

质性分析工具在情报学中的应用

Applying of Qualitative Data Analysis Tools to Information Science

黄晓斌 梁 辰

(中山大学资讯管理学院, 广州 510006)

[摘要] 在调查国内外质性研究文献和相关软件手册的基础上,介绍了质性研究范式及其研究步骤,分析了目前常见的质性分析软件特点,对比了它们在编码和检索、可视化以及结论检验等功能上的差异。通过对国内外图书情报学期刊中相关论文的发表情况进行分析,分别从情报学研究的三大领域——情报内容分布规律、信息用户行为研究和信息系统设计等论述了质性研究工具的适用性,并概括出质性分析软件与内容分析软件间的异同点,最后建议可利用质性分析工具来辅助开展混合法研究。

[关键词] 质性研究 情报学 编码 混合法

[中图分类号] G250.76 [文献标识码] A [文章编号] 1003-2797(2014)05-0004-13

DOI: 10.13366/j.dik.2014.05.004

[Abstract] Based on the relevant studies on qualitative data analysis(QDA) and manuals of QDA tools, the authors summarized the background of research paradigm and procedures of qualitative data analysis. Then five typical tools of QDA were introduced whose functions for coding and query, visualization and conclusion verification were also compared. Furthermore, those papers which used QDA tools and were published in library and information science journals were collected. After demonstrating the applicability of QDA tools in those three fields of research, namely, distribution of information, information user research, information system design, the differences and similarities among QDA tools and content analysis softwares were also concluded. Last but not least, it is suggested that researchers should employ QDA tools to conduct mixed methods researches.

[Key words] Qualitative Data Analysis(QDA) Information science Codes Mixed methods

1 引言

自孔德倡导实证主义以来,实证阶段逐渐被公认为人类智慧发展的最高阶段,而其中适用于通则式解释的定量研究更是成为如今科学研究的主流研究范式^[1]。就情报学研究而言,由于定量数据本身所内含的意义丰富性潜在损失和仅适用于演绎推理研究方式的局限性使得研究者难以开展深入的实证研究,已

在一定程度上制约了情报学相关理论的发展。此外,随着大数据时代的到来,网页内容、电子访谈等非量化、非结构化的文本数据也成为重要研究材料^[2],对这些非量化数据的分析深度、效率和一致性成为影响学者研究成果科学性和可靠性的关键因素。因此如何利用质性研究的方法和工具来归纳和延伸理论体系并进行严谨、规范的实证研究对情报学今后发展具有重要意义。

[作者简介] 黄晓斌,男,教授,博士,研究方向:情报学理论与方法、竞争情报研究、网络信息开发与利用,Email: isshxb@mail.sysu.edu.cn;梁辰,女,硕士生,研究方向:商情分析与竞争情报研究。

2 研究背景

2.1 研究范式

人文社会科学的研究范式可大致划分为质性和量化研究两种。质性研究指在自然情境下研究者通过与研究对象互动等多种方式收集资料,采用归纳推理方法来构建理论、诠释现象的一种研究方法^[3]。量化研究则主要是通过研究者通过调查、实验等客观计量手段收集资料,采用演绎推理方法,借助统计学理论及方法来检验理论正确性的研究方法。两种研究范式的区别主要体现在以下两个方面:

(1) 方法论上的不同。质性研究主要是通过参与式观察、开放式访谈、文本分析等方式来进行研究,重在特定情境下研究对象行为、心理等进行详细记录,主要研究方法有行动研究、案例研究、民族志、现象学、叙事研究等^[4]。而量化研究则主要是结构化问卷调查、封闭式访谈、数学统计等科学、客观的收集和分析数据的方法,重在一般规律的探索及其可推广性^[5]。

(2) 逻辑推理上的不同。质性研究是先收集资料后构建理论,从资料中探索不同节点间的时间顺序、因果、互斥等关系。而量化研究则是先基于理论提出具体假设,通过对实证数据进行统计检验或模拟仿真来验证假设。

2.2 质性研究步骤

质性研究一般要处理大量不同来源的文本资料并从中构架出符合实际情况的理论模型,其操作过程有着严谨规范的要求。本文借鉴 Miles 和 Huberman 的观点将质性研究分为三个步骤^[6]:

(1) 资料简化。首先在初期资料分析阶段要对研究对象的要素进行排序,确定初步的类别和主题阶段,并确定是否需要追加或缩减资料。其次,要对不同类别的质性资料,如访谈记录、视频录像等进行分类、编码、识别核心变量等。其中,编码是该阶段的核心过程。

以扎根理论研究为例,目前公认两种编码范例是社会学家 Glaser 和 Strauss 在 1967 年提出的^[7]和 Strauss 和 Corbin 在 1990、1998 年倡导的操作步骤^[8]。Glaser 和 Strauss(1967)的方法是逐行编码,再按相似度归类,提炼出概念关系。而后来 Strauss 和

Corbin(1990,1998)进一步提出了“开放式编码——轴心式编码——选择式编码”三步编码方法^[9]。

(2) 资料展示。为了架构理论,研究者应注意从概念、主题、模式、事件等发掘有意义的分类,并多采用表格、概念图或其他图像化形式来分析概念之间的关系、展示类之间的关系、构建理论框架。如可通过表格和条形图、柱形图或饼图等图形来呈现编码的频次等;而概念图则可借助链接来建立节点或元素之间的复杂联系,从而形成动态或静态的模型,便于研究者构建理论框架和调整研究思路。

(3) 结论总结/检验。通过分析可得到关于研究的相关初步结论,之后通过参考现有的现场记录或进一步数据收集、关系探索来对其正确性进行检验。首先,编码信度检验,为了比较编码的一致性可通过一致性百分比、Holsti 公式或 Cohen's kappa 信度系数等指标来评估编码质量^[10];其次,基本的编码统计方面可借助代码出现频次、共现频次、偶遇频次、代码精炼工具、编码索引等工具对评估编码条目的重要性和编码方法合理性;此外,由于编码之间的关联关系可能预示着因果关系,假设检验工具可通过有效关联条目来考察所有段落,进一步检验待定的编码条目关联关系是否存在。一般而言,若检索得到关联关系系数超过关系的平均值与标准差之和,则可认为关联关系存在^[11]。

3 典型的质性研究软件

从上个世纪 80 年代以来,质性分析过程的数字化和计算机化逐渐成为一个不可逆转的大趋势^[12]。据不完全统计,目前已有 40 多种质性分析软件^[13-14],主要是由美、英、德等国公司或机构所研发。质性分析软件是指专门以质性研究流派的方式来处理质性资料数据的计算机程序,其功能主要有编码、检索、创建索引、建立关系规则、构建模式或理论、绘制网络构图、初步统计分析等,能有效减轻质性研究者资料查阅、整理和手工编码的负担。

