اسم البحث	توصيف العربية حاسوبياً لغير الناطقين بها -إشكالات وتحديات –
اسم الباحث	أ. د. محمد بن عبد العزيز العميريني
جامعة الباحث	أستاذ اللغة العربية بجامعة محمد بن زايد للعلوم الإنسانية
ملخص البحث باللغة العربية	يُعد التعليم الإلكتروني مفهوماً حديث الظهور على سطح عالم التعليم الحديث، مواكباً لظهور شبكات الإنترنت وانتشارها المذهل في التطبيقات التعليمية المختلفة، وتبرز في الساحة المعلوماتية مصطلحات عدة تحاول وصف هذه الظاهرة التكنولوجية التعليمية حديثة الظهور، منها: التعليم الإلكتروني، والتعليم المحوسب، والتعليم الافتراضي، والتعليم عبر الإنترنت، ونحو ذلك من المصطلحات التي تتباين صياغتها، لكنها تشير في الوقت ذاته إلى جميع أنواع الأنشطة التعليمية التي يتم تصميمها وتنظيمها وإيصالها من خلال شبكة الإنترنت. وأشارت الدراسات المختلفة إلى أهمية توظيف التكنولوجيا في تعليم العربية لغير الناطقين بها، إذ يحقق ذلك مبدأ التفاعل الجيد بين الأستاذ والمتعلّم، وتجاوز مرحلة التلقين والطرح الجاف للمعلومة، كما أنه يحدث تحسناً كبيراً في اتجاهات التلاميذ نحو التعلم. هذا الاتجاه التقني في التعليم يُمكّن من إدارة تعليم اللغة وتعلمها في ضوء برمجيات تعليمية، ومقررات إلكترونية نشطة من أجل إكساب المتعلمين مهارات اللغة، لتحقيق التواصل اللغوي البناء، والتعامل الدقيق مع كل مقوماته. الكلمات المفتاحية: توصيف، الحاسوب، إشكالات، تحديات، توظيف
ملخص البحث باللغة الإنجليزية	Electronic learning (e - learning) has spread widely and internationally to be side by side with the traditional learning. Among these e- learnings are: complicational learning, virtual learning and distance learning. Many educational applications have been developed to be used to teach various types of educational activities that were developed through networks. Many studies show the importance of technology in learning and teaching Arabic to non native arabic speakers. This contributes in creating a positive interaction between teachers and learners. It also helps to transfer teaching Arabic from repetitive to learner center classroom. This new technical direction in education enables Arabic language institutions to produce many electronic programs that help learners to develop new skills and acquire

Classification : Public

Arabic language easily. Key Words: Description,Computer,Problems,Challenges,Function.	

اسم البحث	التصنيف الموضوعي للنص العربي بإستخدام التصنيف غير المتوازن: مواضيع القرآن الكريم نموذجاً
اسم الباحث	أ.د. أكرم محمد زكي –د. بسام عركوك
جامعة الباحث	الجامعة الإسلامية العالمية بماليزيا - جامعة حضرموت
ملخص البحث باللغة العربية	يعد التصنيف الموضوعي من المواضيع الحيوية في الوقت الحاضر بسبب كثرة المواد المنتشرة في الفضاء الالكتروني. حيث يهدف التصنيف الموضوعي إلى وضع عنوان مناسب للنص بناء على معايير ومحددات يتم استخلاصها من النص. وقد وصل التصنيف الموضوعي إلى مراحل متقدمة في اللغة الانجليزية ولا يزال يحتاج إلى المزيد من البحث والتطوير في اللغة العربية. وهناك الكثير من التقنيات والخوازميات التي تساعد في هذا الشأن ومنها تقنيات التصنيف غير المتوازن. وقد تم تطبيق تقنيات التصنيف غير المتوازن على نطاق واسع في مجال التنقيب عن البيانات. حيث يتم استخدامها لتصنيف الفئات غير المتوازنة والتي لا تتساوى في عدد عيناتها. وتكمن مشكلة الفئات غير المتوازنة في أن أداء التصنيف يميل إلى الفئة التي تحوي مزيداً من العينات بينما تحصل الفئة ذات العينات القليلة على أداء ضعيف. يمكن أن تحدث هذه المشكلة في التصنيف القرآني بسبب الاختلاف في عدد الآيات للمواضيع القرآنية المطلوب تصنيفها. صنفت العديد من الدراسات البحثية المواضيع القرآنية والتي اعتمدت على التصنيف التقليدي. وفي المقابل، هناك دراسات قليلة صنفت الموضوعات القرآنية على اساس تقنيات التصنيف غير المتوازن. لذلك، تهدف هذه الدراسة إلى التصنيف الموضوعي للنص العربي (متمثلاً بآيات من القرآن الكريم) باستخدام بعض تقنيات التصنيف غير المتوازن ومقارنة أدائها مع طريقة التصنيف التقليدي. ومن ثم تحديد أفضل طريقة من بين هذه التقنيات في تصنيف آيات القرآن الكريم. استعرضت هذه الدراسة أيضاً الأدبيات التي ستناقش الطرق السابقة التي تعتمد على قواعد البيانات القرآنية والمصنفة حسب أساليب مختلفة. ويليه سيتم تقديم منهجية البحث لتوضيح خطوات ومراحل عملية التصنيف. وأخيراً سيتم الكشف عن النتائج والتجارب

Classification : Public

	لاستخلاص بعض النتائج وإظهار أداء تصنيف المصحف الشريف باستخدام التصنيف غير المتوازن أو بدونه. تعمل هذه الدراسة على تطبيق طرق التصنيف غير المتوازن لتصنيف المواضيع القرآنية غير المتساوية في عدد آياتها. في هذه الدراسة استخدمنا التقينة الأشهر في التصنيف غير المتوازن والتي هي: (SMOTE). حيث تم استخدام العديد من المقاييس لتقييم النتائج التجريبية، والتي أظهرت تحسن في النتائج في أداء التصنيف القرآني عند تطبيق تقنية SMOTE. الكلمات المفتاحية: التصنيف الموضوعي، التصنيف غير المتوازن، مواضيع القرآن الكريم، تقنيات إعادة التوازن، تصنيف النصوص.
ملخص البحث باللغة الإنجليزية	Imbalanced classification techniques have been widely applied in the field of data mining. It is used to classify imbalanced classes that do not have the same number of samples. The problem with imbalanced classes is that the classification performance tends to the majority class that are with more samples, while the class with few samples will obtain performance poorly. This problem can be occurred in the Quranic classification due to the difference in the number of verses of Quranic topics to be classified. Many research studies have classified Qur'anic topics, which are based on the traditional classification. However, there are few studies that classify Qur'anic topics based on the imbalanced classification techniques. Therefore, this paper aims to apply imbalanced classification methods to classify Qur'anic topics that are unequal in the number of their verses. In this research, we used the most famous technique in the imbalanced classification area, which is: (SMOTE). Several metrics were used in this research to evaluate the experimental results. These metrics are Sensitivity, Specificity, G-mean, MCC. The results showed an improvement in the performance of Quranic classification when applying the SMOTE technique. **Keywords**: Theme Classification, Imbalanced Classification, Quranic Topics Re-sampling techniques, Text Classification.

