

搜索引擎与网络学习

20130922

课程目标

提升学习力、思考力

学习力首先是资源获取的能力

信息获取依赖于资源和方法

学习目标

本次课程涉及的搜索引擎

主要介绍google

申明

关于安全搜索

- 系统
- 信息可靠
- 过时
- 品种较少
- 无法赶上快速发展社会的需要
- 快速、高效
- 时效性好
- 准备这次课的主要信息来源

资源

社交网络

推送的信息

搜索方法

RSS

交流分享

不懂的问题，问百度，google

检索是一种重要的学习途径

基本检索

高级检索

谷歌 ncr No Country Redirect

强制跳转到英文网页

baidu

bing

360

国际搜索

国际接轨

最强大的搜索引擎

以大家现场提出的搜索问题为主

互动演示时

充分体现互动课堂的优越性

对老师要重新定位，并非万能

只为大家学到东西，而不是维护老师的权威

如果出现不安全（成人内容）的结果，请忽视之

互联网的普及，信息爆炸

如何整理互联网信息

当知识增加到一定的程度就需要进行索引

图书

图书室

图书馆

归类整理

目录索引

创建文件夹

利用搜索软件

自动创建索引

everything

桌面搜索

根据以上推测，搜索引擎的类型

人工目录型

yahoo

google

baidu

现在基本都是

搜索引擎之父？

利用搜索进行学习 乔纳森·弗莱彻：被遗忘的搜索引擎之父.pdf

启示

坚持

远见

世界上最受欢迎的老师

搜索框正在重塑世界

搜索是最大、最稳、最具颠覆性的创新之一

搜索是最有影响力的学习途径之一

搜索引擎是最好的老师

搜索引擎的重要性

改变了信息获取方式和学习方式

学习

科研

教学

购物

旅游

天气查询

路线的查询

特色景点



搜索引擎出现的历史必然性

搜狗  $SQ=K/T(C)$

- 小吃
- 追踪包裹
- 计算
- SQ = 搜商
- K = knowledge 知识
- T = time 时间
- C = 搜商指数 (社会平均知识获取能力)

简要历史

- 1994.4 杨致远等创办Yahoo
- 1998.9 Google 创立
- 2000.1 李彦宏、徐勇两人创立百度
- 2009年 bing 推出
- 2012.8 360 搜索

搜索引擎发展史

国内搜索引擎的发展趋势



百度第一 超60%

如何准确，高质量的获取信息

SO 接近20%

SO 的广告模式模拟 google，明确标示，百度是混在一起，这点不妥

360告诉我们：机会永远存在，关键看你是否眼看到用户的需求

感悟 yahoo告诉我们适者生存

我们必须根据环境的变化，时代的发展不断改变自己

国内热门的搜索引擎格局

搜索引擎的挑战

排序问题

- 无人能读回数十万万的结果
- google更希望按文章的重要性进行排序
- 百度在排序时，加入了竞价排名的策略
- 引用，其实就是投票
- 与后面课程中的引文分析进行类比
- 多种判断标准

搜索引擎如何帮助我们更好地更快地找到想要的信息

即将到来的搜索引擎变革

- 理解搜索意图
- 直接给出解决方案
- 场景式交互搜索

使用搜索引擎做什么



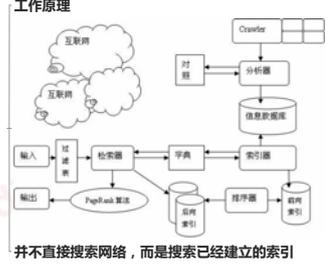
决定搜索成效的因素

- 技巧
- 态度
- 知识或积累

搜索引擎的异同

how search work

google 网页 [点击even more-web search](http://www.google.com/intl/en/inside/search/howsearchworks/theory/)



搜索引擎的工作原理

搜索引擎的异同

- 爬取的范围和数量不同
- 索引的方法不同
- 反馈的结果数量不同
- 排序的方式不同
- 交互的功能不同
- 结果的呈现方式不同

好的搜索引擎应该能够快速帮我们找到最有价值的信息

主界面的介绍

- 搜索框
- 关键词的选取
- 关键词越少，结果越多

搜索引擎的使用入门

逆向思维：要考虑信息发布者可能的描述方式

- 地图 - 20%
- 图书
- 图片
- 视频
- 音乐

头疼
头痛
脑袋痛
大脑不舒服
人体控制中心痛
脖子上的这玩意儿痛

搜索的三重境界

博客

高级搜索介绍 限定?

