DeepStruct: Pretraining of Language Models for Structure Prediction
Chenguang Wang*, Xiao Liu*, Zui Chen*, Haoyun Hong, Jie Tang, Dawn Song
UC Berkeley  Tsinghua University

Traditional Understanding
Born in 1951 in Tbilisi, lago is a Georgian artist.

Structure Understanding
Born in 1951 in Tbilisi, lago is a Georgian artist.

Structure Representation
Joint entity and relation extraction
Born in 1951 in Tbilisi, lago is a Georgian artist.

Named entity recognition
Born in 1951 in Tbilisi, lago is a Georgian artist.

Open information extraction
Born in 1951 in Tbilisi, lago is a Georgian artist.

Autoregressive Training
Output Triples

Input Sentence

Task-agnostic Training Data
• ~ 51M sentences
• ~ 134M entities
• ~ 114M relations (triples)

Multiple Tasks
• 10 tasks
• 28 datasets
• ~ 700K sentences

DeepStruct: Predict *structures* from text

DeepStruct

Predict single words from text

DeepStruct 10B vs GPT-3 175B

Results:
• DeepStruct 10B largely outperforms GPT-3 175B model. (a)
• Achieved state-of-the-art result on 21 of 28 datasets. (b)
• DeepStruct performance increases drastically along with model size. (c)

Scaling Effect

10 tasks and 28 datasets