

# 高校数据资产编目研究

刘豆豆, 焦宝臣

南开大学大数据管理中心, 天津 300071

## 摘要

在网络、数据安全愈发重要的数字经济时代, 在尚缺乏科学合理的适用于高校的数据资产编目模型现状下, 分析了高校数据资产编目的功能和作用, 提出了以数据应用为切入点、以数据服务为目标的高校数据资产编目模型。该模型由相辅相成、互为支撑的数据资产目录建设、数据安全、数据服务三部分构成。基于模型总结了高校数据资产编目五步骤工作框架, 包括数据盘点、数据分类、数据分级、数据标签、数据开放。以南开大学为例, 依据构建的模型和工作框架具体实施数据资产编目工作, 取得了较好的应用效果。这表明提出的模型和工作框架具有较强的可操作性, 能够为各高校开展数据资产编目提供科学可行的技术路线和方法指导。

## 关键词

数据资产编目; 数据分类分级; 数据标签; 数据安全

中图分类号: TP391.4, G647

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2023030

## *Research on data asset cataloging of colleges and universities*

LIU Doudou, JIAO Baochen

Big Data Management Center, Nankai University, Tianjin 300071, China

## *Abstract*

In the era of digital economy in which network and data security are becoming more and more important, there is still no scientific and reasonable data asset cataloging model applicable to colleges and universities. The function of data asset cataloging in colleges and universities was analyzed, and the data asset cataloging model with data application as the breakthrough point and data service as the purpose was put forward. The model was composed of three parts: data asset catalog construction, data security and data service that complement and support each other. The five-step framework of data asset cataloging was summarized: data inventory, data classification, data staging, data labeling, data opening. Nankai University implemented data asset cataloging according to the model and working framework, and obtained good application results. The results show that the model and working framework are operational and can provide scientific and feasible technical route and method guidance for colleges and universities to carry out data asset cataloging.

## *Key words*

data asset cataloging, data classification and staging, data label, data security

## 0 引言

近年来,国家对数据安全愈加重视,相继出台了《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《网络安全标准实践指南——网络数据分类分级指引》等法律法规,明确指出为了保障数据安全,需要对数据进行分类分级管理,而数据分类分级的核心和前提就是要进行数据资产编目。同时,数据资产编目也是数据分类分级管理的重要手段和表现形式,是数据安全的根本要求。

近几年来,数据资产逐渐被广泛关注,尤其在2022年达到前所未有的研究高峰。目前国内外对数据资产编目的研究主要集中在政府、企业、医疗等领域。例如,政府领域,夏义堃等人<sup>[1]</sup>对政府数据资产管理要素框架以及运行模式进行研究;企业领域,Hannila H等人<sup>[2]</sup>从产品组合管理角度对企业数据资产进行分类,以及车伟等人<sup>[3]</sup>提出电力公司数据盘点八步法及构建电力数据资源目录的方法;医疗领域,Shoba U S等人<sup>[4]</sup>指出医院临床数据资产化面临的问题,并提出创建数字化临床记录系统的解决方案。随着数据资产编目研究领域逐渐扩大,2018年由中国信息通信研究院云计算与大数据研究所等编写的《数据资产管理实践白皮书(3.0版)》出台,其涵盖了电信、金融、政务、医疗和工业等相关领域的的数据资产管理,为政府和企业数据资产管理提供了参考。而针对高校的数据资产编目主要在对高校数据治理和高校数据资产管理的整体介绍中提及。例如,彭雪涛<sup>[5]</sup>对美国高校数据治理的研究、贺建虎<sup>[6]</sup>对高校智慧校园数据治理架构的设计。尚缺乏针对教育行业高校

数据资产编目模型和方法的详细研究和权威性文件指导。

高校是一个人员聚集的场所,以人为中心形成了学习、办公、科研、财务、社团、生活等多方面的社会活动,随之产生的数据不仅量大,而且涉及范围广,数据种类多,复杂度高,更新频率快。高校的数据资产结构与政府、行业企业存在一定的差异,因此政务和行业企业的数据资产编目并不完全适用于高校。随着信息化的不断发展,各类微型应用和各种功能的信息系统层出不穷,为了满足这些应用和系统对数据的多样化需求,数据资产目录的建设更加迫在眉睫。挖掘高校的数据资产价值,为高校的管理、决策、发展提供有力的数据支撑,首先需要全方位地了解学校的数据资产,并形成清晰、有序、适合于高校的数据资产目录,方便高校对所属数据资产进行系统性、体系化的管理,同时发挥出数据的内在价值。

教育部部长怀进鹏同志提出“应用为王、服务至上、示范引领、安全运行”的工作要求和思路,强调把业务应用摆在优先突出位置,以应用需求驱动运行平台、安全平台、标准平台和数据资源平台建设,加强内容建设和运营维护,不盲目追求最新技术,切实为师生提供能用、好用的数字化资源。基于此,高校数据资产编目的建设应该以数据应用需求为导向,遵循数据资产目录为应用服务的原则,构建易懂、清晰、全面的高校数据资产目录。

为此,本文从深入分析高校数据资产编目的功能和作用入手,提出了以数据应用需求为导向的数据资产编目思路,研究构建出适用于高校的数据资产编目模型和工作框架,并以南开大学为例,介绍了数据资产编目模型和工作框架的具体落地实施情况。

