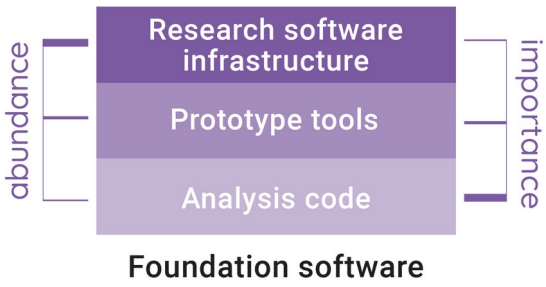
EVERSE - European Virtual Institute for Research Software Excellence

EVERSE aims to establish a framework for research software excellence. The project brings together people to collaboratively design and champion best practices for high-quality, sustainable research software.You can learn more about the project at https://everse.software/.

Reference Framework

EVERSE is building a reference model to link various aspects relevant to research software quality such as technical aspects, FAIR Software, Open Software and sustainability aspects.



RSQkit

We aim to create a community-driven framework that empowers researchers and developers to produce, maintain, and share high-quality research software.

For this we will create the Research Software Quality Kit (RSQkit), integrating tools and services for software quality assessment, and engaging with the research community through pilots, training, and recognition initiatives.

Network

EVERSE will establish a European network for research software quality, fostering collaboration among researchers, developers, and infrastructure providers.

The project will actively engage with five European science clusters, ensuring that the developed framework is tailored to the specific needs of each scientific domain.

- **LAPP** leads work on “Processes and Tools for better Software Quality”.
- **IJCLab** develops chat agents for training, kit browsing and code review.

Interested?
Get in touch with us and join the EVERSE network!

@eosc_everse

company/eosc-everse/

EVERSE-ResearchSoftware

GDPR-compliant sharing of human neuroimaging data using Shanoir

Camille Maumet

▶ **To cite this version:**

Camille Maumet. GDPR-compliant sharing of human neuroimaging data using Shanoir. EOSC France Tripartite Event 2024, Sep 2024, Paris, France. 42, pp.1945 - 1951, 2021. hal-04698499

HAL Id: hal-04698499
https://inria.hal.science/hal-04698499

Submitted on 16 Sep 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.

Distributed under a Creative Commons Attribution 4.0 International License

GDPR-compliant sharing of human neuroimaging data using Shanoir

presented by **Camille Maumet** at **EOSC France Tripartite Event** (Sept. 2024)

A **web-platform** for **medical imaging** data

With the impetus towards a more **open science**, researchers are looking for practical solutions to share more data in support of their findings. In the field of **brain imaging**, European researchers must comply with the General Data Protection Regulation (GDPR)when sharing human neuroimaging data.

We present **Shanoir**, a long-standing open source web-platform for medical imaging data that can be used to share neuroimaging data in compliance with GDPR.

Open Source Software | **Shared governance** | **Compliant with GDPR**

Image credits: Open source CC BY 3.0 by Sofian ID thenounproject; GDPR CC BY 3.0 by Abd Majd thenounproject;

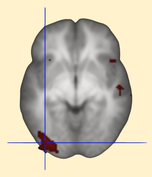
Neurinfo — https://project.inria.fr/neurinfo/

OFSEP Observatoire Français de la Sclérose en Plaques — https://www.ofsep.org/fr/

