EU Cost Action GliMR - European multi-site data integration & large dataset creation for glioma diagnostics

Camille Maumet

▶ **To cite this version:**

Camille Maumet. EU Cost Action GliMR - European multi-site data integration & large dataset creation for glioma diagnostics. EORTC 2021 - EORTC imaging group meeting, Mar 2021, Online, France. inserm-03193804

HAL Id: inserm-03193804
https://inria.hal.science/inserm-03193804

Submitted on 9 Apr 2021

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.

EORTC imaging group meeting March 29, 2021

EU Cost Action GliMR — European **multi-site data integration** & large dataset creation for glioma diagnostics

Camille Maumet

Univ Rennes, Inria, CNRS, Inserm

@cmaumet

Image credits: Sculpture by Malin Bjornsdotter "Cerebia", OHBM Brain Art SIG

Glioma MR Imaging 2.0

- Build a pan-European and multidisciplinary network of **international experts in glioma research, patient organisations, data scientists, and MR imaging scientists**
- Use **advanced MR imaging for improved decision making** in diagnosis, patient monitoring, and assessment of treatment response in clinical trials and clinical practice.


Esther Warnert


Radim Jancalek


Lydiane Hirschler


Kyrre Emblem


Marion Smits


Fran Borovecki

WG1 - Advanced MRI biomarkers for glioma characterisation

Working Group 1 focusses on coordinating the identification and quantification of advanced MRI biomarkers for the application in the field of glioma.

WG2 - Multi-site data integration

Working Group 2 coordinates multi-site data integration and enables the creation of large datasets in glioma diagnostics via the creation of common GDPR- and BIDS-compliant forms and data structures.

WG3 - Clinical translation

Working Group 3 focusses on fostering cross-border information exchange about past and current clinical trials and studies of glioma.

WG4 – Stakeholder relations

Working Group 4 ensures representation of all relevant stakeholders within GliMR and acts as liaison from GliMR to all relevant stakeholders.

WG5 – Dissemination

Working Group 5 disseminates the Action's goals and results in tailored manners to all relevant stakeholders.


Jan Petr

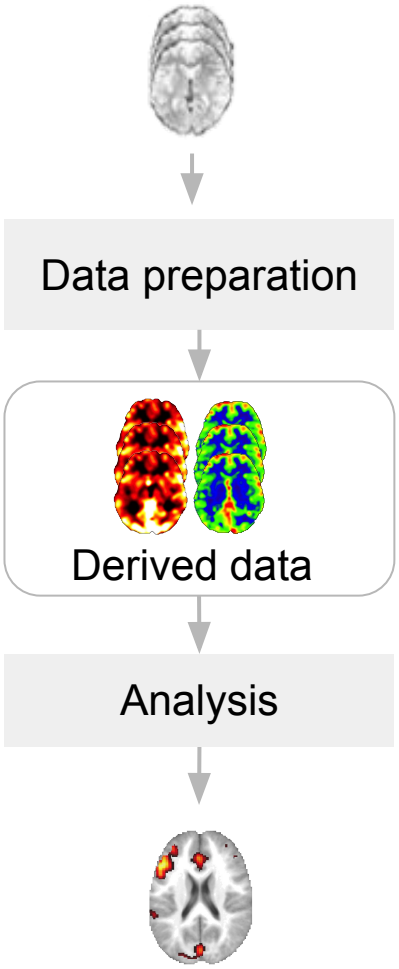

Vera Keil


Patricia Clement


Joana Pinto

https://glimr.eu/

Gliomas research in Europe



- **50,000 new cases** of gliomas per year in Europe
- **Low incidence** rate

>> Limited samples sizes

Image credits: Sculpture by Malin Bjornsdotter "Cerebia", OHBM Brain Art SIG

SCIENCE

A Waste of 1,000 Research Papers

Decades of early research on the genetics of depression were built on nonexistent foundations. How did that happen?

ED YONG MAY 17, 2019

SEAN NEL / SHUTTERSTOCK

In 1996, a group of European researchers found that a certain gene, called *SLC6A4*, might influence a person's risk of depression.

It was a blockbuster discovery at the time. The team found that a less active version of the gene was more common among 454 people who had mood disorders than in 570 who did not. In theory, anyone who had this particular gene variant could be at higher risk for depression, and that finding, they said, might help in diagnosing such disorders, assessing suicidal behavior, or even

Irreproducibility

The Atlantic Science, "A waste of 1000 research papers", Ed Yong.

Why are middle-aged marathon runners faster than twentysomethings?

According to new data from the running app Strava, runners in their 40s are streets ahead of younger rivals

▲ 'Middle-aged runners outperform runners in the 20s ...' Photograph: FatCamera/Getty

According to data released by the running app Strava, middle-aged runners consistently average faster marathon times than their younger rivals, apparently defying the usual rules of athletic performance. Men in the 40-49 age bracket clock an average time of four hours and 17 minutes for a marathon, according to the recent figures. Women in the same age range typically come in at just under the five-hour mark.

Faster in their forties than twenties?

