



TW/C

*RDA Plenary 5, San Diego, CA, March 08, 2015*

# Adoption of RDA DTR and PID in Deep Carbon Observatory Data Portal

Stephan Zednik, Xiaogang Ma,  
John Erickson, Patrick West, Peter Fox,  
& DCO-Data Science Team

Tetherless World Constellation  
Rensselaer Polytechnic Institute





# Outline

- Background & Research questions
  - RDA-DTR, RDA-PIT, DCO Data Portal
- Nature of efforts
  - Basic data type vs. Specific data type
- Approaches
  - Integration architecture vs. Self-contained architecture
- Initial results
- Future works



# Background

- RDA - Data Type Registry (**DTR**) working group
  - Addressed a core issue of data interoperability: to parse, understand, and reuse data retrieved from others
- RDA - Persistent Identifier Information Types (**PIT**) working group
  - Addressed the essential types of information associated with persistent identifiers (PID)
- Deep Carbon Observatory (**DCO**) Data Portal
  - Centrally-managed digital object identification, object registration, metadata management and knowledge graph curation.
  - <http://deepcarbon.net>





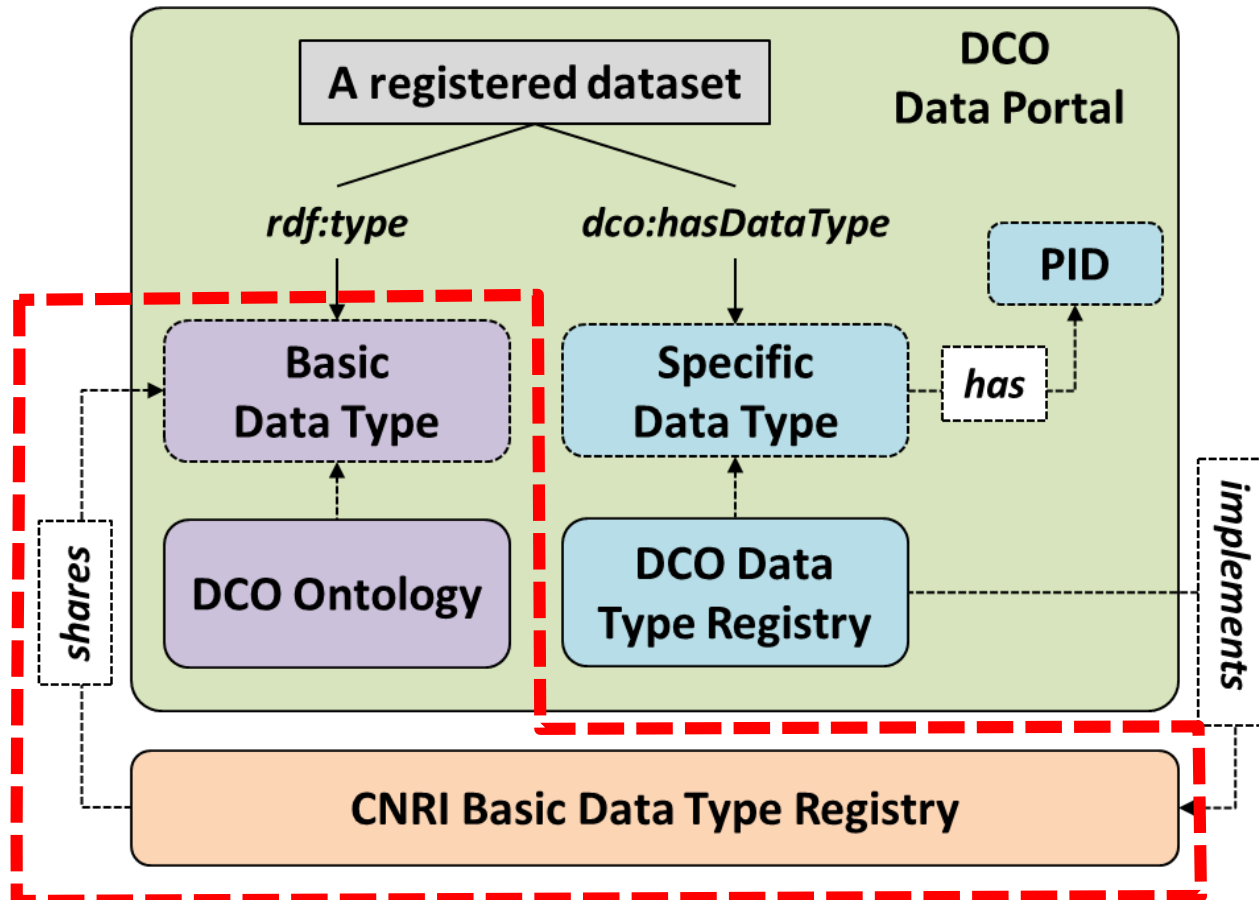
# Research Questions

- Each defined data type needs a stable and resolvable PID
- Provide semantics - meaning and context - to the defined data types
- Annotate datasets with one or more defined data types



