

اللغويات الحاسوبية والمعالجة اللغوية الآلية

Journal of Computational Linguistics & Arabic Language Processing



دور النماذج اللغوية الكبيرة في تطوير (WordNet)
قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين
الشبكات الدلالية

السنة الثالثة - العدد الخامس
ذو القعدة ١٤٤٧ هـ - أبريل ٢٠٢٦ م
Vol. 3, No. 5, Apr, 2026

مجمع الملك سلمان
العالمي للغة العربية
King Salman Global Academy for Arabic Language



اللُّغَوِيَّاتُ وَاللِّسَانُ وَالْمَعْرِفَةُ

Journal of Computational Linguistics & Arabic Language Processing

www.ksaa.gov.sa

مجلة علمية مُحَكَّمة تصدر عن مجمع الملك سلمان العالمي للغة العربية





المشرف العام على المجلة

أ.د. عبدالله بن صالح الوشمي

إدارة التحرير

أ.د. منصور الغامدي

رئيس التحرير

د. عبد الرحمن العصيمي

مدير التحرير

د. هاجر القبيشي

أمانة المجلة

هيئة التحرير

د. سارة الحمود

د. عريب العويشق

أ.د. عبدالعزيز المهيوبي

د. أيمن الغامدي

الهيئة الاستشارية

أ.د. محسن عبدالرزاق رشوان (مصر)

أ.د. إيريك أتويل (بريطانيا)

أ.د. نزار حبش (أمريكا)

أ.د. مصطفى عرسان جرار (فلسطين)

أ.د. عبدالمحسن عبيد الثبتي (السعودية)

أ.د. خالد فؤاد شعلان (مصر)

أ.د. صلاح راشد الناجم (الكويت)

◆ قواعد النشر بالمجلة :

سياسات النشر:

تنشر مجلة اللُّغويات الحاسوبية والمعالجة الآلية للُّغة العربيّة البحوث الأصيلة التي لم يسبق نشرها، والمكتوبة بالُّغة العربية أو الإنجليزية، والدراسات والمُلخصات العلميّة المتخصصة في مجال حوسبة اللُّغة، كاللُّغويات الحاسوبية، والمعالجة الآليّة للُّغة العربيّة، وكل ما له صلة بحوسبة اللُّغة.

تُعرض مسودة الأوراق العلميّة فور وصولها على هيئة التحرير لإجراء تقييم مبدئي لها، وفي حال عدم مناسبتها للنشر تُخطر المجلة الباحث بذلك. أمّا البحوث المقبولة مبدئيًا، فتُرسل إلى اثنين من المحكّمين لتقييمها سرّيًا، وتبذل هيئة التحرير طاقتها في التواصل مع المحكّمين للوصول إلى قرار بشأن كل بحث مقدّم في أسرع وقت ممكن، ويُخطر الباحث بنتيجة التحكيم. وفي حالة قبول الورقة للنشر، فإنّ كافة حقوق النشر تؤوّل للمجلة، ولا يجوز نشرها في أي منفذ نشر آخر ورقّيًا أو إلكترونيًا، دون إذن كتابي من رئيس هيئة تحرير المجلة. وتُرسل نسختان ورقيتان مجانيّتان من المجلة إلى المؤلّف عند نشر ورقته.



Journal information:

◆ بيانات المجلة :

مجلة اللُّغويات الحاسوبية والمعالجة الآلية للُّغة العربيّة	Journal of Computational Linguistics & Arabic Language Processing
مجلة علمية محكمة تصدر عن	A peer-reviewed bi-annual journal issued by
مجمع الملك سلمان العالمي للُّغة العربيّة	King Salman Global Academy for Arabic Language
الرياض - المملكة العربيّة السعوديّة	Riyadh- Saudi Arabia
ردمداً النسخة الورقيّة: ١٦٥٨-٩٤٧٥	ISSN paper version: 1658-9475
ردمداً النسخة الإلكترونيّة: ١٦٥٨-٩٤٨٣	ISSN Electronic version: 1658-9483

تعليمات عامة:

١. تُرسل الورقة المقدّمة للنشر للمجلة بصيغة وورد (word) على بريد المجلة: jclalp@ksaa.gov.sa. على أن تُوجّه جميع المراسلات إلى: رئيس هيئة التحرير- مجلة اللُّغويات الحاسوبية والمعالجة الآلية للُّغة العربيّة- مجمع الملك سلمان العالمي للُّغة العربيّة.

٢. يجب أن تحتوي الورقة على ما يلي:

١-٢- صفحة أولى فيها عنوان الورقة، واسم المؤلف أو المؤلفين، وجهة عملهم، وعناوينهم، وتاريخ تقديم الطلب بنشر الورقة.

٢-٢- الكلمات المفتاحية.

٣-٢- مستخلص الورقة باللُّغتين العربية والإنجليزية -في حدود مئة كلمة لكلٍّ منهما-.

٤-٢- المصادر والمراجع.

٥-٢- الملاحق -إن وجدت-.

٣. يُراعى أن تكون الورقة العلمية مكتوبةً بلغة سليمة -سواء كانت عربية أو إنجليزية-، وأن يتراوح عدد كلماتها بين (٣٠٠٠ - ١٠٠٠٠) كلمة، بما في ذلك المراجع والهوامش والجداول والأشكال والملاحق.

٤. إذا كانت الورقة العلمية مكتوبة باللُّغة العربيّة، فيجب الالتزام بكتابة ما يقابل المصطلحات العربيّة باللُّغة الإنجليزيّة في المرة الأولى التي يظهر فيها المصطلح.

٥. ألا يسبق نشر البحث في أي وعاء علمي -مالم ترهئة التحرير خلاف ذلك-.

المواصفات الفنية للبحث:

- يكون مقاس الصفحة A4، وتُقسَّم إلى عمودين، وتكون المسافة بين العمودين ١ سم، وتكون الهوامش ٥,٢ سم عن يمين ويسار الصفحة، و٥٨,٣ سم من أعلى وأسفل الصفحة.
- تُترك مسافة 5.0 سم في بداية كل فقرة من فقرات المتن.
- يجب الفصل بمسافة سطر واحد بين نهاية محتوى الفقرات داخل المتن وعناوين الفقرات التي تليه.
- يجب ترك مسافة سطر واحد قبل وبعد الجدول أو الشكل والصورة والمتن السابق والتالي لهما.
- يُستخدم الخط (Sakkal Majalla) للنص العربي، والخط (Times New Roman) للنص الأجنبي وبحجم أقل بدرجتين عن الخط العربي. ويكون حجم الخط العربي كالآتي:
 - العنوان الرئيس: ١٤، عريض.
 - العناوين الفرعية: ١٢، عريض.
 - النصوص: ١٢، عادي.
 - الهوامش السفلية: ١٠، عادي.
- يجب ترقيم صفحات البحث.
- يجب ترقيم المعادلات.
- تُرتَّب المصادر والمراجع أبجدياً بنظام APA.
- تُفصل المراجع الأجنبية عن المراجع العربية في قائمة المصادر والمراجع، وتأتي المراجع /المصادر العربية أولاً في الأبحاث العربية والمراجع /المصادر الإنجليزية أولاً في الأبحاث الإنجليزية.

الجداول:

- تُدرج الجداول بعد الفقرة التي ذُكرت بها، وذلك على النحو التالي:
 - إذا كان عرض الجدول ٧,٥ سم يوضع في عمود واحد أسفل الفقرة مباشرة.
 - إذا كان عرض الجدول أكثر من ٧,٥ سم يوضع أسفل الصفحة التي ذكر بها أو أعلى الصفحة التي تليها.
- تُرقّم الجداول بين قوسين في أقصى يمين الصفحة مسبوقة بكلمة «جدول». مثال: جدول (١).
- يُوضع عنوان الجدول في أعلى الجدول مباشرة، في أقصى اليمين.
- يجب ألا يتجاوز عرض الجدول هوامش الصفحة المذكورة في الفقرة السابقة.
- يُكتب مصدر الجدول والتعليقات المصاحبة له -إن وجدت- أسفله.
- تُضبط بيانات الجدول غير الرقمية أقصى يمين الخلية.
- تُضبط بيانات الجدول الرقمية وسط الخلية.
- إذا استمر الجدول على أكثر من صفحة، فيجب تكرار رأسه في الصفحات التالية ووضع عبارة «تابع جدول (...)» في أعلاه.

الصُور والأشكال:

- تُدرج الصور والأشكال بعد الفقرة التي ذُكرت على النحو التالي:
 - إذا كان عرض الصورة ٧,٥ سم تُوضع في عمود واحد أسفل الفقرة مباشرة.
 - إذا كان عرض الصورة أكثر من ٧,٥ سم تُوضع أسفل الصفحة التي ذكر بها أو أعلى الصفحة التي تليها.
- تُرقّم الصور والأشكال، ويوضع الرقم بين قوسين مسبوقة بكلمة «شكل». مثال: شكل (١).

- يُدرج عنوان الشكل أسفل الشكل مباشرةً متوسطًا بين حدوده بحيث لا يتجاوز عنوان الشكل سطرين فقط.
- في حال وجود تعليق مصاحب على الشكل أو الصورة يجب دمجها مع الشكل وحفظه على هيئة صورة لا تقل دقتها عن (300 dpi).

مواصفات كتابة المعادلات:

- تُكتب المعادلات الرياضية باستخدام أحد البرنامجين التاليين : (Math Type) أو (Equation Editor).
- تُرقّم المعادلات، ويوضع الرقم بين قوسين في أقصى يسار العمود مقابل المعادلة على مستوى البحث.

مواصفات كتابة الاقتباس:

- إذا كان النص المقتبس ٣ أسطر أو أقل، فيجب أن يوضع بين علامتي تنصيص «...»، ويليه توثيق النص على النحو التالي: (العائلة، السنة: الصفحة).
- إذا كان النص أكثر من ٣ أسطر، فيجب أن يُميز النص المقتبس بكتابته في منتصف الصفحة، وتُترك من بداية السطر ونهايته هوامش اسم، ويوضع بين علامتي تنصيص «...»، ويُكتب بحجم ١٢ للنص العربي و١٠ للنص الإنجليزي، وينتهي بتوثيق النص على النحو التالي: (العائلة، السنة: الصفحة).

مواصفات كتابة الحواشي:

- «تُدرج الحواشي أسفل كل صفحة، مع ترقيمها آليًا حسب ورودها في كل صفحة».

◆ صفة المجلة:

مجلة اللُّغويات الحاسوبية والمعالجة الآلية للُّغة العربيَّة (Journal of Computational Linguistics & Arabic Language Processing) هي مجلة علمية محكَّمة، نصف سنوية (ISSN:1658-9475 - ISSN:1658-9483)، وتصدر عن مجمع الملك سلمان العالمي للُّغة العربية في الرياض بنسختين: إلكترونية، وورقية (ذات إصدار محدود). وتتيح المجلة الوصول الحر (Open Access) إلى البحوث المنشورة فيها، ولا تتقاضى رسوماً على النشر. وهي مجلة تهتم بنشر الأبحاث العلمية المحكمة في مجال حوسبة اللُّغة، وتشمل: اللُّغويات الحاسوبية (Computational Linguistics)، والمعالجة الآلية للُّغة العربية (Arabic Language Processing) وما يتبعها من تخصصات متَّصلة بها. وللاطلاع على أهداف المجلة وشروط النشر وكل التفاصيل ذات العلاقة يمكنكم زيارة موقع المجلة: jclalp.ksaa.gov.sa

البحوث المنشورة بالمجلة تعبر عن رأي أصحابها، ولا تعبر بالضرورة عن رأي المجمع، ويتحمل مؤلفوها المسؤولية كاملة عن صحة المعلومات والاستنتاجات ودقتها.

محتويات العدد

١٢٤

دور النماذج اللغوية الكبيرة في تطوير (WordNet)
قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين
الشبكات الدلالية

محمد مهيبوب البحيري - جيهان الجرادي

كلمة التحرير

عزيزي القارئ/القارئة - عزيزي الباحث/الباحثة:

يُسعدنا أن نرحب بكم في العدد الخامس من المجلة العلمية المحكمة: "اللُغويات الحاسوبية والمعالجة الآلية للغة العربية"، الصادرة عن مجمع الملك سلمان العالمي للغة العربية، والمتخصصة في مجال حوسبة اللغة العربية.

