

Data FAIRness through Open Science (Proposal)

S. Bertocco, M. Molinaro, G. Taffoni

Groningen, 11-13 October 2019

Summary

- DOI Concepts Summary
- Work done & available services using DOIs
- How we exploit DOIs now
- How we could exploit DOIs in the future

“In computing, a *digital object identifier (DOI)* is a persistent identifier or handle used to *identify objects uniquely*, *standardized* by the International Organization for Standardization (ISO).”

“A DOI aims to be “*resolvable*”, usually to some form of access to the information object to which the DOI refers.

This is achieved by *binding the DOI to metadata* about the object, such as a URL, indicating where the object can be found.”

https://en.wikipedia.org/wiki/Digital_object_identifier



“The DOI for a document remains fixed over the lifetime of the document, whereas its location and other metadata may change. Referring to an online document by its DOI is supposed to provide a more stable link than simply using its URL. But every time a URL changes, the publisher has to update the metadata for the DOI to link to the new URL.”

https://en.wikipedia.org/wiki/Digital_object_identifier



DOIs are in wide use mainly to identify academic, professional, and government information, such as

- *journal articles*
- *research reports*
- *official publications*
- *data sets*

https://en.wikipedia.org/wiki/Digital_object_identifier

CADC DOI Service

GET	/instances	Get a list of DOI statuses
POST	/instances	Initialize a new DOI
DELETE	/instances/{DOI Num}	Delete a DOI instance
GET	/instances/{DOI Num}	Get a DOI instance
POST	/instances/{DOI Num}	Update a DOI instance
POST	/instances/{DOI Num}/mint	Finalize a DOI instance
GET	/instances/{DOI Num}/status	Get the status of a DOI instance
GET	/instances/{DOI Num}/status/public	Get the status of a DOI instance

 At the bottom, there is a section for 'Support Interfaces' with options to 'Show/Hide', 'List Operations', and 'Expand Operations'." data-bbox="175 138 853 901"/>

Swagger UI - Google Chrome <3>

Swagger UI

canfar.phys.uvic.ca/doi/

Digital Object Identifier (DOI) API Doc

CADC DOI Service

The Digital Object Identifier (DOI) Service supports initializing and minting DOIs using DataCite.org.

instances : Main endpoint for the DOI service. Show/Hide | List Operations | Expand Operations

GET	/instances	Get a list of DOI statuses
POST	/instances	Initialize a new DOI
DELETE	/instances/{DOI Num}	Delete a DOI instance
GET	/instances/{DOI Num}	Get a DOI instance
POST	/instances/{DOI Num}	Update a DOI instance
POST	/instances/{DOI Num}/mint	Finalize a DOI instance
GET	/instances/{DOI Num}/status	Get the status of a DOI instance
GET	/instances/{DOI Num}/status/public	Get the status of a DOI instance

Support Interfaces Show/Hide | List Operations | Expand Operations



Object Identifier Servi... x +

cds.nccs.nasa.gov/tools-services/digital-object-identifier-services/

LRZ: How to set u... Previsioni meteo... PostaCertificata concorsi attivi Cloud - Introduzi... Standard Deviat... Federated AAI Co... Astrochannel 2.0 La Vida Es Un Car... DAHUA HACKED:...

Other bookmarks

National Aeronautics and Space Administration
Goddard Space Flight Center

[NCCS](#) | [CISTO](#) | [Sciences and Exploration](#)

Advancing Research and Applications with NASA Climate Model Data

HOME DATA TOOLS & SERVICES NEWS ABOUT

Digital Object Identifier Services (DOI)

Digital Object Identifiers (DOI's) are used to provide a permanent and unique digital identification for any object, digital or physical. The science community has long recognized the importance of citing data in published literature to encourage replication of experiments and verification of results. Authors who try to cite their data often find that publishers do not accept Internet addresses. These addresses are viewed as transient references, frequently changed by the data provider after the paper is published. Digital Object Identifiers (DOIs) were created to avoid this problem by providing a unique and persistent reference to online data. An indicator of the value of DOIs is the fact that they have emerged as the most accepted data identifier in the publishing community.

CDS provides the ability to generate or "mint" a Digital Object Identifier for NASA climate data sets. CDS will also create and maintain a Landing Page that

Object Identifier Services (DOI) | CDS - Google Chrome

Object Identifier Services x +

cds.nccs.nasa.gov/tools-services/digital-object-identifier-services/

LRZ: How to set u... Previsioni meteo... PostaCertificata concorsi attivi Cloud - Introduzi... Standard Deviati... Federated AAI Co... Astrochannel 2.0 La Vida Es Un Car... DAHUA HACKED:...

Other bookmarks

What information is required for the Landing Page?

The Landing Page Meta data could include ([Example CDS DOI landing page](#)):

- A brief description of the dataset.
- How to access the data.
- Data provider contact information.
- Summary, to include a short name, version, data format, spatial and temporal coverage, resolution (latitude, longitude and temporal), and dataset size.
- A link to documentation.
- A description of the variables.

What happens if a DOI Datasets needs to be changed?

Occasionally errors in datasets are discovered and the data must be updated. In this case, a new DOI must be issued to be in compliance with the DataCite standard. CDS will work with the data provider to generate a new DOI and update the Landing Page.

