

●陆 伟 万维雅

基于认知观点的信息检索交互模型*

摘 要 基于情报学认知观点的信息检索交互研究是信息检索研究的一个重要领域。信息检索模型相当复杂,各类生产者、中介和用户的认知结构发生交互,影响着信息检索过程。图3。参考文献5。

关键词 信息检索 情报学 认知观点 知识结构

分类号 G354

ABSTRACT The study of interactivity of information retrieval based on cognitive viewpoint is an important field in the study of information retrieval. Information retrieval model is very complicated, and the cognitive structures of various producers, agents and users interact with each other, affecting the procedures of information retrieval. 3 figs. 5 refs.

KEY WORDS Information retrieval. Information science. Cognitive viewpoint. Knowledge structure.

CLASS NUMBER G354

1 情报学认知观点与信息检索交互

1.1 认知学与情报学

认知科学是研究人类感知和思维信息处理过程的科学,包括从感觉的输入到复杂问题求解,从人类个体到人类社会的智能活动,以及人类智能和机器智能的性质等。认知科学的学科域十分宽泛,甚至有人说它是围绕人类的认知现象形成的一个理论与应用研究学科群,其观点、方法、研究途径等多种多样。当前认知科学以信息加工观点研究认知过程为主流。

情报学与认知科学在研究对象和研究目的方面是相似的,研究领域存在许多重合之处。认知科学强调人的知识和知识结构对其内部心理活动和外部表现的决定性作用;情报学重视人们在信息利用过程中的知识结构和认知能力,强调信息系统和用户双方在认知层而非物理层上交互的重要性。不同的是认知科学的研究焦点是作为信息处理现象的人类认知过程本身;而情报学则基于认知科学的有关理论和方法,侧重于对文献信息处理、组织和利用的探讨。也就是说,情报学研究关于记录了的知识的相关个体现象,认知科学可以具体说明研究的相关现象可能是什么。情报学的目标是为用户提供者和使用者之间所需求信息的有效交流提供便利,信息检索和利用首先或者说从根本上是一项认知行为,

其过程可以以认知的思想观点来处理。也可以说情报学是处理人类认知和认知过程方面的一门学科。

1.2 情报学认知观点

随着认识和研究的深入,一种能处理人、信息和技术关系的情报学一般理论基础将在不远的将来出现,未来情报学的研究领域之一将是认知科学在信息加工、处理和利用中的应用。以往情报学中关于读者或用户研究、交流研究的一些著作,曾在不同程度或潜在地反映、应用了认知观点和方法。进入70年代以后,以认知科学的发展为背景,B. C. Brookes、N. J. Belkin、P. Ingwersen等人明确倡导情报学研究的认知观点。吸取认知科学的基本思想和有关理论、研究信息处理原理、关注对信息的利用和吸收,以创造出支持、改善这种利用、吸收的情报系统和情报服务为目的的研究,代表了情报学领域另一种研究取向——情报学认知观点。

在情报学尤其是信息检索中考察认知观点,有以下意义:一是它能对交互式信息检索过程的本质给予更深刻的洞察,在头脑中有这样一个观点,就能解释信息检索的现在状态和将来研发工作中的优势与局限;二是认知科学为理解个体知识交流和获取的重要方面提供基础,认知观点能详细阐述关于信息检索交互的研究细节。因此,情报学认知观点可以被视做未来情报学研究的一个强大框架,促使信息

* 本文为国家社科基金项目成果之一(项目编号03BTQ021)。

检索领域的重新定义、精炼和改进,以便为信息检索的各种问题提出足够的解决方案。

当今各国学者在情报学认知观点的基本表述上达成共识:任何形式的信息处理,无论是观察对象时知觉的信息处理,还是阅读文本时符号的信息处理,信息处理的主体都以一个范畴或概念体系作为媒介,这一体系也称为世界模型。

所谓世界模型,也称作映象,由知识结构或认知结构组成。这些知识结构或认知结构是由个体和他的社会集体的经历、教育等等决定的。

1.3 信息检索交互

信息检索是情报学的核心领域。对信息检索的研究将是基于认知观点的,兼顾个人和群体的信息需求,研究的一个新领域将是人机对话(HCI)即交互性。关于信息检索,认知观点强调的是交互作用,基于认知的信息检索应该是在交互中实现的。机器不能够感知含义、概念、思考、认知、创造等等,除非在人类的告知或支持下。电脑的低层次处理(尤其是对于书籍和标签)和人类的高层次语义操作之间的定性区别构成了人机交互的核心问题。

