



**CS³
MESH⁴
EOSC**

Connecting European Data



Keeping community-managed data manageable

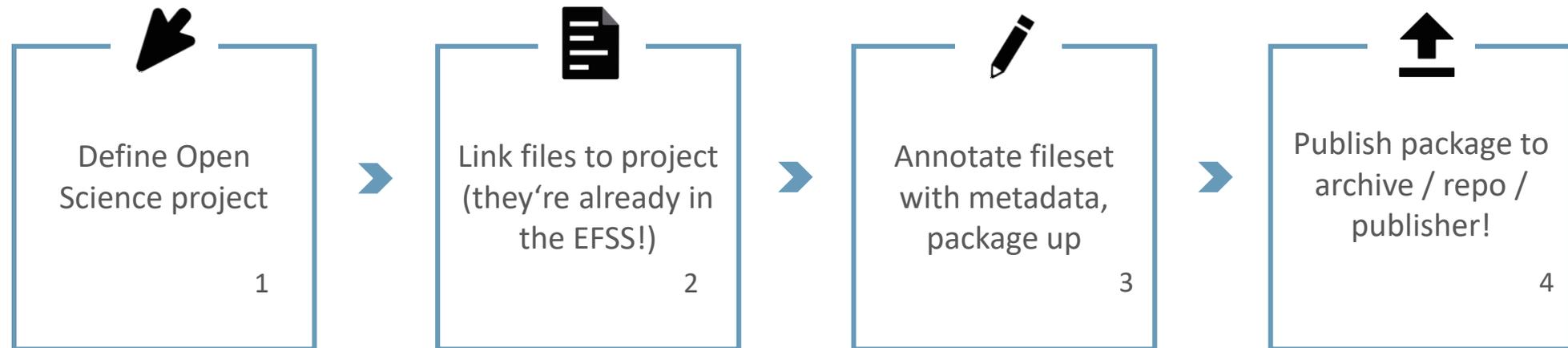
What PARADISEC is teaching CS3MESH4EOSC about user-facing FAIRness

Guido Aben - AARNet eInfrastructure Partnerships Director



CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.

So we have an EFSS full of live research data –
how can we get more value out of that?
FAIR and Open Science seem tantalisingly close!



 OSF **zenodo**

...

SO WE NEED TO READY DATA FOR DEPOSIT –
I.E., “PACKAGE IT UP” AND ANNOTATE DATA



Motivation: package data with maximum useful context

Who ... made it? funded the work?

What ... format are these files? ... is the research about?

Where ... was it collected? ... is it about?

Why ... was it done? ... <link to publication>

How ... were these files created? ... can I repeat that process?

...save future users from guessing...

