HAL
open science

Tools and standards to make neuroimaging derived data reusable

Camille Maumet

To cite this version:

Camille Maumet. Tools and standards to make neuroimaging derived data reusable. Neuroinformatics 2018, Aug 2018, Montreal, Canada. . inserm-01886089

HAL Id: inserm-01886089

https://inria.hal.science/inserm-01886089v1

Submitted on 2 Oct 2018

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.

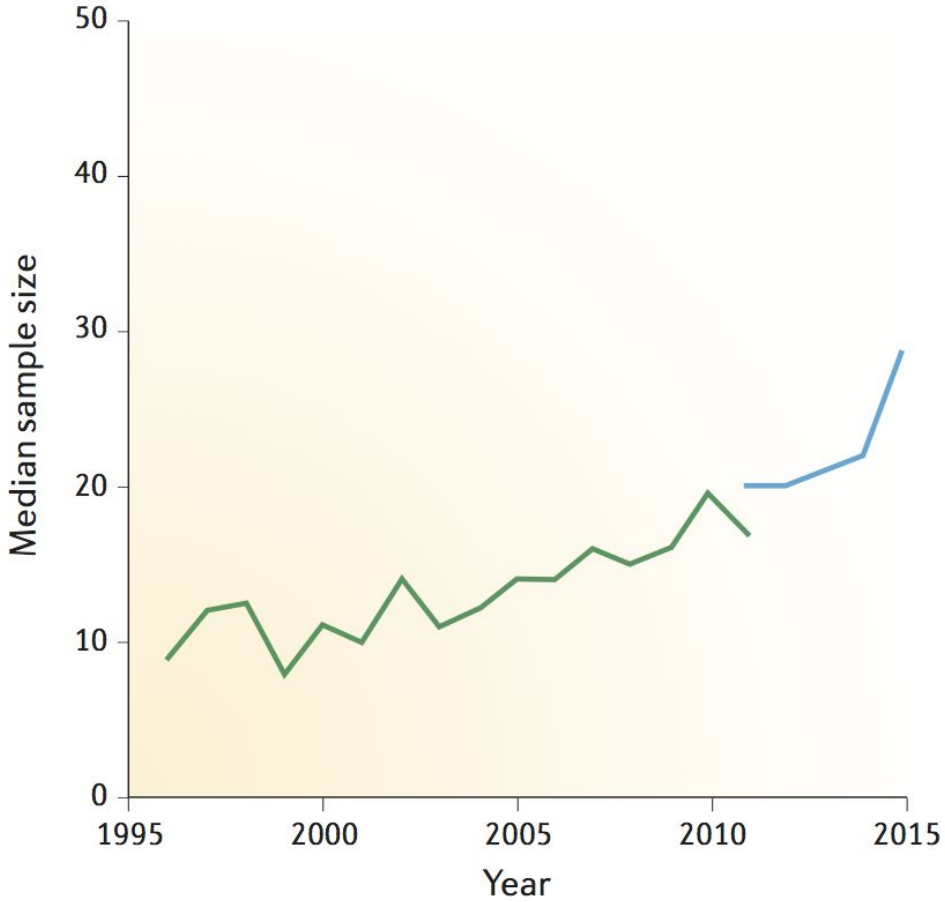
INCF Neuroinformatics 2018 – August 9th, 2018

Tools and standards to make neuroimaging derived data reusable

Camille Maumet @cmaumet

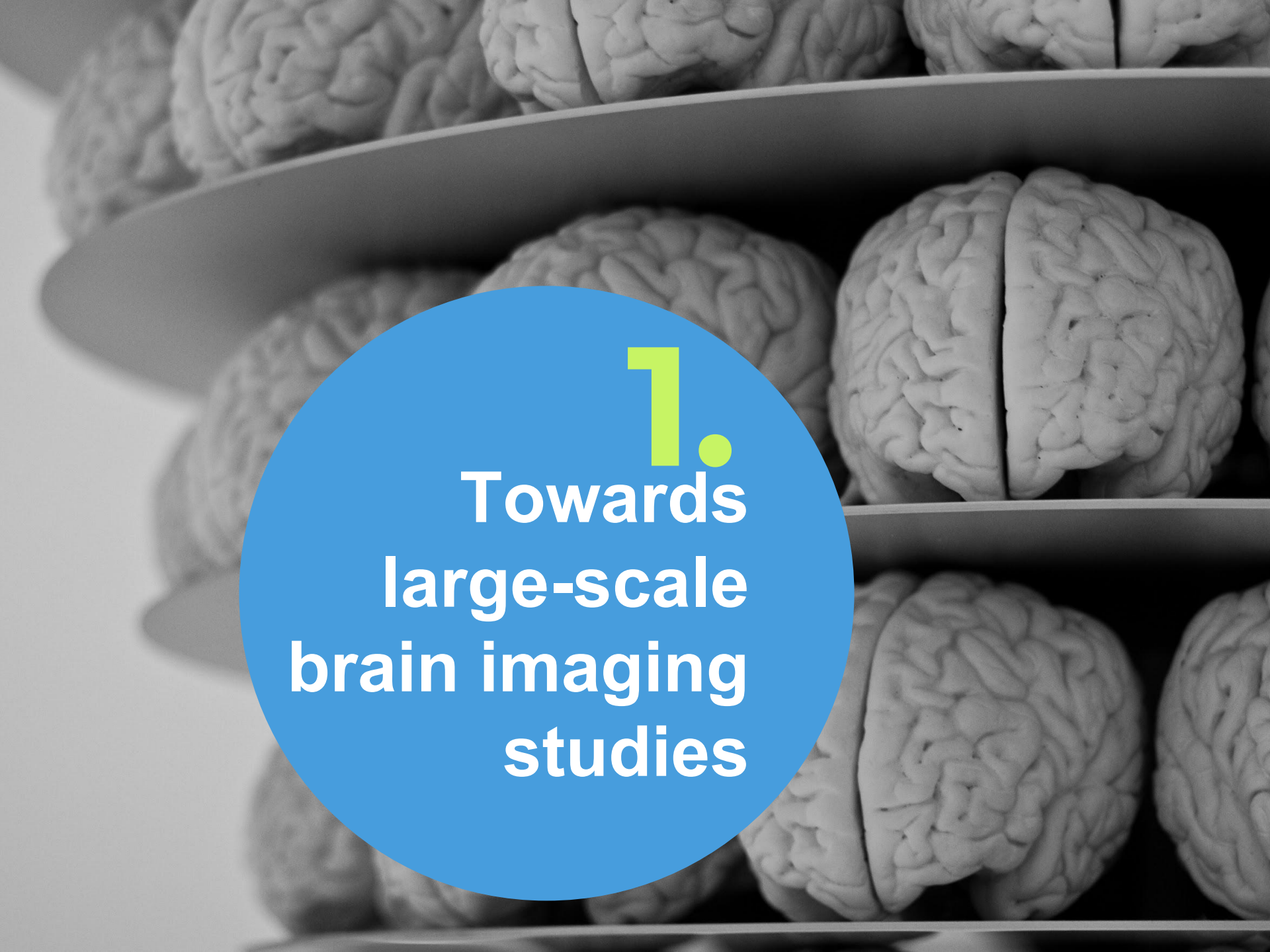
Univ Rennes, Inria, CNRS, Inserm, IRISA UMR 6074, VisAGeS ERL U-1228

Sample sizes in neuroimaging



Poldrack et. al (2017)

2015: 30 subjects / study



1. Towards large-scale brain imaging studies

More and more shared data available

Single study
30 subjects

Consortium
1000 subjects

Cohort
1 000 - 100 000 subjects

OpenNEURO

studyforrest.org

NEUROVAULT

MYCONNECTOME

ABIDE
Autism Brain Imaging Data Exchange

1000 Functional Connectomes Project

NITRC

ADHD 200

OSF

CORR
CONSORTIUM FOR RELIABILITY AND REPRODUCIBILITY

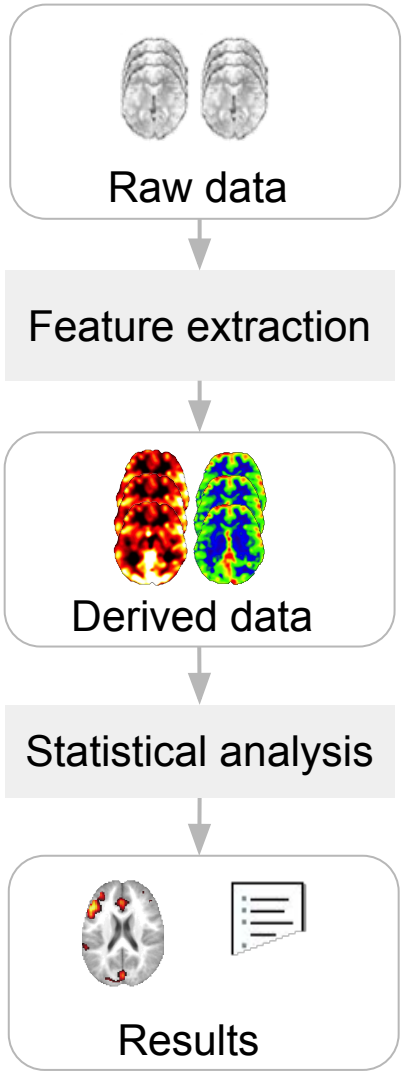
nki

HUMAN Connectome PROJECT

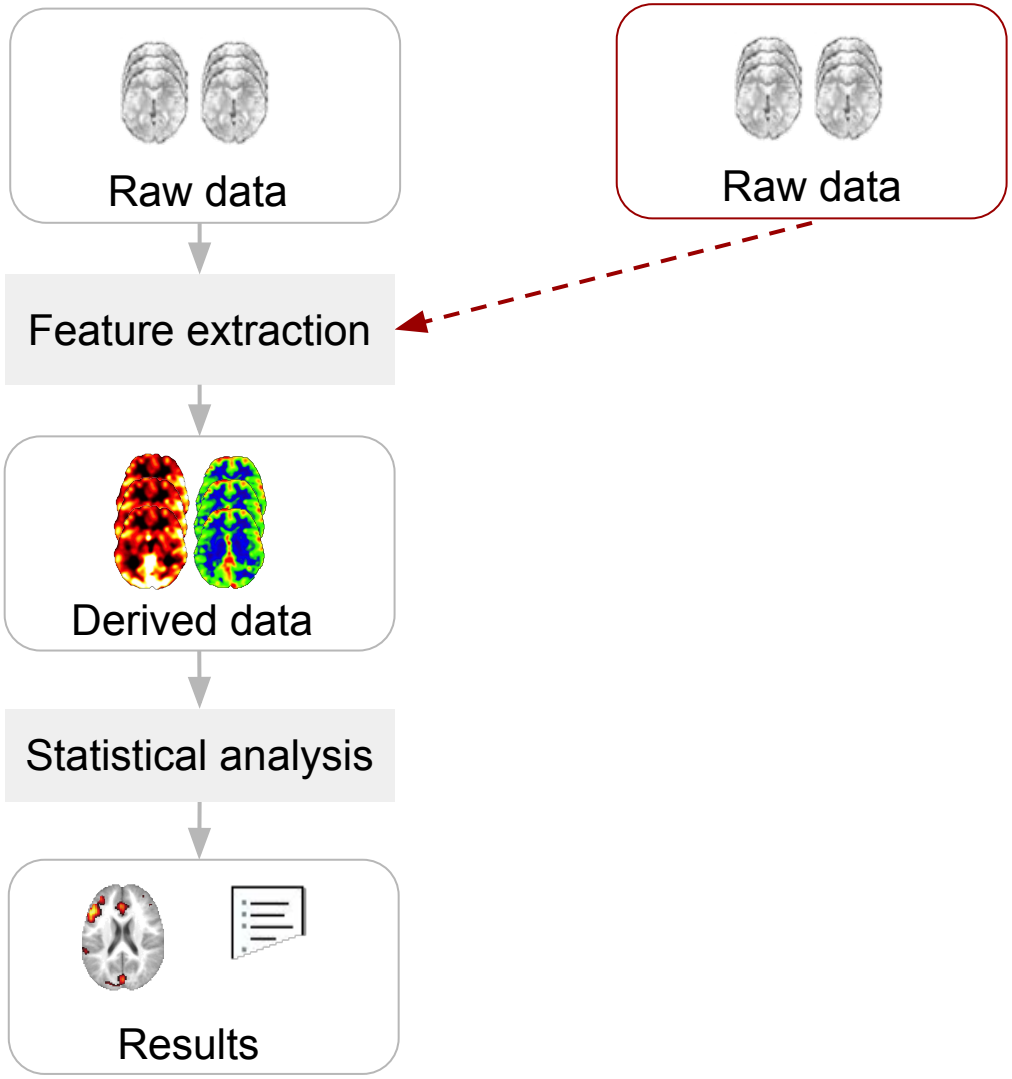
biobank uk

+ Images
+ Homogeneous
- Fewer datasets

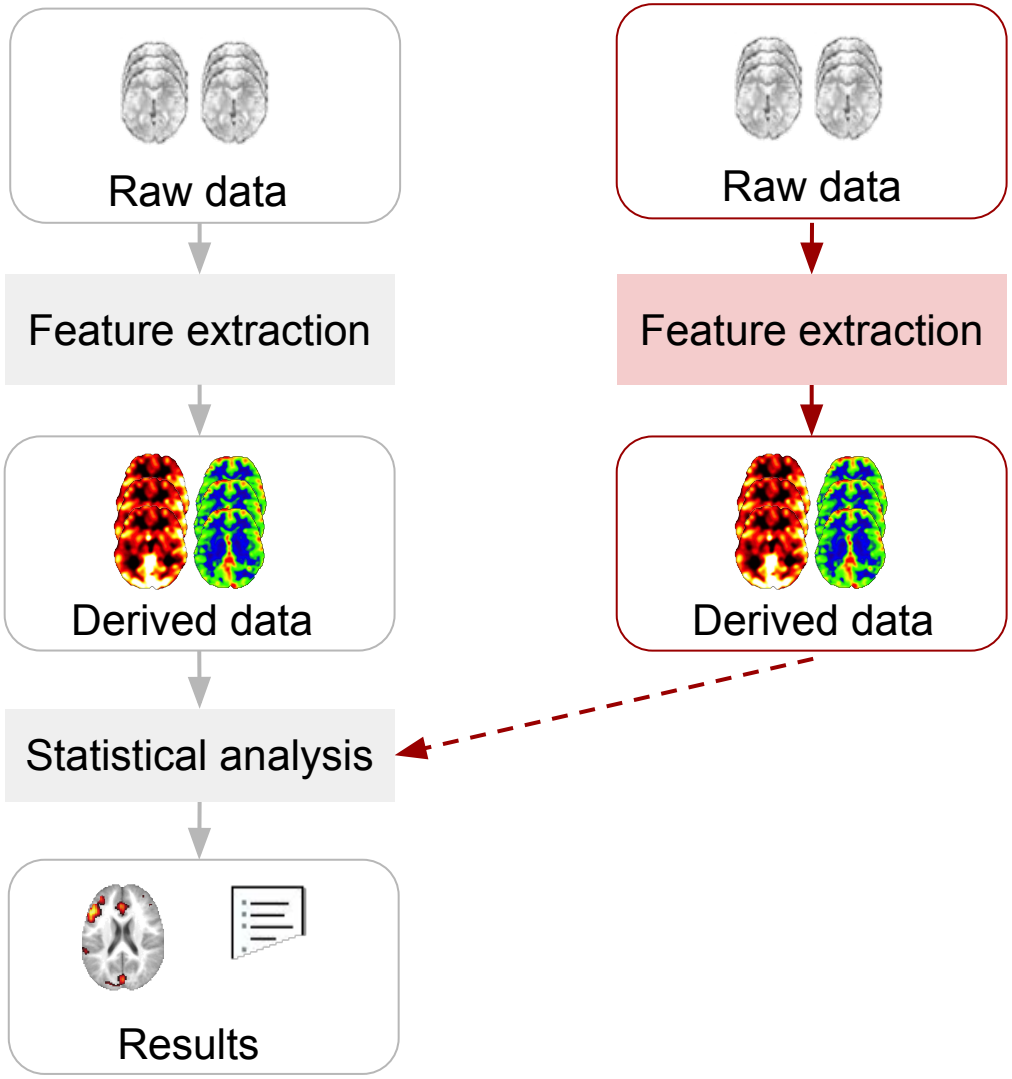
New challenges when working with open data



