ICS 35.030
CCS L 80

中华人民共和国国家标准

GB/T 35274—XXXX
代替 GB/T35274-2017

信息安全技术 大数据服务安全能力要求

Information security technology—Security capability requirements for big data services

（征求意见稿）

（本稿完成日期：2022.02.25）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

××××－××－××发布 ××××－××－××实施

国家市场监督管理总局
国家标准化管理委员会

目次

前言

本文件按照GB/T 1.1－2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》给出的规定起草。

本文件代替GB/T 35274—2017《信息安全技术 大数据服务安全能力要求》，与GB/T 35274—2017相比，本文件主要变化如下：

a) 删除了第 3 章中数据生命周期（见 2017 年版的 3.2）、数据服务（见 2017 年版的 3.3）、数据交换（见 2017 年版的 3.11）、数据共享（见 2017 年版的 3.12）和重要数据（见 2017 年版的 3.13）5 个术语和定义，增加了数据处理（见 3.2）、数据安全（见 3.18）、数据保护（见 3.17）、数据收集（见 3.3）、数据存储（见 3.4）、数据使用（见 3.5）、数据加工（见 3.6）、数据传输（见 3.7）、数据提供（见 3.8）、数据公开（见 3.9）和数据销毁（见 3.10）11 个数据处理活动和安全相关术语和定义，修改了大数据平台（见 3.11，2017 年版的 3.6）、大数据应用（见 3.12，2017 年版的 3.5）、大数据系统（见 3.13，2017 年版的 3.9）、大数据使用者（见 3.14，2017 年版的 3.8）、大数据服务（见 3.15，2017 年版的 3.4）、大数据服务提供者（见 3.16，2017 年版的 3.7）、数据供应链（见 3.19，2017 年版的 3.10）7 个术语和定义的描述；

b) 删除了第 4 章中的总体要求（见 2017 年版的 4.1），取消了原标准的能力要求分级（见 2017 年版的 4.2），重写了标准结构文本描述（见 2017 年版的 4.3），以反映本文件内容结构及覆盖的大数据服务安全能力要求项（见第 4 章）；

c) 删除了第 5 章服务规划与管理（见 2017 年版的 5.4）、数据供应链管理（见 2017 年版的 5.5）和合规性管理（见 2017 年版的 5.6），对第 5 章的策略与规程、组织和人员和资产管理安全能力要求依照信息安全管理体系结构进行了完善（见 5.1、5.2、5.3，2017 年版的 5.1、5.3、5.2）；

d) 重组了第 6 章数据服务安全能力要求，按照数据收集、存储、使用、加工、传输、提供、公开和销毁的数据处理过程明确了组织数据处理活动安全能力要求（见第 6 章，2017 年版的第 6 章）；

e) 增加了第 7 章数据服务安全风险管理能力，按照网络安全等级保护制度和数据安全风险评估工作要求，从风险识别、安全防护、安全监测、安全响应和安全恢复环节，规定了大数据服务提供者的数据服务安全风险管理能力（见第 7 章）。

f) 删除了附录 A 中的大数据服务安全业务目标（见 2017 年版的附录 A3），完善了附录 A 中数据处理活动相关的技术术语（见附录 A，2017 年版的附录 A）。

本文件由全国信息安全标准化技术委员会（SAC/TC260）提出并归口。

本文件起草单位：清华大学、北京大学、中国电子技术标准化研究院、中国信息安全测评中心、中国网络安全审查技术与认证中心、国家计算机网络应急技术处理协调中心、深信服科技股份有限公司、浙江蚂蚁小微金融服务集团有限公司、北京快手科技有限公司、阿里巴巴（中国）有限公司、腾讯云计算（北京）有限责任公司、北京奇虎科技有限公司、华控清交信息科技（北京）有限公司、北京天融信网络安全技术有限公司、上海观安信息技术股份有限公司、北京火山引擎科技有限公司、华为技术有限公司、中国科学院信息工程研究所（信息安全国家重点实验室）、启明星辰信息技术集团股份有限公司、中国软件评测中心（工业和信息化部软件与集成电路促进中心）、北京数安行科技有限公司、长扬科技（北京）有限公司、上海赴源科技服务有限公司、杭州世平信息科技有限公司、上海三零卫士信息安全有限公司、北京信安世纪科技股份有限公司、联想（北京）有限公司、杭州安恒信息技术股份有限公司、成都卫士通信息产业股份有限公司、陕西省信息化工程研究院、上海商汤智能科技有限公司、北京神州

绿盟科技有限公司、浙江大华技术股份有限公司、北京腾云天下科技有限公司、北京山石网科信息技术有限公司等单位。

本文件主要起草人：叶晓俊、谢安明等。

本文件及其所代替文件的历次版本发布情况为：

——2017年首次发布为GB/T 35274—2017；

——本次为第一次修订。

信息安全技术　大数据服务安全能力要求

1　范围

本文件规定了大数据服务提供者的数据服务安全能力要求，包括组织管理安全能力、数据处理活动安全能力和数据服务安全风险管理能力。

本文件适用于指导大数据服务提供者的数据服务安全能力建设，也适用于第三方机构对大数据服务提供者的数据服务安全风险进行评估。

2　规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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 5271　信息技术 词汇

GB/T 22080　信息技术 安全技术 信息安全管理体系 要求

GB/T 22239—2019　信息安全技术 网络安全等级保护基本要求

GB/T 25069—2022　信息安全技术 术语

GB/T 35273—2020　信息安全技术 个人信息安全规范

GB/T 35295—2017　信息技术 大数据 术语

3　术语和定义

GB/T 5271、GB/T 25069—2022、GB/T 35273—2020和GB/T 35295—2017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

具有体量巨大、来源多样、生成极快、宜多变等特征，并且难以用传统数据体系结构有效处理的包含大量数据集的数据。

[来源：GB/T 35295—2017，2.1.1]

3.2

数据处理　data handling

数据操作的系统执行，以实现特定目的的过程，例如数据的数学运算或逻辑运算，数据的归并或分类，文本的操作、存储、检索、显示或打印，数据的挖掘分析、数据可视化等在数据收集、存储、使用、加工、传输、提供、公开、销毁等数据处理活动中的各种数据操作。

[来源：GB/T 5271，01.01.06，有修改]

3.3

数据收集　data collection

根据特定的目的和要求，从一种或多种数据源选择和获取数据，并对数据进行变换、转换、纠错、编码等数据清洗操作，形成数据资产进行存储的数据处理活动。

3.4

数据存储 data storage

将数据持久化保存在硬盘等存储介质中的数据处理活动，以便有效的管理、使用、提供、公开或存档数据资产。

3.5

数据使用 data usage

依据数据权属及数据收集目的，控制组织、人员或信息技术系统等对数据资产进行授权和访问的数据处理活动，包括数据使用前条件控制和数据使用后义务履行等数据处理活动。

注：数据使用一般不改变数据本身，如数据读取、数据排序、数据搜索、数据分类、数据可视化等数据操作，使用过程中需对数据的用途、用法、用量及其目的等进行相应的控制。使用前条件是在使用授权规则进行授权过程中,允许主体对客体访问前必须检验的决策因素集。条件控制可用来检查存在的约束限制,使用权限是否有效,哪些约束限制必须更新等。使用后义务履行是主体在获得对客体的访问权限后必须履行的强制需求。主体在获得权限执行数据使用操作后就应有执行获取这些权限的义务责任。

3.6

数据加工 data processing

通过数据变换、数据转换、数据编码、数据计算、数据压缩、数据分析等数据操作，生成新数据（集）的数据处理活动。

注：数据加工一般会涉及数据的改变，需要先读取数据，并经过变换、转换、编码、分析、挖掘、脱敏等数据操作生成新数据（集）。

3.7

数据传输 data transmission

通过信息通信设备将数据从一个网络节点传送到一个或多个网络节点的数据处理活动。

[来源：GB/T 5271，09.01.02，有修改]

注：网络节点可以是计算机、程序、终端设备、存储器、信息系统等。

3.8

数据提供 data provision

向组织内其他责任主体或其他组织提供所控制的数据资产的数据处理活动。

注：数据提供一般指跨安全域的数据交换、数据共享、数据转让等数据操作，跨组织的数据提供过程还可涉及数据的权益归属、数据跨境安全评估以及个人信息保护或隐私影响评估等事项。

3.9

数据公开 data disclosure

向其他组织、个人或指定范围公开组织所控制的数据资产的数据处理活动，使其可合规地获取所公开的数据。

3.10

数据销毁 data destruction

抹去或销毁存储介质中的数据或销毁存储介质的数据处理活动。

注：数据销毁分为数据删除和介质销毁两种。数据删除是指在所涉及的系统及设备中抹去数据或者覆盖存储的数据，使其不可被检索、访问的状态；介质销毁则通过采用物理破坏、化学腐蚀方法直接销毁存储数据的介质，以达到彻底删除数据的目的，确保数据无法复原。

3.11

大数据平台 big data platform

采用分布式存储和计算技术，提供数据处理功能，支持其大数据应用安全高效运行的软硬件集合，包括监视数据输入/输出、数据处理活动控制等软硬件基础设施及其所控制的数据资产。

