安全+
SECURITY
2025/10 总第 066
技术版 ▶▶ 与安全人士分享技术心得 Share technique experience with security professionals

初探智能制造企业工控安全的体系化建设

国际数据空间的规范机制与理论模型

从“芯”构建信任之钥，开启可信数据空间价值共创之门

网络安全政策导读（热点追踪）

NSFOCUS

本期看点 HEADLINES

主办：绿盟科技
策划：《安全+》编委会
地址：北京市海淀区北洼路4号院绿盟科技园
邮编：100089
电话：(010)6843 8880-5462
网址：www.nsfocus.com

2025/10 总第 066

欢迎您来信nsmagazine@nsfocus.com 与我们交流，分享您的建议和评论。（《安全+》部分图片来源于网络）

安全+ SECURITY+

目录 CONTENTS

卷首语

人工智能正加速成为推动经济社会深刻变革的核心力量。国务院近期印发的《关于深入实施“人工智能+”行动的意见》，围绕科学技术、产业发展、治理能力等重点领域作出全面部署。文件提出，要培育人工智能应用服务商，发展“模型即服务”“智能体即服务”，加快重点领域标准研制和应用场景建设；统筹智能算力布局，支持芯片创新与超大规模智算集群落地；持续加强高质量数据集建设，健全产权和版权制度，推动数据服务产业发展。这些举措为人工智能与网络安全的深度融合创造了新的战略机遇。

置身这一浪潮，绿盟科技以“AI+安全”为锚点，将智能安全能力延伸至卫星通信、无人机、无人驾驶、车联网及AI攻防对抗等新兴领域。2025年针对大模型应用热潮，绿盟科技发布十余款新品，并同步推出大模型安全综合治理方案、场景安全方案、备案评估服务、红队评估服务以及AI智能化渗透系统，全面守护用户智能化转型的可信与安全。

在产品与服务演进上，绿盟科技不断深化升级迭代，在传统产品中持续整合AI能力，并逐步向智能体方向发展。在AI赋能安全运营、数据安全、供应链安全等关键场景中，技术和服务能力已实现全面提升。同时，积极推进安全行业内的合作与能力打通，推动跨领域认证，适配多种国产化硬件和操作系统，确保AI安全相关能力具备良好的弹性、兼容性与包容性。

本期《安全+》将聚焦人工智能与安全融合的前沿探索，同时关注工控安全、可信数据空间建设与政策解读等内容，旨在为读者提供趋势洞察与实践参考，支持在真实业务中更安全、更高效地应用智能技术。

叶晓虎

绿盟科技集团首席技术官

初探智能制造企业工控安全的体系化建设

绿盟科技 工程线 尹亮 郭涛 尹文娟 绿盟科技 湖北代表处 殷陆军

摘要：文章系统研究了智能制造环境下工业控制系统安全体系化建设的理论与实践路径。面对工控安全建设缺乏体系性这一难题，深入分析智能制造环境特征及其带来的工控安全新挑战，提出了包含管理维度、技术维度和运营维度三个维度的体系框架。在具体实践中，从安全治理架构域、安全制度体系域、人员安全能力建设域、工控网络架构域、工控边界防护域、工控主机防护域、工控监测审计域、工控安全管理域、常态运维域、应急响应域十个领域开展建设。企业通过构建工控安全“三维十域”的防护体系，保障生产的安全稳定运行。研究表明，该框架能够有效应对智能制造环境下的工控安全风险，提升企业整体安全防护能力。集团型企业通过体系化建设可使安全事件发生率显著降低，应急响应效率大幅提升。文章为智能制造企业工控安全建设提供了系统性的解决方案和实施路径。

关键词：智能制造 工业控制系统 工控安全 体系化建设 安全框架

引言

工业控制系统作为智能制造的核心基础设施，其安全性直接关系到生产运行的稳定性和连续性。然而，生产制造环境下工控系统面临的安全威胁日益增多，传统的单点防护模式已经难以满足安全需求。

当前，智能制造企业工控安全建设普遍面临三大难题：一是安全管理责任难以落实，没有形成齐抓共管的机制；二是安全技术措施碎片化，缺乏系统性的设计；三是安全运营能力不足，应急响应流程运行不畅。这些问题严重制约了企业工控安全防护水平的提升。因此，建立系统化、体系化的工控安全防护框架，已成为智能制造企业亟待解决的关键问题。

本文基于系统工程理论，结合智能制造特点，提出了包含管理、技术、运营三个维度的工控安全体系框架，探索了智能制造企业工控安全体系化建设的有效路径。

1. 智能制造业务场景概述

典型的生产或制造系统从上到下可划分为外部网络、企业资源层、生产管理层、过程监控层、现场控制层和现场设备层，如图 1 所示。

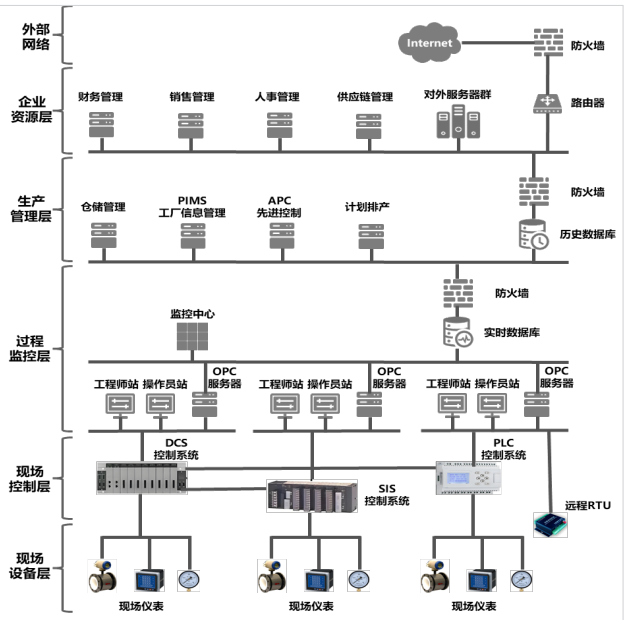


图 1 典型智能制造企业信息化架构图

由图 1 可以看出，以生产管理层为界，向上为通用信息化领域，向下为工业控制系统领域 [1]。生产管理层包括工厂信息管理系 统 (Production Information Management System，PIMS)、先进控制系统 (Advanced Process Control，APC)、历史数据库、计划排产、仓储管理等。过程监控层包括数据采集与监控系统 (Supervisory Control And Data Acquisition，SCADA)、分散型控制系统 (Distributed Control System，DCS)、安全仪表控制系统 (Safety Instrumented System，SIS)、可编程逻辑控制系统 (Programmable Logic Controller，PLC) 的工程师站、操作站、用于过程控制的对象链接与嵌入 (OLE for Process Control，OPC) 服务器、实时数据库、监控中心等 [2]。现场控制层包括数据采集与监控系统 (SCADA)、分散型控制系统 (DCS)、安全仪表控制系统 (SIS)、可编程逻辑控制系统 (PLC) 的控制器。现场设备层包括压力、流量、液位、电量、位移等仪表，以及控制阀、电动阀等执行机构 [3]。

智能制造环境下的业务场景呈现出四点特征。首先，系统架构呈现深度互联特性。工业互联网的快速发展使工控系统与企业管理系统、供应链系统的集成程度不断提高。这种深度互联在提升生产效率的同时，也使攻击路径呈现几何级数增长。其次，数据要素实现互通。在智能制造场景下，数据已成为核心生产要素。从设备传感器采集的实时数据到企业管理系统的决策数据，形成了复杂的数据流动网络，安全管控难度显著增加。再次，技术体系呈现异构融合特点。智能制造企业普遍存在新旧设备共存的状况，这些设备在操作系统、通信协议、计算资源等方面存在显著差异，给统一的安全策略实施带来巨大挑战。最后,威胁态势持续演进升级。攻击手段从传统的病毒、蠕虫发展为更具针对性的高级可持续威胁 (Advanced Persistent Threat，APT)，攻击者越来越多地采用供应链渗透、零日漏洞利用等高级技术 [4]。

2. 智能制造安全风险分析

2.1 工控安全职责不清晰

2024 年 1 月工业和信息化部印发的《工业控制系统网络安全

防护指南》指出：工业企业承担本企业工控安全主体责任，建立工控安全管理制度，明确责任人和责任部门，按照“谁运营谁负责、谁主管谁负责”的原则落实工控安全保护责任。工控安全管理主要由信息化部门主导，面临着众多的合规要求。工控安全工作在落地过程中需要多个业务部门配合，但业务部门往往面临着较大的生产压力，工控安全管理职责难以进行分解。在实际的工业控制系统安全管理实践中，若沿用传统信息安全管理的模式，鉴于工控安全的特殊性，易出现适应性不足的问题，导致安全管理要求难以有效落实。同时，工控安全的工作对人员能力有着较高的要求，工控安全专业人才缺口显著，现有人员技能单一化往往与业务场景复杂化的矛盾凸显。

2.2 工控安全防护存在缺失

在智能制造场景下的工业设备、控制系统及安全产品，往往来自不同厂商，通信协议、数据接口缺乏统一标准，导致安全防护难以同步实施。在边界防护方面，生产系统逐步互联互通，使生产网络横向和纵向的边界模糊，网络安全风险一旦外溢到相邻网络，可能影响生产控制系统的连续稳定运行。在主机防护方面，工业主机往往存在操作系统版本老、漏洞多、基线弱、USB 端口管控缺失、对恶意代码防范能力差等问题。同时，缺乏对工控主机安全软件的保护，容易发生工控安全事件。在安全检测方面，生产网络内可能缺少对威胁入侵检测以及异常行为、网络运行状态的监测。工控安全部门对生产网络中存在的威胁行为、漏洞利用行为、误操作以及违规操作，不能实时感知。工业现场控制操作频繁，操作人员对现场设备下发的指令以及设备如果对接收到的数据缺乏认证手段，相关指令和数据即存在被伪造、篡改、重放的风险。

2.3 工控安全运营碎片化

智能制造企业往往对系统缺乏常态化安全管控机制，可能滋生出多重隐患。在工控资产管理方面，企业普遍存在底数不清的问题。工控资产作为安全防护的核心对象，在工业现场往往分散部署，导致难以精准掌握其数量、类型、位置及运行状态。在监测预警方面，企业普遍缺乏一体化安全监控平台。工控安全部门难以在全局上掌握相关资产的脆弱性是如何分布的、面临的安全威胁有哪些，也难以第一时间判断应该优先对哪些脆弱性进行修补，对哪些威胁进行分析。在应急响应方面，企业普遍存在着流程不畅各自为政的问题。当工控安全事件发生时，不同部门、不同岗位之间缺乏协同的应急响应流程，各部门只能基于局部信息做出决策，导致应急处置呈现碎片化状态。企业跨部门协同机制的缺失，不仅延误了最佳的应对时机，还可能致使事件的影响范围进一步扩大。

3. 工控安全体系化建设基本原则

3.1 顶层设计原则

工控安全管理遵循顶层设计原则，企业需将其纳入战略规划的范畴，由决策层统筹领导，建立跨部门联动机制，确保安全战略与生产业务同步推进。制度保障方面，工控安全部门需构建覆盖研发、生产、运维全链条的安全管理制度体系，细化设备准入、访

问控制、风险评估等技术标准与操作规程，并通过定期评审与动态更新机制保持制度的有效性。责任明确层面，人力资源管理采用业务条线与职能层级双维度划分模式，通过岗位健康、安全与环境管理体系（Health Safety and Enviroment，HSE）手册固化职责边界，将工控安全纳入生产安全考核，形成职责清晰、过程可溯、闭环管理的责任制体系。

3.2 纵深防御原则

工控安全防护遵循纵深防御原则，通过构建边界防护、区域划分、终端安全及生产业务隔离的多层技术体系，形成纵深防御架构。生产网络与办公网络使用工业网闸来实现数据摆渡，构筑第一道防线；同时在生产网内的生产管理层与过程监控层之间部署工业防火墙、入侵检测系统，构筑第二道防线；并在各个生产线内部分区部署工业防火墙、工控安全审计，构筑第三道防线。各防护层通过安全态势感知平台动态联动，结合威胁情报实现攻击路径可视化追踪。同时辅以工业主机防护、漏洞管理及安全审计等主动防御措施，形成“监测—预警—处置”的闭环流程，确保单一防线突破后仍存在多重措施，提升工控系统的业务连续性保障水平。

3.3 持续监控原则

工控安全运营遵循持续监控原则，要求构建多层级、智能化的工控安全监测体系。工控安全部门通过部署工业协议深度解析、设备指纹特征识别及流量探针等技术手段，实现对 PLC、DCS 等控制设备的常态化监测。工控安全运营中心依托安全态势感知平台开展动态风险监测，结合工业威胁情报预警，对异常指令、非授权操作及设备异常状态进行多源数据融合分析，形成可视化的工业安全全景视图。监测机制应具备 7×24 小时不间断运行能力，通过算法实时识别工艺参数偏离、通信功能码异常等早期风险特征，为主动防御提供秒级预警支撑，确保生产环境安全可控。

4. 工控安全体系化建设框架设计

体系化建设是提升智能制造企业工控安全水平的关键路径。基于智能制造场景下存在的安全职责不清晰、技术防护碎片化、工控运营未开展的实际风险，可从管理维度、技术维度、运营维度三个维度构建防护体系。在具体实践中，管理维度包括安全治理架构域、安全制度体系域、人员安全能力建设域，技术维度包括工控网络架构域、工控边界防护域、工控主机防护域、工控监测审计域、工控安全管理域，运营维度包括常态运维域、应急响应域，共十个领域。企业通过构建工控安全“三维十域”的防护体系，保障生产的安全稳定运行。

4.1 管理维度建设

4.1.1工控安全治理架构域

企业通过合理的组织结构设置、人员配备和工作职责划分，可以实现对工控安全工作的全方位管理，充分发挥各部门和各类人员在工控安全工作中的作用。

安全治理架构方面，应建立包含决策层、管理层和执行层的三级安全管理组织。决策层应在企业数字化或信息化委员会中增加

工控安全决策职能，相关岗位负责对工控安全战略进行规划，制定工控安全工作目标，统筹管理工控安全项目的立项、评审和质量控制，保障工控安全所需资源的投入。管理层由信息化管理部门牵头成立工控安全领导小组，负责承接决策层下达的工控安全建设与管理工作，制定工控安全管理制度规范，对生产部门进行工控安全相关的培训，监督检查工控安全规范的执行情况。执行层在各子公司或部门设立工控安全接口人，负责落实企业工控安全各项规章制度，建立工控系统的安全操作细则，开展现场工控安全事件的应急处置，响应各类工控安全风险的整改要求。

4.1.2工控安全制度体系域

工控安全管理制度是所有与工控安全有关的人员必须共同遵守的行为准则。工控安全管理需要建立一套覆盖范围全面、层次结构清晰的管理制度体系，覆盖企业所有工控安全相关活动。工控安全部门在制度建设过程中可参考 ISO27001 四级文件体系。一级文件包括工控安全方针策略，确立工控安全总的指导框架，解决“为什么”做工控安全的问题。二级文件包括规范、程序、管理办法，包括工控安全各子域的具体要求，解决工控安全具体“做什么”的问题。三级文件包括细则、指南、手册，包括工控安全落地的详细操作，解决工控安全“做到怎样”和“怎么做”的问题。四级文件包括记录、表单，是工控安全政策和标准的实施执行结果的痕迹，解决的是“做的结果”的问题。

4.1.3人员安全能力建设域

人员能力建设方面，工控安全部门应组织分层次、多维度地培养工控安全人才。人力资源部门针对全员开展安全意识教育，通过线上课程、培训考核、案例演示等形式强化风险意识。工控安全部门面向技术岗位实施工控安全技能提升，围绕渗透测试、应急响应、攻防对抗等方向开展实战化培训。企业培训部门针对管理层展开安全管理能力提升行动，并进行人员培训，培训内容涵盖战略规划、合规治理等，通过情景模拟、沙盘推演等形式培养工控安全管理机制。在认证方面，人力资源部门建立工控安全认证激励机制，将工控安全领域权威资质纳入人才评价指标，对通过认证人员给予培训经费补贴与职业发展支持。

4.2 技术维度建设

工控安全技术维度主要采用分层防护的设计，如图 2 所示 [5]。

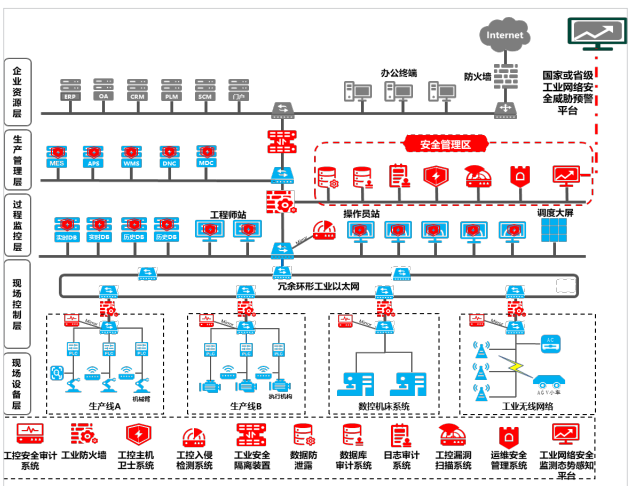


图 2 工控安全分层防护拓扑图

4.2.1工控网络架构域

参考企业普渡模型，按照纵向分层、横向分区的原则进行区域划分。企业整体网络分为管理网和生产网，生产网又划分为生产管理层、过程监控层、现场控制层和现场设备层[6]。在原有架构基础上新建一个独立的安全管理区，对分布在网络中的安全设备进行集中管控。

4.2.2工控边界防护域

在企业管理层与生产管理层之间部署工业网闸，对工业协议报文拆包、内容剥离、安全过滤、单向传输、协议重组等，保障只有生产数据从生产网传输到管理网,禁止管理网违规访问生产网，实现生产网与管理网之间的大边界的安全隔离。

在生产网内部的生产管理层与过程监控层之间部署工业防火墙，实现边界访问控制和安全防护，并对工业通信协议进行深度解析，对途经防火墙的工业协议字段与量值进行检查[7]。

在各个生产线内部的核心控制设备（如 PLC、数控机床等）前端部署工业防火墙，对工业通信协议进行深度解析，对途经防火墙的工业协议字段与量值进行检查，实现对核心控制设备的精准防护。

4.2.3工控主机防护域

在所有工业主机上安装主机卫士系统网络版客户端或将原本安装的单机版客户端替换为网络版，工控安全技术人员通过一键固化、应用白名单、主机加固、外设管控等技术实现对工业主机的安全防护，增强未知威胁防范能力。在安全管理区部署主机卫士管理平台服务器，实现对所有客户端的策略下发、统一管控。

4.2.4工控监测审计域

在各生产线的交换机旁路部署工控安全审计，对流经各产线生产控制系统的网络流量进行审计[8]。对上位机与下位机之间的工业协议进行识别，并进行深度解析，对违规操作、误操作以及关键操作（如下载、上传、组态变更以及 CPU 启停）等进行监测，实时了解生产网络的安全状态，为事后追溯、定位提供证据。

在过程监控层的交换机旁路部署工控入侵检测，对生产网络中存在的异常威胁、漏洞利用行为、恶意攻击进行实时检测。

4.2.5工控安全管理域

4.2.5.1 数据防护设计

在安全管理区部署数据库审计设备，建立数据库操作的风险特征与审计行为的映射规则，对常见的工业实时数据库以及关系型数据库进行审计[9]。

在生产管理层的交换机旁路部署数据防泄露，对出口流量进行内容审计，实现数据敏感信息识别、网络数据泄露监控预警、事件审计及业务分析，防止通信网络中传输、存储、处理的数据信息丢失、泄露或者被篡改、删除。

4.2.5.2 日志审计设计

在安全管理区部署日志审计系统，实现对生产网络的各类网络设备、安全设备、工控设备以及操作系统、数据库、应用系统的日志信息的集中收集与分析。日志审计系统的部署解决了网络设备、操作系统、数据库、业务系统以及安全设备的日常运行过程中各类日志信息未集中收集与分析问题[10]。同时满足《中华人民共和国网络安全法》“日志留存不少于六个月”的要求。

4.2.5.3 漏洞管理设计

工控安全技术人员在安全管理区部署工控漏洞扫描系统，在工控系统上线前或检修时进行安全扫描，适用时期可定期进行扫描，临时接入设备应立即进行安全扫描，及时发现安全漏洞并在

仿真测试环境验证后完成漏洞修补或风险消减工作[11]。

4.2.5.4 安全运维设计

在安全管理区部署运维堡垒机，通过在防火墙上设置 ACL 访问策略或其他技术手段保障运维数据流只能经过堡垒机到达运维对象，对运维过程进行录屏、键盘记录，并定期进行审计，保障远程运维安全。同时通过堡垒机自带的双因素认证功能实现对用户登录账户的双因素认证。

4.2.5.5 态势感知设计

工控安全技术人员在安全管理区部署企业级工业网络安全监测态势感知平台，对部署的安全、网络以及关键业务系统进行操作日志、运行日志以及告警日志的集中采集、泛化和关联分析，并从整体视角进行实时感知、事件分析、预警研判等，同时联动工控安全设备一键封堵、策略下发，提升企业整体工控安全防护预警水平。

企业工业网络安全监测态势感知平台与国家或省级工业安全威胁预警平台对接，落实安全威胁报送与预警处置的工作要求。

4.3 运营维度建设

4.3.1工控安全常态运维域

在日常运维体系建设中，重点围绕安全监控、安全预警、安全设备管理、安全策略管理开展工作。

在安全监控方面，工控安全运营中心常态化监控工控系统运行情况，一旦发生安全事件后，及时分析安全系统及设备的事件日志，进行事件的追踪定位。工控安全运营人员定期通过对所有安全系统及设备的报警日志进行分析，梳理工控安全状况，提出改进建议。

