

知识库构建：提供如下 5 个文件

- 本体-实体关系文件（参见 entity-synonym.txt）

每行第一个词语为本体，tab 之后为若干词语，每个词语为对应该本体的实体，实体之间用空格隔开，例如：

汽车      汽车 车 车子 车辆  
车灯      车灯

- 本体-属性关系文件（参见 attribute-synonym.txt）

每行第一个词语为本体，tab 之后为若干词语，每个词语为对应该本体的属性，属性之间用空格隔开，例如：

价格      价格 价位 价钱 费用  
材料      材料 材质 用料 面料  
配置      配置  
噪音      噪音 噪声

- 实体-实体关系文件（参见 whole-part.txt）

该文件中所有词语均是本体表示。

每行第一个词语为某个实体 E，tab 之后为若干空格隔开的实体 E1,E2,E3… 满足 En 是 E 的组件。例如：

座椅      椅面 椅背 座垫 安全带 皮套  
空调      滤芯 压缩机 空调旋钮 空调按钮 空调开关  
底盘

（“底盘”后面为空，说明该实体不由更小的实体组成）

- 实体-属性关系文件（参见 entity-attribute.txt）

该文件中所有词语均是本体表示。

每行第一个词语为某个实体 E，tab 之后为若干空格隔开的属性 A1,A2,A3… 满足 An 是 E 的属性。例如：

车灯      整体 造型 设计 亮度 外观 效果  
皮套      整体  
空调旋钮      整体 手感 位置 设计

- 属性-描述词-极性关系文件（参见 attribute-description.txt，尚缺实验结果，极性是随机标注的）

该文件中所有属性词均是本体表示。

设属性 A 可以被描述词 P1,P2,…、Z1,Z2,…、N1,N2…描述，其中 Pn 为正向描述，Zn 为中性描述，Nn 为负向描述。则用三行文本表示该属性及其描述词

第一行：A P1 P2 …

第二行：A Z1 Z2 …

第三行：A N1 N2 …

例如：

颜色      鲜艳 好看  
颜色      深 不同  
颜色      单一  
档次      高  
档次  
档次      低 不高

（仍然可能有某类描述词缺失的情况，甚至某个属性没有对应描述词，这种情况保留空位）

评论解析模块：

输入：单条评论语句及知识库相关文件

输出：评论解析文件，每行格式如下

实体 属性 描述词 极性

其中，实体、属性均用其本体词表示，极性用 1、0、-1 分别代表正向、中性、负向。

例如，给出评论语句

“车的性能很不错，价位对我来说能接受。外观看来很大气，又不失年轻活力的动感！内部空间不是很大。”

输出的文件如下（假设下面的搭配知识库中都有的情况下）

汽车 性能 不错 1

汽车 价格 能接受 0

汽车 外观 大气 1

内饰 空间 不是很大 -1

产品画像模块

输入：多条针对同一产品的评论语句及该类产品知识库相关文件

输出：产品画像文件，每行格式如下

实体 属性 p n

其中，实体、属性均用本体表示，p 为输入语料中对该属性正向描述次数，n 为负向描述次数，例如

汽车 整体 1000 400

汽车 油耗 400 500

内饰 空间 300 120

座椅 舒适度 100 40