

有道购物搜索合作商家商品数据提交办法

“合作商家提交商品数据”的服务是商家加入有道购物搜索的一种方式，同时也是一种比较便捷的商品信息更新方式。依照下面说明的方法，您可以轻松制作商品数据文件，获得商品数据接口的服务。

目录

有道购物搜索合作商家商品数据提交办法.....	- 1 -
提供商品数据http接口的方法.....	- 2 -
使用ftp上传商品数据文件的方法.....	- 3 -
附 1：商品数据属性项说明.....	- 4 -
附 2：商品数据文件的XML格式样例.....	- 6 -
附 3：商品数据文件的CSV格式说明.....	- 7 -
附 4：生成csv文档程序示例.....	- 8 -
附 5：FAQ.....	- 9 -
附 6：常见CSV格式检查结果.....	- 10 -

提供商品数据 http 接口的方法

http 接口适合数据量较小，更新频率要求较高的商品数据更新。商家可以开发 http 接口，也可以提供文件的 url 地址，有道购物搜索会在指定时间进行抓取。

Step1: 准备 xml 或者 csv 格式商品数据：按照附 1 的说明准备商品数据，产生的商品数据格式可以采用附 2 的 xml 格式或者附录 3 的 CSV 格式。

Step2: 在线检查CSV格式（如果采用csv格式）：将产生的csv格式的商品数据存为文件，访问网站<http://gouwu.youdao.com/ping/csv>，[检查csv格式是否正确。](#)

Step3: 开发 http 接口：按照下面的 **Http 接口说明**开发 http 接口

Step4: 通知我们：我们会每天凌晨一点抓去您提供的http接口。



Http 接口说明:

1. 接口方式：商家提供商品数据 http 接口，或者商家提交数据文件的 url 地址，有道购物搜索每天凌晨一点抓取。
2. 确定密钥：商家与有道协商，例如“有道”
3. 全量模式：只需要使用接口参数 token=md5(密钥)
 - a) token=md5(有道)= c32054641d3b95ee100e6dc7468e6e58
 - b) url: http://www.**.com/ping?token=c32054641d3b95ee100e6dc7468e6e58
4. 增量模式：需要开始时间 st，结束时间 et，token=md5(st+et+密钥)
 - a) st=2009-06-01, et=2009-06-16,
 - b) token=md5(“2009-06-12009-06-16 有道”)=ffbc290237540726285df8d97dfed54a
 - c) url:
http://www.*****.com/ping?st=2009-06-01&et=2009-06-16&token=ffbc290237540726285df8d97dfed54a

使用 ftp 上传商品数据文件的方法

与 http 接口的方法相比, ftp 的方法适合大量数据的传输, 而 http 接口更适合小数据量的更新。请您根据实际情况选择更新方式。

目前 ftp 只支持 csv 格式的数据文件上传

上传流程:

Step1: 准备 CSV 文件: 按照附 1: 商品数据文件属性项说明和附 2: 商品数据文件格式说明制作 CSV 格式的商品数据文件。

Step2: 检查 CSV 文件: 访问<http://gouwu.youdao.com/ping/csv>, 检查准备的 CSV 数据文件是否符合标准。

Step3: 上传 CSV 文件: 请申请有道购物搜索商品数据服务, 我们的工作人员会为您开通专用的 FTP 帐号和密码, FTP 地址 ftp://61.135.216.125:3210/, 请于每天凌晨 3 点前提交。

Step4: 通知我们: 我们会为您开通商品数据上线流程, 并且每天自动检查最新上传的商品数据文件。

流程图:



特殊说明:

1. 建议第一次上传全部的商品数据文件, 以后每次上传有更新的商品数据即可。
2. 有道购物搜索的商品数据文件格式同时兼容其他商品搜索服务提供商的 CSV 格式商品数据文件。

