

面向信息检索的词汇知识发现

章成志 苏新宁

(南京大学信息管理系 南京 210093)

【摘要】 针对信息检索中词汇知识发现问题,提出面向信息检索的词汇知识发现框架结构,对面向信息检索的词汇的定义、同义词、相关词、排除词等词汇知识获取与挖掘进行详细的介绍,为后续的语义检索研究打下基础。

【关键词】 信息检索 中文信息处理 知识发现 同义词 相关词 排除词 **【分类号】** TP391 G252

Lexical Knowledge Discovery for Information Retrieval

Zhang Chengzhi Su Xinning

(Department of Information Management, Nanjing University, Nanjing 210093, China)

【Abstract】 This paper introduces the problem of lexical knowledge discovery for information retrieval and provides the frame of lexical knowledge discovery. The authors pose the method of words' definition extraction, synonym recognition, relevance words recognition, mutually exclusive words recognition. The work plays an important role for the semantic search.

【Keywords】 Information retrieval Chinese information processing Knowledge discovery Synonym words Relevance words Mutually exclusive words

1 面向信息检索的词汇知识发现总体框架

信息检索中涉及到的词汇知识发现主要包括:命名实体识别、词汇同义词关系的自动识别、文本主题概念或主题词的提取、词汇定义的提取、文本其他元数据的提取等,这些都可以归为信息提取范畴。信息提取完成后,为后续的文本挖掘,如文本摘要、文本篇章结构分析、文本分类等提供了基础。经过信息提取后,可以为信息检索用户提供基于内容的、多层次、不同颗粒度的检索信息。面向信息检索的词汇知识发现总体框架如图 1 所示。图 1 给出了每一类词汇知识发现的大致流程、方法和最终用途。下节将具体介绍其中 4 种主要词汇知识发现的方法。

2 面向信息检索的词汇知识发现方法

2.1 面向信息检索的词汇定义获取

随着网络的高速发展,信息资源也向海量方向不断发展,崭新的术语层出不穷,同时旧的术语概念也被赋予了新的含义。现在人们更加需要不断学习,充实自己,跟上时代的潮流。第一步往往就是要了解那些名目繁多的

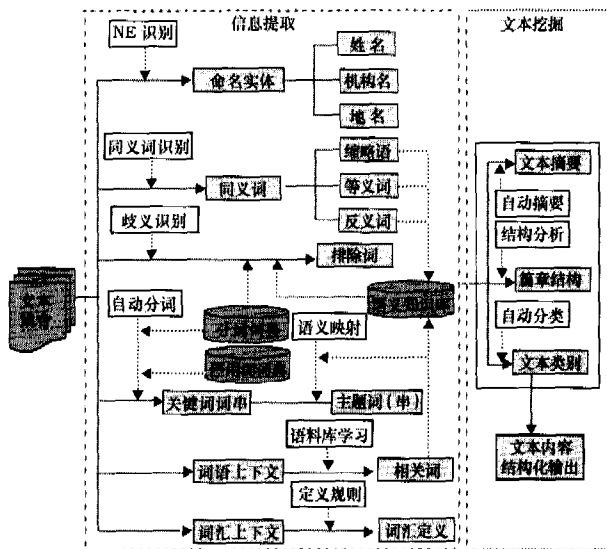


图 1 面向信息检索的词汇知识发现总体框架图

不断变化更新的术语概念,但是字典、词典又满足不了人们的要求,因为它们往往不能及时更新。所以在如此海量的数据中如何快速检索出所需要的术语概念就显得非常必要与重要。本节从定义获取工作流程和实际系统两方面来介绍词汇定义获取的主要方法。

(1) 定义获取工作流程

定义获取系统主要是针对诸如“What is...”“Who is...”的提问从信息库中提取相关定义的一种知识获取系统。其一般的工作流程如下:

①获取与提问有关的信息库。当系统获得用户提出的诸如“What is...”“Who is...”的问题时,定义获取系统需要从特定领域的语料库和/或 Web 文本中查找到相关信息,组成一个信息库。

