WojoodNER 2023:
The First Arabic Named Entity Recognition Shared Task

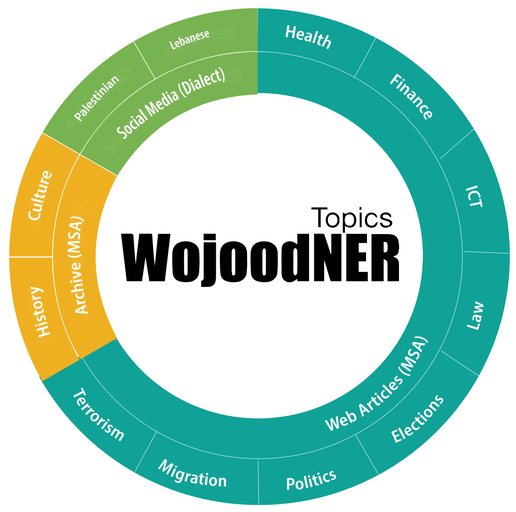
Mustafa Jarrar[1], Muhammad Abdul-Mageed[2,3], Mohammed Khalilia[1],
Bashar Talafha[2], AbdelRahim Elmadany[2], Nagham Hamad[1], Alaa' Omar[1]

[1]Birzeit University, Palestine [2]The University of British Columbia [3]MBZUAI

Task Description

❖ Subtask 1 - FlatNER

❖ Subtask 2 - NestedNER



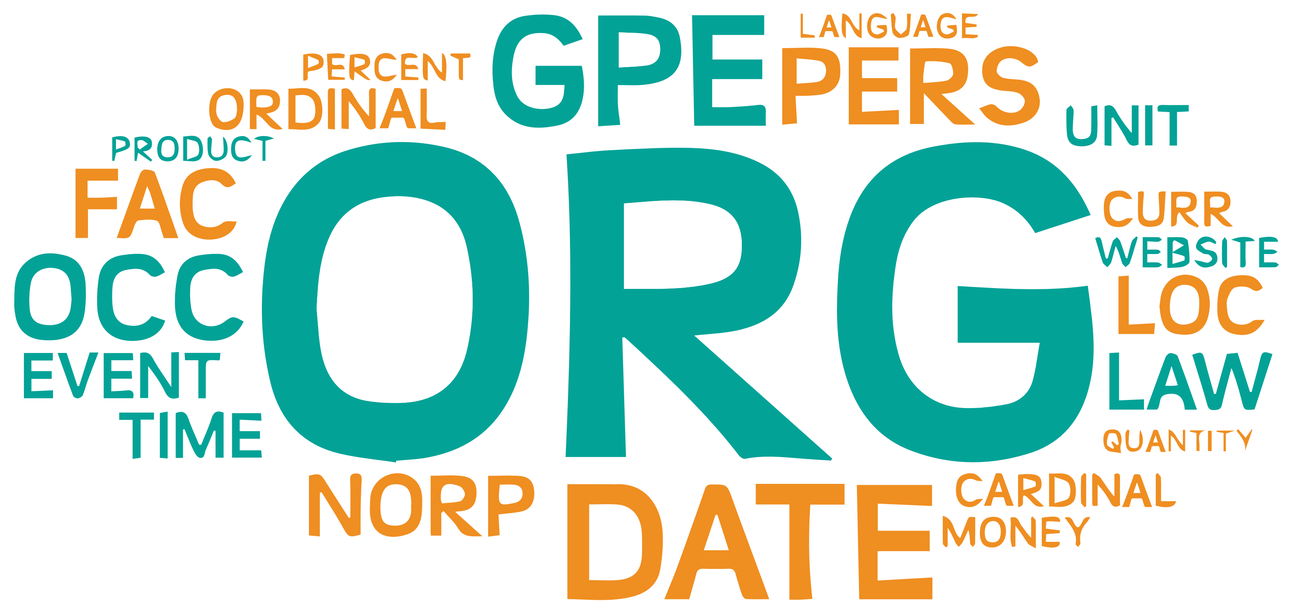
Shared Task Datasets and Evaluation

Datasets

WojoodNER-2023 shared task employs the Wojood corpus as its primary dataset (Jarrar et al., 2022). The Wojood corpus encompasses approximately 550K tokens, spanning both MSA and two Arabic dialects, annotated using 21-entity

Evaluation metrics

The official evaluation metric for subtask1 and subtask2 is the macro-averaged F1 score. In addition to this metric, Precision Recall, and Accuracy is recorded for submissions to both subtasks.

