

科学论文摘要语义增强形式调查研究^{*}

喻琪琛¹, 王晓光^{1,2}

(1. 武汉大学信息管理学院, 武汉 430072; 2. 武汉大学信息资源研究中心, 武汉 430072)

摘要: 摘要是科学论文的重要组成部分, 为检索和预览科学论文提供基础和便利, 对摘要进行语义增强有助于加强摘要的情报功能, 提高读者的信息获取效率。对摘要增强方式进行调查和比较研究可以明确摘要语义增强的类型和发展方向。利用网络调查法对26本样本期刊中的摘要增强形式进行分类统计和比较分析, 进而对不同领域期刊的语义增强方法和模式进行定性研究, 并在此基础上识别摘要语义增强的关键问题。科学论文摘要的语义增强属于知识组织和表示方法的创新, 主要有结构性增强、媒介性增强和表达性增强三种类型。摘要的语义增强不仅有利于提高读者的信息获取效率, 更重要的是满足日益增长的机器读取和理解文章内容, 以及获取科技情报的需要, 代表科学交流系统向富语义出版形态发展的趋势。

关键词: 语义出版; 语义增强; 结构性增强; 媒介性增强; 表达性增强

中图分类号: G250.73

DOI: 10.3772/j.issn.1673-2286.2017.08.002

1 引言

进入大科学时代以来, 科学论文数量呈爆炸性增长。2014年, SCI数据库收录论文总数超过176万篇^[1], 生命医学数据库MEDLINE每年新增50万篇文献。快速增长的科学论文在丰富人类知识资源的同时, 也带来严重的信息过载问题。在流行病学领域, 研究者平均每天用约21小时进行阅读, 才能在该领域保持领先地位^[2]。为便于读者快速获取和筛选学术信息, 科学论文摘要的语义增强问题逐渐引起科学交流系统建设者的关注。

摘要是科学论文的重要组成部分, 在信息搜寻和筛选过程中扮演重要角色。作为“提供文献内容梗概为目的, 不加评论和补充解释, 简明、确切地记叙文献重要内容的短文”^[3], 摘要揭示了一篇论文内含的主要科学情报。通过预览摘要, 读者可快速判断论文价值并决定是否进行阅读和引用, 这对于提高科学交流的效率至关重要。

随着科学交流领域语义出版模式的兴起, 借助各种形式的语义技术改进论文摘要的形式, 提高摘要的语义结构化和可计算化能力, 不仅有利于读者开展传统的略读, 还支持新兴的、借助机器实现的、在摘要自动集成

分析基础上提供的“战略阅读”(strategy reading)^[4]。

目前, 关于摘要语义增强的研究与实践刚开始, 如何平衡摘要语义表现的丰富化和简约化的矛盾是语义出版领域文献建模上的新难题。

2 研究基础

2.1 摘要的基本功能与作用

早在古希腊时期, 人们在处理文献时就采用类似摘要、节录的方式。现代意义上的论文摘要出现于20世纪70年代, 部分期刊将“总结”与“结论”部分从文末转移至正文之前, 形成期刊论文摘要的雏形。

“摘要”的具体定义很多。Swales认为, 摘要是置于学术论文正文之前, 对正文内容和结构的预先指示^[5]; Cleveland等认为, 摘要是知识记录的核心内容, 可作为该篇文章的替代品^[6]; Graetz指出, 摘要能帮助读者节省时间, 读者不用通读全文即可找到文章的各部分, 且预先了解文章结构可以帮助读者更好地深入文章, 如果摘要概括得足够全面, 也许就可以替代全文阅读^[7]; Orasan认为摘要不是文章的一面“镜子”, 而是将读者注意力吸引

^{*} 本研究得到中组部“万人计划”和教育部“新世纪优秀人才”项目资助。

到文章包含的最重要信息上^[8]。通过这些定义发现摘要具有选择性与预示性两个基本属性,即摘要必须选择展示论文的核心内容,使读者借助摘要可预知论文的结构安排及其大概内容。

基于以上两种特征,论文摘要在科学交流过程中的作用可分为两个方面:一是提高检索和阅读效率,读者通过浏览摘要即可获知论文报告的研究对象、方法、结论或创新发现等主要科学情报,进而决定是否需要阅读全文;二是提高科学交流效率,促进学术论文的传播。由于摘要能揭示论文的核心信息与结构,可被视为高度浓缩版的论文,所以在学术期刊审稿时,摘要可作为重要的决策依据,这样有利于提高审稿效率。如 *The British Medical Journal* 的论文平均录用率为5%,很多审稿人是在阅读摘要后就决定退稿的^[9],这样可大幅缩短审稿周期。

