



HAL
open science

Towards reusable derived data in neuroimaging

Camille Maumet

► **To cite this version:**

Camille Maumet. Towards reusable derived data in neuroimaging. MRI Together 2023 - Global workshop on open, reproducible, and inclusive MR research, Dec 2023, Online, Germany. pp.1-53, 2023. hal-04328352

HAL Id: hal-04328352

<https://inria.hal.science/hal-04328352>

Submitted on 7 Dec 2023

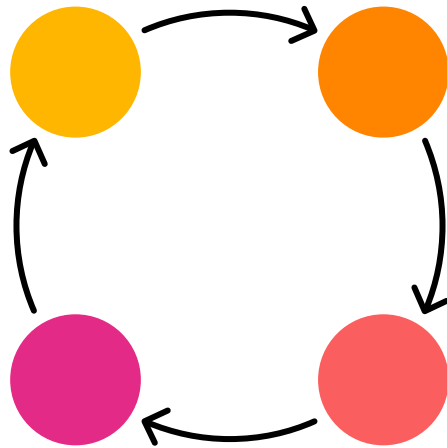
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Session MRI data liberation: Sharing for a better future



Towards reusable derived data in neuroimaging

December 7, 2023

Camille Maumet

Inria, Univ Rennes, Inserm, CNRS, IRISA



FAIR neuroimaging

Findable **A**ccessible **I**nteroperable **R**eusable



Technical solutions: Standards, Databases,
Good practices (version control, etc.)

FAIR neuroimaging

Findable **A**ccessible **I**nteroperable **R**eusable

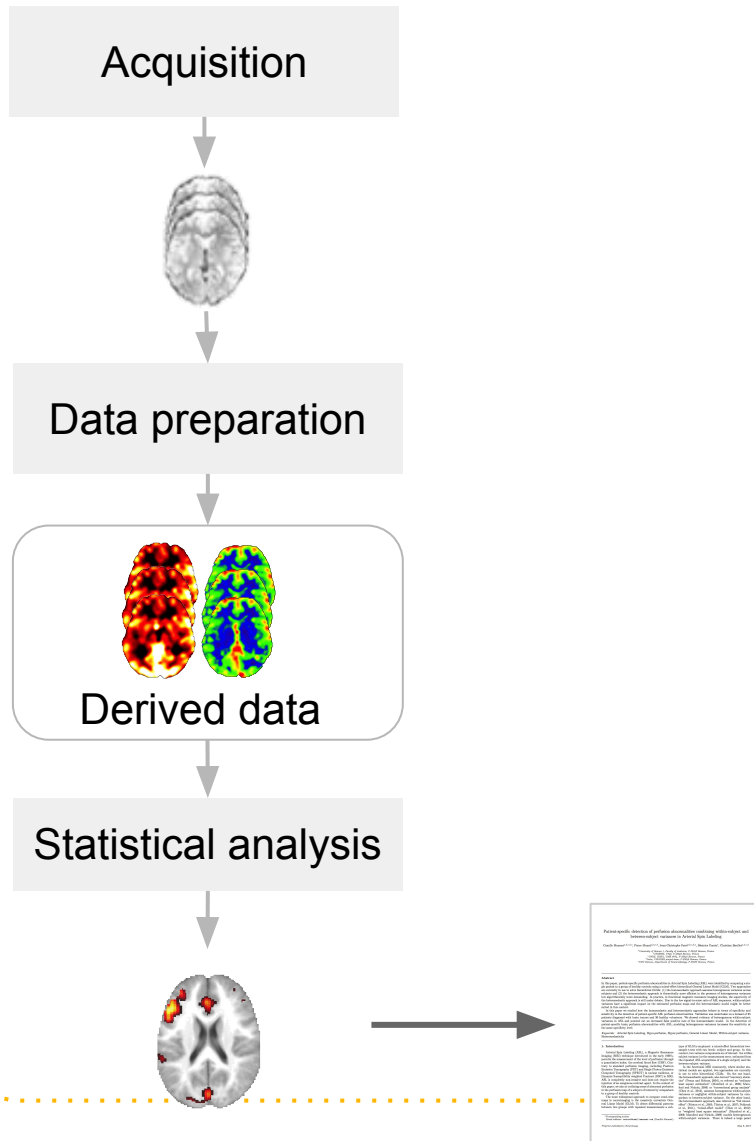


Technical solutions: Standards, Databases,
Good practices (version control, etc.)

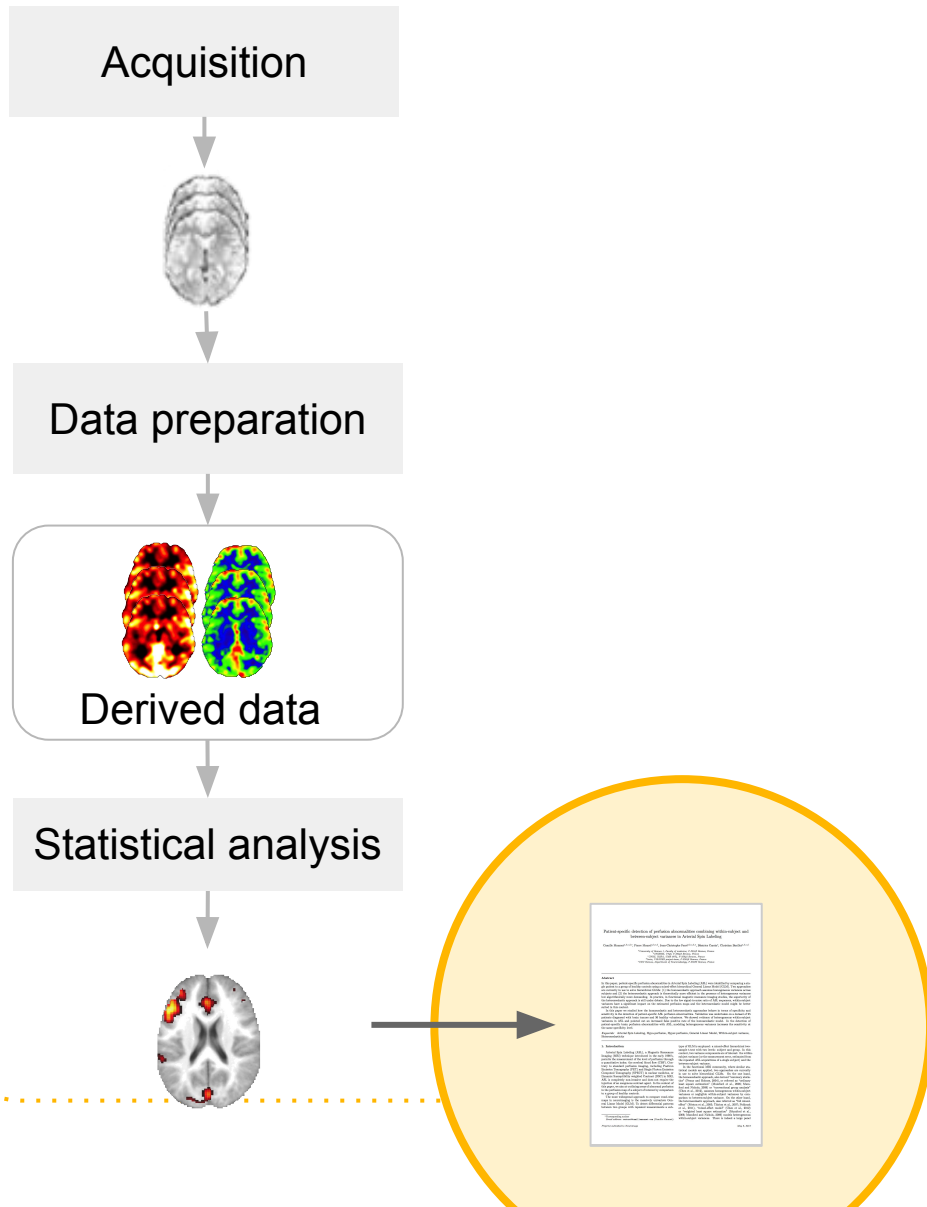
A culture change: towards community-driven research

- Making resources **available to the scientific community**
- **Reusing** existing resources (datasets, tools) that were created by others

Sharing FAIR research products



Sharing FAIR research products

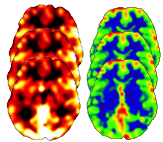


Sharing FAIR research products

Acquisition

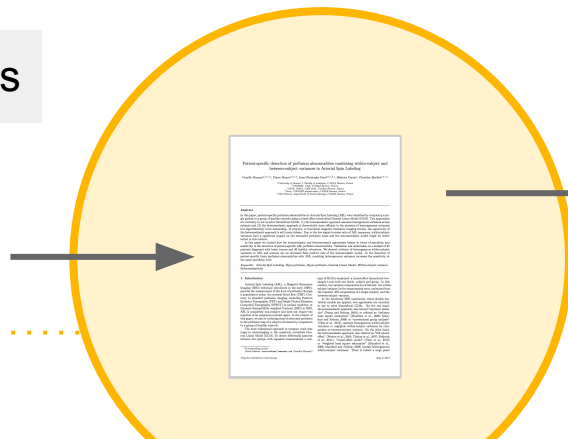
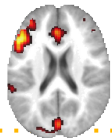


Data preparation

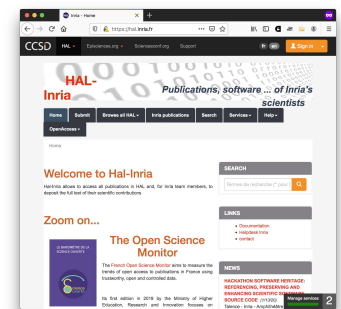


Derived data

Statistical analysis



Biorxiv, arxiv,
HAL,
Pubmed



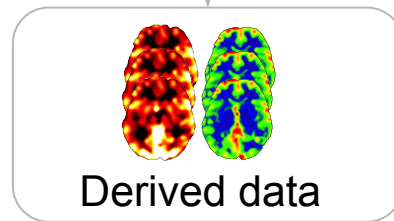
Credits: [Parcels 1 2 & 4 \(CC0\)](#), [Parcel 3 \(CC0\)](#), [Parcel 5 \(CC0\)](#).

