



BDS 国家工程研究中心和天枢智库联合出品

数字安全观察

DIGITAL SECURITY INSIGHT

数据安全专刊 No. 008(总第 241 期)

责编：钟力 zhongli1@360.cn

SECURE THE FUTURE.

导 读

第八期《数字安全观察 数据安全专刊》梳理了2023年上半年的数据安全发展动态，分为政策形势、技术标准、市场趋势、安全事件分析、工程中心研究五个板块，主要内容如下：

政策形势方面，全球均在加快制定并完善数字经济与数据安全相关政策法规。国际方面，欧盟、美国、英国、印度、俄罗斯等国家持续完善数据安全方面的法律政策，并且尤其关注数据跨境传输方面的问题。同时世界各国都着力关注人工智能数据安全风险，强化人工智能领域的监管。国内方面，我国正全方位推进数据安全相关政策法规的健全与落实，据不完全统计，2023年上半年，已发布10余项国家政策法规、10项重点行业政策以及20余项地方政策规章。

技术标准方面，上半年数据安全技术创新不断，多项国家/行业标准陆续颁布。技术创新主要围绕量子计算、数据保密等，其中，面向数据安全的态势感知将成为新热点。此外，据工程中心不完全统计，2023上半年共有超过30项数据安全相关技术标准规范得到制定，多项国家标准围绕个人信息保护、数据安全风险评估等方面进行制定。

市场趋势方面，数据库安全、数据交易安全、API安全、数据要素等是热点话题，受益于发展数据安全产业的政策红利，数据安全市场也将持续增长，2025年数据安全市场天花板有望接近千亿元；同时，业界关于数据交易、数据出境安全评估的产品平台不断涌现，如

贵阳大数据交易所上线全国首个数据产品交易价格计算器、苏州市上线数据出境安全评估申报备案平台等。

安全事件分析方面，上半年数据安全形势依旧严峻，以数据泄露、数据加密勒索、数据合规等为代表的数据安全事件正在对全球数据安全形势造成深远影响。其中，数据泄露事件最为频繁，个人信息首当其冲。由外部攻击引发的数据安全事件类型最多，且78%的攻击源自有组织的犯罪团伙，可见如今的网络攻击都是高度专业化和战略性的，需要格外引起注意。

工程中心研究方面，BDS 国家工程研究中心 AI 安全实验室发布了国内首份《大语言模型提示注入攻击安全风险分析报告》，为国内大模型安全发展提供整体指南。报告指出，提示注入攻击已成大模型安全威胁之首，建议从安全测评、安全防御、安全监测预警等方面，多维度提升大模型的安全性。

目录

第一部分 政策形势总结

(一) 国内政策形势.....	5
1、稳妥推进数字经济建设，明确数据安全产业发展路径.....	6
2、日益健全数据出境安全制度体系.....	7
3、AI 浪潮方兴未艾，数据安全监管势在必行.....	9
4、推动落实数据安全岗位责任.....	10
5、切实加强地方数据安全法制.....	12
(二) 国际政策形势.....	16
1、各国数据安全治理规则不断完善.....	16
2、欧美国家加强对人工智能的监管立法.....	18

第二部分 技术标准动态

(一) 技术趋势.....	19
1、IBM 发布量子安全路线图.....	19
2、Gartner 发布 2023 年安全和风险管理技术采用路线图.....	21
3、隐私计算未来趋势：融合与提升，方案可落地.....	22
4、面向数据安全的态势感知将成为新热点.....	23
5、清华浙大在量子计算破解 RSA 密码方面取得重要突破.....	24
6、美国 NIST 推出物联网数据保护加密算法.....	25
7、使用神经网络，NIST 抗量子算法第四次被破解.....	26
8、美国政府更新零信任成熟度模型，将零信任转型作为长期目标.....	28
(二) 标准动态.....	29
1、国家标准.....	29
2、地方标准.....	31
3、行业标准.....	31

第三部分 市场趋势盘点

(一) 市场洞察.....	33
1、2024 年中国数据库市场规模将达 461 亿元，本土厂商热度持续攀升.....	33
2、数据要素市场十大研判.....	34
3、IDC 发布中国 API 安全市场洞察报告.....	36
4、IDC 发布中国数据安全基础设施管理平台市场洞察报告.....	37
5、2025 年数据安全市场天花板接近千亿元.....	38
6、2026 年中国数字化转型支出规模预计超过 6000 亿美元.....	40

- (二) 业界动态..... 42
 - 1、微软推出 Security Copilot：内置 GPT-4，自动抵御 65 万亿个网络安全威胁..... 42
 - 2、全国首个数据交易领域行业数据指数发布平台上线..... 43
 - 3、全国首个“算力资源专区”正式上线！..... 43
 - 4、上海探路数据交易资产化，国内首个数据交易链问世..... 44
 - 5、贵阳大数据交易所上线全国首个数据产品交易价格计算器..... 45
 - 6、苏州市数据出境安全评估申报备案平台上线..... 46
 - 7、中国移动发布“数联网”，保障数据“可用不可见”..... 46

第四部分 安全事件分析

- (一) 整体态势分析..... 47
 - 1、类型分析..... 47
 - 2、行业分布..... 48
 - 3、起因分析..... 49
 - 4、外部攻击画像..... 51
- (二) 数据泄露事件分析..... 53
 - 1、行业分布..... 53
 - 2、泄露规模..... 54
 - 3、泄露数据类型..... 55
- (三) 勒索攻击事件分析..... 56
 - 1、行业分布..... 56
 - 2、勒索金额..... 58
 - 3、活跃的勒索组织..... 58
- (四) 数据合规事件分析..... 59
 - 1、行业分布..... 59
 - 2、处罚金额..... 61
 - 3、违规原因..... 62
- (五) 总结..... 63

第五部分 工程中心研究

- 《大语言模型提示注入攻击安全风险分析报告》..... 64

一、政策形势总结

2022 年末，《中共中央国务院关于构建数据基础制度更好发挥数据要素作用的意见》（简称“数据二十条”）对外发布，从数据产权、流通交易、收益分配、安全治理等方面构建数据基础制度，提出 20 条政策举措。今年上半年，我国全方位推动“数据二十条”的落实，并持续增强和完善在数据跨境安全约束、数据安全人才及岗位等方面的机制建设，各省、市、地区纷纷争相全面加快数据要素市场的构建，并积极贯彻数据基础制度的落实。

国际上，全球各主要经济体持续完善和加强数据安全法律法规政策体系建设，尤其关注数据跨境传输方面的问题；同时在 ChatGPT 等大语言模型的带动下，各国都着力关注人工智能数据安全风险，强化人工智能监管已是大势所趋。

（一）国内政策形势

今年我国围绕“数据二十条”的落实任务，不断细化完善相关的法规政策。在数据出境安全制度体系、AI 领域数据安全监管、数据安全岗位职责等领域得到了较多的立法加强，反映了这一阶段数据安全领域的热点问题。同时北京、上海、广东等地区积极进行数据安全能力的示范性建设，全国上下的数据安全建设正在加快发展。

1、稳妥推进数字经济建设，明确数据安全产业发展路径

1月13日，工业和信息化部等十六部门联合发表了《关于促进数据安全产业发展的指导意见》，提出了我国数据安全产业的发展目标：到2025年，数据安全产业基础能力和综合实力明显增强；产业生态和创新体系初步建立，标准供给结构和覆盖范围显著优化，产品和服务供给能力大幅提升，重点行业领域应用水平持续深化，人才培养体系基本形成。到2035年，数据安全产业进入繁荣成熟期。为实现该目标，指导意见从提升产业创新能力、壮大数据安全服务、推进标准体系建设、推广技术产品应用、构建繁荣产业生态、构建繁荣产业生态、深化国际交流合作等方面明确发展方向。

2月27日，中共中央、国务院印发《数字中国建设整体布局规划》，指出建设数字中国是数字时代推进中国式现代化的重要引擎，是构筑国家竞争新优势的有力支撑。规划提出要通过构筑自立自强的数字技术创新体系、筑牢可信可控的数字安全屏障强化数字中国关键能力，强调要增强数据安全保障能力，建立数据分类分级保护基础制度，健全网络数据监测预警和应急处置工作体系。

3月10日，党的二十届二中全会通过了《党和国家机构改革方案》，明确要求组建国家数据局，负责协调推进数据基础制度建设，统筹数据资源整合共享和开发利用，统筹推进数字中国、数字经济、数字社会规划和建设等，由国家发展和改革委员会管理。

3月16日，国务院新闻办公室发布《新时代的中国网络法治建设》白皮书，强调网络安全是国家安全的新课题和新内容。中国通过制定《国家安全法》、《网络安全法》、《数据安全法》等法律，系统构建网络安全法律制度，增强网络安全防御能力，有效应对网络安全风险。在数据安全领域，白皮书指出，我们应当立足数据安全工作实际，着眼数据安全领域突出问题，通过立法加强数据安全保护，提升国家数据安全保障能力。

6月6日，国务院发布了《国务院2023年度立法工作计划》，计划中指出，制定网络数据安全条例是完善国家安全法治体系的重要一环，并将在2023年审议国家网信办组织起草的《网络数据安全条例》。

2、日益健全数据出境安全制度体系

1月18日，国务院办公厅转发商务部、科技部《关于进一步鼓励外商投资设立研发中心的若干措施》，支持研发数据依法跨境流动，落实网络安全法、数据安全法、个人信息保护法等有关法律法规要求，加强数据跨境安全管理，保障国家安全和社会公众利益，保护个人信息权益。高效开展重要数据和个人信息出境安全评估，促进研发数据安全有序自由流动。

2月14日，国家互联网信息办公室公布《个人信息出境标准合同办法》，对通过订立标准合同的方式开展个人信息出境的活动作出

具体制度安排，要求自主缔约与备案管理相结合、保护权益与防范风险相结合，保障个人信息跨境安全、自由流动。2021年，《个人信息保护法》规定了安全评估、个人信息保护认证、标准合同等三种数据出境安全管理方式，继《数据出境安全评估办法》和《个人信息保护认证实施规则》颁布实施之后，《个人信息出境标准合同办法》完善了我国个人信息出境管理制度的体系框架。

3月16日，全国信息安全标准化技术委员会发布了国家标准《信息安全技术 个人信息跨境传输认证要求》征求意见稿，规定了个人信息处理者跨境提供个人信息的基本原则、基本要求和个人信息主体权益保障要求，进一步为个人信息跨境传输认证制度提供了具体的合规标准。

5月30日，国家互联网信息办公室发布《个人信息出境标准合同备案指南（第一版）》，对个人信息出境标准合同备案方式、备案流程、备案材料等具体要求作出了说明，指导和帮助个人信息处理者规范、有序备案个人信息出境标准合同。《指南》发布后，上海、山东、浙江等地纷纷公布个人信息出境标准合同备案指引，保障各地备案工作的有序进行。

3、AI 浪潮方兴未艾，数据安全监管势在必行

2023 年上半年，ChatGPT 的成功加剧了生成式人工智能乃至通用人工智能的快速发展，也促使监管部门出台相应措施，加强对人工智能、科技领域数据安全的监督与管理工作。

4 月 4 日，科学技术部发布了《科技伦理审查办法（试行）》，对科技伦理审查的程序、标准、条件等进行规范，对覆盖各领域科技伦理审查作出综合性、通用性规定。其中，《审查办法》在涉及人类研究参与者的科技活动、以及涉及数据和算法的科技活动的要求中，明确了个人信息保护与数据安全的要求，加强了与个人信息保护法、数据安全法等法律法规的衔接。

5 月 23 日，北京市人民政府办公厅印发《北京市促进通用人工智能创新发展的若干措施》，强调在发展通用人工智能的过程当中，创新主体应当加强网络和数据安全管理，落实网络安全、数据安全和个人信息保护主体责任，强化安全管理制度建设和工作落实。鼓励创新主体开展数据安全认证及个人信息保护认证，落实数据跨境传输安全管理制度，全面提升网络安全和数据安全防护能力。

7 月 10 日，国家网信办等七部门联合公布《生成式人工智能服务管理暂行办法》，促进生成式人工智能健康发展，防范生成式人工智能服务风险。《办法》提出促进生成式人工智能技术发展的具体措施，明确训练数据处理活动和数据标注等要求，对生成式人工智能服

务规范进行明确，并要求对生成式人工智能服务实行包容审慎和分类分级监管。

4、推动落实数据安全岗位责任

3月20日，人力资源社会保障部办公厅、中央网信办秘书局、工业和信息化部办公厅颁布《数据安全工程技术人员国家职业标准》，将数据安全工程技术人员根据职业功能划分为数据安全管理工作、数据安全工程规划设计和建设实施、数据安全技术开发与运维、数据安全监测与应急处置、数据安全评估5个方向，并将能力划分为初级、中级、高级，明确了各等级专业技术人员的工作领域、工作内容以及知识水平、专业能力和实践要求。

近年来，首席数据官作为机构统筹管理数据资源的第一责任人，对于推进数据要素市场化配置改革具有关键作用。据不完全统计，截至2023年，全国一共19个省、直辖市、自治区先后出台了首席数据官的制度：

表1 首席数据官制度清单

序号	省份	日期	政策措施
19	四川省	2023年5月29日	《四川省企业首席数据官制度建设指南（试行）（征求意见稿）》
18	江苏省	2023年5月8日	《江苏省数字政府建设二〇二三年工作要点》
17	江西省	2023年4月10日	《关于深化“放管服”改革巩固提升一体化政务服务能力的若干措施》
16	广西壮族自治区	2023年2月4日	《广西壮族自治区首席数据官制度试点工作方案》
15	吉林省	2023年1月31日	《白山市政务服务和数字化建设管理局2022年度法治政府建设工作报告》

14	新疆维吾尔自治区克拉玛依市	2023年1月9日	《克拉玛依市首席数据官制度试点工作方案》
13	天津市	2022年12月13日	《关于建立首席数据官制度的工作方案》
12	湖北省武汉市	2022年12月9日	光谷启动实施“首席数据官”制度
11	云南省迪庆藏族自治州	2022年12月1日	《迪庆州“十四五”电子政务发展规划》
10	福建省厦门市	2022年9月1日	《厦门经济特区数据条例（草案）》
9	湖南省长沙市	2022年8月16日	《支持长沙市深化“放管服”改革和数字化转型的实施方案》
8	安徽省合肥市	2022年8月1日	《合肥市开展首席数据官试点工作方案》
7	北京市	2022年7月27日	《北京市数字经济促进条例(草案)》
6	辽宁省沈阳市	2022年6月1日	沈阳在全省率先试点推行首席数据官制度
5	河南省	2022年4月7日	《河南省国资国企数智赋能三年行动计划（2022—2024年）》
4	黑龙江省	2021年12月30日	《黑龙江省“十四五”数字政府建设规划》
3	上海市	2021年11月25日	《上海市数据条例》
2	浙江省	2021年6月4日	《浙江省数字政府建设“十四五”规划》
1	广东省	2021年4月23日	《广东省首席数据官制度试点工作方案》

