



Metadata

Metadata and Semantics Workshop

Rebecca Koskela, Keith Jeffery, Jane Greenberg, Alex Ball

research data sharing without barriers
rd-alliance.org

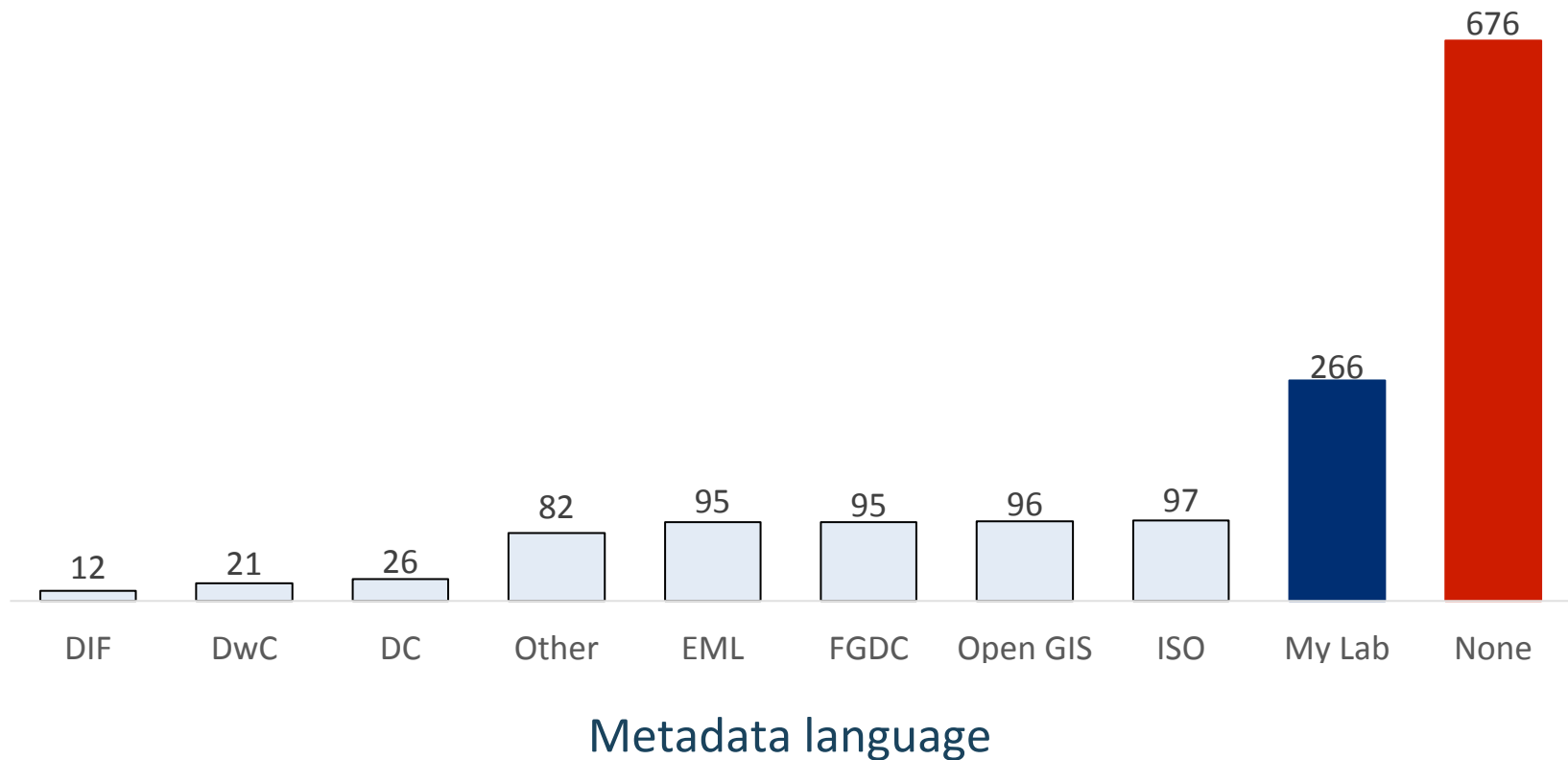
- Metadata Standards Directory Interest Group
 - Global Data Meeting, Washington, DC, October 2012
- Metadata Standards Directory Working Group
 - August, 2013
- Metadata Interest Group
 - First RDA Plenary, Göteborg, Sweden, March 2013
- Metadata in Context Interest Group
 - First RDA Plenary, Göteborg, Sweden, March 2013



