



5th RDA Collaboration Meeting Nottingham, UK  
8th- 9th June 2016

Data Foundations & Terminology (DFT) IG

<https://rd-alliance.org/groups/data-foundations-and-terminology-ig.html>

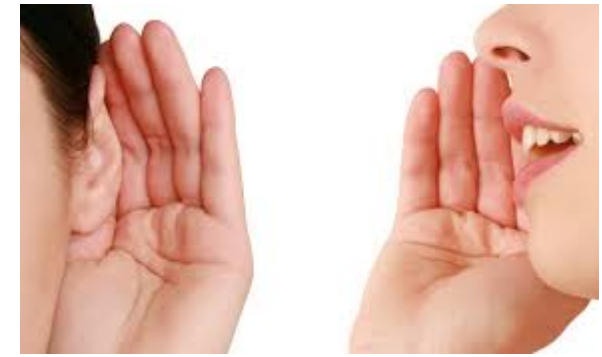
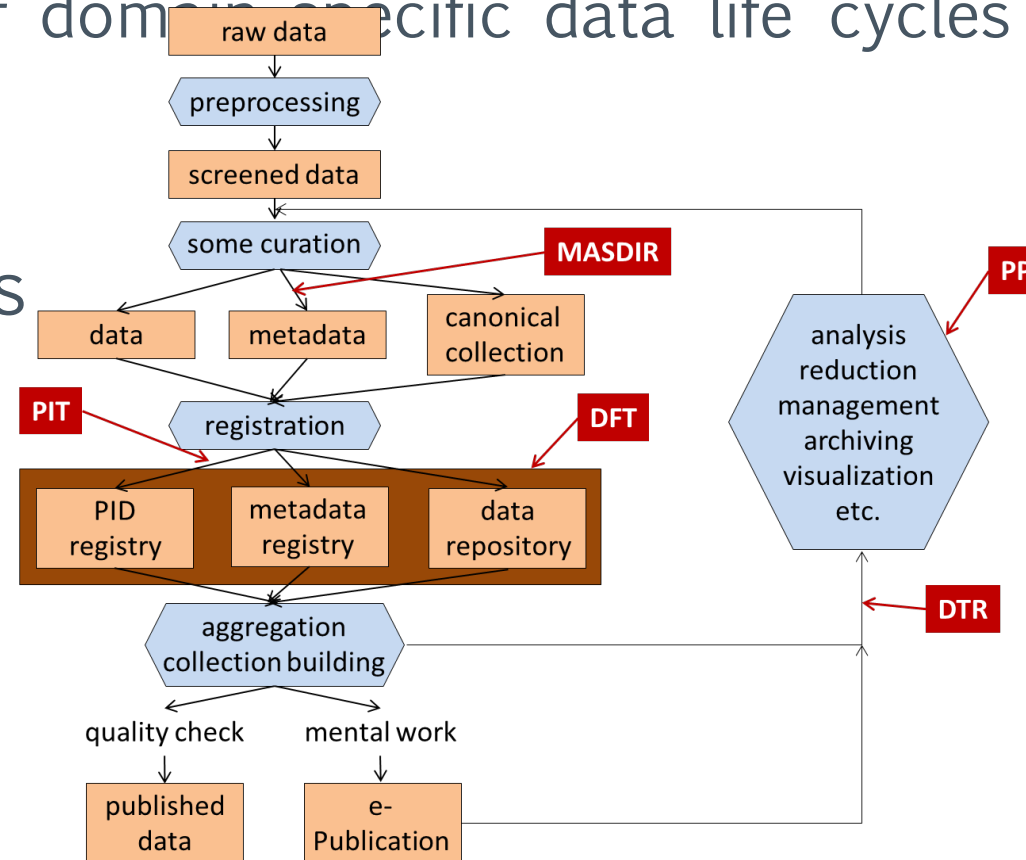
Gary Berg-Cross (Ontolog) & Raphael Ritz:  
Data & Domain Vocabularies and Vocabulary Services:  
Beyond data pilots

