معجم الدوحة التاريخي
للغــة العربيــة

حفل
اكتمال معجم الدوحة التاريخي للغة العربية
ومؤتمر
الذكاء الاصطناعي وخصائص اللغة العربية

22 – 23 كانون الأول/ ديسمبر 2025
الدوحة، قطر

برنامج الحفل | عن المؤتمر | جدول الأعمال | المشاركون | الملخصات | رؤساء الجلسات

حفل

اكتمال معجم الدوحة التاريخي للغة العربية

ومؤتمر

الذكاء الاصطناعي وخصائص اللغة العربية

22 – 23 كانون الأول/ ديسمبر 2025

الدوحة، قطر

برنامج الحفل | عن المؤتمر | جدول الأعمال | المشاركون | الملخصات | رؤساء الجلسات

برنامج الحفل

برنامج حفل اكتمال معجم الدوحة التاريخي للغة العربية

الإثنين 22 كانون الأول / ديسمبر 2025

9:00-8:00	التسجيل
10:30-9:00	حفل اكتمال معجم الدوحة التاريخي للغة العربية
	النشيد الوطني لدولة قطر تلاوة آيات بينات من الذكر الحكيم فيلم قصير عن معجم الدوحة التاريخي للغة العربية كلمات افتتاحية عرض بوابة المعجم المحدّثة تكريم العاملين في المعجم
	حفل استقبال

الورقة المرجعيّة للمؤتمر

شهد العالم في العقود الأخيرة ثورة تقنية غير مسبوقة، مدفوعةً بالتطور السريع في مجالات الحوسبة وتكنولوجيا المعلومات، والتي بلغت ذروتها مع ظهور الذكاء الاصطناعي التوليدي (Generative AI). وقد شكّل الذكاء الاصطناعي في عمومه نقلة نوعية في التعامل مع اللغات الطبيعية، إنتاجا وفهما وتحليلا وترجمة؛ حيث ركزت التقنيات عبر مسارات تطورها على جوانب لغوية مختلفة شملت الترجمة الآلية، والتعرف على الكلام، وتحليل النصوص، وتوليد المحتوى، إلخ. ومع ظهور خوارزميات التعلم العميق (Deep Learning) وتطور الشبكات العصبية الاصطناعية (Neural Networks)، برزت قدرة هذه التقنيات على معالجة اللغات بشكل أكثر كفاءة مما سبق. ولم يكن هذا التطور وليد اللحظة، بل كان تتويجا لمسيرة طويلة من البحث والتطوير بدأت جذورها في منتصف القرن العشرين، وتجلت ملامحها مع نشر آلان تورينج Alan Turing مقالته الشهيرة (Computing Machinery and Intelligence) "الآلات الحاسبة والذكاء"، والتي قدم فيها "اختبار تورينج" (Turing Test) معيارًا لتحديد ما إذا كان يمكن اعتبار الآلة ذكية بناء على قدرتها على محاكاة المحادثة البشرية. وهذا الاختبار لم يضع فقط معيارًا للذكاء الاصطناعي، بل سلط الضوء أيضًا على أهمية اللغة باعتبارها واجهة للتفاعل بين الإنسان والآلة.

وبشكل عام، ترتبط اللغة والتقنية بعلاقة وثيقة ومتبادلة، فمنذ اختراع الكتابة مرورًا بابتكار الطباعة وحتى الوصول إلى العصر الرقمي، كانت اللغة حاضرة بقوة أداةً وموضوعًا للتطور التقني. وفي عصر الذكاء الاصطناعي، برزت تقنيات معالجة اللغات الطبيعية (Natural Language Processing) فرعًا أساسيًّا وأصيلًا من فروع الذكاء الاصطناعي يركز على تمكين الحاسوب من فهم اللغة والتعاطي معها، مما أدى إلى ظهور النماذج اللغوية التوليدية (Generative Language Models) التي تعتمد على بيانات لغوية ضخمة لفهم النصوص البشرية وتوليدها. وقد سلك التطور في مجال الذكاء الاصطناعي اللغوي مسارات متعددة، بدايةً من النماذج الرمزية القديمة إلى النماذج الإحصائية، وصولًا إلى النماذج القائمة على التعلم العميق مثل GPT وغيره.

وقد استفادت اللغة الإنكليزية وعدد من اللغات الكبرى من هذا التطور، مدعومة بمواردها الرقمية، ومعاجمها التاريخية، وتوصيفاتها اللغوية المعيارية، ومراكزها البحثية المتخصصة. وقد نتج عن ذلك ما نراه اليوم من تطور كبير ووفرة في الأدوات والموارد المتخصصة التي شكلت رافعة قوية ومُسرِّعا فعّالا في بناء النماذج اللغوية الضخمة مثل (ChatGPT Geminig Claudeg، وغيرها) وتعزيز كفاءتها؛

حيث يبدو الفرق جليا من حيث الكفاءة بين تعاطيها مع اللغة العربية وهذه اللغات، ليس فقط على مستوى الفهم والتحليل والتوليد والترجمة، وإنما أيضا على مستوى الوعي الثقافي والتحيز الفكري.

لقد أحدَث الذكاء الاصطناعي والنماذج اللغوية الضخمة بذلك تحولاً في إنتاج المعرفة وتوزيعها، ويعيدان تشكيل المشهد المعرفي على المستوى العالمي. وتأتي هذه التطورات مشفوعة بتحيزات لغوية ومعرفية وثقافية كامنة، تَظهر بوضوح خاص في اللغات والثقافات ذات التمثيل الرقمي المحدود، ومنها اللغة العربية التي تواجه تحدّيات عدّة. وتشمل هذه التحديات صعوبات التجزيء (Tokenization)، وتعقيد البنى اللغوية، ونقص البيانات، ومحدودية الموارد الحاسوبية.

ونتيجة لذلك، تواجه الثقافة العربية خطر أن تصبح ضعيفة التمثيل أو أن تُصوّر بصورة غير دقيقة، مما يزيد من التحيزات الثقافية وعدم المساواة في منظومات المعرفة الرقمية. والتصدي لهذه التحديات لا يقتصر على الجوانب التقنية فحسب، بل يبدأ بتمثيل الخصائص اللغوية ويرتبط ارتباطًا عميقًا بالحفاظ على الثقافة، وتعزيز الهوية، وينتهي بالتنمية الاجتماعية والاقتصادية.

والواقع أن للغة العربية في هذا السياق أهمية خاصة؛ فهي لغة عريقة، تمتلك تاريخا غنيًّا وحضارة ثرية ممتدة لأكثر من عشرين قرنا من الزمن؛ وكانت دائما الوعاء الذي نقل وحفظ المعارف الإنسانية في مختلف ميادين العلوم والفلسفة والأدب عبر العصور. ورغم هذه المكانة الحضارية والتاريخية، فإن نصيب اللغة العربية من التطورات التقنية الراهنة، وخاصة في مجال الذكاء الاصطناعي، لا يزال محدودا نسبيا مقارنة باللغات العالمية الكبرى. ويُعزى ذلك إلى عدة عوامل، من أبرزها ندرة الموارد اللغوية الرقمية عالية الجودة، وتنوع المستويات اللغوية المنطوقة والمكتوبة (الفصحى، المعاصرة، واللهجات المحلية)، والمميزات النحوية والصرفية الفريدة، إضافة إلى ضعف الاستثمارات المالية والبشرية في مجال المعالجة اللغوية العربية الآلية بشكل عام.

وانطلاقًا من ذلك، يبدو مشروعًا أن يتولى هذا المؤتمر دعوة الباحثين للإسهام في معالجة الأسئلة الآتية:

1. ما الذي جعل اللغة الإنكليزية -ومثيلاتها- ذات نجاعة عالية في استخدامات الذكاء الاصطناعي؟
2. ما الأسباب التي جعلت لغات أخرى، ومنها اللغة العربية، أقل نجاعة في استخدامات الذكاء الاصطناعي؟

3. كيف تعاملت اللغات المماثلة للغة العربية مع ضعف النجاعة في استخدامات الذكاء الاصطناعي؟
4. هل حقًّا لخصائص اللغة العربية، واللغات المماثلة لها، دورٌ في ضعف النجاعة تلك؟
5. كيف يمكن اللحاق بمستوى اللغة الإنكليزية، ومثيلاتها، في استخدامات الذكاء الاصطناعي؟

لا شكّ أنّ الاهتمام يتزايد في عالمنا العربي والجهود تتضاعف لتجاوز العقبات وتعزيز مكانة اللغة العربية في الفضاء الرقمي وإنتاج الموارد اللغوية اللازمة لذلك. وقد برزت مؤخرا مساهمات عربية مهمة في مجال إنتاج النماذج اللغوية الكبرى مثل نموذج "فالكون 40ب" (Falcon 40B) في الإمارات العربية المتحدة، ونموذج "علام" في المملكة العربية السعودية، ونموذج "فنار" في قطر، ونموذج "دلة " (DALLA) الذي يسعى المركز العربي للأبحاث ودراسة السياسات بالدوحة لتطويره، مستثمرًا البيانات التي يوفّرها مشروعه الرائد "معجم الدوحة التاريخي للغة العربية" ومدونته اللغوية المهيكلة والموثقة والمؤرخة.

وعلى الرغم من أهمية هذه المبادرات العربية إلا أنها لا زالت تواجه تحديات يقتضي تجاوزها جهودًا بحثية وتطويرية مكثفة وشراكات قوية مع إيمان راسخ بأن الخصائص الفريدة للغة العربية، كثرائها الاشتقاقي ومرونتها البلاغية، تقدم فرصًا واعدة لتطوير نماذج ذكاء اصطناعي أكثر دقة وفعالية. ولعلّه من الضروري اعتماد نهج موحد بين الأطراف المعنية في العالم العربي (الدول والمؤسسات) لبناء قواعد بيانات شاملة ومستنيرة ثقافيًا. وسيكون الحوار المركّز والتعاون الاستراتيجي ضروريين لحماية اللغة والثقافة العربية في العصر الرقمي، وضمان المشاركة العادلة في صياغة المستقبل المدفوع بالذكاء الاصطناعي.

وفي هذا السياق، يسعى هذا المؤتمر إلى استكشاف العلاقة المركبة بين خصائص العربية والنماذج اللغوية الاصطناعية، سواء من حيث التحديات أو الإمكانات، في أفق بناء منظومة معرفية عربية متكاملة في مجال الذكاء الاصطناعي اللغوي. وهو يهدف إلى أن يكون منصة لتبادل الخبرات والأفكار بين الباحثين والمطورين والمهتمين، لتعميق البحث في الموضوع، والخروج باقتراحات وتوصيات تسهم في تسريع وتيرة التقدم في هذا المجال الحيوي.

جدول الأعمال

اليوم الأول: الإثنين 22 كانون الأول / ديسمبر 2025

11:40-11:30	**الجلسة الافتتاحية** **رئيس الجلسة: عبدالمنعم حرفان**
1:20-11:40	**جلسة خاصة عن معجم الدوحة التاريخي للغة العربية** **رئيس الجلسة: عبد العلي الودغيري** **عبد الكريم جبل:** فقه الدّلالة وتحولات المعنى: نقض الأوهام الشائعة في الفهم اللغوي ببرهان الشاهد التاريخي في معجم الدوحة **محمد الخطيب:** بنية الشاهد ومنهج التوثيق: معايير صيانة النص في معجم الدوحة **رشيد بلحبيب:** من الاستمداد إلى الإمداد: قضايا الإحياء والتفصيح في ضوء معجم الدوحة **حسن حمزة:** معجم الدوحة والمعجم الفرنسي: رؤية معجمية مقارنة في تأريخ اللغة
3:00-1:20	**الجلسة الأولى: بيانات معجم الدوحة التاريخي للغة العربية في خدمة الذكاء الاصطناعي** **رئيس الجلسة: محمد محمود أحمد محجوب** **عز الدين البوشيخي ومحمد بباه ويحيى الحاج ومحمد رقاس** بيانات معجم الدوحة التاريخي للغة العربية موردا لتطوير تطبيقات الذكاء الاصطناعي العربي **محمد العبيدي:** تعزيز المحتوى الرقمي العربي بالموارد المعجمية في عصر الذكاء الاصطناعي: معجم الدوحة التاريخي للغة العربية أنموذجا **سامر الرشواني وسمية الطنبولي:** تطوير خادم للذكاء الاصطناعي معزز بمعجم الدوحة التاريخي للغة العربية : من أجل فهم أدق لنصوص القرآن والحديث النبوي **أيمن بن نجي وحسن كيسان:** تقييم المعرفة المعجمية في نماذج الذكاء الاصطناعي العربية: دراسة في ضوء بيانات معجم الدوحة التاريخي للغة العربية
3:00	غداء

6:10-4:30 (جلسة موازية)	**الجلسة الثانية: الخصائص اللسانية للعربية وعلاقتها بالنمذجة الآلية** **رئيس الجلسة: مهدي عرار** **محمد غاليم:** عن الذكاء الاصطناعي وخصائص اللغة العربية في الإطار المعرفي **محمد أمين:** الخصائص التركيبية للغة العربية ومعالجتها الآلية **محمد الغازي:** نمذجة التعقيد الصرفي-التركيبي في الذكاء الاصطناعي العربي **محمد ابن سفاج:** العدول عن الرتبة الأصلية للكلمة وتأثيره في عملية الشكل الآلي عبر نماذج الذكاء الاصطناعي العربية (قلم ولسان وفَراسة نماذج)
6:10-4:30 (جلسة موازية)	**الجلسة الثالثة: تطبيقات عربية في أداء نماذج الذكاء الاصطناعي** **رئيس الجلسة: محمد الشيباني** **لمياء هَدْرِيش بلغيث:** تحليل الآراء في العربية الدارجة في عصر الذكاء الاصطناعي: بين التعقيد اللغوي والتحديات التقنية **سليمان بن سالم الشهري:** أداء النماذج اللغوية التوليدية في تمثيل الأمثال الشعبية الخليجية: قراءة تحليلية في المعنى والأسلوب والهوية. **نصر الدين مزاري:** تأثير الثنائية اللغوية (الفصحى والعامي) على أداء نماذج الذكاء الاصطناعي في فهم اللغة العربية: دراسة تجريبية مقارنة بين (CAMeLBERT)، (AraBERT)، و (GPT-3.5) **عبد الحميد صبار:** تحليل المشاعر للهجات العربية: دراسة علمية تجريبية

اليوم الثاني: الثلاثاء 23 كانون الأول / ديسمبر 2025

10:40-9:00	**الجلسة الرابعة: المحاولات العربية في بناء النماذج اللغوية (1)** **رئيس الجلسة: يحيى الحاج** **هند الخليفة:** النمذجة اللغوية الآلية للعربية: دراسة تحليلية شاملة لنماذج اللغة الكبيرة والتحديات اللسانية **أحمد عبد العلي:** المشهد البياني العربي: مسح شامل لمساهمات القطاعات في التدريب المسبق للنماذج اللغوية الكبيرة **رفيف السيد:** نحو تعزيز الذكاء الاصطناعي للغة العربية: رؤية من منظور السياسات **كريم بوزبع:** هل نحتاج إلى نماذج لغوية عربية ضخمة؟ وإن كان الأمر كذلك، فكيف وبأي شكل؟
10:50-10:40	استراحة
12:30-10:50 (جلسة موازية)	**الجلسة الخامسة: المحاولات العربية في بناء النماذج اللغوية (2)** **رئيس الجلسة: محمد باباه** **عز الدين مزروعي:** تعزيز معالجة اللغة العربية: منصة الخليل لأدوات التحليل الصرفي-النحوي وذخائر النصوص المعنونة مفتوحة المصدر **ربيع أمهز ورزان حلبي:** يليل: منصة مفتوحة للتدقيق في اصطفاف النصوص والكلام باللغة العربية من أجل تقييم تطبيقات الذكاء الاصطناعي **فادي زراقط:** تحسين أداء النماذج اللغوية الضخمة وكلف تدريبها وتشغيلها للغة العربية باستخدام تقنية إعادة استعمال الرموز **محمود الحاج:** عرب-جوبز: مدونة متعددة الجنسيات لإعلانات الوظائف العربية

