对电子环境下主题控制系统检索应用的思考

北京大学信息管理系 马张华

讨论内容

在文本检索、关键词检索系统迅速发展的情况下，基于主题控制词表的检索系统还有没有价值？目前的主题标引规则、方法应如何发展，以及研究动向等。

主题控制系统检索应用的思考

一、对词汇控制系统的重新审视

二、控制系统与文本系统性能比较

三、文本检索系统电子环境下的应用以及词汇控制系统差距

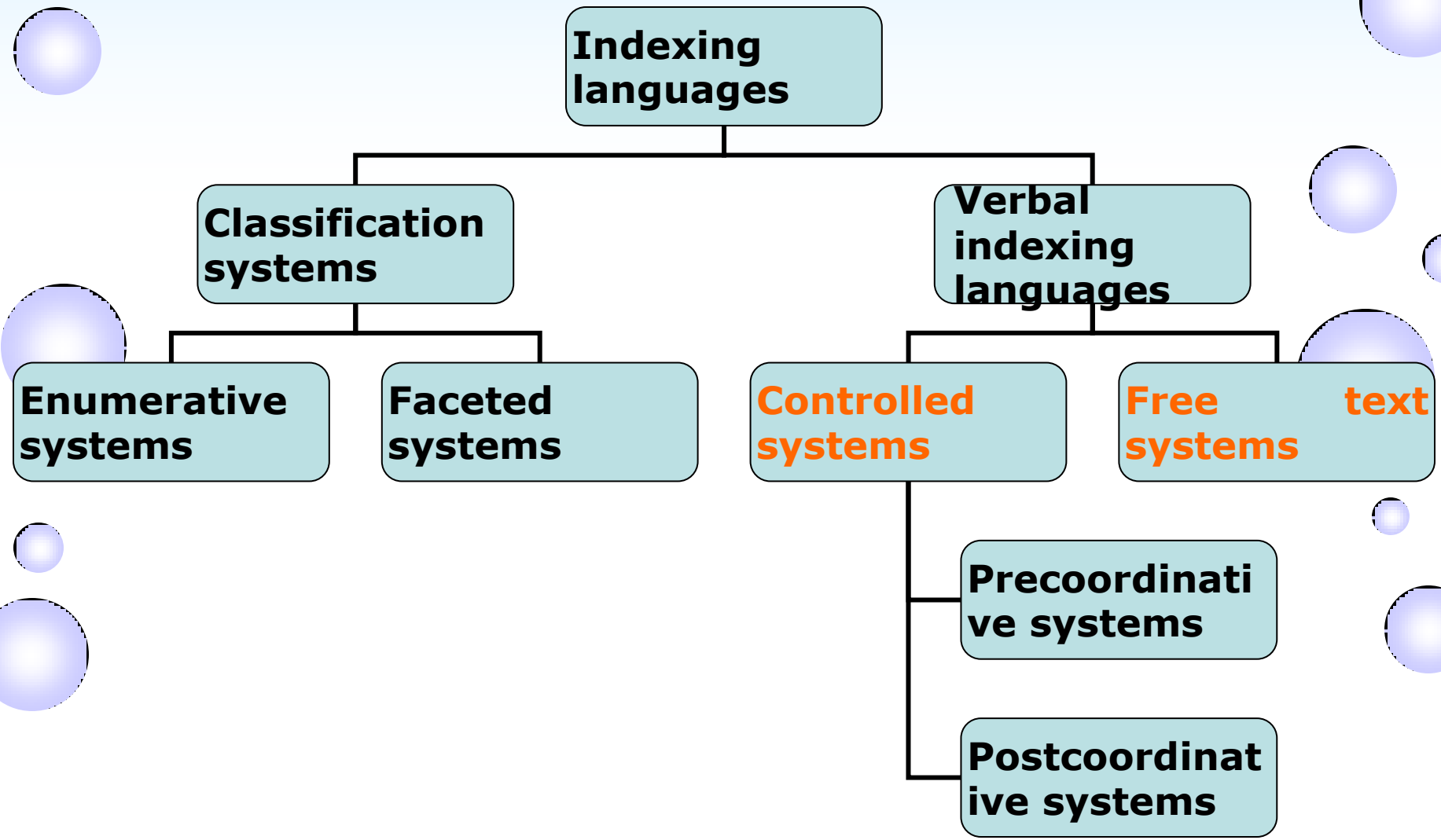
四、关于主题控制系统标引实践和规则改进的思考

一、对词汇控制系统的重新审视

1.1 自然语言特点及其控制的必要性；

1.2 主题分析基础上的标引。

1.1 自然语言特点及其控制的必要性
--简要的标引语言类型区分



1.1自然语言特点及其控制的必要性

词汇控制指根据标引和检索的需要，对自然语言的词汇进行选择、规范并揭示其相关性。原因：

- 词汇量过大——一些词无标引价值。（控制方式：选词）
- 词汇与概念不一一对应：（控制方式：参照、限定、加注）
  - 一义多词：计算机、电子计算机、电脑
  - 一词多义：病毒—医学、计算机
  - 词义含糊：计算机分析—分析计算机、用计算机分析？
- 缺乏明确的结构——自然语言词汇之间关系的多元性和不确定性，不符合检索系统的使用特点。（控制方式：建立参照、多种索引系统）

1.2 主题分析基础上的标引

- 通过主题分析弄清文献有标引价值的主题，有效揭示文献主题内容；
- 结合主题分析的结果按照检索语言及其标引规范，加以标识，有助于提供适用的标识。

对词汇控制系统的重新重新审视

对词汇控制系统的重新审视：

- 1.1自然语言特点及其控制的必要性；
- 1.2主题分析基础上的标引。

常识判断：词汇控制和标引有助于有效揭示和检索文献主题。

二、控制系统与文本系统性能比较

控制语言有没有价值，或词汇控制是不是必要：

- **两者的功能讨论。**检全率、检准率，处理速度、易用性、成本效益。各自的问题。
- **两者对于文献的适用性讨论。**网络资源，论文资源，图书。使用现状。

两者的功能讨论。各自的问题。

- 检全率
- 检准率
- 处理速度
- 易用性
- 成本效益

在一些方面中是相对与互补的关系：如检全率，控制系统可以进行概念检索、相关词扩展，文本系统的标识量大，均是检全因素。应结合具体应用讨论。

比较项目 \ 语言	受控语言	自然语言
检全率	高	低
检准率	低	高
扩检、缩检和改变检索范围	易实现	难实现
检索人员负担	轻	重
面向用户能力	差	好
标引和检索的匹配性	好	差
标引速度	慢	快
标引成本	高	低
对标引人员要求	高	低
专指性	差	好
标引一致性	差	好
自动标引	难实现	易实现
词汇更新	慢	快
词表编制与维护	有	无

