

·信息技术与系统·

# 互文性与信息检索\*

胡立耘

(云南大学公共管理学院 云南昆明 650091)

**摘要:**互文性是人类知识组织的一种重要形式,在某种程度上,信息检索即是揭示、发现信息资源的互文性的过程。回顾图书情报学(LIS)的系统中心——用户中心——阐释学的范式转换历程,互文性在信息检索领域的作用越来越突出。信息检索系统可从营造可视化互文环境、关注检索情境、强调与用户的互动等方面进一步优化。

**关键词:**互文性 信息检索 情境 用户

中图分类号:G254.9

文献标识码:A

文章编号:1003-6938(2011)06-0099-04

## On Intertextuality and Information Retrieval

**Abstract** Intertextuality is an important style of knowledge organization. Information retrieval is a process of displaying and discovering intertextuality. From the history of paradigms in library and information science which transferring from the systems-centered and user-centered to hermeneutical, we can find intertextuality get more attention. The information systems should be improved through building visual intertextual environment, focusing retrieval context, and encouraging the mutual activities between user and systems.

**Keywords** intertextuality; information retrieval; context; user

### 1 互文性的界定

#### 1.1 互文性的定义

互文性(intertextuality)最早由法国文学理论家朱丽娅·克里斯蒂娃提出,她认为互文性是“一篇文本中交叉出现的其他文本的表述”,“已有和现有表述的易位”<sup>[1]</sup>;任何文本均是由引语的马赛克拼构而成的,都是对其他文本的吸收和转换<sup>[2]</sup>。叙事学的代表学者吉拉尔·热奈特<sup>[3]</sup>将互文性定义为“一篇文本在另一篇文本中切实的出现”,符号学家罗兰·巴特<sup>[4]</sup>提出“一切文本都是过去的引文的新织品”,著名的作家博尔赫斯<sup>[5]</sup>甚至指出,“语言本身就是系统的引语”,等等,也都体现了类似的思想。可见,互文性主要是指文本与文本之间所存在的互相指涉、互相参照的特点。狭义的互文性主要针对文学领域的文本,而从广义上理解,互文性普遍存在于人类各知识领域之中。本文将从广义的角度探讨互文性。

#### 1.2 文献的互文性

互文性是人类知识组织的一种重要形式,在依

托口耳相传以传承文化的社会中,史诗、神话、民间故事、寓言等均充满了互文性。人们在分享故事的同时,也参与创造,因而作品不断地产生新的变体,不断地丰富其内涵。一个故事的主题是另一个故事的出发点,一个母题(Motif)在无数的故事中以不同的形象重复出现。在书面文献占主导的现代社会,互文性更为普遍地存在于文化传播、学术传承中。出于不同的目的,文献与文献因引证、参照、反衬、驳斥、模仿、继承等形成的互文性比比皆是。莱布尼茨设想的世界百科知识体系、威尔斯幻想的世界大脑(World brain)、布鲁克斯论述的知识地图、超文本系统发明者尼尔森倡导的文献宇宙(Docuverse),以及拉封丹和奥特勒关于 Mundaneum、布什关于 Memex 的宏伟构想与伟大实践等,在某种程度上都是在通过一定的方式揭示文献之间、知识之间的关联,并努力通过这种关联来获取相应信息。加菲尔德创制的引文索引开启了从文献引证关联的角度进行信息检索的大门,普赖斯从文献的联系规律中展开了科学史的计量研究。可以说,互文性是文献系统的本质。但是,互

基金项目:本文系国家社科基金资助项目“民族地区口承文献的保护与利用研究”(项目批准号:10BTQ027)研究成果之一。

收稿日期:2011-10-11;责任编辑:刘全根

文性并不是一个纯然向以往文献的回溯指引的延宕过程，而是一个基于已有知识系统的求异求变的继承与创新相结合的过程。

### 1.3 互文性的类型

从表现形态上看，互文性包括显性互文性和隐性互文性，显性互文性是通过文献中的引文、注释、参考文献等，指引到其他相关文献，隐性互文性则是通过意象、概念、母题、典故、风格、知识单元等，暗示与该文献有联系的其他文献。显性互文性的极端例子是抄袭，隐性互文性的极致是仿写。

