

·信息工作·

# 交互问答平台的运作机理分析

徐 刚 (武汉大学体育部 湖北武汉 430072)

邓胜利 (武汉大学信息管理学院 湖北武汉 430072)

**摘 要:**交互问答平台是从服务而不是技术的角度来实现问答式服务。在综合各种实际产品应用,并结合其提供的服务的基础上,文章从交互问答平台的构成模块出发,分析交互问答平台的运作流程和技术支撑,同时,从用户和控制角度探讨交互问答平台的管理问题,为全面了解交互问答服务平台运作机理提供参考。

**关键词:**交互式信息服务 交互问答平台 运作机理

中图分类号:G350.7

文献标识码:A

文章编号:1003-6938(2009)04-0088-05

## Operation Mechanism Analysis of Interactive Question and Answer Platform

Xu Gang (Physical Education Department, Wuhan University, Wuhan, Hubei, 430072)

Deng Shengli (School of Information Management, Wuhan University, Wuhan, Hubei, 430072)

**Abstract:** The interactive question and answer platform realizes the interrogatory method service not from the technical but from the service. On basis of synthesizing each actual product application and unifying the service which provides, this article embarks from constitution module of the interactive question and answer platform, analyzes its operation process and technical support. Simultaneously, from the aspect of user and control this article discusses management problems of interactive question and answer platform, which provides a reference for comprehensive understanding its operation mechanism.

**Key words:** interactive information services; interactive question and answer platform; operation mechanism

CLC number: G350.7

Document code: A

Article ID: 1003-6938(2009)04-0088-05

交互问答平台的提出是希望解决用户的自然语言提问和为用户提供直接的答案服务,但它不是从技术的角度出发,而是从服务的角度,充分运用人的智慧来实现这种服务。交互问答平台,是指将具有信息需求的人与能提供帮助的人和信息联系起来的一种服务平台,它基于“提问——回答”这样一种服务表现形式,用户可以以自然语言的方式提问,由其他用户或专家来提供问题的直接答案。<sup>[1]</sup>

### 1 构成部分

交互问答平台的功能主要体现在两个方面:问答和搜索。综合目前推出的各种交互问答平台来看,其主要构成模块包括以下几个。

#### 1.1 提问模块

用户提问窗口一般会在首页显示,现有的提问方

式都基于Web表单形式,用户采用自然语言提问,针对提问有具体的问题设置页面,其中主要包括问题的主题、问题详细说明、问题分类、悬赏分/价格四个部分,但Wondir较为特殊,它只在首页提供了一个问题输入框,提问者可以直接在框中对问题进行较为详细的说明,然后一次性提交。<sup>[2]</sup>问题的主题是对问题的一个概括,为便于回答者更好地理解提问者的需求,提问者可以在“问题详细说明”中对提问的背景进行进一步的解释。值得一提的是新浪的“爱问知识人”在这里提供附件上传功能,可以对问题进行更为详细、全面的说明;交互问答平台对所有的问题进行分类,提问不受限制,为问题选择分类是在平台已有的分类目录下进行选择;采用积分激励制和收费制的交互问答平台在提问的下面还会要求进行悬赏分或价格的设置。

#### 1.2 回答模块

基金项目:本文系教育部人文社会科学研究青年项目“基于Web2.0的交互式信息服务研究”(08JC870009)研究成果之一。

收稿日期:2009-02-20,责任编辑:刘全根

用户可以在浏览问题时直接点击问题的标题来回答,也可以通过分类目录查找到某类下的问题进行回答,Wondir在首页还设置了问题接收框,新的问题可以不断在框中滚动出现,用户直接点击相应的问题即可进入回答页面,“爱问知识人”在个人信息管理模块中设置“找自己感兴趣的问题来回答”的超链接,链接到问题分类页面,用户可以在分类目录中寻找自己感兴趣的问题。<sup>[3]</sup>回答模块负责回答的接收,与提问相似,回答也是采用Web表单的方式接收。在接收的方式上,Wondir提供了较多的选择,包括通过即时消息工具提交、直接发送email给提问者、发送到问题板上三种方式,在回答的同时,用户还可以将该问题转发给好友、查看该问题的搜索结果、添加到个人收藏夹中,其中查看问题的搜索结果既可以看到与该问题类似的其他问答信息,同时也可以看到通过网络搜索到的结果,Wondir内置了元搜索服务。<sup>[4]</sup>新浪“爱问知识人”提供上传附件的功能,以便于更好地回答提问者;其次,同“百度知道”一样,二者都要求回答者提供参考资料的说

