ICS 点击此处添加 ICS 号
点击此处添加中国标准文献分类号

中 华 人 民 共 和 国 国 家 标 准

GB/T XXXXX—XXXX

信息安全技术　大数据安全管理指南

Information Security Technology —— Big Data Security Management Guide

点击此处添加与国际标准一致性程度的标识

（报批稿）

（本稿完成日期：2018 年 2 月 20 日）

XXXX－XX－XX 发布　　XXXX－XX－　实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会　发布

目　　次

前　　言 III
引　　言 IV
1 范围 1
2 规范性引用文件 1
3 术语和定义 1
4 大数据安全管理概述 2
4.1 大数据安全管理目标 2
4.2 大数据安全管理的主要内容 2
4.3 大数据安全管理角色及责任 2
5 大数据安全管理基本原则 3
5.1 职责明确 3
5.2 安全合规 3
5.3 质量保障 3
5.4 数据最小化 4
5.5 责任不随数据转移 4
5.6 最小授权 4
5.7 确保安全 4
5.8 可审计 4
6 大数据安全需求 4
6.1 保密性 4
6.2 完整性 5
6.3 可用性 5
6.4 其他需求 5
7 数据分类分级 5
7.1 数据分类分级原则 5
7.2 数据分类分级流程 5
7.3 数据分类方法 6
7.4 数据分级方法 6
8 大数据活动及安全要求 6
8.1 大数据的主要活动 6
8.2 数据采集 7
8.3 数据存储 7
8.4 数据处理 8
8.5 数据分发 8
8.6 数据删除 9
9 评估大数据安全风险 9
9.1 概述 9

9.2 资产识别 9
9.3 威胁识别 10
9.4 脆弱性识别 10
9.5 已有安全措施确认 10
9.6 风险分析 10
附 录 A （资料性附录）电信行业数据分类分级示例 11
附 录 B （资料性附录）生命科学大数据风险分析示例 13
附 录 C （资料性附录）大数据安全风险 14
参考文献 15

前　言

本标准按照 GB/T 1.1—2009《标准化工作导则第 1 部分：标准的结构和编写》给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由全国信息安全标准化技术委员会（SAC/TC260）提出并归口。

本标准起草单位：四川大学、中国电子技术标准化研究院、清华大学、中国移动有限公司、深圳市腾讯计算机系统有限公司、阿里云计算有限公司、广州赛宝认证中心服务有限公司、中电长城网际系统应用有限公司、腾讯云计算（北京）有限责任公司、华为技术有限公司、成都超级计算中心有限公司、陕西省信息化工程研究院、北京奇虎科技有限公司、北京奇安信科技有限公司、银联智慧信息服务（上海）有限公司、北京华宇软件股份有限公司、中国电子科技网络信息安全有限公司等单位。

本标准主要起草人：陈兴蜀、罗永刚、叶晓俊、上官晓丽、叶润国、杨露、金涛、闵京华、常玲、陈雪秀、胡影、代威、刘小茵、杨思磊、王文贤、李克鹏、赵蓓、王永霞、何军、张丽佳、张勇、郑新华、王建波、金睿、高冀鹏、彭凝多。

引　言

大数据技术的发展和应用影响着国家的治理模式、企业的决策架构、商业的业务模式以及个人的生活方式。我国大数据仍处于起步发展阶段，各地发展大数据积极性高，行业应用得到快速推广，市场规模迅速扩大。在面向大量用户的应用和服务中，数据采集者希望能获得更多的信息，以提供更加丰富、高效的个性化服务。随着数据的聚集和应用，数据价值不断提升。而伴随大量数据集中，新技术不断涌现和应用，使数据面临新的安全风险，大数据安全受到高度重视。

目前拥有大量数据的组织的管理和技术水平参差不齐，有不少组织缺乏技术、运维等方面的专业安全人员，容易因数据平台和计算平台的脆弱性遭受网络攻击，导致数据泄露。在大数据的生命周期中，将有不同的组织对数据做出不同的操作，关键是要加强掌握数据的组织的技术和管理能力的建设，加强数据采集、存储、处理、分发等环节的技术和管理措施，使组织从管理和技术上有效保护数据，使数据的安全风险可控。

本标准指导拥有、处理大数据的企业、事业单位、政府部门等组织做好大数据的安全管理、风险评估等工作，有效、安全地应用大数据，采用有效技术和管理措施保障数据安全。

信息安全技术 大数据安全管理指南

1 范围

本标准提出了大数据安全管理基本原则，指导组织开展大数据安全需求分析、数据分类分级、大数据活动和风险评估等安全管理工作。

本标准适用于各类组织进行数据安全管理，也可供第三方评估机构参考。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 7072—2002 信息分类和编码的基本原则与方法