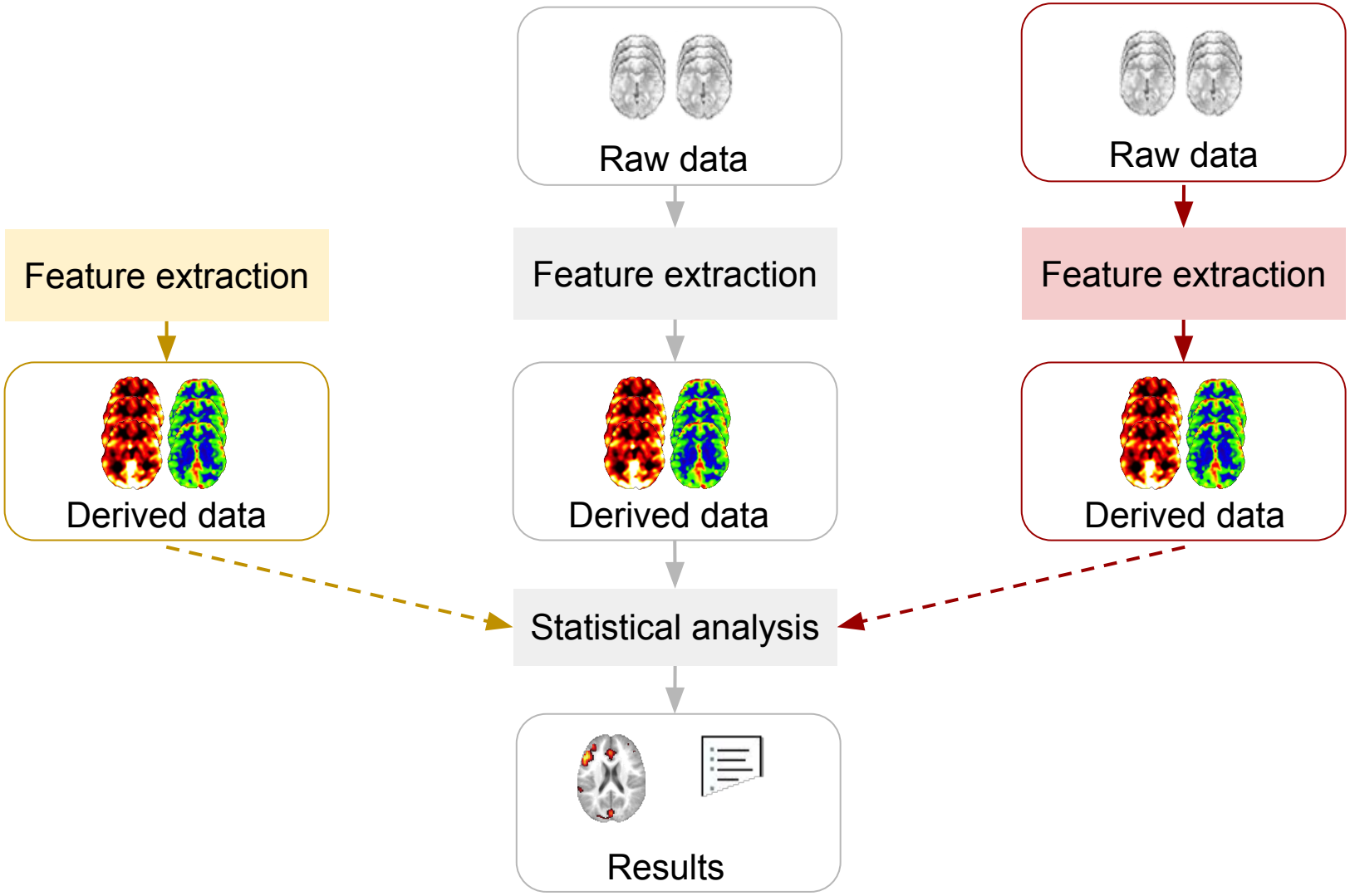
New challenges when working with open data



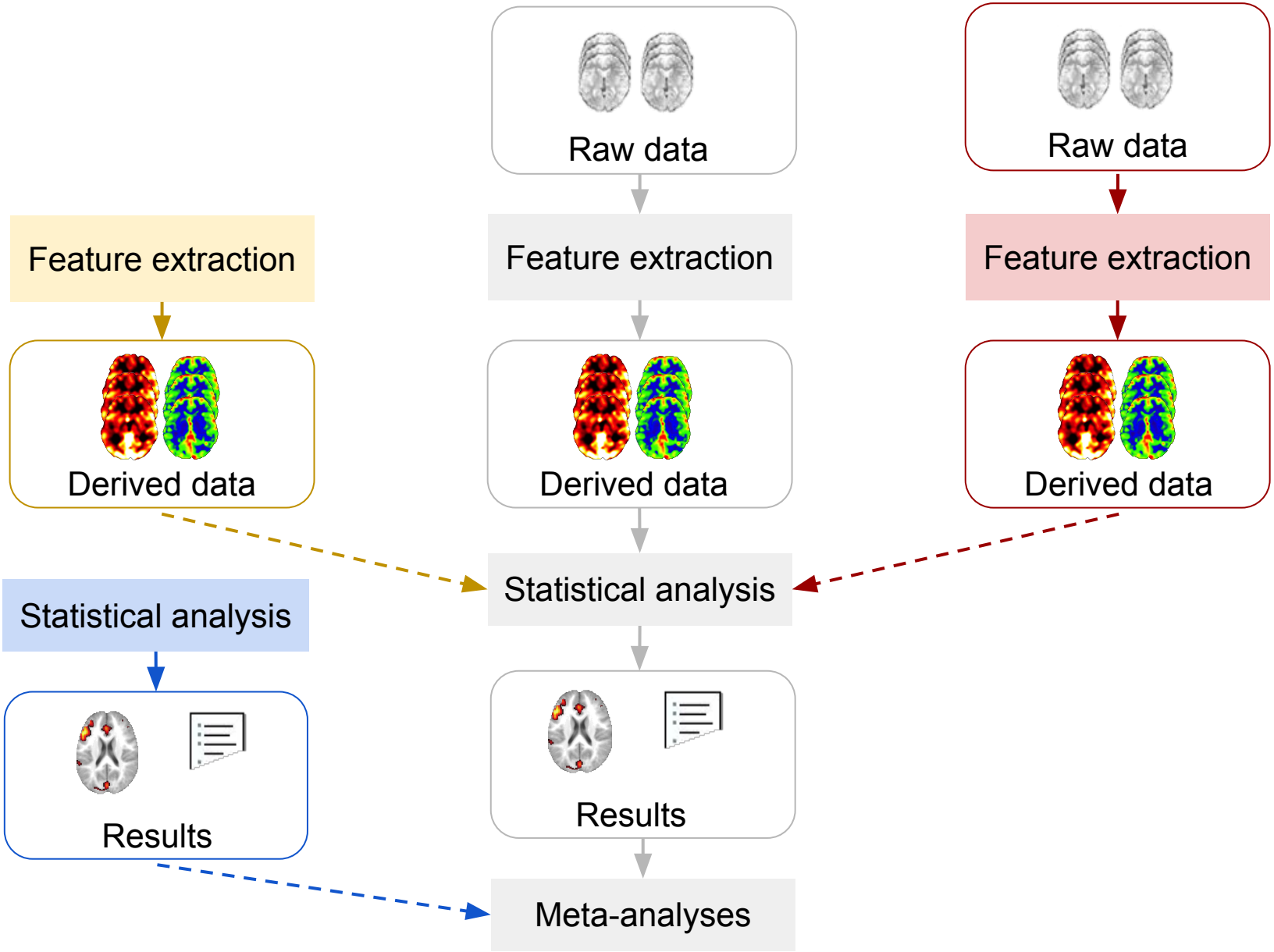
New challenges when working with open data



New challenges when working with open data



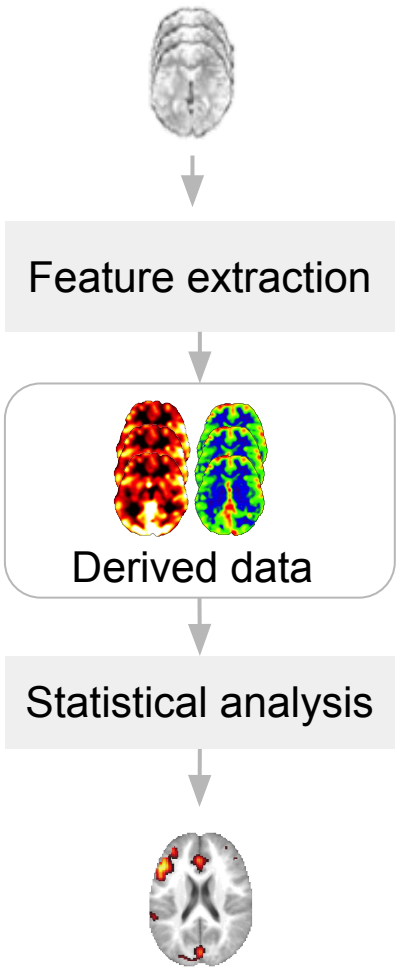
New challenges when working with open data



2.

Analytic variability?

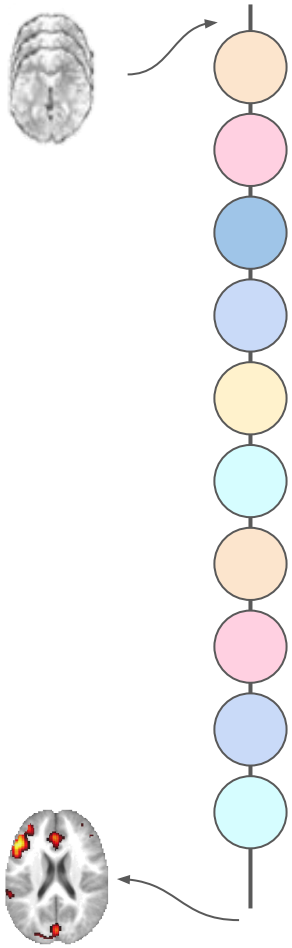
Analytic variability



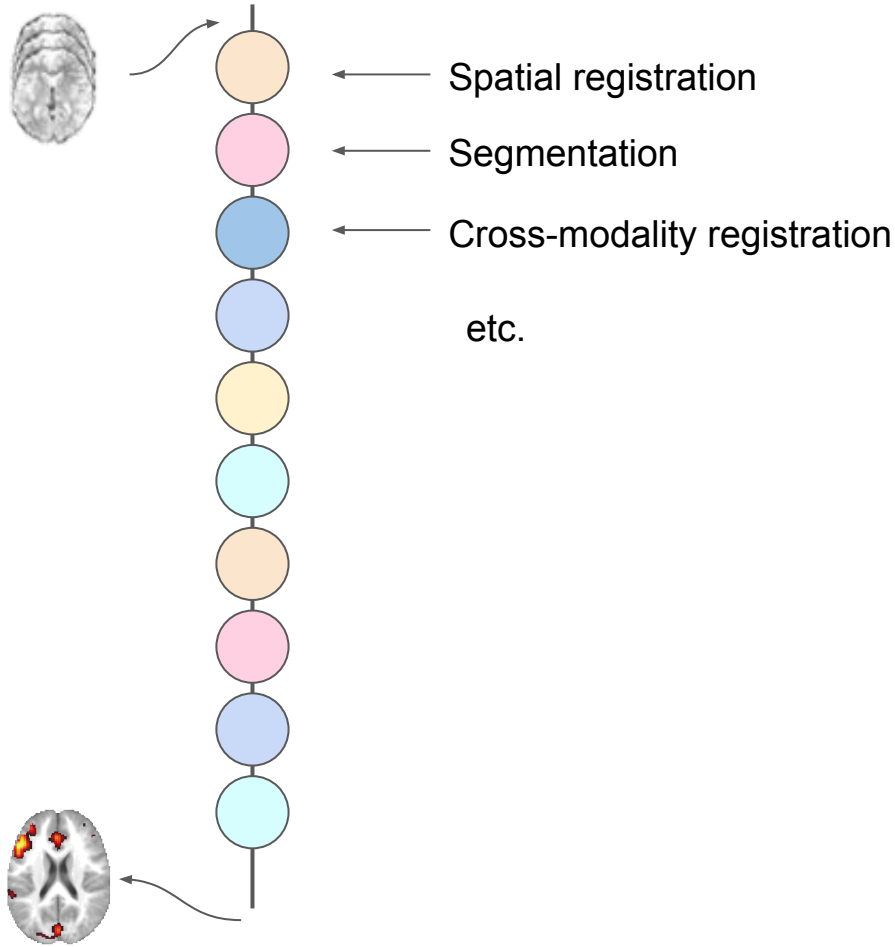
"Different **acceptable** analysis methods"

Carp et al. (2012)