تواصل المجلة مسيرتها العلمية الرصينة في نشر البحوث والدراسات التي تُسهم في تطوير الفهم العميق لتطبيقات الحوسبة في اللغة العربية، وتعمل على رصد أحدث الاتجاهات في مجال الذكاء الاصطناعي اللغوي وأدوات التحليل اللغوي الآلي. كما تُعنى المجلة باستشراف آفاق المستقبل، واستكشاف التحديات والفرص التي تواجه تطوير اللغة العربية رقمياً، بما يعزّز حضورها في فضاء التقنية والمعرفة.

ويأتي العدد الخامس استمراراً لهذا التوجّه العلمي البناء، حاملاً بين صفحاته ثلاثة أبحاث علمية متميزة واستعراضاً لكتاب مهم، تعالج جميعها موضوعات معاصرة ومتنوعة في صميم حوسبة اللغة العربية:

يأتي البحث الأول بعنوان "دراسة مسحية لآليات تقييم النماذج اللغوية الكبيرة"، وهو بحث يستعرض أبرز الممارسات والأطر العالمية لتقييم النماذج اللغوية الضخمة، ويناقش سبل تطوير آليات تقييم مخصصة للغة العربية تراعي خصوصيتها اللغوية والثقافية، مقدماً إطاراً لمقاييس شاملة تضمن موثوقية الأداء.

أما البحث الثاني "أنظمة الترجمة الآلية من العربية إلى الإنجليزية وتقييم بعض برامجها"، فيتجه نحو تقييم ترجمة ثلاثة برامج آلية حاسوبية من خلال تحليل مقارن لجمل عربية تحمل اللبس الصرفي والتركيبى والدلالي، للوقوف على مواطن الضعف واقتراح الحلول اللغوية والتقنية لمعالجتها.

ويأتي البحث الثالث "دور النماذج اللغوية الكبيرة في تطوير (WordNet) قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين الشبكات الدلالية"؛ ليُجري تحليلاً دلاليّاً لتوظيف النماذج اللغوية الكبيرة جنباً إلى جنب مع الخبرة البشرية في بناء شبكة كلمات عربية دقيقة، تستوعب العلاقات الدلالية والتعقيدات الصرفية للغة.

وأخيراً، يقدم العدد استعراضاً علمياً لكتاب "لسانيات المدونات: مقدمة موجزة"، والذي يمثل مدخلاً مهماً يجمع بين الجانبين النظري والتطبيقي في التعريف بلسانيات المدونات، ويسلط الضوء على دور المدونات اللغوية في دراسة المعنى واستخراج المتلازمات اللفظية وتحليل الخطاب الفعلي.

وختاماً، نُجدد ترحيبنا الدائم بتواصلكم المثمر، ومقترحاتكم القيّمة التي تُمثل رافداً أساسياً للارتقاء بمسيرة المجلة وتجويد محتواها العلمي. سائلين المولى عز وجل أن يكلل هذه الجهود بالنجاح، وأن يجعل هذا العدد لبنةً نافعةً تخدم الباحثين، وتُثري ميدان حوسبة اللغة العربية وتقنياتها.

هيئة التحرير

دور النماذج اللغوية الكبيرة في تطوير (WordNet)

قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين الشبكات الدلالية

المعرف الرقمي: <https://doi.org/10.60161/2521-003-005-003>

محمد مهيوب البحيري^{١*} جيهان الجراي^٢

أ قسم اللغات والترجمة، كلية الآداب، جامعة طيبة المدينة المنورة، المملكة العربية السعودية

ب قسم اللغة الإنجليزية، كلية التربية، جامعة تعز، اليمن

نُشر في	قبل في	رُوجع في	استقبل في
١٤٤٧/١١/١٣	١٤٤٧/٠٧/٠٢	١٤٤٧/٠٥/٢٩	١٤٤٧/٠٥/١٢

الملخص

تُجري هذه الدراسة تحليلاً دلاليًا للمفردات العربية ضمن إطار لبناء شبكة الكلمات العربية (Arabic WordNet)، باستخدام النماذج اللغوية الكبيرة (DeepSeek و Gemini و ChatGPT و Grok) مع الخبرة البشرية لإنشاء قاعدة بيانات معجمية دقيقة. تركز الدراسة على معاني الكلمات والعلاقات الدلالية والتعقيدات الصرفية. تعتمد المنهجية على جمع البيانات من المعاجم وقواعد البيانات اللغوية، وتحليل النماذج اللغوية، وتقييم الخبراء اللغويين للتأكد من دقتها في تصنيف المترادفات والمتضادات، وبناء العلاقات الجزئية. أظهرت النتائج قصوراً في تجميع المرادفات وتصنيف المتضادات، لكنها أبرزت قوة برنامج Gemini في التصور، و ChatGPT في الاستخدام السياقي، و Grok و DeepSeek في اللغويات الحاسوبية، مع تفوق Grok في الصرف. تُقدم هذه الدراسة خطوة نوعية نحو بناء شبكة معجمية باستخدام أدوات الذكاء الاصطناعي، مدعومة بالتقييم البشري، مما يُعزز الدقة الدلالية والكفاءة، ويدعم تطبيقات معالجة اللغة الطبيعية.

الكلمات المفتاحية

بيانات معجمية عربية، النماذج اللغوية الكبيرة، الشبكات الدلالية، الذكاء الاصطناعي، شبكة الكلمات العربية.

*الباحث المراسل: محمد البحيري (البريد الإلكتروني: mqassem@taibahu.edu.sa).

The Role of Large Language Models (LLMs) in Developing an Arabic WordNet: Assessing Capabilities and Enhancing Semantic Networks

المعرف الرقمي: <https://doi.org/10.60161/2521-003-005-003>

Mohammad Mahyoob Albuhairey^{A*} Jeehaan Algaraady^B

A Languages & Translation Department, Taibah University, Madina, KSA

B Department of English, Taiz University, Taiz, Yemen

Received on	Revised on	Accepted on	Published on
13-11-2025	20-11-2025	22-12-2025	30-04-2026

ABSTRACT

This study conducts a semantic analysis of Arabic vocabulary within the framework of an (Arabic WordNet), using large language models, (LLMs) (ChatGPT, Gemini, DeepSeek, and Grok) combined with human expertise to create an accurate lexical database. It examines word meanings, semantic relationships, and morphological complexities. The data are collected from dictionaries and linguistic databases, analyzed by LLMs, and evaluated by linguistic experts to ensure accuracy in classifying synonyms and antonym, and in identifying distinct relationships. The results revealed the LLMs shortcomings in synonym classification and semantic grouping, however, Gemini showed visualization capabilities, ChatGPT's showed contextual analysis, and Grok and DeepSeek's showed linguistic analysis, with Grok excelling in morphology. This study is a valuable contribution in building lexical network using artificial intelligence tools, with human evaluation, thereby enhancing semantic accuracy and supporting natural language processing applications.

KEYWORDS

Arabic Lexical Data, Large Language Models (LLMs), Semantic Networks, Artificial Intelligence (AI), Arabic WordNet (AWN).

*Corresponding author: Mohammad Albuhairey (Email: mqassem@taibahu.edu.sa).

١- المقدمة

أحدث الظهور والتقدم المتسارع في تقنيات معالجة اللغة الطبيعية ثورة عارمة في مجال الذكاء الاصطناعي، حيث برزت في السنوات الأخيرة النماذج اللغوية الكبيرة (Large Language Models: LLMs) بقدرات فريدة وغير مسبوقة في محاكاة العقل البشري من خلال فهم اللغة البشرية وتوليدها (Abuhairy et al, 2025; Algaraady & Mahyoob, 2025). وقد أظهرت هذه النماذج أداءً عاليًا ومتميزًا في تطبيقات الذكاء الاصطناعي والمهام اللغوية المختلفة (Rogers, et al, 2020; Brown et al., 2020 Abuhairy et al, 2023; Algaraady & Mahyoob, 2023). ومع ذلك، ما تزال فعاليتها في المجالات المتخصصة، وخاصةً في مجال المعاجم وشبكات الكلمات والأنطولوجيا العربية محدودة، وتتطلب معالجةً وتدريبًا خاصًا لإنتاج تحليل لغويٍّ معمّق يتماشى مع خصوصية اللغة العربية.

تُمثل المعاجم وشبكة الكلمات والأنطولوجيا العربية قاعدةً أساسيةً في تطبيقات اللغويات الحاسوبية كأنظمة الترجمة الآلية المتقدمة، ومحركات البحث الدلالية، وأدوات التحليل النصي المتخصصة، وغيرها، ذلك لاستغراقها في العلاقات الدلالية المعقدة (Navigli & Ponzetto, 2012). تتضمن المعاجم عملية بناء القواميس ودراسة معاني الكلمات وعلاقاتها الدلالية وأنماط استخدامها. وقد أنتجت المعاجم العربية التقليدية أعمالاً ضخمة مثل: لسان العرب والمعجم الوسيط، وغيرها الكثير، كمراجع أساسية لدراسات اللغة العربية (Habash, 2010). في حين تمثل شبكة الكلمات مجموعة من الشبكات الدلالية التي تربط بين المفاهيم، وتشكل شبكة الكلمات أنطولوجيا جزئية بحيث يشكل الارتباط بين مجموعة من الشبكات مكون أكبر وأشمل يغني العديد من استخدامات الذكاء الاصطناعي وتطبيقات الويب الدلالي، بما يسمى بالأنطولوجيا. في حين تعرف الأنطولوجيا بأنها مجموعة من المعارف التي تترابط مكوناتها مشكلةً هيكلًا ضخمًا من المفاهيم المشتركة في حقول تربطها علاقات دلالية بحيث تشمل مختلف فروع المعرفة للغة (السعيد طه، ٢٠١٢؛ دلول، ٢٠٢٠، 2004 Holes).

إن بناء شبكة لغوية عربية أو أنطولوجيا عربية متكاملة يتطلب فهمًا عميقًا للعلاقات الدلالية بين الكلمات، والقدرة على تحديد المرادفات والأضداد، وإدراك السياقات المختلفة لاستخدام الكلمات، بالإضافة إلى فهم التعقيدات الصرفية والنحوية للغة (Navigli & Ponzetto, 2012). ويمهد بناء شبكة كلمات معيارية الطريق لبناء أنطولوجيا متكاملة بحيث تتوافر الجهود لتغطية المحتوى المعرفي الضخم لا سيما أن هذا المحتوى في تنام مستمر. بيد أن بناء شبكة لغوية يشكل تحديًا أكثر تعقيدًا في حالة اللغة العربية، خاصةً في مجال المعالجة الحاسوبية، ويعود ذلك إلى تعقيد بنيتها النحوية، وعلاقاتها الدلالية المتداخلة والمعقدة، مثل: المعاني المتعددة للكلمة الواحدة المرتبطة بالسياق، الفروق الدلالية الدقيقة بين الكلمات المترادفة، بالإضافة إلى نظامها الصرفي الغني والمعقد الذي يعرض الجذور (التي تحمل المعنى الأساسي) والأنماط، وعمليات الاشتقاق الشاملة من الجذور عبر تطبيق أوزان صرفية مختلفة (Navigli & Ponzetto, 2012; Rogers, et al, 2020). يتطلب هذا التعقيد الدلالي والصرفي من النماذج اللغوية الضخمة قدرةً فائقةً على التمييز بين المعاني المختلفة وفهم السياقات الثقافية والتاريخية التي تؤثر على استخدام الكلمات. علاوة على ذلك، يشكل الغموض الإملائي (بسبب غياب التشكيل في النصوص العادية، والتنوع اللهجي الواسع) تحديًا إضافيًا يجعل من تقييم قدرات النماذج اللغوية في بناء المعجمية العربية مهمة بالغة التعقيد (الحلي وآخرون, ٢٠٢٣; أدغيري, ٢٠٢٢; (Farghaly & Shaalan, 2009; Al-Garaady, 2022; Mahyoob & Algaraady, 2018).