12:30-10:50 (جلسة موازية)	**الجلسة السادسة: توظيف الذكاء الاصطناعي في تعليم اللغة العربية** **رئيس الجلسة: خالد الجبر** **إدريس الشرقاوي:** خصائص اللغة العربية وتوظيف تطبيقات الذكاء الاصطناعي في تعليمها للناطقين بغيرها: الميزات والتحديات والحلول **محمد اسماعيلي علوي:** توظيف الذكاء الاصطناعي في تحليل وإعداد الموارد والبيانات اللغوية الموجهة إلى تعليم العربية للناطقين بغيرها: إشكالات منهجية وتطبيقات تربوية **محمد حاج إبراهيم ومجدي حاج إبراهيم:** تعليم اللغة العربية والذكاء الاصطناعي في ضوء السياسات التعليمية في ماليزيا **عبد الله يوسفي:** معالجة الأخطاء التركيبية باستعمال النماذج اللغوية الضخمة
12:40-12:30	استراحة
2:20-12:40 (جلسة موازية)	**الجلسة السابعة: الترجمة إلى العربية في نماذج الذكاء الاصطناعي** **رئيس الجلسة: محمد رقاس** **نضال شمعون:** ثورة الذكاء الاصطناعي في نقل المعرفة، وخصوصية اللغة العربيّة في الترجمة العلمية **الطيب دبة:** محاولة في توصيف عمل العلامات والقرائن في نظام النحو العربي خلفية لسانية ضرورية لمواجهة مشكلات الترجمة الآلية **مراد دياني:** من المطبوع إلى الخوارزمية: الكتابة الرقمية المعززة بالذكاء الاصطناعي ووعود إعادة كتابة الإنتاج المعرفي العربي **إسراء محيسن:** قدرة نماذج اللغة الكبيرة على ترجمة الأمثال العربية إلى الإنجليزيّة (ChatGPT) و(Qwen) أنموذجا

2:20-12:40 (جلسة موازية)	**الجلسة الثامنة: السياقات الثقافية والمعرفية في نماذج الذكاء الاصطناعي** **رئيس الجلسة: حسين الزراعي** **فوزي حراق:** تحديات الخصوصية والأمان في استخدام النماذج اللغوية الضخمة لمعالجة بيانات عربية حساسة **سعد عبد الغفار:** تحليل الانزياحات الدلالية الزمنية للمفاهيم الثقافية في اللغة العربية باستخدام الذكاء الاصطناعي: دراسة تطبيقية على معجم الدوحة التاريخي **بشير نصري :** السياقات الثقافية والمعرفية في نماذج الذكاء الاصطناعي العربية: مقاربة سوسيولوجية ومعرفية
2:30	غداء

المشاركون

الملخصات

إدريس شرقاوي

أستاذ مشارك بكلية اللغات والآداب الحديثة في جامعة ويليام وماري بولاية فيرجينيا، الولايات المتحدة الأمريكية. شغل سابقًا منصب مدير قسم اللغة العربية بالجامعة، كما تولّى الإدارة الأكاديمية لبرنامجي CASA وCLS في المغرب. تتركز اهتماماته البحثية حول الترجمة واللسانيات التطبيقية وتعليم العربية للناطقين بغيرها. من أبرز أعماله: دراسة «استراتيجيات تدريس المكوّن الثقافي للناطقين بغير العربية» ضمن كتاب «اللسانيات التطبيقية وتدريسية العربية» (منشورات المركز المغربي للبحث العلمي وتطوير الكفاءات، دار القلم)، وباللغة الإنجليزية بحث بعنوان: "The Characteristics of The Moroccan Dialect Used in Traditional Oral Tales" المنشور في مجلة Al-Andalus Magreb، جامعة قادش، إسبانيا، العدد 26، سنة 2020.

خصائص اللغة العربية وتوظيف تطبيقات الذكاء الاصطناعي في تعليمها للناطقين بغيرها: الميزات والتحديات والحلول

يحظى الذكاء الاصطناعي باهتمام متزايد في التعليم، إلا أن توظيفه في تعليم العربية للناطقين بغيرها يواجه تحديات ناتجة عن خصائص اللغة العربية وصعوبات معالجتها حاسوبيًّا، خصوصًا في مجال معالجة اللغات الطبيعية (NLP). كما لا تزال الدراسات تفتقر إلى تناول عملي لكيفية تجاوز هذه العقبات عبر منصات الذكاء الاصطناعي. تهدف هذه الورقة إلى إبراز دور الذكاء الاصطناعي في تسهيل تعليم العربية للناطقين بغيرها، مع تقديم وصف لأهم الخصائص اللغوية التي تعيق تطوير التطبيقات الرقمية، واقتراح حلول تعليمية أكثر فاعلية. وبالاعتماد على المنهج الوصفي التحليلي، تتناول الدراسة تطبيقات تعليمية مثل: Padlet Edpuzzleg Quizletg Kahootg Socrativeg، وتبحث في دورها بتمكين المتعلمين من المهارات اللغوية الأربع. وخلصت النتائج إلى أن المنصات المدعومة بالذكاء الاصطناعي قادرة على دعم تعلم اللغة بفاعلية، رغم استمرار التحديات المرتبطة بالمعالجة الآلية للعربية، وأن توظيف الأدوات الحالية بشكل استراتيجي يسهم في تطوير أساليب تعليم أكثر كفاءة. كما تؤكد الدراسة الحاجة إلى تطوير تقنيات الذكاء الاصطناعي مستقبلاً اعتمادًا على فهم أعمق لعلوم اللسانيات العربية لتجاوز ما تبقى من عقبات.

إسراء المحيسن

مدرسة في مركز اللغات في الجامعة الأردنية منذ 2018 – رئيسة شعبة اللغة الإنجليزيّة سابقًا، تحمل درجة الماجستير في الترجمة من جامعة اليرموك وهي حاليًا طالبة دكتوراة في اللغويات والترجمة في جامعة القاهرة في مرحلة إعداد الرسالة. ولديها أبحاث منشورة في الترجمة مثال ذلك: إشكاليات في ترجمة النصوص التقنيّة والترجمة والأسلوبية "رواية عصفور من الشرق أنموذجًا" وغيرها. ومجالات اهتمام الباحثة هي الترجمة والذكاء الاصطناعي واللغويات الاجتماعية.

قدرة نماذج اللغة الكبيرة على ترجمة الأمثال العربيّة إلى الإنجليزيّة: ChatGPT Qweng أنموذجًا

تتناول هذه الدراسة قدرة النماذج اللغوية الكبيرة (LLMs) مثل ChatGPT-5 من شركة OpenAI وQwen-Max 3 من شركة Alibaba على ترجمة الأمثال العربية إلى الإنجليزية دون فقدان دلالتها الثقافية، فالأمثال تعبيرات ثقافية تجسد حكمة المجتمعات وقيمها وخبراتها وتجاربها، مما يجعل من ترجمتها تحديًا. اعتمدت الدراسة المنهج التجريبيّ من خلال إنشاء مجموعة تضم 30 مثلًا (15 بالفصحى و15 بالعامية) تُرجمت بواسطة النموذجين ثم تمّ مقارنتها بالترجمة البشرية. أظهرت النتائج أن ChatGPT كان أكثر دقة في ترجمة الأمثال الفصحى، إذ قدّم مرادفات لفظيّة تُحافظ غالبًا على المعنى الثقافي المقصود، بينما تراجعت دقته في ترجمة الأمثال العامية لاعتماده ترجمات حرفية أو تعبيرات غير مناسبة. أما Qwen فقد واجه صعوبة مع الأمثال الفصحى، لكنه أظهر أداءً أفضل نسبيًا مع العامية. وتؤكد النتائج أهمية دمج الموارد الثقافية واللغوية في تطوير هذه النماذج لتحسين أدائها في الترجمة عبر الثقافات.

عالم أبحاث ومهندس متخصص في معالجة اللغات الطبيعية والتعلم الآلي مع التركيز على التطبيقات المتعلقة باللغة العربية ولهجاتها. يعمل حاليًا ضمن فريق نموذج ALLaM. في السابق، عمل لأكثر من عقد من الزمن في معهد قطر لبحوث الحوسبة. وقبل ذلك، أمضى 13 عامًا في أدوار بحثية متعددة في جامعة ولاية نيو مكسيكو. قام بتأليف والمساهمة في نشر أكثر من 100 ورقة بحثية ومقال علمي محكم.

المشهد البياني العربي: مسح شامل لمساهمات القطاعات في التدريب المسبق للنماذج اللغوية الكبيرة

أدى التقدم في النماذج اللغوية الكبيرة إلى إحداث تحولات جذرية في العديد من المجالات، إلا أن فعاليتها تعتمد بشكل كبير على جودة وتنوع بيانات التدريب. تتميز اللغة العربية بتركيب صرفي معقد وازدواجية لسانية بين العربية المعيارية واللهجات المتعددة، مما يجعل بناء ذخائر بيانات ممثلة مهمة معقدة. يهدف هذا البحث إلى تقديم مسح شامل للقطاعات التي تساهم في توفير بيانات عالية الجودة لتدريب النماذج الكبيرة، مع التركيز على سبعة قطاعات رئيسة: الشبكة العامة، والإعلام والصحافة، الأكاديمية والبحثية، والأدب والكتب، ووسائل التواصل الاجتماعي، والقطاعان القانوني والحكومي، والتعليم. تشير الاتجاهات الحديثة إلى تحول نحو إنشاء مجموعات بيانات اصطناعية لسد الفجوة في الاحتياجات اللغوية للمجالات ذات الموارد المحدودة، باستخدام الترجمة أو المبادرات البحثية. ويخلص البحث إلى أن بناء نماذج كبيرة يتطلب نهجًا استراتيجيًا متعدد القطاعات، يعتمد على الاستفادة من بيانات الشبكة العنكبوتية والإعلام والصحافة لتعزيز المعرفة الأساسية باللغة العربية، ومعالجة بيانات وسائل التواصل الاجتماعي لإثراء فهم اللهجات، واستخدام المجموعات الأكاديمية والأدبية لإثراء العمق اللغوي والثقافي."

دكتوراه في التخطيط والسياسة اللغوية، وخبير لغوي بمعجم الدوحة التاريخي للغة العربية، حائز على جائزة الألسكو-الشارقة للدراسات اللغوية والمعجمية، وله عدد من البحوث المنشورة في الدراسات المعجميّة، والتّخطيط اللّغويّ، منها: "النّسق في الصّناعة المعجميّة: الوسيط أنموذجًا" وبحث بعنوان: "ألفاظ الحضارة في معجم الدّوحة التّاريخيّ للّغة العربيّة - حقل اللّباس نموذجًا"، وكتاب يصدر عن مجمع اللّغة العربيّة اللّيبي، قريبا بعنوان: "التخطيط والسياسة اللغوية في ليبيا".

تقييم المعرفة المعجميّة في نماذج الذّكاء الاصطناعيّ العربيّة: دراسة في ضوء بيانات معجم الدّوحة التّاريخيّ للّغة العربيّة"

يستكشف هذا البحث قدرة النّماذج اللّغويّة العربيّة السياقية القائمة على المحوّلات (transformer-based) على تمثيل المعرفة الدّلاليّة المعجميّة، ورصد التّغيّر الدّلاليّ التّاريخيّ في اللّغة العربيّة، انطلاقًا من افتراض أنّ التّمثيلات الدّاخليّة لهذه النّماذج ترمّز الفروق بين معاني اللّفظ الواحد. تنبع الإشكاليّة من محدوديّة تفسير النّماذج الحاسوبيّة عامّة، وندرة أدوات تقييم عربيّة دقيقة على وجه خاصّ، مقارنة بالنّماذج الإنجليزيّة، وقد اعتمدت الدّراسة منهجًا تفسيريًّا مقارنًا بتحليل تمثيلات أربعة نماذج عربيّة تتبع الهندسة الحاسوبيّة لنموذج (BERT)، هي: (CAMeLBERT-)، (CAMeLBERT-CA)، (ARBERT)، (ArABERT Mix) عبر (852) سياقًا لِلَفظ (شهادة) موزّعة على أربع حقب تاريخيّة، باستخدام القياس متعدّد الأبعاد (Multidimensional Scaling: MDS) أداةً للتّمثيل البصريّ، وقيمة التّمييز المعمّمة (Generalized Discrimination Value: GDV)، وقد أظهرت النّتائج تفوّق (ARBERT) في فصل المعاني داخل الحقبة الواحدة، وكفاءة (CAMeLBERT-Mix) في تتبّع تغيّر المعنى عبر الأزمنة التّاريخيّة المختلفة.

بشير نصري

باحث تونسي ودكتور في علم اجتماع التربية، يعمل أستاذًا باحثًا بجامعة جندوبة في تونس، ويشرف على عدد من البحوث والدراسات العلمية في مجال علم الاجتماع. تتركّز اهتماماته العلمية في علم اجتماع التربية والتنشئة الاجتماعية والتمثلات الاجتماعية، إضافة إلى دراسة مسارات التغير الاجتماعي في تونس بعد سنة 2011. وله منشورات عدة من أبرزها: التنشئة الاجتماعية المدرسية في تونس، وكتاب الرياضة، التنشئة الاجتماعية والتمثلات ضمن مؤلف جماعي، وكتاب الهوية والتمثلات ضمن مؤلف جماعي.

السياقات الثقافية والمعرفية في نماذج الذكاء الاصطناعي العربية: مقاربة سوسيولوجية ومعرفية

تتناول هذه الدراسة العلاقة المعقدة بين الذكاء الاصطناعي والسياقات الثقافية والمعرفية في العالم العربي من منظور سوسيولوجي ومعرفي، مع التركيز على التفاعل الديناميكي بين اللغة والثقافة والتكنولوجيا وتأثير ذلك على تطوير النماذج الذكية. وتنطلق الدراسة من فرضية مفادها أن نجاح الذكاء الاصطناعي العربي يعتمد على فهم معمق للخصوصيات اللغوية والثقافية العربية، والتي تعكس أنماط التفكير، والعادات المعرفية، والإرث الحضاري المتراكم عبر القرون. وتنقسم الدراسة إلى ثلاثة مستويات مترابطة: الأول يركز على التحديات اللغوية المرتبطة بمعالجة اللغة العربية داخل النماذج الذكية، الثاني يستعرض تأثير القيم الاجتماعية والثقافية على تصميم واستخدام هذه النماذج، والثالث يستشرف التحديات المستقبلية المتعلقة بالهوية الرقمية، الخصوصية، والسيادة المعرفية. وتشدد الدراسة على ضرورة أن يقوم الذكاء الاصطناعي العربي على مرجعية ثقافية مستقلة تعزز التمكين المعرفي والحفاظ على الهوية العربية، مع التوفيق بين الأصالة والمعاصرة، وتجنب التقليد الأعمى للنماذج الغربية، لضمان نموذج ذكي عربي متميز وشامل وفعال.

دكتوراه دولة في اللسانيّات؛ أستاذ اللسانيات والمعجمية العربية بمعهد الدوحة للدراسات العليا؛ نائب رئيس المجلس العلمي لمعجم الدوحة التاريخي للغة العربية؛ أستاذ فخري بجامعة ليون2، فرنسا، مركز البحث في اللسانيّات التطبيقيّة؛ كان مديرًا لقسم الدراسات العربيّة، ولمكتب المعجميّة والمصطلح والترجمة العربيّة، ولمركز البحث في اللسانيّات العربيّة؛ له أكثر من مئة بحث في العربيّة واللسانيّات والمعجم والمصطلح والترجمة؛ أشرف على 32 أطروحة دكتوراه و150 رسالة ماجستير؛ نال جائزة ابن خلدون-سنغور في الترجمة.