Sharing FAIR research products

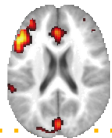
Acquisition



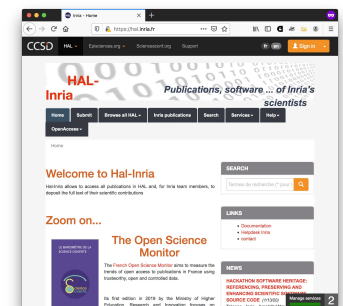
Data preparation



Statistical analysis



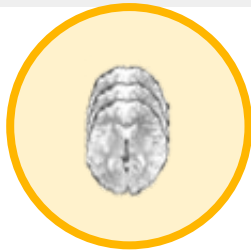
Biorxiv, arxiv,
HAL,
Pubmed



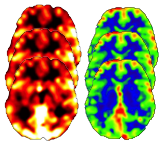
Credits: [Parcels 1 2 & 4 \(CCo\)](#), [Parcel 3 \(CCo\)](#), [Parcel 5 \(CCo\)](#).

Sharing FAIR research products

Acquisition

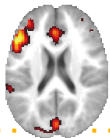


Data preparation

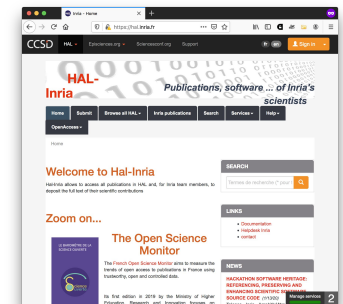


Derived data

Statistical analysis



Biorxiv, arxiv,
HAL,
Pubmed



Credits: [Parcels 1 2 & 4 \(CC0\)](#), [Parcel 3 \(CC0\)](#), [Parcel 5 \(CC0\)](#).

Sharing FAIR datasets



Sharing FAIR datasets



Donders Repository



Sharing FAIR datasets



Donders Repository



Findable we can search for open datasets in each repository

Sharing FAIR datasets



Donders Repository



Findable we can search for open datasets in each repository

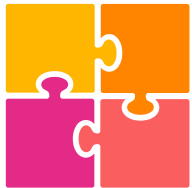


Accessible Direct download (e.g. *Openneuro*) or with Data User Agreement (e.g. *Donders*)

Sharing FAIR datasets



Donders Repository



Interoperable?



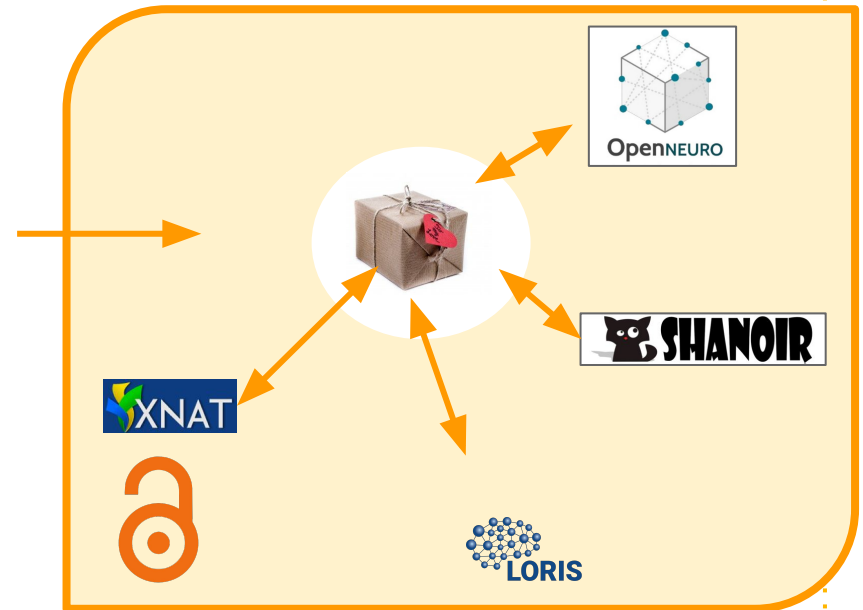
Findable we can search for open datasets in each repository



Accessible Direct download (e.g. *Openneuro*) or with Data User Agreement (e.g. *Donders*)

Standards for neuroimaging data sharing

- Data and **meta-data**
- A common “language”



Brain Imaging Data Structure

```
my_dataset/  
├── participants.tsv  
├── sub-01/  
│   ├── anat/  
│   │   └── sub-01_T1w.nii.gz  
│   ├── func/  
│   │   ├── sub-01_task-rest_bold.nii.gz  
│   │   └── sub-01_task-rest_bold.json  
│   └── dwi/  
│       ├── sub-01_dwi.nii.gz  
│       ├── sub-01_dwi.json  
│       ├── sub-01_dwi.bval  
│       └── sub-01_dwi.bvec  
├── sub-02/  
├── sub-03/  
└── sub-04/
```



The Brain Imaging Data Structure (BIDS): As Easy as Baking a Cake

Patricia Clement, Ghent University Hospital, BE

OpenNEURO

SEARCH SUPPORT FAQ Sign in

A free and open platform for validating and sharing BIDS-compliant [MRI](#), [PET](#), [MEG](#), [EEG](#), and [iEEG](#) data

26,870 Participants 693 Public Datasets

Browse by Modalities

Or Search

SIGN IN Google ORCID

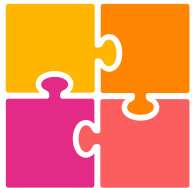
[Gorgolewski et. al, Scientific Data 2016]

<https://openneuro.org/>

Sharing FAIR datasets



Donders Repository



Interoperable?



Findable we can search for open datasets in each repository

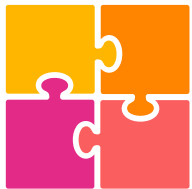


Accessible Direct download (e.g. *Openneuro*) or with Data User Agreement (e.g. *Donders*)

Sharing FAIR datasets



Donders Repository



Interoperable BIDS acts as a common language across tools



Findable we can search for open datasets in each repository

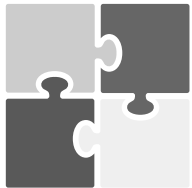


Accessible Direct download (e.g. *Openneuro*) or with Data User Agreement (e.g. *Donders*)

Sharing FAIR datasets



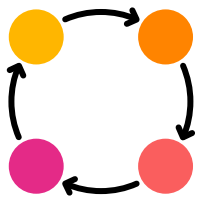
Donders Repository



Interoperable BIDS acts as a common languages across tools



Findable we can search for open datasets in each repository



Reusable Metadata in BIDS were selected to enable reuse. *e.g. onsets of the task for task-based fMRI.*

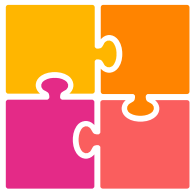


Accessible Direct download (*e.g. Openneuro*) or with Data User Agreement (*e.g. Donders*)

Sharing FAIR datasets



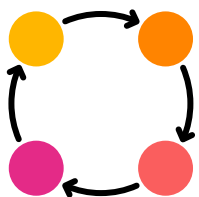
Donders Repository



Interoperable BIDS acts as a common language across tools



Findable we can search for open datasets in each repository



Reusable Metadata in BIDS were selected to enable reuse. *e.g. onsets of the task for task-based fMRI.*



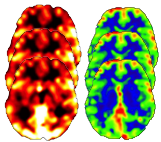
Accessible Direct download (*e.g. Openneuro*) or with Data User Agreement (*e.g. Donders*)

Sharing FAIR research products

Acquisition

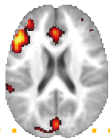


Data preparation

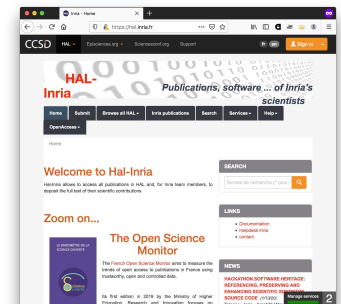


Derived data

Statistical analysis



Biorxiv, arxiv,
HAL,
Pubmed



Credits: [Parcels 1 2 & 4 \(CC0\)](#), [Parcel 3 \(CC0\)](#), [Parcel 5 \(CC0\)](#).