两者功能的比较与思考：功能的相对性与互补性

控制系统与文本系统比较—功能的相对性与互补性

	控制系统	文本系统
检全率	词汇控制增强检全率 标引深度小，降低检全率+	缺乏词间关系控制降低检全率 标引深度大增强检全率
检准率	主题分析基础上标引提高检准率 标引深度小，提高检准率+	全文处理降低检准率 用结合多种因素控制排序提高检准率
处理速度	手工标引及时性弱 (可采用自动标引、人机结合标引改进)	自动处理及时性好+
易用性	可利用语言系统提供检索帮助+	缺乏词汇控制不利于易用性
成本效益	标引费用大、检索费用小（小系统中适用）	标引费用小、检索费用大（大系统中适用）

两者对文献适用性的讨论

- 不同领域应用的情况:
  - 图书。控制系统与相关字段文本检索结合；
  - 论文资源。文本检索是主流，结合部分控制系统；
  - 网络资源。文本检索是主流。
- 影响控制使用的因素主要包括：资源数量与处理能力，成本效益的结合考虑等。

概要结论

- 控制有益于检全、检准、易用性。不利于，处理速度、输入成本。影响控制使用的因素主要包括：资源数量与处理能力；成本效益的结合考虑等。
- 文本系统应引入控制；控制系统则应加强处理能力，降低成本，应用好控制系统的功能。

三、文本检索系统电子环境下的应用以及词汇控制系统差距

- 文本检索系统的改进努力
- 词汇控制系统的努力与差距

文本系统的改进努力：控制的纳入

检索方式：提供简单检索、高级检索、专业检索等检索界面。上述方面文本系统略优；努力提供自然语言检索能力。（引入句法控制、词汇控制）

检索排序：多因素结合提供；多种排序方式的采用；两者差距不大。（引入多因素控制）

检索优化：相关检索帮助，百度，Ask；二次检索；Vivisimo（引入词汇控制）

检索入口的改进，以搜索引擎为例:自然语言检索，还不是智能检索

检索排序显示的发展与改进，以网络为例

- **排序显示的意义：**是提高检准率的重要手段。
- 采用检索匹配加权的形式进行排序显示，可以在保障检全率的同时，将符合检索要求的对象排列在检索结果的前列，提高检准率。

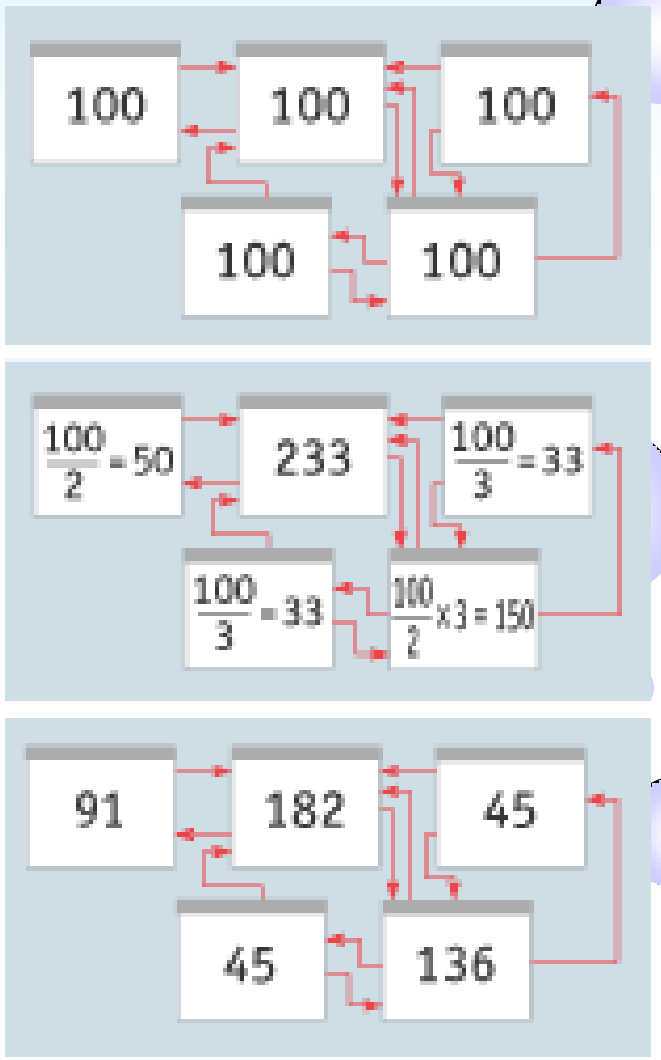
网络系统中检索排序因素的拓展

目前采用作为排序依据的加权方案涉及的因素包括：

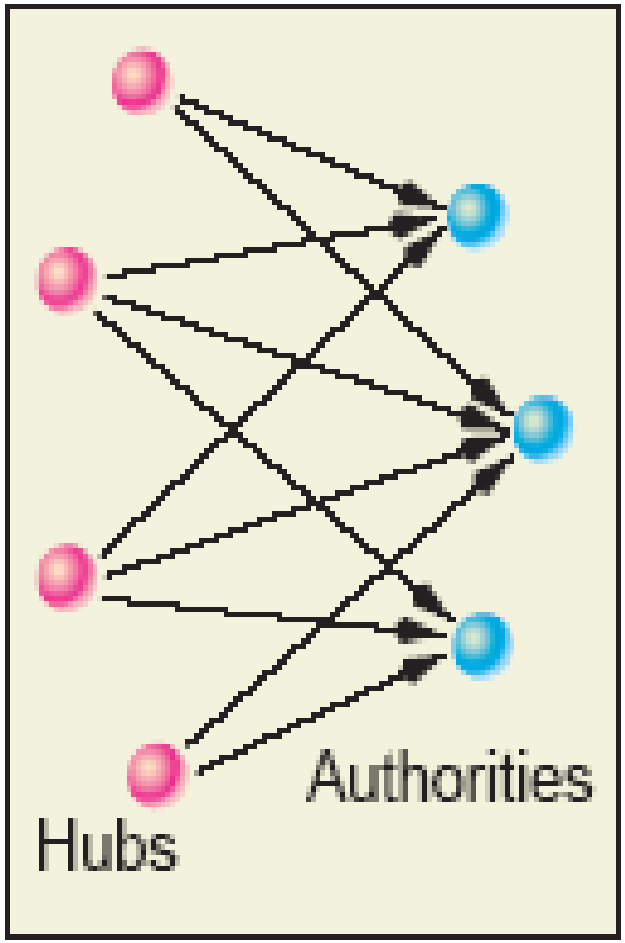
- –网页中查询词匹配数量
- –网页中多个查询词匹配的完备程度
- –匹配单元和分解问题
- –匹配词的接近程度
- –网页中术语的位置e.g. <title>, <h1>, link text, body text
- –本页词频和总词频之比
- –指向本页的锁定文本
- –指向本页的链接分析
- –有时，点击分析
- –对于新网页，结合考虑新鲜度问题
- **关于商业因素**。例如：某些系统如发现检索对象与人为增加检索要素的商业公司网站有联系，则不予排列等。

