



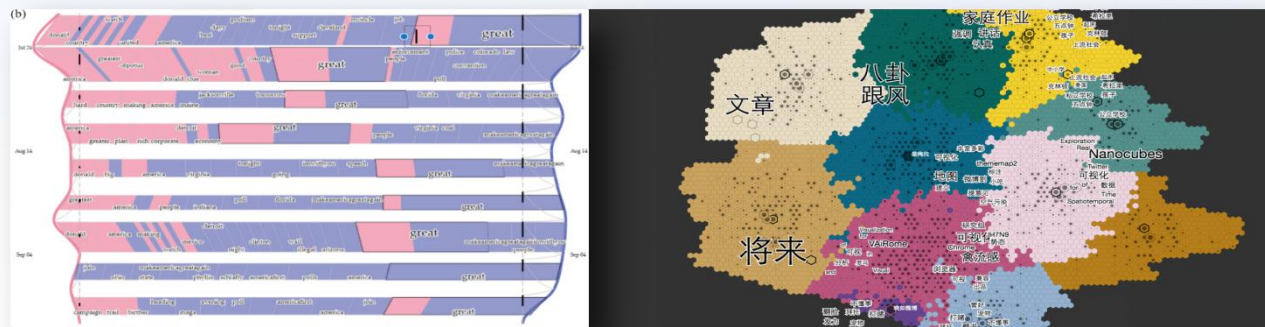
# FDUVIS数字人文之路

---

陈思明 青年研究员, 博士生导师

复旦大学大数据学院

# 大数据可视化与可视分析 - 复旦大学大数据学院, 陈思明



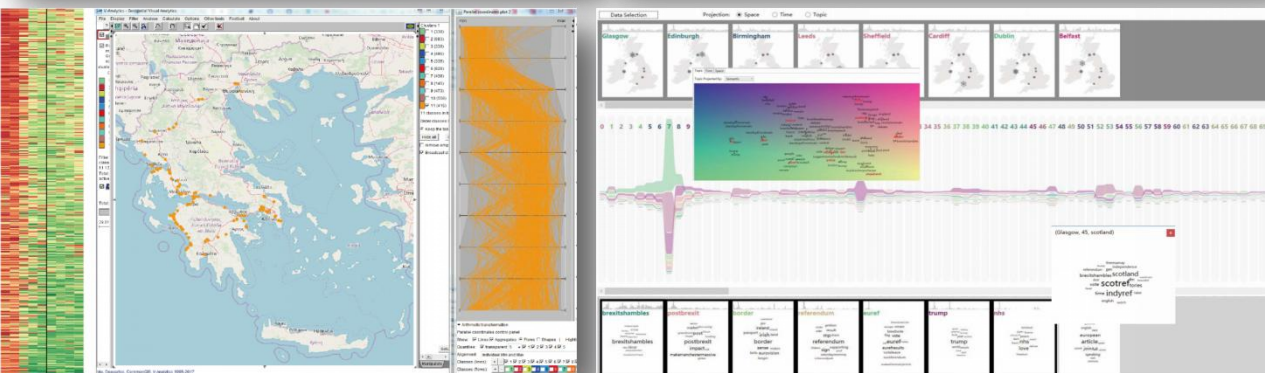
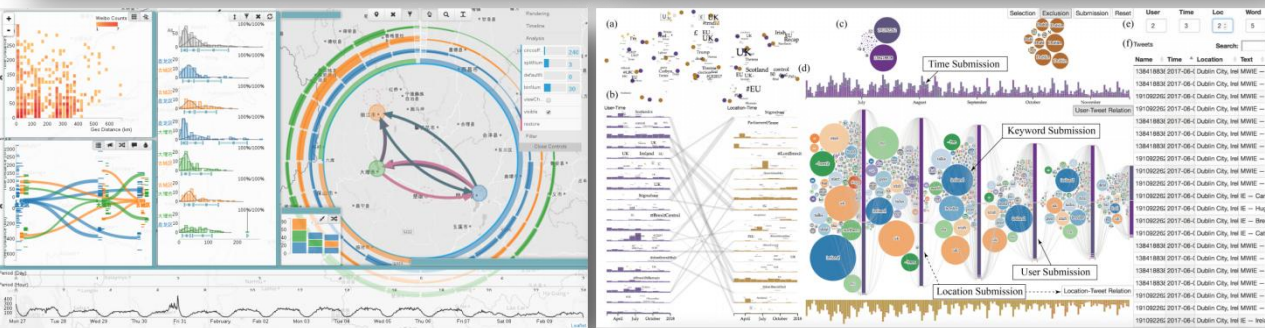
本科：复旦大学：计算机学院

博士：北京大学：智能科学与技术

博士后：德国波恩大学

研究科学家：德国Fraunhofer

青年研究员/博士生导师：复旦大学 大数据学院



扬帆计划获得者、上海市高层次人才。发表论文80余篇，在 IEEE Visualization, IEEE TVCG, ACM CHI, CGF, EuroVis 等重要国际顶级可视化会议以及期刊上发表20余篇文章。担任多个国际会议的组织委员会和程序委员会成员，包括IEEE PacificVis宣传主席，ChinaVis数据分析挑战赛主席，VizSec 程序委员会委员等。他的工作曾获得8次IEEE VAST Challenge数据挑战赛一等奖，以及多个会议最佳论文/海报(提名)奖。

# 实验室简介

---

- 复旦大学可视分析与智能决策实验室（FDUVIS）成立于2020年9月，我们的研究方向涵盖了可视化与可视分析、人机混合智能、用户行为分析、决策支持与数据新闻故事叙述等，并在多个应用领域，包括社交媒体、网络空间安全、时空城市大数据、人文历史数据与金融科技方面有一定的科研成果。研究组秉承着人机协同、共同智能的理念，利用人机交互、数据可视化与机器学习相结合，搭建用户与数据之间的桥梁，让数据及数据背后的意义变得可感知、可理解以及可交互，进而支持人在复杂环境中进行智能决策。
- <http://fduvis.net>