To obtain your DOI or if you have additional questions contact: support@cds.nasa.gov

Share Like 44 Share Tweet

ESCAPE

DOI to identify Data Sets

DOI used to identify datasets is useful for:

- Data citation
- Data Provenance

DOI to identify Software



The screenshot shows a web browser window displaying a GitHub Guide. The browser's address bar shows the URL <https://guides.github.com/activities/citable-code/>. The page title is "Making Your Code Citable" and it is estimated to be a 10-minute read. The main content area contains a paragraph explaining that Digital Object Identifiers (DOI) are the backbone of the academic reference and metrics system. It states that if you're a researcher writing software, this guide will show you how to make the work you share on GitHub citable by archiving one of your GitHub repositories and assigning a DOI with the data archiving tool [Zenodo](#). A "ProTip" at the bottom indicates the tutorial is aimed at researchers who want to

Intro

- [Choosing Your Repo](#)
- [Login to Zenodo](#)
- [Check Repo Settings](#)
- [Create a New Release](#)
- [Minting a DOI](#)
- [Finishing up](#)

DOI to identify Software

Making Your Code Citable · GitHub Guides · Google Chrome

guides.github.com/activities/citable-code/#create

Minting a DOI

Before Zenodo can issue a DOI for your repository, you will need to provide some information about the GitHub repo that you've just archived.

Once you're happy with the description of your software, click the **Publish** button at the bottom of the Zenodo form, and voilà, you've just made a new DOI for your GitHub repository!

Intro

Choosing Your Repo

Login to Zenodo

Check Repo Settings

Create a New Release

Minting a DOI

Finishing up

Discard changes Save Publish

Edit upload

Instructions: (i) Upload minimum one file or fill-in required fields (marked with a red star). (ii) Press "Save" to save your upload for editing later. (iii) When ready, press "Publish" to finalize and make your upload public.

Filename (1 files)	Size	Checksum
arfonsmith/My-Awesome-Science-Software	7 Kb	md5b7d876d3e90d3c8649e3cc3f1d1d

Note: File addition, removal or modification are not allowed after an upload has been published. This is because a Digital Object Identifier (DOI) is registered with CrossRef DataCite for each upload. If you made a mistake please contact us.

Upload type required

Publication

Poster

Presentation

Dataset

Image

Video/Audio

Software

Lesson

Basic information required

Digital Object Identifier 10.5281/zenodo.19630

Optional: Did your publisher already assign a DOI to your upload? If not, leave the field empty and we will register a new DOI for you. A DOI allows others to easily and unambiguously cite your upload.



Zenodo Versioning

Making Your Code Citable · GitHub Guides · Google Chrome

guides.github.com/activities/citable-code/#create

Create a new release

By default, Zenodo takes an archive of your GitHub repository each time you create a new Release. To test this out, head back to the main repository view and click on the **releases** header item.

