等等。采集数据的质量高低会直接影响分析结果的准确性。这是很简单的道理，输入了大量垃圾数据，即使模型再好，用户案例再丰富，输出的结果也是垃圾，甚至还带有误导性。另外，采集数据不足，不足以支撑分析的模型，也很难产生满意的结果。当然，数据也不是越多越好，UEBA 的数据更多会依赖于你期望分析的场景而进行定向选取，更精准的去发现异常痕迹，通常这些数据一般也会与 SOC/SIEM 的数据重叠。

加强数据收集的思路和方法主要有：

● 从用户和实体活动中提取事件数据

可以从数据源本地获取，比如直接或通过现有存储结构中（中央日志管理 CLM，安全信息与事件管理 SIEM），而不应该主要依赖网络数据包作为主要数据源，或者只通过自己部署的代理程序来收集数据。

● 使用更多外部来源的数据

这种数据收集方法应支持结构化和非结构化等多种数据类型，并考虑商业途径、人力资源数据库以及外部威胁情报等数据来源。



2) 数据分析和建模

当前的 UEBA 分析方法和模型，大部分还是基于“规则”建立的，依赖于历史事件、专家经验以及人为设定的阈值，同时也避免不了针对使用环境的不同而调整和优化参数的工作。虽然目前机器学习越来越成熟，而且本身 UEBA 也非常符合统计学的使用范畴，但是对其使用者和调试者而言，仍然需要丰富的数据科学和专业分析经验，具有比较高的门槛，而且对数据量和数据的时间跨度也有一定要求，一般不足百万条的数据量，很难看得出效果。现阶段基于规则的分析方法可能更具优势，对于一般使用者而言，更难利用机器学习或者 AI 再添加自定义或额外的场景。

未来，UEBA 建模技术必然会向机器学习和人工智能方向发展，可期的技术目标有：

- 通过机器学习和统计模型，结合现有的访问策略形成用户行为的基本模型和画像；
- 避免只关注单体的访问行为，而是要结合所有用户及其操作对象的访问，整体进行异常分析；
- 数据建模将向深度机器学习及更智能的分析方向发展。

3) 场景和功能

UEBA 的应用场景非常丰富，但侧重点会更多地集中在内部安全领域，常见的如账号的异常登录（时间、地点、频率），服务器上用户的违规操作（上传、下载、执行敏感指令），终端上的异常动作（发起内网扫描、异常的 DNS 查询请求、大流量的文件传输）等等。独立的 UEBA 方案主要针对以下 5 个应用场景。

● 针对恶意内部人员

监控内部人员和授权的外包人员，针对异常、危险的操作进行监控和识别，而非针对外部黑客攻击。由于内部人员的恶意行为更难以发现，因此 UEBA 产品很难从日志文件中分析出危险行为，还需要更多获取非结构化数据，比如电子邮件、绩效信息以及社交媒体信息等，用以分析内部或外包授权人员的行为，结合人员背景和行为进行分析监测。

● 内部隐患及外部攻击威胁

企业或组织需要能够快速检测攻击行为，比如 APT 攻击及其他未知安全威胁，但是 0day 漏洞的攻击难以发现，而且经常伪装成合法用户进行攻击，如果只使用模型匹配和阈值控制很难监测。UEBA 提供对于这部分攻击的监测，由于该类攻击本身难以识别，UEBA 需要更好的降低误报率，通常需要与 SIEM 工具整合实施。

● 数据泄露

UEBA 在数据泄露场景下的实现需要结合 DLP 产品，基于 DLP 本身对异常行为的检测，结合网络流量和网络节点数据进行更深入的分析，提高对数据泄露威胁的告警质量和优先级。

● 身份和特权访问管理

UEBA 产品通过已经建立的权限访问策略监控用户行为，识别特权用户和应用用户的越权访问行为，也有企业使用 UEBA 清理权限分配过大的用户，实现身份和特权访问管理。

● 威胁优先级

基于安全基线的建立和威胁模型的构建，匹配数据资产和人员角色的访问权限，UEBA 能够指导企业目前存在的安全威胁中哪些级别较高，需要优先处理。



C.2.10 数据泄密防护 (DLP)

C.2.10.1 技术概述

随着信息技术的飞速发展，计算机和互联网已成为日常办公、通信交流和协作互动的必备工具和途径。但是，信息系统在提高人们工作效率的同时，也对数据的存储安全及传输安全提出了更高要求。当前各企事业单位对内网的安全解决方案，还停留在采用防火墙、入侵检测、存储加密系统等保护网络、限制数据访问或全网数据加密的被动、枷锁式防护手段上。

纵观近年来国内外数据泄露事件，互联网接入单位由于内部重要机密通过网络途径泄露而造成重大损失的事件中，只有 1% 是被黑客窃取造成的，而 99% 都是由于内部员工有意或无意的一些泄密行为导致的。

数据泄密防护 (DLP, Data Leakage Prevention) 系统又称为数据防泄露系统，或 DLP 系统。DLP 系统采用内容分析引擎，利用关键字、正则表达式、文件指纹、自然语言处理等数据识别技术，对外发数据进行解析与扫描，实时识别、监控、保护企业或组织的敏感数据；对即将发生、正在发生的泄露敏感数据行为，按照预置策略及时阻断并告警，防止敏感数据传输到外部，有效避免数据泄露带来的安全风险，实现对外发敏感数据的可知、可见、可控的一体化解决方案。

DLP 系统的核心理念是立足于用户内部网络，建立起以数据内容为核心的智能识别策略和集中安全管理为显著特征的全方位数据内容智能识别体系，通过深度内容分析和事务安全关联分析来识别、监视和保护静止、移动和使用中的数据，并能联动其它传统安全产品形成整体数据安全解决方案，满足客户的全方位安全需求。

从防护主体对象的角度划分，DLP 系统可以分为网络 DLP、终端 DLP、邮件 DLP、网页 DLP。

C.2.10.2 DLP 系统框架

DLP 系统是结合了识别、监测、管控三类技术而形成的数据安全系统，其所能实现的功能便是将三类技术要素进行任意组合，对不同的数据内容通过不同的通道（与参与者）进行传输的行为进行区别化的管控。

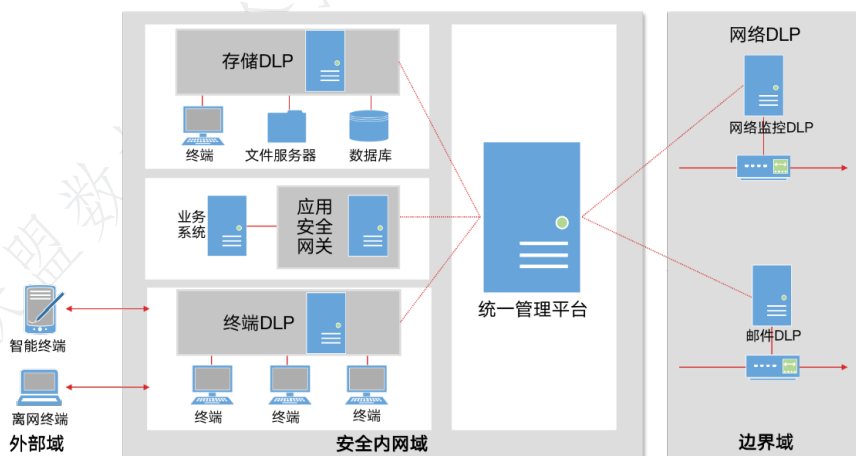


图 82 DLP 系统框架示意图

大部分国内厂商 DLP 系统的各项功能模块都采用了统一的内容识别技术引擎，功能模块的划分往往是根据不同的场景、不同的通道监测技术与管控技术来进行划分的：

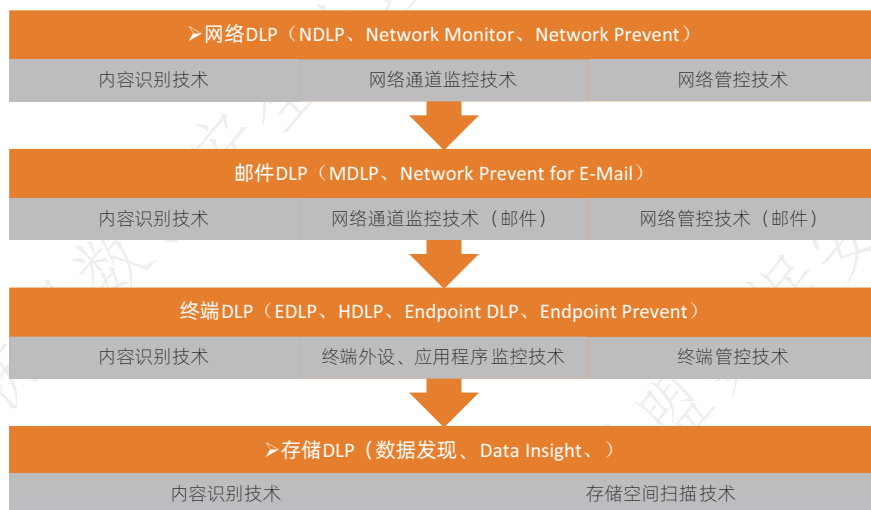


图 83 DLP 系统的功能模块

C.2.10.3 关键技术

1) 常规 DLP 检测技术

● 基于关键字的检测

关键字检测技术是将包含一个或多个关键字或关键字短语的列表中的每个关键字与目标文件进行比较的内容识别技术，关键字支持普通关键字和近邻关键字匹配。

近邻关键字是指如果定义了关键字之间的字间距。如果出现位于该范围内的字，系统也能够检测到相关内容。

● 基于正则表达式的检测

正则表达式是对字符串（包括普通字符（例如，a 到 z 之间的字母）和特殊字符（称为“元字符”）操作的一种逻辑公式，就是用事先定义好的一些特定字符、及这些特定字符的组合，组成一个“规则字符串”，这个“规则字符串”用来表达对字符串的一种过滤逻辑。正则表达式是一种文本模式，模式描述在搜索文本时要匹配的一个或多个字符串，数据防泄露系统支持通过配置正则表达式（如手机号、省份证号等）对外发内容进行检测。

● 基于字典的检测

字典算法是通过字典项来配置敏感信息，每个字典项实际上是一组关键字的集合，只是每个关键字的权重是不一样的，可自定义配置。

● 基于脚本的检测

脚本匹配是使用 Luna 语言编写的脚本，可更精确匹配复杂内容。

● 基于数据标识符的检测

使用数据标识符，可快速实现精确、简短格式的数据匹配。

数据标识符是将模式匹配与数据验证器相结合来进行内容检测的算法。模式类似于正则表达式，但更加有效，因为它们已经过优化，可精确匹配数据。验证器用于进行准确性检查，这种检查着重



于检测范围并确保遵从性。

2) 内容指纹匹配技术

● 索引内容指纹匹配技术

索引内容指纹又称索引数据匹配 (IDM)，是一种对于数据进行模糊匹配的算法技术。该技术主要用于检测各种非结构化储存的文档与样本文档的相似度。

通过对非结构化存储的数据采取分段指纹提取措施，可以实现对待检测文档内容的模糊匹配。文档经过数据预处理之后，获得与内容有关的标准化数据（统一的字符编码、内容标准等），之后采用窗口切割方式进行指纹计算。通过这样的方法可以获得一串压缩率 10% 左右的内容指纹。

对于待检测文档，使用上述的计算方式得到其指纹信息，然后根据自定义的相似度阈值对待检测文档的内容进行模糊匹配，如果达到设置的相似度阈值则认定该文档中的内容满足匹配条件，系统将根据设置采取后续动作。

● 精确内容指纹匹配技术

精确内容指纹又称精确数据匹配 (EDM)，是一种对于结构化数据（各类表格）进行精确匹配的算法技术。该技术主要用于匹配待检测数据是否命中结构化数据中某一行的算法技术。

通过对表格数据的行列进行预分析，计算出最能代表数据复杂性的列进行索引编制，并扩展到每行，以达到快速查询并匹配的效果。

精确数据匹配技术具备以下特性：

- 无序性：对目标中的敏感信息顺序不作要求；
- 精确性：数据必须唯一匹配原始数据的某行；
- 计算效率：具备高速匹配效率。

3) 计算机视觉技术

计算机视觉，是指对图像进行识别判断，并理解其中的含义。围绕数据内容，具备以下技术：

- 光学字符识别 (OCR)：对图片中的文字进行识别辨认；
- 图章识别：对图片中的公章进行识别辨认；
- 物体识别：通过大量样本训练后，对图片中的物体进行识别辨认；
- 标志识别：通过大量样本训练后，对图片中的标志或 Logo 进行识别辨认；
- 人脸识别：通过大量样本训练后，对图片中的人体脸部特征进行辨认。

由于光学字符识别、物体识别、标志识别和人脸识别有很多公开论文进行描述，此处不一一赘述。其中，图章识别技术是经过大量研发投入所具备的独特技术，其计算定位通过以下步骤完成：

- 采用基于机器学习 + 神经网络的标志识别，模糊匹配出圆形图章；
- 文字拉直，对图章区域内的文字进行图片处理，使文字能够水平排版；
- OCR 文字识别，对水平排版的文字进行识别。