# 数字人文

---

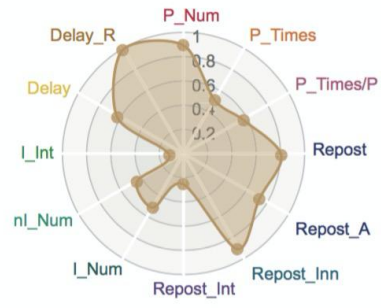
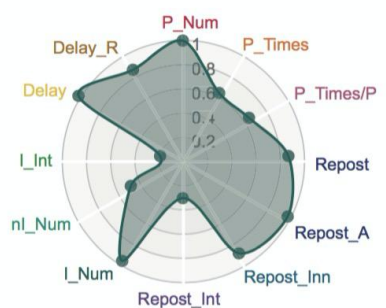
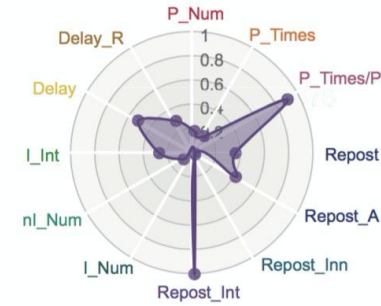
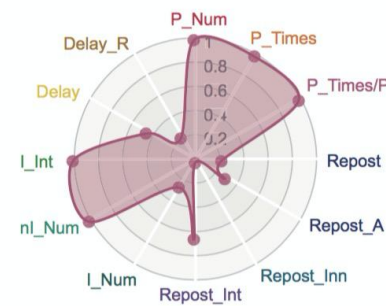
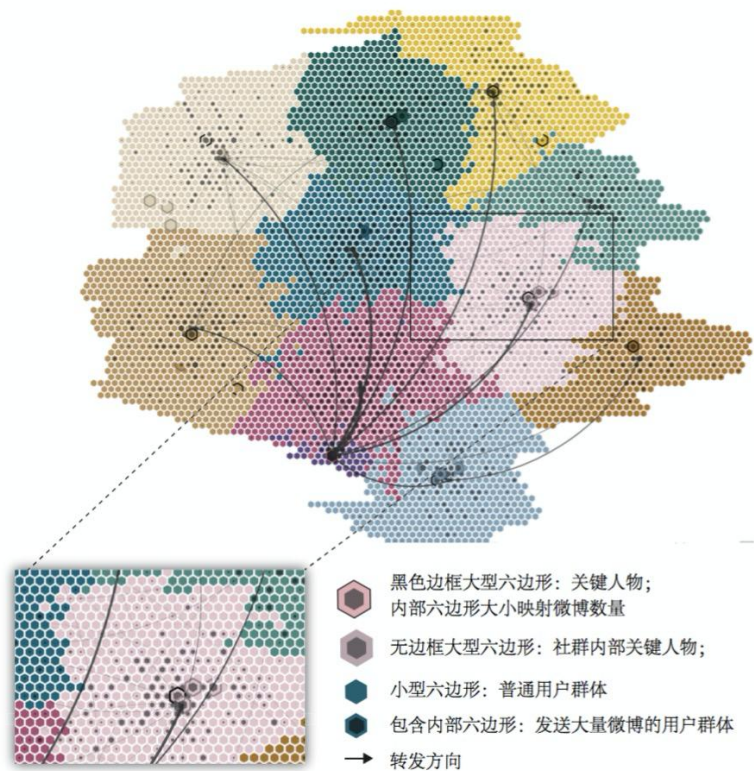
- 可视化+社会学/新闻传播/国际政治
  - 社交媒体传播规律
  - 社交媒体用户画像
  - 舆情演化规律
- 可视化+历史
  - 大数据关联与不确定性推导
  - 名人传记可视化
  - 博物馆可视化交互推荐
- 可视化+中文
  - 音韵学
  - 唐诗宋词





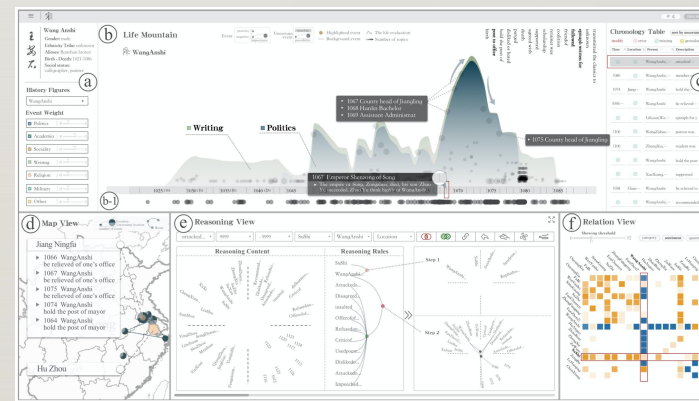
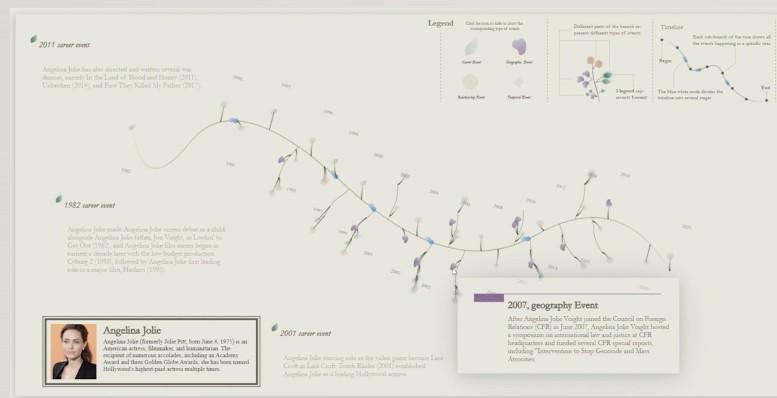
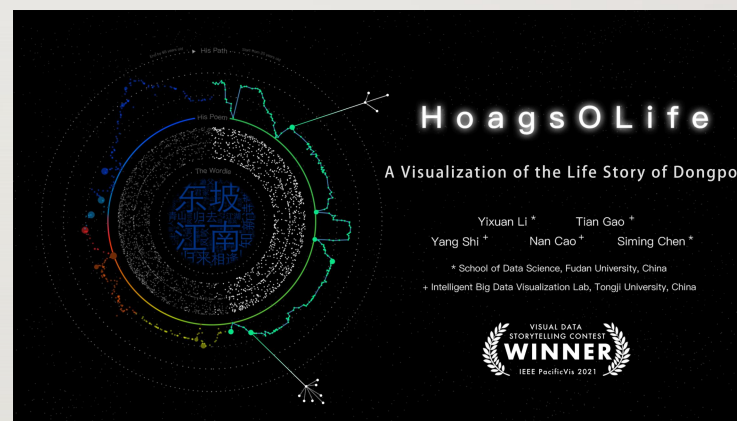
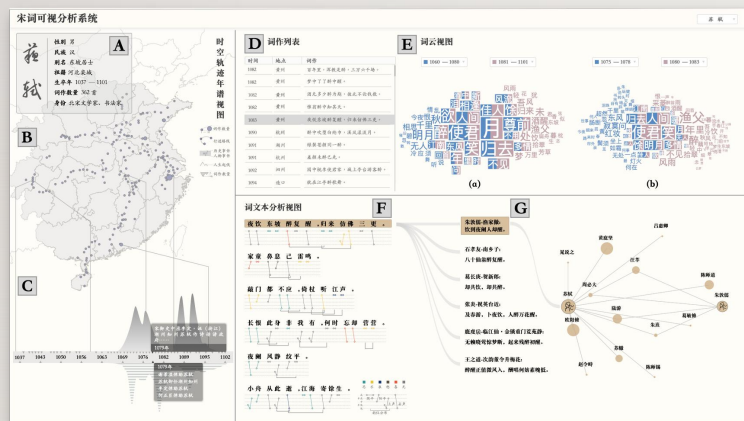


