Towards the (multilingual?) Data on the Web

Ivan Herman

W3C
What we have today: technologies
What we have today: lots of datasets
What the community needs is more deployment

- use cases
- more data
- more linked data
- etc.

It is important that the underlying technology would be seen as stable
W3C’s immediate plans

• Not to concentrate on new technology specifications
• Instead, look at the deployment issues
  • vocabulary definition, usage
  • outreach to different data formats
Vocabulary definitions

• The W3C Community Group structure gives an excellent environment to build vocabularies
  
  • good example: Open Annotation CG

• We would like to greatly extend this practice, possibly offering other tools (e.g., hosting of vocabularies)

• We are considering setting up some sort of a registry with metadata on the vocabularies

  • what would be a good set of metadata on the usability of a vocabulary in a multilingual environment?
Vocabulary validation

• Discussing the possibility of a workshop on vocabulary validation
  • “structural” validation against some schema-like definition
  • “quality” validation on data values, etc.
• Issue: how would one validate multilingual vocabularies?
Reaching out to other types of data

- *Data on the Web* is the really important thing
- Data may be in other formats: table, CSV, etc.
- There is a disconnect between:
  - the Linked Data and the more general Data on the Web worlds
  - the Web Developers’ community and the Linked Data world
Open Data on the Web
23 - 24 April 2013, Google Campus, Shoreditch, London

Call for Participation

"Analysts believe an open data strategy should be a top priority for any organization that uses the Web as a channel for delivering goods and services. Open data strategies support outside-in business practices that generate growth and innovation."


Whether your business has an open data strategy today, or you want to learn from global experts how to publish or consume open data, the World Wide Web Consortium (W3C), the Open Data Institute and the Open Knowledge Foundation invite you to join a discussion about realizing the promise of open data.

Participants will help prioritize W3C’s agenda in the area of data on the Web at a time when a number of working groups have either completed or are close to completing their charters. The Workshop, hosted by Google, is open to all and will take place in the heart of London’s Tech City of Shoreditch.

Main Topics

From Toronto to Taipei, from Washington to Wellington via Westminster, data hubs are being created by public...
Questions arising for such a workshop

- What about the multilinguality of non-RDF, non-Linked Data?
  - how to check
  - how to create
  - is there any way to manage that properly?
  - reaching out to other types of data across languages?
Looking forward to the discussions!