3.1 Nvivo

Nvivo 是澳大利亚 QSR 公司开发的定性分析软件,其名称意指自由自在地对非数值、非结构化数据进行标引、检索和理论化^[15]。Nvivo 最大的优势在于其强大的编码功能,其支持预建式和归纳式两种建立

编码方式。Nvivo 主要是通过节点来容纳编码、相关材料及其来源,其节点类型有以下四类:①自由节点,通常直接从文本中快速建立的编码都是包含在自由节点内,节点之间还没有建立概念上的关联;②树状节点,可通过鼠标拖拉建立节点间彼此上下的阶层关系,使自由节点成为树状节点;③案例节点,可为受访者、组织、产品等实体建立案例节点,借助一套属性集来描述实体特征,并可通过案例册(Case Book)工具查看案例及其属性概况;④关系:呈现项目中任两个对象(节点、文件或链接等)之间的关系,关系类型可自行定义。预建式建立编码可利用预先定义的树状节点间层级关系确定编码体系,而归纳式建立编码则是在多重检视和循环式编码过程中不断地将自由节点进行合并或加入到树状节点结构中以形成体系。编码过程可采用自由编码或快速编码等方式,对于结构化的材料来源还可考虑尝试自动编码功能。在查阅资料和编码的过程中,研究者可借助备注链接、参考链接、注解、超链接四种链接将资料关联整合起来,也可通过集合来对不同对象进行分类,简化资料库、方便资料查询。

Nvivo 的资料展示功能也较为多样。如常用的矩阵编码查询功能,利用表格展示已编码材料中节点、群组或属性值的比较结果。此外,Nvivo 在原有的可基于编码覆盖统计等数据建立图表的可视化功能上,还增加了词频统计的标签云、基于共享相似词、属性值或编码的材料来源或节点形成聚类树等附加展示功能。

而在结论总结方面,Nvivo 提供两种方式供研究者评估编码信度:①使用“编码比较”查询显示团队之间编码一致或不一致的百分比等结果;②采用“kappa 系数”的编码可靠性测量^[16]。为了进一步探索检验节点间的关系,还可生成类似于定量数据的频次表或交叉表等可视化图表,以检验不同问题之间的相互关系。Nvivo 提供了多种类型的链接来描绘因果、相互依存、层次结构、相互矛盾等关系。此外也可将编码分布情况导出到 SPSS 中作定量分析,提高研究客观性。

3.2 Atlas. ti

Atlas. ti 的雏形诞生于德国柏林科技大学在 1989~1992 年间开展的一项跨学科研究项目^[17]。1993

年开始商业化,并提供无使用时间限制但有功能限制的免费试用版。ATLAS. ti 是一款文件兼容性、软件操作友好性较高的软件,其支持各种文本、图像、音频和视频等文件,其中包括动态网页链接、内嵌式对象等非静态文件甚至是谷歌地球及其快照等数据,提供文件关联文档功能以实现多媒体文件和文本文件的同步^[18]。ATLAS. ti 中以阐释单元(Hermeneutic Unit, HU)来记录项目内的数据及操作,支持文件家族功能以辅助研究者按个案属性不同分组管理资料或快捷设定数据检索范围。ATLAS. ti 主要采用归纳式建立编码方式,尤其适用于扎根理论研究^[19]。其编码类目的管理主要是通过编码家族来进行,研究者可在编码家族管理器中设定不同编码之间的对称或非对称链接关系,如相互关联、相互矛盾、部分与整体、因果等。此外,ATLAS. ti 具有超码(super codes)功能,其相当于可定制化的自动编码操作。与普通编码链接的是研究者之前静态关联的参考来源不同,超码存储的是查询“意图”,被点击后其查询操作会自动运行并显示最新查询结果^[20]。同时,ATLAS. ti 还提供布尔、语义、邻近等检索,其中邻近检索能提供重叠、嵌套、顺序性、共现等多种情况的检索结果。

在数据展示方面,ATLAS. ti 不仅像 Nvivo 一样能生成编码共现矩阵,还能利用编码、原始资料、备注等各种对象构建网络图,其中包括链接的关系属性、显示格式等都能通过菜单灵活设置,并支持导出为图片格式或直接复制和粘贴到 Word 文档中。

在结论总结和检验方面,ATLAS. ti 的关系查询功能可计算嵌套(包含与被包含)、重叠(重叠与被重叠)、顺序(先于与后于)六种关系条目数,形成编码关系矩阵^[21],帮助研究者进一步探索因果关系是否存在。对于编码共现关系,ATLAS. ti 内置可自定义停用词词表功能,还提供共现频率矩阵等指标分析功能。上述分析结果均可导出到 SPSS、Excel 中或直接保存为网页。

3.3 QDA Miner

QDA Miner 是由 Provalis Research 公司开发的一个融合质性文本和定量数据分析方法的质性数据分析软件,它包含了定量内容分析、文本挖掘的

WordStat 组件和用于统计分析的 Simstat 组件。基于混合研究软件的定位,QDA Miner 非常适用于定量数据和质性文本庞杂的研究项目,它不仅能导入多种数据格式的定量或质性数据,还可导入其他质性分析软件如 Nvivo,Atlas.ti,HyperResearch 等软件的项目文件。QDA Miner 提供了转录功能使得研究者能对数值型、类别型变量进行格式转换,且所有数值型变量都可即时编辑并生成统计图表。在个案管理方面,QDA Miner 提供了个案描述器和分组工具,可快速生成个案属性的统计结果。在编码过程中,QDA

Miner 从编码自动聚类、编码相似度排序和简化编码拆分或合并步骤三个方面来辅助研究者构建或修改编码类目。首先,QDA Miner 提供了严格、中度、不精确三种可选的聚类提取结果。而对于已经建立类目的编码体系,QDA Miner 可提供具体的相似度结果(见图1)。最左边一栏是类目表,中间一栏是一个大表,从左到右分别是编码名称、同类编码间相似度得分及共有关键词,其他未加入该类的编码与选中的编码之间的相似度及共有关键词。而若编码者想调整现有类目,可通过菜单选择或直接在类目树中通过拖放操作进行调整^[22]。

