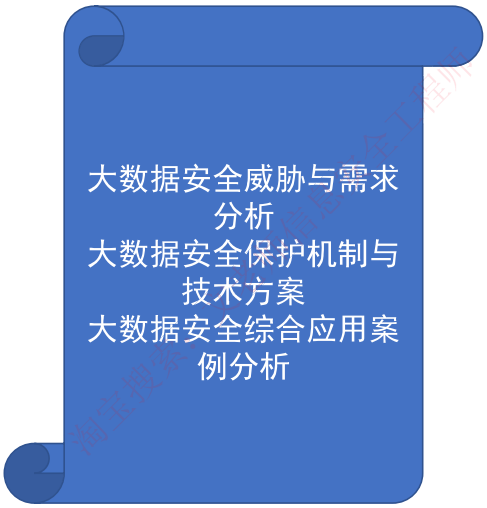
大数据安全需求分析与安全保护工程



文老师软考教育

文老师软考教育

大数据安全威胁与需求分析

◆大数据相关概念发展：一般来说，大数据是指非传统的数据处理工具的数据集，具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低等特征。大数据的种类和来源非常多，包括结构化、半结构化和非结构化数据。

◆有关大数据的新兴网络信息技术应用不断出现，主要包括大规模数据分析处理、数据挖掘、分布式文件系统、分布式数据库、云计算平台、互联网和存储系统。

◆大数据安全威胁分析：

1．“数据集”安全边界日渐模糊，安全保护难度提升。多源、海量、异构、分布存储等大数据新技术导致数据集的安全边界日渐模糊，造成基于网络安全边界的安全防护措施难以完全有效。数据交易和共享促使数据流动日益频繁，静态安全措施难以完全满足数据安全保障要求。复杂分布式计算环境使得网络攻击的影响增大，单个节点遭受侵害影响整个系统安全，例如数据存储设备、域名服务器、认证服务器等。

2. 敏感数据泄露安全风险增大。数据丢失或被盗取，有可能影响国家安全、社会安全和经济安全。

3．数据失真与大数据污染安全风险。攻击者利用数据输入或数据平台缺陷，构造恶意数据并将其注入数据处理系统中，干扰数据处理系统的正常运行或误导计算。典型事例有电商产品的评分、网站访问流量、网页虚假排名等。另一方面，数据获取隐患也会导致人工智能的安全问题。

大数据安全威胁与需求分析

4．大数据处理平台业务连续性与拒绝服务。随着大数据的应用普及，许多关键业务依赖于大数据处理平台的连续稳定运行。例如，电商服务平台、金融服务平台等。

5．个人数据广泛分布于多个数据平台，隐私保护难度加大。

6．数据交易安全风险。数据交易是指数据供方和需方之间以数据商品作为交易对象，进行的以货币交换数据商品，或者以数据商品交换数据商品的行为。数据交易促进商业合作，但也形成潜在的安全风险。如非法数据交易、虚假数据交易、交易服务不完整、交易数据汇聚导致敏感数据泄露、跨境数据流动安全等安全风险。

7．大数据滥用。万物互联，不同数据集之间蕴含潜在关联关系。随着大数据分析技术发展、数据的不断累积，大数据分析可以发现更多、更深入的关联关系。例如，利用大数据技术和不同的生命科学相关大数据，可以开发针对特定人群的生物病毒，容易对群体的生命安全产生重大威胁。综合关联分析微信图片数据、智能手机位置数据，可以识别到自然人，挖掘出个人隐私信息。

大数据安全威胁与需求分析

文老师软考教育

◆大数据安全需求分析：大数据安全需求涉及多个方面，主要内容如下。

1．大数据自身安全。大数据应用依赖于可信的数据。目前，基于数据驱动的安全威胁已经出现，如虚假的数据可以干扰机器学习。大数据安全涉及数据的采集、存储、使用、传输、共享、发布、销毁等全生命周期的多个方面，具体安全包括数据的真实性、实时性、机密性、完整性、可用性、可追溯性。

2．大数据安全合规。建立大数据安全合规管理机制，满足不同国家和地区、行业部门的数据安全政策法规要求。

3．大数据跨境安全。随着跨境电商、跨境交易等国际应用发展，数据跨境流动成为必然。目前，不同国家和地区的数据保护法规对数据跨境流动的要求存在差异性。

4．大数据隐私保护。针对大数据涉及的敏感个人信息．需要相应的隐私保护技术，防止个人敏感数据泄露。

5．大数据处理平台安全。按照数据处理过程，大数据处理平台涉及物理环境、网络通信、操作系统、数据库、应用系统、数据存储。

6．大数据业务安全。大数据产业应用的发展促进了数据流动和共享，需要新的数据安全措施保护数据的安全流动和共享，防止数据扩散、数据滥用问题。需要部署大数据业务安全管理措施，建立数据滥用监测机制、数据受控使用机制，防止数据非法交易及恶意滥用。

