Evaluation Dataset and Graphs

The objective of this research is the performance evaluation of Knowledge Based System(KBS) on OWL2 semantics. This requires dataset of large size with OWL2 semantics for measuring the performance of KBS. Therefore, we are interested in using a large and OWL2 semantic enabled dataset. There are quite number of dataset available for the testing purpose like LUBM dataset [1], DBpedia [2] SP^2B [3] and UOBM [4]. These datasets fulfils the scalability factor but lacks OWL2 semantics.

The UOBM dataset [4] is chosen in our benchmark because its dataset covers most of the OWL semantics in comparison of the surveyed evaluation benchmarks. The dataset is provided with the UOBM benchmark [4]. With the help of UOBM dataset generator, user specified dataset is created. The structure of the dataset belongs to university domain. The missing OWL2 semantics have been added in the dataset.

Benchmark Dataset

The benchmark dataset is analyzed using Protégé plug-in.

Statistics of the proposed university ontology dataset

A detail analysis of the dataset provides the characteristics of evaluated data that are presented in Table 1.

Table 1: Summary Statistics of proposed university ontology dataset

Dataset Characteristics	
Total Number of Triples	2727936
Total Nodes	26878
Total Number of Classes	134
Total Number of Object Properties	45
Total Unique Properties	9

Namespaces

The namespaces used the dataset are mentioned in Table 2.

Table 2: Prefix to URI mapping

Prefix	URI
UB:	http://semantics.crl.ibm.com/univBench.owl#
rdf:	http://www.w3.org/1999/02/22-rdf-syntax-ns#
owl:	http://www.w3.org/2002/07/owl#
rdfs:	http://www.w3.org/2000/01/rdf-schema#

Dataset Scaling

We perform our evaluation on four different sizes of university dataset. These are Dataset 1, Dataset 2, and Dataset3 contains 24K, 240K and 2400K triples respectively.

Dataset Cleaner

In the proposed dataset generator, the redundant triples are identified and eliminated by identifying all the triples, which are associated with object properties characteristics. So that single copy of the triple can be retained.

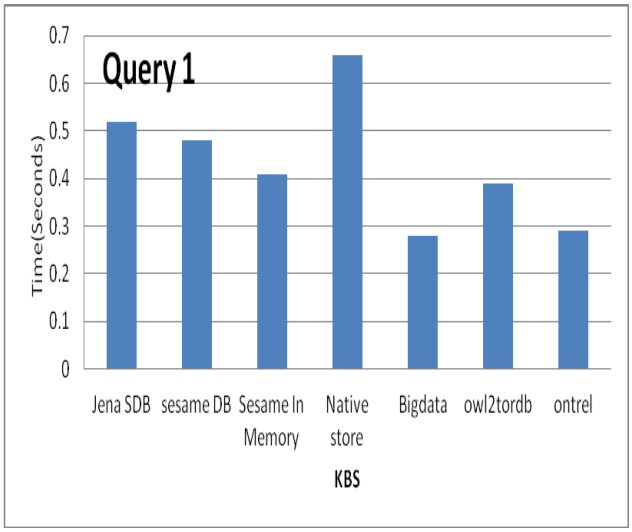
For example, out of the following triples only one triple is used.

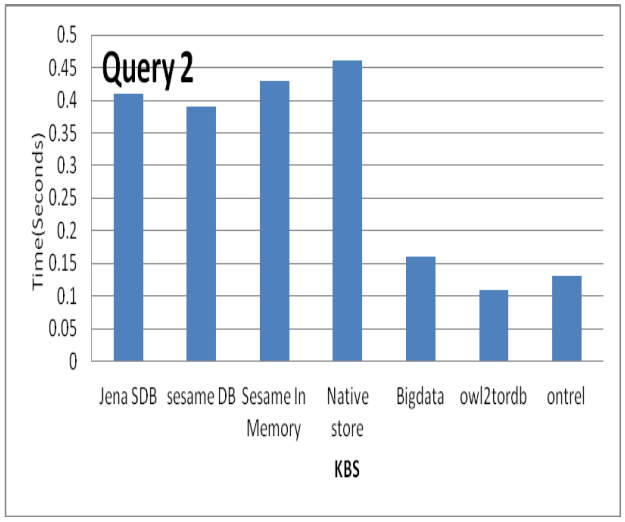
<owl:NamedIndividual rdf:about="http://www.Department10.University0.edu/AssistantProfessor0"> <**isFriendOf** rdf:resource="http://www.Department10.University0.edu/AssistantProfessor8"/>
<owl:NamedIndividual rdf:about="="http://www.Department10.University0.edu/AssistantProfessor8"/ > <**isFriendOf** rdf:resource=" http://www.Department10.University0.edu/AssistantProfessor0""/>

