

Avaliação de Tecnologias de Bancos de Dados Semânticos para a Construção de um Sistema Inteligente de Gestão de Conteúdos de Aprendizagem: Experimentos e Resultados

João Carlos Gluz¹, Fernando Machado¹, Matheus Campezzato Galão¹

¹Pós-Graduação em Computação Aplicada (PPGCA)
Universidade do Vale do Rio dos Sinos (UNISINOS)
Av. Unisinos 950 – 93022-000 – São Leopoldo – RS – Brazil

jcgluz@unisinos.br, fernando22sl@gmail.com, matheusz.cg@gmail.com

Abstract. *The research presented in this paper aims to serve as a basis for the design and construction of an intelligent educational content management system. However, an important issue not yet resolved is the definition of an appropriate database technology able to handle the semantic of these contents, particularly when represented by learning objects. As a result, this work identifies a set of NOSQL databases, which are appropriate candidates for building the semantic repository, defines a methodology to evaluate the performance of these databases and presents initial results of this evaluation.*

Resumo. *A pesquisa apresentada neste trabalho pretende servir de base para o projeto e construção de um sistema inteligente de gestão de conteúdos educacionais. Porém, uma questão importante ainda não resolvida é a definição de uma tecnologia apropriada de banco de dados que seja capaz de tratar da semântica desses conteúdos, em particular quando representados como objetos de aprendizagem. Como resultados, este trabalho identifica que bancos de dados NOSQL são candidatos apropriados para a construção do repositório semântico, define uma metodologia capaz de avaliar o desempenho deste bancos e apresenta os resultados preliminares dessa avaliação.*

1. Introdução

Ambientes digitais educacionais se tornaram o denominador comum da aplicação das Tecnologias de Informação e Comunicação (TIC) na Educação: é praticamente inconcebível imaginar um processo de ensino e aprendizagem mediado por sistemas digitais, sem o suporte desse tipo de sistema. Tais ambientes são classificados ou identificados por termos como Sistemas de Gestão de Aprendizagem (em inglês, LMS – *Learning Management System*). O objetivo mais geral de um LMS é intermediar digitalmente as diversas atividades e tarefas envolvidas em um processo de ensino, permitindo administrar, planejar, documentar, distribuir material e acompanhar digitalmente as etapas e estágios deste processo.

Entretanto, uma vez que conteúdos ou materiais educacionais são uma peça fundamental de qualquer processo de ensino e de aprendizagem (independente da tecnologia utilizada), um LMS também deveria ter a capacidade de ajudar na preparação e criação (autoria) destes conteúdos, de organizar e integrar tais materiais aos currículos e planos de ensino, de distribuir e disponibilizar os materiais aos alunos, de fornecer apoio aos alunos

para facilitar o uso destes materiais, de acompanhar como os materiais estão sendo usados e de fornecer estas informações aos professores e tutores. Neste contexto, é comum separar os objetivos e funções primordialmente relacionados com a gerência de conteúdos e materiais educacionais, em um sistema independente, ainda que possivelmente integrado ao LMS [Irlbeck e Mowat, 2007; Greenberg, 2002]. Os Sistemas de Gerência de Conteúdos de Aprendizagem, mais conhecidos pela sigla em inglês LCMS (*Learning Content Management System*) são justamente os sistemas especificamente voltados para a gestão de conteúdos educacionais digitais [Irlbeck e Mowat, 2007; Greenberg, 2002].

Além de suportar as atividades relacionadas aos conteúdos educacionais nos processos de aprendizagem, seria importante que os LCMS oferecessem suporte inteligente, ajudando a minimizar o trabalho de professores e alunos na condução das tarefas vinculadas a esses conteúdos. Um LCMS deveria ser capaz de suportar de forma inteligente todas as atividades relacionadas aos conteúdos educacionais possuindo conhecimentos precisos sobre o que são e para que servem tais atividades. Um LCMS que atenda esses requisitos pode ser denominado de LCMS inteligente, ou apenas, ILCMS (do inglês, *Intelligent Learning Content Management System*). Atualmente, existem diversos trabalhos e sistemas propostos na literatura que tentam satisfazer essas questões (ver Seção 2). Entretanto, a complexidade de todos os temas envolvidos, incluindo os diversos padrões existentes a respeito destes sistemas e conteúdos, não permitiram até agora criar um ILCMS inteligente e integrado a um LMS.

