

编者按：

2020年9月12日，“数字人文的学术评价体系：定义与规范建构”研讨会在南京大学举行。本次研讨会由南京大学人文社会科学高级研究院数字人文创研中心主办、南京大学数字史学研究中心和南京大学数字人文与超媒体GIS工作室协办，来自历史学、文学、地理学、计算机科学、信息管理学、艺术学、空间科学、图书馆、期刊界等学科领域二十余位专家与会。为相对原始完整呈现本次会议的交流成果，本刊以“汇编”的形式刊登与会学者发言稿的主要内容，以飨读者。

“数字人文的学术评价体系：定义与规范建构” 研讨会发言汇编

王 涛 陈 静 等
吴雪映 / 整理

引言人

王涛（南京大学历史学院）：数字人文的学术、技术与价值评价

现场的各位老师、同学以及线上的朋友们，很高兴用这样一种方式召开一次跟数字人文相关的会议。处于疫情状态之中我们用了这样的技术方式，让我深刻体会到了所谓“不在场的在场”状态，这在很大程度上一方面是在兑现数字人文的一种理念，另一方面也说明技术确实在改变我们做学问和交流的方式。

今天会议的主题看上去很有雄心大志，实际上这只是其发展到某一阶段必然的事情。从学术的意义上来讲，学术评价有两个层面的含义。一是学术工作价值怎么样，好不好，相当于给学术工作打分。二是跟人挂钩的，高校老师的劳动要

体现价值，你的工作能给你挣多少工分，你可以通过这项工作拿到多少得分点。

今天我们要讨论的评价，或者说我自己比较关注的评价是第一种。这跟我的个人经历有关，这几年在数字人文领域中遇到的一些事情引起了我的思考。因为我也忝列在所谓数字人文专家行列中，基于这个身份就经常受到邀请去评价一些稿件，包括学位论文，还有一些对数字人文友好的学术期刊杂志，经常给我发审稿邀请。开始我还挺高兴的，因为这是一种标志，是所谓的同行评议，同行评议是很重要也很有用的一种形式。但是我经过这几年的审稿，突然发现同行在哪里？我算不算真正意义上的同行？因为现在收到的越来越多稿件让我不敢评价，有很多是我并不了解的，比如《左传》《四库全书》，或者某一个很具体的技术领域。我何德何能去做这样的评价？同行评议本身是很传统的学术所采取的一种方式，传统的学术同行评议在专业变得越来越精细化、每个人擅长的领域变得越来越窄的情况下，已经出现问题了，何况是在数字人文这样一个本身是以“跨界融合”为导向的领域。我是不是有资格去给别人辛辛苦苦做的研究品头论足，甚至决定它的生死？这是不是合理的？这让我意识到，在数字人文研究中，同行到底是谁，有没有资格，是一个问题。

还有一个故事，在我们某个群里，前段时间发生了一场激烈争论，这个争论是围绕着可视化而来的。清华大学向帆老师基于CBDB（中国历代人物传记资料库）的人物家谱可视化作品拿到了他们专业领域一个很重要的奖项，但是这样的工作在一些人看来没有价值。这就引起了我的思考。我们有一句俗语“敝帚自珍”，你认为很好的东西在别人看来可能是很一般的东西，这样一种状态在数字人文领域更加突出。这个故事其实让“同行不存在”的状态变得更加极端。向老师的作品得到了传统意义上同行的高度评价，但是在另外一个圈子中被觉得不过如此。那怎么办？

第二个故事其实让我触动挺大，所以才会有今天这样一个议题，就是怎么去评价，特别是进行这种其实非常专业化的学术评价。我觉得至少可以有一些方案。有两种方式的评价，第一种是进行很简单、安全的技术评价。技术的评价是对研究进行一个规范的定义，规范的定义可以基于三个层面。第一个是去看他的资料来源，这是很传统的方式，看这套资料是不是完整可靠，这是第一个前提。因为所有研究，不管你用多么高科技的方法，如果基础不稳，那后面的东西都是有问题的。第二个是方法，我们要去讨论你运用的方法，不管是文本挖掘、社会网络还是各种统计，这个方法是不是很好地回答了你的问题；以及你的方法是不是可以还原，像做实验一样，我们拿到你的材料，用你的方法，可不可以得到你的结果？第三个是看你通过数据和方法得到的解释是不是合理。因为所有的数字

人文研究，用了方法和数据之后，最终会有一个结果，这个结果具有什么样的价值是要经过学者解释的，那么，这个解释是不是合理的？

这三个层面很重要。如果我们认同这样的一种技术评价体系，那么它会非常直接地影响到数字人文学术成果的呈现方式。今天特别邀请了《南京大学学报》主编王浩斌老师，因为我在他那儿做了一个数字人文相关的专栏。我们想兑现这样的方式：数字人文研究相关的成果，除了提交一篇论文之外，还要把你的数据和方法一并提交，而且必须要让读者能够下载到，以检验你的过程是否合理。我们的专栏现在还没有正式发表，我请他来开会相当于督促一下。我们在网上也进行了呼吁，希望能加快编稿速度，让我们在学术规范的建构上有一个《南京大学学报》的标准。我觉得这个标准可以推而广之，对学术共同体构建有很大价值。

第二种评价是价值的评价。数字人文产品或者成果是多元化的，不单纯是一篇论文。不同的成果有不同的评价体系或者评价方法。传统意义上，数据库可能不算正式的研究成果，但是从价值的体系来看，数据库是一个非常重要的事情。比如，CBDB成了数字人文的一个标志性产品或者成果，我们对它的评价其实是基于它的价值。价值的评价为什么显得更加重要？用喝酒的例子来说，我不喝酒，所以我没办法评判酒的好坏。虽然他们都说某个酒好，但是在我看来不过如此，哪怕是茅台。这说明我是一个外行，没办法去评价茅台的价值，如果你去掉茅台的包装给人喝，他也许并不会说这个酒是好酒，在茅台标签上附加有很多其他的价值。

这就让我们意识到，对于数字人文的学术评价需要把一些额外的价值附加到数字人文产品或者成果身上，这是一个增值的过程。当然我们不是要增加虚伪价值，茅台还是有实实在在的价值的，并不是因为喝的人多了才受追捧。从这个意义上来讲，我们做数字人文成果的价值评判，对于整个数字人文的学术发展有很大的帮助。

致辞人

姚远（南京大学社科处、南京大学政府管理学院）

从一个文科的研究者，或者从代表我们学校的人文社科管理部门的角度而言，我觉得这个会恰逢其时，也非常重要。我关注到会议材料上其实写了南京大学两方面的资助，一个是南京大学人文社会科学高级研究院，另一个是南京大学

的国际战略合作伙伴计划，而这两个方面恰恰显示了数字人文的两个特点：跨学科性、高度的国际性。

因为在疫情之中我们看到了对数字人文资源的使用，不仅是国内的，也可能是国际上普遍的。从我自己的学术经历上看，我对两个机构的数字资源有很深的感受：一个是东京大学的东洋文化研究所，从汉籍到艺术史相关资源非常丰富；另一个是哈佛燕京图书馆的资料。今天来之前，我也看了一些相关学校的研究和教学工作，我看到像在哈佛的Peter Bol——著名的中国史专家，面向哈佛的学生开设数字人文的课程，东京大学也开了好几年关于数字人文的课程，而日语中用了另外一个词来表达这个概念：人文情报学，翻译成中文叫做人文信息学，和信息管理有很大联系。所以这样看我们今天的讨论，其实在国际上有很深的研究基础。这对南京大学乃至更广泛的中国人文社科发展有几个方面的意义：

第一是新文科的建设。教育部提出新理科、新工科、新农科、新文科的建设，今天国家的大数据发展战略，对于文科在文化遗产、治国理政中的新作用有更高的要求，数字人文将是一个重要的新着力点。第二是在中华文化保护传承中的重要意义。各位老师做的工作都和我们文化遗产有着直接的关联。第三是它在文明交流互鉴中所起的作用会日益突出，就像之前提到的东京大学和哈佛大学等国外研究机构在研究东亚时，或者说汉学研究中的数字化的努力，中国也应当发挥更大的作用。我们和同行交流，展现中国文化，同其他文明进行对话，数字人文是一个很重要的平台和渠道。

今天讨论的主题是学术评价体系。在高等教育深化改革、破四唯（破除唯论文、唯项目、唯帽子、唯学历）等大背景下，我们对数字人文学术共同体的贡献和人文社科发展的意义应当会有一个新的评价，今天的讨论应当对今后的学术评价工作会有新的推动。从2017年起，南京大学就开始推进关于人文社科成果分类评价制度的试点，我想数字人文今后如何进入到分类评价的探讨或者说认定的范围，应该需要我们学术共同体去共同努力，希望早日把新的探索、新的贡献融入人文社科的综合评价体系之中来。

刘炜（上海图书馆）：机器时代的人文主义：谁有资格来评价？

一、外在与本质

2017年10月25日，在沙特首都利雅得举行的“未来投资倡议”峰会上，由香港汉森机器人技术公司（Hanson Robotics）制造的机器人索菲娅（Sophia）被

沙特政府授予了公民身份，成为世界上第一个受到政府认可的机器公民。这个看似作秀的举动其实蕴涵了一个巨大的跨越，标志着人与机器之间的界限正在被打破。这是否意味着源自启蒙运动的“人类中心主义”价值观受到了冲击？人类主体的范围是否应该重新界定？这些问题虽然还有待观察，但已经提上议事日程。

对数字人文而言，将其定义为数字技术在人文研究中的应用，或者是一个交叉领域，似乎已经远远不够了，人文学科的数字化发展只是其表象，人类数字活动的学科化研究才是其本质。循此方向发展下去，它不仅在颠覆传统人文，而且在创新传统人文。我们现在按传统学科划分的数字人文，难道就是未来数字人文的最终形式？当一个异形的种子刚刚被注入人类体内的时候，建立怎样的检验标准（即评价体系）都是荒谬和无济于事的。

二、载体与内容

当我们谈论数字人文的时候，通常谈论的是数字方法应用于人文研究，很少谈论人文研究对象本身的数字化形态，印刷时代的研究给我们造成一种错觉：似乎文本或文献就是所有人文学科的研究对象。例如，哲学研究的应该是思想观念，而不是概念文本；历史学研究的是历史事实，而不是历史文献；文学可以算是内容与形式统一的范例，但它并不是语言学；语言学，也不只是文本，而是音形义的交互作用。这就造成一个奇特的现象：所有的人文学科在外人看来都是研究文本或文献的，而不是其所宣称的真正的研究对象。人文学者在研究时，需要把文本抽象还原成其研究对象本来的状态，这是在人文学者大脑中进行的动作，人脑就是他们的反应器，人脑的记忆衰减、任意联想以及不确定的多巴胺造成人文学者极端的浪漫性，而比较缺乏的就是科学性。这可以看成是印刷技术的局限。要知道传统的自然科学家除了看书之外，不进行实验研究是不可能得到同行承认的，这也是爱因斯坦并没有因为其最伟大的发现——相对论理论而获得诺贝尔奖的原因。而人文学者完全不必有此顾虑，这是所有人文学科、中国传统理论都享受到的好处。

而当数字技术参与其中之时，情况就完全不同了。数字技术是可以把研究对象及其现象还原为真实对象甚至真实场景的，是可以基于数据进行实验的。这类似于人类发明了火之后，食物的种类大大增加，因为烧煮过程减轻了人类消化系统的负担，等于体外消化了大部分材料，于是人能够一跃上食物链的顶端。对于数字人文来说，不同学科的研究材料产生了很大的不同，大家不仅烧煮的食物不同，连锅碗瓢盆、过程方法都是不同的，你一看就知道哲学家可以用一种叫 N-Gram 操作的“蒸锅”进行观念史研究，而历史学家在调试各种“佐料”的刺

激程度以获得不同的中心度指标，来取得人物之间的聚类关系等，从一开始他们的装备和本领就是不同的，因此到了数字人文时代，可以给不同学科的研究者设计不同的装备，当然这些激动人心的丰富性也会给评价带来难以估量的困难。

三、模型与镜像

数字技术可以让我们在虚拟世界再造一个人文环境，跨越时空，以不同角度和粒度进行观察、模拟甚至操控。这其实是眼下最时髦的一项技术——“数字孪生”，我觉得它在数字人文领域具有广阔的应用前景，一定能大有作为。

“数字孪生”能够给人文研究带来极大的好处，它并不只是让我们能够通过操控和实验，获得细致入微的因果规律，利用机器的“显宏”能力写出史景迁《王氏之死》这类史学小说，或者黄仁宇《万历十五年》这类远读的历史切片，或者帮胡适续完《中国哲学史大纲》；它同时也带来了很大的困惑，究竟怎样的模型（粒度与视角）才能满足不同学科的实验要求？同样的素材可以提取多种对象，每个对象也可以有不同的形态和版本。

这里正好有个例子。昨天有个会议正在举行——“用方法表达思想，历史书写的可能性”，我们就来分析一下用什么方法能够表达思想？可以说当今的时代正在发展各种方法，不仅仅是印刷文本，数字文本可以表达，数字图像可以表达，各种对象关系模型及其系统实现可以表达，任何信息媒体只要受信端能正确（或故意正确）地解码，都算是从认识论层面或多或少表达了思想，书写历史也是一种表达，不同粒度、不同角度、不同分面，尽管认识的模糊性必然导致表达或书写的模糊性，但这不属于我的讨论范围。

我不知道历史学家或人文学者意识到当今数字技术所具有的能力没有。我们现在当然还没有做到大规模建模，为人文学者创造一个可以赖以实验的数字孪生出来，我们偶尔会受到指责，说我们的模型根本不是那么回事；我们也会指责回去，认为他们根本没有弄清楚自己在做什么，以及是怎么做的，这是无法建立一个清晰的模型的。当然，是不是机器学习可以不必打开人文学者研究过程的黑箱，而直接模拟得出结论，有待于计算机技术的进一步突破。对于评价来说，这种情况下我们如何评价？成果的取得是由于模型的精准、人文学者的顿悟还是机器算法的高超？现在都不知道。

四、现状与未来

我们可以做一个思想实验：当机器时代最终到来，所有的研究素材、对象、方法、过程、工具，都数字化了，我们的人文研究，还有必要冠以“数字”人文吗？我们正处在大数据和人工智能时代，当所有的人文活动、素材和方法都依赖

于计算机和网络之后，其实就无所谓数字人文了。就像纸媒时代我们不会说纸张人文一样，人文就是人文，它可能会改变形态，但本质永远是人性。我们可以借助纸本文献研究传统人文，也可以借助数字媒体研究数字人文，或者用数字方法研究传统人文，这些都是数字人文作为学科的表现形态。

五、定义与规范

什么是人文，什么是人文主义、人文精神、人文科学、人文学科？这些概念其实都没有得到清晰的解释，也没有形成共同的理解。我们曾对这些概念进行过一些简单的探讨，请参考文章《数字人文的理论化趋势前瞻》。^①关于什么是数字人文学科，人民大学在申报数字人文硕士学位授予点时，曾下过一个迄今为止最大胆和最清晰的定义，也是我最推崇的：数字人文是计算机学科和人文学科交叉研究所衍生的一个新领域，它以涉及针对计算工具与所有文化产品的交叉领域为研究对象，系统地研究数字与人文相结合的普遍规律和应用方法。具体来说，数字人文学科以人文科学的基本问题为研究对象，以不断发展进步的信息技术和数字技术等为主要工具，以数字资源构建、信息资源管理等数据基础设施建设为基础，以计算分析和案例分析方法等为主要研究手段，通过建立描述学术活动理论、方法和功能的框架以及各种类型的项目实践，探讨数字技术与人文科学跨学科对话中的方法、过程、特征和相互关系，以及数字人文作为一个整体与社会环境之间的互动与关联，并从中探索、归纳和总结出获得成效、提高效率的一般理论、方法和规律，以推动知识创新和服务。

六、当下与终极

虽然我不认为当下有必要甚至有可能建立一个数字人文的学术评价体系，并为其建立学术研究规范，因为现在还为时尚早，理论体系尚无踪影，领域边界也未形成，学术团体此起彼伏，对话还不充分，学术共同体也在飘摇之中；评价对象、评价方法、评价目的都不确定，过早的评价可能限制其发展，而不能达到我们促进领域发展的目的——但我并不反对这样的探索，更不反对这样的会议能够开诚布公地把问题提出来，进行争鸣。

讨论“什么不是数字人文”以及“什么是好的数字人文”都是有积极意义的，能够起到引导作用，而且在寻求评价方法的同时，自身就应该重视指标的可度量性。当然学术评价的终极目的是促进学术发展，最终还是一个定性的事情，同行评议可能是不好的方法当中最好的方法。人文社会科学的任何定量模型，其最初

^① 王丽华、刘炜、刘圣婴：《数字人文的理论化趋势前瞻》，《中国图书馆学报》2020年第3期。

量化数值的获取，都是模拟的、定性的。

七、启蒙与赋权

数字人文可能而且应该成为又一次人文启蒙运动，以数字工具的兴起为契机，以机器语言为共同语言，引导人们走出模拟时代。数字时代应该进一步研究何为人性、何为人文，在面临数字技术的冲击下，哪些是解放人类并促进全面发展的必然趋势，哪些又是别有用心政客用来奴役和控制别人的主张的，其实评价本身也是一种特权。

数字人文研究，必然带来后人文时代即数字主义的兴起。我们当下所有评价的宗旨，是发展数字人文。而数字人文的最终目的，是赋予数字以人性，并因此而保障平等的数字权利。未来我们不仅生产数字，消费数字，因数字而喜乐，用数字获得财富，而且也会困于数字、累于数字，甚至被数字毁灭。无论如何，现在已到了这样一个新的时代，我们可以喊出这样的口号——“天赋数权”。

冯惠玲（中国人民大学信息资源管理学院）：数字人文项目的团队合作（线上）

数字人文引起知识生产方式的改变，像是在学术圈相对平静的湖水当中扔进去一块石子，它泛起的涟漪会波及很多学术领域，或远或近。我们今天探讨学术评价问题，将来它的意义一定是超出数字人文领域的。长期以来那种以刊评文、以量胜出的规则在数字人文出现之后就立不住了，甚至在相当程度上要出局。新的评价规则和方式将是颠覆和创建并举，而且不会是单线和线性的，而是多维和复合的。