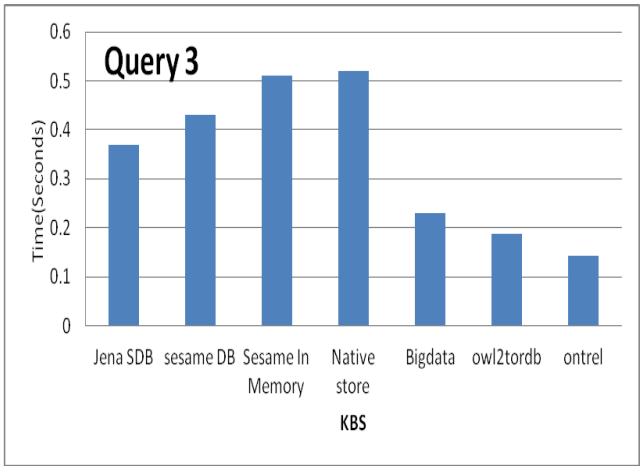
Graphs

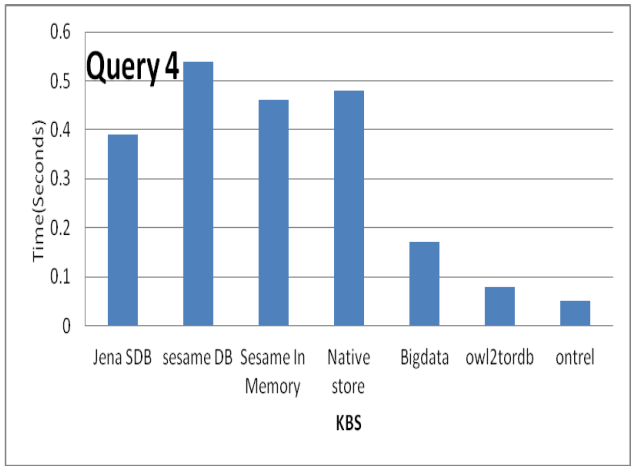
Simple queries

24 K size Dataset

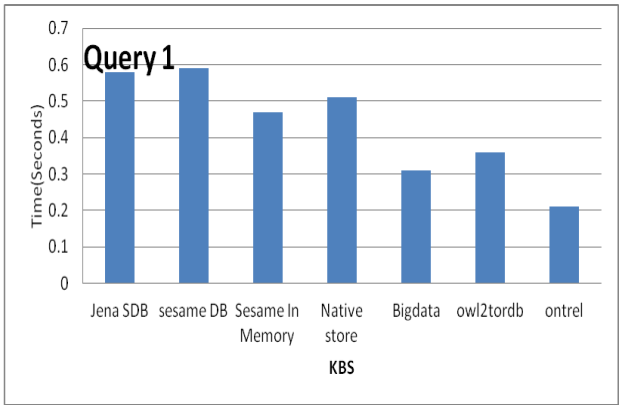


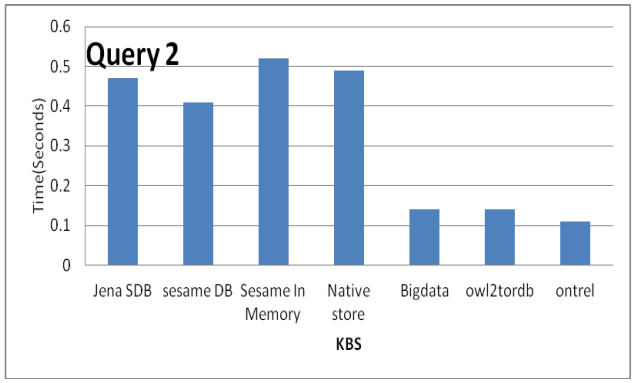