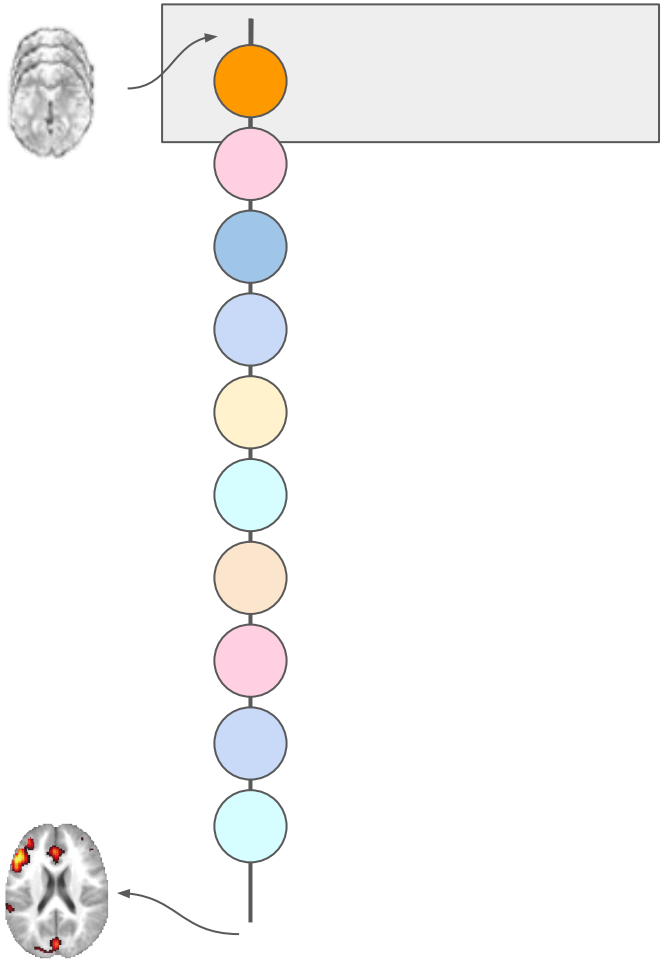
Analytic variability



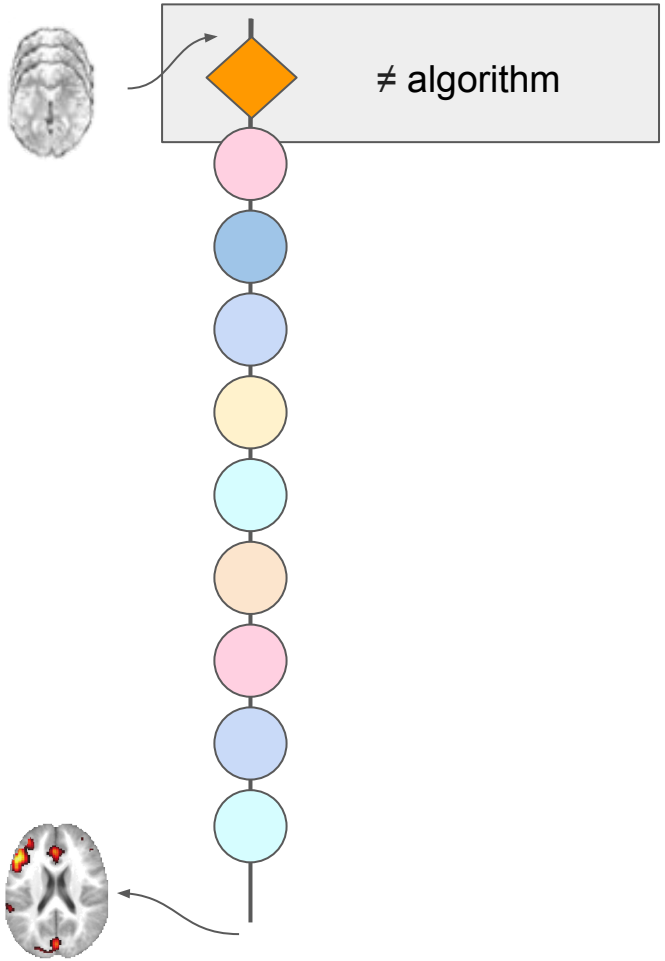
Analytic variability

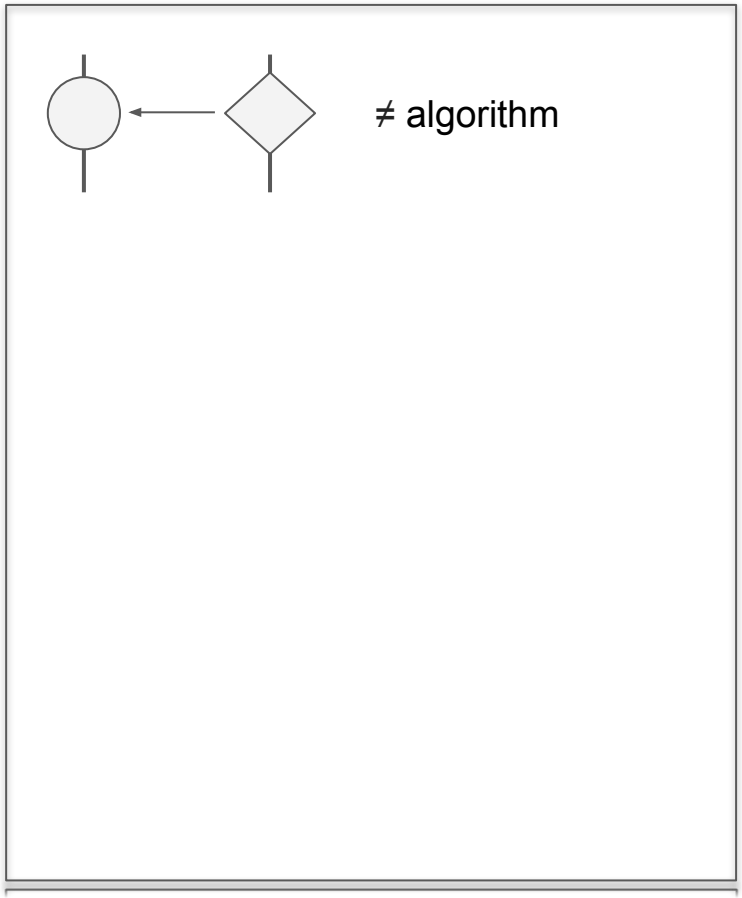


Analytic variability

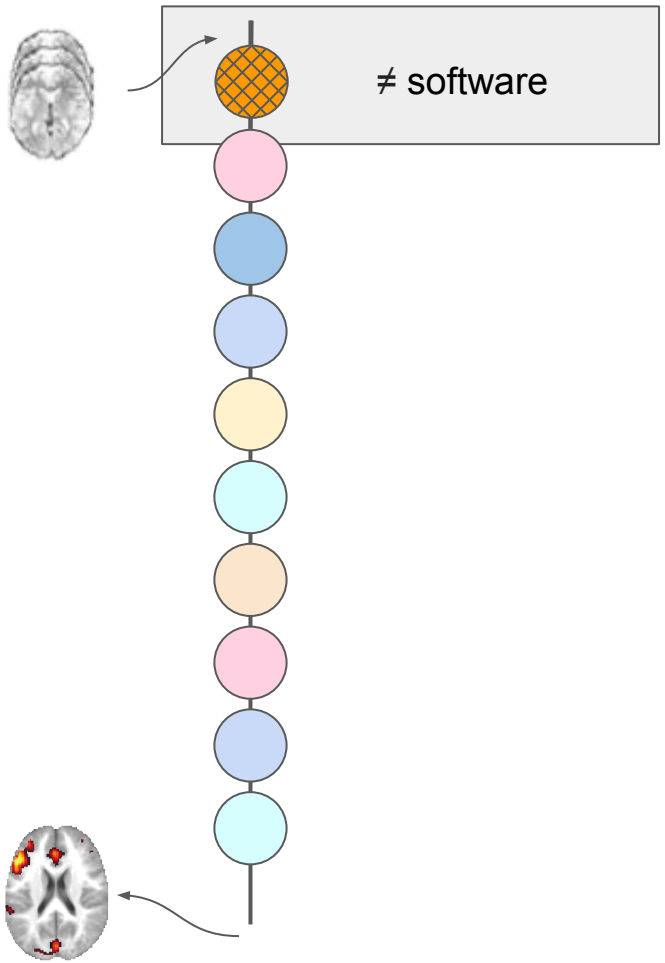


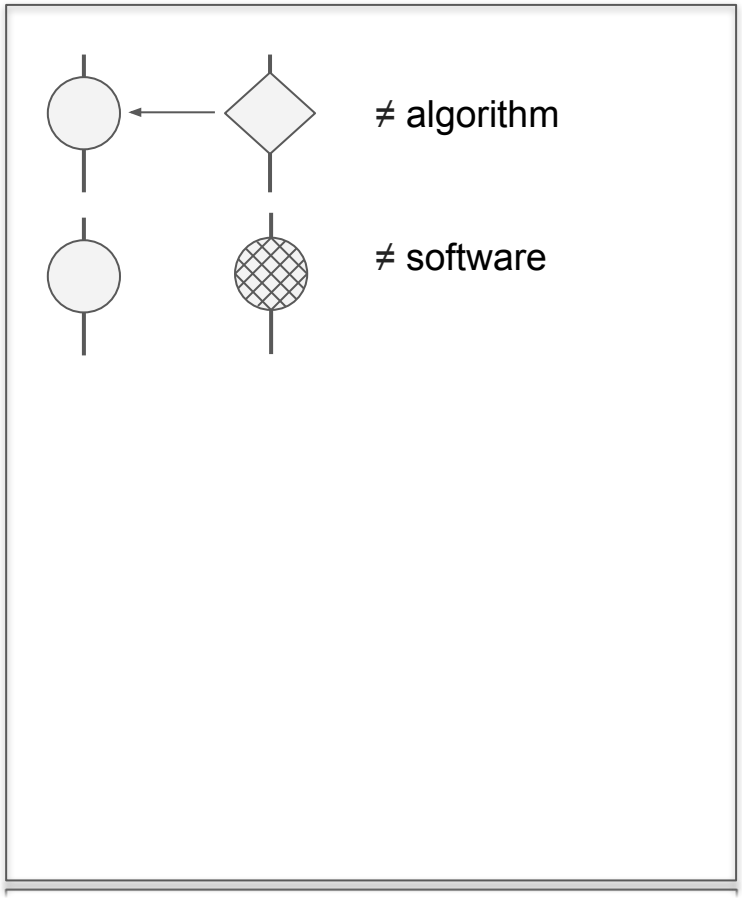
Analytic variability



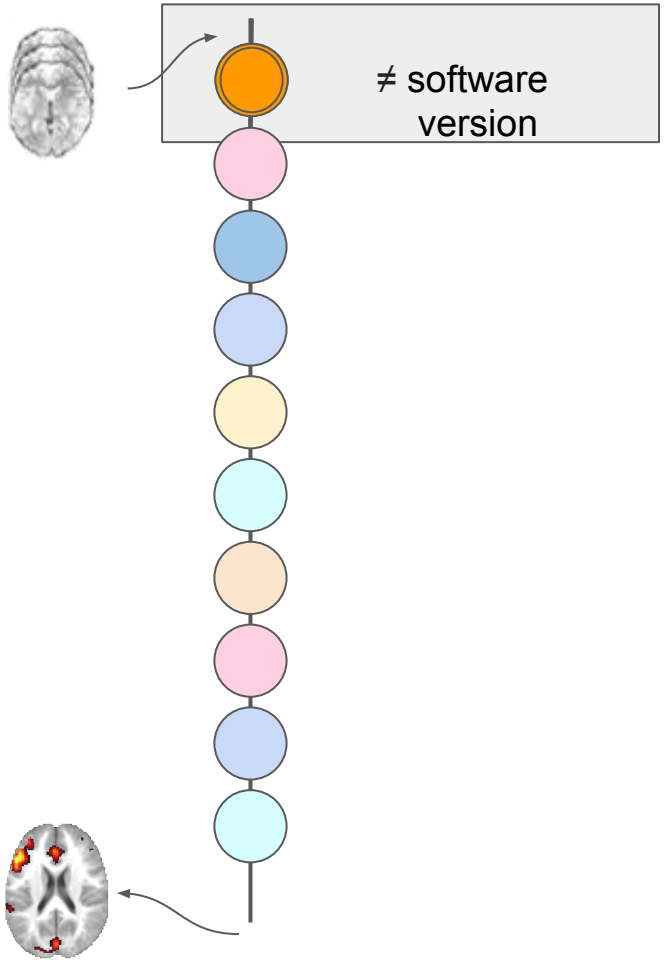


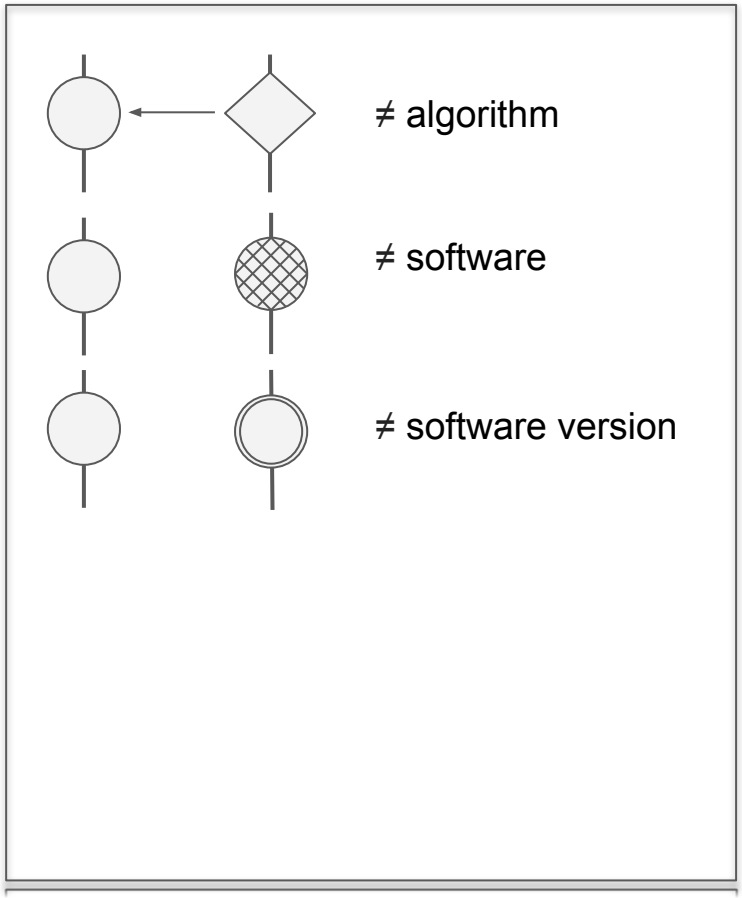
Analytic variability



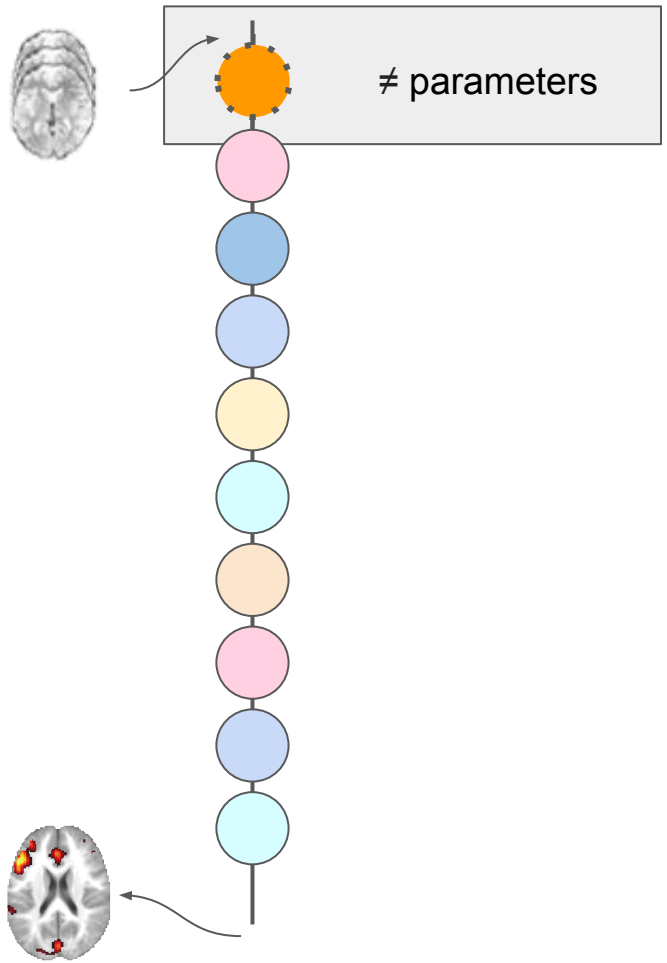


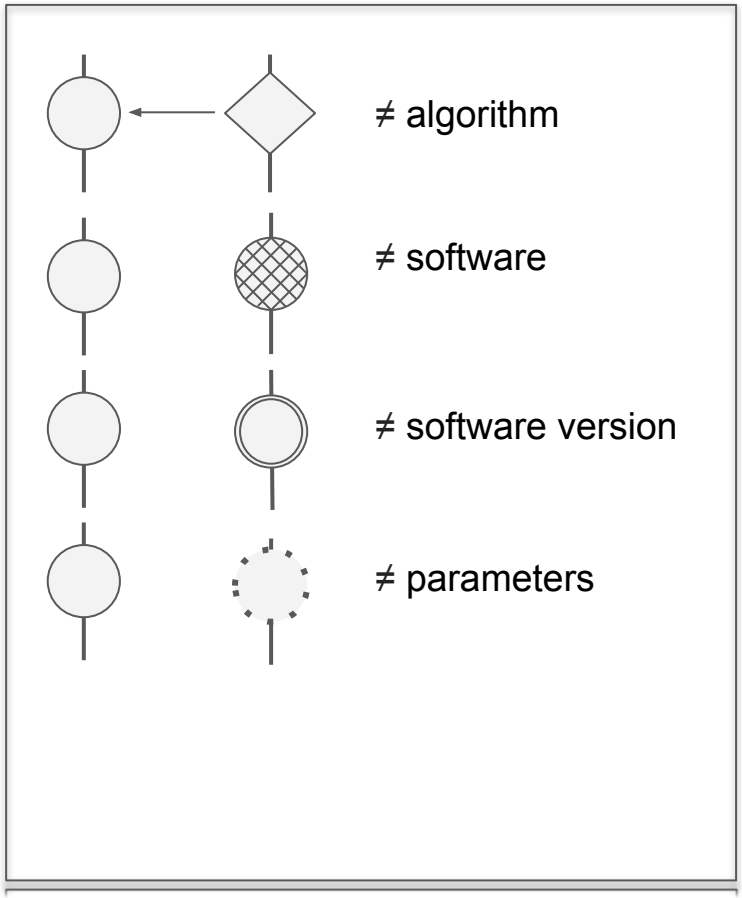
Analytic variability



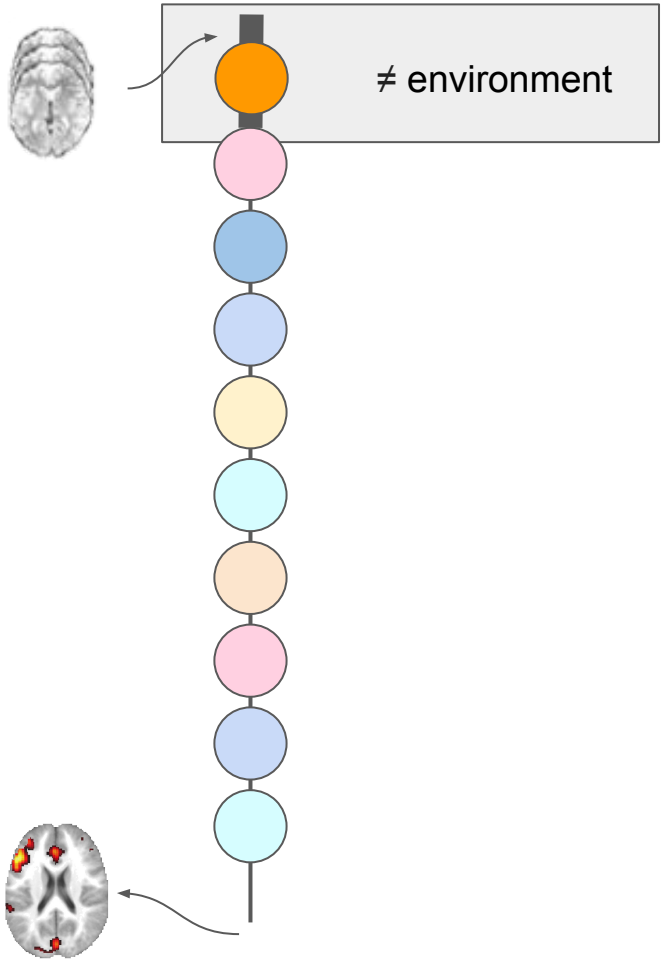


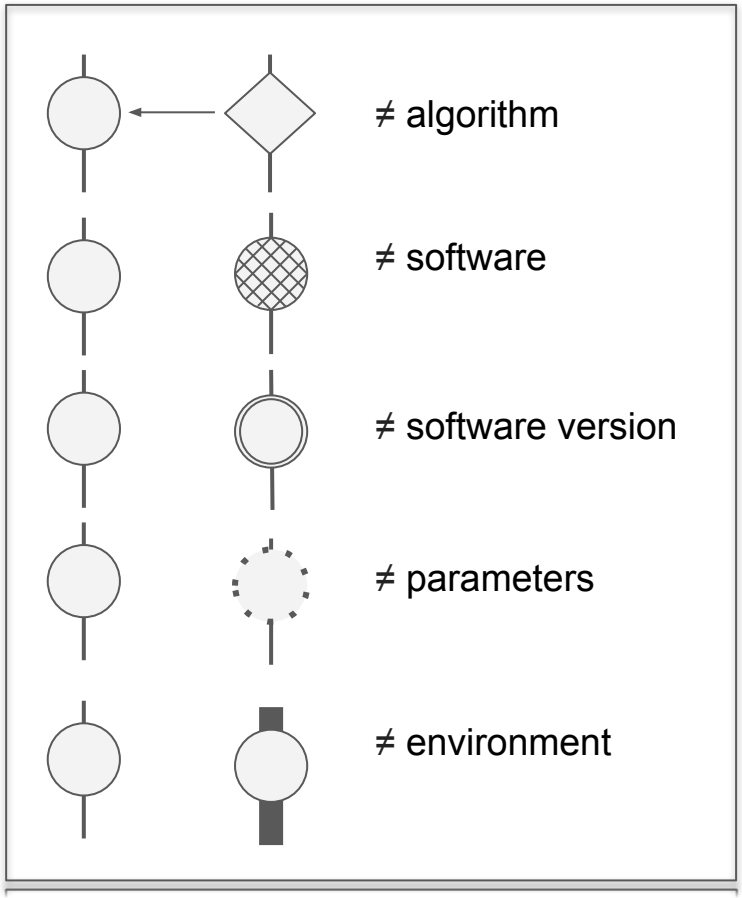
Analytic variability



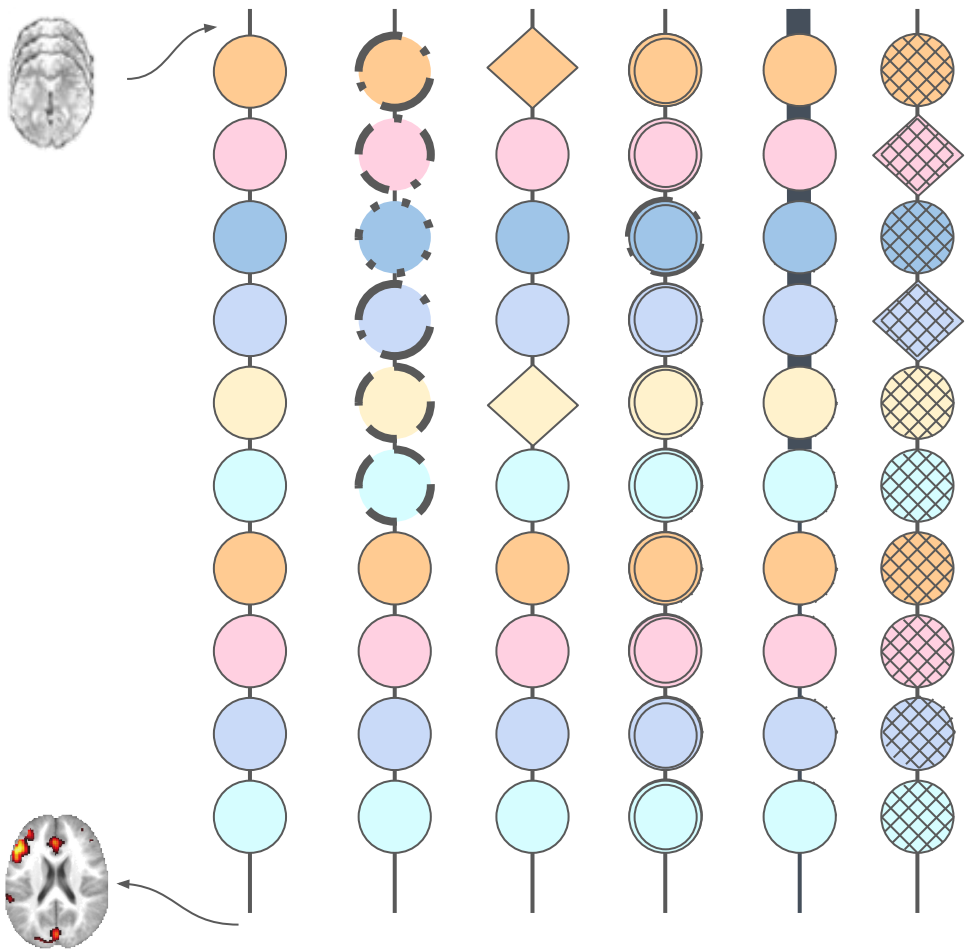


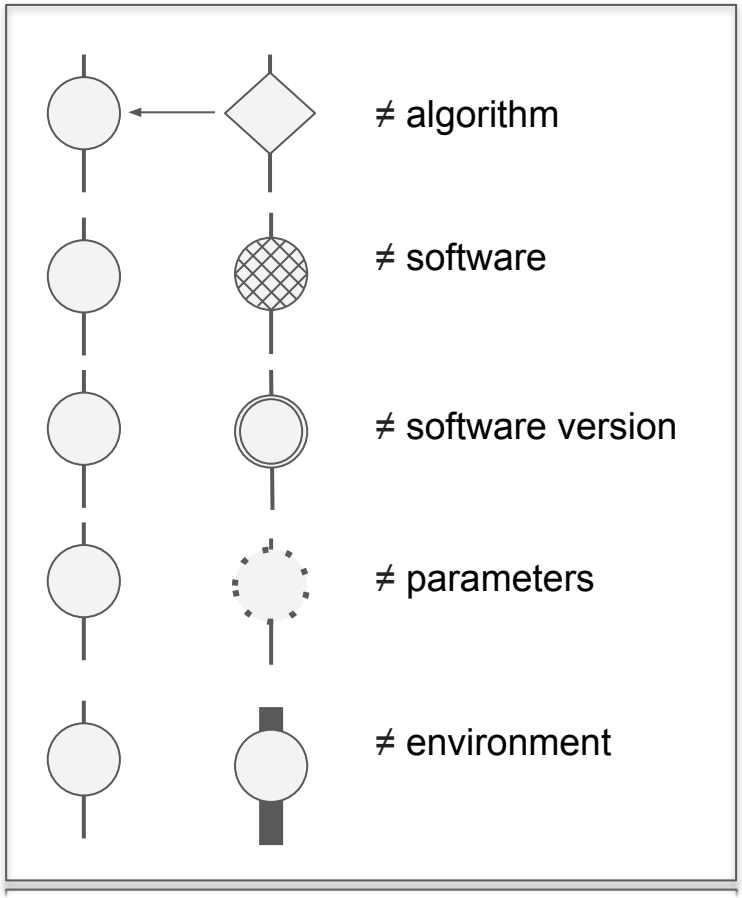
Analytic variability





Analytic variability





Question?

How does analytic variability impact neuroimaging results?

Background