التأريخ في المعجم التاريخيّ الفرنسّي ومعجم الدوحة التاريخيّ للُّغة العربيّة

تعيين التاريخ، أي تعيين تاريخ ظهور مفردات اللغة في النصوص، وظهور دلالاتها فيها عنصرٌ ضروريٌّ من عناصر المعجم التاريخيّ للُّغة، ولا يمكن أنْ يكون المعجمُ تاريخيًّا من دونه. وقد قام معجم الدوحة التاريخيّ للُغة العربيّة بجهد هائل غير مسبوق في تعيين هذا العنصر الذي كان غائبًا في التراث المعجميّ العربيّ كلِّه . غير أنَّ هذا العنصر، على أهمِّيَّته وضرورته، قد لا يكون بالضرورة العنصر الوحيد في التأريخ، بل قد تكون إلى جانبه عناصرُ أخرى، بها يتميّزُ المعجمُ التاريخيُّ من المعجم اللغويّ العامّ؛ فقد توجدُ معاجمُ لغويّةٌ عامّةٌ تعيِّنُ تواريخ ظهور ألفاظها وتواريخ ظهور دلالاتها إلى جانب المعجم التاريخيّ. وتنظر الورقةُ في هذا العنصر، وفي بعض العناصر الأخرى المستخرَجة من المعجم التاريخيّ الفرنسيّ في عمليّة التأريخ.

حسن كيسان

باحث دكتوراه في اللسانيات العصبية-الحاسوبية بجامعة فريدريش ألكسندر إرلنغن-نورنبيرغ بألمانيا، وباحث زائر سابق في معهد ماكس بلانك للسانيات النفسية. تهتم أبحاثه بنماذج اللغة العربية الحاسوبية، والتغيّر الدلالي، وبناء الموارد الرقمية. نشر أعمالاً في اللسانيات الحاسوبيّة والعصبية، مع تركيز خاص على توظيف الذكاء الاصطناعي لخدمة اللغة العربيّة وتعليمها.

تقييم المعرفة المعجميّة في نماذج الذّكاء الاصطناعيّ العربيّة: دراسة في ضوء بيانات معجم الدّوحة التّاريخيّ للّغة العربيّة

يستكشف هذا البحث قدرة النّماذج اللّغويّة العربيّة السياقية القائمة على المحوّلات (transformer-based) على تمثيل المعرفة الدّلاليّة المعجميّة، ورصد التّغيّر الدّلاليّ التّاريخيّ في اللّغة العربيّة، انطلاقًا من افتراض أنّ التّمثيلات الدّاخليّة لهذه النّماذج ترمّز الفروق بين معاني اللّفظ الواحد. تنبع الإشكاليّة من محدوديّة تفسير النّماذج الحاسوبيّة عامّة، وندرة أدوات تقييم عربيّة دقيقة على وجه خاصّ، مقارنة بالنّماذج الإنجليزيّة، وقد اعتمدت الدّراسة منهجًا تفسيريًّا مقارنًا بتحليل تمثيلات أربعة نماذج عربيّة تتبع الهندسة الحاسوبيّة لنموذج (BERT)، هي: (ARBERT)، (ArABERT)، (CAMeLBERT-Mix)، (CAMeLBERT-CA) عبر(852) سياقًا لِلَفظ (شهادة) موزّعة على أربع حقب تاريخيّة، باستخدام القياس متعدّد الأبعاد (Multidimensional Scaling: MDS) أداةً للتّمثيل البصريّ، وقيمة التّمييز المعمّمة (Generalized Discrimination Value: GDV)، وقد أظهرت النّتائج تفوّق (ARBERT) في فصل المعاني داخل الحقبة الواحدة، وكفاءة (CAMeLBERT-Mix) في تتبّع تغيّر المعنى عبر الأزمنة التّاريخيّة المختلفة. الكلمات المفتاحيّة: اللّغة العربيّة، المعجم التّاريخيّ، التّغيّر الدّلاليّ، الذّكاء الاصطناعيّ، النّماذج اللّغويّة السّياقيّة.

ربيع ضاهر أمهز

باحث في الذكاء الاصطناعي التعددي الوسائط في جامعة ستراسبورغ مختبر iCube. سبق له أن شغل منصب رئيس قسم التحول الرقمي في جامعة إيكام ستراسبورغ (ICAM Strasbourg) وقبلها جامعة ستراسبورغ. يحمل شهادة الدكتوراه في المعلوماتية ومعالجة الصور، من École Centrale de Nantes بفرنسا. اهتماماته البحثية تشمل تحليل السلوك البشري في الأنظمة المعقدة، والتفاعل الذكي بين الإنسان والآلة. منشوراته "خوارزمية لاكتشاف الشقوق تلقائيًا في صور الرصف"

يليل: منصة للتدقيق في اصطفاف النصوص والكلام باللغة العربية من أجل تقييم تطبيقات الذكاء الاصطناعي

تعتبر اللغة العربية من أكثر لغات العالم استخدامًا، إلا أن تقنيات الكلام الآلي فيها متأخرة مقارنة باللغات ذات الموارد العالية، بسبب الازدواجية اللغوية (الفصحى واللهجات)، والغنى الصرفي مع سوابق ولواحق/كلتيكات، وتفاوت الكتابة (التشكيل، اختلافات الألف والهمزات، الأرقام العربية والهندية)، إضافةً إلى تشتت الموارد واختلاف تراخيصها وجودتها. ينطلق البحث من فرضية أن منصة متكاملة لاصطفاف النصوص مع الصوت يمكنها تقليل هذه التحديات وتحسين جودة الموارد العربية. استُخدمت منهجية تصميم وتطوير منصة "يليل"، أداة تستقبل نصًا عربيًا وملفًا صوتيًا مطابقًا، وتُرجع طوابع زمنية لكل كلمة بصيغة JSON، مع حساب WER وCER، وتوفير واجهة تفاعلية تُبرز الكلمات تزامنًا مع الصوت. تعتمد "يليل" على خوارزميات المزامنة القسرية، وتتكامل مع موارد عربية وموديلات عالمية لتوليد طوابع دقيقة. أظهرت التجارب فعالية عالية في الاصطفاف والتدقيق، مع بروتوكولات تقييم كمية وبشرية. وبذلك تُسهم "يليل" في سد فجوة جوهرية بالذكاء الاصطناعي العربي، وتوفير مخرجات شفافة قابلة لإعادة الإنتاج تدعم تطوير تطبيقات ASR وTTS والنماذج متعددة الوسائط.

رزان عبدالله حلبي

ريادية وباحثة ووسيطة اجتماعية، مؤسِّسة ومديرة تنفيذية لمنصّة إيليل للكتب العربية المسموعة والأبحاث العلمية في ستراسبورغ. حاصلة على إجازة في العلوم السياسية من جامعة السوربون وماجستير في الدراسات الكولونيالية من جامعة بيرزيت. تُدرّس طلاب الهندسة في جامعة إيكام. نشرت أكثر من 30 مقالاً سياسيًا واجتماعيًا وتركّز أبحاثها على الهوية والانتماء في أوروبا.

يليل: منصة للتدقيق في اصطفاف النصوص والكلام باللغة العربية من أجل تقييم تطبيقات الذكاء الاصطناعي

تعتبر اللغة العربية من أكثر لغات العالم استخدامًا، إلا أن تقنيات الكلام الآلي فيها متأخرة مقارنة باللغات ذات الموارد العالية، بسبب الازدواجية اللغوية (الفصحى واللهجات)، والغنى الصرفي مع سوابق ولواحق/كلتيكات، وتفاوت الكتابة (التشكيل، اختلافات الألف والهمزات، الأرقام العربية والهندية)، إضافةً إلى تشتت الموارد واختلاف تراخيصها وجودتها. ينطلق البحث من فرضية أن منصة متكاملة لاصطفاف النصوص مع الصوت يمكنها تقليل هذه التحديات وتحسين جودة الموارد العربية. استُخدمت منهجية تصميم وتطوير منصة "يليل"، أداة تستقبل نصًا عربيًا وملفًا صوتيًا مطابقًا، وتُرجع طوابع زمنية لكل كلمة بصيغة JSON، مع حساب WER وCER، وتوفير واجهة تفاعلية تُبرز الكلمات تزامنًا مع الصوت. تعتمد "يليل" على خوارزميات المزامنة القسرية، وتتكامل مع موارد عربية وموديلات عالمية لتوليد طوابع دقيقة. أظهرت التجارب فعالية عالية في الاصطفاف والتدقيق، مع بروتوكولات تقييم كمية وبشرية. وبذلك تُسهم "يليل" في سد فجوة جوهرية بالذكاء الاصطناعي العربي، وتوفير مخرجات شفافة قابلة لإعادة الإنتاج تدعم تطوير تطبيقات ASR وTTS والنماذج متعددة الوسائط.

رشيد بلحبيب

أستاذ الدراسات اللغوية، حاصل على درجة دكتوراه الدولة. درّس علوم اللغة العربية في جميع المستويات، في عدد من الجامعات المغربية والعربية. مدير سابق لمركز الدراسات والبحوث الإنسانية والاجتماعية بوجدة. خبير لغوي في معجم الدوحة التاريخي للغة العربية منذ تأسيسه، وعضو مجلسه العلمي، ونائب المدير التنفيذي للمعجم، ورئيس اللجنة العلمية للمعجم حاليا. له إسهامات في التأليف والتحقيق والترجمة، نشر ما يزيد على خمسين بحثا محكما بالعربية والإنجليزية. أشرف وناقش ما يزيد على سبعين رسالة جامعية بين ماجستير ودكتوراه. محكّم معتمد لدى عدد من الجامعات والمجلات العلمية في الوطن العربي وخارجه، في مجالي النشر، والترقيات العلمية. شارك في عشرات المؤتمرات المحلية والإقليمية والدولية.

من الاستمداد إلى الإمداد: قضايا الإحياء والتّفصيح في ضوء معجم الدّوحة التاريخي للغة العربية

تروم هذه الورقة هدفين رئيسين: أوّلًا، إبراز قيمة مصادر الاستمداد التي اعتمدها معجم الدوحة التاريخي للغة العربية، باعتباره جمعًا ثانيًا للغة بفضل تنوّع المتون واتّساعها، وما نتج عنه من مداخل معجمية موثّقة تعرّف بالألفاظ وتاريخها وضبطها ونِسبتها.

وثانيًا، رصد انتقال المعجم من مجرد أداة بحث إلى موضوع للبحث نفسه؛ إذ كشف بعد اكتماله عن إمكانات بحثية واسعة ومشاريع علمية جديدة، أسهمت في مراجعة كثير من المسلّمات اللغوية، واستدراك ما أُقصي من ألفاظ وتراكيب بدعوى الخطأ أو ضعف الفصاحة، وإحياء ما عُدّ مهجورًا، وتقويم ما نُظر إليه بوصفه عاميًا.

ويؤكد ذلك بداية مرحلة جديدة في حياة العربية، قوامها إثراء لغوي متجدد يستند إلى مداخل المعجم ومدوّنته الواسعة.

رفيف السيد

أستاذة الذكاء الاصطناعي وهندسة المعرفة، رئيسة اللجنة الاستشارية للحاضنة التقنية ايفوتيك، ومستشارة دولية في سياسات العلوم والابتكار التكنولوجي واستراتيجيات الذكاء الاصطناعي وتطبيقاته والتحول الرقمي والدبلوماسية العلمية والتنمية والتعليم. درست وحاضرت في عدة جامعات مرموقة: جامعة ساري وجامعة أكسفورد (المملكة المتحدة)، جامعة جورج واشنطن (الولايات المتحدة)، جامعة العلوم والتكنولوجيا وجامعة تسينغهوا (الصين)، وجامعة دمشق . لديها قائمة من الأبحاث العلمية والكتب المنشورة، شاركت في تأسيس عدد من مراكز المعرفة والابتكار الدولية، حاصلة على عدة جوائز دولية للتميز في البحث والإنجاز.

نحو تعزيز الذكاء الاصطناعي للغة العربية: رؤية من منظور السياسات

تهدف هذه الورقة البحثية إلى الوصول إلى فهم أعمق للتحديات التي تعيق تطوير أنظمة قوية وموثوقة للذكاء الاصطناعي التوليدي باللغة العربية، وتناقش آثار هذه التحديات من خلال إجراء تحليل مقارن بين اللغة العربية وعددٍ من اللغات العالمية، وفي مقدمتها اللغة الإنجليزية، من حيث توافر الموارد الرقمية وتنوّع العناصر اللغوية. وتسلّط الدراسة الضوء على الفجوة الرقمية القائمة والأسباب الكامنة وراءها، سواء فيما يتعلّق بحجم الإنتاج البحثي والنشر العلمي باللغة العربية، أو بتطوير مجموعات نصوص عالية الجودة لإغناء المحتوى الرقمي، وتوسيع جهود الترجمة، وتوحيد المصطلحات المتعلقة بها، بالإضافة إلى الأطر والسياسات الداعمة لمعالجة هذه التحديات. وفي الختام، تخلص الورقة إلى مجموعة من التوصيات الموجَّهة إلى صنّاع القرار وواضعي السياسات في العالم العربي، لتعزيز التطوير المستقبلي للذكاء الاصطناعي باللغة العربية. وتشير الورقة إلى أن هذه التحديات أكبر من أن تُواجَه عبر مبادرات متفرقة، بل إنّ مواجهتها تتطلّب تضافرًا وتكاملا للجهود والإمكانيات، من خلال إنشاء تحالف عربي مقترَح يتمتّع بهيكلية مرنة، ويستقطب جميع الجهات المعنية بخلفياتها واختصاصاتها المتنوعة، بهدف بناء البيئة الابتكارية المتكاملة لتشجيع الابتكار في مجال الذكاء الاصطناعي التوليدي باللغة العربية، تعزيزا لدورها ومكانتها في عصر الذكاء الاصطناعي."

سامر رشواني

باحث وأستاذ في مركز دراسات التشريع الإسلامي والأخلاق بجامعة حمد بن خليفة، متخصص في الدراسات القرآنية وتاريخ الفكر الإسلامي المعاصر، والأخلاق الإسلامية. عمل مدرسا في كلية الشريعة بجامعتي دمشق وحلب (2007-2011)، و وزميلاً بمنتدى الدراسات العابرة للأقاليم ببرلين (2011-2013)، وباحثا ومدرسا في معهد الدراسات الإسلامية بجامعة توبنغن. من منشوراته: التفسير الموضوعي للقرآن الكريم دراسة نقدية (2007)، و المناهج الحديثة في الدرس القرآني (2011)، :Behind the Story Ethical Readings of Qurʾānic Narratives, 2024.

