

数字史学项目模式探讨

——以“现代中国城市中的精英、网络与权力”(ENP-China)为例

齐欣欣 / 山东大学文学院

摘要：法国艾克斯-马赛大学的跨学科团队启动了一项名为“现代中国城市中的精英、网络与权力”(ENP-China)的大型项目，旨在运用数字技术研究1830—1949年间中国近代精英的转型过程。项目聚焦精英个体的行为模式，分析他们如何通过建构社会网络来影响社会与政治格局。通过规划缜密的基础建设路径及开发历史文本分析工具等方式，项目团队收集了海量地理和人物数据，建设了开放的数据库平台，并基于这些数据进行了多元研究。该项目在史学视野、数字化成果和基础资源建设三方面展现出自身特色，其建设路径、方法和成果为数字史学项目提供了可参考的范式。

关键词：数字史学 中国近代史 数字人文项目 社会网络分析

引言

数字技术的迅猛发展和向下传播给处理文献材料带来了便利，运用数字技术进行人文研究已然成为国际学术界的重要潮流和趋势，在拥有量化研究传统的历史学领域，这一趋势尤显突出。从早期的简单数据统计，到如今借助GIS、文本挖掘等新技术，“数字”结合传统史学的范式经历了不断更新迭代，使得史学研究边界得以持续拓展。因此，学者逐渐采用“数字史学”这一术语，用以描述数字人文取向的史学研究。在研究内容方面，历史学家也从早期年鉴学派提倡的宏观“社会结构”的分析，转向了“社会网络”等微观世界的精细重构。

数字技术与社会网络理论(Social Network Theory)的联动催生了社会网络

分析 (Social Network Analysis) 这一研究方法, 该方法被广泛运用于历史学者对士绅阶层的研究。士绅, 与欧美学者使用的精英 (Elite) 一词同义, 被视为地方文化的代表, “是封建解体、大一统的专制皇权确立后, 中国传统社会中所特具的一种人物”,^① 在传统中国社会中具有独特的社会地位和功能, 因此该阶层的历史命运向来被研究者所关注。士绅阶层内部交往复杂, 蕴含大量关系型信息, 且相关历史资料详尽, 这使得社会网络分析在士绅研究中具有天然的优势。使用数字方法进行士绅的社会网络分析已成为国内研究的一大趋势, 研究者常基于CBDB数据库, 结合网络分析软件和GIS技术, 对族谱或群体交往情况展开探讨。^② 然而, 目前对精英集团的数字化研究多将关注的时间节点放置在古代, 近代中国转型期间的精英群体却受到忽视。晚清时期, 中国社会经历了“三千年来之大变局”, 不仅旧的士绅集团内部发生了剧烈变迁, 随着通商口岸的开辟, 沿海城市中也出现了新式的城市精英。这一历史变化是中国近代社会现代化转型的透镜, 长久以来备受传统研究者重视, 在当下“数字史学”的研究中亦不应被忽视。

由法国艾克斯-马赛大学负责的研究项目“现代中国城市中的精英、网络与权力”(Elites, Networks and Power in Modern China, ENP-China) 弥补了近代中国精英群体研究在数字史学领域的空白 (官网网址: <https://www.enpchina.eu>, 主页界面参见图1)。该项目计划研究1830年至1949年间中国精英的转型过程。虽然研究时间跨度大且涉及海量历史数据, 但是通过使用数字技术搜集和处理语料, 项目成功地将宏观计划与具体案例相结合, 既实现了数字人文中“远读”和传统研究中“细读”的有机结合, 又通过巧妙设计使得所收集的数据具备可扩展性和延伸性, 属于真正意义上的“创新型”数字人文研究。

本文将对该项目进行深入剖析, 按照“研究计划—建设路径—研究成果”的层次展开介绍, 最后评析项目优势并指出其局限性。希望通过对该项目的分析, 一方面能够丰富对中国近代精英群体和近代中国社会转型的历史认知, 另一方面也为国内数字史学项目的实践提供些许借鉴与启示。

① 费正清、吴晗等:《皇权与绅权 (增补本)》, 上海: 华东师范大学出版社, 2015年, 第1页。

② 钱超峰、杜德斌:《北宋官僚家族网络的空间结构及其演化: 基于CBDB和CHGIS的考察》,《历史地理研究》2019年第2期。刘京臣:《大数据视阈中的明清进士家族研究——以CBDB、中华寻根网为例》,《北京大学学报 (哲学社会科学版)》2019年第4期。

