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

（征求意见稿）

（本稿完成日期：2017 年 3 月 21 日）

XXXX－XX－XX 发布　　XXXX－XX－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

本标准按照 GB/T 1.1-2009 给出的规则起草。

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、中国移动有限公司、深圳市腾讯计算机系统有限公司、清华大学、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、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、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：xxx。

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、商业的业务策略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，从数据收集的角度，数据收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。随着大数据的应用，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。随着大数据的应用和分析，数据价值不断提升，安全受到高度重视。

而拥有大量数据的企业的管理和技术水平参差不齐，有不少企业缺乏技术、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。数据信息泄露的主要途径包括：一是外部攻击者利用系统和平台的漏洞入侵获取数据；二是掌握数据的机构或其合作商内部人员主动泄露数据；三是掌握数据的机构或其合作商内部人员与外部攻击者勾结盗取数据。从数据泄露的途径来看，关键是要加强掌握数据的机构和其合作商的技术和管理能力的建设，加强数据收集、存储、使用、分发等环节的技术和管理措施，制定规范和制度。

本标准指导拥有、处理大数据的政府部门、企业、事业单位、非盈利机构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

本标准为组织的大数据安全管理提供指导，本标准提出大数据安全管理基本原则、大数据安全管理基本概念和大数据安全风险管理过程。本标准提出大数据的数据收集、数据存储、数据使用、数据分发、数据删除等主要阶段的基本概念和管理要求。本标准规范了组织内部不同大数据角色的安全职责。

本标准适用于所有的组织，包括企业、政府部门、非盈利机构等。

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，仅所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 25069——2010 信息安全技术 术语

GB/T 7072——2002 信息分类和编码的基本原则与方法

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

GB/T 31167——2014 信息安全技术 云计算服务安全指南

GB/T xxxxx——xxxx 信息安全技术 大数据服务安全能力要求

GB/T xxxxx——xxxx 信息安全技术 个人信息安全规范

3 术语、定义和缩略语

GB/T 32400-2015中界定的以及下列术语和定义适用于本文件。

3.1 术语和定义

（引用国内在研的大数据标准，根据本标准需求添加所需术语）

3.1.1

大数据 big data

具有数量巨大、种类多样、流动速度快、且特征多变等特性，并且难以用传统数据体系结构进行有效处理的数据集。

注：国际上，大数据的4个特征普遍不加修饰地直接用volume、 variety、 velocity和variability予以表述，并分别赋予了它们在大数据语境下的定义:

a)数量 volume：数据集的规模。

b)多样性 variety：数据来源、领域或类型多样。

c)速度 velocity：单位时间的数据流量。

d)多变性 variability：大数据其他特征，即数量、速度和多样性等特征都处于多变状态。

3.1.2

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，包括监视大数据的存储、输入输出操作控制等大数据服务功能的软硬件集合。

3.1.3

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。

[GB/T 20984-2007]

3.2 缩略语

下列缩略语适用于本文件。

4 大数据安全管理原则

4.1 原则 1 - 职责明确原则

a) 根据数据规模、数据重要性、组织规模等因素，组织可成立安全管理团队，安全管理团队为组织数据及使用安全负责。
b) 组织应明确组织内部不同角色的数据安全管理职责。
c) 组织应明确大数据生命周期各活动的实施主体及安全责任。

4.2 原则 2 - 意图合规原则

对数据的收集、使用需基于法律依据。组织应制定相关流程确保数据的收集和使用方式没有违反任何法律义务，包括法律法规、合同条款等。组织需要确保履行需要承担的内部和外部的责任，包括但不限于：

a) 确保所有数据集和数据流的安全；
b) 正确处理个人信息、重要信息；
c) 实施了合理的跨组织数据保留的策略和实践；
d) 理解数据相关的法律义务，并确保组织履行了这些义务。

