

《军用分类主题映射表》的 总体设计与编制

赵建国, 周健

(中国人民解放军军事科学院, 北京 100091)

摘要:《军用分类主题映射表》是分类主题一体化的检索语言工具,既可以用于军事信息资源的分类标引,也可以用于军事信息资源的主题标引,实现分类与主题相结合的检索方式。本文简述了《军用分类主题映射表》的总体框架,阐述了《军用分类主题映射表》映射与审查过程,叙述了“军用分类主题映射表编制管理系统”功能结构。

关键词: 军用信息检索语言; 分类主题映射; 分类主题一体化

中图分类号: G254

DOI: 10.3772/j.issn.1673-2286.2015.05.005

《军事信息资源分类法》(简称《军分法》)和《军用主题词表》(简称《军表》)是供全军统一使用的两种主要信息检索语言工具。多年来,《军表》和《军分法》在全军图书、情报、档案、资料、文书和出版等领域广泛应用,为军事信息资源的科学组织和有效利用提供重要保障。

随着信息技术的飞速发展,当前信息检索语言分类主题一体化已在信息资源的组织和利用中显示出越来越重要的价值和作用。在《军表》和《军分法》的基础上,编制一部供全军统一使用的分类主题一体化信息检索语言工具,成为我军信息检索语言发展和信息化建设的紧迫任务。全军军用信息检索语言编管会办公室经过反复调研和论证,提交编制《军用分类主题映射表》(简称《映射表》)立项论证报告,并被列为全军军用信息检索语言编制管理工作“十二五”规划以及总装科研计划项目。

2015年3月,由全军军用信息检索语言编管会办公室牵头,总装科技信息研究中心、军事科学院军事图书资料馆、总参某部军用主题词表编管会办公室、国防大学图书馆等单位参加,共同组建的《映射表》编制工作

课题组历时5年,完成《映射表》的编制工作。

课题组调研了当前国内外信息检索语言分类主题一体化的做法和经验,根据两个概念集合之间可相互映射的原理,《映射表》以《军分法》和《军表》为基础,采用概念集合间相互映射原理和技术,研究制订映射规则,将《军表》的全部主题词映射到《军分法》相应类目,实现《军表》与《军分法》的全面融合及一体化集成,形成具有分类主题一体化功能的信息检索语言工具,是实现军事信息资源分类主题一体化同步标引与检索的基础条件和语言保障。

1 编制规则

为确保将《军表》中的主题词映射到《军分法》相应类目的科学性、准确性和一致性,课题组按照映射关系类型制订映射规则。该规则根据概念的内涵与外延等属性以及概念间的逻辑关系,将主题词至类目的映射关系分为类目对应主题词、类目等同标注主题词、类目组配标注主题词、扩展类目主题词、类目隶属主题词等5种关系。

1.1 主题词隶属于类目的基本规则

(1) 只需隶属《军表》中的正式主题词，非正式主题词作为入口词，从属于对其对应的正式主题词，由后台中的编制系统自动隶属。

(2) 一个主题词至少要隶属到1个类目。对具有多种属性，可以隶属多个类目的主题词，类目总数最多不超过3个。

(3) 应根据主题词的本义、本质属性，隶属到最主要的类目。即从分类隶属角度，对主题词的学科属性或事物属性进行分析对比，再根据主次关系，确定隶属类目。

1.2 主题词隶属于类目的细则

(1) 消冗隶属细则。类目对应主题词、扩展类目主题词和类目标注主题词中的类目等同主题词不再当作该类的类目隶属主题词。

例1：“A2121. 11军事著述”的类目对应主题词“军事著作”不再作为类目隶属主题词。

例2：“A00军事思想总论”的类目等同主题词“军事思想”不再作为类目隶属主题词。

(2) 恰当隶属细则。一上位类有时有若干下位类（包括扩展类目），凡下位类的类目隶属主题词不再作为上位类的类目隶属主题词。

例：类目“X9121用户设备”有扩展类目“X9121-71电话机”，其中“X9121-71电话机”的类目隶属主题词“可视电话机”、“录音电话机”不再作为上位类“X9121用户设备”的类目隶属主题词。

(3) 下位主题词隶属细则。具有相同上位词的主题词，并且上位词所隶属类目无下位类时，所有下位主题词应归入该类目。换言之，最下位类目的类目隶属主题词，如果仍有下位主题词，则所有下位词，包括直接、间接下位主题词，应隶属于该类目。

例1：类目“W2311.23地雷部件”再无下位类，其类目隶属主题词“防排装置”有1个下位主题词，故其下位主题词“诡计装置”应隶属于该类目。

例2：类目“V2121. 11-17化学手榴弹”无下位类，其类目隶属主题词“催泪手榴弹”有2个下位主题词，故其下位主题词“爆炸型催泪手榴弹”、“燃烧型催泪手榴弹”应隶属于该类。

(4) 收容类隶属细则。当一个上位类下设有若干具体下位类，同时还设置“其他”、“其他××”、“一般

性问题”、“××一般用语”等收容子类时，该上位类不再收类目隶属主题词。其类目隶属主题词要么归入具体子类，不能归入具体子类的其余主题词应隶属于收容子类。

例：B61军事学边缘学科
=军事科学+边缘学科

B6111军事领导学
军队领导

B6131军事预测学

技术发展预测；军事未来研究；科学预测；
未来学；预测科学

.....