اسم البحث	رقمنة اللغة العربية وتحدّياتُ الانخراط في عالم الذَّكاء الاصطناعي: توظيف الأسنادِ الرقمية في تعليم العربية للناطقين بغيرها
اسم الباحث	أ.د بلقاسم الجطاري
جامعة الباحث	أستاذ في جامعة محمد بن زايد للعلوم الإنسانية
ملخص البحث باللغة العربية	تستعرض هذه الورقة ما يتيحه الإعلام الرقمي من فرص لتطوير تعليم اللغة العربية للناطقين بغيرها، وهي فرص كثيرة واعدة يمكن ملامستها من خلال تأمل تجارب دولية ومؤسسية مختلفة، نجحت في خلق جيل جديد من التعليميات الخاصة باللغات، بل وتغلبت على كثير من العراقيل التي كانت تقف أمام بلوغ غاياتها بهذا الخصوص، كما يمكن الوقوف عليها من خلال تأمل ما يزخر به الفضاء الرقمي من وسائط وإمكانات تفاعلية تسرع التعلمات وتختصر المسافات وتقلص الميزانيات. هذا وتناقش الورقة أيضا، من خلال تتبع منهجي موصول إلى حقلي: تعليمية اللغات واللسانيات الاجتماعية، بعضَ التحديات التي تفرضها الغاية المذكورة، إذ تضع الأصبع على بعض الصعاب التي تعيق بلوغ الأهداف المرسومة، منطلقة في ذلك من خلفية إبستيمية تنظر إلى ثورة الرقميات بعين ناقدة، توازن بين مقادير الكلفة من جهة، وبين مبالغ العائدات المادية والمعنوية والقيمية من جهة ثانية، وذلك بالنظر إلى المتاح الاقتصادي والحضاري للأمة العربية، كما تضع الانفجار الرقمي الحاصل راهنا في مشرح التحري الموضوعي، على نحو تبحث فيه عن السبل الكفيلة بجعل هذا الانفجار أداة للرفع من تنافسية اللغة العربية في سوق اللغات المطلوب تعلمها عبر العالم. **كلمات مفتاحية:** اللغة العربية، الحوسبة، تعليم، لسانيات تطبيقية، ديداكتيك
ملخص البحث باللغة الإنجليزية	This a paper is premised on the theories of artificial intelligence, basically in the achievements of computational linguistics that are cognitively integrated with natural language engineering and cognitive science. The latter is based, in essence, on the findings of research in linguistics, mathematics, psychology and physiology, which are sciences that have known great breakthroughs in recent years. Although we are still in the early stages of big data theories, it has become possible to program the computer with algorithms derived from the image of digitized linguistic blogs to perform many linguistic operations and provide practical solutions to various issues. Indeed, digitization has made it possible to analyse, deduce, classify, and evaluate language statistical data. In line with this, and in order to benefit from experiences in this field, this presentation will be focused on the digitization of the language of Dhad, given the importance it enjoys in the field of language applications via the computer. Key Words:

Classification : Public

Arabic language, computing, education, applied linguistics, didactic.	

اسم البحث	مزالق الترجمة الآلية دراسة تطبيقية على نصوص مختارة
اسم الباحث	أ.د. صديق جوهر
جامعة الباحث	وزارة شؤون الرئاسة
ملخص البحث باللغة العربية	في ظل سياق العولمة أصبحت الاتصالات بين البلدان والثقافات متكررة بشكل متزايد مما اقتضى استخدام بعض التقنيات الإليكترونية للمساعدة في ترجمة اللغات المختلفة من أجل التواصل الإنساني ومنها الترجمة الآلية أو الترجمة بمساعدة الحاسوب. تعتمد دراسة الترجمة الآلية على ثلاثة تخصصات هي اللغويات والرياضيات وعلوم الحاسوب والترجمة الآلية هي فرع من فروع علم اللغة الحاسوبي الذي يستخدم الآليات المحوسبة لترجمة السياق أو المحادثة من لغة بشرية إلى أخرى. لقد تم طرح مفهوم الترجمة الآلية لأول مرة بواسطة الباحث وارين ويفر في عام 1947 بعد عام واحد فقط من تطوير أول جهاز حاسوب رقمي إلكتروني وحاسوبي متكامل. منذ ذلك الحين اعتبرت الترجمة الآلية واحدة من أصعب المهام في مجال معالجة اللغة الطبيعية. وعلى الرغم من أن الترجمة الآلية لاتزال تواجه عوائق متنوعة لا يمكن إنكار أن الترجمة بمساعدة الحاسوب قد حققت تقدماً هائلاً خاصة بعد الاستعانة بتطبيقات الذكاء الاصطناعي وتطوير شركة جوجل وغيرها من الشركات لأنظمة الترجمة العصبية في القرن الحادي والعشرين. في هذا السياق تستطلع هذه الورقة البحثية من خلال استخدام نظريات الترجمة الآلية المعاصرة بوصفها إطاراً منهجياً لتطور أنظمة الترجمة الآلية منذ أربعينيات القرن المنصرم حتى السنوات الأخيرة مع تسليط الضوء على الترجمة الآلية القائمة على البيانات والترجمة الآلية القائمة على الأمثلة والترجمة الآلية الإحصائية والترجمة الآلية العصبية من أجل رصد مزايا وعيوب الترجمة الآلية من خلال التطبيق على نصوص متنوعة ومختارة. خلصت الدراسة إلى تأكيد وجود قصور لا مندوحة عنه في برامج الترجمة الآلية خاصة عند الترجمة من اللغة العربية إلى اللغات الأوروبية. وتؤكد الدراسة عدم قدرة أنظمة الترجمة الآلية على التعامل بكفاءة مع العديد من النصوص خاصة الأدبية والدينية والتراثية والتاريخية. ورداً على الادعاءات المتكررة عن القدرة الخارقة لآليات الذكاء الاصطناعي التي أدخلت على أنظمة الترجمة الآلية نؤكد على صعوبة اختراع مترجم فوري آلي في الوقت الراهن واستحالة الاستغناء عن المترجمين البشريين حتى لو استطاع العلماء إنشاء أنظمة ترجمة آلية عصبية متطورة تطابق الدماغ البشري في المعرفة العامة والقدرة على التفكير لأن العلماء حتى الآن لم يستطيعوا فك شفرات الدماغ البشري نفسه فكيف يمكنهم اختراع آلة تحاكي الدماغ البشري الذي لم يكتشفوا سوى بعض تعقيداته وخباياه؟ الكلمات المفتاحية: الترجمة الآلية - الذكاء الاصطناعي -الترجمة القائمة على القواعد النحوية - الترجمة العصبية - مزالق الترجمة الآلية.
ملخص البحث باللغة الإنجليزية	In the context of globalization, contacts between countries and cultures have become increasingly frequent, which necessitated the use of some electronic technologies to help translate different languages for human communication, including the use of machine translation or computer-assisted translation. The study of machine translation is based on

Classification : Public

three disciplines: linguistics, mathematics, and computer science. Machine translation is a branch of computational linguistics that uses computerized mechanisms to translate context or conversation from one human language to another. The concept of machine translation was first introduced by researcher Warren Weaver in 1947 just one year after the development of the first integrated electronic and digital computer. Since then machine translation has been considered one of the most difficult tasks in the field of natural language processing. Although machine translation still faces various obstacles, it is undeniable that computer-assisted translation has made tremendous progress, especially after the use of artificial intelligence applications and the development of Google and other companies for neural translation systems in the twenty-first century. In this context, this paper explores, through the use of contemporary machine translation theories as a methodological framework, the development of machine translation systems from the 1940s until recently, with a focus on data-based machine translation, example-based machine translation, statistical machine translation and neural machine translation in order to monitor advantages and disadvantages of machine translation by applying MT dynamics to various selected texts. The study concludes by emphasizing the existence of an inevitable shortcoming in machine translation programs, especially when translating from Arabic into European languages. The study confirms the inability of machine translation systems to deal efficiently with many texts, especially literary, religious, cultural and historical texts. In response to the repeated claims about the extra-ordinary potentialities of artificial intelligence mechanisms that have been introduced to machine translation systems, the paper stresses the difficulty of inventing an automated interpreter at the present time and the impossibility of dispensing with human translators even if scientists create advanced neural machine translation systems that match the human brain in general knowledge and ability to think. Since scientists have not yet been able to decipher the human brain itself, how can they invent a machine that simulates the human brain, which they have only discovered some of its intricacies and mysteries?
Keywords: machine translation, artificial intelligence, rules-based translation, neural translation, machine translation scandals.

