



PublicnEUro.eu



OHBM 2024
JUNE 23~27, SEOUL, KOREA

A European repository to share neuroimaging datasets publicly

Dr Cyril Pernet, PhD
Neurobiology Research Unit, Rigshospitalet, DK

novo
nordisk
fonden
Benefiting people and society

publicneuro@nru.dk
cyril.pernet@nru.dk

I have no disclosures.

Brain imaging data as personal data

- EU GDPR Art. 4 (1). Personal data are any information which are related to an **identified or identifiable** (directly or indirectly) natural person.



Re-identification from structural MRI

Conferences > 2019 IEEE 16th International ... ?

Refacing: Reconstructing Anonymized Facial Features Using GANS

Publisher: IEEE

Cite This

Conferences > 2022 International Joint Conf... ?

David Abramian ; Anders Eklund

Refacing Defaced MRI with PixelCNN

Publisher: IEEE

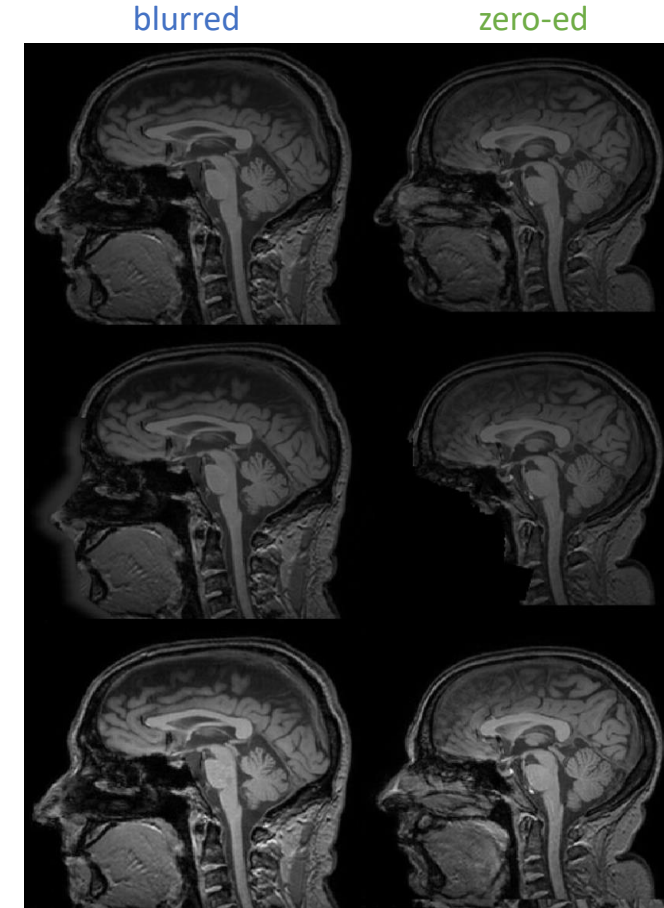
Cite This

PDF

Yaorong Xiao ; William Ashbee ; Vince D. Calhoun ; Sergey Plis

Loss of potential information!! defacing is likely not a solution

- Removing facial features degrades perf of image analysis method
- Prevents biophysical modelling for MEEG
- Defacing biases manual and automated quality assessments



Re-identification from functional MRI

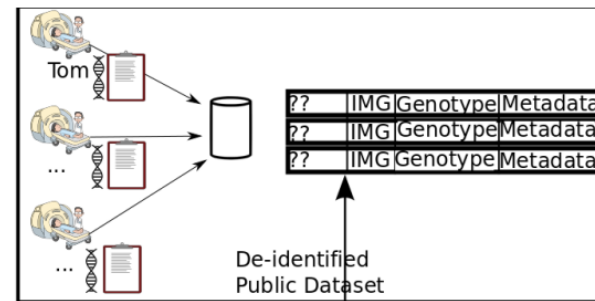
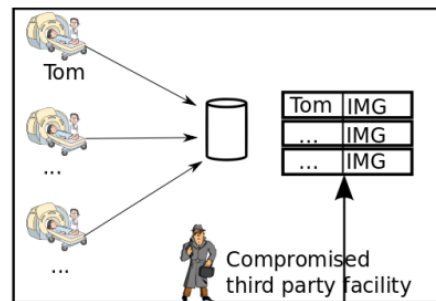
Research Applications Track Paper

