

# NSTL科技文献资源分类体系构建研究\*

邓盼盼 李军莲 冀玉静 任慧玲 葛红梅  
(中国医学科学院医学信息研究所, 北京 100020)

**摘要:** 相对科学的分类体系是实现科技文献资源学科特征有效揭示的重要依据。面向国家科技图书文献中心(National Science and Technology Library, NSTL)资源集成与知识服务的总体需求,通过主体及重点学科揭示、多体系融合、多维同位类设置、多学科列类、新兴学科类目扩展、综合性类目设置、双重语义编码等方法,编制可基本覆盖NSTL主体资源的NSTL科技文献分类体系,共包含61个基本大类,类目深度为4~5级,类目数达5 350个。同时建立与现用核心分类体系的映射关系,支撑资源分类数据规范处理、馆藏目录系统服务等方面的应用。该分类体系可支撑国家重点扶持产业资源保障分析与学科布局评估。

**关键词:** 分类体系; 科技文献; 数据规范; 资源保障; 交叉学科

中图分类号: G25; TP391 DOI: 10.3772/j.issn.1673-2286.2024.08.001

引文格式: 邓盼盼, 李军莲, 冀玉静, 等. NSTL科技文献资源分类体系构建研究[J]. 数字图书馆论坛, 2024, 20(8): 1-8.

国家科技图书文献中心(National Science and Technology Library, NSTL),由中国科学院文献情报中心、中国科学技术信息研究所、机械工业信息研究院、冶金工业信息标准研究院、中国化工信息中心、中国农业科学院农业信息研究所、中国医学科学院医学信息研究所、中国标准化研究院国家标准馆和中国计量科学研究院文献馆9个文献信息机构组成,是基于网络环境的科技文献信息资源服务体系,采集、收藏和开发理、工、农、医各学科领域的科技文献资源<sup>[1]</sup>。学科是信息资源重要的特征之一,相对科学的分类体系是实现资源学科特征有效揭示与服务的重要依据,可广泛用于分类组织、学科导航、筛选过滤、分面检索、科技分析与评价等诸多领域。目前, NSTL尚未建立统一的资源分类体系,成员单位编目采用不同的分类体系或采用同一分类体系的不同版本,造成分类数据分类体系多样化、多版本共存、未作统一规范,不利于整体资源的快速导航浏览和检索,使得资源分类揭示和发现服务受限。新技术条

件下科研工作者对文献信息和知识的需求发生改变,相关领域文献缺乏、获取困难转变为文献过多、难以选择有用文献的难题,在海量资源中快速、高效、准确地定位有价值的科技文献成为关键。网络信息环境下,分类是科技文献的指向标,分类法不再局限于传统分类法理论,用户对象、体系结构、编制方式更趋灵活,向分类主题一体化、人机协同化、兼容与互操作的趋势发展<sup>[2]</sup>。鉴于此,面向NSTL科技文献资源构建一个相对科学、规范、实用的分类体系,有利于对多分类体系共存的资源分类数据进行规范、揭示和组织,具有重要的意义。

## 1 资源分类问题分析

### 1.1 资源分类体系多样化

资源数据采集、编目、存储、组织、服务等各个环

收稿日期: 2024-03-01

\*本研究得到国家科技图书文献中心专项任务“NSTL统一资源分类表数据更新与维护”(编号: 2022XM20)、中国医学科学院医学与健康科技创新工程重大协同创新项目“生物医学文献信息保障与集成服务平台”(编号: 2021-I2M-1-033)资助。

节存在分类体系多样化现象。成员单位在不同历史阶段结合自身学科定位与收藏特点,采用不同的分类体系或采用同一分类体系的不同版本,如《中国图书馆分类法》(以下简称《中图法》)第二版、第三版、第四版、第五版及《中国科学院图书馆图书分类法》第三版等。各单位数据提交至联合目录系统<sup>[3]</sup>后,编目数据中也存在多个版本的分类号。NSTL网络服务系统最终仅采用前几级的部分类目进行资源导航,现用分类体系(《中图法》第四版、第五版)中医学科学和农业科学领域分别有大类“R 医药、卫生”“S 农业科学”,因此导航时仅分别设置了大类“医学、药学、卫生”和“农林牧渔”,颗粒度较粗、揭示不充分、发现服务受限。

## 1.2 资源揭示不充分

通过分析现用分类体系下各类目的资源规模可知,各领域资源规模存在差异。部分过时、冷门或非主体学科无资源或有较少资源,如“K851 考古学”。部分重点发展的学科虽然资源规模大,但类目隐藏在较深层级,无法快速聚焦,如人工智能领域资源规模正随着学科的快速发展和迅速增长,但在现用分类体系中相关资源隐藏在类目的第四级[TP18 人工智能理论subClassOf(父类)TP1 自动化基础理论subClassOf TP 自动化技术、计算机技术subClassOf T 工业技术],细分子类也只有3个,即“TP181 自动推理、机器学习”“TP182 专家系统、知识工程”“TP183 人工神经网络与计算”,不易揭示和利用。