GB/T 25069—2010 信息安全技术 术语

GB/T 31167—2014 信息安全技术 云计算服务安全指南

GB/T 35274—2017 信息安全技术 大数据服务安全能力要求

GB/T 35295—2017 信息技术 大数据 术语

3 术语和定义

GB/T 25069—2010，GB/T 35295—2017，GB/T 20984-2007中界定的以及下列术语和定义适用于本文件。

3.1

大数据 big data

具有数量巨大、种类多样、流动速度快、特征多变等特性，并且难以用传统数据体系结构和数据处理技术进行有效组织、存储、计算、分析和管理的数据集。

3.2

组织 organization

由作用不同的个体为实施共同的业务目标而建立的结构。组织可以是一个企业、事业单位、政府部门等。

3.3

大数据平台 big data platform

采用分布式存储和计算技术，提供大数据的访问和处理，支持大数据应用安全高效运行的软硬件集合，包括监视大数据的存储、输入/输出、操作控制等大数据服务软硬件基础设施。

3.4

大数据环境 big data environment

开展大数据活动所涉及的数据、平台、规程及人员等的要素集合。

3.5

大数据活动 big data activity

组织针对大数据开展的一组特定任务的集合，大数据活动主要包括采集、存储、处理、分发、删除等。

4 大数据安全管理概述

4.1 大数据安全管理目标

组织实现大数据价值的同时，确保数据安全。组织应：

a) 满足个人信息保护和数据保护的法律法规、标准等要求；
b) 满足大数据相关方的数据保护要求；
c) 通过技术和管理手段，保证自身控制和管理的数据安全风险可控。

4.2 大数据安全管理的主要内容

大数据安全管理主要包含以下内容：

a) 明确数据安全需求。组织应分析大数据环境下数据的保密性、完整性和可用性所面临的新问题，分析大数据活动可能对国家安全、社会影响、公共利益、个人的生命财产安全等造成的影响，并明确解决这些问题和影响的数据安全需求。
b) 数据分类分级。组织应先对数据进行分类分级，根据不同的数据分级选择适当的安全措施。
c) 明确大数据活动安全要求。组织应理解主要大数据活动的特点，可能涉及的数据操作，并明确各大数据活动的安全要求。
d) 评估大数据安全风险。组织除开展信息系统安全风险评估外，还应从大数据环境潜在的系统的脆弱点、恶意利用、后果等不利因素，以及应对措施等评估大数据安全风险。

4.3 大数据安全管理角色及责任

4.3.1 概述

组织应建立大数据安全管理组织架构，根据组织的规模、大数据平台的数据量、业务发展及规划等明确不同角色及其职责，至少包含以下角色：

a) 大数据安全管理者：对组织大数据安全负责的个人或团队。大数据安全管理者负责数据安全相关领域和环节的决策，制定并审议数据安全相关制度，监督执行和组织落实业务部门数据安全相关工作。
b) 大数据安全执行者：是执行组织数据安全相关工作的个人或团队。大数据安全执行者负责数据安全相关领域和环节工作的执行，制定数据安全相关细则，落实各项安全措施，配合大数据安全管理者开展各项工作。
c) 大数据安全审计者：负责大数据审计相关工作的个人或团队。大数据安全审计者对安全策略的适当性进行评价，帮助检测安全违规，并生成安全审计报告。

4.3.2 大数据安全管理者的职责

大数据安全管理者的具体职责有：

a) 确定数据的分类分级初始值，制定数据分类分级指南。与提供大数据的业务部门合作，确定数据的安全级别；
b) 综合考虑法律法规、政策、标准、大数据分析技术水平、组织所处行业特殊性等因素，评估数据安全风险，制定数据安全基本要求；
c) 对数据访问进行授权，包括授权给组织内部的业务部门、外部组织等；
d) 建立相应的数据安全管理监督机制，监视数据安全管理机制的有效性；
e) 负责组织的大数据安全管理过程，并对外部相关方（如：数据安全的主管部门、数据主体等）负责。

4.3.3 大数据安全执行者的职责

大数据安全执行者的主要职责有：

a) 根据大数据安全管理者的要求实施安全措施；
b) 为大数据安全管理者授权的相关方分配数据访问权限和机制；
c) 配合大数据安全管理者处置安全事件；
d) 记录数据活动的相关日志。

