



Research Data Alliance Outputs

David Schade

Canadian Astronomy Data Centre

Canadian Virtual Observatory

Canadian Advanced Network for Astronomy Research (CANFAR)

Co-Chair: Research Data Alliance Interest Group

Disciplinary Collaboration Framework





O&A Members

56

Active Organisational & Affiliate members

MEMBERSHIP

Members: 8284

Becoming a member of RDA is simple and open to both individuals and organizations

[Register now](#)

RDA Groups

WG & IGs: 102

Discover what RDA Working and Interest Groups and all other Groups are up to and find out how to join them. [Explore Groups](#)

- [ABOUT RDA](#)
- [GET INVOLVED](#)
- [GROUPS](#)
- [RECOMMENDATIONS & OUTPUTS](#)
- [RDA FOR DISCIPLINES](#)
- [PLENARIES & EVENTS](#)
- [NEWS & MEDIA](#)



RDA 14th Plenary Meeting

Helsinki, Finland
 23 - 25 October 2019
[More Information Here](#)

[NEWS & EVENTS](#)

[RECENT BLOGS](#)

[FOLLOW US](#)

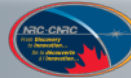


University of Victoria



University of British Columbia

canarie



compute + calcul
CANADA





Outputs and Recommendations

- [Scalable Dynamic-data Citation Methodology](#)
- [Data Description Registry Interoperability Model](#)
- [Basic Vocabulary of Foundational Terminology Query Tool](#)
- [Data Type Model and Registry](#)
- [The FAIRsharing Registry and Recommendations: Interlinking Standards, Databases and Data Policies](#)
- [Persistent Identifier Type Registry](#)



University
of Victoria



University of
British Columbia

canarie



compute + calcul
CANADA





RDA is a place where you could look for relevant outputs that represent useful thinking on your subject



University of Victoria



University of British Columbia

canarie

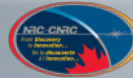
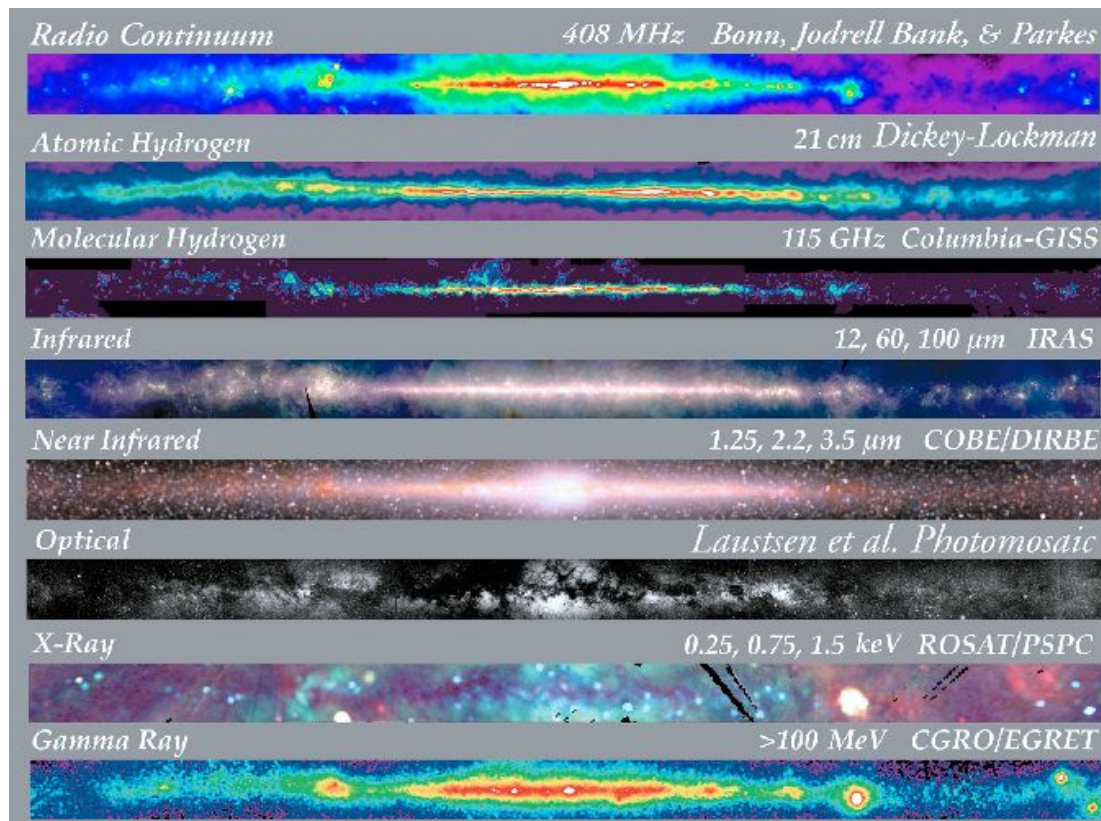


