Received: 29 July 2022 | Revised: 18 October 2022 | Accepted: 18 October 2022

DOI: 10.1111/ejn.15854

SPECIAL ISSUE REVIEW

EJN European Journal of Neuroscience FENS WILEY

Italian, European, and international neuroinformatics efforts: An overview

Alberto Redolfi[1] | **Damiano Archetti**[1] | **Silvia De Francesco**[1] | **Claudio Crema**[1] | **Fabrizio Tagliavini**[2] | **Raffaele Lodi**[3,4] | **Roberta Ghidoni**[5] | **Claudia A. M. Gandini Wheeler-Kingshott**[6,7,8] | **Daniel C. Alexander**[9,10] | **Egidio D'Angelo**[7,8]

[1]Laboratory of Neuroinformatics, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

[2]Scientific Directorate, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy

[3]Functional and Molecular Neuroimaging Unit, IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy

[4]Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy

[5]Molecular Markers Laboratory, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

[6]NMR Research Unit, Queen Square MS Center, Department of Neuroinflammation, UCL Institute of Neurology, London, UK

[7]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

[8]Brain Connectivity Center, IRCCS Mondino Foundation, Pavia, Italy

[9]Centre for Medical Image Computing, University College London, London, UK

[10]Department of Computer Science, University College London, London, UK

Correspondence
Alberto Redolfi, Laboratory of Neuroinformatics, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy.
Email: aredolfi@fatebenefratelli.eu

Funding information
H2020 European Institute of Innovation and Technology, Grant/Award Numbers: 666992, 945539; Italian Ministry of Health, Grant/Award Numbers: JPNDJPco-fuND2-"EDADS"No.825664, RCR2020-23670067, RCR-2021-23671214, RRC-2016-2361095, RRC-2017-2364915, RRC-2018-2365796, RRC-2019-23669119_001; Ministry of Economy and Finance, Grant/Award Numbers: CCR-2017-23669078, CCR-2018-23669122

Edited by: Giuseppe Pignataro

Abstract

Neuroinformatics is a research field that focusses on software tools capable of identifying, analysing, modelling, organising and sharing multiscale neuroscience data. Neuroinformatics has exploded in the last two decades with the emergence of the Big Data phenomenon, characterised by the so-called 3Vs (volume, velocity and variety), which provided neuroscientists with an improved ability to acquire and process data faster and more cheaply thanks to technical improvements in clinical, genomic and radiological technologies. This situation has led to a 'data deluge', as neuroscientists can routinely collect more study data in a few days than they could in a year just a decade ago. To address this phenomenon, several neuroimaging-focussed neuroinformatics platforms have emerged, funded by national or transnational agencies, with the following goals: (i) development of tools for archiving and organising analytical data (XNAT, REDCap and LabKey); (ii) development of data-driven models evolving from reductionist approaches to multidimensional models (RIN, IVN, HBD, EuroPOND, E-DADS and GAAIN BRAIN); and (iii) development of e-infrastructures to provide sufficient computational

power and storage resources (neuGRID, HBP-EBRAINS, LONI and CONP). Although the scenario is still fragmented, there are technological and economical attempts at both national and international levels to introduce high standards for open and Findable, Accessible, Interoperable and Reusable (FAIR) neuroscience worldwide.

KEYWORDS
Alzheimer's disease, computational models, FAIR, international cooperation, neuroinformatics

1 | INTRODUCTION

Neuroinformatics is conceived as a set of advanced information technologies at the service of neurology. Neuroinformatics combines data across all scales and levels of neuroscience in order to understand the complex functions of the brain (Nayak et al., 2018).

The primary outcome of neuroinformatics is the development of computational models to better understand the mechanisms of brain diseases, to measure disease evolution, to support physicians in diagnosis and, ultimately, to evaluate the treatment and drug efficacy (Darwish, 2018; Kruse & Beane, 2018). Neuroinformatics is at the centre of the entire spectrum of neuroscience, from basic to applied clinical research, forming a cyclic loop that helps translate new ideas into breakthrough developments. To make neuroinformatics an essential and functional discipline of neuroscience, it was soon recognised that interoperability among centres worldwide was a fundamental requirement and a high priority.