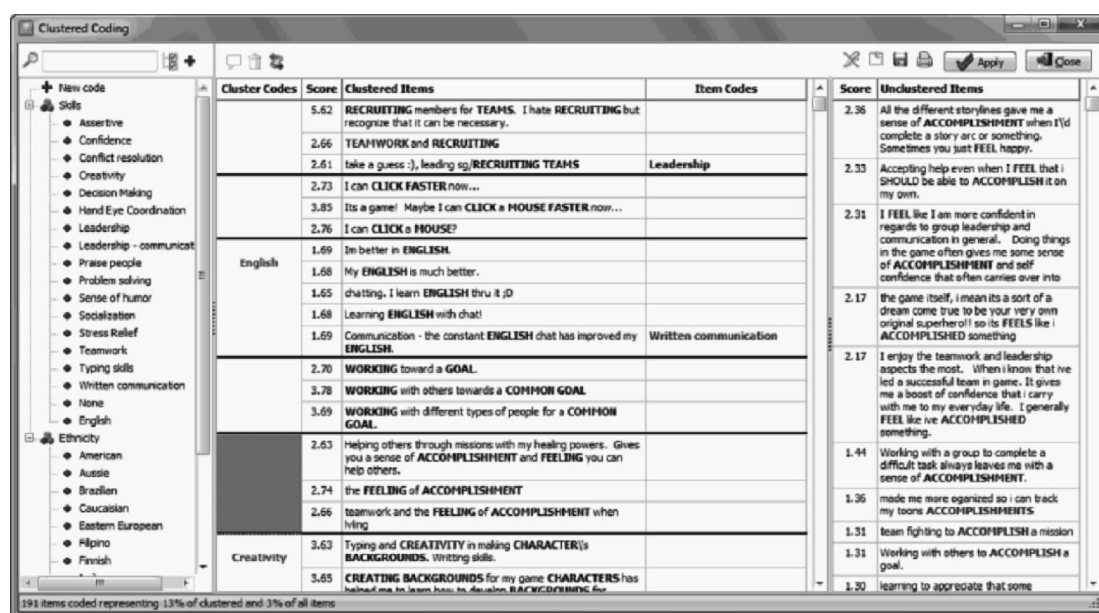


图1 编码类目相似度统计表

QDA Miner 提供了丰富的可视化工具:编码词频矩阵(见图2)、编码变量统计表、多维尺度图、热点地图(见图3)、编码序列图、对应分析图、聚类谱系图及相似度矩阵、气泡图、邻近图、地理标签和时间标签工具等^[23]。

在编码信度检验方面,QDA Miner 内有自由阈值调整、 π 值^[24]、Krippendorff's α 值三种方式^[25]。此外,关于文本分析功能,QDA Miner 中的 WordStat 文本挖掘组件具备朴素贝叶斯、K 最近邻两种基于机器学习的自动文本分类功能,也提供拼写检查。但是目前

QDA Miner 中文兼容性还不太理想,导入中文文档时可能会出现乱码,在使用 WordStat 文本挖掘组件内容分析时也很难准确地实现中文切分词,而且现有的词形归并等文本预处理操作都是针对英文的。在对编码进行初步统计后,研究者可将统计结果保存为数值型变量,再使用统计软件进行分析,如 QDA Miner 配套的 Simstat 组件。该组件集成了聚类分析、多维尺度分析、对应分析、序列分析等统计分析功能,可帮助研究者识别模式和趋势,探索结构性数据,描述、对比和检验假设^[26],但它和 WordStat 组件都需要额外付费。

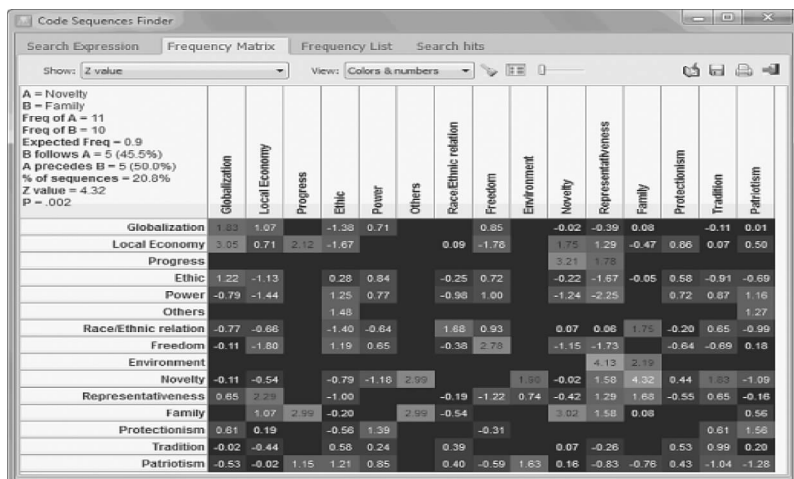


图2 编码词频统计矩阵

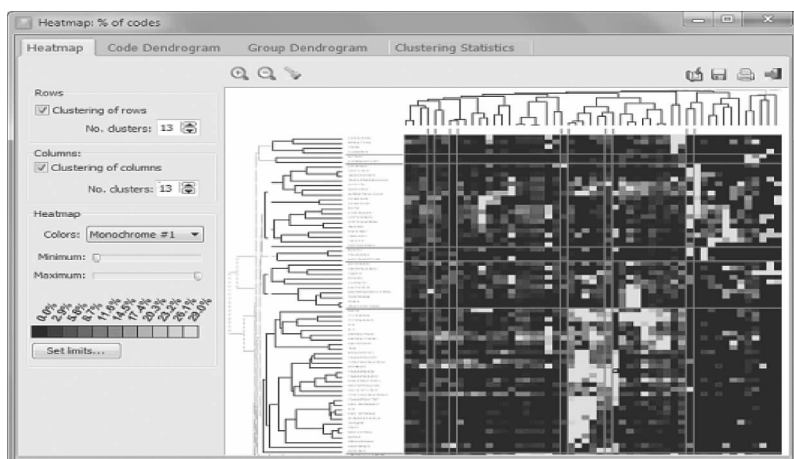


图3 热点地图

3.4 Qualrus

Qualrus 是一个由 Idea Works 公司开发并由美国著名芝加哥学派社会学家 Howard Becker 负责指导的计算机辅助数据分析软件^[27]。虽然 Qualrus 支持的数据格式类型不如 Nvivo 和 Atlas. ti 等主流质性研究软件多,它基本支持常用的几种视频、网页、图像、音频、文本格式资料的导入。Qualrus 自称是新一代质性研究软件,其编码功能特点有:支持人工智能辅助、提供提示性编码等^[28]。在项目初期,Qualrus 会根据用户定义的同义词词表和之前项目的编码习惯提供编码提示,而且软件自身会智能地根据编码者

之前对编码提示的采纳记录不断学习和调整编码提示策略。研究人员可选择采纳或者忽视软件的编码提示,也可更改某些编码提示规则,且软件也会提示相关编码、识别出编码者可能已经忽视的编码间关系等。此外,Qualrus 提供了多种编码辅助功能:①偶遇编码,显示编码间在不同片段中的叠合情况;②编码分类工具,提示编码的相关编码和高阶编码^[29];③提炼工具,查看哪些编码可合并;④编写脚本语言自动编码,针对某特定分析任务按照脚本语言规则撰写程序;⑤编码检索工具,包括布尔操作符和检索工具 QTools。灵活可定制化的编码功能也使得该软件可