compute+calcul
CANADA

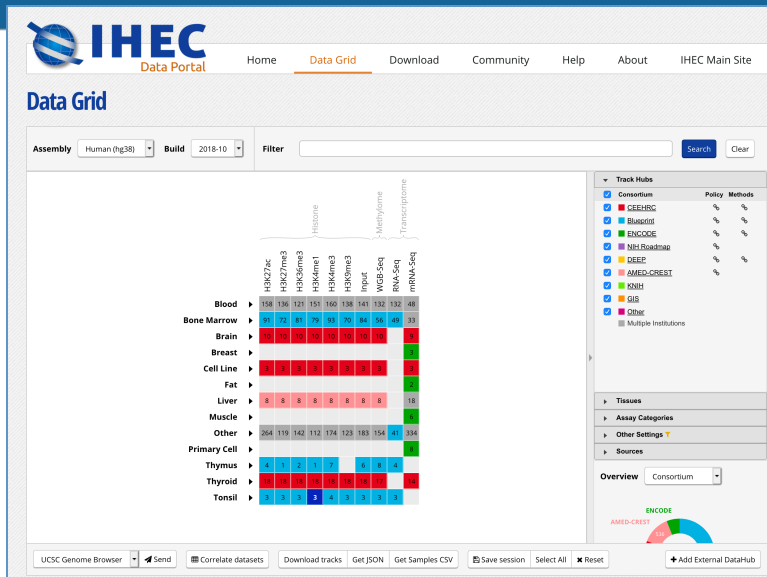


Astronomy is multi-wavelength & multi-facility

The universe cannot be understood without looking across a large range of (wavelength) energy of “light”



Integrative Research is everywhere



Standardization of metadata to support InterOperability is *everywhere*

Integrative Research is everywhere



A screenshot of the IHEC Data Grid web application. The interface is titled 'Ocean Networks Canada Data Search' and is logged in as 'David Schade'. It features a navigation menu with 'Home', 'Data Grid', 'Download', 'Community', 'Help', and 'About'. The main content area is divided into three steps: 'Step 1. Choose Data Source', 'Step 2. Select Data Product', and 'Step 3. View Cart (0 Items)'. On the left, there are 'Data Source Filters' for 'Depth Range' (From/To in meters) and 'Time Period Selection' (From/To in UTC). The 'Data Source Selection' panel shows a tree view with 'Ocean Networks Canada' expanded to show 'Mobile Platforms' and 'Pacific'. A map on the right displays the Pacific Ocean region with various data points marked by colored circles and squares. A 'Select Data Product' button is visible on the map. The bottom of the interface includes a 'Next >>' button and a 'UCSC Genome Browser' link.

Standardization of metadata to support InterOperability is *everywhere*

Integrative Research is everywhere



The image shows two overlapping web browser screenshots. The background screenshot is the IHEC Data Grid interface, which includes a navigation menu with "Home", "Data Grid", "Download", "Community", "Help", "About", and "IHEC Main Site". It also shows a "Data Source Selection" step in a search process. The foreground screenshot is the Polar Data Catalogue search interface, which features a map of the Arctic region and a search form. The search form includes fields for "Latitude", "Longitude", "Start Date", "End Date", and "Word or Phrase". It also has a "Polar Data Catalogue Search" section with instructions on how to use the search tool. The search results section shows a table of coordinates and a "Number of Search Options" of 8.

Standardization of metadata to support InterOperability is *everywhere*

Integrative Research is everywhere



The image displays three overlapping web interfaces illustrating data integration and search capabilities:

- IHEC Data Grid:** Shows a navigation menu with 'Data Grid' selected. It includes a 'Data Source Filters' section with options for 'Depth Range' (From/To in meters) and 'Time Period Selection' (From/To in UTC). A 'Common Property Selection' section is also visible.
- Polar Data Catalogue:** Features a map of the Arctic region with a green line indicating a search path. The interface includes a 'PDC Geospatial Search' header and a 'PDC Search' button.
- NSIDC Glacier Photograph Collection:** Shows a 'Search & Order' form with sections for 'Define Your Search', 'Spatial and Temporal Information', and 'Photographic Information'. The 'Spatial Search' section includes fields for Country, State/Province, Minimum Year, Maximum Year, and Months. The 'Photographic Information' section includes fields for Glacier Name and Photographer Name.

