

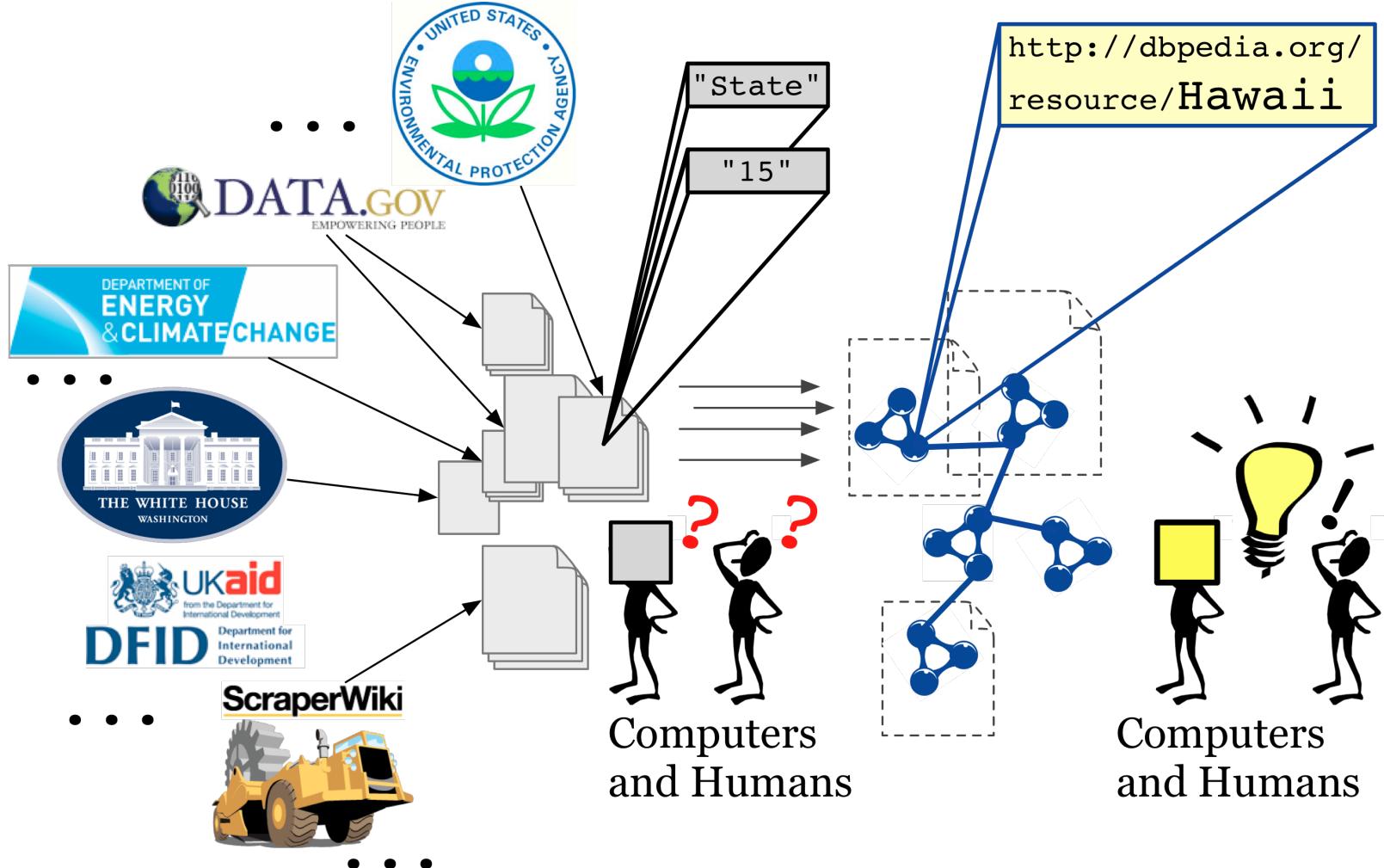


Outline

- Motivation
- *csv2rdf4lod*
 - Quick and dirty: a smart, naïve bootstrap
 - Time and attention: adding enhancements
 - **Distribution and Reproducibility**: version control
 - Finish up Farmers Market example
 - Reproduce
 - Verify
 - Start and Finish RPI Research Center example
 - **Verification**: testing