## 1 高校数据资产编目的功能和作用

在研究高校数据资产编目之前,首先需要明确高校数据资产目录与信息标准的区别。信息标准是依托于国家或者教育部为全国所有高校建设编制的统一的数据标准,涵盖内容广泛、全面,是所有类型高校数据的综合体<sup>[7]</sup>,但是每个学校的情况不同,信息标准中的数据内容并不是每所学校都有所涉及。而数据资产目录反映的则是学校真实、可用的数据情况,目录中包含高校拥有的全量数据资产,支持提供给学校各个学院或职能部门使用。因此,高校数据资产编目在高校数据治理体系中处于承上启下的位置,上承数据集成和标准化,下启数据开放和应用<sup>[8]</sup>。

高校拥有多个职能明确的二级部门,各个二级部门在为学生和教职工服务时会产生相应的数据,各二级部门负责的工作存在一定的独立性,产生的业务数据也具有相应的权威性,同时各二级部门之间也会有部分业务数据的共享<sup>[9]</sup>。因此高校的数据具有来源广、规模大、时效性强的特点。高校数据资产编目工作需要将高校数据资产的这些代表性考虑在内。

高校数据资产编目应具备如下功能和作用。

- 数据资产内容全面且真实可用;
- 综合考虑高校实际业务之间的关联性,建设方便共享的数据资产目录;
- 为数据管理者提供更方便的数据分析和管理的;
- 方便数据使用者快速定位需求数据;
- 方便学校领导全面了解学校数据资产,通过数据挖掘分析为学校发展提供参考依据;

- 提高师生对全校数据治理工作的参与度。

从高校数据资产编目的功能和作用中可以得出,高校数据资产目录主要为领导层和各类应用服务,因此高校数据资产目录的建设要从数据资产的实际应用需求出发,要能够为学校的各类应用提供便捷、清晰、全面的数据服务,并通过数据开放共享,持续赋能高校业务。

## 2 高校数据资产编目模型及工作框架

前文指出当前高校数据资产编目存在的最大问题是:只在数据资产管理和数据治理等过程中被泛泛提及<sup>[10-11]</sup>,并没有提出具体的高校数据资产编目的工作步骤以及编目依据。本文通过分析高校数据资产编目的功能和作用,提出从数据实际应用需求角度构建数据资产目录的思路,并基于该思路的指导,进一步给出高校数据资产编目模型和具体的工作框架。

### 2.1 数据资产编目模型

以实际的数据应用需求为导向,结合高校数据治理相关工作内容,构建高校数据资产编目模型,如图1所示。

图1给出了高校数据资产编目模型,从图1可以看出,该模型主要由三大部分构成:数据资产目录建设、数据安全、数据服务,三部分相辅相成,互为支撑。

#### (1) 数据资产目录建设

在高校数据治理的整个生命周期中,主要有数据集成、数据管理、数据开放3个核心阶段,这3个阶段分别完成数据治理的不同工作,并沉淀出各类数据文档<sup>[12]</sup>。这3个核心阶段对数据资产编目在不同功能方面有重要的指导作用。

