



# Connecting European Data

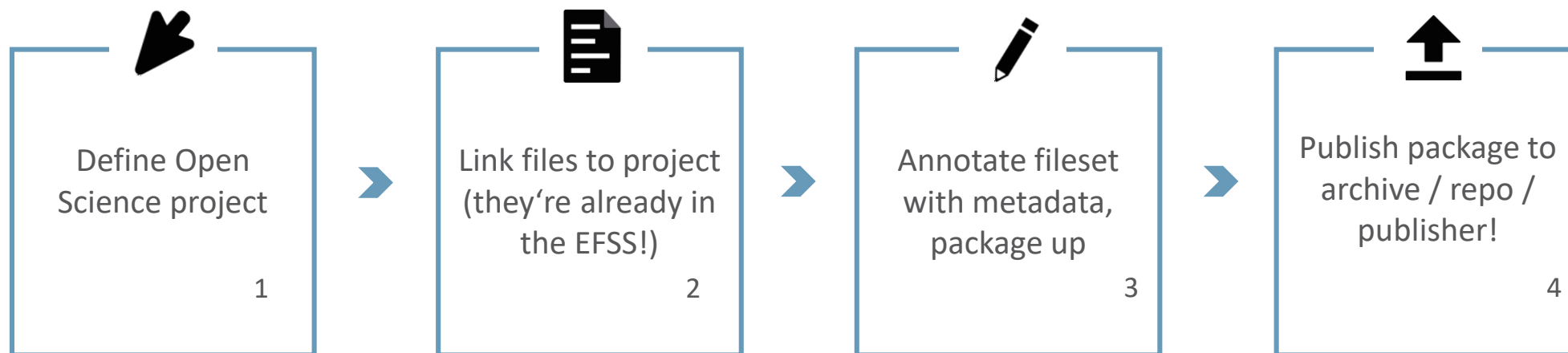


# What PARADISEC is teaching CS3MESH4EOSC about user-facing FAIRness



**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...

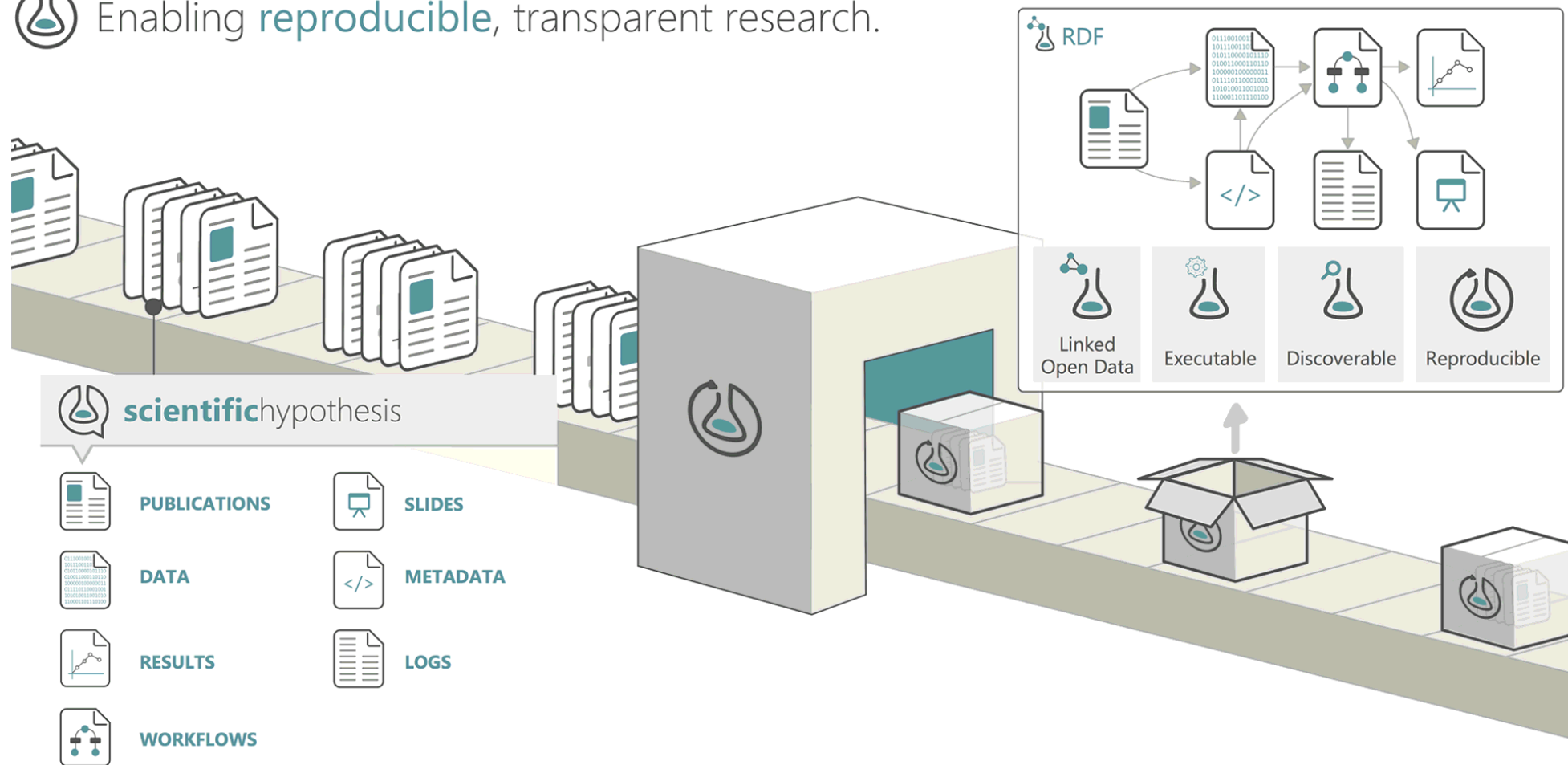


THERE'S A STANDARD FOR THAT! “RO-CRATE”

## 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](http://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](http://www.paradisec.org.au)



NT5-TokelauOf-vid



- ☐ NT5-TokelauOf-vid.txt
- ☒ NT5-TokelauOf-vid.eopas2.txt
- ☐ NT5-TokelauOf-vid.eopas1.txt
- ☐ NT5-TokelauOf-vid.eaf

PLAY (16:26)

Ipiatlak, Etasrik, Ratison. Etasrik ipiatlak mtulep iskei itkos.

There is (a natupu at) Tassiriki, the Radison. Tassiriki has a woman spirit there.

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