Classification : Public

اسم البحث	نحو تطوير برمجيات حرة ومفتوحة المصدر في خدمة تطبيقات المعالجة الآلية للغة العربية
اسم الباحث	أ. د. عمر مهديوي
جامعة الباحث	جامعة مولاي إسماعيل مكناس- المملكة المغربية
ملخص البحث باللغة العربية	يندرج هذا البحث في إطار برمجيات معالجة اللغات الطبيعية عامة، ومعالجة اللغة العربية خاصة، ذلك أن المعالجة الحاسوبية للغات تعدُّ مجالا معرفياً بينيًا، يقع في مفترق الطرق بين اللسانيات والمعلوميات والذكاء الاصطناعي والعلوم المعرفية. وتنحصر أهدافه في بلورة برامج وتقنيات حاسوبية متطورة جدًّا، الغاية منها معالجة المستويات اللِّسانية للِّغات البشرية تحليلاً وتوليدًا. وسعيا منا في تضييق الفجوة الرقمية بين العربية واللغات الأجنبية في مجال الحوسبة اللغوية، فقد خصَّصنا هذا البحث للتعريف بالبرمجيات الحرة ومفتوحة المصدر من جهة، ومن جهة أخرى الوقوف عند دورها باعتبارها آلية لدعم المحتوى الرقمي للغة العربية وتعزيزه على الشابكة، وبالتالي تبيان وظيفيتها في خدمة التطبيقات الحاسوبية للغة العربية عامة. **الكلمات المفاتيح:** برمجيات حرة- مفتوحة المصدر -المحتوى الرقمي -التطبيقات الحاسوبية- اللغة العربية.
ملخص البحث باللغة الإنجليزية	This research falls within the framework of natural language processing software in general, and Arabic language processing in particular, because computer processing of languages is an interdisciplinary field, located at the crossroads between linguistics, informatics, artificial intelligence, and cognitive science. Its objectives are limited to the crystallization of highly advanced computer programs and techniques, the purpose of which is to address the linguistic levels of human languages, analyzing and generating. And in our endeavor to narrow the digital gap between Arabic and foreign languages in the field of linguistic computing, we have devoted this research to defining free and open-source software on the one hand, and on the other hand standing at its role as a mechanism to support and enhance the digital content of the Arabic language on the net, and thus demonstrate its functionality in serving Computational Arabic language applications in general.

Classification : Public

Keywords: free software - open source - digital content - computer applications - Arabic language.	

اسم البحث	التحقق من النطق الصحيح باستخدام تقنية Computer-Aided Language Pronunciation (CALP) دراسة تطبيقية على القرآن الكريم
اسم الباحث	أ.د. محسن عبد الرازق رشوان - د. نعيم محمد عبد الغني
جامعة الباحث	جامعة القاهرة
ملخص البحث باللغة العربية	ينطلق البحث من تعريف تقنية التحقق من الأخطاء، ثم يبدأ في شرح كيفية بناء نموذج معياري للنطق الصحيح، ثم يرصد تقييم تطبيقي حفص وسالم اللذين اعتمدا على تقنية التحقق من النطق الصحيح باستخدام تقنية computer-aided language pronunciation (CALP)، وكان التقييم آليًا أولًا ثم بعدها بمقارنة المخرجات مع العقل البشري، ثم برصد أثر التقنية في التعليم بشكل واقعي عملي، وأخيرًا من خلال تقييم التطبيقين مقارنة بالتطبيقات المشابهة المتاحة الآن على الهواتف المحمولة. **كلمات مفتاحية:** النطق الصحيح – التحقق - التلاوة – أخطاء – الذكاء الاصطناعي – التعليم – التقنية.
ملخص البحث باللغة الإنجليزية	The research starts with the definition of error verification technology, then begins to explain how to build a standard model for correct pronunciation. The research work monitors the evaluation of Hafss and Salem applications, which relied on the technique of verifying correct pronunciation using Computer Aided Language Pronunciation technique (CALP), and the evaluation was automated first, then by comparing the outputs with the human, then by monitoring the impact of technology on education in a realistic and practical way. Finally, the two applications were compared to similar applications now available on mobile phones **Keywords:** correct pronunciation - verification - recitation - errors - artificial intelligence - education - technology.

Classification : Public

اسم البحث	تصنيف صوتيمات اللغة العربية باستخدام نموذج طرف إلى طرف في سبيل كشف أخطاء النطق
اسم الباحث	م. إلهام المفشي - د. أميمة الدكاك
جامعة الباحث	المعهد العالي للعلوم التطبيقية والتكنولوجيا
ملخص البحث باللغة العربية	أصبح تعلم اللغات بمساعدة الحاسوب (Computer-Aided Language Learning (CALL موضوعاً هاماً بسبب الحاجة المتزايدة لتعلم لغات جديدة، وبعد كشف وتحديد النطق الخاطئ جزءاً مكملاً لأنظمة تعلم اللغات. لاقت هذه النظم اهتماماً كبيراً لأنها تمكّن متعلّمي اللغة من تحسين قدراتهم اللغوية دون الحاجة للتواصل مع المختصين اللغويين بشكل مباشر، وذلك بالاستفادة من وسائل التعلّم الحديثة والتقنيات المتطورة. يهدف هذا البحث لتحسين تعرف الصوتيمات phonemes بالاعتماد على نموذج طرف إلى طرف، والذي يلغي الحاجة لعملية المحاذاة القسرية والتقطيع للأصوات يدوياً. قمنا باختبار أداء المنهجية المقترحة والمقارنة مع نتائج بحثنا السابق، وحصلنا على نتائج أفضل بالنسبة لتعرف الصوتيمات. إن نسبة التعرف الجيدة تمكّن من توظيف هذا النموذج في أنظمة تعليم اللغات وتفتح المجال لاستخدامها في تطبيقات أخرى مثل الإملاء الآلي. ركزنا في البحث على اللغة العربية بهدف تقليص الهوة البحثية الموجودة بين التقانات اللغوية الداعمة للّغة العربية ومثيلاتها في اللغات العالمية والتي حققت تقدماً كبيراً في العديد من المجالات. **الكلمات المفتاحية:** تعلّم اللغة بمساعدة الحاسوب، كشف أخطاء النطق، التعرّف على الصوتيمات، شبكات التعلّم العميق، نماذج طرف إلى طرف.
ملخص البحث باللغة الإنجليزية	Computer-aided language learning (CALL) has become an important issue due to the increasing need to learn new languages. Mispronunciation detection and identification are considered important components of language learning systems. These systems have received a great attention because they enable language learners to improve their language proficiency without the need for direct communication with language specialists, using of modern learning methods and advanced technologies. This research aims to improve automatic phoneme

Classification : Public

recognition, using an end-to-end model which eliminates the need of phoneme alignment and segmentation. The performance of the proposed methodology was tested and compared with our previous system results, and we obtained better phoneme recognition rates. These good recognition results enable the use of this model in language learning systems and allow its use in other applications such as speech-to-text systems. The focus of the research was on Arabic language, aiming to reduce the research gap between the linguistic technologies supporting Arabic language and in the other international languages, which have achieved great progress in many fields.