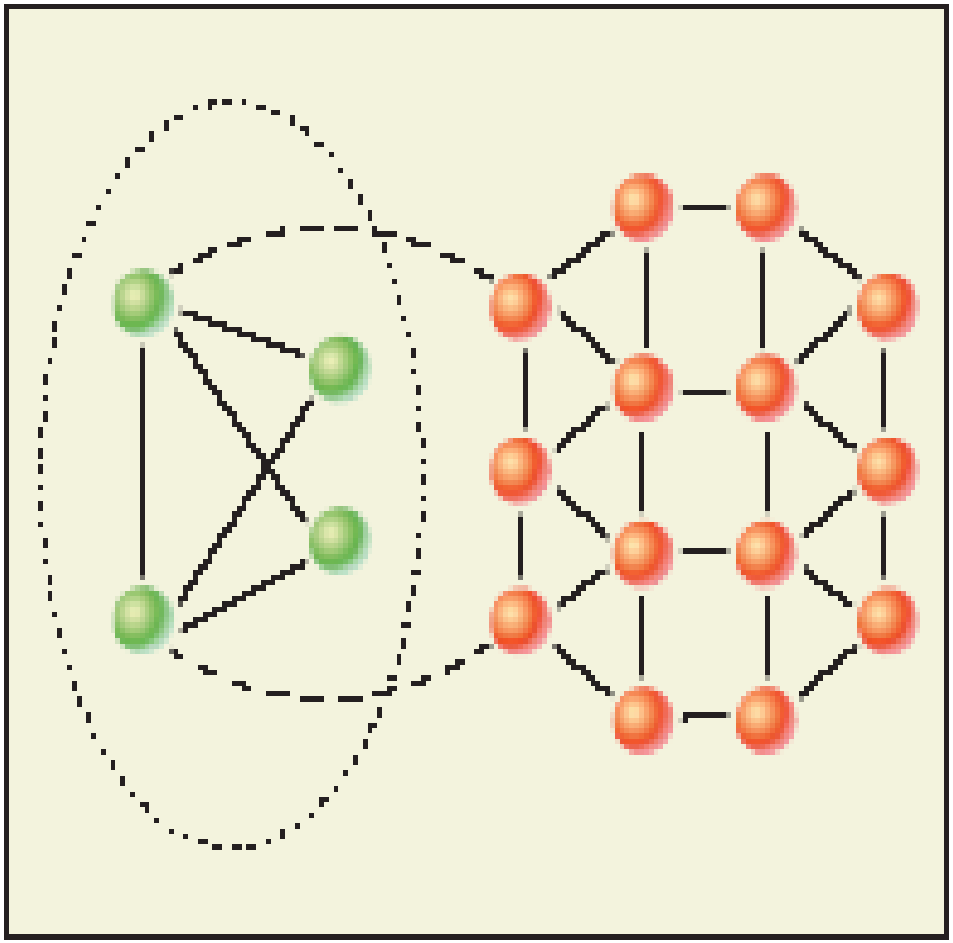
链接控制 —作为重要性测量（略）

- 例:每一网页从100分开始。
- 按入链分数重新计算 。
- 延续计算直到分数不再变化。



网络社区的识别与应用





网络关键词检索技术的特点与传统文本检索的不同

主要表现在：

1. 重视查准因素，忽略检全因素。
2. 结合网络文献的特点， 扩大了检索算法的应用，提高结合多种因素的应用能力。
3. 更加重视易用性。

检索优化的发展

- 检索优化的含义与必要性
- 检索优化的常见形式

检索优化的含义和必要性

检索优化指通过对用户检索提问提出供选择的方案，以交互的方式，优化检索查询，以改进检索结果。

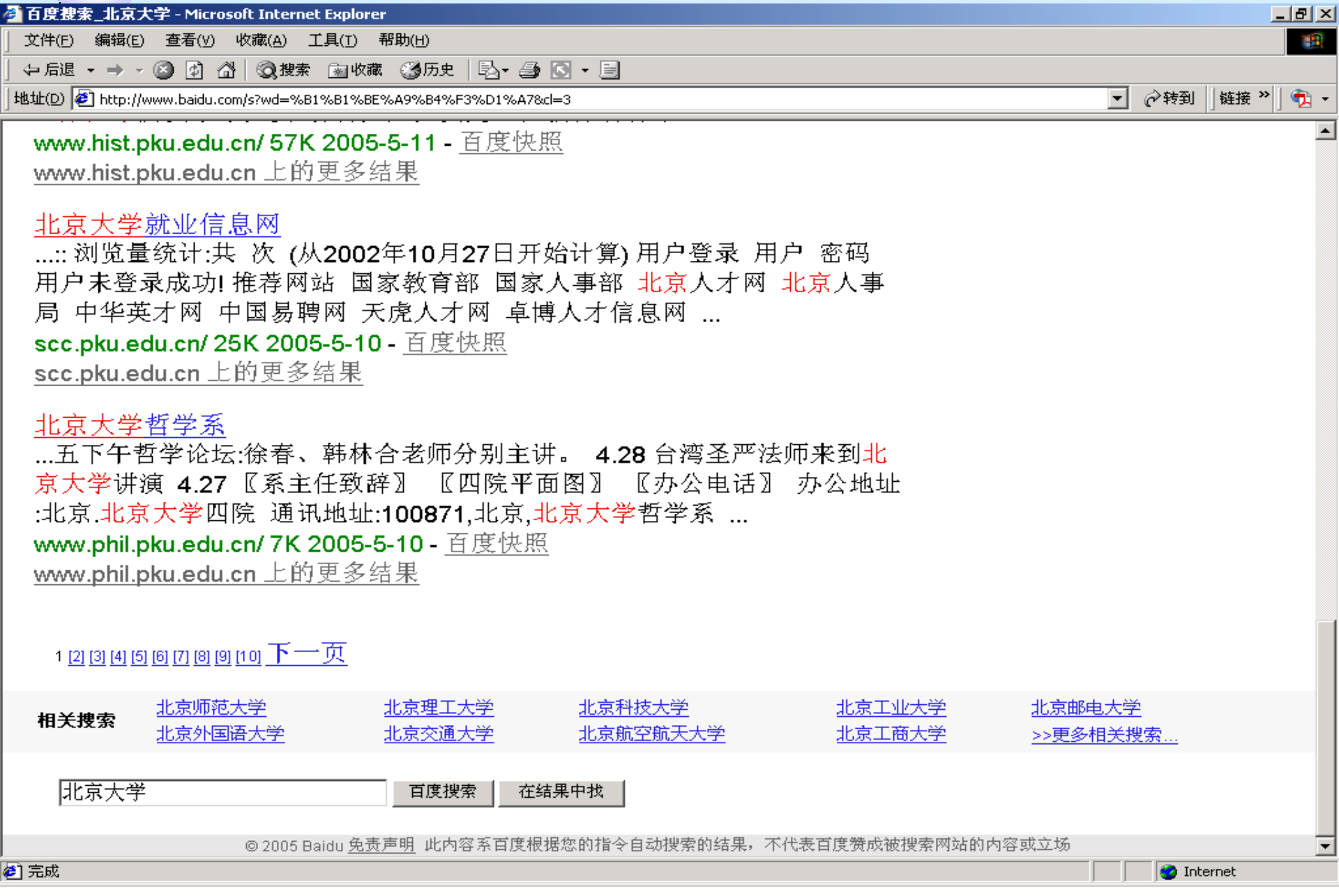
优化的原因：

- 找不到准确表达检索内容的词汇；关键词于进行检索的内容之间可能存在着差距，需要在检索过程中进行调整；
- 表达不够专指，没有确切表达出用户潜在的检索需求。
- 用户不了解逻辑表达式的书写方法，从而影响检索表达，
- 检索深化的问题。检索调查表明，多数检索只用一个词进行。
- 检索调整。需要根据改变检索方向，进行相关查找的问题。
- 多种要素检索。用户很难同时照顾到。

