دولـة الإمارات العـربية المتحدة

جامعة الوصل - دبي
كلية الآداب

فِـكـر ومَـعـرفة

مجلة علمية محكمة سنوية
متخصّصة في العلوم الإنسانيّة والاجتماعيّة

العـدد الرابع
(1446 هـ - 2024 م)

تأسست سنة 2021 م

المشرف على المجلة

أ. د. خالد توكال

نائب مدير الجامعة لشؤون البحث العلميّ

رئيس التَّحرير

د. عبد اللّه طاهر الحذيفي

نائب رئيس التحرير

أ. د. علاء عبد الستار مغاوري

هيئة التحرير

أ. د. على الشبعان – أ. د. سيد عبد الخالق إسماعيل

كلية الآداب

الرؤية والرسالة والأهداف

الرؤية:

تعليم إنساني ابتكاري لمجتمع عالمي.

الرسالة:

تأهيل مخرجات نوعية في برامج البكالوريوس والدراسات العليا، تلبية لاحتياجات سوق العمل المستقبلية في المجتمع الإماراتي والإقليمي والعالمي.

الأهداف:

انطلاقًا من رؤية كلية الآداب ورسالتها فإنها تهدف إلى:

أولًا: إعداد جيل يتمسك بالقيم العربية الإسلامية والمبادئ الإنسانية السامية.

ثانيًا: تقديم مخرجات مؤهلة لخدمة اللغة العربية بحثًا وتدريسًا والسير بها نحو العالمية.

ثالثًا: ترسيخ مبدأ التعايش بين اللغات والثقافات والحضارات.

رابعًا: النهوض بالأدب العربي والانفتاح على الآداب العالمية.

خامسًا: تعزيز وحدة التعليم العام، وتوفير جميع الوسائط المتاحة لتنمية الأداء في اللغة الإنجليزية والحاسوب والبرمجة الآلية للغات.

سادسًا: تأهيل متخرجين أكفاء في كافة تخصصات الكلية.

سابعًا: تشجيع البحث العلمي المتميز في كافة تخصصات الكلية.

كلية الآداب

النشأة والتطور

أنشئت كلية الآداب بناءً على القرار الوزاري رقم: (١٠٧) الصادر من مكتب وكيل الوزارة للشؤون الأكاديمية للتعليم العالي، وزارة التربية والتعليم بتاريخ: ٨ إبريل ٢٠١٩ في شأن الترخيص لجامعة الوصل (Alwasl University) لتصبح جامعة من جامعات التعليم العالي مقرها(دبي) بدولة الإمارات العربية المتحدة.

كانت كلية الآداب قبل ٢٠١٩ جزءا من كلية الدراسات الإسلامية والعربية التي أنشئت سنة ١٩٨٦، وبدأت يومئذ بمرحلة البكالوريوس، ثم أنشئت بها مرحلة الماجستير بشعبتين: اللغة والنحو والأدب والنقد ابتداء من سنة ٢٠٠٢-٢٠٠٣، ثم اكتملت مراحلها الثلاث في سنة ٢٠٠٧-٢٠٠٨ بإنشاء مرحلة الدكتوراه بشعبتيها: اللغة والنحو والأدب والنقد.

يتكون مجلس كلية الآداب من عميد الكلية ورؤساء البرامج الأكاديمية، ويضطلع بمهمة متابعة العملية التعليمية والسير بها نحو الأفضل، والسهر على تحديث البرامج وتهيئة جميع الظروف المواتية لتحسين المخرجات.

أولًا: البرامج الأكاديـمية:

البرامج المعتمدة حاليًّا:

- برنامج البكالوريوس في اللغة العربية وآدابها.
- برنامج البكالوريوس في علوم المكتبات والمعلومات.
- برنامج ماجستير الدراسات اللغوية.
- برنامج ماجستير الدراسات الأدبية والنقدية.
- برنامج دكتوراه الفلسفة في الدراسات اللسانية.
- برنامج دكتوراه الفلسفة في الدراسات الأدبية والنقدية.

مجلة فكر ومعرفة

الرؤية والرسالة والأهداف

الرؤية المجلة:

الريادة في نشر بحث علميّ إنسانيّ ابتكاريّ إبداعيّ.

الرسالة المجلة:

تطوير بحث علمي إنساني مبدع متجذر في أرضية عبقرية الشعب الإماراتي الخاصة، يتميز بالرصانة والموضوعية، متناغم مع حركة الإبداع العلمي العربية والعالمية، يتأثر بها بوعي نقدي متبصر، ويؤثر فيها بعطاء نوعي ذي بصمة متميزة، يخدم حاجات الإنسان وسوق العمل المستقبلية في المجتمع الإماراتي والإقليمي والعالمي.

الأهداف المجلة:

أولًا: تطوير بحث علمي مبدع، يتمسك بالقيم الإسلامية والعربية والمبادئ الإنسانية السامية.

ثانيًا: تقديم بحوث علمية تخدم العلوم الاجتماعية والإنسانية: تطورها وتنشرها وتسير بها نحو العالمية.

ثالثًا: نشر البحوث العلمية المتميزة التي تتعلق بأهم القضايا والمتغيرات المجتمعية وتحليلها واقتراح أفضل الحلول والممارسات.

رابعًا: تأهيل الباحثين الوطنيين المبدعين الأكفاء في كافة تخصصات العلوم الاجتماعية والإنسانية.

خامسًا: تطوير أدوات البحث العلمي المتميز وتعزيز قدرات الباحثين على التنافس في سياق البحث العلمي الجاد.

سادسًا: متابعة الإنتاج العلميّ المتميز الجديد في ميادين العلوم الاجتماعيّة والإنسانيّة.

قواعد النشر

أولاً:

تنشر المجلّة البحوث العلمية باللغات العربية، والإنجليزية والفرنسية؛ تحريرًا أو ترجمةً، على أن تكون بحوثاً أصيلة مبتكرة تتصف بالموضوعية والشمول والعمق، ولا تتعارض مع القيم الإسلامية، وذلك بعد عرضها على محكمين من خارج هيئة التحرير بحسب الأصول العلمية المتبعة.

ثانيًا:

١- يراعى في البحث أن يتميز بالأصالة وأن يضيف إضافة جديدة للعلم والمعرفة، وأن يكون مستوفيًا للجوانب العلمية بما في ذلك عرض الأسس النظرية والأهداف الخاصة للبحث والإجراءات المستخدمة في التحليل وعرض النتائج والمناقشة.

٢- تخضع جميع البحوث المقدمة للنشر في المجلة للشروط الآتية:

٣- ألا يكون البحث قد نشر من قبل، أو قدم للنشر إلى جهة أخرى، وألا يكون مستلًّا من بحث أو من رسالة أكاديمية نال بها الباحث درجة علمية، وعلى الباحث أن يقدم تعهدًا خطيًّا بذلك عند إرساله إلى المجلة.

٤- تقبل البحوث التي تكون جزءا من رسالة جامعية لم تناقش بعد.

٥- لا يجوز للباحث أن ينشر بحثه بعد قبوله في المجلة في مكان آخر إلا بإذن خطي من رئيس التحرير، وإلا تكفل الباحث بسداد التكلفة المالية لتحكيم بحثه خلال الدورة التحكيمية.

٦- يراعى ضبط الآيات القرآنية وكتابتها بالرسم العثماني، وتخريج الأحاديث النبوية الشريفة، إن استشهد بها في البحوث.

٧- يُكتب البحث بمسافات (مفردة)، على ألا يقل عدد صفحاته عن (٢٠) صفحة بواقع (٥٠٠٠) خمسة آلاف كلمة، ولا يزيد عن (٣٠) صفحة في (٧٥٠٠) سبعة آلاف وخمسمائة كلمة، وحجم الخط (١٦) نوع (Simplified Arabic)، وإذا زاد البحث عن

(٣٠) صفحة، فعلى الباحث دفع تكاليف الطباعة للصفحات الزائدة؛ وهي (٥) دولارات عن كل صفحة.

٨- ترسل من البحث نسخة إلكترونية، وفق برنامج "Word ٢٠١٠" وتكتب أسماء الباحثين باللغتين العربية والإنجليزية، كما تذكر عناوينهم ووظائفهم الحالية ورتبهم العلمية، بحسب كشف البيانات المرفق؛ وذلك (بغرض التوثيق الدولي).

٩- يُرفق مع البحث ملخص باللغة العربية (في حدود ١٢٠ كلمة) وآخر باللغة الإنجليزية (في حدود ١٥٠ كلمة)، ويتضمن على الأقل أهداف البحث وإشكاليته، ومنهجهُ وأهم نتائجه، وإسهامات البحث، وخمسة كلمات مفتاحية.

١٠- يُرفق بالبحث الترجمة الكاملة لقائمة المصادر والمراجع باللغة الإنجليزية؛ وذلك بغرض التوثيق الدولي.

١١- ترقم الجداول والأشكال والصور التوضيحية وغيرها على التوالي بحسب ورودها في متن البحث، وتزود بعنوانات يشار إلى كل منها بالتسلسل نفسه، وتقدم بأوراق منفصلة.

١٢- يتبع المنهجية العلمية في توثيق البحوث على النحو الآتي:

- يشار إلى المصادر والمراجع في متن البحث بأرقام متسلسلة آليًّا توضع بين قوسين إلى الأعلى (هكذا:[١] [٢]) وتبين بالتفصيل في أسفل الصفحة وفق تسلسلها في المتن.

- تذكر ببليوغرافيا (معلومات الكتاب) في أول ورود له في البحث على النحو الآتي: اسم المؤلف، عنوان الكتاب، اسم المحقق (إن وجد) أو المترجم، دار النشر، بلد دار النشر، رقم الطبعة يشار إليها بـ (ط) إن وجدت، التاريخ إن وجد وإلَّا يشار إليه بـ (د.ت). أما بحوث الدوريات فتكون المعلومات على النحو الآتي: (اسم المؤلف، عنوان البحث، اسم المجلة، جهة الإصدار، بلد الإصدار، رقم العدد، التاريخ، مكان البحث في المجلة ممثلاً بالصفحات (من...إلى...).

- إذا تكرر المصدر بعد أول إيراد له يُكْتفَى باسم المؤلف وعنوان المصدر، فإن تكرر

مباشرة في الصفحة نفسها يكتب: (المرجع نفسه)، فإن تكرر مباشرة في الصفحة اللاحقة يكتب: (المرجع السابق).

- يشار إلى الشروح والملحوظات في متن البحث بنجمة (هكذا:*) أو أكثر.
- تثبت المصادر والمراجع في قائمة آخر البحث مرتبة ترتيبًا هجائيًّا بحسب اسم المؤلف يليه الكتاب والمعلومات الأخرى.

١٣- يلتزم الباحث بإجراء التعديلات التي يطلبها المحكمون على بحثه وفق التقارير المرسلة إليه، وموافاة المجلة بنسخة معدلة من البحث، وتقرير عن التعديلات التي قام بها.

١٤- يحرص الباحث على تدقيق بحثه لغويًّا، ولا تقبل المجلة بحوثًا غير مدققة لغويًّا.

ثالثًا: الشروط الإضافية على البحوث المترجمة:

١- أن ترفق مع الترجمة المادة المترجمة بلغتها الأصلية.

٢- يرفق مع الترجمة ملخصان أحدهما بالعربية والآخر بالإنجليزية أو الفرنسية، على ألا يتجاوز كل ملخص (١٢٠) كلمة، مع الكلمات المفتاحية.

٣- تكون المادة المترجمة محكمة، أو منشورة في إحدى المجلات المحكمة، أو تكون جزءًا من كتاب محكم.

٤- لا يتجاوز عدد صفحاتها / ٢٠ صفحة / من الحجم العادي (A4) (٦٠٠٠ كلمة) ولا يقل عن / ٧ صفحات / .

٥- المحافظة على النص الأصيل وتفادي الاختزال ما لم يُشَرْ إلى ذلك وبهدف تحسين الترجمة.

٦- أن تكون الجمل مترابطة ومتماسكة وتخدم المعنى المقصود في المادة الأصلية.

٧- يذكر في أول إحالة في الترجمة اسم المؤلف الأصلي مع نبذة عن إسهاماته.

٨- تشتمل الترجمة على مقدمة في سطور تبين الأهمية العلمية للمادة المترجمة، وأهم النتائج المتوقعة.

رابعًا:

١- ما ينشر في المجلة من آراء يعبر عن فكر أصحابها، ولا يمثل رأي المجلة بالضرورة.

٢- البحوث المرسلة إلى المجلة لا تعاد إلى أصحابها سواء نشرت أم لم تنشر.

٣- يخضع نشر البحوث وترتيبها لاعتبارات فنية، بحسب خطة النشر.

٤- يحق للمجلة - عند الضرورة - إجراء بعض التعديلات الشكلية على البحوث المقبولة للنشر دون المساس بمضمونها.

٥- يحق للمجلة نشر البحوث المقبولة إلكترونيًّا، والمشاركة بها في قواعد البيانات والمواقع الإلكترونية.

٦- يزود الباحث بعد نشر بحثه بنسخة إلكترونية (PDF) من العدد الذي نشر فيه بحثه، ومستلة (PDF) لبحثه.

خامسًا: رسوم النشر:

إسهامًا من مجلة فكر ومعرفة في إثراء الحركة البحثية في دولة الإمارات العربية المتحدة بشكل خاص، وكل الأقطار العربية والإسلامية بشكل عام، فإنَّ المجلة لا تحمل الباحثين أية رسوم، إلا ما سبق الإشارة إليه في بند (٧) ثانيًا، وإذا أراد الباحث سحب بحثه بعد حصوله على خطاب القبول فإن عليه دفع تكاليف تحكيم البحث.