B6199其他

作战运筹研究

(5) 上位类隶属细则。当一个上位类目下只设若干具体下位类目，未设任何“其他”、“其他××”、“一般性问题”、“××一般用语”等收容子类时，其类目隶属主题词可归入具体子类的应归入具体子类，不能归入具体子类的其余主题词可隶属于该上位类目。

例：P5131裁军

部分裁军；裁军会谈；裁军问题；局部裁军；全面裁军

P5131.11常规裁军

边境裁军；共同均衡裁军

P5131. 21核裁军

核裁军理论；核均势

(6) 交替类主辅类隶属细则。交替类目不设置类目隶属主题词，应将其隶属于交替注释所指向的“宜入××类目”。换言之，具有交替关系的两个类目，对其各自的类目隶属主题词应先归并、排序、去重，再将其隶属于交替主类，并清空交替辅类中的类目隶属主题词。

例：“[G5121. 8111]维护社会稳定”的交替类目注释“宜入H6141.11维护社会稳定”，类目隶属主题词全部应该归入H6141.11。

(7) 交叉类一致性隶属细则。交叉类目的类目隶属主题词应基本一致。换言之，具有交叉关系的两个类目，对其各自的类目隶属主题词应先归并、排序、去重，再将其分别隶属于两个交叉类目。

例：“H3161.2 9政治工作法规”与“L13政治工作法规、制度”交叉，两个交叉类各自的类目隶属主题词应一致。

(8) 交叉类差异性隶属细则。军事学科与相关学

科的子类互相交叉时, 类目隶属主题词的归属应该有所侧重。一般军事学科概念主题词入军事学科, 相关学科概念主题词入相关学科。

例: “爱国主义、党的建设理论、帝国主义、反对腐败、马克思主义、国际主义、民主建设、社会主义市场经济”等等归入相关学科“Y”有关各类, 而不入“L00 政治工作理论”类。

(9) 通用概念主题词隶属细则。只起聚类、组配或限定作用的通用概念主题词, 一律隶属归入“Z99通用概念”类。综合一般用语, 隶属归入“Z9911”。科技一艘用语, 隶属归入“Z9921”。

例: “办法; 管理”归入“Z9911综合一般用语”, “类型; 设计”归入“Z9921科技一般用语”。

1.3 体系结构

根据《军分法》的体系结构, 《映射表》由“编制说明”、“分类主题词表”、“主题词—分类号对应索引”三部分组成。

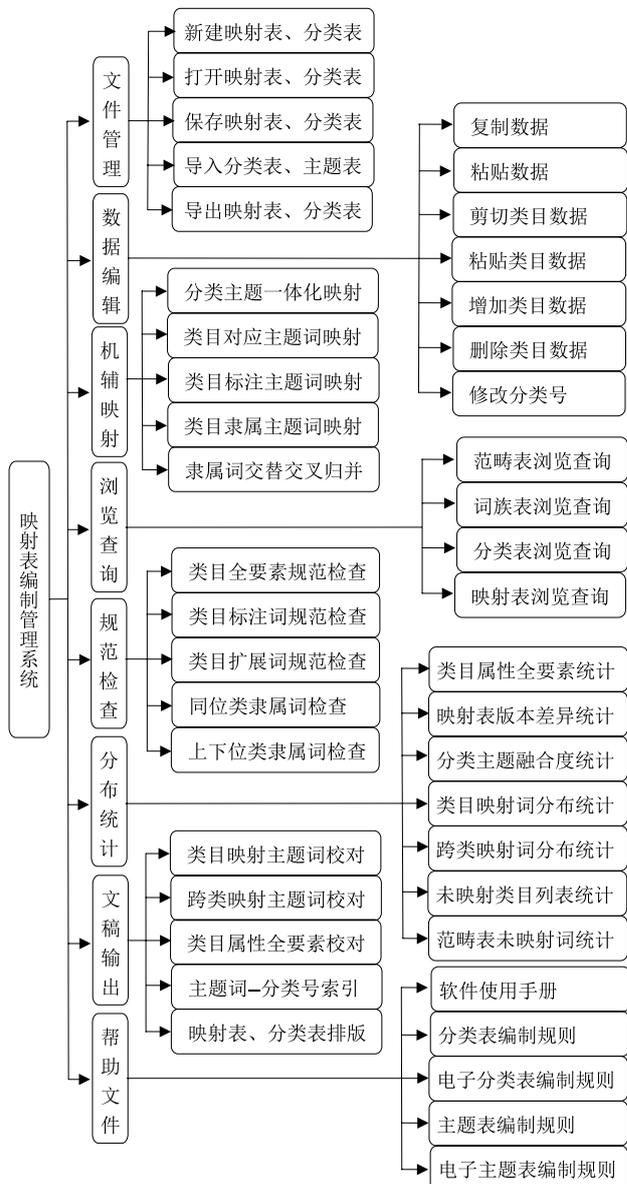
(1) 编制说明: 简述《映射表》的编制目的、指导思想、技术规范、功能特点和使用方法等。有助于全面了解《映射表》及使用方法。