信息检索研究方法经历了传统的、用户导向的和认知的3个阶段,人们正逐渐把注意力从集中在信息的实体和符号上转为理解主题或者接收者上。当信息检索研究积累了足够的实验和分析证据,能够支持更具适应性的认知研究时,信息检索研究开始转向认知信息检索,即信息检索交互。信息检索研究的重点开始变为考虑如何把这些研究成果融入处理交互式信息检索系统的各种假说和模型中去,并能对系统进行测试。

2 信息检索交互模型

2.1 基本认知模型

为了更好地理解信息检索交互中的认知观点,我们可以看图1的基本认知模型。

图1中,“信息检索系统设置”由作为工具手段的结构组成,如信息检索技术和标引规则,代表着设计者对于如何处理信息检索系统中对象的构想。“信息检索系统对象”包括如经标引(描述结构)的文本或图片,以及对文本或图片自身内容的解释。这些描述通过运用“信息检索系统设置”中的相关规则和算法而产生。其他的知识结构被纳入“中介(机制)”,例如训练或指导实施等。

在图的右侧,个体用户在策划信息检索时,有一

定的知识结构、目标和对信息的渴求。在信息检索交互作用这点上,用户的范畴或概念体系(世界模

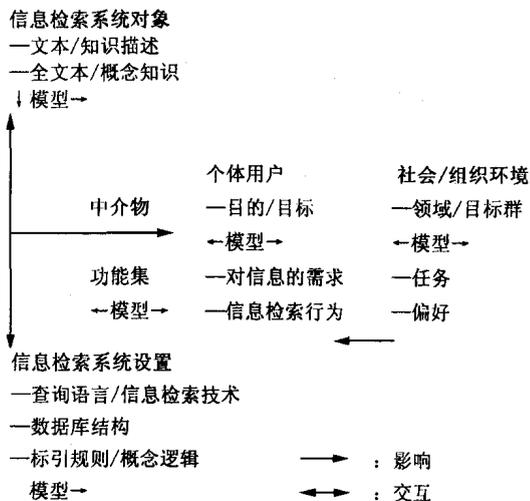


图1 基本认知模型

型)不断地受到他所处社会环境的影响,社会环境包括社会习俗、偏好和依附于地域的集体认知结构等。特定的世界模型引导各个变量的期望值,这是形成交互作用的每个变量所固有的。比如,一种特定的检索语言结构设计是以满足检索者中的特定群体(包括一种特定群体的模式)为目的的。又比如一个用户由于过去的经历,不愿意接近一个特定的服务或人员,过去的经历改变了用户服务或人员的模式。

对于情报学,尤其是信息检索研究,重要的一点是知识结构的个体差异多样性。因此,认知观点建议人们研究位于交互作用变量的表面结构之下的个体世界模型和知识结构的多样性,信息检索的任务是使著者、信息检索系统设计者和索引编制人员的认知结构与信息工作者(中介)和最终用户的认知结构相一致,从而满足当前的信息需求。

2.2 信息检索交互认知模型

最早的信息检索认知模型是建立在图1的基础之上的,并且加入了各种中介知识结构。图2给出的是信息检索交互的认知交流系统的基本模型。

2.2.1 信息空间的主要知识结构

在图2左侧,每个潜在信息的生产者,不管是著者,还是信息检索系统设计者和索引编制人员,都表征着认知模型的一部分。

著者会根据自己的意图、理念和对预期受众的了解,通过转换“工作空间”和“知识状态”(概念知识),以文本或图片等方式生产潜在信息。同样,信

息检索系统设计者会根据所使用的用户模型建立数据库。索引编制人员会根据所使用的用户模型和他人制定的某些索引规则,以某种表示方法为生产出的数据集合编制索引。各种转换最终聚合成检索

者,即信息工作者(中介)或最终用户,能够访问的潜在信息集——信息空间。这种访问既可以针对全文文献进行,如从图书馆书架上或数据库中查找全文文献,也可以在表征潜在信息的各种索引中进行。