ترسل البحوث وجميع المراسلات المتعلقة بالمجلة إلى:

رئيس تحرير مجلّة فكر ومعرفة
ص.ب. ٥٠١٠٦ دبـي - دولة الإمارات العربية المتحدة
هاتف: ٠٠٩٧١٤٣٩٦١٧٧٧
فاكس: ٠٠٩٧١٤٣٩٦١٣١٤
E-mail: fom@alwasl.ac.ae

محتويات العدد

افتتاحية التحرير	١١
البحوث	١٣
الاقتراض اللّغوي ومظاهر التنافس بين العربية ولغات أخرى أ.د. محمّد الصُّحبي البَعَزاوي	١٥
إشكالية المفاهيم الإجرائية في النقد الثقافي: وعي النظرية بين المعجم والاستعمال النقدي د. فتحي أحمد صالح الشرماني – د. أسماء مقبل عوض الأحمدي	٥٣
البورتريه الشِّعري في الشعر الإماراتي شِعر أمل السَّهلاوي أنموذجًا أسماء يوسف الحمادي	٩٣
المصطلح النقدي وشعر النساء: «فحولة الشعراء» للأصمعي أنموذجًا د. سليم قسطي	١٦٩
مساهمة بعض الدول العربية في صناعة معاجم مختصة بالذكاء الاصطناعي، «مقاربة وصفية مقارنة السعودية والإمارات أنموذجًا» أ. د. نوال بومَشْطَـه	٢٠٧
توظيفُ الذكاء الاصطناعيّ في صناعة المعجم العربيّ أ. د. سمر جورج الديّوب	٢٣٧
نحْو أجْرُوميّةٍ للتّأثِيل السَّامي في معجم الشارقة التاريخي د. إبراهيم سعد سنجر	٢٨٣
نحو معجم دلاليّ للمصطلحات الأساسيّة في الخطاب القرآنيّ: قراءة في أربعة ألفاظ من سورة البقرة وفق وجهة نظر أوزتسو في كتابه: (الله والإنسان في القرآن) أ.د. سيد عبد الخالق سيد إسماعيل – د. أحمد محمد بشارات	٣٥٧
المعجميّة العربيّة بين النشأة والمسار والتحوّل: معجم العين ولسان العرب والمعجم التاريخي للغة العربيّة أنموذجًا إيـمان عبد المالك أبو حرب	٣٨٥

افتتاحية التحرير

د. عبد الله طاهر الحذيفي
رئيس التحرير

بسم الله الرحمن الرحيم

منذ اللحظة الأول وجّه هذا العدد أشرعَته صوبَ احتفاءِ جامعة الوصل بدبي بالمنجز الكبير الذي تَوَّجَت به الشارقة معاجمَ اللغة العربية أخيرا بالمعجم التاريخي للغة العربية، وتمثَّل هذا الاحتفاء بأن خصّصت كلية الآداب بالجامعة مؤتمرها الدولي الثالث لـ«المعجمية العربية والدراسات البينية»، الذي أُقيم في نوفمبر ٢٠٢٤. فالتقطت المجلة الثمرةَ اليانعة، أو قل: الطراز الأول من الأبحاث الكثيرة، التي قدمها الباحثون في جلسات المؤتمر، فأضافتها إلى ما كان في خزينتها من الأبحاث المرصودة للنشر، وأُعيد إخضاع الأبحاث المختارة للتحكيم مرة أخرى، لتجريَ في سياق شروط النشر في المجلة، فتمخَّض الحال عن تسعةِ أبحاثٍ تتسم بالأصالة والرصانة والجِدَّة والإضافة، استوعبها هذا العدد.

ولذا يجد القارئ أنّ الباحثين الذين قدموا ثمار جهودهم في هذه الأبحاث قد توجهوا نحو العناية بموقعية بعض المصطلحات والقضايا التي رأوا أنها جديرة بالدرس وأنها على تماسٍّ مباشر مع الدراسات اللغوية والنقدية ودرسوا تطوراتها وامتداداتها في الدرس اللغوي والنقدي العربي والعالمي، وتمكنوا من إثراء كل موضوع بما يستحقه؛ فمن موضوع «الاقتراض اللّغوي ومظاهر التنافس بين العربية ولغات أخرى»، وما احتوى عليه من معالجة لتأثير العربية في عدد من اللغات من ناحية، وتأثرها من ناحية أخرى، إلى «إشكالية المفاهيم الإجرائية في النقد الثقافي: وعي النظرية بين المعجم والاستعمال النقدي»، وما فككه الباحث وأعاد بناءه من مسائل ما تزال تنمو وتتطور في النقد الثقافي الذي عدَّ نفسه بديلا للنقد الجمالي، فإلى موضوع «البورتريه الشِّعري في الشعر الإماراتي، شِعر أمل السَّهلاوي أنموذجًا» وفيه وصف وتحليلي من وجهة نظر أجناسية لشِعْر الشاعرة، من حيث عبور البورتريه من فنّ الرسم والتصوير إلى الشعر، وأشكال البورتريه الشعريّ، وتقنياته... ليأتي موضوع «المصطلح النقدي وشعر النساء: «فحولة الشعراء» للأصمعي أنموذجًا» الذي يذهب إلى تحليل عميق لما كان من تلكّؤ الناقد

العربي القديم عن العناية بشعر النساء، بحجة لِينِه تحت زعم أنّ الشِّعر مرتبطٌ بالفحولة، ومنه إلى موضوع «مساهمة بعض الدول العربية في صناعة معاجم مختصة بالذكاء الاصطناعي، -مقاربة وصفية مقارنة، السعودية والإمارات أنموذجًا-» وفيه يتعرف القارئ على إسهام الدول العربية في صناعة المعاجم المتخصصة بالذكاء الاصطناعي باللغة العربية، كما في معجم البيانات والذكاء الاصطناعي (السعودية)، والمعجم العربي للذكاء الاصطناعي (الإمارات). وما لذلك من قيمة كبيرة في تعزيز مكانة اللغة العربية في ظل تطور التقنية. ليخرج القارئ إلى «توظيفُ الذكاء الاصطناعيّ في صناعة المعجم العربيّ» البحث الذي يهتم بكيفية إنشاء معاجم لغوية عربيّة باستخدام تقنيات الذكاء الاصطناعيّ، وبناء معجم رقميّ للغة العربيّة، يستفيد من التقنيات الحديثة في حل التعقيدات اللغويّة، واللهجية، والمفردات والتراكيب النادرة، وإدارة البيانات، وتصنيف المعاني، واستخدام السياق.. ويأتي بحث «نحْو أَجْرُوميّةٍ للتّأثِيل السَّامي في معجم الشارقة التاريخي» ليتناول البحث في الأصول اللغوية السامية للعربية وأواصر القرابة بينها وبين أخواتها الساميات وتأصيلها... ويذهب البحث التالي إلى اختبار تصوّر بناء معجم دلاليّ للمصطلحات الأساسيّة التي يزخر بها الخطاب القرآني. فيقوم بتحليل أربعة ألفاظ من سورة البقرة هي (جنفٌ، وحنيف، وغُلْف، ورفثٌ) وفق وجهة نظر أوزتسو في كتابه: (الله والإنسان في القرآن)، ثم يأتي بحث «المعجميّة العربيّة بين النشأة والمسار والتحوّل: معجم العين ولسان العرب والمعجم التاريخي للغة العربيّة أنموذجًا» ليضفي على جوِّ هذا العدد ما سبقت الإشارة إليه من ربط العدد بالمعجمية العربية.

وأخيرا تدعو المجلة القرّاء والباحثين وطلبة الدراسات العليا إلى هذه المأدبة الزكية، وقطفِ ثمار جهود العلماء الذين قدموا ويقدمون خير ما لديهم خدمةً للعلم والحقيقة، وإن في متون الأبحاث ما يروي غُلةَ رائدِ المعرفة ومَن يراهم من أهل البحث عن تجويد المنجزات العلمية، وتعميق الدراسات اللغوية والأدبية. ونشير هنا إلى أن موقع الجامعة يحتوي على أعداد مجلة (فكر ومعرفة)، وغيرها من المجلات العلمية الصادة في الجامعة، ويستطيع الباحثون تحميل البحوث التي يريدونها من موقع بوابة الجامعة: https://alwasl.ac.ae/، ورابط المجلة: https://alwasl.ac.ae/ar/fom عبر متصفحات (الإنترنت) مباشرة.

البحـوث

توظيفُ الذكاء الاصطناعيّ في صناعة المعجم العربيّ

Employing Artificial Intelligence in the Making of the Arabic Lexicon

أ. د. سمر جورج الديّوب
أستاذ في قسم اللغة العربية – كلية الآداب والعلوم الإنسانية في جامعة حمص – سوريا

Prof. Samar George Al Dayyou
Professor in the Department of Arabic Language
Faculty of Arts and Humanities, Homs University - Syria

https://doi.org/10.47798/fom.2024.i4.06

ملخص البحث

يركّز هذا البحث على استعراض التحديات، والفرص المرتبطة بإنشاء معاجم لغوية عربيّة باستخدام تقنيات الذكاء الاصطناعيّ، ومعالجة اللغة الطبيعية (NLP). ويهدف إلى تقديم تصوّر شامل لبناء معجم رقميّ للغة العربيّة، يستفيد من التقنيات الحديثة؛ لتوفير معاني المفردات، وتصنيفاتها، وسياقات استخدامها المختلفة.

ويركز البحث أيضًا على الجوانب السلبية، والتحدّيات المتوقّعة، وأهمّها:

تحدّيات التعامل مع التعقيدات اللغويّة، واللهجات، إدارة البيانات الضخمة، ضمان الدقة والتنوع في تحقيق التكامل في معاني المفردات، تقديم أمثلة استخدام شاملة، البحث عن المفردات والتراكيب اللغوية النادرة، التعامل مع اللهجات العربية المتعددة، واختلاف النطق والمعاني، إدارة حجم البيانات الضخمة، تحقيق مستوى عالٍ من الدقة في تصنيف المعاني، واستخدام السياق.

وسيشمل البحث جانبًا نظريًا يبحث في المرتكزات والتحدّيات، وقسمًا عمليًا، ينهض على ما توصّلت إليه المراجع التي حصلنا عليها، والتي ستساعدنا على تقديم تصور للمعجم العربيّ المرتكز إلى الذكاء الاصطناعي، ويحاول أن يضيف إليها. وثمّة دراسات قليلة تناولت استخدام الذكاء الاصطناعي تطبيقيًّا، وتقنيات معالجة اللغة الطبيعية، ومما لا شكّ فيه أنّ بناء المعجم العربيّ باستخدام الذكاء الاصطناعي هو عمل المعجميّ، والمبرمِج معًا.

- الجوانب الإيجابية في توظيف الذكاء الاصطناعي في بناء المعجم العربيّ.

Abstract

This research focuses on reviewing the challenges and opportunities associated with creating Arabic language dictionaries using artificial intelligence (AI) and natural language processing (NLP) techniques. It aims to provide a comprehensive vision for building a digital dictionary of the Arabic language that benefits from modern technologies to provide word meanings, classifications, and various usage contexts.

The research also highlights the negative aspects and expected challenges, most notably:

- Challenges in dealing with linguistic complexities and dialects.
- Managing big data.
- Ensuring accuracy and diversity in integrating word meanings.
- Providing comprehensive usage examples.
- Searching for rare words and linguistic structures.
- Dealing with multiple Arabic dialects and variations in pronunciation and meaning.
- Managing the large volume of data.
- Achieving a high level of accuracy in meaning classification and context usage.

The research will include a theoretical aspect that examines the foundations and challenges, and a practical section based on the findings of the references we have obtained, which will help us provide a vision for an Arabic dictionary based on artificial intelligence and attempt to add to it. There are few studies that have addressed the applied use of artificial intelligence and natural language processing techniques. Undoubtedly, building an Arabic dictionary using

ونسعى في هذا البحث إلى التفصيل في الجوانب الإيجابية الآتية:

- دقّة التحليل اللغوي، وتحسينها.
- التوفير الزمني، والجهد.
- التحديث المستمرّ للمعجم.
- التفاعل مع المستخدِم.
- مسح النصوص القديمة وتحليلها.
- تحسين الترجمة، والفهم اللغوي.
- التوليد النصيّ، وإعادة الصياغة.
- إنشاء معاجم متعدّدة الوسائط.
- التعليم المعزّز.
- إدراج الكلمات الجديدة، والتعامل مع الكلمات المؤقتة.
- تحقيق التوازن بين الاستمرارية والتحديث.
- تطوير المعاجم ثنائية اللغة.
- التعامل مع التعابير المجازيّة.
- مراعاة التنوّع الثقافيّ، واللغويّ.

الكلمات المفتاحية: الذكاء الاصطناعيّ، المعجم العربيّ، الرقميّة، الحوسبة، التقنيات، اللغة الطبيعية.

artificial intelligence is the work of both the lexicographer and the programmer.

Positive Aspects of Employing Artificial Intelligence in Building the Arabic Dictionary.

In this research, we seek to detail the following positive aspects:

- Accuracy and improvement of linguistic analysis.
- Time and effort savings.
- Continuous dictionary updates.
- Interaction with the user.
- Scanning and analyzing old texts.
- Improving translation and linguistic understanding.
- Text generation and rephrasing.
- Creating multimedia dictionaries.
- Enhanced education.
- Including new words and dealing with temporary words.
- Achieving a balance between continuity and updating.
- Developing bilingual dictionaries.
- Dealing with figurative expressions.
- Considering cultural and linguistic diversity.

Keywords: Artificial intelligence, Arabic dictionary, digitization, computing, technologies, natural language.

مقدّمة

الذكاء الاصطناعيّ واللغة الطبيعيّة والهندسة اللغويّة: الاصطلاح والركائز

تعدّ اللغة العربيّة من أغنى اللغات العالميّة من حيث المفردات والتراكيب اللغويّة، وتحتلّ مكانةً مرموقةً في العالَمين: الأدبيّ، والعلميّ. ومع تزايد الاهتمام بتقنيات الذكاء الاصطناعيّ، ومعالجة اللغة الطبيعية (NLP)، يظهر تحدٍّ جديد في كيفية استثمار هذه التقنيات؛ لإنشاء معجم لغويّ شامل، ودقيق للغة العربيّة.

التساؤلات البحثية التي يسعى البحث إلى الإجابة عنها

ما مدى كفاءة تقنيات الذكاء الاصطناعيّ في استخراج المعاني، والعلاقات الدلالية بين الكلمات العربية؟

كيف يمكن للذكاء الاصطناعيّ تحسين عملية تصنيف المفردات العربية وفقًا للسياقات المختلفة؟

ما مدى دقّة تقنيات التعلّم العميق في تحليل الصرف والنحو في اللغة العربية لبناء معجم ذكيّ؟

كيف يمكن دمج تقنيات الشبكات العصبيّة العميقة لتحسين دقة الترجمة المعجميّة بين العربية واللغات الأخرى؟

ما التحديات التي تواجه بناء معجم عربي يعتمد على الذكاء الاصطناعي، وكيف يمكن معالجتها؟

أيّ خوارزميات التعلم الآلي أكثر فعالية في تصنيف الكلمات العربية (مثل: LSTM, BERT, Transformers)؟

كيف يمكن تدريب نماذج معالجة اللغة الطبيعية لتوليد تعريفات دقيقة وموحّدة للكلمات العربية؟

ما تأثير حجم البيانات النصيّة المستخدمة في تدريب أنظمة الذكاء الاصطناعيّ، وجودتها في دقة المعجم الناتج؟ وكيف يمكن تحسين تحليل الجذور، والصيغ الصرفية باستخدام الذكاء الاصطناعي؟ وكيف يمكن دمج المعجم الذكي في التطبيقات اللغوية مثل التدقيق الإملائيّ، والتلخيص التلقائيّ، والترجمة الآليّة؟ وما مدى فعالية استخدام الذكاء الاصطناعيّ في بناء معاجم تخصّصيّة (مثل معجم طبي، قانوني، تقني)؟ وما تأثير المعاجم الذكية في عملية تعليم اللغة العربية للناطقين بغيرها؟

لقد سعينا إلى تقديم تصوّر متكامل للمعجم العربيّ المبني على الذكاء الاصطناعي، وقد لاحظنا اعتماد المراجع التي تمت العودة إليها على الجانب النظري - في أغلبها- وهي مذكورة في البحث.