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

PLAY (16:36)

Ga me itkos. Nagien ipi Lisau.

She is there. Her name is Lisau.

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

NT1-98007-98007A

07A.txt

NT1-98007-98007A.eaf

o esan ga, ipi natupu nigmam.

I spirit) that lives here, it is our natupu.

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

akor runomser mtaki natupu nen kin rusoso ki maarik.

akor is scared of the natupu which they call 'maarik' (mister).

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



1 2 3 4 5 6 31 1

Home Advanced Search Transcription Search

Home Search Tools

Images Audio XML Files

NT1-98007-98007A

PLAY (00:36)

NT1-98007-98007A.txt

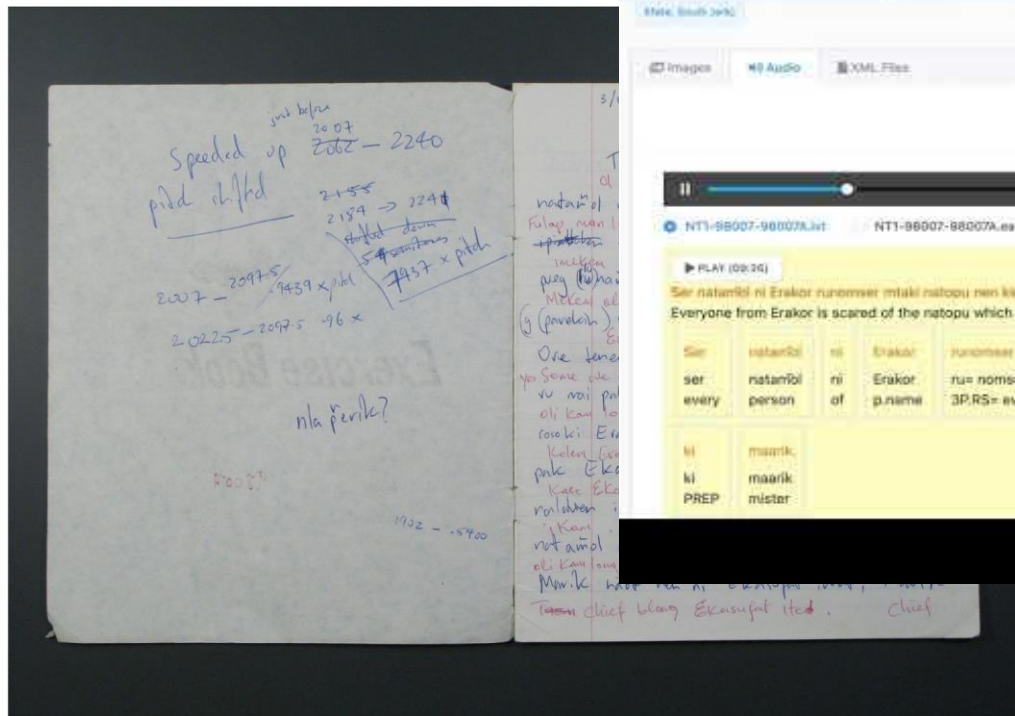
NT1-98007-98007A.eaf

Ser natamoli ni Erakor runomser mtaki natupu nen kin rusoso ki maarik.

