

# 灵玖软件：NLPIR语义挖掘实现“自己”学习

|      |                        |
|------|------------------------|
| 产品名称 | 灵玖软件：NLPIR语义挖掘实现“自己”学习 |
| 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

近年来，大数据如浪潮般席卷全球，并深度改变人们的生活、工作和思维方式。世界上越来越多的国家开始从战略层面认识大数据，在管理治理领域融入大数据思维和技术。在此背景下，我国也应顺应时代发展趋势，契合推进治理能力现代化的时代要求，充分利用大数据提升各个领域的能力及效率。

大数据，其影响除了经济、管理方面的，它同时也能在政治、文化等方面产生深远的影响，大数据可以帮助人们开启循“数”管理的模式，也是我们当下“大社会”的集中体现，三分技术，七分数据，得数据者得天下。

随着移动互联网物联网等新兴IT技术逐渐步入主流，大数据使得数据价值极大提高，无处不在的数据，对信息安全提出了更高要求。同时，大数据领域出现的许多新兴技术与产品将为安全分析提供新的可能性；信息安全和云计算贯穿于大数据产业链的各个环节，云安全等关键技术将更安全地保护数据。大数据对信息安全的要求和促进将推动信息安全产业的大发展。

灵玖软件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。

NLPIR大数据语义智能分析平台主要功能：

- 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。

在大数据背景下，数据挖掘技术已经在各行各业中得到了广泛的应用，所以为了更好地满足应用的需要，我们必须切实加强对其特点的分析，并结合实际需要，切实注重数据挖掘技术的应用，才能促进其应用成效的提升。