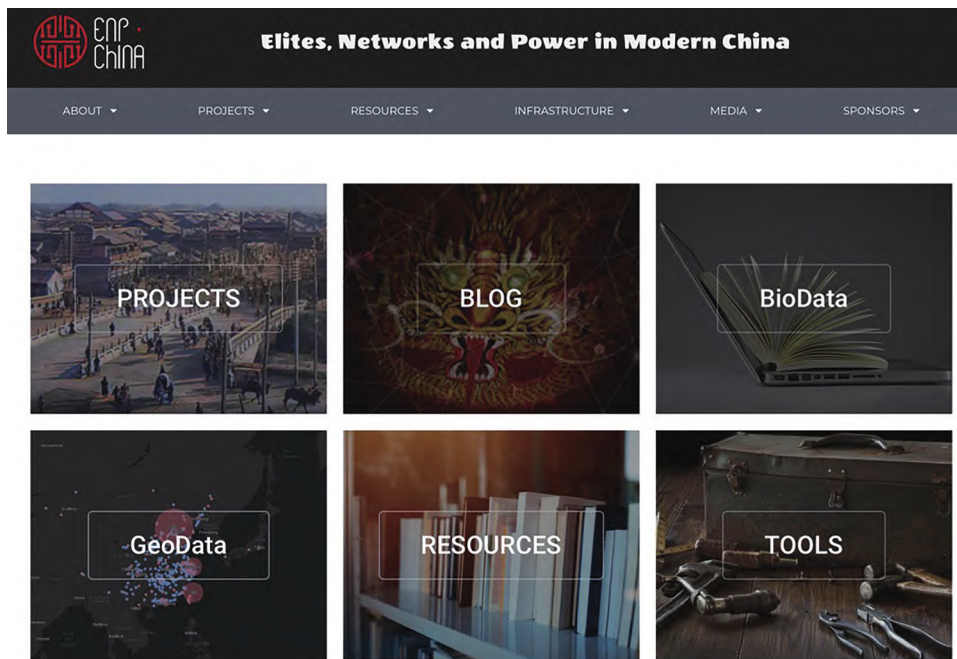


图1 ENP-China 官方网站主页界面（局部）

一、项目概述

（一）研究背景与缘起

在中国被迫进入由西方帝国塑造的世界秩序的一个世纪（1840—1949）里精英群体的性质经历了彻底的转变。由儒家思想培养的文武士大夫组成的少数精英统治集团，于1911年开始逐渐退出历史舞台，权力和社会威望转向了在智力和种族上更为多元化的群体，他们是脱胎于传统士绅的新一批知识分子，包括在环环相扣的跨国网络中运作的中国人和外国人。与此同时，近代中国的重心由乡村转移到城市，城市精英便首先在沿海沿江的开放型经济城市中出现了。

虽然学界对近代城市中的精英已经有了一些有价值的研究成果，如对于商业精英^①、技术精英^②的探讨等。这些研究虽剖析了不同精英群体，却鲜少触及不同精英群体之间是否存在联系、这些联系如何形成，以及这些联系如何让精英在所属城市乃至整个中国范围内塑造了自己的权力。

① 马敏：《官商之间：社会剧变中的近代绅商》，北京：社会科学文献出版社，2022年。

② 巴斯蒂·布律吉埃尔：《近代中国新兴技术精英：工程师的培养及其政治表现，1866—1912》，江穗春译，中国社会科学院近代史研究所政治史研究室、杭州师大浙江省民国浙江史研究中心编：《中国社会科学论坛文集·政治精英与近代中国》，北京：中国社会科学出版社，2013年，第12—31页。

1840年鸦片战争之后,沿海沿江的开放型经济城市加速崛起。经过近一个世纪的发展,到了1930年代,城市内部已经形成了一张丰富的社会网络:一是围绕会馆、同乡会组成的乡缘组织;二是由律师、医生、商会等以职业分工或行业为基础形成的业缘组织;三是与宗教和信仰有关的信缘组织;四是社团组织,包括各种学会、俱乐部等;五是帮会组织,也就是社会正式体制之外的江湖团体。^①近代中国的城市精英大多身处这几个城市网络之中,他们通过与国家权力的密切互动,主导了城市的地方公共事务,又借助学校、传媒和社团掌握了知识和舆论生产,使得知识、资本、舆论和信息在网络中流通。显然,城市精英之间确实构成了复杂关系网络,而这一网络有待深入研究,其组成要素之间如何相互渗透、影响和互动,从而构成一个整体的、流动的“权力网络”,是一个引人深思的问题。

(二) 研究目的

从“精英、权力和网络”这组概念出发,ENP-China项目试图探讨1830—1949年间中国精英的转型过程,重点关注三个现代中国变革的引擎城市区域:北京/天津、广州/香港、上海。具体而言,项目想要解决的问题是:(1)如果从行为者而不是国家机构或社区组织的层面来观察中国的城市社会,哪个群体或个人能够成为“精英”?精英形成和精英网络(重新)部署的模式和过程是什么?(2)精英参与公共行动的载体和模式在不同的城市背景下是如何变化的?这些行动包括以官方身份(行政和军事精英)、自组织协会(如同乡会、商业组织、宗教团体),以及直接(新闻、出版)或间接(电影、广告)承担舆论领袖的角色行事。(3)经济全球化对经济精英在选定的中国城市的转型中所扮演的角色有什么影响?(4)在席卷全国的重大社会政治变动(如辛亥革命、“五四”运动、抗日战争等)中,各种相互竞争的精英组织是如何应对事件、动员组织内个体、管理组织和与其他组织互动的?^②

