国 际 电 信 联 盟

ITU-T X.1751

(09/2020)

国际电信联盟
电信标准化部门

X系列：数据网、开放系统通信和安全性

数据安全 – 大数据安全

电信运营商大数据生命周期管理安全导则

ITU-T X.1751 建议书

ITU-T X 系列建议书

数据网、开放系统通信和安全性

公用数据网	X.1–X.199
开放系统互连	X.200–X.299
网间互通	X.300–X.399
消息处理系统	X.400–X.499
号码簿	X.500–X.599
OSI组网和系统概貌	X.600–X.699
OSI管理	X.700–X.799
安全	X.800–X.849
OSI应用	X.850–X.899
开放分布式处理	X.900–X.999
信息和网络安全	
一般安全问题	X.1000–X.1029
网络安全	X.1030–X.1049
安全管理	X.1050–X.1069
生物测定	X.1080–X.1099
安全应用和服务（1）	
组播安全	X.1100–X.1109
家庭网络安全	X.1110–X.1119
移动安全	X.1120–X.1139
网页安全	X.1140–X.1149
安全协议（1）	X.1150–X.1159
对等网络安全	X.1160–X.1169
网络身份安全	X.1170–X.1179
IPTV安全	X.1180–X.1199
网络空间安全	
网络安全	X.1200–X.1229
反垃圾信息	X.1230–X.1249
身份管理	X.1250–X.1279
安全应用和服务（2）	
应急通信	X.1300–X.1309
泛在传感器网络安全	X.1310–X.1319
智能电网安全	X.1330–X.1339
验证邮件	X.1340–X.1349
物联网（IoT）安全	X.1360–X.1369
智能交通系统（ITS）安全	X.1370–X.1389
分布式账簿技术安全	X.1400–X.1429
分布式账簿技术安全	X.1430–X.1449
安全协议（2）	X.1450–X.1459
网络安全信息交换	
网络安全概述	X.1500–X.1519
漏洞/状态信息交换	X.1520–X.1539
事件/事故/启发式信息交换	X.1540–X.1549
政策的交换	X.1550–X.1559
启发式和信息请求	X.1560–X.1569
标识和发现	X.1570–X.1579
确保交换	X.1580–X.1589
云计算安全	
云计算安全概述	X.1600–X.1601
云计算安全设计	X.1602–X.1639
云计算安全最佳做法和指导原则	X.1640–X.1659
云计算安全实施方案	X.1660–X.1679
其他云计算安全	X.1680–X.1699
量子通信	
术语	X.1700–X.1701
量子随机数发生器	X.1702–X.1709
QKDN安全框架	X.1710–X.1711
QKDN安全设计	X.1712–X.1719
QKDN安全技术	X.1720–X.1729
数据安全	
大数据安全	**X.1750–X.1759**
5G 安全	X.1800–X.1819

欲了解更详细信息，请查阅ITU-T建议书目录。

ITU-T X.1751 建议书

电信运营商大数据生命周期管理安全导则

摘要

ITU-T X.1751建议书对安全漏洞做了分析，并建立了电信运营商大数据生命周期管理安全导则。

随着大数据技术的快速发展，数据的价值大幅提升。大数据给电信业务带来新的机遇。以前，数据在不同的电信业务系统中是孤立和独立管理的。随着大数据业务的建设，数据聚合和融合趋势不可避免。在数据融合聚集过程中，数据在平台和业务流程中流动。数据在其生命周期的不同阶段面临各种安全漏洞。

ITU-T X.1751建议书介绍电信大数据业务的具体特性和数据类别，分析大数据生命周期管理的安全漏洞，并为电信运营商具体提出安全导则。

历史沿革

版本	建议书	批准日期	研究组	唯一识别码*
1.0	ITU-T X.1751	2020-09-03	17	11.1002/1000/14267

关键词

数据生命周期管理，电信大数据业务。

* 欲查阅建议书，请在您的网络浏览器地址域键入URL http://handle.itu.int/，随后输入建议书的唯一ID，例如，http://handle.itu.int/11.1002/1000/11830-en。

前言

国际电信联盟（ITU）是从事电信、信息和通信技术（ICT）领域工作的联合国专门机构。国际电信联盟电信标准化部门（ITU-T）是国际电信联盟的常设机构，负责研究技术、操作和资费问题，并且为在世界范围内实现电信标准化，发表有关上述研究项目的建议书。

