高等院校网络空间安全专业实战化人才培养系列教材

郭启全　丛书主编

数据安全管理与技术

杨正军　王安宇　郭启全　编著

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书围绕数据安全的综合管理与技术实践展开，系统阐述了从理论基础到行业应用的全方位内容。全书以数据全生命周期为核心，结合国内外法律法规与标准规范，探讨数据分类分级、风险评估、安全防护、个人信息保护及数据要素流通等关键议题，旨在为构建安全可靠的数字化生态提供理论指导与实践参考。

本书是高等院校网络空间安全专业实战化人才培养系列教材之一，可作为网络空间安全专业的专业课教材，适合网络空间安全专业、信息安全专业以及相关专业的大学生、研究生系统学习，也适合各单位各部门从事网络安全工作者、科研机构和网络安全企业的研究人员阅读。

图书在版编目（CIP）数据

数据安全管理与技术 / 杨正军等编著. -- 北京 : 电子工业出版社, 2025. 7. -- ISBN 978-7-121-50128-9

Ⅰ. TP309.2

中国国家版本馆 CIP 数据核字第 2025RY1844 号

责任编辑：刘御廷　　文字编辑：刘怡静

印　　刷：

装　　订：

出版发行：电子工业出版社

北京市海淀区万寿路173信箱　　邮编：100036

开　　本：787×1 092　1/16　印张：16.75　　字数：424千字

版　　次：2025年7月第1版

印　　次：2025年7月第1次印刷

定　　价：69.00元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至zlts@phei.com.cn，盗版侵权举报请发邮件至dbqq@phei.com.cn。

本书咨询联系方式：（010）88254569，liuyj@phei.com.cn。

高等院校网络空间安全专业
实战化人才培养系列教材

编委会

序 FOREWORD

在数字化智慧化高速发展的今天，网络和数据安全的重要性愈发凸显，直接关系到国家政治、经济、国防、文化、社会等各个领域的安全和发展。网络空间技术对抗能力是国家整体实力的重要方面，面对日益复杂的网络安全威胁和挑战，按照“打造一支攻防兼备的队伍，开展一组实战行动，建设一批网络与数据安全基地”的思路，培养具有实战化能力的网络安全人才队伍，已成为国家重大战略需求。

一、培养网络安全实战化人才的根本目的

在网络安全“三化六防”（实战化、体系化、常态化；动态防御、主动防御、纵深防御、精准防护、整体防控、联防联控）理念的指引下，网络安全业务越来越贴近实战。实战行动和实战措施都离不开实战化人才队伍的支撑。培养网络安全实战化人才的根本目的，在于培养一批既具备扎实的理论基础，又掌握高新技术和前沿技术、具备攻防技术对抗能力，还能灵活运用各种技术措施和手段，应对各种网络安全威胁的高素质实战化人才，打造“攻防兼备”和具有网络安全新质战斗力的队伍，支撑国家网络安全整体实战能力的提升。

二、培养网络安全实战化人才的重大意义

习近平总书记强调：“网络空间的竞争，归根结底是人才竞争”，“网络安全的本质在对抗，对抗的本质在攻防两端能力较量”。要建设网络强国，必须打造一支高素质的网络安全实战化人才队伍。我国网络安全人才特别是实战化人才严重缺乏，因此，破解难题，从网络安全保卫、保护、保障三个方面加强实战化人才教育训练，已成为国家重大战略需求。

当前，国家在加快推进数字化智慧化建设，本质是打造数字化生态，而数字化建设面临的最大威胁是网络攻击。与此同时，国家网络安全进入新时代，新时代网络安全最显著的特征是技术对抗。因此，新时代要求我们要树立新理念、采取新举措，从网络安全、数据安全、人工智能安全等方面，大力培养实战化人才队伍，加强“网络备战”，提升队伍的技术对抗和应急处突能力，有效应对新威胁和新技术带来的新挑战，为国家经济发展保驾护航。

三、构建新型网络安全实战化人才教育训练体系

为全面提升我国网络安全领域的实战化人才培养能力和水平，按照“理论支撑技术、技术支撑实战”的理念，创新高等院校及社会差异化实战人才培养的思路和方法，建立新型实战化人才教育训练体系。遵循“问题导向、实战引领、体系化设计、督办落实”四项原则，认真落实“制定实战型教育训练体系规划、建设实战型课程体系、建设实战型师资队伍、建设实战型系列教材、建设实战型实训环境、以实战行动提升实战能力、创新实战

型教育训练模式、加强指导和督办落实”八项重大措施，形成实战化人才培养的“四梁八柱”，有力提升网络安全人才队伍的新质战斗力。

四、精心打造高等院校网络空间安全专业实战化人才培养系列教材

在有关部门的大力支持下，具有 20 多年网络安全实战经验的资深专家统筹规划和整体设计，会同 20 多位部委、高等院校、科研机构、大型企业具有丰富实战经验和教学经验的专家学者，共同打造了 14 部技术先进、案例鲜活、贴近实战的高等院校网络空间安全专业实战化人才培养系列教材，由电子工业出版社出版，以期贡献给读者最高水平、最强实战的网络安全重要知识、核心技术和能力，满足高等院校和社会培养实战化人才的迫切需要。

网络安全实战化人才队伍培养是一项长期而艰巨的任务，按照教、训、战一体化原则，以国家战略为引领，以法规政策标准为遵循，以系统化措施为抓手，政府、高校、企业和社会各界应共同努力，加快推进我国网络安全实战化人才培养，为筑梦网络强国、护航中国式现代化贡献我们的智慧和力量！

郭启全

前言 PREFACE

数字化浪潮席卷全球，数据已成为驱动社会经济发展、推动科技进步、提升国家治理能力的关键生产要素。随着大数据、云计算、人工智能等技术的飞速发展，数据的采集、存储、处理、分析和应用能力实现了质的飞跃，为各行各业带来了前所未有的变革与机遇。与此同时，数据安全问题日益凸显，成为制约数字经济健康发展的重大挑战。数据泄露、数据滥用等安全事件频发，不仅侵犯了个人隐私权，也对国家安全、社会稳定、企业运营构成了严重威胁。筑牢数字安全屏障、保障数据安全是实现数字经济健康发展、建设数字中国的重要保障。因此，加强数据安全管理与技术研究，构建完善的数据安全保护体系，已成为时代赋予我们的紧迫任务。

进入新时代，网络安全最显著特征是技术对抗，应树立新理念，采取新举措，立足有效应对大规模网络攻击，认真落实“实战化、体系化、常态化”和“动态防御、主动防御、纵深防御、精准防护、整体防控、联防联控”的“三化六防”措施，按照“打造一支攻防兼备的队伍，开展一组实战行动，建设一批网络与数据安全基地”这条主线，加强战略谋划和战术设计，建立完善的网络安全综合防御体系，大力提升综合防御能力和技术对抗能力。从创新角度出发，按照“理论支撑技术、技术支撑实战”的理念，加强理论创新和技术突破，实施“挂图作战”；从“打造一支攻防兼备的队伍”出发，创新高等院校和企业差异化网络安全人才培养思路和方法，建立实战代人才教育训练体系，加强教育训练体系规划，强化课程体系、师资队伍、系列教材、实训环境建设和培养模式创新，培养网络安全实战型人才。

为了满足培养网络安全实战型人才的需要，郭启全组织成立编委会，共同编著高等院校网络空间安全专业实战化人才培养系列教材，包括《网络安全保护制度与实施》《网络安全建设与运营》《网络空间安全技术》《商用密码应用技术》《数据安全管理与技术》《人工智能安全治理与技术》《网络安全事件处置与追踪溯源技术》《网络安全检测评估技术与方法》《网络安全威胁情报分析与挖掘技术》《数字勘查与取证技术》《恶意代码分析与检测技术》《恶意代码分析与检测技术实验指导书》《漏洞挖掘与渗透测试技术》《网络空间安全导论》。郭启全统筹规划和整体设计全套教材，组织具有丰富网络安全实战经验和教学经验的专家、学者，撰写这套高等院校网络空间安全专业教材，并对内容严格把关，以期贡献给读者最高水平、最强实战的网络安全、数据安全、人工智能安全等方面的重要知识。

《数据安全管理与技术》旨在全面系统地探讨数据安全领域的核心问题，从基础理论、政策法规、技术实践到未来趋势，为读者呈现一幅数据安全技术和管理的全景图。通过深入分析数据安全的基本概念、核心要素、生命周期安全等基础知识，帮助读者建立起对数

据安全的全面认知；同时，结合国内外数据安全管理的实践经验与最新成果，介绍了一系列行之有效的数据安全政策标准和技术手段，为数据安全实践提供有力支撑。

本书由杨正军、王安宇、郭启全编著。在本书的撰写过程中，得到了众多行业内外专家学者的鼎力支持与帮助，感谢刘扬、杨阳、钟子修、孟洁、王志波、孙勇、王佳星、吴连涛、白晓媛、潘无穷等人为本书付出的辛勤努力！

由于作者水平有限，书中难免有不妥之处，敬请读者批评指正。

作者

目录 CONTENTS

第1章

绪论

第 2 章

数据安全管理体系

第 3 章

数据全生命周期安全保护

第 4 章

数据安全关键技术

第5章

个人信息保护

第6章

数据要素流通

第 1 章 绪 论

本章介绍数据和数据安全的基本概念、数据安全核心要素、数据全生命周期安全，数据安全保护制度，数据安全保护与网络安全等级保护的关系，数据安全有关政策和标准，数据安全管理目标、数据安全风险评估，阐述人工智能、云计算、物联网、零信任架构、量子计算等新技术与数据安全的关系，为读者深入学习数据安全管理与技术奠定基础。

1.1 数据和数据安全

科学与技术进步已深刻地影响了全球经济，改变了人们生产和生活的方方面面。20 世纪 50 年代，计算机和数字化技术的出现标志着第三次工业革命的开始，极大地推动了人类社会经济、政治、文化领域的变革，并且深刻影响了人类生活方式和思维方式。

当今世界已经进入第四次工业革命时代。第四次工业革命融合了人工智能（AI）、机器人技术、物联网（Internet of Things，IoT）、3D 打印、基因工程、量子计算和其他技术，在丰富人们生活的同时，也推动了工业的智能化转型，促进了数字世界、物理世界和生物世界的融合。

一项对全球物联网市场的研究显示，2019 年活跃的物联网设备（含计算机、服务器、手机等）数量约为 76 亿台。到 2030 年，全世界的联网设备数量将达到 241 亿台。各类智能终端设备开始出现在人们的生活中，如智能可穿戴设备、AR 眼镜、VR 头戴式设备，甚至是未来的全息投影设备等也将逐渐推广和普及。现代生活中的很多场景依赖于这些智能终端设备的支持。智能导航系统可以在驾车时迅速计算出到达目的地的可选路线；智能手机中的语音助手提供了丰富的人机交互方式，甚至可以代替机主在餐馆预订座位；手机中的社交应用可以识别出照片中的人脸。

随着网络空间和真实空间的进一步交融，物理、数字和生物世界之间的界限变得模糊，或者说是融合。在此场景中，数据的范畴得到了极大的外延。数据量急剧膨胀的同时，数据的重要程度和价值也越来越高，但数据的风险也迅速增长。面对不断变化、飞速演进的数字化世界，系统化地分析、识别数据安全的风险，有针对性地制定和实施各类安全控制措施，设计和实现安全架构，并持续性安全运营，是每个组织保护数据安全的关键。

1.1.1 数据安全基本含义

数据和数据安全是最基础的两个概念，什么是数据？什么是数据安全？从不同的维度，有很多种对于数据的定义。

1．数据的定义

数据（Data）是事实或观察的结果，是对客观事物的逻辑归纳，是用于表示客观事物的、未经加工的原始素材。根据《中华人民共和国数据安全法》（简称《数据安全法》）第三条“本法所称数据，是指任何以电子或者其他方式对信息的记录”，数据是信息的表现形式和载体，可以是符号、文字、数字、语音、图像、视频等。数据和信息是不可分离的，数据是信息的表达，信息是数据的内涵。数据本身没有意义，数据只有对实体行为产生影响时才成为信息。