هناك العديد من الجهود المبذولة في هذا المجال. فعلى سبيل المثال تعد شبكة الكلمات العربية (Arabic Word Net) التي أنجزت خلال الفترة ما بين ٢٠٠٥ و٢٠٠٧ وفق تصميم ومكونات شبكة الكلمات المعتمدة دوليًا (Princeton Wordnet) والتي تعد من أشهر الشبكات العربية، غير أن هذه الشبكة لم تقنع وتلبي تطلعات اللغويين العرب، كونها بنيت اعتمادًا على الترجمة من لغة أخرى، دون مراعاة خصوصية وبنية اللغة العربية، وعدم شمولها على كل مفردات اللغة العربية والعلاقات الدلالية المختلفة بينها، وغيرها من الدراسات التي سنستعرضها في القسم التالي. شكّل هذا القصور والفجوة الواسعة في الأدبيات في إنتاج قاعدة بيانات شاملة لشبكة الكلمات العربية منطلقًا أساسيًا للقيام بهذه

الدراسة، بالإضافة إلى أنه بالرغم من التطور والأهمية المتزايدة للنماذج اللغوية الضخمة؛ إلا أن أدائها في بناء وإثراء المعاجم وشبكة الكلمات والأنطولوجيا العربية لم يُقَيِّم بشكل منهجي باستخدام معايير شاملة ومتخصصة. حيث ركزت معظم الدراسات على تطبيقات أخرى، مثل: أنظمة الترجمة الآلية المتقدمة، ومحركات البحث الدلالية، وأدوات التحليل النصي المتخصصة، والإجابة عن الأسئلة، واستدعاء المعلومات، وتحليل المشاعر، وتصنيف وتلخيص النصوص، وغيرها (Jarrar, 2021)

ومن هذا المنطلق، يقوم الباحثان في هذه الدراسة باستغلال إمكانات وتقييم قدرات النماذج اللغوية الضخمة مثل: Gemini، DeepSeek، ChatGPT، وGrok. في بناء شبكة كلمات عربية شاملة تتضمن عددًا من الجوانب، متمثلة في كلمات مُفصلة تجمع بين المرادفات، والأضداد، والأسماء المتشعبة، والأسماء المختصرة، والإحالات المتقاطعة الموسعة، مع التركيز على تعقيدات الصرف العربي والفروق الدلالية الدقيقة وفقًا لمعايير محددة من أجل توجيه جهود التطوير وتحسين أداء هذه النماذج في التعامل مع العربية. ولم يتم اختيار هذه الجوانب عشوائيًا، بل جاء بعد دراسة عميقة لمتطلبات بناء الشبكة اللغوية والدراسة بخصائص اللغة العربية. وتكمن أهمية الدراسة الحالية في إثراء البحث المتنامي حول معالجة اللغة العربية من خلال تقديم منهجية ومعايير تقييم لبناء شبكة الكلمات. كما تقدم الدراسة توصيات عملية للباحثين والمطورين من أجل تحسين أداء النماذج اللغوية في بناء الشبكات اللغوية والأنطولوجيا العربية.

٢- مشكلة الدراسة

يواجه تطوير شبكة الكلمات العربية (Arabic WordNet) تحديات كبيرة تتعلق بمعالجة تعقيدات اللغة العربية، خاصة الطبيعة الاشتقاقية للجذور اللغوية والفروق الدلالية الدقيقة بين المفردات. تعتمد العديد من الدراسات السابقة على الترجمة من أنطولوجيات أجنبية أو استخدام طرق إحصائية محدودة النطاق، مما يؤدي إلى قصور في الشمولية والدقة، مثل: التجميع غير الدقيق للمرادفات، والتصنيف الخاطئ للأضداد، والثغرات في العلاقات الجزئية والدلالية. تُظهر النماذج اللغوية الكبيرة Gemini، DeepSeek، ChatGPT،

إمكانات واعدة في بناء شبكة المفردات العربية، لكنها تفتقر إلى التكامل مع الخبرة البشرية لضمان الدقة في معالجة السياقات الثقافية واللغوية العربية. هذا يحد من قدرة هذه النماذج على بناء قاعدة بيانات معجمية عربية موثوقة تلبي متطلبات التطبيقات التقنية والتعليمية، مثل: معالجة اللغة الطبيعية، واسترجاع المعلومات، والتعليم اللغوي. بالتالي، هناك حاجة ملحة لتطوير إطار تحليلي يدمج بين قدرات النماذج اللغوية والخبرة البشرية؛ لتعزيز الدقة الدلالية والنطاق المعجمي للشبكة المعجمية العربية.

٣- أهداف الدراسة

- أ. تهدف هذه الدراسة إلى تقييم أداء النماذج اللغوية ChatGPT، Gemini، DeepSeek، Grok في تحليل العناصر المعجمية العربية تحليلاً دلالياً عميقاً من حيث صياغة مجموعات مترادفة دقيقة، وتحديد الأضداد، والمصطلحات الأعم/الأخص، والعلاقات الجزئية، مع التركيز على معالجة تعقيدات الصرف العربي والفروق الدلالية.
- ب. تهدف الدراسة إلى تحديد نقاط الضعف والقوة في مخرجات النماذج اللغوية، واقتراح تصحيحات مستهدفة لتعزيز موثوقية المعجمية العربية المدعومة بالذكاء الاصطناعي.

٤- أسئلة الدراسة

لتحقيق الأهداف المذكورة، صيغت الأسئلة البحثية الآتية:

- أ. إلى أي مدى تتمكن النماذج اللغوية الكبيرة ChatGPT، Gemini، DeepSeek، Grok من صياغة مجموعات مترادفة دقيقة وشاملة للمعجمية العربية مع مراعاة تعقيدات الصرف العربي والفروق الدلالية الدقيقة؟
- ب. ما العيوب الرئيسية في أداء النماذج اللغوية الكبيرة عند تحليل المعجمية العربية؟ وكيف يمكن تصحيح هذه العيوب لتطوير إطار دلالي أكثر موثوقية في سياق شبكة الكلمات العربية؟

٥- الدراسات السابقة

تعد الدراسات السابقة في مجال الأنطولوجيا والشبكات الدلالية للغة العربية (Arabic WordNet) أساساً لفهم التطورات والمنهجيات المستخدمة في هذا المجال الحيوي. وتهدف هذه الجزئية إلى استعراض أبرز الأبحاث التي تناولت بناء الأنطولوجيات وتوسيع الشبكات الدلالية للغة العربية، مع التركيز على التحديات والحلول المقترحة. ستُحلل هذه الدراسات لتقديم رؤية شاملة حول الجهود المبذولة في هذا السياق، مما يمهد الطريق للأبحاث المستقبلية.

ركزت العديد من الدراسات على بناء الشبكات الدلالية للغة العربية، والتي تعد مورداً لغوياً مهماً للمعالجة الآلية للغة. على سبيل المثال، هدفت دراسة السعيد طه (٢٠١٢) "شبكة الكلمات العربية: المنهجية والتطبيق". إلى بناء شبكة دلالية للغة العربية على غرار WordNet الإنجليزية، مع مراعاة الخصائص اللغوية الفريدة للغة العربية، ونظامها المعجمي والدلالي على مستوى المباني والمعاني. وقد قدمت الدراسة منهجية مفصلة لإنشاء هذه الشبكة. مسطرة الضوء على أهمية الشبكة في تعزيز الصناعة المعجمية العربية المعاصرة، خاصةً في المعجمات المتخصصة.

قدمت دلول (٢٠٢٠) إطاراً لبناء أنطولوجيا معجمية شاملة للغة العربية، مع الأخذ في الاعتبار التحديات المرتبطة بتعقيد اللغة العربية وتنوعها. كما استعرضت الدراسة أهمية الأنطولوجيات في تمثيل المعرفة اللغوية، وكيف يمكن استخدامها لتحسين أداء تطبيقات معالجة اللغة الطبيعية. وطورت الحلبي وآخرون (٢٠٢٣) إطار تعلم أنطولوجيا عربية باستخدام تقنية التقابل بين شبكة الكلمات العربية وخصائص مقالات الموسوعة العربية، محققة دقة مفاهيم بلغت ٠,٨٨، وعدد مصطلحات ١٢٨٠٠ بدقة ٠,٥٢ للعلاقات الهرمية، و٠,٧٧ لعلاقات الترافق، و٠,٤٠ لعلاقات التشابه. قدمت أدغيري (٢٠٢٣) منهجية مقترحة لبناء وتطوير أنطولوجيا عربية في مجال اللسانيات الحاسوبية. وسلطت هذه الدراسة الضوء على أهمية بناء أنطولوجيات متخصصة للغة العربية في مجالات محددة، مثل: اللسانيات الحاسوبية. وقدمت منهجية مقترحة لبناء هذه الأنطولوجيات، مع التركيز على مراحل التخطيط، والتصميم، والتنفيذ، والتقييم. وتُعد هذه المنهجية إطاراً عملياً يمكن الاستفادة منه في بناء أنطولوجيات عربية دقيقة ومتخصصة.

وفي إطار الجهود المبذولة لبناء أنطولوجيات رسمية للغة العربية، قدم *Jarrar et al*. (٢٠١١-٢٠٢١) دراسة بعنوان "بناء أنطولوجيا عربية رسمية". وتُعد هذه الدراسة من الأعمال الرائدة في مجال بناء الأنطولوجيات العربية، حيث قدمت منهجية شاملة لبناء أنطولوجيا لغوية رسمية للغة العربية باستخدام ١٥٠ معجمًا. وركزت الدراسة على تمثيل المفاهيم التي تنقلها المصطلحات العربية باستخدام علاقات دلالية رسمية. وتُعد هذه الأنطولوجيا قابلة للوصول والبحث من خلال محرك بحث معجمي. وقدم *Alsudias L. et al*. (٢٠٢٠) أنطولوجيا طبية متخصصة بالأمراض المعدية بطرق إحصائية من مستندات طبية، مدمجة مع معاجم عربية. تُعد عملية ربط الشبكات الدلالية بمصادر معرفية أخرى خطوة مهمة نحو إثرائها وزيادة فائدتها. في هذا السياق، تناولت دراسة *Saif et al*. (٢٠١٧) ربط مجموعات المرادفات في الشبكة الدلالية العربية بمقالات ويكيبيديا باستخدام ميزات أحادية اللغة وثنائية اللغة. وهدفت هذه الدراسة إلى تطوير طريقة لربط مجموعات المرادفات (synsets) في الشبكة الدلالية العربية بمقالات ويكيبيديا المقابلة لها. واستعملت هذه الطريقة ميزات أحادية اللغة وثنائية اللغة للتغلب على نقص المعلومات الدلالية في الشبكة الدلالية العربية. وأظهرت النتائج التجريبية أن الطريقة المقترحة تحقق نتائج واعدة وتتفوق على الطرق الحديثة التي تعتمد على الميزات أحادية اللغة فقط. وقد استخدمت هذه الطريقة لزيادة تغطية الشبكة الدلالية العربية عن طريق إدراج مجموعات مرادفات جديدة من ويكيبيديا. أنشأ *AlMaayah et al*. (٢٠١٦) شبكة الكلمات العربية القرآنية (QAWN) بتوليد مرادفات آليًا من قواميس عربية، وقدم *Ishkewy et al*. (٢٠١٤) أنطولوجيا الأزهرية (أنطولوجيا معجمية عربية) من نصوص القرآن الكريم، بالاعتماد على ربط يدوي بعلاقات معجمية. وقدمت هذه الدراسة أنطولوجيا معجمية للغة العربية، جمعت الكلمات في مجموعات مرادفات (synsets) وسجلت العلاقات الدلالية مثل: "الترادف، والتضاد، وعلاقات النوع (hypernymy)، وعلاقات الجزء (meronymy)، وعلاقات الكل (holonymy)". وتُعد هذه الأنطولوجيا موردًا قيمًا للباحثين في مجال معالجة اللغة الطبيعية العربية، وساهمت في تطوير تطبيقات مثل: استرجاع المعلومات، والترجمة الآلية، وفهم اللغة. اعتمد *Mazari et al*. (٢٠١٢) في دراستهم مفاهيم أنطولوجيا Gold وخصائص نصوص عربية، مستخدمين تقنيات إحصائية لربط المفاهيم بعلاقات هرمية

وغير هرمية. وصمّم AlBarhamtoshy *et al* (٢٠٠٩) شبكة كلمات عربية عبرترجمة شبكة الكلمات الإنجليزية باستخدام قاموس متعدد اللغات، بينما وسّع Rodríguez (٢٠٠٨) شبكة الكلمات العربية شبه آلياً باستخدام قواعد صرفية وقواعدية مع الاستدلال بايزي. وطوّر Alkhalifa *et al* (٢٠١٠) أنطولوجيا عربية مترجمة بإضافة أسماء كيانات من ويكيبيديا، واستخدم Aliane *et al* (٢٠٠٩) أنطولوجيا Gold كنواة لأنطولوجيا عربية مركزية ضمن مشروع الخليل. أنشأ Belkredim *et al* (٢٠٠٨) أنطولوجيا DEAR-ONTO الاشتقاقية، معتمدين على الأفعال السماعية كجذور لتحديد الأصناف المشتقة.