注：平台指由一组子系统和技术形成的软硬件设施组成，通过一些接口和使用模式提供一组一致的功能，任何由它所支持的应用都可以使用平台的功能而不必关心其实现细节。

3.12

大数据应用 big data application

执行数据收集、存储、使用、加工、传输、提供、公开、销毁等数据处理活动，运行在大数据平台，并提供大数据服务的应用系统。

3.13

大数据系统 big data system

包括大数据平台、大数据应用及其控制数据资产的信息技术系统。

3.14

大数据使用者 big data consumer

使用大数据系统的末端个人、组织、其它信息系统或智能终端设备。

3.15

大数据服务 big data service

利用大数据技术开展数据处理，并通过底层大数据平台和上层大数据应用以服务方式为大数据使用者提供有价值的数据处理功能。

注：大数据服务封装了大数据使用者所需的数据处理活动，且采用标准化协议来发现、注册和发布等相关的服务，以便使大数据服务提供者以信息服务的方式交付价值给大数据使用者。

3.16

大数据服务提供者 big data service provider

拥有大数据系统，提供大数据服务的组织。

注：大数据服务提供者是一种拥有或可获得大数据服务所需数据资产的网络运营者，包括附录A中的大数据平台提供者、大数据应用提供者和大数据服务协调者三种角色的数据处理者。

3.17

数据保护 data protection

实施适当的管理、技术或物理手段，以防止未授权的有意或无意地泄露、修改或破坏数据的活动。

[来源：GB/T 5271，09.06.02，有修改]

3.18

数据安全 data security

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。

注：GB/T 5271定义数据安全是指应用于数据保护的计算机安全[08.01.02]

3.19

数据供应链 data supply chain

对大数据服务提供者的数据收集、存储、加工、传输、提供、公开等数据处理活动中涉及上下游组织的数据源及相关数据操作进行计划、协调、操作、控制和优化所需的可用数据资源形成的链状结构。

注：数据供应链的目标是确保大数据服务提供者能在正确的时间，将大数据服务所需的各种数据资源，通过计划、协调、操作、控制、优化等活动，按照正确的服务协议供应给正确的大数据使用者。

4 概述

大数据服务是针对数量巨大、种类多样、流动速度快、特征多变等特性的数据集，通过底层可伸缩的大数据平台和上层多种多样的大数据应用，为大数据使用者提供价值的一种网络信息服务。大数据服务提供者在开展大数据服务时，需要依照数据安全和个人信息相关的法律、法规的规定，并按照网络安全等级保护制度确保大数据系统安全可靠地运行，满足大数据使用者的数据处理活动服务需求。

大数据服务提供者通过落实网络安全等级保护制度，建立、执行并持续改进信息安全管理体系等方式，建设组织管理安全能力；通过识别数据和供应链安全风险，采取风险应对措施对大数据平台、大数据应用及大数据服务中的数据处理活动进行安全防护，提高数据处理活动安全可控能力；通过对大数据服务进行安全监测，发现其使用的网络产品和服务潜在的缺陷和脆弱性，或者威胁国家安全、危害公共利益等大数据服务风险，及时采取恰当的风险处置措施，必要时启动应急响应机制，依法及时向相关主管部门报告，提高数据服务安全风险管理能力。

因此本文件面向有大数据系统和大数据服务所需数据资源的组织，从组织管理安全能力、数据处理活动安全能力和数据服务风险管理安全能力三个方面规定了大数据服务提供者的大数据服务安全能力建设要求，其中：

a) 组织管理安全能力：按照信息安全管理体系要求制定大数据安全策略与规程，从大数据服务组织与人员的安全管理，以及大数据组织数据资产与系统资产管理维度，确保大数据服务提供者的数据安全管理制度满足数据安全合规及数据安全风险管控要求。

b) 数据处理活动安全能力：针对数据收集、存储、使用、加工、传输、提供、公开、销毁等数据处理活动，从大数据平台和大数据应用业务及技术层面实施数据安全保护措施，满足大数据服务中数据处理活动相关的数据安全保护要求。

c) 数据服务安全风险管理能力：按照大数据服务中数据业务流转过程和数据处理活动安全能力要求，从风险识别、安全防护、安全监测、安全响应和安全恢复五个环节建立数据服务风险管理的安全能力，采取风险应对措施确保大数据系统运营中的数据服务及其数据资产始终处于安全保护状态，同时保障大数据服务的可持续性。

大数据服务提供者依据被保护数据资产和系统资产的重要性，以及大数据服务遭受破坏可能导致的危害程度，结合自身的服务模式、服务目标和支撑大数据服务的软硬件设施、大数据平台与大数据应用功能，参考本文件的安全能力要求项进行组织管理安全能力、数据处理活动安全能力和数据服务风险管理安全能力的建设和评估，确保大数据服务中数据业务连续性和数据服务安全风险的可控性。

5 组织管理安全能力

5.1 策略与规程

大数据服务提供者应：

a) 明确数据安全方针、目标和原则，制定相应的数据安全管理制度和大数据服务业务相关的数据安全策略和规程，并将其分发至大数据服务相关的职能部门、岗位和人员；

b) 确保制定的数据安全策略与规程覆盖组织的全部数据处理活动，内容包括目的、范围、方式、岗位、责任、管理层承诺、内外部协调及合规性要求等；

c) 建立供应链安全管理制度和管控规程，以书面协议等方式与供应链上下游组织明确约定交换共享数据的使用目的、数据范围、数量、供应方式、安全责任与义务、保密要求等内容，包括大数据服务相关网络产品和服务筛选机制、筛选指标和评价方法；

d) 制定并实施与数据安全管理制度、数据安全策略和规程、供应链安全管控规程相匹配的大数据平台和大数据应用安全技术机制与措施的实施细则和操作规范；

e) 建立数据安全风险评估和个人信息保护影响评估机制，制定应对数据安全风险的实施细则和操作规范，包括向主管部门报送评估报告的机制；
f) 建立数据安全事件投诉、举报渠道及受理处置流程，公布接受投诉、举报的联系方式、责任人信息等，及时受理、处置数据安全事件投诉举报，对投诉、举报人的相关信息予以保密，保护投诉、举报人的合法权益，必要时公开披露收到的投诉、受理进度以及最终处理结果，接受社会监督；
g) 在组织架构发生重大调整或大数据业务发生重大变化时，或因合并、分立、解散、被宣告破产等原因需要转移数据时，及时评估数据安全管理制度及数据安全策略与规程的适用性，并根据评估效果对数据安全管理制度、数据安全策略和规程和供应链安全管理等文件进行修订；
h) 定期通过 GB/T 22080 信息安全管理体系的认证、GB/T 22239 网络安全等级保护的测评等方式，对组织数据安全管理制度、数据安全策略和规程、大数据系统安全保护措施等进行持续改进，明确大数据服务安全能力持续提升计划和实施机制；
i) 建立落实数据安全法律法规及相关数据安全保护责任考核制度，对敏感个人信息处理、重要数据处理等合规性做出具有法律效力的数据安全声明或承诺；
j) 建立机器可读的数据策略与规程履行机制，对大数据服务过程中涉及角色管理、授权控制、安全审计、资源调配、服务跟踪等策略与规程实施过程进行记录与归档；宜实现自动化的大数据服务策略与规程的执行情况的跟踪、合规审核和服务违法行为的追溯。

5.2 组织与人员

5.2.1 组织与职责

大数据服务提供者应：

a) 成立数据安全管理机构，明确负责大数据服务安全工作的职能部门及岗位的职责，确保组织数据安全管理制度及数据安全策略与规程能够得到有效实施；
b) 明确数据安全负责人，并配备必要的专职人员，负责数据安全管理机构日常的数据安全管理工作；组织最高管理人员宜作为组织数据安全负责人，配备专职的数据安全管理人员和技术人员；
c) 明确大数据系统同步规划、同步建设和同步运营的职能部门及岗位的追责机制，确保大数据系统运营安全风险的可控和大数据服务的业务可持续性；
d) 明确重要数据、敏感个人信息管理等重要岗位的数据安全风险管理要求及安全责任考核要求；宜设置专职的重要数据和个人信息安全风险管理的重要岗位；
e) 明确数据安全风险评估与数据安全应急处置工作的职能部门、职责及追责工作机制，确保能监督管理大数据服务相关数据处理活动的安全；
f) 建立有效沟通和汇报机制，确保数据安全相关人员能够及时了解岗位职责和权力，数据安全事件、隐患等风险信息能够及时汇报给职能部门的岗位人员和负责人；
g) 依照组织监督考核机制，落实执行对大数据安全管理机构及其岗位的监督检查和考核制度，定期对数据安全管理机构、职能部门和安全岗位进行安全责任评估。

5.2.2 角色管理

大数据服务提供者应：

a) 建立大数据安全角色及其授权机制，明确用户角色的分配策略和授权范围；
b) 建立职能部门和安全岗位的角色清单和授权机制，明确大数据服务安全工作的职能部门和安全岗位的角色安全要求；

c) 建立大数据安全角色权限及其人员角色分配和角色授权定期审核机制，及时更新用户角色及角色授权信息；
d) 建立符合大数据系统技术架构和数据处理活动的角色分层授权、职责分离等方面的安全角色管理机制，包括大数据平台及其应用的安全角色协调管理机制；
e) 建立大数据服务上下文可感知的角色启用、停用和禁用的动态管理策略、操作规程和使用机制，包括大数据处理环境下的数据迁移和算法迁移的动态授权和上下文感知的数据访问控制机制。

5.2.3 人员管理

大数据服务提供者应：

a) 制定大数据服务人力资源管理安全策略，明确不同职能部门和安全岗位人员的职责、权力和资质能力等要求；
b) 制定人员招聘、录用、培训、上岗、调岗、离岗、考核、选拔等环节中大数据服务人员安全管理的操作规程；
c) 列出重要岗位清单，并明确重要岗位的人员能力要求，确定相应的考核内容与考核指标，在录用前开展重要岗位人员背景调查，在重要岗位人员调离或终止劳动合同时确保其所拥有的角色和责任都被转移给新的责任人员，并与其签订调岗和离岗保密协议；
d) 建立数据安全责任奖惩管理制度，并按照制度规范对造成损失的人员追责，记录人员责任奖惩信息；
e) 建立第三方人员安全管理制度，对处理个人信息、重要数据等重要岗位的外包人员进行能力考核和聘用审批，定期对第三方人员规范性行为进行安全审查，并按照合规要求签署保密协议；
f) 与所有涉及大数据服务岗位人员签订安全责任协议，人员调离或终止劳动合同时归还组织的软硬件资产，及时变更岗位变动人员的数据处理权限，终止离岗人员的所有数据处理权限。