- **Choice of Analysis Pipeline (functional MRI)**
  Carp (2012)
- **Choice of Analysis Software Version**
  Groenenschild et al (2012)
- **Choice of operating system**
  Glatard et al (2015)
- **Choice of analysis software package**
  Pauli et al (2016)
  Bowring et al (2018) (preprint)

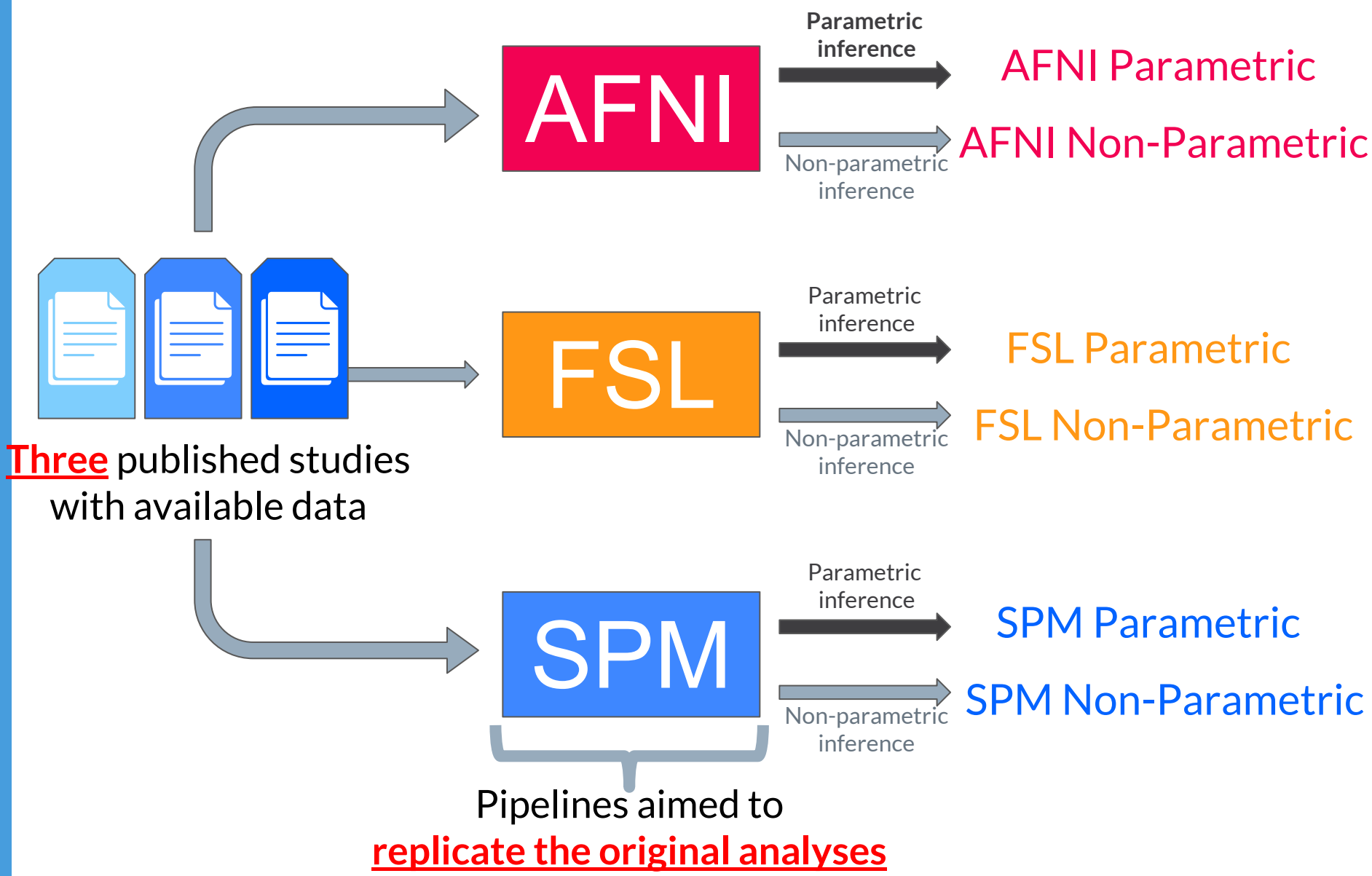
Slide by Alexander Bowring, adapted

Background

- **Choice of Analysis Pipeline (functional MRI)**
  Carp (2012)
- **Choice of Analysis Software Version**
  Groenenschild et al (2012)
- **Choice of operating system**
  Glatard et al (2015)
- **Choice of analysis software package**
  Pauli et al (2016)
  Bowring et al (2018) (preprint)

Joint work with Alexander Bowring and Thomas Nichols

Slide by Alexander Bowring, adapted

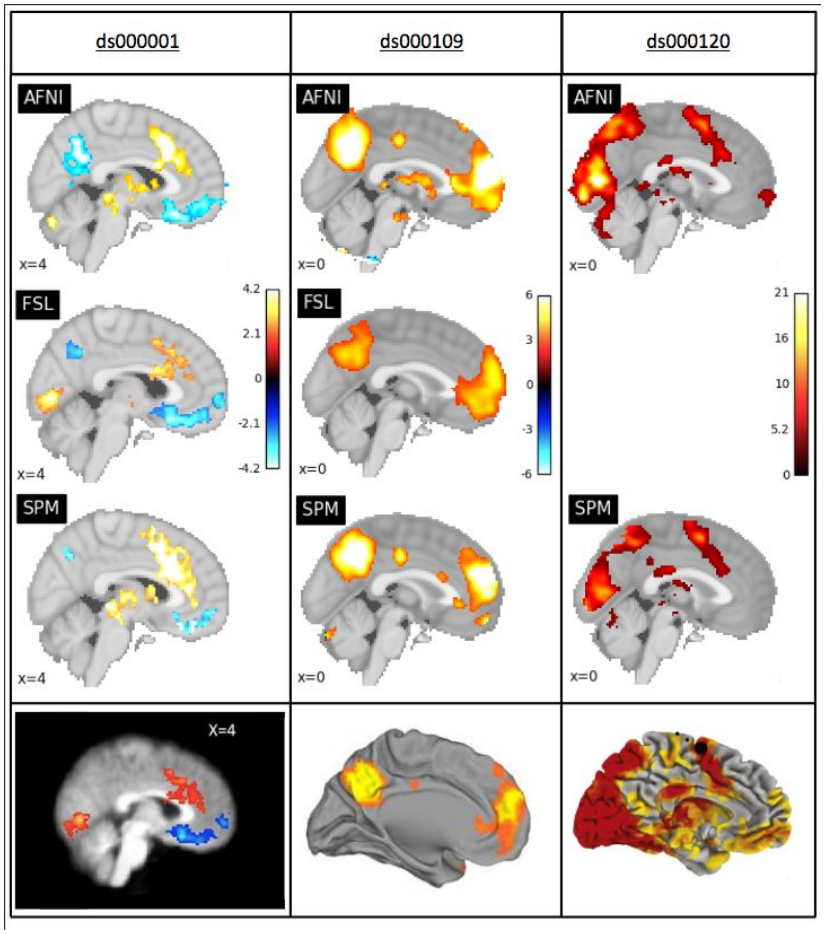


Bowring et al (2018) (preprint)