在现代（1960 年后）计算机系统中，数据以二进制信息单元 0、1 的形式表示，通过计算机系统可以处理。

数据可以分为结构化数据和非结构化数据。结构化数据指符合数据模型的数据，具有明确定义的结构，遵循一致的顺序，并且可以由人或计算机程序轻松访问和使用。结构化数据通常以定义明确的模式存储，如传统的关系型数据库。结构化数据通常是表格形式的，具有明确定义其属性的行和列。SQL（Structured Query Language，结构化查询语言）通常用于管理存储在数据库中的结构化数据。

与之相对应，非结构化数据是未按预定义方式组织或不具有预定义数据模型的数据，因此不适用于主流关系型数据库。对于非结构化数据，存在用于存储和管理的替代平台。非结构化数据通常包含文本、语音、图片、视频、PDF 文档、媒体日志等格式，非结构化数据在 IT 系统中的应用越来越广泛，并可以用于各种商业智能和分析应用程序中。

2．个人数据和敏感个人数据

不是所有的数据都有相同的价值，因此，数据安全主要关注“需要受保护的数据”的安全防护。对一般组织和个人而言，两个相关的概念尤为重要：个人数据和敏感个人数据。

（1）个人数据。个人数据也称个人信息，是与个人身份有关的任何信息。个人数据的范围非常广泛，其含义在不同的国家和地区也略有不同，但个人数据保护的原则类似。值得关注的个人数据的两大类型为个人身份信息（Personally Identifiable Information，PII）和个人隐私信息。美国国家标准与技术研究院（National Institute of Standards and Technology，NIST）在 SP 800—122 行业标准中将个人身份信息定义为“由代理机构维护的有关个人的任何信息，包括：任何可用于区分或追踪个人身份的信息，如姓名、社会安全号码、出生日期和地点、生物特征如指纹和人脸特征；与个人链接或可链接的任何其他信息，如医疗、教育、财务和就业信息。”值得注意的是，在此定义下，互联网用户的 IP 地址为“链接的个人数据”。但是在欧盟，互联网用户的 IP 地址为个人数据。

（2）敏感个人数据。敏感个人数据是个人数据的子集，在不同的国家和地区对其范畴

有不尽相同的定义。一般而言，个人财务数据（如银行卡）、个人健康数据（如基因、病历卡）、个人生物识别特征（如面部特征、指纹）都被认为是敏感个人数据。处理敏感个人数据时，不仅要实施必要的安全防护机制，还需要满足适用的法律法规和监管要求。

3．数据安全

数据安全（Data Security）聚焦于在数据全生命周期过程中保护数据免受未授权的访问与数据损坏，并涵盖一整套相关的标准、技术、框架和流程。《数据安全法》第三条指出，“数据安全，是指通过采取必要措施，确保数据处于有效保护和合法利用的状态，以及具备保障持续安全状态的能力”。

数据安全的主要目的是保护在收集、存储、创建、接收或传输过程的数据。无论使用哪种设备、技术或流程来管理、存储或收集数据，都必须对其进行保护。合规性也是一个主要考虑的因素。

在“互联网 +”飞速普及并逐步渗透到各行各业的今天，所有的政府机构、企业每天都在处理着数据。数据安全领域面对的环境和业务包罗万象。但是，无论是处理大量个人和财务数据的银行系统，还是互联网厂商提供的云服务，亦或是在移动电话上存储用户照片，都是不同形态和内涵的数据的应用。事实上，信息与通信行业的本质，可以理解为聚焦于信息的生命周期，提供相应的服务。而信息的基础载体就是数据。因此，数据安全的目标可以抽象为保障信息生命周期的安全性（机密性、完整性、可用性）及数据处理的合理性（含用户隐私）。

1.1.2　数据安全核心要素

NIST FIPS 199（联邦信息处理标准出版物 199）《联邦信息和信息系统的安全分类标准》中提出，数据安全模型的三个核心要素为机密性（Confidentiality）、完整性（Integrity）与可用性（Availability），简称为 CIA 三角（CIA Triad），如图 1-1 所示。

数据安全模型有助于阐述数据安全和信息安全两个领域的关系。数据安全领域和信息安全领域有着千丝万缕的联系，在某些区分不严格的场合中，这两个术语甚至可以混用。信息安全一般指企业或组织的商业秘密、技术秘密和其他关键信息资产的防护，从一般含义上讲，信息安全更侧重数据安全模型中的机密性。

1．DIKW 金字塔模型

著名的 DIKW（Data, Information, Knowledge and Wisdom，数据、信息、知识和智慧）金字塔模型清晰地阐述了数据、信息、知识和智慧四个术语之间的关系，如图 1-2 所示。

（1）数据。最底层也是最基础的是数据，数据是以原始或未分类的形式表述的一系列事实。如果没有关联或上下文，数据本身的含义并不明确。例如，12122020 是一个数字的序列。如果加上表示日期的上下文，则很容易得到信息：2020 年 12 月 12 日。由此引出第二个概念：信息。

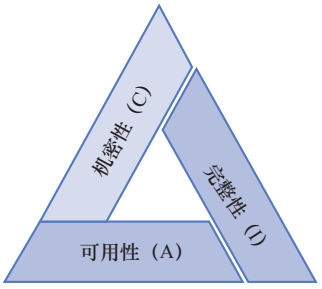

图 1-1　数据安全的三个核心要素

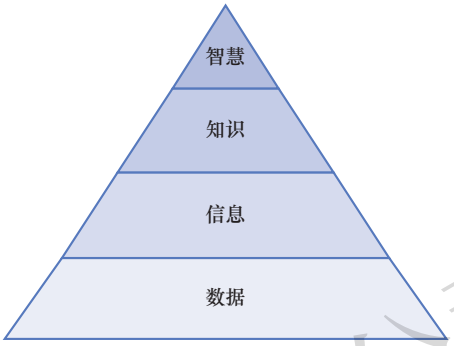

图 1-2　DIKW 金字塔模型

（2）信息。信息是已经排除错误并经过进一步处理的数据，可用于针对特定目的的测量、可视化和分析。取决于不同目的，数据处理可能涉及不同的操作，如组合不同的数据集，确保收集的数据相关且准确等，即基于目的的上下文所关联的数据产生信息。例如，3371.96、3373.28、3347.19，是一串离散的数据点。如果是已知这些数据点是证券市场上证综合指数某几天休市时的数据，则可以创建该特定时间段内数据点的图表来分析上证综合指数的表现。

（3）知识。信息从数据中获取关联，可用于解答“4W”的问题，即 Who（谁）、What（什么）、When（何时）、Where（何地）。不过一旦问题变为 How（如何），则进入更上一层的领域：知识。

从收集的数据中得出的信息与我们的目标如何相关呢？这些信息中的“各部分”如何与其他部分关联以增加更多的意义和价值呢？最重要的是，如何使用信息来实现目标呢？

当信息不仅是对收集的事实的描述，还可以被使用以实现目标时，信息即转变为知识。这些知识通常是组织和个人超越竞争对手的优势。利用信息和知识，发现潜在的或隐含的关联，则称为“见解”。

（4）智慧。当组织或个人使用从信息中获得的知识和见解做出主动决策时，即达到了 DIKW 金字塔模型的顶尖——“智慧”。智慧是知识在决策中的应用。智慧可以回答“Why”开头的问题，也可以解答“What”开头的问题。

根据信息安全和数据安全的概念辨析可知：按照信息和数据的范畴和内涵，数据的范畴和内涵更为通用和抽象；数据安全更聚焦于本领域的业务和技术，而信息安全聚焦于组织的信息甚至是知识的机密性，并聚焦于组织视角的场景与应用。

2．数据安全的三个核心要素

下面详细解释数据安全的三个核心要素的含义。

（1）机密性

机密性指确保只有获得授权的信息访问主体才可以获得指定的信息。信息访问主体可以是实际的用户，也可以是进程、App、服务等。

机密性的定义比较抽象，可以从反方向来直观理解。常见的数据机密性受损的案例包

括：手机丢失导致存储的照片泄露；互联网账户密码泄露导致账户被恶意登录和操作；含有报价等敏感信息的电子邮件被误发给其他人。

在维护信息机密性的过程中采取保护操作，是为了防止敏感信息在传递过程中出现误传或窃取等情况后被泄露。为此，必须使用一定的技术手段，使得对明文真实数据的访问仅限于有权查看该数据的人员。机密性的保护措施一般在信息传递之前实施。但是，还应做好数据泄露预案，包含如何识别数据泄露的数量和类型，并根据这些假想情形制定相应的善后措施。

确保机密性的常见方法包括数据加密、多因素身份认证。当然，对于极度敏感的数据，也可能需要采取物理隔离、非联网设备存储等措施。

除技术手段外，保护数据机密性通常还需要对处理重要数据的人员进行特殊培训。这类培训可以帮助被授权接触重要信息的人员熟悉潜在的风险因素，并实施有针对性的防范措施。

（2）完整性

完整性指确保数据在整个生命周期中不会受到非法的篡改与破坏。

在某些场景下，数据完整性是数据质量的代称，如数据库中不出现逻辑相反的两条记录；在某些场景下，数据完整性是数据损坏的反义词。在数据安全领域，数据完整性旨在防止数据被意外或蓄意地更改。

在设计、实现和应用任何涉及存储、使用和传输数据的系统时，数据的完整性都是需要考虑的关键问题。完整性意味着必须采取行之有效的措施，以确保数据不会在传输过程中被更改，同时也不会被未经授权的人员更改，如在信息机密性受到损害时。由于存储、检索或使用操作而导致的任何数据被意外更改，包括恶意篡改、意外的硬件故障和人为错误，都是数据完整性的故障场景。如果更改是由未经授权访问引发的，则也属于数据完整性的故障场景。

数据完整性的技术从目的上是相同的：确保按预期准确地记录数据，并且在以后检索时，确保数据与原始记录时的数据相同。

常见的确保数据完整性的措施包括文件操作许可控制和用户访问权限控制。此外，在日常的生产过程中，工作组织必须定期采取一些措施来检测由非人为的意外事件（如电磁脉冲或服务器崩溃）所带来的数据变化，这些检测措施包括但不限于检验数据的明文或加密后的校验。除定期检验外，还必须有稳定安全的备份才能确保将受影响的数据恢复到正确的状态。

（3）可用性

可用性指确保合法用户对数据的获取与使用能够得到保障。

任何用于存储、使用和传输数据的系统，应该在用户需要时，及时准确地提供数据。这意味着，存储、使用和传输的业务功能中，用于保护信息的安全控制及用于访问信息的通信通道必须能够正常运行。高可用性系统旨在始终保持可用状态，以防止因断电、硬件故障和系统升级而导致服务中断。确保可用性还涉及防止拒绝服务攻击（Denial of Service

Attack，DoS 攻击），如防止因将大量传入消息发送到目标系统，从而迫使该系统无法针对正常用户提供服务的情况。

严格维护硬件设施、定期执行硬件维修、维护操作系统环境与保持系统始终部署最新的安全补丁与更新是可用性得以维持的重要保障。当然，提供足够的通信带宽并防止出现性能瓶颈也同样重要。在预防措施做足的同时，提供充足且稳定的冗余、故障转移、独立磁盘冗余阵列（Redundant Array of Independent Disks，RAID，简称“磁盘阵列”），甚至高可用性群集等后发性保障措施，可以有效缓解硬件问题发生所带来的后果。最糟糕的情况发生时，快速和自适应的灾难恢复至关重要，而强大的恢复能力需要依靠一套全面的灾难恢复计划（Disaster Recovery Plan，DRP）。

DRP 的风险预估中必须包括不可预测的事件，如不可抗力带来的灾害。为防止此类事件造成数据丢失，备份副本应该尽可能保有多份并存储在地理位置相隔离的地点。除不可抗力外，对于可能遭受的拒绝服务攻击等恶意攻击，DRP 要求部署额外的安全设备或软件（如防火墙和代理服务器），以确保即便在攻击造成了短暂宕机的情况下也能足够迅速地恢复正常的服务。

在数据安全领域，可用性是保障系统的用户能够正常使用系统的关键所在。

CIA 三角的三个核心要素需要权衡一定的冲突。例如，共享驱动器的访问、网络代理服务器的配置、员工发送外部电子邮件的权限等场景，对于企业 IT 系统而言都存在机密性、完整性和可用性的冲突，因此需要不同团队（如网络运营、开发、事件响应和变更管理团队等）的协作以解决问题、应对风险、保障企业内的数据安全。