在安全预警方面，工控安全运营人员及时获取国内外权威漏洞发布的工控安全漏洞信息（如 CVE、CNVD 等），生成漏洞通报报告，漏洞信息包括 CVE 编号、漏洞名称、漏洞描述、受影响的资产、危险等级、修复方式等。

在安全设备管理方面，工控安全运营人员定期对设备配置、日志进行备份，定期对管理范围内的安全系统及设备进行巡检，对设备的物理运行情况、系统运行情况进行检查记录，排除可能存在的安全问题和隐患。

在安全策略管理方面，工控安全运营人员建立工控安全基线，适时对安全设备的配置进行变更调整，如登录口令、登录方式、密码复杂度等;定期对安全设备业务访问策略进行梳理，发现过期、冗余、无效、不明确用途等策略，根据分析结果提供优化调整建议。

4.3.2工控安全应急响应域

为构建科学高效的工控安全事件处置体系，应建立分级响应、逐层递进的三级响应架构，形成工控安全的闭环处置能力。

一级响应（L1 现场处置）作为应急响应的前沿阵地，由经过认证的现场技术团队承担首诊责任，建立“识别—隔离—评估”三步工作法，配置标准化应急处置预案库，各级响应要求技术人员在 30 分钟内完成初步影响评估并启动应急预案，确保初级事件及时得到处置，防止事态扩大。

二级响应(L2 专家支持) 在事件处置难度超出基层处置能力时，应触发专家支援机制。工控安全运营中心迅速召集由工控安全认证工程师、网络安全架构师、业务连续性管理专家构成的技术支持组，技术方面提供漏洞验证、攻击溯源等深度分析，管理方面开展内外部资源协调减少停机带来的影响。

三级响应（L3 厂商协助）针对重大系统性安全事件，建立厂商协同处置程序。预先与核心设备供应商签订 SLA 协议，明确重大故障响应时限。启用证据保护措施，在获取厂商技术支持前完成

证据取证。同时，工控安全运营人员通过工业网络安全监测态势感知平台实现与国家或省级平台事件的报送，及时获取监管部门的指导意见。

5. 总结

智能制造企业工控安全体系化建设是一项系统性工程，涉及管理、技术、运营多个层面。在管理层面，工控安全部门需优化组织架构，明确信息化与生产业务部门权责边界，建立跨部门协同机制，并通过制度流程固化安全要求。在技术层面，工控安全技术人员应重点关注工业协议深度解析和工控安全措施的实施对工控系统实时性的影响，采用工业网闸、工业防火墙等策略构建纵深防御体系，同时随着技术的成熟还可考虑通过信创产品替代降低供应链风险。在运营层面，工控安全运营人员需搭建工业态势感知平台，整合安全数据、生产数据形成动态风险画像，同时应制定应急预案、开展应急演练，推动安全策略持续迭代。通过构建“三维十域”的工控安全防护体系，企业可实现从碎片化防护到整体防护的跃升，有效支撑企业的安全可持续发展。在未来发展中，随着智能制造技术不断演进与应用场景日益复杂，企业需持续关注工控安全动态，不断优化工控安全管理体系，加强人才队伍建设，深化内外部合作交流，方能在数字化时代浪潮中筑牢工控安全防线，实现智能制造的高质量发展[12]。

参考文献

[1] 张莹莹 , 朱智光 . 电力工业控制系统 DDoS 攻击防御技术研究 [C]//2021 年水电和新能源工业控制系统安全技术交流会论文集 .2021:63-68.

[2] 樊佳讯 . 面向工业控制系统的漏洞扫描系统设计与实现 [D]. 哈尔滨 : 哈尔滨工业大学 ,2020.

[3] 闫少勃 . 国产化 PLC 上下位机安全通信技术研究 [D]. 西安: 西安电子科技大学 ,2017.

[4] 马跃强 . 煤炭港口管控一体化系统工控安全防护设计 [J]. 工业信息安全 ,2022(2):75-81.

[5] 文雅玫 , 李建强 , 谢博文 , 等 . 烟草行业工控系统安全监测与管控方案 [J]. 自动化博览 ,2018(11):66-68.

[6] 张杰 . 普渡参考模型在工业化网络中的基本原理及优势探究 [J]. 数字化用户 ,2024(18):91-92.

[7] 伍锦荣 . 工业控制系统网络安全现状及解决方案 [J]. 石油化工自动化 ,2017,53(4):1-5.

[8] 韩正 . 智能制造领域工业控制系统网络安全防护研究 [J]. 网络安全技术与应用 ,2020(11):161-163.

[9] 周品荣 . 城市轨道交通 ACLC 系统信息安全研究 [J]. 信息技术与网络安全 ,2020,39(08):72-75.

[10] 马帅 . 电力监控系统的二次安全防护方案分析 [J]. 集成电路应用 ,2022,39(09):164-165.

[11] 孙易安 , 胡仁豪 . 工业控制系统漏洞扫描与挖掘技术研究 [J]. 网络空间安全 ,2017,8(1):75-77.

[12] 王绍杰 , 霍朝宾 , 田晓娜 . 一个工控系统病毒的处置及思考 [J]. 信息技术与网络安全 ,2018,37(01):5-8.

Grab公司探索基于LLM的数据分类方案

绿盟科技 创新研究院 陈佛忠

摘要：本文介绍 Grab 公司基于大型语言模型（LLM）实现数据分类的实践路径。通过整合 LLM 能力于内部 Gemini 系统，Grab 有效解决了传统分类流程中存在的低效、主观性强和误报率高的问题，实现了列级标签的高效自动生成，显著提升了数据治理效率与安全性。

关键词：大型语言模型 数据分类 数据治理 自动化标签生成

1. 前言

本文将介绍 Grab 公司基于 LLM 实现数据分类的实践案例，来展示LLM在数据分类中的应用场景及其对数据安全的提升效果。

2. 案例背景

Grab 是一家总部位于新加坡的科技公司，成立于 2012 年。Grab 最初作为打车应用起步，现已发展为超级应用平台，提供包括网约车、外卖配送、金融服务和电子支付等多种服务。凭借广泛的业务覆盖，Grab 已成为东南亚地区最大的科技公司之一。

基于如此庞大的业务，Grab 需要保护和管理海量的 PB 级数据，以保障用户、司机及合作伙伴的敏感信息安全，同时提升数据分析效率。为此，Grab 开展了数据分类的实践，并结合 LLM 有效提升了数据分类的准确性和效率。

3. 案例详情

在首次应对数据分类问题时，Grab 通过手动流程对数据库的 schema（数据库模式，负责定义数据库中数据组织和结构，包含表、字段、关系模型等）进行了分类标记。Grab 将敏感度划分为四个等级，从第 1 级（高度敏感）到第 4 级（无敏感信息）。在数百张表的 schema 中，如果其中一张表属于第 1 级，整个 schema 都会被划为第 1 级。手动分类的结果是，约一半的 schema 被标记为第 1 级，并实施了最严格的访问控制。然而，实际上真正属于第 1 级的表非常少，这导致大量非敏感的表格也受到了不必要的严格限制，限制了数据的灵活使用和访问效率。

基于此，Grab 尝试将数据分类进一步细化到表级标记，但在实施过程中发现该方案难以有效执行，主要有两方面原因：一方面，随着数据量和种类的快速增长，表级分类比 schema 级分类耗时更长、成本更高；另一方面，表级手动分类存在较大的主观性差异，不同的打标人员在操作过程中可能会产生不一致的分类结果，影响了分类的准确性和一致性。

为了应对这些挑战，Grab 内部开发了一项名为 Gemini 的服务（Grab 官方宣称其 Gemini 的命名早于 Google 的 Gemini 聊天机

器人），通过整合第三方数据分类服务，实现对数据实体的批量扫描，并自动生成列级和字段级标签，这些标签随后交由数据生产者进行审核确认。在这个过程中，Grab 的数据治理团队提供分类规则，并结合正则表达式分类器和第三方工具中的机器学习分类器，自动识别敏感信息。这一自动化流程大幅提升了数据分类的效率和准确性，简化了手动分类的复杂性。

然而，在自动化标签生成模式的初期，Grab 遇到了大量误报，自动化效果并不理想。Grab 官方分析主要有三个原因：首先，正则表达式分类器在评估过程中导致了过多的误报；其次，第三方数据分类服务的机器学习分类器不允许进行定制化改造，也导致效果不佳；最后，构建内部分类器需要专门的数据科学团队来训练定制模型，需投入大量时间了解数据治理规则，并准备手动标记的训练数据集，这个过程反而增加了团队的工作负担。基于此，Grab 公司希望寻找到一个更佳的方法去实施数据分类。

随着 ChatGPT 的火爆，LLM 也进入了 Grab 的视野中。与传统方法不同，LLM 通过自然语言接口，可以让数据治理人员通过文本提示表达需求，而无须编写代码或训练模型。LLM 的引入使分类过程更加灵活且高效，能够自动处理各种复杂的数据分类任务。基于此，Grab 尝试集成 LLM 的能力来进行数据分类。

如图 1 所示，在 LLM 的方案中，Gemini 系统的架构主要包括数据平台、协调服务、消息队列和分类引擎。数据平台负责管理数据实体并发起数据分类请求到 Gemini；Gemini 负责与数据平台进行通信，创建数据分类任务给到消息队列；消息队列负责安排和分组数据分类任务给到分类引擎；分类引擎目前有两种（第三方分类服务和 GPT3.5），负责执行分类作业并返回结果。

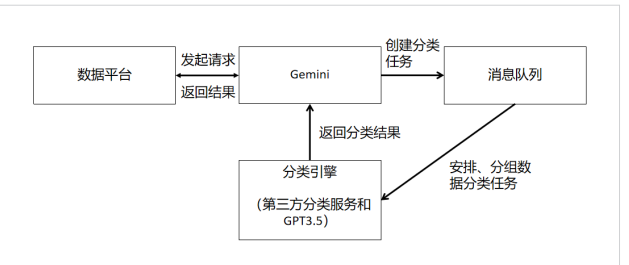


图 1 Gemini 架构图

在该方案中，Grab 希望 LLM 成为列标签生成器，并为每列分配最合适的标签，基于此，Grab 整理了一个标签库，供 LLM 进行分类，表 1 为部分示例。

表 1 标签库示例

列级标签	定义
个人身份证	指可用于唯一识别一个人的外部识别号，应分配给包含“NRIC”“护照”“FIN”“车牌”“社会保障”或类似的列
个人姓名	指的是一个人的姓名或用户名，应分配给包含“姓名”“用户名”或类似内容的列

个人联系方式	指某人的联系信息，应分配到包含“电子邮件”“电话”“地址”“社交媒体”或类似的列
地理散列	指的是 geohash，应该分配给包含“geohash”或类似内容的列

在实际操作中，Grab 发现 LLM 存在两大限制，需要注意。首先是上下文长度限制，GPT-3.5 的上下文长度为 4000 个令牌（约 3000 个单词），输入长度不能超过这一限制。其次是总体令牌限制，输入和输出不能超过设定的令牌配额（目前，所有 Azure OpenAI 模型部署在同一账户下，共享每分钟 240K 个令牌的配额）。这些限制在模型开发和部署中需要特别注意。

基于上述的方案，Grab 大大降低了数据分类的人工工作量并提高了数据分类的准确率。在该方案推出的第一个月内，Grab 已扫描超过 2 万个数据实体，平均每天处理 300—400 个实体。通过自动化标记，Grab 节省了工程师和分析师的大量时间，估计每年总计减少约 360 个人日。工程师和分析师得以专注于核心工作，而非耗时于数据治理。在准确率方面，根据 Grab 官方介绍，对于已确认的表格，平均更改的标签不到一个。在 2023 年 9 月进行的内部调查中，80% 的数据所有者表示，新标记流程可以帮助他们更好地标记数据实体。

取得这样的准确率也不是一蹴而就，Grab 通过不断实践总结出了一些能够有效提升数据分类准确率的方法：

1. 明确要求

任务的要求要尽可能明确，LLM 只会被要求做你要求它做的事情。

2. 少量学习

通过展示交互的示例，以便模型更好地了解它们应该如何响应。

3. 模式执行

利用 LLM 理解代码的能力，明确地向模型提供 DTO（数据传输对象）模式，以便它明白其输出必须符合它。

4. 允许混淆

专门添加了一个默认标签——当 LLM 无法做出决定或感到困惑时，指示它输出默认标签。

4. 小结

通过将 LLM 赋能数据分类流程，Grab 成功实现了从烦琐、易错的手动标记流程到高效、准确的自动化系统的飞跃。Grab 的实践展示了如何通过 LLM 将数据治理流程智能化，增强了对敏感数据的管理能力，确保数据安全。通过对数据的精准分类，Grab 实现了 PB 级数据的自动化管理，提升了工程师和分析师的工作效率。

通过 Grab 的实践可以看出，LLM 在数据分类中的应用带来了显著的效率提升，尤其是在处理大量数据时，自动化分类不仅减少了人工工作量，还提高了分类的准确性。然而，这一技术也存在一定的局限性，例如上下文长度限制和令牌配额问题。

国际数据空间的规范机制与理论模型

绿盟科技 创新研究院 陈佛忠

摘要：在数字经济加速发展的背景下，国际数据空间（IDS）提出以数据主权为核心、去中心化为基础的跨组织数据共享架构。IDS 参考架构模型（IDS-RAM）从业务、功能、信息、过程、系统五层展开，构建标准化的数据交换机制，并通过安全、认证、治理三大视角保障数据的可信流通。体系涵盖身份验证、使用政策、数据生态、语义互操作等关键能力，为构建开放、可信、可控的数据生态系统提供理论与技术支撑。

关键词：国际数据空间 数据主权 去中心化架构 参考架构模型 数据共享治理

1. 引言

随着数字经济的快速发展，数据已成为关键的生产要素。如何在保障数据主权、隐私和安全的前提下，实现跨组织、跨行业、跨国界的数据共享，成为全球关注的焦点。国际数据空间（IDS）作为一种开放、可信的数据共享架构，旨在构建一个以数据主权为核心、技术与治理并重的联邦式数据生态系统。

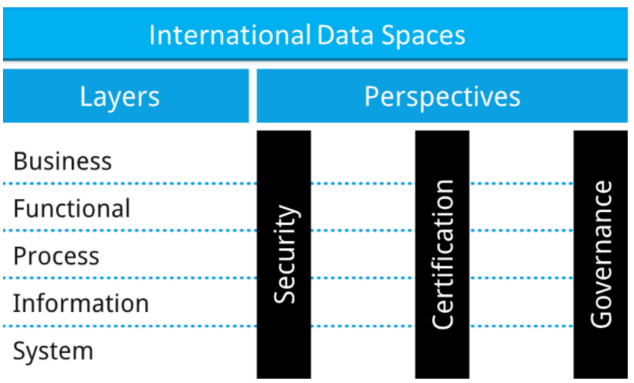
2. 国际数据空间的核心理念

IDS 联合 140 多家企业和研究机构共同推进，致力于建立一个标准化的数据共享框架，使数据提供者在共享数据的同时，仍保有对数据使用的控制权。其核心理念包括：

- **数据主权：**数据所有者对其数据的访问、使用和处理拥有完全的控制权
- **可信交换：**通过技术和制度保障，实现数据在不同参与方之间的可信交换
- **开放互操作：**采用开放标准，确保不同系统和平台之间的互操作性
- **去中心化架构：**避免数据集中存储，降低数据泄露和滥用的风险

3. IDS 参考架构模型（IDS-RAM）五层结构介绍

IDS-RAM 是 IDS 的核心技术框架，定义了构建数据空间的关键组件、交互方式和治理原则。该模型采用五层结构，分别从业务、功能、信息、过程和系统等维度进行设计。



（参考链接：https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0/blob/main/documentation/3_Layers_of_the_Reference_Architecture_Model/3_Layers.md）

3.1 业务层（Business Layer）

业务层定义了数据空间中的参与角色及其交互模式，支持创新的商业模式和数据驱动的服务。主要角色包括：

- **数据提供者（Data Provider）：**拥有并提供数据的组织。
- **数据消费者（Data Consumer）：**使用数据以实现特定目的的组织。
- **数据中介（Data Intermediary）：**促进数据提供者与消费者之间的数据交换。
- **数据空间管理者（Data Space Authority）：**负责数据空间的治理和协调。

3.2 功能层（Functional Layer）

功能层在技术中立的前提下，定义了构建国际数据空间所需的软件功能和服务能力，确保数据共享过程中的信任、安全、互操作性和数据主权。

（参考链接：https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0/blob/main/documentation/3_Layers_of_the_Reference_Architecture_Model/3_2_Functional_Layer/3_2_FunctionalLayer.md）

(1) 信任（Trust）

信任机制是功能层的核心，主要通过以下手段实现：

- **角色管理：**明确参与者的角色及其职责，如身份提供者（Identity Provider）负责创建和验证参与者的身份信息。
- **身份管理：**每个连接器（Connector）必须拥有唯一标识符和有效证书，并能够验证其他连接器的身份。
- **用户认证：**所有参与者需通过认证流程，以建立整个生态系统的信任基础。

(2) 安全与数据主权（Security and Data Sovereignty）

功能层确保数据在交换过程中的安全性和数据主权的实现，主要包括：

- **认证与授权：**连接器需持有有效的X.509证书，通过该证书验证参与者的身份和权限。
- **使用政策与执行：**数据提供者可定义数据使用政策，如禁止数据持久化或转发，连接器负责执行这些政策。
- **可信通信与安全设计：**连接器之间的通信需加密并确保完整性，应用隔离等技术措施用于降低安全风险。
- **技术认证：**核心组件（如连接器）需通过认证机构的技术认证，以确保其符合安全和功能要求。

(3) 数据生态系统（Data Ecosystem）

功能层支持构建一个开放、互联的数据生态系统，主要功能包括：

- **数据源描述：**使用标准化的元数据模型（如RDF/OWL）描述数据源，确保数据的可发现性和可理解性。
- **数据中介服务：**如元数据代理（Metadata Broker）和应用商店（App Store）等，促进数据提供者与消费者之间的连接。
- **词汇管理：**提供共享的本体和词汇，支持语义互操作。

(4) 标准化与互操作（Standardization and Interoperability）

为实现系统间的互操作性，功能层提供：

标准化接口：定义统一的 API 和通信协议，确保不同系统之间的数据交换。

互操作性测试：通过认证流程验证组件的互操作性，确保整个生态系统的协同工作。

(5) 增值应用（Value-Adding Applications）

功能层支持在数据交换过程中引入增值服务，主要包括：

- **数据应用（Data Apps）：**在连接器中运行的数据处理应用，实现数据的转换、分析等功能。
- **应用管理：**连接器需具备应用的部署、管理和隔离能力，确保应用的安全运行。

(6) 数据市场化（Data Marketization）

功能层支持数据的商业化流通，主要功能包括：

- **计费与结算：**支持基于数据使用的计费模型，如按次计费、订阅等。
- **使用限制与治理：**定义数据的使用限制，如地域限制、时间限制等，并通过技术手段强制执行。
- **合规性支持：**确保数据交换过程符合相关法律法规，如GDPR等。

3.3 信息层（Information Layer）

信息层定义了国际数据空间的信息模型（Information Model），提供一个与领域无关的通用语言（Vocabulary），以促进参与者和组件之间的兼容性和互操作性。该模型的主要目的是在分布式参与方组成的可信生态系统中，实现数字资源的（半）自动交换，同时维护数据所有者的数据主权。信息模型支持数据产品和可重用数据处理软件（统称为数字资源）的描述、发布和识别。一旦识别出相关资源，它们可以通过易于发现的服务进行交换和消费。除了这些核心内容，信息模型还描述了国际数据空间的基本组成部分、参与者、基础设施组件和流程。

3.4 过程层（Process Layer）

过程层定义了国际数据空间（IDS）中各组件之间的交互过程，提供了参考架构模型的动态视图。这些过程涉及数据空间的关键价值主张，并涉及业务层中介绍的大多数角色。

(1) 加入过程（Onboarding）

加入过程描述了组织作为数据提供者或数据消费者进入国际数据空间所需的步骤。这包括身份验证、组件认证以及遵守数据空间的治理政策。

(2) 数据提供过程（Data Offering）

数据提供过程涉及数据提供者发布其数据资源的元数据，使其在数据空间中可被发现和访问。这通常通过元数据代理（Metadata Broker）进行，确保数据资源的可发现性和可访问性。

(3) 合同协商过程（Contract Negotiation）

在合同协商过程中，数据提供者和数据消费者就数据的使用政策进行协商，并达成一致。这确保数据的使用符合数据提供者的意图，并保障数据主权。

(4) 数据交换过程（Exchanging Data）

数据交换过程涉及数据在参与者之间的实际传输。这包括数据的加密传输、使用政策的执行以及数据完整性的验证，确保数据在传输过程中的安全性和合规性。

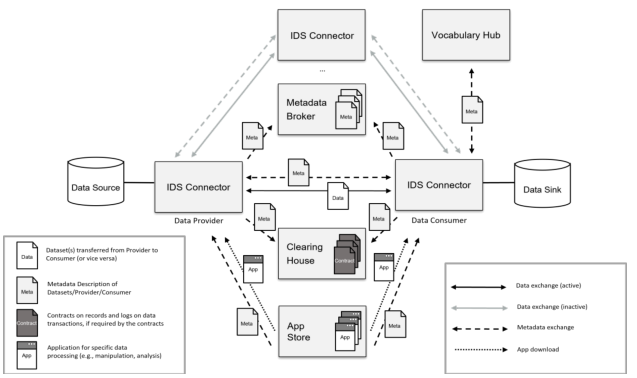
(5) 数据应用的发布与使用（Publishing and Using Data Apps）