这次会议有可能是我们创建新的学术规则和学术秩序的一个重要早期节点。没有科学规范和具有共识的学术评价，数字人文就会缺少正当性，缺少生存和健康发展的条件，在制度性的认可等方面的条件支持都会不充分。而数字人文项目，特别是要做大型长线的项目需要这样的支持和认可。

所以科学周密的评价是急需的。另一方面，评价又需要以实际项目样本作为对象，这样才能对项目的优劣形成实质性的认识。所以现在的数字人文做事情出成果和做评价是互为条件的，这两个条件都不充分是数字人文早期的循环性难题，需要两条线一起推进。我想在我们获得制度性的认可之前，重要的是要在这时候把事情做起来，要做出不同类型的成果，获得一定数量的用户和评价，使它

形成一定的效果和价值。人们针对他的成果和生产过程，评价才更有针对性，也才能够找到它的特性所在。在做的过程中获得成败得失的真实体验，这样有助于我们形成对评价点的思考和需求以及对评价方式的构想。所以今天我想讲的是在数字人文这样一个以合作为特征的项目当中的团队合作问题。绝大多数的数字人文项目需要合作，越是大型的、长线的项目，越是需要在不同领域学者的合作中开展，良性的合作将成为数字人文重要的学术规范之一，也会涉及到合作中的数字人文成果归属区分的认定和保障方法。

去年在敦煌数字人文研习营上，英国伦敦大学学院数字人文中心西蒙·马奥尼（Simon Mahony）教授的讲座就是关于跨界合作的。他讲完之后现场有人提问，说合作其实很困难，如果合作不好怎么办？他回答得非常快，说合作不好就不和他合作。我认为这是一条路，但是把问题简单化了，因为在合作当中，合作者总会有不同的性格、不同的态度、不同的诉求，特别是长时期经历不同的学术训练，通常会有不同的认知思维和行为习惯。如果说两口子实在合不来，那可以分家，可以恢复单身，但是数字人文项目不是那么简单，分了还要找，因为它离不开跨界学者，离不开团队，也离不开差异性的接触。所以合作需要有合作之道，需要有经营合作的办法，数字人文有可能引发学术生产方式的改变，合作就是其中重要的改变之一。

我讲讲自己的一点实践经历，我们这个团队主要在做数字记忆项目，“北京记忆”和“古村落的记忆”已经做了六年多。数字记忆是比较复合性的数字人文产品，综合运用数据库、可视化、数字重建、数字出版、电子文学、电子创意等方法，所以遇到的合作问题比较多。到目前为止，我们大约完成了14个专题，还有六七个专题在建设中，另外还有资源库、嵌入式的专项数字产品等。在这个项目中，参与者大概有一百四五十人，来自不同学科、不同背景、不同身份，还有不同领域的若干家公司，分别承担不同的任务。怎样维系这样一个多元构成团队的运行？怎样让项目能够一直做下去，还能保证尽量提升它的质量？我想到以下三点。

第一，共同的价值认同十分重要，因为我们的数字人文项目是弱功利性的，在这种情况下就需要注重参与人对项目的意义追求，或者叫情怀支撑，探索的乐趣。在数字人文的早期阶段，参与者获得实际利益的可能性比较小，但是工作难度却一点都不小，具有相当的研究性和挑战性。这些参与者做好做差、做快做慢，项目组并没有什么刚性的约束办法。在这种情况下，非常重要的一条就是凝聚志同道合的人，获得参与者对项目意义的认可。我们和每一位研究人员都认真商讨承担数字专题的意愿，请他们做有兴趣、有研究基础的工作，这个兴趣可能

来自两个方面：一个是来自人文，对内容的兴趣；一个是技术应用，希望数字技术能够在人文领域中一展身手。有了这样的兴趣和探索愿望，就有可能把其他的功利需求降低。

比如我们做的冰嬉大典专题网站，是关于清代冬季接近百届的大型冰上活动的一个项目，主持人的研究方向是皇家园林及其活动。她对冰嬉做了好多年研究，能把成果转换数字形式呈现，她非常开心。又如我们的清代陆军衙署项目，一位老师因某种机缘开始研究段祺瑞执政府旧址（也是人大老校区）后，对其中丰富的历史文化产生了浓厚的兴趣，她说愿意为我们的老校舍“铁一号”做任何事情。类似这样的例子，把人文研究的内容转换成数字产品的过程虽然麻烦，但如果看成是自己研究内容或对象在数字平台的一种精彩延伸和扩展，研究者往往就会兴奋不已，甘于投入其中。事实上，在参与者当中认同度高的人，项目就做得比较好，而认同度相对低的人，推动得挺吃力。

第二个是规范运作，特别是充分的事先约定，让大家对项目有规则上的认同，明白责权。和传统的学术项目不太一样，承担数字人文项目的每一个人只是这个项目的一部分，他需要知道自己在项目整体当中的角色、任务、条件和预期的收获。因为只是一部分，那么在最终成果中，知识产权的归属也需要明确，这样可以在比较大的程度上避免责任和利益上的错位冲突，也有利于增强合作的稳定性。为了解决这样的问题，我们在制定了若干基本项目规则的基础上，大部分的合作关系都是以文本的协议或者合同的方式来明确各自的责任。项目组与每一位子课题的主持人，与合作企业以及所有资源授权人都签订了书面合同。对于某些知识产权有多种解决方案的情况，会经过专业人员和核心团队反复讨论，获得共识之后，才去选择一个方案。我们也请了专门的法律顾问，就知识产权归属等问题做出合规的判断，让每个参与者都认同。我们的数字人文项目至今已经和不同的乙方签订了多达六七十份合同，兼顾平台和参与人的利益诉求，用这种方式解决参与者的责权利问题，不仅支持了项目的运行，而且一定程度上也可以避免日后的误解和纠纷。

第三，明确的贡献认可，尊重参与，激励创新，提升参与人的项目归属感。数字人文项目做起来很难，有很大的创新成分。从提出点子到实施，要凝聚很多创造性的劳动，包括智力和体力的付出，最终汇聚到整个项目当中。参与者没有多少来自学校官方的认可，也没多少收益，但还是希望自己的努力和贡献清清楚楚，得到认可。他们尤其不能接受自己的贡献稀里糊涂地融在一个大锅里，甚至被别人占有。所以我们除了合同中明确规定的责权利之外，还会在项目上进行完全的标注，明明白白地予以公开，无论你的身份是什么，承担了哪一部分工作，

都如实标出。我举个例子，我们做的比较大的一个古村落的网站，一共有16项任务，16种角色参与，我们标注了27位参与人和5家公司，其中包括我们的教师、学生以及当地的干部村民等。

学术标注和成果标注也是传统学术规范的一部分，以表明成果的贡献者。但是传统学术标注比较简单，一些作者和主要编辑之外的人可能都被省略了。我们看到的电影一直以来都标注特别多，在片尾有长长的参与人的名单，包括各种杂务人员，甚至是司机，我和一位研究电影的老师讨论过，他说在电影制作中大家不认为人员重要程度不同，国际通用惯例就是都标注。我也看到了国际上的一些讨论，比如我们已经翻译过来的安妮·博迪克（Anne Burdick）的书中专门讨论了数字人文项目的署名问题，认为可以采取类似理工实验室或者表演艺术当中的合作署名方式。因为数字人文成果中，作者概念已经和传统的成果不一样。全署名的好处，第一是尊重和明确每个人的贡献，和知识产权有关。第二是激励责任人创新，精益求精，因为你的名字、工作质量、声望都在这里，一直挂在网上让大家去评价，参与公司也想为自己打一个漂亮的广告，署名有这样的激励效果。

当然大项目、大团队的合作靠这三条是远远不够的，总会遇到各种各样的碰撞，很多差异需要沟通，需要通盘思考，需要换位思考，需要做最优选择，需要有合作的意愿和智慧。我们既然进入了合作的领域，走向跨界合作、团队作战，就要力争把合作这篇文章做好。

姚远（南京大学社科处、南京大学政府管理学院）：数字人文与文化遗产保护

我其实不是数字人文方面专门的研究者，而是在做中国文化遗产保护公共政策方面的研究。我想从这个角度探讨一下如何发挥数字人文的作用。

第一，数字人文与文化遗产保护。不知道各位老师有没有在微信中使用过一个叫做“斯飞坐标”的小程序，它包括全国重点文物保护单位的地理位置和基本信息，是一个探访古迹非常有用的工具。它把5,000多处全国重点文物保护单位都做了标识，效果很好，而且这些都是靠民间的志愿者团队维护的。他们还做了一个叫“斯飞日历”的产品，包括了他们所探访的中国的文物古迹和国外历史遗迹，效果非常好。这给我们带来很大的启发，文化遗产资源如果以数字人文的形式呈现会焕发出非常大的生机。

我们有怎样的文化遗产资源，需要从两方面说起。一方面是不可移动文物。

新中国成立以来进行过三次全国文物普查。1956年开始普查，形成了国家级和省级的第一批文物保护单位，全国一共7,000多处。第二次从1986年开始，我们现存的大多数文物保护单位是从那时候开始设立的，从1986年到2008年，经过20多年的时间，中国编纂了《中国文物地图集》，那时候还没有广泛使用GIS系统，所有的地理位置和基本情况等文物信息是通过地图形式呈现的，耗费了大量人力物力，有些信息并不完全准确，但是一次重要的探索，如果以它为基础进行数字化有非常大的空间。第三次从2007年开始，历时五年，新发现文物古迹达53万处，文物总数达到76万处。这个数量远远高于其他国家和地区的规模，是人类文明历史上最大规模的文化资源遗产的资料收集工程。

另一方面是建设部门进行的历史建筑和街区登记划定工作。从2005年开始，建设部制定城市紫线管理办法，对城市的历史建筑和历史街区进行了划定，它要在地理信息系统上直接划清范围，不是文字上的一个点，而是直接落到了地图上，落到了具体的管理系统中去。2008年国务院颁布《历史文化名城名镇名村保护条例》，各地进行普查监控。南京市就全市的文化资源做了一次大规模的普查建库工作，初步的数字化已经完成，全市记录了数千处。在这样的基础之上，我们对文物保护点、历史建筑、历史文化街区、历史文化名城，已经有从点到面到片的GIS各项信息收集工作，这个数据库极为巨大。

在这两方面的基础之上，我们至少可以做两个方向的工作。第一是依托GIS系统，融入大量已有的研究，比如说地方志中的文物志，全国每一个县都有这样的基础性工作。从历史的、考古的、博物馆的、艺术史的、地理学的，特别是历史地理学的、人文地理的大量研究资料进行连接，基础非常深厚。第二是通过这样的新的数字人文工程，从已有的相关数据库充分对接。比如东洋文化研究所的中国早期资料档案中，大量的对华北的石窟寺古建筑的调查可以匹配起来。建设全国性文化遗产的数字人文工程，其意义非常重大。第一，从研究上讲，比如在数据库建立之后通过关键词的检索，回答学术史上的很多问题，如石窟寺、古城址的变迁，都有巨大意义，因为文物是历史上真实的存在，是最真切记载历史的信息。第二，推动了社会力量参与文化遗产保护工作，公众可以参与监督和日常巡查，这个系统是按图索骥的依据，靠传统地图集是不可能实现的，因为地图集没有电子地图的标识，也没有导航服务。第三，对于政府执法而言，它是一个清晰可查的公开的执法依据，所以对于文化遗产保护本身有很大的现实意义。

再比如从研究上讲，我们之前和北大的历史地理方面专家李孝聪老师谈过，他还关注一个题目——旧县城、旧县治、废弃的县城城址，这些信息在文物普查过程中被发现，如果有这个系统，就可以有助于了解这个城市的变迁以及为什么

会被废弃，给研究城市发展提供线索。

之前我们讨论过做文学地图，对于全国的76万处文物及几十万处历史建筑的巨大规模，如果做全国的文化遗产数字地图等，一开始难度肯定极其巨大。如果先从那些重要的历史名城古都开始试点，我觉得可行。把地图赋予更多文化信息，使它成为历史文化的地图，我觉得可以做到。在这个过程中形成工作标准，在地方政府支持下，更可能在全国进行推广。如果把数字人文工作的个案经验，结合以城市为基础的文化遗产资源的数字化工程，进行更大规模的国家性工程建设，可能使人文研究或者文化遗产保护有重大的突破，因为这个数据库实在是太重要了，它一定是一个国家工程，需要全国性的人力物力投入，应该会得到国家文化文物等有关部门的支持。

朱剑（《南京大学学报》编辑部）：旧瓶与新酒：数字人文评价问题献疑

近年来，数字人文在学界已有了一定的影响，有些成果也引起了广泛关注。虽然如此，但数字人文却尚未被纳入到现行的评价体系里，给从事人文研究的学者特别是青年学者带来了一定的困惑。那么，原因何在？在这里，仅从技术和操作的角度简单地谈些个人看法，或者说，提出一些问题供大家思考。

一、“旧瓶”：现行体制下科研、媒体、评价三者的关系

评价从来不是一个孤立的问题，而是与多方面因素纠缠在一起的。为了讨论的方便，我们先要说清一些关系，然后再来看看数字人文在这些关系中处在一种什么样的位置。

首先是科研与学术评价之间的关系。今天的学术评价，跟学术批评已经没什么关系，主要是为权力部门服务，是权力部门分配学术资源的依据。学术评价之所以为大家重视，正是因为它直接关系到甚至决定了学术资源和利益的分配。它已成为一种组织行为，注重的不是过程而是结果，强调的不是个性而是共性，基本上是关起门来的暗箱操作和一次性的结果。尤其在我们这样一个集中资源办大事的科研体制之下，要成为重点，先得成为评价对象。数字人文学者之所以感到委屈，就是因为连评价对象的身份都不具备，没有被纳入评价体系，就得不到资源和利益，怎能不备感委屈。因此，我们要自问一下：数字人文为什么难成评价对象？

其次是评价改革与回归同行评议的关系。评价如此重要，但评价的现状却难令人满意，评价必须改革已成各界共识。眼下的评价改革有三个动向：第一强调分类评价，第二强调同行评议，第三强调代表作评价。这三者实质上都是回归同行评议。但是要真正回归同行评议，前提是同行评议必须具备公信力和权威性。可是在现实中，同行评议又是如何？同行评议的结果远比定量评价荒唐的事并不鲜见。所以同行评议的真正回归会有一个非常困难的过程。数字人文评价的建构当然也离不开同行评议，且不说同行评议回归之难，还有一个更大的困难拦在我们面前：权威的同行在哪里？

再次是学术媒体与学术评价的关系。媒体在评价中起到的其实是桥梁和平台的作用。无论是对于学术研究，还是对于学术评价来说，学术媒体都是不可或缺的重要成分。抽离了媒体，也就无所谓评价。如果没有媒体，今天做定量评价的机构连基本的数据都没有。共同体评价虽然不会完全依赖数据，但媒体在其中有一个非常重要且不可替代的作用，就是成果展示和民意表达（学术批评）的平台，共同体成员的意志要通过媒体来表达。只要是成果评价或以成果为基础的评价，就离不开发表和评论成果的媒体，至少在目前看，数字人文也不会例外。那么，属于数字人文的媒体又在哪里呢？

今天我们来看科研、媒体和评价三者之间的关系，其实是非常明确的：科研负责生产产品，媒体提供平台，评价给它贴标签。在理论上，这三者之间的关系，不管是媒体还是评价，都应该为学者们的研究服务，而研究什么、如何研究以及最终产品都应该是学者们说了算的事。可是在现实中，这个关系正好颠倒过来。只要科研体制不改变，学术研究也好，学术媒体也好，学术管理也好，离开了评价都是寸步难行。分析了三者关系以后，就可以明白今天的数字人文这个“新酒”面对“旧瓶”——现行评价体制时的尴尬处境了。

二、“新酒”：数字人文何以陷入评价困境

数字人文在现行评价体制中为何会陷入尴尬困境？本次会议手册里对此有一个解释，就是主流学术界对数字人文没有深入的了解，数字人文并没有得到他们的认可。这确实是个直接的原因，但却不是根本的原因。因为我们可以接着追问，主流学术界为什么不感兴趣？我倒觉得手册中的另一段话可能更接近问题的真相：“所谓‘学术评价体系’，不正是强调专业壁垒的传统学术，为了巩固学术阵地而生造出来的外造物吗？数字人文以此为参照系，再造就一个新体系出来，是不是画地为牢？这是否与强调学科跨界与融合的数字人文精神背道而驰？”这段话指出了两点要害：第一，从学术研究的角度来看，专业壁垒与跨学科研究之

间关系的紧张；第二，从学术评价的角度来看，现行的分学科评价体系与跨学科研究的关系的紧张。除此以外，我还想强调，其实还有一个同样很关键问题，那就是传统学术研究成果的形式和传播路径（学术论文、期刊、专著等）的有限性、稳定性，与数字人文成果形式的丰富性、传播平台的多样性之间的紧张，这个特别重要的点，却往往被忽视。

由此我们可以看到，在今天的科研和评价体制下，数字人文陷入评价困境不外乎内外两方面的原因。

首先是数字人文本身的原因，即数字人文的不确定性。数字人文是什么？数字人文是方法，因为有独特的方法进而在未来能形成一种范式。这应该已是共识了。但数字人文是一个新的学科吗？我们大多不认为它是学科，至少在目前它还不是。前面很多专家的发言都强调数字人文的跨学科性，跨学科一定是在至少两个学科的交叉地带产生，是一种不稳定的过渡状态，跨学科的最终走向还是新的学科的产生。新学科是怎么产生的？新学科的建立其实就是一个在交叉地带划界的过程。划界最重要的有两点：第一，独特的研究方法体系的建构，这在数字人文领域似乎没有问题。第二，独特的研究对象的确立，这就存在问题。所以很难说数字人文已经是一个新的学科，它目前还有诸多不确定性。除了方法体系和研究对象外，还有一个在其他交叉学科科学化时并不存在，但却十分重要的一个不确定，就是数字人文的成果形式是什么？我们现在公认的学术成果形式就是论文和专著，可是凭借新技术而生的数字人文所强调的恰恰是对论文和专著形式的突破，但突破后形成的新形式是什么？还不确定。而成果形式的不确定性还必然导致另一个重要的不确定——媒体形式的不确定。传统的媒体形式就是著作和学术期刊。可是数字人文呢？是数字期刊、博客、网站，还是APP、数据库？似乎都是，似乎都不完全是，至少到目前为止，这一切都是不确定的。数字人文的不确定性这一点倒是很确定。