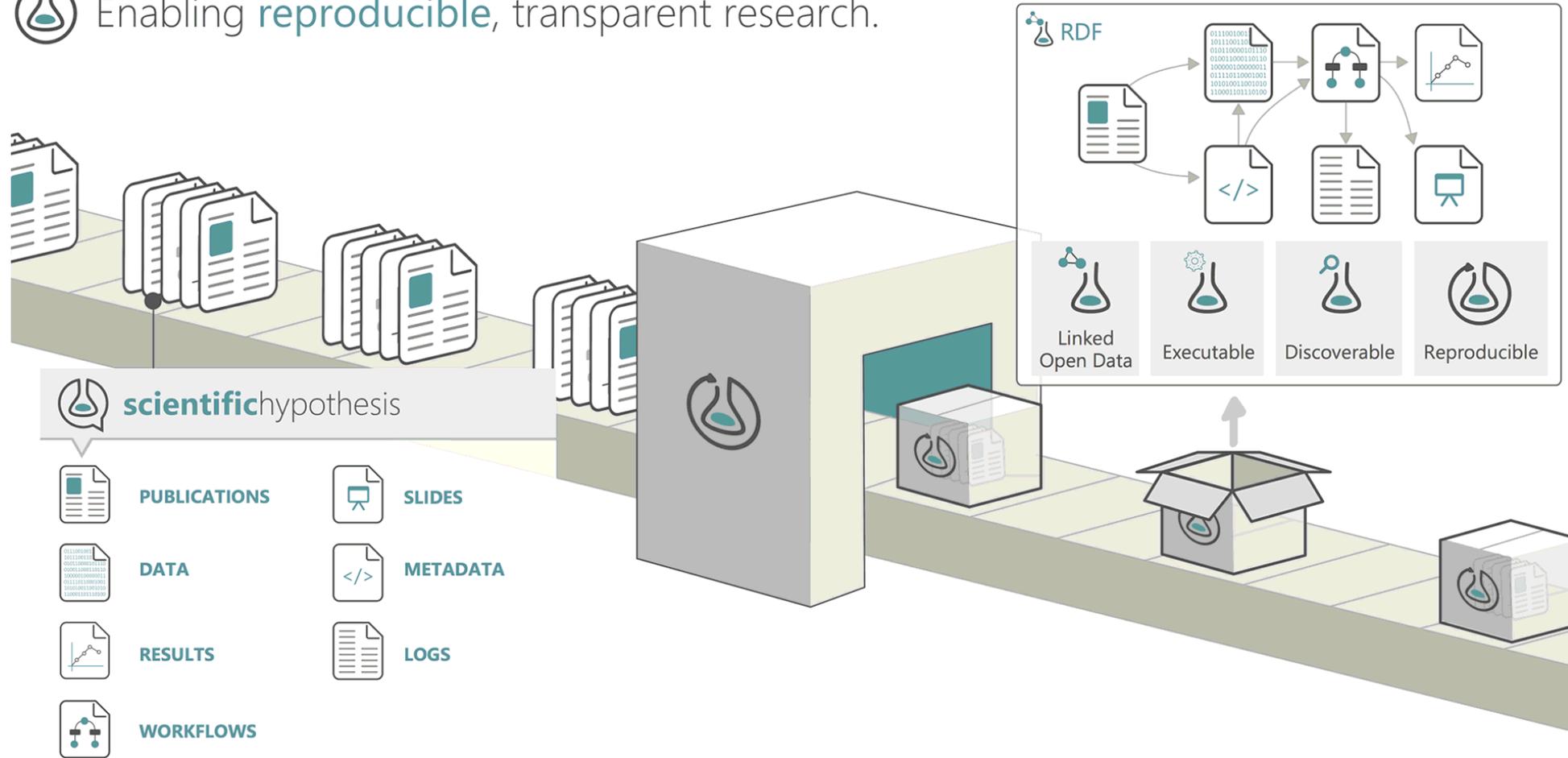


THERE'S A STANDARD FOR THAT! "RO-CRATE"

What is RO-Crate

- Community effort to establish lightweight approach to packaging research data with metadata.
- Schema.org annotations in JSON-LD
- RO-CRATE specification 1.0 released November 2019
- www.researchobject.org
- <https://researchobject.github.io/ro-crate/>

 Enabling **reproducible**, transparent research.



CLEARLY US STORAGE ENGINEERS DIDN'T
COME UP WITH THIS OURSELVES ;)

WE'RE BORROWING FROM AND WORKING
WITH THE FINE PEOPLE AT PARADISEC.



www.paradisec.org.au

NT5-TokelauOf-vid

NT1-98007-98007A



PLAY (16:25)

Ipiatlak, Etasrik, Ratison. Etasrik ipiatlak mtulep iskei itkos. There is (a natopu at) Tassiriki, the Radison. Tassiriki has a woman spirit there.

Ipiatlak, i= piatlak 3S.RS= have	Etasrik, Etasrik p.name	Ratison, Ratison p.name	Etasrik Etasrik p.name	Ipiatlak i= piatlak 3S.RS= have	mtulep mtulep wife
--	-------------------------------	-------------------------------	------------------------------	---------------------------------------	--------------------------

iskei i= skei 3S.RS= one	itkos. i= tok wes 3S.RS= stay there
--------------------------------	---

PLAY (16:36)

Ga me itkos. Nagien ipi Lisau. She is there. Her name is Lisau.

Ga ga	me me	itkos. i= tok wes	Nagien nagi-e -n	ipi i= pi	Lisau. Lisau
----------	----------	----------------------	---------------------	--------------	-----------------

- NT5-TokelauOf-vid.ixt
- NT5-TokelauOf-vid.eopas2.ixt
- NT5-TokelauOf-vid.eopas1.ixt
- NT5-TokelauOf-vid.eaf

07A.ixt NT1-98007-98007A.eaf

o esan ga, ipi natopu nigmam. I spirit) that lives here, it is our natopu.

