

# 数字学术：概念、特征与案例分析\*

曾熙<sup>1</sup> 王晓光<sup>2</sup>

(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:** 数字技术在学术研究中的应用带来了学术研究环境的深刻变化, 催生了“数字学术”的概念。通过文献调查, 本文对数字学术的历史沿革和基本概念进行梳理, 并基于学术生命周期理论对学术活动各环节的数字化特征进行深入分析。通过网络调研, 以“贝尔法斯特团队诗歌|网络”项目为例, 对当前数字学术研究过程与特征进行验证。最后, 通过对数字学术的反思与总结提出数字学术并非无所不能, 研究人员应以开放、包容、客观、冷静的态度看待数字学术热潮。

**关键词:** 数字学术; 数字化; 数字技术; 数据驱动研究范式; 学术交流

**中图分类号:** G250

**DOI:** 10.3772/j.issn.1673-2286.2019.03.001

数字技术的发展与应用为学术研究带来了深刻的变化。一方面, 移动网络、自媒体的发展和数字出版的普及使得互联网中可获取的原生与非原生数字学术资源急速增长; 另一方面, 数字环境下传统研究问题的拓展与新兴研究问题的涌现推动了学术研究视角、研究思维与研究方法的变革。在此背景下, “数字学术”这一概念应运而生。近年来, 国外众多图书馆纷纷开始创建数字学术中心, 开展数字学术研究项目, 并为研究人员提供数字学术支持。这一现象引起了国内外图情界的广泛关注, 有关数字学术的研究层出不穷。然而, 当前关于数字学术的研究主要集中在数字学术中心的空间建设、人员设置、服务方式、服务特色等实践问题上, 对数字学术的历史沿革、基本概念、主要特征等理论问题则鲜有涉及。理论研究的不足制约了数字学术的深入化、系统化发展。因此, 有必要对数字学术的概念进行科学明确的界定, 对其特征进行合理的阐释。

## 1 数字学术的概念

### 1.1 数字学术的起源与发展

数字学术这一概念兴起于20世纪90年代末, 但它并不

是一个全新的事物, 而是继数字人文(digital humanities)、电子科研(e-science)、社会计算(social-computing)、电子学术(e-scholarship)等术语之后, 随着数字技术在学术活动中的应用不断发展成熟起来的<sup>[1]</sup>。事实上, 将数字技术应用于学术研究已经有很长的历史。

1949年, Busa神父利用IBM穿孔机计算机为圣托马斯·阿奎那的神学著作编制索引被认为是最早的数字学术活动<sup>[2]</sup>。1951年, Baglev尝试利用计算机检索代码制作文摘, 这一研究被称作“电子出版物雏形”的诞生<sup>[3]</sup>。1959年, 美国匹兹堡大学建立全文法律信息检索系统<sup>[4]</sup>。1972年, IBM公司开始研制实验型关系数据库管理系统, Nelson和Lancaster提出了未来研究人员在大型计算机网络中存储和检索学术研究文献的美好设想<sup>[5]</sup>; 一些出版商也看到了学术检索的商机, 开始构建学术数据库。20世纪70年代末, 美国的DIALOG系统向公众提供联机全文数据库检索服务<sup>[4]</sup>。1980年, 新闻组Usenet诞生, 用户可以在新闻组上就特定主题进行自由讨论。20世纪80年代初, UMI、OVID、EBSCO等出版商纷纷开始电子出版物的制作与发行; 80年代中期, 随着个人电脑的兴起, 电子邮件开始被广泛用于学术交流。20世纪90年代初, 互联网的应用为来自世界各地的研究人员开展网络协作创作了条件; 90年代末, 研究者开始利用

\*本研究得到教育部人文社会科学重点研究基地重大项目“大数据资源语义表示与组织研究——面向文化遗产领域”(编号: 16JJD870002)资助。

数字技术实现可视化、地理空间展示、模拟空间,以及复杂系统的网络分析<sup>[6]</sup>。进入21世纪,随着大数据、移动互联网、云计算等数字技术的出现,多样化的数字研究方法与研究工具逐渐渗透到学术研究的各个环节,学术研究环境发生巨大的变化:跨时空、跨学科研究问题与各种新兴研究问题不断涌现;数字期刊、书籍、图像、音频、视频、3D模型等形式多样的数字文献层出不穷;文本分析、数据挖掘、GIS数字地图、三维建模等功能丰富的数字研究方法竞相错出;多媒体出版物、微型出版物、纳米出版物、增强型出版物等语义丰富的数字出版物屡见不鲜;开放获取、学术论坛、学术信息门户、学术社交网站等学术传播与交流方式层见迭出。数字学术时

代已然到来。

从数字学术的发展进程可以看出,数字学术是一个仍在不断发展的新兴领域,从简单的索引编制到研究问题、研究材料、研究方法、研究成果与交流方式的巨大改变,数字学术在数字技术的进步与应用中不断改写着其内涵与边界,其概念也在不断丰富。

