

历史研究

上海交通大学新藏地方历史文献的分类法及其依据

赵思渊¹ 汤 萌²

(1. 上海交通大学历史系 上海 200240; 2. 上海交通大学图书馆 上海 200240)

摘 要: 基于上海交通大学新藏地方历史文献整理与研究的需要, 本文讨论地方历史文献分类法的设计及其理论依据。使用内容分析的方法分析近年来历史文献数据库的应用需求及地方历史文献的研究主题, 并说明分类法在文献研究与数据库开发中的意义。针对地方历史文献的分类提出四项原则: 1. 以文书性质为主要依据; 2. 提高批量处理效率; 3. 分类概念间不相容; 4. 提高二次检索效率。依据以上原则提出包含 10 个一级类目的地方历史文献分类法。

关键词: 徽州; 民间文书; 分类法; 数字人文; 数据库

中图分类号: K295; G254 **文献标识码:** A **文章编号:** 1008-7095(2014)03-076-12

DOI: 10.13806/j.cnki.issn1008-7095.2014.03.014

一、问题的提出

目前, 地方历史文献尚无学界共识的定义, 在不同的语境中, 也被称为民间历史文献、民间文书、历史文书档案等。这一文献类型, 主要产生于民间的日常生活, 如果以纸张为物理载体, 则以写本、抄本为主, 不经过出版——即知识再整理的过程, 全部由民间收集的途径发现、获得。这种特点比较接近于档案的性质, 是人群在活动过程中产生的文字以及其他形态材料。

近十年来, 地方历史文献的发现、收藏数量呈井喷态势。徽州、清水江流域、浙南, 是目前发现量最大的三个地域, 总数已超过 100 万件。其中经过前人整理出版的只是极少部分, 得到充分研究的更少。针对这一现状, 历史学界的不少学者呼吁应当大力推进文献的整理、研究, 开创地方历史文献研究的新时代。^①

作为基础工作, 整理出版与数据库开发两方面都是当务之急。两项工作的基础则是收藏文书的整理。针对地方历史文献的搜集、保藏与整理, 近十年来学者们大力提倡“归户”的方式, 以反映文书之间的有机联系。^② 在近年的文书归藏与出版中, 多依照归户的方式排架与编辑, 这逐渐成为学界认同的文献整理方法。

与此同时, 新的问题出现了。第一, 大批民间历史文献入藏后, 虽可以按照归户方式排架,

收稿日期: 2013-12-27

基金项目: 中国博士后科学基金(2013M540354), 上海交通大学校内文科创新项目(12QN08)

作者简介: 赵思渊, 上海交通大学人文学院历史系博士后; 汤 萌, 上海交通大学图书馆馆员。

① 梁勇、郑振满、郑莉:《新史料与新史学——郑振满教授访谈》,《学术月刊》,2012年第4期,155-150页。

② 刘伯山:《徽州文书的整理与遗存》,《徽州文书》第一辑,桂林:广西师范大学出版社,2005,《序》;《徽州文书的遗存与特点》,《历史档案》,2004年第1期,122-126页。

但登记和编目时仍需要对文献的性质进行标示。在徽州文书领域,《徽州文书类目》与《徽州历史档案总目提要》都是针对这种需求而制定。^③ 第二,建设以地方历史文献为主体的全文检索(full-text archives)数据库,提供分类检索功能是学界普遍提出的需求,也是今后地方历史文献整理的大趋势。因此,提出一种适用于数据库检索的分类法是很有必要的。台湾数位人文图书馆(THDL)已经在这方面做出有益的探索。^④

地方历史文献具有特别的性质,需要从对其性质的理解出发,提出新的分类法,作为归藏目录以及数据库检索的参照。当代出版品流通使用的《中图法》以及古籍图书使用的四部分类法、刘国钧《中国图书分类法》具有共同的特点,它们主要是针对有组织的知识载体的分类体系。地方历史文献的特色恰在其非组织化,不能适用于以上分类法。

另一方面,地方历史文献是一个动态的概念,其收藏数量与涉及的内容,都将持续扩增。因此,为地方历史文献编制分类法,需要在设计时提供较高的可扩展性,并不仅仅着眼于已入藏文献。2007年以来,上海交通大学一直在收集主要来自浙江、福建、皖南的地方历史文献,迄今超过30万件。针对这批文献,我们尝试探索分类法的设计及其理论依据,供学界讨论。

二、地方历史文献整理与分类的发展

近十年来,三个区域的发现构成了地方历史文献收藏的主体:徽州、清水江流域、浙南。其中,徽州是地方历史文献发现、收集持续时间最长的区域。从20世纪50年代起开始收藏,至2000年前后,存世量据估算在40万件^⑤,今天这一数字则至少需刷新至60万。徽州文书分别收藏于十几个机构,整理最深入,著录、分类的标准也最复杂。

清水江流域文书是近十年来发现数量增长最快的文书群。它位于贵州的苗族聚居区,杨有庚最早开始收集当地文书,并与武内房司、唐立合作出版文书汇编。^⑥ 张新民、张应强等也先后整理、研究清水江文书,发现数量最初为几千件,至今估计存于民间的可能有50余万件。^⑦ 该文书群的发现和研究,更新了人们对明清以来西南地区商业发展和族群关系的认识。

