

“搜主义”分布式网络资源检索系统使用说明

系统名称：搜主义

开发团队：我狠低调 团队

所属院校：中国石油大学（华东）

计算机与通信工程学院

团队成员：

组长：张海滨

组员：叶瑞森

吴承桀

曹宝琴

高庆金

指导老师：卢清华

# 目录

[“搜主义”分布式网络资源检索系统使用说明 1](#_Toc398454097)

[目录 1](#_Toc398454098)

[一．编写目的 1](#_Toc398454099)

[二．项目背景 1](#_Toc398454100)

[三．运行环境 2](#_Toc398454101)

[四. 环境配置 3](#_Toc398454102)

[4.1 tomcat配置 3](#_Toc398454103)

[4.1.1 配置server.xml 3](#_Toc398454104)

[4.2 test.xml配置 4](#_Toc398454105)

[4.2.1 注意 4](#_Toc398454106)

[4.2.2 说明 4](#_Toc398454107)

[4.2.3 配置test.xml 4](#_Toc398454108)

[4.3 openoffice安装 6](#_Toc398454109)

[4.3.1 安装目录 6](#_Toc398454110)

[4.3.2 安装过程 6](#_Toc398454111)

[4.4 swf安装 9](#_Toc398454112)

[4.4.1 安装目录 9](#_Toc398454113)

[4.4.2 安装过程 9](#_Toc398454114)

[五．使用说明 10](#_Toc398454115)

[5.1 注意 10](#_Toc398454116)

[5.2 项目发布 10](#_Toc398454117)

[5.3 服务器启动 10](#_Toc398454118)

[5.4 开始使用 12](#_Toc398454119)

[5.4.1 输入ip地址进入首页 12](#_Toc398454120)

[5.4.2 全部搜索 12](#_Toc398454121)

[5.4.3 图片搜索 14](#_Toc398454122)

[5.4.4 音乐搜索 16](#_Toc398454123)

[5.4.5 视频搜索 17](#_Toc398454124)

[5.4.6 文档搜索 18](#_Toc398454125)

[5.4.7 高级搜索 21](#_Toc398454126)

[5.4.8 其它 23](#_Toc398454127)

[5.5 服务器注销 24](#_Toc398454128)

[5.5.1 正常注销退出集群 24](#_Toc398454129)

[5.5.2 服务器异常宕机 24](#_Toc398454130)

# 一．编写目的

编写本使用说明的目的是充分叙述本软件所能实现的各个模块的功能及其运行环境，以便使用者了解本软件的使用范围和使用方法。

# 二．项目背景

项目名称：分布式网络资源检索

项目提供者：齐鲁软件设计大赛组委会

开发团队：我狠低调 团队

开发人员：张海滨、叶瑞森、吴承桀、曹宝琴、高庆金

所属学校：中国石油大学（华东）

# **三．运行环境**

操作系统: windows 7

JDK: jdk1.6

TOMCAT: tomcat 6

端口: 8080

test.xml文件位置: %TOMCAT\_HOME%\webapps\test.xml

资源文件位置: %TOMCAT\_HOME%\webapps\resources

# **环境配置**

## tomcat配置

### 4.1.1 配置server.xml

#### 4.1.1.1 文件位置

%TOMCAT\_HOME%\conf\server.xml

#### 4.1.1.2 配置

对 Connector 元素添加属性 URIEncoding , 属性值为 utf-8 。

如下所示：

< Connector

port="8080" protocol

protocol="HTTP/1.1"

connectionTimeout="20000"

redirectPort="8443"

URIEncoding="UTF-8"

/>

## test.xml配置

### 4.2.1 注意

**test.xml文件中请勿包含该符号： &**

### 4.2.2 说明

每个集群只有一台是主服务器，每台服务器上的test.xml文件上根节点的属性，如果是本台服务器是主服务器，则remote,和host的值都是本机的IP；如果是从服务器，则host是本机的IP，remote是远程主服务器的IP。