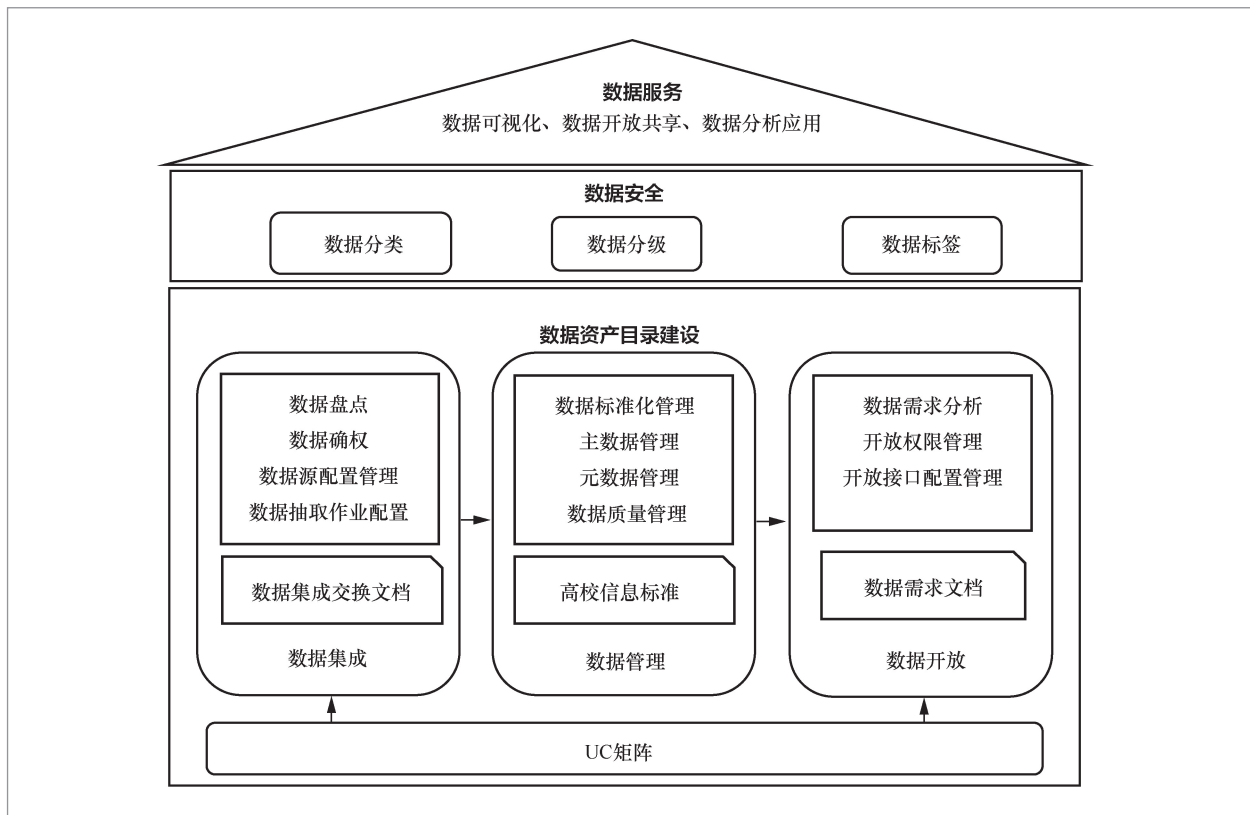


图1 高校数据资产编目模型

● UC矩阵。随着越来越多的数据集成到数据中心，同时数据中心也为学校越来越多的信息系统提供数据支撑，学校的UC矩阵逐步丰富完善起来<sup>[13]</sup>。基于UC矩阵文档可以总体上掌握生产数据的系统以及使用数据的系统，从系统的属性角度掌握学校的数据资产情况以及应用数据需求情况，为数据资产编目提供基础支撑。

● 数据集成。本阶段涉及数据盘点、数据确权、数据源配置管理、数据抽取作业配置等工作。数据集成阶段确定了数据生产部门、数据生产系统以及生产数据，为数据盘点划定了数据范围。此阶段形成的数据集成交换文档，可作为UC矩阵中生产数据系统生产数据的进一步详细说明，基本涵盖了全校所有的数据资产，为数据

资产编目提供了原材料。

● 数据管理。本阶段涉及数据标准化管理、主数据管理、元数据管理、数据质量管理等工作。数据管理阶段会生成更多数据标签，为数据多维度分析提供依据。此阶段会形成高校信息标准（包括管理信息和基础代码），信息标准文档对数据资产目录的建设也有一定的参考价值，可为数据资产编目提供参考依据。

● 数据开放。本阶段涉及数据需求分析、开放权限管理、开放接口配置管理等。数据开放阶段以需求为推动力，可以反过来促进数据集成和数据管理阶段对数据的更新完善。此阶段会形成数据需求文档，可作为UC矩阵中使用数据系统的数据需求的进一步详细说明，为数据资产编目提供目标需求。

### (2) 数据安全

数据安全作用于数据治理的整个生命周期，同时也为数据服务提供安全指导。根据国家相关安全法律法规要求，数据需要进行分类分级管理。数据安全为数据资产目录的分类分级和数据标签建设提供法律依据，同时数据资产目录的建设也为数据分类分级提供实现基础。数据资产编目是数据安全方面的重要实现手段和表现形式，同时数据安全是数据资产编目的重要支撑。

### (3) 数据服务

数据服务就是在各安全法规的指导下，将建设好的数据资产目录提供给下行应用，支撑各类应用的数据需求，最终实现数据可视化、数据开放共享以及数据分析应用。

综上所述，高校数据资产编目模型以UC矩阵为基础，以数据治理过程（数据集成、数据管理、数据开发3个阶段）中形成的重要文档为原材料、参考依据和目标需求，同时作为数据安全的重要实现手段和表现形式，为数据服务提供强有力的支撑。

## 2.2 数据资产编目工作框架

本节根据上述高校数据资产编目模型，结合高校数据相关工作的特性，构建出高校数据资产编目的具体工作框架，详细指出基于应用数据需求的高校数据资产目录建设的工作步骤和过程。高校数据资产编目工作框架如图2所示。

