The North African Journal of Scientific Publishing (NAJSP)
مجلة شمال إفريقيا للنشر العلمي (NAJSP)
E-ISSN: 2959-4820
Volume 2, Issue 4, October - December 2024
Page No: 242-248
Website: https://najsp.com/index.php/home/index
SJIFactor 2024: 5.49 معامل التأثير العربي (AIF) 2024: 0.71 ISI 2024: 0.696

دراسة مقارنة بين محرك الكلمات الرئيسية ومحرك البحث الدلالي

حنان فرج نصيب [1*]، أسماء سالم اعبيد [2]
[1] قسم الحاسوب، كلية الآداب والعلوم قمينس، جامعة بنغازي، بنغازي، ليبيا

Comparative Studying Between the Main Search Engines and The Semantic Search Engines

Hanan faraj Nassib [1*], Asma Salem Abeid [1]
[1] Department of Computer Science, Faculty of Science and Arts Qaminis, University of Benghazi, Benghazi, Libya

*Corresponding author Hanan.fr84@gmail.com *المؤلف المراسل
تاريخ النشر: 2024-12-22 تاريخ القبول: 2024-11-30 تاريخ الاستلام: 2023-09-14

الملخص
نظرا للدور المهم الذي تقوم به استراتيجية تكنولوجيا المعلومات في إنشاء خدمة جديدة للبحث في مكتبات البحث الموحدة، باسم خدمة الاستكشاف على نطاق الإنترنت (Search Engines). حيث تلعب محركات البحث دورا حيويا في مساعدة المستخدمين للبحث والمعرفة والحصول على معلومات موثوق بها وذات مصداقية من الإنترنت، ولكن ليس من السهل العثور على المعلومات الدقيقة من البيانات الضخمة المتاحة على الويب، حيث تتمتع تقنية الويب الدلالية بالقدرة على التركيز على البيانات الوصفية بدلا من بناء الجملة مما جعل محركات البحث الدلالية تبحث عن معنى الكلمات الرئيسية وبالتالي يمكن تحقيق دور فعال للأداء في محركات البحث التقليدية من خلال زيادة دقة المعلومات التي يتم إرجاعها بواسطة استعلام البحث. وفي هذا البحث تمت دراسة مسح لمحركات البحث المعتمدة على النحو ومحركات البحث المعتمدة على الدلالات وتم عرض مقارنة شاملة بينهما ومدى استفادة المستخدم من محركات البحث (Searchengines)، والبحث الدلالي (semantic web).

الكلمات المفتاحية: محركات البحث، البحث الدلالي، الانطولوجيا.

Abstract
Given the important role of the IT strategy to create a new service for researching in libraries of the federated search is known as "Search Engines" scale discovery service. Moreover, the research engine plays a vital role in helping users to research, find out ,and get reliable and credible information from the Internet, but not from easy to find accurate information from huge data available on the web, Semantic Web technology has the ability to focus on metadata rather than syntax, which makes Semantic Search Engines search for the meaning of keywords, and thus an effective role in performance can be achieved in traditional Search Engines by increasing the accuracy of the information returned by the Search query. In this research, a survey of Search Engines based on syntax and Search Engines based on Semantics was

studied, and a comprehensive comparison between them and the extent to which the user benefits from Search Engines and Semantic Search (Semantic Web) was presented.

Keywords: Search Engines; Semantic Web; Ontology.

مقدمة

في ظلال تطور الهائل في تكنولوجيا الاتصالات ومدى حاجتنا إلى المعلومات والبيانات في حياتنا العلمية بشكل كبير وكيفية الحصول على هذه المعلومات عن طريق محركات البحث المختلفة حيث تصبح البيانات الموجودة على الويب ضخمة جدا مما يجعل من الصعب العثور على المعلومات ذات الصلة. حيث أن الغرض الرئيسي من محركات البحث هو السماح للمستخدم بالبحث واسترجاع مستندات الويب مع استعلامات للحصول على المعلومات التي يريدونها، ويعتبر الفرق الرئيسي بين محركات بحث الكلمات الرئيسية (Search Engine) ومحركات البحث الدلالي (Semantic Web) هو استهلاك الوقت حيث يستغرق المستخدم وقت لفرز النتيجة المستردة في محركات البحث الرئيسية ، حيث يمكن للمستخدم الحصول على النتائج ذات الصلة في وقت قصير ،وهي من الخصائص المفيدة جد أو الأكثر أهمية لنظام الويب الدلالي حيث أن محركات البحث القائم على الدلالات لا يوجد فرز لنتائج البحث، وهو أمر مفيد للغايه وهذا أهم ما يميز نظام البحث الدلالي (Semantic Web) .