其次是现行评价体制的原因，数字人文这样一种新生事物，对于今天的评价体系来说，评价它的确是个难题。我们不妨看一看评价的现实情况。先看量化评价。评价机构对评价对象有如下明确的要求：一是公开发表，亦即成果已被固定；二是形式的确定性，因为不确定就无法生成可资比对的数据。所以，从数字人文的特点来看，对其进行量化评价非常困难。那么，引入同行评议可行吗？同行评议同样是需要条件的，至少需要满足的有两个：一是要有自律的和专业的学术共同体，二是要有评价的权力，这需要行政权力部门赋权。在学科正式生成前，这两点得到同时满足都很困难。

由此我们不难明白数字人文评价的困境到底从何而来。

三、从“旧瓶”到“新瓶”：是否需要及如何打造

在这种情况下，要对数字人文评价的前景做一个预测是很困难的，在这里，我只能提供一点思路。

我们首先需要就是否建构数字人文评价体系达成基本共识。这是一个值得思考的问题，因为这样的建构是要冒一定风险的，在一定概率上会出现新建的评价体系“与强调学科跨界与融合的数字人文精神背道而驰”这个我们不愿看到的前景，对于一切都远未定型的数字人文来说，暂时不做评价体系建构，似乎是个明智的选择。但是，事情又不是如此简单。在现行科研体制下，没有评价就没有资源，数字人文又特能“烧钱”，没有丰足的资源，恐怕连安身立命的本钱都没有。可见，我们面临的是一个两难选择。

面对两难选择，我们很难择其一而行之，只能选取兼顾两者的办法。一方面，还是应该建构起数字人文的评价体系，以获取必要的资源；另一方面，我们同时还要设法让这个体系有着足够的弹性，能够包容数字人文可能的成果形式和可能的媒体形式。因此，这必定是一个充满个性或特殊性而有别于传统的评价体系。这样的评价体系建构可以说没有先例可循，它的建构将有待于三方面的共同努力，积极地创造条件。

一是数字人文从成果形式到媒体形式完成迭代，走向相对稳定，并被学术界和管理部门承认和接受。数字人文是跟新技术相关的，没有数字技术，没有互联网技术，就不会有数字人文。从历史上看，每一次堪称革命性的技术变革一定都会带来研究成果形式和媒体形式的变革。今天的学术期刊和以学术论文为主的成果形式都跟现代印刷技术相关。如果数字人文是一种革命性新技术的产物，我们要证明它的革命性，要让别人承认它的革命性，就必须有新的成果形式和新的媒体形式的产生，并逐步走向稳定。这既是衡量今天的数字人文是否成熟的一个重要标志，也是独立的数字人文评价得以建构的前提条件。

二是学术界对学科特征尚不清晰的数字人文的宽容和接纳。无论是宽容还是接纳，都是要有具体对象的，数字人文作为一种新的研究范式或者一个独立的新学科，其存在并得到承认，最重要的是要拿出足以让学术界震惊的颠覆性成果。尽管数字人文在不长的时间里已有相当丰富的成果问世，但成果从形式到性质还比较模糊，大量成果还是工具性的或中间性的，离“颠覆性”尚有不小的距离。而一旦颠覆性成果出现，数字人文作为一个新学科的建立也就水到渠成了。

三是评价体系的设计者对数字人文的认同，并建构起相应的评价理论和可供操作的评价方法。这里的设计者，可以是评价机构，但更应该是数字人文学术共同体。对于数字人文这样一个变动不拘的新生事物的评价，同行评议可能比量化

评价的意义更大，可能性也更大。没有先例可循的数字人文评价更应该要强调评价主体的统一。无论是定性还是定量，都应该由统一的主体——数字人文学术共同体来完成。也就是说，如果要进行量化评价，主体应该是学术共同体，在数据挖掘的方向、数据的解释、算法的确定等方面，学术共同体应该更具有决定权，评价机构能够提供必要的数据就可以了。

显然，以上三方面，无论哪一方面的达成都必定是一个艰难而长期的过程，不可能一蹴而就。而在此之前，数字人文都将处在一个没有独立评价甚至难以纳入评价体系的尴尬境地之中。要结束这一尴尬处境，固然需要多方共同努力，但最有效的途径还在于数字人文学者拿出具有颠覆性的成果来，让学界刮目相看。

张耀铭（《新华文摘》杂志社）：亟待建构数字人文研究成果的评价体系

2020年2月18日，教育部和科技部出台了《关于规范高等学校SCI论文相关指标使用 树立正确评价导向的若干意见》。这个《意见》列出了五项负面清单。一是审慎选用SCI论文量化指标使用评估。二是职称职务评审评定，不把SCI论文相关指标作为直接依据或者聘用的前提条件。第三是解除SCI论文相关指标与资源配置和绩效奖励挂钩的使用。第四是不宜将发表SCI论文数量等指标作为毕业、学位授予的限制性条件。第五是在舆论宣传上不采信、不发布以SCI相关指标为核心编制的排行榜。

《意见》提出的举措与思考，虽然主要是针对的自然科学和工程技术，但是对人文社会科学，包括数字人文，具有重要的参考价值，也将产生重大的影响。我想提五点：第一，政策导向不是完全否定SCI和学术评价，更不是反对发表论文，而是破除SCI至上，鼓励发表高水平、高质量、有创新价值、体现服务贡献的学术研究成果。这对纠正评价异化和急功近利的学术风气、营造良好的学术生态是很有利的。第二，现行的学术评价制度，无论是北京大学的核心期刊，还是南京大学的CSSCI，都是依刊评文，就是根据论文发表在什么级别的刊物上，来确定论文的内容质量。核心期刊的评价体系的许多指标，比如被载量、被引量、影响因子、被重要检索系统收录基金论文比等，是为依刊评文量身定做的，不是为论文的内容质量定做的。

这样的评价指标基本不适合作为评价人文学术和数字技术结合的数字人文的评价。比如数字人文现在围绕人文社会科学领域特定研究所进行的数据库、虚拟

仿真模型、程序、代码设计、数据可视化、数字期刊、数字博客等，这些非传统学术活动及成果，还不能被认定或者转化为现有学术评价体系内可以接受的学术成果。现在认定的成果还是以论文和专著形式呈现的。这对于数字人文研究和科研评价是挑战，也是机遇。挑战在于数字人文的科研成果，不被主流的专业学术评价机构认可，也不被管理层认可。这不仅影响年轻学者的职称职位的晋升，更制约数字人文的可持续发展。机遇在于数字人文同行可积极推进学术共同体建设。数字人文可以先走一步，推进我们的学术共同体建设，另起炉灶，与时俱进，探讨建构数字人文研究成果的评价体系。

第三，数字人文研究以项目为单位，以跨学科为特点，以团队为主体，以数据为基础，以技术为手段，以平台为支撑，与传统的人文社会科学研究的对象、范式和成果性质都有显著的不同，因此应该加强学术评价理论的研究。数字人文界、学术期刊界和文献计量学界暂时可以依托数字人文基础设施建设比较好的高校和图书馆，通力合作，共同攻关，研制出针对数字人文科研的评价指标体系，将有利于数字人文逐渐规范化、科学化，从而获得更好的发展。

第四，根据数字人具有特定属性和特征，实施分类评价的标准和方法，明确不同学科的不同性质和学术评价的不同维度，完善定性与定量相结合的评价体系，特别要高度关注项目制评价，比如项目是否能够在数据收集、内容生成、标准运用、文档记录方面遵循领域内的最佳实践，项目内容的范围、深度形式、呈现方式与出版过程的性质特点，项目迭代更新将会以何种形式持续生存，并实现长期的利用等。

第五，推动数字人文学术共同体建设，完善学术同行评价学术共同体、学术期刊和学术会议三位一体，构成数字人文研究与科研评价的坚强基石。对新的学术模式和组织形式的评价，不仅仅在于专业评价机构的评价，更在于“小同行”对科研学术水平的认可。因此数字人文研究成果的评价应该贯穿于学术会议的评价，专业期刊发表的双向匿名评审，成果发表之后的书评评价与学术评价，让数字人文学术评价回归学术本性，依靠、相信、尊重同行专家的专业判断，充分发挥同行专家在学术标准制定和学术评价过程中的主体性作用。建立评审专家的遴选回避、民主表决、专家信誉和监督制度，维护学者和学术尊严，增强学术同行专家的自律，保障学术同行评议制度的公平正义。

赵思渊(上海交通大学历史系):数字人文需要独立的评价体系吗? (线上)

中国大学中学术评价都有比较成熟的评价机制,如论文、专著、项目基金,甚至包括艺术领域的一些展览,都有比较成熟的评价方法。我们今天在座的这些人曾经都多多少少发表过一些数字人文相关的研究成果,或者拿过相关的项目,也就意味着其实我们正在进行的这些工作是可以纳入到机制里面的。我可以写篇数字人文题目的论文,我也可以做一个数据库,在此基础上申请项目,甚至包括专著或者是教材,看起来好像挑战不是很大。

在当代中国大学的行政逻辑里进行学术评价的时候,我们这些做数字人文研究的人,都可以把自己的研究在某种形式上放到这个框架里面,而并不需要另起炉灶,这是我最开始想这个问题时的想法。不过在这里我想谈一个在中国人文社科研究评价中公认非常重要的指标。我以自己的经验来看,大概2010年之后,数字人文这个概念在中文学术界就已经变得越来越活跃,相关的论文也越来越多,但是如果看国家社科的中标项目的话,题目中包含了数字人文这个概念的国家社科项目,其实比我们预想得少很多,每年都是个位数,而且非常集中在图情学科。论文题目当中包含数字人文这个概念且有国家社科基金资助的,也是要到2017年之后才大幅度增加。但是我相信,在座各位接触数字人文概念远远早于这个时间。如果以基金项目作为观察切入点,数字人文的概念得到认可很晚。而且,它的总量没有我们想象得多,也没那么活跃。

所以我可能跟前面的几位前辈的想法有一点不一样,我觉得在现有的评价机制里,数字人文的研究成果作为一个独立的成果放入这些框架里面是有可能的,但问题是数字人文作为一种话语,获得认可非常困难。我觉得这是一个我自己不能解答的问题。我希望数字人文方法论能够嵌入很多不同学科本来的话语体系中,在不同的学科体系中都能够有办法论证数字人文方法的贡献与价值。

另外,数字人文一定是跨学科的,有不同的人参与,才有可能推进下去。在这种意义上,最好来自不同学科背景的每一个参与者,都能够对各自的学科有所推进,都能有所得,这可能是将数字人文用到更多学科、让更多的领域接受其方法论的一个重要前提。

还有一点比较理想但不容易做到的是,我们看到了非常多的数据库建设项目,或者是广义的动态的概念上的数字人文项目建设,但问题是最后它的成果在哪里?项目在哪里?我们看到了论文或者专著,但是没有看到最后的数据库,或者其他形式的数字人文项目,这些东西是不是有可能以一种更好的形式呈现给学

术界和学术共同体，来评价和使用，可以使数字人文作为一个概念，作为一个方法论，得到更好的评价。比如在英文学界每年会有数字人文的奖项，会有来自学会或者是第三方机构的评价，但是在中文学界好像还不太有。也许可以在目前已有的成熟评价体系之外，再生长出一个更适应数字人文本身特性的辅助评价支撑系统。

赵宇翔（南京理工大学经济管理学院）：跨界合作的数字人文成果评价思考

我想讨论三个问题。第一，跨界合作过程中对数字人文成果评价，到底是对成果的评价，还是对学者的评价，或者这两种都会有。我在思考，数字人文是否有必要，且是否可能建立一个所谓的学术共同体。这几年大家都在推动数字人文的发展，从评价角度来说，事实上很多时候要考虑开展评价的主体，它是否具有合法的效力。我个人感觉，“大同行中的小同行”这个概念比较靠谱。从我们每个人各自学科领域的角度来说，是不是能把数字人文先立起来，是一个比较有意思的问题。比如2017年武汉大学信息管理学院王晓光教授在中国社会科学情报学会下成立了一个数字人文专委会，这事实上是大圈子里面的一个小圈子。再比如近几年图书情报领域的数字人文社科项目数量在增长，评审论文时我们也遇到好多困扰，很多时候文章完全看不懂，没有办法做出评价。这些文章只能退回去，因为我们没有承担审稿工作的能力。我们在思考一个问题，就是学术共同体能否担任起评价的使命和职能。

在这个过程中我们到底是对于研究作评价，还是对于项目作评价，这是两个不同的事。对于数字人文项目来说，很多大型的项目从管理学角度可以称之为矩阵式的模式。所谓矩阵式模式就是大家为了一个共同目标，从不同地方拉人来干活，很多时候干完活就解散了。矩阵式的这种结构一定程度上适合数字人文的项目管理，但矩阵式是基于功利主义的一种模式，不完全适合于研究的评价。对于研究的评价，刚才王涛老师提出了一个比较好的模板，其实更泛地来说，作为数字研究，我们需要考虑两个主要的问题，一个是严谨性，一个是与实践的相关性。数字人文发展很多时候跟图书馆、博物馆、档案馆等这些拥有文化遗产的机构有密不可分的关系。他们是偏重于实践的，理论跟实践之间的对接和桥梁是一个很重要的问题。

第二个问题是对于数字人文认识的进阶和嬗变是否会影响到数字人文成果的

评价。不同学科数字人文的发展进度是不一样的，不同选题上的数字人文的发展也不一样，有些选题已经走得比较快了，而有些还是刚刚起步的阶段。我们经常会遇到一个问题，到底是“借船出海”，还是“造船出海”？对于很多传统人文学科来说，一开始接触数字人文时更多是“借船出海”，它需要一些技术力量去帮助推动研究。但是在这个过程中，不同个性和学习能力的人往往对技术的采纳会产生不同的影响。有些学者有非常强烈的自发学习意识和较强的学习能力，他就会非常快地上手一些新的技术和工具，也慢慢尝试“造船出海”，自己去做这个事情。在这个过程中又会有两个不同的概念：一个叫“轻舟缓行”，一个叫“乘风破浪”。很多时候，尤其数字人文在学科应用的早期，会去采用一些轻量级的工具和方法去做一些尝试。但也有比较有影响力的项目，我们称之为“乘风破浪”。不同的人群、不同的学科流派对于数字人文的接受也有所不同。

第三个问题是数字人文的合作动机和策略，是否会影响成果的评价。跨学科、跨领域合作有许多不同的动机和目的，在早期有一些非常明确的需求，比如历史学者需要做一个工作，可能采用的策略就是缺啥补啥，需要文本挖掘的工具，那就去学这个东西，或者去跟别人合作。也有一些另辟蹊径型的，比如说图书情报领域，大家可能发现这两年这个方向非常热门，是一个发文章和申课题的契机，所以大家会去找机会。也有一些范式转移，可称之为开疆扩土型，事实上它是在做一些比较大的破冰式探索，甚至希望对整个学科的结构都会有一些重新思考。

所以我们不得不去思考一些问题，数字人文合作，尤其是人文学者跟技术学者对接的过程中，是一种约等于外包的概念，比如说人文学者现在外包一个技术的难题给技术的学者，事实上学科位置是不对等的，不对等的情况下学术成果的评价又会如何？所以在这里我们不得不去深入到另外一个问题：数字人文的这种行动者网络如何去调动？这是个开放的问题，现在没有答案。还有一些细节问题，比如技术流派学者到底在什么时间介入，介入的深度广度到底有多少，这对于成果的评价，也会产生一定的影响。

我们借鉴组织行为学中的一些理论，参与合作有两个不同的视角，第一个是工具人视角，第二个是责任人视角。这两种不同的视角在数字人文开展过程当中意义是不一样的。所以从自身角度来说，我一直在做公众科学。刚才姚老师也提到了，他们正在做的文化遗产领域也经常利用大众力量去做一些事情。其实我们现在也在尝试用这种公众科学就是众包的模式去开展相关的数字人文研究，我们想了解大家各自开展的数字人文项目和团队合作，去做一些案例研究，从而探索合作模式中出现的障碍和通行点，找到一些比较可行的方案或者策略去做这样的工作。

徐力恒(香港城市大学中文及历史学系):数字人文研究之余的“数字”和“人文”(线上)

我着眼的主要是高校老师和研究者面对考核评估的实际操作。这只是学术评价体系中的一个方面,是比较世俗的一个问题。

数字人文研究还没有得到充分的制度性认可,属于新兴的跨学科领域,还处于一种尴尬的境地。有几个方面是数字人文学者需要努力去开发建设的,比如学术社群、学术数据、基础设施、学术期刊以及学术传播的机制,这些方面都要草创的情况下,学者面对和参与了很多不直接和数字人文研究挂钩的工作。但是这些工作都是和数字人文相关的,而且也都是学术事务,我把它叫做DH Work。这些都是大家做过的工作,但是这些工作又不必然能直接成为各个单位考核里面比较便于计算的输出成果,我把它叫做DH Output。我们可以换一种思路,把对成果的理解拓宽,看看怎么把DH Work和DH Output联系起来理解。

