

قواعد بيانات القرآن الكريم كأساس للمعجم الآلي الموسع للغة العربية

اعداد: محمد الجاسم .
الرقم الجامعي: ٢٢٣٦١٨ .
لمادة: الحوسبة العربية .
باشراف الاستاذ: حسني المحتسب .

المقدمة

- تحتاج المعالجة الآلية للغة العربية معجمًا موسعًا يحوي كل ألفاظ اللغة العربية الفصحى المستعملة اليوم أو سابقاً.
- وهذا المعجم يجب أن يحوي كل المعلومات عن الألفاظ التي تحتاجها عمليات المعالجة من دلالات مختلفة وصرف وما يجوز أن تستعمل معه الكلمة وما لا يجوز والكلمات المقاربة والمضادة وكل ما يوصل إلى دقة دلالة الكلمة من وسائل.
- وإن دراسة مفردات القرآن الكريم هي خير بداية للوصول إلى المبتغى المطلوب لتصميم المعجم الآلي الموسع للغة العربية.

بين النص القرآني والكتابة العربية اليوم

- التعامل مع النص القرآني يختلف عن التعامل مع نصوص اللغة العربية بما يلي:
- ❖ النص القرآني يأتي مشكلاً في العادة ، والتشكيل يسهل التعامل مع النص من عدة نواحي حيث يزيل الالتباس في معالجة النص الآلية، والصرفية، والنحوية، والدلالية، مع ان اللغة العربية قبل وفي صدر الاسلام لم تكن مشكلة أو منقطة.
- ❖ النص القرآني غالباً ما يكون بالرسم العثماني بخلاف النصوص العربية الأخرى.
- ❖ أيضا يحتوي النص القرآني على علامات ورموز خاصة به ولا تستخدم في نص عربي آخر (الوقف ، والتجويد).
- ❖ في النص القرآني فواصل تفصل بين آية وأخرى ، ولا تكون الآية دائما جملة كاملة.
- ❖ الخطأ في التعامل مع الآيات القرآنية يواجه اعتراضا دينيا واضحا .. فلا بد من الحرص الشديد.

إعداد النص القرآني

- كما ذُكِرَ سابقاً أنّ الخطأ في القرآن الكريم غير مسموح به مهما كان مقداره، ولكن فيما عدا القرآن قد يسمح به بمقدار صغير جداً.
- فاعداد القرآن الكريم لا بد أن يمر بعدة مراحل تدقيق وتأکید على صحة النص وسلامته ...
- ونتيجة لذلك اعتمدت أسس للتعامل مع النص القرآني :
- ❖ اعتماد الرسم العثماني ، و قراءة حفص بن عاصم ومصحف المدينة المنورة.
- ❖ أيضا اعتبار البسملة كأول آية من القرآن الكريم من سورة الفاتحة دون تكرارها في بقية السور من دون داعي(التوبة).

المعالجة الأولية للنص القرآني

- تم اعتماد النص القرآني آية آية .. قيدا مستقلاً (Record) ، أو كل آية في سطر مستقل.
- بين الكلمات في كل آية فراغ فاصل ...
- اما بين آية وأخرى فلانحتاج لفاصل حيث ان كل آية في سطرأ مستقلاً ..
- أدى ذلك إلى عدم الحاجة الى ترقيم الآيات .. فيمكن عد أو ترقيم الآيات عند المعالجة الحاسوبية ببساطة.
- اختيرت العلامة # كعلامة فاصلة بين سورة وأخرى .
- أما علامات الوقف والتجويد فقد اختير لكل علامة رمز حاسوبي يميزها (\$ & % + -) وأيضا لمواقع سجود التلاوة.

من قواعد بيانات القرآن الكريم...

- قاعدة بيانات رسم الكلمات :
 - قاعدة بيانات لرسم المصحف بالرسم العثماني .. فالكلمات في هذه القاعدة هي كلمات الرسم العثماني محذوف منها علامات الوقف والتجويد.
 - هناك كلمات قرآنية تختلف في شكل كتابتها قديما وحديثا منها :
١. استعمال الألف الخنجرية في كثير من الكلمات مثل العالمين (تكتب العلمين).
 ٢. كتابة الهمزة بشكل مختلف عما هو متعارف عليه في الخط الحديث مثل (الذكرين) فهي تكتب اليوم هكذا أذاكرين (مدة فوق الألف).
 ٣. واو الجماعة تكتب أحيانا بالألف أو بدونها كما ان أحيانا تكون هناك واو أصلية مضاف إليها ألف مثل (أدعوا).
 ٤. كتابة بعض الكلمات بشكل خاص مثل الربا تكتب (الربوا).
 ٥. دمج بعض الكلمات بطريقة مغايرة لما متعارف عليه اليوم مثل " يا ابن أمي " (يبنؤم) و " فما للذين " (فمال الذين)
- وعلى هذا فإن أول خطوة في تكوين قاعدة بيانات رسم الكلمات هو إعادة كتابة الكلمات المختلف عليها برسم حديث فتتكون قاعدة البيانات هذه من حقلين الأول هو الرسم العثماني الدقيق والثاني هو الرسم الحديث.