الدراسات السابقة:

قدم الباحثون في هذا الموضوع دراسات سابقة نذكرها على النحو التالي:

1. المقصود بالويب الدلالي، أو مايعرف بالويب ذو الدلالة اللفظية، أو الويب ذو المعنى(SemanticWeb)، والمقصود بذلك هو اعتمادها على برمجيات تمكنها من تعريف المقصود بالبيانات التي تقدمها أوتمثل (Ontology) الشبكة العنكبوتية(الويب)، عن طريق الاستعانة بمايعرف بخرائط المفاهيم الويب الدلالي وهو أحد فروع الذكاء الاصطناعي ويعد ثورة في عالم الويب حيث يسمح للمتصفح أو البرامج الوكيلة بالبحث عن المعلومات، وبالتالي فإن عملية معالجة المعلومات تتم اعتمادًاعلى الحاسبات الآلية بدلا من البشر [1].
2. فقد عرفت (الخليفة) الويب الدلالية أو مايطلق عليها أحياناً (الويب ذات الدلالات اللفظية) أو الويب (ذات المعنى) على أنها شبكة بيانات بالمعنى، أي أنه يمكن البرامج الحاسوبية الخاصة أن تعرف ماذا تعني هذه البيانات [2].
3. إن محركات البحث (Search Engine) هي أدوات شائعة الاستخدام إلى الوصول إلى صفحات النسيج الشبكي العالمي(World Wide Web) لكنها تعود بمئات النتائج والمعلومات التي تكون في الغالب ليست ذات صلة بموضوع البحث الرئيسي مما يجعل الباحث في حيرة ويأس من وجود ما يبحث عنه، فإن استخدام أدوات البحث المتوفرة على الإنترنت بدون استراتيجية بحث محددة تشبه إلى حد كبير من يسير في مكتبة ضخمة بشكل عشوائي يحاول العثور على كتاب معين [3].
4. تعد محركات البحث التقليدية (Search Engine) مفيد جداً في العثور على معلومات حول الموقع والحصول على النتائج خلال فترة زمنية معينة؛ ولكنهم يعانون من عدم معرفته معنى المصطلحات والتعبيرات المستخدمة في صفحات الويب والعلاقة بينها، وفي بعض الأحيان يكون لكلمة واحدة عدة معاني والعديد من الكلمات لها نفس المعني وفي هذه الحالة إذا أراد المستخدم البحث عن كلمة معينة فقد يؤدي إلى حدوث ارتباك ولن يحصل المستخدم على ما يريد الحصول عليه والبحث عنه [4].
5. اعتبر العباسي (الويب الدلالي) إمتداد للشبكة الحالية بحيث تكون للمعلومات معنى محدد، وهذا سيمكن أجهزة الحاسب والبشر بالعمل في تعاون أفضل [5].
6. تعتبر محركات البحث الحديثة مذهلة جدا، حيث تتيح الخوارزميات المعقدة لمحركات البحث أخذ استعلام البحث الخاص بك وإسترجاع نتائج دقيقة حيث قدم معلومات قيمة وسط منجم واسع النطاق من المعلومات والبيانات [6].

المنهجية

تم استخدام المنهج الوصفي التحليلي والذي يؤدي دوراً اكبيراً في شرح مختلف الظواهر كما أنه يتنبأ بالمستقبل، ويفسح المجال أمام الباحثين للدراسة في كافة المجالات، كما يمكن من خلاله استخراج العلاقات بين الظواهر ومن ثم يقوم بتوضيحها.

أهمية الدراسة

تتلخص أهمية الدراسة إلى التعرف على تقنيات محركات البحث (Search Engine) ومحركات البحث الدلالي (Semantic Search Engine)، وكذلك عرض دراسة مقارنة بينهما.