**research data sharing without barriers**  
**rd-alliance.org**

# AGENDA

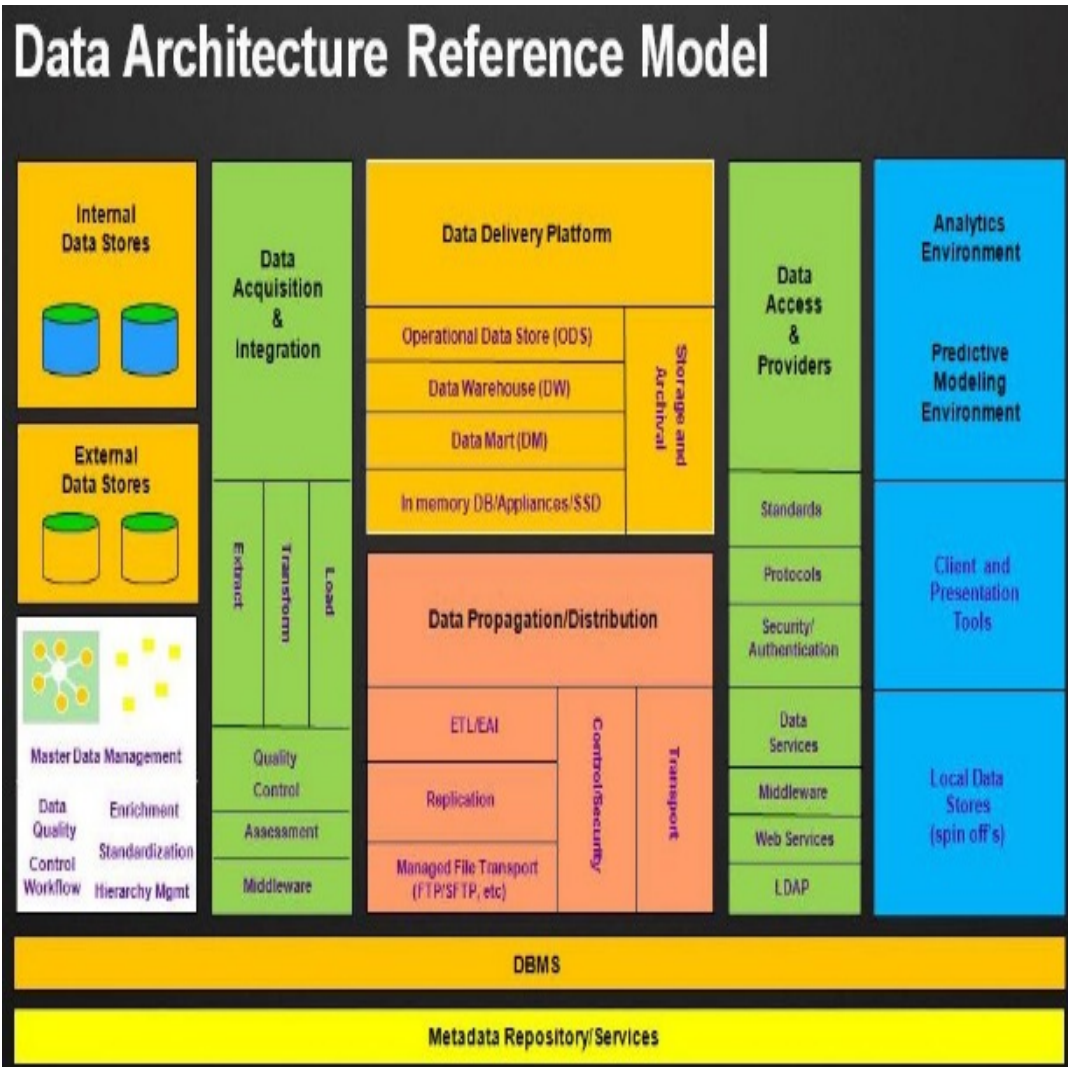
$\pi$

- › Where we are coming from
- › New directions and collaborations
  - Context of domain-specific data life cycles
  - VSIG
- › Some ideas



**Cross-group coordinated** with several RDA WGs, as shown in the Data Fabric Figure on data concepts and relations. Potentially all groups could be engaged in this IG and we with them

# Original Scope: Data Management & Infrastructure: huge range of concepts and terms describing “CyberInfrastructure” & relationships