واللغة العربيّة من اللغة الطبيعية، ونقصد الصورتين الرئيستين لها: المنطوقة، والمكتوبة. وهي محكومة بعوامل التغيير، والتطوّر الذي يصيب أيّة لغة حيّة، ويعدّ الذكاء الاصطناعي وسيلة فعّالة لخدمة اللغة العربيّة بوساطة برامجه الإلكترونية.

وقد ارتبطت اللغة بالحاسوب حتى أصبحا وجهين لعملة واحدة، فتعتمد الدراسات اللغويّة المعاصرة على الحاسوب وبرامجه، ونهضت الدراسات اللسانيّة النظريّة والتطبيقية على تطوير اللغة، وتمّ إحداث فرع في اللسانيات، هو اللسانيات الحاسوبيّة؛ ذلك أنّ اللغة محور العلوم الإنسانية، والحاسوب قمّة التقنيات الحديثة، والتكنولوجيا سمة عصرنا، ومن الطبيعي أن تلتقي اللغةُ الحاسوب.

وتنمو اللغة بنموّ مصطلحاتها؛ لأنّ المصطلح وسيلة تعبّر عن المفاهيم والمدلولات في مستوياتها الصوتيّة، والصرفيّة، والنحويّة، والدلاليّة.

واللسانيات الحاسوبيّة فرع من اللسانيّات التطبيقيّة، وأُطلق عليها تسميات مختلفة: كاللغات الحاسوبيّة، واللغويات المعلوماتية، واللسانيات الآلية، والجامع بين ذلك كلّه أنّها علم يهتمّ باللغة بوصفها أداة طيّعة لمعالجتها في الحاسبات الإلكترونية.

واللسانيات الحاسوبيّة فرع بينيّ، ينتسب نصفُه إلى اللسانيات، وموضوعها اللغة، ونصفه الآخر حاسوبيّ، وموضوعه ترجمة اللغة إلى رموز رياضية يفهمها الحاسوب. فتتألف مبادئ هذا العلم من اللسانيات العامّة بمستوياتها الصوتيّة والنحويّة والدلاليّة، ومن علم الحاسبات الإلكترونية «الحاسوب» ومن علم الذكاء الاصطناعي، وعلم المنطق، وعلم الرياضيات؛ إذ تتناسق هذه العلوم لتشكّل علم اللسانيات الآلي[1].

ومجال معالجة اللغة الطبيعية هو مجال تعليم الآلة لغةَ البشر بوساطة خوارزميات، وأدوات، وتطبيقات تستطيع بها الآلة، أو النظام داخل الآلة معرفة اللغة كما يعرفها اللغويّ المتمرّس في لغته، وهي المعالجة التلقائية للنصّ المكتوب باللغة الطبيعية البشرية «العربية والإنكليزية، والفرنسية، والصينية..» وغيرها بدلاً من اللغات الاصطناعية، مثل لغات البرمجة، والغاية من تلك المعالجة محاولة فهم النصّ، ويطلق عليه بعضهم اسم اللغويات الحاسوبيّة، أو هندسة اللغة الطبيعية[2].

والعمل في ميدان المعالجة الآليّة للغة الطبيعيّة عمل ذو طبيعة هندسيّة، أدواته خوارزميات مبنيّة على رياضيات متقدّمة، وموضوعه اللغة العربيّة.

١- للتوسع في هذا الميدان ينظر: عبد الرحمن بن حسن العارف، توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية "جهود ونتائج" مجلة مجمع اللغة العربية الأردني، ص٥٢.

٢- ينظر: ديانا مانيارد، كالينا بونتشيفا، إيزابيل أوغنتشاين، معالجة اللغة الطبيعية للويب الدلالي، ص٢٣

وتتقاطع معالجة اللغة الطبيعية[١] «Natural Language Processing» مع فروع علميّة مثل اللسانيّات، أو اللغويّات وهندسة الحاسب، والالكترونيّات، وعلم الإحصاء.

وضع جون ماكارتي John McCarthy مصطلح الذكاء الاصطناعي[٢] أوّل مرة عام ١٩٥٦ في جامعة دارتموت Dartmouth في الولايات المتحدة الأمريكيّة. وهو ميدان يهتم بمعالجة اللغات، أو ما يُعرف بالهندسة اللغويّة، فيعالج اللغة الطبيعية بوساطة الحاسوب، ويشمل العلوم المصطلحيّة، والترجمة الآليّة، والاهتمام بمحركات البحث، وغير ذلك، ويقوم على كفاءتين: لسانيّة، وحاسوبيّة، فيعالج الباحث المعالجة الآليّة للغات عن طريق إخضاع الظواهر اللغويّة للتفسير الآليّ[٣].

إنّ اللسانيات الحاسوبيّة دراسة علميّة للغة العربيّة من منظور حاسوبيّ، فحين تدخل اللغة إلى الحاسوب تصبح لغة ذات ركيزة رياضيّة، فتنهض اللغة على رموز، وتجريد ذهنيّ، والرياضيات ذات طابع تجريديّ، رمزيّ.

والذكاء الاصطناعيّ من العلوم الجديدة ذات الخلفيات العلميّة المتعدّدة، يهدف إلى فهم ملكة الإنسان؛ لكي يستطيع استيعاب المعرفة البشرية، ويقوم بإعادة إنتاج بعض مظاهر الذكاء الإنسانيّ من إدراك العلامات اللغويّة وغير اللغويّة، واستيعابها، وتخزينها، وإعادة استخدامها في مقامات جديدة.

١- معالجة اللغة الطبيعية مجال معرفيّ يُسعى به إلى توجيه الآلة إلى فهم اللغة الطبيعية عبر مستوياتها المتعدّدة، ومعالجة وحداتها في هذه المستويات تحليلًا، وتوليدًا، وخلق بيئة تفاعلية قادرة على تحقيق التواصل بين الإنسان والآلة. محمد العطية وآخرون، العربية والذكاء الاصطناعي، ص١٤.

٢- مرّ الذكاء الاصطناعي -بمفهومه الحديث- بثلاثة أطوار رئيسة، تتمثّل في: طور النشأة، وطور التجريب، وطور النهضة التي نلمسها حاليًا.
محمد العطية، وآخرون، العربية والذكاء الاصطناعي، ص١٣.

٣- فيما يتعلق بالمعالجة الآلية للغات ينظر:
Lane Hobson, Hapke Hannes, et al, 2019, Natural Language Processing in Action; Understanding, Analyzing, and Generating Toxteth Python, U S A Marring publication.

وقد اختار مصطلح الذكاء؛ ليدلّلوا على قدرته على محاكاة العقل، وفهم التجربة الشعوريّة لدى الإنسان، وكيفية عمل الذهن في تعامله مع اللغة معرفة واكتسابًا واستعمالاً، فيحاكي اشتغال العقل الإنسانيّ بلغة صوريّة خوارزميّة شبيهة باللغة الاصطناعيّة[1].

ويعدّ الذكاء الاصطناعي Artificial Intelligence أهم نتائج الثورة الاصطناعيّة الرابعة[2]. وقد دخل في ميادين متعدّدة باللغة العربية، كالتطبيقات العربيّة على الهواتف المحمولة، وبرمجتها بحروف عربيّة، والميدان الطبّي، والاقتصاديّ، والتجاريّ، ووسائل التواصل الاجتماعيّ. ومن ذلك المعجم العربي بوساطة الذكاء الاصطناعيّ.

وثمّة تعريفات متعددة للذكاء الاصطناعيّ، فهو ميدان من ميادين علم الحاسوب، ويهتمّ بأتمتة السلوك الذكيّ، وهو دراسة الحاسبات التي تمكّننا من استقبال المعرفة، وتحليلها، والعمل بها، والدراسة التي تعنى بجعل الآلة تقوم بالوظائف التي يقوم بها الإنسان في زمننا الحاليّ بشكل أفضل. وذلك بتطوير الآلة، أو نظام يحاكي السلوك الذكيّ.

إنّه تصميم أنظمة حاسوبيّة بخوارزميّات معينة تحاكي قدرات البشر الفكرية، ويعني ذلك أن تمتلك الآلة قدرة على تقليد العمليات الحركيّة والذهنيّة عند الإنسان، ومحاكاتها، وطريقة عمل عقله في التعبير والاستنتاج والإفادة من

١- ينظر: بلقاسم اليوبي، اللسانيات الحاسوبية: مفهومها -وتطوراتها- ومجالات تطبيقها (استشراف آفاق جديدة لخدمة اللغة العربية وثقافتها)، مجلة مكناسة، ص٤٤.

٢- مرت الحضارة الحديثة بثلاث ثورات صناعية هي: اختراع المحرّك البخاري "١٧٦٠- ١٨٤٠م" وثورة الكهرباء في نهاية القرن التاسع عشر، وظهرت في ستينيات القرن العشرين الحوسبة الرقمية والكمبيوترات المركزية والشخصية والانترنت. ثم ظهرت ثورة الذكاء الاصطناعي.
ينظر: أحمد ماجد، الذكاء الاصطناعي في دولة الإمارات العربية، إدارة الدراسات والسياسات الاقتصادية، ص٤ وما بعدها.
وللتوسع في مصطلح الذكاء الاصطناعي ينظر:
عادل عبد النور بن عبد النور، مدخل إلى عالم الذكاء الاصطناعي، ص٧.

التجارب السابقة، وردود الفعل الذكيّة، فهو مضاهاة لعقل الإنسان، والقيام بدوره.

أما الهندسة اللغويّة[١] فهي تطويع اللغة الطبيعية في نماذج رياضيّة معينة، ثم برمجتها في حواسيب إلكترونية بالإفادة من الذكاء الاصطناعيّ، وعلوم الهندسة، وهندسة المعرفة.

لقد أضحت اللغة ظاهرة حاسوبية قابلة للمعالجة الآليّة، فاللغة ظاهرة إنسانيّة انتقلت إلى الآلة، وتعالج هندسةُ اللغة بوصفها منظومة من الخوارزميّات. أما اللسانيّات الحاسوبيّة فتنطلق من البناء اللغويّ لبناء أنساق صوريّة بهدف ملاءمة التوصيف اللغويّ مع متطلبات الهندسة؛ أي التحكم في النظم بشكل عام[٢]

يعني ما سبق أنّ النتيجة واحدة، وهي تطوير تطبيقات تسهم في انتقالنا

١- الهندسة اللغوية مجال متعدد التخصّصات، يجمع بين علوم الحاسوب، واللغويات؛ لتطوير تقنيات معالجة اللغة الطبيعية (NLP). ويشمل هذا الميدان تصميم الأنظمة والأدوات التي تعالج اللغة البشرية بشكل آلي، وتطويرها، مثل تحليل النصوص، والترجمة الآلية، والتعرف إلى الكلام، وتوليد اللغة الطبيعية.

وتهدف الهندسة اللغوية إلى تمكين الحواسيب من فهم اللغة البشرية، واستخدامها بطرق تتجاوز المستوى الأساسي، مما يساعد في تطبيقات متعددة مثل ومحركات البحث، والمساعدات الصوتية، وأنظمة التعرف إلى الصوت، وتحليل المشاعر في النصوص.

وللحصول على معلومات تتعلق بالهندسة اللغوية ينظر:

«Foundations of Statistical Natural Language Processing» by Christopher D. Manning and Hinrich Schütze,1999, M I T Press:

يقدّم الكتاب يقدم أساسيات معالجة اللغة الطبيعية باستخدام الأساليب الإحصائية، وهو مرجع مهم في هذا الميدان.

«Handbook of Natural Language Processing» edited by Nitin Indurkhya and Fred J. Damerau: 2010, C R C Press.:

يغطي مواضيع متعددة في معالجة اللغة الطبيعية، ويعتبر مصدرًا جيدًا للمعلومات المتعلقة بالهندسة اللغوية.

٢- للتوسع في هذه الفكرة ينظر: عمر مهديوي، مصطلحات الهندسة اللغوية بين الترجمة والتعريب: نحو بناء معجم موحد، مجلة اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد ٧٩

إلى مجتمع المعرفة بلغتنا الأم، الأمر الذي يتطلب إعادة توصيف للغة؛ لتلائم متطلبات التقانة الجديدة.

وثمّة فرق بين الوصف والتوصيف[1]، فقد أُطلِق على ما يُعْمل للإنسان «الوصف»، وما يُعْمَلُ للحاسوب «التوصيف». وبيان الفرق بينهما ماثل في أنّ وصف العربية ما وقع للعلماء العرب من قواعد مستنبطة من الأداء اللغويّ الواقعيّ، وهو مبنيّ في شطر منه على أن المُسْتقبِل يسهم إسهامًا فاعلًا في الحدث التواصليّ، يضاف إلى ذلك ما يتحصّل للإنسان من معرفة بالحدس، والسليقة، والخبرة المعرفية، والتثقُّف، والعرف اللغويّ، والمقام.

وأما التوصيف فإنّه ينظم الوصفّ اللغويّ المجرد، يضاف إليه العناصر التي يتعرّفها الإنسان بالحدس، والسليقة، والقرائن المتعددة: اللفظيّة، والمعنويّة، والموقفيّة. ولما كان الحدس أظهر ما يتكئ عليه الإنسان في تعرُّفه اللغةَ وأداءها، ولما كان الحاسوب يفتقر إلى هذا العنصر البشريّ الخالص، وجب على المُوَصِّف أن يتدارك هذا النقص، ليبلغ بالحاسوب مَبْلَغَ المعرفة الإنسانية باللغة.

فعلى سبيل المثال لا يحتاج الإنسان إلى التفريق بين السوط والصوت، لكنّ الحاسوب بحاجة إلى دليل مستقصى من الدوال لإدراك الفارق بينهما، وكذلك الفرق بين الإنسان والآلة في قولنا: قرأ الطالب الكتاب، فلا يصحّ أن يكون الكتاب فاعلًا مهما كان موقعه في الجملة، حتى لو قلنا: قرأ الكتاب الطالب، أما الحاسوب فيحتاج قائمة من البيانات المعجميّة حول كلّ كلمة، تحدّد سلوكها في الجملة.

