

军队机关公文主题词标引：问题、分析与对策

赵建国, 周健

(军事科学院, 北京 100091)

摘要: 总结了军队机关公文主题词标引的历史作用, 探讨了军队机关公文主题词标引面临的问题与挑战; 论证了军队机关公文主题词标引的必要性, 分析了自然语言与人工语言、关键词检索和主题词检索的各自特点, 提出了改进军队机关公文主题词标引工作的对策。

关键词: 主题词标引; 全文检索; 标引质量; 公文管理; 档案管理; 全寿命管理

中图分类号: G254.2

DOI: 10.3772/j.issn.1673-2286.2015.10.010

军队机关公文是军队机关处理公务中形成的具有法定效力和规范体式的文书, 是军队机关履行职能的重要工具^[1], 包括命令、通令、决定、指示、通知、通报、报告、请示、批复、函、通告、会议纪要等12个文种。

1 军队机关公文主题词标引的历史作用

军队机关应用主题词表标引公文主题词, 起步于20世纪90年代初, 其重要标志是全军第一部大型综合性主题检索语言工具书《军用主题词表》的出版。1992年3月, 中央军委办公厅发布《中国人民解放军机关公文处理条例》, 将主题词正式列入军队机关公文格式的组成要素中。1994年, 《军用主题词表》由总参谋部、总政治部、总后勤部联合颁发全军团以上单位使用。该词表是我军第一部大型综合性主题词表, 对统一规范我军标引、检索用语, 编制各类专业词表, 建立自动化信息管理系统, 实现全军联机检索、信息资源共享具有重要意义和作用。2005年10月, 中央军委发布新的《中国人民解放军机关公文处理条例》, 总装备部发布国家军用标准《军队机关公文格式》, 都重申了军队机关公文标注主题词的规定。

20多年来, 在中央军委的领导下, 在总部机关和全军军用信息检索语言编制管理委员会的大力推动下, 伴随着《军用主题词表》、《军用公文主题词表》的编制完成及军用信息检索语言系列成果的推广应用, 全军

机关公文主题词标引工作全面展开, 并取得了显著的成效: 加快了公文运转和公文内容传达的速度, 受文对象通过阅读主题词就能了解公文传达的中心内容; 完善了公文检索的手段, 弥补了分类检索的不足; 为建立计算机存储与检索信息资源提供了更加科学有效的途径, 提高了文件利用的效率; 加快了档案著录工作的进度, 为开发档案信息资源提供了新的方法和途径。可以肯定, 军队机关公文主题词标引促进了军队文书档案信息资源的开发与应用, 对军队信息化建设起到积极作用。

2 军队机关公文主题词标引面临的问题与挑战

由于种种原因, 军队机关公文主题词的标引实践和组织管理出现了一些问题, 也面临着技术发展和党政机关同行新规的挑战和影响。

2.1 标引实践的问题

同样一篇公文, 对于不同的标引人员, 由于个人理解和认识上的差异以及基本功和责任心的不同, 标引结果也不尽相同, 甚至大相径庭, 导致随意性较大, 乱标、错标、漏标现象普遍存在; 一些人简单地按公文标题字面切分或字面组配标引, 隐含的主题内容没有得到充分的挖掘, 标引的主题词专指度也不够高; 主题词标

引专业性强,对于非专业人员来说费时费力,标引方法繁琐,培训又不及时、不到位,导致标引不规范;由于词表修订难度大,周期长,往往一用十几年,难免挂一漏万,跟不上现代科学技术以及国防和军队建设发展的需要。

2.2 组织管理的问题

部分单位由于组织管理和宣传培训不力,办文人员对主题词标引的重要性认识不清,认为这项工作可有可无,标引时,不能严格按照国家军用标准《军队机关公文主题词标引规则》进行操作。标引工作缺乏严格有效、切实可行的审核制度和问责机制,相当多的单位对标引结果的事后监督不力,甚至完全没有监督,标引工作监督管理不到位的情况普遍存在。

2.3 技术发展的挑战

随着现代信息技术的发展,特别是全文检索技术的出现,针对公文主题词标引,社会上出现了不少议论。有人认为,不论是公文管理系统,还是档案管理系统,通过关键词就可以快速地在后台数据库中模糊查询到与关键词相匹配的文件,基本满足公文、档案检索的需求^[2]。还有人认为,随着办公管理信息系统的不断优化,文档一体化的关键词检索模块已经相当成熟,基本可以摆脱对主题词检索的依赖。用关键词代替主题词检索公文信息成本低廉,更加方便、快捷^[3]。更有人认为,主题词主要用在公文的归档管理和查询检索等再利用环节,公文办理过程中不需要主题词,主题词的存在与否既不影响公文的表达效果,也不会影响公文的执行力,取消主题词标引,应用自然语言检索会减少公文的制作成本,减轻公文制作者的负担,提高公文的时效性^[4]。新技术的发展引发了关于军队机关公文主题词标引前景的热议:关键词检索能否取代主题词检索?在新的技术条件下主题词标引工作还有没有意义?