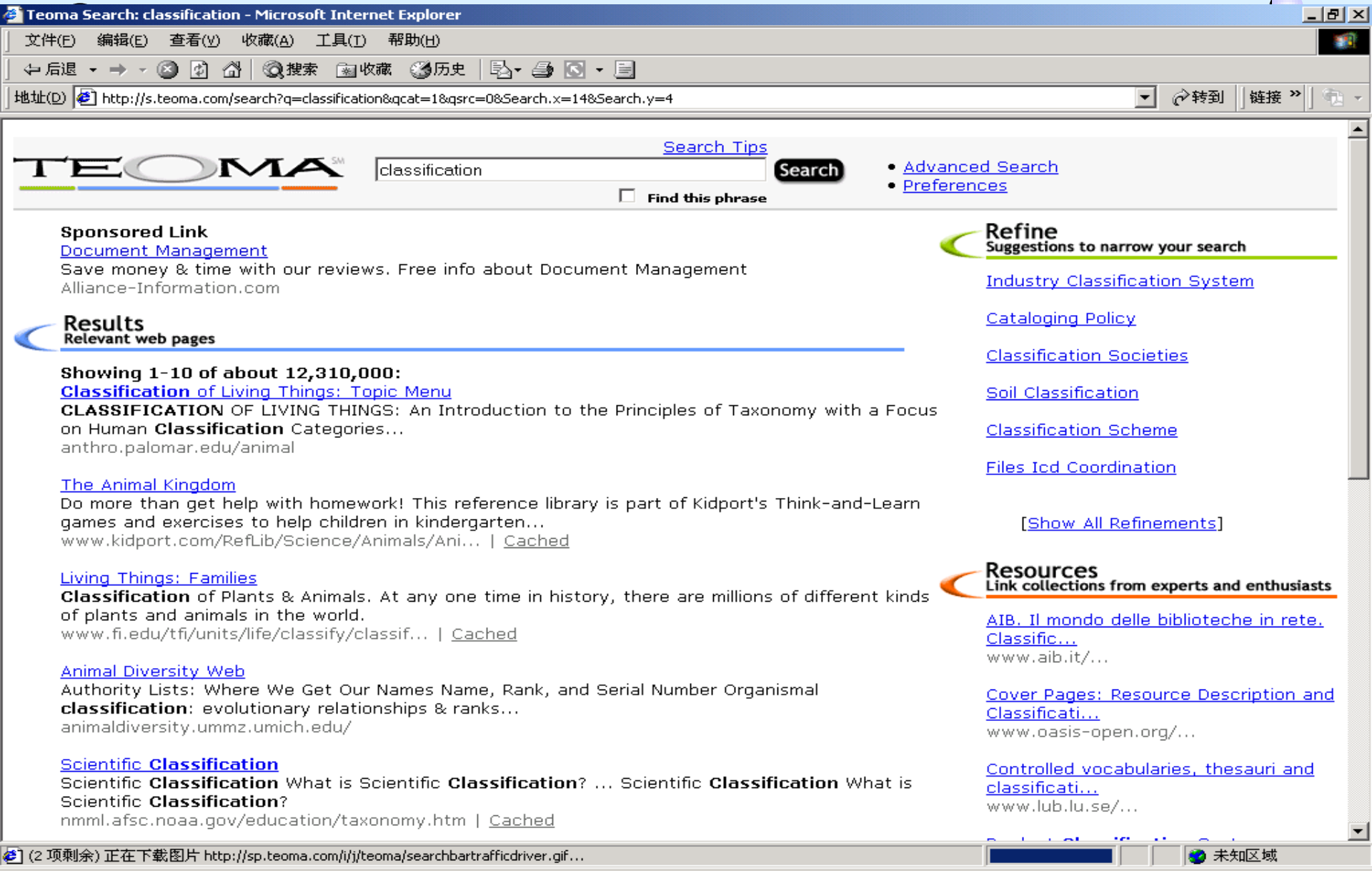
搜索引擎常用检索优化的形式

- 利用用户检索查询，提供检索查询的优化。
- 检索纠错功能。
- 将聚类算法的结果作为二次检索的依据。
- Similar to。

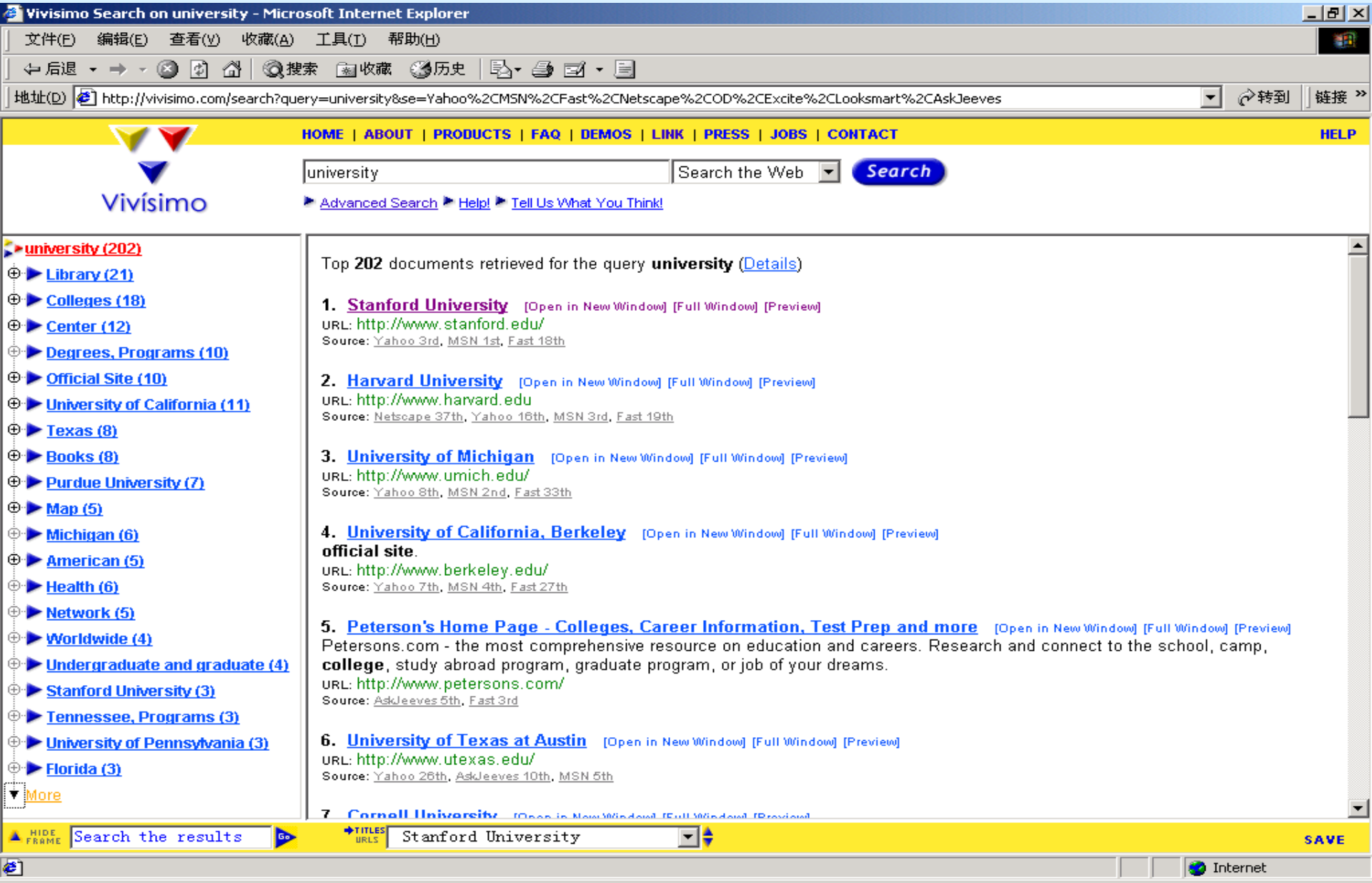
百度的检索优化功能


百度搜索_北京大学 - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏 历史
地址(D) http://www.baidu.com/s?wd=%B1%B1%BE%A9%B4%F3%D1%A7&cl=3
转到 链接
www.hist.pku.edu.cn/ 57K 2005-5-11 - 百度快照
www.hist.pku.edu.cn 上的更多结果
北京大学就业信息网
...:: 浏览量统计:共 次 (从2002年10月27日开始计算) 用户登录 用户 密码 用户未登录成功! 推荐网站 国家教育部 国家人事部 北京人才网 北京人事局 中华英才网 中国易聘网 天虎人才网 卓博人才信息网 ...
scc.pku.edu.cn/ 25K 2005-5-10 - 百度快照
scc.pku.edu.cn 上的更多结果
北京大学哲学系
...五下午哲学论坛:徐春、韩林合老师分别主讲。 4.28 台湾圣严法师来到北京大学讲演 4.27 〖系主任致辞〗 〖四院平面图〗 〖办公电话〗 办公地址:北京.北京大学四院 通讯地址:100871,北京,北京大学哲学系 ...
www.phil.pku.edu.cn/ 7K 2005-5-10 - 百度快照
www.phil.pku.edu.cn 上的更多结果
1 [2] [3] [4] [5] [6] [7] [8] [9] [10] 下一页
相关搜索
北京师范大学 北京理工大学 北京科技大学 北京工业大学 北京邮电大学
北京外国语大学 北京交通大学 北京航空航天大学 北京工商大学 >>更多相关搜索...
北京大学
百度搜索 在结果中找
© 2005 Baidu 免责声明 此内容系百度根据您的指令自动搜索的结果，不代表百度赞成被搜索网站的内容或立场
完成
Internet

