

رؤية

دَا التَّاصِيكِ

مُرْكز الجُوحِ وَنقْدِنَا مَعْلُومَاتِ

لتوثيق اللغة العربية وتطوير التعامل معها آلياً

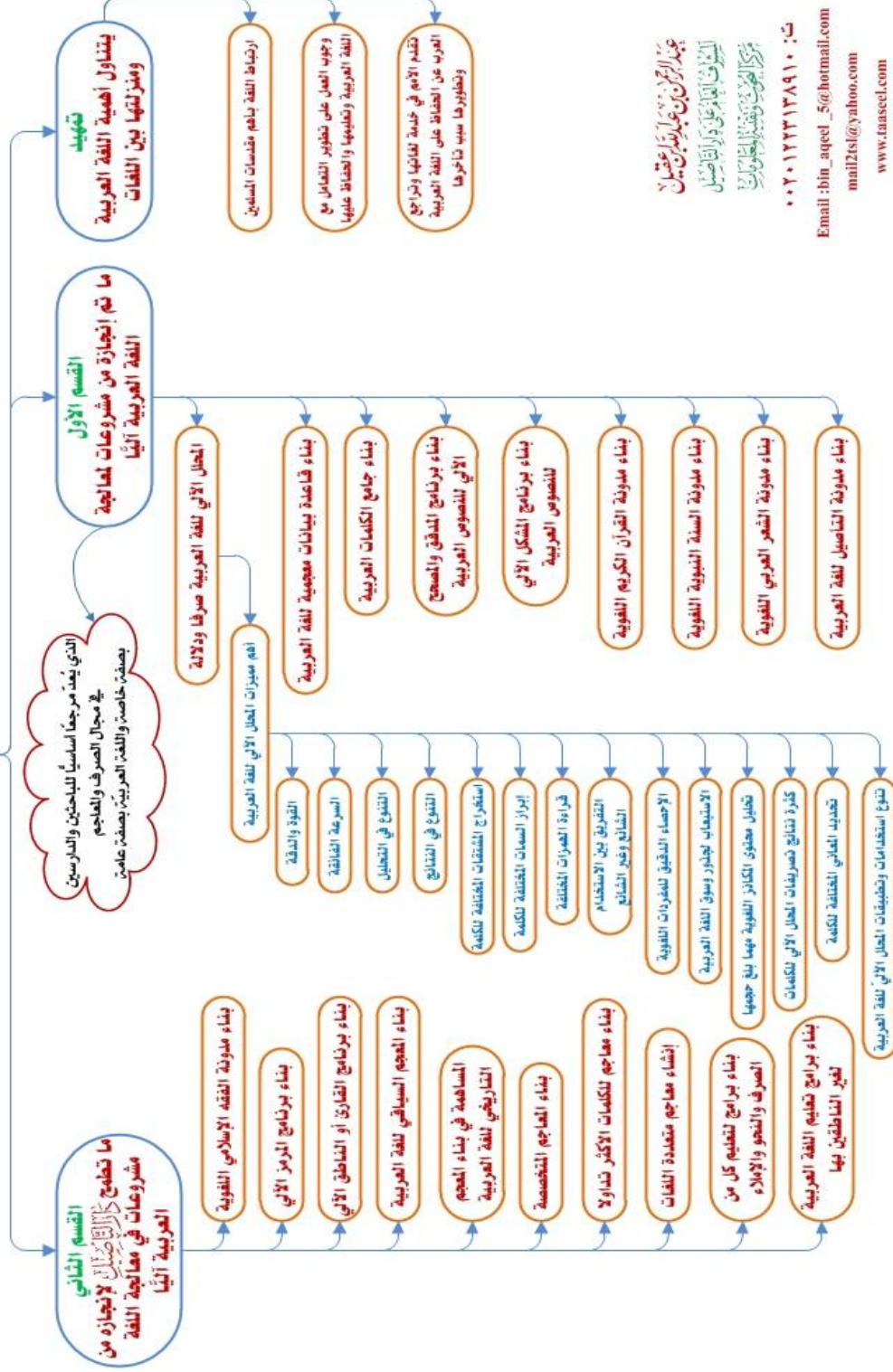
١٤٤١هـ - ٢٠١٩م

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

٢٠١٩ / ١٠ / ٢٩ م

ملخص رؤية مركز البحوث وتقنية المعلومات لتطوير التعامل مع اللغة العربية آتياً

قسم حوسبة اللغة العربية
مركز البحوث وتقنية المعلومات



محمد الزهراني محمد بن عبد الله

مركز البحوث وتقنية المعلومات

www.taaseed.com

ت: ٠٢٠١٢٢٢١٣٨٩١٠

Email: bin_aqeel_5@hotmail.com

mail2ts1@yahoo.com

www.taaseed.com

رؤية إِسْتِزْهَابُ

مركز البحوث وتقنية المعلومات

لتوثيق اللغة العربية وتطوير التعامل معها آلياً

ارتبطت اللغة العربية بأهم مقدسات المسلمين؛ حيث اختارها الله ﷻ لغةً للقرآن الكريم وهو الوحي الذي أوحى به الله لنبيه محمد ﷺ هداية للعالمين، وهي لغة السنة النبوية التي هي جزء من الوحي، بالإضافة إلى ما كُتب من علوم التراث باللغة العربية، فحفظ الله القرآن والسنة النبوية باللغة العربية وحفظ اللغة العربية من التشرذم والضياع الذي هو داء اللغات على مدى التاريخ بالقرآن والسنة النبوية، ومن أجل هذا اكتسبت اللغة العربية خاصية تميزها عن بقية اللغات الحية، وهي ثباتها، وامتناعها عن التغير والضياع كما حصل للغات أخرى وصدق الله العظيم بقوله ﴿إِنَّا نَحْنُ نَزَّلْنَا الذِّكْرَ وَإِنَّا لَهُ لَحَافِظُونَ﴾ [الحجر: ٩] والذكر نزل بلسان عربي مبين.