2023年上半年，首席数据官的建设正逐步从制度架构迈向试点落实。

2月4日，《广西壮族自治区首席数据官制度试点工作方案》印发，方案明确了首席数据官的职责，提出首席数据官是本单位、本辖区的数据治理的指挥者、数据资产的管理者、数据应用的推广者、数据安全的监管者、数据思维的倡导者，既负责顶层设计，也负责跟踪落实、监管推广。

5月8日，江苏省人民政府公布《江苏省数字政府建设2023年工作要点》，强调将在2023年着力完善开放共享的数据资源体系方面，将探索建立省级首席数据官制度。

5月29日，四川省经济和信息化厅发布《四川省企业首席数据官制度建设指南（试行）（征求意见稿）》，文件指出，四川省将在全省建立CDO工作机制，支持企业设立CDO岗位，培育数据管理人才梯队，促进企业数字化发展，助力“数字四川”建设。

5、切实加强地方数据安全法制

（1）北京市

2月2日，《2023年市政府工作报告重点任务清单》印发，其中强调加快数字贸易港建设，加强数据跨境流动国际合作，深入探索数据流动、交易制度规则，并提出强化网络安全、数据安全和个人信息保护。

3月17日，《关于推动北京互联网3.0产业创新发展的工作方案（2023-2025年）》印发，针对互联网3.0监管，提出聚焦互联网3.0内容监管、数据安全、隐私保护、身份可信、资产确权等，加强监管机制和监管模式探索，利用区块链、隐私计算、网络安全、量子加密等新型监管技术，实现对互联网3.0虚拟世界的可监管和可审计，保障互联网3.0数据的安全与隐私，提升对虚拟世界监管的智能化水平。

5月12日,《北京市智能网联汽车政策先行区数据安全管理办法(试行)》发布,填补了国内自动驾驶示范区级数据安全管理的空白,明确了在市自动驾驶办公室统筹指导下,企业负数据安全主体责任,构建了示范区企业数据能力提升及共享机制,厘清了智能网联汽车产业数据安全管理的关键环节,详细梳理了重点数据类型的合规风险,并创新性构建了示范区数据安全能力建设机制。

(2) 上海市

1月5日,《临港新片区国际数据产业专项规划(2023-2025年)》发布,确立“产业+数据”的融合发展模式,形成具有国际影响力的数字产业集群,重点发展数据加工分析、数据安全保障、数据流动交易等国际数据服务;推动企业参与数据跨境安全流动评估,加快智能网联汽车、金融贸易、智能制造、跨境物流、文化创新等不同产业领域的企业启动数据跨境流动安全自评估与申报,加强数据安全风险评估、信息共享、监测预警等技术能力建设。

2月28日,上海市通信管理局宣布开展“浦江护航”——2023年电信和互联网行业数据安全专项行动,提出试点实施电信和互联网行业首席数据官制度、开展重要数据和核心数据识别认定及目录管理、开展电信和互联网行业数据安全风险评估管理、开展常态化数据安全监测预警与通报处置、加强企业数据全生命周期安全管理、加强数据安全能力建设和人才培养六项重点任务。

5月9日,《张江数据要素产业集聚区建设三年行动方案在(2023-2025)》公布,提出要构建与国家级数据交易平台相匹配的数据安全保障体系,建设关键信息基础设施安全防护态势感知系统,建立分层处置、分类计算、动态评估的数据安全管理体系,增强数据安全监测预警和应急处置能力,支持市场化安全资源池及中小企业盾建设,帮助企业守牢网络与数据安全底线。

(3) 广东省

3月2日,《深圳市数据交易管理暂行办法》印发执行,明确了对数据交易主体、数据交易场所运营机构、数据交易标的、数据交易行为、数据交易安全等方面的要求。其中在数据交易安全方面,《办法》强调数据交易场所运营机构应当制定数据安全事件应急预案,对重要系统和数据库进行容灾备份,定期开展数据交易环境安全等级保护测试和渗透测试等数据安全应急演练,提升数据安全事件应对能力。

3月2日,《深圳市数据商和数据流通交易第三方服务机构管理暂行办法》印发执行,围绕数据商和数据流通交易第三方服务机构的业务运行、安全管理、监督管理等方面作出要求。在安全管理方面,《办法》指出数据商应当采取安全保护管理措施,设立安全管理部门,建立健全数据安全分类分级管理、员工访问权限管理、供应商资质管理和内部审计等制度,定期开展安全教育培训。数据商应当建立健全

数据安全监测预警与应急处置机制，及时开展风险评估，制定数据安全应急预案。

4月4日，广东省政务服务数据管理局牵头起草《广东省数据流通交易管理办法（试行）》及系列文件。《办法》指出，数据经纪人是经省数据流通交易主管部门认定，利用行业整合能力，通过开放、共享、增值服务、撮合等多种方式整合利用有关数据，促进行业数据与公共数据融合流通的中介服务机构。规定数据经纪人承担数据安全主体责任，并以此为基点，细化为遴选认定、组织运行、业务管理、安全评估等一系列制度。

（4）其他地区

2月17日，《杭州市公共数据授权运营实施方案（试行）》（征求意见稿）发布。《方案》对公共数据的加工使用主体提出明确要求，主体在满足公共数据安全体系评估结果无高风险项的前提下，应当建立公共数据授权运营内部管理和安全保障制度，具备通过网络安全等级保护三级标准的系统开发和运维实践经验，同时按照《数据安全管理体系认证实施规则》通过数据安全管理体系认证规范数据处理活动。

4月3日，《2023年河南省大数据产业发展工作方案》发布，规定了完善数据基础设施、培育数据要素市场、推动产业链现代化、优化产业发展生态、提升数智治理水平、完善安全保障体系六个重点任务。在安全保障体系方面，《方案》提出，要落实数据分类分级保护

制度，探索建立覆盖数据全生命周期的数据安全体系，建设数据安全态势感知平台，并打造一批 DSMM 贯标标杆企业以便后期推广。

4月27日，《江苏省数字政府建设2023年工作要点》发布，提出了要牢固构建全方位的安全防护体系，包括完善安全管理机制确保政务系统和数据安全边界清晰、职责明确、责任落实，强化安全技术支撑提升网络、系统、数据安全技术防护水平，加强安全监测运营，守牢网络和数据安全底线。

5月3日，台湾立法院通过修正个人资料保护法的草案，将成立台湾数据保护机构，该机构有权处以高达1500万台币的罚款。根据草案，未能实施适当安全措施以防止个人数据被盗窃、篡改、损坏、销毁或泄露的非政府机构可能面临2万至200万台币的罚款。

（二）国际政策形势

2023年上半年，随着ChatGPT等生成式人工智能大模型的横空出世，人工智能侵犯数据隐私的风险也成为各国政策关注的热点，世界各国围绕隐私数据保护、数据跨境传输共享、人工智能安全等领域不断完善立法建设，保护自身数据安全，愈发复杂化的数字地缘战略博弈不断加剧全球数据治理困境。

1、各国数据安全治理规则不断完善

2月9日，欧盟通过《数据法案》（Data Act），消除企业获取数据的障碍以促进创新，保证数据环境的公平性，同时定义公共部分访

问和使用私营部门数据的方式;5月8日,网络安全认证小组(ECCG)宣布对新的《欧盟云服务网络安全认证计划草案》展开审查,要求云服务数据需在欧盟存储和处理,限制欧盟外实体对云服务提供商(CSP)的控制。

2月21日,中国外交部在《全球安全倡议概念文件》中提出,要进一步深化信息安全领域国际合作,共同应对各类网络威胁,构建开放包容、公平合理、安全稳定、富有生机活力的全球网络空间治理体系。

3月31日,美国白宫科技政策办公室(OSTP)发布《促进数据共享与分析中的隐私保护国家战略》,支持使用保护隐私数据共享和分析技术(PPDSA),该技术利用增强隐私技术进行数据分析同时确保隐私安全;6月8日,美国和英国联合发布《二十一世纪英美经济伙伴关系大西洋宣言》,强调两国将侧重在数据传输、人工智能安全和隐私增强技术(PETs)三个领域深化合作。

除了美国和欧盟,其他主要经济体也在积极推动数据安全建设。2月1日,印度宣布推动建立数据大使馆,享有与实体大使馆类似的外交豁免权,解决数据存储和跨境数据流动的问题。3月1日俄罗斯关于向国外传输个人数据的新规定生效,禁止运营商向未对个人数据主体权利提供充分保护的国家或地区传输数据。7月18日,英国颁布《数据保护和数字信息法案》,进一步明确负责任的数据使用,

保护个人权益，减轻企业负担，消除国际贸易壁垒，鼓励和保障自动化决策。

2、欧美国家加强对人工智能的监管立法

美国国家标准与技术研究院（“NIST”）于 2023 年 1 月 26 日发布了人工智能风险管理框架（AIRMF 1.0），概述了构建人工智能相关风险的方法和可信赖的人工智能系统的特征，从治理、映射、测量和管理四个方面帮助组织在实践中解决人工智能系统的风险。4 月 11 日，美国商务部下属国家电信和信息管理局（NTIA）发布“人工智能问责政策”征求意见稿（RFC），针对是否需要 ChatGPT 等人工智能工具实行审查、新的人工智能模型在发布前是否应经过认证程序以及如何更好地建立人工智能问责制等问题征求意见。

英国政府于 3 月 29 日发布《支持创新的人工智能监管方式》（ A pro innovation approach to A I regulation ）政策文件，概述了监管机构应考虑 5 项核心原则安全性和稳健性、透明度和可解释性、公平性、问责和治理、可竞争性和补救性，以促进人工智能的安全和创新使用。

意大利个人数据保护局在 3 月 31 日宣布，即日起禁止使用 ChatGPT，限制 ChatGPT 的开发公司 OpenAI 处理意大利用户信息，并开始立案调查。

加拿大监管机构在 4 月 6 日收到投诉称聊天机器人 ChatGPT 未经同意收集、使用和披露个人信息，隐私专员办公室开始调查 OpenAI 的生成式应用程序 ChatGPT 。

5 月 16 日，法国国家信息和自由委员会（CNIL）发布了最新的《人工智能行动计划》，为监管人工智能系统、保护个人隐私和支持法国及欧洲创新型人工智能系统提供框架和支撑。

6 月 14 日，欧洲议会投票通过《人工智能法案》，禁止实时面部识别，并对 ChatGPT 等生成式人工智能工具提出了新的透明度和风险评估的要求。

二、技术标准动态

上半年数据安全相关技术突破与标准不断更新。技术趋势方面，数据加密、量子计算等创新突破不断，隐私计算技术也进入快速发展阶段。此外，面向数据安全的态势感知将成为上半年热点。标准动态方面，据工程中心不完全统计，2023 年上半年发布并制定数据安全相关标准 30 余项，多项国家标准围绕个人信息保护、数据安全风险评估等方面进行制定。

（一）技术趋势

1、IBM 发布量子安全路线图

关键词：量子安全

IBM 近日在年度 Think 会议上推出了量子安全路线图，以帮助政府机构和大型商业组织过渡到后量子密码学。

美国政府要求联邦机构必须在最后期限之前完成向量子安全加密的过渡。预计企业也将遵循同样的路径，但这是一条漫长而艰难的道路。IBM 开发了一个三阶段解决方案，称为 IBM 量子安全路线图（Quantum Safe Roadmap）。

该路线图包括三个独立的产品，其中两个现已上市，第三个目前是第一代。这三种产品分别是量子安全资源管理器(Quantum Safe Explorer)、量子安全顾问(Quantum Safe Advisor)和量子安全修复器(Quantum Safe Remediator)。这三个产品对应于路线图的发现、观察和转换阶段。

“量子安全资源管理器”使组织能够扫描源代码和目标代码，以查找加密资产、依赖项和漏洞。结果就是 IBM 所说的加密物料清单（CBOM），将所有相关细节信息聚合到一起。

“量子安全顾问”创建加密清单的动态视图以指导修正，并分析加密状况和合规性以确定风险的优先级。

“量子安全修复器”使组织能够部署和测试基于最佳实践的量子安全修正模式，以了解在准备部署量子安全解决方案时对系统和资产的潜在影响。“它允许您使用不同的量子安全算法、证书和密钥管理服务，” IBM 说：“它还可以帮助您实现加密敏捷性，以便您可以快速适应不断变化的政策和威胁，而不会对运营或预算产生重大影响。”

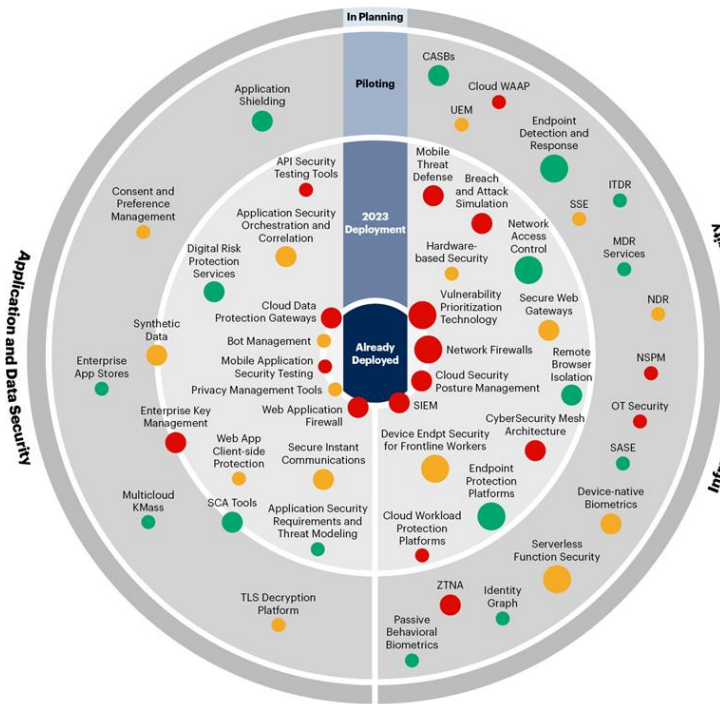
(来源: [GoUpSec](https://www.gartner.com))

2、Gartner 发布 2023 年安全和风险管理技术采用路线图

关键词: Gartner 安全和风险管理技术

Gartner 根据采用阶段、部署风险和企业价值, 绘制了 2022 年至 2024 年间大型全球企业中 49 项安全相关技术的实施情况。安全和风险管理领导者可以使用此信息图来衡量他们的计划并根据主要趋势进行衡量。

Technology Adoption Roadmap for Security and Risk Management 2022-2024



Enterprise Value

The value factor awarded to each technology is based on the analysis of value drivers, including improved speed and agility, enhanced developer experience or productivity, increased cost efficiency or savings, delivery of superior capabilities to business and/or customers, and enabling resilience and reliability.



