

## 国内数字人文视角下语言文学研究热点与趋势分析

### Analysis of the Hot Spots and Trends of Language and Literature Research from the Perspective of Digital Humanities

曹轩铭<sup>1</sup>，何颜菽<sup>1</sup>，武永<sup>2\*</sup>

<sup>1</sup>北京邮电大学 国际学院

<sup>2</sup>北京邮电大学 人文学院

\* wuyong@bupt.edu.cn

**【摘要】** 数字人文是信息时代人文研究的前沿热点，但其在语言文学领域的应用还处于起步阶段。本研究以知网论文为主要数据，采用文献计量方法，使用可视化分析软件VOSviewer，对近15年，数字人文视角下的语言文学研究进行了统计，并分析该领域研究热点的演化路径及发展方向。

**【关键词】** 数字人文；语言文学；研究热点；趋势；VOSviewer

**Abstract:** Digital Humanities is a research hotspot in the field of Humanities in the information age, but its research in the field of language and literature is still in its infancy. From the perspective of Digital Humanities, taking CNKI data as the main source, using bibliometric methods and visual analysis software VOSviewer, this study analyzes the hotspots, evolution path, and direction of the language and literature research.

**Keywords:** digital humanities, language and literature, hot spots, direction, VOSviewer

#### 研究背景和意义

自2009年“数字人文”概念引入中国以来，数字人文领域研究热度逐年升温。近年来，随着相关技术的发展，数字人文也成为语言文学研究中的热点。

数字人文在语言文学领域的研究还处于起步阶段。了解该领域研究热点及其发展过程，对于掌握其研究现状、发展趋势、以及未来开展数字人文视角下的语言文学研究有十分重要的意义。

#### 研究问题

1 数字人文视角下的语言文学研究领域呈现哪些特点？

■ 目前在国内，将数字人文技术和语言文学相结合的研究尚处于起步阶段，研究领域内的高产作者呈现出高度聚集的特点，多集中在高校及大型图书馆。造成这一分布的原因可能是国内研究技术尚不完备，因而仅在多所高精技术机构中有较多应用，研究区域相对聚集。研究领域内的文献发布情况呈现范围广泛，核心相似的特点，不同高被引文献的研究方向不尽相同，但研究过程中使用的核心概念则较相类，出现这一分布的原因可能是国内对相关概念的研究尚不透彻，已开发领域较少，技术限制较大。

2 数字人文视角下的语言文学研究近年来的前沿课题和新兴热点是什么？

■ 文本挖掘和可视化等热点已逐渐常态化，而随着时间的推移，远读、计算批评、图谱构建等成为了时下流行的热点，而这些热点相较于此前更为具体、深入，在技术手段上也更为成熟。因而可以预测，在未来的一段时间内，语言文学研究的热点将会进一步细化，与技术手段的结合也更加紧密和深入。

#### 研究方法

本研究从数字人文视角出发，以知网数据为主要来源，使用文献计量方法，结合可视化分析软件VOSviewer，对近15年数字人文视角下的语言文学研究进行统计，分析该领域的研究热点、方向及演化路径。为了保证计量信息的效度和信度，进行了三轮数据筛选。首先，删除重复文献、书评、综述、会议摘要和新闻等。而后经作者核对和同义关键词替代，再经人工核查的方式剔除访谈、综述、目录等文献，最后确定94篇文献为研究样本，并采用科学知识图谱和内容分析法分析数据文献。知识图谱分析基于VOSviewer v1.6.17绘制数字人文视角下语言文学研究的关键词共现图谱和时序图谱。内容分析基于文献的题目、摘要和内容深入解读研究重点。