Obstacles to wider data sharing

1. Data discoverability challenges
2. Few metadata management tools that support the creation of comprehensive metadata
3. Usability of existing data management tools
4. Availability of best practices and self-paced education materials for scientists and students
5. Paucity of clear and enforceable mandates from funders and publishers
6. Data management insufficiently budgeted/funded
7. Sustainability of key cyberinfrastructure

What metadata standard do you currently use? (Baseline: 2010/2011)



n=1,329. Article:

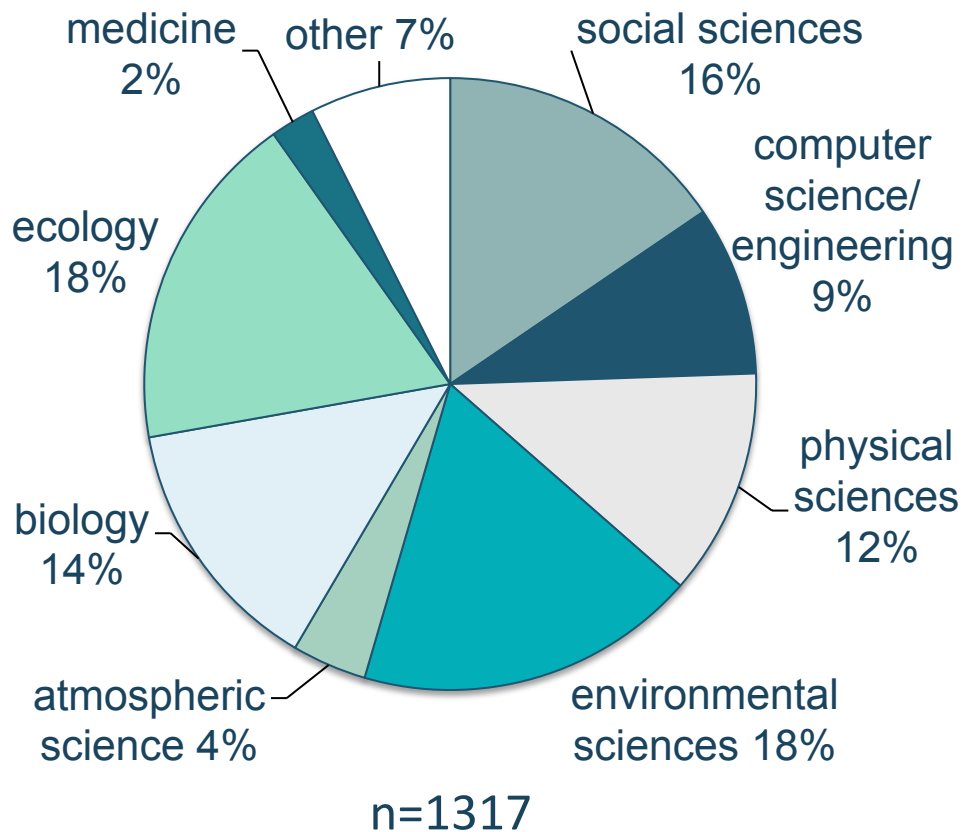
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021101>