Deployment Risk

The risk factor awarded to each technology is based on the analysis of potential risks posed, including cybersecurity risk, talent unavailability, high or unpredictable costs, and technical incompatibility or architectural complexity.



Adoption Phase

The adoption phase is determined by the current deployment plans for a majority of organizations. Technologies placed on the border between phases are on the cusp of moving into the next deployment phase.



Source: Gartner

© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. 781655_C

Gartner

从对技术采用路线图调查数据的分析中得出，**安全和风险管理 (SRM) 领导者正在近期部署应用程序和数据安全解决方案，并为基础设施和边界安全制定更长远的规划。**在调查涵盖的 49 项技术中，51% 处于“部署中”，47% 处于“试验阶段”。

(来源：[安全内参](#))

3、隐私计算未来趋势：融合与提升，方案可落地

关键词：隐私计算

隐私计算是数据隐私保护的重要手段，在大数据时代越来越受到关注。未来，隐私计算的发展趋势将是技术融合，即将多种技术进行整合，形成更加综合和灵活的隐私计算方案。

在技术融合方面，多方安全计算、联邦学习和可信执行环境等技术将会相互协同，互相弥补优劣，提高隐私计算的效率和安全性。

在性能提升方面，未来的隐私计算趋势将是提升性能，通过优化硬件、算法和通信等方面的技术指标，以降低性能损耗，实现更高效的隐私计算。

在方案落地方面，隐私计算厂商需要根据用户的具体业务需求，提供定制化的解决方案，包括技术方案、应用方案和商业模式等。为了实现方案的落地，隐私计算厂商需要与行业合作伙伴紧密合作，共同推动隐私计算技术在具体场景下的应用。同时，隐私计算厂商还需

要提供全方位的技术支持和培训服务，帮助客户快速上手并应用隐私计算技术。

（来源：[“开放隐私计算”公众号](#)）

4、面向数据安全的态势感知将成为新热点

关键词：态势感知 数据安全

伴随着数字化转型，数据在推动数字经济高速发展的同时，数据滥用、数据泄露等安全事件频繁发生，数据安全风险日益凸显。

国际著名咨询机构提出建议，用户必须在整个数据生命周期中规划和缓解管理风险，从而解决数据安全性，隐私性，信任与道德、数据所有权，数据恢复等一系列问题。2021年9月1日，《中华人民共和国数据安全法》正式实施，数据安全上升到我国国家安全战略高度，并加速了数据安全体系建设。

在数据安全体系落地实践中，组织的数据安全风险迫切需要依赖系统化、产品化技术手段实现数据安全全生命周期管控，全面感知、分析、呈现数据安全风险的态势感知应运而生。

面向数据安全的态势感知围绕数据安全全生命周期管理，将数据安全技术、流程、产品和人有机的结合，融合数据资产梳理和发现、数据风险检测识别、数据风险集中分析、数据风险响应处置、数据风险态势呈现等能力，实现数据安全风险可感知、可溯源、可研判、可处置、可呈现，为数据安全运营提供基础。随着数据安全体系建设不

断发展演进，面向数据安全的态势感知势将成为新热点，推动数据安全建设，助力数据安全治理。

（来源：[“数字安全助手”公众号](#)）

5、清华浙大在量子计算破解 RSA 密码方面取得重要突破

关键词：量子计算 RSA 密码

在一项最新研究中，清华大学龙桂鲁、浙江大学王浩华等组成的团队创建了一种算法，仅用 10 个超导量子比特就实现了 48 位因式分解。该团队的最新实验表明，依靠整数因子化的公钥密码技术可能很快就会受到当今原始的 NISQ(含噪声中等规模)量子计算机的攻击。

据研究人员称，该算法是基于经典的 Schnorr 算法——使用格约化来分解整数，同时依靠量子近似优化算法 (QAOA) 来优化 Schnorr 算法中最耗时的部分，以提高因式分解的速度。

研究人员表示，“使用这种算法，我们已经成功地对整数 1961（11 位）、48567227（26 位）和 261980999226229（48 位）进行了因式分解，在超导量子处理器中分别使用了 3、5 和 10 个量子比特。对于 48 位的整数，261980999226229，我们也刷新了真正的量子设备中用一般方法算出的最大整数。”

使用这种算法的近期量子计算机可能能够处理更大的整数分解问题，可能打破广泛用于保护计算机数据和系统的 RSA-2048 加密方案。

(来源: [“安全圈”公众号](#))

6、美国 NIST 推出物联网数据保护加密算法

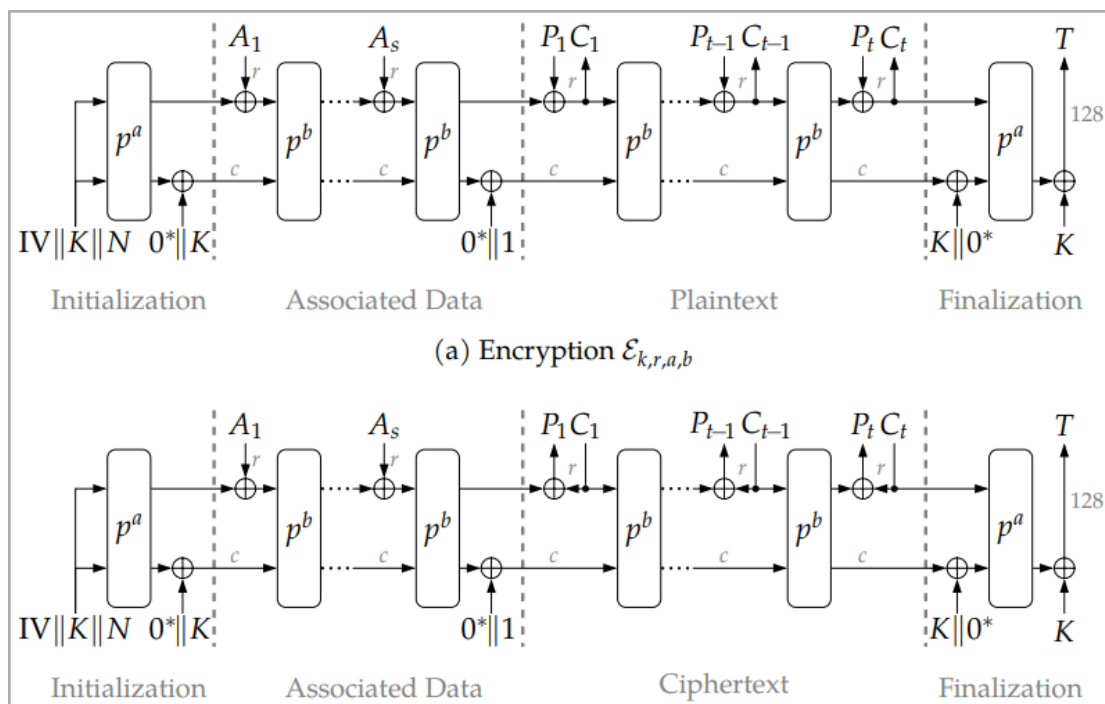
关键词: NIST ASCON

美国国家标准与技术研究院 (NIST) 近日宣布, 名为 ASCON 的认证加密和散列算法系列将成为标准算法, 用于轻量级密码学应用。该算法中标“轻量级密码学”计划, 旨在寻找最佳算法来保护硬件资源有限的小型 IoT (物联网) 设备。

小型物联网设备正变得越来越流行和无处不在, 用于可穿戴技术、“智能家居”应用等。但是, 它们仍然用于存储和处理敏感的个人信息, 例如健康数据、财务详细信息等。也就是说, 实施数据加密标准对于保护人们的数据至关重要。这些设备内部的薄弱芯片需要一种能够以极低的计算能力提供强大加密的算法。

ASCON 被选为提交给 NIST 的 57 项提案中的最佳提案, 领先的密码学家进行了几轮安全分析, 实施和基准测试结果, 以及研讨会期间收到的反馈。整个项目从 2019 年开始, 历时四年。

NIST 表示, 所有 10 名入围者都表现出超越既定标准的卓越性能, 而且没有引起安全问题, 这使得最终选择变得非常困难。ASCON 最终被选为获胜者, 因为它具有灵活性, 包括七个系列、节能、在弱硬件上速度快, 并且短消息开销低。



ASCON 的加密和解密操作模式 (NIST)

NIST 仍然推荐高级加密标准(AES)和 SHA-256 用于一般用途，但是这些不适用于资源有限的小型设备。

(来源: [bleepingcomputer](https://bleepingcomputer.com))

7、使用神经网络，NIST 抗量子算法第四次被破解

关键词：Crystals - Kyber

近日，瑞典皇家理工学院研究团队发表论文，称其提出一种新的神经网络训练方法“递归学习”(Recursive Learning)，并通过周期性循环旋转信息，实现了对美国国家技术标准研究院(NIST)四种抗量子密码安全算法之一 Crystals - Kyber 最高 5 阶掩码的侧信道攻击，以高于 99% 的概率从中恢复了信息位(message bit)。

Breaking a Fifth-Order Masked Implementation of CRYSTALS-Kyber by Copy-Paste

Elena Dubrova, Kalle Ngo, and Joel Gärtner

KTH Royal Institute of Technology, Stockholm, Sweden
{dubrova,kngo,jgartner}@kth.se

Abstract. CRYSTALS-Kyber has been selected by the NIST as a public-key encryption and key encapsulation mechanism to be standardized. It is also included in the NSA's suite of cryptographic algorithms recommended for national security systems. This makes it important to evaluate the resistance of CRYSTALS-Kyber's implementations to side-channel attacks. The unprotected and first-order masked software implementations have been already analysed. In this paper, we present deep learning-based message recovery attacks on the ω -order masked implementations of CRYSTALS-Kyber in ARM Cortex-M4 CPU for $\omega \leq 5$. The main contribution is a new neural network training method called *recursive learning*. In the attack on an ω -order masked implementation, we start training from an artificially constructed neural network M^ω whose weights are partly copied from a model $M^{\omega-1}$ trained on the $(\omega - 1)$ -order masked implementation, and then extended to one more share. Such a method allows us to train neural networks that can recover a message bit with the probability above 99% from high-order masked implementations.

Crystals - Kyber 已被 NIST 选为待标准化的公钥加密和密钥封装机制，同时也被纳入美国国家安全局(NSA)推荐用于国家安全系统的密码算法套件，这使得评估 Crystals - Kyber 对侧信道攻击的抵抗能力变得非常重要：侧信道攻击利用从物理可测量的非主信道获得的信息，例如运行实现的设备的时序或功耗。

瑞典皇家科学院的研究团队开创性的提出一种使用递归学习的神经网络训练方法(Recursive Learning)，他们从人工构建神经网络开始训练，这一网络的权重是由可破解低阶的神经网络复制而来。其核心思想是将能破解低阶($\omega-1$)掩码的神经网络模型作为破解高阶(ω)掩码神经网络模型的基础，然后循环前进至更高阶。

(来源: [安全内参](#))

8、美国政府更新零信任成熟度模型，将零信任转型作为长期目标

关键词：零信任

4月12日消息，美国网络安全和基础设施安全局（CISA）发布《零信任成熟度模型》第二版，对跨多个关键支柱（包括身份、网络、工作负载及数据等）的联邦机构实施指南进行了更新。

第二版更新了政府范围内采用零信任安全架构的关键定义和指标。这套模型是联邦机构在设计并实施零信任架构时可以采取的几种途径之一。其距离 CISA 发布首版零信任成熟度模型已经一年有余。

新版本的成熟度阶段分为四级，除了首版提出的传统、高级、最佳三个阶段外，还增加了初级阶段。CISA 修订了每个阶段的指导标准。此外，新版本不再只是临时文件，其更新了长期目标——支持联邦机构设计和实施零信任架构转型计划，并与白宫管理与预算办公室（OMB）M-22-09“推动美国政府迈向零信任网络安全原则”备忘录保持一致。

(来源: [安全内参](#))

（二）标准动态

1、国家标准

除了《信息安全技术 个人信息跨境传输认证要求》（征求意见稿）外，上半年中多项国家标准围绕个人信息保护、数据安全风险评估等方面进行制定。

3月17日，《信息安全技术 个人信息去标识化效果评估指南》（GB/T 42460-2023）发布，以数据能否或多大概率识别出个人信息主体的程度对个人信息标识度进行分级，从而评估个人信息去标识化的效果。同日，《信息安全技术 电信领域数据安全指南》

（GB/T 42447-2023）发布，给出了开展电信领域数据处理活动的安全原则、通用安全措施，及在实施数据收集、存储、使用加工、传输、提供、公开、销毁等过程中宜采取的相应安全措施，适用于电信数据处理者开展数据安全的建设、评估工作。

5月23日，全国信息安全标准化技术委员会发布了《网络安全标准实践指南——人脸识别支付场景个人信息保护安全要求（征求意见稿）》，《指南》针对室内外区域中的人脸识别支付场景，对人脸识别的服务提供方和实际识别场景中的场所管理方提出了具体的个人信息保护要求，并给出了人脸识别支付主要场景及其个人信息安全风险补充说明。《信息安全技术 移动互联网应用程序（App）个人信息安全测评规范》（GB/T 42582-2023）发布，从个人信息收集、

个人信息存储、个人信息使用、个人信息安全事件处置等多个方面提出了安全要求，以指导对 App 的个人信息安全测评和监管。《信息安全技术 个人信息处理中告知和同意的实施指南》(GB/T 42574-2023)对个人信息处理中的告知和同意适用情形、基本原则、方式进行细化，同时对 App 基本业务功能与扩展业务功能、智慧生活、个性化推送等 13 种场景下的告知和同意提出针对性地落实方案。《信息安全技术 网络安全事件分类分级指南》(GB/T 20986—2023)发布，与 GB/Z 20986-2007 相比，新《指南》更改了“分类方法”的表述，将网络安全事件的分类由 7 类增加至 10 类，并将“信息破坏事件”更名为“数据安全事件”，子类更改为“数据篡改事件、数据假冒事件、数据泄露事件、数据窃取事件、数据损失事件”，同时新增“社会工程事件、数据拦截事件、位置检测事件、数据投毒事件、数据滥用事件、隐私侵犯事件”。