### 4.2.3 配置test.xml

#### 4.2.3.1 文件位置

%TOMCAT\_HOME%\webapps\test.xml

#### 4.2.3.2 配置

对 allresource 元素添加属性 host 和 remote，

host 属性值为：本机IP地址

remote 属性值为：集群主服务器的IP地址

例如：

本机IP：192.168.0.2

主机（集群主服务器）IP：192.168.0.1

则：

<?xml version="1.0" encoding="UTF-8"?>

<allresource host=”192.168.0.2” remote=”192.168.0.1”>

<resourceitem>

                <title>07各地高考作文题</title>     
                <keywords>作文，高考</keywords>   
                <kind>doc</kind>   
                <describe>07各地高考作文题</describe>   
                <date>2007-12-1 8:50:43</date>     
                <url>http://192.168.0.122:8080/resources/07各地高考作文题.doc< /url>     
                <author>李明</author>    
                <publisher>山东师大</publisher>     
        </resourceitem>

<resourceitem>

…

</resourceitem>  
   </allresource>

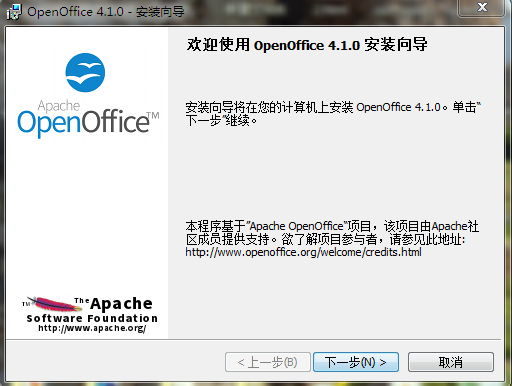
## openoffice安装

### 4.3.1 安装目录

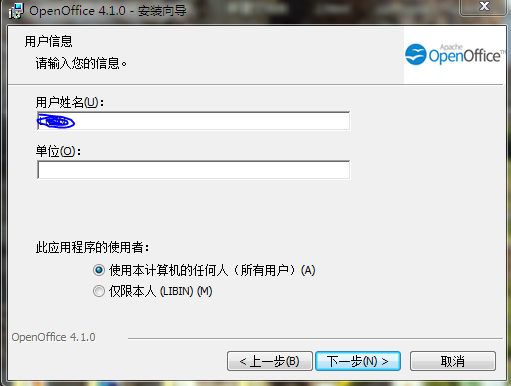
C:\OpenOffice

### 4.3.2 安装过程

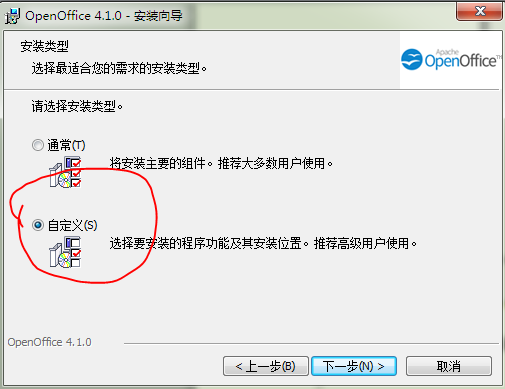
1. 打开Openoffice文件夹。双击setup.exe
2. 点击“下一步”



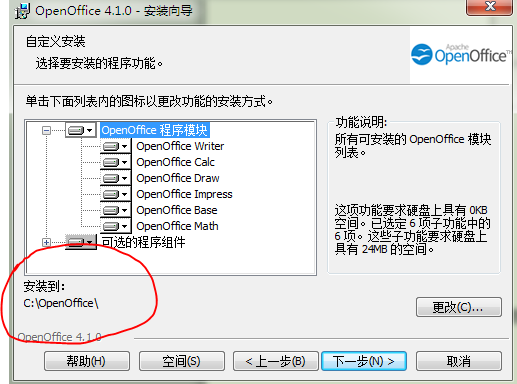
1. 直接点击“下一步”



1. 选择自定义



1. 更改安装路径至：C:\OpenOffice



1. 直接点击下一步



7. 点击“安装”