(三) 研究团队

ENP-China项目受欧洲研究委员会(European Research Council)资助,由法国艾克斯-马赛大学(Aix-Marseille University)的安克强(Christian Henriot)教授领衔,组建了一支强大的跨学科研究团队,包括艾克斯-马赛大

^① 许纪霖:《近代上海城市“权力的文化网络”中的文化精英(1900—1937年)》,《复旦学报(社会科学版)》2012年第6期。

^② 参见ENP-China项目介绍:https://enepchina.hypotheses.org/files/2018/06/ERC_Project-description.pdf, 2024年9月21日。

学亚洲研究所 (Institut de recherches Asiatiques, IrAsia)、语言与语音实验室 (Laboratoire Parole et Language, LPL) 和计算机科学与系统实验室 (Laboratoire d'Informatique et Systems, LIS) 等一流研究机构中的学者, 数据库和数据管理领域的学者, 以及法国社会科学高等研究学院 (EHESS)、法国远东学院 (EFEO) 和中国台湾“中研院”近代史研究所的历史学者。该项目团队也招募历史、计算、数据科学和自然语言处理领域的博士后和历史学博士候选人等年轻学者。^①项目成员的学术活动非常活跃, 经常组织国际活动, 如“精英、知识与权力: 近代中国精英群体的组成与变迁”国际学术研讨会、“数字中国研究新基础”暑期学校等, 与世界各地的学者交流研究成果。

(四) 项目建设过程

ENP-China 项目的基础建设路径由一组相互连接的“技术砖块”(Technological Bricks) 构成, 整个过程错综复杂, 每个步骤都和其他步骤出现双向互动。这一路径最终指向的目标是尽可能地将传统史料转化为“历史大数据”, 这里的“大数据”不仅指历史数据的时间跨度长、范围广,^②也指纸质历史文献数字化后数据

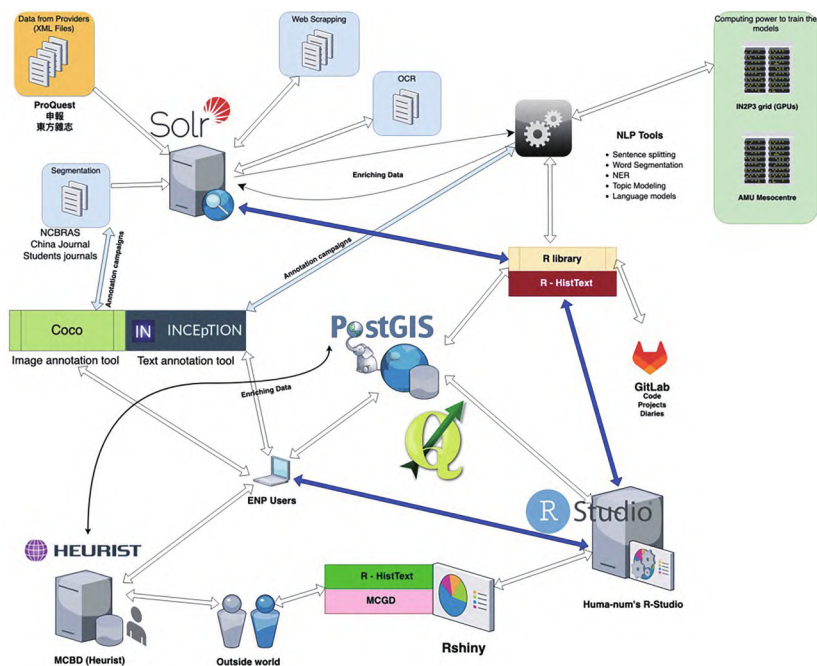


图2 项目基础建设过程

① 参见ENP-China 团队介绍: <https://www.enpchina.eu/our-team/>, 2024年9月21日。

② 段伟文:《大数据透视下的历史与未来》,《史学理论研究》2016年第1期。

发生网络化、开放化和知识结构化等本质变化,^①因此历史大数据能够更好地在数字环境中检索、传播和使用,可以使用搜索引擎等工具进行快速和精确检索,也能够通过各种图表、地图等形式进行可视化。图2为ENP-China官方网站所提供的图示。^②

为方便理解,本文将图中各个相互联结的步骤进行分割和汇总,将其大致分为三部分。

第一部分, **数据采集**。语料将存储到Solr的数据库中,语料收集涉及OCR、网络爬虫和分词技术,资料来源包括他人提供的网页文件(如ProQuest的数据、《东方杂志》、《申报》等)。