2.2 摘要的要素与结构

围绕特定的功能,论文摘要需具备一定要素与结构。这方面研究一直是摘要的重点研究内容,涉及文学、编辑学、图书情报学等多个学科,具体的研究角度包括修辞、语步、语法等多个方面。

在文学领域,论文摘要是一个独立的语类研究对象。摘要结构和必备要素的研究通常从语篇分析入手,分析摘要的语步构成。Swales等认为语步是一组完成特殊任务的文字,可以是一个短语,也可以是一个段落^[10]。从语篇结构角度看,多数论文摘要与学术论文有相同的逻辑结构。学术论文的基本逻辑结构包括导言(introduction)、方法(methods)、结果(results)、讨论(discussion)四个部分,即IMRD模式^[11]。为与论文对应,很多论文摘要也有明确的四个部分,称为四语步模式。此外,还有六语步模式(综述、目的、方法、结果、结论、建议),这在生物医学论文摘要中常见。Salager-meyer认为目的、方法、结果和结论是生物医学论文摘要的必备要素,综述与建议则是可选项^[12]。目前,摘要的四语步结构得到多数学者的认可,这种结构与论文的章节相对应,内容完整,体现论文的论证逻辑。通过研究发现,摘要的要素与结构通常与论文内容结构相对应。

摘要结构的完整性对论文的录用有重要影响。Vaughn等研究发现,摘要结构有欠缺的论文比摘要结构完整的论文被拒的概率高3倍^[13],这意味着摘要

是科学论文的重要组成部分,在科学交流中扮演着重要的角色,具有不可忽视的价值。

2.3 摘要的形式与类型

摘要的要素与结构决定摘要不同的形式与类型。传统论文摘要可分为知识性摘要(也称说明性摘要或描述性摘要)、评价性摘要(也称批评性摘要)与信息性摘要(也称报道性摘要或资料性摘要)三种。科学论文一般采用的是信息性摘要,又可分为非结构式摘要(也称一段式摘要)与结构式摘要两种类型。知识性摘要、评价性摘要和一段式摘要属于文字式摘要发展的第一阶段。随着期刊数量和论文数量的快速增长,以及计算机技术在文献检索中的普及应用,这些摘要形式表现出信息含量小、格式不固定、层次结构不清晰以及不利于文献检索等局限^[14],而逐渐被期刊放弃。

20世纪80年代以来,文字式摘要的发展进入第二个阶段。信息性摘要因为情报含量大,对读者的价值高,逐渐被多数期刊采纳,成为主流形式的摘要类型。与此同时,结构式摘要开始出现,并逐渐在部分自然科学期刊中流行。结构式摘要通过明确的标示,将一段式摘要的语步进行结构化划分,形成段落分离、结构清晰的分段式摘要。这种结构化处理使阅读效率得到提升,摘要的写作质量和蕴含的情报量也得到提升。尽管结构式摘要在信息含量和表达上有明显优势,但在人文社会科学领域推广速度比较缓慢。

随着科学论文的快速增长和出版环境的变迁,传统论文摘要越来越不能满足科研工作者快速获取科学情报的需求和日益增强的机器读取与理解文献需要。近年来,语义出版在学术出版界兴起。Shotton将语义出版阐释为借助语义手段实现科学论文中数据、信息、知识等的识别、抽取、整合以及再利用^[15],以帮助用户获取信息和理解知识。受此思想影响,面向用户需求,实现科学论文内容结构化、对象关联化、知识可视化成为语义出版发展以及构建语义出版物的目标与要求^[2,16]。为实现此目的,语义出版对科学论文的内容及形式进行结构化处理与语义标注。然而,在目前科学论文的语义增强处理中,大多将摘要整体作为一个文本块进行粗粒度标注,而忽略了摘要自身蕴含的丰富信息,且对摘要的处理未形成清晰一致的认知或规范。因而,摘要的语义增强问题逐渐凸显。

3 摘要语义增强的类型

摘要语义增强的目标是借助语义技术实现人与机器获取信息及科学交流效率的提升。具体需要丰富摘要文本的内容与表达,提高机器提取关键科学情报的能力。目前,摘要语义增强的形式大体分为三类,分别是结构性增强、媒介性增强和表达性增强。

3.1 结构性增强

摘要的结构性增强实践较丰富,结构式摘要为典

型代表,其最早由加拿大麦克马斯特大学临床流行病学和生物统计学教授Haynes于1985年提出。*Annals of Internal Medicine*1987年第4期首先采用结构式摘要。根据结构式摘要含有的项式数目,可将结构式摘要分为八项式、五项式、四项式等类型。

