DC TAP

Dublin Core Tabular
Application Profile

https://github.com/dcmi/dctap

DCMI Application Profiles WG

- Karen Coyle (Chair)
- Tom Baker, DCMI
- Phil Barker
- John Huck, University of Alberta
- Ben Reisenberg, University of Washington
- Nishad Thalhath

… and others. Thank you!

A profile

A profile is the definition of a metadata practice that constrains structures, properties and values. A profile is a reuse of vocabulary terms that have been previously defined.

Profile purpose

- Document a community practice
  - For humans
  - For machines
- Input to data creation forms
- Translate to data validation methods
- Manage selection from data sources

TAP functions

- Define metadata structure
- Define constraints
  - cardinality
  - data types
  - value constraints
- Provide readability
  - labels
  - notes
- Note: designed with RDF metadata in mind, but is not limited to RDF metadata

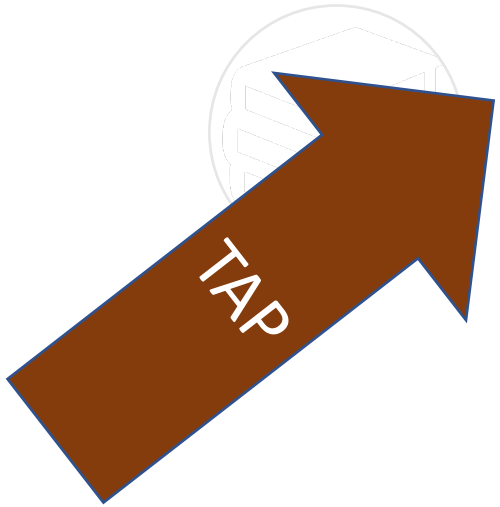
The "meta" problem

1. Vocabulary and structure for a profile

2. A profile that describes one's metadata and constraints

3. Instance data that conforms to the profile

The "meta" problem

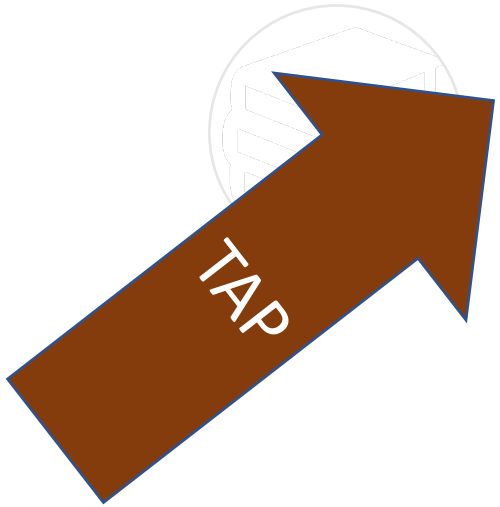


1. Vocabulary and structure for a profile

2. A Profile that describes one's metadata and constraints

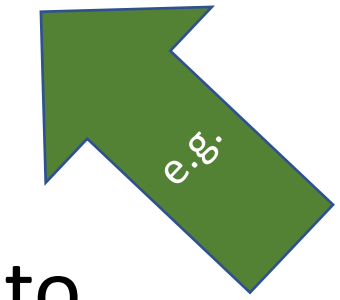
3. Instance data that conforms to the profile

The "meta" problem



1. Vocabulary and structure for a profile

2. A Profile that describes one's metadata and constraints



3. Instance data that conforms to the profile

DC Tabular Application Profile
12 elements

- shapeID
  - shapeLabel
- propertyID
  - propertyLabel
  - note
  - mandatory
  - repeatable
  - valueNodeType
  - valueDatatype
  - valueConstraint
  - valueConstraintType

DC TAP 12 elements in a table

shapeID	shapeLabel	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDataType	valueShape
book	Book	dct:creator	Author	true	true			person
		dct:title	Title	true	false	LITERAL	xsd:string	
		dct:date	Year of publication	false	false	LITERAL	xsd:year	
person	Author	foaf:name	Name	true	false	LITERAL	xsd:string	
		foaf:mbox	Email	false	false	IRI		
		dct:date	Birth year	false	false	LITERAL	xsd:year	

