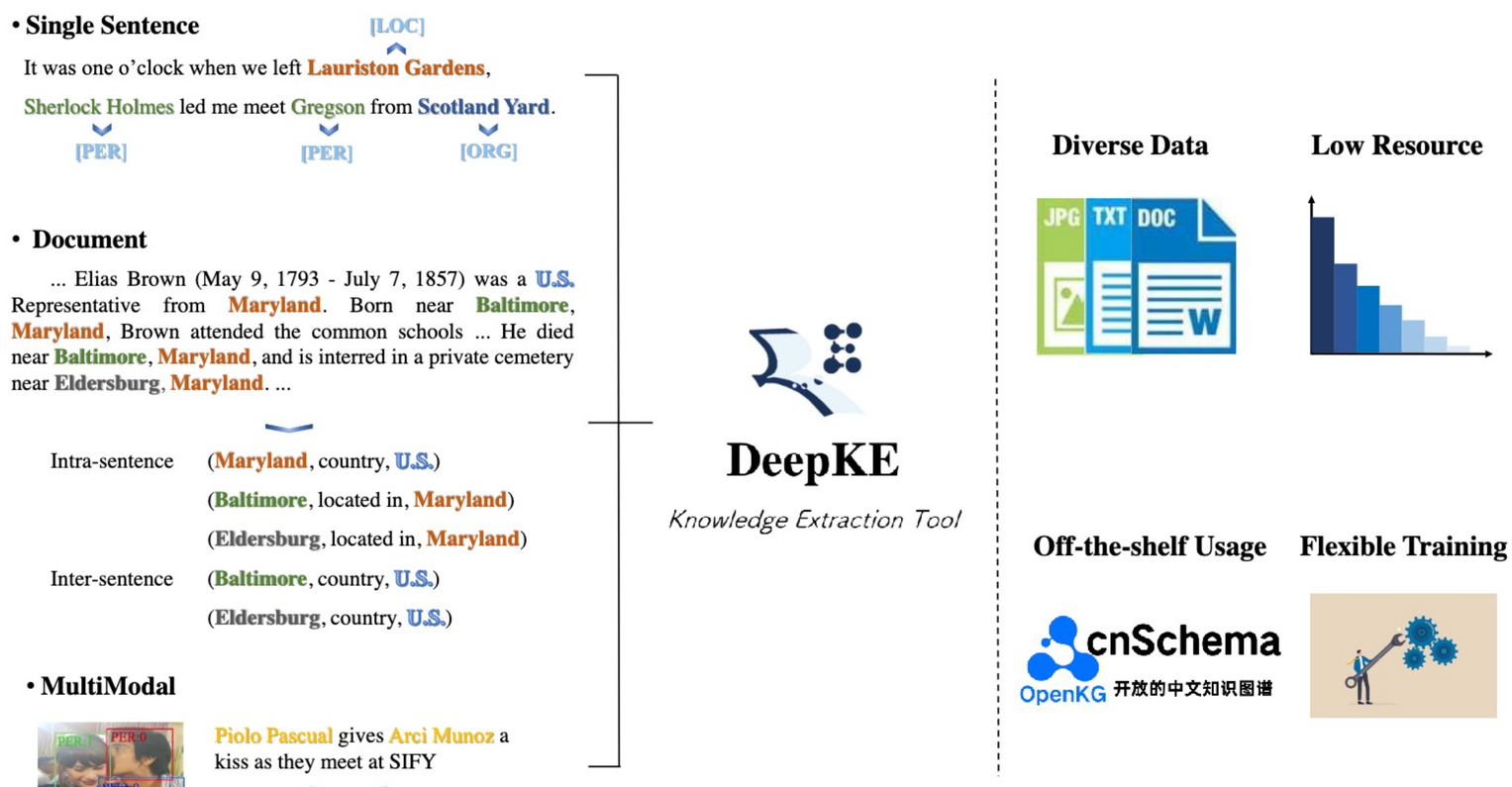
DeepKE: A Deep Learning Based Knowledge Extraction Toolkit for Knowledge Base Population

Ningyu Zhang, Xin Xu, Liankuan Tao, Haiyang Yu, Hongbin Ye, Shuofei Qiao, Xin Xie, Xiang Chen,
Zhoubo Li, Lei Li, Xiaozhuan Liang, Yunzhi Yao, Shumin Deng, Peng Wang, Wen Zhang, Guozhou Zheng
Zhejiang University & AZFT Joint Lab for Knowledge Engine

