



Semantic Web Limitations (RDF 1.1 Starting Points)

Prerana Dahal Sharma



Outline

- Background
- Proposals, Issues & Requirements Presented at W3C Workshop
- Overview on Identified Issues
- Important Issues in Details
- RDF Working Group Next Steps
- Discussion



Background: RDF

- **1999:** first published as a W3C Recommendation
- **2001:** a new Working Group (called "RDF Core") was formed to revise/rewrite the 1999 specification, and add some new features, including datatyped values
- **2004:** the working group completed its work with a set of Recommendations
- practitioners still encounter situations where
 - minor aspects of the current version cause problems
 - the current design is not well explained, and
 - the documents suggest usage patterns which are not currently considered to be good practice
- **2010:** The RDF Next Steps Workshop was held to consider a revision of the 2004 version of RDF



Proposals, Issues and Requirements

- 28 out of 32 extended abstracts or papers were accepted to the Workshop
- A series of presentation based on these papers included different proposals, issues and requirements for the revision of RDF.



Proposals, Issues & Requirements^(a)

- **[Next steps for RDF: Keep the core and pave the cowpaths], by Richard Cyganiak**
- urging caution with respect to possible changes
 - must serve **interoperability**
 - do not **disrupt the network effect**
 - W3C is not the place for R&D
 - focus on areas where RDF is useful
 - can pave the **cowpaths**



Proposals, Issues & Requirements ^(b)

- **[Towards a minor revision of RDF], by Jeremy Carroll**
- identified a number of issues which they would like to see changed.
 - Small Steps Please or not at all
 - If it's not broke, don't fix it
 - **Benefits outweighs the cost**



Proposals, Issues & Requirements[©]

- **[RDF Syntaxes 2.0], by David Beckett**
- identified a number of issues with current RDF Syntax Formats
- Recommendations on Existing Syntaxes
 - **RDF / XML :**
 - low risk to: Add rdf:namespaced elements and attributes such as for graph names, full property URIs
 - **RDFa :** revise using experience with HTML5 world.
 - **Turtle :** align with SPARQL's forked triple pattern syntax.
 - if incompatibility is not a concern
 - **RDF / XML :** remove rdf:ID, property attributes and reification.
 - **N-Triples :** allow Unicode characters.
 - **Turtle :** maybe add named graphs.



Proposals, Issues & Requirements ^(d)

- **[What I want in RDF 2.0], by David Booth**
- identifying a number of issues with the current RDF specification
 - **Standardize a rules language based on SPARQL CONSTRUCT.**
 - Since RDF users already know SPARQL, If such rules language were based strictly on SPARQL CONSTRUCT -- nothing more and nothing less – **the standardization effort would be minimal.**



Proposals, Issues & Requirements ^(d)

- **Standardize an XML Schema-friendly serialization.**
 - It must support named graphs. [TriX](#) is the most obvious candidate.
- **Permit literals as subjects.**
 - Although there are work-arounds, they add unnecessary complication, confusion and debate.
- **Named graphs.**
 - Being standardized will help push tools to uniformly support them.
- **Do something about bnodes**
 - Get rid of them? Maybe provide a standard namespace for minting recognizable, but unique URIs?
- **Standardize turtle, n3 or other human-friendly syntax.**
 - Reading RDF/XML is like trying to read hexadecimal.



Proposals, Issues & Requirements ^(f)

- **[OMG Ontology PSIG], by Elisa Kendall**
- reported on the experiences the OMG Ontology PSIG has with current RDF specification
- Challenges
 - No independent specification of **the common elements of RDF vocabularies & OWL ontologies** that connect them to the web
 - **Some elements**, including documents, local names, namespaces, namespace definitions, and IRIs, could **be collected in a common specification that both language reference**
 - **Common specification for literal and built-in-datatypes(& facets)**, rather than embedding them in the OWL 2 syntax specification



Proposals, Issues & Requirements ^(f)

- **Containers & Collections-** a circular relationship between independent metamodels for RDF and RDF schema would be required to maintain namespace separation, which is not permitted in UML
- Vocabulary & Ontology alignment & mapping is high priority for mapping the semantics of UML & domain specific language models
- **Named graphs & related capabilities** defined should be considered seriously



Proposals, Issues & Requirements ^(f)

- **Need for Standard Interfaces/APIs**
- a number of APIs for accessing RDF/S & OWL Data/KBs
 - Jena, Sesame/Sail, DIG
 - OWL API, OWL Link
- They provide **varying degrees of language coverage**, varying completeness, varying levels of robustness, error handling, explanation support
- **Lack of real standard**, no common way of describing IRIs, documents, local names, namespaces, or additional services from an API perspective
- Organizations **building tools to bridge the UML & Semantic Web**
- Standards must use **multiple, often competing APIs with conflicting jar files**



