Horizon2020
European Union Funding
for Research & Innovation

H2020-INFRAEOSC-2018-2 824087
H2020-INFRAEDI-2018-1 823830
H2020-INFRAIA-2`017-1 730976
H2020-INFRADEV-2019-2 871118
H2020-INFRAIA-2018-1 823827

RO-Crate: Describing and packaging FAIR Research Objects

RO-Crate

Stian Soiland-Reyes
The University of Manchester
University of Amsterdam
https://orcid.org/0000-0001-9842-9718
@soilandreyes

Carole Goble
The University of Manchester
ELIXIR-UK
https://orcid.org/0000-0003-1219-2137
@caroleannegoble

Based on https://youtu.be/pz-MLdI7GLA
Scottish Covid-19 Response Consortium, *2021-03-18*

This work is licensed under a
Creative Commons Attribution 4.0 International License

Open Access | Published: 15 March 2016

The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao & Barend Mons

Scientific Data **3**, Article number: 160018 (2016) | Cite this article

173k Accesses | **2207** Citations | **1794** Altmetric | Metrics

Abstract

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

www.nature.com/scientificdata

SCIENTIFIC DATA

Findable
Accessible
Interoperable
Reusable

https://doi.org/10.1038/sdata.2016.18

FAIR Guiding Principles

https://doi.org/10.1038/sdata.2016.18

Findable

F1. (meta)data are assigned a globally unique and persistent **identifier**

F2. data are described with rich **metadata** (defined by R1 below)

F3. metadata clearly and explicitly **include the identifier** of the data it describes

F4. (meta)data are registered or **indexed** in a searchable resource

Interoperable

I1. (meta)data use a **formal**, accessible, shared, and broadly applicable language for **knowledge representation**.

I2. (meta)data use **vocabularies** that follow FAIR principles

I3. (meta)data include qualified **references** to other (meta)data

Accessible

A1. (meta)data are **retrievable** by their identifier using a standardized communications **protocol**

A1.1 the protocol is **open**, free, and universally implementable

A1.2 the protocol allows for an **authentication and authorization** procedure, where necessary

A2. **metadata are accessible**, even when the data are no longer available

Reusable

R1. meta(data) are richly described with a plurality of accurate and relevant **attributes**

R1.1. (meta)data are released with a clear and accessible data usage **license**

R1.2. (meta)data are associated with detailed **provenance**

R1.3. (meta)data meet domain-relevant community **standards**

tl;dr: machine-readable metadata

Best practices for workflow reproducibility

Methods

(..)

De novo assembly and binning

Raw reads from each run were first assembled with **SPAdes v.3.10.0**[20] with option **--meta**[21]. Thereafter, **MetaBAT 2**[15] (**v.2.12.1**) was used to bin the assemblies using a minimum contig length threshold of 2,000 bp (**option --minContig 2000**) and default parameters. Depth of coverage required for the binning was inferred by mapping the raw reads back to their assemblies with **BWA-MEM v.0.7.16**[45] and then calculating the corresponding read depths of each individual contig with **samtools v.1.5**[46] ('samtools view -Sbu' followed by 'samtools sort') together with the **jgi_summarize_bam_contig_depths** function from **MetaBAT 2**. The QS of each metagenome-assembled genome (MAG) was estimated with **CheckM v.1.0.7**[22] using the **lineage_wf workflow** and calculated as: level of completeness − 5 × contamination. Ribosomal RNAs (rRNAs) were detected with the **cmsearch** function from **INFERNAL v.1.1.2**[47] (**options -Z 1000 --hmmonly --cut_ga**) using the **Rfam**[48] covariance models of the bacterial 5S, 16S and 23S rRNAs. Total alignment length was inferred by the sum of all non-overlapping hits. Each gene was considered present if more than 80% of the expected sequence length was contained in the MAG. Transfer RNAs (tRNAs) were identified with **tRNAscan-s.e. v.2.0**[49] using the bacterial tRNA model (option -B) and default parameters. Classification into high- and medium-quality MAGs was based on the criteria defined by the minimum information about a metagenome-assembled genome (MIMAG) standards[23] (high: >90% completeness and <5% contamination, presence of 5S, 16S and 23S rRNA genes, and at least 18 tRNAs; medium: ≥ 50% completeness and <10% contamination). (...)

