



# 基于开源工具搭建小型搜索引擎

王斌 骆卫华

中科院计算所

2006-5-18



# 提纲

- 搜索引擎体系结构
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题
- 参考文献



# 提纲

- 搜索引擎体系结构 ◀
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题
- 参考文献



# 搜索引擎类型

## ● 目录式搜索引擎



**dmoz** open directory project

In partnership with  
**AOL** search

[about dmoz](#) | [suggest URL](#) | [help](#) | [link](#) | [editor login](#)

Search

[advanced](#)

### Arts

[Movies](#), [Television](#), [Music](#)...

### Games

[Video Games](#), [RPGs](#), [Gambling](#)...

### Kids and Teens

[Arts](#), [School Time](#), [Teen Life](#)...

### Reference

[Maps](#), [Education](#), [Libraries](#)...

### Shopping

[Autos](#), [Clothing](#), [Gifts](#)...

### World

[Deutsch](#), [Español](#), [Français](#), [Italiano](#), [Japanese](#), [Nederlands](#), [Polska](#), [Dansk](#), [Svenska](#)...

### Business

[Jobs](#), [Real Estate](#), [Investing](#)...

### Health

[Fitness](#), [Medicine](#), [Alternative](#)...

### News

[Media](#), [Newspapers](#), [Weather](#)...

### Regional

[US](#), [Canada](#), [UK](#), [Europe](#)...

### Society

[People](#), [Religion](#), [Issues](#)...

### Computers

[Internet](#), [Software](#), [Hardware](#)...

### Home

[Family](#), [Consumers](#), [Cooking](#)...

### Recreation

[Travel](#), [Food](#), [Outdoors](#), [Humor](#)...

### Science

[Biology](#), [Psychology](#), [Physics](#)...

### Sports

[Baseball](#), [Soccer](#), [Basketball](#)...

[Become an Editor](#) Help build the largest human-edited directory of the web





# 搜索引擎类型

## ● 元搜索引擎



# DOGPILE

All the best search engines piled into one.



[Web](#) | [Images](#) | [Audio](#) | [Video](#) | [News](#) | [Yellow Pages](#) | [White Pages](#)

**Go Fetch!**

[Preferences](#)  
[Advanced Web Search](#)

## Web Search Results for "information retrieval"

Search Filter: [Moderate](#)

Best of Breed Results from the  
Top Search Engines! [View Demo](#)

**Best of All Search Engines**

[Google](#)

[YAHOO! SEARCH](#)

[msn Search](#)

[Ask](#)

[More Engines](#)

All Search Engines 1 - 20 of 96 ([About Results](#))

[1](#) | [2](#) | [3](#) | [4](#) | [5](#) | [Next >](#)

1. [Information Retrieval Research - SearchTools Topics](#)

An up-to-date overview of research in the field of **information** retrieval.

[www.searchtools.com/info/info-retrieval.html](http://www.searchtools.com/info/info-retrieval.html) [Found on Google, MSN Search, Yahoo! Search, Ask.com]

2. [UMASS Amherst: Center for Intelligent Information Retrieval](#)

University of Massachusetts research lab focused on efficient access to large, heterogeneous, distributed, text and multimedia databases.

[ciir.cs.umass.edu/](http://ciir.cs.umass.edu/) [Found on Google, MSN Search, Yahoo! Search, Ask.com]

3. [SIGIR: Information Retrieval](#)

"Addresses issues ranging from theory to user demands in the application of computers to the acquisition, organization, storage, retrieval, and distribution ...

[www.acm.org/sigir/](http://www.acm.org/sigir/) [Found on Google, MSN Search, Yahoo! Search, Ask.com]

4. [Information Retrieval](#)

An online book by CJ van Rijsbergen, University of Glasgow.

[www.dcs.gla.ac.uk/Keith/Preface.html](http://www.dcs.gla.ac.uk/Keith/Preface.html) [Found on Google, Yahoo! Search, Ask.com]

### Are you looking for?

[Evaluation of Retrieval S...](#)

[Text Retrieval](#)

[Search Retrieval](#)

[Document Storage And Retr...](#)

[Book Indexer](#)

[What Is Data Retrieval](#)



