BrAPI2RDF - Converting between BrAPI and RDF

Cyril Pommier - Pierre Larmande

Use case Wheat Phenotyping dataset

Phenotypes

Winter wheat (Triticum aestivum L) phenotypic data from the mul
of the INRA Small Grain Cereals Network.

François-Xavier Oury, Emmanuel Heumez, Bernard Rolland, Jérôme Auzanneau, Pierre
Xavier Charrier, Hubert Chiron, Camille Depatureaux, Laurent Falchetto, Olivier Gardet,
Christophe Lecomte, Jean-Yves Morlais, Pierre Pluchard, Didier Tropée, Maxime Trotte
Michel Rousset, Gilles Charmet

Query dataset as a semantic graph.

Or download the dataset as RDF archive.

Abstract

Published 2018 by Portail Data Inra

Back to Form

Search parameter(s):

Geolocation

DATA SETS: 4

Network Data Set :
INRA Wheat Network technologica

Network Data Set :
INRA Small Grain Cereals Network
DOI:http://dx.doi.org/10.15454/1.4489666216568333E12

Network Data Set :
INRA Wheat Network BRC accession (A series)

Trial list | Phenotypic data

Data table view : Default

LEVEL: REPLICATION

1-10 of 81,434 replication | Display 10 results per page

GENOTYPE ID			TREATMENT		
Lot Number	Accession Number	Accession Name	itk	Trial Name	Trial Site
CF99016	29114	CF99016	treated	BTH_Dijon_2000_SetA	Dijon
RE99002	29191	RE99002	treated	BTH_Dijon_2000_SetA	Dijon
SOISSONS	6607	SOISSONS	low inputs	BTH_Dijon_2000_SetA	Dijon
RE9819	37778	RE9819	untreated, no fungicide	BTH_Le_Moulon_2000_SetA	Le Moulon
CF99016	29114	CF99016	treated	BTH_Le_Moulon_2000_SetA	Le Moulon
CF99031	22479	CF99031	low inputs	BTH_Le_Moulon_2000_SetA	Le Moulon

VARIABLE	VARIABLE	VARIAB
rdt: Grain yield at 0% humidity	r: Yield	prec: Earliness
73,8	105,6	
66	120,5	
110,6	142,9	
106,5	102,4	
41,6	104,8	
39,9	75,5	141
64,6	115,4	139

Why RDF for BrAPI?

- Easier integration across sources
  - Germplasm linking at different levels (varieties, accession, hybrid)
  - Linking to literature data through traits
  - Linking/integrating and subseting accross different sources/dataset
  - …
- More complex “chained” queries are easier to perform on graphs
- Validation (model-, type- and value-level)
- One workflow/tool for all BrAPI compliant databases.
- ⟶ Let’s lift BrAPI JSON and add JSON-LD context

