

熊永兰, 张志强, Wei Yongping, 等. 基于科学知识图谱的水文化变迁研究方法探析[J]. 地球科学进展, 2014, 29(1): 92-103, doi:10.11867/j.issn.1001-8166.2014.01.0092. [Xiong Yonglan, Zhang Zhiqiang, Wei Yongping, et al. On a quantitative method for water culture evolution research: Mapping knowledge domains[J]. Advances in Earth Science, 2014, 29(1): 92-103, doi:10.11867/j.issn.1001-8166.2014.01.0092.]

# 基于科学知识图谱的水文化变迁研究方法探析\*

熊永兰<sup>1,2</sup>, 张志强<sup>1</sup>, Wei Yongping<sup>3</sup>, 刘志辉<sup>4</sup>, 程国栋<sup>5</sup>

(1. 中国科学院资源环境科学信息中心, 甘肃 兰州 730000; 2. 西北师范大学地理与环境科学学院, 甘肃 兰州 730070; 3. 澳大利亚墨尔本大学澳中水资源研究中心, 墨尔本 3010; 4. 中国科学技术信息研究所, 北京 100038; 5. 中国科学院寒区旱区环境与工程研究所, 甘肃 兰州 730000)

**摘要:**文化变革(包括知识革新、思想解放、口号宣传等)是水资源管理政策变迁的重要原因。目前针对水文化的相关研究多数采用定性分析方法,难以客观准确地揭示水文化变迁的阶段性特点。作为一种新兴的知识发现方法,科学知识图谱能够全方位揭示知识的来源和发展规律,具有用作一种水文变迁定量研究方法的潜力。通过提取《人民日报》中与水相关的文章关键词,绘制出1946—2012年我国水文化变迁的知识图谱并对图谱解读,分析出我国水文化变迁的轨迹;将水文化变迁与当时的水政策、水利开发阶段和水文极端事件进行关联,发现具有较好的协同性,显示出科学知识图谱是定量研究水文化变迁的有效分析方法。进一步深化报纸普适性、关键词提取准确性以及可视化分析精确性等方面工作是提升基于科学知识图谱的水文化变迁定量研究水平的可靠途径。

**关键词:**水资源管理;水文化;人民日报;科学知识图谱;VOSviewer

**中图分类号:** P349;TV213.4 **文献标志码:** A **文章编号:** 1001-8166(2014)01-0092-12

## 1 引言

当前水资源管理正在经历重大的范式转变。过去研究者与管理人员大都注重从自然系统及工程技术的角度来研究水资源管理问题,而忽视了社会学的作用<sup>[1]</sup>。近年来少数研究人员已开始意识到可持续的水资源管理更应关注社会文化的作用<sup>[2,3]</sup>。自然生态系统变化、社会文化变迁和政策法规演化是主导水资源管理的基本因素,三者相互作用促进水资源管理的发展<sup>[4]</sup>。因此,发展可持续的水资源管理科学与政策体系不仅要认识水生态系统演化的自然规律,更应关注其与社会文化变迁以及水管理

政策发展的动态关系。

关于文化的定义,国内外学者提出了多种表述方式,主要从社会学和人类学的角度来阐述。根据对文化内涵的理解,我们认为水文化应是历史上形成、不断演进且为某一社会群体所认同的关于水的世界观、价值观以及相关的思想和行为范式。水文化变迁主要指人类治水理念的变化,如从除害兴利到可持续利用再到应对气候变化等。以往研究者与管理者都是从历史学、社会学和人类学的角度,采取归纳推演<sup>[2,3,5,6]</sup>、案例分析<sup>[7]</sup>、半结构式访谈<sup>[8]</sup>等定性分析的方法来研究水文化,而仅有少数学者从定量分析的角度来探讨水文化,如 Yip 等<sup>[9]</sup>通过问

\* 收稿日期:2013-11-27;修回日期:2014-01-02.

\* 基金项目:国家自然科学基金重点项目“流域文化变迁与生态演化相互作用对流域水生态政策影响的机理研究——黑河与墨累—达令河对比研究”(编号:91125007);澳大利亚研究委员会发现项目“理解文化和生态对流域退化遏制政策的驱动机制”(编号:DP120102917)资助。

作者简介:熊永兰(1980-),女,副研究员,重庆万州人,主要从事水资源管理研究. E-mail: xiongy@llas.ac.cn

\* 通讯作者:张志强(1964-),男,研究员,甘肃定西人,主要从事生态经济学与可持续发展研究. E-mail: zhangzq@lzb.ac.cn

卷调查和 T 模型来量化可持续文化变化的程度; Harmsworth 等<sup>[10]</sup>利用文化健康指数(Cultural Health Index, CHI)来评估河流健康状况;定量研究水文化变迁的则更鲜见。

文化的社会性决定了文化的传播性,而文化的传播必须有载体。在现代社会中,大众媒介已成为文化传播最重要、最高效的载体。尽管在过去 20~30 年电子媒介得到了快速发展,但报纸同其他媒体一样是新闻报道的主要来源<sup>[11]</sup>。而且,与其他媒体相比,人们更倾向于信任印刷媒体,并且易于吸收其所报道的内容<sup>[12]</sup>。因此,报纸可以作为研究水文化的可靠信息来源。科学知识图谱是近年来科学计量学、信息计量学等领域新兴的研究方法,它将复杂的知识领域通过数据挖掘、信息处理、知识计量和图形绘制来揭示知识来源及其发展规律,并且以图形表达相关领域知识结构关系与演进规律<sup>[13]</sup>。

本研究希望通过利用科学知识图谱的研究方法定量研究水文化变迁,为科学、定量分析水文化提供新的方法。

## 2 基于科学知识图谱的水文化变迁研究方法

### 2.1 基于科学知识图谱的水文化变迁研究框架

科学知识图谱(mapping knowledge domains)是一个新兴的、横跨科学学、信息科学、科学计量学、计算机科学和应用数学等领域的交叉前沿领域,是对知识进行挖掘、分析、分类、导航和可视化(制图)的过程<sup>[14]</sup>。它可以揭示一个知识领域的结构关系和演变过程。知识图谱所描绘的对象主要包括:①从事科学技术活动和作为知识载体的人,包括科学家、技术专家、项目组、研究团体或某一知识领域共同体;②显性或编码化的知识,如论文、专利、所学课程、数据库或类似的应用等;③过程或方法,包括研究问题和解决问题的过程或方法、组织的业务流程以及相关的知识投入等<sup>[15]</sup>。科学知识图谱的基本

分析方法包括引文分析、被引分析、多元统计分析、词频分析以及社会网络分析<sup>[15]</sup>。

水文化常由非结构化的社会符号(如文字)来表达,其载体为大众媒介。因此,一定时期的水文化可由媒体或者文献中特定的词汇来表征,水文化变迁可以通过研究媒体或文献中词汇和词频随时间的变化规律来揭示。除研究对象外,水文化变迁研究与科学知识图谱极具相似性(表 1),我们可以借鉴科学知识图谱的原理与方法来研究水文化变迁。根据 Cobo 等<sup>[16]</sup>关于知识图谱绘制的流程,结合本研究的实际情况,提出了水文化变迁的研究框架(图 1)。该研究框架包括相互关联的 7 个步骤。这些步骤将在以下的 2.2 节至 2.5 节中详细阐述。

### 2.2 数据来源

#### 2.2.1 媒体选择

尽管在过去 20~30 年电子媒介得到了快速发展,但报纸仍然是新闻报道的主要来源<sup>[11]</sup>。报纸能够提供广播媒体所没有的对某个主题的深入报道,因而是研究文化的有效工具。报纸的报道内容也覆盖了新兴社会媒体,如博客和其他数据聚集工具。另外,报纸的权威性和公信力也是其他媒体所不能替代的。报纸也具有很长时期连续存档的特性,可以从历史的角度来分析公众舆论因而可以作为研究水文化的可靠信息来源。