تطوير خادم للذكاء الاصطناعي معزَّز بمعجم الدوحة التاريخي للغة العربية: من أجل فهم أدق لنصوص القرآن والحديث النبوي

ما تزال النماذج اللغوية الضخمة (LLM) المستندة إلى تقنيات الذكاء الاصطناعي تتعثر عند مقاربة النصوص العربية القديمة التاريخية والدينية، كالقرآن الكريم والحديث النبوي. ومن أجل تذليل هذه العقبة، عمدنا إلى منهجية تقوم على تعزيز قدرات نماذج لغوية عربية بمادة «معجم الدوحة التاريخي للغة العربية»، وذلك باعتماد تقنية «التوليد المعزَّز بالاسترجاع(RAG)».. وقد شُيِّدت منهجيتنا البحثية على نظام استدعاء للبيانات يقوم على آلية توجيه تُحلِّل وتُحدِّد مقاصد السائل والمستفهم قبل أن تمُدَّ النماذج اللغوية ببيانات دقيقة مناسبة للسياق. وقد أبانت التجارب أن تعزيز نماذج الذكاء الاصطناعي بمادة معجم الدوحة، المؤلَّفة مما يربو على 8700 جذر لغوي، عن ارتقاء دقة النماذج العربية، مثل فَنار وعلام، من نحو 50% إلى نحو80%. بل إن نموذجا لغويا متقدما مثل«Gemini»حقق تقدما في فهم الدلالات والمعاني السياقية وارتقت دقته من نحو 82% إلى نحو 88%. هذا ناهيك عن القدرة على رصد التطور الدلالي للألفاظ، وتأثيلها في اللغات السامية، وذكر شواهدها في النقوش والنصوص القديمة.

سعد محمد عبد الغفار يوسف

أستاذ البلاغة وتحليل الخطاب بجامعة الوادي الجديد، وأستاذ زائر بجامعة إنديانا (الولايات المتحدة). شغل منصب رئيس قسم اللغة العربية، وخبير تعليم العربية للناطقين بغيرها بوزارة الخارجية المصرية. تشمل اهتماماته: البلاغة وتحليل الخطاب، وتعليم العربية، واللسانيات التطبيقية والجنائية. من أبرز أعماله: Metaphor between the historicity of poetic imagination (Journal of Qur'anic Studies – Q1)، إعجاز البيان القرآني، والبنية الصوتية والدلالية في البيان القرآني (جامعة الملك عبد العزيز، 2025).

تحليل الانزياحات الدلالية الزمنيَّة للمفاهيم الثقافية في اللُّغة العربية باستخدام الذكاء الاصطناعي: دراسة تطبيقيَّة في معجم الدوحة التاريخي

تتناول الدراسة إشكالية غياب نماذج عربية قادرة على تتبع الانزياحات الدلالية للمفاهيم الثقافية عبر حقب زمنية متعددة، مما يحد من فعالية التطبيقات الحاسوبية للغة العربية. ولتجاوز ذلك، طوَّرت الدراسة نموذجًا معرفيًا يعتمد على تقنيات الذكاء الاصطناعي، مثل نماذج التضمين العددي للكلمات، وخوارزميات المعالجة الزمنية للغة (Time-aware NLP)، مستندةً إلى معجم الدوحة التاريخي؛ حيث تمَّ تحويل مفهومي "الحرية، و"الكرم" إلى تمثيلات عددية عبر خمس حقب زمنية من العصر الجاهلي حتى القرن الثالث الهجري، ثم قياس الانزياحات باستخدام مؤشرات كَميَّة أبرزها التشابه الكوني الزاوي (Cosine Similarity) ومؤشر الانزياح الدلالي (SDI)، إضافة إلى خرائط حرارية دلالية. وقد أظهرت النتائج أن مفهوم "الحرية" شهد أعلى معدل انزياح (84%) مقارنة بـ"الكرم" (57%)، ممَّا يعكس حساسيته الزمنية، وتنوع سياقاته الثقافية والسياسية. وقد حاولت الدراسة تقديم إطار تطبيقي مبتكر لدعم الهوية اللغوية العربية في البيئات الرقمية.

سليمان بن سالم الشهري

أستاذ مساعد في قسم تقنية المعلومات بكلية علوم الحاسب والمعلومات بجامعة الإمام محمد بن سعود الإسلامية، ورئيس مركز بحوث علوم الحاسب والمعلومات. عمل سابقًا رئيسًا تنفيذيًا للمركز السعودي للمحتوى الرقمي بدارة الملك عبدالعزيز. يهتم بالذكاء الاصطناعي اللغوي، والعدالة اللغوية، والتحول الرقمي في مؤسسات الذاكرة الثقافية، والتقنيات الخضراء ودعم أهداف التنمية المستدامة.

أداء النماذج اللغوية التوليدية في تمثيل الأمثال الشعبية الخليجية: قراءة تحليلية في المعنى والأسلوب والهوية

يتناول البحث "أداء النماذج اللغوية التوليدية في تمثيل الأمثال الشعبية الخليجية: قراءة تحليلية في المعنى والأسلوب والهوية" كيفية تعامل الذكاء الاصطناعي مع التراث الشفهي الخليجي من منظور لغوي وثقافي. حيث تم تحليل ثلاثين مثلًا من ست دول خليجية عبر نموذج GPT-5 باستخدام منهجية مختلطة تجمع بين التحليل الكمي والكيفي وفق ستة معايير تشمل الفهم الدلالي، والتوظيف الثقافي، والأسلوب، والتطبيق الواقعي، والإسناد الجغرافي، وتفسير المفردات. أظهرت النتائج قصورًا واضحًا في الجوانب الثقافية والمجازية للنموذج رغم أدائه المقبول في الفهم الدلالي، مما يعكس غياب التمثيل العميق للهوية والأسلوب الشعبي في بيانات التدريب. تؤكد الدراسة أن المشكلة ليست في نقص البيانات فقط، بل في ضعف إدماج الأبعاد الثقافية في هندسة النماذج، وتدعو إلى تطوير موارد لغوية خليجية ومعايير تقييم أكثر حساسية للسياق الثقافي حفاظًا على العدالة اللغوية والهوية التراثية في بيئة الذكاء الاصطناعي.

باحثة مساعدة في مركز دراسات التشريع الإسلامي والأخلاق (CILE)، جامعة حمد بن خليفة، شغلت وظيفة باحثة مساعدة في جامعة قطر وحمد بن خليفة. من اهتماماتها العلميّة: الذكاء الاصطناعي، ومعالجة اللغة الطبيعية، واسترجاع المعلومات، والتفاعل بين الإنسان والحاسوب.

تطوير خادم للذكاء الاصطناعي معزَّز بمعجم الدوحة التاريخي للغة العربية: من أجل فهم أدق لنصوص القرآن والحديث النبوي

ما تزال النماذج اللغوية الضخمة (LLM) المستندة إلى تقنيات الذكاء الاصطناعي تتعثر عند مقاربة النصوص العربية القديمة التاريخية والدينية، كالقرآن الكريم والحديث النبوي. ومن أجل تذليل هذه العقبة، عمدنا إلى منهجية تقوم على تعزيز قدرات نماذج لغوية عربية بمادة «معجم الدوحة التاريخي للغة العربية»، وذلك باعتماد تقنية «التوليد المعزَّز بالاسترجاع(RAG)».. وقد شُيِّدت منهجيتنا البحثية على نظام استدعاء للبيانات يقوم على آلية توجيه تُحلِّل وتُحدِّد مقاصد السائل والمستفهم قبل أن تمُدَّ النماذج اللغوية ببيانات دقيقة مناسبة للسياق. وقد أبانت التجارب أن تعزيز نماذج الذكاء الاصطناعي بمادة معجم الدوحة، المؤلَّفة مما يربو على 8700 جذر لغوي، عن ارتقاء دقة النماذج العربية، مثل فَنار وعلام، من نحو 50% إلى نحو80%. بل إن نموذجا لغويا متقدما مثل«Gemini»حقق تقدما في فهم الدلالات والمعاني السياقية وارتقت دقته من نحو 82% إلى نحو 88%. هذا ناهيك عن القدرة على رصد التطور الدلالي للألفاظ، وتأثيلها في اللغات السامية، وذكر شواهدها في النقوش والنصوص القديمة.

أستاذ بجامعة الأغواط في الجزائر وعضو المجمع الجزائري للغة العربية. تتركّز اهتماماته العلمية في اللسانيات النظرية ولسانيات اللغة العربية. أصدر عددًا من المؤلفات المهمة، من أبرزها: اللسانيات وقضايا اللغة العربية (مطبعة رويغي، الأغواط، 2014)، والتفكير السيميائي في اللغة والأدب: قراءة في تراث أبي حيّان التوحيدي (دار عالم الكتب الحديث، الأردن، 2015)، ومبادئ اللسانيات البنوية: دراسة تحليلية إبستمولوجية (دار كنوز المعرفة، الأردن، 2024).

محاولة في توصيف عمل العلامات والقرائن في نظام النحو العربي خلفية لسانية ضرورية لمواجهة تحدّيات الترجمة الآلية

ينطلق موضوع هذه الدراسة من النظر في إمكانية الإفادة من خصائص النحو العربي في مواجهة مشكلات الترجمة من العربية وإليها باستخدام الذكاء الاصطناعي، وفي تذليل الصعوبات الناجم عنها. ومن أجل المضي في الاستجابة لمقتضيات هذه المعالجة حاولت الدراسة أن تقترح توصيفا لعلامات النحو وقرائنه ولآلية عملها في بناء المعنى وفي توجيهه؛ والغرض من هذا التوصيف أن نؤسِّس بنتائجه لصياغة نماذج لسانية رقمية تكون أقدر على فهمٍ أعمق وأدق لمعاني النصوص في مختلِف تراكيبها وسياقاتها، وهو ما يمكّن برامج الذكاء الاصطناعي من تقديم ترجمة آلية مثلى عند النقل من العربية أو إليها.

عبد الحميد صبار

طالب دكتوراه وباحث في مختبر تكنولوجيا المعلومات والنمذجة بكلية العلوم بن مسيك – جامعة الحسن الثاني بالدار البيضاء، المغرب. تتركّز أبحاثه حول الذكاء الاصطناعي، ومعالجة اللغة الطبيعية (NLP)، والرؤية الحاسوبية، والتعلّم العميق، مع اهتمام خاص بتطبيق الذكاء الاصطناعي والنماذج المتقدّمة في حلول الأتمتة الموجّهة للّغات العربية.

تحليل المشاعر للهجات العربية: دراسة علمية تجريبية

تتناول هذه الدراسة تحليل المشاعر في اللغة العربية من منظور علمي مقارن بين ثلاث مقاربات رئيسية هي: التعلّم الآلي (Machine Learning)، والتعلّم العميق (Deep Learning)، والتعلّم بالنقل (Transfer Learning). يهدف البحث إلى تقييم كفاءة هذه النماذج في تصنيف المشاعر داخل نصوص اللغة العربية الفصحى الحديثة (MSA)، باستخدام مجموعة من corpora المتنوعة المأخوذة من منصّات التواصل الاجتماعي ومواقع المراجعات الإلكترونية. أظهرت النتائج أن نماذج التعلّم العميق مثل BiGRU وLSTMs قد حققت أداءً متقدماً مقارنة بالأساليب التقليدية، حيث تجاوزت دقّتها 90%. ومع ذلك، أثبتت نماذج التعلّم بالنقل، وخاصة AraBERT، تفوّقاً واضحاً بدقّة بلغت 94.57%، بفضل قدرتها على استغلال التمثيلات اللغوية المُدرّبة مسبقاً على مجموعات ضخمة من النصوص العربية. تؤكد الدراسة الفعالية العلمية لنماذج التعلّم بالنقل في التعامل مع الثراء اللغوي والتنوّع اللهجي في اللغة العربية، مع الإشارة إلى التحديات البحثية المستمرة مثل قلة البيانات المعرَّفة والحاجة إلى نماذج هجينة تجمع بين الذكاء الاصطناعي المتقدّم والخصائص اللغوية العربية."

أستاذ العلوم اللغوية بكلية الآداب – جامعة طنطا (مصر)، درّس بجامعات طنطا، والإمارات، وليدز البريطانية، والقوميات بالصين. ألَّف 6 كتب مرجعية، و25بحثًا. حقّق كتاب الغريبين للهروي في 10أجزاء. ترجم كتاب «اللسانيات» لـ Jean Atichison. أشرف على 130رسالة، وناقش160. خبير بمجمع اللغة العربية، وعضو لجنة تحقيق التراث بدار الكتب المصرية، وعضو المجلس العلمي لمعجم الدوحة التاريخي للغة العربية."

فقه اللغة وتحوّل الدلالات

يُعنى هذا البحث بدراسة بعض قضايا «فقه اللغة» في معجم الدوحة التاريخي للغة العربية. وفي المقام الأول من هذه القضايا: قضايا معالجة المعنى في هذا المعجم. ومنها: ضوابط تعريف المفردات، وتحرير دلالاتها من الغموض، باستيفاء ذِكر مكوّناتها الدلالية semantic components، وشرحها شرحًا يعتبر صيغها الصرفية، مع التنويه بدور الشواهد في تحرير هذه المعاني. وكذا ينظر البحث فيما قرّ في الدرس اللساني الحديث من تحدُّر الدلالات المجردة abstract من أخرى حسيّة concrete، وأسبقية هذه الأخيرة في الوجود، في ضوء معطيات معجم الدوحة. كما يتناول البحث بالدرس ما يوفّره هذا المعجم من معلومات معجمية، تتيح للدارس أن يسجِّل حركيّة الدلالة عبر الزمن: تعميمًا، وتخصيصًا، وانتقالًا ، بما في ذلك المصطلحات ذات الانتماءات المعرفية المختلفة. الكلمات المفتاحية: -معجم الدوحة- فقه اللغة - تحرير المعنى- الحسي والمجرّد- التغير الدلالي.

عبدالله يوسفي

أستاذ باحث في كلية العلوم القانونية والاقتصادية والاجتماعية السويسي الرباط، جامعة محمد الخامس الرباط-المغرب الوظائف السابقة: إطار سابق في وزارة التعليم العالي / باحث في معهد الدراسات والأبحاث للتعريب / رئيس مختبر البحث لتحليل المعطيات والإعلاميات في العلوم الإنسانية. الاهتمامات العلميّة: المعالجة الآلية للغة العربية / البحث في مجال الاقتصاد القياسي / البحث في مجال التعليم العالي. وله عدد من الأعمال العلمية المنشورة.

معالجة الأخطاء التركيبية باستعمال النماذج اللغوية الضخمة

يعد تصحيح الأخطاء اللغوية ركيزة أساسية في مجال المعالجة الآلية للغات، حيث تشكل صحة الجمل أو صحة قواعد المعطيات اللغوية عاملا أساسيا في تدريب أنظمة الذكاء الاصطناعي عموما، وفي بناء النماذج اللغوية الضخمة على وجه الخصوص. ومن بين أهم الأخطاء اللغوية الشائعة التي ترتكب في بناء الجمل نجد الأخطاء التركيبية أو النحوية (Syntactic Error)، والتي تنتج عادة بسبب عدم توافق أو تناسق الكلمات تركيبيا داخل الجملة.

في هذا المقال سنقوم أولا: باستعراض أدبي لأبرز الأعمال التي صدرت في هذا المجال خلال السنوات الأخيرة، مع التركيز على الأعمال التي استخدمت الأنماط اللغوية الكبيرة وبالخصوص "AraBERT" و"mBERT". ثانيا: سنجري مقارنة بين النظامين "AsynGEC" و"GECToR" للكشف عن الأخطاء النحوية وتصحيحها. من أجل هذه المقارنة قمنا بإعادة إنتاج هذين النظامين الذين دربا على قاعدة المعطيات "QALB" و"Tibyan". هذه القواعد تحتوي على أنماط متنوعة من الأخطاء اللغوية، وقد ساعدت في عملية الضبط الدقيق للنماذج اللغوية الكبيرة المستعملة في هذه المقارنة. هذه الدراسة مكنت من تحديد نقاط القوة ونقاط الضعف في أهم أنظمة التصحيح الآلي للأخطاء التركيبية، الشيء الذي سيمكننا مستقبلا من تطوير نظام خاص بمعالجة هذا النوع من الأخطاء، وذلك بالأخذ بعين الاعتبار نتائج هذه الدراسة.

