

[知识图谱介绍](#)

[RDF 语法](#)

[结构化数据的知识抽取](#)

[半结构化文本的知识抽取](#)

[非结构化文本知识抽取](#)

[命名实体识别](#)

[关系抽取](#)

[关系抽取工具](#)

[无监督方法](#)

[半监督方法](#)

[监督学习](#)

[事件抽取](#)

[知识挖掘](#)

[实体消歧与链接](#)

[知识规则挖掘](#)

[知识图谱表示学习](#)

[知识存储](#)

知识图谱介绍

- [知识图谱入门笔记](#)
- [从零开始构建知识图谱](#)
- [徐增林_知识图谱技术综述](#)
- [知识图谱构建技术综述](#)

RDF 语法

- JSON-LD [JSON-based Serialization for Linked Data](#)
- Turtle [RDF 1.1 Turtle Terse RDF Triple Language](#)
- Turtle中文翻译(自己翻译的, 仅供参考)[RDF 1.1 Turtle 中文翻译](#)

结构化数据的知识抽取

从结构化数据如MYSQL等数据库获取知识得到三元组。

- 直接映射, W3C的 [A Direct Mapping of Relational Data to RDF](#)
- R2RML [R2RML: RDB to RDF Mapping Language](#)
- D2RQ [D2RQ Accessing Relational Databases as Virtual RDF Graphs](#)

半结构化文本的知识抽取

- 面向半结构化文本的知识抽取研究
- 抽取 Web 信息的包装器归纳学习构造
- 中文百科类知识图谱的构建 [zhishi.me](#) [Zhishi.me - Weaving Chinese Linking Open Data](#)
- Wikipedia Mining [Wikipedia as a Corpus for Knowledge Extraction](#)
- Mining Type Information from Chinese Online Encyclopedias
- Depedia 的构建 [DBpedia A Nucleus for a Web of Open Data](#)
- [DBpedia A Multilingual Cross-Domain Knowledge Base](#)
- [DBpedia - A crystallization point for the Web of Data](#)
- [DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia](#)

非结构化文本知识抽取

命名实体识别

- [CRF++的使用](#)
- [使用CRF++做词性标注等序列化任务](#)
- 使用深度学习做命名实体识别 [Neural Architectures for Named Entity Recognition](#)
- [Bidirectional LSTM-CRF Models for Sequence Tagging](#)

关系抽取

关系抽取工具

- DeepDive 官网介绍 <http://deepdive.stanford.edu/>
- DeepDive 开发者毕业论文 [DeepDive: A Data Management System for Automatic Knowledge Base Construction](#)
- [支持中文的 DeepDive](#) 并附有股权交易示例

- [Deepdive实战 抽取演员-电影间关系](#)
- [开放领域关系抽取 OpenIE 包含 TextRunner等](#)
- [TextRunner 论文](#)

无监督方法

- 基于模板类的实体关系抽取，最简单的是[基于触发词的匹配](#)
- 复杂一点的如基于依存句法匹配的,该方法对输入的单据进行依存分析，通过依存分析输出的依存弧判断单句是否为动词谓语句，如果是则结合中文语法启发式规则抽取关系表述。根据距离确定论元位置，对三元组进行评估，输出符合条件的三元组 [基于依存分析的开放式中文实体关系抽取方法](#)
- 基于核的方法典型的为编辑距离核、字符串核、卷积树核等。基于卷积树核的方法以最短路径包含树作为关系实例的结构化表示形式，以卷积树核作为树相似度的计算方法，采用分层聚类方法进行无监督中文实体关系抽取。[基于卷积树核的无指导中文实体关系抽取研究](#)
- 基于聚类的方法，如对共现的实体及它们的上下文进行聚类，最后标记每一个类簇，以核心词汇作为关系表述。如[无监督关系抽取方法研究](#)

半监督方法

- 标签传播算法[标签传播算法理论及其应用研究综述](#)
- 标签传播算法论文[Relation Extraction Using Label Propagation Based Semi-supervised Learning](#)
- 协同训练[基于弱监督学习的海量网络数据关系抽取](#)
- Boot Strapping 算法[基于Boot Strapping的中文实体关系自动生成](#)
- 远程监督 [Distant supervision for relation extraction without labeled data](#)