4.3 原则 3 - 质量保障原则

组织应：

a) 实施适当的措施确保数据的准确性、相关性、完整性和时效性。
b) 建立控制机制定期检查收集和存储的数据的质量。

4.4 原则 4 - 数据最小化原则

组织应采取适当的措施最小化大数据生命周期各活动涉及的数据。

4.5 原则 5 - 责任不随数据转移原则

a) 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。

b) 组织在数据转移前，需对数据进行风险评估，确保数据转移后的风险可承受，方可转移数据。并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。
c) 组织在数据转移前，需确保通过合同或其他诸如强制的内部策略等明确界定了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。

4.6 原则 6 - 最小授权原则

a) 在保证组织业务功能完整实现的基础上应赋予数据活动中各角色最小的操作权限，确保非法用户或异常操作所造成的损失最小。
b) 所有角色只能使用所授权范围内的数据，非授权范围内的数据使用必须进行授权审批。

4.7 原则 7 – 数据保护原则

a) 组织需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) 组织应确保处理大数据处理平台及应用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，确保数据在整个生命周期里，免遭诸如未授权访问、破坏、篡改、泄露或丢失等风险。
c) 组织应解决风险评估和安全检查中所发现的风险和脆弱性，并对数据安全防护措施不当所造成的安全事件承担责任。

4.8 原则 8 – 可审计原则

对数据进行修改、查询、导出、删除等操作时，组织需要记录相应的操作，记录应可追溯可审查。

5 大数据安全管理基本概念

5.1 概述

大数据面临一些新的安全风险，如数据关联分析可能暴露个人隐私、大数据的新特性使得传统的安全机制不能满足安全要求、大数据恶意使用可能会给公共利益、国家安全等带来严重损害等。

大数据安全管理用来实现和维护数据保密性、完整性、可用性、可核查性、真实性和可靠性的过程。大数据安全管理与IT安全管理相似，主要功能包括：

a) 确定组织的数据安全目标、战略和策略；
b) 确定组织的数据安全要求；
c) 识别并分析对组织数据的安全威胁；
d) 识别并分析组织数据安全风险；
e) 规定合适的防护措施；
f) 监督防护措施的实施与运行；
g) 检测并及时响应数据安全事件。

5.2 大数据安全管理方法

组织应系统分析和识别大数据安全要求，并采用系统的安全技术和运行管理措施保护大数据安全。大数据安全管理应包括以下活动：

a) 制定大数据安全策略；
b) 明确组织中大数据安全管理的角色和责任；
c) 风险管理；

d) 配置管理；
e) 变更管理；
f) 应急响应和灾难恢复；
g) 安全意识培训；
h) 大数据平台运行安全管理。

6 制定大数据安全目标、战略和策略

组织应明确拟使用大数据达成的目标，如利用大数据提升组织的竞争力等；应确定实现目标采用的战略；并制定计划实施的策略。

组织宜采用分层目标、战略和策略结构。总体目标指组织根据自身的业务特点建立大数据收集、使用的整体目标、战略和策略。基于总体的大数据目标、战略和策略，组织制定总体的大数据安全目标、战略和策略。各个部门应该根据自身特点，制定适用于部门具体安全需求的数据安全目标、战略和策略。如图1所示。

