

IT4L2020·人文研究的技术之美

暨2020上海图书馆开放数据竞赛训练营

Webinar讲座周

5.25~5.29 中国·上海
上海市图书馆学会

指导单位：上海市经济和信息化委员会
上海市文化和旅游局

主办单位：上海市图书馆学会
文化和旅游研究上海图书馆基地
中国数字人文机构联盟（筹）

承办单位：上海市图书馆学会数字人文委员会
上海市图书馆学会信息技术委员会
上海图书馆历史人文大数据中心

支持单位：上海阿法迪智能标签系统技术有限公司

媒体支持：〈全国报刊索引〉
〈图书馆杂志〉
〈信息资源管理学报〉
〈中国图书馆学报〉
「数字人文研究」公众号
「察言观数」公众号

IT4L2020 · 人文研究的技术之美

暨2020上海图书馆开放数据竞赛训练营

Webinar讲座周 | 5.25~5.29 中国·上海 上海市图书馆学会

5月25日 (14:00~16:00)

主题：关键技术大观

主持：刘炜 (上海图书馆副馆长)

▶ 5月25日

主题：关键技术大观

主持：刘炜（上海图书馆副馆长）

IT4L2020

人文研究的技术之美

暨2020上海图书馆

开放数据竞赛训练营

Webinar讲座周

5.25-5.29 中国·上海
上海市图书馆学会

14:00-14:30

刘炜（上海图书馆副馆长）：人文研究技术大观

14:30-15:15

王鑫（天津大学教授）：数字人文领域的知识图谱构建技术

15:15-15:45

张喆昱（上海图书馆馆员）：上海图书馆开放数据竞赛获奖作品分析

15:45-16:00

互动

IT4L2020

人文研究的技术之美

暨2020上海图书馆 开放数据竞赛训练营

5.25-5.29 中国·上海
上海市图书馆学会

Webinar讲座周



刘炜：人文研究技术大观

上海图书馆上海科学技术情报研究所副馆（所）长。中国图书馆学会阅读推广委员会数字阅读专业委员会主任，国际图联图书馆理论研究专业委员会常务委员，《图书馆杂志》、《图书情报工作》、《大学图书馆学报》等多个专业期刊编委。先后从事图书馆学情报学理论研究、图书馆自动化系统的开发维护、数字图书馆研究和建设等工作，获中宣部四个一批人才、上海市领军人才和上海市劳动模范等称号，享受国务院特殊津贴。