### 1.3 问答信息浏览与检索模块

交互问答平台将各种问题进行主题分类管理,提供基于分类类别的浏览功能,由于交互问答平台大多是综合性的开放服务平台,因此问题涉及的范围广泛。国内交互问答平台的分类设置基本相似,如“百度知道”将问题归为13大类:即生活/休闲/时尚、电脑/数码、健康/医疗、教育/科学、商业/经济、体育/运动、艺术/文学、社会/文化、娱乐/游戏、明星/名人、烦恼/地区;“爱问知识人”分为12大类:即电脑/通讯/互联网、家庭/生活/游戏、科学技术、演艺/娱乐、运动/爱好、社会/文化、商业/经济、教育、健康/医学、旅游/交通、地区问题;在类目层次上,“爱问知识人”达到了三级,而“百度知道”只达到二级类目。国外交互问答平台的分类差别就很大,Wondir类目很多,一级类目就达到34个,并且类目级别达到4级,另外为便于管理,每个类目名的后面括号里都注明其缩写方式,如Animals&Pets简称为ANI,类目的划分也很细,如最后一个类目“Gourps”下面有15个二级类目,点击某个类目,会出现所有与该类目为主题的问题列表,点击问题标题,就可以进入具体的问答信息界面浏览。

另外,交互问答平台还提供问题的检索服务,在提问框中输入问题,点击“搜索答案”就会出现与该问题相关的已回答问题的检索结果列表,直接查看问题的答案或提供参考答案。目前,“爱问知识人”、“百度知

道”、“雅虎知识堂”都提供所有问题的统一搜索。Wondir在提问页面没有提供检索服务,但在用户提问提交之后,会直接出现与你的提问相关的问题检索结果列表,并且还提供网络搜索的结果列表,以供提问者参考,若没有合适的,可以再选择提问。

### 1.4 个人信息管理模块

大部分的交互问答平台要求用户注册才能进行提问和回答,未注册的用户只能浏览问答信息,由于很多交互问答平台都是基于搜索引擎或门户网站提供的,许多用户早已经是某项服务的注册用户,为方便用户和便于管理,各交互问答平台会采取通行证的方式,即只要是服务商的注册用户就可以直接用已有的帐号和密码登录,而无需再次注册。如“百度知道”可以采用“百度贴吧”的帐号和密码登录,新浪和Google都可以采取邮箱用户名和密码登录等。针对注册用户,每个互动问答服务平台为用户提供专门的个人信息管理模块,让用户对个人的基本信息、提问/回答记录进行统一管理,此外还可以享受一些其他的特色服务,如新浪“爱问知识人”提供“我的共享资料”和“我的开放词典”服务,Wondir提供的“新问题提示”服务等。

### 1.5 说明模块

说明模块包括对平台的介绍、平台使用说明、规则说明、服务政策等,具体的就是指如何提问、如何回答、问题处理、如何搜索、服务条款、平台政策等。交互问答平台对提问的范围一般是不加限制的,任何人都可以提问,但为避免出现一些不良信息,还必须制定一定的政策,如对于一些没有意义的提问或是打探他人隐私的提问,平台有权进行删除;对于造成社会危害、带有歧视类的问题禁止提问等。