## 1.3 交叉学科资源分散

传统分类体系中交替类目主要存在以下几种情况:①两个类目名称和含义均相同,如“[P351.1]大气结构”与“P421.3 大气结构”、“[P351.3]大气辐射”与“P422 大气辐射”;②两个类目名称存在差异,内涵相近、相关或更加宽泛,如“P351 高层大气物理”与“P401 大气物理学”、“[P351.2]大气热状态和能源”与“P431<sup>↑</sup>.2 静力和准静力的热力状态和过程、热平衡”、“[P351.4]大气光化反应”与“P402 大气化学”、“[P351.5]大气振荡”与“P433 大气动力学”,这就造成高层大气物理及其子领域的资源分散在多处、无法汇聚;③存在交替类目对应多个宜入类目的情况,如

“[X949]航空、航天安全”宜入V268.6、V328、V468、V528等有关各类。由于宜入类目和交替类目内涵并不完全相同或存在一对多情况,交替类目指示的重点学科资源只能分散在这些相近、相关或宽泛的宜入类目资源中,无法单独聚焦揭示。另外,由于各单位编目规则等差异,部分交叉学科资源归入宜入类目,部分资源归入交替类目,这也造成了交叉学科资源分散、无法集中。

## 1.4 难以揭示新学科资源

学科知识体系及结构有相对稳定的部分,也有不断变化的内容,现用分类体系久未更新,不能体现学科的发展变革、揭示新学科资源。由于在现有分类体系中新兴学科或热门主题没有对应的类目,新学科资源只能归入宽泛的大类或相近的类目,无法充分揭示和快速定位,如物联网相关期刊《IEEE Internet of Things Journal》《Journal on Internet of Things》在现有分类体系中没有对应的类目,只能归至“TM 电工技术”或“TP 自动化技术、计算机技术”这些比较宽泛的类目。

## 2 科技文献分类体系构建思路

分类体系总体构建思路是从现用资源分类体系现状、资源学科分布与揭示、学科发展演变与需求等方面进行分析,明确分类表编制目标与定位。类目提取与融合有两条路径:①根据资源种类和学科分布,遴选分类体系构建所需的核心素材,如学科分类与代码<sup>[4]</sup>、《中图法》分类<sup>[5]</sup>、期刊分类<sup>[6]</sup>、杜威十进制分类<sup>[7]</sup>、医学专业分类<sup>[8]</sup>、美国国立医学图书馆分类<sup>[9]</sup>、国家自然科学基金申请代码<sup>[10]</sup>、国家社科基金项目申报数据代码<sup>[11]</sup>、科技知识组织体系范畴<sup>[12]</sup>等;②明确提炼主体学科及相对稳定的知识内容,从国家战略规划<sup>[13-14]</sup>、领域科技发展报告<sup>[15]</sup>、高校学科专业建设及人才培养布局<sup>[16-18]</sup>等多个层面凝练重点、新兴、交叉和前沿方向,基于多体系融合确定一级类目。在此基础上,分领域形成分类顶层框架,主要包括基础科学、工程技术科学、农业科学、医学科学、人文社会科学5类。根据分类体系素材遴选原则、综合编制原则、类目设置原则等,先通过揭示重点学科、多体系融合、多维设置同位类、多学科列

