

# SPARQL FEDERATION

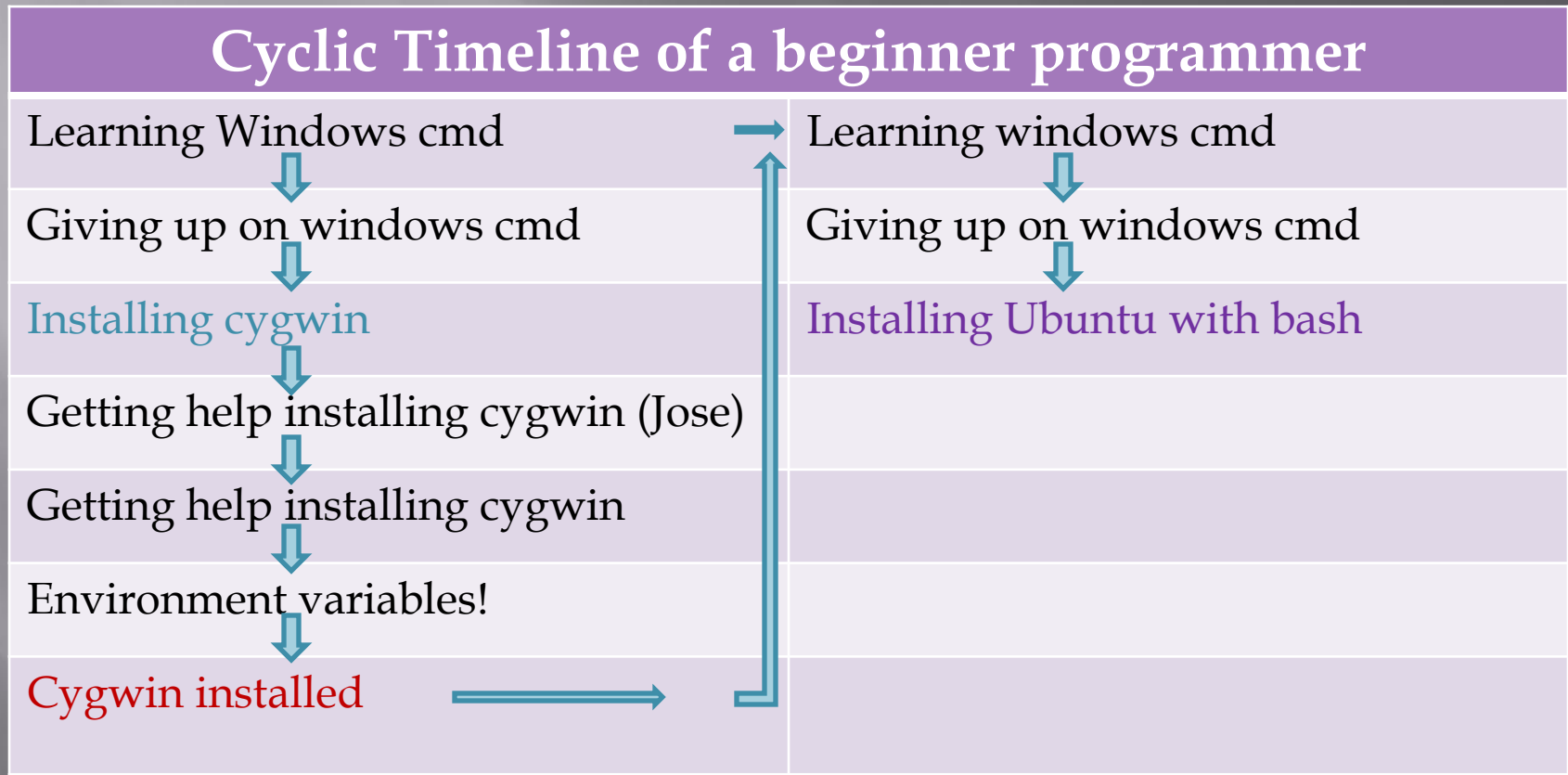
Ray Li

# Introduction to Semantic Web

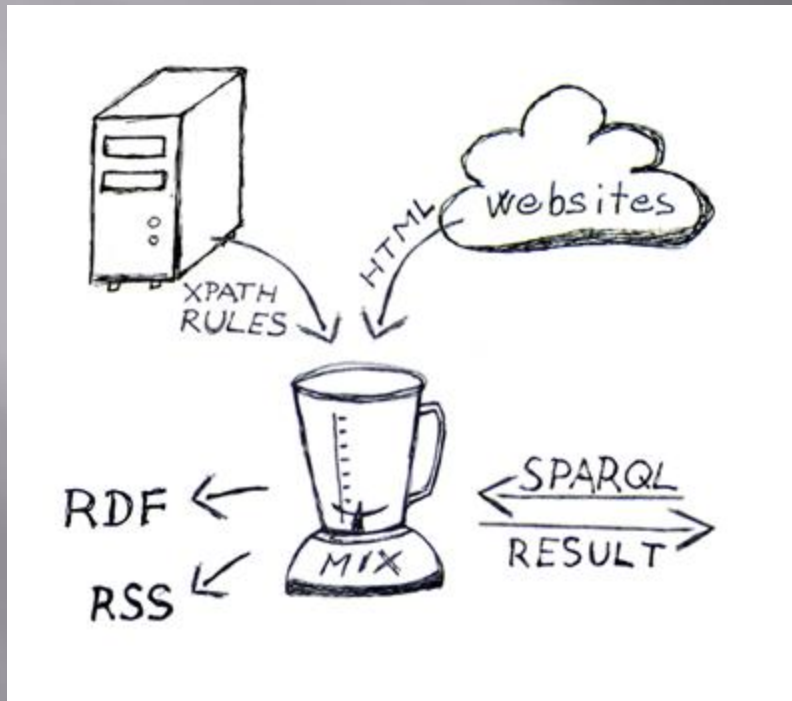


- ▣ AIR
- ▣ N3
- ▣ RDF
- ▣ SPARQL (all 59 pages of its glorious documentation)
- ▣ Very similar to XML

# Installing CWM(RETE) - "python setup.py install"



# SPARQL Endpoints



- ▣ BBC
- ▣ Openlink
- ▣ Revyu (only one that worked)
- ▣ Installed SPARQLWRAPPER successfully!
- ▣ CONSTRUCT, SELECT, ASK, DESCRIBE
- ▣ SPARQL -> Federation

# Endpoint Weaknesses

- ▣ The SPARQL endpoints usually did not implement “DESCRIBE”
- ▣ Also, there was no way of getting meta-data about the data in the rdf store
- ▣ Need standardization of some type of schema or capabilities across the different SPARQL endpoints
- ▣ Tried Querying with html via browser and with python via SPARQLWRAPPER

# Performance

```
initNs=dict(foaf=Namespace("http://  
  /xmlns.com/foaf/0.1/"),  
rev=Namespace("http://purl.org/st  
uff/rev"),  
rdfs=Namespace("http://www.w3.o  
rg/2000/01/rdf-schema"))
```

```
SELECT ?thing ?name ?review  
  ?createdOn ?reviewer  
WHERE {  
  ?thing rdfs:label ?name .  
  ?thing rev:hasReview ?review .  
  ?review rev:reviewer ?reviewer .  
  ?review rev:createdOn  
  ?createdOn.  
},
```

- ▣ <http://revyu.com/sparql>
- ▣ HTML: < 1 sec
- ▣ SPARQLWRAPPER:  
44.947 seconds
- ▣ Definitely needs  
performance optimization
- ▣ MySQL could do a similar  
query in < .1 sec

# DARQ - ARQ, Jena



- ▣ VMWare (ubuntu) ->
- ▣ Voodoo pointer somersaulting across the screen at every twitch of mouse ->
- ▣ Installed CWM (never used it again)
- ▣ Darq-installing -> error messages in German!
- ▣ Ubuntu crashes my computer while cursing at me with ear-splitting beeps.



