

数字人文与开放议程

——西蒙·马奥尼教授访谈

西蒙·马奥尼 陈兴怡 王普宇

摘 要 尽管国际上对开放议程的研究与实践已较为成熟,但在中国,相关研究仍较为匮乏。而作为理念与实践均与之密切相关的数字人文领域,尤其需要提升对开放议程的认识。文章是对伦敦大学学院数字人文荣休教授、数字人文领域开放议程倡导者西蒙·马奥尼(Simon Mahony)教授的访谈整理,集中探讨了开放议程的内容与价值,及数字人文与开放议程的关系。马奥尼教授指出,开放议程的核心在于促进知识的民主化与共享,增强学术研究的透明性与包容性,并鼓励公众参与知识生产;作为开放议程的积极推动者,数字人文学科在全球学术生态中的作用尤为突出,但是不同地区、国家之间的开放议程进度和内容有较大差异,比如英国与中国。马奥尼教授对开放议程的未来发展持乐观态度,认为其将继续推动知识创造与传播的加速,并进一步提升社会的学术参与度。

关 键 词 数字人文; 开放议程; 开放科学

分 类 号 G203; G31

作者简介 西蒙·马奥尼(Simon Mahony),伦敦大学学院信息研究系数字人文荣休教授,北京大学信息管理系客座教授; 陈兴怡(通讯作者),丝路视觉科技股份有限公司北京分公司策展人,Email: cxyamalia@gmail.com; 王普宇,牛津大学工程科学系博士研究生。

0 引言

2021年11月23日,联合国教科文组织发布了《教科文组织开放科学建议书》^[1],将开放科学定义为运动和实践相结合的一体化包容性架构,旨在实现科学知识的公开获取、使用以及复用,通过加强科学合作和信息共享促进科学和社会发展,并向以往未包含在科学界里的社会其他人群分享科学知识的生产、评估和传播进程。数字人文自诞生以来一直是开放精神的坚定倡导者,与开放科学共享着合作、包容、分享、公开、透明等理念,是人文学术中最与之相契合的领域,但这二者,都被置于“开放议程”(Open Agenda)的光

谱中。“开放议程”概念源自多个运动的汇合,这些运动旨在使知识民主化,使学术研究更具包容性和透明度,包括开放获取^[2]、开放科学^①、参与式文化^[3]和数字人文^[4],等等,所有这些运动都提倡共享知识和公众参与研究过程。

在数字人文领域,伦敦大学学院(University College London, UCL)数字人文荣休教授西蒙·马奥尼(Simon Mahony)是“开放议程”的热烈呼吁者和推动者。马奥尼教授不仅在 UCL 数字人文学科的建设和发展中发挥了关键作用,还曾在北京师范大学(珠海)和北京大学等多所高校的研究机构中担任要职,并且是 iSchools 联盟数字人文课程设计委员会委员及多项国际学术机构的成员,在推动数字人文领域的开放与创新方面,拥有独特的见解和丰富的实践经验。在中国工作期间,他曾在 2023 年第五届中国数字人文年会上发表《人文学科开放数据》主题演讲,在“2022 北京论坛”发表《开放议程:开放获取出版》主题演讲,以及在中国人民大学、宁波诺丁汉大学、浙江大学等高校举办的研讨会上发表《开放议程:开放获取、开放出版和开放科学》主题分享,提升了这一概念和理念在中国的认知度。

对于开放议程,国外学术界已形成较为系统的研究框架和理论体系^[5-10],而国内对此议题的研究尚处于初步阶段,相关学术文献较为缺乏,我们希望通过马奥尼教授的访谈,深入展示数字人文领域开放议程的基本议题和发展现状,以此推动在中国的相关讨论。

受访者简介

西蒙·马奥尼(Simon Mahony),伦敦大学学院数字人文荣休教授、北京师范大学(珠海)数字出版与数字人文研究中心教授、北京大学信息管理系客座教授。研究方向为数字人文学科,特别关注教育、传播、信息研究、数字叙事、平等、多样性和包容性(EDI)以及开放议程。代表作有《开放议程:为什么需要开放和透明》(2022),《走向开放和透明,更好地促进知识创造》(2022),《语言和数字人文学科的研究主题》(2020),《数字人文学科中的语言和文化霸权》(2019)等。

1 开放议程与开放科学

Q1:“开放议程”的定义和内容是什么,它与“开放科学”是什么关系?