类、扩展新学科类目、设置综合性类目等进行类目层级关系的重构, 再通过领域专家协助进行基本大类、类目结构、关系、注释、编码系统、映射关系等模型设计与

内容建设, 着力解决重点学科、交叉学科、新兴学科的资源揭示问题, 建设可基本覆盖NSTL主体资源的科技文献分类体系。分类体系构建思路见图1。

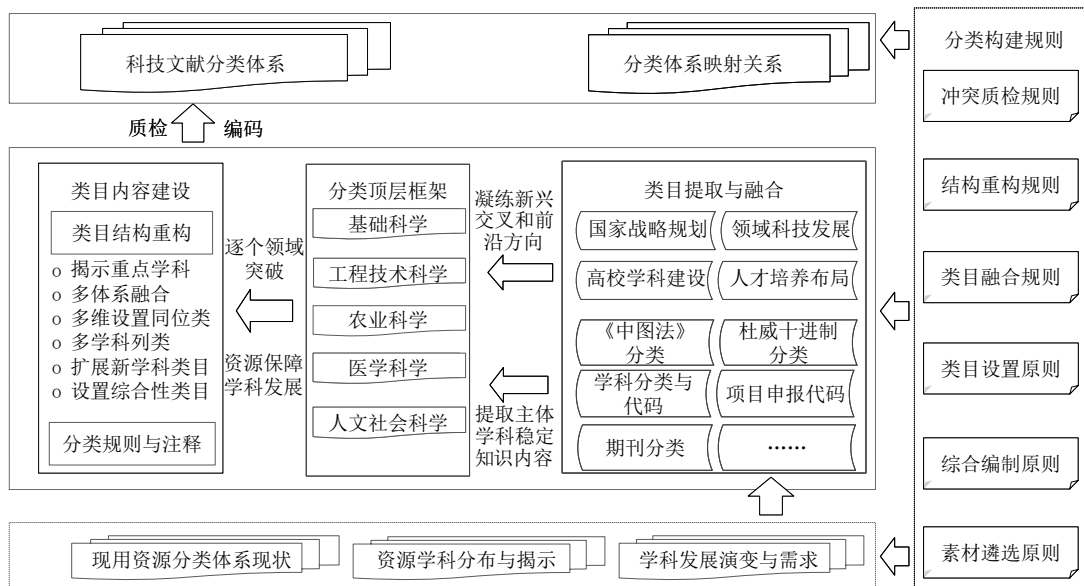


图1 分类体系构建思路

### 3 科技文献分类体系编制规则

#### 3.1 综合编制原则

(1) 文献保障原则。类目设置和编码以满足资源组织和服务的保障需求为基本目标, 列入分类体系的类目应具有一定规模的资源。分类体系应覆盖NSTL主体资源所属学科或主题领域, 考虑特色、交叉学科。

(2) 科学性原则。以学科分类为基础, 从总到分、从一般到具体编制分类体系。基本大类是对知识范畴所作的最概括、最本质的划分。根据学科本质特征及其相互联系, 设置相应类目并建立关系, 组成一个井然有序的分类体系。

(3) 系统完整性原则。分类体系应保证完整性和系统性, 类目间关系合理、正确。

(4) 可扩展性原则。应考虑动态的学科发展及资源变化, 充分揭示新兴学科、事物。

(5) 易用性原则。为了便于应用, 考虑目前编目人员与检索用户的使用习惯, 借鉴现用的多样化分类体系, 继承和延续学科体系中稳定的知识内容, 如同位类逻辑排列顺序。

(6) 合理深度原则。借鉴典型分类体系、主流数据库导航深度, 结合NSTL资源品种编目深度及深层级的资源规模比例, 制定合理的层级深度。基于统计发现, 71.84%的科技期刊、文集汇编、科技报告、会议文献等资源品种编目深度集中在前4级, 28.16%的资源品种编目深度为5~7级。因此, 无论是面向资源组织, 还是为了更高效地进行导航定位、检索聚类, 分类体系编制的层级深度都应尽量控制在4级左右。一级大类实现学科全覆盖, 其余层级纳入重点学科、突出新兴和交叉学科, 各类细分深度可结合各领域资源的分布情况进行适当调整。

#### 3.2 类目设置原则

(1) 类目相对独立。从学科角度进行类目设置时, 所设相应类目应相对独立, 具有独特的研究对象、研究内容、研究方法及其规律等学科基础; 类目可以是成熟稳定的学科, 也可以是有潜力分化为子学科、新学科的重点主题。对于资源量极少或学科关联性较强的特殊情况, 为了保障资源完整性、避免类目冗余, 可将相应学科合并, 进行类目设置。

(2) 类目名称规范。类目名称应当严格准确地表达单个类目的科学内涵和本质属性,应简明、规范、无歧义,具有权威性,兼顾时效性。考虑我国文化特色和汉语语言习惯,尽量复用我国或领域内熟知、当前最新适用的学科或主题名称作为类目名称。

(3) 类目结构一致。同一学科分类体系的类目应体现系统完整性、逻辑相关性及名称一致性,上位类目名称所表达的内涵应基本覆盖下位类目名称的含义,同类型的类目基本保持整体类目名称的一致性。

## 4 科技文献类目构建特点和方法

### 4.1 遵循文献保障原则揭示主体及重点学科

遵循文献保障原则,分类体系应覆盖NSTL实际馆藏文献的学科范围,根据文献主题和规模进行类目设置。首先,分析现用分类体系各类目下的资源规模,提取出资源(包括科技期刊、会议文献、科技报告、文集汇编等)品种总量为5个以上的类目3 000余个,这些类目涉及的学科领域能基本覆盖NSTL主体资源学科。然后,从资源规模和学科分类角度重新梳理分类体系层级结构,将资源较多、隐藏较深的重点学科或热门主题调整至比较靠前的层级,突出重点学科,如:将“TP24 机器人技术”从第四级调至目前分类体系的第二级“33.06 机器人科学与工程”,并扩充类目名称内涵;将冷门或非主体资源学科适当调整至深层级,进行类目合并或保障其有可入的上位类,以保障学科覆盖完整性,如简化人文社会科学类目及下位类,“90.45 文学与历史”合并了语言、文学、艺术、历史等学科。

