高校数据分类分级策略的探讨与实践

张　聪
（深圳大学　信息中心，广东　深圳　518000）

摘　要：随着教育信息化的深度推进，高校数据业务越来越深入师生工作生活，数据安全与个人信息安全问题随之日益突出。通过分析高校数据特点，结合高校数据实际，提出一套切实可行的分类分级策略，并提出针对不同分级数据采取的数据保护方案，旨在保障高校数据的安全性和隐私性。

关键词：数据分类分级；数据安全；分级保护；高校数据
中图分类号：TP274　　**文献标识码：**A　　**DOI：**10.19358/j.issn.2097－1788.2024.06.008
引用格式：张聪. 高校数据分类分级策略的探讨与实践［J］. 网络安全与数据治理，2024，43(6)：53－57.

Discussion and practice of data classification and grading strategy in colleges

Zhang Cong
(Information Center, Shenzhen University, Shenzhen 518000, China)

Abstract: With the deepening of educational informatization, data business in universities has become more and more deeply involved in the work and life of teachers and students, and the problems of data security and personal information security have become increasingly prominent. Through analyzing the characteristics of university data and combining with the reality of university data, this paper proposes a set of feasible classification and grading strategies to ensure the security and privacy of university data, and proposes data protection schemes for different graded data.

Key words: data classification and grading; data security; protection measures with data grading; university data

0　引言

在教育信息化的不断深入发展中，高校信息化业务不断拓展，数据应用范围越来越广泛，业务系统之间的数据流转也越来越紧密，数据的价值得到巨大提升。同时，数据安全事件尤其是个人信息泄露事件频发，使得数据安全成为广受重视的核心问题。

2021年6月10日，《中华人民共和国数据安全法》（下称《数据安全法》）获得全国人大通过，并于9月1日起正式实施。作为我国在数据安全领域的重要法律，该法第二十一条明确规定国家各行业、各领域应建立数据分类分级制度，并执行分类分级的数据保护措施[1]。本文依据高校数据特点及管理实际情况，探讨适合高校的数据分类分级策略。

1　数据分类分级概述

1.1　数据分类分级定义

数据分类和数据分级属于不同的概念。数据分类是指将数据依据相似的属性或功能聚集在一起，目的是便于数据管理及数据安全保护措施的实施；数据分级是指根据数据的重要程度，以及数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用后，对国家安全、经济运行、社会稳定、公共健康和安全、组织及个人利益造成的危害程度，对数据进行定级管理。

1.2　国内外数据分类分级现状

数据分类分级的概念起源于美国，美国政府在安全数据和非涉密敏感数据方面做了分类分级要求。2009年12月29日，时任总统奥巴马签署13256号总统令《美国安全信息分类》（Classified National Security Information, CNSI），将国家安全信息分为绝密（Top secret）、机密（Secret）和秘密（Confidential）三个等级，提到了对于数据分类分级进行培训的要求[2]。2010年10月4日奥巴马签署《受控未分类信息》（Controlled Unclassified Information, CUI），为政府受控非涉密信息进行标记、保护、传播、控制等方面提出了要求。其中，CUI包含了一系列由美国国家标准技术研究院（NIST）出版的出版物，其中出版物199《联邦信息和信息系统安全分类标准》（Standards for Security Categorization of Federal Information

and Information Systems)，从保密性、完整性、可用性三个维度将信息和信息系统分为低、中、高三个等级，并要求采取相应的保护措施[3-4]。欧洲没有对数据进行分类分级，欧洲数据保护法律主要依据欧盟 2018 年推出的《通用数据保护条例》(General Data Protection Regulation，GDPR)，取代 1995 年的《数据保护指令》(Data Protection Directive)，GDPR 强调了个人信息处理者处理数据的原则、义务、应采取的保护措施，着重阐述了对遗传数据、健康数据等敏感个人数据的额外保护策略。该条例主要针对为欧盟国家提供服务的商业机构使用个人数据进行监管，违反 GDPR 的企业予以重罚[5]。调查咨询公司 Forrester Research 在《你的数据有多脏?》的调查报告指出数据分类是 GDPR 合规的重要组成部分，并认为应该依据可识别性、敏感性、稀缺性等维度进行分类[6]。

