# Kanban:

# <https://kanbanflow.com/board/x5dJB8>

# Requirements for Workshop Server

[initial draft]

## User Accounts

ldwinstructor

* Or similar named account
* Installation and configuration (admin) privileges on the server.
* Admin access to Stardog DB

phuseldw

* Attendee user account.
* The same user account+password will be used by each attendee on their own, cloned
* Stardog: Data load, delete, query

## Server Details

OS and Software Applications

* Windows OS
* Minimum 2G RAM. Lowest service tier on AWS (1G RAM) is not sufficient.
* Each servers must be pingable from the other workshop servers, with port open to allow federated query from Stardog Studio between the servers.
* Stardog
* Stardog Studio
* R, RStudio, select R packages
* Windows Git, TortoiseGit , with ability to Git Clone, Update, Push..
* Python or other http server running for localhost with specified home folder.
  + Python added to PATH environment variable.
  + Python http server run at startup and points to the workshop home folder for localhost
* Chrome Browser (to be configured bookmarks)
* Apache Jena (riot, for checking TTL, or other similar tool)
* D3JS libraries
* Custom Graph Editor (JavaScript and associated JavaScript libraries)
* Tim will create the local folder structure used for the exercises, including the clone of materials from GitHub to the local folders.

Additional Configuration

* Remote desktop connection for user account: ' phuseldw '
* Ability to query DBPedia, Wikidata
* Ability to query to the other servers by IP address for federated query.
* Windows OS : Show all file extensions in the explorer view

# Other details

* 1 server needed now to setup as master and begin configuration and testing.
* 2 servers (master + test) needed in late April/Early May to test federated query and finish development of exercises.
* 10 servers needed ~1 week in advance of the session to allow users to test login.
* ~22 servers needed for event. (1 master, 1 instructor, ~20 attendees)
  + Servers for the event must be setup a minimum of 2 days prior to allow testing and listing of IP addresses on handout materials.
  + Attendee servers can be shut down immediately following the session.

Attendees will login to their cloned server using the phusldw account and the IP address of their cloned AWS server.

# Stardog Setup

Add all relevant namespaces to the LDWStudy database.

Include;

schema:, eg: , dbpedia: etc.

# Graph Editor Notes

Nodes at the start

schema:Person

schema:Female

schema:Male

cto:PLACEBO

cto:ACTIVE

dbpedia:Aspirin

with dbpedia:Aspirin already attached to the ACTIVE arm.

link to NCT ID should be predicate:

<http://bio2rdf.org/clinicaltrials_vocabulary:nct-id>

or:

PREFIX ct-voc: <http://bio2rdf.org/clinicaltrials\_vocabuhaselary:>

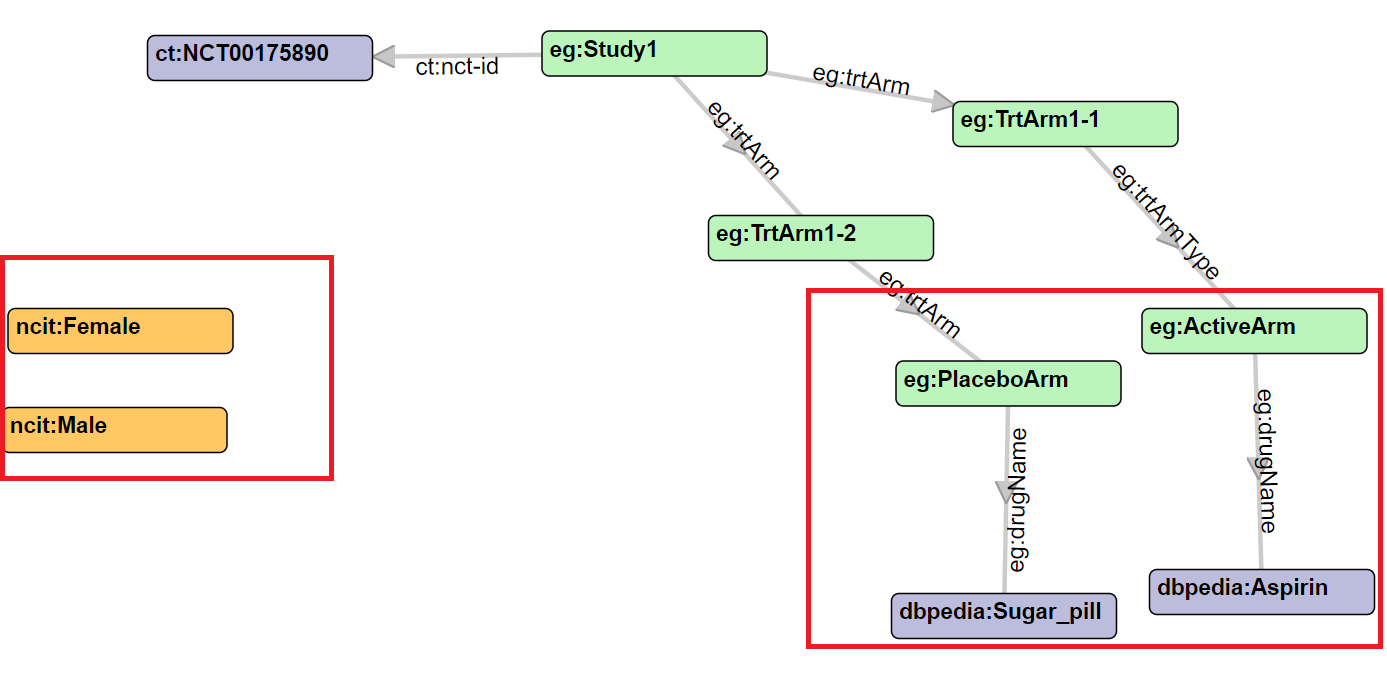
Resulting in:

Foo ct-voc:ncit-id <ID>

TODO:

Consider changing from use of eg: prefix to drugco: or some other.phase

Default graph has the nodes in red:



# MISC

The prefixes in the export from the graph editor must end in a / , as in:

@prefix ct: <http://bio2rdf.org/clinicaltrials/>

To give:

eg:Study1 ct:nct-id ct:NCT00001111

but, can this ct value then be used to query biordf end point????

POSSIBLE FIX.

Go here:

http://bio2rdf.org/clinicaltrials:NCT00175890

Use this instead:

http://identifiers.org/clinicaltrials/NCT00175890>

identifiers.org URI http://identifiers.org/clinicaltrials/NCT00175890

From there go through the same PHASE path to extract Phase.

salicylate

**2019-04-09**

Notes about Wikidata:

Property:P32 Q30612

(instance of) (clinical trial)

From predicates and objects for things that are clinical trials.

SELECT \* WHERE {

?item **wdt:P3098** ?ctID ;

?p ?o .

}LIMIT 100

Find info about a specific Aurora trial:

SELECT \* WHERE {

wd:Q4654994 ?p ?o ;

}LIMIT 100

# Query Notes

Federated Query

See here:

https://stackoverflow.com/questions/45304731/sparql-federated-queries-using-uri-returned-from-a-query-in-another-query

SELECT distinct ?writer ?play ?character ?uri ?type WHERE {

{

?writer lit:writerOf ?play .

?play lit:character ?character .

?character lit:dbpediaUri ?uri .

**BIND (IRI(?uri) AS ?real\_uri)**

}

SERVICE <http://dbpedia.org/sparql> {

**?real\_uri** rdf:type ?type

}

}

WHERE{

{ # My local graph

?studyIRI ct:nct-id ?nctid .

BIND (?nctid AS ?nctidIRI)

}

SERVICE <http://clinicaltrials.bio2rdf.org/sparql>

{

?nctidIRI dcterms:title ?studyTitle ;

ct-voc:phase ?phaseIRI .

?phaseIRI ?p ?phaseName .

FILTER regex(str(?phaseName), "^Phase \\d$", "i")

}

} # End of WHERE statement

PREFIX ctid: <http://identifiers.org/clinicaltrials/>

PREFIX ct-voc: <http://bio2rdf.org/clinicaltrials\_vocabulary:>

PREFIX dcterms: <http://purl.org/dc/terms/>

SELECT ?studyTitle ?phaseName

WHERE{

{ # My local graph

?studyIRI ct:nct-id ?nctid .