Data sharing with Shanoir

Brain images

Directly identifiable information removed



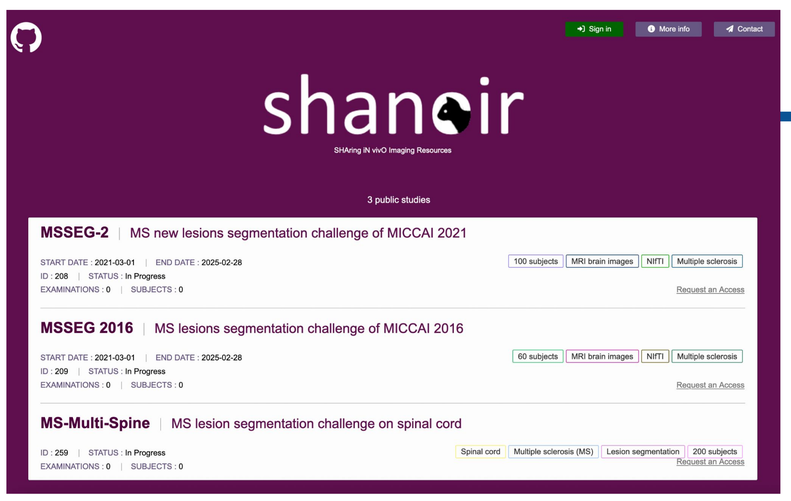
3D image

Non anonymous (**Individualisation**)

Data sharing must comply with **GDPR**

(Open Brain Consent, HBM 2021)

3 steps



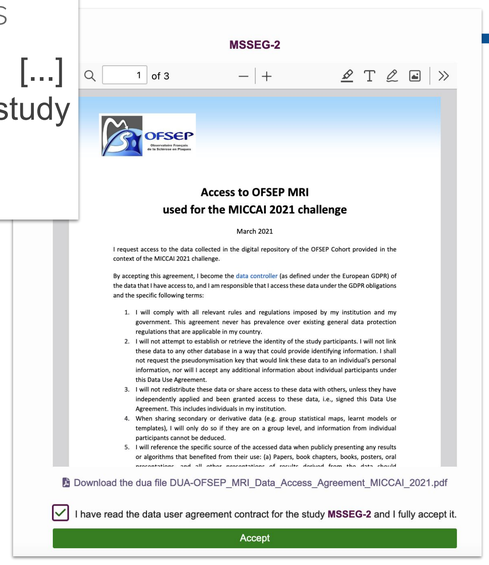
1- List **shared datasets**

https://shanoir.irisa.fr/

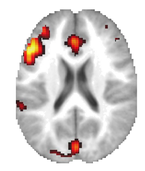
Data User Agreement

Signed by researchers

"[...] I will **not** attempt to [...] **retrieve the identity** of the study participants."



2- Sign the **Data User Agreement**



3- Access the **data**

Further resources

Shanoir is Open Source software: https://github.com/fli-iam/shanoir-ng

Developments are led by a community of members through the **Shanoir Consortium**: https://project.inria.fr/shanoir/consortium/

References

Barillot, C., Bannier, E., Commowick, O., Corouge, I., Baire, A., Fakhfakh, I., Guillaumont, J., Yao, Y., & Kain, M. (2016). Shanoir: Applying the Software as a Service Distribution Model to Manage Brain Imaging Research Repositories. Frontiers in ICT, 3. https://doi.org/10.3389/fict.2016.00025

Commowick, O., Cervenansky, F., Cotton, F., & Dojat, M. (2021). MSSEG-2 challenge proceedings: Multiple sclerosis new lesions segmentation challenge using a data management and processing infrastructure.

Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., Pop, S. C., Girard, P., Améli, R., Ferré, J.-C., Kerbrat, A., Tourdias, T., Cervenansky, F., Glatard, T., Beaumont, J., Doyle, S., Forbes, F., Knight, J., Khademi, A., … Barillot, C. (2018). Objective Evaluation of Multiple Sclerosis Lesion Segmentation using a Data Management and Processing Infrastructure. Scientific Reports, 8(1), 13650. https://doi.org/10.1038/s41598-018-31911-7

Commowick, O., Kain, M., Casey, R., Ameli, R., Ferré, J.-C., Kerbrat, A., Tourdias, T., Cervenansky, F., Camarasu-Pop, S., Glatard, T., Vukusic, S., Edan, G., Barillot, C., Dojat, M., & Cotton, F. (2021). Multiple sclerosis lesions segmentation from multiple experts: The MICCAI 2016 challenge dataset. NeuroImage, 244, 118589. https://doi.org/10.1016/j.neuroimage.2021.118589

Open Brain Consent. (2021). The Open Brain Consent: Informing research participants and obtaining consent to share brain imaging data. Human Brain Mapping. https://doi.org/10.1002/hbm.25351