②提取出定义的一般模式。从特定领域出发归纳出该领域中定义或概念的一般模式,对这些模式进行优化。这是定义获取系统关键的一步,因为只有把这些模式建好后才能提高系统性能和在接下来的抽取定义过程中提高提取到的定义的准确性。

③提取“准定义”。使用②所得出的模式,对信息库进行定义抽取工作,得到一些“准定义”。

④优化整理“准定义”。根据定义的一般特征对准定义进行优化。通过上面的步骤,往往得到一系列关于用户提问的“准定义”,在这一步,系统应该将这些“准定义”进行合并、重组、优化整理得出我们所需的定义。

(2) 现有的定义获取系统模型

现有的定义获取系统模型主要有:

①北京工业大学的基于互联网的术语定义获取系统^[1]

该系统是一个实验性的基于互联网的术语定义获取系统,可以从互联网上查找术语的定义以及与定义有关的内容,方便用户获得新术语以及新技术词汇的定义方面的知识。该系统主要由网页下载和术语定义处理两部分组成。网页下载部分是接受用户查询的术语,然后将其提交给一个商用搜索引擎,将返回的结果页面下载下来,对其进行预处理作为术语定义处理部分要处理的网页文本数据。在术语定义处理部分,系统依据的是一组术语定义模式,称为“正面”模式,对第一部分所得的网页文本进行定义提取工作,系统中使用的术语定义模式是从一定量的科技期刊语料库中人工归纳获取的。同时系统还提供了几个排除模式,用于对提取出的非定义进行排除。从其试验结果看,该系统的运行效率比较高,结果的准确度令人比较满意。但是其在如何消除由正面模式得出的非定义所提供的排除模式过于简单,不利于系统投入实际应用。并且定义提取往往需要注意针对一定领域特定信息的特点,进行模式的识别提取,这样才能够进一步提高系统的正确率,提高系统的性能。

②WebLearn 系统^[2]

该系统主要由5部分组成:搜索引擎,这是一个标准的 Web 搜索引擎;网络爬虫,它将搜索引擎检索到的 Web 网页下载下来,存贮至“Web 页存储池”;显著概念挖掘部分,该部分利用特定的技术从“Web 页存储池”分辨提取出次主题或显著概念;定义提取部分,系统利用给定的模式对“Web 页存储池”中收集的信息进行抽取工作以发现这些网页信息中有关主题、次主题的概念定义;用户界面部分,该部分方便用户与系统交互。在该系统的定义获取部分它使用了一套定义模式并且尝试利用了 HTML 语言中标记标题的一些符号规则进行提取定义。虽然它提出了利用 HTML 语言帮助提取定义,但是其利用的定义模式过于简单,不利于准确的提取定义。

③MSRA 的定义获取排序系统^[3]

提供一个术语,以期获得该术语的定义,该系统使用 Google 引擎提供的搜索定义的功能检索出有关该术语的候选定义,并且根据这些

检索出的定义好坏程度对这些定义进行排序。它不同于传统检索将检索出的定义全部列出来的形式。该系统使用 SVM 作为分类模型, Ranking-SVM 作为顺序的回归模型,根据定义的好坏对候选定义进行排序,再将多余的即不合格的候选定义除去,从而得到符合用户需求的定义。

2.2 面向信息检索的同义词识别

同义词关系为一种特殊词汇语义关系,在信息检索中,通过同义词可进行扩展检索。因此,同义词识别是一项有意义的工作。同义词识别的方法有基于规则的方法、基于互引的方法、基于词语释义向量相似度计算的方法、基于推理的同义词扩张方法、基于混合策略的同义词识别方法等。

(1) 基于规则的方法

利用同义词定义模式解释一个词语,即:同义词定义就是用与一个与被定义项具有相同意义的另外一个词来进行词汇的定义。同义词定义模式含有一些特定的标志词语,例如:“亦称……”,“也称……”等等。