# 搜索引擎类型

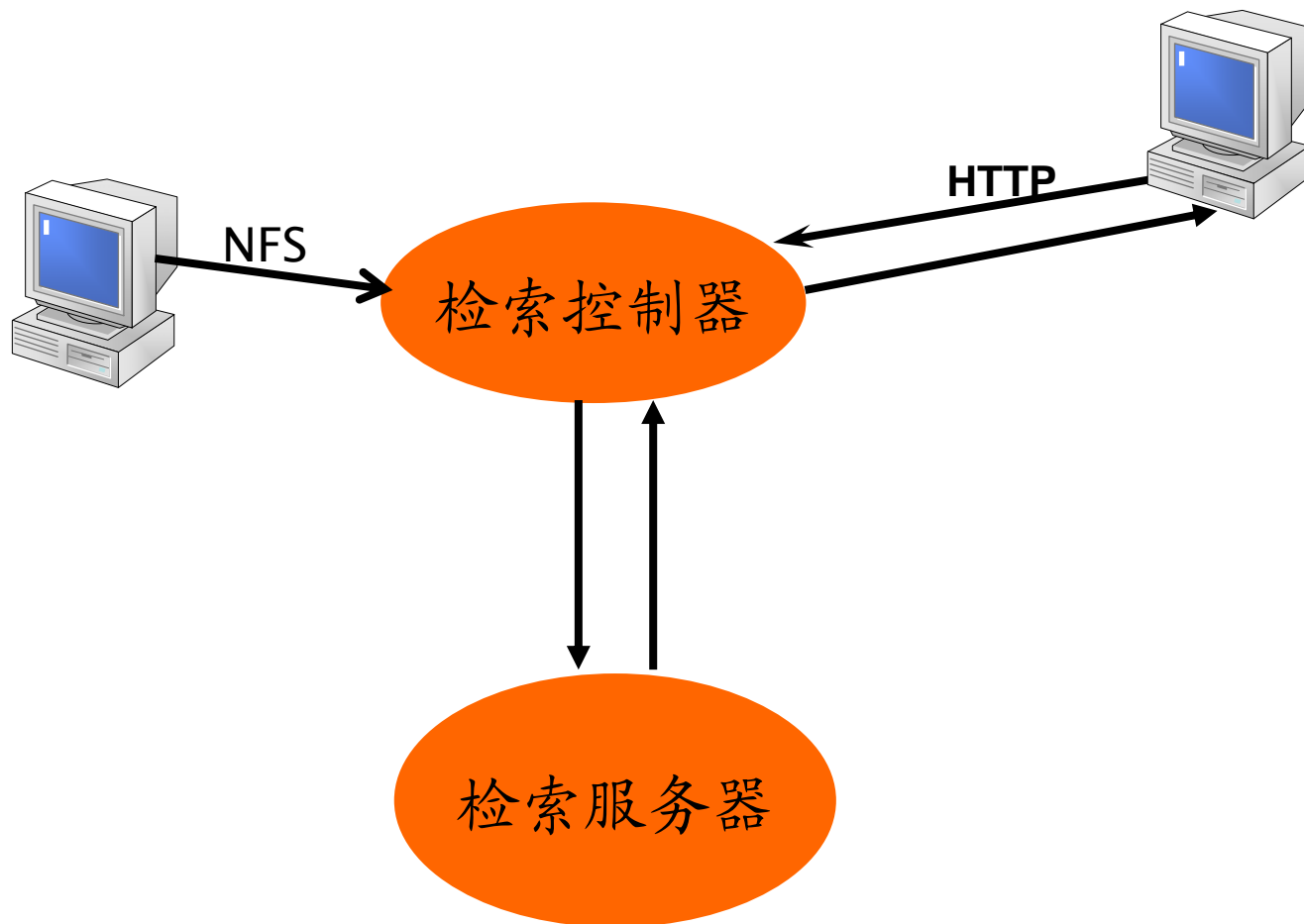
## ● 机器人搜索引擎

- 采集器
- 分析器
- 索引器
- 索引数据库
- 检索器
- 用户接口



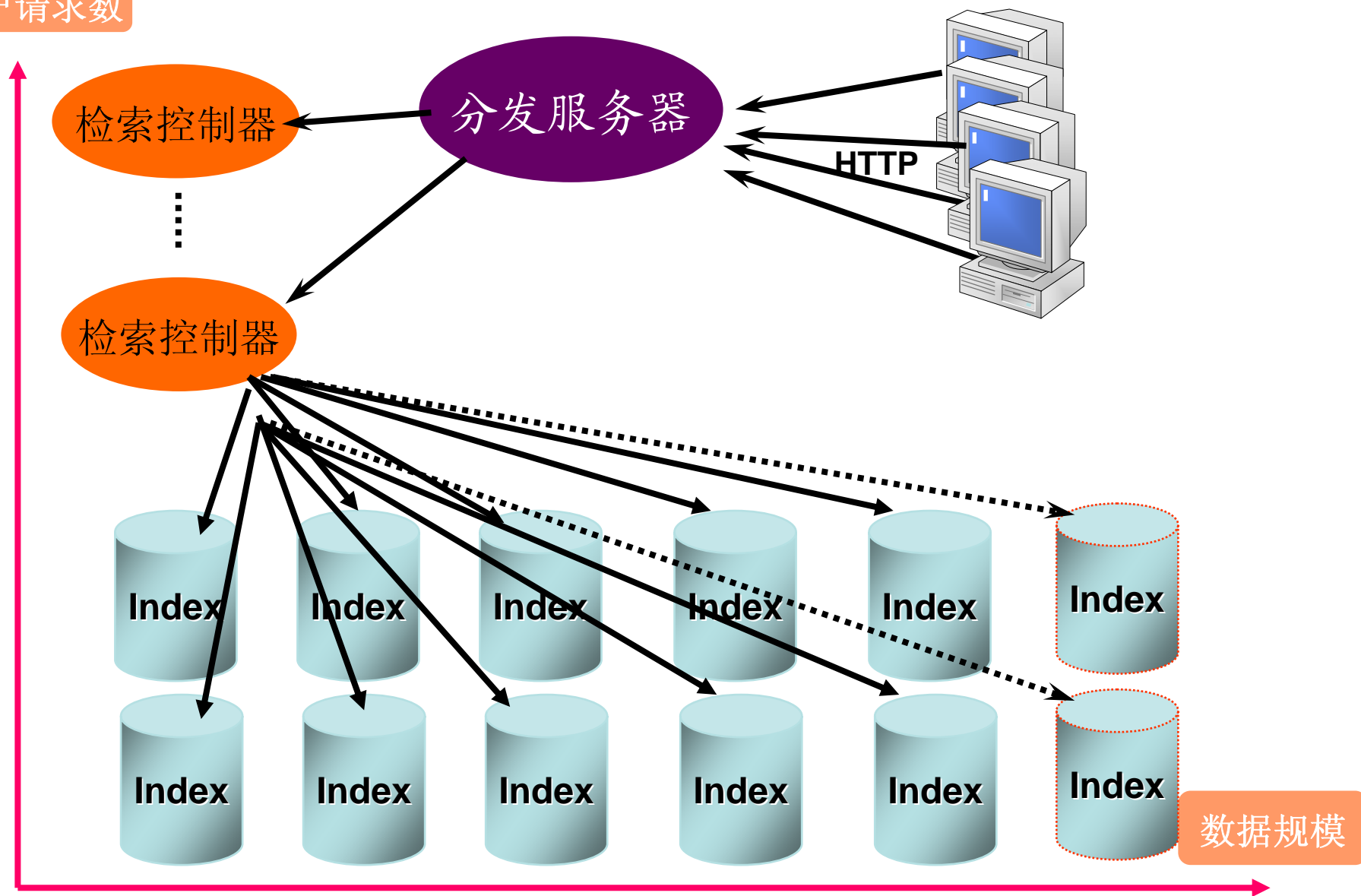


# 分布式搜索引擎拓扑结构



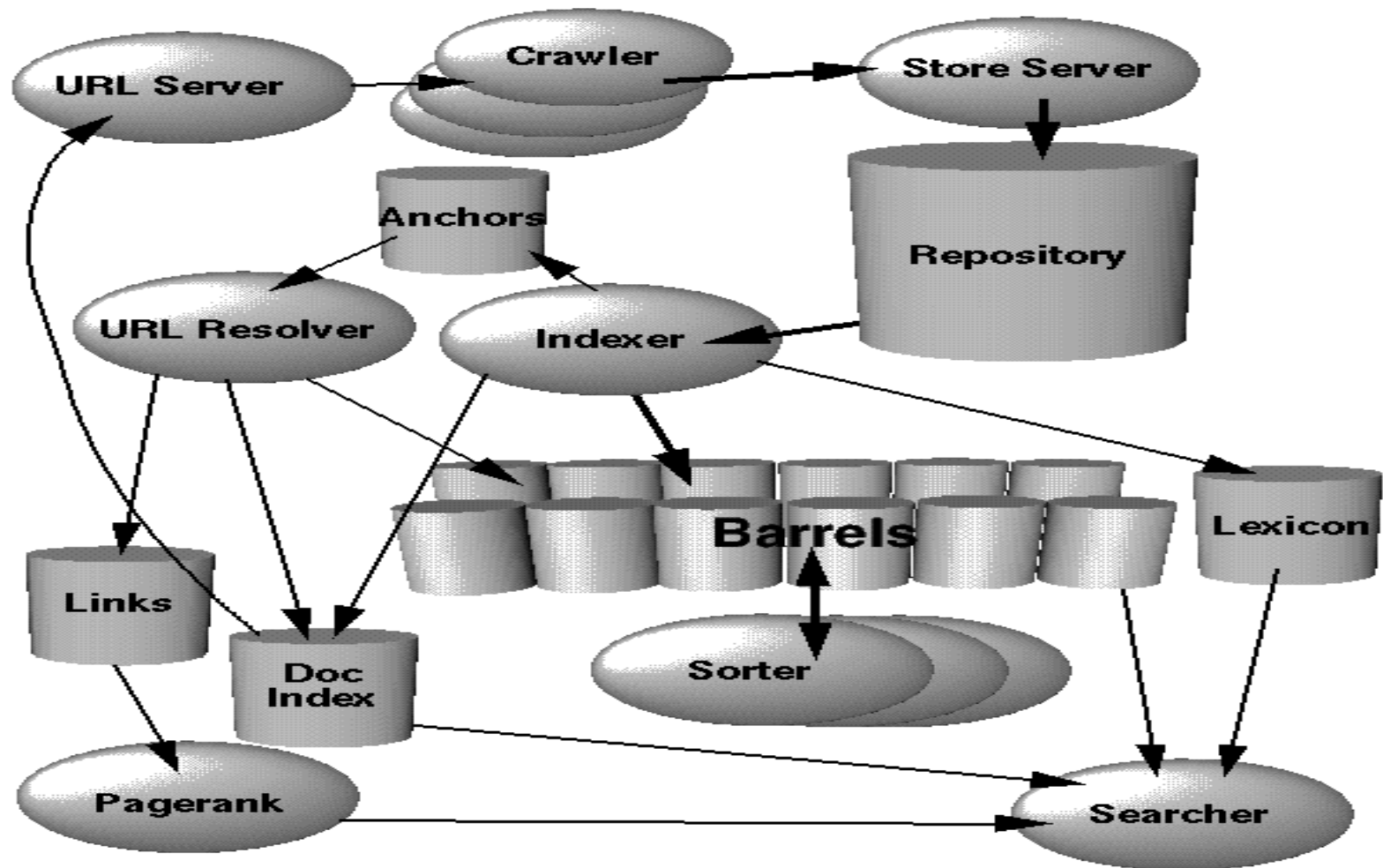


用户请求数





# 一个完整的Web搜索引擎





# 提纲

- 搜索引擎体系结构
- 小型搜索引擎—目标与功能 ◀
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题
- 参考文献



