



Collaboratively building large neuroimaging datasets for Glioma research

Camille Maumet

► To cite this version:

Camille Maumet. Collaboratively building large neuroimaging datasets for Glioma research. ESM-RMB 2020 - European Society for Magnetic Resonance in Medicine and Biology, Oct 2020, Barcelona - Online, Spain. pp.1-16. inserm-02965500

HAL Id: inserm-02965500

<https://inria.hal.science/inserm-02965500>

Submitted on 27 Feb 2024

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.

ESMRMB 2020 - September 30

Collaboratively building **large neuroimaging datasets** for Glioma research



ESMRMB 2020 ONLINE
SEPT. 30 – OCT. 2
CONVERGENCE SCIENCE & EDUCATION

Camille Maumet

Univ Rennes, Inria, CNRS, Inserm

 @cmaumet

 @COST18206

Gliomas research in Europe

- 50,000 primary gliomas / year in Europe
- Low incidence rate



>> Research studies of limited samples sizes

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

(Ir)-Reproducibility

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

Lack of representativity

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

Gliomas research in Europe

- 50,000 primary gliomas / year in Europe
 - Low incidence rate
- >> Research studies of limited samples sizes

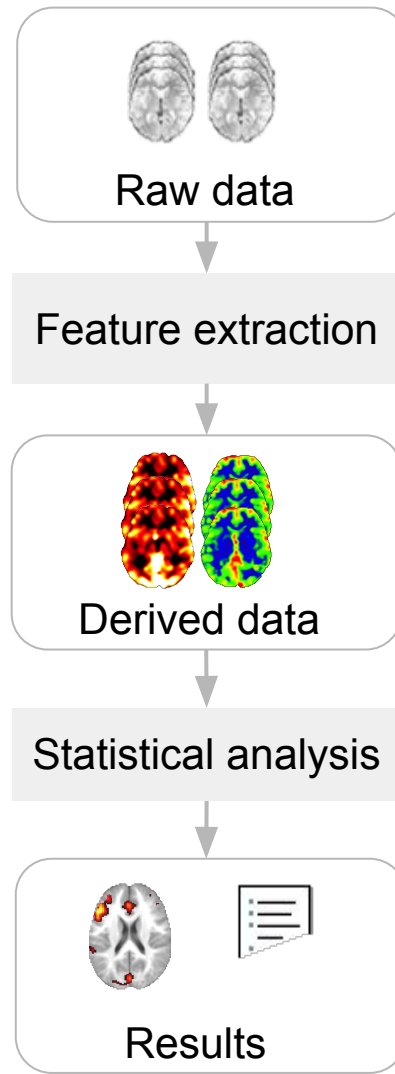


We need **bigger**, more **representative & diverse** samples

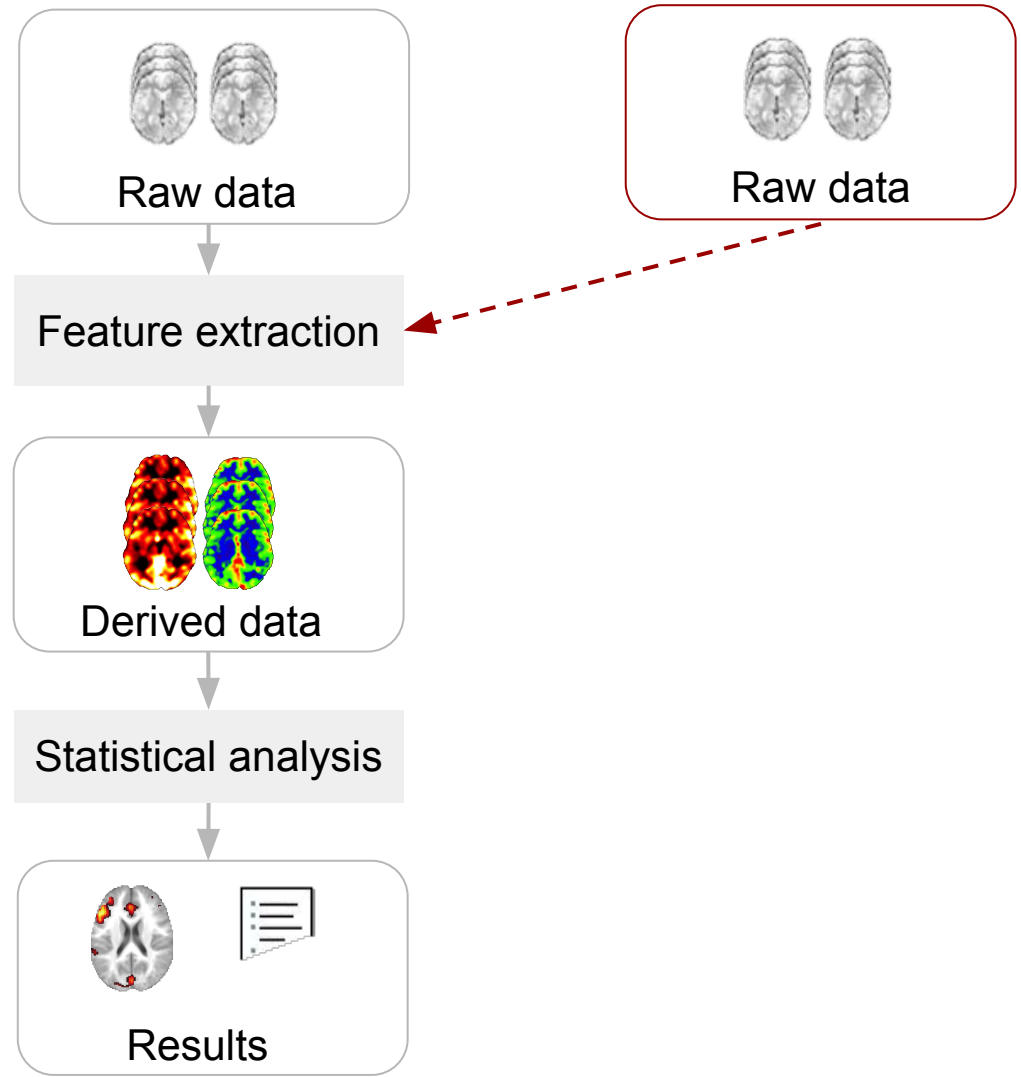


Data integration

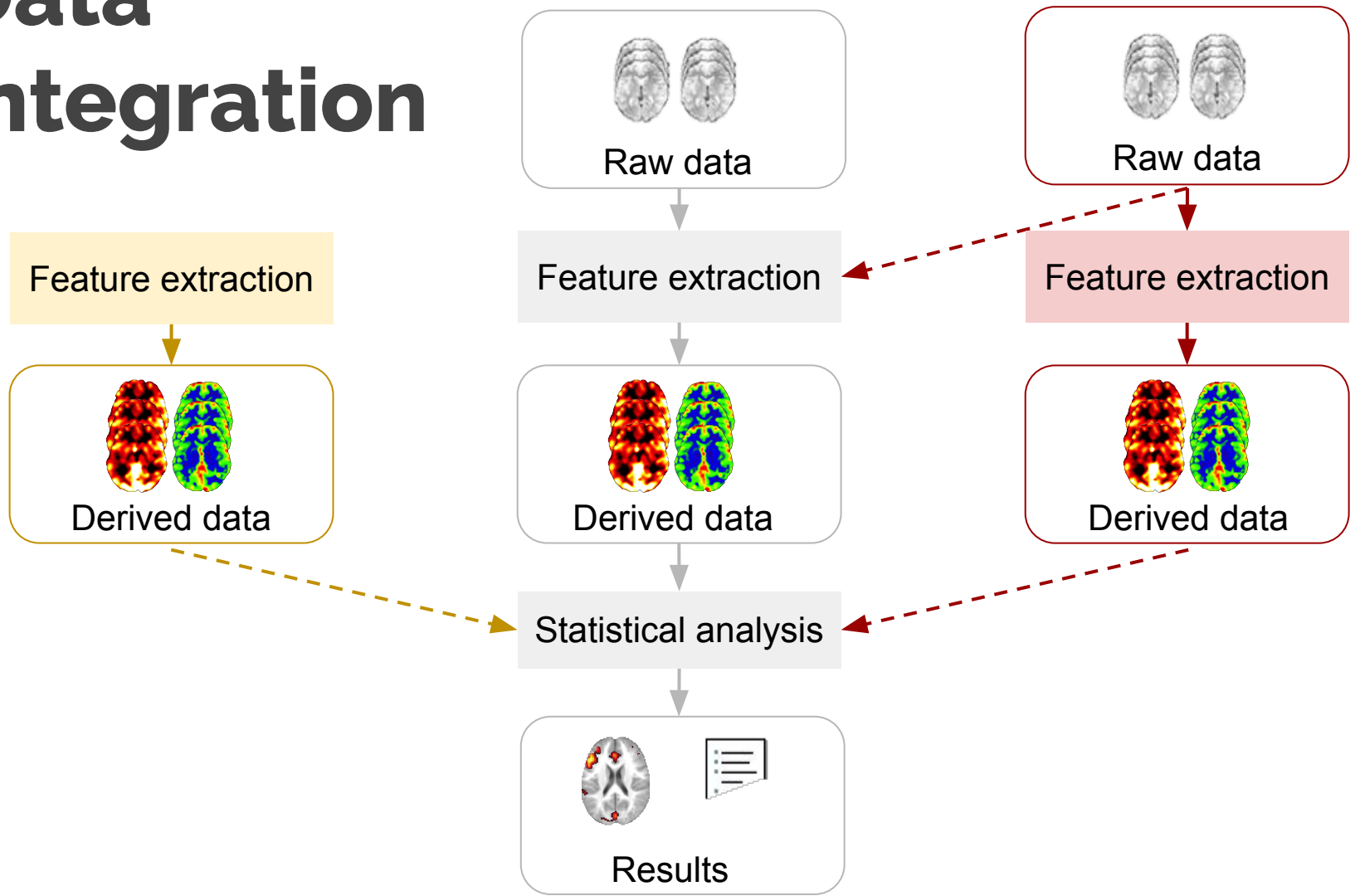
Data integration