Slide by Alexander Bowring

Impact of Analysis Software on Task fMRI Results

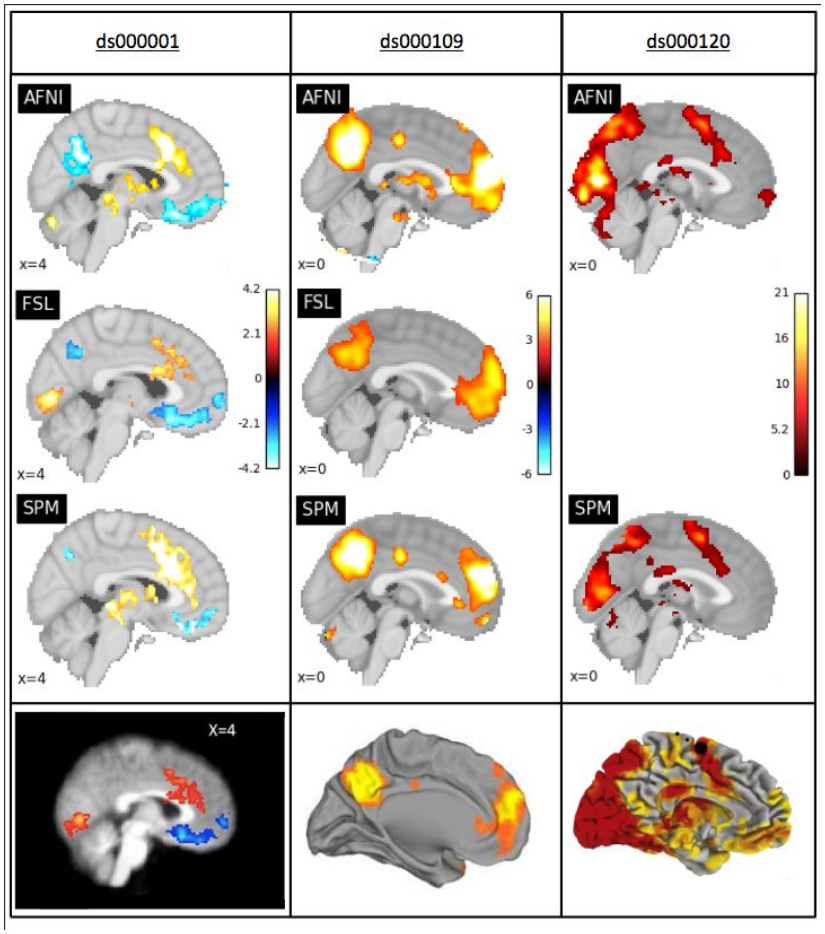
Reproducing the main figure



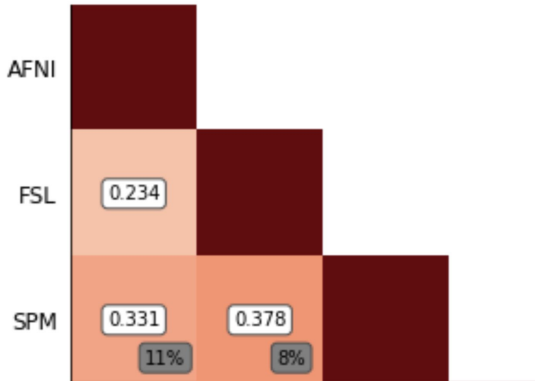
Bowring et al (2018) (preprint)

Impact of Analysis Software on Task fMRI Results

Reproducing the main figure



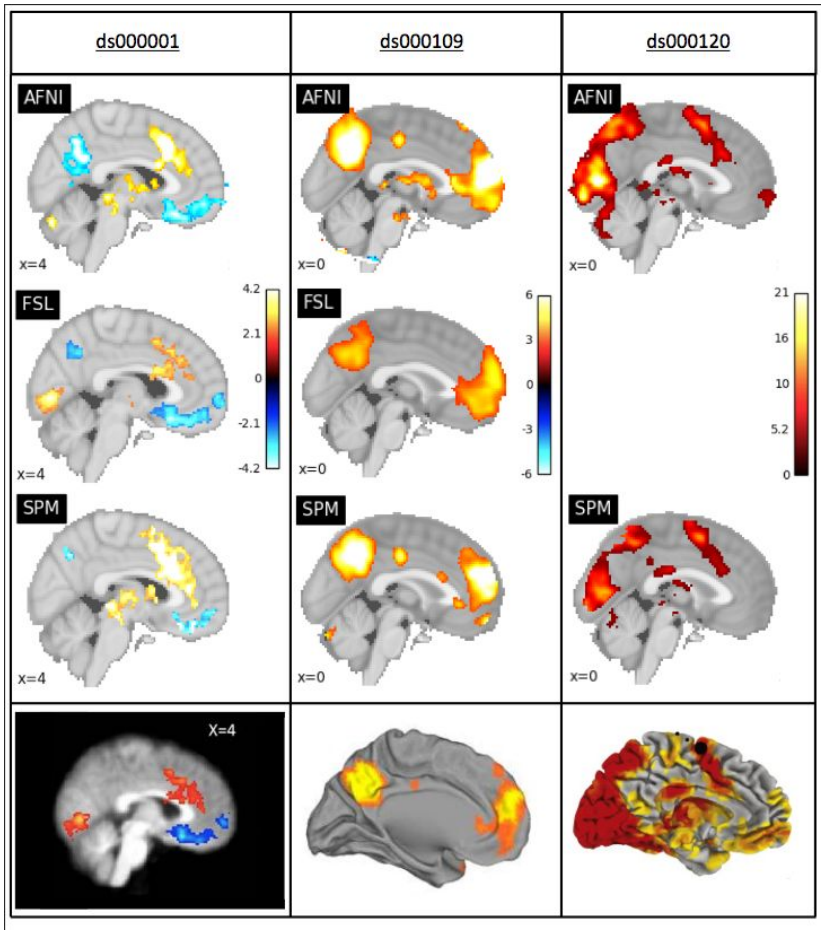
Dice coefficients



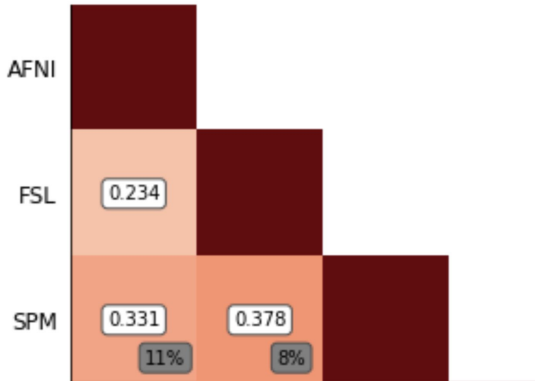
Bowring et al (2018) (preprint)

Impact of Analysis Software on Task fMRI Results

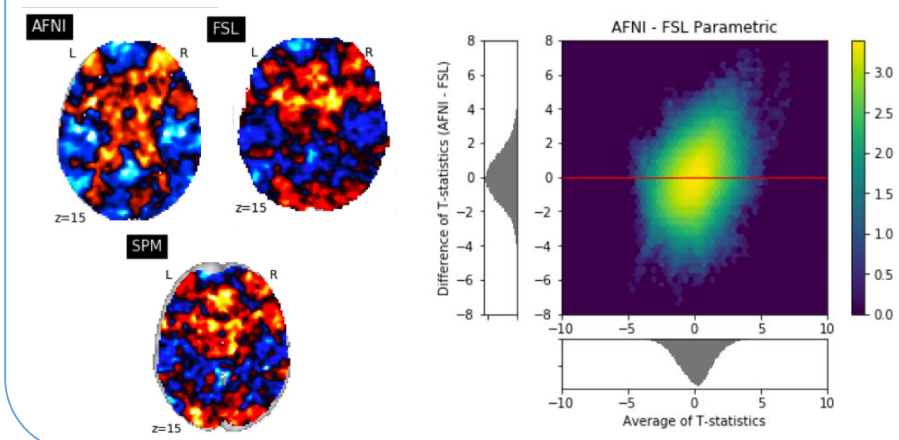
Reproducing the main figure



Dice coefficients



Unthresholded statistics

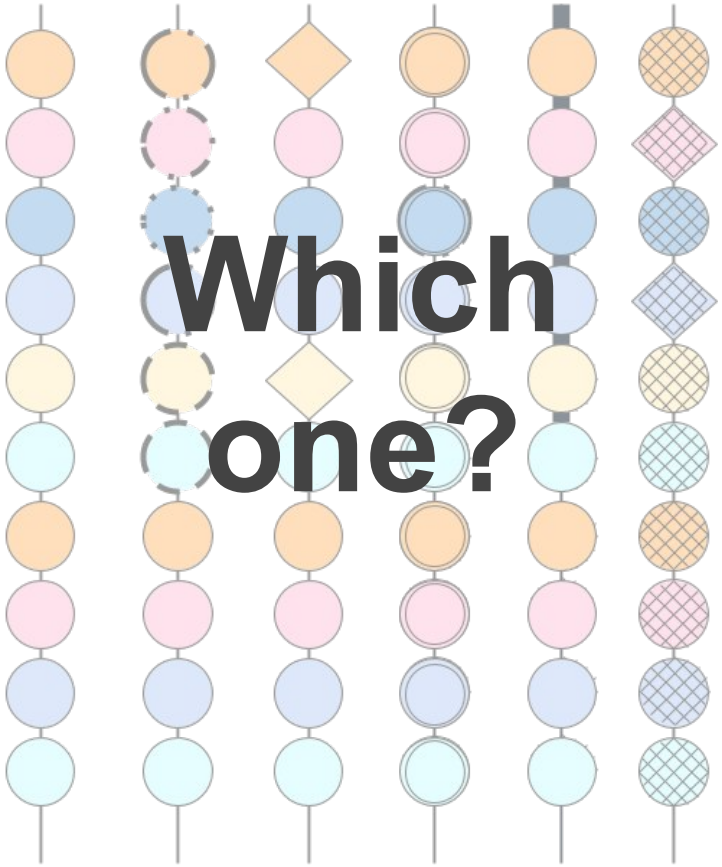


Bowring et al (2018) (preprint)

3.

Transparency

Disclosing analytical details



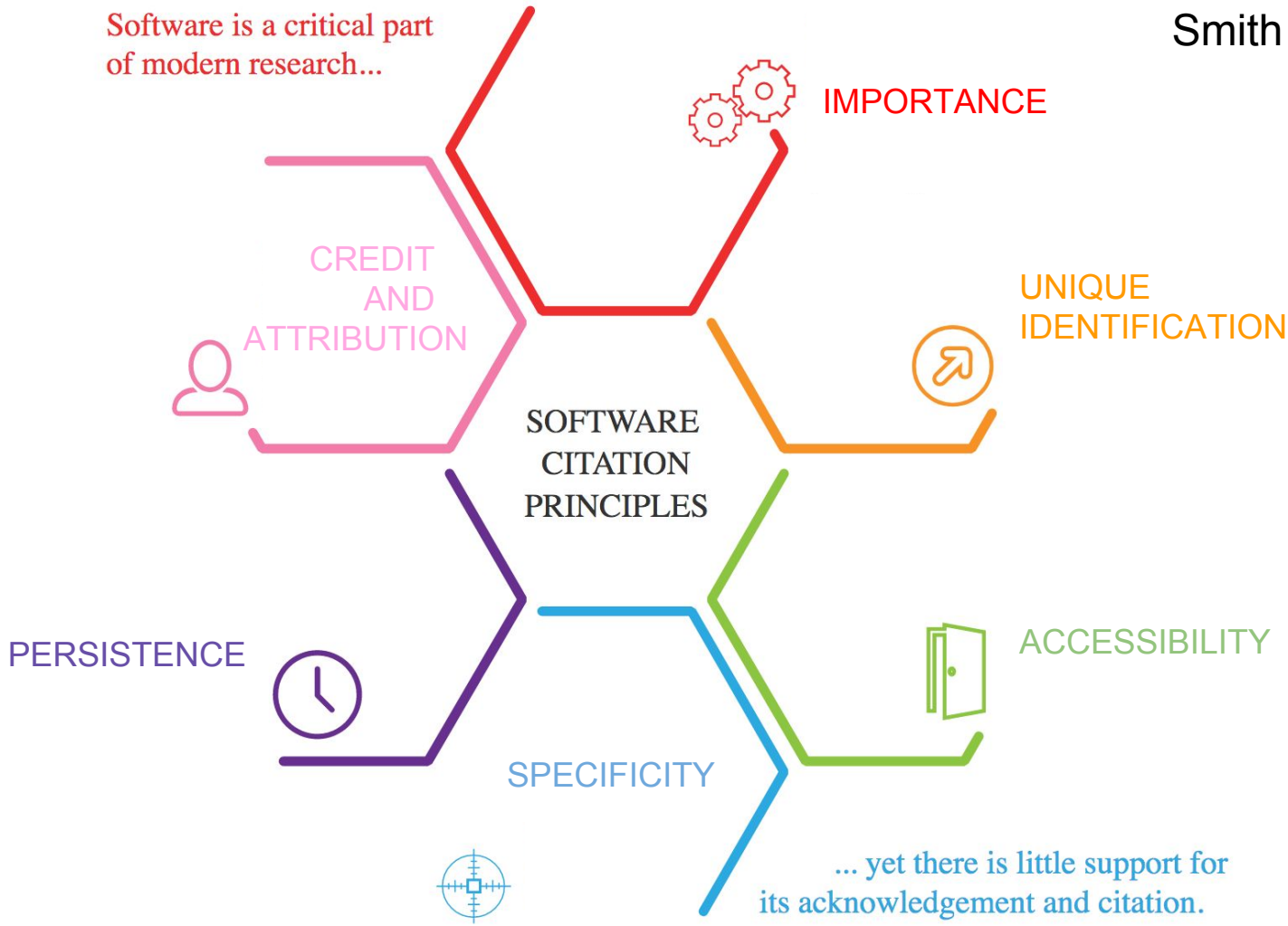
Details about the analysis pipeline are in the code.

In manuscripts, it is best practice to:

- Cite custom code
- Cite libraries and tools

How to cite software? The software citation principles

Smith et al (2016)



https://blog.datacite.org/software-citation-principles/ by Laura Rueda, adapted.

Citing custom code

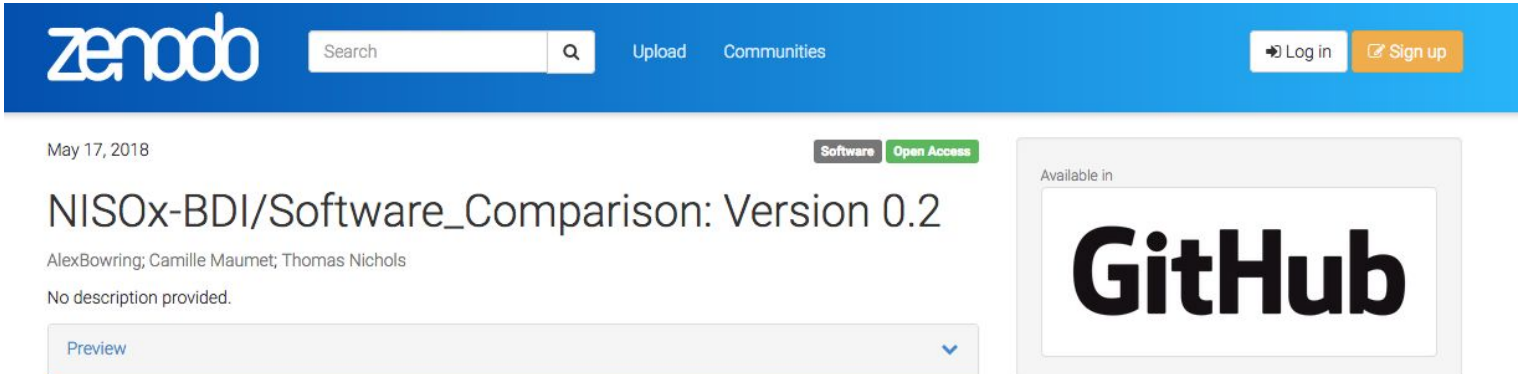
3. Results

Bowring et al (2018)

[...] All analysis scripts, results reports, and notebooks for each of study are available through Zenodo (Nielsen and Smith, 2014) at https://zenodo.org/record/1203654 (Alexander Bowring et al., 2018).

