

DOI:10.12154/j.qbzlgz.2023.05.002

ChatGPT 完成知识组织任务的效果及启示

曹茹烨¹ 曹树金^{2,1}(¹中山大学信息管理学院 广州 510006; ²山东理工大学信息管理研究院 淄博 255000)

摘要: [目的/意义]人工智能技术迅速更迭,ChatGPT 正在驱动各行各业的变革与发展,知识组织环境发生了巨大变化。在此背景下,文章分析 ChatGPT 用于知识组织任务的效能,探索知识组织与新一代 AI 融合的方式方法,对于促进知识组织领域的创新发展具有重要意义。[方法/过程]文章以知识组织体系和实际业务为参考框架,设计了文献著录、标引、本体构建等 8 类实验任务,由 ChatGPT 根据提示完成任务并对其完成效果进行评估,进而思考 ChatGPT 在知识组织领域发挥的作用以及知识组织未来发展的趋向问题。[结果/结论]ChatGPT 具有较强的自适应性,但创造力有限,面对复杂的知识组织任务尚不能直接生成可用的产品。未来知识组织领域的发展需要充分发挥专业人员的创新能力,不断探索知识组织与 ChatGPT 等生成式 AI 深度融合的理论体系和规则方法。

关键词: ChatGPT 知识组织 实验任务

The Effect and Enlightenment of ChatGPT in Completing Knowledge Organization Tasks

Cao Ruyue¹ Cao Shujin^{2,1}(¹School of Information Management, Sun Yat-sen University, Guangzhou, 510006;²Information Management Research Institute, Shandong University of Technology, Zibo, 255000)

Abstract: [Purpose/significance] Artificial intelligence technology is changing rapidly, ChatGPT is driving the transformation and development of all walks of life, and the knowledge organization environment has undergone tremendous changes. In this context, it is of great significance to analyze the effectiveness of ChatGPT for knowledge organization tasks and explore the ways and means of knowledge organization and new generation AI integration to promote innovation and development in the field of knowledge organization. [Method/process] Taking the knowledge organization system and actual business as the reference frame, 8 types of experimental tasks such as document description, indexing, and ontology construction are designed. After ChatGPT completes the task according to the prompts, evaluate its completion effect, and then think about the role of ChatGPT in the field of knowledge organization and the future development trend of knowledge organization. [Result/conclusion] ChatGPT has strong adaptability, but its creativity is limited, and it cannot directly generate usable products in the face of complex knowledge organization tasks. The future development of the field of knowledge organization needs to give full play to the innovative ability of professionals, and constantly explore the theoretical system and rules and methods for the deep integration of knowledge organization and generative AI such as ChatGPT.

Keywords: ChatGPT knowledge organization experimental task

1 引言

ChatGPT自发布以来,迅速掀起新一轮的AI革命,并在不断地自我迭代中,涌现出强大的推理和思维链等能力,GPT-4更是具备极强的泛化能力和适应性。ChatGPT拓宽了人机交互的发生场域,在自然语言处理、智能搜索、情境推理等各类任务场景中展现出了出色的性能。其上线不到半年的时间里,各大互联网企业纷纷布局,微软推出了基于GPT模型的智能搜索引擎New Bing和智能Office插件Copilot,谷歌、百度分别发布大语言模型Bard、“文心一言”。可见,以ChatGPT为代表的生成式AI正在重构产业格局,并将逐渐渗透至各行各业。

信息资源管理学科与作为信息处理方法和技术的的人工智能存在天然的耦合关系。知识组织是信息资源管理的重要分支,实质是通过揭示知识客体中的知识因子和知识关联,实现对知识的序化^[1]。近年来,信息资源管理领域不断推进以语义模型为基础的知识组织研究^[2],而ChatGPT的爆火使知识组织环境又一次发生巨大变化,其在语义理解、知识抽取、知识生成与推荐方面的出色表现,无疑会提高知识组织的效率。近期,学者们纷纷探讨了ChatGPT对知识组织与信息检索模式^[3]、知识组织研究趋向^[4]的影响,思考了ChatGPT与知识组织、知识识别、知识关联的融合式创新^[2],支持AI模型训练的高质量知识组织体系建设^[5],以及场景化应用与融合发展路径^[6]等问题。然而,现有研究大都是基于对ChatGPT核心技术特征和功能地了解,结合知识组织的理论体系、技术发展所作出的一种基于经验的思辨。那么,在知识组织领域具体的、实际的任务中,ChatGPT能够提供的支持及完成的效果如何,能否生成知识组织工具和规则,其创造能力的边界以及由此引发的知识组织未来发展等问题均需要进一步论证。

基于此,本文以知识组织体系为导向,设计若干个实际的任务情景,通过与ChatGPT的持续交互,获取生成内容并检验其完成效果,从而对知识组织与新一代生成式人工智能的深度融合路径以及未来发展方向提供启示。