Sharing FAIR research products

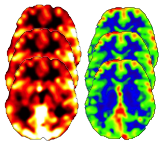
Acquisition



Donders Repository



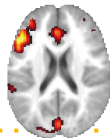
Data preparation



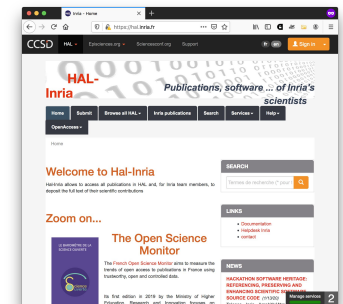
Derived data



Statistical analysis

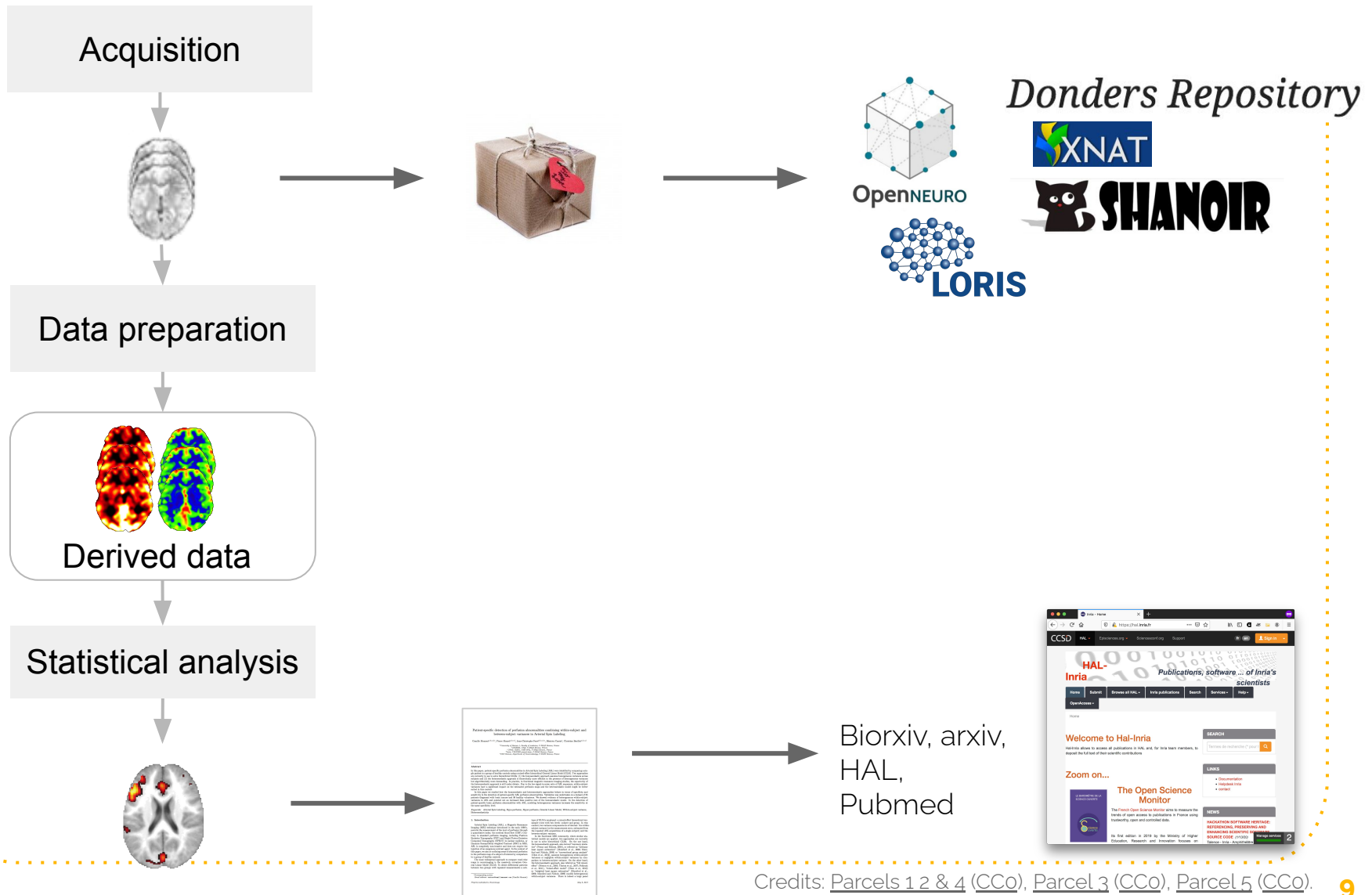


Biorxiv, arxiv,
HAL,
Pubmed

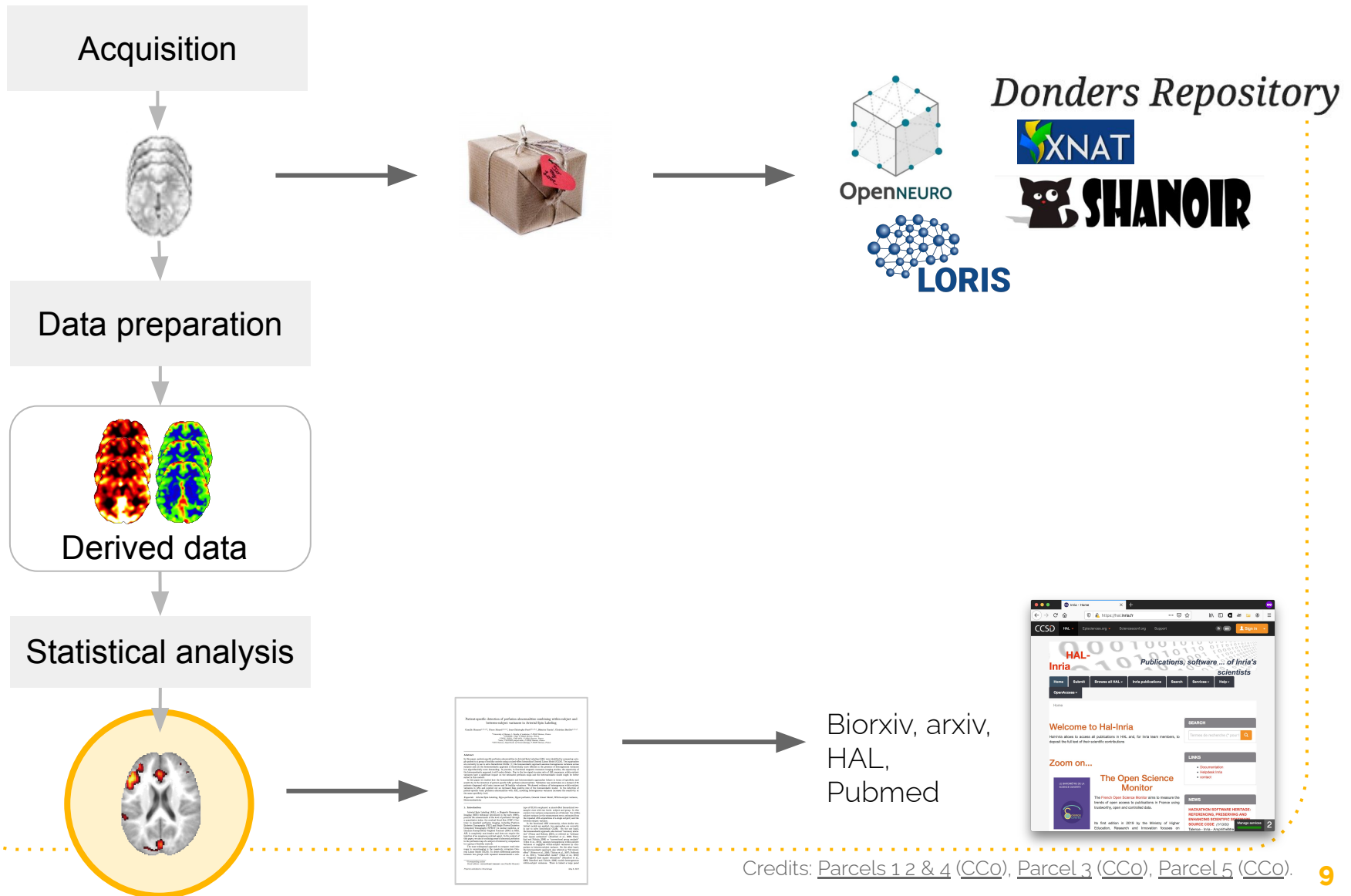


Credits: [Parcels 1 2 & 4 \(CC0\)](#), [Parcel 3 \(CC0\)](#), [Parcel 5 \(CC0\)](#).