自动聚类基础上的检索帮助



自动聚类基础上的二次检索帮助



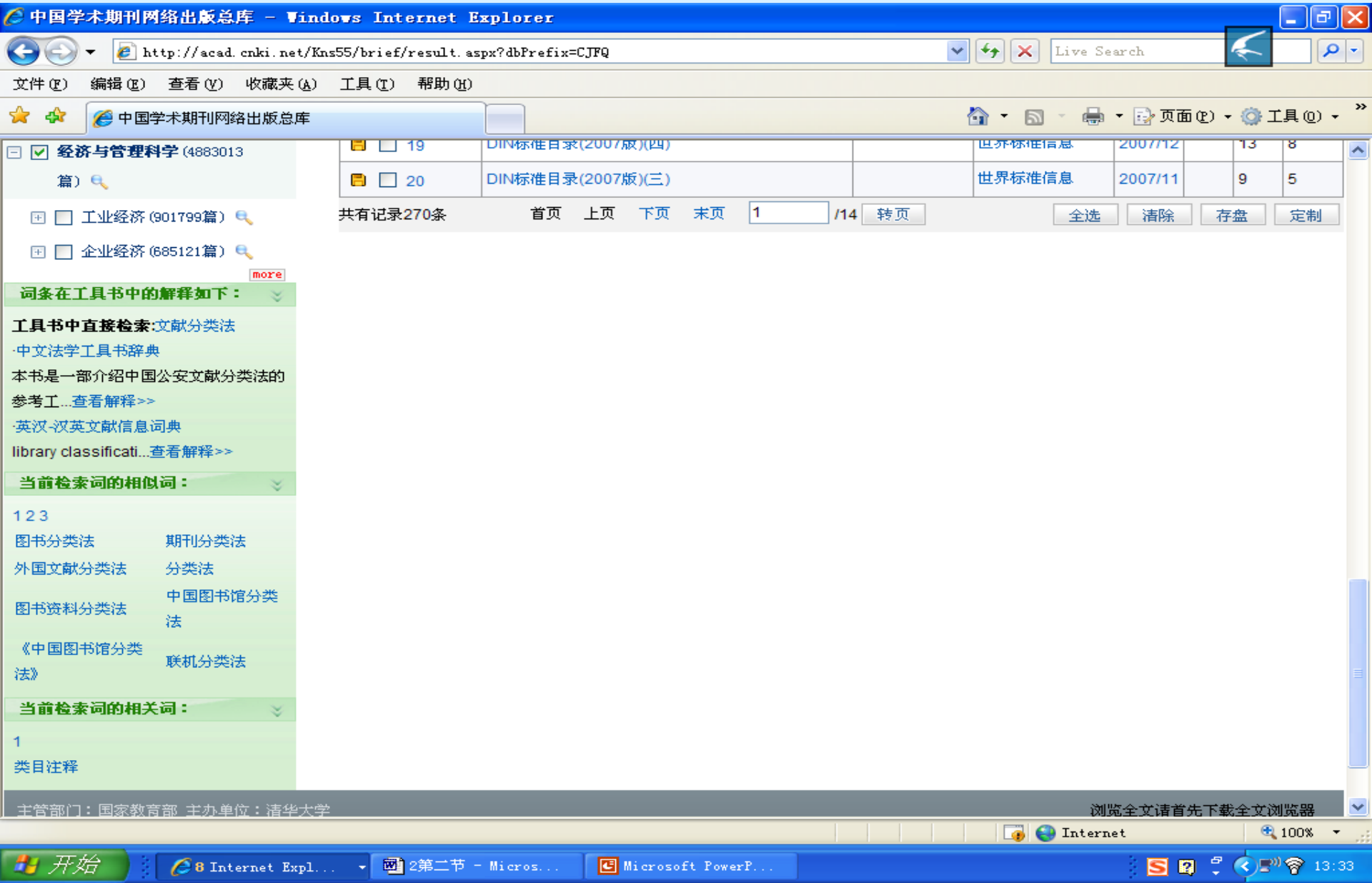
万方检索优化实例


万方数据知识服务平台--论文检索结果 - Windows Internet Explorer
http://s.wanfangdata.com.cn/paper.aspx?f=SimpleSearch&q=%E5%88%86%E7%B1%BB%E6%B3%95&n=10&PID=&CID=
Live Search
文件(F) 编辑(E) 查看(V) 收藏夹(A) 工具(T) 帮助(H)
万方数据知识服务平台--论文检索结果
页面(P) 工具(O)
分类法
检索
高级检索 帮助
论文搜索: 经典论文优先 | 相关度优先 | 新论文优先
共找到6242篇符合条件的论文，以下是1-10
缩小搜索范围
标题
作者
来源 全部
年
论文类型
全部论文
确定
论文类型
期刊论文 (4380)
学位论文 (1141)
外文期刊 (465)
会议论文 (256)
年份
2008- (381)
2005-2007 (2649)
2000-2004 (2734)
1990-1999 (396)
早于1990 (82)
按刊分类
医学、卫生 (1971) 文化、科学、教育、体育 (1332) 工业技术 (1234) 经济 (428) 农业科学 (269) 天文学、地球科学 (175) 语言、文字 (128) 数理科学和化学 (104) 环境科学、安全科学 (85) 政治、法律 (83) 交通运输 (82) 文学 (59) 生物科学 (54) 历史、地理 (48) 艺术 (29) 自然科学总论 (29) 哲学、宗教 (15) 航空、航天 (11) 社会科学总论 (11) 综合性图书 (3) 军事 (1)
添加到导出列表
1 SIMCA分类法与PLS算法结合近红外光谱应用于卷烟纸的质量控制
[期刊论文] 王家俊, 汪帆, 马玲, WANG Jia-jun, WANG Fan, MA Ling - 《光谱学与光谱分析》 2006年10期
应用SIMCA分类法与PLS算法结合卷烟纸的傅里叶变换近红外光谱(FT-NIR)建立了卷烟纸的分类模型,用于卷烟纸的判别分类,效果良好;同时,建立了测定卷烟纸定量、厚度、透气度、水分和灰分等性质的校正模型,其相应的相关系数分别... 查看全文
2 基于核的最小距离分类法的参数选择方法
[期刊论文] 邱潇钰, 张化祥, QIU Xiao-yu, ZHANG Hua-xiang - 《计算机工程》 2008年5期
在基于核函数的最小距离分类方法对数据集进行分类过程中,目标函数的核函数参数选择直接影响分类器的分类成功率.该文提出一种选择应用目标函数来选择适当参数的方法.实验结果表明,与单纯的基于核的最小距离分类法相比,... 查看全文
3 基于改进多信号分类法的异步电机转子故障特征分量的提取
[期刊论文] 方芳, 杨士元, 侯新国, FANG Fang, YANG Shi-yuan, HOU Xin-guo - 《中国电机工程学报》 2007年30期
在基于定子电流信号进行异步电机故障诊断时,转子断条故障特征频率分量常常被电流的基频分量淹没.针对这一情况,该文提出一种新的改进的MUSIC方法来提取这一故障特征频率.MUSIC方法通过特征值分解把自相关矩阵中包含的信... 查看全文
4 平面度误差的快速评定法——测点分类法
[期刊论文] 岳武陵, 吴勇, 苏俊, YUE Wu-ling, WU Yong, SU Jun - 《计量学报》 2007年1期
针对平面度误差判定的最小包容区域法,提出一种新的、快速的实施方法,它将所有测量点分成"高点"、"低
广告服务
论文翻译
数字图书馆论坛
中华临床医师杂志征稿
Internet
100%

同方的检索优化实例


中国学术期刊网络出版总库 - Windows Internet Explorer