根据所面向的读者群的差异,报业市场一般分为小报和严肃报纸 2 类。受商业而不是公众利益的驱动,小报的主要内容是人们感兴趣的故事、娱乐新闻、体育新闻和丑闻。而严肃的报纸则面向国家政治和国际新闻,代表了主流的社会价值取向,是公众利益的体现。因此,在本研究中选择严肃类的报纸作为数据的来源。

《人民日报》是中国关于意识形态的主流喉舌,是传递文化、政策的主流声音,是我国最具影响力和权威性的报纸,在全国传播范围广泛。另外,《人民日报》于 1946 年 5 月发行,是我国发行至今有电子

表 1 水文化变迁研究与科学知识图谱之间的关系

Table 1 Relationship between research on changes of water culture and mapping knowledge domains

	研究对象	研究目标	研究方法	数据源	数据结构
科学知识图谱	科学知识	揭示知识领域的结构关系和演变过程	社会网络分析、词频分析、引文分析、被引分析、多元统计分析	科技文献	结构化
水文化变迁研究	水文化	揭示水文化的演变过程	基于共词的社会网络分析	媒体或者文献	非结构化

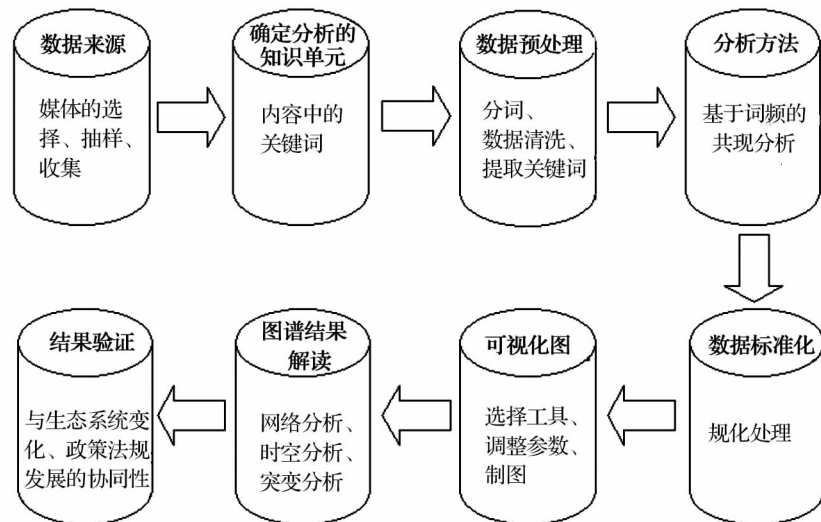


图 1 基于科学知识图谱的水文化变迁研究框架

Fig. 1 Research framework on changes of water culture based on mapping knowledge domains

存储的最早的报纸。因此,本研究选择《人民日报》作为反映中国水文化的主要媒体,并且将 1946—2012 年的报纸作为研究对象,通过研究其有关水的论述和意识形态的报道来反映中国关于水文化的发展脉络。

### 2.2.2 抽样方法

由于对所有年份的全部报纸文章进行分析超出了本文的研究范围,因此采取抽样的方式来进行研究。对于报纸而言,简单随机抽样、自然周抽样和构造周抽样<sup>[17-19]</sup>是 3 种主要的抽样方法。简单随机抽样不能反映媒体内容的周期性特征,而构造周抽样可控制“系统性变化”因素,但该方法忽略了周与周之间的差异,可能会错过重要的“新闻周”(如自 1988 年以来,中国开始实施的“中国水周”活动),因此,本文选择构造周抽样和自然周抽样。

对于每年的报纸,抽取 4 个新闻周,包括 2 个构造周和 2 个自然周。构造周和自然周都分别从每年的上、下半年各抽取 1 个。构造周星期一到星期天分别从上、下半年的 26 个星期一到星期天中随机抽取。自然周的抽样在 1988 年以前随机抽取,在 1988 年以后,根据“中国水周”日期的变化而变化,即 1988—1993 年为 7 月 1~7 日;1994 年以后为 3 月 22~28 日。

### 2.2.3 数据收集

通过定制的网页爬取程序,从“《人民日报》图文数据库”里下载所需日期的报纸,并以.xls 的格式保存。根据抽样的结果,共下载《人民日报》1946—2012 年报纸 2 043 份,文章数量为 148 086 篇。然后

采取人工判读的方法,提取出与水相关的文章,包括涉及水政与水利经济、水资源、农田水利、水土保持、防洪与河道整治、水利管理、环境水利等领域的文章共 2 026 篇。

### 2.3 数据预处理

词频分析方法是文献计量学的传统方法之一,也是科学知识图谱的基本方法。词频分析方法所依据的理论是齐普夫定律(Zipf's law),它揭示了文献中词汇出现频率的分布规律<sup>[20]</sup>。词频分析方法被国内外许多科学计量学研究者应用于学科前沿研究<sup>[15]</sup>。因此,将关键词作为分析的知识单元,采用词频分析和共现分析相结合的方法来研究水文化的变迁。关键词的抽取主要有 2 种方式:全文直接抽取和字段间接抽取<sup>[21]</sup>。不同于科技文章,报纸本身并没有关键词,因此,关键词的抽取只能采取全文直接抽取的方式。首先要对文章进行分词,然后进行词频统计,结合词频—逆向文件频率(Term Frequency-Inverse Document Frequency, TF-IDF)算法提取关键词。

#### 2.3.1 构建分词词典

由于关于水资源或者水利方面的最新词典/叙词表是水利部信息研究所 1998 年编制的《水利水电科技主题词表》,近 15 年来新出现的词并未纳入词表中,而且科技主题词表不能体现媒体语言的特征,所以通过构建专业领域期刊词典,并整合中国科学院计算研究所研制的分词系统 ICTCLAS 5.0 自带词典的方式,来构建适用于《人民日报》的词典。

首先,利用水资源核心期刊列表,选取《水土保

持学报》、《水土保持通报》、《水土保持研究》、《节水灌溉》、《水利学报》、《中国水利》、《水科学进展》、《水利水电科技进展》、《中国农村水利水电》、《水资源保护》、《水生态学杂志》、《人民黄河》等12种主要的水资源核心期刊作为构建专业领域词典的主要来源,从CNKI获取题录数据,并抽取出现这些期刊中的关键词构建专业领域期刊词典;其次,将专业领域期刊词典与ICTCLAS 5.0自带词典整合为自定义专业领域词典。

### 2.3.2 分词

ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System)分词法是由中国科学院计算研究所在多年工作积累的基础上研制出来的,其官方网站为:<http://www.ictclas.org/>。目前,ICTCLAS分词法已经经过了国内和国际权威的公开评测,获得了5万客户的认可,具有综合性能最优的特点。因此,基于自定义专业领域词典,利用ICTCLAS 5.0对语料库中的报纸文章进行分词,分词后共得到64 506个词。

### 2.3.3 数据清洗

对分词结果的清洗,包括去除无意义的单个字、合并相同词、利用停用词表去掉停用词。停用词使用哈尔滨工业大学的停用词表,停用词包括标点、部分连词、语气词、代词等,共506个中文词。另外,汉语短语的语法和语义应该正确,比如“种方法”、“式处理”等是不合理的短语,因为它们不能组成一个合法的语法单位,语义也不完整。数据清洗后得到34 950个词/短语。