适用于不同理论方法——归纳编码和演绎编码。

Qualrus 的不足在于其较差的用户友好性: ①编码操作复杂, 编码时需要多次点击, 且不支持拖放编码操作, 没有编码快捷键, 也不支持自由编码; ②没有默认的自动编码功能, 用户须学习脚本语言编写自动编码规则; ③不能自行调整显示格式, 包括编码的颜色都是软件预先设定的; ④窗口交互性不佳, 如用户只能在编码编辑器中才能实现编码以及快速检索和打印编码所指向的文本片段。

在资料展示方面, Qualrus 内含链接工具, 默认的连接关系有因果、互斥、从属关系三种, 并支持自定义链接关系类型。软件提供了两种编码类目呈现视图——星状图和网络图。

Qualrus 也提供较丰富的结论检验工具, 在研究者采用 IF 和 THEN 语句指定某一假设后, 如统计某两个或以上编码在资料中的共现和非共现情况并生成词频统计矩阵。而且软件也可实现一般的概括统计功能, 如不同资料来源的编码词频统计、特定文本中编码的共现情况等。此外, 软件还具有辅助研究者理论抽样功能^[30]。

3.5 RQDA

RQDA(R-Qualitative Data Analysis) 是由香港城市大学公共与社会行政系黄荣贵博士所研发, 是一个基于 R 语言环境的、辅助研究者进行文本处理分析工作的质性研究工具, 其中文兼容性极好^[31]。虽然与其他几款商业质性分析软件相比, RQDA 小巧很多, 但其也包含了标准的 CAQDA (Computer-Aided Qualitative Data Analysis) 的功能如编码、类目构建、项目管理。首先在文本处理方面, 除了基本的文本导入和分析功能, RQDA 可借助文本挖掘包 (Text Mining, TM) ^[32] 和潜在语义分析包 (Latent Semantic Analysis, LSA) ^[33] 来对构建个性化的词库、专业的中文切分词、词性分析、元数据、生产词条——文档矩阵、语义挖掘等, 提高了文本处理的效率和准确度^[34]。RQDA 采用的是列表式编码组织方式, 但研究者可通过编码类别来实现概念体系的构建^[35]。此外, 利用交叉编码函数生成矩阵可反映不同编码之间的关系, 如包含、重合、邻近、精确等, 也可通过已编码文件函数

来生成矩阵来对比不同文件中编码的分布情况^[36]。

R 语言中丰富的绘图函数也使得 RQDA 成为一款可视化效果较佳、灵活性较强的分析软件。典型的可视化功能有: ①编码类目图, 展示类目及其涵盖的编码之间的关系 (见图 4); ②标签云图, 反映编码频次及其重要性 (见图 5)。此外, 可借助其他 R 包如质性对比分析工具 (Qualitative Comparative Analysis, QCA) 来绘制编码的柱形图、条形图、散点图及其分布矩阵等, 也可遵循 R 语言的代码规则自行撰写绘图函数。

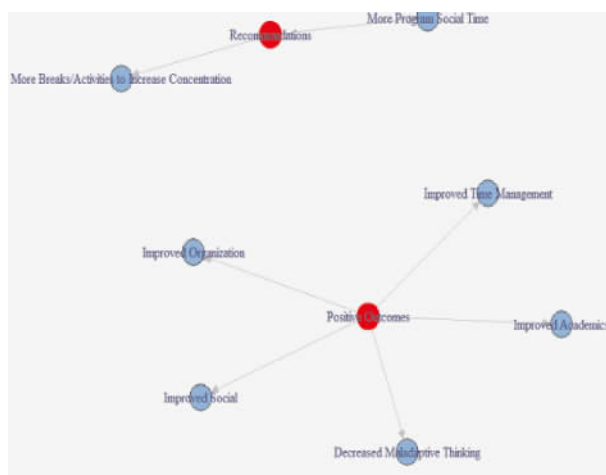


图 4 编码类目图



图 5 标签云图

在结论检验方面,QCA包也可极大地辅助研究者对RQDA分析结果作进一步的对比检验。QCA具有强大的对比功能:①通过个案来对比不同个体特征的差异,如人口统计特征的分布;②通过个案来对比不同数据收集方式中语境的差异,如访谈类型对语境的影响;③通过编码来对比不同编码文本中属性的差异,如正向或负向结果的统计。而在进一步检验结论

时,研究者可采用QCA来分析:①检查数据间的一致性,如查找矛盾、深入分析独立个案;②检验现有理论或假设(证实或证伪);③检验新想法或假设,利用R中的统计分析函数进行数据探索^[37]。

以上所述的几种软件工具各有其特色,为了更好地了解其特点,下面将常见的质性分析软件的各种主要参数进行对比(见表1)。

表1 典型质性研究软件的参数对比

软件名称		Atlas/ti	Nvivo	QDA Miner	Qualrus	RQDA
是否免费		×	×	×	×	√
支持音频、视频		√	√	×	√	×
编码功能	自动编码	√	√	√	√	×
	即点即编	√	√	√	×	×
	快捷键	√	√	√	×	×
	编码合并	√	√	√	√	√
	编码拆分	√	×	√	√	√
	组织方式	家族式(可支持层级式)	层级式	层级式	层级式或非层级式均可	列表式
检索方式		布尔、邻近、语义	布尔、邻近	布尔、条件、示例	布尔、条件	布尔、条件
可视化	统计矩阵	√	√	√	√	√
	类目树	√	√	√	√	√
	网络图/模型	√	√	×	√	×
检验功能	停用词表	√	×	√	×	√
	统计工具	√	√ ¹	√ ²	×	√ ³
	编码间关系类型的统计	嵌套、重叠、顺序	共现	共现、相似	共现	包含、重合、相似、共现
	假设检验	√ ⁴	×	√ ⁵	×	√ ³
运行平台		Windows, Mac OS x, Sun UI-tra	Windows	Windows	Windows	Windows, Linu x / FreeBSD, Mac OS x
中文兼容性		较好	较好	较差	较好	很好

注1: Nvivo 有基本的统计功能但不能进行单个词的词频统计;
注2: QDA Miner 的统计功能广泛,既可实现共生矩阵可视化,也提供一些聚类方法;
注3: RQDA 的统计功能需要用户用 R 语言创建或调用函数;
注4: Atlas. ti 支持聚类和相关假设检验;
注5: QDA Miner 支持序列、聚类、相关、层序、维度、因果等关系检验;