从层次上看，有书目层次的互文性和内容层次的互文性。书目层次的互文性主要揭示文献和文献之间具有的某种关联，这种关联除了是由文献本身或隐或显地指引到其他文献，还应包括因共同的特征（外在特征或内在特征）或其他因素而形成的，并被检索系统或用户显示出来的文献之间的联系。内容层次的互文性则是在文献内或文献间局部内容的相互映照。

从文献形式上看，同一内容的不同媒体之间，同一表达的多种版本之间也存在互文性。

### 1.4 数字互文性

超文本技术等数字技术使数字资源的互文性更为普遍化。Everett<sup>[6]</sup>于2003年提出了数字互文性（digitetxtuality）的概念，认为，新媒体是一个元意义系统，人们可从不同的意义系统中吸收并翻译、转换信息，并加以无尽的重组。信息与通讯技术的发展带来了新的资源生产、传播（共享）、消费模式，解构了作者中心，用户可从不同的路径/方式获得并集中材料，并加以解释、补充，扩展相关资源，对资源进行多种线索和多种形态的组合，数字互文性成为互联网的一个显著特征。

## 2 从互文性的角度看信息检索

### 2.1 信息检索是揭示、发现信息资源的互文性的过程

经典的信息检索是用户需求表达与检索系统中具有相应的标识的文献相匹配的过程，这是一种结构性的检索，而在网络环境下，信息检索更具有功能性，即以丰富的文献关联环境来引导、提示、激发用户发现相关文献，这种方式实际上是对文献互文性的揭示与提供选择的过程。

### 2.2 信息检索是作者预设的互文性与检索系统显示的互文性、用户接受的互文性三者互动的过程

在信息检索中，存在三种不同视角的互文性：

作者预设的互文性：体现在引证关系（单向）、注释（单向）、修辞手段（引语、典故、借用、类比、套语、模拟、戏仿、反讽），等。

检索系统显示的互文性：通过共同特征（外在特征、内在特征、共词）、语义关联（分类、主题词表、本体、概念图）、引证关系（引用与被引用双向）及拓展分析（共引、同被引等）、基于相关反馈的检索模型等加以揭示。

用户接受的互文性：包括语义的（对文献中作者预设的互文性的心领神会，沿着作者设定框架或指引的方向联系到其它文献）、情境的（文献的学科领域、用户的知识背景、所在社会、文化情境等对文献是否关联及如何关联产生影响）、语用的（根据用户需求在检出的文献之间或在检出文献与已知文献之间建立关联）。

三者相互之间有所交叉（见图1）。检索系统可局部地体现作者的互文意图，如引证关系、注释和抄袭；同时，检索系统从语义关联、共同文献特征，或基于统计学意义，发掘隐含的文献互文性，因而可在作者意图之外，进一步揭示文献集合中的互文关系；用户在检索系统提供的互文关系中获得更广的文献视野，同时，在文献阅读中对作者的互文意图根据自己的理解能力、使用情境和实用目的进行认可、选择，并引发联想，进而发现新的与所需求的文献具有互文性的文献。一方面，三者重合得越多，则检索的相关性越好，另一方面，三者完全重合是不可能的，也是没有必要的，用户可能在作者预设的互文性激发下产生更广的互文联想，也常常因检索系统揭示的互文性而发现新大陆，而检索系统也可因用户的互文性反馈而扩展相关文献的广度。三者的交互使用户的信息检索变得更有创造性，并且使一个开放的检索系统更具联想能力。



图1 三种互文性关系图

### 2.3 信息检索的过程同时是多文本交叉阅读的过程

网络环境下,超文本模式的出现导致了链接式非线性阅读(碎读)习惯,全文数据库的比重加大和去中介化的自我服务使检索者的自主性增强。在海量信息面前,用户进行检索时常常是对检出文献进行浏览,选取关键内容,跳读其中的知识模块,检索和阅读的界面越来越模糊。所以,信息检索往往是由多途径查检(即以结构化的检索系统为依托,获得相关文献)、自由浏览(带有一定目的地扫描资料,以了解篇章结构并发现重点段落)、非线性阅读(深度进入文献篇章,实行对关键内容的甄别、取舍,并调动更进一步的检索意图)结合在一起,不仅需利用书目层次的互文性,而且深入到内容层次的互文性。

