



A constraint language for Description Set Profiles

Thomas Baker
Dublin Core Metadata Initiative

2013-09-04: Session on Profiles and Ontologies, DC-2013

2013-09-11: RDF Validation Workshop, Cambridge, MA



“Dublin Core” as a Record Format

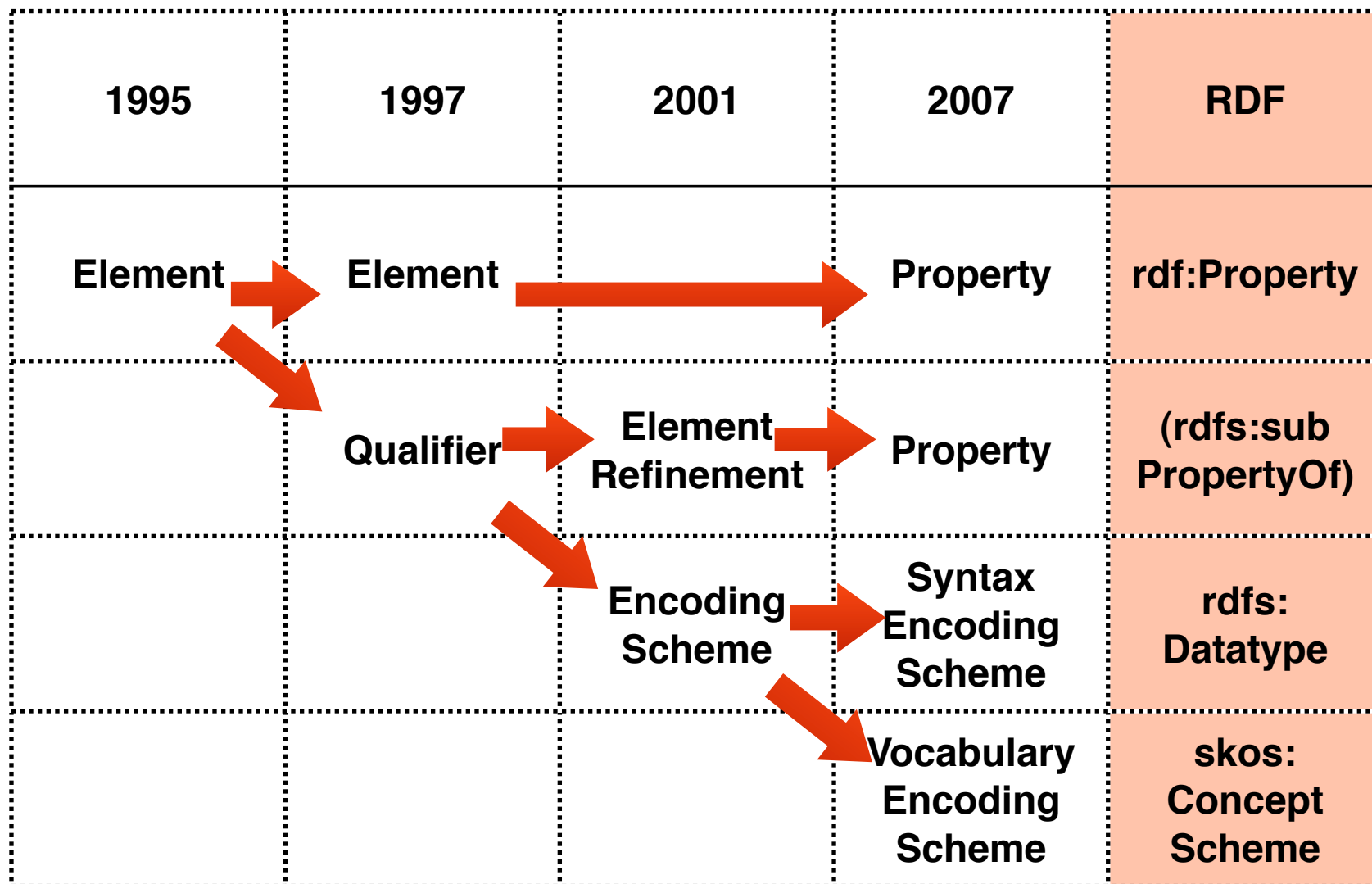
- **1995** workshop goal “simple metadata record”
- **2001** XML format for Open Archives Initiative
- **Today** many still see DC as an exchange format



“Dublin Core” as an RDF Vocabulary

- **1997** RDF organizers at DC Canberra workshop
- **1999** RDF Rec addresses DC requirements
 - DCMI Metadata Terms published as RDF properties
- **2006** Top-10 vocabulary in “Linked Data cloud”

From Record Elements to alignment with RDF





Application Profiles

- **2000:** Customize Dublin Core for specific uses.
 - Mix-and-match terms from different standards
- **But problems:**
 - Implemented incompatibly in HTML, XML, RDF...