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

akor runomser mtaki natopu nen kin rusoso ki maarik. akor is scared of the natopu which they call 'maarik' (mister).

riol riol n	ni ni of	Erakor Erakor p.name	runomser ru= nomser 3P.RS= every	mtaki mtak -ki fear -TR	natopu natopu spirit	nen nen that	kin kin REL	rusoso ru= sos -o -ø 3P.RS= call -TS -3S.O	ki ki PREP	maarik. maarik mister
-------------------	----------------	----------------------------	--	-------------------------------	----------------------------	--------------------	-------------------	--	------------------	-----------------------------



Home Advanced Search Transcription Search About

1 2 3 4 5 6 31 1

Images Audio XML Files

NT1-98007-98007A

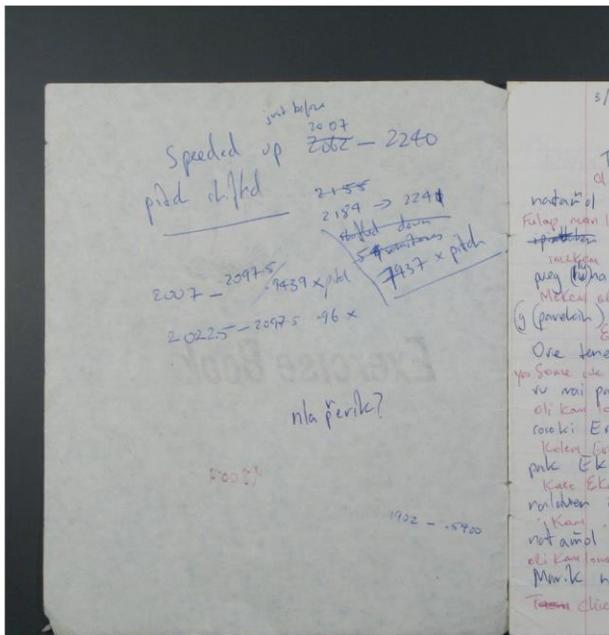


NT1-98007-98007A.ixt NT1-98007-98007A.eaf

PLAY (09:26)

Ser natanioi ni Erakor runomser mtaki natopu nen kin rusoso ki maarik. Everyone from Erakor is scared of the natopu which they call 'maarik' (mister).

Ser ser every	natanioi natanioi person	ni ni of	Erakor Erakor p.name	runomser ru= nomser 3P.RS= every	mtaki mtak -ki fear -TR	natopu natopu spirit	nen nen that	kin kin REL	rusoso ru= sos -o -ø 3P.RS= call -TS -3S.O
ki ki PREP	maarik. maarik mister								



1 2 3 4 5 6

Home Advanced Search Transcription Search About

1 2 3 4 5 6

Images Audio XML Files

NT1-98007-98007A

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

PLAY (09:26)

Ser natanioi ni Erakor runomser mtaki natopu nen kin rusoso ki maarik. Everyone from Erakor is scared of the natopu which they call 'maarik' (mister).