### 2.3.4 提取关键词

在分析学科领域的重点时,一般通过统计关键词绝对频次的方法来分析,但这种方法缺乏横向的可比性。而TF-IDF计算方法不但考虑了关键词在其发生文档中的频率,还考虑了与全部文档中该词频率的关系,这使我们易于分析不同时间段水文化的侧重点。为此,首先通过年度一频次统计构建词/短语的年度分布;其次利用TF-IDF规范化词/短语的年度一频次矩阵,获得年度特色关键词。词/短语在某年度的规范化结果为 $W_j$ :

$$W_j = TF_j \times IDF \quad (1)$$

$$IDF = \log\left(\frac{|D|}{DF_j}\right) \quad (2)$$

式中: $TF_j$ 是关键词 $Term_j$ 在某年度列表中出现频次; $DF_j$ 是包含关键词 $Term_j$ 的年度数量, $|D|$ 是整个时间段的时间跨度。

为了突出各阶段主要研究内容及特色研究内容,本文选择了那些在整个数据集(共875个词)或某个时间段中TF/IDF规范化值 $W_j$ 加权较高(共5 000个词)或在某时间段词频较高的词(共17 423个词)进行加权处理,最后获得22 393个关键词。

## 2.4 数据标准化处理与分析

关键词共现分析法的思想来源于文献计量学的引文耦合与共被引概念,即当2个能够表达某一学科领域研究主题或研究方向的关键词在同一篇文献中出现时,表明这2个词之间具有一定的内在关系,并且出现的次数越多,表明关系越密切、距离越近。利用因子分析、聚类分析和多维尺度分析等多元统计方法,可以进一步按这种“距离”将一个学科内的重要主题词或关键词加以分类,从而归纳出该学科的研究热点、结构与范式<sup>[22]</sup>。在一系列的时间区间里进行比较,可以发现学科的发展变化趋势。本文用这种方法来研究媒体中关键词之间的关系进而归纳出水文化不同时段热点、范式以及发展演化的趋势。

首先,基于已获得的关键词,分时间段构建关键词的共现矩阵。为了体现重要政治事件和水事件在水文化演化中的作用,按照2种方式将67年的《人民日报》划分为不同的时间段。一种方式是依据重大的社会政治事件(如新中国成立、文化大革命、设立经济特区等)和水事件(特大洪水、特大干旱、重要的水政策法规的颁布)将其划分为9个时间段:1946—1949年,1950—1960年,1961—1965年,1966—1976年,1977—1980年,1981—1990年,1991—1997年,1998—2010年,2011—2012年;另一种方式是与中国五年计划相一致,即划分为14个时间段(其中个别阶段,即1946—1952年、1963—1965年,不是实际的五年计划阶段):1946—1952年,1953—1957年,1958—1962年,1963—1965年,1966—1970年,1971—1975年,1976—1980年,1981—1985年,1986—1990年,1991—1995年,1996—2000年,2001—2005年,2006—2010年,2011—2012年。

基于初始共现频次的共现关系分析可以明显突出某一时间段的水文化热点。但某些词会因频次过高而表现出较强的共现关系,而无法突出新型或特色文化。本文提出了一种基于加权的共现关系分析方法,其目的是为了更加有效地同时突出热点文化与新型或特色文化。具体的分析方法如下:

$$WR_{ij} = CoWord_{ij} \times R_{ij} \quad (3)$$

$$R_{ij} = \frac{CoWord_{ij}}{\text{Max}(FRE_i, FRE_j)} \quad (4)$$

式中:  $WR_{ij}$  是关键词加权后的关系强度;  $R_{ij}$  是关键词词频规范化后的关系强度, 其范围为  $[0, 1]$ , 目的是消除部分词的规模影响;  $CoWord_{ij}$  是关键词  $Term_i$  和  $Term_j$  共现的初始频次;  $FRE_i$  和  $FRE_j$  分别是其在某阶段数据集中的出现频次。

## 2.5 制作可视化图谱

选择专门用于绘制知识图谱的、可免费使用的软件 VOSviewer 来构建可视化图谱。VOSviewer 是荷兰莱顿大学 Van Eck 与 Waltman 研发的可视化软件。为了展现地图上的元素, VOSviewer<sup>[23-26]</sup> 使用相似性度量从共现矩阵中创建了相似矩阵, 从而创建一个二维图, 图中元素之间的距离反映其相似性, 并且使用重要的标签, 便于研究人员发现重要的主题。VOSviewer 允许通过标签视图、密度视图、聚类密度视图和分散视图 4 种方式来进行浏览, 为了直观地判别核心主题, 本文选择密度视图方式。通过比较 2 种时间段划分方式的制图结果, 认为按照五年计划的方式划分的结果更能体现出主题的演化, 其具体的可视化结果如图 2 所示 (根据重大事件划分时段的知识图谱在此不再展示)。

## 2.6 图谱结果的解读

根据 VOSviewer 绘制的分阶段水文化图谱, 通过不同的颜色、大小和年轮的不同厚度和视角等来了解核心主题和热点动向。VOSviewer 所展示的关系密度图谱中, 每个节点根据其密度有一种颜色, 红色代表关注较多的主题或领域。一个节点越大, 表示其权重越大, 其颜色越接近于红色。相反, 如果其权重越小, 则其颜色越接近于蓝色。密度视图有助于快速获得图谱中重要内容的概貌。据此, 结合对《人民日报》相关文章的解读, 对于每个阶段的图谱解读见表 2。