浙南文书的大量发现是在2007年之后。上海交通大学地方文献中心率先在石仓(属今丽水市松阳县)发现大量保存完好的契约文书,由此带动了清代乡村地权与手工业研究的小高潮。^⑧ 此后,以石仓为据点,上海交通大学地方文献中心陆续收集浙南民间文书近12万件,目前已转至上海交通大学图书馆保藏整理。另外,藉由编纂《温州通史》的契机,吴松弟、冯筱才、李世众等在温州等地域也发现、收集了大批民间文书。

1. 文书分类的数据库应用:基于内容分析法

历史文献数据库的建设在近五年来成为学界热点。在CNKI数据库中使用主题、篇名、关键词三种检索方式,用词汇“数据库”分别与“徽州文书”、“清水江”、“民间文书”、“历史文献”关

③ 严桂夫编:《徽州历史档案总目提要》,合肥:黄山书社,1996;王钰欣等编:《徽州文书类目》,合肥:黄山书社,2000。

④ 卢家庆:《台湾古契书自动分类与依分类定义契书角色》,台北:台湾大学硕士论文,2008。

⑤ 刘伯山:《徽州文书的遗存与特点》,《历史档案》,2004年第1期,第122—126页。

⑥ 唐立、杨有庚、武内房司编:《贵州苗族林业契约文书汇编(1736—1950)》,东京:东京外国语大学アジア・アフリカ言语文化研究所,2001。

⑦ 张应强、王宗韶编:《清水江文书》,桂林:广西师范大学出版社,2007;张应强:《木材之流动:清水江下游地区的市场、权力与社会》,北京:三联书店,2006;张新民:《走进清水江文书与清水江文明的世界——再论建构清水江学的题域旨趣与研究发展方向》,《贵州大学学报(社会科学版)》,2012年第1期,第40—47页;《清水江文书系列·天柱文书》,南京:江苏人民出版社,2013。

⑧ 曹树基、阙龙兴、潘星辉编:《石仓契约》,杭州:浙江大学出版社,2011。

联检索,检得 68 笔文献,但不仅限于讨论地方历史文献。图 1 统计显示,2007 年以来,学界对历史文献数据库建设的关注显著增多。

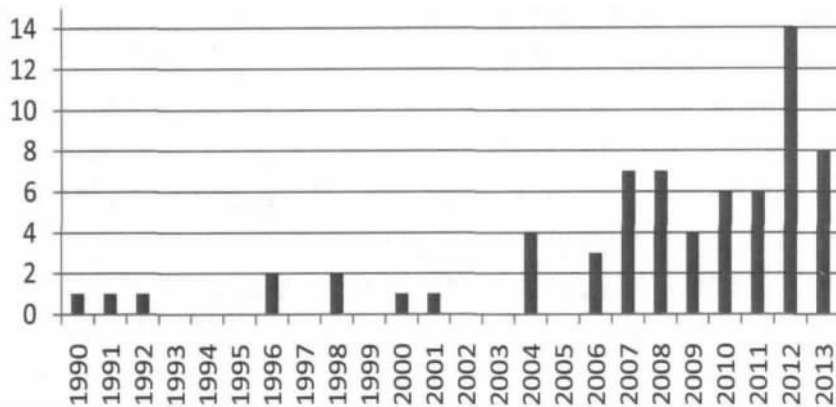


图 1 历史文献数据库建设研究文献发表时间分布

资料来源:CNKI 数据,2013 年 12 月 6 日访问。

统计显示,在这些文献中,67.65%认为需要在历史文献数据库中提供分类检索,17.65%讨论了分类原则(见表 1),14.71%提出了具体的分类体系。已经提出的分类原则有四种:1. 使用《中图法》、《中国档案分类法》或四部法;2. 根据文献内容分类;3. 根据文献物理形态分类;4. 根据文献特点自行编订。

12 篇涉及分类原则的文献中,使用第一种原则的文献计 8 篇,都是讨论地方历史文献数据库建设,其中有的数据库主要内容是古籍、现代出版史料汇编、现代出版的地域研究论著,适用于中图法;有的数据库内容是民国以来档案,因此适用《中国档案分类法》。使用内容分类原则的文献 2 篇,分别讨论徽州文书与清水江文书数据库建设,主要理论依据是《徽州文书类目》提出的分类方法。依据物理形态分类的文献 1 篇,该数据库不仅包含纸质文献,还收录碑刻、民间工艺及影音资料,因此按照载体形态分类。根据文献自身特点提出分类原则的文献 1 篇,研究历史时期农业灾害信息数据库,这是一个专题库,按灾期、灾区、灾种、灾情分类编排资料。

表 1 历史文献数据库建设中分类应用需求分析

文献主题	文献数	占比(%)
未讨论分类法	22	32.35
提出分类应用需求	46	67.65
提出分类原则	12	17.65
提出分类体系	10	14.71
总篇数	68	100.00

资料来源:CNKI 数据库,2013 年 12 月 6 日访问。

10 篇涉及分类体系的文献中,有两篇是在《中图法》中添加号段,改编为适应馆藏的分类法,都是针对以书籍为主的馆藏。7 篇按照自订的内容类别分类,都是针对专题数据库,或是根据该专业领域已有传统分类(灾害史、医学史),或是使用现代概念做一般性分类,分为政治、经济、文化等。两篇讨论民间文书的文献都参照《徽州文书类目》制定按内容分类的分类法。

也就是说,近二十年来国内历史文献数据库发展主要有三种模式:1. 以特定地域为中心,包罗该地域的所有不同类型文献的地方历史文献数据库;2. 针对特定知识领域的专题数据库,如