## 1.2 数字学术的定义

鉴于数字学术尚未完全成熟,其内涵与外延仍在不断演变与扩充,目前学界尚未对其定义形成统一的认识,不同学者对此给出了不同看法(见表1)。

表1 数字学术的定义

来源	观点	侧重点	年份
文献[7]	利用数字证据、数字循证方法、数字研究、数字出版和保存来达到学术研究目标	学术交流	-
文献[8]	数字证据和方法的使用、数字创作、数字出版、数字内容管理和保存,以及对学术成果的数字使用和再利用	学术交流	2013
文献[9]	通过诸如高性能计算、可视化和大数据集处理等技术对大多数学术工作领域进行转化	研究范式	2014
文献[10]	包括但不限于数据分析和可视化,文本编码,计算文本分析,地理信息系统和数字地图,3D建模,数字馆藏和展览,以及数字化与图像、音频、元数据的创建等	学术交流	2016
文献[11]	GIS数字地图、模拟材料数字化成像、数字收藏制作、创建元数据、数据保存、数据策展与管理、三维建模和打印、数据统计分析、数字展览、项目规划、数字出版、项目开发与管理、文本计算与分析、界面设计及可用性、可视化展示分析、数据库开发、技术维护、编码内容、数字学术软件开发等	数字项目全过程	2016
文献[12]	数字学术是利用数字化工具手段和数据进行知识生产和知识扩散的完整过程	知识生产与扩散全过程	2017
文献[13]	贯穿整个研究过程,具体来说,是从结果到过程、从局部到全局、从孤军奋战到团队协作,从单学科到跨学科,从个体意愿到制度意识,从学术孤岛到数字学术生态等	学术研究全过程	2017
文献[14]	不仅包括利用数字工具进行学术研究,也包括在学术活动中利用数字媒介进行信息交流和成果展示	学术研究、交流与展示	2017
文献[15]	利用数字技术和工具,共同创造、发布、利用的学术表达	方法论与研究范式	2018

在数字学术的定义方面,一部分学者强调数字技术与工具对学术活动的介入,另一部分学者则通过列举具体的数字研究技术与研究实践来界定其范围。总体来看,尽管学者们对数字学术的理解各不相同,但都强调数字技术与工具对学术研究的渗透。在数字学术的侧重点方面,有学者聚焦于数字技术对学术交流模式的影响,也有学者侧重于讨论数字环境下研究范式与研究方法的变革,还有学者认为数字学术贯穿于学术研究全过程。本文认为学术研究是一项复杂的系统性工作,各个环节紧密相连,不可分割,数字技术的介

入必将带来学术活动全过程的深刻变化,而不是仅仅影响其中某一个或几个环节。因此,综合上述观点,本文将数字学术定义为利用数字技术与方法介入学术研究全过程的新型研究模式。

## 1.3 相关概念辨析

除数字学术外,数字人文、社会计算、e-science等概念也常被用来形容利用数字技术进行学术研究,厘清这些概念之间的区别与联系对正确认识与理解数字

学术至关重要。

数字人文这一概念源自“人文计算”，是指在数字技术支持下开展人文研究而形成的新型跨学科研究领域<sup>[16]</sup>。社会计算是指面向社会科学的计算理论和方法<sup>[17]</sup>。e-science是一种计算密集型的科学研究模式，强调在分布式研究网络下，利用海量数据集和网格计算开展科学研究<sup>[18]</sup>。尽管概念中并未刻意强调e-science适用的学科范围，但是对e-science项目的调研发现，目前e-science的应用领域主要集中在生物科学、医药卫生、工业技术、物理学、地球科学、天文学等自然科学领域<sup>[19]</sup>。

通过对数字人文、社会计算、e-science和数字学术概念与应用的分析，发现它们都强调数字技术对学术研究的介入。而这些概念的不同之处在于其涵盖的学科范围不同：数字人文专注于文学、历史、艺术、哲学等人文科学；社会计算聚焦于经济、政治、法律、民族、宗教等社会科学；e-science关注自然科学；数字学术的学科范围则更广，自然科学、社会科学与人文科学领域都在其研究范围之内。

## 2 数字学术的特征

数字技术的利用贯穿于学术研究全过程，是数字学术的前提与基础。因此，数字化是数字学术的核心特征。为了更准确、更深入地阐释数字化特征的具体表现，本文从学术生命周期的角度出发，探讨数字技术与工具的介入对学术研究各个环节的影响。英国联合信息系统委员会（JISC）认为学术研究生命周期包括提出问题、组织团队、课题申请、研究探索、成果产出5个环节<sup>[20]</sup>，英国科学与技术设施研究理事会（STFC）提出数字环境下的科研模型包含分析研究趋势、形成思路、设计和组织项目、申请项目、进行实验、收集组织数据、分析数据、发布成果、保存成果9个阶段<sup>[21]</sup>。参照上述模型，本文将从研究问题、研究材料、研究方法、研究成果、学术交流5个方面阐述数字学术的特征。

