

## محركات بحث المادة الصوتية على الإنترنت: دراسة تحليلية لآليات

### البحث والاسترجاع واقتراح مواصفات محرك صوت يدعم اللغة العربية<sup>(\*)</sup>

عرض

أحمد سيد عبد المجيد

ومعاملات البحث الملائمة لهذا الشكل تحديداً، مع محاولة اكتشاف إمكانية استرجاع مثل هذا النوع من الوسائط باستخدام الكلمات العربية، وهل هناك محركات بحث عربية تفيد في استرجاع ملفات الصوت أم لا؟ ومدى حاجة الباحث العربي لمثل هذه الأدوات.

خاصة وقد حظيت مصادر المعلومات على الإنترنت بالجدل والنقاش، حيث دعا فريق إلى عدم معالجة تلك المصادر فنياً بحجة عدم ثباتها وسرعة تحديثها أو حذفها، بينما دعا فريق آخر إلى ضرورة تنظيمها لأهميتها لكل من المكتبة والمستفيد، على أن تُتخذ في سبيل ذلك الإجراءات التي تكفل توثيق بيانات المصدر.

في ظل الكم الهائل من أوعية المعلومات الرقمية المتاحة على الإنترنت وتنوعها، تمثل ملفات الصوت الرقمية مصدراً مهماً من مصادر المعلومات إلى جانب النص والصورة.

ويُعد الصوت ذو أهمية كبرى سواء للأغراض التعليمية أو البحثية أو أغراض إفادة بعض الفئات ذوي الاحتياجات الخاصة كالمعاقين بصرياً، ونظراً لما يتضمنه الصوت من خصائص بإمكانياته المسموعة فهو ذو تأثير فعال في داخل المستفيد أو المتلقي.

من هنا جاءت الفكرة للباحث لتناول طبيعة تنظيم وبحث واسترجاع النتائج من الوحدات الصوتية داخل محركات بحث الويب. وتطور الفكرة أساساً حول آليات

(\*) محركات بحث المواد الصوتية على الإنترنت: دراسة تحليلية لآليات البحث والاسترجاع واقتراح مواصفات محرك صوت يدعم اللغة العربية / أحمد سيد عبد المجيد حسن ؛ إشراف زين الدين محمد عبد الهادي، محمد حسن عبدالعظيم - ماجستير، كلية الآداب، قسم المكتبات والوثائق، ٢٠١٤.

٤. عدم وجود معايير لتنظيم وتكشيف المادة الصوتية توافق متطلبات البحث باللغة العربية على المستوى الرقمي.
٥. افتقار محركات البحث العربية لإمكانية البحث عن المواد الصوتية.
٦. محركات البحث الأجنبية تتعامل مع استرجاع ملفات الصوت بالنص العربي بصورة تفتقر إلى الدقة والكفاءة وعدم المعرفة بخصائص النصوص العربية.
- وقد عملت الدراسة على تحقيق الأهداف التالية: -**

١. التعرف على الإنتاج الفكري المنشور في مجال استرجاع ملفات الصوت الرقمية.
٢. التعرف على خصائص ملف الصوت الرقمي مقارنة بملف الصوت التقليدي.
٣. التعرف على خصائص المادة الصوتية الرقمية، وطبيعتها كوعاء معلومات مسموع مقارنة بالنص الرقمي.
٤. حصر وتقييم آليات عمل محركات البحث المتخصصة في بحث ملفات الصوت.

### وتتلخص مشكلة الدراسة في:

١. زيادة ملفات الصوت الرقمية المتاحة على الإنترنت، مع عدم تطور آليات الوصف الببليوجرافي لها.
٢. عدم تطور محركات البحث بالقدر الكافي لملائمة العمل مع ملفات الصوت الرقمية كنظم للاسترجاع.
٣. أهمية تبني مواصفات ومعايير موحدة خاصة لوصف واسترجاع ملفات الصوت، لما تنفرد به من خصائص ومميزات.

### وقد اكتسبت الدراسة أهميتها

ومبرراتها من: -

١. اهتمام مرافق المعلومات على اختلاف أنواعها باقتناء أوعية المعلومات الالكترونية والتعامل معها من خلال البيئة الالكترونية.
٢. أهمية ملفات الصوت الرقمية كأوعية لنقل المعلومات وسط هذا الكم المتنوع من الإنتاج الفكري باختلاف أشكاله.
٣. تزايد الاهتمام بمحركات البحث الذي وافق الاهتمام المتنامي باستخدام الانترنت ومصادرها المتنوعة.