4 质性分析工具在情报学研究中的相关问题

4.1 目前的应用情况调查

为了解质性研究软件在情报学中的应用情况,本

文通过对国内外情报学期刊进行文献检索,分别对运用质性研究软件进行研究的论文进行年份、期刊分布情况统计。国内情报学期刊相关论文的检索以文中

介绍的五种质性研究软件“Nvivo”、“Atlas.ti”、“QDA Miner”、“Qualrus”、“RQDA”为关键词,在 CNKI 中进行全文检索,发表时间不限(检索日期:2014 年 4 月 1 日),搜索学科领域限定为“情报学、情报工作”。而国外情报学期刊则分别通过 Emerald 中的“Library and Information Studies”、EBSCO(Academic Search Complete)中的“Library, Information Sci-

ence & Technology Abstracts (LISTA) with Full-Text”、PRO Quest 中的“Library and Information Science Abstracts (LISA)”、Taylor&Francis 中的“Information Science”四个数据库对学术期刊论文进行检索,关键词和发表时间设定同上,并对检索结果进行汇总去重最终得到 68 篇相关论文,如表 2 所示。

表 2 运用质性研究软件进行研究的论文在国内外期刊分布统计表

图书情报类期刊刊名	Atlas. ti	Nvivo	QDA Miner	Qualrus	RQDA	总计
Library & Information Science Research	9	18	0	0	0	27
Information Technology for Development	1	5	0	0	0	6
Australian Academic & Research Libraries	0	3	0	0	0	3
Behaviour & Information Technology	0	3	0	0	0	3
Journal of Library Administration	2	1	0	0	0	3
New Review of Academic Librarianship	1	2	0	0	0	3
The Information Society	0	3	0	0	0	3
Journal of Documentation	0	0	0	1	1	2
Journal of Archival Organization	0	2	0	0	0	2
Journal of Library & Information Services in Distance Learning	1	1	0	0	0	2
Science & Technology Libraries	1	1	0	0	0	2
The International Information & Library Review	0	2	0	0	0	2
Journal of Business & Finance Librarianship	0	1	0	0	0	1
Journal of the Society of Archivists	0	1	0	0	0	1
Journal of Web Librarianship	1	0	0	0	0	1
图书情报工作	0	2	0	0	0	2
大学图书馆学报	0	1	0	0	0	1
中国图书馆学报	0	1	0	0	0	1
情报科学	0	1	0	0	0	1
图书馆杂志	0	1	0	0	0	1
现代情报	0	1	0	0	0	1
总计	16	50	0	1	1	68

由论文发表年份分布图(见图 6)可发现,国外期刊从 2001 年开始收录采用质性研究软件进行研究的论文,且每年发表论文数量不断上升,2011 年的发表数量更是达到了 14 篇。而国内的情报学研究则是到了 2010 年才开始出现采用质性研究软件进行研究的期刊论文,随后也处于上升阶段,且发表数量的增长趋势与国外期刊较为一致,但数量上仍有不小差距。从论文期刊分布情况来看,Library & Information Science Research、Journal of Documentation 等影响因子较大的国外期刊都发表了不少以质性研究软件为

工具的论文,其中 Library & Information Science Research 发表的论文数量最多,高达 27 篇。而国内核心期刊中虽然大部分都仅发表了 1~2 篇相关论文,但是其中不乏《中国图书馆学报》、《大学图书馆学报》、《图书情报工作》、《情报科学》等重要核心期刊,说明质性研究工具的采用已受到国内情报学学术界的广泛认可。而从软件使用情况来看,国外期刊的相关论文以 Nvivo、Atlas. ti 为主,同时也有少数论文使用其他的质性研究软件,如在 Journal of Documentation 上发表的两篇质性研究论文分别采用的是 Qual-

us 和 RQDA 软件。而国内期刊的相关论文则基本上都采用 Nvivo 软件进行研究。综上,目前质性研究软件已被国内外情报学界认可,且相关研究论文发表数

量近年来有了一定的提升,但国内情报学界质性研究软件的实证研究仍处于初步发展阶段,应用实例数量较少且采用的软件也较为单一。

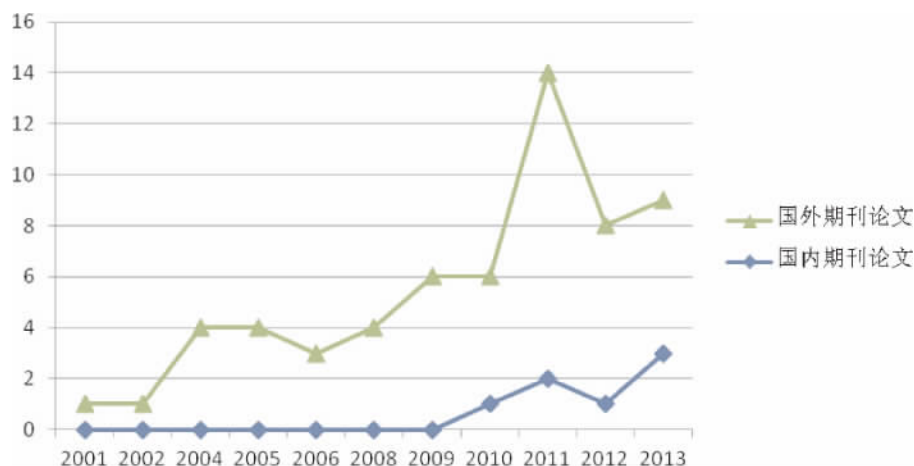


图6 2001~2014年运用质性研究软件进行研究的论文发表年份分布情况

4.2 质性分析软件在情报学研究的适用性

由上述文献调查结果可知,质性研究软件已逐渐受到国内外情报学研究者的重视。Saracevic 曾概括情报学研究主要是回答三个层面的问题:①物质层面的问题,记录的情报中有什么特点和规律;②社会层面的问题,人们如何关联、搜寻、使用情报;③设计层面的问题,如何使获取、记录情报最快速、有效^[38]。在这三个不同细分领域的研究中,质性研究软件都能发挥其独特作用,具有较强适用性。

在物质层面问题的研究中,目前情报学主要是通过计量学的方法来分析文献外部特征,归纳情报的分布特征、变化规律、数量关系等进而推断出学科发展的演变过程及规律、知识结构^[39]。而对于情报内容特征的研究则主要是借助文本挖掘软件来进行内容分析。随着大数据时代的到来,情报的“3V”问题越来越突出,尤其是非结构化的文本数据及多媒体数据已逐渐成为情报的重要来源。而质性研究软件较强的文件兼容性,对文本数据高效简化、编码、链接等处理功能,对概念关系的可视化和检验能力都能极大地辅助情报学研究者对情报内容的分布规律及特点进行