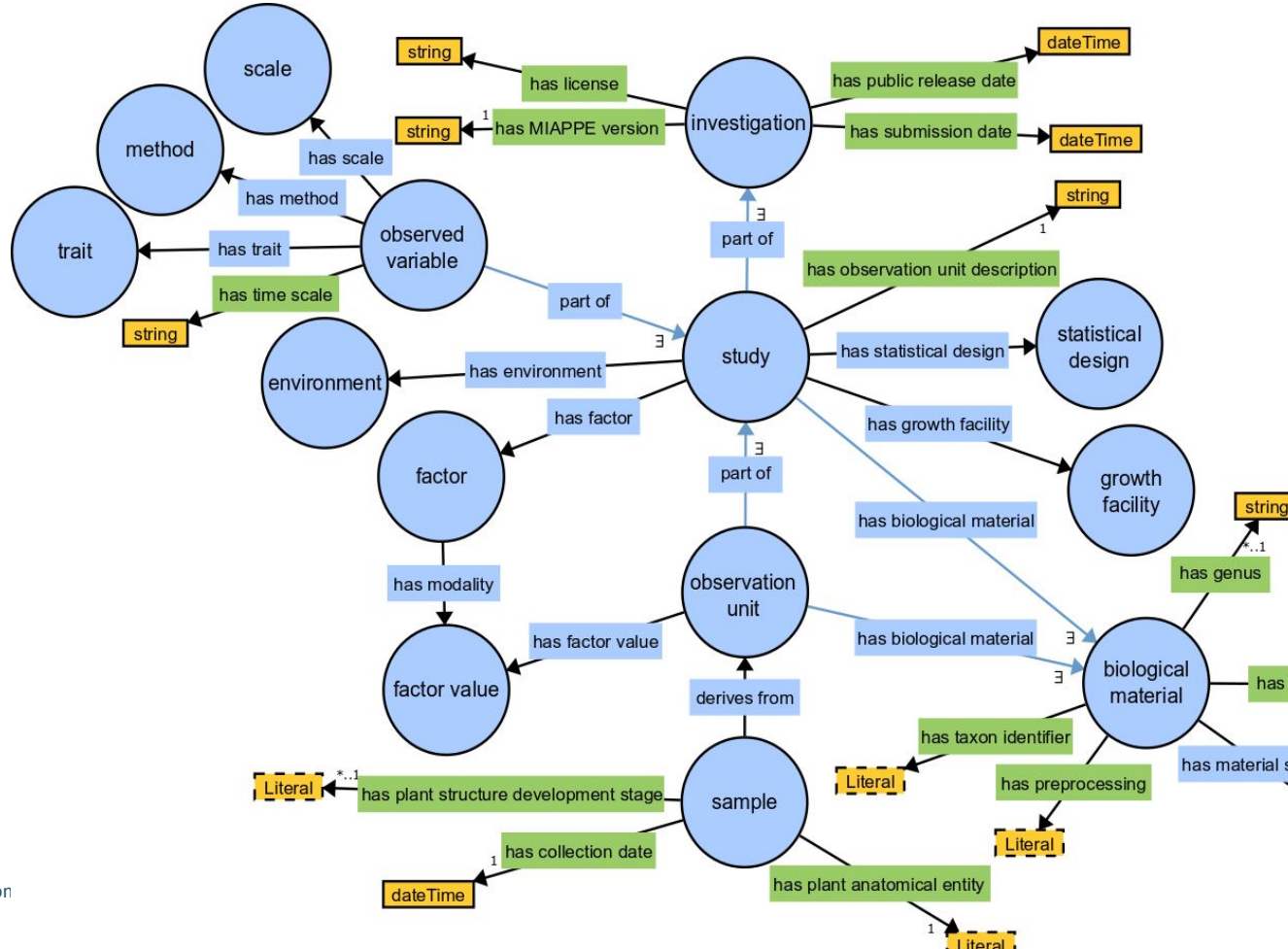
Ontologies

BrAPI types

- "http://brapi.org/rdf/Study"

MIAPPE Ontology

- Plant Phenotyping Experiment Ontology
-
-



JSON-LD basics

A JSON compatible format to map JSON data to RDF triplets

```
{   "@context":
"http://brapi.org/rdf/StudyGermplasm.jsonld",

  "studyDbId":"123",

  "trialName":"myBestTrial",
  "germplasm":[ {
        "germplasmPUI":"http://www.crop-dive
rsity.org/mgis/accession/01BEL084609",
        "germplasmDbId":"382",
        "germplasmName":"Pahang",
        "accessionNumber":"ITC0609",
       ...
    } ]
}
```

```
{
  "@context":
"http://brapi.org/rdf/StudyGermplasm.jsonld",
  "studyPUI":"http://.../studyABC",
  "studyDbId": "123",
  "trialPUI":"http://.../trials#trialX",
  "trialName": "myBestTrial",
  "data":[ {
        "germplasmPUI":"http://www.crop-diversity.or
g/mgis/accession/01BEL084609",
        "germplasmDbId":"382",
        "germplasmName":"Pahang",
        "accessionNumber":"ITC0609",
       ...
    } ]
}
```

JSON-LD basics - context

Mapping and result for Study

```
{
  "@context":
"http://brapi.org/rdf/StudyGermplasm.jsonld",
  "studyPUI":"http://.../studyABC",
  "studyDbId": "123",
  "trialPUI":"http://.../trials#trialX",
  "trialName": "myBestTrial",
  "germplasms":[ {
        "germplasmPUI":"http://www.crop-dive
rsity.org/mgis/accession/01BEL084609",
        "germplasmDbId":"382",
        "germplasmName":"Pahang",
        "accessionNumber":"ITC0609",
       ...
  } ]
}
```

```
{
  "@context": {
    "xsd": "http://www.w3.org/2001/XMLSchema#",
    "brapi": "https://brapi.org/rdf/",

    "Study": "brapi:Study",
    "Trial": "brapi:Trial",
    "Location": "brapi:Location",
    "studyPUI": "@id",
    "studyDbId": "brapi:hasDatabaseIdentifier",
    "studyName": "brapi:hasName",
    "trialPUI": {
      "@id": "brapi:hasTrial",
      "@type": "@id"
    },
    "germplasmPUI": {
      "@id": "brapi:hasGermplasm",
      "@type": "@id"
    }
  }
}
```

BrAPI JSON-LD: ETL approach

Example: github.com/gnpis/BrAPI-extract-index-prototype

1) Extract BrAPI data
2) Transform
   a) URI generation where missing (this creates objects link by URI)
   b) Add JSON-LD context (mapping ontology properties)
   c) Convert to RDF format
3) Load into a triplestore

URI generation

template: {brapi server}/brapi/v1/{brapi object}/{database id}
example URI: https://urgi.versailles.inra.fr/GnpISCore-srv/brapi/v1/study/E012E0

Data lifting

BrAPI calls for individual resources

OK With BrAPI V2

- Trial
- Study `/studies/{studyDbId}`
- Germplasm
- ObservationVariable
- ObservationUnits

Data Lifting

Links between resources

Studies → Germplasm & others

Not natively present in BrAPI JSON (anymore ?)