对我来说,开放议程意味着一切都是开放的,尽可能开放,但必要时封闭,因为有些内容可能是敏感或机密的,并受数据保护法的约束。开放议程包括开放教育、开放获取、开放数据和开放科学,可能还包括开放同行评议和开放源代码。这些都是朝着开放和透明迈进的同一运动的一部分。

在我看来,开放科学只是意味着开放学术(Open Scholarship)。我有古典学教育背景,并扎根于人文学科。

① https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en

从拉丁语“scientia”中可提取“science”一词,意为知识,因此与“scholarship”一词类似,在最广泛的意义上包括了所有人文、社会科学和其他学术领域。欧洲“开放科学云”(European Open Science Cloud)^[11]就是一个例子,它的建立是为了不做任何区分地储存欧洲所有研究数据。

我会避免去定义某事物,支持包容性,而不是严格划分特定学术领域,因为这可能会导致学科间的排斥。当你说某个东西“是”什么时,你也在说它不是什么。通过这种包容的方式,可以促进我们与多学科的协同工作。简而言之,开放科学是开放议程的一部分。

Q2:您是从什么时候开始关注“开放议程”这个议题的? 和您的数字人文研究有关吗?

在伦敦大学学院工作之前,我在伦敦国王学院(KCL)的人文计算中心(现为数字人文系)工作。在广泛使用虚拟学习环境(VLE)前,人文计算中心有独立于KCL机构基础设施的服务器用于项目托管,其中部分服务器的空间专用于储存教学材料。我当时正在教授本科和研究生课程,所有的教学材料也都托管在那。这些材料也都曾存放于开放网络,无需登录——任何知道网址的人都可以访问,并且在服务器更新之前这些材料也一直可用。因为是我自己的知识成果,使用的任何图像也都是我的或公共领域的,所有没有任何版权问题。同时,我还负责了一个线上硕士课程的运营,非常熟悉版权和知识产权问题。这就是我学会开发公开在线托管教学材料的经历,其中也涉及了知识产权类的问题。

但当我转到伦敦大学学院时,我发现这种开放性并不常见,学生需要通过机构登录才能访问我的教学材料。几年后,伦敦大学学院、伦敦国王学院和许多其他大学开始使用 Moodle(<https://moodle.org/>)作为课程管理工具和教学材料托管平台。2011年,一位同事介绍并招募我参与英国联合信息系统委员会(Joint Information Systems Committee, JISC)资助的开放教育资源创建计划。于是我参与了“数字人文开放教育资源”项目(Open Educational Resources for Digital Humanities Project)的开发。这引发了我对开放教育的研究兴趣,几年内我和同事一起举办了研讨会和会议演讲,这让我更广泛地了解了开放议程。

对我来说,开放议程始于教育,即将我的教学材料公开提供给任何希望访问它的人。我会尽我可能根据知识共享许可协议(Creative Commons Licence)免费提供我所有的知识成果。

Q3:您个人在教学、研究过程中做了哪些尝试去推进开放议程,比如您在UCL,这是一个怎样的过程?

我提倡开放议程,并在教学、客座演讲和主题演讲中强调它的重要性。我还专门就这一主题发表了文章^[12,13],并在其他相关主题的文章中介绍了开放议程的原则。同时我也是完全开放获取的《开放人文数据期刊》(*Journal of Open Humanities Data*)的编辑委员会成员^①以及穆尔西亚大学(University of Murcia)历史科学和

① <https://openhumanitiesdata.metajnl.com/>

教学数字杂志《万物流转》(*Panta Rei*)^①——也是完全开放获取——的顾问委员会成员。

走向开放需要学术机构内部的文化变革,我很高兴伦敦大学学院在这方面处于领先地位。例如,请看,英国第一家完全开放获取的大学出版社就是伦敦大学学院出版社(UCL Press)^②,事实上,伦敦大学学院图书馆开放获取页面的网址^③中包含了“开放科学”一词。

2 开放议程与数字人文

Q4:是否存在一个特定的数字人文开放议程?还是说“开放议程”是普遍的,不会因研究领域而异?

如上所述,在我看来,数字人文学科是开放议程的倡导者。在英国,我们支持开放获取与开放出版,鼓励所有学者公开发表文章,并且只有以开放获取形式发表的文章才会被纳入“研究卓越框架”(Research Excellence Framework, REF)系统,这是一个用于评估英国大学的研究质量的系统。开放议程的各个方面及其参与方式可能因研究领域不同而有所差异,但从根本上讲,它关乎开放性和透明度。

Q5:开放议程对于数字人文领域是否具有特别重要性?反过来,数字人文又能为开放科学做些什么?