4.3.4 大数据安全审计者的职责

大数据安全审计者的主要职责有：

a) 审核数据活动的主体、操作及对象等数据相关属性，确保数据活动的过程和相关操作符合安全要求；
b) 定期审核数据的使用情况。

5 大数据安全管理基本原则

5.1 职责明确

组织应明确不同角色和其大数据活动的安全责任。组织应：

a) 设立大数据安全管理者。根据组织使命、数据规模与价值、组织业务等因素，组织应明确担任大数据安全管理者角色的人员或部门，可由业务负责人、法律法规专家、IT安全专家、数据安全专家组成，为组织的数据及其应用安全负责；
b) 明确角色的安全职责。组织应明确大数据安全管理者，大数据安全执行者，大数据安全审计者，以及数据安全相关的其他角色的安全职责；
c) 明确主要活动的实施主体。组织应明确大数据主要活动的实施主体及安全责任。

5.2 安全合规

组织应制定策略和规程确保数据的各项活动满足合规要求。组织应：

a) 理解并遵从数据安全相关的法律法规、合同、标准等；
b) 正确处理个人信息、重要数据；
c) 实施了合理的跨组织数据保护的策略和实践。

5.3 质量保障

组织在采集和处理数据的过程中应确保数据质量。组织应：

a) 采取适当的措施确保数据的准确性、可用性、完整性和时效性；
b) 建立数据纠错机制；

c） 建立定期检查数据质量的机制。

5.4 数据最小化

组织应保证只采集和处理满足目的所需的最小数据。组织应：

a） 在采集数据前，明确数据的使用目的及所需数据范围。

b） 提供适当的管理和技术措施保证只采集和处理与目的相关的数据项和数据量。

5.5 责任不随数据转移

当前控制数据的组织应对数据负责，当数据转移给其他组织时，责任不随数据转移而转移。组织应：

a） 对数据转移给其他组织所造成的数据安全事件承担安全责任；

b） 在数据转移前进行风险评估，确保数据转移后的风险可承受；

c） 通过合同或其他有效措施，明确界定接收方接收的数据范围和要求，确保其提供同等或更高的数据保护水平，并明确接收方的数据安全责任；

d） 采取有效措施，确保数据转移后的安全事件责任可追溯。

5.6 最小授权

组织应控制大数据活动中的数据访问权限，保证在满足业务需求的基础上最小化权限。组织应：

a） 赋予数据活动主体的最小操作权限和最小数据集；

b） 制定数据访问授权审批流程，对数据活动主体的数据操作权限和范围变更制定申请和审批流程；

c） 及时回收过期的数据访问权限。

5.7 确保安全

组织应采取适当的管理和技术措施，确保数据安全。组织应：

a） 对数据进行分类分级，对不同安全级别的数据实施恰当的安全保护措施；

b） 确保大数据平台及业务的安全控制措施和策略有效，保护数据的完整性、保密性和可用性，确保数据生命周期的安全；

c） 解决风险评估和安全检查中所发现的安全风险和脆弱性，并对安全防护措施不当所造成的安全事件承担责任。

5.8 可审计

组织应实现对大数据平台和业务各环节的数据审计。组织应：

a） 记录大数据活动中各项操作的相关信息，且保证记录不可伪造和篡改；

b） 采取有效技术措施保证对大数据活动的所有操作可追溯。

6 大数据安全需求

6.1 保密性

大数据环境下的保密性需求应考虑以下几个方面：

a） 数据传输的保密性，使用不同的安全协议保障数据采集、分发等操作中的传输保密要求；

b） 数据存储的保密性，例如使用访问控制、加密机制等；

c） 加密数据的运算，例如使用同态加密等算法；

d） 数据汇聚时敏感性保护，例如通过数据隔离等机制确保汇聚大量数据时不暴露敏感信息；
e） 个人信息的保护，例如通过数据匿名化使得个人信息主体无法被识别；
f） 密钥的安全，应建立适合大数据环境的密钥管理系统。

6.2 完整性

大数据环境下的完整性需求应考虑以下方面：
a） 数据来源验证，应确保数据来自于已认证的数据源；
b） 数据传输完整性，应确保大数据活动中的数据传输安全；
c） 数据计算可靠性，应确保只对数据执行了期望的计算；
d） 数据存储完整性，应确保分布式存储的数据及其副本的完整性；
e） 数据可审计，应建立数据的细粒度审计机制。

6.3 可用性

大数据环境下的可用性需求应考虑以下方面：
a） 大数据平台抗攻击能力；
b） 基于大数据的安全分析能力，如安全情报分析、数据驱动的误用检测、安全事件检测等；
c） 大数据平台的容灾能力。

6.4 其他需求

大数据安全除了考虑信息系统的保密性、完整性和可用性，还应该针对大数据的特点组织还应从大数据活动的其他方面分析安全需求，包括但不限于：
a） 与法律法规、国家战略、标准等的合规性；
b） 可能产生的社会和公共安全影响，与文化的包容性；
c） 跨组织之间数据共享；
d） 跨境数据流动；
e） 知识产权保护及数据价值保护。