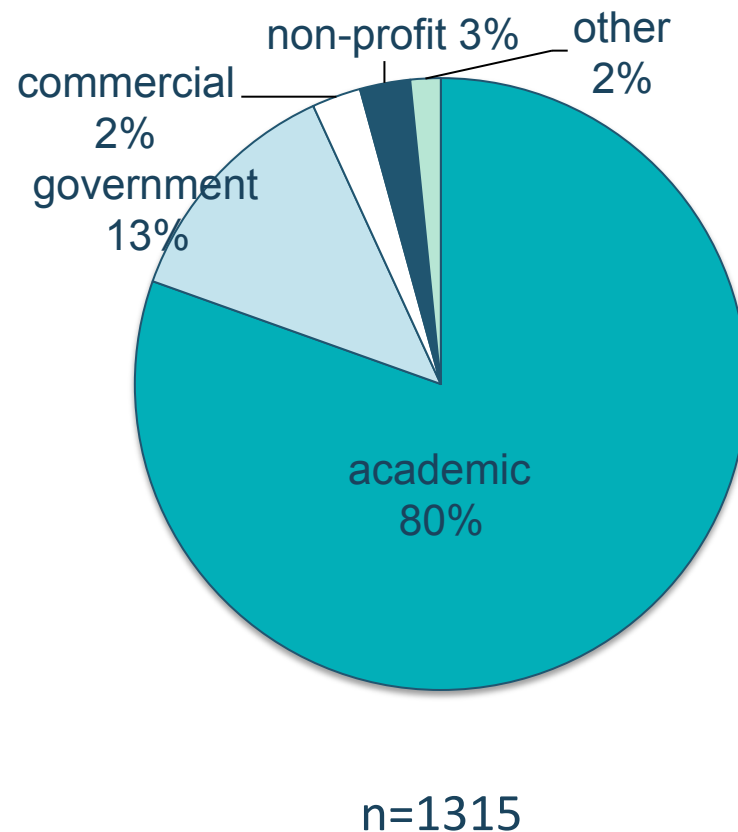
Baseline Assessment: Scientists (2010/2011)