ومن ميزات لغتنا العربية العظيمة قابليتها للنمو والتطور واستيعابها لكل جديد؛ ولهذا

كان من الواجب على العرب والمسلمين العمل على تطوير التعامل معها تقنيا ونشرها وتعليمها والحفاظة

عليها، والدفاع عنها، يقول شيخ الإسلام ابن تيمية رَحِمَهُ اللهُ فِي كتابه: «اقتضاء الصراط المستقيم»

(ص ٢٠٧):

«إن نفس اللغة من الدين، ومعرفتها فرض واجب، فإن فهم الكتاب والسنة فرض، ولا يفهمان إلا بفهم اللغة العربية، وما لا يتم الواجب إلا به فهو واجب». ثم قال: «ثم منها ما هو واجب على الأعيان، ومنها ما هو واجب على الكفاية».

والمطلع على الجهود التي بذلتها الأمم المتقدمة تقنياً في العصر الحديث لتطوير التعامل مع لغاتها وصبغها بالعالمية يجدها قد أصبحت تتعامل مع لغاتها آلياً من خلال مُحلِّلات آلية خاصة بهذه اللغات، أوصلت هذه اللغات للعالمية ويسرت التعامل معها، رغم أنها دون اللغة العربية بكل المقاييس.

والوصول إلى التعامل الآلي مع اللغات العالمية المعاصرة هو الذي بوأ لغات لا تعدل اللغة العربية تاريخًا وثناءً وتطورًا واشتقاقًا مكانة كبيرة وشهرة في هذا العصر .

ومن المعلوم أن التخلف الحضاري للعرب هو الذي أوصلهم إلى العجز في كثير من المجالات ، ومن ذلك أنهم لم يطوروا التعامل مع لغتهم المقدسة حاسوبيا بشكل آلي مما يجلب الأسى للغيورين على اللغة العربية ، وإلى الله المشتكى من تخاذل العرب عن تطوير التعامل الحاسوبي مع لغتهم ، ولنا أن نردد قول المتنبي :

ولم أر في عيوب الناس عيباً كنقص القادرين على التمام

فالعرب - رغم ثراء بعض دولهم الذي يمكنهم من الصرف في وجوه شتى - لم يوفقوا إلى المواءمة بين علوم اللغة العربية وعلوم الحاسب لإيجاد **محلل آلي** للغتهم المقدسة .

ومن هنا جاء اهتمام **إِذَا تَأَصَّلْتَ مُرَكَّبًا بِحُجْرَةٍ وَتَقْنِيًا بِالْمَعْلُومَاتِ** بتطوير التعامل مع اللغة العربية وتوثيقها آليًا ، حيث إن اللغة العربية هي ثلثة الأثافي في منظومة الثقافة في الحضارة الإسلامية مع القرآن الكريم وعلومه والسنة النبوية وعلومها ، وقد وُفِّقْنَا بفضل الله وعونه بعد جهود حثيثة لسد هذه الثغرة وتحقيق قدر كبير مما ينبغي على الأمة القيام به لتطوير التعامل الحاسوبي مع لغتنا العربية ، فله الحمد والشكر .

ولا يخفى على أحد من المعاصرين المهتمين باللغة العربية مدئى **أهمية معالجة اللغة العربية آليًا بواسطة الحاسب الآلي وتطبيقاته** ؛ حيث تمثل المعالجة الآلية للغة أفضل ما تُقدِّمه تقنية المعلومات للباحثين في اللغة العربية .

وقد خطت **إِذَا تَأَصَّلْتَ مُرَكَّبًا بِحُجْرَةٍ وَتَقْنِيًا بِالْمَعْلُومَاتِ** خطوات حثيثة في معالجة اللغة العربية آليًا فأنشأت العديد من المشروعات اللغوية الضخمة المتميزة - ولله الحمد - والعمل جار على تحسينها وتطويرها بشكل دائم ، ومن المشروعات ما أنجزته وانتهت منه ، ومنها ما حددته وتعمل على إنجازه ، وفيما يلي تفصيل ذلك :

القسم الأول: ما تم إنجازه من مشروعات لمعالجة اللغة العربية آلياً:

أولاً: المحلل الآليّ للغة العربية صرفاً ودلالة:

وهو إنجاز لغويّ حاسوبيّ ضخم يُعدّ مرجعاً أساسياً للباحثين والدارسين في مجال الصرف والمعاجم بصفة خاصة واللغة العربيّة بصفة عامة ، ويعد التحليل الصرفي الآلي للغة أوضح النماذج للمعالجة الآلية للغة ، والقاعدة التي لا غنى عنها لبناء البرامج اللغوية الآلية التي يتجلى فيها الذكاء الاصطناعي **Artificial Intelligence** كالصحیح والتشكيل الآليين للنصوص اللذين تم إنجازهما وتطبيقهما على مستوى أعمال مركز البحوث في **إِنَّا لِلَّهِ وَإِنَّا إِلَيْهِ رَاغِبُونَ** ، كما تتجلى تقنيات الذكاء الاصطناعي في النطق الآلي والترجمة الآلية ، والذكاء الاصطناعي هو جعل الآلة تحاكي البشر فتتعلم مما يحيط بها من معلومات حتى تنفذ المهام بكفاءة عالية تشبه كفاءة البشر بهدف مساعدة البشر الذين قد يصيبهم التعب والإرهاق أو الملل الذي لا يصيب الآلة ، وقد أصبحت تطبيقات الذكاء الاصطناعي موجودة في حياتنا اليومية بكثرة سواء من خلال الحاسوب أو الهاتف المحمول أو الصناعة أو الطيران الآلي أو سيارات القيادة الذاتية وغيرها .