7 数据分类分级

7.1 数据分类分级原则

数据分类分级应满足以下原则：
a) 科学性。按照数据的多维特征及其相互间逻辑关联进行科学和系统地分类，按照大数据安全需求确定数据的安全等级。
b) 稳定性。应以数据最稳定的特征和属性为依据制定分类和分级方案。
c) 实用性。数据分类要确保每个类下要有数据，不设没有意义的类目，数据类目划分要符合对数据分类的普遍认识。数据分级要确保分级结果能够为数据保护提供有效信息，应提出分级安全要求。
d) 扩展性。数据分类和分级方案在总体上应具有概括性和包容性，能够针对组织各种类型数据开展分类和分级，并满足将来可能出现的数据的分类和分级要求。

7.2 数据分类分级流程

组织应结合自身业务特点，针对采集、存储和处理的数据，制定数据分类分级规范，规范应包含但不限于以下内容：

a） 数据分类方法及指南；

b） 数据分级详细清单，包含每类数据的初始安全级别；

c） 数据分级保护的安全要求。

组织应按照图1的流程对数据进行分类分级。组织应根据数据分类分级规范对数据进行分类；为分类的数据设定初始安全级别；综合分析业务、安全风险、安全措施等因素后，评估初始安全级别是否满足大数据安全需求，对不恰当的数据分级进行调整，并确定数据的最终安全级别。附录A提供运营商对数据分级的实践案例。

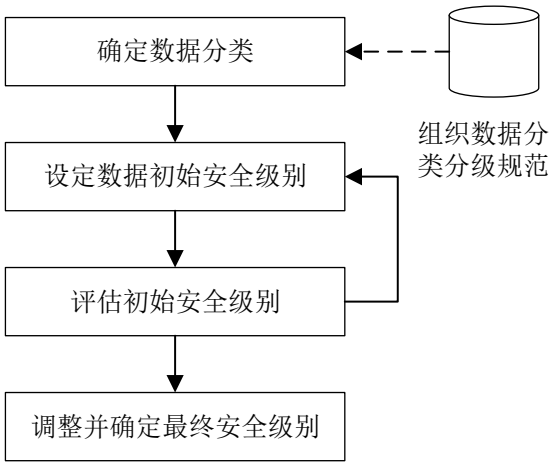

图1 数据分级实施步骤

7.3 数据分类方法

数据分类方法可参照标准 GB/T 7072—2002 中的 6 章实施，由组织根据数据主体、主题、业务等不同的属性进行分类。

7.4 数据分级方法

组织应对已有数据或新采集的数据进行分级，数据分级需要组织的主管领导、业务专家、安全专家等共同确定。政府数据分级参照GB/T 31167-2014中6.3节,将非涉密数据分为公开、敏感数据。个人信息和个人敏感信息参照GB/T 35273—2017中的附录A和附录B执行。

涉密信息的处理、保存、传输、利用按国家保密法规执行。

组织可根据法律法规、业务、组织战略、市场需求等，对敏感数据进一步分级，以提供相适应的安全管理和技术措施。

组织针对不同级别的数据应参照GB/T 30274-2017，选择恰当的管理和技术措施对数据实施有效的安全保护。

8 大数据活动及安全要求

8.1 大数据的主要活动

在数据生命周期中，组织可能参与数据形态的一个或多个阶段，将组织可能对数据实施的操作任务的集合，即活动划分为：数据采集、数据存储、数据处理、数据分发以及数据删除等：

a) 数据采集。指数据进入组织的大数据环境，数据可来源于其他组织或自身产生。
b) 数据存储。将数据持久存储在存储介质上。
c) 数据处理。通过该活动履行组织的职责或实现组织的目标。处理的数据可以是组织内部持久保存的数据，也可以是直接接入分析平台的实时数据流。
d) 数据分发。组织在满足相关规定的情况下将数据处理生成的报告、分析结果等分发给公众或其他组织，或将组织内部的数据适当处理后进行交换或交易等。
e) 数据删除。当组织决定不再使用特定数据时，组织可以删除该数据。

活动和活动之间可能存在数据流，组织应分析各活动中的安全风险，确保安全要求、策略和规程的实施。

8.2 数据采集

8.2.1 数据采集活动的概念

数据采集活动的目标是获得数据，数据采集方式包括但不限于：

a) 网络数据采集。通过网络爬虫或公开API等方式获取数据。
b) 从其他组织获取。通过线上或线下等方式从组织外获取数据。
c) 通过传感器获取。传感器包括温度传感器、电视、汽车、摄像头等公共和个人的智能设备。
d) 系统数据。组织内部的系统在运行过程中采集和产生的业务数据，以及各种系统、程序和服务运行产生的大量运维和日志数据等。