# 小型搜索引擎

## ● 目标

- 理解搜索引擎原理
- 实际搭建一个可运行的实验系统



# 小型搜索引擎

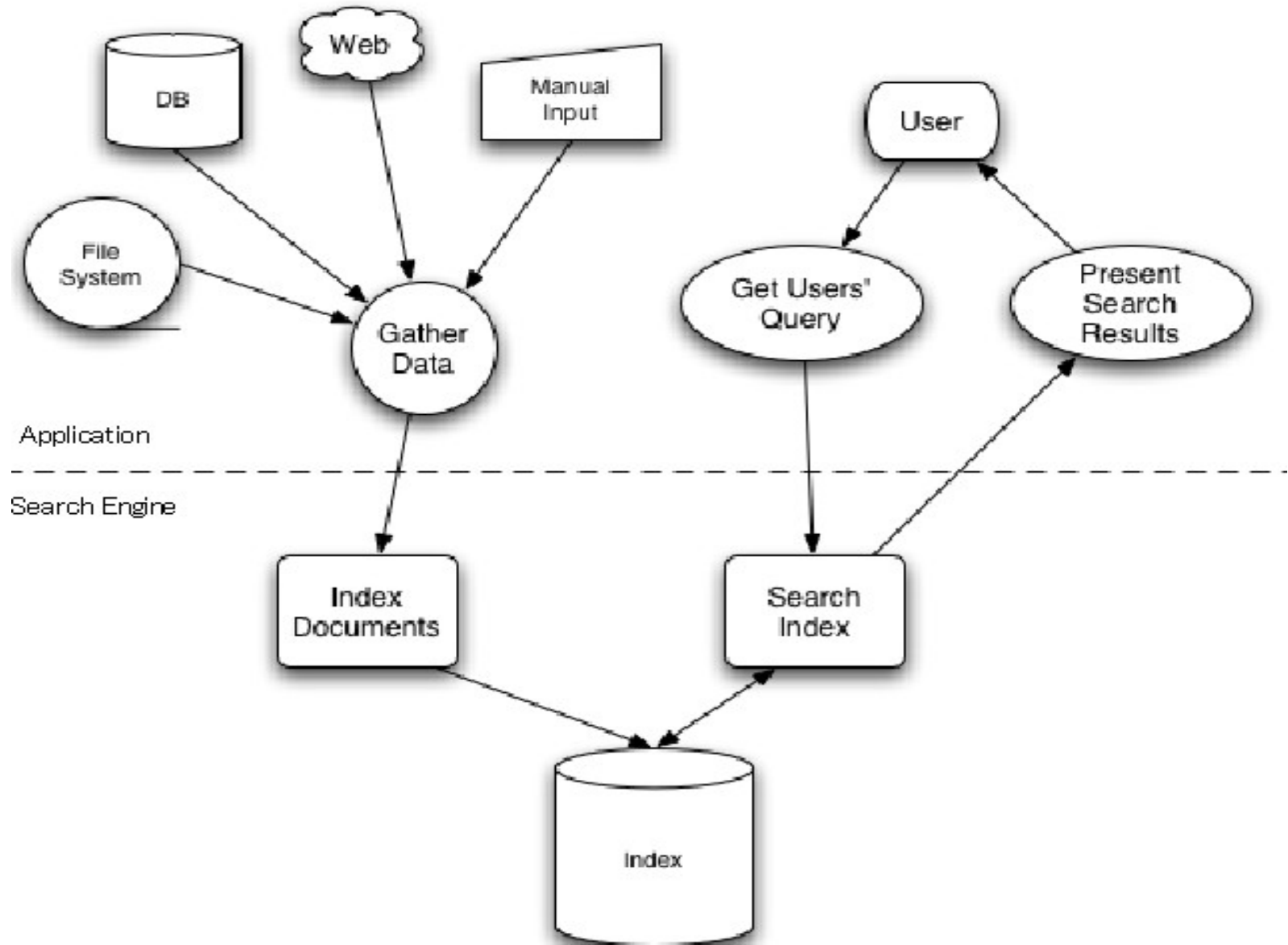
## ● 要实现的功能

- 采集
- 格式转换:HTML2TXT,PDF2TXT,DOC2TXT,...
- 词法处理:分词、词根还原
- 创建索引
- 全文检索



# 小型搜索引擎

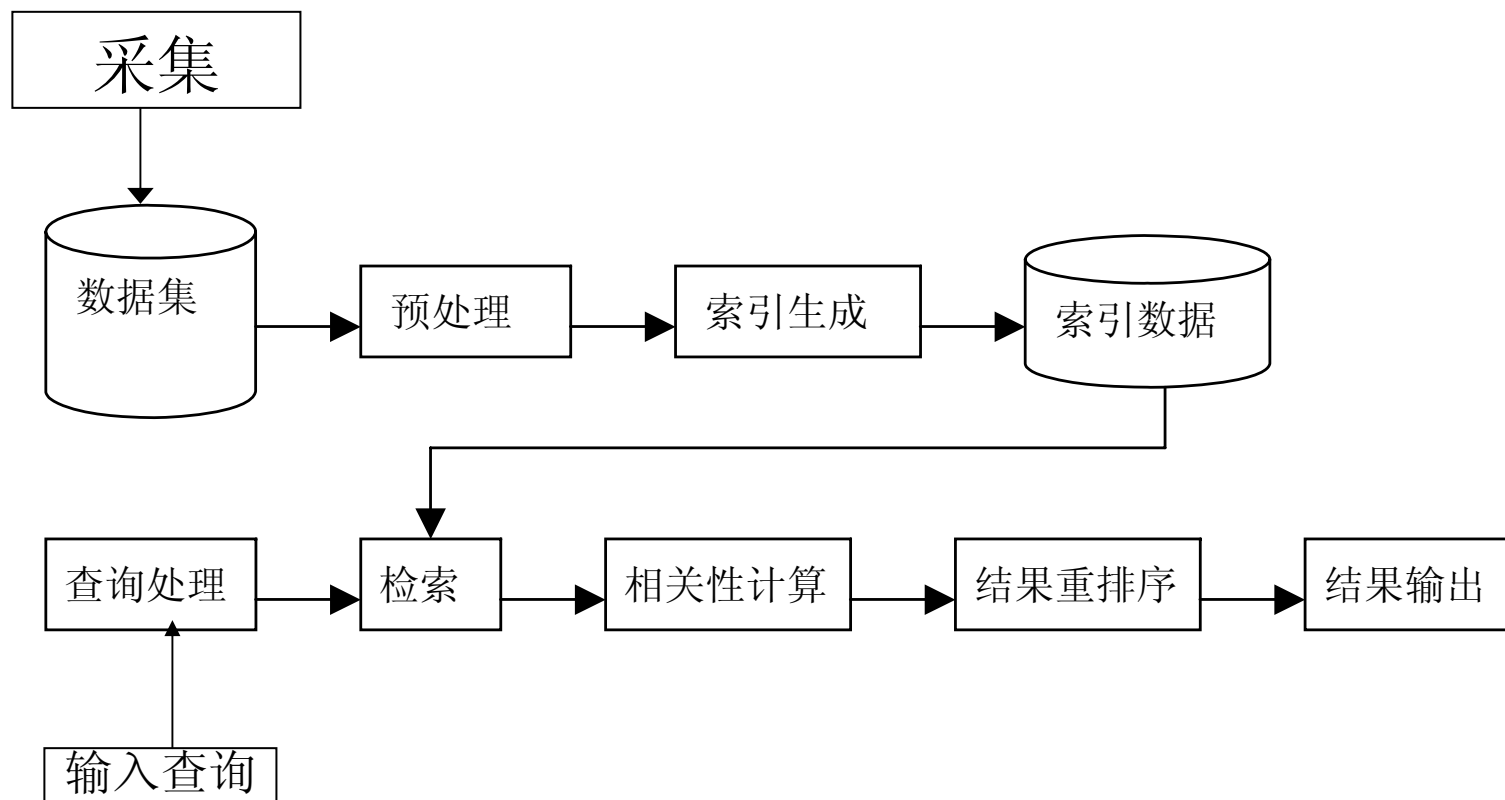