طرق البحث

أنجز هذا البحث بالإعتماد على الدراسات العلمية السابقة في مجال الإنترنت ومحركات البحث المختلفة، حيث تمت دراسة أهم الدراسات والأقرب إلى الموضوع دراسة دقيقة وأخذ التوصيات الموجودة بها بعين الإعتبار.

محتوى البحث:

محركات البحث الكلمات الرئيسية: Keyword Search Engines

تعد محركات البحث التقليدية أدوات شائعة الاستخدام ومفيدة جداً في العثور على المعلومات من الإنترنت والحصول على النتائج خلال فترة زمنية معينة، ولاكن يواجهون مشكلة في عدم معرفتهم معنى المصطلحات والتعبيرات المستخدمة في صفحات الويب والعلاقة بينها، وفي بعض الأحيان يكون للكلمة الواحدة عدة معاني والعديد من الكلمات لها نفس المعني وذلك بسبب زيادة حجم الروابط على صفحات الويب، وفي هذة الحالة إذا أراد المستخدم البحث عن كلمة معينة فقد يؤدي إلى حدوث ارتباك ولن يحصل المستخدم على مايريد البحث عنه.

عيوب محركات البحث الكلمات الرئيسية:Disadvantages Of Keyword Search Engines

- تعتمد تقنية استرجاع المعلومات التقليدية على ظهور الكلمات في المستندات، لذلك فمن الصعب الحصول على نتيجة ذات الصلة.
- الدقة والاستدعاء منخفضة، الدقة هي نسبة المستندات المستردة ذات الصلة للمستخدم للحصول على المعلومات التي يحتاجها، والاستدعاء هو نسبة المستندات ذات الصلة بالاستعلام التي يمكن استرجاعها بسهولة.
- كلمات متعددة المعاني تعني الحصول على عدة معان لكلمة واحدة.
- كلمات مترادفة تعني عدة كلمات لها نفس المعنى.
- ليست كل المعلومات التي يتم فهرستها بواسطة محركات البحث دقيقة أو جديرة بالثقة، مما يؤدى إلى انتشار محتمل للمعلومات المضللة.
- الإفراط في المعلومات حيث يمكن أن تكون وفرة المعلومات هائلة مما يجعل من الصعب على المستخدمين التمييز بين المصادر الموثوقة والمصادر غير الموثوقة.

محركات البحث الدلالية:Semantic Search Engines

محركات البحث الدلالية هي المحركات الذكية التي تبحث عن الكلمات الرئيسية حسب المعنى، ويقدم نتائج دقيقة وذات صلة للاستعلامات المطلوبة بالإضافة إلى أنها تضمن النتائج المتعلقة بمعنى الكلمات الرئيسية التي تم البحث عنها، حيث تستخدم محركات البحث الدلالية مفهوم الانطولوجي حتى تحقق عمليات استرجاع المعلومات ذات المغزى وتحصل على نتيجة عالية الدقة. تعتبر محركات الويب الدلالية بمثابة (Web 3.0) أو امتدادا للويب الحالي والتي تمثل المعلومات من أجل ربط المعلومات في الويب كملفات (HTML وOWL و RDF)

ومن أجل الوصول إلى البيانات المنظمة تم تقديم عدد من محركات البحث الدلالية التي تفهم معنى البيانات وتساعد في عرض نتائج أكثر دقة مقارنة بمحركات البحث الحالية مثل (Hakia, DuckDuckGo Swoogle) [7] .

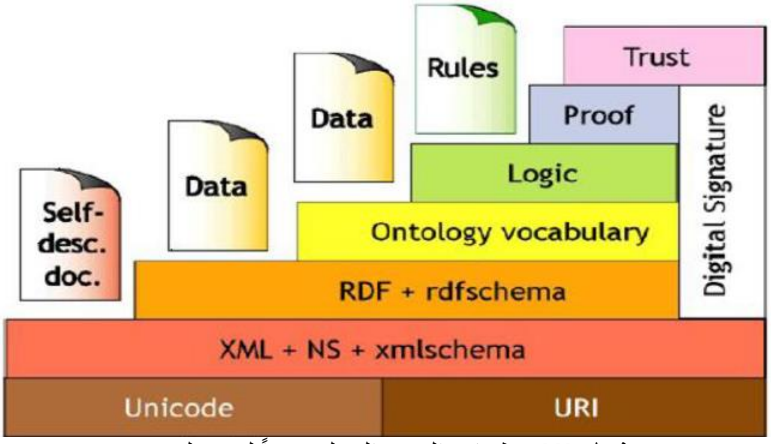
استخدامات الويب الدلالي:

يستخدم الويب الدلالي في تنسيق البيانات في شكل منظم يساعد الآلات على فهم المزيد من المعلومات على الويب، كما يستخدم في العديد من التطبيقات التي لا يمكن تعدادها مثل مجال المكتبات الرقمية،

بالإضافة إلى المجالات التعليمية والطبية والتجارية واللغوية والعديد من المجالات الأخرى، حيث يساعد على اكتشاف أكثر توافق وتكاملا للبيانات من مصادر مختلفة والحصول على نتائج أكثر دقة للمستخدم مقارنه بمحركات البحث على الويب الحالية.

البنية الهيكلية لأنطولوجيا الويب الدلالي:

تعتبر الانطولوجيا أحد أبرز مكونات الويب الدلالي وبناء عليه من الضروري التعرف على التصور العامل بنيته المعمارية والتي تتألف من طبقات وفئات تعكس التسلسل لا لهرمي، وتستغل كلا منها قدرات الطبقات التالية لها. وقد اقترح "بيرنرزلي "Berners-Lee" في عام (2001م). البنية الهيكلية العامة لطبقات الشبكة الدلالية تحت مسمى "Semantic Web Layer Cake" أو " Semantic Web Stack" وذلك على النحو المبين في الشكل رقم (1) [8-10].



شكل (1): طبقات الويب الدلالي وفقاً لبيرنزلي.

وتستخدم هذة الطبقات في هيكل محرك البحث الدلالي والتي يتم وصفة بالتالي:

- نظام الترميز الموحد وروابط المصادر Unicode and URI

يعرف بأنه المستوي الأساسي لتقنية الويب الدلالي التي تستخدم لتحديد موقع المصدر. كما أنه يساعد الويب الدلالي تمثيل المستندات وربطها باللغات البشرية وإحدى معايير ترميز وتمثيل المحتوى لمعظم أنظمة الكتابة في العالم، بالإضافة إلى إمكانية توحيد الرموز والمعادلات الرياضية والعلمية. ويعطي نظام الترميز الموحد الفرصة أمام التمثيل بطريقة صحيحة لمختلف أنواع المصادر التي تحول إلى صورة مقروءة آلياً عبر عملية الرقمنة أو التحويل الرقمي. حيث تضمن روابط المصادر كلاً من "معرفة المصدر الموحد"Universal Resource Identifier" (URI) أو كما يطلق عليه البعض "المعرف العام للمصدر" وكذلك "معرف المصدر دولياً"(IRI)"Internationalized Resource Identifier " وتكمن الفكرة في تعيين عنوان فريد يستخدم لتحديد موقع مصدر المعلومات - أيا كان نوعه - بدقة على الشبكة العالمية.

- لغة التمييز القابلة للامتداد: Extensible Markup Language (XML)

تعتبر بمثابة المعيار وإحدى توصيات تكتل الشبكة العنكبوتية العالمية، وهو نظام لهيكلية المصادر يحدد شكل قياسي متفق عليه لعرض البيانات وإتاحة تمثيل بنية المحتوى.

تستخدم هذه اللغة على نطاق واسع في الويب لعدة أسباب مثل تنسيق النص البسيط والمرن وتحديد البنية المستخدمة لوصف البيانات، كما يلبي تحديات الأعمال الإلكترونية والنشر الإلكتروني، بالإضافة إلى ذلك له دور مهم جداً للتبادل بين أنواع البيانات المختلفة على الويب.