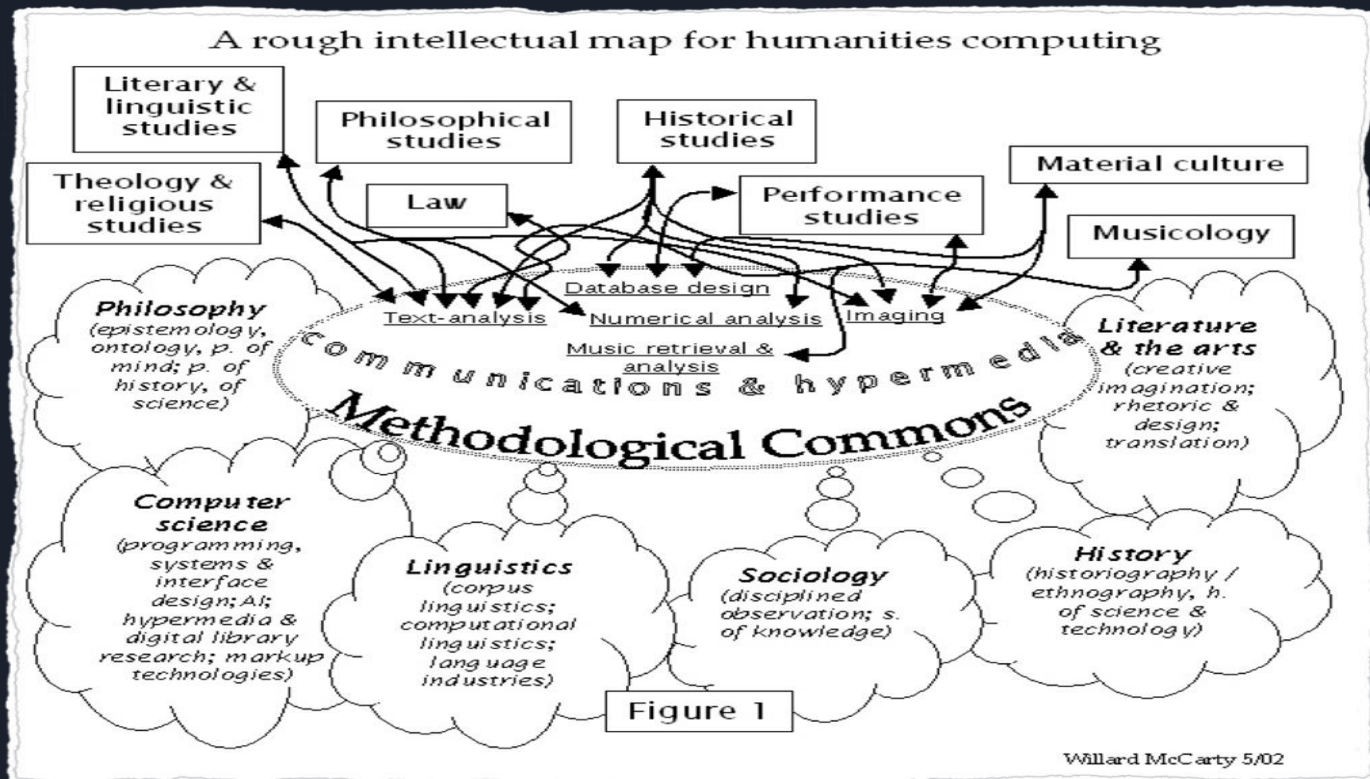


数字人文技术大观

刘炜 上海图书馆
wliu@libnet.sh.cn



人文计算/数字人文肇始



数字人文的技术体系

表1 数字人文技术体系的组成

技术类型	具体技术组成(不限于)	实现功能(学术原语)
数字化技术	扫描, 拍摄, 采样, 捕捉, 图形设计, 3D 建模等技术	收集 Collecting 发现 Discovering
数据管理技术	文本编码, 语义描述, 本体建模, 数据库设计, 多媒体搜索, 语义搜索, 数据看护, 名称实体提取(NER), API 数据服务等技术	收集 Collecting 发现 Discovering 协作 Collaborating
数据分析技术	文本分析(词频、共现、关联、向量、概率), 聚类分类, 主题分析, 内容挖掘, 时序分析, 地理空间分析, 社会关系分析等技术	比较 Comparing
可视化技术	信息美学, 知识地图, 主题图, 关联呈现, 场景模拟, 历史仿真等	比较 Comparing 发布 Delivering
VR/AR 技术	人机交互技术, 脑机界面, 认知技术, 互动测量, 游戏化学习, 计算机竞技等	发布 Delivering 协作 Collaborating
机器学习技术	自动分类, 图像视频音频识别和分析, 个性化服务, 精准推送, 深度学习, 超级计算, 机器绘画, 机器作诗等	发现 Discovering 比较 Comparing 协作 Collaborating