(2) 基于互引的方法

基于互引的同义词识别方法是利用词典中对词汇的注释来计算词汇之间的语义相似度,然后根据语义相似度的大小来识别同义词。它的假设是:如果两个词是同义词,那么在它们的定义中必定含有很多相同的词汇;如果两个词是同义词,那么它们必定经常被同时用来定义同一个词汇。

(3) 基于词语释义向量相似度计算的方法^[4]

将词语的释义语句转换为释义向量,利用向量间的欧氏距离或者向量夹角的余弦值作为词语间相似度,将相似度超过给定阈值的两个词作为候选同义词。

(4) 基于推理的同义词扩张方法

可细分为:利用同义词间存在的性质(对称性、传递性)进行同义词的扩张;基于循环衍生方法的同义词扩张方法^[5]。

(5) 基于混合策略的同义词识别方法

单纯一种方法可能造成识别出来的同义词的置信度不高,结合多种方法来进行投票,保证同义词识别的质量。

2.3 面向信息检索的相关词识别

在信息检索中,由于用户使用的自然语言通常不能考虑到所使用词汇的相关词,很容易造成信息漏检或误检,降低了检索效果,解决这个问题的途径之一便是查询扩展或相关词提示。应用查询式扩展查询扩展或相关词提示可以辅助用户正确表述信息需求,降低信息用户智力负担。此外,在用户查询式的基础上提供相关词,通过检索式的重新构建可以进一步完善检索式,达到扩检和缩检的效果。按照所依据的资源,相关词识别方法可以分为如下三种:

(1) 基于语料库的相关词识别提取方法

从语料库中挖掘相关词,是基于词频共现原理,共现窗口

选定为文档中的自然段落。从大规模语料库中提取相关词,可以转化为依据统计值判断词汇间的紧密程度,如果统计值表明两个词汇间具有很高的紧密度,那么就可以认为它们是相关词。一般来说,两个特征词在语料库中共同出现的频率越高,两者相关性也越高,但是,互信息不仅仅与共现频次有关系,还受语料库大小的影响。词汇间互信息的大小反映了词汇间关联紧密程度,互信息值(MI)越大,关联越紧密,是相关词的可能性越大^[6]。设定一个阈值 R ,若 $MI(w_i, w_j) \geq R$,则 w_i 和 w_j 进入相关词候选集。通过调整阈值 R ,可以限制相关词对的个数。

(2) 基于释义词典的相关词识别方法

将释义词典数据进行数据净化和数据标准化后,形成规范的文本格式,每一行表示一个词条。本实验系统基于释义词典的相关词挖掘算法,只基于第一个假设,为每个词对构建词汇向量空间模型。在构建的向量空间模型中,词汇释义中的每个词作为一个维,用来描述词汇空间。接下来便是判断两个向量之间的相关性,判断相关性的方式很多,这里选择了计算词汇间的 Cos 值^[4]。在向量空间中, Cos 值越大,说明两个向量之间的夹角越小,也就是两者相关的可能性越大。因此,给定阈值 Q ,若 $\text{Cos}(w_i, w_j) \geq Q$,则 w_i 和 w_j 进入相关词候选集。同样,通过调整阈值 Q ,可以限制相关词的个数。

(3) 基于用户查询日志的相关词识别^[7]

以搜索引擎的用户搜索日志作为实验对象。一般同一用户提交的检索词之间更容易具有相关关系,也就是说,如果两个检索词经常共现于一个查询式集合中,那么就可以认为它们是相关的。可以把查询式集合看作段落,将所有用户的查询式作为语料库,借用词共现原理,计算词汇间的互信息,从中抽取相关词对。

(4) 基于混合策略的相关词识别

经过以上三种方法识别出来的相关词需进一步整合,才能应用到实际的检索系统中。相关词识别结果的整合可以通过直接整合或加权整合两种途径实现:

①直接整合:充分信任每种知识源,对于两个词汇,只要有被某个知识源推荐为相关词,则将其收入相关词表。

②加权整合:即给每种知识源附以一定的权重,然后依据组合值对词重新排序,设定阈值,组合值大于阈值的才提取为相关词。

直接整合的优势在于方法简单,便于操作,但是缺少对识别准确度的进一步限制。

笔者利用语料库、释义词典、用户检索日志作为识别相关词的语境,设计并实现了相关词自动提取系统。分析实验结果发现,虽然面向相同的基本词汇集合,但是基于不同语境提取的相关词之间的重复率很低,各个结果间的互补性很强,因此,结果整合非常有必要。在本系统中,通过直接整合途径得到了最后的相关词词表。具体内容见文献[8]。

2.4 面向信息检索的排除词识别

中文自动分词是中文信息处理面临的两大难题,是切分歧义和未登录词问题。其中,交集型歧义切分字段又占全部歧义切分字段的绝大多数^[9]。有研究者将消歧算法用于构建消歧实例库,在自动分词时调用切分实例或规则进行中文的分词^[9,10]。这些方法对提高中文信息检索的质量起到了一定的提升作用,如在全文索引时,降低索引膨胀率,类目在信息检索中的缩检和扩检作用^[10]。由于篇幅限制,本节直接给出几个相关概念和定义。

(1) 几个相关概念

先给出几个相关的说明: Ω :中文字符串集合。 S :中文字符串, $S = c_1 c_2 \dots c_n$,即 $S \in \Omega$ 。 Ψ :分词词典,用于中文分词。 T :已分词的训练语料。

定义 1:交集型歧义切分字段。对于字符串 $S, S = c_1 c_2 \dots c_n, S \in \Omega, S = c_1 c_2 \dots c_n$ 为汉字,如果存在整数 $i_1, i_2, \dots, i_m, j_1, j_2, \dots, j_m (m \geq 2)$,满足:

① $S \notin \Psi$;

② $w_1 = c_{i_1} \dots c_{j_1}, w_2 = c_{i_2} \dots c_{j_2}, \dots, w_m = c_{i_m} \dots c_{j_m}$,且 $w_1, w_2, \dots, w_m \in \Psi$,且 S 中不存在包含 w_1, w_2, \dots, w_m 的词;

③ w_1, w_2, \dots, w_m 构成相互交叉,即: $1 = i_1 < i_2 \leq j_1 < j_2, i_2 < i_3 \leq j_2 < j_3, i_3 < i_4 \leq j_3 < j_4, \dots, i_{m-2} < i_{m-1} \leq j_{m-2} < j_{m-1}, i_{m-1} < i_m \leq j_{m-1} < j_m = n$;则称字符串 S 为交集型歧义切分字段(Overlapping Ambiguity String, OAS)^[9]。

定义 2:伪交集型歧义切分形式。对于给定的交集型歧义切分字段 S ,对于某一形式“ $c_1/\dots/w_i/\dots/c_n$ ”,其在 T 上的出现概率,即: $p(c_1/\dots/w_i/\dots/c_n | T)$,低于给定阈值 θ ,则称该切分形式,即:“ $c_1/\dots/w_i/\dots/c_n$ ”为 S 在 T 上的伪交集型歧义切分形式,其中 w_i 为伪交集型歧义切分形式下的词语之一。

定义 3:准交集型歧义切分字段。对于交集型歧义切分字段 S ,若 S 为复合词,且 S 可加入到词表 Ψ ,此时, S 将不再是严格定义下的交集型歧义切分字段。则将 S 称为准交集型歧义切分字段(Quasi Overlapping Ambiguity String, QOAS)。

定义 4:排除词。对于给定的准交集型歧义切分字段 S ,存在伪交集型歧义切分形式,即 $c_1/\dots/w_i/\dots/c_n$,则称 w_i 与 S 构成排除关系,即 w_i 与 S 互称排除词(Mutually Exclusive Words, MEW)。例如,“电脑科学”为准交集型歧义切分字段,存在伪交集型歧义切分形式“电/脑科/学”,因此“脑科”与“电脑科学”构成排除关系,即,“脑科”为“电脑科学”的排除词。