反过来说,似乎——在西方,我不能代表其他地区发言——数字人文学科已成为开放议程的坚定倡导者。我们大多数人都试图让一切都开放(例如大量开放获取的数字人文学科出版物),但不幸的是,这并不总是可行的。

例如,开放获取出版需要资金来支付文章处理费(Article Processing Charge, APC)。这对没有资金的初级学者以及资金较少的机构学者不利。我在北京师范大学工作时,我用我的研究基金支付APC,但我目前属于伦敦大学学院,担任荣休教授(退休状态),所以我无法获得这方面的资金。所以这也意味着我与北京师范大学合同中剩余的、目前正在审查的文章将不会公开发表。因为我与北京师范大学的合同规定,它们必须发表在SSCI/SCI/A&HCI期刊上,而这些期刊都需要APC才能公开发表。

我希望将来发表的所有文章都提交给开放获取期刊和没有APC的期刊,无论它们在科睿唯安科学网(Clarivate Web of Science)列表中的状态如何。对我来说,开放获取出版比SSCI等更重要。另请参阅我的文章《走向开放和透明以更好地促进知识创造》《开放议程:为什么需要开放和透明》^[12,13]。公开发表的文章加速了研究成果的传播,从而加速了知识的生产。更多的人会阅读和引用这些公开发表的作品,这对引用指标也有好处。

① <https://www.um.es/cepoat/pantarei/>
② <https://www.uclpress.co.uk/>
③ <https://www.ucl.ac.uk/library/open-science-research-support/open-access>

Q6:能否介绍一下您所观察到的不同地区的数字人文开放议程发展状况,比如特征和面临的困难,以及背后的原因。

任何新事物总是会带来问题和反对意见。例如,对于开放教育资源,我们的研究表明,版权是一个主要问题,因为教师往往在分享材料时并不一定知道这些材料的来源。人们对开放获取出版和开放数据也存在类似的反对意见——我在我的文章中提到了这一点^[12],但在西方,这两者似乎都已成为标准做法,因为人们可以看到它们带来的好处——更多的人阅读你的出版物,数据可以重复使用。《开放人文数据杂志》隶属于 Ubiquity Press^①,我们发表研究论文和简短的数据论文,以补充完整论文并推广开放数据集,以期使它们可供重复使用。UCL Press 也是所有内容均可免费下载为 PDF。

英国大学要求教职员工必须公开发表论文,只有开放获取的论文才会被纳入评估,并计入员工晋升。主要资助者,比如英国研究与创新署(UK Research and Innovation, UKRI)、惠康基金会(Wellcome Trust)^②等也都要求他们资助的项目研究数据必须公开。

中国的情况虽略有不同,但中国已采取了一些举措强调开放获取出版和开放数据的重要性。联合国教科文组织有一份关于中华人民共和国开放教育资源的出版物^③,以及一个全球开放获取门户网站^④。中国国家自然科学基金会也曾发布《关于实施资助项目研究论文开放获取的政策声明》^⑤,其中提到公共资助研究的成果属于全社会,因此应免费公开。同样,中国科学院也有一项计划,旨在大规模实现期刊论文的开放获取。北京大学和清华大学等几所中国顶尖大学正在开发开放数据存储库。知识共享署名 4.0 许可证现已提供简体中文和繁体中文版本,以解决与西方地区和文化差异^[14]。

新的举措总是会遭遇反对,但现在人们越来越意识到开放获取出版物的价值以及如何开放其所依据的数据。

Q7:学界之外,存在很多从事数字人文相关工作的商业公司,也还有很多非专业的爱好者在做各种各样的数字人文项目,他们会对开放议程产生怎样的影响?在整个数字人文领域中又扮演怎样的角色?

商业公司在数字人文研究中发挥着重要作用——我们需要合作,特别是在成像、沉浸式虚拟现实和元宇宙等领域。这些领域属于商业开发,大学工作人员不一定具备专业知识或资源。比如 UCL 数字人文的成像实验室就有由商业公司提供的设备,专门用于手稿和珍贵文件的成像,这些公司也会提供相应培训。我还与上海图书馆的数字人文小组合作过,他们也与很多商业公司进行合作,比如上海图书馆新建东馆内的设备和系统。同样,大型基础设施计划通常也依赖商业平台进行开发。

① 一家开放获取出版商,网站:<https://ubiquitypress.com>。

② 维康基金会是英国最大的慈善基金会之一。

③ <https://unesdoc.unesco.org/ark:/48223/pf0000215547>

④ <https://www.goap.info/>

⑤ https://www.eifl.net/sites/default/files/resources/201409/140515-nsfc_oa_policy_english.pdf