深入分析。如 *Library & Information Science Research* 中就有一篇论文是运用 Atlas.ti 软件对用户收藏描述数据进行内容分析,并经过不断对开放式编码进行迭代聚类最终构建出收藏元数据架构(collection metadata schemas)^[40]。

其次,质性研究软件也能帮助情报学研究者开展针对社会层面信息用户行为模式、心理状态变化规律等探索性研究。信息用户一直是情报学研究的核心关注点之一,尤其是用户情报检索、组织、使用行为的特点及其影响因素的作用机制等都是情报学中的重要议题。为了更全面收集信息用户行为的情境因素、更深入地了解用户认知模型、行为模式、心理学模型,研究者需采用个人深度访谈、焦点组访谈、开放式问卷等质性研究方法收集资料^[41],并利用质性研究软件通过扎根理论等归纳推理方法构建出一套关于用户行为模式的理论框架^[42]。目前质性研究软件在国内外情报学研究的应用也主要是集中是在这一领域,如关于信息搜寻行为(Information-seeking Behaviors)^[43-44]、协同探究基于项目的学习(Collaborative Inquiry Project-based Learning)^[45]、决策模型^[46]等

信息用户行为及需求^[47]的研究。

而针对设计层面的问题,目前情报学研究者主要通过信息检索系统的用户实验来探索情报的组织方式、情报系统的设计及其性能表现、情报服务质量评估等问题。无论是对情报系统的信息构建过程、对系统用户使用行为的田野调查还是对情报服务情况的深度访谈,都需要借助质性研究软件来对繁杂的质性资料进行整理,此外还可将初步分析结果如编码词频的统计结果、编码属性的分布情况导出再进行下一步的相关分析、方差检验、对比分析等。尤其在质量评估方面,运用这些质性研究软件可以帮助研究者对用户访谈记录等定性材料进行研究,探索性地了解影响用户使用系统的因素^[48]、全面地评估信息服务的效果^[49]、并支持服务或系统设计的改善^[50]。

4.3 质性研究软件与内容分析软件的对比

由于质性研究软件定位于辅助质性研究者对大量定性资料的整理和分析,因此其十分适用于对文本、音频、视频等资料的内容分析,实现编码、理论构

建、统计及可视化、结论的建议与解释等分析过程。但需要注意的是这里的“内容分析”与“内容分析法”是有差异的。质性研究软件的“内容分析”功能主要是体现其能基于资料内容的深入加工为扎根理论或框架分析等理论探索过程提供理论依据。而“内容分析法”则是作为一种利用推理和比较方法对研究对象的内容特征进行分析的半定量研究方法^[51],侧重的是对信息内容特征的数量统计及语义分析,需要符合客观、系统、定量三大要求^[52]。为了便于今后研究者分析工具的选择,本文对质性研究软件和内容分析软件的异同点进行了总结(见表3)。概括而言,质性研究软件擅长理论构建,而内容分析软件擅长语义统计分析。若研究本身属于质性研究范畴,研究目的在于对尚未明晰的理论进行探索和初步论证,可考虑使用质性研究软件。而当研究者从事的是对某一领域的现状描述、对比分析、趋势推演,需要工具辅助其对文本等材料进行语法单元的统计,则应选择内容分析软件。

表3 质性研究软件与内容分析软件的异同^[53-58]

软件类型		质性研究软件	内容分析软件
共同点	研究对象范围	期刊论文;报纸上的新闻、报道、评论;各类广告、多媒体节目等资源类型,涵盖文本、图像、音频、视频等	
	信度效度检验	信度检验,如一致性百分比、Holsti公式或Cohen's kappa信度系数等;效度仍未有统一实用的评估标准	
不同点	研究方法	质性研究方法,实证的定性研究	半定量研究方法,含定性分析的定量研究
	推理逻辑	①归纳推理,自下而上,研究前无需提出理论假设; ②演绎推理,自上而下,研究前需按照一定理论框架	演绎推理,自上而下,研究前需按照一定理论框架,确定分析的维度及体系
	抽样方式	理论抽样,重视理论饱和度	系统抽样、随机抽样、方便抽样,重视样本容量及其代表性
	分析方法	扎根理论、行动研究、案例研究、民族志等	内容分析法
	分析单元	主要以编码为主,也可对个案、原始文本、文件进行分析,重视个案属性分布情况	主要以编码或其他语法单元(词语、主题、句子、段落等)为主,也可对时空分布等情况进行分析,不划分个案
	编码规则	编码类目之间可以形成分组或层级结构,一个分析单元可以被重复赋予多个编码,在分析过程中不断地调整编码类目体系	编码类目间需类目之间应该是互斥和独立的,一个分析单元只能放在一个类目中,编码类目体系在分析前已确定
	可视化	基本统计结果的图表展示(条形、柱形、散点等);共现矩阵;标签云图;理论框架的网络图	基本统计结果的图表展示(条形、柱形、散点等);共现矩阵;标签云图;社会网络图
	结论检验	根据嵌套、重叠、顺序等关系进行因果关系、矛盾关系判断,通常由质性分析软件完成,也可导出到统计分析软件进一步分析	百分比、平均值、相关分析、回归分析、方差分析、卡方分析、对应分析、主成分因子分析等各种统计分析均可实现,还有社会网络分析、情感分析、流量分析等网络内容分析功能,通常由相应的统计软件完成

4.4 质性研究软件与量化研究软件的结合

质性研究软件与量化研究软件的结合是在混合法(Mixed Methods)逐渐被学术界认可的背景下开始受到关注的。为了消除单一数据来源的偏见和特定研究方法的局限,早在1978年质性研究方法论大师Denzin就提出了“三角互证(Triangulation)”的方法,其中包括:①方法内互证,即同时使用同一研究范式内的不同方法;②方法间互证,即同时使用质性研究方法和量化研究方法^[59]。虽然两种范式的结合早有尝试,但混合法这一说法则是在2003年Tashakkori和Teddle出版《社会和行为研究的混合法手册》一书后才正式被提出,书中将其定义为“混合法研究是一种有着特定哲学假设前提和探索方法的研究设计;是一种质化与量化有机混合的方法论;也是一种在研究中收集、分析、混合质性资料和量化数据的方法。其基本思想在于结合两种范式以更好理解研究问题。”^[60]质性资料可用于解释、阐明、描述、证实或修正量化研究结论,量化数据可用于评估质性研究结论的可推广性^[61]。因此研究者应多尝试采用混合法研究策略,综合运用质性研究软件与量化研究软件对同一研究问题进行分析。

