国内的cms程序基本上都内置了采集功能，国外的几乎没有。看到不少使用drupal的朋友都问到采集，在drupal中采集文章其实也很简单。
怎么采集就不说了，小蜜蜂、火车头，专业的采集程序很多，说说在drupal中怎么采集入库。入库一般两种方式：直接导入数据库，web模拟提交。数据库直导，比较危险，副作用大，不过速度快。模拟提交，稳定安全，速度慢。
采取web模拟提交需要使用到drupal的node\_save函数，这个函数是用来保存节点的，它会自动加载所有模块定义的nodeapi来完成一个节点的保存工作。下面创建一个模块，来实现数据的导入：
第一步，在sites/all/modules下面创建一个文件夹：test。
第二步，定义一个模块，创建一个文件，文件名：test.info，文件内容：

Quote:

name = "test"
version = "6.x-1.x"
core = "6.x"

第三步，创建模块主文件，test.module，内容：

<?php
function test\_menu() {
    $items = array();

    $items['test/set'] = array(
        'title' => 'Photos upload',
        'page callback' => 'test\_set',
        'access arguments' => array('administer nodes'),
        'type' => MENU\_NORMAL\_ITEM,
    );
    return $items;
}

function test\_set() {
   $node = array();
    $node['type'] = 'story';//导入的类型
    $node['title'] = 'test';//节点标题
    $node['name'] = 'admin';//节点作者
    $node['uid'] = '1';
    $node['taxonomy'] = array('',1,1);//这里还不知道是不是这么定义。把节点入库到taxonomy/term1
    $node['promote'] = '1';
    $node['body'] = '22222222222222222222222222222222';//节点正文
    //节点的其它更多内容，分类、评论选项、发表选项、发表时间等，都可以自由定义。
    //print\_r($node);
    $node = (object)$node;
    print\_r($node);
    //$node = node\_submit($node);//提交数组的数据给node\_save。
    node\_save($node);
    if(node\_save){
    echo 'ok';
    }
    else{
    echo 'false';
    }

}
?>

然后启用test模块。启用之后请访问 test/set，看看是不是增加了一个节点。方法就是这样，很简单。可以用火车头采集数据，提交到test/set这个地址来做入库处理。

国内的门户网站大多有如下特点:

　　首页往往会放上一个图片Slide Show，1、2、3、4间隔的变换。

首页只列出文章的标题，按日期排序，不像Drupal默认的那样显示标题和摘要。

首页会划分为各个内容区块，热点新闻，行业新闻，专题等等，各管一摊。

　　Drupal默认安装后，和以上描述的门户网站的风格差的有十万八千里，很多人可能就是因此放弃对Drupal的继续研究，认为洋人的东西不符合国人需求，相反， DeDeCMS之类的国产CMS默认安装就大致符合。

　　实际情况是，Drupal做这样的门户站风格绰绰有余，如下逐条说明：

　　Drupal有大量Image Slide Show的模块，随便找一个，图片Slide show轻松搞定

　　要实现文章列表显示标题的模式，按如下步骤

　　下载views模块，启用

　　进入后台，进入views， 新建一个views

　　修改其中的style， 然后可以到fields区块中点" "， 选中Node title。 更新后，再指定：link to the node，也就是说该标题指向文章的链接,，你也可以根据需要，加上时间显示，作者显示，然后到Sort区块中指定一下排序方式，一般按发布日期的倒序来 排列。

　　最后要实现内容区块，可以下载panels模块，panels可以根据你的需要把首页分割为各种区块模式，三行三列，两行两列，随心所欲。