第二部分, **数据处理**。对于存储在Solr的数据,使用自然语言处理技术进一步整理,过程辅以人工标注(Annotation Campaigns),使用的工具有COCO图片标注工具和文本标注工具。利用标注好的语料训练模型,再将训练后的模型运用于Solr的原始数据,由此达成数据库和数据处理之间的双向互动。

第三部分, **数据分析**。使用R语言进行数据统计和数据分析,分析结果进一步呈现到PostGIS对象-关系型数据库中,也可以用R语言的Shiny包绘制可视化图表。值得一提的是,ENP-China团队基于R语言自主研发了应用程序HistText 1.0,专门用于历史数字语料库中数据的探索和提取,适用于英语、法语和汉语的分析。团队在Gitlab上发布了代码的公开版本,当前的初代版本附带完整的HistText手册,具备对所有功能的描述,并提供现成的脚本示例。^③

处理好的数据将分别存储到自主研发的两个数据库中(图表将用于“现代中国地理空间数据库”),其中“现代中国地理空间数据库”基于Heurist网络数据库接口创建。^④全部数据库中的数据都是开源共享的,用户可以在相应平台中获取。

二、项目成果:专业数据库平台与案例研究

为了深入剖析所收集数据的丰富特征并开展历史分析,ENP-China项目针对

① 祁天娇:《从历史档案到历史大数据:基于威尼斯时光机十年路径的探索》,《中国图书馆学报》2022年第5期。

② 具体请见: <https://www.enpchina.eu/infrastructure/>, 2024年9月21日。

③ Christian Henriot, Cécile Armand, “HistText Manual,” 2023, <https://bookdown.enpchina.eu/rpackage/HistTextRManual.html>, accessed on September 5, 2024.

④ Heurist是一个网络数据库接口,专为人文学科研究而设计。它为关系型研究数据提供了一个后端和前端的解决方案。通过Heurist,研究人员可以创建、管理、分析和可视化各种类型的数据,如地理、文本、音频和视频等。这个平台易于使用,可以方便地定制和扩展,以适应不同领域和项目的需求。

1830年至1949年间的地理数据和人物数据，分别开发了两个数据库。这两个数据库是ENP-China团队研究议程的一部分，初始目的在于支持项目成员利用数据库进行项目研究和个案分析。不过在实际建设中，人物数据库内容已经不仅限于项目预设的“精英群体”范围，而是囊括了几乎所有中国近代历史的参与者。地理数据库也扩展新增了新中国成立后的地理信息。因此，这两个数据库平台为中国近代历史学研究者、爱好者提供了相当丰富的基础资源。ENP-China的官方网站也展示了利用数据库正在进行的案例研究。

（一）现代中国地理空间数据库（Modern China Geospatial Database, MCGD）

该数据库旨在为现代中国的地理空间分析提供一系列数据集，这些数据集包括“主数据集”和两个子数据集“晚清民国”“共和国”。主数据集包含了项目收集的关于现代中国各地的所有数据，共计466,165个条目。“晚清民国”数据集的数据来源是1931年的《中国古今地名大辞典》，列出了1900年至1949年间中国的主要地点。“共和国”数据集的数据则是现在的城市列表。所有MCGD的数据均开源，有使用需求者可以在数据分享平台Zenodo中的ENP-China数据仓库中获取。^①数据以CSV或XLSX格式储存，读者可用Office软件打开，或用Python等工具对其进行更复杂的处理。

使用者可以利用该数据库查找地名相关的地理信息，只需选择所需数据集后将地名键入（QUERY页面）即可。搜索功能（参见图3）可以识别简体、繁体中文，以及现代汉语拼音、威妥玛式拼音，也能够识别中国各地地名的所有变体和音译变体（例如“上海”的各种音译：Shanghae、Changhai、Schanghai等）。除了单个地名搜索，网站还提供了批量查找功能（FILE IMPORT页面，参见图4）。用户需要将要查询的地名列表以CSV文件上传。搜索出的结果将是一条或多条结构化信息数据，包含了地名（中文和拼音的当前名称以及所有已知名称）、所属省份、经纬度等信息，地理坐标的GIS地图也显示在侧。