Esse é o objetivo geral de pesquisa tecnológica deste trabalho. Os resultados de pesquisa apresentados no trabalho servirão de base para a construção de um ILCMS que possa ser integrado aos LMS baseados em software aberto existentes atualmente. Uma questão importante ainda não resolvida no estágio atual dessa pesquisa, que é o objetivo específico deste trabalho, é a definição da tecnologia de banco de dados semântico capaz não apenas de armazenar conteúdos educacionais, mas também de representar a semântica e inferir os conhecimentos relacionados a estes conteúdos, para facilitar os processos de pesquisa, recuperação, autoria e adaptação de conteúdos. Um banco de dados semântico com essas propriedades é peça chave para a construção de um ILCMS.

Atualmente existe um grande incremento das pesquisas e aplicações de bancos de dados que vão além do modelo relacional. Esses bancos, usualmente denominados de NOSQL (do inglês, *Not Only SQL*) estendem ou, as vezes, substituem o modelo relacional por outros modelos de representação de informação. Como hipótese tecnológica, este trabalho identifica bancos de dados NOSQL como candidatos apropriados para a construção do banco de dados semântico.

Dentre as várias modelos semânticos possíveis com bancos NOSQL, a presente pesquisa está centrada na análise dos bancos capazes de representar ontologias OWL na forma de grafos RDF (também denominados de bancos baseados em Triplas RDF). O objetivo específico da pesquisa é definir e aplicar uma metodologia experimental de análise e avaliação de bancos RDF para testar e avaliar a eficácia e desempenho deste bancos quando usados como repositórios semânticos de conteúdos educacionais. O processo de avaliação se constitui nas fases de: a) seleção dos bancos de dados semânticos, b) especificação de metodologia experimental de avaliação bancos semânticos para fins educacionais e c) aplicação da metodologia experimental aos bancos de dados selecionados.

A escolha dos bancos foi baseada nos critérios de: licenciamento livre, popularidade entre os usuários, suporte ao armazenamento nativo de dados em triplas RDF, suporte completo à SPARQL e suporte a inferência OWL. Inicialmente foram selecionados os bancos de dados semânticos RDF4J e Apache Jena, que são bancos de dados de código fonte aberto e exclusivamente RDF com maior ranking nessa classe de bancos, tal como indicado na página <http://db-engines.com/en/ranking/rdf+store> especializada na comparação de tecnologias de bancos de dados. Os experimentos de análise de desempenhos são baseados no Banco de OAs BIOE mantido pelo MEC.

2. Fundamentos Tecnológicos

2.1. Web Semântica

A web semântica, também chamada de Web 3.0, é uma evolução da web tradicional que tem por objetivo definir uma web interpretável por agentes inteligentes, onde é possível relacionar conteúdos (dados) e criar aplicações web sofisticadas indo em direção a inteligência artificial, incluindo, também, conteúdos gerados por máquinas. Os padrões e definições para a web semântica são regidas pela W3C (*World Wide Web Consortium*). O grande benefício da web semântica é que ao mesmo tempo que seres humanos podem navegar em documentos de página web convencionais, agentes inteligentes da web semântica podem processar as anotações legíveis por máquina podendo classificar entidades de dados, descobrir ligações lógicas entre entidades, construir índices e criar páginas de navegação e pesquisa [Sikos, 2015].

2.2. RDF

A RDF (*Resource Description Framework*) é um modelo para descrever recursos disponíveis na web semântica recomendado pela W3C. Através de um arquivo RDF é possível fazer afirmações e definir relações sobre esses recursos. Um recurso pode ser qualquer item (vídeo, texto, música, imagem, entre outros) e deve ser identificado por uma URI (*Uniform Resource Identifier*). As afirmações e relações sobre esses recursos são guardadas na forma de tripla, que é composta por sujeito, predicado e objeto (ver Figura 1).

