

# Google搜索语法

# 本课程仅用于教学用途

中华人民共和国网络安全法

中华人民共和国刑法

社会主义核心价值观

详情百度搜索以上“关键词”+site:gov.cn

# 目录



参考资料



主流搜索



Robots协议



搜索语法



语法实例

## —-参考资料



### 开源工具与资源手册

<https://www.i-intelligence.eu/osint-tools-and-resources-handbook-2018/>



### Google Hacking 渗透性测试者的利剑

<https://book.douban.com/subject/30286814/>



### Google Hacking Database

<https://www.exploit-db.com/google-hacking-database>



### NSA Untangling the Web

<https://www.nsa.gov/news-features/declassified-documents/>

## 二-主流搜索



### 中国

百度, 神马, 360, 搜狗, 头条



### 美国

谷歌, 必应, 雅虎, DuckDuckGo



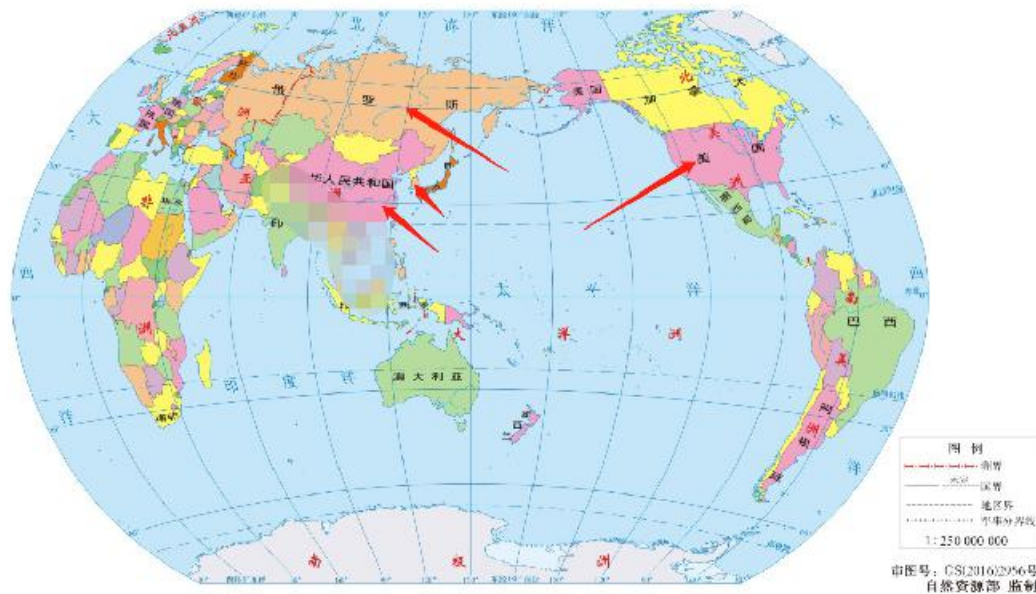
### 俄罗斯

Yandex, Mail.ru



### 韩国

Naver, Daum



### 三-Robots协议

- \$ 结束符（通配符）
- \* 0及以上个字符（通配符）
- ? ?本身，一般后面跟请求参数
- **User-agent: \*** 针对所有爬虫
- **User-agent: Twitterbot** 针对Twitterbot爬虫
- **Disallow: /ab/cd** 爬虫不能爬取URL前缀是/ab/cd的内容
- **Allow: /ab/cd** 爬虫可以爬取URL前缀是/ab/cd的内容
- **Sitemap: /SiteMap.aspx** 站点地图
- **crawl-delay: 1** 每隔1秒爬取一次

### 三-Robots协议

- 1 百度 <https://www.baidu.com/robots.txt>
- 2 知乎 <https://www.zhihu.com/robots.txt>
- 3 淘宝 <https://www.taobao.com/robots.txt>
- 4 携程 <https://www.ctrip.com/robots.txt>
- 5 github <https://github.com/robots.txt>
- 6 NSA <https://www.nsa.gov/robots.txt>

## 四-搜索语法

- 1 北京 天坛 直接搜索，模糊查询
- 2 "北京 天坛" 双引号，精确查询
- 3 北京\*坛 星号，替代0个及以上字符
- 4 北京 -天坛 减号，去除关键词，NOT
- 5 北京 +天坛 加号，添加关键词，AND
- 6 北京 | 天坛 竖杠，或关系，OR

注释：百度限制在38个  
汉字以内，谷歌限制在  
32个单词以内



## 四-搜索语法

- 1 **intitle:北京** 网页标题包含北京
- 2 **intitle:北京 天坛** 网页标题包含北京, 天坛和intitle:北京, 是对等关系
- 3 **allintitle:北京 天坛** 网页标题包含北京与天坛
- 4 **inurl:beijing** 网页链接包含beijing
- 5 **inurl:beijing cn** 网页链接包含beijing, inurl:beijing和cn, 是对等关系
- 6 **allinurl:beijing cn** 网页链接包含beijing和cn
- 7 **intext:北京** 网页正文包含北京, 建议google测试
- 8 **intext:北京 天坛** 网页正文包含北京, intext:北京和天坛, 是对等关系
- 9 **allintext:北京 天坛** 网页正文包含北京和天坛

## 四-搜索语法

- 1 **filetype:pdf** 文档类型是pdf的, xls/doc/ppt均可
- 2 **ext:php** URL后缀是php的
- 3 **cache:www.pjzhang.cn** 网站缓存或网站快照
- 4 **info:beijing** beijing可出现在标题, URL或正文
- 5 **site:pjzhang.cn** 子域名是pjzhang.cn的网站
- 6 **domain:pjzhang.cn** 类似site
- 7 **phonebook:russia** 俄罗斯的开源电话簿



注释: 互联网档案馆  
<https://archive.org/>

## 五-语法实例



**intitle:index.of intitle:index of**



**intitle:后台 inurl:gov.cn**



**intitle: "Index of" intext:log**



**site:cn inurl:jboss filetype:log -github.com**



**inurl:admin/login.php?login= intitle:"CMS"**

注释：谷歌或百度设置中  
有高级搜索框噢

## 五-语法实例



**intitle:"index of /" "\*key.pem"**



**"root.log" ext:log**



**intitle:.\*edu/login**



**site:\*/auth.html intitle:login**



**password mysql site:github.com**

注释：体验盒子GHDB和  
Exploit-db Google  
Hacking

## 六-软件工具

1

WPS演示

2

WPS录屏

3

爱奇艺视频助手

# 共勉

学好英语  
早睡早起  
少看手机  
多多实践



感谢筒子们观看咯!