6-2010

Provenance Management in Parasite Research

Vinh Nguyen
Wright State University - Main Campus, nguyenthikim.2@wright.edu

Priti Parikh
Wright State University - Main Campus, priti.parikh@wright.edu

Satya S. Sahoo
Wright State University - Main Campus

Amit P. Sheth
Wright State University - Main Campus, amit@sc.edu

Provenance Management in Parasite Research

Vinh Nguyen, Priti Parikh, Satya S. Sahoo, Amit Sheth
Kno.e.sis Center, Computer Science and Engineering Department, Wright State University

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.

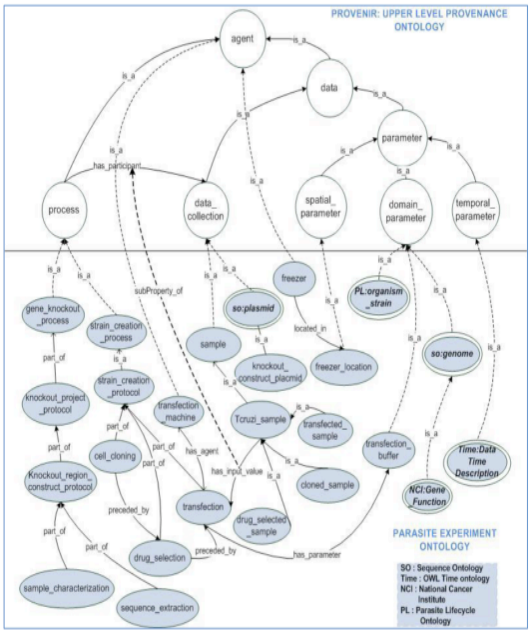
Provenance Management

Four aspects of the provenance management includes:

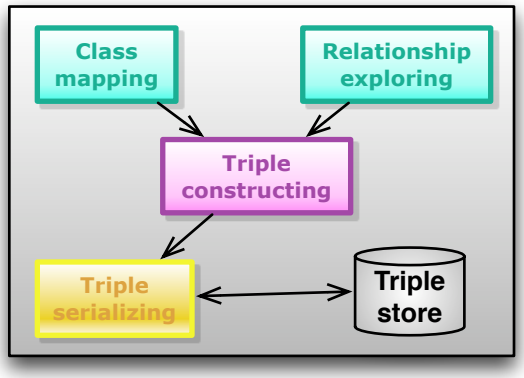
- **Provenance Capture**
- **Provenance Representation**
- Provenance Storage
- Provenance Query Analysis

Provenance Representation

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



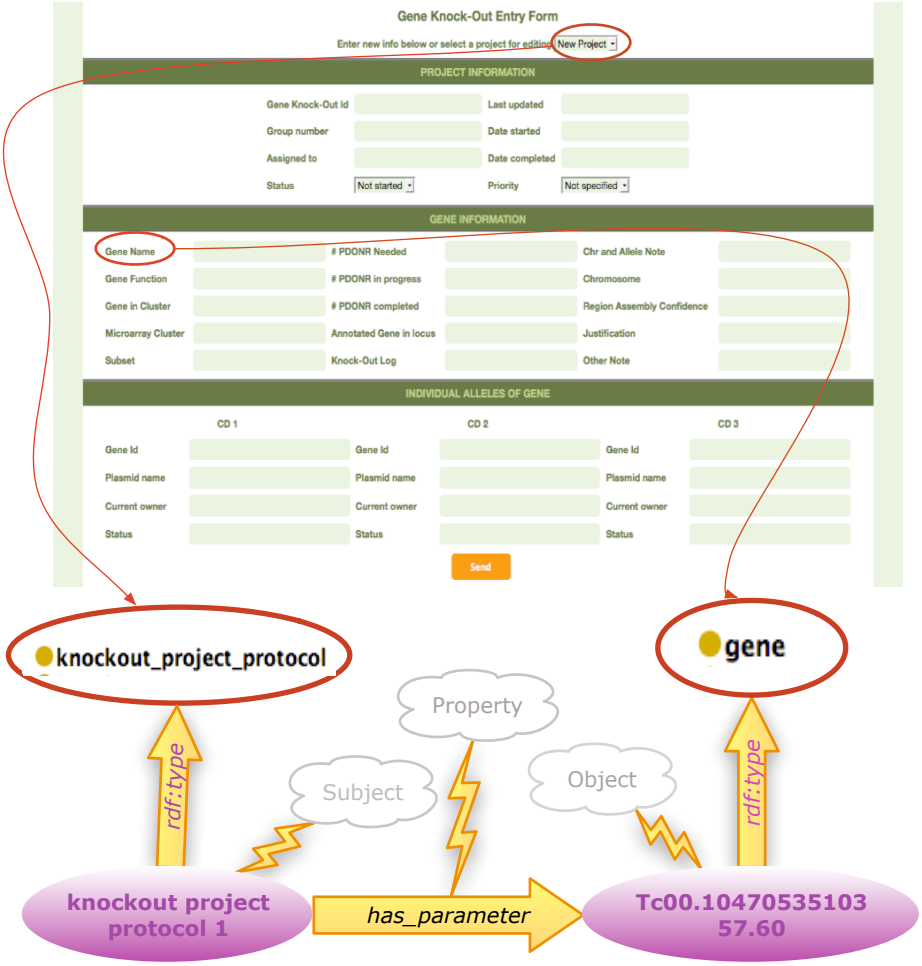
Provenance Capture



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



References

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

Acknowledgement: This work is done as a part of T.cruzi project funded by NIH grant number 1R01HL087795-01A1