搜索设置

and
OR
NOT
""

搜索进阶：基本逻辑关系

site site:ustc.edu.cn
site:edu.cn
site:cn

高效搜索：搜索命令的使用

文件类型
doc
ppt
xls
pdf
txt
inc
exe
rar
.....
filetype

1. 高效检索思路
快速尝试，不断优化
宽泛检索，逐步精炼
精确检索，直达结果
关键的词的变更和添加
与文献检索中，关键词的选取类似
先用非专业词汇检索，逐步逼近

2. 基本检索
搜索结果页面的展示
新闻
图片
网页
分类搜索
专利
学术搜索
book
视频等
search tool选项
google 语言设置会对搜索结果产生影响

3. 高级检索
页面介绍
在搜索结果页面才能设置

AND
OR
- 非
+ 强制搜索

基本搜索命令
词组搜索
"stem cell"
"weight loss"

相关搜索，已经取消
beauty
nokia
如何利用以上命令验证?
方法: ~关键词 -关键词

4. 搜索命令
site:ustc.edu.cn
site:edu.cn
site:cn

site
清华或新加坡有哪些老师跟我玩一样的方向?
《自由星空》视频下载

aptamer
site:tsinghua.edu.cn
aptamer site:ustc.edu.sg

filetype
文件类型
filetype:pdf

intitle intitle:aptamer

高级命令
allintitle

inurl
http://198.82.159.136/view/index.shtml
http://134.29.208.43/view/index.shtml
飞机厂内部

allinurl

intext

link www.ustc.edu.cn

related www.ustc.edu.cn

命令组合 inurl:aptamer
site:com

index of

google应用示例

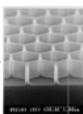
5. 图片搜索
查找原始大图



查找不认识的人



查找某个物体



这人的信息是正确的么?

http://www.wla.hk/ 网站首页 > 关于我们 > 委员会 > 供货商委员会
http://www.wla.hk/html/aboutus/welyuanhui/ProviderCommittee/ 网站首页 > 报告 > 世界奢侈品研究中心
http://www.wla.hk/html/sjzx/WLAResearchCenter/

6. trends
搜索风云榜

可以了解某个关键词被搜索的情况
可以按时间、地域进行统计
aptamer

baidu应用示例

- 1. 基本检索
 - 百度文库
 - 百度知道
 - 百度百科
- 2. 高级检索
 - 可以了解社会热点
- 3. 风云榜
 - 最新趋势
 - 八卦资讯
- 4. 搜索命令
 - |
 -
 - ""
 - 《》
 - filetype
 - site
 - intitle: title:
 - inurl

google baidu bing 360 的比较

- 收录范围
 - google
 - bing
 - 百度 / 中文收录最大
 - 360 / 收录结果偏少, 但质量相对较高
- 搜索命令
 - google较多
 - 百度 SO / 《》
 - inanchor
 - language
 - bing
 - loc
 - prefer
 - contains
- 360
 - 最少
 - ctrl ctrl快捷搜索
- google提供多种精炼功能
- 360提供用户评价 / 广告类似google的机制, 明确标识
- 百度对于文库中的内容直接给出评价, 竞价排名不爽
- 360提供用户评价

小结

- 建立搜索意识
 - 悟+练习
 - 搜索是逐步逼近的过程
- 掌握基本的搜索技巧
 - 基本检索
 - 高级检索
 - 搜索命令
- 掌握利用搜索引擎解决问题的思路
 - 关键词的提炼
 - 检索式的优化, 表达的转换
 - 屏幕写字
 - 逐步逼近的过程
 - 组合利用高级搜索和命令检索
 - 逆向思维, 考虑哪些地方可能有这些信息, 可能会怎么表述

作业及思考题

1. 搜索分享
 - 分享课上未讲到的搜索技巧, 所有可以快速帮你找到特定内容的技巧和经验均可;
 - 作业标题: 学号+搜索分享
2. 牛人搜索
 - 找出某个想学习的主题知识, 找到这方面的牛人和TA发布的学习资料, 并简介搜索过程;
 - 作业标题: 学号+牛人搜索
3. 搜索引擎改进
 - 根据你多年使用搜索引擎的经验, 说说现有搜索引擎的不足之处, 并提出希望改进的方向;
 - 作业标题: 学号+搜索引擎改进
4. 资源分享
 - 分享某些有价值的资源网站, 可以满足学习、科研和生活上特定需求的网站;
 - 作业标题: 学号+资源分享;
5. 课程改进
 - 发现课程中有错的内容, 以及课程中值得改进的地方, 并提出建设性意见;
 - 作业标题: 学号+课程改进;

意见反馈

- 课程报名 如果在先听课的学员希望报名此课程, 请尽快登录<http://mooc.ustc.edu.cn>
- 聊天室, 建议不要在问答中提问 / 每人只能看见自己和主讲者
- 可以在新浪微博 @文献管理 进行互动
- 本课程的核心互动平台
 - 
- 课程论坛
 - <http://mooc.ustc.edu.cn/upload/forum.php?mod=viewthread&tid=28572&extra=page%3D1>
- 汤森路透网站
 - <http://ip-science.thomsonreuters.com.cn/2013year/ustc/courses.html>
- 网易云课堂 爱科技
 - <http://study.163.com/u/scitech#/center/orgCenterCourse>
- 或关注新浪微博 @文献管理 @罗昭锋
- 请积极参与班级QQ群交流, QQ群中的优质内容欢迎发布到论坛。

致谢

-  THOMSON REUTERS 汤森路透
-  网易云课堂
-  XMind <http://www.xmind.net/> 优秀的思维导图软件

感谢所有学员、助教和每一位课程的分享者!