每四年一届的世界电信标准化全会（WTSA）确定ITU-T各研究组的研究课题，再由各研究组制定有关这些课题的建议书。

WTSA第1号决议规定了批准建议书须遵循的程序。

属ITU-T研究范围的某些信息技术领域的必要标准，是与国际标准化组织（ISO）和国际电工技术委员会（IEC）合作制定的。

注

本建议书为简明扼要起见而使用的“主管部门”一词，既指电信主管部门，又指经认可的运营机构。

遵守本建议书的规定是以自愿为基础的，但建议书可能包含某些强制性条款（以确保例如互操作性或适用性等），只有满足所有强制性条款的规定，才能达到遵守建议书的目的。“应该”或“必须”等其它一些强制性用语及其否定形式被用于表达特定要求。使用此类用语不表示要求任何一方遵守本建议书。

知识产权

国际电联提请注意：本建议书的应用或实施可能涉及使用已申报的知识产权。国际电联对无论是其成员还是建议书制定程序之外的其它机构提出的有关已申报的知识产权的证据、有效性或适用性不表示意见。

至本建议书批准之日止，国际电联已收到实施本建议书可能需要的受专利保护的知识产权的通知。但需要提醒实施者注意的是，这可能并非最新信息，因此大力提倡他们通过下列网址查询电信标准化局（TSB）的专利数据库：http://www.itu.int/ITU-T/ipr/。

目录

ITU-T X.1751 建议书

电信运营商大数据生命周期管理安全导则

1 范围

本建议书描述安全漏洞，并建立电信大数据业务生命周期管理导则。本建议书：

– 介绍电信大数据业务特性和数据类别；

– 分析电信大数据业务生命周期管理的安全漏洞；

– 为电信大数据业务具体提出数据生命周期管理的安全导则。

当电信运营商提供大数据业务时，基本前提是已获得用户的明确同意。此外，建议在整个大数据业务过程中为电信运营商提供必要的数据保护措施。

各种数据类别的保护机制不在本建议书的讨论范围内。

2 参考文献

下列ITU-T建议书和其他参考文献的条款，通过在本建议书中的引用而构成本建议书的条款。在出版时，所指出的版本是有效的。所有的建议书和其他参考文献都面临修订，使用本建议书的各方应探讨使用下列建议书和其他参考文献最新版本的可能性。当前有效的ITU-T建议书清单定期出版。本建议书中引用某个独立文件，并非确定该文件具备建议书的地位。

[ITU-T X.1641] ITU-T X.1641建议书（2016年），*云业务客户数据安全导则*。

3 定义

3.1 他处定义的术语

本建议书使用了下列他处定义的术语：

3.1.1 大数据big data [b-ITU-T Y.3600]：一种可能需要在实时约束条件下实现对具有异构特性的海量数据集收集、存储、管理、分析和可视化的范式。

3.1.2 大数据即服务big data as a service (BDaaS) [b-ITU-T Y.3600]：一种云业务类别，当中向云业务客户提供的能力为使用大数据进行收集、存储、分析、可视化和管理数据的能力。

3.1.3 可链接性linkability [b-ISO/IEC 20889]: 一个数据集属性，可通过链接，将一个数据主体的记录与另一个数据集中同一数据主体的记录联系起来。

3.1.4 假名化pseudonymization [b-ISO/IEC 29100]：应用于个人可识别信息（PII）的过程，该过程用一个别名来替代可说明身份的信息。

3.1.5 安全策略security policy [b-ITU-T X.800]：用于提供安全业务的准则集。

3.2 本建议书定义的术语

本建议书定义了下列术语：

3.2.1 数据生命周期data lifecycle：数据生成后的整个生存过程，包括数据收集、数据传输、数据存储、数据使用（涵盖数据分析和可视化）、数据共享和数据销毁。

4 缩写词和首字母缩略语

本建议书使用了下列缩写词和首字母缩略语：

2B 至企业
2C 至消费者
API 应用程序编程接口
APP 应用程序
BDaaS 大数据即服务
BSS/OSS 业务支持系统和运营支持系统
DB 数据库
FTP 文件传输协议
HDFS Hadoop分布式文件系统
IoT 物联网
IP 网际协议
JDBC Java数据库连接
LBS 基于位置的业务
LDAP 轻量目录访问协议
MPP 大规模并行处理器
OSS 运营支持系统
PII 个人可识别信息
REST 表示层状态传输