黄河清等曾对结构式摘要的期刊应用情况进行统计^[17],本文在该调查的基础上,选取18种知名科技期刊作为研究样本,统计分析结构式摘要的类型与要素,如表1所示。

结构式摘要在科技期刊领域应用已超过30年,在形式、内容等方面发展已较成熟。结构式摘要的使用在20

表 1 18种科技期刊的结构式摘要类型及要素

期刊名称	示 例	类 型	摘要包含的要素
<i>Journal of the American Medical Association</i>	http://jama.jamanetwork.com/article.aspx	八项式	重要性、目的、设计、主要成果与方法、结果、结论与相关、实验登记
<i>Lancet</i>	http://www.thelancet.com/journals/lancet	五项式	背景、方法、发现、反应、资助
<i>Bioinformatics</i>	http://bioinformatics.oxfordjournals.org	五项式	动机、结果、获取、联系方式、补充信息
<i>New England Journal of Medicine</i>	http://www.nejm.org/doi/full/10.1056/NEJMoal503877	四项式	背景、方法、结果、结论
<i>Journal of integrative medicine</i>	http://www.jcimjournal.com/jim	四项式	目的、方法、结果、结论
<i>Biomedical journal</i>	http://www.sciencedirect.com/science/journal/23194170	四项式	背景、方法、结果、结论
<i>Journal of Clinical Oncology</i>	http://jco.ascopubs.org	四项式	目的、病患/方法、结果、结论
<i>BMC Bioinformatics</i>	https://bmcbioinformatics.biomedcentral.com	四项式	背景、结果、讨论、结论
<i>Clinical Infectious Diseases</i>	http://cid.oxfordjournals.org	四项式	背景、方法、结果、结论
<i>Neurology</i>	http://www.neurology.org/content	四项式	目的、方法、结果、结论
<i>British Journal of Pharmacology</i>	http://onlinelibrary.wiley.com/doi/10.1111/bph.13456	四项式	背景和目的、实验方式、主要结果、结论与推断
<i>American Journal of Physical Anthropology</i>	http://onlinelibrary.wiley.com/doi/10.1002/ajpa.22952	四项式	目的、材料与方法、结果、讨论
<i>Journal of Clinical Psychiatry</i>	http://www.psychiatrist.com/JCP	四项式	目的、方法、结果、结论
<i>American Journal of Public Health</i>	http://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2015.303003	四项式	目的、方法、结果、结论
<i>Journal of the American Geriatrics Society</i>	http://onlinelibrary.wiley.com/doi/10.1111/jgs.14231	四项式	背景、方法、结果、结论
<i>Astronomy & Astrophysics</i>	http://www.aanda.org/articles/aa/abs/2016/08/aa26766-15/aa26766-15.html	四项式或 五项式	背景、目的、方法、结果、结论
<i>Oxford Journals</i>	http://cid.oxfordjournals.org/	四项	背景、方法、结果、结论
<i>Circulation</i>	http://circ.ahajournals.org	三项式	背景、方法与结果、结论

世纪末期达到顶峰,目前使用的期刊数量增加不明显,大部分期刊仍是沿用传统的一段式摘要形式见图1^[14]。

与一段式摘要相比,结构式摘要有以下优势:情报含量高、表达准确全面;结构清晰,重点突出,有利于读者筛选论文;审稿与编辑加工方便,便于判断文章价值和创新性,有利于提高审稿速度^[18]。结构式摘要也存在一些缺点,主要表现为结构固定、字数较多、写作略烦琐,所以结构式摘要多用于调查、实验和实证等类型的论文,不适合于理论性和综述性论文。

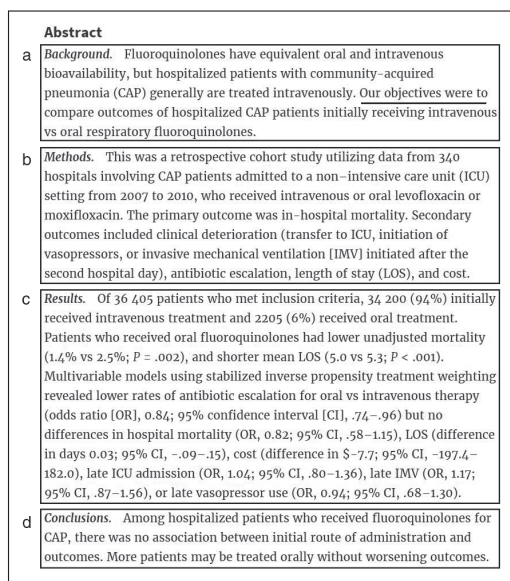


图1 四项语步的结构式摘要示例^[19]

