



Importing and Accessing Data from Spreadsheets and Databases

Josef Hardi

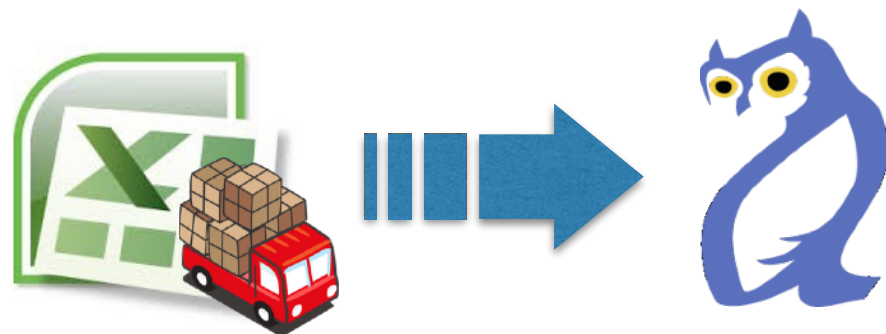
(josef.hardi@stanford.edu)

Protégé Short Course Demo

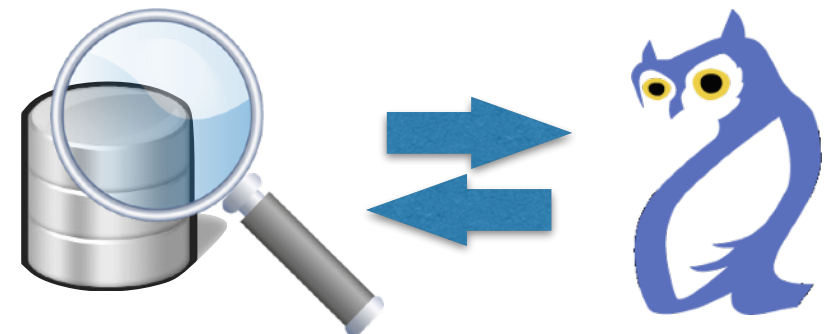
March 31, 2017

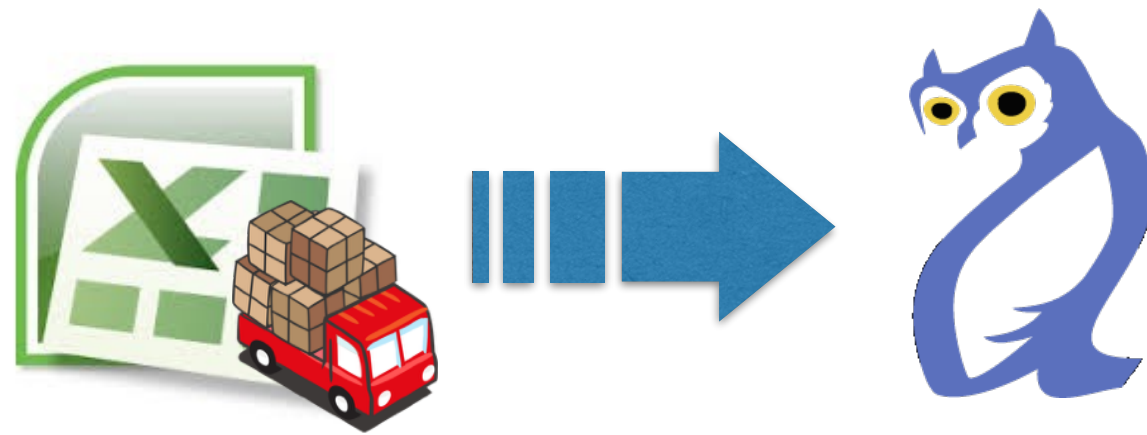


Data Import (Cellfie Plugin)



Data Access (Ontop Plugin)