ومن توفيق الله لنا أن أول تفعيل وتطبيق لهذا المحلل الآلي كان في بناء مدونة القرآن الكريم اللغوية ومدونة السنة النبوية اللغوية اللتين أنجزتا ولله الحمد في

مركز البحوث والتقنية المعلومات **إِنَّا لِلَّهِ وَإِنَّا إِلَيْهِ رَاغِبُونَ** .

ومن أهم ما يتميز به المحلل الآليّ للغة العربية ما يلي:

- **القوة والدقة:** حيث هو أقوى وأدقّ محلل صرفيّ آليّ على الساحة المعلوماتيّة على ندرتها ومحدوديتها وعدم اكتمالها وهو أفضل بمراحل من محلل مايكروسوفت الصرفيّ التي تُعدّ أمّ صناعة البرمجيات في العالم .

- **السرعة الفائقة:** حيث يستطيع تحليل (١٢٠٠) ألف ومائتي كلمة في الثانية الواحدة .
- **التنوع في التحليل:** حيث تتم عملية التحليل في المحلل الآلي على أكثر من مستوى: فيحلل الكلمة غير مشكولة، والكلمة مشكولة شكلاً تاماً، والكلمة مشكولة شكلاً جزئياً .
- **التنوع في النتائج:** حيث يعطي **المحلل الآلي** تشكيلات وقراءات مختلفة للكلمة العربية الصحيحة مصحوبة بجذر الكلمة، وساقها، والسوابق واللواحق الواردة في الكلمة، وقسم الكلم الخاص بها، وكذا الصيغة الصرفية للكلمة .
- **استخراج المشتقات المختلفة للكلمة:** حيث يبرز **المحلل الآلي** المشتقات المختلفة للكلمة فلو أخذنا الفعل (أَسْمَع) لوجدنا أنه يعطينا المشتقات التالية: يُسْمِع، أَسْمِع، أَسْمِع، يُسْمِع مع بياناتها المختلفة من حيث قسم الكلم والجذر والساق والصيغة الصرفية، والسمات المختلفة .
- **إبراز السمات المختلفة للكلمة:** حيث يبرز **المحلل الآلي** بعض السمات الصرفية والمعجمية والنحوية ك: (التذكير والتأنيث)، (البناء والإعراب)، (الإفراد والتثنية والجمع)، (التعريف والتنكير) . . . إلخ، فلو حللنا مثلاً كلمة «كتابا» سنجد تحليل «كِتَابًا» يحمل سمات: «مفرد، مذكر، منصوب، غير مضاف، منون، نكرة»، وتحليل «كِتَابًا» يحمل سمات: «مثنى، مذكر، مرفوع، مضاف، معرف بالإضافة» وهكذا بقية التحليلات الناتجة عن كلمة «كتابا»، هذا فضلاً عن بيانات قسم الكلم والجذر والصيغة الصرفية الخاصة بكل تحليل .
- **قراءة الهمزات المختلفة:** حيث يقوم **المحلل الآلي** بقراءة الهمزات المختلفة، آخذاً بالاعتبار مراعاة اختلاف مدارس رسم الهمزة: (يقرءون، يقرؤون)، (يلجئون، يلجؤون)، (مسؤل، مسؤل).

- **التفريق بين الاستخدام المعاصر وغير المعاصر:** حيث يقوم **المحلل الآلي** بالتفريق بين الاستخدام المعاصر الشائع للكلمات ، والاستخدام غير المعاصر للكلمات ، وكذا الكلمات ذات الرسم الخاص بالقرآن الكريم .
- **الإحصاء الدقيق لمفردات اللغة العربية:** حيث يعطي **المحلل الآلي** إحصائية لمفردات اللغة العربية تجاوزت حتى تاريخه **مائتين وأربعين مليون كلمة** مشكولة غير مكررة .
- **الاستيعاب لجذور وسوق اللغة العربية:** حيث بلغ عدد الجذور التي اعتمد عليها **المحلل الآلي** حتى تاريخه : **اثنى عشر ألفاً وستمائة جذر** ، وقد تجاوزت الجذور المدخلة في **المحلل الآلي** جذور «معجم تاج العروس» التي بلغت : عشرة آلاف وثمانمائة وواحد وتسعين جذراً ، كما بلغ عدد السوق في **المحلل الآلي** حتى تاريخه : **مائتين وسبعين ألف ساق** ^(١) .
- **تحليل محتوى المكانز اللغوية مهما بلغ حجمها:** حيث قام **المحلل الآلي** بتحليل (مدونة التأصيل للغة العربية) التي تحوي : **مليارين ومائتي مليون كلمة** في وقت وجيز بدون أي إشكالات .
- **كثرة نتائج تصريفات المحلل الآلي للكلمات:** حيث يقوم باستيعاب تصريفات الكلمة الواحدة التي قد تصل إلى أن بعض الشقوق يتولد منها **ما يزيد على أحد عشر ألف كلمة** ، وتصل تصريفات ومشتقات بعض هذه السوق بسوابقها ولواحقها إلى ما يقرب من **ثلاثين ألف كلمة** ، انظر مثلاً ساق «يُسْمِع» على موقع **دار التأسيس** .