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

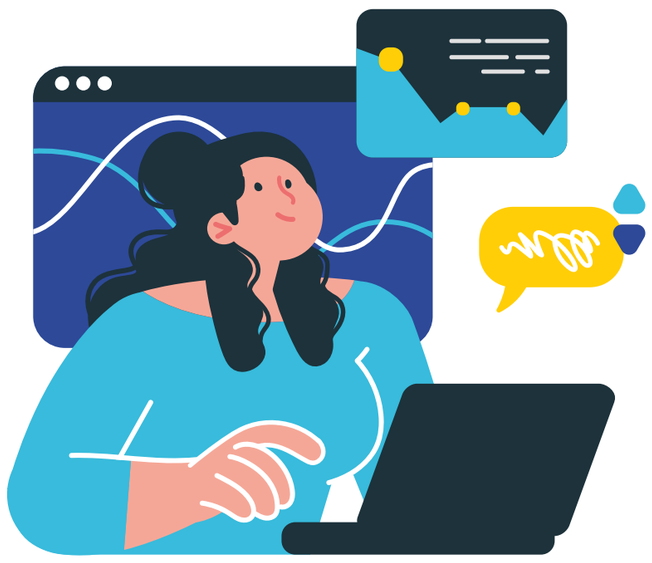
Inria

Shanoir core development is supported by
Shanoir Consortium (InriaSoft) and **FLI-IAM**

FLI France Life Imaging

FAIRCORE4EOSC

Core Components Supporting a FAIR EOSC



Join the Community

Scan the QR Code

faircore4eosc.eu

Twitter

Linkedin

FAIRCORE4EOSC is developing a set of nine EOSC-Core components necessary to enable a FAIR EOSC ecosystem, leveraging the solutions developed in previous projects and closely collaborating with the wider EOSC community in Horizon Europe.

9 EOSC-Core components necessary to enable a FAIR EOSC ecosystem

CAT
EOSC Compliance Assessment Toolkit

DTR
EOSC Data Type Registry

MSCR
EOSC Metadata Schema & Crosswalk Registry

PIDGraph
EOSC PID Graph

PIDMR
EOSC PID Meta Resolver

RAiD
EOSC Research Activity Identifier Service

RDGraph
EOSC Research Discovery Graph

RSAC
EOSC Research Software APIs & Connectors

SWHM
EOSC Software Heritage Mirror

FAIRCORE4EOSC Case Studies

Five user-centric case studies are driving the development and testing of the new components, ensuring they are co-designed and tailored to users needs.

- European Integration of National-level Services
- Climate Change
- Mathematics
- EOSC Service Providers
- Social Sciences and Humanities

Partners

CSC – Coordinator

Australian Research Data Commons

ATHENA Research & Innovation Information Technologies

CERN

CLARIN Common Language Resources and Technology Infrastructure

Consiglio Nazionale delle Ricerche

COMMpla Communication Platforms and Online Solutions

CYFRONET

Data Archiving and Networked Services DANS

DKRZ Deutsches Klimarechenzentrum

FIZ Karlsruhe Leibniz Institute for Information Infrastructure

DataCite Find, Access, and Reuse Data

grnet

GWDG Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Helsingin Yliopisto

Inria

OpenAIRE

Schloss Dagstuhl Leibniz-Zentrum für Informatik

Software Heritage

SURF

The National Library of Finland

Trust-IT Services communicating to markets

University of Warsaw

Funded by the European Union

FAIR-IMPACT

Expanding FAIR solutions across EOSC

Authors/Contributors
Sara Pittonet TRUST-IT
Ingrid Dillo DANS
Clément Jonquet INRAE
Kevin Ashley DCC

EOSC is an integral part of, and supports the European Commission's strategy for realising the European Research Area, in particular the policy priorities of Open Innovation, Open Science and Open to the World and the goal of findable, accessible, interoperable and reusable (FAIR) data[1]. EOSC's realisation is dependent upon national infrastructure such as NFDI.