2 人工智能用于知识组织任务的相关研究

自1929年英国著名的图书馆学家布利斯首次提出“知识组织”的概念至今,知识组织经历了漫长的发展历程。从分类法、主题法,到全文索引技术和语义网的发展,再到知识图谱的出现,知识组织的理论体系、工具、技术手段在不断革新。随着人工智能技术的发展,信息资源管理领域进行了将机器学习和深度学习用于知识组织的诸多尝试。其中,在知识组织的分类任务中,结合机器学习和KOS的自动分类是学者们关注的重点。例如,de Campos和Romero^[7]提出了一个基于贝叶斯网络的模型,利用Eurovoc同义词典中的层次和词汇信息,对文档进行分层分类。李娇等^[8]以《中图法》为分类标准,提出了一种基于多因子算法的自动分类标引方法,并给出自动标引命中的最佳学科分类号。在主题自动标引任务中,梁红兵等^[9]依据现有医学词库的标引规则,设计了一种基于反馈规则学习的主题自动标引方法,可用于医学文献主题词提取以及主副主题词的组配。针对元数据抽取任务,陈淑平和梁东魁^[10]设计了一种基于机器学习的自动抽取方法,通过扫描图书页面的DjVuXML文档,将页面格式、结构作为输入特征,采用有监督机器学习完成抽取。姜霖和王东波^[11]结合神经网络模型和支持向量机分类算法实现了对引文元数据的自动归类和标注。在索引编制任务中,有学者提出了一套结合知识组织理论体系和机器学习的方案,用于多模态信息的索引创建和检索,即使用已有的元数据获取资源对象的高阶特征,使用机器学习提取低级特征,融合之后创建特征索引^[12]。在实现知识组织系统互操作方面,深度学习模型可发挥较大作用,比如有学者以Word2Vec为技术手段,建立了《中国中医药学主题词表》《中国图书馆分类法》与《中文医学主题词表》之间的语义映射^[13]。在本体构建任务中,AI技术被广泛用于领域本体构建,如基于深度神经网络构建汽车领域本体^[14],基于无监督学习的中医古籍本体构建^[15],基于XGB、SGD、SVM等深度学习模型的阿尔兹海默病领域本体构建^[16]等。2012年谷歌推出“知识图谱”的概念,它是一种智能、高效的知识组织方式^[17]。知识图谱构建任务涉及知识抽取、知识融合、知识推理等过程,极大程度地依赖自然语言处理和机器

学习等人工智能技术,比如有研究依托本体模型和 HanLP 自然语言处理工具构建中国近代史知识图谱^[18],基于图注意网络构建中国非物质文化遗产知识图谱^[19],也有研究提出利用迁移学习的数据增强方法构建领域知识图谱^[20]等。

从现有研究来看,人工智能技术已被广泛用于知识组织任务。然而大多是基于分析式 AI 的知识组织问题探索,以判别式建模为主,需要根据具体的任务和知识组织需求选择、优化 AI 模型和算法,一种模型并不适用于所有的知识组织任务场景,这对知识组织专业人员的跨学科素养提出了较高的要求。以 ChatGPT 为代表的生成式 AI 能够在分析判别的基础上生成全新的内容,且具有较强的通用性。那么,随着 AI 技术的跃迁,知识组织是否会更加智能、高效,甚至是全流程化,需要进一步验证,从而为知识组织的未来发展提供导向。

3 研究设计

3.1 实验任务设计

为了回答 ChatGPT 用于知识组织任务的效果以及新一代 AI 技术支持下的知识组织未来发展趋向问题,本研究以知识组织体系和图情机构实际的知识组织业务作为任务场景的设计依据。其中,知识组织体系(又称为“知识组织系统”)是各类知识组织规范和方法的统称,涵盖分类表、主题表、叙词表,以及支持复杂语义描述的语义网、本体、知识图谱等^[21]。图情机构的知识组织业务包括文献资源分类、著录、标引、元数据管理与知识组织系统开发等。围绕以上范畴,本文设计了8类共12个任务,涉及文献著录、分类标引、主题标引、本体构建、知识抽取与知识图谱构建、词表编制、知识组织系统的设计与实现、知识组织系统互操作,这些任务的复杂度不同,具体如表1所示。其中,在一些典型的知识组织任务(如文献分类、知识抽取与知识图谱构建)中又设计了多种情景,以综合、全面地考察 ChatGPT 在多样化需求下的知识组织性能。

表1 知识组织任务清单

任务难度	任务类别	具体任务	提示信息
简单任务	文献著录	任务1:利用 MARC 格式分别对 <i>The Kite Runner</i> , <i>Klara and the Sun</i> , <i>Rehab Science: How to Overcome Pain and Heal from Injury</i> 三本书进行著录	书名与著录要求
	分类标引	任务2:分别利用《美国国会图书馆分类法》和《杜威十进分类法》对 <i>The Kite Runner</i> , <i>Rehab Science: How to Overcome Pain and Heal from Injury</i> 两本书进行分类标引	书名、基本信息与主要内容
		任务3:分别利用《中国图书馆分类法》和《美国国会图书馆分类法》对《石油效应:全球石油危机的背后》一书进行分类标引	书名、基本信息与主要内容
		任务4:利用《中国图书馆分类法》对《川大史学》(全书共七卷)一书进行分类标引	书名、基本信息与主要内容
	主题标引	任务5:利用《医学主题词表》(MeSH)分别对发表于 <i>Lancet</i> 期刊上的学术论文 <i>Lung cancer: current therapies and new targeted treatments</i> , 以及新出版的书籍 <i>Allergic: Our Irritated Bodies in a Changing World</i> 进行主题标引	文献的基本信息和摘要
复杂任务	本体构建	任务6:构建一个芯片领域的简单本体示例,包含一些常见的类和关系,并给出实现代码	限定领域
	知识抽取与知识图谱构建	任务7:中文医学命名实体识别	待抽取的句子(来自阿里云平台中文医疗信息处理评测基准 CBLUE 数据集中的样例)
		任务8:中文医学文本实体关系抽取	待抽取的句子(来自阿里云平台中文医学文本实体关系抽取 CMeIE 数据集中的样例)
		任务9:构建一个心脏病领域的知识图谱示例	限定领域
	词表编制	任务10:编制一个脑机接口领域的词表	(1) 词汇量20个,词表由主表和附表构成,主表按字顺排列,需要给出显示词间关系的参照项,附表需提供机构索引、产品索引。(2) 根据生成结果优化提示,如指定参考的词表,给出特定词汇。
	知识组织系统的设计与实现	任务11:提供一个农史领域的知识组织系统设计方案	根据生成结果连续提问,如限定数据量、数据类型、使用场景等
	知识组织系统互操作	任务12:《美国国会图书馆标题表》与《杜威十进分类法》互操作	使用两类词表对“脂肪代谢障碍”进行标引,并实现在两个系统之间的映射