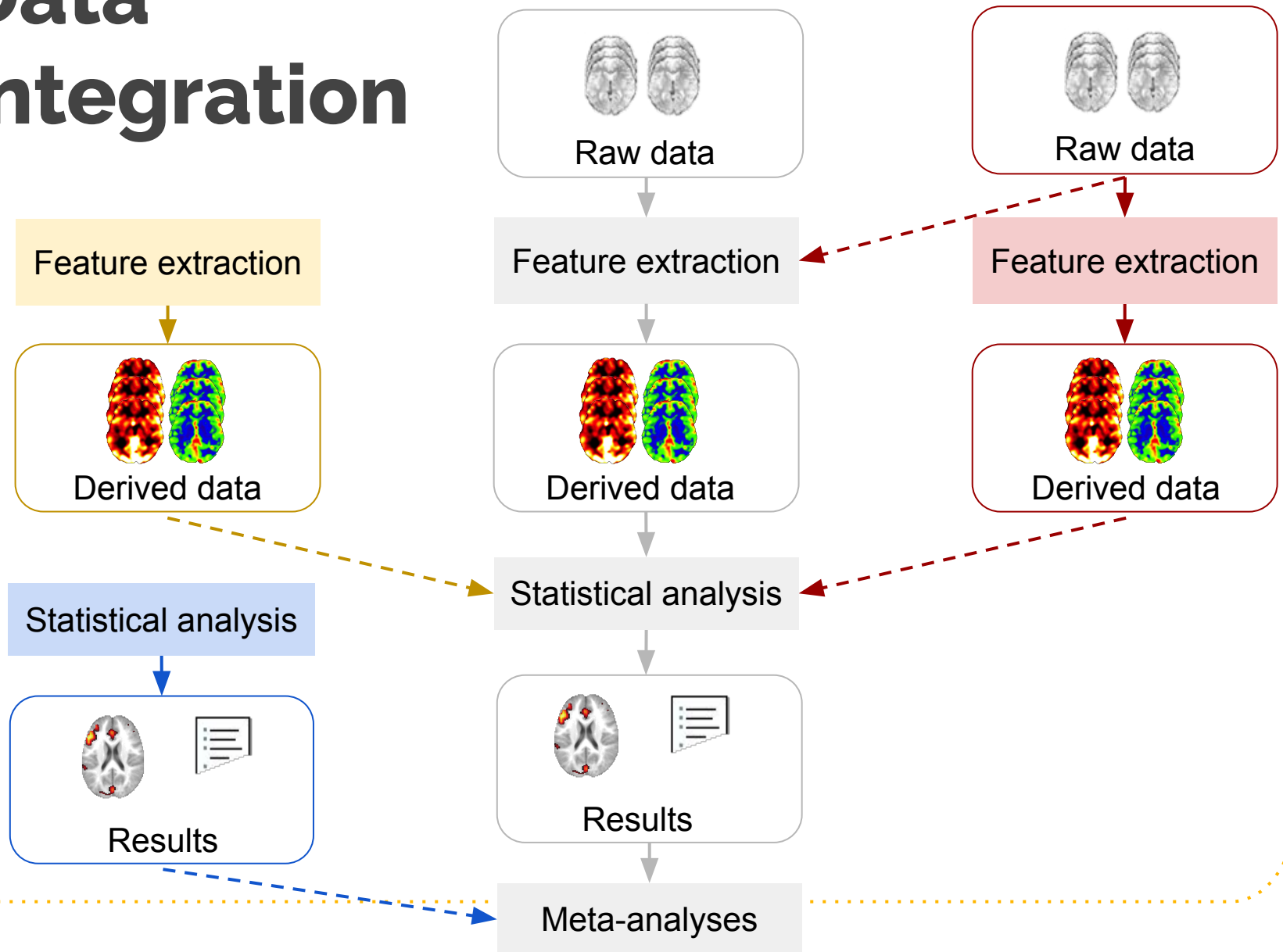
Data integration



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

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



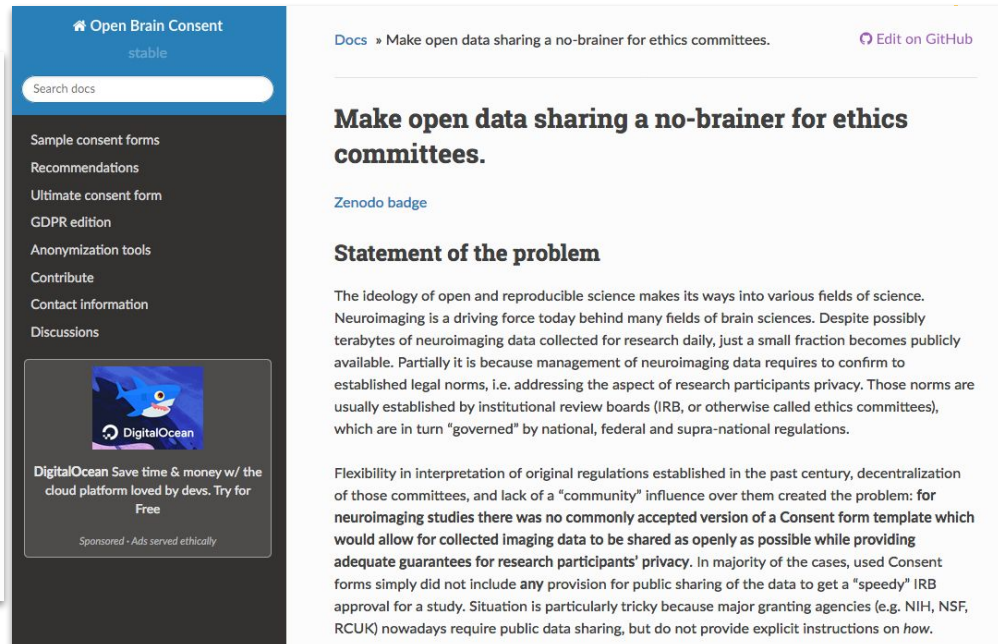
Stephan
Heunis



Yarik
Halchenko

Translations

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

A screenshot of the Open Brain Consent website. The header is blue with the text 'Open Brain Consent' and 'stable'. Below the header is a search bar. The main content area is dark grey and lists various resources: 'Sample consent forms', 'Recommendations', 'Ultimate consent form', 'GDPR edition', 'Anonymization tools', 'Contribute', 'Contact information', and 'Discussions'. There is a DigitalOcean advertisement for cloud storage. On the right side, there is a white box containing the text 'Make open data sharing a no-brainer for ethics committees.' and a 'Zenodo badge' link. Below this is a section titled 'Statement of the problem' with a paragraph of text. At the bottom right, there is a link to 'https://open-brain-consent.readthedocs.io/'.

Open Brain Consent
stable

Search docs

Sample consent forms
Recommendations
Ultimate consent form
GDPR edition
Anonymization tools
Contribute
Contact information
Discussions

DigitalOcean Save time & money w/ the cloud platform loved by devs. Try for Free
Sponsored - Ads served ethically

Docs » Make open data sharing a no-brainer for ethics committees. [Edit on GitHub](#)

Make open data sharing a no-brainer for ethics committees.