数据采集活动主要操作包括但不限于：发现数据源、传输数据、生成数据、缓存数据、创建元数据、数据转换、数据完整性验证等。

8.2.2 安全要求

组织开展数据采集活动时，应：

a) 定义采集数据的目的和用途，明确数据采集源和采集数据范围；
b) 遵循合规原则，确保数据采集的合法性、正当性和必要性；
c) 遵循数据最小化原则，只采集满足业务所需的最少数据；
d) 遵循质量保障原则，制定数据质量保障的策略、规程和要求；
e) 遵循确保安全原则，对采集的数据进行分类分级标识，并对不同类和级别的数据实施相应的安全管理策略和保障措施。对数据采集环境、设施和技术采取必要的安全管控措施。

8.3 数据存储

8.3.1 数据存储活动的概念

数据存储指将数据静态保存在大数据平台，存储的数据包括采集的数据、分析和处理的结果数据等。存储系统可以是关系数据库、非关系数据库等，应支持对不同类型和格式的数据存储，且提供多种数据访问接口，如文件系统接口、数据库接口等。直到数据被彻底删除之前，存储的数据均应由组织提供恰当的安全保护。

组织应充分考虑使用第三方数据存储平台保存数据的安全风险。由于知识产权、法律法规等原因，组织即使能对存储系统中的数据如个人信息或健康数据等进行有效控制，但可能不是数据的拥有者，组织仍需承担数据的存储管理责任。

数据存储活动的主要操作包括但不限于：数据编解码、数据加解密、冷热数据分级存储、数据归档持久存储、数据备份、数据更新、数据访问等。

8.3.2 安全要求

组织开展数据存储活动时，应：

a) 将不同类别和级别的数据分开存储，并采取物理或逻辑隔离机制；

b) 遵守确保安全原则，主要考虑以下几个方面：
   1) 存储架构安全；
   2) 逻辑存储安全；
   3) 存储访问控制；
   4) 数据副本安全；
   5) 数据归档安全；
   6) 数据时效性管理。
c) 建立数据存储冗余策略和管理制度，及数据备份与恢复操作过程规范。

8.4 数据处理

8.4.1 数据处理活动的概念

数据处理活动指通过数据分析和数据可视化等技术从数据中提取信息，提炼出有用知识和价值的系列操作。

数据处理活动的主要操作包括但不限于：数据查询、数据读取、数据索引、批处理、交互式处理、流处理、数据统计分析、数据预测分析、数据关联分析、数据可视化、生成分析报告等。

8.4.2 安全要求

组织开展数据处理活动时，应：

a) 依据个人信息和重要数据保护的法律法规要求，明确数据处理的目的和范围；
b) 建立数据处理的内部责任制度，保证分析处理和使用数据不超出声明的数据使用目的和范围；
c) 遵循最小授权原则，提供数据细粒度访问控制机制；
d) 遵循确保安全原则，主要考虑以下几个方面：
   1) 分布式处理安全；
   2) 数据分析安全；
   3) 数据加密处理；
   4) 数据脱敏处理；
   5) 数据溯源。
e) 遵循可审计原则，记录和管理数据处理活动中的操作；
f) 对数据处理结果进行风险评估，避免处理结果中包含可恢复的敏感数据。

8.5 数据分发

8.5.1 数据分发活动的概念

数据分发活动指将原始数据、处理的数据等不同形式的数据传递给组织内部其他角色、外部实体或公众等。数据分发包括线上或线下等多种方式。

数据分发的原因包括但不限于：

a) 组织内部部门间的数据交换；
b) 为外部生成报告，例如政府部门的统计数据；
c) 企业间的数据交换，为客户提供使用报告等；
d) 数据出售给其他组织；
e) 业务实现需求。

数据分发涉及的主要操作包括但不限于：数据传输、数据导出、数据交换、数据交易、数据共享等。

8.5.2 安全要求

组织开展数据分发活动时，应：

a) 遵循责任不随数据转移原则；

b) 个人信息、重要数据等有处境需求时，应根据相关法律法规、政策文件盒标准执行处境安全评估；
c) 在数据分发前，对数据进行风险评估，确保数据分发后的风险可承受，并通过合同明确数据接收方的数据保护责任；
d) 在数据分发前，对数据的敏感性进行评估，根据评估结果对需要分发的敏感信息进行脱敏操作；
e) 遵循可审计原则，记录时间、分发数据、数据接收方等相关信息；
f) 评估数据分发中的传输安全风险，确保数据传输安全；
g) 提供有效的数据安全共享机制；
h) 建立数据发布的审核制度，严格审核发布信息符合相关法律法规要求。明确数据发布的内容和范围。对发布的数据开展定期审核。

