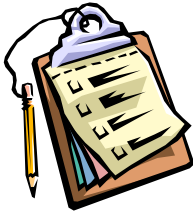


المؤتمر الدولي :
النظم الرياضية لفكر الخليل بن أحمد الفراهيدي وحوسبة اللغة العربية
الأوجه، الآليات، نماذج التفعيل
يومي 06 و 07 ماي 2024م.

تحديات المعالجة الآلية للغة العربية في زمن الرقمنة والذكاء الاصطناعي
محي الدين جودي
جامعة بواتييه - فرنسا

1

خطة العرض



- ❖ الرقمنة والذكاء الاصطناعي ومعالجة اللغات الطبيعية
- ❖ تحديات في مجال المعالجة الآلية للغة العربية
 - ❖ التحليل الصوتي
 - ❖ التعرف الآلي على الكلام العربي
 - ❖ القراءة الآلية للنصوص العربية
 - ❖ الترجمة الآلية للعربية
 - ❖ الدردشة بتقنية الشات بوت
 - ❖ معالجة اللغات العربية الدارجة
 - ❖ توليد اللغة العربية
- ❖ تحليل المشاعر والآراء في النصوص العربية
- ❖ استخلاص المعلومات
- ❖ إنشاء مجموعات بيانات متخصصة باللغة العربية
- ❖ تحدي لغوي وتقني

2

الرقمنة والذكاء الاصطناعي

- ❖ مصطلح "الرقمنة" يشير إلى تحويل العمليات والأنشطة التقليدية إلى شكل رقمي أو استخدام التكنولوجيا الرقمية لتحسينها أو تيسيرها.
- ❖ يشمل هذا التحول استخدام الحوسبة والإنترنت والتكنولوجيا الرقمية الأخرى لتحسين الكفاءة وتحسين الوصول إلى المعلومات والبيانات.
- ❖ **الذكاء الاصطناعي** هو أحد فروع علوم الحاسوب الذي يقوم بدراسة كيفية محاكاة الحاسوب لقدرات الإنسان الحركية والبصرية وتبادل المعرفة فيما بينهما.
- ❖ كما يمكن تعريف الذكاء الاصطناعي بأنه تصميم آلات قادرة على القيام بمهام تتطلب قدرًا من الذكاء البشري
- ❖ من بين المهام : الحفظ والفهم والتواصل والاستدلال والتكيف أو التعلم
- ❖ **تعلم الآلة** هو مجموعة فرعية من الذكاء الاصطناعي الذي تساعد أجهزة الحاسوب على التعلم والتصرف مثل البشر مع تحسين التعلم الذاتي بمرور الوقت.
- ❖ هناك العديد من التطبيقات في مجال الذكاء الاصطناعي منها معالجة اللغات الطبيعية

3

معالجة اللغات الطبيعية

- ❖ معالجة اللغات الطبيعية : فرع من الذكاء الاصطناعي يعنى ببناء برامج لها إمكانية "فهم" لغة طبيعية ما للوصول لهدف محدد.
- ❖ تهتم المعالجة الآلية للغة بدراسة الجوانب الحاسوبية للغة والمشاكل اللسانية التي تواجه هذه المعالجة سواء أكانت هذه اللغة منطوقة أو مكتوبة وبناء نظام معالجة اللغة العربية
- ❖ مهمة معقدة وصعبة، وذلك لصعوبة إدماج المعارف الصوتية والصرفية والنحوية والدلالية في هذا النظام .
- ❖ تتداخل بشكل كبير مع اللسانيات الحاسوبية التي تقدم التوصيف اللغوي المطلوب للحاسوب
- ❖ اللسانيات الحاسوبية أحد فروع اللسانيات التطبيقية تهتم بالاستفادة من معطيات الحاسوب في دراسة قضايا اللسانيات المتعددة مثل: رصد الظواهر اللغوية وفقا لمستوياتها، الصوتية، الصرفية، النحوية البلاغية، والعروضية وإجراء عمليات إحصائية وصناعة المعاجم والترجمة الآلية ،وتعليم اللغات"

4

تحديات في مجال المعالجة الآلية للغة العربية

- ❖ هذه التحديات تمثل أفق البحث الحديث حول المعالجة الآلية للغة العربية في ظل التقدم المستمر في مجالات الذكاء الاصطناعي والتكنولوجيا اللغوية.
- ❖ تسعى هذه البحوث إلى تطوير أدوات وتطبيقات قوية تلبى تحديات التفاعل مع اللغة للعربية في عصر الرقمنة.