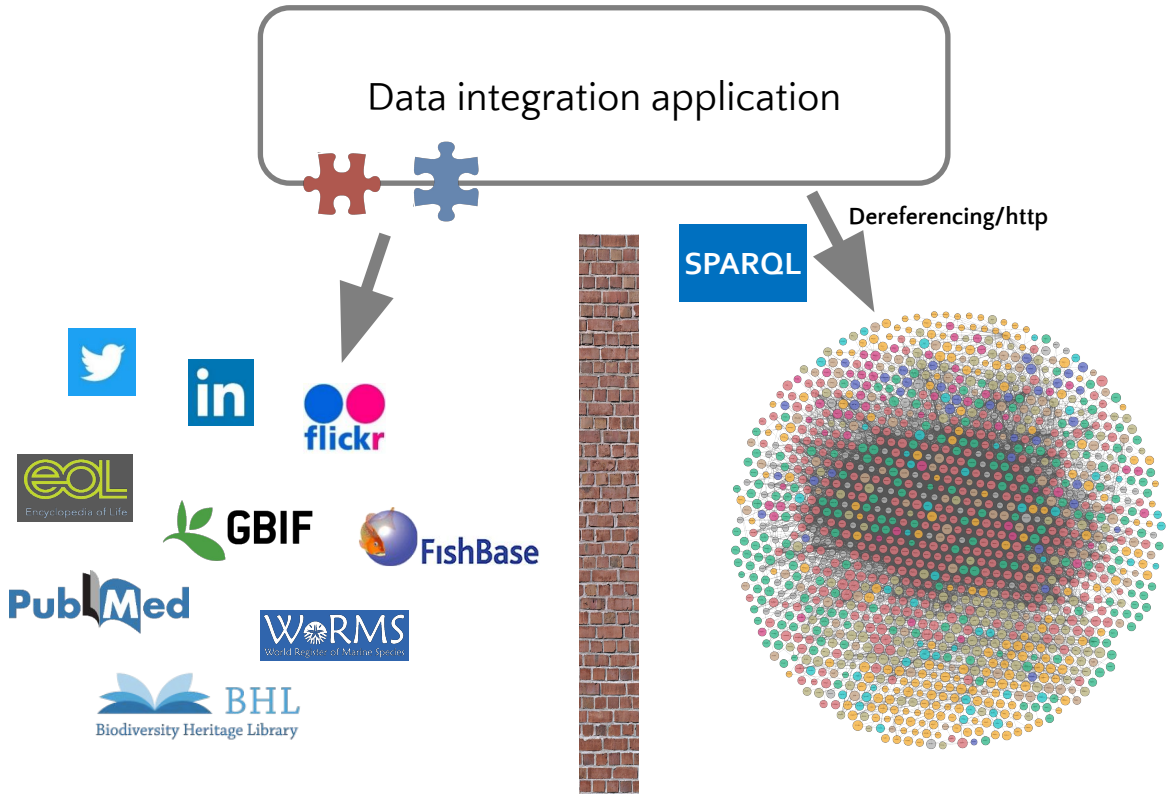
To be added to exchange a whole trial/dataset ?