عز الدين البوشيخي

المدير التنفيذي لمشروع معجم الدوحة التاريخي للغة العربية. حصل على دكتوراه الدولة في اللسانيات من جامعة مولاي إسماعيل بمكناس في المغرب. وشغل منصب أستاذ للتعليم العالي، ونائب العميد في شؤون البحث العلمي والتعاون بالكلية ذاتها، ومديرا لمركز دراسات الدكتوراه، ومديرا للمركز الجامعي لتعليم اللغة العربية وحضارتها، ومديرا لمختبر الدراسات اللسانية والتطبيقات الحاسوبية، ورئيس قسم اللغة العربية والمنسق البيداغوجي لمسلك الدراسات العربية. عمل أستاذًا زائرًا ومحاضرًا بجامعة قطر، ومديرًا لبرنامج اللسانيات والمعجمية العربية في معهد الدوحة للدراسات العليا وأستاذا محاضرا فيه. وهو خبير متعاون مع عدد من المراكز والمنظمات الإقليمية. له عدد من المؤلفات والأبحاث في مجالات اللسانيات والمعجم والمصطلح والترجمة.

بيانات معجم الدوحة التاريخي للغة العربية موردا لتطوير تطبيقات الذكاء الاصطناعي العربي

يُعد ضعف أداء اللغة العربية في تطبيقات الذكاء الاصطناعي مقارنة بالإنجليزية أمرًا شائعًا في الأدبيات، إذ يظهر التفاوت في الترجمة، وتحليل النصوص والمشاعر، وفهم السياقات الثقافية واللغوية، والتعامل مع العبارات المجازية والالتباس اللغوي بمستوياته المختلفة.

يهدف هذا البحث إلى إبراز أهمية بيانات معجم الدوحة التاريخي للغة العربية في تطوير الذكاء الاصطناعي العربي، إذ توفر نحو 300 ألف مدخل معجمي، ومدونة نصية تقارب مليار كلمة، إضافة إلى ببليوغرافيا منظَّمة في قاعدة بيانات مهيكلة.

ويتناول البحث خمسة محاور رئيسية: وصف عام لبيانات المعجم، ودورها في التحليل الصرفي والتركيبي، وإسهامها في التحليل الدلالي وفك الالتباس، وخصائص منصة صناعة المعجم المرتبطة بالذكاء الاصطناعي، ثم عرض أولى محاولات توظيف هذه البيانات في تطوير نماذج الذكاء الاصطناعي للغة العربية.

عز الدين مزروعي

أستاذ التعليم العالي بجامعة محمد الأول في تخصص الرياضيات والإعلاميات. يشغل منصب محرّر مشارك في مجلة جامعة الملك سعود – علوم الحاسب والمعلومات، ونائب مدير مركز الدراسات والبحوث الإنسانية والاجتماعية بوجدة. شغل سابقًا عضوية مجلس أمناء المنظمة العربية للترجمة، ويشارك في هيئات التحكيم لعدد من المجلات العلمية المتخصصة. له أكثر من مئة ورقة بحثية منشورة في مجلات ومؤتمرات عالمية مُحكَّمة.

تعزيز معالجة اللغة العربية: منصة الخليل لأدوات التحليل الصرفي-النحوي وذخائر النصوص المعنونة مفتوحة المصدر

تعد المحللات الصرفية والنحوية من المكونات الأساسية في مجال معالجة اللغات الطبيعية، لاسيما في مراحل المعالجة المسبقة للنصوص. وتمثل إتاحة هذه الأدوات بشكل مفتوح ركيزة أساسية لدعم الابتكار وتسريع وتيرة البحث العلمي في هذا المجال المتنامي بسرعة. وتكتسب هذه الإتاحة أهمية متزايدة في سياق اللغات القليلة الموارد، وعلى رأسها اللغة العربية، التي ما تزال تعاني من محدودية الموارد والأدوات الحاسوبية المتخصصة، سواء على مستوى التحليل أو التوسيم أو المعالجة الآلية للنصوص. في هذا الإطار، تقدم هذه الورقة منصة الخليل لحوسبة اللغة العربية، التي طورها فريق معالجة اللغات الطبيعية بمختبر أبحاث علوم الحاسوب في جامعة محمد الأول بالمملكة المغربية. وتوفر المنصة للباحثين مجموعة متكاملة من الأدوات مفتوحة المصدر للتحليل الصرفي والنحوي، إلى جانب إتاحة تنزيل عدد من الذخائر اللغوية المعنونة التي استعملت في مرحلتي التدريب والاختبار. وتشكل المنصة مرجعا بحثيا مفتوحا يسهم في تطوير تقنيات اللغة العربية وتعزيز مواردها الرقمية والبحثية.

فادي زراقط

مدير وحدة دراسات المجال الرقمي والاجتماعي العربي في المركز العربي للأبحاث ودراسات السياسات، وهو أستاذ مشارك في قسم هندسة الكمبيوتر في الجامعة الأمريكية في بيروت. تركز أعماله البحثية على التفكُّر الآلي والفهم الآلي للغة وتطبيقات ذلك على النصوص العربية والطبية وصوابية النظم، باستخدام أنظمة حل المعضلات المنطقية والتلقين الآلي والذكاء الاصطناعي وتحليل الشبكات. حصل على درجة الدكتوراه في أنظمة حل المعضلات المنطقية وصوابية النظم من جامعة تكساس في أوستن، ودرجة الماجستير والبكالوريوس في هندسة الكمبيوتر والاتصالات من الجامعة الأمريكية في بيروت. عمل في شركات عدة مثل IBM وSun Microsystems في مجال الصواب الآلي وتطوير الانظمة.

تحسين أداء النماذج اللغوية الضخمة وكلف تدريبها وتشغيلها للغة العربية باستخدام تقنية إعادة استعمال الرموز

تقدّم هذه الورقة البحثية منهجية إعادة استخدام الرموز (Token-Reuse) لتحسين أداء النماذج اللغوية الضخمة في لغةٍ هدف؛ اللغة العربية في حالتنا. تتيح هذه المنهجية تخصيص عدد أكبر من الرموز للعربية في نماذج مفتوحة الوزن تستخدم كل منها آلية ترميز مختلفة، مثل تجزيء الجملة(SentencePiece) والترميزعلى مستوى زوج الوحدة الرقمية (Byte-Pair Encoding). وبهذا تُصبح متابعة تدريب النماذج على النصوص العربية أكثر كفاءة وأقل كلفة، لما ينتج عنها من رفع معدلات الخصوبة النصيّة للبيانات العربية، إلى جانب خفض زمن الاستدلال وكلفته. كما توسّع المنهجية المفردات العربية المتاحة للنموذج. ويتبع ذلك تدريبات استعادة الكفاءة والتنقيح، ليتحسّن أداء النموذج في مهام إكمال النصوص والمهام العربية اللاحقة. تفرض مشكلة إعادة الاستخدام تحديات تقنية تتعلق بمواءمة الرموز الجديدة مع المعرّفات الأصلية داخل النموذج. يعرض هذا البحث حلولاً لهذه التحديات عابرة لعائلات المرمّزات المختلفة، ويقدّم النتائج البارزة التي حسنت أداء النماذج مع تقنية إعادة استخدام الرموز في سياق اللغة العربية.

فوزي حراق

أستاذ بقسم علوم الحاسب بجامعة فرحات عباس، سطيف. حاصل على دكتوراه في علوم الحاسب (2011) من نفس الجامعة. عمل في العديد من المؤسسات الأكاديمية والبحثية داخل الجزائر وخارجها. يترأس مجموعة أبحاث "الذكاء الاصطناعي والأنظمة الذكية" بجامعة فرحات عباس سطيف. ويشرف على عدة أطروحات دكتوراه وماجستير. يرأس مشروع أخلاقيات الذكاء الاصطناعي بالجامعة، كما نشر دراسات بارزة في عدة مؤتمرات ومجلات عالمية مرموقة كما عمل مستشارا في عديد الهيئات والوزارات. يشغل الان منصب مستشار بمجمع الملك سلمان العالمي للغة العربية. حصل الدكتور فوزي على عدة تكريمات بما في ذلك جائزة أفضل ورقة بحثية في مؤتمر حوسبة اللغة العربية CITALA 2009.

حماية الخصوصية وأمن النماذج اللغوية الضخمة لمعالجة البيانات العربية الحسّاسة: إطار مرجعي مكيّف مع خصائص اللغة العربية

تشهد النماذج اللغوية الضخمة انتشارًا متسارعًا، لكن توظيفها في التعامل مع البيانات العربية الحساسة يثير مخاوف حقيقية تتعلق بالأمان والخصوصية. فالعربية بطبيعتها الغنية صرفيًا والمتعددة لهجاتها، إضافة إلى غياب التشكيل في أغلب النصوص، تجعل حماية المعلومات الدقيقة أكثر تعقيدًا مما هو متوقع. في هذا الملخص نتناول هذه الإشكالات من منظور لغوي وتقني وأخلاقي، موضحين أن الأساليب التقليدية لحماية البيانات لا تصبح فعّالة ما لم تُكيَّف مع الخصوصيات اللسانية للعربية. انطلاقًا من ذلك، نقترح إطارًا تكامليًا يحمل اسم SecureMed-LLM، يعتمد على تحليل صرفي ودلالي متعدد المستويات بهدف خفض احتمالات كشف الكيانات الحساسة أثناء المعالجة اللغوية. كما نستعرض الجوانب التنظيمية والأخلاقية في البيئات العربية، مع التأكيد على أهمية وضع معايير خصوصية تراعي الهوية اللغوية والثقافية. وتخلص الورقة إلى أن بناء ذكاء اصطناعي عربي موثوق يتطلب إدماج الوعي اللغوي في تصميم أنظمة الأمان. ويسعى هذا الإطار إلى تحقيق توازن عملي بين حماية البيانات والحفاظ على سلامة المحتوى اللغوي المستخدم في مختلف التطبيقات الحديثة.

كريم بوزوبع

أستاذ التعليم العالي بجامعة محمد الخامس بالرباط. شغل عددًا من المناصب الأكاديمية والبحثية، من أبرزها تأسيس ورئاسة مختبر "هندسة اللغة ونمذجة التعلم" بالجامعة، والعمل محاضرًا أول بجامعة روهامبتون في لندن بالمملكة المتحدة، إضافة إلى تأسيس شركة ناشئة متخصصة في الذكاء الاصطناعي ومعالجة اللغة الطبيعية. تتركز اهتماماته العلمية في الذكاء الاصطناعي، ومعالجة اللغات الطبيعية، والأنظمة متعددة الوكلاء، والنظم المعرفية. ومن أبرز أعماله المنشورة تطوير منصة "سَفَر" (safar) لمعالجة اللغة العربية واللهجة المغربية، ومنصة "مُربّع" (murabaa) الرقمية الخاصة ببنية الكلمة العربية.

هل نحتاج إلى نماذج لغوية عربية ضخمة؟ وإن كان الأمر كذلك، فكيف وبأي شكل؟

تتناول هذه الورقة البحثية مسألة ما إذا كان من الضروري بناء نموذج لغوي كبير خاص باللغة العربية، وإذا كان الأمر كذلك، فما هي الشروط المناسبة لذلك. فعلى الرغم من أن هذه النماذج قد أحدثت تحولًا كبيرًا في مجال معالجة اللغة الطبيعية للغات ذات الموارد العالية مثل الإنجليزية، إلا أن اللغة العربية تواجه تحديات لغوية وثقافية وتقنية فريدة. تستعرض الورقة تطور هذه النماذج، ومزاياها وعيوبها، كما تحلل اللسانيات العربية من خلال مناهج قائمة على القواعد والتعلم الآلي التقليدي، وتُظهر أن بعض المهام يمكن تنفيذها دون الحاجة إلى نماذج لغوية كبيرة. وتناقش الورقة متطلبات البيانات وجودة المحتوى واستهلاك الطاقة والاعتبارات الأخلاقية، مع التأكيد على أهمية تطوير نماذج تتماشى مع الثقافة والقيم الأخلاقية العربية. وتختتم بتوصيات لتطوير نماذج لغوية عربية بشكل مسؤول، يوازن بين الأهداف والبيانات والطاقة والأخلاقيات.

لمياء هدريش بلغيث

أستاذة تعليم عالٍ وباحثة بكلية العلوم الاقتصادية والتصرف بجامعة صفاقس، ومديرة مدرسة الدكتوراه منذ 2021، ورئيسة فريق البحث ANLP-RG بمخبر MIRACL، تركّز أبحاثها على استخدام الذكاء الاصطناعي والمعالجة الآلية للغات الطبيعية لتحليل الآراء والمشاعر، والترجمة الآلية، وبناء الأنطولوجيات، والتعرّف على الكلام، وتحليل محتوى وسائل التواصل الاجتماعي لدعم اتخاذ القرار. آخر منشوراتها: لمياء هديش بلغيث، خالد شعلان: التطورات في التعلم الآلي ومعالجة اللغات الطبيعية: الابتكارات والتطبيقات (Springer, 2025).

تحليل الآراء في العربية الدارجة في عصر الذكاء الاصطناعي: بين التعقيد اللغوي والتحديات التقنية

مع التقدم السريع في مجال الذكاء الاصطناعي والاستخدام الواسع لوسائل التواصل الاجتماعي، أصبح تحليل الآراء أمرًا أساسيًا لفهم آراء المستخدمين من خلال تعليقاتهم باللهجات العربية. تستفيد اللغات الكبرى مثل الإنجليزية من محتوى وسائل التواصل الاجتماعي والموارد الرقمية والمعاجم في تدريب نماذج الذكاء الاصطناعي، ما يفسر الأداء المتفوق لهذه النماذج مقارنة بالعربية، ويواجه تحليل الآراء بالعربية تحديات لغوية وتقنية بسبب تنوع اللهجات وتعقيد البنية اللغوية. تقترح هذه الدراسة منهجية مبتكرة تجمع بين الأنطولوجيات والتعلم العميق لتحسين استخراج الآراء، مع التركيز على اللهجة التونسية. تعتمد المنهجية على جمع البيانات من وسائل التواصل الاجتماعي لبناء أنطولوجيات متخصصة وتطبيق التعلم العميق لتحليل الآراء.أظهرت نتائج التقييم قدرة عالية للنماذج على تصنيف الآراء بدقة تزيد عن 90٪ وفهم التعابير المعقدة. كما تناقش الدراسة التحديات الحالية وتقدم توصيات لدعم اللغة العربية في عصر الذكاء الاصطناعي، مع مواكبة أحدث التطورات في هذا المجال.

رئيس مركز الإيسيسكو للغة العربية للناطقين بغيرها بمنظمة العالم الإسلامي للتربية والعلوم والثقافة. وتولىّ سابقًا عددًا من المناصب الأكاديمية والبحثية، من أبرزها: رئاسة قسم اللغة العربية وآدابها، ونائب عميد المعهد العالمي للفكر الإسلامي والحضارة الإسلامية، ونائب عميد الدراسات العليا بكلية معارف الوحي والعلوم الإنسانية في الجامعة الإسلامية العالمية بماليزيا. وتتمحور اهتماماته العلمية حول اللسانيات التطبيقية، ودراسات الترجمة، وتعليم اللغة العربية للناطقين بغيرها. وله عدد من الأعمال المنشورة، من أهمها: التداخل اللغوي في وسائل التواصل الاجتماعي وتأثيره على الهوية اللغوية والثقافية للشباب العربي (جامعة الملك عبد العزيز، جدة، 2023).