Persistence

Accessibility



Software Heritage

Let others cite your code

zenodo

Search | Upload | Communities | Log in | Sign up

May 17, 2018

Software | Open Access

NISOx-BDI/Software_Comparison: Version 0.2

AlexBowring; Camille Maumet; Thomas Nichols

No description provided.

Attribution

Preview

Software_Comparison-0.2.0.zip

- Software_Comparison-0.2.0
  - .gitignore 482 Bytes
  - README.md 7.4 kB
  - ds001
    - AFNI
      - LEVEL2
        - euler_chars.csv 941 Bytes
        - permutation_test
          - euler_chars.csv 933 Bytes
      - SCRIPTS
        - level2.sh 7.7 kB
        - permutation_test.sh 4.8 kB
        - sub-01_level1.sh 4.3 kB
        - sub-02_level1.sh 4.3 kB
        - sub-03_level1.sh 4.3 kB
        - sub-04_level1.sh 4.3 kB
        - sub-05_level1.sh 4.3 kB
        - sub-06_level1.sh 4.3 kB

Files (11.3 MB)

Name	Size	
Software_Comparison-0.2.0.zip	11.3 MB	Preview Download

md5:418825948cf76d92c4943714fcc96af0

Available in

GitHub

Publication date:
May 17, 2018

DOI:
DOI 10.5281/zenodo.1

Unique identifier

Related identifiers:
Supplement to:
https://github.com/NISOx-BDI/Software_Comparison/tree/v0.1

License (for files):
Other (Open)

Specificity

Versions

Version 0.2	10.5281/zenodo.1248809	May 17, 2018
Version v0.1	10.5281/zenodo.1203655	Mar 19, 2018

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.1203654. This DOI represents all versions, and will always resolve to the latest one. Read more.

Share

Citing libraries and tools

Bowring et al (2018)

2.2 Data Analyses

All data analyses were conducted using AFNI 17.0.18 (RRID:SCR_005927, (Cox, 1996)), FSL 5.0.10 (RRID:SCR_002823, (Jenkinson et al., 2012)), and SPM12 v6906 (RRID:SCR_007037, (Penny et al., 2011)).

Citing libraries and tools

2.2 Data Analyses

Bowring et al (2018)

All data analyses were conducted using AFNI 17.0.18 (RRID:SCR_005927, (Cox, 1996)), FSL 5.0.10 (RRID:SCR_002823, (Jenkinson et al., 2012)), and SPM12 v6906 (RRID:SCR_007037, (Penny et al., 2011)).

Specificity

Citing libraries and tools

2.2 Data Analyses

Bowring et al (2018)

All data analyses were conducted using AFNI 17.0.18 (RRID:SCR_005927, (Cox, 1996)), FSL 5.0.10 (RRID:SCR_002823, (Jenkinson et al., 2012)), and SPM12 v6906 (RRID:SCR_007037, (Penny et al., 2011)).

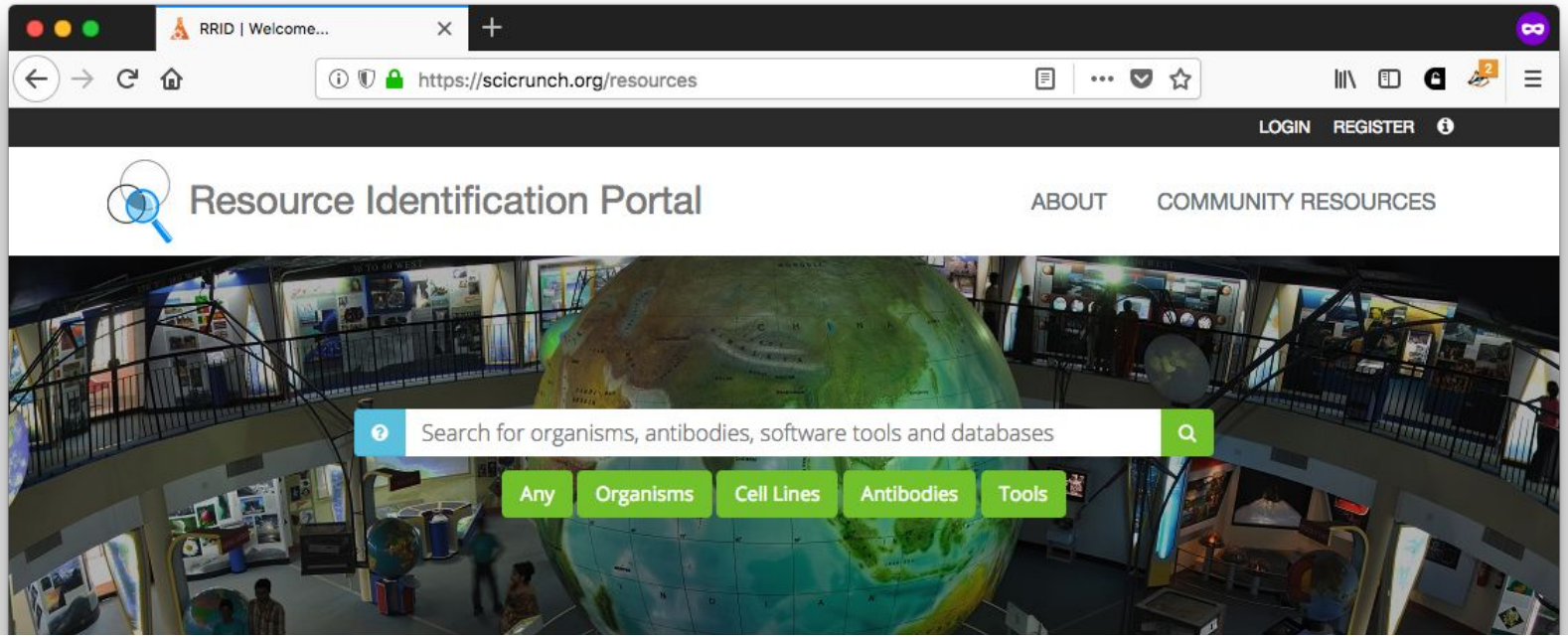
Specificity | Unique identifier

Citing libraries and tools

Bowring et al (2018)

2.2 Data Analyses

All data analyses were conducted using AFNI 17.0.18 (RRID:SCR_005927, (Cox, 1996)), FSL 5.0.10 (RRID:SCR_002823, (Jenkinson et al., 2012)), and SPM12 v6906 (RRID:SCR_007037, (Penny et al., 2011)).



https://scicrunch.org/resources

Bandrowski and Martone (2016)

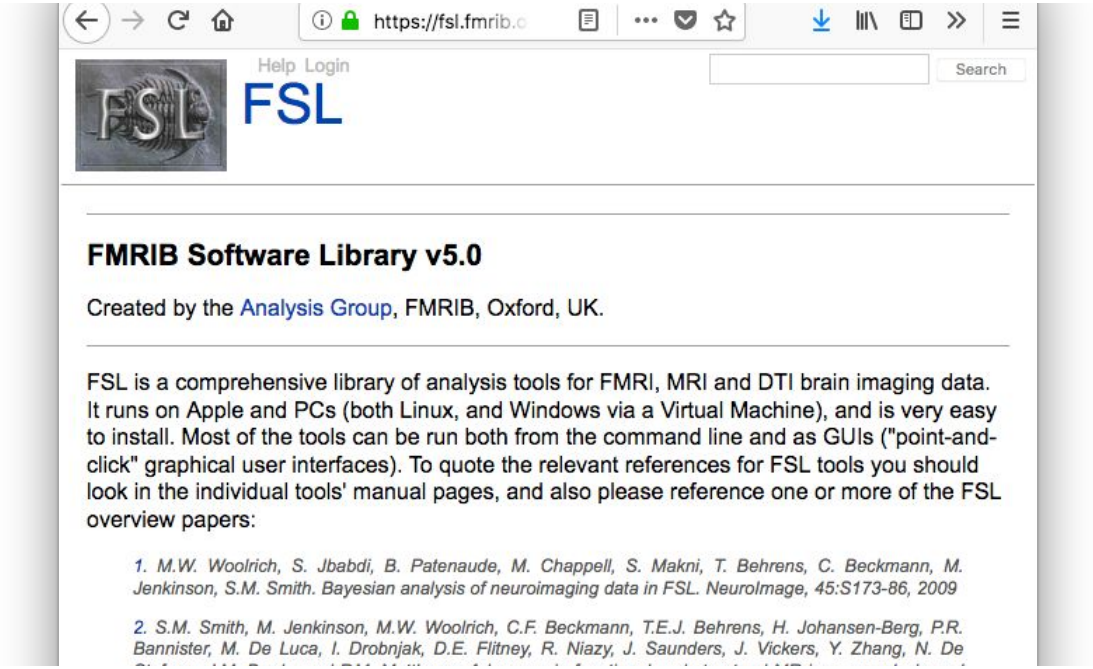
Specificity | Unique identifier | Attribution

Citing libraries and tools

2.2 Data Analyses

Bowring et al (2018)

All data analyses were conducted using AFNI 17.0.18 (RRID:SCR_005927, (Cox, 1996)), FSL 5.0.10 (RRID:SCR_002823, (Jenkinson et al., 2012)), and SPM12 v6906 (RRID:SCR_007037, (Penny et al., 2011)).



https://fsl.fmrib.ox.ac.uk/fsl/fslwiki

Specificity | Unique identifier | Attribution

Citing libraries and tools

2.2 Data Analyses

Bowring et al (2018)

All data analyses were conducted using AFNI 17.0.18 (RRID:SCR_005927, (Cox, 1996)), FSL 5.0.10 (RRID:SCR_002823, (Jenkinson et al., 2012)), and SPM12 v6906 (RRID:SCR_007037, (Penny et al., 2011)).

Also applies to other (non-neuroimaging) tools, repositories and libraries...

2.5 Scripting of analyses and figures

Bowring et al (2018)

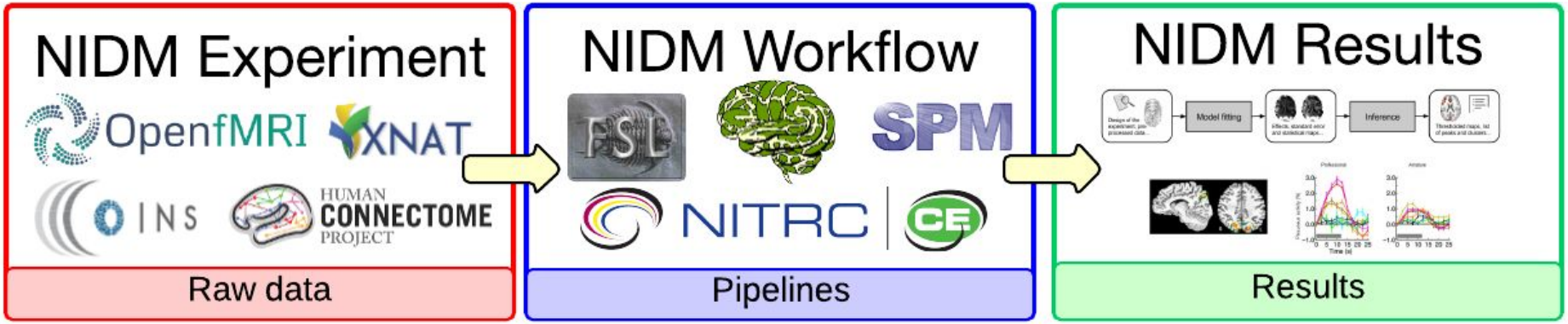
A Python Jupyter Notebook (Kluyver et al., 2016) was created for each of the three studies. Each notebook harvests our results data from Neurovault (RRID:SCR_003806) and applies the variety of methods discussed in the previous section using NiBabel 2.2.0 (Brett et al., 2017), NumPy 1.13.3 (Walt et al., 2011) and Pandas 0.20.3 (McKinney and Others, 2010) packages. Figures were created using Matplotlib 2.1.0 (Hunter, 2007) and Nilearn 0.4.0 (Abraham et al., 2014).

Provenance

“Provenance is information about entities, activities, and people involved in producing a piece of data or thing” — W3C PROV

Machine-readable provenance

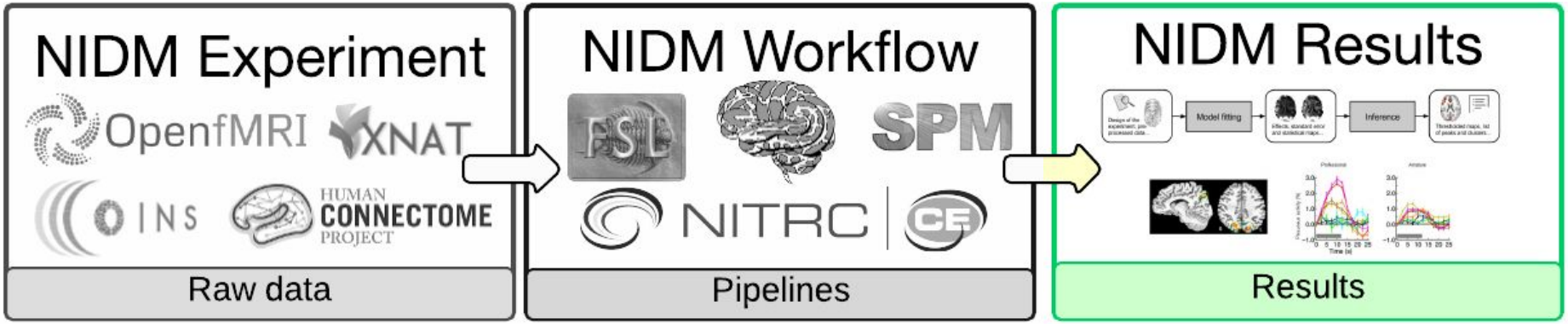
Machine-readable provenance



NIDM: neuroimaging data model

International collaboration with INCF SIG on Reproducibility and Best Practices in Human Brain Imaging

Machine-readable provenance



NIDM: neuroimaging data model

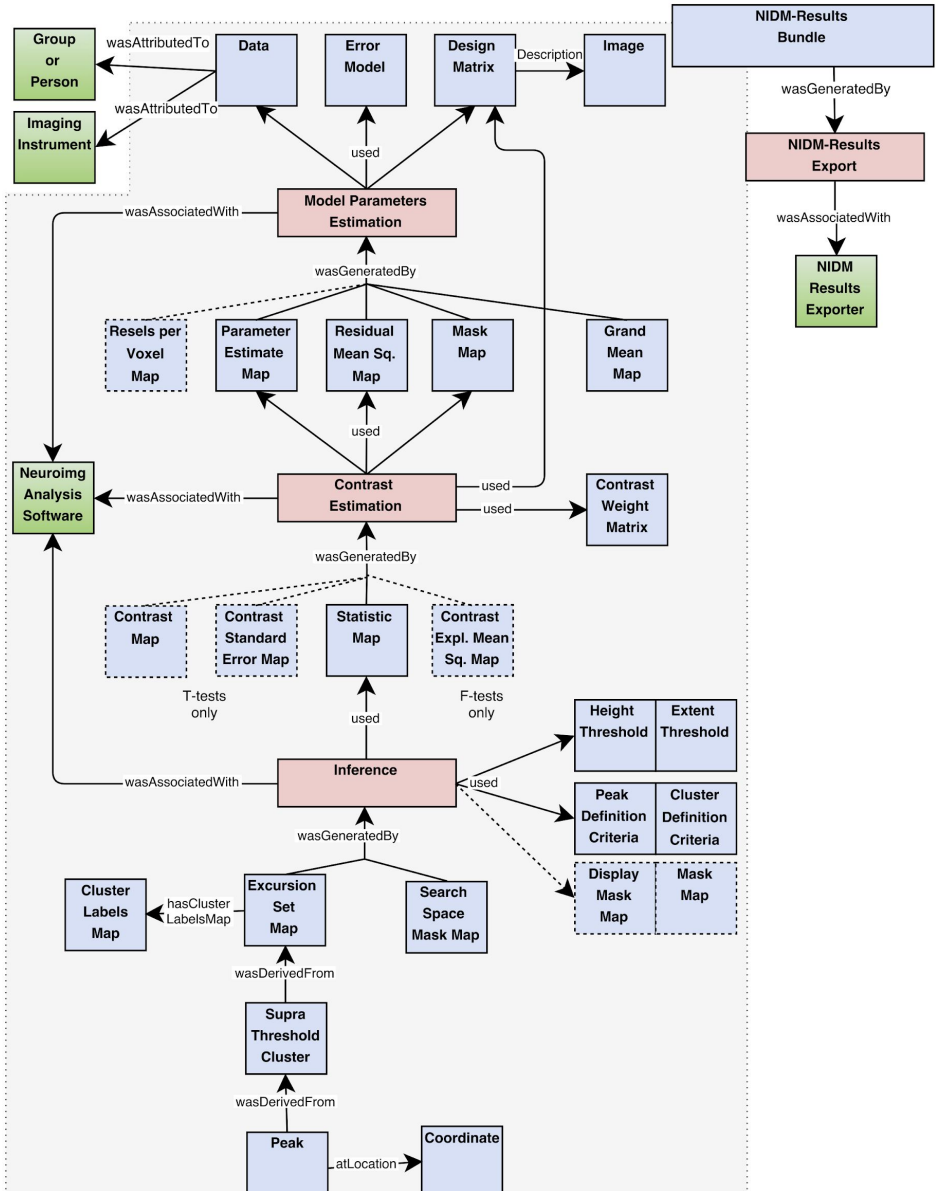
International collaboration with INCF SIG on Reproducibility and Best Practices in Human Brain Imaging.

