

身心语言程式视阈下的信息检索用户认知模型研究

马海群 杨志和

摘 要 当今的科技信息发展环境中,信息检索用户的认知行为和需求心理极为复杂。如何将信息检索中用户的潜意识显现出来,以便更好地服务于检索系统建设,是信息检索研究中的核心难题之一。研究基于流行于欧美国家及港台地区多年的身心语言程式(NLP)理论,从身心、语言和程式三个核心维度出发,构建了信息检索用户的“需求认知、表达与交互模式”、“检索语言认知、表达与交互模式”和“检索目标制定中的心理取向模型”。同时结合 NLP 中的锚理论,分析信息检索用户的“心锚”和相关的“空间锚”,总结出相应的信息检索规律,构建出检索中锚交互模型、集合关系、双因子矩阵模型和表达式,并对相关的假设模型进行数据检验和优化。图 4。表 3。参考文献 15。

关键词 信息检索用户 身心语言程式 信息需求 认知模型

分类号 G251

ABSTRACT With the development of information technology, the user's cognitive behavior in information retrieval becomes more complex. In order to improve the construction of information retrieval system, one of the most important issues for information retrieval research is to find how to explore and show up users' unconsciousness during their retrieval. Based on of the Neuro - Linguistic Programming theory (NLP), this paper defines users' "Mode in Information Demands Recognition, Expression and Interaction", "Patterns in Information Retrieval Language Understanding, Expression and Interaction", "Psychological Orientation Model in Building Search Goal". Meanwhile, with the anchor theory of NLP, this article analyzes the users' "Mind Anchor" and the related "Space Anchor" and constructs then the anchor interaction model, sets, two - factor matrix model in information retrieval. Finally, this article tests and modifies the proposed models. 4 figs. 3 tabs. 15 refs.

KEY WORDS Information searchers. Neuro-Linguistic Programming. Information demands. Cognitive model.

CLASS NUMBER G251

信息需求者都想开辟和拓展个人独特的内心世界,他们时刻想着——怎样选择传入脑里的信息,怎样认知这些信息,怎样储存这些信息,怎样把这些信息与其他存储器中的信息融合,怎样运用它们,等等。在当前的信息环境中,虽然信息检索的内容、形式和环境复杂多变,检索用户的心理特性也千差万别,但是其中的意识形成、思维模式、经验和策略却不会有太大的改变。本文将在前人研究的基础上,结合身心语言程式的相关知识探寻信息检索用户的认知规律。

1 身心语言程式与信息检索

身心语言程式 (Neuro-Linguistic Programming, NLP) 发源于美国加州大学 Santa Cruz 分校,有两位创始人。其中一位叫做理查·班德勒 (Richard Bandler), 主修计算机,但他主要研究人类行为学,对心理学颇有研究;另一位是约翰·葛瑞德 (John Grinder), 语言学家。1975 年理查和约翰将先前的重要研究成果汇集出版了 *The Structure of Magic* (Vol. I and Vol. II), 1976 年他们决定用“Neuro-Linguistic Programming”来命名先前所发现和创造的学问——这标志着 NLP

的诞生^[1-2]。

NLP这一术语已风靡于欧美国家和港台地区多年,目前国内仅有部分商企及其研究者关注,最初被译为神经语言程式(或程序),现在大多译为身心语言程式。它主要被研究者们用于研究语言习得、语言掌握、言语生成、言语理解的神经机制,以及人脑接收、存储、加工和提取言语信息的机制;目前,在欧美及港澳已被广泛应用于语言神经生理机制和言语障碍的神经病理机制研究和治疗中^[3]。其在信息检索用户身上有充分的体现。

我们知道,检索用户是有信息需求的人,其需求是处身于特定信息空间中的个人对信息资源所产生的某种迫切心理,并且在这种迫切心理的驱动下会为获得预期的信息而谋定策略,是用户对需求的特征判定、语言表达,并借助可检系统完成其检索行为的过程。信息检索中,检索用户所处的特定信息空间(检索环境)和其中的思维过程就是NLP中身心的理论范畴;用户为描述其内心的真实图景所采用的语言表达即NLP中语言的理论范畴;而用户为了得到预期的检索结果而采取的一系列检索指令和策略就是NLP中程式的理论范畴。用户选择检索策略的过程可以被称为模式化(Modeling)过程,信息检索中的这种模式化是用户意识(特别是潜意识)发生作用的过程。检索实践中,我们不仅可以将致使这些意识(或潜意识)发生作用的策略重复使用,而且还可以传授给他人。本研究将从身心、语言和程式三个核心维度出发,对信息检索用户的需求认知与表达、检索语言认知与表达以及检索策略形成模式展开研究,并挖掘出其中的一系列规律,构建相应的模型,并且对相关的假设模型进行数据检验和优化。