5 惯例

在本建议书中：

“要求”（is required to）一词指的是一项必须严格遵守的要求，如果宣称遵循本建议书，则不得违反。

“建议”（is recommended）一词指的是一项建议性的、并非绝对需遵守的要求。因此，宣称遵循本建议书时无需提及该项要求。

“禁止”（is prohibited from）一词指的是一项必须严格遵循的要求，如果宣称遵循本建议书，则不得违反。

“可选”（can optionally）一词指的是一项允许的可选要求，不隐含任何建议的意味。本术语无意暗示供应商的实施方案必须提供选项，以及网络运营商或服务提供商可以选择启用该功能。相反地，本术语意味着供应商可以选择提供该功能，并仍宣称遵循本规范。

6 概述

随着其快速发展，大数据的价值得到了巨大提升。大数据给电信运营商带来了新的机遇。电信运营商在所谓的数据仓库系统中长期保存各种不同的数据资源，例如，呼叫、位置、个人数据、移动消费者数据和终端数据。随着大数据的快速生成，电信运营商正在不断创新并投资于发展大数据业务。

电信大数据业务涉及TB级甚至PB级的信息量。生成的数据有多种类型，例如，结构化的、半结构化的和非结构化的。这些数据来源包括私人数据（例如PII）和访问日志数据。攻击者可以将此类数据作为攻击目标。

以前，数据分散于不同的电信业务系统中并独立进行管理。此外，这些系统可位于各种不同的位置上，并由不同的部门来管理。随着大数据业务的发展，电信运营商正在打破部门壁垒，并从各种独立的系统中收集数据。数据融合极大地提高了大数据业务的价值。

在一个大数据业务流程中，数据流经大数据平台并经过各个生命周期阶段，在每个生命周期阶段，信息都面临各种安全威胁和风险。例如，不当的数据收集可导致不当的信息披露。在数据存储阶段，可发生未经授权的访问。在共享信息时，使用敏感数据可造成数据泄漏的风险。因此，有必要对安全漏洞进行分析，并规定电信大数据业务生命周期管理安全导则。

本建议书对安全风险进行了分析，并规定了电信大数据的生命周期管理安全导则。

7 电信大数据业务特性和数据类别

7.1 电信大数据业务特性

越来越多的电信运营商将大数据业务作为其公司创新与发展的重要战略方向。例如，通过构建大数据能力平台或建立专门的运营团队，电信运营商可以推动大数据业务的发展。

电信运营商可以收集与单个用户有关的大量客户数据，例如，用户配置文件、设备、使用情况和位置等。电信运营商可以使用大数据分析技术来利用这些数据开发与各种各样的应用程序有关的业务，例如，零售、医疗保健和智慧城市。电信运营商可以使用这些业务来增强自己的商业能力，也可以将其推销给其他业务领域的第三方服务提供商。

不过，电信运营商用于这些大数据业务的数据广度和类型可以揭示有关个人的惊人数量的详细信息，包括PII、敏感数据（例如宗教信仰或政治背景）和商业秘密。如果电信运营商选择与第三方共享这些数据，则对这一点的考虑就显得尤为重要。因此，电信运营商意识到在其大数据业务的数据生命周期中可能面临威胁并采取安全措施对其用户实施保护就显得非常重要。

7.2 数据类别

在下面的段落中列出了电信运营商手中用户数据的四个主要类别。此外，随着物联网（IoT）及其业务的发展，数据的深度和广度正在进一步扩大。这种扩大给用户信心和安全带来新的风险，电信运营商必须在以下方面加以解决：