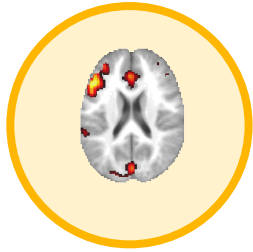
Sharing FAIR research products



Sharing FAIR research products



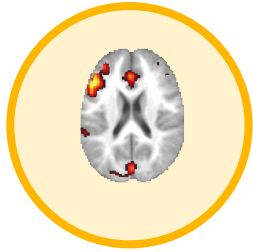
Sharing FAIR results



Sharing FAIR results (task fMRI)



Sharing FAIR results (task fMRI)

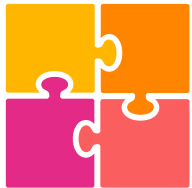
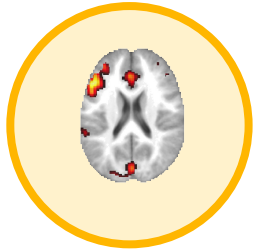


Findable we can search for datasets on NeuroVault

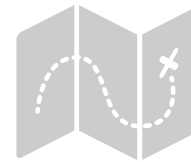


Accessible Direct download

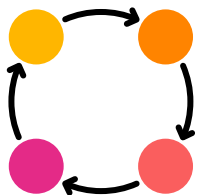
Sharing FAIR results (task fMRI)



Interoperable?



Findable we can search for datasets on NeuroVault

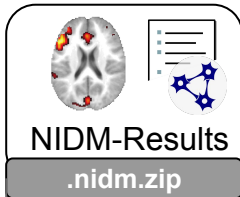


Reusable?



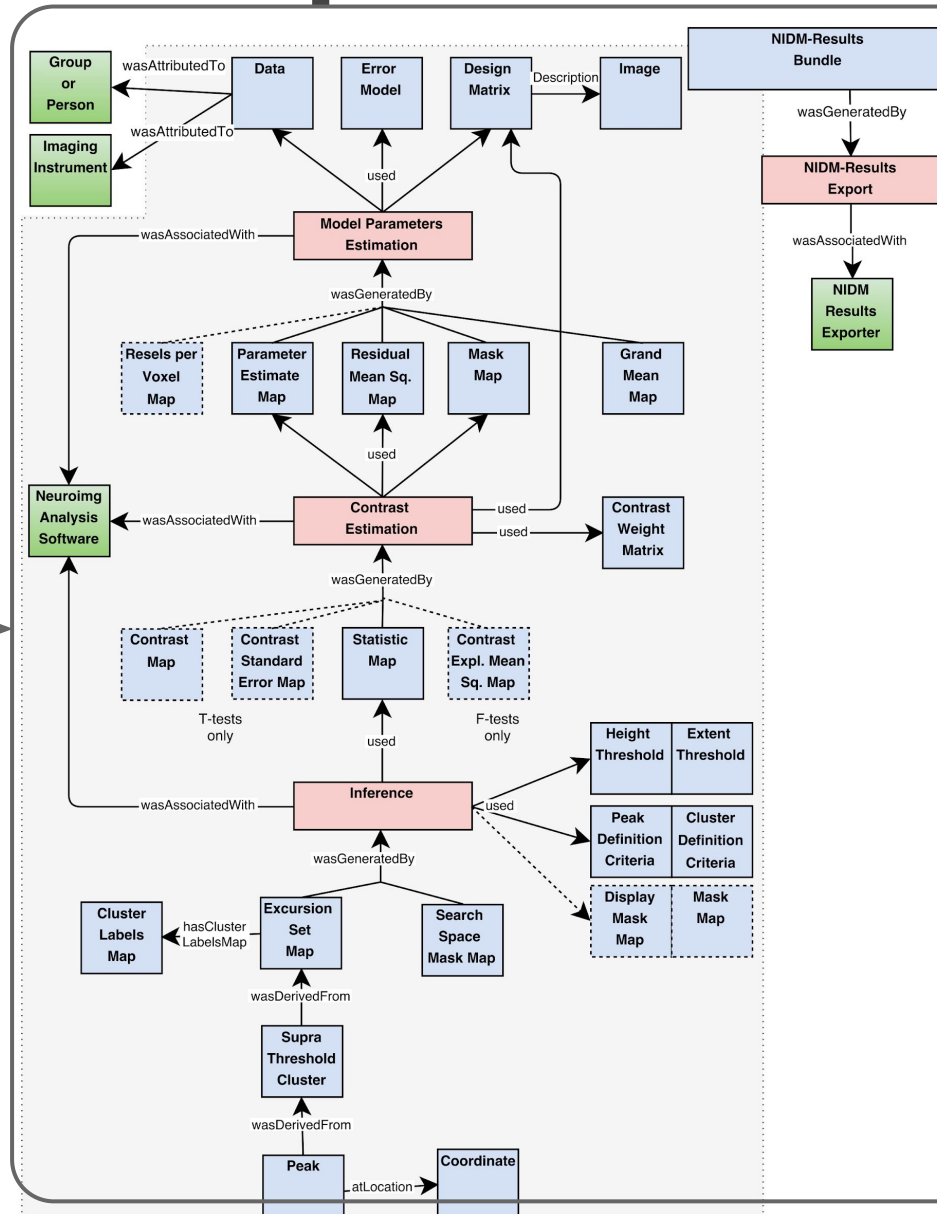
Accessible Direct download

NIDM-Results pack



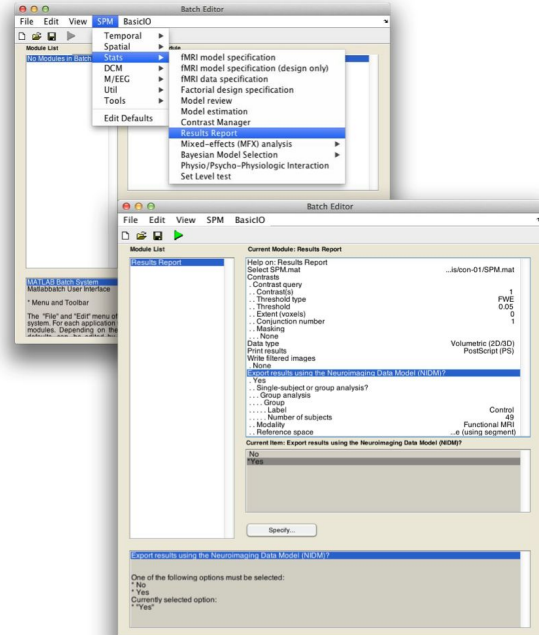
- ClusterLabels.nii.gz
- Contrast.nii.gz
- ContrastStandardError.nii.gz
- DesignMatrix.csv
- DesignMatrix.png
- ExcursionSet.nii.gz
- ExcursionSet.png
- GrandMean.nii.gz
- Mask.nii.gz
- nidm.json
- nidm.ttl
- ParameterEstimate_001.nii.gz
- ParameterEstimate_002.nii.gz
- ResidualMeanSquares.nii.gz
- SearchSpaceMask.nii.gz
- TStatistic.nii.gz
- ZStatistic.nii.gz

[Maumet et al.,
Sci. Data 2016]



NIDM in SPM, FSL & NeuroVault

SPM

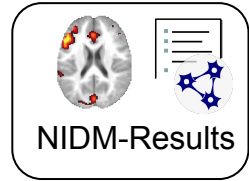


1. Export



```
$ nidmfsfsl fsl_ds107_group 49 -g Control
```

2. Publication on NeuroVault



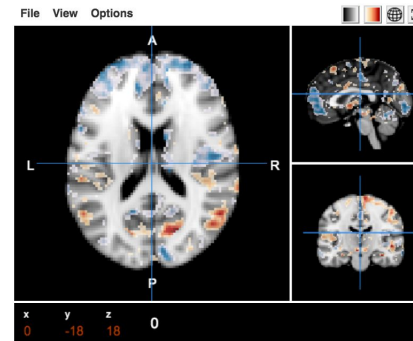
A Correspondence between Individual Differences in the Brain's Intrinsic Functional Architecture and the Content and Form of Self-Generated Thoughts

Contributed by ChrisFiloGorgolewski

Krzysztof J. Gorgolewski, Dan Lurie, Sebastian Urchs, Judy A. Kipping, R. Cameron Craddock, Michael P. Milham, Daniel S. Margulies, Jonathan Smallwood

[Link to the paper](#)

3D View



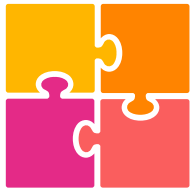
Images Details

Show 7 entries Search:

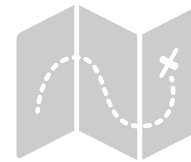
View	ID	Name	Type
	25	fALFF: Past	Z map
	26	fALFF: Future	Z map
	27	fALFF: Past > Future	Z map
	28	fALFF: Positive	Z map
	29	fALFF: Negative	T map
	30	fALFF: Positive > Negative	Z map
	31	fALFF: Social Cognition	Z map

Showing 1 to 7 of 30 entries First Previous Next Last

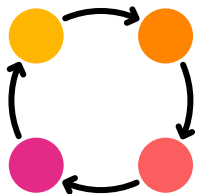
Sharing FAIR results (task fMRI)



Interoperable NIDM-Results acts as a common languages across tools



Findable we can search for datasets on NeuroVault

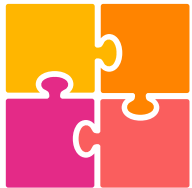


Reusable Metadata in NIDM-Results selected to enable meta-analyses



Accessible Direct download

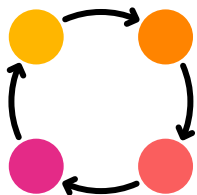
Sharing FAIR results (task fMRI)