### 2.4 信息检索是由用户主导的动态过程

在检索中,用户通过发散性思维不断开拓检索思路,使其检索策略具有多维的检索路径和变幻的检索动机,将信息检索变成多线程的冒险游戏,用户的心智、想象力、知识储备乃至时代观念都是探险的行囊。正是用户在主导着检索意图的调整和检索重心的转移,决定着文献取舍的广度和深度,通过文献之间及文献内容之间的关联来选择、联想、接受所需信息。用户的这种跨文本发现与利用所需信息的方式,消解了文本的形式结构,终结了作者的话语权威,也撇开了检索中介的控制。

## 3 互文性概念引入信息检索的理论基础

### 3.1 系统中心——用户中心——阐释学的范式转换

早在1991年召开的CoLIS1上,来自各国的图书情报学者就提出了图书情报学研究的范式转换。例如,Miksa<sup>[7]</sup>将以往的图书情报学的范式归纳为机构论和信息传播论,并将信息传播论进一步分为系统论和认知论,认为,系统论的重点是信息本身,重点关注如何通过系统使信息可形式化表达和有效传递;认知论则肯定知识是在用户接受、认知和创造的过程中产生的,关注用户接受信息的过程,关注信息如何影响其行为决策。Ingwersen<sup>[8]</sup>指出,图书情报学在不同阶段具有不同的理论取向,如1970年代至1990年代的理性主义取向、1970年代初至1980中叶受到重视的认知观点取向,以及1990年代以来的阐释学取向。理性主义取向认为知识是客观的、可编码和可无损传送的,信息接受者即信息使用者;认知观点取向以N.J.Belkin的知识非常态论(Anomalous

states of knowledge)及B.Dervin的意义建构论(constructivist model of information)等为代表,强调人的认知对信息的决定作用,但忽略了人所存在的社会环境和历史背景;阐释学则既关注文献的具有个人特点的意义建构,也关注文献与社会的互动。自CoLIS1以来,阐释学研究范式不断得到深化。

从信息检索的角度来看,在以系统为中心的范式中,文献的相关性来自用户需求的特征与系统中文献特征的匹配,这种完美匹配(Perfect match)牺牲了用户的隐含需求和未觉醒的需求,以及在检索中可能被激发的新需求。该范式过分简化了信息的意义,将文献中的信息看作科学、客观、可转化成各种形式和可用文献特征作为替代(representation)的,预设了检索系统对文献特征抽取的完善性与准确性;并预设了检索系统用户是有一定的检索知识和检索技术,其需求能为系统设计者和情报机构服务者所预测、激发或引导的,所谓“中等程度的用户”。因而,忽略了用户真实需求与需求表达所存在的差距,忽视了不同的环境下中信息对用户所产生的不同影响,无法掌握用户真正的需求及信息对于用户的意义。在以用户为中心的范式中,引入了认知理论和建构主义思想,认为,用户需求是多层次的、动态的;用户的需求自觉程度、知识空缺状态、所处问题情境、特定典型环境等都可对信息检索与利用行为产生影响;知识并非纯然客观的,知识的解释、理解、建构与用户相关,用户的已有概念体系、知识状态、认知模式,以及心理因素、内在需求、主观判断等值得关注。然而,该范式所强调的并非个体而是群体,通常从用户的生活、职业、问题及解决问题的方式、所处机构,以及文化或社会因素等方面,将用户归入类型化群体加以研究,并且强调信息中介(图书馆员,信息检索人员)的权威,没有充分考虑社会影响的作用。在阐释学研究范式中,则十分关注用户个体的直接参与,无需中介介入,强调互动而非单向信息传递。不仅关注检索系统中信息与用户需求的相关性,而且强调信息与用户使用情境的关联性。

### 3.2 信息检索的情境研究

阐释学强调只有将对整体文本的体会与对单一词句的认识进行交互理解,才能获得真正的意义,同时,理解必须基于理解者个人的视野与历史视野的融合。伽达默尔把视野定义为“我们从一个特殊的有利角度把一切尽收眼底的视觉范围”,认为应理解作



品与个人心理的关系及历史环境中作品之间的含蓄关系<sup>[9]</sup>。也就是说,只有了解单个词与作品整体、作品与所属种类、作品与作者的心理状态等等关系,才能达成对作品的真正理解。这种对文献整体与部分、个体(用户)与文献的关联的研究,类似于信息检索理论中的“情境”研究。