Standardization of metadata to support
InterOperability is *everywhere*

Integrative Research is everywhere

A collage of four overlapping web portals. The top-left portal is the IHEC Data Grid, showing a search interface for Ocean Networks Canada. The middle-left portal is the Polar Data Catalogue, displaying a map of the Arctic region. The middle-right portal is the NSIDC Glacier Photograph Collection, showing a search and order interface. The bottom-right portal is the NIMH Data Archive, displaying a search result for "UIC ACE: Translational Studies of Insistence on Sameness in Autism" with a Venn diagram showing the overlap of Imaging+, Omics, and Clinical data.

IHEC Data Grid
Ocean Networks Canada Data Search
Step 1. Choose Data Source | Step 2. Select Data Product | Step 3. View Cart (0 Items)

POLAR DATA CATALOGUE
PDC Geospatial Search
Arctic | Antarctic

NSIDC National Snow & Ice Data Center
Glacier Photograph Collection
Search & Order

NIMH Data Archive
UIC ACE: Translational Studies of Insistence on Sameness in Autism #1
Collection Title: UIC ACE: Translational Studies of Insistence on Sameness in Autism
Investigators: Ed Cook
Description: The UIC ACE was focused on the genetics, neurobiology, cognitive and affective processes, and pharmacology of insistence on sameness (IS) in autism spectrum disorders (ASD). A large sample of children with self-reported autism spectrum disorder will be screened by the Assessment Core for further...
Data Source: NDAR
Collection Phase: Funding Completed
Collection State: Shared
Blinded Clinical Trial: No

Venn Diagram Data:

Category	Count
Imaging+	16
Omics	89
Clinical	79
Imaging+ & Omics	42
Imaging+ & Clinical	16
Omics & Clinical	79
Imaging+, Omics & Clinical	521

Standardization of metadata to support InterOperability is everywhere

Integrative Research is everywhere



The collage features several overlapping screenshots of scientific data portals:

- IHEC Data Grid:** Shows a navigation menu with 'Home', 'Data Grid', 'Download', 'Community', 'Help', and 'About'. It includes a 'Data Source Selection' step and a 'Depth Range' filter.
- Polar Data Catalogue:** Displays a map of the Arctic region with a search bar and navigation options like 'Arctic' and 'Antarctic'.
- NSIDC National Snow & Ice Data Center:** Features a 'Glacier Photograph Collection' search interface with a 'Search & Order' section.
- NIMH Data Archive:** Shows a search interface with filters for 'Data Source', 'Collection Phase', and 'Funding Source'. It includes a 'Reset Filters' button and a 'Filter' button.
- SCDDB (Scottish Country Dance Database):** Displays a search results page for 'UIC ACE: Translational Studies of Insistence on Sameness'. It includes sections for 'Overview', 'Features', and 'What you can do'. The 'Features' section lists: 'Largest SCD database on the web - contains comprehensive information on dances, formations, people, publications, alliums, recordings, and tunes. Runs rings around the competition!', 'Includes Eric Ferguson's cribs and MinCribs and over 2000 Pilling-style diagrams for dances.', 'Catalogues more than 1000 curated videos for specific dances from YouTube and other sources.', 'Offers "taster" snippets from recordings by many popular musicians and bands.', 'Powerful search functions for dances and recordings.', 'Rate and review dances.*', 'Dance lists make devising programmes and producing cribs and diagram sheets a breeze.*', 'Publish event programmes, class logs, and lists of your favourite dances.*', 'Embed "live" dance lists on your own web pages.', 'Collections allow convenient searching through exactly the publications you own (rather than all of the database)*', and 'The data is actively maintained and kept up to date, and new and exciting features are added all the time!'. The 'What you can do' section states: 'Feel free to browse the database and use the information in it to improve your own and others' enjoyment of Scottish country dancing! In spite of tireless efforts to the contrary, the database probably contains lots of errors and omissions. If you notice anything that is wrong or looks funny, please notify the maintenance team - most conveniently through the "Submit a Correction/Addition" button that is part of every data page. Also, please get in touch if you have a resource that is not in the database at all and that you think ought to be added. We're always on the lookout for more volunteers to help us improve the database. If you are interested in SCD and have a little time to spare, why not dive in? See here for further information. If you would like to use the data in this database for other projects, you can download daily dumps of the database content (or much of it, anyway - we don't include "personal" items like dance lists and collections) to use under a "Creative Commons" license. (* = Function requires free mystrathstepy account)'. Below the main content are four boxes: 'Recent Comments' (Annie-Marie Forsyth on Catch the Wind...), 'Recent Cribs' (Call of the Island (S2, Anderson) from E-cribs; Touring the World (S32, Duggan: From...), 'Recent Lists' (Budget Local Club 22.01.2019 by Ildikó Szokolczai (Jan. 23, 2019, 10:09 a.m.)), and 'Recent Videos' (Slymish House (S32, Ronald: RCCDS Ltd) by YouTube Automatic Uploader; Salute to Noel Lillie (S32, Gertz:)).