http://acad.cnki.net/Kns55/brief/result.aspx?dbPrefix=CJFQ
Live Search
文件(F) 编辑(E) 查看(V) 收藏夹(A) 工具(T) 帮助(H)
中国学术期刊网络出版总库
页面(P) 工具(O)
轻工业手工业(554151篇)
有机化工(491973篇)
more
工程科技II(2517708篇)
建筑科学与工程(571816篇)
电力工业(449802篇)
more
农业科技(2005544篇)
畜牧与动物医学(521537篇)
园艺(356502篇)
more
医药卫生科技(5164307篇)
临床医学(593311篇)
外科学(547712篇)
more
哲学与人文科学(2227415篇)
中国文学(467200篇)
人物传记(315007篇)
more
2.输入内容检索条件：
主题
文献分类法
并且包含
输入检索词
精确
在结果中检索
检索文献
中英文扩展检索
定制或收藏本次检索式
3.您可以按如下文献分组排序方式选择文献：(只对前40000条记录分组)
文献分组浏览: 学科类别 期刊名称 研究资助基金 研究层次 文献作者 作者单位 中文关键词
文献排序浏览: 发表时间 相关度 被引频次 下载频次 浏览频次
每页记录数：10 20 50
摘要显示 列表显示
共有记录270条 首页 上页 下页 末页 1 /14 转页
全选 清除 存盘 定制
序号 篇名 作者 刊名 年/期 被引频次 下载频次 浏览频次
1 科技查新检索中的关键词选择 张柏秋; 吴晓镭 情报科学 2008/09
2 ISO标准(2008年版)馆藏目录(一) 世界标准信息 2008/09
3 《中国图书馆分类法·地震学专业分类表》的编制修订 付桂华 国际地震动态 2008/08 1 3
4 中国国家标准(2008年版)馆藏目录(四) 世界标准信息 2008/08 16 64
5 在《中图法》中增设专业英语类目的探讨 张玉辉 图书馆建设 2008/07 2 5
6 中国国家标准(2008年版)馆藏目录(三) 世界标准信息 2008/07 12 46
7 浅谈大学生如何利用图书馆 李艳明; 李巨伟 黑龙江科技信息 2008/15 18 26
8 中国国家标准(2008年版)馆藏目录(二) 世界标准信息 2008/05 26 104
Internet
100%
开始
Internet Expl...
2第二节 - Micros...
13:29

同方检索优化实例


中国学术期刊网络出版总库 - Windows Internet Explorer
http://acad.cnki.net/Kns55/brief/result.aspx?dbPrefix=CJFQ
Live Search
文件(F) 编辑(E) 查看(V) 收藏夹(A) 工具(T) 帮助(H)
中国学术期刊网络出版总库
页面(P) 工具(O)
经济与管理科学 (4883013篇)
工业经济 (901799篇)
企业经济 (685121篇)
more
词条在工具书中的解释如下：
工具书中直接检索:文献分类法
·中文法学工具书辞典
本书是一部介绍中国公安文献分类法的参考工...查看解释>>
·英汉-汉英文献信息词典
library classificati...查看解释>>
当前检索词的相似词：
1 2 3
图书分类法
期刊分类法
外国文献分类法
分类法
图书资料分类法
中国图书馆分类法
《中国图书馆分类法》
联机分类法
当前检索词的相关词：
1
类目注释
19 DIN标准目录(2007版)(四) 世界标准信息 2007/12 13 8
20 DIN标准目录(2007版)(三) 世界标准信息 2007/11 9 5
共有记录270条 首页 上页 下页 末页 1 /14 转页
全选 清除 存盘 定制
主管部门：国家教育部 主办单位：清华大学
浏览全文请首先下载全文浏览器
Internet 100%
开始
8 Internet Expl...
2第二节 - Micros...
Microsoft PowerP...
13:33