❖ المجالات الرئيسية لمعالجة اللغات العربية: (على سبيل الذكر لا الحصر)

- التحليل الصوتي
- التعرف الآلي على الكلام العربي
- القراءة الآلية للنصوص العربية
- الترجمة الآلية للعربية
- الدردشة بتقنية الشات بوت
- معالجة اللغات العربية الدارجة
- توليد اللغة العربية
- تحليل المشاعر والآراء في النصوص العربية
- استخلاص المعلومات
- إنشاء مجموعات بيانات متخصصة باللغة العربية

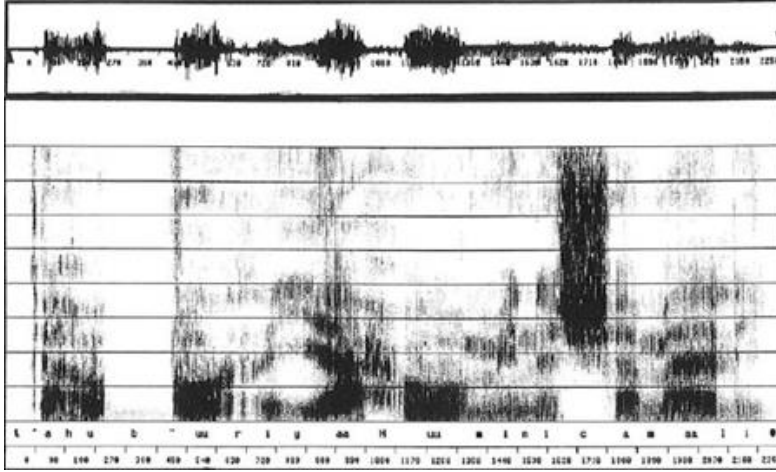
5

التحليل الصوتي

- ❖ التحليل الصوتي هو عملية دراسة الصوت وفهم مكوناته وخصائصه.
- ❖ يشمل هذا التحليل استخدام التقنيات الحاسوبية والإحصائية لفحص الإشارات الصوتية واستخراج المعلومات منها.
- ❖ يُستخدم التحليل الصوتي في مجموعة متنوعة من التطبيقات، منها:
 - التعرف على الكلام: يتضمن فحص النغمات والترددات لتمييز الكلمات والجمل
 - تحسين جودة الصوت: عبر إزالة الضجيج أو تصحيح التشوهات
 - التعرف على الهوية في الكلام أو الأصوات الفردية.
 - تحليل البيانات الصوتية لتدريب نماذج التعلم الآلي
 - تحليل الموسيقى: تحديد النغمات والإيقاع والآلات المستخدمة
 - استخراج المعلومات الصوتية: التعرف على الضجيج أو تحديد الحالة العاطفية
 - التشخيص الطبي: للكشف عن مشاكل صوتية أو تقديم تشخيص لاضطرابات معينة
 - مراقبة البيئة: اكتشاف الأصوات غير المعتادة أو الخطرة.

6

التحليل الصوتي لجملة بالعربية



الإشارة
الصوتية

المخطط الطيفي

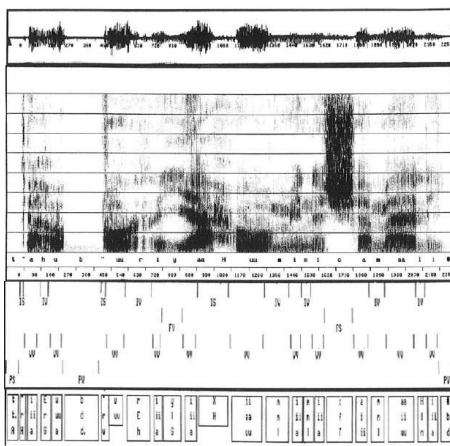
7

التعرف الآلي على الكلام العربي

- ❖ التعرف على الكلام باللغة العربية يشير إلى قدرة الأنظمة الحاسوبية على تحليل وفهم الصوت المنطوق باللغة العربية.
- ❖ يتضمن هذا المفهوم استخدام تقنيات التعلم الآلي ومعالجة اللغة الطبيعية لتحويل الكلام إلى نص أو تحديد الأوامر الصوتية.
- ❖ يمثل تطوير أنظمة فعالة للتعرف على الكلام باللغة العربية تحديًا كبيرًا بسبب اختلاف اللهجات والخصائص الصوتية الخاصة بهذه اللغة.
- ❖ يتطلب التعرف على الكلام جمع الصوت من مصدره، ثم تحليل الأصوات والترددات لفهم الكلمات المنطوقة.
- ❖ يجب أن تكون الأنظمة قادرة على التعامل مع الضوضاء البيئية وتحسين دقة التعرف في ظروف متنوعة.