Article | Open Access | Published: 11 February 2019

A new genomic blueprint of the human gut microbiota

Alexandre Almeida, Alex L. Mitchell, Miguel Boland, Samuel C. Forster, Gregory B. Gloor, Aleksandra Tarkowska, Trevor D. Lawley & Robert D. Finn

Nature **568**, 499–504(2019) | Cite this article

99k Accesses | **132** Citations | **667** Altmetric | Metrics

EMBL-EBI

(..)

Assignment of MAGs to reference databases

Four **reference databases** were used to classify the set of MAGs recovered from the human gut assemblies: **HR**, **RefSeq**, **GenBank** and a collection of **MAGs** from public datasets. HR comprised a total of 2,468 high-quality genomes (>90% completeness, <5% contamination) **retrieved** from both the HMP catalogue (https://www.hmpdacc.org/catalog/) and the HGG[8]. From the **RefSeq** database, we used all the complete bacterial genomes available (*n* = 8,778) as of January 2018. In the case of GenBank, a total of 153,359 bacterial and 4,053 eukaryotic genomes (3,456 fungal and 597 protozoan genomes) deposited as of August 2018 were considered. Lastly, we surveyed 18,227 MAGs from the largest datasets publicly available as of August 2018[13,16,17,18,19], including those deposited in the Integrated Microbial Genomes and Microbiomes (IMG/M) database[52]. For each database, the **function 'mash sketch'** from **Mash v.2.0**[53] was used to convert the reference genomes into a **MinHash** sketch with default *k*-mer and sketch sizes. Then, the Mash distance between each MAG and the set of references was calculated with **'mash dist'** to find the best match (that is, the reference genome with the lowest Mash distance). Subsequently, each MAG and its closest relative were aligned with **dnadiff v.1.3** from **MUMmer 3.23**[54] to compare each pair of genomes with regard to the fraction of the MAG aligned (aligned query, AQ) and ANI.

(..)

https://doi.org/10.1038/s41586-019-0965-1

Why use workflows?

Automation

- Automate computational aspects
- Repetitive pipelines, sweep campaigns

Scaling – compute cycles

- Make use of computational infrastructure & handle large data

Abstraction – people cycles

- Shield complexity and incompatibilities
- Report, re-use, evolve, share, compare
- Repeat – Tweak - Repeat
- Document and visualize analysis steps

Provenance - reporting

- Capture, report and utilize log and data lineage auto-documentation
- Traceable evolution, audit, transparency
- Compare

Findable

Accessible

Interoperable

Reusable

*(**E**xtensible)*

*(**R**eproducible)*

Adapted from Bertram Ludäscher at WORKS2015 https://www.slideshare.net/ludaesch/works-2015provenancemileage

Actual current practices for workflow reproducibility

Article | Open Access | Published: 11 February 2019

A new genomic blueprint of the human gut microbiota

https://doi.org/10.1038/s41586-019-0965-1

Alexandre Almeida, Alex L. Mitchell, Miguel Boland, Samuel C. Forster, Gregory B. Gloor, Aleksandra Tarkowska, Trevor D. Lawley & Robert D. Finn

Nature **568**, 499–504(2019) | Cite this article

99k Accesses | **132** Citations | **667** Altmetric | Metrics

master | **2** branches | **0** tags | Go to file | Add file | Code

alexmsalmeida Update funcs_phy-assoc_fig5b.R `202c12c` on 4 Dec 2019 **120** commits

R	Update funcs_phy-assoc_fig5b.R	9 months ago
pipelines	Update map2ref.sh	2 years ago
scripts	Update parse_checkm.py	2 years ago
LICENSE	Create LICENSE	2 years ago
README.md	Update README.md	10 months ago

README.md

Analysis of Metagenomic Species (MGS)

Scripts used for characterizing metagenome-assembled genomes (MAGs) used in the following publication:

A Almeida, AL Mitchell, M Boland, SC Forster, GB Gloor, A Tarkowska, TD Lawley and RD Finn (2019) A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499–504

Associated data can also be found in our FTP server.