## ● 系统架构







# 简化的搜索引擎流程





# 提纲

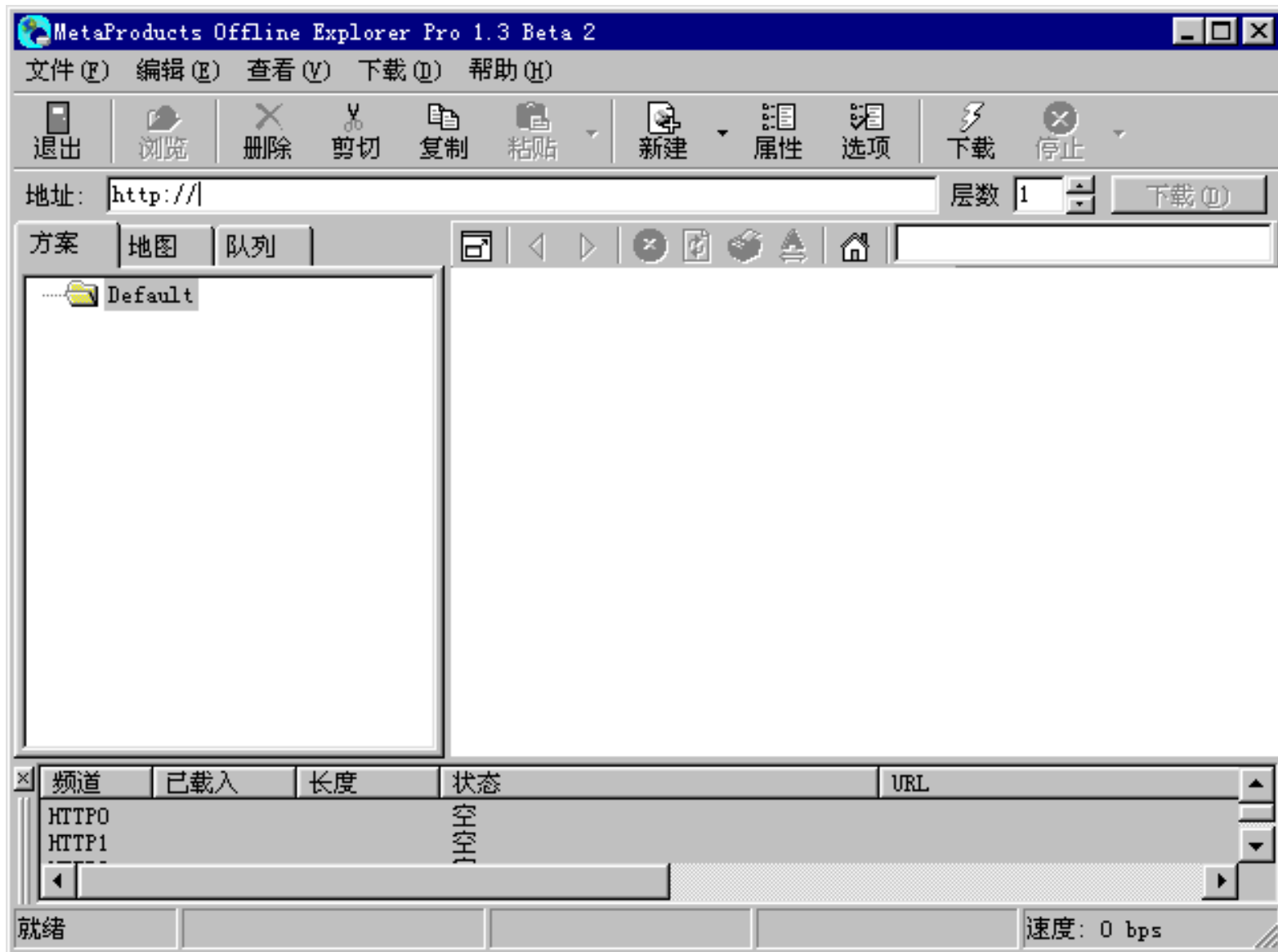
- 搜索引擎体系结构
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介◀
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题
- 参考文献