Keywords: Computer-Assisted Language Learning CALL, Mispronunciation Detection, Phoneme recognition, Deep neural networks, End-to-End Models

Classification : Public

اسم البحث	إشكالات المعالجة الآليّة للغة العربيّة وآفاق تطويرها بنك المشجرات العربيّة أنموذجا
اسم الباحث	د. سندس كرونة
جامعة الباحث	جامعة قرطاج
ملخص البحث باللغة العربية	نهتمّ في هذه الورقة بتقديم بنك المشجرات العربيّة في جامعة بنسلفانيا upenn Arabic Treebank تقديما نقديّا والنظر في إمكانيات تطويره. ويندرج هذا البحث في إطار سعينا إلى النقل المعرفي لمنهجيّة بنك المشجرات بجامعة بنسلفانيا من أجل استثمارها في تطوير بنك المشجرات العربيّة وتجاوز نقائصه. والهدف الأسمى من ذلك تحقيق وسم نحوي آلي للغة العربيّة يراعي خصوصياتها اللسانيّة. ومن أبرز الإشكاليات التي نطرحها في هذا البحث: كيف يمكن تطبيق منهجيّة بنك المشجرات Treebank Methodology في الوسم النحوي الآلي للغة العربيّة مع مراعاة خصوصياتها اللسانية كما قدّمت في نظريّة النحو العربي من جهة والكليّات اللغويّة التي اعتمدت في منهجيّة التحليل الآلي في جامعة بنسلفانيا من جهة ثانية؟ وإلى أيّ مدى تمكّنت منهجيّة بنك المشجرات العربيّة في صيغته الحاليّة من تحقيق هذه المعادلة واجتناب الاختزاليّة التي قد تتعارض مع الوصف اللساني الصحيح؟ وما هي سبل تطوير هذه الطريقة في المعالجة الآلية للغة العربيّة وإعادتها إلى موطنها الأصلي في مراكز البحث والجامعات العربيّة المهتمّة باللسانيات الحاسوبيّة؟ يتكوّن البحث من ثلاثة أقسام : نمهّد في القسم الأوّل بتقديم مشروع بنك المشجرات العربيّة تقديما موجزا من حيث نشأته وأسسه النظرية والمنهجيّة. ثمّ نهتمّ في القسم الثاني بعرض بعض الإشكالات التي واجهها فريق بنك المشجرات العربيّة في مرحلته الأولى بسبب عدم الإلمام الدقيق بخصائص النحو العربي. ونختم عملنا بالإشارة إلى بعض الهنات التي مازالت قائمة في الصيغة الحاليّة لبنك المشجرات العربيّة والتي يمكن تجاوزها بتطوير هذه المنهجيّة في مراكز بحث وجامعات عربيّة مختصّة في المعالجة الآلية للغات. يمثّل هذا البحث محاولة لتوطين منهجية بنك المشجرات بجامعة بنسلفانيا في البحث اللساني العربي في مجال المعالجة الآلية للغة العربيّة وتطوير سبل رقمنتها وتوظيف المدوّنات النصيّة العربيّة المحشّاة annotated في الترجمة الآليّة من اللغة العربيّة وإليها واستخراج المعلومات وتطوير البرمجيات والتطبيقات المعتمدة في تعليم اللغة. **كلمات مفتاحيّة** بنك المشجرات العربيّة –مبادئ النحو العربي –منوال العمل والربط –تحشية –المحلّل الصرفي –المحلّل التركيبي
ملخص البحث باللغة الإنجليزية	In this paper, we are aiming at making a critical presentation of the Penn Arabic Treebank at researching possibilities for its development. This research falls within the framework of our endeavor to transfer knowledge by investing the methodology of the Treebank at the University of Pennsylvania in developing the Arabic Treebank and overcoming its shortcomings to achieve an automated grammatical analysis of the Arabic language that takes into account its linguistic peculiarities. Among the most important problems that we will address: How can the Treebank Methodology be applied, taking into consideration the peculiarities of the Arabic language as presented in the theory of Arabic grammar on the one hand, and the linguistic schools that were adopted in the automated analysis methodology at the University of Pennsylvania on the other hand? To what extent was the methodology of the Arab Treebank in its current form able to achieve this equation and avoid the reductionism that

Classification : Public

may conflict with the correct linguistic description of Arabic syntactic structures? What are the ways to develop the methodology of Arabic Treebank and return it to its original homeland in Arab research centers and universities interested in computational linguistics?

The research consists of three sections: In the first section, we briefly introduce the Arab Treebank project in terms of its creation circumstances and its theoretical and methodological backgrounds. Then, in the second section, we present some of the problems faced by the Arabic Treebank team in its first stage due to a lack of accurate knowledge of the characteristics of Arabic grammar. We conclude our work by referring to some of the challenges that still exist in the current version of the Arabic Treebank, which can be overcome and discuss the possibility of developing this methodology in Arab research centers and universities specialized in the field of Natural Language Processing.

This research represents an attempt to settle the Arab Treebank in Arabic linguistic research in the field of automatic processing of the Arabic language, developing ways to digitilize it, and employing annotated Arabic text blogs in the field of machine translation, information extraction, software development and applications used in language education.

Keywords

Arabic Treebank - principles of Arabic grammar – Government and Binding model - annotation - morphological analyzer - syntactic analyzer.

Classification : Public

اسم البحث	التعابير المسكوكة في اللغة العربية من التوصيف اللساني إلى المعالجة الحاسوبية
اسم الباحث	د. شائق عبد الرحمان
جامعة الباحث	-
ملخص البحث باللغة العربية	يندرج هذا البحث ضمن حقل اللسانيات الحاسوبية، ويتوخى وضع تصور لساني - حاسوبي لبناء قاعدة بيانات لتعابير اللغة العربية المسكوكة وتحويلها إلى نموذج صوري يتم ضبطه والتحكم فيه بواسطة الحاسوب. وتتجلى أهمية البحث في هذه الظاهرة اللغوية، في النقاط التالية: • وجود مواقع إلكترونية غربية اهتمت بحوسبة هذه الظاهرة اللغوية في لغات متعددة، كاللغة الإنجليزية واللغة الفرنسية واللغة الألمانية، في غياب تام لموقع عربي يهتم بجمع ونشر هذه التعابير التي هي جزء لا يتجزأ من متن اللغة العربية. • بناء قاعدة بيانات تضم هذه التعابير قديمها وحديثها، ونشرها على الشبكة العنكبوتية ليعرفها ويتعلمها الناطقون باللغة العربية والناطقون بغيرها؛ فهي جزء مهم من تعلم لغة ما ومعرفة استعاراتها. • المساهمة في إثراء المحتوى الرقمي العربي. ننطلق في معالجة هذه الظاهرة اللسانية آليا من الإشكالية التالية: **كيف يمكن للحاسوب أن يفهم التعابيرالمسكوكة وينتجها؟.** تفرعت عن هذه الإشكالية مجموعة من الأسئلة التي سنجيب عنها، وهي: ما المقصود بالتعابير المسكوكة؟ ما سماتها المميزة؟ ما التوصيف اللساني المقترح لمعالجتها حاسوبيا؟ وما الصعوبات التي ستواجه الحاسوب عند معالجتها؟ وما الحلول المقترحة لحل هذه الصعوبات؟ وكيف سنستثمر نظرية الإطارات في بناء قاعدة بيانات آلية للتعابير المسكوكة تستطيع بواسطتها توليد معطياتها اللغوية على مستويي الفهم والإنتاج؟. وللإجابة عن هذه الأسئلة المتعلقة بالمستويين النظري والتطبيقي للبحث، سنتبع المنهجية التالية: • **الجمع**: جمع متن من التعابير المسكوكة. • **الوصف:** وصف الخصائص اللسانية للتعابير المسكوكة. • **التطبيق:** إعداد نظام معلوماتي يستطيع فك رموز التعابير المسكوكة عند معالجتها آليا. **الكلمات المفتاحية:** التعابير المسكوكة، التوصيف اللساني، اللسانيات الحاسوبية، التمثيل الآلي، قاعدة البيانات، نظرية الإطارات.