5.2.4 培训管理

大数据服务提供者应：

a) 制定数据开发利用和数据安全保护相关的教育培训计划，确保每年组织开展全员数据安全教育培训，并依据培训反馈效果定期对教育培训计划进行审核和更新；
b) 按计划采取多种方式培养数据开发利用人员和数据安全专业人员，包括内部培养、外部培训等，并对包括法律、法规、政策、标准、技术、技能、安全意识等内容的培训结果进行考核、记录和归档；
c) 制定重要数据、敏感个人信息管理等重要岗位的转岗、岗位升级等相应的人员安全能力要求的教育培训计划，并对培训计划、培养方式、培训内容定期审核和更新；
d) 针对涉及重要数据、敏感个人信息等处理的重要岗位的人员，开展大数据服务安全操作技能培训与培训效果实践考核。

5.3 资产管理

5.3.1 数据资产

大数据服务提供者应：

a) 建立数据资产安全管理规范，明确数据资产的安全管理目标和原则；
b) 依据数据分类分级策略，明确数据分类分级的制度和操作规程，以及数据分类分级的变更审批流程和机制；

c) 依据数据资产登记制度，建立数据资产清单，明确大数据服务相关的数据资产的基础属性、分类、级别及相关方的权限和责任等安全属性；
d) 建立数据资产管理平台，包括组织内外部数据资源整合操作规程，实现对数据资产的统一管理；
e) 建立数据资产安全属性自动标记机制，包括安全属性标记策略、标记定义和标记变更机制；
f) 建立数据资产操作审计机制，确保数据资产管理操作行为的可审计、可追溯；
g) 定期审核和更新数据资产安全管理相关的安全管理规范、分类分级策略、操作规程及其数据资产清单等，确保数据资产管理平台中数据的实时性和准确性。

5.3.2 系统资产

大数据服务提供者应：

a) 建立系统资产安全管理规范，明确系统资产安全管理目标和原则；
b) 建立系统资产建设和运营管控措施，明确规划、设计、采购、开发、运行、维护及报废等系统资产管理过程的安全要求，包括内外部人员在任职期内领用和归还系统资产以及在终止任用、合同或协议时归还所使用系统资产的管理制度和机制；
c) 建立系统资产登记制度，形成系统软硬件资产清单，明确系统资产安全责任主体及相关方权责清单，并及时更新系统资产清单信息；
d) 建立系统资产分类和标记规程，使资产标记易于填写和依附在相应的系统资产上；
e) 建立系统资产管理平台，实现对系统资产的统一注册、管理和使用监控等，并具备对系统资产清单、系统权责清单等进行持续更新的能力；
f) 定期审核和更新系统资产管理的安全管理规范、运营管控措施、资产登记制度和系统资产清单等，确保系统资产管理平台中数据的实时性和准确性。

6 数据处理活动安全能力

6.1 数据收集

6.1.1 数据获取

大数据服务提供者应：

a) 制定数据获取操作规程，明确数据的获取源及其数据格式、获取目的和用途、范围、期限和频次，规范数据获取渠道、获取流程和获取方式；
b) 评估数据获取环境、设施和技术工具对数据提供者的网络服务的性能、功能带来的影响，确保数据获取过程不干扰数据提供者的网络服务功能；
c) 定期评估数据获取的范围、流程、频次、渠道、方式等，确保数据获取操作的合规性、正当性和一致性；
d) 在发现可能违反法律、行政法规或者行业自律公约，或者侵犯他人知识产权等合法权益时，立即停止获取数据的行为并采取相应的补救措施；
e) 采取技术和管理措施确保数据获取相关的自动化工具在获得授权后才能收集数据，对超规模、超范围的数据获取等异常行为进行告警，并按照规定及时告知数据提供者；
f) 对涉及重要数据和个人信息自动化获取的场景，具备应对潜在数据泄露风险的技术防范措施，满足合规检查、风险评估和安全事件上报等能力；
g) 跟踪和记录数据获取过程，具备对数据获取操作过程的可追溯能力。

6.1.2 数据标识

大数据服务提供者应：

a) 建立数据识别和标记的操作规程，建立重要数据和个人信息的识别和标记行为的审批管理规程；
b) 采用技术措施对收集的数据进行识别，并依据数据分类分级策略对数据安全属性进行标记；宜配置自动发现和标识重要数据和个人信息的技术工具或服务组件；
c) 配置数据标识管理工具，对数据识别和标记的变更操作进行合规性审核，实现对数据的标识、审核及标识结果使用过程的管理，确保数据识别和标识变更的可追溯性；
d) 定期对数据识别和标记的效果、影响范围和程度变化等数据安全风险进行评估，并采取适当的风险处置措施，降低数据识别和标记带来的安全风险。

6.1.3 数据清洗

大数据服务提供者应：

a) 制定数据变换、转换、纠错、编码等数据清洗的操作规程，明确数据清洗的要求、规则和方法，确保数据清洗操作前后数据间的映射关系；
b) 采取技术手段和管理措施，确保在数据清洗操作中对所获取或清洗操作生成的数据进行保护，包括但不限于衍生数据以及操作日志等；
c) 记录数据清洗操作的行为，在数据清洗完成后对其操作产生的中间或临时数据进行安全删除。

6.1.4 数据加载

大数据服务提供者应：

a) 综合数据量、增长速度、业务需求、数据加载有效性等因素制定不同数据源、跨安全域数据加载的操作规程，明确数据安全加载的要求、规则和方法；
b) 具备对数据加载终端、用户或加载服务组件等进行身份鉴别的能力；宜采取多因素身份鉴别技术，满足数据加载操作的真实性和合法性要求；
c) 采用安全传输技术或数据加密技术实现远程数据加载，保证跨安全域数据的安全传输和加载；
d) 记录数据加载操作过程，在数据加载完成后对数据加载通道缓存的数据进行安全删除；
e) 具备确保数据加载操作过程安全可靠的能力，提供数据加载通道的冗余备份、加载接口的流量过载监控等安全措施。

6.2 数据存储

6.2.1 存储架构

大数据服务提供者应：

a) 建立可伸缩的数据存储架构，提供存储模块的装载与可卸载的存储空间扩展功能，满足大数据平台管控数据资源的可持续增长的可扩展性存储需求；
b) 制定数据存储架构安全管理的操作规程和管控规则，包括外部身份鉴别与授权控制规则、存储数据转移安全规则、存储数据完整性和多副本数据一致性的管理规则等；
c) 建立数据逻辑存储安全管理操作规范，具备对数据逻辑存储结构的分层和分级保护能力，满足不同数据类型、不同数据容量、不同业务需求和不同数据用户的逻辑存储安全管理要求；
d) 建立数据分片和数据分布式存储等逻辑存储和物理存储安全管理操作规范和规则，满足分布式存储数据完整性和保密性要求；
e) 确保数据存储架构具备采用符合国家密码管理部门相关规定的密码技术进行加密存储的能力，且满足不同技术架构层次的加密存储需求；宜提供存储密文数据使用的自动溯源技术机制；

f) 确保数据存储架构具备容灾备份的能力；宜提供数据存储跨机柜、跨机房或跨地域的容错部署和容灾备份的基础设施能力；
g) 落实数据隔离存储和境内存储要求，提供数据分离存储、分布式存储等降低数据安全风险的方法及技术机制，确保在境内收集和产生的重要数据和个人信息存储在境内。

6.2.2 数据副本

大数据服务提供者应：
a) 建立数据存储冗余策略和数据复制、备份和恢复的操作规范，包括数据同城备份与异地容灾备份方案与机制，以满足大数据服务可靠性、可用性等目标；
b) 建立数据冗余强一致性、弱一致性等控制策略与操作规范，满足数据不同一致性水平需求的数据复制等副本多样性和多变性存储管理要求；
c) 建立数据副本的定期检查和更新工作程序，包括数据副本更新频率、保存期限等，，定期对数据副本的一致性进行检测验证，确保副本或备份数据的有效性；
d) 对复制、备份等生成的数据副本执行和数据源同样的安全管控措施，包括访问控制、完整性校验等机制；
e) 建立数据复制、备份和恢复操作的日志记录规范，记录数据复制、备份和恢复等数据副本操作过程，确保副本数据管理过程可溯源；
f) 具备对数据副本存储的多种压缩策略和实现机制，确保压缩后副本数据的完整性和可用性。

6.2.3 数据归档

大数据服务提供者应：
a) 依据大数据业务连续性、数据资产的重要性等需求，建立数据处理各种活动数据的归档存储相关的操作规程；
b) 建立在线或离线的多级数据归档架构，支持大数据服务中各种场景的数据归档要求和海量数据的有效归档和恢复；
c) 建立归档数据的安全策略和管控措施，确保非授权用户不能访问归档数据；
d) 建立归档数据的压缩或加密策略，确保归档数据存储空间的有效利用和安全访问；
e) 定期地采取技术手段和管控措施查验归档数据的完整性和可用性；宜建立归档数据的安全审计与安全删除制度，并指定专人负责归档数据的安全管理。

