Atitit 数据结构化技术总结 tag 抽取

Office pdf 文档抽取

Html文档抽取

Ner抽取

正则表达式抽取（年份，电话等

从百科类结构化抽取