### 2.1 研究问题

研究问题是整个学术研究工作的核心。互联网、大数据等数字技术的出现极大地拓展了学者的研究视野，也带来了各式各样的新兴研究问题。

传统学术研究中，学者可获得的研究数据有限，数据处理能力也有限。因此，多围绕具体、微观的研究问

题展开，通过对具体现象的研究达到以小见大的目的。

数字环境下，互联网与数据库的兴起减轻了学者获取数据与处理数据的负担，拓宽了学者的研究视野，推动了跨时空、跨学科研究问题的出现。从时间维度来看，不同时间、不同来源的多模态数据融合与关联使得全面深入理解事物的历史变迁成为可能。如Alison<sup>[22]</sup>从宏大的历时性角度出发，对一千多年的威尼斯历史档案进行数字化处理与分析，将威尼斯古城千年间的历史演变以动态的数字化形式展现出来。历时性研究盛行的同时，地理信息可视化技术的发展促进了基于海量数据的跨区域、跨文化探索，通过提取数据中的时间特征与空间特征进行历史、地理可视化的跨时空研究不断出现。如Schich等<sup>[23]</sup>绘制了公元前600年到2012年间15万人的出生和死亡地图，通过可视化的形式展示了近三千年欧洲和北美文化中心的形成与迁徙，以及与之相对应的帝国兴衰。此外，文学领域的地理信息可视化研究、考古学领域的GIS结合应用研究，以及Google地图与文学的融合展示等跨学科领域的交叉研究也愈发流行。

除了传统研究问题的不断扩展，数字技术的发展也带来了许多新的研究问题。如大数据、物联网与云计算等技术背景下，涉及海量数据的所有权、控制权、管辖权和使用权的“数据主权”<sup>[24]</sup>问题已经引起各国重视；公民在互联网、社交媒体访问中留下的持久性、全面性且涉及个人隐私的数据“被遗忘权”<sup>[25]</sup>问题引发了诸多讨论；由于网络的虚拟性、匿名性、法治与道德约束有限性等问题导致的“网络暴力”<sup>[26]</sup>问题引起了社会各界的反思；以技术为支撑，以数字为内核，以获取剩余价值、实现资本积累为目的，对生产力和生产关系产生重大影响的“数字资本主义”<sup>[27]</sup>问题已经成为政治经济学的研究热点。此外，“数字孪生”“人工智能伦理学”“数据捐赠”等诸多前所未有的新兴研究问题正随着数字技术的发展不断涌现，极大地扩展了学术研究范围。

### 2.2 研究材料

研究材料是构成学术研究的基本要素，是产生、表现主题的基础<sup>[28]</sup>。数字技术与工具的利用使得研究材料数字化成为数字学术的基本特征。

传统学术研究材料主要包括各类图书、期刊、报纸、档案等纸质文献。数字环境下，从文本的电子录入、图像与表格的扫描与识别到原生数字资源的形成，内容

丰富、形式多样的原生与非原生数字材料已经逐渐成为学术资源的主角。从内容看,数字学术资源涵盖著作、期刊、研究报告、专利、标准、学位论文、学术发现等学术成果数据,教材、工具书、参考书、课件讲稿、教学视频、学术讲座等学术教学资料,书目、文摘、索引等图书情报数据,以及学术会议安排、讲座信息、研究机构、研究人员、研究项目等学术资讯数据等。从媒介形式看,数字学术资源包含文本、图像、音频、视频、3D模型等。除了这些典型的数字学术文献外,规模巨大、类型丰富、变化速度快、价值密度低的大数据也是数字学术资源的重要组成部分,来自各种智能传感设备与计算机系统实时产生的传感数据,政府部分和其他组织公布的开放数据等也逐渐成为重要的数字学术资源。

传统学术研究中,研究者一般通过购买、借阅、租借等方式获取纸质研究材料。然而,纸质文献固有的复本不足、存放空间有限、查找困难、携带不便等问题为学术研究的顺利开展造成了一定的障碍。数字时代,研究材料的获取变得越来越便捷。随着数据库技术的发展,全文数据库、书目库、文摘库、索引库、事实库、事物库等多种数据库开始应用于学术活动中,为世界各地的研究人员提供全天候的资源服务,打破了研究资料获取在时间、空间、复本数量上的限制,极大地提高了资源的利用效率。云计算理念与技术的应用催生了Google Scholar、百度学术等一站式学术文献获取工具,这些工具在数据库的基础上进一步整合了来自全球范围内的图书馆、高校、研究机构存储的学术资源,简化了学者查找和获取研究资源的步骤,为学者提供包括资源检索、查询、获取与共享在内的一站式学术资源服务<sup>[29]</sup>。

## 2.3 研究方法

数字时代的到来产生了海量的数字学术研究材料,为研究人员进行理论拓展与创新、实践验证与推广提供了更广泛的研究空间,也促进了学术研究范式与研究工具的变革。

传统学术研究一般预先提出假设或构想,并利用文献调研、观察、实验等方法收集相关研究材料,通过对材料的分析、思辨与诠释来验证假设。然而,假设的提出往往高度依赖于研究人员的经验、直觉与判断,研究材料的获取与诠释也有较强的主观性与随意性,因此这种理论驱动研究范式在很大程度上限制了学术研

究的精度、深度与广度<sup>[6]</sup>。数字技术的出现为学术研究提供了新的思路与方法。文本挖掘、社会网络分析、复杂网络分析、认知计算、社会感知计算、机器学习、数据挖掘<sup>[30]</sup>等研究方法的利用催生了一种全新的学术研究范式——数据驱动研究范式,即不预先提出假设,而是通过对全面、完整的全样本数据进行分析挖掘来总结过去规律、描述现状、预测未来发展。相比理论驱动研究范式,数据驱动研究范式的研究材料更全面,材料分析更客观。两种研究范式的结合弥补了传统学术研究的不足,提高了学术研究的科学性,极大地扩展了学术研究的精度、深度与广度。