数据应用（Data Apps）是在数据交换过程中提供增值服务的应用程序。数据应用的发布与使用过程涉及数据应用的注册、发现、部署和执行，支持数据处理、分析等功能。

通过上述五个关键过程，过程层为国际数据空间的动态运行提供了指导，确保数据在共享过程中的安全性、合规性和价值实现。

3.5 系统层（System Layer）

系统层定义了国际数据空间（IDS）中各组件的技术实现和交互方式，提供了参考架构模型的物理视图。它将过程层中定义的交互过程映射到具体的技术组件和通信机制，确保数据空间的功能需求得以实现。



参考链接：(https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0/blob/main/documentation/3_Layers_of_the_Reference_Architecture_Model/3_5_System_Layer/media/3.5.0.1_interaction_between_technical_components.png)

(1) IDS 连接器（IDS Connector）

IDS 连接器是参与者与国际数据空间交互的主要接口，负责：

- **数据交换：**从内部数据资源和企业系统向国际数据空间发起数据交换，或接收来自数据空间的数据。
- **元数据提供：**向元数据代理（Metadata Broker）提供数据资源的元数据，支持数据的发现和访问。
- **使用政策执行：**根据数据提供者定义的使用政策，控制数据的访问和使用，保障数据主权。

(2) 元数据代理（Metadata Broker）

元数据代理维护数据资源的元数据目录，支持数据的发现。数据提供者通过连接器向元数据代理注册其数据资源的元数据，数据消费者则可以查询这些元数据以发现所需的数据资源。

(3) 应用商店（App Store）与数据应用（Data Apps）

应用商店提供数据应用的注册、发现和下载服务，支持数据处理和分析。数据应用是在数据交换过程中提供增值服务的应用程序，数据提供者和消费者可以通过连接器从应用商店下载和部署数据应用，以实现数据的处理、转换或分析等功能。

(4) 清算机构（Clearing House）

清算机构记录数据交换的交易信息，支持计费和审计。在数据交换过程中，连接器可以将使用合同和数据交换日志发送到清算机构，以确保交易的透明性和可追溯性。

(5) 身份提供者（Identity Provider）

身份提供者负责参与者和组件的身份验证和认证。在数据交换之前，连接器需要通过身份提供者验证其身份，以建立信任关系。

(6) 词汇中心（Vocabulary Hub）

词汇中心提供共享的词汇和数据模式，支持数据的语义互操作性。连接器可以从词汇中心获取领域特定的词汇和数据模式，以增强数据的语义描述能力。

4. IDS 参考架构模型（IDS-RAM）三大视角概述

IDS IDS-RAM 的架构不仅包括五个层次（业务层、功能层、信息层、过程层、系统层），还引入了三个贯穿所有层次的横向视角：安全视角（Security Perspective）、认证视角（Certification Perspective）和治理视角（Governance Perspective）

4.1 安全视角（Security Perspective）

安全视角是 IDS-RAM 的三大横向视角之一，贯穿于架构模型的所有层级，确保数据空间的安全性和可信性。其核心目标是建立和维护参与者之间的信任关系，保障数据在交换和使用过程中的机密性、完整性和可用性。

(1) 安全架构的核心要素

IDS 的安全架构涵盖以下关键要素：

- **身份与信任管理：**通过身份提供者（Identity Provider）对参与者和组件进行身份验证和认证，建立信任关系。
- **通信安全：**采用加密和安全协议，保护数据在传输过程中的安全，防止数据被窃取或篡改。
- **数据使用控制：**通过使用政策（Usage Policies）控制数据的访问和使用，确保数据提供者对其数据的主权。
- **平台安全：**确保IDS组件的运行环境安全，包括操作系统、硬件和虚拟化层等。
- **应用安全：**确保在IDS中运行的数据应用（Data Apps）符合安全要求，防止恶意行为。

(2) 各层级的安全实现

安全视角在 IDS-RAM 的各层级中都有体现：

- **业务层（Business Layer）：**定义参与者的角色和责任，确保业务流程的安全性。
- **功能层（Functional Layer）：**规定安全功能的需求，如身份验证、访问控制和数据加密等。
- **信息层（Information Layer）：**定义数据的元数据和使用政策，支持数据的安全共享。
- **过程层（Process Layer）：**描述数据交换和合同协商等过

程中的安全要求，确保过程的可信性。

• **系统层（System Layer）：**实现安全机制的技术组件，如IDS连接器、元数据代理和清算机构等。

(3) 安全配置文件（Security Profiles）

IDS 提供了不同级别的安全配置文件，以满足不同场景的安全需求：

• **Base Profile：**提供基本的安全功能，适用于对安全要求较低的场景。

• **Trust Profile：**在Base Profile的基础上，增加了更严格的安全措施，适用于对安全有较高要求的场景。

• **Trust+ Profile：**提供最高级别的安全保障，适用于对安全要求极高的场景，如敏感数据的交换。

(4) 数据使用控制（Usage Control）

数据使用控制是 IDS 安全架构的重要组成部分，允许数据提供者定义数据的使用政策，控制数据在交换后的使用方式。这包括：

访问控制：限制哪些参与者可以访问数据。

使用限制：规定数据的使用目的、使用时间和使用次数等。

追踪与审计：记录数据的使用情况，支持审计和合规性检查。

4.2 认证视角（Certification Perspective）

认证视角是 IDS-RAM 的三大横向视角之一，贯穿于架构模型的所有层级，确保数据空间的可信性和合规性。其核心目标是通过标准化的认证机制，建立参与者之间的信任关系，保障数据在交换和使用过程中的安全性和主权性。

(1) 认证在 IDS 中的重要性

在国际数据空间中，数据主权和数据安全是核心价值主张。为了实现这些目标，参与者必须遵循 IDS 的规则，并提供可靠的信息，证明其遵守了这些规则。认证机制确保了参与者和组件的合规性，从而建立了一个可信的数据共享生态系统。

(2) 各层级的认证要求

认证视角在 IDS-RAM 的各层级中都有体现：

• **业务层（Business Layer）：**定义参与者的角色和责任，确保其业务流程符合IDS的规范。

• **功能层（Functional Layer）：**规定组件的功能需求，确保其实现了必要的安全和合规功能。

• **信息层（Information Layer）：**定义数据的元数据和使用政策，支持数据的安全共享和使用控制。

• **过程层（Process Layer）：**描述数据交换和合同协商等过程中的认证要求，确保过程的可信性和合规性。

• **系统层（System Layer）：**实现认证机制的技术组件，如IDS连接器、元数据代理和清算机构等。

(3) 认证的角色和职责

认证过程涉及以下关键角色：

• **认证机构（Certification Body）：**负责认证过程的监督和质量保证，定义标准评估程序，并监督评估机构的行为。

• **评估机构（Evaluation Facility）：**执行具体的评估任务，验证组件和操作环境是否符合认证标准。

• **申请者（Applicant）：**希望在国际数据空间中运营组件的组织或个人，需要通过认证以获得许可。

只有当评估机构和认证机构都确认所有认证前提条件都已满足时，才会授予认证证书。

(4) 认证的类型

IDS 提供了两种主要的认证类型：

- **操作环境认证（Operational Environment Certification）：**评估部署环境的安全性和合规性，确保其满足 IDS 的要求。
- **组件认证（Component Certification）：**评估IDS组件（如连接器、元数据代理等）的功能和安全性，确保其符合IDS的规范。

4.3 治理视角（Governance Perspective）

治理视角是 IDS-RAM 的三大横向视角之一，贯穿于架构模型的所有层级，确保数据空间的合规性、透明性和数据主权的实现。其核心目标是通过定义角色、功能和流程，建立一个可信赖的数据共享生态系统。

(1) 治理在 IDS 中的重要性

在国际数据空间中，数据主权和数据安全是核心价值主张。为了实现这些目标，参与者必须遵循 IDS 的规则，并提供可靠的信息，证明其遵守了这些规则。治理机制确保了参与者和组件的合规性，从而建立了一个可信的数据共享生态系统。

(2) 各层级的治理要求

治理视角在 IDS-RAM 的各层级中都有体现：

- **业务层（Business Layer）：**定义参与者的角色和责任，确保其业务流程符合IDS的规范。
- **功能层（Functional Layer）：**规定组件的功能需求，确保其实现了必要的安全和合规功能。
- **信息层（Information Layer）：**定义数据的元数据和使用政策，支持数据的安全共享和使用控制。
- **过程层（Process Layer）：**描述数据交换和合同协商等过程中的治理要求，确保过程的可信性和合规性。
- **系统层（System Layer）：**实现治理机制的技术组件，如IDS连接器、元数据代理和清算机构等。

(3) 治理的关键组成部分

治理视角涵盖以下关键组成部分：

- **数据治理模型（Data Governance Model）：**定义数据的所有权、使用权和责任，确保数据的合法使用
- **数据质量管理（Data Quality Management）：**确保数据的准确性、一致性和完整性
- **数据主权保障（Data Sovereignty Assurance）：**通过使用政策和合同，确保数据提供者对其数据的控制权
- **隐私保护（Privacy Protection）：**确保个人数据的保护，符合相关法律法规
- **数据空间实例（Data Space Instances）：**定义特定领域或行业的数据空间实例，支持特定的治理需求
- **IDS 规则手册（IDS Rulebook）：**提供参与者在数据空间中操作的规则和指南，确保一致性和合规性

(4) 治理机制的实现方式

治理机制通过以下方式实现：

- **角色定义：**明确参与者在数据空间中的角色和责任。
- **流程规范：**定义数据交换、合同协商等关键流程的操作规范。
- **政策制定：**制定数据使用、访问控制等方面的政策，确保数据的合法使用。
- **合规性检查：**通过认证和审计机制，确保参与者和组件的合规性。

通过上述治理机制，IDS-RAM 的治理视角确保了数据空间的合规性、透明性和数据主权的实现，支持数据的安全共享和使用。

综上，这三大视角确保了数据空间在设计和实施过程中，能够全面考虑安全性、可信性和治理机制，从而构建一个可信的数据共享生态系统。

2025 Verizon DBIR解读 | 供应链攻击30%+勒索软件44%：边缘设备漏洞与AI滥用催生新风险

绿盟科技 创新研究院 浦明

摘要：Verizon 于 2025 年发布的《数据泄露调查报告》显示，数据泄露事件激增，边缘设备漏洞、生成式 AI 滥用和供应链攻击成为主因。勒索软件占比升至 44%，GenAI 平台引发的内部泄露激增。系统入侵主导攻击模式，基础 Web 应用攻击取代杂项类错误成为第二大威胁行为。外部威胁参与者和间谍动机显著上升，医疗行业首次超越金融成为重点攻击目标。报告显示，攻击技术日益“技术化”与“自动化”，强调以漏洞管理和 AI 使用规范构建防御体系。

关键词：数据泄露 Verizon 勒索软件 供应链攻击 AI 滥用

1. 概述

2025 年 5 月，国际知名咨询机构 Verizon 发布了《2025 年数据泄露调查报告》*(2025 Data Breach Investigations Report)* [1]。该报告综合分析了自 2024 年 5 月至 2025 年年初的 22052 起安全事件，其中 12195 被确认为数据泄露事件，覆盖全球共 139 个国家。这一数据代表了 Verizon DBIR 报告有史以来分析的最高数据泄露事件数量，数据来源主要由 Verizon 威胁研究团队（VTRAC）提供。

2024 年，绿盟科技创新研究院曾对《Verizon 2024 年数据泄露调查报告》[1] 进行深度解读。本文延续该分析框架，聚焦 2025 年 Verizon 报告核心结论，并通过与 2024 年数据的纵向对比展开三层次剖析：

- 报告核心结论：提炼威胁演变的整体趋势；
- VERIS框架四维分析：基于威胁参与者、行为、资产、属性解构数据泄露事件；
- 安全事件三维透视：从成因分类、行业分布、区域分布揭示攻击模式。

通过此结构化解读，旨在通过本文增强读者对数据泄露问题的重视。

2. 2025 Verizon DBIR 报告主要结论

结论 1 2024 年利用漏洞窃取数据的攻击行为增长至 20%，年增长 34%，主要原因为边缘设备和 VPN 的 0 Day 漏洞利用正加速成为主要入侵途径，其占比为 22%，年增率 8 倍，但漏洞修复率仅为 54%，暴露出防护缺口显著。

结论 2 截止到 2025 年年初，勒索软件攻击相比去年新增 37%，占所有安全事件的 44%，2024 年为 32%。但赎金中位数降至 11.5 万美元，2024 年是 15 万美元，64% 的受害者拒付赎金，高于两年前的 50%，可能是赎金金额下降的原因导致。此外，小企业组织更易受勒索软件的攻击。

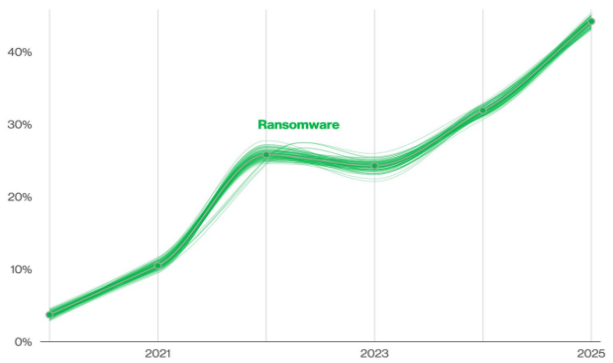


Figure 6. Ransomware action over time in breaches (n for 2025 dataset=10,747)

图 1. 勒索软件攻击趋势图 [1]

结论 3 截至 2025 年年初，72% 的员工使用个人账号访问 GenAI 平台引发内部数据泄露风险激增；与此同时，外部 AI 生成的钓鱼邮件量同比新增 100%，使企业面临双重安全威胁。

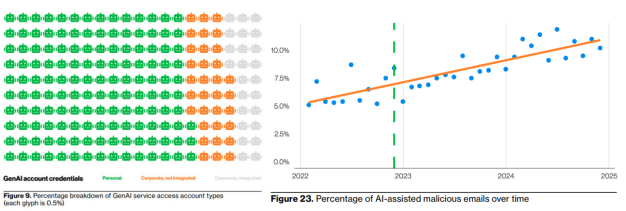


Figure 9. Percentage breakdown of GenAI service access account types (each glyph is 0.5%)

Figure 23. Percentage of AI-assisted malicious emails over time

图 2. GenAI 平台账号登录分布及 AI 生成钓鱼邮件时间分布 [1]

结论 4 2024 年至 2025 年间，涉及第三方的数据泄露事件占比从 15% 上升至 30%，这一增长主要归因于 MOVEit 和 Snowflake 等重大供应链攻击事件引发的系统入侵（System Intrusion）

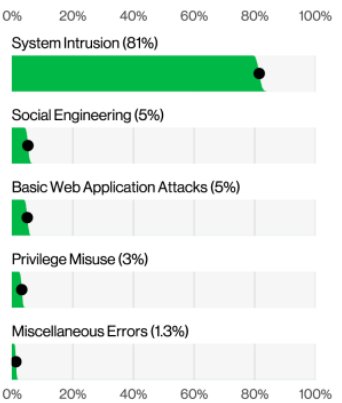


Figure 11. Top patterns in breaches with third-party involvement (n=2,360)

图 3. 涉及第三方数据泄露的攻击行为占比 [1]

结论 5 凭证泄露风险依旧严重，公开代码仓库中可获取的泄露密钥占比高达 50%。具体分布为 Web 应用凭证占 39%，其中 JWT 认证令牌占 66%，云密钥中 Google Cloud API 占比最高 43%，值得注意的是，凭证修复中位数周围长达 94 天，形成持续暴露窗口，使攻击者能够轻易绕过认证机制

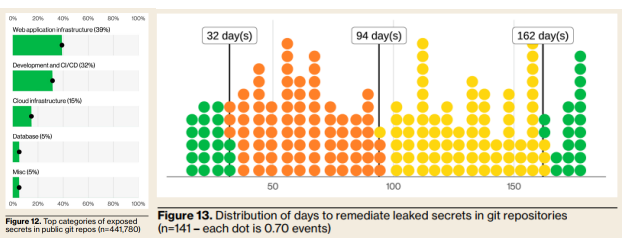


Figure 12. Top categories of exposed secrets in public git repos (n=441,780)

Figure 13. Distribution of days to remediate leaked secrets in git repositories (n=141 – each dot is 0.70 events)

图 4. 公开代码仓库可获取的凭证泄露途径及凭证修复周期 [1]

3．威胁要点分析

Verizon DBIR 报告提供了一套指标 VERIS（The Vocabulary for Event Recording and Incident Sharing）——事件记录和事件共享词汇，旨在提供一种结构化和可重复的方式来描述安全事件的通用语言。VERIS 常用指标包含**威胁参与者（Threat actor）、威胁行为（Threat action）、威胁资产（Threat action）、威胁属性(Threat Variety)、数据泄露事件(Breach)、安全事件(Incident)、外部参与者（External）、内部参与者（Internal）、合作伙伴（Partner）**，具体指标解释读者也可以参考《2024 Verizon DBIR 解读 | 数据泄露转向连接云的第三方软件供应链》一文，此处不再赘述。下面我们首先将重点介绍报告在威胁参与者、威胁行为、威胁资产、威胁属性维度中提到的一些关键数据，之后再进行一些分析与总结，以便读者更全面地理解报告内容。

3.1 威胁参与者维度分析

根据 Verizon 报告的数据分析，2025 年数据泄露事件的主要威胁来源和动机相比 2024 年呈现出以下关键趋势：

- **威胁主体分布有一定变化**

外部参与者仍然是最大风险来源，占比高达 81%，较 2024 年的 65% 有显著上升，且攻击手段以系统入侵（System Intrusion）和社会工程（Social Engineering）为主。内部参与者占比降至 18%，2024 年为 35%，且多数泄露事件源于无心之失，其中杂项类错误（Miscellaneous Errors）的发生频率是特权滥用（Privilege Misuse）的两倍。

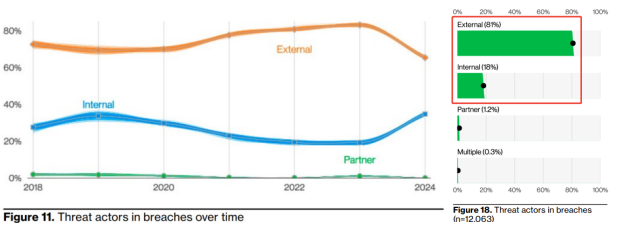


Figure 11. Threat actors in breaches over time

Figure 18. Threat actors in breaches (n=12,063)

图 5. 威胁参与者分布（左 2024 年，右 2025 年）[1]

此外，随着 AI 技术发展迅速，生成式 AI（GenAI）也作为新的威胁参与者正在通过各种方式加剧企业的安全风险，较典型的为数据外泄风险，如企业员工在使用 GenAI 做总结、编码辅助等功能时，常常会上传机密文件和业务代码，从而引发数据泄露。再如员工使用个人邮箱账户绕过企业安全管控的行为。未来可以预见到的是，GenAI 正迅速被集成到手机操作系统中，一些核心功能，如语音助手、相机等默认开启即会大量收集数据，这也同时赋予了新的数据泄露途径，并极大地增加了 BYOD 的风险。

- **攻击动机有新的变化**

金钱驱动仍是主流，与 2024 年情况相似，攻击者主要目标为数据窃取。而间谍动机成为新焦点，2025 年由外部参与者发起的以间谍为目的的泄露事件激增 163%，反映出全球地缘政治紧张局势对网络安全的影响加剧。

3.2 威胁行为维度分析

如图 6 显示了 2024、2025 年报告的威胁行为维度数据对比，数据泄露事件的主要威胁行为呈现显著变化。威胁行为维度上，被窃凭证（Use of stolen creds）、勒索软件（Ransomware）和长尾类别（Other）是导致数据泄露的三大主要威胁。值得注意的是，2025 年长尾类别（Other，44%）占比第一，Verizon 解释这是由于勒索软件和间谍攻击技术复杂化，导致大量细分攻击行为（如 Hacking/Malware）被归入该类别。

2025 年出现三个显著趋势：

一是漏洞利用（Exploit vuln）占比从 10% 上升至 18%，首次超过钓鱼（14%），反映出漏洞武器化的加速趋势；

二是威胁行为类型发生调整——2024 年存在的特权滥用行为在 2025 年消失，同时新增了利用误配置（Exploit misconfig，占 9%）和数据外泄（Export data，占 10%）两类威胁，其共性是通过脆弱性配置实现入侵并窃取数据；

三是针对边缘设备的攻击显著增加：攻击者利用 46% 暴露在互联网边缘的设备漏洞未及时修复的现状，结合这些设备固有的脆弱性，快速发起自动化攻击；并且攻击者利用漏洞从公开到被武器化，约 15 天的速度远快于企业修复的速度，仅 54% 得到修复（虽然 54% 的修复率已经不算低），在防御完成前就实现了入侵和数据窃取。

整体来看，威胁行为正从传统社交工程转向技术性攻击（漏洞利用、配置滥用）迁移，同时攻击技术的复杂化也推动威胁分类体系的动态演进。

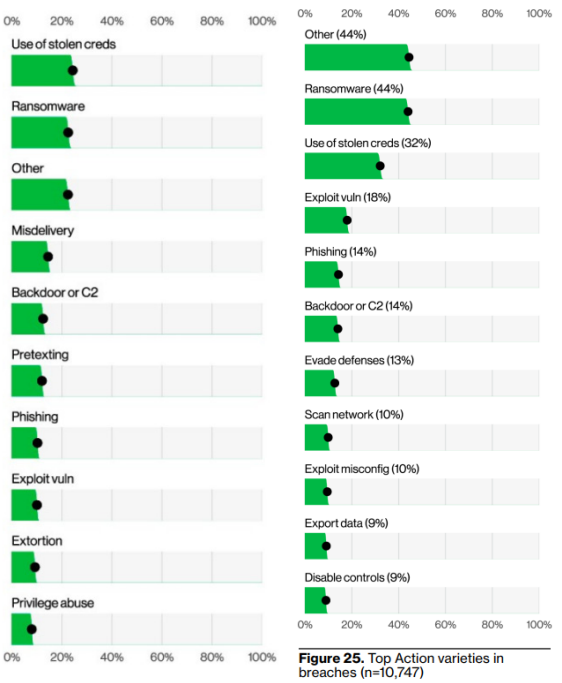


图 6. 威胁行为占比（左 2024 年，右 2025 年）[1]