(2) 排除词识别方法

前面提到,只有当 S 被认定为复合词时,交集型歧义切分字段 S 与 w_i 才有可能构成排除关系。在没有上下文约束的情况下,复合词 S 的正确切分形式不包括切分为“ $c_1/\dots/w_i/\dots/c_n$ ”的情形,例如,“发展中国家”可以切分为“发展中/国家”,“发展中”对“国家”起限定作用,但不可切分为“发展/中国/家”。在有上下文约束的情况下,复合词 S 有可能与周围

的字符再次发生新的交叉关系。例如,句子“这反映了我国的 R&D 活动在发展中国家居中等水平”,其中,“发展中国家”与后续字符“居”、“中”、“等”再次交叉,构成交集型歧义,即“发展中国家居中等”也是交集型歧义切分字段,其正确切分形式为“发展中国家/居/中等”。而在句子“致力于把个性化高品质的设计和一流的工业体系配套有效结合来发展中国家居产业”中,“发展中国家居”为交集型歧义切分字段,其正确切分形式为“发展/中国/家居”。由此可以看出,在有上下文约束下,即真实文本中,只有当包含准交集型歧义切分字段 S 的句子才能切分出复合词 S,才能进一步进行排除词的识别。若该准交集型歧义切分字段与周围字符发生交叉,即成为交集型歧义切分字段,则需要进行歧义消解处理。若该准交集型歧义切分字段 S 不再与任何字发生新的交叉关系,则称该准交集型歧义切分字段为最大准交集型歧义切分字段,笔者借用最大交集型歧义切分字段的定义 (Maximal Overlapping Ambiguity String, MOAS)^[9],给出准交集型歧义切分字段的定义。

定义 5:最大准交集型歧义切分字段。设 $S = c_1 c_2 \cdots c_n$ 为准交集型歧义切分字段,满足:

① $S_{\max} = c_i \cdots c_j (1 \leq i < j \leq n)$, 且 $S_{\max} \in S$;

② S_{\max} 为交集型歧义切分字段;

③ S 中不存在包含 S_{\max} 的更大的交集型歧义切分字段;则称 S_{\max} 为 S 的最大准交集型歧义切分字段 (Maximal Quasi Overlapping Ambiguity String, MQOAS)。

例如,在句子“论法国会计模式对中国会计制度改革的借鉴意义”中,“法国会计”为准交集型歧义切分字段,“法国会”为交集型歧义切分字段,但“法国会计”涵盖了“法国会”,同时不为任何交集型歧义切分字段所包含,因此,“法国会计”是最大准交集型歧义切分字段。

识别最大准交集型歧义切分字段的意义在于:由于最大准交集型歧义切分字段不再与周围任何字符发生新的交叉关系,具有一定的独立性,因此,可以将它们从上下文环境中分离出来,直接作为排除规则,加入到排除词词典中,从而实现前面所提到的功能。

根据 MQOAS 的来源不同,将 MEW 识别分为两类,即基于关键词词典的 MEW 识别方法与基于语料库的 MEW 识别方法。

①基于关键词词典的 MEW 识别方法

该方法以关键词词典为基础。首先,依据语料库从关键词词典中识别出所有 MQOAS,即识别出的 MQOAS 全部为关键词词典中的词语,在这里,笔者利用全切分的方法进行 MQOAS 的识别;然后,对每个 MQOAS 在分词训练语料 T 上检索该 MQOAS,若 T 上存在该 MQOAS,则进行切分形式统计,对于其中某一切分形式“ $c_1/\cdots/w_i/\cdots/c_n$ ”,其在 T 上的出现概率即: $p(c_1/\cdots/w_i/\cdots/c_n | T)$, 低于给定阈值 θ , 则认为 w_i 与该 MQOAS 构成排除词关系,并保存该结果;若在 T 上不存在该 MQOAS,则进行手工切分,人工判别得到排除词,并保存结果。