Interoperable NIDM-Results acts as a common languages across tools



Findable we can search for datasets on NeuroVault

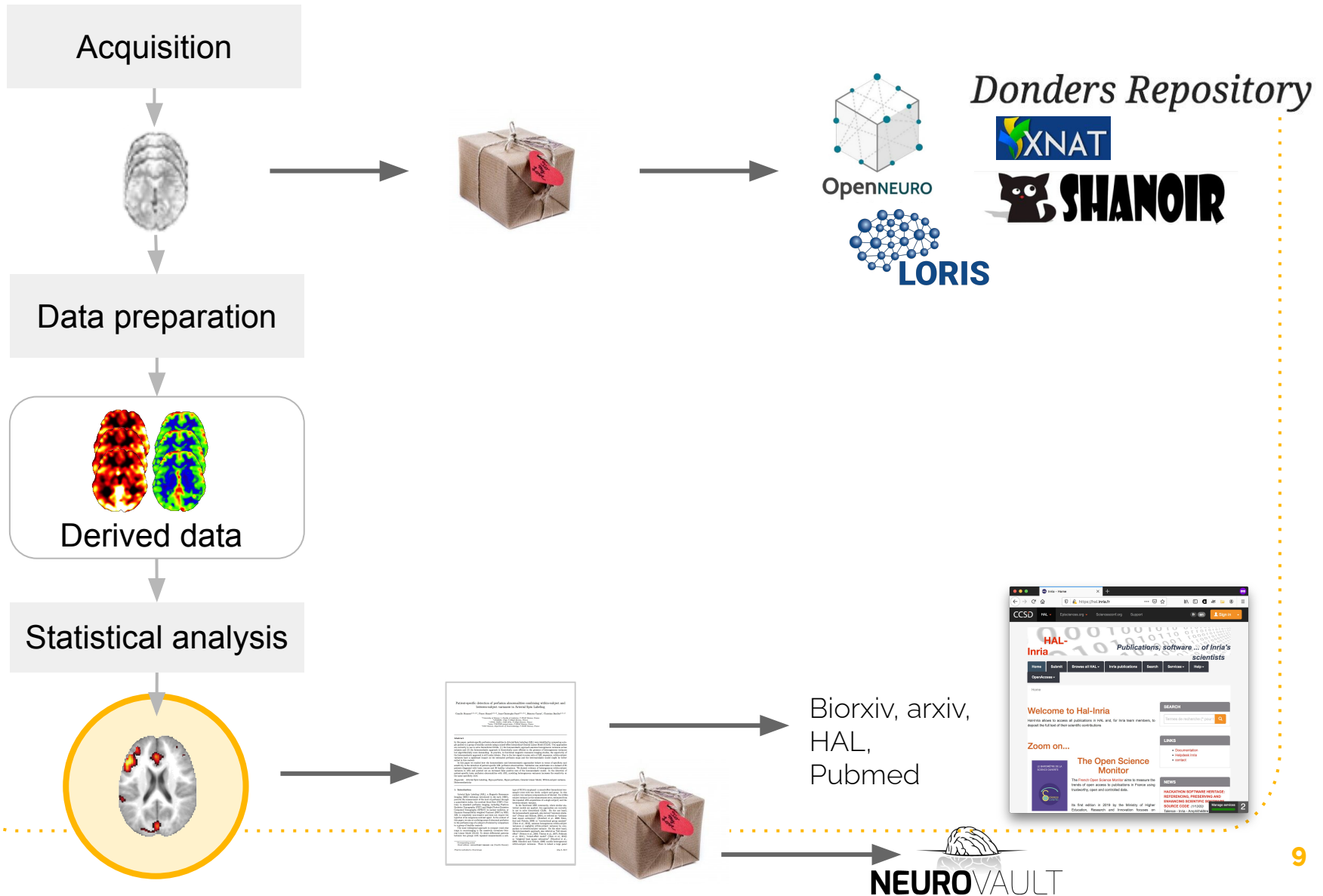


Reusable Metadata in NIDM-Results selected to enable meta-analyses

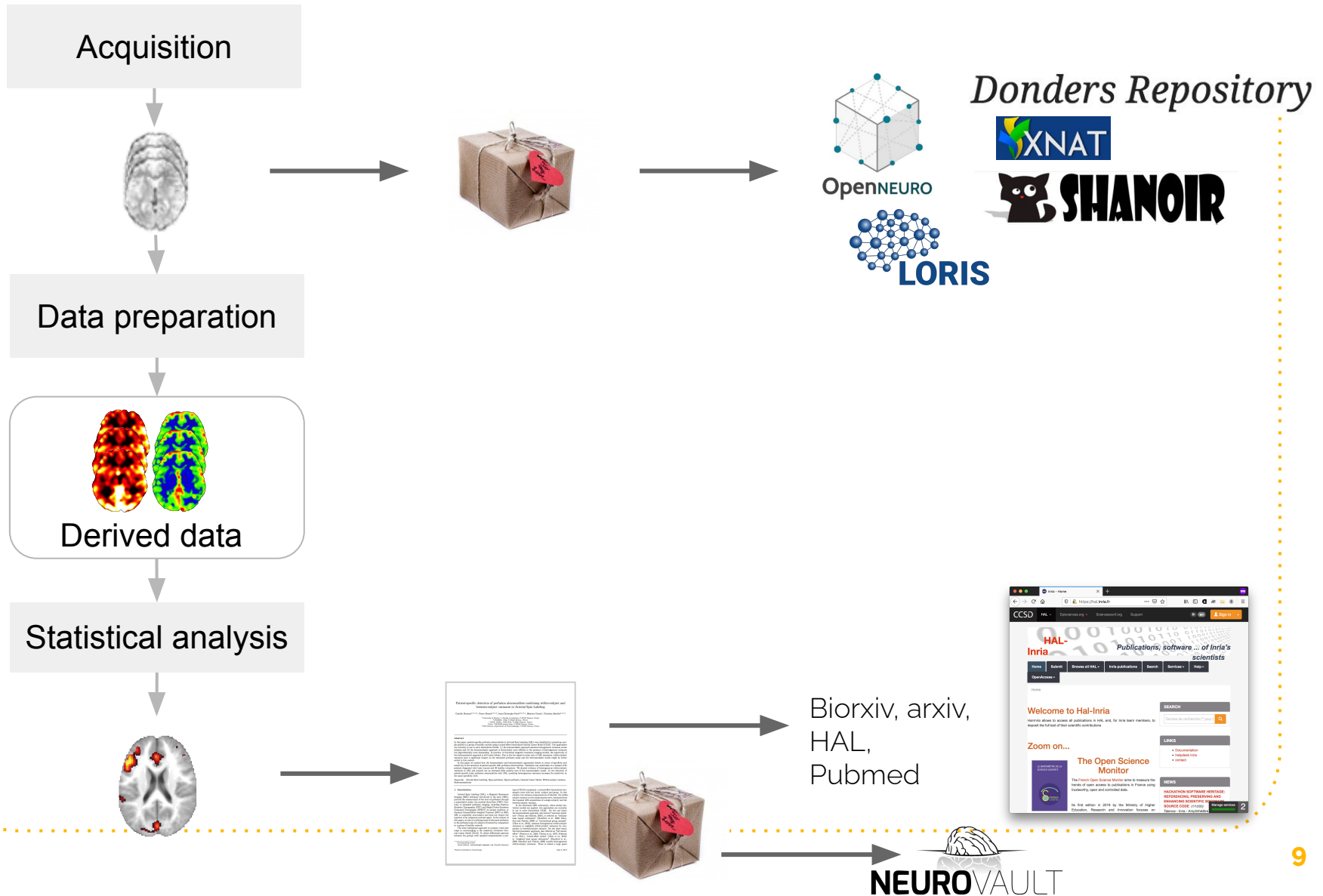


Accessible Direct download

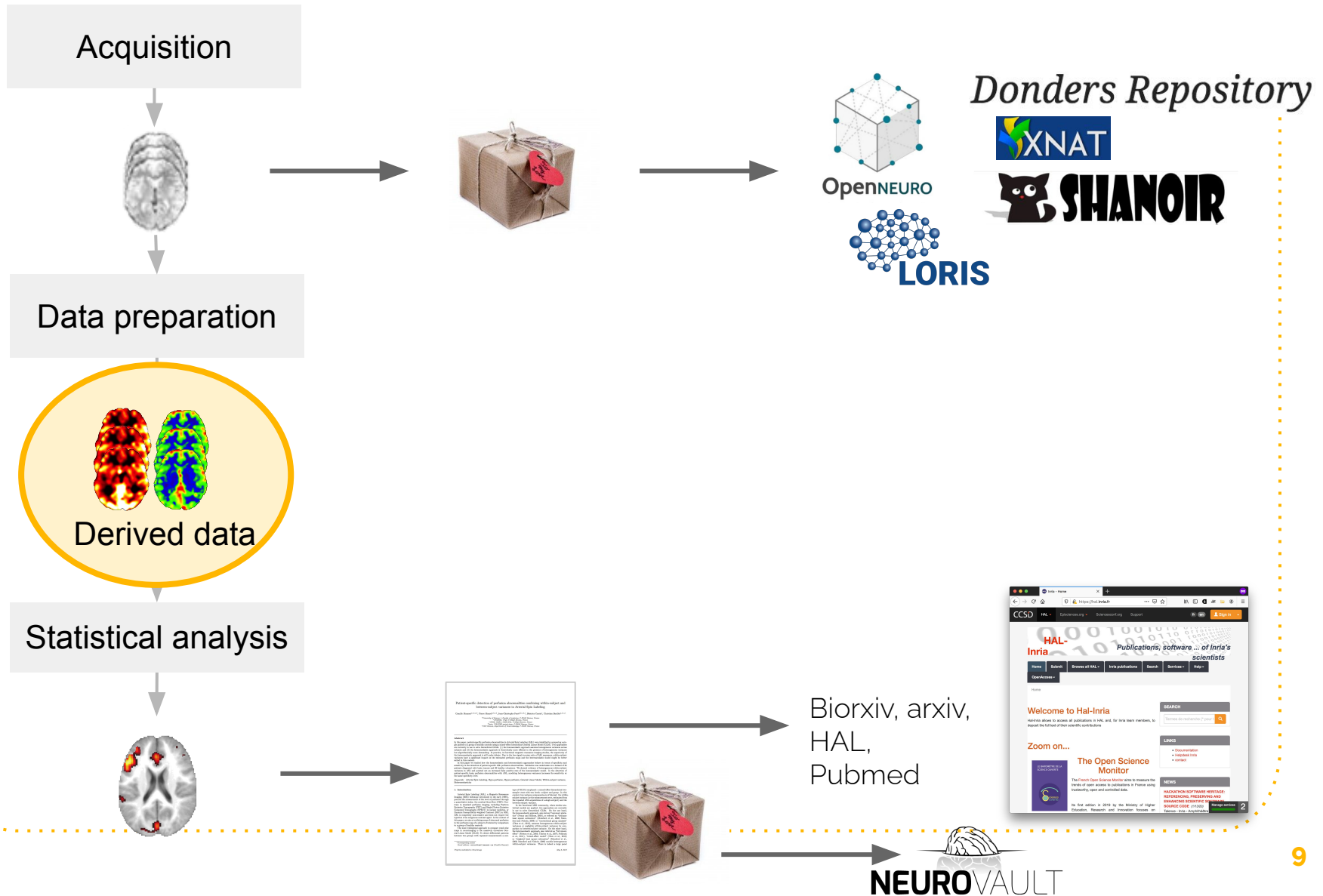
Sharing FAIR research products



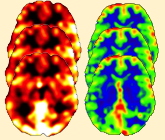
Sharing FAIR research products



Sharing FAIR research products

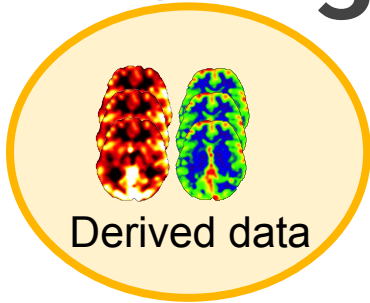


Sharing FAIR intermediary datasets



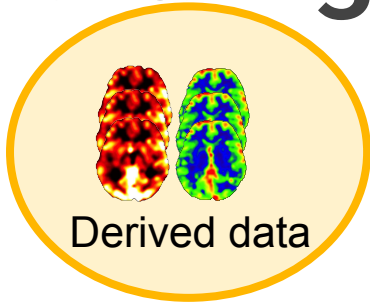
Derived data

Sharing FAIR intermediary datasets



General purpose repositories

Sharing FAIR intermediary datasets



General purpose repositories

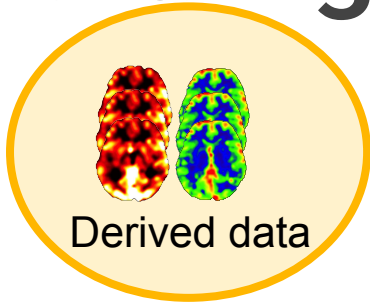


Findable can search for open datasets (less specific queries)

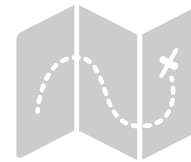


Accessible Direct download or Data User Agreement

Sharing FAIR intermediary datasets



General purpose repositories

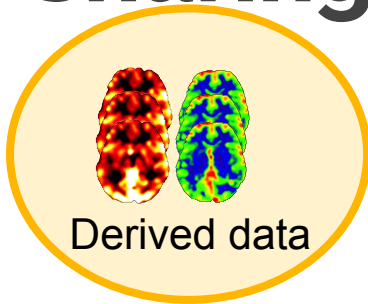


Findable can search for open datasets (less specific queries)

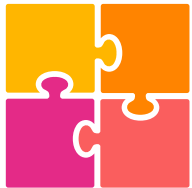


Accessible Direct download or Data User Agreement

Sharing FAIR intermediary datasets



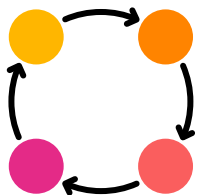