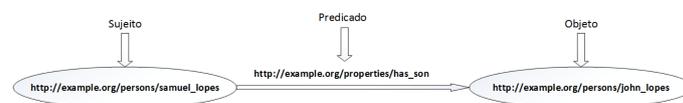


Figura 1. Exemplo de uma tripla RDF

2.3. Ontologias e OWL

Uma ontologia pode ser definida como “uma representação formal do conhecimento por um conjunto de conceitos dentro de um domínio e as relações entre esses conceitos”. É um modelo descrevendo as classes de objetos em um domínio, as relações que possam existir entre os objetos, e as restrições na maneira as propriedades podem ser combinadas. Atualmente os sistemas baseados em ontologias usam a linguagem OWL para a descrição formal das ontologias. Os dados (indivíduos) dessas ontologias são tipicamente armazenados através de triplas definidas na linguagem RDF e acessadas através de consultas da linguagem SPARQL.

2.4. SPARQL

SPARQL (*Simple Protocol and RDF Query Language*) é uma linguagem que permite consultar dados de arquivos RDF, similar à SQL, que é uma recomendação da W3C desde de 2008. Na sintaxe do SPARQL, as variáveis são identificadas pela inicial ? ou por \$ e literais são delimitados tanto por aspas simples ou duplas.

2.5. Bancos de Dados Semânticos

Os bancos de dados NOSQL podem estender ou substituir o modelo relacional, usado nos bancos de dados tradicionais, de diversas formas. Porém no presente trabalho estamos interessados em *bancos de dados semânticos* que sejam capazes de suportar os requisitos da Web-semântica (ver Subseção 2.1) de forma eficiente e com bom desempenho. Consideramos como principal critério que identifica um banco NOSQL como um banco de dados semântico o suporte nativo a representação e armazenamento de dados na forma de triplas RDF. Este é um requisito necessário mas não suficiente. Além disso também é necessário que o banco de dados seja capaz de suportar de forma nativa ou através de *plugin* a inferência baseada nas ontologias OWL.

3. Estado da Arte

A aplicação de tecnologias relacionadas a Web Semântica no projeto de ambientes e sistemas educacionais é um fenômeno atual [Dicheva, 2009; Bittencourt, 2009; Isotani e Mizoguchi, 2008; Silva, 2009]. Essas tecnologias também vem sendo consideradas para tratar questões relacionadas a gestão de conteúdos educacionais, em particular aqueles representados por Objetos de Aprendizagem (OA), desde o início dos anos 2000 [Stojanovic, 2007; Mohan, 2003]. Resultados efetivos começaram a aparecer com os projetos ROSA [Porto, 2007], SLOR [Carrion, 2007], SLCMS [Balog-Crisan, 2008] e a arquitetura baseada em serviços web proposta em [Cobo, 2007]. Porém a rápida evolução das tecnologias da web semântica e dos padrões de metadados para OA, fizeram essas propostas iniciais perder força. Propostas recentes como o projeto OBAA-MILOS [Gluz e Vicari, 2012], OILCMS [Kholief, 2012] e o modelo proposto em [Pahl, 2013] se alinham em termos de tecnologias base, todos prevendo o uso efetivo de ontologias OWL para representação de informações sobre OA, suporte de bancos de dados semânticos RDF/SPARQL como base de repositórios semânticos, além da aderência a padrões de metadados para OA como IEEE-LOM, Dublin Core ou OBAA [Gluz e Vicari, 2011]. Trabalhos mais recentes também buscam enriquecer o repositório semântico buscando e trazendo informações disponíveis na web através de aplicação de tecnologias analíticas (*analytics*) sobre dados ligados (do inglês, *linked data*) [Lama, 2011] ou pelo uso de anotações ontológicas relacionadas as redes sociais [Fernández-Peña. 2013]. Uma tendência recente importante é constatação de que questões sobre desempenho e escalabilidade são agora importantes para a efetiva disseminação da tecnologias de repositórios semânticos [dos Santos, 2014].

4. Metodologia de Avaliação

A medição do desempenho, escalabilidade e estabilidade dos bancos de dados semânticos foi feita em condições controladas de laboratório, através de testes de estresse similares aos empregados na avaliação do desempenho de bancos de dados NOSQL baseados em RDF [Bizer e Schultz, 2009; 2011] e repositórios semânticos de OA que também empregam a mesma tecnologia de banco de dados [dos Santos, 2014; Gluz e Silva, 2014].