# Nature of efforts

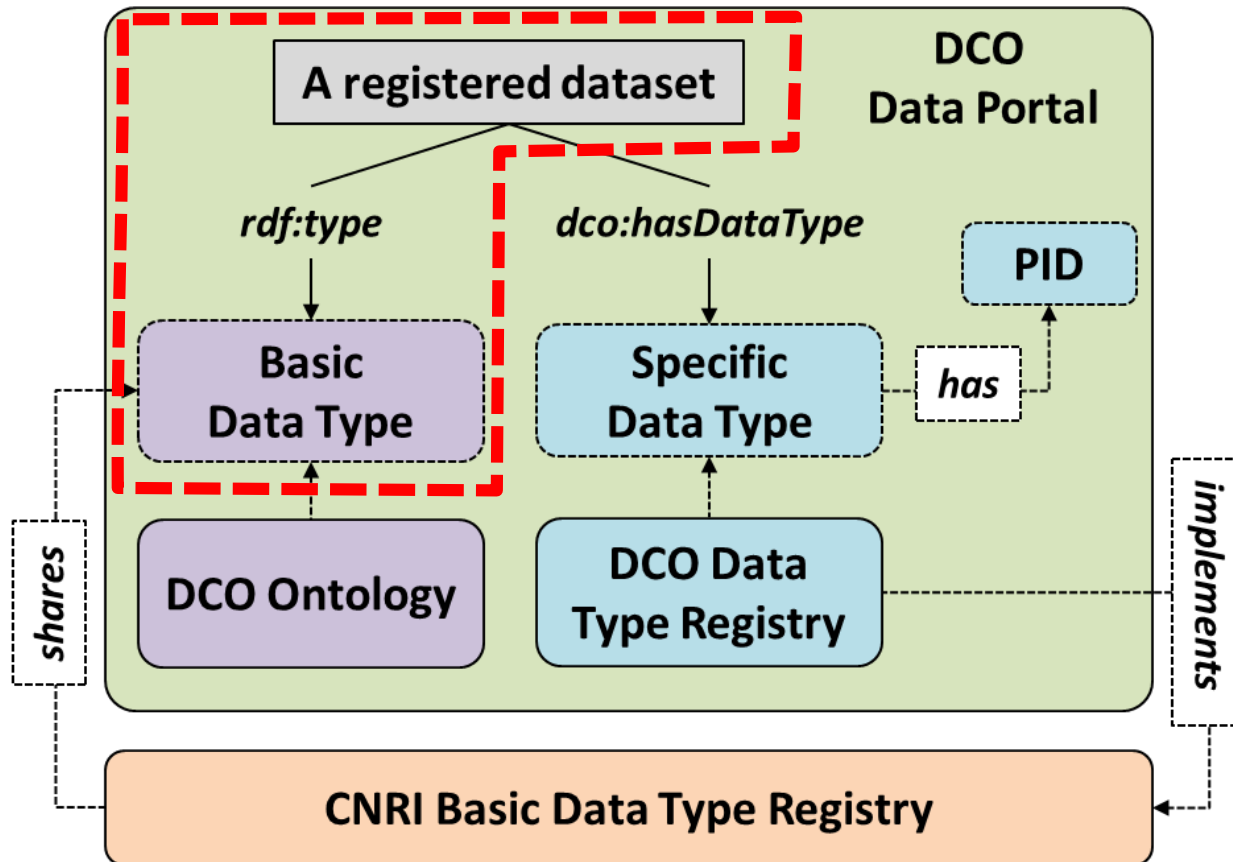
The DTR primitives are comparable to a list of **BASIC DATA TYPE CLASSES** in the DCO ontology, e.g. Dataset, Image, Video, Audio, etc.