3.3 威胁资产维度分析

根据 Verizon 的安全事件分析，数据泄露事件中受影响的主要资产类型包括服务器（云端和本地）、人员（Person）、用户开发环境（User Dev）、媒体（Media）和网络（Network）。其中，服务器资产仍是主要的泄露目标，2025 年占比高达 95%，2024 年为 85%，而人员和用户开发环境的泄露占比基本保持稳定。

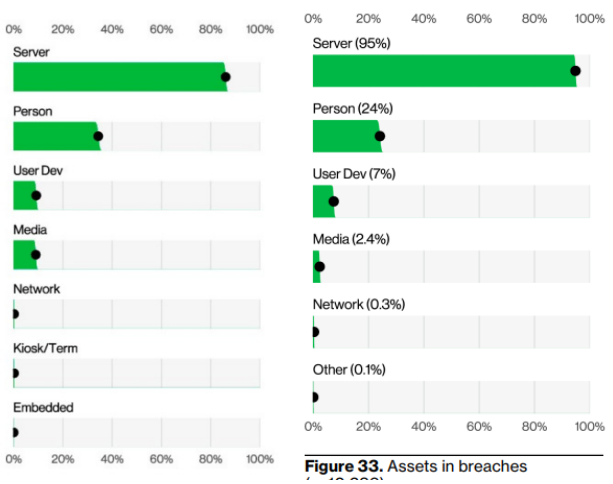


图 7. 威胁资产占比（左 2024 年，右 2025 年）[1]

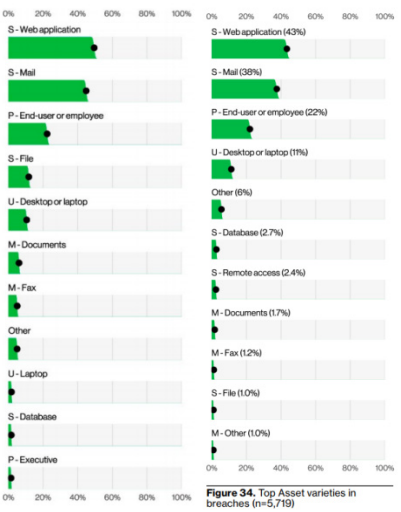


图 8. 威胁资产属性占比（左 2024 年，右 2025 年）[1]

从威胁资产属性来看，Web 应用程序和邮箱服务器是最常被攻击的目标，主要威胁包括凭证窃取和漏洞利用。值得注意的是，2025 年新增了远程访问服务器类别，反映出攻击者正转向利用新的高价值目标。针对此类资产的攻击行为主要集中在漏洞利用（Exploit vuln），并常与后续的勒索软件部署形成关联攻击链。这一趋势表明，随着攻击者偏好的变化，如转向远程访问服务器，数据泄露涉及的资产属性也在动态调整。

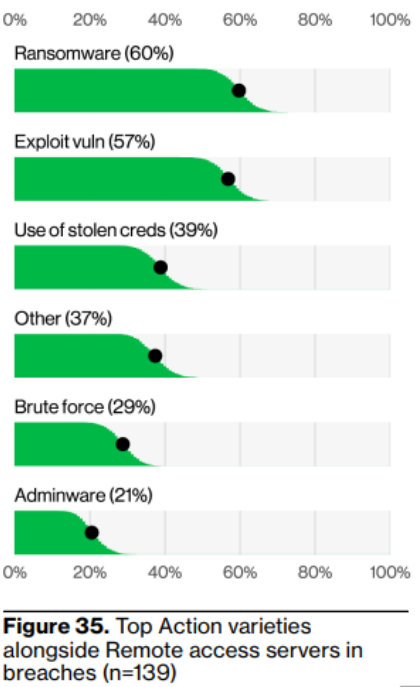


图 9. 针对远程访问服务器资产的攻击行为类别占比 [1]

3.4 威胁属性维度分析

- **Verizon从信息安全三要素（机密性、完整性、可用性）的角度分析，2022—2025年各要素受威胁的频率呈现显著变化，如图10所示：**
- 机密性（Confidentiality）威胁持续上升，从2022年的20%增至2024年的30%，并在2025年大幅跃升至60%，反映出数据窃取和泄露风险的加剧。
- 可用性（Availability）威胁自2022年以来呈线性增长，2025年维持在60%左右，与往年持平，表明勒索软件等破坏业务连续性的攻击仍居高不下。
- 完整性（Integrity）威胁比例波动较大，2022—2024年从40%降至20%，但2025年回升至50%，可能源于数据篡改或供应链攻击的增加。

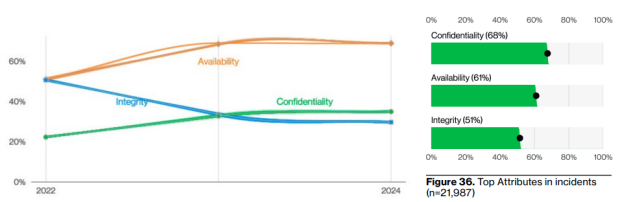


图 10. 信息安全三要素受威胁频率占比（左 2024 年，右 2025 年）[1]

数据泄露类型方面，客户自有数据（Internal）、个人数据（Personal）和凭证数据（Credentials）仍是主要泄露目标。**值得关注的是，2025 年医疗数据（Medical）受勒索软件攻击的占比反超银行数据（Bank）和敏感个人数据（Sensitive Personal），这一变化可能与 2024 年 MOVEit 供应链漏洞的大规模无差别攻击有关，导致医疗行业数据泄露短期激增。**整体来看，攻击者更聚焦于窃取数据（机密性）和破坏服务（可用性），而医疗领域成为新兴的高风险目标。

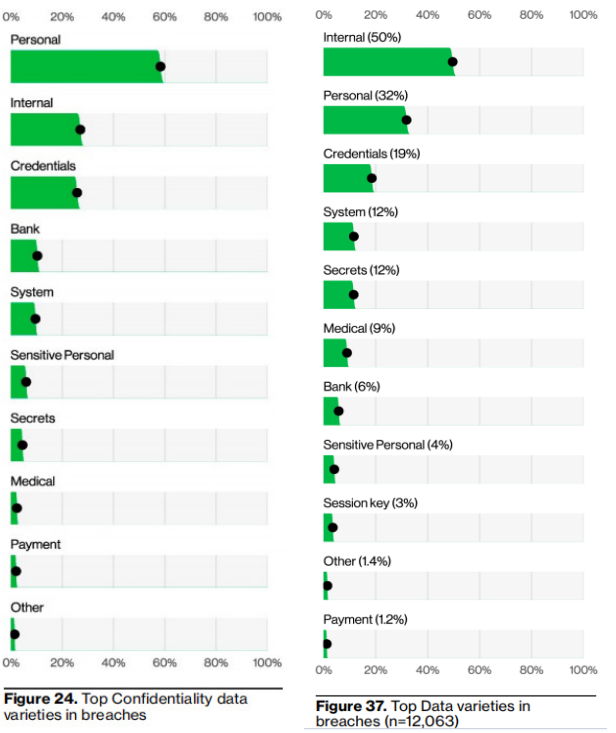


图 11. 数据泄露类型占比（左 2024 年，右 2025 年）[1]

3.5 分析总结

从威胁参与者维度来看，外部威胁参与者占比提升至 81%，

2024 年这一数字为 65%，攻击手段集中在系统入侵和社会工程，并且以间谍为目的的事件暴增 163%，反映出全球地缘政治紧张，黑客组织及犯罪团伙的技术优势正逐步主导威胁态势；与此同时，生成式 AI 的技术发展也导致该技术可被攻击者滥用，进而造成主动泄露（企业员工上传代码或密钥至平台）及被动渗透（AI 集成移动端导致的 BYOD 风险，扩大数据收集面）的两类风险；最后，报告威胁行为维度统计显示，传统社会工程如钓鱼被技术型攻击所反超，如漏洞利用（Exploit vuln）占比从 10% 提升至 18%，并且出现新增的利用配置缺陷（Exploit misconfig，占比 9%）和数据外泄（Export data，占比 10%）两类威胁行为，这两类均依赖系统弱点而非人为欺骗。以上说明，**攻击者的策略向“技术化”与“外部化”迁移，因此防御体系也应从“防人”转向“防技”。**

从威胁属性维度来看，机密性风险占比从 2024 年报告数据的 30% 翻至 60%，相比完整性（50%）和可用性（60%）基本没有变化而言确实高出不少，也进一步说明**攻击者核心诉求的变化，即从“破坏服务”转向“窃取高价值数据”。**

从威胁资产维度来看，**云端 / 本地服务器占比从 85% 提升至 95%，依旧垄断 95% 的泄露事件，且远程访问服务器成为新靶点，常被用于漏洞利用，勒索软件的攻击跳板。**

四．安全事件分析

4.1 事件分类分析

Verizon 将事件分类为八种模式，即基础 Web 应用类攻击（Basic Web Application Attacks），拒绝服务攻击（Denial of Service），丢失和被窃取的凭证（Lost and Stolen Assets），杂项类错误（Miscellaneous Errors 泛指人的因素，无意行为导致安全事件的发生，但丢失设备并不包含在此类中，而属于盗窃类别），权限滥用（Privilege Misuse），社工（Social Engineering），系统入侵（System Intrusion），其他（Everything Else）。笔者分别将 2024、2025 年 Version 报告图从安全事件和数据泄露事件两个维度展示了近年来事件成因随时间推移的变化，如图 12、图 13 所示：

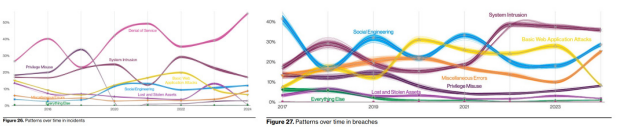


图 12. 2024 Verzion 报告安全事件成因变化，数据泄露事件成因变化 [1]

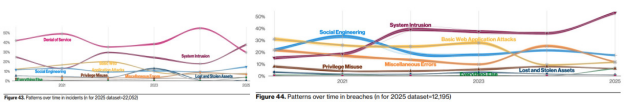


图 13. 2025 Verzion 报告安全事件成因变化，数据泄露事件成因变化 [1]

首先，从安全事件维度来看，自 2018 年以来，拒绝服务攻击持续成为导致事件发生的主要原因。尤其自 2022 年以来，拒绝服务类攻击的频率显著上升，在 2024 年达到最高点，2025 年年初至今有下降趋势，从 60% 降至 30%；另外，是系统入侵类攻击的数据有所变化，系统入侵在 2025 年有所上升，原因是去年的供应链安全事件仍在发酵，且导致安全事件的主要分类为系统入侵。

其次，在数据泄露事件维度方面，从 2017 年起，系统入侵、

社工攻击、杂项类错误和基础 Web 应用类攻击分别成为数据泄露的主要原因。自 2023 年至 2024 年以来，社工攻击和杂项错误模式均显著增加，但 2024 年至 2025 年有轻微下降趋势，社工攻击主要以话术欺诈和钓鱼邮件为主，尤其随着 AI 技术的普及，钓鱼邮件的迷惑性将会越来越高，杂项错误多与发布错误（Publishing error）、配置错误（Miconfiguration）和误投递（Misdelivery）相关，与去年相比前三名有所变化，如图 14 所示：

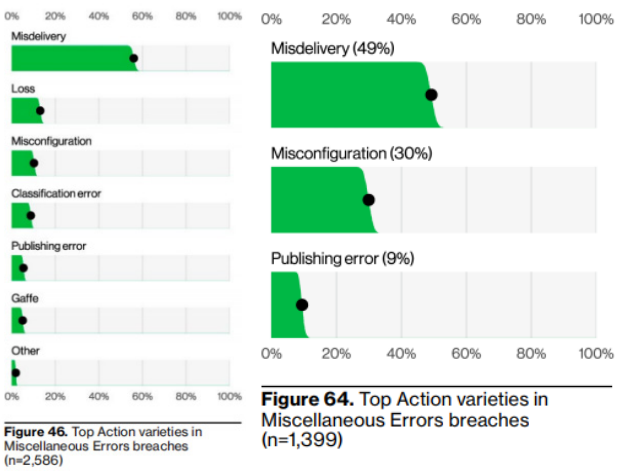


图 14. 导致数据泄露的杂项类错误分类占比（左 2024，右 2025）[1]

4.2 事件行业分布

Verizon 分析了全球发生的 22052 个安全事件，其中有 12195 个被确认为数据泄露事件，按受害者行业和组织规模划分的安全事件和数据泄露数量如图 15 所示：

	Incidents				Breaches			
Industry	Total	Small (1–1,000)	Large (1,000+)	Unknown	Total	Small (1–1,000)	Large (1,000+)	Unknown
Total	22,052	3,049	982	18,021	12,195	2,842	751	8,602
Accommodation (72)	211	52	14	145	121	48	11	62
Administrative (56)	153	107	8	38	145	106	6	33
Agriculture (11)	80	10	3	67	55	10	2	43
Construction (23)	307	151	7	149	252	145	4	103
Education (61)	1,075	116	90	869	851	106	69	676
Entertainment (71)	493	42	12	439	293	37	12	244
Finance (52)	3,336	175	134	3,027	927	162	117	648
Healthcare (62)	1,710	115	153	1,442	1,542	105	132	1,305
Information (51)	1,589	171	76	1,342	784	154	54	576
Management (55)	113	52	3	58	107	52	3	52
Manufacturing (31–33)	3,837	488	74	3,275	1,607	456	42	1,109
Mining (21)	64	27	4	33	52	27	3	22
Other Services (81)	683	87	8	588	583	86	4	493
Professional (54)	2,549	611	95	1,843	1,147	547	75	525
Public Administration (92)	1,422	144	175	1,103	946	124	111	711
Real Estate (53)	339	64	7	268	320	62	6	252
Retail (44–45)	837	170	53	614	419	166	50	203
Transportation (48–49)	361	110	32	219	248	103	25	120
Utilities (22)	358	27	14	317	213	26	10	177
Wholesale (42)	330	260	11	59	319	256	10	53
Unknown	2,205	70	9	2,126	1,264	64	5	1,195
Total	22,052	3,049	982	18,021	12,195	2,842	751	8,602

图 15. 按受害者行业和组织规模划分的安全事件和数据泄露数量汇总 [1]

从数据泄露事件数量来看，受到影响较大的主要行业为：

1. 医疗（位列第二）：事件模式分类上以杂项类错误，系统入侵为主，约占比医疗行业数据泄露事件的 74%

2. 教育：事件模式分类上以系统入侵，社工，杂项类错误为主，约占比教育行业数据泄露事件的 80%，2024 年为 90%

3. 金融：事件模式分类上以系统入侵，社工，基础 Web 类攻击为主，去年的杂项类错误配置攻击消失，约占比金融行业数据泄露事件的 74%，2024 年为 78%

4. 制造业（位列第一）：事件模式分类上以系统入侵，社工，基础 Web 类攻击为主，2024 年的杂项类错误配置攻击消失，约占

比制造行业数据泄露事件的 85%，去年为 83%

5.信息行业：事件模式分类上以系统入侵,基础Web应用类攻击，社工为主，约占比信息行业数据泄露事件的 82%，2024 年为 79%

6. 行政单位（位列第三）：事件模式分类上以系统入侵，社工，基础 Web 类攻击为主，2024 年的杂项类错误配置攻击消失，约占比行政单位数据泄露事件的 78%，2024 年为 91%

从事件分类模式来看，当前安全事件的主要成因集中在系统入侵、社工攻击和基础 Web 类攻击三大领域，这一分布与 2024 年相比保持相对稳定。值得注意的是，在受影响行业分布基本不变的情况下，各行业的安全威胁来源发生了显著变化：原先占比较高的杂项类错误 " 正逐步被 " 基础 Web 类攻击 " 所取代。这一趋势转变表明，由人为操作失误导致的数据泄露事件正在减少，而针对系统漏洞的有组织攻击正在成为网络安全的主要威胁。

与此相比，国内发布的数据泄露报告，如《2025 上半年数据泄露风险态势报告》[4]（由威胁猎人发布）描述了数据泄露事件在不同行业的分布情况。该报告指出数据泄露事件数量 Top10 行业分别为电商、消费金融、银行、快递、证券、社交软件、软件应用、保险、在线票务、汽车品牌。可以看出与国外相比，国内医疗、教育、制造业、行政单位的安全事件相对较少，这也进一步反映出国内政府、传统行业的监管机构措施较严格，数据安全法执行相对严厉，并且相对于互联网、金融、电商等行业，由于金钱驱动力不高，所以攻击者兴趣也较低。

4.3 事件区域分布

图 16 展示了亚太地区（APAC）、欧洲、中东和非洲（EMEA）和北美（NA）三个地区的事件数量分布情况。数据显示，北美地区的事件数量最多，共计 6361 起，而亚太地区地区的事件数量最少，仅为 2687 起。

Region	Frequency	Top patterns	Threat actors	Actor motives	Data compromised
APAC	2,687 incidents, 1,374 with confirmed data disclosure	System Intrusion, Social Engineering and Basic Web Application Attacks represent 97% of breaches	External (99%), Internal (1%) (breaches)	Financial (83%), Espionage (34%) (breaches)	Internal (78%), Other (41%), Secrets (33%) (breaches)
EMEA	9,062 incidents, 5,321 with confirmed data disclosure	System Intrusion, Social Engineering and Miscellaneous Errors represent 89% of breaches	External (71%), Internal (29%) (breaches)	Financial (87%), Espionage (18%) (breaches)	Internal (62%), Personal (49%), Other (37%), Secrets (13%) (breaches)
LAC	657 incidents, 413 with confirmed data disclosure	System Intrusion, Social Engineering and Basic Web Application Attacks represent 99% of breaches	External (100%), Partner (1%), Multiple (1%) (breaches)	Financial (84%), Espionage (27%) (breaches)	Internal (97%), Secrets (27%), Other (24%) (breaches)
NA	6,361 incidents, 2,867 with confirmed data disclosure	System Intrusion, Everything Else and Social Engineering represent 90% of breaches	External (91%), Internal (5%), Partner (5%), Multiple (1%) (breaches)	Financial (95%), Espionage (9%) (breaches)	Internal (49%), Medical (35%), Credentials (23%), Other (17%) (breaches)

图 16. 事件区域分布 [1]

从事件成因方面来看，系统入侵和社工攻击为主要因素，大概占比 90%；从威胁者的角度来看，外部攻击者成为主要威胁；行为动机方面，金融驱动一直是攻击者的主要动机之一。从威胁资产的类型来看，北美主要以内部数据和医疗数据以及凭证为主，而亚太地区则以企业内部数据和凭证为主；从威胁属性来看，亚太地

区主要以被窃取的凭证、勒索软件、漏洞利用为主。

5. 总结

本文对 2025 年 Verizon DBIR 报告进行解读，提炼出报告的主要结论。我们认为 2025 年的报告主要突出的亮点：

生成式 AI 成为双刃剑：生成式 AI 的流行大幅降低了钓鱼邮件、深度伪造语音等社会工程攻击的技术门槛，攻击者可以快速批量构造高欺骗性内容，导致社工攻击效率提升，但同时也会因大量滥用导致数据泄露风险提升。

系统入侵主导攻击模式，基础 Web 应用攻击取代杂项类错误，AI 和漏洞利用的结合组成新的攻击模式：系统入侵仍然是攻击模式中的主流，占比最高，核心为勒索软件和复杂漏洞利用攻击链，基础 Web 应用攻击（如凭证窃取、漏洞利用）取代 2024 年的“杂项错误”，成为第二大威胁行为。AI 和漏洞利用融合的技术将会成为攻击者攻破防护壁垒的关键，企业需构建覆盖漏洞管理，如云服务配置加固、漏洞优先级修复与人员风险管控，如 AI 使用规范等相结合的防御体系。

绿盟科技创新研究院在云上风险发现和数据泄露领域已经开展了多年的研究。借助 Fusion 数据泄露侦察平台，我们已监测到数万个云端暴露资产存在未授权访问的情况，包括但不限于自建仓库、公有云对象存储、云盘、OLAP/OLTP 数据库、大模型组件，以及各类存储中间件等，具体研究内容可参考包括但不限于 DevSecOps 组件，自建仓库、公有云对象存储、云盘、OLAP/OLTP 数据库，大模型组件以及各类存储中间件等，具体研究内容可参考《2023 公有云安全风险分析报告》[5]，《2024 上半年全球云数据泄露风险分析报告》[6]，《全球云上数据泄露风险分析简报》第一期至第五期 [7-11]，云上 LLM 数据泄露风险研究系列 [12-15]。

Fusion 是由绿盟科技创新研究院研发的一款面向数据泄露测绘的创新产品，集探测、识别、泄露数据侦察于一体，针对互联网中暴露的泛云组件进行测绘，识别组件关联的组织机构和组件风险的影响面，实现自动化的资产探测、风险发现、泄露数据分析、责任主体识别、数据泄露侦察全生命周期流程。

图 17 Fusion 能力全景图

Fusion 的云上风险事件发现组件具有如下主要特色能力：

资产扫描探测：通过多个分布式节点对目标网段 / 资产进行分布式扫描探测，同时获取外部平台相关资产进行融合，利用本地指纹知识库，实现目标区域云上资产探测与指纹标记；