Em termos de recursos físicos foi empregado um laboratório de pesquisa em agentes inteligente e bancos de dados semânticos, constituído de um *cluster* experimental de 8 máquinas de alto desempenho interconectados por interfaces Gigabit a um *switch* de alta capacidade, 6 estações de trabalho (PCs) para fins de projeto e desenvolvimento e dois monitores de 40 polegadas para visualização de dados.

A metodologia de experimentação visa avaliar a eficácia e desempenho de cada banco semântico em termos de consultas e atualizações (*updates*) SPARQL. O processo de avaliação se dividiu nas etapas de seleção dos bancos semânticos e aplicação de uma metodologia de avaliação de desempenho aos bancos previamente selecionados.

A escolha dos bancos de dados NOSQL foi baseado nos critérios de: (a) licenciamento livre (pelo menos para fins experimentais e de pesquisa) e disponibilidade de código aberto; (b) popularidade ou maior utilização atual entre os usuários; (c) suporte ao armazenamento nativo de dados em triplas RDF (sem uso de banco de dados relacional para esses fins); (d) suporte completo à SPARQL; (e) suporte para inferência OWL/RDF.

Os critérios (c) e (d) garantem que o banco de dados NOSQL seja um banco RDF, enquanto que o critério (e) garante os aspectos semânticos do banco (ver Subseção 2.5). O critério (a) é apenas uma condição operacional para a realização dos testes e o critério (b) visa identificar quais bancos de dados RDF são mais usados atualmente.

O site DB Engines [DBEngines, 2016] foi instrumental para esse processo de seleção. A Figura 2 mostra as informações sobre o ranking dos bancos de dados semânticos obtidos no mês de maio de 2016 obtidas neste site.

Rank			DBMS	Database Model	Score		
May 2016	Apr 2016	May 2015			May 2016	Apr 2016	May 2015
1.	1.	1.	MarkLogic	Multi-model	9.07	-0.04	-0.83
2.	2.	2.	Virtuoso	Multi-model	2.15	-0.24	-0.91
3.	3.	3.	Jena	RDF store	1.79	+0.05	+0.30
4.	4.	4.	Sesame	RDF store	1.33	+0.01	-0.12
5.	5.	↑ 6.	Algebraix	RDF store	0.67	+0.01	+0.19
6.	6.	↑ 7.	Stardog	Multi-model	0.54	-0.03	+0.30
7.	7.	↓ 5.	AllegroGraph	Multi-model	0.50	-0.08	-0.28
8.	8.	↑ 10.	4store	RDF store	0.24	-0.02	+0.14

Figura 2. Classificação dos Bancos NOSQL RDF no site DB-Engines. Fonte [DBEngines, 2016]

Embora o banco de dados Marklogic apresentasse bons resultados em termos de popularidade, ele foi excluído da seleção por disponibilizar uma licença limitada em tempo e volume de dados para fins de testes, além de possuir restrições explícitas na licença de desenvolvimento [MarkLogic, 2016], quanto à divulgação de testes de desempenho feitos com o banco de dados. Conforme os critérios citados acima, foram selecionados para avaliação de desempenho os bancos de dados: Virtuoso [Virtuoso, 2016], Apache Jena TDB [Apache, 2016] e Sesame/RDF4J [Sesame/RDF4J, 2016].

A metodologia geral de avaliação dos bancos de dados semânticos é derivada da metodologia Berlin de *benchmarks* de bancos de dados RDF e consulta SPARQL [Bizer e Schultz, 2009; 2011], adaptada para a avaliação de desempenho destes bancos aplicados na construção de repositórios semânticos. Estudos recentes [dos Santos et al., 2014; Gluz e Silva, 2014] mostram a importância dessa adaptação para se obter uma análise e

avaliação mais precisa deste tipo de aplicação de bancos de dados semânticos.