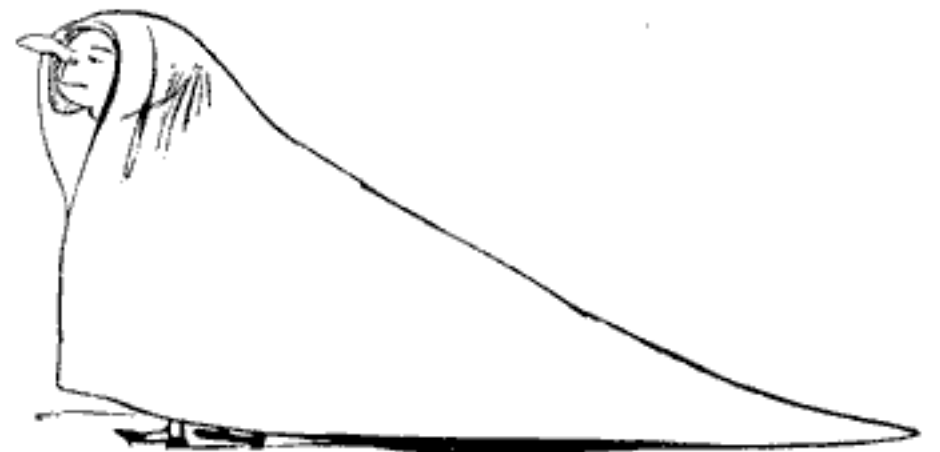




Data Import

Motivation

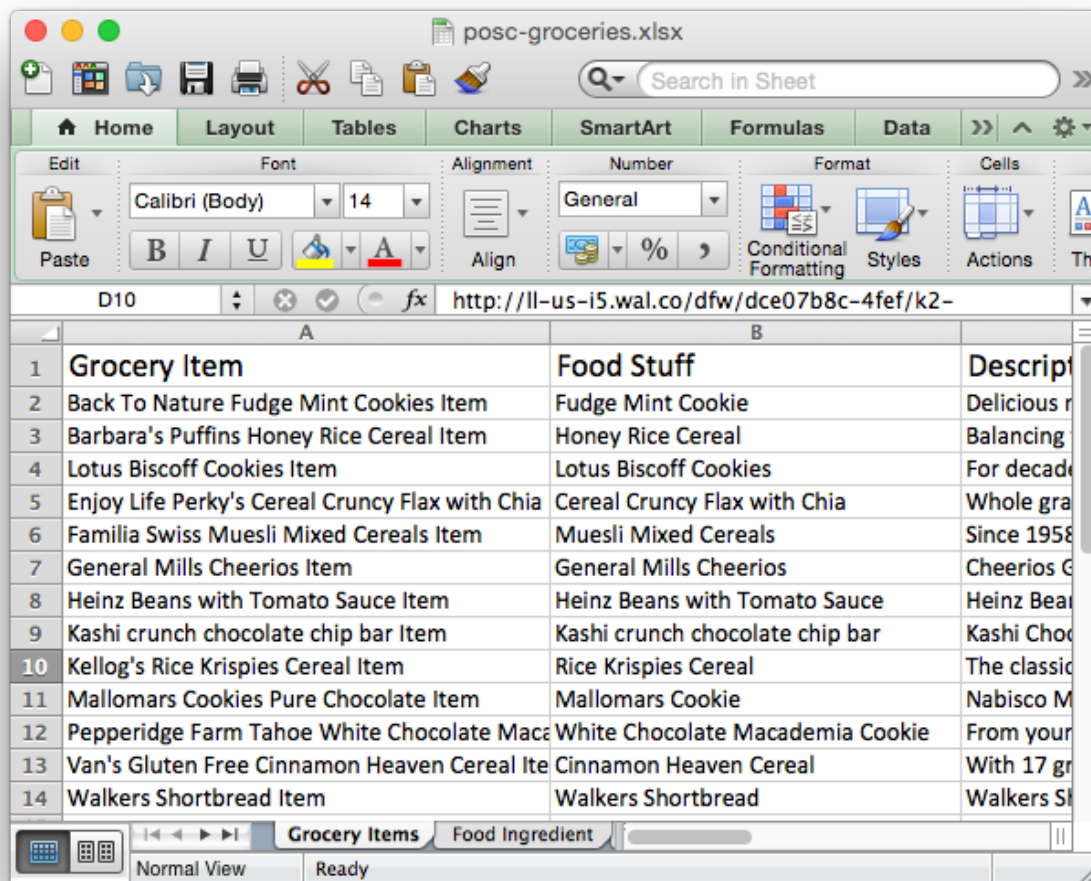
- Very often people store and share knowledge in spreadsheets.
- Benefits of spreadsheets:
 - High efficiency,
 - Flexible formatting,
 - Immediate visualization,
 - Very portable,
 - Everybody knows them!