资产风险发现：通过分布式任务管理机制对目标资产进行静态版本匹配和动态 PoC 验证的方式，实现快速获取目标资产的脆弱性暴露情况；

风险资产组织定位：利用网络资产信息定位其所属地区、行业以及责任主体，进而挖掘主体间存在的隐藏供应链关系及相关风险。

资产泄露数据分析：针对不同组件资产的泄露文件，结合大模型相关技术对泄露数据进行分析与挖掘，实现目标资产的敏感信息获取；

6. 参考文献

[1] https://www.verizon.com/business/resources/reports/dbir/

[2] https://mp.weixin.qq.com/s/RG5AVGbvrRHGfp86UGInKw

[3] https://verisframework.org/

[4] https://www.threathunter.cn/blog/2025

[5]《2023 公有云安全风险分析报告》https://book.yunzhan365.com/tkgd/qdvx/mobile/index.html

[6]《2024 上半年全球云上数据泄露风险分析报告》https://book.yunzhan365.com/tkgd/cltc/mobile/index.html

[7] 全球云上数据泄露风险分析简报（第一期）https://book.yunzhan365.com/tkgd/sash/mobile/index.html

[8] 全球云上数据泄露风险分析简报（第二期）https://book.yunzhan365.com/tkgd/bxgy/mobile/index.html

[9] 全球云上数据泄露风险分析简报（第三期）https://book.yunzhan365.com/tkgd/xyih/mobile/index.html

[10] 全球云上数据泄露风险分析简报（第四期）https://book.yunzhan365.com/tkgd/xbin/mobile/index.html

[11] 全球云上数据泄露风险分析简报（第五期）https://book.yunzhan365.com/bookcase/wxjf/index.html

[12] 云上 LLM 数据泄露风险研究系列（一）：基于向量数据库的攻击面分析 https://mp.weixin.qq.com/s/5jndWjm_yMEXY0E-W369NQ

[13] 云上 LLM 数据泄露风险研究系列（二）：基于向量数据库的攻击面分析

https://mp.weixin.qq.com/s/KZsGvmyE6WtspDb5ZvNKVg

[14] 云上 LLM 数据泄露风险研究系列（三）：开源大模型应用的攻击面分析

https://mp.weixin.qq.com/s/ADHC4e03ymaPe5ifZ7aODA

[15] 开源大模型推理软件的攻击面分析：云上 LLM 数据泄露风险研究系列（四）

https://mp.weixin.qq.com/s/-hHPcWM71kW--c51GoT4qw

从“芯”构建信任之钥，开启可信数据空间价值共创之门

绿盟科技 营销线 李永松

摘要：在数字经济加速发展的背景下，数据已成为核心生产要素，但“数据孤岛”和安全隐忧严重制约其价值释放。可信数据空间应运而生，作为集成“技术—机制—生态”的基础设施，通过可信计算、国密算法、统一规则等手段，实现数据“可用不可见、合规共享”，推动数据要素高效流通与多方价值共创。绿盟科技以“从芯构建信任”为起点，打造端到端可信体系，助力构建安全、可控、可持续的数据新生态。

关键词：可信数据空间 数据流通 国密算法 隐私计算 价值共创

1. 引言：数字时代的数据困境与机遇

在数字经济蓬勃发展的当下，数据已成为驱动各行业发展的关键生产要素，被誉为新时代的“石油”“黄金”。随着信息技术的飞速发展，数字经济正以前所未有的速度改变着我们的生活和经济模式。数据作为数字经济的核心要素，其价值越发凸显。2024 年我国数字经济规模占 GDP 比重已超过 40%，数据的流通与共享在其中发挥着关键作用。从宏观层面看，数据驱动的决策模式正在改变政府的治理方式，提升公共服务的效率和质量;从微观层面讲，企业通过对数据的深度挖掘和分析，能够实现精准营销、优化生产流程、创新产品和服务，从而在激烈的市场竞争中占据优势。

然而，数据资源的“沉睡”与“孤岛化”问题长期存在，成为制约数字经济发展的重要瓶颈。政府部门因数据合规和安全管控要求“不愿”共享，企业因技术壁垒和标准缺失难以互通，社会因监管滞后和隐私滥用风险不敢开放。这些问题如同一道无形的枷锁，禁锢了数据的流动与价值释放。数据泄露、滥用等安全事件时有发生，严重威胁企业和个人的利益。根据相关统计，近年来全球范围内的数据泄露事件呈逐年上升趋势，造成的经济损失高达数十亿美元。数据流通面临着存储管理风险、数据泄露风险、数据滥用风险、数据篡改风险等安全问题，这些风险不仅阻碍了数据要素价值的充分释放，也使数据拥有者对数据流通心存顾虑。

在这样的背景下，构建可信数据空间成为打破数据流通困境、释放数据要素价值的关键。可信数据空间作为一种新型的数据流通利用基础设施，通过技术、规则和生态的共同作用，确保数据的安全性、可靠性和高效性，为数据要素的市场化配置提供了有力支撑。国家也高度重视可信数据空间的建设，发布了一系列政策文件，如《可信数据空间发展行动计划（2024—2028 年）》，明确提出到 2028 年，我国将建成 100 个以上可信数据空间，形成一批数据空间解决方案和最佳实践。

2. 可信数据空间：数据要素价值共创的新引擎

2.1 可信数据空间的定义与内涵

可信数据空间是基于共识规则，联接多方主体，实现数据资源共享共用的数据流通利用基础设施，是数据要素价值共创的应用

生态，是支撑构建全国一体化数据市场的重要载体 。它以“可信”为核心，通过一系列技术、规则和机制，保障数据在流通和使用过程中的安全性、可靠性和合规性，打破数据孤岛，促进数据要素的自由流动和价值释放。

2.2 可信数据空间建设的五个核心要素

可信数据空间的建设是一个复杂的系统工程，需要综合考虑多个核心要素，包括应用场景、数据资源、生态、规则机制和技术系统。

1. 应用场景：应用场景是可信数据空间建设的出发点和落脚点。只有明确了具体的应用场景，才能有针对性地整合数据资源、构建生态体系、制定规则机制和选择技术系统。应用场景可以涵盖政务、金融、医疗、教育、工业等多个领域，如医疗可信数据空间的政府决策、公卫预警、智能诊疗、智能药研、三医联动等应用场景；企业可信数据空间的供应链、碳足迹管理、物流与仓储等应用场景。

2. 数据资源：数据资源是可信数据空间的核心资产。丰富、高质量的数据资源是实现数据价值的基础。数据资源可以来自政府部门、企业、社会组织和个人等多个渠道，包括结构化数据、半结构化数据和非结构化数据等多种类型。在建设可信数据空间时，需要对数据资源进行全面治理和加工，建立数据目录和数据标准，确保数据的一致性和可用性。

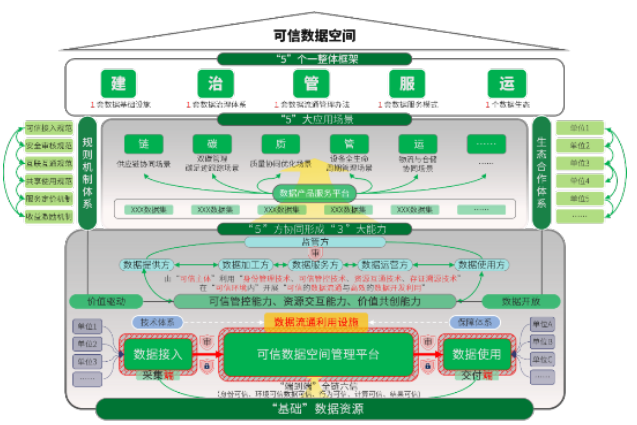
3. 生态：生态是可信数据空间可持续发展的关键。一个健康的生态系统需要吸引数据提供方、数据使用方、技术服务方、运营方等多方主体参与，形成互利共赢的合作关系。通过建立开放的合作平台和激励机制，促进生态各方的协同创新和价值共创，推动可信数据空间的不断发展壮大。

4. 规则机制：规则机制是可信数据空间运行的保障。它包括数据接入规则、数据使用规则、数据安全规则、收益分配规则等，明确了生态各方的权利和义务，规范了数据的流通和使用行为。规则机制的制定需要充分考虑法律法规、行业标准和市场需求，确保其合法性、合理性和可操作性。

5. 技术系统:技术系统是可信数据空间的支撑。它包括数据接入、存储、计算、安全等多个方面的技术，为数据的流通和使用提供技术保障。在选择技术系统时，需要综合考虑技术的先进、可靠、安全、成本、可扩展等，确保其能够满足可信数据空间的业务需求。

2.3 可信数据空间的整体架构

结合可信数据空间的核心要素，总结提炼出可信数据空间的设计框架，如下图所示：



1. 厘清基础：首先厘清已有基础，着重整合已有“数据资源”和“生态资源”，为整个可信数据空间提供坚实的数据和生态基础。

2. 顶层设计：规划搭建“建、治、管、服、运”整体框架，从宏观层面指导可信数据空间的建设、治理、管理、服务和运营。它明确了可信数据空间的发展战略、目标定位、组织架构和运营模式，确保可信数据空间的建设和发展符合国家战略和市场需求。

3. 规则机制：制定可信接入规则，确保参与可信数据空间的各方主体身份可信、数据来源合法;建立安全审核规范,对数据的传输、存储和使用进行严格的安全审查；明确互联互通规范，促进不同可信数据空间之间的信息共享和业务协同；规范共享使用规范，规定数据的使用范围、方式和期限；制定服务定价机制，合理确定数据服务的价格；建立收益激励机制，激发生态各方参与可信数据空间建设和运营的积极性。

4. 生态系统：汇聚数据提供方，提供丰富的数据资源；数据加工方，对原始数据进行清洗、标注、分析等处理；数据服务方，为数据的流通和使用提供技术支持和服务；数据运营方，负责可信数据空间的日常运营和管理；数据使用方，利用数据进行业务创新和价值创造;数据监管方，对可信数据空间的运行进行监督和管理，确保其合规运营。通过各方的协同合作，形成一个完整的可信数据空间生态系统。

5. 应用场景：通过规则机制体系和生态合作伙伴一起价值共创出应用场景。根据不同行业和领域的需求，结合数据资源和技术能力，开发出具有实际应用价值的场景，如联合执法、三医联动等，实现数据的价值变现。

6. 数据服务：结合应用场景需求，利用流通设施构建可信可控安全数据流通服务能力。它提供数据共享、数据交易、数据托管、数据分析等多种服务,为应用场景的实现提供技术支持和服务保障。

7. 三大能力：结合流通设施构建可信数据空间“可信管控能力、资源交互能力、价值共创能力”三大能力。可信管控能力确保数据的安全性和可信性；资源交互能力实现数据的高效流通和共享；价值共创能力推动数据的创新应用和价值创造。

8. 流通设施：构建“端到端”全链可信的数据基础设施，包括数据存储设施、数据计算设施、数据传输设施、数据安全设施等。它为数据的流通和使用提供物理支撑，确保数据能够在可信数据空间中安全、高效地流动。

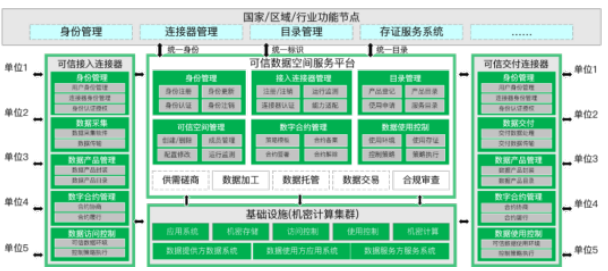
2.4 国家标准与要求

国家高度重视可信数据空间的标准化建设，发布了一系列相关标准和要求,为可信数据空间的建设和发展提供了指导和规范。《国家数据基础设施建设指引》明确了可信数据空间在国家数据基础设施中的重要地位和作用，提出了可信数据空间的建设目标、原则和重点任务。《可信数据空间 技术架构》则从技术、流程、安全等多个方面，对可信数据的建设和运营提出了具体的标准和要求。

这些标准和要求强调了“多个统一”，即统一身份、统一标志、统一目录、统一监管。统一身份确保参与可信数据空间的各方主体身份真实可信；统一标志实现数据的唯一标志和精准定位；统一目

录方便数据的查询和管理；统一监管保障可信数据空间的安全和合规运行。同时，强调全国一盘棋，要求各地在建设可信数据空间时，要遵循国家统一的标准和要求，加强统筹协调和资源整合，避免重复建设和数据孤岛的形成。

2.5 可信数据空间服务平台和可信接入连接器设计



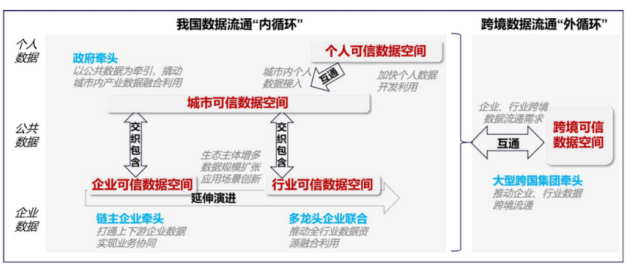
绿盟科技积极响应可信数据空间建设的号召，推出了可信数据空间服务平台和可信接入连接器，为可信数据空间的建设提供了有力的技术支持和解决方案。

绿盟科技可信数据空间服务平台依据国家标准设计，拥有六大核心能力，包括身份管理、接入连接器管理、目录管理、可信空间管理、数字合约管理、数据使用控制等。还扩展有供需磋商、数据加工、数据托管、数据交易、合规审查等能力。

可信接入连接器则是实现数据安全接入可信数据空间的关键设备。它采用了“芯片级”的硬件与软件融合的安全架构，并对接入的设备和用户进行严格的身份认证和权限管理，确保只有合法的设备和用户才能接入可信数据空间。连接器还具备数据加密、传输加密、访问控制等安全功能，有效防止数据泄露和篡改，保障数据在传输过程中的安全性。通过可信接入连接器，企业和机构可以将内部的数据安全地接入可信数据空间，实现数据的流通和共享，同时保护自身的数据安全。

2.6 厘清可信数据空间建设的核心要素

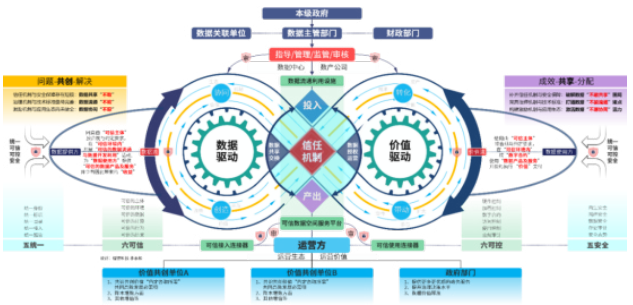
1. 厘清五类可信空间建设的逻辑关系：



（来源：可信数据空间发展联盟《可信数据空间科普问答》）

企业可信数据空间是基础，通过企业内部的数据流通和共享，提升企业的运营效率和创新能力；行业可信数据空间是纽带，促进产业链上下游企业之间的数据协同和创新，推动行业的发展；城市可信数据空间是支撑，整合城市内的各类数据资源，提升城市的治理水平和服务质量；个人可信数据空间是补充，保护个人数据隐私的同时，实现个人数据的价值变现；跨境可信数据空间是拓展，促进国际间的数据交流和合作，推动数字经济的全球化发展。这五类可信数据空间相互关联、相互促进，共同构成了可信数据空间的整体架构。

2. 厘清规划建设的核心要素：



（1）信任机制：信任机制是可信数据空间建设的核心。通过建立可信的身份认证、数据加密、访问控制等技术手段，以及完善的法律法规和监管体系，确保数据的安全性和可靠性，增强数据提供方和使用方之间的信任。

（2）价值共创：实现数据驱动与价值驱动双循环。以数据为驱动，挖掘数据的潜在价值，为业务创新提供支持；以价值为驱动，根据市场需求和业务目标，引导数据的流通和应用，实现数据的价值最大化。通过双循环机制，激发各方参与可信数据空间建设的积极性和创造性。

（3）持续运营：注重运营生态、运营价值。建立健全的运营管理体系，吸引和培育优质的生态合作伙伴，打造互利共赢的生态环境。同时，通过不断优化数据服务和应用场景，提升数据的价值和影响力，实现可信数据空间的可持续发展。

（4）保障能力:落实“五统一、六可信、六可控、五安全”。“五统一”即统一身份、统一标志、统一目录、统一接入、统一监管；“六可信”指身份可信、数据可信、行为可信、服务可信、算法可信、结果可信;“六可控”包括访问可控、使用可控、流向可控、风险可控、审计可控、溯源可控；“五安全”涵盖数据安全、网络安全、应用安全、平台安全、物理安全。通过这些保障能力的建设，确保可信数据空间的安全、稳定运行。

3. 厘清数据运营逻关系：

[参考：中国信息通信研究院政务服务中心 王跃，莫莉娟，苏娜 公共数据授权运营体系整体视图、建设策略与 部委应用研究，数据与计算发展前沿（中英文），2024，6 (5)]

持续开展数据授权运营工作是建设后核心工作。数据授权运营涉及组织机构众多、流程复杂、用户多样，需要升级现有的数据治理体系、管理办法、服务模式，以及数据生态。信任机制是关键，管理流需要确保数据流与价值流在双边正向激励合规有序安全运行。

2.7 可信数据空间的建设路径

可信数据空间的建设是一个渐进的过程，需要遵循“试点—示范—推广”的建设路径，逐步推进。

1. 试点：选择具有代表性的行业、企业或地区开展试点工作。在试点的过程中，充分结合当地的实际情况和需求，探索适合的建设模式、技术方案和运营机制。通过试点，积累经验，发现问题，为后续的示范和推广提供参考。

2. 示范：在试点成功的基础上，打造一批具有示范效应的可信数据空间案例。这些案例应在技术创新、应用创新、模式创新

等方面具有突出表现，能够为其他地区和行业提供借鉴和启示。通过示范案例的宣传和推广，提高社会各界对可信数据空间的认识和认可，激发更多主体参与可信数据空间的建设。

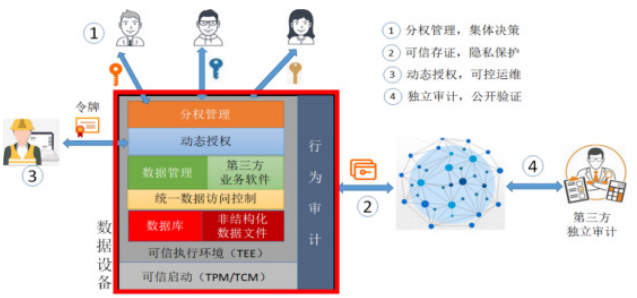
3. 推广：将试点和示范阶段取得的成功经验和模式进行复制和推广，在更大范围内推进可信数据空间的建设。加强政策引导和支持，鼓励各地结合自身特点，积极开展可信数据空间建设，形成全国范围内的可信数据空间网络，促进数据要素的自由流通和价值释放。

3. 从“国芯”构建信任：关键技术与创新突破

3.1 技术架构

绿盟科技基于可信根(TCM\TPM \TPCM)、可信执行环境(TEE)、国密算法（SM4)、统一访问使用控制、数字合约、行为审计、区块链等技术融合创新，形成了一套完善的技术架构。这套架构犹如一座坚实的“保险箱”，为可信数据空间的建设奠定了稳固的基础。

在这个架构中，可信根作为信任的源头，为整个系统提供了初始的信任锚点。可信执行环境则为数据的处理和计算提供了一个安全、隔离的空间，确保数据在运行过程中的安全性和保密性。国密算法作为我国自主研发的加密算法，为数据的加密和解密提供了强大的技术支持，保障了数据的机密性和完整性。统一访问使用控制则对用户的访问行为进行严格的管理和控制，确保只有合法的用户才能访问和使用数据。数字合约则通过智能合约的形式，实现了数据交易和使用的自动化和规范化，提高了数据流通的效率和安全性。行业审计则对数据的使用和流通进行全面的审计和监督，确保数据的使用符合法律法规和行业规范。区块链技术则通过去中心化的分布式账本，实现了行为审计数据的不可篡改和可追溯，增强了数据的可信度和安全性。



3.2 国产可信环境

可信启动（TCM\PCM\TPCM）

基于可信根的可信启动技术是保障系统安全启动的关键。以TCM（可信密码模块）为例，它作为可信根的一种实现方式，内置了多种安全功能,如密码运算、密钥管理、数字签名等。在系统启动时，TCM 首先对 BIOS 进行完整性度量，确保 BIOS 没有被篡改。然后，BIOS 再对操作系统内核进行度量，以此类推，形成一条信任链。只有当链上的每一个环节都通过验证，系统才能正常启动。这种层层验证的机制，就像给系统穿上了一层坚固的铠甲，有效防止了恶意软件在系统启动阶段的入侵，保障了系统的安全性和可信性。

可信执行环境（TEE）

基于可信执行环境 TEE 的可信隔离空间技术，为数据的处理

提供了一个安全的“保险箱”。在 TEE 中，数据和代码被隔离在一个与普通执行环境完全隔离的安全区域内运行。这个区域拥有独立的内存管理、处理器资源和安全机制，外界无法直接访问其中的数据和代码。例如，当进行金融交易时，交易数据可以在 TEE 中进行加密处理和验证，即使操作系统被攻击，交易数据也能得到有效保护，避免了数据泄露和篡改的风险，为数据的安全处理提供了坚实的保障。

3.3 国密算法

可信环境结合 SM4 形成加密虚拟机、加密容器，构建成机密计算。SM4 作为我国自主研发的对称加密算法，具有高效、安全的特点。在加密虚拟机中，SM4 算法对虚拟机中的数据进行加密，使得即使虚拟机的镜像被窃取，攻击者也无法获取其中的明文数据。加密容器则利用 SM4 算法对容器内的数据进行加密保护，确保数据在容器间传输和存储过程中的安全性。通过这种方式，机密计算实现了数据在计算过程中的全程加密，有效保护了数据的机密性，为数据的安全计算提供了可靠的技术手段。

3.4 机密计算

机密虚拟机 \ 机密容器

在云计算环境中，机密虚拟机和机密容器为用户提供了安全的计算环境。机密虚拟机利用可信执行环境技术，将虚拟机的内核和应用程序运行在一个受保护的安全区域内，对虚拟机中的数据进行加密存储和处理。机密容器则通过容器编排技术和加密机制，实现了容器内应用程序和数据的安全隔离和加密保护。无论是大型企业的核心业务计算，还是中小企业的灵活应用部署，机密虚拟机和机密容器都能为其提供安全、高效的计算服务，保障数据在云端的安全运行。