我国在 2018 年国务院办公厅发布《科学数据管理办法》，对科学数据率先进行分类分级管理，金融、铁路、通信等许多领域也出台了分类分级指南，并采取相应措施进行保护。2021 年全国人大通过了《数据安全法》和《个人信息保护法》，确立了全国范围数据分类分级的法律基础[7]。在此之后，全国信息安全标准化技术委员会秘书处于 2021 年 12 月制定《网络安全标准实践指南—网络数据分类分级指引》(下称《指引》)[8]，广东省网络空间安全协会于 2022 年制定《高校数据分类分级指南》(下称《指南》)[9]，成为高校数据分类分级的重要依据。2024 年全国网络安全标准化技术委员会发布《数据安全技术 数据分类分级规则》(下称《规则》)(GB/T 43697－2024)，为全国数据分类分级工作确定统一方法，给出各类数据的识别指南[10]。

1.3 数据分类分级对高校的重要性

数据分类分级对高校的重要性体现在以下几个方面：

(1) 促进数据安全精细化管理。建立数据分类分级制度，实行数据分类分级保护措施，是《数据安全法》和《个人信息保护法》的内在要求。通过对重要性更高的数据进行更高程度保护，有利于加深数据安全认识，优化数据安全防护资源的分配，降低敏感核心数据发生安全事故的概率。

(2) 促进数据治理和质量提升。高校在进行数据分类分级过程中了解自身数据资产特点及生命周期流程，了解自身数据存在的缺陷和问题，有助于高校针对性地提出数据治理和数据质量提升计划，改善自身数据情况。

(3) 规范数据要素流动和交易。数据分类分级保护措施要求依据分类分级结果，对数据流通、交易、提供行为进行不同程度的监督，有利于规范数据要素流动、交易和提供行为，使得高校对自身数据流转情况更为清晰，以作出针对性的制度、流程和人员安排。

(4) 规范数据全生命周期管理。通过将数据分类分级保护措施深入到数据全生命周期管理中，可使得高校更清晰自身数据产生、储存、流转、提供、销毁等全生命周期的运转流程，为高校信息化业务提高行政效率、减少多头填报作出贡献[11]。

2 高校数据分类分级存在的误区

2.1 分类分级对象不清晰

部分高校对于应进行分类分级的对象不明晰，导致数据分类分级覆盖面不足。《指南》中对高校数据的定义是“高校在开展或辅助开展教育活动过程中以电子或者其他方式对信息的记录，产生环节包括但不限于收集、存储、使用、加工、传输、提供、公开等”[9]。这意味着高校在进行或辅助进行教学活动中涉及到的所有数据，包括人员信息、教学信息、科研信息、资产信息等都应属于高校数据。除此之外，还应包括：

(1) 开展或辅助开展教学活动涉及到的业务办理过程中生成的数据。如审批人、审批时间、申请原因、申请事由等。

(2) 办理业务中以电子或其他形式形成的非线性数据。如教学视频、申请表格、教学图片、扫描件等。

(3) 开展或辅助开展教学活动的归档数据。

(4) 支撑教学活动进行的业务系统数据，包括业务系统参数、日志等。

2.2 分类分级认识不清晰

部分高校对于数据分类分级的作用不明确，混淆数据分类和数据分级两个概念。《数据安全法》指出数据分类分级的目的是要执行“分级数据保护”的措施。《规则》中明确数据分类的目的是便于管理，数据分级的目的是保护数据安全[10]。《指引》中指出数据分类分级应当以“分类管理，分级保护”为原则[8]，二者含义不同，目的不同，不可相互替代，因此数据分类和数据分级依据的维度和标准也完全不同。

2.3 分类分级权责不清晰

部分高校不清楚数据分类分级的制定流程，认为都由信息化部门确定，业务部门信息化素养不足，对数据安全一知半解甚至完全不了解。事实上数据安全是由数据生命周期的相关方共同维护，分类分级工作也需要各相关方共同参与，信息化部门依据专业能力统筹、统合数据分类分级结果。分类分级制定过程中应与数据生产方、使用方充分交流，了解数据应用场景、应用规模、数据量大小以及发生数据安全事故造成的影响，科学制定数据分类分级方案。