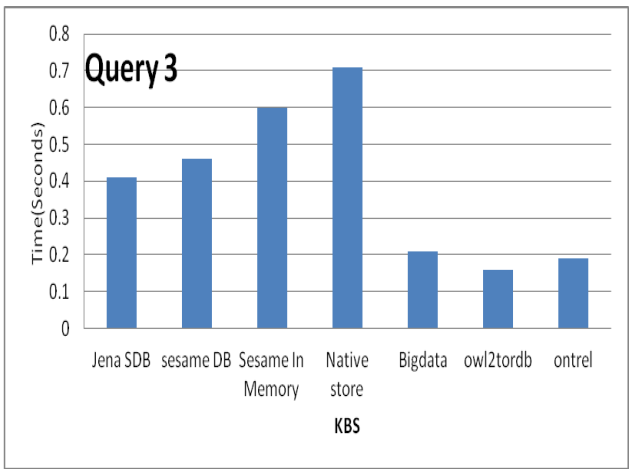


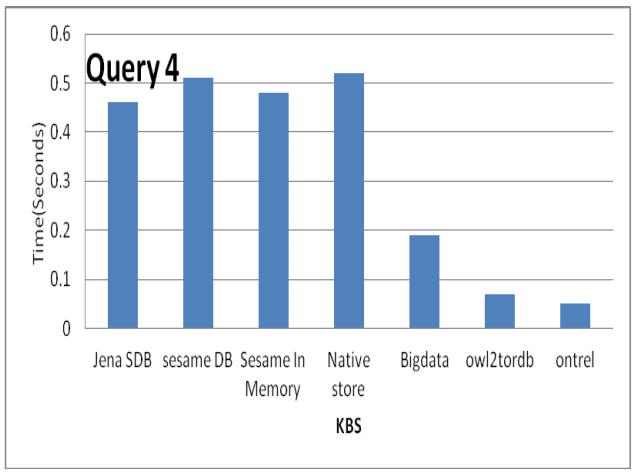


240 K size dataset

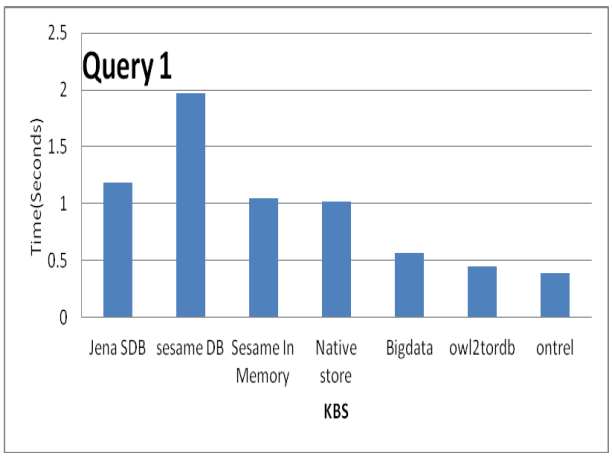


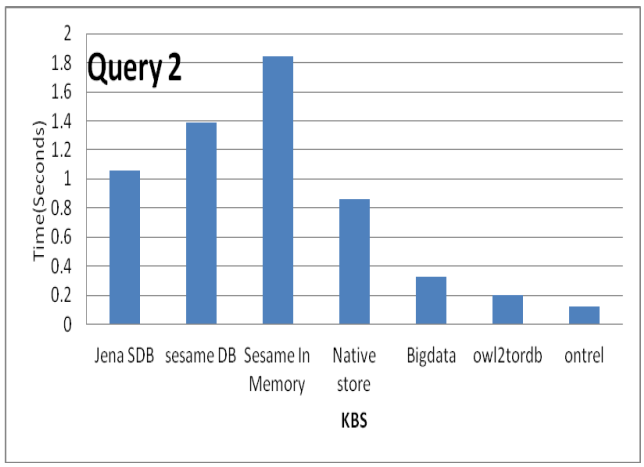


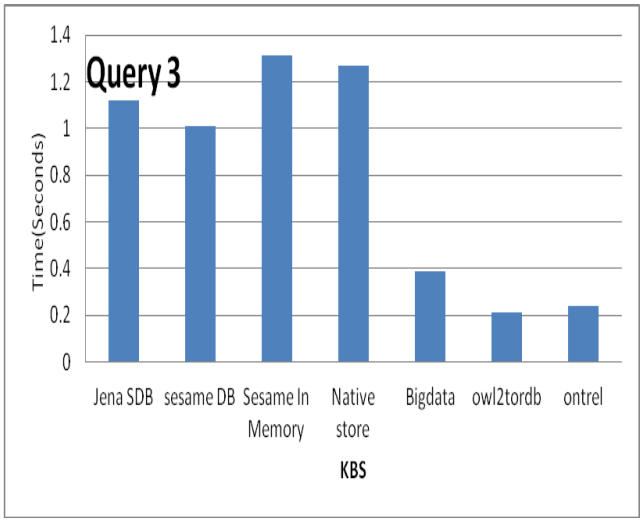