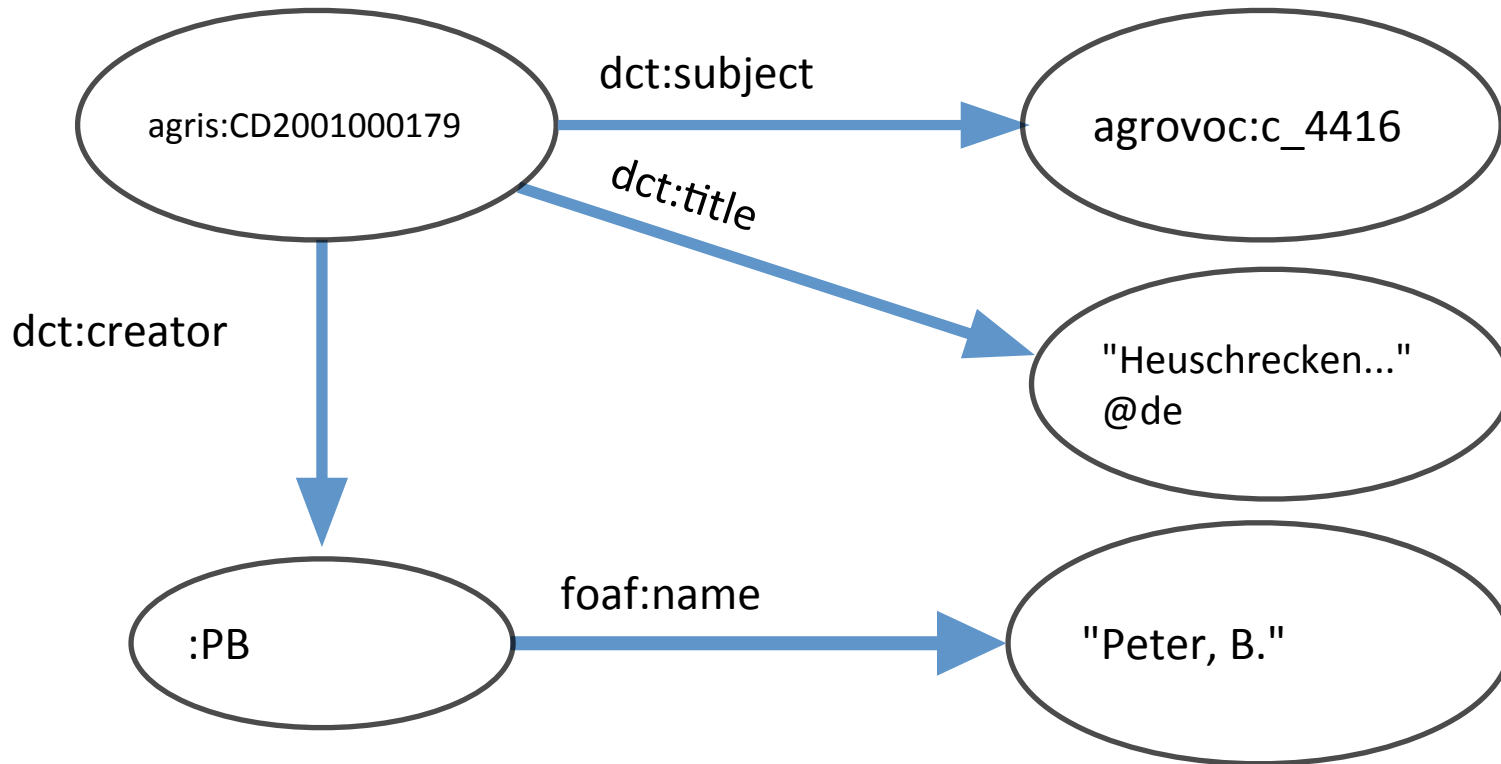
Abstract Model (DCAM)

- **2003-2007** Development of DCAM
 - Common target for mapping from HTML, XML...
 - Mapped to RDF to the extent possible