2.4 党政机关公文处理新规的影响

党政机关公文主题词标引,始于1986年的国务院文件,同样取得一些成效。然而,由于主题词标引过程中存在着类似军队遇到的诸多问题,考虑到全文检索技术的发展,党政机关公文主题词标引环节终于被取消:2012年4月,中共中央办公厅、国务院办公厅联合发

布《党政机关公文处理工作条例》;2012年6月,国家质量监督检验检疫总局、国家标准化管理委员会联合发布《党政机关公文格式》——这两项新规都取消了标引主题词的要求。随着党政机关公文主题词标引工作的终止,军队机关公文主题词标引工作的命运开始受到关注:党政机关取消公文主题词标引的新规对现行《中国人民解放军机关公文处理条例》的修订会产生什么样的影响?取消公文主题词标引是否人心所向、大势所趋?军队机关公文主题词标引究竟何去何从?这个问题也给我们主持修订国家军用标准《军队机关公文主题词标引规则》工作造成困扰:是按照总装计划正常修订这一规则,还是等军委办公厅再次修订《中国人民解放军机关公文处理条例》,明确军队机关公文是否继续标引主题词后再决定是否对这一国家军用标准进行修订?在没有得到职能部门明确答复的情况下,我们选择了前者。

3 分析与对策

3.1 军队机关公文主题词标引的必要性

针对有些人对军队机关公文主题词标引工作重视不够,认为可有可无的情况,我们首先从理论上对这项工作的必要性进行了重新思考和定位。

3.1.1 军队机关公文主题词标引是科学组织与共享利用文书档案的重要条件

军队信息化建设的一个重要环节就是各类数据库建设。首先要将信息资源有序化处理,然后才能实现信息资源的快速检索和有效利用。军队机关公文的主题词标引工作是军队机关按主题内容有序化处理公文的首要条件,也是日后快速查找、高效利用公文的重要基础。这项工作建设军队机关公文数据库、军队档案数据库、军事信息资源数据库不可缺少的重要工作。可以说,军队机关公文主题词标引是科学组织、管理和共享利用公文的重要环节和前提条件。

3.1.2 军队机关公文主题词标引是提高文书档案全寿命管理周期整体水平的重要保证

从军队机关公文起草、成文局部环节说,表面上看,取消主题词标引工作,减少机关公文成文的要素、

环节以及工作量,可以提高机关公文处理的工作效率。但是,从机关公文全寿命管理周期的全局,特别是后期维护、应用等环节看,如果机关公文取消主题词标引,当机关公文藏入档案馆室、归入公文数据库、档案数据库、信息资源数据库时,从事档案管理、数据库管理的工作人员将面临两种选择:要么不标引,直接用关键词检索;要么从头做起,对公文档案进行标引。由于关键词检索具有很大的局限性,查全、查准率低,不标引肯定影响档案的检索和使用效率。如果进行主题词标引,需要重新熟悉档案的内容,付出更大的艰辛,才能按主题内容实现机关公文的有序化管理工作。可见,在军队机关公文形成之初标引主题词,虽然与公文起草单位、个人的直接关系不太大,但对公文建档、计算机管理能带来事半功倍的效果。

3.1.3 军队机关公文主题词标引符合文书档案有序化处理的客观规律

公文管理是档案工作的基础,要做好档案工作必须首先对公文处理过程进行超前控制,实行政档一体化管理。我们认为,军队文书档案有序化处理的最佳时机始于公文形成阶段,最佳人选是公文起草办文人员。这是因为:(1)再没有人比公文起草办文人员更了解该公文所表达的主题内容,更容易选择中心内容、重点段落,更容易选定标引词。(2)如果机关公文起草、成文时未标引主题词,公文入藏各级档案馆室、归入各类数据库时,各级档案馆室、各类数据库从事信息资源管理的工作人员势必要对同一份公文,多次重复标引主题词,既浪费时间、降低效率,又可能人为造成因人而异、因时而异的标引结果,加大主题词标引的差异性,也极有可能降低公文主题标引的质量。(3)如果机关公文起草、成文时标引了主题词,各级档案馆室、各类数据库工作人员只需从信息资源维护管理、检索利用的专业角度对公文标引主题词的结果重新审核、质量把关。

总之,军队机关公文主题词标引能提高文书档案检索效率,在机关公文全寿命管理周期全过程中,整体提高文书档案维护管理、共享利用水平。因此,军队机关公文主题词标引工作理应继续进行。从全军军用信息检索语言编管会办公室的职能角度来讲,我们也希望军队机关公文主题词标引工作得到可持续发展。