→ File Archive (RO Crate) Use case

SPARQL Micro-Services
Querying Web APIs with SPARQL

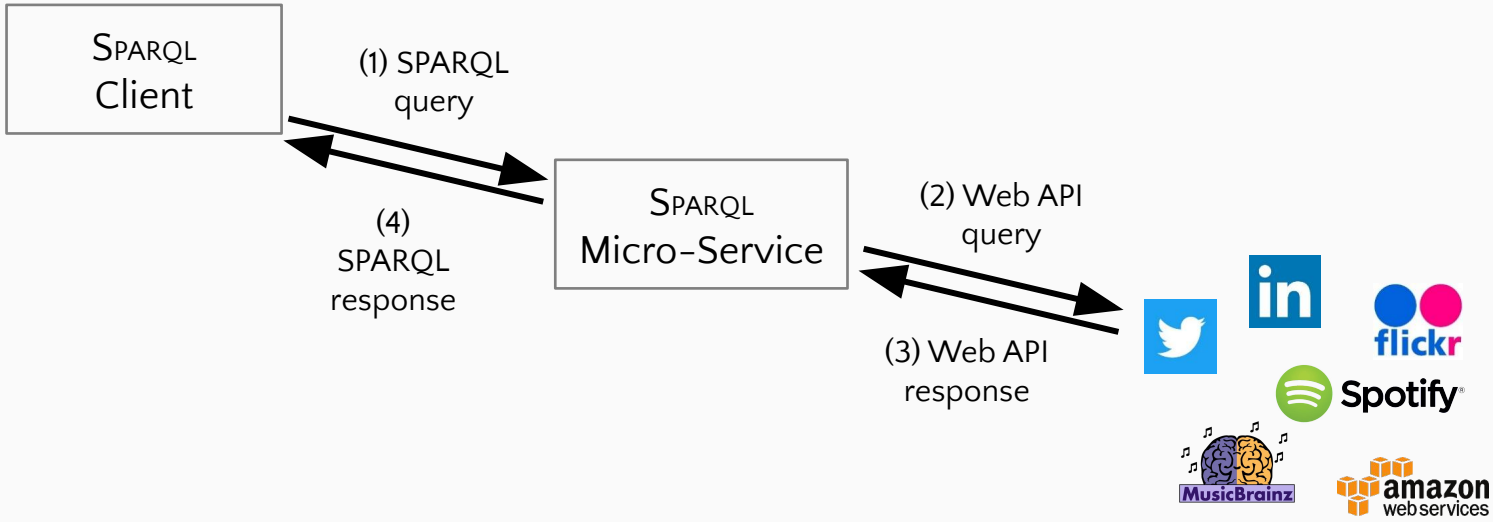
Courtesy of Franck MICHEL

How to leverage Web APIs and Linked Data at the same time?



The SPARQL Micro-Service Architecture

Lightweight method to **query a Web API with SPARQL**
(*and assign dereferenceable URIs to Web API resources)*



A SPARQL µ-service is a configurable SPARQL endpoint whose arguments delineate the graph being queried.

Endpoint: http://example.org/flickr/getPhotosByTag?tag=bridge

```
SELECT * WHERE {
    ?photo a schema:Photograph;
      schema:name       ?title;
      schema:contentUrl ?img.
}
```

Arguments passed as HTTP parameters

Endpoint: http://example.org/flickr/getPhotosByTag_**sd**

```
SELECT * WHERE {
    ?photo a schema:Photograph;
      schema:keywords   "bridge";
      schema:name       ?title;
      schema:contentUrl ?img.
}
```

Arguments passed in the graph pattern

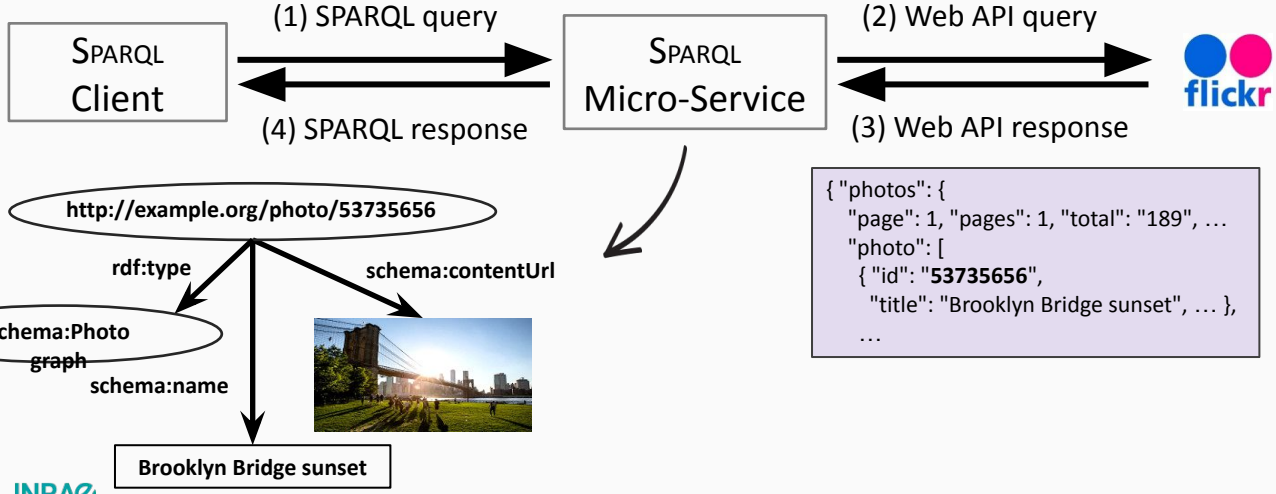
Into the details - Arguments passed as HTTP parameters

```
SELECT * WHERE {
  ?photo a schema:Photograph;
    schema:name       ?title;
    schema:contentUrl ?img.
}
```

http://example.org/flickr/getPhotosByTag?tag=bridge

https://api.flickr.com/services/rest/?method=flickr.photos.search&

format=json&per_page=100&tags=bridge&...