A adaptação impactou na metodologia Berlin nos pontos mais “customizáveis” dessa metodologia. A metodologia Berlin não obriga que os dados RDF estejam conformes a uma dada ontologia OWL, porém na metodologia proposta neste trabalho se estabeleceu que as triplas RDF devem seguir a estrutura definida pela ontologia OBAA [Gluz e Vicari, 2011] por ser capaz de representar informações (metadados) de conteúdos educacionais compatíveis com os formatos de metadados de OA: LOM, Dublin-Core e OBAA. Para se ter um conjunto de dados realista para testes relacionados à repositórios de conteúdos educacionais, estes dados serão extraídos de repositórios de OA existentes atualmente e convertidos para representação RDF de acordo com o modelo OWL da OBAA. Seguindo a abordagem de [Gluz e Silva, 2014] serão utilizados para dados de testes de desempenho os metadados dos objetos armazenados no repositório BIOE [<http://objetoseducacionais2.mec.gov.br/>] que atualmente possui quase 20.000 OAs. Como cada OA pode ter mais de 20 diferentes tipos de metadados armazenados no repositório, cada um representado por pelo menos uma tripla RDF, isso implica em um volume considerável de triplas para fins de experimentos de desempenho. Os procedimentos de execução de testes de consultas e atualizações SPARQL serão baseados em cenários vinculados ao uso de OA. Os cenários de avaliação de desempenho definidos em [Gluz e Silva, 2014] foram adaptados para permitir a avaliação do impacto trazido pelos processos de inferência nas consultas.

Por outro lado, o modelo de métricas de desempenho adotado na avaliação de bancos para repositórios semânticos é o mesmo definido na metodologia Berlin [Bizer e Schultz, 2010], reforçando a generalidade dos resultados obtidos com os teste de desempenho.

5. Experimentos e Resultados

Na etapa inicial da experimentação, os metadados de todos os OA contidos no BIOE foram lidos e convertidos para o formato RDF/OWL definido pela ontologia OBAA v. 3.0. No total 19.969 OAs foram identificados e baixados do repositório BIOE. A conversão para triplas RDF compatíveis com a ontologia OBAA v3.0 resultou em um total de 489.763 triplas. Para a carga e conversão dos objetos do BIOE foi desenvolvida uma aplicação em JavaScript+HTML5, rodando em NodeJS, capaz de localizar, baixar e converter para o formato de triplas da ontologia OBAA (ver Figura 3).

Para a condução dos experimentos de avaliação de desempenho em laboratório, foram utilizadas duas máquinas com a seguinte configuração de hardware: processador Intel i5, 4GB de memória RAM, HD 500GB SATA, com sistema operacional Linux Ubuntu 16.04. Os seguintes softwares foram utilizados nos experimentos: Java 8, update 66, banco de dados Apache Jena v3.3 e banco de dados RDF4J v2.2.

A medição do desempenho, escalabilidade e estabilidade dos bancos de dados semânticos foi feita em condições controladas de laboratório, através de experimentos de medição de estresse desenvolvidos especificamente para avaliar a eficácia e desempenho de cada banco de dados em termos de consultas e atualizações SPARQL sobre conteúdos educacionais representados na forma de objetos de aprendizagem. Tais experimentos foram baseados em métodos consolidados de avaliação de bancos RDF [Bizer e Schultz, 2011], adaptados para os requisitos específicos dos conteúdos educacionais. O presente

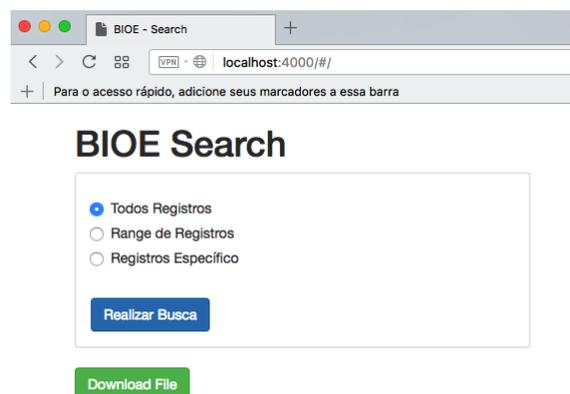


Figura 3. Aplicação para conversão de objetos BIOE para OBAA. Fonte Autores

trabalho apresenta os resultados dos experimentos, identificando as melhores soluções tecnológicas candidatas para construção do ILCMS.