首先是数字人文的研究数据,没有数据就没有办法做DH研究。我们研究数字人文的时候,会参与到不少数字人文数据的生成,或者是校订、利用的工作,但是我们所处的人文学科框架下,很多同行普遍对数据的引用或者标注不规范。比如引用数据和原始文献上的区分,还不是特别明确。其实其他学科在这方面已经做得比较成熟,他们确立的一些规范和标准是值得参考的。数据引注的一些原则性共识,也不限于自然科学,其实社会科学在这个方面也比人文学科走得更前。使用数据的规范还需要数字人文学者的推动和建立。那么疑问就来了,谁来做?从我这边的情况来说,会有项目或者出版机构要求的规范来牵动这方面的发展。香港官方的研究基金机构在项目申请里要求明确说明数据的储存是什么样的情况,数据是不是公开,它是不是生成出来的数据,会不会跟其他的学者去共享,等等,也包括使用数据写出来的这些论文和学术成果是不是会公开或者是出版,会不会采用open access等。这方面的标准和规范慢慢会建立起来,真正要影响整个人文学界的使用还需要很长时间。

第二个方面是项目团队的合作。有的项目维持了很长时间,很多学者做了很多工作。但是从这些项目直接产出,转化为可计算的成果往往只是项目工作的一部分,就算发了研究论文,也往往是多人共同署名。摊分到每个人的工分只是三分之一,不会很多。

第三是基础设施建设,这方面还处于进行式。理论上基础建设做得好会便利所有学者,但是在那个愿景实现之前,学者还要投入很多精力去参与基础建设的讨论、协调它的标准等。我和以前CBDB的一些同事甚至会觉得,这样的工作

其实要有专业人士去做才会比较有效率，这方面也需要数字人文学者投入很多精力。

第四个方面是学术社群，培养新一代数字人文学者。如果没有这种社群，大家单打独斗就会很累，也很不容易，所以社群需要我们参与经营。但是数字人文和其他比较成熟的学科领域相比，没有那么多正式的学会或者委员会。一些培训活动或者教大家怎么使用数字人文工具和方法的教学，往往不是正式的大学课程，或者常规化的活动。它也是有意义的，可是在正式的框架里不容易获得承认。

第五个是学术传播。刚才朱剑老师也谈了学术媒体的重要性和数字人文新型成果的一些形式，我自己参与得比较多的是“零壹Lab”微信公众号的工作。2016年以来，我和公众号的同仁开始征集数字人文相关的内容去编辑发布，起到学术传播的效果，也得到一个数字人文奖项（DH Awards 2019），但是像这样的工作在单位不太容易得到承认，它是不是能算入成果呢？比如说香港一些大学有一项要求，大学老师要做一些知识转移（knowledge transfer）的工作。这方面我自己也在摸索。

第六个是新的发表园地的创立。一些新的期刊最近创立了，比如清华大学的《数字人文》、中国人民大学的《数字人文研究》，我想这些期刊的主事者肯定会有更深的体会。创办这些期刊需要投入很多精力，才能走到今天。

我提到的这六个方面的DH Work，在学术单位里面怎么能得到考核？我没有简单的答案，也不参与制定标准。我可以分享一些自己的应对方式，就是把这些工作转化为可以算数的研究成果。比如最直接的方法就是写成论文，写的论文尽量在各个不同领域发，写给不同对象看，让大家看到数字人文学者已经做了这样的工作，也就是尽量嵌入到各个学科和领域之中。这些工作如果没有带着意识转化为成果的话，可能做完就过去了。它可能起到一些实际的建设作用，但是对于自己本身的考核没起作用就很可惜了，而且从正面角度看，如果当下都能写出论文的话，也是介入学术界的方式，尤其作为一个新兴的领域，这种介入是亟待进行的。

当然，从个人在各个学校单位里实际的角度看，这也是增加一点工分的办法，做过的这些DH Work不至于变成无用之功，我想还是很必须的。最后我想给大家介绍一下，2020年7月出版的英文期刊《中国历史学刊》（*Journal of Chinese History*）的数字人文专号，它包括这几个方面的内容：导言、数字人文综述、研究论文，以及多篇数字人文基本工具的介绍。这个专辑的内容很前沿，参与专题写作的也都是比较重要的学者，但期刊还比较新，所以还没有进入SSCI

名单。从机构的考核方面来看，在SSCI至上的风气下，在上面发文章其实只算很少的业绩。总而言之，我今天谈到的DH Work在个人的考核里能够被承认，或者能够转化为DH Output的，到目前为止还是非常少。要把数字人文工作一一转化为可被计算的工分“数字”，挑战仍然很大。

姜文涛（浙江大学国际联合学院人文社会科学研究中 心）：什么是好的数字人文？

今天讨论“什么是好的数字人文？”我自己从事比较文学和文艺理论研究，从2016年开始与美国威斯康星大学麦迪逊校区从事现代中国文学研究的戴安德（Anatoly Detwyler）一起在《山东社会科学》杂志上组织“数字人文：观其大较”专栏，有七期了，刊发了不少国内外较为前沿的数字人文文学文化研究方面的文章。

我想从数字人文文学文化研究的学术组织经验，以及文学研究的学科史和属性方面，针对“什么是好的数字人文”这个问题提出七方面的意见，仅供参考。^①

第一是人文学术意识。我们需要继承人文阐释的悠久传统，才能继往开来，才有可能创造出新的知识生产形式和社会文化形式。我们相信数字人文会提出有价值的人文问题，会以自己的方式推动人文科学的发展。但这些问题并不会凭空产生，基础仍然是人文学的阐释传统。增强人文研究，需要具备人文学科的学术训练和数字人文基本方法训练这两方面的准备工作。这是开展好的数字人文研究需要具备的前提。

第二是社会科学的素养。数字人文研究的许多方面与社会科学密不可分，比如统计学、心理学、社会学等。这些学科中量化研究与质性研究之间的关系讨论、方法和实践已经有很多了。好的数字人文研究者应该具备一定的社会科学研究素养，有意识地从这些学科汲取方法论上的能量，以防止停留于数字炫技或者过度人文阐释的层面。当然这并不意味着人文研究要成为社会科学方法的附庸或者殖民地。相反，好的数字人文研究，应该对所研究的问题及所使用的研究方法在本学科的学科史意义上有充分的了解。

第三是学术史和学科史的视野。有价值的学术问题的组织和提出，一定是具有学术史和学科史视野的，是建立在学术史对该问题或相关问题已有讨论的基础

^① 参见姜文涛：《什么是好的数字人文》，《澳门理工学报》2019年第4期。此文被《人大复印资料·社会科学总论》全文转载，可在刊物网页或知网下载。

之上的，数字人具有天然的跨学科和学科融合潜质。然而我们需要注意到，所有成功的跨学科研究无不建立在对本学科严谨的学术训练基础之上。从学者学术生涯生存策略上来讲，对学术史、学科史的重视也是很重要的。就文学研究而言，前两年发生的一场有关数字人文文学研究的学术争论中，美国圣母大学的笄章难（Nan Z. Da）教授重提语言修辞和阐释的复杂性，以基本的统计学为原则，从实证层面讨论了所谓的计算文学研究中较为典型的学术研究成果，通过指出其中许多的技术问题、逻辑错误和概念的缺陷，论证了计算文学研究的方法和理论前提并不适合用来分析文学史和语言学的复杂性。这场反对和捍卫数字人文文学研究的辩论还在进行中，希望能对数字人文文学研究的发展起到推动作用。需要强调的是，如果没有人文学科研究方法和学科史方面系统的训练，而只通过统计方法和计算技术进行炫技般的学术表演，那也许并不符合基本的学术研究标准，更遑论好的数字人文学术研究。

第四是数字方法和技能。数字人文对于从事数字人文研究的学者提出了新的技能要求，包括掌握一些编程或者统计学的知识。所以在理想的情况下，数字人文研究者应该学会编程语言，这是一种脚本编写语言的素养。在脚本编写的时候，某个概念也会逐渐成型，这个过程是思索数字人文研究问题很重要的一个侧面。

第五是数据意识。如果认为数字人文就是做数据库，这肯定是对二者都不了解的错误看法。实际上文本与数据都是需要讨论的理论问题，并非不证自明。正如在文学研究中，并不存在原始的文献学意义上的文本一样，也不存在所谓的原始数据。最近清华大学《数字人文》第二期刊登的一篇文章《什么是数字化？——一项考古学研究》作者邦妮·麦（Bonnie Mak）以著名的早期英文图书在线数据库（Early English Books Online, EEBO）为例，来理清数字化信息与底本之间的关系，阐明数字化信息在哪些方面见证了当前对过去的认识，并被应用到相应的学术研究之中。这是一项很有意思的针对数据库问题展开的知识考古学研究，希望能对思考相关的问题有所推动。

第六是反思性。任何一项研究都有局限性，并不存在任何解决了所有问题、具备了上帝视角的研究，并没有无中生有的重大科学研究理论发现。在《数字人文》第三期所刊登的另一篇文章《机器学习与人文视角》中，特德·安德伍德（Ted Underwood）认为机器学习提供的眼光并不具有经验的客观性，也并非机器计算出的东西。相反，是人类建立了算法，算法发现的一切都反映了人类的视角，都反映了算法的创造者的视角。机器只是简单地帮助我们看得更远更广，以一种新的、令人惊讶的方式扩大我们观察文学的视角。知识的生产是一种渐进的

不断修正的过程。所以数字人文研究应该坦言本身研究的数据基础设施前提、采用的方法和步骤，以及这些所带来的局限性，或者说未来研究可以修正的空间。

第七是开放性与协作性。好的数字人文研究应该具备反思性、技能性和数据的公开性，敞开整个研究过程和步骤，展示可重复的人文研究成果，是一种民主化的知识劳动。同时因其规模庞大，一项数字任务研究常常需要多个研究者合作完成，这就需要各个成员放弃自己的文化偏见，具备开放与协作的特性，这也是新的学术人格与知识伦理产生的过程。

以上所罗列的七点，从我本人所属学科等方面出发，只是抛砖引玉，未必全都合宜。期待更多学者积极参与相关讨论与实践，共同审慎而乐观地推动健康的数字人文研究，同时也促进人文学术整体的良性发展。

桑海（澳门理工学院）：数字人文与学术生态重构

数字人文犹如一种被投进池塘的新鱼，它能否活下去，能否活得好，能对这个池塘的生态产生怎样的影响，都是值得关心的问题。对当前学术生态不满意，甚至认为存在某种程度的生态危机，是许多学术界人士的共同感受。从系统的角度看，反馈是一个关键因素，而学术系统的反馈枢纽集中在评价环节，评价会直接影响到学术生产、学术出版、学术资源分配等。因而如果学术生态失调，很可能是评价出了问题。

讨论数字人文与评价的关系，至少包含这样两个层面：

一是对于数字人文研究成果如何评价。

（一）数字人文的发展对评价提出了新要求

数字人文是一种新的研究形态，呈现与一些现有学术评价体系不相兼容的新特点（比如观念和与研究范式与传统研究差异较大，在论文评审、发表、摘引、评论等环节处于不利位置；甚至成果形式未必以论文形态呈现，而可能是某种数据库、可视化网站、软件或算法等），因此有改良现有评价标准的诉求，甚至呼唤一种新标准或体系。

从数字人文内部视角看，需要讨论“什么不是数字人文”“什么是好的数字人文”（《澳门理工学报》2019年第4期“笔谈”），将似是而非的伪劣数字人文研究识别出来（建立门槛和规范），并努力建立数字人文研究的基本标准（所谓评价体系），为数字人文争得一席之地，证明研究的学术价值。在评价问题的背后，其实是数字人文作为一种新兴的研究范式或准学科，自我定位和获得承认的

渴望和焦虑。

（二）数字人文评价不可能外在于人文学术评价

尽管数字人文有一些明显不同于传统研究的新特点，但数字人文并非“法外之地”，其研究成果一定要得到人文学术界特别是传统学科的接受和认可。这种接受和认可是双向的：一方面是数字人文需要自我调适，对接传统学科自身的学术脉络，真正解决推动学科发展的疑难问题；另一方面是数字人文需要影响传统学科，给人文学术带来新的理念和活力。

二是数字人文对现有学术评价体系可能发生怎样的影响。

（一）现有的学术评价的问题所在

一言以蔽之，评价的现状存在着种种颠倒的状况，至少可从以下几个方面来看：

一是重论文（专著）评价而轻其他。当前学术界存在严重的“唯论文”倾向，不但轻视教学和人才培养，轻视在线发表等，更无法涵盖数据库建设、工具开发等新学术方式，这不但与数字人文桀骜难容，也与人文学科的发展方向——新文科不相适应。

二是在论文评价中，重所发期刊而轻论文本身。于是乎，学术评价与学术期刊评价之间几乎画上了等号。这种“以刊评文”的方式存在严重的逻辑错位，并带来一系列的焦虑。

三是在期刊评价中，重量化评价而轻同行评议。目前的量化评价体系并不完善，还不能取代同行评议。同行评议虽然也存在诸多令人不满意之处，但仍是论文评价和期刊评价不可替代的基础。学术研究的趋势一方面是不不断细化，另一方面是不不断跨学科，不同学科之间评价标准的差异，以及同行概念的日益模糊，使得同行评价也面临很多新的问题。同行评议若想真正复兴，必须与量化评价实现有效的结合。

四是在量化评价内部，重影响因子而轻其他指标。在Web2.0时代，以引用行为为基础的影响因子，应当拓展到社交行为中“被获取、传播和利用”的多元化增值目标，因此计量学者在大约十年前提出了替代/补充计量学（Altmetrics）。

五是在影响因子内部，重引用数据数量而轻学科间差异和引用动机。近年来CSSCI开始重视“分类评价”，引用行为背后隐含的不同动机也越来越被关注。

（二）数字人文对评价的可能影响

1. 数字人文作为理念

数字人文对新技术的吸纳和涵容，或许会在某种程度上改变人文学术的一些基本观念，学术评价的内涵和方法都会趋于丰富多元。比如在成果形式上，除

了论文之外，数据库、软件工具、研究过程等的价值，也开始被慢慢意识到，从而有可能以某种方式进入评价的视野。我倾向于认为，数字人文不是学科，而是一种人文学术的范式转变，或许有一天数字人文的内涵会与人文学术的概念趋于一致。

2. 数字人文作为方法或工具

在数字人文学者因评价体系的不友好而感到委屈时，往往忽略了一个事实，即学术评价的主流方法是量化评价，这与高举量化大旗的数字人文有异曲同工之妙。如果数字人文所倡导的“远读”、数据分析、文本挖掘、深度学习等间接阅读和分析是有效的，对量化评价又为何不能有新的期待呢？在弥漫着后人类情绪的学术前景中，与其把赌注押在通过人的改变来恢复同行评议的公信力，也许还不如期待量化评价与同行评议的更好结合。

先说同行评价。从乐观的角度看，数字人文的学术共同体或许正在形成之中，当下国内主要的数字人文从业者，其优势是有新锐的气息，有比较紧密的联系和较好的合作基础。最主要的是，由于这个领域比较新且未学科化，壁垒和包袱比较少，还处在抱团取暖的阶段。当然困难也是很明显的，除了同行评议常被诟病的主观性、人情因素等外，典型的困境还包括：面对包罗万象、涵盖多个学科甚至溢出学科的数字人文成果，评价者往往力不从心；再有就是数字人文方法论与学科传统之间的某种冲突，造成评价差异过大；等等。基于上述困境，我认为解决评价问题，归根到底还要靠新技术。

再说量化评价。在量化评价领域，虽然很多人认为影响因子是“虽不完美但不可替代的评价指标”，但以引文统计为核心的影响因子毕竟太单一了，或许还需要更多的算法来支持，建立更加立体、全面和有效的评价体系。在这个意义上，对影响因子理论和方法的优化，特别是替代/补充计量学（Altmetrics）仍可抱有一定期待。

评价也体现在媒体对学术成果的发表上面。从媒介变革角度看，随着文献储存和出版的数字化，对于学术资源的稀缺性和可获得性的忧虑，逐渐被信息过载带来的烦恼所代替。应对信息过载的策略本质上可归为两类：搜索和过滤。期刊数据库已使搜索变得轻而易举；学术出版则需承担起过滤功能。在学术期刊网络化潮流中，特别值得注意的是“中国高校系列专业期刊”或“域出版”，这是一种将期刊论文筛选后按专业或专题在网络平台发布的模式。目前其正酝酿向“平台型媒体”（Platisher）转型，成为既是有“过滤器”的媒体，也是可以孕育新学术生态的互动开放平台。

学术出版将从传统出版延伸到数字出版、语义出版、数据出版、增强出版等

新出版形态；自觉地将出版与科研相融合，支持数字人文等学术新范式；运用大数据分析、深度学习和自然语言处理等新技术，提供知识和信息的解决方案和个性化服务。

胡恒（中国人民大学清史研究所）：数字人文研究与现行学术评价机制嫁接的可能路径（线上）

数字人文近些年在国内外日益成为一股强劲的学术潮流，吸引越来越多学科的进入与学者们的注意，日渐显示出其旺盛的生命力。但这一新兴方向繁荣发展背后，一系列迥异于传统论著型的数字人文成果无法被纳入现行学术评价机制，成为困扰数字人文研究者的重大难题。越来越多的学者从促进数字人文健康发展的角度，尝试性地发出为数字人文建立独特学术评价机制的呼吁。不仅中国，即使是在数字人文发展较早的英美诸国，也都存在类似的困扰。爱尔兰都柏林三一学院的Poul Holm教授等人在对世界各地数字人文发展调研的基础上，着重将“学术界的奖励机制并未纳入数位出版”单独予以论述，并不忧虑地写道：“许多想要追寻学术生涯的未来DH研究者，可能会看到他们自己被迫要在两个世界中竞争——人文学的数位世界和传统世界——同时要担负在两个世界中都没能出类拔萃的风险。”纽约州立大学布法罗分校的数字人文研究社群专门设立了关于数字人文学术评价的页面，既收集了来自数字人文研究者的声音，也涵盖了来自学术评价部门包括研究机构和雇佣机构的观点。欧美建立了多元化的学术评价机制、存在运行良好的学术共同体及一系列指导数字人文评价指南的地方，数字人文的评价尚且存有争议，对于中国缺乏可靠的同行评价，更多依赖于量化指标的学术评价机制来说，对数字人文进行客观准确且可操作、可持续的学术评价显然更具难度。

