RDFa
How and Why

Ralph R. Swick
World Wide Web Consortium

22 January 2010

A Web of Information

- The current Web represents information using

  - natural language

  - graphics, multimedia, page layout

- Humans can process this easily

  - can deduce facts from partial information

  - can create mental associations

  - can deal with a variety of media

    - (see Web accessibility guidelines)

W3C Semantic Web

Data in HTML Documents

Data Exposed with MarkUp

- microformats *http://microformats.org/*
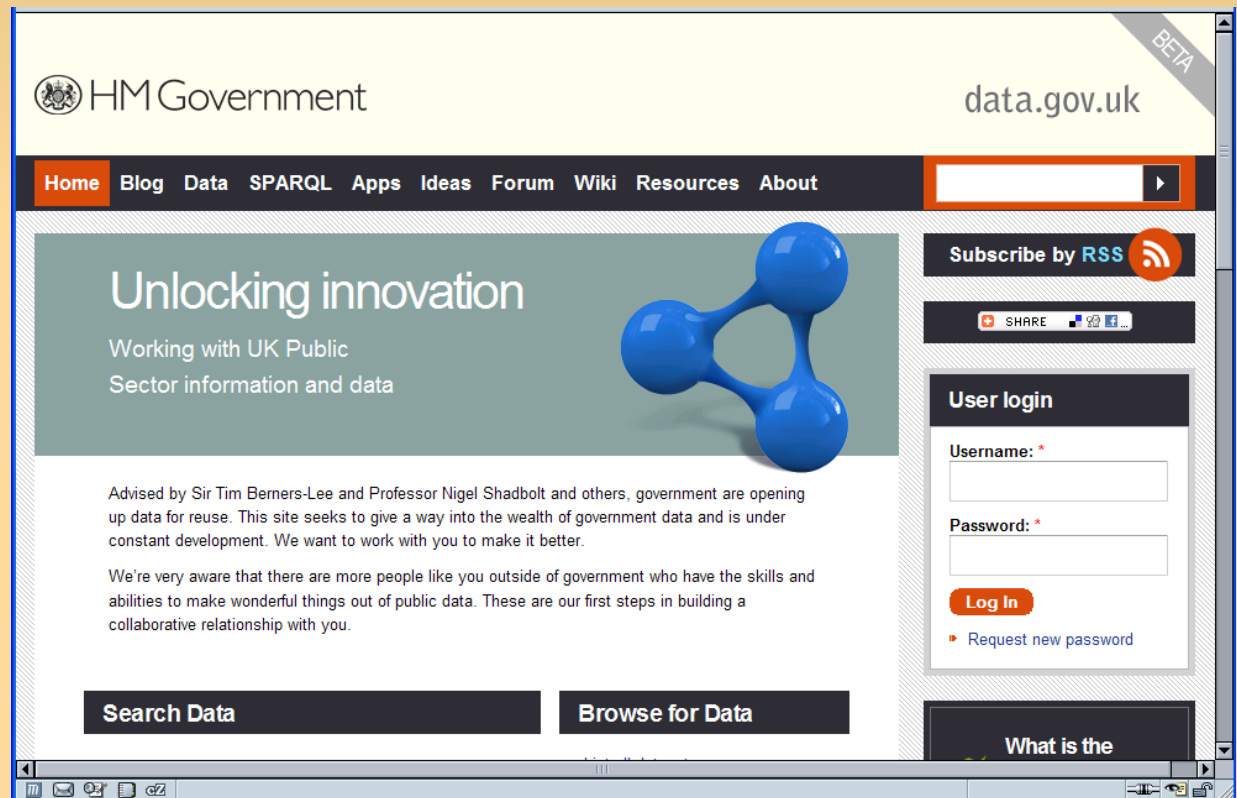
- 9+ sets of terms and relations (formats):
  hCalendar, hCard, rel-license, rel-nofollow, rel-tag, VoteLinks, XFN, XMDP, XOXO

- each of the formats has a custom interpreter

- semi-formal community agreement process for creating new formats (terms and relations)

- extensible only through community agreement

Special-purpose "smart" portals

- Various types of "portals" are created (for a journal on-line, for a specific area of knowledge, for specific communities, etc)

- The portals may:

  - integrate many data sources

  - may have access to specialized domain knowledge

- Goal is to provide a better local access, search on the integrated data, reveal new relationships among the data

A Web of Linked Data

- Encourage *all* data to be published in machine-useable formats

- *Any* data

- see, e.g. *data.gov.uk*

credit: Richard Cyganiak, *LOD cloud*

As of July 2009

W3C Semantic Web

Extensible Structured Data

- *RDF* - Resource Description Framework
  3 syntaxes: RDF/XML, RDFa, RDF/N3*

- A data modeling framework that takes full advantage of the Web

- Self-describing data

- Objects, terms, and relations are all identified using URIs

- Publish *any* data by publishing definitions of your terms and relations

* only RDF/XML and RDFa are W3C Recommendations at this time

Semantic Web data integration

1. Map the various data onto an abstract data representation

   ▪ make the data independent of its internal representation

2. Merge the resulting representations

3. Make queries on the whole

   ▪ queries that could not have been done on the individual data sets

Expose your data as a set of _relations_
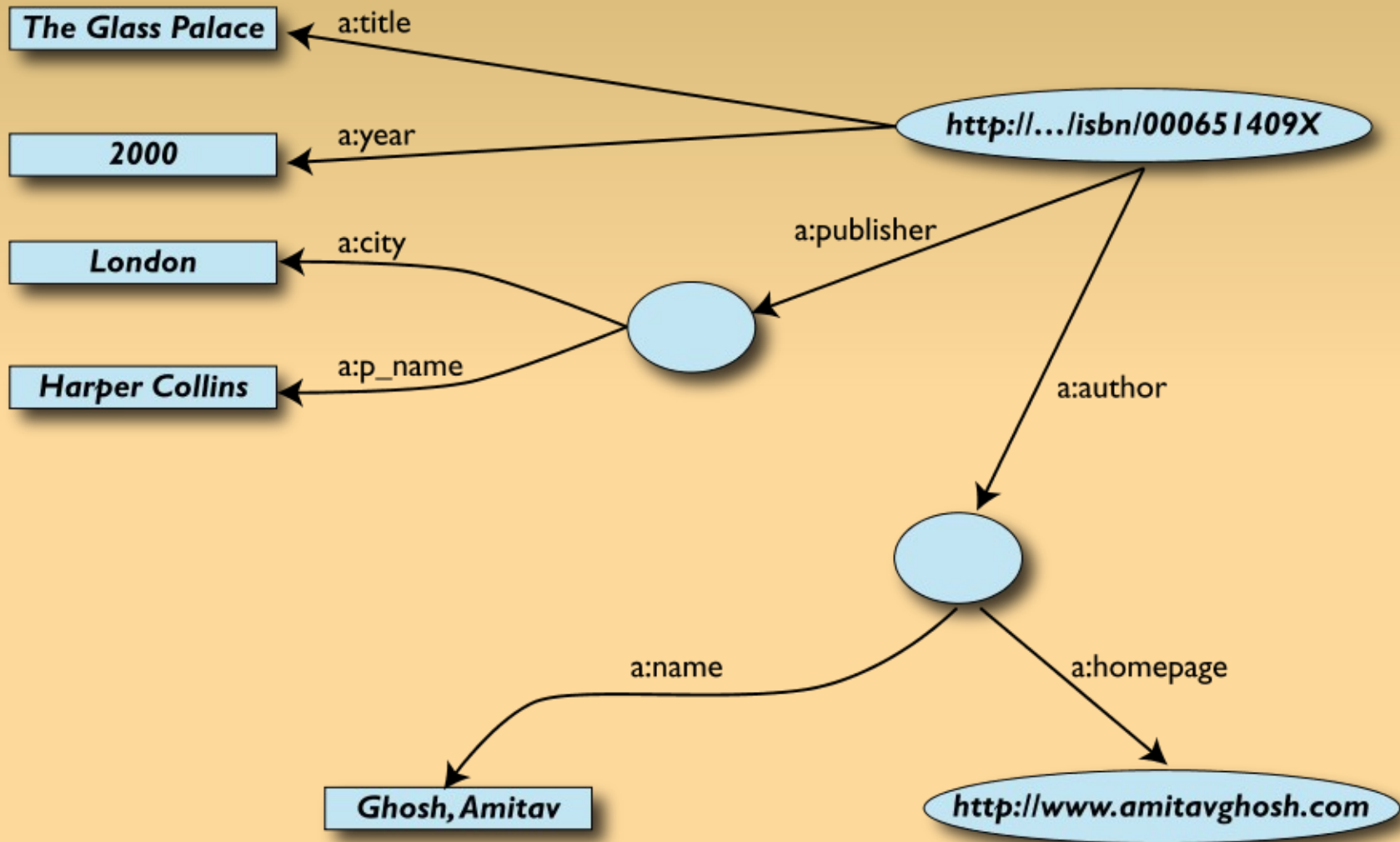
Merge data sources by matching URIs
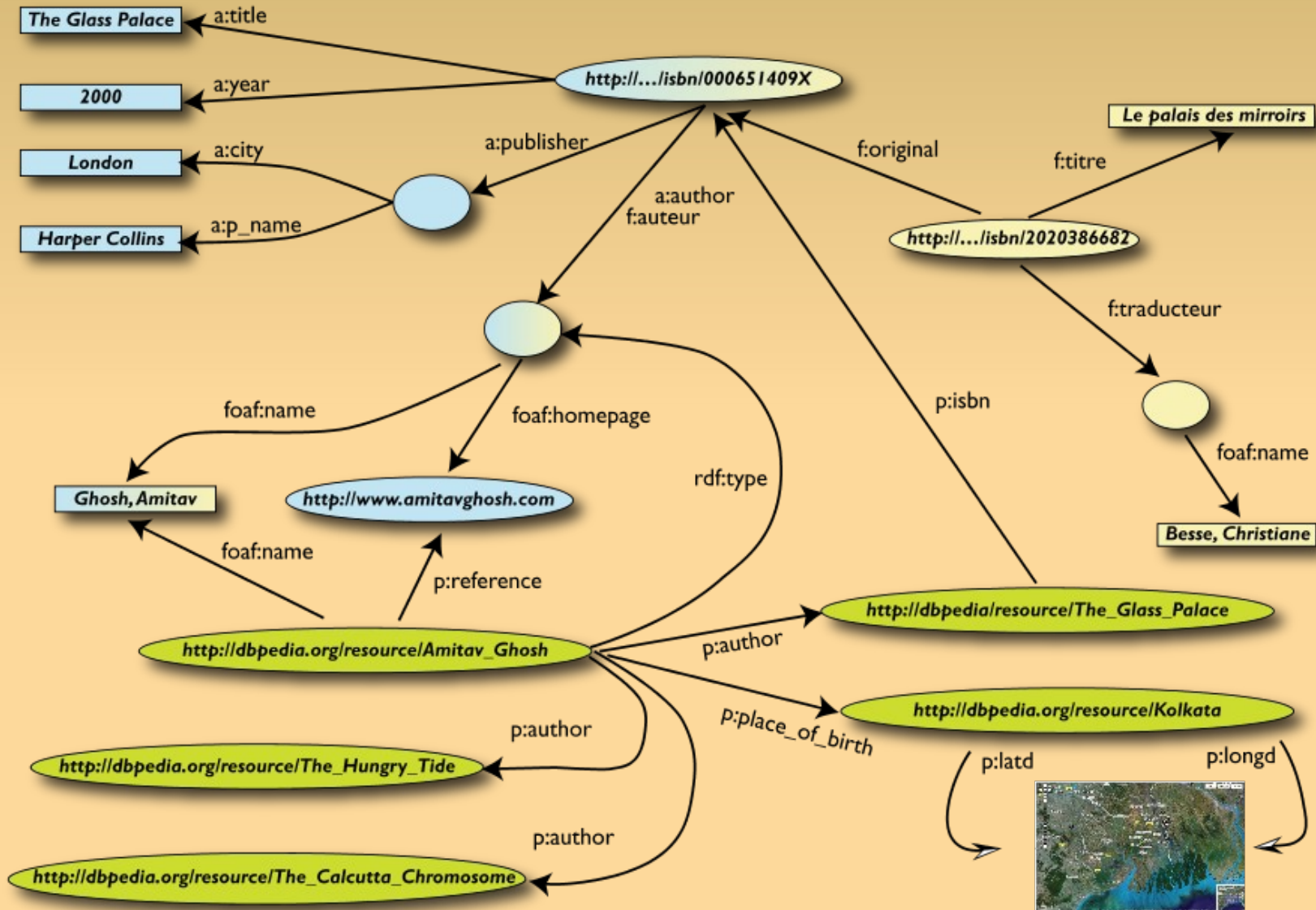
LOD cloud: One big merged graph

Why RDF and RDFa?