Proposals, Issues & Requirements ^(g)

- **[RDF and XML: Towards a Unified Query Layer], by Axel Polleres**
- reported requirements from a number of SPARQL users regarding a **combined XML and RDF querying**
 - One Unified RDF/XML/RDB Query Layer to combine XQuery+SPARQL+GRDDL+SQL ?
 - Enable **optimisation across layers**
 - query each source format in native language instead of multistep transformation via “narrow” interfaces (e.g. SPARQL, SPARQL-Result/XML, Xquery/XSLT)
 - Enable declarative view (a la Relational Algebra for core fragment of the language)
 - **Tighter integration** of various source formats that populate the (Semantic) Web
 - **needs a cross-activity effort in W3C?** Semantic Web + XML + others?



Proposals, Issues & Requirements ⁽ⁱ⁾

- **[Contextualized RDF Importing]**, by Jie Bao
- highlighted the usefulness of **contexts and import facilities** for RDF
- Adding **contexts** to RDF
 - Based on Named Graph
 - Add **two constructs**
 - `rdf:context`
 - `rdf:constructs`



Proposals, Issues & Requirements ⁽ⁱ⁾

- **Need for annotations** of RDF statements by Axel Polleres
- Our Claim:
 - RDF needs agreement **on representation and semantics for important annotation domains e.g. time, provenance, trust**
- Representational Issues:
 - several options (reification, N-quads, TriG/X)
 - reification the only standards compliant thus far, sub-optimal
- Semantics of annotations:
- Proposal: Annotated RDFS
 - allows arbitrary ordered annotation domains
 - give them a semantics on top of RDFS
 - live side-by-side with non-annotated RDF
 - SPARQL(1.1) compatibility



Proposals, Issues & Requirements ^(k)

- **[RDF: Back to the Graph], by Peter F. Patel-Schneider**
- presented his motivation to revise the **RDF semantics towards a pure data structuring language**
- A suggestion to move RDF back to a data-structuring language
 - How to make this all work out while old-style RDF remains?
 - Add new MIME types for
 - RDF as data,
 - RDF encoding RDFS,
 - RDF encoding OWL, and
 - OWL ontology in XML.
 - Retain existing RDF MIME type for uncategorized RDF



Proposals, Issues & Requirements ^(k)

- **The New RDF Order**
- RDF is a general-purpose data structuring language,
 - not just triples and graphs.
- Names, etc., in RDF still have their SW characteristics,
 - global, etc.
- Semantics come in at the RDFS or OWL level.
- RDFS and OWL don't have to handle all of RDF.
 - Might not handle, e.g., quads, trees, named graphs.
 - Might not be extensions of current RDF semantics.
- Is this much different from the current situation?
 - There are incomplete systems for RDF(S).
 - Many RDF(S) systems are unsound, e.g, on owl:sameAs.
 - Not all OWL 2 handles all of RDF.
 - Not all OWL 2 is an extension of RDF semantics.



Proposals, Issues & Requirements ⁽ⁿ⁾

- **[RDF Isolation API], by James Leigh**
- presented a **RESTful API for isolating changes to RDF data sets** and a proposal to make **RDF Lists first-class objects**.
- Restful API
 - Need an API to:
 - Manage RDF Services
 - Manage Multiple RDF Store States
 - Managing Queries
 - Describe Relationships between Services
- RDF Lists
 - Ordered lists as **an RDF term**
 - Treat Ordered lists **as a single term**
 - All RDF format must be able to serialize ordered lists
 - May use `rdf:first/rdf:rest` vocab
 - Parsers should **parse lists as term**
 - This included `rdf:first/rdf:rest` vocab
 - Writers should use **compact list syntax**



Proposals, Issues & Requirements ⁽⁰⁾

- **[When owl:sameAs isn't the same: An Analysis of Identity Links on the Semantic Web], by Ivan Herman**
- initiated a discussion on the usage of owl:sameAs and on what alternatives could be used.
 - Two things are not identical but simply closely related in some manner.
 - Two URIs do refer to the same thing and all properties do hold of both URIs, but that we cannot re-use the URI in a different context.
 - With some characteristics being formalized (e.g., like skos:related that is defined to be **symmetric but not transitive**)
 - Without any formal semantics, with only guidelines for usage (a bit like rdf:value)



Proposed Work Items

- Standardize Model for Graph Identification
- Modify Semantics to Support Graph Identification
- Switch to Improved Inference Rules
- Apply Fixes to known Spec Errors
- Standardize a JSON RDF Syntax
- Make Turtle a W3C Standard
- Add Graphs to Turtle
- Add Graphs to RDF/XML
- Revise Blank Node Semantics



Proposed Work Items

- Specify Linked Data Style of RDF
- Weakly Deprecate some RDF/XML Features
- Define alternatives to owl:sameAs
- Weakly Deprecate some Data Model Feature
- Namespace Profiles
- Weakly Deprecate some RDF Semantics Features
- Have Explicit Support for Annotations
- Align RDF Semantics with SPARQL
- Improve rdf:List Support in RDF/XML
- Explain how to determine What a URI Means
- Allow Literal as Subjects



Identified Issues

- Two major categories were “**RDF Core**” and “**RDF Infrastructure**”
- RDF Core issues were treated as a higher priority for a potential Working Group to consider.
- RDF Infrastructure issues were elevated for discussion only if they were seen as urgently important for immediate treatment by a Working Group;
- Issues under this category were not felt to be fundamental for a new RDF group, and could be tackled by possibly other groups or the community at large.