O experimento de desempenho das consulta (*query*) de informações do banco teve como objetivo medir o tempo de resposta do banco para consultas SELECT em SPARQL relativas a triplas RDF individuais e também a todas as triplas RDF de um dado OA. Os testes foram realizados através da simulação de clientes do banco de dados, que acessam os dados do banco através de consultas SPARQL enviadas por meio de web services ao banco. Os clientes simulados foram desenvolvidos em Java, utilizando a interface API Rest para acessar os web services. Na aplicação de simulação, os clientes são executados concorrentemente através de *threads* distintas. Cada uma dessas *threads* representa um cliente capaz de executar consultas SPARQL que usa a interface API Rest para enviar as consultas diretamente ao banco em análise.

Inicialmente foram definidas séries de experimentos de desempenho de consulta começando com a simulação de 300 clientes, com cada cliente executando 500 consultas, e avançando em passos de adição de 200 clientes por vez, para identificar os limites e capacidades do banco, ou seja, testes de desempenho com 500, 700, 900 e 1100 clientes. Entretanto nem o RDF4J nem o Apache/Jena executando no servidor disponível no laboratório foram capazes de tratar experimentos com SELECTs acima de 500 clientes simulados. Mesmo com 600 clientes não foi possível medir o desempenho. Para reduzir as variáveis do experimento e minimizar qualquer possível interferência do próprio software de simulação na capacidade do servidor, este software foi executado em um servidor separado com a mesma configuração do servidor que estava executando o bando de dados. As conexões e consultas SPARQL foram feitas através da rede local.

Cada um dos experimentos de desempenho foi executado em duas modalidades: com o mecanismo de inferência OWL desativado e com o mecanismo de inferência OWL ativado.

A Figura 4 mostra as consulta utilizada quando o mecanismo de inferência está desativado. O painel da esquerda (a) dessa Figura mostra a consulta usada para avaliar o desempenho relativo apenas a triplas RDF sem considerar a estrutura de um OA. Note a seleção *filter(regex ...)* nesta consulta, sugerida pela metodologia Berlin, que exige muitos recursos de processamento. O painel da direita (b) mostra a consulta

realizada no caso de um OA. Nesta consulta o identificador do OA na base de dados (bioeID na Figura 4) é selecionado aleatoriamente para cada consulta gerada.

As consultas quando o modo de inferência está ativado são similares às apresentadas na Figura 4, a diferença está na incorporação ao campo WHERE do SELECT da consulta a uma tripla RDF que deve ser obtida por inferência inversa de uma propriedade de objetos OWL¹.

<pre> 1 PREFIX obaa: <http://obaa.unisinos.br/obaa3.owl#> 2 3 select ?suj ?obj ?description 4 where { 5 ?suj ?pred obaa:LearningObjectMetadata. 6 ?suj obaa:requirement ?obj. 7 ?obj obaa:description ?description 8 filter(regex(?description, ".mp3")) 9 }</pre>	<pre> 1 PREFIX obaa: <http://obaa.unisinos.br/obaa3.owl#> 2 3 select ?pred ?obj 4 where { 5 obaa:bioeID ?pred ?obj 6 }</pre> <p style="text-align: right; margin-right: 20px;">Aleatório</p>
--	--

Figura 4. (a) Consultas de triplas individuais (b) Consultas de OAs. Fonte Autores

A Figura 5 mostra os resultados dos experimentos de desempenho de consultas: o quadro da esquerda mostra os resultados com 300 clientes simulados e o quadro da direita com 500 clientes simulados. O desempenho foi medido em consultas respondidas por segundo (Q/S) e em OAs retornados por segundo (OA/S). Note que a Figura 5 não diferencia os experimentos realizados com e sem a inferência. Isso ocorre porque não houve diferença significativas (apenas na segunda casa decimal) nos resultados entre as consultas selecionando triplas inferidas, de consultas selecionados triplas explicitamente declaradas (*asserted*). Conforme pode ser observado na Figura 1, nos experimentos realizados até o momento o banco que tem se saído melhor foi o Apache Jena. A avaliação do banco de dados Virtuoso ainda está em curso.

