

Nlpir Parser智能挖掘平台敏感信息过滤

产品名称	Nlpir Parser智能挖掘平台敏感信息过滤
公司名称	灵玖中科软件（北京）有限公司
价格	.00/个
规格参数	
公司地址	北京市海淀区苏州街49-3号5层11号
联系电话	010-62648216

产品详情

敏感词、文字过滤是一个互联网网站必不可少的功能，如何设计一个好的、高效的过滤算法是非常有必要的。随着现在互联网应用的发展，信息的传播速度快，营造好的环境尤其的需要，但一些敏感信息的出现严重的影响了国家和个人的信息安全。

敏感信息指不为公众所知悉,具有实际和潜在利用价值,丢失、不当使用或未经授权访问对社会、企业或个人造成危害的信息。包括：个人隐私信息、业务经营信息、财务信息、人事信息等。

个人意思信息：个人基本资料、身份鉴别信息、个人资产信息、交易信息、宗教信仰、政治倾向、种族、血缘、基因、性生活等资料皆属于敏感个人信息；

业务运营信息：销售业绩、合同信息、管理决策信息、收发文、会议纪要、工作报告等；

财务信息:日常入账信息、税务信息、财务报表等；

人事信息：薪酬信息、员工基本资料信息、绩效考核信息等；

除国家秘密信息和可以公开信息意外的信息。

Nlpir Parser智能挖掘平台敏感信息过滤系统是灵玖软件经过不断努力研发和总结出的，融合了自然语言理解、网络搜索和文本挖掘的技术，可以导入大批量用户业务敏感的关键词列表，实现对内存与文件的实时智能扫描，生成命中的敏感关键词、敏感类别与权重等信息。

Nlpir

Parser智能挖掘平台敏感信息过滤系统在对敏感信息过滤时，首先需要对过滤的内容进行如下预处理：

消重：对需过滤内容中重复的敏感知识进行消重。

格式化：自动处理过滤内容无效字符，如“（空格）”、“.”“*”等信息。

拼音替换：将预处理内容中出现的拼音进行替换。

交叉歧义识别：对过滤内容进行中文分词，涉及交叉歧义的能够准确识别。

其次根据敏感词汇的类型对敏感内容进行过滤计算，将敏感信息的过滤结果划分为无敏感信息、高危敏感信息和疑似敏感信息三种，其中：

高危敏感：涉国家安全、违法信息和重大社会问题等可确定的敏感信息。

疑似敏感：可能引发敏感事件的信息，如谈论国家政治/体制问题、工商干部打人等可疑的敏感信息。

Nlpir Parser智能挖掘平台敏感信息过滤系统应用

1、敏感内容信息过滤

可以设定敏感关键词，针对文章内容、信息关键词进行扫描，可以对敏感信息、事件、人物等信息进行剔除或抓取，适用于网站、出版、上网行为管理。通过针对特定关键词设定，可以过滤大量垃圾信息，净化互联网环境，适用于信息网站、各种论坛的信息整理。

2、敏感账号扫描

通过设定一个或一组敏感账号进行扫描，可以跟踪相关账号的传播路径、权重、受众群体及社会反响等互联网信息，适用于产品、人物跟踪。

3、特定情报实时发现

可以快速便捷地匹配大量自定义的业务关键词，智能发现违反国家法律法规以及侵犯用户权益的内容，确保国家、社会与个人的信息内容安全。