وتقوم اللغة العربية على مكوّنين رياضيين غير موجودين في كثير من لغات العالم، هما: الجذر، والوزن. والجذر هو البنية الأساسيّة للكلمة، ويتولّى الوزن وضع هيكلها العام، يوزّع الحركات على حروف الكلمة، ويوزع المورفيمات

١- ينظر: نهاد الموسى: العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، ص٦٩- ٧٠

التي تضاف إلى الجذر، وهذا ما جعل اللغة العربية لغة انصهارية، وهي ميزة تختلف عن اللغات الأخرى التي تعدّ إلصاقية.

تجمع الهندسة اللغويّة اللسانيات وعلم الحاسوب، وترتبط التكنولوجيا بالعلم والمعرفة، وتقوم الهندسة اللغوية على ركائز. منها:

- فهم اللغة الطبيعية التي نستخدمها.
- استخدام أساليب التعلم الآلي لتحليل الكلمات، وتفسيرها، وإنشاء عبارات لفهم نيّة المستخدِم، أو مشاعره.
- استخدام تقنيات البرمجة اللغويّة العصبيّة للمهام، كفهم اللغة الطبيعية N L U وتوليد اللغة الطبيعية N L G والترجمة الآليّة، والتعرّف إلى الكلام، وتحليل المشاعر....

ويكون التحليل النحوي باستخدام خوارزميات لتحديد بنية الجمل، وتحليلها؛ لفهم كيفية تركيب الجملة، وفهم المعنى الكامن وراء الجمل والكلمات.

أما التحليل الدلالي فهو عملية فهم الكلمات والجمل في سياقها، ويمكن له فهم نياتنا بشكل أفضل، والاستجابة وفقًا لذلك.

٢- المعجم العربيّ والذكاء الاصطناعيّ:

يتمّ تصوّر المعجم العربيّ القائم على الذكاء الاصطناعيّ بأنّه معجم ناجم عن دمج معاجم متعددة، تعزى فيه الشروحات إلى أصحابها، ويكون له تحديثات مستمرّة.

ولا يكون القيام بهذه المهمّة من قبل أشخاص، بل من قبل مجموعة ذات تخصّص بالمعجمية واللسانيّات الحاسوبيّة. ولا يتقيد هذا المعجم بحجم معين

لإتقان تخزينه على وسائط كبيرة الحجم، ويكون قابلًا للتعديل والحذف والإضافة بناء على رغبة منشئه، يتميز بالدقة في البحث، وفي إيجاد المعلومة المطلوبة بسرعة، ولا يتقيد بطريقة واحدة في البحث.

واللغة العربية -شأن سائر اللغات- عماد الشخصية العربيّة، وتجسيد حيّ لمعارفنا وخبراتنا، وتعاني اضطرابًا ناجمًا عن اضطراب الناطقين بها فكريًا وثقافيًّا، واجتماعيًّا، واقتصاديًّا؛ وذلك بسبب كثرة العاميّات، وابتعادها عن الفصحى من جهة، وطغيان بعض اللغات الأخرى من جهة ثانية، وذلك بسبب تأثير البيئة الثقافية والفكرية؛ إذ تتأثّر اللغة بالسياق الثقافيّ، والفكريّ لمتحدّثيها، فكلما كان هناك استقرار فكريّ، وثقافيّ، انعكس ذلك إيجابًا على اللغة، والعكس صحيح.

ويؤدي تعدّد اللهجات العامية، وتباعدها عن الفصحى إلى ضعف الترابط اللغويّ، مما يسهم في اضطراب اللغة، وعدم ثباتها على قواعد محدّدة.

ويؤثّر طغيان اللغات الأجنبية، سواء بسبب الاستعمار التاريخي أو العولمة الحديثة، في بنية اللغة الفصحى، ويؤدي إلى تراجعها لصالح اللغات الأخرى.

ونحتاج إلى إكمال المعاجم التراثية برصد دقيق لألفاظ الحضارة، والمصطلحات العلمية في فروع المعرفة المختلفة حتى تكتمل صورة اللغة العربية عبر الزمن بهدف إعداد معجم يعالج صرفيًا واشتقاقيًا، تحليلًا وتوليدًا للجذور والبنيات المشتقة منه، ومعالجة المفردات البسيطة، والمركبة.

ويشتمل الذكاء الاصطناعي على خاصيّة الوسائط المتعددة في عرض المادة اللغويّة بحسب الحاجة إليها كالأصوات، والصور، والفيديوهات، وعرض الكلمات كتابة ونطقًا، وأصل الكلمة، واستعمالها، ومعانيها، واستعمال وسائل الشرح المشهورة كالتفسير بالمغايرة، والتفسير بالمصاحبة، والتفسير بالسياق، والتفسير بالصورة.

ويتمّ فيه تسهيل العودة، والاستعمال للمداخل المعجميّة المراد البحث عنها، وتعديل المعلومات المطلوبة في زمن قصير، ويرى الباحث الكلمات، ويسمعها، فهو معجم سمعيّ، بصريّ.

ويتّسم بالتفاعلية؛ أي أن يكون المستخدِم قادرًا على الاقتراح، والنقد، والإضافة. ويختصر الإحصاء الحاسوبيّ جهد سنوات، ويساعد علماء البلاغة في تحديد شروط الفصاحة، وتصحيح بعض الأحكام الخاطئة التي وردت في كلام العرب، وتمييز الدخيل من المعرب، وتنمية مهارات اللغة العربية بالاعتماد على إطار لسانيّ نظريّ قادر على توصيف الظاهرة اللغوية[1].

ويمكن للذكاء الاصطناعيّ إذا ما زودناه بالمعلومات الكافية أن يساعد في تحديد المعاني بدقّة، ويكتشف الاختلاف اليسير في العبارات المتشابهة. ويحفظ هذا الذكاءُ المادة اللغويّة، ويخزّنها رقميًّا بمساحات كبيرة، ويحدّث البيانات المدخلة حاسوبيًا، ويتمّ التمكين من استرجاع نصّ لغويّ محدّد من بين أعداد ضخمة من النصوص المدخلة، ويحقّق المنهجيّة، والموضوعيّة، والدقّة في البحث اللغويّ.

ويحدّد المعنى الدقيق لأيّ مصطلح، وترجمته، والمعاني المغايرة له، ويستطيع تحليل العلاقة بين جذر الكلمة وصيغها الصرفية، ويوضح المدلولات التي تبنى عليها صياغة مفردات المعجم العربيّ.

إنّه معجم دقيق لا يكتفي بوضع المقابل العربيّ للمصطلح الأجنبيّ، بل يعرّفنا بهذا المصطلح، فثمّة تزايد سريع في علم المصطلح مع توافر الحاسوبية، وعصر المعلومات، والاقتصاد المبنيّ على المعرفة.

١- ينظر: محمد مدور، اللسانيات الحاسوبية وأثرها في تطوير اللغة العربية، أعمال ملتقى وطني حول اللغة العربية والتقانات الحديثة، ص٤٤- ٤٨.

وتعدّ اللغة العربية من أكثر اللغات قابليّة للتمثيل الحاسوبيّ، وذلك بسبب:

الطبيعة الجبريّة الرياضية للغة العربية، بسبب انتظام مستوياتها الصوتيّة، والصرفيّة، والإعرابيّة، والصلة القويّة بين المعنى، والمبنى.

صغر عدد الجذور قياسًا بتعدّد المفردات.

وجود فائض لغويّ يمثّل مصدرًا للتعدّد، والتبادلات التركيبيّة التي تمثّل إحدى العقبات الأساسية لمعالجة اللغة آليًّا[1].

وبناء على ما تقدّم يجب القيام بما يأتي:

- إعادة توصيف اللغة العربية صرفيًّا، ونحويًّا، ودلاليًّا بشكل يناسب منطق الحاسوب.
- إنجاز أنموذج لساني يمكّن من تمثيل مكوّنات اللغة من فعل، واسم، وحرف.
- إنجاز أنموذج خوارزميّ بلغة الحاسوب.

ويمكن للذكاء الاصطناعيّ القيام بعمليات متعددة في وقت واحد، ويمكن له أن يعمل لوقت غير محدود، وأن يحتفظ بالقدرة نفسها على الدقة، والأداء في مدة العمل كلها.

١- للتوسع في هذه الفكرة ينظر:
نبيل علي، اللغة العربية والحاسوب، عالم الفكر، المجلد ١٨، العدد٣، ص١٠٣- ١٠٤

لكنّ ثمّة عقبات تواجه هذا العمل، منها:

- وقوع اللبس في النصّ العربيّ غير المشكول، وتمايز الأبنية بتمايز الأصوات، مثل «نظر، نظم»، وتمايز المفردات في اللغة بأصواتها، كالتمايز الصوتي بين «فائدة وعائدة».
- ثمّة صعوبة في ضبط تحوّلات اللغة، وعلائق الكلم، وأسرار النظم، وأصول البيان.

٣- تصوُّر عن بناء المعجم العربي على المستوى الصرفيّ باستخدام الذكاء الاصطناعيّ:

تعدّ المعالجة الآليّة للغة العربية متأخرة نسبيًا عن اللغات الأجنبية، مما أحدث فجوة رقميّة بين لغة الضاد واللغات المتقدّمة تقنيًّا، ومعرفيًّا.

وإنشاء معجم لغويّ باللغة العربية باستخدام الذكاء الاصطناعيّ مشروع مثير يمكن أن يسهم في تعزيز المعرفة اللغويّة، والبحث الأكاديميّ، والتطبيقات العمليّة.

ومن المفروض أن تتمّ المعالجة على مستويات، منها المستوى الصرفي الذي يعنى بدراسة سلوك الأفعال العربية؛ للتعرف إلى الأفعال المتصرّفة، وكشف أصولها، وتدقيقها إملائيًّا.

ويحتل علم الصرف مكانة مهمّة في المعالجة الآلية للغة العربية؛ ذلك لأنّه يربط عناصر المنظومة اللغويّة، ويوصل إلى تحليل صرفيّ ناجح، مرهون بالبدء الصحيح، فثمّة علاقة قويّة بين التحليل الصرفي وبناء المعجم العربيّ بوساطة الذكاء الاصطناعيّ.

وتبدأ عملية تصريف الأفعال العربية بالفعل الثلاثيّ الصحيح السالم، ويصرّف في الأزمة كلّها، مع المتعدّي، والمبني للمعلوم، والمبني للمجهول، وتُذكر خصائص الفعل.

ويمكن العمل على أيقونة خيارات؛ ليظهر خيار حركة بنية الفعل، والبحث في المعجم، أما خانة الأزمنة فيظهر فيها الفعل متصرّفًا في حالاته كلها، ويمكن اختيار الزمن الذي نريده.

ويكون هنالك أيقونة خاصة بالضمائر، وأيقونة خاصة بنوع الفعل من جهة اللازم والمتعدي، والرباعي، والخماسي، والسداسي منه[١].

ويعود تحليل الكلمة إلى الجذر «Root» وذلك بردّ الكلمة إلى حروفها الأصلية، والجِذع «Stem» وهو صورة الكلمة المعيّنة بعد تجريدها من السوابق واللواحق، والفرع «Lemma» وهو المقابل الصرفيّ المجرّد للوحدة المعجميّة في الكلمة، ويأتي على صورة الماضي المسند الغائب، والزائدة «Affix» وهو مورفيم لاحق يسبق جذع الكلمة في بدايته، فيكون سابقة «Prefix» أو يلتحق به في نهايته، فيكون لاحقة «Suffix»[٢]

ونأخذ مثالًا على ذلك الفعل «سيغادرون»، فجذره «غ د ر»، وجِذعه «يُغادر»، وفرعه «غادر، وزائدته السابقة «سـ»واللاحقة «ون».

وتوجّه الآلة مدلولات الكلمات المجرّدة من الضبط، وترتّب احتمالات معالجة الوحدات الصرفيّة بحسب دورانها في واقع اللغة، وتُدرّب مدوّنة لغوية على المعالجة العميقة.

١- للتوسع في الحديث عن العلاج الصرفي للغة العربية ينظر:
عمر مهديوي، اللسانيات الحاسوبية واللغة العربية، إشكالات وحلول، ص٤٥.

٢- للتوسع ينظر د. محمد عطية وآخرون، العربية والذكاء الاصطناعي، ص١٠٣.

فلدينا عبارة «المؤتمر الدوليّ»على سبيل المثال، هي أقل من جملة، تشكّل وحدة تركيبية كاملة حين ترد في سياق، وجملة «المؤتمر الدولي وسيلة ناجعة لتطوير الحوار»وهي جملة كاملة الأركان، ركناها المبتدأ والخبر، وتمثّل وحدة تركيبية حين ترد في سياق لغوي مكتمل.

أما الفقرة «المؤتمر الدوليّ وسيلة ناجعة لتطوير الحوار؛ إذ يعدّ ملتقى للباحثين الجادّين»فهي كيان تركيبيّ، حيّزه أكبر من الجملة، وتمثّل مجموعة جمل مترابطة حول فكرة معيّنة.

٤– ركائز المعجم العربيّ بوساطة الذكاء الاصطناعيّ[١]:

يجب أن يتسم المعجم العربيّ المبنيّ على الذكاء الاصطناعيّ بقدرات متعدّدة، من أهمّها:

- **تحليل النصوص**: يتمّ استخدام تقنيات معالجة اللغة الطبيعية (NLP) لتحليل النصوص العربيّة، واستخراج المفردات والمعاني من مصادر متعدّدة مثل الكتب، والمقالات، والمواقع الإلكترونية. ويتمّ فصل النصوص إلى وحدات

١- لإنشاء معجم لغوي باللغة العربية باستخدام الذكاء الاصطناعي، يمكن الرجوع إلى المراجع الآتية:
- كتاب "اللغة العربية والذكاء الاصطناعي" للدكتور علاء طعيمة: يقدم مشاريع عملية متنوعة في مجالات معالجة اللغة الطبيعية للغة العربية باستخدام تقنيات التعلم الآلي والتعلم العميق.
رابط الكتاب: [كتاب اللغة العربية والذكاء الاصطناعي] (https://dlarabic.com)
- مقالة «الذكاء الاصطناعي واللغة العربية: أهم التحديات وتقنيات معالجتها» من إم آي تي تكنولوجي ريفيو: تناقش التحديات التي تواجه معالجة اللغة الطبيعية للغة العربية مثل الشح في الموارد وتعدد اللهجات.
رابط المقالة: [الذكاء الاصطناعي واللغة العربية: أهم التحديات وتقنيات معالجتها]
(https://technologyreview.ae)
- مقالة «من أفضل أدوات الذكاء الاصطناعي التوليدي باللغة العربية»من إم آي تي تكنولوجي ريفيو: تستعرض مجموعة من النماذج والأدوات المتقدمة المستخدمة في معالجة اللغة الطبيعية للغة العربية، مثل نموذج «نور».
رابط المقالة: [أفضل أدوات الذكاء الاصطناعي التوليدي باللغة العربية]
(https://technologyreview.ae)

لغويّة صغيرة مثل الكلمات، والجمل، والفقرات.