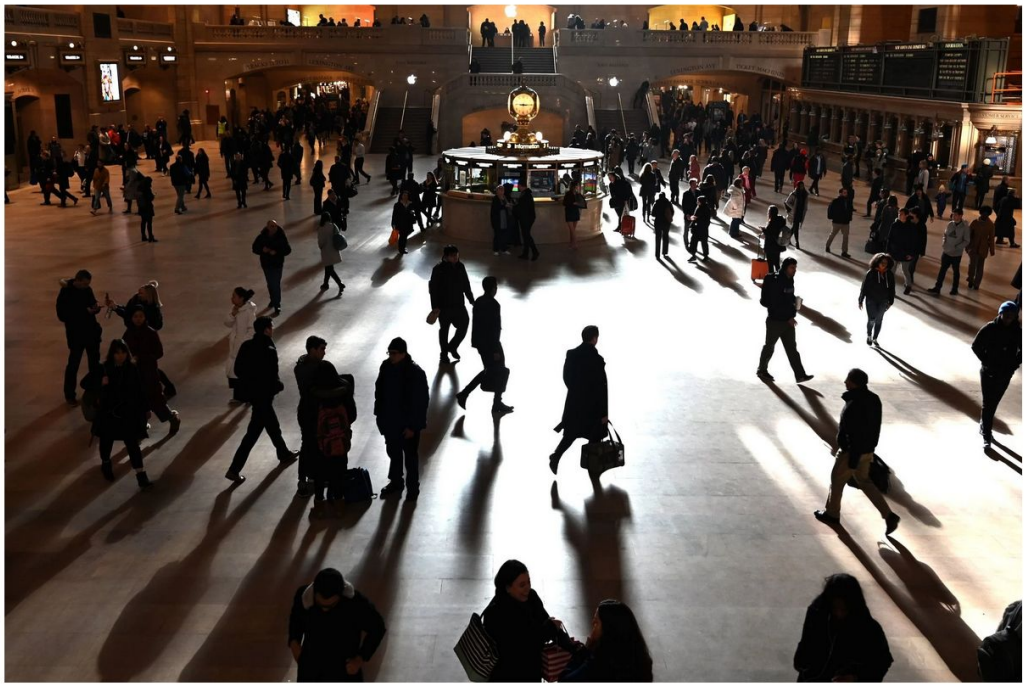
Selection bias

The Guardian, "Why are middle-aged marathon runners faster than twenty somethings?". Oliver Balch

The New York Times

Many Facial-Recognition Systems Are Biased, Says U.S. Study

Algorithms falsely identified African-American and Asian faces 10 to 100 times more than Caucasian faces, researchers for the National Institute of Standards and Technology found.

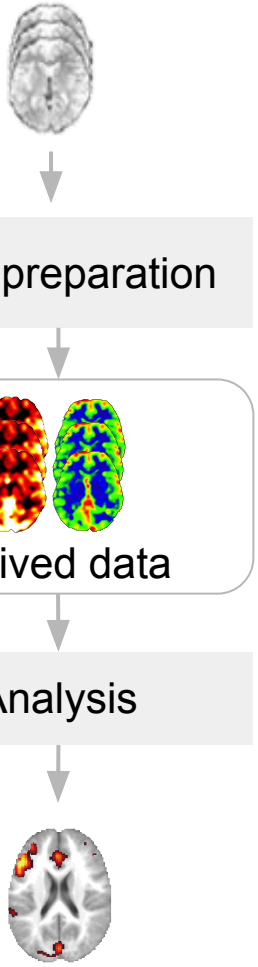


Morning at Grand Central Terminal. Technology for facial recognition is frequently biased, a new study confirmed. Timothy A. Clary/Agence France-Presse — Getty Images

Algorithmic injustice

The New York Times, "Many Facial-Recognition Systems Are Biased, Says U.S. Study", By Natasha Singer and Cade Metz, Dec. 2019