1 branch 0 releases 2 contributors

Create new file Upload files

Initial commit

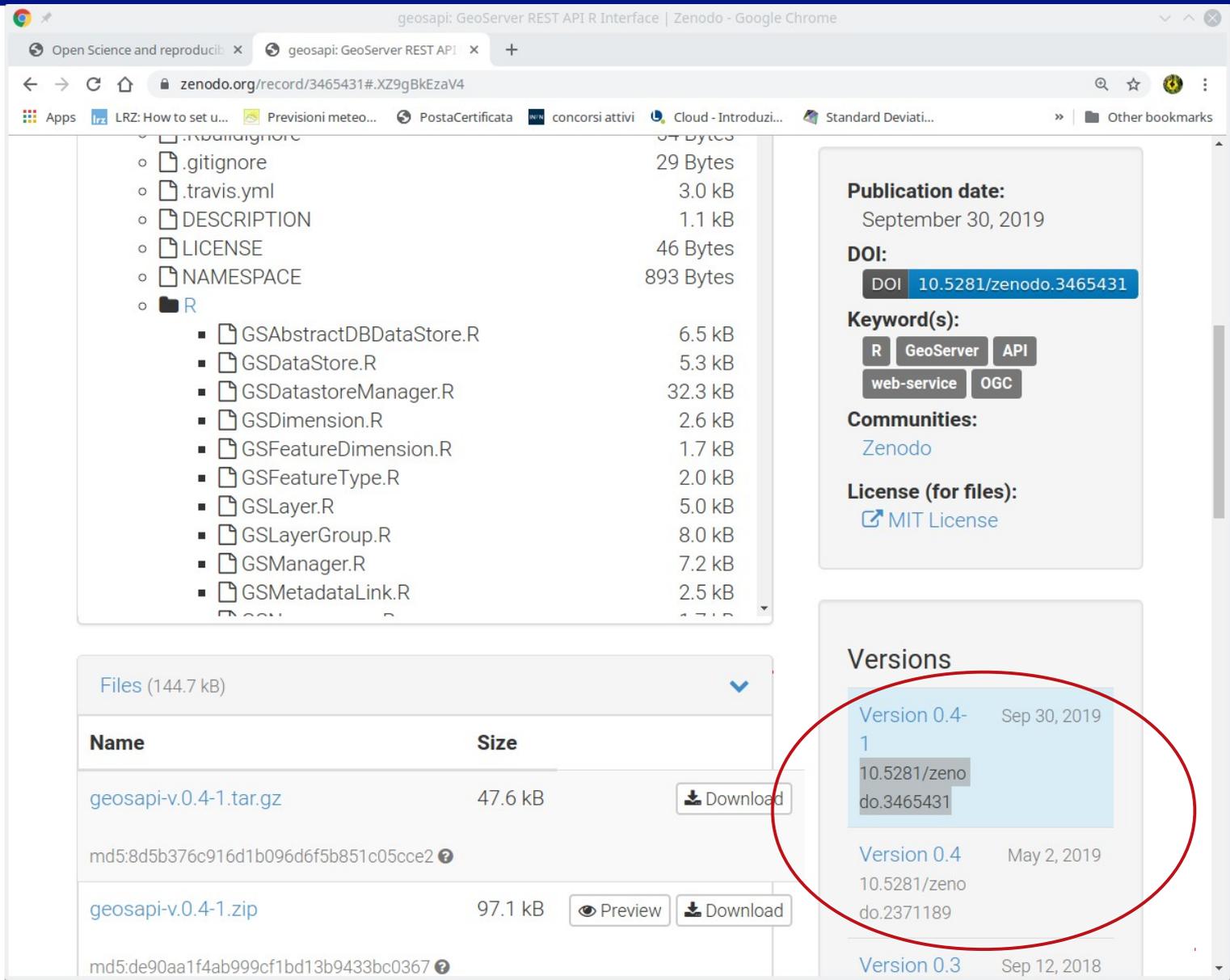
Update README.md

- Intro
- Choosing Your Repo
- Login to Zenodo
- Check Repo Settings
- Create a New Release**
- Minting a DOI
- Finishing up

The screenshot shows a Zenodo record page for the software package 'geosapi: GeoServer REST API R Interface'. The page is displayed in a Google Chrome browser window. The Zenodo logo is at the top left, followed by a search bar and 'Upload' and 'Communities' links. 'Log in' and 'Sign up' buttons are on the right. The record details include the date 'September 30, 2019', 'Software' and 'Open Access' tags, the title 'geosapi: GeoServer REST API R Interface', the author 'Emmanuel Blondel', and the URL 'https://cran.r-project.org/package=geosapi'. A file preview section shows a zip file 'geosapi-v.0.4-1.zip' and its contents: '.Rbuildignore' (54 Bytes), '.gitignore' (29 Bytes), and '.travis.yml' (3.0 kB). On the right, statistics show 3,189 views and 286 downloads, with a 'See more details...' link. Below this, it is noted as 'Indexed in OpenAIRE'. A 'Publication date:' field is also visible.



Software Versioning & DOI



geosapi: GeoServer REST API R Interface | Zenodo - Google Chrome
 zenodo.org/record/3465431#.XZ9gBkEzav4

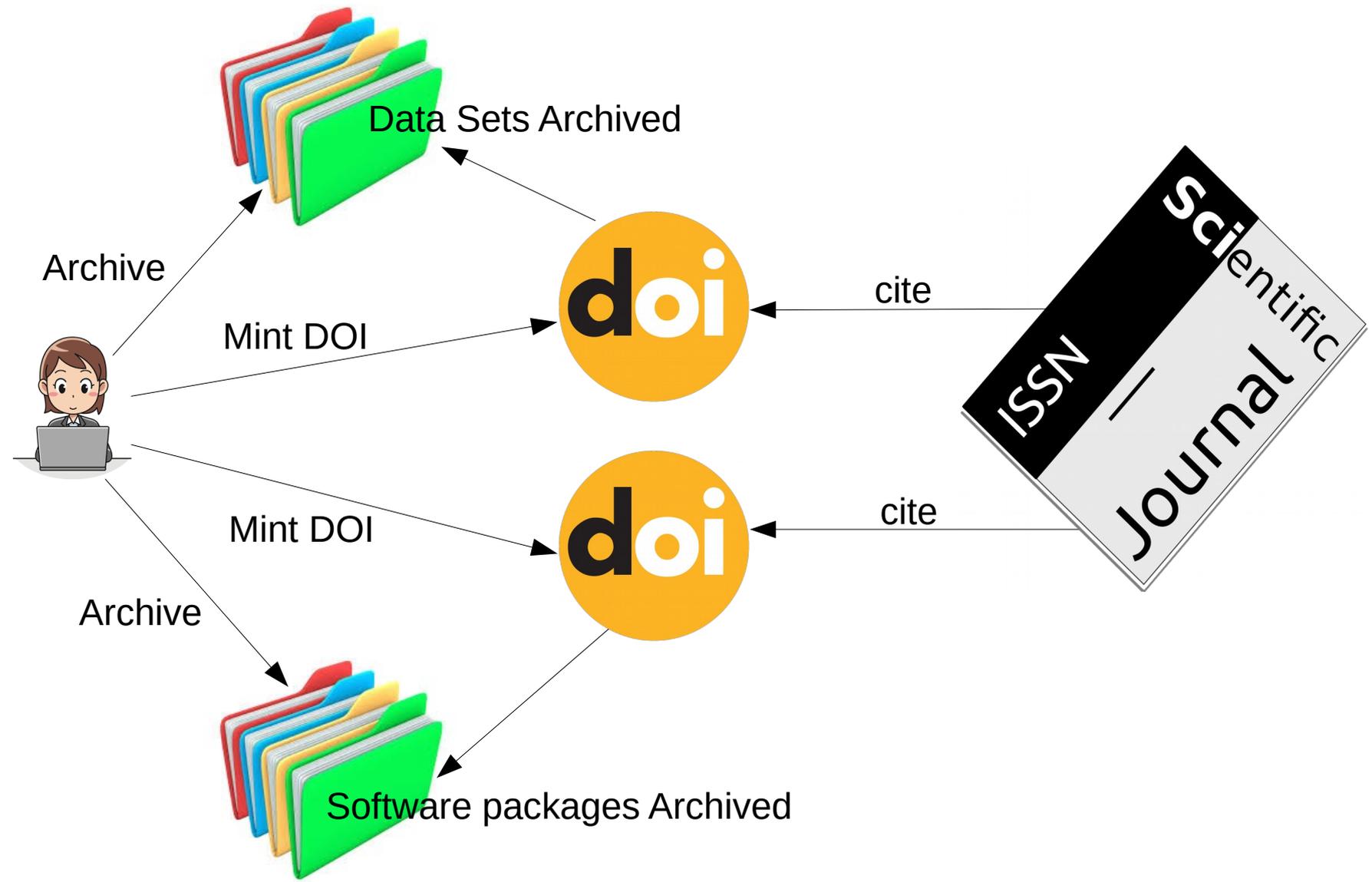
Publication date: September 30, 2019
DOI: DOI 10.5281/zenodo.3465431
Keyword(s): R, GeoServer, API, web-service, OGC
Communities: Zenodo
License (for files): MIT License