2 身心理念下的用户需求认知、表达与交互模式

信息检索用户的信息需求是用户的一种心理能动活动。作为表现检索用户高级需求的一种形式,它是一个取决于各种主客观因素的动态量值^[4]。主观视角的检索用户信息需求是一

个非常复杂的概念,与个人特性、能力、文化程度、生活环境、从业领域等息息相关,涉及心理学、行为科学、社会科学等理论与方法,在此不予展开。

信息检索中,信息需求的判定是置身于特定检索环境中的检索用户的一种心理活动,这种信息需求心理的核心维度主要有:

(1)潜在信息需求。这是信息需求用户的原始心理状态,处于该状态中的信息检索用户还不能完全地确定自身的信息检索具体范围、真正意图以及途径。此时,用户的需求心理处于抑制状态,其主观能动性还没有被具体的激活源(或事件)所调动。

(2)感知信息需求。当信息检索用户对具体问题的认知有所深化时,用户的主观能动性就会被特定的问题所激活。这时,检索用户会为获取到与特定问题相关的信息而采取主动的认知行为。

(3)言表信息需求。处于这一层次的信息检索用户将用书面(或口头)语言的形式将自己的信息需求表达出来,并会付诸检索行为。

事实上,人的潜意识比意识具有更大的能量和影响力。然而,在具体的信息检索中,我们却常常不自觉地忽视了这个可以利用的资源。结合用户身心语言程式的相关理论,检索用户在信息检索中完全可以把潜意识变成意识,并且可以将这两个方面有机结合起来,这也符合马斯洛的人类需求层次理论(生理需求、安全需求、情感和归属需求、尊重需求、自我实现需求)——当人们的基层需求得到满足时,就会迈入到更高的需求层次中。从具体的信息检索用户的意识活动视角出发,宏观视角的用户信息检索需求认知、表达与交互模式如图1所示。

信息检索中,用户心中的潜在信息需求一般无法具体化,并且用户自身也无法确定其具体的内容形式。但是,在某一检索目标的驱动下,用户便会对内心深处的潜在信息需求进行感知,自然就会产生规范化和规模化的信息需求集。接着,用户会根据特定的语言逻辑和表达习惯将这些信息需求集用特定的语言以“自认为准确”的方式表达出来,并依据身边能利用

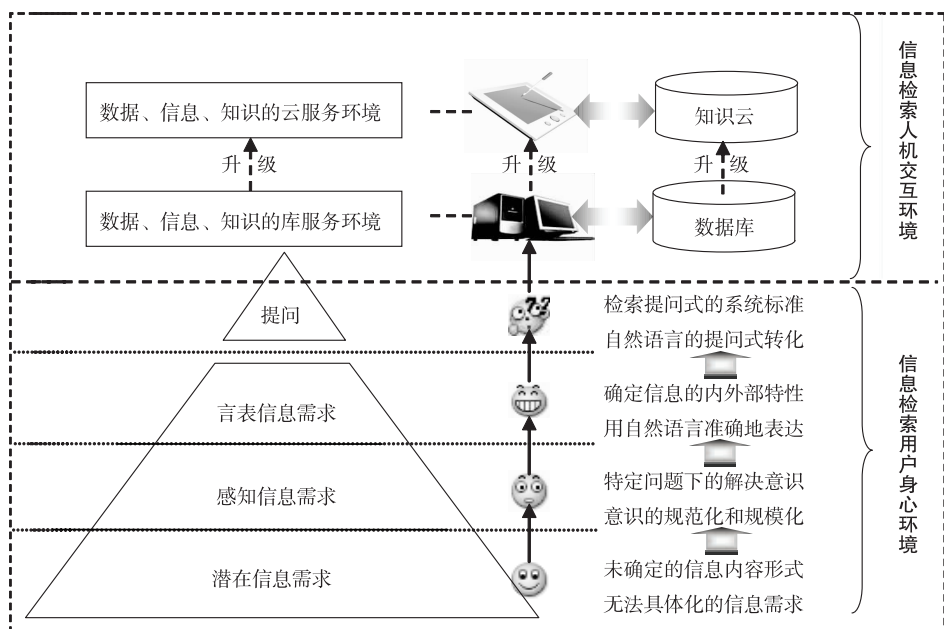


图1 用户检索需求的认知、表达与交互模式

的数据、信息、知识服务平台将其信息需求表达进行系统化、标准化和形式化,同时与这些数据、信息、知识服务系统进行交互。此过程中,用户最终会以提问的形式将源于内心的潜在信息需求集合表达出来,并与检索系统交互以满足其需求。支撑该过程的是用户需求认知与表达的身心环境,以及为满足其需求所依赖的人机交互环境。目前,这些能满足检索用户特定信息需求的数据统一存储于数据库中。但是,在“云计算”理念下,将来这些存储有用户特定信息需求的数据、信息及知识集合的“数据库服务系统”势必会升级为“知识云服务系统”。