8

التعرف الصوتي لجملة بالعربية



الإشارة الصوتية

المخطط الطيفي

التقسيم إلى فئات صوتية كبيرة
الحركات والحروف الشديدة والرخوية
وأخرى
المهموس
والمجهور

التعرف الصوتي

تهب الريح من الشمال

9

القراءة الآلية للنصوص العربية

❖ القراءة الآلية للنصوص العربية تشير إلى استخدام التكنولوجيا لفهم ومعالجة النصوص المكتوبة باللغة العربية.

❖ يشمل تطبيق التقنيات المتقدمة في معالجة اللغة الطبيعية والذكاء الاصطناعي لتحليل النصوص واستخراج المعلومات منها.

❖ من بين التطبيقات الشائعة للقراءة الآلية للنصوص العربية:

➤ تحويل النص إلى صوت : Text-to-Speech يتيح هذا التطبيق للأنظمة الحاسوبية تحويل النص المكتوب باللغة العربية إلى كلام مسموع، مما يسمح بتوليد النصوص بشكل صوتي.

➤ التعرف الضوئي على النص : OCR يُستخدم في تحويل النصوص الموجودة في الصور أو المستندات الممسوحة ضوئياً إلى نص قابل للتحريك، مما يتيح معالجة النصوص العربية الموجودة في الوسائط الرقمية.

10

الترجمة الآلية للعربية

- ❖ ترجمة نص من لغة معينة إلى لغة أخرى باستخدام التقنيات المعلوماتية المتقدمة بما في ذلك الحاسبات وبرامج وطرق الذكاء الاصطناعي
- ❖ هدف الترجمة الآلية الحصول على النص المترجم بآلية كاملة دون تدخل بشري بجودة عالية
- ❖ يعد تحسين أنظمة الترجمة الآلية من العربية إلى اللغات الأخرى مجالاً رئيسياً للبحث، لأن ذلك من شأنه تسهيل التواصل بين المتحدثين باللغة العربية ومتحدثي اللغات الأخرى.
- ❖ توجد العديد من نظم الترجمة الآلية للعربية، وقد تطورت هذه النظم على مر السنوات بفضل التقدم في مجال تكنولوجيا المعالجة اللغوية الطبيعية والذكاء الاصطناعي.

11

الدرشة بتقنية الشات بوت

- ❖ **الدرشة** : يشير إلى سلسلة مناقشة يتبادل فيها مستخدمو الإنترنت الرسائل على الفور
- ❖ **الشات بوت** : برامج مدمجة في أنظمة الدردشة لأغراض مختلفة، مثل: خدمة العملاء ، والتسويق، والشراء والبيع ، ومؤخراً الحصول على المعلومات.
- ❖ انتشرت التطبيقات القائمة على تقنية Chatbot في العديد من الشركات والمواقع الإلكترونية التي تعتمد على الاتصال مثل Facebook و: Twitter وTelegram وغيرها ؛
- ❖ من بينها تطبيق ChatGPT الذي طورته شركة OpenAI الأمريكية.

Mahieddine Djoudi (2023). Fonctionnement et utilisation de ChatGPT, Webinaires en Sciences & Technologies Avancées (WAST'2023), 13 mai 2023. ASCA, <https://www.youtube.com/watch?v=0WPvWneqTMk>

12

معالجة اللغات العربية الدارجة



❖ معالجة اللغات العربية الدارجة تشير إلى تطوير وتطبيق التقنيات والأدوات التي تهدف إلى فهم ومعالجة اللهجات واللهجات العربية غير القياسية والمستخدم في الحياة اليومية وفي العمليات التواصلية اليومية.

❖ هذه اللغات الدارجة قد تختلف عن اللغة العربية القياسية في البنية والمفردات والنطق، ويتطلب التعامل معها تقنيات متخصصة.

❖ معالجة اللغات العربية الدارجة تتطلب فهماً عميقاً للتباينات اللغوية والثقافية في العالم العربي، بالإضافة إلى تطبيق تقنيات متقدمة في معالجة اللغة الطبيعية لتحقيق فهم دقيق وفعال للنصوص والأصوات.