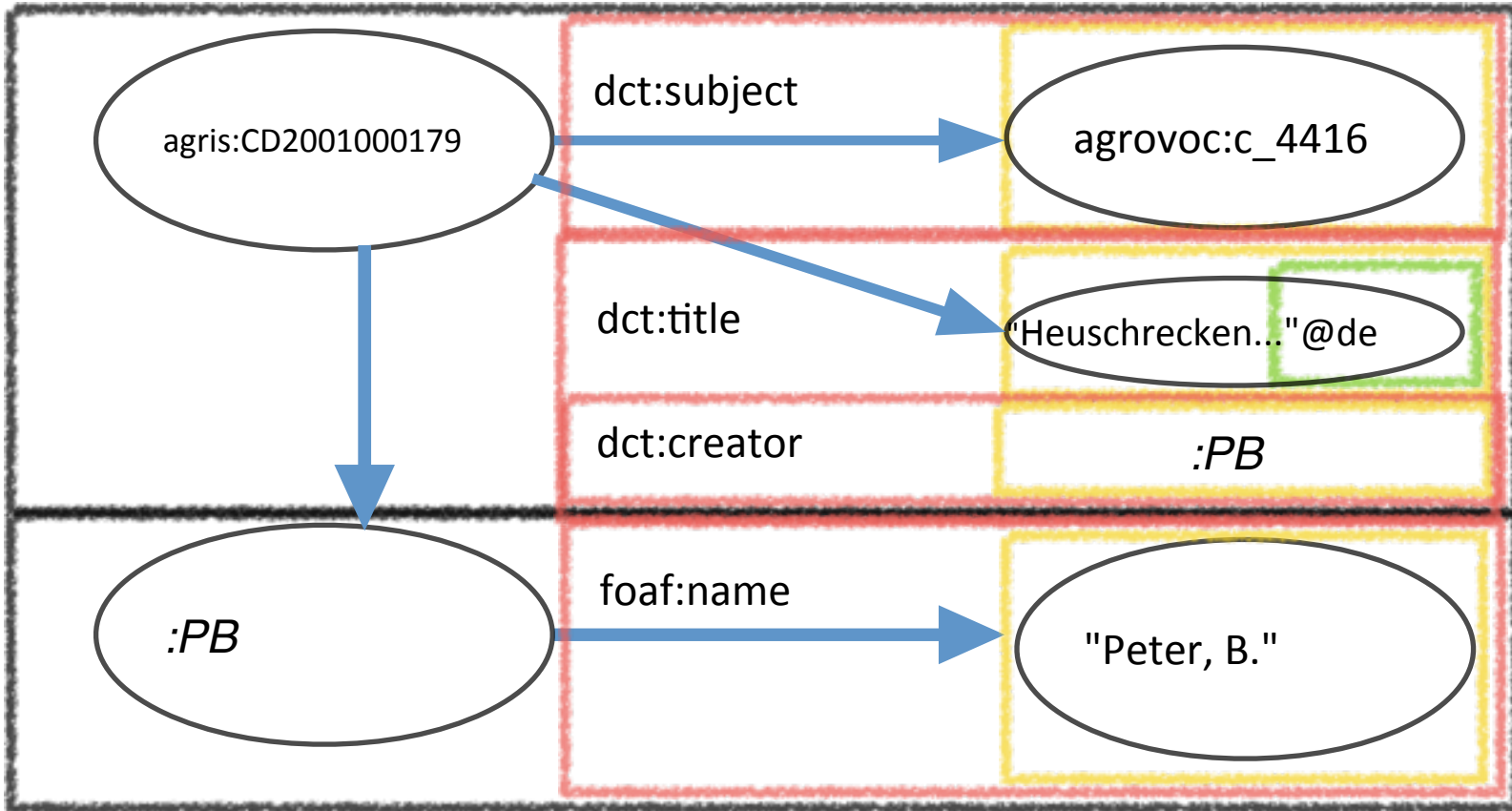


Bridging two mindsets

- Orientation to ***Record Formats***
 - Bounded sets of fields to be “filled in” with information
- Orientation to ***Graphs***
 - Unbounded webs of information connected by statements



Subject	Predicate	Object
agris:CD2001000179	dct:subject	agrovoc:c_4416k
agris:CD2001000179	dct:title	"Heuschrecken..."@de
agris:CD2001000179	dct:creator	:PB
:PB	foaf:name	"Peter, B."



Viewing a graph in terms of a “record”



agris:CD2001000179

dct:subject

agrovoc:c_4416

dct:title

"Heuschrecken"

de

dct:creator

:PB

:PB

foaf:name

"Peter, B."

Seeing components that can be validated.



Description Set

Description

Described Resource URI

Property URI

Value URI

Property URI

Value String

Lang

Property URI

Value ID

Description

Value ID

Property URI

Value String

Abstracting those components to a generalized model.



Description Set ←

Description ←

Non-literal ←

Property URI

Value URI

Value String

Lang

Vocabulary Encoding
Scheme URI

Literal ←

Value String

Lang

Grouping constructs with no equivalent in RDF.



Property URI

Described Resource URI

```
<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/" >
  <rdf:Description rdf:about="http://agris.fao.org/resource/CH2001000179">
    <dcterms:title>Heuschrecken brauchen ökologische Ausgleichsflächen</dcterms:title>
    <dcterms:subject rdf:resource="http://aims.fao.org/aos/agrovoc/c_4416" />
    <dcterms:creator rdf:nodeID="PB" />
  </rdf:Description>
  <rdf:Description rdf:nodeID="PB">
    <foaf:name>Peter, B.</my:name>
  </rdf:Description>
</rdf:RDF>
```

Value URI

Value String

**Abstract Model
components embedded
in application syntaxes**

Expressed as triples

Subject	Predicate	Object
agris:CD2001000179	dct:subject	agrovoc:c_4416k
agris:CD2001000179	dct:title	"Heuschrecken..."@de
agris:CD2001000179	dct:creator	:PB
:PBS	foaf:name	"Peter, B."



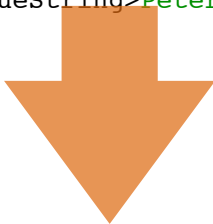
Property URI

Described Resource URI

```
<?xml version="1.0" encoding="UTF-8" ?>
<dcds:descriptionSet xmlns:dcds="http://purl.org/dc/xmlns/2008/09/01/dcds-xml/">
  <dcds:description>
    dcds:resourceURI="http://agris.fao.org/resource/CH2001000179">
      <dcds:statement dcds:propertyURI="http://purl.org/dc/terms/title">
        <dcds:literalValueString>Heuschrecken brauchen ökologische Ausgleichsflächen</dcds:literalValueString>
      </dcds:statement>
      <dcds:statement dcds:propertyURI="http://purl.org/dc/terms/subject"
        dcds:valueURI="http://aims.fao.org/aos/agrovoc/c_4416k">
        <!-- Reference to value using local identifier -->
      </dcds:statement>
      <dcds:statement dcds:propertyURI="http://purl.org/dc/terms/creator"
        dcds:valueRef="PB" />
    </dcds:description>
    <!-- Description of value using local identifier -->
    <dcds:description dcds:resourceId="PB">
      <dcds:statement dcds:propertyURI="http://xmlns.com/foaf/0.1/name">
        <dcds:literalValueString>Peter, B.</dcds:literalValueString>
      </dcds:statement>
    </dcds:description>
  </dcds:descriptionSet>
```

Value URI

Value String



Expressed as triples

Subject	Predicate	Object
agris:CD2001000179	dct:subject	agrovoc:c_4416k
agris:CD2001000179	dct:title	"Heuschrecken..."@de
agris:CD2001000179	dct:creator	:PB
:PBS	foaf:name	"Peter, B."