(2) 分类主题词表: 是从分类体系指向主题词的一体化分类主题表。内容包括《军分表》24个基本大类、4个通用复分表以及各级类目对应的主题词。其款目构成要素包括: 分类号、类名(包括类目对应主题词、扩展类目主题词)、类目注释以及类目标注主题词、类目隶属主题词等, 该表既是一部增加主题词的新版《军分表》, 又是一部以《军分表》体系组织而成的《军表》的分类索引, 是从分类角度进行分类号和主题词一体化标引和检索的工具。

(3) 主题词—分类号对应索引: 是从主题词指向分类号的对应索引。借鉴主题法原理, 将分类表中类目对应主题词、扩展类目主题词、类目标注主题词、类目隶属主题词等与分类号相对应, 从而形成主题词字顺索引。

2 “映射表编制管理系统” 框架结构

《映射表》是一项大型复杂的数据知识工程, 为了减少大量繁杂的人工劳动, 提高编表的效率和质量, 课题组研制开发了“映射表编制管理系统”。其框架结构如下:



该系统由文件管理、数据编辑、机辅映射、浏览查询、规范年检查、分布统计、文稿输出、帮助用户八个部分组成, 其功能完备、操作简便, 实现了《军表》与《军分法》之间的多元、多维关联映射, 可从多个角度对《映射表》质量进行量化评估, 能够对《映射表》全要素进行可视化维护管理, 为完成《映射表》编制任务提供计算机技术保障。

3 映射与审核

3.1 映射工作

在完成编制《映射表》各项准备的基础上, 课题组

依据各参编单位的业务工作领域进行映射任务分工,通过“映射表编制管理系统”开展分编工作。《军表》中有正式主题词55105条,经过近1年时间的分编,已将约46000条词映射到相应类目,同时发现约9000条词在《军分法》中找不到相应类目,无法实现映射。

通过对约9000条无法实现映射主题词进行具体分析,主要原因在于:一是其中约6000条是《军表》收选的军事学科以外的相关学科主题词,而《军分法》中没有设置“相关学科”的大类;二是其中约2500条是《军表》收选的各军事学科中的一般概念主题词,而在《军分法》相应学科中没有设置“一般性问题”类目;三是其中余约500条是《军表》收选的各军事学科中的专指概念主题词,在《军分法》中也没有设置相应类目。

为实现《军表》与《军分法》的全面融合映射,经反复研究,对《军分法》进行适当调整修订。一是在《军分法》的基本大类中增加了Y 相关学科,该大类中下设二级类目22个、三级类目153个、四级类目127个、五级类目9个、扩展类目8个;二是在《军分法》各军事学科中新增二级类目15个、三级类目71个、四级类目126个、五级类目6个、扩展类目233个,其中“一般性问题”类目15个。通过修订和调整《军分法》,较好地完成《军表》8万余条主题词到《军分法》8千余个类目的全面融合映射。

3.2 审核工作

《映射表》分审工作结束后,聘请了4名曾参加或指导过《军表》和《军分法》编制工作的军内外检索语言专家参加总审,强化《映射表》总审工作力度。

(1) 个人审查与集体讨论相结合

专家个人审查中发现的问题及提出的修改意见,须经总审组集体讨论后,方可进行实际修改,避免因专家个人对问题的认识和理解不同而出现偏差。

(2) 全面审查与“映射表编制管理系统”专项审查相结合

“映射表编制管理系统”具有从多个角度对《映射表》质量进行专项检查功能,总审组在对《映射表》进行全面审查的同时,充分利用该系统具有的专项检查功能,发现和解决具有共性、可预见性的问题,取得明显的成效。

