

**THIS IS MY**

**"IT'S ONLY WEDNESDAY" FACE**

First... how are you feeling about all of this?

# Queries from yesterday

## The common name of the hosts

PREFIX efo: <<http://www.ebi.ac.uk/efo/efo.owl#>>

PREFIX sio: <<http://semanticscience.org/resource/>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

```
select ?label
```

```
where {
```

```
  ?s a          efo:EFO_0001067 .
```

```
  ?s sio:has-participant ?p .
```

```
  ?p a          sio:host .
```

```
  ?p rdfs:label  ?label
```

```
} LIMIT 100
```

# More queries: what pathogens go with what hosts?

PREFIX efo: <<http://www.ebi.ac.uk/efo/efo.owl#>>

PREFIX sio: <<http://semanticscience.org/resource/>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

select ?path ?host

where {

  ?inf sio:has-participant ?h .

  ?h a sio:host .

  ?h rdfs:label ?host .

  ?inf sio:has-participant ?p .

  ?p a sio:pathogen .

  ?p sio:has-identifier ?id .

  ?id rdfs:label ?path .

}

# Species/Host Matrix

PREFIX efo: <<http://www.ebi.ac.uk/efo/efo.owl#>>

PREFIX sio: <<http://semanticscience.org/resource/>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

```
select ?hostname ?pathname
```

```
where {
```

```
  #within the course, every record is a named graph
```

```
  graph ?ghost {
```

```
    ?o a sio:measuring . # ?o would be ".../obs_12345
```

```
    ?o sio:has-participant ?infect .
```

```
    ?infect sio:has-participant ?host .
```

```
    ?host a sio:host .
```

```
    ?host rdfs:label ?hostname .
```

```
    ?infect sio:has-participant ?pathogen .
```

```
    ?pathogen a sio:pathogen .
```

```
  }
```

```
  graph ?pathogen {
```

```
    ?pathogen rdfs:label ?pathname
```

```
  }
```

```
}
```

# Retrieve metadata

PREFIX dc: <http://purl.org/dc/elements/1.1/>

PREFIX course: <http://training.fairdata.solutions/DAV/home/LDP/gofair/>

PREFIX ldp: <http://www.w3.org/ns/ldp#>

PREFIX efo: <http://www.ebi.ac.uk/efo/efo.owl#>

PREFIX sio: <http://semanticscience.org/resource/>

```
select distinct ?license ?author where {
```

```
  graph course: {
```

```
    ?s dc:title ?title .
```

```
    ?s dc:creator ?author .
```

```
    ?s dc:license ?license .
```

```
    ?s ldp:contains ?record .
```

```
  }
```

```
  graph ?record {
```

```
    ?obs a sio:measuring .
```

```
    ?obs sio:is-located-in ?location .
```

```
    ?location rdfs:label "Spain"@en
```

```
  }
```

```
}
```

# Metadata

**Mark D. Wilkinson**

FBBVA Chair in Biotechnology

Isaac Peral Distinguished Researcher

Center for Plant Biotechnology and Genomics

Universidad Politecnica de Madrid

*mark.wilkinson@upm.es*

# FAIR is about Data, *but even more about Metadata*

Why would I say that?

Why is FAIR Data not enough? Which principles are mostly about Metadata?

Under what circumstances would FAIR Data be impossible/impractical/useless?



# FAIR is about Data, *but even more about Metadata*

Why would I say that?

Why is FAIR Data not enough? Which principles are mostly about Metadata?

Under what circumstances would FAIR Data be impossible/impractical/useless?

*When it comes to FAIR Data, be smart about it*

# FAIR is about Data, *but even more about Metadata*

Why would I say that?

Why is FAIR Data not enough? Which principles are mostly about Metadata?

Under what circumstances would FAIR Data be impossible/impractical/useless?

*When it comes to FAIR Data, be smart about it*

*But when it comes to FAIR METADATA, go to the limit!*

What kinds of Metadata are there?

What kinds of Metadata are there?

Intrinsic - example?

Contextual - example?

Provenance - example?

Quality/Process - example?

Statistical - example?

Structural - example?

Administrative - example?

# What kinds of Metadata are there?

Intrinsic - example?

Contextual - example?

Provenance - example?

Quality/Process - example?

Statistical - example?

Structural - example?

Administrative - example?

FAIR “best-practices” would say that you should include all of these in your metadata!  
(this is why I say that FAIR is **even more** about Metadata - it takes more planning)

# Planning your Metadata:

The dataset we are using in this course doesn't have Intrinsic metadata (anymore)

Intrinsic metadata is tightly tied to the data - often generated automatically by the instrument/pipeline that created the data

**PRINCIPLE:** Mostly I2 ([\(Meta\)data use vocabularies that follow FAIR principles](#)) and R1.3 ([R1.3. \(Meta\)data meet domain-relevant community standards](#))

**MY OPINION:** It is generally best to leave intrinsic metadata “intact” and make it available in its native format. There are many tools that consume these kinds of metadata. BUT: This does not preclude you from creating FAIR representations for the purposes of discovery/interoperability/query

**FAIR VOCABULARIES:** DICOM, for example, has its Vocabulary made available in FAIR format: <http://bioportal.bioontology.org/ontologies/SEDI>

# Planning your Metadata:

Contextual metadata may be tied to a specific piece of data, or may be tied to a dataset, or even a set of datasets (e.g. an entire funded project)

→ We may need to create multiple contextual metadata records (“record” or sections of larger records)

**PRINCIPLE:** F2 ([F2. Data are described with rich metadata](#)) I2 ([I2. \(Meta\)data use vocabularies that follow FAIR principles](#)), and R1.3 ([R1.3. \(Meta\)data meet domain-relevant community standards](#))

**MY OPINION:** Data that cannot be found is reuseless. Data that can be found and cannot be understood/interpreted is reuseless. Data that can be found, and understood, but not assessed for its relevance is reuseless.

**FAIR VOCABULARIES:** EDAM (Bioinformatics operations, data types, formats, identifiers and topics); Disease-specific (e.g. NCI Thesaurus); Anatomy (e.g. Foundational Model of Anatomy - FMA)

# Planning your Metadata:

Provenance metadata is usually attached to all levels - specific piece of data, a dataset, or a set of datasets (e.g. an entire funded project)

→ We will need to create multiple provenance metadata records

**PRINCIPLE:** R1.2 ([R1.2. \(Meta\)data are associated with detailed provenance](#)), F2 ([F2. Data are described with rich metadata](#)) and R1.3 ([R1.3. \(Meta\)data meet domain-relevant community standards](#))

**MY OPINION:** Data of unknown origin is reuseless. Data that cannot be properly cited, *should not* be used.

**FAIR VOCABULARIES:** Dublin Core (author, title, etc.) and Data Catalogue (DCAT); EDAM (Bioinformatics operations, data types, formats, identifiers and topics)



# Planning your Metadata:

Quality and Process metadata is usually attached to a specific piece of data or a dataset arising from a single process/run

→ We *may* need to create multiple Quality/Process metadata records

**PRINCIPLE:** R1.2 ([R1.2. \(Meta\)data are associated with detailed provenance](#)), F2 ([F2. Data are described with rich metadata](#)) and R1.3 ([R1.3. \(Meta\)data meet domain-relevant community standards](#))

**MY OPINION:** Where possible/relevant, quality measures should be made available to allow assessment of the conformity of a dataset (e.g. is there an ISO standard? A Data Seal of Approval? etc.)

**FAIR VOCABULARIES:** Image and Data Quality Assessment Ontology: <http://bioportal.bioontology.org/ontologies/IDQA>; [Data Quality Ontology](#) - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4817336/>

# Planning your Metadata:

Statistical metadata is usually attached to other (usually quality/process) metadata, in order to provide a way of evaluating the quality measures recorded in that metadata

→ *We may* need to create multiple Statistical metadata records (in this course, we will not create any... but in your own datasets you probably will need this)

**PRINCIPLE:** R1.2 ([R1.2. \(Meta\)data are associated with detailed provenance](#)), F2 ([F2. Data are described with rich metadata](#)) and R1.3 ([R1.3. \(Meta\)data meet domain-relevant community standards](#))

**MY OPINION:** Is the dataset representative, or biased? What is the bias within the dataset? (e.g. contains 95% female, 5% male)

**FAIR VOCABULARIES:** Depends on the metric - probably EDAM or a more domain-specific vocabulary

# Planning your Metadata:

Structural metadata is associated with both data and metadata - sometimes directly, and sometimes “out of band” (e.g. in HTTP headers)

→ We need to find a way to communicate Structural metadata about both data and metadata

**PRINCIPLE:** Mostly I1 - the structure/syntax should be as machine-readable as possible.

**MY OPINION:** FAIR is largely about supporting mechanized access. Vanilla HTML is not a good option to support machines

**FAIR VOCABULARIES:** Dublin Core (‘format’) and Data Catalogue (DCAT) have fields related to the data structure; EDAM (data types, formats); and the HTTP Protocol (MIME-type) provides out-of-band structural metadata

# Planning your Metadata:

Administrative metadata is (in my experience) most frequently associated with datasets or data catalogues, e.g. dealing with who has access under what conditions.

→ We will need to create Administrative metadata at the catalogue and/or dataset level (though these might be identical)

**PRINCIPLE:** Mostly R's (license, citation), and the A's (authorization)

**MY OPINION:** You **must not** use data that does not have a license, period! You must not use data that you cannot cite. FAIR requires that the process by which you may become authorized to obtain data is explicit - no exceptions.

**FAIR VOCABULARIES:** Dublin Core ('format') and Data Catalogue (DCAT) have fields related to the data structure; There is some work (ongoing) related to Consent ontologies.

What does FAIR Metadata  
“look like”?

# “Skunkworks”

Task: Build a prototype



# Skunkworks Participants



Mark Wilkinson

Michel Dumontier

Barend Mons

Tim Clark

Jun Zhao

Paolo Ciccarese

Paul Groth

Erik van Mulligen

Luiz Olavo Bonino da Silva  
Santos

Matthew Gamble

Carole Goble

Joël Kuiper

Morris Swertz

Erik Schultes

Erik Schultes

Mercè Crosas

Adrian Garcia

Philip Durbin

Jeffrey Grethe

Katy Wolstencroft

Sudeshna Das

M. Emily Merrill

# The Hourglass Concept

We want a large ecosystem of apps that use FAIR Data





# The Hourglass Concept

We want to support a wide range of source providers



# The Hourglass Concept

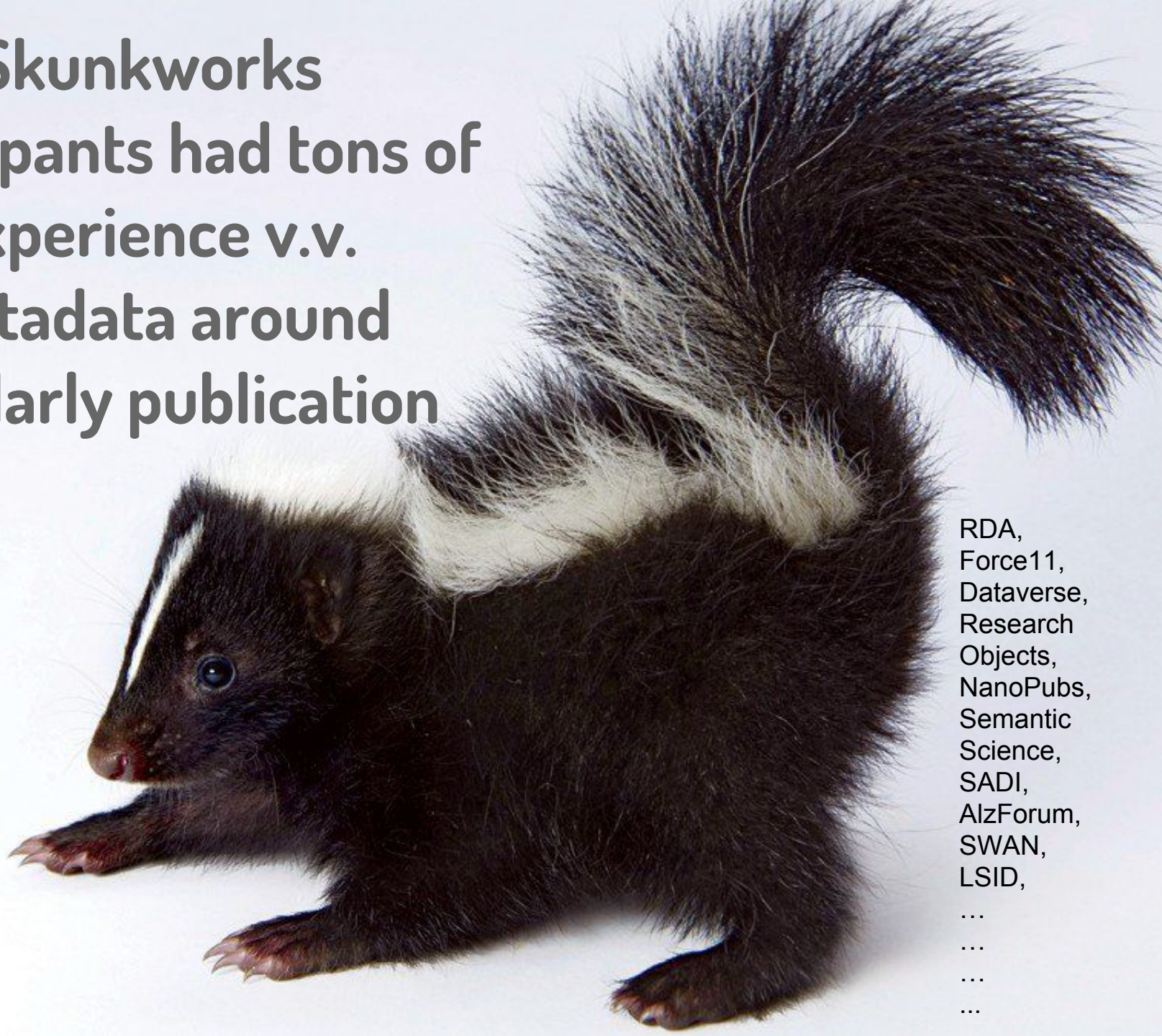
The FAIR solution between them must be THIN!