8. 完成

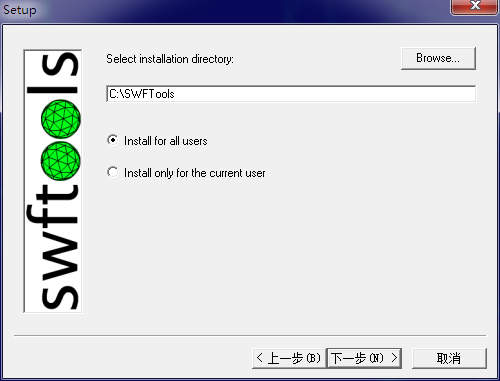
## swf安装

### 4.4.1 安装目录

C:\SWFTools

### 4.4.2 安装过程

1. 打开SwfTools文件夹。双击 “swftools-0.9.2.exe”
2. 点击下一步
3. 更改安装路径至：C:\SWFTools



1. 连续点击下一步或完成直至安装完成。

# 五．使用说明

## 5.1 注意

**测试的前提是局域网内每台服务器都能够互相访问， ping 都能够正常。**

## 5.2 项目发布

将项目部署文件夹下的DRRS.war包拷贝到%TOMCAT\_HOME%\webapps目录下，即完成发布。

## 5.3 服务器启动

假设五台服务器参与测试，IP地址分别为：192.168.0.1、192.168.0.2、192.168.0.3、192.168.0.4、192.168.0.5。其中选192.168.0.1充当集群主服务器，其余为普通的资源服务器。

首先先配置电脑IP，然后在五台服务器test.xml中根节点的属性分别配置如下：

**主服务器test.xml**

<allresource remote="192.168.0.1" host="192.168.0.1">

<resourceitem>

**从服务器1 test.xml**

<allresource remote="192.168.0.1" host="192.168.0.2">

<resourceitem>

**从服务器2 test.xml**

<allresource remote="192.168.0.1" host="192.168.0.3">

<resourceitem>

**从服务器3 test.xml**

<allresource remote="192.168.0.1" host="192.168.0.4">

<resourceitem>

**从服务器4 test.xml**

<allresource remote="192.168.0.1" host="192.168.0.5">

<resourceitem>

主服务器首先启动，随后其他服务器陆续启动。主服务器启动后将每隔一定时间监听其他服务器宕机情况。如果有服务器宕机，主服务器发现后，将更新自己和其他服务器的注册表。

主服务器正常退出，将选举一台作为主服务器，继续能为用户提供查询任务。

用户打开浏览器，登入任意一台主机，例如

<http://192.168.0.1:8080/DRRS/>

就进入了主页，即可开始查询。

## 5.4 开始使用

(为保证最佳浏览效果，建议在谷歌浏览器中使用)

### 5.4.1 输入ip地址进入首页

打开浏览器，输入地址http://ip:8080/DRRS，如 <http://localhost:8080/DRRS>，进入到搜索主页



### 5.4.2 全部搜索

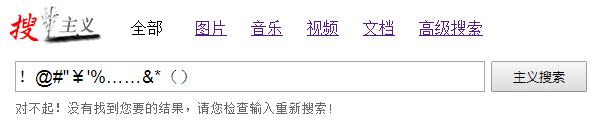
全部搜索即综合搜索，根据您键入的关键进行全部资源的搜索匹配，结果中包含文档、图片、音乐、视频等所有格式的资源



结果资源的排序按照与关键字匹配度进行排序，排序越靠前说明匹配度越高。对于各种格式的资源，在点击搜索结果后会执行对应文件应有的操作，如点击文档类的结果，会自动进行文档预览，对视频或音乐结果，会进行播放，对图片会进行大图浏览，详细操作后面分模块介绍。