2.4 分类分级流程不清晰

部分高校认为分类分级是一次性工作，完成后就束之高阁，万事大吉。数据分类分级应遵循分类（级）规划、分类（级）准备、分类（级）审批、分类（级）实施、结果评估、维护改进等步骤。分类分级制定完成后应针对不同分级制定并落实数据保护措施，定期评估分类分级效果，持续改进分类分级方案。

3 高校数据分类分级方案

3.1 高校数据资产特点分析

高校作为非营利的高等教育机构，数据存储和流动有其特点。

（1）数据量不大，数据涉及领域主要为教育、科研领域。高校在校师生规模的数据量级一般在数千到数万之间，资产规模的数据量级在数万到数十万之间。这样的人员和资产规模决定了高校数据量级不会非常庞大，除流水外如成绩、课程、项目经费等数据规模量级在千万以内。数据领域偏向教育、科研范畴，受教务、科研相关数据标准影响较大。

（2）存在大量过程类数据。办理业务以申请－审核类为主，高校多数业务是各教学部门通过部门秘书或师生个人发起，经学院领导、业务部门老师审批通过后成为既定事实。在这一业务流程中生成的数据及其日志多数是过程数据，仅有最终的结果为结果数据。

（3）数据类型多样。除了传统的教学、科研、人员信息等线性数据外，还包括视频、图片、文档、日志等非线性数据，以及完成归档的数据。

（4）数据流转的核心是人员基本信息，对外提供数据较少。师生是学校业务的行为主体，因此与其相关的个人信息是数据流转的核心，占据数据流转总量的 80% 以上，如学生学号、姓名、性别、所属学院、年级、班级、专业等是常用的个人信息属性。与人员相关联的信息如课程、选课、成绩等信息也经常流动。另外高校数据对外提供的情况较少，主要都是在内部流动。

3.2 高校数据分级分类原则

依据《规则》和《指引》的要求，数据分类分级遵循以下原则[9-10]：

（1）合法合规性。高校数据分类分级应遵守国家和教育行业的法律法规、指南、标准等。

（2）适用性。遵循本校实际，确保分类分级标准的可执行性。

（3）科学性。分类分级规则应分别根据多个维度综合考虑制定，规则应保持相对稳定。

（4）就高从严。数据集的定级不应低于其中数据项的最高级别，需考虑数据规模的变化。

（5）动态更新。数据分类分级的规则和结果不是一成不变的，应该按照学校的安全形势要求不断调整。

3.3 数据资产识别和梳理

在进行分级分类之前，应先对学校数据资产进行梳理，确定数据分类分级的范围。数据包括但不限于各类业务系统存储的信息、日志、附件、多媒体文件等。

3.4 数据分类方案

《规则》指出数据分类应以先行业领域、再业务属性的方式分类。高校数据绝大部分属于教育行业数据，因此可依据业务域进行分类。以深圳大学为例，如图 1 所示，高校数据包括人事域、教务域、研究生域、学工域、科研域、资产域等，并可再细分为数据集。每个业务系统可根据主要业务范围划分至不同业务域下，如表 1 所示。数据集的数据项内容可根据业务使用过程的聚合程度来决定，也可以根据数据的存储结构来决定。

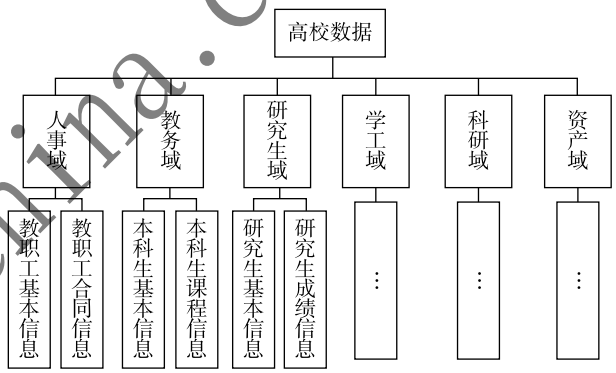


图 1 高校数据分类图（示例）

表 1 业务域主要业务系统及主管部门（示例）

业务域	主要业务系统	主管部门
人事域	人事系统 ……	人力资源部
教务域	教务系统 教师评学系统 ……	教务部
研究生域	研究生系统 成绩自助打印系统 ……	研究生院
学工域	学工系统 宿管系统 违纪处分系统 ……	学生部武装部
科研域	科研系统 科研质量管理平台 ……	科学技术部 社会科学部
资产域	房产管理系统 国有资产管理系统 ……	设备与国有资产管理部 后勤保障部