تُعد عملية استخراج العلاقات الأنطولوجية من النصوص العربية تحديًا مهمًا في مجال معالجة اللغة الطبيعية. في هذا الصدد، قدم Al Zamil & Al-Radaideh (٢٠١٤) دراسة حول "الاستخراج التلقائي للعلاقات الأنطولوجية من النص العربي". اقترحت هذه الدراسة منهجية لاستخراج العلاقات الأنطولوجية من النص العربي، مع التركيز على الميزات الدلالية للغة العربية. وقدمت نموذجًا رسميًا لاستخراج العلاقات الأنطولوجية، ووضحت كيفية استخدام الأنماط الدلالية المعجمية لتحديد هذه العلاقات. كما أظهرت النتائج التجريبية أن المنهجية المقترحة تحقق نتائج واعدة في أتمتة عملية استخراج العلاقات الأنطولوجية من النصوص العربية. وتُعد الأنطولوجيات المعجمية مكونًا أساسيًا في العديد من تطبيقات معالجة اللغة الطبيعية.

١- منهجية الدراسة

تتبع المنهجية المستخدمة في هذه الدراسة نهجًا تحليليًا وصفيًا لدراسة المفردات العربية ضمن إطار شبكة الكلمات العربية (Arabic WordNet) باستخدام النماذج اللغوية الكبيرة (ChatGPT، Gemini، DeepSeek، Grok)، مع دمج الخبرة البشرية لتعزيز الدقة والشمولية. بدأت الدراسة بجمع بيانات من معاجم عربية موثوقة (مثل: لسان العرب والمكنز الكبير والمعجم الوسيط وغيرها) لتحديد المعاني المتعددة والعلاقات الدلالية (مرادفات، أضداد، مصطلحات أعم/أخص) واللغوية (مشتقات صرفية). كُلفت النماذج اللغوية بتحليل هذه المفردات؛ لإنشاء شبكات معجمية تشمل مجموعات مترادفة (Synsets)، مع التركيز



دور النماذج اللغوية الكبيرة في تطوير (WordNet) قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين الشبكات الدلالية

على معالجة تعقيدات الصرف العربي والفروق الدلالية الدقيقة. أُجري تقييم مخرجات النماذج من خبراء لغويين لقياس دقتها في تجميع المرادفات، وتصنيف الأضداد، والعلاقات الجزئية، مع مراجعة السياقات في نصوص عربية متنوعة لضمان الموثوقية. دُمج التحليل البشري لتصحيح أوجه القصور، مثل: التجميع غير الدقيق والتصنيف الخاطئ وغيرها، كما سيتم ذكرها في الأقسام التالية. وتُحقق من النتائج عبر مقارنة السياقات في النصوص العربية ومراجعة التحليلات؛ لضمان موثوقية الشبكة المعجمية وملاءمتها للتطبيقات التعليمية والحاسوبية.

٧- التحليل والمناقشة

٧-١- النماذج اللغوية وبناء شبكة الكلمات العربية (WordNet Arabic)

تتناول هذه الدراسة مقارنة تحليلية بين أربعة نماذج ChatGPT و Gemini و DeepSeek و Grok في بناء شبكة الكلمات العربية (Arabic WordNet) لبعض الجذور والكلمات العربية على سبيل المثال "كتب"، و"ضرب"، و"علم". حيث طُلب من النماذج الأربعة بناء شبكة دلالية للجذور كمرحلة أولى. وفي المرحلة الثانية أُجريت هندسة للأمر المقدم للنماذج الأربعة بتفاصيل دقيقة تتعلق ببناء شبكة معجمية تفصيلية متكاملة. يهدف التحليل إلى تقييم قدرات كل نموذج في تقديم شبكة معجمية شاملة ودقيقة للمعاني المتعددة، والعلاقات الدلالية، والصرفية لهذه الجذور، مع التركيز على مدى التفصيل، والدقة، والشمولية، ونقاط القوة والضعف لكل نموذج، وتحديد النموذج الأكثر فعالية في تقديم مخرج أفضل ومقارنته بالأداء البشري.

٧-٢- المرحلة الأولى لعملية البناء والتحليل بدون هندسة الأوامر

٧-٢-١- تحليل قدرات النموذج ChatGPT

يقدم نموذج ChatGPT تحليلاً مفصلاً ومنظماً للجذور الثلاثية، مع تقسيم واضح للمعاني المتعددة (Senses) لكل جذر. بالنسبة لـ "كتب"، يحدد النموذج ثلاثة معانٍ رئيسة (الكتابة، التأليف، الفرض / القدر)، مع تقديم مجموعات مرادفات (Synsets) لكل معنى،

مثل: "كتب، دون، سطر" لمعنى الكتابة. كما يوفر أمثلة توضيحية، مثل: "كتب رسالة إلى صديقه"، مما يعزز فهم السياق كما هو موضح في الشكل ١. بالنسبة لـ "ضرب"، يقدم النموذج أربعة معانٍ (الضرب الجسدي، ضرب المثل، الطرق، الضرب الرياضي)، مع مرادفات مثل: "ضرب، لطم، صفع" وأمثلة سياقية. أما بالنسبة لـ "علم"، فإنه يحدد أربعة معانٍ (المعرفة، التدريس، المعرفة كاسم، وضع علامة)، مع مرادفات وأمثلة دقيقة. يتميز هذا النموذج بتقديم هيكلية هرمية واضحة للعلاقات الدلالية (مرادفات، وأمثلة، وتفسيرات)، مما يجعله مناسباً لفهم العلاقات بين الكلمات بسهولة. ومع ذلك، يفتقر النموذج إلى تحديد الأضداد أو العلاقات الأعم/الأخص بصرحة، مما يحد من شموليته مقارنة بنماذج أخرى.

شكل (١) تحليل وبناء شبكة الكلمات للجذر "كتب" باستخدام نموذج ChatGPT



٧-٢-٢- تحليل قدرات النموذج Grok

النموذج Grok يركز على تحليل الجذور الثلاثية (كتب، ضرب، علم) من منظور صرفي ودلالي، مع تقديم هيكلية مشابهة لشبكة الكلمات. يوفر النموذج تفاصيل حول الأجزاء النحوية (فعل، اسم، صفة)، والمرادفات، والمصطلحات الأعم/الأخص، مثل: "التواصل" كمصطلح أعم لـ "كتب" و "لَکَم" كمصطلح أخص لـ "ضرب". كما يشمل النموذج مفاهيم مرتبطة (مثل: القراءة والكتابة لـ "كتب") وعلاقات صرفية، مثل: مشتقات "كتاب" و "مكتوب". يتميز النموذج Grok بتقديم رسوم بيانية توضح العلاقات بين الكلمات، مما يسهل تصور الشبكة الدلالية. ومع ذلك، فإن النموذج يفتقر إلى أمثلة سياقية واضحة، مما قد يقلل من فائدته في التطبيقات العملية. كما أن تركيزه على العلاقات الصرفية قد يجعله أقل شمولية في تغطية المعاني المتعددة مقارنة بالنموذج ChatGPT.

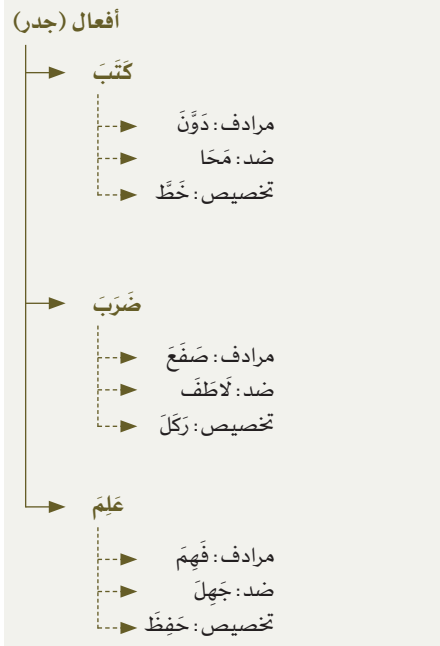
شكل (٢) تحليل وبناء شبكة الكلمات للجذر "كتب" باستخدام النموذج Grok

المجموعة ١: كتب (الكتابة التسجيل)	١. كتب
- الكلمات: كتب (كتابة)، سَجَل (تسجيل)، دون (تدوين) - الجزء التحوي: فعل. - التعريف: إنتاج أو تسجيل نص أو معلومات. - مصطلح اعم: تواصل، توثيق. - مصطلحات أخص: - خط (خط / الكاليفرافيا). - ألف (تأليف). - الأشكال المرتبطة: - كتاب (كتاب، اسم). - كتابة (كتابة، اسم). - مكتوب (مكتوب صفار) - مفاهيم مرتبطة: القراءة والكتابة التأليف النص.	- الجذر: ك-ت-ب، مرتبط بالكتابة أو التسجيل. - الجزء النحوي: فعل. - المعنى: كتابات تسجيل، تأليف. - مرادفات: سجل (التسجيل)، دَوْن (الدوين)، مع اختلاف حسب السياق. - الأشكال المرتبطة: - اسم: كتاب (كتاب)، كتابة (كتابة). - أفعال: يكتب (مضارع)، مكتوب (مفعول). - مصطلح أعم: فعل التواصل أو التسجيل. - مصطلح أخص: خط (خط / كاليفرافيا)، ألف (تأليف). - مفاهيم مرتبطة: القراءة والكتابة التأليف التوثيق.

٧-٢-٣- تحليل قدرات النموذج DeepSeek

يقدم النموذج DeepSeek تحليلاً مميزاً يركز على العلاقات الدلالية المتقدمة، بما في ذلك المرادفات، والأضداد، والمصطلحات الأعم/الأخص. على سبيل المثال، لـ "كتب"، يحدد مرادفات مثل: "دون، صاغ، سطر" وأضداد مثل: "محا، حذف"، مع مصطلحات أعم (أنشأ) وأخص (خط، نسخ). بالنسبة لـ "ضرب"، يوفر أضداداً مثل: "لاطف" ومصطلحات أخص مثل: "ركل"، بينما يحدد لـ "علم" أضداداً مثل: "جهل، نسي" ومصطلحات أخص مثل: "حفظ". ومع ذلك، فإن النموذج يقتصر على معنى واحد رئيسي لكل جذر، مما يقلل من قدرته على التعامل مع تعدد المعاني مقارنة بالنماذج Gemini و ChatGPT. كما أن غياب الأمثلة السياقية يحد من فائدته في التطبيقات التعليمية.

شكل (٣) تحليل وبناء شبكة الكلمات للجذر "كتب" باستخدام النموذج DeepSeek



١. كَتَبَ (الفعل: كتب - "يُدَوِّن، يُسَجِّل")

- المرادفات: دَوَّنَ، صَاغَ سَطَّرَ، أَلَفَ.
- الأضداد: مَحَا، حَذَفَ، شَطَبَ.
- تعميم (كلمة أشمل): أَنْشَأَ (أنشأ نصاً)، أَعَدَّ (أعد وثيقة).
- تخصيص (كلمة أضيق): حَطَّ (باليد)، نَسَخَ (أعاد الكتابة)، طَبَعَ (بالآلة).
- مفاهيم مرتبطة:
- القِرَاءَةُ (الفعل المقابل)، الوَثِيقَةُ (النتيجة)، القَلَمُ (الأداة).