6.2.4 数据保留

大数据服务提供者应：
a) 建立数据存储时效性管控措施和操作规程，确保按照法规要求和主管部门的数据保存时长要求对相关数据留存操作予以记录；
b) 明确数据的访问、存储、删除等存储数据操作的有效期，具备数据存储授权时效性与数据保留周期的管理和控制能力；宜配置数据时效性组件实现数据保留有效期合规性的自动化检测；
c) 建立对过期的存储数据、相关备份与归档数据的安全删除方法和技术机制，限期删除已完成数据处理目的、达到存储期限的数据，并能够将数据安全删除的验证结果告知相关方；
d) 因收购、兼并、重组、破产或其他合同等因素，导致其数据服务转让或终止，确保其备份或归档的冗余数据，包括操作日志等衍生数据的保留时长符合相关规定；
e) 具备不同时效性的数据分层存储管理机制与技术；宜按照时效性实现数据在各分层间的自动迁移，确保大数据使用者能高效地访问和获取存储数据。

6.2.5 密钥存储管理

大数据服务提供者应：

a) 建立数据存储架构的密钥管理操作规范，对密钥的产生、分发、存储、使用、更新、归档、撤销、备份、恢复和销毁等进行密钥全生命周期的管理；宜具备密钥集成管理、密文数据透明处理等密钥存储管理能力；
b) 建立支撑多层存储加密需求的密钥管理操作规范，支持商用推荐密码算法的密钥管理互操作性等相关标准规范的密钥服务。

6.2.6 多租户数据存储管理

大数据服务提供者应：

a) 制定多租户数据存储隔离操作规程，包括租户应用上线、下线的数据迁移操作规程；
b) 建立多租户数据安全存储技术机制，提供数据分割、存储区切割等技术机制来隔离多租户数据资源；宜依据不同租户空间内的用户需求，提供用户定制化的租户数据安全保护技术机制；
c) 建立多租户数据复制、备份、归档等存储冗余信息保护技术机制与操作规范；
d) 建立多租户数据可用性保障策略和技术机制，包括但不限于故障转移、多副本自动管理等。

6.3 数据使用

6.3.1 授权管理

大数据服务提供者应：

a) 明确数据的管理责任主体，包括数据主体、数据控制者或数据处理者及其相关参与者的权利的约束和限制条件，确保数据的使用不能损害相关权利人的合法权益；
b) 建立数据使用授权管理操作规范与技术机制，包括多层次主动防御机制和数据安全保护措施的授权管理方法；
c) 建立数据使用审计记录和受保护的数据使用日志的分布式存储机制和管控措施，具备对潜在违约数据使用者的责任识别和处置能力；
d) 建立收集和汇聚数据的使用权利体系，包括建立在合法正当、知情同意等原则基础上确定的原始数据收集目的、用途等底层权利，和汇聚数据控制者的数据财产权等权利；
e) 具备对违约责任、缔约过失责任、侵权责任等数据使用风险的分析和应对处置能力。

6.3.2 访问控制

大数据服务提供者应：

a) 按照数据处理活动和大数据技术架构，综合主体角色与信用等级、数据分类分级与业务需要、数据使用时效性等因素，采用相应的访问控制模型，实施大数据使用者身份标识与鉴别、数据访问权限分配等策略，实现数据使用相关授权的管控；
b) 利用大数据系统的分布式层次访问控制技术实施大数据使用者身份标识与鉴别、数据访问控制等策略，并提供集成外部数据资源的数据访问控制的接口；
c) 建立面向大数据系统不同技术层次的数据访问接口、管理工具和管理命令的管控措施和技术机制，包括访问控制时效的管理和验证、应用接入数据存储的合法性和安全性取证机制等；
d) 建立访问控制管控操作审计记录及其日志数据的分布式存储机制和管控措施，具备对完整的数据使用和访问操作进行记录、管理、安全审计和事件归因等能力。

6.3.3 数据展示

大数据服务提供者应：

a) 建立数据展示操作规范，对数据的展示范围、内容、方式等进行安全评估，确定数据展示的必要性和安全性；宜具备根据大数据使用者权限在大数据技术架构不同层次自动展示相应安全级别数据的能力；

b) 在展示重要数据和敏感个人信息时，采用数据脱敏等技术，并通过防截屏、屏幕水印等控制措施，降低数据展示时的数据泄露风险；

c) 在数据展示完成后，及时安全删除本地缓存或展示通道缓存中的数据，包括数据展示操作产生的中间或临时数据。

6.4 数据加工

6.4.1 分布式计算

大数据服务提供者应：

a) 建立分布式计算节点间可信连接策略和操作规范，采用节点认证、可信计算等技术机制来确保大数据服务处理节点接入的真实性；

b) 建立分布式计算节点和用户安全属性的周期性确认机制，确保预定义分布式计算安全策略存储和数据加工相关算法在各节点实施技术机制的一致性；

c) 建立分布式计算过程中数据文件鉴别、用户身份鉴别的策略和规范，确保分布式计算数据和主体行为的可信性；

d) 建立分布式计算过程中不同节点的数据副本的更新检测机制，确保数据副本的完整性、一致性和真实性；

e) 建立分布式计算中数据泄露控制技术机制，防止除计算结果之外的信息被泄露，包括但不限于数据分布式处理过程中的调试信息、日志记录、缓存数据、中间数据等；

f) 制定数据分布式计算节点的服务组件自动维护策略和管控措施，包括虚假节点监测、故障节点自动修复的技术机制，避免云计算环境或虚拟计算环境下潜在的对数据处理算法、机器学习模型等算法与模型迁移执行的网络攻击。

6.4.2 大数据分析

大数据服务提供者应：

a) 建立大数据分析相关数据源的数据获取、汇聚及使用操作规范，明确大数据分析相关数据获取、汇聚及使用方式、访问接口、授权机制等；

b) 在进行大数据分析前，对多源数据的汇聚开展数据安全评估，建立多源数据聚合、关联分析等容易生成如重要数据的大数据分析过程中的数据资源操作规范和安全控制实施规范；

c) 对大数据分析结果进行合规性评估，避免其包含可恢复的重要数据和个人信息，控制大数据分析结果的数据提供和公开活动，降低分析结果使用中存在的重标识等数据安全风险；

d) 制定个人信息保护影响评估规程和报告模板，进行大数据分析的个人信息保护影响评估，并对个人信息的分析过程和结果进行记录；

e) 在利用算法推荐服务或进行自动化决策时，遵循公平公正、公开透明、科学合理和诚实信用等原则；宜建立多方参与的算法推荐服务治理机制，具备对第三方提供集成算法的审核能力；

f) 定期审计大数据分析过程记录和分析结果的存储、使用、提供等数据处理活动的合规性，具备对大数据分析结果质量和分析算法执行真实性进行数据溯源的能力；

g) 具备构建大数据分析过程及数据血缘图谱的能力，根据血缘关系对大数据分析数据及其衍生数据实施一致的管控。

6.4.3 密文计算

大数据服务提供者应：

a) 明确需要进行密文计算的数据资产，可自动或人工选择相适应的密码协议和机制，对密文计算过程进行保护；
b) 具备保存密文计算日志存证的能力，确保密文计算行为的可追溯性；
c) 具备对密文数据进行搜索、排序、计算等透明处理的能力；
d) 具备验证密文计算数据真实性、正确性的能力。

6.4.4 数据脱敏

大数据服务提供者应：

a) 建立数据脱敏管理操作规范，明确数据脱敏规则、脱敏方法和使用限制，包括脱敏后数据的可恢复性评估机制；
b) 明确梳理出需要脱敏的数据资产，制定不同分类分级数据的脱敏处理流程；
c) 配置数据脱敏的技术工具或服务组件，支持如泛化、抑制、干扰等数据脱敏技术；宜具备非结构化数据进行脱敏处理的能力，以及对脱敏处理效果的验证能力；
d) 支持在数据脱敏时保留其原始数据格式和特定属性，以满足基于脱敏数据的开发与测试要求；
e) 对数据脱敏处理过程相应的操作进行记录，以满足数据脱敏处理安全审计要求。

6.5 数据传输

大数据服务提供者应：

a) 区分安全域内、安全域间的数据传输场景，建立安全域内、安全域间不同场景的数据传输安全策略和操作规程；
b) 建立大数据系统运营安全策略相应的安全控制措施，如安全通道、可信通道、数据加密等，保证大数据系统中传输数据的保密性和完整性；
c) 建立大数据系统中数据传输接口安全管理操作规范，包括跨安全域的数据安全传输密钥管理和加密操作接口规范；
d) 具备在构建传输通道前对两端主体身份进行标识和鉴别的能力，包括跨域传输数据前应对传输双方的安全级别进行评估的能力，避免将高安全等级数据传输到低安全等级的安全域；
e) 具备对传输数据的完整性进行检测的能力以及相应的恢复控制措施；
f) 具备对数据传输安全策略的变更进行审核和监控的能力，包括对通道安全配置、密码算法配置、密钥管理等保护措施的审核及监控的能力；
g) 建立数据传输链路的冗余、恢复机制，保证数据传输链路的可靠性，并采用断点续传、超时重新连接等技术机制保障数据传输任务的可靠性。

6.6 数据提供

6.6.1 组织内提供

大数据服务提供者应：

a) 建立组织内跨安全域数据提供安全操作规范，明确数据提供涉及的职能部门和岗位相关的用户职责和权限，保证组织内数据提供安全策略的有效性；
b) 审核组织内跨安全域提供数据的应用场景及其数据内容，确保数据接收方的数据使用和数据加工没有超出大数据服务提供者授权的数据使用范围；

c) 对数据提供过程进行监控，并采用数据加密、安全通道等管控措施提供数据，定期评估数据提供通道的安全性；宜利用专业数据提供工具或组件提供自动化安全保障；
d) 制定数据提供安全审计策略和审计日志管理操作规范，记录数据提供活动日志，为数据提供相关安全事件的处置、应急响应和事后调查提供证据支撑；
e) 在数据提供完成后对数据提供通道缓存的数据进行安全删除。