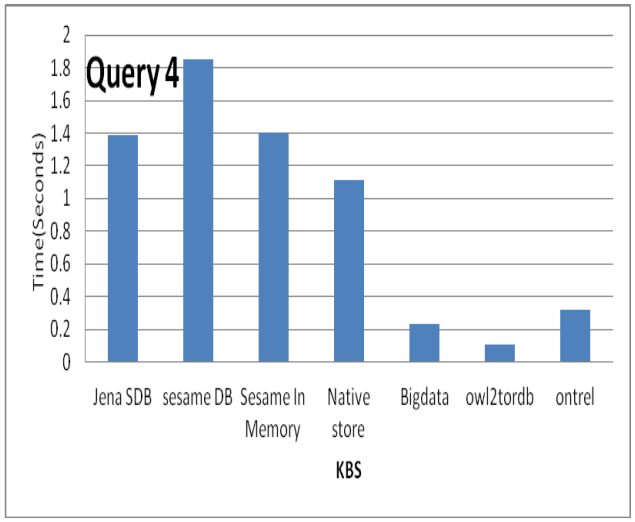


2400 K size Dataset



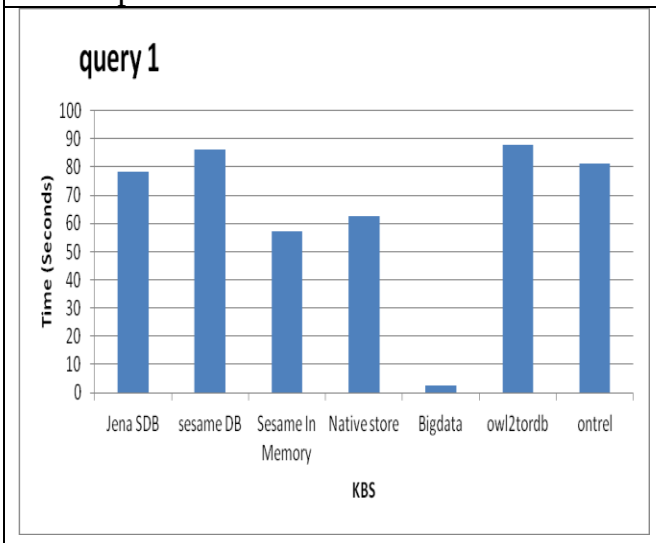


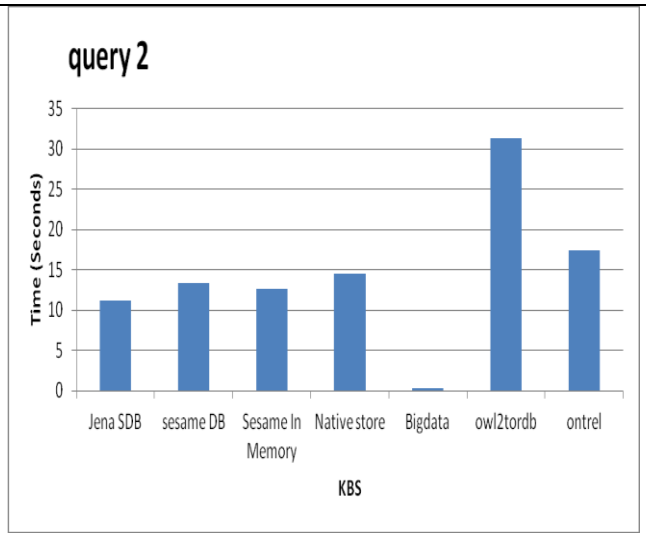


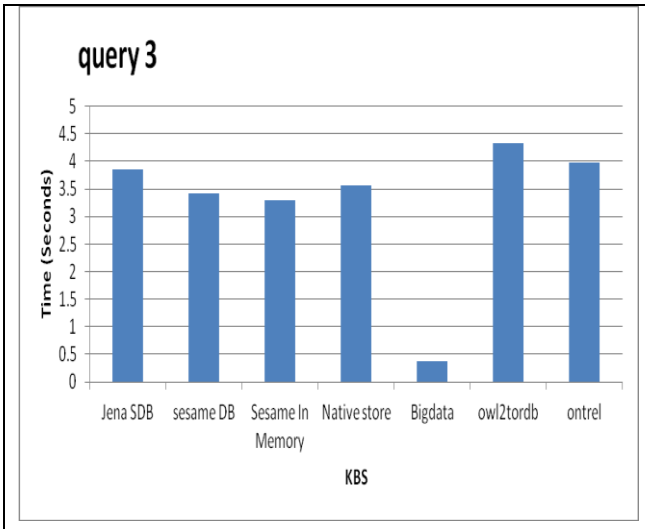