- Allow any data to be published

- Data provider can define new terms and relations as necessary

- Standard representation format for all data requiring only one interpreter

An RDF vocabulary example: SKOS

- ## Simple Knowledge Organization System
  http://www.w3.org/TR/skos-reference

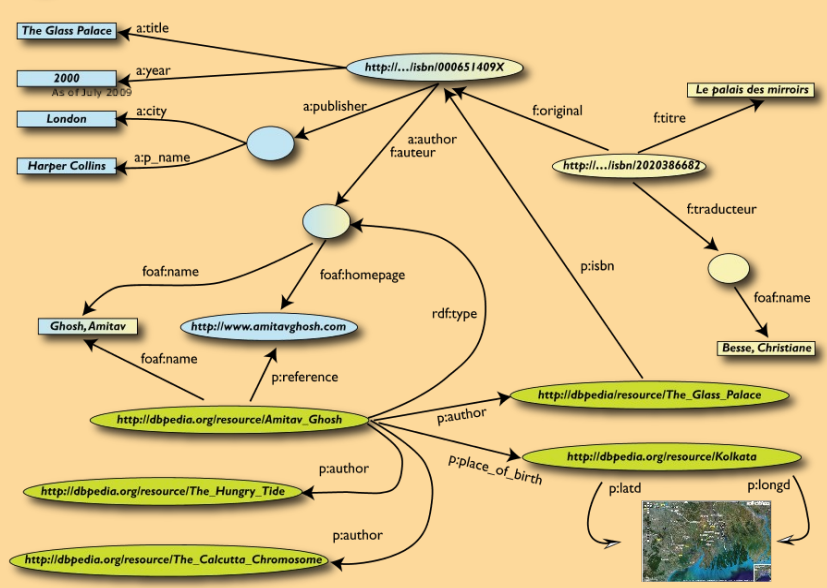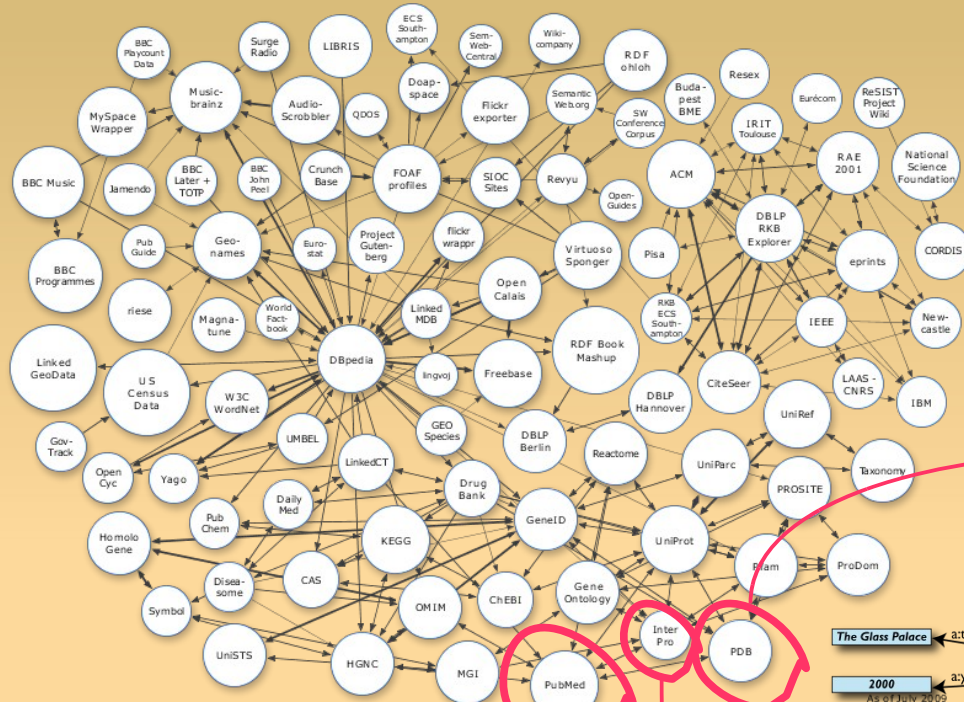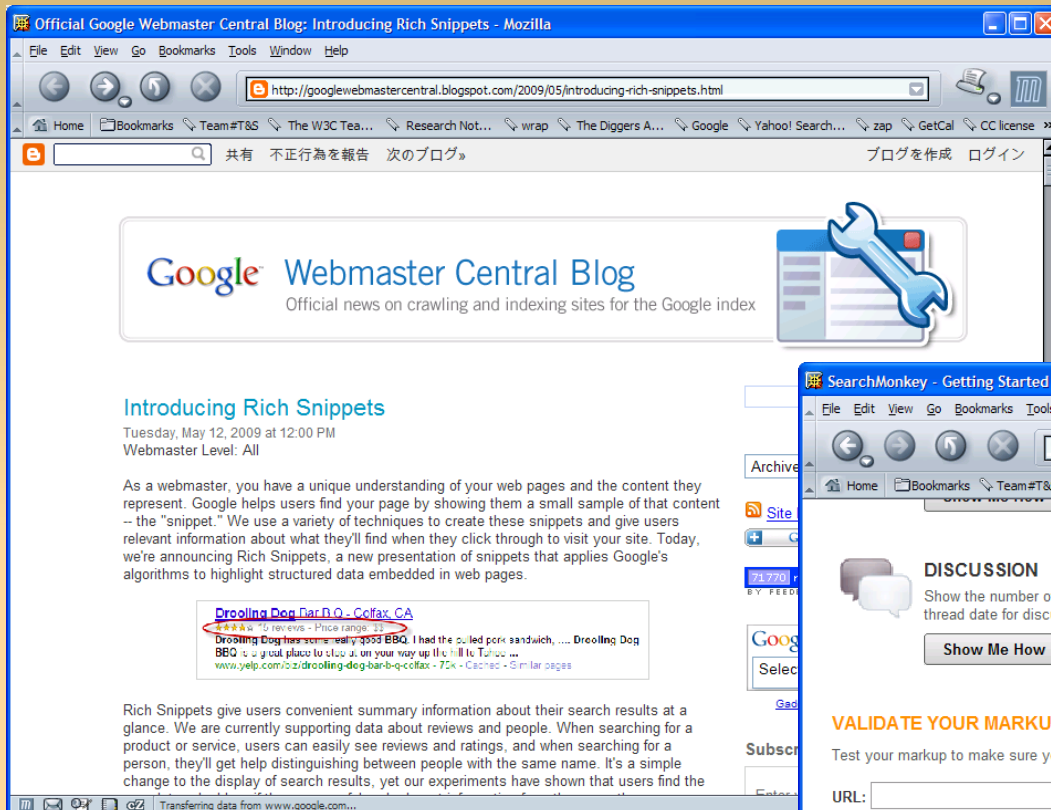- ## A vocabulary for declaring term taxonomies
  broader, narrower

