Design for simple application profiles

Tom Baker and Karen Coyle
Dublin Core Metadata Initiative

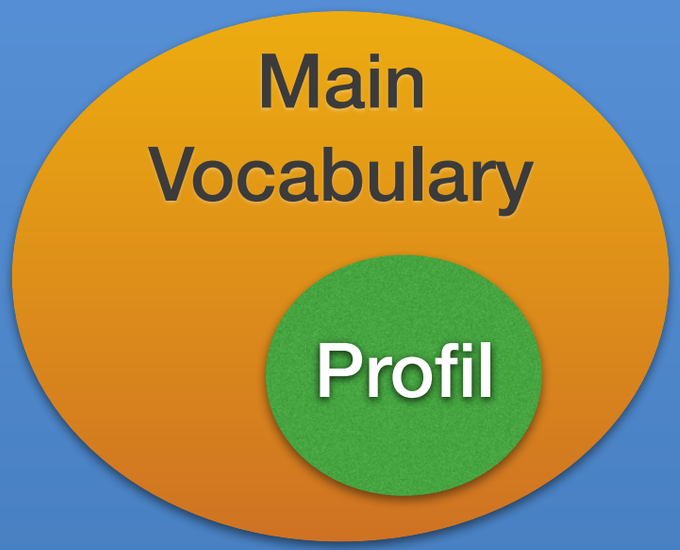
DCMI Application Profiles

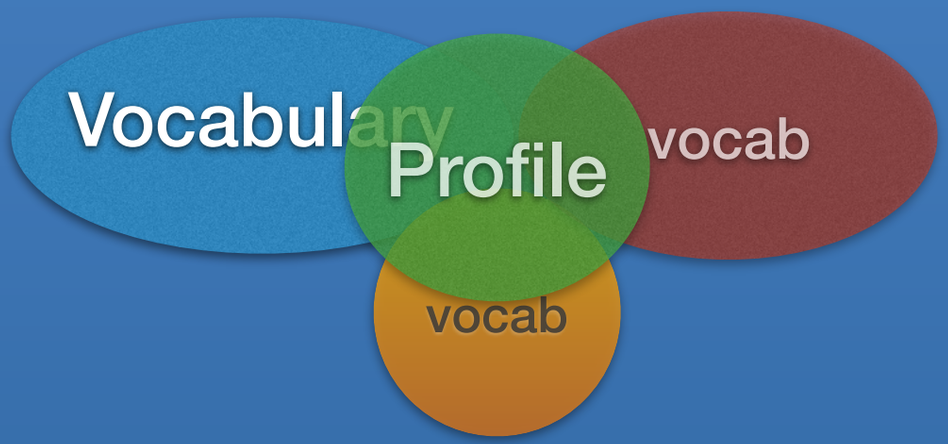
- Application Profiles (1999): how metadata terms from different vocabularies are combined and constrained in community-specific metadata
- Description Set Profile Constraint Language (2008) - “templates and constraints”
- DCMI Application Profiles Interest Group - 2019