因此,从身心语言程式的“身心”理论范畴去定义,信息检索用户的身心就是处于用户特定检索环境中的用户判定信息需求、谋定检索策略时的思维模式的总称。在此过程中,信息检索用户可以通过思维、视觉或听觉等获取到相关的信息,然后谋定信息检索策略并将其保存到脑以及其他存储设备中。

3 语言理念下的用户检索语言认知、表达与交互模式

信息检索中,用户在内心深处同自身进行内部对话的同时,还与除自己以外的他人进行交流。检索用户不仅需要用特定的语言来表达自身内心深处的信息需求,还需要与周围世界进行有效交互。其中所运用的语言不仅包括图像、文字、声音、体态、手势等有形语言,还包括思维、信念、习惯等无形语言,其目的只有一个,就是通过语言更好地显现(或表达)出内心深处的真实图景。信息检索用户与人(包括自身和他人)或机器之间进行对话以及交互时所采用的语言方式就是NLP的“语言”理论范畴。

在具体的信息检索中,用户常用的语言是自然语言。目前在搜索引擎领域已经出现了很多的自然语言检索系统,比如大连尤里卡软件开发公司的尤里卡中文搜索引擎,它是基

于自然语言理解技术的智能中文搜索引擎,用户可以使用日常语言表达来检索。1983 年 Salton 和 McGill 在出版的名为 *Introduction to Modern Information Retrieval* 一书中就断言道:“语言学的方法对信息检索来说是必不可少的。”^[5] 的确如此。目前,世界范围内的自然语言处理技术研发已取得了卓越的成就,例如:语义网、智能信息检索已成为研究热点,大力研究、开发和应用个性化、人性化的用户友好型人一机接口等等。另外,已经有一些该方面的产品投放市场,如翻译系统、文本鉴定系统、自然语言接口等。

结合自然语言处理的层次划分(通常分为语音学、形态学、词汇学、修辞学、语法学、语义和语用六个层次)^{[4]274},信息检索用户的自然语言形式可分为:

(1)以语音学为主,修辞学、词汇学和形态学为辅的声音。由于技术瓶颈的原因,目前检索用户运用声音来实施检索的语音(或音频)检索系统目前还处于研发阶段。一旦其投入市场,势必会有大批的目标群体来择用这种利用

声音就可以实现检索目标的检索系统。

(2)以词汇学、语义学、语用学、语法学为主,形态学和修辞学为辅的书面文字。目前的信息检索中,用户最为常用的就是利用词语进行检索表达。

(3)以形态学为主,修辞学等修饰成分为辅的图形。在信息检索中,该种检索方式主要运用于公安侦查系统中。

(4)最后一种潜在的用户语言就是由声音、书面文字和图形的任意混合来表达。这仅仅是一种预期,或许在将来人们会设计出一种检索系统,该系统允许人们可以利用特定的文字、语音、语言或图片的任意组配对某一特殊的命题进行检索。

当然,在具体的信息检索过程中,信息检索用户的检索语言认知、表达模式并非以单一的语言形式出现,而是集抽象认知和具体控制于一体的复杂交互集,但并非不可认知。结合认知过程的探索与检索系统实践,我们可以从宏观的视角将信息检索用户的检索语言认知、表达与交互过程归结为如图 2 所示的模式。