3.2 实验效果评估方法

在本研究设计的知识组织任务中,文献著录、分类和主题标引均属于简单任务,大多任务本身有明确的答案,可作为 ChatGPT 应用效果评估的依据。具体的评估方法如下:(1)书目著录和分类标引:从综合性学术搜索平台(如读秀学术搜索)、书目数据库(如 Ebook Central)以及百科平台(如维基百科)中获取文献的外部属性特征、信息内容特征、分类号。关于分类号,中文文献一般按照《中国图书馆分类法》赋予,英文文献一般被赋予 LCC(美国国会图书馆分类法)分类号和 DDC(杜威十进分类法)分类号。在 ChatGPT 生成文献著录和分类标引的结果后,与学术资源数据库提供的文献信息进行对比,以评估 ChatGPT 任务完成的效果。(2)主题标引:所选实例为医学类文献,可根据生物医学类数据库中提供的 MeSH terms(医学主题词),判断 ChatGPT 主题标引结果的准确性。除此以外,本研

究还涉及对新产生的文献资源的著录和标引任务,学术搜索平台、图书馆馆藏目录或数据库中尚未对其进行描述和揭示,对这类任务效果的评估则依靠专家经验进行判断。

本体构建、知识抽取与知识图谱构建、词表编制、知识组织系统设计与开发、知识组织系统互操作都属于复杂任务,并没有确切的答案,知识组织的结果是在特定标准和规范约束下的多种可能性集合。针对此类任务,同样通过邀请从事知识组织研究和相关工作的专家学者,根据自身专业知识和实践经验,对 ChatGPT 完成的效果进行评估。

4 结果与讨论

4.1 著录与标引结果分析

在文献著录任务中,请求 ChatGPT 著录的三本书分别出版于 2003 年、2021 年和 2023 年。相较于近期出版的图书,网络中关于早期出版的书籍有更多的描述。使用三个实例的原因如下:一是个例不具普遍性;二是为了检验 ChatGPT 生成的内容是否会受限于少量的相关信息而影响准确性;三是考察 ChatGPT 对新产生的文献资源描述的效果。结果显示,一方面 ChatGPT 对已有的旧文献资源的著录字段中,主要责任者、出版

发行项、载体形态项、图书馆分类号基本正确,但 ISBN 号等字段错误,如图 1(左)所示,并且对出版于 2003 年和 2021 年的两本图书著录结果的准确率相差无几。在与 ChatGPT 的连续对话中获悉,作为一个语言模型,它无法直接访问真实的图书馆目录或者其他数据库信息,其语料主要来自于出版商、书店或者其他公共来源,这正是 ChatGPT 在文献著录任务中存在偏差的主要原因。另一方面,ChatGPT 在利用 MARC 格式对 *Rehab Science: How to Overcome Pain and Heal from Injury* 一书进行著录时,仅能给出 MARC 格式的样例,如图 1(右)所示。因为该书出版于 2023 年 5 月 30 日,而 ChatGPT 的语料截止到 2021 年,其知识库中没有关于该书的详细信息,且因其无法访问互联网或最新的出版目录而不能提供实时的图书编目信息。由此可以推测,如果 ChatGPT 等生成式 AI 能够访问公开的出版书目数据或其他相关数据,它完成图书编目和著录任务的效果或许会更好。目前,OpenAI 已为 ChatGPT 增加了插件使用功能,赋予 ChatGPT 使用工具、联网、计算的能力^[23],这也会较大程度地提高文献著录的准确率。可以预见的是,以 ChatGPT 为代表的生成式 AI 将有助于提高传统编目和著录工作的效率。

在分类标引任务中,ChatGPT 基于《美国国会图书馆分类法》(LCC)和《杜威十进分类法》(DDC)对外文书

LDR	00000nam a2200000 a 4500
001	123456789 // 替换为实际的书目控制号码
003	0CoLC
005	20230614000000.0
008	230614s2003 xx 000 0 eng d
020	1a 9781594480003 // 替换为实际的ISBN
040	1a DLC 1b eng 1e rda 1c DLC
050	0 1a PS3608 .0832 1b K58 2003
082	04 1a 813/.6 1b 21
100	1 1a Hosseini, Khaled
245	10 1a The Kite Runner / 1c Khaled Hosseini.
264	1 1a New York : 1b Riverhead Books, 1c 2003.
300	1a 371 pages ; 1c 24 cm
336	1a text 1b txt 12 rdacontent
337	1a unmediated 1b n 12 rdamedia
338	1a volume 1b nc 12 rdacarrier
650	0 1a Fathers and sons 1v Fiction.
650	0 1a Betrayal 1v Fiction.
650	0 1a Social classes 1v Fiction.
651	0 1a Afghanistan 1x History 1y 20th century 1v Fiction.
700	1 1a Hosseini, Khaled 1e author.

LDR	00000nam a2200000 a 4500
001	123456789 // 替换为实际的书目控制号码
003	0CoLC
005	20230614000000.0
008	230614s2023 xx 000 0 eng d
020	1a 9781234567890 // 替换为实际的ISBN
040	1a DLC 1b eng 1e rda 1c DLC
050	00 1a RD734 1b .A1 2023
082	00 1a 617.1/027 12 23
100	1 1a Author's Last Name, Author's First Name
245	10 1a Rehab Science : 1b How to Overcome Pain and Heal from Injury / 1c
264	1 1a Place of Publication : 1b Publisher's Name, 1c 2023.
300	1a x, 200 pages ; 1c 23 cm
336	1a text 1b txt 12 rdacontent
337	1a unmediated 1b n 12 rdamedia
338	1a volume 1b nc 12 rdacarrier
500	1a Includes index.
650	0 1a Rehabilitation 1x Methods.
650	0 1a Pain 1x Treatment.
650	0 1a Wounds and injuries 1x Treatment.
700	1 1a Author's Last Name, Author's First Name 1e author.

