

“是什么”与“怎么办”

——“定义数字人文”研讨会综述

陈 静

摘 要 为了回答“数字人文是什么”以及“如何定义数字人文”，有必要对数字人文的核心内涵以及可能存在的边界进行厘清，尤其是从中国语境出发提出反思路径。来自爱丁堡大学、英国伦敦大学学院、南京大学、浙江大学、上海图书馆、山东大学、北京大学、上海大学、清华大学、中国社会科学院等机构的 20 多位学者于 2019 年 11 月 8 日齐聚南京大学，围绕“数字人文是否需要被定义”“数字人文与组织机构联盟”“数字人文与教学”“数字人文与学术期刊及评价机制”“数字人文与文学、史学、哲学等的关系”“数字人文与大数据、人工智能间的关系”“数字人文与计量研究法：相同或相异”等议题展开了激烈讨论，并围绕“为什么需要定义数字人文”“如何定义数字人文”以及“数字人文的拓展可能”形成了富有建设性的共识。

关键词 数字人文；数字人文教学；学术评价；研讨综述

分类号 G203

作者简介 陈静，南京大学艺术学院副教授，南京大学人文社会科学高级研究院数字人文创研中心成员，Email: cjchen@nju.edu.cn。

0 背景

会议伊始，来自爱丁堡大学的梅丽莎·特拉斯(Melissa Terras)和爱德华(Edward Vanhoutte)围绕《数字人文读本》(*Defining Digital Humanities*)一书的撰写及编辑过程，讨论了英美语境中“数字人文”的定义，并提出了他们的看法。梅丽莎认为，关于数字人文定义的争论在西方持续了多年，然而时至今日仍没有明确答案，甚至这个概念因为其复杂性而无法被定义了。爱德华称，在《数字人文学术》(*Digital Scholarship in the Humanities*)多年的编辑工作中，他看到了大量学科合作的案例，这是数字人文的独特之处，也是数字人文作为新的研究领域的潜力所在。他们以往对中国的数字人文研究情况所知不多，因此很希望能够了解中国学者是如何看待数字人文，以及中国数字人文的特殊性何在。

会议召集人之一、南京大学陈静介绍了会议主题的背景。从定义的角度而言，“数字人文”作为概念与作为对象在内涵与外延上是不一致的。按照某些学者的回顾，英文中的“Digital Humanities”概念始于 2001 年的某次小范围讨论；作为一个研究领域，其历史则被追溯至 1940 年代及“人文计算”。在中国，情况则更为复杂：一方面，作为由西方语境打造的词语，“Digital Humanities”被译介至华语世界的时候，因为时空的差异，译法不尽相同；另一方面，从学科溯源的角度来看，中文语境中虽然没有类似“人文计算”的前期概念或如布萨神父般的“传奇”人物，却有着与英美数字人文同型异构的历史脉络，因此，梳理中国数字人文的历史并对其定义也就

更有难度。然而,历史与定义作为独立性与合法性依据,无论是在数字人文原产地欧美,还是在后发地中国,都是无法回避甚至是迫切需要回答的问题。时至今日,学界尚未就此达成清晰的共识,原因除去以上谈到的历史与移植因素外,还有数字人文自身的复杂性和“定义”这个行为本身所具有的风险性和限定性。“数字人文”在字面上具有的宏大叙事性,往往容易使人忽略其跨学科性、实践性所带来的多样性和复杂性,比如数字人文与“数字图书馆”“数字学术”“电子文学”,甚至“数字艺术”“数字遗产”等人文各领域中已有概念之间的关系。与此同时,定义往往需要确保一定的普适性,会排除无法满足条件的个别、特殊案例,因而具有排他性,这对新生事物可能会造成一定的消极影响,因此,“如何定义数字人文”成为重要议题。