1.1.3 数据生命周期安全

伴随着数据规模剧增、大数据时代来临，物联网的持续发展与演进，数据生命周期管理（Data Lifecycle Management，DLM）变得越发重要。越来越多的设备正在世界上的每个角落产生海量的数据，识别不同数据的整个生命周期，并对数据进行适当的安全防护，同时维持数据的使用便捷，显现出其重要意义。

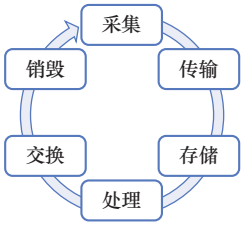

图 1-3　数据生命周期

按照国家标准 GB/T 37988《信息安全技术 数据安全能力成熟度模型》的定义，数据生命周期（Data Lifecycle），即数据从其最初采集（或创建），到传输、存储、处理、交换（传输），直到其使用寿命结束并被销毁的生命周期阶段序列，如图 1-3 所示。

1．数据生命周期的安全防护机制

数据生命周期管理（DLM）是管理整个数据生命周期的数据流的过程。在整个数据生命周期中准确地识别和保护数据，特别是敏感数据，是数据安全的基础要求。基于数据生命周期的分析还可以确定应用安全控制的位置，并选择合适的安全机制。不同数据类型的不同生命周期阶段，其每个阶段的安全防护关键点也不相同。

（1）在数据的采集阶段，不同的数据类型的采集方式可谓千差万别。例如，数字温度计采集办公室的室内温度数据，电信设备生成用户每次通话的计费账单，手机同步照片到云存储的创建方式都各有不同。但从数据安全的基本原则角度，在数据的采集阶段，应该对数据进行识别并做出合理的分类，还应明确该类数据的保护要求。

（2）在数据的传输阶段，可通过数据传输加密保障数据的机密性，通过网络基础设施及网络层数据防泄露设备的备份建设，实现网络的高可用性，从而保证数据传输过程的稳定性。

（3）在数据的存储阶段，与创建阶段类似，数据的存储方式也有很大的差异，可能是在本地存储设备上存储，或者采用分布式存储、云存储。本阶段应该实现必要的数据访问控制，以确保数据受到合适的保护，从而降低数据被泄露、被篡改和被破坏的风险。

（4）在数据的处理阶段，数据安全防护的关键点是将安全控制机制应用于使用中的数据。一般而言，应能够监控对数据特别是敏感数据的访问，并应用各类安全控制以确保数据的安全。

（5）数据的交换阶段，是数据安全控制中最薄弱的环节。而对于数据这一信息载体而言，类比于传统的实体商品，数据在交换时才能体现价值。因此在一定程度上，数据安全和数据价值会产生冲突。适度的安全控制对于平衡数据安全和数据价值十分重要。

（6）在数据的销毁阶段，需要关注不同法律法规对于数据归档的约束性要求。当数据的归档时限已过，应根据监管要求删除数据。此时，数据安全的关注重点是采取合理的数据销毁机制，以确保数据无法被恢复。即使数据并非以传统的文件类型存储，也需要考虑相应的销毁机制。

上述为数据生命周期的 6 个典型阶段。需要注意的是，不同的数据类型其生命周期阶段也不完全一致。比如，对于数字温度计的温度数据，以数字温度计的视角，只有采集和交换两个阶段。

2．数据的三种状态及其安全防护机制

随着数据访问界面不断复杂，攻击路径不断增加，攻击者的攻击策略变得越来越体系化和系统化，对数据安全管理而言，数据安全的各个方面都变得非常重要——从安全的数据存储、使用和传输，到访问控制，以及有效的密钥管理，如果有一个环节易于受到攻击，则会破坏整体安全机制的有效性。

若要应对来自上述多个维度的风险，则需要一种全面的、以数据为中心的安全防护方法。也就是说，应在数据生命周期的所有环节关注保护数据本身，而不是仅关注网络、应用程序或服务器。

数据在存储（At Rest）、使用（In Use）和传输（In Motion）三种状态时，都会面临独特的安全威胁和挑战，因此要针对这三种状态下的数据进行保护，如图 1-4 所示。

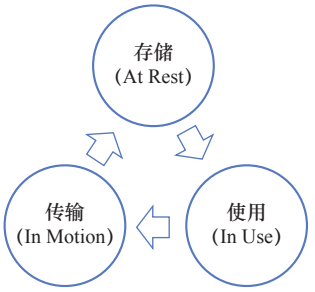

图 1-4　数据生命周期状态

（1）数据的存储

数据在硬盘驱动器上存储时处于静止状态。静态形式存储的数据指位于硬盘驱动器、外置存储、云存储或磁带机等位置的数据，又称静态数据。在这种相对安全的状态下，传统的分层防御机制可以起到作用。首先，是存储位置所在的物理设施，如机房、办公室的保护。其次，是基于外围的防御措施，如防火墙、入侵检测和防御系统的保护。再次，是基于端点的防御措施，如防病毒软件、端点策略防护、补丁管理等。最后，是数据本身的防御措施，包括加密和访问控制两大类型。

加密分为驱动器加密和文件加密。加密硬盘驱动器能防止硬盘丢失、被盗带来的数据安全风险，对安全销毁场景也有作用。但多数硬盘加密技术有缺陷，开机后磁盘内容会被解密，数据使用状态下不再生效。因此，部分含敏感数据的文件应采用应用级别的文件加密。对于结构化数据存储，很多数据库发行版本支持数据库、表甚至列级加密，可针对敏感数据和个人数据进行细粒度防护。

其他存储安全措施也可能对数据安全有帮助，如将不同的数据分类存储在单独的位置，以减少攻击者获得足够信息而进行欺诈或其他犯罪的可能性。

上面描述的是传统的主机存储、移动存储和服务器存储。对于云存储，还要关注其他方面的要求。例如，存储空间必须是在合同、服务水平协议和法规允许的地理位置，存储的数据必须要保证包括所有的副本和备份，以防止因数据丢失而造成的损失。例如，存储和使用受欧盟《通用数据保护条例》约束的电子健康记录，可能对数据拥有者和云服务提供商都是一种挑战。将数据存放在云端后要考虑其可靠性、保密性及完整性，考虑供应商的安全权限管理措施是否完善，对存储的数据是否进行安全管理，以及数据的存放格式是否合理。所以，若要保证在这一过程中的数据安全，就需要对数据进行额外的完整性保证、数据加密及数据隔离等措施。

综上所述，静态数据通常被认为是相对安全的数据状态。在此状态下，基于网络和物理边界的解决方案可以被视作第一道防线。根据数据本身的用途和敏感性，还可以添加额外的防线。数据的存储也需要注意应遵从相关的法律法规的约束。

（2）数据的传输

虽然数据在静止状态更容易得到完善的保护，但数据一直处在静止状态是无法带来真正的价值的。数据通过传输以提供给其他需求方，从而实现数据的共享和价值。数据在此

阶段往往最容易受到攻击。

云数据通过网络、进程通信等方式传输给其他的客户和虚拟服务以供其使用，由于网络的开放性，数据不能在没有安全控制的情况下进行传输。数据的创建者要考虑是否对数据进行管理权限的设置。所以若要保证传输数据的安全，就要使用数据加密技术以同时维护数据传输过程中的机密性及传输后的完整性。

相较于静态数据，传输中的数据更容易受到攻击，无论是通过公共网络或专用网络传输，还是利用其他传输机制。图 1-5 是几种典型的数据传输机制示例。

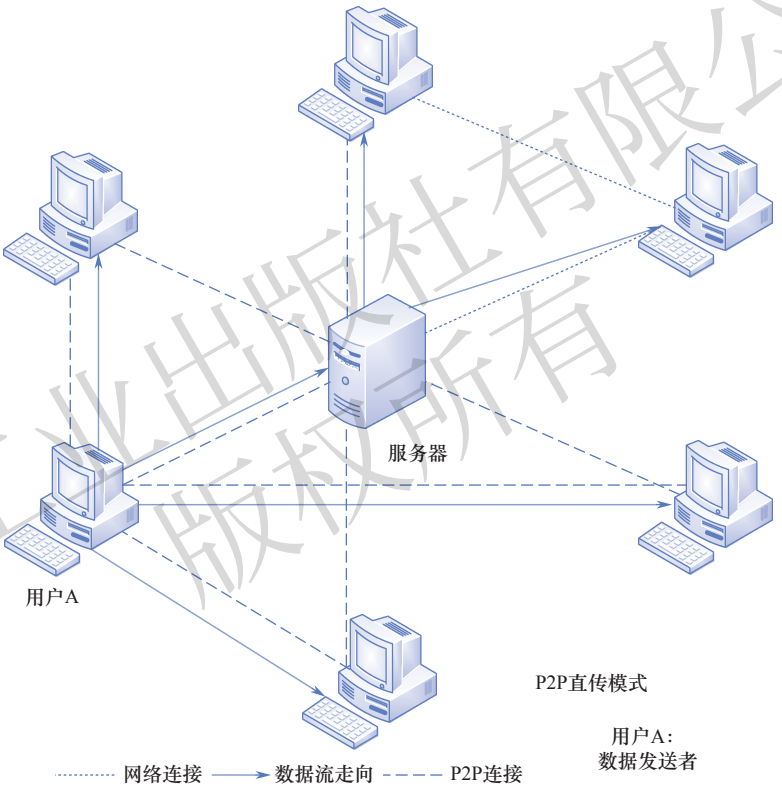

图 1-5　数据传输机制示例

保护传输中的数据的一般做法是数据加密。应选用合适的密码算法与加密实践，从而将传输加密当作一道有效的防线。

（3）数据的使用

对于静态数据的防护，不仅是最易于解决的，而且是普遍受到关注的重点。对处于传输状态的数据予以防护，是最易于理解的情形。与之形成对比的是，正在使用中的数据的安全性，却最容易被忽视，也是攻击者最易于突破的环节。使用中的数据比静态数据更容易受到攻击，因为根据数据的可用性，需要这些数据的人员必须可以访问它们。当然，有权访问数据的人员和设备越多，在某些时候数据落入错误的人员的手中的风险就越大。确保使用中的数据安全的关键是尽可能严格地控制访问并结合一种或多种类型的身份认证，

以确保用户使用的不是已被窃取的身份或凭据。

组织还需要能够跟踪和报告相关信息，以便检测可疑活动、诊断潜在威胁并主动提高安全性。例如，由于一定次数的失败登录尝试而被禁用的账户可作为系统受到攻击的警告信号。

经验证明，如果存在对攻击者而言有价值的数据，一定要识别出可以访问数据的每一个入口，并对其加以保护。仅将加密局限在数据生命周期中的一部分是一种危险的行为。保护存储、传输和使用这三种状态中的数据是至关重要的。

在数据生命周期的三种不同状态下，面对的挑战不同，采取的防御手段也不同。图 1-6 归纳了数据生命周期各种状态下的常见保护机制。后续章节会进一步介绍。

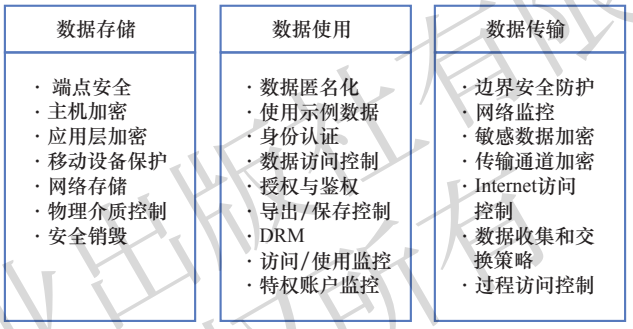

图 1-6　数据生命周期各种状态下的常见保护机制

1.2 数据安全保护制度

数据安全是网络安全保护的重要方面，数据安全保护制度与网络安全等级保护制度、关键信息基础设施安全保护制度密切相关。2021 年 9 月 1 日起实施的《数据安全法》，标志着我国将数据安全保护的政策要求，利用法律进行了明确和固化，以法律形式确立了数据安全保护制度，为各地区、各部门加强数据安全保护提供了法律保障。