① 访问网址：<https://zenodo.org/communities/enp-china/>，2024年9月21日。



图3 MCGD 查询单个地名界面（局部）

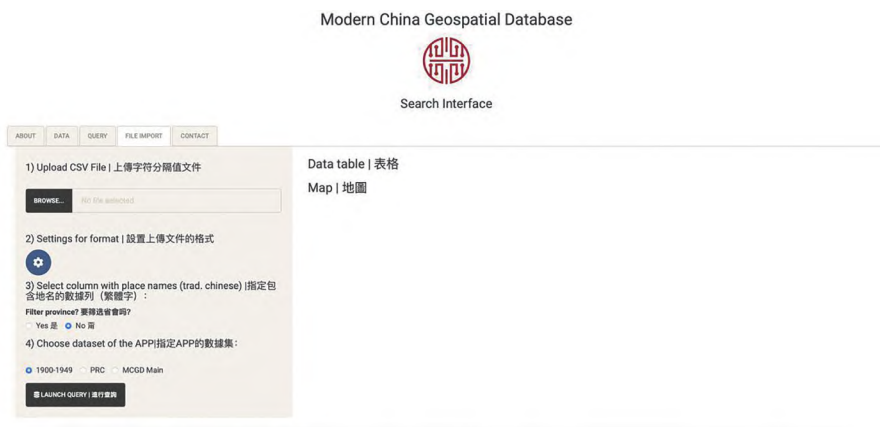


图4 MCGD 批量查询地名界面（局部）

（二）现代中国人物传记数据库（Modern China Biographical Database, MCBDB）

现代中国人物传记数据库旨在收集晚清民国时期（1830—1949）活跃在中国的所有历史人物。其中收录的信息有六大来源，分别是：年谱和名人录、传记词典、报刊、档案、网络数据和其他印刷资料（私人日记等）。所收录的信息内容围绕人物个体展开，也包括机构（任何类型的组织：部委、俱乐部、公司等）、地点（任何命名的住宅区）和事件（任何形式的个人或集体行动）。这些信息既包括“历史信息”也包含“历史数据”，为了更好地建设数据库，需要将二者进行区分。因此MCBD数据库包含结构化的“历史数据”，也包括由非结构化文本

字符串组成的“历史信息”。

整个MCBD的结构较为复杂，从抽象结构来看包含了11个部分，可以形象地理解为描述一个人物的11种属性：

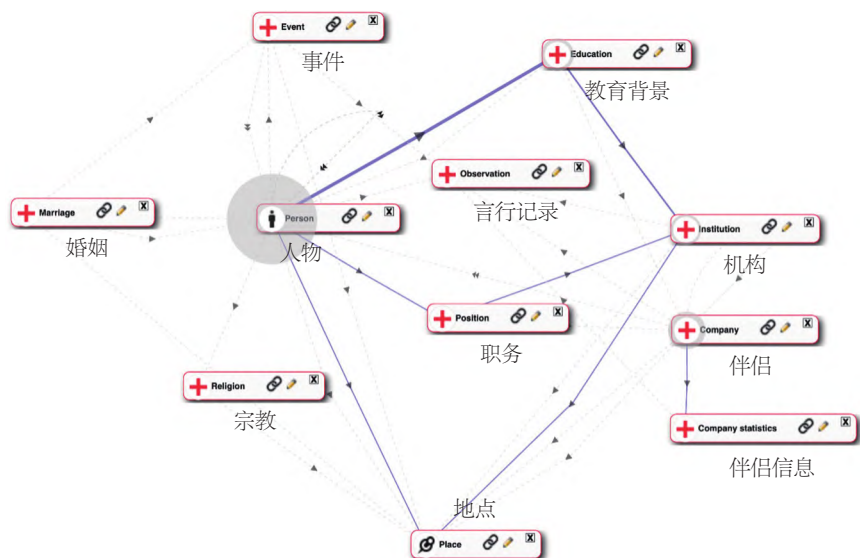


图5 MCB D结构示意图

具体的数据库由五大部分构成：人物、职业及社会生活、个人生活、事件和次级记录（孩子、所属机构等）。每个组成部分又包含了若干表格，例如人物部分由姓名、出生/死亡日期等分表格构成。总体来看，数据库结构呈现出“信息树”样态。

MCBD的建设还处于不断收集数据的过程之中，所以用户目前不能导出全部数据，但是团队会将已有的部分数据集以XML文件的形式定期导出到Zenodo中ENP-China的数据存储库中以供所需。

使用MCBD的中文检索功能可以同时查询多种类型字段，包括全名、姓氏、字号、机构名称以及任何术语字段和出生日期，但需要注意的是它只支持繁体中文。以检索“周树人”为例，检索结果出现了数据库中收录的关于该条目的全部信息。于是可以得知周树人的笔名、出生日期和出生地等信息。通过结果中依次列出的周树人的生平、教育背景和职业活动，能够获知其曾就读于日本仙台医学专门学校，于1920—1926年担任北京大学的教师，曾加入中国左翼作家联盟并成为主席团成员等事迹。不仅如此，结果信息中所有的超链接都可以点击查看，例如点击“绍兴”可以得知数据库中收录的与地点绍兴相关的数据，包括其