对于“情境”,不同的学者从各自的研究出发,采用了不同的术语表达,如 context、setting、situation、environment<sup>[10]</sup>等。在 CoLIS2 中,S. Brier<sup>[11]</sup>提出网络符号学(Cybersemiotics),用阐释学的方法分析社会系统与知识范畴对于知识组织和信息检索的影响;J. Hansson<sup>[12]</sup>则采用阐释学来分析文本与事实、图书馆分类系统与社会话语(social discourse)的关系。2005年,有学者<sup>[13]</sup>就情境(Context)在信息表达、信息搜索、信息相关性、信息处理、文献分析等的方面的意义展开了全面讨论。在信息检索中,用户依托特定的情境来选择相关文献并建构意义,这些情境包括用户对该主题文献的已知程度、用户对目标文献的相互关系的理解、检索专题的知识范畴、用户检索时所处的问题状态、用户所属文化的特点、其它用户的影响等。

## 4 互文性概念引入信息检索系统的意义

### 4.1 全面营造检索系统的可视化互文环境

检索环境对检索动机具有推进与催化作用。检索系统营造出可视化互文环境,可使用户从检索系统的“暗箱”中检索(Retrieval)信息的同时,能在特定相关文献的范围中浏览(Browse),进而定位到知识单元进行阅读。有利于用户既有鸟瞰全貌的视角,发现相关文献(书目层次的互文性),又具有“鱼眼”功能,准确定位并强调文献的相关细部(内容层次的互文性)。在丰富的互文环境中,扩展用户的文献知见面,提升用户的信息视域(Information Horizon)<sup>[14]</sup>,增加信息偶遇(information encountering)的机会<sup>[15]</sup>,使用户在预期之外发现新的相关信息。

目前的检索系统中,正在逐步强调文献关联的显性化。例如,与目标文献同步显示参考文献、相关文献、施引文献,实现检索平台及跨库的文-文关联;各种引用和共引关系的可视化显示;以 Tag clouds 的方式聚集主题和社区;以图或表格将检出文献从机构、专题、刊名等角度聚类显示;在 FRBR 的理念下实现 OPAC 中同一作品的不同表现形式、不同版本、

乃至不同单件的关联;可悬窗式激活相关词条的在线知识词典等。有可视化工具还具有对文献的互文关联程度的评价作用。

信息检索系统应进一步加强文献互文性的可视化,例如,将 web of Science、Scopus 等引文分析可视化的相关举措整合到全文数据库中、将 OPAC 的基于 FRBR 理念的做法拓展到多种媒体的全文检索平台中,将学科知识图谱、主题图普遍运用到检索系统中,加强内容层次的互文性的揭示等,从而实现文献互文性的多形式、多维度、多层次立体显示。

### 4.2 在检索系统中充分考虑检索情境

上文指出,系统中的文献不仅具有基于共同特征或引证关系的各种客观的相互联系,而且具有与情境相关的潜在的可联系性。在检索系统中关注检索情境,如,知识范畴、文化背景、术语体系、学者共同体等对于信息检索的作用,可使检索系统在营造丰富的互文环境的同时,有效避免一切皆互文的泛化。分类显示、专题服务等在某种程度上可突出在特定学科、产业、专题等范畴中文献的关联。领域本体、应用本体可更为明晰地显示特定情境中人、事、时间、概念、地点的关联。

在一些文化敏感的数据库(Culturally Sensitive Database)建设中,更应考虑文献在特定情境下的关联,例如在开发各民族的本土知识数据库(Indigenous Knowledge Databases)时,数据库常常包括大量的从口头表达收集而来形成各种形式的文献资源(书面、视频、音频文献),文献中技能与传说、传统与想象、风俗与信仰杂糅,并且包含与科学分类系统不同的动植物分类知识、与学术命名不同的名称表达,一些文献具有特定场所、时间及群体适用性限制。因此,应从其文化的独特视角对知识进行关联,以符合民族所惯用的方式进行信息组织。

### 4.3 强调检索系统互动性,促进用户主体性的发挥

在参考咨询理论中有著名的三 M:材料(Material),方法(Method),思维(Mind),这三者也是信息检索所必不可少的。在网络环境下的富营养信息海洋中,检索系统如何理解、激发、显现和利用用户的发散性思维尚未引起足够的重视。是用户的选择与接受将文献具有情境依赖的悬而未决的互文性落实,因此,将用户接受的互文性纳入检索系统设计中应是检索系统优化的一个方向。