6.6.2 跨组织提供

大数据服务提供者应：

a) 建立跨组织数据提供的安全影响分析和风险评估操作规程，涉及重要数据共享、转让或委托处理时应与数据接收方通过合同、协议等形式明确双方的数据安全保护责任和义务，确保数据接收方具备不低于数据提供者的数据安全保护能力；
b) 在数据处理者依法向其他处理者提供其处理的个人信息时，向个人信息主体告知接收方的名称或者姓名、联系方式、处理目的、处理方式和个人信息的种类，并征得个人信息主体的同意；
c) 建立跨组织数据提供的安全事件应急处理机制，发生重大事件时，及时终止数据提供，并要求数据接收方按 6.8 的要求销毁已接收的数据；
d) 在向境外组织提供个人信息和重要数据前，按照合规性的要求组织开展数据出境安全评估；确实要向境外提供重要数据和达到相关规定数量的个人信息时，需上报主管部门，并通过法律法规要求的数据出境安全评估；
e) 督促和监督数据接收方加强数据安全管理，发现其未落实合同规定要求和责任时督促数据接收方及时整改，必要时终止数据提供活动，并按 6.8 的要求销毁已接收的数据；
f) 通过合同、协议等形式与数据接收方建立责任管理机制，对接入的数据接收方发生兼并、重组、破产时，要求数据接收方继续履行相关数据安全保护义务；对于没有能力继续履行相关义务的，要求数据接收方以合同中约定的形式返还、删除其接收和产生的数据，签订数据删除确认书；
g) 建立机器可读的数据提供格式规范，具备对嵌入的第三方数据提供自动化工具进行安全监测的能力，当发现超出双方约定的数据处理活动时及时停止数据接入服务；
h) 建立必要的数据出境管控技术机制和措施，如要求数据接收方提供数据安全保护能力的证明，配置数据脱敏处理使用不可逆的加密算法等；
i) 在委托处理、向其他组织提供、公开或向境外提供个人信息时，进行个人信息安全影响评估，并对处理情况进行记录，个人信息安全影响评估报告和处理情况记录至少保存 3 年，共享、交易、委托处理重要数据的审批记录、日志记录至少保存 5 年；
j) 定期对数据接收方的数据销毁机制进行验证，确保跨组织数据提供到期后其在数据接收方无数据残留，签订数据删除确认书。

6.7 数据公开

6.7.1 数据发布

大数据服务提供者应：

a) 建立数据主动公开管理制度和操作规范，明确发布大数据使用者的权利和义务；
b) 建立数据发布的管理措施与机制，包括数据发布的方式、范围，以及数据内容审核及审批制度；
c) 提供数据发布清单，包括发布数据摘要、数据格式、更新频率等内容，以及使用条件等；
d) 提供发布数据的访问接口及格式规范，确保用户能高效的获取发布数据；
e) 具备对待发布数据进行重要数据及个人信息的检查能力，确保涉及重要数据、个人信息等数据不向社会公开发布；宜具备自动化识别、检测和实时数据脱敏处理等能力；

f) 定期审核发布数据资源使用报告，评估数据公开发布带来的数据安全风险；
g) 在涉及用户数量巨大、数据处理活动对社会影响大时，定期发布社会责任报告，包括数据保护措施、发生的安全事件及应对处理情况等内容。

6.7.2 在线访问

大数据服务提供者应：
a) 建立外部组织在线访问数据的授权操作规范和数据使用操作规范，包括访问申请的登记、审核、审批、办理、归档等工作制度，明确在线数据访问申请管理的审核内容；
b) 建立外部组织在线访问数据的数据使用责任追究操作规范，包括在线访问中安全事件应急保障处理流程及管控措施；
c) 建立数据在线访问目录库和外部组织在线访问数据的渠道，包括数据访问接口及格式规范，如提供机器可读的可扩展标记语言格式，确保用户能高效地访问可公开的数据资源；
d) 通过大数据平台服务组件实现外部组织在线访问数据的申请的登记、审核、审批等管理功能，提供包括用户身份鉴别、访问服务组件的互认等安全管控功能；
e) 采用自动和人工审计相结合的手段对重要数据和个人信息等高风险数据的在线访问操作进行监控；宜具备对异常或高风险数据访问操作的自动化识别和实时预警能力。

6.8 数据销毁

6.8.1 数据删除

大数据服务提供者应：
a) 建立数据删除操作规范，建立法律法规要求的重要数据或个人信息多人和多级的操作模式，明确大数据服务停止运营、用户账户注销、用户申请数据删除等场景的数据删除操作规程；
b) 建立物理删除和逻辑删除的数据删除方法和技术，明确不同类别和级别的数据删除方式和技术要求，如基于安全策略、基于分布式杂凑算法等分布式存储的删除技术；
c) 建立不可逆数据删除机制，配置必要的数据删除工具，确保能根据业务场景需求以不可逆方式删除相关的数据及其衍生的各种冗余数据；
d) 在用户主动提出删除请求，注销账户或者变更、撤销授权时，以及因使用自动化采集技术等无法避免采集到的非必要个人信息，在约定时间对个人信息进行删除或进行匿名化处理，确保个人信息及相关数据不可访问或不可重标识；
e) 主动、及时删除已实现数据处理目的或者实现处理目的不再必要的数据，包括数据处理过程中备份数据、衍生数据及操作日志数据等；
f) 建立数据删除效果评估和复核机制，定期检查已被删除的数据是否还能访问；
g) 监督数据删除操作及其删除效果反馈过程，包括已共享或者已被其他用户使用的数据的删除技术管控措施；宜通过大数据平台提醒数据管理人员及时发起对共享数据的删除。

6.8.2 介质管理

大数据服务提供者应：
a) 建立大数据系统存储介质访问和使用管理规范，对存储介质进行标记，明确存储介质的数据类型，采取有效的介质净化技术和操作规程对存储介质进行净化；
b) 依据介质存储内容的重要性明确磁介质、光介质和半导体介质销毁方法和机制；
c) 按照法律法规和标准规范销毁存储介质，使用经认证的物理销毁设备或请具备资质的单位对存储重要数据和个人信息的介质设备进行物理销毁；

d) 制定对存储介质进行销毁的监管措施，确保对销毁介质登记、审批、交接等介质销毁过程监控；
e) 建立介质管理系统，确保对存储介质的使用和传递过程进行全程跟踪，对介质访问、使用、销毁等过程进行记录和审计，并定期对销毁记录及介质销毁效果进行检查。

7 数据服务风险安全管理能力

7.1 风险识别

7.1.1 数据安全风险识别

大数据服务提供者应：

a) 采用接口扫描、流量分析、业务监测等方式构建数据资产的识别能力，及时对数据处理活动相关的数据源信息、数据资产等进行标识，并更新数据资产目录；
b) 对数据处理活动组件进行威胁及脆弱性识别，分析其对大数据服务安全的影响，形成数据安全风险分析报告，必要时将风险应对的建议措施提交给安全管理机构审批；
c) 具备基于数据安全风险分级的数据资产的分类分级能力，能通过技术手段对识别的数据资产分类分级等安全属性进行自动化标记；
d) 在涉及重要数据和个人信息的数据处理活动时，能生成数据安全风险分析报告和风险应对的建议措施，并按规定提交相关主管部门；
e) 配置或开发识别大数据平台数据处理活动的组件，以及大数据应用身份信息等业务元数据管理组件，识别包括开源程序在内的大数据平台和大数据应用的脆弱性和威胁信息；
f) 针对应用于关键信息基础设施的大数据平台提供定制化的数据处理和大数据系统运营安全控制措施，且通过主管部门认可的第三方评估机构的检测、评估和认证；
g) 建立数据处理活动安全风险知识库，对大数据服务的过程和大数据系统运营行为等日志进行汇聚和融合分析，自动生成大数据服务涉及的数据安全风险信息；
h) 针对数据提供、数据发布、流程审批、钓鱼邮件、勒索病毒等安全风险重点场景，定期进行数据处理活动安全管控措施的隐患排查和治理，提升对数据安全风险应对措施的识别能力；
i) 定期自行开展或邀请第三方专业机构开展数据安全风险识别和分析工作，按规定保存数据安全风险分析报告，包括识别风险应对情况记录等信息。

7.1.2 供应链安全风险识别

大数据服务提供者应：

a) 定义数据服务供应链上下游组织的数据交换共享的格式规范、接口规范，约定数据提供和数据获取方式等技术或管理措施，确保大数据系统中供应链的数据处理活动组件及流通数据管控操作的透明性和可识别性；
b) 对数据服务供应链上下游组织的网络产品和服务的脆弱性和威胁，包括适用法规和合同要求进行识别，形成供应链风险分析报告，必要时将风险应对的建议措施报送给安全管理机构审批；
c) 及时向大数据服务供应链上下游的组织，发布所识别的供应链风险及应对风险的建议措施，确保数据供应链中数据收集、数据提供和数据公开中数据安全风险可控；
d) 基于主管部门的监管要求，定期对供应链上下游组织的数据处理活动中的数据安全风险、支撑大数据服务的网络产品和服务安全技术和措施有效性进行分析和评估。