- إطار وصف المصدر Resource Description Framework (RDF)

هي لغة بسيطة للتعبير عن نماذج البيانات المترابطة على الويب، وهيكل قياسي لوصف العناصر الرئيسية لمصادر المعلومات الرقمية من العنوان والمؤلف والمحتوى وإدارة المعلومات المتعلقة بحقوق النشر وغيرها، وذلك بصورة تكون قابلة للمعالجة والتفسير من جانب أنظمة المعلومات، ويساعد ذلك على

تحقيق التوافق التركيبي المتبادل ووصف دلالات المحتوى ويمكن التعبير عن هذا النموذج من خلال قواعد (XML) نفسها مما يتيح تبادل البيانات ودمجها من مصادر متعددة.

- لغة انطولوجيا الويب Ontology Vocabulary

يتم استخدام (Ontology) لوصف بيانات الويب الدلالي ولغة موجهه لمعالجة المحتوى الرقمي بصورة دلالية والتسمية الكاملة لها "Ontology Web Language" وتبني أعلى أو على رأس ملف "إطار وصف المصدر". وتستخدم في تحسين الاداء لإعطاء طريقة موحدة لتسهيل العلاقات بين الأجزاء المختلفة، حيث أن الانطولوجيا هي الطريقة التي توفر قواعد ومفردات مشتركة للبيانات المتاحة على الشبكة.

- التوقيع الالكتروني والتشفير Digital Signature

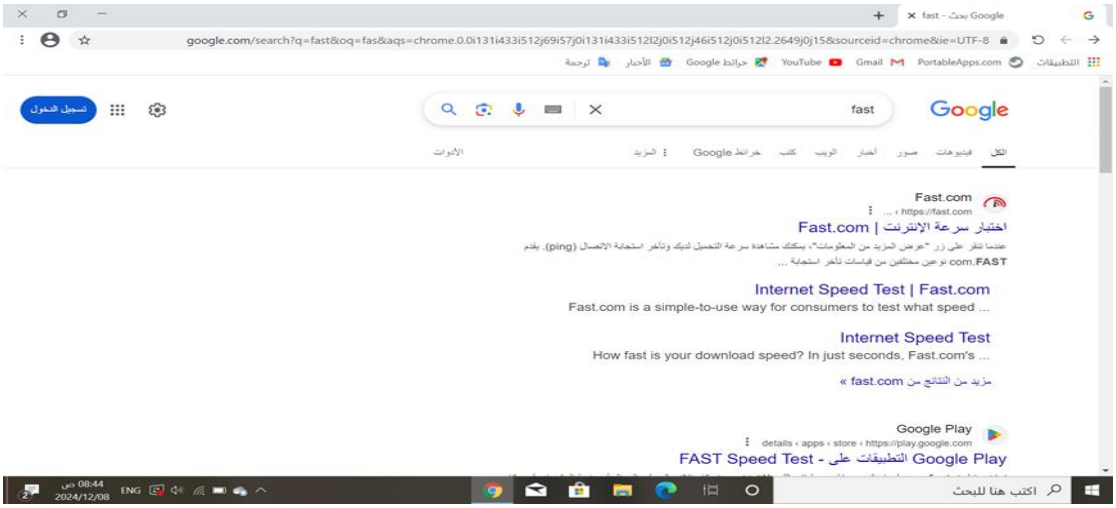
عرفت المنظمة الدولية للتوحيد القياسي (ISO)(International Organization for Standardization) التوقيع الالكتروني بأنه "تحويل مشفر لوحدة البيانات بحيث يسمح بتحديد مصدر المحتوى ومدى استقامة ووحدة البيانات من خلال حمايته من أ تزوير أو تزييف".

- المنطقية والإثبات والمصداقية Logic, Proof and Trust

هي الطبقة الأخيرة من طبقات الويب الدلالي التي تتبع طبقة الانطولوجي، وتعتبر المنطقية طبقة توفر دلالات لكافة البيانات المتاحة وتعمل على تعظيم قدرة التطبيقات على الإستفادة من القواعد والضوابط اللغوية والدلالية في بيئة محتوى ومصادر الويب الدلالي. ويمكن النظر إلى المنطقة باعتبارها أساس تمثيل المعرفة.