5 月 26 日，全国信息安全标准化技术委员会发布了《网络安全标准实践指南—网络数据安全风险评估实施指引》，给出了网络数据安全风险评估思路、工作流程和评估内容，可用于指导数据处理者、第三方机构开展数据安全评估，也可为有关主管监管部门组织开展检查评估提供参考。

2、地方标准

3月28日，全国首个“数据资产价值与收益分配评价模型”标准在青岛发布。根据此标准开发设计的评价模型，对数据流通过程中产生的价值数据以可追溯且不可篡改的形式记录在区块链上，在上链数据可信的基础上进行量化计算、价值评价，并将评价结果作为数据资产评估的依据，为数据资产融资和入表做准备。

3月29日，上海市地方标准《公共场所人脸识别分级分类应用规范》（征求意见稿）正式发布，《规范》在明确公共场所人脸识别应用基本原则的基础上，对公共场所不同人脸识别应用场景进行分类，根据人脸识别应用目的、底库规模等风险因素进行分级，并基于分级分类针对性地提出了应用和管理要求。

4月7日，天津市市场监督管理委员会发布了《网络数据安全监督检查规范》，《规范》从安全合规管理、数据分类分级、终端数据安全等18个方面提出了79个数据安全检查要求，并对监督检查流程、检查方式以及检查结果进行了规定。

3、行业标准

2月1日起，由中国网络安全产业联盟拟定的《数据安全和个人信息保护社会责任指南》正式施行。《指南》从组织治理和内部管理，合规性、创新型和价值体现，公平运行、竞争与合作，消费者权益保护，公益参与和社会发展五个方面对组织在网络安全领域应当负有的

社会责任进行了详细的说明，并给出了数据安全和个人信息保护社会责任评价方法和社会责任报告模板。

4月7日，国家市场监督管理总局召开专题新闻发布会介绍《快递电子运单》和《通用寄递地址编码规则》两项国家标准的相关情况，《快递电子运单》国家标准设立专门章节，强化个人信息保护内容，其中包括禁止显示完整的个人信息，推荐对个人信息进行全加密处理，以及规范个人信息相关内容的读取权限等；《通用寄递地址编码规则》国家标准顺应我国邮政业数字化转型发展需要，提出了通用寄递地址编码的编码原则、编码规则和编码维护要求。

5月22日，工业和信息化部发表了《工业领域数据安全标准体系建设指南（2023版）》（征求意见稿），目标到2024年，初步建立工业领域数据安全标准体系，有效落实数据安全要求，基本满足工业领域数据安全需要，推进标准在重点行业、重点企业中的应用，研制数据安全国家、行业或团体标准30项以上；到2026年，形成较为完备的工业领域数据安全标准体系，全面落实数据安全相关法律法规和政策制度要求，标准的技术水平、应用效果和国际化程度显著提高，基础性、规范性、引领性作用凸显，贯标工作全面开展，有力支撑工业领域数据安全重点工作，研制数据安全国家、行业或团体标准100项以上。

三、市场趋势盘点

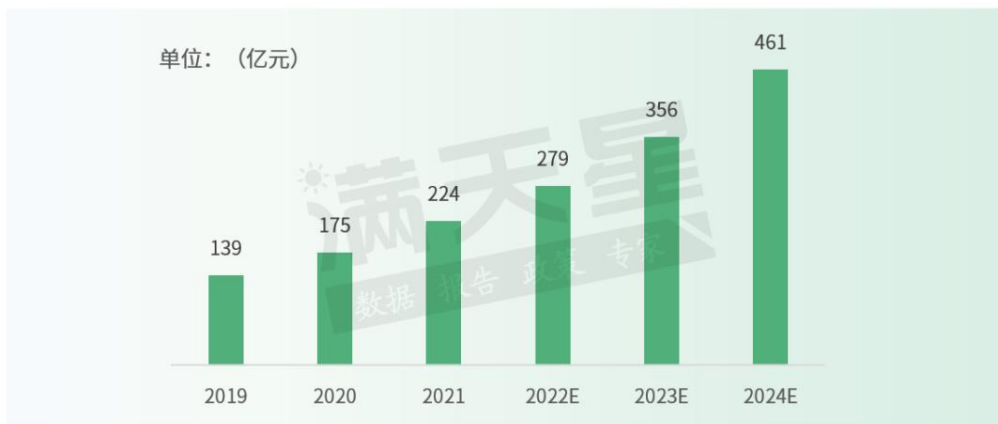
2023 上半年，数据库安全、数据交易安全、API 安全、数据要素等是热点话题，受益于发展数据安全产业的政策红利，数据安全市场也将持续增长。同时，业界关于数据交易、数据出境安全评估的产品平台不断涌现。

（一）市场洞察

1、2024 年中国数据库市场规模将达 461 亿元，本土厂商热度持续攀升

关键词：数据库

中国数据库管理系统市场保持快速增长趋势，预计到 2024 年将达到 461 亿元。从技术架构来看，集中式关系型数据库仍然为当前市场的主流，但云数据库及分布式数据库也成为众多厂商研发的重点以及用户部署的新方向。



数据来源：赛迪顾问

在未来的数据库选型过程中，有 53.8% 的调研对象对于数据库产品的安全可靠性的要求依然很高，其中大多为金融机构和政府。结合访谈结果来看，相关单位认为随着行业的发展，数据库产品的安全性逐渐得以保证，未来最需要考虑的因素应该是不同产品之间的兼容性以及业务系统和数据的平滑、无损迁移。

（来源：[安全内参](#)）

2、数据要素市场十大研判

关键词：数据要素

上海数据交易所研究院对数据要素市场发展做了十大研判，内容如下：

研判一：数商破圈成为年度关键词。2023 年“数商”将破圈成为年度关键词，数商不仅将成为创新创业的重要赛道和政府支持相关产业的重要门类，也将成为彰显时代特色的代名词。

研判二：数据交易机构分化重组。仅在 2022 年，就有湖南、无锡、福建、郑州、苏州、广州、深圳、杭州等 8 地数据交易机构揭牌成立。2023 年新设数据交易场所将变得困难，既有的数据交易场所将围绕全国统一要素市场目标进行功能分化和结构重组，形成国家-区域-行业-场外的多层次数据要素市场体系。

研判三：场内交易发挥基础设施功能。2023 年将重点投入建设具有基础设施功能的数据交易场所，充分发挥场内交易的规范引导作用。

研判四：数据要素市场的逻辑和规则更为清晰。2023 年数据要素市场的逻辑和规则将更为清晰，各类主体将在更为清晰的市场逻辑、更为明确的流通交易规则和创新容错规则中，让数据交易不再“棘手”。

研判五：数据要素流通标准体系逐步健全。2023 年数据要素相关标准体系将逐步健全，形成分类齐全、层层递进的标准体系。

研判六：数据资产登记和入表走向现实。2023 年数据资产登记和入表将分两步走实现，全国统一的数据资产登记机构将依托数据交易机构建立，在确权登记的前提下推动数据资产入表的第二步，从而为更为广阔的数据资产应用服务提供基础。

研判七：数据产权结构性分置走向落地。2023 年将见证各层级各类制度和规范文件密集出台，推动数据产权结构性分置从理念走向落地。

研判八：公共数据授权运营激活场内交易。2023 年公共数据产品优先乃至应当进场交易将成为各地政策方向，通过公共数据产品激活场内交易，盘活带动整个数据要素市场。

研判九：数据跨境流通打通数字贸易外循环。2023 年数据跨境流通将实现更大发展，与东盟、“一带一路”国家之间的数据跨境流通将成为突破口，打通数字贸易外循环。

研判十：区块链+可信隐私计算赋能数据要素流通。2023年区块链技术将被真正应用于数据交易系统中，可信隐私计算将继续迭代，为数据交易机构实现的更为广泛的数据要素流通提供安全可信的交付环境。

（来源：[安全内参](#)）

3、IDC 发布中国 API 安全市场洞察报告

关键词：IDC API 安全

IDC 认为，API 作为数据流转和使用的重要通道，承载着十分重要的责任。同时，API 的多样性、复杂性在不断增加，传统基于网络和主机边界安全的防护技术无法充分应对云计算和微服务技术下不断弹性部署的业务安全需要，许多用户在攻击事件发生后才意识到 API 风险。因此，API 资产的全面梳理和安全防护成为市场的迫切需求，API 的安全建设也成为企业数字化创新的基础保障。

IDC 将 API 安全定义为专门为保护 API 通信免受误用、滥用和漏洞利用而设计的解决方案，其所提供的功能包括 API 资产发现、验证和执行，动态和自适应的流量监测和模式分析、检测和阻止威胁，例如恶意软件、漏洞利用、代码注入、机器人流量、DDoS 攻击、欺诈和滥用等。目前 API 安全多以 API 安全网关、API 安全管理平台等产品形式进行交付。

IDC 认为，API 的安全防护市场在中国还处于初步发展阶段，产品和技术能力还需进一步加强。目前，国内众多网络安全厂商、数据安全厂商、云计算服务商、专业的 API 安全厂商纷纷布局 API 安全市场，且各个厂商结合自己的技术积累和行业优势推出了各具特色的产品和解决方案。

（来源：[“IDC 咨询”公众号](#)）

4、IDC 发布中国数据安全基础设施管理平台市场洞察报告

关键词：IDC 数据安全基础设施管理平台

IDC 定义下的数据安全基础设施管理平台是一个进行数据安全管理的底层平台，其从数据的发现与分类分级出发，是集成了数据合规治理、数据安全访问治理、敏感数据管理、数据防泄漏、数据加密、数据脱敏等多种数据安全产品能力的统一安全监测、管理、运营平台。该平台作为数据安全的基础设施防护底座，可不断集成并模块化多种数据安全能力。

IDC 认为，目前我国的数据安全基础设施管理平台市场方兴未艾，处于初步发展阶段，众多技术服务提供商已经意识到数据安全能力融合的大趋势，开始将其数据安全能力模块化、原子化，结合平台统一管理优势，帮助用户从数据安全单点建设走向体系化建设。IDC 预测，未来，数据安全基础设施管理平台将逐步发展成为各组织数据安

全建设的基础设施，在最终用户的数据安全建设体系中起到“统一管理、指挥调度”的重要作用。

IDC 结合当前数据安全市场发展现状及未来趋势，为技术买家提供以下几点建议：

- 做好数据安全管理体系建设的顶层设计尤为重要。
- 数据资产发现和分类分级是基础能力。
- 数据安全基础设施管理平台需要与众多品牌的安全产品、能力组件对接联动，需要技术提供商能够通过更加便捷的工具、智能化的流程实现安全信息的汇聚与协同，降低产品对接的复杂度。
- 业务上云已经成为企业数字化转型的重要途经，如果企业计划或者已经拥抱云计算，则必然要考虑数据安全基础设施管理平台与云环境的适配。
- 企业应通过数据安全教育和培训加强员工数据安全保护意识，了解数据安全违规行为需要面临的法律责任，降低内部人为因素造成的数据风险。

(来源：[安全内参](#))

5、2025 年数据安全市场天花板接近千亿元

关键词：数据安全

在我国数据安全防护和数据开发利用并重的数据安全监管格局下，数据安全市场正从传统以数据承载环境为中心的“系统视角”向以数据全生命周期流转为中心的“业务视角”转变，逐步演进为独立的赛道。在此背景下，浙商证券分析认为数据安全赛道市场具有广阔成长空间，2025年市场天花板接近千亿元，2019-2025年复合增长率可达67%。

数据安全监管框架：防护与开发并重，未来2-3年落地可期。2020年6月发布的《数据安全法》（草案）预示着我国进入数据安全战略全面落地时期，受到政策实施、试点开展、以及产业指引三方面因素驱动，我们认为数据安全未来2-3年落地可期。

数据安全市场发展趋势：从“系统视角”到“业务视角”，追求更高回报率。未来数据安全市场有望沿着三个趋势发展：1) **更高的安全投资回报**，即从事后、外挂式的数据安全防护演变至涵盖事前事中事后、与业务紧耦合式的数据安全防护；2) **更合规的数据开发**，即借助隐私计算等技术手段，在保障数据“可用不可见”的前提下实现数据共享；3) **更轻量化的数据安全改造**，即在补齐存量系统数据安全短板时尽量减少对前端业务的影响，降低改造成本和复杂度。

数据安全市场空间：考虑“数据二十条”，中短期186亿元，长期近千亿元空间。若从中国数据总量在全球数据量的占比来看，考虑数据量增加和数据安全渗透率的提升，我们预计2025年数据安全市场潜在空间有望达到820亿元，相较于2019年复合增长率为67%。若从

“数据二十条”中数据要素基础制度带来的潜在增量市场来看，我们预计中短期（数据要素市场探索建设期，预计 2023-2025）市场增量需求约为 **186 亿元**，远期（数据要素市场成熟落地期，预计 2025-2030）市场增量需求接近千亿元。

安全厂商布局：各有侧重，自有产线布局 and 战略投资并行。通过对部分安全厂商在数据安全领域布局的梳理，我们发现：1) 网安厂商在数据安全领域的布局节奏自 2020 年开始显著加快，且布局侧重点大致分为整体解决方案以及聚焦特定行业两类；2) 头部安全厂商大多从自身优势领域切入数据安全赛道进行落地，随后进行扩展布局；3) 除基于自身行业和技术优势发布自有产品及解决方案外，头部厂商还通过战略投资创业公司的方式进行布局。

数据安全投资框架。需求侧，建议短期跟踪政策变化，长期关注数据交易带来的商业模式变化；供给侧，建议关注厂商布局产品结构和产品矩阵协同，产品布局前瞻或行业聚焦深入的公司有望获得更高胜率。

（来源：[东方财富网](#)）

6、2026 年中国数字化转型支出规模预计超过 6000 亿美元

关键词：数字化转型

IDC 最新数据显示，到 2026 年，中国数字化转型支出规模预计超过 6000 亿美元，五年复合增长率将达到 17.9%，增速位于全球前列。



硬件主导中国数字化转型支出，但软件和云部署模式支出增长更快。在中国数字化转型支出中硬件支出在五年预测中占比最大，未来五年接近五成投资份额将流向硬件市场。中国发改委公布的《关于数字经济发展情况的报告》指出我国应适当提前布局数字基础设施，夯实中国数字经济发展的基础。加强数字基础设施布局如传统基础设施数字化、智能化改造，提高应用基础设施水平，系统优化算力等规划已成为硬件支出的重要投资方向。IDC 预计，到 2026 年中国数字化转型硬件市场支出将超过 3000 亿美元。