7.2 安全防护

7.2.1 区域边界防护

大数据服务提供者应：

a) 对大数据服务涉及的区域边界进行划分，制定区域边界数据流转的安全策略和操作规程；
b) 对跨区域边界的数据使用、提供、公开等数据处理活动防护措施进行检查，确保数据安全保护责任不随数据跨区域边界转移而改变，在不同区域间的数据安全策略实施效果保持一致；
c) 制定区域边界安全防护设施策略和规程更新维护规则，并采用必要的手段或管控措施确保更新后区域边界安全策略与规程得到实施；
d) 在涉及重要数据和个人信息的跨区域边界访问时，对应用接口或服务接口等访问主客体进行双向身份鉴别，保证不同区域边界之间的安全管控措施有效并与风险程度相适应；
e) 建立基于主客体属性和区域边界上下文的逻辑访问控制措施及机制，建立跨区域边界数据动态访问控制机制。

7.2.2 计算环境防护

大数据服务提供者应：

a) 建立数据感知、保护、预测、响应等一体化的多层次计算环境安全防御体系，采取相关的安全管控措施确保大数据系统满足网络安全等级保护制度的纵深防御要求；
b) 制定数据安全风险管理相关的数据服务安全基线配置清单，启用、禁止或限制大数据平台和大数据服务管理平台特定的功能、端口、协议或服务；
c) 建立终端智能设备、第三方或开源系统与服务组件等计算设施接入约束规范，采用技术工具对大数据服务的接入设备、服务组件及系统等计算设施安全属性进行管理；
d) 制定计算环境安全初始化策略，包括数据存储等大数据服务模块自启动检查机制，确保大数据服务数据在各种故障重启后的数据一致性和完整性；
e) 制定满足可靠性与可用性的计算环境垂直扩展、水平扩展策略，提供海量数据或复杂类型数据高效处理方法及其安全保护技术与机制；
f) 提供上下文感知的细粒度授权管理和访问控制功能，如依据资产安全属性和用户属性设置授权规则和访问控制措施等，以支持分布式计算环境中大数据分析及人工智能算法迁移的安全性；
g) 建立统一身份管理平台，支持用户证书、账户、授权、认证、审计等安全元数据统一管理；宜为处理重要数据和个人信息用户授权管理提供独立的身份管理物理服务器；
h) 具备分布式用户身份鉴别、授权控制，安全审计的关联巡检功能，能根据法律法规要求及业务需求对大数据服务元数据进行核查，禁用非法账号、闲置账号、过期账号及彼此间的关联关系。

7.2.3 数据操作防护

大数据服务提供者应：

a) 制定供应链中数据流转安全管控策略，通过技术机制对大数据系统以及供应链中外部系统间的数据加载、数据交换、数据共享等数据流转操作进行控制；
b) 在大数据应用、供应链中关联业务组件下线以及智能终端设备退网时，通过技术手段执行规范的数据转移、转存或删除操作；
c) 响应数据主体对于个人信息或数据转移至指定数据接收方的请求，在符合主体要求条件时，提供个人信息和数据转移的途径；
d) 建立数据高风险操作清单及其管控措施，如在人工进行数据高风险操作时进行双人、双账户鉴别后协作完成，以及在程序进行高风险操作时通过基于密码技术鉴别的接口进行实现等；
e) 部署数据防泄露、数据脱敏、个人信息去标识化等安全功能组件，具备防范使用数据抓取、数据分析等从网络层和应用层获得重要数据和个人信息的能力；
f) 因业务确需向境外提供个人信息时，在通过国家主管部门组织的风险评估的基础上，要求境外

接收方提供适当的安全保证能力证明。

7.2.4 数据接口防护

大数据服务提供者应：

a) 制定数据服务接口安全控制策略，明确规定使用接口的安全限制和安全控制措施，如身份鉴别、授权策略、访问控制、数字签名、时间戳、安全协议等；
b) 建立数据服务接口清单，明确数据服务接口安全规范，包括接口名称、接口参数、接口安全要求等，具备对接口不安全的输入参数进行限制或过滤的能力，并提供异常处理功能；
c) 具备对重要数据和个人信息等服务接口的调用进行记录、汇聚和集中存储的能力，并具备通过大数据分析技术进行风险识别和安全性分析的能力。

7.2.5 威胁信息分析

大数据服务提供者应：

a) 具备对数据处理活动相关的威胁情报数据的收集、分析和利用能力，掌握大数据服务涉及数据处理活动所面临的威胁信息；
b) 具备对数据资产、系统脆弱性、安全事件等监测能力，对引起大数据服务安全态势发生变化的安全要素进行获取，对安全事件的发展趋势进行预测，对大数据服务安全态势进行多维度展示；
c) 具备对威胁情报数据的关联分析，并将威胁情报数据向大数据服务安全检测防御规则或控制措施的转化能力，并及时在大数据平台和大数据应用中进行实施；
d) 与专业机构建立威胁情报数据共享机制，持续提高数据安全风险和供应链安全风险应对处置和防范能力；
e) 采用自动化技术机制，对多源异构威胁情报数据进行归并、融合和分析，并进行主动、协同式的数据安全威胁检测、预警和应急处置。

7.3 安全监测

7.3.1 数据处理监测

大数据服务提供者应：

a) 建立数据处理活动及其服务接口的安全监管措施，具备对数据处理活动及其服务接口的访问进行自动化监控和应急处置的能力；
b) 建立数据处理活动的监测规则和安全基线，能根据预定义的阈值对数据处理活动异常行为进行告警，并展示数据处理活动发生的位置、操作以及数据处理活动的风险及威胁等信息；
c) 具备对数据处理状态和存储数据访问进行监测的能力，及时发现违规数据访问等数据使用行为，记录并安全存储数据处理状态和数据处理活动的监控日志；
d) 建立针对重要数据和个人信息流转监测机制，并采取技术措施及时发现对这些数据的违规操作或数据泄露等安全风险；
e) 部署数据防泄漏实时监控工具，对异常或高风险的跨区域边界的交换、共享等数据提供行为进行实时监控，发现异常时可自动阻断数据传输；
f) 建立重要数据和个人信息异常处理上报机制，当重要数据或个人信息发生泄露、非法篡改等情况时，及时采取处置措施，并按规定及时告知用户，向主管部门上报。

7.3.2 系统运行监测

大数据服务提供者应：

a) 建立大数据系统运行监测平台，支持分布式计算节点的处理器、内存、磁盘、网络流量等计算资源状态及其支持数据服务运行状态的统一监测；
b) 对分布式计算节点的用户身份鉴别和数据访问情况进行采集、存储和安全性分析，具备对计算节点接入、用户账号及数据使用等安全风险的监测和分析能力；
c) 具备对大数据系统越权访问、高频访问、恶意操作等运行异常行为的发现、记录、统计和分析能力，并按重要程度和影响程度对系统运行异常行为进行分级告警；
d) 具备对数据供应链、数据委托处理、第三方服务组件等重要业务相关应用和涉及重要数据和个人信息处理过程相关应用的监测能力，对系统运行状况及其数据流转情况等进行监测；
e) 建立大数据系统硬盘健康检测、多副本冗余、纠删码冗余、数据自动修复等机制，确保大数据系统及其管控数据的完整性和可用性。

7.3.3 服务持续监测

大数据服务提供者应：

a) 依据 5.1 策略与规程的要求，制定大数据服务持续监控机制，确定数据服务接口、业务连续性、供应链安全等服务水平协议监测指标和频率，对大数据服务安全措施进行持续监测；
b) 跟踪和控制大数据服务相关的网络产品和服务及其数据访问接口，及时对影响国家安全风险等相关网络产品和服务组件进行终止、重启等操作；
c) 对监测累积的大数据服务监测记录进行关联和分析，定期向安全管理机构汇报大数据服务及其数据资产的安全状态；
d) 使用自动化的监控工具维护大数据服务监控信息的准确性、真实性，确保数据安全风险和供应链安全风险的应对措施有效并持续的得到实施。

7.3.4 安全检查评估

大数据服务提供者应：

a) 定期对大数据安全控制措施进行检查，并在爆发网络攻击、重大安全漏洞时，及时按照上级主管部门要求、国家安全机构建议开展专项安全检查；
b) 跟踪大数据安全和个人信息保护法律法规和管理规定，结合大数据系统实际情况及时制定、完善组织内部的数据安全检查评估内容，定期开展通用和专项安全检查评估工作；
c) 在法律法规修订、组织业务重组、组织业务模式或运行环境发生重大变更、或发生重大数据安全事件时，及时地对数据安全保护措施进行安全影响评估；
d) 采取模拟攻击方式对数据处理活动及其组件进行安全风险评估，持续提升风险监测的能力。

7.4 安全响应

7.4.1 应急预案管理

大数据服务提供者应：

a) 建立大数据服务应急预案，包括应急组织机构与职责、数据安全事件分类分级、监测与预警、应急处置流程、保障措施等内容，且应急预案经本单位审核通过后，按要求报主管部门备案；
b) 开展大数据服务应急预案培训，培训内容覆盖全流程的数据处理活动、关键业务及数据安全应急所需的安全应对控制措施；
c) 建立大数据服务应急预案演练计划，保存演练记录和演练总结报告，在大数据系统本身、外界环境发生重大变化时，对应急预案进行更新，保留审核发布记录；

d) 与主管部门、第三方安全机构及其他相关职能部门间，建立大数据服务应急处理、协调、沟通渠道；
e) 建立大数据服务安全应急响应最佳实践知识库，包括数据泄露、数据篡改、数据破坏、网络勒索等不同类型的安全事件及处置办法，并用于应急响应培训及演练计划。