# D-Map (Information Diffusion Map) 算法设计与视觉映射

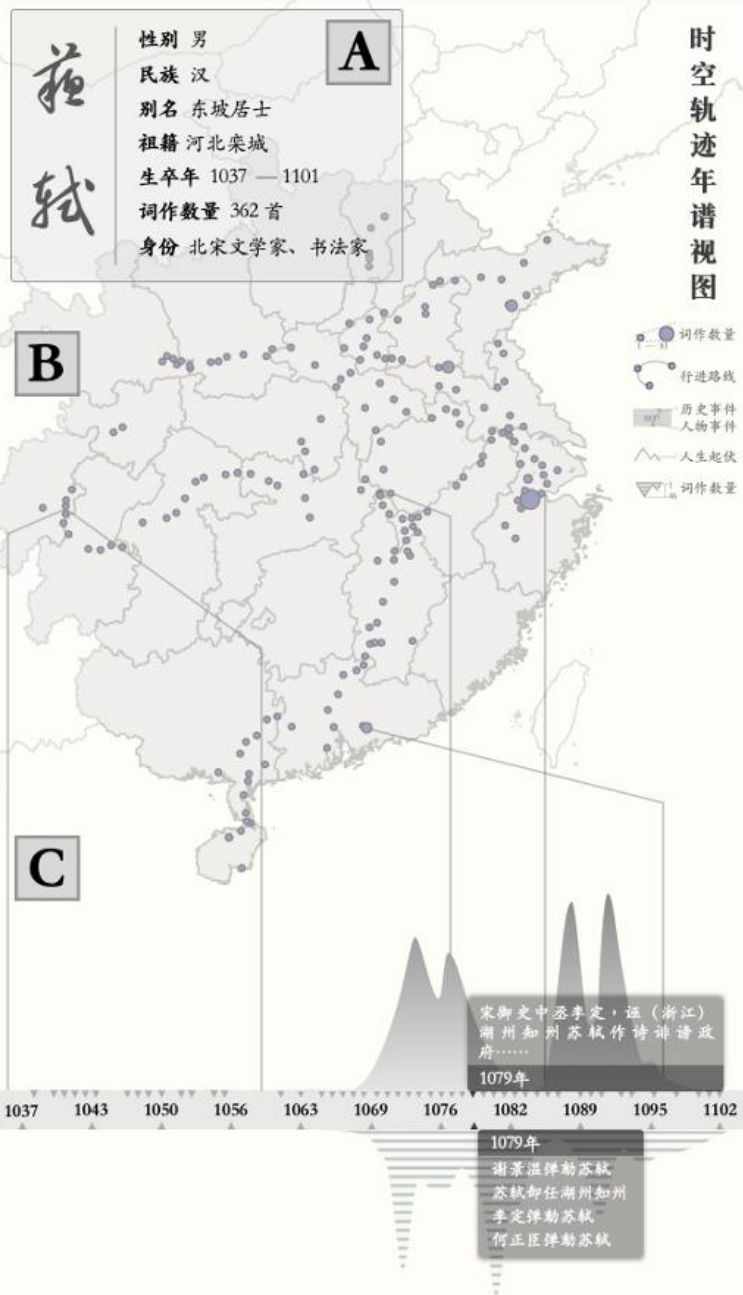


D-Map的颜色映射。四个系列的颜色分别映射每个社群的4个高维空间, 12种高维属性。雷达图展示了选择社群的每个维度上值的比较。

# 数字人文可视化

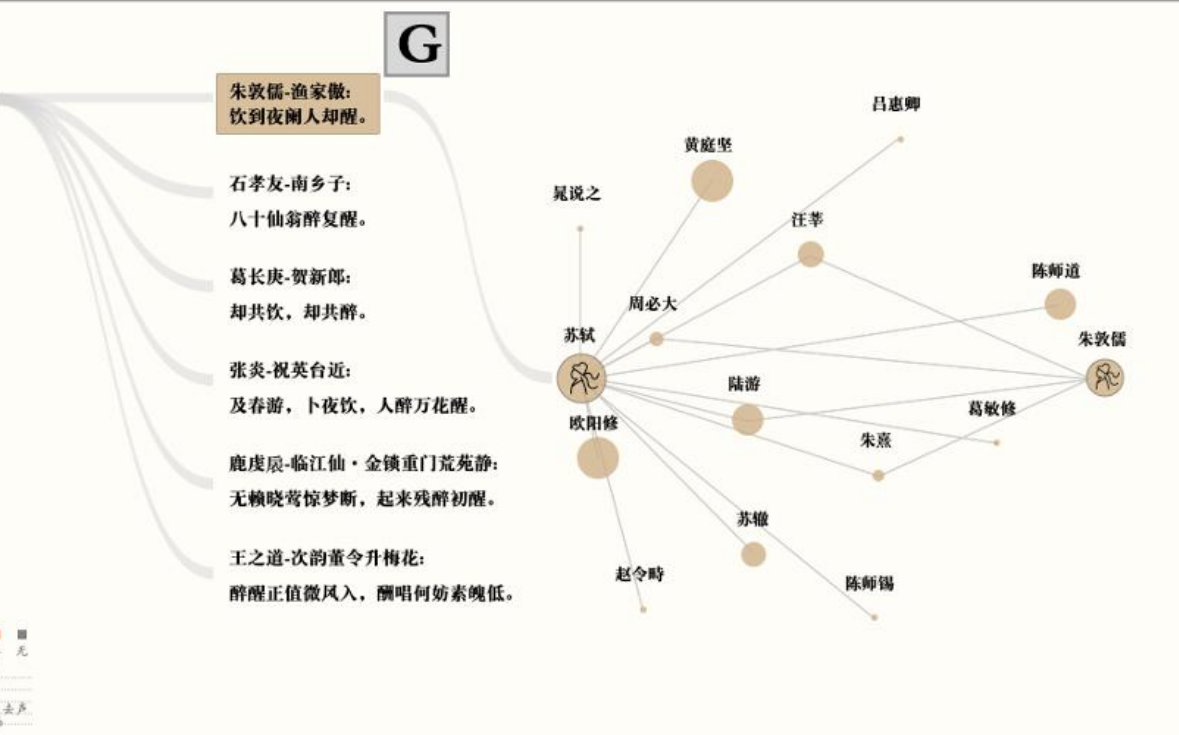
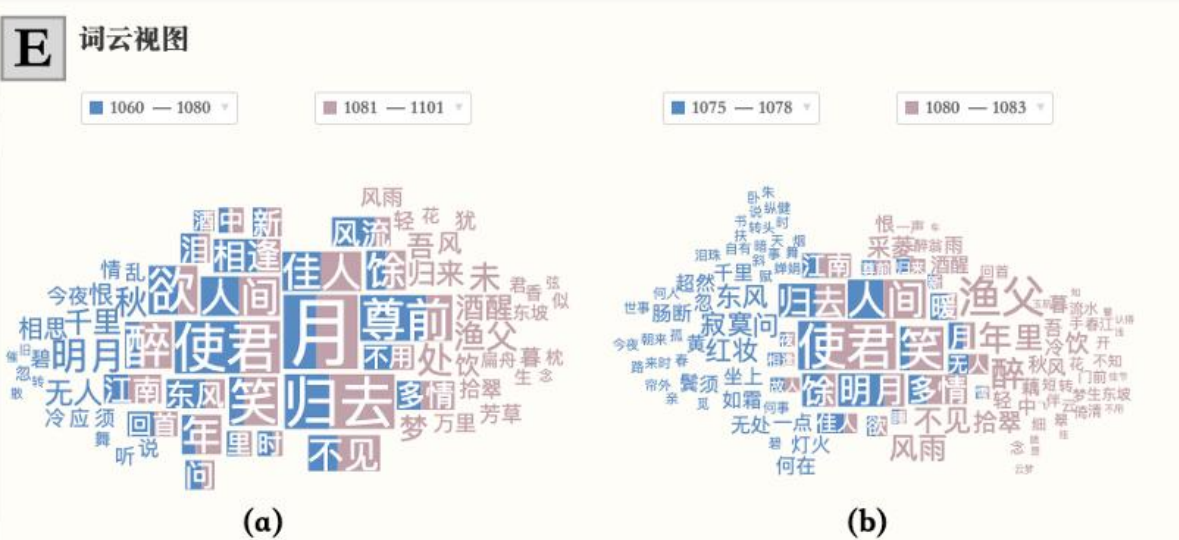