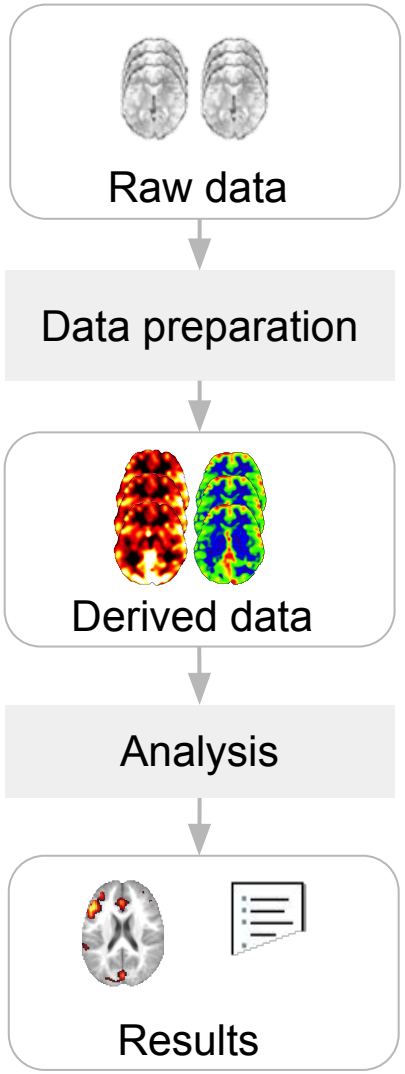
A brain imaging study



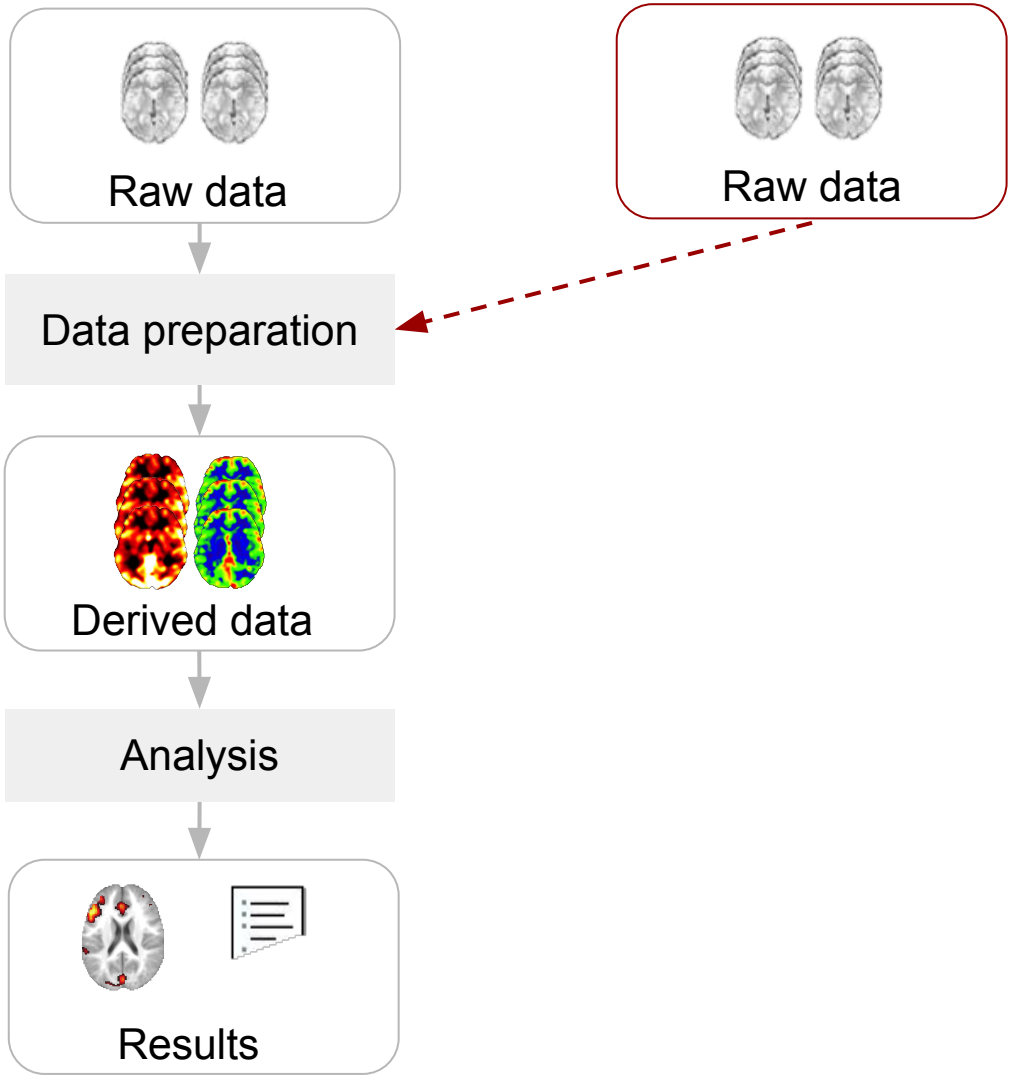
As a **community**, we need **bigger** & more **representative** samples