文本系统的改进努力：控制的纳入

检索方式：提供简单检索、高级检索、专业检索等检索界面。上述方面文本系统略优；努力提供自然语言检索能力。（引入句法控制、词汇控制）

检索排序：多因素结合提供；多种排序方式的采用；两者差距不大。（引入多因素控制）

检索优化：相关检索帮助，百度，Ask；二次检索；Vivisimo（引入词汇控制）

。

文本检索系统的改进努力

文本控制的特点：

- 采用后控的方式；
- 多方面，多角度，词法、句法；
- 多因素；

文本控制的不足：

- 词汇控制不严格；
- 一些控制的方式仍有待优化、改进，如自动聚类，检索语句切分等，仍然在发展探索中

总体评价：

- 作了大量努力，有明显效果。

控制系统的努力与差距

- **控制系统加强处理能力，降低成本**的努力：
  - 联合编目；
  - 自动标引试验。
- **控制系统的不足：**
  - 检索语言能力的应用与开发不足，优势未得到发挥。

检索语言能力的应用与开发的差距

可在词表和标引数据基础上提供，而未提供的功能包括：

检索入口方面：入口词检索；以浏览形式提供词表词的问题；相关词的提供问题。

检索优化方面：主题检索帮助方面可以提供的，如结合结合分类等提供；相关主题词提供；分类的二次检索，结合主题标题形式的二次检索帮助等。

词汇控制系统的差距何在

A，主要是检索端或检索应用方面的差距。

B，检索端的重要性：功能是通过检索界面实现的，未实现的功能只是潜在能力；且无法在应用基础上进一步改进。

C，两者性能各有优缺点。但自然语言系统努力改进，控制系统改进不力，检索端成为短板。

控制系统检索端差距的原因

- 对检索语言应用端的重视不够，停留在检索语言编制和标引阶段；
- 缺乏电子环境下应用的研究；
- 与计算机软件编制人员沟通不够；
- 应用基础上的改进不够----持续发展意识不够等。

应汲取网络、文本数据库等的发展，结合主题语言的特点改进。

维基百科的分类界面-检索界面的多样性

比较基础上的思考

- 词汇控制是有价值的，文本检索系统改进的手段之一是引入词汇控制；
- 基于词汇控制的检索系统的不足不是词汇控制造成的，而是检索应用的开发不充分的缘故，
- 目前控制系统的检索界面应向文本系统学习，结合控制语言的特点加以开发。

四. 关于主题控制系统标引实践和规则改进的思考

基本看法：

- 标引方法和规则是根据应用需要确定的，应结合电子环境下的实践发展、改进和调整。
- 在电子环境下检索系统的探索中，图书馆书目检索系统、文献数据库系统、网络检索系统正经历一个后者向前者学习，超过前者，前者反过来学习后者的过程。
- 不仅要向国外的同行学习，**而且要善于向网络、文献数据库的检索发展学习；但向网络学习并非全盘否定自己。**

理论、方法、规则的改进问题

想到的一些问题：

- 检索应用方式的优化改进问题；
- 使用方式以及相应规则的调整问题，比如说：
  - 是不是建立标题；轮排还要不要？与标引规则。
  - 还要不要控制，自由词的应用问题，入口词的问题；
  - 特定主题类型标引规则的调整问题；
  - 词表的应用问题，如作为切分工具；
  - 词表系统的构建层次问题，如：wordnet—关键词—叙词
  - 一检索系统中不同特点检索系统之间的结合和分工问题。
  - 不同系统之间兼容与互操作问题。
- 其他问题，如主题标引中中文分面公式问题。
- MARC格式的适用性和改造问题，灵活性问题，如轮排的处理；XML语言应用问题。

控制系统的检索应用的改进问题

想到的几个基本功能，如：

- 检索入口界面词表浏览功能的提供；
- 入口词检索功能的采用，自然语言检索方式的加强；
- 检索优化功能的开发，如结合控制语言的二次检索功能，包括标题词浏览等，相关词的提供等；
- 一检索系统中不同特点检索方法之间的结合和分工问题。
- 不同系统之间兼容与互操作问题。
- 其他方法的引入等。如用户因素等。

结合使用方式的相应标引处理规则的思考或调整，如：

- 是不是建立标题；还要不要轮排？（文本系统缓存中的先组标题保留）
- 自由词的应用与控制问题，结合检索词的入口词拓展问题；
- 特定主题类型标引规则的调整问题；如地区，文学、艺术，特殊文献类型等；
- 主题标引中中文分面公式问题；
- MARC格式的适用性和改造问题，灵活性问题，如轮排的处理；

是否建立标题、轮排

关于轮排模式。国内采用轮替法轮排：

A B C D

B A C D

C A B D

D A B C

例：电子计算机—硬磁盘—生产工艺

硬磁盘—电子计算机—生产工艺

使用“：”、“，”连接的主题词，轮排中随原连接的词移动。

例：小说—语言学：美学

语言学：美学—小说

美学：语言学—小说

联结主题的轮排，必要时对连接词应作适当调整。

轮排模式在检索优化中有使用价值。

特定主题类型标引规则的调整问题；如地区，等

地区应该标引中**直接标引和间接标引的应用效果以及规则的设置。（类似的如时代标引以及应用等）**

城市地理—广州

地方志—浙江—桐乡

雕塑—罗马

旅游指南－德国－科隆

主题标引中中文分面公式问题；

- 国内主题分面公式归纳为：主体因素（研究对象等中心主题概念）、方面因素或限定因素（成分、材料、方法、过程、条件、状态、尺度、性质等对主体因素研究方面（角度）的说明或限定因素）、空间因素、时间因素、文献类型等。
- 与阮冈纳赞,P;M:E’S.T；轮（round），层（level），相（phase）等概念构成的比较完善的主题分析系统之间的差异。

通用引用次序

- 阮冈纳赞的五个基本范畴

一 本体（personality）- 物质（material）- 动力（energy）-空间（space）-时间（time）

轮（round）、层（level）、相（phase）：

同一范畴中，如主题特征不止一个，则可分析出二层本体、三层本体，二层物质、三层物质，用Ｐ２、Ｐ３，Ｍ２、Ｍ３表示。

如动力后再出现本体、物质，则称为第二轮本体、第二轮物质，用２Ｐ、２Ｍ表示。

如，“合金对直升飞机起落架的加工”可标引为：

直升飞机——起落架——加工——车刀——合金钢

Ｐ１　　Ｐ２　　Ｅ　　２Ｐ　　２Ｍ

其他一些与检索语言应用相关的问题

- 一检索系统中不同特点检索系统之间的结合和分工问题。
- 不同系统之间兼容与互操作问题。
- 词表的应用问题，如切分工具、用以自然语言检索等；
- 词表的层次问题，如：wordnet—关键词—叙词；
- 标识语言如XML语言应用问题， XML语言可以了解一些。
- 关于ontology。