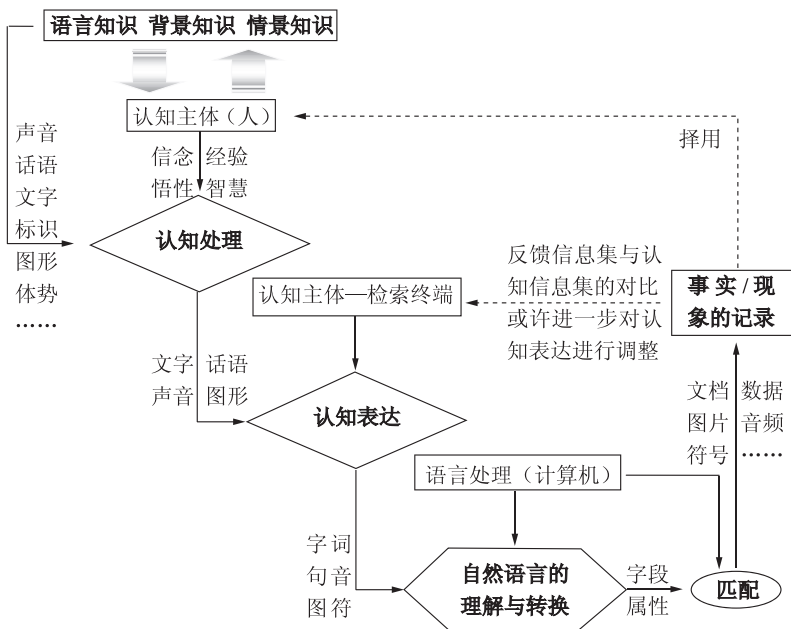


图 2 用户检索语言认知、表达与交互模式

自然语言是用户检索时最为常见的表达。对于一些具备专业检索知识技能的用户来讲,还有可能直接应用跨语言、受控语言或是规范化的其他语言形式来进行检索表达。当然,与检索表达息息相关的重要因素还有检索语词之间的相互组配关系。目前检索词间的组配关系主要有:布尔逻辑组配、截词组配、限制组配、位置组配、聚类组配、模糊组配、概念组配等。关于检索词的组配在此不细述,具体可参见赖茂生、贺德方等人的文献^{[4]147-171 [6]}。

4 程式理念下的用户检索策略谋定模式

结合 NLP 中“程式”的理论我们可以明确:信息检索用户的程式就是在检索过程中用户为获得预期的结果(信息)所采取的策略。所采取的特定信息检索策略在特定的检索系统中就会得到相应的系统反馈结果,现实中,这种结果有时是信息检索用户需要的,有时会不符合用户的需求^[7]。而用户在实施检索前,总对检索结果有美好预期,决定其预期的不仅有其内在的信息需求、检索系统环境,而且还有一整套指导个人与环境发生作用的策略。

事实上,在信息检索过程中,当用户确定了

信息需求后,处于身心环境中的需求表达就要发生转移,重心转向依赖特定的人机交互环境来完成其需求的系统实现。用户需求的系统实现过程就是用户与系统的博弈过程,是用户借助特定的信息检索系统进行适应性选择的过程,即用户针对特定信息检索系统,确定信息检索的途径与检索用词,明确检索词间的逻辑关系与查找步骤,实施检索、并择用相关信息的过程。目前,国内外影响较大的信息检索策略主要出自于兰卡斯特(F. W. Lancaster)、贝茨(M. J. Bates)、查尔斯·鲍纳(Charles Bourne)等西方学者之手,其中影响最大的是美国人查尔斯·鲍纳提出的五大联机信息检索策略^{[4]132-134 [6]324 [8]}。

在具体的信息检索策略制定过程中,用户的核心任务主要有:

(1)检索目标的制定。检索目标的制定是一个贯穿于信息检索始终的过程,根据 *Towards on Ecology of Mind* 一书的作者 Gregory Bateson 和 NLP 国际培训师 Robert Dilts 的心理取向模型^[9-11],我们将对信息检索用户的检索目标形成过程中的心理取向进行研究。