## 2 运作流程

用户在向其他网民提问之前,可以先进行答案的搜索,平台将给出类似的已回答的问题列表,当用户对这些仍不能满意时,可以进行提问,详细描述问题,然后提交给系统,问题将发布在平台上,等待其他用户来回答,最终提问者从回答中获取答案。所有的问答信息最终将进入知识库,供用户提问搜索。这是所有交互问答平台所共有的基本流程(见图1)。但由于具体发展的水平不一样,各个阶段的处理又有所不同。国内的交互问答平台目前在智能性方面还不足,在用户提问之后不能自动进行检索服务,对问题进行查重,或提供相似的已回答问题以供用户参考,只有在用户进行“搜索答案”时,才能进行知识库的搜索。国外的Wondir在这方

面发展的较完善,当用户提问后,首先,他们将获得一个结果页面,就像其他任何的搜索引擎搜索的结果一样,系统自动为问题进行分类,如果用户觉得不合适,还可以自己更改类目。该页面包括已回答的相似问题列表、Web资源列表以及回答者的匹配结果。用户可以对这些搜索结果进行浏览,以快速获得答案。若已有的资源不能使用户满意,则用户可以选择向其他用户询问。Wondir对提问者问题进行回答者的自动匹配,匹配结果也显示在结果页面,用户可以通过点击“Ask now”的链接,向回答者发送email,等待用户的回答。另外,提问还被发送到站点的不同区域,如按照分类类别来组织问题的可检索知识库,主页中的问题接收框(它显示了最新的一些提问),向申请提供该方面问题提示的用户发送email。

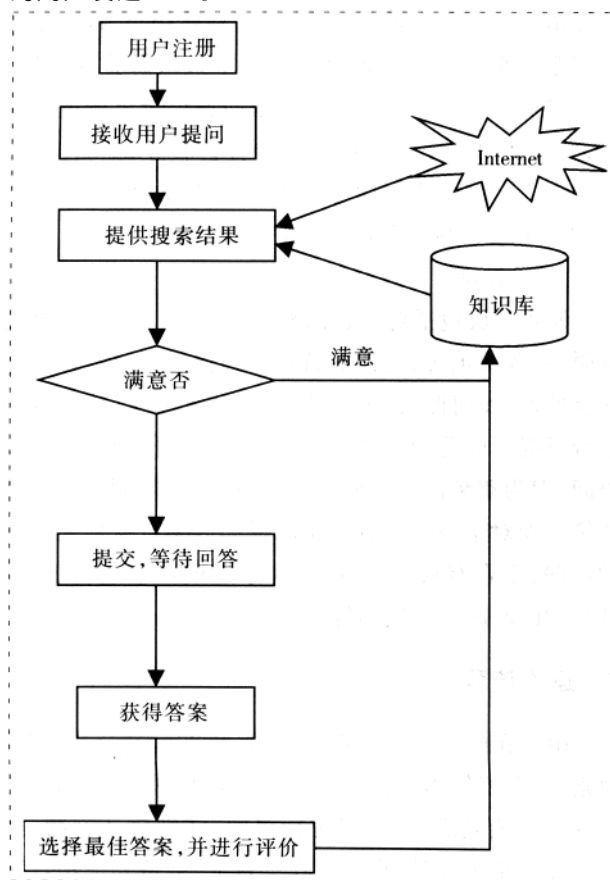


图1 交互问答平台的运作流程

### 3 技术支撑

#### 3.1 自动分类技术

交互问答平台在接收用户提问时,根据用户问题的主题自动选择类别,类别的产生是在交互问答平台已有的类目体系下选择的,包括选择一级类目和二级类目等。目前,交互问答平台在提供自动分类的同时,

也提供人工分类方式,即由用户自己为问题选择类别,在认为系统自动分类与自己的实际需求不相符合时,也可以对系统的自动分类进行手工更改。另外,针对大量的问答信息,平台将建立知识库,对所有的提问和回答进行存储,提供进一步的检索功能。交互问答平台才刚处于起步阶段,但从已有的统计来看,用户的需求量巨大,这必将形成大量的问答知识,采用人工进行知识加工和整理显然是不够的,必须利用计算机进行自动标引、自动分类和自动排序等,形成一个有序的知识库。