### 4.2 基于多体系融合搭建分类顶层框架

在遴选的分类素材中,《中图法》分类<sup>[5]</sup>、期刊分类<sup>[6]</sup>、杜威十进制分类<sup>[7]</sup>等综合性图书分类法覆盖范围较广泛,虽未能展现新兴学科,但可用于获取学科体系及知识内容中相对稳定成熟的部分;医学专业分类<sup>[8]</sup>、美国国立医学图书馆分类<sup>[9]</sup>、科技知识组织体系范畴<sup>[12]</sup>等特定对象或技术领域的分类法颗粒度较细;国家自然科学基金申请代码<sup>[10]</sup>、国家社科基金项目申报数据代码<sup>[11]</sup>等科研项目代码则更新较快,能够体现科

学技术领域研究热点;学位授予和人才培养一级学科<sup>[16]</sup>、研究生教育学科专业<sup>[17]</sup>、高等学校本科专业<sup>[18]</sup>等学科分类则对已经发展为学科的主题进行了名称规范,类目名称多以“学”结尾,更符合学科的名称要求,可成为最后的类目名称参考。多个分类体系的类目及关系融合,有利于保障学科领域的覆盖完整性,除了主体学科外,也能兼顾部分新兴、热门学科。

由于现用分类体系层级深度达10级,为了缩短资源揭示路径,基于合理深度原则,将层级体系控制在4级左右。基于资源分布提取覆盖主体资源的3 000余个类目及其上层类目作为基础类目框架;将其与深度为3级的学科分类与代码、国家自然科学基金申请代码等进行全量融合,基于类目名称、词典等进行类目匹配,识别可以合并归一的类目,如“P941.6 极地”“D0615 极地科学”“D010105 冰冻圈地理学”“1706080 极地科学”,融合为一个类目后,分别继承来源类目的下位类目如“北极”“南极”。由于多个分类体系的层级深度不同,当融合后存在关系冗余冲突时,则保留比较扁平化的层级关系,如“14035 电磁学subClassOf(父类)140 物理学”与“O441 电磁学subClassOf O44 电磁学、电动力学subClassOf O4 物理学”,融合后的层级关系为“电磁学subClassOf物理学”“电磁学subClassOf电磁学、电动力学subClassOf物理学”,“电磁学”与“电磁学、电动力学”既存在上下位关系又存在兄弟关系,为了扁平化重构层级结构,仅保留“电磁学subClassOf物理学”,这样融合形成的层级关系有可靠的来源作为重要依据。通过计算机辅助进行分类体系自动融合后,由各领域专家结合各类目下的资源分布情况进行分类体系融合结果确认和层级结构重构,在保证系统完整性的前提下进行冷门学科类目的删减或合并。

### 4.3 多维设置同位类构建扁平化层级结构

类目层级关系以资源汇聚为目的,并非严格意义上的隶属关系。一方面,结合资源及学科发展等重构类目层级结构,如将深层级的“工程材料”升为一级类目“21 材料科学与工程”;另一方面,在符合科学逻辑的前提下,为了分类体系扁平化并更充分地进行资源揭示,允许从多个角度进行同位类设置,并服从同一个上位类,由上位类直接划分出来的下位类可以是子学科或该学科的研究方面。理论上,划分出的各子类应相互排

斥,但考虑到体系深度和学科间的交叉或相关性,同位类间也允许存在部分重叠或交叉现象,如“土木建筑材料”分别基于材料类型和材料功能进行了子类划分(见图2)。考虑到用户习惯和易用性,采用惯用次序与逻辑次序相结合的同位类排序方法,以求类目简洁、实用。

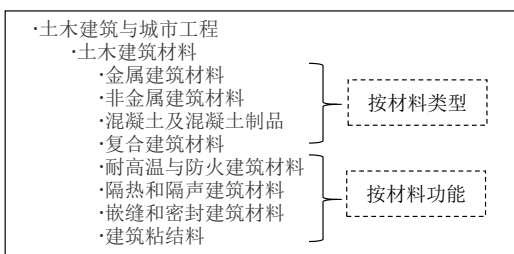


图2 多维设置同位类示例

#### 4.4 多学科列类汇聚交叉学科资源

为了解决由交替类目或编目规则造成的资源分散问题,采用多学科列类设置亦入类目,汇聚交叉学科资源。交叉学科或主题从不同角度划分可以服从多个学科体系,可以在多个学科体系下设置内涵相同的类目,这些类目称为亦入类目,具有相同的ID和类目名称,具有多个上位类、类号,在不同的学科体系下能够实现资源互通,增加资源揭示路径。亦入类目示例见表1,其中“制药化学”是“第二部分 工程技术科学”领域中“30 化学工程与技术”学科与“第四部分 医学科学”领域中“78 药学”学科的远距离交叉。