在混合法研究中,质性研究软件的作用主要体现在两个方面:理论探索和数据交互。理论探索作用体现在研究者利用质性研究软件遵循编码规则、合并主题、形成概念和类型、构建概念网络图、检验初步结论及理论升华等步骤,实现质性研究和量化研究间互补、完备、发展、拓展、证实、补偿和多样化的目的^[62]。尤其是在Nvivo和Qualrus这些适用于归纳和演绎两种推理方式的质性分析软件,可应用于同时性或相继性^[63]的混合方法研究设计中。而像Atlas.ti偏适用于扎根理论方法的软件则仅适用于先质性后量化的研究路径。而数据交互则指研究者综合利用量化统计软件和质性研究软件实现质性资料与量化数据之间的交互,即将质性研究的编码词频统计、共现矩阵、个案属性概况等结果导入到统计软件中进行进一步的假设检验和相关、回归、对应、聚类等统计分析。目

前大部分质性研究软件都支持数据导出功能,而且在数据统计方面,QDA Miner和RQDA都属于其中的佼佼者。但相对单一研究范式的项目而言,混合法研究更为复杂,在研究设计及分析结果的一致性和效度检验方面的要求也会更加苛刻,其研究结果必须同时满足质性研究和量化研究的信度和效度标准,其中借助质性研究软件规范编码过程是保证质性研究信度的关键步骤之一。此外,要体现混合法的价值,除了需要从两种方法得到丰富结论外,更重要的是结合两者的研究结论提出一个理论性陈述、叙述或故事——元推论(Meta-inferences)^[63]。

5 结语

质性研究是一种在自然情境中收集各种质性资料,运用现象学、阐释学的理论和方法对社会现象进行整体性探究、通过归纳演绎构建理论的研究方法^[64]。与定性研究主要思辨的研究过程不同,质性研究具有规范的资料编码和理论推导过程,目前已被广泛应用到社会学、管理学、心理学等领域。本文综述了科学的研究范式及质性研究步骤,并具体介绍了五种典型的质性研究软件。由于质性研究软件一般都具备编码、可视化、初步统计分析等功能,可在研究过程中发挥重要作用:①兼容多种格式资料,便于资料的抽取和整合,提高简化资料的速度。②提供可视化界面及交互操作功能,辅助概念、类别及其中关系的提炼,协助理论架构的调整;③生成编码、文件、个案属性的初步统计结果,以供研究者对结论的进一步探索和检验。而且质性研究软件在情报学中具有较强的适用性,能应用于情报内容分布规律、信息用户行为模式、认知模式和心理模式和信息系统设计及服务评估等多个研究领域。但是需要注意的是,质性研究方法和内容分析法是相似但截然不同的两种研究方法,学者在选择研究工具过程中需要明辨质性研究软件和内容分析软件间的异同点。随着三角互证法的兴起,质性研究软件凭借着其特有的理论探索和数据交互功能将成为研究者开展混合法研究的重要工具。

参考文献

- 1 艾尔·巴比. 社会学研究方法[M]. 邱泽奇译. 北京: 华夏出版社, 2000: 24-26, 94-96
- 2 Bampton R, Cowton CJ. The E-Interview[C]. In Forum: Qualitative Social Research, 2002
- 3 陈向明. 教育行动研究中如何使用质的方法[J]. 基础教育课程, 2005, 4: 25-7
- 4 陈向明. 质性研究方法与社会科学研究[M]. 北京: 教育科学出版社, 2000: 13-22, 53-55
- 5 希尔·弗曼. 如何做质性研究[M]. 重庆: 重庆大学出版社, 2009: 4-7
- 6 Miles MB, Huberman AM. Qualitative data analysis: An expanded sourcebook[M]. Sage, 1994: 1, 105, 273
- 7 Glaser BG, Strauss A L. The discovery of grounded theory: Strategies for qualitative research[J]. Chicago: Aldine, 1967, 17(4): 364
- 8 Anselm S, Juliet C. Basics of qualitative research: Grounded theory procedures and techniques[M]. London: Sage Publications, 1990: 85-103
- 9 Strauss A, Corbin J. Basics of qualitative research: Techniques and procedures for producing grounded theory[M]. London: Sage, 1998: 58-61
- 10 Cohen J. A Coefficient of Agreement for Nominal Scales[J]. Educational and Psychological Measurement, 1960, 20(1): 37-46
- 11 Malina MA, Selto FH. Communicating and controlling strategy: an empirical study of the effectiveness of the balanced scorecard[J]. Journal of Management Accounting Research, 2001, 13(1): 47-90
- 12 Dohan D, Sanchez-Jankowski M. Using computers to analyze ethnographic field data: Theoretical and practical considerations[J]. Annual Review of Sociology, 1998, 24: 477-98
- 13 胡萍. 质性分析工具的比较与应用研究[D]. 湖南: 湖南师范大学, 2012
- 14 夏传玲. 计算机辅助的定性分析方法[J]. 社会学研究, 2007, 5: 148-63
- 15 王宏仁. 作为一种思考的鹰架, 质性分析辅助工具介绍: Nvivo 7.0[J/OL]. [2007-02-14]. http://www.pomine.com/software/NVivo/documents/Nvivo7_doz_2007.pdf
- 16 International CQ. NVivo 10 Getting Started Guide – Chinese QSR International[EB/OL]. [2012-06-01]. <http://download.qsrinternational.com/Resource/NVivo10/Whats-new-in-NVivo-10-Chinese.pdf>
- 17 张浓芳. 国外常用 CADQAS 软件介绍[J]. 软件导刊, 2010, 9(8): 11-3
- 18 Friese S. Atlas. ti6 速成教程[EB/OL]. [2014-01-10]. http://www.atlasti.com/uploads/media/Quick_Toura6_zh_01.pdf
- 19 Pope C, Ziebland S, Mays N. Qualitative research in health care: Analysing qualitative data[J]. British Medical Journal, 2000, 320: 114
- 20 Hwang S. Utilizing Qualitative Data Analysis Software A Review of Atlas. ti[J]. Social Science Computer Review, 2008, 26(4): 519-27
- 21 于春玲, 等. 中国情境下成功品牌延伸影响因素的案例研究[J]. 管理世界, 2012(6): 147-162
- 22 Corp PR. QDA Miner User's Manual[EB/OL]. [2014-01-12]. <http://provalisresearch.com/Documents/QDAMiner40.pdf>
- 23 Friese S. Which package: comparisons and strengths[EB/OL]. [2014-01-11]. http://www.quarc.de/software_overview_table.pdf
- 24 Scott WA. Reliability of content analysis: The case of nominal scale coding[J]. Public Opinion Quarterly, 1955, 19: 321-325
- 25 Haves AF, Krippendorff K. Answering the Call for a Standard Reliability Measure for Coding Data[J]. Communication Methods and Measures, 2007, 1(1): 77-89
- 26 Corp Q. QDA Miner More Features [R/OL]. [2014-01-14]. <http://www.quantrus.com/qdafeatures2.html>
- 27 Works TI. Qualrus: The intelligent qualitative analysis program [R/OL]. [2014-01-12]. <http://www.qualrus.com/>
- 28 Silver C, Fielding N. Using computer packages in qualitative research[M]. New York: Sage, 2008: 339-341
- 29 Bowers Jr J, Swingle A. A Presentation on Qualrus[J/OL]. [2014-01-12]. <http://www.iup.edu/assets/0/347/351/5139/5141/16A6CA70-E279-4D41-BCFE-868592A89D2F.pdf>
- 30 Works TI. Qualrus Manual[J/OL]. [2014-01-13]. <http://www.ideaworks.com/download/qualrus/QualrusManual.pdf>
- 31 Huang R. RQDA: R-based qualitative data analysis. R package version 01-9; 2010
- 32 Feinerer I. Introduction to the tm Package Text Mining in R [EB/OL]. [2013-12-01]. <http://www.dainf.ct.utfpr.edu.br/~kaestner/Mineracao/RDataMining/tm.pdf>
- 33 Wild F. An R Package for Latent Semantic Analysis [EB/OL]. [2014-01-15]. <http://dspace.learningnetworks.org/bitstream/1820/966/14/2007-03-29Isatel-2007package.pdf>
- 34 Ye B. 用定性数据分析包 RQDA tm 进行文本挖掘[EB/OL]. [2014-01-14]. https://r-forge.r-project.org/scm/viewvc.php/*