用户谋定信息检索策略的过程中,为了确定其具体的检索目标,信息检索用户会产生相应的规律性心智活动,其模型如图3所示。

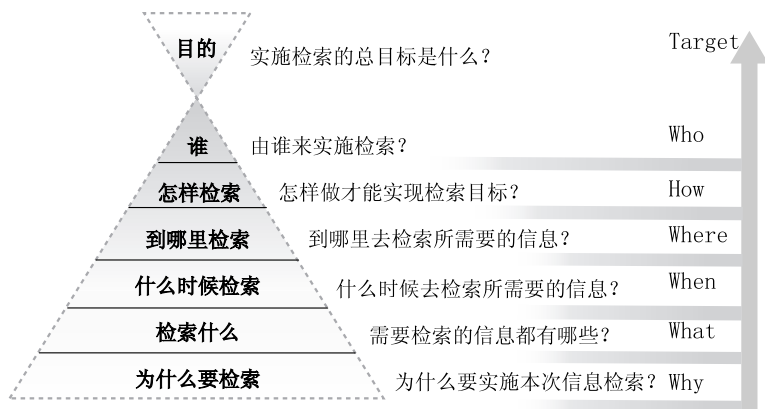


图3 用户检索目标制定中的心理取向模型

在谋定信息检索策略的过程中,如果用户对图3的层次有清晰的把握,并且能确保对上述

层次中的所有问题都极具确定性和及时性,那么其所谋定出的检索策略将与特定的检索目标、检索环境以及检索能力构成最有力的正反馈。此时,该用户与系统间就处于互为影响力最大的状态,也就是常说的巅峰状态。但在检索中,这种状态并非持续越久越好,如果超过适度范围,则不仅会导致检索用户时间和成本的浪费,而且还会接收到冗余的信息。

(2)检索策略的重构。对于信息检索用户而言,均无力改变相对稳定的信息检索系统,而只能选择改变自己的需求认知、言表以及检索策略。检索用户无法控制其所身处的检索环境,而只能改变自身对检索环境的态度。如果用户能从所处的信息检索系统环境中查阅到对自身的和生活和发展有利的信息,则会为满足其特定的信息需求而努力。这种努力主要表现在两方面:第一,用户会为了实施检索行为而积累支付相关费用的能力;第二,用户会为了得到较为理想的特定信息而进行多次尝试,以便构建出一种“自认为最合理”的检索表达或指令。而“用户多次尝试,以构建出一种‘自认为最合理’的检索表达或指令”正是身心语言程式里“重构”的理论范畴。

用户为了得到预期的信息而重新构建检索表达或指令的行为始终会贯穿于整个检索过程。在具体的信息检索中,即便是针对同一检索任务,不同的检索用户也会不断地重构出某一能迎合自身心理的、独特的、“自认为最合理”的检索表达或指令。信息检索用户的建构(Frame)和重构(Reframe)能力使信息检索系统得以发挥其大众化服务的功能,其中起关键作用的是检索用户的心态。例如:当某一检索用户对检索系统反馈的结果不满意时,就会沮丧或气愤;然而,当检索用户把所有的反馈看作是验证其检索策略、检验其检索表达或指令的合理性的途径时,无论反馈的结果如何,都是用户通过检索系统进行学习和提高的过程。

(3)检索中的信念。结合人类学和心理学,人的信念会成为习惯并随时间发展。信息检索中,用户的信念很多时候被当作真理而有意无意地左右着其检索行为。事实上,无论是何种

检索行为都和检索用户的信念息息相关。用户的这种信念可能是从父母或朋友那里学来的,也可能是自身成长过程中的经验和感悟。

在信息检索中,当我们对检索用户进行深入分析时,总能发现一系列其所“自认为的”或坚持的核心信念。当然,这些信念可能只对独特的个体有用,但我们也可以从宏观的视角去探讨信念本身或其变体。这些核心的信念有:没有一个人比我对该检索系统了解得更为透彻;我能随时改变我的检索指令;我想要什么,系统就能反馈给我什么;我已经支付了费用,有权检索到我想要的信息或相关的反馈结果一定会有用;即检即所得,即得即所需;信息检索系统反馈了所需的信息,这个检索系统足够人性化,等等。但是,事实是:信念终究是情绪化了的观点,无论信念源自哪里,它都不是真正的事实。这就要求信息检索用户正确看待潜藏于自身心中的信念,用成熟的心智面对信息检索系统环境。

(4)检索中的锚设定。信息检索过程中,用户既然对某种信息集抱有预期和信念,那么用户就会靠潜意识里的某种联想来选择一种独特的、有利于实现其检索目标的心理情境。用户潜意识里的这种反应实际上就是在自身与触发事件之间建立了一个心理情境,这种心理情境是身心语言程式中心锚(Anchor)的范畴。只要我们把图2和图3与系统论和控制论的感应与效应原理相结合,这点并不难理解。此过程中,触发事件就是检索目标和可以利用的检索系统。而图3所示的“用户检索目标制定中的心理取向模型”表明,检索目标的形成是用户在某种信息需求驱动下所产生的某种认知过程。至此可明确:信息检索用户的心锚主要涉及用户自身、用户认知和表达、认知语言三个核心要素。

我们知道,用户必须依赖特定的检索系统环境,才能实现其检索目标。于是,抽象的用户信息需求认知和表达(心锚范畴)最终要依赖具体的信息检索系统空间。当进行人一机交互时,检索用户借助特定的检索空间以求满足其信息需求的过程,正是身心语言程式的空间锚