表1 亦入类目示例

亦入类目1	亦入类目2	交叉学科类型
11 地质学 11.51 海洋地质学 亦入: 09.21	09 海洋科学 09.21 海洋地质学 亦入: 11.51	同领域学科, 近距离交叉
13 生物科学 13.09 遗传学 13.09.45 植物遗传学 亦入: 13.30.09	13 生物科学 13.30 植物生物学 13.30.09 植物遗传学 亦入: 13.09.45	同领域学科, 内部交叉
30 化学工程与技术 30.57 制药化学 亦入: 78.06	78 药学 78.06 制药化学 亦入: 30.57	不同领域学科间, 远距离交叉

#### 4.5 基于学科发展扩展新兴学科类目

已有学科的深耕细作促使学科分工不断细化、深化,形成更为精细化、差异化的学科类别;新的理论、

技术、方法与社会需求紧密结合,相互作用,推动着学科的分化、新生与交叉融合。随着学科的新生、交叉与发展演变,用户关注度及资源保障规模也发生着相应的变化。为了更好地进行资源揭示,学科资源保障与分类体系也应遵循学科发展规律及应用需要不断革新,从国家顶层发展战略规划、领域科技发展、高校学科建设、专业设置与人才培养、资源专家推荐等各个层面,凝练新兴、交叉和前沿方向,挖掘和细化有潜力的热门学科或主题,突出重点学科重点领域前瞻布局,扩展新类目,使得分类体系跟上学科发展。同时,标识表示新兴学科或主题类目,可以为新学科资源的汇聚和保障分析奠定基础。新类目示例见表2。

表2 新类目示例

领域	新类目示例
基础科学	量子群和量子代数、算子理论与算子代数、微纳光学与光子学、波导光学与集成光、空间遥感、全息实景导航、量子导航定位、三维数字成像、水下导航定位、海底资源开发、海洋灾害与防灾减灾
工程技术科学	纳米材料、深地资源开采、机械微纳制造、机械与装备增材制造、机械仿生制造、机械仿生学、微机电工程、机器人科学与工程、3D打印建筑、核安全工程、智慧城市、新农村、城市群、新能源汽车
农业科学	农业能源工程、农业装备工程、智能农业装备、农机自动驾驶及监控、农用航空智能作业、智能灌区建设、智能农业机器人、农作物机械化生产、农业水资源安全利用、设施农业、智慧农业、农业转基因生物安全学、作物基因工程育种、作物良种繁育学、作物生物育种、植被恢复与生态重建、人工林定向培育、退化土地造林、能源林培育、碳汇造林、农业资源与环境、节水农业、生态农业
医学科学	群医学、护理智能机器人、危急重症护理学、再生医学、纳米医学、移动医学和远程医学、神经医学工程、细胞工程学、生物药学、海洋药理学、医学技术、康复治疗技术学、呼吸治疗学、基因组学

#### 4.6 综合性类目收录组合多学科资源

领域通用性类目或者多个学科组合的类目(如“自然科学总论”“数理科学和化学”“天文学、地球科学”“医药、卫生”)下存在一定规模的综合性资源,也确实存在一些收录文献范围较广泛的资源,涉及多个相关学科,如《Science》涉及生物学、化学、物理学、地球科学、计算机科学等领域,《Nature》涵盖生命科学、自然科学、临床医学等领域。当前分类体系中的一级大类基本相对独立,考虑到应用的需求和综合性资源的归属,分别为5个部分分别设置综合性类目,即“基础科学综合”“工程技术综合”“农业科学综合”“医学科学综合”“人文社会科学综合”。