他相关历史人物、红十字会这样的社会组织等，充分实现了各类信息之间的网络连接。

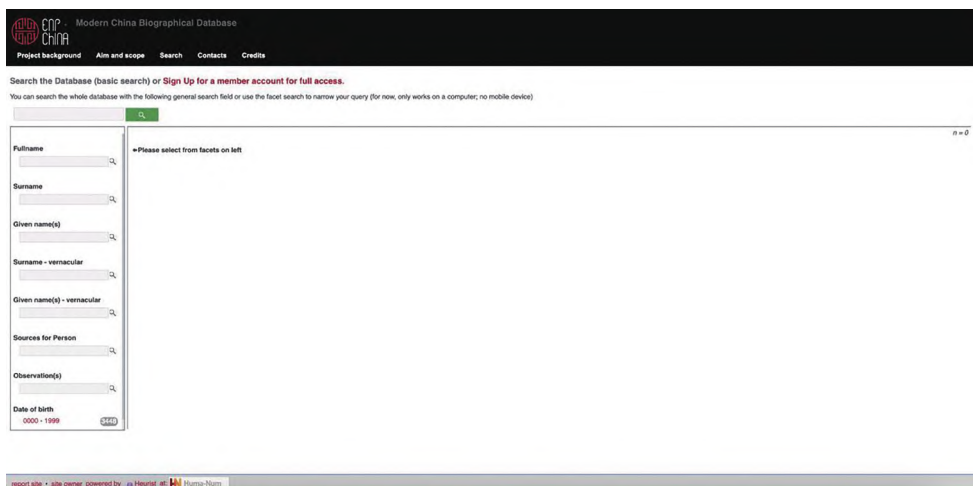


图6 MCBDB 检索界面

（三）不同群体的案例研究

ENP-China的官网提供了八个参考案例。^①从研究对象上来看可以划分为四类研究，每类各由两个案例构成。第一组研究精英社团组织，包括“山姆大叔在中国”（Uncle Sam in China）和“近代中国的知识分子社区与全球网络”（Intellectual Communities and Global Networks in Modern China）两个案例。前者通过社团聚焦美国文化在华传播的两个主要渠道——海外教育和中国的美式社交俱乐部，后者则研究皇家亚洲文会北华支会内的个体或集体，及其出版系统如何通过跨国编辑、作者、读者和学科的日益多样化来构建学术团体。第二组关注精英教育与日常生活，“一个世纪的国家活力”（A Century of National Vibrancy）案例研究精英接受教育与组织基础教育的过程；“近代温州地方精英的日常生活（1830—1949）”（The Everyday Life of Local Elites in Modern Wenzhou [1830-1949]）案例则聚焦温州地方精英的生活故事。第三组深入研究项目所收集的精英数据，“生成历史语料库中的事件链图”（Generating Graphs of Event Chains in Historical Corpora）实践项目试图通过从大量数据中生成事件图进一步优化各种历史事件之间的联系；“近代中国的精英、媒体与舆论”（Elites, the Press and Public Opinion in Modern China）案例则以《申报》数据为例研究现代媒体的

^①ENP-China 案例简介，参考：<https://www.enpchina.eu/projects/>，2024年9月21日。

发展对民国精英之间社会关系的影响。最后一组聚焦精英官僚群体,“国民党军事精英的群体传记分析”(Group Biographical Analysis of the KMT Military Elite)剖析军事精英;“专家、权力与社会:现代中国技术统治的崛起”(Experts, Power and Society: The Rise of Technocracy in Modern China)则探讨新兴的技术官僚阶层。

三、项目特色

(一) 提供历史学研究新视角

对中国本土士绅的讨论已有几十年的历史,不过国内学者的研究普遍将视野框定在“乡土中国”之内,关注的是中国传统乡村的精英阶层如何应对现代社会变迁,而ENP-China则将研究视野完全放置在中国近代新兴的城市之中,将研究对象选取为随着中国城市发展应运而生的新一代精英阶层。由于精英的主要活动领域处于私人领域和国家之间的中间地带,ENP-China仍然使用了西方研究中国城市社会时使用的“公共领域”这一传统概念。对于该概念的阐释通常以国家和区域为主体,离不开“国家与社会”这一分析框架,^①而ENP-China则对此作出了挑战:不同于从国家机构或社区组织的层面来观察中国的城市精英,而将其作为精英权力的来源,它将精英视为个体或集体行动者,试图发现这些近代中国城市中的精英如何反过来通过制度和非正式/正式社会网络的行动,成为影响社会和政治的决定性因素。

近代中国社会的转型,最突出的表现就是传统文化和现代文化之间的张力,因此研究群体或个体在中西文化上作出的选择不可避免。而且不论是逐步走向现代化的城市区域,还是城市中的精英阶层,都受西方刺激和影响颇深。因此ENP-China在研究中增设了一重跨文化视角,将海外在华精英俱乐部以及精英们接受的海外教育也纳入考察范围,试图弥合西方与中国之间的研究裂隙,在不预设何方为主导力量的前提下,研究成分更为多元的精英群体或个人。这种跨文化的视角补充了中国近代社会转型本应具备的多元复杂形态。

(二) 多元丰富的数字成果

ENP-China是一个大型科研项目,在研究所关注的一百余年长时段跨度中,中国社会发生了翻天覆地的变化,因此人物、事件众多。在该项目的网站平台上,展示了该项目目前的全部研究成果,大者有对广泛数据的收集和处理,小者