Data integration

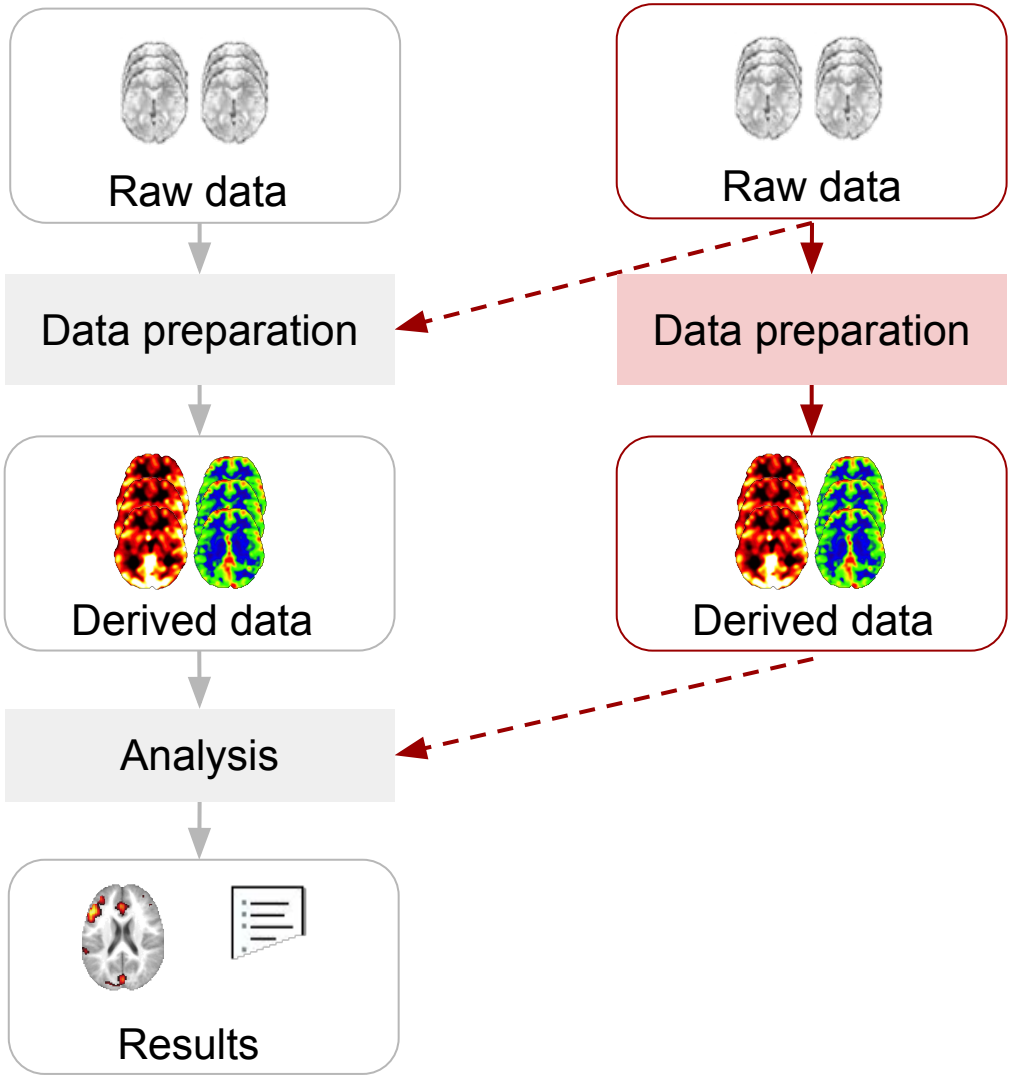
Data integration



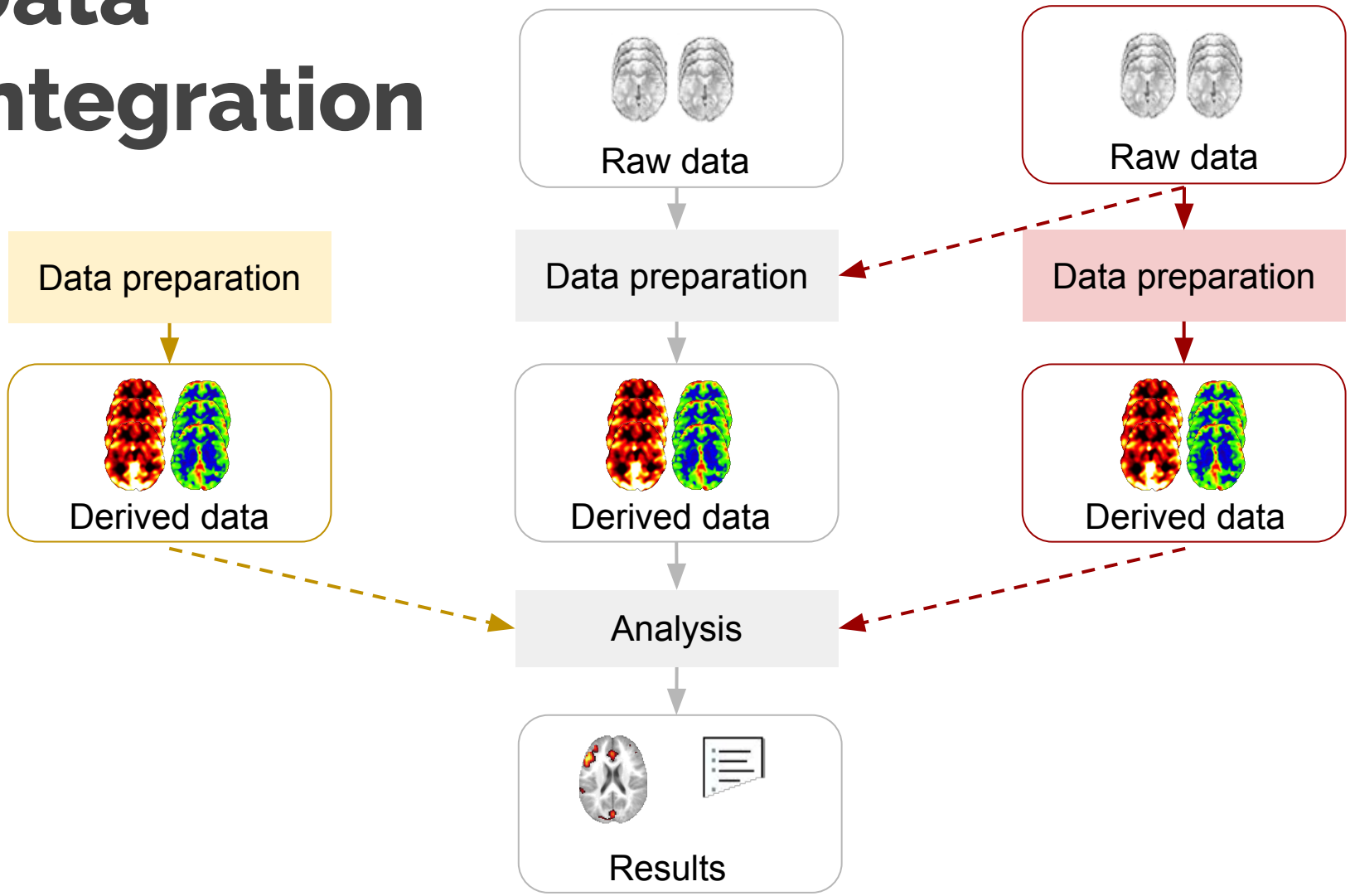
Data integration



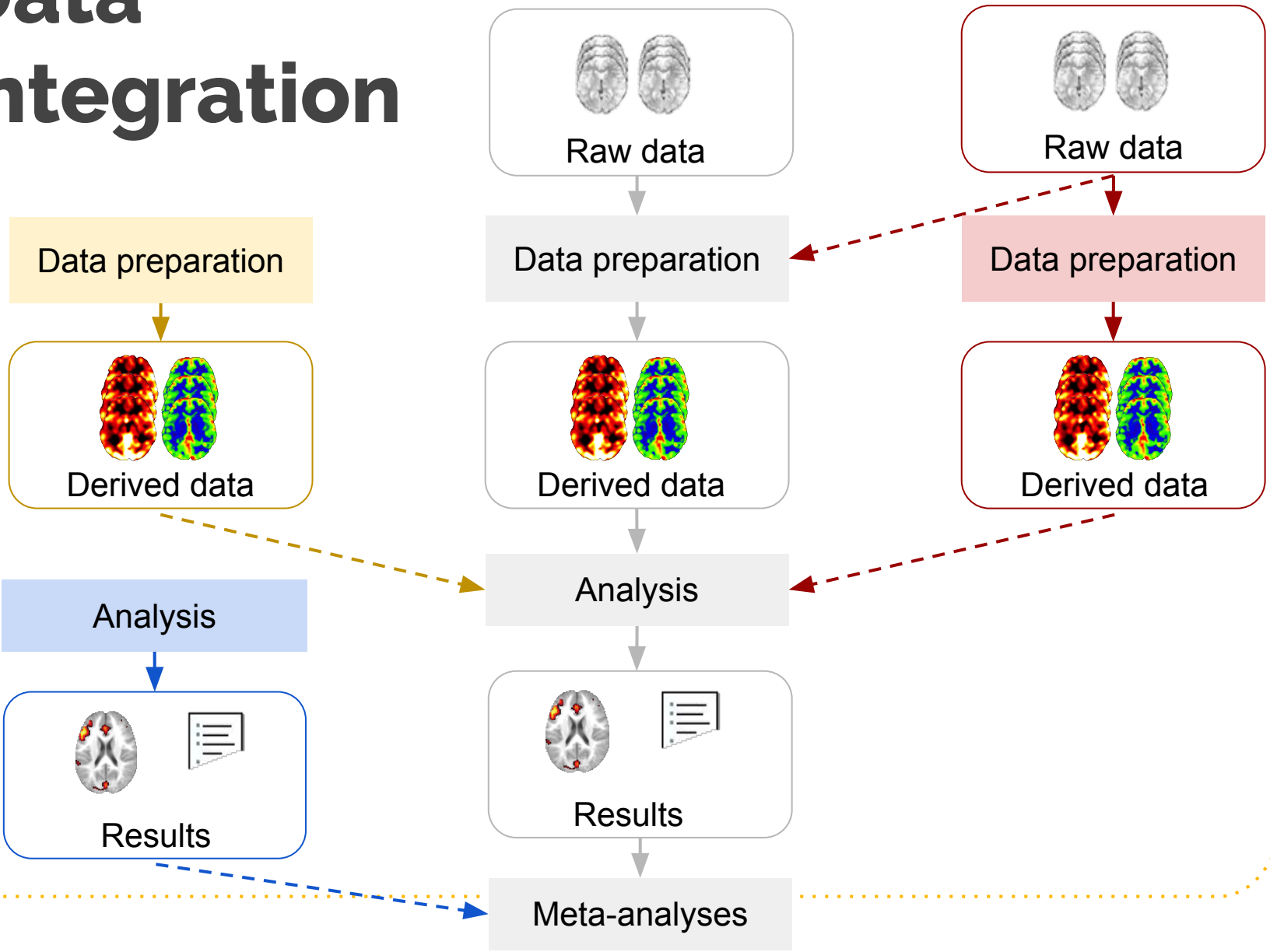
Data integration



Data integration



Data integration



GliMR WG2 **Multi-site data integration**

WG2 - Multi-site data integration

Working Group 2 coordinates multi-site data integration and enables the creation of large datasets in glioma diagnostics via the creation of common GDPR- and BIDS-compliant forms and data structures.

