

## 基本资料

姓名： 刘焕勇

出生日期： 1993.10

手机： 13718796807

联系方式： lhy\_in\_blcu@126.com

## 工作经历

2017.05-至今      中国科学院软件研究所

**NLP 研发工程师**

2017.05-至今

职责业绩： 负责课题项目组自然语言处理相关任务的方案制定，在职期间：

- 1、完成了金融领域事件抽取
- 2、完成了金融领域知识图谱的构建
- 3、完成了金融领域事件驱动模型的建立
- 4、完成了面向金融领域的情感分析模型
- 5、完成了面向金融领域的自动问答系统
- 6、完成了钢铁行业期货情绪与价格实时监测模型
- 7、完成了公司新闻文本分类与新闻推荐

## 教育经历

北京语言大学（ **2014.09 - 2017.07** ）

语言学及应用语言学      硕士

哈尔滨师范大学（ **2010.09 - 2014.07** ）

信息与计算科学      本科

## 工作技能

- 1、 工作语言为 Python，代码见：<https://github.com/liuhuanyong>
- 2、 熟悉汉语词法，句法，擅长基于词法与句法的规则式自然语言处理
- 3、 在信息抽取、情感分析、知识图谱构建上积累一定工作经验，对知识表示有浓厚兴趣
- 4、 能够使用 Scrapy 等工具完成大规模数据采集，使用 Elasticsearch 进行大规模文本检索
- 5、 能够使用 Scikit-learn, Genism, Tensorflow, Keras 等进行机器学习/深度学习的自然语言处理任务
- 6、 能够使用 Mongodb, Mysql, Redis, Neo4j 等数据库

## 项目经历

### 产业链新闻分类与新闻推荐

2018.01-2018.03

项目职务： 自然语言处理工程师

项目描述： 针对构建起的产业链（包括一级行业节点，二级产品节点与公司节点），基于领域特征词构建与过滤的方式，进行产业链节点新闻的挂载与分类；基于文本相似度计算的方法，实现相关新闻的推荐。

项目职责： 1、产业链节点的构建  
2、产业链特征词的构建与扩充  
3、节点挂载与分类的规则制定与实现  
4、基于文本相似度计算的新闻推荐

项目业绩： 产业链文本挂载与分类是进一步得到产业链节点情绪的前提，目前基于产业链节点情绪的效果，验证了模型的有效性。基于文本相似度计算的新闻推荐目前已经成功上线。

### 金融领域自动问答系统

2017.11-2018.01

项目职务： 自然语言处理工程师

项目描述： 基于积累的产业链数据，包括结构化数据（交易数据，情绪数据），新闻文本等非结构化数据，采用基于词法与句法规则相结合的信息抽取方式，提供面向自然语言问句的自动问答服务。

项目职责： 1、问答数据库的设计与整理  
2、问题类型的设计  
3、问题的解析与答案生成模块的研发

项目业绩： 作为项目负责人，负责各项事宜。目前该系统已经成功上线。

### 金融领域情感分析

2017.07-2017.10

项目职务： 自然语言处理工程师

项目描述： 针对金融领域中的钢铁行业，为了找出行业新闻反映出的市场情绪与实际期货（螺纹钢、铁矿石、焦炭）等期货价格情况的关系，研制出面向钢铁行业的情感分析模型。以历史新闻以及历史期货行情数据进行对比回测，验证了模型的可行性。

项目职责： 1、行业资讯的采集，完成行业资讯 300W,公司资讯 1000W 的采集  
2、情感方案的制定。篇章级情感方案的制定，对象级情感方案的制定  
3、领域情感词，领域实体词的收集，扩充。  
4、行业情绪与期货情绪的历时对比与实时监测

项目业绩： 研制的系统在历时的对比中表现很好，目前以及成功上线。

### 对外汉语自适应学习系统研发

2017.05-2017.06

项目职务： 平台负责人

项目描述： 面向对外汉语复句学习者进行开发形成的系统。该系统针对对外汉语学习者的复句学习，采用自适应算法进行练习题以及学习者学习情况的更新，并基于此提供相应的练习题推荐。

项目职责： 项目负责人，负责平台系统的各项事宜。

项目业绩： 完成平台的开发，最终交于北京语言大学大数据与语言教育研究所使用。

<b>语言政策领域知识图谱构建</b>		<b>2017.01-2017.04</b>
项目职务：	平台负责人	
项目描述：	基于现行语言政策领域文本以及知网语言政策领域学术文本，挖掘语言政策领域实体与实体关系，构建起语言政策领域知识图谱；基于语言政策领域知识图谱，开发形成语言政策领域知识图谱检索平台，语言政策领域知识计量分析平台。	
项目职责：	平台独立开发者，完成项目的各项事宜。	
项目业绩：	作为硕士毕业课题，顺利完成知识图谱的构建，并以平台系统的方式落地。该系统最终交于国家语委语言政策研究所使用。	
<b>面向 A 股上市公司的金融知识图谱构建</b>		<b>2016.05-2017.02</b>
项目职务：	自然语言处理工程师	
项目描述：	该项目面向 A 股上市公司，通过收集公司 F9 基本信息、财务信息等结构化数据，通过抓取金融领域文本，采用事件抽取方法，形成上市公司事件数据，将结构化数据与非结构化数据进行融合，采用 Neo4j 数据库进行存储，并基于该图数据库，形成知识图谱查询，钻取，驱动链的构建等应用服务。	
项目职责：	1、结构化数据的采集 2、公司新闻的采集 3、金融领域事件的制定，事件的抽取 4、金融领域知识图谱 Schema 的设计与存储 5、知识图谱应用系统的设计	
项目业绩：	作为该课题的负责人，负责课题的各项事宜。	
<b>语言计算实验室（LC-LAB）研发</b>		<b>2016.08-2016.09</b>
项目职务：	独立开发人	
项目描述：	面向领域语言的词语计算的综合 NLP 处理平台，作为社会语言计算的一个实验平台。基于 NLP 技术进行社会内容计算，包括语言计算技术，词语、热点、人物、文学与职位计算。	
项目职责：	平台独立开发人，负责平台的整体设计、技术选型以及全程开发	
项目业绩：	已成功上线，增强了国家语言资源监测与研究中心的科研实力。	
<b>历时语义云平台(SCP)研发</b>		<b>2015.08-2015.11</b>
项目职务：	平台负责人	
项目描述：	基于历年报纸语料，进行词语的词义演变计算与考察	
项目职责：	平台负责人，负责平台的数据提取与技术架构设计	
项目业绩：	已经成功上线，包括全局互信息搭配词，Word2vec 相关词聚类	
<b>中华传统文化元素计算平台(CCC) 研发</b>		<b>2015.07-2015.10</b>
项目职务：	主要参与者	
项目描述：	基于古典名著中提取传统文化载体元素，面向平面媒体与微博类网络媒体进行流通计算	
项目职责：	主要参与者，负责平台的数据处理工作	
项目业绩：	完成了传统文化元素的流通计算与可视化检索	