- ## A vocabulary for declaring human-readable labels *in multiple languages* for each term
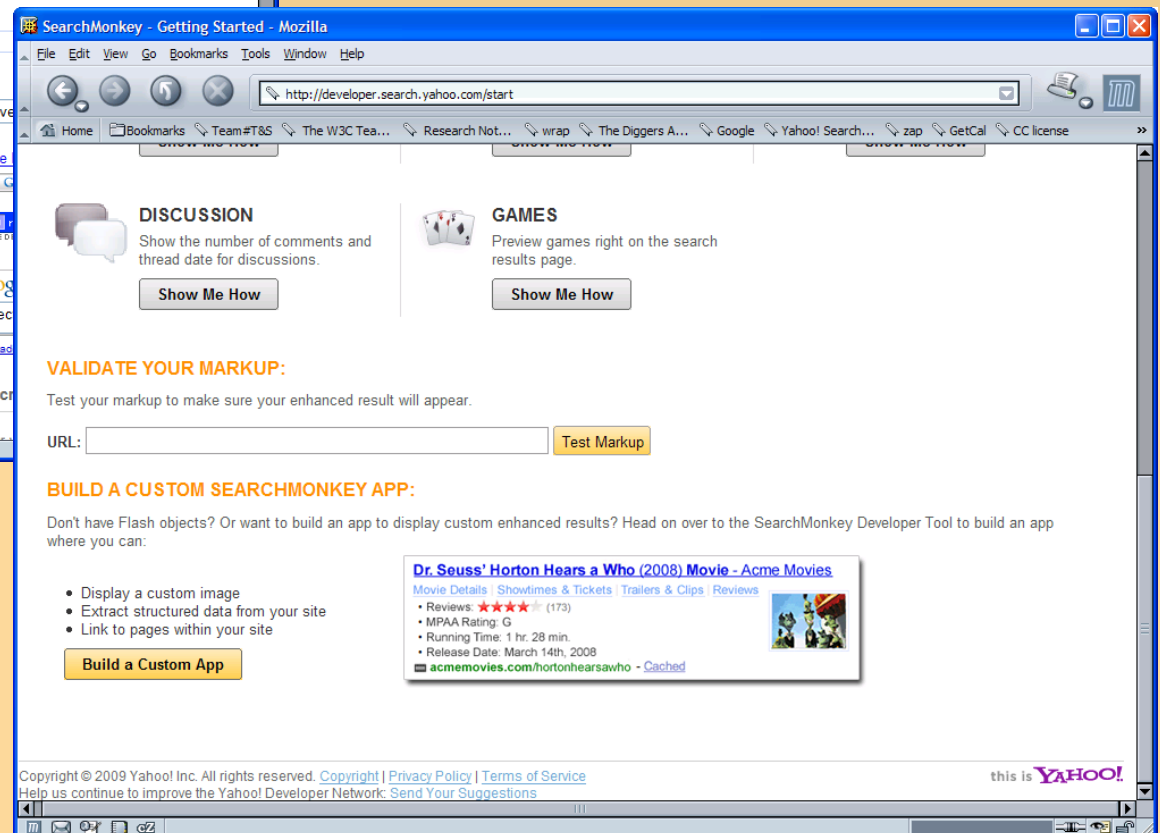
  > skos:prefLabel " 東 "@ja-Hani ;
  > skos:prefLabel " ひがし "@ja-Hira ;
  > skos:altLabel " あずま "@ja-Hira ;
  > skos:prefLabel " ヒガシ "@ja-Kana ;
  > skos:altLabel " アズマ "@ja-Kana ;
  > skos:prefLabel "higashi"@ja-Latn ;
  > skos:altLabel "azuma"@ja-Latn .

