

## What is provenance?

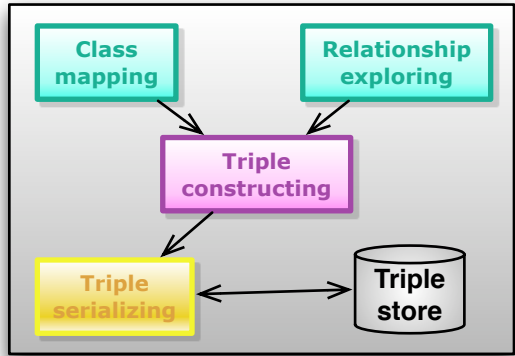
Provenance, from the French word “provenir”, describes the lineage or history of a data entity. Provenance is critical information in scientific applications to verify experiment process, validate data quality and associate trust values with scientific results.

## Provenance Management

Four aspects of the provenance management includes:

- Provenance Capture
- Provenance Representation
- Provenance Storage
- Provenance Query Analysis

## Provenance Capture



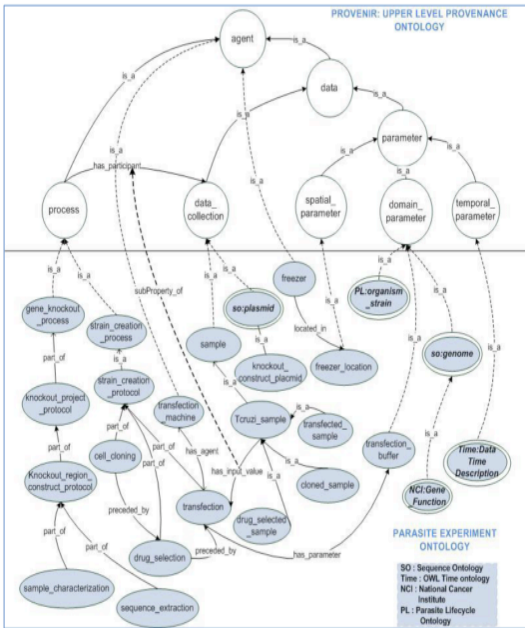
Provenance information collected via Web forms was previously stored in relational database (RDB) then converted into RDF format using ETL.

In this project, we create new Web forms to capture the provenance metadata according to PEO ontology and generate output data in RDF directly.

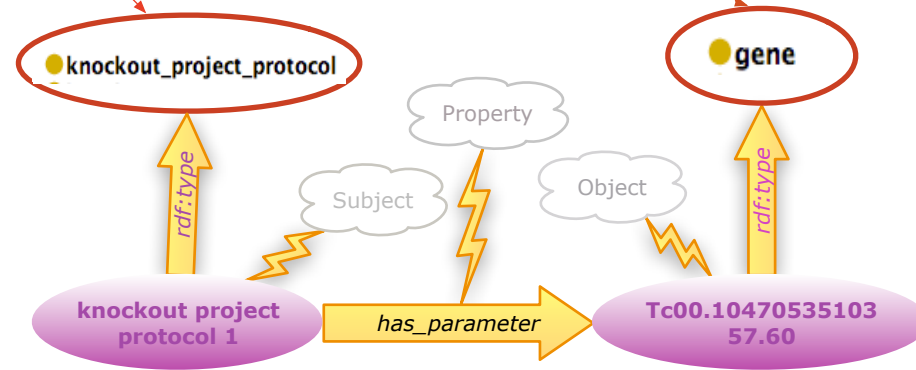
## Application in Parasite Research

## Provenance Representation

The Parasite Experiment Ontology (PEO) models the provenance information associated with GKO and SP experiment protocols by extending the Provenier Ontology.



The screenshot shows a web form for entering gene knock-out information. It is divided into sections: PROJECT INFORMATION, GENE INFORMATION, and INDIVIDUAL ALLELES OF GENE. A 'New Project' dropdown is highlighted in red. The GENE INFORMATION section includes fields for Gene Name, Gene Function, Gene in Cluster, Microarray Cluster, and Subset, along with counts for PDONR stages and other notes.



## References

[1] Tcruzi project page: <http://wiki.knoesis.org/index.php/Trykipedia>