## 4.7 双重编码揭示领域和语义信息

编码即标记符号,可以用数字、字母、特殊的符号或它们之间的组合来表示,常用编码形式为纯数字或数字、符号、字母混合构成顺序码、随机码。为了揭示更加丰富的语义信息、面向不同的应用场景,该分类体系采用类号和类目唯一标识符双重编码机制。①采用数字加符号、层累标记制的类号编码机制,每两位数字表示一个层级,级别之间采用“.”分隔。这种编码机制便于显示类目之间的隶属、并列关系,可以明确类目在层级结构中的位置和先后次序,有利于应用前端类目编排和展示。编码时采取空号法间隔编号,依次为01、03、06、09、12等,未来可根据学科发展进行类目细分和新类目增补,类目体系结构具有可扩充性和灵活性。②由于亦入类目会存在两个或多个类号,在类目关系存储及集成应用时会造成一定的数据冗余。为了便于分类数据存储、支撑资源组织和保障布局分析等应用,同时采用类目唯一标识符,由揭示领域信息的1个字母和5位数字顺序码组成:基础科学(01~19)采用字母Z,从Z00001开始顺序编码;工程技术科学(20~49)采用字母G,从G00001开始顺序编码;农业科学(50~69)采用字母N,从N00001开始顺序编码;医学科学(70~89)采用字母Y,从Y00001开始顺序编码;人文社会科学(90~99)采用字母S,从S00001开始顺序编码。这种面向计算机的编码模式中,每个类目,包括亦入类目,有且仅有一个唯一标识符,有利于对类目进行技术标识和区分,减少数据存储的冗余度,在资源编目、存储和服务过程中可以快速准确定位资源领域及归属机构。

## 4.8 分类表类目结构

类目结构由类目、关系、属性构成。类目包括类目唯一标识符、类目名称、类号;关系包括等级关系、并列关系、相关关系、亦入类目及与其他分类体系间的映射关系;属性主要包括注释和来源信息,其中规则注释标明了类目的内涵及使用方法。分类表类目结构示例见表3,其中:“11 地质学”具有类号11和唯一标识符Z00923;“11 地质学”与“11.24 矿物学”为等级关系,互为上下位类;“11.24 矿物学”与“11.27 岩石学”为并列关系,互为同位类;“11.36 水文地质学”亦入类号为“12.33 水文地质学”;“11.45 地震地质学”有相关类目“10.01.09.12 地震与地球构造”;“11.39 工程地质学”具有规则注释“动力地质及工程地质作用(风化、地面沉降)、区域工程地质等入此”。

“11.39 工程地质学”具有规则注释“动力地质及工程地质作用(风化、地面沉降)、区域工程地质等入此”。

表3 分类表类目结构示例

类目	结构示例
11 地质学	类号11, 类目唯一标识符Z00923
11.24 矿物学	与“11 地质学”为等级关系, 互为上下位类
11.27 岩石学	与“11.24 矿物学”为并列关系, 互为同位类
11.30 矿床学与矿相学	映射关系: P61 矿床学[来源:《中图法》分类]; TD11 矿床学[来源:《中图法》分类]; D0205 矿床学[来源:国家自然科学基金申请代码]; 1705024 矿床学与矿相学[来源:学科分类与代码]
11.36 水文地质学	亦入类目: 12.33 水文地质学
11.39 工程地质学	规则注释: 动力地质及工程地质作用(风化、地面沉降)、区域工程地质等入此
11.45 地震地质学	相关类目: 10.01.09.12 地震与地球构造

## 5 科技文献分类体系构建结果与应用

### 5.1 分类体系构建结果

目前已建成的NSTL科技文献分类体系(见表4)共分为5个领域、61个基本大类,深度为4~5级,类目数达5 350个。其中:基础科学一级类包含13个基本大类,类目数为1 317个;工程技术科学一级类包含21个基本大类,类目数为2 008个;农业科学一级类包含11个基本大类,类目数为667个;医学科学一级类包含11个基本大类,类目数为1 089个;人文社会科学一级类包含5个基本大类,类目数为269个。为了简化应用,本分类体系在使用时不再进行复分、仿分。

### 5.2 分类体系映射与应用

为了实现资源分类数据的规范化、标准化,依托现有术语平台和工具,基于复用类目来源、关系继承、词形特征及相似度等方法,建立与《中图法》第四版、第五版等现用核心分类体系的映射关系。为避免数据冗余,每个类目仅建立最有效的映射关系,推荐顺序依次为精确匹配、近似匹配、上位匹配。理想情况下,源类目均建立与目标类目的映射关系,从而有助于历史数据的规范;针对其他未映射源类目,基于词形特征或相似度进行映射关系推荐<sup>[9]</sup>。分类体系映射示例见图3,其中:设置“26.24 机械与装备制造”类目时,参考了源类

目“TH16 机械制造工艺”，二者名称和内涵有差异，因此基于类目参考源建立映射关系“TH16 机械制造工艺SM（近似匹配）26.24 机械与装备制造”；基于关系继承规则，二者可以继承彼此的下位概念，建立上位匹配“TH166 计算机集成制造BM（上位匹配）26.24 机械与装备制造”或“TH16 机械制造工艺NM（下位匹配）26.24.15 机械与装备增材制造”。目前分类体系电子版<sup>[20]</sup>通过服务系统面向NSTL及各成员单位提供服务，支持类目多维检索、树结构导航、类目使用反馈、新类目推荐等。