图1 ChatGPT对 *The Kite Runner*(左)和 *Rehab Science: How to Overcome Pain and Heal from Injury*(右)的著录结果

籍的分类标引对象有两类,分别是已有的旧文献和新产生的文献资源,结果如图2和图3所示。将ChatGPT对The Kite Runner分类标引的结果与维基百科以及图书馆书目数据库中的信息进行对比,发现分类号完全一致。同时,参考LCC和DDC的分类体系并邀请分类编目的专家对图3(即ChatGPT对Rehab Science: How to Overcome Pain and Heal from Injury分类标引的结果)进行评价,发现虽然类号和类名的对照存在部分偏差,但分类号的赋予较为合理、准确,如RD701-811实际对应的类名为Orthopedic surgery, RD792-811是Orthopedic surgery下的Physical rehabilitation。RM695-893实际对应的类名为Physical medicine. Physical therapy具体包括exercise, thermotherapy, electrotherapy等。由此可以得出,ChatGPT基于LCC和DDC对外文图书的分类标引准确率较高,并且从对新产生的文献资源的标引结果可以推断ChatGPT是在学习了LCC和DDC的标

引规则后自动生成的分类号,而非从已有的网络信息中提取。在任务3中请求ChatGPT基于LCC对《石油效应:全球石油危机的背后》一书进行分类标引,与基于《中国图书馆分类法》标引的结果进行对比,结果如图4所示。由于在实际的标引工作中,极少会依据LCC对中文图书进行分类,ChatGPT的原始训练语料中可能不存在类似案例,因此它仅能依据对LCC规则的学习推理并生成类号。从结果来看,ChatGPT依据中图法为《石油效应:全球石油危机的背后》一书赋予的分类号完全错误(正确的分类号应为F416.22),TP319为“专用应用软件”而非石油工业。而ChatGPT依据LCC给出的具体分类号虽不完全准确,但大类类号与实际类名较为一致,整体分类结果相对合理。此外,ChatGPT基于中图法对多卷书的标引结果同样有较大偏差且未能依总论复分表进行复分。由此证明ChatGPT对中图法标引规则的学习是不足的,根本原因是在ChatGPT的

训练数据中中文语料是不足的。为减少单次实验带来的随机性误差,针对同一任务进行了多次重复提问,发现ChatGPT生成内容并不唯一,但任务完成的效果无甚差别。从多次实验中所给出的不同结果也可进一步证明,ChatGPT是在学习了现有的分类标引规则后自动生成的分类号而非从已有网络资源中直接提取,但因语料不足或信息缺失等原因导致准确率不高。

在主题标引任务中,Lung Cancer: Current Therapies and New Targeted Treatments是2017年发表的学术论文,通过与PubMed数据库中提供的MeSH terms进行对比,发现ChatGPT依据《医学主题词表》(MeSH)对该论文的标引结果具有较高的准确率,提供的主题词与MeSH terms完全一致的词超过半数。Allergic: Our Irritated Bodies in a Changing World是2023年5月出版的新书,ChatGPT依据医学主题词表(MeSH)为其赋予的主题词包括Allergies、Environmental Exposure、Biolog-



根据题目给出的信息,将《The Kite Runner》归类于以下类别和分类号:

美国国会图书馆分类法:

PS3608.O832 K58 2003 (美国文学分类)

杜威十进分类法:

813/.6 (美国小说分类)

以上是根据《The Kite Runner》的内容和出版年份给出的分类号,可以根据实际情况选择合适的分类号。

图2 ChatGPT对The Kite Runner的分类标引结果



根据《美国国会图书馆分类法》(Library of Congress Classification),将《Rehab Science: How to Overcome Pain and Heal from Injury》进行分类标引如下:

- R - Medicine
 - RD701-811 - Orthopedic Surgery and Sports Medicine
 - RM695-893 - Physical Medicine and Rehabilitation

根据《杜威十进分类法》(Dewey Decimal Classification),将《Rehab Science: How to Overcome Pain and Heal from Injury》进行分类标引如下:

- 610 - Medicine and Health
 - 615.82 - Physical Therapy and Rehabilitation
 - 616.0472 - Pain Management and Rehabilitation