②基于语料库的 MEW 识别方法

该方法以语料库为基础。首先依据分词词典,在中文语料上利用全切分方法识别出所有最大交集形歧义 MOAS;然后对每个 MOAS 判别当前 MOAS 是否为复合词,在这里,复合词的判别可以人工判别或采用统计方法进行辅助判别;若为复合词,则在分词训练语料 T 上检索该 MOAS,若 T 上存在该 MOAS,则进行切分形式统计,对于其中某一切分形式“ $c_1/\cdots/w_i/\cdots/c_n$ ”,其在 T 上的出现概率即: $p(c_1/\cdots/w_i/\cdots/c_n | T)$, 低于给定阈值 θ , 则认为 w_i 与该 MOAS 构成排除词关系,并保存该结果;若在 T 上不存在该 MOAS,则进行手工切分,人工判别得到排除词,并保存结果。

笔者对《经济日报》1983 - 2003 年语料进行了缩检测试。测试方法为:从排除词词典中随机选取 100 条排除词记录,将排除词作为检索词,如“本体”,进行全文检索,考察缩检效果。测试结果为:缩检前,这 100 个查询词返回的文档总数为 425 350 条,平均返回文档数为 425 条,但缩检后,返回文档总数变为 28 753 条,平均返回文档数变为 29 条,参见图 2。很明显,排除词起到了明显的缩检作用,提高了检准率。

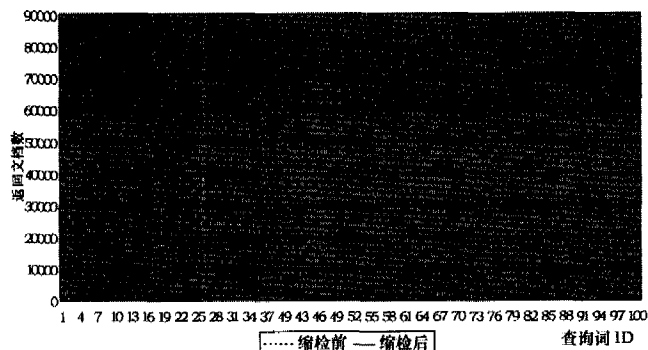


图 2 查询词缩检前后返回文档数

3 结 语

本文提出在传统的词语匹配的基础上,进行初步语义匹配的尝试,即:针对智能信息检索中词汇知识发现问题,提出面向智能信息检索的词汇知识发现框架结构,对面向智能信息检索的词汇的定义、同义词、相关词、排除词等词汇知识获取与挖掘进行详细介绍,为后续的语义检索研究打下基础。后续的工作包括:进一步扩大词汇知识发现的范围;将挖掘出的各种词汇知识整合到实际的信息检索中进行测试;利用文本挖掘和自然语言处理技术进行本体构建的尝试;在进行词汇的定义、同义词、相关词、排除词等词汇知识获取与挖掘后,将这些知识进行整合,用于信息检索系统当中。

参考文献:

- 1 许勇, 荀恩东, 贾爱平, 宋柔. 基于互联网的术语定义获取系统. 中文信息学报. 2004, 18(4): 37 - 43
- 2 Bing Liu, Chee Wee Chin, Hwee Tou Ng. Mining topic - specific

- concepts and definitions on the web. In: Proceedings of the 12th international conference on World Wide Web, Budapest, Hungary, 2003:251-160
- 3 Xu J, Cao Y, Li H, Zhao M. Ranking definitions with supervised learning method. In: Proceedings of 14th International Conference on World Wide Web, Chiba, Japan, 2005:811-819
 - 4 Salton G, McGill M J. Introduction to Modern Information Retrieval. New York: McGraw-Hill Book Co., 1983
 - 5 王云. 查找同义词和相关词的循环衍生法. 核情报工作与研究, 1997(2):7-8
 - 6 罗盛芬, 孙茂松. 基于字串内部结合紧密度的汉语自动抽词实验研究. 中文信息学报, 2003, 17(3):9-14
 - 7 崔航, 文继荣, 李敏强. 基于用户日志的查询扩展统计模型. 软件学报, 2003, 14(9):1593-1599
 - 8 章成志, 苏兰芳, 苏新宁. 基于多语境的相关词自动提取系统的设计与实现. 现代图书情报技术, 2006, (9):23-28, 80
 - 9 孙茂松, 左正平, 邹嘉彦. 高频最大交集型歧义切分字段在汉语分词中的作用. 中文信息学报, 1999, 13(1):27-34
 - 10 Mu Li, Jianfeng Gao, Changning Huang et al. Unsupervised Training for Overlapping Ambiguity Resolution in Chinese Word Segmentation. In: Proceedings of the Second SIGHAN Workshop on Chinese Language Processing, Sapporo, Japan, 2003:1-7
 - 11 苏新宁, 杨薇. Web 环境下全文检索系统开发平台的设计. 情报科学, 2002, 20(4):402-404, 415
- (作者 E-mail: zcz51@citiz.net)