Camille Maumet

Cyril Pernet

Data Privacy

Data infrastructure

Data portability

Image credits: "Privacy lock", "Transport map", "Mini bus"

Neuroimaging data sharing

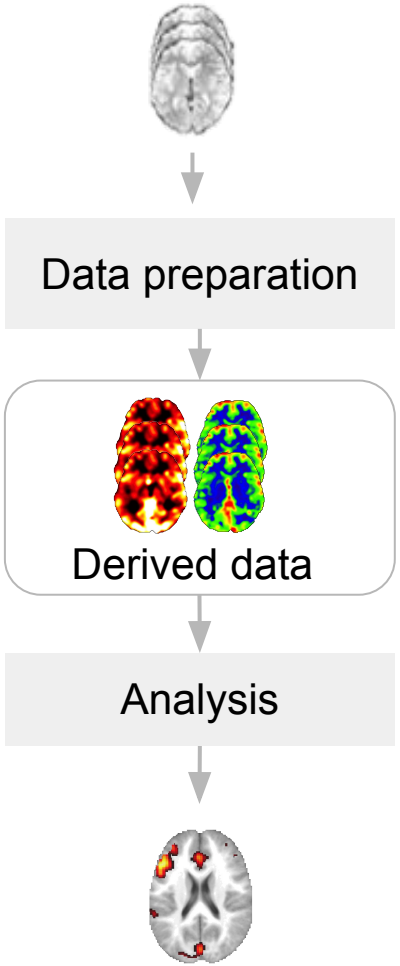
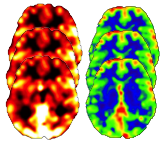


Image credits: Parcels 1 2 & 4 (CC0), Parcel 3 (CC0), Parcel 5 (CC0).

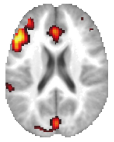
Neuroimaging data sharing

Data preparation



Derived data

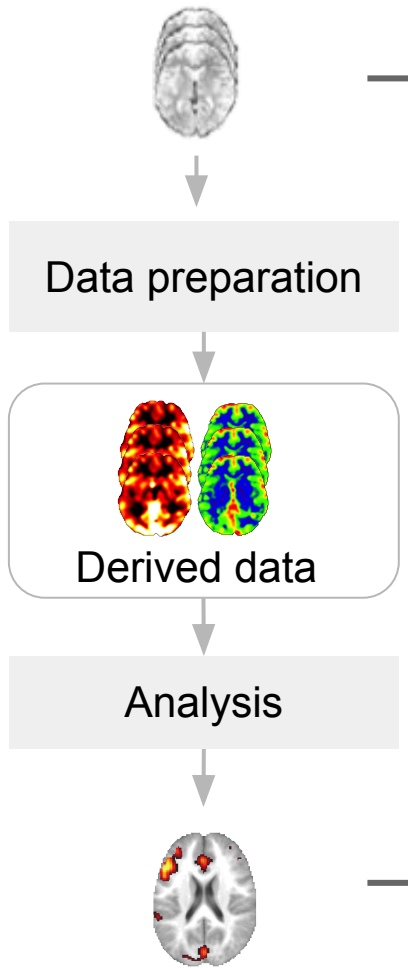
Analysis



HAL, BiorXiv
Pubmed

Image credits: Parcels 1 2 & 4 (CC0), Parcel 3 (CC0), Parcel 5 (CC0).

Neuroimaging data sharing



HAL, BiorXiv
Pubmed

Image credits: Parcels 1 2 & 4 (CC0), Parcel 3 (CC0), Parcel 5 (CC0).

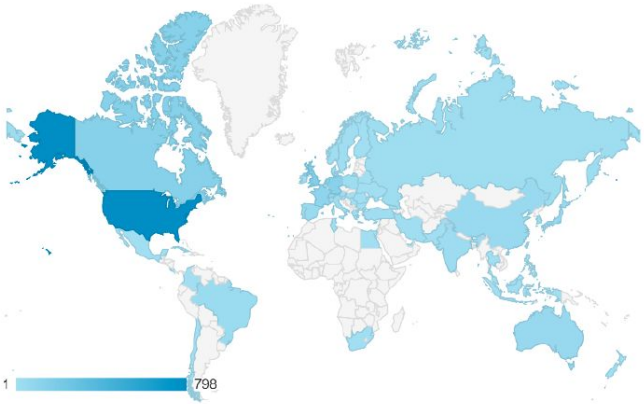
Brain Imaging Data Structure (BIDS)