1) 从电信运营商的业务支持系统和运营支持系统（BSS/OSS）生成的数据，包括用户身份、通话时间、通话目标、通信账单、业务类型甚至终端类型；

2) 从电信运营商的运营支持系统（OSS）生成的数据，主要是用户行为数据，包括通过移动互联网、聊天、游戏和上网产生的数据；

3) 基于用户之基于位置的服务（LBS）信息的数据，与类别1)和类别2)不同，该数据与用户的实际位置紧密相关，可用于企业营销、人口流动、公共安全和城市规划；

4) 在IoT场景中生成的至业务（2B）或至消费者（2C）数据，包括有关“物”和“人”的大数据 – 这些数据在医疗保健、可穿戴设备和智能家居领域具有重要价值。

有关事物的大数据包括：从仪表收集的水、电、气读数；从传感器收集的气候和污染数据；以及有关资产装运的跟踪数据。

关于人的大数据包括：关于健康的数据；个人财务数据；以及采购历史数据。

注意：根据数据最初的分类方式，某些数据的可链接性可能会给用户带来比预期更大的影响。例如，结合其他数据对其进行分析的匿名化的类别4)数据仍可使个人是可识别的，例如，类别1)中的PII。因此，即使在没有立即将数据归类为PII或其他个人数据的情况下，在确定如何保护数据安全时，也建议将可链接性作为一个关键考虑因素。