Name	Size
geosapi-v.0.4-1.tar.gz	47.6 kB
geosapi-v.0.4-1.zip	97.1 kB

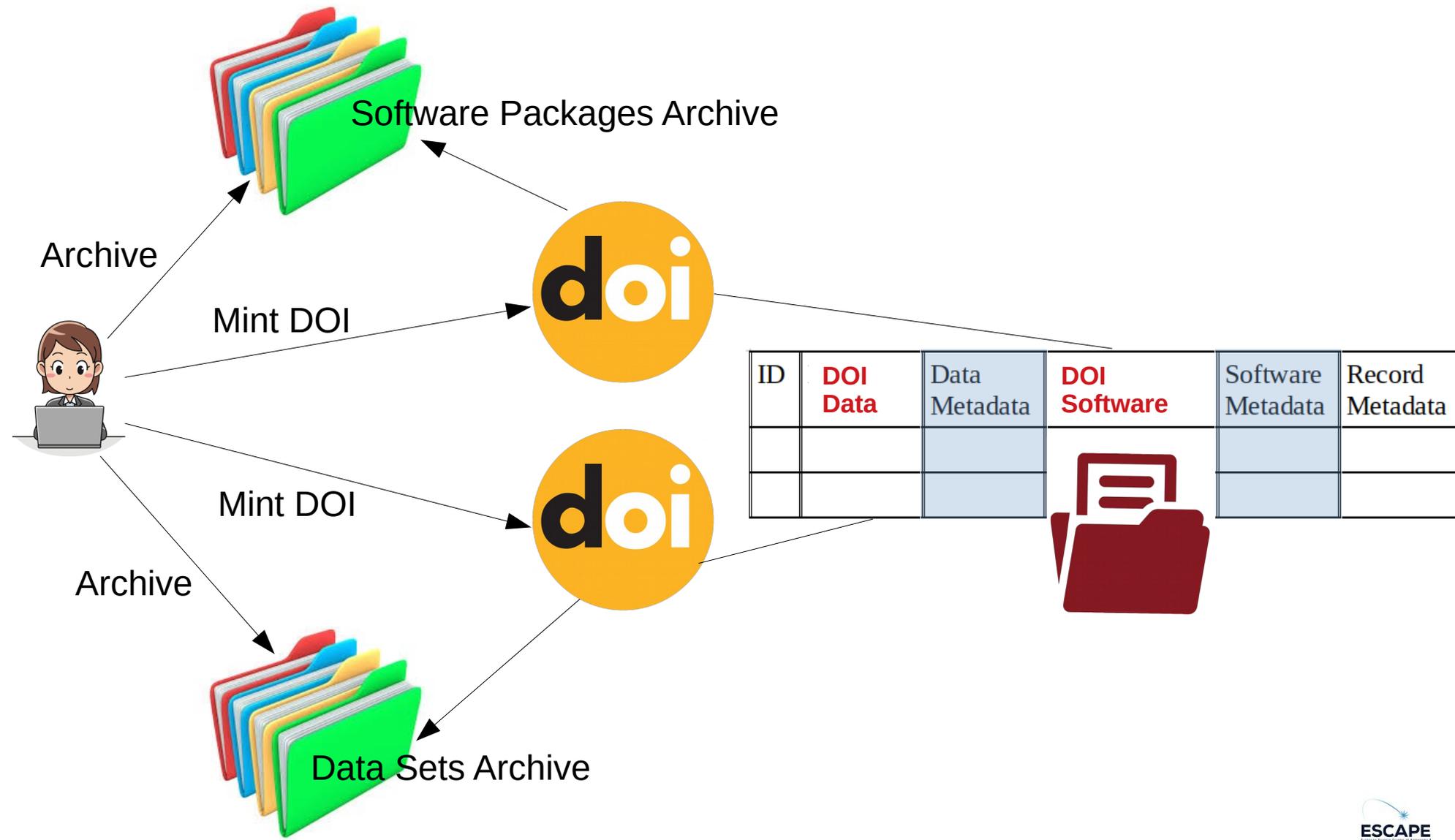
Versions
 Version 0.4-1 Sep 30, 2019
 10.5281/zenodo.3465431
 Version 0.4 May 2, 2019
 10.5281/zenodo.2371189
 Version 0.3 Sep 12, 2018