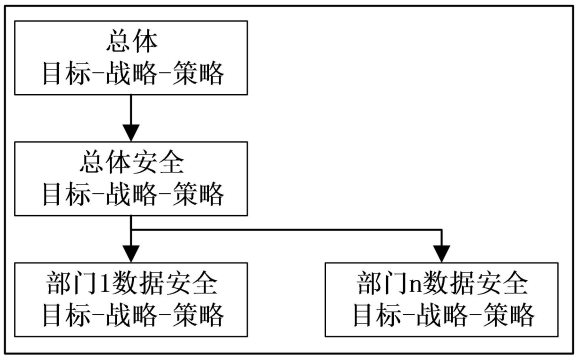

图1 目标、战略和策略层次结构

数据安全目标、战略和策略可以使用自然语言阐述，也可以使用机器语言表示，应该包括以下几个方面：

a) 保密性；
b) 完整性；
c) 可用性；
d) 可核查性；
e) 真实性；
f) 可靠性。

目标、战略和策略要规定组织的数据安全等级、接受风险的阈值和应急需求。

7 明确大数据安全管理角色与责任

7.1 概述

大数据安全管理角色包括组织内部的安全管理团队和职能部门。安全管理团队对组织的大数据安全全面负责。职能部门是根据业务需求对数据进行收集、分析或使用的具体部门，负责数据收集、分析或使用等的技术实现。职能部门对本部门收集或使用的数据安全负责，细化数据在收集、分析或使用等阶段的安全要求，并推动落实。

7.2 数据安全管理团队的职责

安全管理团队的具体职责有：

a) 应确定各种数据的分类分级初始值，制定数据分类分级指南；
b) 应综合考虑相关的法律法规、政策、标准、大数据分析技术当前水平、组织所处行业特殊性等，综合评估数据安全分析，制定数据安全基本要求；
c) 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；
d) 负责组织的大数据安全管理过程，并对外部相关方（如：国家安全的主管部门、数据主体等）负责；
e) 对于组织的数据使用，大数据安全管理团队具有相应的权力、职责和管理责任。

7.3 职能部门的职责

职能部门在履行其职能时会生成、收集不同数据，持久保存数据并进行分析。职能部门可能涉及一个或多个大数据主要阶段，根据涉及的阶段履行相应安全职责。

职能部门需要配合安全管理团队来保障数据安全，职能部门的主要职责有：

a) 确定本部门数据的最终分级；
b) 根据本部门涉及的大数据主要阶段，明确和细化本部门数据在收集、存储、使用等过程中的具体安全要求，并有效实施；
c) 配合安全管理团队处置安全事件；
d) 根据要求安全使用数据。

7.4 明确大数据主要活动实施部门安全管理责任

7.4.1 大数据主要活动基本概念

从数据进入组织的大数据平台开始，以数据被删除结束，数据主要包含数据收集、数据存储、数据使用、数据分发以及数据删除几个活动，即：

a) 数据收集：该阶段使数据进入组织的大数据生态环境，比如保存在大数据平台。
b) 数据存储：该阶段指将数据持续存储在存储介质上。
c) 数据使用：组织通过该阶段的活动作出基于数据的决策，以便更好地履行组织的职责或实现组织的目标。使用的数据可以是组织内部持久保存的数据，也可以是以数据流方式接入分析平台的实时数据流。
d) 数据分发：组织可以在满足相关规定的情况下将数据使用活动中生成的报告、分析结果等分发给其他组织，也可以将组织内部的数据适当处理后销售给其他组织。
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。

不同活动之间可能存在数据流，从而存在一定的安全风险，组织需要确保安全策略、规程和安全要求到位，满足大数据的安全保护要求。

7.4.1.1 收集

收集活动包括数据获取或创建过程，数据收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) 从其他组织获取数据。通过线上或线下等方式获得数据。
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。组织内部的系统运行过程中产生的业务数据，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。

收集活动的主要操作包括：发现数据源、收集数据、生成数据、缓存数据、创建元数据、数据转换、数据验证、数据清理、数据聚合等。

7.4.1.2 存储

数据存储指将数据持久保存在大数据平台，存储的数据包括采集的数据、分析结果数据等。存储系统可以是关系数据库、非关系数据库等，应支持对不同数据类型和数据格式的数据存储，且提供多种数据访问接口，如文件系统接口、数据库接口等。直到数据被删除之前，存储的数据均可被组织合规使用。

在某些情况下，组织将使用第三方的数据存储平台保存数据，此时组织失去对存储基础设施的部分控制权，组织应充分考虑这种方式存在的安全风险。

组织即使能对存储系统中的数据进行有效控制，但可能并不是数据的拥有者。组织不是数据控制者的原因有：知识产权、法律问题（如个人信息或健康数据处理相关法律）等。这种情况下，组织仍需承担数据的管理责任。