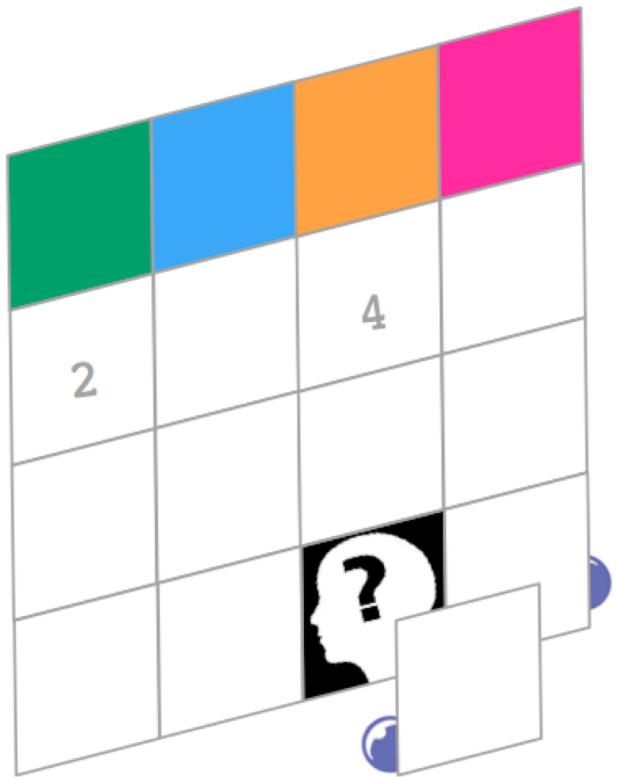


Wanted: Integration





csv2rdf4lod-automation



timrdf / **csv2rdf4lod-automation**

Source Commits Network Pull Requests (1)

Switch Branches (1) Switch Tags (0) Branch List



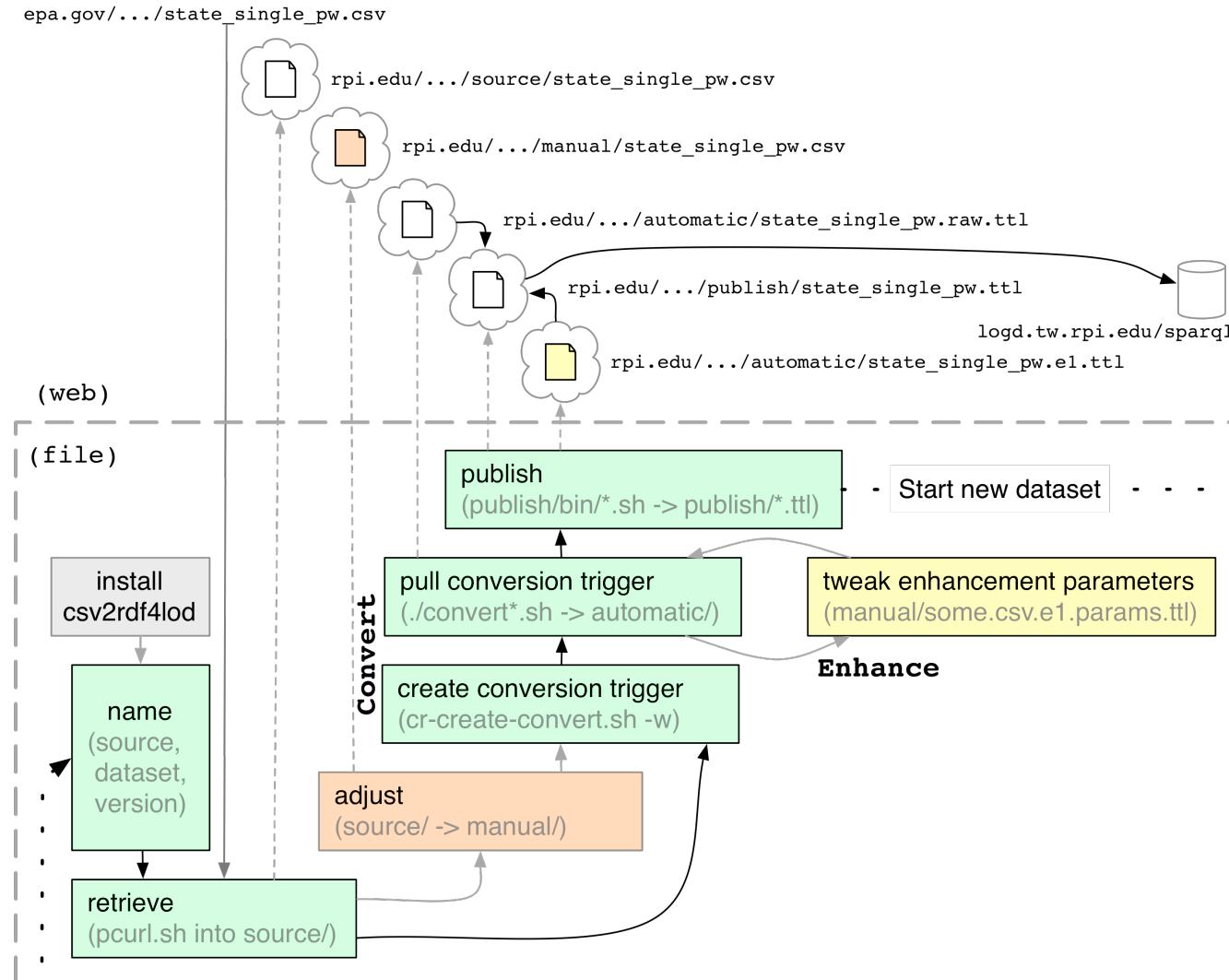
```
$ ./retrieve.sh  
$ vi manual/data.e1.params.ttl  
$ ./convert-foreign-aid.sh  
$ publish/bin/publish.sh
```