General purpose repositories



Interoperable?



Findable can search for open datasets (less specific queries)



Reusable?



Accessible Direct download or Data User Agreement

Sharing other FAIR datasets

1- Look for existing standards and databases:



2- You can help develop the next standards

BIDS Derivatives

<https://bids.neuroimaging.io/>

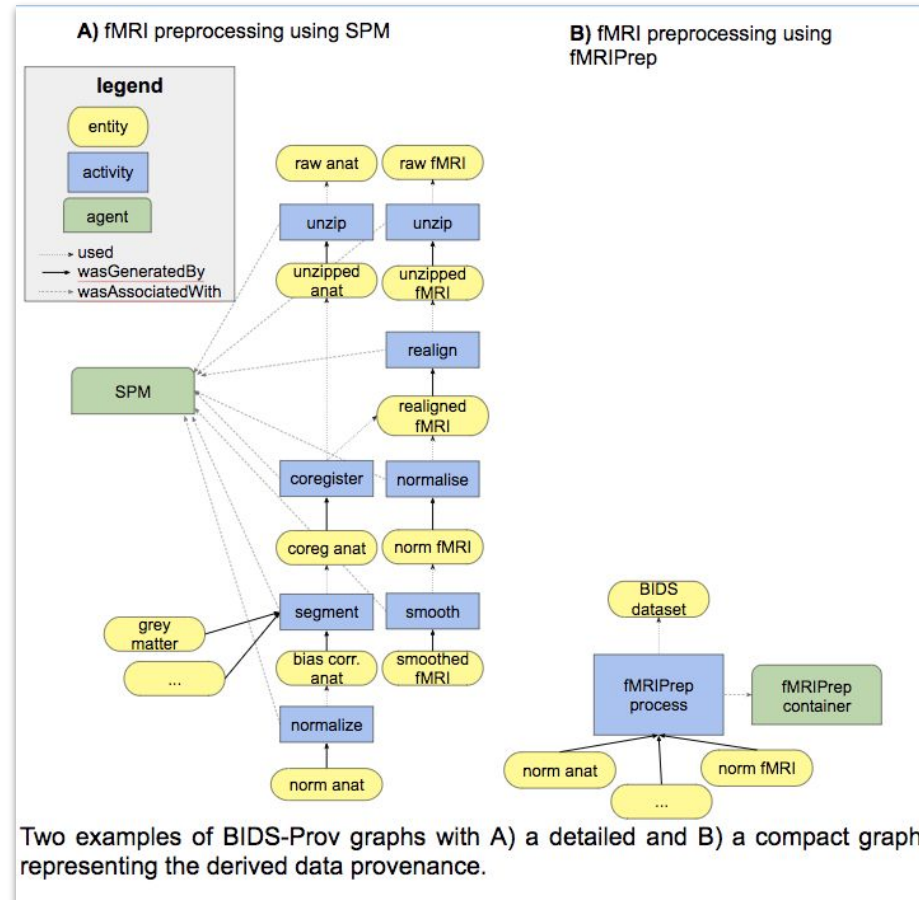
Derivatives are outputs of common processing pipelines, capturing data and meta-data sufficient for a researcher to understand and (critically) reuse those outputs in subsequent processing. Standardizing derivatives is motivated by use cases where formalized machine-readable access to processed data enables higher level processing.

- BEP011 The structural preprocessing derivatives
 - BEP012 The functional preprocessing derivatives
 - BEP013 The resting state fMRI derivatives
 - BEP016 The diffusion weighted imaging derivatives
 - BEP017 Generic BIDS connectivity data schema
 - BEP021 Common Electrophysiological Derivatives
- And more!

BIDS Provenance



Satra Ghosh



<https://github.com/bids-standard/BEP028> BIDSprov

Provenance Use Case 1 **An unambiguous description of the methods**



Provenance Use Case 1 **An unambiguous description of the methods**



Which release of SPM was used?