医学史、灾害史等;3. 针对特定文献形态的专题数据库,如民间文书、碑刻、牌坊等。设置分类检索的方案则有四种:《中图法》分类;内容分类;物理形态分类;文献性质分类。数据库类型与分类方法之间的关系(见表2)。

表2 历史文献数据库分类原则适用性与应用案例

数据库 分类原则	地域文献数据库		专业知识数据库		特定文献形态数据库	
	适用性	应用	适用性	应用	适用性	应用
《中图法》分类	○	○	○	—	—	—
内容分类	○	○	○	○	○	○
物理形态分类	○	○	—	—	○	—
文献性质分类	—	—	○	○	○	—

注:“适用性”指某一分类原则是否可以对某种数据库进行分类检索,“应用”指某分类原则是否已经应用于某种数据库的案例。“○”表示可适用或有应用案例,“—”表示不可适用或无应用案例。

以上用文献计量的方法说明,历史文献数据库建设中,设置分类检索是一种主流的应用需求,使用什么样的分类原则与分类体系,与数据库封装内容密切相关。目前民间文书数据库建设的主流分类检索是参照《徽州文书类目》制定的内容分类法,下文将检验内容分类方式应用于地方历史文献的适用性。

2. 地方历史文献分类法的早期发展:一个比较分析

20世纪90年代初,徽学界提出了两种分类方法。《徽州文书类目》是以中国社会科学院历史研究所、经济研究所收藏徽州文书为基础,编制文书目录。该分类法为三级结构,每个文书编号的第1位代表文书类型,2、3位表示朝代,4、5位表示年号,实际可视作两个复分表。一级分类基本参照周绍泉先生的研究修改而成^⑨,依据历史学研究主题分类。

《徽州历史档案总目提要》是安徽省档案馆、南京大学图书馆、中国社会科学院等单位收藏徽州文书编制的联合目录。该分类法没有说明分类层级,实际为二级分类,亦未设置复分表。虽然在目录制订的说明中,根据文书产生的来源,分为政府来源和社会活动来源两个方面介绍,但分类表却设置为7个类别:政务、土地、赋税、工商、宗法、文化、邮政,其下共有二级分类90个。

《类目》与《总目提要》是大陆学界较早编制的民间文书目录与分类体系,他们都是根据文献叙述内容进行分类,代表了当时大陆学界的理解。比较分析两种目录,有助于理解地方历史文献整理的发展,也可以说明文献叙述内容作为分类依据的不足。

表3 《徽州文书类目》、《徽州历史档案总目提要》一级分类比较

《徽州文书类目》	《徽州历史档案总目提要》
土地关系与财产文书	土地类
赋役文书	赋税类
商业文书	工商类
宗族文书	宗法类
官府文书	政务类
教育与科举文书	文化类
会社文书	邮政类
社会关系文书	
其他文书	

资料来源:《徽州文书类目》目录;《徽州历史档案总目提要》目录。

^⑨ 周绍泉:《徽州文书的分类》,《徽州社会科学》,1992年第2期,第137—144页。

两种分类法一级分类指向内容相同的有 5 个,其中“商业文书”与“工商类”之间不完全重合。分类的涵括性方面,因为设置了“其他文书项”,《徽州文书类目》优于《徽州历史档案总目提要》。但“社会关系文书”这一项从分类逻辑考虑,是不合理的。如果将“社会关系”理解为社会活动中的人、组织之间的关系的的话,一级分类的其他项目都属于“社会关系”,因此“社会关系”这一分类名称是不精准的。

“社会关系文书”下属的二级分类 15 个:

8.1 兰谱;8.2 信札;8.3 乡规乡约;8.4 伏罪甘罚文约;8.5 投主服役文约;8.6 甘结(社会交往中的义务保证);8.7 公益活动文约;8.8 劝息约、劝息杜讼文约;8.9 限约;8.10 术数文书;8.11 婚嫁文书;8.12 投师文书;8.13 卖身约;8.14 同心赴讼合同;8.15 雇佣文约。

《类目》没有说明将这些内容置入同一个分类项内的理由,从内容看,8.1、8.2 与其他 13 项没有共同点,8.4、8.8、8.9、8.14 与诉讼有关,8.3、8.6、8.7 与村落生活有关,8.10、8.11 与人生仪式有关,8.5、8.12、8.13、8.15 与人身约束有关。

也就是说,8. 社会关系文书所属各二级分类并不具有共同点,不符合将相似事物集群的分类法原则。笔者推测设置这一分类项的原因,是将所有已知文献内容中由社会活动所产生,但无法置入前 7 项分类的内容合并为一个一级分类。

《提要》的一级分类均有确切含义,前 6 项分类也具有足够弹性,可以涵括大部分文书内容。但“邮政类”作为一级分类与前 7 项比较而言太过单薄。“邮政类”包含 4 个二级分类:书信、邮政回执、公文封、腰牌,后三者都应属于“政务类”,没有单列为一级分类的必要,书信则似乎不完全能被“邮政”这一词汇所概括。

两书相同的一级分类中,划分二级分类的逻辑也不相同。以土地类为例,《类目》主要根据土地交易形式划分,但一种普遍存在的土地交易形式即“加找”似乎没有列入。《提要》根据标的物与文献形态划分,实际上同一份契约的交易内容经常同时包含田地、房屋、坟墓等,这类情况如何处理,《提要》也没有予以说明。

