

甬图文萃

第二辑

宁波市图书馆
宁波市图书馆学会 编



宁波出版社

图书在版编目(CIP)数据

甬图文萃/陈宁雄等主编. —2版. —宁波: 宁波出版社, 2002.12

ISBN 7-80602-174-4

I. 甬... II. 陈... III. 图书馆工作—文集
IV. G25-53

中国版本图书馆 CIP 数据核字(2003)第 060582 号

责任编辑:吴百星 陈金菊 王伯城

封面设计:陈 杰

书名题字:谢灼华

书 名:甬图文萃第二辑

作 者:陈宁雄等

出版发行:宁波出版社

(宁波市苍水街 79 号 邮政编码:315000)

印 刷:宁波银行印刷厂

开 本:850×1168 毫米 32 开

字 数:211 千

印 张:8.5

版 次:2002 年 12 月第 1 版第 1 次印刷

印 数:1~1000 册

书 号:ISBN 7-80602-174-4/G·78

定 价:20.00 元

甬图文萃编辑委员会

主任：陈宁雄

编委：(按姓氏笔画为序)

庄立臻 何玉娟

吴百星 陈金菊

袁芳芳 谭龙敏

目 录

改革与发展

统一认识 形成合力 加快宁波文化设施建设步伐	陈宁雄(1)
数字图书馆中的知识发现	韩惠琴 刘柏嵩(6)
论数字图书馆	沈冠武(20)
对数字图书馆作品著作权的若干思考	陈金菊(25)
图书馆的改革与展望	李张敏(33)
未来远程图书馆的发展态势	陈 俭(39)
浅谈家庭藏书与社会文明	龚新曼(53)
把握“入世”机遇 拓展图书馆的社会职能	邬雪琴(57)
请给农民读者应有的国民待遇	余巨平(62)
提高档次 提升品位	钱 俊(68)
图书馆与地区经济发展	顾正葆(73)

管理与人才

论转型期公共图书馆的人才培养	袁芳芳(76)
浅谈公共图书馆的人力资源发展	王黎黎 袁志伟(84)
宁波市公共图书馆人才资源分析与对策	袁芳芳(89)
以馆员为中心提高图书馆管理水平	鲁黎明(96)
对图书馆实施轮岗制的几点思考	何继奋(101)
信息时代图书馆员素质浅谈	吴月珍(105)
浅谈图书馆员的职业道德建设	俞红国(108)
试论当代图书馆员的继续教育	俞国玉(114)

技术与创新

公共图书馆服务的新观念和新形式	庄立臻(121)
-----------------------	----------

知识经济·科教兴国·图书馆·····	陈金菊(127)
图书馆网站建设“十要”·····	俞国玉(133)
网络时代图书馆期刊工作变革浅谈·····	冯晓霞(138)
公共图书馆期刊信息资源开发利用新论·····	王丽萍(146)
文献利用值·····	常益民(154)
象山县图书馆自动化建设之路·····	王雨芬(160)

服务与提高

浅论网络环境下图书馆的信息服务·····	庄立臻(162)
基于 Internet 的读者服务·····	陈 俭(169)
图书馆的信息咨询服务·····	苏 薇(178)
强化为科技兴农服务·····	陈志雅(181)
浅谈少儿读者活动·····	马霞维(186)
浅谈社区阅览室的基本服务功能·····	何玉娟(189)
制约中学图书馆藏书利用率的因素及提高对策·····	罗尧敏(192)
学习《宣言》抓住时机 服务教改·····	瞿嘉福(199)
试论中学图书馆在素质教育中的作用·····	陆秀菊(203)
构建具有宁波党校特色的图书资料馆藏体系·····	夏根娣(206)

探讨与研究

有关国家《古籍著录规则》商榷二题·····	王元庆(211)
《中国图书馆分类法》“K20”类目体系析疑·····	魏海波(215)
评欧阳健《中国神怪小说通史》·····	张颖 陈速(226)
宁波古刊本举要·····	骆兆平(238)
天一阁古籍纸的分析及护养经验浅谈·····	李大东(246)
宁波高教园区图书馆模式分析·····	郑启明 蒋和法(251)
初中阅读指导课应定为校本课程·····	王明之(256)