ويتمّ ذلك بتطبيق يمكن له تحليل النصوص الطويلة، واستخراج الكلمات المفتاحيّة، وتعريفاتها. ومن التقنيات المستخدمة في هذا الميدان:

- **تحليل البيانات الضخمة** (Big Data Analysis): تستخدم أدوات تحليل البيانات الضخمة لجمع البيانات النصية، ومعالجتها من مصادر متنوعة مثل الإنترنت، وقواعد البيانات، والكتب الإلكترونية. ويمكن للذكاء الاصطناعيّ تحليل هذه النصوص؛ لاستخراج الكلمات، والتعابير الجديدة، وكذلك لفهم تردّد الكلمات، واستخداماتها في سياقات مختلفة.
- **خوارزميّات معالجة اللغة الطبيعية** (NLP): تشمل تقنيات معالجة اللغة الطبيعية أدوات مثل تحليل المشاعر، والتعرف إلى الكيانات المذكورة (NER)، وتحليل الموضوعات. وتستخدم هذه التقنيات لفهم المعاني الضمنيّة للنصوص، والتعرف إلى الأنماط اللغويّة المعقّدة.

وتساعد هذه العملية على اكتشاف الكلمات الجديدة، ولاسيما تلك الكلمات الواردة على الشابكة، ويمكن تضمينها في المعجمات؛ لضمان تحديثها باستمرار. ويمكن تحليل استخدام الكلمات في مختلف السياقات، ويستطيع الذكاء الاصطناعي تحديد الاتجاهات اللغوية الناشئة والتغيرات في الاستخدام اللغويّ. ويمكن أن يساعد ذلك على فهم تطور اللغة عبر الزمن.

- **التعلّم العميق**: هو مجموعة فرعية من التعلم الآليّ، يركّز على دراسة الآلات التي يمكن لها التعلّم، وتطويرها. ويتيح التعلم العميق تحويل البيانات الأوليّة إلى متّجهات، وإدخالها إلى الشبكة من غير الحاجة إلى استخراج هذه الميزة يدويًا من البيانات، وتُستخرج بشكل أفضل من الخبير من خلال

مجموعة طبقات مختلفة في هيكلها بناء على البيانات التعليمية المتعلقة بالمشكلة[١].

يتمّ تدريب نماذج التعلم العميق[٢] على قواعد اللغة العربية، والأنماط اللغوية؛ لتوليد تعريفات دقيقة، ومفصّلة للكلمات. ويتمّ تحليل الكلمات باستخدام خوارزميّات التعلم العميق؛ لاستخراج معانيها، وسياقات استخدامها. ويمكن للذكاء الاصطناعيّ تعلّم الأنماط اللغويّة من النصوص، واكتشاف معانٍ جديدة للكلمات بناءً على السياق الذي ظهرت فيه.

- **التصحيح الآليّ**: يتمّ تطوير نظم تصحيح لغويّ، تعتمد على الذكاء الاصطناعيّ؛ لتحسين جودة النصوص، وتقديم اقتراحات؛ لتصحيح الأخطاء النحويّة، والإملائيّة.

- **الترجمة الآليّة**: يتمّ تعزيز تقنيات الترجمة الآليّة بين العربية واللغات الأخرى؛ لتوفير معاجم ثنائيّة اللغة، تعتمد على الذكاء الاصطناعي.

وتستخدم نماذج الترجمة الآلية مثل Google Translate وDeeply لتحويل النصوص من لغة إلى أخرى بدقة عالية.

وتساعد تقنيات الترجمة الآلية في تطوير معاجم ثنائية اللغة من خلال تحليل النصوص في لغات مختلفة، وتحديد المعاني المكافئة للكلمات. وتحسين خدمات الترجمة الفوريّة، بتحليل النصوص بدقّة، وسرعة.

١- ميلاد وزان، التعلم العميق: المبادئ، والمفاهيم، والأساليب، ص٢٧، ٢٩.

٢- ثمة دراسات تهدف إلى تحسين نماذج التعلم العميق؛ لاستخراج المعاني، وتصنيف النصوص العربية. ويمكن العثور على أوراق بحثية تناقش استخدام نماذج مثل BERT وGPT في السياقات العربية.
ومن الأمثلة على ذلك: «Arabic Text Classification Using Deep Learning Techniques» والتي تتناول كيفية تطبيق تقنيات التعلم العميق على النصوص العربية.

- **التحليل الدلاليّ متعدّد اللغات** (Multilingual Semantic Analysis): تشمل هذه التقنية تحليل العلاقات الدلالية بين الكلمات في لغات مختلفة، مما يساعد في تطوير معاجم ثنائيّة اللغة.
- **تحليل الصوت والنص** (Speech and Text Analysis): تستخدم هذه التقنية لتحليل النصوص المنطوقة، وتحويلها إلى نصوص مكتوبة، مما يسهل إدراج الكلمات، والتعابير المنطوقة في المعجم.
- **تحليل النصوص متعدّدة اللغات**: تساعد تقنيات تحليل اللغات في تحليل النصوص متعددة اللغات، وفهم العلاقات بين الكلمات في لغات مختلفة، مما يسهم في تطوير معاجم شاملة، ومتعددة اللغات.

وتحتاج أنظمة الترجمة الآلية إلى التعامل مع الفروقات الثقافيّة، واللغويّة بين النصوص المختلفة؛ لضمان دقة الترجمة. وتتطلب أنظمة الترجمة الآلية تحسينًا مستمرًا للتعامل مع التغيرات اللغوية، والتطورات الجديدة في اللغة.

- **التصنيف والفهرسة**: يتمّ استخدام الذكاء الاصطناعي لتصنيف الكلمات، وفهرستها بناءً على معانيها، واستخداماتها المختلفة في النصوص، مما يسهل عملية البحث، والوصول إلى المعلومات.
- **التفاعل الصوتي**: يتمّ تطوير واجهات صوتية تفاعلية تمكِّن المستخدمين من البحث عن معاني الكلمات من خلال الأوامر الصوتية، باستخدام تقنيات التعرّف إلى الصوت.

وذلك بالمساعدات الصوتية (Voice Assistants) إذ تستخدم هذه المساعدات مثل Siri وGoogle Assistant لتقديم المعلومات اللغوية والنصية بشكل تفاعلي.

والتطبيقات الذكية (Smart Apps) التي تشمل تطبيقات المعاجم التفاعليّة التي تعتمد على تقنيات التعلّم العميق، والتفاعل الذكيّ؛ لتوفير معلومات لغويّة دقيقة.

وتستخدم تقنيات الواقع المعزّز (Augmented Reality) لتقديم المعلومات اللغويّة بشكل تفاعليّ ومرئيّ، مما يسهِّل على المستخدمين فهم المعاني، واستخدام الكلمات في السياقات المختلفة.

وتساعد الواجهات التفاعلية في تسهيل الوصول إلى المعاجم، والمعلومات اللغوية من خلال البحث الصوتيّ، والتفاعليّ. وتوفّر الواجهات التفاعليّة تجربة مستخدم محسّنة بتقديم معلومات دقيقة، وسهلة الوصول بطريقة تفاعلية.

ويمكن استخدامها في تطبيقات التعليم؛ لتقديم دروس لغويّة، تفاعليّة تساعد في تحسين تعلم اللغة، وفهم المعاني. لكنها تحتاج إلى القدرة على التعامل مع التعقيدات اللغويّة، وفهم المعاني الضمنيّة للنصوص. ويجب أن تكون قادرة على التكامل مع الأنظمة الأخرى، وتقديم معلومات دقيقة ومتّسقة.

- **التعليم الإلكتروني**[1]: يتمّ تصميم تطبيقات تعليميّة تفاعليّة تعتمد على الذكاء الاصطناعيّ؛ لتعليم المفردات، والقواعد اللغويّة بطريقة مبتكرة، وفعّالة.

والتعلم الآليّ هو أحد فروع الذكاء الاصطناعيّ الذي يسمح للأنظمة بالتعلّم، والتحسّن من البيانات من دون أن تكون مبرمجة بشكل صريح. ويستخدم لتحليل البيانات اللغويّة، واستخراج القواعد، والأنماط اللغويّة بتلقائيّة.

١- ثمّة أبحاث تتناول كيفية إنشاء أدوات تعليمية تعتمد على الذكاء الاصطناعي؛ لتسهيل تعلم اللغة العربية. من ذلك: «Interactive AI-based Tools for Arabic Language Learning» الذي يشرح كيفية تطوير تطبيقات، وأدوات تعليميّة تساعد المتعلمين على اكتساب اللغة بشكل أكثر فعاليّة.

ويستخدم التعلم الخاضع للإشراف (Supervised Learning) في إنشائه، ويتضمن تدريب النماذج على مجموعة من البيانات المعلمة، حيث تكون البيانات متبوعة بالنتائج المرجوّة. ويمكن استخدام هذه الطريقة لتدريب النماذج على التعرف إلى الاستخدام الصحيح للكلمات، وتصحيح الأخطاء اللغويّة.

والتعلم غير الخاضع للإشراف (Unsupervised Learning) وفي هذه الحالة، تحاول النماذج العثور على الأنماط في البيانات من غير توفير نتائج مرجوة سابقًا. يستخدم هذا الأسلوب لاكتشاف الفئات اللغويّة، أو تجميع الكلمات المتشابهة دلاليًا.

- **التعلم المعزّز** (Reinforcement Learning): ويشير هذا المصطلح وصف التقنية التي تسمح بمزج واقعي متزامن لمحتوى رقمي من البرمجيّات، والكائنات الحاسوبية بشكل ثنائيّ، أو ثلاثيّ الأبعاد مع العالم الحقيقي[1].

ويمكن استخدامه لتحسين التفاعلات مع المستخدمين بتحليل ردود الفعل، وتعديل النهج بناءً على التفاعل.

ونأخذ مثالًا على ذلك: تحديد القواعد النحوية، فيمكن للنماذج المدرّبة باستخدام التعلم الآليّ تحديد القواعد النحوية، وتصحيح الأخطاء اللغويّة بتلقائيّة، مما يساعد في تحسين جودة النصوص اللغويّة.

- **تحسين التعريفات**: يمكن استخدام التعلّم الآلي لتحليل السياقات المختلفة التي تُستخدم فيها الكلمات، مما يساعد في توفير تعريفات أكثر دقّة، وشموليّة في المعاجم.

١- ثمّة دراسات متعدّدة تتحدّث عن الواقع المعزّز من أهمها:
Tobias Hollerer, Dieter Schmalstieg,: "Augmented Reality: 2015, Principles and Practice» Addison-Wesley Professional

- **التصنيف الدلاليّ**[١]: تستخدم تقنيات التعلّم الآليّ لتصنيف الكلمات بناءً على معانيها، وعلاقاتها بالكلمات الأخرى، مما يساعد في تنظيم المعجم العربيّ بشكل أفضل. ويستخدم للتصنيف الدلالي ما يأتي:
- **النماذج الدلالية** (Semantic Models): تستخدم لتمثيل المعاني اللغويّة، وتحليل العلاقات بين الكلمات. وتشمل هذه النماذج تقنيات مثل تضمين الكلمات (Word Embeddings) والشبكات الدلاليّة.
- **تحليل المكونات الأساسية** (Principal Component Analysis – PCA): ويستخدم لتحليل البيانات وتقليل الأبعاد، مما يساعد في تحديد الأنماط والعلاقات الدلاليّة بين الكلمات.
- **الشبكات العصبية الدلالية** (Semantic Neural Networks): تستخدم لفهم العلاقات المعقّدة بين الكلمات والمعاني اللغويّة.

وتؤدي هذه العملية إلى تحسين تنظيم المعجم العربيّ: فيمكن للتصنيف الدلالي تحسين تنظيم المعجم من خلال تصنيف الكلمات بشكل أكثر منطقيّة ودقة، مما يسهل على المستخدمين العثور على الكلمات ومعانيها.

ويساعد التصنيف الدلالي في تطوير معاجم ثنائية اللغة، أو متعددة اللغات بفهم العلاقات بين الكلمات في لغات مختلفة.

ويمكن استخدام التصنيف الدلالي في تحليل النصوص الأدبية؛ لتحديد المواضيع الرئيسة، والعلاقات بين الشخصيات، والأحداث.

١- «بناء المعجم الرَّقمي العربي في ضوء اللِّسانيَّات الحاسوبيَّة» (https://www.asjp.cerist.dz/en/article/147910).

- «نظرة معمقة حول معالجة اللغة العربية الطبيعية في الذكاء الاصطناعي» https://saal.ai/ar/an-insight-to-arabic-natural-language-processing-nlp-in-ai-today/).

وتشمل تقنية تحليل السياق[1] (Context Analysis) تحليل النصوص؛ لفهم كيفيّة استخدام الكلمات في سياقات مختلفة، مما يساعد في توفير تعريفات دقيقة، وشموليّة.

وتستخدم نماذج اللغة (Language Models) مثل GPT-٤ لتحليل النصوص، وفهم السياقات المختلفة، مما يساعد في تقديم أمثلة دقيقة لاستخدام الكلمات.

وتستخدم تقنية تحليل العلاقات اللغويّة (Relation Extraction) لاستخراج العلاقات بين الكلمات في النصوص، وتحليل كيفية تفاعلها بعضها مع بعض.

وثمّة صعوبات في هذا الميدان، تتمثّل في التعامل مع التعابير المجازية التي يمكن أن تكون تحديًا لأنظمة التصنيف الدلاليّ، فتحتاج الأنظمة إلى فهم المعاني الضمنيّة، والمعقّدة.

وتحتاج الأنظمة إلى التعامل مع التنوع الثقافيّ، واللغويّ الكبيرين في النصوص، مما يتطلب مرونة وفهمًا عميقًا للمعاني المختلفة.

ومن الصعوبات الحاجة إلى بيانات عالية الجودة؛ إذ تحتاج أنظمة التعلم الآليّ إلى بيانات تدريب عالية الجودة؛ لتحقيق نتائج دقيقة. وقد يكون الحصول على هذه البيانات تحديًا في اللغة، أول اللهجة.

1- Foundations of Statistical Natural Language Processing

- لمؤلّفيه: Christopher D. Manning و Hinrich Schütze

يعد هذا الكتاب مقدمة شاملة لمعالجة اللغة الطبيعية باستخدام الأساليب الإحصائية. ويغطي موضوعات مثل نماذج اللغة، وتحليل السياق، وتمثيل الكلمات. ويركز على الأسس النظرية والتطبيقات العملية، مما يجعله مرجعًا مهمًّا للباحثين في هذا الميدان.