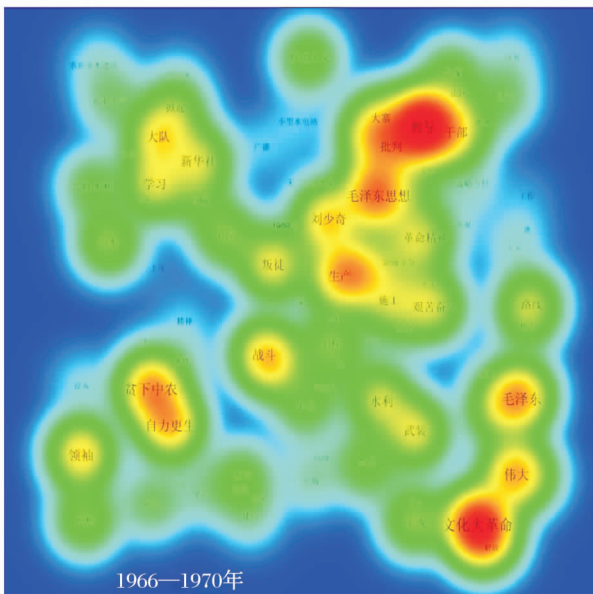
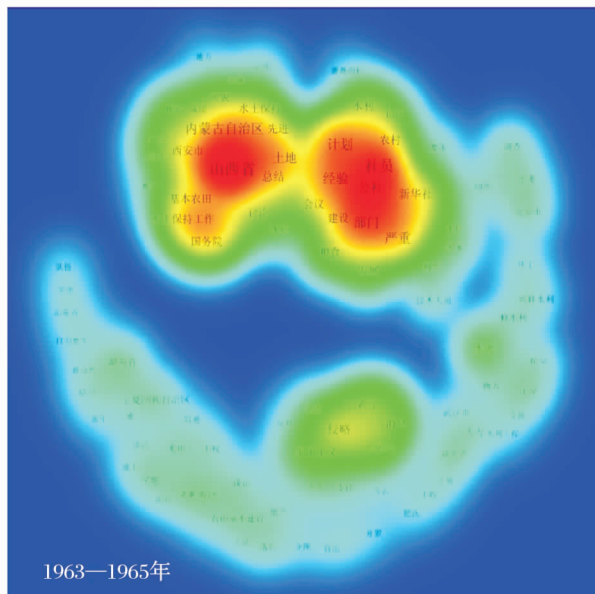
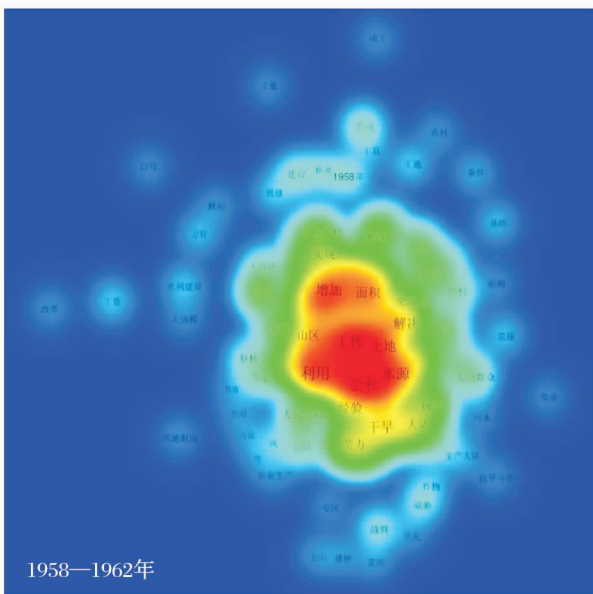
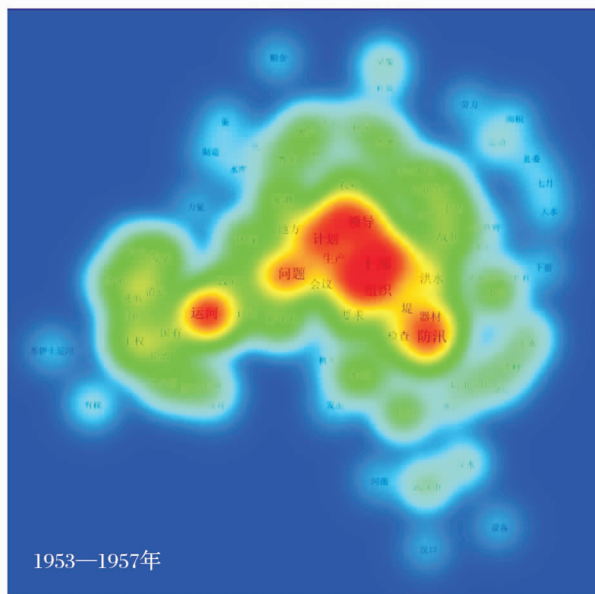
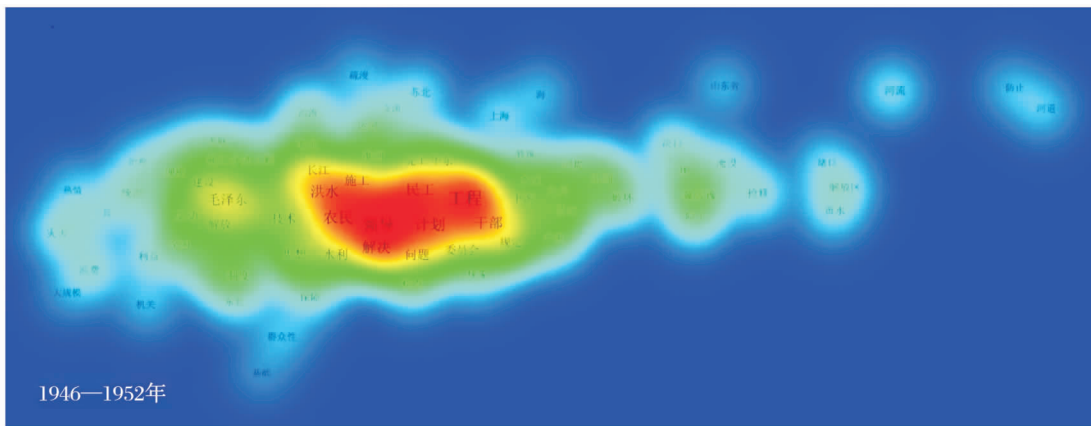
## 3 基于科学知识图谱的水文化变迁研究方法的检验