表 4 《徽州文书类目》、《徽州历史档案总目提要》土地类二级分类比较

《徽州文书类目》	《徽州历史档案总目提要》
土地、房屋、耕牛等事产买卖文书	田契
税契凭证	地契
土地、房屋、耕牛等事产典当文书	山契
清业清白合同	坟契
土地、房屋、耕牛等事产租佃文书	房契
批契及财产处置文书	佃仆文约
对换文约	置产簿
买卖、典当山林力盆和田地皮文书	盘保簿
买卖、典当山林力盆和田地皮文书	收租簿
共业合同	丈量图
退契、借契、还契	业主执照
认契	
借贷文书	
分业合同	

资料来源:《徽州文书类目》目录;《徽州历史档案总目提要》目录。

3. 数据库环境下的分类原则

分类法是依据某种逻辑聚合相似事物。^⑩ 针对地方历史文献资料群,可以建立的语法型(syntactical)集合关系有三种:物理形态、文书性质、文书叙述内容。《类目》与《提要》都是编制

^⑩ Brian B. *Theory of Library Classification*, London: Clive Bingley Ltd. 1979, p9.

目录过程中形成的分类法,着眼于馆藏文献的组织,因此都采用文书内容的逻辑设置分类。上文分析已经说明,按照文书叙述内容分类,同一件文书可能被置入不同分类位置,如果应用于批量编目会出现混乱。另外,本文制定分类法,是针对一个不断扩展的资料群,因此需要分类法具有可扩展性,不能仅针对已入藏文献。从这一出发点考虑,文书叙述内容作为分类逻辑也是不适用的。

物理形态是操作最为简便的一种分类逻辑,广义的地方历史文献包含多种不同物理形态,可归纳为四种:纸质、其他纤维制品(主要是纺织物、木制品)、不可移动物、影音资料。分类法是为编目与后续研究提供参照,在史学研究中,多数研究议题不涉及对文献物理形态的讨论,因此这不是一种有效的分类逻辑。

文书的性质,是指文书产生时被其书写者所赋予的意义,及其在特定的时空背景中承载的功能。在中国文献学的传统中,编制目录、设置部类的原则一向是“辨章学术、考镜源流”,即根据书籍承载的知识的学术脉络设置部类,由此反映书籍的功能与性质。^① 地方历史文献是一种非组织化的知识载体,不存在学术传承脉络。但民间文书有特定的产生源流和社会功能,这是本文设置分类法的首要依据。此外,本文编制文献分类法是以数据库应用为导向,基于这种考虑,分类法的制定有如下原则:批量处理时的效率、分类概念间的不相容性、二次检索的效率。

所谓批量处理的效率,是指在大规模处理文献时,人工判定快速有效,并且今后能应用于机器判定。要满足这一要求,分类法中就不能包含需要人工通读全文才能获取的信息,或者机器无法通过一组关键词识别的信息。分类概念间不相容性是分类法定义的自然延伸,是文献整理不紊乱的基本前提。研究者使用关键词在数据库中检索时,要依靠分类法对检索结果进行排列和二次检索,因此分类集合的设置需要考虑相关研究领域的重要主题。

后文列出的分类法,正是根据以上原则编制的,如何验证该分类法的适用性呢?批处理效率利用编目抽检结果检验。以出版的《石仓契约》第三辑为例,在上海交通大学新藏地方历史文献资源编目登记中,设置7个栏位,其中资源类型即分类法,应用下文将要提出的方案。表5、表6统计显示,资源类型栏登记错误仅1个,说明该方案在编目登记中有效。

表5 上海交通大学新藏民间历史文献编目登记质检统计

登记栏位数	登记条目数	总错误数	总错误率(%)	资源类型错误数	资源类型错误率(%)
7	2296	21	0.91	1	0.04

资料来源:《上海交通大学新藏地方历史文献编目登记》;《上海交通大学新藏地方历史文献编目质检统计》,上海交通大学图书馆,内部资料。错误率=错误数/栏位数*100%

表6 上海交通大学新藏地方历史文献编目登记错误统计

	错误数	占比(%)
描述栏	5	23.81
事由栏	0	0.00
标的栏	14	66.67
金额栏	0	0.00
附注栏	1	4.76
资源类型栏	1	4.76
文献材质栏	0	0.00
合计	21	100.00

资料来源:《上海交通大学新藏地方历史文献编目登记》;《上海交通大学新藏地方历史文献编目质检统计》,上海交通大学图书馆,内部资料。

① 张舜徽:《中国文献学》,郑州:中州书画社,1982,136页。

分类概念的不相容性从两方面予以确保,第一是对每个一级类目的定义给予明确的边界;第二采用列举排除法,列举说明每个一级类目中可能发生混淆的内容。检验二次检索效率采用关键词抽样法。地方历史文献主要应用于历史学研究,本文选择 5 个以地方历史文献为主要领域的研究机构,从中选取 17 名学者,搜集他们 2000 年以来发表的以地方历史文献为主要材料的论文,提取研究主题,编制民间历史文献研究关键词列表。在编制分类法时,一级、二级分类应当在不违背以上原则的情况下涵盖该关键词列表。

表 7 地方历史文献研究关键词提取来源

机构	选取学者数(人)	选取论文数(篇)
中山大学历史人类学研究中心	8	94
厦门大学民间历史文献研究中心	3	26
香港中文大学历史系	3	17
上海交通大学地方文献研究中心	2	24
中国社会科学院徽学研究中心	1	7
合计	17	168