在本建议书中，建议依据安全导则在数据生命周期的所有阶段上对这四类数据进行安全化。

8 电信大数据业务中的数据生命周期

电信大数据业务的数据生命周期主要由六个阶段组成：数据收集；数据传输；数据存储；数据使用；数据共享；数据销毁。数据传输阶段可以涉及若干阶段。

电信大数据业务中的数据生命周期如图1所示。

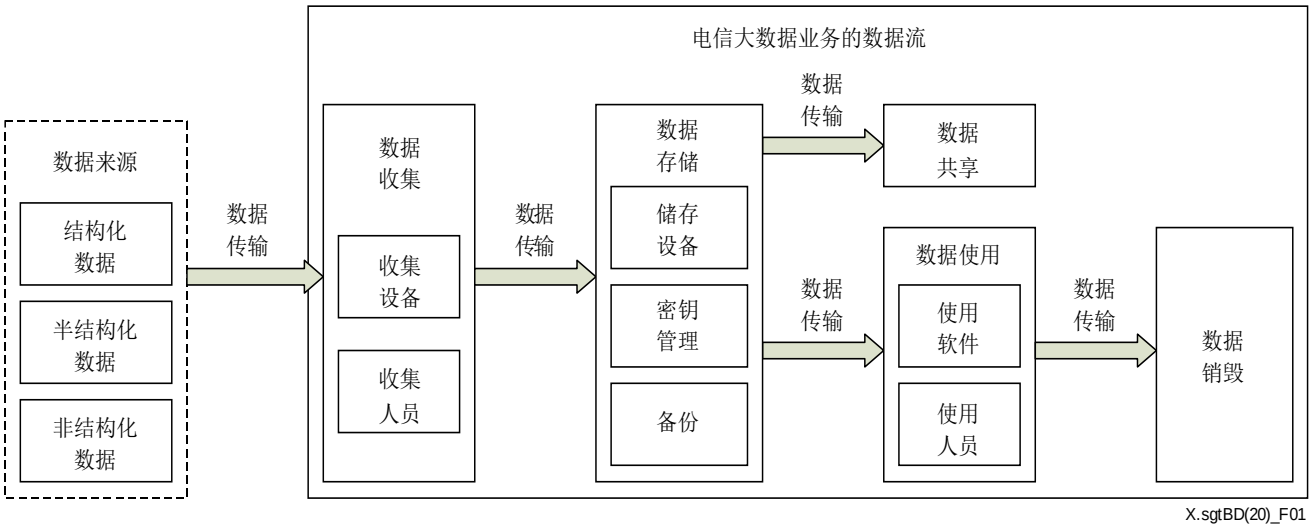

图1 – 电信大数据业务数据的数据生命周期

在电信大数据业务的生命周期中，数据收集是开始，数据销毁是结束。在收集后，可以传输、存储、使用和共享数据。数据传输可以发生在不同阶段之间，例如，在收集后，可以

将数据传输给专门的存储设备；在使用中，可以将数据从存储设备传输给使用软件或实体；在使用后，要求销毁过期或无用的数据。

数据收集：使用专门的数据收集设备或实体，将不同类型和类别的数据收集到一个特定的目录或临时的目录中进行存储。

数据传输：该过程涉及从收集设备到存储设备以及从存储到使用和共享的数据传输。有时，数据销毁阶段也涉及数据传输。

数据存储：数据存储在专用的设备中，例如，数据库（DB）、分布式文件系统和磁盘阵列。对重要数据，还需要进行加密和数据备份。

数据使用：数据分析软件和应用程序访问并处理数据，以提供各种各样的大数据业务。在该过程中可涉及各种各样的个人敏感数据。

数据共享：数据服务提供商或数据所有者与其他提供商或第三方共享自己的数据分析处理结果，或甚至是源数据。

数据销毁：对过期的数据、尤其是存储重要或敏感信息的设备中的那些数据，要求通过专门针对此目的的安全机制来完全销毁。

9 电信大数据业务数据生命周期中的安全漏洞

在电信大数据业务过程中，数据流向业务的每一步。安全漏洞可来自流程内或流程外。内部安全漏洞与设备和系统有关，例如，收集设备、存储设备和使用设备。会因配置错误或使用不当而导致外部安全漏洞。与数据有关的漏洞的描述可参见[b-ITU-T X.1040]。

9.1 数据收集阶段

9.1.1 设备和系统的安全漏洞

设备和系统易受病毒或木马攻击或感染，从而导致安全问题。

9.1.2 配置和管理的安全漏洞

1) 人员管理

数据收集设备可能会被不当操作或未经授权使用，而导致数据泄漏的风险。

2) 设备管理

收集设备的恶意或错误配置可导致未经授权的收集和数据泄露。

数据从位于不同位置并由不同部门管理的许多不同业务系统中累积。恶意节点可闯入设备群集以执行未经授权的数据收集操作。

3) 数据管理

过度收集与所声明之使用目的无关的数据，可能会导致数据泄漏。

在收集阶段，临时的数据存储区（例如，文件传输协议（FTP）服务器的随机路径）不受收集人员的控制，这会增加数据泄露或未经授权更改的风险。

9.2 数据传输阶段

9.2.1 配置和管理的安全漏洞

1) 人员管理

管理员未经授权修改传输端口或传输配置，可能会导致数据泄漏。

2) 设备管理

收集的数据通过不安全的信道进行传输，将使数据易于遭受窃听或更改。

传输接口未经认证，因此会发生错误连接或恶意连接。

不同节点之间的传输机制不能保护数据的机密性和完整性，这会导致数据泄漏并增加更改、窃听和拦截的风险。

3) 数据管理

数据传输过程中缺乏机密性保护可能会导致数据被拦截或数据泄漏。

数据传输过程中缺乏完整性保护可能会导致恶意操作或数据破坏。

数据传输过程中缺乏可用性保护可能会导致数据被拦截或篡改。

9.3 数据存储阶段

9.3.1 设备和系统的安全漏洞

没有部署防病毒软件或过期的安全软件会导致敏感数据的泄露和中毒。

错误配置或未维护的安全软件会使数据得不到保护，这可导致数据集的泄露、中毒和敌方攻击。

9.3.2 配置和管理的安全漏洞

1) 人员管理

访问控制措施配置不当，会使存储的数据暴露于未经授权访问或更改的风险下。由于可链接性，被授权访问某些存储的数据集的个人能够推断出未经授权进行访问的PII或其他个人数据。

2) 设备管理

关系数据库、Hadoop分布式文件系统（HDFS）和大规模并行处理器（MPP）数据仓库共存于一个存储环境中。权限管理不善会导致未经授权的访问和数据泄露。

许多非结构化的数据分别存储在不同的节点中，这使得难以实施相同的安全策略。安全策略不一致或冲突可导致安全保护的缺失和数据泄漏。

3) 数据管理

不对存储的数据进行全部或部分加密，这在发生被动或主动数据泄露的情况下会增加用户暴露的风险。

对存储的数据进行加密；不过，与加密算法或密钥管理有关的漏洞会增加存储的数据被未经授权访问或更改的风险。

存储期间未对数据的完整性采取积极主动的保护措施，那么即使在数据收集后，也会使数据暴露于未经授权修改的风险下。

存储的数据包括已过时的数据或者与声明之使用目的不直接相关和不十分必要的数据。随着存储数据数量的增加，在发生任何被动或主动数据泄露的情况下，用户暴露的风险将会增加。