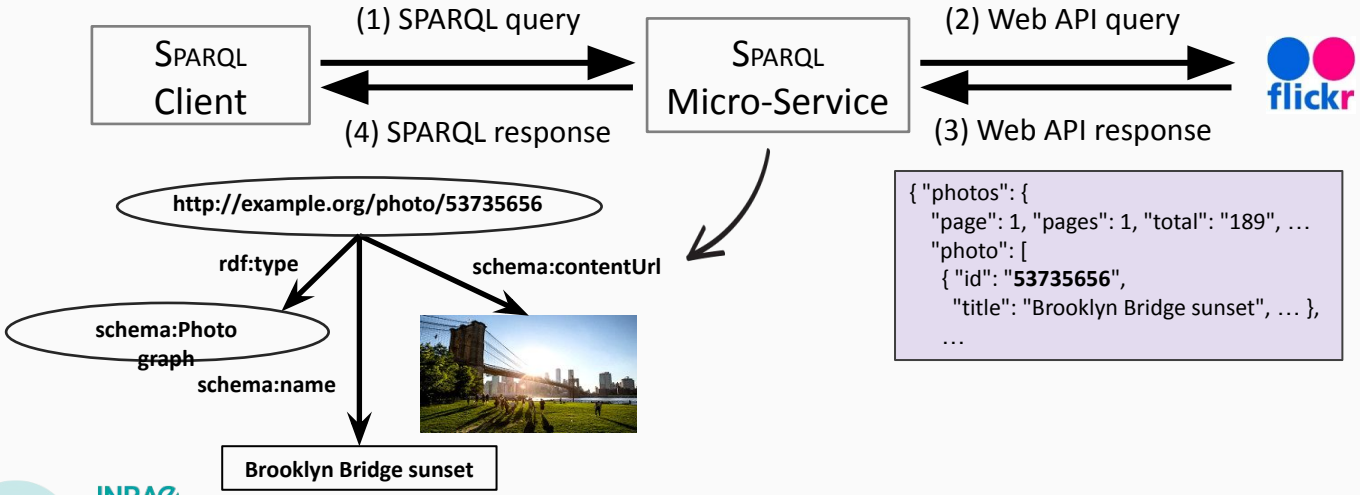


Into the details - Arguments passed in the graph pattern

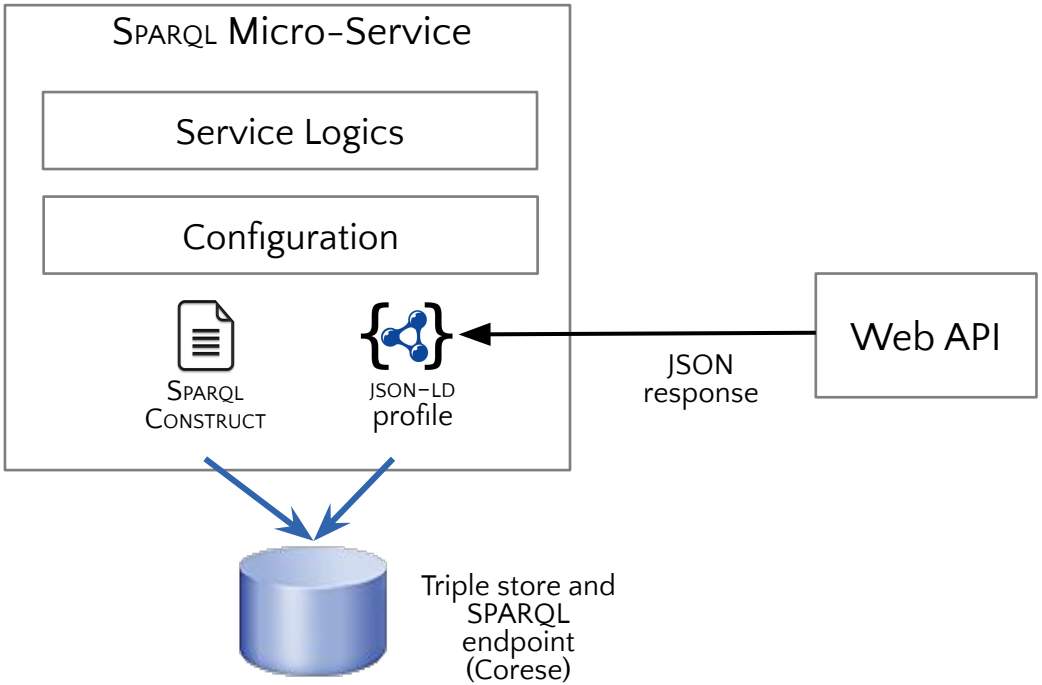
```
SELECT * WHERE {
  ?photo a schema:Photograph;
    schema:keywords   "bridge";
    schema:name       ?title;
    schema:contentUrl ?img.
}
```