此外，软件市场增长最高。2021-2026 年软件市场五年复合增长率 (CAGR) 超过 20%，增速超过数字化转型整体增速的 30%。IDC

定义下的软件市场主要由应用开发与部署（Applications Development and Deployment）、应用（Applications）、系统基础架构软件（System Infrastructure Software）等软件市场组成。

（来源：[安全内参](#)）

（二）业界动态

1、微软推出 Security Copilot：内置 GPT-4，自动抵御 65 万亿个网络安全威胁

关键词：Security Copilot

北美时间 3 月 28 日，微软召开首届“Microsoft Secure”大会，并宣布推出网络安全产品——Microsoft Security Copilot。

据悉，Security Copilot 将目前最强大语言模型 GPT-4 内置在产品中，并与微软拥有 65 万亿个网络安全威胁的安全模型库相结合使用，为企业、个人用户提供网络安全、恶意代码防护、隐私合规监控等生成式自动化 AI 服务。

多数传统网络安全软件经常需要数小时甚至数天才能对安全事件做出反应，而企业、个人用户在 Security Copilot 帮助下，可实现“分钟级”安全事件响应、评估、防御，以保证数据资产安全。

（来源：[AIGC 开放社区](#)）

2、全国首个数据交易领域行业数据指数发布平台上线

1月13日，全国首个数据交易领域的行业数据指数发布平台上线，广州数据交易所此次携手广东南方财经控股有限公司、广州航运交易所、中国联合网络通信有限公司广东省分公司、广东电网有限责任公司、广东省交易控股集团有限公司、广州交易集团有限公司、广州人才集团有限公司、全联集采水产品（广东）股份有限公司、北京百度网讯科技有限公司等9家数据商联合发布**80余项行业数据指数**，涵盖财经金融、能源电力、交通旅游、智慧城市、船运船舶、医药健康、知识产权、公共资源、农业水产、人力管理等**10个行业**，将帮助市场了解行业运行情况和**发展态势**，洞悉未来**市场发展趋势**。

（来源：[广东省政务服务数据管理局](#)）

3、全国首个“算力资源专区”正式上线！

关键词：算力 交易所

近日，由贵阳大数据交易所和贵州省算力科技有限责任公司合作共建的“算力资源专区”正式上线。“算力资源专区”以贵州枢纽算力运营调度平台为抓手，旨在推动算力跨地域、跨业务、跨平台集中高效调度，实现在算网资源层面的统一管理、编排和调度。目前，该专区涵盖了通用算力、智能算力、超算算力，目前共接入算力 29.7Pflops，存力 84P，运力 500G，同时还提供配套的安全资源和增值服务。

（来源：[人民网](#)）

4、上海探路数据交易资产化，国内首个数据交易链问世

关键词：数据交易

3月3日，由大数据流通与交易技术国家工程实验室与上海数据交易所正式启动国内首个数据交易链的建设工作，这也是国内数据流通交易领域的新一代基础设施建设项目。

上海数据交易所数据交易系统建立了登记、挂牌、交易、交付、清结算和凭证发放六大业务环节，通过建立数据交易链，利用区块链存证和智能合约等技术使这些业务环节更加安全、高效和透明。区块链技术将数据交易系统的业务环节上链，大大提高数据交易的效率、安全性和透明度，使得交易参与主体互信互认，为各方从登记、确权到交付的交易全过程提供安全保障，体现交易所监管客观公正。

数据交易链的建设目标是构建一个技术自主可控、以平台生态完善为核心的联盟链技术体系，从而在数据产品流通交易中提供登记确权、存证防伪、数据溯源、交易监管等功能，进一步支撑数据安全合规高效流通使用，解决数据产品流通领域中权属确定、可信流通、分布式交易等多维度的难点。数据交易链将基于国内自主可控的开源区块链底链技术、智能合约开发技术、数据隐私保护技术、跨链信息互通技术等先进技术依据统筹设计、分期实施的策略建设而成。

（来源：[个人信息与数据保护实务评论](#)）

5、贵阳大数据交易所上线全国首个数据产品交易价格计算器

关键词：数据交易

为探索多样化、符合数据要素特性的定价模式和价格形成机制。近日，在国家发改委价格监测中心的指导下，贵阳大数据交易所上线全国首个数据产品交易价格计算器。



参考成熟要素市场价格机制，基于《数据产品成本评估指引 1.0》等规范，从价格形成原理出发，结合数据要素特性，该产品通过建立估价模型，以数据产品开发成本为基础，综合考量数据成本、数据质量、隐私含量等多重价值修正因子对于数据产品价格的影响，并基于预计的商业模式和市场规模，评估计算数据产品价格，为数据交易买卖双方议价提供参考，补全“报价—估价—议价”价格形成路径中的关键环节，促进数据要素高效配置、公平交易和自由流动。

(来源：[贵阳大数据交易所](#))

6、苏州市数据出境安全评估申报备案平台上线

关键词：数据出境安全评估

为保障数据要素安全跨境流动,服务苏州数字经济发展,依据《数据出境安全评估办法》和《江苏省数据出境安全评估申报工作指引(第一版)》,苏州市委网信办在“苏商通”门户网站、移动端 APP 推出苏州市数据出境安全评估申报备案平台,为企业“一站式”申报数据出境安全评估开通便捷通道,提供工作指引、申报备案、异议申报、申报咨询、政策动态五类服务。

(来源: [新浪财经](#))

7、中国移动发布“数联网”,保障数据“可用不可见”

关键词：数据流通

4月14日,中国移动发布新型可信数据流通基础设施数联网(DSSN)及相关产品,并与6家数据交易所达成战略合作,探索构建新型可信数据流通基础设施。据悉,数联网(DSSN, Data Switching Service Network)是中国移动依托“连接+算力+能力”信息服务体系,基于隐私计算、区块链、低代码开发等核心技术构建的跨行业、跨区域、跨领域的下一代可信数据流通基础设施,实现了数据“可用不可见”。

(来源: [中国工信新闻网](#))

四、安全事件分析

2023年已经过半，在过去的半年中，以数据泄露、数据加密勒索、数据合规等为代表的数据安全事件正在对全球数据安全形势造成深远影响：

以数据泄露为核心的安全问题日益凸显，其中以个人隐私数据泄露事件和企业商业机密泄露事件居多，对企业利益和个人权益造成严重危害；数据加密勒索攻击无差别地影响着全球各个行业领域，其攻击手段不断升级，复杂性和多样性持续增长；在鼓励数据要素流通的同时，全球合规监管力度也在持续加强，众多数据安全违法处罚事件引发普遍关注，数据保护及合规实现问题的重要性再一次警醒各行各业。

（一）整体态势分析

本部分主要基于大数据协同安全技术国家工程研究中心发布的《全球数据安全观察》中收录的数据安全事件，以数据泄露、勒索攻击、数据合规三种事件类型为维度进行统计和总结分析。

1、类型分析

图1描绘了2023年上半年数据安全事件类型的分布情况，从图中可以看出，数据泄露事件占统计事件总数的64%，在所有类型中占比最高；勒索攻击事件次之，占比为21%；数据合规事件占比为

12%，主要指由于数据违法违规问题而被监管机构处罚的事件；其他类型事件占比为 3%，主要包括数据滥用、人为恶意删除等。

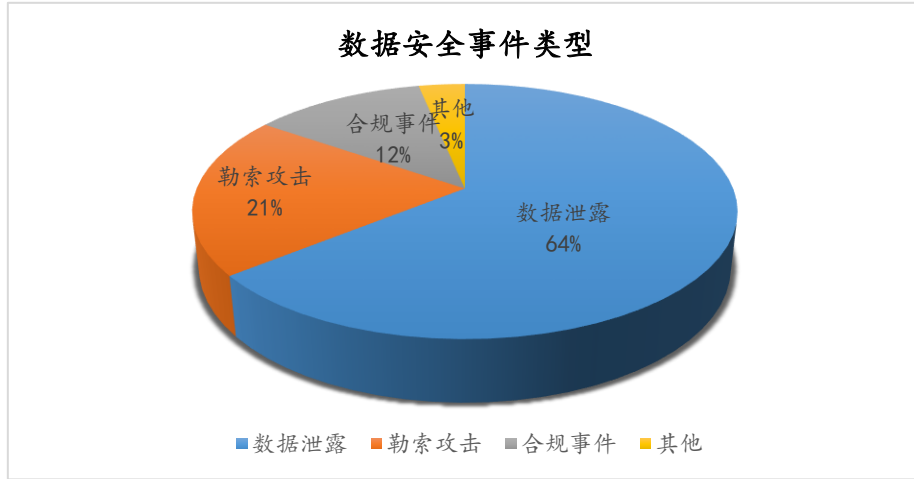


图 1 数据安全事件类型分布

2、行业分布

本报告中对事件行业的划分，主要参考了《国民经济行业分类》（GB/T 4754-2017）。对 2023 年 1-6 月份《全球数据安全观察》中收录的事件所属行业进行统计，结果如下：

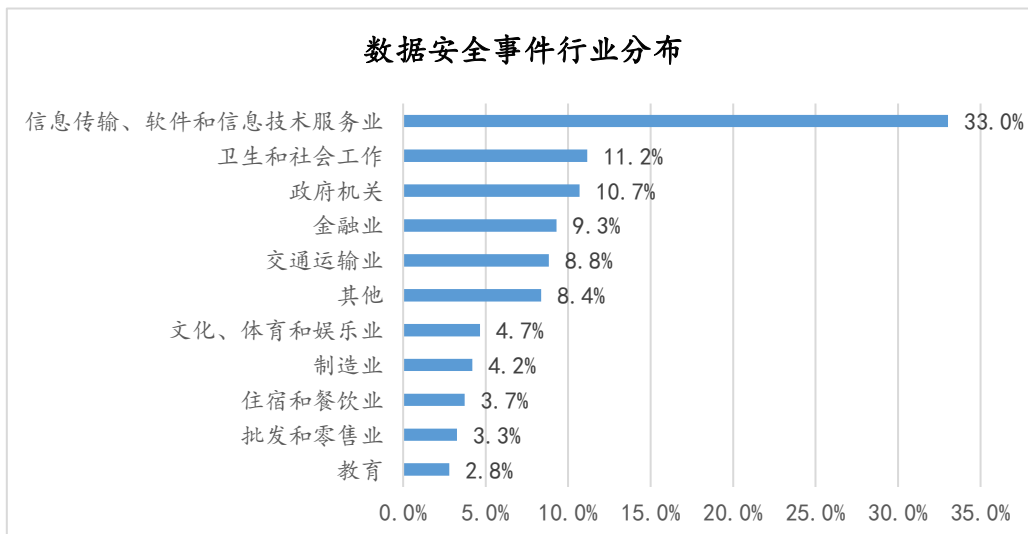


图 2 数据安全事件行业分布情况

可以看出，数据安全事件高发的行业涵盖信息传输、软件和信息技术服务业，卫生和社会工作行业，政府机关等，其中 33.0% 的数据安全事件发生在信息技术服务业，在所有行业中占比最高，这显然与该行业大量接触与处理数据密切相关。

卫生和社会工作行业和政府机关分列二、三位，占比分别为 11.2%、10.7%，可见各类网络基础设施的数据仍然存在较大风险。

金融业、交通运输业也受到了许多网络黑客的觊觎，安全事件发生占比均在 9% 左右；而文化、体育和娱乐业、制造业、住宿和餐饮业、批发和零售业的占比较低，均在 3%-5% 区间；教育业在本次分析中占比最低，为 2.8%。

3、起因分析

深入剖析数据安全事件发生的具体原因（注：此处主要指数据泄露和勒索攻击事件，数据合规事件的原因将在下文单独分析），可将事件起因主要归类为以下几类：外部攻击、缺乏安全措施、安全漏洞、配置错误、第三方合作伙伴泄露、内鬼、内部人员操作失误等。

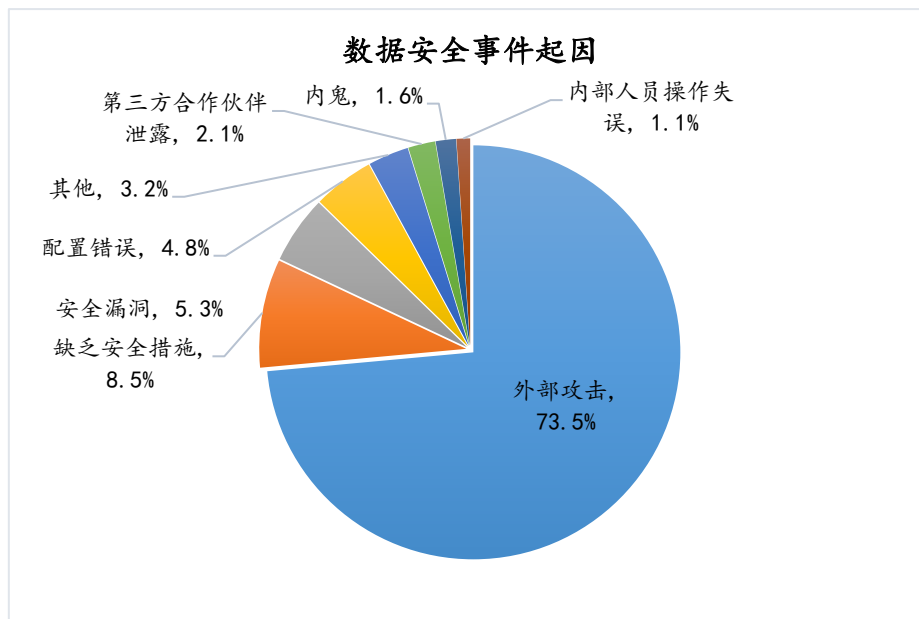


图3 数据安全事件发生原因分布

如图3所示，在本次报告的样本中，由外部攻击导致的数据安全事件占比高达73.5%，遥遥领先其他因素。

缺乏安全措施、安全漏洞和配置错误也是引发数据安全事件的重要原因，在本报告中分别占比8.5%、5.3%、4.8%。其中，缺乏安全措施主要指数据库或服务器未做任何安全措施，例如加密或访问控制等，而导致的数据资源公开暴露；安全漏洞类型主要包括软件本身安全漏洞、安全机制缺陷、数据库漏洞等；配置错误则主要包括服务器配置不当、云存储配置错误等。