日益复杂的研究问题、类型多样的海量研究数据催生了多种功能丰富、高效易用的数字化研究工具。根据研究工具的载体形式,可以将其划分为硬件设备与软件工具。当前在学术领域应用较广泛的硬件设备主要包括高性能计算机、3D打印机、3D扫描仪、虚拟现实眼镜、全景相机、智能传感器等,这些设备的应用使得学术数据获取与计算变得更加便捷。学术研究软件工具的类型则更加丰富,EndNote、Mendeley、Zotero等文献收集、组织、管理与共享软件,TEI、GATE、DocuSKY等文本编码与标注工具,Content DM、Mirador、Digital Scanners等图像内容管理与分析工具,ABBYY Fine Reader、ATLAS.ti、Oxygen等文本挖掘工具,ArcGIS、GRASS、GoogleEarth、Story Maps等地理空间系统分析工具,TensorFlow、RapidMiner等机器学习工具,UCINET、Gephi、NVivo等数据可视化工具,Scalar、Omeka、Drupal等数字出版工具,Dspace等数据管理与发布工具的发明与使用为学术研究活动带来了极大的便利,使学术研究达到事半功倍的效果。

## 2.4 研究成果

数字环境下,研究人员利用丰富的数字化方法对数字研究材料进行研究与加工,形成的研究成果同样呈现出数字化特征。

传统学术研究成果一般以纸质的专业书籍、期刊论文等方式呈现,随着出版技术与工具的发展,数字学术出版物开始出现并逐渐普及,占据了学术出版的半壁江山。数字学术出版物从无到有、从单一到多元的发展过程,同样也是其载体形式不断变化与更新的过程。最早电子出版物载体是20世纪80年代出现的软磁盘;

随后,只读光盘成为电子出版物的主流载体<sup>[4]</sup>;90年代中期,无须物理存储载体、直接通过网络阅读的首批互联网期刊开始发行<sup>[5]</sup>;90年代后期,国内外开始出现专门的手持电子书阅读器。2014年,大多数学术出版已经转移到线上,超过90%的期刊都可以在线获取<sup>[5]</sup>。

除了载体形式的变化,研究成果的内容表现形式也在不断丰富。早期的数字出版物多是直接对纸质文献进行数字化处理得到的PDF文档。随后,集成了文字、图像、音频和视频等要素的多媒体文献,3D模型,动漫,软件,游戏,信息系统,可视化图表等多种形式逐步被应用于学术研究成果的内容呈现。进入21世纪,文本结构理论与本体技术在出版领域的应用催生了大批规范定义论文内容组件及其语义关系的学术论文内容本体,促进了学术研究成果的结构化表达。2009年,Shotton<sup>[31]</sup>提出了语义出版的概念。众多出版机构纷纷利用语义网、XML技术、自然语义处理、本体等多种先进数字技术开发了诸如增强型出版物、纳米出版物、微型出版物等语义丰富、关联关系明确的语义出版物模型。语义出版物的出现与发展不仅有利于提高读者的知识获取与理解效率,还能提高研究成果的机器可读性,促进知识单元的机器自动处理与重组再利用,实现更高级的知识服务<sup>[32]</sup>。

## 2.5 学术交流

从1665年世界上第一本科技期刊*Journal Des Scavans*诞生,传统学术交流模式已有三百多年的历史<sup>[33]</sup>。进入20世纪中期,数字技术与工具的介入极大地丰富了学术交流的渠道与方式,使其呈现出数字化的特点。

传统学术交流模式主要包括面对面对话、学术研讨会、学术演讲、学术沙龙、书信往来等直接交流方式和学术书籍、期刊出版等间接交流方式<sup>[34]</sup>。其中,直接交流受众规模有限、时间经济成本较高,间接交流也存在出版周期长、互动性不足、费用高等问题。数字技术的普及催生了电子邮件、即时通信工具、数字出版、网络资源发布和学术社交网络等多种基于互联网的学术交流方式。电子邮件、即时通信工具等直接交流方式通过网络将世界各地的人连接在一起,交流双(各)方不仅能即时免费传递文本、图片等信息,还能通过语音通话、视频通话达到面对面交流的效果。间接交流方式也各有所长,数字图书与期刊出版是当前学术界最重要的交流方式,内容丰富、获取便捷的数字出版物已经成

为学术研究的基础。学术网站、学术博客、学科信息门户、wiki资源和其他网络学术资源的发布是数字出版的有力补充,它最大程度地打破了学术交流的时空和篇幅限制,扩展了学术交流的受众规模和范围。学术论坛、网络新闻组、学术评论与社交平台等学术社交网络是互动性较强的学术交流方式。研究人员可以借助学术社交网络进行一对多、多对多的学术交流,达到取长补短、激发灵感、启迪智慧的作用。

