

## Text Data Collection



~10k article  
abstracts from  
Wikipedia

## Annotation Phase

LLM-based  
annotation  
(zero-shot)

rule-based  
verification

## Initial Annotations



5114 samples  
3899 entity types  
~60k entities  
8924 relation types  
~42k triples

## Annotation Post-Processing

LLM-based  
triple-verification

## Final Annotations



5010 samples  
3466 entity types  
~59k entities  
7103 relation types  
~30k triples