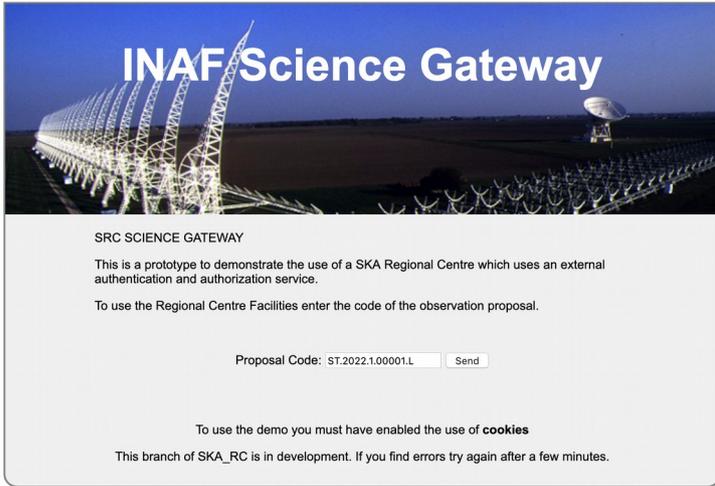
Scientific citations



Link Data & Software



Reproducibility



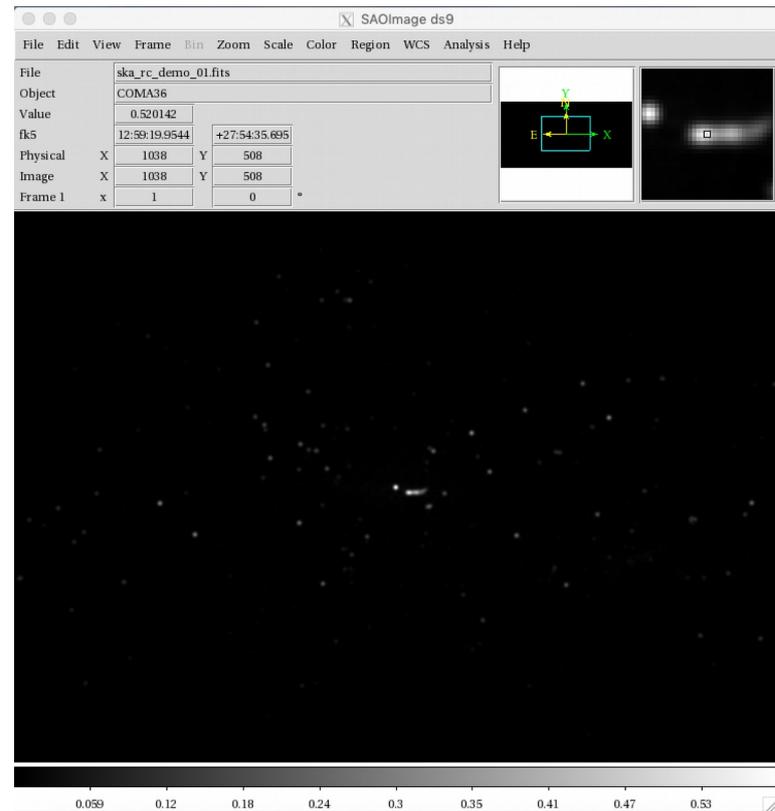
Query

ID	DOI Data	Data Metadata	DOI Software	Software Metadata	Record Metadata