统一认识 形成合力 加快宁波文化设施建设步伐

陈宁雄

宁波市政府提出了“扩大规模、完善设施、优化环境、提高品位”的城市建设的新思路,这一思路得到了全市上下的热烈拥护。进入九十年代以来,宁波市的经济建设和社会各项事业都得到了应有的发展,先后建设了体育、医疗、广电、职教四大中心,改造了宁波剧院,新闻文艺中心也在建设之中。然而宁波的社会事业发展仍严重滞后于经济的发展,尤其是文化基础设施建设更显落后。

人们说:“城市的地位靠经济,城市的品位靠文化”,根据国际上城市发展的规律,建设标志性文化设施是城市发展的必然趋势。标志性文化设施是一个城市的脸面,是推进城市化进程的重要标志。去过欧美考察的人一般都到过美国的国会图书馆、大都会博物馆,法国的密特朗图书馆、蓬皮杜艺术中心,澳大利亚的悉尼歌剧院等公共文化设施。这些建筑以其外观的独特壮丽和丰厚的文化内涵,成为这些城市的象征性符号和“灵魂”。宁波目前承担上述公共性文化设施功能的相应为包玉刚图书馆、天一阁博物馆内的史迹陈列馆、逸夫剧院和市群艺馆,但这些设施远远不能承担起副省级城市的文化标志功能。在文化部对全国同类城市的考核中,我市上述四项设施没有一个达到最起码的三级标准,可以说,宁波至今还没有真正意义上的图书馆、博物馆、大剧院和群众艺术中心。市图书馆规模已从建造初的全省第一下降到第五位,在十五个副省级城市中倒数第二,在苏锡常、杭宁温六城市中倒数

第一。

自去年下半年以来,我市的文化基础设施建设日益引起从市领导、“两会”代表到普通老百姓的高度关注。市委黄兴国书记就文化建设问题作了重要讲话,加快文化基础设施建设也成为“两会”讨论的热点之一。我市至今拿不出一个令宁波人骄傲的标志性文化设施,这已直接影响了宁波城市综合实力在全国同等城市中的排位,影响了我市的经济建设和改革开放的环境,也使宁波城市的品位大打折扣,城市的集聚、辐射能力受到极大削弱。

最近,我参加了市政协教文卫体委员会的宁波文化设施现状调查,并考察了珠江三角洲部分城市的文化设施建设。大家形成了如下一些想法和建议。

一、升华认识,放远眼光,高度重视文化设施建设。论经济实力,论历史文脉,宁波这座城市早就应该有几个标志性文化建筑,但为何到目前为止还是“千呼万唤不出来”呢?我们认为,关键的问题是在于搞城市建设中没有“文化品位”这一观念。长期以来,我们在抓城市建设和发展中的价值判断标准主要是看国民生产总值和经济增长速度以及拓马路、盖高楼,而忽视了城市的文化形象塑造。以至当今天我们回眸这座城市的轮廓时,就不由得感慨起它缺少鲜明的文化个性,显出几分苍白。

在新千年到来之际,我们对社会发展、城市建设的认识需要世纪升华,需要放远眼光。宁波要提前实现现代化,不能只看GDP,不能只看人均收入,还要看文化、看环境、看综合实力。当今国际上知名的城市都有着比较发达的文化事业和信息业,拥有众多的图书馆、博物馆、剧场、电影院等文化设施。有人说过,一个地区、一个民族、一个国家,它在世界文明史上的地位,最终是由它的文化发展达到什么程度来决定的。要建设现代化城市,文化设施、文化品位、文化氛围等方面是必不可少的。目前,上海、广州、深圳等城市,已经把文化设施的建设、文化事业的发展放到一个非常突出

的位置,如广州市市长林树森在今年的一次市委工作会议上用了1/3的讲话篇幅,纵论“城市环境、城市文化和经济工作的关系”,表示广州要树起自己鲜明的文化风骨。鉴于这样深刻的认识,广州在最近制定的“三年一中变”的规划中,对标志性文化工程做了具体规划,将由政府和民间共同出资数十亿人民币建设一系列水准达到国际标准的文化建筑。为此,我们希望市委、市政府和各级有关部门能站在面向现代化、面向世界、面向未来的高度,用长远的战略眼光来通盘规划城市建设中标志性文化建筑的实施,以使我们的这座城市更富文化品位,更有诗情画意。