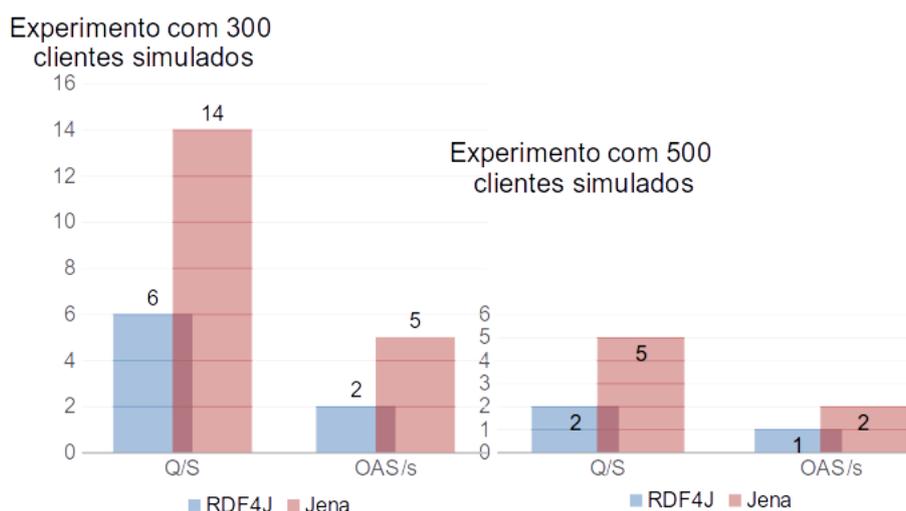


Figura 5. Resultados dos experimentos de desempenho de consultas. Fonte Autores

¹Nestes casos a base tem a asserção de que o objeto OWL O_1 está relacionado ao objeto O_2 pela propriedade P , mas como a ontologia OBAA define várias propriedades como inversas, então se P é uma dessas propriedades ela tem uma propriedade inversa P^{-1} que deve ser obtida por inferência. Neste caso pode-se fazer a consulta se O_2 está relacionado a O_1 pela propriedade P^{-1} .

6. Conclusão

Este trabalho identificou três candidatos para serem utilizados como bancos de dados NOSQL capazes de lidar com a semântica de conteúdos educacionais representados sob a forma de OA. Também foram definidas as diretrizes principais do processo de avaliação de desempenho destes bancos e apresentados os resultados preliminares dessa avaliação. Apesar do banco Virtuoso ainda estar sendo avaliado, os resultados vem consolidando a escolha pelo Apache Jena como uma boa ferramenta para a criação de um repositório semântico de conteúdos educacionais. Entretanto, o resultado mais relevante obtido até agora possivelmente é o fato da inferência OWL não causar, pelo menos nas condições testadas, um atraso ou perda de desempenho. Isso é fundamental para o tratamento semântico: sem inferência lógica (Lógica Descritiva no caso), o tratamento da "semântica" se torna muito limitado, praticamente inefetivo.

Referências

- Apache (2016) *Página oficial do produto*. Disponível em: <<http://jena.apache.org/index.html>>.
- Balog-Crisan,R.; Roxin, I. (2008) Semantic Learning Content Management System. *Procs. of IADIS Int. Conf. on e-Learning*.
- Bittencourt, I. I.; Costa, E.; Silva, M.; Soares, E. (2009) A computational model for developing semantic web-based educational systems. *Knowledge-Based Systems*, v. 22, n.4, p.302–315.
- Bizer, C.; Schultz, A. (2009) The Berlin SPARQL Benchmark. In: *Int. J. on Semantic Web & Inf. Systems*, Vol. 5, Issue 2, Pages 1-24.
- Bizer, C.; Schultz, A. (2011) *Berlin SPARQL Benchmark (BSBM) Specification - V3.1*. Disponível em <<http://wifo5-03.informatik.uni-mannheim.de/bizer/berlinsparqlbenchmark/spec/index.html>> .
- Bizer, C.; Schultz, A. (2010) *Berlin SPARQL Benchmark (BSBM) - Benchmark Rules* Disponível em <<http://wifo5-03.informatik.uni-mannheim.de/bizer/berlinsparqlbenchmark/spec/BenchmarkRules/index.html>>
- Carrion, J.S.; Gordo, E.G.; Sanchez-Alonso, S. (2007) Semantic learning object repositories. *Int. J. of Cont. Eng. Educ. and Life Long Learn.*, V.17, I.6.
- Cobo,J.L.; Arroyo, S.; Sicilia, M. Sánchez, S. (2007) A Semantic Web Service Architecture for Learning Object Repositories. In: Li, E.Y. (Ed) *Adv. in Elect. Business*, V. 2.
- DBEngines (2016) *DB-Engines Ranking of RDF Stores*. Disponível em <<http://db-engines.com/en/ranking/rdf+store>>.
- Dicheva D. et al. (Eds.) (2009) *Semantic Web Technologies for e-Learning*. IOS Press.
- Fernández-Peña, F.; Ramos-Moreno,Y.; Ponce-Toste, Y. (2013) Semantic Enhancement of Learning Object Repositories. *Procs. of Eureka-2013*.
- Gluz, J. C.; Silva, L. R. J. (2014) Ontological semantic search of learning objects: experiments and results. *Anais SBIE 2014*.

