ArabicNLU 2024:
The First Arabic Natural Language Understanding Shared Task

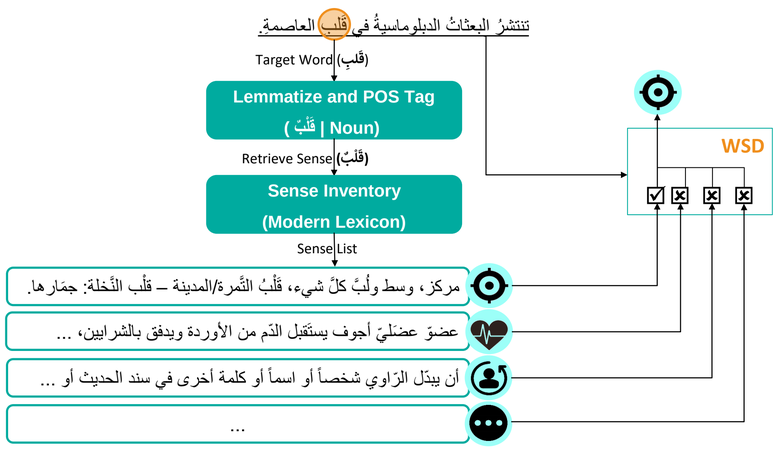
Mohammed Khalilia[1] Sanad Malaysha[1] Reem Suwaileh[2] Mustafa Jarrar[1],
Alaa Aljabari[1] Tamer Elsayed[3] Imed Zitouni[4]

[1]Birzeit University, Palestine [2]Hamad Bin Khalifa University, Qatar [3]Qatar University, Qatar [4]Google, USA

Task Description

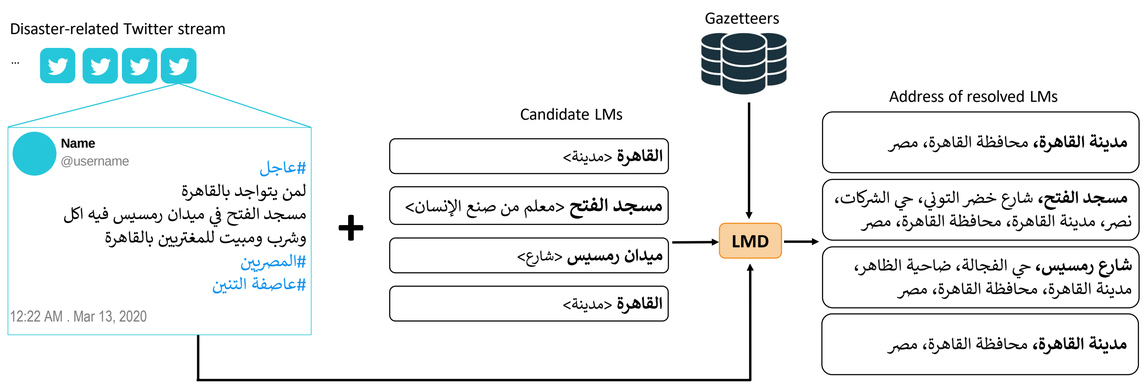
Subtask 1 – WSD

Word Sense Disambiguation (WSD): disambiguate the semantics of polysemous words, stems from their multiple meanings.



Subtask 2 - LMD

Location Mention Disambiguation (LMD): Resolving Location Mentions (LMs) in texts and linking them to toponyms in geo-positioning databases



Shared Task Datasets

Subtask 1 Word Sense Disambiguation

Best: 78%

Baseline: 84%

SALMA Corpus

34k sense-annotated tokens

Subtask 2 Location Mention Disambiguation

Best: 95%

Baseline: 62.7%

IDRISI-DA Corpus

3900 annotations

Natural Language Understanding

WSD Dataset (SALMA)

The first sense-annotated corpus for Arabic. It contains:

- 1,440 sentences.
- 34K tokens (8,760 unique tokens with 3,875 unique lemmas).
- A total of 4,151 senses.

LMD Dataset (IDRISI-DA)

The first Arabic LMD dataset. It contains:

- 2,869 posts from diverse dialects.
- 3,893 location mentions, of which 763 are unique.
- Across 7 countries.

Shared Task Teams and Results

Teams

- **WSD:** 35 unique registered teams, only 3 teams submitted.
- **LMD:** 25 unique registered teams, only 2 teams managed to submit.

Team	Affiliation	Task
Pirates	Nile University	1
Rematchka	Cairo University	1, 2
Upaya	SCB DataX	1, 2

Results

Results of Subtask 1 – WSD

Rank	Team	Accuracy
	Baseline	84.2%
1	Upaya	77.8%
2	Pirates	70.8%
3	Rematchka	57.5%

Results of Subtask 2 – LMD

Rank	Team	MRR@1	MRR@2	MRR@3
1	Rematchka	94.97%	95.00%	95.00%
2	Upaya	59.94%	59.94%	59.94%
	Baseline	57.24%	63.96%	64.28%

Open Challenges

- **WSD and LMD are challenging tasks!!** Need for more datasets and more research.
- **Arabic dialectics** should be supported in WSD and LMD.
- **LLMs performed badly** - compared to classification architectures, for Arabic LMD.
- Maybe **Arabic-tailored LLMs are needed**!!.