从图2可以看出，高校数据资产编目工作框架核心步骤主要包括数据盘点、数据分类（数据资产域划分和数据资产子类划分）、数据分级、数据标签、数据开放。

第一步：数据盘点。

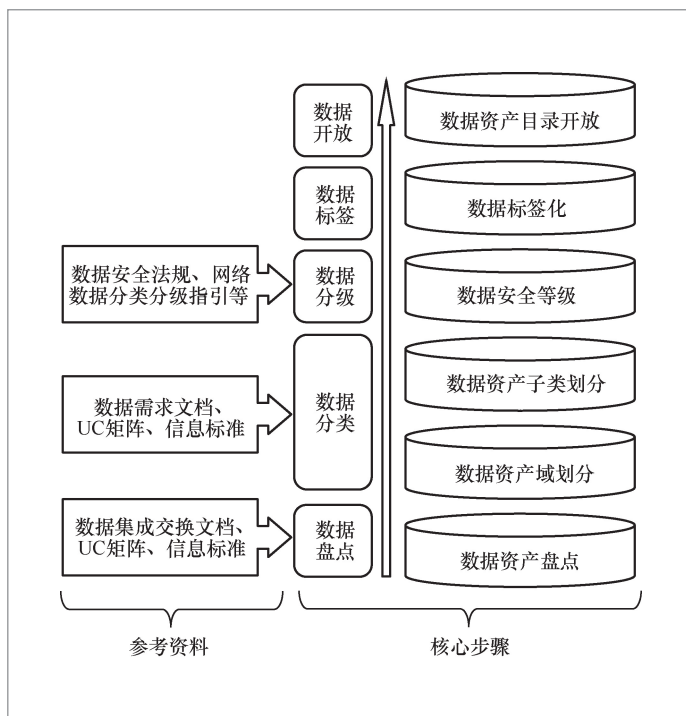


图2 高校数据资产编目工作框架

数据资产盘点是做数据资产编目的重要前提，将信息标准作为参考，将数据集成交换文档和UC矩阵作为支撑，对整体数据进行梳理盘点，可以掌握学校全量的数据资产情况。

第二步：数据分类。

数据分类又细化为数据资产域划分和数据资产子类划分两个子步骤。

- 数据资产域划分。通过对数据需求进行分析，归纳出数据资产大类，形成数据资产域。依据是数据需求文档和UC矩阵。建立数据资产域时要遵循的原则包括数据要全量、层次要清晰、命名要规范、分类不交叉、稳定可扩展。

- 数据资产子类划分。在数据资产域的基础上，通过对数据需求文档中各字段需求热度的统计分析，将数据资产域分别进行资产的细化，建设数据资产子目录。资产细化过程可以根据数据字段的需求热

度来划分,将每个大类涉及的数据表及字段进行汇总、统计、分析,得到数据字段需求热度排序,根据数据需求热度划分出数据资产子类。

第三步:数据分级。

依据《中华人民共和国数据安全法》《网络安全标准实践指南——网络数据分类分级指引》相关要求,同时结合高校数据开放共享实际需求,高校数据安全等级的划分可从多个维度确定,如数据开放程度、数据影响对象及危害程度等。

- 《网络安全标准实践指南——网络数据分类分级指引》中,将对国家安全、公共利益、组织合法权益、个人合法权益4个影响对象的危害程度从低到高分成一般数据、重要数据、核心数据3个级别。其中高校的数据基本上归属于一般数据,一般数据又可以按照数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用,对个人、组织合法权益造成的危害程度,从低到高分为1级、2级、3级、4级共4个级别。

- 随着高校数据在各二级部门的确权,高校的数据从数据中心在全校内进行开放共享,也需要有一定的审批权限,需要根据数据开放程度设定安全级别,以保证数据适宜地开放给合适的人群,因此高校数据还可以从数据开放程度从低到高分1级、2级、3级共3个级别。

数据安全等级的划分可依托于数据资产目录,同时数据安全等级也可为数据资产目录的建设提供更多维度的参考。数据安全等级的划分也为数据开放提供有效依据。

第四步:数据标签。

数据资产目录是高校数据的主分类,但并不能满足多维度、多样化的数据划分需求。因此需要将其他维度通过数据标

签的形式给数据做标记<sup>[14]</sup>。数据标签是从多个不同角度对资产的描述,可以有效地扩充数据的应用角度,辅助数据资产目录对数据进行检索和提炼服务,是对数据目录的一种补充和优化。数据标签和数据资产目录两者相互配合、互为支撑,为高校的数据应用提供便捷、灵活、精准的服务。

第五步:数据开放。

数据资产目录的建设以数据应用为目标,因此数据资产目录的建设最终要为数据开放服务。数据资产目录的管理需要依托于可视化的数据资产管理平台,支持目录的扩建、修改、删除等操作;同时数据开放也需要提供可视化的平台,为用户提供数据目录预览、查询、筛选功能,支持数据脱敏,支持提供数据视图、应用程序接口(application program interface, API)、数据清洗与整合下行作业(data clean integrator, DCI)等类型数据接口。

随着高校信息化的不断发展,涉及的数据范围及领域将会不断地发生变化,数据资产目录处于一个动态调整的状态,这就涉及类目增、减等变更问题。其中包含以下两方面:第一,要明确的是当新的数据进入数据中心并有对外提供数据服务的要求时,需要增加资产目录;第二,新增的目录如果是新的数据源并涉及新增数据领域,需要调整一级目录,如果只是在某一数据源头中增加了一些新的数据,那么只调整数据子类即可。因为数据资产的划分粒度是细化到字段的,所以字段归属哪个大类和小类分别以标签的形式体现,在可视化界面上则体现为资产目录。通过调整数据标签可调整字段的归属类别。需要特别提出的是通常情况下不会通过减少数据资产目录保证学校数据的完整展现。

### 3 应用实例

南开大学依托前文提出的数据资产编目模型以及具体的工作框架,完成了数据资产编目工作,并将数据资产目录应用于全校的数据公共服务,取得了较好的应用效果。下面详细介绍南开大学具体开展的资产编目工作。

#### 3.1 数据资产盘点

学校的数据资产是由学校内各个部门产生的,因此依据学校数据中台中的UC矩阵文档和数据集成交换文档,从数据产生部门和数据生产系统进行逐一梳理,整合汇总全校数据资产。南开大学部分数据资产梳理结果见表1。

#### 3.2 数据资产域划分

依据UC矩阵中的使用数据系统以及数据需求文档开展数据资产域划分。首先对每个使用数据系统的数据需求详细情况进行梳理,并对需求内容进行统计汇总,然后根据统计汇总情况对数据资产进行粗略分类,将相互关联的数据归为一类<sup>[15]</sup>,最终得到数据资产目录大框架,即数据资产域。南开大学数据资产域划分结果如图3所示。

#### 3.3 数据资产子类划分

数据资产域划分完成后,对每个资产域的数据表及字段进行汇总、统计、分析,得到数据字段需求热度排序,根据数据需求热度及属性进行数据资产子类划分。部分教职工数据字段需求热度统计情况见表2。