由第三方合作伙伴原因导致的安全事件占比2.1%；内部层面，由组织内鬼和内部人员操作失误导致的事件占比较小，分别为1.6%和1.1%。

4、外部攻击画像

外部攻击已成为影响数据安全的头号威胁，本节将从外部攻击中攻击者的身份和攻击方式两个角度对外部攻击事件的格局进行分析。

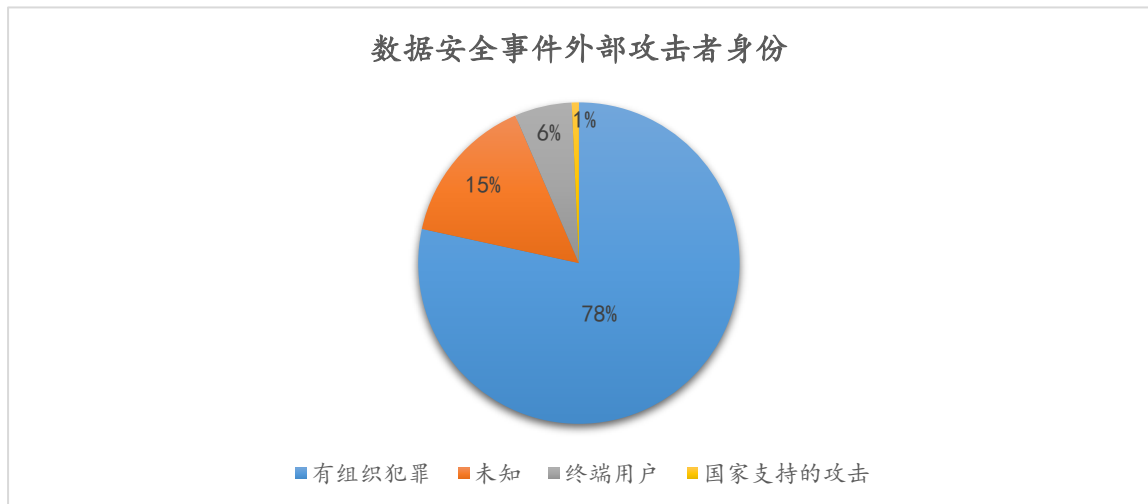


图 4 数据安全事件外部攻击者身份分布

如图 4，从外部攻击者的身份来看，黑客团伙在网络攻击中扮演着最主要的角色，在本次报告中所有由外部攻击引发的事件中，有 78% 的攻击源自有组织的犯罪团伙；另外，15% 的攻击源自未知攻击者，6% 源自终端用户（此处指区别于黑客团伙的个体终端使用者），1% 源自国家支持的攻击。需要注意的是，许多由国家支持的网络攻击事件往往不被公开，因此可能无法被报道和统计。

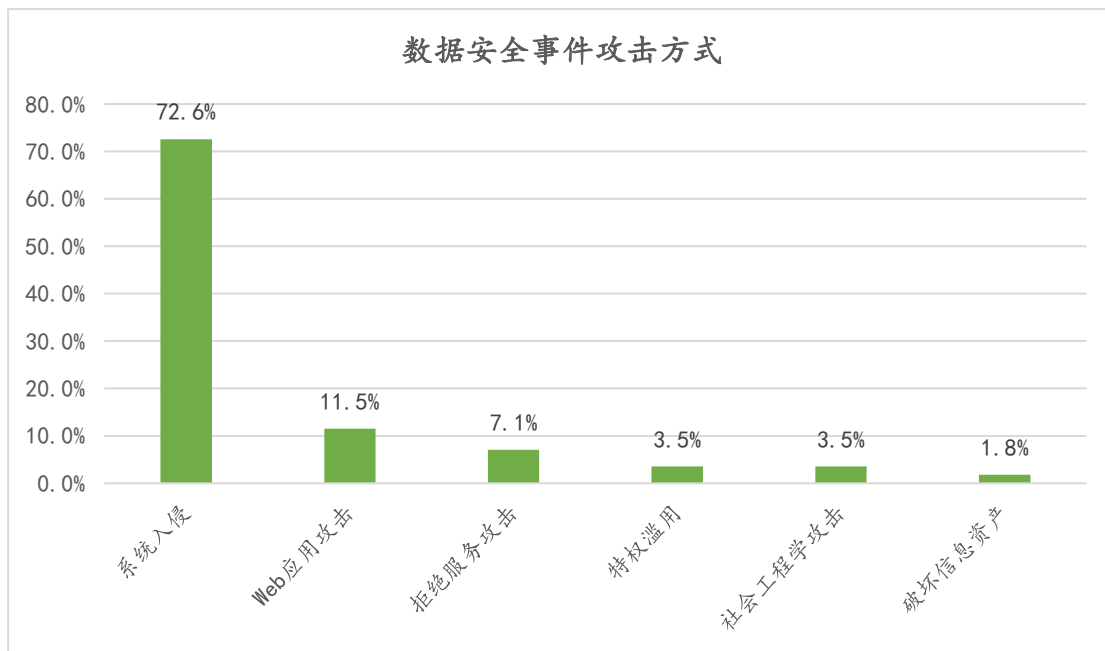


图 5 数据安全事件外部攻击方式分布

（注：排除了数据集中部分攻击方式未知的事件）

如图 5，从外部攻击的方式来看，系统入侵是攻击者最偏爱的攻击手段。本次报告中 72.6% 的外部攻击采取了系统入侵的方式，系统入侵主要指利用恶意软件和/或未经授权访问来实现攻击目标，例如在服务器上部署恶意软件。

Web 应用攻击和拒绝服务也是攻击者常用的攻击手段，分别占据 11.5% 和 7.1% 的比例。其中前者主要针对的是 Web 应用程序，在初始入侵之后，除了窃取数据以外没有大量其他操作；而后者主要包括网络层攻击和应用层攻击，旨在损害网络和系统的可用性。

其他的攻击方式包括特权滥用(由未经批准或恶意使用合法特权引起)和社会工程学攻击(针对个人心理特征促使其做出可能破坏数据安全的行为)，均占比 3.5%；破坏信息资产攻击（恶意删除/转移数据使其不可用）在本次报告中占比最小，为 1.8%。

（二）数据泄露事件分析

从上半年的情况看，数据泄露依然是发生频率最高的数据安全事件类型，本部分将从行业分布、泄露规模、泄露数据类型三个角度对数据泄露事件态势进行分析。

1、行业分布

2023年上半年，数据泄露事件发生最严重的行业为信息传输、软件和信息技术服务业，占比达29.0%。该类型包括互联网和相关服务、软件和信息技术服务、电信、广播电视和卫星传输服务等。行业的数字化特征决定了其生产活动中必然产生海量的数据，如何更安全地存储、传输、使用数据是业内人士需要不断重视和思考的问题。

卫生和社会工作行业、金融业数据泄露也较为严重，在所有类别中分居二、三位，占比分别为12.3%和11.6%。前者主要包括医院、医疗机构以及健康服务中心等组织，其往往包含个人医疗信息等敏感数据，不法分子可能利用这些数据对患者进行诈骗、网络钓鱼等活动以谋取利益；后者主要包括银行、证券、保险、金融、邮政储蓄等，往往涉及银行账户、财产信息等具有较高价值的敏感数据。

交通运输业、政府机关和文化、体育和娱乐业分别以10.1%、8.0%、7.2%的比例紧随其后；本次报告中制造业发生的数据泄露事件频率最低，为1.4%。

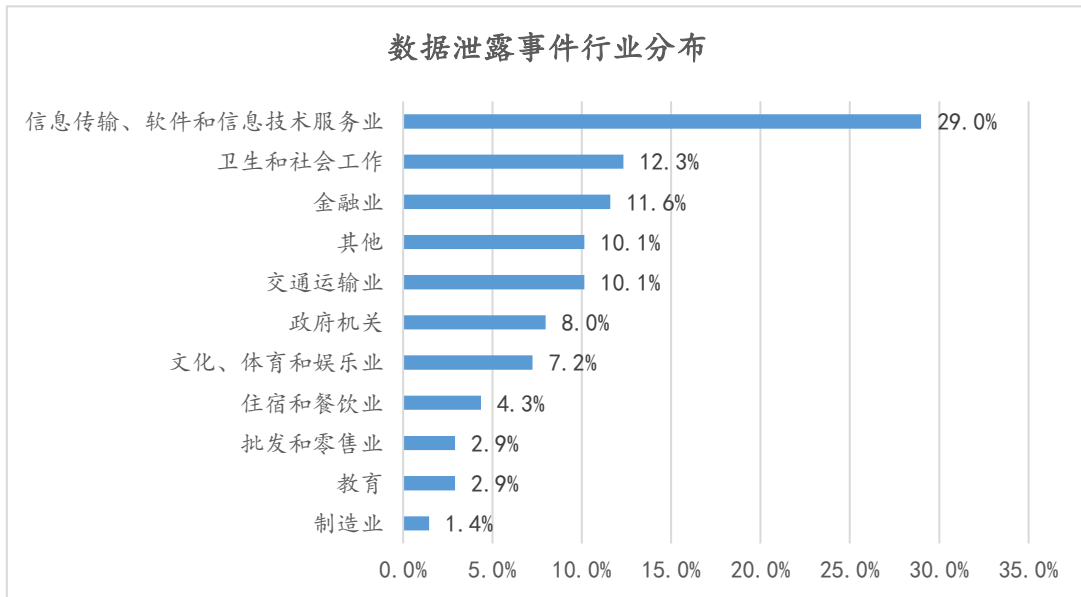


图 6 数据泄露事件行业分布情况

2、泄露规模

如图 7 所示，从数据泄露的条数看，数据泄露条数在 100 万-500 万区间的比例最高，为 31.5%；在 10-100 万区间的次之，占比 27.4%；量级在 1000 万以上的大规模数据泄露事件占比也达到了 16.5%。

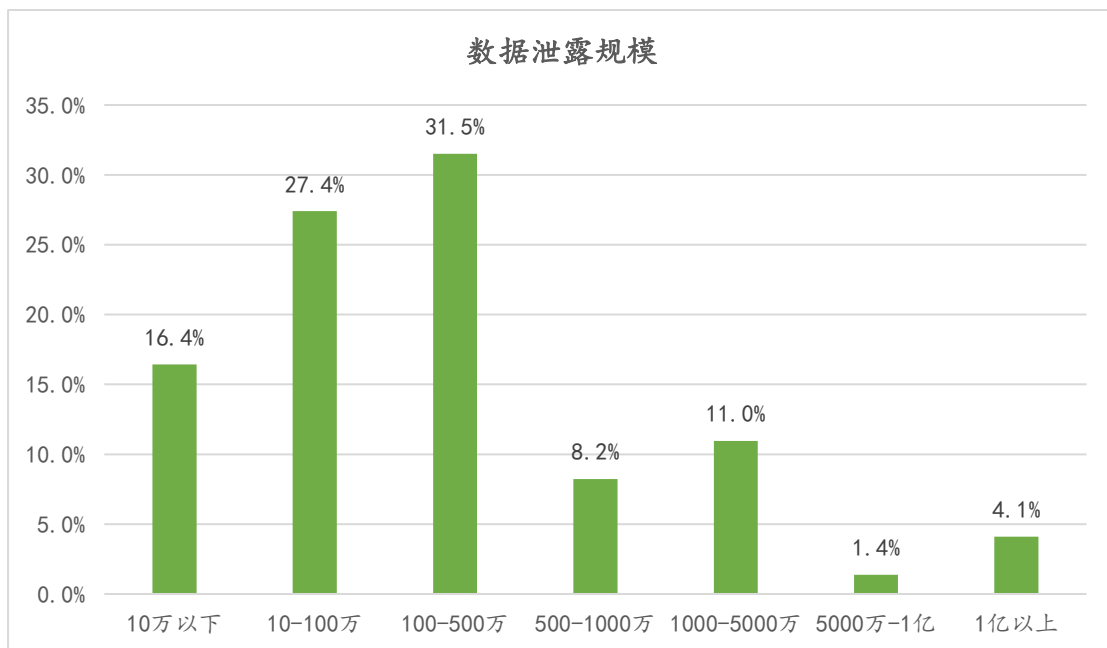


图 7 数据泄露条数占比

如图 8，从数据泄露的量级来看，数据泄露量在 1TB 以上事件占比达到了 12%，在 100GB 以上的事件更是占比过半，达到了 56%。能够造成大规模数据泄露的普遍都是大型企业或行业巨头，应具备较高的安全建设水平，由此可见当前数据安全威胁形势的确愈发严峻。

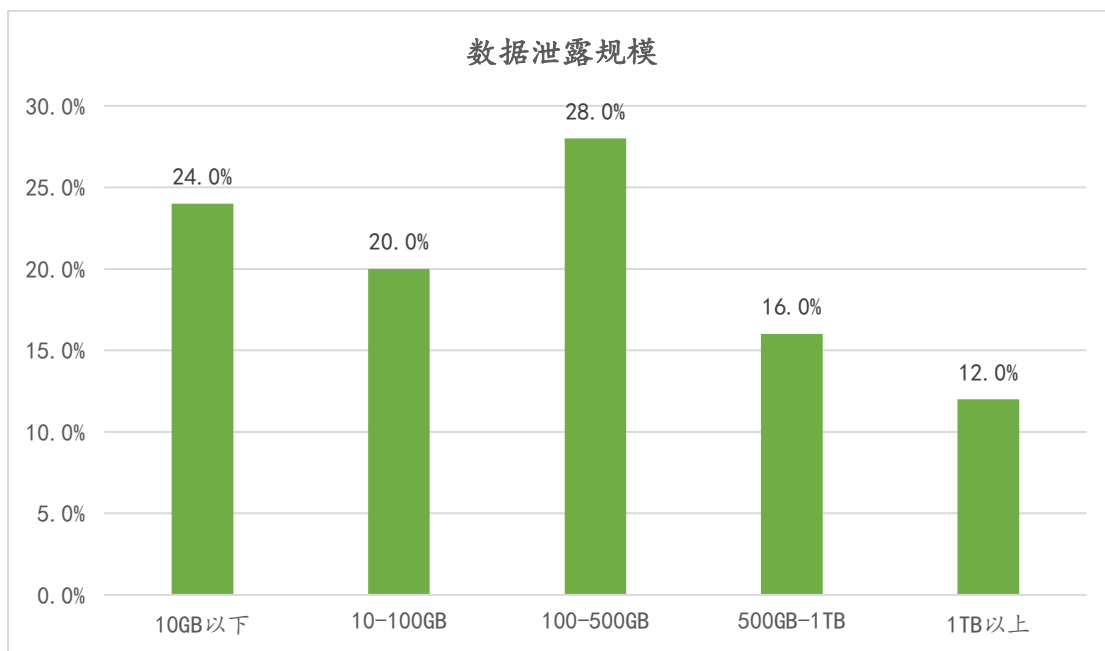


图 8 数据泄露量级占比

3、泄露数据类型

从泄露的数据类型来看，个人信息泄露最为严重，68.1%的数据泄露事件涉及到了个人信息，主要包括姓名、电话、身份信息、电子邮件、地址等；内部资料也是常见的被泄露信息，有 22.5%的泄露事件涉及此类数据，主要包括内部机密、企业商业机密信息、技术机密信息等数据。