## D 词作列表

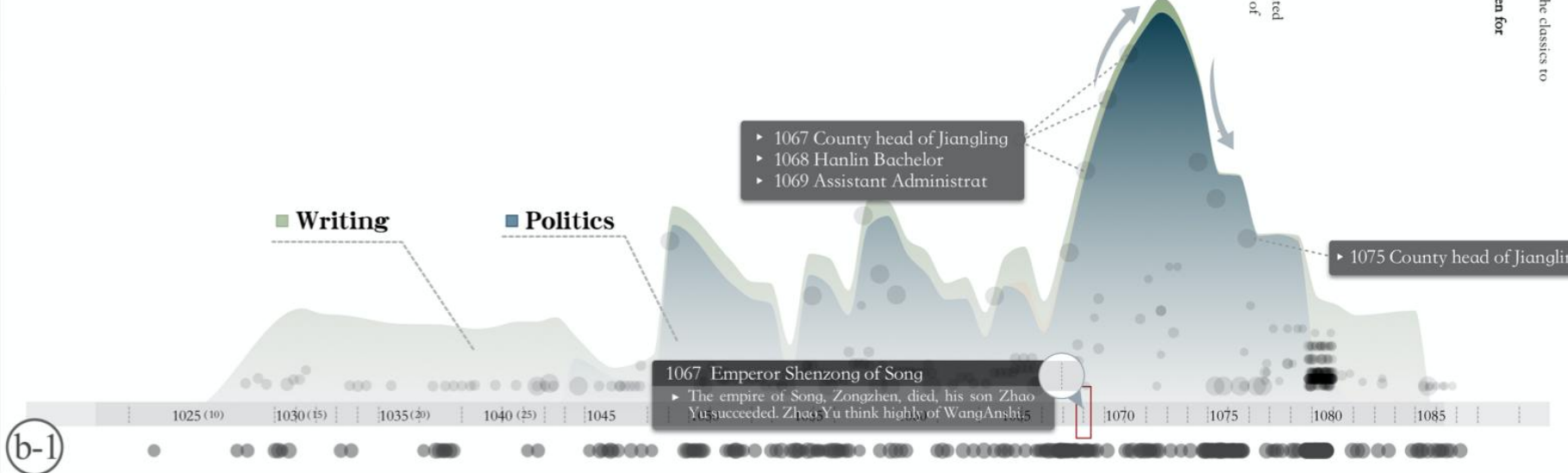
时间	地点	词作
1082	黄州	百年里，浑教是醉，三万六千场。
1082	黄州	梦中了了醉中醒。
1082	黄州	酒无多少醉为期，彼此不论钱数。
1082	黄州	惟翁醉中知其天。
1083	黄州	夜饮东坡醉复醒，归来仿佛三更。
1090	杭州	醉中吹堕白纶巾，溪风漾流月。
1091	湖州	绿鬓苍颜同一醉。
1091	杭州	羞颜未醉已先。
1092	泗州	园中桃李使君家，城上亭台游客醉。
1094	造口	犹在江亭醉歌舞。



Wang Anshi  
 Gender: male  
 Ethnicity Tribe: unknown  
 Aliases: Banshan laoren  
 Birth - Death: 1021-1086  
 Social status: calligrapher, painter

**(b) Life Mountain**  
 WangAnshi

Event: positive, negative, importance  
 Uncertain event: timeline, possibility  
 Highlighted event, Background event, The life evaluation, Number of topics



**Chronology Table** (sort by uncertainty)

Time	Location	Person	Description
		WangAnshi,...	attacked...
1080		WangAnshi,...	member of
1074	Jiang...	WangAnshi	hold the...
1056--		WangAnshi	be relieved...
		LiKuan,Wa...	epitaph for y
1100		WangZishao...	patron was
1100		ZhengXia,...	student was
		WangAnshi	hold the post...
		XueXiang,...	supported
1058	Chan...	WangAnshi	be relieved to
		WangAnshi,...	recommended

**(d) Map View**

**Jiang Ningfu**

- ▶ 1066 WangAnshi be relieved of one's office
- ▶ 1067 WangAnshi be relieved of one's office
- ▶ 1075 WangAnshi be relieved of one's office
- ▶ 1074 WangAnshi hold the post of mayor
- ▶ 1064 WangAnshi hold the post of mayor

**Hu Zhou**

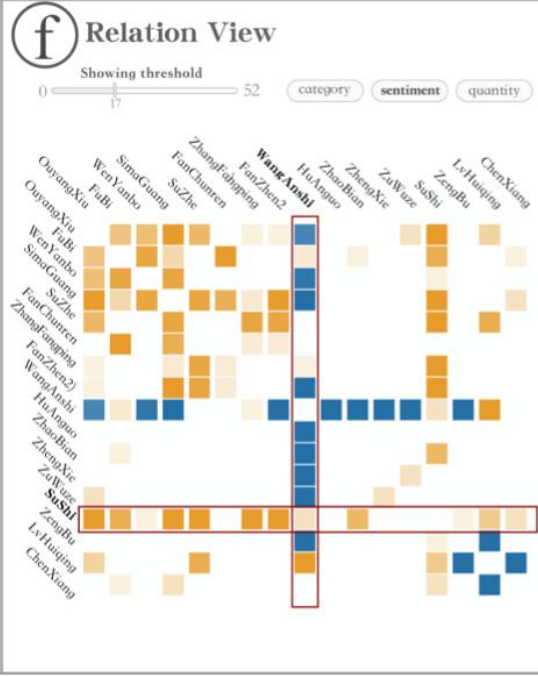
**(e) Reasoning View**

Reasoning Content: SuShi, WangAnshi, Location, attacked..., 9999, -9999

Reasoning Rules: SuShi, WangAnshi, Criticof, Offerofof..., Disliked..., Attacked..., Impeached...

Step 1: WangAnshi... Attacked..., Disagreed..., insulted, Offerofof..., Refusedan..., Criticof, Usedpoetr..., Disliked..., Attacked..., Impeached...

Step 2: WangAnshi... Attacked..., Disagreed..., insulted, Offerofof..., Refusedan..., Criticof, Usedpoetr..., Disliked..., Attacked..., Impeached...





*Visual Reasoning for Uncertainty  
in Spatial-temporal Events of Historical Figures*





### Legend

Click the icon to hide or show the corresponding type of events



Career Event



Geographic Event



Relationship Event



Temporal Event

Different parts of the branch represent different types of events



1 legend represents 1 event

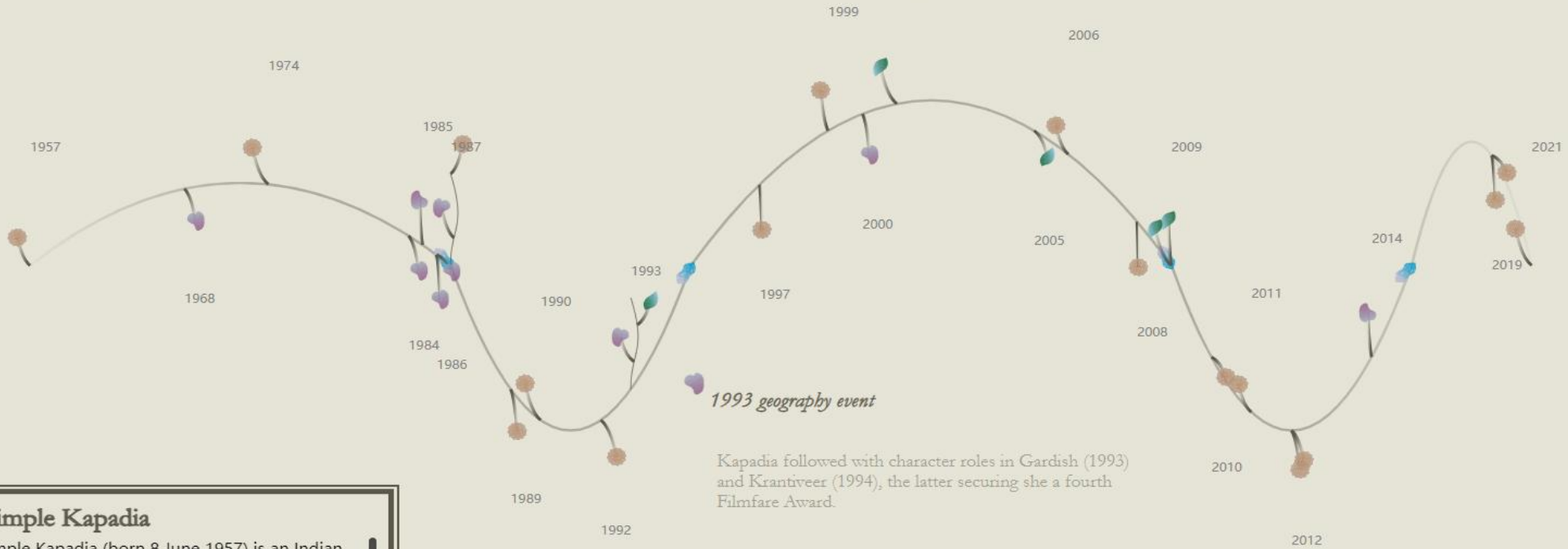
Timeline:

Each sub-branching of the vine shows all the events happening in a specific year

Begin

The blue-white node divides the timeline into several stages

End



### Dimple Kapadia

Dimple Kapadia (born 8 June 1957) is an Indian actress predominantly appearing in Hindi films. Born and raised in Mumbai by wealthy parents, she aspired to become an actress from a young age and received her first opportunity through her

# 数字博物馆的交互

## 博物馆场景



**公众用户：**来自不同教育背景的博物馆观众，他们在探索主题和数据分析领域几乎没有专业知识，希望从探索过程中获得见解，但没有要实现的特定目标。

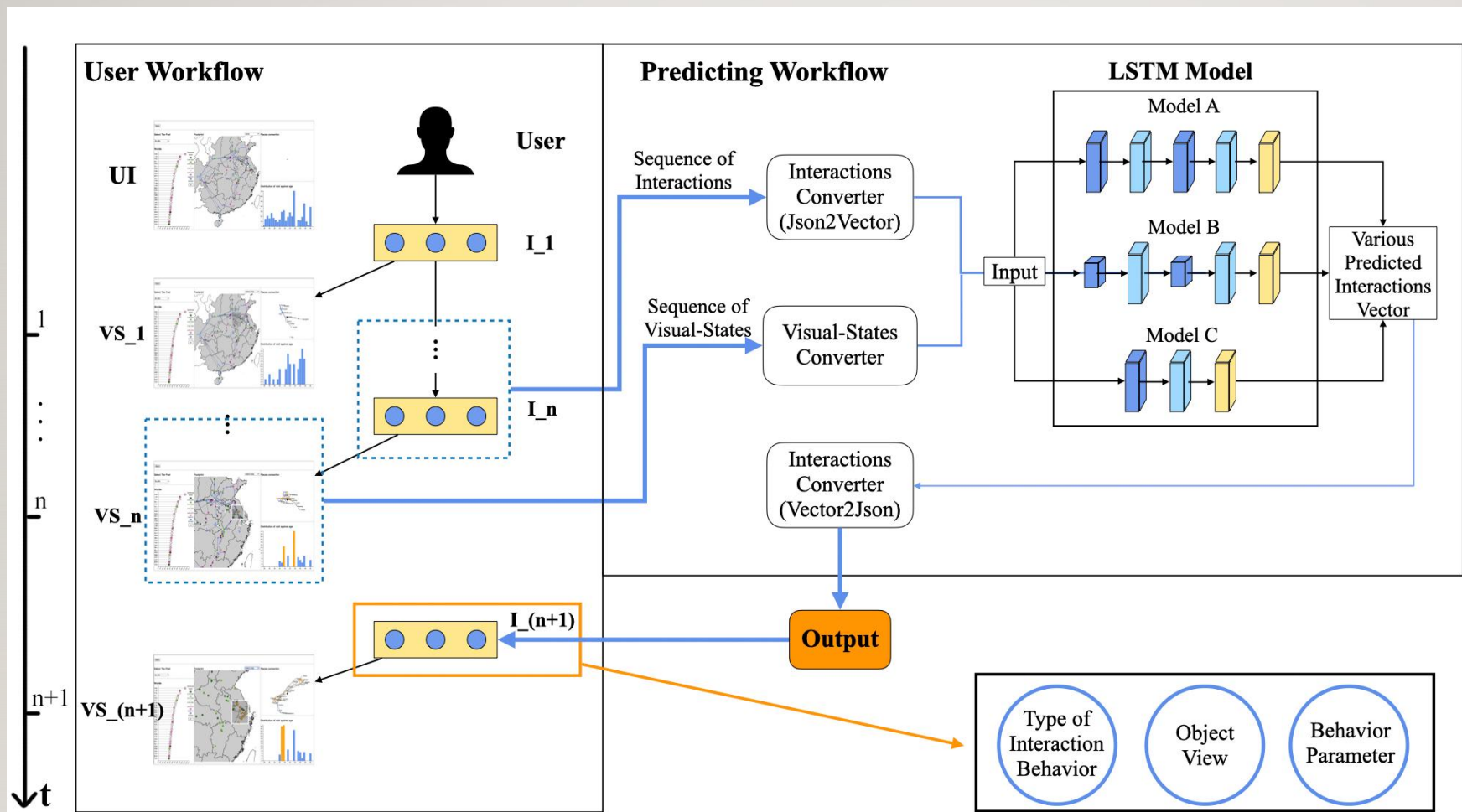
## 研究挑战

**观众背景知识缺乏：**面对具有多视图和交互类型的系统，用户必须决定采用哪种交互类型以及探索哪些数据。然而，这类**决策依赖于用户的背景知识和数据分析能力**。

**多视图和交互类型：**由于背景知识的缺乏，观众可能倾向于专注某一个视图和一种交互类型，导致**单调的数据探索**，以及交互式可视化系统的**利用率不足**。

**没有外部指导：**大多数展览设计没有外部指导。观众通常只有在遇到障碍时才会参考系统的说明或指南。而此类指示可能**无法提供持续的帮助或考虑用户的实际交互**，因此很难改善用户的交互体验。

# 模型框架



模型的基本工作流程

## 模型部件:

交互转换器

视觉状态转换器

LSTM模型

## 模型训练:

模型初始化

模型更新

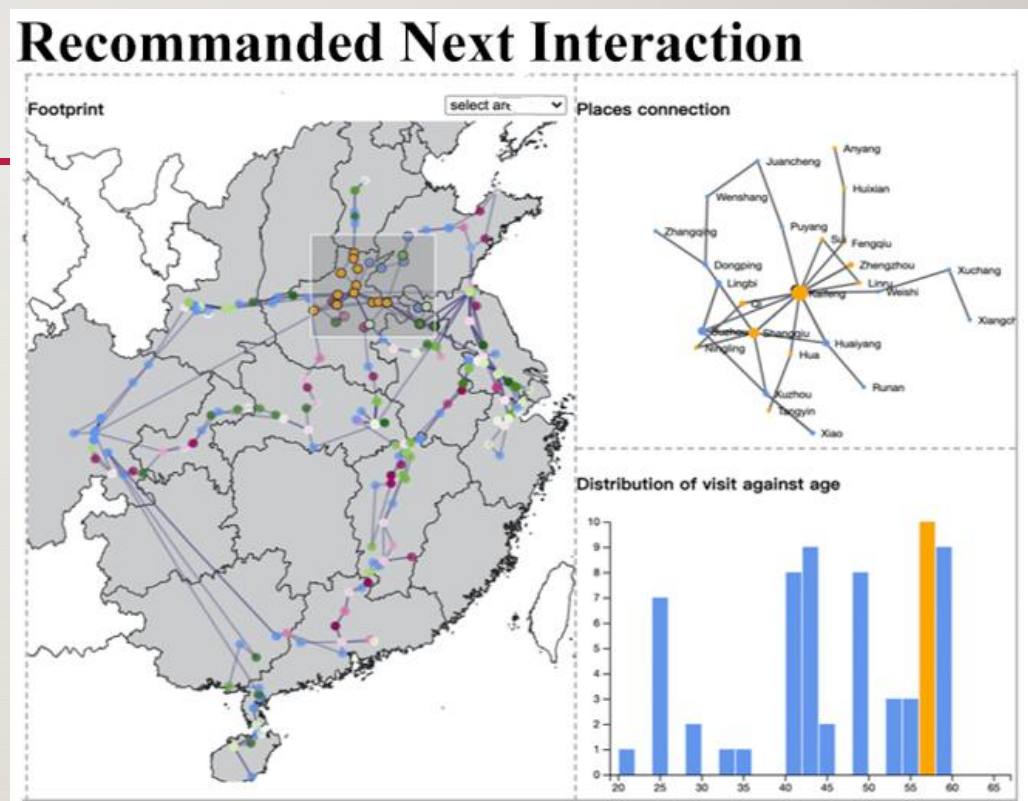
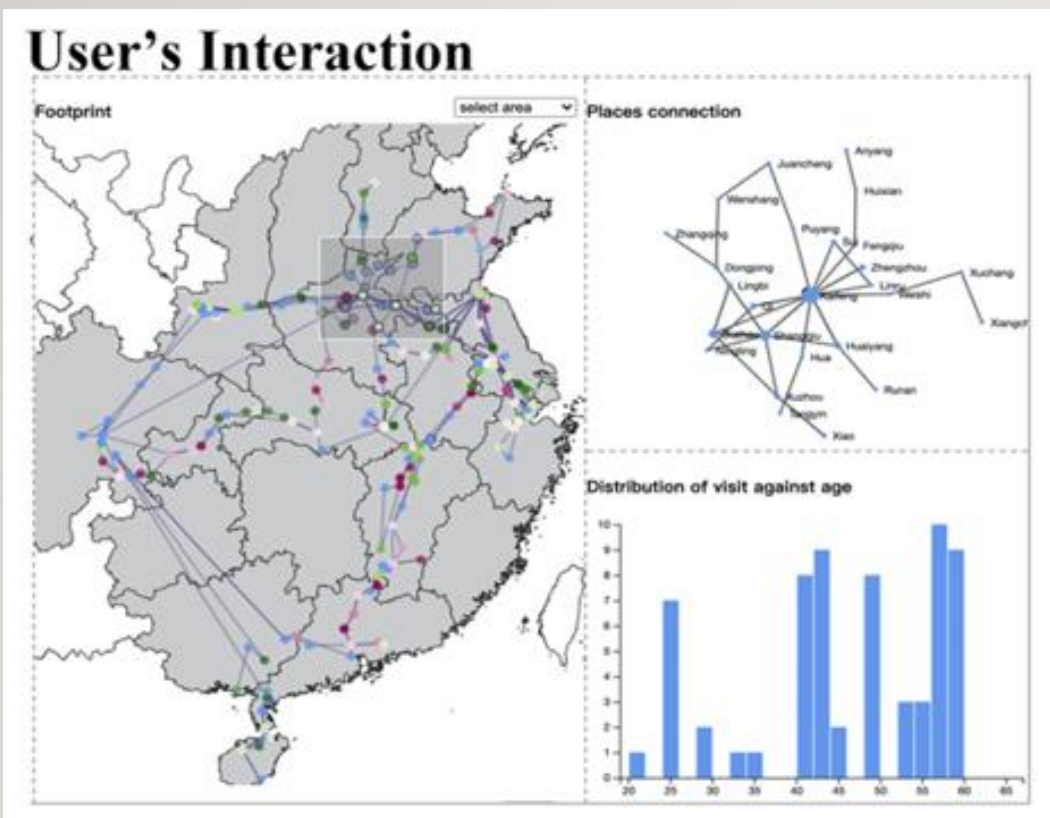
## 人机交互:

交互和视觉状态捕获

推荐结果呈现



# 模型框架 – 推荐结果呈现 (R4)



模型输出  
数值向量

解码

具体的  
交互行为

反馈

系统界面中  
高亮显示

# OLDVIS ONLINE

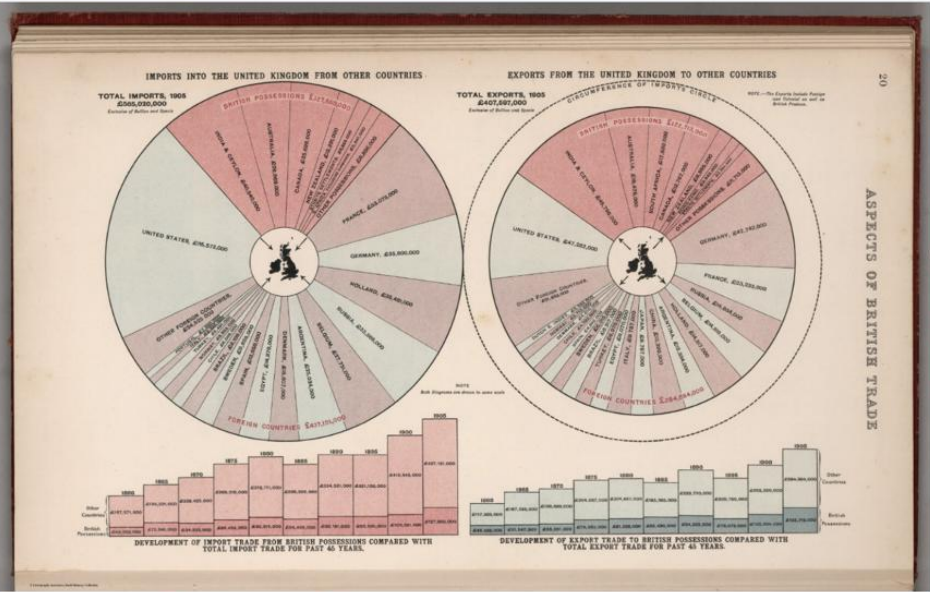
Selectors

UnLabeled Labeled Unsure Search

Entries

1/1 in page labeled #entries: 25563

29. Statistical Diagram: Aspects of British Trade.



author: Bartholomew, J. G.  
 year: 1907  
 source: David Rumsey Map Collection  
 language: English  
 tags: Statistical Atlas, Statistical Diagram  
 view metadata copy metadata url google

Vis  Not Vis  
 Map  Not Map  
 Text  Not Text  
 Unsure  Confident

previous 1 entry next 1 entry goto first unlabeled



Dataset & Code: [github.com/oldvis](https://github.com/oldvis)  
 Paper: [arxiv.org/abs/2308.16053](https://arxiv.org/abs/2308.16053)







