CORE
Provided by Nature Precedings
Metadata, citation and similar papers at core.ac.uk

NUI Galway
OÉ Gaillimh

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

Digital Enterprise Research Institute
DERI

A Linked Data Approach to Help Identify Therapeutic Targets for Cancer

Helena F. Deus, Diogo F. Veiga, Pablo R. Freire, John N. Weinstein, Gordon B. Mills and Jonas S. Almeida*
**helena.deus@deri.org*

Objective

To integrate data derived from biological experimental results with public biomedical data sources such as Diseasome, DrugBank and KEGG.

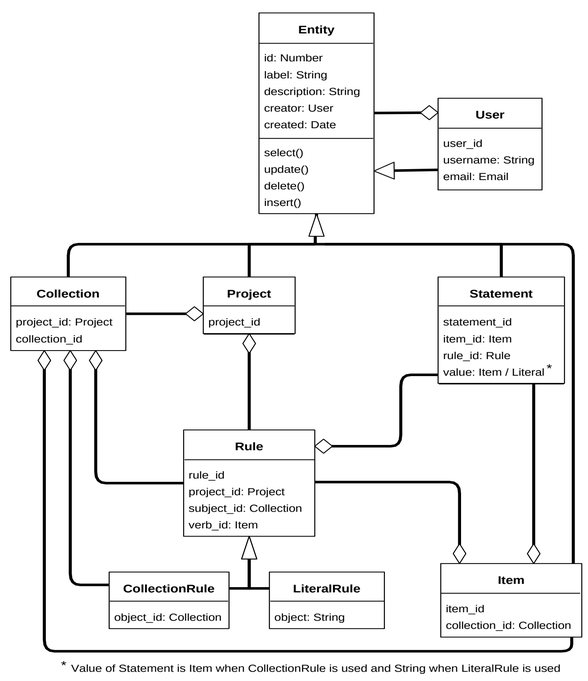
Challenges

Integrating the biological knowledge needed to identify cancer "driver" mutations remains a challenge due to:

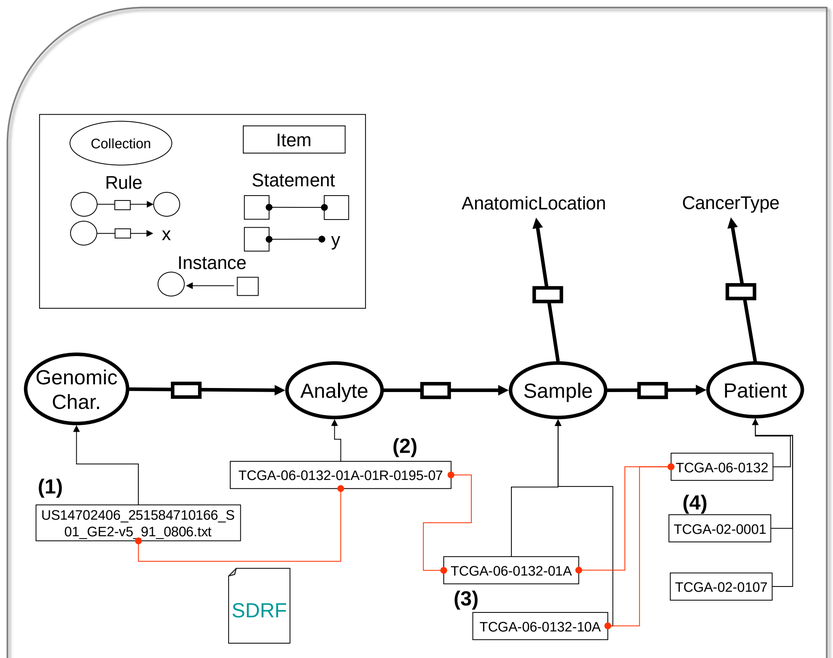
1) Heterogeneous experimental data
2) Distributed biomedical data sources
3) Controlled access clinical information

Methods

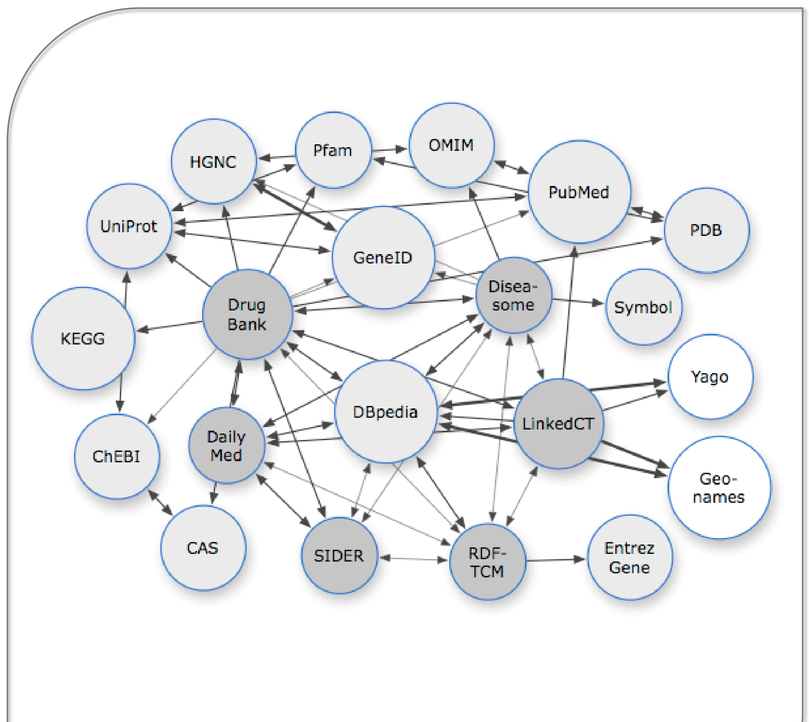
Our approach to help identify therapeutic targets for cancer required the interoperation between three components:



a) The S3DB object model was used as a knowledge organization system for experimental results.