不完整的数据备份和恢复机制可能会导致数据不可用。

9.4 数据使用阶段

9.4.1 设备和系统的安全漏洞

网络化或基于云的数据分析软件存在漏洞，这些漏洞将使暴露于该软件的数据处于危险中。

数据分析软件的可视化应用程序存在可被攻击者利用的安全问题。

9.4.2 配置和管理的安全漏洞

1) 人员管理

如果用户或可视化应用程序缺少认证和授权功能，则伪造的用户或可视化应用程序可以获得可视数据，而导致数据泄漏。

2) 设备管理

缺少对数据使用设备的身份认证会导致未经授权的设备加入大数据平台。后果是，会出现数据泄漏或业务不可用问题。

3) 数据管理

当由可视化应用程序使用时，包括任何PII和敏感数据在内的个人数据都无法适当匿名或掩盖，这可能增加了可识别性。

可视化数据是高度可链接的，这增加了推断身份或其他个人数据的可能性。

数据分析过程产生的输出数据未被存储在一个安全环境中，或者保留的是已过时或与使用目的不直接相关或不必要的输出数据。随着所保留数据的数量增加，在发生任何被动或主动数据泄露的情况下，用户暴露的风险也会增加。

在该阶段生成的数据日志没有被安全地存储，在发生数据泄露的情况下，这可能被用来推断个人数据。

9.5 数据共享阶段

9.5.1 设备和系统的安全漏洞

数据共享设备和系统易受病毒或木马攻击或感染，从而导致安全问题。

9.5.2 配置和管理的安全漏洞

1) 人员管理

共享数据用户的安全职责和能力指定不当，导致数据保护不当和数据泄漏。

2) 设备管理

数据共享的范围和边界不明确；此外，缺乏安全控制措施，导致敏感数据直接暴露于合作伙伴。

数据共享信道缺乏安全保护，而导致泄露或篡改重要数据。

3) 数据管理

数据共享日志记录不完整，则一旦发生安全事件，就很难确定原因。

9.6 数据销毁阶段

在该阶段结束时，如果尚未完全销毁数据，则敏感数据可能会恶意人员恢复，而导致数据泄漏。

9.7 安全漏洞与数据生命周期的关系

安全漏洞出现在大数据业务的不同生命周期阶段。表1概述了安全漏洞与大数据业务的数据生命周期之间的关系。

在表1中，单元格中的字母“Y”（是）表示该阶段存在安全漏洞。

表1 – 安全漏洞与数据生命周期之间的关系

漏洞		生命周期阶段					
		数据收集	数据传输	数据存储	数据使用	数据共享	数据销毁
设备漏洞		Y		Y	Y	Y	
配置和管理漏洞	人员管理	Y	Y	Y	Y	Y	
	设备管理	Y	Y	Y	Y	Y	
	数据管理	Y	Y	Y	Y	Y	Y