**Skunkworks  
participants had tons of  
experience v.v.  
metadata around  
scholarly publication**



**Skunkworks  
participants had tons of  
experience v.v.  
metadata around  
scholarly publication**



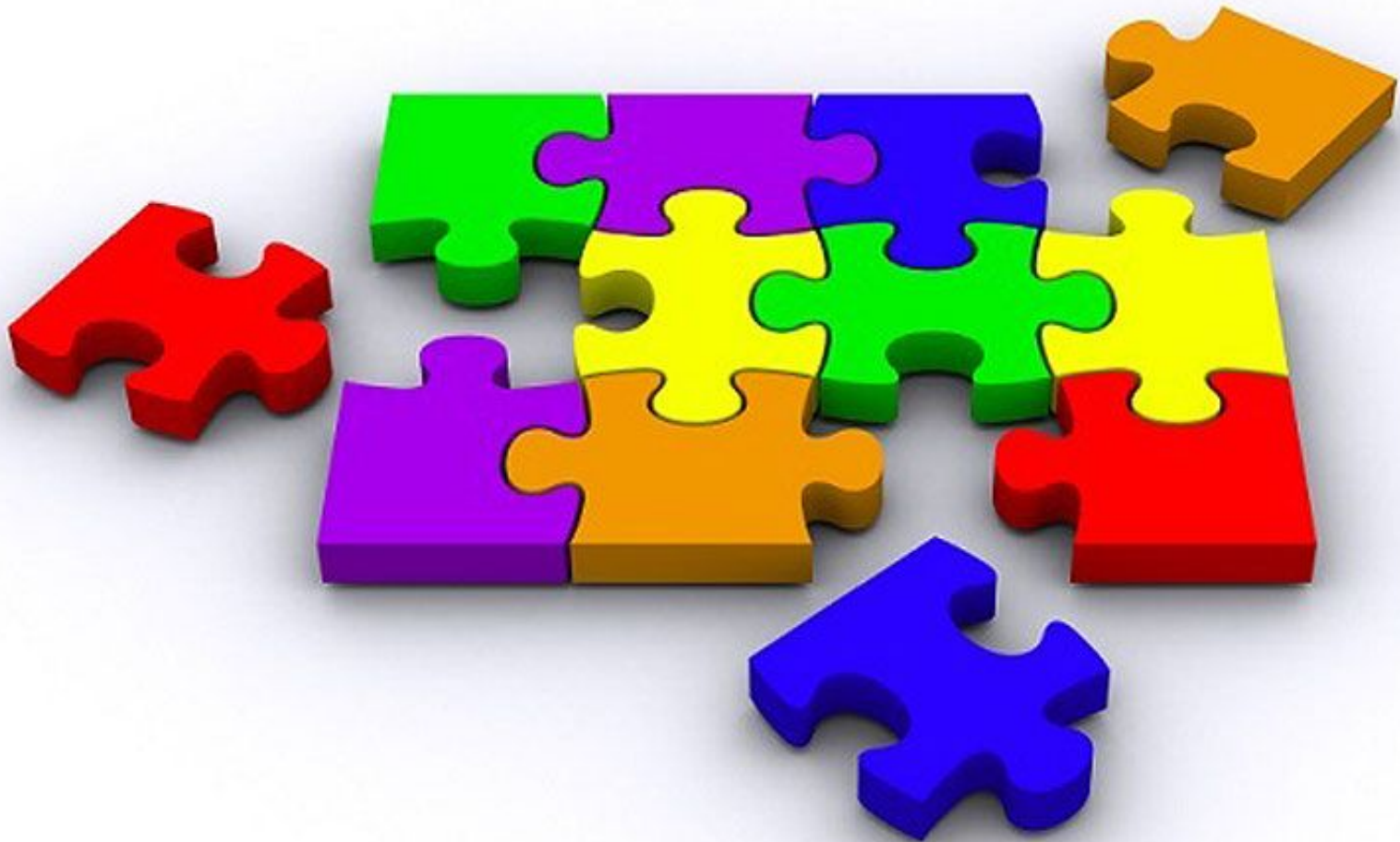
RDA,  
Force11,  
Dataverse,  
Research  
Objects,  
NanoPubs,  
Semantic  
Science,  
SADI,  
AlzForum,  
SWAN,  
LSID,

...  
...  
...  
...

**There was very little  
disagreement  
about F,  
about A,  
or about R**



# The “I” is the big problem



# The “I” is the big problem



**Interoperability is  
Hard!!**

# Keeping the history brief

A series of teleconferences led to the concept of putting metadata into an iterative set of ~identical “containers”





# Skunkworks Hackathons

The “containers of containers of containers” idea was elaborated by the belief that we should also reject any solution that required a new API

ProgrammableWeb.com already catalogues  
**>16,000** different Web APIs

# Skunkworks Hackathons

The “containers of containers of containers” idea was elaborated by the belief that we should also reject any solution that required a new API

ProgrammableWeb.com already catalogues  
**>16,000** different Web APIs

**APIs DO NOT MAKE YOU INTEROPERABLE!**

# Skunkworks Hackathons

The “containers of containers of containers” idea was elaborated by the belief that we should also reject any solution that required a new API

{ REST }

# Skunkworks Hackathons

The “containers of containers of containers” idea was elaborated by the belief that we should also reject any solution that required a new API



{ REST }

Should we talk about what REST is?

# Skunkworks Hackathons

Are there existing standards that are

{ REST }

And have the properties of



?



# Linked Data Platform 1.0

W3C Recommendation 26 February 2015

**This version:**

<http://www.w3.org/TR/2015/REC-ldp-20150226/>

**Latest published version:**

<http://www.w3.org/TR/ldp/>

**Latest editor's draft:**

<http://www.w3.org/2012/ldp/hg/ldp.html>

**Test suite:**

<https://dvcs.w3.org/hg/ldpwg/raw-file/default/tests/ldp-testsuite.html>

**Implementation report:**

<https://dvcs.w3.org/hg/ldpwg/raw-file/default/tests/reports/ldp.html>

**Previous version:**

<http://www.w3.org/TR/2014/PR-ldp-20141216/>

# LDP

## Useful Features



Uses machine-accessible standards and representations, following a REST paradigm



Defines the concept of a "Container" - a machine-actionable way to represent repositories, data deposits, data files, data points, and their metadata



Defines HTTP-resolvable URIs for each of these containers



Uses a widely accepted standard (DCAT) to relate metadata to data → machine-actionable data mining

# The FAIR Accessor

In incremental detail



# What can we describe with FAIR Accessors?

FAIR Accessors provide a machine-actionable, structured,

REST-oriented way to publish Metadata

about a wide range of scholarly “entities”

# What can we describe with FAIR Accessors?

Warehouses (e.g. EBI)

Databases (e.g. UniProt)

Repositories (e.g. Zenodo, INRA-URGI Wheat Repo, UniProt)

Datasets (e.g. output from a workflow)

Research Objects (data a/o workflow a/o results a/o publications)

Data “slices” (e.g. the result of a database query)

Data Records (e.g. image, excel file, patient clinical record)

Other...

# What does a FAIR Accessor “look like”?

Container  
**Resource**



```
<FAIR metadata/>

Contains

MetaRecordResource1
MetaRecordResource2
MetaRecordResource3

...
```

MetaRecord  
**Resource3**



```
<FAIR metadata/>

foaf:primaryTopic Record R

dcat:Distribution_1
  Source URL_U1
  format rdf+xml
dcat:Distribution_2
  Source URL_U2
  format application/xml
```

# The FAIR Accessor

<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

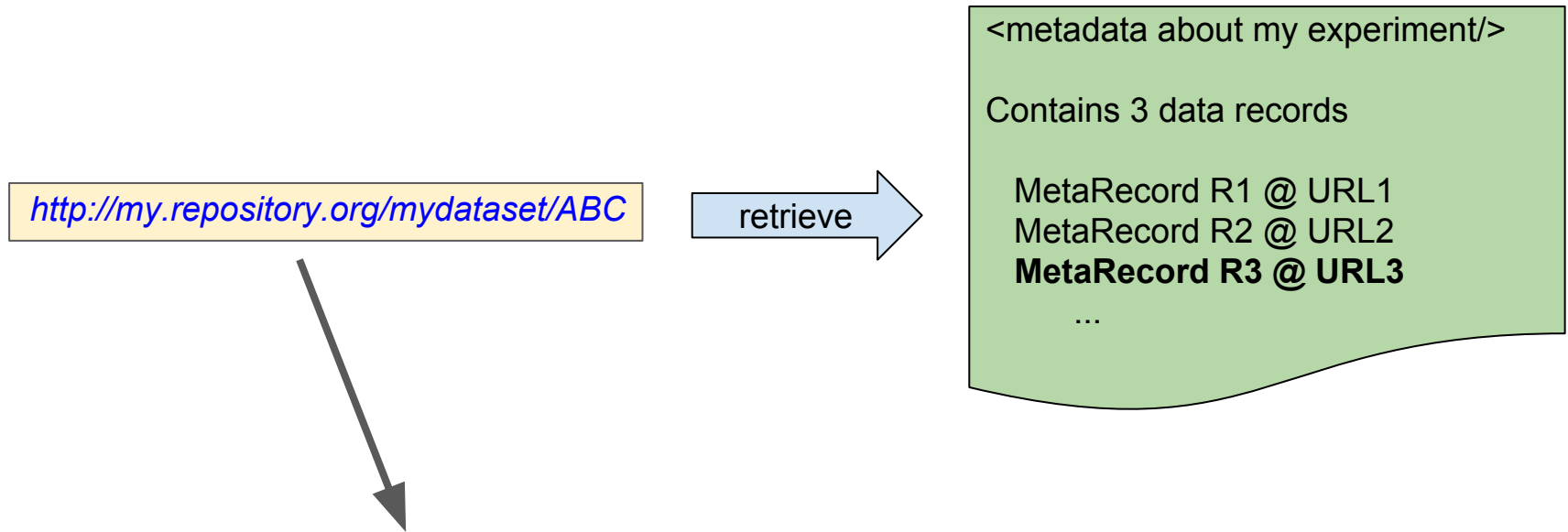
MetaRecord R1 @ URL1

MetaRecord R2 @ URL2

**MetaRecord R3 @ URL3**

...

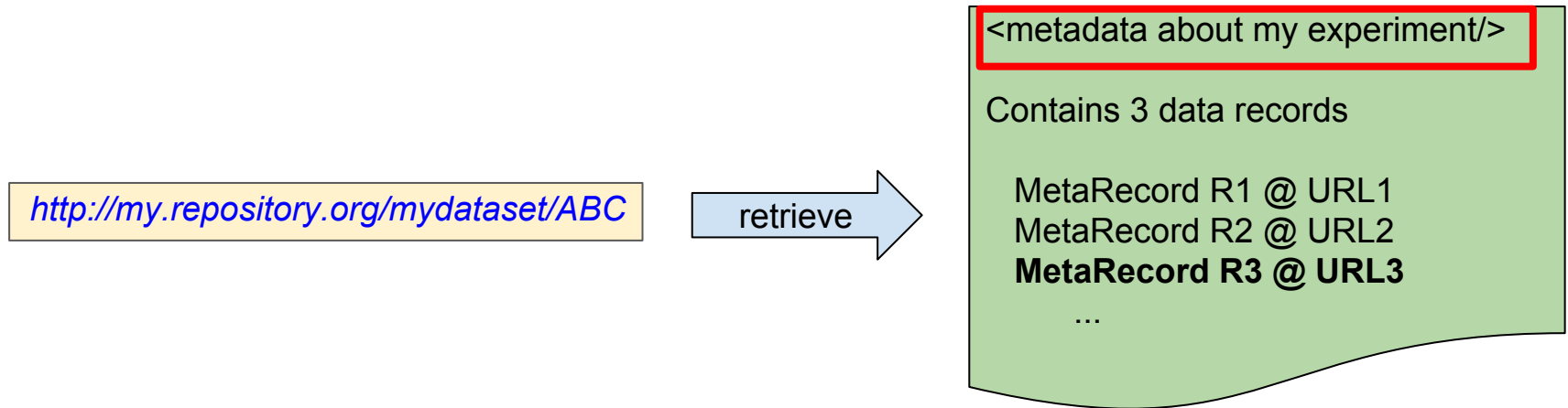
# The FAIR Accessor



The Container Resource

(remember, in REST, a “Resource” is ~a URL that identifies information ... that information is in a particular “state”)

# The FAIR Accessor



“Metadata” simply means “information about something”

In this case, it is information about the experiment.  
What was the hypothesis? What technologies did I use?  
What protocols? What organism/strain?

**Who to cite!!!!**

(i.e. your materials and methods... written in a formal,  
machine-readable manner)

# The FAIR Accessor

<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1

MetaRecord R2 @ URL2

**MetaRecord R3 @ URL3**

...

A link to the **metadata** for data record #3

<http://my.repository.org/mydataset/ABC/URL3>

# The FAIR Accessor

<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

Distribution\_1

Download @ URL\_D1

format text/csv

Distribution\_2

Download @ URL\_D2

format application/excel



# The FAIR Accessor

<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

Distribution\_1

Download @ URL\_D1

format text/csv

Distribution\_2

Download @ URL\_D2

format application/excel

e.g. What machine/device? What filters/cutoffs?  
What algorithms with what settings/parameters?

**Data Usage License!!!**

# The FAIR Accessor

<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

Distribution\_1

Download @ **URL\_D1**

format text/csv

Distribution\_2

Download @ URL\_D2

format application/excel

At **URL\_D1** I can retrieve the data in CSV format

# The FAIR Accessor

<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

Distribution\_1  
Download @ URL\_D1  
format text/csv

Distribution\_2  
Download @ **URL\_D2**  
format application/excel

At **URL\_D2** I can retrieve the data in MS Excel

F<sub>indable</sub> A<sub>ccessible</sub> I<sub>nteroperable</sub> R<sub>eusable</sub>



<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1

MetaRecord R2 @ URL2

**MetaRecord R3 @ URL3**

...