Classification : Public

ملخص البحث باللغة الإنجليزية	This study falls within the field of computational linguistics. It intends to develop a computational linguistic visualization to build a database of Arabic idioms and convert them into a graphic model that is computer-tuned and controlled. The importance of this study is reflected in the following points: • The existence of Western websites that are interested in computerizing this linguistic phenomenon in multiple languages, such as English, French and German, while the absence of an Arabic website that is interested in collecting or publishing these idioms. • Build a database of old and modern idioms and publish it on the Web so that Arabic and non-Arabic speakers know and learn them, as they are an important part of language learning. • Contribute to enriching Arabic digital content. The study's main problem is: How can a computer understand and produce idioms? To answer this question, we first have to answer some other sub-questions, including: What is meant by idioms? What are their distinctive features? What is the suggested linguistic description for its computational processing? What are the difficulties that the computer will face when dealing with idioms? What are the proposed solutions to solve these difficulties? How will we invest frame theory in building an automated database of idioms through which they can generate their linguistic data at the understanding and production levels? To answer these questions, we will follow the following methodology: • Data collection: collect a corpus of idioms. • Description: describe the linguistic characteristics of idioms. • Application: develop an information system that can decode collocations when processed automatically. Keywords: Idioms, linguistic characterization, computational linguistics, automated representation, database, frame theory.

Classification : Public

اسم البحث	إطار عامّ مرجعيّ لتوظيف التقنيَّة في تعليم اللُّغة العربيَّة وتعلُّمها
اسم الباحث	د. عيسى صالح الحمادي
جامعة الباحث	جامعة محمد بن زايد للعلوم الإنسانية
ملخص البحث باللغة العربية	تُعَدّ التغيرات المتسارعة الناجمة عن التقدُّم العلميّ والتِّقنيّ سمةً مميّزةً لهذا العصر؛ لذا أصبح من الضروري مُواكَبة العَمليَّة التربويَّة بمختلف جوانبها لهذه التغيرات، ومواجَهة المشكلات والتحدّيات التي قد تنجم عنها، متمثلةً في: كثرة المعلومات، وزيادة أعداد الطلاب، ونقص المعلّمين المؤهَّلينَ، وبُعد المسافات، وازدياد الحاجة إلى التعليم، وقد أدت هذه التغيرات إلى ظهور أنماط عديدة للتعليم الفرديّ، أو الذاتيّ، ولمواجَهة هذه التغيرات؛ ظهرت مفاهيم: التعليم المبرمَج، والتعليم باستخدام الحاسوب، والتعليم عن بُعد. إن توظيف التقنيات في التعليم أضحى خيارًا إستراتيجيًّا لكل دولة تُقدِّر -بوعي- ما للقطاع التربويّ من أهميَّة بالغة في ازدهار المجتمع وتطوره التنموي، وبديلًا لا يمكن الاستغناء عنه ليس في الظروف الاستثنائيَّة، لكن -أيضًا- لبناء جيل جديد قادر على الاستفادة من معطيات الثورة الصناعيَّة الرابعة، يمتلك من المعارف والمهارات التي تُمكِّنه من الانخراط في وظائف الغد. والتعليم إذا استَثمَر التقنيات الحديثة، ووظَّفَها في مختلف جوانب العَمليَّة التعلُّمِيَّة تخطيطًا، وإدارة، وتدريسًا، وتقويمًا يمكن أن يحقق أهدافه إذا توافرت الشروط لإنجاحه؛ حتى يمكن صياغة سياسات علميَّة مناسِبة للتعامل مع المستقبل. والبحث الحالي يُركِّز على بناء إطار مرجعيّ لتوظيف التقنيَّة في تعليم اللُّغة العربيَّة وتعلُّمها، ويشمل إعداد منظومات معايير: البيئة الرَّقْميَّة لتعليم اللُّغة العربيَّة، ومعلم اللُّغة العربيَّة الرَّقْميّ، ومتعلم اللُّغة العربيَّة الرَّقْميّ، ومعايير تدريس فنون اللُّغة الرَّقْميَّة، والدليل المقترَح لمُعلِّمي اللُّغة العربيَّة ومُتعلِّمها؛ لتوظيف التقنيَّة في تعليم اللُّغة العربيَّة وتعلُّمها، في ضوء الإطار العامّ المرجعيّ، ومعايير التقويم الإلكترونيّ في اللُّغة العربيَّة. الكلمات المفتاحية: الإطار المرجعيّ. توظيف التقنيَّة. اللُّغة الرَّقْميَّة. البيئة الرَّقْميَّة. البيئة الرَّقْميَّة. فنون اللُّغة الرَّقْميَّة. الكفاءة الرَّقْميَّة. العصر الرَّقْميّ.
ملخص البحث باللغة الإنجليزية	The rapid changes resulting from scientific and technical progress are a distinctive feature of this era. Therefore, it has become necessary to keep pace with the educational process in its various aspects to these changes. To face the problems and challenges that may result from it, represented by: the abundance of information, the increase in the number of students, the lack of qualified teachers, the distance, and the increased need for education. These changes have led to the emergence of many patterns of individual or self-education, and to confront these changes; Concepts emerged: programmed education, computer education, and distance education. Employing technology in education has become a strategic choice for every country that consciously appreciates the importance of the educational sector in the prosperity and development of society, and an indispensable alternative not in exceptional circumstances, but also - to build a new generation capable of benefiting from the data of the industrial revolution. Fourth, he possesses the knowledge and skills that will enable him to engage in the jobs of

Classification : Public

tomorrow. Education, if it invests in modern technologies and employs them in various aspects of the learning process in planning, management, teaching, and evaluation, can achieve its goals if the conditions are met for its success so that appropriate scientific policies can be formulated to deal with the future. The current research project focuses on building a frame of reference for employing technology in teaching and learning the Arabic language, and it includes the preparation of standards systems: the digital environment for teaching Arabic, the digital Arabic language teacher, the digital Arabic language learner, the grammatical standards for teaching Arabic, and the Linguistic Art Guide; To employ technology in teaching and learning the Arabic language, in light of the general frame of reference and the standards of electronic assessment in the Arabic language. **Key Words:** reference frame. Technology employment. The digital language. The digital environment. Numerical Language Arts. Numerical efficiency. The digital age	

اسم البحث	تمثيل النصوص العربية وتصنيفها بناء علي نهج ثنائي متواليات الحروف الهجائية
اسم الباحث	فاطمة الغنام
جامعة الباحث	معهد بحوث الالكترونيات، القاهرة، مصر
ملخص البحث باللغة العربية	نظرا للتنامي المتزايد في أعداد الملفات عبر شبكة الانترنت أصبحت عملية التصنيف الالي للنصوص أحد الموضوعات البحثية المهمة في مجال معالجة اللغات الطبيعية. يستخدم التصنيف الالي للنصوص في كثير من التطبيقات مثل إدارة المحتوى ، البحث السياقي ، التنقيب عن الآراء ، تحليل مراجعة المنتجات، تصفية الرسائل غير المرغوب فيها ، واستخراج المشاعر النصية. لتصنيف مجموعة من النصوص آليا، يلزم أولاً تحويل النصوص الي صورة رقمية مناسبة لخوارزميات التعلم الآلي. الطريقة الشائعة لاستخراج سمات النصوص وتمثيلها رقميا هي حقيبة الكلمات. لكن تتمثل المشكلة الرئيسة لهذه الطريقة في الأبعاد العالية لمساحة متجه السمات والذي يتناسب تناسبا طرديا مع أعداد الكلمات الموجودة بالمستندات ، فضلاً عن الحاجة إلى أدوات وبرامج تعتمد على معالجة اللغة الطبيعية والتي تمثل تحد كبير في حالة اللغة العربية. تشرح هذه الورقة نهجا جديد لاستخراج سمات النصوص وتمثيلها رقميا يعتمد علي الحروف الهجائية. هذا النهج يتبني الفرضية المعروفة باسم رمزية الصوت التي تنص على الارتباط الظاهر بين متواليات صوتية معينة ومعاني معينة في الكلام. النهج المتبع في هذه الورقة يقدم اسهامين رئيسين في مجال تصنيف النصوص. أولاً، أظهرنا إمكانية استخدام سمات ثابتة تستند إلى الحروف الهجائية دون الحاجة إلى المفردات الموجودة بالمستندات؛ وهذا يساهم بشكل كبير في تقليل الأبعاد العالية لمساحة متجه السمات. ثانيًا، لا يتطلب هذا النهج إستخدام أدوات وبرامج معقدة لمعالجة اللغة الطبيعية، ولا سيما اللغات عالية التصريف والاشتقاق كما في اللغة العربية. لقد أثبت العمل الحالي القدرة على تصنيف مجموعات من