然而在我看来，就目前而言，单独建立数字人文学术评价机制还为时过早。

第一，数字人文属于计算机科学与传统人文学科相结合的“交叉学科”，但并不具有唯一性。近些年交叉学科如雨后春笋般出现，这反映出经历数百年现代学科分化之后的新一轮逆转，学科融合的趋势大大增强。2020年8月，国务院学位委员会新增交叉学科为我国第14个学科门类正是一个标志性事件，美国早在2000年前后就在学科专业目录（Classification of Instructional Programs）中设置了交叉学科。如为数字人文特设一套评价机制，则类推起来，其他学科亦可同样强调各自学科特性而产生N种评价机制。学术评价机制之所以成立，恰恰就在于

它要为多元、分散和差异化的学科体系找寻到相对能够为多数学科、学人所接受的评价标准，尽管会对不同学科、不同成果的特殊性予以考量，但普适性仍是最重要的考量要素。

第二，数字人文的内涵非常宽泛。如Brett Bobley所说，“它是一个围绕着技术与人文学术诸多不同活动的伞状概念”，就目前国内数字人文业已参与的学科而言，已涉及到历史、文学、哲学、艺术学、图书馆学、计算机等领域，且其外延正处于剧烈变动时期，充满了不确定性。以数字人文研究学者来源之多元、学科领域跨度之大、话题与方法差异之明显，谁来评价，谁能够评价，都将是一个极其棘手的问题。

第三，从利用数字及定量方法较为成熟的临近学科的经验来看，数字当然对学科发展有所改变，但对各学科关切的基本问题和提问方式恐怕改变还不是太大。项洁教授在清华大学《数字人文》创刊会上提到：“当数字化和定制工具成为‘直升机’，数位学者可以轻松鸟瞰文献全貌后，真正的发现却依然要依赖于学者自己的经验和对问题的深入把握。”事实上，回到各自学科，与笼统、宽泛的“数字人文界”相比，这才可能是真正意义上的“同行”，也才具备学术评价的知识储备。既然要回到各自学科，这就意味着传统的学术评价机制并非完全不能容纳数字人文的评价问题。

在这三个大前提下，目前为数字人文建立单独学术评价机制的条件并不成熟，并且数字人文是否有必要建立单独评价机制，不仅取决于学理本身，也取决于本学科从业人员的规模及业已取得的学术成绩是否足以支撑，从这一角度而言，目前尚不够成熟。因此，就当下数字人文发展而言，如何策略性地、建设性地寻找数字人文发展与现行学术评价机制的衔接，更好地扮演一个“建设者”“融入者”而不是“革命者”的角色，也许是一种更为稳健的策略。学术评价机制亦非一成不变，随着数字人文研究成果的增多和学术影响的扩大，学术评价机制亦会随之调整，只有深度参与和融入，才有逐渐改善的希望，使之朝向有利于数字人文的方向发展。

谈论所谓数字人文发展受到现有学术评价机制的约束，这固然是问题的一面，但问题的另一面在于数字人文真的与现有学术评价机制格格不入，以至要推倒重来吗？现有学术评价机制提供的空间是否已被充分利用了？在我看来，并不尽然，数字人文研究仍然大有可为。

一、发展研究导向型数字人文项目。我们始终要承认单单依靠新技术并不能与新贡献划上等号，数字人文项目的价值与意义绝不在于追求技术的新颖别致。到目前为止，就我熟悉的史学而言，我认为数字人文设问的方式多数依然来自传

统史学研究。2020年4月24日《中国社会科学报》所刊出的《理性看待跨学科研究热潮》一文，在援引《泰晤士报高等教育副刊》对学者的采访中提到：“很多跨学科研究项目看上去很热闹，很容易受到关注，并得到短期资助，但是，如果这些研究不能有所产出，这股热潮终将散去。”当下数字人文引发极大关注，甚至显得“虚火旺盛”之时，这些提醒弥足珍贵。俗云“基础不牢，地动山摇”，在各学科围绕教学与研究资源竞争时，数字人文的发展不能只在概念、理论和方法上打转，而应该在启动数字人文项目之初优先考虑研究性产出问题，要有实实在在的研究成果和数字人文项目，只有这样，数字人文才能在学术界站稳脚跟，也才能在此基础上“拆东墙补西墙”，匀出资源做一些为传统学术评价机制尚无法容纳的“情怀式”的数字人文项目。如何将数字人文项目与现有学术研究无缝对接，既蕴含数字人文的新意，也能适应传统学术研究与发表的需求，就变得非常关键。我非常赞同南京大学梁晨教授在《江海学刊》上发表的《量化数据库：“数字人文”推动历史研究之关键》一文所提出的，将量化数据库的建设与研究作为当下数字人文建设的重点所在。

二、着力建设高质量的数字人文基础设施。这是许多研究者所公认的目前国内数字人文领域的薄弱环节。一个新的学科在兴起之初，理论方法及学科界定的讨论自然有其必要，但过度沉溺于此，显然偏离了方向，应当将主要精力放在基础设施建设上，如此方能行稳致远。目前存在的主要问题是：（1）好的数字人文项目不多；（2）可公开共享的数字人文资源有限，多数被列入基金支持或宣传报道的数字人文项目往往雷声大雨点小，“只听楼梯响，不见人下来”，既不能公开共享，数据质量亦无法检验，这样即使做出一些数字人文研究成果，也无法验证，不符合科学性的要求；（3）不少数字人文项目无法满足各相关学科对数据质量的要求。总体来说，我们对数据库等数字人文项目错误的容忍度偏高。

三、跨学科研究当中，尊重彼此学科最基础的方法论变得很重要。数字人文的优势在于它与人文比更数字，与数字比更人文，但有一个前提：任何研究都需尊重所涉学科最基础的方法论。我们有些遗憾地看到现有一些数字人文研究成果完全忽略人文学科最基础的知识背景与方法论。如果一项研究利用了新的技术手段，但并不尊重所涉学科的基础理论，得出令人惊愕的结论，那么它对数字人文的健康发展所带来的恶劣影响是巨大的。

四、拓展数字人文研究发表平台。一方面，要创办数字人文研究的期刊方阵，并争取进入核心期刊目录，以尊重和融入现有学术评价机制的模式在一定程度上解决数字人文研究评价问题。目前中国人民大学创办了《数字人文研究》期刊，清华大学与中华书局合办了《数字人文》辑刊，华中师范大学创办了《大数

据与中国历史研究》辑刊等，但还不够。另一方面，还可以设立若干数字人文民间基金和奖项，开展数字人文学界与产业界的跨领域合作等。

五、改善数字人文的教学。近些年，随着国家对本科、研究生教学的日渐重视，教学评价在学术评价体系中占据了越来越高的比重，最近教育部又在推进“新文科”建设，这对数字人文来说，正是一个良好的发展契机。据目前所知，南京大学、北京大学、清华大学、中国人民大学、天津大学等都开设了与数字人文相关的课程，且已取得显著成效，有的还获得了一些奖励，应该继续推广。

当然，上述想法主要还是基于当下学术发展大环境而对数字人文健康发展所做的策略性、战术性考量。不同分支领域的数字人文的评价问题也许并不相同，我所谈的主要是基于历史学特别是中国史方向的数字人文评价问题。我相信，学术在发展，一个多元的基于学术贡献的学术评价新体系终将建立，而数字人文的评价问题也将迎刃而解。

申斌(广东省社会科学院历史与孙中山研究所): 荒原野草与“权法”: 数字人文学术评价逻辑的一点浅见(线上)

数字人文的成果形式是全新的，但是数字人文成果由人产生，其学术评价由人进行，评价的后果也由人承担，而人性不是全新的。所以在具体讨论数字人文的学术评价之前，我想回顾一下学术评价的形成逻辑，这有助于我们厘清今天究竟是在什么意义上讨论数字人文的学术评价。

从本质上来说，学术评价的根本意义，在于帮助人类筛选出有价值的研究成果，进行知识积累，以让人生活得更美好。真正有价值的学术成果决不会被永远埋没，不但后来的学者会在追求真理的研究之路上去追随它，而且也必将被运用于现实。学术研究和社会现实的长期发展，会自然筛选出有价值的研究成果。因此，最可靠、最合理的评价者其实是时间。反过来说，就是当下的（共时的）学术评价会存在误差，甚至出现误差的几率很高。无论理科还是文科，无论社会科学还是人文学科，无论古今中外，根据当下认识对很久以前的研究成果进行重评的事例太多了。今天认为好的，四十年后可能就被认为一点价值也没有甚至是误导性的；今天无人留意的甚至被认为差的，四五十年后可能被奉为经典。

但同时代的学术评价是一直存在的，因为这是自然的和必然的。在人文学研究职业化之前，以及已经职业化但从业人数较少的时候，学术评价不是由“研究的行政管理部门”或者“研究者学术共同体”有意识进行的，而是在学者们通过

阅读论著、研究写作、授课、公开和私下交流，自然而然地形成、表述出来的。这种对人的综合学术水平及某方面研究造诣的评价，表现为非正式的“口碑”。在当时，这种口碑在很大程度上是具有公信力，可以服众的。目前被广泛认可和采用的“同行评议”制度就是这种非正式口碑的制度化产物，只不过增加了一些“匿名”的形式以防蔽。

但是，同行评议制度的有效性是有适用限度的。就范围来说，需要限定于小领域内部，因为当代学术研究高度分化，评价所需的知识非常专门；就功能来说，同行评议首先是用来维护学术规范，排除掉“不合格”的研究，其次才是遴选出优秀成果。当成果分属相距较远的领域时，同行评议其实很难从真正学理意义上对其分出高下。

随着学术研究职业化、职业研究者从业人数激增，以及政府对科研投资规模的扩大，基于形式或量化指标的评价出现了。它的诞生不单是管理者的需要，也是科研职业化后的内在要求。量化指标评价依据引文率、权威期刊、核心期刊等几乎不涉及专业研究内容的标准，所以非常适合于行政管理，可以让管理者花最小成本、以最客观的方式来进行跨领域的学术研究投入产出比较评判，这是其外在动因。而学术研究职业化，意味着从业者要靠做研究获取报酬维持生活。于是，不同领域的研究者需要彼此竞争以获得科研经费资助。这时小领域同行评价就失去公信力了。学术共同体也需要寻找一个可以平衡各方利益，可以服众、有公信力的评价办法（其主要功能在于遴选优秀，而非把合格关）。至少看起来更客观的各种量化指标承担了这一角色。今天我们常常批评科研管理部门唯量化指标是从，但千万不要忘记其实这背后也有来自学者内部的需求。量化指标评价虽然存在弊端，但仍是现实中最能服众的一种评价方式。

纯粹学问意义上的学术评价是自然形成的，一直都在进行。我们大概没人否认CHGIS、CBDB、MARKUS、清代粮价资料库、中国基本古籍库、《四库全书》全文电子版在中国史学发展上的重大价值，对它们不足的批评也一直都有。从长远看，我们不需要担心没有评价标准会导致混乱，不利于数字人文发展。时间会自然地把浑水摸鱼的东西淘汰掉，冲洗得毫不留痕迹，大浪淘沙最后留下具有长远学术意义的真金。

所以我们这里讨论的，其实是需不需要一套能被学术行政管理部门认可的评价指标体系。之所以今天各方都很急切地想讨论这个问题，是因为一方面研究者甚至是科研机构在数字人文上投了不少精力、时间、金钱，却因为没有被管理部门认可的指标体系而难以获得相应的认可与回报（薪酬、职称、评奖、人才帽子、基地牌子等），从而不但影响了自己继续投入的热情，也制约了科研经费的

获取与科研人才的吸引,进而不利于数字人文项目及整个方向的发展。另一方面,科研管理部门也开始投入了不少资金建设数据库,但绝大多数数据库建成、项目结项后就成了僵尸数据库,既没有向社会开放,也没有人员和经费去做后期维护、更新,更不必说追踪完善了。从学术知识积累、普及、传播角度看,经费颇有打了水漂的感觉。所以管理部门也希望有个指标体系以便于评价这种新型学术成果,并且探索一种新的资金支持方式,促进数字人文形式的成果积累和完善。

明确了我们需要什么样的评价体系和为什么需要之后,可能有些问题讨论起来就更清楚了。我有以下四点意见。

首先,对于数字人文这样的新兴领域而言,“荒原野草”才是目前最好的状态。有太多的制度诱导,反而可能限制了数字人文潜在的无限可能性。没有利益诱导的时候,可能浑水摸鱼的人还少一些,完全从纯学术兴趣出发的人更多些。真正意义上的学术创新,往往是几个人出于好玩的心态做出来的。凡是可以被体系化、计划化、规范管理的,都只能是对有着充分确定性的科学原理的工程性应用。就我所知,比如厦门大学郑振满老师的莆田历史人文地理信息系统,就是花了长达二十多年时间完成的,其间没有直接的支持。目前大陆史学界做数字人文最投入的学者如王涛、潘威等都在此过程中,靠着数字人文产品以外的、传统的论文等成果(当然与数字人文有密切关系)获得了职业生涯上的成功。所以数字人文在管理部门那里没有被认可的评价体系,虽然会给数字人文发展带来很多麻烦(最主要的是对大型数字人文项目缺乏可持续的经费支持),但从长远看并不会真正阻碍数字人文的发展。

其次,目前开始思考建设一套能被学术行政管理部门认可的评价指标体系也是有意义的,只是作为从事数字人文的研究者,一定要很清楚这是一个为了让数字人文在现有科研体制下活得好一点的“权法”,不能太当真,以为能设计出一套能很好地评价数字人文工作及产品价值的标准体系。评价标准的设计思路一定不能是“如何评出最好的数字人文成果”,而是如何“否定掉那些肯定不合格的数字人文成果”。换句话说,这个评价体系是划底线用的,不是设定“优等生”标杆用的。因为数字人文是新兴领域,我们实在不知道未来哪些研究成果会有多大价值,所以需要尽可能地保持多元开放态度。而一旦设定出一套“优等生”认定标准,那么是在无形中通过利益诱导扼杀其他潜在的可能发展趋向。

再次,关于具体措施,现阶段探索一套完全独立的全新评价标准恐怕是不现实也不必要的,扩充既有的不同科研成果评价体系,把数字人文成果分类纳入相近的成果评价体系更为可行。比如数字期刊、数字博客,可以在期刊、文章评价

中给它们一点空间；可视化网站可以参照博物馆、展览等评价；数据库、程序其实提供的是一个研究平台和科研工具，可以参考理科实验室和试验技术方法的评价，等等。对于需要持续的资金、人力支持的大型数字人文项目，则可以参照类似国家实验室的管理办法，根据其数据、功能的扩展，用户使用情况和体验，定期动态考评。

最后，关于评价中的回归学科与跨学科的矛盾，从宏观看，这是个时代过渡性问题。目前的数字人文还只是数字技术在传统人文研究中的应用，除传媒研究外，多数还是把前数字时代各种各样的非数字的信息人为数字化以后，再开发各类平台、工具进行研究。在这种状态下，其实我们要评价的数字人文成果，绝大多数还是可以、甚至必须放回各自的学科脉络才能评价的。可是再过三四十年，人文社会科学首要研究对象是数字化生存的人和人的数字活动，处理的主要是原生数字信息，恐怕那时整个人文学科、社会科学的定义可能都会有很大改变。我们今天大概无法预见这将是怎样的改变，那时所有学科可能都数字人文化了，也可能都消融成为全新的数字时代的人文学；或许未来人文学依赖的主要会是像BAT这些公司的技术和平台，研究资助者也可能更多元。所以，对未来保持开放，可能也是以“权法”看待评价标准的题中之义。

李江（南京大学信息管理学院）：科学学视角的学术评价

我题目中的“科学学”是一个术语，Science of Science，是一个新兴的研究方向。我跟大家分享五个问题。第一个问题，学术评价指标的使用准则。下图来源于1978年我们这个领域的一个专业期刊*Scientometrics*第一期的一篇文章，作者是Francis Narin。在过去几十年里，我们一直认为这张图是使用学术评价指标的准则和指南。

图里指标的分布，右侧轴是客观程度，上边的轴是接近于被评价对象真实状态的程度。我们评价的目的是揭示被评价对象的真实状态。指标的分布，最右侧的是加权发文量，被引次数，高被引论文，一直到左上角那边的是非结构化的访谈或者叫同行评议。越靠右侧的是越客观的，都是定量指标。而越靠上方是越主观的，同时它是越接近于被评价对象的真实状态，比定量的指标更接近于我们想要的东西，但是这有个前提，这是一个规范的同行评议。有人会问，这两者一看就不能兼得。我们有没有可能用中庸的思想去选一个折中的方案，选择兼顾客观和接近于被评价对象真实状态的两个指标的方案，比如右上方的这些指标可不可

以？这从来不是学术评价追求的目标。

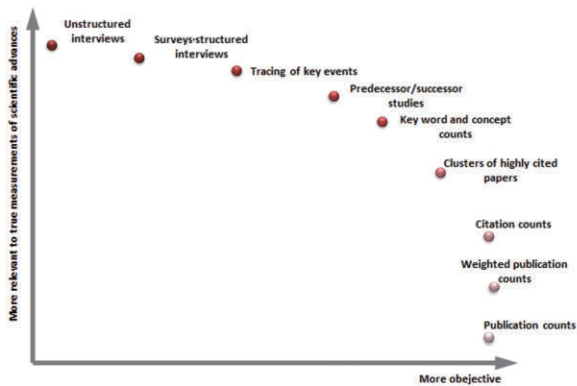
第二个问题，同行评议。1665年，英国皇家学会创办了一个期刊叫做《哲学汇刊》。我们通常认为这是科学领域的第一本学术期刊。这个期刊创刊的时候建立了一套来稿登记、同行评议发布和荐稿

的制度，这是所有学术期刊现在沿用的一套制度。最初的同行评议制度跟我们现在看到的不一样，一直到一百多年后，大概在1770年代才开始有了规范的同行评议。我们对同行评议的认知通常来自于自己的经历，如果我们的项目和论文在送给同行去审的时候遭遇了不公，我们对于同行评议是持有偏见的，会有各种各样的情绪，导致我们错误地认为同行评议不是一种合理的学术评价机制。但事实上同行评议是最接近于被评价对象真实状态的。但是从这个图上来看，同行评议就一定是最好的指标了吗？它不是完美的。