(١) ساق الكلمة : هي المفردة اللغوية مجردة من السوابق واللواحق وتحمل دلالة صرفية واحدة بحيث إذا اقتطع منها جزء صارت إما مفردة خطأ لا تدل على شيء ، أو أخذت دلالة صرفية أخرى ، مثال : (كتابة) مصدر مجرد على وزن فعالة ، لو اقتطعت منها التاء المربوطة صارت (كتاب) فأخذت دلالة صرفية جديدة ، هي دلالة الاسم بدل المصدر المجرد في (كتابة) ، وبوزن صر في جديد هو (فعال) .

● **تحديد المعاني المختلفة للكلمة:** حيث يقوم **المحلل الآلي** بتحديد المعاني المعجمية المختلفة للكلمة، والحقل الدلالي لها، وما يرتبط بها من علاقات دلالية، ويجري العمل على استكمال هذه الميزة من ميزات المحلل وكادت أن تتم على مستوى سوق المحلل.

● **تحديد العلاقة بين كل سوق المحلل من حيث الترادف والتضاد:** حيث يقوم **المحلل الآلي** بإعطاء المترادفات والمتضادات المختلفة لكل كلمة من كلمات اللغة، ويجري العمل أيضا على استكمال هذه الميزة من ميزات المحلل.

● **تنوع استخدامات وتطبيقات المحلل الآلي للغة العربية:** حيث يُعدّ **المحلل الآلي** أداة غاية في الدقة يُعتمد عليها في تطبيقات عديدة منها ما يلي:

١- **بناء العديد من المعاجم اللغوية المختلفة.**

٢- **مراجعة المعاجم العربية الموجودة على الساحة ووضع استدراقات وتصويبات عليها.**

٣- **استرجاع المعلومات بسرعة ودقة.**

٤- **تيسير التعليم عن طريق الحاسوب.**

٥- **بناء البرامج الحاسوبية المتخصصة في اللغة العربية**، على سبيل المثال لا الحصر: (المدقق والمصحح الإملائي الآلي للنصوص العربية، والمشكل الآلي لنصوص اللغة العربية).

ولما كان المعجم رافداً مُهمّاً من روافد اللغة والمعرفة ويُمثل ذاكرة الشعوب، فقد اعتنى به اللغويون العرب من حيث المحتوى والهيكل من قديم الزمان، فازدهرت الصناعة المعجمية العربية وتعدّدت مدارسها.

ومع انتشار استعمال الحاسوب والشبكة العنكبوتية (الإنترنت) وظهور ما يسمى «مجتمع المعرفة وعصر المعلومات» برزت على الساحة تطبيقات لغوية حاسوبية جديدة ومتنوعة.

هذه التطبيقات تعتمد في غالبها على المعجم ، وقد شهد المعجم الحاسوبي تطورًا كبيرًا على صعيد الهيكلية ومحتوى المداخل ، وكذلك الخدمات المتطورة التي يُمكن أن يُقدّمها للمستخدم مستفيدا في ذلك من الإمكانيات الهائلة التي يوفرها الحاسوب من طاقة تخزين للمعلومات المعجمية وسرعة البحث عنها واسترجاعها ، والقدرة على معالجة البيانات متعددة الوسائط ، مما سيسهم في تطوير الصناعة المعجمية العربية الحديثة بسرعة فائقة ، وأصبحت لا تترى في هذا العصر في الدول المتقدمة تطبيقات لغوية حاسوبية دقيقة إلا إذا كانت مبنية على **محلل آلي** قوي ذي دقة عالية .

ومن هنا كانت قناعة **كازالتأصيل مركز البحث وتقنية المعلومات** في إنشاء **المحلل الآلي للغة العربية** ، حيث إنه القاعدة الأساسية التي ستبنى عليها أعمال توثيقية جلييلة توثق اللغة العربية بالإضافة إلى برامج ومحركات بحث لخدمة اللغة العربية تضعها تقنيا ومعاصرة في مستوى مقدمة اللغات العالمية الحية .

ثانيا - بناء قاعدة بيانات معجمية للغة العربية :

وهي مجموعة من البيانات المنظمة توفر سهولة الاستعلام والاستعادة ، وتعدّ هذه القاعدة إحدى الركائز التي يبنى عليها برنامج المحلل الآلي ، وهي بمثابة الركيزة الأساسية التي يُعتمد عليها في بناء أي برنامج لمعالجة اللغة العربية آليا .

ثالثا - بناء جامع الكلمات العربية^(١) :

وهو هيكل معرفي ضخم عبارة عن شبكة من العلاقات اللغوية **Lexical Semantic Relations** الصرفية والدلالية والنحوية والمعجمية .

(١) وهو ما يطلق عليه في الغرب شبكة الكلمات العربية **Arabic Word Net** .

والأصل في **جامع الكلمات العربية** أن يتكون من قواعد بيانات معجمية تتلاحم عناصرها باستخدام علاقات لغوية مختلفة .

وقد سعى فريق بحثي في جامعات برينستون وبرشلونة ومانشستر إلى بناء شبكة للكلمات العربية على غرار شبكة كلمات برينستون لكلمات اللغة الإنجليزية إلا أن العمل توقف نتيجة نقص الموارد المتاحة للفريق ، ولم يبلغ عدد كلمات شبكة الكلمات العربية التي عملوا عليها سوى (١٧٥٦١) كلمة ، ولم تكن كلمات تلك الشبكة عربية بالمفهوم الدقيق بل كانت ترجمة لبعض أجزاء كلمات شبكة كلمات برينستون ، ولم يتصد للقيام بهذا العمل أيٌّ من مجامع اللغة العربية أو كليات ومعاهد تدريس اللغة العربية^(١) .

