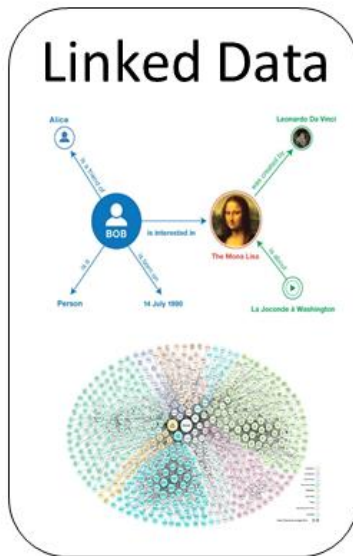


+



=



Active working group members:

James Biard (NCEI/NOAA)

Jonathan Yu (CSIRO, Australia)

Mark Hedley (UK Met Office)

Adam Leadbetter (Marine Institute, Ireland)

Acknowledgements

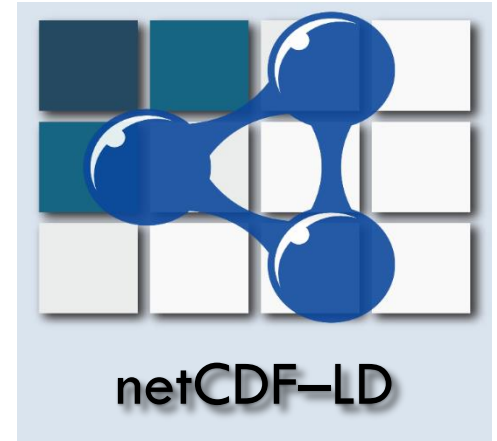
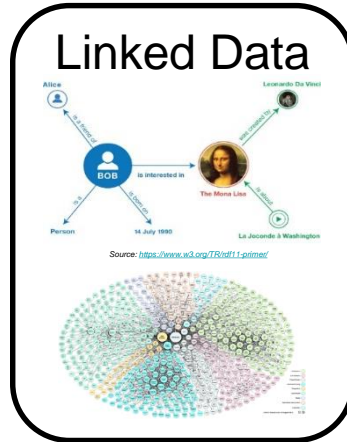
Nick Car & Alex Ip (Geoscience Australia)

Kelsey Druken (NCI Australia)

Sean Arms (UCAR)

[Early contributors]

Why Linked Data for netCDF?



+

=

Great uptake & tools.

Many conventions.

Lots of data in
netCDF!

Way to interlink, discover,
and integrate data on the
web.

Rapidly growing galaxy of
information spanning many
disciplines.

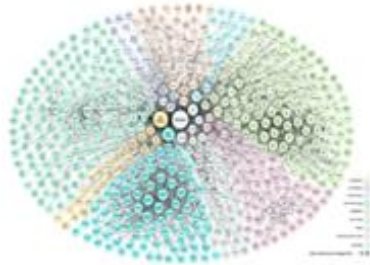
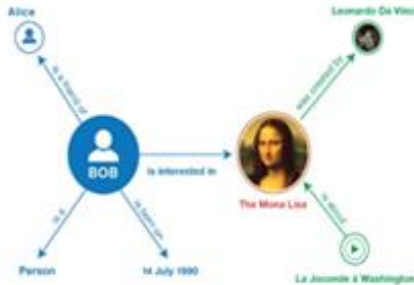
Linked Data profiles –
JSON-LD, CSVW.

Recipe for constructing
Linked Data from netCDF
files.

Define conventions and
validate data.

Enable linking to other
resources.

Linked Data



Recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge ***on the web.***

Standard format ...

Reachable ...

Relationships between data ...

