

برمجان اللغة العربية مجمع الملك سلمان العالمي للغة العربية

المسار: تحدي المعجم

المجال: استخراج الأصل المعجمي Lemma (الكلمة الأصلية التي تمثل المدخل المعجمي)

اسم المشارك: حمادة عبيد أحمد إبراهيم

نبذة

تطوير تطبيق يمكنه تحليل الكلمات بصورها المختلفة واستخراج جذرها اللغوي والساق التي يمكن البحث بها في المعاجم.

المشكلة

تكون الكلمة الواحدة عند الاستعمال في صور متعددة، فقد تكون محلاة بالألف واللام أو دونهما أو بها سوابق لاصقة كحروف الجر، أو لاصقة مثل الضمائر وغيرها (برمجان، ببرمجان، وبرمجان، فبرمجان، كبرمجان، لبرمجان، البرمجان، بالبرمجان، فالبرمجان،... إلخ)، وللبحث عنها في المعاجم لا يمكن البحث بكل هذه الصور لأن مداخل المعاجم إما أن تعتمد على جذر الكلمة أو ساقها (برمجان)، لذا فهناك إشكال في البحث عن الكلمة بسوابقها ولواحقها ولا بد من إرجاعها إلى جذرها أو ساقها (برمج، برمجان).

الحل

الحل المقترح (ليس بتقنيات الذكاء الاصطناعي) هو بناء قاعدة بيانات تشمل كل أبنية الكلمة العربية بالإضافة لكل جذور العربية، ثم وضع خوارزمية دقيقة ومحكمة يمكنها تحليل أي كلمة عربية وإرجاع ساقها وجذرها من خلال قاعدتي بيانات الجذور والأبنية حتى يمكن التعرف على معناها في المعاجم المختلفة.

البيانات

قاعدة بيانات الجذور (من صنع المطور)

قاعدة بيانات الأبنية (من صنع المطور)

القيمة المضافة

إمكان الوصول آليا لمعاني الكلمات مهما كانت صورتها (مجردة، بسوابق فقط،
بلواحق فقط، بسوابق ولواحق) بالبحث عنها في المعاجم بعد استخراج جذرها
وساقها.