About

Analysing Metagenomic Species (MGS)

Readme

MIT License

Releases

No releases published

Packages

No packages published

Languages

R 70.7% | Python 18. | Shell 10.8%

```
mkdir "${path}/tmp"

# clean output files before running
rm -rf "${path}/checkm_output"
rm -f "${path}/checkm.log"
rm -f "${path}/bins_"*
rm -f "${path}/marker_file"

# checkm tree
bsub -M 85000 -n 8 -o "${path}/checkm.log" -J "checkm_tree_${path}" "checkm

# checkm tree_qa
bsub -o "${path}/checkm.log" -M 5000 -J checkm_tree_qa_${path} -w "ended(che

# checkm lineage_set
bsub -M 5000 -o "${path}/checkm.log" -J checkm_lineage_set_${path} -w "ended

# checkm analyze
bsub -M 50000 -n 8 -o "${path}/checkm.log" -J checkm_analyze_${path} -w "end

# checkm qa
bsub -M 10000 -o "${path}/checkm.log" -J checkm_qa_${path} -w "ended(checkm_
```

Enabling reproducible, transparent research.

https://www.researchobject.org/

https://doi.org/10.1016/j.future.2011.08.004

Many Objects are the Outcomes of Research

All are first class citizens and are required to make research FAIR+R

Each object has its own metadata and repositories

Scattered Reporting and Reading

Contextualised
Active, Unified
Semantic linking

Buried in a PDF
figure

De-contextualised
Static, Fragmented
Lost Semantic linking
Rebuild to be reproducible

Encapsulated content and references to external resources

integrated view over fragmented resources using **PIDs** and **metadata**

The RO package has its **own metadata**, can be registered and **deposited** in its own right, unpackaged and accessed, activated and reproduced if appropriate

self-describing objects, chiefly **metadata**

PID

RO Metadata file

Structured metadata about the RO and content

type
id
description
datePublished
…

author

organisation

license

RO Content

files

directories

links to web resources

Archive file format / packaging system

BagIt, zip
OCFL, Git

self-describing objects, chiefly **metadata**

PID

RO Metadata file

Structured metadata about the RO and content

type
id
description
datePublished
...

author

organisation

license

Linked Data approach

RO Content

type, id
description
datePublished
creator
size
format ...

image file

directory of data

links to web resources

https://github.com/o/script

https://zenodo.org/record/3541888

Bagit, zip
OCFL, Git

Archive file format / packaging system

RO-Crate Structure

https://w3id.org/ro/crate/1.1

RO-Crate

TABLE OF CONTENTS

1 RO-Crate Metadata File (`ro-crate-metadata.json`)
2 RO-Crate Website (`ro-crate-preview.html` and `ro-crate-preview_files/`)
3 Payload files and directories
4 Self-describing and self-contained

The structure an *RO-Crate* MUST follow is:

```
<RO-Crate root directory>/
|  ro-crate-metadata.json    # RO-Crate Metadata File MUST be present
|  ro-crate-preview.html     # RO-Crate Website homepage MAY be present
|  ro-crate-preview_files/   # MAY be present
|   | [other RO-Crate Website files]
|  [payload files and directories]  # 0 or more
```

JSON with a flat list of:

- **Root Dataset** (the RO-Crate itself)
- **Data** Entities (e.g. files, dirs, DBs)
- **Contextual** Entities (e.g. people)

Objects **reference** each other by `@id`

TABLE OF CONTENTS

1 Contextual vs Data entities
2 Identifiers for contextual entities
3 People
4 Organizations as values
5 Contact information
6 Publications via citation property
7 Publisher
8 Funding and grants
9 Licensing, Access control and copyright
  a Metadata license
10 Extra metadata such as Exif
11 Places
12 Subjects & keywords
13 Time
14 Thumbnails

RO-Crate 1.1
Introduction
Terminology
RO-Crate Structure
Metadata of the RO-Crate
Root Data Entity
Data Entities
Contextual Entities
Provenance of entities
Workflows and scripts
Appendix

RO-Crate

http://www.researchobject.org/ro-crate/

RO-Crate gives **Best Practice** Guidance on existing metadata standards.