SIGMOD '21, June 20–25, 2021, Virtual Event, China

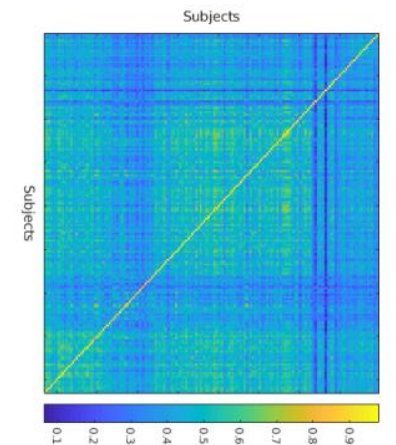
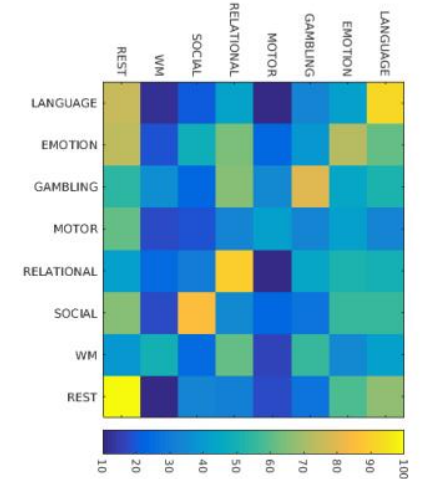
De-anonymization Attacks on Neuroimaging Datasets

Vikram Ravindra
ravindr@purdue.edu
Purdue University
West Lafayette, Indiana

Ananth Grama
ayg@cs.purdue.edu
Purdue University
West Lafayette, Indiana



Can we identify Tom and extract his genomes?



Re-identification from PET

NeuroImage 258 (2022) 119357



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Face recognition from research brain PET: An unexpected PET problem

Christopher G. Schwarz^{a,*}, Walter K. Kremers^b, Val J. Lowe^a, Marios Savvides^c,
Jeffrey L. Gunter^a, Matthew L. Senjem^{a,d}, Prashanthi Vemuri^a, Kejal Kantarci^a,
David S. Knopman^e, Ronald C. Petersen^e, Clifford R. Jack Jr.^{a,#}, the Alzheimer's Disease
Neuroimaging Initiative[#]