With the ambitious goal to realise an EOSC of FAIR data and services, the FAIR-IMPACT project supports the implementation of **FAIR principles** across scientific communities and research outputs at European, national and institutional level through a matrix of coordination, engagement and support activities.

[1]EOSC Strategic Research and Innovation Agenda (SRIA) 1.1 - Nov 2022

Practices, policies and tools towards FAIR data management

FAIR-IMPACT supports the implementation of FAIR-enabling practices, tools and services across scientific communities at a European, national and institutional level.

We do this by connecting knowledge across scientific domains on **persistent identifiers (PIDs), metadata and ontologies, metrics, certification and interoperability aspects** via a community-led approach.

We work with research institutes and organisations to identify practices, policies, tools and technical specifications towards a FAIR data management cycle starting from real-life Use Cases on:

- social sciences & humanities
- photon & neutron sciences
- life sciences & agri-food
- environmental sciences

FAIR-IMPACT Open Calls For Support

FAIR is not a binary concept and cannot be achieved by any single action or actor. FAIR is a shared journey involving several stakeholders over the entire research lifecycle.

Through our series of open calls, FAIR-IMPACT raises awareness and supports the uptake of existing and emerging tools, solutions and approaches to begin or progress FAIR-enabling journeys. Three key stakeholder groups are being supported: **research organisations, repositories & data services providers, and national level initiatives.**

Route 1: Guidance and one-to-one support

- **Practical advice and expert guidance** on key FAIR-enabling topics.
- **One-to-one support** to:
  - Undertake a supported self-assessment of current FAIR-enabling practices
  - Develop a FAIR implementation action plan
  - Devise an engagement strategy

Route 2: Support to test specific tools, approaches and solutions

- A dedicated **support package** with workshops, expert guidance and advice from FAIR-IMPACT project partners and resource providers help the participants to implement a tool or approach.
- **Financial support** to enable participation, payable by invoice upon successful completion of the support action.

The Third Open Call Route 2 opens on 30 September till 30 November

Learn more

FIF FAIR implementation framework of resources

by EOSC | FAIR-IMPACT

To support the adoption of FAIR enabling tools, methods and resources, FAIR-IMPACT developed a FAIR Implementation Framework (FIF). A key component of the Framework is a catalogue of existing FAIR-enabling tools, approaches and solutions available to support FAIR Implementation in a practical sense.

Browse the online Catalogue and/or suggest a new resource

The FAIR-IMPACT Synchronisation Force workshop series

Three online synchronisation workshops bringing together FAIR and EOSC ecosystem projects, infrastructures, and initiatives in Europe to assess the progress of work related to FAIR-IMPACT focus areas and SRIA objectives. Each workshop will result in a report.

The 2024 edition of the Synchronisation Force Workshop is taking place from 3 September to 6 November, with a final plenary session on 7 November.

Synchronisation Force ONLINE WORKSHOP

2024 Edition

3 September - 7 November

Scan QR Code and learn more

FAIR-IMPACT

Metadata and Ontologies work-package outcomes & use cases

WP4

Greater and more harmonised use of **semantic artefacts** throughout the EOSC ecosystem, leading to semantic interoperability **within and between disciplines.**

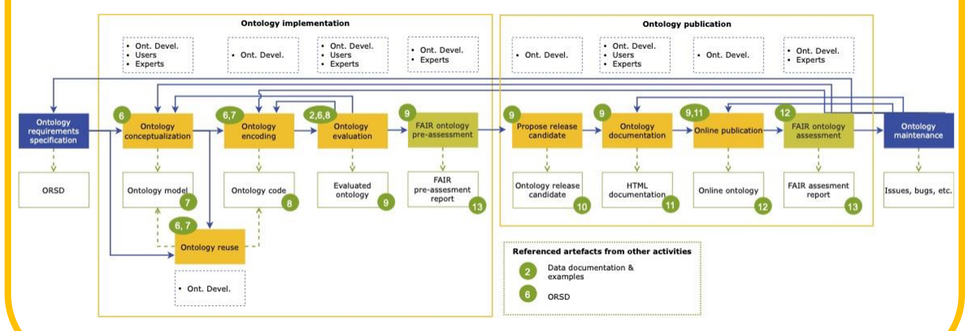
… gathers, synthesises and disseminates the materials needed to **federate the approach to metadata and ontologies** at various organisational and technical levels within European Open Science Cloud (EOSC).