Cellfie Plugin

- The Cellfie plugin is useful for importing ontology structure and instances from Excel spreadsheets.
- The application uses a robust mapping standard to translate tabular data into OWL axioms.
- The plugin is maintained by the Protégé team.

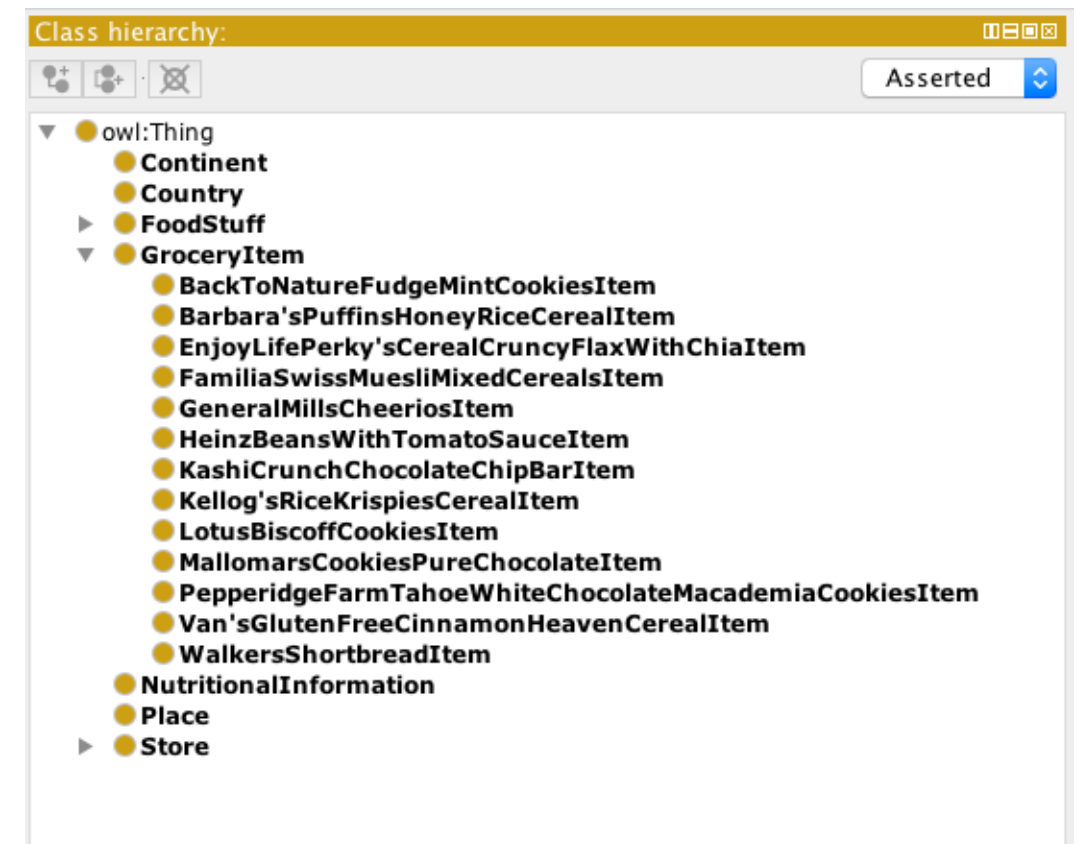

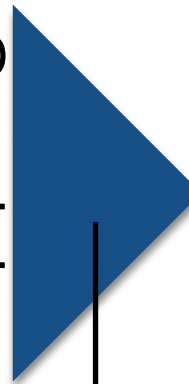
Preview



The screenshot shows a spreadsheet application window titled 'posc-groceries.xlsx'. The spreadsheet has three columns: 'Grocery Item', 'Food Stuff', and 'Description'. It contains 14 rows of data, including items like 'Back To Nature Fudge Mint Cookies Item', 'Barbara's Puffins Honey Rice Cereal Item', and 'Lotus Biscoff Cookies Item'.