Collection of interrelated data → Linked Data

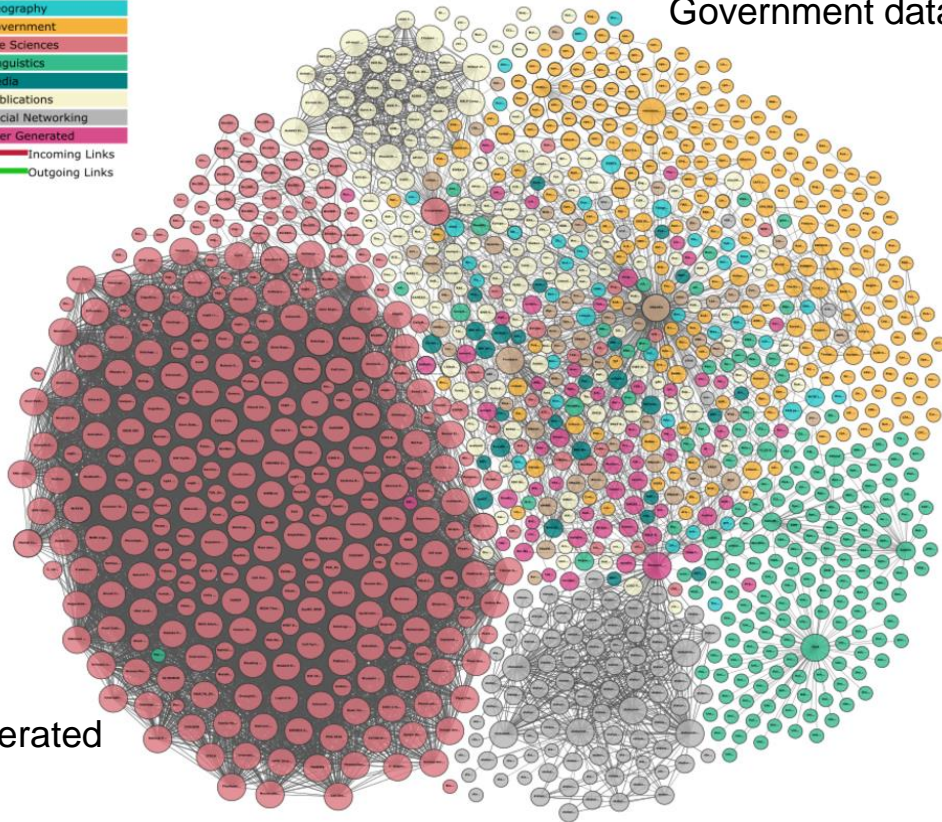
Key concept: Give each *thing* in the data an individual identity or URI

Last updated: 2017-08-22



Publications

Government data



Geo

User generated

32 billion triples in 2014

...

192 billion+ triples in 2017

See <http://stats.lod2.eu/stats>

Can we plug netCDF/HDF data in?

Motivations

1. Interoperability

Handle multiple metadata standards/conventions
e.g. combining CF with project conventions

2. Reuse and discovery

Leverage web resources and Linked Data tools to ***enhance discovery*** across large collections of files (e.g. represent separate files as graphs)

Journey so far

2014-2015

Conceived during eReefs in discussions with Ethan Davis and Adam Leadbetter



[DOI: 10.1007/978-3-319-15994-2_9](https://doi.org/10.1007/978-3-319-15994-2_9)

Collaboration with Mark Hedley (UK Met.)

EarthCube
Advancing CF for
geosciences workshop
(Boulder, CO), May 2016

WG proposal
endorsed by CF
community

Drafting OGC Standard
netCDF-Classic-LD

OGC
TC NZ

EC CF
meeting
Sept 2017



Fall meeting
Poster
Dec 2016



netCDF-LD
WG and
telecons
(Aug 2016)

EU: UK Met office,
Marine Inst. Ireland
AU: GA, NCI
US: NOAA, NCICS
(notably Jim Biard)

Piloting encoding rules,
tools and tests on GitHub



Design principles

1. Work with current netCDF files

Design a simple mechanism to that works with existing netCDF files as-is to encode in a Linked Data friendly format.

2. Allow consistent & precise naming of each *thing* in netCDF/HDF metadata

Implies introducing new syntax (compatible with netCDF / HDF) to build URIs for each attribute name and property value

Enable consistent way to link to references, e.g. model, instrument, etc.

3. Provide useful Linked Data enabled outputs (RDF, schema.org)

Overview

Tools

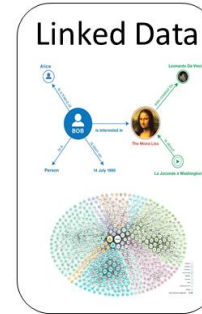
Syntax (aliasing, prefixes)

Supporting registries

Next steps



+



=



Tools



Python libraries (Github bald repo) - (bald = binary array linked data)