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

Distribution\_1

Download @ URL\_D1

format text/csv

Distribution\_2

Download @ URL\_D2

format application/excel

F<sub>indable</sub> A<sub>ccessible</sub> I<sub>nteroperable</sub> R<sub>eusable</sub>



<http://my.repository.org/mydataset/ABC>



```
<metadata about my experiment/>  
  
Contains 3 data records  
  
MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
MetaRecord R3 @ URL3  
...
```

## Globally-unique Identifiers for data and metadata

<http://my.repository.org/mydataset/ABC/URL3>



```
<metadata about the record/>  
  
primaryTopic Record R3  
  
Distribution_1  
  Download @ URL_D1  
  format text/csv  
Distribution_2  
  Download @ URL_D2  
  format application/excel
```

F<sub>indable</sub> A<sub>ccessible</sub> I<sub>nteroperable</sub> R<sub>eusable</sub>



<http://my.repository.org/mydataset/ABC>



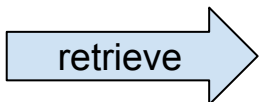
**<metadata about my experiment/>**

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

**Metadata that can be searched (e.g. Google)**

<http://my.repository.org/mydataset/ABC/URL3>

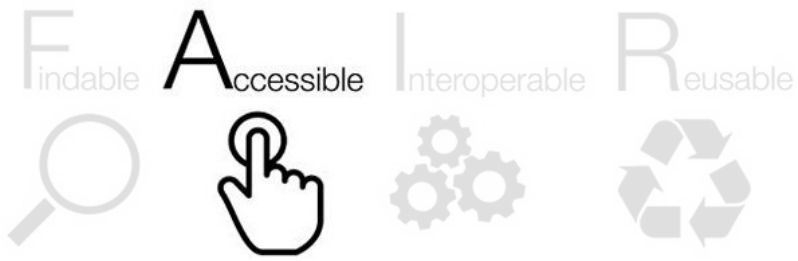


**<metadata about the record/>**

primaryTopic **Record R3**

Distribution\_1  
Download @ URL\_D1  
format text/csv

Distribution\_2  
Download @ URL\_D2  
format application/excel



<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

**Clear access protocol (in this case, simply the Web!)**

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

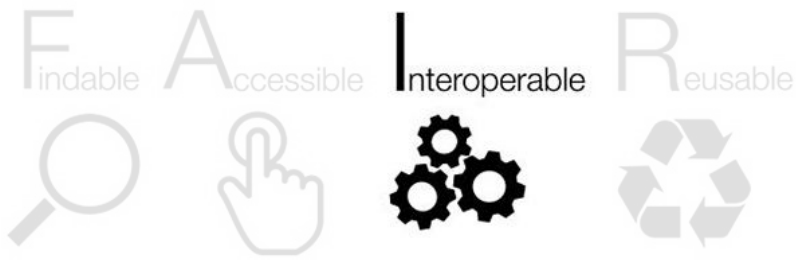
primaryTopic **Record R3**

Distribution\_1

Download @ URL\_D1  
format text/csv

Distribution\_2

Download @ URL\_D2  
format application/excel



<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

**Metadata uses a machine-readable syntax and ontologies  
(I will show you an example of this this later...)**

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

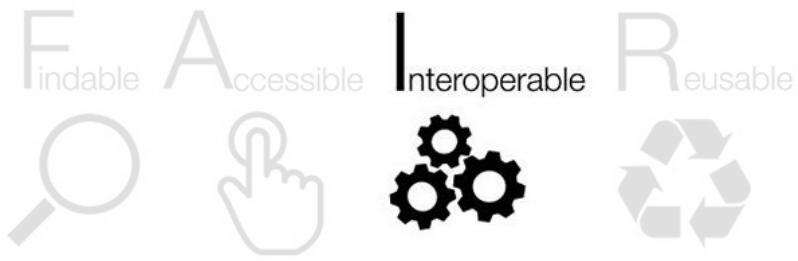
Distribution\_1

Download @ URL\_D1  
format text/csv

Distribution\_2

Download @ URL\_D2  
format application/excel





<http://my.repository.org/mydataset/ABC>

retrieve

<metadata about my experiment/>

Contains 3 data records

MetaRecord R1 @ URL1  
MetaRecord R2 @ URL2  
**MetaRecord R3 @ URL3**  
...

**These data structures conform to a global standard  
Called "DCAT" (the Data Catalog standard)  
A lot of software can interpret these data structures.**

<http://my.repository.org/mydataset/ABC/URL3>

retrieve

<metadata about the record/>

primaryTopic **Record R3**

Distribution\_1

Download @ URL\_D1  
format text/csv

Distribution\_2

Download @ URL\_D2  
format application/excel

# What does a FAIR Accessor “look like”?

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>

Contains

MetaRecordResource1  
MetaRecordResource2  
**MetaRecordResource3**  
...



<http://my.repository.org/mydataset/ABC/URL>  
3

HTTP GET

<FAIR metadata/>

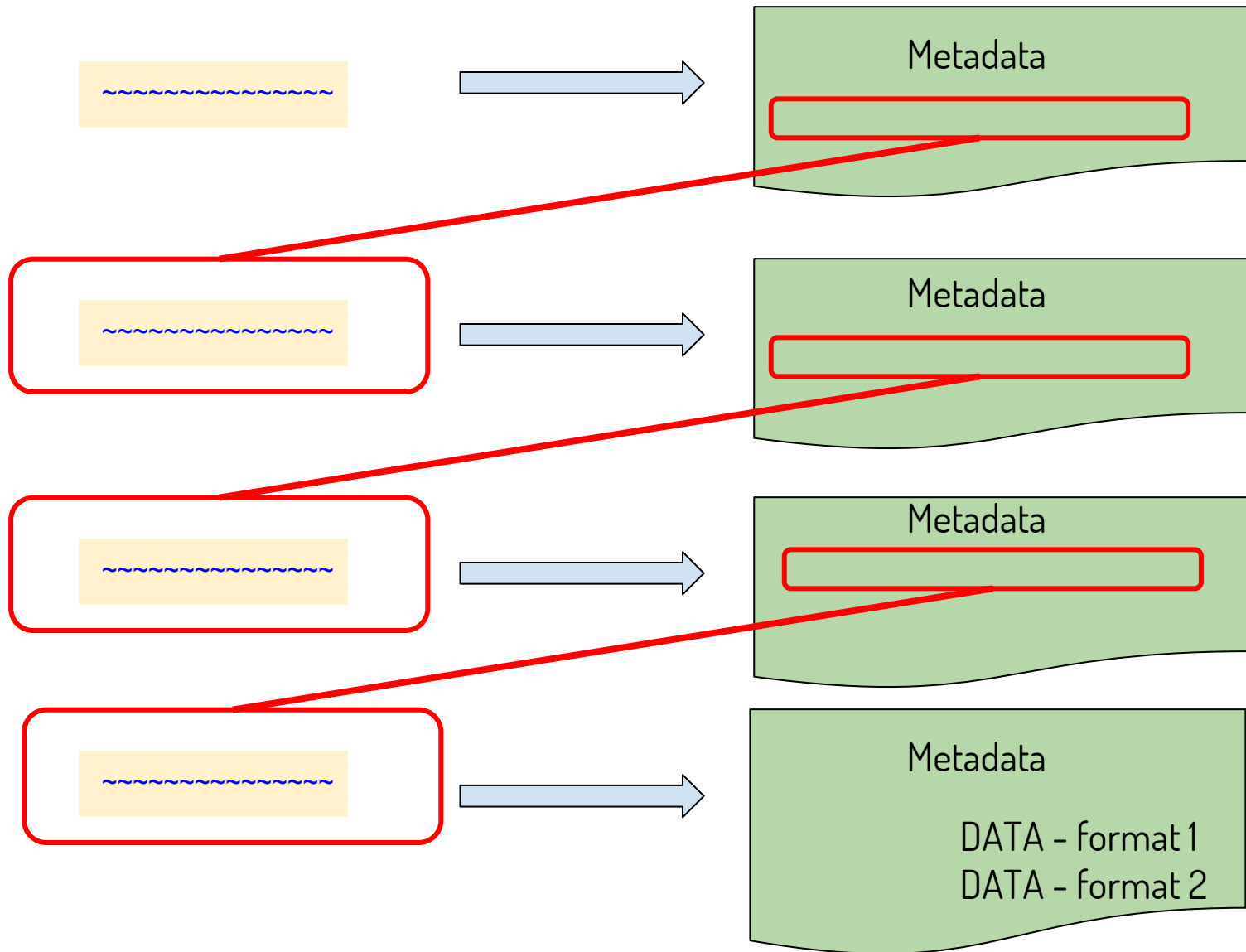
foaf:primaryTopic **Record R**

dcat:Distribution\_1  
Source URL\_U1  
format rdf+xml

dcat:Distribution\_2  
Source URL\_U2  
format application/xml



# Or you may add additional layers...



# Features of the FAIR Accessor

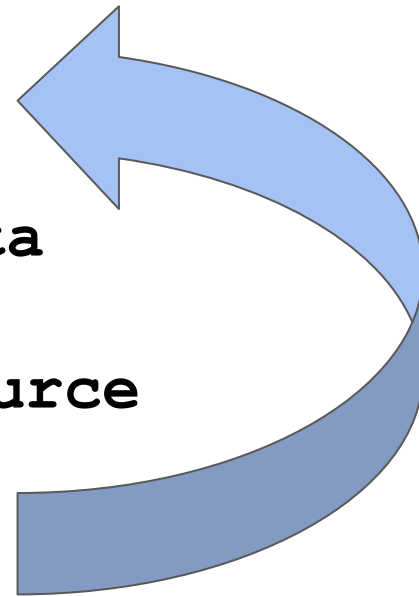
1: There is no API

**GET**

**Interpret the Metadata**

**Select the desired Resource**

**GET**



# Features of the FAIR Accessor

## 1: There is no API

**GET**

ANY Web agent can explore/index a FAIR Accessor  
(e.g. Google)

An agent that understands globally-accepted vocabularies  
can explore it “intelligently”

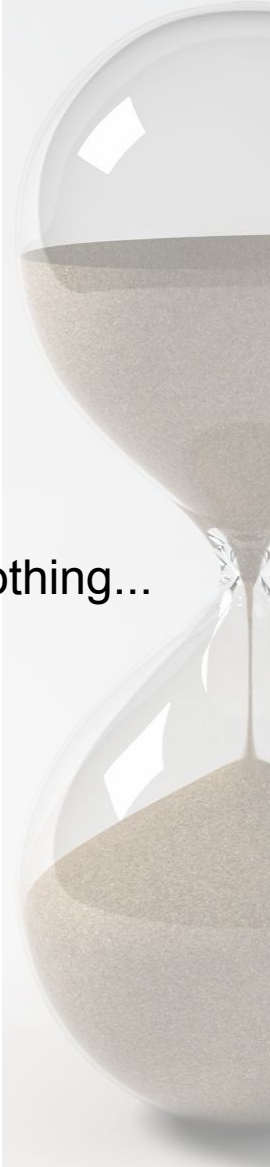
**GET**



# Features of the FAIR Accessor

1: There is no API

It's difficult to get thinner than nothing...



# Features of the FAIR Accessor

## 2: Identifiers for unidentifi-ed/-able things



HTTP GET

<FAIR metadata/>

*This is the **ArrayExpress** query  
I did for paper doi:10/1234.56*

Results:

MetaRecordResource1  
MetaRecordResource2  
**MetaRecordResource3**

...



# Features of the FAIR Accessor

## 2: Identifiers for unidentifi-ed/-able things



<FAIR metadata/>

*This is the **ArrayExpress** query  
I did for paper doi:10/1234.56*

Results:

MetaRecordResource1

MetaRecordResource2

**MetaRecordResource3**

...

Should assist with reproducibility and transparency





# Features of the FAIR Accessor

## 3: A predictable “place” for metadata

**Different “kinds” of metadata have distinct ontological types, and distinct document structures. There is no ambiguity regarding what the metadata is describing - a repository or a record.**

<http://my.repository.org/mydataset/ABC>



Repository metadata

MetaRecordURL

<http://my.repository.org/mydataset/URL3>



PrimaryTopic: record 1A445

Record Metadata...

DATA - format 1

DATA - format 2



# Features of the FAIR Accessor

## 3: Symmetry & predictable path to citation

The record metadata contains an “upward” link to the Repository-level metadata, which should contain license and citation information

<http://my.repository.org/mydataset/ABC>



Repository metadata:  
**Cite: doi:10/8847.384**  
**License: cc-by**

<http://my.repository.org/mydataset/URL3>



Part of dataset ABC  
Metadata...

DATA - format 1  
DATA - format 2



# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>

Contains

<<184 Records>>

Contact Mark Wilkinson  
For more information about  
These records



# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>

Contains

<184 Records>>

Contact Mark Wilkinson  
For more information about  
These records

**CONFIDENTIAL**



# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>  
Contains

**MetaRecordResource3**

<http://my.repository.org/mydataset/URL3>

HTTP GET

<FAIR metadata/>

foaf:primaryTopic **Record R**

dcat:distribution

**<<NONE>>**



# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>  
Contains

**MetaRecordResource3**

<http://my.repository.org/mydataset/URL3>

HTTP GET

<FAIR metadata/>

foaf:primaryTopic **Record R**

dcat:distribution

**CONFIDENTIAL**  
<<NONE>>



# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>

Contains  
**MetaRecordResource3**  
...

<http://my.repository.org/mydataset/URL3>

HTTP GET

<FAIR metadata/>

foaf:primaryTopic **Record R**

dcat:Distribution\_1  
Source URL\_U1  
format rdf+xml



# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

<http://my.repository.org/mydataset/ABC>

HTTP GET

<FAIR metadata/>

Contains  
**MetaRecordResource3**  
...

OPEN  ACCESS

<http://my.repository.org/mydataset/URL3>

HTTP GET

<FAIR metadata/>

foaf:primaryTopic **Record R**

dcat:Distribution\_1  
Source URL\_U1  
format rdf+xml





# Features of the FAIR Accessor

## 4: Granularity of Access/Privacy/Security

Thin solution - if it's private, ***do nothing! Literally!***



# The Real Thing

A working FAIR Accessor  
Serving a “Slice” of UniProt

## A real-world scenario...

You are publishing a paper describing the evolution of proteins in the RNA Processing machineries of the fungus *Aspergillus nidulans*.

You want to be a good scholarly publisher interested in transparency and reproducibility

So you must describe, in detail, the inclusion/exclusion criteria for selecting proteins for your dataset

(today, this is generally done either in the text of the paper, or not at all...)

