The LDBC benchmarks will span four main areas of Linked Data management: complex query execution, transactionality in graphs, RDF inference and RDF support for ETL/data integration. They will:

- target hard problems and choke-points that mostly affect users and hence slow down the uptake of RDF and graph database technology,
- encourage innovative performance and scalability improvements that directly benefit users,
- be open, community generated, liberally licensed (open-source/creative commons) and target real-world usage scenario,
- become the de facto standard for publishing performance results and enable objective judgements about the performance and functionality of competing vendor offerings.

The Linked Data Benchmark Council will found a non-profit organization which broad international membership from the graph and RDF data management industry, that will work in the same spirit as the Transaction Processing Council (TPC), which has established set of benchmarks for relational database management systems, widely accepted by the industry. It will be responsible for:

- specifying benchmarks, benchmarking procedures and verifying/publishing results,
- providing an auditing service for certifying results published by vendors for benchmarks endorsed by LDBC,
- training auditors for its benchmarking, creating a long lasting business model for auditing benchmark results.

LDBC will bring together researchers and RDF and graph database vendors to develop and implement the first suite of open, fair and vendor-neutral benchmarks for measuring the performance of RDF and graph databases.
**Scope**

Collaborative Project 2012-2015 in Information and Communication Technologies

The main objective of the Linked Data Benchmark Council (LDBC) is the development of benchmarks for different technology areas including core data management (query processing, query optimisation, transactions), graph analysis, and data integration and reasoning. More specifically, LDBC aims at the

- Development of new benchmarks that will spur research & industry progress in large-scale graph and RDF data management.
- Establishment of an industry-neutral entity (the LDBC non-profit organization, with international members beyond the partners of this EU project) for developing graph and RDF benchmarks, auditing benchmark results, and publishing audited results.

The LDBC audience includes IT professionals considering to use these emerging technologies, researchers in both the database and semantic web research communities, and data management technology vendors.

The ultimate goal of LDBC is allow IT users to better compare RDF and graph database technology products among themselves and competing technologies such as relational database systems, and also spur further innovation in RDF and graph database technology by making improvements measurable.

For more information, please visit the LDBC web page [http://ldbc.eu](http://ldbc.eu).

**Target Audiences**

The target audiences of LDBC that will comprise the core of the LDBC foundation as well as benefiting from and using the project results in the areas of technology, market and education are:

**Technology Users:** This group includes both private and commercial users of RDF and graph databases that will use or integrate this technology for the benefits it has over traditional relational database management techniques.

**Researchers:** This category includes a broad range of researchers from those who focus on graph-shaped data representations, query languages and optimisations, all the way to researchers from other fields who use this technology.

**Technology Vendors:** This group is made up of commercial developers of RDF and graph database software components. It also includes vendors who sell the software they produce as well as those who sell only services around their (open-source) products.

**Project Summary**

Massive amounts of graph shaped data coming from a variety of applications related to social and telecommunication networks are published nowadays. The W3C LOD initiative has boosted the publication and interlinkage of a large number of datasets with billions of RDF triples resulting to the Linked Open Data Cloud.

A significant number of commercial RDF and graph database systems exist but at the present time, there is no independent authority for developing benchmarks to compare these and neutrally assessing benchmark results, through industry-strength auditing.

The LDBC Project tackles the above issues by establishing the Linked Data Benchmark Council, an independent council that will outline the EU project beyond 2015 for developing graph and RDF benchmarks, auditing benchmark results, and publishing audited results.

**Technical User Community**

LDBC will continuously involve technology users through its Technical User Community (TUC) for the purpose of:

- defining real-world usage scenarios,
- identifying performance and usage problems,
- providing datasets, queries, rule-sets, code fragments,
- participating in LDBC Benchmark Task Forces to define new benchmarks.

The main incentive for the end users and data owners to participate in the TUC is to influence the benchmark development agenda of the LDBC, and thereby influence innovation in RDF and graph data management.

The application areas where LDBC is involving users in its TUC are:

- integrating and querying Life Science data,
- dynamic publishing for ontology driven, rich content,
- managing and analyzing social network data, e.g. for marketing,
- analyzing the network behavior in computer or telecommunication networks.

The research agenda of this EU project is aimed at developing a set of initial benchmarks that will spur research & industry progress in large-scale graph and RDF data management for different technology areas including core data management, graph analysis, data integration and reasoning. The research groups involved are experienced in creating RDF and graph database systems so as to ensure that the benchmarks will test systems ‘where it hurts’ -- directing benchmark tests specifically at certain technological ‘choke-points’.

This includes setting challenges that will lead to progress in:

- scalability, storage, indexing and query optimization techniques for RDF and graph database solutions beyond Terabyte scales,
- quantitatively and qualitatively assess different solutions for data integration,
- computationally cheaper reasoning in RDF engines.

http://ldbc.eu