非专业爱好者,我们称之为“公众科学”或“公民科学”(citizen-science),也为研究计划做出了巨大贡献。想想许多依赖于志愿工作的众包项目,例如伦敦大学学院的“转录边沁”项目(Transcribe Bentham)^①和 Zooniverse^②。文化遗产领域的许多工作也由积极参与的社会公众承担。^③

商业应用和公众科学在开放议程中都存在无数可能性。公众科学平台本质上是开放的,任何人都可以访问。商业产品也允许公开发布研究材料,前提是这项工作由大学或博物馆工作人员进行。

Q8:现在越来越多的人文学科学者在尝试使用 ChatGPT 等生成式人工智能来辅助研究,包括建立数据库、数据标注、数据分析等基本工作,你认为它会取代目前数字人文领域的常用工具吗?会成为以后数字人文研究的发展趋势吗,是否可以认为是数字人文内部将发生的一种变革?这对开放议程又会带来怎样的影响?

这是一个很难回答的问题,只有时间才能告诉我们它会产生什么影响。我不知道是否有学者正在使用生成式人工智能来做数据库、数据标注、数据分析等基本工作。生成式人工智能依赖于它可以访问的数据,这些数据通常来自网络或训练数据集。我曾经出于好奇使用过 ChatGPT 看看它能做什么,但是输出的答案比较笼统,缺乏一定的深度,所以我没有用它进行其他更深入的研究。

在我看来,人工智能应该成为信息素养教学的一部分。比起展示它能做什么,更重要的是关注它不能做什么。它可以作为教育者工具包的一部分,而不是研究人员工具包的一部分。因为对于所有生成式人工智能(Generative AI)、大型语言模型(Large Language Models, LLMs)以及更广泛的人工智能而言,一个重要的问题就是信任(或缺乏信任)。欧盟委员会的白皮书(2020 年)明确指出了这一点:关于人工智能——欧洲追求卓越和信任的方法^④。及欧洲议会的《人工智能法案(2024 年)》^⑤,以及英国科学创新与技术部发布的《人工智能监管:一种有利于创新的方法》^⑥。在中国政府,《生成式人工智能服务管理暂行办法》于 2023 年 8 月生效,几乎完全侧重于生成式人工智能^⑦,预计人工智能的国家和行业标准将于 2026 年出台^⑧。

Q9:您在很多场合都谈开放议程的问题,那么您对数字人文开放议程的发展是持乐观态度还是悲观态度?影响这一进程的当下和未来因素有哪些?

如上所述,现在我们看到并理解了开放议程的好处。总体而言,我很乐观,但这一概念需要学术界和出版

① <https://www.ucl.ac.uk/bentham-project/transcribe-bentham>
② <https://www.zooniverse.org/>
③ <https://www.theguardian.com/science/article/2024/may/27/hobbyist-archaeologists-identify-thousands-of-ancient-sites-in-england-deep-time-digventures-national-trust>
④ https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en
⑤ [https://www.europarl.europa.eu/RegData/seance_pleniere/textes_adoptes/definitif/2024/03-13/0138/P9_TA\(2024\)0138_EN.pdf](https://www.europarl.europa.eu/RegData/seance_pleniere/textes_adoptes/definitif/2024/03-13/0138/P9_TA(2024)0138_EN.pdf)
⑥ https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1176103/a-pro-innovation-approach-to-ai-regulation-amended-web-ready.pdf
⑦ https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm
⑧ https://english.www.gov.cn/news/202407/02/content_WS668403b6c6d0868f4e8e8c85.html

商的接受和支持。以前的模式是学者需要为出版付费,然后大学图书馆必须支付订阅费才能访问这些出版物,但这种模式已经在逐渐失效。同时,我们需要清楚,我们这些在大学或与学术机构有联系的人处于特权地位。开放获取和开放数据不同于普遍获取。在访问方面,富人和穷人之间仍然存在数字鸿沟,比如许多地区的互联网接入很少或成本高昂。还有许多其他问题阻碍了普遍获取,例如语言问题、在线发布论文或数据的可访问性、连通性——并不是所有地方都像我们一样享有数据的连通性,以及内容过滤和阻止的问题。

此外,我们生活在一个需要保护网络免受企业利益侵害的时代。那些为了个人目的而试图分裂我们而不是团结我们的人;那些人的目的是操纵舆论,让我们怀疑自己对某种情况的了解和理解,从而造成混乱。^[15]这些都是我们在互联网社会中面临的危险。教育的开放和透明是(避免这些危险的)关键。当我们撰写文章或生产数据时,我们必须清楚了解我们用来支持任何观点的证据来源。因此我们需要完整和透明的文档,以便我们对信息和数据充满信心。我们也必须准备好质疑我们所提供的信息的真实性和可靠性,寻找潜在偏见、利益冲突和平衡观点的证据^[15]。