http://example.org/flickr/getPhotosByTag_sd

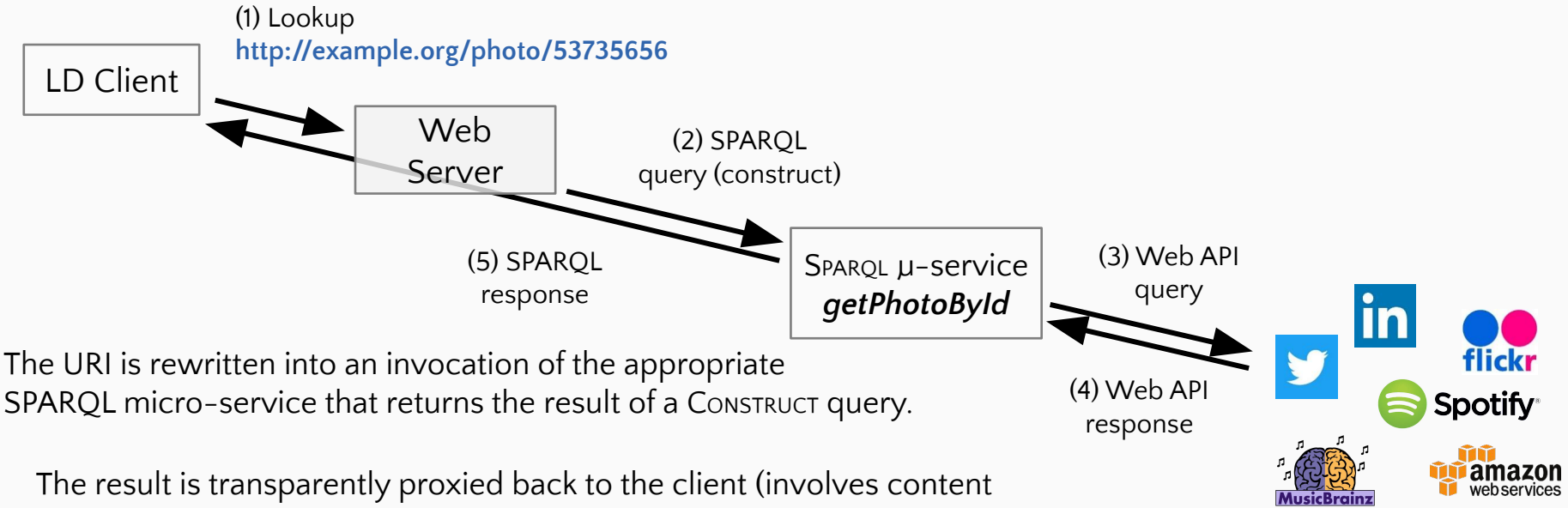
https://api.flickr.com/services/rest/?
method=flickr.photos.search&
format=json&per_page=100&tags=bridge&...



Translating the Web API response to RDF



Bridging Web APIs and the Web of Data

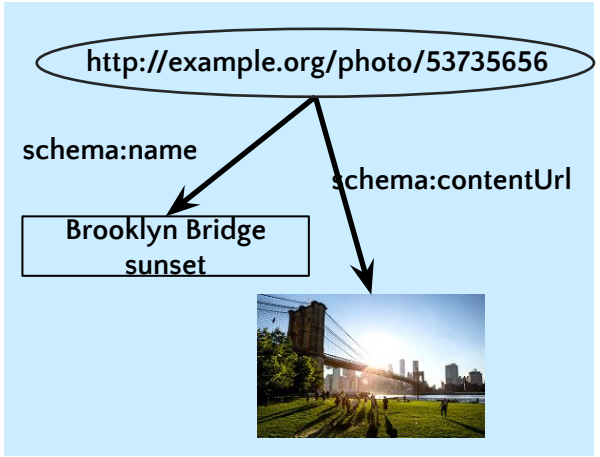


The URI is rewritten into an invocation of the appropriate SPARQL micro-service that returns the result of a CONSTRUCT query.

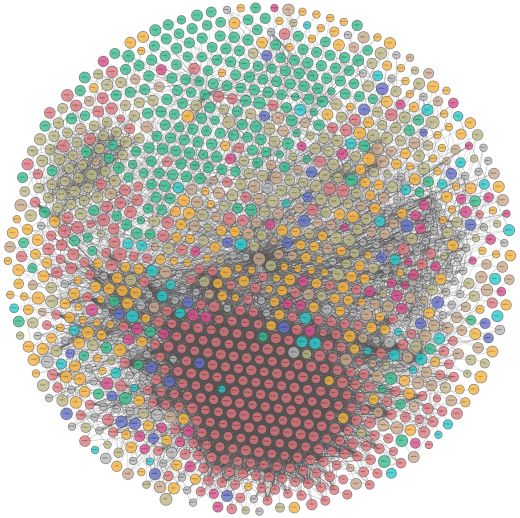
The result is transparently proxied back to the client (involves content negotiation).