[Zenodo badge](#)

Statement of the problem

The ideology of open and reproducible science makes its ways into various fields of science. Neuroimaging is a driving force today behind many fields of brain sciences. Despite possibly terabytes of neuroimaging data collected for research daily, just a small fraction becomes publicly available. Partially it is because management of neuroimaging data requires to confirm to established legal norms, i.e. addressing the aspect of research participants privacy. Those norms are usually established by institutional review boards (IRB, or otherwise called ethics committees), which are in turn "governed" by national, federal and supra-national regulations.

Flexibility in interpretation of original regulations established in the past century, decentralization of those committees, and lack of a "community" influence over them created the problem: for neuroimaging studies there was no commonly accepted version of a Consent form template which would allow for collected imaging data to be shared as openly as possible while providing adequate guarantees for research participants' privacy. In majority of the cases, used Consent forms simply did not include any provision for public sharing of the data to get a "speedy" IRB approval for a study. Situation is particularly tricky because major granting agencies (e.g. NIH, NSF, RCUK) nowadays require public data sharing, but do not provide explicit instructions on how.

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

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)

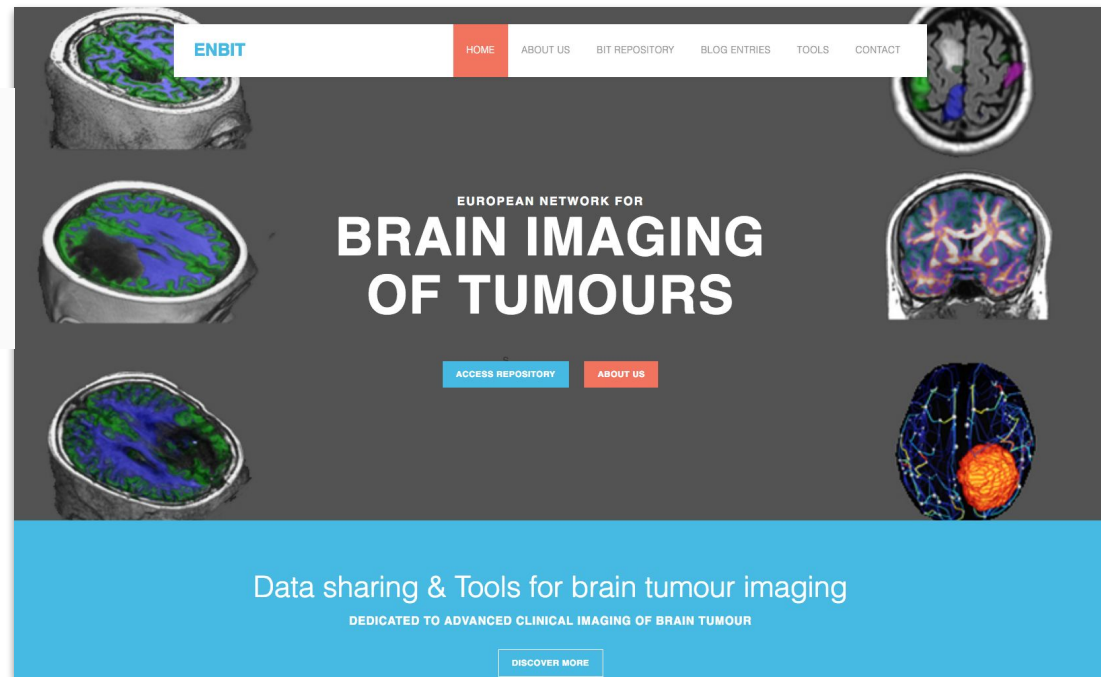
Collaboration with **ENBIT**



Daniele
Marinazzo



Cyril
Pernet



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!



Guiomar
Niso



Melanie
Ganz



Robert
Oostenveld



Russ
Poldrack



Kirstie
Whitaker

Collaboratively building **large neuroimaging datasets** for Glioma research

Join GliMR Working Group 2!

➤ glimr.cost.wg2@gmail.com

Thank you!



**ESMRMB 2020 ONLINE
SEPT. 30 – OCT. 2
CONVERGENCE SCIENCE & EDUCATION**

Credit: Presentation template by [SlidesCarnival](#), adapted

 **@cmaumet**