二、列入规划,分步实施,加快文化设施建设步伐。今年是“十五规划”制订的重要一年,而从目前我市城市总体规划的框架来看,文化功能的凸现不尽明显,尤其是建设一批现代化文化设施的布局没有在总体规划中得到应有的重视和反映,这可以说是一个极大的遗憾。由此带来了最近在选址何处时的举棋不定与“临时抱佛脚”的被动状态。我们认为,规划是实施的前提。只有将拟议中的设想,经过慎重论证列入未来五年经济和社会发展的规划之中,并依法加以保障,才能确保预期的宏伟蓝图实现。为此应尽快制订城市总体规划的子规划——《文化设施建设和布局规划》,从宁波城市定位高度,明确建设对象、划定相应地块,提出完成期限,这是文化建设的当务之急。我们意见,在“十五”期间至少建成两个、启动一个、设计一个,力求高质量,高品位,我们在视察中感到城区留给文化的空间已经不多了。此外《文化设施建设和布局规划》也要象三江六岸规划、月湖景区改造一样,广泛征求专家和市民的意见,并集中一定的财力物力,组织精干班子予以操作和实施。

三、以人为本,完善功能,成为标志性文化设施。文化设施建设必须坚持以人为本、为民造福和可持续发展、一切从实际出发的原则。在选址上,一要建在市民便于参与的地段;二要建在视野开

阔、易于创造景观的地段；三要处在交通便捷、利于人流集散的地段等。在功能上，要高水平、高标准，努力为市民提供优美、整洁、安全、舒适的环境，使人们在认知中得到身心的休闲，在休闲中得到心情的陶冶，从而达到市民素质的极大提高。在服务上，要立足现代化，考虑多样性，尽可能为群众提供品位高雅、功能完善的文化设施。总之，我们不仅应把优秀的作品奉献给人民，还要让广大群众走进自己的文化乐园时充满自豪。

四、设计精美，质量第一，建设标志性文化设施。作为现代化港口城市、国家历史文化名城、长江三角洲南翼经济中心城市、国家优秀旅游城市，必须有一流的文化设施。如一流的大剧院、音乐厅，一流的图书馆、博物馆，一流的艺术中心等。要做到一流，首先建筑的设计要高起点，保证几十年不落后，可采用国际招标的方式遴选；其次建筑的施工要高质量，要实施全方位的监理；再次在管理和服务上要达到星级宾馆的水平，使之成为代表宁波城市形象的文化建筑艺术精品，体现城市的气魄与胸怀。不能再象当时建造宁波市图书馆那样，在十五年里两次新建，最终还是一个“无级馆”。文化设施的建设应该同时顾及眼前利益和长远利益，与城市发展的战略意图相吻合。

五、解放思想，引入机制，拓展建设筹资渠道。要学习上海、广州、深圳等地的成功经验和做法，解放思想，政策保障，动员社会各方力量，共同参与投资建设和经营大型文化设施。为此我们认为文化设施建设所需资金可从以下五个途径解决：

（一）以政府投入为主，建设公益性文化设施。党的十四届六中全会关于精神文明建设的《决议》指出：“图书馆、博物馆、科技馆是公益性事业单位，要给予经费保证”、“大中城市要重点建设好图书馆、博物馆”。因此这些公益性文化设施应以政府投入为主建设，不宜搞市场操作。

（二）政府和社会力量共同出资，按照市场经济规律和机制来

兴办文化设施。文化基础设施建设要走出原来文化局办文化的管理模式。如上海大剧院由广电集团管理经营;广州市政府划拨的市中心一块 70 亩土地,报业集团要建造一个三万平方米的图书馆和两万平方米群艺馆作置换。当然,政府要给他们政策上的支持。

(三)以引进外资和利用社会资金为主,建设一批上档次的文化娱乐设施,如歌舞厅、娱乐城、电影院和文化广场等。可在充分论证的基础上,向社会公布一批文化设施项目,通过招投标的形式,鼓励社会各界共同参与项目建设经营。

(四)以政府支持和文化部门自筹相结合为主,改造一批原有的文化设施,如专业剧团、文化馆、文保点等。我市专业剧团走场团合一道路,应该说是一种有益的尝试,使文化单位发展文化产业成为可能,增加了事业发展的后劲。对此,政府尚需作进一步的扶持和推动。

(五)以县(市)、区和乡(镇)政府投资为主,建设一批合理分布城乡的公益性文化设施,如文化馆(站)、图书馆(室)等。我们了解到深圳的村级图书馆,已达到 200 平方米、2 万册藏书标准的有 67 个,今年准备再达标 50 个。而我市乡镇图书馆真正能达到这个水平的只有三、四家。相比之下,我市的农村文化建设任重道远。