另外，如图所示，搜索结果显示页面中搜索框下面会有提示找到结果的条数和执行查询所消耗的时间。

如果没有找到结果也会有如图的提示

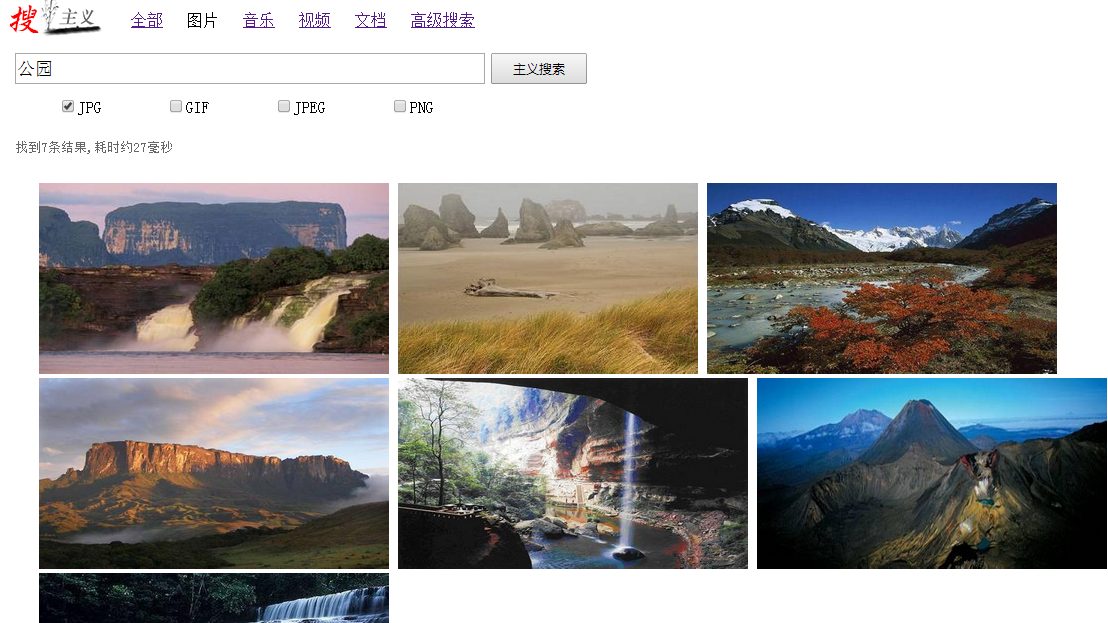


### 5.4.3 图片搜索

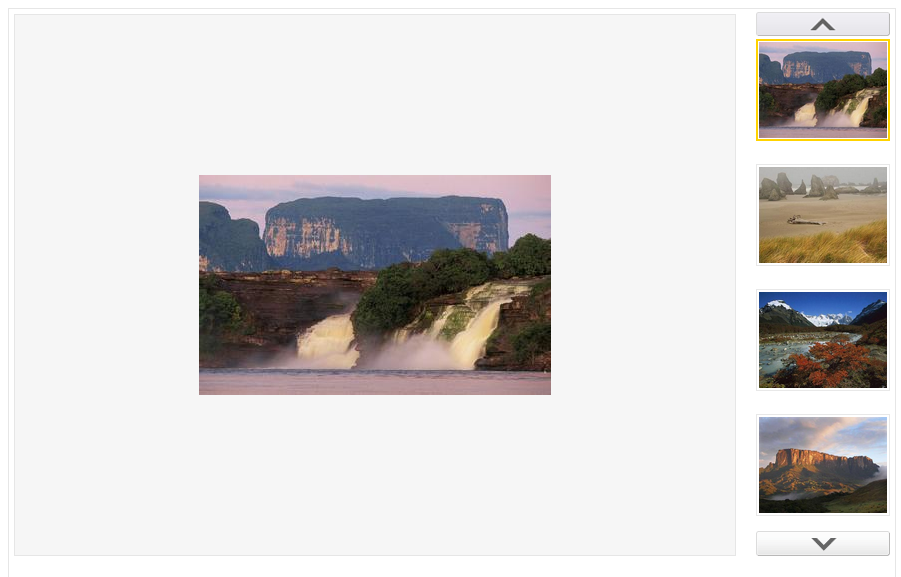
点击导航菜单里面的“图片”菜单跳到图片检索主页



在这个页面中进行检索时会有格式选择多选框，您可以选择您想要的格式进行针对行搜索。如果不选格式，则默认为所有格式的图片。



在结果中，如果您点击任何一张图片，会跳转到大图浏览模式



您可以在大图中根据鼠标箭头方向，点击显示下一张或上一张，也可以在右部分导航菜单中选择想要浏览的图片。

### 5.4.4 音乐搜索

点击导航“音乐”跳转到音乐搜索主页面