^① 邓京力:《“国家与社会”分析框架在中国史领域的应用》,《史学月刊》2004年第12期。

则包括对诸多个案的深入探究,从整体来看多元且丰富,有些成果甚至已经溢出了该科研项目最初的研究范围。

可以说,ENP-China真正将数字人文研究“致广大而尽精微”的特点落实到实践之中,为数字历史研究提供了一定的范式。目前国内数字史学领域项目缤纷多样,但是部分研究缺少学术专业性,这主要表现为一些学者过分关注数字处理技术,而忽视了深入的人文阐释,使得项目成果难以自证其价值。Enp-China项目则将数据处理与史学思辨充分结合,在每个完整的研究过程中,都包括了“预设问题—搜集资料—数据分析—诠释”四个步骤,这要求研究者不仅要亲历清洗、处理数据的全过程,也要有足够的史学素养,能够提出假设、建构研究框架,保证问题是历史学界关心的。ENP-China的研究实践启示数字史学研究不应让专业史学研究向卖弄“数字”噱头让步,而应该“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”,^①唯有如此,才能为该领域的繁荣发展建立扎实的“数字技能+史学修养”基础。

(三) 推动资源在开源共享中可持续发展

数据库的建设和开发是需要耗费极大人力、物力成本的宏大工程,因此出于利润考虑,大多数数据库都需要使用者花钱购买,但ENP-China项目团队对数据则显得格外慷慨。项目过程中产生的所有数据集都在Zenodo的ENP-China数据储存库中提供,^②PEERS网站中也登载ENP-China团队在其数字历史实验室发表的系列研究成果、方法论论文等。^③ENP-China的官方网站“Resources”栏目展示了该团队的首部著作《知识、权力和网络:近代中国转型中的精英》(*Knowledge, Power, and Networks: Elites in Transition in Modern China*),其中收录的文章可被视为数据库的使用指南,写作者详细展示了如何利用ENP-China提供的丰富数据和工具来研究史料。^④本文在此选取其中四个代表性案例作简要介绍,以彰显ENP-China在历史研究中发挥的价值。《中层精英:共产主义革命前夕的上海中国银行的中层管理人员和银行专业人员》^⑤聚焦1949年上海中国银行的中层员工,借助ENP-China的档案资料,深入剖析了他们的职位、教育背景、

① 王涛:《“数字史学”:现状、问题与展望》,《江海学刊》2017年第2期。

② 访问网址: <https://zenodo.org/communities/enp-china/records?q=&l=list&p=1&s=10&sort=newest>, 2024年9月21日。

③ 访问网址: <https://www.peers.press/e/P9s7QlDIQeUUaCFfqlO4Lltq5Z2/collections>, 2024年9月21日。

④ ENP-China官网“Resources”页面: <https://www.enpchina.eu/resources/>, 2024年9月21日。Cécile Armand, Christian Henriot, and Sun Huei-min eds., *Knowledge, Power, and Networks: Elites in Transition in Modern China*, Leiden, The Netherlands: Brill, 2022.

⑤ Brett Sheehan, “Middling Elites: Middle Managers and Bank Professionals at the Shanghai Bank of China on the Eve of the Communist Revolution,” *Knowledge, Power, and Networks*, pp. 83-112.

收入、社交网络及性别构成。研究揭示,中层管理人员普遍接受新式教育,并在银行内部通过培训和晋升机制获得认可。这表明尽管等级制度在民国时期的银行业中依然占据重要地位,但现代知识和技能的获取对于中层管理人员的晋升同样关键,同时反映出银行中层管理人员的专业化进程与晚清金融机构职业化进程的相似性。此外,研究观察到20世纪初银行规模的大幅扩张也扩大了中层管理人员的队伍,同时与全球现代化相关的知识和技能也变得重要。《具有中国特色的外国俱乐部:上海扶轮社与语言政治》^①一文,通过数据库中的名人录和档案资料,重构了扶轮社员的生活轨迹和人脉网络,揭示了中国扶轮社员在教育、职业和生活方式上的“美国化”倾向。然而,英语的主导地位和社会活动组织者对外国联系的依赖,使扶轮社在吸引中国精英时面临困境。上海西区扶轮社的成立,虽为本土化尝试,但存在时间短暂,未能带来显著变革。这一案例揭示出外国俱乐部在中国本土化过程中的挑战与困境。《留美学生:民国初年的教育网络和新商业形式》^②利用MCBD数据库和ENP-China提供的《中国学生月刊》(*Chinese Students' Monthly*)、机构档案(中华教育文化基金会、中国第二历史档案馆等)和中美报纸等材料,结合网络分析技术,研究了留美学生回国后的社会活动轨迹,分析出他们如何运用国外知识服务中国社会或追求个人目标。这一研究勾勒出留学生的社会网络、职业选择及其对近代中国社会发展的贡献。《政治权威之间的游走:1949年至1966年洛克菲勒华裔生物和化学学者及其职业轨迹》^③通过分析洛克菲勒大学研究员数据库中的信息以及多重对应分析(MCA)结果,探讨了这些学者出国学习后的职业路径及其异同。