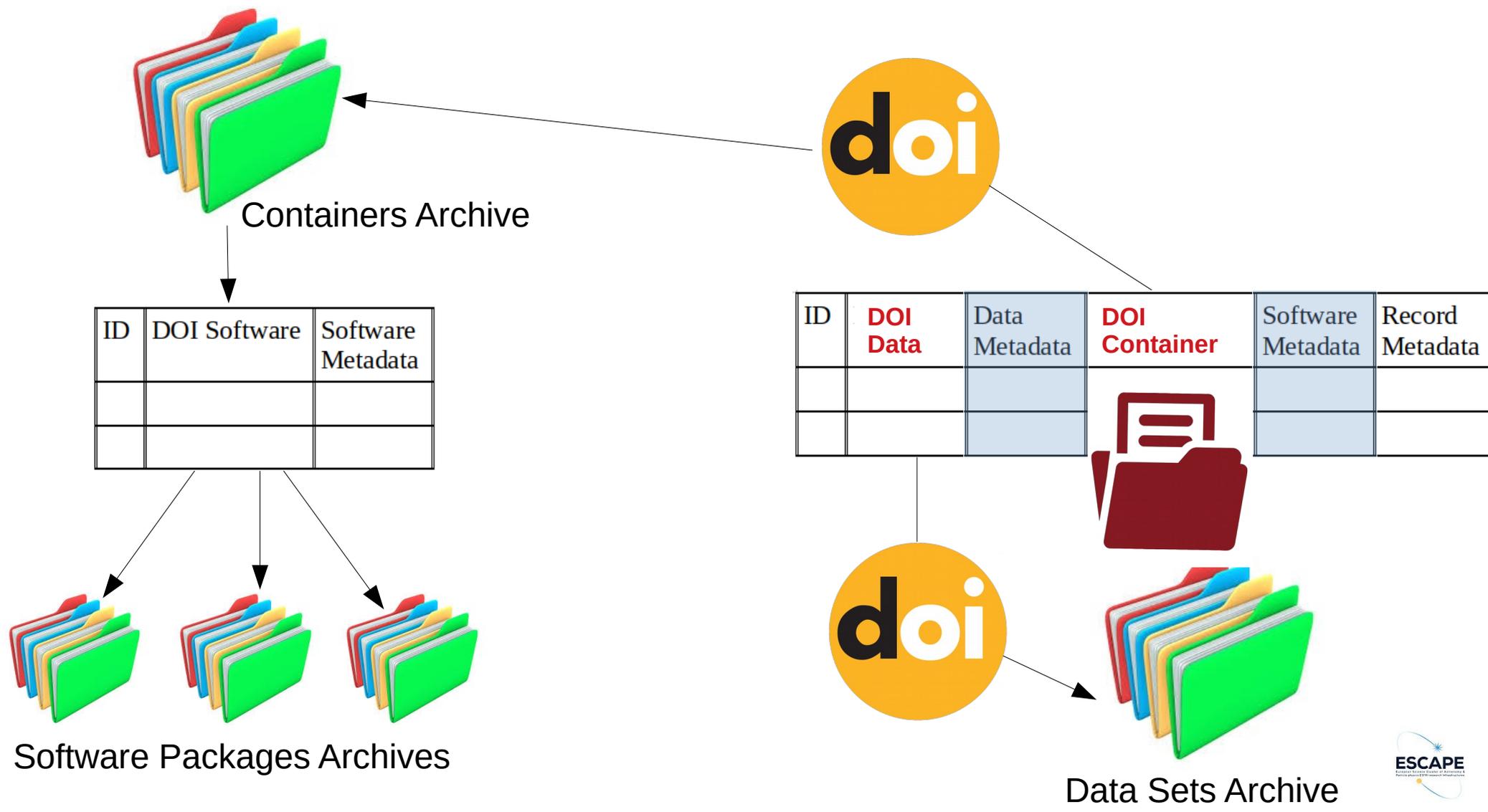
Provide virtual machine
 mounting data
 installing software



Re-produce

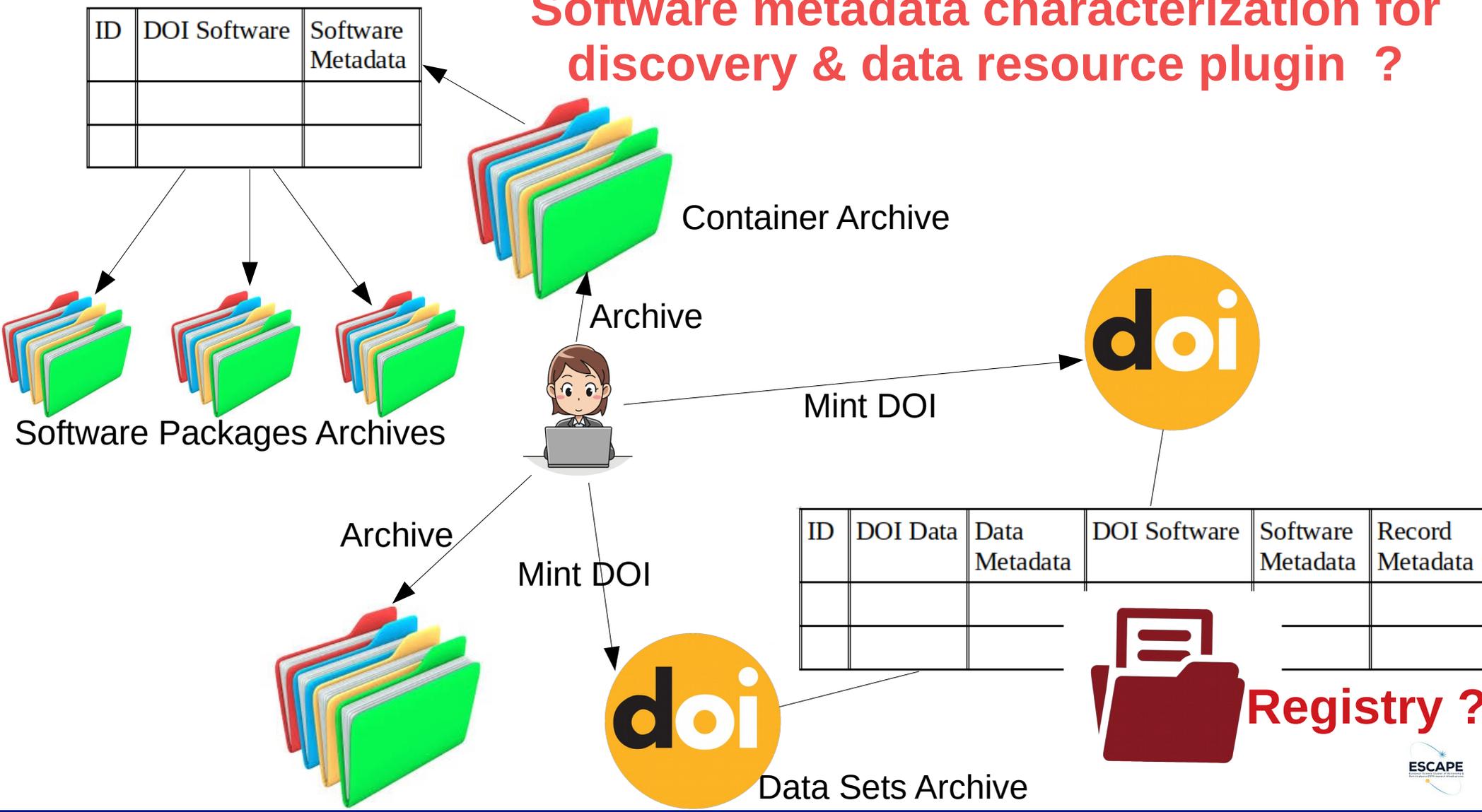


Software Archiving



What do we (IVOA) need for?