❖ **نجوم** أول محرك للتفكير باللغة العربية باستخدام الذكاء الاصطناعي - بالفصحى والدارجة

د. مراد بوعاش ود. فؤاد بوسطوان <https://search.nojoom.ai/>

13

توليد اللغة العربية

❖ عملية توليد اللغة العربية تتضمن إنشاء نصوص أو جمل باللغة العربية بشكل تلقائي باستخدام الحوسبة. يتم ذلك عبر استخدام تقنيات معالجة اللغة الطبيعية والذكاء الاصطناعي.

❖ هذه التقنيات يمكن أن تكون مفيدة في عدة سيناريوهات، بما في ذلك:

➤ تلخيص المحتوى، كتابة المقالات، إنشاء التعليمات، وإنشاء المحتوى التسويقي.

➤ ترجمة النصوص من لغة مصدرية إلى اللغة العربية

➤ توليد الشرح والتوضيح للمفاهيم التقنية أو العلمية، وهو مفيد في التعليم عبر الإنترنت

➤ توليد نصوص أدبية أو شعرية باللغة العربية، وهو مجال مثير للاهتمام في الذكاء الاصطناعي.

➤ توليد الأسئلة والإجابات في مجالات مثل التعليم والتدريب.

➤ توليد نصوص تفاعلية تستجيب لاستفسارات المستخدمين أو لتوجيههم في التطبيقات والمواقع على الإنترنت.

14

تحليل المشاعر والآراء في النصوص العربية

- ❖ يعد تحليل الآراء والمشاعر في النصوص العربية أمرًا بالغ الأهمية لفهم الاتجاهات والآراء العامة على وسائل التواصل الاجتماعي والمنتديات عبر الإنترنت والمقالات الإخبارية.
- ❖ التعرف على القطبية في تعليقات الصحافة الإلكترونية الجزائرية
- ❖ القدرة على تحديد مدى إيجابية أو سلبية الآراء في تعليقات الصحافة المكتوبة، وذلك في سياق دراسة التوجه العام للقراء ومدى تفاعلهم مع الأخبار والأحداث.
- ❖ استعمال خوارزميات تعلم الآلة التي تحتاج إلى كم من المعلومات على شكل مكنز يتم تدريب الخوارزميات عليه
- ❖ قمنا بجمع يدويا كم من التعليقات من موقع أحد الصحف الجزائرية لتشكيل المكنز الخاص بعملنا.
- ❖ ثم قمنا باستخدام طريقتين مشهورتين في هذا المجال هما:
 1. شعاع الدعم الآلي Support Vector Machines
 2. المصنف البايزي الساذج Naive Bayesian Classifier
- ❖ مع عدة خيارات في الإعدادات لكل منهما.
- ❖ كما تم حساب معيار الدقة (Precision) والاسترداد (Recall) والصحة (Accuracy) لكل حالة.
- ❖ مقارنة بأعمال مشابهة فقد حصلنا على نتائج تنافسية -- دكتوراه لهشام رحاب 2020

15

استخلاص المعلومات

- ❖ استخلاص المعلومات باللغة العربية يشير إلى عملية استخراج المعلومات الهامة والمفيدة من النصوص العربية بشكل تلقائي أو يدوي.
- ❖ يعتبر هذا العمل جزءًا أساسيًا من مجال معالجة اللغة الطبيعية ويشمل عدة تحديات وتقنيات.
- ❖ بعض النقاط المهمة حول استخلاص المعلومات باللغة العربية:
 - يتطلب استخلاص المعلومات تحليلًا دقيقًا للنصوص العربية، بما في ذلك الفهم السياقي والبنية اللغوية.
 - يمكن أن يتم استخلاص المعلومات لأغراض متنوعة، مثل تلخيص النصوص، واستخراج الأحداث الرئيسية، وتحليل المشاعر، واستخراج المعلومات الكمية.
 - يتضمن استخلاص المعلومات في اللغة العربية استخدام تقنيات التعلم الآلي ومعالجة اللغة الطبيعية لتحقيق أداء دقيق وفعال.
 - يتعامل الباحثون مع تحديات لغوية مثل التعامل مع اللهجات المتنوعة، وتحليل اللغة العربية الدارجة، وفهم الاختلافات الثقافية في اللغة.
 - يعتمد استخلاص المعلومات باللغة العربية على فهم السياق الثقافي واللغوي لتحقيق أفضل فهم وتفسير للنصوص.
 - يعتبر استخلاص المعلومات أحد أسس الاحترافية في التعامل مع اللغة العربية على وسائل التواصل الاجتماعي والإنترنت.
 - يعزز استخلاص المعلومات باللغة العربية من فعالية تحليل وفهم النصوص العربية، مما يساهم في تحسين جودة المحتوى وتعزيز تفاعل المستخدمين العرب.