数据作为关键要素和国家重要战略资源，在国家经济发展和社会进步中越来越发挥基础性和全局性作用。《数据安全法》是数据领域的基础性法律，也是国家安全领域的一部重要法律，旨在提升国家数据安全的保障能力和数字经济的治理能力，规范数据处理活动，保障数据安全，促进数据开发利用，保护个人和组织的合法权益，维护国家主权、安全和发展利益。

1.2.1 《数据安全法》主要内容

《数据安全法》是我国实施数据安全监督管理的基础性法律，目的是提升国家数据安

全的保障能力和数字经济的治理能力。《数据安全法》阐明了数据安全与发展的关系，明确了未来数据治理的方向。《数据安全法》共 7 章 55 条。第一章为总则，第二章为数据安全与发展，第三章为数据安全制度，第四章为数据安全保护义务，第五章为政务数据安全与开放，第六章为法律责任，第七章为附则。

《数据安全法》第七条规定，国家保护个人、组织与数据有关的权益，鼓励数据依法合理有效利用，保障数据依法有序自由流动，促进以数据为关键要素的数字经济发展；第九条规定，国家支持开展数据安全知识宣传普及，提高全社会的数据安全保护意识和水平，推动有关部门、行业组织、科研机构、企业、个人等共同参与数据安全保护工作，形成全社会共同维护数据安全和促进发展的良好环境；第十一条规定，国家积极开展数据安全治理、数据开发利用等领域的国际交流与合作，参与数据安全相关国际规则和标准的制定，促进数据跨境安全、自由流动。

《数据安全法》的内容可以总结归纳为如下要点。

（1）开展数据领域国际交流与合作，参与数据安全相关国际规则和标准的制定，促进数据跨境安全、自由流动。

（2）全面加强数据开放利用，推进技术和标准体系建设，建立健全数据交易管理制度。

（3）建立分类分级数据保护制度，形成集中、统一、权威的数据安全机制，建立数据安全应急处理机制、数据安全审查制度、数据安全出口管制以及根据实际情况采取数据投资贸易反制措施等，健全数据安全基础制度，以数据安全保障数字中国建设、推动数字经济高质量发展。

（4）明确数据安全保护义务，落实数据保护责任，加强数据安全风险监测和评估。

（5）国家机关政务数据要建立健全数据安全管理制度，落实数据安全保护责任，及时、准确公开政府数据，构建统一、规范、互联互通、安全可控的政务数据开放平台，推动政府数据开放利用。

数据处理者应重点关注如下内容。

1．重视数字经济

国家以法律形式保护数据应用和发展，体现了数字经济发展的极端重要性，为国家经济发展注入了强大动力。《数据安全法》体现了国家保护个人和组织依法、合理使用数据，促进数据依法有序自由流动和数字经济发展的坚定态度。加快数字化发展，推进数字产业化和产业数字化，推动数字经济与实体经济深度融合，是我国经济发展的基本国策。

2．高度重视数据安全

国家高度重视数据安全，加强综合治理和安全监管，切实保障国家数据安全。近年来，非法采集、加工、使用以及数据窃取、泄露等问题突出，严重危害国家安全、社会稳定和经济发展。各地区、各部门可以采用多种方式，积极宣传数据安全的法律、政策，调动有关部门、行业组织、科研机构、企业、个人等社会各界和社会力量的积极性，共同参与数据安全保护工作，共同维护数据安全。

3. 发展和安全并重

数据应用和数据安全是数据的两个重要方面，数据处理者需要找好平衡点，既不能过度强调应用而忽视安全，也不能因为重视安全而阻碍数据应用发展。为了解决这对矛盾，各地区、各部门要积极探索，该管好的重要数据、核心数据应保护好，一般数据应被积极应用，促进经济发展。

4. 加强国际合作

国家重视有关数据工作的国际交流合作，各地区、各部门应积极与有关国家、地区、组织等在数据安全治理、数据开发利用等领域开展国际交流与合作，积极参与数据安全相关国际规则和标准的制定，促进数据跨境安全、自由流动，推动世界数字经济发展，掌握国际数据治理的主动权和话语权。

1.2.2 我国数据安全管理的职责分工

《数据安全法》提出，维护数据安全，应当坚持总体国家安全观，建立健全数据安全治理体系，提高数据安全保障能力。各地区、各部门按照上述原则，依据职责分工，配合开展数据安全工作。

1. 中央国家安全领导机构统筹协调国家数据安全工作

数据安全与网络安全有着如下密不可分的关系。

《数据安全法》第五条规定，中央国家安全领导机构负责国家数据安全工作的决策和议事协调，研究制定、指导实施国家数据安全战略和有关重大方针政策，统筹协调国家数据安全的重大事项和重要工作，建立国家数据安全工作协调机制。

本条明确了中央国家安全委员会是国家数据安全工作的最高领导机关，负责国家数据安全工作的决策和议事协调，研究制定和指导实施重大战略、方针政策，统筹协调重大事项和重要工作，领导各单位、各地区开展数据安全工作。数据安全工作上升到国家安全最高领导机构负责，体现了数据安全工作的极端重要性，也充分体现了国家对数据安全工作的高度重视。国家安全委员会办公室建立国家数据安全工作协调机制，落实任务分工，承担日常统筹协调、督促落实等工作。

2. 各地区各部门按照法律法规的规定在职责范围内承担数据安全监管职责

《数据安全法》第六条规定，各地区、各部门对本地区、本部门工作中收集和产生的数据及数据安全负责。工业、电信、交通、金融、自然资源、卫生健康、教育、科技等主管部门承担本行业、本领域数据安全监管职责。公安机关、国家安全机关等依照本法和有关法律、行政法规的规定，在各自职责范围内承担数据安全监管职责。国家网信部门依照本法和有关法律、行政法规的规定，负责统筹协调网络数据安全和相关监管工作。

本条明确了各地区、各部门的职责。

（1）各地区、各部门按照数据安全法规定和国家有关政策要求，组织本地区、本部门开展数据安全工作，对本地区、本部门工作中收集和产生的数据及数据安全负责，承担数据安全的主体责任。

（2）工业、电信、交通、金融、自然资源、卫生健康、教育、科技等主管（包括监管）部门，包括能源、水利、公检法、司法、发改、财政等重要行业部门，承担本行业、本领域数据安全主管（监管）职责，结合法律和国家政策，出台行业规范，指导并组织本行业、本领域开展数据安全工作。

（3）公安机关、国家安全机关等依照本法和有关法律、行政法规的规定，在各自职责范围内承担数据安全的国家监管职责，开展威胁情报工作，侦查打击危害数据安全的违法犯罪活动，保卫数据安全。

1.2.3　数据安全保护与网络安全等级保护的关系

数据安全与网络安全有着如下密不可分的关系。

（1）《数据安全法》第二十七条规定，利用互联网等信息网络开展数据处理活动，应当在网络安全等级保护制度的基础上，履行数据安全保护义务。法律将网络安全等级保护制度延伸到数据安全保护领域，表明了网络安全等级保护制度与数据安全保护制度的密切关系，进一步确立了网络安全等级保护制度的基础性作用和地位。

（2）数据安全保护是网络安全等级保护和关键信息基础设施保护的重要内容。《中华人民共和国网络安全法》（简称《网络安全法》）以及网络安全等级保护制度、关键信息基础设施安全保护制度，均对数据安全保护提出了明确要求。数据通常存储并运行在网络系统中，是网络系统和关键信息基础设施不可分割的重要组成部分，同时，关键信息基础设施中存储处理大量国家重要数据和核心数据。按照《数据安全法》和有关政策要求，数据分为一般数据、重要数据和核心数据，其中，重要数据应落实第三级等级保护要求，核心数据应落实第四级等级保护要求和关键信息基础设施安全保护要求。

（3）数据安全的特殊性。存储和运行在网络系统中的数据，其安全保护和管理与网络安全保护是一个整体、密不可分。从保护措施、保护手段、管理措施、保卫措施等方面，网络和数据安全通常按照整体布局和统筹安排。但数据安全具有其特殊性，体现在数据可以交易、买卖、销毁、出境、跨平台跨系统跨境传输等，需要在这些环节和活动中，对数据安全的特殊性予以高度重视，制定相关政策、标准规范，采取相应的管理和技术手段、措施，解决数据全生命周期的安全问题。

1.3　数据安全法律政策和标准

在数据安全与隐私保护维度，存在一系列的国内外标准。一些标准是特定国家或地区的准入门槛，称为“强制认证标准”。企业在面向这些国家和地区提供服务时，除应满

足当地法律法规的要求外，还需要满足强制认证标准的要求。这个动作也称“标准遵从”。另外一些标准是推荐性的，可以作为参考。部分标准还提供认证功能，以认证一个组织、组织所提供的产品或服务满足特定的安全水平。

1.3.1 数据安全法律政策

为推进数据安全建设，我国在法律上、制度政策上逐渐形成数据安全保护体系。

2016 年 8 月，中共中央网络安全和信息化委员会办公室发布《关于加强国家网络安全标准化工作的若干意见》，在发展高端制造业和《促进大数据发展行动纲要》等国家战略的要求下，推进关键信息基础设施保护、大数据安全、网络安全等领域的标准研究和制定工作。

2017 年《网络安全法》正式实施，网络运行安全和关键信息基础设施安全、网络安全等级保护制度等工作逐步加强。

2021 年 3 月，国家互联网信息办公室联合四部门共同发布《常见类型移动互联网应用程序必要个人信息范围规定》，对过度利用个人信息的现象进行整治。

2021 年 4 月，《关键信息基础设施安全保护条例》出台，我国大力推进以 5G、人工智能、工业互联网、数据中心为代表的新型基础设施建设，为实现数据安全奠定基础。

2021 年，《数据安全法》与《中华人民共和国个人信息保护法》（简称《个人信息保护法》）正式实施，数据安全与个人信息保护工作迎来了新阶段。

1.3.2 数据安全标准

1. 国际数据安全标准

国际标准化组织（International Standards Organization，ISO）和国际电工委员会（International Electrotechnical Commission，IEC）是两个权威的国际标准组织。SC 27 是 ISO 和 IEC 两个组织的联合委员会 JTC 1 下属的安全技术分委会，于 1990 年成立，负责信息安全领域国际标准、技术报告和技术规范的制定。SC 27 的标准化活动包括处理信息安全、网络安全和隐私的通用方法、管理系统要求、技术和相关准则。

SC 27 下设五个工作组和两个管理组，其组织架构和工作内容见表 1-1。数据安全和隐私保护作为横向业务领域，与上述各个工作组的工作内容、国际标准都有相关性。

表 1-1　SC 27 组织架构和工作内容

SC 27		工 作 内 容
管理组	SWG-M	SC 27的总体管理
	SWG-T	SC 27的横向项目管理

（续表）

SC 27		工 作 内 容
工作组	WG 1	信息安全管理系统（ISMS）
	WG 2	密码学和安全机制
	WG 3	安全评估、测试和标准
	WG 4	安全控制和服务
	WG 5	身份管理与隐私技术

SC 27 发布的 ISO/IEC 27000 系列标准涵盖信息安全管理的最佳实践建议。这些国际标准提供了信息安全的顶层框架，包含组织信息风险管理流程中涉及的法律、物理和技术控制措施。

ISO/IEC 27001 是 ISO/IEC 27000 系列信息安全标准中的一篇，描述了建立、实施、维护和持续改进信息安全管理系统（ISMS）的要求。符合该标准要求的组织可以选择在成功完成审计后由认可的认证机构进行认证。ISO/IEC 27001 要求管理层完成以下几项工作。

（1）系统地检查组织的信息安全风险，综合考虑威胁、脆弱性和影响。

（2）设计并实施一套连贯而全面的信息安全控制措施和其他形式的风险处理办法（如风险规避或风险转移），以应对那些被认为不可接受的风险。

（3）采用一个总体的管理程序，以确保信息安全控制措施持续满足组织的信息安全需求。

安全领域的管理标准主要是 ISO 27005：“信息安全风险管理”。实施指南类标准主要是 ISO 27110：“网络安全框架实施指南”。安全控制措施设计类的标准主要是 ISO 27002：“信息安全控制的实践准则”。