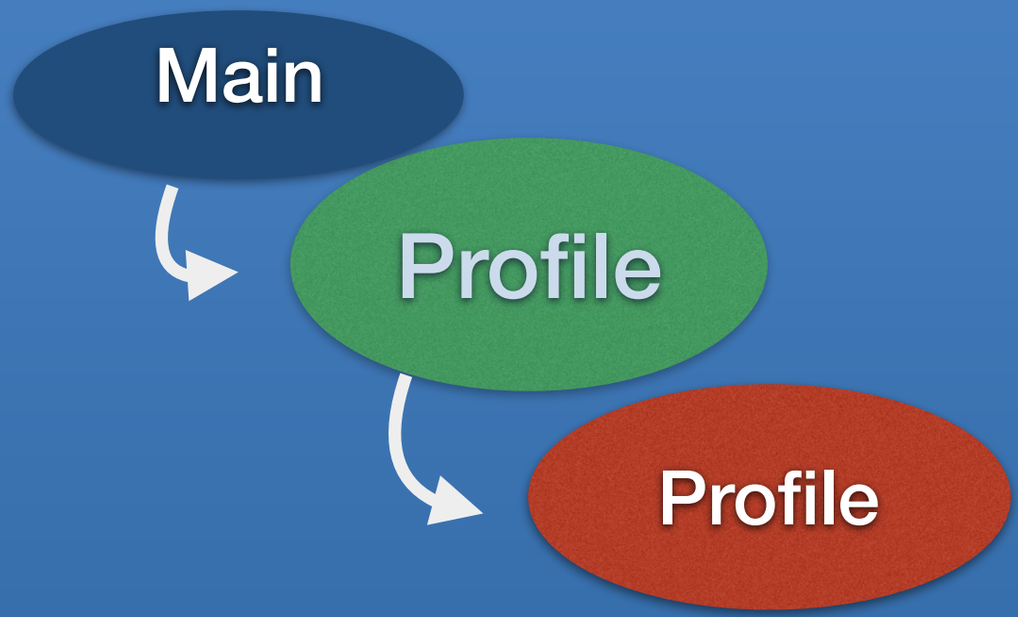
Profile functions

- For documenting community consensus
- For data creation
- For analyzing and validating data
- For mixing/selecting data from different sources
- For ingesting data from others
- For retrieval or display of different views

What is a profile?