Everyone from Erakor is scared of the natopu which they call 'maarik' (mister).

```

<FORM><AUDIO start="3.992" end="20.202"/></FORM>
<FORM><AUDIO start="20.202" end="23.85"/></FORM>
<FORM><AUDIO start="23.85" end="29.036"/></FORM>
<FORM><AUDIO start="29.036" end="33.801"/></FORM>
<FORM><AUDIO start="33.801" end="42.482"/></FORM>
<FORM><AUDIO start="42.482" end="48.277"/></FORM>
<FORM><AUDIO start="48.277" end="53.786"/></FORM>
<FORM><AUDIO start="53.786" end="62.765"/></FORM>
<FORM><AUDIO start="62.765" end="67.58"/></FORM>
<FORM><AUDIO start="67.58" end="76.076"/></FORM>
<FORM><AUDIO start="76.076" end="84.542"/></FORM>
<FORM><AUDIO start="84.542" end="95.18"/></FORM>
<FORM><AUDIO start="95.18" end="101.74"/></FORM>
<FORM><AUDIO start="101.74" end="114.34"/></FORM>
<FORM><AUDIO start="114.34" end="119.0"/></FORM>
<FORM><AUDIO start="119.0" end="126.58"/></FORM>
<FORM><AUDIO start="126.58" end="133.579"/></FORM>
<FORM><AUDIO start="133.579" end="139.422"/></FORM>
<FORM><AUDIO start="139.422" end="147.507"/></FORM>
<FORM><AUDIO start="147.507" end="155.084"/></FORM>
<FORM><AUDIO start="155.084" end="157.911"/></FORM>
<FORM><AUDIO start="157.911" end="171.409"/></FORM>
<FORM><AUDIO start="171.409" end="185.343"/></FORM>
<FORM><AUDIO start="185.343" end="191.461"/></FORM>
<FORM><AUDIO start="191.461" end="195.845"/></FORM>
<FORM><AUDIO start="195.845" end="211.8"/></FORM>
<FORM><AUDIO start="211.8" end="219.788"/></FORM>
<FORM><AUDIO start="219.788" end="224.07"/></FORM>
<FORM><AUDIO start="224.07" end="232.987"/></FORM>
<FORM><AUDIO start="232.987" end="240.561"/></FORM>
<FORM><AUDIO start="240.561" end="248.959"/></FORM>
<FORM><AUDIO start="248.959" end="255.484"/></FORM>
<FORM><AUDIO start="255.484" end="257.93"/></FORM>
<FORM><AUDIO start="257.93" end="261.03"/></FORM>
<FORM><AUDIO start="261.03" end="264.311"/></FORM>
<FORM><AUDIO start="264.311" end="270.181"/></FORM>
<FORM><AUDIO start="270.181" end="281.638"/></FORM>
<FORM><AUDIO start="281.638" end="290.507"/></FORM>
<FORM><AUDIO start="290.507" end="300.547"/></FORM>
<FORM><AUDIO start="300.547" end="324.679"/></FORM>
<FORM><AUDIO start="324.679" end="335.473"/></FORM>
<FORM><AUDIO start="335.473" end="342.75"/></FORM>
<FORM><AUDIO start="342.75" end="348.013"/></FORM>
<FORM><AUDIO start="348.013" end="356.797"/></FORM>
<FORM><AUDIO start="356.797" end="360.307"/></FORM>
<FORM><AUDIO start="360.307" end="364.723"/></FORM>

```



Home
Advanced Search
Transcription Search
About

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

Collections
538

Items
14,815

Contributors
3,352

Universities
54

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.

View Collection: →

View Item: / →

Simple Search

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.

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/TJLK05

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

Recently Updated Content

Fijian conversation

Fj (speed 7 1/2) side 1 (15:30) "S nikau & Tinapati: Conversation S nikau: Narrative" side 2 (5:54) "S nikau: short text and Texts, Milner 1 & 2 page 126"

Mwakai (aka Mongol) language recordings

Recordings of Mwakai stories

Paama

Paama recording

Vivinei di ari ka ngeta tatamana

A story recorded in a fieldmethods class, told by Glorious Oxenham.

laai story

Motu Field Methods 2018 ANU

These recordings were collected during a 4 week intensive field methods class held at The Australian National University. The speakers were Deveni Temu and Erue Stephens.

Health Check
CoEDL Labs

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

Must match

efate

Name

and

item

Type

+

Should match

+

Must not match

+

Search

Total 17 1 2

1. Item

Efate dialects

2 page typed manuscript. -- Efate dialects. Erakor. English Listings showing verb and noun corresponding. Unidentified informant. English definitions in Capell's hand.; Date of recording unknown.

</view/AC2/VEFAT401>

2. Item

Spirax notebook, N.H.I. (Efate).

Manuscript, mostly pencil. In Capell's text (fading text). i. Tongoa story (pp 1-3) ii. Makura 2nd Reader (4-17) iii. Emae 2nd Reader (50-57) [Stories, notes etc, blank pages 18-49; pages 58-69 blank] ...

</view/AC2/VFN14>

3. Item

Home Advanced Search Transcription Search

About



search

keyword search phrase search

Total 10000 1 2 3 4 5 6 1000

1. bee DET 3sg-hang

</view/EAG1/S7?transcription=EAG1-S7-wad20110712IMKfrogstory.eaf&begin=288.03&end=291.11>

2. place DET 3sg-empty

</view/EAG1/S7?transcription=EAG1-S7-wad20110712IMKfrogstory.eaf&begin=292.19&end=293.47>