存储活动的主要操作包括：数据编解码、数据加解密、数据持久存储、数据备份、数据更新、数据访问等。

7.4.1.3 使用

数据使用活动包括利用数据预处理、数据分析和数据可视化等技术从原始数据中提取信息，提炼出有用知识，支撑组织根据数据作出合理的决策等操作。

数据预处理指对收集的数据进行提取、转换和去噪等处理。原始数据格式和数据类型众多，通过数据提取和转换操作，复杂的数据可以转换为简单的结构化数据，便于对数据进行分析。数据去噪过程删除噪声数据，避免对分析过程产生不利影响。数据分析指从大数据中抽取有用信息或发现有价值的模式的过程。可视化指通过使用统计图、统计表、报告等多种方式展示数据的过程。数据可视化帮助组织更好地理解数据。

使用活动的主要操作包括：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。

7.4.1.4 分发

数据分发活动将原始数据、处理过的数据、分析的结果等不同形式的数据传递给外部实体或组织内部的其他部门。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) 需要为外部生成报告，例如政府机关；

c) 企业与企业间的数据交换、客户需要使用报告等；

d) 数据被出售给一个广告代理或调查公司；

e) 数据是该组织发布的业务的一部分，例如业务数据；

数据分发涉及的主要操作包括：数据传输、数据脱敏、数据交换、数据交易、数据共享。

7.4.1.5 删除

删除活动指删除组织的大数据平台或租用的第三方大数据存储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。

数据需要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使用。

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) 满足客户要求删除其数据的要求。但可能存在法律法规需要保留数据，如和健康相关的数据。

数据删除活动的主要操作包括：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。

7.4.2 大数据主要活动职责

7.4.2.1 数据收集

实施部门应：

a) 定义采集数据目的和用途，明确数据采集源和采集数据范围。
b) 遵循意图合规原则，确保数据收集的合法性、正当性和必要性，且只采集满足业务所需的最小数据集。
c) 遵守质量保障原则，制定数据质量保障的策略、规程和要求。
d) 对数据采集环境、采集设施和采集技术采取必要的安全管控措施。
e) 遵循数据保护原则，对收集数据进行分类分级标识，并对不同类别和级别的数据实施相应的安全管理策略和保障措施。

7.4.2.2 数据存储

实施部门应：

a) 首先对存储的数据进行分类分级，非涉密数据根据本标准的分级要求进行分级。
    1) 数据应先分类再分级；
    2) 不同类别的数据应分开存储，并采取物理或逻辑隔离机制；
    3) 组织根据自身需求，可对组织数据进行内部分类和分级，例如将敏感数据进一步划分为一般敏感和重要敏感数据。
b) 遵守数据保护原则，按照标准 GB/T XXXXX-XXXX 中 6.3 的要求，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。

7.4.2.3 数据使用

实施部门应：

a) 依据国家个人信息和重要数据保护的法律法规要求建立数据使用正当性原则，明确数据使用和分析处理的目的和范围。
b) 建立数据使用的内部责任制度，保证在数据使用声明的目的和范围内对受保护的数据进行使用和分析处理。
c) 遵守最小授权原则，提供细粒度访问控制机制，限定数据使用过程中可访问的数据范围和使用目的。
d) 遵守数据保护原则，按照标准 GB/T XXXXX-XXXX 中 6.4 的要求，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) 遵守可审计原则，记录和管理数据使用操作。
f) 对数据分析结果的风险进行合规性评估，避免分析结果输出中包含可恢复的敏感数据。

