EC2 AMI LinkedData

Idafen Santana Pérez
OEG-DIA
Facultad de Informática
Universidad Politécnica de Madrid
isantana@fi.upm.es

Nandana Mihindukulasooriya
OEG-DIA
Facultad de Informática
Universidad Politécnica de Madrid
nandana@delicias.dia.fi.upm.es

Boris Villazón-Terrazas
OEG-DIA
Facultad de Informática
Universidad Politécnica de Madrid
bvillazon@fi.upm.es

ABSTRACT

We present the process that has been followed for the development of an application that makes use of several heterogeneous datasets that are related to .....

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Miscellaneous; E.2 [**Data storage representations**]: Linked Representations

General Terms

Design, Experimentation

Keywords

linked data, linked open data

1. INTRODUCTION

Linked Data principles are being adopted by an increasing number of data providers, getting as a result a global data space on the Web containing billions of RDF triples [].

Domains library, government, statistics, but no Computer Science related datasets .... why?

2. AMAZON EC2

Today the advancements of cloud computing has made it possible for us to have different environments for computing without having the need to have a dedicated infrastructure for each of the different environments we require. One of the main advantage of virtualized environment is that you can recreated the same environment as many times as you want with a minimal effort. This put the computer users lives at ease because they can use a virtual machines for different temporary tasks without having the trouble to change the configuration of the personal machines. This ability can be very useful in scenarios like reproducing environments for software testing, to try out software before installing them on a physical machine i.e. as a staging machine.

If it is possible to use already prepared virtual machines which matches the requirements, for example, the architecture i.e. it is 32bit or 64bit, the operation system, storage and memory requirements it would save a lot of time and effort of people trying to prepare a computing environment for some given task. In other words, if we could facilitate the search and discovery of virtual machine images in a systematic way, we could build a lot of semantic application about this which will automate this process and make the use of cloud computing for these tasks much easier.

3. AMAZON EC2 LINKED DATA LIFE CYCLE

Here we briefly describe the proces

4. APPLICATIONS

Our goal of this project is to create a LinkedData dataset about Amazon virtual machine images. Thanks to this dataset, we have been able to build the following applications.

4.1 VMI Finder Desktop Client

This desktop client can run on a computing environment and gathers the data about the computing environment that it is executing. Then it will create a SPARQL query based on the information it gathered from the environment and connect to the SPARQL endpoint which exposes the ec2ld. The query result will provide a set of Amazon Machine Image ids which matches the computing environment it runs. Users can manually modify the information that has been automatically collected by editing, adding, or removing any information as they wish.

4.2 VMI Finder Webapp

Users can provide a set of properties and the webapp will create SPARQL query based on that information to find the AMIs fulfills the characteristics defined by the aforementioned parameters. Advanced used can directly enter SPARQL queries using the EC2LD ontology.

4.3 VMI Agent Desktop client

This client can run inside a virtual machine and collect all the information about a virtual machine and given the virtual machine, the agent can publish those data so they will be used in the future searches of the virtual machine images. This will enrich the dataset with a lot additional

I-SEMANTICS Triplification Challenge 2012 Graz, Austria

information that are not normally available using only the Amazon API.

4.4 VMI Finder Webpage widget

Once configured with certain properties, this widget can communicate with the SPARQL endpoint and dynamically fetch the amazon virtual machine images which match the widget configuration. This can be a good addition to the software providers i.e. they can include this widget in the download pages so that users know in which public virtual machine images they can try this software.

5. CONCLUSIONS AND FUTURE WORK

In this paper we have presented an application that makes use of several ...

6. ACKNOWLEDGMENTS

This work has been supported by the R&D project Seals. We would like to kindly thanks

7. REFERENCES