Classification : Public

	المستندات النصية العربية بنجاح. حيث أظهر ما يقرب من 80٪ توفيرًا في مساحة متجه الملفات وتحسنًا في الأداء بنسبة 2٪ مقارنة بأفضل النتائج الأخري المسجلة لتصنيف مجموعة بيانات الجزيرة الإخبارية. الكلمات المفتاحية: معالجة اللغات الطبيعية، تصنيف النصوص، تمثيل النصوص رقمياً ، الوثائق العربية، الحروف الهجائية ، خوارزميات التعلم الآلي، رمزية الصوت.
ملخص البحث باللغة الإنجليزية	In text classification, Texts must be converted into numerical representations that the learning algorithms can understand. The high dimensions of vector space and the requirement for language-specific tools are two major issues with the widely utilized bag of words technique. In the present study, a new alphabetic approach is used to construct feature terms and applied it to classify Arabic documents. The proposed approach has two main contributions to text classification area. First, we have demonstrated the possibility of using constant feature terms that are based on the standard alphabet without the need for the documents vocabularies; this definitely helps in reducing the dimensions of the vector space for large corpus. Second, it does not require natural language processing tools. The current work has proved the ability to classify collections of Arabic text documents successfully. It showed approximately 80% savings in vector space and 2% performance improvement compared to the best recorded results on Aljazeera News dataset. Keywords: Natural language processing, Text classification, Text representation, Arabic documents, Alphabet, Machine learning algorithms, Phonemic symbolism.

Classification : Public

اسم البحث	"الأصوات العربية" وتعليمها للناطقين بغيرها تطبيق Arabic Phonetics نموذجًا على الهواتف المحمولة والحواسيب اللوحية IPhone/IPad
اسم الباحث	د. محمد الغليض
جامعة الباحث	جامعة عين شمس
ملخص البحث باللغة العربية	غالبًا ما يواجه الطلاب الناطقين بغير اللغة العربية الفصحى المعاصرة تحديات في اكتساب النظام الصوتي للغة العربية؛ مما قد يشكل غالبًا تحديًا للمعلمين أثناء عمليات انتقال هؤلاء الطلبة من اللغة الأم إلى اللغة المستهدفة أثناء التعلم. على المستوى الصوتي ، هناك عدد محدود من أبحاث تحليل التباين بين اللغتين العربية والإنجليزية ، وندرة مثل هذه الأبحاث بين العربية والصينية والفرنسية والألمانية والهوسا والإندونيسية والروسية والإسبانية والتركية والأردية. يحتاج المتحدثون بهذه اللغات إلى تطبيق يساعد في تعلم الصوتيات العربية وعلى وجه الخصوص الأصوات التي ليس لها مكافئ في لغاتهم. من ناحية أخرى، نتيجة لوباء Covid-19، هناك رغبة في تطوير أدوات التعلم التكنولوجي لاستخدامها في أي مكان بدلاً من الفصول الدراسية العادية. نتيجة للعمل المكثف لإيجاد حل لهذه الحاجة وبالاستعانة بأدوات المنهج الوصفي، قمنا بتطوير تطبيق جوال يملأ هذه الثغرة. في هذه الورقة، سيتم تقديم الحلول لهذه التحديات من خلال تطبيقنا للأصوات العربي ACOCO. هذا التطبيق تم تصميمه لتحسين نطق اللغة العربية لغير الناطقين بها. يركز ACOCO على "العربية الدولية" ، بما في ذلك العربية من الشرق الأوسط وشمال أفريقيا. التطبيق متوفر حاليًا على متجر آبل: https://apps.apple.com/app/id1482546188 . ختامًا، توصي الورقة بالعمل على تقييم التطبيق وتقويمه؛ للوقوف على جوانب القصور، ومحاولة معالجتها بطريقة منهجية تتفق وأساليب صناعة التطبيقات والأصدة اللغوية، وطبيعة اللغة العربية. الكلمات الافتتاحية: تعليم اللغة العربية بوصفها لغة ثانية، اللسانيات الحاسوبية، التطبيقات اللغوية، علم الأصوات العربية.
ملخص البحث باللغة الإنجليزية	Non-native students of Modern Standard Arabic often face challenges in acquiring the phonology system of Arabic. Processes of transfer from mother tongue to the target language often take place in learning which could constitute a challenge for educators. At the phonological lever, there is limited number of contrastive analysis researches between Arabic and English and scarcity of such research between Arabic and Chinese, French, German, Hausa, Indonesian, Russian Spanish, Turkish and Urdu. Speakers of these languages need a program that helps in learning Arabic phonemes and in particular the sounds they do not have in their languages. On the other hand, as a result of Covid-19 pandemic, there is an urge to develop

technological learning tools to be used at anyplace rather than regular classroom setting. As a result of extensive work on finding solution for this need and by using a descriptive-analytical method, we have developed a mobile app that fills this lacuna. In this paper, solutions to these challenges will be presented and our Arabic phonetic App will be introduced. The ACOCO App. is designed to improve Arabic pronunciation of non-native speakers of Arabic. ACOCO focuses on 'international Arabic', including Arabic from The Middle East and North Africa. It is available on Apple store: https://apps.apple.com/app/id1482546188. Recommendations of the study: Further work on evaluating and refinement of ACOCO, to identify any shortcomings, and to address them in a systematic manner consistent with the methods of Apps processing and corpus linguistics. This would include observing further the modern structure and use of the Arabic language. Keywords: Teaching Arabic as a second language, computational linguistics, linguistic applications, Arabic phonology.	

اسم البحث	استكشاف أداء نموذج التميمي المحسن للتوسيم النحوي الآلي على سور من القرآن الكريم
اسم الباحث	د. أفراح التميمي
جامعة الباحث	جامعة الإمام محمد بن سعود الإسلامية
ملخص البحث باللغة العربية	نموذج التميمي المحسن للتوسيم النحوي نموذج مطور من نموذج أولي يستعمل خوارزمية الحقول العشوائية المشروطة Conditional Random Fields (CRF) إحدى خوارزميات تعلم الآلة الموجه المستعملة لبناء نماذج احتمالية تمييزية تتنبأ بوسوم كلمات متتالية مع أخذ السياق بعين الاعتبار. وقد درب النموذج على مدونة من 60 ألف كلمة تقريبا موسمة يدويا بأقسام الكلام. ويستعمل هذا الموسم النحوي مجموعات وسوم تنطلق من نظرية تمام حسان في تقسيم الكلام العربي. وتأتي مجموعة الوسوم الأساسية فيه على سبعة أقسام هي: الأسماء – الأفعال – الصفات - الضمائر- الظروف -الأدوات - الخوالف، بالإضافة إلى التقسيمات غير اللغوية التي يستلزمها التحليل النصي الحاسوبي، وهي: علامات الترقيم – الاختصارات - الكلمات الأجنبية - الرموز. وتعتمد هذه الورقة على المنهج التجريبي لاستكشاف أداء نموذج التميمي المحسن للتوسيم النحوي بمجموعة الوسوم الأساسية على سور من القرآن الكريم؛ سعيا لتوسيم القرآن الكريم به كاملا. وأظهرت النتائج بعد توسيم تلك السور بهذا النموذج، ثم مراجعتها يدويا، بأن الأفعال شكلت ما نسبته 18% تقريبا من المدونة العينة، وأن النموذج نجح في توسيم 1846 فعلا من 2081 فعلا، محققا نسبة من الصحة بلغت 89%. إن هذه النسبة قد تجعل من أهمية تحسين أداء وسم الأفعال حسب خصوصية القرآن الكريم أولوية في مشاريع توسيم القرآن الكريم مستقبلا، وفي تحسين أداء النموذج في جميع مجموعات وسومه. **الكلمات الدلالية** التوسيم النحوي؛ النموذج المحسن؛ مجموعة وسوم؛ تدريب؛ اختبار؛ خوارزمية؛ مدونة