附 1：商品数据属性项说明

必需项是必须填写的，如有遗漏会报格式错误；

推荐项可以为用户提供更多的信息，同时也有助于有道购物搜索为贵公司提供服务；

注：请认真填写“必需项”，将会影响商品在搜索中的结果！

必需项：

商品名称

链接地址

图片地址

销售价格

商品描述

图书 ISBN

库存状态

商品品牌

推荐项：

ID

商品型号

规格

商品分类

条形码

市场价格

有道第一级分类

发货城市

属性项“链接地址”的特殊说明

1. 商品链接地址是商品的唯一标识，请不要在同一份文件中多次提交同一件商品。
2. 该商品在贵公司网站上的最简洁 url 地址。如 <http://www.shop.com/product.php?id=1> 和 <http://www.shop.com/product.php?id=1&uid=2> 和指向同一页面，您只需要提交 <http://www.shop.com/product.php?id=1>。
3. 同一页面里有多种商品的情况：
 - a) 因“规格”不同：请填写“规格”属性项
 - b) 因“发货城市”不同：请填写“发货城市”属性项
 - c) 如果有“链接地址”重复但是属性项相同的情况，将视为格式错误。

表格一：feed 文件的属性项说明

	名称	说明	示例
必需项	商品名称	该商品在贵公司网站上的名称	诺基亚 N95 8G
	链接地址	该商品在贵公司网站上的最简洁 url 地址。 如 http://www.shop.com/product.php?id=1 和 http://www.shop.com/product.php?id=1 &uid=2 和指向同一页面，您只需要提交 http://www.shop.com/product.php?id=1。	http://www.shop.com/product_1.html http://www.shop.com/product.jsp?pid=1
	图片地址	该商品图片的 url 地址，请用绝对路径	http://www.shop.com/images/pid_1.jpg
	销售价格	该商品在贵公司网站上的人民币价格	20.00
	商品描述	该商品的 HTML 格式，或者纯文本（txt）格式描述	2007 年度最佳畅销书 诺基亚机皇N95
	图书音像产品的 ISBN 或 ISRC 号码	对于图书音像类，此处为该商品的 ISBN 或 ISRC 号码；对于其他类别商品，置空。	978753424444 7504721913
	库存状态	“有货”、“无货”、“下架”。当贵公司再也不销售此产品，并且对应网页已经不可访问时，请标明下架。	有货 无货 下架
	商品品牌	商品品牌	诺基亚，雅斯兰黛
推荐项	ID	该商品在贵公司内部序列号，用来作为商品区分标识。	如：1，2，3 如：432，5436
	商品型号	该商品的型号	N95
	规格	商品规格信息	如化妆品：50ml 如硬盘：80GB
	商品分类	该商品在贵公司网站上的分类，分级使用两个英文大于号“>>”进行分隔	新潮电子>>GPS 模块
	条形码	即产品外包装上由厂商印制的 UPC 编码	6937010300346
	市场价格	该产品在市场上的人民币价格	100.00
	有道第一级分类	该商品在有道分类中所属的第一级分类，右侧列出了目前有道支持的 9 个大类商品。如果不属于这 9 种分类，请标明“其他”。	美容化妆 电脑数码 图书音像 母婴用品 家电 家居用品 户外/运动 食品/保健品 充值卡 其他
	发货城市	该商品在贵公司的发货城市	北京

附 2：商品数据文件的 XML 格式样例

```
<results>
  <!-- 参数 -->
  <params>
    <count>2</count>
  </params>
  <!-- 商品数据 -->
  <dataset>
    <product id="0">
      <name>诺基亚 e71</name>
      <url>http://www.***.com/item/300618/</url>
      <imgurl>http://img.***.cn/79/58/206223.jpg</imgurl>
      <price>2100.00</price>
      <marketprice>3740.00</marketprice>
      <desc>诺基亚最新手机 e71</desc>
      <available>true</available>
    </product>
    < product id="1">
      <name>诺基亚 5800</name>
      <url>http://www.***.com/item/300619/</url>
      <imgurl>http://img.***.cn/79/58/206224.jpg</imgurl>
      <price>2000.00</price>
      <marketprice>3000.00</marketprice>
      <desc>诺基亚最新手机 e71</desc>
      <available>true</available>
    </product>
  </dataset>
</results>
```

附 3：商品数据文件的 CSV 格式说明

有道目前支持CSV格式的商品数据文件。CSV格式是普遍支持的一种数据表单格式，其的详细定义请参考：http://en.wikipedia.org/wiki/Comma-separated_values。

简单要求如下：

1. 商品之间使用换行符“\n”间隔。

2. 同一件商品的各个属性值之间使用逗号“,”间隔。

■ 例如：

网易	有道购物搜索
----	--------

■ Csv 中的一行：

网易,有道购物搜索

3. 如果商品的属性值中有英文逗号“,”或者换行符“\n”，则将该属性值用英文双引号“”包括起来。

■ 例如：

网易	有道,购物搜索
----	---------

■ Csv 中的一行：

网易,"有道,购物搜索"