Interoperability is challenged by the increasing amount of information, also known as Big Data, that has involved neuroscience over the past two decades (Calhoun, 2015). The volume, velocity and variety (3Vs) currently available for data acquisition are the cause of this phenomenon. Brain researchers today are able to collect more data in a few experiments than researchers could have collected in an entire career a generation ago, and this is predicted to expand exponentially over the next decade (Van Horn, 2021).

Interoperability has been gradually advanced by the International Neuroinformatics Coordinating Facility (INCF), because the lack of standards and best practices for multimodal information exchange makes sophisticated analysis impossible. The INCF recently agreed on a set of community developed guidelines to ensure that neuroscience data, and all digital objects, are findable, accessible, interoperable and reusable: they are known as the FAIR principles (Abrams et al., 2021). This represents a tremendous methodological effort for the development of neuroinformatics e-infrastructures. By adopting the

Abbreviations: AD, Alzheimer's Disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; AI, Artificial Intelligence; AIBL, Australian Imaging Biomarkers and Lifestyle Study of Ageing; AWS, Amazon Web Services; BIDS, Brain Imaging Data Structure; Brain-CODE, Brain Institute Centre for Ontario Data Exploration; BSB, Brain Scaffold Builder; CONP, Canadian Open Neuroscience Platform; CPU, Central Processing Units; CSA, Coordination and Support Actions; DICOM, Digital Imaging and COmmunications in Medicine; DL, Deep Learning; EBM, event-based model; EBRAINS, European Brain ReseArch INfraStructure; EC, European Commission; E-DADS, Early Detection of Alzheimer's Disease Subtypes; EDF, European Data Format; EEG, electroencephalogram; EHR, electronic health records; ELISA, enzyme-linked immunosorbent assay; ESFRI, European Strategy Forum on Research Infrastructures; EuroPOND, European Progression Of Neurological Disease; FAIR, Findable, Accessible, Interoperable and Reusable; FDG, 18F-FluoroDeoxyGlucose; FET, Future Emerging Technologies; GAAIN, Global Alzheimer's Association Interactive Network; GB, GigaBytes; GPPM, Gaussian Process Progression Model; GPU, Graphics Processing Unit; HBD, Health Big Data; HBP, Human Brain Project; HL7-FHIR, Fast Healthcare Interoperability Resources; HPC, High Performance Computing; IBI, International Brain Initiative; ICT, Information and Communication Technologies; IDA, Image and Data Archive; IEEE, Institute of Electrical and Electronics Engineers; INCF, International Neuroinformatics Coordinating Facilities; INFN, National Institute for Nuclear Physics; IRCCS, Scientific Institutes of Hospitalization and Care; IT, Information Technology; IVN, National Virtual Institute; JPND, Joint Programme Neurodegenerative Disease research; LEASPy, LEArning Spatiotemporal Patterns in Python; LONI, Laboratory of Neuro Imaging; MEF, Ministry of Economy and Finance; MIP, Medical Informatics Platform; ML, Machine Learning; MoH, Ministry of Health; MRI, Magnetic Resonance Imaging; NIfTI, Neuroimaging Informatics Technology Initiative; NIH, National Institute of Health; NLP, Natural Language Processing; NMF, Nonnegative Matrix Factorisation; PET, Positron Emission Tomography; PI, Principal Investigator; PNNR, Italian National Recovery and Resilience Plan; RAM, Random Access Memory; REDCap, Research Electronic Data Capture; RIN, IRCCS Network of Neuroscience and Neurorehabilitation; SOP, standardised operational procedures; SuStaIn, Subtype and stage inference; TB, TeraBytes; TVB, The Virtual Brain; USC, University of Southern California; VIP, Virtual Imaging Platform; VRRS, Virtual Reality Rehabilitation Systems; XNAT, Extensible Neuroimaging Archive Toolkit.

FAIR principles, other neuroscientists would be able to share their data to enhance the knowledge and improve artificial intelligence (AI)-based transfer learning, which is becoming the cutting-edge technology for many health funding agencies (Italian Ministries, European Commission [EC] and US National Institutes of Health [NIH]) to invest in.

In order to face the issue of interoperability and ensure efficient collaboration, neuroscientists rely on infrastructures for data sharing and analysis, which in the case of neuroinformatics often take the form of web-based platforms.