تعليم اللغة العربية والذكاء الاصطناعي في ضوء السياسات التعليمية في ماليزيا

اعتمدت الحكومة الماليزية في صياغة سياساتها التربوية واللغوية مجموعةً من القرارات لتطوير جميع المناهج الدراسية، بما في ذلك مادة اللغة العربية، ودمجها تدريجيًا بالتقانة الحديثة وتقنيات الذكاء الاصطناعي في مختلف المراحل الدراسية. وقد جاءت هذه الدراسة للنظر في هذه السياسات التربوية ومناقشة متطلبات إدماج الذكاء الاصطناعي في تعليم اللغة العربية. وكشفت الدراسة عن الإستراتيجيات الرئيسة التي اعتمدتها الجامعات الماليزية لتشجيع معلمي اللغة العربية على استخدام تقنيات الذكاء الاصطناعي في فصولهم الدراسية؛ وهي توفير البنية التحتية الرقمية والأجهزة الإلكترونية اللازمة، ونشر ثقافة الوعي بالذكاء الاصطناعي عبر عقد دورات توعوية وتدريبية لتطوير مهارات التعامل مع الذكاء الاصطناعي، ومؤتمرات وندوات ولقاءات علمية تعنى بالذكاء الاصطناعي، فضلاً عن تخصيص منح بحثية لدعم البحث العلمي المرتبط بالذكاء الاصطناعي والتقانة الحديثة. وقد توصلت الدراسة إلى أن التجربة الماليزية في استخدام الذكاء الاصطناعي لم تنضج تمامًا، ولا تزال بحاجة إلى تحسين وتجويد، وبخاصة فيما يتصل بتأهيل المعلمين للتعامل إيجابيًا مع تقنيات الذكاء الاصطناعي واستخدامها في فصولهم الدراسية.

محمد ابن سفاج

باحث بسلك الدكتوراه، الوظائف: أستاذ اللغة العربية بالسلك الثانوي التأهيلي ومدرب نماذج الذكاء الاصطناعي، من اهتماماته العلمية: تدريسية اللغة العربية- التدقيق اللغوي- الذكاء الاصطناعي- تحقيق المخطوطات، من أعماله المنشورة: غاية المرام في مثلث الكلام لابن جابر الهَوَّاري (دراسة وتحقيق) - أبو جعفر الرعيني الغرناطي أول من ألف في المثلثات اللغوية القرآنية - الحسن الثاني شاعرًا.

العدول عن الرتبة الأصلية للكلمة وتأثيرُه في عملية الشكل الآلي عبر نماذج الذكاء الاصطناعي العربية (قلم ولسان وفَراسة نماذج)

يسعى هذا البحث إلى معالجة قضية الرتبة النحوية ومدى تأثيرها في جودة مُخرجات نماذج الشكل الآلي العربية، محاولا الإجابة عن إشكالية هي: كيف تؤثر الرتبة النحوية في عملية الشكل الآلي؟ وما الحلول التقنية واللسانية التي تحسّن الشكل الآلي وتجعله دقيقا وصحيحًا؟ ويُفترض أن تغير الرتبة النحوية داخل النص المُدخل يسبب للنماذج الشاكلة خللا في إخراجه مشكولا شكلا سليمًا، وأن للرتبة دورا مهما في إدراك النماذج الآلية للوظيفة النحوية والموقع الإعرابي للكلمة. وللتأكد من الفرضية، تمت دراسة ظاهرة الرتبة النحوية والنمذجة الآلية للشكل وفق المنهجين الوصفي والتحليلي، كما تم تفعيل المنهج التجريبي على مجموعة من المدخلات النصية غير المشكولة والمتغيرة رتبتُها، واختبار ثلاثة نماذج عربية تقوم بشكلها، وقد خلص البحث إلى أن الشكل الآلي عبر نماذج الذكاء الاصطناعي العربية المتخصصة في التدقيق اللغوي يعاني قصورا وضعفا كبيرا في مراعاة الرتبة النحوية الفرعية للكلمة، وأن هناك حلولا تقنية ولسانية يمكن أن تُتجاوز بها هذه المشكلة.

محمد اسماعيلي علوي

أستاذ التعليم العالي، ومدير مركز اللغات بجامعة السلطان مولاي سليمان المغرب. متخصص في اللسانيات. مدير مختبر الدراسات الأدبية واللسانية والديداكتيكية بنفس الجامعة. مشارك في العديد من المؤتمرات الدولية. عضو لجنة تحكيم جائزة مجمع الملك سلمان العالمي للغة العربية في دوراتها الثلاثة الأولى. خبير منظمة الإيسيسكو في مجال تدريب أساتذة العربية. فائز بجائزة الشارقة للدراسات اللغوية والمعجمية في دورتها السادسة 2023.

توظيف الذكاء الاصطناعي في تحليل وإعداد الموارد والبيانات اللغوية الموجهة إلى تعليم العربية للناطقين بغيرها: إشكالات منهجية وتطبيقات تربوية

تنطلق هذه الورقة من افتراض مفاده أن الذكاء الاصطناعي قادر على إعداد وتوليد مواد لغوية مناسبة لمتعلمي اللغة العربية الناطقين بغيرها، وتحترم المعايير المحددة في الأطر العالمية لتعليم اللغات. وتبقى الإشكالية المركزية من كل هذا هي: كيف يمكن توظيف الذكاء الاصطناعي في تحقيق هذا الهدف للتحقق من الفرضية، والإجابة عن السؤال المركزي أعلاه؟ وقد قسمنا الورقة إلى قسمين كبيرين؛ قسم نظري حددنا فيه أهم الأفكار والأسس النظرية التي تمكننا من فهم الانتقال من التطبيقات الحاسوبية في تعليم اللغات إلى عصر الذكاء الاصطناعي وإمكاناته التي يمكن الاعتماد عليها في تعليم اللغات وتعلمها. وقسم تطبيقي حددنا فيه بعض التطبيقات المتخصصة في تعليم اللغة العربية للناطقين بغيرها والمهام التي تقوم بها وتمكن المدرس من إعداد مواد لغوية مفيدة ومناسبة لأهدافه التربوية. كما قدمنا تطبيقات عملية اعتمدنا عليها في إعداد نماذج حقيقية من هذه المواد الموجهة لفائدة طلاب العربية من المستوى المتقدم؟

باحث مصري، أستاذ اللغويات بجامعة الأزهر، وخبير لغوي مشارك بمعجم الدوحة التاريخي للغة العربية، رأس اللجنة العلمية بالمعجم (2023 - 2024) ودرب عددًا من الفرق المعجمية، من مؤلفاته: ضوابط الفكر النحوي. البنية العاملية ومنازل الكلام في العربية: نحو لسانيات لتحليل الخطاب القرآني. رؤية الأشياء كما هي: مقاربة لسانية في تفكيك خطاب التحيز، حصل على جائزة كتارا للبلاغة القرآنية (2021) وجائزة الألكسو عن الدراسات العربية والمعجمية (2022).

بنيةُ الشاهد ومنهجُ التوثيق: معايير صيانةِ النصِّ في معجم الدوحة

أرسى معجم الدوحة بنيانَه على ثلاث ركائزَ كبرى تضبط مساره، وتوجّه منطقه العلمي: أولها: حرمةُ الشاهد المعجمي؛ فالنصّ يُصان من التصحيف والتحريف، عبر: مقابلة الروايات، وضبطُ مبنى الشاهد، وترجيح القراءة الأوثق، والتحاكم إلى معهود العرب في كلامها. ثانيها: حوكمة الدلالة؛ فلا نتعامل مع معجمٍ ساكن، بل مع لغة تجري في الزمن؛ الكلمة التي تبدأ حسّيّة في وصف الرمل والمطر، قد تنتهي عقليةً في كتب الكلام والفلسفة، أو تُصبح مصطلحًا دقيقًا عند الفقهاء والأطباء؛ فكان علينا أن نتتبّع هذا الامتداد الواسع، وأن نُعيد بناء السيرة الدلالية لكلّ لفظ. ثالثها: الحيادُ المعجمي؛ إذ يلتزم الخبير المعجمي حدودَ ""المعطى النصي"" مستقرئًا دلالتَه الداخلية، دون إقحام معان غريبة عنه، أو تأويلاتٍ مذهبيّةٍ أو ثقافيّةٍ خارجةٍ عن سياقه. ومن خلال هذه الثلاثية تمكن معجم الدوحة من ألا يمنح لأي لفظ موطنًا فيه إلا بعد التحقّق من سلامته رسمًا، وثبوت شاهده نصًّا، وانسجامه دلالة."

محمد العبيدي

رئيس وحدة الببليوغرافيا والمدوّنة في معجم الدوحة التاريخي للغة العربية، وأحد مؤسسي المشروع ونائب مديره التنفيذي، وعضو المجلس العلمي. يهتم بالبحث في الدراسات اللغوية والدلالية والصناعة المعجمية. وله عدد من الأعمال المنشورة، من أبرزها: دلالة السياق في القصص القرآني، والبحث الدلالي عند الشوكاني، ومن قضايا المعنى في صناعة المعجم التاريخي، والاقتراض اللغوي والتواصل الحضاري في ضوء معجم الدوحة التاريخي، والمعجم اللغوي التعليمي.

تَعزيز المحتوى الرّقميِّ العربيِّ بالموارِدِ المعجميّة في عَصرِ الذكاءِ الاصطناعيِّ معجم الدّوحة التّاريخيِّ للّغة العربيّة أنموذجًا

يواجه المحتوى الرقميّ العربيّ تحدِّياتٍ جوهريَّةً تتمثَّل في محدوديَّةِ حجمِه وضعفِ نوعيَّته مقارنةً باللغاتِ العالميَّةِ الأخرى؛ ويُعزى هذا الضعفُ إلى غيابِ بنيةٍ تحتيَّةٍ لغويَّةٍ رقميَّةٍ متكاملةٍ قادرةٍ على دعمِ تطبيقاتِ الذّكاءِ الاصطناعيّ. ومن هنا تنبعُ أهميَّةُ الموارِدِ المعجميَّةِ بوصفِها قاعدةَ بياناتٍ لغويَّةٍ شاملةٍ يمكنُ استثمارُها في تعزيزِ المحتوى الرقميّ العربيّ وتطويرِ المعالجةِ الآليَّةِ لِلّغة. ويُعدّ معجمُ الدوحةِ التاريخيُّ للغةِ العربيةِ أنموذجًا رائدًا في هذا المجال، بوصفِه أوّلَ مشروعٍ معجميٍّ عربيٍّ يُؤرِّخُ للألفاظِ ومعانيها عبرَ مدوَّناتٍ نصِّيَّةٍ ضخمةٍ تمتدُّ من النقوشِ العربيةِ القديمةِ حتى العصرِ الحديثِ. ويتَّسمُ المعجمُ بالتوسيمِ الصرفيِّ والدلاليِّ الدقيقِ، والتأصيلِ التاريخيِّ للألفاظِ والنظائرِ الساميةِ، والتوثيقِ الزمنيِّ للسياقاتِ النصِّيَّةِ، ما يجعلُه موردًا مهمًّا لتدريبِ الذكاءِ الاصطناعيِّ وتطبيقاتِه المختلفةِ. وتخلُصُ الدراسةُ إلى أنَّ معجمَ الدوحةِ التاريخيَّ يُسْهِمُ في تعزيزَ المحتوى الرقمي العربي، وإمدادِ الذكاءِ الاصطناعيِّ بمعطياتٍ مُؤرَّخةٍ ومُوَسَّمةٍ.

أستاذ التعليم العالي في اللسانيات. جامعة ابن زهر، المغرب. نائب مدير مختبر القيم والمجتمع والتنمية بجامعة ابن زهر، حاليا. اهتماماته العلميّة: التركيب الصرفي للغة العربية. اللسانيات الحاسوبية والإنسانيات الرقمية. السياسة اللغوية للعصر الرقمي. اللسانيات السامية والأمازيغية المقارنة. بعض أعماله المنشورة: Elghazi, M., & Elkabche, M. J. (2025). "The Future of the Arabic Language: Economic, Civilizational, and Religious Perspectives." In I. Oumrani (Ed.), Linguistics and Economics: Studies in Existing and Possible Relationships. Dar Konoz, Jordan. كتاب اللسانيات والعلاقات الدولية (2021). دار ركاز للنشر والتوزيع، الأردن. (ردمك: 978-9923-44-011-7).

نمذجة التعقيد الصرفي-التركيبي في الذكاء الاصطناعي العربي: نموذج "الحالة الحدية" لأداة النفي "ليس" وتعميماتها

تواجه النماذج اللغوية للذكاء الاصطناعي (مثل LLM) تحديات مهمة في التعامل مع بعض التراكيب الصرفية العربية المعقدة والتي تصنف ضمن "الحالات الحدية" في مجال هندسة البرمجيات والحواسيب؛ إذ غالبًا ما تفشل في فهم الأدوات الوظيفية ذات السلوك المزدوج، مما يؤدي إلى أخطاء في التحليل وتوليد النصوص. تركز هذه الورقة على أداة النفي "ليس" باعتبارها حالة دراسية تنطبق عليها صفات الأفعال من حيث التصريف، صفات الحروف من حيث وظيفتها النحوية. وتقترح هذه الدراسة نموذجًا تحليليًا جديدًا، على عكس النظرة التقليدية، يعتبر "ليس" كلمة مركبة نتجت عن اندماج تاريخي بين أداة النفي "لا" والفعل "أيسَ".هذا النموذج التفكيكي لا يقتصر على "ليس" وحدها، بل يوفر إطارًا منهجيًا لنمذجة أدوات النفي الأخرى كحزمات من السمات، مما يحسن دقة أنظمة الذكاء الاصطناعي. ويساهم هذا النهج في تطوير محللات صرفية ونحوية وأنظمة توليد نصوص أكثر كفاءة. وتخلص الدراسة إلى أن التحليل اللساني العميق هو حجر الزاوية لبناء نماذج ذكاء اصطناعي قادرة على فهم الفروق الدقيقة للغة العربية ومعالجتها بفعالية.

أستاذ اللسانيات بجامعة مولاي إسماعيل/المدرسة العليا للأساتذة. مدير مختبر "الدراسات اللسانية والأدبية والديداكتيكية"، ومدير سابق للمدرسة العليا للأساتذة. باحث متخصص في اللسانيات العامة، وعلم التركيب وقضايا المعجم والدلالة. من آخر أعماله المنشورة: "المعجم بين السياقات التركيبية والسياقات التداولية"، "النظر التركيبي في درس اللغة "، "المعجم العربي: قضايا ومقاربات لسانية".

الخصائص التركيبية للغة العربية ومعالجتها الآلية

يكشف هذا البحث عن العلاقة العميقة بين الدراسات النظرية اللسانية والتطبيقات العملية للمعالجة الآلية للغة، مبرزًا أن وصف أي لغة، ومنها العربية، ليس محايدًا، بل هو انعكاس للإطار النظري المعتمد. ومع تطور الذكاء الاصطناعي، القادر على معالجة اللغة إحصائيًا دون الرجوع إلى القواعد التقليدية، يطرح سؤال أساس: ما جدوى اللسانيات النظرية اليوم؟يجيب البحث بالتمييز بين هدفين متكاملين: هدف اللسانيات يتمثل في فهم اللغة بوصفها ظاهرة بيولوجية مرتبطة بالعقل البشري، بينما تهدف المعالجة الآلية إلى بناء تطبيقات ذكية قادرة على محاكاة اللغة وإنتاجها. ورغم اختلاف الغايات، يؤكد البحث أهمية التكامل بين المجالين، إذ تساعد النظريات اللسانية في توجيه النماذج الحاسوبية لحل مشكلات لغوية معقدة، فيما تتيح المعالجة الآلية اختبار الفرضيات اللسانية تجريبيًا. ويقترح البحث نموذجًا هجينًا يجمع بين قوة المحولات الإحصائية ودقة الأنظمة القاعدية لتطوير معالجة أكثر كفاءة للغة العربية.