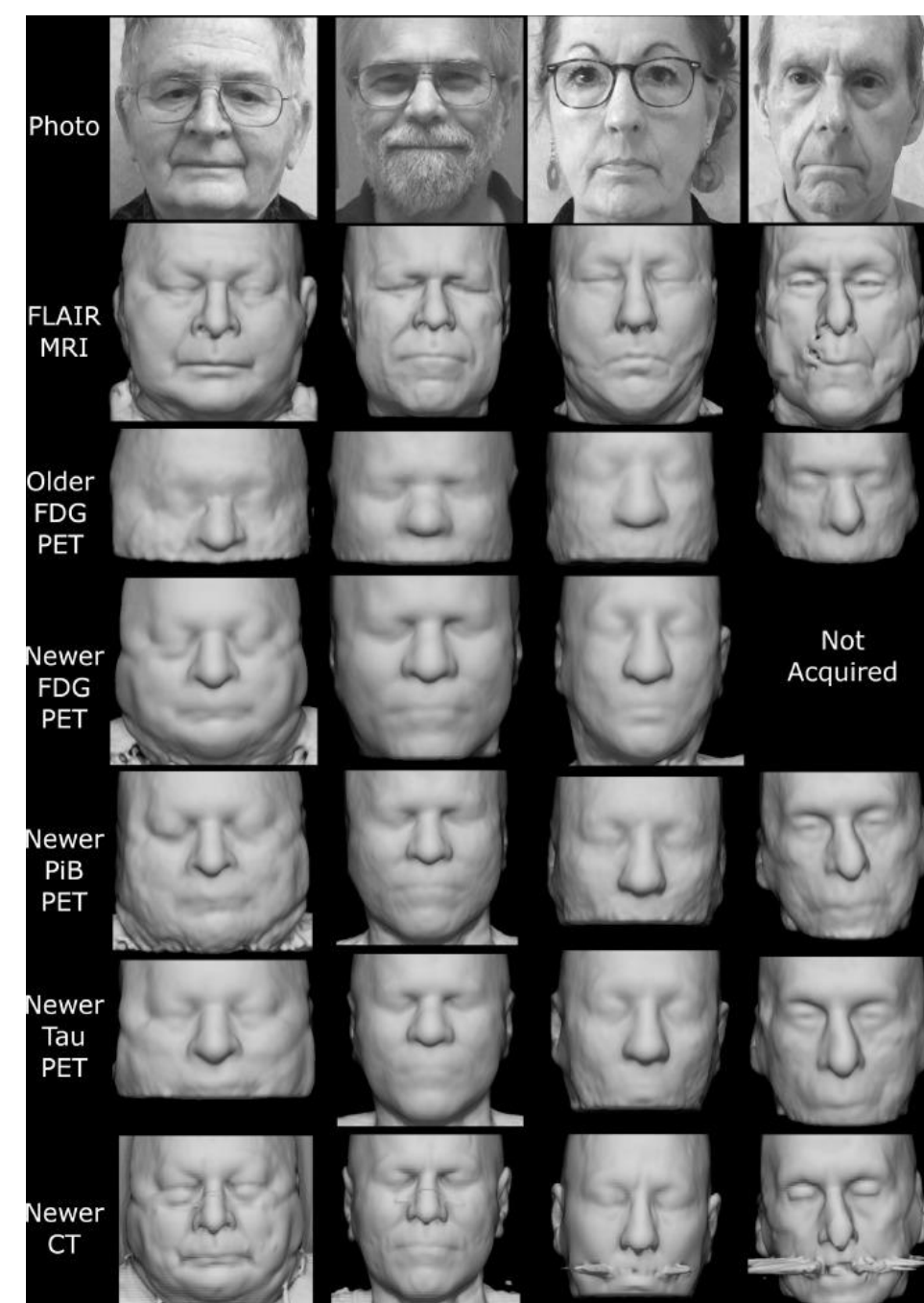
^a Department of Radiology, Mayo Clinic, Rochester, MN, USA

^b Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA

^c CyLab Biometrics Center and Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

^d Department of Information Technology, Mayo Clinic, Rochester, MN, USA

^e Department of Neurology, Mayo Clinic, Rochester, MN, USA



Re-identification from EEG and MEG



NeuroImage

Volume 58, Issue 2, 15 September 2011, Pages 620-629



Reliability of ERP and single-trial analyses ☆

Carl M. Gaspar^a , Guillaume A. Rousselet^a, Cyril R. Pernet^b



Pattern Recognition Letters

Volume 95, 1 August 2017, Pages 37-43

Permanence of the CEREBRE brain biometric protocol

Maria V. Ruiz-Blondet , Zhanpeng Jin, Sarah Laszlo

Article | [Open Access](#) | [Published: 29 September 2021](#)

Brief segments of neurophysiological activity enable individual differentiation

[Jason da Silva Castanheira](#), [Hector Domingo Orozco Perez](#), [Bratislav Misic](#) & [Sylvain Baillet](#)

[Nature Communications](#) **12**, Article number: 5713 (2021) | [Cite this article](#)

7459 Accesses



RESEARCH ARTICLE | [Open Access](#) |

Discovering heritable modes of MEG spectral power

Eemeli Leppäaho, Hanna Renvall, Elina Salmela, Juha Kere, Riitta Salmelin, Samuel Kaski