(Spatial Anchors)范畴,这时检索用户将从心中原始的、抽象的信息需求(心锚范畴)出发,结合特定的信息检索系统空间(空间锚范畴),确定清晰的检索目标,同时谋定明确的检索策略。最终,用户通过与检索系统空间的交互获得特定的信息反馈,以满足其原始的信息需求。

事实上,图书馆学五定律的提出者阮冈纳赞(1963年),以及其后结合数字环境对“阮冈纳赞图书馆学五定律”的涵义进行延伸的 Noruzi (2004年)、Cloonan 和 Dove (2005年)等人无不立足于对当时的图书馆用户和信息检索者的需求认知、检索环境的宏观把握,才得以总结出经典定律体系^[12-15]。他们提出的经典定律正是对“锚”理论的成功实践:用户与资源之间存在着映射关系,这种映射关系正是规律性的直观反映。以下将以用户为中心,从“锚”的视角对检索中的认知规律进行探索。

5 锚视角下的用户信息检索认知规律

基于以上的信息检索用户认知分析以及“锚”探索,可以得出以下认识并由此推演出一些相关的论断:

(1)信息需求来源于人们对自身信息欲望的认知(理性或感性)过程,然而人的欲望是无限的,所以信息需求可以时刻产生,认知行为可以随时进行;

(2)信息检索中,只要讨论“有或无信息需求”,就可以肯定用户事先已展开了相关的认知活动;

(3)在信息检索中,不仅存在认知到的无需求和有需求的信息,而且一定存在用户未曾认知到的无需求和有需求的信息;

(4)信息检索用户说无信息需求和有信息需求,只是针对当时自己所认知到的有限信息集而言;

(5)信息检索用户最终择用的信息只会是检索系统反馈出的信息,并且肯定是用户认知到的、有需求的信息;

(6)检索系统所反馈出的信息并不等于用户认知到的有需求信息集,也不等于用户未认

知到的无需求信息集,更不等于它们二者的集合,而是等于用户“认知到”的所有“有需求的信息集”和“无需求的信息集”加上用户“未认知到”的所有“有需求的信息集”与“无需求的信息集”。

推论1:信息检索用户在得到一定的信息检索集,并在其中进行择用的同时,也遗漏了一些用户来不及认知或认知所不能及的有用(有需求)信息集。

推论2:检索系统反馈的信息集中,还有检索用户未曾认知到的信息,所以聪明的检索者会将每一次检索当作全新的认知过程。

若将上面的规律整合,结合空间集合理论,我们可以进一步得出如下结论:

需求空间 $\subseteq$ 认知空间

择用空间 $\subseteq$ 检索空间

我们知道,人的需求是欲望驱动下的认知结果之一,且用户对信息的择用是其借助系统环境实施检索后的行为;因而,我们可以进一步将上面的空间关系推演为:

需求空间 $\subseteq$ 认知空间 $\subseteq$ 欲望空间

择用空间 $\subseteq$ 检索空间 $\subseteq$ 系统空间

令需求空间、认知空间、欲望空间、择用空间、检索空间、系统空间和效用空间中的信息集分别为集合W、集合K、集合N、集合U、集合R、集合S、集合M,此时可以得出如图4所示的信息检索中的锚交互关系,对应空间中的集合关系将为:

$$N \subseteq K \subseteq W; U \subseteq R \subseteq S;$$

$$M = (N \cap K \cap W) \cap (U \cap R \cap S)$$

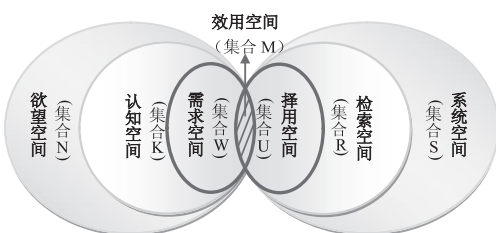


图4 信息检索中的锚交互模型

6 锚交互下检索用户认知假设模型构建

基于系统和欲望内涵的宽泛性,我们将对

空间范围进行约束,将其局限于认知空间、需求空间、择用空间和检索空间中,则我们可以构建出如表 1 所示的信息检索双因子矩阵。

表 1 信息检索双因子矩阵

锚 锚		认知空间		需求空间		检索空间		择用空间	
		认知	未认知	需求	无需求	检出	未检出	择用	未择用
认知空间	认知								
	未认知								
需求空间	需求	m_{11}	m_{10}						
	无需求	m_{01}	m_{00}						
检索空间	检出	n_{11}	n_{10}	n_{11}	n_{10}				
	未检出	n_{01}	n_{00}	n_{01}	n_{00}				
择用空间	择用	r_{01}	r_{10}	v_{11}	v_{10}	w_{11}	w_{10}		
	未择出	r_{01}	r_{00}	v_{01}	v_{00}	w_{01}	w_{00}		