为问题进行自动分类有利于回答者更好地理解提问者的出发点,同时也有利于对问答信息的存储和检索,是基于问题分类组织的交互问答平台的主要支撑技术之一。

#### 3.2 高级检索技术

较早推出的台湾Yahoo! 奇摩知识+提供“进阶搜寻”,从关键词、搜寻范围、指定分类、问题状态、发问时间、正面评价6个方面进行限定检索,提供的是布尔逻辑检索的“and”检索功能,其检索范围限定在问答知识库内。而实际上,在用户提问后,系统不仅要同时对问答知识库进行检索,同时也应该结合网络资源,进行相关匹配检索,将搜索引擎的功能充分融合在其中。为保证搜索的全面性和准确性,交互问答平台应采用元搜索引擎技术。

交互问答平台Wondir实现了Web资源的搜索,采用的就是元搜索引擎技术,它通过与AllTheWeb、About.com等搜索引擎合作,当用户提问问题时,Wondir负责对用户的问题转发给各个搜索引擎,各搜索引擎通过自己的检索接口进行搜索,最终在Wondir中显示的Web资源搜索结果列表是将各搜索引擎检索到的信息进行查重、排序形成的结果,每条检索结果将显示网页的标题、部分摘要以及该条信息的来源,如来自AllTheWeb等。

高级检索技术将网络资源纳入检索的范围,方便用户获取更多的信息,在交互问答平台中直接提问获取答案,而不需要再到各个搜索引擎中进行检索,为用户提供了“一站式”服务。采用元搜索技术,免去了管理和维护庞大的索引数据库而需要花费的巨大精力,交互问答平台可以集中于对自己积累的知识库进行管理。

#### 3.3 信息发布与推送技术

交互问答平台作为一种服务平台,主要是以Web页面的形式为用户提供相互交流的窗口,同时为便于用户利用和吸引更多的用户参与,交互问答平台向用户提供信息发布和信息推送服务。



信息发布可以分为HTML文件类型发布和数据库查询信息发布两种。在主页发布的“精彩推荐”、“主题知识”、“精彩回答”、动态问题接收框等就属于HTML文件类型发布。另一个途径就是给用户提接口实现数据库查询功能。对比分析国内外交互问答平台发现,国内的交互问答平台由于才刚刚起步,更多是对问答信息的二次组织,偏重于信息发布,如对回答比较精彩的问答进行特别推荐,配以图片的形式,这种做法对于交互问答平台能够起到很好的宣传效果,吸引更多的用户参与。国外的交互问答平台在检索功能方面更加强大,不仅提供对知识库的查询,还充分利用网络资源,利用元搜索引擎对网络信息进行搜索和综合。

交互问答平台在提供信息发布服务的同时,也针对不同用户的兴趣和爱好进行推送服务,即通过RSS订阅和email方式向用户主动推送相应的新信息。交互问答平台提供对各分类类目的订阅,用户需要下载和安装一个RSS阅读器,然后通过平台提供的RSS标记来订阅感兴趣的内容,在RSS阅读器中添加一个频道,复制XML地址即可,RSS阅读器能自动地以指定间隔时间检查服务器,收取新的内容。

另外,国外的交互问答平台还提供电子邮件推送服务,Wondir通过一个“New Question Alerts”进行问题的推送,用户首先需要在平台进行注册,在其个人模块“My Wondir”中用户可以对“New Question Alerts”进行设置,用户可以输入关键词或词组进行问题内容的限制,并设置发送的时间频率,服务器定期扫描用户需求文档,提取用户需求特征信息提供给数据库进行检索,并将检索结果以email的形式发送给用户。