Authors
Clément Jonquet (INRAE)
Nina Grau (INRAE)
& WP4 participants

… **semantic artefacts (SA)** – a broader term to include ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards – in EOSC.

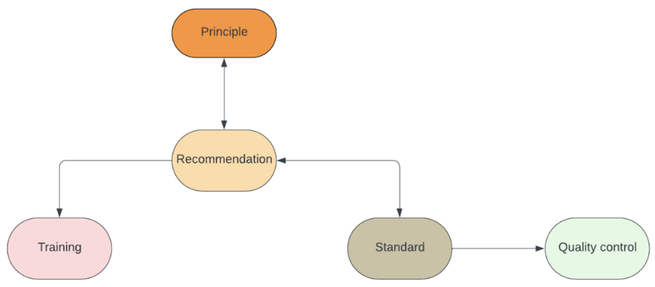
Semantic Artefact « *FAIR-by-design* » methodology

- Built on the **Linked Open Terms methodology**
- Implement **FAIR guidelines** and develop an harmonised mechanism to describe SA with the **MOD ontology**



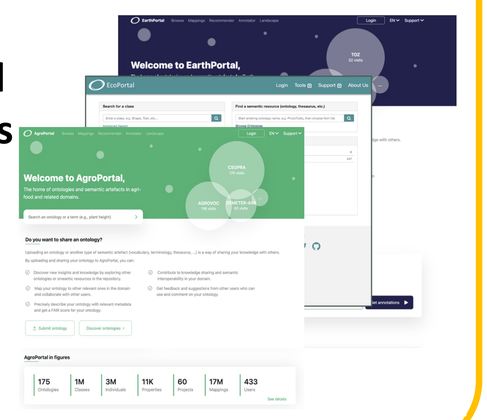
Semantic Artefact Governance

- Built on a community survey
- Review and analyse SA governance **components**
- Propose 3 SA governance **models**



Semantic Artefact Catalogues

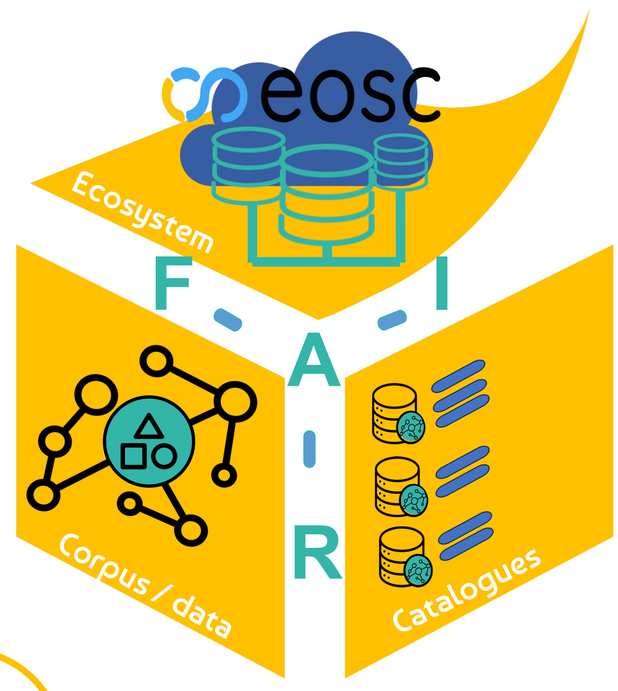
- Built on existing tools and technology (e.g., AgroPortal, OntoPortal)
- Support new communities
- Review **173**
- **catatogues and 15 technologies**
- Develop **FAIR -enabling criteria**



Did you know that we are developing Semantic Artefact framework for EOSC ?