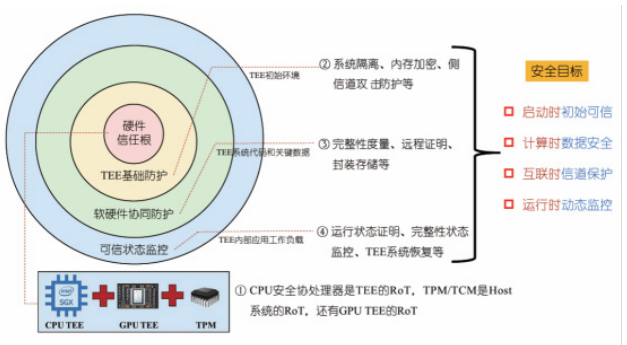
机密存储

机密存储采用加密技术对存储的数据进行加密处理，确保数据在存储过程中的机密性和完整性。例如，使用 SM4 算法对数据进行加密后存储在硬盘或云端存储中，只有拥有正确密钥的用户才能解密读取数据。同时，通过数据冗余和备份技术，保证数据的可靠性，防止数据丢失。这种机密存储方式，为数据的长期存储和管理提供了安全保障，让数据拥有者无须担心数据在存储过程中的安全问题。

机密计算

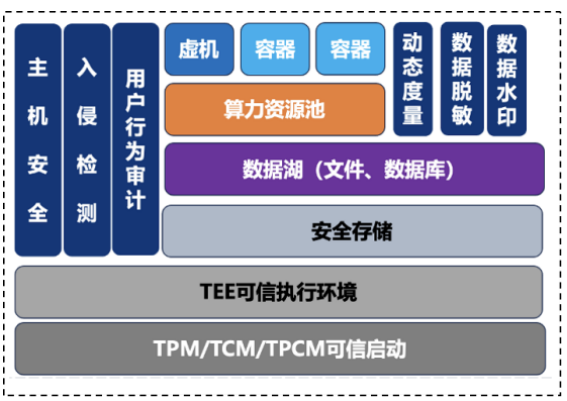
机密计算在计算过程中对数据进行加密处理，使计算过程中的数据始终以密文形式存在。即使计算节点被攻击，攻击者也无法获取到明文数据。通过同态加密、多方安全计算等技术，机密计算能够在不泄露数据内容的前提下进行数据的计算和分析，为数据的安全计算提供了创新的解决方案。例如，在医疗领域，通过机密计

算技术可以对患者的医疗数据进行安全分析，挖掘数据中的潜在价值，同时保护患者的隐私。



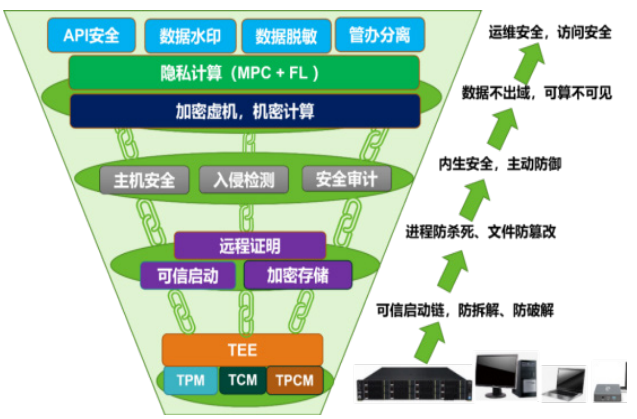
（来源：中国科学院软件研究所 冯登国 机密计算：进展与展望，中国计算机学会通讯，CNCC2024 特邀报告）

3.5 轻量化的一体化安全能力



绿盟科技在可信数据空间服务平台、机密计算集群、可信接入连接器的技术架构中轻量化融入了主机安全、入侵检测、用户行为审计、数据脱敏、数据水印、API 安全等传统安全能力。主机安全防护软件可以保护主机系统免受恶意软件的攻击；入侵检测系统能够实时监测网络流量，及时发现并阻止入侵行为；数据脱敏技术对敏感数据进行处理，使其在不影响业务使用的前提下降低数据泄露的风险；数据水印技术则为数据添加不可见的标志，以便在数据被泄露时能够追踪溯源；API 安全则保障了数据接口的安全性，防止 API 被攻击和滥用。这些传统安全能力的轻量化融入，就像为可信数据空间增添了一道道坚固的防线，进一步提升了可信数据空间的整体安全性。

3.6 隐私计算



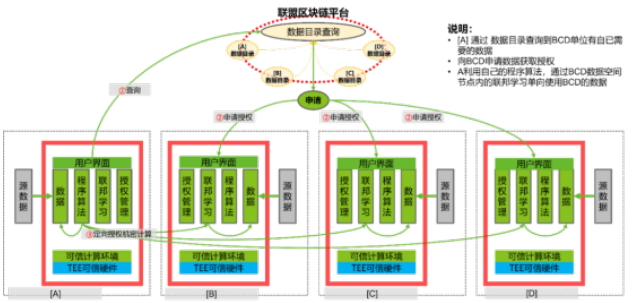
绿盟科技的隐私计算融合了多方安全计算(MPC)、联邦学习(FL)和可信执行环境（TEE）等多种技术的综合架构。通过将多方安全计算 MPC 与 TEE 结合，可以在提升计算效率的同时，进一步增强数据的安全性。通过在 TEE 中运行联邦学习算法，可以确保数据在计算过程中的安全性和隐私性。例如，在医疗领域，不同医院可以通过联邦学习技术联合训练疾病诊断模型，而无须将患者的医疗数据集中存储和共享。TEE+MPC 和 TEE+FL 融合的隐私计算技术的应用，为数据的流通和合作提供了更加安全、隐私保护的方式，促进了数据的价值挖掘和利用。

3.7 区块链

区块链技术在可信数据空间中发挥着重要作用。它通过去中心化的分布式账本，实现了数据的不可篡改和可追溯。在数据流通和共享过程中，每一次数据的操作都被记录在区块链上，形成一条完整的交易记录。审计人员可以查看区块链上的记录，但无法篡改其中的数据。这种特性增强了数据的可信度和安全性，使得数据的来源和流转过程更加透明。例如，在供应链金融中，通过区块链技术可以记录货物的流转信息和资金的交易记录，确保交易的真实性和可靠性，为供应链金融的发展提供了有力的支持。

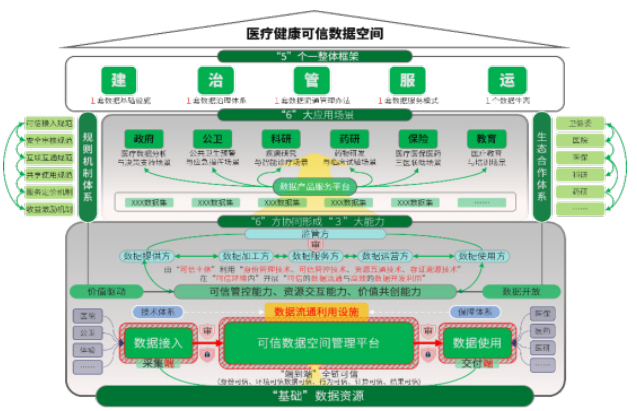
4. 可信数据空间建设探索与实践经验

案例 1：科学数据可信空间



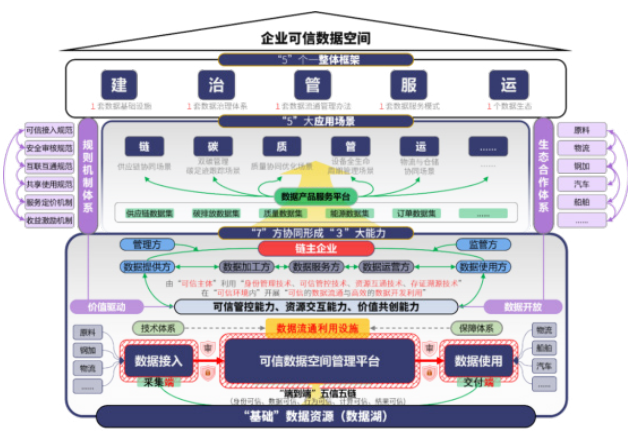
科学数据可信空间致力于为科研人员提供一个安全、可信的数据共享和协作平台。通过采用可信执行环境（TEE）+ 联邦学习（FL）+ 数字合约来访问使用授权数据，可以实现“数据可用不可见，原始数据不出域”。同时，该空间还提供了数据溯源和审计功能，方便科研人员对数据的来源和使用情况进行追踪和验证，为科学研究的可靠性提供了有力保障。

案例 2：医疗健康可信数据空间



医疗健康可信数据空间的建设对于提升医疗服务水平、推动医学科研发展具有重要意义。它汇聚了医疗机构、药企、医保部门等多方的数据资源，包括患者的电子病历、医疗影像、临床检验结果等。通过隐私计算技术，在保护患者隐私的前提下，实现了医疗数据的流通和共享。以卫健委为例，依托绿盟自研可信数据空间能力底座，建设医疗健康行业可信数据空间，实现医疗机构高价值数据集中汇聚开发利用，应用分布式连接隐私计算模式，实现“敏感数据不出域，数据可用不可见”，促进数据合规高效流通使用。通过该空间，药企可以远程使用真实的临床数据，开展药物研发和临床试验；科研机构可以进行大规模的医学数据分析，探索疾病的发病机制和治疗方法；医保部门可以实现精准的医保报销和费用控制，提高医保基金的使用效率。

案例 3：企业可信数据空间



企业可信数据空间是企业实现数字化转型、提升竞争力的重要支撑。它以龙头企业为核心，协同上下游企业，实现了供应链数据的高效流通和共享。以某企业为例，通过搭建企业可信数据空间，与上下游供应商、经销商等建立了紧密的数据连接。供应商可以实时了解企业授权的生产计划和零部件需求，及时调整生产和供货计划，提高供应链的响应速度和协同效率。

5. 未来展望：可信数据空间的无限可能

随着数字经济的持续发展和技术的不断创新，可信数据空间的未来充满了无限可能，将在更多领域发挥重要作用，为经济社会的发展带来深远影响。

在人工智能领域，可信数据空间将成为数据驱动创新的重要支撑。人工智能的发展离不开大量高质量的数据，可信数据空间能够汇聚多方数据资源，为人工智能模型的训练提供丰富的数据来源。通过可信数据空间，科研机构和企业可以获取到来自不同领域、不同行业的海量数据，从而训练出更加智能、精准的人工智能模型。这些模型将在医疗诊断、智能驾驶、金融风险预测等领域得到广泛应用，推动各行业的智能化升级。

从“芯”构建信任之钥，开启可信数据空间价值共创之门，是数字时代发展的必然趋势。让我们积极拥抱这一变革，以创新为驱动，以实践为导向，共同探索可信数据空间的无限可能，开创数据价值共创的新时代，让数据这一关键生产要素在可信数据空间的舞台上绽放出更加耀眼的光芒，为经济社会的发展带来更加美好的未来。

从法律依据到执法案例，看网络安全领域的“安全负责人”制度

绿盟科技 总裁办 张皓天 刘伯英

摘要：随着《个人信息保护负责人信息报送工作的公告》发布，安全负责人制度正逐步走向实操与执法阶段。本文系统梳理了网络安全、数据安全、个人信息保护等五类责任人的设立依据、职责要求与处罚风险，并通过典型案例展示制度落地的执法趋势，为企业理解合规边界、完善内控机制提供参考。

关键词：安全负责人 网络安全法 数据合规 执法案例 个人信息保护

2025 年 7 月 18 日，国家互联网信息办公室发布《关于开展个人信息保护负责人信息报送工作的公告》，要求根据《个人信息保护法》第五十二条、《个人信息保护合规审计管理办法》第十二条规定，处理 100 万人以上个人信息的个人信息处理者，应当向所在地设区的市级网信部门履行个人信息保护负责人信息报送手续。未按照《个人信息保护法》《个人信息保护合规审计管理办法》等法律法规规章的规定履行信息报送手续的，依照有关法律法规规章的规定处理。

此前，网络安全负责人与数据安全负责人的落实和上报，该文件也一定意义上说明我国网络安全法律定义的第三个安全责任人制度，开始正从纸面规定走向执法实践。

本文梳理了当前我国网络安全领域中具备明确法律效力，可以依据条款进行处罚的文件中，规定的网络安全领域三个负责人的法律依据、要求与职责；并整理了近年通用行业领域关于未明确安全负责人相关的执法案例，供读者参考。

1. 设立安全负责人的法律依据与触发条件

我国网络安全领域已形成以《网络安全法》《数据安全法》《个人信息保护法》为核心的监管框架，三部法律均将明确安全责任人作为企业的核心合规义务。

整理目前“硬法”性质的合规文件，关于网络安全，数据安全和个人信息保护的三个负责人，共有如下 5 种用词，涉及 6 个文件。

1.1 网络安全负责人

《中华人民共和国网络安全法》第二十一条，“国家实行网络安全等级保护制度。网络运营者应当按照网络安全等级保护制度的要求，履行下列安全保护义务，保障网络免受干扰、破坏或者未经授权的访问，防止网络数据泄露或者被窃取、篡改：（一）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”。

1.2 数据安全负责人

《中华人民共和国数据安全法》第二十七条，“重要数据的处理者应当明确【数据安全负责人】和管理机构，落实数据安全保护责任”。

1.3 个人信息保护负责人

《中华人民共和国个人信息保护法》第五十二条，“处理个人信息达到国家网信部门规定数量的个人信息处理者应当指定【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”。

《个人信息保护合规审计管理办法》第十二条，“处理 100 万人以上个人信息的个人信息处理者应当指定【个人信息保护负责人】，负责个人信息处理者的个人信息保护合规审计工作”。

1.4（关基）专门安全管理机构的安全管理负责人

《中华人民共和国网络安全法》第三十四条，“除本法第二十一条的规定外，关键信息基础设施的运营者还应当履行下列安全保护义务：（一）设置【专门安全管理机构和安全管理负责人】，并对该负责人和关键岗位的人员进行安全背景审查”。

《关键信息基础设施安全保护条例》第十三条，“运营者应当设置专门安全管理机构，并对【专门安全管理机构负责人】和关键岗位人员进行安全背景审查”。

1.5 网络数据安全负责人

《网络数据安全管理条例》第三十条，“重要数据的处理者应当明确【网络数据安全负责人】和网络数据安全管理机构”。

综上所述，从目前通用领域的硬法角度，所有网络运营者（网安法定义为网络的所有者、管理者和网络服务提供者）均需确定【网络安全负责人】；

重要数据处理者应明确【数据安全负责人】/【网络数据安全负责人】（网络安全负责人 + 数据安全负责人）；

处理 100 万人以上个人信息的个人信息处理者应当指定【个人信息保护负责人】；

关键信息基础设施需确定【专门安全管理机构的安全管理负责人】。

同时，应当注意不同行业下可能有更细一步的通用要求，需要查询相关行业规定，可能比通用硬法要求更严格。如，金融行业的《征信业务管理办法》第三十四条规定，个人征信机构、保存或者处理 100 万户以上企业信用信息的企业征信机构，需要设立【信息安全负责人】和【个人信息保护负责人】，由公司章程规定的【高级管理人员】担任。

2. 安全负责人任职要求与职责

2.1 任职要求与职责

《网络安全法》《数据安全法》《个人信息保护法》未对安全负责人任职要求和职责提出细粒度的明确要求。

《网络数据安全管理条例》第三十条中，要求【网络数据安全负责人】应当【具备网络数据安全专业知识和相关管理工作经历】，由【网络数据处理者管理层成员担任】，【有权直接向有关主管部门

报告网络数据安全情况】。

《个人信息保护合规审计管理办法》第十二条中，要求【个人信息保护负责人】负责个人信息处理者的个人信息保护合规审计工作。

《个人信息保护合规审计管理办法》附件中的《个人信息保护合规审计指引》则对个人信息保护负责人提出了更明确的要求，可供参考：

二十二、对个人信息处理者指定的个人信息保护负责人履职情况进行合规审计的，应当重点审查下列事项：

（一）个人信息保护负责人是否具有相关的工作经历和专业知识，熟悉个人信息保护相关法律、行政法规；

（二）个人信息保护负责人是否具有明确清晰的职责，是否被赋予充分的权限协调个人信息处理者内部相关部门与人员；

（三）个人信息保护负责人在个人信息处理重大事项决策前是否有权提出相关意见和建议；

（四）个人信息保护负责人是否有权对个人信息处理者内部个人信息处理的不合规操作进行制止和采取必要的纠正措施；

（五）个人信息处理者是否公开个人信息保护负责人的联系方式，并将个人信息保护负责人的姓名、联系方式等报送保护部门。

未来关于安全负责人的要求可能进一步提高。《网络安全等级保护条例（征求意见稿）》要求第三级以上网络的运营者对网络安全管理负责人和关键岗位的人员进行安全背景审查，落实持证上岗制度。

与第一章提到的类似，在不同行业的细分领域规定中，对于各类责任人任职要求与职责也有更细一步的划分。如《工业和信息化领域数据安全管理办法（试行）》中有如下要求：

工业和信息化领域重要数据和核心数据处理者，还应当：

（一）建立覆盖本单位相关部门的数据安全工作体系，明确数据安全负责人和管理机构，建立常态化沟通与协作机制。本单位法定代表人或者主要负责人是数据安全第一责任人，领导团队中分管数据安全的成员是直接责任人。

2.2 需要背景审查

网安法中有要求对关基【专门安全管理机构的安全管理负责人】进行【安全背景审查】。

《网络数据安全管理条例》中，掌握有关主管部门规定的特定种类、规模的重要数据的网络数据处理者，应当对网络数据安全负责人和关键岗位的人员进行【安全背景审查】，加强相关人员培训。审查时，可以申请公安机关、国家安全机关协助。

3. 未设立负责人的处罚与实际案例

1.《网络安全法》规定的【网络安全负责人】，主要是警告罚款；

《网络安全法》第五十九条规定，不履行本法第【二十一条】、第二十五条规定的网络安全保护义务的，由有关主管部门责令改正，给予警告；拒不改正或者导致危害网络安全等后果的，处一万元以上十万元以下罚款，对直接负责的主管人员处五千元以上五万元以下罚款。

2.《网络安全法》规定的【关基专门安全管理机构的安全管理负责人】，主要是警告罚款；

《网络安全法》第五十九条规定，关键信息基础设施的运营者不履行本法第三十三条、【第三十四条】、第三十六条、第三十八条规定的网络安全保护义务的，由有关主管部门责令改正，给予警告；拒不改正或者导致危害网络安全等后果的，处十万元以上一百万元以下罚款，对直接负责的主管人员处一万元以上十万元以下罚款。

未设置该责任人同样属于关保条例中第三十九条中的条款范围。如违反，由有关主管部门依据职责责令改正，给予警告；拒不改正或者导致危害网络安全等后果的，处十万元以上一百万元以下罚款，对直接负责的主管人员处一万元以上十万元以下罚款：

3.《数据安全法》规定的【数据安全责任人】，从警告罚款到暂定业务吊销执照，追究刑事责任；

数安法第四十五条规定，开展数据处理活动的组织、个人不履行本法【第二十七条】、第二十九条、第三十条规定的数据安全保护义务的，由有关主管部门责令改正，给予警告，可以并处五万元以上五十万元以下罚款，对直接负责的主管人员和其他直接责任人员可以处一万元以上十万元以下罚款；拒不改正或者造成大量数据泄露等严重后果的，处五十万元以上二百万元以下罚款，并可以责令暂停相关业务、停业整顿、吊销相关业务许可证或者吊销营业执照，对直接负责的主管人员和其他直接责任人员处五万元以上二十万元以下罚款。

违反国家核心数据管理制度，危害国家主权、安全和发展利益的，由有关主管部门处二百万元以上一千万元以下罚款，并根据情况责令暂停相关业务、停业整顿、吊销相关业务许可证或者吊销营业执照；构成犯罪的，依法追究刑事责任。

4. 网络数据安全管理条例规定的【网络数据安全负责人】，主要是警告罚款到停业吊销执照；

第五十七条规定，违反本条例第二十九条第二款、第三十条第二款和第三款、第三十一条、第三十二条规定的，由网信、电信、公安等主管部门依据各自职责责令改正，给予警告，可以并处五万元以上五十万元以下罚款，对直接负责的主管人员和其他直接责任人员可以处一万元以上十万元以下罚款；拒不改正或者造成大量数据泄露等严重后果的，处五十万元以上二百万元以下罚款，并可以责令暂停相关业务、停业整顿、吊销相关业务许可证或者吊销营业执照，对直接负责的主管人员和其他直接责任人员处五万元以上二十万元以下罚款。

5.《个人信息保护法》规定的【个人信息保护负责人】，从警告，罚款，没收违法所得到暂停业务，吊销执照，同时附带了禁业处罚。

未设置个保负责人属于未履行个人信息保护义务行为。同时属于个人信息保护合规审计管理办法中第十八条个人信息处理者、专业机构违反本办法规定的，依照《中华人民共和国个人信息保护法》

《网络数据安全管理条例》等法律法规的规定处理；构成犯罪的，依法追究刑事责任。

因此法律责任均为个保法第六十六条规定，违反本法规定处理个人信息，或者处理个人信息未履行本法规定的个人信息保护义务的，由履行个人信息保护职责的部门责令改正，给予警告，没收违法所得，对违法处理个人信息的应用程序，责令暂停或者终止提供服务；拒不改正的，并处一百万元以下罚款；对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