<https://github.com/binary-array-ld/bald>



Command line tools (in development):

nclddump

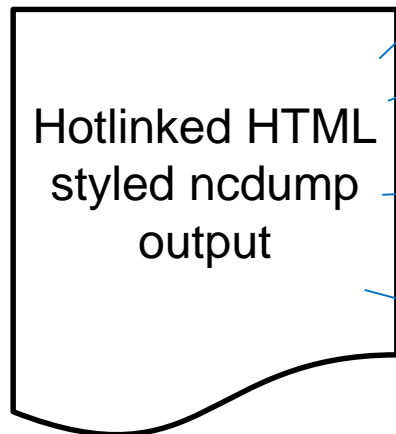
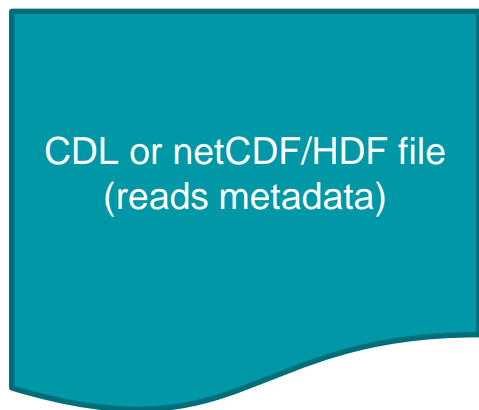


nc2rdf

Demos

ncldump

```
$ python ncldump.py example.cdl
```

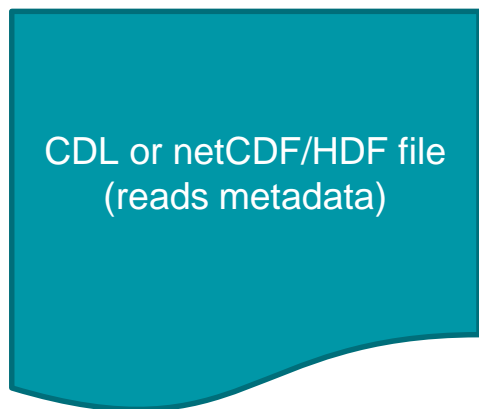


Information on
web pages and registries

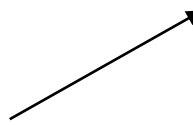
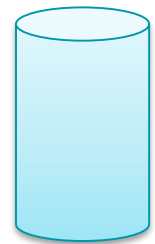
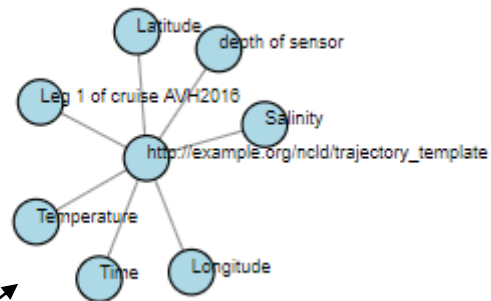
Example on <https://binary-array-ld.github.io/netcdf-ld/#examples>

nc2rdf

```
$ python nc2rdf.py example.cdl
```



(visualisation representation)



Load into
triple store DB
for semantic queries

Binary Array LD Syntax (for netCDF and HDF)

Methods to encode or process nc/hdf for translating to RDF / Linked Data ready

Aliasing

Lookup table for 'well-known' or declared mappings

Can be explicit or implicit

Pros: Easy to convert current nc files

Cons: Resolving clashes

e.g. `title` → `acdd:title`

(netcdf)

(RDF)