# HoagsOLife

A Visualization of the Life Story of Dongpo

Yixuan Li \*

Tian Gao +

Yang Shi +

Nan Cao +

Siming Chen \*

\* School of Data Science, Fudan University, China

+ Intelligent Big Data Visualization Lab, Tongji University, China



# 01 作品概述



专业的人文研究者

计算机科学研究者



共同打造从上古《诗经》、中古《广韵》到近代《中原音韵》的古汉语典籍韵部流变系统，项目链接与之相关的古籍页面。

为**专家学者**和**文化爱好者**提供**第一款交互式可视化**的汉语语音演变的数字人文平台。

<http://124.220.133.205:41082/>

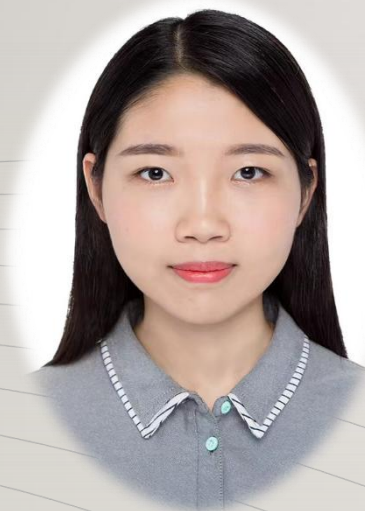
# 01 fduvis 跨学科团队成员-文理跨界

文科博士在读

前端大神

美学大师

架构设计与后端



郑伟莉

复旦大学  
汉语言文字学

3篇C刊录用



梅书鸿

复旦大学  
应用统计

国际性奖项



蓝钰萱

西北大学  
设计

国际性奖项



陈玄

复旦大学  
大数据学院

助理研究员

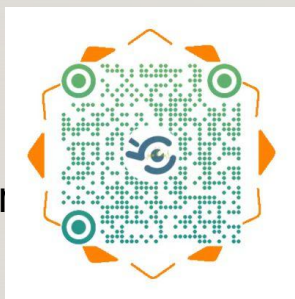


# 复旦大学可视分析与智能决策实验室 (FDUVIS)

[HTTP://FDUVIS.NET/](http://fduvvis.net/)

- 复旦大学可视分析与智能决策实验室 (FDU-VIS) 成立于2020年9月, 我们的研究方向涵盖了可视化与可视分析、人机混合智能、用户行为分析、决策支持与数据新闻故事叙述等, 并在多个应用领域, 包括社交媒体、网络空间安全、时空城市大数据、人文历史数据与金融科技方面有一定的科研成果。

- 联系方式:
- [simingchen3@gmail.com](mailto:simingchen3@gmail.com)



The background features a central horizontal bar with a color gradient from red to yellow. Above and below this bar are decorative wavy lines in a light gray color, resembling sound waves or stylized mountains. The main title is centered on the yellow part of the bar.