传统学术期刊具有出版周期长、费用高等问题。数字技术与工具的介入带来了学术出版流程的优化和出版成本的缩减,然而,数字期刊与图书的价格却在不断飙升。针对这一情况,20世纪90年代末,学术界兴起了旨在促进学术文献免费获取,推动学术信息广泛交流的开放获取(Open Access, OA)运动。近年来,开放获取的理念席卷出版界、图书情报界和学术界,取得了丰硕的成果。截至2019年3月,开放获取期刊目录(Directory of Open Access Journals)共收录来自129个国家和地区的12 877种OA期刊<sup>[35]</sup>;在OpenDOAR注册的OA存储库已达4 062个<sup>[36]</sup>。开放获取运动突破了学术文献的价格与使用权限障碍,对促进学术交流、提高学术资源的公共利用率和研究效率意义重大。

## 3 数字学术研究项目案例分析

上文基于学术生命周期阐述了数字学术的数字化特征在学术研究各环节的具体表现。通过对当前主要的数字学术研究中心网站及相关文献进行调研,本文选取埃默里大学数字学术中心的“贝尔法斯特团队诗歌|网络”项目为例进行案例分析,对数字学术的研究过程与核心特征进行验证。

### 3.1 项目研究问题

1963年,时任贝尔法斯特女王大学英语讲师的Hobsbaum组织了一个由学生、教师和作家组成的写作研讨会,与会成员每周举办一次会议并在会上阅读、分享、讨论他们的作品(以诗歌为主)。贝尔法斯特研讨会持续了9年,期间偶尔中断,最终由于北爱尔兰政治动荡于1972年完全停止。期间有多个成员出版了他们的第一批作品,并在之后成为著名诗人。

有关贝尔法斯特研讨会的成员名单、成员交往程度、成员活动范围的认识存在很多分歧。为解决这一问

题,埃默里大学数字学术中心联合校内外多个研究机构开展了“贝尔法斯特团队诗歌网络”项目,对贝尔法斯特研讨会相关的历史文献进行数字化处理,并通过文本挖掘、语义标识等手段进行研讨会成员、诗人社交网络、诗人地理空间的数据挖掘与可视化展示。

### 3.2 项目研究材料

本项目的研究材料主要包括来自Stuart A. Rose手稿档案与珍本图书馆、贝尔法斯特女王大学图书馆、大英图书馆、美联社、画廊出版社、维克森林大学出版社、其他知识库馆藏以及个人收藏的贝尔法斯特团队历史文献信息,具体来看其囊括了研讨会中阅读的诗歌草稿,团队成员的简历、传记,会议记录,团队成员发表的诗歌,团队成员间的通信记录,有关贝尔法斯特团队及其成员的新闻报道和照片等多种类型的数字化材料。多来源、多类型的数字化研究材料互为补充、互为印证,使得全面深入地挖掘研讨会成员名单、诗人交往情况与地理位置等史实成为可能。研究材料的跨时空在线获取节省了大量的时间与精力,为研究的进一步开展奠定了坚实的基础。

### 3.3 项目研究方法

本项目并未预先作出假设,而是采用数据驱动研究范式,通过对研究材料的整理和组织,综合采用Python、Django、exist-db、d3.js、EAD/TEI、Twitter Bootstrap、DBpedia Spotlight、oxygen XML Editor、VIAF、GeoNames等多种开源数字技术与自行开发的TEI与EAD增强软件对研究材料中的人物、机构、地理位置等实体进行识别与标记,并通过建立实体的外部链接来增强数据的语义内涵,为用户提供交互式可视化体验。

具体来看,该项目主要从以下角度深入挖掘了研讨会历史信息。

(1) 贝尔法斯特团队成员名单。该模块通过对研讨会历史文献进行挖掘,提取出与研讨会直接联系的人员,通过网络图展示出来。网络图中的每个节点代表每位团队成员,节点之间的连线代表成员之间有直接联系。因此,该网络图还能显示团队成员之间的社交关系。本模块提供了两幅网络图,其中图1展示了参与研讨会的所有成员<sup>[37]</sup>,图2分阶段展示了1963—1966年及1966—1972年两个阶段的成员<sup>[37]</sup>(1966年Hobsbaum