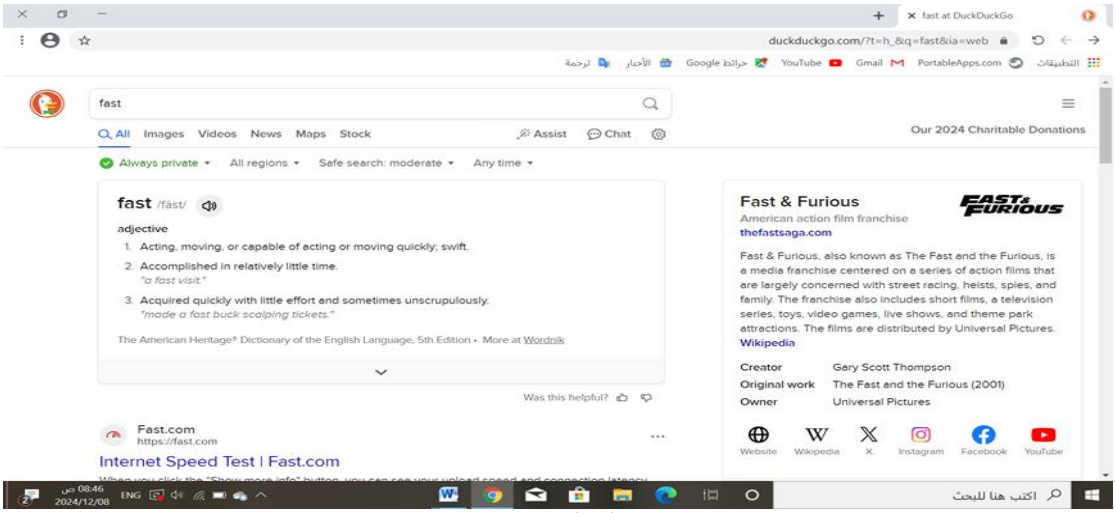
المقارنة

لمقارنة محرك الكلمات الرئيسية ومحرك البحث القائم على الدلالات تتم كتابة استعلام بحث مماثل في كل من محركات البحث مثل Google كمحرك بحث للكلمات الرئيسية، DuckDuckGO كمحركات دلالية هنا يتم استخدام الكلمات متعددة المعاني للاستعلام عن كلمة "Fast".



شكل (2): محرك البحث Google.

في الشكل أعلاه يوفر محرك بحث Google نتيجة بحث أساسية لمرادفات الكلمة "Fast". وليقدم نتائج دقيقة، حيث لم يتم ذكر معني أخر وهو أحد عيوب محركات البحث عن الكلمات الرئيسية.



الشكل (3): محرك البحثDuckDuckGo.

في الشكل أعلاه باستخدام محرك البحث الدلالي يتم استرداد النتائج الأكثر صلة ودقة لاستعلام معين. ومن ثم يتضح من الشكل 2 والشكل 3 أن Google غير قادر علي التعامل مع الكلمات متعددة المعاني، بينما يقدم محرك البحث الدلالي DuckDuckGo النتائج بكل المعاني الممكنة لكلمة "Fast".

جدول (1): يوضح المقارنة بين محرك الكلمات الرئيسية ومحرك البحث الدلالى.

محركات البحث الرئيسية	محركات البحث الدلالي
▪ تعتبر من المحركات التقليدية التي تقوم بإرجاع النتائج حسب سياق الإستعلامات.	▪ تستخدم أسلوب الوصول للمعلومات الدقيقة وذات صلة لاستفسارات المستخدمين.
▪ لايعتمد على كلمات التوقف وعلامات الترقيم والنتائج غير دقيقة	▪ يعتمد على كلمات التوقف وعلامات الترقيم، حيث تؤثر هذة العلامات على نتائج البحث.
▪ لغات HTML,XML هي اللغات الأساسية المستخدمة لإنشاء صفحات الويب.	▪ لغات OWL, RDF هي اللغات الأساسية المستخدمة لإنشاء صفحات الويب.
▪ يتم البحث بالضبط اعتمادا على الكلمات الموجودة في موقع الويب والتي يحددها الباحث.	▪ يسعي إلى توفير الدقة في المعلومات التي يتم إرجاعها من خلال فهم معني الكلمات الرئيسية المتعلقة بما يرغب به الباحث
▪ يحاول توسيع الإستعلام باستخدام الكلمات الأساسية بدلا من إستخدام المنهجية.	▪ يحاول الوصول إلى العلاقات بين الكلمات الرئيسية باستخدام علم الإنطولوجيا.