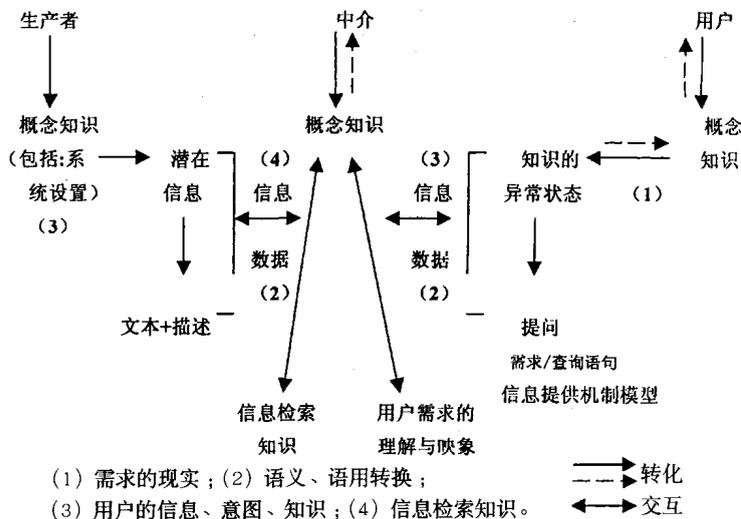


图 2 信息检索交互认知模型

总的来说,主要是如图 3 所示的这些信息生产者参与图 2 左半部分的信息检索交互活动,他们都有各自特定的认知模型、工作空间和知识状态。

生产者(Generator):	负责(Responsible for):
著者(总体而言)	潜在概念信息
索引系统设计者	表示规则
信息检索技术设计人员	信息检索技术或指令语言
索引编制人员和维护人员	信息实际表示和数据更新
系统/数据库设计人员	数据库结构和功能
系统生产者	检索者模型、选择策略

图 3 信息生产者参与的信息检索交互活动

某些情况下,如较小的内部信息检索系统中,除了著者,所有的生产者都可以是同一个人或同一个团队。信息检索实验系统的索引系统设计者、信息检索技术设计人员和索引编制人员往往会是一个人或一个团队,例如在部分匹配技术实验中就常常如此。图 2 所示的模型从认知观点出发,解释了这类实验在限定的实验框架内取得相当不错结果的原因:实验没有纳入最终用户和中介知识结构,并且很少使用不同的知识结构,这使得结构匹配变得很容易。

2.2.2 中介使用的主要知识结构

图 2 所示模型的中间部分是中介的认知模型,包括中介在信息检索活动中的工作空间和实际的知识

状态。该模型主要有两种知识结构:

(1)信息检索知识。它包括:关于信息检索系统设置的知识,即各种信息检索系统和信息源、信息检索技术、软件功能、数据库结构、表示方法和规则、实际应用的索引编制方法等;关于信息检索过程的知识,即信息检索方案(包括系统选择)、信息检索策略、逻辑安排和系统提问等。

(2)概念知识。它包括:领域和领域任务的知识,包括主题与情感领域、概念与概念关系、范例结构、用户查找行为与偏好等;关于意图、期望、想象、评价、计划等方面的知识。

在信息检索期间,“信息检索知识结构”与“概念知识结构”发生动态交互,主要目的是使之适应用户的要求。此外,“概念性知识”与信息检索系统存储的文献概念内容(即文本和索引款目)有关。中介人员试图通过这种生产者、用户和中介三方交互的方式,在用户的概念知识、中介自身的概念性知识、“信息空间”之间建立起联系。中介人员的期望、想象等同样在其信息检索行为中起重要作用。

作为三方交互的第三方,中介拥有构建当前用户提问模型(相当于图中的理解与用户需求)和用户模型(相当于用户描述,即用户偏好和行为知识)的能力。

对于人工中介,就用户的查找行为、偏好、期望和意图以及用户的信息检索和领域概念知识而言,

用户模型是一种动态的心理构成。在专门的领域中,新入行的检索中介人员在语义记忆里建立起用户模型需要花费一段时间,尤其是当他对相应学科领域不熟悉时;如果检索中介人员本身掌握有关学术行为的“元”知识通常会缩短这个过程,由于是自己熟悉的领域,在其语义记忆中会自动建立相应的领域用户模型,并形成相当详尽的概念结构图。相反,对于通才型的检索中介人员来说,比方说公共图书馆员,是决不可能形成既能覆盖极其庞大的“信息空间”又能穷尽用户概念的高度结构化的概念结构图的。因此,与专门领域的检索中介人员相比,通才型的检索中介人员的认知状况受到的限制要大得多。由于要考虑到每一个潜在用户,通才型人员采用和构建的用户模型一般更具普适性,他们常常利用自身的信息检索知识和信息检索系统来弥补自己的不足。检索人员,尤其是通才型的检索人员,必须经常通过访问信息检索获取新知识以适应新的情况。而非人工中介则不得不依靠信息检索系统设计者的构想和预期,也就是设计人员的认知结构来实现。在信息检索期间,人工中介的认知结构和工作空间会发生变化,而非人工中介的知识结构则是固定不变的,尽管它们肯定也包含适应性规则。