8.6 数据删除

8.6.1 数据删除活动的概念

数据删除活动指组织删除自有或租用的大数据平台上的数据及其副本。如果数据来自外部实时数据流，还应断开与实时数据流的链接。

数据被删除的原因包括但不限于：

a) 为了减少数据泄露的风险。避免数据被不适当的分发或处理；
b) 删除不相关或不正确的数据。数据与最初使用目的不再相关，或数据不正确；
c) 业务完成后的数据删除处理。数据业务完成服务目标，不再需要保存相关数据；
d) 满足客户的数据删除要求。法律法规要求保留的数据除外。

数据删除活动的主要操作包括但不限于：删除元数据、删除原始数据及其副本、断开与外部实时数据流的链接、删除数据的访问接口、不可恢复的数据销毁等。

8.6.2 安全要求

组织开展数据删除活动时，应：

a) 删除超出数据留存期限的相关数据，对留存期限有明确规定的，按相关规定执行。
b) 依照数据分类分级建立相应的数据删除机制，明确需要进行数据销毁的数据、方式和要求，明确销毁数据范围和流程；
c) 遵守可审计原则，建立数据删除策略和管理制度，记录数据删除的操作时间、操作人、操作方式、数据内容等相关信息。

9 评估大数据安全风险

9.1 概述

组织应参照GB/T 20984-2007开展风险评估工作，并关注大数据环境下安全风险评估的特点。

9.2 资产识别

组织开展资产识别时，应关注大数据的资产特点，包括但不限于：

a) 个人信息；
b) 重要数据；
c) 大数据可视化算法与软件；
d) 大数据分析算法与软件；
e) 大数据处理框架，如流处理框架、交互式处理框架、离线处理框架；

f） 大数据存储框架，如分布式文件系统、非关系型数据库等；
g） 大数据平台计算资源（如CPU、内存、网络等）管理框架等。

9.3 威胁识别

组织开展威胁识别时，应关注大数据环境下的威胁特点，包括但不限于：
a） 潜在的不利因素
- 潜在攻击方具有的资源、技术能力、动机等。常见的攻击方有个人、组织、国家等；
- 潜在攻击方窃取、利用和滥用数据的意图；
- 大数据访问、存储和处理所需资源；
- 直接访问数据或窃取数据的风险；
- 发起攻击、恶意利用大数据的成本与收益。

b） 恶意利用所需的科学专业知识和技能
- 数据和结果分析需要使用的技能、专业知识；
- 数据使用和结果分析需要的技术和设备；
- 利用系统脆弱性需要的技能、技术和知识。

c） 数据出境威胁

9.4 脆弱性识别

组织开展脆弱性识别时，应关注大数据环境下的脆弱性特定，包括但不限于：
a） 大数据存储、处理等基础软件和基础设施的脆弱性；
b） 大数据相关系统的脆弱性。

9.5 已有安全措施确认

组织应对已采取的安全措施的有效性进行确认。安全措施的选择可以参考GB/T 35274—2017。

9.6 风险分析

组织应采用适当的方法与工具确定威胁利用脆弱性导致大数据安全事件发生的可能性，综合安全事件所作用的大数据资产价值及脆弱性的严重程度，判定安全事件造成的损失对国家安全、社会公共利益、组织和个人利益的影响。

附 录 A
（资料性附录）

电信行业数据分类分级示例

A.1 数据分类

依据支撑电信业务的业务支撑域系统（B域）、网络支撑域系统（O域）、管理信息域系统（M域）、信令/DPI数据系统、业务管理平台等五大领域的数据，电信大数据分为以下四类。