7．大数据安全运营。建立大数据运营安全机制，如大数据分类分级、大数据安全服务、大数据平台的安全维护。

大数据安全保护机制与技术方案

◆大数据安全保护机制：大数据安全保护是一个综合的、复杂性的安全工程，涉及数据自身安全、数据处理平台安全、数据业务安全、数据隐私安全、数据运营安全以及数据安全法律政策与标准规范。
◆围绕大数据的安全保护，常见的基本安全机制主要有数据分类分级、数据源认证、数据溯源、数据用户标识和鉴别、数据资源访问控制、数据隐私保护、数据备份与恢复、数据安全审计与监测、数据安全管理等。

◆大数据自身安全保护技术：大数据自身安全是指有关数据本身的安全问题，如数据的真实性、数据的完整性、数据的机密性、数据的准确性等。

◆大数据平台安全保护技术：大数据平台涉及物理环境、网络通信、操作系统、数据库、应用系统、数据存储等安全保护。通常采用安全分区、防火墙、系统安全加固、数据防泄露等安全技术用于保护大数据平台。

◆大数据业务安全保护技术：大数据业务安全主要包括业务授权、业务逻辑安全、业务合规性等安全内容。其中，业务授权主要基于角色的访问控制技术，按照业务功能的执行所需要的权限进行分配。业务逻辑安全针对业务流程进行安全控制，避免安全缺陷导致业务失控。业务合规性是指业务满足政策法规及安全标准规范要求。敏感数据安全检查、系统安全配置基准数据监控等技术常用于解决业务合规性安全需求。

大数据安全保护机制与技术方案

文老师软考教育

◆大数据隐私安全保护技术：隐私是指与个体相关的非公开的信息。隐私保护成为大数据时代新的安全需求。针对个人信息安全保护，国家颁布了《信息安全技术个人信息安全规范》（于2020年10月1日实施）等法规政策及标准规范。围绕隐私保护，主要的技术有数据身份匿名、数据差分隐私、数据脱敏、数据加密、数据访问控制等。

◆大数据运营安全保护技术：大数据运营安全是指大数据平台及数据的运行维护及数据资源经营过程的安全。大数据平台及数据的运行维护包括大数据处理系统的安全维护、安全策略更新及安全设备配置、数据资源容灾备份、安全事件监测与应急响应等。网络入侵检测、网络安全态势感知、网络攻击取证、网络威胁情报分析、安全堡垒机等技术常用于大数据平台运维安全保护。

◆数据资源经营过程安全涉及数据使用、数据交易、数据跨境流动等安全问题。数据脱敏、数据监控、数据安全网关等常用于数据经营安全保护。

◆大数据安全标准规范：大数据安全标准规范有利于提升数据安全整体保障能力。全国信息安全标准化技术委员会在2016年成立大数据安全标准特别工作组，主要负责制定和完善我国大数据安全领域标准体系，组织开展大数据安全相关技术和标准研究。目前，已制定的国家标准主要有《信息安全技术个人信息安全规范》《信息安全技术大数据服务安全能力要求》《信息安全技术大数据安全管理指南》《信息安全技术数据交易服务安全要求》《信息安全技术个人信息去标识化指南》等。