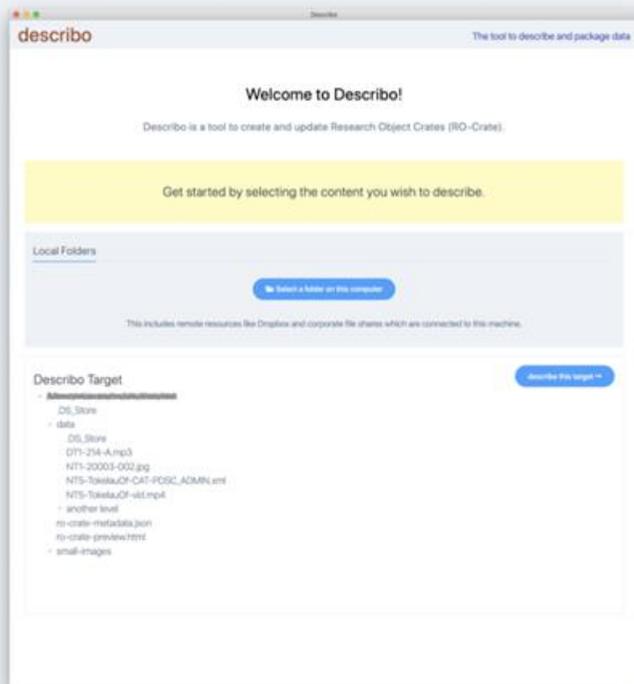
3. not.exist

</view/EAG1/S7?transcription=EAG1-S7-wad20110712IMKfrogstory.eaf&begin=293.9&end=294.61>

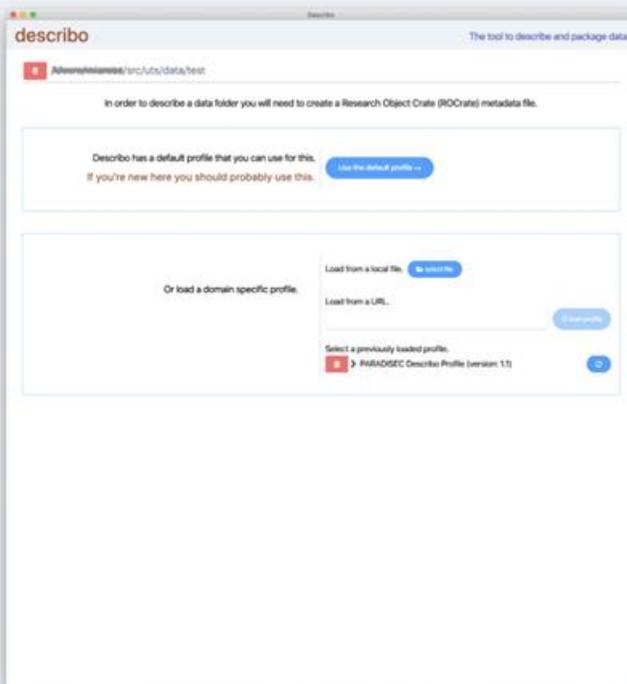
Describo

The tool to describe and package data.

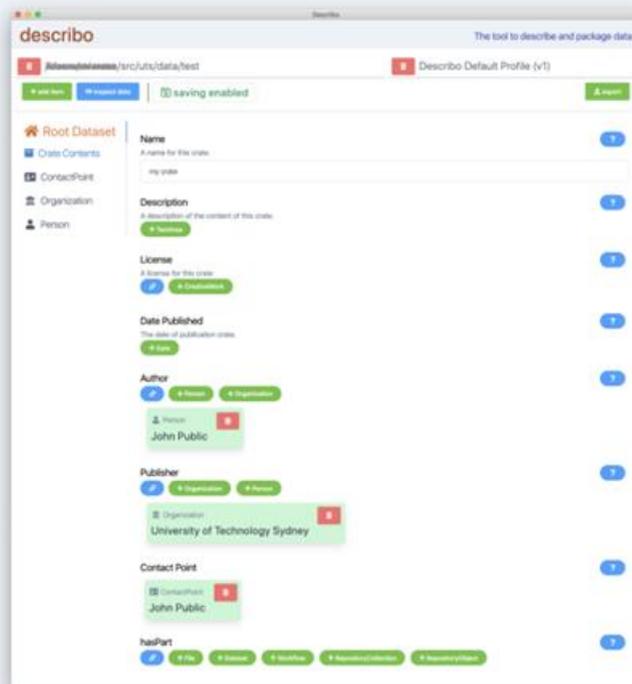
Describo is a tool to create and update **Research Object Crates (RO-Crate)**



Select a folder to describe.



Select a profile.



