# 概述

# 使用方式

// 1. 初始化

$ch = curl\_init();

// 2. 设置选项，包括URL

curl\_setopt($ch, CURLOPT\_URL, "http://www.doucube.com");

// 3. 执行并获取HTML文档内容

$output = curl\_exec($ch);

// 4. 释放curl句柄

curl\_close($ch);

## curl\_getinfo()函数

获取最后一次传输的相关信息。PHP Manual有详细信息。

在不同的环境中有不同的效果，在curl\_exec();后会获取当前连接是否成功，成功返回200，失败返回对应的代码例如302,404；

不设置参数时，返回的是一个数组。

参数：  
 CURLINFO\_HTTP\_CODE

最后一个收到的HTTP代码；

## Curl\_setopt参数

CURLOPT\_URL

指定请求的URL；

CURLLOPT\_HEADER

设置为0表示不返回HTTP头部信息

CURLOPT\_RETURNTRANSFER

设置0直接输出，1是存入内存(curl\_exec的返回值是文本流)

# 匹配文本

//开启

$ch = curl\_init();

//配置

curl\_setopt($ch, CURLOPT\_URL, "http://www.baidu.com/");

//设置exec保存html代码

curl\_setopt($ch, CURLOPT\_RETURNTRANSFER, 1);//不自动输出内容

$aa = curl\_exec($ch);

//正则匹配

//$aa中可能会有多行空格或其他字符，使用preg\_match可能匹配不到。

//匹配标签比较困难，直接匹配其中的href比较简单。

preg\_match\_all("/<title.\*>(.\*)<\/title>/",$aa,$bb);

## 相关正则

//匹配a标签中的href

/<[a|A].\*?href=[\'\"]{0,1}([^>\'\"]\*).\*?>/

# 设置重定向

可能可以实现登陆限制的网站。如DMM

# GET请求

如何使用CURL来发送GET请求，发送GET请求的关键是拼装格式正确的URL。

不用特别配置相关参数。

# POST请求

主要设置以下2项

-----------------------------------

//定义要提交的POST数据

$post\_data = array(

“sss”=>”sss”,

);

------------------------------------

// 我们在POST数据哦！

curl\_setopt($ch, CURLOPT\_POST, 1);

// 把post的变量加上

curl\_setopt($ch, CURLOPT\_POSTFIELDS, $post\_data);

//设置post的下标名称

curl\_setopt($ch, CURLOPT\_POSTFIELDS, “data=”.$post\_data);

# 文件上传

文件上传与POST过程一样。主要是POST数据不同。

$post\_data = array (

// 要上传的本地文件地址。添加@符号。

"upload" => "@C:/doucube/www/test.zip"

);

-----------------------------

// 我们在POST数据哦！

curl\_setopt($ch, CURLOPT\_POST, 1);

// 把post的变量加上

curl\_setopt($ch, CURLOPT\_POSTFIELDS, $post\_data);

# 多次请求

1、创建多个cURL资源

$ch1 = curl\_init();

$ch2 = curl\_init();

1. 分别设置对应配置

Curl\_setopt($ch1, CURLOPT\_URL, “http://www.aa.com”);

Curl\_setopt($ch2, CURLOPT\_URL, “http://www.bb.com”);

# PHP爬虫