Linked Data by stealth
Developer: ”It’s just JSON” (example-driven)

Base vocabulary: schema.org

Extensible: re-use formal ontologies (e.g. FHIR)
..or make ad-hoc terms

Any entity can be described as little or detailed as needed

URI-based identifiers: Click to see more

RO-Crate

RO-Crate 1.1
- Introduction
- Terminology
- RO-Crate Structure
- Metadata of the RO-Crate
- Root Data Entity
- Data Entities
- Contextual Entities
- Provenance of entities
- Workflows and scripts
- Appendix

Background
Community
Examples
Outreach and Publications
Profiles
RO-Crate In Use
Specification
Tools

Publisher

The Root Data Entity SHOULD have a publisher property. This SHOULD be an Organi
MAY be a Person.

```
{
  "@id": "https://doi.org/10.5281/zenodo.1009240",
  "@type": "Dataset",
  "name": "Sample dataset for RO-Crate v0.2",
  "publisher": {
    "@id": "https://ror.org/03f0f6041"
  },
  "temporalCoverage": "2017"
},

{
  "@id": "https://ror.org/03f0f6041",
  "@type": "Organization",
  "identifier": "https://ror.org/03f0f6041",
  "name": "University of Technology Sydney"
},
```

JSON-LD

Standard Web Mark-up

schema.org

Funding and grants

To associate a research project with a Dataset, the *RO-Crate JSON-LD* SHOULD con
the project using type Organization, referenced by a funder property. The project `Org`
SHOULD in turn reference any external funder, either by using its URL as an `@id` or v
Entity describing the funder.

> **TIP**
> To make it very clear where funding is coming from, the *Root Data Entity* SHOULD
> funders directly, as well as via a chain of references.

Summary: RO-Crate in a nutshell

Practical lightweight approach to packaging research **data entities (any object)** with metadata

Aggregate **files** and/or **any URI-addressable content**, with contextual information to aid decisions about re-use:
Who What When Where Why How.

Web Native Machine readable. Human readable. Search engine friendly. Familiar.

Extensible and Incremental: add additional metadata; nested and typed by their profile.

Open Community effort

http://www.researchobject.org/ro-crate/

Tooling

- Describo interactive **desktop application** to create, update and export RO-Crates for different profiles. (~ *beta*)
  - CalcyteJS is a **command-line** tool to help create RO-Crates and HTML-readable rendering (~ *beta*)
  - ro-crate - **JavaScript**/NodeJS library for RO-Crate rendering as HTML. (~ *beta*)
  - ro-crate-js - utility to render HTML from RO-Crate (~ *alpha*)
  - ro-crate-ruby **Ruby** library to consume/produce RO-Crates (~ *alpha*)
  - ro-crate-py **Python library** to consume/produce RO-Crates (~ *planning*)

These applications use or expose RO-Crates:

- Workflow Hub imports and exports Workflow RO-Crates
- OCFL-indexer NodeJS application that walks the Oxford Common File Layout on the file system, validate RO-Crate Metadata Files and parse into objects registered in Elasticsearch. (~ *alpha*)
- ONI indexer
- ocfl-tools
- ocfl-viewer
- Research Object Composer is a REST API for gradually building and depositing Research Objects according to a pre-defined profile. (RO-Crate support *alpha*)

https://www.researchobject.org/ro-crate/implementations.html

User facing

https://uts-eresearch.github.io/describo/

Infrastructure facing

Software libraries

https://www.npmjs.com/package/ro-crate
https://github.com/ResearchObject/ro-crate-ruby
https://pypi.org/project/rocrate/

[Marco La Rosa, Peter Sefton]

PARADISEC

Home About us Our Collections Deposit Resources Support Us Donate

https://mod.paradisec.org.au/

Pacific and Regional Archive for Digital Sources in Endangered Cultures

A race against time to digitise analog records of materials from endangered cultures from all over the world.