表1 南开大学部分数据资产梳理结果

数据部门	来源系统	数据表
人事处	人力资源管理服务平台	人员基本信息 离退休信息 专业技术职务 行政职务信息 ...
教务处	创达多媒体系统	多媒体刷卡记录 多媒体扫码记录 ...
	教学管理信息系统	学生选课数据 学生课表数据 学生基本数据 学籍基本数据 学生成绩数据 ...
研究生院	研究生信息管理系统	研究生基本信息 研究生学籍基本信息 学籍异动 研究生课程基本信息 ...
研工部	金智研工部系统	研究生奖学金 研究生贷款 学生奖励信息 研究生学生干部 ...
...	...	...
...	...	...

依据字段需求热度情况对数据字段进行分组,例如需求热度大数量级、需求热度小数量级、相同需求热度字段等。以教职工数据字段需求热度统计为例,可从需

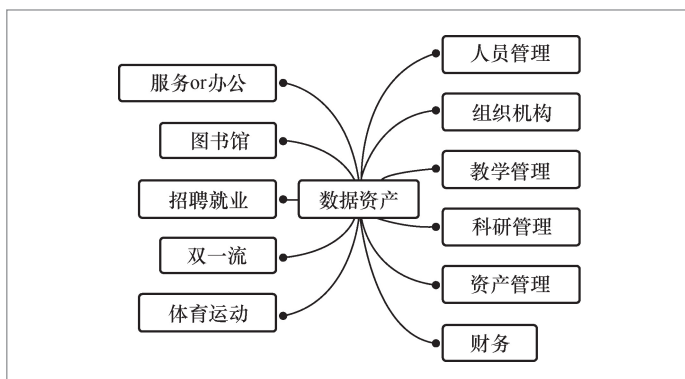


图3 南开大学数据资产域划分结果

表2 部分教职工数据字段需求热度统计情况

字段	字段描述	计数项: 字段需求热度	字段	字段描述	计数项: 字段需求热度
DWDM	所在单位	62	MZ	民族	8
JGH	教工号	56	ZGXW	最高学位	8
XM	姓名	32	ZZMM	政治面貌	7
XB	性别	23	TXDZ	通信地址	6
RYZT	人员状态	18	YJXKM	一级学科	6
RYFL	人员分类	17	GJDQM	国籍	5
DZYG	邮箱	14	XMPY	姓名拼音	5
ZYJSZW	职称	14	ZGXML	学历	5
SJHM	手机号	13	BYSJ	毕业时间	4
SFZJH	身份证件号	12	BYYX	毕业院校	4
SFLB	身份类别	11	XRZW	现任职务	4
JG	籍贯	10	...	...	...

求热度大于等于10和需求热度小于10两大类进行分析。需求热度大于等于10的字段主要分为基本信息和职称职务,在需求热度小于10的字段中,除去基本信息和职称职务相关数据,同时考虑到需求热度均为个位数,可划分为同一小组,得到学位学历相关字段。结合全量的教职工数据字段需求热度统计(因涉及字段过多不便全量截图),除了基本信息、职称职务信息、学位学历信息,还可划分出通信信息、教师工作量信息等需求子类。最终得到教职工的数据资产子目录,如图4所示。

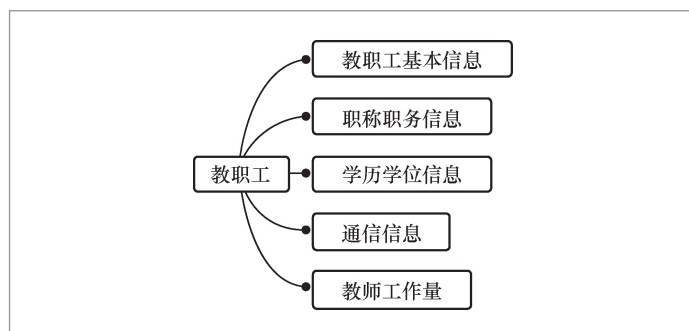


图4 教职工的数据资产子目录

其他各大类数据以同样的方式进行分类细化,同时可根据数据的实际情况进行更多级的目录建设,最终形成高校数据资产目录。

### 3.4 数据安全等级

从数据开放程度维度,南开大学根据数据确权管理以及数据可开放情况<sup>[16-17]</sup>对数据资产设立3个安全等级。

1级:不需要审批就能在全校范围内开放。

2级:审批或脱敏后可在全校范围内或对特定人群开放。

3级:暂不在全校范围内开放。

从数据影响对象及危害程度维度,南开大学对数据资产设立4个安全等级。

1级:数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用,不会对个人合法权益、组织合法权益造成危害。

2级:数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用,可能对个人合法

权益、组织合法权益造成轻微危害。

3级：数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用，可能对个人合法权益、组织合法权益造成一般危害。

4级：数据一旦遭到篡改、破坏、泄露或者非法获取、非法利用，可能对个人合法权益、组织合法权益造成严重危害，但不会危害国家安全或公共利益。

每个字段都有多个维度安全等级的标注，其中涉及由多个字段计算得到的字段的安全等级默认按照就高从严原则定级，同时也支持人为干预调整安全等级。

### 3.5 数据标签化

南开大学在数据资产目录之外，同时给数据设立了多维度数据标签，数据标签包括数据分类、数据安全等级、数据权属、业务属性、数据存储、技术属性等，如图5所示。数据资产目录结合数据标签可为