检索系统应提供用户与检索系统的深度交互,

由用户主导检索过程,将用户接受的互文性作为反馈,动态更新并以可视化的方式显示新的文献关联。除了在线显示用户推荐相关文章这种直接有效的方式外,还可显示用户检索路线图,便于用户回访。同时,挖掘用户的检索过程,构成概念链路、知识链路和文献链路,并进行统计归纳,显示基于情境的文献内及文献间的某一篇章、语词、概念的相互关联,从而加强检索系统互文性揭示的深度与广度。以少数民族本土数据库的设计为例,可依托数据库所具有的非线性、可组合的开放结构,由用户从一系列的文献或多种资源形式中选择所需的特定文献或知识片段,并根据自己的理解框架重组它们之间的联系,进一步丰富检索系统中文献资源的互文性。再如,挖掘特定专业领域学者在检索中的文献取舍行为,以了解用户的文献利用行为是如何为学术范式、信息保障环境所影响,文献的关联又如何突显了特定领域的学术特点和思考方式,为专门领域知识服务模式的设计提供新的思路。此外,还可通过虚拟社区,促进用户与用户、用户与作者的交互,从用户信息资源共享的角度来拓展文献的互文性。

#### 参考文献:

- [1][3] (法)蒂费纳·萨莫瓦约.邵炜译.互文性研究[M].天津:天津人民出版社,2003:3,19.
- [2]Kristeva J. dialogue and novel [A].Moi T. The Kristeva Reader[M].New York: Columbia university press,1986:37.
- [4] (法)罗兰·巴特.屠友祥译.文之悦[M].上海:上海人民出版社,2002:96.
- [5] (阿根廷)豪·路·博尔赫斯.王永年,陈泉译.博尔赫斯全集·小说卷[M].杭州:浙江文艺出版社,1999:442.
- [6] Everett, A. Digitextuality: Theses on Convergence Media and Digital Reproduction [A]. Everett A, Caldwell J. New Media: Theories and Practices of Digitextuality[M]. London: Routledge, 2003: 3-28.
- [7] Miska F L. Library and information science: two paradigms [A]. Vakkari P, Cronin B. Conceptions of Library and information science [M]. London and Los Angeles: Graham Taylor, 1992: 229-252.
- [8] Ingwersen, P. Conception of information science [C]. Vakkari P, Cronin B. Conceptions of Library and information science. London and Los Angeles: Graham Taylor, 1992: 299-312.
- [9] (联邦德国)H·R·姚斯, (美)R·C·霍拉勃.周宁,金元浦译.接受美学与接受理论[M].沈阳:辽宁人民出版社,1987:323.
- [10] Chang S L, Lee Y. Conceptualizing context and its relationship to the information behavior in dissertation research process[J]. 图书馆学与资讯科学, 2000, 26(2): 5.
- [11] Brier S. Cybersemiotics: A new Paradigm in analyzing the Problems of Knowledge Organization and Document Retrieval [A]. Ingwersen P, Pors N O. CoLIS 2: Second International Conference on Conceptions of Library and Information Science; Integration in Perspective [M]. CopenhagenDenmark: Royal School of Library and Information Science. 1996: 23-43.
- [12] Hansson J. In My Mind's Eye - In Search of the Mimetic Relation Between a Library Classification System and its Social Discourses[A]. Ingwersen P, Pors N O. CoLIS 2: Second International Conference on Conceptions of Library and Information Science; Integration in Perspective [M]. Copenhagen, Denmark: Royal School of Library and Information Science, 1996: 99.
- [13] Crestani F, Tuthven I. Context: Nature, Impact, and Role—5th International [A]. Conference on conceptions of Library and Information Science [C]. Berlin Heidelberg: Springer, 2005: 1-253.
- [14] Sonnenwald D H. Information Horizons [A]. Fisher K E, Erdelez S, McKechnie L. Theories of information behavior, Medford [M]. Newjersey: Information today, Inc, 2005: 192-197.
- [15] Erdelez S. Information encountering: A conceptual framework for accidental information discovery [A]. Vakkari P, Savolainen R, Dervin B. Information seeking in context—Proceedings of an international conference on research in information needs, seeking, and use in different contexts [M]. Tampere, Finland; Los Angeles, CA: Taylor Graham. 1996: 412-421.

作者简介: 胡立耘(1966-),女,云南大学公共管理学院教授,硕士生导师。