الجمهورية العربية السورية
جامعة البعث
كلية الهندسة المعلوماتية
قسم هندسة البرمجيات ونظم المعلومات

تقنيات الويب الدلالي و البحث في الويب

دراسة أعدّت لنيل درجة الماجستير في الهندسة المعلوماتية باختصاص هندسة البرمجيات و نظم المعلومات

إعداد

م. أسماء عبد الرحمن شعار

إشراف

د. رانيا لطفي

أستاذ مساعد في قسم هندسة البرمجيات ونظم المعلومات ـ جامعة البعث

٢٠١٥م-١٤٣٦هـ

ملخص

تناولنا في هذا البحث دراسة إمكانية تحسين أداء استرداد المستندات بمساعدة تقنيات الويب الدلالي على أن يتم الاحتفاظ ببنية الفهرس التقليدي وبدون خسارة فوائد نماذج الاسترداد التقليدية التي تتميز بسرعتها لكنها غير دقيقة لأنها لا تستغل المعنى، وذلك بهدف دمج فوائد البحث التقليدي مع تحسين دقة البحث بالاعتماد على الدلالة الموصّفة ضمن أنطولوجي تخص مجال معين. وعلى وجه التحديد حاولنا إثبات ما يلي:

" إن فهرسة النصوص ضمن الإعلانات (المصنفة وفقاً لأحداث محددة) بالإضافة إلى الوسوم الدلالية التي تمثل ثلاثيات مستخرجة وثلاثيات مستنتجة بالاستدلال من شأنه أن يحسن بشكل هام من أداء الاسترداد."

ومن أجل ذلك قمنا بتصميم وبناء نموذج أولي لتطبيق ويب دلالي كان بمثابة محرك بحث متخصص في مجال معين قادر على التعامل مع المعنى لمعلومات البحث من خلال نمذجة مناسبة لمعرفة المجال ويتضمن كل مفاهيم وتقنيات الويب الدلالي من استخراج المعلومات إلى استرداد المعلومات بالإضافة إلى الاستدلال، ومن ثم إضافة نكهة دلالية في وقت الفهرسة من أجل تعزيز الفهرس بمعلومات مستخرجة ومستنتجة تكون بمثابة بيانات وصفية مرافقة للمستندات، حيث تم الاعتماد على بنية وتقنية الفهرسة التقليدية واستغلال عناصر الأنطولوجي وقت البحث قدر الإمكان.

وقد أثبتت عملية التقييم جدوى هذا النظام بنسبة استرداد تصل إلى (٨٧.٠٨%) بالمقارنة مع البحث التقليدي (٣٩.٢٦%)، وبالتالي تكون البنية المشروحة مناسبة لتحسين دقة البحث التقليدي.

Abstract

In this dissertation, we try to handle the studying of possibility of improving the documents retrieving performance with the help of the Semantic Web technologies, while preserving the structure of traditional index and without losing the benefits of traditional retrieving models that are characterized by quickness but are not accurate because they do not take the advantage of meaning. That was an aim for benefiting from the advantages of traditional search in addition to improved Search accuracy Depending on the semantics described within particular domain ontology. Specifically we tried to prove the following:

"The indexing of the text within announcements (classified according to specific events) in addition to Semantic tags that represent extracted triples and inferred triples would significantly improve the performance of the retrieving."

To prove that purpose, we have designed and built a prototype of Semantic Web application that was a search engine specialist in a particular field and is capable of dealing with the meaning of the search information through appropriate knowledge modeling. This application includes all the Semantic Web concepts and techniques from information extraction to information retrieving in addition to inference. Then adding semantic indexing flavor at indexing time in order to enhance the index with extracted and inferred information serve as metadata accompanying to the documents, where we relied on the traditional index structure and technology and exploitation of the ontological elements at search time as much as possible.

The feasibility evaluation process of this prototype system has proven retrieving up to (87.08%) compared with the traditional search (39.26%), and thus the described structure is suitable to improve the accuracy of the traditional search.