# Nature of efforts (cont.)

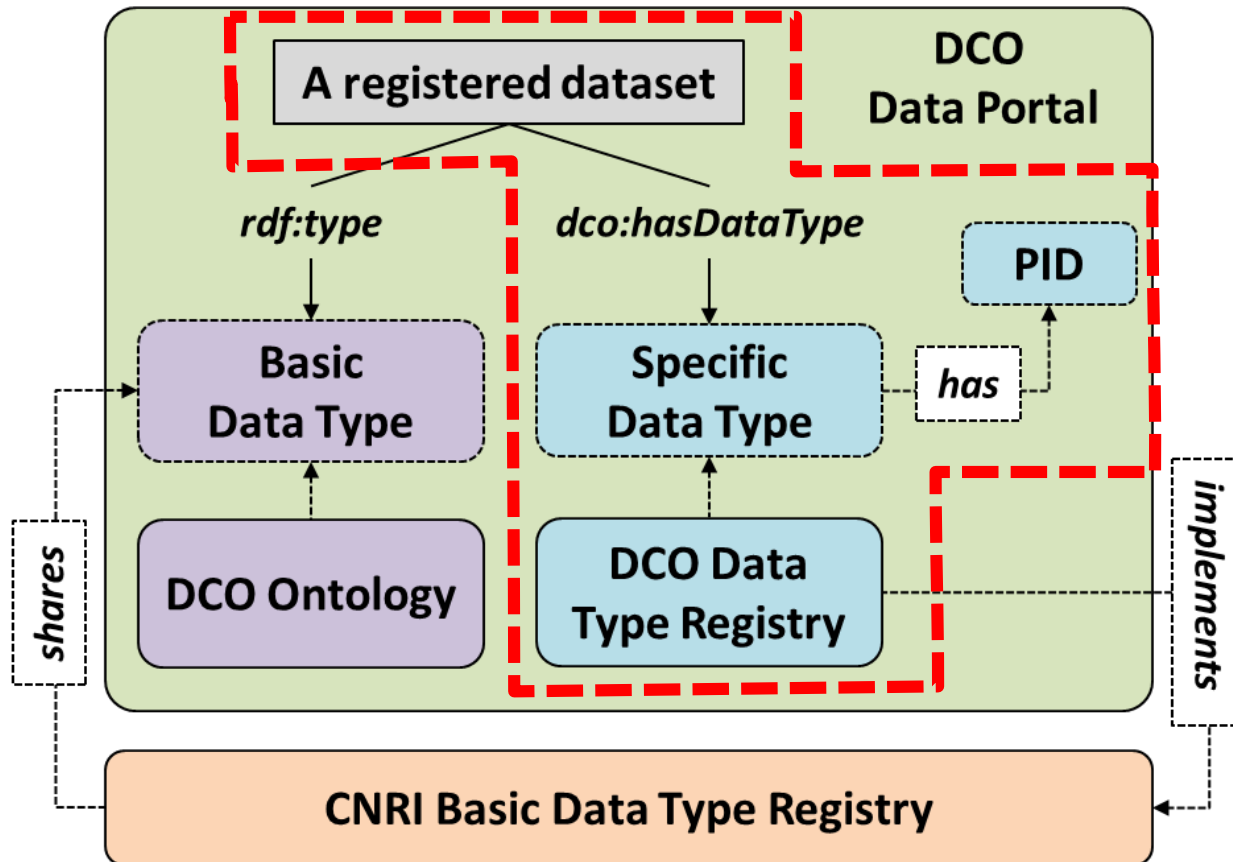
A registered DCO dataset is asserted as an instance of one of those basic data type classes.





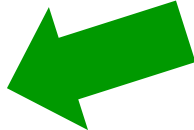
# Nature of efforts (cont.)

It is possible to further annotate the dataset with the **SPECIFIC DATA TYPES** defined within a DTR, and each data type has a unique PID.





# Possible DCO-DTR Approaches

- An integration architecture
  - DCO Data Portal is built on the VIVO platform
  - DTR and DCO-VIVO as separate knowledge bases
  - DCO-VIVO uses DTR API to access data type information
- A self-contained architecture  *We are currently working on this*
  - To have the functionality of DTR completely within the DCO Data Portal
  - Need to modify the DCO Ontology, e.g. add a class `dco:DataType` and collect properties associated with it