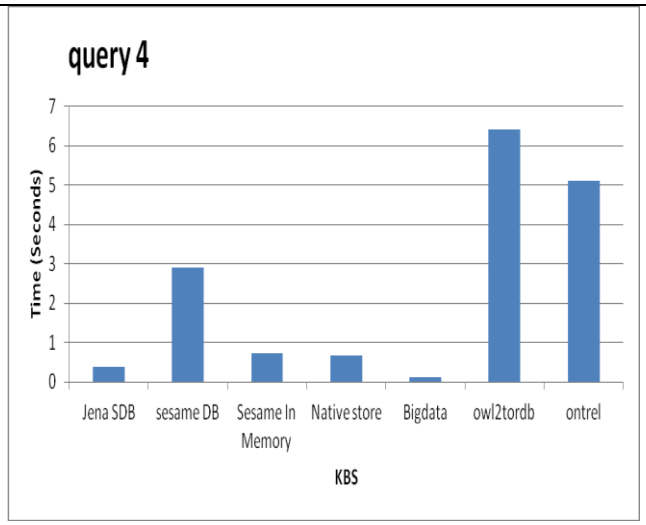
Complex queires

24 K triples

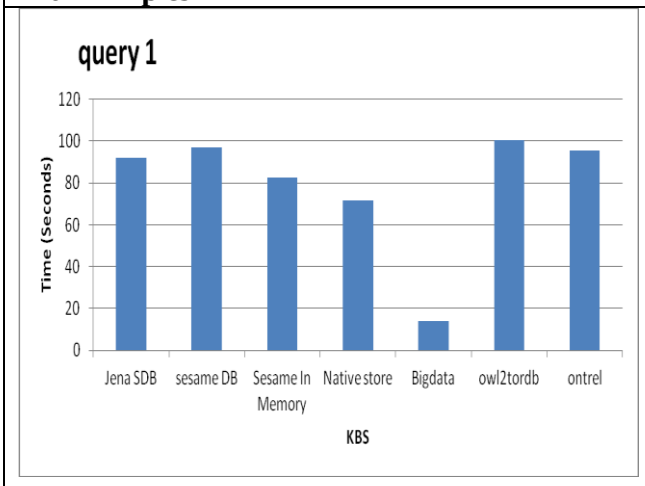


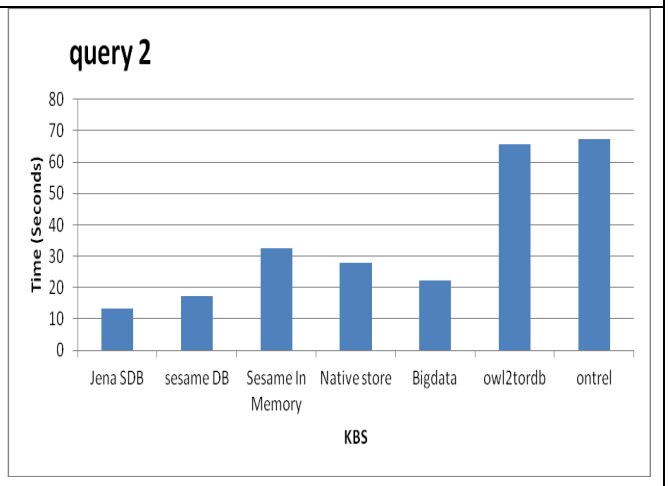


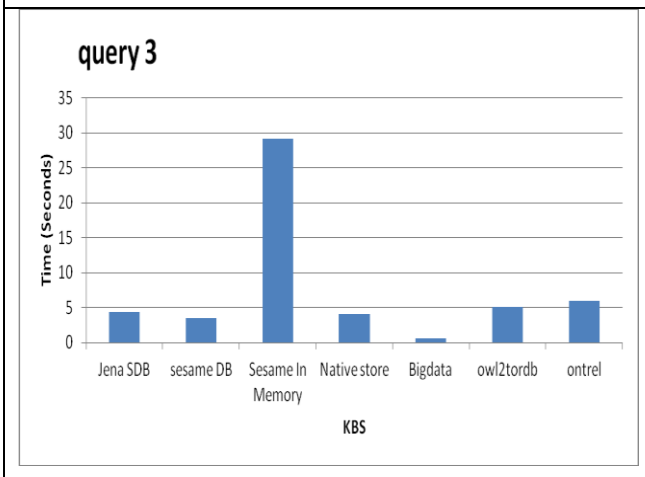


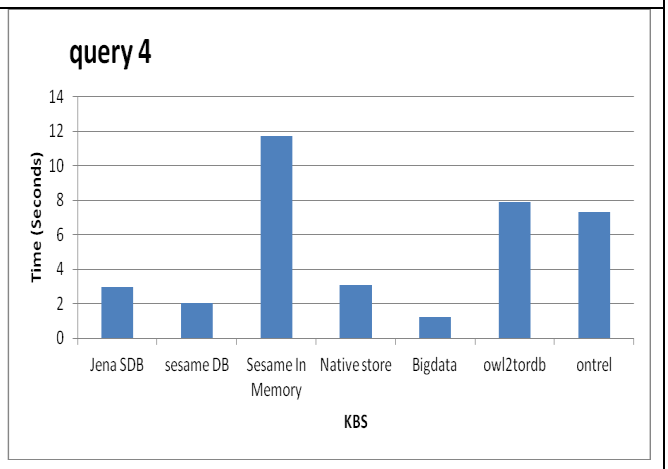