محمد بباه

يعمل خبيرا حاسوبيا ورئيسا لوحدة الحوسبة بمشروع معجم الدوحة التاريخي للغة العربية. حاصل على شهادة الدراسات العليا المعمقة في التحليل الرقمي ومعالجة الإشارات (2005) وشهادة الدكتوراه في المعالجة الآلية للغات الطبيعية من جامعة محمد الأول بالمغرب(2013). شارك في بعض مشاريع البرمجيات الآلية للغة العربية كبرنامج المحلل الصرفي مفتوح المصدر (برنامج الخليل) والمشكّل الآلي مفتوح المصدر بالتعاون مع المنظمة العربية للتربية والثقافة والعلوم (ألكسو) ومدينة الملك عبد العزيز للعلوم والتقنية.

بيانات معجم الدوحة التاريخي للغة العربية موردا لتطوير تطبيقات الذكاء الاصطناعي العربي

يُعد ضعف أداء اللغة العربية في تطبيقات الذكاء الاصطناعي مقارنة بالإنجليزية أمرًا شائعًا في الأدبيات، إذ يظهر التفاوت في الترجمة، وتحليل النصوص والمشاعر، وفهم السياقات الثقافية واللغوية، والتعامل مع العبارات المجازية والالتباس اللغوي بمستوياته المختلفة.

يهدف هذا البحث إلى إبراز أهمية بيانات معجم الدوحة التاريخي للغة العربية في تطوير الذكاء الاصطناعي العربي، إذ توفر نحو 300 ألف مدخل معجمي، ومدونة نصية تقارب مليار كلمة، إضافة إلى ببليوغرافيا منظَّمة في قاعدة بيانات مهيكلة.

ويتناول البحث خمسة محاور رئيسية: وصف عام لبيانات المعجم، ودورها في التحليل الصرفي والتركيبي، وإسهامها في التحليل الدلالي وفك الالتباس، وخصائص منصة صناعة المعجم المرتبطة بالذكاء الاصطناعي، ثم عرض أولى محاولات توظيف هذه البيانات في تطوير نماذج الذكاء الاصطناعي للغة العربية.

محمد حاج إبراهيم

أستاذ مشارك بكلية دراسات اللغات الرئيسة في جامعة العلوم الإسلامية الماليزية. شغل سابقًا منصب منسق وحدة اللغة العربية بالجامعة، كما عمل مترجمًا ومذيع أخبار في إذاعة "هنا ماليزيا" التابعة لراديو وتلفزيون ماليزيا. تتركّز اهتماماته العلمية في البلاغة والنقد وتعليم اللغة العربية للناطقين بغيرها. ومن أبرز أعماله المنشورة: 1- تصميم نموذج التعليم الإلكتروني عبر تطبيقات زيب 2.0 للغة العربية للناطقين بغيرها لأغراض خاصة، سلسلة الدراسات التخصصية في مجال اللغة العربية للناطقين بغيرها للإيسيسكو ، 2- تطبيق تكنولوجيا الذكاء الاصطناعي تشات جي بي تي ومدى إثرها في تعليم اللغة العربية في الجامعات الماليزية، سلسلة الدراسات التخصصية في مجال اللغة العربية للناطقين بغيرها للإيسيسكو.

تعليم اللغة العربية والذكاء الاصطناعي في ضوء السياسات التعليمية في ماليزيا

اعتمدت الحكومة الماليزية في صياغة سياساتها التربوية واللغوية مجموعةً من القرارات لتطوير جميع المناهج الدراسية، بما في ذلك مادة اللغة العربية، ودمجها تدريجيًا بالتقانة الحديثة وتقنيات الذكاء الاصطناعي في مختلف المراحل الدراسية. وقد جاءت هذه الدراسة للنظر في هذه السياسات التربوية ومناقشة متطلبات إدماج الذكاء الاصطناعي في تعليم اللغة العربية. وكشفت الدراسة عن الإستراتيجيات الرئيسة التي اعتمدتها الجامعات الماليزية لتشجيع معلمي اللغة العربية على استخدام تقنيات الذكاء الاصطناعي في فصولهم الدراسية؛ وهي توفير البنية التحتية الرقمية والأجهزة الإلكترونية اللازمة، ونشر ثقافة الوعي بالذكاء الاصطناعي عبر عقد دورات توعوية وتدريبية لتطوير مهارات التعامل مع الذكاء الاصطناعي، ومؤتمرات وندوات ولقاءات علمية تعنى بالذكاء الاصطناعي، فضلاً عن تخصيص منح بحثية لدعم البحث العلمي المرتبط بالذكاء الاصطناعي والتقانة الحديثة. وقد توصلت الدراسة إلى أن التجربة الماليزية في استخدام الذكاء الاصطناعي لم تنضج تمامًا، ولا تزال بحاجة إلى تحسين وتجويد، وبخاصة فيما يتصل بتأهيل المعلمين للتعامل إيجابيًا مع تقنيات الذكاء الاصطناعي واستخدامها في فصولهم الدراسية.

محمد رقاس

خبير حاسوبي في معجم الدوحة التاريخي للغة العربية. حاصل على الدكتوراه في علوم الحاسب من جامعة محمد الأول بالمغرب وماجستير في هندسة المعلوميات. يتركز بحثه في المعالجة الآلية للغة العربية، خصوصا منصّات بناء المعاجم والمعاجم التفاعلية. ساهم في تأليف كتاب "نحو معجم تاريخي للغة العربية". ومن أبحاثه: "الهيكل الحاسوبي للمعجم التاريخي للغة العربية، التقنيات الحاسوبية في خدمة المعجم التاريخي للغة العربية"، "AMELIORATION OF THE INTERACTIVE DICTIONARY OF ARABIC".

بيانات معجم الدوحة التاريخي للغة العربية موردا لتطوير تطبيقات الذكاء الاصطناعي العربي

يُعد ضعف أداء اللغة العربية في تطبيقات الذكاء الاصطناعي مقارنة بالإنجليزية أمرًا شائعًا في الأدبيات، إذ يظهر التفاوت في الترجمة، وتحليل النصوص والمشاعر، وفهم السياقات الثقافية واللغوية، والتعامل مع العبارات المجازية والالتباس اللغوي بمستوياته المختلفة.

يهدف هذا البحث إلى إبراز أهمية بيانات معجم الدوحة التاريخي للغة العربية في تطوير الذكاء الاصطناعي العربي، إذ توفر نحو 300 ألف مدخل معجمي، ومدونة نصية تقارب مليار كلمة، إضافة إلى ببليوغرافيا منظَّمة في قاعدة بيانات مهيكلة.

ويتناول البحث خمسة محاور رئيسية: وصف عام لبيانات المعجم، ودورها في التحليل الصرفي والتركيبي، وإسهامها في التحليل الدلالي وفك الالتباس، وخصائص منصة صناعة المعجم المرتبطة بالذكاء الاصطناعي، ثم عرض أولى محاولات توظيف هذه البيانات في تطوير نماذج الذكاء الاصطناعي للغة العربية.

محمد غاليم

أستاذ بجامعة محمد الخامس. كاتب عام سابق لمعهد التعريب بنفس الجامعة، رئيس تحرير مجلة أبحاث لسانية. من مجالات اهتمامه: اللسانيات المعرفية، وفلسفة اللغة، واللسانيات التطبيقية. من أعماله: تصورات نحيا بها، دور المعرفة النواة في تأسيس اللغة، دار كنوز المعرفة، الأردن (2025)؛ عن إمكان التواصل وحدوده، في ضوء المبادئ المعرفية وفلسفة العمل، دار رقيم، الرياض (2025)؛ في اللسانيات التطبيقية وقضايا اللغة العربية، دار قناديل، بغداد (2025).

عن الذكاء الاصطناعي وخصائص اللغة العربية في الإطار المعرفي

يتناول البحث التضافر الحيوي المثمر الذي يجب أن يقوم ويتقوى في المجال العربي بين الذكاء الاصطناعي ومقتضيات بناء نماذجه، من جهة، واللسانيات المعرفية وما تكشف عنه من بنيات وخصائص كلية في نسق اللغة العربية، من جهة ثانية. ونرى تطوير ذلك كفيلا بتحقيق إدماج أكثر سلاسة للغة العربية في المجال الدولي الواسع للنمذجة الذكية الاصطناعية. كما يبين البحث أن تطوير نماذج ذكية اصطناعية أقوى وبمستويات محاكاةٍ بشرية أعمق يتطلب فهم الأسس المعرفية للذكاء وبلورة معايير دقيقة لقياسه. وأن من أبرز هذه الأسس المعرفية الجوهرية مبادئ أنساق المعرفة النواة التي تميز الذكاء البشري الذي يجب أن يشكل الإطار المرجعي الضروري للفهم ولمعايير القياس.

محمود الحاج

مدير مجموعة أبحاث معالجة اللغة الطبيعية في جامعة ڤين وأستاذ مشارك في كلية الهندسة وعلوم الحاسوب بجامعة فين في هانوي، فيتنام. يشغل حاليا منصب باحث زائر في جامعة لانكستر في المملكة المتحدة. حصل على درجة الدكتوراة من جامعة إسكس، حيث ركز بحثه على التلخيص متعدد الوثائق للغة العربية. تتضمن اهتماماته البحثية التلخيص، واستخراج المعلومات، ومعالجة اللغة الطبيعية المالية، وتطوير الموارد للغات محدودة الإمكانات.

عرب-جوبز: مدونة متعددة الجنسيات لإعلانات الوظائف العربية

عرب-جوبز هي مدونة واسعة النطاق ومُتاحة للعامة، تتألف من إعلانات وظائف عربية تم جمعها من مصر، الأردن، المملكة العربية السعودية، والإمارات العربية المتحدة. تحتوي المدونة على أكثر من ٨٥٤٦ إعلانًا وظيفيًا وأكثر من ٥٥٠ الف كلمة، مما يعكس التنوع الإقليمي واللغوي والاجتماعي-الاقتصادي في سوق العمل العربي. تتضمن هذه الدراسة تحليلات حول تمثيل النوع الاجتماعي، والتنوع اللهجي، والبنية المهنية، وتكشف عن فروقات منهجية بين الجنسين في كل من استخدام اللغة وإمكانية الوصول إلى الوظائف. نعرض في هذا العمل سلسلة من التجارب تشمل تقدير الرواتب وتطبيع تصنيفات الوظائف باستخدام نماذج لغوية كبيرة، بالإضافة إلى مهام معيارية لاكتشاف التحيز القائم على النوع الاجتماعي وتصنيف المهن. وتُظهر النتائج قابلية استخدام النماذج اللغوية الكبيرة في تحليل النصوص العربية، وتُشكل أساسا لأبحاث مستقبلية حول الإنصاف والتمثيل في معالجة اللغة العربية.

مراد دياني

محرّر علمي بالموسوعة العربية، ورئيس تحرير دورية استشراف للدراسات المستقبلية، المركز العربي للأبحاث ودراسة السياسات. حاصل على الدكتوراه في الاقتصاد من جامعة ستراسبورغ . تتركز أبحاثه الاقتصادية في مجالات اقتصاد المعرفة، إضافةً إلى بحوث تهمُّ نظريات العدالة والديمقراطية. شارك في العديد من المؤتمرات العلمية الدولية، وصدر له العديد من الأبحاث في مجلات محكّمة، وفي كتب جماعية بالعربية والفرنسية والإنكليزية. وقد صدر له عن المركز العربي العديد من الكتب المؤلّفة، والمحرّرة، والمترجمة، من آخرها مناهج البحث في العلوم الاجتماعية والإنسانية.

من المطبوع إلى الخوارزمية: الكتابة الرقمية المعزّزة بالذكاء الاصطناعي ووعود إعادة كتابة الإنتاج المعرفي العربي

تُسائل هذه الورقة البحثية الإمكانات الكامنة في ثورة الكتابة الرقمية المعزّزة بالذكاء الاصطناعي، أو "الثورة الرابعة في الكتابة"، للإسهام في درء العجز المعرفي العربي المُستدام، من خلال إعادة تشكيل القرائية، والتأليف، والتأويل، والتداول، على غرار التأثيرات البنيوية العميقة لثورة الطباعة المبكّرة. منهجيًا، تعتمد الورقة على التحليل التاريخي المقارن لتتبّع تطور القرائية عبر التدرّج من الشفهية إلى الكتابة، والطباعة، والرقمية، والذكاء الاصطناعي، مع التركيز على الزمانية، والتسلسل، والسببية البنيوية. وتحاجج الدراسة بأن الكتابة المعززة بالذكاء الاصطناعي تهجن الشفهية والقراءة والكتابة، وتمكّن الإنتاج المشترك بين الذكاء البشري والذكاء الاصطناعي، وتفتح قنوات بديلة ترخّص بتلافي السلطوية، والأبوية، والرقابة، وتخلق مساحات خطاب موازية، وتخفّض الحواجز أمام تأويل المعرفة، وإنتاجها، وإعادة إنتاجها، وتداولها، ومن ثمّ، تعيد تشكيل السلطة المعرفية والسلطة السياسية. كما تُبرز أنّ هذه الإمكانيات مشروطة بقرائية الذكاء الاصطناعي، وإمكانات توسيع قواعد البيانات والوصول إليها، والانفتاح المؤسسي، لأنها تحمل في الآن ذاته بذور ترسّخ المراقبة، والتحيّز الخوارزمي، والسلطوية الرقمية. وتخلص الورقة إلى أن التحول المدفوع بثورة الكتابة الرقمية المعزّزة بالذكاء الاصطناعي غير حتمي، ومحكوم بالتبعية للمسار.

نصرالدين مزاري

أستاذ محاضر بجامعة زيان عاشور بولاية الجلفة بالجزائر، صحفي سابق ومقدم أخبار بالإذاعة الجزائرية، مسؤول سابق للبيداغوجيا والدراسات، وعضو اللجنة العلمية بقسم الإعلام، وعضو مخبر الدراسات التاريخية والإنسانية، من اهتماماته: تحليل الخطاب الإعلامي والأدبي، اللغة العربية والتعليم العالي، والذكاء الاصطناعي، من كتبه ومقالاته الدولية: المهارات ما وراء المعرفية باستخدام الذكاء الاصطناعي في التعليم العالي، اللغة العربية في الإعلام، تشريعات الإعلام، أخلاقيات الذكاء الاصطناعي، خوارزميات التوصية واستهلاك الأخبار والمضامين.

تأثير الثنائية اللغوية (الفصحى والعامي) على أداء نماذج الذكاء الاصطناعي في فهم اللغة العربية: دراسة تجريبية مقارنة بين CAMeLBERT، AraBERT، وGPT-3.5

تهدف هذه الدراسة إلى محاولة سد فجوة بحثية تتعلق بفهم تأثير الثنائية اللغوية على أداء نماذج الذكاء الاصطناعي في معالجة اللغة العربية، وتفترض الدراسة أن التباين بين اللغة الفصحى والعامية يؤدي إلى تدهور جوهري في أداء النماذج، خاصة المدربة على الفصحى فقط، وباستخدام منهج تجريبي، تم اختبار ثلاثة نماذج لغوية بارزة: AraBERT، وCAMeLBERT، وGPT-3.5، على مهام تصنيف المشاعر والتعرف على الكيانات المسماة باستخدام مجموعات بيانات مصنفة يدويًا من كلا المستويين اللغويين، وكشفت النتائج عن انخفاض حاد في أداء نموذج AraBERT على النصوص العامية بنسبة تصل إلى 37.6%، مما يؤكد قصور النماذج المتخصصة في الفصحى، وفي المقابل، أظهرت النماذج الهجينة والمتعددة اللغات، مثل CAMeLBERT، مرونة أكبر مع انخفاض أقل في الأداء لم يتجاوز 12.1%، وتؤكد هذه النتائج أن دمج العامية في بيانات التدريب ليس خيارًا تكميليًا، بل ضرورة حتمية لتطوير نماذج ذكاء اصطناعي فعالة وقادرة على التعامل مع التنوع اللغوي للغة العربية، مما يمهد الطريق لتطبيقات أكثر شمولية.