# 采集工具

## ● Offline Explorer

- 小型网页采集器
- 性质：商业软件
- 作者：Metaproducts
- 平台：Windows





# 采集工具

## ● wget

- 小型网页采集器
- 性质：GNU GPL
- 作者：Hrvoje Niksic
- 平台：Linux



# 采集工具

## 🌟 wget

- 用法
- `wget [option] [URL]`
- 举例：

```
wget -r -l4 -t0 http://www.ict.ac.cn/
```

```
wget -t 0 -w 31 -c -B
```

```
ftp://user@password:ftp.mtgroup.ict.ac.cn/smt -i  
filelist.txt -o down.log &
```



# 采集工具

## ● Larbin简介

- 全功能的网页爬虫
- 性质：GNU GPL
- 作者：Sébastien Ailleret
- 网址：<http://larbin.sourceforge.net/index-eng.html>
- 版本：V2.6.3
- 平台：Linux/FreeBSD
- 后续支持：Nutch



# 采集工具

## ● Larbin简介

### ■ 功能特点

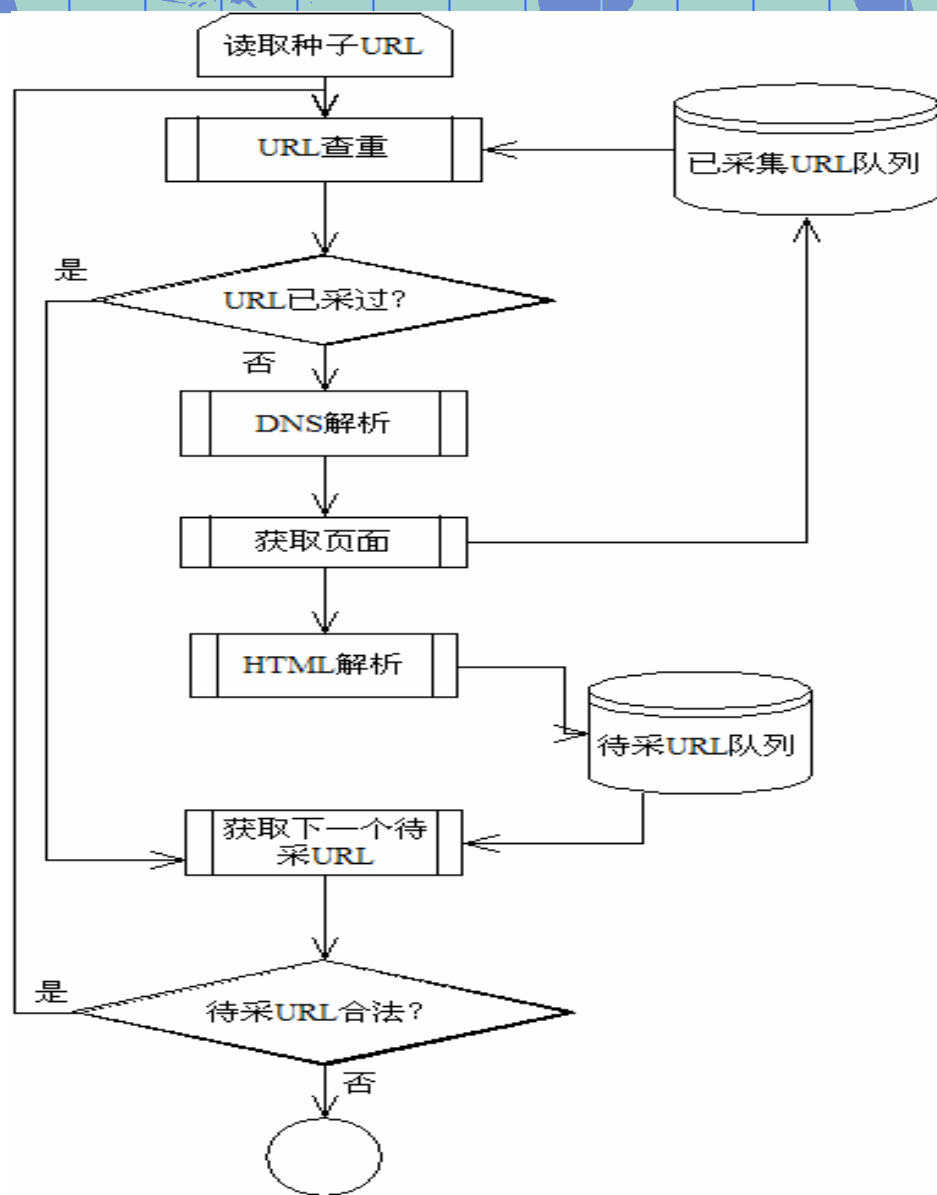
- 高速，低资源占用
- 从种子站点开始全网采集
- 支持远程输入站点采集
- 限定域/站点/文件类型采集
- 现场保存和续采
- 网页去重
- 状态检测与显示





# 采集工具

## ● Larbin采集流程





# 采集工具

## ● Larbin简介

- 运行

`./larbin [-c larbin_conf] [-scratch]`

- 结果存放

`save/dxxxx/fxxxx`  $\Leftrightarrow$  `index`

- 运行状态查看

`http://server_run_larbin:8081/`



# 采集工具

## ● 配置与部署

- 配置文件

- 外部环境变量设置: Larbin.conf

- 内部变量设置:

- options.h

- types.h



[Larbin.conf]

- From, UserAgent: robot协议
- httpPort, inputPort: 管理与交互
- pagesConnexions, dnsConnexions: 并发采集占用的网络资源
- depthInSite: 每个站点的采集深度
- noExternalLinks: 限定站点
- waitDuration: 采集间隔
- Proxy: 代理
- StartUrl: 种子Url
- limitToDomain: 限定域, 如cn, uk等
- forbiddenExtensions: 排除文件类型



[options.h]

- #define DEFAULT\_OUTPUT // do nothing...
- #define SIMPLE\_SAVE // save in files named save/dxxxxxxx/fyyyyyy
- #define MIRROR\_SAVE // save in files (respect sites hierarchy)
- #define STATS\_OUTPUT // do some stats on pages
- #define SPECIFICSEARCH //Set up a specific search
- #define FOLLOW\_LINKS //follow links in pages
- #define LINKS\_INFO //associate to each page the list of its sons
- #define NO\_DUP //suppress duplicate pages
- #define MAXBANDWIDTH 200000



[types.h]

- #define hashSize 64000000 //max number of urls that can be fetched
- #define timeoutPage 30 // default time out
- #define maxPageSize 100000
- #define saveDir "save/" // if we save files, where files are stored
- #define indexFile "index.html" //for MIRROR\_SAVE



