How can DDI make the most of RDF?

EDDI15 Conference – Franck Cotton / Guillaume Duffes - Insee

Contents

1. Brief overview of RDF
2. DDI and RDF: the premises
3. RDF vs XML: any value added for DDI?
4. DDI4 and RDF
5. RDF in the statistical community
6. RDF and DDI at Insee
7. Conclusion

Brief overview of RDF

- RDF is defined in the W3C specification as follows: "*Resource Description Framework (RDF) is a foundation for processing metadata; it provides interoperability between applications that exchange machine-understandable information on the Web*.[…] *The broad goal of RDF is to define a mechanism for describing resources that makes no assumptions about a particular application domain, nor defines (a priori) the semantics of any application domain.*"

Brief overview of RDF

- The basic model consists of three types of elements represented below:

DDI and RDF: the premises

- In its early versions, DDI already dealt with vocabularies, e.g the Dublin Core elements which have RDF as main implementation today.
- Imported ***a priori*** as an external XML Schema (own namespace). Limits…
- What if another framework allowed to reference ***a posteriori*** as many vocabularies as needed → RDF

- XML serialised as XML Schemas → strongly typed and binding...
- But also a closed model: if one property is missing, it cannot be added from another XSD schema.
- The only solution would be: extend the schema → backward compatibility impossible.

RDF vs XML: any value added for DDI?

- RDF on the contrary allows to combine different vocabularies.
- If a property is missing, then certainly a more specific vocabulary exists and defines it, use it!
- Well known and defined RDF vocabularies are now legions: Dublin Core, FOAF, SKOS, PAV, PROV, etc.

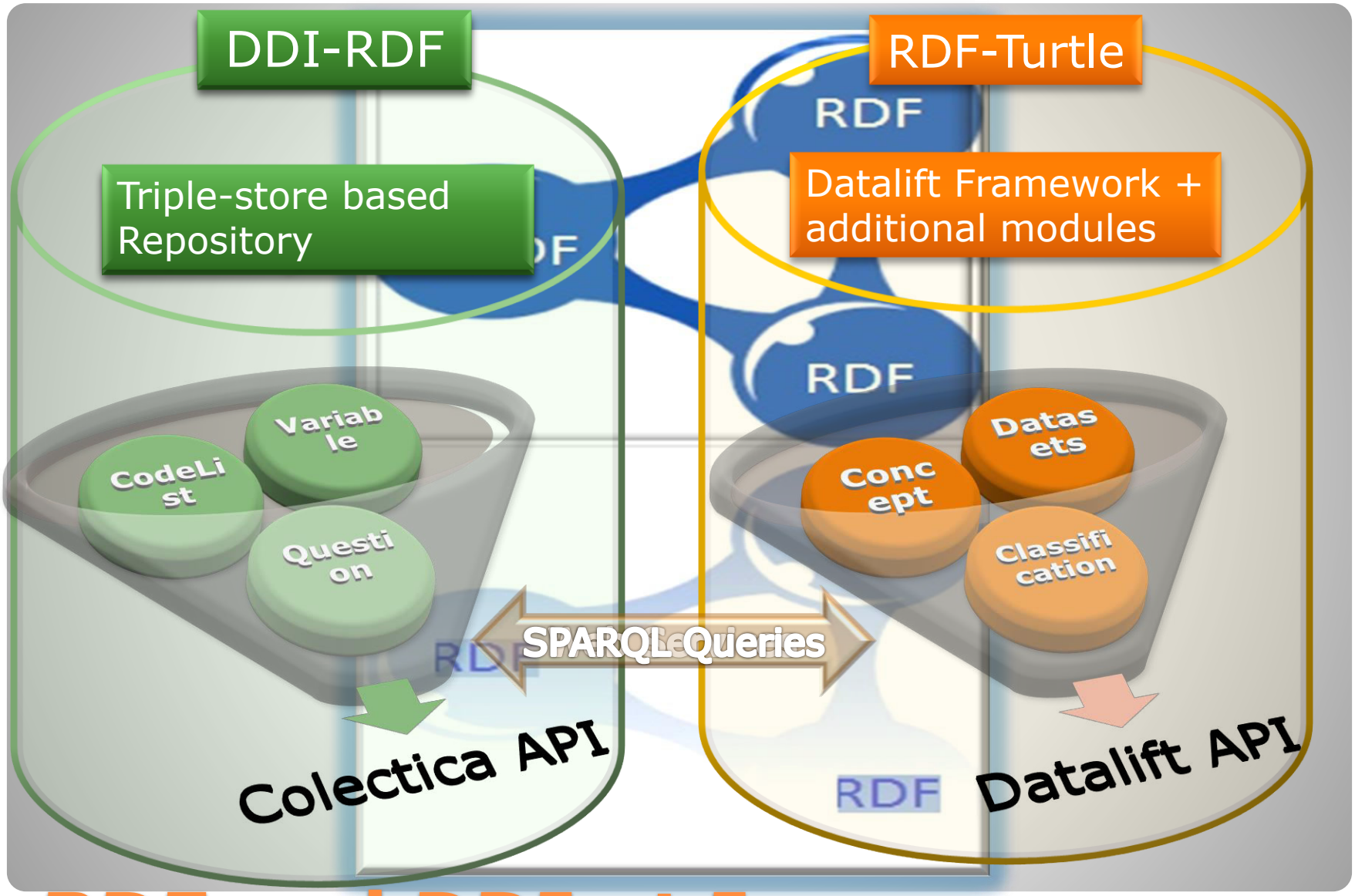
RDF vs XML: any value added for DDI?

DDI4 and RDF

- DDI4 is documented according to functional views.
- DDI4 should make the most of existing vocabularies in functional views (e.g a concept in DDI should be a SKOS concept)
- RDF comes along now with RDF databases called Triple Stores. Mature commercial or open-source solutions exist.

RDF in the statistical community

- Various initiatives in the statistical community are under progress:
  - HLG-MOS (UNECE): Work on Linked Open Metadata
  - One Work Package in the DIGICOM project (European level)
  - Continuing development of the RDF Data Cube
  - SemStats.



RDF and DDI at Insee

How to implement in RDF the following ddi:ConceptualVariable takesMeaningFrom skos:Concept?

```
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .
@prefix gsim:    <http://rdf.unece.org/gsim/v2#> .
@prefix ddi:     <http://rdf.ddialliance.org/ddi#> .

SELECT ?cncptvarlab ?cnpt WHERE
{
?cncptvar gsim:takesMeaningFrom ?cnpt .
?cnpt a skos:Concept .
?cnptvar rdfs:label ?cncptvarlab .
?cncptvar a ddi:ConceptualVariable .
}
```

Demo

RDF and DDI at Insee

Conclusion

- RDF enables interoperability between applications:
  - that deal with machine-actionable information on the Web
  - or that manage linked information internally
- RDF is then tailor-made for DDI.
- How would both DDI serialisations evolve in parallel when one is much more “pluggable” than the other?