类别	子类及范围
（A类)用户身份相关数据	**（A1）用户身份和标识信息：**（A1-1）自然人身份标识、（A1-2）网络身份标识、（A1-3）用户基本资料、（A1-4）实体身份证明、（A1-5）用户私密资料 **（A2）用户网络身份鉴权信息：**（A2-1）密码及关联信息
（B类)用户服务内容数据	**（B1）服务内容和资料数据：**（B1-1）服务内容数据、（B1-2）联系人信息
（C类)用户服务衍生数据	**（C1）用户服务使用数据：**（C1-1）业务订购关系、（C1-2）服务记录和日志、（C1-3）消费信息和账单、（C1-4）位置数据、（C1-5）违规记录数据 **（C2）设备信息：**（C2-1）设备标识、（C2-2）设备资料
（D类)企业运营管理数据（企业运营管理数据依据其商业价值，分为 “核心”、“重要”、“一般”、“公开”四类数据。）	**（D1）企业管理数据：**(D1-1)：企业内部核心管理数据、(D1-2):企业内部重要管理数据、(D1-3)：企业内部一般管理数据、(D1-4)：市场核心经营类数据、(D1-5)：市场重要经营类数据、(D1-6)：市场一般经营类数据、(D1-7)：企业公开披露信息、（D1-8）：企业上报信息 **（D2）业务运营数据：**(D2-1)：重要业务运营服务数据、(D2-2)：一般业务运营服务数据、(D2-3)：业务运营服务数据、(D2-4)：数字内容业务运营数据 **（D3）网络运维数据：**(D3-1)：网络设备及IT系统密码及关联信息、（D3-2)：核心网络设备及IT系统资源数据、(D3-3)：重要网络设备及IT系统资源数据、(D3-4)：一般网络设备及IT系统资源数据、(D3-5)：公开网络设备及IT系统资源数据、(D3-6)：公开网络设备及IT系统支撑据 **（D4）合作伙伴数据：**(D4-1)：渠道基础数据、(D4-2)：CP/SP基础数据*

A.2 数据分级

依据个人信息保护需求和电信业务运行需要，电信大数据分级如下表所示：

类别	定位	子类及范围

第4级	极敏感级	（A1-4）实体身份证明、（A1-5）用户私密资料、（A2-1）用户密码及关联信息、(D1-1)企业内部核心管理数据、(D1-4)市场核心经营类数据、(D3-1)网络设备及IT系统密码及关联信息、（D3-2）核心网络设备及IT系统资源类数据
第3级	敏感级	（A1-1）自然人身份标识、（A1-2）网络身份标识、（A1-3）用户基本资料、（B1-1）服务内容数据、（B1-2）联系人信息、（C1-2）服务记录和日志、（C1-4）位置数据、(D1-2)企业内部重要管理数据、（D1-5）市场重要经营类数据、（D1-8）企业上报信息、（D2-1）重要业务运营服务数据、（D3-2）重要网络设备及IT系统资源类数据、（D4-1）渠道基础数据、（D4-2）CP/SP基础数据
第2级	较敏感级	（C1-3）消费信息和账单、（C2-1）终端设备标识、（C2-2）终端设备资料、（D1-3）企业内部一般管理数据、(D1-6)市场一般经营类数据、（D2-2）一般业务运营服务数据、(D3-3)一般网络设备及IT系统资源类数据、(D3-6)：网络设备及IT支撑数据
第1级	低敏感级	（C1-1）业务订购关系、（C1-5）违规记录数据、（D1-7）企业公开披露信息、(D2-3)业务运营服务数据、（D2-4）数字内容业务运营数据、(D3-5)：公开网络设备及IT系统资源类数据

附 录 B
（资料性附录）

生命科学大数据风险分析示例

表 B-1 展示了以下 3 个场景下生命科学大数据风险评估案例：

- 场景 1：使用生物大数据来设计针对特定人群的病毒
- 场景 2：误导传染信息、传染病监视系统
- 场景 3：利用大数据技术破坏现有的病原体检测能力

表 B-1 3 种典型场景下生命科学大数据风险分析

应用场景			场景 1	场景 2	场景 3
风险分析项	攻击方能力要求		具有充足资源的组织或国家、内部人员。攻击方技术先进、具有大量攻击所需资源、数据和所需软件。	熟悉计算机的个人、组织等。	技术上先进、能访问所需软件和数据、具有所需的资金支持。
	利用数据仓库、软件、网络基础设施的脆弱性的能力要求		能使用公开访问的数据和分析软件。	报告机制和数据库的开放访问	开放访问数据和一些分析软件
	专业知识和技能要求	利用系统的脆弱性	无	报告系统的访问，无特殊的技能要求	无
		使用大数据分析来设计有害的生物代理	需要特殊的技能：微生物基因组学、分子生物学、生物信息学	无	需要特殊技能：生物信息、分子生物学、微生物基因
后果	经济、政治体系、社会、健康、环境和农业方面产生严重后果		后果很严重，造成人员伤亡或引发疾病、受到当地或国际社区的攻击、合规问题。	后果为中到高：损害人员健康、农业、环境等。	后果严重：不能检测危险的病原体感染。
	是否具有足够的应对措施		无足够应对措施。可能的技术措施有：访问控制、组织的内部措施、个人的措施。	无。IP 地址跟踪。	无。IP 地址跟踪、访问控制、组织的内部措施。
风险发生的概率			低	中	高

附 录 C
（资料性附录）

大数据安全风险

1. 大数据恶意使用给个人信息保护或国家安全带来损害

由于缺乏风险评估所需的必要信息，评估类似大数据等新兴技术的风险比较困难。随着大数据技术的进步、采集信息的不断丰富、数据共享标准的制定，大数据分析可以发现更多、更深入的关联关系。