图3 ChatGPT对Rehab Science: How to Overcome Pain and Heal from Injury的分类标引结果

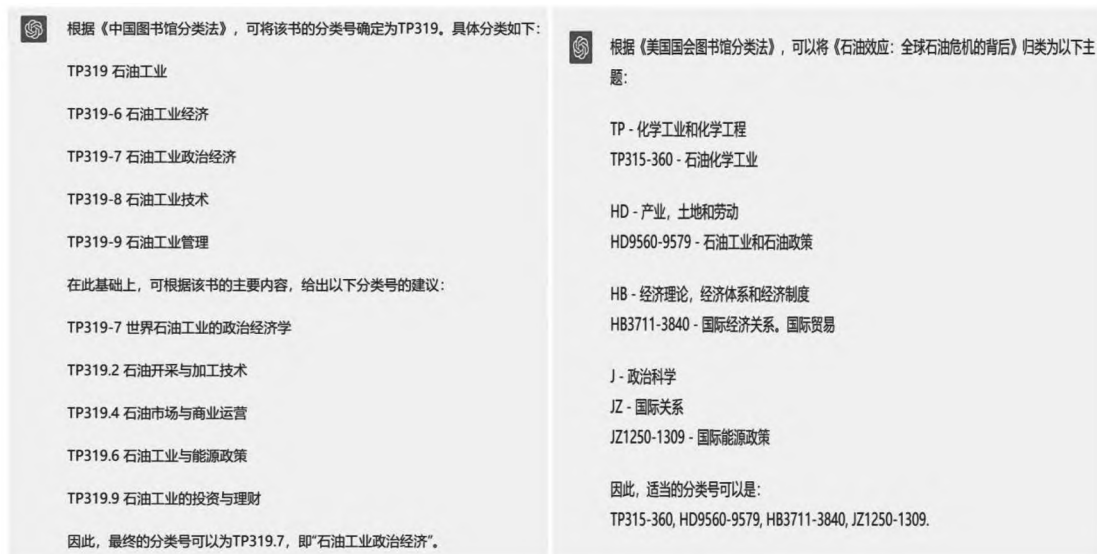


图4 ChatGPT对《石油效应：全球石油危机的背后》一书的分类标引结果(左：基于《中国图书馆分类法》，右：基于LCC)

ics、Immunotherapy、Air Pollution等,经过专家评价认为该标引结果较为准确。由此说明,ChatGPT对已有旧文献以及新产生的文献资源的主题标引效果均较好。

4.2 本体与知识图谱构建结果分析

本体和知识图谱构建都是复杂的系统工程,需要对专业领域的重要术语、类的等级体系和属性等有全面的了解,并借助一定的工具和方法,由多方协作完成。ChatGPT的出现能否简化甚至颠覆传统的构建方案需要通过实际任务来验证。

在完成本研究设计的本体构建任务时,ChatGPT能够准确地给出合理的构建思路、方法、工具以及功能实现的详细代码。围绕该任务,ChatGPT提供了芯片领域的一个本体示例,包括芯片的型号、制造商、功能、规格及其关系的知识描述,并提供了基于RDF的本体建模实现代

码,运行结果如图5(局部示意图)所示。从任务完成效果来看,ChatGPT能够自动构建本体。但针对具体的领域和具体的项目需求,如果没有预先定义的本体结构、模式、元素和属性,仅靠ChatGPT并不能生成完全符合需求的本体,所以在实际任务中可将其作为方案和技

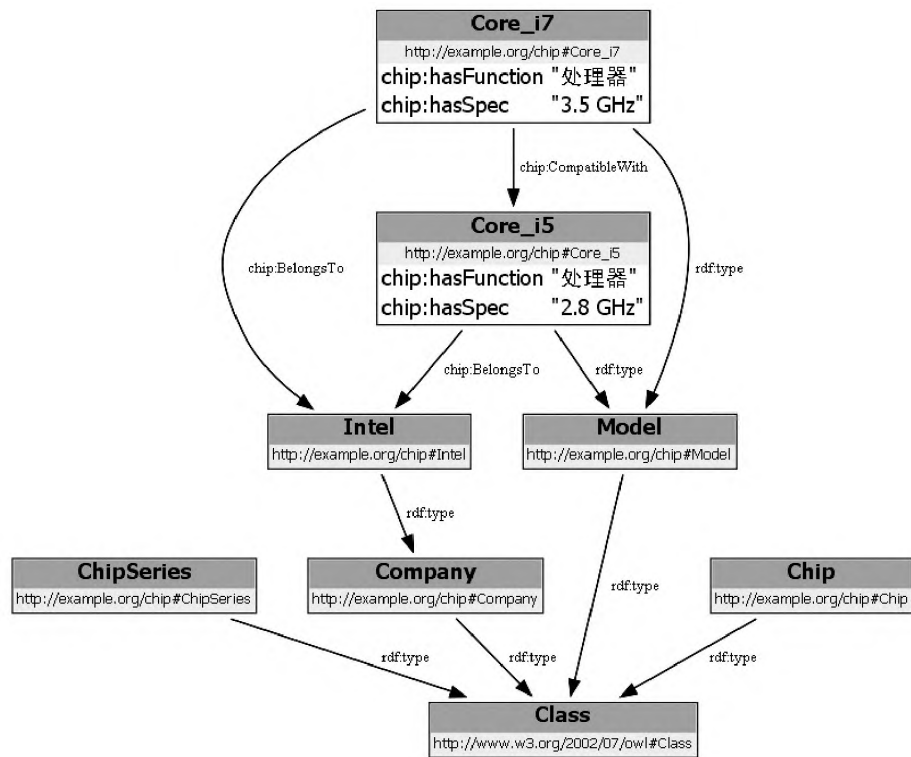


图5 ChatGPT构建的本体示例运行结果(局部)

术支持的辅助工具,并且需要在其构建的本体原型基础上,由专家进行质量评估,进一步补充和完善领域本体。

在知识抽取和知识图谱构建任务中,以阿里云平台的中文医疗信息处理评测基准 CBLUE 数据集和中文医学文本实体关系抽取 CMcIE 数据集中的句子分别作为命名实体识别和实体关系抽取的对象,以原有标注作为 ChatGPT 知识抽取效果评价的标准,部分结果如表2所示。对比来看,由于预先并未向 ChatGPT 指定需要抽取的实体和关系类型,其抽取结果与原有标注不完全一致,但结果准确并涵盖了原有标注中遗漏的实体和关系。除此以外,在关系抽取任务中,ChatGPT 能够实现对单句中多组关系的抽取,准确性较高。为避免单一领域造成的实验结果偶然性,进一步请求 ChatGPT 完成了科技文献的科研实体识别、金融和交通等其他领域的知识抽取任务,效果较好,但抽取出的部分知识元具有通用性而在专业领域并无实质意义,说明 ChatGPT 缺乏全面深入的领域知识。

针对知识图谱构建任务,知识抽取仅是其中一个关键环节,为了检验 ChatGPT 对知识图谱构建整个系统工程的胜任能力,本研究请求 ChatGPT 构建了一个

表2 ChatGPT完成的知识抽取任务(部分)