# Local SPARQL Endpoint



Beginning of conversion to Linux  
(Ubuntu)

Staring at the boot menu for a week

Jose magic!

**!Installed!**

\*bash\* (cd, ls, kill, man, nano, gedit,  
apt-get install!...)

DARQ-installation failed, suggested  
that I try it in windows cmd

**Installation "successful", but none of  
the test cases or examples worked...**

Give up, and install our own local  
SPARQL endpoint (ARC2) and hooked  
up a URL mask, everything go go go.



# Pimpin' out the SPARQL Endpoint

## BEFORE

ARC SPARQL+ Endpoint (v2009-03-05)

[This interface](#) implements [SPARQL](#) and [SPARQL+](#) via [HTTP Bindings](#).

Enabled operations: select, construct, ask, describe, load, insert, delete, dump

Max. number of results : 250

```
SELECT * WHERE {
  GRAPH ?g { ?s ?p ?o . }
}
LIMIT 10
```

Options

Output format (if supported by query type):

jsonp/callback (for JSON results)

API key (if required)

Show results inline:

Change HTTP method: [GET](#) [POST](#)

## AFTER

**DIG ARC SPARQL+ Endpoint at W3C**  
v2009-03-05

TUESDAY, JULY 14, 2009

**SPARQL Federation Project in DIG**

[This interface](#) implements [SPARQL](#) and [SPARQL+](#) via [HTTP Bindings](#).

Enabled operations: select, construct, ask, describe, load, insert, delete, dump (Max. number of results : 250)

```
SELECT * WHERE {
  GRAPH ?g { ?s ?p ?o . }
}
LIMIT 10
```

Change HTTP method: [GET](#) [POST](#)

Database:

Output format:

Jsonp/callback:

API key:

Show results inline:

*Last Updated: 07-14-2009 By: Ray, Yotam*

URL: <http://dig-sparql.vze.com/>

# RDFlib -> ARC2

- ▣ RDFlib (installs easily in bash on Ubuntu)
- ▣ Works perfectly, but its MySQL database *does not* port to ARC2
- ▣ ARC2 – PHP, and MySQL
  - Accepts only proprietary MySQL data, but parses RDF and N3
  - Parses the N3 from CWM but will not do RDF translations
- ▣ Current SPARQL endpoint also supports/simulates multiple database querying.
- ▣ ARC2 also has a pretty good support community and a mailing list, if you have any questions/inquiries.

# Federation Engine!



- Research papers on SPARQL Federation

- Unknown math symbols

- Yotam magic!

- Read lots more papers

- Plan out Federation logic:

Security

- User Identity
- Query history
- Query time frame

Optimization

- Transparency
- Physical
- Content-filtering

Protocols

- Endpoints Capabilities
- Public Policies
- Private Policies

# My expectations coming into the project

- ▣ Work/speak with the people at W3C
  - IE error: *“object expected! at line <#randomly generated number> “*
  - Other various aggravating CSS browser bugs.
- ▣ Database architecture and planning
- ▣ A new perspective of the web from a researcher's point of view



# What I learned

- ▣ How to troubleshoot OS installation on a 7yr-old dell machine (floppy disks!), Ubuntu.
- ▣ *<very important>* How to install poorly documented, and essentially uninstallable packages *</very important>*
- ▣ A much better grasp of database management and optimization: mysql architecture and a little SPARQL
- ▣ New ways of looking at information on the web (via networked RDF graphs - very interesting and powerful)
- ▣ A whole lot about solid state-hard drives, various computer hardware/OS architecture, and really expensive but nevertheless fascinating Cycling accessories from Jose
- ▣ *That I'm not the only one who wants to give the head of the IE development team a piece of his mind.*



# Other things I've learned



- ❑ Innovative ways to solve seemingly impossible problems.

# Concluding word

- ▣ I would like to thank all the faculty, resourceful grad students, fellow UROPs, and of course my supervisor: Lalana for all the guidance, help, and humor you've all imparted to me. I found this summer to be very productive, and I really enjoyed working as part of the team.