In Italy, neuroinformatics is playing a prominent role in helping national health institutions to share data and resources, and improving health standards in the country. Since 2017, the Ministry of Health (MoH) and the Ministry of Economy and Finance (MEF) have committed to fund Italian neuroinformatics platforms and coordinate actions to improve neuroscientists' cooperation and competitiveness (Nigri et al., 2022).

Europe (EU) has driven neuroinformatics from the outset with the support of coherent and strategic approaches that have enabled the design and development of dozens of e-infrastructures and mobilised nearly €20 billions of investment across the EU over the past 20 years.

Internationally, the United States and Canada have helped shape the neuroscience research landscape by identifying, developing and supporting new and existing e-infrastructures to provide researchers with first class shared resources to better understand and potentially solve large-scale problems such as neurodegeneration or dementia through multicentre collaborations (Mueller et al., 2005).

Neuroinformatics platforms for data acquisition and development of data-driven models have evolved at all three geographic scales, with overlaps and incompatibilities. Thus, in this dynamic scenario, there is an increasing need for standards to make international cooperation more efficient, with the ultimate goal of transparent and accessible open science (Vicente-Saez & Martinez-Fuentes, 2018).

We will here review the most significant neuroimaging-focussed neuroinformatics initiatives, platforms and networks that have been or are being carried out in Italy, Europe, and at the global level, highlighting their objectives, architectures, the services they offer and the type of data they rely on.

A glossary of some of the specialist terms used in the review is provided in Box 1.

Box 1 | Glossary:

- Algorithm: a set of instructions for accomplishing a specific task implemented in a single software
- Artificial intelligence: the theory and development of computer models capable of performing tasks that normally require human intelligence, for example, speech recognition, visual perception, decision making, translation between languages and more
- Blockchain: an emerging technology used to create innovative and secure solutions in various sectors, including health care. A Blockchain network is used in health-care system to preserve and exchange patient data through hospitals, diagnostic laboratories and physicians
- Computational model: an algorithm or pipeline capable of simulating and studying complex systems using mathematics, physics and computer science
- Container: a virtualised runtime environment used in application development. Typical examples of containers are Docker or Singularity. They are typically used to build, run and deploy applications that are isolated from the underlying physical hardware
- Database: structured information stored in a computer system in such a way that it can be easily viewed or manipulated
- Data lake: an easily accessible, centralised repository for large amounts of unstructured data. Data lakes can contain hundreds of terabytes or even petabytes of raw data
- Deep learning: a type of machine learning based on artificial neural networks in which the computational model gradually extracts more and more information from the data. Deep learning networks are created using programming languages (Python, C++) and specialised libraries (TensorFlow and PyTorch)
- Hadoop: an open-source distributed processing framework that manages data processing and storage for big data applications in scalable clusters of servers
- High Performance Computing (HPC): a type of computing infrastructure that uses

supercomputers and computer clusters to solve advanced computational problems
- Machine learning: the use and development of computer models that are able to learn and adapt without following explicit instructions, using algorithms and statistical models to analyse and draw inferences and rules from patterns data
- MapReduce program: a programming model from Google for processing huge data sets on large clusters of servers
- Question answering bot: also known as QABot, is an artificial intelligence tool capable of searching for strings, patterns or meanings to provide answers to questions asked by humans in a natural language
- Pipeline: also known as a workflow, is a software implementation with a well-defined input and output. A pipeline may consist of one or more algorithms and other software steps drawn from one or more software tools that may also generate intermediate data
- Web server: a computer that runs websites. The web server's job is to store, process, and deliver web pages to users. This communication is done using the secure hypertext transfer protocol (HTTPS)
- Single sign on: also known as SSO, is a user authentication service that permits to use one set of login credentials (e.g. name and password) to access multiple applications

1.1 | The neuroinformatics playground

Computational neuroscience generates complex data that must be stored efficiently in large quantities. This is especially true for raw data and neuroimaging-derived data, which can require several GBs per single subject. Specific informatics tools have been developed for each data typology. Among the most widely used are Research Electronic Data Capture (REDCap), (XNAT) and LabKey.