1 为什么需要定义数字人文

讨论开始后,山东大学邱伟云明确指出,数字人文应该被定义。他认为,数字人文领域如果与人工智能、大数据、计量研究等有关,那就有必要解释为什么是用“数字人文”概念而不是“智能人文”概念,以及数字与人文的结合究竟具有什么样的意义;此外,在定义缺失的情况下,是很难对在这个领域内从事学术研究的群体进行描述的,也就是说,无法界定“数字人文社群”的意义和价值。南京大学王涛则从个人经验出发,提出定义的内涵因所涉及的问题和语境而不同,因此讨论具体内容(在目前来看)意义并不大。南京大学殷曼婷认为,定义的必要性还体现在其所揭示的问题上:当下关于数字人文的讨论,很大程度上还是在一个传统人文学的框架内,但实际上,人们都能体会到数字技术所带来的冲击,以及给人文学科带来的危机感,而传统的人文艺术更关注的是人以及人的主体性问题,因此数字人文是否更应该关注的是数字技术所带来的冲击下人的可能性?是否更应该关注身处其中的人的状态与心理?浙江大学姜文则提出了不同看法,他认为数字人文虽然是在数字时代才出现的概念与领域,但从人文传统来看,实际上是人文学从19世纪发展到现在的选择,这种选择出现的基础是基于印刷技术而形成的印刷人文学、高等教育系统以及人文研究,因此,对于数字人文的研究也应该回到其人文学的语境中去探究,去看数字人文的史前史,既要考虑到人文传统的内在发展脉络,同时还要兼顾方法论上的创新。南京大学裴雷则从基础设施建设的角度提出了新的看法,即数字人文是一种基础性的研究,而且不应该“数字”与“人文”必须集成于一身的:数字(Digital)部分,是基础公共设施;人文(Humanities)部分,则是人文学者具体实践的内容。从这个意义上看,数字人文的主体是谁并不是很重要,更重要的是研究什么样的问题,可以把相关的研究分成三种类型:第一种是有趣的研究,这往往是学科中的新问题,而且是创新性的问题,出现以后可能会有很多参与者;第二种是有用的研究,不管是对学科有用,还是对社会有用,同样会有很多参与者;第三种就是有为的研究,因为某些明确的目的,尽管目前或许看不到实际意义,但还是需要做的,数字人文目前大概还属于这一种。所以在鼓励学科发展的意义上,无需对数字人文进行定义,但在实践的层面,尤其是教学的层面,向社会及学生进行解释的时候,数字人文的定义就非常重要了。南京大学陈刚则基于历史地理学科的发展提供了新视角,指出尽管历史地理自有其历史,且已经与历史学研究结合,利用计算机及数字技术开展了一些研究,比如地理系统与考古资料结合考察六朝时期的墓葬情况、复原历史城市空间等,但在发展中也出现了一些问题,比如历史信息化对应的研究领域较为具体,所面向的也主要是理工科学生,多学科合作层面比较欠缺。而近年在与历史学者的合作中,他发现数字人文可以拓展研究的视域,要求研究者在具备专业素养的基础上,还能对历史问题有所贡献,更要经得起人文学者的批评。

2 如何定义数字人文

在有关如何定义数字人文方面,学者们则主要从“什么不是、什么是”“数字性与人文性”两个方向,对数

字人文的定义路径进行了讨论。前者的提出是针对无法准确定义的权宜之计,以反向思考的方式提供了另外一种路径。邱伟云指出,国外数字人文学界关于数字人文的定义已经进行了十多年探讨,目前已经进入了第二阶段,即讨论“什么不是数字人文”。王涛则进一步地对“什么不是”进行了阐发:重复性的、介绍性的、没有人文问题的数字人文研究,不是或者说不是好的数字人文研究;数字人文研究必须强调以研究为导向,以人文问题为旨趣,方法仅仅只是一种解决问题的路径。而梁晨指出,目前有种对数字人文的批评,认为数字人文因为使用了科学方法而不具有人文性,事实上,这种观点的出发点本身就值得反思:是否只有一种规范的、通用的人文研究方法?为什么强调方法就没有人文性了呢?人文研究一定要关注人文性么?传统人文研究总是强调解释,强调个人对历史、事件的理解,并且这种解释需要具有故事性和人文性,但数字人文是数据驱动的,是“硬”的,是基于“发现”的,就像其团队用大量时间和精力进行了《缙绅录》的数据整理,过程漫长,但结果很可能就是一张图表,这张图表难道就没有人文性么?梁晨对此提出质疑,并认为历史学家更应该强调“史诗”,用更好的方法无限地去接近文本中的史实,更好地处理文本;如果说历史学家需要懂更多的语言去收集史料,那么计算机在今日也是一种语言,因此,方法本身其实并不能决定事物的本质,更重要的还是要回到研究什么问题上去。北京大学王军则从数字化、数据的角度,谈到了对定义数字人文的理解。他认为,数字人文的发展有两个基本趋势,一个是数字化工具和方法对人文研究产生影响,一个是人文研究在数字环境中发展,而数字环境就体现在工具和方法上。但数字工具和方法的局限性也是很明显的,它们无法对人本身的独特价值进行解答。数字人文需要对人工智能这样的技术有所了解,同时也需要对数字环境下的人有所认识,即一方面需要知道人工智能这样的算法是依赖于数据集与计算能力的;另一方面,从人的角度去看这些数据集和计算能力,就会发现数据集也存在偏差,计算能力和算法的提升也受到人的行为、判断的影响。如果借用吴国盛教授的观点,即“liberal arts”与“humanities”是有区别的,数字人文在前者看来就是有问题的,只有在后者的意义上才是成立的,这就是因为后者是学科意义上的。数字人文给科学和人文的结合提供了一个平台,为学科融合、学科交流提供了基础,打破了学科边界,因此数字人文不仅仅是工具,而且还是一种新的视角,会引发研究范式的变化,也正是在这个意义上,定义是非常困难的,数字人文本身因其革命性和流动性,具有反定义的内在气质。殷曼婷接着指出,王军提到人文的分层定义和由技术带来的范式革命很重要。人文可以从三个层次理解,人文关注人、关注个人,也关注普遍人性。个体是社会的底层,而人文研究往往更多的是关注作为群体的人和普遍的人性,比如史观,这是一种对社会发展背后逻辑的理解,涉及普遍价值观的建构;而在数字时代,出现了建模观念,这是一种新的世界观架构,可能导致普遍价值观的变化,因此,传统人文学科必须要面对变化的问题,必须面对当代语境中的技术如何改变人的问题,如数据集,就不仅仅是数据的分类,也是人的观念的分类,一种群体的分类是会作用到群体自身的。