Identified Issues

○ RDF Core

- Graph identification (Named graphs)
- Turtle
- Specifying “follow your nose”
- JSON
- Atom
- XSLT-friendly XML
- XML Schema-friendly XML
- Skolemize bnodes
- Disentangle RDF/RDFS namespaces
- URI->IRI X Binary RDF
- n-ary predicates
- Revise semantics
- List construct
- RDFa style profiles in, e.g., Turtle
- Weakly deprecate some features



Identified Issues

- RDF Infrastructure
 - Rule-based querying
 - Identity for non-RDF resources
 - Evolution (provenance)
 - Provenance vocabularies
 - Annotations
 - Standard APIs
 - Unified RDF/XML/RDB query layer
 - Change vocabularies
 - Isolation API
 - Identity vocabulary



Findings: Issues of Importance

- Standardize Model for Graph Identification
- Modify Semantics to Support Graph Identification
- Fix the inference rules in Semantics
- Apply Fixes to known Spec Errors and Shortcomings, including considering postponed issues from the previous RDF Working Group
- Standardize a JSON RDF Syntax
- Make Turtle a W3C Standard
- Add Graphs to Turtle



Standardize Model for Graph Identification

- Produce a W3C Recommendation which provides for interoperability for selected use cases for reification, named graphs, graph literals, annotations, etc.
- **Why ?**
 - widely used by the community
 - part of SPARQL already
 - numerous use cases
 - clarify confusion in implementation
- **Why not?**
 - adds complication and may not solve the issue nevertheless
 - complicates the RDF model (potentially)
 - risks with backward compatibility should be assessed (e.g., syntax)
 - does it need standardization?



Standardize Model for Graph Identification

○ Proposals

- Named graphs, provenance and trust , Jeremy Carroll, Christian Bizer, Patrick Hayes, Patrick Stickler, WWW 2005, <http://www.w3.org/2009/12/rdf-ws/p613.pdf>
- <http://www4.wiwiwiss.fu-berlin.de/bizer/SWTSGuide/carroll-ISWC2004.pdf>



Standardize Model for Graph Identification

o Likely Technical Issues

- mutual roles of quads vs. singleton named graphs vs. named graphs
- extension the RDF(S) semantics?
- new RDF(S) terms? `rdf:Graph`, `rdf:subGraphOf`, `rdf:equivalentGraph`, etc.
- syntax (TRIG, n3)
- graph inclusion, can named graphs share triples
- whether blank nodes can be shared among multiple graphs
- whether blank nodes can be used as graph names
- named graphs do not fully replace reification
- relationships to SPARQL



Modify Semantics to support Graph Identification/Named Graph

- **Semantics for the Next Steps**
 - Updating the semantics to handle extensions added to RDF, e.g, named graphs. This could be very tricky for the current style of the RDF semantics, particularly if there is interesting intended meaning to capture.
- **Why ?**
 - Whatever the rationale is for the extension.
- **Why not?**
 - Tricky semantics may be needed.
- **Proposals**
 - Go to the unique-model semantics, where much would be easier.
 - Require that any extension come with a semantics.
- **Likely Technical Issues**
 - If the RDF and RDFS semantics become unique-model semantics then any extension may become simple.
 - Issues from carroll et al. include special interpretation of graph names and a built-in subGraphOf predicate.
 - A further issues with named graphs is interactions, if any, between multiple graphs.



Fix the inference rules in Semantics

- **Inference Rules**
 - Fix the inference rules in the semantics, as they are currently incomplete. (implements basic inference, basic forward chaining, extended RDF syntax)
- **Why ?**
 - It's a "bug".
- **Why not?**
 - None known.
- **Proposals**
 - Fix them
- **Likely Technical Issues**
 - None known.
- **People Interested in Doing The Work**
 - The work has been done, what remains is editorial



Apply Fixes to known Spec Errors and Shortcomings

- **RDF/XML and RDF Concepts Errata**
 - Apply RDF/XML and RDF Concepts spec errata
 - Typos, errata folded in, clarifications.
- **Why ?**
 - Make the RDF specs match the latest URI work in IRIs
- **Why not ?**
 - Not sure of the implication of the IRI change.
- **Proposals**
 - Revise the specifications to globally substitute the term *RDF URI Reference* with an up-to-date reference to IRIs



Standardize a JSON RDF Syntax

- A Specification for a way to serialize RDF graphs in JSON.
- Should include consideration of adding profiles to remove the need for a bunch of namespaces – eg. Twitter annotations and Facebook open graph.
- Make it friendlier at the top of the document to avoid scaring user.
- Suggest a survey of existing work and a community building "event" or process to bring alignment since this seems urgent to start soon.