# 提纲

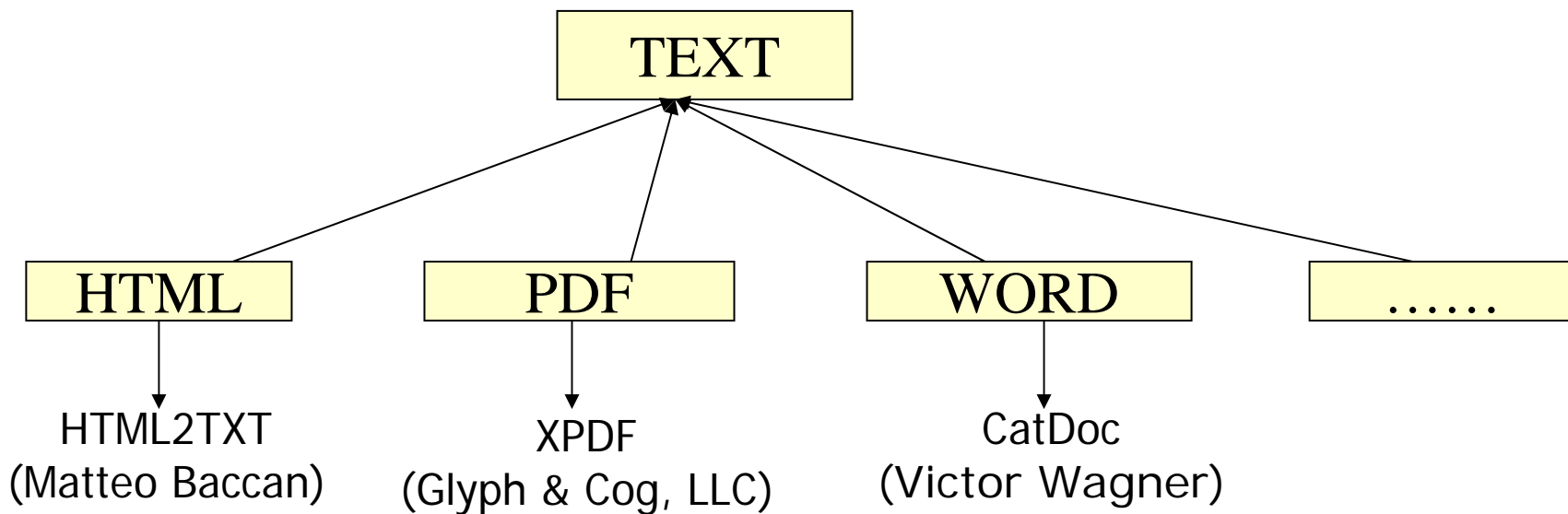
- 搜索引擎体系结构
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块◀
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题
- 参考文献





# 数据分析与预处理工具

## ● 格式转换工具





# 数据分析与预处理工具

## ● 格式转换工具

- HTML2TXT

接口: `void dfConvert( char * Buffer )`

- XPDF

接口: `class PDFDoc`

- CatDoc

接口: `int analyze_format(FILE *f)`



# 数据分析与预处理工具

## ● 中文分词工具—ICTCLAS

- 中文分词、命名实体识别工具
- 性质：自然语言处理开放资源许可证
- 作者：张华平
- 网址：[http://www.nlp.org.cn/project/project.php?proj\\_id=6](http://www.nlp.org.cn/project/project.php?proj_id=6)
- 平台：Windows/Linux
- SIGHAN第一届国际中文分词比赛Bakeoff第一名



# 数据分析与预处理工具

## ● 中文分词工具—ICTCLAS

### ■ 功能

- 中文分词
- 词性标注
- 中文命名实体识别：人名、地名、机构名



# 数据分析与预处理工具

## ● 中文分词工具—ICTCLAS

- API

初始化: `bool ICTCLAS_Init()`

运行:

`ICTCLAS_ParagraphProcess(char  
*sParagraph, char *sResult)`

退出: `bool ICTCLAS_Exit()`

- Configure.xml



# 数据分析与预处理工具

## ● 词根还原库—Snowball

- 用于IR的词根还原库
- 性质：开源
- 作者：Martin Porter
- 网址：<http://snowball.tartarus.org/index.php>
- 平台：Windows/Linux



# 数据分析与预处理工具

## 词根还原库—Snowball

### ■ API

#### ■ 初始化:

```
struct sb_stemmer * sb_stemmer_new(const char * algorithm,  
    const char * charenc)
```

#### ■ 运行

```
const sb_symbol * sb_stemmer_stem(struct sb_stemmer *  
    stemmer, const sb_symbol * word, int size)
```

#### ■ 退出

```
void sb_stemmer_delete(struct sb_stemmer * stemmer)
```



# 提纲

- 搜索引擎体系结构
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介◀
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题
- 参考文献





# 全文检索工具

## ● Lucene简介

- 完整、高效、易用、可扩展的IR库
- 性质：Apache License
- 作者：Doug Cutting
- 网址：<http://lucene.apache.org/>
- 版本：Lucene Java 1.9
- 平台：跨平台
- 支持：Apache Jakarta项目



# 全文检索工具

## ● Lucene特点

- 基于Java实现
- 可扩展、高性能索引工具：
  - $>20\text{M/s@P-M } 1.5\text{G}$
  - 动态内存  $\leq 1\text{M}$
  - 增量式索引
  - 索引数据大小  $\approx (20\% \sim 30\%) \times \text{源数据大小}$



# 全文检索工具

## ● Lucene功能

- 结果排序—最好结果优先
- 强大的查询表达式处理功能—短语、通配符、模糊查询等
- 分字段检索
- 指定日期范围检索
- 根据字段排序
- 支持多索引检索与结果合并
- 支持更新与检索同时进行



# 全文检索工具

## ● Lucene的其他语言版本

- [CLucene](#) - Lucene implementation in C++
- [dotLucene](#) - Lucene implementation in .NET
- [Lucene4c](#) - Lucene implementation in C
- [LuceneKit](#) - Lucene implementation in Objective-C (Cocoa/GNUstep support)
- [Lupy](#) - Lucene implementation in Python
- [NLucene](#) - another Lucene implementation in .NET (out of date)
- [Zend Search](#) - Lucene implementation in the Zend Framework for PHP
- [Plucene](#) - Lucene implementation in Perl
- [KinoSearch](#) - a new Lucene implementation in Perl
- [PyLucene](#) - GCJ-compiled version of Java Lucene integrated with Python via SWIG
- [MUTIS](#) - Lucene implementation in Delphi
- [Ferret](#) - Lucene implementation in Ruby