7.4.2.4 数据分发

实施部门应：

a) 遵守责任不随数据转移原则，对数据分发后产生的数据安全事件承担必要的安全责任。

b) 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，方可分发数据，并通过合同明确数据接收方的数据保护责任。
c) 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。
d) 遵守可审计原则，记录时间、分发需求、数据接收方等相关信息。
e) 提供有效的数据共享访问控制机制，明确不同机构或部门、不同身份与目的的用户的权限，保证访问控制的有效性。
f) 建立大数据公开的审核制度，严格审核发布信息符合相关法律法规要求。明确数据公开内容、权限和适用范围，信息发布者与使用者的权利与义务。定期审查公开发布的信息中是否含有非公开信息，一经发现，立即删除。
g) 评估数据传输安全风险，明确数据传输安全要求。

7.4.2.5 数据删除

实施部门应：

a) 立即删除超出收集阶段明确的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) 在删除数据可能会影响执法机构调查取证时，采取适当的存储和屏蔽措施。
c) 依照数据分类分级建立相应的数据销毁机制，明确销毁方式和销毁要求。
d) 遵守审计原则，建立数据销毁策略和管理制度，明确销毁数据范围和流程，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。

8 管理大数据安全风险

8.1 概述

风险管理主要包括以下四种不同的活动：

a) 在总体安全策略环境内确定适合于组织的大数据风险管理战略；
b) 根据风险评估结果，选用适当的防护措施；
c) 形成安全策略，必要时更新总体安全策略；
d) 根据批准的安全策略，制订安全计划以实现保护措施。

8.2 评估大数据风险

大数据风险评估可关注以下内容：

a) 安全事件发生的概率
    1) 实施不利行为的因素
        – 潜在攻击方具有的资源、科学与技术专长、访问设施与设备的能力、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、存储和分析所需资源；
        – 直接访问数据或窃取数据的概率；
        – 发起攻击、利用大数据技术、基础设施和数据集的经济能力；
        – 攻击的成本与收益。
    2) 系统的脆弱点
        – 大数据存储、处理等基础软件和基础设施的脆弱性；
        – 识别和限制数据访问和使用分析技术的能力；

– 大数据相关系统的脆弱性。

3) 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、技术专长和知识。

b) 后果

1) 事件发生的后果

– 造成的经济损失、名誉损失等；

– 持续时间；

– 影响的规模；

– 事件恢复的时间、代价等。

2) 存在的应对措施

– 应对脆弱性的措施；

– 防止或减轻后果的措施；

– 防止或控制滥用大数据和大数据技术的措施。

附录C给出了生命科学大数据风险评估的具体案例。

8.3 选择安全保护措施

8.3.1 选择措施

根据8.2节的风险评估方法，评估大数据的安全风险，选择使风险降低到可接受水平的保护措施。选择安全保护措施包含以下步骤：

a) 分类分级组织大数据；

b) 选择适当的保护措施。

组织应采用数据分类分级保护的方法对数据先分类，后分级，最后实施分级保护。

8.3.2 数据分类分级

8.3.2.1 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客观存在的逻辑关联进行科学和系统化的分类，按照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目下要有数据，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，能够实现各种类型数据的分类和分级，以及满足将来可能出现的数据类型和安全需求。

8.3.2.2 数据分类方法

数据分类方法宜参照标准 GB/T 7072—2002 中的 6 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。

8.3.2.3 数据分级方法

组织应对已有数据或新收集的数据进行分级，数据分级时需要组织的业务部门领导、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执行,将非涉密数据分为公开、敏感数据。个人数据按照GB/T ******《个人信息安全规范》中的**，识别和确认个人敏感信息。

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

8.3.2.4 数据分级保护要求

组织应按照图2所示的数据分级步骤分级数据。附录A提供运营商对数据分级的实践案例。

涉密信息的处理、保存、传输、利用按国家保密法规执行。

组织应根据搜集、存储和使用的数据范围，结合自身行业特点制定组织的数据分类分级规范，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。