قد أخذت **إِذَا تَوَضَّعْتَ** على عاتقها الدخول في هذا المشروع اللغوي الضخم فقامت بعون الله ثم بما لديها من فريق مؤهل من اللغويين والمطورين والمبرمجين ، وبما تملكه من قواعد بيانات مختلفة للغة العربية ببناء هذا الجامع لكلمات اللغة العربية التي تجاوزت كلماته : **مائتين وستين مليون كلمة مشكولة وغير مكررة** .

وتكمن أهمية جامع كلمات اللغة العربية في أنه يساعد على :

● **تنمية المحتوى الإلكتروني للغة العربية** على الشبكة العنكبوتية سواء كان حُرًّا أو مُخَرَّجًا^(٢) .

● **تطوير محركات البحث العربية** .

(١) انظر كتاب «علم اللغة والتقنيات المعاصرة» تأليف د . المعتز بالله السعيد .

(٢) المقصود بالحر هو ما يمكن الاستفادة منه عن طريق الاطلاع ، مع إمكانية نسخ المحتوى والاستفادة منه ، والمخزن هو ما يمكن من الاطلاع عليه والاستفادة منه فقط دون إتاحة عملية النسخ .

- إيجاد مورد لغوي وأداة للبحث والتطوير في ميداني اللغة والمعاجم .
 - إيجاد وسائل لتطوير صناعة المعاجم بما تحويه من المعاني والعلاقات الدلالية بين المفردات .
 - المساهمة بالتكامل بين المحلل الآلي وأدواته المتفرعة عنه ومدونة التأسيس للغة العربية في صناعة المعجم التاريخي المنشود للغة العربية .
 - إيجاد أداة فعالة للبحث والتطوير في مجالات تطبيقات الذكاء الاصطناعي ومعالجة اللغات الطبيعية ، والويب الدلالي الذي يطلق عليه الأنطولوجيا .
- وذلك لأن وجود جامع معياري للكلمات العربية يفتح الطريق أمام بناء أنطولوجيا^(١) معرفية عربية كاملة للغة العربية ، وهو مشروع طموح^(٢) .

رابعا - بناء برنامج المدقق والمصحح الآلي للنصوص العربية :

وهو برنامج يقف على الأخطاء الإملائية داخل النصوص العربية ويقوم بتصويبها بشكل آلي (بأسماؤهم : بأسمائهم) ، كما أنه يقف على الأخطاء المتعلقة بالكلمات المشكولة بنية وإعرابا ، ويقف على الأخطاء النحوية المتعلقة بالرسم الإملائي كما لو دخلت «أن» الناسخة على كلمة مرفوعة بالواو ، على سبيل المثال : (أن المسلمون . . .) ، وفوق هذا يقف على الكلمة الصحيحة لغوياً ، ولكنها خطأ في سياقها ويقوم بتصويبها **وقد تم إنجاز هذا البرنامج ، ولله الحمد ،** ويتم تطويره وتحسينه بشكل مستمر .

(١) الأنطولوجيا هي هيكل معرفي يتكون من خلال الترابط بين الشبكات الدلالية للكلمات فيبرز مفاهيم مختلفة تعطي شكلا من أشكال التمثيل المعرفي .

(٢) نقلا عن كتاب «علم اللغة والتقنيات المعاصرة» تأليف د . المعتز بالله السعيد ، بتصرف .

خامسا - بناء برنامج المشكل الآلي للنصوص العربية :

وهو برنامج يقوم بتشكيل النصوص العربية بشكل آلي بحيث يستطيع التعامل مع جميع النصوص العربية لمختلف الفنون ، وقد أنجز ووصلت دقته إلى درجة مرضية لنصوص اللغة العربية حتى تاريخه ، بينما بلغت دقة المشكل لنصوص الحديث النبوي (٩٨٪) ثمانية وتسعين في المائة ولله الحمد .

والعمل جار على استكمال تطوير برنامج المشكل الآلي ليعطي نتائج ذات دقة أكبر .

سادسا - بناء مدونة القرآن الكريم اللغوية :

هي قاعدة بيانات تم فيها فك اللبس الصرفي والدلالي لنصوص القرآن الكريم ، وبناء محرك بحث آلي على مستويات عدة : على مستوى رقم السورة ، وعلى مستوى رقم الآية داخل السورة ، وعلى مستوى الكلمة ، وعلى مستوى المعنى ، وعلى مستوى قسم الكلم ، وعلى مستوى الصيغة الصرفية ، وعلى مستوى المشتقات ، وعلى مستوى الموضوع ، مما يمكن أن يُطلق عليه مسمى «مدونة القرآن الكريم اللغوية» ، ومن خلال هذه المدونة نستطيع استخراج إحصاءات ومعلومات مختلفة سواء على مستوى النص القرآني كاملا ، كمعرفة الجذور ومشتقاتها التي اختارها الله ﷻ لإبلاغ رسالته للعالمين من خلال القرآن الكريم ، أو على مستوى السورة الواحدة لعدد الجذور المستخدمة فيها ، وكذا معرفة عدد أقسام الكلم والصيغ الصرفية ، وعدد استخدام الفعل المضارع أو الماضي معلوماً أو مجهولاً أو فعل الأمر ، وغير ذلك كثير من الإحصاءات التي لا تخلو من لمحة إعجاز لدى المتدبرين فيها ، وقد تم الفراغ منها في إصدارها الأولي ولله الحمد ، ويجري العمل على البلوغ بها إلى أعلى درجة من الاستيعاب الكامل والإتقان .

سابعا - بناء مدونة السنة النبوية اللغوية :

وهي قاعدة بيانات يتم فيها فك اللبس الصرفي والدلالي لنصوص الحديث النبوي ، وبناء محرك بحث آلي على عدّة مستويات منها ما هو : على مستوى الحديث كاملا ، أو على مستوى المتن فقط ، أو السند فقط ، أو على مستوى الكلمة ، أو على مستوى المعنى ، أو على مستوى قسم الكلم ، أو على مستوى الصيغة الصرفية أو على مستوى المشتقات ، أو على مستوى الموضوع ، مما يمكن أن يطلق عليه مُسمى : «مدونة السنة النبوية اللغوية» .