# 全文检索工具

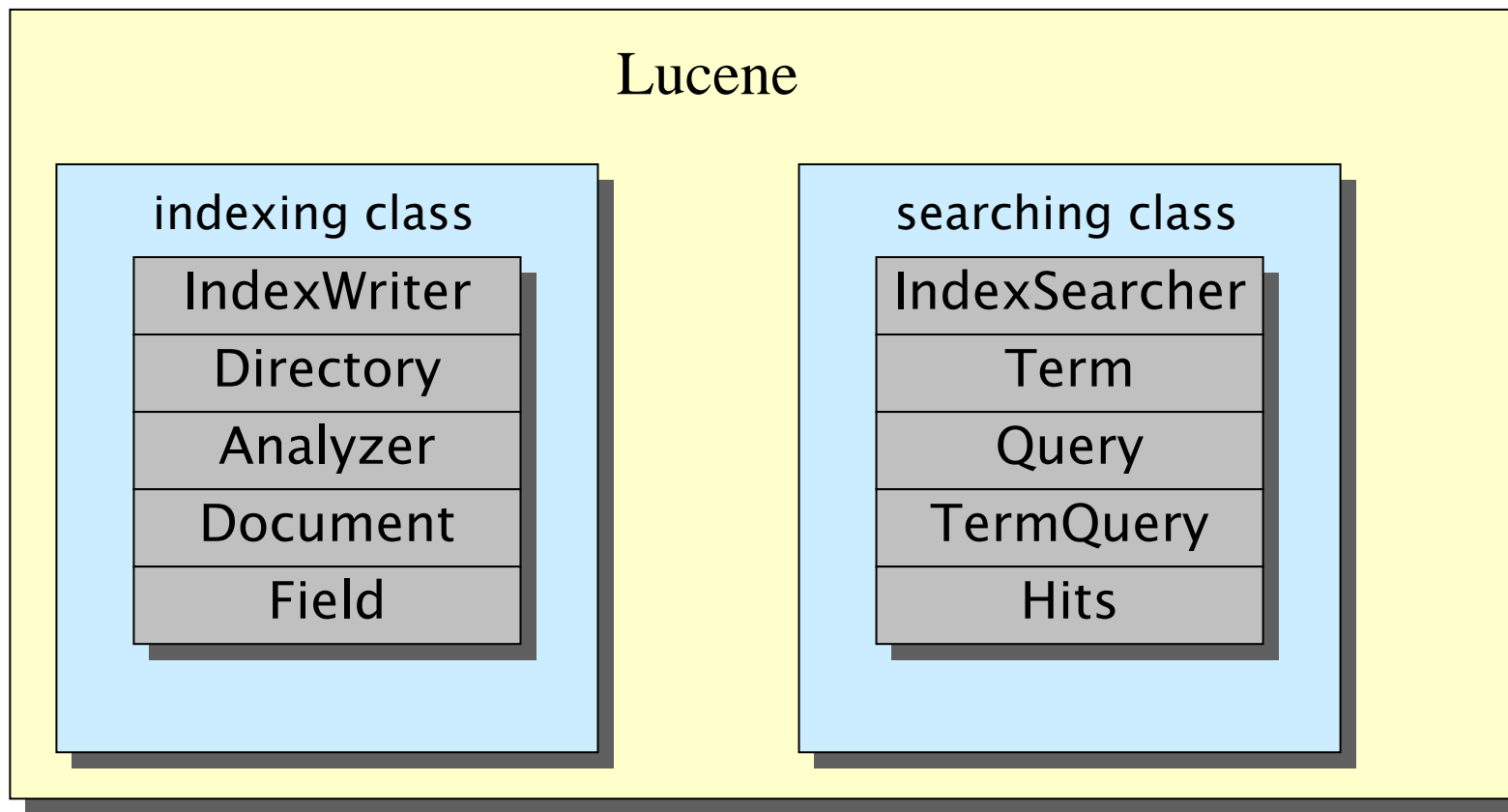
## ● Clucene简介

- Lucene的C++实现
- 作者: Ben van Klinken
- 协议: LGPL
- 网址: <http://clucene.sourceforge.net/>
- 平台: Windows/Linux/FreeBSD/...
- 版本: 0.9.11(unstable), 0.9.10(stable)
- 与Lucene接口基本一致
- 速度快于Lucene



# 全文检索工具

## ● Lucene结构





# 全文检索工具

## ● Lucene的核心接口

### ■ 索引器

- IndexWriter: 创建与更新索引数据
- Directory: Lucene索引的位置
- Analyzer: 从文本中提取要建索引的符号
- Document: 字段的集合，可视为虚拟的文档（网页、Email或文本）
- Field: 索引字段，用来建索引的基本单位



# 全文检索工具

## ● Lucene核心接口

### ⚙ Field使用说明

Field method/type	Analyzed	Indexed	Stored	Example usage
Field.Keyword(String, String) Field.Keyword(String, Date)		✓	✓	Telephone and Social Security numbers, URLs, personal names Dates
Field.UnIndexed(String, String)			✓	Document type (PDF, HTML, and so on), if not used as a search criteria
Field.UnStored(String, String)	✓	✓		Document titles and content
Field.Text(String, String)	✓	✓	✓	Document titles and content
Field.Text(String, Reader)	✓	✓		Document titles and content





# 简单示例一索引

[indexing implementation]

```
void IndexFiles(char* path, char* target)
```

```
{  
    StandardAnalyzer analyzer;  
    IndexWriter* writer = new IndexWriter(path, &analyzer);  
    indexDocs(writer, path);  
    writer->optimize();  
    writer->close();  
    delete writer;  
} // IndexFiles
```



# 简单示例一索引

[indexing implementation]

```
void indexDocs(IndexWriter* writer, char* directory)
```

```
{
    DIR* dir = opendir(directory);
    dirent* de = readdir(dir);
    while (de != NULL)
    {
        if (de.isdir())
        {
            indexDocs(writer, de->d_fullpath);
        }
        else {
            Document* doc = new Document();
            doc->add( *Field::Text(_T("path"), de->d_fullpath) );
            doc->add( *Field::Text(_T("contents"), de->d_text) );
            writer->addDocument(doc);
        }
        de = readdir(dir);
    }
}
```



# 全文检索工具

## ● Lucene的核心接口

### ■ 检索器

- IndexSearcher: 打开并在索引数据中查找
- Term: 用于检索的基本单位，形式为(字段，值)
- Query: 检索表达式类型
- TermQuery: 最基本的Query类型，返回在指定字段包含指定关键词的文档
- Hits: 检索结果集



# 简单示例一检索

[indexing implementation]

```
void SearchFiles(const char* path, char* userquery)
{
    IndexReader* reader = new IndexReader(path);
    StandardAnalyzer analyzer;
    Query* q = QueryParser::parse(userquery,_"contents",&analyzer);
    Hits* h = s->search(q);
    for (int i=0; i<h->length(); i++)
    {
        Document* doc = h->doc(i);
        printf("Result%d: %s\n", i, doc->get("path"));
    }//for
    delete reader;
} // IndexFiles
```



# 提纲

- 搜索引擎体系结构
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现 ◀
- 实用化问题
- 参考文献



# 小型搜索引擎的设计与实现

- 运行采集器
- 构造索引器
- 构造检索服务器
- 构造用户交互接口



# 小型搜索引擎的设计与实现

## ● 运行采集器

### ■ 典型设置

```
[larbin.conf]
```

```
From luoweihua@ict.ac.cn
```

```
UserAgent larbin_2.6.3
```

```
#noExternalLinks
```

```
waitDuration 30
```

```
depthInSite 5
```

```
startUrl http://www.sina.com.cn/
```



# 小型搜索引擎的设计与实现

## 运行采集器

### 典型设置

[options.h]

```
#define SIMPLE_SAVE
```

```
#define FOLLOW_LINKS
```

```
#define NO_DUP
```

```
#define MAXBANDWIDTH 200000
```

[types.h]

```
#define saveDir "save/"
```





# 小型搜索引擎的设计与实现

## ● 运行采集器

- ❏ `./configure`

- ❏ `gmake`

- ❏ `./larbin`



# 小型搜索引擎的设计与实现

## ● 构造索引器

① Preprocess Webpages(Web Page Path, New Data Path, DocId);

