

نقد ورقة علمية من حيث أهمية البحث

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

Nov 2025



الورقة البحثية المنتقدة

Enrichment of Arabic WordNet Using Machine Translation and Transformers

Link: <https://aclanthology.org/2023.icnlp-1.36.pdf>

Chapter
01

تهدف الورقة إلى إغناء (Arabic WordNet AWN) اعتمادًا على:

- Princeton WordNet (PWN)
- أنظمة الترجمة الآلية (Google, Reverso, MT5)
- نماذج التحويلات (Transformers) مثل BERT و AraBERT

المخرجات الأساسية:

- توليد ~ 90,127 Synsets مرشحًا
- دقة تحقق يدوي لعينة 1000 $\approx 75.4\%$ Synset
- تطوير بوابة ويب لمراجعة الخبراء

تؤكد الورقة على أن:

- WordNets مورد أساسي لمعالجة اللغة الطبيعية (NLP)
- إنشاء WordNet يدويًا مكلف وطويل
- AWN فقير مقارنةً بـ PWN من حيث التغطية

تقدّم العمل على أنه:

- حلّ عملي لتوسيع AWN
- خطوة نحو تحسين دعم اللغة العربية في تطبيقات NLP

تلبية حاجة حقيقية في مجتمع اللغة العربية:

• نقص الموارد المعجمية الدلالية القياسية للغة العربية

التركيز على مورد له استعمال واسع عالميًا:

• WordNet معيار معروف في الكثير من تطبيقات NLP

حجم الإضافة المحتملة كبير:

• Synsets 90,127 مرشحة

• تقدير بإمكانية الحصول على ~68,000 Synsets صحيحة بعد التحقق

وجود آلية للتحقق البشري:

• بوابة ويب تسمح بمراجعة الخبراء، ما يزيد من موثوقية المورد

نقد أهمية البحث من حيث وضوح "المشكلة البحثية"

- المشكلة مُعرّفة بشكل عام فقط:
 - "AWN صغير ومكلف التوسع يدويًا"
- لا يوجد تحديد كمي واضح للفجوة:
 - ما هي التغطية الحالية لـ AWN في مجالات معيّنة (أفعال، أسماء، مجالات دلالية)؟
 - ما هي التأثيرات السلبية الحالية على تطبيقات NLP العربية؟
- لا يتم ربط المشكلة بسيناريوهات تطبيقية محددة:
 - مثل: تأثير نقص AWN على أنظمة السؤال والجواب، الترجمة، أو التحليل الدلالي

شريحة 7: نقد أهمية البحث من حيث الإضافة التطبيقية والتأثير

الأثر الإيجابي المحتمل:

- توسيع AWN يمكن أن يخدم:
 - أنظمة الاسترجاع الدلالي
 - أنظمة السؤال والجواب بالعربية
 - تطبيقات تعليم اللغة والترجمة

لكن:

- لا توجد تجارب تطبيقية تُظهر:
 - تحسن أداء نظام فعلي قبل/بعد استخدام AWN المُعْتَمَد
- التقييم محصور في:
 - نسبة صحة الكلمات/التعريفات/الأمثلة في عينة Synset 1000
- لا يتم ربط النتائج بمقاييس معيارية في مهام NLP (مثل: تحسن F1، أو دقة تصنيف، إلخ)

- الاعتماد الكبير على الترجمة الآلية:

- استخدام Google Translate و Reverso مع أن الأخطاء معروفة خاصة في السياقات المتخصصة
- التركيز على التشابه الدلالي (cosine similarity) لا يكفي وحده لضمان جودة معجمية عالية

- نسبة الدقة:

- دقة إجمالية 75.4% تعني أن حوالي ربع المخرجات غير صحيحة
- هذا يؤثر تساؤلاً عن مدى جدوى إدخال آلاف الـ Synsets إلا بعد جهد تحقق بشري كبير

خلاصة النقد من حيث أهمية البحث

الورقة تتناول موضوعاً مهماً في سياق دعم اللغة العربية في NLP

أهمية البحث:

- **قوية من حيث:**
 - الحاجة العملية إلى توسيع Awn
 - توظيف تقنيات حديثة ومعالجة حجم كبير من البيانات
- **لكنها أضعف من حيث:**
 - غياب تجارب تطبيقية على أنظمة حقيقية
 - الاعتماد الكبير على الترجمة الآلية
 - تقديرات متفائلة للأثر بدون بيانات تنفيذية كاملة



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY



Thank You