Standardization of metadata to support InterOperability is everywhere

For example:

Data Citation of Evolving Data



Recommendations of the Working Group on Data Citation (WGDC)

Andreas Rauber, Ari Asmi, Dieter van Uytvanck and Stefan Pröll

Revision of October 20th 2015

I. MAKING DATA CITABLE

These WGDC recommendations enable researchers and data centers to identify and cite data used in experiments and studies. Instead of providing static data exports or textual descriptions of data subsets, we support a dynamic, query centric view of data sets. The proposed solution enables precise identification of the very subset and version of data used, supporting reproducibility of processes, sharing and reuse of data.

Goals of this WG are to create identification mechanisms that:

- allows us to identify and cite arbitrary views of data, from a single record to an entire data set in a precise machine-

A. Preparing the Data and the Query Store

Prepare existing data sources and provide the required infrastructure, which is needed for implementing the query based approach.

- **R1 – Data Versioning:** Apply versioning to ensure earlier states of data sets can be retrieved.
- **R2 – Timestamping:** Ensure that operations on data are timestamped, i.e. any additions, deletions are marked with a timestamp.
- **R3 – Query Store Facilities:** Provide means for storing queries and the associated metadata in order to re-execute them in the future.

B. Persistently Identify Specific Data Sets

When a data set should be persisted, the following steps need to be applied:



University
of Victoria



University of
British Columbia

canarie



compute+calcul
CANADA



Recommendation on data citation

- Relies on executing a query that returns exactly the dataset to be cited
- Saves storage
- Versions, timestamps of data and queries

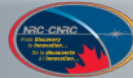


University
of Victoria



University of
British Columbia

canarie



compute + calcul
CANADA



Recommendation on data citation

I view this as high-level work that presents obstacles to practical implementation

- The assumption is that all datasets that you want to publish/cite can be represented as a query (or set of queries)
 - Not all published data exists in queryable systems
 - Datasets are distributed across systems
 - Astronomers find data from queries (sometimes) and then apply arbitrary selections

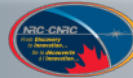


University
of Victoria



University of
British Columbia

canarie



CADC/CCDA
McGill University
Canada



compute + calcul
CANADA





RDA is a place where you could look for relevant outputs that represent useful thinking on your subject



University of Victoria



University of British Columbia

canarie



compute + calcul
CANADA



Disciplinary Collaboration Framework Interest Group

Ultimately, a major goal of the Research Data Alliance is increasing the power of ***cross-disciplinary research***

The universe cannot be understood by studying a single element



University
of Victoria



University of
British Columbia

canarie



compute + calcul
CANADA



Metadata: The principal challenge

Data models: the metadata and metadata structures that characterize a data collection

Development of cross-discipline Data Models is the principal challenge



University
of Victoria



University of
British Columbia

canarie



compute + calcul
CANADA





Generic Stream

- Very light metadata
 - limited “Findability”
- Little Integration or harmonization of metadata
 - limited accessibility
- Mix of content: data, literature, links, documents
- Not developed by researchers
- Provide useful service
- Support obligation to publish data

Research-centric Stream

- Metadata that has sufficient depth for the research
- Data Integration
- Management of data types and links to appropriate applications
- Developed by researchers themselves
- **Integral part of the advancement of research**

Disciplinary Collaboration Framework Interest Group

There are successful examples ***cross-disciplinary research*** and ***Integrative cross-disciplinary Research Data Management systems***

Let's learn from those successes

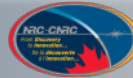


University
of Victoria



University of
British Columbia

canarie



compute + calcul
CANADA



END



University
of Victoria



University of
British Columbia

canarie



compute+calcul
CANADA