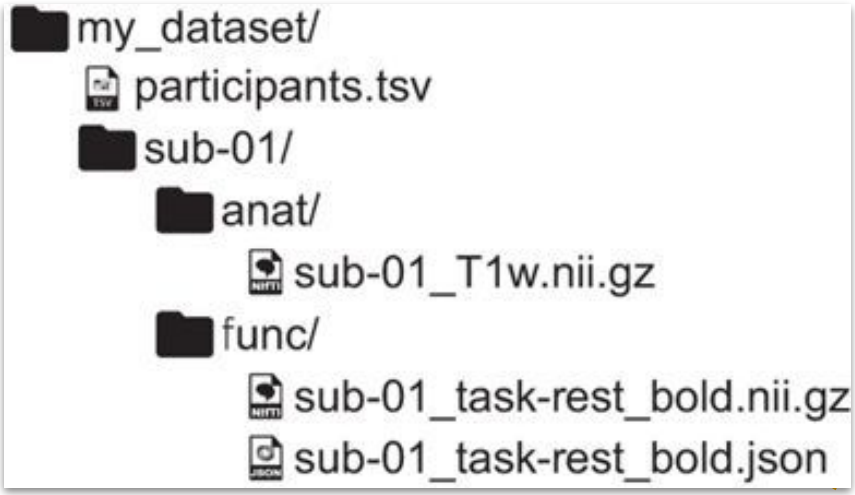
- Used in over 60 labs around the world
- Adopted by: FCP-INDI, Developing Human Connectome, SchizConnect and Donders Data repository.
- Extensions: MEG, iEEG, EEG

Krys Gorgolewski

(Gorgolewski et al., Sci. Data 2016)

Traffic to bids.neuroimaging.io





Slide by R. Poldrack & K. Gorgolewski (CC BY), adapted.

Data portability

How differences in acquisition parameters can be conveyed with enough details to support future post-processing?

- **Extend the international BIDS standard** in setting up the structures for sharing advanced MR imaging data.

Brain Imaging Data Structure (BIDS):

A global community of 200+ members!

BIDS
BRAIN IMAGING DATA STRUCTURE


Guiomar Niso


Melanie Ganz


Robert Oostenveld


Russ Poldrack

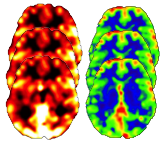

Kirstie Whitaker

Image credits: "Mini bus"

https://bids.neuroimaging.io/

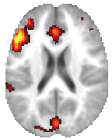
Neuroimaging data sharing

Data preparation



Derived data

Analysis



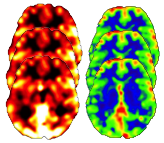
HAL, BiorXiv
Pubmed

Image credits: Parcels 1 2 & 4 (CC0), Parcel 3 (CC0), Parcel 5 (CC0).

Neuroimaging data sharing

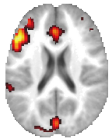
OpenfMRI
XNAT
SHANOIR
LORIS
OpenNEURO

Data preparation



Derived data

Analysis



HAL, BiorXiv
Pubmed

Image credits: Parcels 1 2 & 4 (CC0), Parcel 3 (CC0), Parcel 5 (CC0).

Data privacy

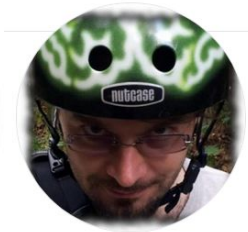
How to inform participants, collect consent for data sharing, and ensure de-identification of data and anonymity of metadata in accordance with GDPR?

- Review existing standards on data privacy

Open Brain Consent

(Bannier et al., HBM 2020)

Stephan Heunis



Yarik Halchenko

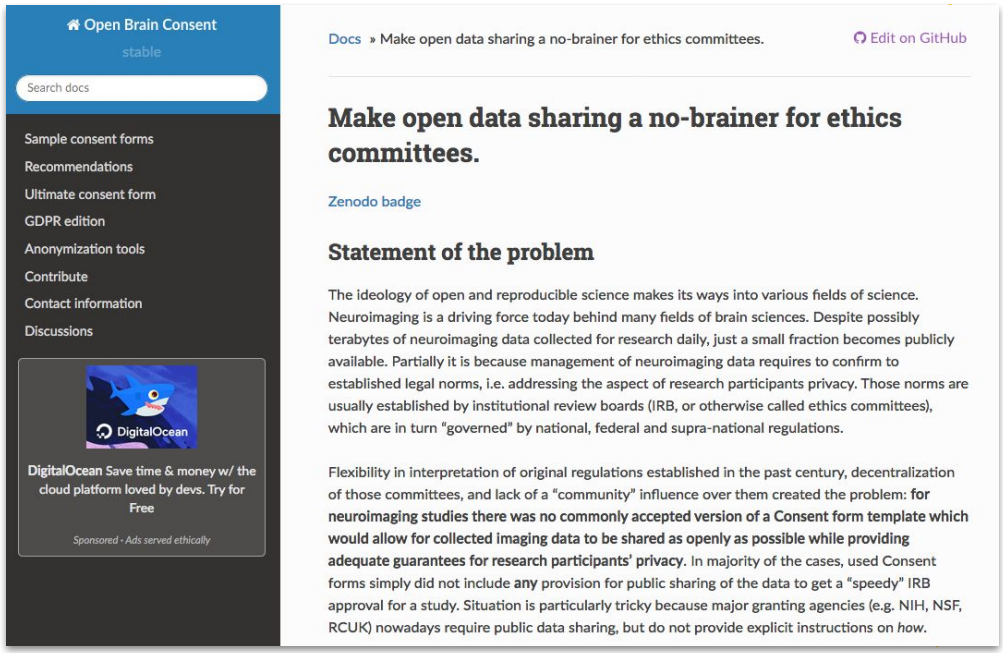


Image credits: "Privacy lock"

https://open-brain-consent.readthedocs.io/

Data privacy with GDPR

Consent form

Direct link

To be approved and signed by participants

"[...] While the collection, use and storage of your data are done for the purpose of conducting the study to which you are currently participating, **these data might also be used for other future research projects in the field of medical and cognitive neuroscience**. [...]"