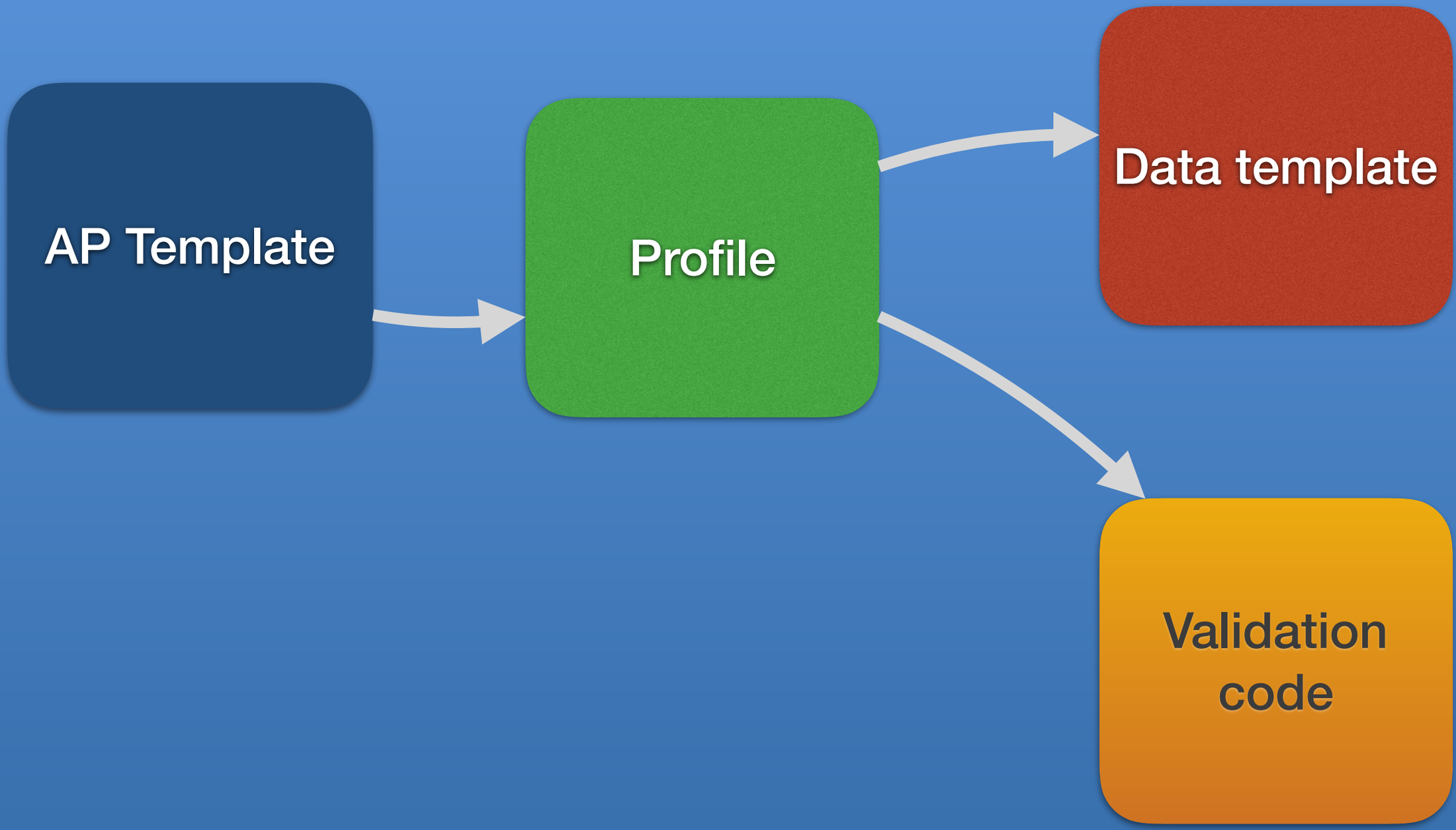




DCMI AP Interest Group requirements

- Use **tabular** format to help non-coders create simple profiles
- "Core"
- Technology-agnostic concepts
- Convertible into machine-actionable representations
- Compatible with validation standards (ShEx, SHACL, XML schemas…)

Dublin Core AP Project



Dublin Core AP Project

AP Template

Common features of application profiles

- Selected vocabulary terms
- Constraints on terms (cardinality, rules)
- Constraints on values
- May have more than one entity

How simple is simple enough? Too simple?

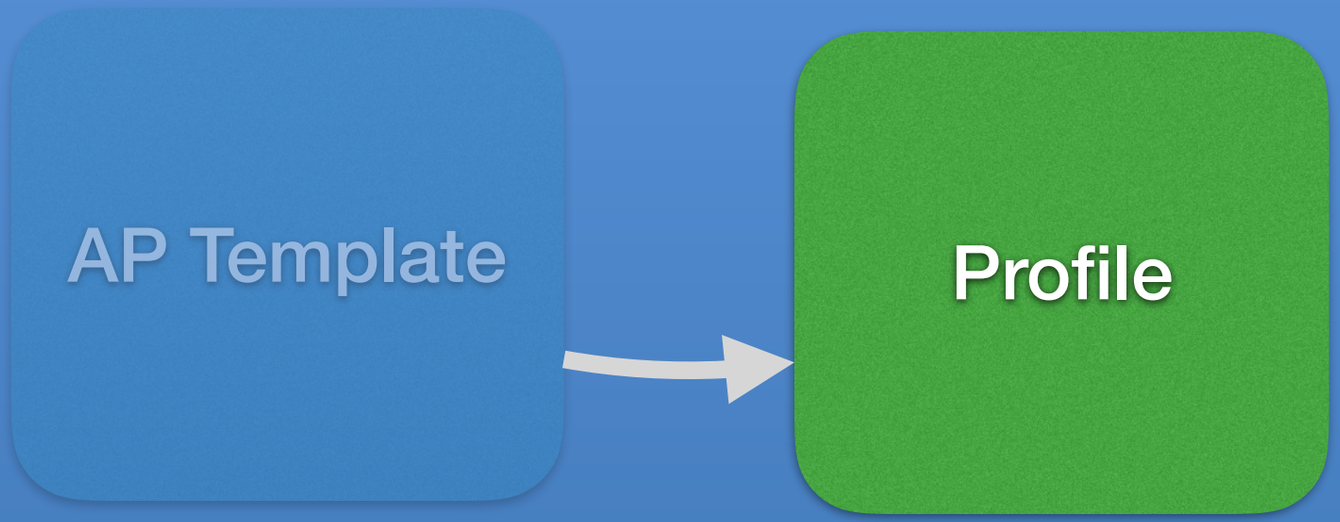
Entity_ID	Entity_label	statementID	property	propLabel	cardinality	valuetype	annotation