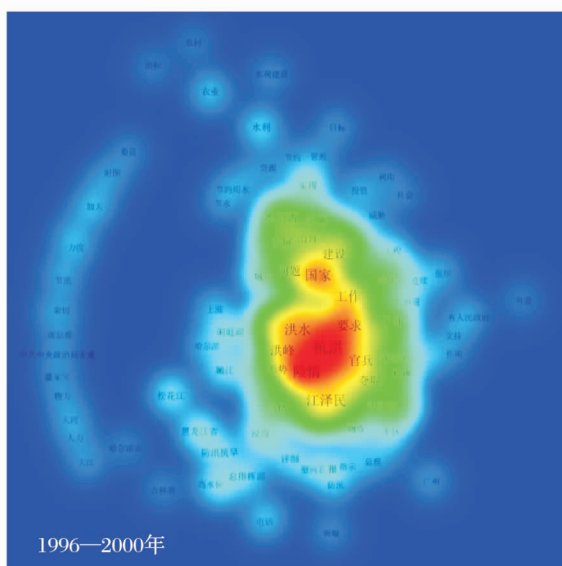
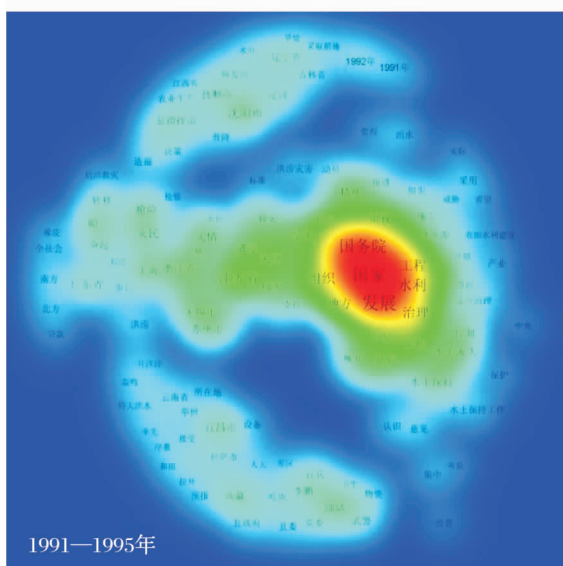
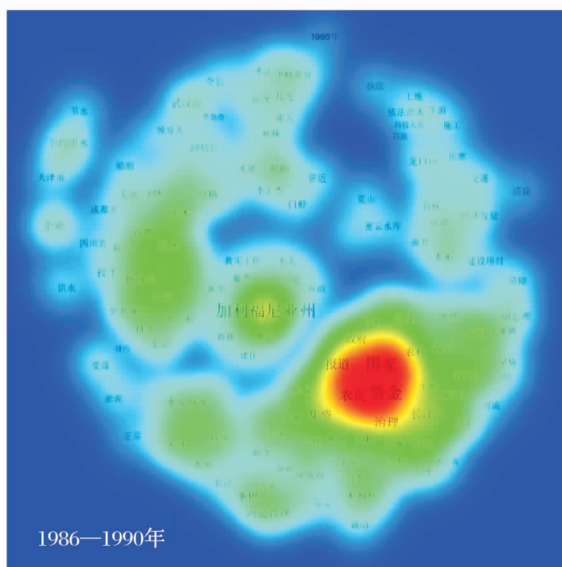
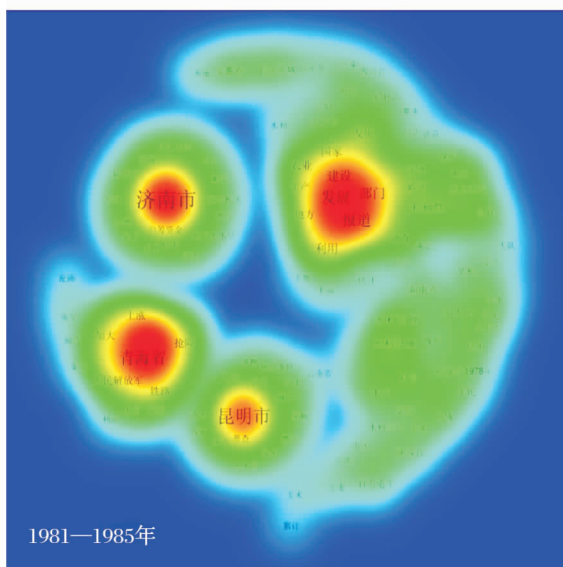
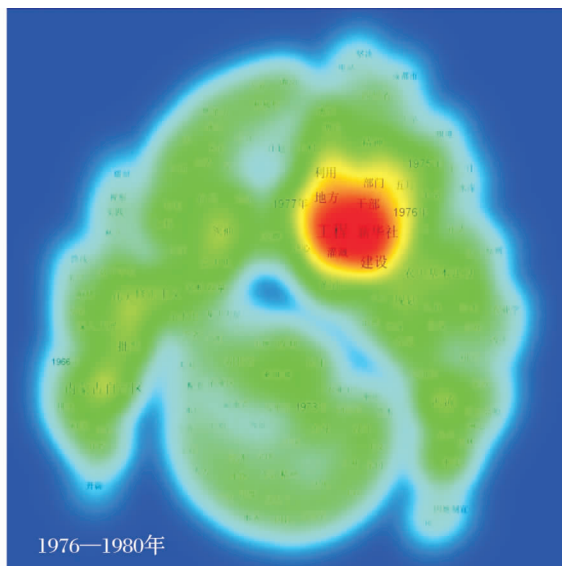
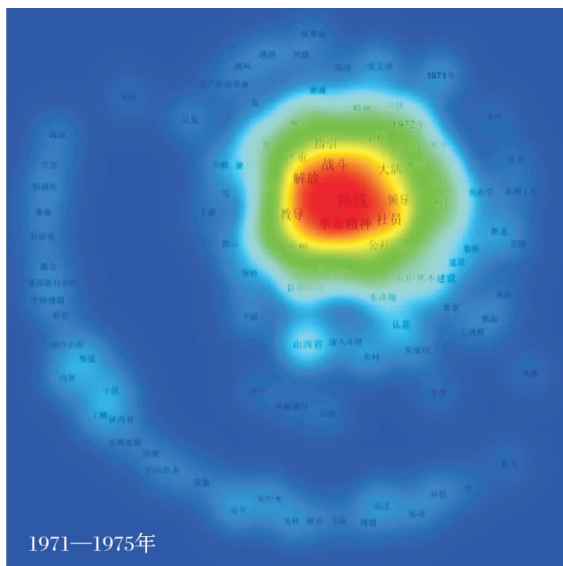
基于科学知识图谱的水文化变迁研究框架得出的我国水文化变迁的趋势是否正确, 换句话说, 该方法是否有效, 还有待于进一步检验。新闻报道因其导向作用而反映和影响着舆论和政策的形成<sup>[27-29]</sup>。媒体为了凸显某些问题的重要性, 往往会对这些问题进行特殊对待, 比如更加频繁的报道和放在版面中更加突出的位置<sup>[30]</sup>。媒体、舆论(文化)和政策之间存在着协同效应<sup>[31]</sup>。Norgaard<sup>[4]</sup> 认为, 在社会生

态共进化系统中, 社会文化变迁与自然生态系统变化、政策法规演化等具有协同性。水文的极端事件(洪水、干旱以及近年出现的气候变化)是流域水资源系统作为一个特殊的社会生态共进化系统的主要特征, 虽然因时因地而异, 防洪抗旱一直是水资源管理的两大主要功能。多数情况下, 投资是政策的具体体现。因此, 本文用我国不同时期的主要水政策法规、水利投资重点以及洪灾旱灾情况来检验本文得出的基于科学知识图谱的水文化变迁趋势正确与否(表 3)。

水资源的开发利用主要是满足社会经济发展 5 个方面的需求: 饮水保障、防洪安全、粮食供给、经济发展和生态环境<sup>[32]</sup>。饮水保障、防洪安全和粮食供给是水利开发的基础阶段, 主要是为了满足人们安全性的需求; 经济发展需求主要是为了满足人们对高物质生活的需求; 而生态环境需求主要是为了满足人们健康环境、资源可持续利用的需求, 这是更高层次的需求。从表 3 中可以看出, 由《人民日报》反映的水文化变迁与我国主要的水政策法规具有很好的协同性。建国初期, 我国的水政策法规主要集中在防汛抗洪和治理淮河流域。新中国成立后开始实施治河工程, 如荆江分洪工程、淮河治理工程等, 这主要是因为 1950 年淮河流域发生特大洪水, 造成严重水灾, 在此之后毛泽东对根治淮河进行了 4 次批示, 1950 年 10 月, 政务院做出了《关于治理淮河的决定》, 确定了“蓄泄兼筹, 以达根治之目的”的治淮方针。而这一时期的水文化主题正是以抗洪和防洪为目标的治河工程。

从“一五”开始到“六五”期间, 我国的水政策法规开始转向以农田水利建设与管理 and 水土保持为主。毛泽东提出“水利是农业的命脉”, 因此, 新中国成立以后, 国家开始兴修水利发展农业。“一五”期间, 水利部就向中共中央提交了一系列农田水利工作的报告, 如《中央水利部党组关于农田水利工作会议的综合报告》、《中共中央同意水利部党组〈关于华北五省农田水利工作会议纪要的报告〉》、《中共中央、国务院关于今冬明春大规模地开展兴修农田水利和积肥运动的决定》等。“二五”期间, 相关部委制定的农田水利方面的政策有《水利部、交通部关于公路沿线兴修农田水利工程需注意事项的联合通知》、《中央转批农业部和水利电力部关于加强水利管理工作的十条意见》、《中共中央关于抗旱备荒的指示》等。3 年自然灾害时期和文化大革命期间, 国家制定的农田水利方面的政策法规较少。