第一个缺陷，从文献上来看，前人的研究表达了这样的观点，同行评议能识别出差的，但识别不出好的，同行评议经常拒绝开创性贡献。举个例子，阿克洛夫的一篇文章《柠檬市场理论》初期投稿时频频被拒，到第三次投稿的时候，审稿人居然有点调侃地说，如果这个观点是对的，那么经济学就要改写。然而阿克洛夫2001年获得了诺贝尔经济学奖。

同行评议第二个不完美的地方，在于不靠谱的同行评议让三分之一的基金评审处于随机分配的状态。John Ioannidis提出，项目不要申报，改为按人分配经费。按人分配有三种。第一，人均分配，经费预算总额由所有符合资格的人平均分配。第二，随机分配，跟北京汽车摇号一样。第三，按人以往的研究业绩分配。这三个方案都不是按项目申报的。这很不靠谱吗？不一定，我做了一个测算，自科基金项目的“中奖率”大概是20%，按照4年连续申报，不规定两年申报不了要取消一年的资格，那平均报4年也能中一次，中一个项目能管4年，所以这个概率其实也差不多，还可以完美地回避一些非常关键的一直被诟病的问题，如性别歧视、年龄问题、关系问题等，甚至还能极大减轻同行评议的评审负担，所以确实有机构在这么干。

同行评议也遇到了一些危机。如今的发表论文数量在大幅增长，三十年前发



表的国际论文大概一年几千篇，现在一年是二十万篇。但中国的审稿人有多少？这么多稿子来了，可是审稿人没有大量增加，同行评议的危机就在于少数审稿人承担了极大的审稿任务，导致审稿质量下降。所以尽管我们认为同行评议是最接近被评价对象真实状态的评价方案，可是它依然有很多不完美的地方。

第三个问题，定量评价指标。刚才在Francis的图里定量评价指标都在靠右侧，非常客观。可是离被评价对象的真实状态比较遥远，但是它受欢迎的一个很重要原因是，在一个关系社会里我们通常会倾向于用一个相对客观的能够说服更多人的方案来做学术评价。所以在过去30年，我们做了大量的定量评价。经济学里面有一个定律，叫古德哈特定律，是一项社会指标或者经济指标，一旦成为一个用以指引宏观政策制定的既定目标，那么该指标就会丧失它原本具有的信息价值，这个指标就是GDP。我们所有的指标跟GDP现在的处境是一模一样的。我们可以找到一个指标来替代期刊影响因子吗？替代不了，它只是不一样，谁也替代不了谁。

2012年，有一个《旧金山宣言》，它的内容比较多，其总的建议是禁止使用基于期刊的定量指标，如以期刊影响因子作为评价单篇论文质量的替代指标，来衡量科学家个人的研究贡献，以及在人员聘用晋升或资助决策中直接使用。截止到现在，有2,035个机构在上面签名，放弃使用期刊影响因子，还有一万多个个人签名。第二，《莱顿宣言》提出了学术评价的十个准则，最重要的一个是量化的评估应该支持而非取代专家评审。第三个宣言是2019年下半年欧洲的一些国家发起的一个倡议，叫做《关于学术交流语言多样化的赫尔辛基倡议》。它最重要的一个观点是在科研评价与基金资助中，要促进语言多样性，保护当地语言。

第四个问题是面向真实世界的影响力评价。我们前面看到所有的指标，像publications, citations, journal impact factors，这些都是面向学术界的，这些评价的投票人都来自学术界。所以评价应该面向真实世界，它不仅应该包括学术影响力，还包括社会影响力、经济影响力、政策影响力。随后Gunnar Sivertsen及其团队开始做面向真实世界的学术影响力评价的研究。

第五个问题，关于数字人文的学术评价，这似乎应该是重点。但遗憾的是，我觉得我不懂数字人文，我也不应该对数字人文的学术评价指手画脚，所以我把这一块留白了，希望大家如果有问题的话一起讨论。

裴雷（南京大学信息管理学院）：数据引证与评价：方法、工具与挑战

我是图书情报专业的，学术训练就是从学术交流开始的，我们学的是米哈伊洛夫的学术交流模型。在图书情报的学科视野中，学术评价是学术交流系统里面的一种正式的反馈环节，最终是为了促进学术交流。今天我想跟大家交流三个方面的问题。

第一是关于数据成果如何进入学术评价体系。我们也谈了一些可行的路径，最直接或最有效的就是引证体系。如果数据成果能够被引证、传播和发现，那么这样的成果事实上能够进入到学术评价体系，虽然不一定能从价值层面对成果进行准确评价，但无疑提供了一类数据成果学术评价的可操作的方式。

第二是图书情报领域在数据引证实践过程中的一些观察和思考。我们为什么要去推动数据引证？最朴素的动机是为了更有效地推动数据重用。数据重用和数据引证是一种可行的成果后评价机制。与之对应，同行评议或者期刊评审是前评价机制，即在发布之前怎么评价它的价值。引证提供了一种简单、不完备、客观度很高，但是不一定接近真实的价值评价的工具体系。

那么，为什么要管理数据？为什么要做数据链接？为什么要做数据的发现？为什么要去重用数据？因为数据本身的价值，它的再利用或者说是被转化利用，能产生新的学术成果。我们做数字人文，不是为了做一个人文项目自己玩的，而是让更多人做出更多的成果。其实抱怨更来自自然科学领域，因为自然科学领域有大量做数据的人，比如陈刚老师做GIS，地理里面有大量的数据，这些数据工作的成果如何去认定？所以在*Nature*上也有一篇文章讲到，我们都在重视谁有闪亮的观点发现，却忽视了真正在做数据的这些人的努力，数据的价值非常高，但是没有得到充分重视。

第三个就是缺乏数据引证的标准形式。现在图书情报学者也在做数据引证研究、软件引证研究或者叫工具引证研究。他们做了很多基础性的工作，但是这些工作要么由于使用者没有引证意识，要么没有引证规范，不会引用，被忽视了，这是我们面临的一个现实问题。有些论文作者也会去做标记，但标记不规范。

总结起来，有这样几种标记的方法。我们做一个静态网页去标记这个数据是出自哪一篇文章或哪一本书，所标记的引文出处作为一条属性信息就代表了一个数据结果。还有一种方案叫OECD方案，它会对数据赋一个唯一的DOI号，这种DOI号事实上是一种数据出版，可以定位到一个数据集产品。此外，还有用URL

的、用网址的、用嵌入式的，以及最难做的动态数据的引证。而数字人文项目很多是交互性的动态项目，不一定能够准确引证到数据。因为数据引证的根本目的，一方面是评价数据的结果，另一方面是要指引使用者能够再次发现这个数据集。如果只给一个DOI号很难达到数据重现或重新发现目的，所以这是一个现实的技术性难题。也有一些行业协会在推动自身的数据引证结构，还有一些数据机构也在做，比如很多数据存储机构会在数据存储过程中生成数据集的句柄，这个句柄具有指向作用，可以形成一种引证结果。我们也在做这种工作，在此基础上也做出了一些基础性的工作，就是有大量的数据集句柄，它构成了一种数据出版平台，或者是一种数据的索引工具，这就是从数据的索引问题到数据索引工具的一个变化。

这里要给大家介绍几个比较有代表性的，比如英国的Digital Curation Center对所生成的数据形成大量的引证链接的一种标识符，DANS也做数据存档的引证标识。此外还有期刊，现在全球共有1,370多种数据期刊，它们不仅要求提供论文成果，而且需要提供论文数据集。这些论文数据集会形成大量再传播机制，这都是一些有意义的探索。一些大型数据出版集团也做了一些引证工具，比如汤姆森路透的DCI (Data Citation Index)。在数据引证这个环节，通过这样的处理，虽然不完备，但是已经基本实现了数据的发现和一定层面的评价功能。而且数据科学和学术出版领域的学者会有一些相应的评价分析方法介入，可以形成一种新的知识再发现和再挖掘。

怎样推动数据引证真正进入学术交流体系，不论是世界论文索引标准里，还是中国的论文索引标准里，一定要有数据引证标准格式和标准规范的出现。对此我们呼吁了很久，但是数据集的引证标准规范还没有出来。我们也去做了一些跟踪，比如谷歌数据检索，在技术层面是可以实现唯一标识符的，但是这种唯一标识符如果不符合学术交流的标准规范的话，其实也是一个无效的加工和学术传播体系，现在的学术期刊很有可能觉得引证不规范就会把它拒掉，这是一个问题。

第三，引证也存在一些难点。第一是在技术层面上，比如动态数据、交互性数据的引证和发现问题。第二是引证标准还不能够完全作为一种事实性标准，存在于现有的学术交流系统里。所以这造成了研究成果没有被评价或被重视，比如在数据引证上还有一个共同作者的问题——首发作者、汇聚作者、可视化作者等。将来如果出现引证目录特别长，需求文件格式都没法做的情况，应该怎么去标识各个作者的价值？我们一个学生在做一个叫数据出版的创业项目，想用区块链的方式确定学者的数据首发权，这对于引证来讲应该是很好的一个基础。

如果要做数据引证，至少有三个层面的问题需要去研究。第一是数据引证如

何做到规范化。第二是要有大量的实证，经实际研究证明这种评价体系是有效评价。研究者在数据使用过程中最担心的是数据的真实性和来源的权威性问题，如果有一套引证体系能够解决此类问题，对数据的再复用就非常有价值。第三是数据引证行为问题研究。

总之，目前迫切需要解决的问题，第一是标准；第二是共识，行业跟行业之间差异大。武汉大学王晓光老师统计过人文社科重大项目，大多都是文献整理和数据整理，但是成果在哪里？数据整理不开放不服务，它就没有价值的延续和评价。怎样去利用真实的学术价值，究竟是发布发表还是被大家再利用，这是我们需要去推动的。未来需要做三件事情：一是数据引证结构进入国家学术标准，这是最基础的。二是需要一整套让数据流转的数据发现和利用的体系。三是需要大家有一定的共识来共建数据标准，有了数据标准，就可以进行开放科学，用众包的形式和集成的方式做项目，那未来的数字人文的价值可能会更加容易得到大家的传播和接受。

潘威（云南大学历史地理研究所）：数字人文与历史地理信息化结合的探索与困惑（线上）

历史地理信息化已经发展了二十年左右，但是在很多方面，特别是西部地区的研究没有任何改变。我自己做的是一些实际的工作，我的探索和困惑也和我们团队的工作相关。第一个问题就是DH和HGIS如何产生价值。第二是我们团队在探索两者融合中进行了哪些尝试。第三是我们现在的工作中一些想不明白的地方。

首先，无论历史学还是处在相对边缘位置的历史地理学这样一个小领域，学科的边界正在淡化，特别是数字人文浪潮起来以后，创新产生在边缘，这些年能够让大家记住的漂亮研究都是交叉的产物。

第二，历史地理信息化的四个组成部分实际上都需要和DH进行深度融合。比如CBDB中的社交网络和CHGIS的地名数据其实已经初步实现了整合，开创了一些范例，后续的发展很值得期待。能够看到实体地理空间和抽象的社交网络空间，两者的融合比对，让我们重新认识历史上发生的事件。在我自己的研究中，第一，它提取史料中的时空信息更为便捷。历史气候史料，将中国的气候变化过程和整个东亚、整个世界半球范围进行对话，特别是基于文献的历史气候的工作路径和第四纪地质、全球气候变化这些领域用的数据进行对话，就必须扩大史料

数据的来源范围,提升数据质量。如何更有效地从大量的史料中提取数据,技术的知识体系就变得非常重要,特别是这种方法要便捷准确,光靠人工的办法来做肯定不行。第二,它使得定位历史地理现象更为高效,有更多的时间进行线上的分析。实际上历史自然地理研究一直到2010年左右,相对于20世纪七八十年代的工作好像进展不大,我们的工作更多就是用软件把过程复原出来,分析做得还很不够。第三,建立更为广泛的环境和各要素之间的关系。以往很多自然地理学者的工作是对曲线,单纯的对曲线就能够反映气候社会的关系吗?我觉得这肯定是不够的,如果两者能够结合的话,对于这个问题的认识,可以分区域、分要素来看,这样能够更好地实现对人际关系的认识。第四,建立跨学科的学术共同体,拓展历史自然地理研究成果的受众范围。我们团队的探索路径是什么?第一,我们现在的工作就是以GIS技术为核心,融合数据库和多种波谱分析手段。我举两个例子,比如说我们的近三百年东亚台风研究尝试用一些DH方法,分析中日韩越这几个国家古代文献中的台风信息,接下来在GIS层面上去重建运动路径,最后做了一个历史台风数据管理系统。第二,就是黄河方面的研究,黄河的有关材料量很大,中文材料就很多,近些年我们又发掘出了很多外文史料,包括英文、日文、法文的一些关于黄河的调查记录、老旧地图和水利图等以往用得很少的材料,在这些材料中,我们十分需要DH技术和HGIS技术的结合。台风研究和黄河变迁研究都有一个空间上的问题,都必须把它放到具体的时空环境下才能去谈它的变迁,才能分析变迁的原因和影响。

我们进行了史料本身的文本结构分析,还做了一些网络分析,不同时段会呈现出不同结构。我们再将这些分析结果用GIS平台展现,这能够辅助理解清代治河档案的特征,也便于在庞杂的史料体系中找到我们最感兴趣、最值得去研究的问题。

但是GIS也存在问题。第一,历史文献当中的地理位置记载实际上是不清楚的,即便有了GIS,也有大量的问题没有解决,需要借助GIS与现在的地理建构空间关系。第二,历史图像和文字材料的处理,特别是两者的结合,需要DH和HGIS两者结合的手段。第三,一部分历史文献中的记录是零碎的、局部的,并不是一种系统的变化,我们今天是用水文学知识、地貌学知识,靠过去的那套黄河记录体系,肯定是对不上的。如果我们要去做一个整体情况时空过程,就特别需要GIS手段辅助,如果是史料把握就特别需要DH手段,然后把两者有机结合起来。

在这个工作中,我们有点困惑。第一,我们本来认为使用了DH技术,资料管理和使用将更为便利,能够将研究人员从繁琐的资料查找工作中解放出来,

但是实践证明，借助DH和GIS等信息化手段的过程中，我们并没有缩短单篇论文的时间成本。由于要对文本进行大量的矢量化的前期处理，工作时间和强度其实明显在延长。同时单篇论文经费成本在不断提高，从2010年的1,500元左右，到现在上涨到7,500元，对经费的依赖特别大，我们就要不停地申报课题，最后会发现原本要做的那个东西究竟在哪里？

第二，在人水关系的长时段演变这个问题上，通过营造历史场景，使研究者能够重新认识史料记载，发现史料背后的故事。我们回顾团队十多年的研究成果，比如我们正在努力构建数字历史和平台，尝试将很多元素整合，但是我们已经得到的认识和理解，好像都不是必须通过这些技术手段才能获得。

第三，在历史地理学视野下讨论环境人文互动关系的深入研究，特别是制度社会工程环境之间多维关系的机制好像没有发挥什么作用。比如现在历史气候学界一个最好的关于气候与社会关系的研究大概就是方修琦老师的模型。可是即便是这个模型实际上也有很多问题，因为在这个体系中，似乎认为气候变化就必然会导致一个什么结果。可是实际上只要我们去做一些实质性的研究，就会发现气候传导到社会中，路径是多样的，其实是有很多的制动阀或者转向阀在发挥作用，可是究竟在什么位置，目前来看仍旧不清楚，那么究竟该怎么做？我们也不清楚，这个工作方案拿不出来。

第四，现在广泛认可的成果都有几个特点，如名校牵头、大型跨学科团队运作、著名学者助推。对于西部地区的工作者来说，我们的团队规模没有影响力，我们的学校平台没有那么高，相对于东部地区来说我们有天然的劣势。在这样的情况下，中小团队怎么跟上这个潮流？广大的一般院校，不是特别有名的大学该怎么办？如果我们不注意中小规模团队和个人信息化工作的培养，这一大部分的力量是不是就要被排斥在DH浪潮之外？

第五，我们真的需要一套数据规范吗？当然规范的数据是有意义的，对有效的学术交流还有数据体系的形成是非常重要的，但是目前来看规范和规则会不会影响小方向、小团队的一些探索性工作？如何平衡规范和个人创新探索之间的关系？

何捷（天津大学建筑学院）：文本空间映射（mapping）的知识组织及其学术潜力

建筑学城乡规划风景园林专业里面的一些工作，其实有好多问题，可以给大家的讨论作一个例子。另外中间也隐含了一个问题，就是空间问题在数字人文里究竟怎么去认识。那天看到《风景园林》杂志上四川大学的同行发表了一篇文章，我们团队也在做类似的东西。它的题目没有数字人文，但是关键词里赫然写着数字人文，也算比较少见的方向的工作。这一年我审阅了很多篇发表在核心期刊《中国园林》《风景园林》的与数字人文相关的论文，无一例外都有比较多的问题，下面这两个是比较典型的。我自己做这个工作其实很难，它好像变成一个很火热的途径了，究竟是真的火热，还是大家觉得这是一个套路，是一个投机的方式，可以跟大家讨论一下。

那篇论文我看不太懂，我就用笨办法自己做了一下，发现问题更多了，这里面有算法的一些问题，还包括古今意义的不明，包括远读和精读的关系。我们想不出来他这个东西是怎么做出来的，假如数字人文是可重复、可追溯、可验证的，这个事就很好解决。我们用很多方法去试验它的数据集都不成功，实验也没法去重现，因为唐诗还是一个比较复杂的事，涉及词汇共现问题、情感分析问题，把这些情感和词汇 map 到关中空间里面有好多问题，没法实现。现在大家喜欢用诗词古文或者其他文学作品来讨论地方景观认知的问题，这在理论上是可以的，但究竟怎么做还是一个比较大的问题。