省委、省政府提出建设“文化大省”号召已经半年多了,我市为迎接第七届中国艺术节建造宁波大剧院已成定局,但已列入《政府工作报告》的图书馆、博物馆的建设任务不能因此而耽搁。我们相信,标志性文化设施的建设一定能给宁波的城市面貌和城市形象带来新的巨大变化。我们的后代一定会记住为宁波城市文化基础设施建设作出贡献的有功之臣,包括我们的决策者和建设者。

(2001 年 7 月)

数字图书馆中的知识发现

韩惠琴 刘柏嵩

1 引 言

数字图书馆的功能是组织数字化信息及其技术进入图书馆并提供有效服务^[1]。从根本上说,数字图书馆是为了解决网络环境下数字化信息的组织、查询与服务问题。我们处在一个信息爆炸的时代,因特网上信息资源呈几何级数增长,导致几个问题出现:(1)信息过载,各种信息太多,难以消化;(2)由于信息过多,用户难于寻找所需的信息,资源迷向;(3)网络环境下使用各种不同的语言,而传统的信息检索系统几乎都是基于单语言环境;(4)信息更新速度较快,具有较强的实时性。在这种环境下,如何将相关信息实时提供给用户?

目前数字图书馆是大量的数字化馆藏、数据库远程访问、全文 Web 资源链接和文档提交的结合体(也称为 Hybrid Library)。它希望提供类似于传统图书馆的良好的信息服务,如有序化的组织、结构化存储、高效的检索等。但是在网络环境下,数字化收藏并不具备这些特点。如何组织数字图书馆中的数字化信息,如何提供有效服务,这是当前数字图书馆研究的难点之一。传统的信息检索或信息访问,主要是根据用户提供的查询条件从文档数据库中检索出相关的文档信息,已完全不适用于数字图书馆环境。面对“被数据淹没,却饥饿于知识”的挑战,基于人工智能的数据开采和知识发现技术应运而生,并被广泛应用。本文针对数字图书馆特

点,提出几种将知识发现技术应用于数字图书馆的信息发现和信息提供的方案。

2 知识发现过程和特点

知识发现(Knowledge Discovery)是近几年兴起的一个极有发展前途又有广泛应用前景的新领域,其研究的主要目标是采用有效的算法,从大量现有或历史数据集合中发现并找出最初未知、但最终可理解的有用知识,并用简明的方式显示出来。发现了知识可用于信息管理、查询优化、决策支持等,还可以用于数据自身的维护。目前人民大学的王珊教授和电子系统工程研究所的李德毅院士等国内专家在知识发现方面作了大量卓有成效的工作。而且,国际上知识发现在健康保险、电子商务、金融、科学数据分析等领域中已有许多成功的应用。当前,知识发现的对象也正从关系型或事务数据库的结构化转向半结构化和无结构化的领域,如超文本文档和多媒体数据等。

知识发现一般包含如下几个步骤:(1)理解相应的问题领域;(2)准备相关数据子集;(3)发现模式(数据挖掘);(4)所发现模式的后处理;(5)应用发现结果。

知识发现的主要技术和方法包括:人工神经网络技术、决策树方法、规则归纳法、各种聚类技术等。其主要知识形式为:数据间依赖性知识、分类规则、关联规则、聚类规则、描述性知识、概要性知识和变化与或偏差知识等。

知识发现技术一开始就是面向应用的。它不仅要完成特定数据的检索、查询、调用,而且要对这些数据进行统计、分析、综合和推理,以指导实际问题的求解,发现一事件和其他事件的相互关联,甚至利用已有的数据对未来的活动进行预测。如美国 NBA 的教练,可利用 IBM 公司提供的知识发现工具决定临场如何替换队员。需要说明的是,这里所说的知识发现,不是要求发现放之四海

而皆准的真理,也不是要去发现崭新的自然科学定理和纯数学公式,更不是计算机定理证明。所有发现的知识都是相对的,是有特定前提和约束条件且面向特定领域的。如在数字图书馆中,根据用户访问的模式、学科范围等,及时提供给用户相关的信息,就是知识发现的一种形式。