Describe your data.

```

▼ metadata:
  name: "PARADISEC Describo Profile"
  version: 1.1
▼ items:
  ▼ Collection:
    ▼ metadata:
      about: "A profile to describe a PARADISEC collection."
      allowAdditionalProperties: false
    ▶ inputs: [-]
  ▼ Item:
    ▼ metadata:
      about: "A profile to describe a PARADISEC item."
      allowAdditionalProperties: false
    ▶ inputs: [-]
  ▼ typeDefinitions:
    ▶ License: {-}
    ▶ Language: {-}
    ▶ enabledCoreTypes: [-]
    ▶ dataPacks: [-]

```

Define your own profile. Describo is extensible in every way including overriding inbuilt types and defining which types are available to the user.

Author

+ Person + Organization

Person John Public

Entity type definitions like Person, Organization and many others.

Publisher

+ Organization + Person

Organization University of Technology Sydney

Create and attach entities to properties.

Name

Access for this crate

my name

Built in types - collect text input from a user.

Description

A description of the content of this crate

my description

Built in types - collect text input from a user using a text area. Great for descriptions!

Edit: Language -

Lookup data of type Language

English

- Antigua and Barbuda Creole English
- Bahamas Creole English
- Belize Kriol English
- Chinese Pidgin English
- English
- Ghanaian Pidgin English
- Grenadian Creole English
- Guyanese Creole English

Load data packs and allow the user to select a pre-defined JSON-LD object

Date Published

The date of publication crate.

+ Date

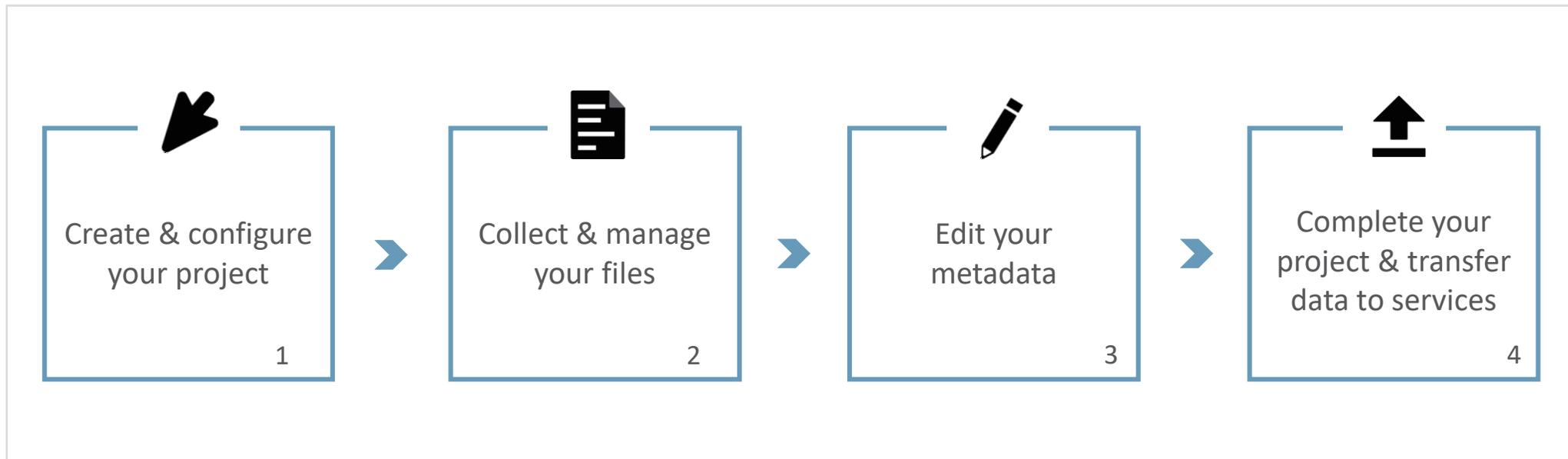
Pick a date

<< 2020 May >>

Sun	Mon	Tue	Wed	Thu	Fri	Sat
26	27	28	29	30	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31	1	2	3	4	5	6

Built in types - collect an ISO date from a user.

And finally, adding a researcher-facing workflow UX that allows them to pick, pack, and publish in one go: ScieboRDS





**CS³
MESH⁴
EOSC**

Connecting European Data

Thank you!
Discover more on...

 [cs3mesh4eosc.eu](https://www.cs3mesh4eosc.eu)

 [company/cs3mesh4eosc](https://www.linkedin.com/company/cs3mesh4eosc)

 [CS3org](https://twitter.com/CS3org)

 [CS3MESH4EOSC Project](https://www.youtube.com/channel/UCHKcZEKmqXjCvc3MLFjFxbw)

<https://www.youtube.com/channel/UCHKcZEKmqXjCvc3MLFjFxbw>



CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.