用户提供灵活的数据组织方式。

### 3.6 数据资产目录开放

南开大学建有校级的全量数据中台，依托于全量数据中台的数据资产管理子系统进行数据资产目录、数据安全等级和数据标签的管理，同时依托于全量数据中台的数据开放子系统对数据目录进行全校范围内的开放共享。

南开大学数据开放系统除以上功能外，还推出了一些常用的数据应用套餐供数据使用者选择，数据应用套餐直接省去了数据检索过程，数据使用者可直接申请获取数据接口。南开大学常用的数据应用套餐见表3。

### 3.7 数据资产编目实施效果

南开大学数据资产目录已基本建设

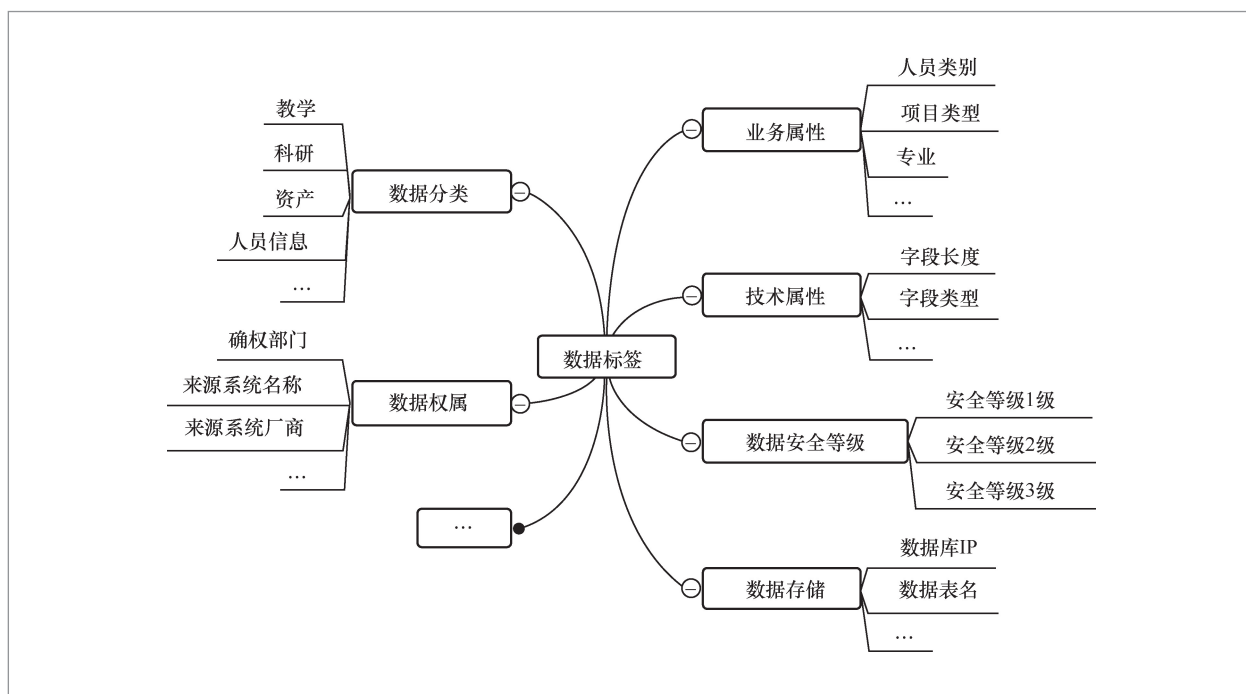


图5 南开大学数据标签

表 3 南开大学常用的数据应用套餐

套餐名称	涉及数据集
系统建设基础数据	本科生基本信息、研究生基本信息、教职工基本信息、组织机构数据
在校在籍学生	在校在籍本科生、在校在籍研究生、在校在籍留学生
在岗职工	在岗各类教职工数据、在岗各类其他人员数据
教师考核	教师工作量、科研项目、科研成果、出版物、获奖
学生评优	学生成绩、学生奖助勤数据、体育运动

完成,目录涵盖了人员信息、组织机构、就业、宿管、资产、财务、一卡通、体育部、双一流、科研、图书馆等大类,如图6所示。通过数据开放子平台对数据资产目录进行开放,已为学校30余个业务系统提供数据视图/表接口、API、DCI等数据接口。

相较于没有数据资产目录,数据资产目录建成后的数据管理和开放共享有以下几个突出优势。

- 更方便数据使用者了解学校数据的真实情况。建设数据资产目录前按照信息标准维护数据,信息标准中的数据表及

字段存在大量学校实际不使用的字段,不了解数据真实情况的话,容易造成数据误解。在构建数据资产目录后,所见即所得,数据资产目录中都是学校实际拥有的数据资产,不存在空字段、失效字段等,给数据使用者带来便利。

- 数据资产预览和检索更方便。之前数据使用者通过与业务部门或者数据中心线下沟通数据情况来确认需求数据。数据资产目录建设完成后给数据使用者提供了可视化的平台,并且数据资产目录支持模糊查询和多条件联合查询(数据目录、数