2. 国内数据安全标准

近年来，数据安全类国标发布数量逐渐增多，这反映了数据安全在国家安全战略中的地位不断提升，以及社会对数据安全保护需求的日益增长。

截至 2024 年，我国已经基于《网络安全法》《数据安全法》《个人信息保护法》出台了多项数据安全领域国家标准、行业标准、地方标准和团体标准，如下。

（1）国家标准《数据安全技术 数据分类分级规则》（GB/T 43697）不仅给出数据分类分级的通用规则，为数据分类分级管理工作的落地执行提供重要指导，同时针对重要数据制定了识别指南。

（2）《信息安全技术 步态识别数据安全要求》（GB/T 41773）规定了步态识别数据在收集、存储、传输、使用等过程中的安全要求，适用于步态识别数据处理者规范数据处理活动。

（3）医疗健康大数据安全要求团体标准强调了医疗健康大数据在收集、存储、传输和使用等各个环节的安全要求。它要求确保数据的真实性和完整性，并采用安全可靠的采集

设备和方法。此外，还提出了身份认证和访问控制的要求，以及数据备份和灾难恢复的建议。

1.4 数据安全管理

伴随着各行各业的数字化转型，数据范畴越来越多，数据量级越来越大，数据传输的速度越来越快，数据的价值将会越来越高。而数据资源区分于传统资源的重要特征是复制和篡改的低成本，流动迅速，这些都对数据的安全构成了重大的挑战。

为了应对日益复杂的环境，特别是网络环境的开放性和复杂性、攻击者的日益专业化和组织化，各类组织需要采用以数据为中心的安全治理框架，按照“识别、保护、检测、响应、恢复”的方法论，针对敏感数据和关键数据，基于其生命周期流程，实现完善的端到端的保护。

1.4.1 数据安全管理目标

数据安全管理目标可分为“数据合规”、“风险管理”和“充分利用”三个维度。数据合规是数据安全管理的底线要求，包括数据安全、数据跨境合规和数据隐私合规等内容；风险管理是数据安全管理要面对和解决的主要问题，而数据的充分利用，有助于最大化的发挥数据的价值，促进组织目标的达成。数据的“充分利用”与业务场景、业务目标相关，在本书的后面章节深入讨论。

（1）在数据合规维度，基于对业务与场景的识别，对各类数据的识别，数据所面临风险的识别，进一步推进到组织对于不同等级的数据安全风险设置不同的管理政策，制定相应的策略，以明确数据合规的管理目标。

例如，通过业务与场景识别，可以确认企业产品和服务的业务形态、所提供服务的国家和地区、业务的交互关系和关键路径。基于此，可以识别出数据的类型，对数据做准确的分级分类，识别数据的流向和数据生命周期的关键阶段。对于新型数据，如 AI 模型数据、AI 训练集数据，还可以提炼出数据特征，以匹配可能的风险。

（2）在风险管理维度，核心是识别、评估和处理各类信息系统的使用，会给组织的数据或关键信息资产带来的风险。风险的评估更像是商业概念，重点是考虑哪些数据资产会影响组织的盈利能力，或者哪些数据资产的泄露会给组织带来巨大的经济损失。然后再查看这些数据资产面临哪些风险和威胁。

基本的数据安全风险评估涉及三个因素：数据资产的重要性、威胁的严重程度及系统脆弱性。利用这些因素，可以评估数据安全风险，即组织蒙受商业损失的可能性。风险评估很难量化，但可以直观地用下述公式来表示：

$$\text{风险} = \text{资产} \times \text{威胁} \times \text{脆弱性}$$

上述公式还有另一种直观的理解，任何数字乘以 0 的结果都是 0。如果威胁很大，系统也脆弱，但资产价值为 0，那么就不需要额外的保护。

从数据风险评估模型可以得出，安全事件发生的可能性与资产的脆弱性及威胁出现的频率相关，而安全事件造成的损失与资产的价值和资产的脆弱性相关。

更进一步的评估可以基于描述安全事件发生的可能性与安全事件的损失的风险值评估矩阵进行，如图 1-7 所示。基于对风险的评估，组织可以基于自己的风险管理要求和偏好，定义相应的风险管理的目标。

事故发生的可能性	安全事件的损失				
	极低	低	中	高	极高
极低	0	1	2	3	4
低	1	2	3	4	5
中	2	3	4	5	6
高	3	4	5	6	7
极高	4	5	6	7	8

图 1-7　风险值评估矩阵

1.4.2　数据安全风险评估

系统化的数据安全风险评估方法，如图 1-8 所示。数据安全风险评估，主要围绕数据处理者的数据和数据处理活动，对可能影响数据保密性、完整性、可用性和数据处理合理性的安全风险进行分析和评价。

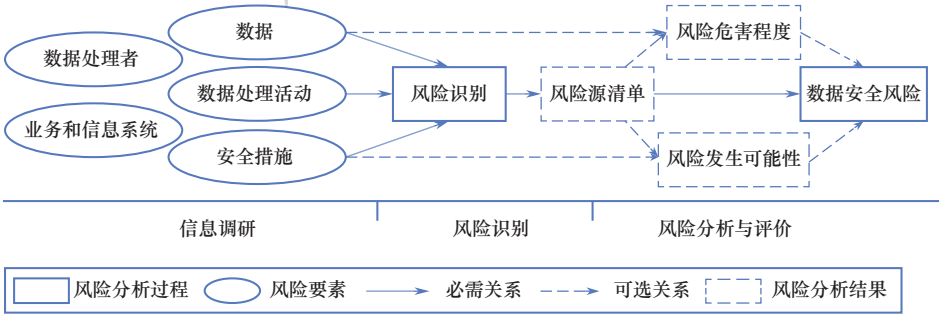

图 1-8　数据安全风险评估

数据安全风险评估主要包括信息调研、风险识别、风险分析与评价等。

（1）信息调研：对可能影响数据安全风险的要素的情况进行调研，包括数据处理者、业务和信息系统、数据资产、数据处理活动、数据安全防护措施。掌握数据处理者、业务和信息系统基本情况，梳理涉及的数据资产和数据处理活动，了解采取的数据安全防护措施情况，掌握被评估对象或同行业相关数据安全事件历史发生情况。

（2）风险识别：基于信息调研情况，从数据安全管理、数据处理活动安全、数据安全

技术、个人信息保护等方面进行数据安全风险识别，识别评估对象现有安全措施完备性并对其有效性进行验证，识别可能存在的风险源。

（3）风险分析与评价：通过分析风险类型、风险危害程度和可能性，评价风险等级。

1.4.3 兼顾安全与发展平衡

在数据安全管理中，一个常见的误区是“将数据放入保险箱”，越安全越好。诚然，将数据放入“保险箱”，可以实现最大化的安全，风险也最小。但是，在数字经济时代，数据只有充分流动才能最大化发挥出其价值。而数据的正确、有序、高效地流动，本质上隐含了数据安全的要求。所以，需要辩证地看待数据安全治理，在制定数据安全管理的目标时，兼顾发展与安全的平衡。

在明确数据安全管理目标之后，即可在实践中定义和优化组织的数据安全管理架构。

1.5 数字化生态安全

在信息化和数字化时代，数据是国家的基础性、战略性资源，数据安全事关国家政治安全和社会稳定。大数据的双刃剑作用突显，如何管控好大数据对国家安全带来的风险与挑战，如何科学平衡数据安全与数据流通应用之间的关系，如何构建数据安全综合治理体系，确保重要数据和公民个人信息安全，是必须解决的重要问题。

1.5.1 数据安全面临的威胁风险

信息化和数字化时代，面临的数据安全风险是多样的，如下。

（1）数字化建设面临的最大威胁是网络攻击。数字化建设面临的最大威胁是网络攻击，数字化在促进我国经济发展的同时，给国家经济安全、社会运行和治理带来了风险与挑战，因此，要高度重视数据安全工作。

（2）数字化时代的网络安全呈现出风险普遍化、集聚化等特征，风险防控难度加大。数字安全风险遍及所有场景，网络安全边界发生重大变化，网络世界和物理世界相融合，万物均可互联，每个设备都可成为攻击点，任何对网络设施的攻击都可以变成对物理世界的危害。

（3）关键数据资源成为网络攻击的重点。数字化时代，数据成为重要生产要素，数字化发展带来大量数据安全问题。产业数字化、公共服务数字化、社会治理数字化使得政务、商务数据和个人信息爆发式增长、海量汇聚，数据资源量大、多源、跨平台流动和开放共享，成为网络入侵攻击和窃密的重点目标，关键数据资源、商业秘密和个人隐私泄露风险突出。

（4）网络攻击专业化智能化程度提高。数字化时代新形态网络广泛应用，网络安全边界逐渐模糊，网络接入设备多样化，除了手机、电脑、可穿戴设备，还有车联网、智能制造、智能家居等设备，各种网络安全漏洞、后门层出不穷，网络攻击切入点更加多样，攻击的专业化、智能化、自动化水平更高，常规的网络安全管理、监测、防护措施和设备设施面临挑战。

（5）数据安全保障不足。数据安全保护任务繁重，但政府有关部门、重要行业、央企、互联网企业的数据安全专门机构和专业人员明显不足，大多数单位负责数据安全工作的是兼职人员。一些数字化建设项目没有配套的数据安全建设经费，数据安全保护经费难以得到保障。

1.5.2　数据安全存在的问题隐患

当前，存在的数据安全问题隐患主要有以下几点。

（1）数据大集中大流动大应用，客观上造成数据保护困难、网络攻击容易。我国数据安全问题隐患突出，数据遭攻击、窃取、破坏等事件高发。

（2）数据安全意识差，数据重应用轻安全的老问题依然突出，数据安全保护仍然滞后于数字化建设与应用。缺乏创新性措施，主管责任、主体责任、监管责任、第三方服务责任等四方责任落实不到位。

（3）数据安全综合防御体系和治理体系尚未建立。数据安全防护缺乏整体性、综合性措施，整体合力未形成。数字化时代的网络攻击、数据窃取更加精准，隐秘性、破坏性强，在应对智能化、专业化网络攻击、数据窃取等方面，传统的防护手段和措施难以适应数字化时代要求。

1.5.3　数字化生态安全保护

我国加快推进数字经济、数字政府、数字中国建设和数字化转型，本质是打造数字化生态，建设数字社会。数字化生态由基础要素（网络、系统、平台、数据、技术等）支撑，数据流打通相关领域和行业，数据资源得到有效利用，构成数字化生态。数字化生态建设与自然界生态建设类似，是一个复杂的系统工程，具有“脆弱性、风险性、长期性和复杂性”等特性。数字化建设面临的最大威胁和风险是网络攻击，因此，开展数字化生态建设，要与数字化生态安全建设“同步规划、同步建设、同步运行”，确保数字化生态的基础要素安全、数据流通安全、数据应用安全。

有关数字化生态安全建设的详细介绍，请见高等院校网络空间安全专业实战化人才培养系列教材中的《网络安全保护制度与实施》一书。

1.6 新技术与数据安全的关系

随着互联网到物联网再到智联网的演进，泛在智能连接的社会即将到来。无论是工业、农业还是人们的生产生活，都更多地依赖于数字化的安全。在新兴的场景中，数据安全从攻击和防御两个维度都呈现螺旋式上升的趋势，数据安全领域也在不断地创新、深化、扩展。

1.6.1 人工智能与数据安全

人工智能技术的突破，对数据安全的影响显著。人工智能基于对数据的分析和洞察，在应用领域实现智能和自动化的决策。现代的人工智能通常与机器人技术、物联网等深度融合，应用于不断增长的数据类型和庞大的数据量的场景中。数据是人工智能全场景中最有价值的资产。在人工智能的全生命周期中，数据以不同的形态存储、转换、转移和处理。

1. 人工智能和数据安全的交叉维度

人工智能和数据安全这两个领域紧密相关，其交叉点存在于以下维度。

（1）AI 的数据安全：聚焦人工智能全生命周期中涉及的各类数据的防护。例如，部署阶段需要考虑 AI 模型的机密性和完整性防护，训练阶段需要考虑原始数据和训练集的安全防护。

（2）AI 的隐私保护：聚焦数据采集最小化原则的应用，AI 的隐私数据访问控制的实现，以及差分隐私等各类隐私增强技术与 AI 的结合。

（3）AI 用于处理个人数据的安全：聚焦隐私数据的个人权利在人工智能的实现。

2. 人工智能和数据安全的内在关系

（1）人工智能系统需要数据安全。在人工智能系统的完整性、机密性和可用性维度，以及隐私保护和防滥用方面，需要安全功能的支撑。而此类安全功能的缺失，将导致人工智能系统本身的安全风险。