First published: 01 January 2019 | <https://doi.org/10.1002/hbm.24454> |

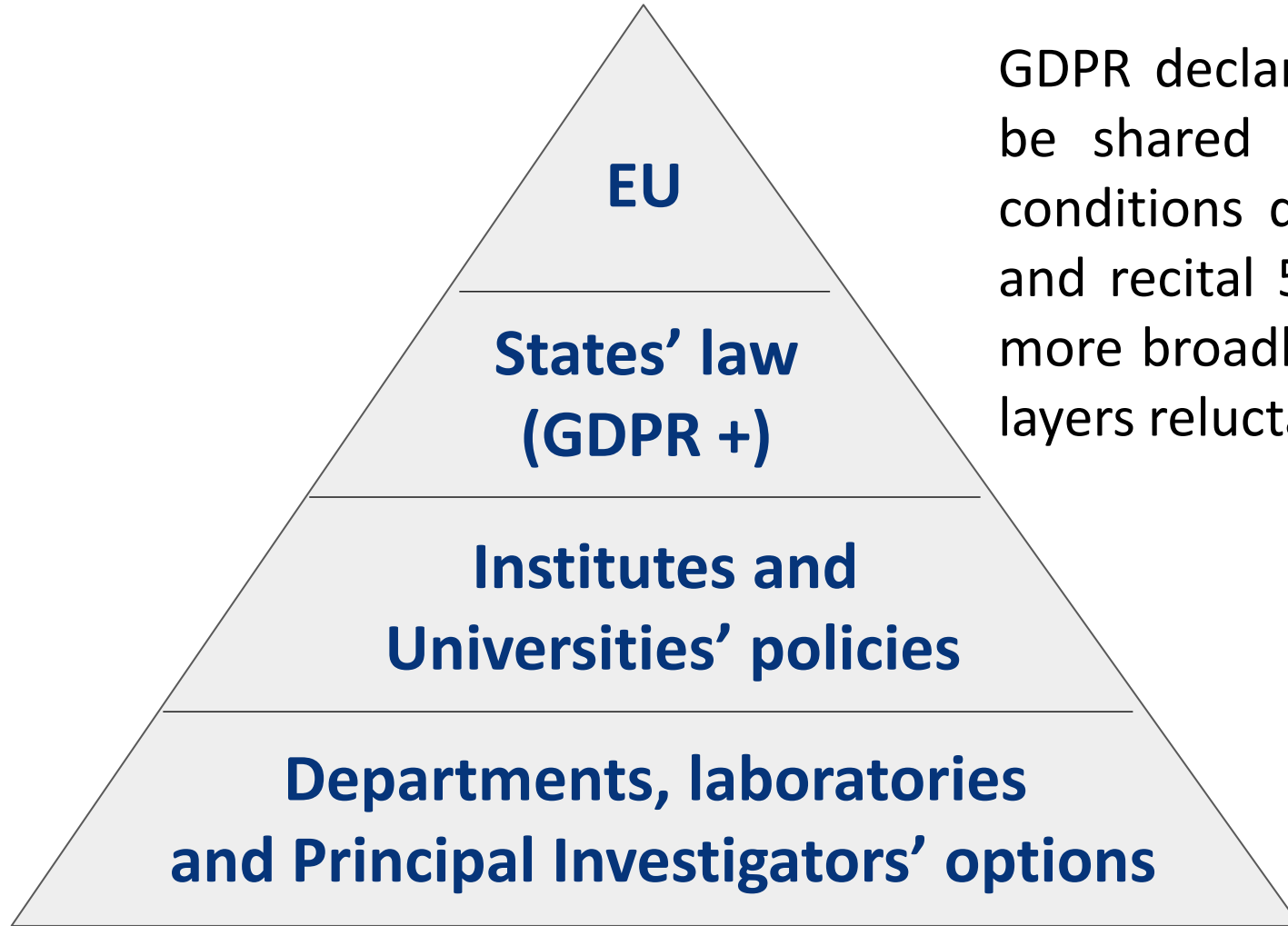
good discriminability, stability over time, and tightly correlated with molecular expression/genes

Brain imaging data as personal data

- EU GDPR Art. 4 (1). Personal data are any information which are related to an **identified or identifiable** natural person.
- There are only 2 types of data: personal or anonymous.
- For biomedical imaging data, features in the data (e.g. face, ears but also neural fingerprints), and metadata (age, sex, clinical profile) are identifiable. At least we know we can single out individuals, which can be seen as sufficient for being personal data.
- Note: Pseudonymization (the process of removing personal information) does not change the status of the data.