Software metadata characterization for discovery & data resource plugin ?



Acknowledgment



This presentation benefits support from the ESCAPE project, funded by the European Commission, Horizon2020 programme (grant n. 824064).

Questions

?

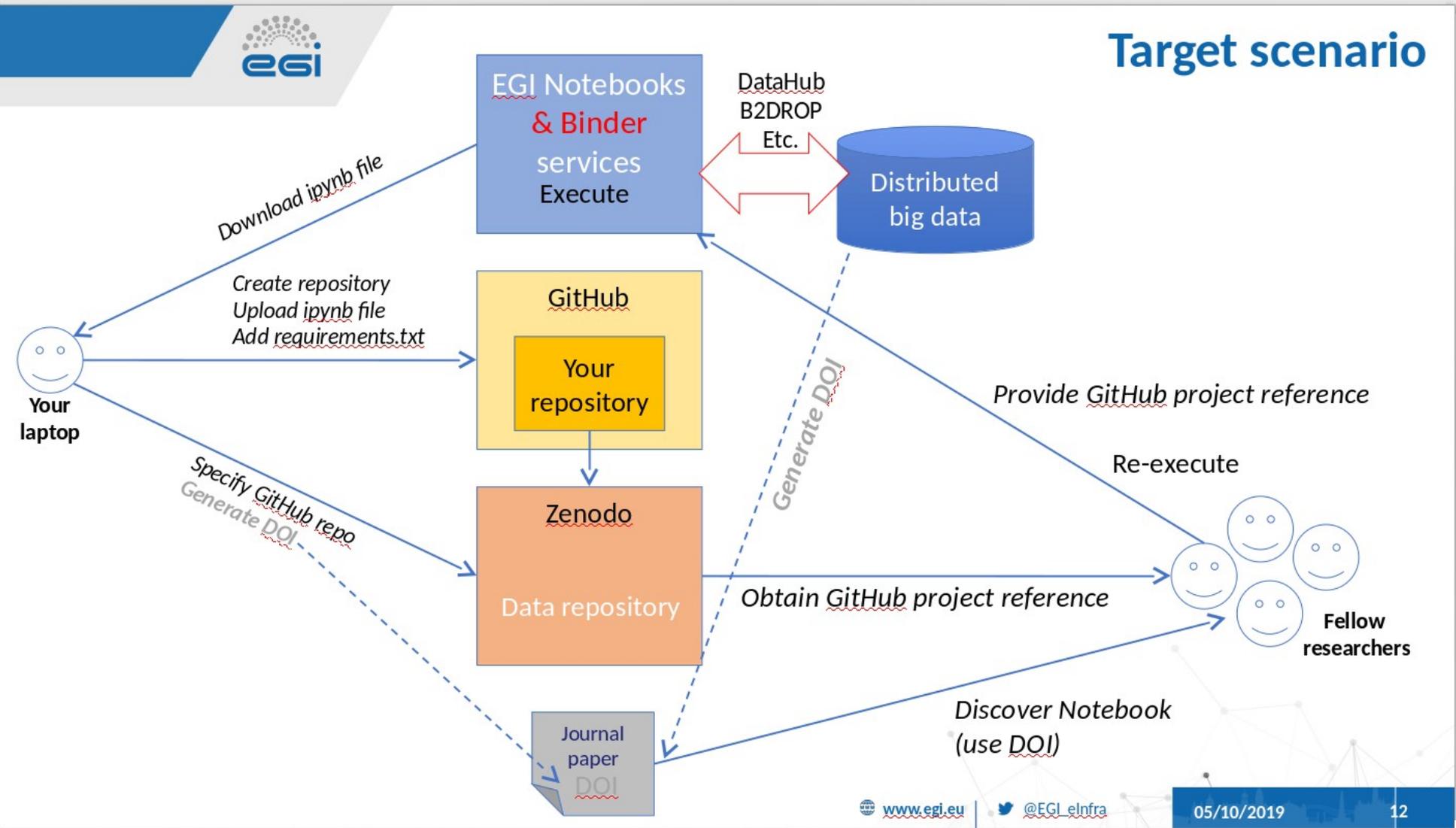
<https://guides.github.com/activities/citable-code/>

“Sharing your data and software on Zenodo”
<https://doi.org/10.5281/zenodo.802100>

“Open Science and reproducibility”
<https://doi.org/10.5281/zenodo.3470835>

“Data analysis with Jupyter Notebook for Open Science”
<https://indico.egi.eu/indico/event/4431/session/12/contribution/36/material/slides/0.pptx>