2.2.3 用户的知识结构

图2的右侧列出了用户检索时涉及到的部分知识结构。

“概念知识”表示用户的当前知识状态,它会转化为特定情况下的思想状态——“问题空间”,用户在这之中识别他对知识的缺乏,如果不能通过自身思考填满这个问题空间,用户的问题空间会转化为“不确定状态”。“问题空间”和“不确定状态”即是模型中的“知识的异常状态”。仅当产生不确定状态时,用户才会有外部信息需求。用户可能会根据其知识状态所固有的“信息提供机制模型”向中介或信息检索系统提问,提问可以明确地表达为描述词,或者在此基础上表达为更充分的信息需求或问题。如果模型中的中介是人工中介,提问就会被检索人员所感知,这样提问就会改变检索人员的知识状态。提问由此转化为信息,因为信息的定义是接收者知识状态的改变。如果是非人工的中介系统,提问语句将会是有待系统处理的数据,系统会根据实现的前提和知识结构处理提问。

中介的主要任务是在信息检索系统的信息空间和用户的问题空间之间进行匹配。假定信息检索系统能够为用户提供一些潜在信息,用户就会使用它,从而减少或消除用户的不确定状态,并把用户的问题空间和知识状态转化为信息。在信息检索交互期间,

这种匹配过程可能会反复出现。在某个时刻,用户的信息需求可能会得到满足,或者是其实际问题得到部分解决,或者说用户的情感体验基本得到满足。

3 结论

信息检索的研究方法到达认知阶段,意味着人们对信息检索的研究转向关注个体的认知、情感、动机等所有参与信息检索交互的变量都固有的心理过程,此外还涉及到检索活动周围的社会环境等。这样一来,信息检索模型就变得相当复杂,各类生产者、中介和用户的认知结构发生交互,影响着信息检索过程。无论用户掌握的主题知识是深还是浅,无论用户对中介功能或实现这些功能的信息检索系统的结构了解程度如何,信息检索系统设计都必须努力适应这些认知结构。基于认知观点的信息检索交互模型的提出与完善,为满足信息检索系统设计的这方面要求提供了基本的理论框架。

现在已经出现了一些基于认知观点的信息检索中介系统的设计模型和原型系统,如“智能”信息检索专家系统设计——IR,为用户提供专门知识支持的非推论性的信息检索设计——Bookhouse,基于知识的信息检索辅助设计——KIRA,辅助用户建模设计——Euromath,“智能”用户模型构造设计——IR-NL,支持检索过程的推理设计——OAKDEC等等。这些研究成果表明,当今世界各国研究者的研究不仅是简单地了解用户和中介人员的心理行为,还要了解中介功能和自动中介的任务,在基于认知观点的信息检索交互模型的基础上,通过构造更理想的用户模型、系统模型和中介模型,在信息提供者 and 信息需求者之间架起沟通的桥梁。

参考文献

- 1 祝朝安. 情报学理论问题. 情报理论与实践, 1995(2)
- 2 贺颖等. 认知观给情报科学带来的新思维. 情报杂志, 2003(8)
- 3 Peter Ingwersen. Information Retrieval Interaction. Taylor Graham Publishing, 1992
- 4 Jonathan Back, Charles Oppenheim. A model of cognitive load for IR: implications for user relevance feedback interaction. Information Research, 2001(1)
- 5 Raquel Navarro-Prieto, Mike Scaife, Yvonne Rogers. Cognitive Strategies in Web Searching. Human Factors & the web, 1999(6)

陆伟 武汉大学信息资源研究中心讲师, 博士。通信地址: 武汉市。邮编 430072。

万维雅 武汉大学信息管理学院硕士研究生。通信地址同上。(来稿时间: 2004-07-14)