ICS 35.040
L 80

中华人民共和国国家标准

GB/T 37964—2019

信息安全技术
个人信息去标识化指南

Information security technology—
Guide for de-identifying personal information

2019-08-30 发布　　2020-03-01 实施

国家市场监督管理总局
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。

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/TC 260)提出并归口。

本标准起草单位:清华大学、启明星辰信息技术集团股份有限公司、浙江蚂蚁小微金融服务集团有限公司、阿里巴巴(北京)软件服务有限公司、北京奇安信科技有限公司、北京天融信网络安全技术有限公司、中国科学院软件研究所、中国软件评测中心、上海计算机软件技术开发中心、北京数字认证股份有限公司、西安电子科技大学、湖南科创信息技术股份有限公司、中国电子技术标准化研究院、陕西省信息化工程研究院。

本标准主要起草人:金涛、谢安明、陈星、白晓媛、郑新华、刘贤刚、陈文捷、刘玉岭、宋鹏举、赵亮、宋玲娓、叶晓俊、王建民、方明、裴庆祺、潘正泰。

引　言

在大数据、云计算、万物互联的时代，基于数据的应用日益广泛，同时也带来了巨大的个人信息安全问题。为了保护个人信息安全，同时促进数据的共享使用，特制定个人信息去标识化指南标准。

本标准旨在借鉴国内外个人信息去标识化的最新研究成果，提炼业内当前通行的最佳实践，研究个人信息去标识化的目标、原则、技术、模型、过程和组织措施，提出能科学有效地抵御安全风险、符合信息化发展需要的个人信息去标识化指南。

本标准关注的待去标识化的数据集是微数据（以记录集合表示的数据集，逻辑上可通过表格形式表示）。去标识化不仅仅是对数据集中的直接标识符、准标识符进行删除或变换，可以结合后期应用场景考虑数据集被重标识的风险，从而选择恰当的去标识化模型和技术措施，并实施合适的效果评估。

对于不是微数据的数据集，可以转化为微数据进行处理，也可以参照本标准的目标、原则和方法进行处理。例如针对表格数据，如果关于同一个人的记录有多条，则可将多条记录拼接成一条，从而形成微数据，其中同一个人的记录只有一条。

信息安全技术
个人信息去标识化指南

1 范围

本标准描述了个人信息去标识化的目标和原则，提出了去标识化过程和管理措施。

本标准针对微数据提供具体的个人信息去标识化指导，适用于组织开展个人信息去标识化工作，也适用于网络安全相关主管部门、第三方评估机构等组织开展个人信息安全监督管理、评估等工作。

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 25069—2010 信息安全技术 术语

3 术语和定义

GB/T 25069—2010 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

以电子或其他方式记录的能够单独或与其他信息结合识别特定自然人身份或反映特定自然人活动情况的各种信息。

[GB/T 35273—2017，定义 3.1]

3.2

个人信息主体 personal data subject

个人信息所标识的自然人。

[GB/T 35273—2017，定义 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，无法识别个人信息主体的过程。

[GB/T 35273—2017，定义 3.14]

注：去除标识符与个人信息主体之间关联性。

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。

3.5

聚合数据 aggregate data

表征一组个人信息主体的数据。

注：例如各种统计值的集合。