离开贝尔法斯特女王大学,团队成员也不断更替)。

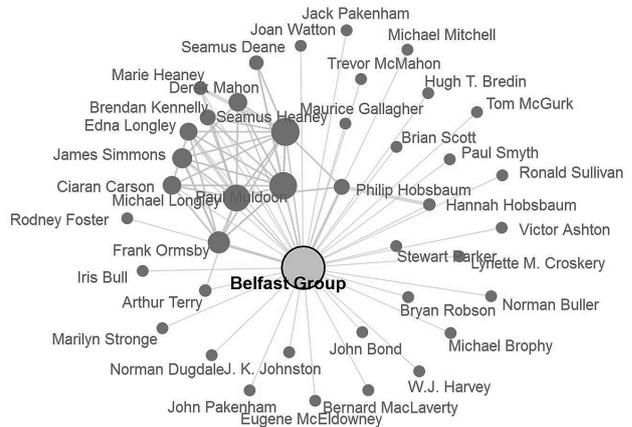


图1 贝尔法斯特团队成员网络图<sup>[37]</sup>

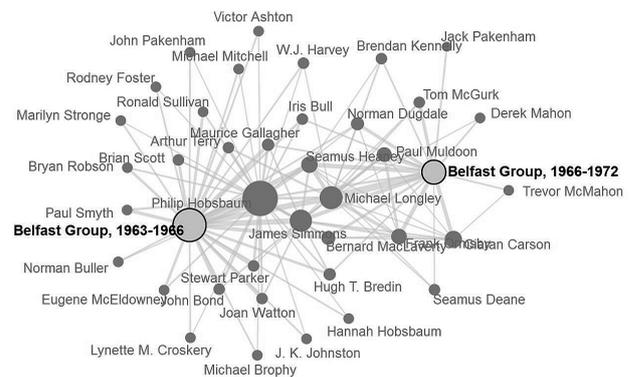


图2 贝尔法斯特1963—1966年及1966—1972年团队成员网络图<sup>[37]</sup>

(2) 贝尔法斯特团队成员之间的社交网络(见图3)<sup>[37]</sup>。该模块与第一模块使用了相同的数据,只是在侧重点与可视化方式上有所区别。该模块以余弦图的方式展示团队成员之间的社交关系(主要包括通信关系、婚姻关系、作品提及关系等)。其中,弦的厚度表示成员之间的交流频率,弦越粗说明成员之间的交流越密切;弦的颜色表示交流双方的主动程度。如表示Philip Hobsbaum和Hannah Hobsbaum社交关系的弦的颜色为紫色<sup>[37]</sup>(代表Philip),这是因为研究材料显示Philip写给Hannah的信件更多,或者Philip在作品中提及Hannah的次数更多。

(3) 贝尔法斯特团队成员地理位置(见图4)<sup>[37]</sup>。该模块主要通过对研讨会历史文献中的地理位置数据进行提取,并与具体的人物进行关联,以地图与图标的形式展示团队成员的地理位置。其中,文档图标表示诗人在诗歌作品中提到的地点,人像图标表示诗人居住

的地点，“人像+文档”图标则表示诗人居住过且在作品中提到过的地点。读者点击图标则可以跳转到相应的历史文献。

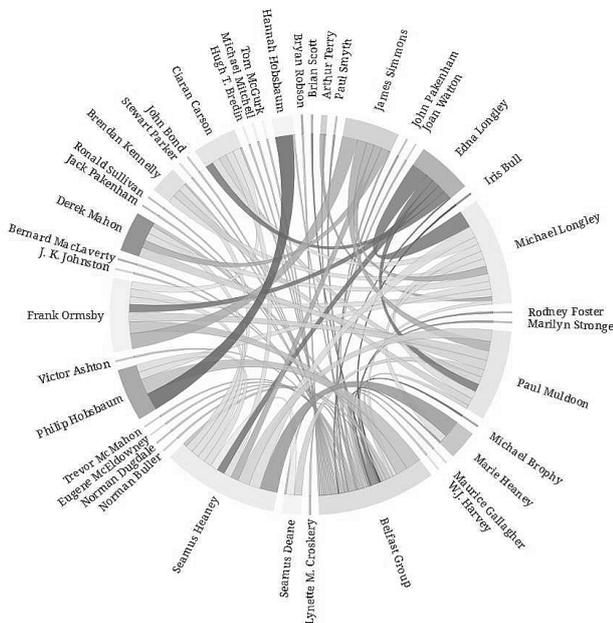


图3 贝尔法斯特团队成员社交网络<sup>[37]</sup>



图4 贝尔法斯特团队成员地理位置<sup>[37]</sup>

数据驱动研究范式的使用，研究材料的深入挖掘使得用传统理论驱动研究范式无法解决的问题得以解决。文献中人物、社交关系和地理位置等实体的识别、提取与语义增强为准确认定贝尔法斯特团队的成员名单与团队中诗人的社交网络、活动范围提供了保障，网络图、余弦图、地理信息图等多种形式的可视化图像更是直观清晰地展示了项目的研究结论。可见，数字化研究方法已经成为实现更准确、更科学、更深入学术研究的有力支撑。

### 3.4 项目研究成果

通过对相关历史文献的深入分析与挖掘，该项目取

得了多种数字化研究成果。如详细介绍项目研究过程，包含文本、图片和大量外部链接的项目介绍网页；介绍项目研究背景、研究细节的长篇随笔文章；展示研究结论的互动式可视化图像；介绍研究过程与研究结论的期刊论文；用于增强TEI和EAD的数字内容自动语义增强软件等。这些内容丰富、形式多样的研究成果多角度、全方位、深层次地展现了贝尔法斯特研究项目从立项到研究结论发布的全过程，有助于读者深入了解项目的研究情况，为类似研究提供理论、实践与技术支持。

### 3.5 项目学术交流方式

该项目采取直接交流与间接交流相结合的方式进行学术传播与交流。其中，直接交流方式主要包括研究团队进行的学术演讲，2013—2015年，该项目的负责人、技术顾问等分别在数字人文会议、美国档案工作者协会年会等多个学术会议中进行演讲。间接交流方式包括通过网页发布项目研究情况，通过可视化展示方案与读者进行互动，通过OA期刊发表期刊论文等。学术演讲有利于增强受众与研究者的互动，项目介绍网页能帮助来自全球各地的受众随时了解项目执行情况，研究结论的可视化表示提高了阅读的易读性和趣味性，开放获取论文的发表打破了学术论文的价格与使用权限障碍。总体来看，直接交流与间接交流相结合的多渠道交流模式极大地扩展了研究项目的受众群体，增强了研究成果的易获取程度，提高了该项目的交流效率。