表4 NSTL科技文献分类体系

领域	基本大类
基础科学	基础科学综合、数学、物理学、化学、天文学、地理学、测绘学、大气科学、海洋科学、地球物理学和化学、地质学、水文学、生物科学
工程技术科学	工程技术综合、材料科学与工程、计量和标准、仪器仪表、矿业工程、冶金工程与金属、机械工程、能源科学与动力工程、核科学与技术、石油与天然气工程、化学工程与技术、电工技术、电子与通信工程、自动化与计算机信息科学、轻工与纺织、土木建筑与城市工程、水利工程与水资源、交通运输工程、航空与航天、环境与安全、军事技术与武器工业
农业科学	农业科学综合、农业工程、植物保护学、农艺与作物学、园艺学、林学、畜牧学、草业与草原、兽医学、水产学、农业资源与环境
医学科学	医学科学综合、公共卫生与预防医学、基础医学、临床医学、口腔医学、中医学、特种医学、护理学、药学、医学技术、生物医学工程
人文社会科学	人文社会科学综合、政治、法律与军事、经济与金融、管理科学、科学传播与知识服务

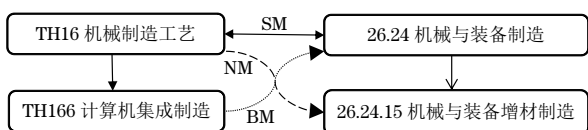


图3 分类体系映射示例

NSTL联合目录系统汇聚了NSTL及其9个成员单位采集的理、工、农、医各领域印本、电子和开放获取资源的书目信息。对于历史数据，遵循尽量全自动转换、减少人工干预原则，基于分类体系及与《中图法》第四版、第五版的映射关系，将资源原始编目类号与《中图法》第五版类号进行匹配：如完全相同，则基于与《中图法》第五版的映射关系，直接赋予资源NSTL类号；如果出于复分等原因没有完全相同的类号，则将资源原始编目类号从右往左依次去掉1个字符进行匹配，直至可以精确匹配到《中图法》第五版类号，然后基于

映射关系赋予对应的NSTL类号；如果取至最后1位仍没有对应类号，则标识为“类号可能有误”。对于增量数据，在进入系统时，基于各单位使用分类版本（《中图法》第四版、第五版），针对停用类目、引导性类目等进行提示，并进行分类自动转换赋类。这样，历史数据和增量数据中的每个资源均拥有NSTL类号，后续将随着业务流进入集成与服务等环节。此外，NSTL科技文献分类体系也支撑了NSTL馆藏目录服务系统<sup>[21]</sup>的资源导航、资源检索、结果聚类与学科资源保障分析等服务。

## 6 结语

为更好满足资源集成与知识服务的总体需求，本研究借鉴国内外资源分类的先进成果及技术，以学科分类为基础、资源保障为重点，遵循科学性、可扩展性、易用性等原则，编制了可基本覆盖NSTL主体资源的NSTL科技文献分类体系，建立了与现用核心分类体系间的映射关系，着力解决了重点学科、交叉学科、新兴学科的资源揭示问题，以促进资源数据的标准化、规范化，支撑资源组织与服务。未来将探索新兴学科资源发现、国家重点扶持学科与技术攻关方向的资源保障评估等应用。该分类体系的构建规则与方法同样适用于其他同类型分类体系的编制，其中重点学科、交叉学科、新兴学科相关类目可为其他分类体系更新提供参考。

## 参考文献

- [1] 曾建勋. 基于发现系统的NSTL用户服务体系思考[J]. 情报杂志, 2020, 39 (11): 134-138.
- [2] 赵建国, 韩丽影. 分类表与主题表一体化集成研究:《军用分类主题映射表》集成构建[J]. 情报理论与实践, 2016, 39 (5): 86-89, 73.
- [3] 葛红梅, 丁道劲, 徐晶晶, 等. NSTL联合目录资源描述发展与思考[J]. 数字图书馆论坛, 2020 (8): 22-28.
- [4] 国家质量监督检验检疫总局, 中国国家标准化管理委员会.《学科分类与代码》国家标准第2号修改单: GB/T 13745—2009/XG2—2016[S]. 北京: 中国标准出版社, 2016.
- [5] 国家图书馆《中国图书馆分类法》编辑委员会. 中国图书馆分类法[M]. 5版. 北京: 国家图书馆出版社, 2010.