2008: Description Set Profile Constraint Language (DSP)



```
Description Set [template]
  Description [template]
    Statement [template]
      Property [constraint]
        <http://purl.org/dc/terms/subject>
      VocabularyEncodingSchemeURI [constraint]
        <http://aims.fao.org/aos/agrovoc>

    Statement [template]
      Property [constraint]
        <http://purl.org/dc/terms/title>
      MinOccurs [constraint]
        1
      MaxOccurs [constraint]
        1

    Statement [template]
      Property [constraint]
        <http://purl.org/dc/terms/creator>

  Description [template]
    Resource Class [constraint]
      <http://xmlns.com/foaf/0.1/Person>
    Statement [template]
      Property [constraint]
        <http://xmlns.com/foaf/0.1/name>
```

- Data using this Description Set Profile describes:

Resource with *exactly one title*,

Value of *subject* is an *AGROVOC URI*.

Authors are members of the *class Person* and have *names*.



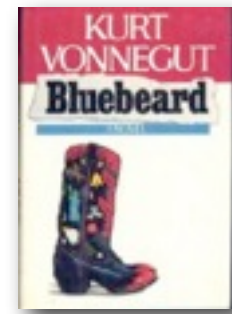
Templates and Constraints

- **Description Template constraints:** Identifier, Standalone, Minimum occurrence, Maximum occurrence, Resource class membership
- **Statement Template constraints:** Minimum occurrence, Maximum occurrence, Type
 - **Property constraints:** Property list, Sub-property
 - **Literal Statement Template constraints:**
 - **Literal value constraints:** Literal list, Literal language, Literal language list, Syntax Encoding Scheme, Syntax Encoding Scheme list
 - **Non-literal Statement Template constraints:**
 - **Non-literal value constraints:** Description template reference, Class membership
 - Value URI constraints: Value URI occurrence, Value URI list
 - Vocabulary encoding scheme (VES) constraints: VES occurrence, VES list
 - Value string constraints: Minimum occurrence, Maximum occurrence (plus all other constraints that apply to literal values - see above)

Work

```
{
  "items" : [
    {
      "id": "http://bibfra.me/work/u2",
      "title" : "Bluebeard",
      "titleRemainder" : "a novel",
      "type" : "Book"
      "classification" : "025.3 M465f 2008",
      "author" : "http://bibfra.me/auth/person/kurt\_vonnegut",
      "subject": [ "http://bibfra.me/auth/subject/foo",
                  "http://bibfra.me/auth/subject/bar" ],
      "language": "english",
      "summary": "Broad humor and bitter irony collide in this fictional
autobiography of Rabo Karabekian, who, at age seventy-one, wants to be left alone on his
Long Island estate with the secret he has locked inside his potato barn. But then a
voluptuous young widow badgers Rabo into telling his life story—and Vonnegut in turn
tells us the plain, heart-hammering truth about man’s careless fancy to create or destroy
what he loves.",
      "hasInstance": "http://bibfra.me/instance/u2-1"
    }
  ],
  {
    "id": "http://bibfra.me/instance/u2-1",
    "type": "HardcoverBook",
    "edition": "1st",
    "volumes": "1",
    "pages": "300",
    "date" : "1987-09-02",
    "publisher" : "http://bibfra.me/auth/org/delacorte\_press",
    "ISBN": "9780385295901",
    "image": "http://ecx.images-amazon.com/images/I/51vbo8zwJlL.\_SL160\_.jpg"
  }
]
```

```
{
  "id": "http://bibfra.me/instance/u2-1",
  "type": "HardcoverBook",
  "edition": "1st",
  "volumes": "1",
  "pages": "300",
  "date" : "1987-09-02",
  "publisher" : "http://bibfra.me/auth/org/delacorte\_press",
  "ISBN": "9780385295901",
  "image": "http://ecx.images-amazon.com/images/I/51vbo8zwJlL.\_SL160\_.jpg"
}
```



Instance

Templating Descriptions

Work

```

"items" : [
  {
    "id": "http://bibfra.me/w
    "title" : "Bluebeard",
    "titleRemainder" : "a nov
    "type" : "Book"
    "classification" : "025.3
    "author" : "http://bibfra
    "subject": [ "http://bibf
      "h
    "language": "english",
    "summary": "Broad humor a
autobiography of Rabo Karabekian, wh
Long Island estate with the secret h
voluptuous young widow badgers Rabo
tells us the plain, heart-hammering
what he loves.",
    "hasInstance": "http://bibfra.me/instance/u2-1"
  },
  {
    "id": "http://bibfra.me/i
    "type": "HardcoverBook",
    "edition": "1st",
    "volumes": "1",
    "pages": "300",
    "date" : "1987-09-02",
    "publisher" : "http://bib
    "ISBN": "9780385295901",
    "image": "http://ecx.imag
  }
]