（2）人工智能系统需要符合隐私保护的监管要求。尤其是原始数据和训练数据中包含个人数据的场景。匿名化和差分隐私等隐私增强技术与人工智能技术的结合，仍然是学术界和工业界研究和实践的热点。

（3）使用人工智能处理个人数据，存在安全、隐私与可信赖等诸多方面的风险，因此需要仔细地评估和应对这些风险。例如，GDPR 要求的删除权和修正权，在人工智能场景数据流转中的处理会比较复杂。此外，AI 用于处理个人数据时，还需要考虑公平、道德，以及用于涉及个人的自动决策的可解释性。

3. 人工智能对数据安全的影响

人工智能系统引入了传统的信息技术系统不具备的复杂性。在合规层面，人工智能系

统引入了个人数据控制者和处理者的边界复杂性。在组织层面，参与设计和部署人工智能系统的角色和人员要远远多于传统的信息技术系统，如数据科学家、统计学家、算法工程师等。在业务层面，人工智能系统和现有的业务流、数据流和工作流存在着错综复杂的交互。在技术层面，人工智能系统如何安全及可信地处理个人数据，仍处于探索阶段，并没有明确的处理原则与最佳实践方案。在实现层面，人工智能系统还可能依赖于第三方的软件或代码，因此引入了供应链的复杂性。

总体而言，人工智能对于数据安全的影响取决于以下几个因素。

（1）构建和部署 AI 系统的方式。

（2）部署 AI 系统的组织的复杂度。

（3）现有风险管理能力的成熟度。

（4）AI 系统处理个人数据的性质、目的、范围和背景。

制定和实施安全控制措施时，也要综合分析和考虑上述因素。

4．人工智能赋能数据安全

人工智能的飞速发展，对于网络安全领域和数据安全领域都有极大的促进作用。作为攻击者和防御者都可以使用的“武器库”，现有的网络安全和数据安全的解决方案、系统和工具需要不断地更新，以更好地符合实时、主动的数据保护的需要。目前人工智能在网络和数据安全领域的常见应用场景如下。

（1）数据保护

数据保护涉及保护敏感信息免遭数据丢失和损坏，以及保护数据并确保其可用性和符合监管要求。AI 工具可以帮助组织分类敏感数据、并通过监控数据传输以及防止未经授权的访问或泄露来改善数据保护。AI 还可以实现更好的数据加密、匿名化和假名化，以保护静态和传输中的数据。显然，和人工处理相比，人工智能可以自动适应威胁形势并全天候持续监控，使组织能够领先于新出现的网络威胁。

（2）端点安全

端点安全指保护计算机、服务器和移动设备等端点免受网络安全威胁。人工智能可以通过持续监控端点的可疑行为和异常来检测实时安全威胁，从而改进现有的端点检测和响应（EDR）解决方案。机器学习算法还可以帮助识别和减轻高级端点威胁（例如无文件恶意软件和零日攻击），以防止它们造成危害。

（3）云安全

人工智能可以通过自动识别影子数据、监控数据访问异常以及在威胁发生时向网络安全团队告警，以帮助保护混合云环境中的敏感数据。

（4）风险控制和欺诈识别

随着网络攻击和身份盗窃变得越来越普遍，各类组织需要方法来保护其用户的数据和资产。人工智能通过自动分析交易数据来识别欺诈。此外，机器学习算法可以实时适应新的和不断演变的威胁，不断提高其欺诈检测能力。

（5）网络安全自动化

安全编排、自动化和响应（SOAR）是许多组织用来简化安全操作的解决方案。AI 可以与 SOAR 平台集成，使日常任务和工作流程自动化。这种集成可以加快事件响应速度，并让安全团队有时间专注于更复杂的问题。

（6）身份和访问管理

身份和访问管理（IAM）工具管理用户对数字资源的访问和使用。其目标是阻止黑客入侵，同时确保每个用户都拥有恰好能满足其任务的准确权限。人工智能驱动的 IAM 解决方案可以通过提供基于角色、职责和行为的细粒度访问控制来改进流程，进一步确保只有授权用户才能访问敏感数据。人工智能还可以通过使用机器学习来分析用户行为模式并实现根据个人用户的风险级别而变化的自适应身份验证措施来增强身份验证流程。

（7）漏洞管理

漏洞管理是持续发现、缓解和解决组织 IT 基础设施和软件中的安全漏洞的过程。人工智能可以根据潜在影响和漏洞利用可能性，自动对漏洞进行优先排序，从而增强传统漏洞扫描程序的功能。这有助于组织首先解决最关键的安全风险。人工智能还可以自动化补丁管理，以及时减少网络威胁的发生。

1.6.2 云计算与数据安全

在各组织的数字化转型的进程中，云化转型是一个关键的考量因素。云计算，通常称为“云”，指通过互联网等网络提供存储、服务器和软件的托管服务。云转型有助于企业加速业务部署，员工可以随时随地访问资源，改善团队协作并降低管理和运维成本。与此同时，“云”中的数据安全是一个关键的挑战。

云数据安全的主要目标是：确保数据安全和隐私保护，处理多家云服务供应商的数据安全问题，用户、设备和软件的访问控制。

因为云资源访问的便利性，如果没有适当的安全策略，收集的数据就会时刻处于泄露的危险之中。许多人认为，只有大型企业才会成为网络攻击的受害者。实际上，根据数据泄露的研究报告，中小企业同样是恶意行为者的核心目标。云安全领域缺乏投资的组织将面临巨大的问题，包括可能遭受数据泄露以及在管理敏感客户数据时无法保持合规。

1. 云化部署的四种模型

云化部署常采用四种不同的模型，如表 1-2 所示。

表 1-2 云化部署的四种模型

部署模型	描　述
公有云	公有云基础设施由第三方服务提供商托管，并由多个租户共享。每个租户都控制托管在云中的自己的账户、数据和应用程序，但基础设施对所有租户都是通用的。这种模式成本低，风险高。一个账户的泄露会使所有其他账户面临风险

（续表）

部署模型	描　　述
私有云	私有云也称为单租户部署模式，其中基础设施通过私有云提供，并由一个租户独家使用。在此模型中，云资源可以由组织或第三方供应商管理。 此部署模型的好处是它为各个组织提供了较高的控制级别，增强的安全性并确保合规性，使其成为处理敏感信息的组织最常用的模型。显然，私有云部署和运维成本最高
混合云	混合云将公有云和私有云整合到一个共享环境中。 这种部署模式的最大优势在于其灵活性和性能
多云	多云部署利用多种公有云服务。这些服务通常由不同的计算和存储解决方案组成，但各云服务平台都提供多种选项，以构建多云基础设施

大多数组织使用第三方云服务提供商，如华为云、阿里云、腾讯云等，来托管其数据和应用程序。云数据安全需要这些云服务提供商与其客户共同承担责任。

2．常见的云服务模型及其保护责任

可靠的云服务提供商具有强大的安全性，可以保护自己免受攻击，但是云服务使用方的安全措施也非常重要。如果使用方的组织内存在安全配置错误、特权访问漏洞或人为错误，攻击者可能会从恶意端点侵入移动到组织的云工作负载中。因此，需要明确安全策略，制定安全措施，提升组织内部的安全意识、安全能力。

对于四种常见的云服务模型，其保护基础设施、业务和数据的责任划分有显著不同，“责任共担模型”描述了常见的责任划分，如表 1-3 所示。

表 1-3　常见的云服务模型及其保护责任

	SaaS	PaaS	IaaS	本地部署
应用程序配置	客户	客户	客户	客户
身份和访问控制	共担	共担	客户	客户
应用程序数据存储	云服务商	共担	客户	客户
应用程序	云服务商	客户	客户	客户
操作系统（OS）	云服务商	云服务商	客户	客户
网络流控制	云服务商	云服务商	共担	客户
主机基础设施	云服务商	云服务商	云服务商	客户
物理安全	云服务商	云服务商	云服务商	客户

（1）软件即服务（SaaS）场景，客户在现有的云服务提供商中可以找到满足需求的服务。此模型的大部分责任由客户委托给云提供商，并从提供商规模化的安全优势中获益。

（2）平台即服务（PaaS）场景，客户与云提供商共享的责任界面存在显著差异。典型的设计是，云提供商负责硬件基础设施和操作系统，客户负责平台上开发的应用程序的安全性，以及端点、用户和网络安全。

在不确定使用哪种云服务模型时，建议优先考虑 PaaS 而不是 IaaS。因为 IaaS 模式需要更多的开发工作量和部署成本。

（3）基础设施即服务（IaaS）场景，云提供商负责提供处理、存储和网络等基础设施，供用户构建应用程序。用户负责基础设施上安装的任何应用程序的安全性（例如操作系统、应用程序、中间件）。实际场景可能有所不同，从完全由客户负责，到云供应商提供工具来帮助管理操作系统（OS），甚至代表客户使用这些工具。

如果选择 IaaS 模型，可以在企业软件架构中，识别和替换适合使用托管服务的组件，如存储、身份和访问管理、数据流控制、日志记录和监控。

3．云数据安全内容

（1）云安全数据目录

保护云数据的第一步是了解组织在云中拥有哪些数据。因此，采用强大的数据目录技术非常重要。数据目录可以发现所有数据以及有关数据的所有信息，无论这些数据位于托管或非托管数据资产、数据缓存、数据管道、大数据环境还是影子数据（未知数据存储）。

除了发现数据，云安全数据目录还应对数据进行分类和编目，以获取确定如何保护数据所需的信息，这意味着揭示从数据类型和记录到敏感度和所有者的所有内容。任何有价值的数据目录解决方案都应该自动、异步地完成所有这些工作，并且不需要用户事先了解或付出额外的努力。挑战在于找到所有数据，即使用户不知道数据存在的地方。要做到这一点，云数据安全解决方案必须是 100% 自主工作的，不需要安装代理或连接器，也不需要了解访问凭据。

（2）数据安全态势管理

通过云安全数据目录，组织已经知道了自己拥有哪些数据以及这些数据在哪里，下一步就是实施政策，规定如何保护数据，并确定既定政策与现有安全态势之间的差距，这些差距可能会使敏感数据面临风险。从暴露和访问到保留期限和加密，所有内容可以由数据安全态势管理（DSPM）平台管控和验证。DSPM 的核心是一个策略引擎，它可以检测和警告数据安全策略违规行为，然后提供指导性补救措施。这使安全团队能够评估安全态势，优先处理对业务构成最大风险的数据，并补救实际上使敏感数据面临风险的问题。

与云安全态势管理工具不同，DSPM 策略关注数据，而不关注数据驻留在哪个基础架构上。这些策略以敏感度为基础，以数据为中心。

（3）云数据访问控制

云环境中的访问非常复杂。寻求全面了解其数据安全风险的组织需要了解哪些实体可以访问哪些数据。数据访问控制功能可以实现数据访问的可视化，并为数据安全从业者提供所需的信息，以便他们探索数据和实体之间的连接方式。

例如，假设 DSPM 发现违规行为，即允许第三方供应商访问，从而过度暴露敏感数据，审计人员此时会问，该供应商还能访问什么，或者还有谁可以访问这些敏感数据——CDAC 可以帮助审计人员直观地了解这两个问题的答案。或者，假设数据检测和响应（DDR）发出有关某台机器中存在恶意活动的警报，那么调查人员可能想知道该机器上具有特定身份的恶意行为者访问了哪些云数据。CDAC 能够直观地了解并简化到可能暴露的具体数据。

（4）数据检测和响应

全面云数据安全策略的最后一个组成部分是了解当前哪些活动表明敏感数据可能受到攻击。数据检测和响应（DDR）监控并警告正在进行的实时活动。与扩展检测和响应（XDR）或端点检测和响应（EDR）等类似工具不同，DDR 了解哪些数据是关键的。然后，它使用机器学习来检测围绕这些关键、最敏感数据的异常活动，从而使警报更加具体，并大大降低误报率、漏报率，减少安全运维的成本。

4．云数据安全的特点

在云数据安全领域，涌现出一些新的概念和解决方案，并且逐步得到关注。云安全联盟（CSA）在“什么是云数据安全”一文中提及云安全数据目录、数据安全态势管理（DSPM）、云数据访问控制（CDAC）、数据检测和响应（DDR）等创新方案。