- Gluz, J. C., Vicari, R. M. (2011) Uma Ontologia OWL para Metadados IEEE-LOM, Dublin-Core e OBAA. *Anais do SBIE 2011*.
- Gluz, J. C. ; Vicari, R. M. (2012) An Agent-Based Infrastructure for the Support of Learning Objects Life-Cycle. *Procs. of ITS 2012*, Crete.
- Greenberg, L. (2002) LMS and LCMS: What's the difference? *Learning Circuits*, Dec.
- Harman, K.; Koohang, A. (Eds.). (2007). *Learning Objects: Standards, Metadata, Repositories, & LCMS*. Inf. Sci. Press.
- Irlbeck, S.; Mowat, J. (2007) Learning Content Management System (LCMS). In Harman, K.; Koohang, A. (Eds.) *Learning Objects: Standards, Metadata, Repositories, & LCMS*. Inf. Sci. Press.
- Isotani, S.; Mizoguchi, R.; Bittencourt, I.; Costa, E. (2008) Web 3.0 - Os Rumos da Web Semântica e da Web 2.0 nos Ambientes Educacionais. *Anais do XIX SBIE* (SBIE 2008), Fortaleza.
- Kholief, M.; Nada, N.; Khedr, N. (2012) Ontology-Oriented Inference-Based Learning Content Management System. *Int. J. of Web & Semantic Technology (IJWesT)* V.3, N.3, July.
- Lama, M.; Vidal, J.C.; Otero-García, E.; Bugarín, A.; Barro, S. (2011) Semantic Linking of a Learning Object Repository to DBpedia. *Procs. of 11th ICALT*.
- MarkLogic. (2016) *Developer License - Enterprise NoSQL Power for Developers*. Disponível em <<http://developer.marklogic.com/free-developer>>.
- Mohan, P.; Brooks, C. (2003) Learning Objects on the Semantic Web., 2003. *Procs of 3rd IEEE Int. Conf. on Advanced Learning Technologies*.
- Pahl, C.; Javed, M.; Abgaz, Y.M. (2013) Ontology Evolution for Learning Content Management Systems. *Procs of EdMedia 2013*.
- Porto, F.; Moura, A.M.C.; da Silva, F.J.C.; Fernandez, A.P. (2007) The ROSA project: leveraging e-learning to a semantic layer. *Int. J. Know. and Learn.*, V.3, N.1.
- dos Santos, H.; Carrillo, G.; Cechinel, C.; Ochoa, X. (2014) Towards the use of Semantic Learning Object Repositories: Evaluating Queries Performance in two Different RDF Implementations. *Bull. of IEEE Techn. Comm. on Learn. Techn.*, V.16, N.4.
- Sesame/RDF4J. (2016) Página oficial do produto. Disponível em <<http://rdf4j.org/about/>>.
- Sikos, L.F. (2015). *Mastering Structured Data on the Semantic Web*. Editora Apress.
- Silva, M.; Barros, H.; Veras, D.; Pacca, H.; Ibert, I.; Barros, E.; Silva, (2009) A. Modelando um Sistema Educacional de MMC sob a perspectiva da Web Semântica. *Anais do XX SBIE* (SBIE 2009).
- Stojanovic, L.; Staab, S.; Studer, R. (2001) eLearning based on the Semantic Web. *Procs of WebNet2001*.
- Virtuoso. (2016) *Página oficial da distribuição open-source do produto*. <https://github.com/openlink/virtuoso-opensource>