- Demographics

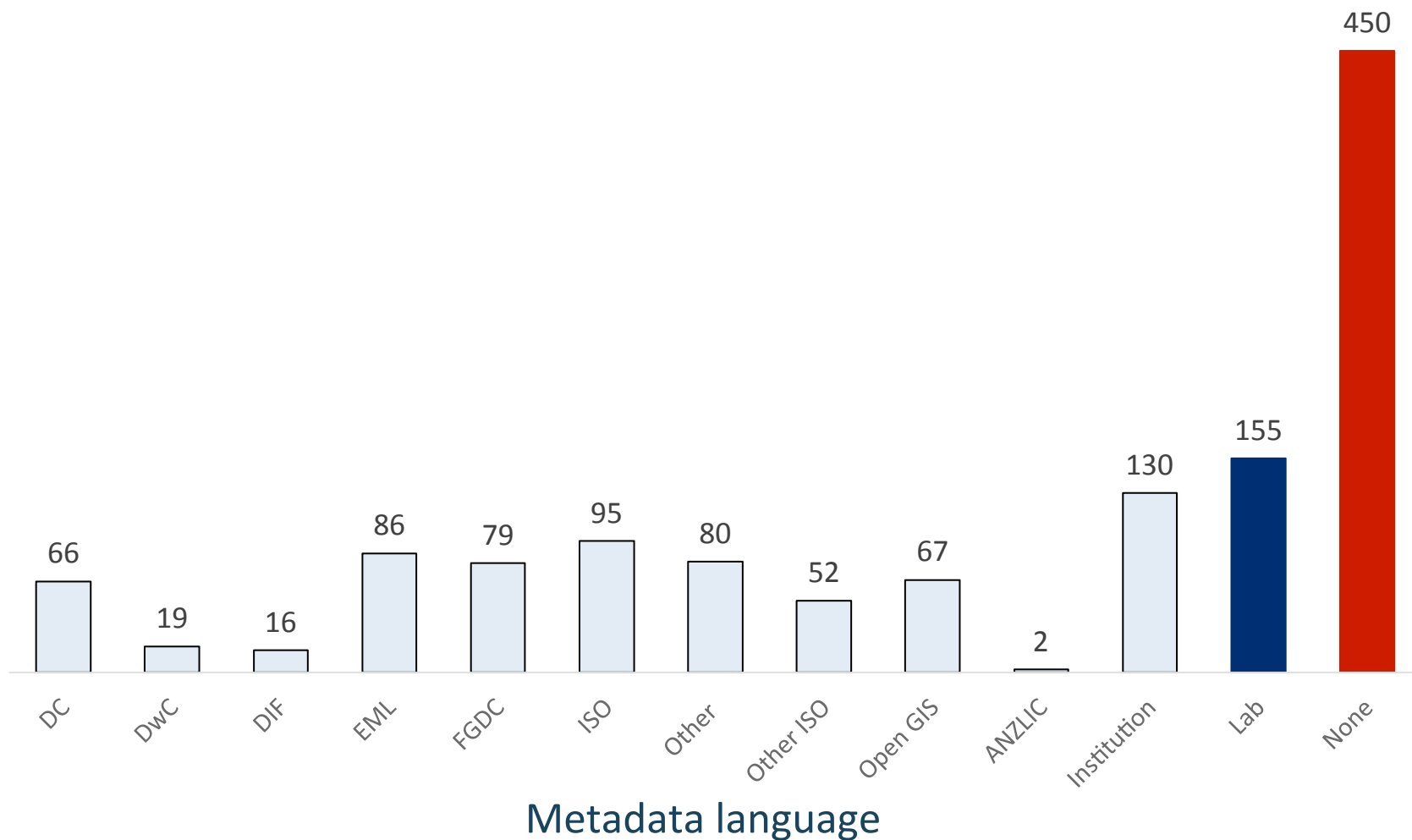
Discipline



Work Sector

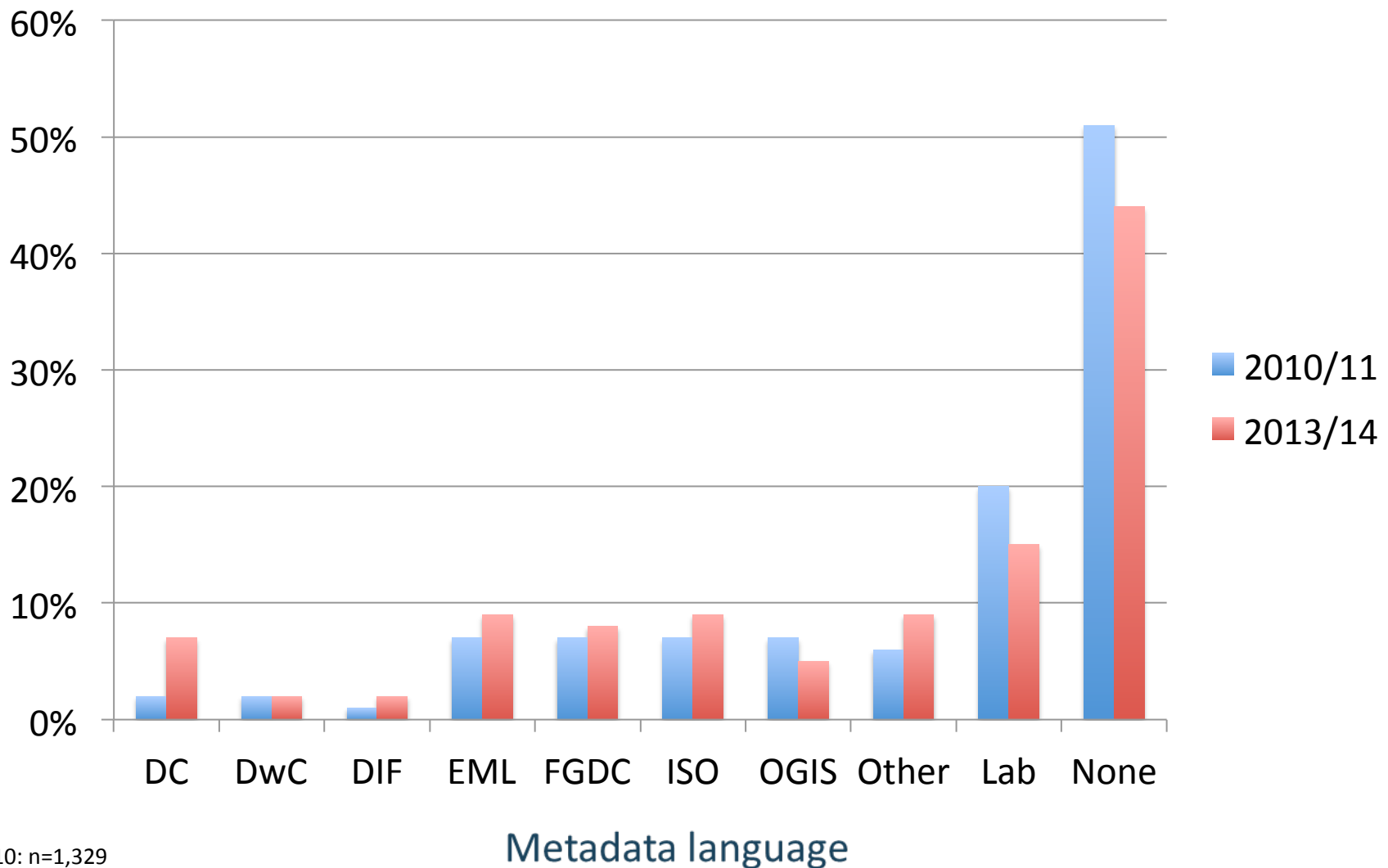


What metadata standard do you currently use? (Follow Up: 2013/2014)



n=1,015


What metadata standard do you currently use? (Baseline & Follow Up Comparison)



2010: n=1,329

2013: n=1,015

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021101>

 **DCC** because good research needs good data

Search

Home	Digital curation	About us	News	Events	Resources	Training	Projects	Community	Tailored support
------	------------------	----------	------	--------	-----------	----------	----------	-----------	------------------

[Home](#) > [Resources for digital curators](#) > Disciplinary Metadata

In this section

Briefing Papers

How-to Guides & Checklists

Developing RDM Services

Curation Lifecycle Model

Curation Reference Manual

Policy and legal

Data Management Plans

Tools

Case studies

Repository audit and assessment

Standards

[Disciplinary Metadata](#)

DIFFUSE

Publications and presentations

Roles

Question journals

Disciplinary Metadata

While data curators, and increasingly researchers, know that good metadata is key for research data access and re-use, figuring out precisely what metadata to capture and how to capture it is a complex task. Fortunately, many academic disciplines have supported initiatives to formalise the metadata specifications the community deems to be required for data re-use. This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them.

For those disciplines that have not yet settled on a metadata standard, and for those repositories that work with data across disciplines, the General Research Data section links to information about broader metadata standards that have been adapted to suit the needs of research data.

