Vocabulary and Extensions

Peter Mika
Yahoo!
Agenda

• What is schema.org today?
• Solutions for extensions
• Developing extensions
• Case studies
  – Learning Resources Metadata Initiative (AEP)
  – Technical Publishing (Microsoft)
  – rNews (IPTC)
  – GoodRelations (Martin Hepp)
• Discussion
What is a schema?

• A set of classes
  – Each with a set of attributes
    • Each attribute has an expected type (range)
      – The value may be a primitive type OR another item of given type(s)
      – Primitive types: Boolean, Date, Number, Text

• An attribute may apply to multiple classes (domains)
  – Same attribute used in multiple classes

• Inheritance among classes
  – Multiple inheritance on some classes

• Nothing else!
  – E.g. no required attributes, no cardinalities etc.
Current vocabulary

- Represents common interests of Web search users
  - Creative works: CreativeWork, Book, Movie, MusicRecording, Recipe, TVSeries ...
  - Non-text objects: AudioObject, ImageObject, VideoObject
  - Event
  - Person, Organization
  - Place, LocalBusiness, Restaurant ...
  - Product, Offer, AggregateOffer
  - Review, AggregateRating

- Not covered
  - Long tail, e.g. antique coins or banana stickers
  - Domains of expert knowledge, e.g. biochemistry
Defining and using extensions

• Specialization
  – Supported by schema.org

• Vocabulary mixing
  – Not supported by schema.org

• RDF/OWL-based extensions
  – Not supported by schema.org
Specialization

• Extend an existing class, property or enum by appending a slash (/) and the more specialized term
  – Person/Engineer
  – Person/Engineer/ElectricalEngineer
  – musicGroupMember/leadVocalist
  – EBook/KindleBook

• Search engines treat such class/property the same as the base class/attribute
  – Search engines do not understand the meaning of the extension itself
Vocabulary mixing

• Create and host your own schema
• Mix properties from different namespaces

```html
<div itemscope itemtype="http://schema.org/Person">
  <span itemprop="name">Jane Doe</span>
  <span itemprop="http://xmlns.com/foaf/0.1/age">42</span>
</div>
```

• Limitations
  – No co-typing in microdata
• Additional attributes are ignored by search engines
RDF/OWL-based extensions

• schema.org ontology is available in OWL
• Extend schema.org using subClassOf/subPropertyOf or map to other ontologies using equivalentClass

• Not supported by search engines
Discussion forum

• W3C Web Schemas Task Force
  – Part of the Semantic Web Interest Group (SWIG)
  – Chaired by R.V. Guha
  – Open to the public
  – No patent commitment
  – public-vocabs@w3.org

• Replaces existing discussion forum
Process

• Develop and discuss
  – Gather broad industry support or adoption
• Release
  – Waiver for patent rights (in line with schema.org)
  – CC BY (attribution)
• Incorporate into schema.org
• Support in search engines and other tools
Discussion