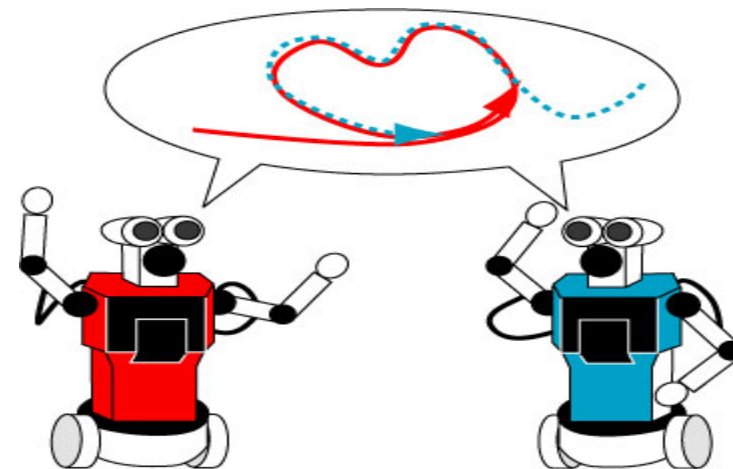
Shared Conceptualizations  
Communicated by Vocabularies are  
Needed for Understanding

## Challenges:

Knowledge is a socially distributed & fluid phenomena

Heterogeneous categorizations and varied methods for organizing information

Lack of agreed upon standard metadata & definitions.



# Where we are coming from: Original Objectives for IG

1. Leverage existing DFT WG products, tools and approach but improve these:
  - Increase in Terms & Initial definitions in TeD-T
  - new requirements coming out of groups just completed or nearing completion,
  - support as part of adoption.
2. Focus on facilitating community discussion on core concepts
3. Update current synthesis model and stabilize the effort for subsequent use.
4. Facilitate definition development
  - Potential adopters feedback
  - additional use case scenarios

Main page  
Add Term  
Add Category  
RDF Export of terms  
Browse Term Collection  
All Terms - Hierarchical  
All Terms - List  
List by scope  
Recent populated terms  
Ted-T Graph  
Help  
Tutorial  
Tools  
Upload file  
Special pages  
Printable version

All pages  
Display pages starting at:   
Display pages ending at:   
Namespace: (Main) ☐ Hide redirects

Access	Access Workflow	Access a repository
Access control list	Active Collection	Active Data
Add a retention period	Addition of access controls	Administrative metadata
Aggregation	Architecture	Attribute
Authentication	Authenticity metadata	Authoritative source
Authorize a deposition	Bit Sequence	Bit Stream
Blueprint	Canonical Data Collection	Catalog
Cataloguing	Checksum	Choosing a storage location
Citable Data	Citation Metadata	Collection
Collection Management Identification	Components	Concept
Conceptual/Logical/Physical Level	Container	Content Interpretation
Content Re-use	Content Replication	Context Information
Corpuse	Create derived data products	Curation Workflow
Darwin Core	Data	Data Acquisition
Data Aggregate	Data Catalog	Data Citation

Definition:

Explanation:

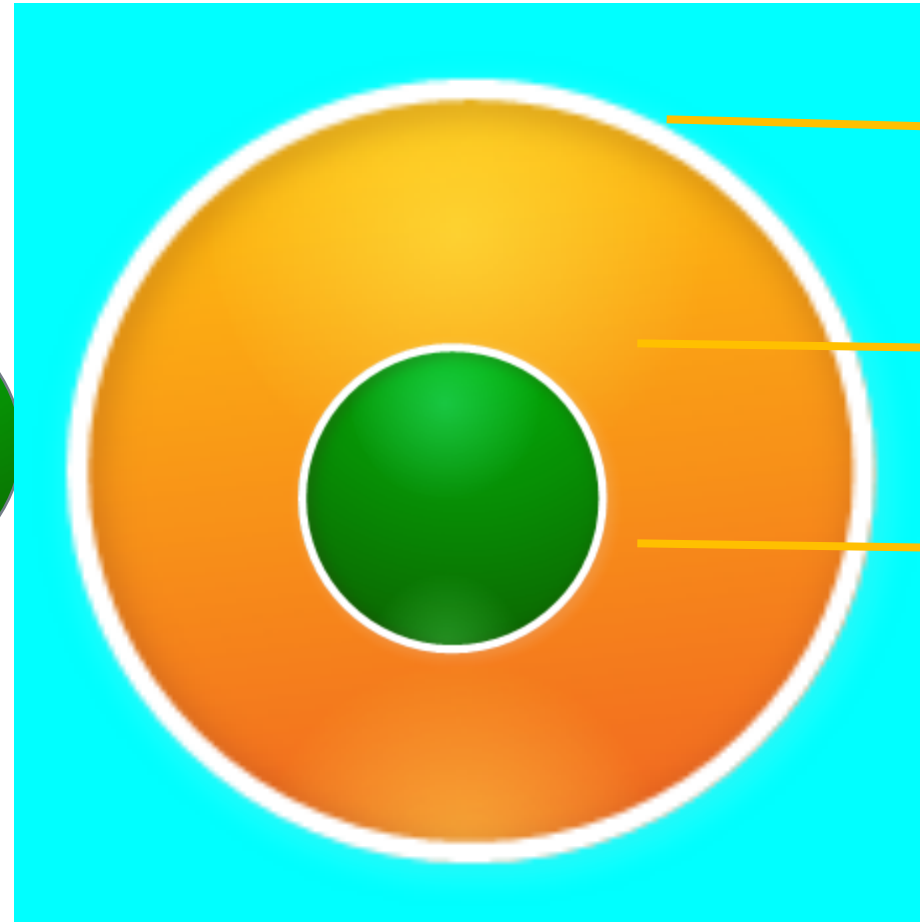
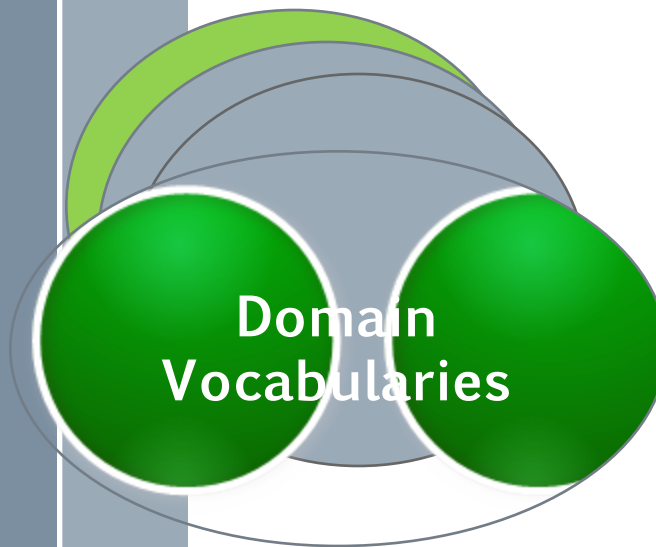
Examples:

References:

Scope:

Status:

# Where we are Going - Building out Vocabulary in Stages



Expand vocabulary as Groups provide input and **collaborate** on them

Firm, update & "connect" concepts

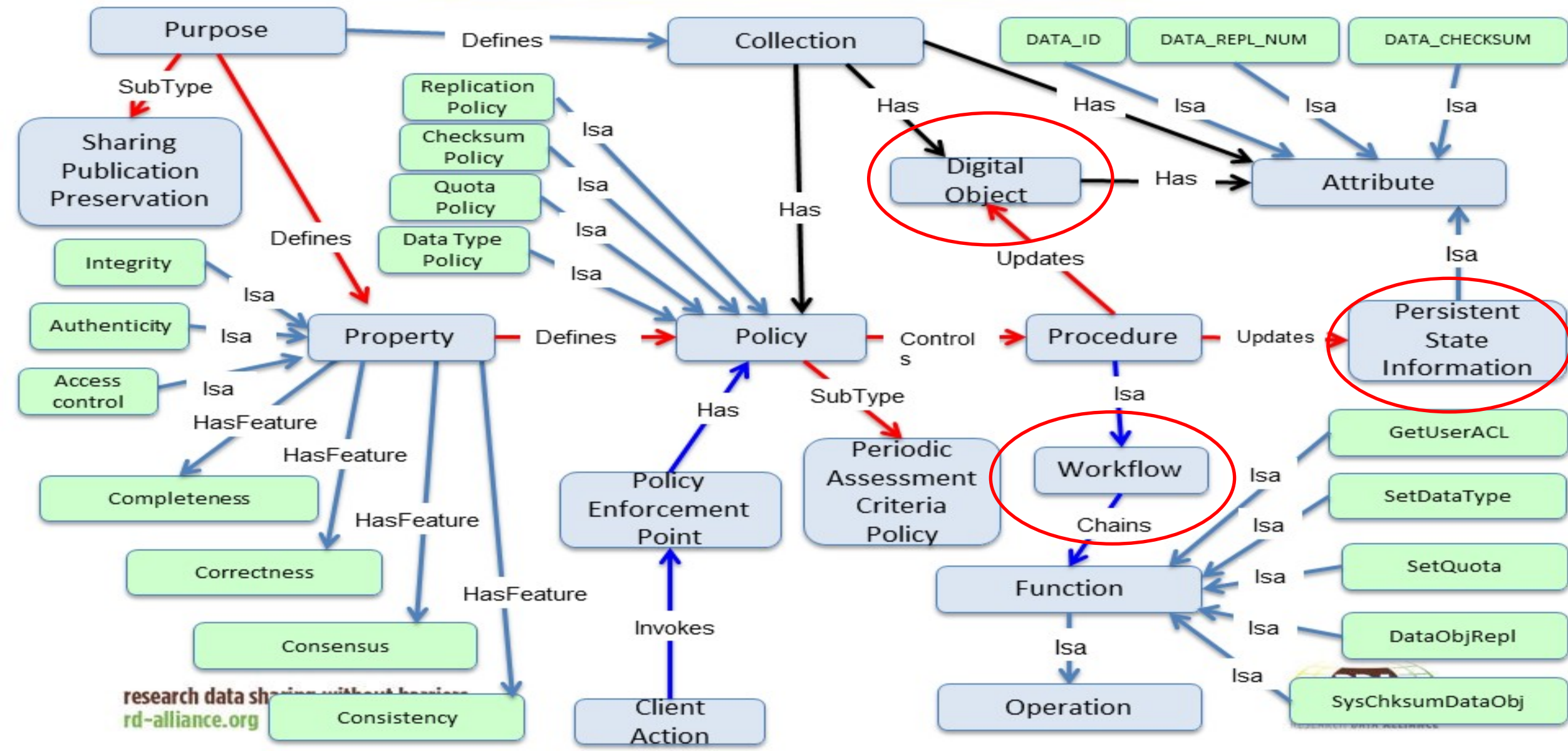
Improved consensus & documentation – **Metadata, collections, ....**



# Policy Components - Conceptual Fundamentals

## Policy-based Data Management Concept Graph

4



# Collaboration & P8 plans

$\pi$

- › We are collaborating with the VSIIG to improve our terminological services such as providing synonyms and “automated ”mappings to other vocabularies
- › We are working with RDF forms to provide URIs to terms and connect to the LinkedData world.
- › This might be leveraged to add more semantics to the effort.
- › We also have plans for a Domain Vocabulary BoF at P8 to begin discussion on possible leveraging work for their vocabulary needs
  - Registering, standardizing, mapping, harmonizing, formalizing
  - dealing with the hierarchy within a vocabulary, and

# Exercising Use case for Vocabulary Services

- › **Goal:** Export RDA DFT vocabulary to Research Vocabularies Australia (RVA)
- › Register this to test the value of RVA for RDA DFT (Data Foundations & Terminology) (exercises RVA APIs)
- › To evaluate the use of Semantic Media Wiki for starting a vocabulary then exporting as RDF for ingest  
**Status:** Complete, next step complete RDF conversion so synonyms are available
- › **Contact:** Gary Berg-Cross



