

促进主题词表的社会应用

主题词表即叙词表,是以特定的结构集合展示经过规范化处理的优选词和非优选词及其词间语义关系,作为标引和检索的术语控制工具。主题词表自诞生以来得到了很好的应用。例如, EI叙词表支撑Engineering Village文献平台的检索、浏览等功能;联合国粮食及农业组织编制的AGROVOC叙词表包含21种语言,在农业方面得到广泛应用;美国NASA叙词表在美国航空航天领域得到较好应用。

随着互联网技术高速发展,大数据时代对主题词表提出了新的应用需求,不仅在图书情报机构中应用,更需要在跨界的相关机构中应用。例如,当前发达国家将主题词表应用于电子政务环境下政府信息的描述和提取中,利用主题词表建立知识库的自动分类系统,实现知识本体意义上概念关系的智能推理和语义聚类;应用于电子商务环境和大型企业的知识管理领域中,利用主题词表对海量信息的知识点进行自动标注,实现不同颗粒度的智能查询、领域监测、知识挖掘、领域知识聚类应用;在医学、教育领域主题词表也应用于词语切分、信息抽取、聚类、词频统计、情感分析等文本信息处理基础工作,利用主题词表进行英汉对照索引,实现英汉跨语言检索和辅助机器翻译功能。这些应用无疑都存在一个需要精确定位信息的较大的社会需要。

同时,必须强化主题词表的应用场景建设,以网站应用、平台应用、系统应用和嵌入式应用的形式,适应从纸质媒介向网络媒介转移,适应新的专业应用环境,适应不同专业背景知识下的人员,适应快速更新、合作开发和维护的应用情形。迫切需要对主题词表应用场景进行深入的分析 and 提炼,探讨促进主题词表社会应用的机制和方法,确定主题词表的社会应用模式,提出保证主题词表社会应用长效实施和持续维护的合作方案,发挥主题词表的社会价值及经济价值。

大数据时代,社会的不同需求促进着主题词表的发展,而主题词表的更新也需要更加适应社会需要。知识组织工作者需要从社会需求的角度出发,不断地完善主题词表的结构和功能,为主题词表注入新的血液和活力,加强数字环境下主题词表的生命力,促进主题词表的应用呈现新的气象。

■ 周晓英