```

id	URI	"must be http://bibfra.me... "
title	String	"must be one"
titleRemainder	String	
type	String	from controlled vocabulary
classification	String	
author	URI	
subject	URI	
lang	String	
summary	String	
hasInstance	URI	"one and only one"

"hasInstance": "http://bibfra.me/instance/u2-1"

```

  {
    "id": "http://bibfra.me/i
    "type": "HardcoverBook",
    "edition": "1st",
    "volumes": "1",
    "pages": "300",
    "date" : "1987-09-02",
    "publisher" : "http://bib
    "ISBN": "9780385295901",
    "image": "http://ecx.imag
  }
]

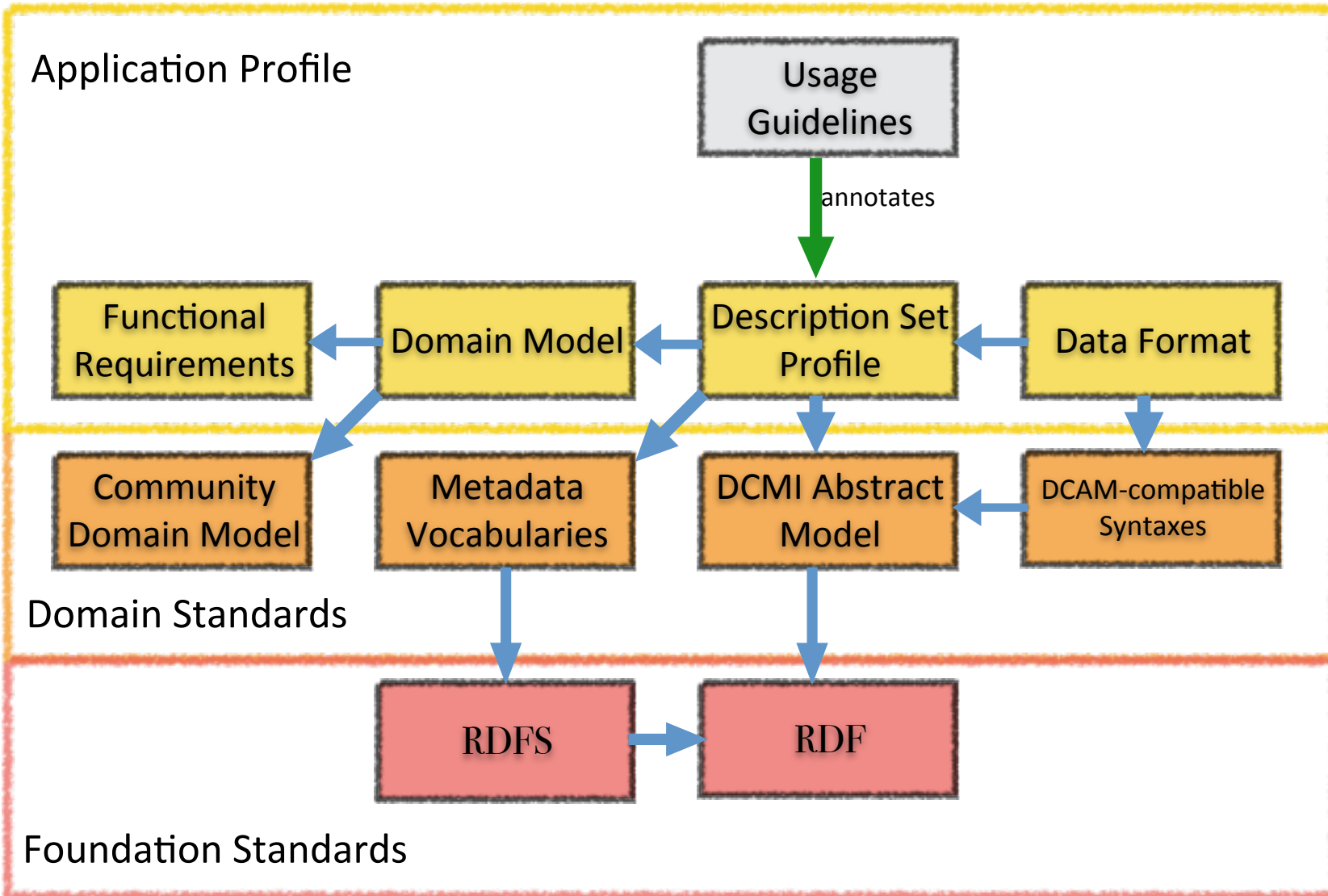
```

Instance

id	URI	"must be http://bibfra.me... "
type	String	from controlled vocabulary
edition	String	
volumes	String	integer
pages	String	integer
date	Date	use Bibframe date-time format
publisher	URI	
ISBN	String	must be valid ISBN
image	URI	