<http://purl.org/twc/id/software/csv2rdf4lod> a doap:Software .

<http://purl.org/twc/vocab/conversion/Enhancement> a owl:Class .

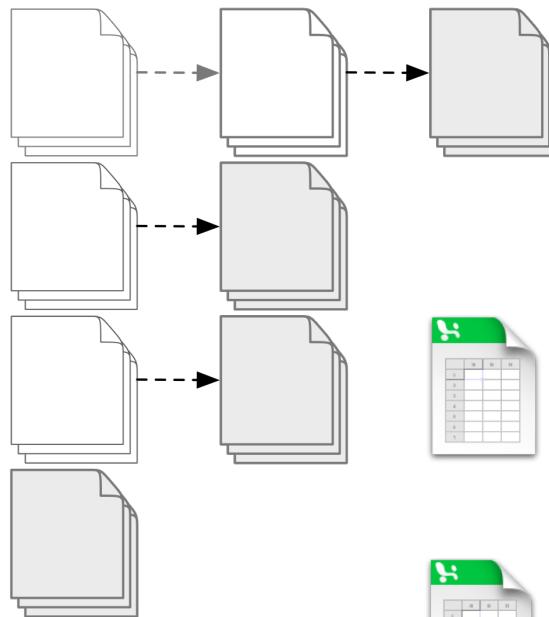
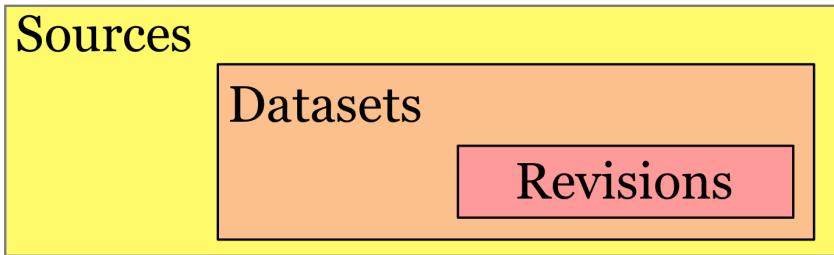


csv2rdf4lod's Integration Stages





A smart, naïve bootstrap: Contextualizing Data



who?	source:	"data-gov"
what?	dataset:	"1095"
when?	version:	"2011-Aug-02"
		"HI"

who?	source:	"data-gov"
what?	dataset:	"4383"
when?	version:	"2011-Sep-27"
		"Hawaii"