任务	句子	原标注结果	ChatGPT抽取结果
命名实体识别	尿潴留者易继发泌尿系感染。	疾病:尿潴留、泌尿系感染	疾病:尿潴留、泌尿系感染
	根据临床表现有发热、剧咳、肺部体征少,X线胸片表现相对较明显,提示肺炎支原体感染。	临床表现(症状体征):发热、剧咳、肺部体征少;医学检验项目:肺部体征;医疗程序(检查程序/治疗或预防程序):X线胸片;疾病:肺炎支原体感染	症状:发热、剧咳、肺部体征少;检查:X线胸片;疾病:肺炎支原体感染
	用免疫学方法检测某种病原体的特异抗原很有诊断价值,因其简单快速,常常用于早期诊断,诊断意义常较抗体检测更为可靠。	医疗程序:免疫学方法检测、抗体检测	检查方法:免疫学方法、抗体检测;检测物质:特异抗原;病原体:病原体
实体关系抽取	麻风病##证据##证据评分 麻风病的接触发病率;有中等质量的证据显示,与安慰剂比较,单一剂量的利福平在降低与麻风病患者接触的人员1-2年麻风病发病率方面更有效。	麻风病(疾病)-预防-利福平(其他)	麻风病-降低发病率-利福平
	腹主动脉瘤#####案例#2 一名55岁的男性,有高血压(药物控制良好)和吸烟史,因下腹剧痛持续2天到初级医疗保健医处就诊。	腹主动脉瘤(疾病)-就诊科室-初级医疗保健医处(其他)	腹主动脉瘤-相关-高血压;腹主动脉瘤-相关-吸烟史;腹主动脉瘤-症状-下腹剧痛;腹主动脉瘤-就诊-医疗保健医处

关于心脏病领域的知识图谱简单示例。在此过程中,ChatGPT 提供了思路方法(包括确定本体结构、收集数据、建立本体、图谱构建与存储)、构建结果(三元组)、存储和可视化的工具,以及实现代码。由此预见,ChatGPT 在知识组织领域细粒度的知识挖掘和知识关联任务场景中有极大优势。

4.3 词表及其他知识组织系统设计与实现结果分析

在词表编制任务中,预先设定了词表所属领域、词汇量、词表构成、款目词排列方式等框架。结果显示,ChatGPT 编制的词表中有较多参照项缺失,未能很好地揭示词间关系,效果较差。通过与 ChatGPT 的连续对话获悉,其编制词表的依据主要是领域知识、语言学原则、领域专家意见和已有的词汇资源。但从任务完成效果来看,ChatGPT 对已有词表的编制规则学习不足。为进一步验证该结论的合理性,本研究为 ChatGPT 构造了更明确的任务场景,即提供特定的词汇,请求 ChatGPT 依据《中国分类主题词表》编制词表,揭示词间关系,结果如表3所示。与原词表相比,ChatGPT 遗漏了分类号,并且所生成的词间关系与词表中原有的词间关系存在较大偏差。可见,领域术语的多样性和复杂性,使得 AI 模型无法很好地识别和理解,导致词表的准确性和完整性受到影响。因此,短期内在专业领域或特定应用场景中人工编制词表仍有必要,这不仅有助于提高词表的可解释性和可控性,同时也能够为 AI 模型的训练提供丰富的语料基础。

在知识组织系统的设计与实现任务中,首先 Chat-

表3 ChatGPT编制词表与原词表对比结果

《中国分类主题词表》原词表(部分)	ChatGPT编制的词表
超音速流 Y 超音速流动 超音速流动 0354.3 D 超音速流 Z 流动(力学) C 超音速风洞 C 超音速进气道 C 喉管 超音速风洞 V211.74 Z 风洞 C 超音速流动 C 吹吸式风洞 C 二元风洞 超音速进气道 V228.7 Z 进气道 C 超音速流动 C 高超音速进气道	超音速流 D 超音速流动 超音速流动 Y 流动(力学) 超音速风洞 Y 流体力学试验 超音速进气道 Y 航空发动机

GPT生成了一个整体的设计方案,主要步骤包括确定目标用户和任务场景、建立本体和词汇表、设计分类体系和标签系统、建立知识图谱、设计搜索和推荐系统以及建立知识共享平台。当进一步限定数据源、信息资源类型和使用场景之后,ChatGPT提供了更加详细的指导,比如如何对文本数据进行标注、分类,如何将多媒体数据与文本数据进行关联,如何从不同维度设计检索和浏览的功能等。在与 ChatGPT 的连续交互中发现,ChatGPT 可以提供知识组织系统设计与实现的方案,辅助使用者进行整体的规划和布局,甚至具体到工具的使用、代码实现。但 ChatGPT 生成方案的详略程度、参考价值与用户提问的方式、提示的策略有很大关系。

在知识组织系统的互操作任务中,ChatGPT 给出了两种方案:一是使用关键词和主题进行映射。将“脂肪代谢障碍”与 LCSH(美国国会图书馆标题表)中的相关主题词进行匹配,再对应到 DDC 中的相关分类号;二是使用分类号实现互操作。因为 LCSH 的主标题及部分复分词后附有 LCC 的分类号^[23],ChatGPT 提出可以先查找 LCSH 中与“脂肪代谢障碍”相关的分类号,然后映射到 DDC 中的分类号。经多轮测试以及与 ChatGPT 的互动得知,ChatGPT 生成 KOS 互操作方案的依据是已有的词表和国际标准化组织(ISO)制定的一些词表互操作标准。由此说明,语料准备工作对充分发挥 ChatGPT 的能力以处理知识组织任务至关重要。

5 结论与建议

5.1 研究结论

在对 ChatGPT 完成知识组织任务的效果进行分析和评估后,主要得出以下结论:

(1) ChatGPT 具有较强的自适应性,可以根据知识组织的具体任务(如分类和主题标引、本体构建等),基于训练数据自适应地生成内容,适用于不同场景下的知识组织研究与实践。此外,ChatGPT 具备的思维链能力能够为知识组织领域基于知识关联的推理任务注入强大动能。

(2) ChatGPT 能够自动完成简单的知识组织任务,如编目、标引、元数据生成等。其工作机制是基于对现

有的分类表、主题词表、术语表等知识组织系统的标引和资源描述规则进行学习后自动生成的结果。但由于训练数据的不足或不全面,以及信息缺失或歧义等问题使自动分类、主题标引结果的准确率不高。由此说明,高价值语料是生成式 AI 的基础。充分发挥生成式 AI 在知识组织领域的应用价值,首先需要做的便是语料准备工作(尤其是中文语料)。

(3) 对于复杂的知识组织任务,如本体构建、知识图谱构建、词表编制,并不能单纯依靠 ChatGPT 自动实现。即在生成知识组织领域的产品、工具和规则的现实任务中,ChatGPT 能力是有限的。首先,无论是词表编制还是知识库构建,都需要深入理解特定领域的概念、术语和知识体系,以及领域内的实体、属性和关系等元素,这需要领域专家的专业知识和经验来指导,使知识组织系统具有更高的可解释性和可控性;其次,知识组织系统的功能需要与任务场景相结合,用户需求的表达和反馈是开发、优化知识组织系统的关键,而对用户需求的理解和解读需要知识组织领域专业人员的思维和判断力。

(4) ChatGPT 的创造力是基于现有的、成熟的知识组织领域的知识、标准、规范而来,并不会超出现有的知识范围。在知识组织系统设计与实现的任务中,ChatGPT 所给出的设计方案,是参考了已有的标准、文献、专业知识和实践经验后整合的结果。它对于减少个体与个体之间的信息不对称问题,面向知识组织个体研究者或实践工作者解决具体问题有相当的优势,然而对于一个专业领域的贡献却是微乎其微。

(5) 知识组织与 AI 融合要强调提示工程的重要性。在知识组织的复杂任务中,ChatGPT 可以为知识组织者提供方案、生成阶段性知识组织产物(比如抽取的知识元、赋予的分类号等)以及代码设计等技术支持。面对庞大的、复杂的知识组织工程,ChatGPT 尚不能直接生成一个产品,但可在项目实施的各个环节发挥作用。而充分发挥生成式 AI 的功能,提高知识组织工作效率的关键是有效利用 Prompt 来提问。这对知识组织领域的研究人员和工作者提出了新的挑战,如何根据期望输出,为 ChatGPT 类 AI 模型预设角色,更好地组织问题输入是能否得到有用答案的前提。那么,与生成

式AI的交互规范、模式、逻辑或许将成为知识组织领域的新研究问题。

(6)知识组织领域的未来发展趋势必然是与ChatGPT等生成式AI的深度融合。知识组织研究和实践工作将在生成式AI的支持下更加智能和高效。与此同时,传统的知识组织工作仍然有存在的价值,如专业领域词表的编制,知识组织系统标准和规范的研制等,这些是大语言模型训练的语料来源和基础。

(7)ChatGPT类生成式AI在简单、规范化、程序化的知识组织任务中的胜任力,对知识组织专业人才核心能力的培养提出了挑战。操作能力将不再是知识组织人才培养的重点,而创造能力的培养会越来越重要。

5.2 知识组织领域未来发展建议

基于研究结论,针对知识组织领域未来的研究与实践、知识组织人才核心能力培养、知识组织领域专家与AI的有效协作提出如下建议:

(1)需要开展知识组织领域语料库的开发与利用工作。将集成各领域知识和专家经验的成熟的知识组织系统(词表、本体、知识库等)作为丰富的语料来源,根据生成式AI的需求标注训练数据,使AI深入学习知识组织领域的理论基础、方法、规范体系,以提高生成内容的准确性。

(2)探索知识组织与ChatGPT等生成式AI深度融合的理论体系、规则方法以及实现路径。鉴于现阶段生成式AI有限的创造力,知识组织领域的创新与可持续发展仍极大地依赖研究人员和专家智慧,结合ChatGPT类生成式AI的涌现能力和思维链,可加速领域创新进程。

(3)要注重提示工程,探索如何针对知识组织领域的特定任务开发和优化提示,以使用户有效地使用ChatGPT类大语言模型完成各种应用和研究课题。同时,针对不同复杂度的知识组织任务,要分级、分类地探索与ChatGPT等生成式AI的融合方式,提高AI在不同任务场景下的适用性。

(4)拓展ChatGPT等生成式AI在知识组织领域的应用空间,在科学研究、图情机构实际的知识组织业务中,以及其他行业的知识库开发工作中充分调用此类

AI模型,不断探索新的融合模式。在此过程中总结如何更好地解释AI模型的决策过程和预测结果,以增强模型的可解释性、可信度和可靠性。

(5)在知识组织人才核心能力培养方面,要立足于“智慧驱动+场景驱动+创新驱动”的发展定位,着重培养以下能力:①适应能力:适应知识组织业务的智慧化转型,具备与AI协作开展知识组织工作的能力,能够有效地将AI语言模型用于知识组织领域的各种研究和实践任务中。②知识能力:掌握知识组织的基础理论、技术方法、标准规范,能够深度理解AI生成的知识组织结果,并具备解释AI决策过程和预测结果的能力。③场景化能力:识别和拓展AI赋能下的知识组织应用场景的能力,以及对用户需求的理解和解读能力等。④知识组织系统设计能力:对知识组织系统规则、工具、标准、规范的设计与开发能力。此外,理论创新能力、对AI生成内容准确性和完整性的评估能力、跨领域知识组织能力等也是知识组织人才能力培养的重点。