搜索与图片搜索类似，就不再复述了。



在结果界面中，简化了一切不必要的装饰，留住最核心的功能，一目了然，只需在结果中一点，就能实现播放！从首页到听到想要的歌曲，只需点击2次鼠标，简单，才最美！

### 5.4.5 视频搜索

视频搜索首页



根据关键字与格式得到结果



这只是一个结果展示界面，需要点击链接，即可实现播放---->



### 5.4.6 文档搜索

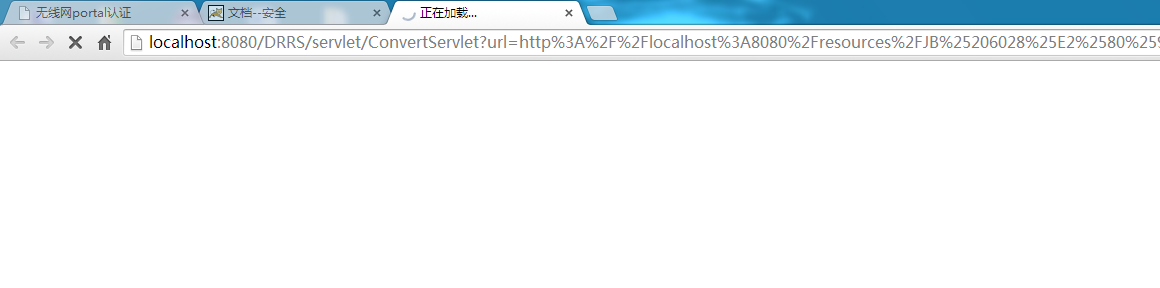
进入到文档搜索首页



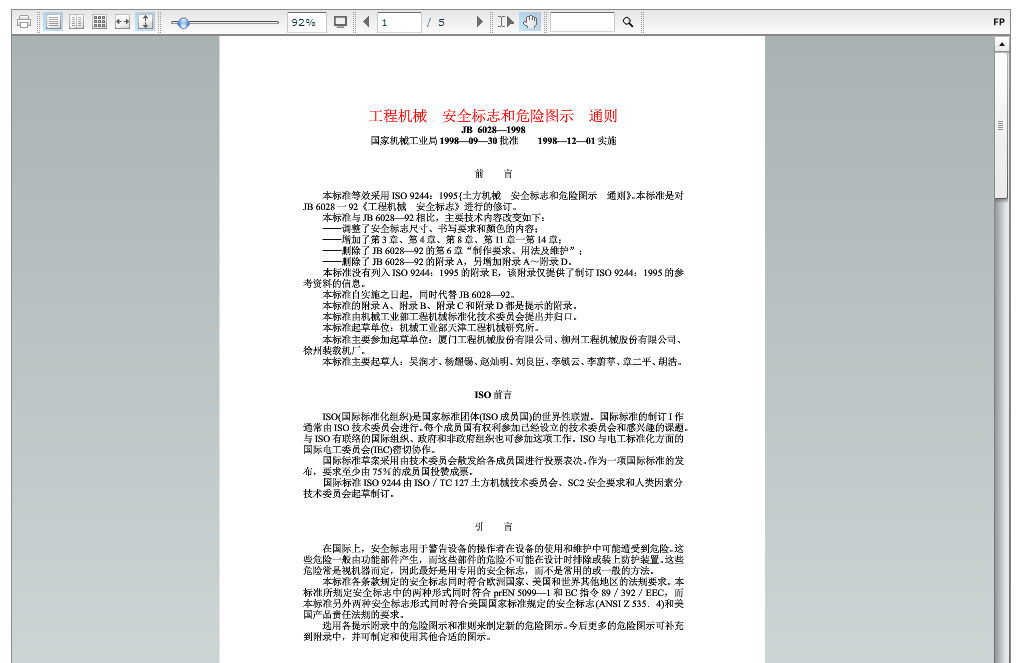
与其它搜索不同的是，文档预览分为普通检索和全文检索，普通检索即根据输入的关键字和xml里面所描述的信息进行匹配得出结果，全文检索则是在普通检索的基础上，再对文章的内容进行关键字匹配得出结果。