注: a背景部分, 简要介绍当前论文研究主题的背景和意义; b方法部分, 简要介绍实验的方法、过程、材料、步骤等; c结果部分, 简要介绍实验所得数据、新发现等内容, 在摘要中占比最大; d结论部分, 简要介绍根据实验结果进行逻辑分析得出的结论。

3.2 媒介性增强

经过媒介性增强的摘要称为富媒体式摘要。目前, 常见的富媒体式摘要主要有两种类型, 一种是图形摘要, 另一种是视频摘要。

(1) 图形摘要。图形摘要是利用作者提供的、表达重要实验成果或科学发现的图表作为摘要内容的一种摘要形式^[20], 如图2所示。大部分采纳图形摘要的期刊也要求作者提交文字摘要或文字说明, 图表部分与文字部分互为补充。图形摘要弥补文字摘要在信息表达和传递上的不足, 在知识信息表达上更加直观, 被广泛应用于化学、生物等自然科学领域, 能表达复杂的、难以用文字描述的科学知识。

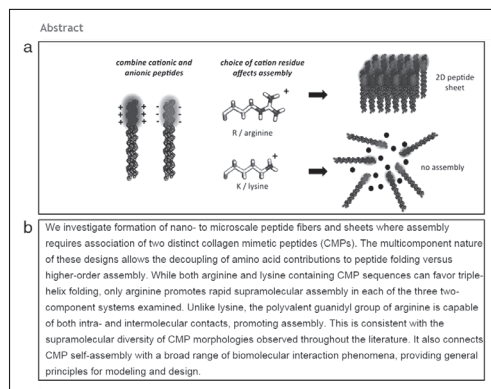


图2 图形摘要示例^[22]

注: a论文摘要的图形部分, 用图形表现两种化合物的化学反应过程; b论文摘要的文字部分, 用文字描述研究过程及其发现。

(2) 视频摘要。视频摘要也是近年来兴起的一种新型摘要形式, 它利用一段简短的视频向读者展示科学研究的过程、步骤、发现等难以用文字准确描述的内容^[21], 如图3所示。视频摘要可视化地揭示某种动态过程, 传递的信息量远大于文字描述。视频形式的内容在论文摘要出现反映出科学研究工具和过程的数字化进程已经相当深入, 很多科学研究的记录从开始就是数字化的, 直接把原始的数字记录发布出来, 比用文字转录更有利于提高科学交流的效率。

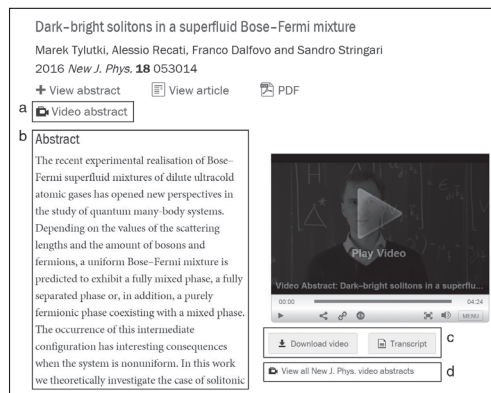


图3 视频摘要示例^[23]

注: a以“摄像机”小图标对视频内容进行标示; b文字摘要部分, 对文章的研究背景、方法、结论进行介绍; c视频内容可作为独立的文档下载, 同时附带视频内容的文字说明; d浏览该期刊所有的视频摘要。

富媒体式摘要以图表/视频与文字相结合的方式, 将摘要内容要素进行多媒体表达, 呈现单一媒介形式的文字和文本无法表示的重要信息和知识。如研究人员直接将实验过程和结果通过视频的方式展现给读者, 使摘要构成要素中的方法和结论部分更加直观。富

媒体代表学术期刊的重要发展方向,随着数字阅读终端的兴起,富媒体式摘要将会越来越常见。

3.3 表达性增强

论文摘要的语义增强既要便于读者阅读,也便于机器读取内容和抽取知识。与媒介性增强相比,结构性增强更符合机器读取内容的需求,但是粒度仍不够细致。为进一步浓缩摘要和提高机器对论文关键科学情报的抽取能力,出版商提出两种新颖的摘要形式,分别是结构化数字摘要(Structured Digital Abstract, SDA)和亮点摘要。