4) 高级语义分析检测技术

高级语义检测技术是通过文本基本处理、高层语义分析进行语义分析。文本基本处理是语义分析的基础。主要包括中文分词、语言模型、特征权重、核心词、关键词提取等一系列的方法，其中中文分词是中文语义分析的基石。高层语义分析主要包括词向量和句向量分析、主题模型、深度神





经网络以及文本分类等技术。

C.2.10.4 核心功能

1) 敏感数据识别

采集动态扫描与静态分析等方式梳理企业数据资产，形成数据资产清单。按照数据的价值、敏感程度、影响主客体和分发范围等维度，对数据资产进行分类分级，根据数据资产密级高低定义安全管控定义。

采用关键字、正则表达式等基础规则匹配方法以及基于自然语言处理的智能识别方法识别网络、终端、邮件等主体中的敏感数据，根据管控策略执行监控、阻断、审计等管控行为。

2) 特殊人员标识

针对组织机构特殊人员管控要求，数据防泄露系统可以按照组织架构、用户账号、姓名等设置黑名单、放行白名单、审计白名单。

● 管控黑名单

用于标记组织架构中，如待离职、安全风险值较高人员、被劫持账号等。凡是黑名单成员的外发行为，一律阻断，不允许外发数据。

● 放行白名单

用于标记组织架构中本身信息不被审计留存的人员，如纪检、监察等特殊成员。凡是放行白名单成员的外发行为，数据防泄露系统采取放行审计留存措施。

● 审计白名单

用于标记组织架构中因职业、工作本身涉及大量敏感信息的人员，如董事会秘书等。凡是审计白名单成员的外发行为，数据防泄露采取一律审计后放行措施。

3) 管控策略配置

用户结合企业的管理制度、组织架构、特殊要求，设置不同的匹配规则和匹配策略根据敏感数据管理的制度和手册，整理出违规行为处置流程、风险策略和事件策略，预先输入到数据防泄露中，提高安全事件处置效率。

数据防泄露系统内置常见数据基础识别规则，如：身份证号、身份证地址、银行卡号、税号等，用于采取关键字、正则表达式等方法匹配的内容识别。高级的数据防泄露系统具备基于自然语言的敏感数据识别的能力，可以组合关键字、正则表达式、人工智能模型等多种数据识别方式，大幅提高了系统敏感数据识别能力。

4) 监控保护审计

数据防泄露采用内容分析引擎，实时监听网络传输数据，分析网络信息包的副本，检测传输数据是否违反策略，自动记录敏感数据的使用情况。根据预先制定的安全策略，对敏感数据的扩散范围进行审计，对于审计出来的风险事件自动触发异常报警。

5) 违规行为阻断

数据防泄露系统发现违规外发等行为，自动根据安全管控策略，阻断违规行为，防止敏感信息



泄露，触发审计与告警。对于需要审批的违规行为，系统对其阻断后将自动进入审批流程，并且系统根据审批的结果来判断是否放行或者始终阻断。

6) 安全态势评估

数据防泄露系统，根据数据安全风险评估模型，实时汇总各项指标数据，实时计算指标数值，根据模型计算规则，综合各项指标，实时评估风险范围和风险级别，及时掌握数据安全风险态势。根据风评报告模板，基于风险态势各项指标的总体数据与局部数据，自动生成风评报告，提升用户数据安全管控效率。

C.2.10.5 应用实践

1) 终端安全管理系统

对入网后的终端设备还需进行安全加固，建立统一的安全基线，确保接入网络的终端设备是安全、可信的终端。主要通过补丁更新、防病毒软件安装、系统配置等技术，一方面实现终端标准化管控，一方面实现对终端行为的审计（包括敏感数据识别、敏感数据分类、数据分级阶段、敏感数据分布、敏感数据外发、敏感数据追踪、终端水印管理、屏幕录像审计、终端离线管控、终端外发审批、泄密追溯等功能）。

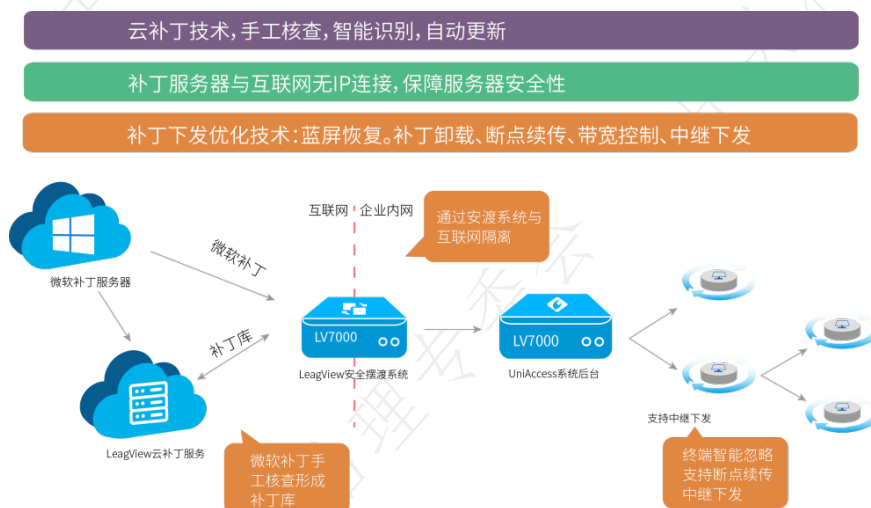


图 84 终端安全管理系统

2) 邮件安全网关

通过邮件外发数据是常见的泄密手段，所有的泄密事件中很多都是出自这个环节上，建议使用邮件防泄密网关来对邮件外发内容进行扫描，发现违规外发情况，可以根据策略设定，向管理人员发送告警信息（以弹出消息框或者邮件形式）。

首先，部署邮件防泄密网关，使得所有邮件外发都会经过邮件防泄密网关。同时在统一管理平台上对邮件防泄密网关下发数据防护策略。

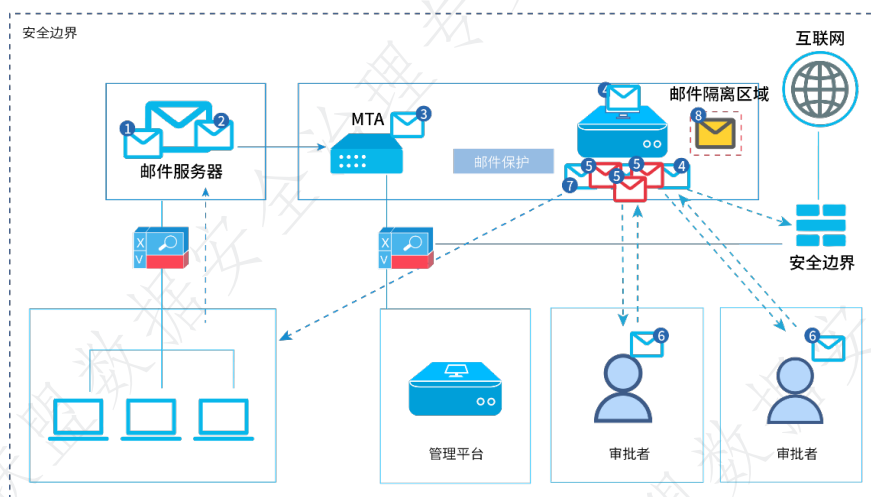


图 85 邮件安全网关

3) 文档发布 / 数据交换平台

数据只有在不断的交换、传输的过程中，才能体现自身的价值；因此数据流转环节一定要“疏、通结合”，才能实现真正的安全可控。达到安全控制的同时，要给用户提供更加高效的数据交换通道：文档发布系统、数据交换系统。

同时，文档发布系统还可以与企业内部的 OA 系统进行第三方对接，提供 Web Service 接口进行联动，在终端用户访问 OA 系统时（静态页面、动态页面、办公文档）附加相应的水印信息。对于极端情况下泄露后的图片，可有相应技术手段对其泄露源头进行溯源，以便事后取证。

4) 大数据分析平台

数据安全态势感知一般分为四个层次，数据采集层、数据处理层、数据分析层、风险感知层。可实现数据采集、数据处理、数据分析、风险呈现、数据流转、数据外发、数据追溯等功能。

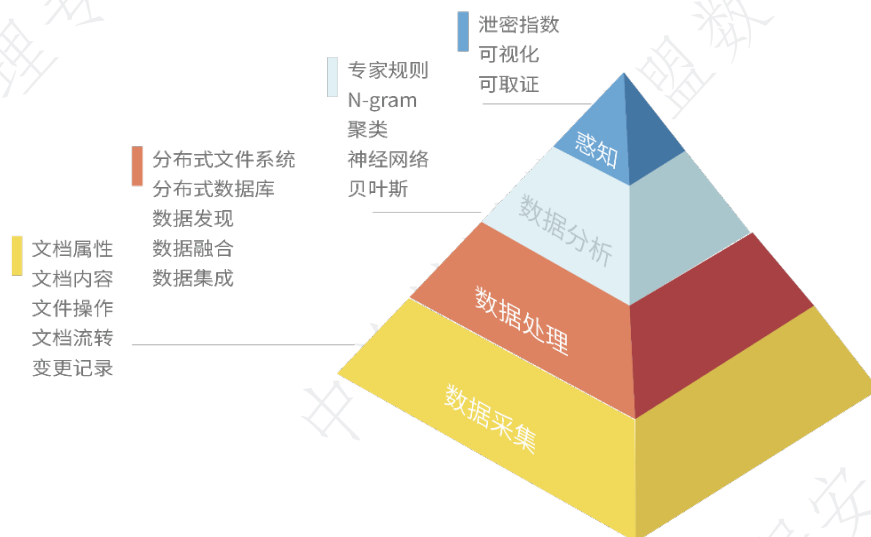


图 86 大数据分析平台示意图

C.2.11 数据备份保护

C.2.11.1 技术概述

现如今，云计算、物联网、大数据、人工智能等新兴技术推动各类组织数字化转型，技术革新带来诸多收益的同时，也加速着信息基础设施的迭代更新。现代化的数据中心不再是传统单纯的 X86 物理或虚拟化架构，而是进入公有云与私有云混合、传统应用与云原生应用混合、X86 环境与国产化混合、稳态与敏态混合的全新 IT 时代。

全新的 IT 环境可能包含各种物理环境、云环境、容器、大数据环境等，在动态、复杂的混合环境下，灾备技术作为生产数据与应用安全的重要保障，已广泛应用于各行业用户。而广大用户的生产数据保护的对象已从文件、数据库、操作系统扩展到虚拟机、云数据、容器等，数据与业务保护的诉求也从传统的数据备份和数据恢复扩展到应急容灾、副本管理、数据归档等，如下图 87 所示。



图 87 新 IT 环境下的多样化灾备需求

为保障生产数据和应用的安全，一般会采用多种灾备技术，如备份、容灾、归档等，传统灾备模式通常采用多套灾备系统堆叠，每套灾备系统独立实现备份或容灾或归档功能，无法满足复杂环境的需求，因此创新的灾备架构既要能够解决复杂环境带来的脆弱性问题，同时具备多种灾备技术的统一管理，满足动态、敏捷和全面的要求。

C.2.11.2 功能架构

随着灾备技术的演进与发展，数据与业务保护的架构可充分发挥云计算弹性、池化、资源综合利用的特点，利用分层分级数据捕获、系统高可用、分布式等相关灾备技术，建立统一架构、覆盖不同场景和不同灾备诉求的灾备平台，实现灾备资源弹性扩展和灾备服务能力的自由编排。

平台架构如下图所示：

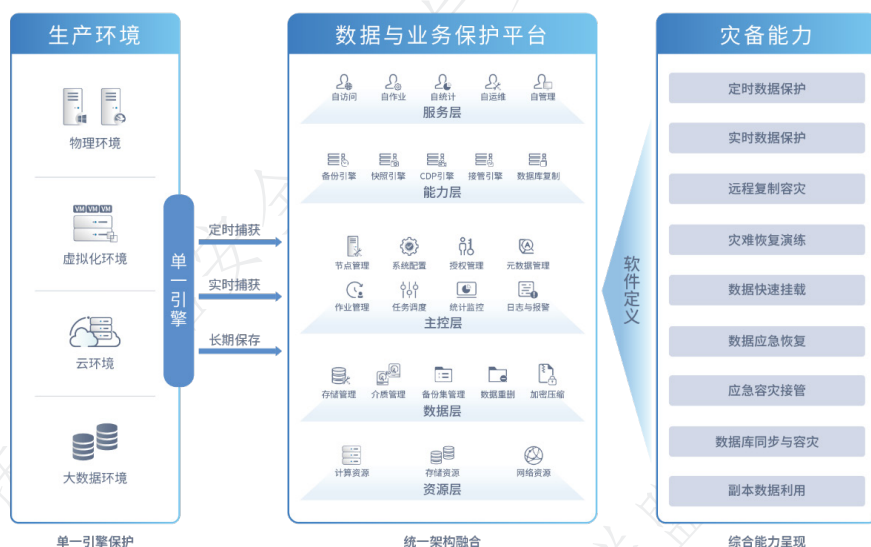


图 88 新一代数据与业务保护平台架构图

C.2.11.3 关键技术

相对传统灾备系统，其关键技术包括：

1) 单一引擎全面保护

为规避传统灾备系统针对不同 IT 环境、数据类型安装多个灾备引擎存在的模块化管理、资源占用、兼容性等一系列问题，平台采用单一灾备引擎获取数据，全面覆盖云、大数据、虚拟化、X86、小型机、国产化等不同 IT 环境和不同灾备对象的统一灾备或分级灾备保护的需求。