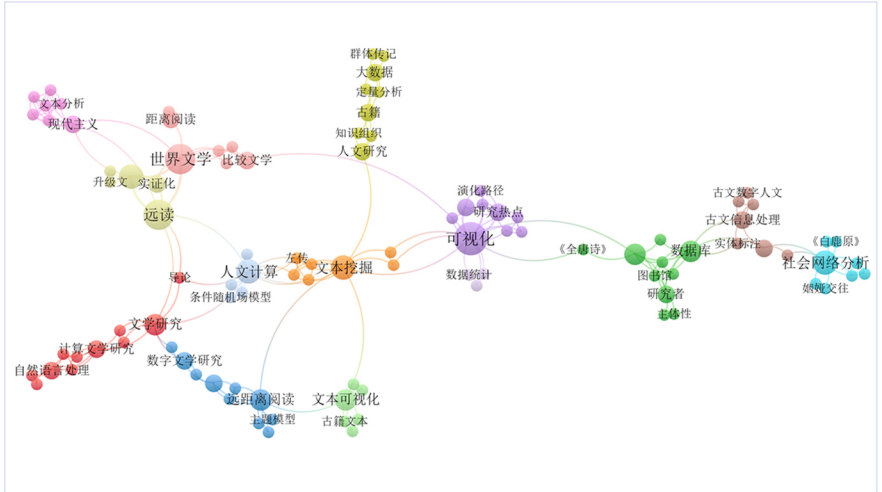


## 调查结果

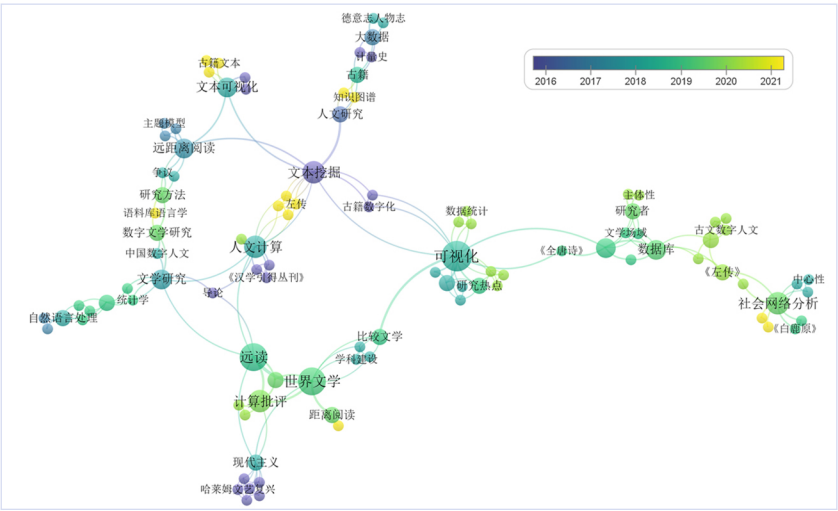
1 从发文总量来看，国内学者的论文呈大幅增长趋势。2013-2015年期间发文5篇，2016-2018年期间发文24篇，2019-2021年期间发文69篇。

2 从研究机构来看，数字人文研究团队主要聚集在各高校及大型图书馆，中文文献研究工作者主要集聚在北京大学（中文系）、南京农业大学、南京大学、上海图书馆、南京师范大学等机构。

3 从作者发文量来看，王东波（11篇）、李斌（7篇）、赵薇（7篇）等几位作者发出的高被引文献较多。就研究领域而言，这些高产作者主要从事数字人文视域下古代文学、语言文学学习模式、数字人文技术革新等方面的研究。



4 从研究前沿来看，在文献研究时，通过VOSviewer对94篇文献中频繁重复出现的277个关键词构成的集合进行共被引聚类分析，最终形成由34个具有统计意义的聚类组成的聚类视图，每个聚类都代表一个研究前沿热点。为挖掘每个前沿热点的历时演变路径，我们绘制了277个关键词形成的聚类网络的标签视图，并按照时间顺序通过颜色深浅标示出关键词聚类的热点时期。根据聚类分析结果，同时结合时间线标签视图，我们对34个聚类分析热点做进一步合并，最终发现计算批评、时间序列分析、知识图谱、关联数据、信息资源开发等研究前沿。



### 研究不足与展望

本研究仅对国内的数字人文发展情况进行了调查，并在此基础上进行了一定的研究分析，但是尚未分析国际上的数字人文发展情况及其在语言文学中的发展趋势。结合国际发展趋势对这一问题进行深入研究考虑势必使我们的研究更为完善、全面，同时也具备更广的涵盖范围和一定的前瞻性。

## 部分参考文献

黄水清、王东波和何琳（2015）。以《汉学引得丛刊》为领域词表的先秦典籍自动分词探讨。《图书情报工作》，59(11)，127-133。

戴安德、姜文涛和赵薇（2016）。数字人文作为一种方法：西方研究现状及展望。《山东社会科学》，2016(11)，26-33。

欧阳剑（2016）。面向数字人文研究的大规模古籍文本可视化分析与挖掘。《中国图书馆学报》，42(02)，66-80。

郭英剑（2018）。数字人文：概念、历史、现状及其在文学研究中的应用。《江海学刊》，2018(03)，190-197+239。

赵薇（2021）。数字时代人文学研究的变革与超越——数字人文在中国。《探索与争鸣》，2021(06)，191-206+232-233。