240 K Triples

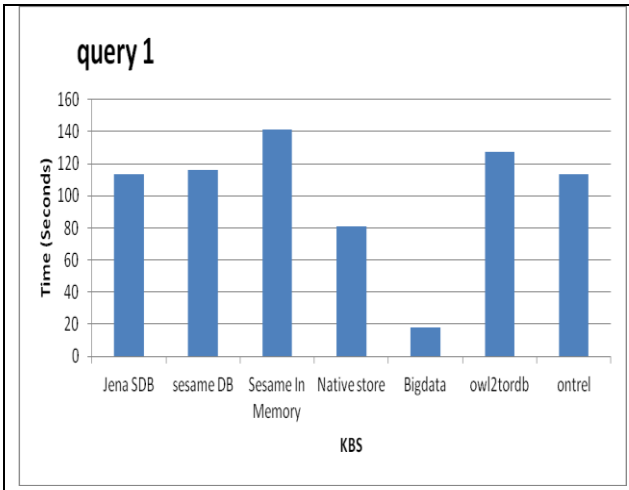


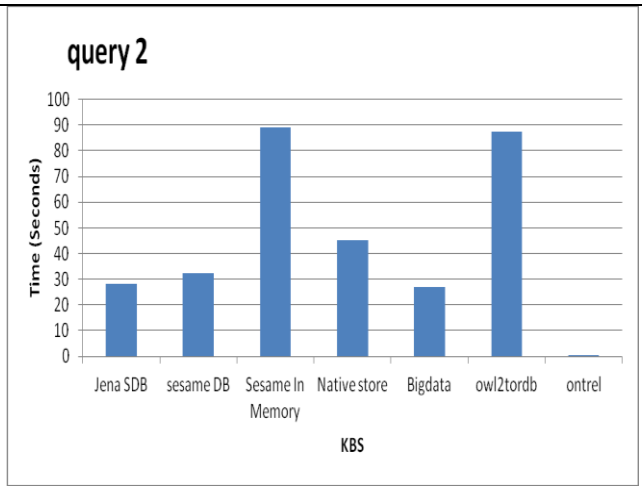


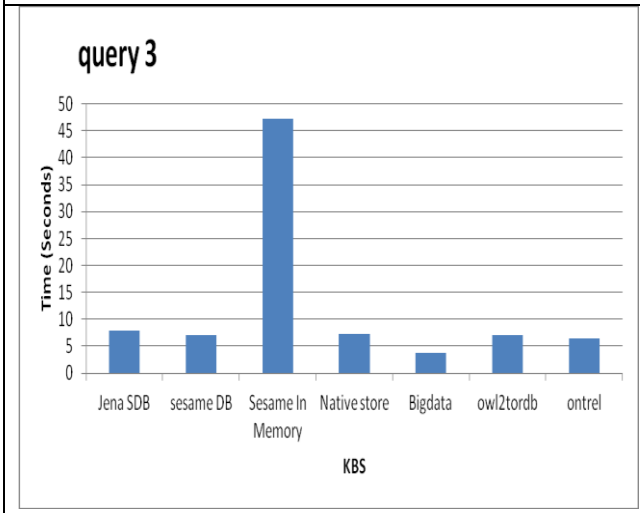


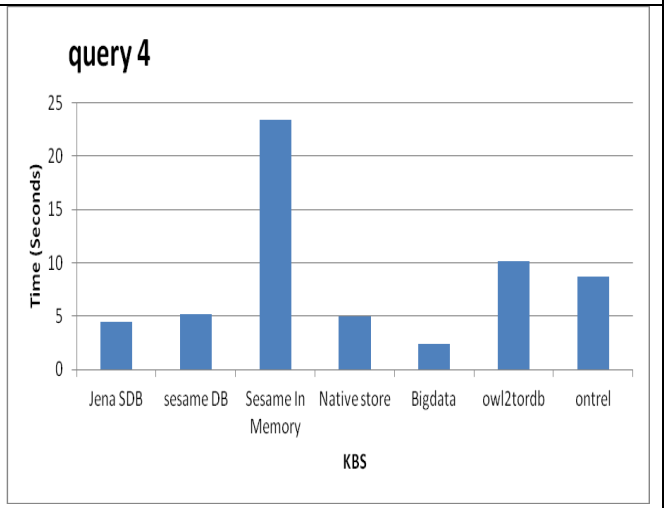


2400 K Triples