ومن خلال هذه المدونة نستطيع استخراج إحصاءات ومعلومات مختلفة سواء على مستوى نصوص كتب الحديث النبوي كلها كمعرفة الجذور ومشتقاتها التي هدى الله رسوله ﷺ لاختيارها تبليغا لرسالته للعالمين ، أو على مستوى كتاب واحد من كتب السنة أو على مستوى عدة كتب .

ومن خلال هذه المدونة نستطيع كذلك تبيان عدد أقسام الكلم والصيغ الصرفية ، وعدد استخدام الفعل المضارع أو الماضي معلومًا أو مجهولًا أو فعل الأمر ، وغير ذلك كثير من الإحصاءات التي لا تخلو من لمحة إعجاز لدئ المتدبرين وقد تم الفراغ منها في إصدارها الأولي ولله الحمد ويجري العمل على البلوغ بها إلى درجة الاستيعاب الكامل والإتقان .

ثامنا - بناء مدونة الشعر العربي اللغوية :

وهي قاعدة بيانات تحوي قرابة مليونين ومائتي ألف بيت من الشعر تشمل ما دون ووصل إلينا من شعر منذ العصر الجاهلي حتى العصر الحديث ، معزواً كل بيت من أبيات الشعر فيها إلى قائله ، مع بيان تاريخ وفاة الشاعر ، وكذا تصنيف تلك القصيدة أو غيرها من حيث الغرض الشعري والبحر الشعري التي نُسجت على منواله ، وقد تم فيها فك اللبس

الصرفي لجميع أبيات الشعر وجر العمل على فك اللبس الدلالي ، كما تم بناء محرك بحث آلي يبحث على مستويات عدة منها : على مستوى اسم القصيدة أو اسم الشاعر ، وعلى مستوى البيت ، وعلى مستوى الكلمة أو التركيب ، وعلى مستوى قسم الكلم ، وعلى مستوى الصيغة الصرفية ، وعلى مستوى المشتقات وهكذا .

ومن خلال هذه المدونة الشعرية نستطيع استخراج :

- إحصاءات ومعلومات مختلفة لعدد الجذور المستخدمة سواء على مستوى المدونة الشعرية كلها أو على مستوى كل فترة زمنية محددة تاريخياً أو وفقاً للترتيب التاريخي لورود الكلمة ، أو على مستوى شعر شاعر معين أو على مستوى القصيدة الواحدة .
 - معرفة عدد أقسام الكلم والصيغ الصرفية الموجودة في المدونة .
 - معرفة عدد مرات استخدام الفعل المضارع أو الماضي معلوماً أو مجهولاً أو فعل الأمر .
 - إمكانية رصد التطور الصرفي والدلالي الذي طرأ على مفردات اللغة أو نصوصها .
- وقد تم بفضل إنجاز هذه المدونة ويجري العمل على تدقيقها وتحسينها .

تاسعا - بناء مدونة التأصيل للغة العربية :

وهي عبارة عن مجموعة ضخمة مهيكلة من النصوص اللغوية العربية الكاملة المكتوبة المحوسبة ، وتتمتع بمحرك بحث عالي الكفاءة وواسع الإمكانيات .

وقد روعي في التخطيط العلمي والتنفيذ لاختيار نصوص هذه المدونة ما يلي :

- ١- أن تكون مُمثلة لقطاع إقليمي كبير من الدول الناطقة باللغة العربية ، فيراعى التنوع المكاني أو الجغرافي ، بحيث يتم اختيار النصوص من معظم البلدان العربية إن لم يكن كلها ، ومن غير الدول العربية ما أمكن ذلك طالما أن النص المكتوب عربيّ .

٢- أن تمثل - بشكل حقيقي وواقعي- أنماط استخدام اللغة العربية المعاصرة والقديمة في أنحاء العالم العربي ، فيُراعى فيها التنوع الأسلوبي بحيث تكون النصوص لكُتّاب ومؤلفين مختلفين .

٣- أن تمثل التنوع التاريخي في محتواها بداية من أقدم النصوص العربية المكتوبة في العصر الجاهلي حتى العصر الحديث .

٤- أن تمثل التنوع الموضوعي للنصوص بحيث تتوافر فيها النصوص الشرعية من قرآن وسنة والنصوص الأدبية من شعر ونثر، والمصطلحات الطبية ، والقانونية ، والزراعية ، وغيرها .

٥- أن تكون المدونة متنوعةً من حيث فئات النصوص ومحتواها .

٦- أن تحوي المدونة المعلومات الخاصة بالنصوص المدخلة مثل المصدر، ونوع النص ، وتاريخ نشوء النص ، واسم المؤلف ، وتاريخ وفاته .

٧- أن تغطي المدونة مصطلحات العلوم الحديثة : الطب ، والهندسة ، والكيمياء ، والفيزياء ، وغيرها .

٨- وقد يسر الله **لدار التأسيس** إدخال ثلاثة وعشرين ألفاً ومائتي مرجع والعمل جارٍ للوصول بالمراجع لخمسين ألف مرجع .

ومن أهداف إنشاء مدونة التأسيس للغة العربية :

- أن تكون المدونة مَورداً ثَمَوياً للباحثين عامةً والباحثين اللغويين خاصةً؛ لتفيد في رصد ووصف التطورات التي طرأت على الكلمة العربية صرفاً ودلالة من خلال الاستخدام الواقعي والتاريخي لكلمات اللغة العربية .