Entity

Statement

Value

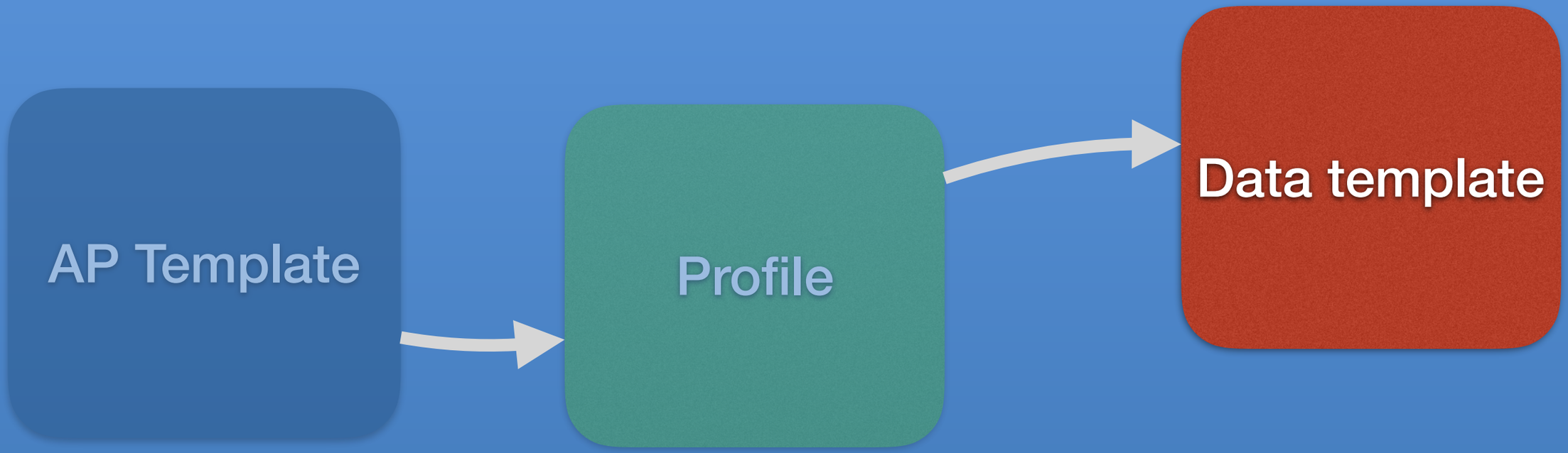
Dublin Core AP Project



Sample profile

Entity_name	Entity_label	Statement_ID	Property	Property_Label	Cardinality	Value_type	Annotation
book	Book	creator	dct:creator	Author	0,-1	@person	Author is not required; no limit on the number
		title	dct:title	Title	1,1	literal	Each book must have a title
		pubDate	dct:date	Year of publication	1,1	xsd:year	Only the year, 9999
person	Person	name	foaf:Name	Name	1,1	literal	Each person has one name
		email	foaf:mbox	Email	0,1	URI	Email is optional but only one allowed
		birthDate	dct:date	Birth year	0,1	xsd:year	Only the year, 9999

Dublin Core AP Project



Possibly like ...