وقد تعاني النماذج التحيّزَ إذا كانت البيانات التدريبية غير متوازنة، أو تحتوي على تحيّزات لغويّة معينة، مما قد يؤثر في دقة النتائج.

ويمكن استخدام تقنية تحليل البيانات اللغوية[١]؛ لاستخراج الأنماط اللغويّة، واستخداماتها المختلفة في النصوص العربيّة، مما يمكِّن من تحديث المعجم العربيّ، وتطويره باستمرار.

ويمكن التعرف إلى اللهجات[٢] وذلك بتدريب نماذج للتعرف إلى اللهجات العربيّة المختلفة، وإضافتها إلى المعجم، مما يساعد في توسيع نطاق استخدامه؛ ليشمل مختلف المناطق الناطقة بالعربيّة.

٥- أهداف المعجم العربي وركائزه باستخدام الذكاء الاصطناعيّ:

يهدف المعجم العربيّ المستند إلى الذكاء الاصطناعيّ إلى ما يأتي:

- تجميع قاعدة بيانات لغويّة شاملة، بجمع كمية كبيرة من النصوص العربيّة، وتحليلها.

١- ثمّة دراسات حول كيفية استخدام الذكاء الاصطناعي لتحليل التغيرات اللغوية في اللغة العربية عبر العصور، وتوثيقها.
ومن الأمثلة على ذلك: «Tracking Linguistic Evolution in Arabic Using AI and NLP Techniques» الذي يبحث في تطور اللغة العربية، وكيفية تتبع هذه التغيرات باستخدام تقنيات الذكاء الاصطناعيّ.

٢- ثمّة أبحاث تركز على تحليل اللهجات العربية المختلفة، وفهمها باستخدام تقنيات معالجة اللغة الطبيعية.
ومن الأمثلة على ذلك: «Dialect Identification and Modeling in Arabic Natural Language Processing» الذي يناقش تحديات التعرف إلى اللهجات، والتعامل معها في النصوص العربيّة.
ونعثر على هذه الدراسات من خلال قواعد البيانات الأكاديميّة مثل Google Scholar، IEEE Xplore، SpringerLink، وغيرها من المصادر البحثية. وباستخدام الكلمات المفتاحية مثل «Arabic NLP، Arabic Text Analysis، Arabic Lexicon AI» نصل إلى الدراسات ذات الصلة.

- تصنيف المفردات[1]: بتحليل المفردات العربية بناءً على السياق والمعاني، وتصنيفها تبعًا لذلك.
- التعرف إلى التغيرات اللغويّة: أي اكتشاف التغيّرات، والتطورات في اللغة عبر الزمن.
- تحليل السياق: باستخدام الذكاء الاصطناعي؛ لفهم السياقات المختلفة للكلمات.
- توفير أدوات تعليمية: أي إنشاء أدوات تفاعليّة لتعليم اللغة العربية.

يتمّ أولًا جمع البيانات، كجمع نصوص من مصادر متعدّدة مثل الكتب، والمقالات، والصحف، والمواقع الإلكترونية. ويتم استخدام المعاجم الموجودة بوصفها بيانات أوليّة.

وبعد جمع البيانات تأتي خطوة معالجتها، ويتم ذلك بإزالة النصوص المكرّرة، وتصحيح الأخطاء الإملائيّة، والعناصر غير الضروريّة، وتقسيم النصوص إلى كلمات، وجمل مع الاحتفاظ بالسياق، وتحويل النصوص إلى صيغة قابلة للتحليل باستخدام نماذج الذكاء الاصطناعيّ. ويساعد ذلك في ضمان دقّة المعجم العربيّ، وجودته العالية.

ثمّ يمكن -تبعًا لما سبق- بناء أنموذج لغويّ، ويمكن استخدام نماذج جاهزة مثل BERT أو GPT، وهي من نماذج التعلم العميق، وتدريبها على البيانات العربيّة. ومن ثمَّ يتمّ تدريب النموذج على النصوص العربيّة لفهم العلاقات بين الكلمات، ومعانيها في السياقات المختلفة.

١- من الدراسات حول كيفية استخدام تقنيات الذكاء الاصطناعي لإنشاء معاجم رقميّة تحتوي على مفردات عربية مع تفسيراتها، وسياقاتها المختلفة.
«Building a Comprehensive Arabic Lexicon Using AI Technologies» التي تستعرض الأساليب والتقنيات المستخدمة في بناء معجم لغويّ شامل.

ثمّ تأتي مرحلة التعرّف إلى الأنماط، وذلك باستخدام خوارزميات التعلم الآليّ «Machine Learning Algorithms» للتعرف إلى الأنماط اللغويّة، وتصنيف الكلمات.

وتأتي مرحلة إنشاء قاعدة بيانات المعجم العربيّ، فتجمع الكلمات، وتعريفاتها، ومرادفاتها، وتوضع أمثلة على استخداماتها. وتصنّف الكلمات بحسب الفئات النحويّة: «اسم، فعل، صفة، ....»، فتنظّم الكلمات بحسب النوع، والجذر، والاشتقاق، وتحدّد المعاني المختلفة للكلمات بناءً على السياق الذي وردت فيه، وتوفّر أمثلة استخدام لكل كلمة؛ لتوضيح معناها في السياقات المختلفة.

وتحلّل نتائج النموذج، وتحسّن دقّته وذلك بالتغذية الراجعة[١]، وتهني استخدام ملاحظات المستخدمين؛ لتحسين المدخلات، ويحدَّث النموذج دوريًّا ببيانات جديدة؛ لتحسين أدائه.

وتصمّم واجهة المستخدِم، بشكل يسمح للمستخدمين بالبحث عن الكلمة ومعانيها بسهولة، باستخدام تطبيق ويب، أو تطبيق موبايل. وتضمن التفاعليّة التفاعل السلس مع المستخدمين، مثل الاقتراح التلقائيّ للكلمات في أثناء الكتابة.

١- التغذية الراجعة هي عملية نقل المعلومات حول نتائج أداء معين، أو تأثيراته إلى الجهة التي قامت بذلك الأداء بهدف تحسين السلوك، أو الأداء في المستقبل، أو تعديلهما. وتعدّ التغذية الراجعة جزءًا مهمًّا من عمليات التعلّم والتطوير في عدد من المجالات، مثل التعليم، والعمل، و، والإدارة. هناك نوعان رئيسان من التغذية الراجعة:

١. التغذية الراجعة الإيجابية: تعزز السلوك الجيد، أو الأداء الفعال بالتشجيع والإشادة، مما يزيد من احتمالية تكرار هذا الأداء في المستقبل.

٢. التغذية الراجعة السلبية: تهدف إلى تصحيح السلوك أو الأداء غير الفعال، أو تعديلهما بتقديم النقد البناء والنصائح، مما يساعد على تحسين الأداء في المستقبل.

ينظر في هذا الميدان:

- Hattie, J., & Timperley, H. (2007). «The Power of Feedback». Review of Educational Research, 77(1), 81-112.

ويتمّ ذلك بتطبيق يتيح للمستخدم إدخال كلمة، والحصول على تعريفها، وتصريفها، وأمثلة على استخدامها.

ويمكن تطوير واجهة تفاعليّة تمكن المستخدمين من البحث عن معاني الكلمات باستخدام الأوامر الصوتيّة. على سبيل المثال، يمكن للمستخدم أن يسأل: «ما معنى كلمة ‹ابتكار›؟» ويرد النظام الذكي بتقديم المعنى، والمرادفات الممكنة للكلمة.

ويمكن تطوير أدوات تعليميّة تفاعليّة مثل الألعاب التعليميّة، واختبارات المفردات، والمعاجم الإلكترونية.

ولتحقيق ذلك تستخدم التقنيات الآتية:

- التعلم العميق (Deep Learning) لبناء نماذج تتعلّم، وتحلّل الأنماط اللغويّة.
- معالجة اللغة الطبيعية (NLP): باستخدام مكتبات مثل NLTK وspacey، يضاف إلى ذلك مكتبات خاصة باللغة العربية مثل farasa لاستخراج النصوص، وتحليلها.
- نماذج تعلم الآلة: أي استخدام مكتبات مثل TensorFlow وPyTorch لتدريب النماذج.
- قواعد البيانات الكبيرة (Big Data): باستخدام قواعد بيانات مثل MongoDB أو PostgreSQL لتخزين البيانات والمعاني، ومعالجة كميّات ضخمة من البيانات النصيّة.
- تطوير الواجهة: استخدام HTML, CSS, JavaScript وإطارات العمل

مثل React أو Vue.js. ويمكن لتطبيقات الويب، والهاتف المحمول أن يوفرا واجهة سهلة للمستخدمين.

٦- المحاولات الجادة لبناء معجم عربيّ باستخدام الذكاء الاصطناعي:

هناك بعض المحاولات والمشاريع التي تهدف إلى استخدام التكنولوجيا، والذكاء الاصطناعيّ في هيكلة المعاجم العربية القديمة، وصياغتها. ومنها:

مشروع المعجم العربي المتكامل[١] (MIA):

هو مشروع ضخم يهدف إلى إنشاء معجم عربيّ، شامل، يشمل كلمات اللغة العربية منذ البدايات القديمة حتى العصر الحديث. يستخدم المشروع تقنيات التعلم الآليّ، والذكاء الاصطناعيّ؛ لتحليل الكلمات والتعابير العربية، وتصنيفها، ولتوليف النصوص، وإعادة صياغتها بدقّة، وموضوعيّة.

ويعدّ هذا المشروع واحدًا من أكبر المشاريع في مجال بناء المعاجم العربية باستخدام التكنولوجيا الحديثة. ومع أن النتائج المحددة لم تُنشر بعد، لكنّ الجهود المبذولة تهدف إلى توفير معجم شامل يستخدم الذكاء الاصطناعيّ؛ لتحسين دقّة المعلومات اللغويّة.

مشروع «معجم الدوحة التاريخي للغة العربية»[٢]:

هو أحد أبرز المشاريع التي تعمل على تحويل المعاجم القديمة والنصوص التاريخية إلى صيغ إلكترونية قابلة للبحث. هذا المشروع يهدف إلى تسجيل تاريخ استعمال الألفاظ العربية منذ ظهورها أول مرة، وتوثيق تحوّلاتها الدلالية والبنيوية

١- لتفاصيل أكثر عن هذا المشروع يمكن زيارة موقع مشروع المعجم العربي، والمعجم التاريخي للغة العربية.

٢- لمزيد من المعلومات حول هذا المشروع يمكن زيارة موقعه الرسمي [موقع معجم الدوحة التاريخي] (https://www.dohadictionary.org)

عبر العصور المختلفة. ويُعتبر معجم الدوحة التاريخي أحد المبادرات الرائدة في هذا الميدان، إذ يستند إلى مجموعة واسعة من المصادر العربية القديمة والحديثة، ويعتمد على تقنيات رقميّة حديثة لجعل المعاجم والنصوص أكثر فعالية وسهولة في الوصول

مشروع «نور العرب»[1] (Noor Al-Arab):

يهدف هذا المشروع إلى تطوير أدوات، وتقنيات لتحليل النصوص العربية التقليديّة، والقديمة، وفهمها بما في ذلك الشعر، والأدب، والنصوص الفلكيّة والدينيّة، باستخدام تقنيات الذكاء الاصطناعيّ؛ لتحسين جودة التحليل والتصنيف.

ولم تُعلن نتائجه بعد علنًا، ولكن النهج المتبع يظهر وعدًا كبيرًا في تطبيقات البحث الأكاديميّ، والتقنيّ.

مشروعات البحث الأكاديمية:

تقوم جامعات ومراكز بحثيّة متعدّدة بمشاريع لاستخدام الذكاء الاصطناعيّ في تحليل اللغة العربيّة القديمة، وتحديث المعاجم، مما يشمل تطبيقات متعدّدة من استخراج المصطلحات، وتصنيفها إلى تحسين الترجمة، والفهم اللغوي.

وتمثّل هذه المحاولات خطوات مهمّة نحو تحديث المعاجم العربية القديمة، وتطويرها باستخدام التكنولوجيا الحديثة، وتظهر كيفية دمج التقنيات الذكيّة في ميادين البحث اللغويّ، والثقافيّ؛ لتحقيق نتائج دقيقة، ومفيدة للمجتمع العلميّ، والعاملين في الحقل الثقافيّ، والتراثيّ.

١- للتعرف إلى المشروع يمكن العودة إلى رابطه:
https://www.tii.ae/ar/news/technology-innovation-institute-announces-launch-noor-worlds-largest-arabic-nlp-model

وتختلف نتائج المشروعات البحثيّة الأكاديميّة بحسب الهدف، والنهج المتّبع، ومعظمها لم تصل بعد إلى مرحلة الاستخدام التجاريّ، أو الانتشار الواسع. ومع ذلك، تسهم في تطوير التقنيات، والأدوات اللغويّة، والثقافيّة التي قد تؤثر بشكل إيجابي في مستقبل بناء المعاجم، وتحديثها.

وثمّة محاولات جادة[١] في سبيل المضيّ في طريق بناء هذا المعجم، ففي سابقة على مستوى المعاجم العربية أطلق مكتب الذكاء الاصطناعي[٢] في دولة الإمارات العربية «المعجم العربي للذكاء الاصطناعيّ»، مستهدفًا المهتمّين بمصطلحات الذكاء الاصطناعيّ باللغة العربية. وتسعى هذه المبادرة إلى تعزيز مكانة اللغة العربية في قطاعات الذكاء الاصطناعيّ، وتوضيح المصطلحات، وتقليل الأخطاء اللغويّة، وسدّ الفجوة بين الذكاء الاصطناعيّ واللغة العربيّة، كما تشجّع على مشاركة المجتمع باقتراح مصطلحات باللغة الإنجليزية، وسيعمل مكتب الذكاء الاصطناعي مع المتخصّصين لترجمتها، ووضع تعريفات لها، ونشرها في المعجم الرقميّ.

المعجم التاريخيّ للغة العربية في الشارقة:

ثمّة تجربة رائدة في بناء المعجم التاريخيّ للغة العربية في الشارقة[٣]، ويسهم

١- يضاف إلى هذه المشاريع بعض المشاريع البحثية مثل تلك الواردة في كتاب "اللغة العربية والذكاء الاصطناعيّ" للدكتور علاء طعيمة؛ إذ يقدّم حلولًا مبتكرة لتحديات المعالجة اللغويّة. ويتناول الكتاب تطبيقات عمليّة متعدّدة مثل تحليل المشاعر، وتوليد النصوص، وتصنيف المواضيع باستخدام تقنيات التعلم العميق، والتعلم الآلي.
وتسهم هذه الجهود في تطوير معاجم لغويّة عربيّة أكثر دقّة وشمولية، مما يفتح آفاقًا جديدة لتطبيقات الذكاء الاصطناعيّ في مجالات التعليم، والترجمة، والتفاعل البشري-الآلي. ومع استمرار التطور التكنولوجيّ، يتوقع أن يشهد هذا المجال مزيدًا من الابتكارات التي تعزّز من فهم اللغة العربية، واستخدامها في التطبيقات الحاسوبيّة المتقدمة.