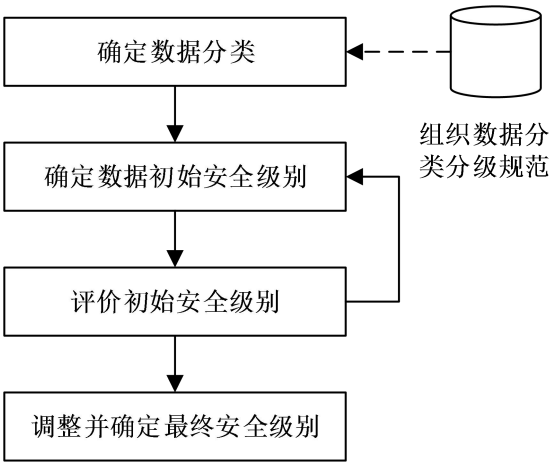

图2 数据分级实施步骤

8.4 制订安全计划

大数据安全计划阐述大数据的安全要求、已采用或拟采用的安全保护措施。大数据安全计划应包括以下内容：

a) 大数据平台安全体系结构和设计；
b) 数据收集、使用的目的、范围、手段等；
c) 数据分类分级及安全要求；
d) 大数据不同阶段的保护措施，大数据支撑平台的保护措施；
e) 数据不同角色的职责分配；
f) 大数据安全相关的安全意识培训等。

9 管理大数据平台运行安全

大数据平台运行安全是大数据安全的基础，大数据平台的运行安全管理目的是确保大数据平台安全持续满足要求。大数据平台运行安全应满足标准GB/T XXXXXX-XXXX中7.4的要求。

组织应制定重大变更管理流程，主要活动包括：

a) 标识重大变更；
b) 重大变更的安全风险评估；
c) 重大变更申请及批准流程；
d) 重大变更中不同角色的职责。

大数据平台的重大变更包括但不限于以下变更：

a) 鉴别（包括身份鉴别和数据源鉴别）和访问控制措施的变更；
b) 数据存储实现方法的变更；
c) 大数据平台中软件代码的更新；
d) 备份机制和流程的变更；
e) 安全措施的替换、撤除；
f) 已部署商业软硬件产品的替换；
g) 数据源的变更，如增加、删除数据源；
h) 分析算法、分析方式的变更。

附录 A 电信行业数据分类分级示例

a) 数据分类

将全公司业务支撑域系统（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，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*

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系统支撑数据
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 国家基础数据

国家基础数据包含以下五类：

- **自然资源和空间地理基础数据**：包括基础地理与区划综合信息、遥感影像综合信息、全国自然资源综合信息、全国遥感资源环境动态监测综合信息、自然灾害监测预警和突发事件应急反应综合信息、资源安全动态评估预警综合信息、可持续发展和地区经济综合信息、生态环境评估综合信息、重大基础设施及生态工程监测综合信息。
- **人口基础数据**：人的状况和基础性信息,主要涉及包括教育、公安、民政、劳动和社会保障、人口和计划生育等信息。
- **法人单位基础数据**：与法人单位密切相关的数据，它代表了法人单位的基本状态与特征，具有跨业务系统共享需求基础。法人单位基本信息包含组织机构代码、组织机构名称、机构类型、机构住所、法定代表人姓名、经营或业务范围、注册资本或开办资金金额、注册资本或开办资金币种、成立日期、注册或登记机构名称、注册或登记号等。
- **宏观经济基础数据**：包括金融、税收、统计等基础信息以及消费、投资、进出口以及经济运行、节能减排、知识产权等方面的业务信息。
- **文化信息基础数据**：以国家物质和非物质文化遗产信息、少数民族传统文化、国家重要文物、国家档案信息等为主要内容的数据。

附录 C 生命科学大数据风险分析示例

表 C-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 C-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
概率	攻击方		具有充分资源的组织或国家、内部人员。这些攻击方技术先进、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。
	数据仓库、软件、网络基础设施的脆弱性		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
	需要专业知识和技能	能利用系统的脆弱性	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			不太可能	近期是可能的	有可能

附录 D 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、收集信息的不断丰富、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 7072——2002 信息分类和编码的基本原则与方法

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

[14] 贵州省《政府数据 数据分类分级指南》（试行）

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

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

[17] 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.

[18] 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.