资产名称	所属主目录	资产编码	资产属性	资产类型	所属副目录数	所属副目录	所属标签数	所属标签	创建时间	状态
基本	人员信息/教职...		结构化	动态裁...	1	人员信...	0		2022-...	草稿
图书借...	图书馆/图书借...		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
读者欠...	图书馆/读者信息		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
读者存...	图书馆/读者信息		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
读者逃...	图书馆/读者信息		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
读者预...	图书馆/读者信息		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
读者委...	图书馆/读者信息		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
读者信息	图书馆/读者信息		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
图书借...	图书馆/图书借...		结构化	动态裁...	1	图书馆...	0		2022-...	草稿
图书借...	图书馆/图书借...		结构化	动态裁...	1	图书馆...	0		2022-...	草稿

图 6 南开大学数据资产目录

据标签以及字段取值), 数据使用者能够自主地精准定位所需数据。

- 支持多视角数据目录展示形式。数据资产目录有默认展示形式, 如果默认模板和业务需求理念不一致, 可结合数据标签, 支持数据使用者自主灵活配置所需的目录形式。

- 设有多个常用数据应用套餐, 以供数据使用者更便捷地获取所需数据。

- 通过数据安全等级控制, 从之前的粗犷式数据开放转型为从数据开放内容、开放数据量、开放对象等多层面的精细化管理, 做到了开放有依据、共享有审批。

- 数据接口安全管理化。数据接口管理平台化, 实现接口账号管理、开放对象授权、开放时间授权、多形式数据脱敏等管理, 南开大学数据开放接口配置和数据开放对象配置分别如图7、图8所示。

The screenshot shows a configuration window titled '查看' (View) with a close button 'X'. It is divided into two sections: '配置接口信息' (Configure Interface Information) and '配置接口参数' (Configure Interface Parameters). Under '配置接口信息', there are fields for '接口名称' (Interface Name) set to '教室基本信息', '信息目录分类' (Information Directory Category) set to '教学管理 / 教学', and '更新频率' (Update Frequency) set to '每天'. Under '配置接口参数', there are fields for '开放截止日期' (Open End Date) set to '2023-10-01' and '数据来源' (Data Source) set to '职能部门 / 教务处'. At the bottom, there are radio buttons for '开放对象' (Open Target): '对所有用户公开' (Open to all users), '部分用户公开, 点击选择' (Open to some users, click to select), and '不公开' (Not open). A '描述' (Description) field is also present at the bottom.

图7 南开大学数据开放接口配置

The screenshot shows a configuration window titled '开放对象' (Open Object). It has three tabs: '人员' (Personnel), '机构' (Institution), and '机构管理员' (Institution Administrator). The '人员' tab is active. On the left, there is a search box '请输入关键字' (Please enter keywords) with a dropdown menu showing '第三方厂商' (Third-party vendor), '教学科研' (Teaching and research), and '职能部门' (Functional department). In the center, there is a '备选列表' (Candidate list) with columns for '用户身份' (User identity), '用户角色' (User role), and a search input '请输入姓名' (Please enter name). Below this is a table with columns '用户名' (Username), '姓名' (Name), '用户角色' (User role), and '用户身份' (User identity). The table contains one row with a checkbox, a blurred name, a blurred role, and '个人用户' (Individual user). On the right, there is an '已选列表' (Selected list) with a '查询' (Query) button and a '删除已选' (Delete selected) button. Below the table, there is a row with a checkbox, a blurred name, '个人用户' (Individual user), and '教职工' (Faculty).

图8 南开大学数据开放对象配置

南开大学是国家“双一流”重点建设的综合性研究型大学,学校数据资产在一定程度上体现了国内高校的数据资产特性。虽然本文提出的数据资产编目模型和工作框架并没有在其他高校实际落地使用,但是数据资产目录的建设思路、模型及工作框架对其他高校具有一定的借鉴作用。

## 4 结束语

在教育信息化、数字校园大环境下,高校数据资产潜在的巨大价值逐渐被大家认可。数据资产价值的体现就是能够持续赋能各类应用和信息系统,同时支撑各种要求的数据分析,为领导层和管理者提供有效的决策依据,推动高校业务不断发展。因此基于数据应用需求的数据资产目录建设成为数据治理工作中的重要环节。当前对于高校数据资产编目工作,各高校都处于各自摸索状态,尚缺乏权威的指导性文件和通用标准。

在了解政务、企业数据资产编目原理的前提下,结合高校数据资产特性,通过分析高校数据资产编目的功能和作用,结合网络安全和数据安全,以应用需求为切入点,构建了适用于高校的数据资产编目模型和具体的工作框架。高校数据资产编目模型由数据资产目录建设、数据安全、数据服务三部分相辅相成,互为支撑;同时形成了五步骤工作框架,包括数据盘点、数据分类、数据分级、数据标签、数据开放。该模型和工作框架在南开大学数据资产编目工作中具有良好的实用效果。因此该模型和工作框架可为各高校开展数据资产编目相关工作提供科学依据和方法指导。