# 文藻声韵

古汉语典籍韵部流变系统

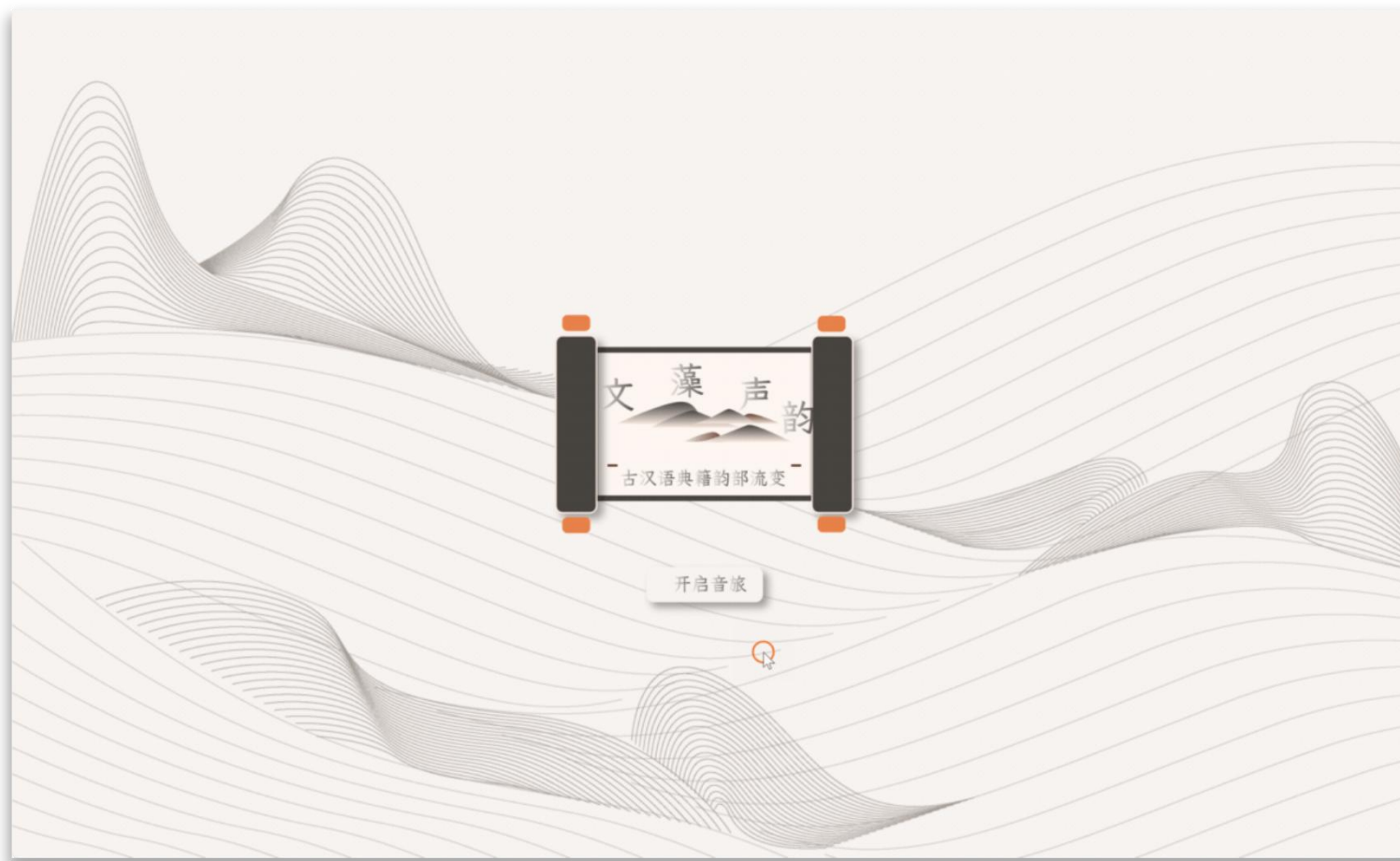


# ONE

作品概述



# 01 作品概述



专业的人文研究者

计算机科学研究者

共同打造从上古《诗经》、中古《广韵》到近代《中原音韵》的古汉语典籍韵部流变系统，项目链接与之相关的古籍页面。

为**专家学者**和**文化爱好者**提供**第一款交互式可视化**的汉语语音演变的数字人文平台。

<http://124.220.133.205:41082/>



# 01 fduvis 跨学科团队成员-文理跨界

文科博士在读

前端大神

美学大师

架构设计与后端



郑伟莉

复旦大学

汉语言文字学

3篇C刊录用



梅书鸿

复旦大学

应用统计

国际性奖项



蓝钰萱

西北大学

设计

国际性奖项



陈玄

复旦大学

大数据学院

助理研究员

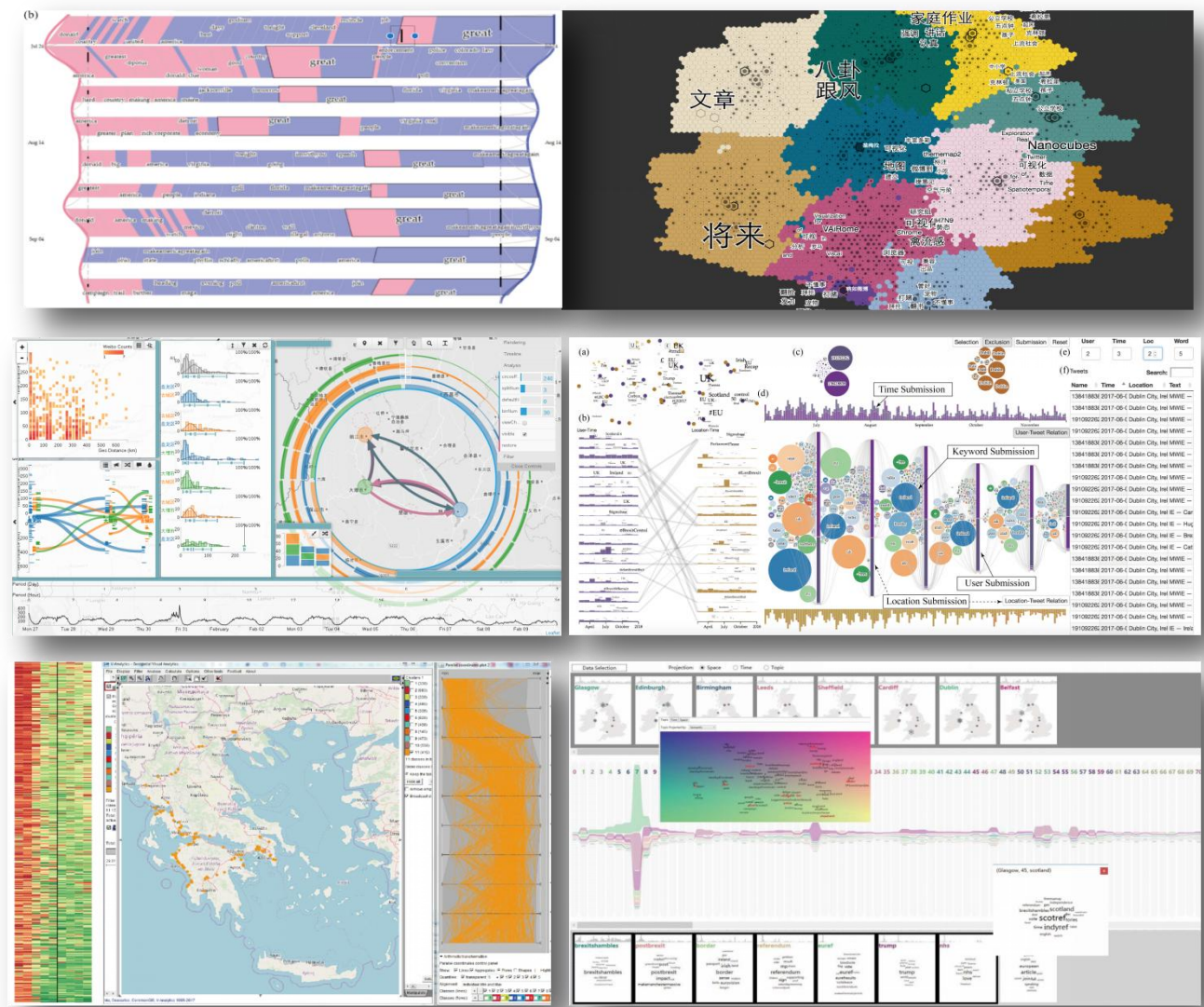
# 01 fduvis 指导教师--陈思明



研究科学家：德国Fraunhofer IAIS

青年研究员/博士生导师：复旦大学大数据学院

扬帆计划获得者、上海市高层次人才。发表论文100余篇，在IEEE Visualization, IEEE TVCG, ACM CHI, CGF, EuroVis等重要国际顶级可视化会议以及期刊上发表30余篇文章。担任多个国际会议的组织委员会和程序委员会成员，包括三大可视化会议PC主审，IEEE PacificVis宣传主席，ChinaVis数据分析挑战赛主席，VizSec程序委员会委员等。他的工作曾获得8次IEEE VAST Challenge数据挑战赛一等奖，以及多个会议最佳论文/海报（提名）奖。





# TWO

问题与挑战



# 古籍探寻的潜在价值

中国古籍卷帙浩繁，我国现存约20万种古籍，超过5000万册（件）。上海图书馆古籍数据库藏137,655册。



## 古籍探寻的任务

**古籍探寻的现状**---目前的ocr、书目统计、电子化数据库并不足够。

**古籍探寻的要求**---由表及里，除了浅层信息，我们还需要探索古籍的中华文明传承价值。



## 古籍探寻的难度

**大数据技术与传统人文两个领域的跨学科合作**

大数据技术注重穷尽性、关联性、探索性。传统人文注重经验、严谨，结论与古文对证。需要结合大数据技术的优势与传统人文的优势。



