2015-02-23 breakout: Reconciliation services - Things to things

UNEDITED NOTES

facilitator: Rob Sanderson

Things to Things

Dynamic

- provenance
- annotations as nanopub

Degrees of sameness

Outside our domain?

- no need to be perfect

Event stream

Use case – authorities

Deleted resources – tombstones

Entities joining/splitting

Alternate identities

Services/tools

- certainty and risk aversion
- more data = better, = more uses ?

(new group)

Sameness?

Reconcile at work level

Context – what we use the data for!

Reconcile most discriminant attributes

Similarity pocky ? for sameness

- certainty?
- risk aversion
- discovery vs. inference/data management
- schema:samas – avoiding the alt itoces ?

Indexing & uncertainty

- perfect as enemy of good enough
- sameAs external link as vote for URI
- distinctness different for different types of entity

special cases are special: Bible, Mark Twain

(new group)

Services like ref. cleaning

- identify auth. we trust
    - e.g., VIAF
- map between authorities
    - devil in details for mapping
    - mechanism to address errors in auth!
    - disambiguate

"Authority" – what do we really mean?

- Wikipedia/media wiki as authority

- Open URI resolver?
- endorsement of approach by adopters
- divide and conquer easy bits
- start with under-served domains
- cheaper going forward
- market will decide on URIs to use
- vocab/facet reconciliation
- places & dates, concepts with multilingual

(new group)

Ontology alignment

Global auth & global contexts

Dynamic entities : join/split

Pseudonym – Europe vs. US

Local assertion = local URI

Pre-existing identifiers

Link all the things

Subjects broader issue than people

- sameness degree?
- certainty?

Tools to produce the data?

- languages and character sets
- different to strings to things
- service to publish alignments
- provenance important for trust