信息推送服务方便用户及时获取自己感兴趣的信息,同时也提高了用户对平台的关注。这也是交互问答平台吸引用户积极参与的有效方法之一。

### 3.4 即时消息技术

即时消息(IM)已成为受到广泛欢迎的交流方法,简单地说,IM就是两个或更多的人之间的实时联机交流。这种媒介使得人们可以通过网络进行文本信息的快速交流来实现“聊天”。通过向特定的窗口(即由IM软件建立的“聊天室”)键入信息,一个用户就可以与另一个用户“聊天”。信息几乎同时出现在另一个用户的显示器上,接收者阅读后键入答复。

IM给交互问答平台提供了另外一种寻求问答服务的途径。除了快捷、方便、可获取之外,这种即时消息服务能够让提问者和回答者更好的沟通,及时获取答案。如果说交互问答平台连接了信息需求者和信息提供

者,那么即时消息技术的应用就更进了一步。许多用户已经熟悉了这个程序,并且在自己的计算机上安装了这种程序。它安装起来也十分容易。安装有Java的网络浏览器可使用网络版的AIM Express (Express),无须下载。这种程序对带宽的要求不高,运行很快。目前IM主要产品有AOL Instant Messenger, ICQ, Microsoft Messenger, Yahoo! Messenger等。

Wondir交互问答平台实现了用户之间的即时交流,它采用了Jabber即时消息(IM)系统。<sup>[5]</sup>Jabber使用的是客户端——服务器端的系统架构,而不是其他一些即时消息系统使用的客户端——客户端的系统架构。所有从一个客户端发给另一个客户端的Jabber信息和数据都必须通过服务端。任何一个客户端都可以通过商议与另一个客户端自由地建立一个直接连接,但这些连接必须先通过一个客户端——服务端形式进行协商,才能建立。Jabber在设计上很大程度地沿袭了Internet上最成功地消息系统email。Jabber可以在一个使用共同协议的服务器组成的分布式网络上提供通信,连接这个网络的客户端,可以像接收消息一样发送消息给同一个服务器或其他Internet上的服务器上的用户。不过,尽管email是一个存储——转发系统,但Jabber转发信息却是实时的,因为Jabber服务器知道一个用户什么时候在线,这也是即时消息的核心所在。

此外,“新浪爱问知识人”也曾尝试提供专家视频答疑,它借用新浪的聊天室开展用户与专家之间的实时交流,但仅仅是作为活动来开展,是有选择性地应用。

即时消息的应用避免了用户等待回答的时延,当有合适的答案提供者在线时,用户可以要求进行即时消息问答,直接获取答案。

## 4 运作管理

各互动问答平台具有主体上相同的运作流程,它仅仅围绕“提问——回答”这样一种目标来进行,绝大多数的互动问答平台需要用户首先注册才能成为提问者/回答者,这有利于对用户进行控制,同时也能有针对性地提供更多个性化服务。在问题的解决过程中,互动问答平台首先提供搜索服务,基于已有的问答知识库和高级的元搜索引擎技术,提高了问题的解决效率,避免重复提问。<sup>[6]</sup>互动问答平台是以用户为中心的服务平台,因此对于用户的提问,最终由用户对问题进行处理。在整个流程中,平台主要需要对用户进行管理,包括注册、个人信息管理等;此外还需要对问题进行控制,包括问题的审核与发布、问题处理、问题存档。

#### 4.1 对用户的管理

用户是互动问答平台的主要支撑,提问和回答都需要用户的参与,因此聚集人气是互动问答平台必须注重的。对用户的管理主要体现在两方面:用户认证和用户个人信息管理。用户认证是通过提供注册/登录来实现的,百度知道、新浪爱问等均实行了通行证的方式,即已成为百度或新浪的用户可以直接根据已有的个人用户名和密码进行登录互动问答平台,新的用户需要注册,为个人设置用户名、密码以及联系的电子邮箱地址等信息;用户个人信息管理是通过提供专门的个人信息管理模块来实现的,以个人中心、我的档案等来命名的模块。