Grocery Item	Food Stuff	Description
Back To Nature Fudge Mint Cookies Item	Fudge Mint Cookie	Delicious r
Barbara's Puffins Honey Rice Cereal Item	Honey Rice Cereal	Balancing
Lotus Biscoff Cookies Item	Lotus Biscoff Cookies	For decade
Enjoy Life Perky's Cereal Cruncy Flax with Chia	Cereal Cruncy Flax with Chia	Whole gra
Familia Swiss Muesli Mixed Cereals Item	Muesli Mixed Cereals	Since 1958
General Mills Cheerios Item	General Mills Cheerios	Cheerios G
Heinz Beans with Tomato Sauce Item	Heinz Beans with Tomato Sauce	Heinz Bear
Kashi crunch chocolate chip bar Item	Kashi crunch chocolate chip bar	Kashi Choc
Kellog's Rice Krispies Cereal Item	Rice Krispies Cereal	The classic
Mallomars Cookies Pure Chocolate Item	Mallomars Cookie	Nabisco M
Pepperidge Farm Tahoe White Chocolate Maca	White Chocolate Macademia Cookie	From your
Van's Gluten Free Cinnamon Heaven Cereal Ite	Cinnamon Heaven Cereal	With 17 gr
Walkers Shortbread Item	Walkers Shortbread	Walkers Sh