The research objects that we are dealing with:

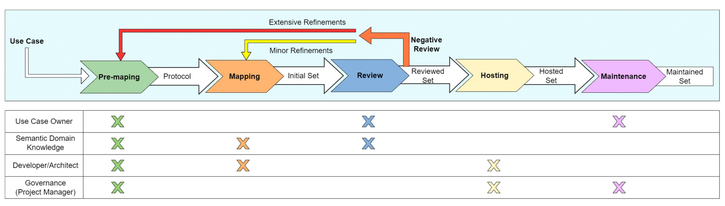
- **Semantic Artefact**
- **Mapping & crosswalk**
- **Research Software**



Mappings

RESEARCH DATA ALLIANCE

- Built on community survey & workshops
- Review and analyse **practices and needs**
- **FAIR mappings** specification



Research Software

eosc | FAIRCORE4EOSC

- Built on use cases
- Review **software metadata landscape**
- Guidelines RS metadata standard

Intrinsic metadata	**Extrinsic metadata**
Refers to the information that is inherently embedded within a software source code artifact. It includes metadata files that are captured in the main source code directory (such as README file, codemeta.json file, etc.)	Refers to the information that is external to the software source code artifact. It includes metadata elements that provide context, provenance, and additional information about the software.

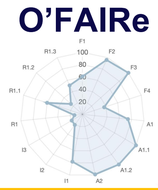
Used in data catalogues

- Built on multiples use cases
- Foster **use** of semantic artefacts into discipline **data repositories**
- Develop **generic connectors between data repos and SAC**

Semantic Artefact FAIRness assessment

- Built on recognised FAIR tools and method
- **Implement and foster SA FAIRness** among various communities

O'FAIRe



Deliverables and milestones

- M4.1 - Semantic artefact governance models: example of community practices
- D4.1 - Semantic artefact governance models and disciplinary approaches for inclusion within EOSC
- M4.2 - Processes & tools to engineer FAIR semantic artefacts
- D4.3 - Specification of shared metadata description of semantic artefacts and their catalogues including common reference API
- M4.4 - Review of Semantic Artefact Catalogues and guidelines for serving FAIR semantic artefacts in EOSC
- D4.4 - Guidelines for recommended metadata standard for research software within EOSC

Contact:
clement.jonquet@inrae.fr

Follow us on our channel!

Newsletter · Twitter · Youtube · LinkedIn

Recherche Data Gouv in the EOSC landscape

EOSC Tripartite 2024

***Recherche Data Gouv* is a complete ecosystem for sharing and opening research data, aiming to:**

- → Leave no research team without any solution to open or share their data
- → Make support a central element of the system
- → Recognise, federate and expand existing initiatives

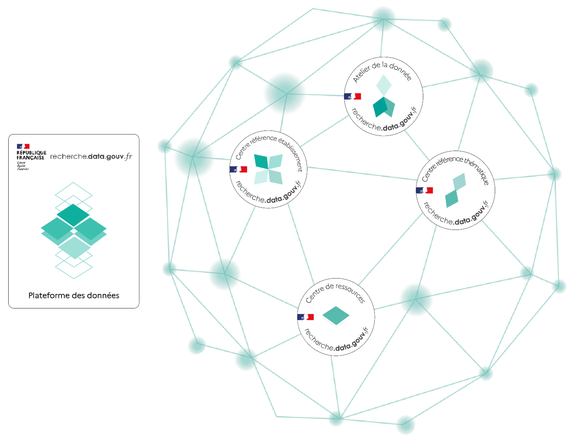
For a more cumulative, robust and reproducible science that is transparent and accessible to all

What

National Research Data Platform

A sovereign solution for depositing, publishing, reporting, discovering, accessing and reusing data through:

- ✓ A national repository with a generic collection and institutional spaces
- ✓ A registry to search for data published in the repository itself or other external repositories



National Federation of Competence Centres

A support offer on research data management through the data life cycle, close to research teams and organised:

- ✓ Geographically, providing generalist expertise
- ✓ Thematically, providing expertise in given scientific fields
- ✓ Globally, supporting common tools and best practices

Make National and European levels converge in *Recherche Data Gouv* strategy

How

Contribute to the setup of the European network of trustworthy data repositories

Recherche Data Gouv is involved in the FIDELIS project from 2025 onwards.