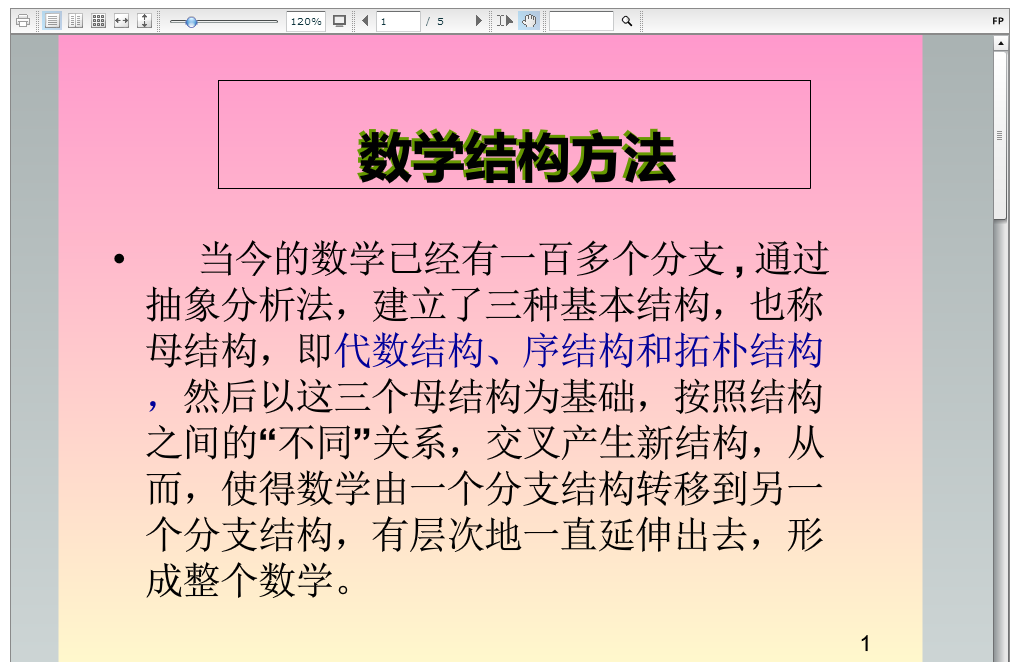
点击结果，则进入行文档在线预览



对于文件比较大的word或ppt在线预览可能会转化的时间比较长，出现这个界面时说明后台正在进行文档转化，请耐心等待….



终于出来结果了，预览比例是前5页，在这个界面您可以放大缩小界面等操作，根据顶部菜单进行相应的操作。以下是对ppt进行预览



### 5.4.7 高级搜索

进入到高级搜索界面



输入想要搜索的信息，进行精确检索。高级搜索将不会对关键字进行分词操作，所以只有和输入信息完全匹配才展示结果。



接下来的操作与全部搜索结果类似。如果要返回重新搜索请点击页面“不满意，重新搜索”连接返回到高级搜索主页。

### 5.4.8 其它

1.对于输入的关键字太长，系统会自动截取前38个字符，所以输入前请控制语句长度。



2.对于未输入任何有效字符，即未输入或者输入全为空格，则自动刷新页面。

## 5.5 服务器注销

### 5.5.1 正常注销退出集群

主服务器退出集群，将选举一台为新的主服务器，同时更新所有服务器的注册表，最后主服务器还将会更改test.xml中的根节点属性，将remote属性值改为最新选举出的IP，方便下次加入时无需再配置test.xml。

从服务器退出集群，将向主服务器提出注销请求，主服务器将会更新自己注册表和通知其他服务器的注册表。

### 5.5.2 服务器异常宕机

默认主服务器不存在异常宕机，此功能由于时间原因，有待完善，不过思路是这样的，当服务器宕机的时候，其他服务器坚持到主服务器宕机后，将根据注册表，从中选出IP最大的那个作为主服务器，因为每台服务器上的注册表IP都是一样的，且最大的IP也是确定的。

从服务器发生宕机时，主服务器将定时监测到，此时，主服务器首先将更新自己的注册表，同时也将通知其他服务器更新自身的注册表，要是此时有查询任务，因为主服务器监测机制问题，及每隔一定时间监测，以至于不能及时发现，导致查询可能会出现异常，时间可能会稍微慢一点，但是只要主服务器发现了宕机情况，查询功能则恢复正常效率。