المسترجعة وتحليلها، ثم قياس مدى فعالية هذه النتائج مع أساليب الاسترجاع الآلي، وخصائص اللغة العربية واقترح معايير ومواصفات لمحركات بحث صوت تدعم اللغة العربية.

أما عن أدوات جمع البيانات فهي:

### بحث الإنتاج الفكري:

استخدمت هذه الأداة لجمع الجانب النظري من الدراسة والتي تتعلق ببناء الويب، وعمل محركات البحث في جمع مصادرها وطريقة عملها، إلى ما يتعلق بملفات الصوت الرقمية وخصائصها.

### المعايشة مع الإنترنت:

وذلك من خلال جلسات الاتصال المباشر مع الانترنت ومراقبة أساليب البحث والاسترجاع داخل محركات البحث وتقييم عمل هذه الأساليب ونتائجها.

وقد جاءت الدراسة في مقدمة وأربعة فصول، يمكن استعراض فصول الدراسة بإيجاز فيما يلي:-

**الفصل الأول:** يتناول أنواع المواد الصوتية، وتطبيقاتها في مختلف المجالات. بالإضافة لتطور الشكل الرقمي لتلك الملفات حتى أصبحت أحد وسائط المعلومات المهمة على

٥. التعرف على الخصائص البحثية والاسترجاعية لمحركات الصوت محل الدراسة، حتى يتسنى اختيار أفضلها وفقاً للاحتياجات البحثية.

٦. تحديد معايير واصفات تكشف واسترجاع ملفات الصوت الرقمية.

٧. دراسة خصائص اللغة العربية في كشف ملفات الصوت واسترجاعها.

٨. اقتراح مواصفات محرك صوت يدعم اللغة العربية.

وقد اعتمدت الدراسة على استخدام

### مناهج البحث التالية :

#### المنهج الوصفي

وذلك للتعرف على محركات بحث الصوت محل الدراسة وإمكانياتها، حتى يتسنى إجراء المنهج التجريبي عليها لاكتشاف آليات البحث فيها، ونواحي القصور داخلها، لاقتراح محرك بحث صوت يدعم اللغة العربية.

#### المنهج التجريبي

وذلك من خلال استخدام مجموعة من الواصفات مع محركات بحث الصوت محل الدراسة واستخلاص النتائج

الإنترنت. واختتم الفصل بالحديث عن محركات بحث الصوت.

**الفصل الثاني:** يتناول هذا الفصل عمليات الإعداد الفني لوسيط المعلومات الصوتي من الناحية المادية والموضوعية، وذلك من خلال أشهر المعايير والقواعد السائدة مثل قواعد الفهرسة الأنجلوأمريكية AACR2، ومعيار MARC21، ومعيار دبلن كور Dublin Core، وأخيراً معيار MPEG7. ثم تطرق الحديث إلى التحليل الموضوعي للبيانات الصوتية والذي انقسم إلى تحليل نصي، وتقني، ونصي تقني. وانتهى الفصل بالحديث عن استرجاع ملفات الصوت الرقمية بشقيه البحث الحر، والبحث بالتصفح.

**الفصل الثالث:** تناول هذا الفصل دراسة عدد من محركات بحث الصوت من خلال التعريف بها، والتعرف على مدى توافر عناصر الوصف الببليوجرافي داخل تسجيلاتها، ومستوى الكشف وطرقه المتبعة في تحليل محتوى المعلومات الصوتية. ثم تناول واجهات هذه المحركات وطرق عرض النتائج داخلها، فضلاً عن سماتها وخدماتها وانتهاءً بآليات البحث والاسترجاع.

**الفصل الرابع:** يتضمن هذا الفصل المواصفات المقترحة لمحرك بحث الصوت الداعم للغة العربية، من خلال عدد من العناصر تجلت للباحث بعد الانتهاء من الفصل الثالث. تتمثل في كشف ملفات الصوت الرقمية والذي يتضمن المعالجة الفنية، واستخلاص الكلمات المفتاحية من محتوى ملف الصوت، مروراً بمراحل الكشف باستخدام اللغة العربية وتفادي مشكلاتها؛ وطريقة عمل المحرك، وانتهاءً بمحرك الاستفسار وأساليب البحث المقترحة، ثم ترتيب النتائج وعرضها.