(3) 综合审查与专业审查相结合

总审组在对一些专业性较强的学科进行审查时,遇到疑难问题或较生疏的学科,及时向有关专业单位的专家请教或直接邀请他们参加审查工作,先后对《映射表》进行两次全面审查、修改,多次专项检查、核准,显著提高《映射表》的整体质量。

3.3 试标引工作

《映射表》总审工作结束后,总装科技信息研究中心资料馆、军事科学院军事图书资料馆、国防大学图书馆等单位,利用《映射表》,对馆藏文献进行试标引。总装科技信息研究中心资料馆对3217篇中文军事科技期刊论文进行标引,军事科学院军事图书资料馆对10余种刊物的4000余篇文章进行标引,国防大学图书馆对1500余种军事图书及自建数据库中2800余篇军事期刊论文进行标引。

4 功能与特点

《映射表》运用分类主题一体化技术实现了《军表》与《军分法》的全面融合,其整体功能大于《军分法》和《军表》两者功能之和。这样,将主题语言直观性优势引入分类语言,标引人员直接通过主题词查找到文献信息所属相应的类目,用户也不需要系统掌握分类法,直接用主题词,就可以通过《映射表》进入分类系统,按知识体系检索到所需要的文献信息,实现真正意义上的分类主题一体化,满足全军机关、院校、科研部门和文献管理机构对军事信息资源管理的迫切需求,提高文献信息标引和检索效率。同时,也为军用信息检索语言的本体语义研究摸索经验,以推动军队信息化应用的总体水平。

(1) 通过分类表和主题表的映射,建立分类主题一体化工具,同时进行分类号、主题词的标引,简化操作程序,降低标引难度,提高标引质量,节省人力和物力成本。

(2) 便于分类检索、主题检索以及分类主题两者组成的混合检索,提高查全率和查准率。

(3) 为实现计算机辅助标引,以及对军事信息资源的科学组织和开发利用,打下坚实数据基础。

(4) 为实现多学科或专业检索语言的兼容互换,建立多学科的分类主题映射表奠定基础。

(5) 为《军表》和《军分法》的修订提供依据。

5 结论与展望

《映射表》是我军首部大型、综合性分类主题一体化检索语言, 将《军表》的全部主题词对应到《军分法》的相应类目, 实现《军分法》与《军表》的全面融合和一体化集成, 是进行各类军事信息资源科学组织与有效利用的基础性工具。同步研发的“映射表编制管理系统”, 解决《军表》与《军分法》之间的多元、多维关联映射关系, 创新了分类主题一体化映射的思路、途径和技术手段, 从多个角度对《映射表》质量进行量化评估, 实现《映射表》全要素的可视化维护和管理。随着《映射表》在军事信息资源组织和服务领域的应用, 《映射表》将推进军事信息资源的主题和分类自动标引, 简化标引程序, 降低标引难度, 提高标引质量和效

率。同时, 解决军事信息资源的回溯分类或主题重组难题, 极大地提升主题分类一体化检索质量和效率。当然也需要进一步优化版面格式、完善《索引表》组配限定和《复分表》主题映射, 尽快组织研制开发“分类主题一体化机辅标引系统”, 实现信息资源分类主题同步标引、检索服务。

参考文献

- [1] 赵建华.《军用主题词表》的发展现状及展望[J].情报理论与实践,1996(2):20-21
- [2] 赵建华,陈庆平,周健,等.《军事信息资源分类法》的总体设计与技术特色[J].中国索引,2005(4):34-39.

作者简介

赵建国, 男, 1960年生, 军事科学院全军军用信息检索语言编管会办公室副主任、研究员, 研究方向: 信息组织、信息检索语言。

周健, 男, 1963年生, 硕士, 军事科学院全军军用信息检索语言编管会办公室副研究员, 研究方向: 信息组织、信息检索语言, E-mail: zj1963@tom.com。

Mapping Table of Military Classification and Thesaurus: Design and Compilation

ZHAO JianGuo, ZHOU Jian

(PLA Academy of Military Science, Beijing 100091, China)

Abstract: Mapping Table of Military Classification and Thesaurus is a retrieval languages tool of integration of classification and subject method. It can be used both in classification indexing & retrieval and subject indexing & retrieval of military information resources. It can also be used in the process of one-time classification and subject indexing & retrieval. This paper briefly introduces the overall design of Mapping Table of Military Classification and Thesaurus, discussed the frame structure of Compilation and Management System for Mapping Table of Military Classification and Thesaurus, elaborated the mapping and review process and put forwards the conclusion and prospect.

Keywords: Information Retrieval Language of Military; Classification and Thesaurus Mapping; Integration of Classification and Subject

(收稿日期: 2015-04-20; 编辑: 王立学)