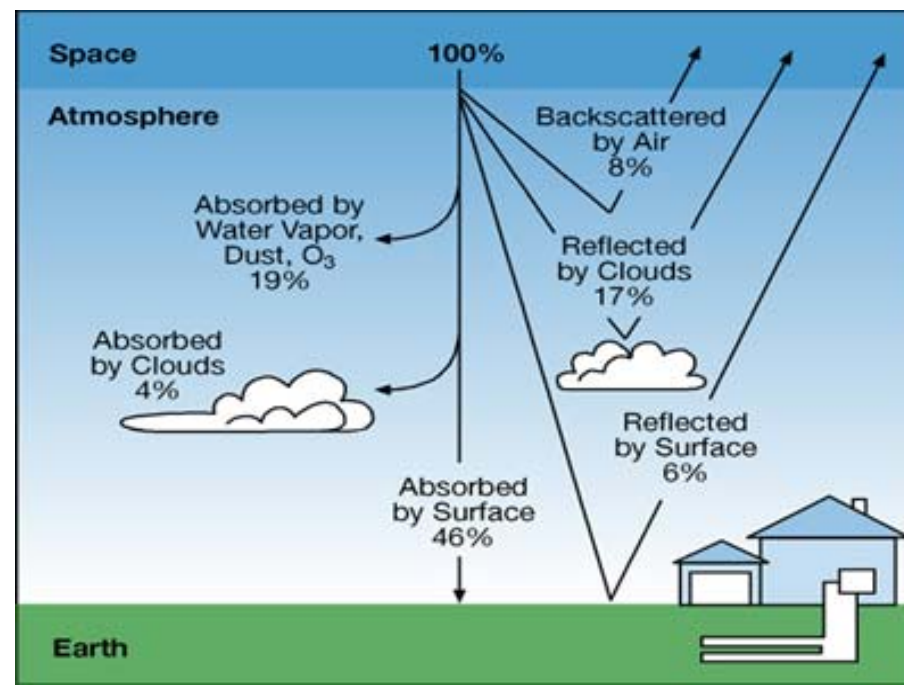
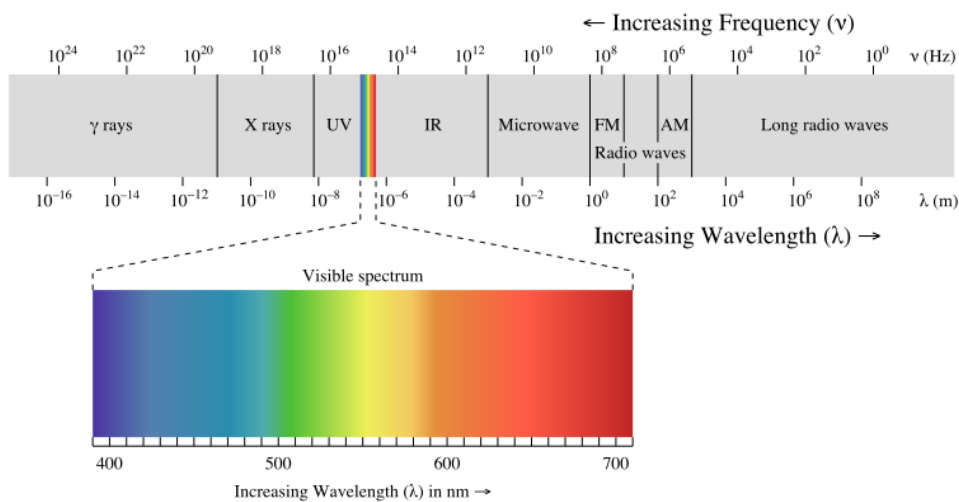
# Some Observations and Ideas

$\pi$

- › Getting consensus is hard...requires sustained collaboration is & a common vision
- › The trend is to leverage agreed upon conceptualizations & augment it with a formalized, digital representational of this knowledge allowing some degree of automated processing.
- › There is much discussion of “registering syntax and declare semantics.” Human understandable “semantics” is captured in readable definitions and can be part of metadata documentation.
- › But formal semantics requires more careful thought and discipline to be useful for semantic search and interoperation.

