Linking Things on the Web

A Pragmatic Examination of Linked Data for Libraries, Archives and Museums

http://arxiv.org/abs/1302.4591

Ed Summers
ehs@pobox.com
Lead w/ the Demo!
Open Graph Concepts
Core Concepts › Open Graph › Open Graph Concepts

Open Graph
Open Graph helps people tell stories about their lives with the apps they use. It provides developers with the opportunity to deeply integrate their app into the core Facebook experience, which enables distribution and growth.

Highly Customized
Watching? Running? Cooking? Open Graph allows people who use your app to share rich stories about what they are doing in it or with it.

Distribution
Open Graph stories appear in the news feed and ticker of friends and on the person’s timeline. They are eligible for more engaging presentation formats and greater distribution.

Re-engagement
Friends discover and rediscovers apps through the stories friends share in news feed and the content they highlight on their timeline.

Topics

Actions
Objects
Aggregations
The Type Hierarchy

Here is the entire hierarchy in a single file. Types that have multiple parents are expanded out only once and have an asterisk.

**DataType**
- Boolean
- Date
- DateTime
- Number
  - Float
  - Integer
- Text
- URL

**Thing**
- additionalType, description, image, name, url
- CreativeWork: about, accountablePerson, aggregateRating, alternativeHeadline, associatedMedia, audience, audio, author, award, awards; comment, contentLocation, contentRating, contributor, copyrightHolder, copyrightYear, creator, dateCreated, dateModified, datePublished, discussionUrl, editor, encoding, encodings, genre, headline, inLanguage, interactionCount, isFamilyFriendly, keywords, mentions, offers, provider, publisher, publishingPrinciples, review, reviews, sourceOrganization, text, thumbnailUrl, version, video
- Article: articleBody, articleSection, wordCount
- BlogPosting
  - NewsArticle: dateline, printColumn, printEdition, printPage, printSection
  - ScholarlyArticle
    - MedicalScholarlyArticle: citation, publicationType
- Blog: blogPost, blogPosts
- Comment
- Diet: dietFeatures, endorsers, expertConsiderations, overview, physiologicalBenefits, proprietaryName, risks
- ExercisePlan: activityDuration, activityFrequency, additionalVariable, exerciseType, intensity, repetitions, restPeriods, workload
4 Rules
1. Use URIs as names for things.
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL).
4. Include links to other URIs. so that they can discover more things.
"Barack Obama".
Given that:
A has property P
A = B

Therefore:
B has property P
<http://creativecommons.org/ns#license>
<http://creativecommons.org/licenses/by/3.0/us/> .
<http://creativecommons.org/ns#license>
<http://creativecommons.org/licenses/by-sa/3.0/> .

The "Fix"

<http://data.nytimes.com/47452218948077706853.rdf>
<http://creativecommons.org/ns#license>
<http://creativecommons.org/licenses/by/3.0/us/>.

<http://data.nytimes.com/47452218948077706853>
<http://www.w3.org/2002/07/owl#sameAs>
By design a URI identifies one resource. We do not limit the scope of what might be a resource. The term "resource" is used in a general sense for whatever might be identified by a URI. It is conventional on the hypertext Web to describe Web pages, images, product catalogs, etc. as "resources". The distinguishing characteristic of these resources is that all of their essential characteristics can be conveyed in a message. We identify this set as "information resources."

(Jacobs & Walsh, 2004).
Are these Information Resources?

- A book
- A clock
- The clock on the wall of my bedroom
- A gene
- The sequence of a gene
- A software
- A service A namespace
- An ontology
- A language
- A number
- A concept
This specification does not limit the scope of what might be a resource; rather, the term "resource" is used in a general sense for whatever might be identified by a URI. Familiar examples include an electronic document, an image, a source of information with a consistent purpose (e.g., "today's weather report for Los Angeles"), a service (e.g., an HTTPtoSMS gateway), and a collection of other resources. A resource is not necessarily accessible via the Internet; e.g., human beings, corporations, and bound books in a library can also be resources. Likewise, abstract concepts can be resources, such as the operators and operands of a mathematical equation, the types of a relationship (e.g., "parent" or "employee"), or numeric values (e.g., zero, one, and infinity). (RFC 3986)
RDF is broken if it cannot describe the Web in its entirety. The whole point in changing the model of the Web from a collection of Universal Document Identifiers that can be used to retrieve documents to one where Uniform Resource Identifiers are used to identify resources that can be accessed in the form of a representation through a uniform interface was so that we could accurately model how people have been actively using the Web since computational hypertext was introduced in 1993. (Fielding, 2002)
3. USE HTTP METHODS AS ROY INTENDED

4. REPRESENTATIONS FOR THIS
What To Do?

• Focus on the URL namespaces that libraries, archives and museums create.
• Improved RDF vocabularies that account for Web Architecture.
• HTML Renaissance
• Tools: editors, publishers, indexers.
Image Credits

- http://www.w3.org/History/1989/proposal.html
- https://developers.facebook.com/docs/concepts/opengraph/
- http://schema.org/docs/full.html
- http://www.dlib.indiana.edu/~jenlrile/metadatamap/
- http://commons.wikimedia.org/wiki/File:Capuchin_monkeys_sharing.jpg
Image Credits (Continued)

- http://inkdroid.org/rage14/emails/
- http://inkdroid.org/rage14/proposal/
- http://www.flickr.com/photos/psd/8271699529