mappings

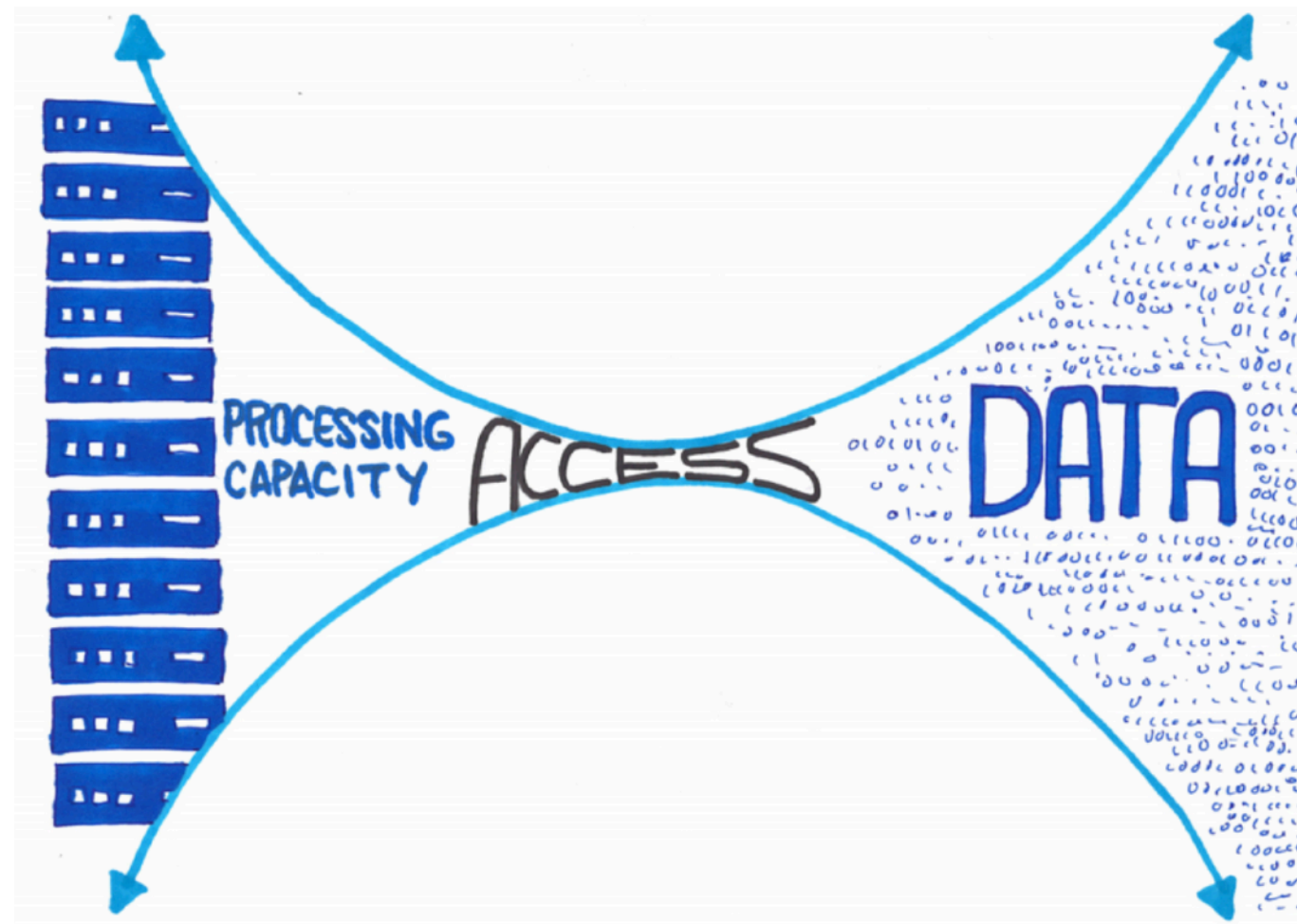


MappingMaster
DSL



Data Access

Motivation



- **Data access bottleneck:** only the skillful personnel own the access.
- **Data integration problem:** different data presentations but the meaning is the same.

Ontop Plugin

- The Ontop plugin is useful for querying data stored in SQL databases by using the vocabularies of an ontology.
- The system is based on the concept of Ontology-based Data Access (OBDA).
- The plugin is maintained by the KRDB Research Center in Italy.

Preview

The image shows a preview of a Protege ontology editor interface. It consists of two main windows: 'groceries' and 'grocery-ontology'.

groceries (http://protege.stanford.edu/ontologies/groceries) : [Users/jhardi/Documents/Work/Short Course Mar-2017/Ontop Demo/_wiki/posc-groceries.owl]

Active Ontology x Entities x Individuals by class x DL Query x Ontop SPARQL x Ontop Mappings x

Datasource manager Mapping manager Mapping Assistant - BETA

Mapping editor:

Mapping manager

+ Create - Remove Copy

Walkers Shortbread

{UPCA} a :WalkersShortbreadItem ; :productName {GROCERY_ITEM} ; :price {PRICE}^^xsd:decimal ; :isSoldByStore {STORE_ID} .

select * from CATALOG where GROCERY_ITEM = 'Walkers Shortbread'

Familia Swiss Muesli

{UPCA} a :FamiliaSwissMuesliMixedCerealsItem ; :productName {GROCERY_ITEM} ; :price {PRICE}^^xsd:decimal ; :isSoldByStore {STORE_ID} .

select * from CATALOG where GROCERY_ITEM = 'Familia Swiss Muesli Mixed Cereals'

General Mills Cheerios

{UPCA} a :GeneralMillsCheeriosItem ; :productName {GROCERY_ITEM} ; :price {PRICE}^^xsd:decimal ; :isSoldByStore {STORE_ID} .

select * from CATALOG where GROCERY_ITEM = 'General Mills Cheerios'

Kashi Crunch Chip Bar

{UPCA} a :KashiCrunchChocolateChipBarItem ; :productName {GROCERY_ITEM} ; :price {PRICE}^^xsd:decimal ; :isSoldByStore {STORE_ID} .

select * from CATALOG where GROCERY_ITEM = 'Kashi Crunch Chocolate Chip Bar'

Mallomars Cookies

{UPCA} a :MallomarsCookiesPureChocolateItem ; :productName {GROCERY_ITEM} ; :price {PRICE}^^xsd:decimal ; :isSoldByStore {STORE_ID} .

select * from CATALOG where GROCERY_ITEM = 'Mallomars Cookies Pure Chocolate'

