

High-resource

Hundreds of millions of documents online, large datasets, large Wikipedia, a lot labelled data

Medium-resource

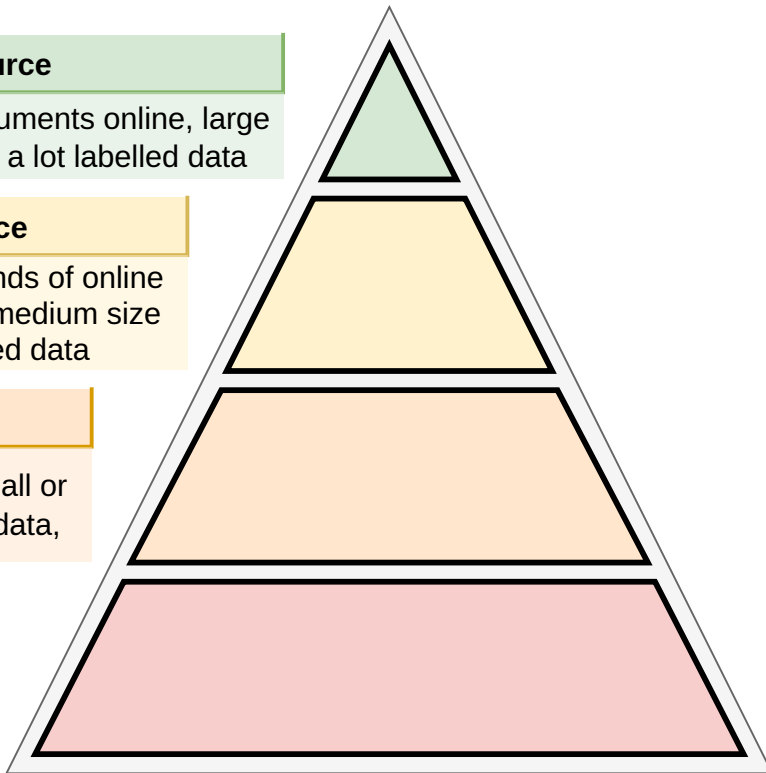
Millions to Hundredthousands of online documents, parallel data, medium size Wikipedia, few labelled data

Low-resource

Few documents online, small or no Wikipedia, no labelled data,

Zero-resource

No documents online, no Wikipedia, no labelled data



NLP research has mostly focused on these languages

Need to apply NLP to these languages