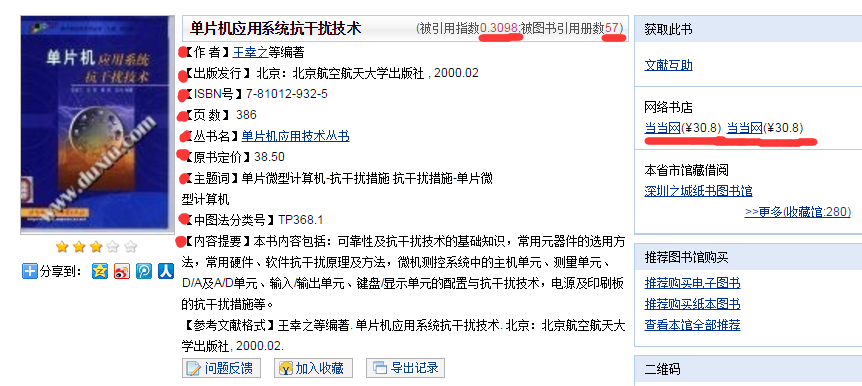
需要采集的网站：http://www.duxiu.com/



可以通过“进入体验版”取得网站的入口cookie post网址时需要跟随

我提供关键词列表，或者入口网址列表 **如：**手册、指南、单片机



上面这张图红色的是我需要的字段，

还有网址中红色这三个字段我也需要

http://book.duxiu.com/bookDetail.jsp?dxNumber=**000000031488**&d=**8B69666DD650D32C47BD08AE446B2AD3**&fenlei=**181704080101**&sw=51%E5%8D%95%E7%89%87%E6%9C%BA

网址的源代码中这个红色的也需要

<input type="hidden" name="f[0].ssid" id="ssidfav" value="**10171408**"/>

<img style="border:1px solid #bababa " src="http://unicover.duxiu.com/coverNew/CoverNew.dll?iid=6564636B6467636B6D6A5A98A9AC9DA95A653133323935373134" border=0 />

**入库**：mysql ,分表，根据fenlei=**181704080101** 这个分类的前两位数进行分表入库。

**代理**，我本来想用阿布代理https://www.abuyun.com/，当然你们有更好的也行，

我采这些内容是用来卖书的，如果你们能同时做成一个图书网站更好，或者介绍一个朋友做，我这需求也简单，网站主要就是搜索，列表显示，购物车，订单提交，微信、QQ会员或者手机等第三方登录，后台有一个商品修改，订单管理 大约就是这样，微信和支付宝扫码支付。

还有一个网站也是一样的内容，如果这个简单些也可以彩这个 http://search.szdnet.org.cn/