What's New in RDF 1.1

SemTechBiz
June 2013

http://www.w3.org/2013/Talks/0603-rdf11

Sandro Hawke, W3C Staff
sandro@w3.org
@sandhawke
Overview

1. Stability and Interoperability
2. Non-XML Syntaxes
   - Turtle (and friends)
   - JSON-LD
3. Datasets
   - “Quads”, “Named Graphs”, “Reification”, ...
4. Improved data types
5. Miscellaneous
6. Q & A
Part 1: Standards

- RDF developed 1997-1999
- RDF “rearticulated” 2001-2004
  - Added datatypes
  - Abstract model, multiple syntaxes in theory
- SPARQL 1.0 2004-2008
- SPARQL 1.1 2009-2013
- RDF 1.1 Working Group: 2011-2013
Chartering

- RDF Next Steps (Workshop, June 2010)
- Future of RDF (Survey, August 2010)
- Charter:
  - Turtle
  - JSON
  - “Multiple Graphs” (Datasets)
  - (smaller things)
Lots of “please don't touch anything”

...but still some pain points:

- XML
- Reification
- Collections/Containers
- ...

Fix things without breaking anything!
Participation

(That's the workshop. Working group mostly meets on phone.)
Part 2: Beyond XML

- **In theory**, RDF didn't have to look like this

```xml
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
</rdf:RDF>
```

- **XML was created for marking-up (annotating) text.**
@prefix contact: <http://www.w3.org/2000/10/swap/pim/contact#>.  
@prefix miller: <http://www.w3.org/People/EM/contact#>.

miller:me a contact:Person;
  contact:fullName "Eric Miller";
  contact:mailbox <mailto:em@w3.org>;
  contact:personalTitle "Dr.".
Turtle Features

- **Elegant Syntax**
  - Pretty easy to read
  - Pretty easy to write
    - Nice shortcuts like [...] and (...) 
  - (Still has namespaces)
- **No confusion with XML**
  - Really matches RDF model
- **Very widely deployed de facto standard**
Changes During Standardization

(Changes from what?)

- Write 18.0 instead of 18. for decimals
  - Previously:  <a> <b> 1 . <c> <d> <e> .
  - Now okay:  <a> <b> 1. <c> <d> <e> .
  - But maybe people are used to it now. We'll see.
- More allowed characters, escaping rules
  - OK: foo:bar:baz   (colon in local part, for FB OGP)
  - OK: part:33445   (local part starting with digit)
W3C Candidate Recommendation

- If you have a Turtle parser:
  - Read over the spec (it's about 20 pages, with lots of examples)
  - See how your code fares against the test suite
  - Send us your results
- Make sure your software can read Turtle
- Consider writing it instead of RDF/XML
Very simple!

but needs a “context”
Determining Context

- Context can come from:
  - Value of @context property of the JSON object
  - URL dereference of value of @context property
  - URL dereference of HTTP header value
  - Application Specific

- Normal JSON can be JSON-LD
  - (if the receiver knows the @context)

- JSON-LD-API provides functions to convert among various useful forms (with & without context)
No problem for Web Apps

Promising (experimental) for data publication

```html
<script type="text/turtle">
...
</script>

<script type="application/ld+json">
...
</script>
```
Fate of RDF/XML?

- Not deprecated
- Valuable deployment (eg Adobe XMP)
- But it's now it's clear, RDF =
  - RDF/XML
  - RDFa
  - SPARQL
  - Turtle
  - JSON-LD
  - … etc
Part 3: Datasets

- Old challenge in RDF. How do you:
  1. Fetch or construct two or more RDF documents
  2. Keep them separate
  3. Say something about them in RDF?

- Reification?
  - 1999 vocabulary for talking about RDF triples in RDF triples. Never fully specified
Triples are in zero or more “named graphs”

```sparql
# some “default graph” triples
<s> <p> <o1>.

GRAPH <g1> {
    # some triples in graph “g1”
    <s> <p> <o1>
}

GRAPH <g2> {
    # some triples in graph “g2”
    <s> <p> <o2>
}
```
Alternate View: Quads

- Formally speaking these are different:
  1. Default graph + a set of <name, graph> pairs
  2. A set of triples and quads

- BUT the specs advise against noticing the difference (ie empty named graphs), so you can think about it either way.
Why Default Graph?

- For folks who don't want the named graphs
  - Just looking for triples? Here you go....

- How to use it?
  - Might be used for metadata (my favorite)
  - Might be union of all named graphs
  - Might be something else

- So how is an application entitled to use it?
  - Depends on how it got it (what link or protocol)
Graph names do not necessarily denote the associated graph. It's just structure; semantics depend on vocabulary of use in that dataset.

GRAPH `<#phase1>` { eg:item7 eg:weight 31 }


We don't specify what `<#phase1>` denotes.

- `vt:begin` and `vt:end` just use it (indirectly) to state properties of some triples.
Blank Nodes in Datasets

- Blank nodes can be shared among graphs
  - Surprising to some people
  - Needed for some use cases (e.g., separation of inference)
- Blank node as Graph Names?
  - Maybe. See ISSUE-131
Syntax

- Examples above use subset of SPARQL
- TriG
  - Very similar, except no GRAPH keywords and default graph in {...}.
- N-Quads
  - 3 or 4 IRIs/literals on a line
- Others?
Part 4: Changes to Data Types

- In RDF 1.0:
  `<s> <p> “Hello World”.`
  was not the same triple as:
  `<s> <p> “Hello World”^^xsd:string.`
- If you ever stumbled on this, it was painful.
- In RDF 1.1 they are the same triple.
A data type for indicating some text is an HTML fragment:

```html
<s>  <p> '''
     <span class="nav">  
       <a href="/">Home</a>
     </span>
   '''^^rdf:HTML.

Consider instead of plain string or rdf:XMLLiteral
Other Datatype Details

- rdf:XMLLiteral changed to match what's usually implemented
- Specific list of datatypes suggested, including new XSD types like durations
- Language-tagging formalized slightly differently
- Ill-typed literals (eg: “a”^^xs:int) now make the graph logically inconsistent
- Now all literals have a datatype
- Semantics are now in RDF, not RDFS
Part 5: Miscellaneous

- Editorial changes (rewrites)
  - rdf-concepts
  - rdf-mt (semantics)
  - … expecting Primer
- Standard way to Skolemize blank nodes
  - http://.../.well-known/genid/...
Collections and Containers

- We're not deprecating rdf:Seq or rdf:List (!!)
- But Turtle and JSON-LD both have special support for “pure” rdf:List structures
- I hope implementations will optimize for that.
  - That is, lists which can be losslessly serialized in turtle using the ( … ) list syntax.
Part 6: Q & A

- See http://www.w3.org/2010/rdf-wg
- All of our drafts, issues, email, etc are public
- We'll answer all email
  - public-rdf-comments@w3.org