BOOK

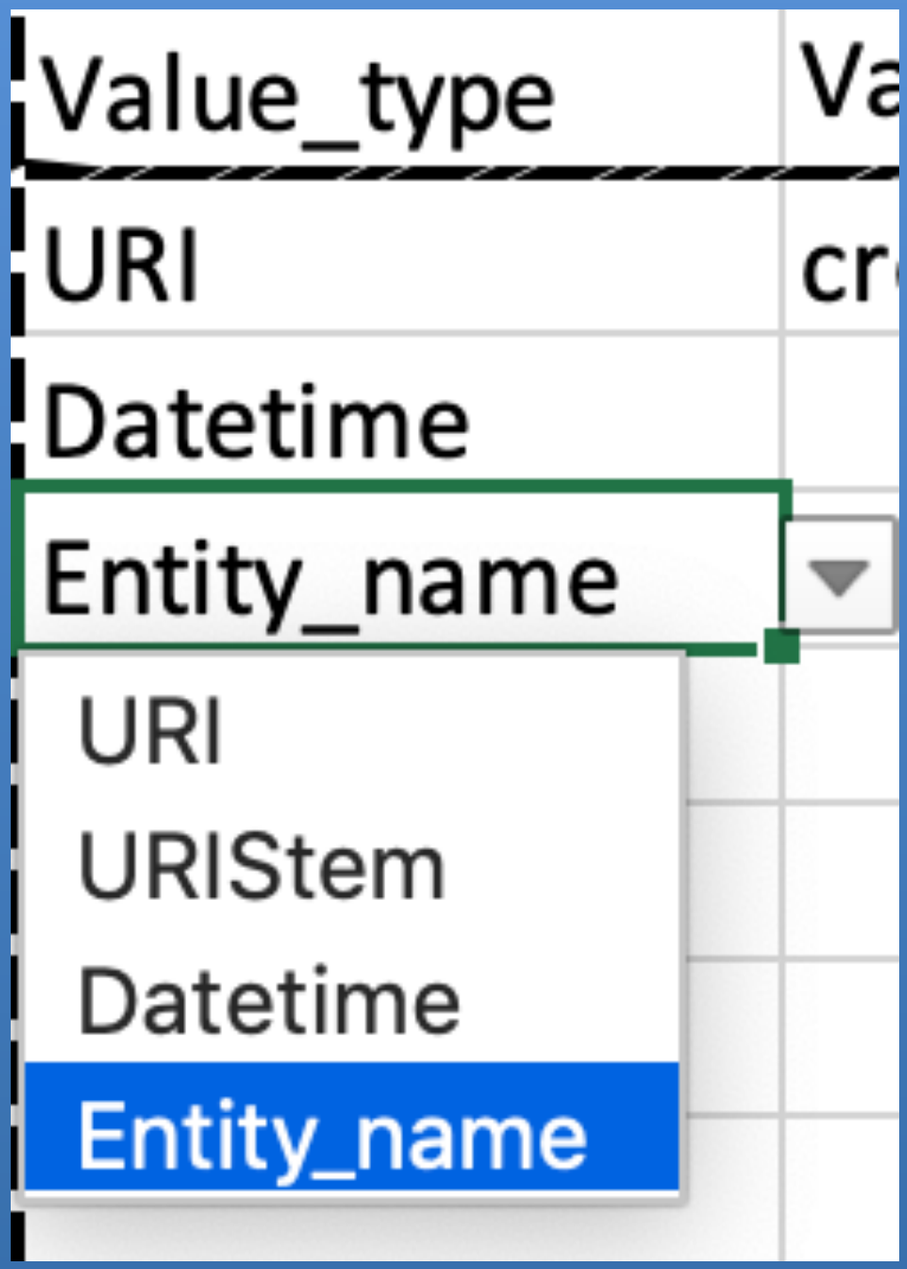
Author

Title

Date

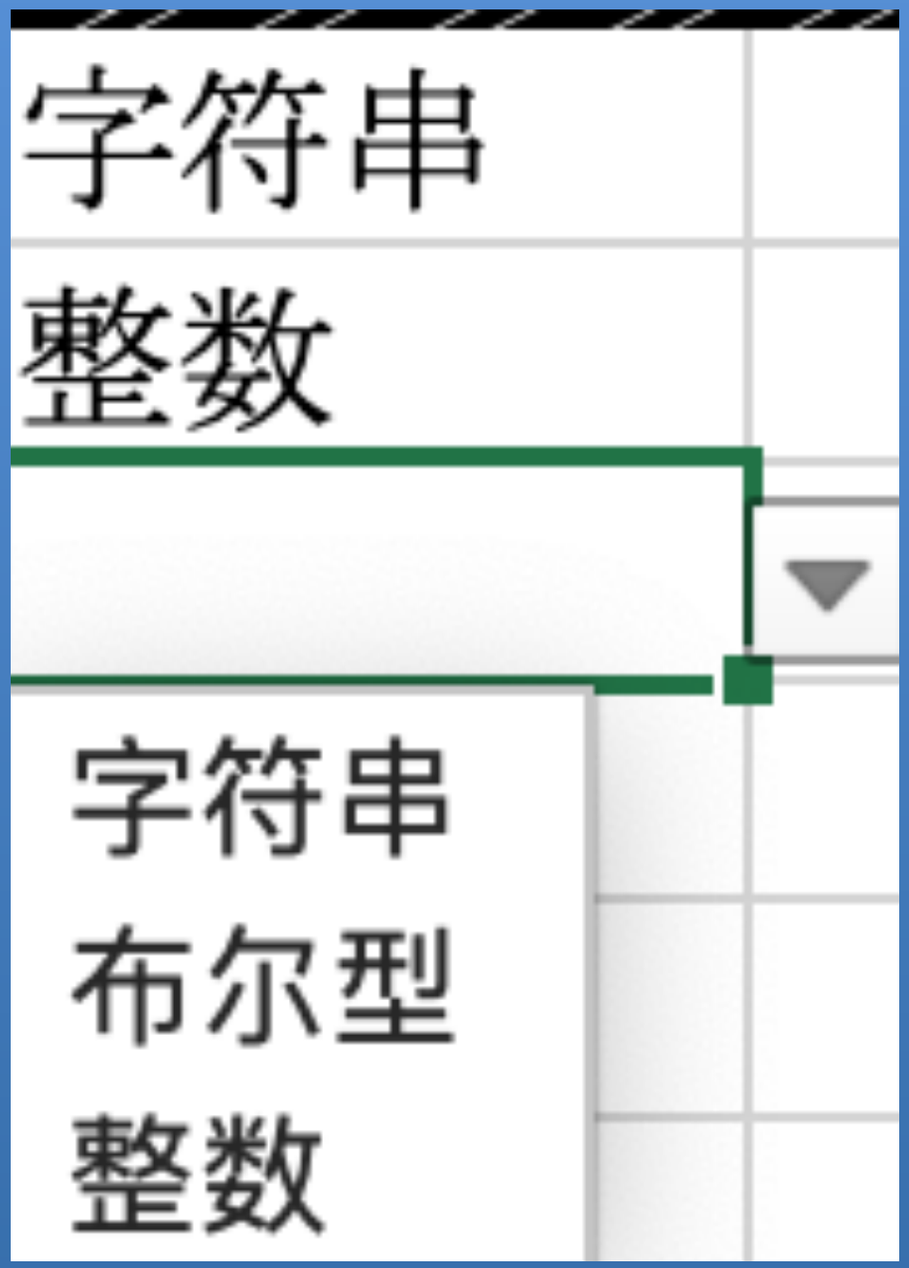
Defining values

Value_type
Community-specific pick lists?



RDA DMP	A	B
uri	URI	URI
string	literal	Literal
date	xsd:date	
date-time		DateTime
complex nested data structure	entity	Entity_name
		URIStem
term controlled vocabulary		

Value_type
Community-specific pick lists?



