

## “数字史学”：现状、问题与展望

发布日期：2017-09-13

**摘要：**人文社科学界对数字人文、大数据等概念的兴趣日益浓厚，大家都在热情洋溢地讨论着这些高端话题，但作为理性与自觉的学者，应冷静地做出判断与回应，特别是关于方法论上的思考，甚至一些“破坏性”的反思，将有助于我们在这个研究路径上越走越远。从历史学的角度反思数字人文，可以发现数字史学关注网络的使用，以数据库建设为主，与数字人文存在不同。造成这种状况的原因是数字史学强调“数据驱动”，没有完成史学研究的“论证驱动”任务。在学术研究已经实现“数字化生存”的今天，数字史学的方法与态度，能够为世界史研究提供站到与国际同行同一起跑线的机会。

**关键词：**数字史学；论证驱动；世界史；

人文社科学界对数字人文、大数据等概念的兴趣日益浓厚，大家都在热情洋溢地讨论着这些高端话题，似乎以为数字人文的“金矿”能够带来像20世纪初那样的语言学转向，以“信息转向”的方式将人文学科的学术脉络引到一个全新的天地。严格来讲，数字人文的学术史并不算短，20世纪80年代就出现了“人文计算应用”的概念，致力于用电脑技术研究传统的人文课题，但“数字人文”成为真正获得普遍认可的概念与学术“流派”，应该是2004年之后的事情。这样看来，数字人文又是一个彻底的新生事物，它是否能够带来学术转向的奇效，或者这个转向的到来还遥遥无期，都是一个未知数。

当然，我们不希望数字人文是昙花一现的虚假繁荣。2015年以来，大数据、数字人文成为许多学术会议上的高频词。愈是在这个时刻，我们作为理性与自觉的学者，愈是要冷静地做出判断与回应，特别是在方法论上的思考，甚至一些“破坏性”的反思，将有助于我们在这个研究路径上越走越远。

“数字人文”是一个过于宏观的概念，不同学者有不同的预设。这里实际上存在一个悖论。“数字人文”从字面上看，具有跨学科的本性冲动，它将信息技术投入到人文社科领域，注定了需要不同学科背景的研究者联合攻关。所以，数字人文的方兴未艾，不是在消解传统学术，而是需要同传统学术、传统学科有机融合，共同构建学术新阵地。

在西学的脉络中，数字人文在文学领域的展开更早，影响更大，有学者甚至开始讨论数字人文的“黑暗面”。在数字人文的语境下梳理“数字史学”，或许能够让我们看到更多隐含的信息。

### 数字人文与数字史学

数字史学有数字人文的共性，但其作为历史学的个性特征更值得我们关注。只有在深入了解这种差异后，我们才能够更加清楚“数字史学”能够带来什么样的研究成果，能够切入什么样的研究课题，而不再是泛泛而谈“数字人文”“大数据”的概念。

数字史学重视网络的利用与展示，这也是很早就有学者提出“网络史学”的原因。历史学家利用互联网向同行传递和展示史料，将文献放置到互联网以便史学爱好者、历史教师、学生能够利用它们，这些行为已经被史学界认可为专业的学术活动，成为数字史学的重要内容，是一条迥异于宽泛意义上数字人文的路径。

当然，“要么在线，要么落伍”的宣言并不意味着互联网能够与数字史学画上等号。互联网只是恰好因为平台优势而成为一种媒介而已，不过在这个现象背后有一些值得玩味的问题。数字史学对网络的强调让我们看到了它与“公众史学”存在诸多交集。所以，如何让专业历史知识成为“普通群众能够看得见、听得着、读得懂并能解读的历史”，对此，数字史学有着强烈的引导大众的动力，可望提供多元化的历史教学形式。

数字史学就目前看，研究工作的重点还停留在数据库建设的阶段，比较成熟的研究方法以GIS的运用为主。与“数字文学”相比，文本分析在数字历史领域运用还比较少，这背后有许多客观原因，最主要的因素是史料的数字化程度还比较低，效果也达不到大规模进行文本挖掘的程度。传统的历史研究重视的一手文献往往以档案、手稿等形式出现。这些文献的数字化远没有达到丰富的程度，而且进行数字化使用的OCR技术，还存在诸多缺陷，需要大量的人工校对，其准确度才能达到进行研究的要求。

所以在从事数字文学的专家已经在讨论如何分析文本的时候，数字史学的专家还要首先解决史料数字化的问题。在这样的前提下，数据库建设具有完善基础设施的重要意义。问题在于，数据库建设是非常烧钱的项目，如果没有足够的资金支持，难以做出高质量、有使用价值的数据库。