7.4.2 安全事件处置

大数据服务提供者应：
a) 建立安全事件处置操作规程，明确安全事件的处置方法，包括不同安全事件分类分级、启动条件及所需的资源，不同类别、级别事件的响应、处置和报告流程；
b) 明确大数据系统遭受破坏时，恢复关键数据业务和全部数据业务的预期处理时间，并在大数据系统发生故障、受到损害或发生中断时，在指定的时间内完成关键业务处理活动的恢复；
c) 及时向可能受影响的部门和人员、供应链相关方、与安全事件相关的其他组织通报安全事件；
d) 在发生有可能危害重要数据或危害国家安全等关键业务的安全事件时，立即组织研判，在规定时间内形成安全事件报告，按规定及时向安全管理机构上报；
e) 在安全事件涉及个人信息时,及时告知受影响个人信息主体，涉及大量个人信息的，按规定及时向有关主管部门报告；
f) 在安全事件处置完成后，及时调查安全事件的直接原因和间接原因、经过、责任，评估安全事件造成的影响和损失，总结安全事件防范和应急处置工作的经验教训，提出处理意见和改进措施，形成安全事件处置报告，报告经安全管理机构审批后向主管部门报告。

7.4.3 事件归因分析

大数据服务提供者应：
a) 制定安全事件归因数据溯源策略和机制，以及归因溯源数据安全存储与使用的管理制度；
b) 跟踪和记录数据处理活动及其相关的数据服务，确保能对数据处理活动事件进行溯源，包括对记录的使用、提供和公开等数据处理活动进行细粒度安全审计；
c) 支持根据时间顺序、攻击来源、攻击对象等对数据处理活动和大数据服务中的攻击行为进行关联分析，提高数据安全和供应链安全事件溯源分析能力；
d) 采用数字水印、数据标记、可信存证或血缘分析等技术保障数据处理过程的可追溯性；
e) 对归因溯源数据进行备份或归档,并采取技术手段对重要数据和个人信息处理活动的溯源数据进行安全保护；
f) 采取技术机制和管控措施保证事件归因溯源数据的完整性,确保通过溯源数据能重现数据处理活动的全过程；
g) 建立基于溯源数据的数据业务合规性审核机制,并依据审核结果改进大数据服务相关的访问控制、合规性保障等数据处理活动安全策略及其安全技术机制。

7.5 安全恢复

7.5.1 数据恢复

大数据服务提供者应：
a) 识别需要备份和归档的大数据服务数据，制定数据备份、归档与恢复的计划；
b) 配置必要的数据备份、归档与恢复工具，采用技术手段记录数据备份、归档与恢复过程，确保数据备份、归档与恢复过程可溯源；

c) 定期对备份和归档数据的可用性、完整性和一致性进行检测，对数据恢复的安全风险进行分析，具备对数据恢复情况的检测能力；
d) 定期测试数据备份、归档和恢复安全管控措施，具备对备份和归档数据的恢复重建能力；
e) 建立数据备份、归档与恢复团队，结合业务需求和安全目标，定期或按需开展数据恢复实战演练。

7.5.2 业务恢复

大数据服务提供者应：

a) 建立大数据系统容灾备份和灾难恢复的管理制度，明确系统在建设规划、运行维护、应急响应和灾难恢复中需要满足的业务连续性的安全运行要求；
b) 制定大数据系统灾难恢复预案，明确大数据系统灾难恢复的范围和目标、灾难切换规程和操作手册、灾后重建运行操作指南，定期组织灾难恢复预案的培训和演练；
c) 建立灾难恢复运行监控平台，具备容灾备份系统的有效性验证能力，及时发现容灾备份系统的运行故障；
d) 开展大数据系统灾难恢复演练，根据演练情况修订灾难恢复预案等业务连续性计划，确保系统容灾备份和灾难恢复预案的有效性；
e) 建立异地容灾备份和恢复操作规范，配置系统容灾备份和灾难恢复专业团队，确保灾备中心具备按照业务系统重要性和其影响面要求及时接管大数据系统的数据处理活动的能力。

附 录 A
（资料性）
大数据服务模式和大数据用户角色

A.1 大数据服务模式

在互联网环境下，基于分布式计算和大数据技术的大数据服务存在数据自营模式、数据租售模式、数据中台模式、数据众包模式、数据外包模式、虚拟数仓模式、数据湖模式等多种商业模式。不同商业模式下对大数据平台的基础设施资源选择、计算资源调度与管理、支持数据处理活动的大数据应用服务组件的部署等控制范围不同，大数据服务提供者在为大数据使用者提供大数据服务时应具备的大数据系统安全能力也不同。

大数据系统主要由大数据平台和大数据应用两部分组成。大数据平台对多种数据源进行聚合和融合、分布存储、并行计算和分析处理，使用多种协议标准化各大数据应用之间的数据接口、编程接口和数据提供者接口之间的映射，封装不同类型数据实体操作，以及对数据处理过程中的异常进行安全处理，使大数据使用者可以高效透明地访问或使用大数据系统中的多源数据，以通用的、可互操作的、灵活的使用模式管理这些海量、异构、快速变化的数据资源。大数据应用为数据提供者和大数据使用者提供数据收集、存储、使用、加工、传输、提供、公开，直至销毁等数据处理功能。

按照信息基础设施、大数据技术框架与大数据应用支撑的信息技术层次结构，大数据平台服务可细分为通用计算资源/计算设施服务、大数据平台服务和大数据应用支撑服务（图A.1所示）。通用计算资源/设施服务一般由计算硬件或虚拟机计算资源、网络通信资源和数据存储资源等组成，它为大数据平台与大数据应用服务提供可伸缩的计算、通信和存储基础设施；大数据平台服务主要为海量、复杂、异构、动态变化的大数据分析及各种智能数据处理算法提供高效可扩展的数据存储和数据处理服务；大数据应用支撑服务主要指通过集成领域业务和数据模型、数据挖掘与分析套件、大数据可视化套件、集群自动化调度、面向应用的数据处理活动等大数据应用领域相关的服务组件和开发接口，用以简化大数据应用开发和部署。

从大数据平台基础设施、数据计算和分析能力、应用支撑接口支持等业务模式和部署实施方式等维度，可以将大数据平台提供者的大数据平台服务分为四种类型：

——**核心大数据服务**：大数据平台提供者采用虚拟化技术、云计算技术、数据仓库技术或数据湖技术，向大数据应用提供可扩展存储结构、分布式计算、内存计算等通用服务能力，以及数据存储管理和数据快速交互式分析处理能力。核心大数据服务提供关系、键值、文档、半结构化、文本、流数据等多种结构化数据，以及图像、音频、视频等类型的非结构化数据的组织和访问服务，提供面向海量数据的分布式数据存储服务和可伸缩的数据并行处理和分析与可视化服务，并具备与其它开源的通用基础设施、分布式计算资源进行数据存储和计算交互的能力，提供大数据应用开发接口和支持工具，以支持大数据应用提供者通过这些开发接口组合使用核心大数据服务功能来构建和部署他们所需的大数据应用服务。

——**高性能大数据服务**：面向批量数据分析、流式数据分析、海量数据联机事务处理等高可靠、高性能、高可用、可伸缩大数据存储和计算服务需求，在核心大数据服务基础上通过向下集成核心大数据服务所需的服务器、存储与网络设备、虚拟化软件等通用基础设施和计算资源，减少大数据服务基础设施部署和运维管理复杂度，简化大数据服务过程中的性能故障诊断及服务优化等问题。具备高性能大数据服务的大数据平台提供者一般都为大数据使用者提供集成的服务

器、存储设备、操作系统、虚拟化管理软件、数据管理系统以及一些为数据查询、处理、分析用途而预先安装的数据服务组件，并提供应用编程接口、数据访问服务等应用开发环境和系统健康监测服务，大数据应用开发者需使用这些个性化的编程接口与服务组件开发相应的大数据应用程序。