3 知识发现在信息发现中的应用

3.1 Web 数据开采

数字图书馆的核心是信息,除了来自本馆收藏的电子文献外,目前大部分信息来自 Internet。对 Internet 的数据开采是数字图书馆信息的一个重要来源。Internet 上数据源是无序的、非结构的,数据之间存在冗余,数据源具有很强的动态性和多样性,这些特点导致了信息获取的困难。Web 数据开采就是针对这些问题而发展起来的一项新技术。其实现过程包括 Web 信息的采集、文档的识别与分类等。

一般来说,Web 信息的采集可利用已有的 Robot 技术。Web 数据开采的难点在于对采集后的信息进行文档自动分类和自动摘要。

(1)文档自动分类:文档分类是指根据文档的内容或属性,将大量的文档归到一个或多个类别的过程。其关键是构造一个分类模型,并利用此分类模型将未知文档映射到给定的类别空间。分类器的构造一般采用机器学习方法、神经网络方法等。国外对文档分类技术的研究已开展多年,并在邮件分类、电子会议、信息过滤等方面得到了较广泛的应用。国内对中文文档自动分类的研究主要有上海交大的王永成、南京大学的邹涛、山西大学的曹焕光等。其中南京大学研制的 CTDCS 系统,作为“面向电子报刊电子图书馆的网络出版发行系统”的一个子系统,目前已基本达到实用化水平。与此相关的还有关键词的自动抽取,用一个从文本中抽

取的关键词集合在一定程度上来代表文本的语义,然后借此进行分类。

(2)自动摘要:著名的搜索引擎 YAHOO 选取网页首部一定数量的句子作为摘要的代用品,中国经济信息网采用人工方法编制文摘。目前上海交大王永成教授采用直接从原文中选取字串生成摘要的方法^[4],其基本思路是:首先扫描全文,对不同标识符后面的文本字串赋予不同的权值。如对<TITLE>后面的文本字串赋予最高的权值,然后根据字串的属性信息来调整字串的权值。在文本字串长度大于一定阈值的情况下,对中英文字串采用不同的方式调整权值,然后根据权值大小挑选字串,生成文字流畅且具备一定质量的自动摘要。目前该系统原型已在中国经济信息网试用。

Web 数据开采有以下两种实现形式:

3.1.1 网络智能体

智能体是一个具有控制问题求解机理的计算单元,指模拟人类行为和关系、具有一定智能并能够自主运行和提供相应服务的程序。随着网络技术的发展,可以让智能体在网络中移动并完成某些功能。这就是网络智能体的思想。

与智能搜索引擎(3.4)和智能浏览器(4.6)相比,用于网络的智能体可在空闲时间工作,在不需要用户监督的情况下可以昼夜不停地运行。

一个成功的智能体系统是由 PC Magazine 推荐的 Browser Buddy,它是一个用于组织和链入 Web 页面的基于规则的智能体。经过整夜的信息搜索以后,它会在早上给用户提供一个服务清单。在它的导引下,用户可以迅速进入那些需要长时间交互才能到达的页面。它使用神经网络而不是关键词来识别信息的模式。用户使用类似自然语言的描述将智能体限制在一个概念区域中,然后智能体在用户的教导下不断训练,直到能有效地找到用户感兴趣

的文档。其核心是 Cambridge Neurodynamics 开发的动态推理引擎。它综合使用了神经网络与模糊逻辑,广泛地应用于网络数据挖掘与知识发现。

除了通用的智能体,目前还有一系列用于在 WWW 上寻找特定信息的专用智能体软件,如搜索在线新闻、CD 唱片、CAI 课件和教材信息等。

在数字图书馆的信息发现中,网络智能体可在以下方面改善信息服务的质量和水平:

(1)信息导引。根据用户的需求,代替用户寻找网上信息或告诉用户所需的信息到网上何处寻找。

(2)用户问题解惑。根据网上的信息资源解答用户提出的某方面问题。

(3)信息过滤。从网上信息中筛选出符合用户条件的信息,并以合适的形式(如全文、摘要、标题)呈现给用户;