Bridging Web APIs and the Web of Data

SPARQL μ-service
getPhotosByTag



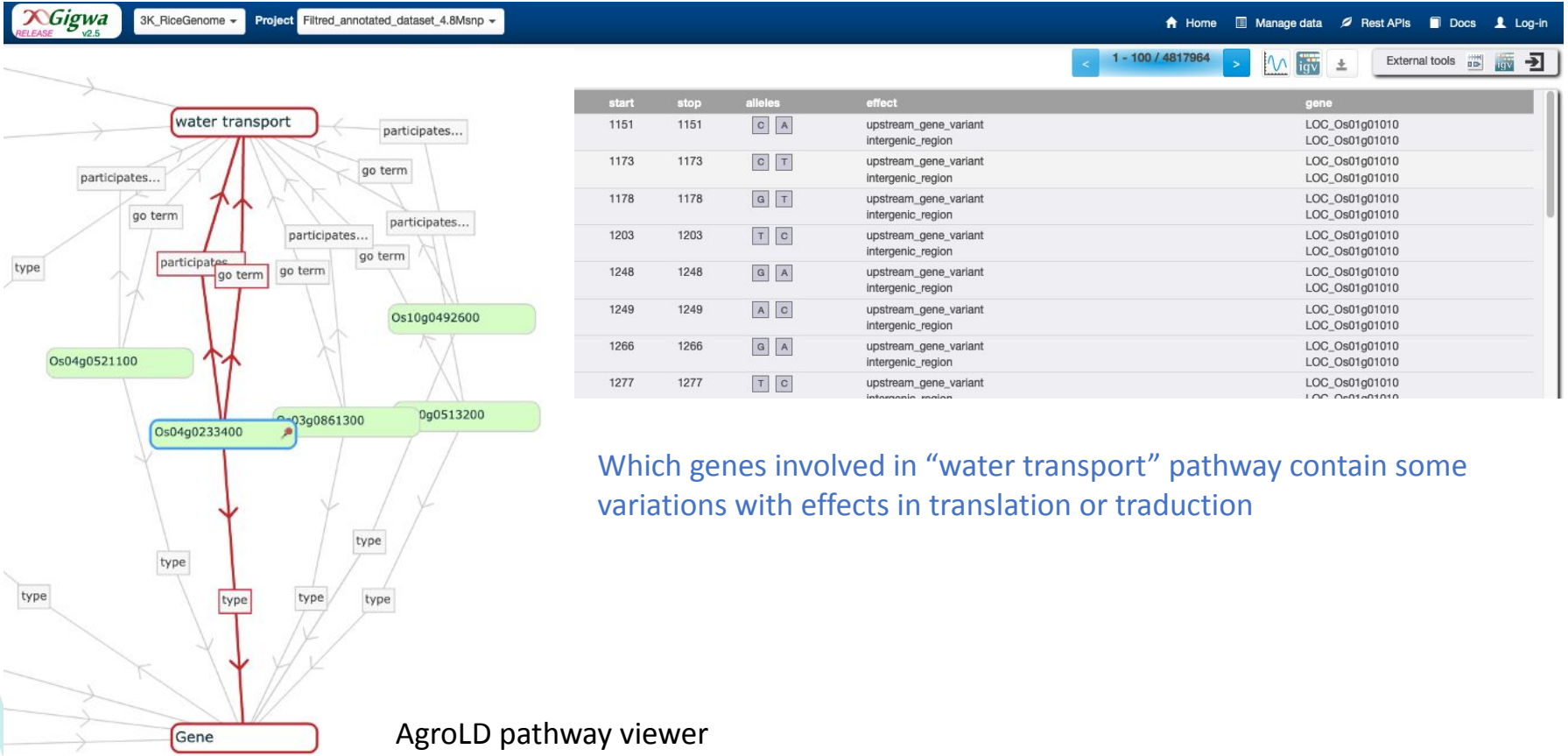
unlock



Assign dereferenceable URIs to Web API resources

Expose into the Web of Data resources previously locked in a silo

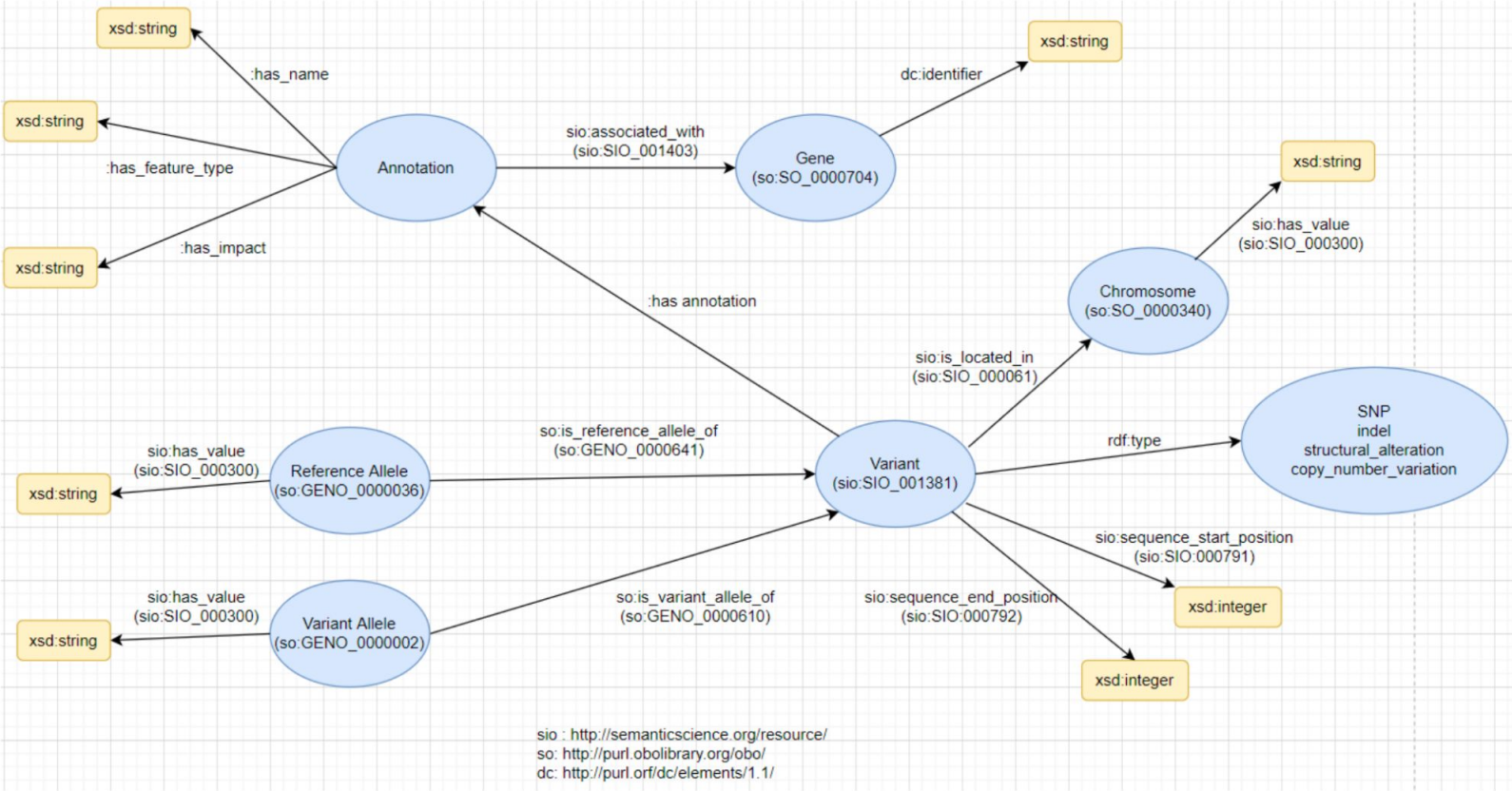
Use case with Gigwa and AgroLD



Which genes involved in “water transport” pathway contain some variations with effects in translation or traduction

AgroLD pathway viewer