- [6] 国家图书馆《中国图书馆分类法》编辑委员会. 中国图书馆分类法期刊分类表[M]. 3版. 北京: 国家图书馆出版社, 2012.
- [7] Dewey decimal classification[EB/OL]. [2023-02-02]. <https://www.dewey.org/webdewey/login/login.html>.
- [8] 中国图书馆分类法编辑委员会, 中国医学科学院信息图书馆. 中国图书馆分类法·医学专业分类表[EB/OL]. [2023-02-02]. <http://cmesh.imicams.ac.cn/index.action?action=index>.
- [9] National Library of Medicine. National Library of Medicine classification[EB/OL]. [2023-02-02]. <https://classification.nlm.nih.gov/>.
- [10] 国家自然科学基金委员会. 国家自然科学基金申请代码[EB/OL]. [2023-02-02]. <https://www.nsf.gov.cn/publish/portal0/tab1440/>.
- [11] 全国哲学社会科学工作办公室. 国家社科基金项目申报数据代码表[EB/OL]. [2023-02-02]. <http://www.nopss.gov.cn/n1/2023/0525/c219561-32694765.html>.
- [12] 鲜国建, 孙巍, 赵瑞雪, 等. 科技知识组织体系范畴协同构建系统设计及实现[J]. 数字图书馆论坛, 2014(11): 26-31.
- [13] “中国工程科技2035发展战略研究”项目组. 中国工程科技2035发展战略·综合报告[R]. 北京: 科学出版社, 2019.
- [14] 中华人民共和国中央人民政府. 国务院关于印发《中国制造2025》的通知[EB/OL]. [2023-02-02]. [https://www.gov.cn/zhengce/content/2015-05/19/content\\_9784.htm](https://www.gov.cn/zhengce/content/2015-05/19/content_9784.htm).
- [15] 中国医学科学院. 中国医学科技发展报告(2021)[R]. 北京: 科学出版社, 2022.
- [16] 国务院学位委员会第六届学科评议组. 学位授予和人才培养一级学科简介[M]. 北京: 高等教育出版社, 2013.
- [17] 中华人民共和国中央人民政府. 国务院学位委员会 教育部关于印发《研究生教育学科专业目录(2022年)》《研究生教育学科专业目录管理办法》的通知[EB/OL]. [2023-02-02]. [https://www.gov.cn/zhengce/zhengceku/2022-09/14/content\\_5709785.htm](https://www.gov.cn/zhengce/zhengceku/2022-09/14/content_5709785.htm).
- [18] 中华人民共和国教育部. 教育部关于公布2022年度普通高等学校本科专业备案和审批结果的通知[EB/OL]. [2023-04-04]. [http://www.moe.gov.cn/srcsite/A08/moe\\_1034/s4930/202304/t20230419\\_1056224.html](http://www.moe.gov.cn/srcsite/A08/moe_1034/s4930/202304/t20230419_1056224.html).
- [19] 贾君枝, 崔西燕. 人物本体词表之间的互操作及分类体系构建[J]. 情报学报, 2019, 38(7): 731-741.
- [20] NSTL科技文献学科分类表[EB/OL]. [2024-02-23]. <http://cmis.imicams.ac.cn/nstl/>.
- [21] 国家科技图书文献中心馆藏目录服务系统[EB/OL]. [2024-02-23]. <https://catalogue.nstl.gov.cn/#/home>.

## 作者简介

邓盼盼, 女, 硕士, 助理研究员, 研究方向: 知识组织与服务。

李军莲, 女, 博士, 研究馆员, 通信作者, 研究方向: 医学信息资源组织与利用、医学知识组织系统建设与应用, E-mail: lijunlian@imicams.ac.cn。

冀玉静, 女, 学士, 副研究馆员, 研究方向: 医学知识标注、医学术语学。

任慧玲, 女, 硕士, 研究馆员, 研究方向: 医学知识组织、医学资源建设。

葛红梅, 女, 硕士, 副研究馆员, 研究方向: 图书馆信息组织、图书馆元数据建设。

## Construction of NSTL Scientific and Technological Literature Resource Classification System

DENG PanPan LI JunLian JI YuJing REN HuiLing GE HongMei

(Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing 100020, P. R. China)

**Abstract:** A relatively scientific classification system is an important basis for effectively revealing the disciplinary characteristics of scientific and technological literature resources. In response to the overall demand for resource integration and knowledge services of the National Science and Technology Library (NSTL), this article develops the NSTL scientific and technological literature classification system that basically covers the main resources of NSTL and consists of 61 basic categories with a depth of 4-5 levels and a total of 5 350 categories. The construction methods include main and key discipline disclosure, multi-system fusion, multidimensional homologous classes setting, multidisciplinary classes setting, emerging discipline categories expansion, comprehensive classes setting, and dual semantic encoding, etc. The classification system establishes a mapping relationship with the current core classification system, supporting the application of standardized processing of cataloging classification data and collection system services. It supports the analysis of resource security and disciplinary layout in national key areas.

**Keywords:** Classification System; Scientific and Technological Literature; Data Specification; Resource Guarantee; Interdiscipline

(责任编辑: 王玮)