عنوان البحث : Querying DBpedia Using HIVE –QL

ملخص النشر: *Wseas; 14th International Conference on Telecommunications and Informatics (TELE-INFO '15), Sliema, Malta, August 17-19, 2015*

أسماء المشاركين: Ahmed Slama Ismail, Haytham AL-Feel, Hoda Mokhtar Omar

Abstract

DBpedia is considered one of the main data hubs on the web nowadays. DBpedia extracts its content from different Wikipedia editions, which dramatically increase day after day according to the participation of new chapters. It became a big data environment describing 38.3 million things with over than 3 billion facts. This data explosion affects both the efficiency and accuracy of the retrieved data. From this point of view, we propose a new architecture to deal with the DBpedia using big data techniques in addition to the Semantic Web principles and technologies. Our proposed architecture introduces HIVE-QL as a query language for DBpedia instead of the SPARQL Query Language, which is considered the backbone of the semantic web applications. Additionally, this paper presents the implementation and evaluation of this architecture that minimizes the retrieval time for a query in DBpedia.

ملخص البـــــــحث :

تعتبرالموسوعة المنظمة الديبيديا واحدة من أهم مصادر المعلومات فى الوقت الحاضر على شبكة الانترنت. وتستخرج الديبيديا محتوياتها من مختلف الفصول الخاصة بموسوعة الويكيبديا الامر الذى يشكل زيادة كبيرة يوما بعد يوم تبعاً لمشاركة فصول جديدة من الديبيديا. وبذلك اصبحت تمثل واحدة من بيئه البيانات كبيرة الحجم والتى تصف ما يقرب من ٣ مليون شى مع اكثر من ٣٨.٣ بليون من الحقائق. .مما لاشك فيه ان هذه الزيادة االكبيرة فى حجم البيانات يؤثر بشكل ما فى كلا من كفاءة ودقة البيانات التى يتم الاستعلام عنها. لذلك فمن منطلق هذه الرؤيه فإننا نقدم مقترح عمل بتنفيذ هيكل بنائى جديد لكى تتعامل مع الديبيديا بإستخدام تقنيات البيانات كبيرة الحجم بالإضافة الى تقنيات الويب الدلالى. لذلك فان التصميم والتطبيق للهيكل البنائى للبحث يقدم لغة الاستعلام (HiveQl) للاستعلام عن معلومات الديبيديا كبديل عن لغة الاستعلام (SPARQL) والتى تعتبر العمود الفقرى لتطبيقات الويب الدلالى. بالاضافة الى ذلك يقدم هذاالبحث كيفية تصميم و تنفيذ و تقييم هذا الهيكل ،ومن النتائج وباللإختبارات بتضح بالتجربة العملية أن مقترحنا يقلل من وقت استرجاع المعلومات عند الاستعلام من الديبيديا