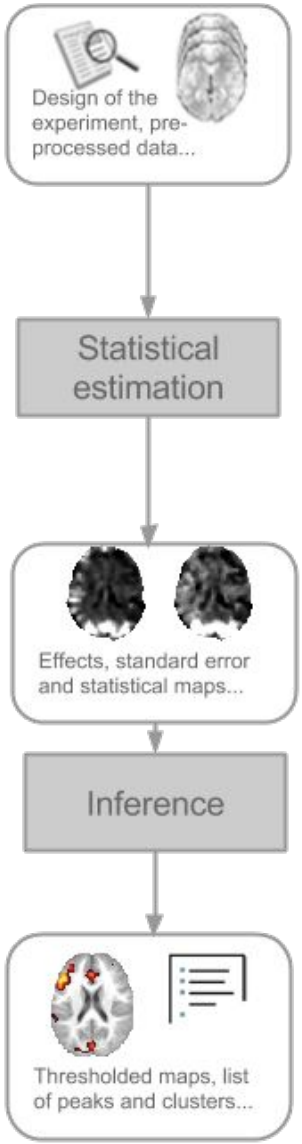
NIDM-Results → fMRI, VBM results

Harmonised model and provenance.

Joint work with INCF NIDASH working group - https://github.com/incf-nidash/

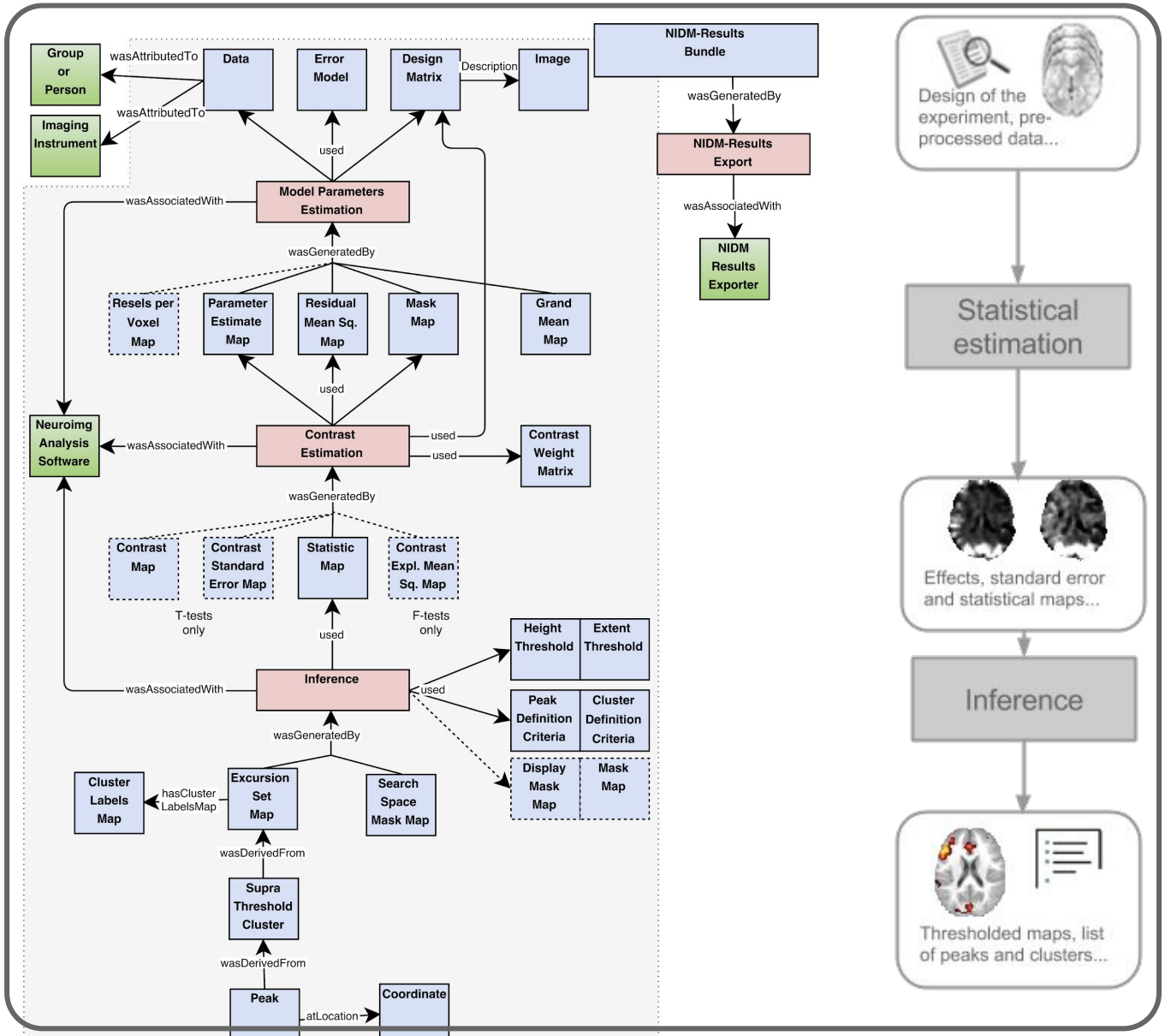
NIDM-Results in brief

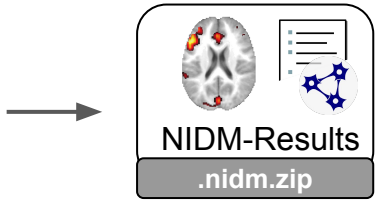




Maumet et. al, (2016)

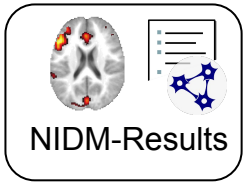
NIDM-Results in brief

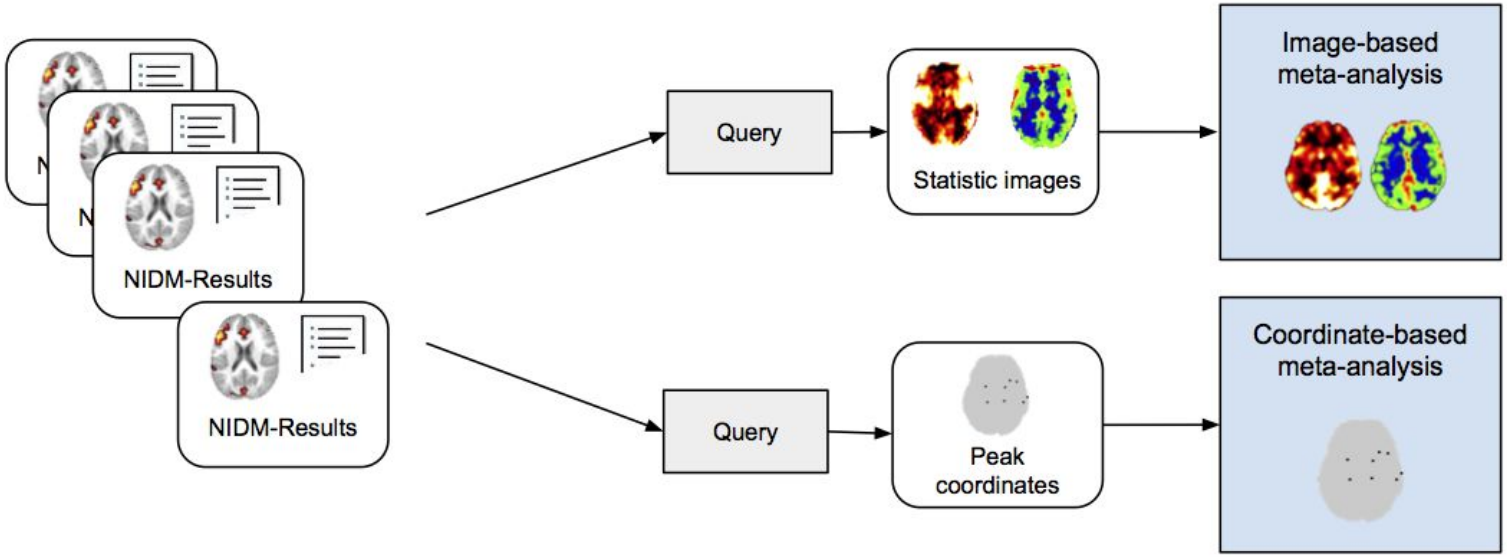




Maumet et. al, (2016)

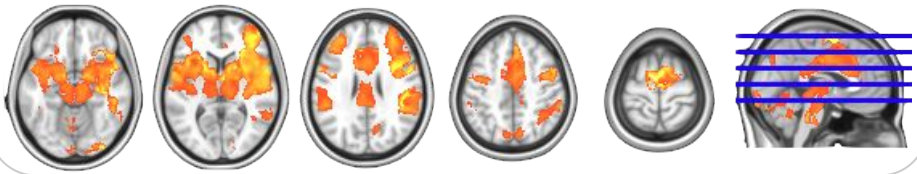
Derived data reuse with NIDM-Results: meta-analysis





Coordinate-based meta-analysis

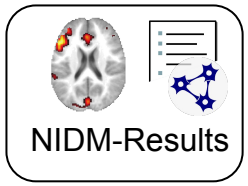
Image-based meta-analysis

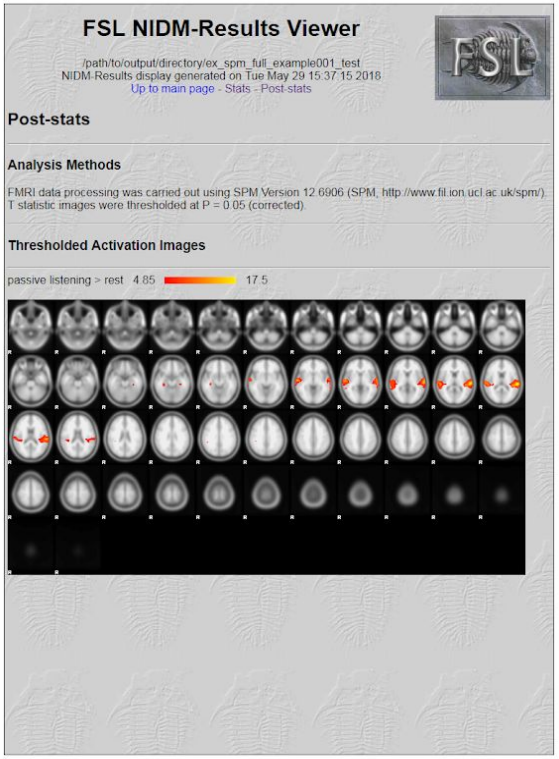


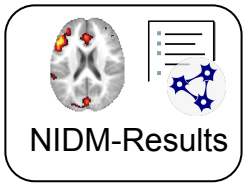
Scripts available at http://github.com/incf-nidash/nidmresults-paper

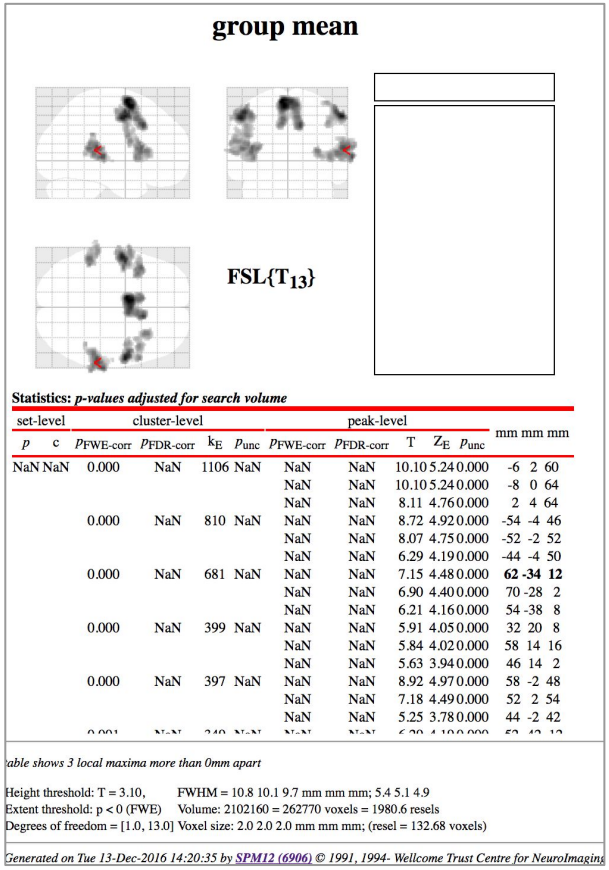
Maumet et. al, (2016)