② Build\_Index\_By\_Lucece\_IndexWriter(New\_Data\_Path);



# 小型搜索引擎的设计与实现

## ❶ 数据预处理

```
void Preprocess_Webpages(string Web_Page_Path, string New_Data_Path, uint& DocId)
{
    DIR* dir = opendir(Web_Page_Path);
    dirent* de = readdir(dir);
    while (de != NULL)
    {
        if (de.isdirectory())
        {
            Preprocess_Webpages(de.getfullpath(), New_Data_Path, DocId);
        }else{
            Preprocess_EachPage(de.getfullpath(),New_Data_Path,DocId);
        }
        de = readdir(dir);
    }
    closedir(dir);
}
// Preprocess_Webpages
```



# 小型搜索引擎的设计与实现

## ● 数据预处理

```
Preprocess_EachPage(string Page_Path, string New_Data_Path, uint& DocId)
{
    int Type = GetFileType(Page_Path);
    switch(Type)
    {
        case HTML: Htm2Txt(GetFileText(Page_Path), NewText);
                    break;
        case PDF: Pdf2Txt(GetFileText(Page_Path), NewText);
                    break;
        case DOC: Doc2Txt(GetFileText(Page_Path), NewText);
                    break;
        default: break;
    }
    string SegText;
    ICTCLAS_Paragraph(NewText, SegText);
    SnowBall(SegText);
    StoreResult(SegText, DocId++);
} // Preprocess_Webpages
```



# 小型搜索引擎的设计与实现

## ● 构造检索服务器

```
void SearchServer()
{
    InitServer();
    StartSocket(server_socket);
    while (true)
    {
        socket client_socket = accept(server_socket);
        beginthread(Search Daemon, client_socket);
    } //while
    CloseServer();
} //SearchServer
```



# 小型搜索引擎的设计与实现

## 检索

```
void Search_Daemon(socket client_socket)
{
    GetUserQueryBySocket(buffer);
    DecodeUserQuery(buffer, UserQuery);
    Search_Result_By_Lucene(UserQuery, AllHits);
    EncapsulateResult(AllHits, Result_String);
    ReturnResult(Result_String, client_socket);
} // Search_Daemon
```



# 构造用户交互接口

- 用户接口
  - Web浏览器界面
  - 用户查询请求处理程序



# 构造用户交互接口

## ● Web浏览器界面

### ■ 输入查询表达式

搜索结果

包含以下**全部**的字词

包含以下**任何**一个字词

**不包括**以下字词

10 项结果

搜索





# 构造用户交互接口

- Web浏览器界面

- 结果输出



[网页](#) [图片](#) [资讯](#) [论坛](#) [网页目录](#) [更多»](#)



A Survey on Web Information Retrieval Technolo

搜索

[高级搜索](#)  
[使用偏好](#)

☒ 搜索所有网页 ☐ 搜索所有中文网页 ☐ 搜索简体中文网页

网页

约有3,150,000项符合A Survey on Web Information Retrieval Technologies的查询结果，以下是第1-10项。（搜索用时 0.15 秒）

### [A Survey on Web Information Retrieval Technologies](#) 的图书搜索结果



[Advanced Web Technologies and ...](#) - 936 页

[Web Technologies Research and ...](#) - 1082 页

### [A Survey On Web Information Retrieval Technologies - Huang ...](#) - [ [翻译此页](#) [BETA](#) ]

Retrieving **information** from the **Web** is becoming a common practice for internet users. However, the size and heterogeneity of the **Web** challenge the effectiveness of classical **information retrieval** techniques. In this paper, we report the ...

[citeseer.ist.psu.edu/336617.html](#) - 26k - 2006年5月15日 - [网页快照](#) - [类似网页](#)

### 文档 [A Survey On Web Information Retrieval Technologies](#) 上载者 ...

文档, [A Survey On Web Information Retrieval Technologies](#). 上载者, sunjian. 原作者, Lan Huang (State University of New York at Stony Brook). 加入时间, 2002-09-03 17:11:05. 文档介绍 ...

[www.nlp.org.cn/docs/download.php?doc\\_id=78-3k](#) - [网页快照](#) - [类似网页](#)

### [中文自然语言处理开放平台](#)

垃圾邮件”的智能分析、过滤和Rough集讨论(1664 Downloads, 2002-10-08 12:35:56); [A survey of IR and filtering Methods](#)(904 Downloads, 2002-09-03 19:07:26); [A Survey On Web Information](#)



# 构造用户交互接口

## ● 用户查询请求处理程序

- CGI或脚本语言（ASP,PHP,JSP,etc）

- 功能

- ❶ 获取用户查询式：把用户通过Form输入的查询语句封装发送给检索服务器
- ❷ 显示结果：从检索服务器获取结果，缓存并分页呈现给用户



# 构造用户交互接口

## ❶ 获取用户查询

```
$Query_and = $_Get["query_and"];
```

```
$Query_or = $_Get["query_or"];
```

```
$Query_not = $_Get["query_not"];
```

```
$Total_Query = GenerateTotal($Query_and,$Query_or,$Query_not);
```

```
$ClientSock = new ClientSocket($Address, $Port);
```

```
$ClientSock->send($Total_Query);
```



# 构造用户交互接口

## ② 显示结果

```
$Result_String = $ClientSock->receive();  
$AllResults = DecodeResult($Result_String);  
Foreach ($AllResults as $EachResult)  
{  
    Echo "DocId : " . $EachResult->docid . "<br>";  
    Echo "DocPath:" . $EachResult->docpath . "<br>";  
} //Foreach
```



# 提纲

- 搜索引擎体系结构
- 小型搜索引擎—目标与功能
- 采集工具Larbin简介
- 数据分析与预处理工具—分词工具ICTCLAS及相关格式处理模块
- 全文检索工具Lucene简介
- 基于开源工具包的小型搜索引擎的设计与实现
- 实用化问题◀
- 参考文献



# 实用化问题

- 系统架构
- 数据更新策略
- 格式识别与转换
- 多语种识别与转换
- 有针对性的优化策略
- 结果排序与优化
- 负载均衡
- 缓存策略
- .....



# 参考文献

- Sergey Brin and Lawrence Page. The Anatomy of a Large-Scale Hypertextual Web Search Engine. <http://www-db.stanford.edu/~backrub/google.html>
- Lan Huang. A Survey on Web Information Retrieval Technologies. <http://citeseer.ist.psu.edu/336617.html>
- cneie.com. 中文搜索引擎技术揭密：系统架构. <http://www.cneie.cn/news/20064/2006429161655.html>
- 卢亮. Larbin 一种高效的搜索引擎爬虫工具. <http://www.wespoke.com/archives/000879.php>
- Martin Porter. snowball manual. <http://snowball.tartarus.org/compiler/snowman.html>
- Doug Cutting, Otis Gospodnetic , etc. Overview of Apache Lucene. <http://lucene.apache.org/java/docs/>
- ERIK HATCHER, OTIS GOSPODNETIC. Lucene in Action. Manning Publications Co. Greenwich CT.





谢谢！