

NLPIR大数据语义智能文本挖掘多功能系统

产品名称	NLPIR大数据语义智能文本挖掘多功能系统
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

当今时代，大数据已经渗透到每一个行业和业务职能领域：政务管理、金融风控、教育科研、媒体出版、公共安全等，成为重要的生产因素。对海量数据的挖掘与运用，将预示着新一波生产率的增长与企业行业竞争优胜劣汰的制胜武器。

NLPIR大数据语义智能分析平台针对大数据内容采编挖搜的综合需求，融合了网络精准采集、自然语言理解、文本挖掘和语义搜索的最新研究成果，先后历时十八年，服务了全球四十万家机构用户，是大数据时代语义智能分析的一大利器。

NLPIR大数据语义智能挖掘平台，针对大数据内容处理的需要，融合了网络精准采集、自然语言理解、文本挖掘和网络搜索的技术，提供了客户端工具、云服务、二次开发接口。开发平台由多个中间件组成，各个中间件API可以无缝地融合到客户的各类复杂应用系统之中，可兼容Windows，Linux，Android，Maemo5, FreeBSD等不同操作系统平台，可以供Java，C，C#等各类开发语言使用。

NLPIR大数据语义智能分析平台十三大功能：

- 1、精准采集：对境内外互联网海量信息实时精准采集，有主题采集(按照信息需求的主题采集)与站点采集两种模式(给定网址列表的站内定点采集功能)。
- 2、文档抽取：对doc、excel、pdf与ppt等多种主流文档格式，进行文本信息抽取，信息抽取准确，效率达到大数据处理的要求。
- 3、新词发现：从文本中挖掘出新词、新概念，用户可以用于专业词典的编撰，还可以进一步编辑标注，导入分词词典中，提高分词系统的准确度，并适应新的语言变化。
- 4、批量分词：对原始语料进行分词，自动识别人名地名机构名等未登录词，新词标注以及词性标注。并可在分析过程中，导入用户定义的词典。
- 5、语言统计：针对切分标注结果，系统可以自动地进行一元词频统计、二元词语转移概率统计。针对常用的术语，会自动给出相应的英文解释。

6、文本聚类：能够从大规模数据中自动分析出热点事件，并提供事件话题的关键特征描述。同时适用于长文本和短信、微博等短文本的热点分析。

7、文本分类：根据规则或训练的方法对大量文本进行分类，可用于新闻分类、简历分类、邮件分类、办公文档分类、区域分类等诸多方面。

8、摘要实体：对单篇或多篇文章，自动提炼出内容摘要，抽取人名、地名、机构名、时间及主题关键词；方便用户快速浏览文本内容。

9、智能过滤：对文本内容的语义智能过滤审查，内置国内最全词库，智能识别多种变种：形变、音变、繁简等多种变形，语义精准排歧。

10、情感分析：针对事先指定的分析对象，系统自动分析海量文档的情感倾向：情感极性及情感值测量，并在原文中给出正负面的得分和句子样例。

11、文档去重：快速准确地判断文件集合或数据库中是否存在相同或相似内容的记录，同时找出所有的重复记录。

12、全文检索：支持文本、数字、日期、字符串等各种数据类型，多字段的高效搜索，支持AND/OR/NOT以及NEAR邻近等查询语法。

13、编码转换：自动识别内容的编码，并把编码统一转换为其他编码。

大数据不仅仅是一种技术，更是一种思维!大数据并不高高在上，它就在生活的方方面面。