2) 灾备能力融合

为确保灾备架构的灵活性、敏捷性，实现灾备能力可编排，灾备平台由统一资源层、数据层、控制层和服务层等构成，融合数据备份、数据快照、实时捕获、数据库复制、长期保存、容灾接管等灾备技术，通过软件授权控制方式，可灵活组合不同层级的灾备保护能力，满足业务工作负荷需要。

3) 数据多元化利用

为最大化利用数据，降低副本数据存储空间，采用副本数据管理技术实现了从副本数据采集、安全传输、安全存储、数据索引、数据呈现到各种数据利用的统一管理，提高了数据的价值。

灾备平台可面向用户生产环境提供定时备份、实时备份、应急接管等功能，覆盖包括本异地灾备、云上数据保护、云容灾、海量文件备份、防勒索病毒等场景，还可以进一步通过副本数据管理技术，满足用户开发测试、模拟演练等场景需求。

C.2.11.4 应用场景

灾备平台可运用在多个应用场景，具体如下：

1) 本异地灾备场景

对于不同生产环境的用户，灾备平台可提供备份、数据复制、应用接管、异地数据同步等功能，提供本异地的数据级和应用级灾备，满足等保合规、数据和业务保护需求。

2) 云上数据保护场景



为保障云上数据安全，面向云虚拟主机、数据库数据、文件等，提供云上数据备份与恢复、数据可用性校验，可以将云上数据备份至本地或云端，实现云上数据保护。

3) 云容灾场景

对于业务上云的用户，通过构建云下或异构云间的容灾，采用实时数据复制、应急接管等，当云上业务中断时，可在本地或异构云上实现业务连续。

4) 海量文件备份场景

面对用户海量非结构化数据的增长，基于灾备平台分布式可扩展架构，实现海量备份任务的并发、数据的负载均衡，提升海量数据备份效率。

5) 防勒索病毒场景

灾备平台可提供 CDP 持续数据保护，当用户生产主机感染勒索病毒后，可在灾备平台选择任意时间点恢复数据，可有效防止勒索病毒攻击导致的数据丢失。



D.2020 年以来重大数据安全事件分析

D.1 勒索软件持续流行，对不同国家不同行业造成的损失案件不计其数

勒索软件（RAN）是一种恶意软件，通过感染用户的操作系统，采用加密用户的数据或拒绝用户访问设备等方式，使用户数据资产或计算资源无法正常使用，以此向用户勒索钱财换取解密密钥或恢复对设备的访问，赎金形式包括真实货币或虚拟货币。

2021 年 3 月，计算机巨头宏碁（Acer）受到勒索软件 REvil 攻击，勒索赎金高达 5000 万美元，创下勒索软件赎金的新纪录。宏碁是一家成立于台湾省的电脑制造商，主要生产笔记本电脑、pc 主机、电脑显示器、平板电脑等其他数码产品。宏碁是全球第六大计算机制造巨头，在欧洲、非洲的市场占有率常年保持第一排名，而在美国市场也同样占据可观的份额。此次勒索攻击始作俑者 REvil 团队在其网站上公开宣布 Acer 被攻破，并放出了被盗文件的相关图片证据，这些图片包括了公司财务表格、银行账户余额信息、银行往来记录等机密数据。据国外某技术团队表述，此次攻击可能是因为攻击者攻击了 Acer 的一台 Exchange 服务器，才导致数据遭到挟持。

2021 年 5 月，美国最大的成品油管道运营商 Colonial Pipeline 遭受到勒索软件攻击，被迫关闭其美国东部沿海各州供油的关键燃油网络。这极大影响了美国东海岸燃油等能源供应，美国政府宣布进入国家紧急状态。Colonial 是美国炼制油品输油管线的龙头业者，每天从美国墨西哥湾沿岸的炼油厂，运输 250 万桶的汽油、柴油、飞机燃油，至人口稠密的美东地区，如纽约、华盛顿、亚特兰大等地，管线总长逾 5500 英里。美国联邦调查局（FBI）发布声明，确认“黑暗面”勒索软件对 Colonial Pipeline 网络遭到袭击负有责任。

2021 年 5 月，俄语黑客组织 REvil 攻击了我国国内某地产百强企业，并对该地产公司的重要文件进行了窃取及加密，导致内部大量系统被攻陷，无法展开正常工作，大约 3TB 以上的数据，并索要 400 万美元作为赎金。其中包含有关雇员的 25,000 条数据，身份证、护照和其他机密文件，同时还有公司的所有财务报表，所有施工计划等等。该黑客组织为了证明数据被盗的真实性，公布了几张雇员的个人数据以及照片作为证据，并上传至博客。

2021 年 8 月，全球 IT 咨询巨头财富 500 强公司埃森哲遭遇了 LockBit 勒索软件组织的勒索软件网络攻击。一个名为 LockBit 2.0 的勒索软件组织声称从埃森哲窃取了 6TB 的数据，并要求支付 5000 万美元的赎金。埃森哲在受到攻击后发布了电子邮件声明，称其在事件发生后立即控制并隔离了受影响的服务器，并从备份中完全恢复了受影响的系统。然而，在 10 月份提交给 SEC 的文件中，埃森哲披露了一些客户系统遭到破坏，攻击者窃取和泄露了私密的公司数据。埃森哲作为全球知名的咨询公司，同时也是网络保险提供商，此次的攻击不仅对其数据和系统造成了破坏，也对其网络保险业务的开展带来了负面影响。

2021 年 12 月，澳大利亚州政府披露，由于外部薪资软件提供商 Frontier Software 的系统遭到勒索软件攻击，隶属南澳大利亚州政府的 8 万名员工的敏感个人信息遭到泄露，包括姓名、出生日期、税号、家庭住址、银行账户明细、发薪期、报酬等内容。此事件唯一没有影响到的是教育系统，因为教育系统没有使用 Frontier Software 提供的软件。



如今的勒索软件团伙已经学会了瞄准组织的关键业务系统，服务中断的每一秒都是巨大的损失，受害者要想尽快恢复运营，就不得不向勒索软件团伙妥协支付赎金。而在很多已支付赎金的情况下，勒索仍在继续，包括被再次勒索、数据仍被暴露在暗网上或无法检索他们的数据。

勒索病毒以防为主，用户应注意日常防范措施，针对重要的文档和数据，一定要定期做好数据备份；提高安全运维人员的职业素养，定期进行木马病毒查杀；提高人员安全意识，不点击来源不明的邮件、网站，不安装来源不明的软件，不插拔来历不明的存储介质等。而一旦中招，最佳方法不是屈服于勒索软件的要求，而是从备份中恢复系统和数据，并向执法和数据保护机构发出事件警报。其余的挣扎都是徒劳的，你的无能为力只会让勒索软件参与者变得越来越多，并且为他们提供勒索动力。

D.2 网络攻击，热点事件背后看不见硝烟的战场

近年来，黑客组织出于政治、名誉甚至是利益的目的，纷纷精心策划各种形式的网络入侵计划，多数攻击活动具有极强的隐蔽性和针对性，通常会运用受感染的各种介质、供应链和社会工程学等多种手段实施先进的、持久的且有效的威胁和攻击，从医疗信息、账户凭证、个人信息、企业电子邮件到企业内部敏感数据等等，不一而足。

2021 年 4 月，黑客利用从暗网购买了一名联合国雇员的登录账号，成功入侵联合国的计算机网络，并窃取到大量可用于攻击联合国组织的机构内部数据。据悉，泄露凭证属于联合国专有项目管理软件 Umoja 上的某个账户，导致联合国内部数据泄露的缘由，就是泄露账户未开启双因素身份验证，而这本是一项基础安全功能。黑客掌握联合国系统访问权的最早日期为 4 月 5 日，截至 8 月 7 日他们在联合国网络上仍然保持活跃。据报告该事件的 Resecurity 公司称，在最近这次入侵中，黑客试图找出关于联合国计算机网络架构设计的更多信息，并入侵了 53 个联合国账户，泄露账号在暗网的售价也仅为 1000 美元。联合国因未启用双因素身份验证的 Umoja 账户所导致的黑客入侵事件再一次佐证了内控安全不到位会引发严重的安全事故。当前，组织在抵御外部攻击的同时，要提防内部管理人员因操作失误、恶意操作、越权操作等问题而卷入数据泄露、运维事故的漩涡。

2021 年 5 月，据日本公共广播公司 NHK 报道，黑客攻击了富士通公司开发的信息共享平台 ProjectWEB，非法访问和窃取了日本中央省厅和重要基础设施企业数据。富士通开发的 ProjectWEB 软件被日本政府机构和企业广泛使用。富士通指出，攻击者获得了对使用 ProjectWEB 项目的未经授权的访问，并窃取了一些客户数据，受影响的日本政府包括国土、基础设施、运输和旅游部、内阁秘书处、成田国际机场等。

2021 年 8 月，T-Mobile 系统遭到了网络犯罪攻击，数百万客户、前客户和潜在客户的数据因此泄密，泄露的信息包括姓名、驾照、政府身份证号码、社会保障号码、出生日期、T-Mobile 充值卡 PIN、地址和电话号码。不法分子利用了解技术系统的专长以及专门工具和功能，访问了该公司的测试环境，随后采用蛮力攻击及其他方法，进入到了含有客户数据的其他 IT 服务器。T-Mobile 表示将向所有可能受到影响的人提供为期两年的免费身份保护服务（迈克菲的身份窃取防护服务），为后付费客户提供账户接管防护服务。T-Mobile 是一家跨国移动电话运营商，是德国电信的子公司，属



于 Freemove 联盟。T-Mobile 在西欧和美国运营 GSM 网络，并通过金融手段参与东欧和东南亚的网络运营。该公司拥有 1.09 亿用户，是世界上较大的移动电话公司之一。对于网络犯罪分子来说，这类公司具有较高价值。通信公司有义务保护好客户信息，需要在数据安全方面做更多功课。

2021 年 3 月，能源巨头壳牌公司遭遇黑客攻击，攻击者利用了安全厂商 Accellion 的文件传输程序 FTA 的零日漏洞，访问了一些个人数据以及属于壳牌利益相关方和子公司的数据。黑客以公开壳牌公司的敏感数据相要挟，要求壳牌公司支付赎金。Accellion FAT (File Transfer Appliance) 是美国 Accellion 公司开发的一个文件传输程序，它可以帮助企业使用一个本地或托管的私有云传输大的敏感的文件。利用 Accellion 漏洞攻击的核心问题是零日漏洞 (0-day)，零日漏洞是指被发现后立即被恶意利用的安全漏洞，在厂商发布安全补丁之前，可能已经给客户造成了巨大的损失。近年来利用零日漏洞的攻击事件在逐渐增加，虽然目前不能完全防范零日漏洞攻击，但是，科学完善的防御体系能有效减少被零日攻击的机率，以及降低零日攻击造成的损失。

2021 年 5 月，位于美国俄亥俄州费尔菲尔德进行 DNA 检测服务的 DNA 诊断中心 (DDC) 披露了一起数据泄露事件，黑客设法访问了用户的高度敏感的个人数据，包括支付卡数据。其中超过 210 万 (2102436) 名客户/用户的敏感个人和财务数据被黑客窃取。此外，金融账号被盗，平台账号密码也被盗。虽然信用卡信息可能已经过时，但全名和账户密码的暴露足以让客户成为其他平台后续攻击的受害者。

网络攻击者通过破译口令、IP 欺骗、DNS 欺骗等非法方式获得非法权限对被攻击的主机进行非授权操作，其手段多样、手法隐蔽，给用户造成巨大的经济损失，威胁社会和国家安全。因此，组织的数据安全管理面临着更高的要求和挑战，必须建立严格的安全保障体系，加强对数据访问权限的管控，对敏感数据进行加密处理，在网络中层层设防，让攻击者无缝可钻，无计可施。

D.3 内部员工泄露和删除数据事件频发

安全专家统计表明，组织在遭遇数据泄露事件时，有百分之八十的概率是出现在内部人员身上。内部人员利用合法获得的访问权对组织信息系统中的数据的机密性、完整性以及可用性造成负面影响，包括恶意的和非恶意行为。产生威胁的内部人员包括管理人员、管理员以及普通职员，包括离职人员泄露数据、为泄私愤删除数据、员工安全意识淡薄泄露数据等类型。

2022 年，葛兰素史克 (GlaxoSmithKline) 在美国马里兰州联盟法院指控前员工丹尼斯·布鲁克斯 (Denise Brooks) 窃取商业机密，据 GSK 提供的消息，被告于 2022 年 1 月 13 日“突然辞去”了 GSK 质量体系负责人的职务，令人意外的是，布鲁克斯在辞职当日大摇大摆地进入了 GSK 在马里兰州罗克维尔的工厂，他居然让工厂的 IT 人员把大量的商业机密从她的旧笔记本电脑拷贝到新的电脑上，而这些忙碌的 IT 员工并不知道布鲁克斯此时已经离职。在被 GSK 抓住把柄之后，这位员工无视了“多次要求归还 GSK 商业机密信息的请求”，忍无可忍的 GSK 最终选择将其告上法庭。公司认为，布鲁克斯非法拥有、转让和可能滥用机密将会“不可挽回地损害”公司利益。