(1) 结构化数字摘要。结构化数字摘要最早由Gerstein和Seringhaus提出,其目的是加强科学文本与科学数据的关联,以便于文本挖掘和知识发现^[24]。Gerstein等认为SDA应当包含三个要素:一是翻译表,包括文章中出现的各种实体的名称(如蛋白质、基因等),这些实体名称可通过实体链接的方式与特定知识库中的条目建立关联;二是由领域本体表示的重要结果清单(如蛋白质间的反应等);三是相关标准编码,用于表示获取实验方法^[24]。

在以上理念指导下,FEBS Letters联合分子间联合作用数据库对结构化数字摘要进行应用尝试,重点对论文中出现的蛋白质及其化学反应进行描述^[25]。语义出版概念提出者Shotton也曾进行过SDA实验^[26]。但这种摘要形式还相当少见,只在少数数字知识基础设施较为完善的领域有所应用(见图4)。

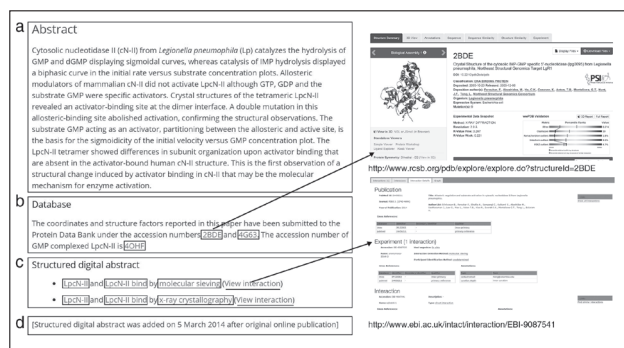


图4 结构化数字摘要应用示例^[27]

注: a传统的一段式摘要部分,对论文背景、方法、目的、结论等要素加以描述; b数据库信息部分,描述当前文章使用到的数据在数据出版服务平台上的代码; c结构化数字摘要部分,用极其简洁的文本描述蛋白质及其化学反应,通过实体链接的方式,将文本中的链接指称与特定知识库或者本体条目进行链接; d结构化数字摘要的编辑说明部分,表明SDA是在2014年3月5日原始论文在线出版后添加的。

(2) 亮点摘要。为进一步浓缩摘要,提高科学交流效率,部分出版商尝试提出亮点摘要。其一般不超过5个句子,每个句子不超过25个单词,由作者提供并与其他形式的摘要内容同时发布,如图5所示。

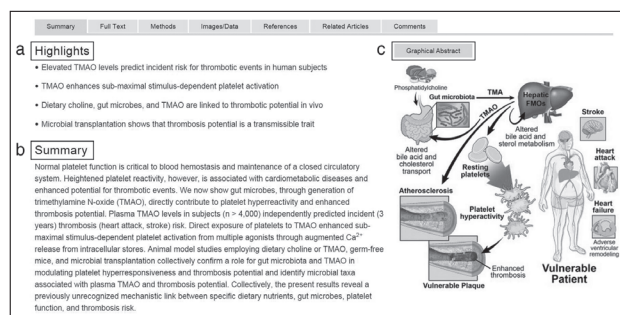


图5 亮点摘要示例图^[28]

注: a亮点摘要部分,用四句话介绍论文最有价值的观点和发现; b传统摘要部分,用一段叙事性文字描述当前研究的目的、过程、意义、结论等内容; c图形摘要部分,用可视化的示意图描绘文字摘要中提及的反应过程。

亮点摘要简单明了,揭示论文中最重要的科学情报。从句法形式看,亮点摘要中的句子多为陈述句和简单句,基本是断言或命题,便于机器的读取和知识抽取,为开展实体链接和RDF转换,以及制作纳米出版物提供便利。

从语义上看,亮点摘要中的句子涉及研究活动的多个方面,大体可分为四类,如表2所示。这些句子是IMRD架构下文章各部分的核心句。从全局作用看,亮点摘要与一段式摘要存在一定的冗余重复,是出版商为高级知识服务作准备的一种数据耕耘策略,其是利用专家而非机器提取论文中的关键情报,为未来的知识发现提供数据基础。

4 摘要语义增强面临的问题

4.1 摘要的简约性与丰富性

摘要揭示长篇论文关键信息的一种策略,可帮助读者决定是否要精读论文。简约性是摘要的基本属性。传统的一段式摘要虽然有明显的简约性,但很容易遗漏重要信息,尤其是方法要素部分^[29],所以对摘要进行语义增强,通过多种途径完善摘要的形式十分必要。但是,摘要的语义增强通常会增加摘要篇幅,这与其简洁性要求有一定矛盾。一般而言,一段式英文摘要篇幅约