Home Advanced Search Transcription Search About

Welcome to the catalog of the **PARADISEC** collection.

View Collection: collection id →

View Item: collection id / item id →

Collections	Items	Contributors	Universities
538	26,755	5,876	66

Simple Search

search both item's and collection's

Name Description Contributor

Subject Language Content Language

keyword search phrase search

Simple wildcard searches are supported. Try adding '*' to match zero or more characters or '?' to match a single character.

PARADISEC has been developed since 2003 by a consortium of Australian researchers from the University of Sydney, the University of Melbourne and the Australian National University. See the PARADISEC website for further details.

PARADISEC (Pacific And Regional Archive for DIgital Sources in Endangered Cultures) curates digital material about small or endangered languages. The catalog entry for an item is usually written by the depositor, and some are more detailed than others. In the case of collections that we have digitised from deceased researchers, we do the best we can to describe the records, but often there is little information available. By placing these items in the collection we hope that other researchers will enrich the descriptions as they use the material. We believe the rights in the material presented in this collection have been cleared by the depositors. Please let us know if you think that is not the case for any particular item.

This installation is a Proof of Concept demonstrating a new generation catalog operating as a single page application interacting with an OCFL on-disk repository. Search capability is provided provided via Elastic Search. Read the about page to find out more.

sample items in this demonstrator

Collections

AC1 NT1 NT3 NT5 NT6 WDVA1 SN1

Items

AC1/001 AC1/002 NT1/005 NT1/20003 NT1/2004

NT1/98001 NT1/98007 NT1/98009 NT1/98017

NT3/nafakoron NT3/SESAKE NT5/TokelauOf NT6/2005V001

PAMBU/DOC1014 SN1/MM20130708Museum WDVA1/MIR_07

WDVA1/TJILK05

Note that only the items listed above have content loaded. For all other items only the metadata is available in this demonstrator.

Cultural Heritage: A data curation service for endangered languages: 500,000 files in 28,624 items and 574 collections

long term preservation and accessibility of research data objects

https://arkisto-platform.github.io/

Standards Based

Scaleable

arkisto

Sustainable

Recordings in South Efate

https://mod.paradisec.org.au/

Open (subject to agreeing to PDSC access conditions)

Item: NT1 / 98007
Collection: NT1
Contributors:

- Nick Thieberger (collector)
- Corin Bone (operator)
- Nick Thieberger (recorder)
- Kalsarap Namaf (speaker)
- Iokopeth (speaker)
- John Maklen (speaker)
- Waia Tenene (speaker)

NT1-98007. Text #047 (speaker is John Maklen. Tex title: Mantu the flying fox and Erromango); Text #049 Kotkot (speaker is John Maklen); Te... show more

hide item metadata

Date Created: 27/09/2012
Last Update: 17/05/2018
Born Digital: No
Cite As:
Dialect:
Digitised On: 31/12/2000
Discourse Type:
External: No
Ingest Notes:
Language As Given: Nafsan
Originated On: 1998-10-03
Originated On Narrative:
Private: No
Received On:

RO-Crate

1:43 / 45:29

NT1-98007-98007A.ixt NT1-98007-98007A.eaf

PLAY (09:22)

Natopu nen kin ito esan ga, ipi natopu nigmam.
The natopu (local spirit) that lives here, it is our natopu.

Natopu	nen	kin	ito	esan	ga,	ipi	natopu	nigmam.
natopu	nen	kin	i= to	esan	ga	i= pi	natopu	nigmam
spirit	that	COMP	3S.RS= stay	place	3S.	3S.RS= be	spirit	1P.POS

Text title: Asarat (speaker is John Makler

Images Audio XML Files

1 2 3 4 5 6 ... 31

Data and Method Thematic Commons

Sharing data, tools and workflows in the cloud

13 EU Life Science Research Infrastructures

https://www.eosc-life.eu/

Workflow + data + provenance interchange, stewardship, recording dependencies -> portability & reuse/reproducibility

RO-Crate as interchange format

WorkflowHub

Download RO Crate

Example of setting up a simulation system Version 2

View on GitHub | Request Contact | Unsubscribe | Download RO Crate

Actions

Work-in-progress

CWL version of the md_list.cwl workflow for HPC.