Tabular format

shapeID	shapeLabel	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDataType	valueShape
book	Book	dct:creator	Author	true	true			person
		dct:title	Title	true	false	LITERAL	xsd:string	
		dct:date	Year of publication	false	false	LITERAL	xsd:year	
person	Author	foaf:name	Name	true	false	LITERAL	xsd:string	
		foaf:mbox	Email	false	false	IRI		
		dct:date	Birth year	false	false	LITERAL	xsd:year	


propertyID
dct:creator
dct:title
dct:date
dct:publisher
Only propertyID is required

Add data type

propertyID	valueDataType
dct:creator	xsd:string
dct:title	xsd:string
dct:date	xsd:year
dct:publisher	xsd:string

Include labels and notes

propertyID	propertyLabel	note
dct:creator	Author	name in natural order
dct:title	Title	take title from the cover
dct:date	Date of publication	year only

Cardinality

propertyID	valueDataType	mandatory	repeatable
dct:creator	xsd:string	true	true
dct:title	xsd:string	true	false
dct:date	xsd:year	true	false
dct:publisher	xsd:string	false	false

Creating shapes for "things" described in the metadata

shapeID	shapeLabel	propertyID	valueDataType	valueShape
book	Book	dct:creator		person
		dct:title	xsd:string	
		dct:date	xsd:year	
person	Author	foaf:name	xsd:string	
		sdo:birthDate	xsd:year	

Shape

shapeID	shapeLabel	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDataType	valueShape
book	Book	dct:creator	Author	true	true			person
		dct:title	Title	true	false	LITERAL	xsd:string	
		dct:date	Year of publication	false	false	LITERAL	xsd:year	
person	Author	foaf:name	Name	true	false	LITERAL	xsd:string	
		foaf:mbox	Email	false	false	IRI		
		dct:date	Birth year	false	false	LITERAL	xsd:year	

A group of statement constraints that share a subject node and are identified with the same shapeID.

Statement constraint

shapeID	shapeLabel	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDataType	valueShape
book	Book	dct:creator	Author	true	true			person
		dct:title	Title	true	false	LITERAL	xsd:string	
		dct:date	Year of publication	false	false	LITERAL	xsd:year	
person	Author	foaf:name	Name	true	false	LITERAL	xsd:string	
		foaf:mbox	Email	false	false	IRI		
		dct:date	Birth year	false	false	LITERAL	xsd:year	

A statement constraint consists of a property and any rules that constrain the property and its value.

"Book" shape

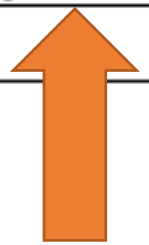
shapeID	shapeLabel	propertyID	valueDataType	valueShape
book	Book	dct:creator		person
		dct:title	xsd:string	
		dct:date	xsd:year	
person	Author	foaf:name	xsd:string	
		sdo:birthDate	xsd:year	

"Person" shape

shapeID	shapeLabel	propertyID	valueDataType	valueShape
book	Book	dct:creator		person
		dct:title	xsd:string	
		dct:date	xsd:year	
person	Author	foaf:name	xsd:string	
		sdo:birthDate	xsd:year	

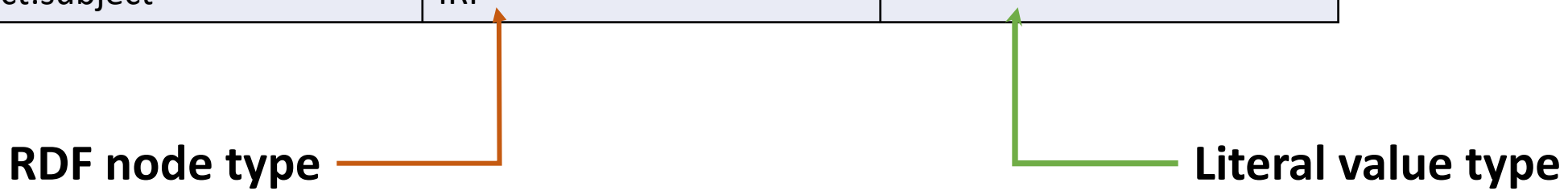
Linking shapes

shapeID	shapeLabel	propertyID	valueDataType	valueShape
book	Book	dct:creator		person
		dct:title	xsd:string	
		dct:date	xsd:year	
person	Author	foaf:name	xsd:string	
		sdo:birthDate	xsd:year	