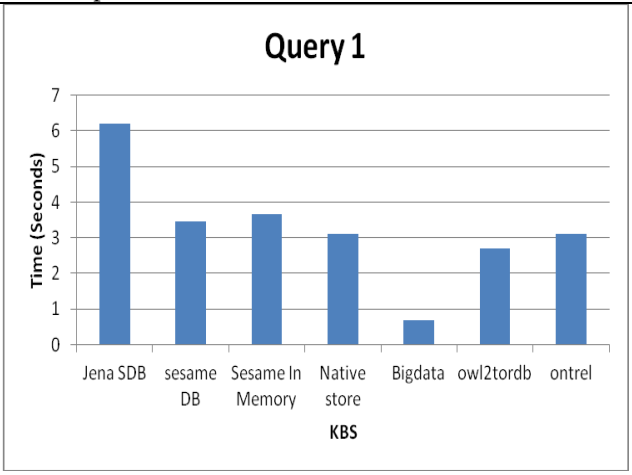


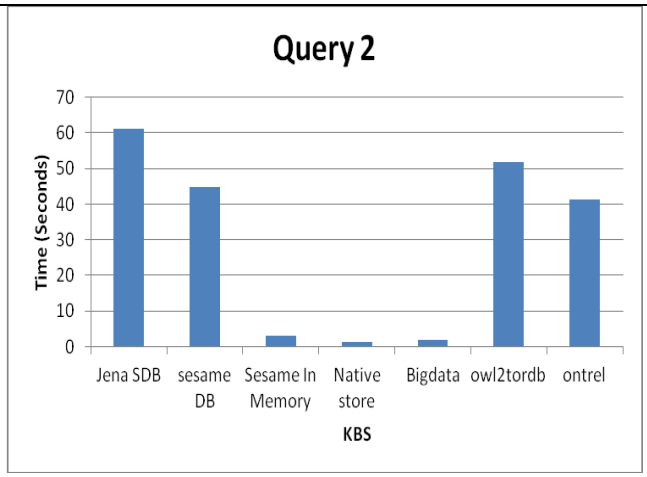


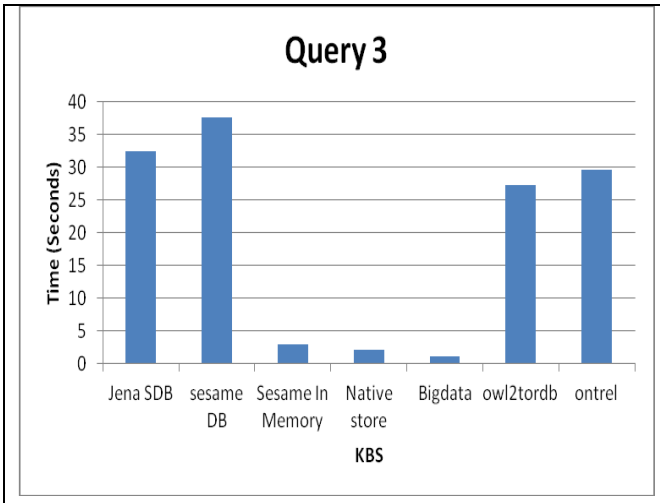


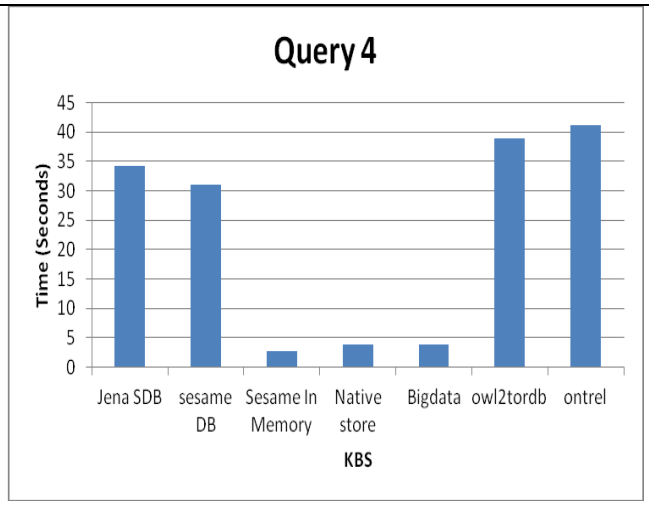
Object properties characteristics pattern queries