数字化程度不足的状态，也决定了学者使用数字工具的倾向。一方面这跟历史学科自20世纪70年代以来的“空间转向”紧

密相关，学者们开始采用“边界与疆域”“边疆与路口”“中心与边缘”等概念来解释历史问题。另一方面在于GIS的工具相对成熟，使用的门槛较低，能够切合历史地理的研究课题，容易被传统学者接纳。

## 从数据驱动到论证驱动

在许多场合，当数字史学的项目用漂亮的可视化工具得到呈现，志得意满的讲述者描绘为得到这些数据如何辛苦后，他们往往被追问的问题是：然后呢？这些形象的呈现，说明了什么、解释了什么？是否经得起质疑？类似的问题似乎还没有进入到数字史学专家的研究视野。从史学的专业角度看，许多数字史学的项目，都属于未完成品，它们往往只提供了材料，但对材料缺乏必要的深度解读。

造成这种状态的重要原因有两个：

第一，数字史学的许多参与者是基于对技术的兴趣加入进来的。数字史学仍然有一定门槛，需要学者对数字工具有一定了解，甚至最好能够成为勒华·拉杜里所谓的会编程的历史学家。而且，参加数字人文论坛的专家，有不同的学科背景，他们只能就技术问题展开讨论，在具体研究面向上，可能提不出专业的质疑，无法让数字史学的项目获得专业分析上的精进。

第二，数字人文研究所强调的“数据驱动”可能偏离了史学研究的路径。传统史学研究是论证驱动的，擅长分析问题，而数字史学研究者在数据驱动的思维指导下，花费大量精力去转换、提取、清洗数据，以为数据的获取就是研究本身，显然背离了史学研究的精髓。即便在公众史学领域大展拳脚的数字史学项目，也由于公众史学注重普及的特性，而在问题的论证上存在先天不足，包括像“历史与新媒体中心”参与的项目，也有学术性欠缺的遗憾。然而，只要历史学还要致力于学术问题的分析与解决，或者像有些学者指出的那样，新一代史学家应该更加关注“出思想”，那么数据驱动就远远无法满足历史学的专业需求，数字史学专家若仅停留在“数据”的层面，这个学派的未来就难以为继。

在许多西方学者看来，2008年金融危机的爆发，带来了传统人文学科的衰退，数字人文的概念被认为是解救疲软人文学科的一剂良药。不论这两种现象是否存在内在的关联，我们不希望看到数字史学遭遇20世纪70年代“计量史学”曾经受到的打击。

数字史学的路径要越走越宽，需要从云端走入学术研究的日常。最核心的是，在有针对性的研究课题中，将数据驱动与论证驱动结合起来，以真正的问题意识为指引，而不是为了卖弄数字工具的噱头，才能带来更专业的成果呈现。

## 数字史学给中国世界史研究带来发展机遇

经过几代人的努力，中国的世界史研究已经取得了丰硕成果。但无法否认的是，不论是从研究的深度，还是研究的思想性，优质的世界史研究成果还非常稀缺；特别是与国际同行相较，中国学者的主体性欠缺，在世界史研究领域缺乏公认的代表性成果。但在学术研究已经实现“数字化生存”的今天，数字史学的方法与态度，能够为世界史研究者提供站在与国际同行同一起跑线的机会。

我们对数字史学的直接感受就是，研究者能够获得大量在线数据库，为研究提供了海量文献，获取资料的便捷程度是老一辈世界史学者无法企及的。借助智能化的搜索技术，年轻一代的史学工作者，能够在很短的时间内，达到老一辈史学家经过几十年“上穷碧落下黄泉”的辛苦才能实现的对史料的把握程度。更值得关注的是，数字时代的史料形态也得到了扩充。比如学术数据库中题录的价值。我们利用大数据处理工具，在海量题录中提取了有价值的信息，可以获得对学术史的理解，甚至尝试预测未来研究的趋势，比如题录的信息揭示，在德国史研究领域，历史与记忆的问题会成为热点。

当然，充分而扎实的史料只是提供了一个方面。数字史学的态度还在为我们的研究提供全新的研究思路，从而让世界史的学术研究呈现多元化、跨界化的趋势。

无论如何，随着数字史学的方法与理念在中国学术界不断得到认可，会吸引越来越多的中国学者利用全新的史料与方法进行世界史研究。基于大数据与数字工具的人文研究，有可能改变质性研究过于倚重主观感悟的特征；虽然定性研究包括历史学在内的人文学科最为迷人的地方，但数字史学的方法将让历史学的科学属性得到进一步延伸。我们相信，在不久的将来，当更多从事世界史研究的中国学者主动适应数字史学的态度，使用相关方法，越来越多的原创性研究成果会涌现出来，并且达到与国际同行对话的程度。

(来源：《江海学刊》2017年第2期 作者：王涛)

(作者单位：南京大学历史学院)