2021 年，一家纽约信贷联盟的前雇员朱莉安娜·巴里尔承认，她被解雇后曾未经授权进入前公司的计算机系统，40 分钟内销毁了 2 万份文件，多达 21GB。被删除的文件包括客户的抵押贷款申请





和该金融机构的反勒索保护软件相关文件。她的举动不仅对客户造成了伤害，也给该信贷银行带来了巨大的困扰和安全风险。

2021 年，上海市一名 29 岁的程序员录某未经公司许可，在离职当天，私自将即将上线的京东到家平台系统代码全部删除，构成破坏计算机信息系统罪，被判处有期徒刑 10 个月。

2021 年 12 月 8 日，桂林市公安局发布通报称：12 月 7 日，卫某、黄某违反规定转发流调表，导致涉疫人员敏感信息泄露，造成不良的社会影响。卫某、黄某 2 人的行为侵犯了公民的个人隐私，桂林市公安局依据《中华人民共和国治安管理处罚法》相关规定，依法分别给予卫某妮、黄某鹏行政处罚。疫情人员信息泄露的事件已发生多起，主要原因都是由于内部关键工作人员的保密意识淡薄，对敏感信息不明确，且数据管理不到位；而机构对敏感信息无技术保护，可被随意发布分享至互联网社交软件。

虽然组织将大量资源集中用于应对缓解外部威胁行为者，但内部风险可能对组织构成更大的威胁。组织内部的员工引起的数据安全事件，是所有事件中代价最高昂且最难检测到的事件。内部人员相比外部攻击者有显著的“优势”，他们不仅熟悉组织的政策、程序和技术，而且对内部的弱点了如指掌。恶意内部人员甚至可能是安全设备的策略配置者，因此内部威胁相对于外部威胁，检测极其困难，他们可以在实施恶意行为时规避相应的检测机制，事后删除日志以逃避追溯。组织除了对员工进行技术和制度制约，定期进行保密培训，提高员工的保密意识也尤为重要。

D.4 非法数据交易造成数据泄露

近年来，非法获取、买卖公民个人信息的案件时有发生，被贩卖的个人信息规模之大，触目惊心，已经严重影响到个人的生活，甚至人身安全。

2020 年 1 月，网曝中国电信超 2 亿条用户信息被卖。据相关法院裁判文书显示，“2013 年至 2016 年 9 月 27 日，被告人陈亚华从号百信息服务有限公司（为中国电信股份有限公司的全资子公司）数据库获取区分不同行业、地区的手机号码信息，并提供给被告人陈德武，而陈德武则以人民币 0.01 元/条至 0.02 元/条不等的价格在网上出售，获利达 2000 余万元，涉及公民个人信息 2 亿余条。”另根据中国电信官方数据统计，截至 2016 年底，中国电信移动用户数达 2.15 亿户。也就是说，涉及到的 2 亿余条个人信息几乎是当时中国电信的全部客户信息。

2021 年 1 月江苏镇江丹阳警方成功侦破一起公安部督办的侵犯公民个人信息案，涉及 10 多个省市，抓获犯罪嫌疑人 30 名。该团伙采用境外聊天工具和区块链虚拟货币收付款，共贩卖个人信息 6 亿余条，违法所得 800 余万元。由于该案件侵犯对象广，犯罪嫌疑人摒弃传统的网络通讯工具和银行卡转账的收付款方式，采用境外聊天工具和区块链虚拟货币，犯罪手段新颖，社会危害严重，该案被公安部挂牌督办。

湛江银保监分局于 2021 年 1 月对建设银行湛江市分行存客户信息安全管理不到位的违规行为开出两张罚单，对该行做出罚款 20 万元的行政处罚。此外，涉事人王某因泄露客户信息，被禁止从事银行业工作 1 年。裁判文书以及相应罚单显示，2017 年至 2018 年期间，王某将共计 31465 条客户信息出售至贷款公司，获利 3.6 万。而上述信息则被相关贷款公司用于拨打电话并推销贷款业务信息，





从事不正当竞争。最终，王某被法院判处有期徒刑八个月。王某所在的建行湛江分行也因信息安全管理不到位被罚 20 万。

防范个人信息非法买卖，除了加大对非法收集、泄露、出售个人信息行为的打击力度，更要努力做到事前防范，提升自身的安全防范意识，尽力做到不使个人信息从自己手里泄露。对于所察觉的信息“黑市”交易行为及其人员，也应及时向执法部门举报，以公民群体力量织就严密的防范网络。

D.5 商业间谍窃取竞争对手商业秘密

不同于好莱坞《007》、《谍中谍》等影片塑造的高光亮丽间谍形象，商业间谍又被称为“工业白蚁”、“新的寄生虫”与“工业强盗”等，其作用主要是窃取他人的商业秘密和经济技术情报，比一般的间谍更加隐秘、更为广泛、更善伪装，甚至就在你我身边。

FBI 揭露了两位前 GE（通用电气）员工公然窃取公司知识产权和商业机密的行为。最初，参与高度复杂涡轮机制造的性能工程师德利亚通过自身权限下载了涉及该器械的机密性文件，并说服 IT 部门某员工帮助他获取了成本模型、提案和合同等文件。利用这些信息，他们成立了一家竞品公司，不仅在同产品上削弱了 GE，也在 FBI 进行调查的同时正常运行了多年。

2021 年 12 月 29 日，美国田纳西州一家提供直接面向消费者的虚拟护理服务公司对三个竞争对手提起诉讼，声称这些公司企图窃取他们的患者信息。MDSave 在 11 月发现 Sesame 和 Tripment 窃取了其网站的数据，并谎称他们与 MDSave 的供应商有直接合同。与此同时，Sesame 和 Tripment 的网站列出了许多与 MDSave 网站上的供应商名单相同的程序。

2021 年 2 月，最高人民法院知识产权法院宣判了人民法院史上判决赔偿额最高的侵犯商业秘密案件。在嘉兴中华化工公司等与王龙集团公司等侵害技术秘密纠纷上诉案中，公开宣判被诉侵权人王龙集团公司等盗用香料“香兰素”技术秘密，被判赔偿技术秘密权利人嘉兴中华化工公司等 1.59 亿元。

商场如战场，对于竞争激烈的商场而言，能够随时掌握竞争对手的动态，自然可以从容应对，立于不败之地。若以正当渠道收集数据自然无可厚非，但有些商家罔顾职业道德和法律，安排商业间谍窃取竞争对手的商业机密数据，给企业造成的经济损失也越来越大，有些企业最后甚至被迫倒闭破产。作为企业经营者，从企业管理的角度，对重要的商业秘密必须采取严格的保密制度，对关键岗位的人员要签订保密协议；从员工的角度，忠于自身职业操守，牢记“莫伸手，伸手必被捉”，严守企业的商业秘密，还要做有心人，时刻提高警惕、注意留心观察、强化保密意识，敢于、勇于、善于与侵犯企业商业秘密的行为作斗争。

D.6 系统或权限配置不当造成数据泄露

2021 年 7 月 29 日，Comparitech 研究人员报告，一个含有估计 3500 万个人详细信息的神秘营销数据库泄露在网上，居然未设密码。该数据库包括姓名、联系信息、家庭住址、种族以及众多的人口统计信息（包括爱好、兴趣、购物习惯和媒体消费等）。相关样本显示，大多数记录与芝加哥、



洛杉矶和圣迭戈这些大城市的居民有关。据 Comparitech 声称，凡是拥有网络浏览器和互联网连接的人都可以访问数据库全部内容，里面含有的信息可用于有针对性的垃圾邮件和诈骗活动以及网络钓鱼。Comparitech 网络安全研究团队在 6 月 26 日发现了该数据库，尽管使出了浑身解数，还是无法确定该数据库归谁所有。该公司联系了托管该数据库服务器的亚马逊网络服务（AWS），要求撤下数据库，不过，该数据在 7 月 27 日之前仍可以访问。

2021 年 8 月，vpnMentor 的研究团队发现，B2B 营销公司 OneMoreLead 将至少 6300 万美国人的私人数据存储在一个不安全数据库中，该公司任由此数据库完全敞开。该数据库包含列出的每个人的基本个人身份信息数据，以及有关其工作和雇主的类似数据和信息。这些信息很可能被提供给注册其 B2B 营销服务的客户或顾客。vpnMentor 看到了数据库中大量的 .gov 和纽约警察局电子邮件地址，这让黑客有可能渗透到原本安全的高级政府机构。vpnMentor 表示，政府和警察部门成员的私人数据如同从事犯罪活动的黑客眼里的金矿，可能导致重大的国家安全事件，使公众严重丧失对政府的信任。据 vpnMentor 称，姓名、电子邮件地址和工作场所信息暴露在任何拥有网络浏览器的人面前。

2021 年 2 月，福特汽车被曝存在一个较为严重的安全漏洞，黑客可通过该漏洞访问其配置错误的 Pega Infinity 系统，从而获取包括客户数据库、员工记录、内部票证等在内的专有数据信息。黑客还可以借此执行账户接管操作。该漏洞由 Robert Willis 和 break3r 发现，并得到了 Sakura Samurai 组织成员 Aubrey Cottle、Jackson Henry 和 John Jackson 的进一步验证和支持。该问题是由 CVE-2021-27653 漏洞引起的，它是一个信息泄露漏洞，存在于福特公司配置不当的 Pega Infinity 客户管理系统中。研究人员表示，泄露的数据资产包含敏感的个人身份信息、客户和员工记录、财务账号、数据库名称和表、OAuth 访问令牌、内部支持票、组织内的用户个人资料等。

德国专业话筒和耳机制造巨头森海塞尔（Sennheiser）将超过 28000 名客户的个人数据暴露在配置错误的 Amazon Web Services（AWS）服务器上。根据针对德国音频设备制造商森海塞尔的一份报告表述，森海塞尔在网上留下了一个不安全的亚马逊网络服务（AWS）服务器。该服务器存储了大约 55GB 的超过 28000 名森海塞尔客户的信息。

2021 年，Safety Detectives 在 10 月 13 日披露，巴西电商集成平台 Hariexpress 泄露了超过 610 GB 数据。此次事件是由于 Elasticsearch 服务器配置错误导致的，其中包含了超过 1751023279 条用户记录，包括客户的全名、家庭和送货地址、电话号码和账单明细。还暴露了卖家的全名、电子邮件和公司 / 家庭地址、电话号码和公司 / 税号（CNPJ/CPF）。错误配置的 Elasticsearch 服务器未进行加密，更没有密码保护。

服务器互联网暴露与配置问题是造成数据泄露与黑客入侵窃取数据的重要原因，组织应重视重要服务器的安全防护能力，采取及时更新软件补丁、定期进行漏洞扫描与检查等安全措施。

D.7 第三方人员或软件漏洞造成数据泄露

目前，许多企业会将系统开发、运维以及部分关键业务外包给第三方，包括服务提供商、合作伙伴、供应商等。但据来自 SecureLink 和 Ponemon 研究所共同发表的一份报告显示，51% 的组织经历过由第三方造成的数据泄露。无论是由于服务提供者所管理的在线资源配置不当、还是不安全的第三方



软件，或是与第三方不安全的通信渠道，第三方人员的安全意识不足等原因，如果组织并未有足够地关注与防范，那么与第三方合作可能会使组织面临巨大的风险。

2020 年，约克大学（University of York）透露，由于第三方服务提供商遭遇网络攻击，导致其数据遭到泄露。第三方云服务提供商 Blackbaud 受到指责。根据学术机构的说法，姓名、头衔、性别、出生日期、学生编号、电话号码、电子邮件地址、物理地址和领英的个人资料记录可能已被窃取。此外，课程信息、获得的资格证书、课外活动的细节、职业、雇主、调查反馈，以及校友记录和筹款活动都可能被曝光。

2021 年 3 月，ParkMobile 发现一起与第三方软件漏洞有关的网络安全事件。调查发现，其基本的用户信息被人访问，包括车牌号、电子邮件地址、电话号码和车辆昵称。在少数情况下，邮寄地址也被访问。该公司还发现加密的密码被访问，但读取这些密码所需的加密密钥并未被访问。ParkMobile 表示，它使用先进的散列和加入随机字符串（salting）技术对用户密码进行加密，以此保护用户密码。ParkMobile 表示，用户应考虑更改密码，作为另一道预防措施；信用卡或停车交易历史记录未被访问；它并不收集社会保障号码、驾照号码或出生日期。ParkMobile 称：“作为美国较大的停车应用软件，用户的信任是我们的重中之重。请放心，我们认真对待保护用户信息安全的责任。” ParkMobile 是在北美颇受欢迎的移动停车应用，用来显示街头可用的停车位 Parkmobile 还支持应用内支付停车费，即用户进入符合要求的距离之后可以在手机上为车位付费。不过需要特别注意的是，该功能只面向 ParkmobilePro 付费用户开放。它还能提供停车费折扣、路边援救以及临时优惠活动。在给用户带来方便的同时，其安全性也需要进一步加强。

2021 年，据美联社报道，管理着全球主要航空公司（超过 400 家航空公司，包括星空联盟和 OneWorld 会员）的机票处理和常旅客数据的 IT 服务公司——SITA（全球信息技术公司）宣布服务器被黑客入侵，全球多家航空公司的乘客数据遭到泄露。近十几家航空公司已经通知乘客，由于黑客侵入了 SITA 的乘客服务系统（PSS），乘客的一些数据已经被侵入者窃取。乘客服务系统负责处理从机票预订到登机的交易数据和业务。受影响的旅客总数仍不清楚，但至少超过 210 万。