《现代图书情报技术》特邀专栏组稿

《现代图书情报技术》是中国科学院主管、中国科学院文献情报中心主办的计算机信息管理技术方面的学术性刊物。刊物拥有清晰的定位,即以跟踪技术的研究、应用、交流为主体,服务于广大信息技术人员。

本刊从2004年起开设不定期栏目——《特邀专栏》,每一期专栏集中发表关于某个特定方面的技术研发与应用的研究型文章,汇集科研成果、聚焦研究前沿。

1 《特邀专栏》目的与定位

对于学术期刊而言,高质量的稿件始终是刊物发展的关键所在。因此,编辑部在广泛组稿的同时,也希望透过业界专家的支持,合作策划重大选题,集中组织优秀稿件,系统深入进行报道。

2 《特邀专栏》操作办法及流程

(1)本栏目特邀国内外知名专家、学者、教授担任专栏主编,专栏的设立一般由期刊的策划编辑和特邀专栏主编沟通,根据国内外图书情报技术学科的发展需要提出选题。

(2)选题一旦确定后,由特邀专栏主编承担稿件的组织,审核并撰写前言。一期特邀专栏一般为3-5篇文章为宜。稿件组织过程中,策划编辑将与特邀专栏主编进行定期的沟通,及时掌握稿件的撰写情况,并对稿件的撰写提出适当的建议和意见。

(3)稿件经特邀专栏主编审核通过,提交给编辑部。后期由策划编辑负责与作者的联系沟通及安排出版等事宜。

(4)专栏的选题一旦确定后,将确定基本时间表。一般的操作周期为3-5个月。以正式确定特邀专栏题目为起始点,在1个月内确定约请论文的作者和题目,3个月内确定初稿,5个月内确定采用稿。

(5)对于拟定录用的特邀专栏稿件,本刊将减免发表费,并支付稿费。稿件一旦发表,编辑部将及时赠与样刊。

3 《特邀专栏》稿件内容要求

(1)深入反映本专栏选题方向的前沿研究成果或重大应用成果,侧重理论研究、技术分析、系统论证或设计等,注意理论与实践相结合。

(2)特邀专栏稿件应该主要是原始性和原创性研究论文,也可以有一篇综述性论文,但综述性论文必须可靠地覆盖该方向的原始核心文献。

(3)文章按照严谨的学术文章体例写作,即明确扼要地界定研究问题,简要说明研究方法,系统精炼地描述国际国内发展状况,进而详细地描述作者自身研究工作的技术线路及研究结果。

(4)特邀专栏的一系列文章应注意覆盖专栏选题所涉及的各个研究方向和多个研究单位,充分覆盖可能存在的多种观点和技术线路。

(5)充分承认前人/别人的工作,充分引证所参考引用的文献(尤其是本研究工作中的原始核心文献和国内最先出现的研究文献),严格遵守著录规范。

4 《特邀专栏》稿件格式要求

(1)论文版式请参照本刊网站“下载专区”中“论文模板”。

(2)多个作者时,请注明通信作者,并注明各个作者的单位。

(3)每篇稿件以6-8千字为宜(按篇幅字数计算,包括图表)。

2007年本刊《特邀专栏》的组稿工作已开始启动,欢迎广大专家、学者给予支持、帮助!