Pepperidge Macademia Cookies

{UPCA} a :PepperidgeFarmTahoeWhiteChocolateMacademiaCookiesItem ; :productName {GROCERY_ITEM} ; :price {PRICE}^^xsd:decimal ; :isSoldByStore {STORE_ID} .

select * from CATALOG where GROCERY_ITEM = 'Pepperidge Farm Tahoe White Chocolate Macademia Cookies'

Safeway Store

{STORE_ID} a :Safeway ; :storeName {STORE_NAME} ; :address {STORE_ADDRESS} ; :geo:lat {LAT}^^xsd:decimal ; :geo:long {LON}^^xsd:decimal .

select * from STORE where STORE_BRAND = 'Safeway'

Whole Foods Store

{STORE_ID} a :WholeFoods ; :storeName {STORE_NAME} ; :address {STORE_ADDRESS} ; :geo:lat {LAT}^^xsd:decimal ; :geo:long {LON}^^xsd:decimal .

select * from STORE where STORE_BRAND = 'Whole Foods'

Mapping count: 11 Search: [] Enable filter

To use the reasoner click Reasoner > Start reasoner

grocery-ontology (http://protege.stanford.edu/ontologies/grocery-ontology) : [Users/jhardi/Documents/Work/Short Course Mar-2017/Ontop Demo/_wiki/posc-groceries.owl]

Active Ontology Classes Object Properties Individuals by class Ontop Mappings Ontop SPARQL

ontop query editor:

Query Editor

PREFIX : <http://protege.stanford.edu/ontologies/grocery-ontology/>

PREFIX owl: <http://www.w3.org/2002/07/owl#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>

PREFIX xml: <http://www.w3.org/XML/1998/namespace#>

PREFIX foaf: <http://xmlns.com/foaf/0.1/>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT * WHERE {

?x a :GroceryItem ; :productName ?name ; :price ?price ; :isSoldByStore [a :Safeway] .

FILTER (?price < 4.00)

}

Execution time: 0.046 sec - Numb... Show: [] All [] Shor... Attach Prefi... Execute Save Changes

x	name	price
<http://protege.stanford.edu/ontologi...	"Walkers Shortbread"	"3.81"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Back To Nature Fudge Mint Cookies"	"2.63"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"General Mills Cheerios"	"3.00"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Pepperidge Farm Tahoe White Chocol...	"2.53"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Kashi Crunch Chocolate Chip Bar"	"3.36"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Lotus Biscoff Cookies"	"2.33"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Kashi Crunch Chocolate Chip Bar"	"2.89"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Walkers Shortbread"	"3.24"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Familia Swiss Muesli Mixed Cereals"	"2.34"^^xsd:decimal
<http://protege.stanford.edu/ontologi...	"Familia Swiss Muesli Mixed Cereals"	"3.43"^^xsd:decimal

Export to CSV...