3 数字人文的学科深度

数字人文的定义与跨学科合作密切相关,这一方面涉及其边界问题,另一方面也涉及深度问题。清华大学严程指出,数字人文是人文学科的重组和融合。她结合在清华大学的实际教学经验指出,从宽泛的意义上看待人文,就会发现使用古籍数据库、数据处理、数据分析等新的方法和视角都是很好的学习方式,可以帮助学生打破人文学科之间的壁垒,这些方法不仅仅节省体力与时间,还能促进产出真正有价值的成果。上海图书馆的刘炜则从其从事文献整理工作的多年经验提出,传统人文研究方法在数字时代依然是有效的。文献研究者在数字时代的任务依然是保存人文价值,保存这个时代的每一张纸片,其关于人文研究的思考和问题也是有价值的,但文献学从本质上讲就是跨学科的,是为其他学科提供资料和证据。对比现在的技术,就会发现之前因为技术和工具的限制,文献是平面的、缺少细节的,对深层次内容和演变过程的研究都忽略了,但现在

的数字方法有能力把这些内容尽可能多地、全面地保存下来,从而支持研究者从专业知识的角度提出更多的问题,开展更加深入的研究。中国社会科学院文学研究所的赵薇则从“深度合作何以可能”这一问题出发,进行了比较系统的思考。她认为目前数字人文的跨领域合作体现了数字人文的深度,与此同时,怎么合作、以什么样的方式合作其实是非常重要的;目前很多跨学科的研究是以技术为导向的,用一些可能在非人文领域已经比较成熟的技术方法来开展人文研究,但其想要解决的或者已经发现的问题,在人文领域中并非一个能够成立的问题,甚至是背离某些基础事实的,那这样的研究价值何在?即使是跨学科合作,也是需要一定内部规范和价值体系的,而这需要人文和数字两方面的学者共同推动。人文学者需要了解技术实现路径,要有一定的“现场感”,才可能在实践的意义明白究竟技术能做到什么、做不到什么,也才可能取得有价值的数字人文研究成果。

4 数字人文与方法

关于“数字人文是什么”的讨论,一直都与量化研究、数据分析、人工智能等技术和方法论问题密切相关。上海交通大学葛岩和秦裕林从社会科学的视角,提出了他们关于人文研究和量化方法的观点。葛岩指出,从认识论的角度而言,计算社会科学、量化史学很早就出现了,强调以统计的方法来解释历史问题;而在人文研究中,基于推理和想象的人文与基于实验和量化的科学都是有意义的。科学认识论中的试错与验证是基于概率的,同时还需要量化的证据。在人文研究中,既要满足推理,又要提供量化证据,是非常困难的,但现在数字人文的出现,代表了一种新的趋势。他们推测两者的结合会在将来实现一些目前人们做不了的事情,因此不妨在现在树立起这面大旗,在学术领域中占据一席之地。梁晨从事量化史学多年,他从个人经历谈起——作为经济学出身的学者,在历史研究中应用量化方法是非常自然的事情。他认为,数字人文与量化方法相比,更加广阔,首先,它结合了很多元素,比如明确的工具意识、明确的方法论意识;其次,将数据分析与经验解释结合,用技术的方法解决传统史学的问题,且需要在史学领域内获得认可。人文学者与计算机专家合作,双方的诉求不同、关切点不同,对人文学者而言,更重要的是如何利用计算方法更加有效地处理史料、理解史料,大数据在这方面有明显的缺陷,往往会忽略特殊性问题,但历史研究非常注重时代以及那个时代的史料,因此他这样的历史学者同样关注当下以及当下处理史料的方式,并抱有一种警惕的态度。