3.2 关键词检索不能取代主题词检索

关键词和主题词分别是自然语言和人工语言的典型代表。针对有些人认为在新的技术条件下,关键词检索基本可以取代主题词检索,军队机关公文主题词标引已经失去意义的观点,学术界对自然语言和人工语言的各自特点进行过反复研究,认为关键词检索暂时无法取代主题词检索。这是因为自然语言虽然简单易用,无需规范,成本低廉,方便灵活,能及时反映新事物、新概念,但是无法控制同义词、近义词,不能揭示彼此间关系,难以实现扩检、缩检和相关检索,影响查全、查准率。自然语言检索系统中有许多词与检索需求看似高度相关,所表达的内容却无关宏旨^[5]。而人工语言虽然结构复杂,易用性不强,词汇管理成本高,维护和更新难度大,标引负担重,具有一定的滞后性,但它能简单明了又比较专指地表达文献信息和检索需求的主题概念,语词与概念一一对应,可以控制同义词、近义词和多义词,显示概念之间的等级关系,便于进行上、下位扩检和相关检索,查全、查准率高。关键词检索特别是全文检索虽然无需任何标引,检索者用自己认为合适的关键词在文本中就能匹配查找,看似成本低、处理速度快、易学易用,但其检索功能单一,适用范围有限,不能满足多样性的检索需求,查全率和查准率不高^[6]。实际上单纯的全文检索系统只具备计算机辅助文本浏览功能,并不具备正规的主题检索功能^[7]。正是由于关键词检索和主题词检索各有利弊,短时间内它们不可能彼此取代,只能互相补充,互相融合。

3.3 应对策略

既然理论上军队机关公文主题词标引工作没有失去意义,技术上关键词检索暂时取代不了主题词检索,为解决军队机关公文主题词标引过程中出现的矛盾和问题,确保这项工作的可持续发展,我们研究了以下对策。

3.3.1 军队机关公文主题词标引统一使用“军用主题词表辅助标引系统”

为解决军队机关公文主题词标引结果不统一、不规范,随意性大,乱标、错标、漏标现象普遍存在,以及简单按公文标题字面切分或组配标引,隐含的主题内

容得不到有效挖掘等问题, 全军军用信息检索语言编管会办公室组织相关部门研制开发了“军用主题词表辅助标引系统”, 并在国家军用标准《军队机关公文主题词标引规则》中明确规定: “军队机关公文主题词标引, 统一使用‘军用主题词表辅助标引系统’”^[1]。该系统是一个采用现代信息处理技术, 模拟人工标引思路和流程, 辅助用户进行主题标引的新型检索语言工具。可根据标引工作的实际需求, 采用标题转换、段落转换、全文转换三种策略, 自动从文献中提取主题概念, 并将其转换为标引词, 供用户选择, 标引效能显著提高; 系统设有方便、快捷的词表查询功能; 可联动查询和显示用户拟选标引词的全方位语义信息, 并可按语义关联进行跳转扩充查词, 保障用户准确、迅速选定标引词。

3.3.2 完善机辅标引算法, 提高标引专指度

针对标引专指度不够的问题, 我们通过组织改进、更新概念转换算法, 提高机辅标引软件推荐主题词的全面性、专指性, 提升军队机关公文主题词标引深度和自动化水平。经过多年努力, 全军已有14部专业主题词表能够用于计算机辅助标引。全军军用信息检索语言编管会办公室正在组织维护完善机辅标引软件, 审查修改专业主题词表, 以期达成全部23部军用专业主题词表的机辅标引, 届时, 《军用主题词表》及23部军用专业主题词表的专指度将大幅提升, 机辅标引自动化水平也有望明显改善。与此同时, 我们为全军各级机关和文献管理部门提供23部军用专业主题词表的机辅标引服务保障。

3.3.3 建立军用信息检索语言网络交互服务平台, 实现常态化培训和主题词表动态管理

为解决非专业人员普遍感到军队机关公文主题词标引工作专业性强, 标引方法繁琐, 费时费力, 难以掌控, 培训不及时、不到位的问题, 以及词表修订难度大, 周期长, 往往一用十余年, 难免挂一漏万, 跟不上现代科技和军队建设发展需要的现状, 全军军用信息检索语言编管会办公室正在组织构建“全军军用信息检索语言网络交互服务平台”, 为全军各级机关、文献管理部门提供军用信息检索语言系列化、标准化的数百种服务内容。包括5项机辅编制和应用软件工具, 2项培训手册, 4项印刷版、电子版检索语言编制标准, 9项检索

语言应用标准, 2项公文和档案管理标准等。该平台将提供军用信息检索语言的用户、管理者、专家学者3者间交流互动的窗口和渠道, 通过视频、动画等生动活泼、形式多样的常态化在线培训, 达成检索语言的用户化、网络化。与此同时, 在军网上及时发布检索语言最新数据, 实现《军用主题词表》和军用专业主题词表的动态化维护管理, 彻底改变修表周期长、“十年磨一剑”的状况。

