

2022年5月16日下午，由中国人民大学信息资源管理学院主办的“国际档案学研究前沿”系列讲座第三期暨海外名师讲堂第一讲在线上顺利举行。该系列讲座是70周年院庆系列活动之一。

本期邀请荷兰阿姆斯特丹大学档案学教授Charles Jeurgens主讲“应用人工智能解锁历史数据：以激活荷兰东印度公司文件为例 (Artificial Intelligence as a tool to unlock historical data: Unsilencing Dutch East India Company Records as an example)”，信息资源管理学院副教授马林青博士主持。

Charles Jeurgens博士，现任是荷兰阿姆斯特丹大学档案学系教授、系主任，荷兰国家档案馆战略顾问，拥有社会和经济史及档案学学位。1995年以来在荷兰若干个省及国家档案馆工作过，2004-2016年任莱顿大学档案学专业教授，自2016年起任阿姆斯特丹大学档案学系教授。

讲座主要从以下三个方面展开：荷兰档案学发展的背景、“激活荷兰东印度公司 (VOC) 档案”项目介绍以及人工智能在VOC历史档案中的具体应用。

讲座开始，Charles博士先简要介绍了荷兰的相关情况及荷兰档案学发展历程。



Brief history of the Dutch archival landscape

- Professional archival professionals since 1802 (State archivist)
- 1891: establishment of Society of Dutch Archivists
- Dutch Manual (1898)
- Archives law: 1918, 1962, 1995, (In preparation: 2024)
 - 1 National Archive (The Hague)
 - 12 Provincial Archives
 - Approximately 100 municipal/regional archives
- Several Private Archives: International Institute for Social History, ATRIA, NIOD, IHLIA, The Black Archives, etc



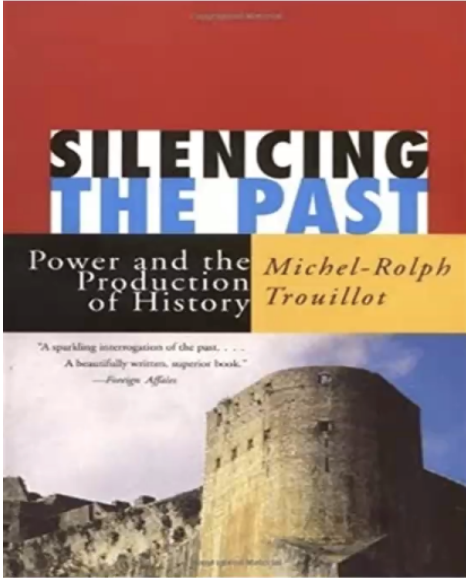
紧接着，Charles博士对“激活荷兰东印度公司档案”项目进行了介绍。该项目旨在通过应用数字方法探索获取历史档案的新途径，重点关注档案去殖民化问题。该项目选取东印度公司历史档案中的遗嘱档案作为研究对象，因为之前的档案工作者仅将这些遗嘱档案中立遗嘱的男性的名字标注出来以供检索，这些档案中所涉及到的女性及土著居民是无法被检索到的，这些人被埋没在这些历史档案中，成为“沉默的人”，而该项目就是要让这些被“沉默的人”能被看到、被发现。Charles博士通过对Michel-Rolph Trouillot的著作《Silencing the Past》中对历史形成中的“沉默时刻”的解读，探究档案工作者和档案机构在此过程中的角

色定位，并强调档案工作者不仅决定了哪些文件能成为档案，而且也决定了哪些文件可以通过著录、索引及相关技术的开发和应用能被获取。在此基础上展开了对东印度公司档案中所包含的“沉默”内容的介绍。

Recording

正在讲话: Charles Jeurgens

Decolonising and Unsilencing



Moments of silencing in the production of history:

1. The moment of **fact creation**
Recording of information: who decides what is recorded, which words are used etc.
2. The moment of **fact assembly**
Which recorded information is admitted to become part of the archives)
3. The moment of **fact retrieval**
What information is used from the archives to compile a narrative)
4. The moment of **retrospective significance**

What interests us most here is the role archivists and archival institutions have played in creating or perpetuating silences.

Archivists not only determine to a significant extent which records are admitted in the archives, but also how **retrieval of information** from the documents is facilitated via archival descriptions, indexes and other technologies designed and used to expedite access to the records from the past.

最后，Charles博士具体介绍了**应用过程**，主要包括：通过HTR手写文本转录软件对东印度公司的遗嘱档案进行转录；构建标注类型，确定其中的实体，实体的属性，及属性的值；对遗嘱档案进行标注，发现在被标注的档案中有61%的档案里包含有女性的名字，30%包含有土著居民的名字。这些标注有利于这些“沉默”人和信息被检索、被利用。该项目将来还希望能有更多利益相关方的参与，以及能提供符合伦理的负责任的获取。Charles博士特别强调我们应注意人工智能应用中的过程透明性问题，提醒我们当我们用标注的语料去训练算法时，应思考这些标注的语料是谁标注的，出于什么目的标注的等问题，以避免人工智能算法带来的歧视。

在提问环节，Charles博士就大家基于讲座内容及相关研究中的问题进行了耐心解答，参会者也对Charles博士的精彩分享表示了感谢。

视频下载链接：

https://pan.baidu.com/s/10yMZ_k_W5Bpcaql1hVu1Cg

提取码：35li

撰稿：张晨文

中国人民大学
数字人文研究中心

Research Center for
Digital Humanities of RUC

公众号账号: rucdh2019

网址: <http://dh.ruc.edu.cn>

邮箱: rucdh@ruc.edu.cn