- **الاستفادة من المدونة في بناء المعجم التاريخي لألفاظ لغة العربية** وبناء المعجم السياقي والمعجم المتخصصة والمعجم متعددة اللغات ، وكذا المعجم المبنية على نظرية الاستعمال الواقعي للغة العربية ، فمن خلال هذه المدونة يُمكن معرفة مدى شيوع استخدام الكلمة ، ومعرفة المعاني المختلفة للكلمة ، ومعرفة تطورها الصرفي والدلالي عبر تاريخ استخدامها .

وقد تم الفراغ بحمد الله من إنجاز هذه المدونة بصيغة مرضية وبلغت كلماتها مليارين ومائتي مليون كلمة ، ومن المخطط له أن تبلغ كلماتها ثلاثة مليارات كلمة ، وقد تم ربطها بالمحلل الآلي للغة العربية ، ويجري العمل على تحسينها وتجويد نصوصها .

ومما تسعى دار التأسيس لتنفيذه في هذه المدونة بصيغتها النهائية هو:

- أن تكون كل كلمات المدونة محللة تحليلا صرفياً بحيث يتم اختيار التحليل الصرفي المناسب للكلمة داخل السياق الواردة فيه .
- أن تكون كلمات المدونة محللة تحليلا نحويًا بحيث يبين الموقع الإعرابي لكل كلمة داخل السياق الواردة فيه .
- أن تكون الكلمة داخل المدونة محللة تحليلا دلاليًا بحيث يتم اختيار المعنى المناسب للكلمة داخل السياق .

هذا ويمكن البحث في المدونة على مستوى : الكلمة الواحدة ، أو جزء من الكلمة ، أو التركيب ، وكذا الكلمة مشكولة أو غير مشكولة ، والصيغة الصرفية ، والجذر ، وقسم الكلم ، والساق ، والكلمة ومشتقاتها ، والكلمة وتصريفاتها .

القسم الثاني: ما تطمح دار التأصيل لإنجازه من مشروعات معالجة اللغة العربية آلياً:

إضافة إلى المشروعات الضخمة المتميزة التي أنجزتها **دار التأسيس**، والتي سبق ذكرها، هناك مشروعات أخرى قوية تخطط **دار التأسيس** في الشروع في تنفيذها مستقبلاً، وقطعت في التخطيط لبعضها شوطاً لا بأس به، وتطمح لإنجازها بعون الله، ومنها:

١- بناء مدونة الفقه الإسلامي اللغوية:

وهي قاعدة بيانات متخصصة في ألفاظ الفقه الإسلامي، يتم فيها فك اللبس الصرفي والدلالي لنصوص الفقه الإسلامي، وبناء محرك بحث آلي على عدّة مستويات، أطلق عليها مُسمى: «**مدونة الفقه الإسلامي اللغوية**»، ومن خلال هذه المدونة نستطيع استخراج إحصاءات ومعلومات وأحكام وتصنيفات مختلفة.

٢- بناء برنامج الرمز الآلي:

الرمز الآلي هو برنامج حاسوبي يقوم بعملية فك اللبس الصرفي والدلالي آلياً بعدما يتم تعليمه من خلال ترميز يعنى (بفك اللبس الصرفي والدلالي) لنصوص **مدونة التأصيل للغة العربية**، أو أي نص عربي وستتم الاستفادة من برنامج الرمز الآلي في فك لبس أي نصوص جديدة يتم إضافتها للمدونة اللغوية، حيث إن المدونة قابلة لزيادة النصوص فيها.

٣- بناء برنامج القارئ أو الناطق الآلي:

وهو برنامج من المفترض أن يقوم بقراءة أي لفظة من ألفاظ اللغة العربية بشكل آلي، وسيُعتمد في بنائه على برنامج المشكل الآلي، ويمكن أن يوضع ذلك البرنامج عند إنجازه وتطبيقه مع جميع المعاجم الإلكترونية بحيث ييسر على غير الناطقين باللغة العربية تعلم نطق ألفاظ اللغة العربية وقراءتها وستكون لهذا البرنامج إفادة للمكفوفين ولغير العرب.

٤- بناء المعجم السياقي للغة العربية:

المعجم السياقي هو معجم يحتوي فيه المفردة على الخصائص الصرفية والنحوية ، وطريقة الكتابة (الشكل الإملائي) ، والمعاني المختلفة مع أمثلة وشواهد لمختلف الاستعمالات فهو كالمعجم التقليدي الذي يضم مفردات اللغة مقرونة بمعانيها ، ولكنه يزيد عليه بأنه يذكر المفردة بمعناها المطلوب ، إضافة لشواهد من نصوص مختلفة تتضمن هذه الكلمة بمعانيها المختلفة .

٥- المساهمة في بناء المعجم التاريخي لألفاظ اللغة العربية:

وهو معجم يتتبع الكلمة المفردة منذ ظهورها تاريخياً حتى العصر الحالي ذاكراً معاني الكلمة المختلفة حسب سياقها ، راصداً ما طرأ عليها من تطور صرفي ودلالي ، وكذا استعمالها أو عدم استعمالها ، وهو ما أسس مجمع اللغة العربية في القاهرة من أجله منذ ما يقارب مائة عام ، ولم يتمكن من إنجازه أو البدء به لضخامته ولاعتبارات عدة .