GDPR, states, institutions and labs



GDPR declares that within EU personal data can be shared (Art 1), it also states that general conditions do not apply to research data (Art 5 and recital 50), i.e. scientific data can be shared more broadly without purpose limitation (despite layers reluctance).

Each state has a specific implementation adding different restrictions. Each institution has specific policies and Data Protection Officers interpret the law and those policies differently.

Our Solution

Allows diversity by design – share the data your way!

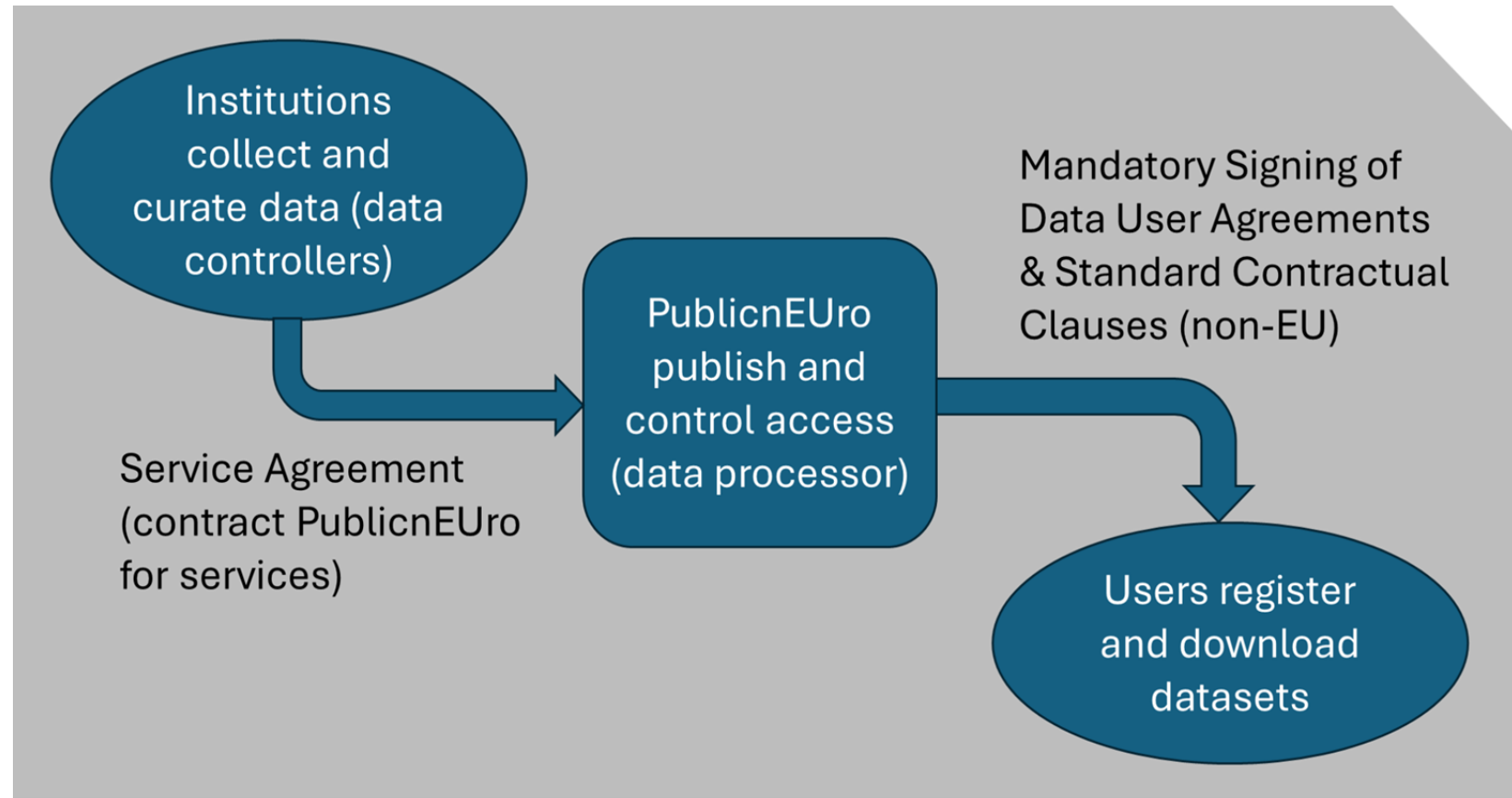


Figure 1: overview of the data workflow (boxes and arrows) and legal tools around it

Each data set has data user agreements that must be signed before getting access. PubicnEUro allows all combinations of sharing by identifying each user.