16

إنشاء مجموعات بيانات متخصصة

- ❖ إعادة إحياء وتحديث **مشروع الذخيرة العربية** للدكتور عبد الرحمان الحاج صالح
- ❖ إنشاء مجموعات بيانات متخصصة باللغة العربية يتطلب الاهتمام بالتفاصيل وتحديد الغرض الرئيسي لاستخدام هذه المجموعات.
- ❖ قد تكون هذه المجموعات مفيدة لتدريب نماذج الذكاء الاصطناعي في مجالات مثل التعلم الآلي ومعالجة اللغة الطبيعية.
- ❖ يجب أن تكون هذه المجموعات متوفرة بشكل عادل ويتم التعامل مع البيانات بشكل آمن وخصوصي.
- ❖ كما يُفضل توثيق المجموعة بيانات بشكل جيد لضمان فهم صحيح للغرض والاستخدام المستقبلي.
 - أزواج من الجمل باللغة الإنجليزية وترجماتها إلى العربية
 - جمل بالعربية مصنفة بناءً على المشاعر المعبر عنها (سعادة، حزن، إلخ)
 - أسئلة باللغة العربية وإجابات عليها
 - مجموعة بيانات تحتوي على نماذج صوتية ونصوص توافقها
 - جمل مقروءة لاستعملها في التعرف الآلي

17

تحدي لغوي وتقني : النحت

- ❖ **النحت** : بناء كلمة جديدة من كلمتين أو أكثر أو من جملة بحيث تدل على المعنى المراد
- ❖ **الخليل بن أحمد الفراهيدي** هو أول من اكتشف ظاهرة النحت في اللغة العربية
- ❖ حين قال: إن العين لا تأتلف مع الحاء في كلمة واحدة لقُرْبٍ مخرجيهما، إلا أن يشتقَّ فعل من جمع بين كلمتين
- ❖ مثل (حي على) كقول الشاعر: أقول لها ودمع العين جار ألم يحزنك **حيلة** المنادي

الإسم	المعنى المراد
بسملة	بسم الله الرحمن الرحيم
سبحة	سبحان الله
حولقة	لا حَوْلَ ولا قُوَّةَ إلا بالله
حَيْعَلَة	حيَّ على الصلاة، حيَّ على الفلاح
سمعة	سَمِعَ اللهُ لَمَنْ حَمَدَهُ
حسيلة	حَسْبِيَ اللهُ

18

أهمية النحت في عصر الرقمنة

❖ وسيلة من وسائل تنمية اللغة وتكثير مفرداتها؛

❖ كثير الاستعمال في اللغات الأخرى

Tanzanie, Informatique, Progiciel, Brexit, Courriel, Webinaire, Chatbot, ...

❖ امتدادا للمشروع الذخيرة العربية

❖ إنشاء مجموعات بيانات للكلمات المنحوتة في الميادين العلمية المختلفة

الكلمة	منحوتة من	الكلمة	منحوتة من
برماجي	بر + بحر + جو	بَرِيذَال	بريد + إلكتروني
ركمجة	ركوب + الموجة	حِينَبَات	حيوان + نبات
لأخلاقِي	لا + أخلاق	فُونَفْسَجِي	فوق + بنفسجي
الْكُرِيضَة	الكرية + البيضاء	قَبْتَارْخِي	قبل + تاريخ
زَمَان	زمن + مكان	شَنْكُوتِيَة	شبكة + عنكبوتية
كَهْرْمَائِي	كهربائي + مائي		

19

بعض المراجع

❖ الحاج صالح، عبد الرحمن. 1999. مشروع الذخيرة اللغوية العربية وأبعاده العلمية والتطبيقية. اللغة العربية، مج. 1999، ع. 2، ص. 32-49. <https://search.emarefa.net/detail/BIM-950793>

❖ المعالجة الآلية للغة العربية <https://www.asjp.cerist.dz/en/article/161257>

❖ الذكاء الاصطناعي واللغة العربية <https://ask2learn.net/ai-and-arabic-language/>