资料来源:CNKI 数据库,2013 年 12 月 6 日访问。

从以上来源共提取关键词 840 个,其中指涉具体事物的地名、人名、时代、文献名等根据依文书性质分类的原则予以剔除,其余的关键词是抽象概念或专有名词,共计 399 个,其中词频最高的 10 个关键词如表 8。

表 8 地方历史文献研究关键词词频表(抽象概念与专有名词类)

关键词	词频(次)
宗族	12
市镇	9
族谱	8
水利	7
城乡关系	6
民间信仰	5
赋役	5
地方社会	5
赈济	4
卫所	4

资料来源:CNKI 数据库,2013 年 12 月 6 日访问。

以上 10 个关键词可分为 4 组:A 民间信仰、城乡关系、地方社会;B 宗族、市镇、卫所;C 水利、赋役、赈济;D 族谱。A 组是社会科学意义上的分析概念,B 组是指涉具体社会组织、政治组织与聚落形态的专有名词。按照文书性质这一首要原则,以上都不适合作为分类依据。C 组是政治制度及其对应的社会活动,构成文献产生的来源,也表达文献的社会功能,适用为本文的分类依据,但水利、赈济作为单独类别包含文献较少,并且在批量处理时难以辨识,因此在后文分类中没有采用。D 组族谱本身就是一种独立的文献形态,其来源与社会功能都很明确,易于辨

识,在后文作为一个独立类别。

三、基于上海交通大学收藏地方历史文献的分类法设计

根据以上原则与考量,设置一级分类10个:契约、账本、赋役、行政、诉讼、信函、日用类书与工具书、家礼、宗教、戏剧。下面对各类进行详细说明和细目设置。

A. 契约。契约是徽州文书中最早发现,亦是存量最大的一类文献,据估算占已发现徽州文书的60%左右。^⑫在对明清民间文献的研究中,狭义的契约一般指土地、房屋或其他物品交易双方或多方订立的交易完成证明。“契”的初始含义是“大约”,段玉裁解释为“约,取缠束之义”,又引《周礼》:“《大郑》云:书契,符书也。《后郑》云:书契,谓出予受入之凡要。凡簿书之最目,狱讼之要辞,皆曰契。”又加按语:“今人但于买卖曰文契。”^⑬由此我们可以把握清人对“契约”语义理解的三个要点,即“缠束”、“出予受入”和“买卖”。按照现代语义理解,这三个要点可以解释为:约束性、财产或权利的转移、经济行为。相关研究也证明,在清代社会,相互合意达成约束是契约一词的主要含义。^⑭因此,根据现存地方历史文献的形态以及传统上对“契约”本义的理解,我们将广义的契约定义为在一项经济行为中,参与者双方或多方对财产或权利的转移进行认可及约束的文本。在这一定义下,涵盖的内容应当分类为5个二级分类:A1各种散件契约(红契与白契)、A2抄契簿、A3合同、A4收据、A5分家书(分关书、阉书、继书、遗嘱)。

需要特别说明的是分家书系统。在之前的徽州文书分类系统中,将分家书系统归类为家族/宗族文书类,但在归档实际操作中,如果单独设置家族文书分类,会造成相当多文献面临重复分类的问题。从文献来源角度看,无论作为分家的分关书、阉书或者财产继承的继书、遗嘱,都是对一个家庭/家族财产分配方案的认定,符合本稿对契约的定义,并且也符合契约“缠束”与“出予受入”的本义,因此也归入契约类。

B. 账簿。根据现有研究,早期财政、经济记录登记文本的称谓是“籍”或“簿”,“帐”字被引申指代会计文本发生于南北朝时期,到唐代,“帐”已经广泛地指代“连续地记录和反映各项经济业务的籍簿。”^⑮根据元人孔齐的《静斋正直记》记载,元代时“帐目”一词已经明确地指称非官方的财产收支登记^⑯,这与现代汉语对帐目的理解已经非常接近了。自元代以来,“账”、“帐”、“簿”等词汇都具有了财产登记文书的含义。因此在本稿中,将账簿定义为对经济行为中产生的数据进行登记、统计、核算的文献。

传统账簿的产生来源可以分为家户和非家户单位两种。在账本定义与分类中作为核算单位与账本产生来源的家户,仅指主干家庭与核心家庭,而不包括宗族,使用家户而非家庭一词,强调的是家庭作为经济核算与生产的单位。

作为一种经济活动的记录,我们从两个维度考虑其功能、属性:经济活动内容与会计流程。账簿在经济活动中所扮演的角色有两大类,一类是商业活动的记录。“商业”一词在这里特别强调经由贸易、交换实现的盈利活动,其经营主体是一个商业组织,手工业生产、金融借贷则不涵盖在内。在这类记录中,财产的交流、流动是记录的重点。另一类是非商业活动,财产的增减是记录的重点。应对两种不同类型的经济活动,账簿记录内容各不相同。

^⑫ 徐国利:《徽州文书的理论研究与整理方法》,《中国社会科学院研究生院学报》,2005年第4期,第104—111页。

^⑬ [清]段玉裁:《说文解字注》,上海:上海古籍出版社,1981,第493页。

^⑭ 岸本美绪:《契约文书的研究》,滋贺秀三编:《中国法制史——基本资料の研究》,东京:东京大学出版会,1993,第759—805页。

^⑮ 郭道扬:《中国会计史稿上册》,北京:中国财政经济出版社,1982,第151、191、258、298页。

^⑯ [清]孔齐:《静斋正直记》,上海:上海古籍出版社,2002,卷三,第397页。

商业账簿主要记录经营主体的日常收支、贸易往来、资产清理。在传统账簿书写习惯中,以上三种功能常常写作“收支”、“客簿”、“盘货”等字样。非商业账簿主要记录的是财产收支与财产登记。在传统账簿书写习惯中,常使用“置产簿”、“租谷簿”等字样。置产簿,可理解为乡村社会背景下的土地财产等,因为在这种经济生活中,土地是主要的固定资产类型;租谷簿,可理解为农业生产的收入登记,地租是农业经济中最主要的收入来源。另外,在社会经济史研究传统中,也常将家户登记收支情况的账簿称为“家计账”。