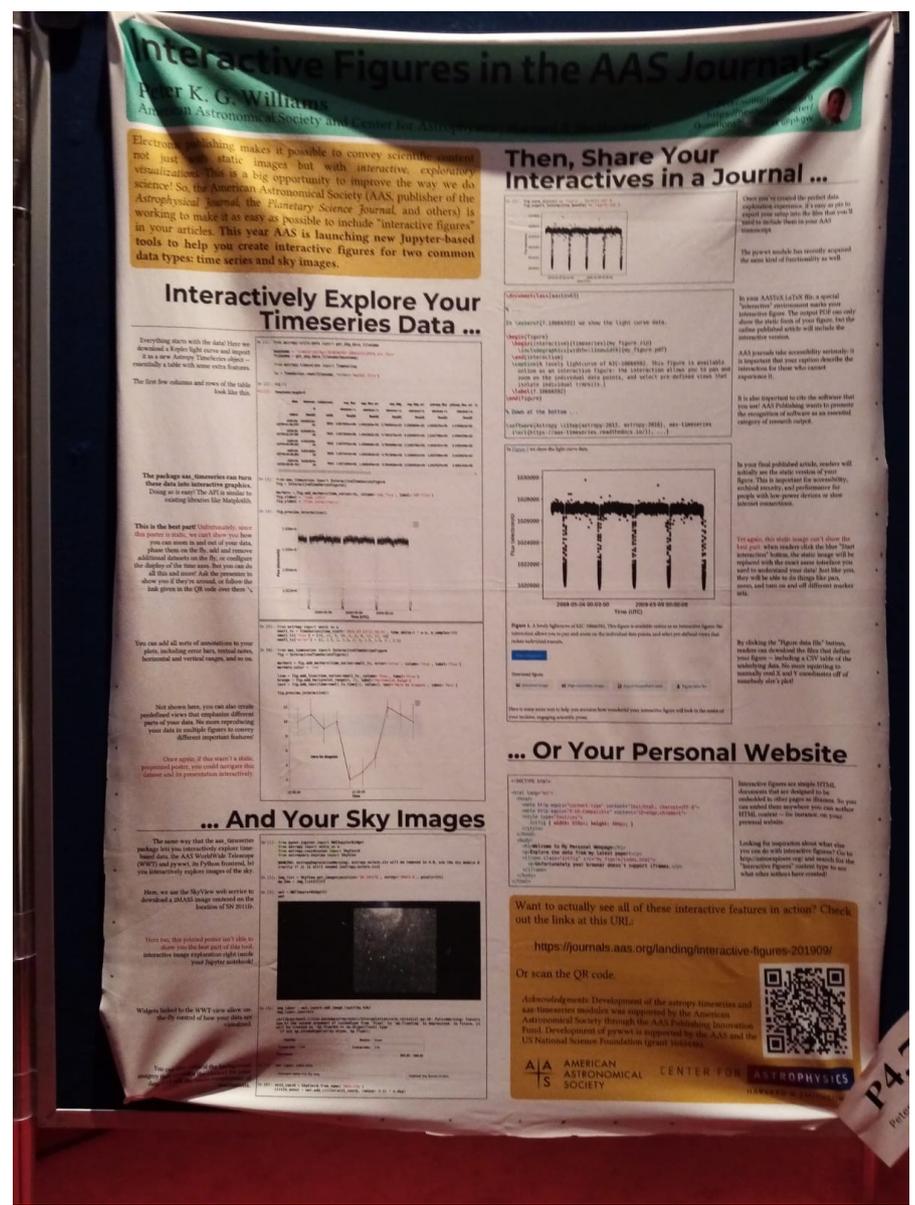
From EGI Conference 2019



<https://indico.egi.eu/indico/event/4431/session/12/contribution/71/material/slides/0.pptx>



From ADASS 2019



Poster 4.7
 Peter Williams —
 Interactive Figures in the AAS
 Journals

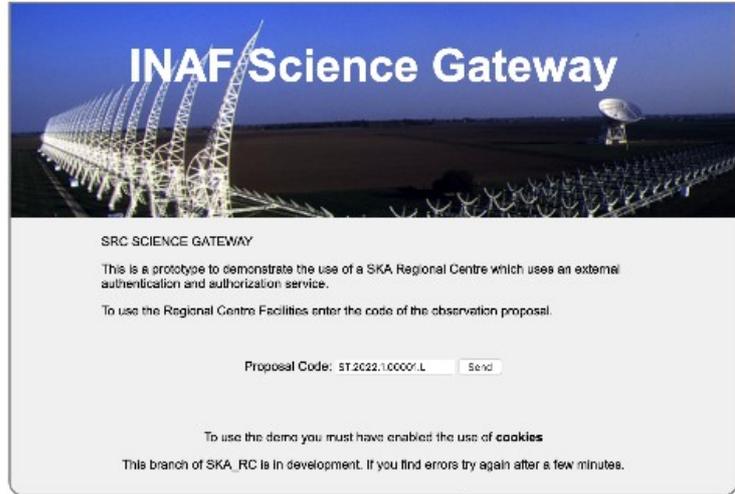
Backup slides

Steps to obtain Software DOI

- Choose the repository to assign the DOI
- Login to Zenodo
- In github authorize Zenodo to access the repo
- Create a release in Zenodo
- Acquire the DOI

Computation

Prototype use case for SKA RC



Science Gateway Catalog
 IN: proposalID
 OUT: IVO-ID, Data_Center, files,
 ivo://authority.org/path?groupID



Data Center (Archive+Computation)
 Token exchange
 Data access and computation

Grouper/GMS
 IN: groupID, user_unique_ID
 OUT: authorization

Authentication

User unique Identifier