有前款规定的违法行为，情节严重的，由省级以上履行个人信息保护职责的部门责令改正，没收违法所得，并处五千万元以下或者上一年度营业额百分之五以下罚款，并可以责令暂停相关业务或者停业整顿、通报有关主管部门吊销相关业务许可或者吊销营业执照；对直接负责的主管人员和其他直接责任人员处十万元以上一百万元以下罚款，并可以决定禁止其在一定期限内担任相关企业的董事、监事、高级管理人员和个人信息保护负责人。

近年来，我国网络安全和数据保护领域的执法行动不断强化，安全责任人的个人责任追究走向台前，处罚案例数量明显增加，覆盖行业范围持续扩大。

涉及【未确定网络安全负责人】，通常为产生安全事故后进行处罚，部分场景涉及一案双查

案例一：某农村商业银行股份有限公司

2025 年 7 月，某农村商业银行股份有限公司被通报存在三项违规行为，被中国人民银行酒泉市分行处以警告，罚款 23.95 万元：其中，违法行为涉及未按规定确定网络安全负责人、采取防范计算机病毒的技术措施、健全全流程数据安全管理制度、明确数据安全负责人或管理机构、向行业主管部门定期报送风险评估报告。

案例二：某网络科技公司网站链接到非法网站

2024 年 11 月，湘西州互联网信息办公室依法对永顺县某网络科技公司作出行政处罚。经调查核实，该网络科技公司运营的网站，未履行网络安全保护义务，未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，导致网站内有链接点击后跳转至非法网站的危害网络安全后果。其行为违反了《中华人民共和国网络安全法》第二十一条第一款之规定。湘西州互联网信息办公室根据《中华人民共和国网络安全法》第五十九条之规定，对该公司作出警告和罚款人民币 10000 元，对直接负责人罚款人民币 5000 元的行政处罚。

案例三：某物业服务公司网站被入侵植入黑链

2024 年 3 月，长垣市某物业服务有限公司网站遭黑客攻击入侵，网页被植入境外博彩网站“暗链”。

经查，该公司未制定内部网络安全管理制度和操作规程，未

确定网络安全负责人，网络安全意识淡薄，网站长期无人维护，导致其网络代码出现漏洞，遭到黑客攻击劫持并被植入“暗链”，对网络空间产生了极大威胁。长垣市公安局根据《中华人民共和国网络安全法》第五十九条第一款之规定，对该公司处以警告并责令限期整改的行政处罚。

案例四：江西某技工学校网站被黑客控制

2023 年 10 月，江西省吉安市公安局工作发现，当地某技工学校网站遭黑客攻击控制。

经查，该学校未按规定制定内部网络安全管理制度和操作规程，未确定网络安全负责人，未采取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，未按规定留存相关的网络日志。

江西省吉安市公安局依据《网络安全法》第二十一条、第五十九条之规定，责令该技工学校限期整改，并给予警告。

案例五：四川某医院因网络攻击瘫痪

2023 年，四川省某医院遭受网络攻击，造成全院系统瘫痪。公安机关迅速调集技术力量赶赴现场，指导相关单位开展事件调查和应急处置工作。经调查发现，该医院未制定内部安全管理制度和操作流程,未确定网络安全负责人,未采取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，导致被黑客攻击造成系统瘫痪。公安机关根据《中华人民共和国网络安全法》第二十一条和五十九条之规定，对该院处以责令改正并警告的行政处罚。

涉及【未明确数据安全责任人】，既有执法检查，又有事故后的一案双查

案例一：安庆某房地产公司未履行数据安全保护义务被处罚

2023 年 5 月，安庆市太湖县公安局网安大队、经济开发区派出所民警在对太湖县某房地产公司进行数据安全检查过程中，发现该公司数据安全意识淡薄，未建立数据安全管理制度和操作规程，未对员工开展数据安全教育培训，未对采集到的居民个人信息采取加密措施。

同时，该公司未明确数据安全负责人和管理机构，对存放业主用户数据的办公电脑未采取任何安全防护措施，未采取必要技术措施保障公司数据安全，未履行数据安全保护义务，导致大量数据信息面临泄露的重大风险

太湖县公安局对该公司未履行数据安全保护义务的违法行为，依法处以警告并对直接责任人员处罚款人民币 10000 元的行政处罚。

案例二：隆回某置业公司内鬼贩卖业主信息，未履行数据安全保护义务被处罚

2023 年 5 月，隆回县公安局网络安全保卫大队在查办一起侵犯公民个人信息案件中，发现某小区业主信息泄露线索，随即于 2023 年 5 月上旬联合隆回县局三阁司派出所对该小区的开发公司

邵阳市某置业公司发起“一案双查”。

经查，邵阳市某置业公司数据安全意识淡薄，未建立健全全流程数据安全管理制度，未开展数据安全教育培训，未采取相应的技术措施和其他必要措施保障公司的数据安全，导致该公司开发楼盘的所有业主信息被本公司工作人员多次贩卖。同时，该公司未明确数据安全负责人和管理机构，对存放置业业主用户数据的办公电脑未采取任何安全防护措施，未履行数据安全保护义务

涉及【未明确个人信息保护安全责任人】，从 2025 年起开始关注

案例：上海网信部门处罚一批医疗服务类互联网企业 管理制度、安全防护、存储环节三大类问题值得关注

2025 年 4 月 28 日，近期，上海市网信办在专项执法行动中发现，一批医疗服务类互联网企业（主要从事医疗软件开发与维护、医疗服务培训、数字健康服务等）未依法履行网络安全、数据安全保护义务，所属系统存在网络安全漏洞，被境外 IP 访问并窃取，发生个人信息泄露情况，反映出部分医疗服务类互联网企业存在个人信息制度不规范不健全、安全防护不严密、存储不合规等问题，上海市网信办根据相关法律法规对一批医疗服务类互联网企业予以行政处罚。

现将部分典型问题通报如下：

管理制度方面。部分医疗服务类互联网企业未按照《个人信息保护法》等相关法律法规要求，建立健全个人信息保护内部管理制度。检查中发现，这批被处罚的企业普遍未制定网络数据安全管理制度和操作规程，未明确安全负责人或管理机构，未制定数据分类分级管理、数据访问权限管理、应急预案等制度，网络日志留存不足 6 个月等问题。

4. 总结

网信部门的执法规则体系日渐完善，如 2023 年 6 月 1 日实施的《网信部门行政执法程序规定》，2025 年 5 月 1 日起施行的《个人信息保护合规审计管理办法》（其中第五条特别明确了专项审计），将于 2025 年 8 月 1 日起实施的《网信部门行政处罚裁量权基准适用规定》，以及执法活动的常规化，如 2025 年 6 月 30 日网信办发布的《涉企行政检查事项清单》。

应该说，我国数据合规正在不断进入“实质合规”，执法的法治化、信息化程度越来越高。为此，我们认为，本次个人信息保护负责人的报送工作，不仅仅是落实《个人信息保护法》的要求，也是开启了我国数据合规实质性合规工作新阶段的序章。

参考文献

[1] 锦略法律科技 . 重要! 处理 100 万人以上个人信息企业需在 8 月 29 日前报送负责人信息 [EB/OL]. (2025-07-28)[2025-07-29]. https://mp.weixin.qq.com/s/IEZwtInzoJMHS_k6nyYATw.

综上，我国以《网络安全法》《数据安全法》《个人信息保护法》为基础，系统构建了安全负责人制度的法律。从顶层立法确立制度框架，到协同明确履职要求，再到责任追究贯通组织与个人，形成了“有制度、有职责、有问责”的全链条监管体系。企业及相关主体须依法指定安全负责人，切实履行法定职责。

绿盟网络空间安全仿真平台，助力网安人才高质量培养

绿盟科技 营销线 马跃强 绿盟科技 内蒙古代表处 暴宁

摘要：在数字化时代，网络安全人才已成为国家竞争核心资源。绿盟科技基于多年实战经验，打造“绿盟网络空间安全仿真平台”，集实训、演练、科研于一体，覆盖工控、5G、无人机等全场景。平台具备高仿真、智能化等技术优势，广泛落地于教育、能源、运营商、金融等行业，助力人才培养与攻防能力提升。未来，绿盟科技将持续拓展新兴领域，推动网安人才高质量发展。

关键词：网络安全仿真平台 人才培养 攻防演练 高仿真

1. 背景

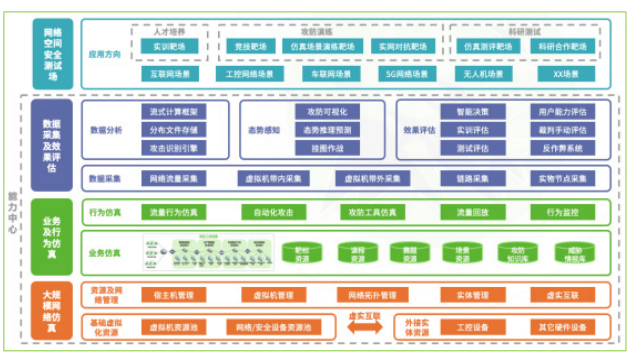
数字化时代下，网络安全作为网络强国、数字中国的底座，是我国数字化转型发展中不可或缺的要素。在网络安全实战上，攻防对抗背后都是人与人的较量，习总书记明确指出：“人才是第一资源，网络空间的竞争，归根结底是人才的竞争”；2019 年 9 月，工信部发布《关于促进网络安全产业发展的指导意见（征求意见稿）》，意见中指出：“把网络安全职业人才队伍日益壮大作为发展目标之一”；2023 年 1 月，国家能源局印发《2023 年电力安全监管重点任务》，任务要求：“推进国家级电力网络安全靶场建设，组织开展年度攻防演练。”

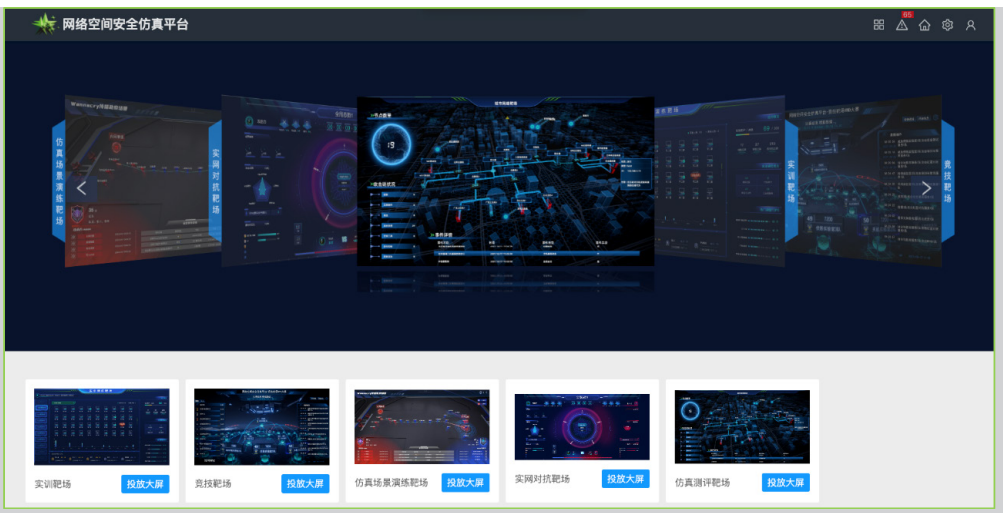
由此看出，我国网络安全人才问题受到了空前的重视，而网络安全靶场，作为高质量网安人才培养的最佳途径，已经上升到国家战略高度。

2. 绿盟科技网络安全靶场

绿盟科技凭借多年对安全领域的深入理解与前沿探索，打造出集实训、竞技、科研测试等能力为一体的网络安全靶场平台——绿盟网络空间安全仿真平台。

绿盟网络空间安全仿真平台，通过虚拟化、虚实结合、数字孪生、安全编排、行为及流量仿真、效果评估、智慧安全知识图谱、人工智能等技术，构建互联网、5G、工控、无人机等仿真场景，并对场景中生成的用户和攻防行为进行评估分析，满足用户对人才培养、安全竞赛、应急演练、实战对抗、安全测评、技术研究等方面的需求。




网络空间安全仿真平台
实训靶场
投放大屏
竞技靶场
投放大屏
仿真场景演练靶场
投放大屏
实网对抗靶场
投放大屏
仿真测评靶场
投放大屏


Wannacry传播勒索场景
攻击
攻击服务器
攻击服务器
攻击服务器
60 分
红队
队员：张三、李四
45 分
蓝队
队员：王五、赵六
安全事件列表

绿盟网络空间安全仿真平台具有以下特点：

（1）全场景:覆盖实训、竞技、攻防对抗、APT 应急、仿真测试、科研等技术领域及工控、无人机、卫星等新领域，满足用户的不同业务场景需求；

（2）高仿真：通过虚实结合、数字孪生、知识图谱等技术，对全场景下的设备、装置、系统等进行 1:1 仿真，同时结合自动化网络攻击工具和脚本，实现高逼真的网络攻击模拟；

（3）智能化：基于深度学习、ATT&CK 等技术，在仿真平台内实现智能决策和智能化攻击,可提升演训水平、降低演练难度和成本。

同时，绿盟网络空间安全仿真平台，在教育、能源、运营商、金融等领域，进行了大量落地实践，并取得了一定的应用成果。

3. 行业落地案例

在教育行业：以实训 + 课题研究为重点的靶场，通过内置多元化的理论和实验课程，帮助学生建立扎实的网络安全知识体系，以及各种网络安全攻防演练和漏洞挖掘的动手能力。此外，高校不仅服务于日常实验教学工作，还承载着多种教学实践和课题科研应用场景，绿盟网络空间安全仿真平台不仅提供丰富的课程体系，还为科研人员提供先进的实验平台和工具，支持教师们开展网络安全领域的前沿课题。

在能源行业：以攻防演练为重点的靶场，通过内置理论赛题、CTF 赛题、CFS 赛题、AWD 赛题以及 APT 应急场景，可快速提升竞技双方的实战能力。绿盟网络空间安全仿真平台具有完备的攻防演练能力，能够根据实际场景，自定义各种赛题，同时还可以对一场竞赛进行全生命周期管理，包括竞赛前、竞赛中、竞赛后三个阶段，支持多种形式的态势展示、过程监控和环境运维，并对选手提交的答案进行实时评估，同时对竞赛整个过程进行反作弊监控。

在运营商行业：以人才培养、攻防演练和攻防技术试验验证为重点的靶场，通过体系化的实操实训课程和能力考核机制设计，让学员能够由浅入深地学习掌握网络攻防技能。基于靶场自身强大的异构网络虚实结合场景仿真能力，开展红蓝对抗演练，让学员在实战中锻炼和提升自己的技能水平。通过丰富的漏洞库、情报库、场景库、流量仿真库、攻防工具库等资源，为用户开展网络安全技术、攻防技战术研究提供多元化的能力支撑。

在金融行业：以人才培养和仿真测评为重点的靶场，通过体系化、进阶式的攻防实战课程体系设计，让学员能够系统性地学习和掌握攻防技战术和攻防工具的应用，有效提升学员在攻防竞赛、攻防演练和安全运营保障等工作中的技能水平。通过靶场的漏洞仿真环境构建，用户能够开展验证、评估、修复于一体的漏洞安全运营管理工作，缓解并降低漏洞安全风险，有效提升网络安全运营管理水平。

4. 展望

未来，绿盟科技将继续在网络安全靶场投入，重点在工业互联网、无人机、卫星、车联网等创新领域进行布局，为更多的行业用户提供一站式解决方案，助力网安人才高质量培养。

网络安全政策导读（热点追踪）

绿盟科技 总体技术部 林涛

栏目说明：

本专栏基于绿盟科技政策研究团队在网络安全政策法规方面的日常跟踪，筛选国内外当期热点政策法规文件，并重点结合网络安全产业发展，对其内容和影响等进行简要分析。

更多内容敬请关注微信公众号：“网络安全罗盘”和“绿盟科技”。

国内篇

1.《国家网络身份认证公共服务管理办法》

【内容概述】5 月 23 日公安部、网信办等六部门联合发布。《办法》共 16 条，主要规定了四个方面内容：一是明确了国家网络身份认证公共服务及网号、网证的概念、申领方式；二是明确了使用国家网络身份认证公共服务的效力、应用场景；三是强调了国家网络身份认证公共服务平台、互联网平台等对数据安全和个人信息保护的责任；四是对未成年人申领、使用国家网络身份认证公共服务作出特殊规定。《办法》自 2025 年 7 月 15 日起施行。

https://www.mps.gov.cn/n6557558/c10087550/content.html.

【专家看法】

（一）背景分析

我国网络身份应用实践中，有几个较为紧迫的问题。一是互联网平台对用户个人信息的超范围采集问题；二是用户使用不同互联网服务时，需要重复注册的问题；三是对于互联网平台采集个人信息行为缺乏统一的监督平台问题等。而推进这些问题的解决，无疑应成为《办法》出台的最根本考量。

（二）主要内容

《办法》主要对涉及的三类主体相关行为进行了规定。一是自然人用户，主要规定了如何申请和使用“网证”“网号”；二是互联网平台（互联网服务提供方），主要规定了互联网平台应提供同等服务、保护用户个人信息等义务；三是公共服务平台（国家网络身份认证平台）及其建设运营者的保护用户个人信息义务。此外，还规定了主管部门的有关职责。总体来看，对第一类和第三类主体的相关规定占比较大。

（三）重要变化

与《办法》（征求意见稿）相比，有两处重要变化。

一是突出国家网络身份认证公共服务的自愿性。体现在：不再明确规定相关部门“推广应用和监督管理”的职责（第三条），有助于减少社会对于“强制推广使用”的疑惑；增加应当保留其他合法验证方式的规定（第六条），避免造成“二选一”的误读。

二是加强网络和数据安全保护。增加了公共服务平台的安全合规要求，包括网络运行安全、事件报告、应急预案等，以此强化服务平台自身对于用户信息保护的能力和水平。

（四）影响思考

一是“自愿使用”与“规模应用”的博弈。从互联网用户的角度看，采用公共服务固然能更加安全便捷，但面临短期内采用国家网络身份认证服务的互联网应用数量尚少，而且用户个人信息也面临暴露面增加的风险。如何解决这一关键矛盾，就成为后续制度实施需要关注的核心问题之一。

二是从产业影响来看，《办法》的出台实施，将进一步强化行业对个人信息保护的重视，或将对个人信息处理合规审计、个人信息保护咨询规划等带来促进。同时，也将进一步推动现行个人信息相关保护制度，如个人信息审计、数据安全评估等的衔接和协同发展。

2.《政务数据共享条例》

【内容概述】6 月 3 日国务院发布。《条例》旨在推进政务数据安全有序高效共享利用，提升政府数字化治理能力和政务服务效能，全面建设数字政府。《条例》包括总则、管理体制、目录管理、共享使用、平台支撑、保障措施、法律责任及附则共 8 章 44 条。《条例》自 2025 年 8 月 1 日起施行。

https://www.gov.cn/zhengce/content/202506/content_7026294.htm?sessionid=947902253.