值得注意的是，该项目的研究背景中提到“学术界对贝尔法斯特研讨会在培养北爱尔兰新一代诗人中的重要作用这一问题存在争论，希望通过该项目解决这一争论”。其中难度可想而知：参加研讨会一定能对诗人的创作有促进作用吗？这些促进作用体现在哪里？如何量化？即使利用最先进的技术也难以回答这些问题。在项目的研究过程与成果中均未提到这一争论，究竟是研究团队未进行相关研究，还是研究过但未能得出可靠结论，笔者无法得知。但是这一案例也充分说明数字学术并非无所不能。

## 4 总结

通过对数字学术的起源、发展和定义进行梳理，本文将数字学术定义为利用数字技术与方法介入学术研究全过程的新型研究模式。在此基础上，提出数

字学术的核心特征是数字化,并基于学术生命周期理论深入阐释了数字化特征在学术研究各环节的具体表现。通过对“贝尔法斯特团队诗歌|网络”项目的案例分析进一步验证了数字学术的数字化特征。

作为一个仍在发展与演变的新兴概念,数字学术通过扩展研究问题范围,提供易获取的海量研究材料,创新学术研究范式与高效研究工具,丰富学术成果的内容与形式,提高学术交流效率等方式为研究人员开展学术研究带来了诸多便利。但是,数字学术也存在很多局限。如数据驱动研究范式下的学术研究对模型与算法依赖性强,选择不同的模型与算法很可能得出截然不同甚至相反的结论;人类情感的复杂性和语言的歧义性带来的研究材料解读难题目前无法通过技术解决;研究人员多年积累的问题提出、分析与解决的经验无法使用数字技术替代;“远距离阅读”“文学模式识别”等文本量化分析仅能分析词频、语法等表层特征,无法对文学作品的情感与精神内涵进行深层次解读等。

总之,面对数字学术的热潮,学者既要以开放包容的态度接受新观念、新事物,也要冷静客观地看待其局限性,在学术研究中综合利用数字技术与传统方法,达到扬长避短的效果。

## 参考文献

- [1] 刘兹恒,涂志芳. 数字学术环境下学术图书馆发展新形态研究——以空间、资源和服务“三要素”为视角[J]. 图书情报工作, 2017, 61(16): 15-23.
- [2] HOCKEY S. The History of Humanities Computing [M] // SCHREIBMAN S, SIEMENS R, UNSWORTH J. A Companion to Digital Humanities. Oxford: Blackwell Publishing Ltd, 2004: 1-19.
- [3] 谢新洲. 数字出版技术 [M]. 北京: 北京大学出版社, 2002: 12.
- [4] 徐丽芳. 数字出版: 概念与形态 [J]. 出版发行研究, 2005(7): 5-12.
- [5] BERRY D M, FAGERJROD A. Digital Humanities: Knowledge and Critique in a Digital Age [M]. Cambridge: Polity Press, 2017: 114-115.
- [6] 欧阳剑. 大数据视域下人文数字人文研究 [J]. 图书馆杂志, 2018, 37(10): 61-69.
- [7] Wikipedia. Digital Scholarship [EB/OL]. [2019-01-18]. [http://en.Wikipedia.org/wiki/Digital\\_Scholarship](http://en.Wikipedia.org/wiki/Digital_Scholarship).
- [8] RUMSEY A S. New-Model Scholarly Communication: Road Map for Change [EB/OL]. [2019-01-22]. <http://uvasci.org/institutes-2003-2011/SCI-9-Road-Map-for-Change.pdf>.
- [9] LYNCH C. The “digital” scholarship disconnect [J]. Educause Review, 2014, 49(3): 10-15.
- [10] MULLIGAN R. Digital scholarship support in ARL member libraries: an overview [EB/OL]. [2019-01-20]. <http://www.arl.org/focus-areas/scholarly-communication/digital-scholarship/digital-scholarship-support/3978-digital-scholarship-support-in-arl-member-libraries-an-overview>.
- [11] ARL. SPEC Kit 350: Supporting Digital Scholarship (May 2016) [EB/OL]. [2019-02-02]. <https://publications.arl.org/Supporting-Digital-Scholarship-SPEC-Kit-350/9>.
- [12] 谢蓉,刘炜. 数字学术与公众科学: 数字图书馆新生态——第十三届数字图书馆前沿问题研讨班会议综述和思考 [J]. 大学图书馆学报, 2017, 35(1): 6-10.
- [13] 任树怀. 数字学术发展动向及大学图书馆的角色 [J]. 上海高校图书情报工作研究, 2017, 27(4): 5-6.
- [14] 李迪. 高校图书馆数字学术空间构建研究 [J]. 图书馆学刊, 2017, 39(11): 89-93.
- [15] 王贤. 美国伊利诺伊大学香槟分校图书馆数字学术服务及启示 [J]. 图书情报工作, 2018, 62(11): 143-150.
- [16] 刘炜,叶鹰. 数字人文的技术体系与理论结构探讨 [J]. 中国图书馆学报, 2017, 43(5): 32-41.
- [17] 王飞跃. 社会计算的定义与应用 [C] // 中国科学技术协会学术部. 新观点新学说学术沙龙文集20: 社会能计算吗. 北京: 中国科学技术出版社, 2009: 6-18.
- [18] e-Science [EB/OL]. [2019-02-20]. <https://en.wikipedia.org/wiki/E-Science>.
- [19] 孙坦. 数字化科研——e-Science研究 [M]. 北京: 电子工业出版社, 2009: 396-397.
- [20] Jisc. How Jisc is helping researchers [EB/OL]. [2019-02-19]. <http://www.jisc.ac.uk/whatwedo/campaigns/res3/jischelp.aspx>.
- [21] 张晓林. 研究图书馆2020: 嵌入式协作化知识实验室? [J]. 中国图书馆学报, 2012, 38(1): 11-20.
- [22] ALISON A. The ‘time machine’ reconstructing ancient Venice’s social networks [J]. Nature, 2017, 546(7658): 341-344.
- [23] SCHICH M, SONG C, AHNY Y, et al. A network framework of cultural history [J]. Science, 2014, 345(6196): 558-562.
- [24] 杜雁芸. 大数据时代国家数据主权问题研究 [J]. 国际观察, 2016(3): 1-14.
- [25] 吴飞,傅正科. 大数据与“被遗忘权” [J]. 浙江大学学报(人文社会科学版), 2015, 45(2): 68-78.