4. 如果商品的属性值中有英文双引号“””，则将该属性值中每个英文双引号替换成两个英文双引号，并将该属性值整体用英文双引号“””包括起来。

■ 例如：

网易	有道,""购物""搜索
----	-------------

■ Csv 中的一行：

网易,"有道,""购物""搜索"

注意：

1. 文件名规范：一级域名+时间戳+“.csv”，例如 youdao090107.csv
2. 生成 csv 程序，如附 3：生成 csv 文档程序示例（JAVA）
3. Csv 格式检查页面：<http://gouwu.youdao.com/ping/csv>

附 4：生成 csv 文档程序示例

1. 伪代码：

Print 第一行属性项名称

For each 商品 do

For each 属性项 do

If 不是第一个属性项

Print(“;”)

End if

value=属性项的值中每个双引号替换成连续两个双引号

Print(“”+ value +””)

End for

End For

2. JAVA 程序：

```
import java.io.IOException;
```

```
import java.io.Writer;
```

```
public class CsvWriter {
```

```
    public static final String SEPERATOR = ",";
```

```
    public static final String BRACER = "\"";
```

```
    public static final String NEWLINE = "\n";
```

```
    public static final String REGEX_SPECIAL = ".*[,\\n\"].*";
```

```
    public static boolean write(Writer out, String[] valuesInLine) throws IOException {
```

```
        int count = 0;
```

```
        for (String value: valuesInLine) {
```

```
            if (count != 0)
```

```
                out.append(SEPERATOR);
```

```
            count++;
```

```
            if (value.contains(SEPERATOR) || value.contains(BRACER)
```

```
                || value.contains(NEWLINE)) {
```

```
                if (value.contains(BRACER)) {
```

```
                    value = value.replace(BRACER, BRACER + BRACER);
```

```
                }
```

```
                out.append(BRACER + value + BRACER);
```

```
            } else
```

```
                out.append(value);
```

```
        }
```

```
        out.append(NEWLINE);
```

```
        return true;
```

```
    }
```

```
}
```

附 5: FAQ

1. 问：一定要提供 csv 格式的数据文件吗？
答：是的。目前有道购物搜索只支持 csv 格式的数据文件。
2. 问：必须严格按照 csv 文件格式制作数据文件吗？
答：是的。这样可以在每个列值中保留包括换行符 “\n”、英文双引号 “””，英文逗号 “,” 等各种符号和字符。
3. 问：csv 文件格式需要特殊处理的符号包括哪些？
答：换行符 “\n”、英文双引号 “”” 和英文逗号 “,”。详细处理方法参见附录 2。
4. 问：是否将 csv 数据文件上传到 ftp，就会被收录？
答：不是。只要 csv 文件格式通过，并且技术人员开通对应商家的商品数据服务，上传的 csv 数据文件才会被收录。
5. 问：每次上传，是上传全部数据，还是只上传更新的数据？
答：两者皆可。建议只上传更新的数据。
6. 问：提交的商品数据什么时候可以上线？更新频率如何？
答：每天凌晨上线。
7. 问：FTP 地址？
答：ftp:// 61.135.216.125:3210/
8. 问：CSV 格式检查网址？
答：<http://gouwu.youdao.com/ping/csv>

附 6：常见 CSV 格式检查结果

网址：<http://gouwu.youdao.com/ping/csv>

1. 成功：

csv 文件解析成功！共解析得到 n 种商品

2. 错误：

错误类型	错误信息	错误原因和解决方法
上传	上传文件编码方式不支持	没有自动检测到编码, 建议使用 gbk 或者 utf8 编码
	上传文件读写错误	文件 IO 错误
	读取上传文件错误	文件格式错误
文件解析	同一个 feed 文件中 domain 必须唯一	
	url 重复, 并且相同城市、相同规格	
	csv 文件第一行属性项名称错误: 为空	
	该行属性项太少, 请参考第一行属性项名称, 行号: 1	
	必需项为空 行号: 1 名称: 商品名称 详细信息:	
	库存状态错误 行号: 1 值: 没有 详细请参见相关说明文档	请使用“有货”, “无货”, “下架”等值
	库存状态为空 行号: 1 名称: 诺基亚 6270 详细信息:	
故障	初始化失败	请与有道购物相关技术人员联系
	上传失败	