全球许多企业都在适应远程和混合工作环境，这意味着有更多的接入点容易受到威胁。实施云数据安全有六大优势。

（1）更高的可视性：强大的云数据安全解决方案可帮助组织保持对数据的可视性，包括组织拥有的数据类型、数据所在位置以及任何特定时间的访问者。

（2）更安全的数据：云存储通过为传输中的数据添加多层高级加密，帮助企业实现安全的数据传输、存储和共享。

（3）云数据合规性：解决方案中的安全程序旨在不断满足各种合规性要求。云数据丢失防护（DLP）可以帮助发现、分类和匿名化敏感数据，以避免违规。

（4）更容易地备份和恢复：通过自动执行数据备份并标准化数据备份流程，云数据安全可以监控这些备份并排除潜在阻塞。如果确实出现挑战，灾难恢复可以在几分钟内恢复和还原数据和应用程序。

（5）高级事件检测：许多云数据安全解决方案都提供最新的安全功能和工具，包括人工智能和内置安全分析。这些附加功能有助于扫描可疑活动，尽早提醒安全团队并尽早消除可能的威胁。

（6）降低组织成本：云数据安全有助于降低总体拥有成本以及运营和管理责任。团队可以实施威胁分析和威胁警报的自动化流程。

1.6.3　物联网与数据安全

随着经济的发展和社会的进步，物联网应运而生。根据 Statista 机构的统计和预测，到 2030 年，全球预计将有 500 亿个物联网设备，可以形成一个庞大的设备互联网络，涵盖从智能手机到厨房家电的所有领域和场景。

一般而言，物联网（IoT）指连接物理对象，也就是说“物”的网络。物联网通过传感器、软件和网络连接等技术，将设备同其他设备和系统连接并交换数据。

计算机技术和通信技术的结合构成了计算机网络，即互联网。互联网和移动通信技术的结合构成了移动互联网。而移动互联网、感知技术的叠加构成了物联网。物联网组成技

术如图 1-9 所示。

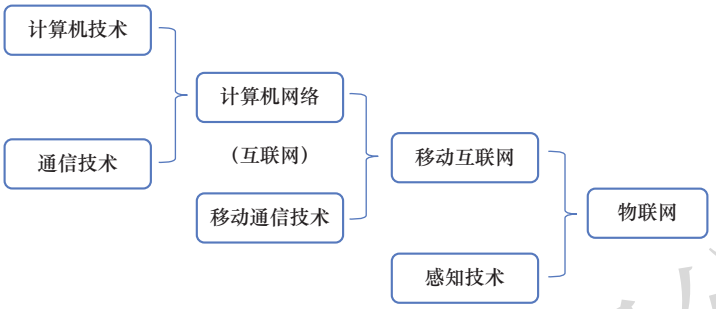

图 1-9　物联网组成技术

1．物联网的功能

物联网的功能可以抽象地概括为捕获数据、传输数据、处理数据和采取行动。

（1）捕获数据：通过传感器，物联网设备从其所处的环境中捕获数据。数据可以是简单的（如温度），也可以是复杂的（如实时视频、音频）。在物联网设备上也可以执行简单的数据预处理或数据分析，如数据采样、数据超出阈值之后生成告警信息等。

（2）传输数据：物联网设备利用可用的网络连接，根据配置，通过公共或私有网络传输捕获的数据或预处理的数据。

（3）处理数据：对来自物联网网络内设备的聚合数据进行分析，提取数据表达的信息，并根据信息分析行动方案。

（4）采取行动：基于对数据、信息的分析和洞察，支持下一步的行动或商业决策。

可以简单地将物联网的网络架构分为感知层、网络层和应用层三个层级，如图 1-10 所示。在感知层，泛在的物联网传感器从所在环境中捕获目标对象的数据。在网络层，物联网设备通过有线网络、无线网络等方式，将数据直接或通过物联网网关传递到云端。在应用层，云服务基于对数据的处理和分析，生成基于数据的结果、趋势、洞察，并在需要时自主或由人工采取下一步行动。

2．物联网安全

物联网安全指物联网硬件、软件及其系统中的数据受到保护，不因偶然的或恶意的原因受到破坏、更改、泄露，物联网系统可以连续、可靠、正常地运行，物联网服务不会中断。物联网安全包括解决或缓解物联网中的安全威胁的技术手段或管理手段。

物联网的整个生态系统中，设备、网络、数据，都存在着安全风险。因此，物联网安全包含数据的安全、网络的安全、设备的安全。其中，网络与设备的安全，其主要目标也是数据的安全。数据的完整性、全面采集、安全传输及有效保护是物联网安全的基础。

IEEE 于 2017 年 2 月发布的“物联网安全最佳实践”文章中指出，物联网的安全分为三个层面的问题：设备的安全问题，网络的安全问题，数据的安全问题。其中最为主要的数据安全问题是物联网安全问题的本质和核心。在万物互联互通的物联网中，信息直接与

物理世界相连接，在带来极大便利性的同时也带来了巨大的风险。国家信息安全，生产安全，甚至个人私密信息，都因为其在物联网上运行而增添了被盗取、拦截和更改的风险。基于这些风险，对数据的保护必不可少。

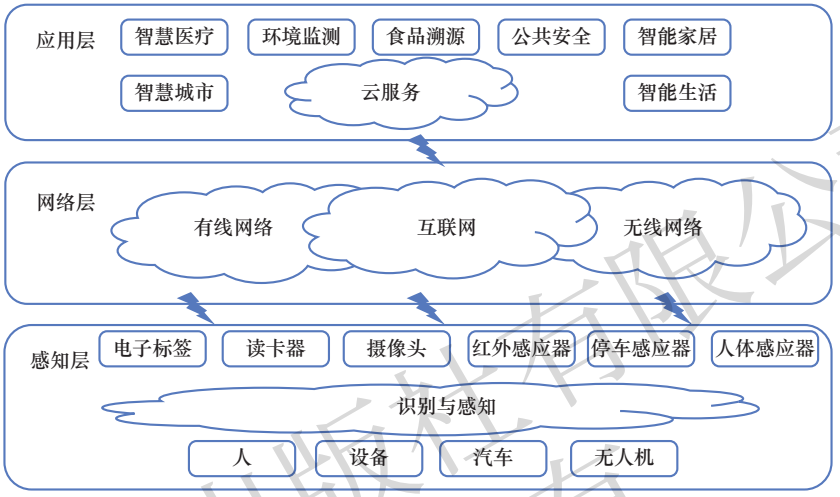

图 1-10　物联网架构示意图

物联网的设备数量、协议异构性和网络复杂性的结合，给物联网安全带来严峻的挑战。仅从物联网设备的层面看，虽然安全性基本可控，但是仍存在不正确的配置和缺乏更新等问题。因为消费者往往缺乏安全意识，消费级物联网设备存在一些挑战。而从物联网网络的维度，或者在工业物联网的场景中，安全问题更为复杂。2021 年 4 月，针对物联网安全的数据统计显示：

（1）不到 20% 的物联网安全风险专业人员了解他们组织所使用的所有或大部分物联网设备；

（2）76% 的物联网风险专家认为他们组织的物联网安全态势使他们容易受到网络攻击；

（3）56% 的组织没有其使用的物联网设备的完整清单，64% 的组织没有其物联网应用的完整清单；

（4）2020 年，在新冠疫情等综合因素导致恶意软件总量下降 39% 的前提下，物联网恶意软件攻击增加了 30%。

3．物联网的安全风险

基于物联网的本质特点，存在如下几个突出的安全风险。

（1）部署的复杂性增加了远程暴露和外部攻击面。物联网设备并不像大部分 IT 设备那样部署在物理上独立的机房中，而是可能放置在办公室开放空间、工厂、家庭等各种设备的拥有者无法控制的位置。因此，设备本身和网络连接都存在特别大的攻击面，给予攻击者远程与设备互动的机会。物联网安全必须考虑大量的入口点，以保护资产。

（2）物联网设备的资源限制，导致其缺乏计算能力来整合安全功能（如传输加密）或

安全产品（如防火墙）。有些设备甚至几乎不具备与其他设备连接的能力，例如，仅支持蓝牙技术的物联网设备。2020 年，一位网络安全专家利用一个蓝牙漏洞，在不到 90 秒的时间内入侵了一辆特斯拉 Model X 的系统。其他依靠 FOB（无线）钥匙打开和启动的汽车也可能因类似的原因而遭遇攻击。

（3）升级和维护的复杂性。物联网设备制造商发布补丁和更新的速度通常比生产通用操作系统和其他通用软件的机构慢。物联网设备在数周或数月内未打补丁和运行过时的软件是很常见的。甚至在一些组织中，一些设备从未得到软件或固件升级。

（4）物联网设备的终端消费者或操作者缺乏安全知识或安全意识。物联网设备本身有很多安全措施，但终端消费者往往是任何物联网安全系统中最薄弱的一环。物联网设备体积小、易于携带，已经成为消费者日常生活和工作离不开的伙伴。如果算上智能手机，大多数人都有自己的物联网设备。问题是大多数人没有意识到安全风险，如经常将设备连接到未知网络。

（5）行业对于物联网安全缺乏洞察和远见。在数字化转型的过程中，汽车、物流和医疗健康等行业大规模选用物联网设备，以提升生产力和效率。与此同时，物联网的数据泄露风险尚未得到这些行业的充分重视，没有得到充分的、必要的资金和资源投入。

4．物联网与数据安全的关系

在物联网中，数据的形态发生了较大的改变，数据的速度与往日相比也不可同日而语。此外，物联网设备形态及部署方式的复杂性，为数据安全和隐私保护带来了巨大挑战。

当前，物联网设备存在近端、无线和广域网互联等多种连接场景，以用于传输各种不同的数据。展望未来，物联网将更加独立于人类的干预，通过人工智能在物联网中的深度应用，物和物可能会更多、更主动地沟通。在如此复杂的连接和交互场景，数据的机密性、完整性和隐私保护问题变得更加复杂。

识别和应对物联网数据安全的挑战，还应从最基础的数据识别和分类入手，应考虑的问题包括系统中有哪些物联网设备，如终端设备、网关等，在这些设备上及传输过程中，有哪些数据需要保护？例如，敏感数据包含录音、视频、地理位置和运动健康数据等，需要做机密性和完整性防护。对于物理世界的数据，如传感器采集的环境温度，未必需要机密性防护，但可能需要完整性防护（防篡改）和可用性保护。

1.6.4　零信任架构与数据安全

现代社会的新闻中充斥着大量的对网络攻击和用户数据泄露的报道，而且其中不乏对 IT 业界知名大机构的报道。面对恶意的、复杂的、有组织的和新型的网络攻击，使用单点的安全解决方案，如防火墙、反病毒软件、数据泄露防护（DLP）等来应对并不现实。因此，以组织或网络的视角，构建统一的安全架构非常有必要。

1．分层防御

在网络、产品和解决方案的设计中，最常使用“分层安全”（Layered Security）或“纵深防御”（Defense in Depth）的原则。这些设计采用多种工具、多种技术来阻止或缓解

攻击，并提供全面的风险和安全管理。

这种分层的安全架构常常被形象地比喻为分层防御、纵深防御或洋葱式防御，如图 1-11 所示。

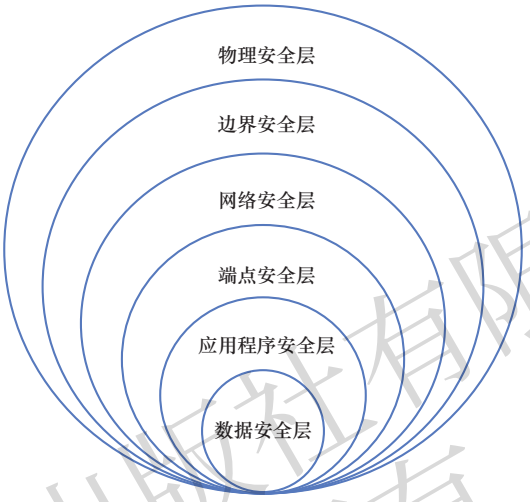

图 1-11　分层防御示意图

传统的网络安全架构通过逐层构筑防御来最终保证数据的安全。在现代的复杂网络环境中，人员可能不可靠，网络边界可能不清晰，应用可能不被信任，系统可能有漏洞或后门，这些都可能损害到最终的保护目标——数据的安全性。