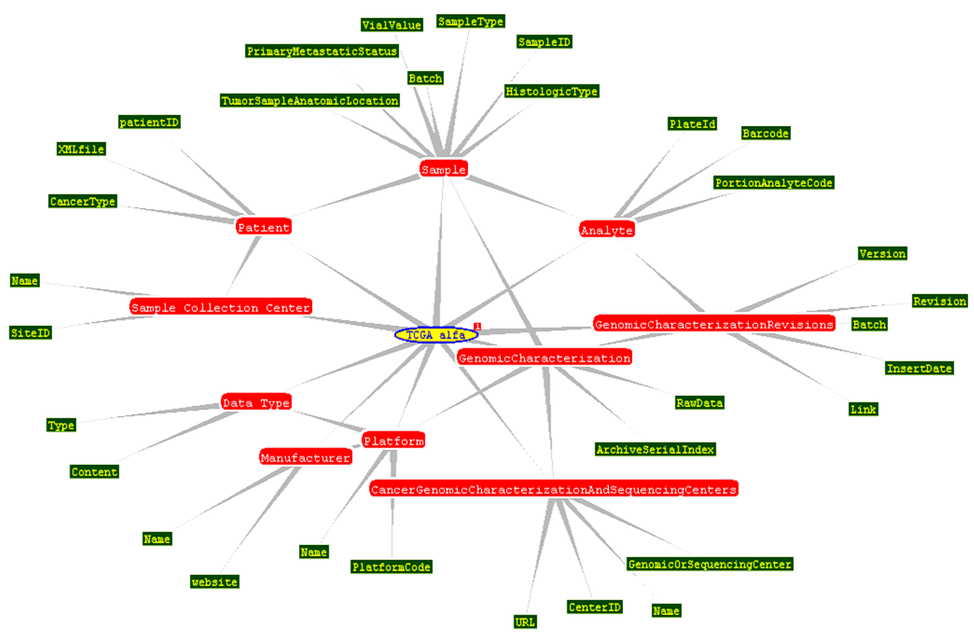
b) The TCGA (The Cancer Genome Atlas) data model was identified by domain experts using S3DB.



c) The Linked Open Drug Data (LODD) Cloud was used to link TCGA experimental results and biological assertions.

Results

Linking experimental results with biological knowledge in the LODD cloud enabled biology domain experts to easily identify relevant cancer "driver" mutations.




```
PREFIX diseasome: <http://www4.wiwiss.fu-berlin.de/diseasome/resource/diseasome/>
PREFIX : <http://ibl.mdanderson.org/TCGA/>
SELECT ?AnalysisLabel ?geneSymbol ?geneProduct ?AssociatedDisease WHERE {
  SERVICE <http://ibl.mdanderson.org/TCGA/sparql.php> {
    ?Analysis :R1057 ?AnalysisGenes .
    ?Analysis rdfs:label ?AnalysisLabel .
    ?AnalysisGenes :R1044 ?geneSymbol .
    ?AnalysisGenes :R1033 ?geneProduct .
  }
    ?Genes rdfs:label ?geneSymbol .
    ?DiseaseID diseasome:associatedGene ?Genes .
    ?DiseaseID diseasome:name ?AssociatedDisease .
}
```

Query S3DB/TCGA dataset

Query diseasome dataset

S3DB/TCGA query results

Diseasome query results

AnalysisLabel	geneSymbol	geneProduct	AssociatedDisease
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	NPHP4	nephronophthisis 4	Nephronophthisis
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	NPHP4	nephronophthisis 4	Senior-Loken_syndrome
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	NPHP4	nephronophthisis 4	Nephronophthisis 4, 606966
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	NPHP4	nephronophthisis 4	Senior-Loken syndrome 4, 606996
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	REN	renin	Hyperproreninemia
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	REN	renin	Renal_tubular_dysgenesis
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	REN	renin	Hyperproreninemia
Exploratory Analysis of the Copy Number Alterations in Glioblastoma Multiforme	REN	renin	Renal tubular dysgenesis, 267430

S3DB Model published in Public Library of Science (PLoS) One, BMC Bioinformatics and Nature Biotechnology
TCGA Ontology published in Journal of Biomedical Informatics

sfi Science Foundation Ireland
ENTERPRISE IRELAND
Ireland's EU Structural Funds Programmes 2007 - 2013
Co-funded by the Irish Government and the European Union
EUROPEAN REGIONAL DEVELOPMENT FUND
IRCSET Investing in People and Ideas
Irish Research Council for Science, Engineering and Technology
SEVENTH FRAMEWORK PROGRAMME

Enabling **Networked** Knowledge

Nature Precedings : doi:10.1038/npre.2011.6169.1 : Posted 29 Jul 2011