RDF(s) model for Gigwa



Example retrieving variant from Gigwa

https://gigwa.ird.fr/gigwa/rest/brapi/v2/variants?end=1152&referenceDbId=3K_RiceGenom e§1&start=1000

start position : 1000

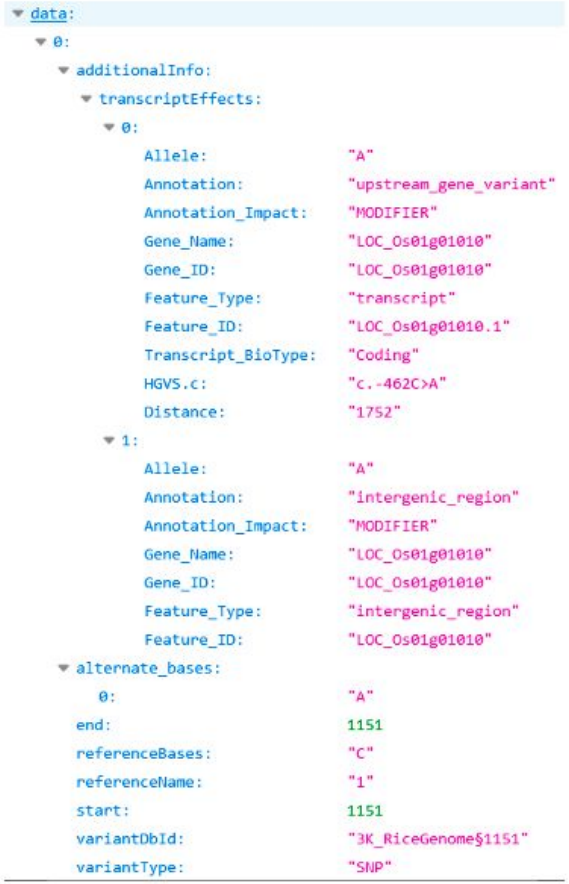
end position : 1152

referenceDbId: 3K_RiceGenome§1

2 micro-services were created

getVariants : *get all variants between start and end pos on a chr*

getVariantsWithAnnotations : *get annotations from a variant*



Thanks for your attention