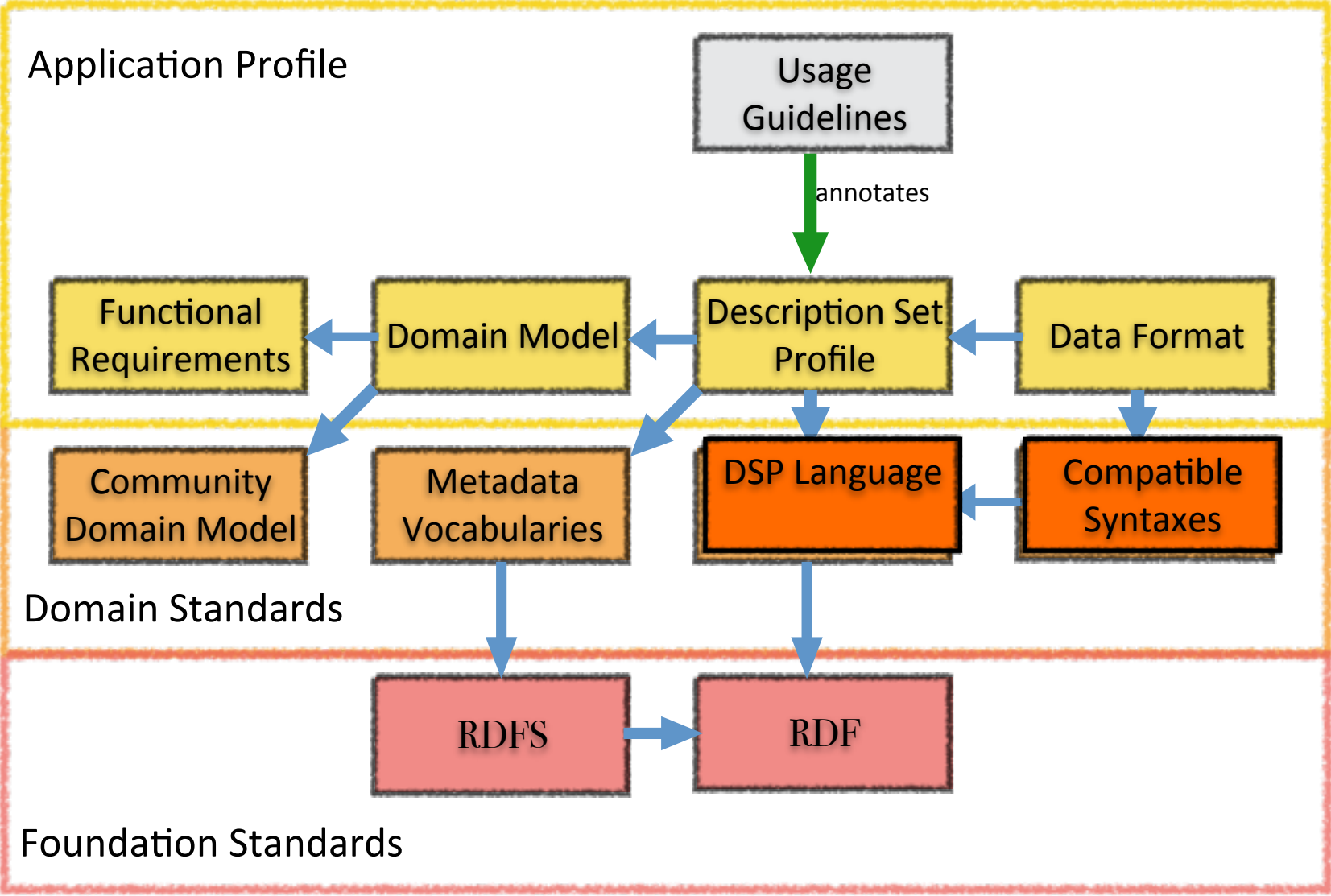


Singapore Framework for Dublin Core Application Profiles (2007)