医疗信息、凭据信息和银行账户信息属于敏感程度较大的信息，分别有 8.7%、5.8%、4.3%的泄露事件涉及到了这三类信息，其中凭

据信息主要包括账号密码、口令信息、证书及其他证明身份的信息等。

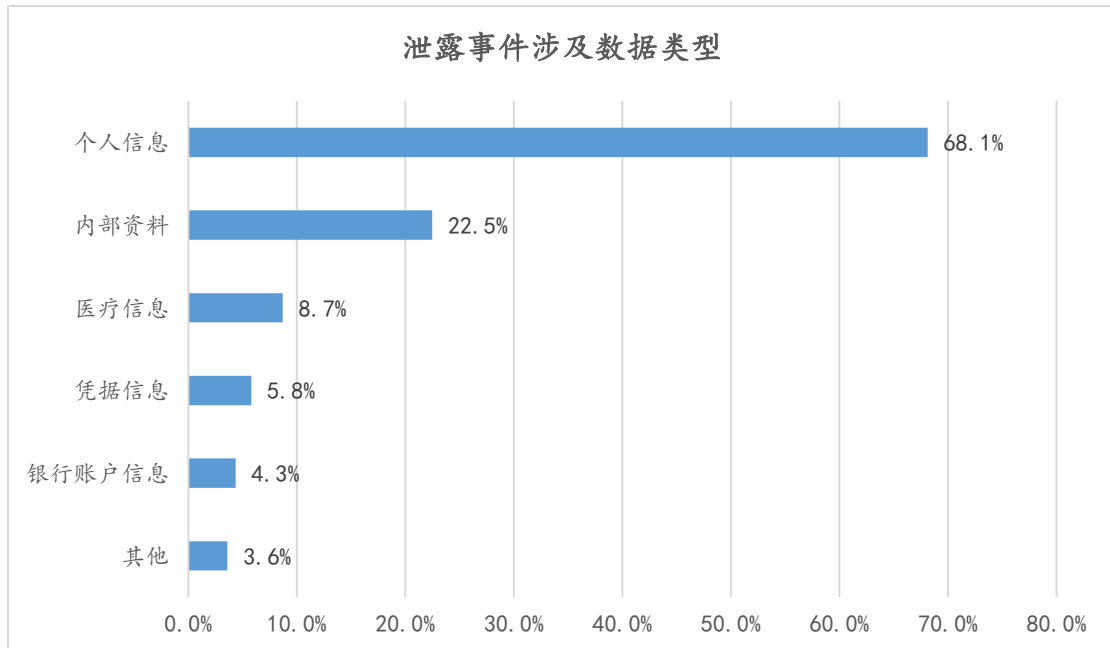


图 9 泄露数据类型分布

(注：部分事件涉及多种类型数据，故总和大于 100%)

(三) 勒索攻击事件分析

勒索攻击又称为“赎金木马”，是指网络攻击者通过对目标数据强行加密，导致企业核心业务停摆，以此要挟受害者支付赎金进行解密。近年来，勒索攻击已成为数据安全领域最具威胁的攻击手段之一，并且呈逐年递增的趋势。针对 2023 年上半年勒索攻击态势，本章从行业分布、赎金要求、活跃勒索多个角度进行了研究分析。

1、行业分布

图 10 是勒索攻击事件所属行业的分布情况：由图可知，受到勒索攻击影响最大的是信息技术服务业，占比达到 25%，该行业往往需

要直接或间接向数量庞大的用户群体提供服务，一家软件公司开发的服务可能被多家大型企业同时使用，因此成为勒索团伙的重点目标；政府机关以 22.7% 的占比在遭受攻击次数中排名第二，政府机构所拥有的内部资料和公民/职员信息对勒索团伙有着高度吸引力；制造业发生的勒索事件占比 13.6%，在所有行业中排在第三位，由于制造业产业链的特殊性，该行业遭遇勒索攻击时，很可能面临生产线停摆、运营中断等问题，恢复生产需要的资源和成本远高于其他行业，因此容易招致期望牟取暴利的勒索团伙；针对医疗卫生行业和交通运输业的攻击分别占据了 11.4% 和 9.1% 的比例，其中后者相比往年明显增多，勒索组织正在更多地渗透航空公司、铁路公司等关乎旅客出行的公共交通事业。

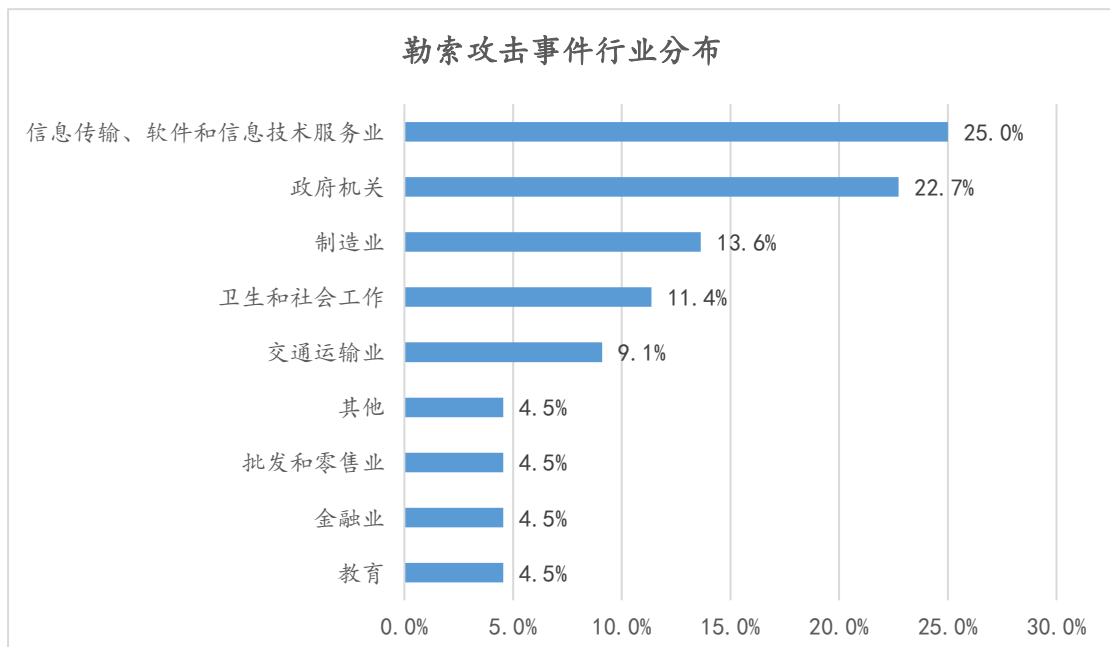


图 10 勒索攻击事件行业分布情况

2、勒索金额

勒索金额方面，如图 11 所示，勒索金额在 10 万美元以下和 100 万-500 万美元之间的事件在所调查的勒索数额中占比最高，均为 29.4%；索要 500 万美元以上大额赎金的事件总占比达到了 17.7%。

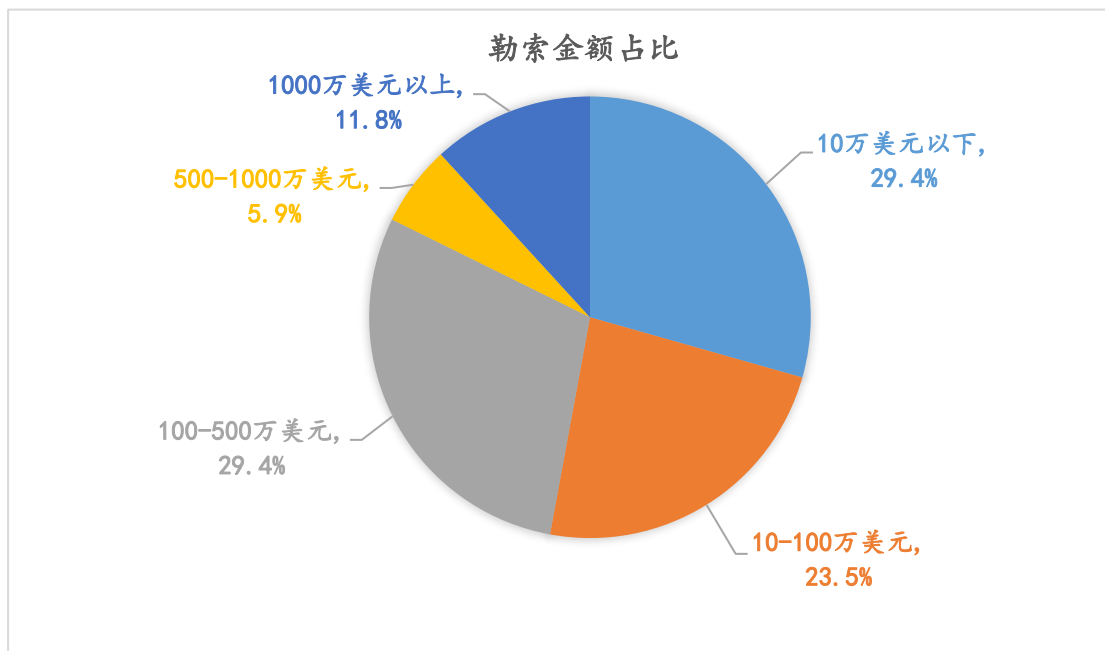


图 11 勒索金额情况统计

由于勒索组织往往依据企业规模来确定勒索金额，大型企业遭受恶意攻击后一般会被勒索更高额度的赎金，从而拉高了整体赎金规模水平。另外，由于多数被勒索的企业机构拒绝披露勒索金额，实际数据很可能比本报告记录的样本数据更高。

3、活跃的勒索组织

在本报告统计的勒索攻击事件中，攻击活动最为频繁的是 AlphV/BlackCat 组织，占勒索攻击事件的 13.6%；LockBit 和 Play 组

织紧随其后，以 11.4% 的占比并列排在活跃度第二位；Vice Society 排名第四，占比达 6.8%。

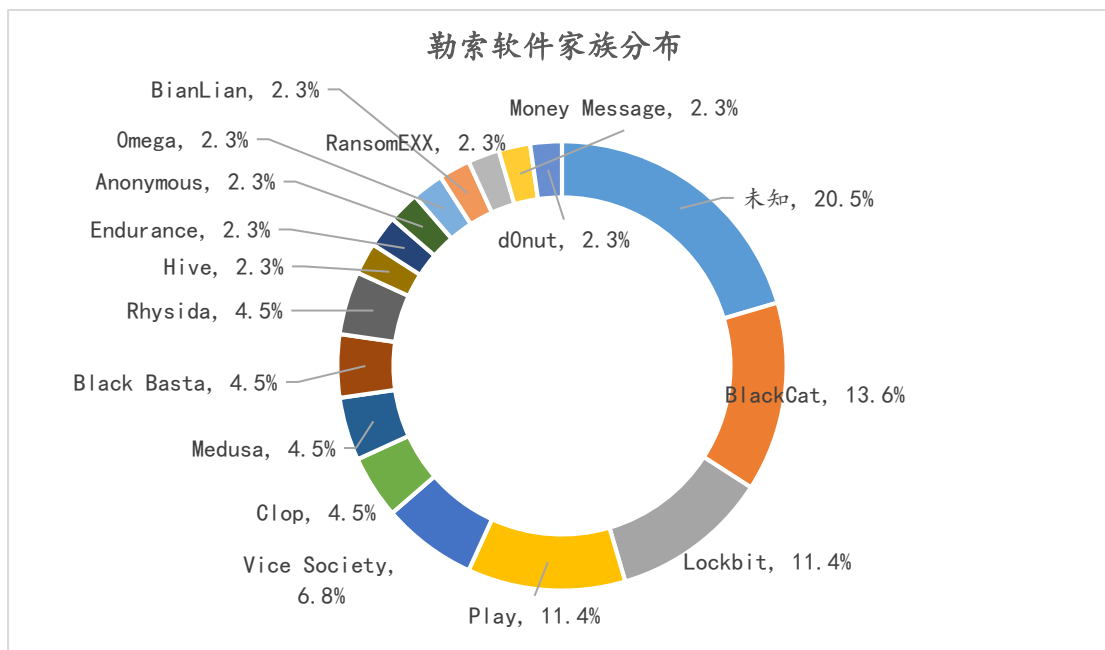


图 12 勒索软件家族情况统计

(四) 数据合规事件分析

数据安全监管力度的逐步加大对企业数据合规提出了更高要求，企业需要及时调整数据合规思路。下文将从行业分布、处罚金额、违规原因三个角度梳理分析上半年数据合规事件的格局。

1、行业分布

图 13 的统计数据显示，绝大多数的数据安全合规事件集中在信息传输、软件和信息技术服务业，占比高达 69.2%。这主要是因为信息技术相关行业的主体一般需要为庞大的用户群体提供各类复杂业

务,其中往往涉及处理各类个人敏感信息,面对日趋严格的国家监管,操作稍有不范就可能触碰数据安全的“红线”。

住宿与餐饮业、金融业、卫生与社会服务业以 7.7%的占比排在并列第二位,住宿行业常常需要收集客户的个人身份信息以提供更适合用户需求的服务;金融业在管理用户理财产品和资产情况时也往往需要涉及处理敏感信息;医疗行业更是负有保护患者隐私的重大责任。这几类行业都伴随着较高的数据违规风险,随着各国政府在司法和行政监管方面针对个人信息保护的重视程度和处罚力度不断提升,多家相关企业因个人信息保护问题被监管机构施以重拳。