Classification : Public

The enhanced Tamimi's model for grammatical tagging is an improved model from an initial model that uses Conditional Random Fields (CRF). It is one of the supervised machine learning algorithms used to build discrimination probability models that predict a sequence of words taking context into account. This model has trained on a corpus that is consisted of approximately 60,000 tokens manually annotated with parts of speech. It uses tagsets based on Tammam Hassan's theory of parts of speech. The basic tagset is seven parts of speech: nouns, verbs, adjectives, pronouns, adverbs, particles, and exclamations, and non-linguistic categorical required by automated textual analysis: punctuations, abbreviations, foreign words, and symbols. This paper depends on the experimental method to explore the performance of the enhanced automated Tamimi's model in a sample of the Holy Quran; to tag the Holy Qur'an in its entirety with all tagsets of the model. After tagging the sample corpus with this model and correcting manually, the results show that verbs constituted approximately18% and the model correctly tagged 1846 verbs from 2081, achieving a percentage of accuracy it stood at 89%. This percentage may make the importance of improving verb tags according to the specificity of the holy Qur'an a priority in future Qur'an tagging projects, and in improving model performance. Keywords POS tagging; enhanced model; tag set; training set; test set; algorithm; corpus	**ملخص البحث باللغة الإنجليزية**

اسم البحث	« نَـمْـذَجَةُ التَّمَثُّلاتِ الصَّرفيَّة العربية وتعليميتها بواسطة الحاسوب في ضوءِ تحليلٍ لِأُمْثِلَةٍ نَصِّيَّة تطبيقية »
اسم الباحث	أ.د. حُسَام الدِّين سَمِير عبد العال محمد
جامعة الباحث	جامعة محمد بن زايد للعلوم الإنسانية
ملخص البحث باللغة العربية	يَنْدَرِجُ هذا البَحْث ضِمْن المُعَالَجَة الآليَّة للغةِ العربية، فَيُقَدِّم طريقةً ومُقْتَرَحاً مِنْ خلالِ تَصَوُّرٍ آليّ يحتوي معلوماتٍ لإنجازِ تحليلٍ صرفيّ للنَّصّ العربيّ. لقد قامت في العصرِ الحَاضِرِ بُحُوثٌ صَرْفِيَّةٌ تَقْتَضِي أُسُسًا دِرَاسِيَّة حَدِيثَة، كما أَنَّهَا تدعو إلى فَحْص الظَّوَاهِر الصَّرفية العربية واختبار مُكَوِّنَات النِّظَام اللُّغَويّ. ومنها دراسة Christian Gaubert (كِرِيسْتيان جوبير)، حيثُ تُمَثِّلُ تلك الدِّرَاسَاتُ مَرْحَلَةً مُهِمَّةً في النَّظَرِيَّة الصَّرفيَّة العَامَّة، وبخاصةٍ مناهجها المبتكرة في وصف التغييرات الصَّرفية في اللغة العربية هذا مِنْ نَاحِيَة. وَمِنْ نَاحِيَةٍ أُخرى يَكْتَسِبُ هذا الموضوعُ أهميةً كبيرةً في الدِّرَاسَات اللُّغَويَّة، حيث جِدَّة المنهج الوَظِيفيّ المُطَبَّق فيها، وبلورة مفاهيم صرفية ونحوية جديدة. كما يهدف البحثُ إلى تقديم دراسة علمية تطبيقية لاستخدام الذَّكاء الاصطناعي في معالجة الصَّرف العربي. ومن أهم نتائج الدراسة الإجابة عن جُملةٍ من التساؤلات حول طرائق الوصول البنيويّ إلى ماهيات النصوص العربية، وكيفيات مقاربتها صرفيًّا، وأشكال التحليل المُسْتَخْدَمَة، وأنماط الأبنية الصرفية، والمنهج النَّسقيّ المُوَظَّف، ورموزه، وعلاماته، وتسلسل وسائطه الحَاسوبية المتعدِّدَة. وما مدى أثر رُؤْيَة كِرِيسْتيان جوبير في إسهام المنهج التَّطْبيقيّ الحَدِيث في تقريبِ ظواهر الصَّرف العربيّ للمُعَالَجَة الآليَّة ؟ وكيف يُنْظَرُ إلى تَحْلِيلِ النَّصّ العربيّ المَشْكول وغير المَشْكول، والتَّفْرِيق بينَ الحَالاتِ التي يعتورها اللَّبْس البنيويّ، وَمَا الطَّرائق المُعْتَمَدَة في تحليلِ بِنْيَةِ المُرَكَّبَاتِ الصَّرفِيّة الإفْرَادِيَّة والسِّيَاقيّة ؟ **<u>الكلمات المفتاحية</u> :** نَـمْـذَجَةُ التَّمَثُّلاتِ، التَّحْلِيلِ الصَّرْفِيّ الآلِيّ، كِرِيسْتِيَان جُوبِير، نُصُوصِ اللُّغَةِ العَرَبِيَّة، نَمَاذِج تَطْبِيقِيَّة.
ملخص البحث باللغة الإنجليزية	This study is intended to introduce the attempt by the French linguistic researcher Christian Gaubert in a research of the phenomena of an automatic morphological analysis of Arabic texts through the modern morphological applied theory in analytic geometry of the structure of the Arab word, the extrapolation of our reading of its terms, the framing of its concepts and the discussion of its recent vision of the Arab potential in its morphological structure, offering that vision for the benefit of the Arab reader, seeking to manage a scientific discussion of what came in this attempt, and directing what appeared in it according to the modern morphological field. **<u>Keywords :</u>** Modeling and teaching Arabic, Automatic morphological analysis, Christian Gaubert, Arabic texts, analytical descriptive study, applied examples.