我们也在尝试。我们正在试着处理诗歌的数据集文本，有关关中的诗歌，原来的文章里分了几大类，我把写景的单独挑出来，最后得出结论：关中就是长安，但凡写到关中的，基本上还是在写长安。我们也把这些共现的问题重新考量，看地点与地点之间共现关系，以不同的视角去看，我们现在也没有完全理清清楚这个逻辑，从数据本体到分析逻辑、方法论，用现代文本的做法去做唐诗文本是完全不搭界的。另外就是 context 的问题，像长安即关中这个问题，我们一开始觉得可以不谈长安，但是长安是跑不了的一个事实。我二十多年前做汉长安的一些历史研究，对这个东西比较熟悉，因为在长期的权力发展过程中，关中的景物变成一个很有象征性意义的东西。它不是真的写景，好多意象都是权力的象征。

我们一直在讨论空间的问题，但空间究竟是什么？空间是不是只是一个地理的问题？空间和地方的关联性，怎么去认识象征问题？我们充分体验到了复

杂性，以前觉得指代是比较明确的事情，但后来发现空间地方性更强的时候，mapping本身可能就不是直接对应的关系。空间尤其是这种地方问题是不是真的可以mapping到一个实体上，这在理论上可能需要进一步讨论，也是空间人文需要去考虑的。我们也做了一些其他项目，比如上海钱庄，就是mapping不了的。空间的网络，mapping里的属性，家族对钱庄的掌握，以及家族里面拥有人和经理人的关系，在空间上是不是真的存在？可能map不是真实的，也可能是一些关联的、意象的或者结构性的。其实可以通过数字方法建立DH知识体系，这可能从另外一个角度更好地发展了数字人文的内涵。

陈刚（南京大学地理与海洋科学学院）：时空GIS基础框架建设的基本概念及规范化问题思考

我围绕目前自己开展的工作，谈一下历史地理信息系统建设的时空数据框架及其规范问题。随着信息技术飞速发展，“数字人文”理念不断深入，促发中国历史地理学及信息化研究领域取得长足进步；其一，涌现出新型研究成果，产生了一批具有重要学术影响力、普惠学林的历史GIS基础数据库及信息系统；其二，推动了学科发展，产生新的学科增长点，壮大了学科研究队伍；其三，在大数据时代，历史地理学领域开始革新研究方法与研究理念，开始产生新思考和新探索，并促进了跨学科合作与国际学术交流。数字人文本身是一个跨领域的松散学术联盟，比如我们南京大学几位从事数字人文研究的老师，虽然来自不同学科，但我们有共同点，即认为数字技术可以帮助大家在各自研究领域提出新问题、开展新研究与产出新成果。

同时，以GIS为工具的数字人文研究得到学界关注，这与近年来在人文社科领域，以空间为维度或导向的学术研究的大量涌现密不可分。在历史学领域，Knowles（2000）率先阐述了历史学研究中“空间转向”（Spatial Turn）的意义与历史地理信息系统的应用，包弼德（Peter K. Bol）指出历史学已进入“空间转向”时期，空间成为理解历史过往之现代思考的核心词汇。台湾“中央研究院”（以下简称“中研院”）王汎森院士（2014）也说：“在GIS工具的帮助下，历史研究者可以从空间的角度去思考问题。”

近年来，台湾“中研院”范毅军研究员提出倡议，藉由历史地理信息系统（Historical GIS，简称HGIS）建设和推进数字人文领域的广泛合作，研建虚拟时空框架，进而发展空间人文，“重回历史现场”的学术倡议与构想，进而提出

“Geo-Humanities(Spatial Humanities)=HGIS + Digital Humanities”的基本看法。

我赞同范先生的意见,并认为空间人文研究的核心是发展新一代的历史地理信息系统,集成多源时空数据,构建时空GIS基础框架(平台)。2018年,由包弼德教授主持,在上海哈佛中心举办“International Conference on Cyberinfrastructure for Historical China Studies”,其目的就是倡导学界为中国研究构建统一的数字基础设施,提供一个可共享可交互的数据平台,包括数据库、数据工具和研究理念等。

事实上,目前在历史GIS领域产生了很多数据库成果,但是这些数据成果可能存在以下问题:大多缺乏统一规划,缺乏统一的数据设计,数据冗余大、数据错漏多、数据更新与维护难,数据交换与共享困难、形成“信息孤岛”,多数数据库随着软硬件环境变化,难以进一步升级与维护,以及存在重复建设等问题;一些数据库的生命周期短,开放度不够、利用率不高,特别是随着项目任务结束,缺乏后期维护与升级,成为无人问津的死库。因此,构建基于统一时空基准的历史地理基础信息平台(时空GIS基础框架),成为新时期历史地理信息化建设的重要问题。

在这一方面,我们不仅看到了复旦大学中国历史地理研究中心的努力,也看到了台湾“中研院”、哈佛大学等研究机构所取得的重要成就。随着互联网及云计算技术的成熟,基于时空GIS基础框架理念研制新一代HGIS已成为重要发展方向。SinicaView是台湾“中研院”研发的新一代HGIS,基于统一时空架构的4D GIS平台,构建了兼具社会服务与跨领域学术研究应用的综合型历史地理信息服务平台。哈佛大学地理分析中心建立的数字化合作研究平台——WorldMap,基于亚马逊云环境,运用开源和开放存取模式,便利了学者们在互联网上使用基础地理信息服务,进行历史地理信息的可视化查询、数据存取、合作共享和在线制图。

在我看来,新一代HGIS的核心是时空GIS基础框架(STGI)。STGI是多源、多媒体历史地理数据的软硬件集成环境/平台,同时也包括数据获取、加工、分析、交换及Web服务所涉及的标准、技术、设施、机制等的总称。这个框架由几大部分组成,第一个是数据库,第二个是数据目录和交换体系,第三个是在此系统构建的公共服务体系,以及服务于整个信息共享服务及应用的标准体系。总结STGI的特征,包括统一时空基准、统一时空数据模型、统一数据分类与编码、统一元数据与交换标准、统一Web信息服务、统一用户环境和统一工具集。说起来好像比较轻松,但实际上是一件比较复杂的系统性工程。它的意义在于弥补数字鸿沟,消除信息孤岛,促进学术公平,提升研究水平和推进跨界融合。题外

话,我认为跨界是非常难的事情,对学者而言,风险很大,这就妨碍了数字人文的健康发展。在我看来,数学人文发展,需要融入各个学科,特别是融入到各学科的方法论、学术评价体系里面,而不是构建一个独立的数字人文学科或体系。

当然,构建时空GIS基础框架也存在着一定挑战,我以历史GIS研究为例来说明。首先,历史地理信息往往是包括“人、事、时、地、物”五要素的时空整合信息,具有时空模型复杂、历史地理要素空间定位难、时空精度差、信息不确定性强、数据多源异构等多方面问题。其次,历史地理数据主要来自古文献及古地图的考证、解读、定量与定位,数据的可获得性、完整性、可靠性、歧义性、时段性等与现代地理数据存在很大差异,从而对历史地理信息的时空定位提出新的难题,历史地理数据的不确定性与时空尺度问题、古今地理数据的时空配准与整合应用等已经成为历史地理信息化研究的基础科学问题。

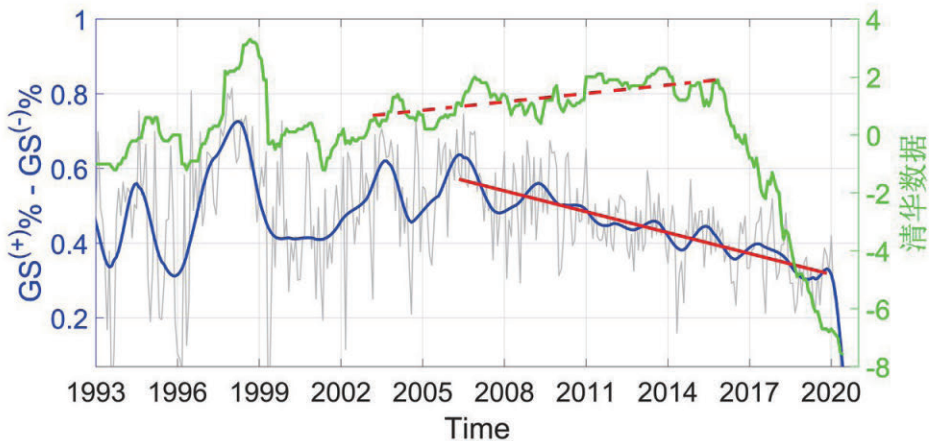
今天报告的第二部分,简单汇报一下我们过去的实际工作。我们团队长期关注南京历史地理研究,特别是六朝建康历史地理信息化研究,现在展示的“六朝建康历史地理信息系统”,在2014年就大体研制完成。它的研究意义在于把南京六朝时期的城市空间用GIS技术复原或再现出来。众所周知,建康是一座消失在历史烟尘中的美丽古都,留存至今的大体只有有限的历史文献及考古挖掘材料,这样的研究存在困难与挑战。我们团队最早是从历史文献与考古资料入手,从《建康实录》数字化开始,同时收集与整理了大量历史文献、近现代以来的南京城市六朝考古发掘与研究文献300多篇(部)。在此基础上,结合正史、地方史志及其他学术文献等的比勘分析,借助野外踏勘及ArcGIS制图技术,与南京大学历史学院张学锋教授合作,在现代城市电子地图基础上编绘六朝建康城市复原图,作为六朝建康历史地理研究的基础工作底图。同时,考证与定位六朝考古遗址(墓葬、石窟、石刻等遗存)约250余处,按不同时代及遗址类型进行分类表示,绘制了六朝建康城市考古遗址专题地图,进而研制专题数据库与六朝建康历史地理信息化展示平台,相关成果获得学界一定好评。此外,我们还利用无人机、GPS等设备,对南京周边六朝石刻进行野外数据采集与三维重建,构建了六朝陵墓石刻三维展示系统。近几年来,我们实验性地开展了六朝人物GIS建设,这是在CBDB项目的影 响下开始的,是基于统一时空GIS基础框架来研制的,并以《高僧传》文本为基础,做了文本数字采集、时空数据模型设计、数据建库与GIS可视化等,当然这是一个探索性工作,还有待在学界的批评与指导下开展深入研究。

高剑波（北京师范大学地理科学学部）：数字人文的计算及可能带来的影响

准确定义数字人文其实很困难。我个人的理解是，任何信息在计算机里面都是数字，人文进入计算机中就变成了数字人文。我觉得计算本身以及计算结果在数字人文的评价体系里应该有很重要的位置。假如是纯粹的学者个人研究，不在意如何评价，他可以做任何事情。但是作为学术圈子还是要有所选择，所以评价体系需要考虑这个领域到底可以干点什么事情。

首先是数据。我个人对做数据库很敬佩，但并不是任何一个数据库都有价值，所以还需好好考察。另外一个功能就是教化大众，在美国好的大学会有口号，说它要影响整个城市，但在中国，北大、清华、南大好像不把它当作是最重要的任务——我们要影响社会。但是我觉得数字人文还是需要把提升大众的人文艺术修养和素质作为主要功能。第三是影响其他领域和产业链。第四是能够帮助解决重大实际问题。数据库到底应怎么评价，一个很好的例子就是CBDB。它刚出来的时候是我的博士后研究里面的数据，通过绘制唐代各个时期的网络图，去看它里面的关系，比如说红色的是Kinship随时间变化，蓝色的中央网络大小随时间变化，王朝繁荣时期中下层往往可以上升到中央网络，所以那个时期政治文化名人辈出。到后期王朝衰败就融不进去了。

这些都表示这个数据库挺好，但是它到底是不是够好，或者说有没有改进余地？其实这个数据库是严重缺失很多信息的。这几乎是数字人文里最好的数据库，还有这么大的缺陷，其他数据库到底会怎么样？所以我觉得假如要评价一个数据库的质量，还是要看能不能做计算。要是一算，发现没什么东西，就要打很多折扣。但是不光是评价，怎样改进才是更重要的。



这是用一个清华大学的国际关系数据库（中国与大国关系数据库）外事互访数据构造的。通过这个数据库可以发现中美关系的变化，一直上升，直到2016年突然下降，但是要是用全球的国际新闻数据库会发现其实美国金融危机以后两国关系就已经变坏了。2018年以后贸易战关系应该是变差，但是清华数据库显示的趋势（黑色曲线部分）反而上升。说明这一清华数据库其实并不能用来反应双边关系走向。他们花了很多时间心血，但还是不行，甚至连趋势都反了。

另外一个问题是艺术性的问题。比如之前提到的树状家族谱系项目成果，获得了计算机领域顶会艺术类论文最佳论文奖，但不能把它当作是对历史的精确描述，因为它的精确性是基于数据库的好坏，数据库本身就有很大缺失，文章用技术展示的不能被当做是历史表述。

前一段时间Lazer等人在*Science*上发表了一篇文章《计算社会科学的困境和挑战——五点建议》，提到计算社会科学的最大困境就是海量论文虽已发表，但尚未能上升至科学的高度，而只是用了数据和方法。他们强调可以从五个方面提升，包括加强协作、完善新型数据基础设施、注重伦理、法律与社会影响、重组大学组织结构、解决实际问题。第五点更重要，应该强调帮助解决重大实际问题，影响产业界，帮助创造新理论。混沌和分形理论是复杂性科学的两大基石。假如数字人文想要产生重大影响，还得好好想想怎么样产生重大理论。

我建议，年轻老师研究历史的时候，最好能够跟研究现实融合起来，光研究历史很难产生重大的现实影响。但是你要是研究现实，没有历史理论则不厚重，所以要把它们结合起来。第二个就是尽可能多利用计算，有新问题以后去找其他领域合作，很容易创造新东西、新方法。但是必须找到愿意听你好好描述问题的人，才能有效合作。

我认为最要紧的是产生重大影响，没有重大影响的话，怎么评价都没有意义。

议程（清华大学人文学院）：数字人文背景下的古典文献新秩序

我的题目在众多题目当中比较特别，数字人文是一个很新的内容，我来讨论它和古典文献之间的关系，也是为了唤起关注。

我想先讲一个故事。我读古典文献专业时，非常崇拜的一位老师要评教授，我想这么厉害的老师肯定没有问题，结果他没评成。这位老师整理的很多古籍不算成果，因为文科处的老师说他不是第一作者，第一作者是古籍的原作者，整理

者是第二作者，所以不算成果。学文献的人都知道这个说法很可笑。原作者早已作古，没法去领成果奖。整理者花了很多功夫，然而在这个评价体系中它是失效的，对于传统文献学界坐十年冷板凳的倡议来说，这是背道而驰的。

今天会上这么多老师呼吁数据公开、算法公开，把数据库和底层数据公开，就像把自己整理的文献拿出来和大家共享，但是对我而言整理文献连成果都不算，出版还得自己掏钱，还不一定能申请上资助的情况下，我把它拿出来共享，写了一篇文章之后，它就变成了公共资源，那么对于学者的劳动来说，这是不是有问题的？所以它在研究中带来了一些新问题。古典文献学传统的成果评价就是有问题和困境的。目前困境得到了一定程度上的解决，比如最近有对冷门绝学的鼓励，用一些资金项目和政策引导，来使这些传统的学者看到一点希望。

数字领域也有很多新的问题。1956年达特茅斯会议之后，大家都在思考机器如何能像人一样学习，像人一样产出研究成果，像人一样创作。清华大学孙茂松老师团队开发的九歌作诗系统，输入“数字人文”四个字，让它做一首诗，最后一句是“文场扰扰是何方”。我们就有一个问题，这首诗写得很好，我把它发表了，那么这首诗的知识产权和署名权是九歌还是孙茂松老师，抑或孙茂松老师团队？九歌其实是学习了过去30万首古典诗词，然后才能够作这些诗，那么它是不是集体创作？这是一个疑问。放到我们学者身上，学者的知识产权是需要被尊重的。机器的知识产权问题也涉及到人、团队集体劳动的产权认定的问题。另外在协作的成本上，多学科共建的成果在不同学科当中，有的是SCI，有的是CSSCI，有的是A刊。或者我们最后也没作出一篇论文来，什么都没发出来，我们做了一个数据库，这个数据库可不可以用来评教授，可以给谁评教授？谁算第一作者，谁算第二作者？这听起来很功利，今天大家做数字人文的人都是有很多情怀热忱在支撑，但是在座的有这么多年学者、未来学者和同学们希望沿着这条路走下去，我们的前辈指出一个方向，我们努力做，为大家趟出一条路，然后后面的同学们终于可以用它做毕业论文、学位论文了，可以被承认了。如果这样的话，就是一个很好的尝试。

我们就来看看评价有没有可能性。我站在古典文献和自己做过的数字人文研究仅有的一点交叉的基础上，提出一点看法。我刚才说的整理不算成果，也引起了很大争议。其实古典文献的笺、注、疏等工作一向都被认为非常重要。如果能够被重新肯定，其实是传统的复苏，是非常有意义的回溯。放到数字人文里面，连笺、注、疏等工作都可以被认定了，那么数据库整理、文献数据化、基础设施建设，是不是也可以被认定？另外交叉学科领域中，统计学、计算机学方面有很多专利、算法工具、数据成果，都可以作为参与贡献认定。

另外，很多老师提到同行评议不好，但是我想外行评议肯定更不好。有一些已经成熟的工具平台的数据库已经相当成熟，很多商业机构都拿它卖钱了，那么学者能不能够拿它出成果？现在理工科都讲学科转化，我们文科是不是有转化问题？是不是可以把它转化成申请项目的基础、职称或者评价的基础。

关于基于问题和成果的领域内对话，即哪个学科的问题，交给哪个学科去讨论。我的一篇拙作研究清代女性作家之间的人物交流网络，讨论清代人物之间的交流，这是一个非常传统的问题。虽然我使用了数字的方法，但解决的是传统学科的问题，我就进入传统学科的评价体系，将其发表在传统学科的期刊，接受传统学者的评议。

另外还想提一点，刚才几位老师提到清华大学和中华书局合办的《数字人文》，在座的大部分老师都是我们期刊共同的哺育者，是编委或重要的编辑力量、支持者和审稿人。我们找相关的学术委员来评审，交叉学科则找两三位委员，共同看同一个问题。2020年是《数字人文》的创刊年，现在第四期也编完了。其实这里面所有人面临的问题和我们今天讨论的一样，没有评价机制，也没有很好的薪酬机制，都在义务劳动，但大家仍然怀着一腔热情在工作帮助我们。所以我们尽所能找有资格的审稿人，即使发表不了也能够给作者一个交代，告诉作者如何做相关的研究，给年轻作者以支持。我们有一个“未来学者”专栏，都是学生投稿，我们也很愿意给他们反馈，即使不能在期刊发表，但至少可以得到激励。