**وانتهت الدراسة بمجموعة من النتائج كان من أهمها:-**

1. عدم وجود محرك بحث عربي أو معرب يدعم بحث ملفات الصوت الرقمية وفق متطلبات اللغة العربية.
2. ضعف الإنتاج الفكري العربي حول المتطلبات الفنية والوظيفية لإنشاء محركات بحث تدعم اللغة العربية نظرياً وتطبيقياً.
3. عدم توافر معايير موحدة لإصدار ملفات الصوت في شكلها الرقمي، مع تباين عناصر الميئاتاداة المصاحبة لملفات الصوت الرقمية من صيغة لأخرى.

وقد اقترح الباحث مجموعة من التوصيات كان من أهمها:-

توصيات تتعلق باستخدام اللغة العربية:

١. تطوير الدراسات اللغوية المرتبطة باستخدام الحاسب الآلي نظرياً وعملياً.
٢. دعم استراتيجيات ومعايير صناعة المحتوى العربي الرقمي.
٣. تطوير خوارزميات البحث باللغة العربية بما يحقق كفاءة وسرعة الاسترجاع.

توصيات تتعلق بمحركات البحث العربية:

١. الاهتمام بإنشاء محركات بحث عربية أو حتى معربة تفي بمتطلبات المستخدم العربي سواء بالنسبة للمحتوى العربي أو الأجنبي، فضلاً عن كسر الحاجز اللغوي في البحث عن المحتوى العربي المتناثر على الويب.
٢. إنشاء مجموعات بحثية كبيرة من المتخصصين في المجالين اللغوي والتقني، لتكتمل منظومة عمل قادرة على بناء محرك بحث عربي.
٣. أهمية احتواء محركات البحث العربية على المعاجم اللغوية والمعالجات

الصرفية والمحلات النحوية والمصححات الإملائية.

٤. الاهتمام بوضع معايير إنشاء محركات البحث العربية برمجياً ولغوياً.

٥. دعم محاولات إنشاء محركات البحث العربية المتخصصة موضوعياً، ونوعياً لاسيما محركات بحث الوسائط المتعددة.

توصيات تتعلق بالإنتاج الفكري العربي:-

١. ضرورة دعم الدراسات العربية الأكاديمية حول الاستخدام الآلي للغة العربية، ومتطلبات إنشاء محركات البحث العربية فنياً ووظيفياً.
  ٢. الاهتمام ببحث ودراسة النواحي الفنية والوظيفية للملفات الصوتية، خاصة ذات المحتوى العربي، وإمكانية تطويرها لتتوافق وبيئة الاستخدام العربية.
- توصيات تتعلق بملفات الصوت:-

١. توحيد معايير إنتاج ملفات الصوت الرقمية، وبياناتها المصاحبة لضمان سهولة التنظيم والبحث والاسترجاع.

٢. الاهتمام بالتسجيلات الصوتية العربية كمصدر لا غنى عنه من مصادر

- المعلومات، وتحويل المتوفر منها في الشكل التقليدي إلى الشكل الرقمي. والمعلومات:-
٣. دعم استخدام ملفات الصوت الرقمية في المكتبات ومرافق المعلومات، نظراً لما تلعبه من دور كبير بالنسبة للمعاقين بصرياً.
٤. دعم تجارب إنشاء محرك بحث صوتي عربي يدعم اللغة العربية ويقدم خدماته للمستفيدين بصفة عامة، والمعاقين بصرياً بشكل خاص.
٥. الاهتمام بدعم آليات بحث اللغة العربية داخل محركات البحث الأجنبية، لاسيما ما يدعم بحث ملفات الصوت الرقمية موضوع الدراسة.
١. الاهتمام بتدريس محركات البحث بصفة عامة، ومحاولات إنشاء المحركات العربية بصفة خاصة داخل أقسام المكتبات والمعلومات.
٢. ضرورة تدريس مشكلات الاستخدام الآلي للغة العربية، نظرياً وعملياً بأقسام المكتبات والمعلومات، لتوفير الخلفية الكافية لدى أخصائي المعلومات حول تلك المشكلات وسبل معالجتها.