٧-٢-٤- تحليل قدرات النموذج Gemini

يقدم النموذج Gemini تحليلاً شاملاً ومنظماً يشبه النموذج ChatGPT، لكنه يتجاوز في شمولية العلاقات الدلالية. لـ "كتب"، يحدد ثلاثة معانٍ (الكتابة، التأليف، القضاء)، مع مرادفات، مصطلحات أعم/أخص، وعلاقات معجمية (مثل: كاتب، كتابة). بالنسبة لـ "ضرب"، يغطي أربعة معانٍ (الضرب الجسدي، ضرب المثل، السفر، الضرب الرياضي)، مع إضافة أضداد مثل: "تلطف" للمعنى الجسدي. لـ "علم"، يحدد ثلاثة معانٍ (التعلم، الوعي، التعليم)، مع أضداد مثل: "جهل" و"غفل". يتميز النموذج Gemini بتقديم رسم بياني مفصل للعلاقات، مع تغطية شاملة للعلاقات المعجمية والدلالية. هذا يجعله الأكثر شمولية في التعامل مع تعدد المعاني والعلاقات بين الكلمات. ومع ذلك، قد تكون تعقيداته الزائدة عائقاً في التطبيقات البسيطة.

شكل (٤) تحليل وبناء شبكة الكلمات للجذر "كتب" باستخدام النموذج Gemini

١. **كُتِبَ - To Write**
 - **مجموعات المعنى (Synsets):**
 - المعنى ١: القيام بفعل الكتابة.
 - الشرح: تشكيل الحروف، الكلمات، أو الرموز على سطح باستخدام قلم أو أداة أخرى.
 - مرادفات: دُون، سَطَّر، حَزَّر.
 - اسم أعم (Hyponyms): فعل (القيام بفعل)، (بمعنى أوسع: عبر، تواصل).
 - اسم أعم (Hyponyms):
 - نسخ.
 - ألف.
 - خط.
 - مصطلحات ذات صلة (علاقات معجمية):
 - كاتب (اسم فاعل).
 - كتابة (اسم مصدر).
 - مكتوب (اسم مفعول، وقد يأتي بمعنى رسالة).
 - المعنى ٢: تأليف أو إنشاء عمل أدبي أو كتاب.
 - الشرح: إبداع عمل مكتوب، غالبًا ما يكون ذا طول كبير أو مخصصًا للنشر.
 - مرادفات: ألف، أنشأ.
 - اسم أعم (Hypernym): أبداع، أنتج.
 - مصطلحات ذات صلة:
 - مؤلف.
 - كتاب.
٢. **علاقات تعدد المعاني (Polysemy):**
 - استخراج أسهم من عقدة "كتب" إلى عقدة "كتب - المعنى ١: الكتابة"، و"كتب - المعنى ٢: التأليف"، "نب - المعنى ٣: القضاء/الأمر".
 - وبالمثل، استخراج أسهم من "ضرب" إلى معانيها الأربعة، ومن "علم" إلى معانيها الثلاثة.
٣. **علاقات المرادفات (Synonymy):**
 - داخل كل مجموعة معنى (Synset)، ستكون هناك روابط بين الكلمة الرئيسية والمرادفات.
 - مثال: من "كتب - المعنى ١" إلى "دُون"، "سَطَّر"، "حَزَّر".
 - مثال: من "علم - المعنى ١" إلى "إلى" "تعلّم"، "استوعب"، "أدرك".
٤. **علاقات اسم الأعم (Hyponymy):**
 - سهم يتجه من المعنى الأصغر إلى المعنى الأعم.
 - مثال: من "كتب - المعنى ١: الكتابة" سهم إلى "فعل"، (أو "تواصل").
 - مثال: من "ضرب - المعنى ١: توجيه ضربة" سهم إلى "اعتدى".
 - مثال: من "علم - المعنى ١: التعلم" سهم إلى "اكتسب".
٥. **علاقات اسم الأعم (Hyponymy):**
 - سهم يتجه من المعنى الأعم إلى المعنى الأصغر (عكس Hypernymy).
 - مثال: من "كتب - المعنى ١: الكتابة" سهم إلى "نسخ"، "ألف"، "خط" (هذه في أنواع الكتابة).
٥. **علاقات الكلمات ذات الصلة (Lexical Relations):**
 - ستربط هذه العلاقات الكلمات المشتقة أو ذات الصلة بمعانيها الأصلية.
 - مثال: من "كتب - المعنى ١: سهم إلى "تنب" (اسم فاعل)، "كتابة" (اسم مصدر)، "مكتوب" (اسم مفعول).
 - مثال: من "ضرب - المعنى ١: سهم إلى "ضارب".
 - مثال: من "علم - المعنى ٢: سهم إلى "عالم"، "معلومة".
٦. **علاقات الضد (Antonymy):**
 - إذا كان هناك ضد مباشر لمعنى ما، فسيظهر سهم يربط بينهما.
 - مثال: من "علم - المعنى ١: التعلم" سهم إلى "جهل".
 - مثال: من "ضرب - المعنى ١: توجيه ضربة" سهم إلى "جهل".

٣-٧- مقارنة بين النماذج

- الشمولية في تغطية المعاني: النموذج ChatGPT و Gemini يتفوقان في تغطية المعاني المتعددة (٣-٤ معاني لكل جذر)، بينما يقتصر النموذج DeepSeek على معنى واحد رئيسي، والنموذج Grok يغطي معاني متعددة، ولكنه أقل تفصيلاً. النموذج Gemini يقدم أعلى مستوى من التفصيل بفضل شمولية العلاقات (مرادفات، أضداد، أعم/أخص).
 - العلاقات الدلالية والصرفية: النموذج Gemini و DeepSeek يتفوقان في تقديم الأضداد والمصطلحات الأعم/الأخص، بينما النموذج ChatGPT يركز فقط على المرادفات والأمثلة. النموذج Grok يقدم علاقات صرفية قوية، ولكنه يفتقر إلى الأضداد.
 - الأمثلة السياقية: النموذج ChatGPT يتفوق في تقديم أمثلة سياقية واضحة، مما يجعله مفيداً للتطبيقات التعليمية. النموذج Gemini يقدم بعض الأمثلة ضمناً، بينما النموذج Grok و DeepSeek يفتقران إلى هذه الأمثلة.
 - التصور البصري: النموذج Grok و DeepSeek يقدمان رسوماً بيانية، مما يجعلهما مناسبين للحوسبة اللغوية. النموذج Gemini يقدم تصوراً بيانياً أكثر تفصيلاً، بينما النموذج ChatGPT يعتمد على هيكلية هرمية نصية.
 - سهولة الاستخدام: النموذج ChatGPT هو الأكثر وضوحاً وبساطة، مما يجعله مناسباً للمبتدئين. النموذج Gemini، رغم شموليته، قد يكون معقداً بسبب كثرة العلاقات. النموذج Grok و DeepSeek يتطلبان معرفة تقنية لفهم الرسوم البيانية.
- يعد النموذج Gemini الأكثر شمولية ودقة في تحليل شبكة الكلمات العربية، حيث يغطي المعاني المتعددة، والعلاقات الدلالية (مرادفات، أضداد، أعم/أخص)، والعلاقات المعجمية بشكل مفصل. ومع ذلك، قد تكون تعقيداته عائقاً في التطبيقات البسيطة. النموذج ChatGPT مناسب للتطبيقات التعليمية بفضل أمثلته السياقية ووضوحه. النموذج Grok و DeepSeek مفيدان في الحوسبة اللغوية بسبب الرسوم البيانية، لكنهما أقل شمولية في تغطية المعاني.

٤-٧- البيانات المفقودة في النماذج الأربعة لتحليل شبكة الكلمات العربية وكيفية الحصول على شبكة كلمات متكاملة

تتناول هذه الدراسة العوامل المفقودة في النماذج الأربعة ChatGPT و Gemini و DeepSeek و Grok لبناء وتحليل شبكة الكلمات العربية (Arabic WordNet) مع اقتراح حلول لسد هذه الفجوات وبناء شبكة متكاملة. يقع التركيز على تحديد النقاط المفقودة في كل نموذج، واقتراح تحسينات، ومناقشة الخطوات اللازمة للحصول على شبكة مثالية بدمج نقاط القوة في كل نموذج.

٧-٤-١- البيانات المفقودة في كل نموذج

٧-٤-١-١- النموذج ChatGPT

- أ. الأضداد: النموذج ChatGPT لا يحدد الأضداد لأي من المعاني (مثل: "محا" أو "حذف" لـ "كتب"، أو "جهل" لـ "علم")، مما يحد من قدرته على تقديم شبكة دلالية كاملة.
- ب. العلاقات الأعم/الأخص (Hypernyms/Hyponyms): لا يوفر النموذج مصطلحات أعم (مثل: "تواصل" لـ "كتب") أو أخص (مثل: "خط" أو "نسخ") بصراحة ووضوح، مما يقلل من عمق التحليل الدلالي.
- ج. العلاقات المعجمية المتقدمة: لا يشمل النموذج علاقات مثل: الأسماء المشتقة (كاتب، ضربة) أو العلاقات بين المجالات الدلالية (مثل: ارتباط "كتب" و "علم" في سياق التعليم).

٧-٤-١-٢- الحلول المقترحة:

- إضافة قسم للأضداد لكل معنى، مثل: "جهل" لـ "علم" أو "تلطف" لـ "ضرب"، مع أمثلة توضيحية.
- تضمين مصطلحات أعم وأخص، مثل: "تواصل" كمصطلح أعم لـ "كتب" و "خط" كمصطلح أخص؛ لتعزيز الشبكة الدلالية.

- إدراج العلاقات المعجمية مثل الأسماء المشتقة (كاتب، كتابة) والعلاقات بين المجالات (مثل: ارتباط "كتب" و "علم" في التعليم).

٧-٤-١-٢- النموذج Grok

- أ. الأمثلة السياقية: لا يقدم النموذج أمثلة سياقية (مثل: "كتب رسالة إلى صديقه")، مما يجعل التحليل نظرياً ويصعب تطبيقه في التعليم.
- ب. الأضداد: لا يشمل النموذج الأضداد، مما يحد من تغطية العلاقات الدلالية.
- ج. تعدد المعاني المحدود: رغم تغطية معانٍ متعددة، إلا أن النموذج لا يحدد جميع المعاني الممكنة (مثل: "الضرب الرياضي" لـ "ضرب" أو "وضع علامة" لـ "علم").
- د. العلاقات بين المجالات الدلالية: لا يبرز النموذج العلاقات بين الجذور (مثل: ارتباط "كتب" و "علم" في التعليم).

٧-٤-١-٢- الحلول المقترحة:

- إضافة أمثلة سياقية لكل معنى لتوضيح الاستخدام العملي، مثل: "ضرب الباب بقوة" لـ "ضرب".
- تضمين الأضداد، مثل: "محا" لـ "كتب" و "جهل" لـ "علم"؛ لإكمال الشبكة الدلالية.
- توسيع تغطية المعاني المتعددة، مثل: إضافة "الضرب الرياضي" لـ "ضرب" و "وضع علامة" لـ "علم".
- إدراج علاقات بين المجالات الدلالية، مثل: ارتباط "كتب" و "علم" في سياق التعليم أو "ضرب" و "علم" في سياق الاستكشاف.

٧-٤-٣- النموذج DeepSeek

- أ. تعدد المعاني: يقتصر النموذج على معنى واحد رئيسي لكل جذر (مثل: "كتب" كتدوين فقط)، مما يجعله يتجاهل معانٍ أخرى مثل: "التأليف" أو "القضاء" لـ "كتب".
- ب. الأمثلة السياقية: لا يوفر النموذج أمثلة سياقية، مما يقلل من فائدته في التطبيقات التعليمية.