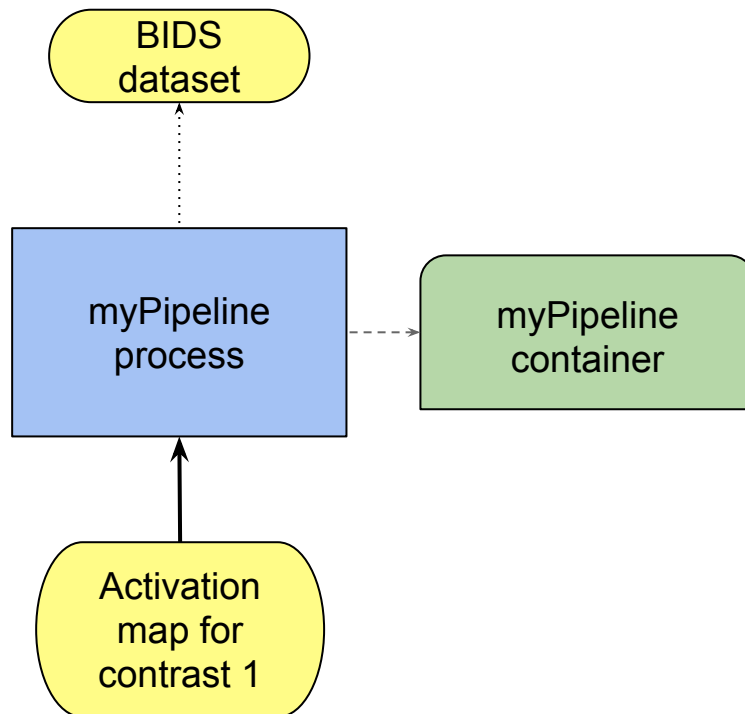
Which registration algorithm was selected?

How was the brain mask computed?

Provenance Use Case 1 **An unambiguous
description of the methods**

Can we
reproduce
exactly the same
results
(same data, same
methods)?

Provenance Use Case 1 **An unambiguous description of the methods**



Can we
**reproduce
exactly the same
results**
(same data, same
methods)?

Provenance Use Case 2 **Dealing with
unwanted variability**

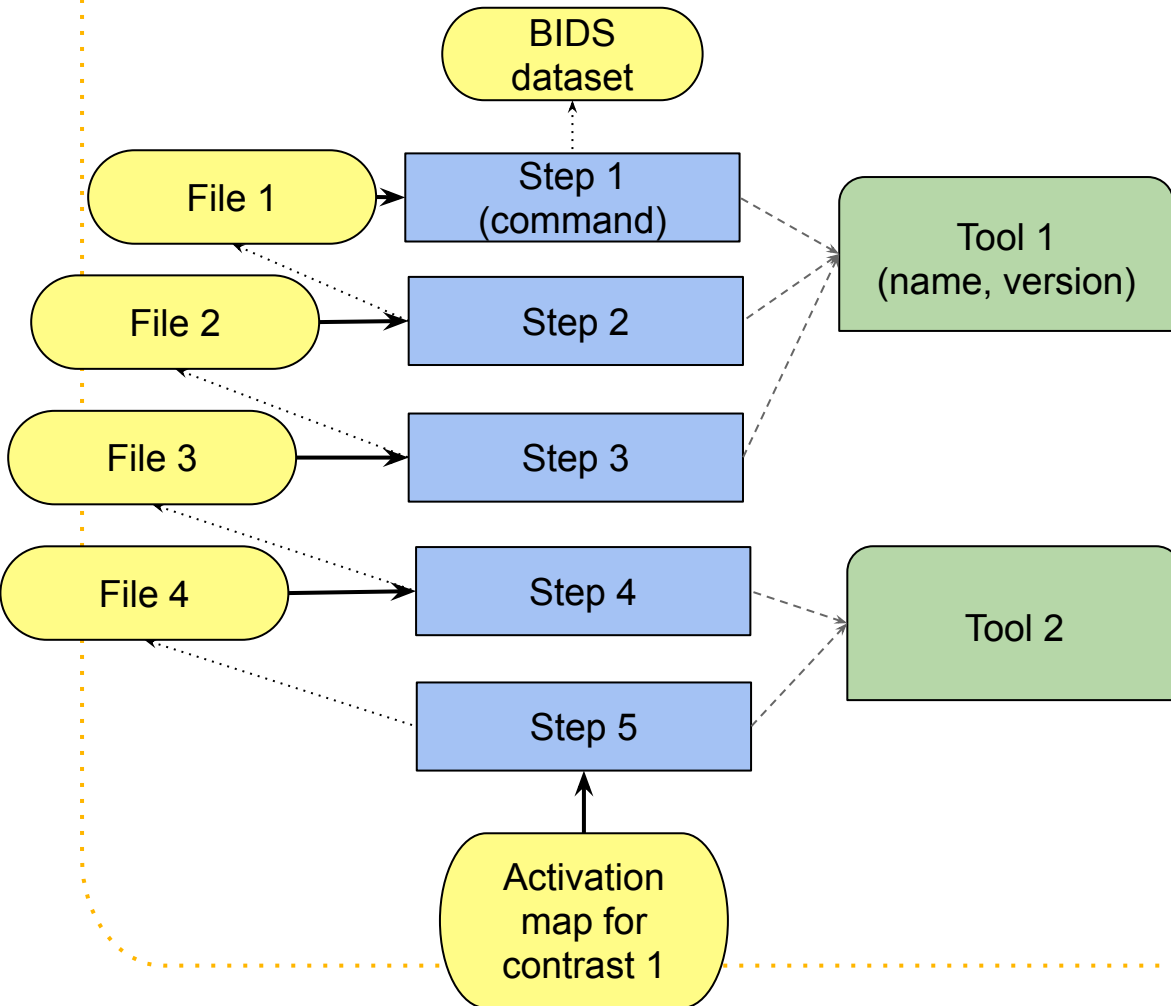
A previously
widely-used
**parameter is no
longer considered
as best practice**

Provenance Use Case 2 **Dealing with unwanted variability**

A previously widely-used **parameter is no longer considered as best practice**

A bug was found in a given tool **version**, can we retrieve all corresponding results?

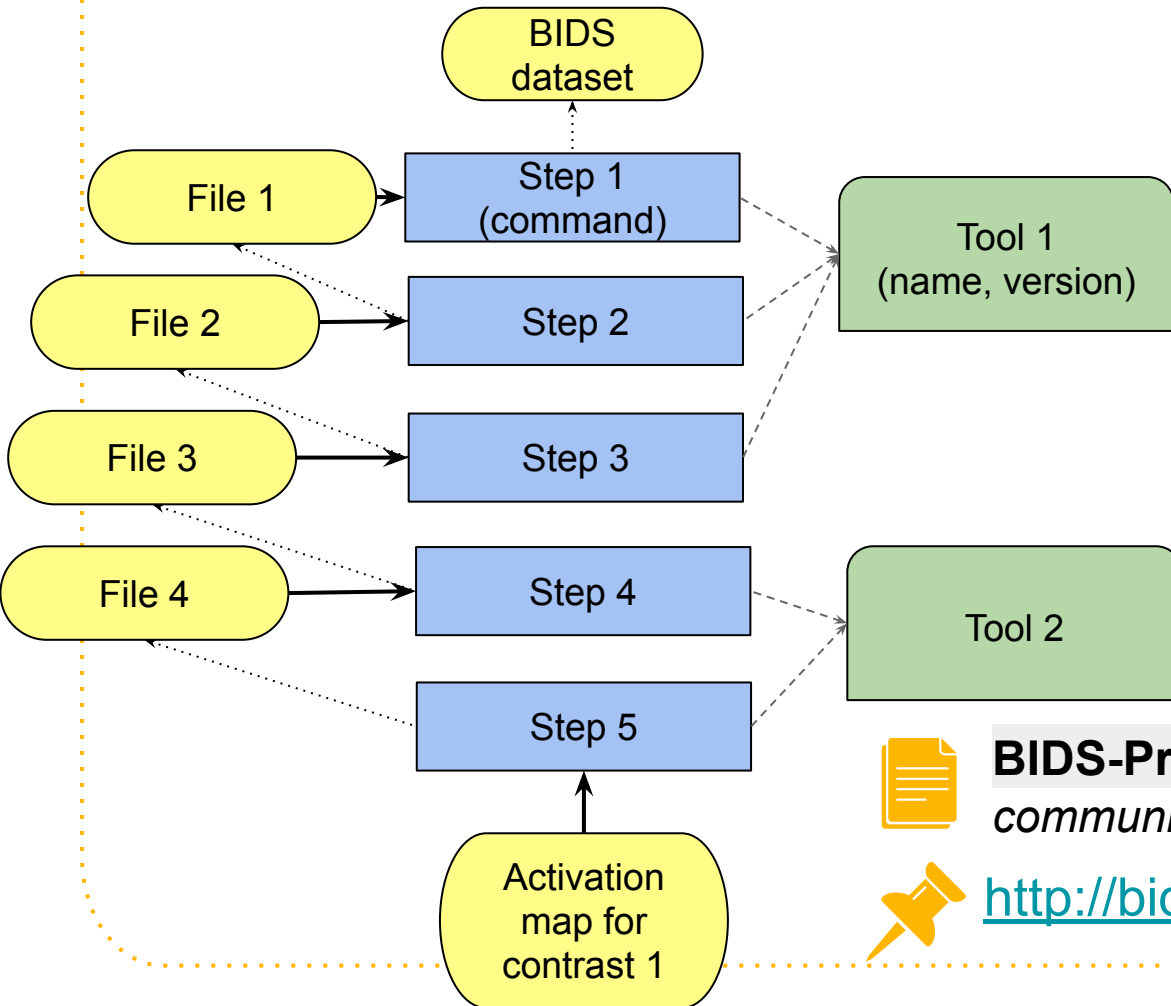
Provenance Use Case 2 **Dealing with unwanted variability**



A previously widely-used parameter is no longer considered as best practice

A bug was found in a given tool version, can we retrieve all corresponding results?