Contexts on the Web

Contextual Depth Entity URIs

1	http://dbpedia.org/resource/Hawaii
0	http://lodt.rpi.edu /id/us/state/ Hawaii
-1	http://lodt.rpi.edu /source/ data-gov / Hawaii
-2	http://lodt.rpi.edu /source/ data-gov /dataset/ 1095 /typed/state/ HI
-2	http://lodt.rpi.edu /source/ data-gov /dataset/ 1095 /value-of/state_name/ HI
-3	http://lodt.rpi.edu /source/ data-gov /dataset/ 1095 /version/ 2011-Aug-02 / thing_1

Global Context
Local Context



"Who?"
"What?"
"When?"
-->

who?	source:	"data-gov"
what?	dataset:	"1095"
when?	version:	"2011-Aug-02"

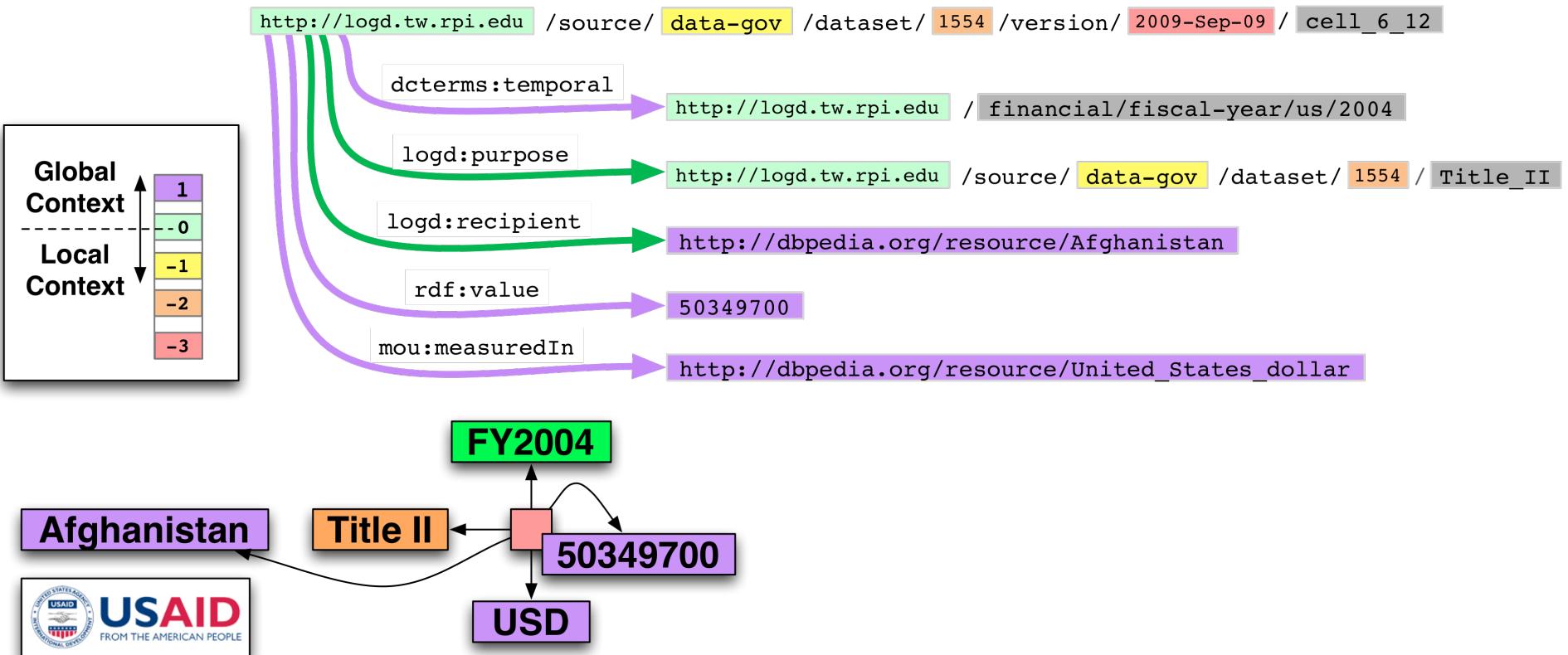


Using Multiple Contextual Depths



country_name	program_name	FY2004	FY2005
Afghanistan	Title II	50349700	57233800
Albania	Child Survival and Health	1000000	150000

United States Foreign Aid





Example Enhancement Parameters

```
<http://lodg.tw.rpi.edu/source/data-gov/dataset/4383/version/2010-Oct-22/
  a conversion:LayerDataset, void:Dataset;

  conversion:base_uri          "http://lodg.tw.rpi.edu"^^xsd:anyURI;