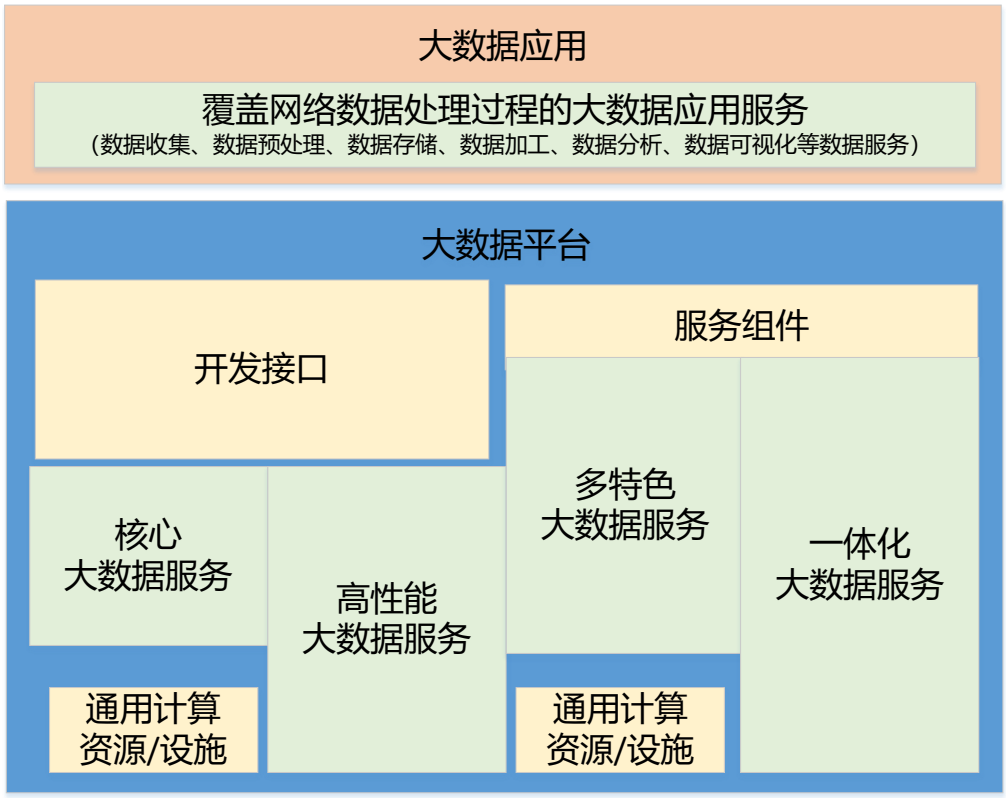

图 A.1　大数据服务类型和大数据系统组成结构

——**多特色大数据服务：**面向金融、电信、能源、交通、电子政务等不同领域大数据服务共性需求，在核心大数据服务基础上通过向上集成领域相关的大数据分析与挖掘算法、业务数据模型、数据处理过程相关数据服务等面向应用的多种特征的数据业务服务，并提供包括应用编程接口、数据存储适配器等面向应用领域增强的特色大数据服务组件和构件。提供这种特色服务的大数据平台提供者一般在核心大数据服务中集成了大数据应用相关的数据服务基础功能，即提供支持大数据应用、具备领域特征的大数据建模、管理、处理和分析服务，使大数据应用提供者可借助这些特色大数据服务组件快速构建其大数据应用，启用大数据领域相关的特色服务。

——**一体化大数据服务：**大数据平台提供者在核心大数据服务基础上向下集成可扩展的大数据基础设施和通用计算资源，向上集成面向应用领域的数据采集、数据存储、数据分析、数据可视化等多特色大数据应用服务组件，一体化大数据服务平台提供者将核心大数据服务拓展为性能好、宜部署、大数据平台与大数据应用基础功能特色兼备的大数据存储和计算平台服务，为大数据使用者提供可扩展和完整的一站式大数据平台与应用支撑服务。

大数据服务提供者应通过大数据系统冗余和数据复制、备份等高可用解决方案，保证大数据服务水平协议的实现。例如基于云计算技术的大数据基础设施服务应参照 GB/T 31168《信息安全技术 云计算服务安全能力要求》等有关信息安全标准规范落实相关的网络安全责任，确保大数据服务的可扩展、可伸缩等可用性需求。此外，大数据服务提供者应该部署相关的服务安全管控组件，实时地监控大数据服务中数据主体、数据拥有者、大数据使用者及系统服务组件对大数据处理过程的各种属性，以保证实现

大数据服务安全目标。

大数据服务提供者应依据其大数据应用服务目标和支撑大数据服务的大数据平台应担当的角色和责任，选择大数据平台相应的服务类型，按照GB/T 22239-2019《信息安全技术 网络安全等级保护基本要求》等网络安全保护制度相关标准制定相应的安全策略和安全管理制度，部署合适的安全防护措施，并从数据服务安全风险管理角度识别大数据服务安全能力现状，分析与大数据安全目标的差距，选择适合组织数据处理活动业务和大数据服务的安全控制措施；在此基础上通过安全监测等技术措施持续改进大数据安全控制措施和制定大数据服务安全能力提升计划，可持续的保证大数据服务安全目标。

A.2 大数据用户角色

A.2.1 角色分类

本文件参照GB/T 35589—2017，将大数据用户分为数据提供者、大数据使用者、大数据平台提供者、大数据应用提供者和大数据服务协调者五种角色，分别对应于技术参考模型的数据提供者、数据消费者、大数据框架提供者、大数据应用提供者和系统协调者。大数据服务提供者包括大数据平台提供者、大数据应用提供者和大数据服务协调者三种角色。本文件主要从大数据服务提供者的组织管理安全能力、数据处理活动安全能力和数据服务风险管理安全能力三方面提出了他们在组织层面、业务层面和风险管理层面的网络安全能力要求。

A.2.2 数据提供者

数据提供者将组织外部公共网络数据资源、外部合作企业私有数据资源、内部数据资源或大数据服务提供者系统运行过程中的各种日志、事件等系统行为数据资源进行抽象和建模，按照国家和行业数据安全标准与规范对这些数据源数据进行采集和整合后引入到大数据平台中，供大数据平台和大数据应用发现、访问、转换和分析这些数据资源。依据数据来源不同，数据提供者可进一步分为外部公共数据资源提供者、外部组织私有数据提供者、内部数据提供者、机器或系统数据提供者等数据提供角色。

A.2.3 大数据平台提供者

大数据平台提供者提供必要的网络、计算、存储等大数据服务所需的IT运行环境资源，和必要的基础设施应用程序开发接口或服务组件，以支持大数据组织、存储、分析和基础设施部署和运维管理，响应大数据应用提供者提出的大数据服务请求。大数据平台提供者应通过大数据平台提供的身份标识与鉴别、授权与访问控制、密文处理与密钥管理、安全审计与数据溯源等安全功能保护所处理数据的保密性和完整性、数据处理活动的真实性、数据处理过程中个人信息保护等数据安全目标。鉴于大数据复杂性、多样性、快速变化等特点，为上层大数据应用提供分布式、可扩展的数据存储管理和分布式并行数据处理服务是大数据平台提供者的核心目标。大数据平台提供者需要通过大数据平台实现对大数据服务相关存储资源和计算资源的高效管理和有效调度。因此，大数据平台提供者可进一步分为大数据基础设施服务提供者、大数据存储管理服务提供者和大数据应用支撑服务提供者。

A.2.4 大数据应用提供者

大数据应用提供者将整合的大数据资源及其应用组件以软件服务方式部署到大数据平台上，并通过应用终端安全接入、数据分类分级、输入数据验证等安全策略配置和安全控制措施，给大数据使用者提供安全的数据组织、存储、加工、分析和可视化服务。大数据应用提供者可分为数据、技术和服务三种产业链角色，其服务安全能力依赖于其商业模式所处的角色和义务，需要依据数据处理过程各阶段中所

负责的数据处理活动进行定义，确保满足数据安全和个人隐私保护需求。大数据应用提供者应采用机器学习、数据挖掘等人工智能技术帮助大数据使用者开展诸如精准营销等各种分析与咨询服务。

A.2.5 大数据服务协调者

大数据服务协调者规范和集成组织大数据服务所需的大数据平台和各类大数据应用数据业务活动，配置和管理大数据平台和大数据应用支撑安全功能组件，以构建一个安全风险可控的大数据服务生态系统，确保大数据应用的各项数据服务能在大数据平台上安全高效地正确运行。大数据服务协调者负责为大数据服务算力分配对应的物理或虚拟节点，调度数据处理活动所需要的计算和存储资源，确保为大数据使用者提供的大数据服务质量达到服务水平协议要求，并通过算力资源的自动化按需调度和数据供应链的按需协调，保证大数据服务可用性；通过不同的安全技术手段和安全措施，构筑大数据服务安全防护体系，实现覆盖硬件、软件和上层应用的算力、模型和数据的安全保护；按照网络安全等级保护制度要求，从安全物理环境、安全通信网络、安全区域边界、安全计算环境等方面来保证大数据服务平台的安全性；依照业务连续性等相关要求通过配置合理的大数据平台和数据容灾框架，提升大数据服务资源灾备和恢复能力。

A.2.6 大数据使用者

大数据使用者可以是一个真实的终端用户或组织机构，也可以是一个信息技术系统。它使用大数据平台提供者或大数据应用提供者提供的数据和服务，实现其业务目标。大数据应用和大数据平台提供给大数据使用者的数据应经过数据合规性控制、数据脱敏等安全控制措施，并通过授权和访问控制，以保证数据保密性、完整性和可用性。大数据使用者的数据服务安全要求一般通过与大数据服务提供者的服务契约和服务水平协议体现，大数据服务水平协议中应规范性描述大数据应用服务各方面的安全属性，包括输入/输出等数据完整性和保密性属性，服务安全约束和响应时间等服务质量约束，以及在数据业务层面的诸多服务质量属性，如涉及的业务规则、数据依赖关系、时间/人员消耗可用性等。服务水平协议中还要规范描述大数据服务参与方相关的关系，如服务间依赖关系、数据服务和数据资源约束关系、数据服务和应用组件间关系、服务消息间关系等。

参 考 文 献

[1] GB/T 22081—2016 信息技术 安全技术 信息安全控制实践指南
[2] GB/T 24353—2009 风险管理 原则与实施指南
[3] GB/T 35589—2017 信息技术 大数据 技术参考模型
[4] GB/T 39786—2021 信息安全技术 信息系统密码应用基本要求
[5] 中华人民共和国全国人民代表大会常务委员会,中华人民共和国网络安全法，2016年11月7日。
[6] 中华人民共和国全国人民代表大会常务委员会,中华人民共和国数据安全法，2021年6月10日。
[7] 中华人民共和国全国人民代表大会常务委员会，中华人民共和国个人信息保护法， 2021年8月20日。
[8] 国务院关于印发促进大数据发展行动纲要的通知，国发〔2015〕50号，2015年9月5日。
[9] 国务院办公厅关于运用大数据加强对市场主体服务和监管的若干意见，国发〔2015〕51号，2015年7月1日。
[10] 国家互联网信息办公室,《网络安全审查办法》，2021年12月28日。
[11] 国家互联网信息办公室,《网络数据安全管理条例》（征求意见稿），2021年11月14日。
[12] ISO/IEC 27002 Information security, cybersecurity and privacy protection — Information security controls
[13] NIST Special Publication 800-53 Security and Privacy Controls for Information Systems and Organizations