RDA DMP	A	B
uri	URI	URI
string	literal	Literal
date	xsd:date	
date-time		DateTime
complex nested data structure	entity	Entity_name
		URIStem
term controlled vocabulary		

Value_type
Usable in conversion scripts

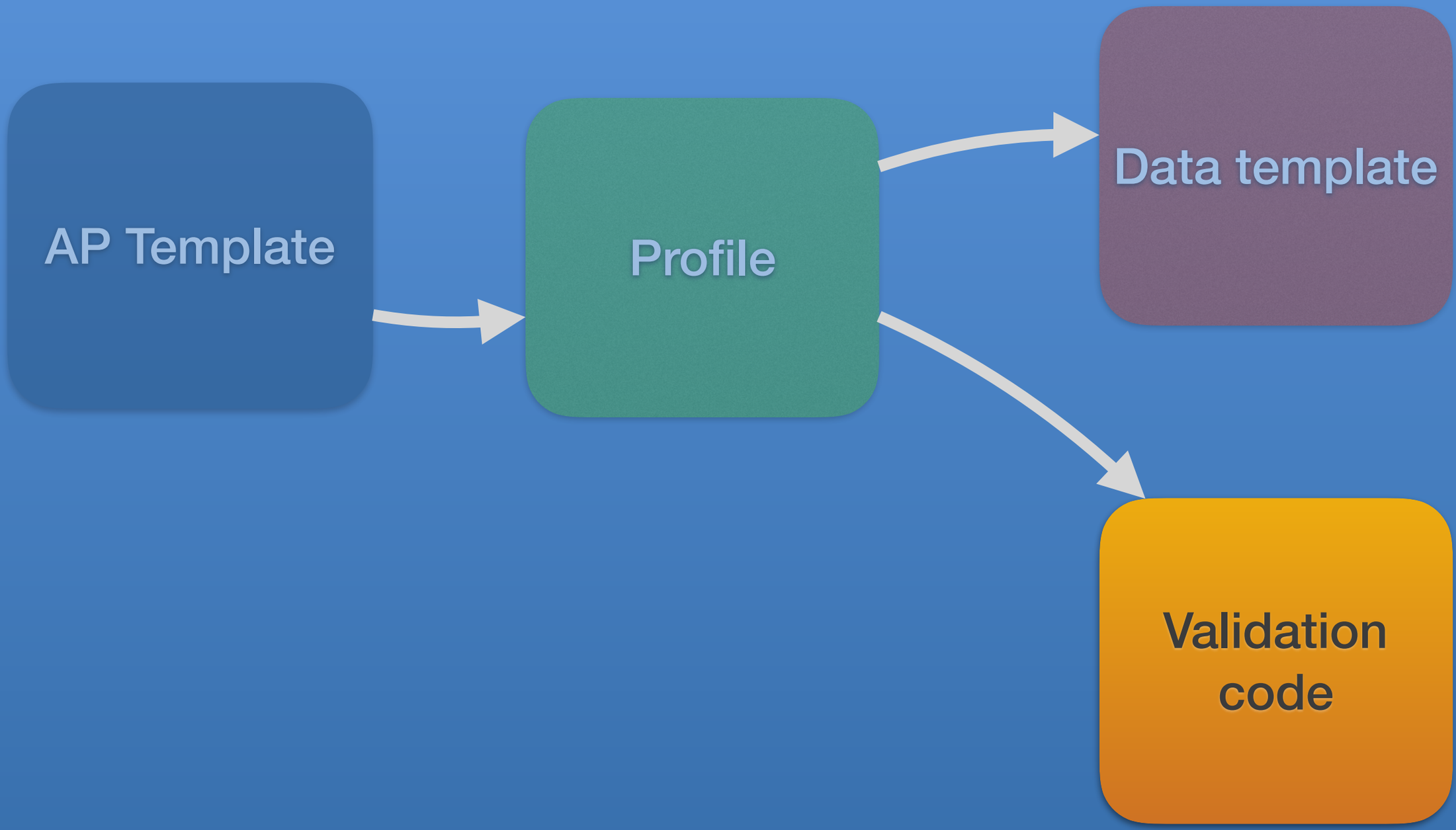
```
33      if value_type == "DateTime":
34          vtype = " xsd:dateTime"
35          schema.append(f"    {property}{vtype}{card} ;\n")
36      elif value_type == "URIStem":
37          value = value + "~"
38          schema.append(f"    {property} [{value}] {card} ;\n")
```