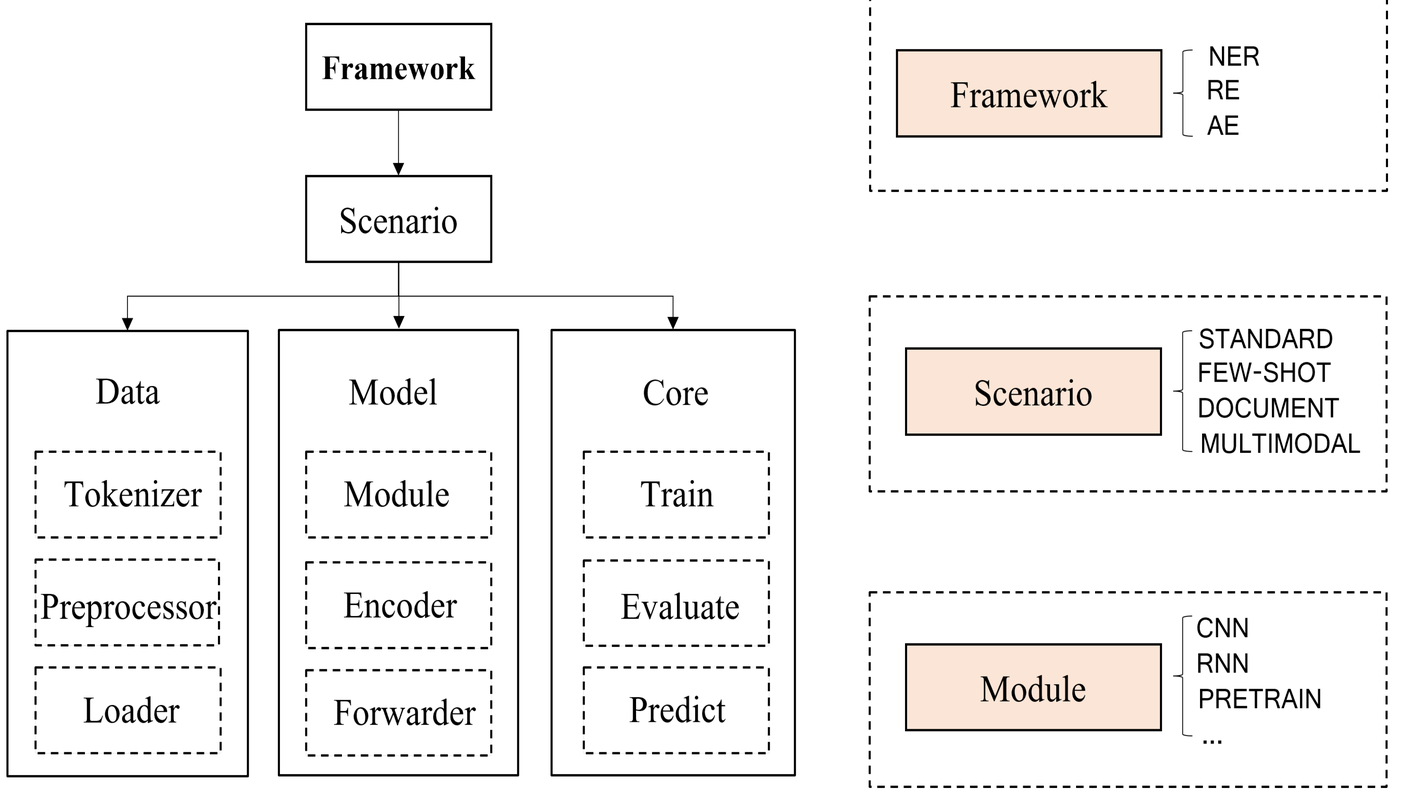
Characters of DeepKE

- An open-source and extensible knowledge extraction toolkit
- Implements various information extraction tasks, including **named entity recognition** (NER), **relation extraction** (RE) and **attribute extraction** (AE)
- Compatible with diverse English and Chinese relation extraction datasets
- Support many classic neural network models and advanced methods



- Contain **cnSchema-based Off-the-shelf Models (Chinese)**: DeepKE-cnSchema (NER / RE)
- Provide the online demo system, detailed tutorials and documentation

Design of DeepKE



- Three modules with one unified framework : Data, Model and Core
- **Data**: preprocessing and loading input data; a bridge between data files and models
- **Model**: including submodules of diverse NNs, encoders and forwarders; implementing NNs in special tasks and scenarios
- **Core**: model training, evaluation, prediction and different loss functions

Usage of DeepKE

- Different tasks and scenarios with a unified API including standard supervised RE, few-shot RE (*KnowPrompt* [1]), standard supervised NER, few-shot NER (*LightNER* [2]), document-level RE (*DocuNet* [3]) , multimodal NER and RE (*IFAformer*)
- **Step1:** Go to the *example* folder in the DeepKE GitHub Repo.
  Step2: Modify **yaml* to customize hyperparameters
  Step3: Run *train.py* for training and *predict.py* for prediction
- Execute real-time extraction in the online demo system and show results as knowledge graphs and tabels

GitHub

Demo

DeepKE
Knowledge Extraction Tool

[1] Chen X, Zhang N, Xie X, et al. Knowprompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction [C]. WWW 2022.
[2] Chen X, Lei L, Shumin D, et al. LightNER: A Lightweight Tuning Paradigm for Low-resource NER via Pluggable Prompting [C]. COLING 2022.
[3] Zhang N, Chen X, Xie X, et al. Document-level relation extraction as semantic segmentation [C]. IJCAI 2021.