

北京数字人文特别兴趣小组

201911 研讨会报告

研讨会议题

人文背景学者与非人文背景学者如何在数字人文研究中合作

研讨会旨趣

近几年数字人文在中国发展迅速，讲座、工作坊、论坛等成为介绍、讨论数字人文的重要形式。随着数字人文基本知识、方法论和工具的普及，如何利用这些知识、方法论、工具切实解决人文领域的学术问题，如何推动数字人文知识、方法论、工具本身的发展，逐渐成为研究者越来越迫切的需求。我们希望以例会形式在北京举办的“北京数字人文兴趣小组研讨会”（Digital Humanities Special Interest Group – Beijing Seminar，下文简称 DHSIG 研讨会）正是对此需求的积极回应。

DHSIG 研讨会的目标为：定期对数字人文研究工具、方法、思潮等热点议题组织圆桌讨论。讨论不设主讲者，只设置引导讨论的主持人 / 主持小组。视所有参与者为平等的研讨者，会前向参与者分发相关材料，以便讨论者在会前了解每次议题的基本背景。为促进学生积极发言，每个环节的议题我们先邀请学生发言，之后所有与会者再参与讨论。以期通过研讨会成员的讨论，对此议题形成共同结论。会后以共同署名倡议的形式向公众发表这一结论（这一结论内部可能由多种互补，甚至对立的论点组成）。我们希望这一结论或在纷繁的讨论中，为数字人文研究者对某一议题提供一系列可操作的方案；或将当前的某些探讨、探索面向未来推进。

我们的倡议

关于“人文背景学者希望用数字人文的方法解决什么问题”的倡议

我们认为：首先，人文背景学者希望数字人文方法能帮助自己完成一些机械耗时的学术工作。

第二，人文背景学者希望数字人文方法能够突破纸质载体对内容呈现的限制。

第三，人文背景学者希望数字人文方法能将某些复杂和繁琐人文研究方法自动化。例如在古籍整理中，自动校勘、自动过录、自动引文核对（溯源）、不同历法精确对应、自动生成谱系树，以及对印章及古文字释读方面的帮助，等等。

第四，人文背景学者希望数字人文方法能帮助自己对某些经验判断进行定性描述。例如对于古籍版本的某些经验性知识的定性尝试。

关于“数字人文研究中人文背景学者和非人文背景学者在交流中需要注意什么”的倡议

我们认为：第一，非人文背景学者进行人文研究探索，或人文背景学者进行非人文研究探索时，一定要遵循探索领域的学术规范。

第二，我们倡议和数字人文相关的领域，如文学、历史学、哲学、数学、计算机科学、统计学等，建立面向其他领域同事的专有名词表，以便其他领域学者了解自己领域的专业术语。

第三，我们倡议在数字人文合作中，合作多方基于平等协商同意的前提下，可以有策略地安排署名。

第四，我们倡议对于多方合作的研究或项目，需要用其他领域合作者能理解的方式向对方介绍自己领域在此研究或项目中的现状。

关于“数字人文研究中人文背景学者和非人文背景学者应该受到哪些训练”的倡议

我们倡议：首先，对数字人文感兴趣的人文背景学者应该学习编程语言（推荐 Python），以便了解程序能面向人文做什么。第二，对数字人文感兴趣的人文背景学者应该学习正则表达式。如果畏难于正则表达式，至少应该掌握 Word 的高级查找、替换方法（wildcards）。第三，系统掌握 ANSI, GB, GBK, BIG-5, UTF-8, Unicode 文本编码知识。第四，了解基本的 Excel 公式，如 vlookup, if, left, right 等。

我们倡议对古代中国研究感兴趣的非人文背景学者应当阅读荣新江教授的《学术训练与学术规范》。

关于“如何评估数字人文项目和研究成果”的倡议

我们认为：首先，对数字人文项目和研究成果的评价应当侧重其对人文领域的学术贡献。

第二，工程性数字人文论文不应该被低估。我们倡议对工程性数字人文论文评估的参考角度如下：1. 面向人文需求的准确性；2. 内容的系统性；3. 工程方法的可复用性；4. 面向人文背景学者的可用性和易用性；5. 数据或者工具是否有面向人文背景学者的可复核性（反之，面对人文背景学者，它是否只是一个“黑箱”）；6. 与其他项目的兼容性；7. 在人文研究中的引用率。