3.5 数据分级方案

《规则》和《指南》中都将数据分为核心、重要、一般三个级别，《指南》还将一般数据又分为三个级别：内部敏感级（一级）、有条件外部共享级（二级）、其他（三级），并给出了每个级别的建议划分标准。在分级维度上，可从群体、重要性、精度、深度、规模等方面综合考量。结合高校数据资产的特点，笔者制定下列数据分级规则：

（1）以数据发生泄露、篡改、丢失、非法利用等安全事故带来的危害作为主要依据判断重要性，数据共享的范围也应纳入判断范围，不作共享的数据可分入低重要级，非敏感性系统数据可纳入一般数据的级别。

（2）以数据描绘个人或业务主体的精确程度为依据判断精度。可依据本校实际情况定义高精度和低精度。

（3）以数据存储、流动量大小为依据判断规模。《指南》中分别设置了1万、10万个人信息和1 000、1万个人敏感信息作为重要数据、敏感数据的指标性门槛。在实践中应当考虑高校本身数据规模情况制定适合的规模门槛。以深圳大学为例，个人基本信息和个人身份信息在万级的存储量，因此个人基本信息纳入重要数据级别，个人身份信息纳入核心数据级别。而个人基本信息以外的数据不少是十万量级，部分数据在百万级，可考虑以十万以上规模作为重要数据的考虑门槛，百万以上规模作为敏感数据的考虑门槛等。

（4）个人信息应单独考虑定级。个人信息受到《个人信息保护法》约束，有单独的处理方式和规定，应在分类分级和数据保护措施中有所体现。个人信息定级应遵循统一的定级规则，依据群体、规模等属性确定不同数据集内的个人信息定级。

确定数据的分级后依据就高从严的原则，以不低于数据集中数据项最高级别的原则进行定级。数据定级流程如图2所示。

最终数据分级如表2所示。

3.6 数据分级保护措施

数据分级的目的是为了在促进数据流动的情况下保护重要数据的安全，因此必须制定科学、适当的保护策略，将数据分级的作用落到实处。笔者认为可循以下方面进行数据保护：

（1）访问控制。依据不同分级设定不同的访问控制策略，包括设置网络及服务器防火墙白名单、交换机访问策略、客户端账号控制等。

（2）角色和权限控制。依据不同分级设定不同的权限控制策略，主要方法为设定基于角色和角色组的切合实际的角色管理体系，通过精细化的权限管理控制各类人员访问、修改数据的权利，对于包含高级别重要数据的业务系统，应设置专门角色管理、配置权限，并定时评估角色权限配置的科学性。

（3）容灾备份和演练。依据不同分级设定不同的容灾备份和恢复演练策略，设置定时演练，分级越高容灾频度越高，演练要求也越高。

（4）安全保障技术手段。依据不同分级设定不同的加密、脱敏、水印、签名等策略，对级别高的数据应视情况要求进行动态脱敏，并针对数据全生命周期进行精细化安全保障技术手段管理。

（5）评估和审查。依据不同分级设定不同的数据安

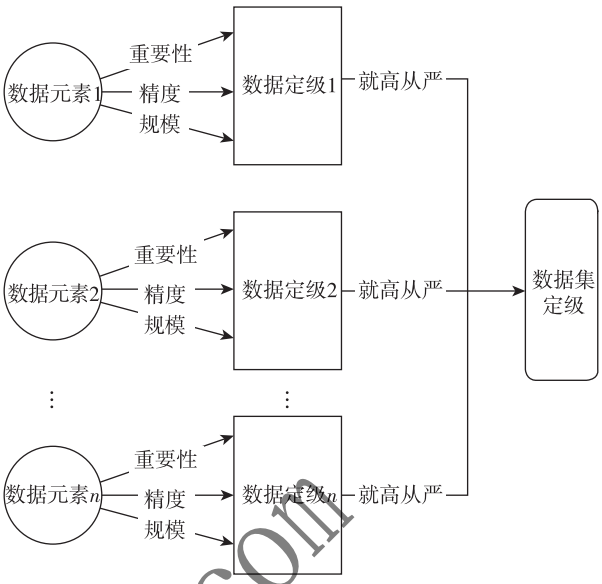


图2 数据定级流程图

表2 高校数据分级（示例）

数据项	安全级别	数据集	安全级别
学工号	重要		
姓名	重要		
学分绩点	核心	本科生成绩信息	核心
等级制成绩	核心		
……	……		
课程号	一般		
课程名	一般		
课程英文名	一般	本科生课程信息	一般
课程分类	一般		
……	……		
……	……	本科生个人基本信息	重要
……	……	本科生个人身份信息	核心
……	……	……	……