❖ تحديات المعالجة الآلية للغة العربية- <https://search.emarefa.net/ar/detail/BIM-1253161-%D8%AA%D8%AD%D8%AF%D9%84%D8%A7%D8%AA-%D9%84%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D9%85%D8%B9%D8%A7%D9%84%D8%AC%D8%A9-%D8%A7%D9%84%D8%A2%D9%84%D9%8A%D8%A9-%D9%84%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9>

❖ الذكاء الاصطناعي في الصوت العربي (2) : التحديات- <https://ae.linkedin.com/pulse/%D8%A7%D9%84%D8%B0%D9%83%D8%A7%D8%A1-%D8%A7%D9%84%D8%B5%D8%B7%D9%86%D8%A7%D8%B9%D9%8A-%D9%81%D9%89-%D9%84%D8%B5%D9%88%D9%8A%D8%A9-%D9%88%D8%A7%D9%84%D9%87%D9%88%D9%8A%D8%A9-%D9%88%D8%A7%D9%84%D8%AA%D8%A7%D9%84%D8%AA%D8%AF%D9%8A%D8%A7%D8%AA-ala-a-suleiman-7dotc>

❖ تحديات وجود الذكاء الاصطناعي باللغة العربية- <https://maaal.com/2022/09/%D8%AA%D8%AD%D8%AF%D9%84%D8%A7%D8%AA-%D9%88%D8%AC%D9%88%D8%AF-%D8%A8%D8%A7%D9%84%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B0%D9%83%D8%A7%D8%A1-%D8%A7%D9%84%D8%A7%D8%B5%D8%B7%D9%86%D8%A7%D8%B9%D9%8A-%D8%A8%D8%A7%D9%84%D8%BA%D8%A9-%D8%A7>

❖ مقارنة بين الذكاء الطبيعي والذكاء الاصطناعي ... المعالجة الآلية للغة العربية، جهود وتحديات <https://revue.ummo.dz/index.php/pla/article/download/2722/1803>

❖ اللغة العربية بين الانتماء والهوية والتحديات المستقبلية في ... <https://khubabaa.com/wp-content/uploads/myfiles/%D8%A7%D9%84%D9%84%D8%BA%D8%A9-%D8%A8%D9%8A%D9%86-%D8%A7%D9%84%D8%A7%D9%86%D8%AA-%D9%85%D8%A7%D8%A1-%D9%88%D8%A7%D9%84%D9%87%D9%88%D9%8A%D8%A9-%D9%88%D8%A7%D9%84%D8%AA%D8%A7%D9%84%D8%AA%D8%AF%D9%8A%D8%A7%D8%AA-%D8%A8%D8%A7%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B0%D9%83%D8%A7%D8%A1-%D8%A7%D9%84%D8%A7%D8%B5%D8%B7%D9%86%D8%A7%D8%B9%D9%8A-%D8%A8%D8%A7%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9>

❖ هل حان الوقت لتتكلم الآلات باللغة العربية؟- <https://www.wamda.com/ar/memakersge/2016/05/%D9%87%D9%84-%D8%AD%D8%A7%D9%86-%D8%A7%D9%84%D9%88%D9%82%D8%AA-%D8%A8%D8%A7%D9%84%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B0%D9%83%D8%A7%D8%A1-%D8%A7%D9%84%D8%A7%D8%B5%D8%B7%D9%86%D8%A7%D8%B9%D9%8A-%D8%A8%D8%A7%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9>

❖ اللغة العربية ومواكبة تحديات العصر الرقمنة <https://www.qf.org.qa/ar/stories/keeping-the-arabic-language-relevant-in-the-digital-era>

❖ اللغة العربية في ظل الذكاء الاصطناعي- <https://www.aljazeera.net/culture/2023/9/11/%D8%A7%D9%84%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9-%D8%A7%D9%84%D8%B0%D9%83%D8%A7%D8%A1-%D8%A7%D9%84%D8%A7%D8%B5%D8%B7%D9%86%D8%A7%D8%B9%D9%8A-%D8%A8%D8%A7%D9%84%D8%BA%D8%A9-%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9>

❖ العين: الخليل بن أحمد الفراهيدي، تحقيق مهدي المخزومي وإبراهيم السامرائي، بغداد 1980 - 1985 م. النحت في العربية قديما وحديثا - المجلد 1 - الصفحة 116 - جامع الكتب الإسلامية

20