3.3.4 建立健全严格的审核和问责制度

针对标引工作缺乏严格有效、切实可行的审核制度和问责机制, 有关部门对标引结果的事后监督不力的问题, 我们认为可以通过标引结果的审核和问责制度建设来解决。这项工作是确保军队机关公文主题词标引和检索质量的一项重要举措。审校的主要内容是对主题词标引结果进行检查, 看其是否达到了标引的针对性、客观性、专指性和一致性。要求标引人员严格遵循标引工作程序和相关规则, 减少标引误差。标引词的选择要准确、全面、无歧义、无遗漏, 主题词登录格式要规范。建议职能部门把标引质量管理纳入到办文人员年度考评的内容中去, 各单位建立岗位责任制, 对胡标、乱标者及时予以惩戒; 对监督不力、严重失察者追究其事后责任。

4 结语

军队机关公文主题词标引工作自身及其管理审核存在一些问题, 同时面临新技术和党政机关取消公文主题词标引的挑战和影响, 但公文主题词标引的必要性没有动摇。在相当长的一段时间里, 自然语言检索不能取代受控语言检索的地位和作用。从文档一体化管理的角度看, 军队机关应该继续做好公文主题词标引工作。作为全军军用信息检索语言编制管理研究的常设主管机构, 我们对此有充分的信心。这项工作的最终命运取决于军队文书档案部门对于关键词与主题词的检索效率对照评估结果以及决策部门的权衡和综合考量。

参考文献

- [1] 《〈中国人民解放军机关公文处理条例〉问答》编写组.《中国人民解放军机关公文处理条例》问答[M].北京:解放军出版社,2005:11.

- [2] 王茂跃.取消公文主题词的真正原因是什么?——兼与郑彦离等先生探讨[J].档案管理,2014(6):68-70.
- [3] 马国金.党政公文处理规范统一后机关公文的形势变化[J].秘书,2012(9):34-35.
- [4] 姜性诚.话说公文的主题词[J].应用写作,2011(6):1.
- [5] 张琪玉.检索标识的专指度[J].江西图书馆学刊,2006(2):5-6.
- [6] 张琪玉.全文检索与索引[J].图书馆杂志,2007(11):3-5.
- [7] 张琪玉.关于自然语言检索问题[J].图书馆论坛,2004(12):211-213,145.
- [8] 赵建华,贾国文,赵建国等.GJB2418.5A-2006军队机关公文主题词标引规则[M].北京:总装备部军标出版发行部,2006:10.

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制管理委员会办公室副主任、研究员,研究方向:信息检索语言编制、管理与应用研究。
周健,男,1963年生,硕士,军事科学院全军军用信息检索语言编制管理委员会办公室副研究员,研究方向:信息检索语言编制、管理与应用研究, E-mail: 13439973495@163.com。

PLA Administrative Documents Descriptors Indexing: Problems, Analyses and Countermeasures

ZHAO JianGuo, ZHOU Jian
(PLA Academy of Military Science, Beijing 100091, China)

Abstract: Concludes the historical function of PLA administrative documents descriptors indexing. Discusses the problems and challenges consists in PLA administrative documents descriptors indexing. Analyses each features of natural languages and artificial languages, key word retrieval and descriptors retrieval. Demonstrated the necessity of PLA administrative documents descriptors indexing and put forwards the countermeasures to improve the work.

Keywords: Descriptors Indexing; Full-text Retrieval; Indexing Quality; Official Documents Management; Archives Management; Full-life Management

(收稿日期: 2015-09-01; 编辑: 王立学)

■ 书 讯 ■

《汉语主题词表》(工程技术卷)

《汉语主题词表》自1980年问世以后,经1991年进行自然科学版修订,在我国图书情报界发挥了应有的作用,曾经获得了国家科学技术进步二等奖。为了适应网络环境下知识组织与数据处理的需要,2009年由科学技术信息研究所主持,并联合全国图书情报界相关机构,完成了《汉语主题词表(工程技术卷)》的重新编制工作。

全书共收录优选词19.6万条,非优选词16.4万条,等同率0.84。在体系结构、词汇术语、词间关系等方面进行改进创新。为了方便工程技术领域不同专业用户使用,《汉语主题词表》(工程技术卷)按专业分13个分册出版,同时建立《汉语主题词表》服务系统,提供在线概念检索和辅助标引服务,通过可视化技术展示各类概念关系,是图书馆、档案馆、出版社、期刊杂志社、文献信息中心等专业工作者及科研、教育及工程技术领域人员必备的参考书。

《汉语主题词表(工程技术卷)》已于2014年由科学技术文献出版社出版,全书2300余万字,总定价3880元,可分册购买。