相比组织内部的数据安全管理，对第三方的风险管理更具有挑战性，对拥有成百上千供应商的组织来说，更是如此。对第三方合作伙伴的风险管理，涉及从合同签订之前到合同执行过程中一直到合同完成之后整个生命周期，并且在数字化环境下，需要得到领导充分的重视和支持，经多个业务和职能部门的协同来完成，过程更需要科学的方法。

D.8 疫情之下，远程办公导致数据泄露事件频发

Zoom 在 2020 年 4 月频繁登上数据泄露新闻，共有四起安全和隐私相关事件被媒体爆出：先是擅自将用户数据通过 SDK（软件开发工具包）共享给 Facebook（“SDK 共享信息事件”）；后又因用户分组功能的缺陷，把用户邮箱地址、头像暴露给使用同一邮箱域名的其他陌生用户（“用户分组事件”）；又被发现有一万五千多个 Zoom 用户会议视频在亚马逊云服务器上处于公开可浏览状态（“视频泄露事件”）；还被黑客撞库攻击，使黑客获取了五十多万万个 Zoom 账户的账户名和密码（“黑客撞库事件”）。正如媒体所评论的，新冠疫情期间用户量和股价都猛涨的 Zoom 正在被全球的黑客放





在显微镜下找漏洞。

自 2020 年 1 月新冠疫情在全球爆发以来，由于疫情波及范围广、持续时间长，对正常的商业运营造成严重影响。在这种特殊的背景下，远程办公，如在线会议、即时通信、文档协作、基于功能扩展模块实现的协同办公等，成为企业在疫情期间的刚需，但并非所有企业，包括远程办公产品和服务的供应商，已经准备好应对激增的流量。Zoom 发生的上述每个事件都折射出企业在远程办公场景下在网络安全和数据保护方面需要关注的问题。

此外，大多数的数据泄露事件为个人信息的数据泄露，组织保护措施不力导致数据泄露将触犯法律，个人隐私保护的安全合规性不容忽视。

D.9 人脸信息泄露，对个人生活造成危害

近年来，随着信息技术飞速发展，人脸识别逐步渗透到人们生活的方方面面。同时人脸识别技术所带来的个人信息保护问题也日益凸显。一些经营者滥用人脸识别技术侵害自然人合法权益的事件频发，引发社会公众的普遍关注和担忧。

上海小鹏汽车销售服务有限公司购买了具有人脸识别功能的摄像设备 22 台，全部安装在旗下门店，涉及 5 个直营店及 2 个加盟店，开通系统账号 8 个，2021 年 1 月至 6 月期间，共计收集上传人脸照片 431623 张。通过算法对面部数据进行识别计算，以此进行门店的客流统计和客流分析，包括进店人数统计、男女比例、年龄分析等。收集消费者面部识别数据，并未经得消费者同意，也无明示、告知消费者收集、使用目的。上海小鹏汽车销售服务有限公司被徐汇区市场监督管理局罚款 10 万元。截至案发，上海小鹏汽车销售服务有限公司已拆除上述门店内的人脸识别摄像设备，上传的人脸照片已进行删除。

2021 年 2 月，江海区检察院在排查中发现，辖区 7 个住宅小区物管公司在未向行政主管部门申请并审核验收的情况下，擅自安装并投入使用“人脸信息识别”门禁系统，且已完成大部分小区业主的人脸信息收集。小区物管公司涉嫌侵害不特定多数人的个人信息安全，被江海区检察院提起行政公益诉讼。

2021 年 3 月 16 日，肥西县市场监管局接到举报人举报，称安徽佳创房地产开发有限公司在徽创佳兆业未来城售楼部未经举报人同意，通过摄像头收集其人脸信息，举报人提供了售楼部工作人员赵某某通过微信向举报人发送的举报人人脸收集照片。4 月 1 日，肥西县市场监管局联合肥西县公安局网安大队对该公司营销中心进行现场检查，发现营销中心内部有网络摄像头 4 部，在其办公电脑中发现名为“佳兆业人脸识别风控管理平台”（网址）的应用程序，其中存有大量人脸拍摄图片、与人脸相匹配的客户姓名、到访时间等信息。

随着人脸识别及其关联技术的发展，“刷脸进门”“刷脸打开”“刷卡支付”等现象已成为人们的生活常态。在此情况下，相较于身份证号码等其他个人敏感信息而已，公民对人脸信息被收集、存储的现象早已“习以为常”，对信息被泄露也已“见怪不怪”，而恰恰是这种“习以为常”“见怪不怪”削弱了人们对人脸信息安全的保护意识。随着《个人信息保护法》、《个人信息安全规范》等法律法规的出台和实施，为政府监管和公民维权提供了法律依据，在法律层面为人脸信息的保护





工作提供了指引。同时个人也应提高自我保护意识、增强自我保护能力。

D.10 儿童数据泄露问题日趋严重

近年来，在“眼球经济”“流量为王”的互联网经济背景下，未成年人特别是儿童因其年龄小、阅历浅、自我保护意识较弱，相关个人信息易泄露，受到网络骚扰、网络欺诈等侵害的案件时有发生。检察机关发现相关网络运营商、互联网企业存在疏于保护或违法获取使用未成年人个人信息问题，未充分履行企业的社会责任。

2021年3月11日，浙江省杭州市余杭区检察院诉国内某知名短视频公司（下称“某公司”）侵犯儿童个人信息民事公益诉讼案，经杭州互联网法院出具调解书后结案。某公司在开发运营该公司APP的过程中，未以显著、清晰的方式告知并征得儿童监护人有效明示同意允许注册儿童账户，并收集、存储儿童个人信息。在未再次征得儿童监护人有效明示同意的情况下，向具有相关浏览喜好的用户直接推送含有儿童个人信息的短视频，同时也没有采取技术手段对儿童信息进行专门保护。检察机关就该案提出了停止侵权、赔礼道歉、消除影响、赔偿损失等诉求，某公司均无异议，目前已针对存在问题全面开展整改。

一款具有10万+的销售记录的儿童智能手表，通过抽奖等伪装二维码将恶意程序轻松植入手表，后台可以实时定位、多次收集位置信息。孩子每次抽奖恶意程序就自动把手表里的重要信息如位置、通讯录、通话记录打包实时发送出去。玩完抽奖游戏孩子下楼玩耍，仍然可以实时定位孩子，轻松圈定孩子的活动范围，孩子的一举一动随时被掌握。

儿童是特殊群体，需要给予特殊保护。2019年10月1日，《儿童个人信息网络保护规定》正式实施，针对中华人民共和国境内通过网络收集、存储、使用、转移、披露不满十四周岁的儿童个人信息进行规范。《个人信息保护法》中也将不满十四周岁未成年人的个人信息作为敏感个人信息加以保护。真正实现儿童个人信息网络保护是多方参与、齐抓共管的过程，既依赖于法律规范、政府监管，也依赖于儿童监护人、行业组织以及社会公众在其中发挥的作用。





E.2020 年以来典型数据安全相关法律案件分析

E.1 滴滴因数据安全问题，被启动网络安全审查

2021 年 6 月 30 日，滴滴在美国纽交所挂牌上市。滴滴赴美上市引发监管关注，2021 年 7 月 2 日，根据国家互联网信息办公室的公告，为防范国家数据安全风险，维护国家安全，保障公共利益，依据《国家安全法》《网络安全法》，网络安全审查办公室按照《网络安全审查办法》，对“滴滴出行”实施网络安全审查。为配合网络安全审查工作，防范风险扩大，审查期间“滴滴出行”停止新用户注册。后续，国家互联网信息办公室发布了两则通告，认为“滴滴出行”、“滴滴企业版”等滴滴旗下 APP 存在严重违法违规收集使用个人信息问题，要求应用商店下架对应 APP。

数据安全问题也对滴滴在美上市的后续活动产生了较大影响，2021 年 12 月 3 日，滴滴出行发布公告，“经认真研究，公司即日起启动在纽交所退市的工作，并启动在香港上市的准备工作。”

滴滴事件展现了国家对于数据安全问题的高度重视，也表明数据安全已成为境内企业境外上市过程中不可回避的重要问题。为了给企业境外上市中的数据保护工作提供清晰的指引，证监会于 2022 年 4 月 2 日公布了《关于加强境内企业境外发行证券和上市相关保密和档案管理工作的规定（征求意见稿）》，旨在进一步加强境内企业境外发行上市相关保密和档案管理工作，明确上市公司信息安全责任，维护国家信息安全。因此，境内企业应高度重视数据安全和个人信息保护工作，落实数据安全保障责任，避免因境外上市或其他数据处理活动而影响国家安全。

E.2 科勒卫浴违规收集人脸信息，遭“3·15 晚会”曝光并被罚款 50 万元

科勒卫浴旗下上海家饰佳科勒卫浴专卖店安装了具备人脸识别功能摄像头，一旦顾客进入科勒卫浴门店，人脸信息就会被捕捉、记录。科勒卫浴获取客户人脸数据的行为，并未对消费者进行任何提示。

在 2021 年央视“315 晚会”节目中，曝光了科勒卫浴上述违规收集人脸信息的行为。2021 年 3 月 16 日，科勒卫浴针对“315 晚会”中的报道发表声明称，科勒卫浴高度重视晚会中提及的问题，并第一时间展开整改工作，目前已安排相关门店连夜拆除该摄像设备。

2021 年 7 月 26 日，上海市静安区市场监督管理局作出《行政处罚决定书》，认定科勒卫浴在利用摄像设备收集消费者人脸信息时，并未取得消费者的明示或授权同意，违反《消费者权益保护法》第二十九条第一款之规定，依法作出罚款人民币 50 万元的决定。

科勒卫浴案例表明，违规收集个人信息不仅会使企业遭到媒体曝光，影响企业声誉的损失，也会遭到行政处罚。对于收集人脸信息等敏感个人信息的活动，应只有在具有特定的目的和充分的必要性，并采取严格保护措施的情形下才可以开展。而且，处理敏感个人信息还应取得个人的单独同意。建议企业可以做好个人信息分类管理，遵守《个人信息保护法》等相关法律法规对处理个人信息及敏感个人信息的规定。



E.3 某科技公司利用爬虫窃取简历数据，被判 4000 万元罚款

某科技公司成立于 2014 年，主要经营招聘工具软件和大数据分析等业务。2015 年至 2019 年间，该公司组建专门爬虫技术团队，在未取得求职者和平台直接授权的情况下，秘密爬取国内主流招聘平台上的求职者简历数据。

针对上述行为，海淀区检察院对某科技公司、王某某等人提起诉讼。案件审查发现，涉案数据中具有爬虫特征的个人信息共 2.1 亿余条，某科技公司获取上述数据后，对数据进行重整，并用于开发产品意图谋利，期间，某科技公司爬虫技术团队负责人欧某某，私自将公司窃取的简历数据对外出售，个人非法获利人民币 30 余万元。

该案件最终的判决为，某科技公司、王某某等人侵犯公民个人信息，某科技公司被判处有期徒刑人民币 4000 万元，被告人王某某被判处有期徒刑七年，罚金人民币 1000 万元，其他被告人均被判处相应刑罚。本案对被告单位判处的罚金数额、对被告人判处的刑期和罚金数额，均系近年来全国同类案件判罚最重案例。

爬虫是互联网企业获取信息的技术手段之一，但是近年来，因爬虫行为而构成著作权侵权、不正当竞争，甚至触犯刑法的案例不在少数。本案为企业通过爬虫手段收集数据敲响了警钟，企业在进行爬虫前，应从爬取行为、爬取内容、以及爬取用途等方面进行合规性审查，避免爬虫行为被认定为违法犯罪行为的风险。

E.4 淘宝 12 亿用户数据遭爬取，危及用户个人信息安全

2020 年 8 月，淘宝（中国）软件有限公司报警称，2020 年 7 月 6 日至 13 日之间，有黑产通过 mtop 订单评价接口绕过平台风控批量爬取加密数据，爬取字段量巨大，7 月 6 日至 13 日之间平均每天爬取数量 500 万，爬取内容包括买家昵称、用户评价内容、昵称等敏感字段。经淘宝网站排查，逯某有重大作案嫌疑。

经司法鉴定，逯某通过其开发的软件爬取淘宝客户数字 ID、淘宝昵称、手机号码等淘宝客户信息共计 11.8 亿条，逯某将其爬取信息中的淘宝客户手机号码通过微信文件的形式发送给黎某使用，共计 1971.26 万条。黎某在收到手机号码后将信息数据导入“微信加人”软件，加微信好友成功后，由公司员工负责发送广告链接，用户在广告群中购买商品，公司从中获得佣金。通过上述模式，黎某开设的浏阳市泰创网络科技有限公司在一年间获得经营额超过 395 万元，违法所得超过 34 万元。最终，法院认为被告人逯某受雇于被告人黎某，二人违反国家规定，非法获取公民个人信息，情节特别严重，其行为均已构成侵犯公民个人信息罪。基于此，法院判处黎某有期徒刑三年六个月，并处罚金 35 万元；判处逯某有期徒刑三年三个月，并处罚金 10 万元。