- ج. العلاقات المعجمية المشتقة: لا يغطي النموذج الأسماء المشتقة (مثل: كاتب، ضربة) بشكل كامل.
- د. التكامل مع معاني أخرى: لا يربط النموذج بين المعاني المتعددة للجذر الواحد أو بين الجذور المختلفة.

٧-٤-١-٣-١- الحلول المقترحة:

- توسيع التحليل ليشمل جميع المعاني المتعددة، مثل: "التأليف" و"القضاء" لـ "كتب"، مع مرادفات وأضداد لكل معنى.
- إضافة أمثلة سياقية، مثل: "علم الأطفال الحروف" لـ "علم"، لتوضيح الاستخدام.
- تضمين الأسماء المشتقة والعلاقات المعجمية، مثل: "كاتب" و"كتابة" لـ "كتب".
- إدراج علاقات بين المعاني المتعددة للجذر الواحد وبين الجذور المختلفة لتعزيز الشبكة الدلالية.

٧-٤-١-٤- النموذج Gemini

- أ. الأمثلة السياقية المحدودة: رغم شمولية النموذج، إلا أنه لا يقدم أمثلة سياقية واضحة لكل معنى، مما قد يحد من فائدته في التعليم.
- ب. التعقيد الزائد: يقدم النموذج تفاصيل كثيرة (مثل العلاقات المعجمية والدلالية)، مما قد يكون مربكاً للتطبيقات البسيطة.
- ج. العلاقات بين المجالات الدلالية: لا يبرز النموذج بوضوح العلاقات بين الجذور (مثل: "كتب" و"علم" في التعليم).

٧-٤-١-٤-١- الحلول المقترحة:

- إضافة أمثلة سياقية واضحة لكل معنى، مثل: "ضرب مثلاً في التضحية" لـ "ضرب".
- تبسيط العرض للتطبيقات البسيطة من خلال تقديم نسخة مختصرة من التحليل.
- تضمين علاقات بين المجالات الدلالية، مثل: ارتباط "كتب" و"علم" في التعليم.

للحصول على شبكة كلمات متكاملة، يجب دمج نقاط القوة في النماذج الأربعة: شمولية النموذج Gemini في تغطية المعاني والعلاقات، وضوح النموذج ChatGPT في الأمثلة السياقية، وقدرات النموذج Grok و DeepSeek في التصور البياني. يتطلب ذلك إضافة الأضداد، والمصطلحات الأعم/الأخص، والأمثلة السياقية، والعلاقات بين المجالات الدلالية، مع توفير تصور بصري وبياني. النموذج المقترح يوازن بين الشمولية والبساطة، مما يجعله مناسباً للتطبيقات الأكاديمية، والتعليمية، والحاسوبية. والمعالجة البشرية للحصول على شبكة معجمية متكاملة هي الأساس، كون النماذج اللغوية المقترحة دُرِبت في الأساس وبصورة كبيرة على اللغة الإنجليزية وبعض اللغات الأوربية، أما اللغات الأخرى وخصوصاً العربية فالبيانات غير كافية، والحل هو إثراء النماذج ببيانات عربية شاملة، والتدخل من خبراء العربية؛ حتى يضمن التحليل الشامل لكل ما يتعلق بالكلمة العربية لتغطية شبكة الكلمات.

٧-٥- المرحلة الثانية لعملية البناء والتحليل مع هندسة الأوامر

يُقدّم هذا القسم تحليلاً شاملاً ودراسةً كاملةً لما يتعلق بأساسيات بناء الشبكات العربية للكلمات بعد عجز النماذج اللغوية عن تغطية الأجزاء الأساسية للمفردة العربية في المرحلة الأولى، ركز الباحثان على المرجعيات اللغوية الأساسية التي تستخدم لبناء شبكة الكلمة العربية في هندسة الأوامر كمرحلة ثانية من البناء والتحليل. سيعرض هذا الجزء مدى تحسن أداء النماذج بعد هندسة الأوامر المقدمة لها من قبل الباحثين في بناء وتحليل الشبكات اللغوية العربية. في المثال الآتي حُلّ الجذر نفسه الذي حُلّ بدون هندسة الأوامر، وهو تحليل الجذر "كتب" في شبكة الكلمات العربية (Arabic WordNet)، مع التركيز على تغطية جميع المعاني الممكنة، والعلاقات الدلالية واللغوية، والأنطولوجيا، والتعريفات مع أمثلة سياقية. نُظِم التحليل وفقاً للعناصر المطلوبة: المجموعات المتشابهة (Synsets)، والأشكال الأساسية (Lemmas)، والعلاقات الدلالية واللغوية، والأنطولوجيا، والتعريفات مع الأمثلة.

المجموعات المتشابهة (Synsets)

تُغطي المجموعات المتشابهة المعاني المختلفة لكلمة "كتب" كفعل، مع تحديد المرادفات لكل معنى.

١. Synset الكتابة (الفعل)

- الجزء النحوي: فعل
- المرادفات: كتب، دون، سطر، حرر
- التعريف: تسجيل أو كتابة نص أو رموز على سطح باستخدام أداة مثل القلم.
- مثال: كتب رسالة إلى صديقه.

٢. Synset التأليف (الفعل)

- الجزء النحوي: فعل
- المرادفات: كتب، ألف، صاغ
- التعريف: إنشاء أو إنتاج عمل أدبي أو علمي مكتوب.
- مثال: كتب رواية جديدة نالت إعجاب النقاد.

٣. Synset القضاء/القدر (الفعل، مجازي)

- الجزء النحوي: فعل
- المرادفات: كتب، فرض، قدر
- التعريف: جعل شيء مفروضًا أو مقدرًا، غالبًا في سياق ديني أو شرعي.
- مثال: كتب الله عليه الجهاد.

الأشكال الأساسية (Lemmas)

تشمل الأشكال الأساسية المتغيرات الصرفية، والمشتقات، والتعبيرات متعددة الكلمات المرتبطة بـ "كتب"، مع تسوية التشكيل (بدون تشكيل أو تنوين).



دور النماذج اللغوية الكبيرة في تطوير (WordNet) قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين الشبكات الدلالية

● الأشكال الأساسية (الفعل):

- كتب (الماضي، الثالث المفرد المذكر)
- يكتب (المضارع)
- مكتوب (اسم المفعول)
- اكتب (الأمر)

● المشتقات:

- كتاب (اسم، بمعنى المصنف أو النص المكتوب)
- كتابة (اسم مصدر، بمعنى فعل الكتابة)
- كاتب (اسم فاعل، الشخص الذي يكتب)
- مكتب (اسم مكان، مكان الكتابة)
- مكتبة (اسم مكان، مكان تخزين الكتب)
- كتيب (اسم تصغير، كتاب صغير)
- كتابي (صفة، متعلق بالكتابة)

● التعابير متعددة الكلمات:

- كتاب مقدس
- كتاب مدرسي
- كتاب دراسي
- كتابة يدوية
- كتابة أدبية
- كتابة رسمية

العلاقات الدلالية (Semantic Relations)

علاقات شمولية (Hypernym / Hyponym)

● مصطلحات أعم (Hypernyms)

- معنى الكتابة: تواصل (الكتابة شكل من أشكال التواصل)
- معنى التأليف: إبداع (التأليف شكل من أشكال الإبداع)
- معنى القضاء: أمر (القضاء شكل من أشكال الأمر أو التقدير)

● مصطلحات أخص (Hyponyms)

- معنى الكتابة: خط (كتابة باليد)، نسخ (إعادة كتابة نص)، طبع (كتابة بالآلة)
- معنى التأليف: إنشاء (كتابة نص إبداعي قصير)، تصنيف (إنتاج كتاب علمي)
- معنى القضاء: لا توجد مصطلحات أخص شائعة.

علاقات الجزء-والكل (Part-Whole)

● الكتاب (كاسم):

- الأجزاء: صفحة، فصل، غلاف، فهرس.
- الكل: مكتبة (الكتاب جزء من مجموعة المكتبة)

● الكتابة (كعملية):

- الأجزاء: حرف، كلمة، جملة.
- الكل: نص، وثيقة.

العلاقات السياقية

● كلمات مرتبطة في السياق:

- قلم (أداة الكتابة)
- ورق (سطح الكتابة)
- حبر (مادة الكتابة)



دور النماذج اللغوية الكبيرة في تطوير (WordNet) قاعدة بيانات معجمية عربية: تقييم القدرات وتحسين الشبكات الدلالية

- قراءة (الفعل المقابل للكتابة)
- نص (نتيجة الكتابة)
- مؤلف (الشخص الذي يؤلف)
- دار نشر (المؤسسة المرتبطة بالتأليف)

العلاقات اللغوية (Lexical Relations)

الترادف (Synonymy)

- الكتابة: كتب، دون، سطر، حرر
- التأليف: كتب، ألف، صاغ
- القضاء: كتب، فرض، قدر

التضاد (Antonymy)

- لمعنى الكتابة: محا، حذف، شطب
- لمعنى التأليف: لا أضداد مباشرة (لأنه إبداعي)
- لمعنى القضاء: لا أضداد مباشرة (لأنه سياق ديني / مجازي)

الاشتقاق الصرفي (Morphological Derivation)

● من الجذر ك-ت-ب:

- كتاب (اسم)
- كتابة (اسم مصدر)
- كاتب (اسم فاعل)
- مكتوب (اسم مفعول)
- مكتب (اسم مكان)
- مكتبة (اسم مكان)
- كتيب (اسم تصغير)

الإحالات المتقاطعة (Cross-References)

● مفاهيم ذات صلة:

- علم (في سياق التعليم، حيث الكتابة أداة لنقل المعرفة)
- قراءة (الفعل المكمل للكتابة)
- نشر (مرتبط بالتأليف)
- وثيقة (نتيجة الكتابة)

الأنطولوجيا (Ontology)

الهيكل الهرمي:

- فعل ← تواصل ← كتابة ← كتب (المعنى ١)
- فعل ← إبداع ← تأليف ← كتب (المعنى ٢)
- فعل ← أمر ← قضاء ← كتب (المعنى ٣)
- اسم ← شيء مادي ← مصنف ← كتاب
- اسم ← مكان ← مكتبة ← مكتب

العلاقات الرأسية:

- مفهوم أعلى: تواصل (للكتابة)، إبداع (للتأليف)، أمر (للقضاء)، شيء مادي (للكتاب)
- مفهوم أدنى: خط، نسخ (للكتابة)، إنشاء، تصنيف (للتأليف)

التعريفات والأمثلة (Definitions & Usage Examples)

أ. الكتابة:

- التعريف: فعل تسجيل نص أو رموز على سطح باستخدام أداة.
- مثال: كتب الطالب ملاحظاته في الدفتر.



دور النماذج اللغوية الكبيرة في تطوير
(WordNet) قاعدة بيانات معجمية عربية:
تقييم القدرات وتحسين الشبكات الدلالية

ب. التأليف:

- التعريف: إنتاج عمل أدبي أو علمي مكتوب.
- مثال: كتب كتابًا عن تاريخ الأندلس.

