Integration “In the Large”

Position paper for the W3C Workshop on Data and Services Integration

Ora Lassila & Ian Oliver
Nokia Location & Commerce
2011-10-21
Some background…

• This is a **position paper**, mostly about questions, not answers

• Our main goals are the following:
  A. to present some **observations** about system and data integration (based on both our own research\(^1\) as well as product development\(^2\) experience)
  B. incite some **discussion** on how to move forward and produce solutions to what (at least) we perceive as problems

---

1. at Nokia Research Center
2. as part of the team developing the platform for Nokia’s “Ovi Services”, a suite of mobile online services
“Existential Crisis” of the Semantic Web...?

• Semantic Web was conceived as “integration and interoperability” technology

• It is all grown up now: The main technical pieces are in place

  BUT...

• What about our dream, our vision of being able to ontologically model the world? Has it been realized, can it be realized?
“Existential Crisis” of the Semantic Web...?

• Prescriptive approaches to the world are known to fail
  – rather, the Semantic Web is very much intended to be **descriptive**
• “Global ontology” not achievable
  – the broader the scope, the **weaker** or more complex the ontology
  – (some of us always knew that)
• Not just a technical challenge...
## Hierarchy of information scales (cf. mapping)

1. **Mapping scalar objects**, units of measure, etc.
   - e.g., UNIX date → ISO 8601 date
   - Mostly syntactic, yet often offered as “semantic transformations”
   - **THIS IS NOT A PROBLEM!**

2. **Mapping structured objects**
   - e.g., ovi:Person → facebook:Person
   - Doable, particularly if semantics on both sides are already a good match, still this may lead to “subsetting”, making round-trips difficult

3. **Mapping application data models**
   - or ontologies) onto other applications’ models
   - e.g., Ovi Services → Facebook
   - Achieving bijective and transitive mappings much harder, also much of the semantics is embodied in applications’ “business logic”

4. **Mapping entire cultural “contexts”**
   - e.g., US → France → Finland
   - note: finland:Café ≠ france:Café
   - Is it even possible…? Very difficult, but perhaps not entirely hopeless

---

Attempts to solve the problem

• Many attempts to formalize information flow, mapping of semantics, etc. (unsurprisingly often based on category theory)
  − Barwise & Seligman, Goguen, Gärdenfors, Sowa, etc.
  − provides the mathematical basis in that it clearly allows us to understand why things are hard…
  − unclear how the real world fits in

• Perhaps more close to Semantic Web technologies, work on ontology mapping and ontology matching is promising
  − translating ontologies is one of the key mechanisms that allows Semantic Web to work in the first place
  − unclear how this works “in the large”
Integration experiment: M3 [Oliver 2009]

• Larger systems constructed from very loosely coupled smaller components
  – components have their own local semantics, own logic
  – free to “interpret” data from other sources using own local semantics (M3’s notion of “semantic mapping”)
  ⇒ not “real data integration” in the commonly accepted sense

• No notion of an “application”, just data
  – however, data (and its semantics) not enough, we also need formalization of “actions” (i.e., processing)

• This is possibly a more natural way of developing “semantically aware systems” (cf. Goguen, Barwise, Seligman, et al.)
Where do we go from here?

• Current Web architecture, especially the “Semantic Web stack”, offers a good basis for building a higher-level framework:
  – representation: RDF, OWL
  – mapping: RIF, SPARQL, GRDDL
  – what about “services”…?

• Can a formal framework be defined that addresses issues of semantic mapping and reconciliation of differences in semantics?
  – ostensibly, “yes” – the mathematics is difficult, though
  – also practical problems (social, organizational, etc.)
One final question…

• (The most important one, in my mind)

• Can we move information systems closer to **how humans behave**?
  – partial “understanding” between parties, middle ground between complete interoperability and catastrophic failure
  – local spaces, local understanding, partial information interchange?