上述用户数据泄露事件，危害了用户的个人信息安全，也反映出互联网平台风控机制的漏洞。互联网平台企业应完善自身的数据安全保护机制，通过采取有效的组织措施和技术措施，保护用户的个人信息。



E.5 圆通员工泄露用户数据，引发负面舆情

2020年7月，圆通速递有限公司河北省区内部员工与外部不法分子勾结，利用员工账号和第三方非法工具窃取运单信息，导致40万条个人信息泄露，泄露信息的内容包括发件人及收件人的地址、姓名、电话。圆通速递在安全检测过程中发现、报案，并协助调查，相关犯罪嫌疑人已于2020年9月落网。

圆通员工泄露数据的行为侵害了个人隐私。2020年11月16日起，境内多家媒体刊载相关报道，“圆通内鬼租售账号导致40万条个人信息泄露”等相关话题引发社会关注。2020年11月19日，上海市网信办网安处会同青浦区网信办、青浦公安分局、区商务委、区邮政管理局约谈了圆通公司，责令要求圆通公司认真处理员工违法违纪事件，做到信息对称、及时公开、正面应对，加快建立快递运单数据的管理制度。

圆通内部员工泄露信息，体现了该公司的数据安全内部管理制度存在严重问题。为防止数据泄露，公司应履行《网络安全法》、《数据安全法》以及《个人信息保护法》下的合规义务，制定内部管理制度和操作规程，合理确定个人信息处理的操作权限，并定期对从业人员进行安全教育和培训。

E.6 多家银行因网络安全及个人信息保护合规问题，被处以高额罚款

针对个人信息保护及网络安全问题，包括中国银行保险监督管理委员会、中国人民银行及各分支机构在内的主管机构对部分金融机构侵害消费者金融信息安全的行为进行处罚，如：

2021年1月29日，中国银行保险监督管理委员会对中国农业银行股份有限公司作出行政处罚，罚款420万元，原因在于中国农业银行股份有限公司存在下述违法违规行为：（一）发生重要信息系统突发事件未报告；（二）制卡数据违规明文留存；（三）生产网络、分行无线互联网络保护不当；（四）数据安全管控较粗放，存在数据泄露风险；（五）网络信息系统存在较多漏洞；（六）互联网门户网站泄露敏感信息。

2020年10月20日，中国人民银行长春中心支行发布的行政处罚信息公示表显示，中国农业银行股份有限公司吉林市江北支行因侵害消费者个人信息依法得到保护的权利；违反反洗钱管理规定，泄露客户信息，受到警告处分并被罚款1223万元。

2020年10月19日，中国人民银行德阳市中心支行发布的行政处罚信息公示表显示，建设银行德阳分行存在侵害消费者个人信息依法得到保护的权利的违法行为，处以罚款1406万元。

金融领域是强监管的领域，网络安全与个人信息保护也是监管机构关注的重点。因此企业不仅应遵守包括《个人信息保护法》等在内的法律法规，也应遵守金融领域监管机构的相关规范。金融机构应完善内部治理，严格按照合法、正当、最小必要原则收集、使用和保管客户信息，充分保障个人在个人信息处理活动中的权利。同时，对于重要的业务承载系统，也应加强安全防护，避免发生信息安全事件。



F. 主要数据安全相关标准汇总

针对国家出台的数据安全相关法律法规,围绕数据安全和个人隐私保护,国家及金融、电信、工业、互联网等重点行业颁布了系列技术标准,对指导各行业组织合法、合规地开展数据安全治理工作,促进数据充分利用、有序流动和安全共享,推动数字经济发展具有重要意义。

F.1 国家标准汇总

序号	标准号	名称	实施时间	简介
1	GB/T 35273-2020	信息安全技术 个人信息安全规范	2020.10.01	针对各类组织的个人信息处理活动,规范了开展收集、保存、使用、对外提供等个人信息处理活动应遵循的原则和安全要求。涉及的主要法律法规:《网络安全法》第 41-43 条,《个人信息保护法》等。
2	GB/T 39725-2020	信息安全技术 健康医疗数据安全指南	2021.07.01	给出了健康医疗数据控制者在保护健康医疗数据时可采取的安全管理和技术措施。涉及的主要法律法规:《数据安全法》第 6 条、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)等。
3	GB/T 39335-2020	信息安全技术 个人信息安全影响评估指南	2021.06.01	提出了个人信息安全影响评估的基本原理和实施流程,适用于各类组织自行开展评估,或主管部门开展检查评估。涉及的主要法律法规:《个人信息保护法》第 55、56 条等
4	GB/T 39477-2020	信息安全技术 政务信息共享 数据安全技术要求	2021.06.01	规定了政务信息共享过程中的数据安全技术要求以及相关基础设施的安全技术要求。涉及的主要法律法规:《数据安全法》第 42 条等。
5	GB/T 38667-2020	信息技术 大数据 数据分类指南	2020.10.01	提供了大数据分类过程及其分类视角、分类维度和分类方法等方面的建议和指导。涉及的主要法律法规:《数据安全法》第 21 条。
6	GB/T 37932-2019	信息安全技术 数据交易服务安全要求	2020.03.01	规定了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,可用于对数据交易服务机构进行安全评估。涉及的主要法律法规:《数据安全法》第 33 条等。
7	GB/T 37964-2019	信息安全技术 个人信息去标识化指南	2020.03.01	提出了个人信息去标识化的目标和原则,给出了去标识化过程和管理措施指导,适用于各类组织开展个人信息去标识化工作。涉及的主要法律法规:《个人信息保护法》第 51 条等
8	GB/T 37973-2019	信息安全技术 大数据安全管理指南	2020.03.01	提出了大数据安全管理基本原则,可指导组织开展大数据安全需求分析、数据分类分级、大数据风险评估等安全管理工作。涉及的主要法律法规:《数据安全法》第 27、30 条等。
9	GB/T 37988-2019	信息安全技术 数据安全能力成熟度模型	2020.03.01	给出了组织数据安全能力的成熟度模型架构,适用于组织开展数据安全能力建设,以及对组织数据安全能力进行评估。涉及的主要法律法规:《数据安全法》第 18、27 条等。涉及的认证:“数据安全能力成熟度认证”等。



序号	标准号	名称	实施时间	简介
10	GB/T 36073-2018	数据管理能力成熟度评估模型	2018.10.01	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能力成熟度分级模型及其评估方法。涉及的主要法律法规：《数据安全法》第 18、22、30 条。
11	GB/T 35274-2017	信息安全技术 大数据服务安全能力要求	2018.07.01	规定了大数据服务提供者应具有的基础安全能力、大数据服务生命周期安全能力以及大数据服务平台与应用相关的系统服务安全能力。涉及的主要法律法规：《数据安全法》第 27 条等。

F.2 行业标准汇总

序号	标准号	名称	实施日期	简介
1	JR/T 0098.8-2012	中国金融移动支付 检测规范 第 8 部分：个人信息保护	2012.12.12	规定了移动支付个人信息保护的内容管理、组织管理、访问控制和个人信息生命周期管理 4 个方面的基本要求以及检测细则。
2	YD/T 3797.1-2021	云服务用户数据保护能力评估方法 第 1 部分：公有云	2022.04.01	规定了公有云云计算服务提供者在提供云计算服务时应具备的用户数据安全保护能力要求和评估方法进行规范，包括事前防范能力、事中保护能力和事后追溯能力。
3	YD/T 3954-2021	云服务用户数据保护能力参考框架	2022.04.01	规范了云计算服务提供者在提供云计算服务时应具备的用户数据安全保护能力，包括事前防范能力、事中保护能力和事后追溯能力。
4	YD/T 3956-2021	电信网和互联网数据安全评估规范	2022.04.01	规定了数据安全评估实施流程和数据安全相关管理及技术措施的评估要点。
5	YD/T 3865-2021	工业互联网数据安全保护要求	2021.07.01	规定了工业互联网数据安全保护的范畴及数据类型、工业互联网数据重要性分级与安全保护等级划分方法，规定了低/中/高重要性数据在数据产生、传输、存储、使用、迁移及销毁阶段的具体安全保护要求。
6	YD/T 3867-2021	基础电信企业重要数据识别指南	2021.07.01	给出了基础电信企业重要数据的定义、识别规则、识别方法和重要数据安全保护实施指导，并给出了基础电信企业重要数据示例。
7	JR/T 0223-2021	金融数据安全 数据生命周期安全规范	2021.04.08	规定了金融数据生命周期安全原则、防护要求、组织保障要求以及信息系统运维保障要求，建立覆盖数据采集、传输、存储、使用、删除及销毁过程的安全框架。
8	YD/T 3797.2-2020	云服务用户数据保护能力评估方法 第 2 部分：私有云	2021.01.01	规定了云服务提供商在提供私有云平台时应具备的用户数据安全保护能力要求，包括事前防范能力、事中保护能力和事后追溯能力。
9	YD/T 3800-2020	电信网和互联网大数据平台安全防护要求	2021.01.01	准规定了大数据平台安全防护要求，包括大数据平台基础设施安全要求、平台安全要求、数据安全要求和安全管理要求。



序号	标准号	名称	实施日期	简介
10	YD/T 3801-2020	电信网和互联网数据安全风险评估实施方法	2021.01.01	提供以数据为核心保护对象的面向应用场景的风险评估方法论，规定了电信网和互联网数据安全风险评估的基本原则、基本要素及各要素之间的关系、实施流程、风险分析模型与方法。
11	YD/T 3802-2020	电信网和互联网数据安全通用要求	2021.01.01	本标准规范了数据采集、传输、存储、使用、开放共享、销毁等数据处理活动及其相关平台系统应遵循的原则和安全保护要求，包括组织保障、制度建设、规范建立等管理性要求，以及规范执行相关配套技术性要求。
12	YD/T 3806-2020	电信大数据平台数据脱敏实施方法	2021.01.01	本标准规定了电信大数据平台数据脱敏的指导原则、典型流程，对不同场景下的脱敏实施方法、脱敏算法选择给出了建议。
13	YD/T 3813-2020	基础电信企业数据分类分级方法	2021.01.01	规定了基础电信企业数据分类分级原则、数据分类工作流程和方法，数据分级方法，并给出基础电信企业数据分类分级示例。
14	YD/T 3736-2020	电信运营商大数据安全风险及需求	2020.10.01	规定了电信运营商大数据通用安全风险及相关需求，包括安全风险、安全需求、防护架构、技术要求等方面的内容。
15	YD/T 3746-2020	车联网信息服务用户个人信息保护要求	2020.10.01	规定了车联网信息服务用户个人信息保护的信息内容分类、敏感性分级和分级保护要求。
16	JR/T 0197-2020	金融数据安全 数据安全分级指南	2020.09.23	给出了金融数据安全分级的目标、原则和范围，以及数据安全定级的要素、规则和定级过程。
17	YD/T 3644-2020	面向互联网的数据安全能力技术框架	2020.07.01	定义了互联网行业的数据安全能力技术框架，该框架基于互联网行业的业务特点，在分析了覆盖数据全生命周期的数据安全威胁的基础上，提炼了数据安全能力建设的技术框架，并对框架内容进行了详细的描述。
18	JR/T 0171-2020	个人金融信息保护技术规范	2020.02.13	将个人金融信息按敏感程度、泄露后造成的危害程度，从高到低分为 C3、C2、C1 三个类别；同时，规定了个人金融信息在收集、传输、存储、使用、删除、销毁等生命周期各环节的安全防护要求，从安全技术和安全管理两个方面，对个人金融信息保护提出了规范性要求。
19	YD/T 3327-2018	电信和互联网服务 用户个人信息保护技术要求 即时通信服务	2019.04.01	规定了即时通信服务的用户个人信息保护技术要求。
20	YD/T 3367-2018	移动浏览器个人信息保护技术要求	2019.04.01	规范了移动浏览器个人信息的类型，分析了安全威胁，确定了相应的安全防护目标和技术要求。
21	JR/T 0158-2018	证券期货业数据分类分级指引	2018.09.27	给出了证券期货业数据分类分级方法概述及数据分类分级方法的具体描述，并就数据分类分级中的关键问题处理给出建议。



序号	标准号	名称	实施日期	简介
22	YD/T 3105-2016	电信和互联网服务 用户个人信息保护技术要求 电子商务服务	2016.10.1	规定了电子商务服务的用户个人信息及相关权益的保护要求。
23	YD/T 3106-2016	电信和互联网服务 用户个人信息保护技术要求 移动应用商店	2016.10.1	规定了移动应用商店服务的用户个人信息保护要求及与用户相关权益保护要求。
24	YD/T 3082-2016	移动智能终端上的个人信息保护技术要求	2016.07.01	规定了移动智能终端上的个人信息的类型，根据其不同的处理环节，结合每类信息的特点，对相应类型的个人信息保护提出具体的技术要求。
25	YZ/T 0147-2015	寄递服务用户个人信息保护指南	2016.01.01	规定了寄递服务用户个人信息的组成、信息保护的基本原则及信息保护的具体要求。
26	YD/T 2781-2014	电信和互联网服务 用户个人信息保护 定义及分类	2014.12.24	规定了电信和互联网服务用户个人信息保护的术语和定义、保护范围、信息内容和分类。
27	YD/T 2782-2014	电信和互联网服务 用户个人信息保护 分级指南	2014.12.24	规定了电信和互联网服务用户个人信息保护的分级对象和分级方法。