第三，我们倡议数据、模型、算法代码的公开和开源。公开和开源有助于对数字人文研究结果的重现和复核，可以避免重复工作。我们倡议基于 Github 开放，并在 readme.md 中描述使用方法，并说明版权。考虑到本土项目、中心等自身利益，我们能理解不愿公开和开源的学术项目，但是仍然强烈希望对于可以开放开源的内容尽量开放开源。

倡议者（按姓名音序）

陈佩辉，北京大学哲学系博士

高树伟，北京大学中文系博士

何孟纯，人民出版社编辑

史睿，北京大学中国古代史研究中心副主任

王宏甦，中国历代人物传记资料库项目经理

王军，北京大学信息管理系教授

王林旭，北京大学信息管理系硕士

汪伟，北京大学哲学系博士

徐力恒，香港城市大学中文及历史学系助理教授

徐阳，北京大学历史学系硕士

严程，清华大学历史系助理研究员

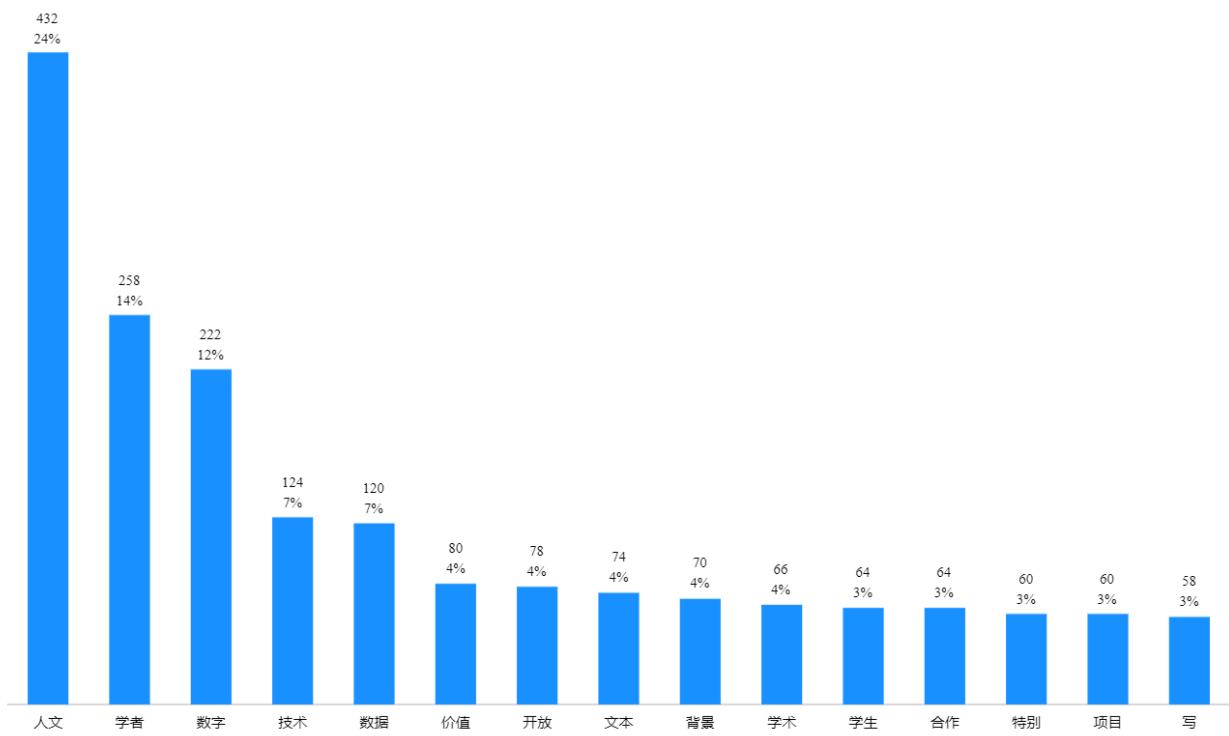
杨浩，北京大学哲学系助理教授

攸兴超，北京斗拱科技有限公司与北京大学数据分析研究中心联合工作室

张力元，北京大学信息管理系博士

朱本军，北京大学信息管理系博士

会议录音高频词表



附件一：实务手册

https://github.com/DHSIG/dhsig_bj_files/blob/master/Seminar%20references/201911/201911%20DHSG%20-%20Beijing%20Seminar%20Practice%20Manual.docx

附件二：参考文献

https://github.com/DHSIG/dhsig_bj_files/tree/master/Seminar%20references/201911/references