FIDELIS
A European network of Trustworthy Digital Repositories

Join the EOSC Competence Centres network

Recherche Data Gouv leads the first French national Competence Centre within the Skills4EOSC project.

Skills 4 eosc NETWORK
TRUSTED COMPETENCE CENTRE

Why

- **Contribute** to the definition of "trustworthy"
- **Share and harmonise support** best practices
- **Deal with technical and semantic interoperability**, through a better use of controlled vocabularies
- **Develop our national strategy** coherently within the European landscape
- **Enhance international visibility** of French research data

- **Train the trainers** and gain expertise
- **Benefit from a large panel of services, trainings and training material**
- **Harmonise our practices** and build something together
- **Share best practices** and organisation
- **Work on shared issues** to identify common solutions
- **Be more efficient** and gain time

Points of attention

Take into account the ecosystem's complexity
- → Each of *Recherche Data Gouv*'s bricks has its own organisation and needs
- → Many of its stakeholders have their own EOSC strategy

Communicate along the way
- → EOSC awareness is heterogeneous internally
- → Our EOSC strategy only makes sense if understood and endorsed by all

Synchronise with competent and legitimate stakeholders
- → We need to work together to ensure that the diversity of the French landscape doesn't lead to fragmentation

Next step

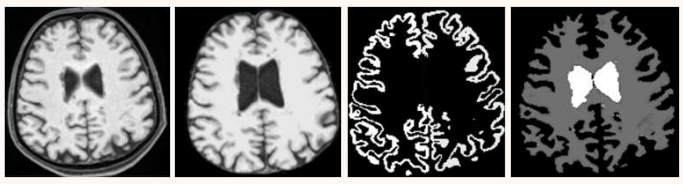
Join the EOSC nodes build-up phase and offer *Recherche Data Gouv*'s resources and services through a future EOSC Node of the EOSC Federation

The Virtual Imaging Platform

Laboratoire CREATIS, INSA Lyon, Université Claude Bernard Lyon 1, UMR CNRS 5220, Inserm U1294, Université Jean Monnet

What VIP provides

- Scientific applications as a Service
  - ✓ Web portal: https://vip.creatis.insa-lyon.fr/ [1]
  - ✓ 25+ medical imaging applications available as a service



Example of white/grey matter brain segmentation with Freesurfer on VIP
Credits: Berardino Barile and Dominique Sappey-Marinier, CREATIS

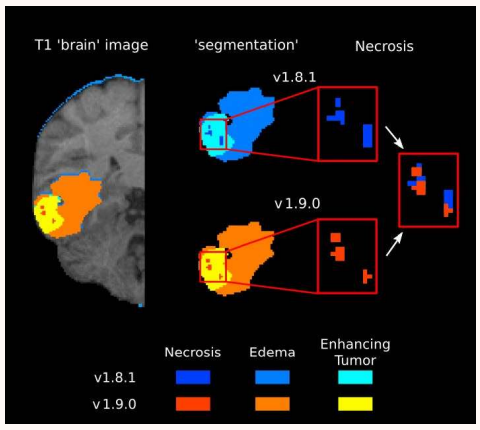
- Transparent access to computing resources
  - ✓ 110+ CPU years (EGI biomed VO) used from 09/2023 to 09/2024
- Large community
  - ✓ 1500+ registered users
  - ✓ 77 publications with results obtained on VIP
- Open and reproducible science
  - ✓ Zenodo, DOIs, Containers, Boutiques

Turning applications into services

- Boutiques [2]
  - ✓ Describe, publish and execute applications across platforms
  - ✓ http://boutiques.github.io/
- Integrate a new application in VIP in 3 main steps
  - ✓ Build a container image containing the scientific application
  - ✓ Create a Boutiques descriptor of the application
  - ✓ Send them to the VIP team