# The query that returns the relevant proteins

WHERE

```
{  
  ?protein a up:Protein .  
  ?protein up:organism ?organism .  
  ?organism rdfs:subClassOf taxon:162425 .  
  ?protein up:classifiedWith ?go .  
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .  
  bind(replace(str(?protein),  
              "http://purl.uniprot.org/uniprot/", "", "i") as ?id)  
}
```

# The query that returns the relevant proteins

```
WHERE
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .
  bind(replace(str(?protein),
              "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}
```

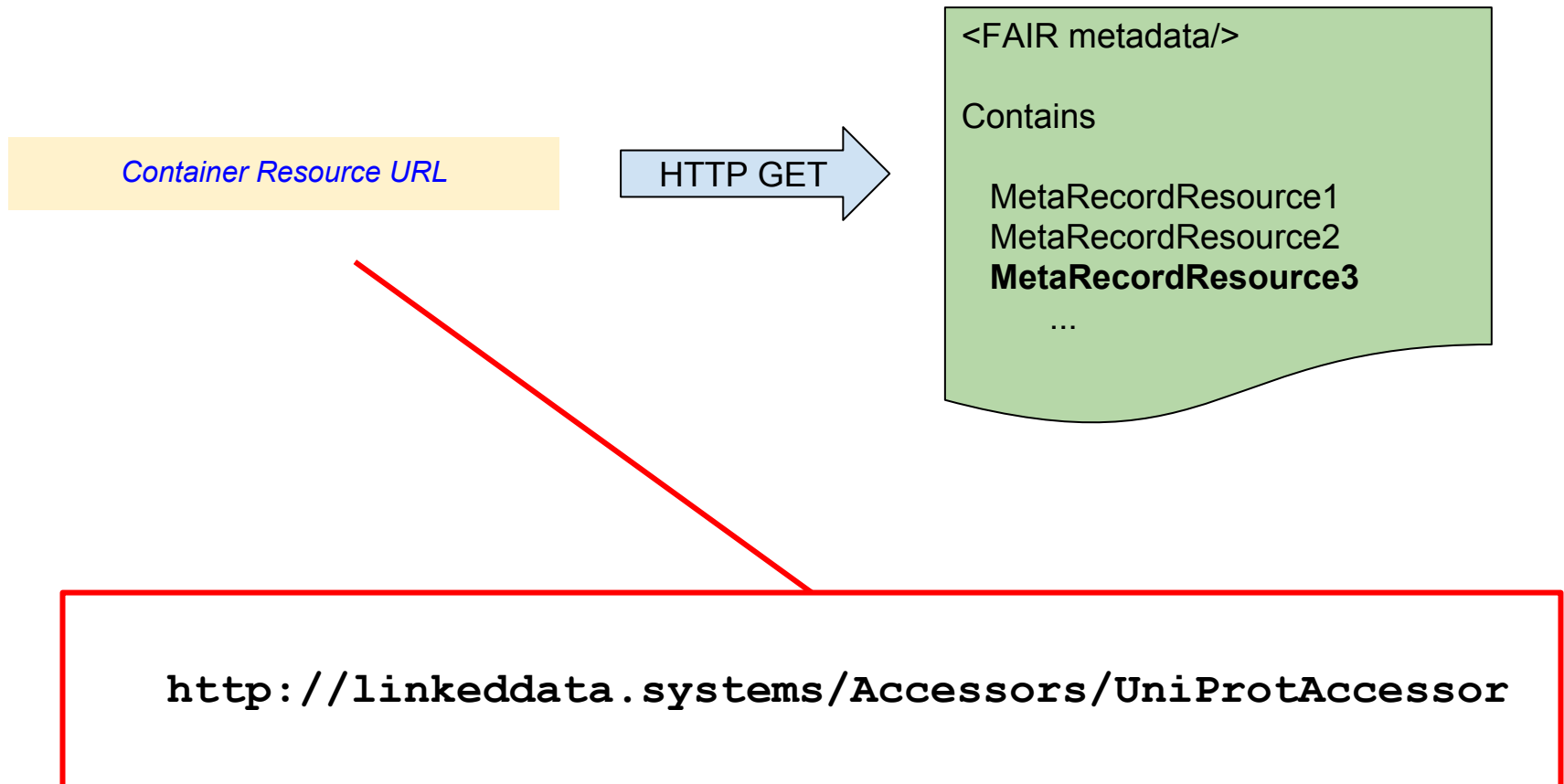
***NCBI Taxonomy:  
Aspergillus nidulans***

# The query that returns the relevant proteins

```
WHERE
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .
  bind(replace(str(?protein),
              "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}
```

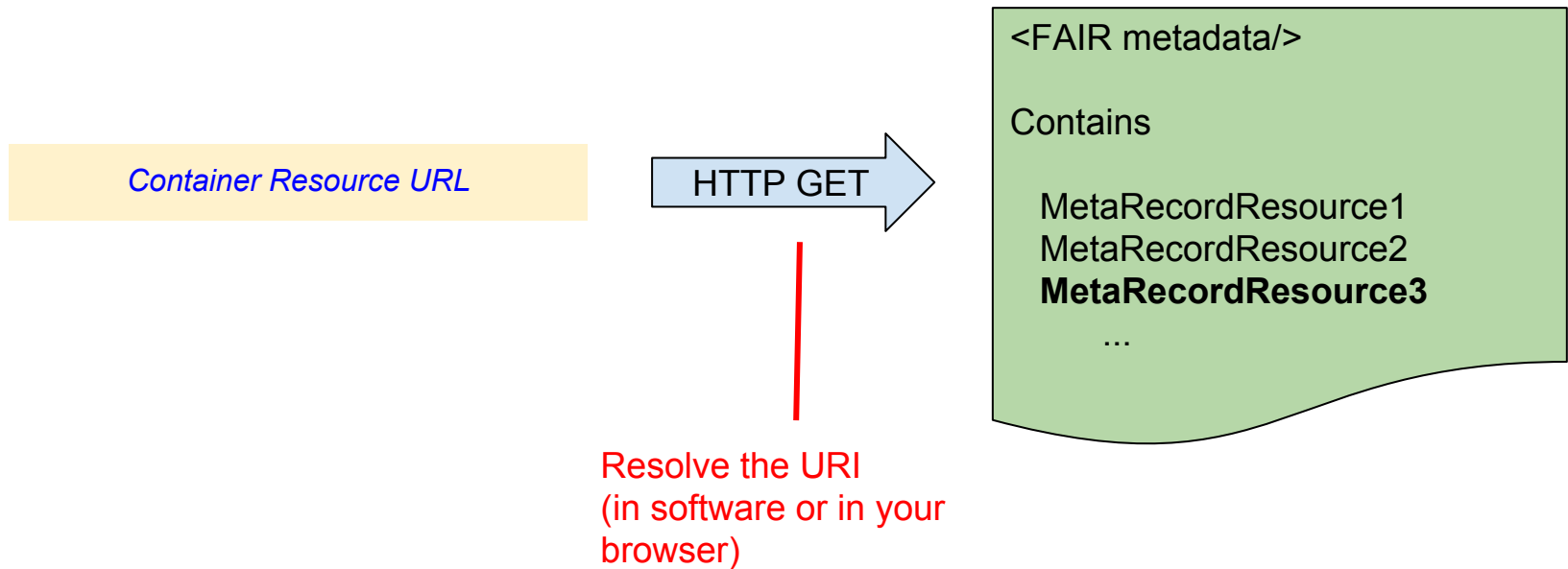
**Gene Ontology:  
RNA Processing**

# Create and publish a FAIR Accessor for that query



This is a working example, if you want to follow along in your own browser...

# Create and publish a FAIR Accessor for that query



`http://linkeddata.systems/Accessors/UniProtAccessor`



# Create and publish a FAIR Accessor for that query

*Container Resource URL*

HTTP GET

<FAIR metadata/>

Contains

MetaRecordResource1

MetaRecordResource2

**MetaRecordResource3**

...

Returns a page of metadata (in this example, in RDF)

**UniProt Slice FAIR Accessor -  
Aspergillus RNA Processing  
proteins**

creator  
language  
license  
title  
authored By  
entities  
term has Principal  
Investigator  
type

[wilkinsonlab.info/](http://wilkinsonlab.info/)  
eng  
cc by nd4.0  
UniProt Slice FAIR Accessor - Aspergillus RNA Processing proteins  
0000 0002 9699 485X  
412  
Dr. Mark Wilkinson

contact Point  
description

**Dataset**  
**Basic Container**  
**Collection**  
**Wilkinson.rdf**

Takes a SPARQL query of the UniProt database specific to proteins and their GO annotations related to RNA Procssing proteins in Aspergillus and makes it a FAIR Accessor source. The precise query is:

```
PREFIX up:<http://purl.uniprot.org/core/>
PREFIX taxon:<http://purl.uniprot.org/taxonomy/>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?id

WHERE
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .

  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}
```

identifier  
keyword

**Uni Prot Accessor**  
Aspergillus nidulans  
Aspergillus  
Proteins  
RNA Processing

landing Page  
language  
publisher  
theme  
contains

[uniprot.org/](http://uniprot.org/)  
en  
[wilkinsonlab.info/](http://wilkinsonlab.info/)  
**RNA Processing conceptscheme.rdf**  
C8UZX9  
C8UZY5  
C8V0B4  
C8V0M2  
C8V0I17

**UniProt Slice FAIR Accessor -  
Aspergillus RNA Processing  
proteins**

**creator**  
**language**  
**license**  
**title**  
**authored By**  
**entities**  
**term has Principal  
Investigator**  
**type**

[wilkinsonlab.info/](http://wilkinsonlab.info/)  
**eng**  
**cc by nd4.0**  
UniProt Slice FAIR Accessor - Aspergillus RNA Processing proteins  
**0000 0002 9699 485X**  
412  
Dr. Mark Wilkinson

**contact Point**  
**description**

**Dataset**  
**Basic Container**  
**Collection**  
**Wilkinson.rdf**

Takes a SPARQL query of the UniProt database specific to proteins and their GO annotations related to RNA Procssing proteins in Aspergillus and makes it a FAIR Accessor source. The precise query is:

```
PREFIX up:<http://purl.uniprot.org/core/>  
PREFIX taxon:<http://purl.uniprot.org/taxonomy/>  
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>  
SELECT distinct ?id
```

```
WHERE
```

```
{  
  ?protein a up:Protein .  
  ?protein up:organism ?organism .  
  ?organism rdfs:subClassOf taxon:162425 .  
  ?protein up:classifiedWith ?go .  
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .  
  
  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)  
}
```

**identifier**

**Uni Prot Accessor**

Note that this Metadata is about ME! I am the creator of this dataset, and may be credited for it.

UniProt Slice FAIR Accessor -  
Aspergillus RNA Processing  
proteins

creator wilkinsonlab.info/  
language eng  
license cc by nd4.0

title UniProt Slice FAIR Accessor - Aspergillus RNA Processing proteins  
authored By 0000 0002 9699 485X  
entities 412  
term has Principal Investigator Dr. Mark Wilkinson

type Dataset  
Basic Container  
Collection

contact Point Wilkinson.rdf  
description Takes a SPARQL query of the UniProt database specific to proteins and their GO annotations related to RNA Procressing proteins in Aspergillus and makes it a FAIR Accessor source. The precise query is:



```
PREFIX up:<http://purl.uniprot.org/core/>  
PREFIX taxon:<http://purl.uniprot.org/taxonomy/>  
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>  
SELECT distinct ?id
```

```
WHERE  
{  
  ?protein a up:Protein .  
  ?protein up:organism ?organism .  
  ?organism rdfs:subClassOf taxon:162425 .  
  ?protein up:classifiedWith ?go .  
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .  
  
  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)  
}
```

identifier

Uni Prot Accessor

## UniProt Slice FAIR Accessor - Aspergillus RNA Processing proteins

creator  
language  
license  
title  
authored By  
entities  
term has Principal Investigator  
type

[wilkinsonlab.info/](http://wilkinsonlab.info/)

eng

cc by nd4.0

UniProt Slice FAIR Accessor - Aspergillus RNA Processing proteins

0000 0002 9699 485X

412

Dr. Mark Wilkinson

Dataset

Basic Container

Collection

Wilkinson.rdf

contact Point  
description

Takes a SPARQL query of the UniProt database specific to proteins and their GO annotations related to RNA Procassing proteins in Aspergillus and makes it a FAIR Accessor source. The precise query is:

This is how I selected the data for my experiment

→ reproducibility!!

```
PREFIX up:<http://purl.uniprot.org/core/>
PREFIX taxon:<http://purl.uniprot.org/taxonomy/>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?id

WHERE
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .

  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}
```

identifier

Uni Prot Accessor

```
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?id
```

```
WHERE
```

```
{
```

```
  ?protein a up:Protein .
```

```
  ?protein up:organism ?organism .
```

```
  ?organism rdfs:subClassOf taxon:162425 .
```

```
  ?protein up:classifiedWith ?go .
```

```
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .
```

```
  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
```

```
}
```

**identifier**

**keyword**

**landing Page**

**language**

**publisher**

**theme**

**contains**

**Uni Prot Accessor**

Aspergillus nidulans

Aspergillus

Proteins

RNA Processing

[uniprot.org/](http://uniprot.org/)

[en](#)

[wilkinsonlab.info/](http://wilkinsonlab.info/)

[RNA Processing conceptscheme.rdf](#)

[C8UZX9](#)

[C8UZY5](#)

[C8V0B4](#)

[C8V0M2](#)

[C8V0U7](#)

```
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>  
SELECT distinct ?id
```

WHERE

```
{  
  ?protein a up:Protein .  
  ?protein up:organism ?organism .  
  ?organism rdfs:subClassOf taxon:162425 .  
  ?protein up:classifiedWith ?go .  
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .  
  
  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)  
}
```

identifier

keyword

landing Page

language

publisher

theme

contains

**Uni Prot Accessor**

Aspergillus nidulans

Aspergillus

Proteins

RNA Processing

[uniprot.org/](http://uniprot.org/)

en

[wilkinsonlab.info/](http://wilkinsonlab.info/)

[RNA Processing conceptscheme.rdf](#)