من قواعد بيانات القرآن الكريم...

- قاعدة بيانات الجمل القرآنية :
- كما ذكر أن الجمل القرآنية قد تكون أكثر من آية لجملة واحدة أو أكثر من جملة في آية واحدة أي ان الجمل ستكون متداخلة بين الآيات.
- فلذلك كان لابد من وضع علامات تفصل بين الجمل القرآنية ، فعلامات الوقف والتجويد قد يشير بعضها إلى نهاية الجملة .
- وقد تختلف مواقع تجزئة النص إلى جمل بحسب فهم من يقوم بها من المفسرين :
- مثال على ذلك تقسيم بداية سورة البقرة على طريقتين :

@الم
ذَلِكَ الْكِتَابُ لَا رَيْبَ
فِيهِ هُدًى لِّلْمُتَّقِينَ@

• ويمكن أن تحزأ كالآتي:

@الم
ذَلِكَ الْكِتَابُ لَا رَيْبَ فِيهِ
هُدًى لِّلْمُتَّقِينَ@

- وهناك بعض المواضع التي تشير إليها علامات خاصة من علامات الوقف والتجويد التي تبين إمكان الوقف على أي من الكلمتين (فيه أو هدى) بشكل متبادل (أي أحدهما فقط) . ولكل من الصيغتين إعراب مختلف لكل جملة.

من قواعد بيانات القرآن الكريم...

- قاعدة بيانات الصرف :

- في قاعدة بيانات الصرف يتم تجزئة الكلمات القرآنية إلى لواصقها الأولية والنهائية والتشكيل، وتكون التجزئة ٤ لواصق أولية كحد أعلى و٢ لواصق نهائية وجذع الكلمة وتشكيله وتشكيل آخر الكلمة.
- بناء على ذلك يحجز ٩ حقول للكلمة، بعض الأمثلة :
- ❖ كلمة " أفبالباطل " تحتوي ٤ لواصق أولية هي " أ - فَ - يَ - ال " وجذعها هو باطل يعقبه كسرة في الأخير .
- ❖ كلمة " فسيكفيكمهم " فتحتوي على لواصق أولية عددها ٢ هما " فَ - سدَ " والجدع يكفي واللواصق الأخيرة هي " كَ " و " هم " والتشكيل على الجذع هو السكون (غير الظاهر على الياء) وتشكيل آخر اللواصق هو السكون أيضا.

من قواعد بيانات القرآن الكريم...

- قاعدة بيانات اللفظ :
- قاعدة بيانات اللفظ تحتوي على ما يأتي :
 ١. قاعدة بيانات القراءات القرآنية:
فالمعروف أنه على الرغم من التزام كل القراء بالاعتماد على الرسم العثماني إلا أن هناك اختلافات طفيفة بينهم ويمكن جمع هذه الاختلافات في قاعدة بيانات تبين ذلك .
 ٢. قاعدة بيانات الألفاظ الشاذة:
وهي كلمات خاصة تلفظ بصيغة معينة في بعض القراءات دون أخرى .
 ٣. قاعدة بيانات الأصوات:
فالمعروف إن وحدة الكلام المنطوقة هي الأصوات، وليس الحروف وتحتوي اللغة العربية على أصوات لا توجد في غيرها كما أن أصوات التجويد وخاصة الممدود ولفظ النون الساكنة والميم الساكنة والإدغام والإمالة والتخفيف تنطق بشكل معين . وتجميع رموز هذه البيانات يجعلها مفيدة للاستعمال في دراسات وأبحاث اللغة العربية .
 ٤. قواعد البيانات الصوتية:
أصبح بالإمكان تجزئة الصوت بالحاسوب وتقطيعه وإعادة تركيبه والتعرف عليه.
فإذا ما خزن صوت نموذجي لقارئ ما وكان هذا النموذج وافيًا وقطع إلى أجزاء مناسبة فإن بالإمكان إعادة تركيب هذه الأجزاء لقراءة نص آخر مغاير للنص الأصلي لكنه يحتوي على تلك القطع والأجزاء وبصوت القارئ نفسه .

وأخيرا

- إضافة الى ما ذكر سابقا .. ان معالجة القرآن الكريم تساعد كثيرا في تكوين المعجم الآلي الموسع للغة العربية ، حيث ان القرآن يحوي كثير من الأمثلة التي تساعد في تكوين المعجم الآلي .. و لكن ليس كافيا لتغطية كل الكلمات والمعاني في اللغة العربية فلذلك يجب توسيع القواعد المستخرجة منه لكي تشمل كل الكلمات والمعاني في اللغة العربية .



هل لدى أحد أي سؤال ➤