G. 国内外相关理论与介绍

G.1 微软的整体化数据治理方法 DGPC

2010 年，微软公司的可信计算研究部门发表研究论文提出了一套以隐私、保密和合规为目标的整体化数据治理方法 DGPC (Data Governance for Privacy, Confidentiality and Compliance)。DGPC 是为了帮助企业或组织能够以统一、跨学科的方式来实现以下三个目标，而避免让组织内不同部门各自独立解决三个割裂的问题：

- 1) 传统的 IT 安全方法侧重于 IT 基础设施，通过边界安全与终端安全进行保护，而 DGPC 的重点应该是加强对存储数据的保护，并随着基础设施移动，提供持续性保护；
- 2) 隐私相关的保护措施必须超越与安全重叠的隐私保护措施，包括向重点获取、保护和执行客户对如何及何时收集、处理或向第三方共享数据的行为提供保护措施；
- 3) 数据安全和数据隐私合规责任需要通过一套统一的控制目标和控制行为，并经过合理化处理，得到履行和兑现。

G.1.1 DGPC 的组成

DGPC 与组织现有的 IT 管理和控制框架（如 COBIT）以及 ISO/IEC 27001/27002 和支付卡行业数据安全标准（PCI DSS）等安全标准协同工作，由组织的人员、流程和技术三个核心能力领域组成。

G.1.2 开展风险评估的三项基础要素

G.1.2.1 信息生命周期

为了识别安全风险并选择合适的技术措施和行为来保护机密数据，组织必须首先了解信息如何在整个系统中流动，以及信息如何在不同阶段被多个应用程序和人为了不同目的被访问和处理。

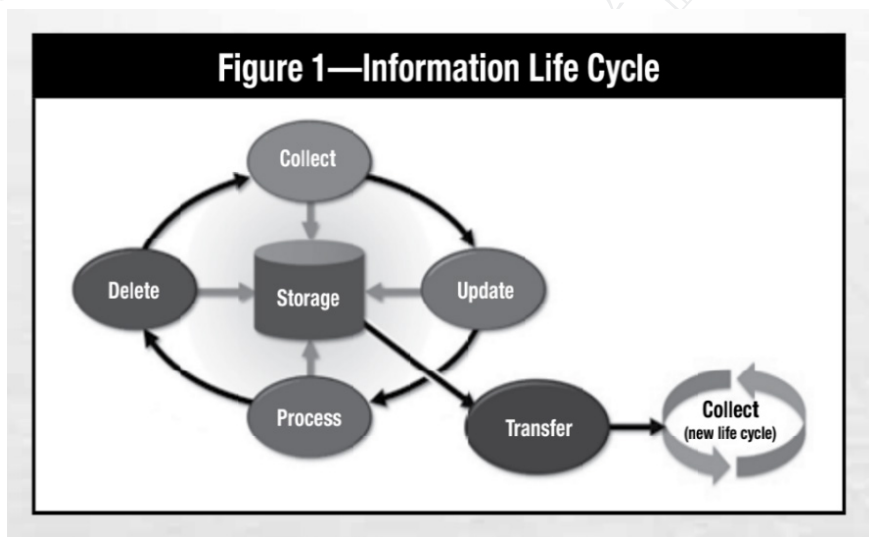


图 89 信息生命周期



G.1.2.2 四个技术领域

组织还需要系统评估保护其数据机密性、完整性和可用性的技术是否足以将风险降低到可接受的水平。以下技术领域为此任务提供了一个参考框架，包括安全的基础架构、身份和访问控制、信息保护、审计和报告、信息保护、审计和报告。

G.1.2.3 风险 / 差距分析矩阵和风险评估

风险 / 差距分析矩阵可帮助组织识别并解决现有保护工作的缺失：针对特定数据流中的隐私、机密和合规威胁的数据安全，该矩阵能提供数据现有和未来的保护技术、措施和行为，形成统一视图。每一行描绘信息生命周期中的一个阶段；矩阵中的前四列表示一个技术领域，而最右边的列表示控制行为，这些行为措施必须在信息生命周期的每个阶段满足四种数据隐私和保密原则的要求。

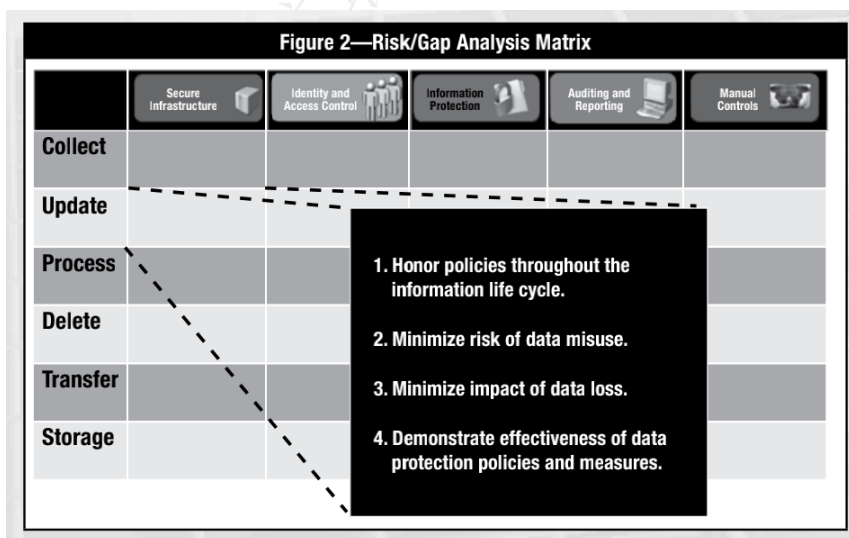


图 90 风险 / 差距分析矩阵

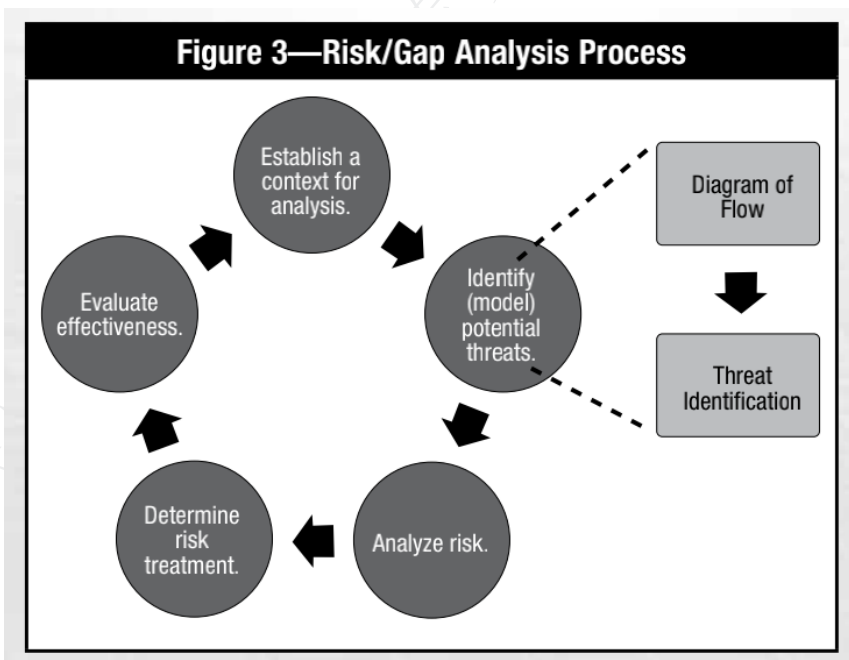


图 91 风险 / 差距分析过程

风险 / 差距分析矩阵为组织提供了一个强大的风险评估和缓解工具，组织利用该矩阵开展风险评估能够帮助组织找出现有保护措施中的不足并进行相应调整，这一过程包含五个步骤。

- 建立风险分析

这涉及定义数据流的业务目的，需要了解如何使用数据以及涉及哪些业务系统；并确定业务流程中的隐私、安全和合规目的。

- 执行威胁建模

大多数威胁建模技术只关注安全威胁，建模过程涉及两个阶段。

- 创建数据流的图表化呈现

微软的产品团队和咨询服务机构通常使用数据流图（DFD），并添加“信任边界”。如上图所示，信任边界是用来分隔业务实体和 IT 基础架构领域（如网络或管理域）的边界。在这种情况下，客户将 PII 提供给应用服务器，应用服务器将其存储在由云提供商管理的服务器中。每个事务都记录在由管理与应用服务器同一主机的日志服务器中。每次机密数据需要跨越信任边界时，对于安全性、策略、流程或实施的要求可能会发生变化，因此将会在图 92 中识别出威胁。

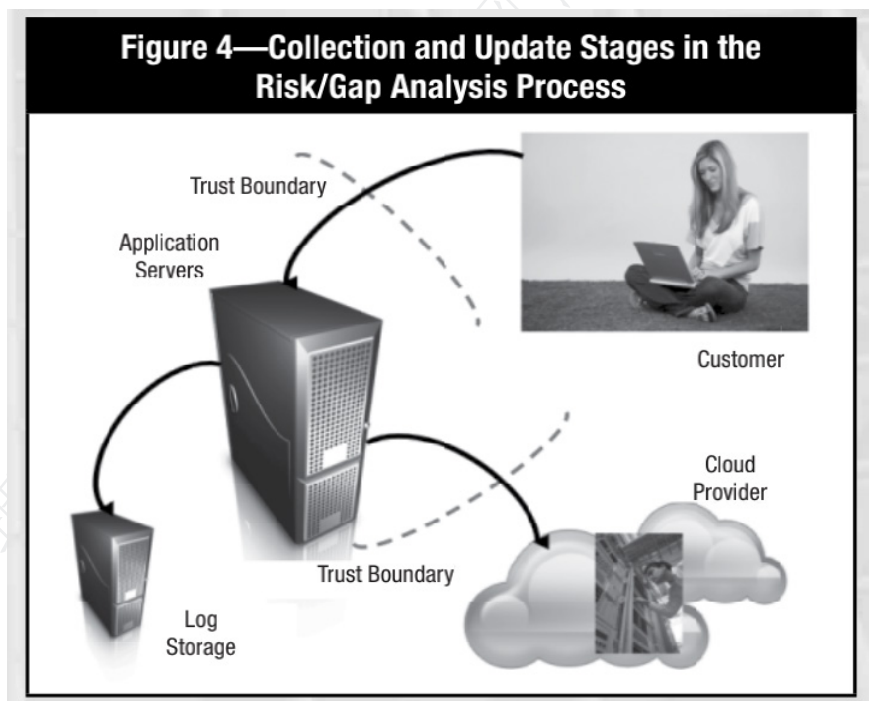


图 92 收集和更新风险 / 差距分析过程的阶段



Figure 5—Threat Identification	
Threat Type	Specific Threat
Choice and consent	Options have to be displayed clearly in order to obtain appropriate consent.
Access and correction	Customer is not able to view/modify personal information.
Accountability	Customer PII is not properly classified.
Compliance	Compliance reports are not defined; escalation path to business owners is not specified.
Information protection	Customer information is sent in the clear, over an unauthenticated channel.
Data quality	Quality depends on customer; no threat is identified.

图 93 风险类别

➤ 威胁列举

威胁不仅限于攻击者或技术威胁，还可以指代任何可能违反四项数据隐私和机密性原则的任何内容。组织应该使用这些原则来定义威胁类别(如上图所示)，再将其应用于数据流所产生的输出示例。这些类别的确切定义将取决于该组织的独特政策和适用的行业、地区和法律合规框架。

● 分析风险

大多数组织已经采取了一些措施来确保数据安全性和隐私性，正如其现有控制框架和信息安全管理信息系统所指定的那样。为了完成这一步骤，组织应首先统计现有的保护控制、技术和行为。然后，对于风险/差距分析矩阵中的每个单元格，确定哪些控制措施、技术和活动支持遵守四项隐私和保密原则。当矩阵的相应单元格中已识别现有保护措施未解决的威胁，并评估了相关风险时，此步骤即告结束。

● 确定缓解措施

在矩阵的适当单元格中，组织应列出使各项风险达到可接受水平所需的额外控制措施、技术和活动，然后评估每项风险的成本/收益。当组织决定是否以及如何降低，转移每个确定的风险时，这一步骤就结束了。

● 评估缓解措施的有效性

如果存在不可接受的风险，组织应审查前面步骤的结果并重新启动整个评估周期。

Figure 6—Identified Mitigation Measures					
	Secure Infrastructure	Identity and Access Control	Information Protection	Auditing and Reporting	Manual Controls
Collect/Update	Servers are on regular OS and App. Patch cycle, and up to date in malware signatures.	All transactions to take place on authenticated communications.	Choices are displayed and consent is obtained as per MPSD guide.	All material transactions are to be logged as per logging framework.	The escalation path for issues is defined.
	Incoming data are correctly classified and tagged as per customer choice and consent.		Transaction log data are encrypted in transit and at rest. All material customer transactions arrive over encrypted communication channel.	Communications channel and log servers are monitored. Failover process to local log servers in processor facilities is up and running. Alerts and alert recipients are defined and operational. Access and use reports, along with recipients and delivery schedules, are defined.	