  conversion:source_identifier  "data-gov";
  conversion:dataset_identifier "4383";
  conversion:version_identifier "2010-Oct-22";

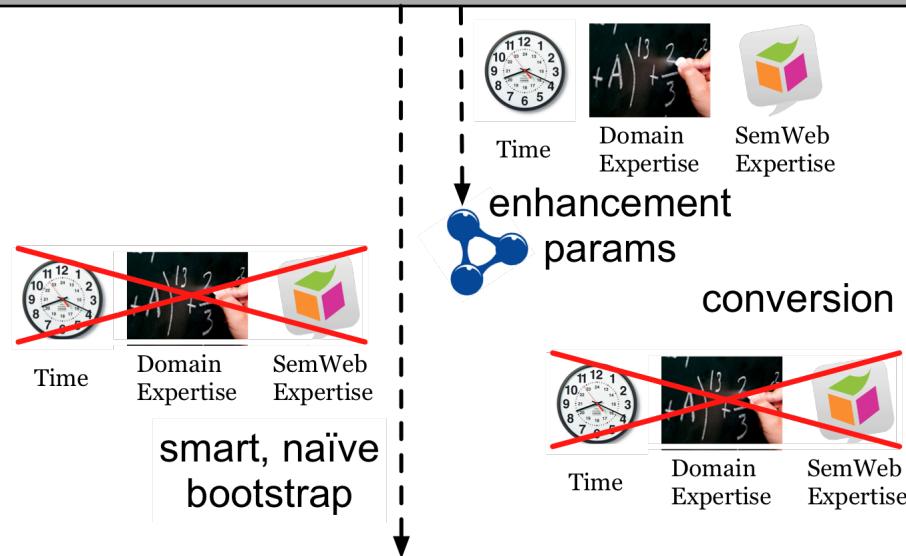
  conversion:conversion_process [
    a conversion:EnhancementConversionProcess;
    conversion:enhancement_identifier "1";

    conversion:enhance [
      ov:csvCol           1;
      ov:csvHeader        "LocAddState";
      conversion:bundled_by :address_bundle;
      conversion:domain_name "FarmersMarket";
      conversion:equivalent_property con:stateOrProvince;
      conversion:comment "";
      conversion:range     rdfs:Resource;
      conversion:range_name "State";
      conversion:links_via <http://www.rpi.edu/~lebot/lod-links/state-fips-dbpedia.ttl>,
                            <http://www.rpi.edu/~lebot/lod-links/state-fips-geonames.ttl>,
                            <http://www.rpi.edu/~lebot/lod-links/state-fips-govtrack.ttl>;
      conversion:subject_of dcterms:identifier;
    ];
  ];
}
```



