

A Collaborative Grammatical Annotation Tool, BoAT

Furkan Akkurt – furkan.akkurt@bogazici.edu.tr

Department of Computer Engineering, Boğaziçi University, Istanbul, Turkey
1st UniDive Training School, Chişinău, Moldova, 8-12 July 2024

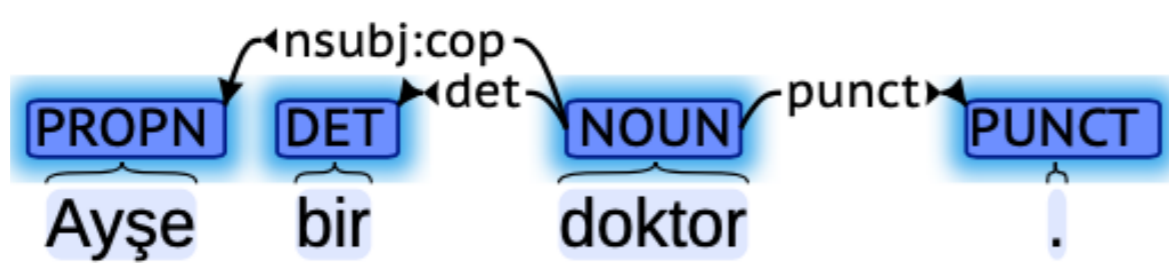
Annotation tool

- Editing all fields, including metadata

ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	MISC
1	Ayşe	Ayşe	PROPN	PROPN	–	3	nsubj:cop	–
2	bir	bir	DET	DET	–	3	det	–
3	doktor	doktor	NOUN	NOUN	–	0	root	–
4	.	.	PUNCT	PUNCT	–	3	punct	–

Annotating the sentence *Ayşe bir doktor.* ‘Ayşe is a doctor.’

- Dependency graph visualization



Dependency graph of the sentence above

- Validation reports

Errors
ID 3: 'SpaceAfter=No' is missing in the MISC field of node #3 because the text is 'doktor.'

Validation report of the sentence above

- Collaboration
- Advanced search

5 results found.
Query was feats: Evident=Nfh in the treebank *ud-turkic*

Sentence ID	Text	Annotate	See annotations
1	Bizimki yazlıktaki çocuk odasını apar topar derleyip toplamış, mutfağın masasını da elden çıkarmıştı.		
8	Sude üç saattir ofiste yokmuş, Ayşe de evde değilmiş.		
13	Hırkam evde miymiş?		
15	Sonuçlar açıklanmış, değil mi?		
16	Ayşe bu kitabı okumamış değil.		

Searching for the feature Evident=Nfh

- Single modality
- Autocompletion

Future work

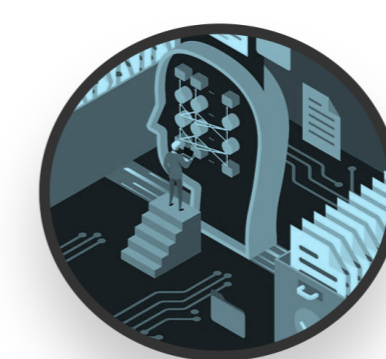
- Inter-annotator agreement
- Progress visualization, reminders
- Rules for automatic annotation
- More advanced search possibly based on dependencies

Questions

- Treebank comparison
- What fields to automate? *initially*
- Searching published treebanks
What if not valid?
- What other features to add?
Is there a green hat idea?
Open-source LLM integration?

Links

- github.com/furkanakkurt1335/boat
- furkanakkurt5827.space



Funded by the
European Union

UniDive

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