监督学习

- pipeline方法。采用Classification by Ranking CNN(CR-CNN)模型，可以学习深度学习在关系抽取中的应用方式，如position embedding 这种。[Classifying Relations by Ranking with Convolutional Neural Networks](#)
- 关系抽取中的特征介绍，包含三大类(contextual and lexical features、nominal rol affiliation、pre-existion relation features)/八小类(lexical features、hypernyms from wordNet、dependency parse、PropBank parse、FrameNet parse、nominalization、predicates from TextRunner、nominal similarity derived from

the Google N-Gram data set)特征: [UTD Classifying Semantic Relations by Combining Lexical and Semantic Resources](#)

- Pipeline 方法, 基于Attention CNN模型 [Relation Classification via Multi-Level Attention CNNs](#)
- Pipeline 方法, 基于Attention-BLSTM模型 [Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification](#)
- Joint 方法, 基于 LSTM-RNN模型 [End-to-End Relation Extraction using LSTM on Sequences and Tree Structures](#)
- 远程监督与深度学习结合, 采用注意力机制取筛选有效实例 [Distant Supervision for Relation Extraction with Sentence-level Attention and Entity Descriptions](#)

事件抽取

- 对时间抽取做了一个综合的介绍, 将其分为元事件抽取和主题事件抽取。其中元事件表示一个动作的发生或状态的变化。主题事件包括一类核心事件或活动以及所有与之直接相关的事件和活动, 可以由多个元事件组成。对于元事件抽取包含基于模式匹配的元事件抽取和基于机器学习的元事件抽取。对于主题事件抽取, 包含基于事件框架的主题事件抽取和基于本体的主题事件抽取两种。它的这个分类和我之前接触到的不大一样, 仅供参考。 [事件抽取技术研究综述 2013 年](#)
- 介绍神经网络在事件抽取方便的综述文章。同时对事件抽取的定义和ACE任务做了简单介绍。其中事件抽取的子任务包含触发词识别、事件类型分类、论元识别、角色分类。而触发词识别和事件类型分类又可以归结为事件识别。论元角色识别和角色分类可归为论元角色分类。根据学习方式的不同, 可以将事件抽取分为基于流水线模型(先事件识别, 后论元角色识别, 论元角色分类的输入是上一步识别出的触发词和所有候选实体)的事件抽取和基于联合模型的事件抽取。而后, 对于ACE任务, 事件识别变为基于词的34类多元分类任务, 角色分类变为基于词的36类多元分类任务。 [神经网络事件抽取技术综述](#)
- [The stages of event extraction](#)
- [Refining Event Extraction through Cross-document Inference](#)
- [Joint Event Extraction via Structured Prediction with Global Features](#)
- [Incremental Joint Extraction of Entity Mentions and Relations](#)
- [Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks](#)
- [Leveraging FrameNet to Improve Automatic Event Detection](#)
- [Improving Information Extraction by Acquiring External Evidence](#)

知识挖掘

实体消歧与链接

- Entity Linking with a Knowledge Base Issues, Techniques and Solutions
- A Generative Entity-Mention Model for Linking Entities with Knowledge Base
- Graph Ranking for Collective Named Entity Disambiguation
- Large-Scale Named Entity Disambiguation Based on Wikipedia Data
- Learning Entity Representation for Entity Disambiguation
- Learning to Link Entities with Knowledge Base
- Leveraging Deep Neural Networks and Knowledge Graphs for Entity Disambiguation
- Using TF-IDF to Determine Word Relevance in Document Queries

知识规则挖掘

- 关联规则挖掘综述
- Random walk inference and learning in a large scale knowledge base
- Variational Knowledge Graph Reasoning

知识图谱表示学习

- Collaborative Knowledge Base Embedding for Recommender Systems
- Improving Learning and Inference in a Large Knowledge-base using Latent Syntactic Cues
- Jointly Embedding Knowledge Graphs and Logical Rules
- Knowledge Graph Embedding by Translating on Hyperplanes
- Knowledge Graph Representation with Jointly Structural and Textual Encoding
- Knowledge Representation Learning with Entities, Attributes and Relations
- Learning Entity and Relation Embeddings for Knowledge Graph Completion
- Translating Embeddings for Modeling Multi-relational Data

知识存储

- [图数据库 中文第二版](#)
- [Neo4j权威指南](#)
- [实战 将数据存进Neo4j](#)