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

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

ki	maarik.
ki	maarik
PREP	mister



1 2 3 4 5 6

859-1" ?&gt;100TYPE TEXT PUBLIC "-//ONRS-LACITO//DTD ArchiPage//EN" "http://lacito.Archivage.vjf.cnrs.fr/archives/AL/ERK"&gt;&lt;HEADER&gt;&lt;TITLE&gt;&lt;/TITLE&gt;&lt;SOUNDFILE href="NT1-98007-A"/&gt;

end="3.992"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="3.992" end="20.202"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="20.202" end="23.85"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="23.85" end="29.036"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="29.036" end="33.801"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="33.801" end="42.482"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="42.482" end="48.277"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="48.277" end="53.786"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="53.786" end="62.765"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="62.765" end="67.58"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="67.58" end="76.076"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="76.076" end="84.542"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="84.542" end="95.18"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="95.18" end="101.74"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="101.74" end="114.34"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="114.34" end="119.0"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="119.0" end="126.58"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="126.58" end="133.579"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="133.579" end="139.422"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="139.422" end="147.508"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="147.508" end="155.084"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="155.084" end="157.911"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="157.911" end="171.409"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="171.409" end="185.343"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="185.343" end="191.461"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="191.461" end="195.845"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="195.845" end="211.8"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="211.8" end="219.788"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="219.788" end="224.07"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="224.07" end="232.987"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="232.987" end="240.561"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="240.561" end="248.959"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="248.959" end="255.484"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="255.484" end="257.93"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="257.93" end="261.03"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="261.03" end="264.311"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="264.311" end="270.181"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="270.181" end="281.636"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="281.636" end="290.507"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="290.507" end="300.547"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="300.547" end="324.679"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="324.679" end="335.473"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="335.473" end="342.75"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="342.75" end="348.013"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="348.013" end="356.797"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="356.797" end="360.387"/&gt;&lt;/S&gt;

&lt;/FORM&gt;&lt;AUDIO start="360.387" end="364.723"/&gt;&lt;/S&gt;



[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/SESASE 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.

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>

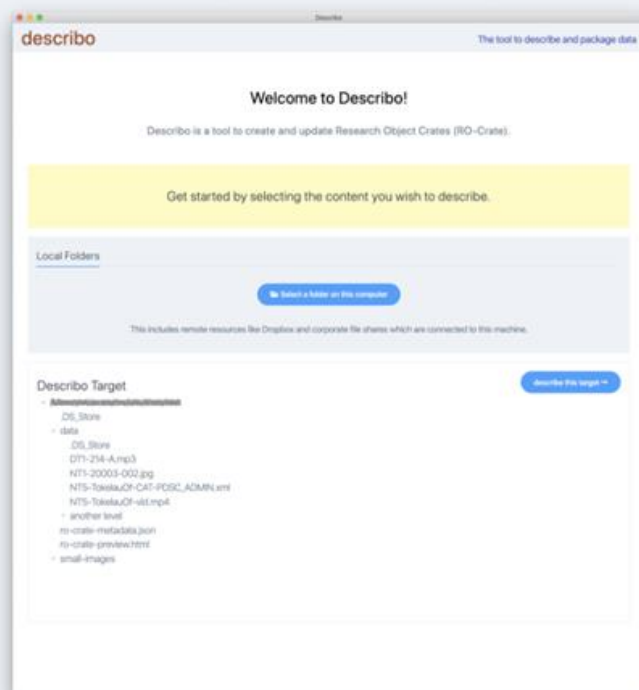
### 3. not.exist

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

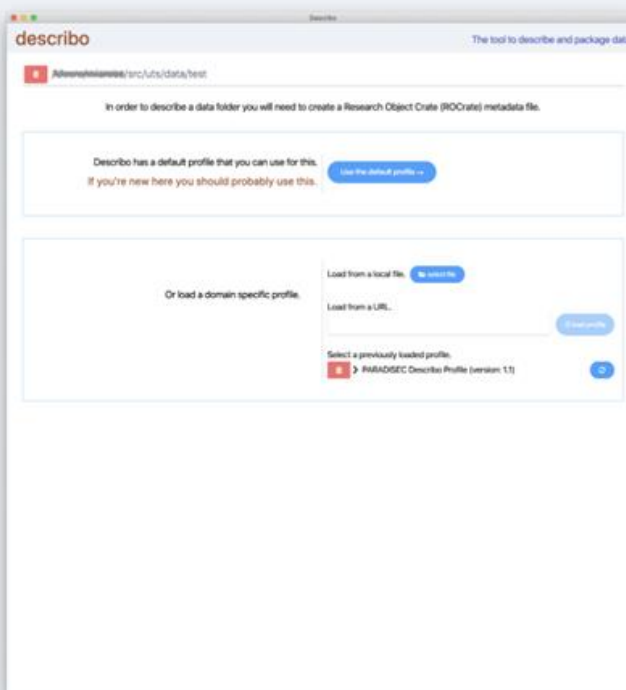
# Describo

The tool to describe and package data.