RDFa usage example:
Enhanced Search Results

Google *rich snippets*

Yahoo! *SearchMonkey*

RDFa Data Publishing examples

Bitmunk:
digital media commerce

London Gazette:
official and regulatory information re-use

The RDFa HTML5 Challenge

- RDF uses URIs to identify objects, terms and relationships

- RDF/XML uses XML namespaces to shorten these URIs to acceptable XML element and attribute names

- RDFa (XHTML) also uses XML namespaces to shorten these URIs for compactness (and ease of authoring)

- Not yet consensus on namespace mechanisms for HTML5

RDFa Working Group

- Proposed Working Group to address RDFa use inside HTML5

- W3C Advisory Committee Call for Review: *Semantic Web Activity Proposal; RDFa Working Group*

- Call for Review ends 26 January

Many Real Semantic Web Examples

- RPI Data-gov wiki

  http://data-gov.tw.rpi.edu/wiki/The_Data-gov_Wiki

- Contributed Case Studies and Use Cases

  http://www.w3.org/2001/sw/sweo/public/UseCases/

- Collection of more examples

  http://www.w3.org/People/Ivan/

- Semantic Web Challenge

  http://challenge.semanticweb.org/

Acknowledgements

- These slides are available on:

  http://www.w3.org/2010/Talks/0122-RDFa-rrs

- Thanks to Ivan Herman,

  W3C Semantic Web Activity Lead

- And the Semantic Web community

  http://www.w3.org/2001/sw/

  http://www.w3.org/standards/semanticweb/