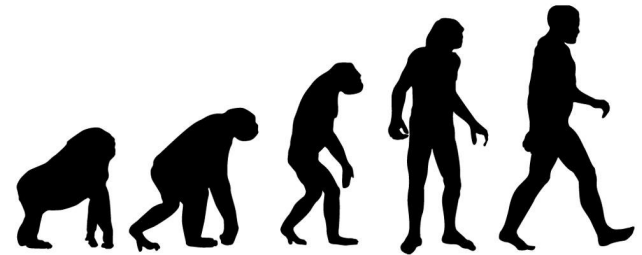
[C8UZX9](#)

[C8UZY5](#)

[C8V0B4](#)

[C8V0M2](#)

[C8V0I7](#)



PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>  
SELECT distinct ?id

```
WHERE  
{  
  ?protein a up:Protein .  
  ?protein up:organism ?organism .  
  ?organism rdfs:subClassOf taxon:162425 .  
  ?protein up:classifiedWith ?go .  
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .  
  
  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)  
}
```

identifier	<a href="#">Uni Prot Accessor</a>
keyword	Aspergillus nidulans Aspergillus Proteins RNA Processing
landing Page	<a href="#">uniprot.org/</a>
language	en
publisher	<a href="#">wilkinsonlab.info/</a>
theme	<a href="#">RNA Processing conceptscheme.rdf</a>
contains	C8UZX9 C8UZY5 C8V0B4 C8V0M2 C8V0I17



**THIS IS IMPORTANT!**  
**REMEMBER THIS!!!**





162425	type	Concept
	pref Label	Aspergillus nidulans
162425	type	Concept
	pref Label	Aspergillus nidulans
GO 0002098	type	Concept
	pref Label	tRNA wobble uridine modification
GO 0004479	type	Concept
	pref Label	methionyl-tRNA formyltransferase activity
GO 0004808	type	Concept
	pref Label	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase activity
GO 0006378	type	Concept
	pref Label	mRNA polyadenylation
GO 0006388	type	Concept
	pref Label	tRNA splicing, via endonucleolytic cleavage and ligation
GO 0006400	type	Concept
	pref Label	tRNA modification
GO 0030488	type	Concept
	pref Label	tRNA methylation
GO 0019988	type	Concept
	pref Label	charged-tRNA amino acid modification
GO 0016436	type	Concept
	pref Label	rRNA (uridine) methyltransferase activity
GO 0016434	type	Concept
	pref Label	rRNA (cytosine) methyltransferase activity
GO 0031119	type	Concept
	pref Label	tRNA pseudouridine synthesis
GO 0031167	type	Concept
	pref Label	rRNA methylation
GO 0034227	type	Concept
	pref Label	tRNA thio-modification
GO 0070038	type	Concept
	pref Label	rRNA (pseudouridine-N3-)-methyltransferase activity
GO 0034470	type	Concept
	pref Label	ncRNA processing

```

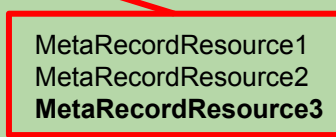
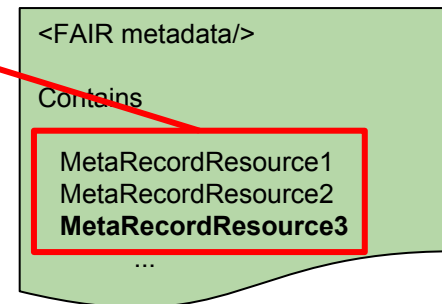
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?id

WHERE
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .

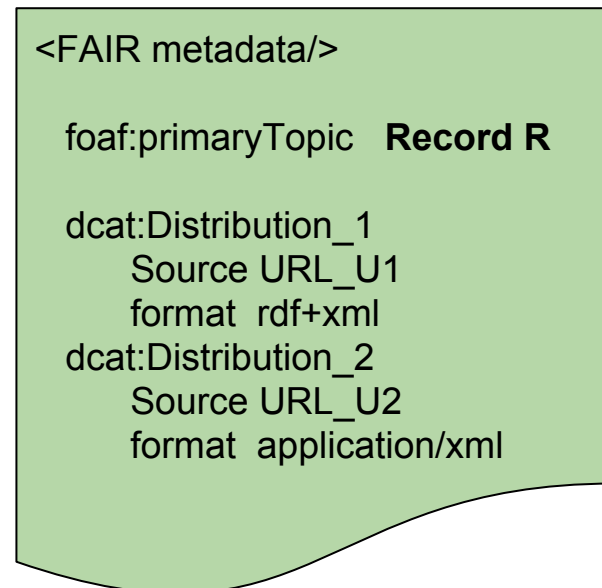
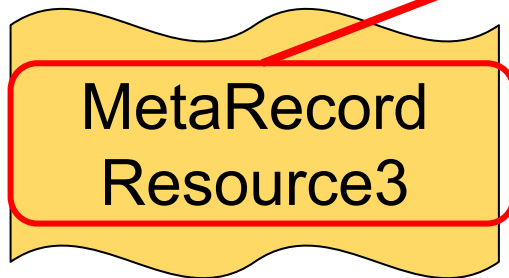
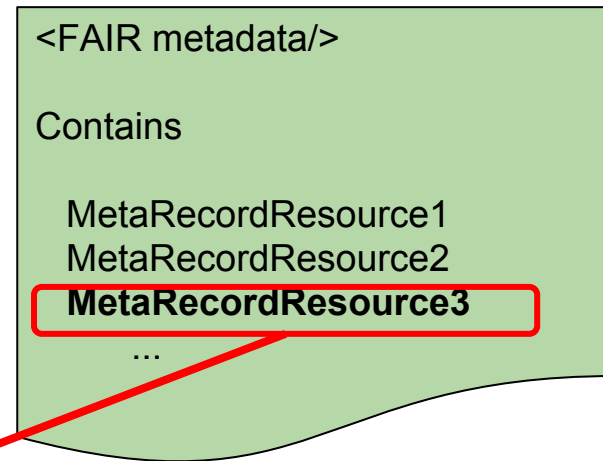
  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}

```

identifier	<b>Uni Prot Accessor</b>
keyword	Aspergillus nidulans Aspergillus Proteins RNA Processing
landing Page	<a href="http://uniprot.org/">uniprot.org/</a>
language	en
publisher	<a href="http://wilkinsonlab.info/">wilkinsonlab.info/</a>
theme	<a href="#">RNA Processing conceptscheme.rdf</a>
contains	<b>C8UZX9</b> <b>C8UZY5</b> <b>C8V0B4</b> <b>C8V0M2</b> <b>C8V0I7</b>



# Step down to individual Record metadata



# Step down to individual Record metadata

MetaRecord  
Resource3

HTTP GET

```
<FAIR metadata/>

foaf:primaryTopic  Record R

dcat:Distribution_1
  Source URL_U1
  format  rdf+xml

dcat:Distribution_2
  Source URL_U2
  format  application/xml
```



Software calls HTTP GET on the URL  
representing the MetaRecord Resource  
for the desired record in the Container

(or just click on it, or type it into your browser)



UniProt Protein C8VBH1	bibliographic Citation	The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212
	creator	UniProt Consortium
	language	eng
	license	3.0/
	title	UniProt Protein C8VBH1
	Version	UniProt release 2016_09
	in dataset	Uni Prot Accessor/
	contact point	contact
	description	Exonuclease, putative (AFU_orthologue AFUA_2G05560)
	distribution	<a href="http://fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&amp;predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Fclassified%2Ffragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&amp;predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Forganism">fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&amp;predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Fclassified%2Ffragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&amp;predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Forganism</a>
	...	C8VBH1.html
	identifier	C8VBH1.rdf
	keyword	C8VBH1 Annotation Aspergillus nidulans Aspergillus Functional Annotation GO Gene Ontology Proteins RNA Processing
	landing page	uniprot.org
	language	en
	publisher	uniprot.org
	page	sparql
	primary topic	uniprot.org/ C8VBH1
C8VBH1	...	Source7968453A 9A8F 11E6 A83B A5E65C07C3DD
Mappings7968453A 9A8F 11E6 A83B A5E65C07C3DD	logical Source	POMap7968453A 9A8F 11E6 A83B A5E65C07C3DD
	predicate Object Map	Subject Map7968453A 9A8F 11E6 A83B A5E65C07C3DD
	subject Map	Source7968453A 9A8F 11E6 A83B A5E65C07C3DD
Mappings7968453A 9A8F 11E6 A83B A5E65C07C3DD	logical Source	POMap7968453A 9A8F 11E6 A83B A5E65C07C3DD
	predicate Object Map	Subject Map7968453A 9A8F 11E6 A83B A5E65C07C3DD
	subject Map	Subject Map27968453A 9A8F 11E6 A83B A5E65C07C3DD
Object Map7968453A 9A8F 11E6 A83B A5E65C07C3DD	parent Triples Map	Subject Map27968453A 9A8F 11E6 A83B A5E65C07C3DD
Object Map7968453A 9A8F 11E6 A83B A5E65C07C3DD	parent Triples Map	Object Map7968453A 9A8F 11E6 A83B A5E65C07C3DD
POMap7968453A 9A8F 11E6 A83B A5E65C07C3DD	object Map	organism
	predicate	Object Map7968453A 9A8F 11E6 A83B A5E65C07C3DD
POMap7968453A 9A8F 11E6 A83B A5E65C07C3DD	object Map	classified With
	predicate	Mappings7968453A 9A8F 11E6 A83B A5E65C07C3DD
Source7968453A 9A8F 11E6 A83B A5E65C07C3DD	Mapping	Triple Pattern Fragments
	reference Formulation	fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Forganism
	source	Mappings7968453A 9A8F 11E6 A83B A5E65C07C3DD
Source7968453A 9A8F 11E6 A83B A5E65C07C3DD	Mapping	Triple Pattern Fragments
	reference Formulation	fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Fclassified%2F
	source	data 1179
Subject Map27968453A 9A8F 11E6 A83B A5E65C07C3DD	class	<a href="http://identifiers.org/taxon/(TAX)">http://identifiers.org/taxon/(TAX)</a>
Subject Map27968453A 9A8F 11E6 A83B A5E65C07C3DD	template	data 1176
	class	<a href="http://purl.obolibrary.org/obo/(GO)">http://purl.obolibrary.org/obo/(GO)</a>
Subject Map7968453A 9A8F 11E6 A83B A5E65C07C3DD	template	organism
	class	<a href="http://identifiers.org/uniprot/(ID)">http://identifiers.org/uniprot/(ID)</a>
Subject Map7968453A 9A8F 11E6 A83B A5E65C07C3DD	template	data 0896
	class	<a href="http://identifiers.org/uniprot/(ID)">http://identifiers.org/uniprot/(ID)</a>
	template	application/rdf+xml
	format	application/x-turtle
fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Fclassified%2F	text/html	
	type	Projector
		Dataset
		dataset
		Distribution
	download URL	<a href="http://fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&amp;predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Fclassified%2F">fragments?subject=http%3A%2F%2Fidentifiers%2Eorg%2Funiprot%2FC8VBH1&amp;predicate=http%3A%2F%2Fpurl%2Euniprot%2Eorg%2Fcore%2Fclassified%2F</a>

## The document that is returned

<FAIR metadata/>

foaf:primaryTopic up:C8UZX9

dcat:Distribution\_1  
Source URL\_U1

format rdf+xml

dcat:Distribution\_2  
Source URL\_U2

format application/xml

**UniProt Protein  
C8UZX9**

**bibliographic Citation**

The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212

**creator**

UniProt Consortium

**language**

**eng**

**license**

**cc by nd3.0**

**title**

UniProt Protein C8UZX9

**in Dataset**

**Uni Prot Accessor/  
contact**

**contact Point**

**contact**

**description**

KRR1 small subunit processome componentKRR-R motif-containing protein 1

**distribution**

**C8UZX9.rdf**

**C8UZX9.html**

**C8UZX9**

**identifier**

**keyword**

Annotation

Aspergillus nidulans

Aspergillus

Functional Annotation

GO

Gene Ontology

Proteins

RNA Processing

**landing Page**

**uniprot.org**

**language**

**en**

**publisher**

**uniprot.org**

**page**

**sparql**

**uniprot.org/  
C8UZX9**

**primary topic**

**C8UZX9**

...

**format**

application/rdf+xml

**type**

**Dataset**

**Dataset**

**Distribution**

**download URL**

**C8UZX9.rdf**

**format**

text/html

**type**

**Dataset**

**Distribution**

**download URL**

**C8UZX9.html**

**C8UZX9**

**C8UZX9.rdf**

**C8UZX9.html**

**UniProt Protein  
C8UZX9**

<b>bibliographic Citation</b>	The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212
<b>creator</b>	UniProt Consortium
<b>language</b>	eng
<b>license</b>	cc by nd3.0
<b>title</b>	UniProt Protein C8UZX9
<b>in Dataset</b>	<b>Uni Prot Accessor/ contact</b>
<b>contact Point</b>	KRR1 small subunit processome componentKRR-R motif-containing protein 1
<b>description</b>	<b>C8UZX9.rdf C8UZX9.html C8UZX9</b>
<b>distribution</b>	Annotation Aspergillus nidulans Aspergillus Functional Annotation GO Gene Ontology Proteins RNA Processing
<b>identifier</b>	<b>uniprot.org</b>
<b>keyword</b>	en <b>uniprot.org sparql uniprot.org/ C8UZX9</b>
<b>landing Page</b>	<b>primary topic</b>
<b>language</b>	...
<b>publisher</b>	<b>format</b>
<b>page</b>	application/rdf+xml
	<b>type</b>
	<b>Dataset Dataset Distribution</b>
	<b>download URL</b>
	<b>C8UZX9.rdf</b>
	<b>format</b>
	text/html
	<b>type</b>
	<b>Dataset Distribution</b>
	<b>download URL</b>
	<b>C8UZX9.html</b>



Note the change in metadata focus!

This metadata is about the UniProt Record (not about Mark Wilkinson).

The record described in this metadata was created by UniProt, so the citation and authorship information is now THEIRS, not MINE.