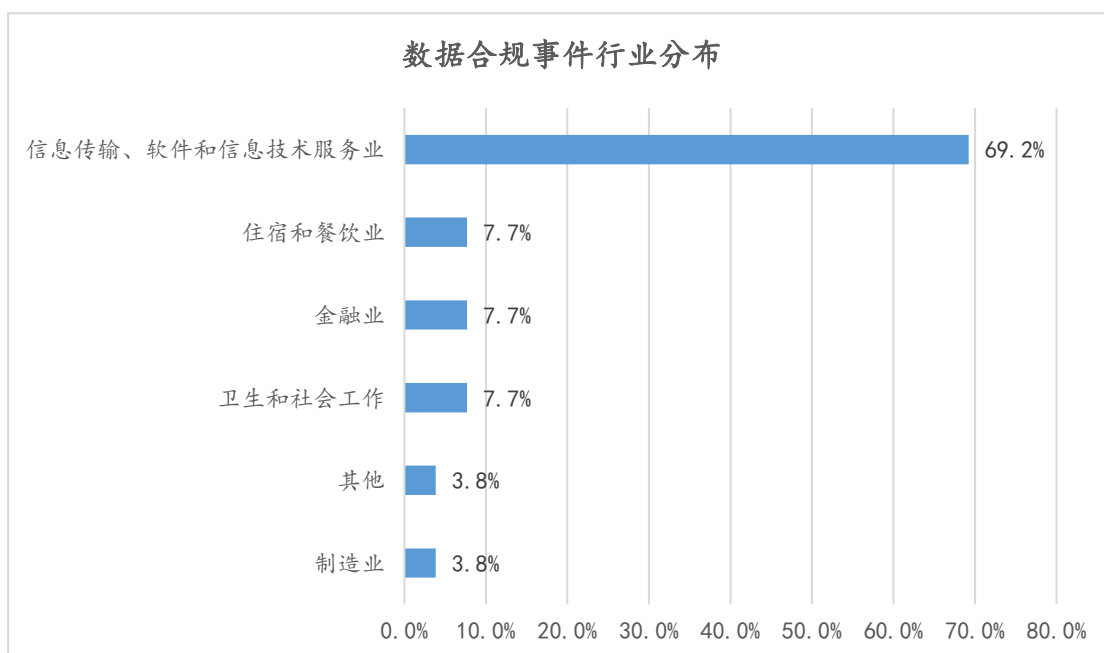


图 13 数据合规事件行业分布情况

2、处罚金额

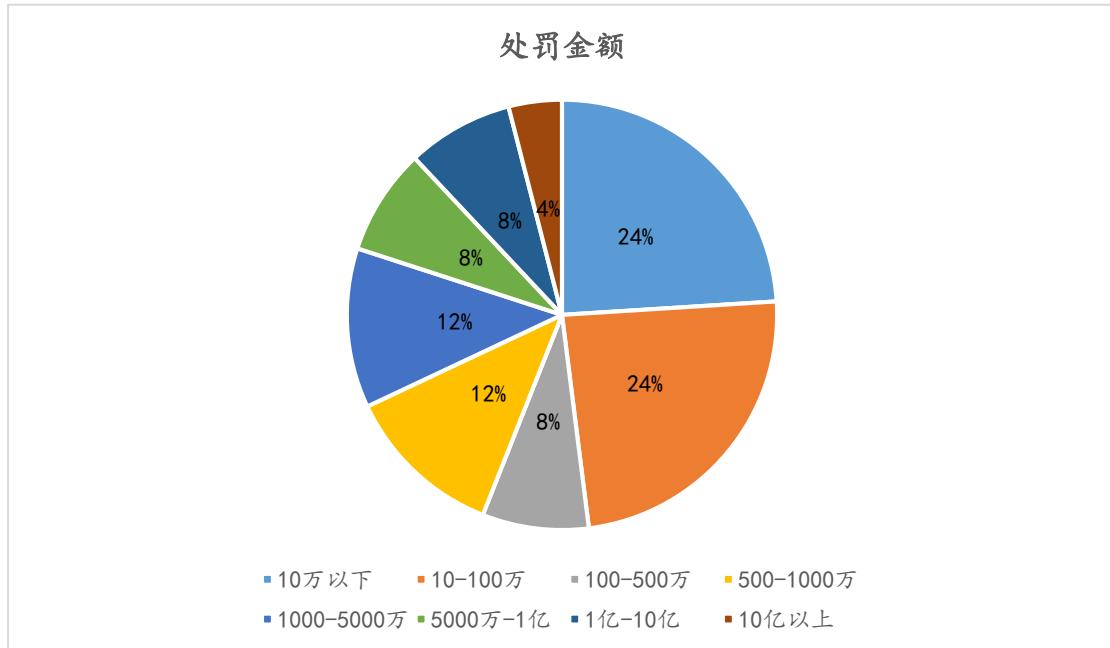


图 14 数据安全合规事件罚金统计

从图 14 已披露的罚款金额来看，罚金在各个区间的分布比较平均，100 万以下的事件占 48%，但 1 亿元以上的大额罚款事件占据了 12% 的比例。对许多企业来说，巨额罚金所带来的影响不仅局限于经济效益，还有负面舆情所带来的附加效应。

3、违规原因

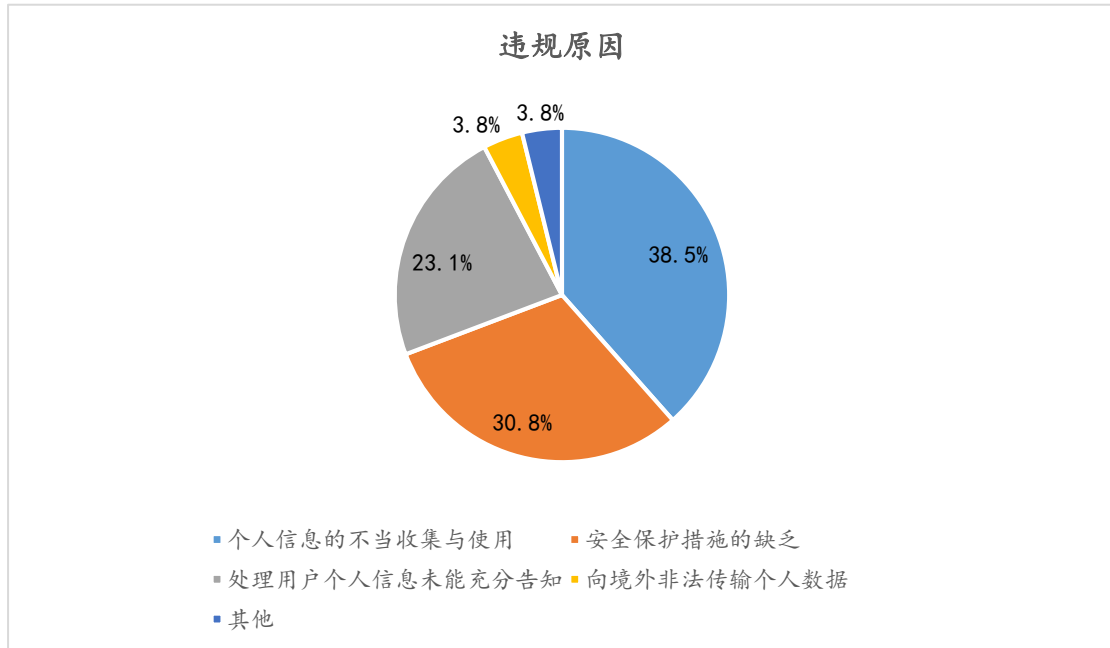


图 15 数据安全合规事件违规原因统计

如图 15，本报告涉及的数据安全合规事件中，违规原因主要包括：个人信息的不当收集与使用（38.5%）、安全保护措施缺乏（30.8%）、处理用户个人信息未能充分告知（23.1%）、向境外非法传输个人数据（3.8%）等。

可以看出，处罚多集中于企业对于用户数据的违规处理，“个人信息的不当收集与使用”“处理用户个人信息未能充分告知”两项数据之和便占据了所有数据合规事件的 62.6%，这也体现了各国政府在执行数据监管工作中的重点。未能对敏感数据进行恰当保护也是企业遭到处罚的重要原因，相关企业应加强责任意识 and 主体意识，主动搭建好数据安全的技术防线。

（五）总结

2023 年上半年，数据安全形势依旧严峻，数据泄露事件最为频繁，大部分数据安全事件可归咎于外部攻击，多数攻击来自有组织成体系的黑客团伙，系统入侵、Web 应用攻击等传统攻击手段占据主导地位，为数据使用者敲响了构建完善防御体系的警钟。

数据泄露事件频发，个人信息首当其冲，内部资料紧随其后，可见在处理泄露风险较大的数据类型时应当加倍谨慎；勒索攻击方面，BlackCat 和 LockBit 等勒索组织最为活跃，各行各业应未雨绸缪，做好对勒索软件特别是活跃勒索软件的防御工作，避免经营活动受到影响；数据合规方面，多数处罚源自于违规收集或处理用户信息，其中不乏各类高额罚单，警示相关企业需戒除侥幸心理，怀履深渊、临薄冰之心。

总体来看，信息技术服务业仍然是各类数据安全事件的重灾区，从业人士需要加强数据安全意识和合规意识，通过搭建体系的管理机制和有效的防护措施来保障企业运营过程中数据全生命周期的安全与处理合规。

五、工程中心研究

《大语言模型提示注入攻击安全风险分析报告》

伴随着大模型等人工智能产业的快速发展，新的安全风险挑战也接踵而至。大数据协同安全技术国家工程研究中心 AI 安全实验室发布国内首份《大语言模型提示注入攻击安全风险分析报告》，为国内大模型安全发展提供整体指南。报告指出，提示注入攻击已成大模型安全威胁之首，建议从安全测评、安全防御、安全监测预警等方面，多维度提升大模型的安全性。报告引言如下：

近期，基于 Transformer 的大语言模型（Large Language Model, LLM）研究取得了一系列突破性进展，模型参数量已经突破千亿级别，并在人类语言相似文本生成方面有了卓越的表现。目前已有多个商业化大模型发布，如 OpenAI 推出的 GPT 系列、Google 推出的 T5 和 PaLM，以及 Meta 推出的 OPT 等大语言模型等。特别是 OpenAI 推出的 ChatGPT，由于其强大的理解与生成能力，在短短 2 个月内就突破 1 亿用户量，成为史上用户增长速度最快的消费级应用程序。为了应对市场冲击，谷歌也推出了 BARD 聊天机器人，Meta 则开源了 LLaMA 模型。国内各大企业、高校和研究机构也纷纷进入大模型领域，推出了一系列对话大模型，包括百度文心一言、360 智脑、讯飞星火、商汤商量、阿里通义千问、智源悟道、复旦 MOSS、清华 ChatGLM 等。

大语言模型正在各个应用领域引起巨大的变革，并已经在搜索、金融、办公、安全、教育、游戏、电商、社交媒体等领域迅速普及和应用。例如微软将 GPT4 应用于必应搜索引擎和 Office 办公软件，而谷歌把 PaLM2 等模型应用在 Workspace 办公套件、Android 以及 Bard 聊天机器人。

然而，伴随着大语言模型迅速普及和应用的同时，也衍生出一系列严重的安全风险，并引发了多起安全事件。如 OpenAI 曾经默认将用户输入的内容用于模型训练，从而导致了多起隐私数据泄漏事件。据媒体报道，亚马逊公司发现 ChatGPT 生成的内容中发现与公司机密非常相似的文本。韩国媒体报道称，三星公司在引入 ChatGPT 不到 20 天内就发生三起涉及机密数据泄漏的事故，其中两起与半导体设备有关，一起与会议内容有关。据网络安全公司 Cyberhaven 的调查，至少有 4% 的员工将企业敏感数据输入 ChatGPT，而敏感数据占输入内容的 11%。

此外，大模型系统近期也被相继爆出多个安全漏洞。例如，ChatGPT 的 Redis 客户端开源库的一个错误，导致 1.2% 的 ChatGPT 付费用户个人信息泄露，包括聊天记录、姓名、电子邮箱和支付地址等敏感信息。随后，OpenAI 网站又被爆出 Web 缓存欺骗漏洞，攻击者可以接管他人的账户，查看账户聊天记录并访问账单信息，而被攻击者察觉不到。360 AI 安全实验室近期还发现大模型软件 LangChain 存在任意代码执行的严重漏洞。

总体而言，目前大语言模型面临的风险类型包括提示注入攻击、对抗攻击、后门攻击、数据污染、软件漏洞、隐私滥用等，这些风险可能导致生成不良有害内容、泄露隐私数据、任意代码执行等危害。在这些安全威胁中，恶意用户利用有害提示覆盖大语言模型的原始指令实现的提示注入攻击，具有极高的危害性，最近也被 OWASP 列为大语言模型十大安全威胁之首。

报告对面向大语言模型的提示注入攻击和防御技术展开研究，并通过构建数据集对大语言模型的提示注入攻击安全风险进行了测评。首先，系统分析了面向大语言模型的提示注入攻击和防御技术，并验证了相关技术的有效性。在提示注入攻击方面，对直接注入攻击和间接注入攻击两种方式进行了分类，涉及目标劫持攻击、提示泄露攻击、越狱攻击等。在提示注入攻击防御方面，从大语言模型输入侧、输出侧两端对相关技术进行分析，涉及提示过滤、提示增强等。其次，构建了包含 36000 条的提示注入攻击验证数据的数据集，覆盖了 3 类典型攻击方法和 6 类安全场景，用于对大语言模型的提示注入攻击风险测评。然后，对 OpenAI GPT-3.5-turbo、谷歌 PaLM2 以及 UC Berkeley 等高校团队开源的 Vicuna-13B 共 3 个典型的大语言模型进行了测评，测评结果显示，本报告中构造的数据集能分别以 79.54%、77.41%、67.24% 的成功率实现 3 类模型的攻击。这 3 类大语言模型一定程度上代表了目前商业和开源大语言模型的最先进水平，因此测评结果具有代表性。最后，对本报告工作进行总结，并对未来工作进行了展望，

提出应该加强提示注入攻击防御技术研究，提高对攻击的检测能力和效率。

报告可以为大语言模型厂商、相关的开发者以及研究人员提供参考，以构建更加安全可信的大语言模型。另外，基于报告形成的测评能力，大数据协同安全技术国家工程研究中心 AI 安全实验室将通过“安全大脑国家新一代人工智能开放创新平台”对外提供大语言模型提示注入攻击风险安全测评服务。

（完整报告获取链接：nelab-bdst.org.cn）

《数字安全观察》产品系列

- 每周动态：政策法规/行业动向/安全事件/技术趋势
- 深度分析：政策解读/行业洞察/市场预测/事件分析
/技术前瞻/策略建议/国际智库精编
- 国防专刊：网空战略/力量建设/科装动态/空天态势
- 数据安全专刊：政策形势/安全事件/安全研究/大咖观点

总编辑：杜跃进

执行编辑：张义荣、钟力

本期编委：唐会芳、王雨薇、陈璐、徐家骏、赵易初、李昊、徐明启

如有反馈 邮件请至 nelab@360.cn、dipperresearch@360.cn