Reasoner state out of sync with active ontology [x] Show Inferences

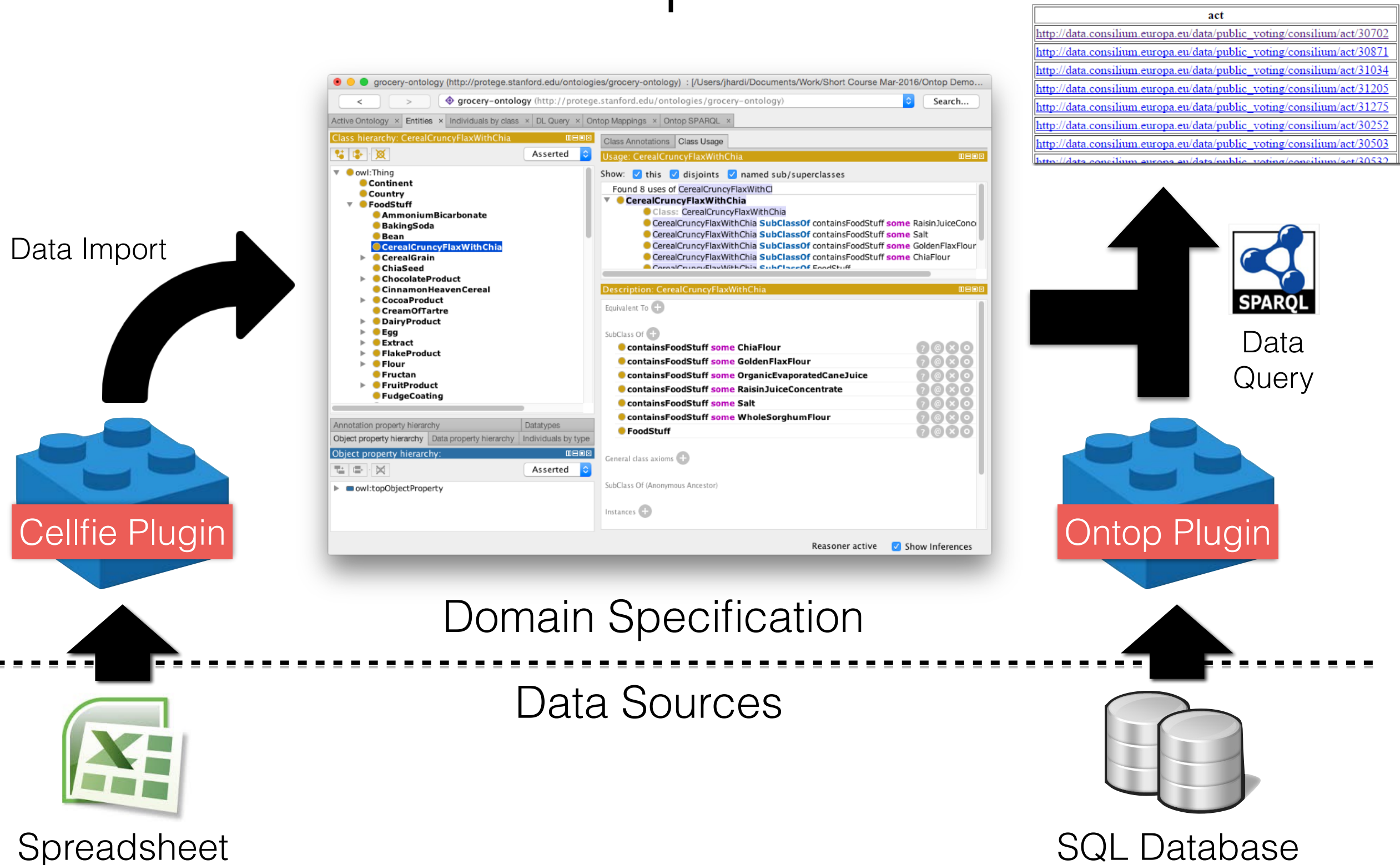
Start the Demo Engine

- Introduce the mapping structure and syntax
- Get familiar with the plugin UI
- Run some examples



Conclusion

The Pipeline



Questions?



Resources

- **M.J. O'Connor, C. Halaschek-Wiener, M. A. Musen**
Mapping Master: A Flexible Approach for Mapping Spreadsheets to OWL 9th International Semantic Web Conference (ISWC), Shanghai, China, 2010.
- **D. Calvanese, B. Cogrel, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, and G. Xiao**
Ontop: Answering SPARQL Queries over Relational Databases on Semantic Web Journal, 2016.
- **Botoeva, E., Calvanese, D., Cogrel, B., Rezk, M., & Xiao, G.** OBDA Beyond Relational DBs: A Study for MongoDB. Description Logics, 2016.

Resources

- **Cellfie Tutorial:** <https://github.com/protegeproject/cellfie-plugin/wiki/Grocery-Tutorial>
- **Ontop Tutorial:** <https://github.com/protegeproject/ontop-tutorial/wiki/Grocery-Tutorial>
- **MappingMaster Syntax:** <https://github.com/protegeproject/mapping-master/wiki/MappingMasterDSL>
- **Ontop Mapping Syntax:** <https://github.com/ontop/ontop/wiki/ontopOBDAModel>
- **Ontop Performance Report:** <http://ontop.inf.unibz.it/performance/>