借助高质量的数据集和配套使用教程,ENP-China项目组在最大程度上保证了自己建设的数据库能被合理使用,让数据库具备了超越检索功能的学术潜能。研究者可以根据自己的研究需求下载数据,进行多种方式的数字人文研究探索。通过单独运用或组合运用,数据可以用于社会网络分析、空间分析和所有类型的可视化。由于MCBD源数据库中的每个文档都带有自己的可识别ID,保持了与文章全文的联结,研究者也可以使用自然语言处理工具进行文本分析。如此一来,ENP-China的数据不再处于封锁状态,而是犹如源源不断的活水,流向世界各地

①Cécile Armand, "Foreign Clubs with Chinese Flavor: The Rotary Club of Shanghai and the Politics of Language," *Knowledge, Power, and Networks*, pp. 231-257.

②Peter E. Hamilton, "The American-Returned Students: Educational Networks and New Forms of Business in Early Republican China," *Knowledge, Power, and Networks*, pp. 258-288.

③Yi-Tang Lin, "Navigating between Political Authorities: Chinese Rockefeller Fellows in Biology and Chemistry and Their Career Trajectories from 1949 to 1966," *Knowledge, Power, and Networks*, pp. 289-321.

的使用者，这恰恰彰显了数字人文基础设施建设的深远价值。

四、结语

ENP-China项目本质上是跨学科的数字人文研究，它植根于历史研究，但大量借鉴社会学和政治学的概念和方法，从数据获取到分析的始末也离不开数字技术的充分运用。除了为数字史学研究范式的建构提供有益参考，ENP-China建构的地理信息和人物信息两类大型语料库，以及自主开发的历史文本分析工具，也都成为其他研究者可参考和使用的宝贵资源。此外，ENP-China项目也通过实践探索推动了计算方法的更新，在方法论层面上探索了数据挖掘和机器学习如何更好地适配于史学研究。

通过上述综合分析，可以认为ENP-China项目是数字史学自身合法性的有力证明，即“建立在与过去研究方法不同的巨观与复杂计算之上，不断地寻找与提出只有运用数字技术才能提出的史学问题并给出答案”。^①史料是史学研究的基础，“大数据时代”的来临对历史学的重要影响就体现在对史料的重新发掘和认识上。ENP-China项目借助数字化历史资料来研究现代中国的阶段性变化，在数字技术的助力之下，它一方面从迄今尚未开发的资源中提取超大规模的历史信息，另一方面对现有数字资源进行重塑和整合，最终使得仅靠传统方法难以利用的资料在大数据方法处理下重新焕发生机，为世人展现出更细腻、丰富的历史图景。

从项目官方网站展示的成果来看，仍有一些不足之处亟待改进。首先，网页内容过于庞杂，导致部分内容出现重复，且部分超链接存在错误，这需要开发人员进一步优化和完善。其次，两大数据库平台的功能尚显基础，目前停留于“搜寻”功能阶段。为了充分发挥其潜力，可以进一步研发量化分析和可视化等高级功能，或建立专题数据库，深入挖掘海量信息之间的关联，从而最大限度地提升数据库的使用价值。尽管存在这些细微瑕疵，但ENP-China项目作为跨学科研究的代表，已经展现出了其巨大的潜力和价值。随着跨学科研究团队中每个学者的不断精进和雕琢，相信该项目未来会产出更多杰出的数字人文成果，为历史学研究挖掘出新的可能性。

^① 邱伟云：《验证、修正、创新：数字史学方法的三重功能》，《南京大学学报（哲学·人文科学·社会科学）》2019年第2期。

Exploration of Project Modes on Digital History: Example of “Elites, Networks, and Power in Modern China” (ENP-China)

Qi Xinxin

Abstract: An interdisciplinary team from Aix-Marseille University in France has launched a large-scale project named “Elites, Networks, and Power in Modern China” (ENP-China), aimed at exploring the transformation process of China’s modern elite during the period from 1830 to 1949 using digital techniques. This project focuses on the behavioral patterns of individual elites, analyzing how they influenced social and political landscapes through the construction of social networks. Through meticulous planning of infrastructure development paths and the development of historical text analysis tools, the project team has collected vast amounts of geographical and biographical data, established open database platforms, and conducted diverse research based on these data. This project demonstrates its unique characteristics in three aspects: historical perspective, digital outcomes, and basic resource construction. Its construction path, methods, and achievements provide a referential paradigm for digital historical projects.

Keywords: Digital History; Modern Chinese History; Digital Humanities Project; Social Network Analysis

(编辑: 许可)