淘宝爬虫使用说明

作者：崔逸卿

1. 压缩包内有allTypes文件夹以及downloadInfo\_taobao.py脚本。allTypes文件夹中已经按照家具类型、风格形成了文件层次关系。downloadInfo\_taobao.py则是根据allTypes文件夹中的内容爬取信息的工具。
2. 因为淘宝也对用户的访问频率作了限制，所以建议将allTypes中的各种家具类型分为好几块分给不同的人去爬。
3. 脚本使用方法：只要将downloadInfo\_taobao.py拷贝到和家具类型名同一路径下即可。例如，将downloadInfo\_taobao.py拷贝至allTypes目录内。直接python downloadInfo\_taobao.py运行脚本即可开始爬取。

有任何问题可以随时联系我。cuiyiqing@sensetime.com