C8UZX9  
C8UZX9.rdf

C8UZX9.html

**UniProt Protein  
C8UZX9**

<b>bibliographic Citation</b>	The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212
<b>creator</b>	UniProt Consortium
<b>language</b>	eng
<b>license</b>	cc by nd3.0
<b>title</b>	UniProt Protein C8UZX9
<b>in Dataset</b>	<b>Uni Prot Accessor/ contact</b>
<b>contact Point</b>	KRR1 small subunit processome componentKRR-R motif-containing protein 1
<b>description</b>	
<b>distribution</b>	<b>C8UZX9.rdf</b> <b>C8UZX9.html</b> <b>C8UZX9</b>
<b>identifier</b>	Annotation Aspergillus nidulans Aspergillus Funcinal Annotation GO Gene Ontology Proteins RNA Processing
<b>keyword</b>	
<b>landing Page</b>	<b>uniprot.org</b>
<b>language</b>	en
<b>publisher</b>	<b>uniprot.org</b>
<b>page</b>	<b>sparql</b> <b>uniprot.org/ C8UZX9</b>
<b>primary topic</b>	
...	
<b>format</b>	application/rdf+xml
<b>type</b>	<b>Dataset</b> <b>Dataset</b> <b>Distribution</b>
<b>download URL</b>	<b>C8UZX9.rdf</b>
<b>format</b>	text/html
<b>type</b>	<b>Dataset</b> <b>Distribution</b>
<b>download URL</b>	<b>C8UZX9.html</b>



**Container  
Resource**

**Symmetrical Link  
back upward to the Accessor  
Container, for additional  
metadata**

**C8UZX9  
C8UZX9.rdf**

**C8UZX9.html**



**UniProt Protein  
C8UZX9**

<b>bibliographic Citation</b>	The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212
<b>creator</b>	UniProt Consortium
<b>language</b>	eng
<b>license</b>	cc by nd3.0
<b>title</b>	UniProt Protein C8UZX9
<b>in Dataset</b>	<b>Uni Prot Accessor/</b>
<b>contact Point</b>	<b>contact</b>
<b>description</b>	KRR1 small subunit processome componentKRR-R motif-containing protein 1
<b>distribution</b>	<b>C8UZX9.rdf</b> <b>C8UZX9.html</b>
<b>identifier</b>	<b>C8UZX9</b>
<b>keyword</b>	Annotation Aspergillus nidulans Aspergillus Funcinal Annotation GO Gene Ontology Proteins RNA Processing
<b>landing Page</b>	<b>uniprot.org</b>
<b>language</b>	en
<b>publisher</b>	<b>uniprot.org</b>
<b>page</b>	<b>sparql</b> <b>uniprot.org/</b> <b>C8UZX9</b>
<b>primary topic</b>	<b>C8UZX9</b>

<FAIR metadata/>

foaf:primaryTopic **Record R**

dcat:Distribution\_1  
Source URL\_U1  
format rdf+xml

dcat:Distribution\_2  
Source URL\_U2  
format application/xml

<b>C8UZX9</b>	...
<b>C8UZX9.rdf</b>	<b>format</b> application/rdf+xml <b>type</b> Dataset Dataset Distribution
<b>C8UZX9.html</b>	<b>download URL</b> C8UZX9.rdf <b>format</b> text/html <b>type</b> Dataset Distribution
	<b>download URL</b> C8UZX9.html

**UniProt Protein  
C8UZX9**

**bibliographic Citation**

The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212

**creator**

UniProt Consortium

**language**

**eng**

**license**

**cc by nd3.0**

**title**

UniProt Protein C8UZX9

**in Dataset**

**Uni Prot Accessor/  
contact**

**contact Point**

**contact**

**description**

KRR1 small subunit processome componentKRR-R motif-containing protein 1

**distribution**

**C8UZX9.rdf**

**C8UZX9.html**

**identifier**

**C8UZX9**

**keyword**

Annotation

Aspergillus nidulans

Aspergillus

Functional Annotation

GO

Gene Ontology

Proteins

RNA Processing

**landing Page**

**uniprot.org**

**language**

**en**

**publisher**

**uniprot.org**

**page**

**sparql**

**uniprot.org/  
C8UZX9**

**primary topic**

**C8UZX9**

**C8UZX9**

**C8UZX9.rdf**

...

**format**

application/rdf+xml

**type**

**Dataset**

**Dataset**

**Distribution**

**download URL**

**C8UZX9.rdf**

**C8UZX9.html**

**format**

text/html

**type**

**Dataset**

**Distribution**

**download URL**

**C8UZX9.html**

Two ways to retrieve the record - RDF or HTML

(in REST-speak, two Representations  
of that Resource)

**UniProt Protein  
C8UZX9**

**bibliographic Citation**

The UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-D212

**creator**

UniProt Consortium

**language**

eng

**license**

cc by nd3.0

**title**

UniProt Protein C8UZX9

**in Dataset**

**Uni Prot Accessor/  
contact**

**contact Point**

**contact**

**description**

KRR1 small subunit processome componentKRR-R motif-containing protein 1

**distribution**

**C8UZX9.rdf**

**C8UZX9.html**

**C8UZX9**

**identifier**

**keyword**

Annotation

Aspergillus nidulans

Aspergillus

Functional Annotation

GO

Gene Ontology

Proteins

RNA Processing

**uniprot.org**

**landing Page**

**language**

en

**publisher**

**uniprot.org**

**page**

**sparql**

**uniprot.org/  
C8UZX9**

**primary topic**

...

**format**

application/rdf+xml

**type**

**Dataset**

**Dataset**

**Distribution**

**C8UZX9.rdf**

**download URL**

**format**

text/html

**type**

**Dataset**

**Distribution**

**C8UZX9.html**

**download URL**

Note that this metadata record is somewhat more FAIR, than what you can (easily) retrieve from UniProt itself!

e.g. the UniProt record does not include the citation or license information - you have to manually surf around the UniProt Web page to find that.

So the Accessor makes UniProt's already notably FAIR data, even more FAIR (with respect to "R")

**C8UZX9**

**C8UZX9.rdf**

**C8UZX9.html**

# How FAIR are we now?

What does the Accessor give us?

# What we have achieved



F

We have created a FAIR record for something - i.e. a slice of a database - that was, historically, un-recordable and un-identifiable in any formal way.



F + R

Accessors are a standard approach to providing human & machine accessible metadata to facilitate appropriate discovery (contextual, biological), proper usage (license) and proper citation for any kind of data.



F + A

The discovery, accessibility, and drill-down/up behaviors do not require any novel API, rather simply rely on global Web standards; this allows them to be indexed by existing Web search engines

# What we have achieved



The metadata itself uses machine-accessible syntaxes, and widely adopted ontologies and vocabularies, thus easily integrates with other metadata



Accessors provide a lightweight means to protect privacy while still providing the maximum degree of transparency possible



Accessors can be static, or dynamic. i.e. we can provide template Accessor file(s) that are edited in Notepad, then published together with the data; or Accessors can dynamically generate their output from code (e.g. layered on a database server)

# Building our own Metadata record

Objective: Build a “FAIR Accessor” for the Triples that you published yesterday.

Tools we will use:

- 1) SKOS Concept Scheme Builder
  - a) “SKOS Play!” <http://labs.sparna.fr/skos-play/convert>
  
- 2) Linked Data Platform Server (***Virtuoso*** running on my own Google Cloud machine)
  - a) <http://training.fairdata.solutions/DAV/home/LDP/gofair/>
  
- 3) FAIR Data Point editor (this afternoon with Luiz)
  - a) <https://fair-course.fair-dtls.surf-hosted.nl/editor>

```
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?id
```

```
WHERE
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .

  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}
```

identifier	<a href="#">Uni Prot Accessor</a>
keyword	Aspergillus nidulans Aspergillus Proteins RNA Processing
landing Page	<a href="#">uniprot.org/</a>
language	en
publisher	<a href="#">wilkinsonlab.info/</a>
theme	<a href="#">RNA Processing conceptscheme.rdf</a>
contains	C8UZX9 C8UZY5 C8V0B4 C8V0M2 C8V0I17



**THIS IS IMPORTANT!**  
**REMEMBER THIS!!!**  
**Now we will create one.....**



# SKOS Concept Schemes and SKOS Play!

Consideration #1: build a SKOS Concept Scheme that describes our dataset

- I usually do this using my own boutique software (more power)
- Today, we will use an easy online tool to reduce complexity (less power)

SKOS Play is hosted at: <http://labs.sparna.fr/skos-play/convert>

It provides an Excel template that you edit, then upload

The output is a SKOS Concept Scheme in RDF-Turtle format

Surf to that page now....

### Where is the Excel file you want to convert ?

- In one of the included example**  
Example 1 (simple exemple, in english) ▼  
Download example: **Example 1 (simple exemple, in english)** **DOWNLOAD THIS**
- In a local file on my computer**  
excel2skos-HRB-simple.xlsx Change Remove  
*(Supported extensions : .xls or .xlsx - OpenOffice is not supported !)*
- On the web**  
http://...  
*(a link to an excel file available online)*
- In a Google spreadsheet**  
Login with your Google account (using the link above) to convert a file from your Drive !

### What is the default language of the labels ?

Enter/select language

# SKOS Play! Excel template

<u>ConceptScheme</u> URI	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl">http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl</a>	
dct:title	Pest Observations Dataset	
dct:description	FAIR Data representation of a Pest observation dataset	
URI	<u>skos:prefLabel</u>	<u>skos:definition</u>
<a href="http://purl.obolibrary.org/obo/NCIT_C16270">http://purl.obolibrary.org/obo/NCIT_C16270</a>	Agriculture	The practice of cultivating the land or raising stock, as well as the work, business, or study of farming
<a href="http://semanticscience.org/resource/SIO_010414">http://semanticscience.org/resource/SIO_010414</a>	pathogen	A pest or pathogen
<a href="http://www.ontotext.com/proton/protonext#Food">http://www.ontotext.com/proton/protonext#Food</a>	food	food is something you eat
<a href="http://opendata.aragon.es/def/ei2a#Crop">http://opendata.aragon.es/def/ei2a#Crop</a>	Crop	An agricultural plant
<a href="http://www.disit.org/km4city/schema#AgricultureAndLivestock">http://www.disit.org/km4city/schema#AgricultureAndLivestock</a>	Agriculture	Of interest to agriculture

# SKOS Play! Excel template

These fields can be whatever you want them to be

<b>ConceptScheme URI</b>	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl">http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest Scheme.ttl</a>	
<b>dct:title</b>	Pest Observations Dataset	
<b>dct:description</b>	FAIR Data representation of a Pest observation dataset	
<hr/>		
<b>URI</b>		
	<b><u>skos:prefLabel</u></b>	<b><u>skos:definition</u></b>
<a href="http://purl.obolibrary.org/obo/NCIT_C16270">http://purl.obolibrary.org/obo/NCIT_C16270</a>	Agriculture	The practice of cultivating the land or raising stock, as well as the work, business, or study of farming
<a href="http://semanticscience.org/resource/SIO_010414">http://semanticscience.org/resource/SIO_010414</a>	pathogen	A pest or pathogen
<a href="http://www.ontotext.com/proton/protonext#Food">http://www.ontotext.com/proton/protonext#Food</a>	food	food is something you eat
<a href="http://opendata.aragon.es/def/ei2a#Crop">http://opendata.aragon.es/def/ei2a#Crop</a>	Crop	An agricultural plant
<a href="http://www.disit.org/km4city/schema#AgricultureAndLivestock">http://www.disit.org/km4city/schema#AgricultureAndLivestock</a>	Agriculture	Of interest to agriculture
<hr/>		



# SKOS Play! Excel template

These fields will be selected when we use the Ontology Lookup Service (OLS) at EBI or LOV database

<u>ConceptScheme</u> URI	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl">http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl</a>	
dct:title	Pest Observations Dataset	
dct:description	FAIR Data representation of a Pest observation dataset	
URI	<u>skos:prefLabel</u>	<u>skos:definition</u>
<a href="http://purl.obolibrary.org/obo/NCIT_C16270">http://purl.obolibrary.org/obo/NCIT_C16270</a>	Agriculture	The practice of cultivating the land or raising stock, as well as the work, business, or study of farming
<a href="http://semanticscience.org/resource/SIO_010414">http://semanticscience.org/resource/SIO_010414</a>	pathogen	A pest or pathogen
<a href="http://www.ontotext.com/proton/protonext#Food">http://www.ontotext.com/proton/protonext#Food</a>	food	food is something you eat
<a href="http://opendata.aragon.es/def/ei2a#Crop">http://opendata.aragon.es/def/ei2a#Crop</a>	Crop	An agricultural plant
<a href="http://www.disit.org/km4city/schema#AgricultureAndLivestock">http://www.disit.org/km4city/schema#AgricultureAndLivestock</a>	Agriculture	Of interest to agriculture

# SKOS Play! Excel template

These fields are at our discretion; however, generally speaking, you should copy the information from the original ontology

<b>ConceptScheme URI</b>	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl">http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest_Scheme.ttl</a>	
<b>dct:title</b>	Pest Observations Dataset	
<b>dct:description</b>	FAIR Data representation of a Pest observation dataset	
URI	skos:prefLabel	skos:definition
<a href="http://purl.obolibrary.org/obo/NCIT_C16270">http://purl.obolibrary.org/obo/NCIT_C16270</a>	Agriculture	The practice of cultivating the land or raising stock, as well as the work, business, or study of farming
<a href="http://semanticscience.org/resource/SIO_010414">http://semanticscience.org/resource/SIO_010414</a>	pathogen	A pest or pathogen
<a href="http://www.ontotext.com/proton/protonext#Food">http://www.ontotext.com/proton/protonext#Food</a>	food	food is something you eat
<a href="http://opendata.aragon.es/def/ei2a#Crop">http://opendata.aragon.es/def/ei2a#Crop</a>	Crop	An agricultural plant
<a href="http://www.disit.org/km4city/schema#AgricultureAndLivestock">http://www.disit.org/km4city/schema#AgricultureAndLivestock</a>	Agriculture	Of interest to agriculture

Let's go!



https://www.ebi.ac.uk/ols/search



# Ontology Lookup Service

[Home](#) [Ontologies](#) [Documentation](#) [About](#)

OLS > Search

agriculture

Exact match  Obsolete terms

## Filters

### Term type

Filter by type

class **92**

individual **1**

property **1**

### Ontologies

Filter by ontology

## Search results for *agriculture*

[Previous](#) Showing **1** to **10** of **94** results [Next](#)

### [Agriculture](#) **NCIT:C16270**

[http://purl.obolibrary.org/obo/NCIT\\_C16270](http://purl.obolibrary.org/obo/NCIT_C16270)

The practice of cultivating the land or raising stock, as well as the work, business, or study of farming.