表 2 亮点摘要语句类型示例

类 型	示例句
概括研究目标和创新点	Our approach takes into account both direct and indirect effect. We propose a new index to quantify SSRN downloads
指明研究对象和研究方法	We analyze the expert-based ranks in the Finnish publication forum. We analyze data for about 30 000 authors and 367 000 papers
表达作者观点和科学发现	Book chapters are not a suitable unit to use alongside with journal articles in bibliometric studies. We find that papers whose abstracts contain more frequently used words tend to receive slightly more citations
指明研究意义和研究价值	Results may suggest policies and actions to guide us towards a more effective scientific research. The methodology is applied to the Italian case but can be employed elsewhere

200词, 四项结构式摘要通常有400词, 若采用六项式摘要或者多媒体摘要, 其内容更多, 完整阅读一篇文章摘要的时间会更长, 这在一定程度上背离了摘要的初衷, 也成为部分学术期刊拒绝采用增强摘要的原因。

如何在增强摘要的语义丰度和保持其简约性间找到平衡点, 是创新论文摘要形式的关键。相比较而言, 结构性增强和表达性增强比媒介性增强更符合摘要的初衷, 它们利用浓缩的结构、清晰的关键句或文章中已经存在图形、图像, 借助结构性标签或实体链接, 在没有明显增加摘要内容长度的情况下提高内容的“模块化”程度, 以及与外部知识的关联性, 满足读者选择性阅读(如只看摘要的方法部分)或者扩展阅读(如通过超链接查看某个蛋白质关联本体库)的需要。这与语义出版的目标一致, 即通过语义化处理提高文档的可获取性、可读性和可理解性。媒介性增强利用图像和音视频媒介表达数字科学知识, 发挥其在语义表达多样性方面的优势, 也增强摘要的信息含量, 但并不适用于所有类型的论文, 应用空间相对有限。

4.2 面向检索还是面向阅读

在科学交流系统视角下, 论文摘要在印刷时代主要用于著录和检索目的, 由学科编辑和专业图书馆员制作, 到逐渐演化成论文的必要组成部分, 编写责任人也转移到作者。进入数字时代, 论文摘要依旧存在, 但其作用场景发生变化。在期刊出版网站上(如中国知网、Elsevier), 读者通常先利用文章标题、摘要和关键词进行检索, 然后对检索结果进行浏览, 再通过阅读论文摘要判断是否阅读全文。在这个过程中, 摘要的检索价值逐渐让位于筛选价值, 即标题、关键词和全文检索对摘要检索的替代性很高, 摘要已不再是必要的检索点, 其

检索意义略有弱化, 但摘要作为论文全文的预览文本, 其筛选和选择意义越来越突出, 在整个科学交流系统中的价值也没有弱化。

摘要价值的细微变化, 对语义出版系统的开发提出新要求。对摘要进行语义增强必须考虑用户的信息搜寻、决策和阅读过程。对传统信息检索而言, 三种形式的语义增强都不会提高资源的发现能力, 但对于语义检索而言, 摘要的语义增强为后期语义检索提供更多的发挥空间。如可实现方法检索、基于内容的图像检索和基于内容的视频检索, 这在一定程度上可提高检索的精准度和用户的信息获取效率。所以, 不管是用于阅读还是检索, 摘要语义增强都有其必要性, 而且语义增强策略也必须考虑读者的基本需求。

4.3 面向读者还是面向智能代理

从应用主体看, 读者是摘要内容的最终消费者。目前, 常见的三种语义增强形式也主要面向读者, 以自然语言为主, 用于提高读者的阅读效率。这种趋势在未来也难以变化, 因为摘要作为一种信息浓缩的方法和手段, 基本目的就是辅助读者更快地了解文献内容。随着科学交流系统向语义网环境的迁移, 智能代理将越来越多, 它们被设计用来对摘要内容进行读取、理解、分析、挖掘、集成和整合, 以满足读者不同的信息需求(如战略阅读^[30])。在这种情况下, 读者需要借助智能代理快速发现某个学科领域的宏观知识(如基因编辑方法的发展历史与趋势), 要求智能代理能够“阅读”和“理解”论文摘要, 进而通过知识集成、融合与挖掘, 发现宏观知识。

