

قضايا بحثية 1 - التكليف 2

Prepared by: **Abobakr Mohammed Baras**
ID: **253002656**

OCT 2025





الورقة البحثية:

Arabic WordNet and the Challenges of Arabic

Paper Link: <https://aclanthology.org/2006.bcs-1.2/>

Chapter
01

العنوان: Arabic WordNet and the Challenges of Arabic

- العنوان واضح ويُبرز المشروع (Arabic WordNet) والمشكلة العامة (تحديات اللغة العربية).
- نقاط نقد:
 - لا يذكر الفترة الزمنية أو حالة المشروع (مستوى النسخة أو التقدم).
-

Arabic WordNet is a lexical resource for Modern Standard Arabic based on the widely used Princeton WordNet for English (Fellbaum, 1998). Arabic WordNet (AWN) is based on the design and contents of the universally accepted Princeton WordNet (PWN) and will be mappable straightforwardly onto PWN 2.0 and EuroWordNet (EWN), enabling translation on the lexical level to English and dozens of other languages. We have developed and linked the AWN with the Suggested Upper Merged Ontology (SUMO), where concepts are defined with machine interpretable semantics in first order logic (Niles and Pease, 2001). We have greatly extended the ontology and its set of mappings to provide formal terms and definitions for each synset. The end product would be a linguistic resource with a deep formal semantic foundation that is able to capture the richness of Arabic as described in Elkateb (2005). Tools we have developed as part of this effort include a lexicographer's interface modeled on that used for EuroWordNet, with added facilities for Arabic script, following Black and Elkateb's earlier work (2004). In this paper we describe our methodology for building a lexical resource in Arabic and the challenge of Arabic for lexical resources.

نقاط نقد:

- المستخلص وصفي لكن يفتقر إلى أرقام/نتائج كمية (كم عدد synsets المكتملة؟ مدى التغطية الفعلية؟).
- لا يذكر منهجية التقييم أو معايير النجاح بوضوح.
- يفتقد توضيح مساهمة البحث مقارنة بالأعمال السابقة.

المقدمة تعطي خلفية قوية عن WordNet و EuroWordNet.
نقاط نقد:

- تحتاج إلى ربط أوضح بين الخلفية والفجوة (لماذا AWN ضرورة الآن؟).
- يمكن تقصير بعض الفقرات التاريخية لصالح إبراز أسئلة البحث بوضوح.

- نقص معاجم عربية موحدة ومربوطة شبكيًا يُسهّم في تطبيقات ال NLP.
 - صعوبة التعامل مع الصرف والجذور في النصوص بدون حركات.
 - الاعتماد على WordNet الإنكليزي يخلق تحيزاً؛ المقال يقترح SUMO كبديل.
- نقطة نقدية: الفجوة مبنية جيداً، لكن المقال يحتاج إلى إبراز قياسات تُظهر حجم المشكلة عملياً.

- بناءAWN متوافق مع PWN/EWN.
- ربط الـsynsets بـSUMO لإنشاء أساس دلالي رسمي.
- تطوير واجهة للقاموس والمحررين ودعم Unicode/العربية.
- الوصول لتغطية تقارب 10,000 synsets.

- ما غطاه المؤلفون: SUMO، EuroWordNet، PWN، أعمال التصريف العربية (مثل Buckwalter)، وأدوات تحليل الصياغة.
- نقد:**

- مراجعة قوية لكن يمكن تحديثها بإشارات أحدث (أدوات تعلم عميق، موارد موازية، تقييمات معيارية).
-
- تفتقد مناقشة مقارنة منهجية مع أعمال عربية أخرى إن وُجدت.

النهج المستخدم:

- بناء مرجعي يُبنى يدوياً حول Common Base Concepts ثم التوسيع الآلي/اليدوي.
- استخدام 17 إجراء استدلالي لاشتقاق اقتراحات عربي-إنكليزي، ثم تحقق يدوي.
- ربط بالمفاهيم في SUMO، وتطوير واجهة تحرير ودعم قاعدة بيانات (MySQL).

نقد:

- الاعتماد الكبير على العمل اليدوي يثير تساؤلات عن القابلية للتوسع والسرعة.
- يحتاج فصلاً أو جدولاً يوضح الـ 17 إجراءً بالخطوات والأمثلة.
-

الورقة البحثية تناقش الصعوبات الموجودة لبناء wordnet للغة العربية لذلك لا توجد نتائج واضحة غير استخلاص الباحثين عن أفضل الطرق لانجاز هذا العمل



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY



Thank You