依据以上理解,将账簿分为两大类:商业账簿与非商业账簿,商业账簿设置二级分类三个:商业收支账、往来账、盘货账。非商业账簿倘单一从经济功能的角度考虑,设置家计收支账、置产簿、收租簿三个分类即可。但是,本文认为有两种特殊的非商业账簿必须独立分类,才能将所有账簿文献包含在这一编目内。这两种文书是会簿和征信录。

会簿,是明清时代人们为某目的建立一个“会”或“社”时产生的文书。费孝通较早描述了“会”在乡村中存在的形态:摇会、徽会和广东票会,并将“会”定义为“集体储蓄和借贷的机构。”^{①⑦},在上海交通大学已收藏的浙南、徽州文书中存有数量丰富的会书,并已进行了深入研究。^{①⑧}

中岛乐章罗列了会社文书包含的范围,“以宗族和村落祭祀、庶民金融、商业、公益事业等为目的,纷纷结成了各种‘会’组织,记录‘会’组织的设立意趣书、运营规约、收支决算账簿类(会簿)、“会”参加权的买卖、典当契约等多种史料。”^{①⑨}也就是说,会社文书包含两大内容,第一是会社设立的约定及其权利转移的证明,第二是会社经营过程中的收支记录。

从现存会书的形态来看,订立会约的会书与公布会社账目的账本常常联为一个整体,或者毋宁说,会书与会社账本应当被当作一个整体看待。而这个作为整体存在的文献形态,其主要功能是账本。基于这样的考虑,会书暨会社账本列入账本类。

另外,在迄今上海交通大学收藏整理的徽州文书中,尚未发现可明确定义为征信录的文献,但征信录是徽州文书乃至明清地方历史文献中重要的一类文献,因此在分类法中设立二级分类。征信录取《中庸》“虽善无征,无征不信”之义,王振忠将其定义为“公开财务收支,以昭信实的档案。”^{②⑩}也就是说征信录文献的主要内容是某一团体的财务收支的记录,符合本稿对账本的定义。不过,征信录的内容除财务收支,往往亦包括宗族世系以及对某一与该团体相关事件的文献辑录,因此也带有家谱以及专门志的性质。本稿是讨论文献编目与分类法,因此从文献来源及产生的角度,将其归类为账本,其利用价值则不局限于此。

根据以上分析,将账本分为如下二级分类:B1 商业账簿(商业收支账、往来账、盘货账)、B2 家计账(家计收支账、置产簿、租谷簿)、B3 会簿、B4 征信录。

C. 赋役。本稿对赋役类文献的定义是政府对纳税人进行赋税、徭役的征收、登记过程中产生的文献。这一大类分为以下二级分类:C1 田赋缴纳记录(串单、易知由单、田赋征收单)、C2 土地登记记录(都图文书、土地清丈单、土地陈报单)、C3 徭役登记记录(黄册、百眼图、里甲文书)、C4 归户册、C5 实征册、C6 税额转移记录(过税票、推税票、收税票,统称为税票)、C7 盐税、C8 厘金缴纳记录、C9 商业税及其他税种缴纳记录。

①⑦ 费孝通:《江村经济》,上海:上海古籍出版社,2007,第 202—206 页。

①⑧ 章毅:《祀神与借贷:清代浙南定光会研究——以石仓〈定光古佛寿诞会簿〉为中心》,《史林》,2011 年第 6 期,第 66—77 页。

①⑨ 中岛乐章著、高飞、郭万平译:《明代乡村纠纷与秩序:以徽州文书为中心》,南京:江苏人民出版社,2006,第 39 页。

②⑩ 王振忠:《清民国时期徽州征信录及其史料价值》,“江南与海外交流”国际学术研讨会,上海:复旦大学历史系,2008。

对这一分类需要解释的是归户册、实征册和赋役合同。归户册是政府对每个纳税户拥有的土地及其相应额征田赋的登记记录,因此它事实上属于一种田赋登记记录。但是这一文献系统在地方历史文献的归户中具有重要意义,因此单列为一个二级分类。基于同样的原因,实征册也单列为一个二级分类,实征册是政府对每个纳税户当年应缴纳税额的登记记录。另外,在土地交易或其他场合中,交易双方会订立纳税义务归属的合同,这类合同是关于赋役事务的,但它不符合政府对纳税人的关系,而是私人关系的文献,因此列入契约大类中的合同二级分类。这也与之前徽州文书的分类法略有不同。

另外,黄册、里甲文书、都图文书、盐税等在上海交通大学藏地方历史文献中尚未发现,但这些都是明清赋税制度的重要部分,此单列为二级分类。都图文书是新近发现的文书类型,主要内容是对徽州各县的都、图、图所属村落、花户进行说明。这类文献的产生来源及其性质尚不能完全断定,既有可能是收税胥吏为收税目的编制,也有可能是用于指导商人和一般农户的类书。目前的研究倾向于它是为税收目的而编制的文献,因此暂且归类在赋役类中。

