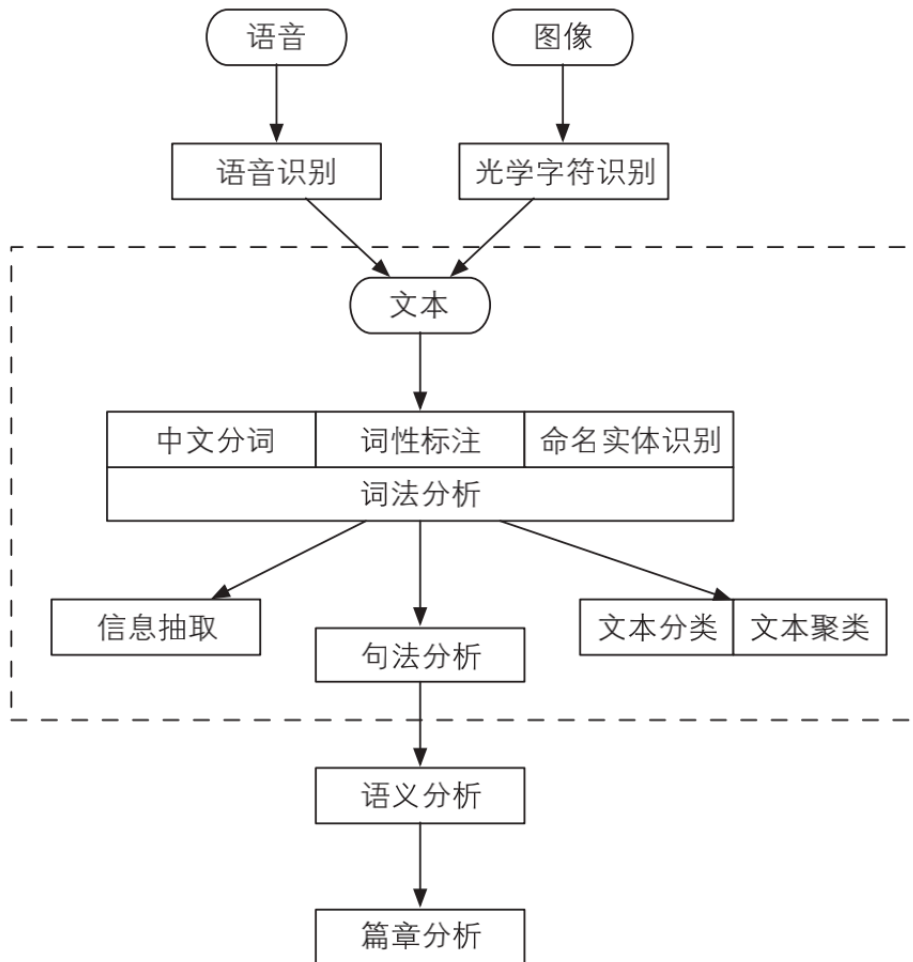


自然语言处理入门

自然语言处理（**Natural Language Processing, NLP**）是一门融合了计算机科学/人工智能以及语言学的交叉学科。这门学科研究的是如何通过机器学习等技术，让计算机学会如何处理人类语言，乃至实现终极目标——理解人类语言或人工智能。

自然语言处理的层次

如下：



自然语言处理的流派



开源工具

指标	NLTK	CoreNLP	LTP	HanLP
功能	词法分析、句法分析、语义分析、文本分类等	词法分析、句法分析、语义分析、关系抽取、指代消解、文本分类等	词法分析、句法分析、语义分析	词法分析、句法分析、关键词句提取、文本分类等
语言	Python	Java 等	C++、Python 等	Java、Python 等
速度	慢	较慢	较快	特别快 ^①
内存	占用多	占用特别多	省内存	省内存
精度	较准	准	准	较准
插件	无	Lucence	无	Lucence 和 Hadoop 等
社区 ^②	8000	6400	2100	143 100
授权	Apache-2.0	Apache-2.0	商用需付费	Apache-2.0