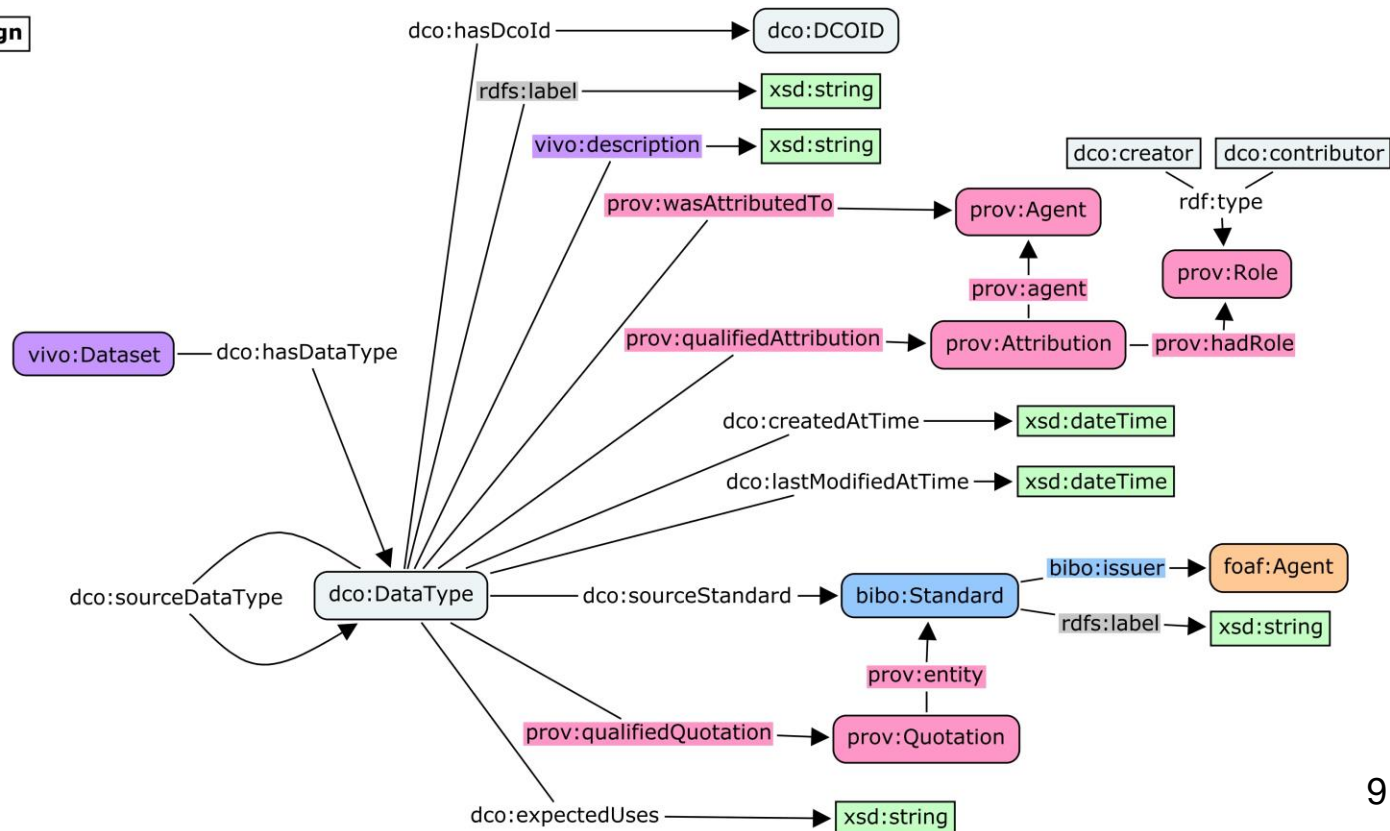
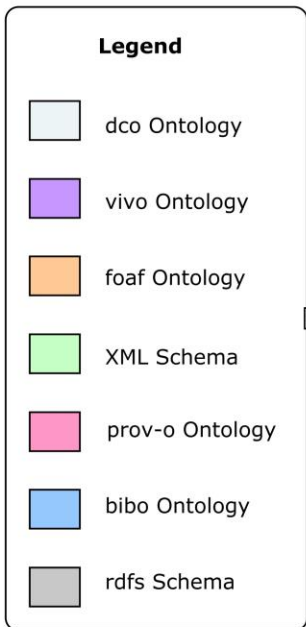




# Initial results

- Updates to the DCO Ontology:
  - A new class `dco:DataType`. Each specific data type is an instance of it
  - An object property `dco:hasDataType` linking a dataset and a data type
  - A collection of other classes and properties associated with `dco:DataType`

## DCO-DTR: A Schematic Design





# An example



Admin Panel [Edit this individual](#)

Verbose property display is **off** | [Turn on](#)