D. 诉讼。指诉讼案件过程中产生的文献。阿风根据前人研究,将明清徽州诉讼文书分为三类:官府文书,包括诉讼卷宗与官府发出的诉讼文书;“抄帖招文”与“执照”;民间文书,包括与诉讼相关的合同文书,以及民间收存、整理的诉讼文书资料。^②

从文书来源看,第一类文书是由官府制作、发出的文书;第二类是由民间制作,但得到官府认可的文书;第三类则是由民间自行编辑的文书。阿风的分类法符合本稿根据文书来源设计分类法,以及分类项不相容的原则,因此本稿同意阿风的分类法。但是,根据本稿对契约类文书的定义,所有的合同文书均应归入契约类,因此阿风分类法中归入民间文书的诉讼合同,应归入契约类。

阿风在以上三大类之外,根据诉讼文书的保存形态,又将诉讼文书分为9个小类。即诉状;立案帖文、票稿;传讯票;勘按结状与复命禀文;保状;提讯名单;供状与堂谕(看语);甘结;其他。

诉状是民人上告时书写的状纸,可归入第三类。清代诉状有严格的格式,并且呈递后会附有官员的批语,但仍应视作民间书写的文书。立案帖文、票稿是官府同意立案,并作出相应指示的文书,应归入第一类。传讯票是官府传讯案件相关人等前来官府的文书,传讯票有固定的格式,可以由于诉讼案件发出,也可能根据其他事由发出,因此我们认为所有传讯票应统一归入行政类文书。结状与复命禀文都是官府委派人员(里、保长,差役等)对案件情况的报告,应归入第一类。保状是诉讼双方在官府指定的店铺(即歇家)等待诉讼时,歇家呈递的保证书,这类文书有固定格式,且属于诉讼制度一部分,应归入第一类。提讯名单是官府提掉诉讼双方的文书,显然属于第一类。供状是诉讼双方在堂上的供词,看语则是官员作出的相应判决词。两种内容通常写在一张文书上,应归入第一类。甘结是诉讼双方同意判决,同意结案的文书,也是诉讼制度规定的最后一环,属于第一类。

以上是针对单件文书的分类处理,如果是汇编成册的诉讼文书,则应根据它的编纂者确定其分类,如果是由官府汇编存档的诉讼案卷,则应归入第一类,如果是民间自行编纂保存的诉讼过程记录,则应归入第三类。

这样,诉讼文书分为三个二级分类:D1 官府诉讼文书、D2 抄招、D3 民间诉讼文书。需要注意的是,徽州文书中存在一类教导讼师或普通人诉状书写格式以及诉讼技巧的文献,这类文献应当归入日用类书与工具书类。这类文献通常也会描述一个完整的案件,包括事由、人名、时间、地点等,因此容易被误认为是实际诉讼中产生的文献,这是文献整理中需要认真鉴别的。

E. 行政。这一大类的定义是政府处理赋役与诉讼以外的事务过程中产生的文献。这一类包含三个二级分类:E1 保甲、E2 政府布告与文告、E3 执照。从文献来源和社会功能考虑,C 赋

^② 阿风:《明清徽州诉讼文书的分类》,《徽学(第五卷)》,合肥:安徽大学出版社,2008,第252—280页。

役与 E 行政都是由官方发出的,实现政权控制人民的目的而产生的文献,应归为一个大类“官文书”。但处理赋役事物的文献存量很大,又是一个重要的研究主题,从提高二次检索效率的角度考虑,将赋役单列为一个类别。

F. 家谱。家谱是徽州文书中研究史最悠久的一类文献,这类文献的定义也较少歧义。除了常见的统宗谱、分支谱以及家规、族约之外,现存文献中还可以看到一些单页或数页的,只包含数代家族世系信息的世系图谱,这类文献既可能是编制宗族总谱的前期文献,也可能是根据宗族总谱抄录的用于某一家族分支祭祀的文献,在徽州当地,习称为“祖宗本”或“祖宗簿”,亦单列为一个二级分类。这一类包含三个二级分类:F1 家谱、F2 家规族约、F3 祖宗簿。

G. 信函。上海交通大学收藏文献中信函存量亦不少,这一类文献不需定义,分为两类:G1 信函、G2 信函留底。

H. 日用类书与工具书。20 世纪 90 年代之后日用类书受到徽州研究者的重视,其中最早受到重视的是介绍水路行程和经商知识的类书,研究者命名为“商书”。^② 日用类书这一名称,最早由酒井忠夫使用^③,在此之前,尚有仁井田陞称之为“日用百科全书”。^④ 但对于日用类书的总体性定义迄今少有,仅王振忠在新著中对日用类书作了归纳概括。^⑤ 王振忠认为“日用类书原先主要是供文人雅士的日常行事之用,明代后期发展成为一般民众皆可利用的民间日用类书。”除商业外,已发现的日用类书尚有对书信格式、诉状格式、称谓礼仪等各种专门事务进行介绍、指导的类型。本稿将日用类书定义为对某一商业与日常生活专门事务进行介绍、指导的文献。虽然如《商路规程》等类书是以商业指导为主,不过也常常包含其他内容,并且批量处理时难以迅速辨认,所以将类书统归为一类。另外,徽州文书中尚存有不少辞书、俗字典、书写格式指导书等文献,从功能上说,它们与日用类书接近,都是用于指导商业或日常生活中的具体事务,本稿将这些文献归类为工具书,将日用类书与工具书归为一大类。这一大类下分为如下二级分类:H1 类书、H2 辞书工具书。