RDF nodes and literal property values

propertyID	valueNodeType	valueDataType
dct:creator	BNODE	
dct:title	LITERAL	xsd:string
dct:date	LITERAL	xsd:year
dct:subject	IRI	



Value constraints

shapeID	shapeLabel	propertyID	valueDataType	valueShape	valueConstraint
book	Book	dct:creator		person	
		dct:title	xsd:string		
		dct:date	xsd:year		
		dc:subject	xsd:string		History
person	Author	rdf:type			foaf:Person
		foaf:name	xsd:string		
		sdo:birthDate	xsd:year		

Value constraint types

Value constraint types

- Beyond single constraint strings
  - IRI stems (value from this vocabulary)
  - pick lists (strings or URIs or URI stems)
  - patterns
  - language tags

valueConstraint	valueConstraintType
http://id.loc.gov/	uriStem
/^[0-9]{1,2}-?[0-9]{0,2}$/	pattern
red,green,blue	picklist
@en,@fr,@sp	languageTag

Some other details

How to handle multiple values in a table cell

- IRI, BNODE
- red,blue,green
- @en,@fr,@sp

- Multiple values are "OR" functions
- AND and NOT are more difficult and currently out of scope

Some other details

Boolean values

- true/false
- 1/0

Other uses would need to be conveyed to potential users of the profile

Close, but unresolved

Recording namespaces and their prefixes

prefix	namespace
dc	http://purl.org/dc/elements/1.1/
foaf	http://xmlns.com/foaf/

dc:subject
foaf:name

Close, but unresolved

TAP Vocabulary

```
<http://example.org/dctap#shapeLabel>
  a rdfs:Property ;
  rdfs:label "Shape label"@en ;
  skos:definition "A brief human-readable label for the shape."@en ;
  rdfs:isDefinedBy <http://example.org/dctap#>  .

<http://example.org/dctap#propertyID>
  a rdfs:Property ;
  rdfs:label "Property identifier"@en ;
  skos:definition "The IRI of a vocabulary term defined in an RDF-compatible vocabulary."@en ;
  rdfs:isDefinedBy <http://example.org/dctap#>  .

<http://example.org/dctap#propertyLabel>
  a rdfs:Property ;
  rdfs:label "Property label"@en ;
  skos:definition "A brief human-readable label for the property."@en ;
  rdfs:isDefinedBy <http://example.org/dctap#>  .
```

https://github.com/dcmi/dctap/blob/main/dctap.ttl

Unresolved but needed functions

i18n

- How can we internationalize the table (and its contents)?

Out of scope (for now)

Interaction between properties or shapes (and, or, not)

- e.g. foaf:name OR (foaf:familyName AND foaf:givenName)
- e.g. mandatory Professor shape or TA shape per course
- some can be done with structure but that's limited

Tabular designs requiring multiple "sheets"

Out of scope (for now)

Open/closed

- Open = other properties or shapes are allowed
- Closed = only what is in the template is allowed
- Shape level? profile level? both?

Some DC TAP resources

Github repository: https://github.com/dcmi/dctap/

Examples: https://github.com/dcmi/dctap/tree/main/examples

Please participate

Open and respond to issues on github
Join the mailing list
Offer data to be profiled
Offer profiles to be TAP'd

Questions? Comments?

shapeID	shapeLabel	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDataType	valueShape
book	Book	dct:creator	Author	true	true			person
		dct:title	Title	true	false	LITERAL	xsd:string	
		dct:date	Year of publication	false	false	LITERAL	xsd:year	
person	Author	foaf:name	Name	true	false	LITERAL	xsd:string	
		foaf:mbox	Email	false	false	IRI		
		dct:date	Birth year	false	false	LITERAL	xsd:year	

https://github.com/dcmi/dctap

Thanks, everyone!

https://github.com/dcmi/dctap/