数字人文学科为我们提供了做到这一点的关键工具,数字工具和数字素养应该成为每门学术课程的核心部分。

参考文献

- [1] UNESCO. 教科文组织开放科学建议书[R/OL]. [2024-09-09]. https://unesdoc.unesco.org/ark:/48223/pf0000379949_chi.
- [2] Budapest Open Access Initiative (2002) [J/OL]. Journal of Library, Archives and Information Science, 2012(2). <https://doi.org/10.4403/jlis.it-8629>. DOI:10.4403/jlis.it-8629.
- [3] JENKINS H. Convergence culture: where old and new media collide [M]. New York: New York University Press, 2006: 20.
- [4] GOLD M K. Debates in the digital humanities [M]. Minneapolis: University Of Minnesota Press, 2012.
- [5] ELSE H. Radical open-access plan could spell end to journal subscriptions [J/OL]. Nature, 2018, 561(7721): 17-18.
- [6] LAAKSO M, MATTHIAS L, JAHN N. Open is not forever: a study of vanished open access journals [J/OL]. Journal of the Association for Information Science and Technology, 2021, 72(9): 1099-1112.
- [7] ROWLEY J, JOHNSON F, SBAFFI L, et al. Academics' behaviors and attitudes towards open access publishing in scholarly journals [J/OL]. Journal of the Association for Information Science and Technology, 2017, 68(5): 1201-1211.
- [8] SPEZI V, WAKELING S, PINFIELD S, et al. Open-access mega-journals: the future of scholarly communication or academic dumping ground? A review [J/OL]. Journal of Documentation, 2017, 73(2): 263-283.
- [9] SUBER P. Open access [M]. Cambridge: The MIT Press, 2019.
- [10] WAKELING S, CREASER C, PINFIELD S, et al. Motivations, understandings, and experiences of open - access mega - journal authors: results of a large - scale survey [J/OL]. Journal of the Association for Information Science and Technology, 2019, 70(7): 54-768.
- [11] European Open Science Cloud (EOSC) | Shaping Europe's digital future [EB/OL]. (2023-12-14) [2024-09-28]. <https://digital->

strategy.ec.europa.eu/en/policies/open-science-cloud.

[12] MAHONY S. Toward openness and transparency to better facilitate knowledge creation[J/OL]. Journal of the Association for Information Science and Technology, 2022, 73(10):1474-1488.

[13] MAHONY S. The open agenda: why openness and transparency are needed[J/OL]. Information Matters, 2022, 2(5). <https://informationmatters.org/2022/05/the-open-agenda-why-openness-and-transparency-are-needed/>.

[14] PETERS D. Our 4.0 license suite is now available in simplified and traditional Chinese[EB/OL]. (2020-01-02) [2024-09-29]. <https://3.130.221.114/2020/01/02/cc-licenses-in-simplified-traditional-chinese/>.

[15] Mahony S, CHEN Q. Concerns about the role of artificial intelligence in journalism, and media manipulation[J/OL]. Journalism, 2024, 0(0). <https://doi.org/10.1177/14648849241263293>.

Digital Humanities and Open Agenda

Simon Mahony Chen Xingyi Wang Puyu

Abstract Digital humanities, as an active promoting force of the Open Agenda, shares the same core values with Open Science: freedom, collaboration, inclusion, sharing, openness, and transparency. Digital technologies not only provide the technical foundation for realizing freedom, openness, and sharing, but they also transformed human thinking and the paradigms of scientific research. The widespread application of emerging technologies, such as artificial intelligence, has exerted a significant impact on the disciplinary literacy, research paradigms, and trends of digital humanities. Simon Mahony, Emeritus Professor of Digital Humanities at University College London (UCL), has been active in the field for over 20 years. As a contributor, researcher, and educator in digital humanities, he analyzes the development of this field from a comprehensive perspective. Through three in-depth interviews with Professor Mahony, this article discusses the spirit of openness in digital humanities, starting from the Open Agenda and Open Science, with a particular focus on the development and future of digital humanities in the era of artificial intelligence. By conducting a comparative analysis of domestic and international discussions surrounding the Open Agenda, this article aims to uncover the various focal points and its potential development in China, providing new perspectives for digital humanities' spirit of openness and exploring its expansive prospects in the new era.

Key words digital humanities; open agenda; open science