大数据安全综合应用案例分析

◆阿里巴巴大数据安全实践：从业务、数据和生态三个层面来保护数据安全与隐私。

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

2．数据安全管控。此大数据安全平台基于数据业务链路构建了全面的数据管控体系，主要包括数据加工前、数据加工中、数据加工后、数据合规等方面的数据安全管控，如表所示。

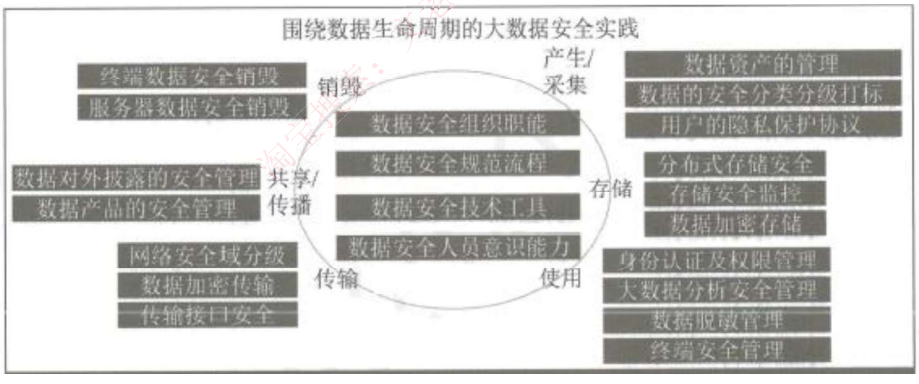
数据处理阶段	数据安全措施
数据加工前	数据来源合法性评估、数据用途合理性评估 数据相关方权属评估、消费者隐私评估 商业秘密合规性评估、数据标签规范性管理 数据分级分类的管理
数据加工中	多租户隔离控制、统一元数据管理 数据探索容器、数据建模风控 数据脱敏控制、代码可信审核 数据画像风控、Pn防护（K匿名和差分） 统一身份与鉴权、访问监控与审计
数据加工后	数据血缘安全控制、 数据营销渠道管控 数据效果回流风控、数据染色取证追溯 数据加密平台服务、数据环境安全容器

大数据安全综合应用案例分析

文老师软考教育

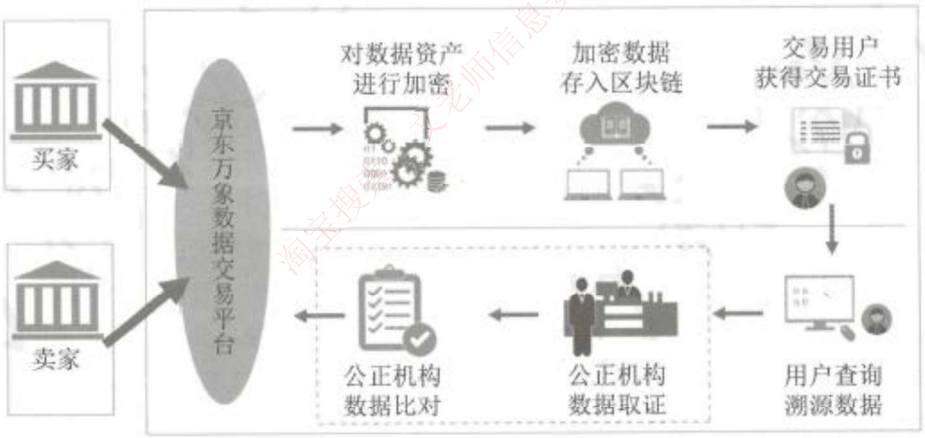
3．生态安全管控。通过对数据ISV的准入准出、基于垂直化行业的标签体系建立以及数据生态的市场管理机制建立，确保业务和安全间找到有效的平衡点。

◆阿里巴巴形成了以数据生命周期为中心的大数据安全管理理念，以数据为中心，围绕数据生命周期，对组织机构的数据进行安全保障，有效地控制了数据安全风险，提升了公司自身及生态伙伴的数据安全能力，促进了生态内数据资源的流通与共享，更大地发挥了数据的价值。其中，数据安全能力成熟度模型从组织建设、制度流程、技术工具、人员能力、数据生命周期通用安全等方面评估大数据安全能力成熟度，以便明确大数据安全保障能力的提升方向。



大数据安全综合应用案例分析

◆京东大数据安全实践：如图所示，京东万象数据服务平台利用区块链技术对流通的数据进行确权溯源，数据买家在数据服务平台上购买的每一笔交易信息都会在区块链中存储起来，数据买家通过获得交易凭证可以看到该笔交易的数字证书以及该笔交易信息在区块链中的存储地址，待买家需要进行数据确权时，登录用户中心进入查询平台，输入交易凭证中的相关信息，查询到存储在区块链中的该笔交易信息，从而完成交易数据的溯源确权。



大数据安全综合应用案例分析

文老师软考教育

◆上海数据交易中心安全保护：制定规制+技术的模式，即交易规则和技术共同保障交易安全。
◆其中，公布的上海市数据交易准则有个人数据保护原则、数据互联规则、流通数据处理准则、流通数据禁止清单、交易要素标准体系。

◆个人数据保护原则的主要内容如下：
◆告知同意原则。初始收集的数据被再次使用或再处理时，如果再使用的目的与初始目的不一致的，则须告知数据主体并取得其同意。如果数据主体不同意的，不得对该个人数据进行任何使用或处理。
◆禁止公开原则。在任何情形下均不得擅自公开、向第三人提供带有身份标识的个人数据。
◆选择退出原则。任何使用个人数据进行的推销、推介和广告活动，应当给予接受人以退出选择。
◆数据正确原则。数据持有人应使个人数据符合处理目的的要求，必要时及时更新和更正。
◆维护权益原则。成员应设立个人隐私投诉机制，积极响应和解决个人投诉。
◆应急补救原则。一旦出现个人数据泄露事故，成员应当及时通知有关个人并采取补救措施，在隐私地位无法恢复时应及时给予赔偿。