例如通过关联分析用户在社交网站中写入的信息、智能手机显示的位置信息等多种数据，可以识别到自然人，挖掘出个人信息。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，给该群体的生命安全产生重大威胁。

2. 数据交易增加数据管理难度

由于大数据交换和交易具有便捷、快速、隐蔽的特性，监管大数据在不同数据控制者的处理过程非常困难。当发生数据泄露或个人数据保护问题时难以定位事件源，即责任方难以追溯，是其中最大的风险之一。

3. 数据不准确给机构的利益带来损失

网络的数据并非都可信，这主要反映在伪造的数据和失真的数据两个方面。有人可能通过伪造数据来制造假象，进而对数据分析人员进行诱导；或者数据在传播中逐步失真。这可让大数据分析和预测得出无意义或错误的结果，给组织、国家带来重大损失。

4. 大数据增加访问控制实现难度

访问控制是实现数据受控共享的有效手段，由于大数据可能被用于多种不同场景，其访问控制需求十分突出。难以预设角色，实现角色划分.由于大数据应用范围广泛，它通常要为来自不同组织或部门、不同身份与目的的用户所访问，实施访问控制是基本需求。然而，在大数据的场景下，有大量的用户需要实施权限管理，且用户具体的权限要求未知。面对未知的大量数据和用户，预先设置角色十分困难。

同时，难以预知每个角色的实际权限。面对大数据，安全管理员可能无法准确为用户指定其可以访问的数据范围，而且这样做效率不高。

5. 数据聚集增加遭受网络攻击风险

为了从数据中挖掘有价值的信息，需要将不同的数据源进行汇聚和关联分析。数据汇聚增加了遭受网络攻击的风险。大数据系统本身是一个复杂系统，使得大数据系统不可避免存在一些安全脆弱点。成功的网络攻击导致数据被窃取、破坏后造成更加严重的损失。

6. 数据共享的安全风险

很难预先知道安全分享数据，才能既保证敏感信息不被泄漏，又保证数据的正常使用。真实数据不是静态的，并且随着时间的变化而变化。数据规模在不断增加、分析技术不断发展，很难准确评估数据共享的风险。因此很难对数据进行充分的访问控制，存在敏感信息泄露的风险。

参 考 文 献

[1] NIST Special Publication 1500-1, NIST Big Data Interoperability Framework: Volume 1, Definitions Final Version 1, September 2015

[2] NIST Special Publication 1500-2, NIST Big Data Interoperability Framework: Volume 2, Big Data Taxonomies Final Version 1, September 2015

[3] NIST Special Publication 1500-4, NIST Big Data Interoperability Framework: Volume 4, Security and Privacy Requirements Final Version 1, September 2015

[4] NIST Special Publication 1500-6, NIST Big Data Interoperability Framework: Volume 6, Reference Architecture Final Version 1, September 2015

[5] Editor draft of ISO/IEC 20547-1, Big data - Reference architecture - Part 1: Framework and application process, May 2, 2016

[6] 4th Working Draft of ISO/IEC 20547-3, Big data - Reference architecture - Part 3: Reference architecture, January 23, 2016

[7] 2nd Editors Draft of ISO/IEC 20547-4, Big data - Reference architecture - Part 4: Security and Privacy Fabric, January 5, 2016

[8] Committee Draft of ISO/IEC 20546, Big data – Overview and Vocabulary, March, 2016

[9] ITU-T Y.3600, Big data – Cloud Computing based requirements and capabilities, November, 2015

[10] ENISA Big Data Security – Good Practices and recommendations on the security of Big data systems, December, 2015

[11] ICO Big data and data protection Version 1.0, July 28, 2014

[12] GB/T 20529.1——2006 企业信息分类编码导则 第一部分：原则与方法

[13] 贵州省《政府数据 数据分类分级指南》（试行）

[14] ISO/IEC DIS 38505-1, Information Technology — Governance of IT — Part 1: The application of ISO/IEC 38500 to the governance of data

[15] A Joint AAAS-FBI-UNICRI Project, National and Transnational Security Implications of Big Data in the Life Sciences

[16] GB/T 19715.1-2005/ISO/IEC TR13335-1:1996. 信息技术 信息技术安全管理指南 第1部分：信息技术安全概念和模型.

[17] GB/T 19715.2-2005/ISO/IEC TR13335-2:1997. 信息技术 信息技术安全管理指南 第2部分：管理和规划信息技术安全.

[18]GB/T 35273—2017 信息安全技术 个人信息安全规范

[19]GB/T 20984—2007 信息安全技术 信息安全风险评估规范