10 电信大数据业务数据生命周期的安全导则

本建议书描述了详细的机制，该机制适用于第7.2节中描述的所有数据类别。

10.1 数据收集阶段

10.1.1 设备和系统的安全导则

要求设备和系统上用于数据收集的软件是最新的，且没有任何公开的漏洞。

要求用于数据收集的设备和系统需安装安全软件，例如，与设备操作系统兼容的用于抵御病毒和恶意软件的安全软件。

10.1.2 配置和管理的安全导则

1) 人员管理

建议告知用户收集哪些数据及其原因，并在收集开始之前获得用户的明确同意。

进行数据收集时，需要对收集人员进行认证和授权。

2) 设备管理

建议对数据收集的安全机制和对策做出规定，例如，对收集设备进行严格的认证和授权。建议数据类别应遵循收集原则。

要求设立大数据业务以实现预期的业务目的。要求确定与收集以实现使用目的密切相关的和必不可少的数据。

要求规定数据收集设备、收集信道、数据格式、收集过程和收集方法。

要求对收集设备和收集人员执行访问认证；要求提供对异常收集行为的检测和告警。

要求严格限制临时数据存储区，例如，禁止未经授权地将数据从该区导出给其他的存储资源，建议对更改存储区进行授权。

建议采用广泛使用并经可信第三方测试的加密算法，对收集的数据（包括元数据）的传输进行保护。

安全地管理和存储加密密钥，并在可能的情况下，优先采用具有前向保密性的加密协议。

3) 数据管理

建议根据其重要性和敏感性确定所收集数据的各种分类。

要求记录日志并对以下异常行为发出告警：

– 重复收集和传输超过设定的阈值时；
– 收集过程中传输被中断时；
– 超过设定的存储容量阈值时。

10.2 数据传输阶段

10.2.1 配置和管理的安全导则

1) 人员管理

防止管理员随意修改传输端口或接口配置参数。

2) 设备管理

要求通过端到端加密信道来进行数据传输。

传输接口提供认证功能，以防止发生恶意连接。

3) 数据管理

建议在数据传输阶段实施数据的机密性和完整性保护。

建议在传输过程中及时发现数据完整性受损；建议在发现错误后实施必要的措施来将其恢复。

10.3 数据存储阶段

10.3.1 设备和系统的安全导则

要求存储设备和系统上的软件是最新的，且没有任何公开的漏洞。

要求存储设备安装最新的安全软件。

10.3.2 配置和管理的安全导则

1) 人员管理

要求实施对用户或应用程序的访问控制，例如，密钥网络认证协议、Kerberos（麻省理工学院开发的一种安全认证系统）和细粒度授权。

建议将重要的数据操作集成到多人操作室控制模式中，以使单个人不能对重要的数据拥有完整的操作权限，例如，对其进行批输出、复制、销毁、发布和使用。

2) 设备管理

授权方法必须将访问限制在以下人员，即针对所收集的数据，由他/她来执行所声明之使用目的是必不可少的。必须设置权限以防止个人访问超出其履行特定之职责所绝对需要的更多数据，并考虑到因任何存储数据集之间的可链接性以及个人之间的单独权限而推断出个人数据的可能性。

3) 数据管理

a) 数据最小化

建议限制数据的存储，包括在数据使用阶段进行的处理结果输出，以便仅保留与声明的使用目的相关且必要的数据。

建议根据必须保留数据以达成其所声明之使用目的的最短可能时间，为所有数据设定明确的最长保留期限。

b) 数据加密存储

加密存储对确保重要数据的机密性而言是必要的。建议支持分层数据加密模型；根据数据保密等级，使用不同的安全存储机制。

建议采用被广泛部署并经可信第三方测试的加密算法。安全地管理和存储加密密钥，并在可能的情况下，优先采用具有前向保密性的加密协议。

c) 数据完整性保护

建议提供完整性检测机制，以确定因数据存储而造成的数据损坏和丢失。

建议在检测到错误后实施必要的措施，以恢复数据完整性[ITU-T X.1641]。

必须保留记录对存储数据进行任何更改的审计日志。必须安全地存储日志，并建议捕获和报告对其进行更改的任何尝试。

d) 数据备份与恢复

建议提供完整的数据备份和恢复机制，以确保数据的可用性和完整性。

10.4 数据使用阶段

10.4.1 设备和系统的安全导则

建议设备和系统上的软件是最新的，且没有任何公开的漏洞。

建议设备和系统安装最新的安全软件。

10.4.2 配置和管理的安全导则

1) 人员管理

无论应用程序使用哪种接口（例如，资源表示状态传输应用程序编程接口（REST API）和Java数据库连接性（JDBC）），建议提供统一的认证，以供不同的（数据使用）应用程序用来访问大数据平台。详细的认证机制可以是Kerberos（麻省理工学院开发的一种安全认证系统）、轻量目录访问协议（LDAP）或其他。