Prefixing

Kinda like namespacing

Pros: Easy to convert conformant files

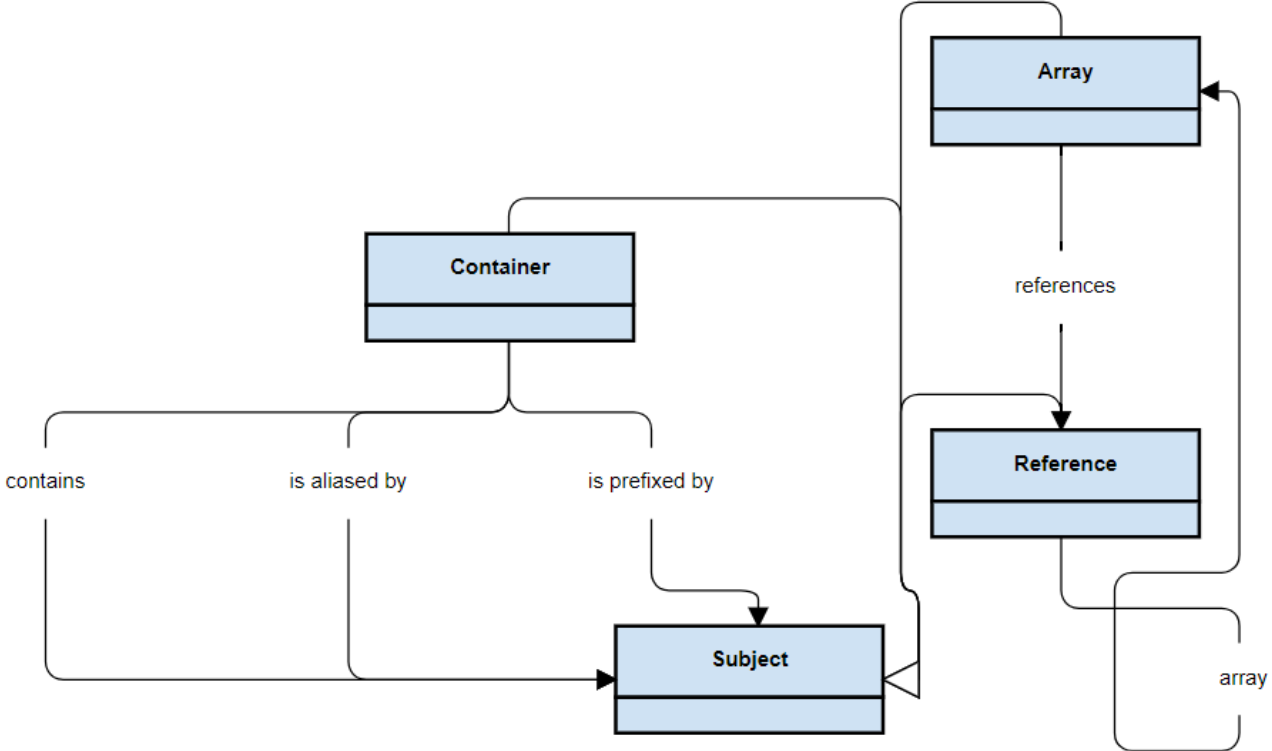
Cons: Current files need tweaking

e.g. `acdd__title` → `acdd:title`

(netcdf)

(RDF)

Binary Array Linked Data (BALD) model



<http://binary-array-id.net/latest?classView=true>

Aliasing example

variables:

```
int variable(pdimm0, pdimm1) ;  
    variable:SDN_ParameterDiscoveryCode = "BactTaxaAbundSed" ;
```

```
int cfvariable(pdimm0, pdimm1) ;  
    cfvariable:standard_name = "air_temperature" ;
```

```
// global attributes:  
    :isAliasedBy = "alias_list" ;
```

}
Variable metadata
}

Aliasing example – adding context (explicit aliases)

variables:

```
int alias_list ;  
    alias_list:SDN_ParameterDiscoveryCode =  
"http://vocab.nerc.ac.uk/isoCodelists/sdnCodelists/cdicrCodeList.xml#SDN\_ParameterDiscoveryCode" ;  
    alias_list:BactTaxaAbundSed = "http://vocab.nerc.ac.uk/collection/P02/current/BAUC/" ;  
    alias_list:standard_name = "https://def.scitools.org.uk/CFTerms/standard\_name" ;  
    alias_list:air_temperature = "http://vocab.nerc.ac.uk/collection/P07/current/CFSN0023/" ;  
  
int variable(pdimm0, pdimm1) ;  
    variable:SDN_ParameterDiscoveryCode = "BactTaxaAbundSed" ;  
  
int cfvariable(pdimm0, pdimm1) ;  
    cfvariable:standard_name = "air_temperature" ;  
  
// global attributes:  
    :isAliasedBy = "alias_list" ;
```

Mappings

Variable metadata

Aliasing example – RDF representation

```
<example> a bald:Container ;  
  bald:contains <variable>, <cfvariable> .
```

...

```
<variable> a bald:Array ;
```

```
  ns1:SDN_ParameterDiscoveryCode
```

```
    <http://vocab.nerc.ac.uk/collection/P02/current/BAUC/> ;
```

```
<cfvariable> a bald:Array ;
```

```
  ns2:standard_name
```

```
    <http://vocab.nerc.ac.uk/collection/P07/current/CFSN0023/> .
```