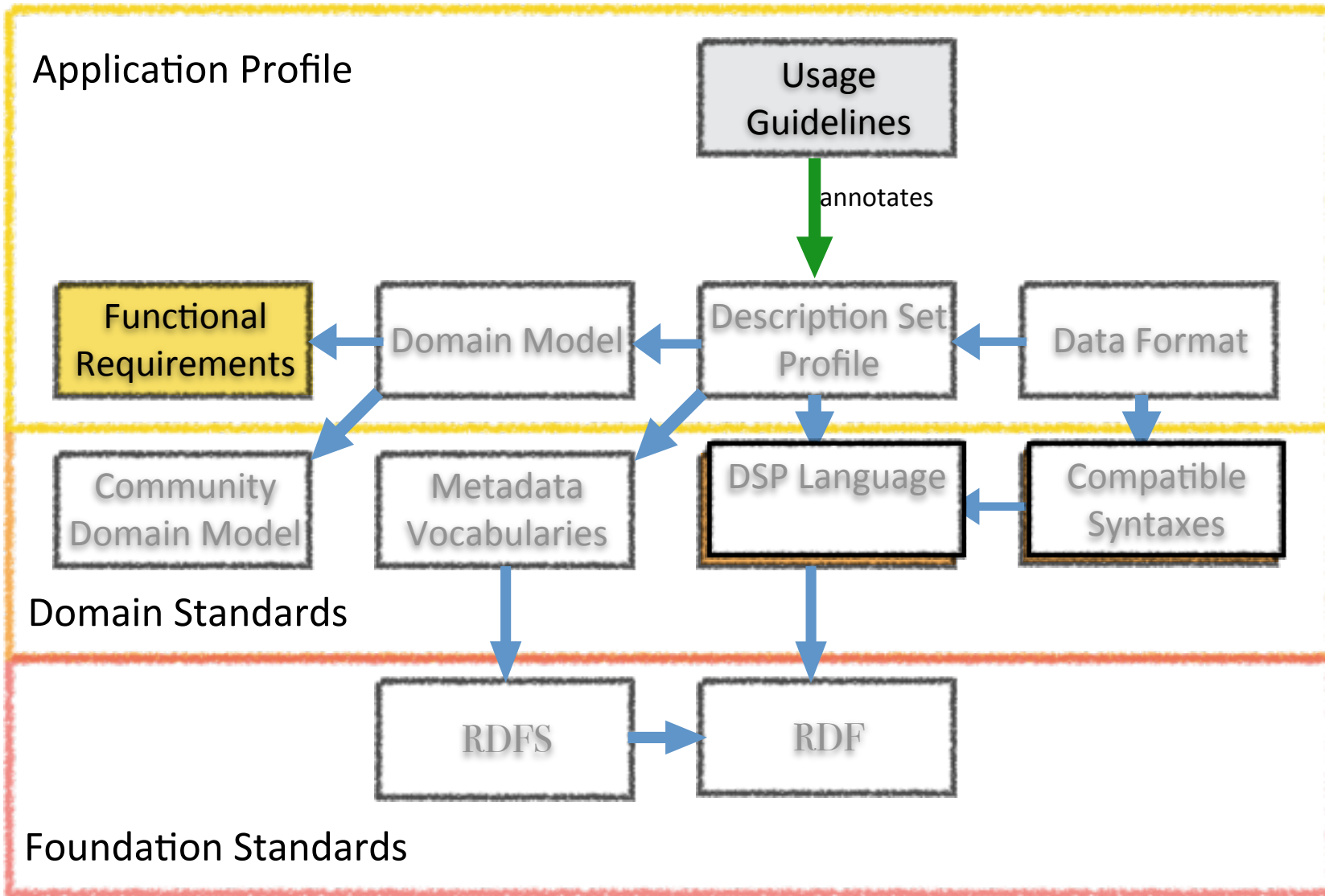




An updated Description Set Profile templating language?