◆流通数据处理基本原则主要如下：
◆保护个人权益原则。保护个人隐私和其他合法权益。
◆诚实守信原则。遵守各种自律规范，忠实履行承诺和协议。
◆保护正当数据权益原则。尊重他人的数据收集和处理劳动成果，维护公平的数据利用秩序。
◆数据安全原则。保障数据收集、存储、传输和使用各个环节的安全，防范数据泄露的风险。

大数据安全综合应用案例分析

◆华为大数据安全实践：华为大数据分析平台FusionInsight基于开源社区软件Hadoop进行功能增强，提供企业级大数据存储、查询和分析的统一平台。平台的安全措施主要分析如下。
1．网络安全。FusionInsight集群支持通过网络平面隔离的方式保证网络安全。

2．主机安全。通过对FusionInsight集群内节点的操作系统安全加固等手段保证节点正常运行，包括更新最新补丁、操作系统内核安全加固、操作系统权限控制、端口管理、部署防病毒软件等。

3．用户安全。平台提供身份认证、权限控制、审计控制等安全措施，防止用户假冒、越权、恶意操作等安全威胁。其中，FusionInsight的身份认证使用LDAP作为账户管理系统，并通过Kerberos对账户信息进行安全认证；权限控制基于用户和角色的认证统一体系，遵从账户／角色RBAC（基于角色的访问控制）模型，实现通过角色进行权限管理，对用户进行批量授权管理，降低集群的管理难度；FusionInsight审计日志中记录了用户操作信息，可以快速定位系统是否遭受恶意的操作和攻击。

4．数据安全。平台提供集群容灾、备份、数据完整性、数据保密性等安全服务，以保证用户数据的安全。

大数据安全综合应用案例分析

文老师软考教育

◆科学数据安全管理：《科学数据管理办法》部分具体要求如下：
◆第二十五条涉及国家秘密、国家安全、社会公共利益、商业秘密和个人隐私的科学数据，不得对外开放共享；确需对外开放的，要对利用目的、用户资质、保密条件等进行审查，并严格控制知悉范围。

◆第二十六条涉及国家秘密的科学数据的采集生产、加工整理、管理和使用，按照国家有关保密规定执行。主管部门和法人单位应建立健全涉及国家秘密的科学数据管理与使用制度，对制作、审核、登记、拷贝、传输、销毁等环节进行严格管理。对外交往与合作中需要提供涉及国家秘密的科学数据的，法人单位应明确提出利用数据的类别、范围及用途，按照保密管理规定程序报主管部门批准。经主管部门批准后，法人单位按规定办理相关手续并与用户签订保密协议。

◆第二十七条主管部门和法人单位应加强科学数据全生命周期安全管理，制定科学数据安全保护措施；加强数据下载的认证、授权等防护管理，防止数据被恶意使用。对于需对外公布的科学数据开放目录或需对外提供的科学数据，主管部门和法人单位应建立相应的安全保密审查制度。
◆第二十八条法人单位和科学数据中心应按照国家网络安全管理规定，建立网络安全保障体系，采用安全可靠的产品和服务，完善数据管控、属性管理、身份识别、行为追溯、黑名单等管理措施，健全防篡改、防泄露、防攻击、防病毒等安全防护体系。
◆第二十九条科学数据中心应建立应急管理和容灾备份机制，按照要求建立应急管理系统，对重要的科学数据进行异她备份。

大数据安全综合应用案例分析

◆支付卡行业数据安全规范：在国际上，支付卡行业数据安全标准(PCI-DSS)是PCI安全标准委员会制定的数据安全规范。PCI-DSS的规范目标在于严格控制对支付卡持卡人数据的处理、存储和传输，以保障银行卡用户在线交易的安全。PCI-DSS按每年交易量将商家分为四个等级，对不同等级的商家提出不同强度的安全要求。PCI-DSS适用于所有涉及信用卡支付的企业。PCI-DSS包括以下6大类要求：

◆构建和维护安全的网络；
◆保护持卡人数据；
◆维护漏洞管理程序；
◆实施严格的存储控制措施；
◆定期监控和测试网络；
◆维护信息安全策略。

谢谢！