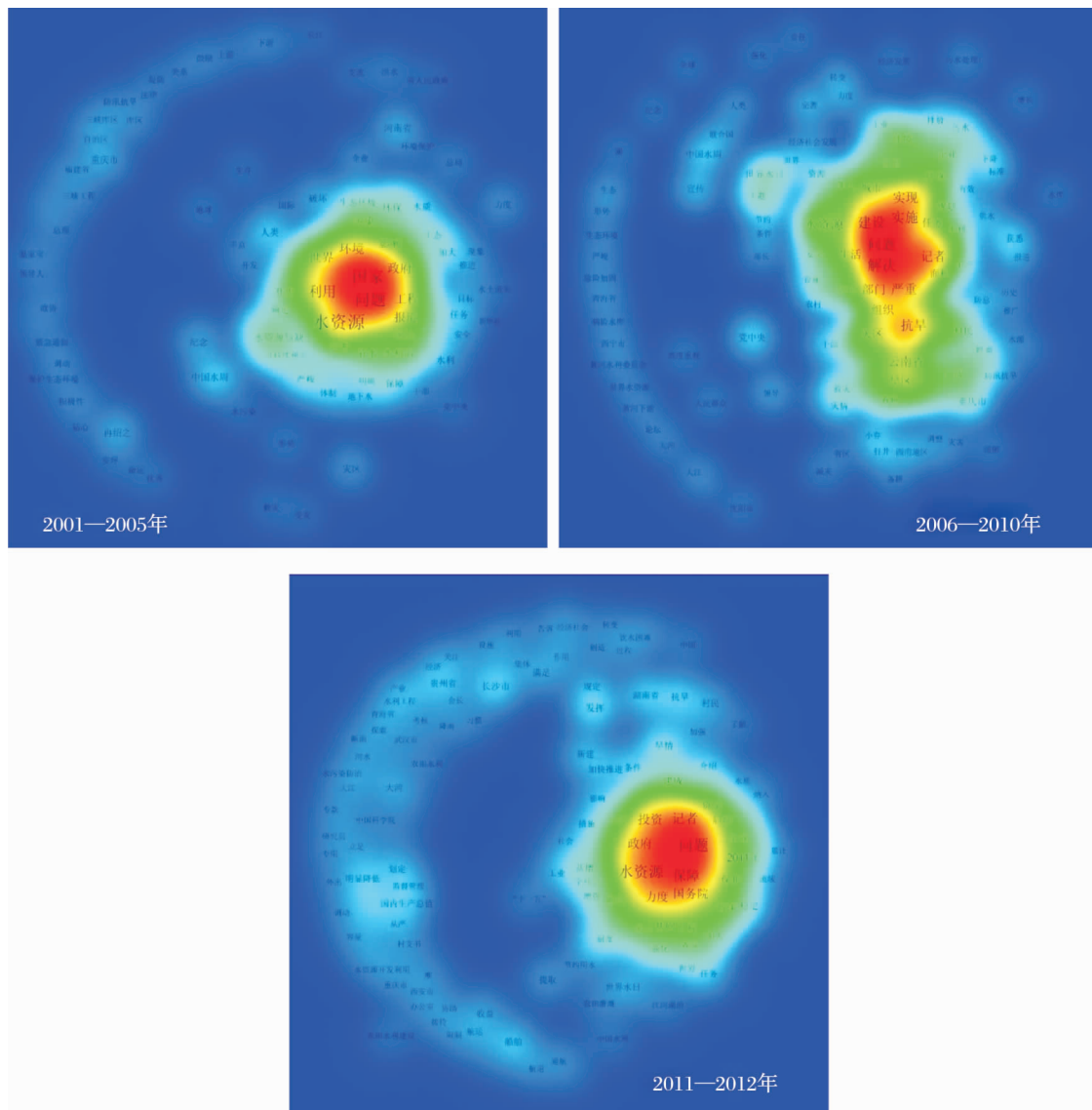


图 2 《人民日报》水文化变迁图谱 (1946—2012 年)

Fig. 2 Diagram on changes of water culture based on People's Daily (1946-2012)

表 2 基于科学知识图谱的中国水文化变迁轨迹

Table 2 Trajectories on changes of water culture in China based on mapping knowledge domains

阶段划分	热点或核心主题	热点或核心主题的体现
1946—1952 年	抗洪 治河工程	工程、计划、领导、农民、解决、洪水、长江、民工 毛泽东、建设、解放、荆江分洪工程
1953—1957 年	农业灌溉与农业生产 防洪	计划、领导、干部、组织、生产、农业、灌溉、问题 防汛、器材、堤、检查、洪水
1958—1962 年	苏伊士运河的主权问题 抗旱	运河、政府、讨论、主权、国有、通航 工作、利用、公社、水源、经验、干旱、抗旱、解决
1963—1965 年	兴修农田水利 农田水利建设	增加、面积、兴修水利、农民、劳动力 社员、公社、部门、建设、经验、计划、农村、水利、严重
1966—1970 年	各省的水土保持工作 农业学大寨和兴修水利 文化大革命	陕西省、内蒙古自治区、土地、西安、水土保持、基本农田、水土保持工作 毛泽东思想、教导、干部、批判、大寨、灌溉、生产、水利、大队、学习、兴修水利、治水 文化大革命、解放、伟大、毛泽东、贫下中农、自力更生、领袖



续表2

阶段划分	热点或核心主题	热点或核心主题的体现
1971—1975年	文化大革命时期的农田基本建设	路线、革命精神、社员、大队、解放、战斗、教导、思想、农业生产、农田基本建设、生产
1976—1980年	农业灌溉和农田基本建设	工程、建设、灌溉、利用、干部、地方、农田基本建设
1981—1985年	水利工程即水电站的建设	发展、建设、部门、利用、农业、发电、水电站
	济南市的水短缺问题	济南市、地下水水位、枯竭、断流、泉群、自筹资金
	青海省的防洪	青海省、抢险、上涨、加大、铁路、堤防、雨量
1986—1990年	昆明市的水资源普查	昆明市、普查、找水、增添、预测、涌水、储量
	河流治理,包括污染、防洪和河道管理	国家、资金、治理、农民、长江、海河、污染、水质、防汛、大河、河道管理、治水
1991—1995年	水土流失与治理	国家、发展、工程、水利、治理、水土流失、水土保持、面积、控制、综合治理
	江苏、浙江、上海等地洪灾	组织、地方、受灾、救灾、灾区、抗灾、江苏省、浙江省、上海
1996—2000年	抗洪抢险	抗洪、险情、洪水、洪峰、官兵、要求、江泽民、汛情
	水资源管理	国家、水资源、建设、问题、治理、发展、实施、节水
2001—2005年	水资源问题	水资源、问题、国家、环境、利用、工程、节水、水资源短缺、污染、环保、生态环境
2006—2010年	水资源管理	问题、解决、建设、实施、实现、水资源、规划、任务、生活、环保、应对、环境
	西南地区的旱灾	抗旱、组织、严重、灾区、云南省、旱区、广西
2011—2012年	水资源管理	水资源、保障、问题、国务院、力度、政府、投资、政策、保护、目标、基础、强化、建成、效益、加大

表3 1946—2012年分阶段水文化主题、水政策法规和水文极端事件

Table 3 Subject of water culture, water policies and extreme hydrological events by different periods during 1946-2012

时段划分	核心主题	主要的水政策法规	水利投资重点	洪灾与旱灾
1946—1952年	抗洪;治河工程(防洪)	防汛抗洪;治理淮河	防洪除涝	洪灾
1953—1957年 ("一五"期间)	农业灌溉与农业生产;防洪;苏伊士运河的主权问题	农田水利工作;治理黄河; 水土保持	防洪除涝	洪灾
1958—1962年 ("二五"期间)	抗旱;兴修农田水利	水利工程建设;抗旱防汛	水库与水电	旱灾
1963—1965年 (调整巩固充实提高时期)	农田水利建设;各省的水土保持工作	水土保持;水产;黄河治理; 水电	水库与水电	洪灾、旱灾
1966—1970年 ("三五"期间)	农业学大寨和兴修水利;文化大革命	N/A	N/A	N/A
1971—1975年 ("四五"期间)	文化大革命时期的农田基本建设			旱灾、洪灾
1976—1980年 ("五五"期间)	农业灌溉和农田基本建设	灌溉;农田水利工程管理;水土流失治理	水库与水电	旱灾
1981—1985年 ("六五"期间)	水利工程即水电站的建设(主要用于农业发展);济南市的水短缺问题;青海省的防洪;昆明市的水资源普查	水电工程管理;水污染防治	水库与水电	旱灾
1986—1990年 ("七五"期间)	河流治理,包括污染、防洪和河道管理	河道管理;内河航运; 《水法》	防洪除涝	旱灾
1991—1995年 ("八五"期间)	水土流失与治理;江苏、浙江、上海等地的洪灾	河道管理;取水许可管理;水土保持;水生态环境防治与保护;供水管理	水库与水电	洪灾
1996—2000年 ("九五"期间)	抗洪抢险;水资源管理	水利工程管理;水土保持;水生态环境防治与保护;防洪	防洪除涝	洪灾、旱灾
2001—2005年 ("十五"期间)	水资源问题(主要是生态环境和水短缺问题)	水利工程管理;水土保持;水价;水污染防治	防洪除涝	洪灾、旱灾
2006—2010年 ("十一五"期间)	水资源管理(主要是对水污染的治理和规划);西南地区的旱灾	水利工程管理;流域综合规划;水价;污染监控与防治	防洪除涝	洪灾、旱灾
2011—2012年	水资源管理(主要体现在政策和发展目标方面)	水资源管理;流域管理	N/A	洪灾、旱灾