(4)知识发现。从网上大量的原始数据中,挖掘出能反映其中规律的知识提供给用户。

3.1.2 智能信息捕捉器

另一种实现方式是智能信息捕捉器。它依赖于成熟的 Caching Web 技术,立足于某一领域的信息需求,以自动捕捉、采集和整理领域所需信息为目的,如记者自觉地搜集指定范围的信息,并整理汇编、发布或反馈信息一样^[11]。对应于机器人管理中的捕捉需求事务处理,在信息捕捉的整个过程中承担自动捕捉需求的分类设置、信息源确定和信息搜索路径选择。用户的一个捕捉信息需要按照领域知识规则进行自动分类后,建立相应数据库。根据定义的捕捉需求查找信息源登记表,然后根据信息源提供的信息粒度筛选信息源。信息源确定后,再根据模型算法,计算确定搜索路径,并自动优选最佳搜索路径,按逻辑式自动组织搜索路径,按逻辑式自动组织搜索关键字。可同时对多个特定领域内

的信息捕捉。其主要功能是过滤无用冗余信息、智能概念抽取、生成信息概要等。

3.2 多语种信息发现

要真正实现全球知识共享,就必须跨越地理和语言的障碍获取信息^[11]。多语种信息发现(Multilingual Information Discovery)已引起信息检索、自然语言处理、机器翻译、语音处理、人机交互多个领域专家的兴趣。目前主要进行多语种电子文档获取、集成机器翻译和多语种信息检索系统方面的研究。作为数字图书馆项目一个专门工作组,MIDAS已在以上三个方面取得阶段性成果。

3.3 跨学科协同检索

众所周知,大多数研究课题需要检索多个数据库并跨多个学科领域。虽然学术或商业机构为各个数据库提供相应的基本检索工具,但在操作简单的 Web 时代,用户需要去熟悉每一种检索界面和操作方法,显然不合时宜。需要通过 Web 环境将这些数据库集成起来,采用知识发现中的方法,提供给用户统一(All in One)的检索界面,并跨学科查询返回统一结果。

对网络信息和专用数据库的查询,由银盘信息公司(Silver Platter Information)提供的 Knowledge Cite Library^[12],其检索结果排列按六项标准进行:(1)位置,在主题位置则权重高;(2)词的临近度;(3)出版日期,越新则权重越高;(4)完整性;(5)词频;(6)词重。该系统特点是采用著名的 Alta Vista 作为系统核心,采用“主题同步”检索过程,检索结果的超文本链接采用美国国会图书馆主题标头方式。其功能是向近 70 个数据库并发请求,同时可与图书馆的公共检索(OPAC)、馆际互借、文档提交和电子资源相连,针对多任务向用户提供一个无缝(Seamless)检索界面。

而对于书目数据库查询,目前较理想的工具还是 Z39.50 客户系统,如 BOOKWHERE 2000 等。

3.4 智能搜索引擎

搜索引擎的目的是帮助用户寻找资源。在 Internet 环境下,其典型实现是基于关键词匹配的信息检索机制。据专家评测,目前主要的搜索引擎返回的相关结果的比率不足 45%,其原因是现有的 Internet 搜索引擎拥有极少量的知识,并且是面向最一般的用户模型。不划分知识领域,不对用户建立任何描述以及使用关键词匹配的交互方式,都限制了搜索引擎的使用效率。由此提出了智能搜索引擎的概念。智能搜索引擎可以预期用户的需求,并可有效地抑制关键词的多义性。如 Eastman、Kodak 都在尝试在它们的 Intranet 中使用更加智能化、知识化和专业化的搜索引擎。目前比较成功的智能搜索引擎有 FSA, Eloise, InFind, Compass Ware 和 FAQFinder。

Arthur Andersen 的 FAS 和 Eloise 专门用于搜索美国证券交易委员会的 Edgar 商业数据库。这两个系统中都内嵌了特定领域中的商业知识,并使用了推断-证明式的自然语言理解技术。InFind 是一个类似于元搜索引擎的产品,但支持自然语言、模糊检索,可将结果自动去重、聚类,并可返回对所有搜索引擎并发查询的结果,其检索结果根据分类排列。

另一个类似的产品是芝加哥大学人工智能实验室开发的 FAQFinder。它是一个具有问答式界面的智能搜索引擎。在获各用户问题后,它查询 FAQ 文件以给出比较合适的回答。FAQFinder 的内核由五个互联联系的技术环节构成:

(1)基于统计方法建立 FAQ 文件。这是由 FAQFinder 中的工具 Smart 完成的。

(2)用一个由简单名词和动词短语构成的文法树分析用户的查询,以得到一个用于支持内容匹配的描述。

(3)问题识别者(Question Recognizer)操作文法树以辨识问题从属的类别。

(4)使用语义网分析与概念匹配技术,找出与用户查询最近似