SEEK ID: https://workflowhub.eu/workflows/98?version=2

Creators and Submitter

Creators

Not specified

Submitter

Douglas Lowe

License

Apache Software License 2.0

Activity

Views: 40 **Downloads:** 0

Created: 29th Jan 2021 at 16:37
Last updated: 29th Jan 2021 at 16:56
Last used: 4th Feb 2021 at 08:12

Tags

This item has not yet been tagged.

Add your tags

Attributions

None

- graph.svg
- md_list.cwl
- ro-crate-metadata.json
- ro-crate-preview.html


```
  "license": "Apache-2.0",
  "identifier": "https://workflowhub.eu/workflows/98?version=2",
  "url": "https://workflowhub.eu/workflows/98/ro_crate?version=2",
  "isBasedOn": "https://github.com/douglowe/biobb_hpc_cwl_md_list
  "sdPublisher": {
    "@id": "https://orcid.org/0000-0002-1248-3594"
  },
  "sdDatePublished": "2021-02-04 11:33:41 +0000",
  "creativeWorkStatus": "Work-in-progress"
},
{
  "@id": "md_list.cwl",
  "@type": [
    "File",
    "SoftwareSourceCode",
    "ComputationalWorkflow"
  ],
  "programmingLanguage": {
    "@id": "#cwl"
  },
  "url": "https://github.com/douglowe/biobb_hpc_cwl_md_list/blob/
  "image": {
    "@id": "graph.svg"
  },
  "contentSize": 10444
```


https://workflowhub.eu/workflows/98

FAIR Digital Objects

Actionable knowledge unit
Digital butterfly – digital twins

courtesy Dimitris Koureas
Coordinator DiSSCo EU Research Infrastructure

Specimen object image
courtesy of Alex Hardisty

https://doi.org/10.3897/rio.6.e54280

Specimen Data Refinery

Workflows to Digitise Natural History Specimens

FAIR Digital Objects -> Packaged + Actionable

SYNTHESYS+ Synthesis of Systematic Resources a DiSSCo project

DiSSCo Distributed System of Scientific Collections

FAIR Digital Object Framework
Open Digital Specimen

+

EOSC-Life

researchobject

elixir

Workflow Infrastructure
RO-Crate

Same conceptual workflow; multiple executable flavours

Workflow RO-Crate

Embedding into Infrastructures & Standards

EGI-ACE data spaces for
Earth Science researchers

Reliance

Research Lifecycle Management technologies for Earth Science Communities and Copernicus users in EOSC

ROHUB

zenodo

OpenAIRE

CS³MESH⁴EOSC

Connecting European Data

FAIR and GDPR compliant data storage and sharing fabric Science Mesh using Cloud Services for Synchronization and Sharing

Standards Based

Scaleable

arkisto

Sustainable

MENDELEY DATA

Seven Bridges

exchange between genomics platform and repository.

BioCompute Objects

FDA

standardise and share analyses generated from genome sequencing.

IEEE 2791-2020

MACHINE-PROCESSABLE

Standards

EXAMPLES

LOW TECH

Incremental

Multi-platform

Graceful degradation

Commodity tooling

Technology Independent

Keep it simple

Developer friendliness

Just enough complexity / standards

- sufficient extra benefits from what already exists…without compromising developer entry-level experience so they do their own thing

Just Enough Linked Data Just In Time

- simplifications instead of generalizations

Retain Linked Data benefits

- querying, graph stores, vocabularies, clickable URIs)

Plus the developer needs

- documentation, examples, libraries, tools, community …

Limited flexibility frees up developers

Familiarity is important for uptake

Join us!