注:N/A表示无统计数据。2000年之前的洪灾与旱灾根据《中国水利年鉴2001》之《20世纪中国水利大事年表》统计,2000年之后根据《中国水旱灾害公报》(2006—2011)统计。水利投资重点则由水利投资比重最大的类型来反映,数据来源于《中国水利统计年鉴2010》

文化大革命之后,国家又制定了灌溉和农田水利工程管理方面的政策法规。水土保持是有效改善农业生产基础条件和生态环境、促进农业增产和农民增收的有效途径。从建国开始,国家就开始大力推行水土保持工作。1952年中央人民政府发布了《关于发动群众继续开展防旱、抗旱运动并大力推行水土保持工作的指示》、1957年发布了《中华人民共和国水土保持暂行纲要》,从1980年开始,国家陆续出台了流域层面综合治理水土流失的政策法规。在建国后的30多年中,农业灌溉与农业生产、兴修农田水利、农田水利、农田基本建设等方面的主题是每个五年计划时期水文化的核心主题,而水土保持这一主题则贯穿在农田基本建设当中。

“七五”以来,我国的水资源管理从供水管理向需水管理转变<sup>[33]</sup>,包括调整经济产业结构和用水结构、采取节水措施、控制污染等,因此,颁布了相应的政策法规来促进这一管理方式的转变,如黄河实施的水量合理分配制度、各流域取水许可管理权限、水价政策、重要流域水污染防治规划等。水文化的核心主题也由农田水利建设与管理转向河流治理、面向水短缺和生态环境的水资源管理。2011年中央一号文件《中共中央国务院关于加快水利改革发展的决定》的发布标志着我国的水利事业进入了一个新的发展阶段。我国将实行最严格的水资源管理制度,大力发展民生水利,凸显水利保障经济安全、生态安全和国家安全的作用。相应的水文化核心主题也体现在水资源管理政策和发展目标、规划等方面。

从表3可以看出,水库水电建设和防洪除涝从建国到21世纪初都是我国水利投资的重点,与此同时,由《人民日报》体现的主题也集中在防洪和以灌溉、水库和农村小水电为主的农田基本建设方面。“十五”以来,由《人民日报》体现的主题开始转向水生态环境和水资源管理,而我国的水利投资重点尽管仍然是防洪除涝,但对供水的投资力度加大,对水保及生态也开始投入。表3列出了自建国以来我国的主要洪旱灾害。《人民日报》对洪旱灾害的报道与这些灾害事件出现的时间具有较好的契合性。

## 4 讨论与结论

本文提出了一个基于科学知识图谱定量研究水文化变迁的方法。它通过提供一种可视化的视角和方法快速、定量地研究水文化的发展阶段和趋势。通过对图谱的解读,发现水文化的演化与当时的政策、水利发展阶段和水文极端事件的发生具有一致

性,因此,可以利用科学知识图谱的方法来定量研究水文化。同时也表明《人民日报》是宣传党和国家有关水资源利用方针和政策的主要工具,凸显了其“环境监视”的职能,在一定程度上反映了其对水文化的导向作用。本文定量分析的结果为下一步构建政策、生态和水文化的协同演化模型奠定数据基础,为发展可持续的水资源管理科学与政策体系提供理论依据。

但是,科学知识图谱方法是以静态的图谱形式揭示隐含在基础知识中的动态结构信息,其研究对象是没有客观空间结构关系的抽象信息,因此,其可视化结果的可信度不仅取决于数据样本的完整性、绘制技术的完备性,还取决于领域专家对图谱解读的深度。本研究只是对科学知识图谱在水文化领域应用的一种探索,还存在着一些问题值得进一步探讨:第一,本文的数据来源是《人民日报》,《人民日报》是否能完全代表中国的水文化?新闻对制度的结构性偏向以及政治人物观点的主导地位都是公认的<sup>[34~36]</sup>。将来我们将研究更多的主流报纸。第二,科学知识图谱都是基于结构化的数据,报纸的内容是非结构化的数据,将其转化为结构化数据的方法是采用分词,然后提取关键词的方式。一方面,由于汉字语言的特殊性,分词的准确性有待提高;另一方面,基于词频的TF-IDF算法无法体现词在句中的位置信息和语义关系(如主语、谓语、宾语等),判断词的重要性的能力还有待提升。第三,关于科学知识图谱的可视化软件很多,而且都主要是针对英文文献,VOSviewer 是否是最适合开展类似研究的工具?第四,图谱的解读是绘制知识图谱的最终目的,目前对知识图谱的解读主要依靠领域专家对知识领域了解的广度和深度,如何保证解读的准确性?这些问题都需要通过进一步的研究和探索来解决,以提升和完善定量研究水文化的方法。

## 参考文献(References):

- [1] Ehrlich P R. Culture evolution and the human predicament[J]. *Trade in Ecology and Evolution*, 2010, 24: 409-412.
- [2] Tàbara D, Ilhan A. Culture as a trigger for sustainability transition in the water domain: The case of the Spanish water policy and the Ebro River Basin[J]. *Regional Environmental Change*, 2008, 8: 59-71.
- [3] Pahl-Wostl C, Tàbara D, Bouwen R, et al. The importance of social learning and culture for sustainable water management[J]. *Ecological Economics*, 2008, 64(3): 484-495.
- [4] Norgaard R B. Development Betrayed: The End of Progress and A Co-Evolutionary Revisioning of the Future[M]. London: Rout-