REDCap (https://www.project-redcap.org) (Harris et al., 2009), developed at Vanderbilt University in 2004, has a user-friendly web-based interface and allows researchers to collect alphanumeric data such as sociodemographic, clinical and neuropsychological information into relational databases (MySQL or MariaDB). XNAT (https://www.xnat.org) (Marcus et al., 2007) was developed by the Neuroinformatics Research Group at Washington University. It is an open-source software platform specifically designed to manage imaging data through a PostgreSQL database and provides an online image viewer that supports multiple neuroimaging formats. LabKey (https://www.labkey.com) (Nelson et al., 2011), originally developed in 2003 at the Fred Hutchinson Cancer Research Center as a platform for proteomic data management, is now an end-to-end platform for full molecular data integration that uses relational databases (PostgreSQL or MSSQL). LabKey is implemented in Java and runs on the Apache Tomcat web server.

In the last two decades, AI and ML paradigms have become ubiquitous in various aspects of biomedical research, and neuroinformatics is no exception (Vu et al., 2018). The heterogeneous, albeit not fully harmonised, but representative cohorts of clinical and research data archived in the aforementioned databases have enabled AI and ML to build several types of data-driven models: (i) progression models, used to simulate disease evolution and predict future disease behaviour, attempting to answer specific questions such as: 'how long will it take the pathology to blow up?' (Heo et al., 2019; Oxtoby & Alexander, 2017); (ii) classifiers, used to assign a predefined diagnostic category to a patient and to answer the question: 'is the patient suffering from Alzheimer's Dementia (AD) or Parkinson's Dementia?' (Clark et al., 2016); and (iii) cluster analysis tools, used to group subjects into non-predefined diagnostic categories based on common underlying biomarker patterns (Ulfenborg et al., 2021). Recently, deep learning (DL) has emerged as an AI technique that has made great leaps forward in solving complex neuroscience problems such as image processing in computer vision (LeCun et al., 2010) or Natural Language Processing (NLP) (Lee et al., 2020). The main drawbacks are the need for a huge amount of data for training and the availability of powerful computing resources.

The models and algorithms developed in neuroinformatics can be used to build platform-based tools for the research and clinical communities. In the first case, neuroinformatics platforms provide neuroscientists with powerful tools for group analysis. In the second case, neuroinformatics platforms provide physicians with a single-case tool, to contrast each patient's data against a validated normative distribution representative of the general population, or a second-opinion service to confirm the physician's diagnosis or prognosis (De Francesco et al., 2021).

In recent years, several architectures have been developed in neuroinformatics to host data, algorithms

and pipelines (Figure 1). The main ones are (i) centralised, where each partner has the ability to upload its data to a web server that contains, stores, processes, develops and provides access to the data and final trained models; (ii) distributed, where each partner builds its computational model using its local dataset and exploits a web server to distribute pre-trained models—the data are never shared in this architecture; and (iii) federated (McMahan et al., 2017), where raw data are kept securely at each institution, and only the aggregated data and final trained models are exchanged on the web server—this is useful for sharing the knowledge built with the partners without sharing all data directly (Abdulrahman et al., 2021).

All architectures have some pros and cons. The distributed and federated architectures, unlike the centralised ones, mitigate the risk of data movement and minimise privacy issues. Unlike the federated architecture, the distributed architecture cannot improve learning based on the other institutions' models (Campos et al., 2022). The federated architecture requires the development and maintenance of multiple data sharing policies.

1.2 | Ten years of disease investigation with neuroinformatics platforms

In this section, we will review the main neuroimaging-focussed neuroinformatics platforms and initiatives.

1.3 | Italian national level

1.3.1 | RIN

The Neuroscience and Neurorehabilitation Network (RIN) is the largest Italian research initiative in this field. RIN was established in 2017 by the Italian MoH to promote collaboration among the Scientific Institutes of Hospitalization and Care (IRCCS). RIN focusses on the research areas of genomics, neuroimmunology, tele-neurorehabilitation and neuroimaging with neuroinformatics as the basis for building the different platforms. RIN includes 30 Italian IRCCS hospitals (https://www.reteneuroscienze.it/en).