将表 1 的信息检索双因子矩阵和集合理论相结合,我们可以确定如下的一系列概念:

(1)对信息检索用户有效用的信息就是由用户认知到的、有需求且由系统检出后用户最终择用的信息集。

若用集合 A 表示,则:

$$A = m_{11} \cap n_{11} \cap r_{11} \cap u_{11} \cap v_{11} \cap w_{11}$$

若定义:信息效度是检索出的信息对用户的有用程度,且用 E 表示。则针对某一主题进行检索后的信息效度可表示为:

$$E = \frac{m_{11} \cap n_{11} \cap r_{11} \cap u_{11} \cap v_{11} \cap w_{11}}{n_{11} \cup n_{10} \cup n_{11} \cup n_{10}}$$

(2)信息检索系统对用户需求的满足程度,是指用户认知到的检索系统反馈信息集对用户认知到的信息需求的满足程度。

若用 S 表示,则某一信息检索系统对用户特定信息需求的满足程度为:

$$S = \frac{n_{11}}{m_{11}}$$

(3)对用户来讲,信息检索系统的查准率为系统检出的、且对用户有需求的信息集与系统所检出的全部信息集之间的比例。

若用 P 表示信息检索系统查准率,则对于用户来讲,信息检索系统的查准率为:

$$P = \frac{n_{11} \cap u_{11}}{n_{11} \cup n_{10} \cup n_{11} \cup n_{10}}$$

(4)对用户来讲,信息检索系统的查全率为系统“检出、且用户对其有需求”的信息集,与所有“检出与未检索出的、且用户对其有信息需求”的信息集之间的比例。

若用 R 表示信息检索系统查全率,则对于用户来讲,信息检索系统的查全率为:

$$R = \frac{n_{11} \cap u_{11}}{n_{11} \cup n_{01} \cup n_{11} \cup n_{01}}$$

7 模型检验

本文以“云计算在图书情报与数字图书馆中的应用”为检验主题,借助中国知网“学术文献总库”文献检索平台,根据检索者的身心判断,提取出检索系统所提供的部分与“云计算”相似或相关的检索词为主题进行精确检索时(检索日期:2010/10/29),共得到检索系统反馈的所有记录 3293 条。但以“云计算”为直接主题进行检索时,检索系统直接生成的记录为

2349 条,其中检索者判断的有关云计算与图书馆关联研究主题的相关记录为 158 条。但是依据中国知网的学科分类,检索系统直接反馈的与“图书情报与数字图书馆”直接相关的记录仅有 78 条。

本文以陶新权等作者撰写的“云计算在图书馆中的应用”为例,最终得到的 78 条记录中有 11 条与“云计算在图书馆中的应用”一文不相关,而该作者最终择用的相关记录仅为 4 条。则可以确定出如表 2 所示的数值矩阵。

表 2 用户检索的双因子数值矩阵

锚 锚		认知空间		需求空间		检索空间		择用空间	
		认知	未认知	需求	无需求	检出	未检出	择用	未择用
认知空间	认知								
	未认知								
需求空间	需求	3293	—						
	无需求	2271	—						
检索空间	检出	2349	—	67	11				
	未检出	944	—	80	—				
择用空间	择用	4	—	4	—	4	—		
	未择出	74	—	74	11	74	—		

注:用“—”表示缺省项。其中:①无论使用何种检索平台,认知空间中“未认知”项为缺省;②正常情况下,检索过程中绝大多数用户都希望那些无需求的信息没有被检出,所以“无需求,未检出”项为缺省;③用户不可能也没有必要择用那些无需求的信息集,所以“择用,无需求”项为缺省;④用户一定不会、也不能择用那些未检出的信息,所以“未检出,未择用”项为缺省。

针对“云计算在图书馆中的应用”这一主题来讲,本次检索中,该检索系统所反馈出的信息集对用户的有效程度为:

$$E = \frac{m_{11} \cap n_{11} \cap r_{11} \cap u_{11} \cap v_{11} \cap w_{11}}{n_{11} \cup n_{10} \cup n_{11} \cup n_{10}} = 4/2349 = 0.001703$$