I. 家礼。“家礼”一词本于《朱子家礼》,《四库全书提要》断《家礼》并非出自朱熹之手,而现代学者多倾向认定《家礼》为朱熹之作。^⑥ 无论如何,《家礼》是宋以降庶人礼仪代表性文本,也是对普通民众日常礼仪影响最大的规范性文本。^⑦ 《家礼·序》中描述了“家礼”的内涵:

凡礼有本、有文,自其施于家者言之,则名分之守、爱敬之实,其本也。冠婚丧祭,仪章度数者,其文也。其本者,有家日用之常,礼固不可以一日而不修;其文,尤皆所以纪纲人道之始终,虽其行之有时,施之有所,然非讲之素明,习之素熟,则其临事之际,亦无以合宜而应节,是亦不可以一日而不讲且习焉者也。^⑧

可知,家礼是对庶民成年礼、婚礼、丧葬、祭祀的礼仪的规范。在上海交通大学藏文书中,存在大量婚书、聘书、礼单、命理单、选吉期书、坟墓选风水单,以往研究者多根据自己的需要,将文献做不同分类。本稿认为,用“家礼”这一来自传统典籍的概念涵盖这些文献是最为适宜的。据

② 陈学文:《明清时期江南的商品流动与水运业的发展》,《浙江学刊》,1995 年第 1 期,第 31—37 页。

③ 酒井忠夫:《元明時代の日用类书とその教育史的意义》,《日本の教育史学》,1958 年第 1 期,第 67—94 页;《明代の日用类书と庶民教育》,林友春编:《近世中国教育史研究》,东京:国土社,第 62—74 页。

④ 仁井田升:《中国法制史研究奴隶农奴法·家族村落法》,东京:东京大学出版会,1962,第 741—829 页。

⑤ 王振忠:《明清以来徽州村落社会史研究》,上海:上海人民出版社,2011,第 51—52 页。

⑥ 束景南:《朱熹家礼真伪辨》,《朱子学刊》,1993 年第 5 期;陈来:《朱子家礼真伪考》,《北京大学学报(哲学社会科学版)》,1989 年第 3 期,第 115—122 页;杨志刚:《中国礼仪制度研究》,上海:华东师范大学出版社,2005,第 193—195 页。

⑦ 杨志刚:《中国礼仪制度研究》,第 206、238 页。

⑧ [宋]朱熹:《家礼》,台北:台湾商务印书馆股份有限公司,2008,《序》。

此,本稿将家礼文献定义为民众进行婚姻、丧葬、祭祀等礼仪过程中产生的文献。分类为:I1 婚书(聘书、礼单、选吉期单、两造命理单)、I2 丧葬文书(葬礼礼单、坟墓选风水单)、I3 个人命理单、I4 家族礼仪文书。

J. 宗教。地方历史文献中的宗教类文献主要就是科仪书,如有晚清基督教传教文献,也归入这一大类。宗教唱本、宝卷归为戏剧文献。罗经、风水地理书、命理推算指导书等按照典籍使用四部分类法归类。分为两类:J1 科仪书、J2 基督教文书。

K. 戏剧。戏剧文献包含三个二级分类:K1 戏曲抄本、K2 宗教唱本、K3 歌册。

L. 教育考试。用于教育与科举考试的文献归类为教育考试大类。分为三个二级分类:L1 启蒙书、L2 制义书、L3 科举试卷。制义书与科举试卷的区别是:制义书是科举考试写作的指导书,科举试卷是考试中实际写作的试卷或练习范文的辑录。

M. 医药。医药类书籍应按照四部分类法归类,但民间文献中有一些相关材料值得注意。一是单张的药方,这可能是为病人实际开具的药方,还有的是由民间宗教的仪式专家开具的,反映民间社会医疗与宗教之间的关系。另一类是医书抄本,通常没有完整的格式,以口诀等形式为民间提供医疗知识,可以认为和职业医家、儒医编纂的医书、医案属于不同的知识传播系统。这类文献目前可以分类两类:M1 药方、M2 医书抄本。

以上就是本稿对地方历史文献的全部分类。另外,应在每一大类下,都设置一个0号二级分类,供未来这一大类中新发现的尚无法归入已有二级分类的文献归类,直到可以确定其性质或建立新的二级分类。

据此,本文提出针对地方历史文献分类的四项原则:1. 分类主要依照文书性质;2. 确保批量处理的效率;3. 分类概念间不相容;4. 确保二次检索效率。依据以上原则,本文制订了包含10个类别的地方历史文献分类法,并利用编目抽检结果、研究文献计量等方法检测这一分类法的适用性。分类法可应用于数据库开发与文献编目整理,也可以作为学者进行量化研究与比较研究的分析工具。

Classification of Historical Archives in Private Holding: On the Collections of Shanghai Jiao Tong University

ZHAO Si-yuan¹, TANG Meng²

(1. History Department; 2. Library, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: The classification of the database and research of historical archives in private holding were examined. The application requirements of the database of historical documents in recent years and the research topics of historical archives in private holding were analyzed by using content analysis. The significance of classification in archive research and database development was discussed. Four principles on classification of historical archives were proposed, based on which a classification system and database development of historical archives in private holding in the Library of Shanghai Jiao Tong University was proposed.

Key words: Huizhou; historical archives in private holding; classification; digital humanities; database