



Easy RDF and SPARQL for LAMP systems

presented by:
Akilah McIntyre
& EE Durham

ARC RDF Store

ARC RDF Store

- What is ARC RDF Store (ARC2) and Why do we need it?
- What features are included in ARC?
- What can I do with it?
- Who else is using it?
- Sample applications of ARC

What is ARC RDF Store?

- ARC is a flexible RDF system for semantic web and PHP practitioners.
- It's free, open-source, easy to use, and runs in most web server environments
<https://github.com/semsol/arc2/wiki>
- 5 minute install and you are up and running with your own RDF parser and SPARQL endpoint
<http://localhost:8090/arc2-starter-pack/endpoint.php>

- **Parsers:**
RDF/XML, N-Triples, Turtle, SPARQL + SPOG, Legacy XML, HTML tag soup, RSS 2.0, Google Social Graph API JSON
- **Serializers:**
N-Triples, RDF/JSON, RDF/XML, Turtle, SPOG dumps
- **RDF Storage (using MySQL):**
SPARQL SELECT, ASK, DESCRIBE, CONSTRUCT, + aggregates, LOAD, INSERT, and DELETE
- **SPARQL Endpoint Class**

Why do we need it?

- WE WANT RAW DATA NOW”
 - Sir Tim Berners-Lee, Semantic Web and WWW founder, from 2009 TED Talk “Tim Berners-Lee on the next Web”,
http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html
- Background:
 - Linked Data
<http://linkeddata.org/>
 - The Linking Open Data cloud diagram
<http://richard.cyganiak.de/2007/10/lod/>

- So now you have all this data – use & access it

- SPARQL

<http://en.wikipedia.org/wiki/SPARQL>

<http://www.w3.org/TR/rdf-sparql-query/>

What features are included in ARC?

- Loading and processing RDF, OWL, etc
- ARC RDF Store::Remote
- ARC RDF Store::Local (MySQL)
- Data manipulation

<https://github.com/semsol/arc2/wiki>

What can I do with it?

- There is TONS of remote data
 - Cannot expect to keep up with changing Data (too much of it)
- Data owners keep up with their own data:

<http://www.obofoundry.org/>

<http://www.linkedmdb.org/>

<http://www.data.gov/catalog/raw>

<http://data.gov.uk/data>

<http://www.fedscope.opm.gov/>

- Web-accessible data stores

NOTE: not all endpoints do allow remote retrieval of their data -- many only allow access from their own site, or to download info for local processing ex: data.gov, dbpedia etc

<http://www.w3.org/wiki/SparqlEndpoints>

<https://wiki.base22.com/display/btg/Semantic+Web+SPARQL+end-points>

- ARC RDF Store::Remote

<https://github.com/semsol/arc2/wiki/Remote-Stores-and-Endpoints>

```
/* ARC2 static class inclusion */
include_once('path/to/arc/ARC2.php');

/* configuration */
$config = array(
    /* remote endpoint */
    'remote_store_endpoint' => 'http://example.com/sparql',
);

/* instantiation */
$store = ARC2::getRemoteStore($config);
```

- So, optimal solution would be remote processing of remote data

but: SLOW! Dependent on uncontrolled external resources.

- Remote example 1:

http://localhost:8090/arc2-starter-pack/remote_query_movies.php

- Remote example 2:

http://localhost:8090/arc2-starter-pack/remote_query_gene.php

- ARC RDF Store::Local

<https://github.com/semsol/arc2/wiki/Using-ARC%27s-RDF-Store>

```
/* ARC2 static class inclusion */
include_once('path/to/arc/ARC2.php');

/* configuration */
$config = array(
    /* db */
    'db_host' => 'localhost', /* optional, default is localhost */
    'db_name' => 'my_db',
    'db_user' => 'user',
    'db_pwd' => 'secret',
    /* store name (= table prefix) */
    'store_name' => 'my_store',
);
/* instantiation */
$store = ARC2::getStore($config);
```

- Take data from anywhere (data.gov, movie data sites, gene data sites, etc.)
- Load it into local RDF store
- Process and run local application
FAST! Dependent on controlled internal resources.
- Local query example
- [http://localhost:8090/arc2-starter-pack/
PeriodicSample_local.php](http://localhost:8090/arc2-starter-pack/PeriodicSample_local.php)

Who else is using it?

- http://drupal.org/project/sparql_ep
- *The RDF SPARQL Endpoint module indexes the RDF data publicly available on a Drupal site into an ARC2 RDF store. It also provides a SPARQL endpoint via the ARC2 SPARQL interface.*
- Semantic MediaWiki RDF store connector: RDFIO
- http://semantic-mediawiki.org/wiki/SPARQL_and_RDF_stores_for_SMW
- *This extension extends the RDF import and export functionality in Semantic MediaWiki by providing import of arbitrary RDF triples (not only OWL ontologies, as was the case before), and a SPARQL endpoint that allows write operations.*

Sample application of ARC2

- Statistical data consolidation, analysis and review:

Data-gov: investigating open government datasets using semantic web technologies

<http://data-gov.tw.rpi.edu/wiki/ARC2>

- Demo:

Comparing US (USAID) and UK (DFID) Global Foreign Aid

<http://data-gov.tw.rpi.edu/demo/linked/aidviz-1554-10030.html>

Current State ARC2 Development

- As of January 2011, current development work on ARC2 has ceased.
- This is due to two factors:
 - Lack of funds
 - The ever-growing RDF specifications are too complex to efficiently.
- ARC2 is still available
- Current Google Group
 - <http://groups.google.com/group/arc-dev>

Q & A