Provenance Use Case 3 **Exploring factors impacting analytical variability**



Are my conclusions dependant on the **size of the smoothing kernel** used at Step 5?

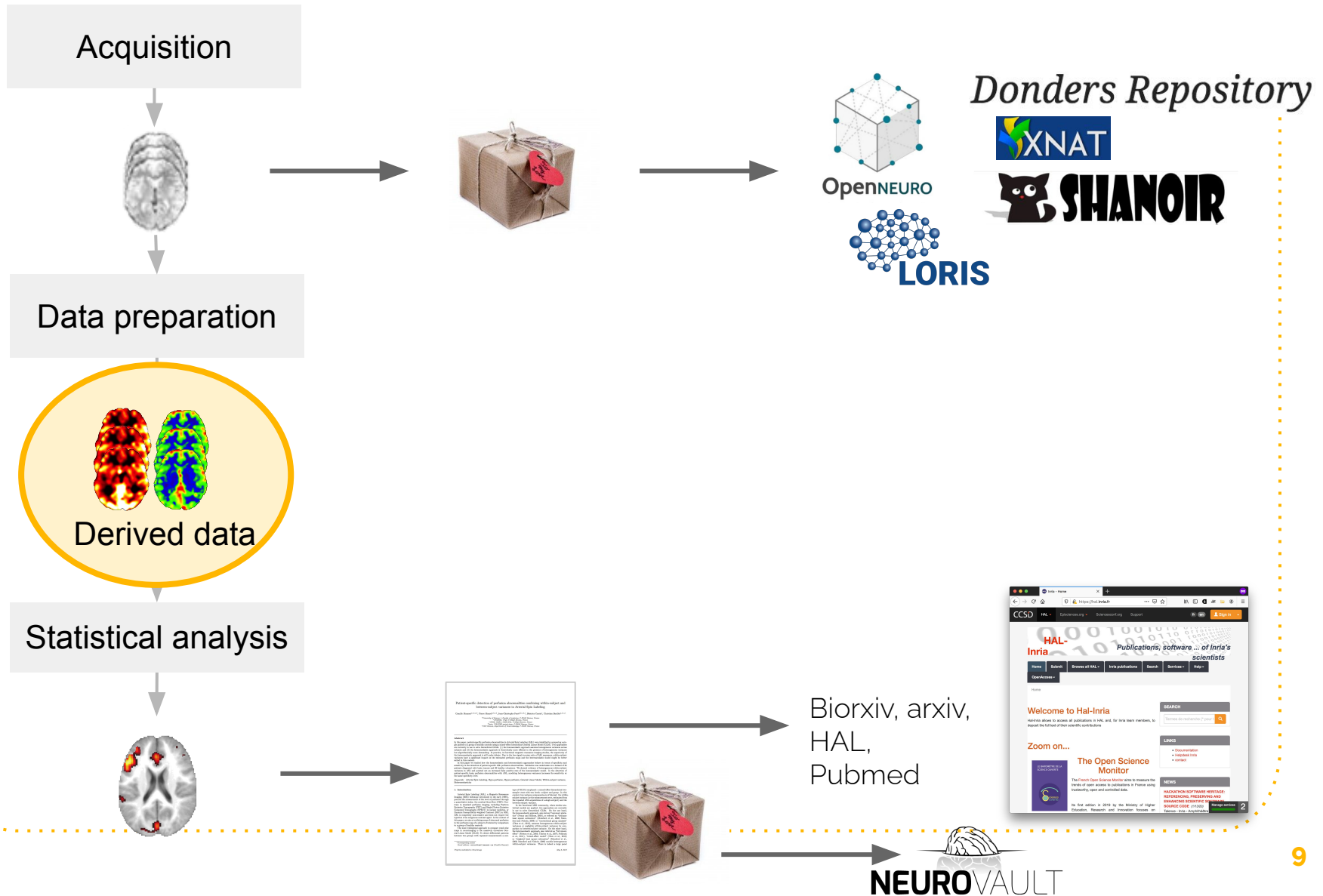


BIDS-Prov will soon be open for community feedback!

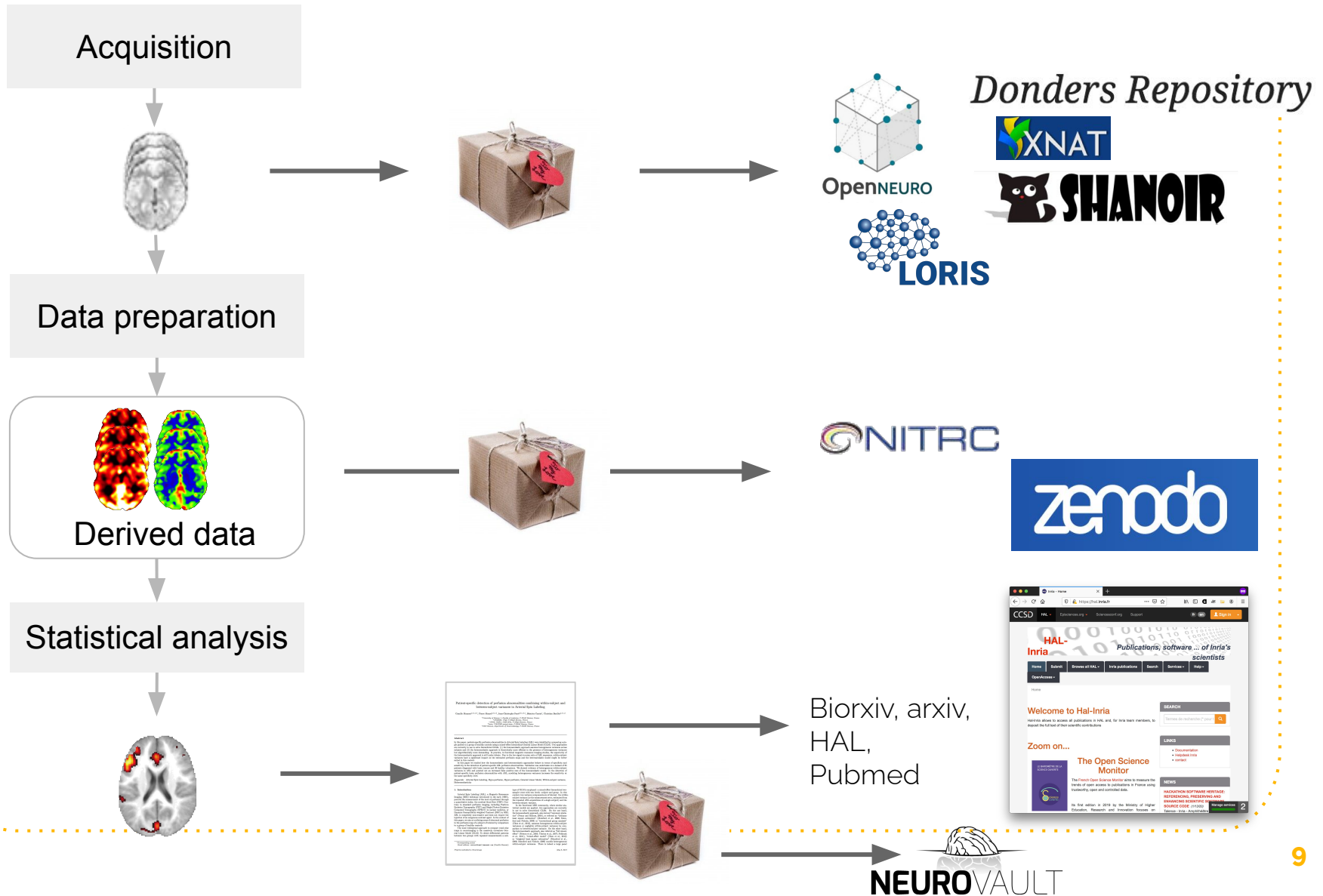


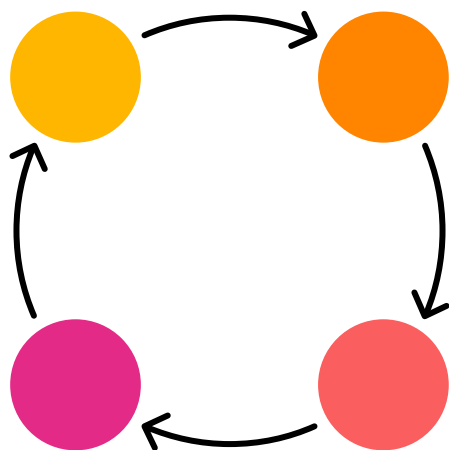
<http://bids.neuroimaging.io/bep028>

Sharing FAIR research products



Sharing FAIR research products





Reusing neuroimaging
derived data?

Camille Maumet

Thank you!

 @cmaumet.bsky.social

 @cmaumet@fediscience.org