我的核心的观点是，评价的基础是我们做的这些事在新材料、新视角和新方法方面，对学界和学术研究，甚至对将来的学科发展，包括新文科，是不是有价值。如果我们获得评价，可以证明我们是有价值的，当然没有评价是不是就没有价值，它可能是超前的未来的价值，但我还是认为如果要尊重学术劳动和这份热情，很重要的是评价和认可。它可能是多方面的，一方面首先有一群先行者可能最初没有获得评价，但是今天成为业内人士之后，我们是有权同行评价或者给予支持的。另外方法和工具比如数据库已经非常成熟了，为什么学者不能拿它作为一个成果？最后我想说的是，学术评价是对学者的尊重，学术工作应该以好奇心为导向，不是以评价为导向。

邱伟云（山东大学历史文化学院）：术语、图像、规模——论数字人文研究成果的书写范式转移

我今天的标题是术语、图像和规模，会说到遇到的一些问题或困难。我最早做数字人文研究是在2010年。在项洁老师主持的数位典藏与数位人文研讨会上，我发表了以清季外交史料为例做关键词的研究，当时我接受到的评议是：强调统计数字的新写作风格，使得历史人物与故事消失。有评议人认为，我的报告与王汎森演讲形成强烈对比，开发历史研究新领域需要抛弃悠久传统的历史叙事写作风格吗？现在这个问题有逐步消减的趋势。我最近有一篇文章，得到的评议友善了很多：作者常年在数字人文领域生根结果，不只长于使用资料库检索，而且活用数字资讯工具进行截取计算，从而进行推论。

对于新的研究方式，在文哲研究领域已经比十年前更为宽容，这是一件非常开心的事情。我的博士论文原本是想用数字方法做平等概念研究，但是那时我的导师说，你用数字方法做平等观念研究，出来找不到教职，所以我就用观念史的研究方法探讨平等观念。2013年写完博论，2015年出版。我在2013年有篇文章用数字方法研究平等观念，但是这篇文章2018年才发表，因为评价相对友善了，所以我觉得数字人文的环境是在变好的。

我归结为三个变化。第一是同行的形成。在十年前可能没有所谓的DH学者，所以我们的评审可能是历史学家或者是计算机学家，但是到2020年，已经有DH学者出现，有同行产生。第二是环境的友好。我记得项洁老师讲过一个故事，他2009年在台湾大学办第一届数位典藏与数位人文研讨会的时候，常常去跟中文系的老师聊天，中文系的老师不太理他，说我不要这个东西。但现在很多中文学界或历史学界的老师会主动去问项洁老师我们能怎么合作，环境在变友好。第三是可期的未来。现在有了数字人文刊物，很多期刊都有数字人文专栏出现。过去我们当学生的时候，找不到C刊或者重要期刊发表数字人文的论文，但现在有了，未来学者们都可以在这个领域放心地生根了，因为有园地可以让你们发展。

纸本形态是印刷时代的学术规范，未来的数字时代，其形态必须是一种电子媒体，但是在进入电子媒体评价成果前的一个过渡点，我们要如何让纸本期刊的评价体系稍微弹性一点，稍微能够接受数字人文一点，即需探索过渡时期的一些可以调整的方式。

第一个是在过去的文本探勘的研究中，常常遇到术语概念的问题。计算机学者审论文，会问概念史是什么意思，语言转向什么意思。让我们解释，我们就

会加很多注脚。其次是计算机学术语。人文学者要我们解释什么叫文本勘查，什么叫数据驱动，什么叫做数字化、数据化或者智能化，我们又必须要对人文学者的审查委员进行解释。统计学的术语，什么叫转折点分析、集群分析，我们也必须要对人文学者或者计算机学者解释。所以在这样的交叉领域下，尤其在过渡阶段，当各领域学者还没有形成一种共同语言的时候，我们要对这些进行解释。但是如果概念变成了数字人文或者是人文研究学界、计算机学界和统计学界能够共分共享的时候，我们就会省略刚才的术语问题。

第二个问题是图像阐释。图像阐释问题很重要，人文学者对于视图能力，长期以来都缺乏培养。葛兆光先生曾经讲过，我们应该更为看重图像领域，图像研究有四个重点：色彩、构图、布局位置、变形。图像学也被用于数字人文研究中的图像可视化。葛兆光先生举例说，图像色彩背后有它的寓意，图像变形也有意义。从这里可以看到数据可视化的重要性，我们要培养识读图像的素养。之前我有一篇论文，插图本来是彩图，每一个颜色代表一个集群，但是编辑部老师说他们没有资源彩色印刷，所以把这张图变成了黑白，整个分群结构就没办法在图像显现。

我在香港中文大学《新亚学报》发表了一篇文章，其中有两张插图，编辑部的老师说页数太多，第一个图可不可以删掉，只放第二个图，因为第二个图是精简的，跟研究主题相关。但是我跟老师解释第一个图也不能删，因为它是一个层层定性、层层精简的过程，要让读者看到总体数据图像。图像在数字人文成果中产生的问题让我们意识到数字人文中的图像数据可视化非常重要。

第三个是规模大小问题，牵扯到阐述详略和图表多寡。这里我想提到两个小问题：第一，如果是一个小题目、一个略微阐释或一个应用少量图表的研究能不能成为一个好的数字人文。第二，如果是大题目，阐释详细，图表很多，是否能够成为印刷期刊的成果评价系统所喜。数字人文最主要特色就是共读方法，它是结合宏观远读跟微观精读，再从宏观脉络中把握问题，又能让学者发挥学养，深刻探究问题，并解释宏观脉络背后意义规律的一种综合方法。

所以在这样的以共读方法为特征的研究范式下，它很难是一个小规模的文章。我们除了需要对数字方法跟数据进行阐释之外，还需要对人文概念进行阐释，这就会使整个数字人文的研究非常繁复，所以在这个情况下，我要归结到最后一个问题，就是数字人文的成果有两行，我们到底要尊重同行的评议，还是编辑部的意见。同行评议会认为你的研究只有图表，人文层次不够好，编辑部的意见说，我们期刊最多字数就是三万，这个五万字论文是否能够精简到三万字左右？在这样的大命题跟一个必须详细阐述的研究范式下，我们如何调和同行评议

认为人文解释、统计解释、计算解释不够，同时纸本期刊希望精简页数或精简图像的意见？

在印刷时代纸本期刊的过渡时期，要解决术语问题、图像规模问题，但当我们到了数字时代电子媒体时期，代表性的例子是2015年斯坦福大学图书馆出版《荒漠生辉》，把整本著作做电子媒体呈现，包含跟主题相关的语音、视频、照片文字，以及GIS都在互动式的媒体里呈现。但是就像斯坦福大学图书馆馆长杨继东先生所说，互动式的著作在经费用完之后就没了，每一本都花了非常多钱。于是我们要去思考，当没有办法立即建立一种新型数字人文成果评价体制的时候，需要扩充既有评价体系而不断创新体系思考，对此我们现在可以进行讨论，或者希望期刊界进行弹性接纳，这样才能帮助数字人文发展。

赵薇（中国社会科学院文学研究所）：计算文学研究——评价还是批评？谁来评价，何时评价？（线上）

数字人文内部的差异太大，无法一概而论，所以我只能说说我熟悉的——和文学研究有关的“计算文学研究”，这样一种被归纳出来的研究形态，或者说“称谓”。我想以笄章难对李友仁一篇研究性论文的质疑为例，来讨论笄章难《以计算的方法反对计算文学研究》，这样一个由貌似领域“内部”的专家提供的颇具迷惑性的批评文本所引发的争议和其实质所在。

在今年《数字人文》第二期李友仁的翻译文章中，作者使用一种非常常见的数据分类办法，即层次聚类的方法，看它能不能把三种文学类型：所谓“明清小说”“野史”和“正史”分开，进而借助于多维度尺度测量（MDS）和主成分分析（PCA），看有可能是什么因素区分了它们。在笄章难看来，这个问题似乎出在了它的数据预处理阶段，即由于作者将每本书做了万字数据块的切分，他的PCA上每个点代表一个万字数据块，而非整本书，所以在比较最常见的1,000个词时，作者已做出极为相似的数据点，使得来自每种体裁的点彼此非常接近。这段指责看上去不仅逻辑连贯而且十分厉害，但是仔细分辨，每一句话都相当模糊，无法确指。比如“极为相似的数据点”，谁和谁相似？后面她贴了两幅对比图，意为凸显：如果用全本的话，就没有那么多数据了吧，这无异于指责研究者故意制造数据。然而，是这样吗？首先这两幅图都来自于李友仁的文章，也就是说，李友仁并不惮于公开全本数据降维后的PCA，相反，正是靠这幅图，他其实强调了小说和正史的区分更加分明了。李友仁文本的预处理方式并非是为

了莫须有的“同质化”数据，而是标准化数据的一个必要步骤。

在进行聚类之前，需要在全部文本水平上选择高频特征，考察它们在各个文本片段中的频率，以此来构造向量空间索引，再通过测量文本向量间的距离，进行降维投射分析，这已经是数据检索和文本相似性研究的通行做法。竺章难挑战这一点是没有太大意义的。因为在这个过程中，无论使用全本还是“万字片段”，第一步都是文本的向量化。由于各个作品体量相差悬殊，为了保证词频计算的科学性，在测欧式距离之前要先做等值切割（即确定window大小），然后才能把各片段所包含的、全部文本范围中的高频特征词的分值计算出来，从而在多维空间里为文本片段定位。这里，批评者并没有给予详细解释，而是轻率地使用了“均质化”这个似是而非的说法，但无论是采用万字片段还是每本全书，最后的结果都是分开了，也就是说，还是有一个分布的大体趋势摆在那里的。那么，切割文本的确是恰好弥补了可视化的数据稀疏问题，但这不能成为批评者的口实。这种放大文体信号的做法在文体测量学中并不始于作者。关键问题是这么做究竟为了什么，不能就此把手段等同于目的，这是十分tricky的地方。

事实上，竺章难要想推翻对李友仁数据这种解释的“科学”方法，是提出更站得住脚的解释，并给予计算和分析上的佐证，但是她无法提供数据例证，更无法提出比作者更有利的对主成分的解释。她只能用文学研究的“定见”来帮助自己：“野史和正史的差异在于内容而非正式语言的使用，而和小说的接近则是因为它们同样来自于道听途说。”李友仁的结论意味深长：“历史修辞的结构影响，仍然是通过语言的正式性（语域）来施加的。”可见研究的重点并不在于不同标签下的文本可以被分开，更在于他们被什么分开，为何沿不同的主成分轴落下。细查plot loading后发现，恰恰是切割后的片段，显示了白话/文言向度分布上的细微差异。这当然是一种阐释，是作者破解算法黑箱的努力。

也许此二人根本分歧还在于，竺章难不认为N元语法的出现频率（或者此类文本表示模型）可以表征这些区分因素，李友仁则不仅相信，还认为在古代汉语中，单字同样是重要的语法和语义单位，这里面体现的其实是对词语和文类、风格之间关系理解的根本分歧；另一方面则是对古代汉语的认识偏差（尽管受限于技术进展，基于一元语法的分词方法仍待商榷。但即便如此，李友仁有自己的认识，参见其与郭英德先生访谈的注释）。遗憾的是，竺章难没能正面出击，而是揪住所谓“均质化”一点，这背后当然是一种批判策略。可以说她基本上还是从自身领域的经验来横加批评的，这种否定自然是在彰显一种等级秩序。

增大语料单位、放大文类信号的做法由来已久，我们也在对民国时期的报刊语料做文类区分和识别，此前也试过相近的方法，虽然没有使用一元语法，没有

切分，使用的是余弦相似度，但并不妨碍我们得出近似的结论。也许，正是由于和他有着近似目的，我们更容易理解为什么要这样做，也很清楚做法出自何处。事实上从斯坦福实验室的量化形式主义，到Matthew Jockers，再到李友仁，人们一直在探索，这里面当然有探索风格、文体、文类理论的意图，然而只有自己做起来，才会知晓其来路，才有可能理解“尝试”的意义和价值。不熟悉这个线索，就无法放在一个类似于“学科史”的脉络里来进行纵深的考察和评估。

最后，我认为有必要逐步建立起真正的同行评价。同行评价的意义究竟是什么？什么样的人才能称为“同行”？笄章难是李友仁的同行吗？应该不是。她看上去是懂技术的文学专家，但不是懂技术的传统学科“权威”就能随意评价数字人文的研究，评价的人也应是实践者，且必须有相关实践经验。实际上，她才是非常厉害的“特洛伊木马”，她的反应正代表了大众对此类研究的一种“常识性”反弹，而公众对这样一个有问题的批评文本的反应，除了再次体现了人文阐释领域对实证研究和统计表征十分有限的认识度，说到底还是由于真正意义上“同行评价”的缺位所致。那么，我们应该理睬她吗？是的，而且必须。但是，时候未到。苏真（Richard Jean So）在《数字人文》创刊会的发言中提到，数字人文在北美已经过了激烈争论阶段，更多的人选择不回应，不争论，不炒作，埋头做事，让成果说话，当越来越多的里程碑式的作品出来了，无谓的争论会失去意义。

在数字人文每一个次分支领域，现阶段真正在踏踏实实做的人太少了。而数字人文又是一个探索性极强的领域，探索的意思是，它可能是试错性报告，像斯坦福小册子的很多研究一样，可能一时间未必能产生什么成熟成果，更不要说“重磅成果”。对于大量认真探索的研究，请“评价体系”和急于评判的人网开一面，给予其必要的时间和耐心。同时，另一方面，对于一些通认的学术规范上的问题，我们又必须严谨起来。在《数字人文》辑刊的编辑过程中，对真正体现了跨学科努力的文章，我们都是非常珍视的，但是也不难看到，越是有创新的、大跨度的研究，越容易面临一些疑问，这些疑问不仅来自编辑，更来自由多学科组成的外审专家们。而归纳起来，最多的疑问可能来自所谓测量效度方面，这也是笄章难很多批评的出发点。但问题是，效度的缺失与否，是可以通过效度检验来评估的，而不是一句话就能够断是非的。总之，我的核心观点是，在这个时刻，逐步建立起真正意义上的同行评议是非常重要的。

陈静（南京大学艺术学院）：数字时代的学术共同体，存在吗？

我的汇报分两部分，第一部分简单说一下在评价中遇到的一些问题。多年来我的研究方向多与数据有关，不少人会好奇你在做什么，我觉得在研究的问题上，应该允许有不同的专业角度介入。所以我想举一个例子，就是我们最近在做的一个事情——中国传统色彩的研究。

我们在做中国传统色彩知识研究的时候发现了一个问题，色彩研究涉及到很多内容，比如文献研究的部分，我们做了明清文献中关于色彩名词的梳理，也用到了数据分析的方法，但是它同时又是一个基于文物和传统手工艺的实验研究。

我们希望把所有跟色彩有关的知识做成一个数据库，有知识图谱，这涉及图情还有非遗理论，还涉及到系统搭建、交互，包括要做一些文本处理和质量控制，我们自己还开发了一个小程序，专门去做人工和计算机算法的准确性评估。同时还在和化学系的老师合作做植物染材的化学分析，但是化学分析同时还要做以前文物残片染料的分析，所以又跟文保复原的老师合作。另外我们要做一些数字化的转化，做数据采集，所以我们又跟摄影和图像学的老师合作，包括国标文件的规则制定，有了图片以后要做色彩管理，进行色彩测量，我们又跟可视化和色彩管理的老师合作。这个项目又要用实际的染材染出来，我们又染了一个月的颜色，还要跟服饰和工艺的老师合作，然后要做图像分析怎么办？里面的纹样怎么办？所以一个项目八个领域，跟不同的团队合作。

但我不知道得出的结果放在哪个学科，如果要发论文，我要把结果拆成八份去发。但如果我想综合性地去评价研究价值，谁来做评价？数字人文是一个大帐篷，不同的学科放在一个帐篷下面，我们怎么来定义它，谁都只能看到它局部的时候，我们怎么从宏观和整体性上来评价。我个人碰到这个问题，而且最大的问题就是直接影响到了我的成果发表和职称评定，我也不知道怎么办，求教大家。

第二部分，我就各位发言内容，汇报一下总体感受。今天我受益良多，做了一个笔记，想跟大家讨论一下。第一个内容其实我们在讨论的是一个生产的问题，中间牵扯到共享的术语、学术的同行、共享的学术边界，这是一整个模块的问题。第二个内容是学术发表、讨论和渠道。数字人文研究需要有固定的渠道，或者临时的场合，像我们这种以会议为形式聚合的一个临时的场合。第三个内容就是有学术成果、期刊、教学公共服务平台，包括数据库。朱剑老师也明确提出来，学术成果不同于学术批评，学术批评在一定程度上是缺席的，学术评价也不等同于学术批评，我们有评价标准，有评价机制，还有一套体制。

总结来说，第一个就是大家最关心的成果问题——成果形式，成果影响力，什么算成果。第二个就是同行是否存在，学术共同体是否存在。接着是同行评议怎么办。我们现在没有同行评议，是因为期刊权力取代了同行评议的存在，导致评价的独立性不存在。我今天感受最大的一点是评价的反馈机制。理想的评价有助于学术发表和发展，但是现在的评价没有。国内同行评议的时候，有多少人是在认真帮别人做学术？

说了这么多大家都很无奈。第一，“法外之地”是没有的。第二，传统领域里面都有很多成果得不到承认，凭什么要求DH得到承认？第三就是无法摆脱现存学术体制。应对方法——我觉得要先有同行，抱团很重要，哪怕是“小同行”；再有评价，这是核心。我们现在都是在讨论，有评价以后能圈出一个圈子来，其实不是的。靠人才会有评价，有了评价才会有标准。有了评价标准，再考虑评价渠道的问题。我反向思考的大概是这么简单的几点。

（编辑：王波）