

### 3 个主要的字典：

title\_dict

tags

word\_dict

## 一、标题字典 title\_dict

键：1、2、3、4、5……

值：每个标题的内容

```
>>> title_dict
{1: '\uffeff我校2篇研究生学位论文获2005年上海市研究生优秀成果\n', 2: '我校8位教授被聘为本科专业全国教学指导委员会委员\n', 3: '学校党政领导亲临机关检查指导开学工作\n', 4: '“创新型国家建设与高等教育发展”研讨会在\n', 5: '学校举行“温馨家园·快乐学子”新春联欢会\n', 6: '我校隆重举行百年校庆庆标揭幕仪式\n', 7: '市科教党委副书记翁铁慧与我校医疗器械学院留校学生共迎新春\n', 8: '学校举行2006年迎新春团拜会\n', 9: '我校召开2003-2004年度校先进表彰大会\n', 10: '第二届上海市辅导员论坛分论坛在我校召开\n', 11: '上海理工大学百年校庆标志征集作品群众投票评选\n', 12: '学校召开2005年度部门工作总结交流大会\n', 13: '学校校园网管理中心荣获上海教科网十年建设先进集体荣誉称号\n', 14: '元旦献辞\n', 15: '学校召开学位点建设工作会议\n', 16: '市科委基础研究处领导来我校城建学院调研\n', 17: '国家新闻出版总署副署长柳斌杰教授为我校学生作学术报告\n', 18: '注重培养素质 强化就业指导——学校召开“大学生就业”专题工作会议\n', 19: '我校“污水回用处理装置微型生物生态学研究示范工程”项目通过验收\n', 20: '以百年校庆为契机，共谋区校发展大计——我校党委与杨浦区区委常委举行中心组联组学习\n', 21: '我校参建的“上海生态建设示范楼”荣获“2005年全国十大建设科技成就奖”\n', 22: '上海高校社科部主任工作会议在我校召开\n', 23: '我校十两位教师获市科教系统“模范佳偶”称号\n', 24: '我校召开先进性教育巩固和扩大整改成果并进行“回头看”工作动员大会\n', 25: '我校2005年度精神文明十佳好事评选揭晓\n', 26: '以学习、贯彻、遵守、维护党章为主线抓好学校党风廉政建设——学校召开中心组理论学习会传达中纪委六次全会精神\n', 27: '我校第十次学位点申报取得突破性的成果\n', 28: '校领导慰问春节加班的教职工\n', 29: '\n'}
```