Benefits of using PublicnEUro

- Data hosting (Accessible)
 - Metadata enhancement and searchability (Findable)
 - Unload the administrative burden
-
- data access is managed for you
 - non-EU paperwork can just as easily be subcontracted
 - as 'paper' heavy or lightweight as needed (subcontract signing of DUA and SCC or have oversight)



PublicEUro hosting and security (ready)

- **Public not open:** there is a public register of the data and access is controlled (that public register is the open metadata catalogue)
- Tight control from regulation (GDPR) perspective, 'entrusted' admin access
- Fully identified users are given time-limited access tokens to datasets. All transactions are recorded and available for audits.
- State-of-the-art secured data centre for hosting (computerome DK - ISO27001)



PublicnEUro Findability and Accessibility (WIP)

- Data are secured behind firewall - the filesystem maps data using URL/Datalad hashes
- Metadata are open <https://github.com/Public-nEUro/DataCatalogue> + DOI
- DataLad serves data and DataCat serves metadata - future work will allow data to be also findable on OpenNeuro and metadata browsable and searchable with Neurobagel just as regular open datasets.



Thank you for your attention

Come and find me at poster 2216 to discuss how we can help you share your data

PublicnEUro: a European platform to share neuroimaging datasets publicly



Dr Cyril Pernet, Dr Melanie Ganz & Prof Gitte Knudsen
Neurobiology Research Unit, Rigshospitalet, Denmark

<https://public-neuro.eu>
publicneuro@nru.dk



Background: Data sharing using a web platform is becoming integral to the *research life cycle*. Not only does data sharing allow reproducing analyses, a tenet of experimental research, it also allows deepening analysis of existing datasets, combining data, meta-analysing, and asking outright new questions. Because *brain imaging data (MRI, PET, EEG)* can be seen as personal data, this activity is challenging for EU-based researchers who must comply with the **General Data Protection regulation**. Here we introduce PublicnEUro: a platform for EU-regulation-compliant data sharing.

Public not Open: We introduce the notion of public data, that is data which are publicly declared (open metadata, doi, etc) but not openly available (i.e. they are access controlled).

FAIR and GDPR compliant:

- Datasets are shared using the institution data user agreements ('share your way')
- Users must sign those to gain access, ensuring GDPR compliance at the national and institutional levels (Figure 1).
- Datasets are highly structured using the *Brain Imaging Data Structure*, the international standard for organising and naming, ensuring interoperability and reuse.

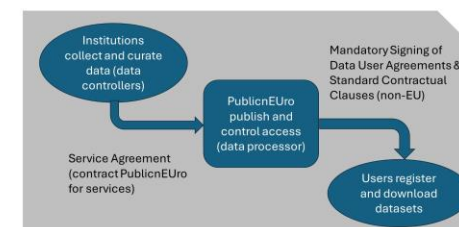


Figure 1: overview of the data workflow (boxes and arrows) and legal tools around it



Security at each data point:

- Dataset deposit is made into a sandbox by data controllers and curated by 'entrusted' PublicnEUro collaborators.
- Datasets are stored on secured servers on Computerome Denmark (GDPR-compliant cloud and HPC platform).
- Fully identified users are given time-limited access tokens to datasets. Data access are recorded and available for audits.

By (i) allowing institution-specific data user agreements, (ii) verifying users' identities, and (iii) controlling data access, one can bring data to users.

Current work aims at increasing findability, re-usage, and integration, working with Datalad and the Datalad Catalogue, OpenNeuro, and Neurobaguel.

novo
nordisk
fonden
Benefiting people and society



DataCat
GENERATE A USER-FRIENDLY DATA BROWSER FROM STRUCTURED METADATA

OpenNEURO





OHBM 2024

JUNE 23~27, SEOUL, KOREA