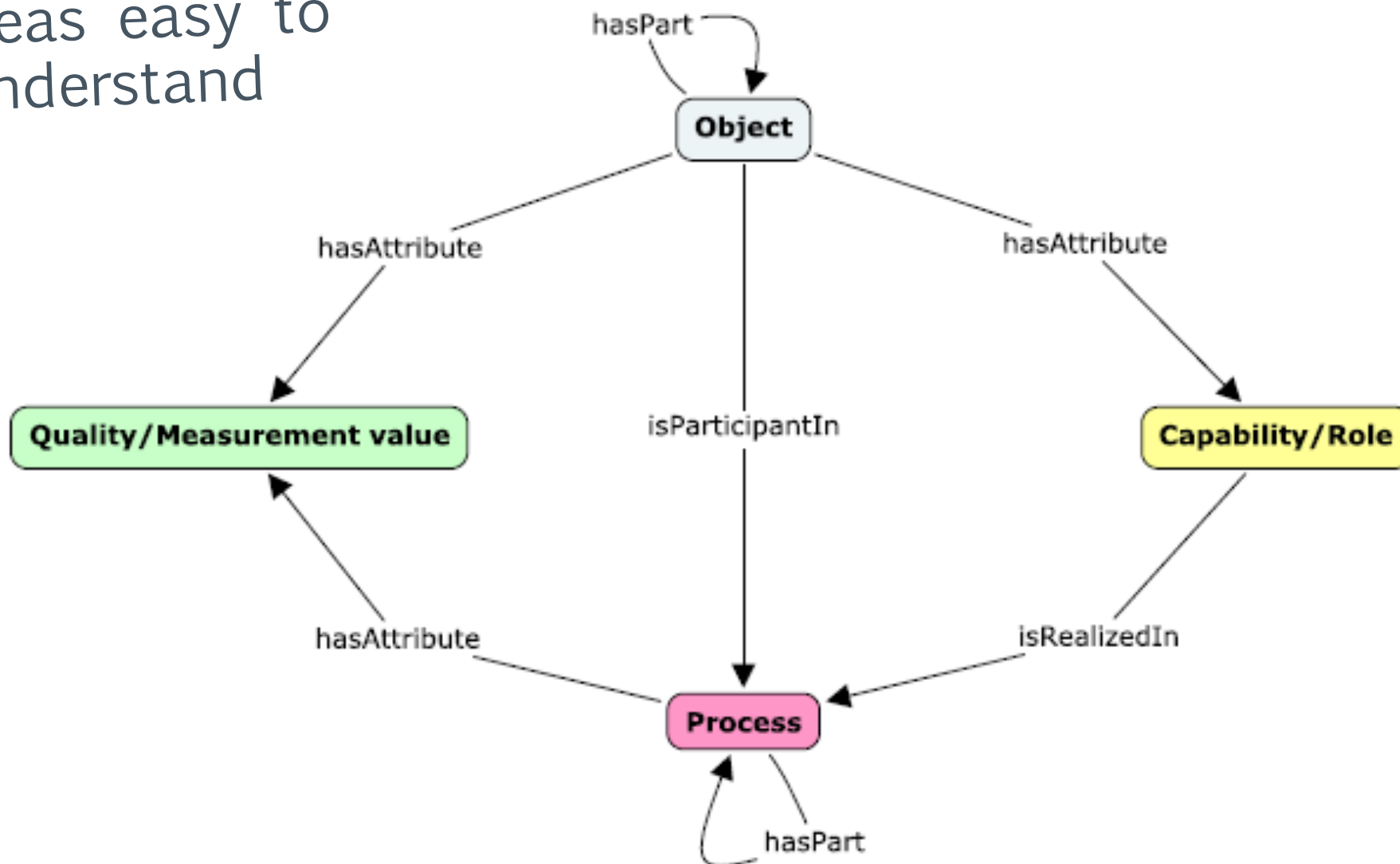
## Example: Radiation-Related Names

atmosphere\_clouds\_radiation~incoming~shortwave\_\_absorbed\_energy\_flux  
atmosphere\_clouds\_radiation~incoming~shortwave\_\_absorptance  
atmosphere\_clouds\_radiation~incoming~shortwave\_\_reflectance  
atmosphere\_clouds\_radiation~incoming~shortwave\_\_reflected\_energy\_flux  
atmosphere\_clouds\_radiation~incoming~shortwave\_\_transmittance  
atmosphere\_clouds\_radiation~incoming~shortwave\_\_transmitted\_energy\_flux  
atmosphere\_clouds\_radiation~outgoing~longwave\_\_emittance  
atmosphere\_clouds\_radiation~outgoing~longwave~downward\_\_energy\_flux  
atmosphere\_clouds\_radiation~outgoing~longwave~upward\_\_energy\_flux



# The SemanticScience Integrated Ontology (SIO)

- › Ideas easy to Understand



# SIO use cases