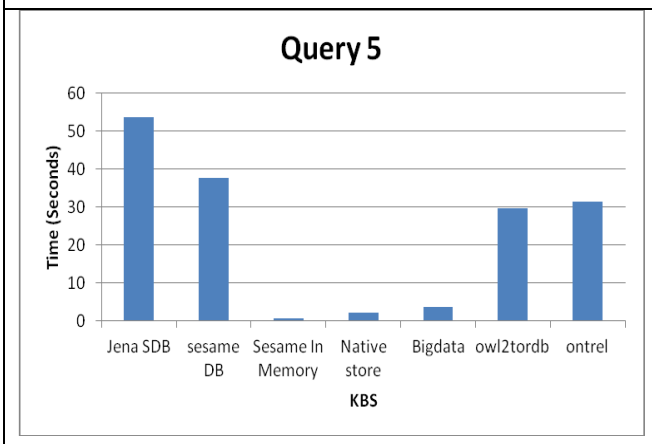
24 K triples

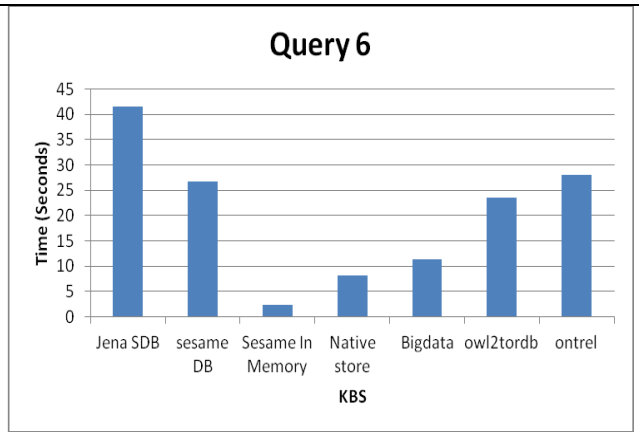


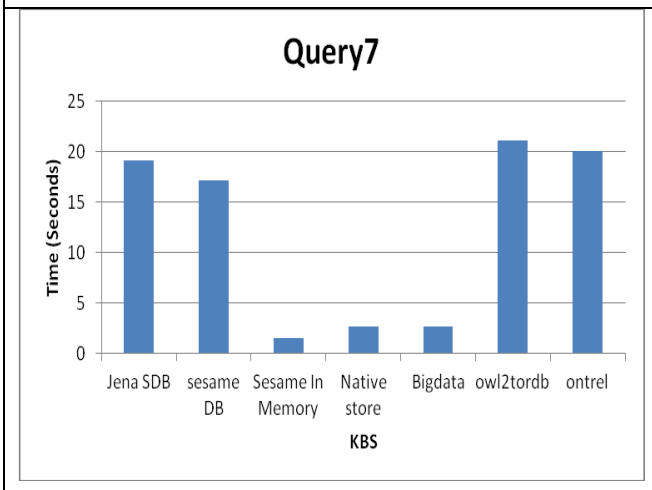


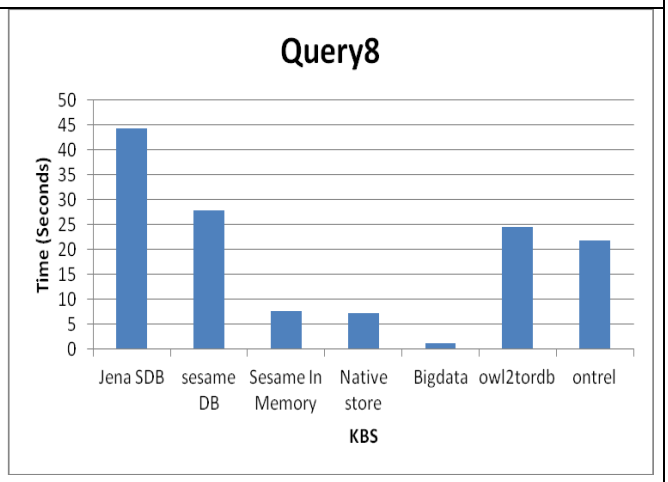

Query 3
Time (Seconds)
0
5
10
15
20
25
30
35
40
Jena SDB
sesame DB
Sesame In Memory
Native store
Bigdata
owl2tordb
ontrel
KBS

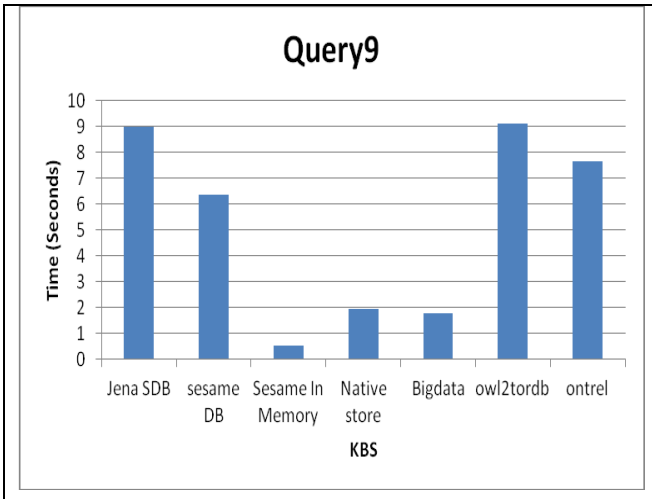

Query 4
Time (Seconds)
0
5
10
15
20
25
30
35
40
45
Jena SDB
sesame DB
Sesame In Memory
Native store
Bigdata
owl2tordb
ontrel
KBS


Query 5
Time (Seconds)
0
10
20
30
40
50
60
Jena SDB
sesame DB
Sesame In Memory
Native store
Bigdata
owl2tordb
ontrel
KBS


Query 6
Time (Seconds)
0
5
10
15
20
25
30
35
40
45
Jena SDB
sesame DB
Sesame In Memory
Native store
Bigdata
owl2tordb
ontrel
KBS


Query7
Time (Seconds)
0
5
10
15
20
25
Jena SDB
sesame DB
Sesame In Memory
Native store
Bigdata
owl2tordb
ontrel
KBS


Query8
Time (Seconds)
0
5
10
15
20
25
30
35
40
45
50
Jena SDB
sesame DB
Sesame In Memory
Native store
Bigdata
owl2tordb
ontrel
KBS



REFERENCES

1. Guo, Y., Pan, Z., & Heflin, J. (2005). LUBM: A benchmark for OWL knowledge base systems. Web Semantics: Science, Services and Agents on the World Wide Web, 3(2), 158-182.

2. Owens, Alisdair, and Nick Gibbins. "Effective benchmarking for RDF stores using synthetic data." (2008).

3. Schmidt, Michael, et al. "SP2Bench: a SPARQL performance benchmark." 2009 IEEE 25th International Conference on Data Engineering. IEEE, 2009.

4. Ma, L., Yang, Y., Qiu, Z., Xie, G., Pan, Y., & Liu, S. (2006). Towards a complete OWL ontology benchmark (pp. 125-139). Springer Berlin Heidelberg.