2- © 2022 Artificial Intelligence, Digital Economy & Remote Work Application Office | United Arab Emirates. All Rights Reserved

٣- رابط الموقع الالكتروني للمعجم التاريخي للغة العربية بالشارقة: http://www.almojam.org/

الذكاء الاصطناعي بشكل فعّال في بنائه، ويعتمد المعجم على نظام ذكيّ يعزّز فهم اللغة العربية بشكل أعمق، ويسهم في معالجة البيانات، وتحليلها بكفاءة عالية، ويساعد الآلة على التفاعل بذكاء مع قواعد بيانات اللغة العربية.

٧– تحديات بناء معجم عربيّ شامل بالذكاء الاصطناعيّ:

يتطلب بناء معجم لغويّ شامل التعامل مع كميّات ضخمة من البيانات النصيّة، مما يستلزم بنية تحتيّة تقنية قويّة، وإدارة فعّالة للبيانات، وتحقيق مستوى عالٍ من الدقّة في تصنيف المعاني، واستخدام السياق.

وتحتاج أنظمة التعرف إلى النصوص المنطوقة إلى تحقيق دقّة عالية في التحويل؛ لضمان تحويل النصوص الصوتيّة بشكل صحيح.

وتتنوّع اللهجات العربية تنوّعًا كبيرًا، ويشكل هذا الأمر تحديًا كبيرًا في جمع النصوص، وتحليلها. وقد تكون هناك صعوبة في التمييز بين المعاني المختلفة للمفردات بناءً على اللهجة.

وقد تعتمد دقّة المعجم، وموثوقيته على جودة المصادر النصيّة المتاحة، وتنوّعها مما قد يحدّ من تغطية بعض المفردات، أو التراكيب اللغويّة النادرة.

وقد نواجه تحدّيات في تحقيق التنوّع الكامل في معاني المفردات، وتقديم أمثلة استخدام شاملة على الرغم من دقّة تقنيات الذكاء الاصطناعي

ويتطلّب معالجة كميات كبيرة من البيانات النصية التعامل مع قضايا الخصوصيّة والأمان، بخاصة إذا كانت البيانات مأخوذة من مصادر خاصّة، أو حساسة.

٨- مخرجات مشروع بناء المعجم العربيّ:

يخرج مشروع بناء معجم شامل بالاعتماد على تقنية الذكاء الاصطناعيّ بما يأتي:

- **معجم لغويّ رقميّ شامل**: يتضمن تصنيف المفردات، ومعانيها، وأمثلة استخدامها في سياقات مختلفة.
- **واجهة بحث متقدمة**: تتيح للمستخدمين البحث عن الكلمات بسهولة وفهم معانيها.

ولتحسين البحث والاستخراج عمليًا -بالاستعانة بمهندس معلوماتية- يتم اعتماد خوارزمية، وهي: نموذج BERT للبحث الدلالي، وذلك تدريب نموذج BERT على قاعدة بيانات تحتوي على نصوص عربية ومعجمية، وتطوير واجهة بحث تستخدم النموذج المدرّب لفهم استفسارات المستخدمين، وتقديم نتائج دقيقة.

- **أدوات تعليمية تفاعليّة**: لدعم تعلم اللغة العربية بطريقة مبتكرة، وفعّالة.
- **تحديثات دوريّة**[1]: لضمان مواكبة المعجم العربيّ للتطورات اللغوية الحديثة. وذلك بتقنيات مثل: التنقيب عن النصوص (Text Mining) وتشمل هذه

١- فيما يتعلق بالتنقيب عن النصوص (Text Mining) ينظر:

- Mihalis Anthopoulos, Alexandros Balkas. 2007. *Text Mining: Classification, Clustering, and Applications*. Springer. USA.

- وفيما يتعلق بالتحليل الزمني (Temporal Analysis) ينظر:

- Marion Frederick Huber. 2005. *Foundations of Time Series Analysis and Prediction Theory*. John Wiley & Sons. USA.

- وفيما يتعلّق بنظم التوصية (Recommender Systems) ينظر:

- G. Shani, A. Gunawardana. 2011. *Recommender Systems: An Introduction*. Cambridge University Press. United Kingdom.

التقنية استخراج المعلومات من النصوص الضخمة، وتحليلها؛ لاكتشاف الأنماط اللغويّة، والكلمات الجديدة.

والتحليل الزمني (Temporal Analysis) ويستخدم لتحليل التغيرات في استخدام الكلمات عبر الزمن، وتحديد الكلمات التي تصبح شائعة، أو تتراجع في الاستخدام.

نظم التوصية (Recommendation Systems) ويمكن استخدامها لتقديم اقتراحات للتحديثات اللغوية بناءً على الأنماط المكتشفة في النصوص الحديثة.

ويمكن لأنظمة الذكاء الاصطناعي تحديد الكلمات الجديدة، والشائعة بسرعة، مما يساعد في تضمينها في المعاجم؛ لتبقى محدّثة.

وباستخدام تحليل السياق يمكن تحسين الأمثلة المقدّمة لاستخدام الكلمات في المعاجم، مما يجعلها أكثر واقعيّة ودقّة.

- **تحليل الاستخدام اللغوي**: تساعد تقنيات الذكاء الاصطناعيّ في تحليل كيفية تغير استخدام الكلمات، والتعابير في مختلف المجالات، مثل الإعلام، الأدب، والعلوم.

وتواجه هذه العملية معوقات مثل: التعامل مع الكلمات المؤقتة، بعض الكلمات قد تكون شائعة لمدة قصيرة فقط. وتحتاج الأنظمة إلى القدرة على التمييز بين الكلمات التي ستكون لها ديمومة في الاستخدام، وتلك التي ستتلاشى بسرعة.

ويجب تحقيق توازن بين تحديث المعجم العربيّ باستمرار، والحفاظ على استقراره، وموثوقيّته.

- **تحليل النصوص**، باستخدام خوارزمية الشبكات العصبية المتكررة (RNN) لتحليل السياق اللغوي.

ويكون ذلك بتجميع مجموعة كبيرة من النصوص العربية الأدبيّة والعلميّة، وتدريب نموذج RNN على هذه النصوص؛ لفهم الأنماط اللغوية، واستخدام النموذج المدرّب لتحليل النصوص الجديدة، واستخراج المعلومات المعجمية مثل الجذور، والأوزان، والمعاني المختلفة للكلمات.

فلو طلبنا تحليل كلمة كتاب -مثلًا- باستخدام الذكاء الاصطناعيّ لوجدنا مرحلة البحث والاستخراج؛ إذ يبحث نموذج BERT عن «كتاب»في قاعدة بيانات المعجم العربيّ، ويعرض النتائج ذات الصلة مثل المعاني المختلفة، والاستخدامات، والأمثلة.

ثم تأتي مرحلة تحليل النصوص بنموذج RNN فيحلل نصوصًا تحتوي على كلمة «كتاب»، ويستخرج السياقات المختلفة، مثل «كتاب علمي»، «كتاب أدبي»، «كتاب تاريخي».

وتأتي مرحلة التنقيح والتحديث، بنموذج Transformer الذي يجمع البيانات الحديثة التي تحتوي على كلمة «كتاب»من الشابكة، ويقترح إضافة معانٍ جديدة، واستخدامات حديثة للمعجم.

ويمكن تحسين النماذج اللغوية بفضل قدرات الذكاء الاصطناعيّ في التعلم التلقائيّ، وتحسين الأداء التنبّؤي للتعبيرات والتراكيب اللغوية، ويسمح التعلم الآلي بالتكيف السريع مع التطورات في استخدام اللغة العربيّة، مما يضمن تحديث المعاجم، وتطويرها باستمرار لتلبية احتياجات المستخدمين الحديثة.

ويمكن استخدام تقنيات الذكاء الاصطناعيّ لتوليف النصوص القديمة،

وإعادة صياغتها بشكل يجعلها أكثر قراءة، وفهمًا للأجيال الحديثة، مع الحفاظ على الدقّة، والأصالة اللغوية.

- **تحليل المشاعر**[1] **بوساطة الذكاء الاصطناعيّ**: وهو عملية استخدام تقنيات التعلّم الآليّ، ونماذج معالجة اللغة الطبيعية (NLP) لتحديد المشاعر التي يعبر عنها النص، وفهمها. وتعتمد هذه العملية على سياق الكلام؛ لفهم ما إذا كانت المشاعر إيجابية، أو سلبية، أو محايدة.

فتُجمع النصوص التي تحتاج إلى تحليل، مثل التعليقات على وسائل التواصل الاجتماعيّ، وتقييمات العملاء، والرسائل الإلكترونية، والمقالات، وتزال العناصر غير الضروريّة مثل الرموز الخاصّة، والتكرارات، والإعلانات، لضمان جودة البيانات المستخدمة في التحليل، وتقسم كلمات النص إلى كلمات فردية (Tokenization) ويقسم النص إلى جمل لفهم السياق الكامل.

ولتحقيق ذلك يُعتمد نموذج قائم على قاموس مشاعر يحتوي على كلمات تصنّف إلى إيجابية أو سلبية، وتدريب النموذج على مجموعة من البيانات المسبقة التي تحتوي على نصوص مصنّفة، واستخدام تقنيات مثل الشبكات العصبيّة، أو شجرة القرار، أو نموذج بايز Bayes.

وتحليل النصوص في سياقها لفهم الدلالات والمعاني، فالكلمة نفسها يمكن أن تحمل مشاعر مختلفة بناءً على السياق الذي تأتي فيه. مع الأخذ بعين الاعتبار

١- هنالك فجوة بين ذكاء الإنسان وذكاء الآلة، فالآلة قادرة على القراءة، أو الكتابة، لكنها لا تستطيع محاكاة المشاعر الداخلية، كالحب والغيرة والألم والحزن، ويمكن أن تخضع الآلة إلى اكتشاف وجود بعض المشاعر من غير أن تتولد المشاعر فيها عبر قرائن لغوية، وغير لغوية قابلة للقياس.
يمكن العودة إلى الدراسة الأكاديمية لتعرّف آلية عمل تحليل المشاعر:
«From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales» by Saif M. Mohammad:

- نشر في مؤتمر Association for Computational Linguistics (ACL 2011) في يونيو ٢٠١١.

العوامل الخارجية كالزمان، والمكان، والأحداث التي تؤثر في تفسير النصوص.

ويتمّ تقييم أداء النموذج، وتحسينه باستمرار من خلال مقاييس مثل الدقّة، والاسترجاع، والدقّة المتوسطة.

ومن صعوبات العمل على تحليل المشاعر فهم التعابير المجازية والساخرة، والتعامل مع اختلافات اللهجات واللغات، وحمل الكلمة الواحدة معاني مختلفة تبعًا للسياق.

وتفيد عملية تحليل المشاعر في فهم مشاعر العملاء تجاه منتجات أو خدمات معينة، ورصد مشاعر الجمهور تجاه أحداث، أو مواضيع معينة، وتحسين الردود التلقائيّة وفهم شكاوى العملاء بعمق أكبر.

٨- خاتمة وتوصيات

إن إنشاء معجم لغوي باللغة العربية باستخدام الذكاء الاصطناعي يمكن أن يحدث ثورة في طريقة فهمنا، واستخدامنا لغتنا. وسيسهم هذا المشروع في دعم الباحثين، والمتعلمين، والمستخدمين اليوميين بتوفير أداة قوية، وشاملة؛ للتفاعل مع اللغة العربية بطريقة أكثر ذكاءً، وفعالية على الرغم من التحدّيات الموجودة.

ويمثل هذا العمل خطوة نحو دمج الذكاء الاصطناعيّ في تطوير المعاجم العربية، مما يساعد على تحسين دقتها، وشموليتها باستخدام تقنيات التعلم الآليّ، ومعالجة اللغة الطبيعية، ويمكن تحقيق تقدّم كبير في هذا الميدان، مما يسهم في الحفاظ على اللغة العربية وتطويرها؛ لمواكبة العصر الرقمي.

إن دور الذكاء الاصطناعيّ، وبخاصة التعلّم العميق، في إنشاء المعاجم يمثل نقلة نوعية في كيفية تطوير هذه الأدوات اللغوية الأساسيّة. من جمع البيانات، وتحليلها، إلى تدريب النماذج، واستخراج التعريفات، وتحديث المفردات، ويسهم الذكاء الاصطناعي في تحسين دقة المعاجم، وشموليّتها، وجعلها مواكبة التطورات اللغوية.

واتضح لنا أنّ إنشاء معجم باستخدام التعلم العميق -على سبيل المثال- يتضمن الإفادة من النماذج والتقنيات من معالجة اللغة الطبيعية (NLP) لإنشاء معجم، أو توسيعه تلقائيًا. ويتمّ ذلك بتحديد نوع المعجم الذي نريد إنشاءه، هل هو معجم أضداد، أو مرادفات، أو مصطلحات في ميدان معين، أو معجم ثنائي اللغة.

ثم تجمع البيانات النصية من مصادر متنوعة، وتحضّر البيانات للتدريب، ويقسم النصّ إلى كلمات، أو عبارات، وتحوّل الكلمات إلى أحرف صغيرة، وتزال علامات الترقيم، وغيرها، وتختزل الكلمات إلى جذرها الأساس، ويتم

اختيار نماذج التعلّم العميق المناسبة، ثم تأتي مرحلة تضمين الكلمات Word-2Vec أو Glove أو Fast Text لفهم معاني الكلمات، وعلاقاتها. ومرحلة نماذج اللغة: BERT أو GPT أو T5 لإنشاء التعريفات، والأمثلة، والترجمات.

وتنشأ النماذج المدرّبة بالاستعانة بالاختصاصيين، وتستخدم النماذج المدربة لإنشاء إدخالات المعجم، فيستخدم نموذج اللغة؛ لإنشاء تعريفات لكل كلمة، كاستخدام نماذج ثنائية اللغة، أو تقنيات الترجمة الآلية.

وتأتي عملية المراجعة البشرية للمدخلات؛ لتحسينها، وبعدها تأتي مرحلة التقييم التلقائيّ، وتوفر واجهة برمجة التطبيقات للوصول والتكامل، وتطوير تطبيق ويب أو هاتف محمول للمستخدمين النهائيين، ويتمّ تحديث المعجم العربيّ ببيانات جديدة، ويعيد تدريب النماذج بشكل دوري.