- checkout* /www/Using_RQDA2tm.pdf? revision = 330&root = rqda
- 35 Huang R. 定性分析软件 RQDA 简明使用手册 (PDF in Chinese) [EB/OL]. [2014-01-15]. <http://homepage.fudan.edu.cn/rghuang/files/2012/04/RQDA-intro.pdf>
 - 36 Varner L, Carter A, Furter RD, H. R Qualitative Data Analysis (RQDA) Package: A Free Qualitative Analysis Tool[C]. In Skill-Building Workshop presentation, Minnesota, USA: American Evaluation Association, 2012
 - 37 Huang R. What is RQDA and what are its features? [R/OL]. [2014-01-15]. <http://rqda.r-forge.r-project.org/>
 - 38 Tefko S. Information Science. Encyclopedia of Library and Information Sciences, Third Edition[M]. Taylor & Francis; 2009: 2570-85
 - 39 王曰芬, 路菲, 吴小雷. 文献计量和内容分析的比较与综合研究[J]. 图书情报工作, 2005, 49(9): 70-73
 - 40 Stvilia B, Jørgensen C. User-generated collection-level metadata in an online photo-sharing system[J]. Library & Information Science Research, 2009, 31(1): 54-65
 - 41 高淑琴, 罗金增. 论质性研究在图书馆学情报学研究中的应用[J]. 图书馆学研究, 2010(12): 2-4
 - 42 韩正彪, 周鹏. 扎根理论质性研究方法在情报学研究中的应用[J]. 情报理论与实践, 2011, 34(5): 19-23
 - 43 Agosto DE, Hughes-Hassell S. People, places, and questions: An investigation of the everyday life information-seeking behaviors of urban young adults[J]. Library & Information Science Research, 2005, 27(2): 141-63
 - 44 Laplante A, Downie JS. The utilitarian and hedonic outcomes of music information-seeking in everyday life[J]. Library & Information Science Research, 2011, 33(3): 202-10
 - 45 Chu SKW, Tse SK, Loh EKY, Chow K. Collaborative inquiry project-based learning: Effects on reading ability and interests[J]. Library & Information Science Research, 2011, 33(3): 236-43
 - 46 Agosto DE. A model of young people's decision-making in using the Web[J]. Library & Information Science Research, 2002, 24(4): 311-41
 - 47 Toms EG, O'Brien HL. Understanding the information and communication technology needs of the e-humanist[J]. Journal of Documentation, 2008, 64(1): 102-30
 - 48 李武, 杨琳. 科研人员接受 OA 知识库的影响因素分析——一项基于先导访谈的探索性研究[J]. 大学图书馆学报, 2010(3): 98-103
 - 49 Roszkowski MJ, Baky JS, Jones DB. So which score on the LibQUAL +TM tells me if library users are satisfied? [J]. Library & Information Science Research, 2005, 27(4): 424-39
 - 50 Shachaf P, Horowitz S. Are virtual reference services color blind? [J]. Library & Information Science Research, 2007, 28(4): 501-20
 - 51 赵丹信. 图书情报领域中内容分析法研究进展与趋势[J]. 图书馆学研究, 2008(2): 6-8
 - 52 邱均平, 邹菲. 国外内容分析法的研究概况及进展[J]. 图书情报知识, 2004(6): 6-8
 - 53 Myers MD. Qualitative research in information systems[J]. Management Information Systems Quarterly, 1997, 21: 241-242
 - 54 任立肖, 沙勇忠. 网络内容分析研究[J]. 情报理论与实践, 2005, 28(5): 523-526
 - 55 邱均平, 邹菲. 关于内容分析法的研究[J]. 中国图书馆学报, 2004, 30(2): 12-17
 - 56 朱亮, 孟宪学. 文献计量法与内容分析法比较研究[J]. 图书馆工作与研究, 2013, (6): 64-66
 - 57 成波, 黄晓斌. 国内外网络内容分析应用现状研究[J]. 图书情报工作, 2007, 51(9): 45-50
 - 58 陈向明. 扎根理论的思路和方法[J]. 教育研究与实验, 1999(4): 58-63
 - 59 Denzin NK. The research act: A theoretical introduction to sociological methods[M]. New Jersey: Transaction Publishers, 1970: 301
 - 60 Tashakkori A, Teddlie C. Handbook of mixed methods in social & behavioral research[M]. New York: Sage, 2003: 9
 - 61 Sieber SD. The integration of fieldwork and survey methods[J]. American Journal of Sociology, 1973, 78(6): 1335-1359
 - 62 Venkatesh V, Brown SA, Bala H. Bridging the qualitative-quantitative divide: guidelines for conducting mixed methods research in information systems[J]. MIS Quarterly, 2013, 37(1): 21-54
 - 63 Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research[M]. Thousand Oaks, CA: Sage, 2011: 54-59
 - 64 陈向明. 质性研究的新发展及其对社会科学研究的意义[J]. 教育研究与实验, 2008(2): 14-18

(收稿日期: 2014-05-08)