Search by Discipline



[Biology](#)



[Earth Science](#)



[General Research Data](#)

Contributions from ...	Standards
Syracuse University	ClinicalTrials.gov Protocol Data Element
Purdue University	Definitions (DRAFT)
Libraries	FITS (Flexible Image Transport
National Center for Ecological Analysis and	System) GenDMS
Synthesis	Genome Metadata
DICE Center, UNC-CH	GEOSS Standards
National Snow and Ice Data Center, Univ. of	Registry Journeau
Colorado	Observ-OM
ResearXis-Discinnet	Observations and Measurements
University Medical Center Groningen	Open Archives Initiative
Commonwealth Scientific Industrial	PROV
Research Organisation	QuDEx
Open Microscopy Environment	Resource Metadata for the Virtual
Rensselaer Polytechnic Institute	Observatory
Newcastle University	WCS - World Coordinate System
UK Data Archive	The Gulf of Mexico Research Initiative
US Virtual Astronomical Observatory /	Information and Data Cooperative (GRIIDC)
Space Telescope Science Institute	
EDINA, The University of Edinburgh	
U. S. Geoscience Information Network	
Met Office, UK	
IMOS/AODN Integrated Marine Observing	
System/Australian Ocean Data Network	
NRCan/GeoConnections	
	<h2>Survey results</h2> <p><i>continued</i></p>



- The only difference between metadata and data is mode of use
- Metadata is not just for data, it is also for users, software services, computing resources
- Metadata is not just for description and discovery; it is also for contextualisation (relevance, quality, restrictions (rights, costs)) and for coupling users, software and computing resources to data (to provide a VRE)
- Metadata must be machine-understandable as well as human understandable for autonomicity (formalism)
- Management (meta)data is also relevant (research proposal, funding, project information, research outputs, outcomes, impact...)