RO-Crate

https://www.researchobject.org/ro-crate/community

Team

The *RO-Crate* team is:

- Peter Sefton https://orcid.org/0000-0002-3545-944X (co-chair)
- Stian Soiland-Reyes https://orcid.org/0000-0001-9842-9718 (co-cha
- Eoghan Ó Carragáin https://orcid.org/0000-0001-8131-2150 (emeritu
- Oscar Corcho https://orcid.org/0000-0002-9260-0753
- Daniel Garijo https://orcid.org/0000-0003-0454-7145
- Raul Palma https://orcid.org/0000-0003-4289-4922
- Frederik Coppens https://orcid.org/0000-0001-6565-5145
- Carole Goble https://orcid.org/0000-0003-1219-2137
- José María Fernández https://orcid.org/0000-0002-4806-5140
- Kyle Chard https://orcid.org/0000-0002-7370-4805
- Jose Manuel Gomez-Perez https://orcid.org/0000-0002-5491-6431
- Michael R Crusoe https://orcid.org/0000-0002-2961-9670
- Ignacio Eguinoa https://orcid.org/0000-0002-6190-122X
- Nick Juty https://orcid.org/0000-0002-2036-8350
- Kristi Holmes https://orcid.org/0000-0001-8420-5254
- Jason A. Clark https://orcid.org/0000-0002-3588-6257
- Salvador Capella-Gutierrez https://orcid.org/0000-0002-0309-604X
- Alasdair J. G. Gray https://orcid.org/0000-0002-5711-4872
- Stuart Owen https://orcid.org/0000-0003-2130-0865
- Alan R Williams https://orcid.org/0000-0003-3156-2105
- Giacomo Tartari https://orcid.org/0000-0003-1130-2154
- Finn Bacall https://orcid.org/0000-0002-0048-3300
- Thomas Thelen https://orcid.org/0000-0002-1756-2128
- Hervé Ménager https://orcid.org/0000-0002-7552-1009
- Laura Rodríguez Navas https://orcid.org/0000-0003-4929-1219
- Paul Walk https://orcid.org/0000-0003-1541-5631
- brandon whitehead https://orcid.org/0000-0002-0337-8610
- Mark Wilkinson https://orcid.org/0000-0001-6960-357X
- Paul Groth https://orcid.org/0000-0003-0183-6910
- Erich Bremer https://orcid.org/0000-0003-0223-1059
- LJ Garcia Castro https://orcid.org/0000-0003-3986-0510
- Karl Sebby https://orcid.org/0000-0001-6022-9825
- Alexander Kanitz https://orcid.org/0000-0002-3468-0652
- Ana Trisovic https://orcid.org/0000-0003-1991-0533
- Gavin Kennedy https://orcid.org/0000-0003-3910-0474
- Mark Graves https://orcid.org/0000-0003-3486-8193
- Jasper Koehorst https://orcid.org/0000-0001-8172-8981
- Simone Leo https://orcid.org/0000-0001-8271-5429
- Marc Portier https://orcid.org/0000-0002-9648-6484
- Paul Brack https://orcid.org/0000-0002-5432-2748
- Milan Ojsteršek https://orcid.org/0000-0003-1743-8300
- Bert Droesbeke https://orcid.org/0000-0003-0522-5674

Meetings

The RO-Crate team try to meet at least monthly in a telcon, see the rolli
in details and minutes.

See also recent and upcoming events.

Mailing list

The RO-Crate Community has a low-traffic public RO-Crate mailing list,
and general enquiries.

Note that most discussion about the RO-Crate specifications typically h
community meetings.

Slack Chat

We have an informal Slack chat channel #ro-crate on seek4science.slac

Code of Conduct

Contributors to the RO-Crate community are expected to comply with c
ensure an open and inclusive environment.

Contributing

To suggest changes, improvements or issues, use the GitHub repository
https://github.com/ResearchObject/ro-crate - if you are new to GitHub c
appreciate the GitHub guides like Hello World, MarkDown and How to c

You are welcome to join us!

Open Source

This specification and documentation is Open Source and licensed as A
see https://www.apache.org/licenses/LICENSE-2.0 for details.