#### 4.2 对问题的控制

对问题的控制包括问题审核与发布、问题处理和问题存档。互动问答平台以表单的方式接收用户的提问,提问必须经过管理人员的审核,没有违反平台制定原则的问题将在平台中发布;问题处理是指在规定的期限内,仍然由其他用户来回答问题,如果提问者在规定期限内没有对问题进行处理(选择最佳答案、通过提高悬赏延期问题有效时间、设置投票和选择无满意答案),且该问题有一个或以上的回答,则问题会过期并进入“已解决”状态,过期问题如果不处理,则用户无法

进行新的提问;问题存档是在问题处理结束后,将问题和相应的答案存入数据库中,形成问答知识库。

参考文献:

- [1]邓胜利.基于用户体验的交互式信息服务[M].武汉:武汉大学出版社,2008:24-26.
- [2>Welcome to My Wondir![EB/OL].[2006-08-20].<http://www.wondir.com/wondir/jsp/index.jsp>.
- [3]“爱问”推出中国大百科在线搜索服务[EB/OL].[2006-06-30].<http://tech.bokee.com/55/2005-08-03/380053.html>.
- [4]John Battelle.Grokking Wondir[EB/OL].[2006-06-27].<http://battellemedia.com/archives/001058.php>.
- [5]Jabber技术概况[EB/OL].[2008-12-20].<http://jerryes.cnblogs.com/archive/2005/12/13/296435.html>.
- [6]Wade Roush. What's the Best Q&A Site?We put Yahoo Answers, Amazon's Askville, and rival question-and-answer services to the test[J]. MIT Technology Review, December 22, 2006.

作者简介:徐刚(1956-),男,武汉大学体育部资料室馆员,研究方向:信息管理;邓胜利(1979-),男,武汉大学信息管理学院讲师,研究方向:信息管理理论与应用。

(上接第82页)译名建议[J].科技资讯,2006(16):231-232.

[5]龙丽,庞弘.基于Copyleft理念的Creative Commons发展[J].图书馆学研究,2008(8):93-97.

[6]Charles R. Beitz. The Moral Rights of Creators of Artistic and Literary Works [J]. The Journal of Political Philosophy. 2005,13(3):344.

[7][EB/OL].[2009-01-13].<http://service.wanfangdata-com.cn/Announce.aspx>.

[8][EB/OL].[2009-01-13].<http://www.cqvip.com/Copy-right/copyr.shtml>.

[9]李武,刘兹恒.一种全新的学术出版模式开放存取出版模式探析[J].中国图书馆学报,2004(6):66-69.

[10]Removing the Barriers to Research: An Introduction to Open Access for Librarians [EB/OL].[2009-01-13].<http://eprints.rclis.org/1027/1/acrl.htm>.

[11]黄吉瑾,张心全.当Copyright遇上Copyleft[J].前沿,2004(11):201-205.

[12]Doyle H., Gass A., Lappin D., Kennison R.. Analysis of Open Access Landscape[J].图书情报工作,

2005(12):6-13.

[13]In Oldenburg's Long Shadow: Librarians, Research-Scientists, Publishers, and the Control of Scientific Publishing [EB/OL].[2006-05-20].<http://www.arl.org/arl/proceedings/138/guedon.html#iva>.

[14]马红.比较OA期刊与OA知识库实现途径的差异分析[J].现代情报,2008(2):203-204,208.

[15]E-LIS Copyright Issues [EB/OL].[2009-01-12].<http://eprints.rclis.org/copyright>.

[16][EB/OL].[2009-01-12].<http://dlist.sir.arizona.edu/information.html>.

[17][EB/OL].[2009-01-12].<http://caltechlib.library.caltech.edu/information.html>.

[18][EB/OL].[2009-01-12].<http://www.ozone-db.org/frames/home/license.html>.

[19]苏建华.论文抄袭识别技术研究综述[J].数字图书馆论坛,2007(11):61-64.

作者简介:范贤容,(1985-),女,中南大学医药信息系硕士研究生,研究方向:医院信息系统;韩欢(1985-),女,中南大学医药信息系硕士研究生,研究方向:信息资源管理。