Data user agreement

Direct link

To be approved and signed by researchers reusing the data

"[...] 2. I **will not attempt to establish or retrieve the identity of the study participants**. I will not link these data to any other database in a way that could provide identifying information. [...]

3. **I will not redistribute these data** or share access to these data with others, unless they have independently applied and been granted access to these data, i.e., signed this Data Use Agreement. This includes individuals in my institution.
[...]"

Translations

- Bosnian
- Czech
- German
- Greek
- Spanish
- Finnish
- French
- Italian
- Dutch
- Norwegian
- Turkish

French version

Anne Hespel

Elise Bannier

Image credits: "Privacy lock"

Data privacy with GDPR

Consent form
Direct link

To be approved and signed by participants

"[...] While the collection, use and storage of your data are done for the purpose of conducting the study to which you are currently participating, **these data might also be used for other future research projects in the field of medical and cognitive**

Data user agreement
Direct link

To be approved and signed by researchers reusing the data

"[...] 2. I **will not attempt to establish or retrieve the identity of the study participants**. I will not link these data to any other database in a way that could provide identifying information. [...]

3. **I will not redistribute these data** or share access to these data with others unless they have

11:10 Roxana Albu:
Re-use of data in GDPR era

Translations

- Bosnian
- Czech
- German
- Greek
- Spanish
- Finnish
- French
- Italian
- Dutch
- Norwegian
- Turkish

French version

Anne Hespel

Elise Bannier

Image credits: "Privacy lock"

Data infrastructure

Which tools and databases are available and how to use them to automate multi-site integration?

- Review existing data sharing infrastructure (in Europe)

Collaboration with **ENBIT**

Daniele Marinazzo

Cyril Pernet

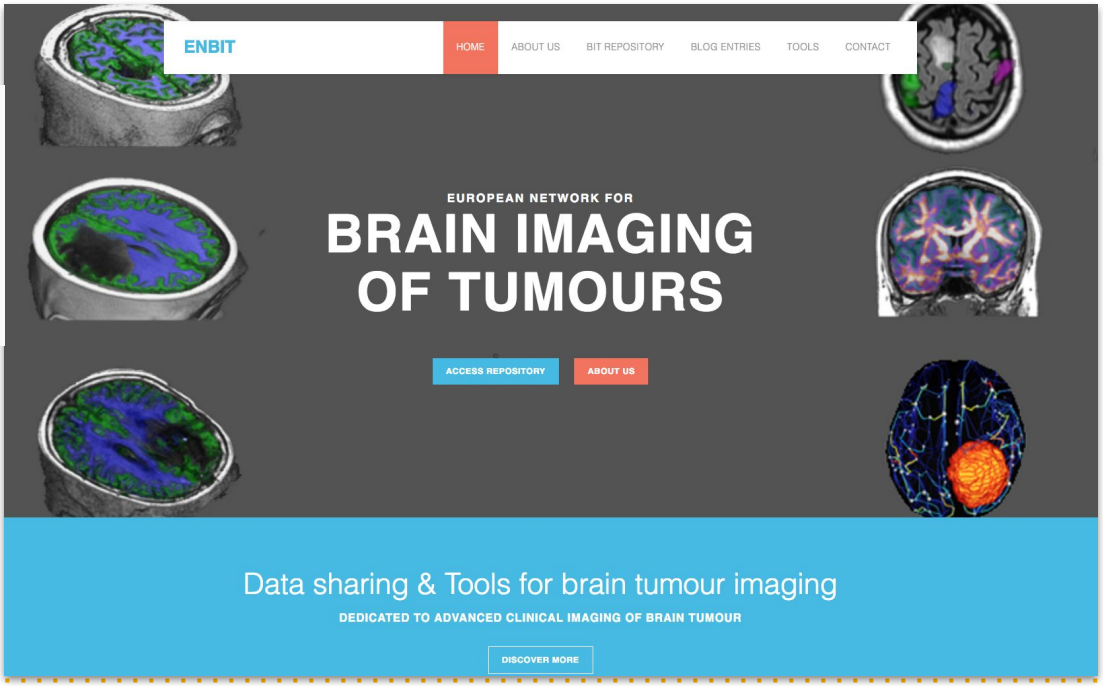


Image credits: "Transport map"

https://en-brain-imaging-of-tumours.github.io/website/

EU Cost Action GliMR — European multi-site data integration & large dataset creation for glioma diagnostics

Camille Maumet

Thank you!

Credit: Sculpture by Malin Bjornsdotter “Cerebia”, OHBM Brain Art SIG; Presentation template by SlidesCarnival, adapted

@cmaumet