# THREE

数据分析与可视化设计





# 03 视觉探索工作流程

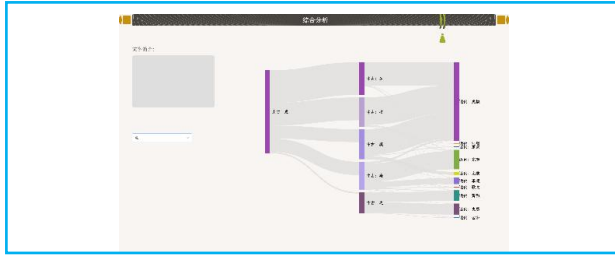
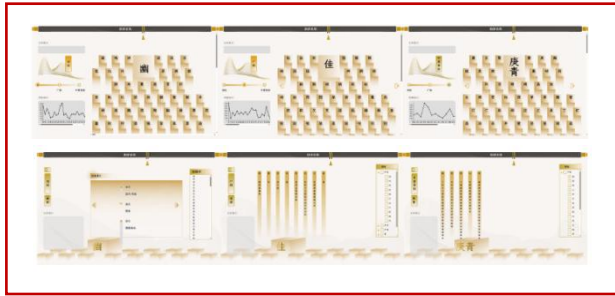
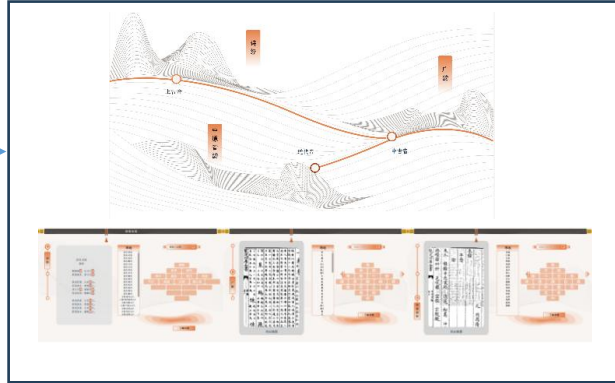
首页：背景知识  
+系统入口

语音全览

韵部系联

文本挖掘

综合流变分析



文本可视化，分古汉语三个阶段

文本挖掘，分古汉语三个阶段

历时分析

上古《诗经》、中古《广韵》、近代《中原音韵》独立呈现、深度挖掘文本信息，进行统计。

综合三个阶段分析，呈现同一韵部的字在不同时代的变化，链汇古今。

由简到难  
层层递进

# 知识普及性

- 1、对汉语历史文化的传播与知识普及，调动起公众对汉语文化的关注与热爱。
- 2、了解汉语音韵的基本知识，并从中感受中华文化的魅力。

# 专业性

- 1、可视化呈现上古《诗经》、中古《广韵》、近代《中原音韵》的全部材料，并且按照学科专业术语分门别类，方便学者取用。
- 2、展示汉语历史的演变源流和音变过程，为学者提供学科研究的新思路。

# FOUR

技术实现





## BACKEND



Node.js  
16.20.2  
LTS

Express



Express.js



Nodemon



REST API

RESTful API

中文繁简转换: chinese-conv

## FRONTEND



Framework: Vite



Vue.js + TypeScript + Sass



UI: Element plus



Visualization: D3.js

AXIOS

HTTP Client: axios



GitHub



PM2

Process manager: PM2

## PLATFORM

F I V E


作品创新





以音韵学为例，针对“音韵学”逐渐成为“冷门绝学”的痛点，打破音韵学为“口耳之学”的局限性，调动视觉感受，增加用户交互式文本挖掘，探索汉语历史发展的魅力。





# 感谢聆听

文藻声韵 - 古汉语典籍韵部流变系统