Uncertain about your data or provenance?

Use

W3C PROV

To model your uncertainty!

What is W3C PROV?

Since April 30th 2013, the PROV family of specifications is a W3C Recommendation!

This means that PROV is the new standard for expressing and interchanging provenance on the Web.

PROV includes a data model, constraints and two serializations!

Modeling Uncertainty

In PROV ➔ Bundles model coarse-grained provenance of provenance (PoP)

In UP ➔ attributes model fine-grained uncertain provenance and provenance of uncertain things.

prefix up <http://semweb.mmlab.be/ns/up/>

Lightweight approach

Attributes for modeling of uncertain provenance

up:assertionConfidence: confidence (between 0 and 1) assigned to a provenance statement by the asserter

up:assertionType:
  - up:HumanAsserted
  - up:Incomplete
  - up:MachineGenerated
  - up:Future
  - up:MachineCollected
  - up:Trusted
  - up:Complete
  - up:Untrusted

Example 1: Provenance Reconstruction

entity(ex:document1)
entity(ex:document2)
entity(ex:document3)
wasDerivedFrom('d1', ex:document3, ex:document1, [up:assertionConfidence="0.6", up:assertionType="up:MachineGenerated"])
wasDerivedFrom('d2', ex:document3, ex:document2, [up:assertionConfidence="0.9", up:assertionType="up:HumanAsserted"])

Example 2: Named Entity Recognition

entity(ex:document)
entity(ex:namedEntities, [prov:type="prov:Collection"])
activity(ex:NER)
wasDerivedFrom(ex:namedEntities, ex:document, ex:NER)
eentity(dbpedia:New_York)
eentity(dbpedia:Joe_Biden)
eentity(ex:New_York, [up:contentConfidence="0.6"])
eentity(ex:Joe_Biden, [up:contentConfidence="0.8"])
specializationOf(ex:New_York, dbpedia:New_York)
specializationOf(ex:Joe_Biden, dbpedia:Joe_Biden)
hadMember(ex:namedEntities, ex:New_York)
hadMember(ex:namedEntities, ex:Joe_Biden)

Visit http://semweb.mmlab.be for our other Web-related research!