إنّ ثمّة تقدّمًا في ميدان استخدام التكنولوجيا، والذكاء الاصطناعي في بناء المعاجم العربية القديمة، وتحديثها لكنّه يعدّ ميدانًا ناشئًا، ومستمر النموّ. ومن المتوقع أن تستمرّ هذه المشاريع في تطوير النتائج، والأداء بمرور الوقت، مما يعزز دور التكنولوجيا في الحفاظ على التراث اللغويّ، والثقافيّ العربيّ.

ويخرج البحث «بناء المعجم العربي بوساطة الذكاء الاصطناعي» بالتوصيات الآتية:

- جمع البيانات اللغوية من مصادر متعددة مثل الكتب، المقالات، الإنترنت، والوثائق الأخرى، وتنظيمها؛ لتوفير قاعدة بيانات شاملة، يمكن استخدامها لتدريب نماذج الذكاء الاصطناعي.
- استخدام تقنيات التعلّم العميق؛ لتطوير نماذج قادرة على فهم اللغة العربية، وتحليلها بلهجاتها ومستوياتها المختلفة، وتدريب النماذج على كميات كبيرة

من البيانات لضمان دقة النتائج، وقدرتها على التعامل مع التنوع الكبير في اللغة العربية.

- العمل على تحسين أنظمة الترجمة الآلية؛ لتكون قادرة على تقديم ترجمات دقيقة، وفعّالة بين اللغة العربية، واللغات الأخرى.
- إنشاء أدوات تساعد في تحليل النحو، والصرف وفهم تركيب الجمل العربية بشكل صحيح، ومعالجة التحديات الخاصة باللغة العربية مثل التشكيل، والإعراب، والتفريق بين الكلمات المتشابهة في الكتابة والمختلفة في المعنى.
- تطوير تطبيقات تعتمد على الذكاء الاصطناعيّ في مجالات مختلفة مثل التعليم، والأدب، الإعلام، والأعمال، وتعزيز التعاون بين الباحثين، والمؤسسات الأكاديمية، والتجارية؛ لتبادل الخبرات، والمعرفة، وتحقيق تقدم أسرع، وأفضل في هذا الميدان.
- توفير أدوات، ومنصات مفتوحة المصدر للإسهام في تسهيل البحث، والتطوير في ميدان المعاجم العربية باستخدام الذكاء الاصطناعي، وإنشاء أنظمة؛ لتقييم الأداء بشكل دوري لضمان استمرار التحسين، والتطوير المستمر للنماذج، والأدوات المعتمدة.
- إعداد جيل من الخبراء في الذكاء الاصطناعيّ يكون على مستوى من الدراية بأسس علوم اللغة العربية، واللسانيات، والتحديات التي تواجه الطرفين. وإعداد جيل من أساتذة علوم اللغة واللسانيات العربية يكونون على دراية جيدة بمبادئ علم الحاسوب، وركائز عمله، ومبادئ الذكاء الاصطناعي، وقضايا حوسبة اللغة الطبيعية، ومعالجتها آليًا.

تلك التوصيات تساعد في توجيه البحث نحو تحقيق نتائج فعّالة، ومفيدة في مجال بناء المعجم العربي باستخدام الذكاء الاصطناعيّ.

مصادر البحث ومراجعه

- العارف، عبد الرحمن بن حسن: ٢٠٠٧، توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية «جهود ونتائج» مجلة مجمع اللغة العربية الأردني، العدد ٧٨، عمّان، كانون الأول.
- عبد النور، عادل بن عبد النور: ٢٠٠٤، مدخل إلى عالم الذكاء الاصطناعي، مدينة عبد الملك بن عبد العزيز للعلوم التقنية، المملكة العربية السعودية.
- عطية، محمد، والسعيد، المعتز عبد الله، وراغب، أحمد، وعبد الغني، نعيم: ٢٠١٩، العربية والذكاء الاصطناعي، ط١، مركز الملك عبد الله بن عبد العزيز الدولي، الرياض، المملكة العربية السعودية.
- علي، نبيل: ١٩٨٧، اللغة العربية والحاسوب، عالم الفكر، المجلد ١٨، العدد٣، الكويت.
- ماجد، أحمد: د.ت، الذكاء الاصطناعي في دولة الإمارات العربية، إدارة الدراسات والسياسات الاقتصادية، الإمارات العربية المتحدة.
- مانيارد، ديانا، بونتشيفا، كالينا، أوغنتشاين، إيزابيل: ٢٠١٩، معالجة اللغة الطبيعية للويب الدلالي، ترجمة: خالد بن عبد الرحمن الميمان، دار وجوه للنشر والتوزيع، الرياض.
- مدور، محمد، اللسانيات الحاسوبية وأثرها في تطوير اللغة العربية، أعمال ملتقى وطني حول اللغة العربية والتقانات الحديثة.
- مهديوي، عمر: ٢٠١٨، اللسانيات الحاسوبية واللغة العربية، إشكالات وحلول، ط١، دار كنوز المعرفة، عمان.
- مهديوي، عمر: ٢٠١٨، مصطلحات الهندسة اللغوية بين الترجمة والتعريب: نحو بناء معجم موحد، مجلة اللسان العربي، المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، العدد ٧٩.
- موسى، نهاد: ٢٠٠٠، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت.

- وزان، ميلاد: د. ت، التعلم العميق: المبادئ، والمفاهيم، والأساليب، ترجمة د. علاء طعيمة، نسخة إلكترونية مجانية.
- اليوبي، بلقاسم: ١٩٩٩، اللسانيات الحاسوبية: مفهومها- وتطوراتها- ومجالات تطبيقها (استشراف آفاق جديدة لخدمة اللغة العربية وثقافتها)، مجلة مكناسة، العدد ١٢، ص٤٤.

المراجع الأجنبية:

- Artificial Intelligence, Digital Economy & Remote Work Application Office | United Arab Emirates. All Rights Reserved.
- Lane Hobson, Hapke Hannes, et al. 2019. Natural Language Processing in Action: Understanding, Analysing, and Generating Text with Python. USA: Manning Publications.
- Tobias Hollerer, Dieter Schmalstieg. 2015. Augmented Reality: Principles and Practice. Addison-Wesley Professional.
- Nitin Indurkhya, Fred J. Damerau, eds. 2010. Handbook of Natural Language Processing. CRC Press.
- Christopher D. Manning, Hinrich Schütze. 1999.. Foundations of Statistical Natural Language Processing. MIT Press.
- Hattie, J., & Timperley, H. 2007. «The Power of Feedback». Review of Educational Research., 77(1).
- Saif M. Mohammad. 2011. «From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales». Published in .ACL 2011 Conference (Association for Computational Linguistics) in June 2011.
- Mihalis Anthopoulos, Alexandros Balkas. 2007. Text Mining: Classification, Clustering, and Applications. Springer. USA.
- Marion Frederick Huber. 2005. Foundations of Time Series Analysis and Prediction Theory. John Wiley & Sons. USA.
- G. Shani, A. Gunawardana. 2011. Recommender Systems: An Introduction. Cambridge University Press. United Kingdom.

المواقع الإلكترونية:

- الذكاء الاصطناعي واللغة العربية: أهم التحديات وتقنيات معالجتها:
(https://technologyreview.ae)
- أفضل أدوات الذكاء الاصطناعي التوليدي باللغة العربية:
(https://technologyreview.ae)
- د. نبيل طعيمة، كتاب اللغة العربية والذكاء الاصطناعي:
(https://dlarabic.com)
- رابط الموقع الإلكتروني للمعجم التاريخي للغة العربية بالشارقة:
(http://www.almojam.org)
- رابط الموقع الإلكتروني لمعجم الدوحة التاريخي للغة العربية:
(https://www.dohadictionary.org)
- رابط مشروع نور العرب الإلكتروني:
(https://www.tii.ae/ar/news/technology-innovation-institute-announces-launch-noor-worlds-largest-arabic-nlp-model)
- مقال «بناء المعجم الرَّقمي العربي في ضوء اللِّسانيَّات الحاسوبيَّة»د. محمد يزيد سالم:
(https://www.asjp.cerist.dz/en/article/147910)
- مقال «نظرة معمقة حول معالجة اللغة العربية الطبيعية في الذكاء الاصطناعي»:
https://saal.ai/ar/an-insight-to-arabic-natural-language-processing-nlp-in-ai-today/

المراجع العربية (Romanized References):

- Al-ʿĀrif, ʿAbd al-Raḥmān ibn Ḥasan. 2007. Tawẓīf al-Lisāniyyāt al-Ḥāsūbiyyah fī Khidmat al-Dirāsāt al-Lughawiyyah al-ʿArabiyyah: «Juhūd wa-Natāʾij», Majallat Majmaʿ al-Lugha al-ʿArabiyyah al-Urdunī, ʿAdad 78, ʿAmmān, Kānūn al-Awwal.
- ʿAbd al-Nūr, ʿĀdil ibn ʿAbd al-Nūr. 2004. Madkhal ilā ʿĀlam al-Dhikāʾ al-Iṣṭināʿī, Madīnat ʿAbd al-Malik ibn ʿAbd al-ʿAzīz lil-ʿUlūm wa-l-Taqniyyah, al-Mamlakah al-ʿArabiyyah al-Suʿūdiyyah.
- ʿAṭiyyah, Muḥammad, al-Saʿīd, al-Muʿtaz ʿAbd Allāh, Rāghib, Aḥmad, wa-ʿAbd al-Ghanī, Naʿīm. 2019. Al-ʿArabiyyah wa-l-Dhikāʾ al-Iṣṭināʿī, Ṭ1, Markaz al-Malik ʿAbd Allāh ibn ʿAbd al-ʿAzīz al-Dawlī, al-Riyāḍ, al-Mamlakah al-ʿArabiyyah al-Suʿūdiyyah.
- ʿAlī, Nabīl. 1987. Al-Lugha al-ʿArabiyyah wa-l-Ḥāsūb, ʿĀlam al-Fikr, al-Mujallad 18, al-ʿAdad 3, al-Kuwayt.
- Mājid, Aḥmad. n.d. Al-Dhikāʾ al-Iṣṭināʿī fī Dawlat al-Imārāt al-ʿArabiyyah, Idārat al-Dirāsāt wa-l-Siyāsāt al-Iqtiṣādiyyah, al-Imārāt al-ʿArabiyyah al-Muttaḥidah.
- Māniyārd, Dīyānā, Būnshīfā, Kālīnā, Ūghānshāyn, Īzābīl. 2019. Muʿālajat al-Lugha al-Ṭabīʿiyyah lil-Wīb al-Dalālī, Tarjamah: Khālid ibn ʿAbd al-Raḥmān al-Mīmān, Dār Wujūh lil-Nashr wa-l-Tawzīʿ, al-Riyāḍ.
- Madūr, Muḥammad. Al-Lisāniyyāt al-Ḥāsūbiyyah wa-Atharuhā fī Taṭwīr al-Lugha al-ʿArabiyyah, Aʿmāl Multaqā Waṭanī Ḥawla al-Lugha al-ʿArabiyyah wa-l-Taqānāt al-Ḥadīthah.
- Mahdīwī, ʿUmar. 2018. Al-Lisāniyyāt al-Ḥāsūbiyyah wa-l-Lugha al-ʿArabiyyah: Ishkālāt wa-Ḥulūl, Ṭ1, Dār Kunūz al-Maʿrifah, ʿAmmān.
- Mahdīwī, ʿUmar. 2018. Muṣṭalaḥāt al-Handasah al-Lughawiyyah bayna al-Tarjamah wa-l-Taʿrīb: Nḥwa Bināʾ Maʿjam Muwaḥḥad, Majallat al-Lisān al-ʿArabī, al-Munẓamah al-ʿArabiyyah lil-Tarbiyyah wa-l-Thaqāfah wa-l-ʿUlūm, Maktab Tansīq al-Taʿrīb, al-ʿAdad 79.
- Mūsā, Nihād. 2000. Al-ʿArabiyyah: Nḥwa Twaṣīf Jadīd fī Ḍawʾ al-Lisāniyyāt al-Ḥāsūbiyyah, al-Muʾassasah al-ʿArabiyyah lil-Dirāsāt wa-l-Nashr, Bayrūt.
- Wazzān, Mīlād. n.d. Al-Taʿallum al-ʿAmīq: Al-Mabādiʾ, wa-l-Mafāhīm, wa-l-Asālīb, Tarjamah: Dr. ʿAlāʾ Ṭuʿaymah, Nuskhah Elektrūniyyah Majjāniyyah.
- al-Yūbī, Bilqāsim. 1999. Al-Lisāniyyāt al-Ḥāsūbiyyah: Mafhūmahā - wa-Taṭawwurātuhā - wa-Majālāt Taṭbīqihā (Istishrāf Āfāq Jadīdah li-Khidmat al-Lugha al-ʿArabiyyah wa-Thaqāfatihā), Majallat Maknāsah, al-ʿAdad 12, ṣ44.

المراجع الإلكترونية (Romanized Websites References)

- Al-Dhikāʾ al-Iṣṭināʿī wa-l-Lugha al-ʿArabiyyah: Aham al-Taḥaddiyāt wa-Taqnīyāt Muʿāljatihā. Available at: https://technologyreview.ae
- Afḍal Adawāt al-Dhikāʾ al-Iṣṭināʿī al-Tūljīdī bi-l-Lugha al-ʿArabiyyah. Available at: https://technologyreview.ae
- Ṭuʿaymah, Nabīl. Kitāb al-Lugha al-ʿArabiyyah wa-l-Dhikāʾ al-Iṣṭināʿī. Available at: https://dlarabic.com
- Al-Muʿjam al-Tārīkhī lil-Lugha al-ʿArabiyyah - Sharjah. Available at: http://www.almojam.org
- Al-Muʿjam al-Tārīkhī lil-Lugha al-ʿArabiyyah - Dūḥah. Available at: https://www.dohadictionary.org
- Mashrūʿ Nūr al-ʿArab al-Iliktrūnī. Available at: https://www.tii.ae/ar/news/technology-innovation-institute-announces-launch-noor-worlds-largest-arabic-nlp-model
- Sālim, Muḥammad Yazīd. Bināʾ al-Muʿjam al-Raqmī al-ʿArabī fī Ḍawʾ al-Lisāniyyāt al-Ḥāsūbiyyah. Available at: https://www.asjp.cerist.dz/en/article/147910
- Naẓrah Maʿmīqah Ḥawla Muʿālajat al-Lugha al-ʿArabiyyah al-Ṭabīʿiyyah fī al-Dhikāʾ al-Iṣṭināʿī. Available at: https://saal.ai/ar/an-insight-to-arabic-natural-language-processing-nlp-in-ai-tod