The Broader Picture: Virtualisation of e-Research through Metadata

12

Complete cohort of researchers, research managers, innovators, media

User Model

interaction with data, processing, persons

providing what the user requires

Processing Model

representing research

Data Model

representing ICT

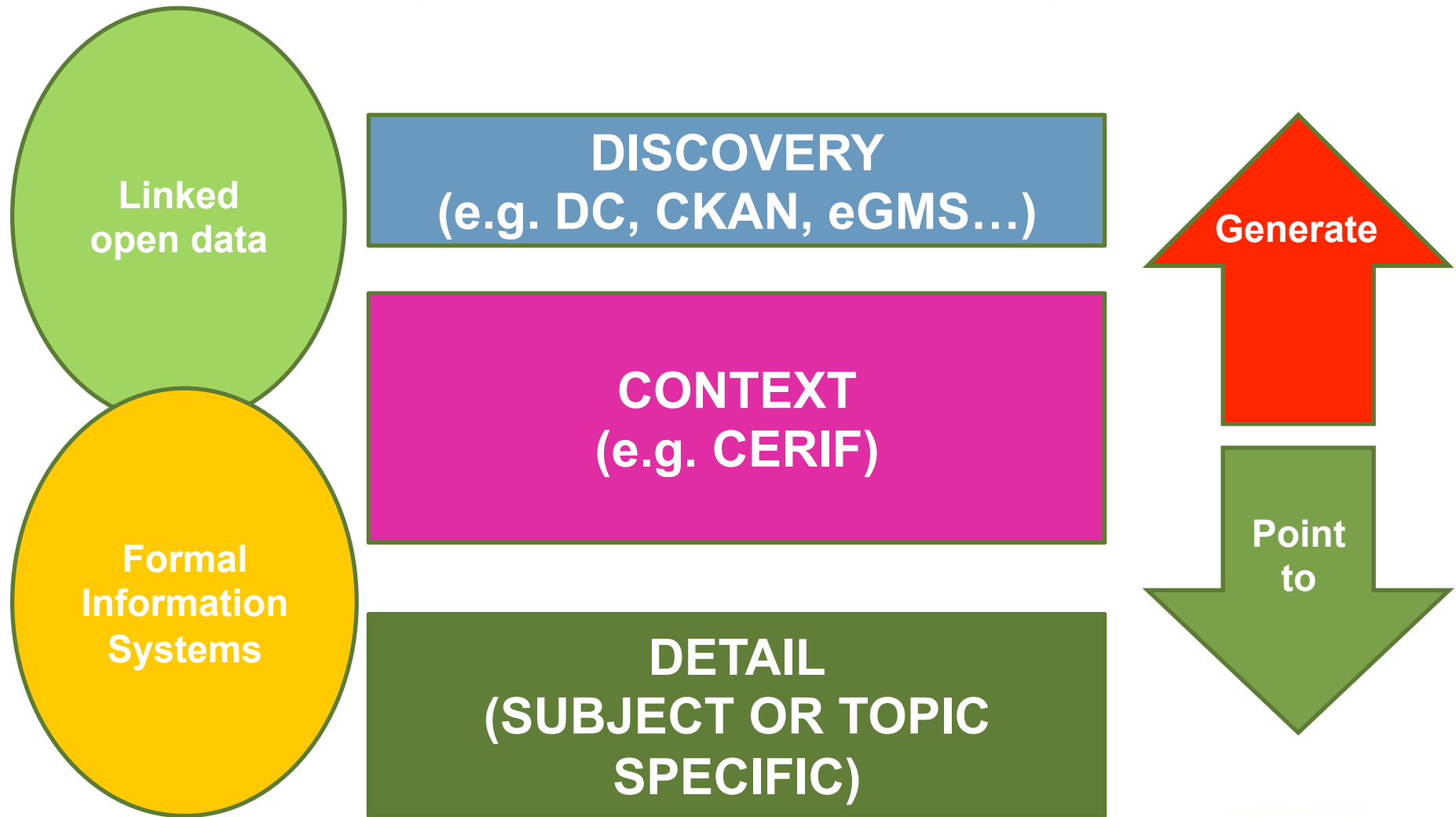
Resource Model

We are talking about this

Complete ICT environment for research

The Broader Picture: Research Data and Open Government Data

13



involves not only metadata groups but all RDA

- Use cases into repository (DICIG)
- Standards into MSDWG directory (MSDWG)
- Analyse for commonalities and differences (MIG)
- Propose canonical metadata 'packages' for 'purposes' (MIG)
- Validation of 'packages' (domain groups)
- Provision of converters (this is a problem!)
- Move to standardisation of 'packages' (RDA)

RDA Collaboration Group Feedback on Metadata

- All groups represented demonstrated the variety and complexity of Metadata
- Types, Terminology, PIDs: formal structures in metadata
- Policy: key-value pairs; relationships carried by metadata
- Domain: many attributes, specific, some commonality, need for interoperation
- Repositories: quality assured by metadata
- Brokering: metadata (and especially good quality metadata) makes brokering feasible

Group Requirements

- All groups represented stressed need for working with Metadata Groups in general require:
 - Advice on what standard(s) to use
 - Assistance in metadata standards
 - Convergence to common metadata model
 - Cannot be done globally even in any one domain
 - But possibly can achieve at discovery and contextual level
 - Interoperation among many metadata models
 - Requires best possible metadata
 - And brokering technology for matching/mapping
 - Then generation of converters for data instances
 - Common metadata packages
 - Avoids the $n*(n-1)$ problem
 - Superset (at level of granularity)

- Syntax (metadata standards structures – what they cover)
 - Objects/entities and properties/attributes
- Semantics (terms in metadata standards – what they mean)
 - Relationships between terms including multilinguality
- Temporal information
 - Relationships not base information
 - Provides the temporal interval when the assertion is true
- Integrity
 - Referential (represent dependencies)
 - Functional (all attributes depend uniquely on the unique ID)
- Represented in some form of first order logic
 - Allows induction and deduction – saves input and permits brokering
 - Performance

Metadata at RDA P5

Monday, March 9

- Joint Session: MIG, MSDWG, DIC, RDP
 - <https://rd-alliance.org/joint-session-metadata-groups>

Tuesday, March 10

- Metadata Plenary Session
 - <https://rd-alliance.org/Metadata-Plenary-Session>

Wednesday, March 11

- Environment-related WGs & IGs Plenary Session
 - <https://rd-alliance.org/Environment-related-Plenary-Session>