- ledge Press, 1994.
- [5] Wei Tianjiao. Development of China's water culture concepts viewed from water management[J]. *Journal of China University of Mining & Technology (Social Science)*, 2011, 13(2): 130-134. [尉天骄. 从水管理看中华水文化理念的发展[J]. 中国矿业大学学报:社会科学版, 2011, 13(2): 130-134.]
- [6] Zheng Hao. Historical changes of water culture in Tarim River Basin[J]. *Water Resources Development Research*, 2006, (4): 54-59. [郑浩. 塔里木河流域水文化历史变迁[J]. 水利发展研究, 2006, (4): 54-59.]
- [7] Allon F, Sofoulis Z. Everyday water: Cultures in transition[J]. *Australian Geographer*, 2006, 37(1): 45-55.
- [8] Head L, Muir P. Changing cultures of water in eastern Australian backyard gardens [J]. *Social & Cultural Geography*, 2007, 8(6): 889-905.
- [9] Yip R C P, Poon C S. Cultural shift towards sustainability in the construction industry of Hong Kong[J]. *Journal of Environmental Management*, 2009, 90: 3 616-3 628.
- [10] Harnsworth G R, Young R G, Walker D, et al. Linkages between cultural and scientific indicators of river and stream health [J]. *New Zealand Journal of Marine and Freshwater Research*, 2011, 45(3): 423-436.
- [11] Levinsen K, Wien C. Changing media representation of youth in the news—A content analysis of Danish newspapers 1953-2003 [J]. *Journal of Youth Studies*, 2011, 14(7): 837-851.
- [12] Roznowski J L. A content analysis of mass media stories surrounding the consumer privacy issue 1990-2001[J]. *Journal of Interactive Marketing*, 2003, 17(2): 52-69.
- [13] Noyons E C M, Moed H F, Van Raan A F J. Integrating research performance analysis and science mapping [J]. *Scientometrics*, 1999, 46(3): 591-604.
- [14] Shiffrin R M, Börner K. Mapping knowledge domains [J]. *PNAS*, 2004, 101(Suppl. 1): 5 183-5 185.
- [15] Liu Zeyuan, Chen Yue, Hou Haiyan, et al. The Method and Application of Mapping Knowledge Domain[M]. Beijing: People Publishing House, 2008: 5, 37. [刘则渊, 陈悦, 侯海燕, 等. 科学知识图谱:方法与应用[M]. 北京:人民出版社, 2008: 5, 37.]
- [16] Cobo M J, López-Herrera A G, Herrera-Viedma E, et al. Science mapping software tools: Review, analysis, and cooperative study among tools[J]. *Journal of the American Society for Information Science and Technology*, 2011, 62(7):1 382-1 402.
- [17] Riffe D, Aust C F, Lacy S. The effectiveness of random consecutive day and constructed week sampling in newspaper content analysis[J]. *Journalism Quarterly*, 1993, 70: 133-139.
- [18] Hester B, Dougall E. The efficiency of constructed week sampling for content analysis of online news[J]. *Journalism & Mass Communication Quarterly*, 2007, 84(4): 811-824.
- [19] Song Y Y, Chang T K. Selecting daily newspapers for content analysis in China: A comparison of sampling methods and sample sizes[J]. *Journalism Studies*, 2012, 13(3): 356-369.
- [20] Qiu Junping. Informetrics (V) fifth lecture word frequency distribution of literature information—Zipf's Law [J]. *Information Studies: Theory & Application*, 2000, 23(5): 396-400. [邱均平. 信息计量学(五)第五讲 文献信息词频分布规律——齐普夫定律[J]. 情报理论与实践, 2000, 23(5): 396-400.]
- [21] Wang Yuefen, Song Shuang, Miao Lu. Application study of co-occurrence analysis in knowledge service[J]. *New Technology of Library and Information Service*, 2006, (4): 29-34. [王曰芬, 宋爽, 苗露. 共现分析在知识服务中的应用研究[J]. 现代图书情报技术, 2006, (4): 29-34.]
- [22] Qiu Junping, Zhou Chunlei, Yang Siluo. Review and reflection of the information science research in china in the past 30 years (III): An analysis of the development stages and the trends of the information science [J]. *Library & Information Studies*, 2009, 2(2):1-9. [邱均平, 周春雷, 杨思洛. 改革开放30年来我国情报学研究的回顾与展望(下)[J]. 图书情报研究, 2009, 2(2): 1-9.]
- [23] Van Eck N J, Waltman L. How to normalize cooccurrence data? An analysis of some well-known similarity measures[J]. *Journal of the American Society for Information Science and Technology*, 2009, 60(8):1 635-1 651.
- [24] Van Eck N J, Waltman L. VOSviewer: A Computer Program for Bibliometric Mapping[R/OL]. ERIM Report Series Reference No. ERS-2009-005-LIS. 2009, Available at SSRN. [2013-09-15]. <http://ssrn.com/abstract=1346848>.
- [25] Van Eck N J, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping [J]. *Scientometrics*, 2010, 84: 523-538.
- [26] Van Eck N J, Waltman L, Noyons E C M, et al. Automatic term identification for bibliometric mapping[J]. *Scientometrics*, 2010, 82(3): 581-596.
- [27] Bengston D N, Fan D P, Celarier D N. A new approach to monitoring the social environment for natural resource management and policy: The case of US national forest benefits and values[J]. *Journal of Environmental Management*, 1999, 56(3): 181-193.
- [28] Neuendorf K A. The Content Analysis Guidebook [M]. Thousand Oaks, CA: Sage Publications, 2002: 205.
- [29] Wanta W, Golan G, Lee C. Agenda setting and international news: Media influence on public perceptions of foreign nations [J]. *Journalism & Mass Communication Quarterly*, 2004, 81(2): 364-377.
- [30] Zhu J H, Blood D. Media agenda setting theory [M] // Kovacic B, ed. *Emerging Theories of Human Communication*. New York: State University of New York Press, 1997: 88-114.
- [31] Trumbo C. Longitudinal modeling of public issues: An application of the agenda-setting process to the issue of global warming [J]. *Journalism Mass Communication Monographs*, 1995, (152):1-57.
- [32] Wang Shucheng. The five orders of society and economy development which water resource must satisfy [J]. *Guangxi Water Resources & Hydropower Engineering*, 2000, (3): 1-3. [汪恕诚. 水利满足社会与经济五个层次[J]. 广西水利水电, 2000, (3): 1-3.]

- [33] He Chansheng. Watershed science and water resources management[J]. *Advances in Earth Science*, 2012, 27(7): 705-711, doi:10.11867/j.issn.1001-8166.2012.07.0705. [贺缠生. 流域科学与水资源管理[J]. 地球科学进展, 2012, 27(7): 705-711, doi:10.11867/j.issn.1001-8166.2012.07.0705.]
- [34] Bennett W L. An introduction to journalism norms and representations of politics[J]. *Political Communication*, 1996, 13: 373-384.
- [35] Cook T E. *Governing with the News: The News Media as A Political Institution* (2nd Edition. Chicago) [M]. Chicago: University of Chicago Press, 2005.
- [36] Lawrence R G. Game-framing the issues: Tracking the strategy frame in public policy news [J]. *Political Communication*, 2000, 17: 93-114.

## On A Quantitative Method for Water Culture Evolution Research: Mapping Knowledge Domains

Xiong Yonglan<sup>1,2</sup>, Zhang Zhiqiang<sup>1</sup>, Wei Yongping<sup>3</sup>, Liu Zhihui<sup>4</sup>, Cheng Guodong<sup>5</sup>

(1. *The Scientific Information Center for Resources and Environment, Chinese Academy of Sciences, Lanzhou 730000, China*; 2. *College of Geography and Environment Science, Northwest Normal University, Lanzhou 730070, China*; 3. *Australia-China Centre on Water Resources Research, Department of Infrastructure Engineering, University of Melbourne, Melbourne 3010, Australia*; 4. *Institute of Scientific and Technical Information of China, Beijing 100038, China*; 5. *Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China*)

**Abstract:** Cultural transformation (including knowledge innovation, ideological emancipation and slogan publicity, etc.) is one of the most important drivers of changes in water management policies. Earlier studies into water culture were mainly based on qualitative methods. However, using only qualitative methods, it is difficult to reveal the stage characteristics of water culture evolution objectively and accurately. Mapping knowledge domains, a rising knowledge discovery method, can holistically reveal the sources of knowledge and the laws of its development. This paper is devoted to exploiting its potential as a quantitative method for researches targeting at water culture evolution. By extracting the key words related to water from People's Daily, knowledge domains on water culture evolution in China during 1946-2012 are mapped and interpreted, based on which the water culture evolution is unveiled. By comparing the water culture evolution with changes in water policies, stages of water resource development and extreme hydrological events, it is found that there exist clear relationships. This convinces us that mapping knowledge domains could play an effective role in quantitative research on water culture evolution. Further steps towards better quantitative researches on water culture evolution using this method include the enhancement of the universality of newspapers, improvement on the accuracy of key words extraction, and increase in the precision of visualization analysis.

**Key words:** Water resource management; Water culture; People's Daily; Mapping knowledge domains; VOSviewer.