摘要的语义增强提高了摘要的规范化和形式化水平, 有利于智能代理理解数据。在这点上, 摘要的语义

增强与一些新型语义出版物(如纳米出版物^[31]、微型出版物^[32])的设计目标一致,都是通过知识表示和知识组织形式的创新,满足智能代理对知识计算的需求,以支撑数据挖掘和自动的知识发现。

4.4 摘要语义增强的发展趋势

摘要结构性增强与媒介性增强提高了摘要内容的表现形式与表达能力,表达性增强为实现知识计算提供基础。摘要语义增强的进一步发展需要这三种不同增强方式的结合,在细粒度结构性增强的基础上,借助标签云、标签树等可视化方式,使媒介性增强更多维与立体。同时,借助对摘要中知识实体的标注,提升表达性增强的能力,使其更符合更高层次的知识重组与知识关联要求。

摘要语义增强在提升摘要规范化和形式化水平的同时,也使得摘要更便于被机器与人理解,实现摘要功能的完备。摘要是科学论文正文内容的浓缩,与正文存在关联和索引关系,所以,未来有必要将摘要内容与细粒度正文内容进行映射,建立语义链接,方便读者对正文内容进行精确定位。

5 结论

作为科学论文必不可少的组成部分,摘要对于科学成果的发表、论文的获取和文档的阅读十分重要。伴随语义网的发展,对论文及其摘要进行语义增强已经成为发展语义出版系统、提高科学交流效率的重要工作之一。

本文定性分析了论文摘要语义增强的目标、需求和影响因素,并通过调查研究,首次总结提炼出结构性增强、媒介性增强和表达性增强三种增强类型,并对比分析三种增强形式的差异和特点。论文摘要的语义增强,代表科学论文和科学交流系统向富语义出版框架发展的趋势。越来越多的期刊开始语义增强实践,但不同形式的语义增强效果如何,以及不同学科领域的接受情况都值得未来从用户角度进行评估研究。

参考文献

[1] 2015年中国科技SCI论文统计结果[EB/OL].[2016-04-05].<http://info.howsci.com/447.html>.

- [2] 徐丽芳,丛挺.数据密集、语义、可视化与互动出版:全球科技出版发展趋势研究[J].出版科学,2012,20(4):73-80.
- [3] 中国知网.国标文摘编写规则(GB 6447-86)[EB/OL].[2016-04-05].<http://xuewen.cnki.net/CJFD-LXYB199605036.html>.
- [4] RENEAR A H,PALMER C L.Strategic reading,ontologies,and the future of scientific publishing[J].Science,2009,325(5942):828.
- [5] SWALES J M.Genre analysis:English in academic and research settings[M].[S.1.]:Cambridge:Cambridge University Press,1990.
- [6] CLEVELAND D B,CLEVELAND A D.Introduction to indexing and abstracting[M].[S.1.]:Santa Barbara:Libraries Unlimited,1983.
- [7] GRAETZ N.Teaching EFL students to extract structural information from abstracts[C]//International Symposium on Language for Special Purposes.1982:23.
- [8] ORASAN C.Patterns in scientific abstracts[C]//Proceedings of Corpus Linguistics 2001 Conference.2001:433-443.
- [9] 牛桂玲.中外学术论文中英文摘要语料库的创建及应用[M].北京:知识产权出版社,2013.
- [10] SWALES J,FEAK C B.Abstracts and the writing of abstracts[M].Ann Arbor:The University of Michigan Press,2009.
- [11] BRUCE N J.Rhetorical constraints on information structure in medical research report writing[C]//ESP in the Arab World Conference.Birmingham:University of Aston,1983.
- [12] SALAGER-MEYER F.Discoursal flaws in medical English abstracts: a genre analysis per research-and text-type[J].Text-Interdisciplinary Journal for the Study of Discourse,1990,10(4):365-384.
- [13] VAUGHN H D L,ONWUEGBUZIEN A J.Quality of abstracts in articles submitted to a scholarly journal:a mixed methods case study of the journal research in the schools[J].Library & Information Science Research,2010,32(1):53-61.
- [14] 张春芳,孔燕.结构式摘要的形成、发展及现状——兼谈医学论文摘要的英译[J].医学与哲学,2011,32(17):77-79.
- [15] SHOTTON D.Semantic publishing:the coming revolution in scientific journal publishing[J].Learned Publishing,2009,22(2):85-94.
- [16] 王晓光,陈孝禹.语义出版的概念与形式[J].出版发行研究,2011(11):54-58.
- [17] 黄河清,韩健,张鲸惊,等.中外科技期刊英文摘要文体格式的变化及建议[J].中国科技期刊研究,2015,26(2):143-151.
- [18] 张春芳,刘雪立.国外结构式摘要研究回顾——以James Hartley的研究为例[J].中国科技期刊研究,2012,23(1):56-61.
- [19] BELFORTI R K,LAGU T,HAESSLER S,et al.Association between initial route of fluoroquinolone administration and outcomes in patients hospitalized for community-acquired pneumonia[J/OL].Clinical Infectious Diseases,2016,63(1):1-9[2016-04-05].<https://doi.org/10.1093/cid/ciw001>.