Prefix example – ereefs running example

variables:

```
float eta(time, j, i) ;  
  eta:units = "metre" ;  
  eta:long_name = "Surface elevation" ;  
  eta:standard_name = "sea_surface_height_above_sea_level" ;  
  eta:medium_id = "ocean"  
  eta:scaledQuantityKind_id = "sea_surface_elevation"  
  eta:substanceOrTaxon_id = "ocean_near_surface"
```

Variable metadata

Prefix example – ereefs with prefixes added

variables:

```
float eta(time, j, i) ;  
    eta:units = "metre" ;  
    eta:cf__long_name = "Surface elevation" ;  
    eta:cf__standard_name = "cfsn__sea_surface_height_above_sea_level" ;  
    eta:ereefs_medium_id = "feature__ocean"  
    eta:ereefs_scaledQuantityKind_id = "property__sea_surface_elevation"  
    eta:ereefs_substanceOrTaxon_id = "feature__ocean_near_surface"
```

Variable metadata

Prefix example – added prefix mappings

variables:

```
int prefix_list;
```

```
    prefix_list:cf__ = https://def.scitools.org.uk/CFTerms/
```

```
    prefix_list:cfsn__ = http://mmisw.org/ont/cf/parameter/
```

```
    prefix_list:feature__ = "http://environment.data.gov.au/def/feature/" ;
```

```
    prefix_list:property__ = "http://environment.data.gov.au/def/property/" ;
```

```
    prefix_list:ereefs__ = "http://registry.it.csiro.au/sandbox/ncl/ereefs-attributes/" ;
```

Prefix Mappings

```
float eta(time, j, i) ;
```

```
    eta:units = "metre" ;
```

```
    eta:cf\_long\_name = "Surface elevation" ;
```

```
    eta:cf\_standard\_name = "cfsn\_sea\_surface\_height\_above\_sea\_level" ;
```

```
    eta:ereefs\_medium\_id = "feature\_ocean"
```

```
    eta:ereefs\_scaledQuantityKind\_id = "property\_sea\_surface\_elevation"
```

```
    eta:ereefs\_substanceOrTaxon\_id = "feature\_ocean\_near\_surface"
```

Variable metadata

Supporting registries

SciTools Registry

Browse

About

Admin ▾

Advanced ▾

Search

Submit

Jonat

List all registers




Filters

Category

SciTools ✕

Owner

Entity

Name	Notation	Description	Status
 ACDD	ACDD	Vocabulary of terms used in the Attribute Conventions Dataset...	experimental
 CFTerms	CFTerms	Vocabulary of terms used in the CF conventions for netCDF fil...	experimental
 net cDF	NetCDF	Vocabulary of terms used in the netCDF User Guide.	experimental

Developed by [Epimorphics Ltd](#)

<https://def.scitools.org.uk/>

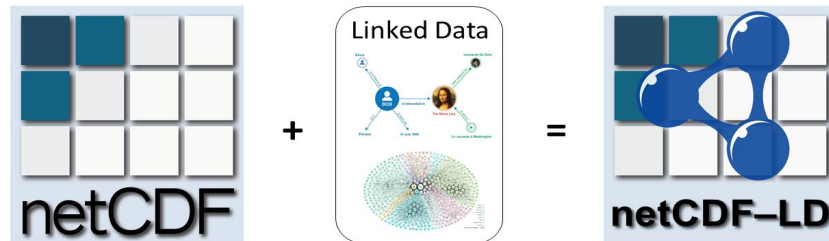
Next steps

Push netCDF standardisation via OGC forward

Looking for collaborators to work on demonstrators

- Interoperability experiments support enhanced discovery of netCDF data
- Link controlled vocabularies with netCDF data
- Want to contribute?

Thanks



Active working group members:

James Biard (NCEI/NOAA)

Jonathan Yu (CSIRO, Australia)

Mark Hedley (UK Met Office)

Adam Leadbetter (Marine Institute, Ireland)



GitHub

Python libraries (bald = binary array linked data)

<https://github.com/binary-array-ld/bald>

<https://github.com/opengeospatial/netCDF-Classic-LD>



<http://tinyurl.com/netcdf-ld>

Demo

<http://waterinformatics-ext1-cdc.it.csiro.au/ncl-d-demo/>