**Ontology:** [NCI Thesaurus OBO Edition](#) **NCIT**

### [Agriculture](#) **OMIT:0001811**

[http://purl.obolibrary.org/obo/OMIT\\_0001811](http://purl.obolibrary.org/obo/OMIT_0001811)

**Ontology:** [Ontology for MIRNA Target](#) **OMIT**



# Ontology Lookup Service

Home | Ontologies | Documentation | About

OLS > Search

agriculture

Exact match  Obsolete terms

## Filters

### Term type

Filter by type

class **92**

individual **1**

property **1**

### Ontologies

Filter by ontology

## Search results for *agriculture*

Previous Showing 1 to 10 of 94 results Next

### **Agriculture** NCIT:C16270

[http://purl.obolibrary.org/obo/NCIT\\_C16270](http://purl.obolibrary.org/obo/NCIT_C16270)

The practice of cultivating the land or raising stock, as well as the work, business, or study of farming.

**Ontology:** [NCI Thesaurus OBO Edition](#) NCIT

### **Agriculture** OMIT:0001811

[http://purl.obolibrary.org/obo/OMIT\\_0001811](http://purl.obolibrary.org/obo/OMIT_0001811)

**Ontology:** [Ontology for MIRNA Target](#) OMIT

https://www.ebi.ac.uk/ols/search

The screenshot shows the OLS interface for the term 'Agriculture'. The breadcrumb path is 'OLS > NCI Thesaurus OBO Edition > NCIT > NCIT:C16270'. The term 'Agriculture' is displayed in large blue font. Below it is the URI: [http://purl.obolibrary.org/obo/NCIT\\_C16270](http://purl.obolibrary.org/obo/NCIT_C16270). The definition is: 'The practice of cultivating the land or raising stock, as well as the work, business, or study of farming'. Below the definition are synonyms: 'Agriculture'. There are also options for 'Tree view', 'Term history', 'Conceptual Entity', 'Occupation or Discipline', and 'Technology'. A red box highlights the term, URI, and definition. Red arrows point from labels on the right to these elements.

Label

URI

Definition

**SKOS Play Excel Sheet:**

URI	<u>skos:prefLabel</u>	<u>skos:definition</u>
<a href="http://purl.obolibrary.org/obo/NCIT_C16270">http://purl.obolibrary.org/obo/NCIT_C16270</a>	Agriculture	The practice of cultivating the land or raising stock, as well as the work, business, or study of farming

Lather, rinse, repeat

URI	<u>skos:prefLabel</u>	<u>skos:definition</u>
<a href="http://purl.obolibrary.org/obo/NCIT_C16270">http://purl.obolibrary.org/obo/NCIT_C16270</a>	Agriculture	The practice of cultivating the land or raising stock, as well as the work, business, or study of farming
<a href="http://semanticscience.org/resource/SIO_010414">http://semanticscience.org/resource/SIO_010414</a>	pathogen	A pest or pathogen
<a href="http://www.ontotext.com/proton/protonext#Food">http://www.ontotext.com/proton/protonext#Food</a>	food	food is something you eat
<a href="http://opendata.aragon.es/def/ei2a#Crop">http://opendata.aragon.es/def/ei2a#Crop</a>	Crop	An agricultural plant
<a href="http://www.disit.org/km4city/schema#AgricultureAndLivestock">http://www.disit.org/km4city/schema#AgricultureAndLivestock</a>	Agriculture	Of interest to agriculture

Save Excel File as  
**Pest\_Scheme.xlsx**

Now, we need to decide where the resulting SKOS file will be published....

# Your decision really depends on your own compute facilities

I am now going to show you how to use a Linked Data Platform server (the technology we discussed earlier - the Russian Dolls - that forms the basis of the FAIR Accessor)

W3C Recommendation



Linked Data Platform 1.0

W3C Recommendation 26 February 2015

The LDP Specification describes a universally-applicable approach to  
Storing and retrieving Linked Data (triples)

It follows (almost) the REST design philosophy

It's surprisingly easy to understand and use!

I have set-up a public LDP  
Server for you

<http://training.fairdata.solutions/DAV/home/LDP/gofair/>

You are welcome to play with this!  
(it will disappear after the course)

For the moment, please just watch  
because I want to explain the LDP  
REST behavior “the hard way”  
(Luiz will show you a tool this afternoon)



Follow redirects

Execute

Please send me  
RDF in Turtle  
format

Which web address? ("Resource")

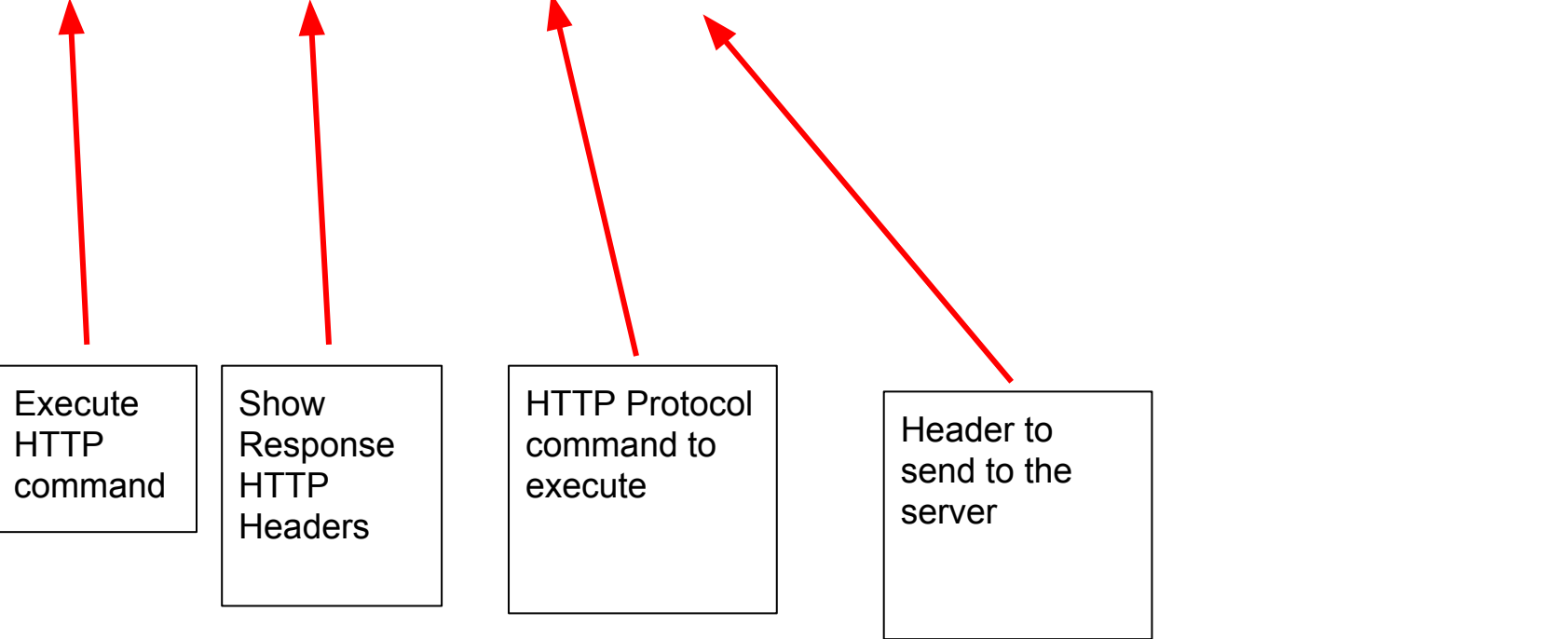
**curl -L -i -X GET -H "Accept: text/turtle" http://....**

Execute  
HTTP  
command

Show  
Response  
HTTP  
Headers

HTTP Protocol  
command to  
execute

Header to  
send to the  
server



```
curl -L -i  
  -X GET  
  -H "Accept: text/turtle"  
  "http://training.fairdata.solutions/DAV/home/LDP/gofair/"  
  
| less
```

```
HTTP/1.1 200 OK
Date: Thu, 07 Mar 2019 14:29:39 GMT
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
X-Frame-Options: SAMEORIGIN
Accept-Ranges: bytes
Allow: GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH
Vary: Accept,Origin,If-Modified-Since,If-None-Match
MS-Author-Via: DAV, SPARQL
Accept-Patch: application/sparql-update
Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
Link: <?p=1>; rel="first"
Link: <?p=1>; rel="last"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair,meta>; rel="meta"; title="Metadata File"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair,acl>; rel="acl"; title="Access Control File"
X-SPARQL-default-graph: http://training.fairdata.solutions/DAV/home/LDP/gofair/
ETag: "dd2ce6661f9e91df0b51363cef286259"
Content-Type: text/turtle
Content-Length: 263866

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix ns1: <http://training.fairdata.solutions/DAV/home/LDP/gofair/> .
@prefix ldp: <http://www.w3.org/ns/ldp#> .
ns1:obs_2147365908 rdf:type ldp:Resource .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
ns1:obs_2147365908 rdf:type rdfs:Resource .
@prefix ns4: <http://www.w3.org/ns/posix/stat#> .
ns1:obs_2147365908 ns4:mtime 1551964987 ;
ns4:size 2783 .
ns1:obs_2147365880 rdf:type ldp:Resource ,
rdfs:Resource ;
ns4:mtime 1551965977 .
```

```
HTTP/1.1 200 OK
Date: Thu, 07 Mar 2019 14:29:39 GMT
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
X-Frame-Options: SAMEORIGIN
Accept-Ranges: bytes
Allow: GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH
Vary: Accept,Origin,If-Modified-Since,If-None-Match
MS-Author-Via: DAV, SPARQL
Accept-Patch: application/sparql-update
Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
Link: <?p=1>; rel="first"
Link: <?p=1>; rel="last"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair,meta>; rel="meta"; title="Metadata File"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair,acl>; rel="acl"; title="Access Control File"
X-SPARQL-default-graph: http://training.fairdata.solutions/DAV/home/LDP/gofair/
ETag: "dd2ce6661f9e91df0b51363cef286259"
Content-Type: text/turtle
Content-Length: 263866
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix ns1: <http://training.fairdata.solutions/DAV/home/LDP/gofair/> .
@prefix ldp: <http://www.w3.org/ns/ldp#> .
ns1:obs_2147365908 rdf:type ldp:Resource .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
ns1:obs_2147365908 rdf:type rdfs:Resource .
@prefix ns4: <http://www.w3.org/ns/posix/stat#> .
ns1:obs_2147365908 ns4:mtime 1551964987 ;
ns4:size 2783 .
ns1:obs_2147365880 rdf:type ldp:Resource ,
rdfs:Resource ;
ns4:mtime 1551965977 .
```

```
HTTP/1.1 200 OK
Date: Thu, 07 Mar 2019 14:29:39 GMT
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
X-Frame-Options: SAMEORIGIN
Accept-Ranges: bytes
Allow: GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH
Vary: Accept,Origin,If-Modified-Since,If-None-Match
MS-Author-Via: DAV, SPARQL
Accept-Patch: application/sparql-update
Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
Link: <?p=1>; rel="first"
Link: <?p=1>; rel="last"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair,meta>; rel="meta"; title="Metadata File"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair,acl>; rel="acl"; title="Access Control File"
X-SPARQL-default-graph: http://training.fairdata.solutions/DAV/home/LDP/gofair/
ETag: "dd2ce6661f9e91df0b51363cef286259"
Content-Type: text/turtle
Content-Length: 263866
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix ns1: <http://training.fairdata.solutions/DAV/home/LDP/gofair/> .
@prefix ldp: <http://www.w3.org/ns/ldp#> .
ns1:obs_2147365908 rdf:type ldp:Resource .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
ns1:obs_2147365908 rdf:type rdfs:Resource .
@prefix ns4: <http://www.w3.org/ns/posix/stat#> .
ns1:obs_2147365908 ns4:mtime 1551964987 ;
ns4:size 2783 .
ns1:obs_2147365880 rdf:type ldp:Resource ,
rdfs:Resource ;
ns4:mtime 1551965077 .
```

I want to create a new Container  
For my SKOS Pest Scheme

```
$ cat container.ttl
```

```
@prefix ldp: <http://www.w3.org/ns/ldp#> .
```

```
<>    a ldp:Container,    ldp:BasicContainer .
```

```
curl -k -i  
-X POST "http://training.fairdata.solutions/DAV/home/LDP/gofair/"  
-u gofair:gofair  
-H "Slug: SKOS"  
-H "Content-Type: text/turtle"  
--data-binary @container.ttl
```

```
HTTP/1.1 201 Created
Date: Fri, 08 Mar 2019 08:49:06 GMT
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
X-Frame-Options: SAMEORIGIN
Accept-Ranges: bytes
Location: http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/
Allow: GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH
Vary: Accept,Origin,If-Modified-Since,If-None-Match
MS-Author-Via: DAV, SPARQL
Accept-Patch: application/sparql-update
Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,meta>; rel="meta"; title="Meta
data File"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,acl>; rel="acl"; title="Access
Control File"
ETag: "8b13e48bc32b8ae50c39e43fd6228f98"
Content-Type: text/turtle
Content-Length: 184
```



# HTTP Response Code 201 - Created

Tells us that we just created a new Container

```
HTTP/1.1 201 Created
```

```
Date: Fri, 08 Mar 2019 08:49:06 GMT
```

```
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
```

```
X-Frame-Options: SAMEORIGIN
```

```
Accept-Ranges: bytes
```

```
Location: http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/
```

```
Allow: GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH
```

```
Vary: Accept,Origin,If-Modified-Since,If-None-Match
```

```
MS-Author-Via: DAV, SPARQL
```

```
Accept-Patch: application/sparql-update
```

```
Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json
```

```
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
```

```
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
```

```
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,meta>; rel="meta"; title="Meta data File"
```

```
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,acl>; rel="acl"; title="Access Control File"
```

```
ETag: "8b13e48bc32b8ae50c39e43fd6228f98"
```

```
Content-Type: text/turtle
```

```
Content-Length: 184
```

# HTTP Response Header "Location"

Tells us the URL for that new Container  
(it obeyed the Slug naming suggestion)