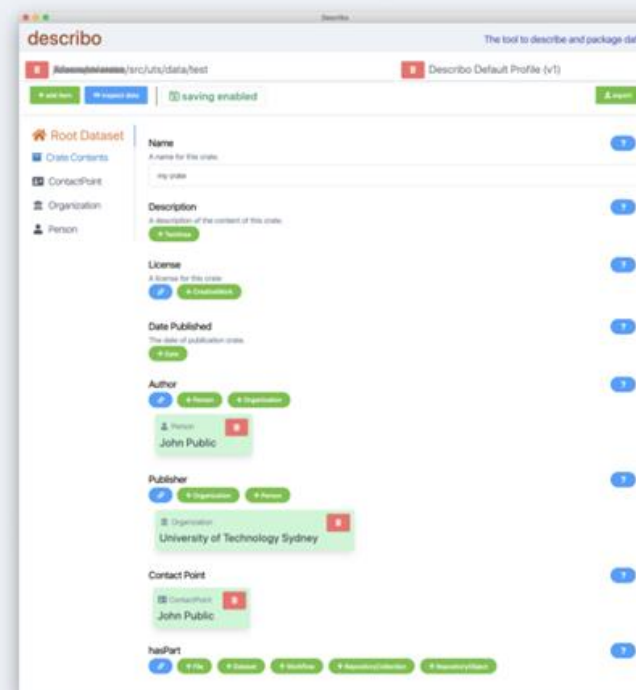
Describo is a tool to create and update **Research Object Crates (RO-Crate)**



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.

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

Name  
A name for this crate  
my name

Built in types - collect text input from a user.

Description  
A description of the content of this crate  
+ Add new

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

Date Published

The date of publication crate.

+ Date

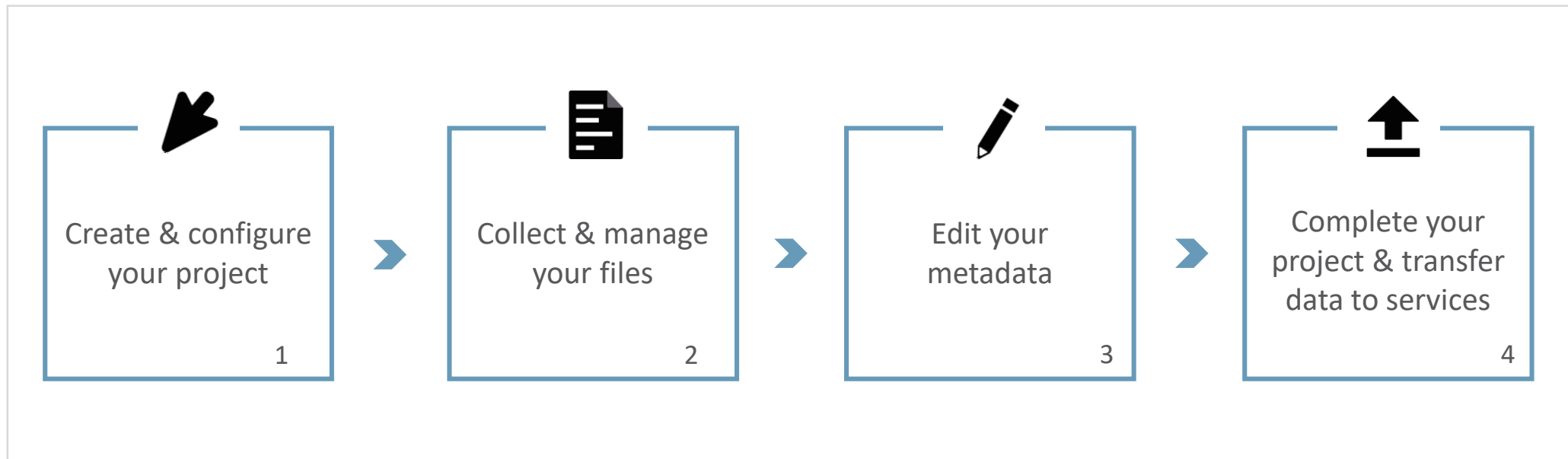
Pick a date

<< 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



...



**Thank you!**  
Discover more on...

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

 [company/cs3mesh4eosc](https://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**.