XML记录北大图书馆网站示例

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-
   syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.0/"
xmlns:dcq="http://purl.org/dc/qualifiers/1.0/">
<rdf:Description about="http://www.lib.pku.edu.cn">
<dc:title>Peking University Library</dc:title>
<dc:coverage>P.R.China - Beijing</dc:coverage>
<dc:coverage>1902-</dc:coverage>
<dc:creator>Peking University Library</dc:creator>
<dc:format>text/html</dc:format>
<dc:publisher>Peking University Library</dc:publisher>
<dc:date>1996-10-25</dc:date>
```

<dc:description>Library homepage, brief introduction, electronic resources, OPAC, User Guide, News, digital library, Inter-library Loan, FAQ, Navigations, Focus, User training program, CALIS, CAI</dc:description>

<dc:identifier>http://www.lib.pku.edu.cn</dc:identifier>

<dc:language>chi</dc:language>

<dc:relation>http://www.lib.pku.edu.cn/enhtml/index.htm</dc:relation>

<dc:type>Text data</dc:type>

<dc:type>text/html; charset=gb2312</dc:type>

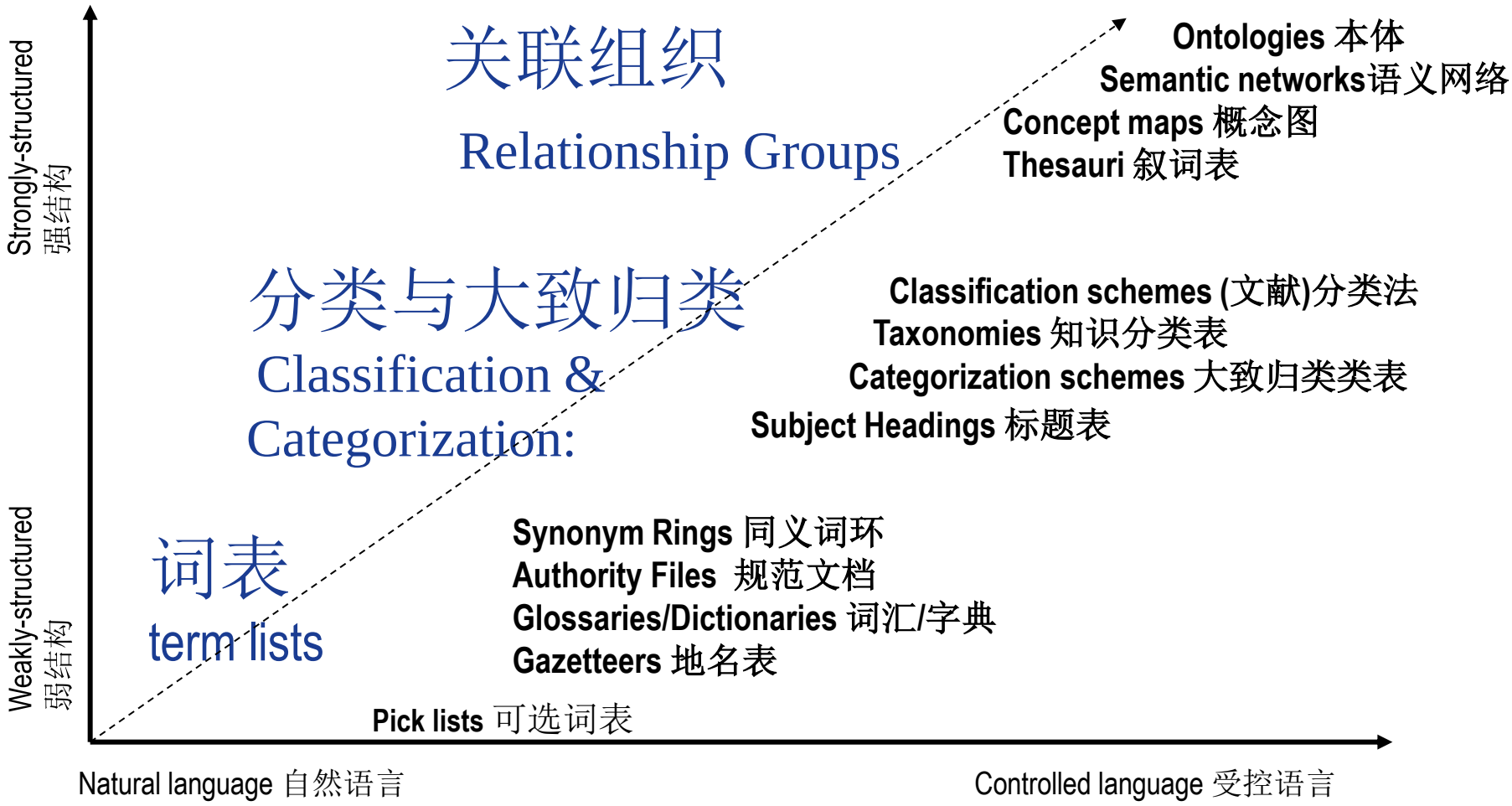
</rdf:Description>

</rdf:RDF>

Title	Peking University Library
Identifier.URI	http://www.lib.pku.edu.cn
Type.OCLCg	Text data
Type text/html;	charset=gb2312
Coverage.spatial	P.R.China - Beijing
Coverage.temporal	1902-
Creator.namePersonal	Peking University Library
Date.created	1996-10-25
Description	Library homepage, brief introduction, electronic resources, OPAC, User Guide, News, digital library, Inter-library Loan, FAQ, Navigations, Focus, User training program, CALIS, CAI
Format	text/html
Language.ISO639-2	chi
Publisher	Peking University Library
Relation.hasVersion	http://www.lib.pku.edu.cn/enhtml/index.htm

OCLC 协作联机资源编目系统（CORC）中元数据的显示界面

知识组织系统 (KOS) 一览



资料来源：曾蕾：受控语言标准最新进展

理论、方法、规则的改进问题

想到的一些问题：

- 检索应用方式的优化改进问题；
- 使用方式以及相应规则的调整问题，比如说：
  - 是不是建立标题；轮排还要不要？与标引规则。
  - 还要不要控制，自由词的应用问题，入口词的问题；
  - 特定主题类型标引规则的调整问题；
  - 词表的应用问题，如作为切分工具；
  - 词表系统的构建层次问题，如：wordnet—关键词—叙词
  - 一检索系统中不同特点检索系统之间的结合和分工问题。
  - 不同系统之间兼容与互操作问题。
- MARC格式的适用性和改造问题，灵活性问题，如轮排的处理；XML语言应用问题。
- 其他问题，如主题标引中中文分面公式问题；

谢谢！