建议提供细粒度的授权，以供不同的应用程序用来访问大数据平台，包括以下方法：

a) 通过用户名、网际协议（IP）地址和应用程序（APP）名称的细粒度授权来访问大数据平台；

b) 细粒度授权来访问存储资源，例如，数据仓库软件Hive、开源的非相关分布式数据库Hbase、HDFS和数据库；

c) 通过数据库或文件系统的不同操作（例如，SELECT（选择）、INSERT（插入）和CREATE（创建））的细粒度授权；

d) 有关数据导入和导出权限的细粒度授权；

e) 有关访问HDFS文件和目录的细粒度授权。

2) 设备管理

建议在数据使用阶段提供恶意活动监视和执法机制。

建议提供安全设计跟踪，以测试数据使用是否适当，确保符合已建立的安全策略和操作程序，协助进行损害评估，并就安全控制、安全策略和数据使用程序方面的任何修改提出建议。

建议一种安全审计策略，以考虑要求记录哪些有关数据使用情况的信息、在什么条件下以及用于安全审计信息交换的语法与语义规范。

3) 数据管理

为了保护敏感数据，在数据使用阶段需要进行数据假名化。将对数据假名化的详细导则进行统一，并在随后的数据共享阶段中予以考虑。

建议对敏感数据的使用进行审计，并按[ITU-T X.1641]中的规定生成审计日志。

10.5 数据共享阶段

10.5.1 设备和系统的安全导则

建议用户设备和系统上的软件是最新的，且没有任何公开的漏洞。

建议设备和系统安装最新的安全软件。

10.5.2 配置和管理的安全导则

1) 人员管理

必须告知用户将与第三方共享哪些数据，包括元数据和作为数据使用阶段所执行过程之输出结果的任何数据，以及这些第三方指的是谁。在共享任何数据（包括输出数据）之前，必须征得用户的明确同意。

2) 设备管理

建议对数据导出行为进行控制。

当与外部业务共享数据时，建议限制数据的使用，以避免转售数据。

建议在相关的利益攸关方（例如，在其间传输数据的运营商）之间对安全保护机制进行协商，并纳入一个有关共享数据传输、存储、访问、销毁的安全策略以及一旦出现共享数据泄露时的备份方案。

3) 数据管理

数据假名化指的是用字符或数据隐藏原始PII和敏感数据的过程。目的是保护PII和敏感数据。

可以使用不同的假名化算法来配置不同的应用程序。建议通过数据假名化来：

a) 支持动态添加和删除假名化算法；

b) 支持细粒度的假名化，这意味着管理员可以为数据库的某个特定的表或列配置假名化；

c) 使用公共算法，以避免第三方专用算法；

d) 避免显著影响业务连续性和系统性能。

10.6 数据销毁阶段

10.6.1 数据管理的安全导则

数据必须在固态状态下擦除和覆盖。

在删除数据后，建议完全清除系统中的资源存储空间，例如，文件、目录和数据库记录，不要存在恢复的可能。

参考书目

[b-ISO/IEC 29100] ISO/IEC， Information technology – Security technique – Privacy framework，2011。

[b-ITU-T X.800] ITU-T X.800建议书（1991年），CCITT应用的开放系统互连（OSI）安全体系结构。

[b-ITU-T X.1040] ITU-T X.1040建议书（2017年），用于电子商务业务数据生命周期管理的安全参考架构。

[b-ITU-T Y.3600] ITU-T Y.3600建议书（2015年），大数据 – 基于云计算的要求及能力。

[b-ISO/IEC 20889] ISO/IEC 20889:2018, Privacy enhancing data de-identification terminology and classification of techniques。

[b-ISO/IEC 29100] ISO/IEC 29100:2011, Information technology – Security techniques – Privacy framework。

ITU-T系列建议书

系列A	ITU-T工作的组织
系列D	资费及结算原则和国际电信/ICT的经济和政策问题
系列E	综合网络运行、电话业务、业务运行和人为因素
系列F	非话电信业务
系列G	传输系统和媒介、数字系统和网络
系列H	视听及多媒体系统
系列I	综合业务数字网
系列J	有线网络和电视、声音节目及其他多媒体信号的传输
系列K	干扰的防护
系列L	环境与ICT、气候变化、电子废物、节能；线缆和外部设备的其他组件的建设、安装和保护
系列M	电信管理，包括TMN和网络维护
系列N	维护：国际声音节目和电视传输电路
系列O	测量设备的技术规范
系列P	电话传输质量、电话设施及本地线路网络
系列Q	交换和信令，以及相关的测量和测试
系列R	电报传输
系列S	电报业务终端设备
系列T	远程信息处理业务的终端设备
系列U	电报交换
系列V	电话网上的数据通信
系列X	**数据网、开放系统通信和安全性**
系列Y	全球信息基础设施、互联网协议问题、下一代网络、物联网和智慧城市
系列Z	用于电信系统的语言和一般软件问题

瑞士印刷
2021年，日内瓦