Standardize a JSON RDF Syntax

- **Why ?**
 - Allows web authors (Javascript, HTML5, ... developers) more easily to use rdf data with existing tools and techniques.
 - Multiple JSON formats and implementations (some interoperable) already exist showing interest in this work.
- **Why not?**
 - Current JSON formats are not aligned - different approaches - making it JSON-user friendly versus making it familiar to existing RDF users.
 - Needs some R&D and alignment.
 - Risk that the result would be some standard that would not be adopted if it was not 'web author' friendly.
- **Proposals**
 - Possible starting points include: Tails RDF JSON, RDFi, JSON-LD, and JRON
- **Likely Technical Issues**
 - Should support named graphs if they are added to the rdf model.



Make Turtle a W3C Standard

- Decide **the syntax stack of how Turtle themed languages fit together** (N-Triples, any future N-Quads, Turtle, maybe N3) including how the media types work
- A **specification for Turtle**, generally compatible with existing systems which **read and write** it.
- Some **syntax extensions** :
 - allowing **raw date / date time literals to improve validation and ease of use.**
- making this the **recommended RDF syntax.**



Make Turtle a W3C standard

- **Why?**
 - It is in widespread use - in tutorials, W3C docs and code.
- **Why not?**
 - another new syntax - all syntaxes must be aligned and capable of encoding the same models.
 - May need new media type (but the current one is not IETF approved).
 - It is in widespread use - possibly cannot make major changes.
 - May need a new name for the named graph format. Qurtle (Dave B)
- **Likely Technical Issues**
 - Whether to include date.
 - Better formal explanation of mapping from model to syntax.
 - Alignment with SPARQL formats
 - Errata...
 - May need two MIME types for turtle doc that encodes 1 graph only and turtle doc that encodes multiple graphs (sparql dataset) because one needs to know in advance whether to stream an incoming document into a graph or a dataset.



Add Graphs to Turtle

- A specification for an extension to Turtle which includes **support for graph metadata**
- **Why ?**
 - Provide support for the named graph model changes
 - Align with SPARQL dataset work.
 - Alignment with other serializations if/when they support named graphs.
- **Why not?**
 - Concern that one wants to know when a document has one graph, versus may have many graphs.
- **Proposals**
 - Trig and N-Quads.
 - ('{' and '}' to group triples into multiple graphs and
 - to precede named graphs by their names

(Note: Trig is not a true superset of Turtle or N-Quads)
- **Likely Technical Issues**
 - Should this be a superset of Turtle?
 - Expect this to be a different mime type to Turtle, maybe a different named spec.



Recommended Next Steps

- a strong demand for **a few features to be added** in a compatible manner.
- W3C should consider **chartering an RDF Working Group** at the earliest convenience to address those issues.
- W3C should use the workshop **summary table as guidance** in the production of the working group charter.
- the charter should also list some of the work items of the summary table as **“time permitting”**, i.e., to be addressed if sufficiently motivated participants rapidly develop a design acceptable to the Working Group.



Discussion

- Why certain features were included and others were less prioritized?
- Selection Criteria:
 - Important issues were determined by consensus
 - RDF Core vs. RDF Infrastructure
 - Impact on existing systems: For each feature there was a lot of discussion on whether these could be introduced without breaking existing systems. In a number of cases, proposals were abandoned simply because it would impact these systems too much.
 - Widespread acceptance and existing deployment: (despite there being no standard) Example: Turtle standardization or named graphs.



References

- W3C Workshop on RDF:
<http://www.w3.org/2009/12/rdf-ws/#section6>
- W3C Workshop-RDF Next Steps Report:
<http://www.w3.org/2009/12/rdf-ws/Report.html>
- Presentation on Need for annotations of RDF statements:
http://axel.deri.ie/presentations/20100626W3C_RDF_NS_RDFneedsAnnotations.pdf
- RDF Back to the Graph:
[http://ect.bell-labs.com/who/pfps/talks/rdf/rdf-next-steps-pfps.html#\(7\)](http://ect.bell-labs.com/who/pfps/talks/rdf/rdf-next-steps-pfps.html#(7))
- Ordered Lists as Fundamental concepts:
<http://www.w3.org/2009/12/rdf-ws/slides/rdflist.pdf>