Resource URI: <http://info.deepcarbon.net/individual/n5989>

DCO Ontology [Dataset](#)

DCO ID [11121/1490-1829-4848-8082-CC](#)

Websites

The basic data type

```
dco:dcoOntology
rdf:type vivo:Dataset .
```

Overview

Research

Contact

Related Documents

Time

Identity

Additional Document Info

Bib Mapping

Other

View All

Has data type [+](#)

[Resource Description Framework](#)


keywords [+](#)

The specific data type

```
dco:dcoOntology
dco:hasDataType dco:RDF .
```



# An example (cont.)

 **Admin Panel** [Edit this individual](#) Verbose property display is **off** | [Turn on](#)  
Resource URI: <http://info.deepcarbon.net/individual/n47205>

Resource Description Framework | **Data Type**

**DCO ID** [11121/1305-5365-5550-2605-CC](#)

Research Identity Other View All

Has source standard +

[W3C - Resource Description Framework \(RDF\)](#)

source data type +

- [JSON-LD](#)
- [N-Quads](#)
- [N-Triples](#)
- [N3](#)

```
dco:RDF
  a dco:DataType .
dco:DataType
  a owl:Class .
```

Each registered object, such as a data type, has a unique DCO ID, which is resolvable by the global Handle System



# Future works

- More **use case analyses** relevant to data types in the DCO community
- **Refine the schema** for the annotation and provenance of specific data types
- A faceted '**data type browser**' on the DCO Data Portal
- **Interoperability** between DCO specific data types and data types registered in other DTRs
  - The integration architecture between DCO-VIVO and DTR

*Thank you!*





# TWOC

- Backup slides



# Update the DCO Dataset Browser

1-10 of 101 < > 10 ▾

**Years**

2015 (4)  
2014 (20)  
2013 (1)  
2005 (1)

**DCO Authors**

Basharina, L A (4) ⓘ  
Borisov, O G (1) ⓘ  
Chenicheri, Lakshmi (1) ⓘ  
Dmitriev, L B (1) ⓘ  
Erickson, John (2) ⓘ  
Etiopo, Giuseppe (1) ⓘ  
Fox, Peter (1) ⓘ  
Gnanasekar, Sumithra (1) ⓘ  
Gorelchik, V I (1) ⓘ

**Other Authors**

**Data type**  
-XXXXXXXX  
---XXX  
----XXXXXXXXX  
----XXXXXXXXX  
-XXXXXXXX  
---XXXXXX  
-----XXXXXXXXXX

**A. Di Muro 1**  
DCO ID: 11121/7240-4014-1206-8279-CC  
Project: DECADE - CO2 Soil Degassing at Piton de la Fournaise  
Distributions:  
2014-08-18\_17:22:16 (Direct access: A. Di Muro 2 Most recent ERUPTION FOURNAISE 21 june 2014.jpg)

**A. Di Muro 2**  
DCO ID: 11121/7051-5194-5749-5165-CC  
Project: DECADE - CO2 Soil Degassing at Piton de la Fournaise  
Distributions:  
2014-08-18\_17:25:19 (Direct access: A. Di Muro satellite image Reunion Island.jpg)

**Abiotic methane in serpentinites sampled in Italy, Greece, Portugal and Turkey**  
DCO ID: 11121/1379-5903-3822-5833-CC  
Communities: Deep Energy Community  
Authors: Etiopo, Giuseppe  
Distributions:  
2014-05-15T17:10 (Direct access: default-file-name)

**Active Crater of Mutnovsky Volcano (1960-1964, Kamchatka, Russia)**  
DCO ID: 11121/4251-3184-7546-9088-CC  
Communities: Reservoirs and Fluxes Community  
Authors: Pronin, A A; Kirsanov, I T; Vakin, E A

**Active Fumaroles of the Piton de la Fournaise (Reunion Island)**  
DCO ID: 11121/9...  
Communities: Res...  
Authors: Kirsanov...

**An investigation of the CO2 degassing from the Piton de la Fournaise**  
DCO ID: 11121/6013-5737-8927-7988-CC  
Communities: Reservoirs and Fluxes Community  
Authors: Ivanov, I Z

**Background and Strategy for the Data Management of the Deep Carbon Observatory**  
DCO ID: 11121/8597-5508-3283-9133-CC  
Project: Deep Carbon Observatory Data Management  
Distributions:  
2014-08-11\_12:58:52 (Direct access: DCO\_DMP\_v5.pdf)

**Data type as a facet in dataset searching and browsing**