2．零信任安全架构

零信任安全架构是对纵深防御架构的补充和增强。具体而言，零信任安全架构将网络防御从网络边界转移到资源或数据。

零信任的理念一直存在于网络安全和信息安全业界。2004 年，耶利哥论坛提出了基于网络位置限制隐式信任的思想，并提出了“去边界化”的概念。后来，John Kindervag 在 Forrester 机构工作时，发明了零信任一词。这项工作包括零信任概念和网络架构模型，并改进了在耶利哥论坛上讨论的概念。

零信任架构（Zero Trust Architecture，ZTA）策略指不再基于系统的物理或网络位置（局域网或因特网）授予系统的隐式信任的策略。当需要资源时才授予对数据资源的访问权，并在建立连接之前执行对用户或设备的身份认证。

随着远程办公、跨组织协作、云转型等场景的普及，企业内网这个概念进一步模糊化。因此，基于网络分段和网络隔离的防御有其局限性，ZTA 的重点是保护资源不受非授权的访问。

3．零信任安全架构与数据安全的关系

零信任架构是一种端到端的网络和数据安全方法，包括身份、凭证、访问管理、操作、终端、宿主环境和互联基础设施。零信任是一种侧重于数据保护的架构方法，初始的

重点是将资源访问限制在那些“需要知道”的人身上。传统上，组织的网络专注于边界防御，内网的授权用户可以广泛地访问资源。因此，网络内未经授权的横向移动一直是组织面临的最大挑战之一。

在本质上，零信任架构提供对信息系统和服务的精细化访问决策。也就是说，授权和批准的主体（用户 / 计算机）可以访问数据，但不包括所有其他主体，即攻击者。进一步，“数据”可以推广到“资源”，零信任架构的范围相应扩展到对打印机、计算资源、物联网执行器的访问。

典型的资源访问控制架构如图 1-12 所示。主体（用户或计算机）请求访问资源，由访问控制体系的策略决策点（Policy Decision Point，PDP）和相应的策略执行点（Policy Enforcement Point，PEP）授予访问权限。访问控制体系需要完成身份认证、授权和鉴权，并且隐式信任区域需要尽可能缩小。

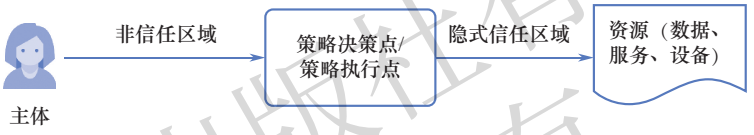

图 1-12　典型的资源访问控制架构

以机场的访问控制案例为例。所有旅客均需出示身份证或护照等“身份凭据”，以及用有效机票作为“认证凭据”，通过机场安检点（类比于 PDP/PEP）进入候机区。乘客可以在候机区内闲逛，所有乘客都有一个共同的信任级别。在这个模型中，隐式信任区域是公共的候机区。但是，登机需要再次身份认证，凭借特定航班的有效机票才能上飞机。而要进入机场工作区，则需要额外的身份认证手段，如门禁卡。根据所请求访问的“资源”不同，需要的访问控制方式也不同。

基于访问控制的设计理念，在 PDP/PEP 之后，将默认信任对资源的访问。为了使 PDP/PEP 尽可能细致，隐式信任区必须尽可能小。零信任架构提供了技术和能力，以允许 PDP/PEP 更接近资源。其思想是对网络中从参与者（或应用程序）到数据的每个业务流都进行身份验证和授权。

1.6.5　量子计算与数据安全

世界正在朝着量子革命前进。量子计算机可提供前所未有的颠覆性的处理能力。在密码学和密钥管理领域，基于量子理论与量子计算的特殊性，既存在理论上的完美保密方案，也面临着前所未有的挑战。

1．量子计算对数据安全的正面影响

将量子计算用于密码学领域，称为量子密码学。量子密码学聚焦密码学场景的关键问题，特别是量子密钥分发场景的理论和实践问题。不同于传统的对称和非对称密码学技术

高度依赖数学，量子密码学更像是物理学。它利用光子（粒子 / 光波）及其内在属性来开发理论上“坚不可摧”的密码系统。通过在光链路上发送光子来传递密钥，依赖在不干扰系统的情况下无法测量系统的量子状态这个量子物理学的典型特征，可以确保理论上不可窃听的高度安全性。

量子密码学主要研究量子密钥分发（Quantum Key Distribution，QKD）。从理论上讲，这是一种无须身份验证的密钥交换的可靠方式。

QKD 建立在公钥交换系统的基础上，利用了单个光子的量子纠缠的奇异特性。目前正在探索的 QKD 系统一般部署在标准的光纤电缆上，但是并不使用它们来发送数据信号，而是用于发送单独的光子。

由于这些单独的光子与保留在发送方系统中的光子纠缠在一起，因此对它们的任何拦截都会导致波函数坍缩（量子物理学中的术语），从而使发送方识别出其通信可能被拦截或被监听。

我国在 QKD 领域的研究和实践方面走在世界的前列，中国科学技术大学潘建伟院士及其团队主导建设了连接北京和上海的专用量子通信光缆。2019 年秋天，美国的量子交易所（Quantum Xchange）宣称建设了连接纽约市的金融机构和新泽西州的数据中心的美国第一个商业量子分发网络。该网络可以通过现有光纤网络交换量子密钥。

截至目前，QKD 的应用仍然相当小众。即使不计算独立光纤网络的可能支出，单独的发射机和接收机的成本也都可能超过 10 万美元。另一方面，QKD 系统也存在比较明显的局限性。QKD 依赖交换纠缠的单个光子，因此无法切实地用作通信系统。相反，它将仅限于交换加密密钥。同时，绝大多数威胁并非是在传输链路上拦截加密密钥并直接解密密文的。

除加密密钥的交换场景外，在设备上执行的加密过程、完整性保护过程或身份验证过程是 QKD 更大的脆弱点。而量子密钥分发系统无法为这些场景提供有效的安全性防护。

在英国政府发布的《量子安全技术》（*Quantum Security Technologies*）白皮书中，不鼓励使用 QKD，称其似乎引入了新的潜在攻击途径，对硬件的依赖性不符合成本效益，QKD 的有限范围使其不适合应对未来的挑战，而后量子密码学则是更好的选择。在更多的业务场景需要量子网络之前，QKD 可能仍然是一个小众的解决方案。

2．量子计算对数据安全的负面影响

与此对应的是研究量子计算的发展对于现有密码学和密钥管理技术的负面影响，从而确定量子计算时代仍然安全的密码算法和密钥管理技术。这个研究方向一般称为后量子密码学（Post Quantum Cryptography，PQC）或量子安全的密码学（Quantum-safe Cryptography）。

互联网、大部分计算机和通信系统的安全性能都依赖安全和有效的加密算法，特别是通信双方的身份认证、完整性校验、传输加密等算法。例如，迪非－赫尔曼密钥交换算法（Diffie-Hellman Key Exchange）、RSA 数字签名算法，以及使用共享密钥进行加密的 AES 对称加密算法。近年来，椭圆曲线密码学等非对称算法也得到了越来越多的应用。如果在

未来几十年中在量子计算方面取得突破，那么上述非对称算法的安全性很可能会面临重大风险。

本质上，RSA 和 Diffie-Hellman 算法的安全性基于计算的困难度。例如，RSA 基于大整数的因式分解问题和离散对数的计算困难问题。大整数的因式分解问题是指，计算机可以很迅速地计算出两个非常大的素数的乘积，但是很难将一个非常大的合数分解为两个质因数。

1994 年，美国数学家 Peter Shor 发明了 Shor 算法，该算法可以利用理想的量子计算机求解大整数因子分解，即具备理论上破解 RSA 和 Diffie-Hellman 算法的能力。Shor 算法可以在近似多项式的时间内（近似 log$N$，N 是输入的合数的长度）完成质因数分解。如果可以构建足够大规模的量子计算机来运行 Shor 算法，则可以非常有效地分解大整数，进而使得 RSA 和 Diffie-Hellman 的相关算法被彻底攻破。这对互联网安全会造成巨大的影响。

Matteo Mariantoni 曾在 2014 年预测，15 年以内，花费 100 万美元构建的量子计算机可以破解所有现存的密码算法。值得庆幸的是，截至 2020 年 10 月底，尚无法确认实用的量子计算机什么时候能够成为现实。因此，基于公钥的密码算法并不会面临马上被破解的风险。

理想的量子计算机仅能将对称密钥的有效长度减半。基于足够长度的对称密钥的加密算法，如 AES-256，在理论上的量子时代仍然安全，短期内无须考虑替换。

应用于 AES-128 的 Grover 算法需要大约 2^{64} 次迭代计算，而这些迭代无法有效地并行化。由于量子计算机运行非常慢（按照每秒的操作数而言），非常昂贵，并且难以从发生故障的量子计算机转移量子状态，因此，即使量子计算机集群也很难成为对称算法的实际威胁。根据 NIST PQC（后量子密码学）标准化项目中的评估标准，AES-128 和 SHA-256 均具有量子抗性。

3．量子时代的密码算法演进趋势

目前，学术界和工业界对量子计算机展开了相关的研究，目标是设计和开发后量子时代仍然安全的算法。这一系列量子安全算法的目标是在经典计算机而非量子计算机上运行，以抵抗量子计算机的攻击。部分算法基于除整数分解和离散对数以外的其他数学难题设计，如基于格（Lattice-based）的密码体制。其中，走在这项研究前列的是美国国家标准与技术研究院（NIST）。

美国国家标准与技术研究院于 2016 年启动后量子算法的征集，在多轮标准化流程之后，NIST 选择了四种作为后量子密码学（PQC）算法进行标准化：CRYSTALS–KYBER，以及三种数字签名方案：CRYSTALS–Dilithium、FALCON 和 SPHINCS+（无状态基于哈希的签名）。这些算法以 FIPS 203、204、205 的形式发布，预计将在 2024 或 2025 年正式发布为推荐标准。此后，新的标准化算法可能会被添加到 X.509、IKEv2、TLS 和 JOSE 等安全协议中，并在各个行业中得到应用。

IETF 等标准组织和行业组织也在研究后量子密码算法在通信协议中的应用。例如，IETF 密码论坛研究小组完成了两种有状态的基于散列的签名算法 XMSS 和 HSS/LMS 的标准，并与 NIST 合作进行标准化。扩展的 Merkle 签名方案（eXtended Merkle Signature Scheme，XMSS）和 Leighton-Micali 签名（Leighton-Micali Signature，LMS）是目前可用于生产系统的后量子密码算法，可应用于设备固件更新时的签名校验等场景。

学术界和工业界的最新报告称，未来十年内，能够大规模破解现有密码学的量子计算机极不可能产生。人们也普遍认为，量子计算机不会对对称算法构成大的威胁。IETF、3GPP 等标准化组织和各行业都在密切关注 NIST PQC 算法标准化的结果。

量子计算机很可能会对某些行业产生很大的颠覆性影响，但可能在几十年内不会对非对称密码学构成实际威胁，也很可能永远不会对对称密码学构成实际威胁。

需要在很长一段时间内保护信息的组织应该开始考虑后量子密码学。而对于民用的密码学使用场景，当前采纳的椭圆曲线密码学算法和 RSA 算法在较长的一段时期内仍可以放心使用。

习　题

1．什么是数据？什么是数据安全？
2．简述 DIKW 金字塔模型。
3．简述数据安全模型的三个核心要素。
4．数据生命周期包含哪些关键的环节？
5．简述数据的三种状态及其安全防护机制。
6．数字安全的管理分为哪几个维度？
7．数据安全风险评估包括哪些内容？
8．云化部署常采用的四种模型分别是什么？
9．零信任安全架构的策略是什么？
10．物联网安全分为哪三个层面的问题？
11．数据安全管理目标分为哪三个维度？
12．基本的数据安全风险评估涉及哪三个因素？
13．人工智能如何赋能数据安全？
14．分析云计算与数据安全的关系。
15．分析物联网与数据安全的关系。
16．分析零信任架构与数据安全的关系。
17．分析量子计算与数据安全的关系。