图 94 评估缓解措施

G.2 高德纳的数据安全治理框架 DSG

Gartner 对数据安全治理的基本定义是“数据安全治理绝不仅是一套用工具组合而成的产品级解决方案，而是从决策层到技术层，从管理制度到工具支撑，自上而下、贯穿整个组织架构的完整链条。组织内的各个层级需要对数据安全治理的目标和宗旨达成共识，确保采取合理和适当的措施，以最有效的方式保护信息资源。”。

Gartner 的 DSG 框架建议的数据安全治理框架如下图所示，自顶向下依次从五个层面开展和执行。DSG 框架特别强调，数据安全治理工作需要从顶层的制定策略开始，而不能从部署安全保护工具开始。

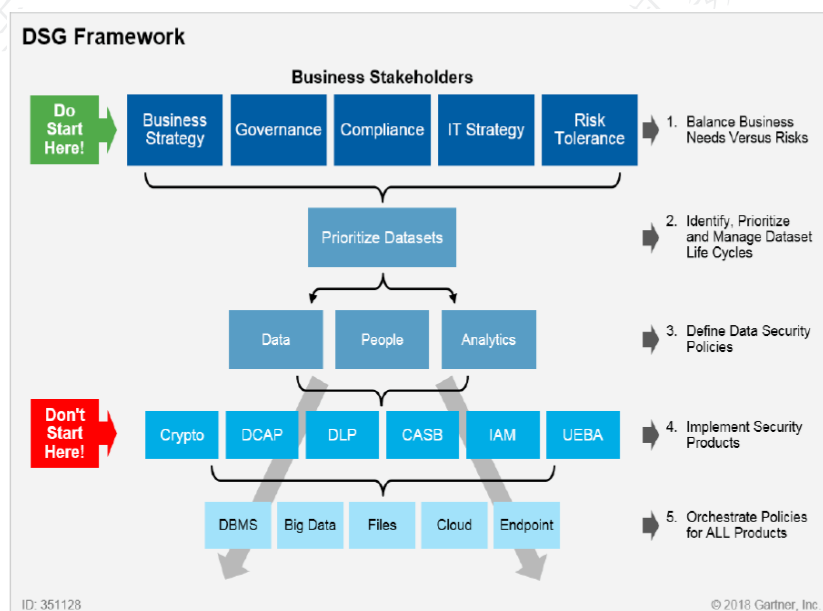


图 95 DSG 框架



● 平衡业务需求与风险

综合分组织战略、组织管理制度、合规要求、IT 战略、风险容忍度，制定符合业务发展与风险管控相平衡的数据安全治理策略。

● 数据梳理和管理数据的生命

数据梳理的工作内容是对数据进行分类分级，包括结构化数据和非结构化数据，目的是识别组织内的敏感数据，为下一步实施分级管控做好准备。

管理数据的生命周期，是分析评估数据在其生命周期内的各个环节面临的安全风险。Gartner 定义的数据生命周期包括获取创建、存储、分析、演变、归档、终结六个环节。

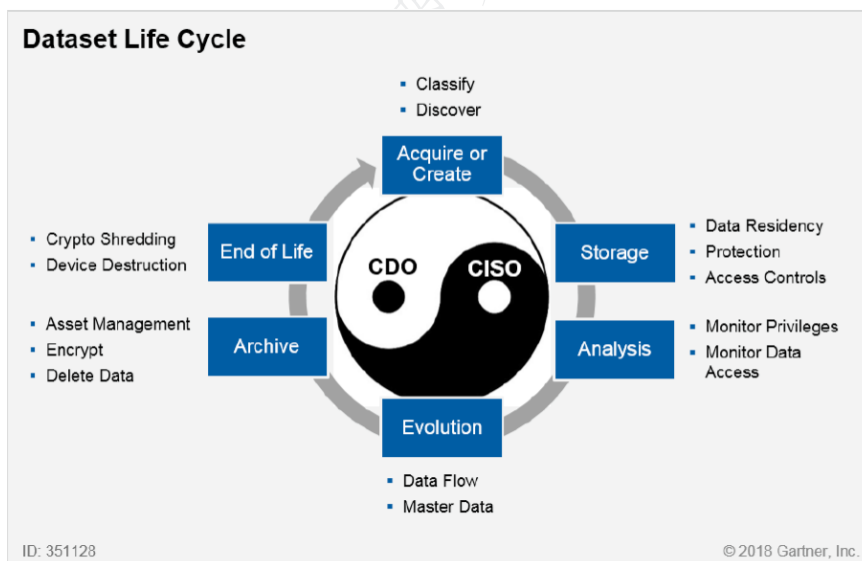


图 96 DSG 数据生命周期

● 制定数据安全控制措施

在数据分类分级的基础上，明确被保护的数据对象的访问控制规则。基于此制定不同级别的数据安全策略，包括数据安全保护的制度和方针、数据安全保护的组织架构和职责分工、数据安全保护的工作流程、数据安全保护的技术和工具等。

● 部署数据安全技术和产品

根据设计的技术管控手段，部署数据安全保护工具，如加密、数据库审计、数据防泄露、云安全代理、统一身份认证、用户行为分析等。

● 管理所有数据

将数据安全管控手段全面推广到各种类型的敏感数据，包括数据库、大数据、电子文件、云、终端等。

● 持续自适应风险与信任评估

为了实现对数据安全风险进行长久可持续地适应、评估和管控，Gartner 在 DSG 框架的基础上进一步提出了 CARTA (The continuous adaptive risk and trust assessment) 方法。CARTA 方法建议从阻止 (Prevent)、探测 (Detect)、响应 (Response)、预测 (Predict) 四个方面设计数据安全的管控措施，并通过四个方面的循环往复，持续优化管控体系。

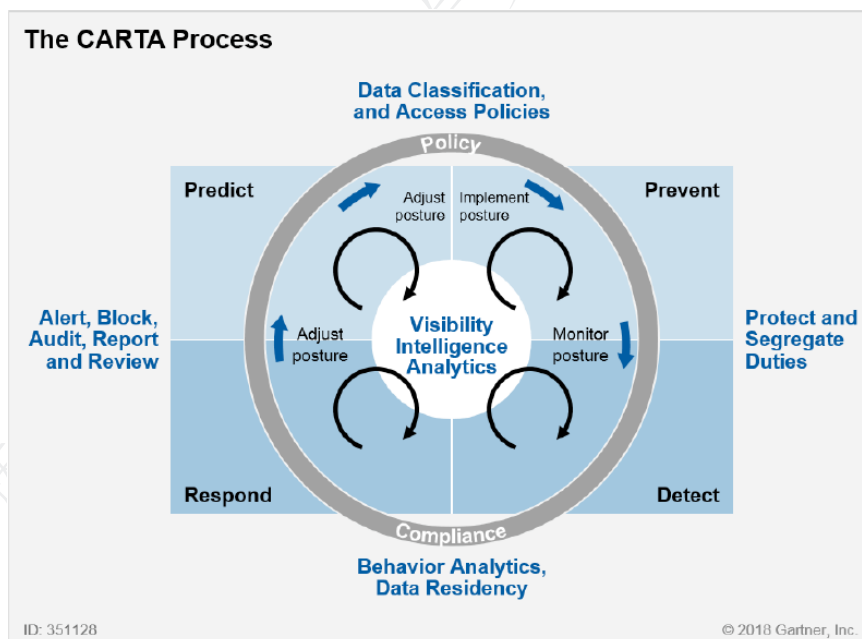


图 97 持续自适应风险与信任评估方法

G.3 DSMM 能力成熟度评估模型

国标《信息安全技术 数据安全能力成熟度模型》（GB/T37988-2019）给出了一个全方位评估企业或组织机构的数据安全能力成熟度的三维立体模型——DSMM。该模型将数据安全的能力成熟度分成从 1 到 5 共五个等级（由弱到强依次为非正式执行、计划跟踪、充分定义、量化控制、持续优化），分别从组织建设、制度流程、技术工具、人员能力四个方面对通用安全和构成数据生命周期（数据安全过程）的六个阶段（采集、传输、存储、处理、交换、销毁）及其进一步细分的三十个过程域内各自达到的数据安全能力成熟度等级进行评估。

DSMM 适合用来作为评估组织数据安全能力的方法和标准，亦可在组织开展数据安全能力建设的过程中被用作参考目标和依据。基于该标准的数据安全能力评估结果，可以鼓励数据在同等安全能力水平的组织间安全有序流动，或者流向能力水平更高的组织，避免数据流向低安全能力的组织。

数据安全治理和 DSMM 的区别主要体现在以下方面：

- 数据安全治理是以数据的分类分级为核心，进行安全策略的设定；DSMM 是以数据的生命周期为核心，寻求安全策略的覆盖。
- 数据安全治理体系化的建立数据使用或服务的人员角色，根据角色对数据使用的主要场景，提供建议的安全措施以及所要使用的安全工具；而 DSMM 更多的是提供一种评估方法，看数据使用的过程中，企业是否定义了明确的控制措施，并无具体化的方法指导。
- 相比 DSMM 完全围绕数据生命周期展开，数据安全治理通常反对在数据生命周期中不区分化地采用安全措施，而是更强调根据数据使用场景和数据使用需求，针对性地选用适宜的安全措施。比如在开发测试环境，需要采用脱敏的技术获得高仿真数据，在这种场景下，数据安全治理就会弱化推荐审计、管控和加密等措施。



● 数据安全治理将安全体系的构建，明确归纳为安全政策的制定，技术支撑平台的建设，对安全政策执行有效性的监督，安全政策的改善这一持续循环过程；而 DSMM 对安全成熟度进行了等级化分级，将持续改善定义为了最高级别，不像在数据安全治理理念中将持续改善视作标准组成部分之一。

● 数据安全治理本质上是一种方法论，帮助企业如何迅速构建一套行之有效的数据安全体系；而 DSMM 是一种评估方法，像一个考试一样，帮助企业或监管机构评价和了解企业或组织的当前安全能力状况。

从某种意义上讲，数据安全治理和 DSMM 之间是相辅相成、互助互补的关系。一方面，DSMM 能够为数据安全治理在事前提供开展工作的目标和方向，在事后提供对治理效果的评估方法；另一方面，数据安全治理又为追求 DSMM 描述的高级数据安全能力目标和弥补 DSMM 发现的数据安全短板提供了一套切实、可操作且行之有效的方法论。

目前 DSMM 已在 10 多个领域的 50 多家机构开展了落地实践，涵盖政府、银行、证券等行业。如在 2017 年 7 月，贵州省和阿里巴巴签订战略合作协议，发起成立全国第一个以数据安全为目标的“大数据安全工程研究中心”，其理事单位汇聚了中国信息安全认证中心、国家信息技术安全研究中心、中国电子技术标准化研究院等一大批国家级机构。同年 11 月，大数据安全工程研究中心落地贵阳市经开区；今年 4 月，在贵阳市政府的支持下，15 家当地企事业单位开始开展 DSMM 试点测评。

G.4 信通院 DSG 框架

当前我国各企业数据安全治理能力水平参差不齐，由于缺少以标准为基础的度量方法以及贴近产业现状的实施指南，企业无法准确衡量自身的数据安全治理能力建设情况与监管要求及公众期待的差距，进而无法明确自身的提升需求和目标。

针对这些问题，中国信通院以“准确度量企业的数据安全治理能力现状，合理规划数据安全治理能力提升路径”为目标，推出了“数据安全治理能力评估”（以下简称“DSG 评估”）。DSG 评估作为一套方法论体系，可以用于企业的数据安全治理体系建设参考，作为一套度量准则，可以用于考察企业的数据安全治理能力现状，作为一套改进指南，可以用于指导企业从组织、制度、技术、人员等方面有目标、有路径、分阶段的能力提升。

DSG 评估框架共包括数据安全战略、数据全生命周期安全、基础安全三部分。

1) 数据安全战略

包括数据安全规划和机构人员管理两个能力项，主要考虑从企业的顶层规划方面提出要求，为数据安全治理能力的建设搭框架、配人手。

2) 数据全生命周期安全

包括数据采集安全、数据传输安全、存储安全、数据备份与恢复、使用安全、数据处理环境安全、数据内部共享安全、数据外部共享安全、数据销毁安全在内的九个能力项，通过对数据全流过程进行规范和约束以有效降低数据安全风险。

3) 基础安全

包括数据分类分级、合规管理、合作方管理、监控审计、鉴别与访问、风险和需求分析、安全





事件应急等七个能力项，主要从数据安全的保障措施上进行定义和要求。

DSG 评估等级包括基础级、优秀级、先进级三个级别。基础级关注核心业务的安全治理能力及技术支撑情况，优秀级关注全业务流程的安全治理能力及技术手段的实施，先进级要求具备量化评估及优化改进机制。

综合来说，DSG 具备更加聚焦、容易操作、与时俱进、市场机制几个特点。DSG 评估专注于数据安全的相关内容，提供明确的评估方法，实操性强，并且紧跟立法要求和技术趋势，通过市场化的 DSG 评估，可以促进行业自下而上的向好发展。

从行业的角度，通过 DSG 评估可以详细了解行业数据安全治理能力发展现状，从企业的角度，DSG 评估可以帮助企业总结现状问题并发现差距，同时能够根据企业特点，推荐最佳行业实践，给出针对性的优化建议。此外，还可以通过参与数据安全治理相关的研讨交流，对外输出自身实践，对内吸收业内优秀案例，实现治理经验的提升。