ReproVIP

- French ANR JCJC project (grant N°ANR-21-CE45-0024-01)
  - ✓ https://reprovipgroup.pages.in2p3.fr/documentation
  - ✓ Aim: evaluate and improve the reproducibility of scientific results
  - ✓ Focus on computational reproducibility

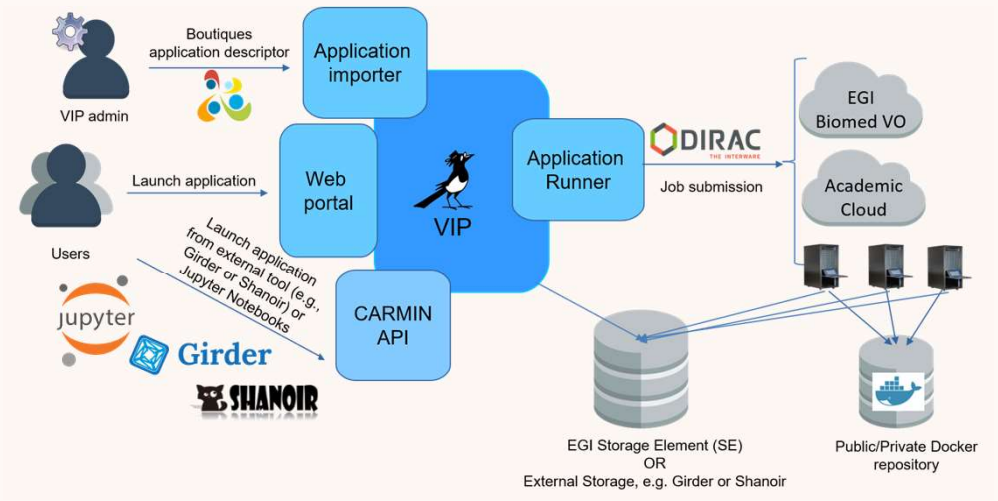


Differences in tumor segmentation outputs obtained with two different versions of Brain Tumor Segmentation (BraTS) pipeline on the same input image, as presented in [3].

Scientific Challenges

- Research teams work on solving a scientific hard problem
  - ✓ Solutions are evaluated on a given set of data according to the guidelines given by the organizers
- Challenges MICCAI 2016 & 2021 [4]
  - ✓ Partners: OFSEP, Empenn, FLI
  - ✓ VIP as execution platform
- MSSEG-2 https://portal.fli-iam.irisa.fr/msseg-2/
  - ✓ Automatic segmentation of tissues and lesions in MRI brain scans
  - ✓ 31 pipelines integrated in VIP for the challenge (24 teams)

VIP ecosystem



Organization

- Source code
  - ✓ Cecill-B license
  - ✓ https://github.com/virtual-imaging-platform
  - ✓ Archived on Software Heritage
- Platform developed, deployed and maintained at CREATIS
  - ✓ Multiple collaborations

References

1. Tristan Glatard, et al. 2013, « A Virtual Imaging Platform for Multi-Modality Medical Image Simulation ». *IEEE Transactions on Medical Imaging 32* (1): 110-18, https://hal.science/inserm-00762497v1
2. Tristan Glatard, et al. 2018, «Boutiques: a flexible framework to integrate command-line applications in computing platforms». GigaScience, Volume 7, Issue 5, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6007562/
3. Morgane Des Ligneris, et al. 2023,. « Reproducibility of Tumor Segmentation Outcomes with a Deep Learning Model ». International Symposium on Biomedical Imaging (ISBI), https://hal.science/hal-04006057v1.
4. Olivier Commowick, et al. 2018. «Objective Evaluation of Multiple Sclerosis Lesion Segmentation using a Data Management and Processing Infrastructure». Nature Publishing Group, https://www.nature.com/articles/s41598-018-31911-7

LABEX PRIMES UNIVERSITÉ DE LYON
anr

CONTACT
vip-support@creatis.insa-lyon.fr