Classification : Public

اسم البحث	مستقبليات اللغة العربية في فضاء الويب الثالث (WEB 3) والعوالم الافتراضية (الميتافيرس)
اسم الباحث	أسامة عبدالوهاب ريس
جامعة الباحث	مجمع اللغة العربية بالخرطوم
ملخص البحث باللغة العربية	منذ بداية العقد الثالث من القرن الحادي والعشرين تطورت الفضاءات الرقمية واندمجت العديد من التقانات الرائدة من الذكاء الإصطناعي والتكنولوجيا المكانية وتقنيات سلاسل الكتل، والحوسبة السحابية لتشكل هذه الفضاءات الرقمية الجديدة والتي بدورها صنعت ما أصبح يعرف بالويب الثالث، وعوالم افتراضية واعدة أصبحت تسمى بالميتافيرس. حيث أفضت هذه التطورات إلى إضافة فضاءات وأبعاد جديدة للحضارة الإنسانية ونشأ جيل جديد من الشباب (الجيل زد) انغمس في هذه الفضاءات والأبعاد مما أدي لإبتعاده عن الحياة التقليدية. بدأت هذه الثورة الجديدة بالألعاب ثم تطورت في أزمة الكورونا العالمية بحيث خدمت قضية العمل والتعليم عن بعد، والتجارة الإلكترونية وأصبحت الآن تتسع لممارسة جميع النشاطات البشرية. بذلك أفضى هذا التحول إلى نقلة حضارية نوعية مما يتطلب تدخلات غير تقليدية وصناعة تموضع مستقبلي مواكب. كما يتطلب إجراء نقلة مفاهيمية في جل المعارف. يغطي طيف البحث عدة موضوعات تبدأ من الحوار بين الإنسان والآلة لتصل حتي تطبيقات الذكاء الإصطناعي في قضايا اللغة العربية والإدراك الآلي. يسعى هذا البحث إلى تعميق المعرفة حول طرق بناء تفاعل أكثر ذكاءً بين الإنسان والآلة باستخدام اللغة الطبيعية. وقد وسعت نطاقه التطورات الأخيرة في معالجة اللغة الطبيعية. علاوة على ذلك، أصبح من الممكن الآن ليس فقط الاستفادة من التفاعل السلبي المقيد بين الإنسان والآلة بل تم تطوير قنوات ذكية أكثر تفاعلية حيث يكون للذكاء البشري والذكاء الإصطناعي (Artificial Intelligence AI) فرصة للتفاعل. تقدم الورقة نموذجا إطاريا لتفاعل اللغة العربية في هذه الأبعاد وستعمل الورقة على استكشاف مستقبليات اللغة العربية في هذا الإطار مع النظر في الآفاق والتحديات المصاحبة. وتقدم الورقة تعريفا بمجموعة من المعارف المستحدثة في مجال التفاعل المحوسب باللغة العربية والإدراك الآلي والتفاعل بين الإنسان والآلة وتحديات الويب الثالث والميتافيرس. ثم تبحر في الآفاق المستقبلية في حوسبة اللغة العربية في الفضاء الثالث والعوالم الافتراضية ومن ثم تقدم نموذجا لصناعة إطار ومنصة لتطوير برمجيات قادرة على إنتاج وفهم اللغة في صورتيها الصوتية و المرئية (الكلام والكتابة) في الفضاء السيبراني بتطوراته المتسارعة. وتقدم توصيات بمجالات البحوث المستقبلية التي انفتحت فيما بين التخصصات اللغوية والرقمية. **الكلمات المفتاحية** العوالم الإفتراضية (ميتافيرس)، اللغويات الحاسوبية العربية، إدراك اللغة، الويب الثالث، نماذج اللغة، الذكاء الاصطناعي، معالجة اللغة الطبيعية، العالم الدلالي.
ملخص البحث باللغة الإنجليزية	Since the beginning of the third decade of the twenty-first century, cyber spaces have evolved and many leading technologies from artificial intelligence, spatial technology, block chain technologies, and cloud computing have merged to form these new cyber spaces. That in turn created what has been known as Web 3.0, and the promising virtual worlds have been coined to Metaverse. These developments led to the addition of new spaces and dimensions to human civilization, and a new generation of young people arose immersed in these spaces and dimensions. That led to the distance and depart of this Z generation from the traditional life.

Classification : Public

This new revolution began with games, then developed in the global Corona crisis, serving the cause of work, distance education, and electronic commerce, and has now expanded to the practice of all human activities. This transformation led to a qualitative civilized shift, which requires non-traditional interventions and a future positioning process. It also requires a conceptual paradigm shift in knowledge.

The spectrum of the research covers several topics, starting from the dialogue between man and machine, up to the applications of artificial intelligence in issues of the Arabic language and machine perception. This research seeks to deepen knowledge about ways to build smarter human-machine interaction using natural language. Its scope has been extended by recent developments in natural language processing. Moreover, it is now possible not only to benefit from restricted passive human-machine interaction but more interactive smart channels have been developed where human intelligence and artificial intelligence (AI) have a chance to interact.

The paper presents a framework model for the interaction of the Arabic language in these dimensions. The paper will explore the future of the Arabic language in this context, while considering the prospects and the accompanying challenges. The paper introduces a set of concepts in the field of computerized interaction in the Arabic language, automatic perception, human-machine interaction, and the challenges of the third web and metaverse. It then navigates the future horizons of computing the Arabic language in WEB 3.0 and virtual worlds, and then presents a model for creating a framework and a platform for developing solutions capable of producing and understanding the language in its audio and visual forms (speech and writing) in cyberspace with its rapid developments. It provides recommendations for future research areas that have opened up between the linguistic and digital disciplines.

Keywords

Metaverse, Arabic Computational Linguistics Language Perception, WEB 3.0, Language Models, Artificial Intelligence, Natural Language Processing, Semantic Verse.

Classification : Public

اسم البحث	تصور مقترح لبناء معجم رقمي للمتصاحبات اللفظية فكرة تطبيق ذكي
اسم الباحث	د. بدرية براك بنيه العنزي
جامعة الباحث	جامعة الأميرة نورة بنت عبدالرحمن
ملخص البحث باللغة العربية	يهدف هذا البحث إلى تقديم تصور مقترح لتطبيق ذكي لمعجم رقمي لغوي للمتصاحبات اللفظية في اللغة العربية المعاصرة، يستهدف أبناء العربية ومتعلميها لغة ثانية؛ لتحقيق مقاصد منها: تعزيز هوية اللغة العربية، وتنمية الكفاية اللغوية، بالإضافة إلى زيادة الثروة اللفظية، ومساعدة الفئة المستهدفة للارتقاء بمستواها اللغوي وإثراء رصيدها المعرفي، وتمكينها من فهم اللغة العربية وممارستها وفق المواقف اللغوية المختلفة. وكذلك الإسهام في إثراء المحتوى العربي على الإنترنت بمشروع نوعي يخدم اللغة العربية ومتعلميها، ويضاهي المعاجم اللغوية في اللغات الأخرى، إذ لا يوجد حتى اليوم معجم تكاملي بهذه الكيفية، مما يؤكد حاجة الميدان اللغوي لمثل هذه المشروعات النوعية كون تعليم اللغة العربية أصبحت حالة ملحة تفرضها طبيعة التحولات التي نشهدها داخليا وعربيا ودوليا فيما يسمى باقتصاديات المعرفة، والتنمية المعرفية المستدامة، والهوية اللغوية، والأمن اللغوي وغيرها. وتعتمد فكرة التطبيق على برامج اللغويات الحاسوبية وأدواتها التحليلية، ويعتمد بناؤه على مراحل علمية منهجية تتمثل في الجمع والوضع المعجمي والإخراج النهائي (تطبيق ذكي). **كلمات مفتاحية** تطبيق ذكي – معجم رقمي – المتصاحبات اللفظية – المدونات اللغوية – الكلمة الأساس – المدخل المعجمي
ملخص البحث باللغة الإنجليزية	This paper aims to present a proposed vision to design a smart app for a digital lexicon of collocations in contemporary Arabic language. The smart app targets Arabic native speakers and learners of Arabic as a second language to achieve several purposes including: Enhancing the identity of the Arabic language and developing linguistic competence. In addition to increasing verbal strength. Besides helping the target group to improve their linguistic level, enrich their knowledge, and

Classification : Public

enable them to understand and practice the Arabic language according to different linguistic situations. Also contributing to the enrichment of the Arabic content on the internet with a qualitative project that serves the Arabic language and its learners, and is comparable to linguistic lexicons in other languages, as there is no such integrated lexicon to date. This confirms the need of the linguistic field for such qualitative projects because teaching the Arabic language has become a crucial situation imposed by the nature of the transformations that we are witnessing internally, regionally, and globally. By which it is called knowledge economies, sustainable cognitive development, linguistic identity, linguistic security, and others. The idea of the app depends on computer linguistics programs and its analytical tools. The designing of the app depends on systematic scientific stages represented in plurality, lexical status, and final output (smart app). *Keywords:* Smart app, digital lexicon, collocations, linguistic corpuses, core word, lexical entry.	

Classification : Public