BIND (?nctid AS ?nctidIRI)

}

SERVICE <http://clinicaltrials.bio2rdf.org/sparql>

{

{

?s ?p ctid:NCT00175890 ;

dcterms:title ?studyTitle ;

<http://bio2rdf.org/clinicaltrials\_vocabulary:phase> ?phaseIRI .

?phaseIRI dcterms:title ?phaseName .

}

LDWorkshop

Can you determine Phase from the NCT graph?

To find phase:

http://bio2rdf.org/clinicaltrials:NCT00175890

|  |  |
| --- | --- |
| [phase [clinicaltrials\_vocabulary:phase]](http://bio2rdf.org/describe/?url=http%3A%2F%2Fbio2rdf.org%2Fclinicaltrials_vocabulary%3Aphase) | * [Phase 3 [clinicaltrials\_resource:8357418e2694434468870b487644532d]](http://bio2rdf.org/describe/?url=http%3A%2F%2Fbio2rdf.org%2Fclinicaltrials_resource%3A8357418e2694434468870b487644532d) |

Query the clinical trials graph to get:

1. Study Title
2. Study Phase

Query the drugbank graph to get:

1. Drug description

All the studies are on Aspirin. So include that in the graph at the start and link to it.

Find clinical trial Phase

SELECT ?phase ?o

WHERE

{

<http://bio2rdf.org/clinicaltrials:NCT00175890> <http://bio2rdf.org/clinicaltrials\_vocabulary:phase> ?phase .

?phase ?p ?phaseName .

}

This query:

SELECT ?phase ?phaseName

WHERE

{

<http://bio2rdf.org/clinicaltrials:NCT00175890> <http://bio2rdf.org/clinicaltrials\_vocabulary:phase> ?phase .

?phase ?p ?phaseName .

FILTER regex(str(?phaseName), "^Phase \\d$", "i")

}

Wikidata: Aspirin

https://www.wikidata.org/wiki/Q18216

test at: <https://dbpedia.org/sparql>

Prefixes in dbpedia:

<http://eo.dbpedia.org/sparql?nsdecl>

prefix dbpedia-owl: <http://dbpedia.org/ontology/>

SELECT \*

WHERE{

dbpedia-owl:Drug rdfs:label ?drugName

}LIMIT 10

FILTER ON THIS to ASPIRIN

----

<https://www.wikidata.org/wiki/Q4654994> - From the paper by KF

Wikidata Identifier for it. Immutabbility .

ClinicalTrials.gov + EudraCT + ….?

<https://www.wikidata.org/wiki/Q4654994>

All Clinical Trials in Wikidata

SELECT \* WHERE {

?item wdt:P3098 ?ctID .

}

Aspirin Studies

* One assigned phase.
* Recruiting

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | NCTID | Sex All | Aspirin | Placebo |
| 1 | NCT02906761 | Y | Y | Y |
| 2 | NCT03405974 | Y | Y | Y |
| 3 | NCT02301286 | Y | Y | Y |
| 4 | NCT02467582 | Y | Y | Y |
| 5 |  |  |  |  |
| 6 | NCT02813824 | Y | Y | Y |
| 7 | NCT03356769 | Y | Y | Y |
| 8 | NCT03079999 | Y | Y | Y |
| 9 |  | Y |  |  |
| 10 |  | Y |  |  |
| 11 | NCT02647099 | Y | y | y |
| 12 |  | Y |  |  |
| 13 |  | Y |  |  |
| 14 |  | Y |  |  |
| 15 |  | Y |  |  |
| 16 |  | Y |  |  |
| 17 |  |  |  |  |
| 18 |  | Y |  |  |
| 19 |  | Y |  |  |
| 20 |  | Y |  |  |
| 21 |  | Y |  |  |
| 22 |  |  |  |  |
| 23 |  | Y |  |  |
| 24 |  | Y |  |  |
| 25 |  | Y |  |  |
|  |  | Y |  |  |