【专家看法】

《条例》对政务数据共享所涉及的管理协调、权责要求等重要问题做出了体系化规定，是我国数据共享制度要求在政务数据领域的具体落地。其在彰显政务数据在整个数据资源体系中的重要位置的同时，也能为其他领域、行业数据共享的规范化、法治化提供重要法规依据。

（一）《条例》的三个重要问题

一是适用范围

《条例》适用于政府部门之间的政务数据共享及相关的安全监管，而非面向全社会。这是《条例》与《政府信息公开条例》的最大区别之一。此外，就政务数据的范畴来看，其与政府信息、公共数据资源等概念相比，在具体含义、承载形式、共享 / 公开目的等方面也存在较多差异，需要注意区分。

二是管理机构：一个重要部门

《条例》规定“国务院政务数据共享主管部门”负责统筹推进全国政务数据共享工作。对于该共享主管部门的具体指代，条例未给出进一步明确。而从部委职能定位以及我国政务数据管理工作实践来看，此职责通常由国务院办公厅承担，其下设有专门的政务服务机构，各地方也大多在各自办公厅（局）下设政务数据管理的专职部门。

从《条例》规定来看，各级数据共享部门除了履行常规监督职责外，还有两项重要职责，即明确同级政府的数源部门，以及审核同级政府部门报送的政务数据目录。而国家层面的“国务院数据共享主管部门”则还承担制定“政务数据目录编制标准规范”

的总体职责。

三是管理抓手：两个重要机制

对于推进政务数据共享的常态化管理，《条例》明确了“目录管理”“平台支撑”两个重要工作机制。

“目录管理”是指“政务数据目录”。该目录包括国家级、地方级两类，目录由各级政府部门按统一标准制定，制定时需开展保密风险、个人信息保护影响等评估，并经其部门负责人审核；制定后还需报同级共享主管部门审核。

“平台支撑”是指“国家政务大数据平台”。根据《条例》规定，该平台应是“全国一体化政务大数据体系”的重要载体和组成部分，并共同隶属于国家“数据基础设施”的范畴。这对于明确政务数据共享平台、数据基础设施之间的衔接和协同关系具有重要意义。

（二）《条例》对安全的要求

《条例》将政务数据共享的主体分为四类：政务共享管理部门、数据提供部门、数据需求部门、平台建设部门。结合政务数据共享工作流程，相应涉及安全相关的要求可以大致分为以下几类。

一是保密风险、个人信息保护影响等评估。该要求主要面向政务数据共享目录编制单位，包括数据提供部门、数据需求部门等。

二是督促落实安全管理责任。主要是授权政务共享管理部门按照“谁管理谁负责、谁使用谁负责”的原则，督促落实政务数据共享过程中的主体管理责任。

三是对数据的安全防护和风险监测。主要面向数据提供、数据需求部门，具体要求包括数据分类分级制度、保障政务数据安全的技术和其他必要措施、事件预案、应急处置等。

四是政务数据平台安全。主要面向政务数据共享平台建设管理单位，具体要求主要涉及平台的安全稳定运行、平台上政务数据的安全等。

（三）影响和后续关注

《条例》将于 2025 年 8 月 1 日起施行。“政务数据目录编制标准规范”作为各级政务数据目录编制的统一依据，无疑会受到政府部门乃至社会各界的重点关注。

对于网络和数据安全行业而言，除了关注《条例》提出的各项安全要求可能带来的潜在市场机会，也需重点强化对制度之间衔接、关联的学习和理解。尤其是对《条例》与全国一体化大数据中心、国家数据基础设施体系等战略规划关系的把握，这对于业务布局和市场开拓都有重要的指导意义。

3.《2025 年护航新型工业化网络安全专项行动方案》

【内容概述】7 月 1 日工信部发布。《方案》明确了三大重点任务方向：一是在提高企业防护水平方面，工信部门将梳理建立 2025 年工业领域网络安全防护重点企业清单，并指导清单企业落实工业互联网安全分类分级管理，特别将提升重点车联网平台安全防护能力，并指导企业根据车联网平台级别，采取相应的安全技术措施，定期开展网络安全符合性评测和风险评估。二是在增强工业控制系统产品安全能力方面，深化工业控制系统网络安全评估，研究制定评估实施指南。推动 PLC、DCS 等重点工业控制产品网络安全标准研制，并探索开展自愿性网络安全检测认证。三是在创新赋能模式方面，持续开展“安全深度行”活动，推广应用优秀案例与方案，并探索

网络安全保险新模式。组织全国性政策标准宣贯等。

https://wap.miit.gov.cn/jgsj/waj/wjfb/art/2025/art_62e699f6183e4e4fbac700a1a1f0bc86.html?sessionid=-748476320

【专家看法】

（一）政策脉络简析

1. 新型工业化战略

新型工业化是实现中国式现代化的重大战略路径。新型工业化道路最早由党的十六大提出并历经发展完善，党的二十大明确了到 2035 年“基本实现新型工业化”的战略目标。2023 年国家专门召开全国新型工业化推进大会，全面推进新型工业化目标任务的贯彻落实。

新型工业化的要求主要包括产业创新、制造业数字化转型、工业绿色发展、产业结构优化升级、提升产业链供应链韧性和安全水平等方面。

2.2024 年“护航新型工业化网络安全专项行动”方案

从政策关联信息来看，工业和信息化部 2024 年曾印发过《护航新型工业化网络安全专项行动方案》（以下简称《2024 方案》）。部分省市还据此开展了工业互联网安全相关行动，如《天津市护航新型工业化网络安全专项行动实施方案》（2024.7）、《北京市 2024 年工业互联网安全深度行活动实施方案》（2024.7）等。但《2024 方案》文件本身未见公开发布。

（二）《方案》回答了“是什么”：明确了新型工业化的网络安全内涵框架

如何界定新型工业化网络安全，是开展相关工作需要解决的首要问题。对此，《方案》明确将新型工业化网络安全细化分解为“企业网络安全防护”“工控系统产品安全能力”“工业领域网络安全服务”三大体系。其中：

“企业网络安全防护”是从对重点企业的外部防护入手，强化对新型工业化主体的网络安全监管和保护。

“工控系统产品安全能力”是从对重点企业的内部安全入手，强化应用和供给两个层面的安全能力。

“工业领域网络安全服务”是从对重点企业的网络安全外部赋能入手，调动外部专业力量加强对重点企业的网络安全服务和协助。

（三）《方案》回答了“怎么做”：明确了新型工业化网络安全的年度重点工作

首先，在“企业网络安全防护”方面，《方案》明确了 2025 年度强化企业网络安全防护的三项重点工作。一是各地确定 2025 年“工业领域网络安全防护重点企业清单”，并按照《工业互联网安全分类分级管理办法》完成定级备案；二是组织开展《工业互联网安全》系列国家标准的贯标达标试点工作；三是各地督促指导车联网相关企业开展车联网网络安全防护定级备案、网安测评和风险评估工作。

其次，在“工控系统产品安全能力”方面，《方案》明确了 2025 年度强化企业网络安全能力的两项重点工作。一是按照《工业控制系统网络安全防护指南》要求，对重点行业企业开展工业控制系统网络安全评估工作；二是开展重点工业控制产品的网络安全标准研制、检测认证工作。

再次，在“工业领域网络安全服务”方面，《方案》明确了

2025 年度强化对企业网络安全服务赋能的四项重点工作。一是支持专业机构开展重要工业控制系统识别、网络安全现场诊断、风险监测、漏洞修复、通报处置等服务；二是地方主管部门加强企业网络安全风险在线监测、点对点威胁通报；三是持续开展安全深度行、网络安全保险方案遴选等护企惠企活动；四是组织开展形式多样的全国性新型工业化网络安全政策宣贯。

（四）思考

难点思考。当前正处于专项行动的推进阶段（7-10 月），相关工作实施成效无疑将成为各界关注的焦点。而事关各方积极性的诸如“支持”和“鼓励”措施如何落地、对典型案例遴选有哪些具体标准等问题，或更成为影响专项工作成效的关键要素。

行业机会。对于网络安全行业而言，《方案》的发布实施也进一步明确了新型工业化对网络安全的重点需求。包括工业互联网安全定级备案和测评、工业控制系统网络安全评估、现场诊断、风险监测、漏洞修复、通报处置等。此外，诸如重点工业控制产品网络安全标准研制等工作，也可为行业赢得一定市场先机。

4.《国家数据局综合司关于公布 2025 年可信数据空间创新发展试点名单的通知》

【内容概述】7 月 16 日国家数据局发布。按照《国家数据局综合司关于组织开展 2025 年可信数据空间创新发展试点工作的通知》（国数综资源〔2025〕46 号）工作安排，经申报推荐、专家评审、名单公示，确定了 2025 年可信数据空间创新发展试点名单。公布名单显示，2025 年可信数据空间创新发展试点项目共 63 个，其中城市可信数据空间方向 13 个，行业可信数据空间方向 22 个，企业可信数据空间方向 28 个。

https://www.nda.gov.cn/sjj/zwgk/tzgg/0716/20250716205501357780564_pc.html

【专家看法】

（一）时间线回顾

本次试点于 2025 年 4 月 3 日 -24 日面向各省、自治区、直辖市及计划单列市、新疆生产建设兵团数据管理部门、有关中央企业开展了申报材料征集。此后经答辩和评审，于 7 月 8 日 -14 日对初选名单进行了公示。从最终公布的名单来看，与公示名单保持一致。

（二）试点特点分析

一是试点聚焦三类可信数据空间。按照试点通知及最终试点名单的情况来看，本次试点只围绕企业数据空间、行业数据空间、城市数据空间三个类型，而不是全面试点各种不同类型的可信数据空间。这也体现了可信数据空间建设的分阶段、分步骤推进的思路。

二是试点主体的性质具有多样性。入选的三类可信数据空间试点单位，除了城市可信数据空间入选单位有一定特殊性（多为属地数据运营平台公司）之外，企业可信数据空间、行业可信数据空间试点的入选主体，涵盖了科研平台、国有企业、民营企业等多种性质。这有利于调动各类不同主体积极性、发挥不同主体的特点和优势，有助于保持数据空间发展的活力。

三是强调试点的过程管理。按照试点要求，入选为试点仅仅是试点工作的开始而非终结。在为期 2 年的试点期内，试点单位还将接受国家数据局牵头开展的动态评估和动态管理，有推进不力或滞后等情况时将面临警示惩戒或约谈，若存在严重不及预期、重大风险或验收不合格情形的，还将被取消试点资格。可见，试点

过程实际也是“双刃剑”“试金石”，不仅考验试点单位的建设推进工作，也同样是对管理单位监管、评估能力的考验。

（三）试点工作的发展走势分析

首先，从政策关联来看，可信数据空间或成为数据基础设施建设的“桥头堡”。按照《国家数据基础设施建设指引》，国家数据基础设施的核心就是“数据流通利用基础设施”，而“可信数据空间”则是“数据流通利用基础设施”的核心和关键构成要素，是打造数据流通关键技术、产品集成方案和有效协同的载体平台。因此，其试点推进对于整个数据基础设施工作具有重要实践价值。

其次，试点或将在深化的基础上进一步拓展领域。试点工作的依据是《可信数据空间发展行动计划（2024—2028 年）》（国数资源〔2024〕119 号）。该行文件作为我国推进可信数据空间建设发展的纲领性文件，不仅设定了到 2028 年建成 100+ 可信数据空间的发展总目标，还规划了包括企业、行业、城市、个人、跨境在内的 6 类可信数据空间的培育。

可见，本次发布的 63 个可信数据空间试点，无论在数量上还是在涵盖类型上，都有进一步深化拓展的必要和客观需求。未来，个人可信数据空间、跨境可信数据空间无疑将成为扩大试点的重要领域方向。

国外篇

1. 欧盟推出欧洲漏洞数据库，以提高其数字安全性 (EU launches a European vulnerability database to boost its digital security)

【内容概述】5 月 13 日欧盟委员会发布。欧洲漏洞数据库（European Vulnerability Database，EUVD）由欧盟网络与信息安全局（ENISA）推出，是欧盟范围内统一的漏洞信息共享平台，其面向政府、私营企业、ICT 供应商及安全研究人员等开放访问。

https://digital-strategy.ec.europa.eu/en/news/eu-launches-european-vulnerability-database-boost-its-digital-security

【专家看法】

（一）背景概述

近年来，网络安全威胁日益严峻，漏洞管理成为保障网络安全的关键环节。欧盟为了提升网络安全标准，加强漏洞管理与风险防控，推出了欧洲漏洞数据库（EUVD）。该数据库是欧盟《网络与信息系统安全指令 2》（NIS2）的产物，旨在提供透明、可信且可操作的安全漏洞数据。其整合来自欧盟成员国、行业威胁研究和其他数据库的信息，确保来自多个来源的公开信息高度互联，增强欧盟内部的信息共享，提高欧盟在漏洞管理与风险防控方面的整体能力。

（二）漏洞库主要情况

EUVD 的主要功能包括：一是提供漏洞信息，能够提供全面、及时、准确的漏洞信息，包括高危漏洞和已被利用的漏洞，使企业和机构能够迅速采取缓解措施，增强对网络安全威胁的响应速度和效率。二是整合多源信息，能够利用开源软件 Vulnerability-Lookup 快速关联公开数据库、各国 CSIRT 发布的通告、厂商的补丁建议等多个已知来源的漏洞信息，从而实现信息的高度互联。三是实时更新漏洞信息，EUVD 几乎能够实时更新漏洞信息，并在网站顶部突出显示关键漏洞和已利用的漏洞，方便用户及时获取最新信息。四是提供三个仪表盘视图，三个仪表盘分别对应严重漏洞（Critical）、已被利用漏洞（Exploited）和欧盟计算机安全事

件响应小组（CSIRTs）的网络协同处理漏洞（Coordinated），方便用户根据需求进行筛选和查看。五是详细记录漏洞信息，每条漏洞记录均包含漏洞描述、受影响产品与服务、漏洞等级、利用方式，以及可用的修复或缓解方案，形成了从“风险预警”到“解决方案”的闭环，进而缩短企业响应周期。

（三）影响简析

该数据库的发布正值全球对通用漏洞披露 MITRE CVE（Common Vulnerabilities and Exposures）项目未来资金和可持续性存在不确定性的时期。它的建立为欧盟提供了一个独立且互补的漏洞信息平台，或有助于减少对单一系统的依赖，增强区域网络安全的自主性和韧性。

2. 美国《持续强化国家网络安全并修订第 13694 号和第 14144 号行政令》*(Sustaining Select Efforts To Strengthen the Nation’s Cybersecurity And Amending Executive Order 13694 and Executive Order 14144)*

【内容概述】6 月 6 日美国白宫发布。该行政令对拜登政府时期的 14144 号行政令（《关于加强和促进国家网络安全的行政令》）和奥巴马政府时期的 13694 号行政令（《关于阻断从事重大恶意网络活动人员财产的行政令》）进行了修订，旨在通过聚焦针对外国网络威胁的关键防护措施和强化安全技术实践来加强美国的网络安全防御体系。主要修订内容涉及推进安全软件开发、防范网络互联劫持行为、发展后量子密码技术、更新加密协议、调整 AI 网络安全工作重点、颁布网络安全政策、限定网络制裁适用范围等方面。

https://www.whitehouse.gov/presidential-actions/2025/06/sustaining-select-efforts-to-strengthen-the-nations-cybersecurity-and-amending-executive-order-13694-and-executive-order-14144/

【专家看法】

（一）政策背景

一是为应对全球网络安全威胁。近年来，网络攻击手段持续演进且愈发复杂，针对关键基础设施、政府机构、企业及个人的网络攻击事件频繁发生，对其国家安全、经济发展与社会稳定造成威胁。同时，美国认为俄罗斯、伊朗、朝鲜等国的网络威胁也相当严重，指控这些国家通过国家支持的黑客组织对美国发动网络攻击，威胁了美国的国家网络安全。

二是现有政策存在滞后。第 13694 号行政令由奥巴马于 2015 年签署、第 14144 号行政令由拜登政府于 2025 年 1 月签署，随着网络威胁形势的变化，其部分内容已难以有效应对当前复杂的网络态势，或部分条款被认为过于宽泛或存在执行漏洞，未能达到预期的效果。特朗普政府认为，现有行政令在应对新兴技术威胁（如 AI 驱动的攻击、量子计算对加密的威胁）和执行等方面存在不足，因此，亟须对现有行政令进行修订。

（二）主要内容

一是推进安全软件开发。指示联邦政府推进安全软件开发，要求在 2025 年 8 月 1 日前，美国商务部同国家标准与技术研究院（NIST）组建网络安全联盟，并制定指南，以展示基于《安全软件开发框架》*（Secure Software Development Framework）*的最佳实践。

二是防范网络互联劫持行为。指示各部门和机构采取行动，阻止网络互联劫持行为，以保护网络通信的安全性和可靠性，防止恶意攻击者通过劫持网络流量获取敏感信息或干扰网络服务。

三是发展后量子密码技术。指示各部门和机构在后量子密码方面采取行动，防范下一代计算架构可能的威胁。要求在 2025 年 12 月 1 日前，国土安全部(DHS)、网络安全和基础设施安全局(CISA)等部门联合发布并定期更新后量子密码产品类别清单，国家安全局(NSS)、预算管理办公室（OMB）等部门协同制定传输层安全协议 1.3 版以及后续版本。

四是更新加密协议。要求采用新的加密协议，以提升数据传输和存储的安全性，确保敏感信息在数字化环境中的保密性和完整性。

五是调整 AI 网络安全工作重点。要求将 AI 网络安全工作的重点转向漏洞识别和管理，而非审查，以更有效地利用 AI 技术来发现和修复网络安全漏洞，提高网络防御能力。

六是颁布网络安全政策。指示颁布网络安全政策，包括机器可读的政策标准和物联网信任政策规定，确保美国民众了解其个人和家用设备符合基本的安全工程原则。

七是限定网络制裁适用范围。将网络制裁的适用范围限定于国外恶意行为者，防止其被国内政治对手滥用，并明确制裁不适用于与选举有关的活动，以避免网络制裁措施在国内政治中的不当使用。

（三）影响

该行政令是对美国网络安全战略的重要调整。对美国自身而言，通过推动安全软件开发、防范网络劫持、发展后量子密码技术等措施，旨在强化美国国家网络安全关键能力。对国际关系而言，行政令明确将俄罗斯、伊朗、朝鲜等国列为网络威胁源，无疑会加剧国际关系的紧张局面，进而影响国际间的贸易、技术等方面的合作。

3.《赢得人工智能竞赛：美国人工智能行动计划》*(Winning the AI Race: America’s AI Action Plan)*

【内容概述】

7 月 23 日美国白宫发布。该文件是对特朗普政府“第 14179 号”行政令中关于“制订一项人工智能行动计划”的指示要求的贯彻落实，旨在通过全面战略部署，确保美国在全球人工智能领域的绝对主导地位。该文件围绕“加速人工智能创新、建设美国人工智能基础设施、领导国际人工智能外交与安全”三大支柱，提出了 90 余项联邦政策行动建议。主要政策包括：出口美国人工智能产品、推动数据中心的快速建设、促进创新和采用、维护前沿模型中的言论自由等。

https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

【专家看法】

（一）背景简述

2025 年 1 月，特朗普政府上台后，将人工智能领域的竞争列为国家安全核心议程，迅速推翻了拜登政府时期“安全优先”的人工智能监管思路，转向“创新主导”战略。

其核心举措包括四项：一是撤销拜登政府时期“有害的人工智能政策”；二是宣布实施“星际门”投资计划，建设美国新一代 AI

基础设施；三是发布新的人工智能新政总纲领《关于消除美国在人工智能领域领导地位的障碍的行政命令》（“第 14179 号”）；四是调整既有人工智能政策与行动，拟借助资源倾斜和政策松绑，力图推动美国 AI 行业发展，保持美国在人工智能全球竞争中的领导地位。

（二）内容要点

《行动计划》明确了“创新”“基础设施”以及“国际外交与安全”三大支柱，主要内容包括：

支柱一：加速人工智能创新。《行动计划》将解除监管约束视为刺激创新的重要手段，力图通过多管齐下的政策工具重塑创新环境、降低创新门槛。该部分提出了 15 项联邦政策行动，包括：消除繁文缛节和苛刻监管，确保前沿人工智能保护言论自由和美国价值观，鼓励开源和开放权重人工智能，推动人工智能应用，构建世界级科学数据集，投资人工智能可解释性、控制性和稳健性突破，构建人工智能评估生态系统等等。

支柱二：建设美国人工智能基础设施。《行动计划》将人工智能发展依赖的硬件基础列为投入重心，从能源、数据、芯片、劳动力多方入手，使人工智能硬件基础建设获得空前政策倾斜；同时还对人工智能基础设施安全作出明确规定。该部分提出了 8 项行动计划，包括：在保障安全的同时简化数据中心、半导体制造设施和能源基础设施的审批流程，发展与人工智能创新速度相匹配的电网，恢复美国半导体制造业，为军事和情报界建设高安全性数据中心，培养人工智能基础设施所需的熟练劳动力，加强关键基础设施网络安全，推广安全设计的人工智能技术和应用，提升联邦政府应对人工智能事件的成熟能力等。

支柱三：在国际人工智能外交和安全中发挥领导作用。《行动计划》提出构建以美国技术为核心的全球人工智能生态网络，通过输出包括硬件、模型、软件、应用及标准在内的完整技术栈，与愿意加入美国人工智能联盟的国家建立合作，还特别提到对抗中国影响力。该部分提出了 7 项行动计划，包括向盟友和伙伴出口美国人工智能，在国际治理机构中对抗中国影响力，加强人工智能计算出口管制，堵塞现有半导体制造出口管制中的漏洞等。

（三）特点与影响

《行动计划》表面高举“自由创新”旗帜，实则充满政策悖论。

一是“区别对待”的创新支持政策。在美国国内，特朗普政府承诺解除监管约束，却将政策红利过度向头部企业集中，挤压了人工智能初创企业的生存空间和创新动力。

二是“路人皆知”的安全关注。《行动计划》对人工智能安全的关注重点明显偏向供应链安全和地缘竞争风险，打压潜在竞争对手的意味远高过对技术发展和应用安全本身的考量。

三是“唯我独尊”的国际图谋。在国际层面，核心矛盾在于既想图谋全球技术霸权，却刻意割裂最大市场中国；表面倡导“技术开放”与“全球联盟”，却处处构建歧视性出口管制，而即使“放行”部分技术产品的出口，却难洗清“跟踪定位”和“远程关机”等各种嫌疑。

总的来看，该计划的实施或将带来两方面的影响。一方面可能在短期内促进美国人工智能技术的发展，但因监管真空与市场结构扭曲的隐患存在，进而导致人工智能伪创新等新问题出现。另一方面，美国试图通过技术绑定盟友，迫使其他国家站队，或将进一步深化中美技术脱钩，进而加深全球人工智能生态的分裂与对抗。

大模型安全

绿盟安全评估和防护解决方案

提供全生命周期安全守护

融合多维安全能力

大模型备案
红队服务赋能

1 AI安全评估系统 AI-SCAN

- 生成内容合规评估
- 模型幻觉评估
- 模型越狱评估
- 提示词攻击评估
- 角色逃逸评估
- 模型反演攻击
- 模型后门扫描
- 组件漏洞检测
- .....

模型评估

2 AI安全一体机 AI-UTM

- 小模型赋能安全
- 内置安全围栏
- 输入输出合规检测
- 敏感内容安全代答
- 大模型key管控
- 大模型token管控
- 敏感数据防护
- 大模型安全审计

模型防护

大模型应用网关

3 大模型应用防火墙 WAF-LLM

- Web流量防护
- 提示词注入防护
- 供应链防护
- 内容安全防护
- API 防护

风险识别

应用防护

模型防护

deepseek　火山引擎　OpenAI　阿里云百炼　360智脑　阶跃星辰

文心千帆　Gemini　腾讯混元　Moonshot AI　智谱·AI　华为 盘古大模型

自建大模型

本地部署大模型场景　调用公网模型场景　大模型一体机场景　安全审计监管场景

THE EXPERT
BEHIND GIANTS
巨人背后的专家

多年以来，绿盟科技致力于安全攻防的研究，
为政府、金融、运营商、能源、交通、科教文卫等行业用户和各类型企业用户，
提供具有核心竞争力的安全产品及解决方案，帮助客户实现业务的安全顺畅运行。
在这些巨人的后面，他们是备受信赖的专家。

客户支持热线：400-818-6868

NSFOCUS 绿盟科技

车联网安全研究报告

「2025版」

THE EXPERT
BEHIND GIANTS
巨人背后的专家

多年以来，绿盟科技致力于安全攻防的研究，
为政府、金融、运营商、能源、交通、科教文卫等行业用户和各类型企业用户，
提供具有核心竞争力的安全产品及解决方案，帮助客户实现业务的安全顺畅运行。
在这些巨人的后面，他们是备受信赖的专家。

客户支持热线：400-818-6868

NSFOCUS绿盟科技