One-time cost

```
"Hawaii", "Alii Garden Market Place",
"75-6129 Alii Drive", "Kailua-Kona",
"96740", "-155.9819183", "19.61436844"
```



```
ds4383:thing_1367
  raw:column_1 "Hawaii";
  raw:column_2 "Alii Garden Market Place";
  raw:column_3 "75-6129 Alii Drive";
  raw:column_4 "Kailua-Kona";
  raw:column_5 "96740";
  raw:column_6 "-155.9819183";
  raw:column_7 "19.61436844" .
```

```
ds4383:thing_1367
  con:preferredURI
ds4383:farmersMarket_1367 .
```

```
ds4383:farmersMarket_1367
  a ds4383_vocab:FarmersMarket;
  con:address :address_1367;
  dcterms:title "Alii Garden Market Place";
  wgs:lat -155.9; wgs:long 19.6 .
```

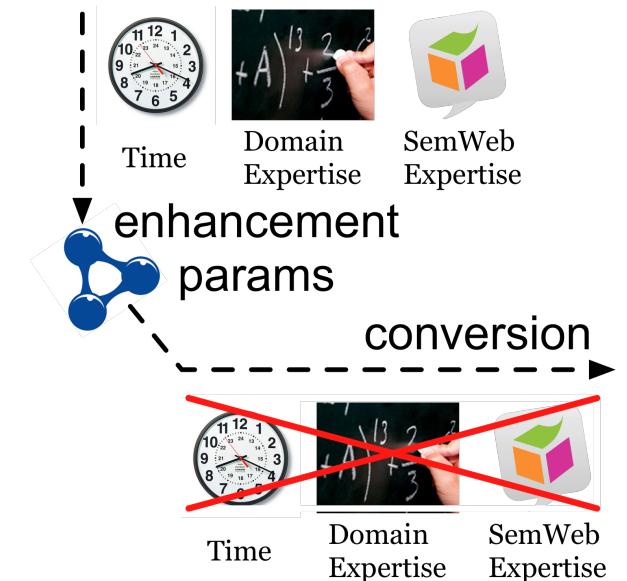
```
:address_1367
  a con:Address;
  con:stateOrProvince typed_state:Hawaii;
  con:street           "75-6129 Alii Drive";
  con:city              "Kailua-Kona";
  con:zip               "96740" .
```

```
typed_state:Hawaii
  a ds4383_vocab:State;
  dcterms:identifier "Hawaii";
  rdfs:label          "Hawaii";
  owl:sameAs
    <http://sws.geonames.org/5855797/>;
    govtrackusgov:HI,
    dbpedia:Hawaii .
```



Declarative Enhancement Parameters (in RDF)

- Transformation terminology common across all datasets
 - High-level specification; small file size
 - Analogous to RDFS/OWL
 - “Easier” to inspect new enhancements
- No software tailored to specific datasets
- Automatically reproducible
- “Straightforward” provenance
 - Small, common, exercised black box controlled by familiar parameters
- Queriable
 - SPARQL for datasets according to how they were transformed.
 - Filter data by dataset transformations





Outline

- Start and Finish RPI Research Center example
 - Reconstruct (and verify) the RDF
 - Layer 1 versus Layer 2 (cell-based qb)
 - Subclass enhancement
 - Templates to consolidate People
 - LOD-Linking from SPARQL query
 - SPARQL Named graph organization
 - Publish/bin/*



Named Graph Organization

Conversion VoID Metadata:

```
http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers
    a conversion:AbstractDataset, void:Dataset
    void:subset
http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers/2011-Oct-17
http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers/2011-Oct-18
http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers/2011-Nov-02
    a conversion:VersionedDataset, void:Dataset
    void:dataDump
<http://logd.tw.rpi.edu/source/data-rpi-edu/file/research-centers/
 /2011-Nov-02/conversion/data-rpi-edu-research-centers-2011-Nov-02.ttl>
```

SPARQL Endpoint: <http://logd.tw.rpi.edu/sparql>

```
INSERT http://logd.tw.rpi.edu/source/data-rpi-edu/file/research-centers/
        version/2011-Nov-02/conversion/data-rpi-edu-research-centers-2011-Nov-02.ttl
INTO   http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers/2011-Oct-17

http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers/2011-Oct-18
http://logd.tw.rpi.edu/source/data-rpi-edu/dataset/research-centers/2011-Nov-02
http://logd.tw.rpi.edu/source/twc-rpi-edu/dataset/instance-hub-people
```