数字人文典型技术应用 五大金刚

数据化

实体识别
图像检索
本体设计
关联数据

可视化

包括VR/AR/XR
3D建模

文本分析

OCR
文本挖掘
词频统计
情感分析

时空分析

GIS
历史地图
时光机

社会网络 分析

关系识别
推理检索

数字人文典型技术应用 五大金刚

- 数字人文技术体系
- 文本图像与社会网络
- 领域与案例
- **GIS应用**
- 可视化

▶ 5月26日

主题：文本、图像与社会网络

主持：王丽华（上海大学副教授）

14:00-14:30

许鑫（华东师范大学教授）：学派视域下的多重关系网络构建与解析
——以清代汉学派为例

14:30-15:00

陈柏彤（上海大学讲师）：文本数据主题分析与演化分析

15:00-15:30

陈涛（上海图书馆高级工程师）：IIIF在文化遗产图像资源中的深度应用

▶ 5月27日

主题：领域与案例

主持：韩春磊（上海图书馆信息处理中心主任）

14:00-14:30

杨继东（斯坦福大学东亚馆馆长）：数字人文时代的学术出版
——斯坦福大学出版社的尝试

14:30-15:00

王涛（南京大学教授）：历史书写方式的数字人文转向

15:00-15:30

张书华（2019竞赛一等奖团队代表，清华大学图书馆馆员）：
以景点和人物为中心的多数据融合小程序——丹红印记开发介绍

15:30-16:00

互动

▶ 5月28日

主题：地理信息系统

主持：夏翠娟（上海图书馆研究员）

14:00-14:30

潘威（云南大学副教授）：GIS技术在西北传统水利秩序研究中的应用

14:30-15:00

张学宝（西安云图信息技术有限公司总经理）：
HGIS在人文领域的应用实践及技术选型

15:00-15:30

夏翠娟（上海图书馆研究员）：面向历史人文研究的时空数据基础设施建设

▶ 5月29日

主题：数字人文可视化

主持：张磊（上海图书馆正高级工程师）

14:00-14:30

魏记（华为技术专家）：华为河图助力数字人文

14:30-15:00

向帆（清华大学副教授）、朱舜山（首席工程师）：
疫情的流图——以设计实验理解视觉表达和可视化技术

15:00-15:30

杨冠灿（中国人民大学讲师）：京剧脸谱的数据分析与可视化

15:30-16:00

王军（北京大学教授）：从人文计算到可视化——数字人文的发展脉络梳理

16:00-16:30



概念：数字人文

Digital humanities (DH) is an area of scholarly activity at the intersection of computing or digital technologies and the disciplines of the humanities. It includes the systematic use of digital resources in the humanities, as well as the reflection on their application. DH can be defined as new ways of doing scholarship that involve collaborative, transdisciplinary, and computationally engaged research, teaching, and publishing. It brings digital tools and methods to the study of the humanities with the recognition that the printed word is no longer the main medium for knowledge production and distribution.

https://en.wikipedia.org/wiki/Digital_humanities



概念：数字人文

数字人文（DH）是计算或数字技术与各人文学科交叉融合的学术领域。它包括在人文学科中系统地使用数字资源，并对其应用进行反思。DH可以被定义为新的学术研究方式，涉及到合作、跨学科和计算参与的研究、教学和出版。它将数字工具和方法引入到人文学科的研究中，并认识到印刷品不再是知识生产和传播的主要媒介。

https://en.wikipedia.org/wiki/Digital_humanities



概念：数字人文

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。

-----引自：中国人民大学新增数字人文硕士点论证材料

数字人文组成





如何能获奖？

- 洞悉“如何开发数字人文项目”
- 创意：充分展示技术的人文之美
- 我有很好的创意，没人来实现怎么办？
- 我有很好的技术，但是不知道做什么怎么办？
- 不提倡的做法
- 怎么做容易获奖？



总结

数字人文，什么都行

创意优先，初步实现

展示资源，宣传特色

发挥所长，弘扬文化



谢谢！

IT4L2020

人文研究的技术之美

暨2020上海图书馆 开放数据竞赛训练营

5.25-5.29 中国·上海
上海市图书馆学会

Webinar讲座周



王鑫：数字人文领域的知识图谱构建技术

天津大学智能与计算学部教授、人工智能学院副院长，中国计算机学会高级会员、信息系统专委会秘书长、数据库专委会委员、中国中文信息学会会员、中国人工智能学会会员，国际期刊 Knowledge-Based Systems 副主编，主要从事知识图谱数据管理与机器学习、大规模图数据库、大数据分布式处理等方向研究，在国内外学术期刊和会议上发表论文 80 多篇，主持国家自然科学基金项目、国家重点实验室开放课题等多项科研项目。

IT4L2020

人文研究的技术之美

暨2020上海图书馆
开放数据竞赛训练营

5.25-5.29 中国·上海
上海市图书馆学会

Webinar讲座周



张喆昱：上海图书馆开放数据竞赛获奖作品分析

上海图书馆馆员，研究方向为开放数据、公共数字文化。

IT4L2020·人文研究的技术之美

暨2020上海图书馆开放数据竞赛训练营

Webinar讲座周

5.25~5.29 中国·上海
上海市图书馆学会

指导单位：上海市经济和信息化委员会
上海市文化和旅游局

主办单位：上海市图书馆学会
文化和旅游研究上海图书馆基地
中国数字人文机构联盟（筹）

承办单位：上海市图书馆学会数字人文委员会
上海市图书馆学会信息技术委员会
上海图书馆历史人文大数据中心

支持单位：上海阿法迪智能标签系统技术有限公司

媒体支持：〈全国报刊索引〉
〈图书馆杂志〉
〈信息资源管理学报〉
〈中国图书馆学报〉
「数字人文研究」公众号
「察言观数」公众号