- org/10.1093/cid/ciw209.
- [20] Elsevier. Graphical abstracts[EB/OL].[2014-05-10].<http://www.Elsevier.com/journal-authors/graphical-abstract>.
- [21] CELLPress.Viedo Abstract Guidelines[EB/OL].[2016-05-10].
<http://www.cell.com/video-abstract-guidelines>.
- [22] PARMAR A S,JAMES J K,GRISHAM D R,et al.Dissecting electrostatic contributions to folding and self-assembly using designed multicomponent peptide systems[J/OL].JACS,2016,138(13):4362-4367[2016-03-11].<http://pubs.acs.org/doi/abs/10.1021/jacs.5b10304>.
- [23] TYLUTKI M,RECATI A,DALFOVO F,et al.Dark-bright solitons in a superfluid bose-fermi mixture[J/OL].New Journal of Physics, 2016,18(5)[2016-05-09].<http://iopscience.iop.org/article/10.1088/1367-2630/18/5/053014/meta>.
- [24] GERSTEIN M,SERINGHAUS M,FIELDS S.Structured digital abstract makes text mining easy[J].Nature,2007,447(7141):142.
- [25] LICATA L,BRIGANTI L,PELUSO D,et al.MINT,the molecular interaction database:2012 update[J].Nucleic Acids Research,2012, 40(1):857-861.
- [26] SHOTTON D.Semantic publishing:the coming revolution in scientific journal publishing[J].Learned Publishing,2009,22(2):85-94.
- [27] SRINIVASAN B,FOROUHAR F,SHUKLA A,et al.Allosteric regulation and substrate activation in cytosolic nucleotidase II from Legionella pneumophila[J].Febs Journal,2014,281(6):1613-1628.
- [28] CELL.Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk[EB/OL].[2016-03-10][2016-02-11].[http://www.cell.com/cell/abstract/S0092-8674\(16\)30113-1](http://www.cell.com/cell/abstract/S0092-8674(16)30113-1).
- [29] HARTLEY J,SYDES M,BLURTON A.Obtaining information accurately and quickly:are structured abstracts more efficient?[J]. Journal of Information Science,1996,22(22):349-356.
- [30] RENEAR A H,CAROLE L P.Strategic reading,ontologies,and the future of scientific publishing[J].Science,2009,325(5942):828.
- [31] 吴思竹,李峰,张智雄.知识资源的语义表示和出版模式研究——以 Nanopublication为例[J].中国图书馆学报,2013,39(4):102-109.
- [32] CLARK T,CICCARESE P N,GOBLE C A.Micropublications:a semantic model for claims,evidence,arguments and annotations in biomedical communications[J].Journal of Biomedical Semantics,2014,5(1):28.

作者简介

喻琪琛,女,1993年生,硕士研究生,研究方向:数字出版、语义出版。

王晓光,男,1978年生,博士,教授,博士生导师,研究方向:数字信息资源管理、信息组织与知识服务、语义出版,E-mail: wxguang@whu.edu.cn。

Towards the Semantic Enhanced Forms of Scientific Paper Abstract

YU QiChen¹, WANG XiaoGuang^{1,2}

(1. School of Information Management, Wuhan University, Wuhan 430072, China;

2. Center for Studies of Information Resources, Wuhan University, Wuhan 430072, China)

Abstract: As an important part of scientific paper, abstract provides convenience for readers to access and preview scientific papers. Semantic enhancement of paper abstract contribute to information access. The investigation and comparative study classifies the semantic enhanced forms of paper abstract. Based on 26 journals, the semantic enhanced forms of paper abstracts are analyzed, the key issues and advantages are also recognized and compared as well. Semantic enhancement of paper abstract is a kind of innovation of knowledge organization and representation. There are three categories: structural enhancement, media enhancement and expressive enhancement. Semantic enhancement paper abstract not only helps to improve the reader's information access efficiency, but also meets the growing needs of machine readable content.

Keywords: Semantic Publishing; Semantic Enhancement; Structural Enhancement; Media Enhancement; Expressive Enhancement

(收稿日期: 2017-08-09)