- [26] 百度百科. 网络暴力(社会现象) [EB/OL]. [2019-02-20].  
<https://baike.baidu.com/item/%E7%BD%91%E7%BB%9C%E6%9A%B4%E5%8A%9B/4707152?fr=aladdin>.
- [27] 宋平. 马克思主义视域下的数字资本主义 [J]. 佳木斯大学社会科学学报, 2019, 37 (1): 16-20.
- [28] 徐融. 毕业论文写作(文科类) [M]. 北京: 中国商业出版社, 2002: 37-38.
- [29] 王秀秀, 武和平. 基于“云计算”的数字学术资源整合策略与服务模式研究 [J]. 电化教育研究, 2012, 33 (6): 72-74, 93.
- [30] 马费成, 张瑞, 李志元. 大数据对情报学研究的影响 [J]. 图书情报知识, 2018 (5): 4-9.
- [31] SHOTTON D. Semantic publishing: the coming revolution in scientific journal publishing [J]. Learned Publishing, 2009, 22 (2): 85-94.
- [32] 王晓光, 宋宁远. 语义出版物的内容组织架构研究——基于纳米出版物和微型出版物的比较分析 [J]. 出版科学, 2017, 25 (4): 20-27.
- [33] 孙华. 数字学术交流发展研究初探 [J]. 农业网络信息, 2015 (12): 112-114.
- [34] 左梁. 信息化情况下学术交流形式的创新 [C] // 吉林省科学技术协会. 第十一届中国科协年会论文集. 长春: 吉林省科协, 2009: 14-17.
- [35] DOAJ. DOAJ (Directory of Open Access Journals) [EB/OL]. [2019-03-22]. <https://doaj.org/>.
- [36] OpenDOAR. OpenDOAR Statistics [EB/OL]. [2019-03-22]. [http://v2.sherpa.ac.uk/view/repository\\_visualisations/1.html](http://v2.sherpa.ac.uk/view/repository_visualisations/1.html).
- [37] Belfast Group Poetry|Networks. Networks & Maps [EB/OL]. [2019-03-22]. <https://belfastgroup.ecds.emory.edu/network/>.

## 作者简介

曾熙, 女, 1995年生, 硕士研究生, 研究方向: 数字人文、知识服务, E-mail: zengxi\_321@163.com。

王晓光, 男, 1978年生, 教授, 博士生导师, 研究方向: 语义出版、知识组织、数字资产管理、数字人文, E-mail: whu\_wxg@126.com。

Digital Scholarship: Concept, Characteristics and Case Analysis

ZENG Xi<sup>1</sup> WANG XiaoGuang<sup>2</sup>

(1. School of Information Management, Wuhan University, Wuhan 430072, China; 2. Center for Studies of Information Resources, Wuhan University, Wuhan 430072, China)

Abstract: The application of digital technology in academic research has brought profound changes in the academic research environment and research process, which has given birth to the concept of “digital scholarship”. Through literature survey, this paper sorts out the historical evolution and basic concepts of digital scholarship, and analyzes the digitalization characteristics in each link of academic research based on academic life cycle theory. Through the network research, the “Belfast Group Poetry | Networks” project is taken as an example to verify the current digital academic research process and characteristics. Finally, the digital scholarship reflection and summary proposed digital scholarship is not omnipotent. Researchers should treat the digital scholarship craze in an open, inclusive, objective and calm manner.

Keywords: Digital Scholarship; Digitization; Digital Technologies; Data-Driven Research Paradigm; Academic Exchange

(收稿日期: 2019-03-01)