



中华人民共和国国家标准

GB/T 42460—2023

信息安全技术 个人信息去标识化效果评估指南

Information security technology—
Guide for evaluating the effectiveness of personal information de-identification

2023-03-17 发布

2023-10-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	I
引言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 个人信息去标识化效果分级	3
5 个人信息去标识化效果评估流程	3
6 评估实施	4
6.1 评估准备	4
6.2 定性评估	5
6.3 定量评估	5
6.4 形成评估结论	5
6.5 沟通与协商	5
6.6 评估过程文档管理	5
附录 A (资料性) 直接标识符示例	6
附录 B (资料性) 准标识符示例	7
附录 C (资料性) 准标识符识别	8
附录 D (资料性) 基于 K 匿名模型的去标识化效果评估示例	10
参考文献	15

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》的规定起草。

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专利的责任。

本文件由全国信息安全标准化技术委员会(SAC/TC 260)提出并归口。

本文件起草单位：清华大学、中国电子技术标准化研究院、北京大学、绿盟科技集团股份有限公司、上海二零卫士信息安全有限公司、中国软件评测中心、北京天融信网络安全技术有限公司、蚂蚁科技集团股份有限公司、阿里巴巴(北京)软件服务有限公司、北京市政务信息安全保障中心、深圳市腾讯计算机系统有限公司、北京百度网讯科技有限公司、中国人民银行数字货币研究所。

本文件主要起草人：金涛、王建民、周晨炜、谢安明、张峰昌、陈磊、查海平、赵亮、王龔、叶晓俊、屈劲、白晓媛、李媛、刘巍然、刘俊河、洪爵、宋玲妮。

引 言

GB/T 35273 提出了个人信息去标识化的要求,明确了个人信息去标识化处理的环节和场景,GB/T 37964 就如何开展个人信息去标识化活动给出了指导。经去标识化处理后的个人信息并不能完全实现匿名化,仍存在重标识的风险,需结合应用场景进行去标识化效果评估。

本文件旨在依据个人信息能多大程度上标识个人身份(即标识度)进行分级,用于评估个人信息去标识化活动的效果。个人信息基于标识度分级,有利于个人信息分级别探讨适用场景和安全管理要求,更有利于个人信息的使用和保护。根据国内外相关研究及实践成果,附录中给出了可供参考的计算方法和阈值推荐。

信息安全技术

个人信息去标识化效果评估指南

1 范围

本文件提供了个人信息去标识化效果分级与评估的指南。

本文件适用于个人信息去标识化活动,也适用于开展个人信息安全管理、监管和评估。

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 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 37964—2019 信息安全技术 个人信息去标识化指南

3 术语和定义

GB/T 25069—2022、GB/T 35273—2020、GB/T 37964—2019 界定的以及下列术语和定义适用于本文件。

3.1

个人信息 **personal information**

以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息。

注:不包括匿名化处理后的信息。

[来源:GB/T 35273—2020,3.1,有修改]

3.2

个人信息主体 **personal information subject**

个人信息所标识或者关联的自然人。

[来源:GB/T 35273—2020,3.3]

3.3

去标识化 **de-identification**

通过对个人信息的技术处理,使其在不借助额外信息的情况下,无法识别或者关联个人信息主体的过程。

[来源:GB/T 35273—2020,3.15]

3.4

微数据 **microdata**

一个结构化数据集,其中每条(行)记录对应一个个人信息主体,记录中的每个字段(列)对应一个属性。

[来源:GB/T 37964—2019,3.4]