该检索系统对用户检索有关“云计算在图书馆中的应用”文献这一特定需求的满足程度为:

$$S = \frac{n_{11}}{m_{11}} = 2349/3293 = 0.713331$$

针对“云计算在图书馆中的应用”这一主题,该系统的查准率为:

$$P = \frac{n_{11} \cap u_{11}}{n_{11} \cup n_{10} \cup n_{11} \cup n_{10}} = 67/2349 = 0.028523$$

针对“云计算在图书馆中的应用”这一主题,该系统的查全率为:

$$R = \frac{n_{11} \cap u_{11}}{n_{11} \cup n_{01} \cup n_{11} \cup n_{01}} = 67/3293 = 0.020346$$

结合本次实验,可推知:

- (1)用户在“未认知”状态下,不会产生任何形式的“有需求”或“无需求”;
- (2)用户在“未认知”状态下,允许检索系统反馈产生“意外”,即得出意想不到的与需求相关的检索结果,但用户不能也不可能对系统“未检出”的与需求相关的信息作如此反应;
- (3)“未认知”状态无法上升到择用空间,即当面临“择用”或“不择用”某些信息时,用户预先肯定对所有的信息集进行了认知;
- (4)“无需求”顺应了“未检出”,即检索用户是理性的,自然不会“择用”那些“无需求”的信息集;既然是“无需求”的信息,系统就没有必要检出,这种“无需求”、“未检出”交互状态构成了一个完美的零空间。

(5)系统“未检出”,自然就谈及不到“择用”与“未择用”。
因而可知:

$$\begin{aligned} \{m_{10}\} &= \{m_{00}\} = \{n_{00}\} = \{r_{10}\} = \{r_{00}\} = \{u_{00}\} \\ &= \{v_{10}\} = \{w_{10}\} = \{w_{00}\} = \emptyset \end{aligned}$$

则检索双因子矩阵可优化为表3的形式。

表3 信息检索双因子优化矩阵

锚 锚		认知空间		需求空间		检索空间		择用空间	
		认知	未认知	需求	无需求	检出	未检出	择用	未择用
认知空间	认知								
	未认知								
需求空间	需求	m_{11}	—						
	无需求	m_{01}	—						
检索空间	检出	n_{11}	n_{10}	u_{11}	u_{10}				
	未检出	n_{01}	—	u_{01}	—				
择用空间	择用	r_{11}	—	v_{11}	—	w_{11}	—		
	未择出	r_{01}	—	v_{01}	v_{00}	w_{01}	—		

与传统的信息检索定律相比,本研究所建立的模型以用户为主导,深化传统检索中“相关”与“不相关”的旨义;并且,若用本文的模型去研究文献检索,则能够更加具象地描述“相关文献”、“命中文献”以及“择用文献”的内涵,有助于开拓集合论视角下的信息检索系统研究领域。另外与检索线性模型相比,空间、集合视角下的用户信息检索研究无需确定线性基点,只需要找各因素之间的关联性(无论正相关还是负相关)就可以建立模型,展开探索。

8 总结

本文用身心语言程式的理念研究了信息检索中用户的需求认知、表达、交互模式以及寓于其中的策略,从锚的视角总结出了信息检索中的一系列规律,得到了相关的空间、集合描述模型,并且对模型进行了假设和检验。在现实中,除了上述的检索用户普遍认知规律外,还会有一些生理、利益、习惯和经验方面的规律。另外,在探寻了信息检索用户认知规律后,有必要

对用户的检索行为进行探究,但限于选题及篇幅,在此不予展开。探究用户的检索认知和行为规律,言其可行是因为无论用户的检索行为如何复杂多变,其目的和属性是相对稳定的。复杂多变的是检索系统中的信息以及用户对检索系统的态度,而相对稳定的是人们为满足其信息需求而实施检索时的心智模式、表达方法与检索策略。

参考文献:

[1] 李中莹. 重塑心灵: NLP——一门使人成功快乐的学问[M]. 修订版. 北京: 世界图书出版公司, 2006: 1-24.

[2] Knight S. NLP Solutions: How to model what works in business and make it work for you[M]. London: Naperville, Ill: Nicholas Brealey, 1999: 1-41, 60-88.

[3] [英] 休·奈特. NLP入门[M]. 郑日昌, 译. 上海: 上海人民出版社, 2006: 1-4.

[4] 赖茂生, 等. 计算机情报检索[M]. 第2版. 北京: 北京大学出版社, 2006: 128.

[5] Salton G, McGill M J. Introduction to modern information retrieval[M]. New York: McGraw Hill, 1983: 286-288. (下转第99页)