LOME: Large Ontology Multilingual Extraction
Patrick Xia*, Guanghui Qin*, Siddarth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White, Benjamin Van Durme

**Architecture**
- Raw Text
- Multilingual Tokenization
- FrameNet Parser
- Multilingual coreference resolution
- Fine-grained hierarchical entity typing
- Temporal relation prediction
- Third-party systems (e.g. relation extraction)

**Entities**
- **the rabbit (animal)**
  - Mentions: “The rabbit ate a carrot.”, “兔子喝了水”
- **a carrot (food)**
  - Mentions: “The rabbit ate a carrot.”
- **水 (food)**
  - Mentions: “兔子喝了水”

**Events**
- **Ingestion: 吃了**
- **Ingestor: 兔子**
- **Ingestible: a carrot**

**FrameNet**
- BIO tagger to find spans + MLP typing module for labeling + XLM-R
- First to report on full FrameNet parsing, SOTA on frame identification:
  - Yang and Mitchell (2017)
  - Hermann et al. (2014)
  - Peng et al. (2018)
  - This work

**Coreference**
- Neural coreference resolution model + XLM-R + given spans
- Trained and evaluated multilingually:
  - Arabic
  - Chinese
  - Dutch
  - English
  - Italian
  - Russian
  - Spanish

**Typing**
- Coarse-to-fine type decoder + XLM-R + Borda Voting for cluster-level labels
- Benchmarked on multiple datasets:
  - BBN
  - FIGER
  - UltraFine

**Time**
- Real-valued event pair model + XLM-R
- Benchmarked on both English (monolingual) and Chinese ACE:

**Benchmarks**

**Demo**
- Website: [https://nlp.jhu.edu/demos/](https://nlp.jhu.edu/demos/)
  - 1. Docker Hub: Concrete-based outputs
  - 2. Web Demo

Hover mouse to view event spans (in blue boxes) to see their arguments. Argue by the coref model.

**Website**

**Overview**
- Sample inputs:
  - The rabbit ate a carrot.
  - 兔子喝了水

**Diagram**
- Flowchart of architecture and component descriptions.

**Chinese Temporal Relation (F1)**
- Majority Class
- Zero-shot
- Chinese-trained