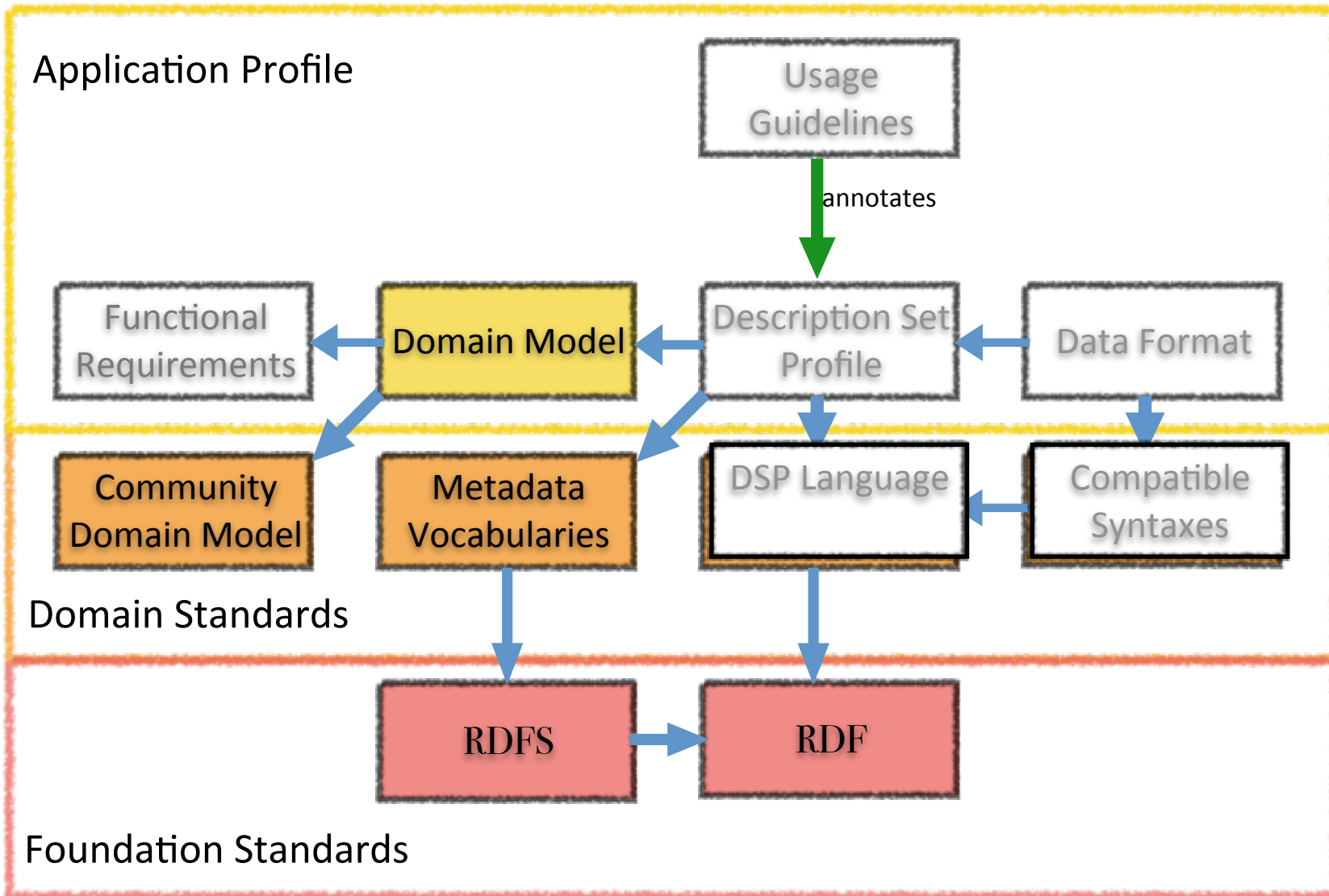
Application-specific Context





Model “reality” with a light touch

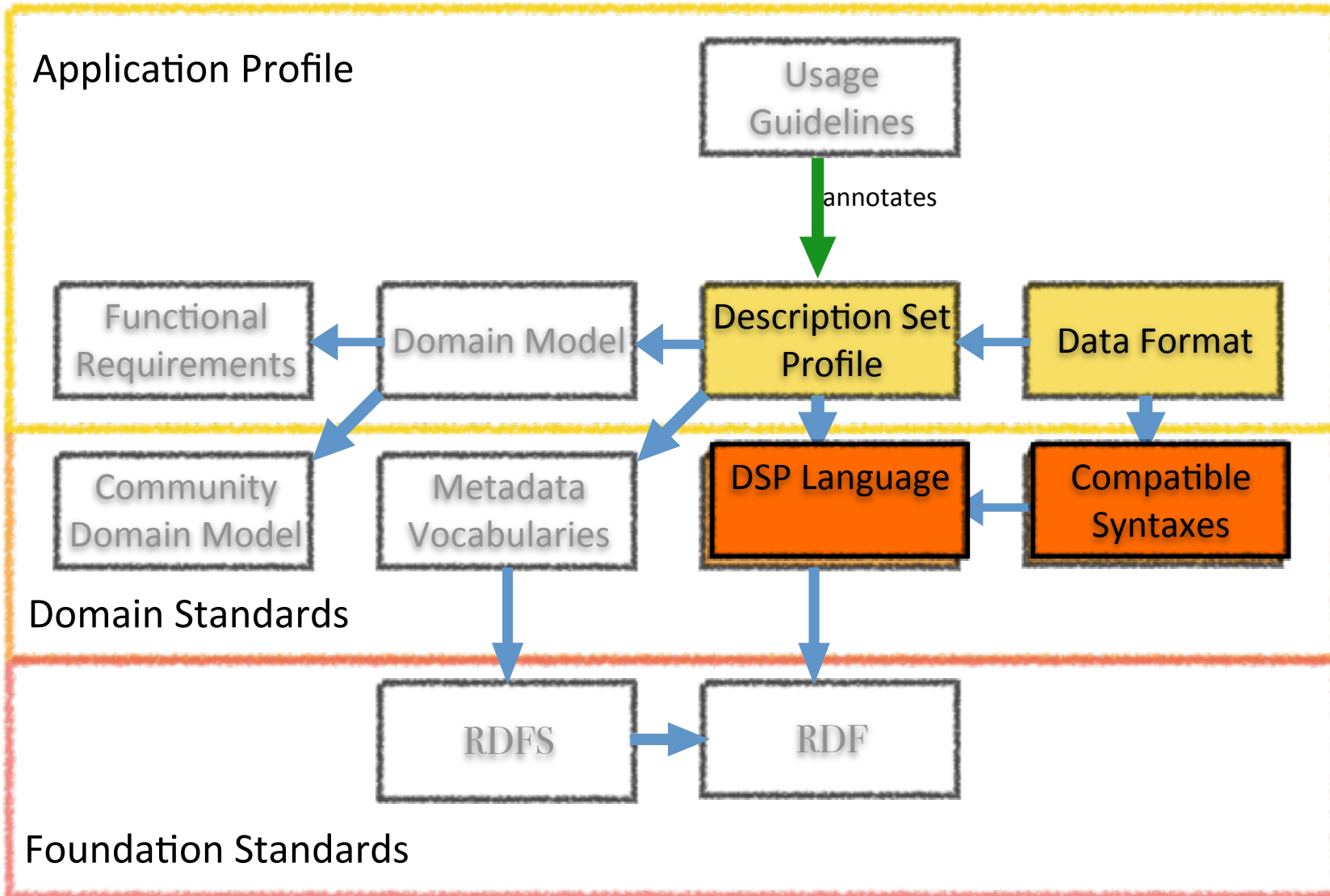
Principle of Minimal Semantic Commitment





Constrain the Data

For consistency and quality control





Good Linked Data

- Design for peaceful coexistence
 - Expect unexpected uses
- Emphasize pragmatic control of data quality
 - Constraints on data rather than sweeping generalizations about the world.
- Keep it simple
 - Simple model have larger audiences



Requirements for application profiles

- Ordinary people can read and author
- Nice if constraints are reusable by machines
- Nice if constraints usable for data validation



tom@tombaker.org