全评估和个人信息合规审查频度。存有高级别数据的业务系统应该接受频度更高的安全评估和审查。

4 结论

数据分类分级作为《数据安全法》的法定义务，是高校必须完成的工作。通过将数据分类分级，也促进高校了解自身数据现状，利于数据管理和数据质量提升。目前国内的数据分类分级仍处于探索阶段，如何通过科学定级，做到促进数据流通共享和保障数据安全的平衡，仍然是一个重大课题。本文提出一套切实可行的高校数据分类分级方案，以期在符合相关法律法规和标准的情况下做出符合本校实际的分类分级，为本校数据安全保驾护航。

参考文献

[1] 高磊，赵章界，林野丽，等. 基于《数据安全法》的数据分类分级方法研究 [J]. 信息安全研究，2021，7（10）：933－940.

[2] The President Executive Order 13526. Classified National Security Information [EB/OL]. [2023－12－19]. https://www.archives.gov/isoo/policy-documents/cnsi-eo.html.

[3] 完颜邓邓，陶成煦. 美国政府数据分类分级管理的实践及启示 [J]. 情报理论与实践，2020，12（43）：172－177.

[4] 卢锐恒，许晓耕，白雪珺，等. 敏感个人信息分类分级研究 [J]. 安全与通信保密，2023（4）：46－56.

[5] EUR-Lex 32016R0679 [EB/OL]. [2023－12－20]. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX：32016R0679.

[6] KHATIBLOO F. How dirty is your data [EB/OL]. [2023－12－02]. https://www.forrester.com/report/How-Dirty-Is-Your-Data/RES73121.

[7] 严炜炜，谢顺欣，潘静，等. 数据分类分级：研究趋势、政策标准与实践进展 [J]. 数据科学，2022（9）：2－12.

[8] 全国信息安全标准化技术委员会秘书处. 网络安全标准实践指南—网络数据分类分级指引（TC260/PG－20212A）[S]. 2021.

[9] 广东省网络空间安全协会. 高校数据分类分级指南（T/GDCSA 014－2022）[S]. 2022.

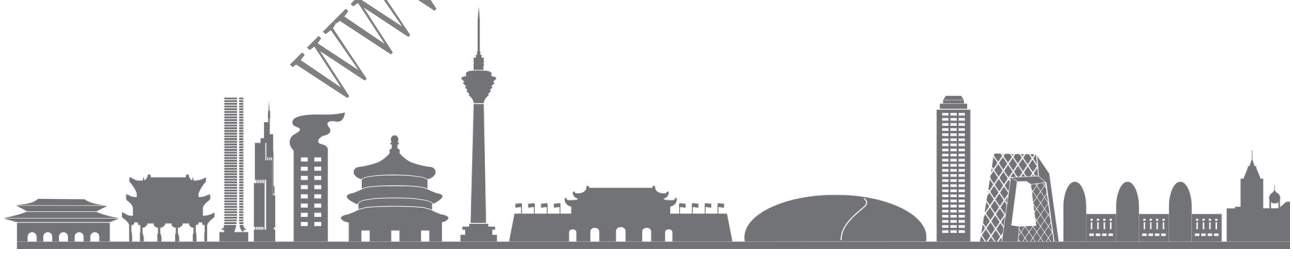
[10] 全国网络安全标准化技术委员会. 数据安全技术 数据分类分级规则（GB/T 43697－2024）[S]. 2024.

[11] 廉康，刘力铭，罗伟雄，等. 高职院校的数据分类分级方法探索 [J]. 电子元器件与信息技术，2023，7（8）：128－131.

（收稿日期：2024－03－21）

作者简介：

张聪（1988－），男，硕士，工程师，主要研究方向：数据治理、大数据、数据安全、数论。



版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部