刘炜则从六个方面综述了数字人文与大数据、人工智能的关系。第一,数字人文与人文并不是对立的关系:人文性并不能与技术脱离,没有数字技术,人文自身也无法孵化出数字人文。第二,大数据、人工智能是数字技术的新阶段,从数据管理和数据处理的角度而言,为人文研究提供了基础服务。第三,数字人文中的技术与计算、建模、环境模拟、方法论有关,但并不是所有的人文研究都需要应用技术、都需要加以量化的。第四,数字人文还在发展的初期,目前方法论上带来的颠覆性还不是很明显,更重要的是人文学自身,会在吸收了技术变革之后衍生发展出新的变化。第五,数字人文中存在很多问题有待解决,比如数据滥用、诚信版权、伦理、环保、可持续和商业应用等。第六,图档博机构的作用是作为文化机构保存人类文明,但未来的文明是数字的,该怎么办?裴雷则提出,图书馆恰恰提供了基础设施,并且不仅仅是硬件设备,更是基础数据、关联数据、规范数据的共享;人文数据准确来说还不算大数据,但大数据所代表的数字思维在人文研究中确实很有帮助。

5 数字人文的跨学科教学

在会议的最后阶段,学者们还针对数字人文跨学科合作教学进行了讨论。目前,南京大学、中国人民大学、北京大学、清华大学等多所国内高校都已经开设了面向本科生和研究生的数字人文课程。王涛作为国内

第一位面向研究生开设数字人文课程的教师,介绍了授课的一些经验。他指出,这门课虽然是面向历史系学生的,但更多的是从跨学科的角度去强调理念的培养和技术的学习,不仅仅是为了让学生意识到软件的实际操作可能解决文本分析的什么问题,更多地是促使学生形成一种从技术思维看待问题的方式,比如使用 GIS 系统的目的是培养从空间观念看问题的意识;数字人文的教学内容也包含批判思维和批判方法,这与传统人文学科的教学是一致的。天津大学博士研究生马昭仪和南京大学博士研究生谢丽萍也从自身学习的经验指出,数字人文所提倡的跨学科教学虽然在学科内部,比如与本专业的老师交流时会遇到一些困难,但其提供的视野是非常难得的,同时还提供了新的理解和解释路径,从而开辟新的研究领域,使自己与同年级同学相比,有一些新的研究成果。

尽管此次会议讨论议题广泛,甚至到最后也并未就“数字人文是什么”达成共识,但与会学者认可了数字人文中的“数字”与“人文”两者结合的重要性和相互关系的价值;同时也认为在当前的数字环境和学术发展趋势下,数字人文将会是非常具有潜力的领域,而在其发展的初期,不做范围上的界定,不划定边界与领域,而是从核心议题和焦点问题出发,对其进行内部的约定,并就其成果进行一种价值上的评定,以“什么是好的数字人文”来正向肯定、引导数字人文研究,将会是更具有意义与价值的“定义”。

"What is" and "How to do" —Summary of "Defining Digital Humanities" Seminar

Chen Jing

Abstract In order to answer what digital humanities is and how to define digital humanities, the core connotations and possible boundaries of "Digital Humanities" are clarified, especially from the Chinese context, a reflection path is proposed. More than 20 scholars from the University of Edinburgh, University College London, Nanjing University, Zhejiang University, Shanghai Library, Shandong University, Peking University, Shanghai University, Tsinghua University, Chinese Academy of Social Sciences and other units on November 8, 2019 gathered at Nanjing University to focus on "Digital Humanities needs to be defined" "Digital Humanities and organization alliances" "Digital Humanities and teaching" "Digital Humanities and academic journals and evaluation mechanisms" "Digital Humanities and literature, history, philosophy, etc." "The relationship between Digital Humanities and big data, artificial intelligence" "Digital Humanities and measurement research methods: the same or different" and other topics. They had heated debates and discussions around these issues. And a constructive consensus was formed into "Why need to define Digital Humanities" "How to define Digital Humanities" and "The possibility of Digital Humanities expansion".

Key words Digital Humanities; Digital Humanities Teaching; Academic Evaluation; Seminar Summary