RDA 10 Montréal

feedback

André Schaaff

Centre de Données astronomiques de Strasbourg

DCP Session

IVOA, Santiago de Chile, 20-22/10/2017

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Astronomy ESFRI & Research Infrastructure Cluster

H2020-Astronomy ESFRI and Research Infrastructure Cluster (Grant Agreement number: 653477).

Purpose

- A summary of selected sessions of RDA Montréal, Plenary 10, September 2017

IG Federated Identity Management

- See GWS session 1 slides about a dedicated feedback FIM4R & RDA IG FIM

IG Repository Platforms for Research Data

- “collection of experiences with selecting, implementing, and using specific research data repository platforms / products across institutions and domains”
- Participation to the case statement
- Presentations and discussions around Fedora, Sauvera, Islandora, Dataverse, KIT DM, Dash, etc.
- …

WG Research Data Repository Interoperability

- “will establish standards for interoperability between different research data repository platforms focusing on machine-machine communication with the primary goal of enabling migration, replication and cross-repository discovery”
- Participation to the primer (a deliverable of the WG) by describing IVOA VOSpace
- Recommendation under writing (example: Packages MUST contain a metadata file datacite.xml with metadata following the DataCite Metadata Schema)

WG Provenance Patterns

- “focus on finding, detailing and recommending best practices for provenance representation and management”
- IVOA Provenance work is now known in the WG
- Orientation more to ontologies
- Should be followed by the IVOA

BoF The Digital Representation of Scientific Units of Measure

- “representation of digital scientific units is fragmented with a number of projects/ontologies/approaches to address this issue” + https://units.unf.edu : funded by NIST
- Identify the different current unit representations that communities have implemented => done for VOUnits
- => to follow if it becomes an IG

IG Software Source Code

- “discuss issues on management, sharing, discovery, archival and provenance of software source code, review and revise metadata for describing and discovering source code, develop guidelines for managing, describing and publishing software source code, collect and publish use cases of current examples and practices”

IG Software Source Code (2)

- Survey following the Montréal session (M. Gruenpeter)
- People interests
  - cite software
  - recover software
  - data needs software (without software, data is basically useless)
  - discover software
  - provide framework for better software discovery and research
  - reuse software (with environment)
  - manage software
  - preserve software
  - software as a first class research product
  - describe properly software
  - PID for software
  - identify and incorporate better practices for software
  - software provenance

IG Software Source Code (3)

- Use cases

- discover software by searching for specific interest
  - domain/area
  - algorithm/ functionality
  - data provided (software producing a certain result with a particular data set)
  - environment (software used in particular environment -distro, compiler, etc.)
  - conditions for use, reuse and modification
- publish/ deposit/ archive software with associated metadata
- link software artefact to its context:
  - data
  - people / authorship
  - funding
  - dependencies
  - built form

IG Software Source Code (4)

- Use cases (2)

- cite software and give due credit
- integrate software to other workflow
    - reproduce software
        - discover dependencies and environment needed

- Ontologies

- most don't or use data ontologies
- Datacite
- CodeMeta
- DublinCore
- package management (NPM, gemspec, PYPI)

IG Software Source Code (5)

- Properties needed
  - PID
  - maintainer email
  - link to compiled version
  - repository retrieval link
  - data input/output expected
  - authorship & affiliation
  - version
  - description
  - references
  - origin source (for provenance)
  - type
  - description / algorithms / problem solved
  - language
  - revisions+ dates
  - funders
  - is documented & documentation link
  - terms of use / license
  - dependencies
  - compiler
  - environment (compiles / run on)
  - status
  - examples
  - related_to (relations to other software)
  - publisher
  - tests link & test data

IG Software Source Code (6)

- Advantages for structured and linked data
  - help scientists discover software
  - better research
  - better connection with data
  - better credit (by linking to authors)
  - backward and forward linking
  - better ecosystem
- An interest at IVOA level
- I will participate to the teleconferences

IG Preservation Tools, Techniques and Policies

- Aim could be resumed as “Finding a way to bridge the gap between researchers and repositories”
- PresQT (Preservation Quality Tool) workshop before RDA

BoF Systems, Technologies and Data Flows

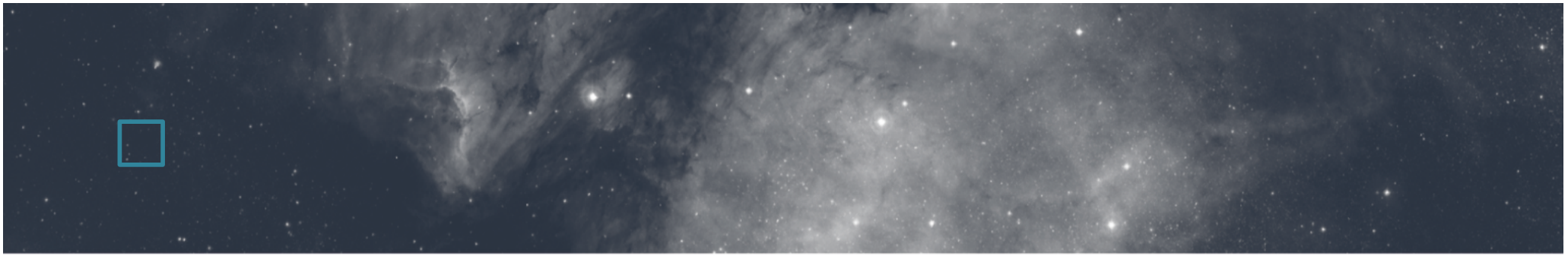
- Comments

WG Array database

- Initially a use case to define in this WG
- Comments

Conclusion

- No dedicated IG / WG to our domain
- But many opportunities to participate and to get involved



H2020-Astronomy ESFRI and Research Infrastructure Cluster (Grant Agreement number: 653477).