## 二、分词后的字典 tags

键：1、2、3、4、5……

值：每个标题分词后的关键词列表

```
>>> tags
{1: ['研究生', '我校', '2005', '优秀成果', '学位', '论文', '上海市'], 2: ['本科专业', '我校', '聘为', '教学', '委员', '指导', '教授', '委员会', '全国'], 3: ['检查指导', '党政领导', '开学', '亲临', '机关', '学校', '工作'], 4: ['高等教育', '创新型', '研讨会', '建设', '国家', '发展'], 5: ['联欢会', '新春', '学子', '温馨', '家园', '快乐', '学校', '举行'], 6: ['揭幕仪式', '我校', '庆标', '校庆', '隆重', '举行', '百年'], 7: ['我校', '翁铁慧', '迎新春', '科教', '留校', '医疗器械', '党委', '学院', '书记', '学生'], 8: ['团拜会', '迎新春', '2006', '学校', '举行'], 9: ['2003', '2004', '我校', '表彰大会', '年度', '先进', '召开'], 10: ['论坛', '我校', '辅导员', '第二届', '上海市', '召开'], 11: ['上海理工大学', '校庆', '征集', '评选', '百年', '投票', '标志', '群众', '作品'], 12: ['2005', '交流网', '总结', '先进集体', '荣誉称号', '学管', '中心', '荣获', '十年', '学校'], 13: ['校园网', '上海', '建设'], 14: ['元旦', '献辞'], 15: ['学位点', '学校', '召开', '建设', '工作'], 16: ['科委', '我校', '城建', '调研', '学院', '领导', '基础', '研究'], 17: ['学', '报告', '我校', '柳斌杰', '素', '太', '学生', '注重', '强化', '培养', '就业', '学校', '召开', '会议', '工作'], 18: ['就业', '指导', '专题', '工作'], 19: ['项目', '通过', '验收'], 20: ['以', '百年', '校庆', '为契机', '共', '谋', '区', '校', '发展', '大计', '我校', '党委', '与', '杨浦区', '区委', '常委', '举行', '中心组', '联组', '学习'], 21: ['我校', '参建', '的', '上海', '生态', '建设', '示范', '楼', '荣获', '2005', '全国', '十大', '建设', '科技', '成就', '奖'], 22: ['上海', '高校', '社科', '部主任', '工作会议', '在我', '校', '召开'], 23: ['我校', '十', '两位', '教师', '获', '市', '科教', '系统', '模范', '佳偶', '称号'], 24: ['我校', '召开', '先进性', '教育', '巩固', '和', '扩大', '整改', '成果', '并', '进行', '回头看', '工作', '动员', '大会'], 25: ['我校', '2005', '年度', '精神文明', '十佳', '好事', '评选', '揭晓'], 26: ['以', '学习', '贯彻', '遵守', '维护', '党章', '为主线', '抓好', '学校', '党风廉政', '建设', '学校', '召开', '中心组', '理论', '学习', '会', '传达', '中', '纪委', '六', '次', '全', '会', '精神'], 27: ['我校', '第十', '次', '学位', '点', '申报', '取得', '突破', '性', '的', '成果'], 28: ['校', '领导', '慰问', '春节', '加班', '的', '教职', '工'], 29: []}
```

### 三、倒排列表 word\_dict

键：索引关键词

值：对应的标题 ID

```
>>> word_dict
{'科技': [21], '素质': [18], '联欢会': [5], '整改': [24], '论坛': [10], '署长': [17], '学校': [3, 5, 8, 12, 13, 15, 18, 26], '精神文明': [25], '项目': [19], '申报': [27], '示范': [19, 21], '抓好': [26], '开学': [3], '学子': [5], '杨浦区': [20], '两位': [23], '团拜会': [8], '迎新春': [7, 8], '生态': [21], '上海市': [10], '总结': [12], '学习': [20, 26], '交流': [12], '先进性': [24], '第十次': [27], '指导': [2], '研讨会': [4], '论文': [1], '会议': [15, 18, 22], '年度': [9, 12], '校领导': [28], '生物': [19], '征集': [11], '新闻出版': [17], '学术报告': [17], '教师': [23], '举行': [20], '遵守': [26], '为主': [26], '教科网': [13], '党风廉政': [26], '评选': [11, 25], '处理器': [19], '2006': [8], '上海': [22], '微型': [19], '十大建设': [21], '成果': [24, 27], '校庆': [6, 11, 20], '大计': [20], '教职工': [28], '巩固': [24], '研究': [19], '贯彻': [26], '高等教育': [4], '扩大': [24], '党章': [26], '中纪委': [26], '工作': [24], '隆重举行': [6], '荣获': [13, 21], '亲临': [3], '管理中心': [13], '动员大会': [24], '2003': [9], '2004': [9], '十佳': [25], '社科部': [22], '翁铁慧': [7], '六次': [26], '基础': [16], '聘为': [2], '校园网': [13], '慰问': [28], '参建': [21], '获市': [23], '书记': [7], '群众': [11], '就业指导': [18], '中心组': [20, 26], '成就奖': [21], '理论': [26], '注重': [18], '委员': [2], '先进': [9], '学位点': [15, 27], '召开': [12, 15, 18, 22, 24, 26], '本科专业': [2], '研究生': [1], '荣誉称号': [13], '就业': [18], '称号': [23], '柳斌杰': [17], '常委': [20], '建筑': [21], '2005': [1, 12, 21, 25], '创新型': [4], '第二届': [10], '高校': [22], '十年': [13], '生态学': [19], '家园'
```