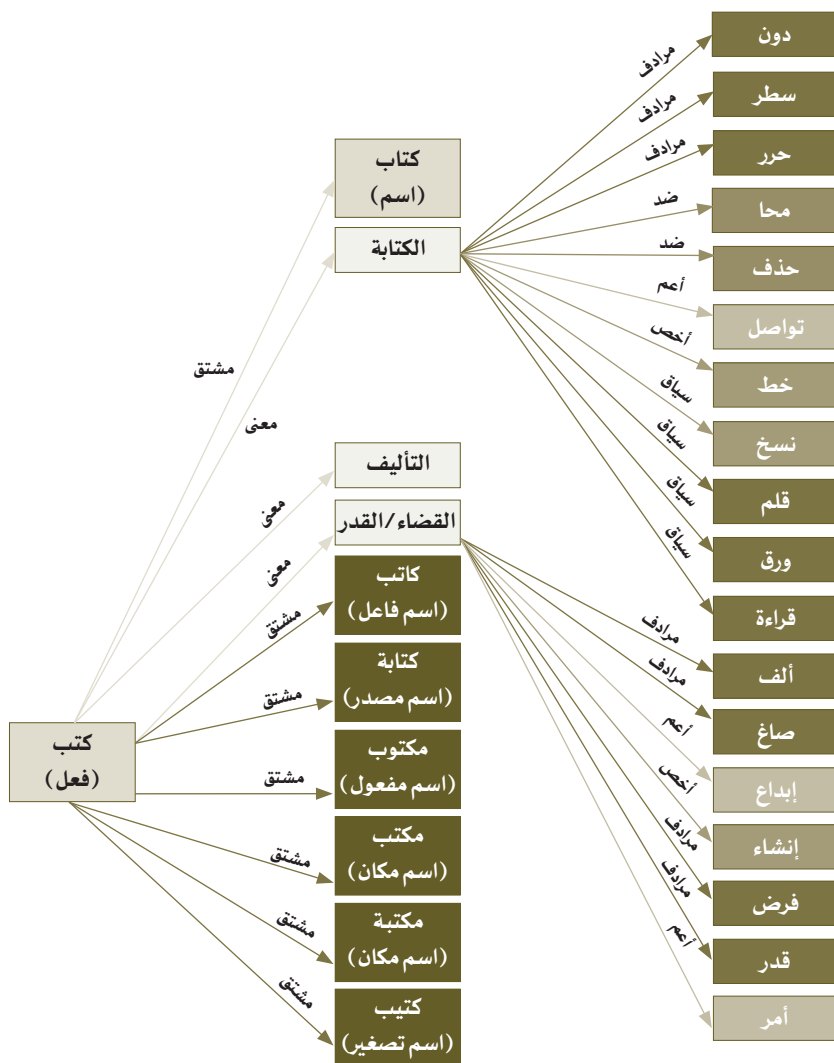
ج. القضاء/القدر:

- التعريف: جعل شيء مفروضًا أو مقدرًا في سياق ديني أو شرعي.
- مثال: كتب الله له الشفاء بعد مرض طويل.

د. الكتاب (اسم):

- التعريف: مصنف مكتوب يحتوي على نصوص أو معلومات.
- مثال: اشترى كتابًا جديدًا عن الفلسفة.

شكل (٥) الرسم البياني الشبكي اللغوي والمعجمي لكلمة "كتب" في المرحلة الثانية



- يُظهر الرسم البياني العلاقات بين "كتب" (كفعل واسم)، ومعانيها، والمشتقات.
- الألوان تميز العلاقات: الأزرق (مرادف)، الأحمر (ضد)، الأخضر (أعم)، الأرجواني (أخص)، البرتقالي (مشتق)، الرمادي (سياق).

يقدم التحليل الشبكي اللغوي والمعجمي لكلمة "كتب" دراسة شاملة تتناول جميع أبعاد الكلمة في شبكة الكلمات العربية، حيث يغطي المعاني المتعددة (الكتابة، التأليف، القضاء/القدر) والأشكال الأساسية (مثل: كتاب، كاتب، مكتبة)، مع التركيز على العلاقات الدلالية واللغوية. يتميز التحليل بتقديم مجموعات مترادفة (Synsets) واضحة لكل معنى، مثل: "كتب، دون، سطر" للكتابة، مع تعريفات دقيقة وأمثلة سياقية (مثل: "كتب رسالة إلى صديقه") تعزز فهم الاستخدامات. كما يشمل التحليل علاقات شمولية (مثل: "تواصل" كمصطلح أعم و"خط" كمصطلح أخص)، وعلاقات الجزء والكل (مثل: صفحة وكتاب)، والعلاقات اللغوية مثل: الترادف ("دون، سطر")، التضاد ("محا، حذف")، والاشتقاق الصرفي (كاتب، مكتوب). التصور البصري يوفر هيكلية مرئية تسهل تتبع العلاقات، مما يجعل التحليل أداة قيمة للحوسبة اللغوية والتطبيقات التعليمية. في الملحق الأول يبين التصور البصري والرسم البياني في المرحلة الأولى والثانية.

يبرز التحليل أيضًا من خلال الأنطولوجيا الهرمية التي توضح موقع "كتب" ضمن مسارات دلالية (مثل: فعل ← تواصل ← كتابة)، مما يعزز فهم العلاقات الرأسية بين المفاهيم. كما يربط الإحالات المتقاطعة "كتب" بمفاهيم مثل: "علم" و"قراءة"، مما يظهر التفاعل بين المجالات الدلالية مثل التعليم. ومع ذلك، يمكن تحسين المخرجات بإضافة المزيد من الأمثلة السياقية لتغطية تنوع الاستخدامات الثقافية أو الإقليمية، وتوسيع العلاقات بين المجالات الدلالية لتشمل مفاهيم إضافية مثل: "نشر" أو "توثيق". عامةً، يوفر التحليل إطارًا قويًا لفهم "كتب" في سياقاتها اللغوية والدلالية، مع إمكانية التوسع لتحليل جذور أخرى أو تطبيقات متقدمة في معالجة اللغة الطبيعية.

٨- التوصيات

- دمج نقاط القوة في النماذج الأربعة: يُوصى بدمج شمولية النموذج Gemini، ووضوح الأمثلة السياقية في النموذج ChatGPT، وقدرات التصور البرمجي في النموذجين Grok وDeepSeek لإنشاء إطار تحليلي موحد يغطي غالبية الجوانب الصرفية والدلالية والسياقية للمفردات العربية.

- **هندسة الأوامر:** يوصى بالتركيز على هندسة الأوامر المدخلة للنماذج اللغوية وصياغتها بفعالية من خبراء لغويين ذوي خبرة في العلاقات الدلالية والصرفية والنحوية.
- **توسيع قاعدة البيانات اللغوية:** ينبغي توسيع قاعدة بيانات Arabic WordNet لتشمل المزيد من الجذور اللغوية والمعاني الإقليمية، مع الاستفادة من مصادر نصوص متنوعة مثل: الكتب الدينية والأدبية وقواعد بيانات لتعزيز السياقات.
- **تحسين أداء النماذج اللغوية:** يُوصى بتدريب النماذج اللغوية الكبيرة (خاصة Grok) على بيانات عربية موسعة لتحسين دقة تجميع المرادفات، وتصنيف الأضداد، والعلاقات الجزئية، مع التركيز على معالجة الفروق الدلالية الدقيقة.
- **تطوير أدوات تصور متقدمة:** يُقترح تطوير أدوات تصوّر تفاعلية باستخدام بعض التقنيات الذكية الحديثة لدعم التطبيقات الحاسوبية والتعليمية، مما يسهل فهم العلاقات المعقدة في الشبكة المعجمية.
- **الشراكة اللغوية بين الخبراء وأدوات الذكاء الاصطناعي:** يوصى بتطوير برامج تدريب مشتركة تجمع بين الخبراء اللغويين ومتخصصي الذكاء الاصطناعي؛ لتحسين أداء النماذج اللغوية الكبيرة في تحليل المفردات العربية، من خلال إنشاء آليات مراجعة بشرية دورية لتصحيح الأخطاء في تجميع المرادفات، وتصنيف الأضداد، والعلاقات الدلالية، مما يعزز دقة الشبكة المعجمية العربية ويضمن تلبية متطلبات السياقات الثقافية واللغوية المتنوعة.
- **تعزيز التطبيقات العملية:** يُوصى بتطبيق الإطار الدلالي في مجالات مثل: التعليم اللغوي، ومعالجة اللغة الطبيعية، وتطوير أنظمة استرجاع المعلومات، مع إجراء دراسات مستقبلية لتقييم أثر هذه الشبكات على تحسين فهم اللغة العربية في السياقات الثقافية والاجتماعية.

٩- الخاتمة

يُظهر البناء الشبكي اللغوي والمعجمي للمفردات العربية ضمن إطار شبكة الكلمات العربية (Arabic WordNet) باستخدام النماذج اللغوية الكبيرة (ChatGPT، Gemini، DeepSeek، Grok) قدرة هذه النماذج على تقديم تحليلات تتناول المعاني المتعددة، والعلاقات الدلالية (مثل:

المرادفات، الأضداد، المصطلحات الأعم/الأخص)، والعلاقات اللغوية (مثل: المشتقات الصرفية). اتضح أن النموذج Gemini يتفوق في الشمولية بتغطيته الواسعة للعلاقات الدلالية واللغوية، بينما يبرز النموذج ChatGPT بوضوح الأمثلة السياقية التي تدعم التطبيقات التعليمية. النموذجان Grok وDeepSeek يقدمان دعمًا قويًا للحوسبة اللغوية، لكنهما يعانيان من نقص في الأمثلة السياقية والتغطية الكاملة للمعاني المتعددة. تكشف الدراسة عن قصور في النماذج مثل: التجميع غير الدقيق للمرادفات، والتصنيف الخاطئ للأضداد، والثغرات في العلاقات الجزئية، لكنها تؤكد تفوق Grok في التعامل مع تعقيدات الصرف العربي والفروق الدلالية، مما يجعله أداة واعدة لتطوير قاعدة بيانات معجمية عربية موثوقة. يمثل هذا العمل خطوة نوعية نحو بناء شبكة معجمية عربية مدعومة بالذكاء الاصطناعي، حيث يوفر إطارًا دلاليًا متينًا يعزز الدقة الدلالية والنطاق المعجمي للمفردات العربية، مع تقليل الوقت والجهد المطلوبين لتطوير الموارد اللغوية. أظهرت الدراسة أن الهندسة البشرية للأوامر المقدمة للنماذج طورت بشكل فعال أداء النماذج في بناء شبكة الكلمات العربية (Arabic WordNet)، لاسيما في معالجة تعقيدات الصرف العربي والفروق الدلالية الدقيقة. ولتعزيز الكفاءة، وتقليل الوقت والجهد، ورفع الإنتاجية اللغوية بدقة عالية، يُعد دمج الخبرة البشرية مع أدوات الذكاء الاصطناعي ضرورة حتمية. يتيح هذا التكامل الاستفادة المثلى من قدرات النماذج اللغوية في معالجة البيانات آليًا، مما يوفر موارد معجمية موثوقة تدعم التطبيقات التقنية المتنوعة، مثل معالجة اللغة الطبيعية، والتعليم اللغوي، واسترجاع المعلومات، مع الحفاظ على الدقة العالية والمرونة في تلبية متطلبات السياقات الثقافية واللغوية العربية.

المراجع العربية

- الحلي، سندس محمد. المحاييري، محمد مازن. والعباس، فيصل. (٢٠٢٣). إيجاد أنطولوجيا عربية بالتقابل بين شبكة الكلمات العربية والموسوعة العربية. مجلة جامعة دمشق للعلوم الهندسية، ٣٩ (٣)، ١٧٧-١٩٣.
- أدغيري، أمينة. (٢٠٢٢). منهجية مقترحة لبناء وتطوير أنطولوجيا عربية مجال اللسانيات الحاسوبية أمودجا. مجلة حوسبة اللغة العربية، ٢ (١)، ١٨-٣٩.
- طه، المعتز بالله السعيد. (٢٠١٢). شبكة الكلمات العربية المنهجية والتطبيق. مجلة مجمع اللغة العربية بدمشق، ٨٧ (٢)، ٤٢٣-٤٥٠.
- دلول، إيمان صبحي. (٢٠٢٠). أنطولوجيا الدلالة المعجمية للغة العربية: نحو مقارنة منهجية لسانية. مجلة الميادين للدراسات في العلوم الإنسانية، ٢ (٤)، ٨-٣٤.

رومنة المراجع العربية

Bibliographic information in Roman script

- Al-Halabi, S. M., Al-Muhayri, M. M., & Al-Abbas, F. (2023). Iejad ontology Arabia bil-taqabul bayna Shabakat al-Kalimat al-Arabia wa-al-Mawsu'a al-Arabia [Finding an Arabic ontology by matching the Arabic WordNet and the Arabic Encyclopedia]. Damascus University Journal for Engineering Sciences, 39(3), 177-193.
- Adghiri, A. (2022). Manhajia muqtaraha li-bina' wa-tatwir ontology Arabia: Majal al-lisaniyat al-hasubiya unmuzhajan [A proposed methodology for building and developing an Arabic ontology: Computational linguistics as a model]. Journal of Arabic Language Computing, 2(1), 18-39.
- Taha, M. S. (2012). Shabakat al-kalimat al-Arabia: Al-manhajia wa-al-tatbiq [Arabic WordNet: Methodology and application]. Journal of the Arabic Language Academy in Damascus, 87(2), 423-450.