Derived data inspection with NIDM-Results: SPM- and FSL-like views





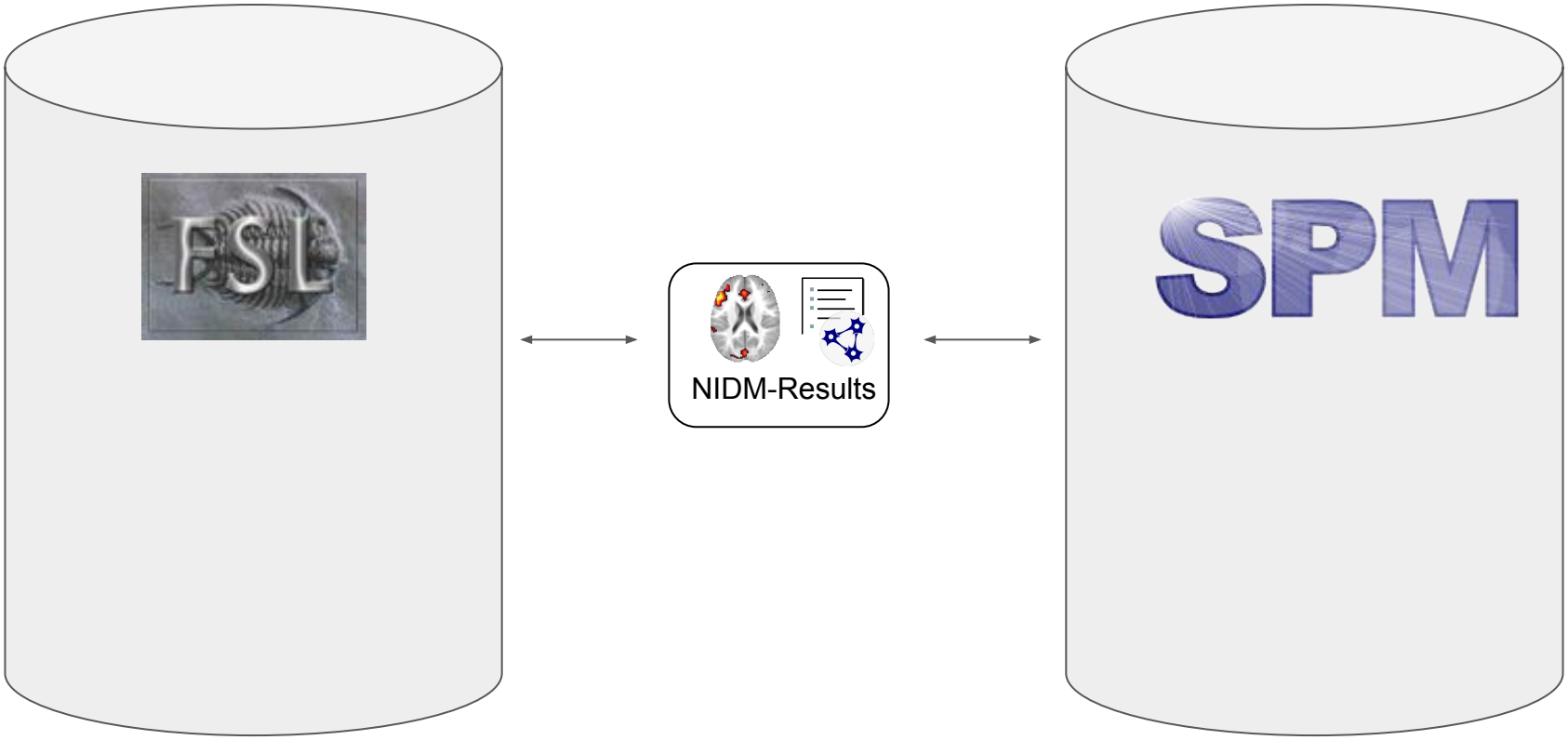




https://github.com/incf-nidash/nidmresults-spmhtml
https://github.com/incf-nidash/nidmresults-fslhtm
Maullin-Sapey et al (2017)

Screenshots by Tom Maullin-Sapey

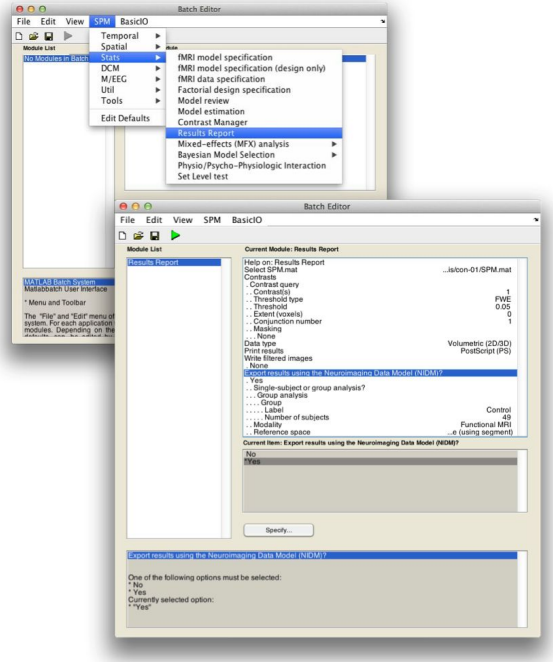
Derived data inspection with NIDM-Results: SPM- and FSL-like views



How to use NIDM-Results for my study?

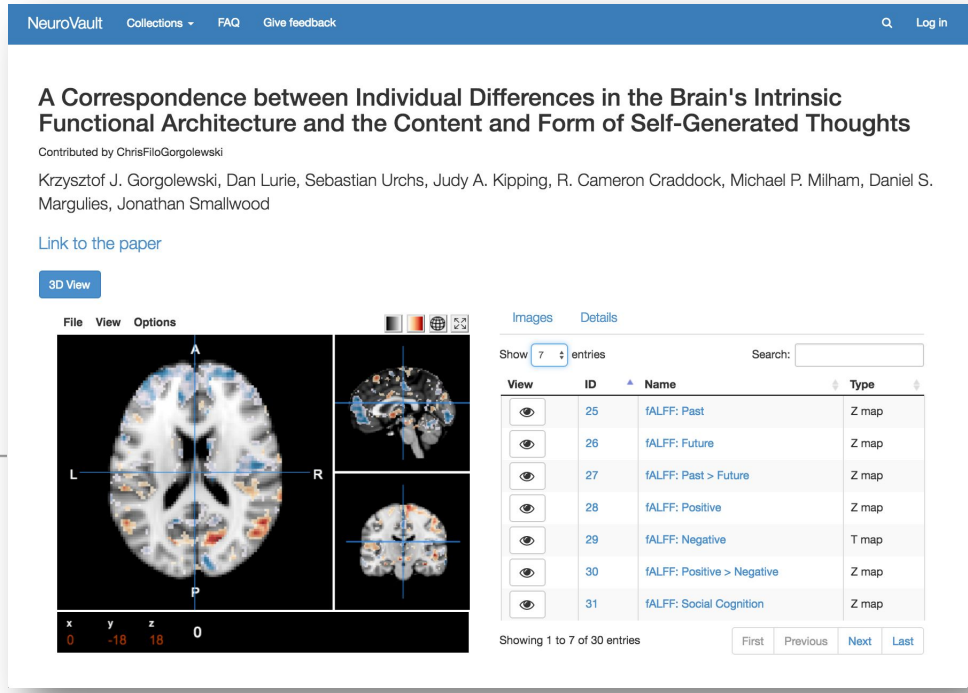
NIDM-Results

SPM



1. Export

2. Publication on NeuroVault



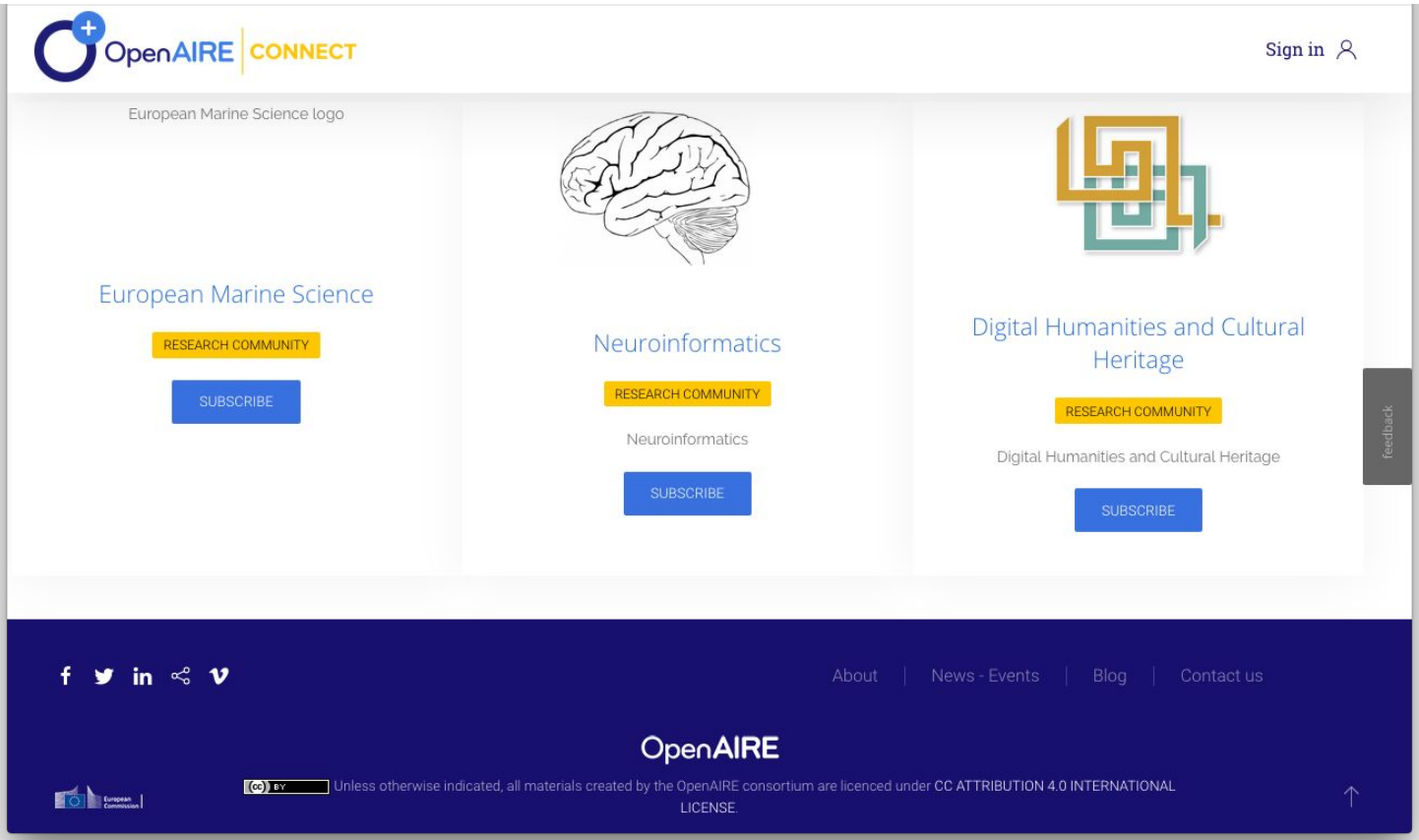
FSL

```
$ nidmfsl fsl_ds107_group 49 -g Control
```

https://github.com/incf-nidash/nidmresults-spm
https://github.com/incf-nidash/nidmresults-fsl

One step further: linking it all together!

"Gather, view and monitor all research outcomes"

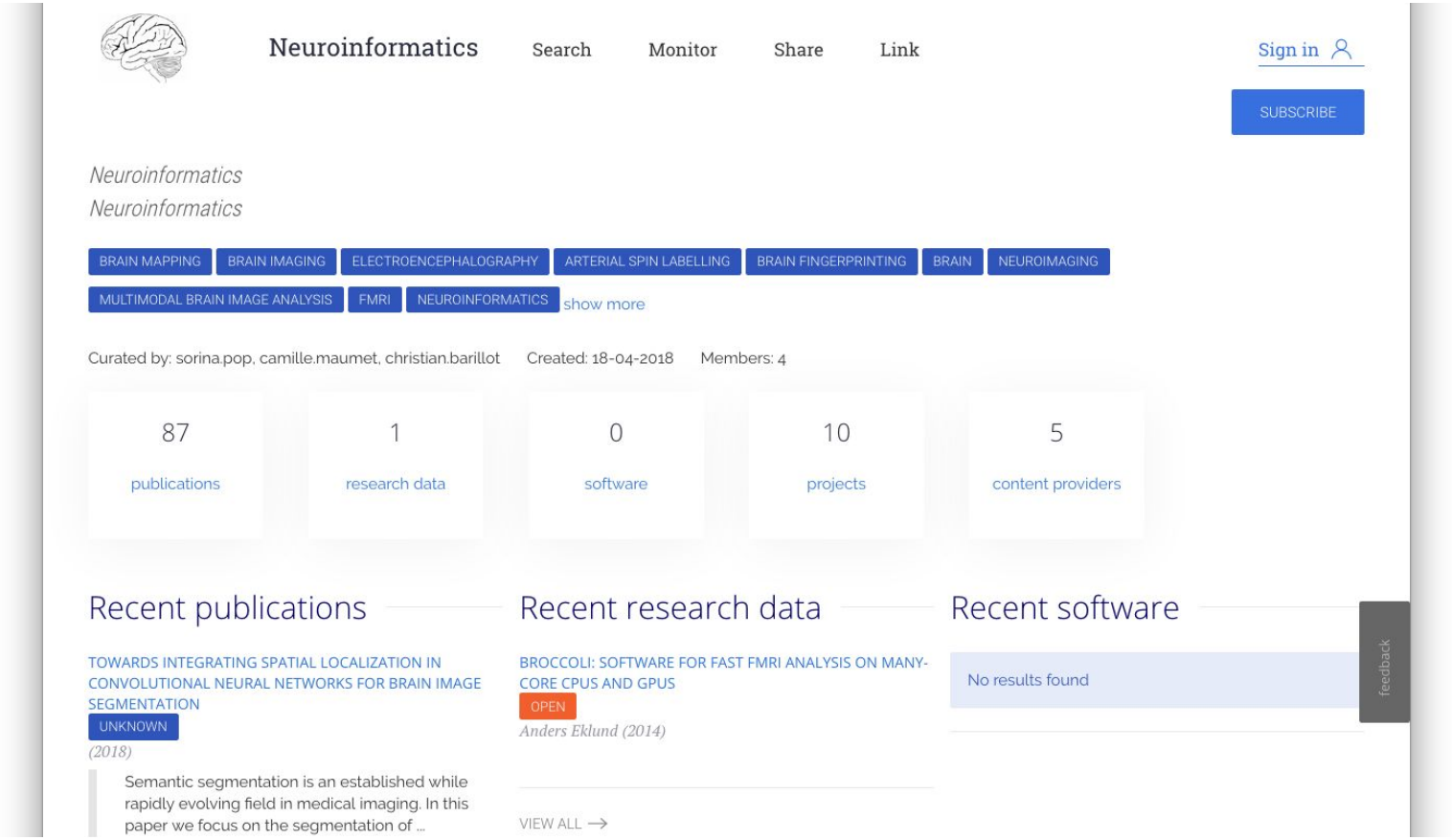


https://beta.connect.openaire.eu/

Principe (2018)

One step further: linking it all together!

"Gather, view and monitor all research outcomes"



https://beta.connect.openaire.eu/

Principe (2018)

Many ways to contribute

You can...

- Make your research outputs available and citable
- Cite existing software
- Use derived data standards and tools
- Contribute to the development of new specifications and their ecosystems
- Make datasets in neuroimaging data repository discoverable

Join the NIDASH community! https://github.com/incf-nidash/nidm-specs

OpenAIRE-connect portal: https://beta.ni.connect.openaire.eu/ and guidelines for data repository managers: https://www.openaire.eu/openaire-guide-for-repository-managers

References

R.A. Poldrack et al, Nat. Rev. Neurosci. (2017). doi:10.1038/nrn.2016.167.

J. Carp, Front. Neurosci. 6 (2012) 1–13.

E.H.B.M. Gronenschil et al, PLoS One. 7 (2012) e38234.

T. Glatard et al, Front. Neuroinform. 9 (2015) 12.

R. Pauli et al, Neuroinform. 10 (2016) 24.

A. Bowring et al, bioRxiv. (2018) 285585. doi:10.1101/285585.

A.M. Smith et al, PeerJ Comput. Sci. 2 (2016) e86.

A.E. Bandrowski and M.E. Martone, Neuron. 90 (2016) 434–436.

T. Maullin-Sapey et al, OHBM (2017).

C. Maumet et al, Scientific Data. 3 (2016). doi:10.1038/sdata.2016.102.

K.J. Gorgolewski et al, Front. Neuroinform. 9 (2015). doi:10.3389/fninf.2015.00008.

P. Príncipe et al, International Open Repositories Conference, 2017.

Photo and template credits

Brains on slide 4: Neil Conway - link
Transparency on slide 29: Public domain (CC0) - link
Presentation template by SlidesCarnival, adapted.

Acknowledgements

NIDM working group

Tibor Auer, Alexander Bowring, Samir Das, Fariba Fana, Guillaume Flandin, Satra Ghosh, Tristan Glatard, Chris Gorgolewski, Karl Helmer, Dorota Jarecka, David Keator, Tom Maullin-Sapey, Nolan Nichols, Tom Nichols, Smruti Padhy, Jean-Baptiste Poline, Vanessa Sochat, Jason Steffener, Jessica Turner.

Neuroimaging software teams

Gang Chen, Richard Reynolds, Robert Cox (AFNI), Mark Jenkinson, Matthew Webster, Paul McCarthy, Eugene Duff, Steve Smith (FSL), Guillaume Flandin (SPM).

Meta-analysis datasets Tracey group at FMRIB.

Thank you! To all INCF NIDASH task force members.

Thank you!

@cmaumet
camille.maumet@inria.fr

Join the NIDASH community:
https://github.com/incf-nidash/nidm-specs