يوضح الجدول أعلاه الفرق الأساسي بين الكلمات الرئيسية ومحرك البحث القائم على الدلالات وهو مفيد لدراسة أهمية البحث القائم على الدلالات على محرك البحث عن الكلمات الرئيسية. ومن خلال مقارنة محركي البحث، أتضح أن محركات البحث الدلالية تتمتع بمزايا أكثر من محرك البحث عن الكلمات الرئيسية.

الخلاصة

يوفر الويب الحالي طريقة سهلة لمشاركة المعلومات عبر الإنترنت، مما يجعل حجم البيانات على الويب ضخما تدريجيا، حيث أن محركات البحث ذات الأساس الدلالي تتمتع بمزايا أكثر من محركات البحث الرئيسية من حيث دقة الحصول على النتائج حيث تساعد المستخدمين في العثور على المعلومات على الويب بسرعة وسهولة ودقة، وهناك نوعان من محركات البحث: محركات البحث عن الكلمات الرئيسية ومحركات الويب الدلالية.

محركات الكلمات الرئيسية التي تعتبرها محركات البحث هي محركات البحث الأساسية للويب لكن لا يمكنهم العثور معلومات دقيقة لاستعلامات المستخدم لأنها تعتمد على بناء جملة الكلمات الرئيسية، حيث

تعمل محركات البحث الدلالية على حل المشكلة من خلال النظر في معني الكلمات الرئيسية واسترداد النتائج ذات الصلة دلاليا ويعتمد محرك البحث الدلالي على تقنية الويب الدلالي التي تساعدهم على فهم المفهوم وتساعد الآلات على فهم المعلومات ومعالجتها. في هذا البحث تم تقديم لمحة عامة عن تكنولوجيا الويب، وعرض مقارنة شاملة لأثنين من محركات البحث الأكثر شيوعا.

المراجع:

[1] P. Amrit and W. Tim, “Navigating the Post-API dilemma | Search Engine Results Pages present a biased view of social media data,” arXiv [cs.IR], 2024.

[2] Pranave, S. Shree, V. K. Reddy Challa, and N. Panda, “System search service implementation based on a custom lexical search,” Procedia Comput. Sci., vol. 235, pp. 1548–1557, 2024.

[3] M. Khaleel, A. Jebrel, and D. M. Shwehdy, “Artificial intelligence in computer science,” Int. J. Electr. Eng. and Sustain., vol. 2, no. 2, pp. 01–21, 2024.

[4] T. Sultana, A. K. Mandal, H. Saha, M. N. Sultan, and M. D. Hossain, “Intent identification by semantically analyzing the search query,” Modelling, vol. 5, no. 1, pp. 292–314, 2024.

[5] Y. S. Negi and S. Kumar, “A comparative analysis of keyword- and semantic-based search engines,” in Intelligent Computing, Networking, and Informatics, New Delhi: Springer India, 2014, pp. 727–736.

[6] A. C. S. Sheela and C. Jayakumar, “Comparative study of syntactic search engine and semantic search engine: A survey,” in 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 2019.

[7] M. Khaleel, Y. Nassar, and H. J. El-Khozondar, “Towards utilizing Artificial Intelligence in scientific writing,” Int. J. Electr. Eng. and Sustain., pp. 45–50, 2024.

[8] K. Vayadande, P. A. Bailke, V. J. Nandeshwar, R. Kumar, and V. R. Dange, “A comparative study of different techniques of text-to-SQL query converter,” How Machine Learning is Innovating Today’s World. Wiley, pp. 353–365, 17-Jun-2024.

[9] H. Xiong et al., “When search engine services meet large language models: Visions and challenges,” IEEE Trans. Serv. Comput., vol. 17, no. 6, pp. 4558–4577, 2024.

[10] A. Saeed, N. Dhanda, A. S. Rao, and R. Verma, “AI-Enabled Semantic Web,” in 2024 2nd International Conference on Disruptive Technologies (ICDT), 2024, pp. 1136–1141.