Dalloul, I. S. (2020). Ontology al-dalala al-mu'jamia lil-lugha al-Arabia: Nahwa muqaraba manhajia lisania [Ontology of Arabic lexical semantics: Towards a linguistic methodological approach]. *Al-Mayadeen Journal for Humanistic Studies*, 2(4), 8-34.

المراجع الأجنبية

Al-Barhamtoshy, H., & Al-Jideebi, W. (2009, December). Designing and implementing Arabic WordNet semantic-based. In *the 9th Conference on Language Engineering* (pp. 23-24).

AlMaayah, M., Sawalha, M., & Abushariah, M. A. (2016). Towards an automatic extraction of synonyms for Quranic Arabic WordNet. *International Journal of Speech Technology*, 19(2), 177-189.

Alsudias, L., & Rayson, P. (2020, May). Developing an Arabic infectious disease ontology to include non-standard terminology. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 4842-4850).

Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journals*, Special Issue on CALL,(9), 3-17.

Alkhalifa, M., & Rodríguez, H. (2010). Automatically extending named entities coverage of Arabic WordNet using Wikipedia. *International Journal on Information and Communication Technologies*, 3(3), 20-36.

Al-Garaady, J., & Mahyoob, M. (2025). Understanding User Perceptions of DeepSeek: A Mixed-Methods Sentiment and Thematic Analysis. *Available at SSRN* 5172367.

Albuhairy, M., Al-Garaady, J., & Alblwi, A. (2023). A proposed framework for human-like language processing of ChatGPT in academic writing. *International Journal of Emerging Technologies in Learning (iJET)*, 18, (14).

Al Zamil, M. G., & Al-Radaideh, Q. (2014). Automatic extraction of ontological relations from Arabic text. *Journal of King Saud University-Computer and Information Sciences*, 26(4), 462-472.

- Albuhairy, M. M., & Algaraady, J. (2025). DeepSeek vs. ChatGPT: Comparative efficacy in reasoning for adults' second language acquisition analysis. *Humanities and Educational Sciences Journal*, 44, 864–883.
- Al-Garaady, J. (2022). An analysis of Yemenis' responses and sentiments on social media towards the emergence of the COVID-19 pandemic. *Humanities and Educational Sciences Journal*, 27, 589-607.
- Aliane H., Alimazighi Z. Ahmed Cherif M. "Al –Khalil: 2009. The Arabic Linguistic Ontology Project", Semantic web and Arabic Language Team, in Computer Science Department, University, Algiers.
- Algaraady, J., & Mahyoob, M. (2025). Exploring ChatGPT's potential for augmenting post-editing in machine translation across multiple domains: challenges and opportunities. *Frontiers in Artificial Intelligence*, 8, 1526293.
- Belkredim, F. Z., & Meziane, F. (2008). DEAR-ONTO: a derivational Arabic ontology based on verbs. *International Journal of Computer Processing of Languages*, 21(03), 279-291.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.
- Farghaly, A., & Shaalan, K. (2009). Arabic natural language processing: Challenges and solutions. *ACM Transactions on Asian Language Information Processing (TALIP)*, 8(4), 1-22.
- Habash, N. Y. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool Publishers.
- Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Georgetown University Press.
- Ishkewy, H., Harb, H., & Farahat, H. (2014). Azhary: An Arabic Lexical Ontology. *International Journal of Web & Semantic Technology*, 5(4), 71.
- Jarrar, M. (2021). The Arabic ontology—an Arabic wordnet with ontologically clean content. *Applied ontology*, 16(1), 1-26.



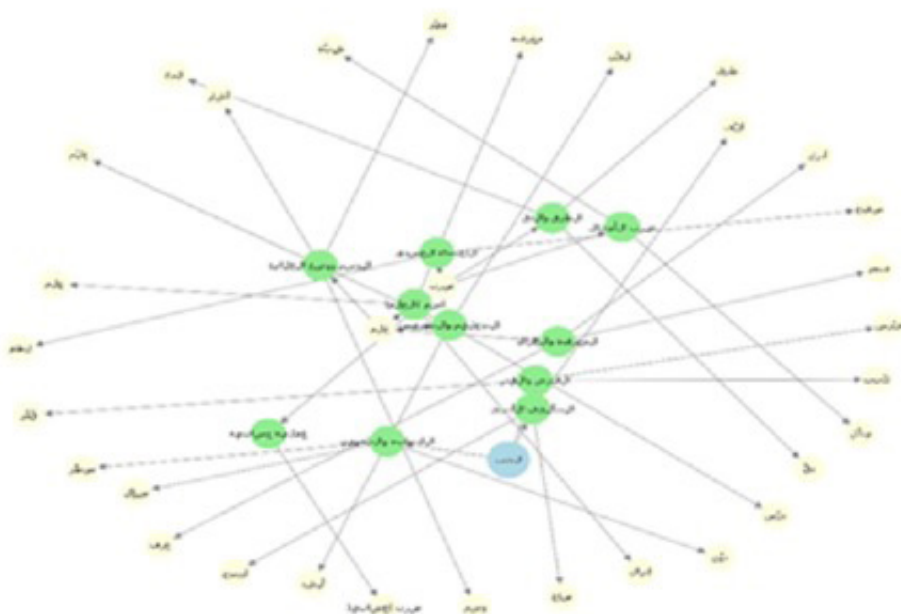
- Mazari, A. C., Aliane, H., & Alimazighi, Z. (2012). Automatic Construction of Ontology from Arabic Texts. In *ICWIT* (pp. 193-202).
- Mahyoob, M., & Algaraady, J. (2018). Towards developing a morphological analyser for Arabic noun forms. *International journal of Linguistics and computational Applications*, 5(3), 45-51.
- Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial intelligence*, 193, 217-250.
- Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A primer in BERTology: What we know about how BERT works. *Transactions of the association for computational linguistics*, 8, 842-866.
- Rodríguez, H., Farwell, D., Ferreres, J., Bertran, M., Alkhalifa, M., & Martí, M. A. (2008). Arabic WordNet: Semi-automatic Extensions using Bayesian Inference. In *LREC*.
- Saif, A., Ab Aziz, M. J., & Omar, N. (2017). Mapping Arabic WordNet synsets to Wikipedia articles using monolingual and bilingual features. *Natural Language Engineering*, 23(1), 53-91.

نبذة تعريفية عن الباحث أو الباحثين:

Author Biodata Dr. Mohammad Mahyoob Albuhairey is an Associate Professor of Computational linguistics at the Department of Languages & Translation, Taibah University, Madina, KSA. He holds his PhD in Computational Linguistics. His research interests include Computational Linguistics, AI, Disabilities, Digital Media Studies, Language Acquisition, Corpus Linguistics and e-learning.	بيانات الباحث: محمد مهيووب البحيري، أستاذ مشارك في اللغويات الحاسوبية بقسم اللغات والترجمة في جامعة طيبة بالمدينة المنورة، بالملكة العربية السعودية. حصل على درجة الدكتوراه في اللغويات الحاسوبية. من جامعة عليجار الإسلامية في الهند، عام ٢٠١٥. تدور اهتماماته البحثية حول اللغويات الحاسوبية، والذكاء الاصطناعي، والإعاقة، ودراسات الإعلام الرقمي، واكتساب اللغة، والمدونات اللغوية، والتعليم الإلكتروني.
معرف أوركيد (ORCID) : 0000-0002-6664-1017	
Email: eflu2010@gmail.com	

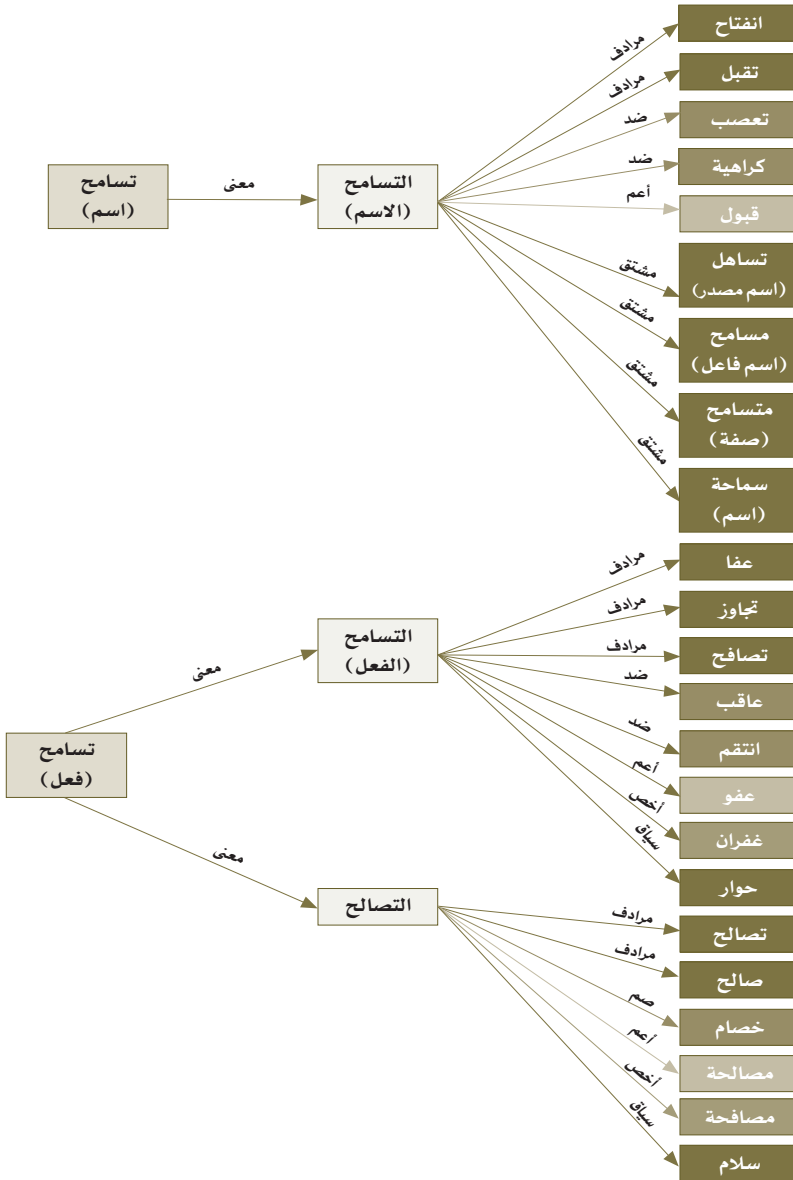
Author Biodata Dr. Jeehaan Algaraady is an Assistant Professor of Computational Linguistics at the College of Education, Taiz University, Yemen. She received her PhD in Computational Linguistics. Her research interests include Computational Linguistics, Data Mining, AI, Language Disorders, Theoretical Linguistics, and e-learning.	بيانات الباحثة: جيهان الجرادي، أستاذ مساعد في اللغويات الحاسوبية بقسم اللغة الإنجليزية في جامعة تعز، باليمن. حاصلة على درجة الدكتوراه في اللغويات الحاسوبية من جامعة عليجار الإسلامية في الهند عام ٢٠١٨. تدور اهتماماتها البحثية حول اللغويات الحاسوبية، والذكاء الاصطناعي، والاضطراب اللغوي، والتقيب في البيانات، والتعليم الإلكتروني.
معرف أوركيد (ORCID) : 0000-0003-3901-4648	
Email: Jihan.amu@gmail.com	

الرسم البياني الشبكي للكلمات "كتب ورسم وعلم" قبل هندسة الأوامر



ملحق (ب)

الرسم البياني الشبكي لكلمة "تسامح" بعد هندسة الأوامر





مجمع الملك سلمان
العالي للغة العربية

King Salman Global Academy for Arabic Language



الأكاديمية العالمية
للغة العربية

Journal of Computational Linguistics & Arabic Language Processing