(6)未来知识组织领域专家和ChatGPT等技术之间需要进行有效协作,以更好地完成知识组织工作。首先,领域专家需要参与数据标注以及AI知识组织结果的监督和反馈过程,为利用生成式AI完成知识组织任务提供高质量的训练数据,并利用专家知识和经验对数据标注以及AI生成内容的质量进行评估、反馈和优化,提高利用AI开展知识组织工作的准确性;其次,专家需要不断探索、定期更新和补充新的领域知识,交由生成式AI进行训练,以保证知识组织工作的及时性;此外,专家应基于对知识组织服务对象的需求和使用场景的了解,参与Prompt优化和人机交互设计的过程,为知识组织工作者利用AI完成知识组织任务提供更高效、易用的途径。

直面ChatGPT等生成式AI带来的机遇与挑战,探索与新一代人工智能深度融合的路径,是知识组织未来发展的重要任务。本研究基于实验案例的分析与建议,旨在为新一代人工智能背景下的知识组织研究与实践以及专业教育等未来发展问题提供参考。然而,研究本身也存在一些局限性:任务设计未能覆盖知识组织体系的方方面面,研究结论不够全面。未来将从

更多的维度设计多样化、深层次的知识组织任务,以揭示生成式AI用于知识组织任务的潜在优势和问题,丰富研究结论。

参考文献

- [1] 蒋永福,付小红.知识组织论:图书情报学的理论基础[J].图书馆建设,2000(4):14-17.
- [2] 曾建勋.迎接ChatGPT即将带来的颠覆性变革[J].农业图书情报学报,2023,35(1):9-11,27-28.
- [3] 陆伟,刘家伟,马永强,等.ChatGPT为代表的大模型对信息资源管理的影响[J].图书情报知识,2023,40(2):6-9,70.
- [4] 曹树金,曹茹烨.从ChatGPT看生成式AI对情报学研究与实践的影响[J].现代情报,2023,43(4):3-10.
- [5] 赵瑞雪,黄永文,马玮璐,等.ChatGPT对图书馆智能知识服务的启示与思考[J].农业图书情报学报,2023,35(1):29-38.
- [6] 蔡子凡,蔚海燕.人工智能生成内容(AIGC)的演进历程及其图书馆智慧服务应用场景[J].图书馆杂志,2023,42(4):34-43,135-136.
- [7] de Campos L M, Romero A E. Bayesian network models for hierarchical text classification from a thesaurus[J]. International Journal of Approximate Reasoning, 2009, 50(7): 932-944.
- [8] 李娇,黄永文,罗婷婷,等.基于多因子算法的自动分类研究[J].数据分析与知识发现,2020,4(11):43-51.
- [9] 梁红兵,杨铭魁,黄晓.基于反馈规则学习的医学文献主题自动标引方法[J].计算机工程,2003(11):174-176.
- [10] 陈淑平,梁东魁.基于机器学习的扫描图书元数据自动抽取研究[J].现代情报,2013,33(6):45-48.
- [11] 姜霖,王东波.引文元数据的自动发现和标注方法研究——以外文引文为例[J].数据分析与知识发现,2017,1(1):47-54.
- [12] Macfarlane A, Missaoui S, Frankowska-Takhari S. On machine learning and knowledge organization in multimedia information retrieval[J]. Knowledge Organization, 2020, 47(1): 45-55.
- [13] 郭思成,李纲,周华阳.基于Word2Vec的医学知识组织系统互操作研究——以词表间语义映射为例[J].情报理论与实践,2019,42(9):160-165,176.
- [14] 陈晓燕,贾珊,何有世.基于深度学习的领域本体构建研究——以汽车领域为例[J].情报杂志,2020,39(1):174-178.
- [15] 刘涛,舒红平.基于无监督学习的中医古籍文献本体构建方法研究[J].电子技术与软件工程,2016(12):169.
- [16] Bangyal W H, Rehman N U, Nawaz A, et al. Constructing domain ontology for alzheimer disease using deep learning based approach[J]. Electronics, 2022, 11(12):1890.
- [17] 李涛,王次臣,李华康.知识图谱的发展与构建[J].南京理工大学学报,2017,41(1):22-34.
- [18] 曾桢,赵浩宇.基于文献的中国近代史知识图谱构建与实证研究[J].数字图书馆论坛,2022(4):35-42.
- [19] Tao F, Hao W. Research of Chinese intangible cultural heritage knowledge graph construction and attribute value extraction with graph attention network - ScienceDirect[J]. Information Processing & Management, 2022, 59(1): 102753.
- [20] 钱玲飞,崔晓蕾.基于数据增强的领域知识图谱构建方法研究[J].现代情报,2022,42(03):31-39.
- [21] 贺德方.国内外知识组织体系的研究进展及应对策略[J].情报学报,2010,29(6):963-972.
- [22] 前瞻网.ChatGPT又又又更新了!推出插件功能,可直接联网与5000+应用交互[EB/OL].[2023-04-20].<https://baijiahao.baidu.com/s?id=1761210585039995250&wfr=spider&for=pc>.
- [23] 赵晓红,宋晓丹.《美国国会图书馆标题表(LCSH)》的使用方法[J].图书馆建设,2003(6):51-53.

[作者简介]曹茹烨,女,1992年生,中山大学信息管理学院博士研究生。

曹树金,男,1962年生,山东理工大学特聘教授,中山大学信息管理学院教授、博士生导师(通讯作者)。

收稿日期:2023-05-24

欢迎订阅

2024年《情报资料工作》杂志

- 中国社会科学情报学会学报
- CSSCI来源期刊
- 全国中文核心期刊
- 中国社会科学院AMI核心期刊
- “复印报刊资料”重要转载来源期刊
- 邮发代号82-22 全年定价288元