<http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/>

```
HTTP/1.1 201 Created
Date: Fri, 08 Mar 2019 08:49:06 GMT
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
X-Frame-Options: SAMEORIGIN
Accept-Ranges: bytes
Location: http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/
Allow: GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH
Vary: Accept,Origin,If-Modified-Since,If-None-Match
MS-Author-Via: DAV, SPARQL
Accept-Patch: application/sparql-update
Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,meta>; rel="meta"; title="Meta
data File"
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,acl>; rel="acl"; title="Access
Control File"
ETag: "8b13e48bc32b8ae50c39e43fd6228f98"
Content-Type: text/turtle
Content-Length: 184
```

```
$ curl -k -i -u gofair:gofair -H "Accept: text/turtle" -X GET
"http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/"
```

```
HTTP/1.1 200 OK
```

```
Date: Fri, 08 Mar 2019 10:50:36 GMT
```

```
Server: Virtuoso/07.20.3230 (Linux) x86_64-generic_glibc25-linux-gnu
```

```
X-Frame-Options: SAMEORIGIN
```

```
Accept-Ranges: bytes
```

```
Allow: GET, HEAD, POST, PUT, DELETE, OPTIONS, PROPFIND, PROPPATCH, COPY, MOVE, MKCOL, LOCK, UNLOCK, TRACE, PATCH
```

```
Vary: Accept, Origin, If-Modified-Since, If-None-Match
```

```
MS-Author-Via: DAV, SPARQL
```

```
Accept-Patch: application/sparql-update
```

```
Accept-Post: text/turtle, text/n3, text/nt, text/html, application/ld+json
```

```
Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"
```

```
Link: <http://www.w3.org/ns/ldp#BasicContainer>; rel="type"
```

```
Link: <?p=1>; rel="first"
```

```
Link: <?p=1>; rel="last"
```

```
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,meta>; rel="meta"; title="Metadata File"
```

```
Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS,acl>; rel="acl"; title="Access Control File"
```

```
X-SPARQL-default-graph: http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/
```

```
ETag: "8b13e48bc32b8ae50c39e43fd6228f98"
```

```
Content-Type: text/turtle
```

```
Content-Length: 431
```

```
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
```

```
@prefix ns1:     <http://training.fairdata.solutions/DAV/home/LDP/gofair/> .
```

```
@prefix ldp:     <http://www.w3.org/ns/ldp#> .
```

```
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
```

```
ns1:SKOS
```

```
    rdf:type      ldp:BasicContainer , ldp:Container , rdfs:Resource , ldp:Resource .
```

```
<http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/>
```

```
    rdf:type      ldp:BasicContainer , ldp:Container .
```



**Advanced options**

Select the resulting RDF serialization Turtle ▾

Automatically generate SKOS-XL for labels

Get result in a zip file

Generate graph files for Openlink Virtuoso



The file is downloaded as **Pest\_Scheme.ttl** (RDF turtle format)

**\$ cat Pest\_Scheme.ttl**

Now we have to  
upload the Scheme  
to this location

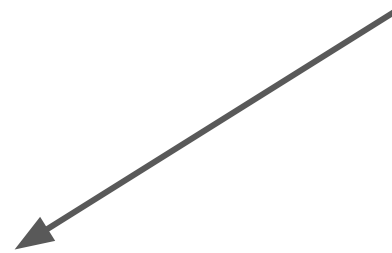
```
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

```
<http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl> a skos:ConceptScheme;  
dct:description "FAIR Data representation of a Pest observation dataset"@en;  
dct:title "Pest Observations Dataset"@en;  
skos:hasTopConcept <http://opendata.aragon.es/def/ei2a#Crop>, <http://purl.obolibrary.org/obo/NCIT\_C16270>,  
  <http://semanticscience.org/resource/SIO\_010414>, <http://www.disit.org/km4city/schema#AgricultureAndLivestock>,  
  <http://www.ontotext.com/proton/protonext#Food> .
```

```
<http://purl.obolibrary.org/obo/NCIT\_C16270> a skos:Concept;  
skos:definition "The practice of cultivating the land or raising stock, as well as the work, business, or study of farming"@en;  
skos:inScheme <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl>;  
skos:prefLabel "Agriculture"@en;  
skos:topConceptOf <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl> .
```

```
<http://semanticscience.org/resource/SIO\_010414> a skos:Concept;  
skos:definition "A pest or pathogen"@en;  
skos:inScheme <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl>;  
skos:prefLabel "pathogen"@en;  
skos:topConceptOf <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl> .
```

...  
...  
...



```
curl -k -i -X POST "http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/"
-u gofair:gofair
-H "Slug: Pest_Scheme.ttl"
--data-binary @Pest_Scheme.ttl -H "Content-Type: text/turtle"
```

HTTP/1.1 100 Continue

Date: Fri, 08 Mar 2019 11:14:15 GMT

HTTP/1.1 201 Created

Date: Fri, 08 Mar 2019 11:14:15 GMT

Server: Virtuoso/07.20.3230 (Linux) x86\_64-generic\_glibc25-linux-gnu

X-Frame-Options: SAMEORIGIN

Accept-Ranges: bytes

Location: http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl

Allow:

GET,HEAD,POST,PUT,DELETE,OPTIONS,PROPFIND,PROPPATCH,COPY,MOVE,MKCOL,LOCK,UNLOCK,TRACE,PATCH

Vary: Accept,Origin,If-Modified-Since,If-None-Match

MS-Author-Via: DAV, SPARQL

Accept-Patch: application/sparql-update

Accept-Post: text/turtle,text/n3,text/nt,text/html,application/ld+json

Link: <http://www.w3.org/ns/ldp#Resource>; rel="type"

Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl,meta>; rel="meta"; title="Metadata File"

Link: <http://training.fairdata.solutions/DAV/home/LDP/gofair/SKOS/Pest\_Scheme.ttl,acl>; rel="acl"; title="Access Control File"

ETag: "4826edeb1474d9c4397deebae880daff"

Content-Type: text/turtle

Content-Length: 199

**We can browse to it...**



# DONE!

**Nota bene: I just showed you the **most difficult** way to interact with an LDP server**

(i.e. raw HTTP messages built at the command-line)

There are, of course, LDP client libraries in every language that will do the hard work for you!

The point was to show you that there is no “black magic” here. FAIR invented nothing! We use only 100% standard Web messaging to achieve the FAIR goals.

# My own LDP Client (in Ruby) that you saw earlier

```
require 'ldp_simple'

dc = RDF::Vocabulary.new("http://purl.org/dc/elements/1.1/")
client = LDP::LDPClient.new({
  :endpoint => "http://training.fairdata.solutions/DAV/home/LDP/gofair/",
  :username => "gofair",
  :password => "gofair"})

top = client.toplevel_container

  # add metadata to our container

top.add_metadata([

  ["http://training.fairdata.solutions/DAV/home/LDP/gofair/",
  dc.author,
  "Mark Wilkinson"],

  ["http://training.fairdata.solutions/DAV/home/LDP/gofair/",
  dc.title,
  "A slide of insect pest grazing observations for the EU"]

])
```

# Once metadata is added to the container, we can SPARQL it

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX course: <http://training.fairdata.solutions/DAV/home/LDP/gofair/>

select ?title ?author where {
  course: dc:title ?title .
  course: dc:creator ?author .
}
```

<b>title</b>	<b>author</b>
"A slide of insect pest grazing observations for the EU"@en	"Mark Wilkinson"@en

Keep adding metadata until...

Note that this Metadata is about ME! I am the creator of this dataset, and may be credited for it.

UniProt Slice FAIR Accessor -  
Aspergillus RNA Processing  
proteins

creator	wilkinsonlab.info/
language	eng
license	cc by nd4.0
title	UniProt Slice FAIR Accessor - Aspergillus RNA Processing proteins
authored By	0000 0002 9699 485X
entities	412
term has Principal Investigator	Dr. Mark Wilkinson
type	Dataset Basic Container Collection
contact Point description	Wilkinson.rdf Takes a SPARQL query of the UniProt database specific to proteins and their GO annotations related to RNA Procressing proteins in Aspergillus and makes it a FAIR Accessor source. The precise query is:



```
PREFIX up:<http://purl.uniprot.org/core/>
PREFIX taxon:<http://purl.uniprot.org/taxonomy/>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?id
```

```
WHERE
```

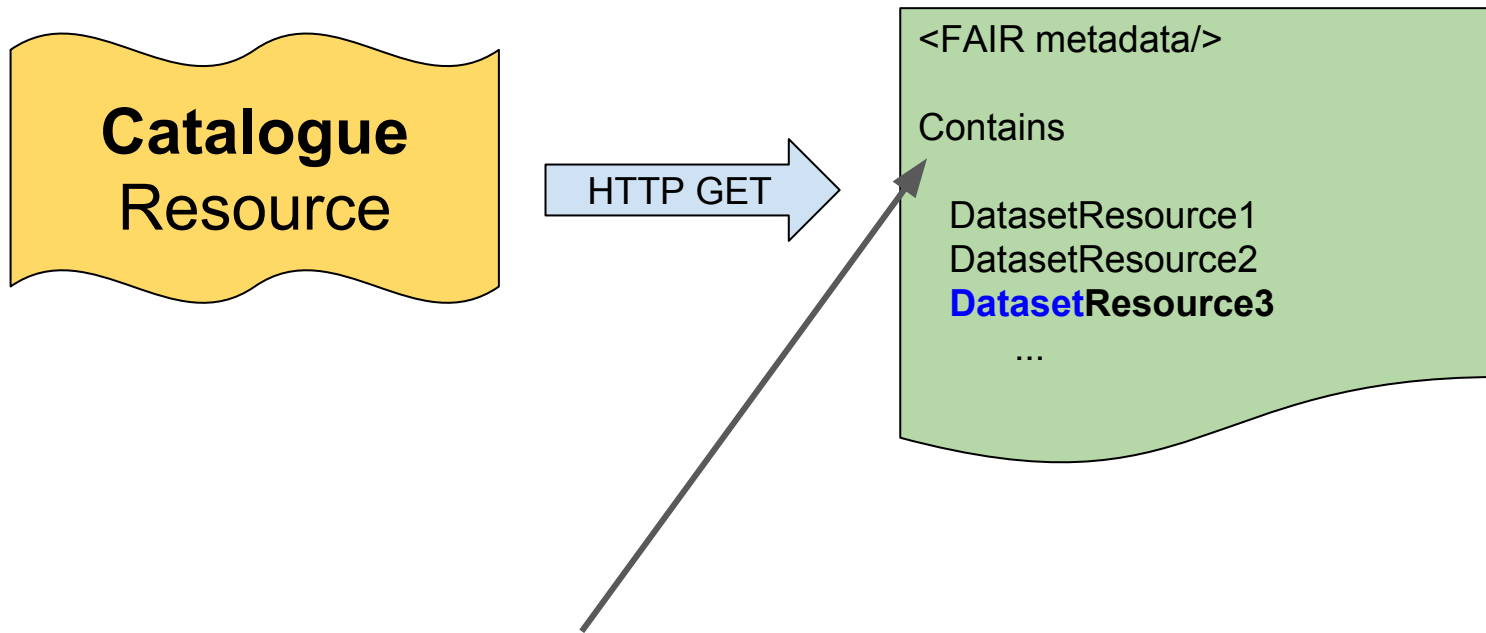
```
{
  ?protein a up:Protein .
  ?protein up:organism ?organism .
  ?organism rdfs:subClassOf taxon:162425 .
  ?protein up:classifiedWith ?go .
  ?go rdfs:subClassOf* <http://purl.obolibrary.org/obo/GO_0006396> .

  bind(replace(str(?protein), "http://purl.uniprot.org/uniprot/", "", "i") as ?id)
}
```

identifier

Uni Prot Accessor

# What does a FAIR Accessor “look like”?



? did we do this?  
? when ?

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX course: <http://training.fairdata.solutions/DAV/home/LDP/gofair/>
PREFIX ldp: <http://www.w3.org/ns/ldp#>
```

```
select ?title ?author ?record where {
course: dc:title ?title .
course: dc:creator ?author .
course: ldp:contains ?record
}
```

title	author	record
"A slide of insect pest grazing observations for the EU"@en	"Mark Wilkinson"@en	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/species_290307346">http://training.fairdata.solutions/DAV/home/LDP/gofair/species_290307346</a>
"A slide of insect pest grazing observations for the EU"@en	"Mark Wilkinson"@en	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365908">http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365908</a>
"A slide of insect pest grazing observations for the EU"@en	"Mark Wilkinson"@en	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365909">http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365909</a>
"A slide of insect pest grazing observations for the EU"@en	"Mark Wilkinson"@en	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365910">http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365910</a>
"A slide of insect pest grazing observations for the EU"@en	"Mark Wilkinson"@en	<a href="http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365911">http://training.fairdata.solutions/DAV/home/LDP/gofair/obs_2147365911</a>

That triple was created automatically by the LDP Server when we POSTed the data records into the container URL

# The FAIR Data Point Editor will create a (simple) FAIR Accessor for you

Editor

Repository

Catalog Dataset Distribution

Show optional fields

Title

Has version

Publisher

Publisher Name

Theme taxonomy

Build Share

Select a field to read more about it.

RDF preview

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix dct: <http://purl.org/dc/terms/>.
@prefix dcat: <http://www.w3.org/ns/dcat#>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix fdp: <http://rdf.biosemantics.org/ontologies/fdp-o#>.
@prefix datacite: <http://purl.org/spar/datacite/>.

<> rdf:type dcat:Catalog.
```



**Enter Luiz!**

PREFIX dc: <<http://purl.org/dc/elements/1.1/>>  
PREFIX course: <<http://training.fairdata.solutions/DAV/home/LDP/gofair/>>  
PREFIX ldp: <<http://www.w3.org/ns/ldp#>>  
PREFIX efo: <<http://www.ebi.ac.uk/efo/efo.owl#>>  
PREFIX sio: <<http://semanticscience.org/resource/>>

```
select ?conceptscheme ?sub ?pred ?obj where {  
  graph course: {  
    ?s dc:title ?title .  
    ?s dc:creator ?author .  
    ?s dc:license ?license .  
    ?s ldp:contains ?container .  
  }  
  graph ?container {  
    ?record a ldp:Container .  
    ?record ldp:contains ?skos .  
  }  
  
  graph ?skos {  
    ?conceptscheme a <http://www.w3.org/2004/02/skos/core#ConceptScheme>  
  }  
  graph ?conceptscheme {  
    ?sub ?pred ?obj  
  }  
}
```