نضال شمعون

أستاذ الفيزياء النظريّة في المعهد العالي للعلوم التطبيقيّة والتكنولوجيا بدمشق حاصل على دكتوراة في الفيزياء من أكسفورد، وماجستير في الرياضيّات من كامبردج، ودبلوم الإلكترونيّات من السوربون، وشهادة الهندسة من المدرسة المركزيّة باريس. متخصّص في فيزياء الجسيمات وعلوم الكون؛ له حوالي 50 ورقة بحثية، وحاز منحَ مؤسسة همبولدت الألمانيّة والأكاديمية الصينية للعلوم. عضو في مجلس إدارة الجمعية العربية للفيزياء والمجمع اللغوي بدمشق، ومترجم لكتب علمية، مثل "الكون الأنيق".

ثورة الذكاء الاصطناعي في نقل المعرفة، وخصوصية اللغة العربيّة في الترجمة العلميّة

تتناول الدراسة العلاقة المعقدة بين الذكاء الاصطناعي واللغة العربية، مركزة على تحديات وفرص الترجمة العلمية الآليّة. وتبتدئ بتحليل العوامل الكامنة وراء الهيمنة الحالية للإنجليزية في بيئة الذكاء الاصطناعي، والتي تشمل وفرة الموارد والكتلة الحرجة من الباحثين. في المقابل، تسلط الضوء على التحديات الهيكلية التي تواجه العربيةَ، كنقص البيانات عالية الجودة، وندرة المتخصصين، والاعتماد على الحلول الأجنبية، بالإضافة إلى الخصائص اللغوية الفريدة التي تزيد من تعقيد المعالجة الآلية، كالاشتقاق، وإشكالية التشكيل، وازدواجية اتجاه الكتابة، وتعدّد اللهجات. تستعرض الدراسة الحلولَ التقنية، خاصّةً النماذج العصبونية المتقدمة، التي تسمح بمعالجة قضايا كالفصل الذكي بين النص العربي والرموز العلمية اللاتينية داخل المستندات المختلطة. كما تطرح رؤيةً لاستغلال الذكاء الاصطناعي في التكيّف مع الفروقات المصطلحية بين الدول العربية، لضمان تلاؤم الترجمة مع السياق الثقافي والإقليمي للمُتلقّي. وأخيراً، تؤكد الدراسة على نموذج التكامل بين الآلة والمترجم البشري كضمانة للدقة العلمية والجودة اللغوية، مع توصياتٍ لتعزيز مكانة العربية في المشهد الرقمي، عبر توحيد الجهود وبناء الموارد اللغوية وتعزيز التعاون بين التخصّصات.

هند بنت سليمان الخليفة

أستاذة تقنية معلومات في جامعة الملك سعود ورئيسة مجموعة "إيوان" البحثية. تُعدّ ضمن أعلى 2% من العلماء استشهادًا على مستوى العالم. وقد نشرت أكثر من 200 ورقة بحثية بمعامل 42، إلى جانب براءتي اختراع أمريكيتين. ألّفت كتبًا عن معالجة اللغة العربية الطبيعية والذكاء الاصطناعي، وحازت على جوائز دولية عدة. كما تشغل مناصب قيادية في لجان علمية دولية ومحلية.

النمذجة اللغوية الآلية للعربية: دراسة تحليلية شاملة لنماذج اللغة الكبيرة والتحديات اللسانية

تشهد معالجة اللغات الطبيعية العربية تطورًا متسارعًا مع ظهور نماذج اللغة الكبيرة، والتي تُعيد تشكيل المشهد التقني للذكاء الاصطناعي اللغوي. تقدم هذه الدراسة التحليلية مراجعة منهجية شاملة لـ34 دراسة أساسية تغطي الفترة من 2019 إلى 2025، مركزة على تطوير وتقييم نماذج اللغة الكبيرة المخصصة للغة العربية ولهجاتها المتنوعة. تكشف النتائج عن تقدم ملحوظ في التقنيات المعمارية، حيث تم تطوير 36 نموذجًا رئيسيًّا موزعًا بين نماذج التشفير (22 نموذج) ونماذج فك التشفير (10 نماذج) والنماذج المختلطة (4 نماذج). رغم الإنجازات المتحققة، تواجه المنطقة العربية تحديات جوهرية تشمل التركيز المفرط على العربية الفصحى، ونقص البيانات عالية الجودة للهجات العربية، وضعف مستوى الشفافية والوصولية في النماذج المطورة.

يحيى محمد الحاج

خبير حاسوبي حاصل على شهادة الدكتوراه في علوم الحاسب. عمل باحثا في مختبرات أوربية مختلفة. درَّس في كلية الحاسب بجامعة الإمام بالسعودية تسع سنوات، انخرط خلالها في مختلف الأنشطة العلمية. شارك مع نخبة من الباحثين العرب في إطلاق مبادرة لتعزيز استخدام اللغة العربية في مجال الحاسب، تمخض عنها مؤتمر دولي ينعقد باستمرار ومجلة تابعة له. حصل على منحة للتميز البحثي من الصندوق العربي للإنماء بالكويت، قبل أن يلتحق بالمعجم.

بيانات معجم الدوحة التاريخي للغة العربية موردا لتطوير تطبيقات الذكاء الاصطناعي العربي

يُعد ضعف أداء اللغة العربية في تطبيقات الذكاء الاصطناعي مقارنة بالإنجليزية أمرًا شائعًا في الأدبيات، إذ يظهر التفاوت في الترجمة، وتحليل النصوص والمشاعر، وفهم السياقات الثقافية واللغوية، والتعامل مع العبارات المجازية والالتباس اللغوي بمستوياته المختلفة.

يهدف هذا البحث إلى إبراز أهمية بيانات معجم الدوحة التاريخي للغة العربية في تطوير الذكاء الاصطناعي العربي، إذ توفر نحو 300 ألف مدخل معجمي، ومدونة نصية تقارب مليار كلمة، إضافة إلى ببليوغرافيا منظَّمة في قاعدة بيانات مهيكلة.

ويتناول البحث خمسة محاور رئيسية: وصف عام لبيانات المعجم، ودورها في التحليل الصرفي والتركيبي، وإسهامها في التحليل الدلالي وفك الالتباس، وخصائص منصة صناعة المعجم المرتبطة بالذكاء الاصطناعي، ثم عرض أولى محاولات توظيف هذه البيانات في تطوير نماذج الذكاء الاصطناعي للغة العربية.

رؤساء الجلسات

حسين الزراعي

خبير لغويّ أوّل في معجم الدّوحة التاريخيّ. أستاذ التعليم العالي في اللسانيات، رئيس قسم اللغة العربية الأسبق في كلية الآداب جامعة صنعاء (2009-2011)، حاصل على الدكتوراه من جامعة محمد الخامس (2003)، ودرجة الأستاذية من جامعة صنعا (2015). عمل في جامعة صنعاء(2004-2011) وجامعة الملك خالد (2011-2013)، وله أكثر من 20 بحثاً وكتاباً في مجالات لسانية مختلفة، في التركيب والدلالة والمعجم والصرف والتداوليات وغيرها. له في المُنجَز المعجميّ: نظرية في (بناء الكلمة وتحليلها)، كتاب منشور بالجزائر (2011). والمعجم النّسقي المحوسب، منشور في مجلة أم القرى 2012. رئيس فِرَق إعداد المعاجم المدرسية (2008): مركز البحوث العربي بدولة الكويت، إشراف مركز البحوث والتطوير التربوي باليمن.

خالد الجبر

حاصل على الدّكتوراة في النّقد والبلاغة (الأردنيّة 2002م)، والماجستير في الأدب والنّقد (1992م)، والبكالوريوس (النّجاح 1987م). درَّس في جامعاتِ فلسطين والأردن وقطر 25 عامًا، رأَسَ قسم العربيّة بجامعة البترا، وعَمَدَ كلّيّة الآداب بالجامعة الإسلاميّة. مقرّر اللّجان: الاستشاريّة، والتّنفيذيّة لمعجم الأدباء، والتفرّغ الإبداعيّ (الأردن). أصدرَ 21 كتابًا منها 9 محكَّمة، و25 بحثًا منشورًا. كتبَ 300 مدخل لموسوعة "أعلام العلماء والأدباء العرب والمسلمين". أعدَّ وقدَّم 5 برامج تلفزيونيّة. التحق بمعجم الدّوحة بتاريخ 2016. وهو عضو وحدة الاعتماد واللجنة العلميّة.

 عبد العلي الودغيري	حاصل على دكتوراه الدولة في العلوم اللغوية، وتابع دراساته العليا في جامعتي الصُّوربون بباريس وجامعة محمد بن عبد الله بفاس، وجامعة محمد الخامس بالرباط. عمل أستاذًا للعلوم اللغوية متخصّصاً في الدراسات المعجمية، بجامعتي محمد بن عبد الله بفاس ومحمد الخامس بالرباط (أكدال). تولى منصب رئيس الجامعة الإسلامية العالمية بالنيجر (تابعة لمنظمة المؤتمر الإسلامي) من 1994 إلى 2005، ومدير مؤسسة علال الفاسي بالرباط من 2006 إلى 2008.
 عبد المنعم حرفان	حصل على دكتوراه الدولة في اللسانيات من جامعة سيدي محمد بن عبد الله بفاس (المغرب). عمل بها أستاذا باحثا، ورئيس قسم اللغة العربية ومدير مختبر للبحث اللساني: "الأبحاث المصطلحية والدراسات النصية"، تولى سابقا إدارة تحرير مجلة ورقات، وعضوية هيئة تحرير المجلة الدولية "التواصل اللساني"، وعضو هيئة تحرير مجلة آفاق أدبية. تتركز اهتماماته البحثية على الدراسة اللسانية لظواهر اللغة العربية الصرفية والتركيبية والنحوية والمعجمية. له عدد من المؤلفات والأبحاث المنشورة في مجالات اللسانيات والصرف والمعجم والنحو. التحق بمعجم الدّوحة التّاريخيّ للّغة العربيّة خبيرا متعاونا منذ العام 2015 حتى 2025.

رؤساء الجلسات

 محمد الشيباني	أستاذ تعليم عال بجامعة صفاقس - الجمهوريّة التونسّية. تقلد عدة وظائف من بينها: رئيس لقسم اللّغة العربيّة بكليّة الآداب والعلوم الإنسانيّة بصفاقس، عضو بمجلسها العلميّ وبالمجلس العلميّ لمخبر مقاربات الخطاب، عضو بالمجلس العلميّ لفريق ترجمة الدّراسات اللّسانيّة والمعجميّة بالمركز الوطنيّ للترجمة بتونس، رئيس تحرير لمجلّة "أصداء الجامعة" (جامعة صفاقس)، وعضو هيئة تحرير "مجلّة الفكر الجديد". التحق بمعجم الدوحة التّاريخي سنة 2015، وهو حاليّا خبير لغويّ أوّل ورئيس وحدة الدّراسات والأبحاث فيه. له إصدارات عدة، منها: أسس تولّد معاني الكلام وتأويلها في النّظريّة البلاغيّة العربيّة (2021).
 محمد بباه	يعمل خبيرا حاسوبيا ورئيسا لوحدة الحوسبة بمشروع معجم الدوحة التاريخي للغة العربية. حاصل على شهادة الدراسات العليا المعمقة في التحليل الرقمي ومعالجة الإشارات (2005) وشهادة الدكتوراه في المعالجة الآلية للغات الطبيعية من جامعة محمد الأول بالمغرب(2013). شارك في بعض مشاريع البرمجيات الآلية للغة العربية كبرنامج المحلل الصرفي مفتوح المصدر (برنامج الخليل) والمشكّل الآلي مفتوح المصدر بالتعاون مع المنظمة العربية للتربية والثقافة والعلوم (ألكسو) ومدينة الملك عبد العزيز للعلوم والتقنية.

 محمد رقاس	خبير حاسوبي في معجم الدوحة التاريخي للغة العربية. حاصل على الدكتوراه في علوم الحاسب من جامعة محمد الأول بالمغرب وماجستير في هندسة المعلوميات. يتركز بحثه في المعالجة الآلية للغة العربية، خصوصا منصّات بناء المعاجم والمعاجم التفاعلية. ساهم في تأليف كتاب "نحو معجم تاريخي للغة العربية". ومن أبحاثه: "الهيكل الحاسوبي للمعجم التاريخي للغة العربية، التقنيات الحاسوبية في خدمة المعجم التاريخي للغة العربية"، "AMELIORATION OF THE INTERACTIVE DICTIONARY OF ARABIC".
 محمد محمود أحمد محجوب	حصلَ عَلى دكتوراه الدّولةِ في اللّسانياتِ منْ جامعةِ محمد بنِ عبدِ اللهِ بفاس–المغربِ. التحقَ بجامعةِ نواكشوط عضوا في هيئة التدريس. درَّسَ سابقًا في جامعةِ الإماراتِ، والمعهد العالي للدراسات والبحوث الإسلامية (موريتانيا). أدارَ مركزَ (البحوثِ والإنماءِ) بمنطقةِ أبوظبي التّعليميةِ 2005-2006، وشَغَلَ منصبَ الأخصائيِّ الأوّل لمناهجِ اللغةِ العربيةِ في مجلسِ أبوظبي للتّعليمِ 2012-2015. له إسهاماتٌ نقديةٌ ولغويةٌ وتربويةٌ متنوّعةٌ. شاركَ في العديد من المؤتمرات والندوات العلمية كان من آخرها: ندوة "قضايا اللغة العربية واللسانيات التطبيقية - الدوحة (2017)". من بحوثِهِ المنشورةِ: (مفهومُ اللغةِ بينَ سوسير وتشومسكي)، (والنَّسقُ المصطلحيُّ البلاغيُّ)... انضمّ إلى الفريقِ المركزيّ لمعجمِ الدّوحةِ التاريخيّ في 2015/04/01م، ومن ضمنِ أنشطتهِ فيه تقديمُه دورةً في "الصناعةِ المعجميّةِ" لفريق المعالجة المعجمية في موريتانيا 2017م.

رؤساء الجلسات

 مهدي عرار	دكتوراه في اللسانيات والعلوم اللغوية، وأستاذ كرسي الشيخ تميم بن حمد آل ثاني للغة العربية وآدابها، حاصل على جوائز عالمية وعربية متعددة. له (16) كتابا منشورا بيروت، وما يزيد على 45 بحثا علميا منشورا في مجلات علمية ومؤتمرات محكمة، وقد عمل نائب رئيس جامعة بيرزيت المساعد، للدراسات العليا والبحث العلمي، وأستاذا معتمدا في جامعة جورج تاون، واشنطن، ورئيس قسم اللغة العربية في جامعة الشارقة، وعميد كلية الآداب في جامعة بيرزيت.
 يحيى محمد الحاج	حاصل على شهادة الدكتوراه في علوم الحاسب. عمل باحثا في مختبرات أوربية مختلفة. درَّس في كلية الحاسب بجامعة الإمام بالسعودية تسع سنوات، انخرط خلالها في مختلف الأنشطة العلمية. شارك مع نخبة من الباحثين العرب في إطلاق مبادرة لتعزيز استخدام اللغة العربية في مجال الحاسب، تمخض عنها مؤتمر دولي ينعقد باستمرار ومجلة تابعة له. حصل على منحة للتميز البحثي من الصندوق العربي للإنماء بالكويت، قبل أن يلتحق بالمعجم.

معجم الدوحة التاريخي
للغــة العربيــة