Mand	Repeat
yes	no
no	no
yes	yes

0,-1
1,1
1,1
1,1

0..1

Dublin Core AP Project



Example: Wikidata entity schema from tabular profile

painting (E130)

language code	label	description	aliases	edit
en	painting	surface artistically covered with paint	paintings	edit
nl	schilderij			edit

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

start = @<painting>

<painting> {
    # instance of painting
    wdt:P31 [wd:Q3305213] ;
    # inception
    wdt:P571 xsd:dateTime ? ;
    # location
    wdt:P276 . + ;
    # title
    wdt:P1476 . + ;
    # collection
    wdt:P195 . + ;
    # creator
    wdt:P170 @<creator>+
}

<creator> {
  wdt:P31 [wd:~] ;
}
```

Main page
Community portal
Project chat
Create a new Item
Create a new Lexeme
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Print/export
Download as PDF

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page

Entity_name	Entity_label	Property	Property_label	Cardinality	Value	Value_type	Annotation
painting	Painting	wdt:P31	Is a	1,1	wd:Q3305213	URI	Instance of "painting"
		wdt:P571	Date of inception	0,1		DateTime	
		wdt:P276	Location	1,-1			
		wdt:P1476	Title	1,-1			
		wdt:P195	Collection	1,-1			
		wdt:P170	Creator	1,-1	creator	Entity_name	
creator	Artist	wdt:P31	Is a	1,1	wd:	URIStem	

Entity_name	Property	Property_label	Mand	Repeat	Value	Value_type	Annotation
painting	wdt:P31	Is a	yes	no	wd:Q3305213	URI	Instance of "painting"
	wdt:P571	Date of inception	no	no		DateTime	
	wdt:P276	Location	yes	yes			
	wdt:P1476	Title	yes	yes			
	wdt:P195	Collection	yes	yes			
	wdt:P170	Creator	yes	yes	creator	Entity_name	
creator	wdt:P31	Is a	yes	no	wd:	URIStem	

```
print(prefixes, "\n")
for line in schema:
    print(line, end="")
print("\n", end_matter)
```

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

start = @<painting>

<painting> {
    # 'Is a' instance of "painting"
    wdt:P31 [wd:Q3305213]  ;
    # 'Date of inception'
    wdt:P571 xsd:dateTime? ;
    # 'Location'
    wdt:P276 . + ;
    # 'Title'
    wdt:P1476 . + ;
    # 'Collection'
    wdt:P195 . + ;
    # 'Creator'
    wdt:P170 @<creator>+ ;
}

<creator> {
    # 'Is a' instance of "creator"
    wdt:P31 [wd:~]  ;

 }
```

painting (E130)

language code	label	description	aliases	edit
en	painting	surface artistically covered with paint	paintings	edit
nl	schilderij			edit

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

start = @<painting>

<painting> {
    # instance of painting
    wdt:P31 [wd:Q3305213] ;
    # inception
    wdt:P571 xsd:dateTime ? ;
    # location
    wdt:P276 . + ;
    # title
    wdt:P1476 . + ;
    # collection
    wdt:P195 . + ;
    # creator
    wdt:P170 @<creator>+
}

<creator> {
  wdt:P31 [wd:~] ;
}
```

Main page
Community portal
Project chat
Create a new Item
Create a new Lexeme
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Print/export
Download as PDF

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page

Lots of questions

- How simple is too simple? How complex is too complex?

- Should we define a standard set of basic value types or leave those to communities?

- What parts of a profile need to be citably identified in order to be re-used within a profile?

General conclusions

Many reasons, many requirements - hard to create a "universal" template

Simple template is plausible, complex templates require specific user knowledge

Basic guidelines for simple profiles would be helpful

Links

- http://dublincore.org/groups/application_profiles_ig/
- https://github.com/dcmi/dcap
- https://github.com/dcmi/dcap/tree/master/prototypes/

Thank you

tom@tombaker.org
kcoyle@kcoyle.net