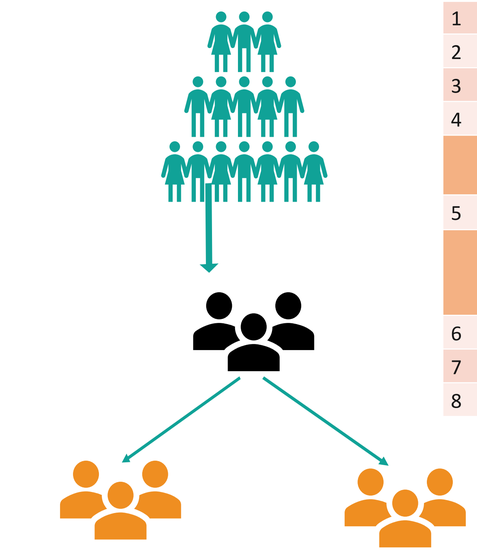


Shared Task Teams & Results

Teams

- 45 unique team registrations.
- 12 teams submitted for FlatNER and Nested-NER.
- 35 submissions for FlatNER from eleven teams and 22 submissions for NestedNER from eight teams

Team	Affiliation	Task
Alex -U 2023 NLP	Alexandria University, Pakistan	1,2
AlexU -AIC		1,2
AlphaBrains	University of Gujrat,	1,2
ARATAL	IPSA	1
El -Kawaref	German University in Cairo	1
ELYADATA	ELYADATA	1,2
Fraunhofer IAIS	Fraunhofer IAIS	1
LIPN	LIPN, Université Paris 13	1,2
Lotus	MBZUAI	1,2
R 00	Jordan University of Science and Technology	1,2
Think NER	Ulm University	1,2
UM6P & UL	Mohammed VI Polytechnic University	1,2



Results

- All the models submitted to the shared task adopt the transfer learning approach, leveraging pre-trained models trained on various data sources.
- The top-performing models addressed the challenge of identifying nested entities of the same type

Rank	Team	F1	Precision	Recall
1	ELYADATA	93.73	93.99	93.48
2	UM6P & UL	93.03	92.46	93.61
3	AlexU -AIC	92.61	92.10	93.13
4	LIPN	92.45	92.31	92.59
	Baseline-I (ARBERTv2)	91.68	91.01	92.35
5	Think NER	91.4	90.03	92.82
	Baseline-II (AraBERTv2)	91.06	90.74	91.38
6	Alex-U 2023 NLP	90.01	89.39	90.63
7	AlphaBrains	88.84	88.45	89.23
8	Lotus	76.02	82.19	70.72

Results of Subtask 2 - NestedNER

Rank	Team	F1	Precision	Recall
1	LIPN	91.96	92.56	91.36
2	El-Kawaref	91.95	91.43	92.48
3	ELYADATA	91.92	91.88	91.96
4	Alex-U 2023 NLP	91.80	91.61	92.00
5	Think NER	91.25	90.76	91.73
6	ARATAL	91.13	90.49	91.77
7	UM6P & UL	91.13	90.70	91.57
8	AlexU -AIC	91.13	91.33	90.92
	Baseline-I (ARBERTv2)	89.20	88.32	90.09
	Baseline-II (AraBERTv2)	87.33	86.00	88.00
9	AlphaBrains	87.15	87.45	87.58
10	Lotus	83.39	80.90	86.04
11	R00	76.99	76.67	77.31
12	Fraunhofer IAIS	64.45	65.53	63.40

Results of Subtask 1 - FlatNER

Downloads and Demo

https://sina.birzeit.edu/wojood/

Public (data, code, demo)