## 参考文献:

- [1] 夏义堃, 管茜. 政府数据资产管理的内涵、要素框架与运行模式[J]. 电子政务, 2022(1): 2-13.  
XIA Y K, GUAN X. The connotation, element framework and operation mode of government data asset management[J]. E-Government, 2022(1): 2-13.
- [2] HANNILA H, SILVOLA R, HARKONEN J, et al. Data-driven begins with DATA; potential of data assets[J]. Journal of Computer Information Systems, 2022, 62(1): 29-38.
- [3] 车伟, 赵申. 供电企业数据盘点与数据目录构建研究[J]. 机电信息, 2019(36): 1-3.  
CHE W, ZHAO S. Research on data inventory and data catalogue construction of power supply enterprises[J]. Mechanical and Electrical Information, 2019(36): 1-3.
- [4] SHOBA U S, GOURI S. Digitisation of clinical information in the health sector—impediments and solutions—an overview[J]. Indian Journal of Science and Technology, 2011, 4(3): 255-258.
- [5] 彭雪涛. 美国高校数据治理及其借鉴[J]. 电化教育研究, 2017, 38(6): 76-81.  
PENG X T. Data governance in American universities and its reference[J]. e-Education research, 2017, 38(6): 76-81.
- [6] 贺建虎. 高校智慧校园数据治理架构设计与实现[J]. 信息技术与信息化, 2022(6): 157-160.  
HE J H. Design and implementation of data governance architecture for smart campus in colleges and universities[J]. Information Technology & Informatization, 2022(6): 157-160.
- [7] 王海亮, 刘亚翹. 高校数字化校园信息标准建设的研究[J]. 中北大学学报(社会科学版),

- 2015, 31(3): 116-120, 124.
- WANG H L, LIU Y H. Information standards construction based on digital campus of universities[J]. Journal of North University of China (Social Science Edition), 2015, 31(3): 116-120, 124.
- [8] 李雨霏, 刘海燕, 闫树. 面向价值实现的数据资产管理体系构建[J]. 大数据, 2020, 6(3): 45-56.
- LI Y F, LIU H Y, YAN S. Construction of a value-oriented realization of data asset management system[J]. Big Data Research, 2020, 6(3): 45-56.
- [9] 宋锴业, 徐雅倩, 陈天祥. 政务数据资产化的创新发展、内在机制与路径优化: 以政务数据资产管理的潍坊模式为例[J]. 电子政务, 2022(1): 14-26.
- SONG K Y, XU Y Q, CHEN T X. Innovative development, internal mechanism and path optimization of government data capitalization—taking Weifang model of government data asset management as an example[J]. E-Government, 2022(1): 14-26.
- [10] 戴炳荣, 闭珊珊, 杨琳, 等. 数据资产标准研究进展与建议[J]. 大数据, 2020, 6(3): 36-44.
- DAI B R, BI S S, YANG L, et al. Research status quo and suggestions on data assets standardization[J]. Big Data Research, 2020, 6(3): 36-44.
- [11] 叶雅珍, 刘国华, 朱扬勇. 数据资产化框架初探[J]. 大数据, 2020, 6(3): 3-12.
- YE Y Z, LIU G H, ZHU Y Y. An initial exploration on framework of data assetization[J]. Big Data Research, 2020, 6(3): 3-12.
- [12] 魏建行, 刘远志, 罗超, 等. 基于数据中台的高校数据治理体系研究[J]. 信息技术与信息化, 2022(6): 98-101.
- WEI J X, LIU Y Z, LUO C, et al. Research on university data governance system based on data center[J]. Information Technology & Informatization, 2022(6): 98-101.
- [13] 傅柱, 徐绪堪, 何劲, 等. 基于U/C矩阵的概念设计知识管理应用探究[J]. 情报理论与实践, 2018, 41(12): 144-148, 154.
- FU Z, XU X K, HE J, et al. Knowledge management application of conceptual design based on U/C matrix[J]. Information Studies: Theory & Application, 2018, 41(12): 144-148, 154.
- [14] 李思男, 李宁, 李战怀. 多标签数据挖掘技术: 研究综述[J]. 计算机科学, 2013, 40(4): 14-21.
- LI S N, LI N, LI Z H. Multi-label data mining: a survey[J]. Computer Science, 2013, 40(4): 14-21.
- [15] 陈树年. 网络信息分类法研究[J]. 现代图书情报技术, 2002(3): 54-57, 80.
- CHEN S N. Study on Internet information classification[J]. New Technology of Library and Information Service, 2002(3): 54-57, 80.
- [16] 周健雄, 梁孟. 基于国家标准要求的政务数据共享安全架构研究[J]. 信息安全与通信保密, 2021, 19(6): 24-31.
- ZHOU J X, LIANG M. Research on government data sharing security architecture based on national standards[J]. Information Security and Communications Privacy, 2021, 19(6): 24-31.
- [17] 盛小平, 郭道胜. 科学数据开放共享中的数据安全治理研究[J]. 图书情报工作, 2020, 64(22): 25-36.
- SHENG X P, GUO D S. Research on data security governance in open sharing of scientific data[J]. Library and Information Service, 2020, 64(22): 25-36.

## 作者简介



刘豆豆(1990-),女,南开大学大数据管理中心工程师,数据管理科科长,主要研究方向为信息与通信工程。



焦宝臣(1982-),男,博士,南开大学大数据管理中心高级工程师,数据管理科科长,主要研究方向为计算机应用技术。

收稿日期: 2023-02-24

通信作者: 焦宝臣, bcjiao@nankai.edu.cn

基金项目: 2022年度高等教育科学研究规划课题(No.22XX0102); 2021年中国高校产学研创新基金(No.2021KSA01001); 河北省重点研发计划项目(No.22370701D)

**Foundation Items:** Scientific Research Planning of Higher Education in 2022(No.22XX0102), Industry-University-Research Innovation Fund of China Universities in 2021(No.2021KSA01001), Key Research and Development Projects in Hebei Province(No.22370701D)