وبفضل من الله فإن كازالتأصيل من خلال مركز البحث والتقنية للمعلومات تمتلك أهم الأدوات اللازمة لإنجاز هذا المعجم التاريخي وأهمها :

- **المحلل الآلي لألفاظ اللغة العربية صرفاً ودلالة :** الذي يشتمل على كل مفردة من مفردات اللغة مزودة ببياناتها الصرفية ومعانيها المختلفة القديمة والحديثة ، وقد ذكرت تفاصيله فيما سبق .
- **مدونة دار التأصيل للغة العربية :** وهي قاعدة معلومات تفاعلية للنصوص العامة للغة العربية التي يمكن من خلالها إيجاد الشواهد على تطور الكلمات المفردة للغة العربية ومعانيها تاريخياً ، مما يسهل على المتخصصين رصد التطور التاريخي دلالياً وصرفياً ومعرفة التطورات التي مرت بها مفردات اللغة العربية حتى هذا العصر .
- **المدقق والمصحح الآلي لألفاظ اللغة العربية .**
- **المشكل الآلي لألفاظ اللغة العربية .**

٦- بناء المعاجم المتخصصة:

القيام ببناء معاجم في تخصصات لغوية منها:

● معجم غريب ألفاظ القرآن الكريم .

● معجم غريب ألفاظ الحديث .

● المعجم التاريخي لألفاظ الجرح والتعديل .

وقد تم بحمد الله إنجاز هذه المعاجم بصيغتها الأولى .

وسيتم العمل مستقبلاً بعون الله على معاجم أخرى منها:

● معجم المترادفات والمتضادات .

● ومعجم التعبيرات الاصطلاحية .

وصولاً إلى غيرها من المعاجم المتخصصة التي تزيد الاستفادة من لغتنا العربية .

٧ - بناء معاجم للكلمات الأكثر تداولاً في اللغة العربية على مستويات عدة:

تبدأ من المرحلة الابتدائية مروراً بالمرحلة المتوسطة والثانوية وكذا المرحلة الجامعية والكلمات الشائعة في الصحافة وغير ذلك .

٨- إنشاء معاجم متعددة اللغات:

وهي معاجم تحوي أهم مفردات اللغة العربية الحية بالإضافة إلى ترجماتها المختلفة إلى اللغات العالمية الحية ، بالإضافة إلى المعاجم المتعددة اللغات التي تساعد على نشر اللغة العربية والتعامل معها والتعرف عليها عالمياً ، وبالتبع الاطلاع على الدين الإسلامي ومنتجات الحضارة الإسلامية .

٩- بناء برامج لتعليم كل من الصرف والنحو والإملاء والتدريبات التطبيقية المترتبة عليها:

سيمكن **المحلل الآلي للغة العربية** وقرينته **مدونة التأسيس للغة العربية** المتخصصين من بناء برامج تعليمية حاسوبية يتفاعل الطالب معها ، وتعتمد هذه البرامج في محتواها على الجداول والرسوم ذات الدلالات المختلفة التي تُساعد على توصيل المعلومة وتثبيتها في عقل الطالب ، وسيراعى في إنشاء هذه البرامج أن تتناسب مع جميع الأعمار من حيث عدد المفردات اللغوية ومعانيها وكذا أساليبها الكتابية ، وبهذا ينجذب الطلاب إلى تعلم اللغة وإتقانها ويُقبلون عليها ، ولا يخفى على أحد في هذا العصر أهمية الاستعانة بالتقنيات الحديثة في العمليات التعليمية .

١٠- بناء برامج تعليم اللغة العربية لغير الناطقين بها:

سيمكن **المحلل الآلي للغة العربية** وقرينته **مدونة التأسيس للغة العربية** المتخصصين من بناء برامج تعليمية حاسوبية يتفاعل معها الطالب الذي لا يعرف اللغة العربية مع البرامج التعليمية وصولاً إلى تعلم اللغة وإتقانها كتابةً ونطقاً ، مما يُساعد على نشر اللغة العربية بين المسلمين وغيرهم .

هذا ، وما ذكر سابقاً ليس إلا أمثلة لكيفية تطوير التعامل مع لغتنا العربية الخالدة وصولاً إلى إنزالها منزلتها التي تليق بها في هذا العصر .

وبهذا يتبين أن إنشاء **المحلل الآلي للغة العربية** وقرينته **مدونة التأسيس للغة العربية** وما يتبعها من أدوات بالإضافة **لمدونات اللغة العربية المتخصصة في القرآن والسنة النبوية والشعر** وترجمة ألفاظ اللغة العربية الحية للغة الإنجليزية وغيرها من اللغات العالمية وغيرها ، كل ذلك يعد خطوات إستراتيجية ونقل حضارية نوعية لتسهيل التعامل مع اللغة العربية آلياً ، وحصراً مفرداتها وتوثيقها وتسجيل توثيقها دولياً ، وجعلها معاصرةً ومواكبةً لتطورات الحضارة المعاصرة المعتمدة على كبر على البرامج الحاسوبية التي سببت انتشار وتطور اللغات العالمية المعاصرة .

وكل هذا هو مما ينبغي أن تتوجه إليه جهود الدول والهيئات والجامعات والمؤسسات
وأفراد الباحثين المتخصصين باللغة العربية وعلومها وأن يحظى بالعناية والاهتمام من الذين
مكن الله ﷻ لهم وعيا وعلما ومكانة ومالا .

يسر الله لنا إتمام ما نصبو إليه من تطوير للتعامل مع لغتنا العربية آليا وتيسيرها وتوثيقها
ووضعها في مكانها الحضاري المعاصر .

وبالله التوفيق ومنه العون والسداد وله الحمد والشكر .

عبد الرحمن بن عبد الله بن عقيل

القاهرة في: يوم الأحد: ٢٢/٠٤/١٤٤١هـ

الموافق: ١٩/١٢/٢٠١٩م

المشرف العام على دار التأصيل

مركز البحث والتقنية المعلومات