The goal of the genomic section is to share and consult clinical and genomic data to establish a national

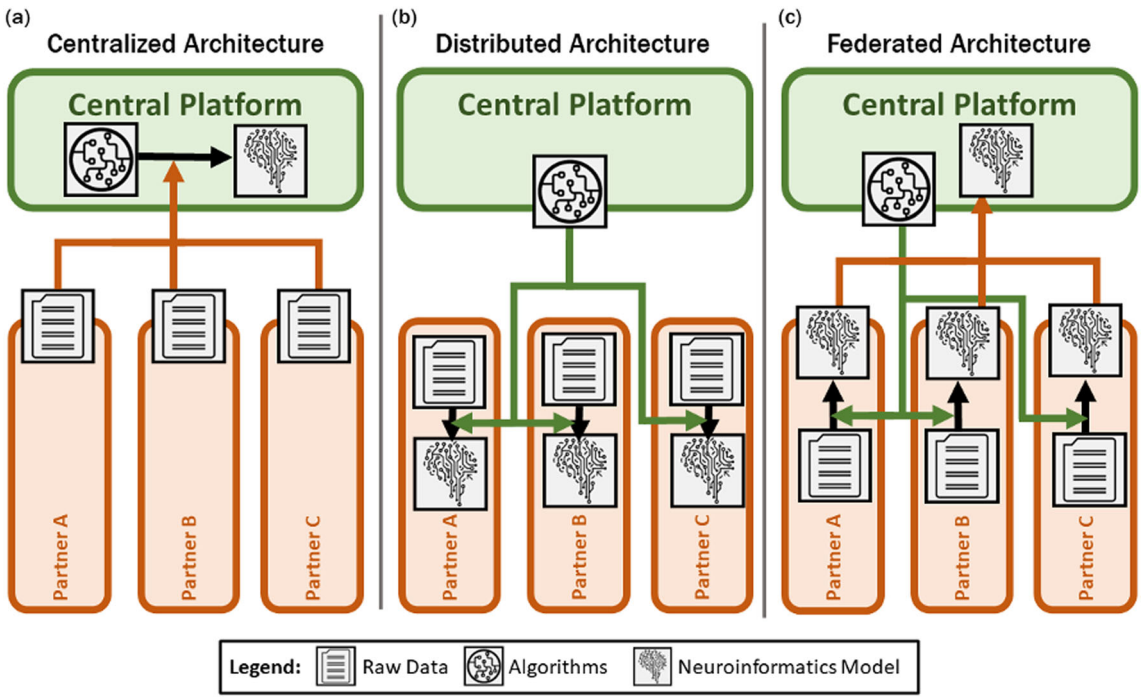

FIGURE 1 Schematisation of the architectures of different neuroinformatics platforms. (a) Centralised architecture: Partners share their data on a central platform where data are pooled and processed via algorithms to inform neuroinformatics models. (b) Distributed architecture: Algorithms or pretrained models are distributed to partners from a central platform. Neuroinformatics models are then informed separately by each partner on the basis of the partner's in-house data. (c) Federated architecture: Same as distributed architecture, but separated models and/or aggregated data are then shared by partners in order to build a unified neuroinformatics model on the central platform. Green boxes indicate neuroinformatics platforms (central platform), and orange boxes indicate single partner facilities (local platform).

repository using the same standardised operational procedures (SOPs). The neuroimmunology branch of RIN is developing SOPs for standardisation of conformational antibodies, immunohistochemistry, enzyme-linked immunosorbent assays (ELISA) and blots adopting the external quality assessment scheme validation (Braga et al., 2018). In the area of tele-neurorehabilitation, each participating IRCCS uses virtual reality rehabilitation systems (VRRS) (Olivieri et al., 2013). Tele-neurorehabilitation is supposed to provide digital tools based on tablets, interactive software and virtual games to improve patient monitoring, out-of-hospital rehabilitation, diagnosis and prevention. A recent study from the RIN network evaluated the efficacy of face-to-face cognitive VRRS and provided preliminary evidence to support individualised treatment and telerehabilitation delivery for cognitive rehabilitation (Manenti et al., 2020).

Three neuroinformatics platforms (https://www.neugrid2.eu/, https://arianna.pi.infn.it/it and https://ebrains.eu/service/medical-informatics-platform/) are in use for neuroimaging (Nigri et al., 2022). Several tools and computational models for quality assessment, image analysis, and ML analyses have been developed (Redolfi et al., 2020) or are under scientific validation. To measure the stability of high-field Magnetic Resonance Imaging (MRI) acquisitions, three types of phantoms (FUNSTAR [https://www.goldstandardphantoms.com/products/funstar/], ACR small [American College of Radiology, 2018], ACR Large [Chen et al., 2004]) have been collected. To date, over 380 scans per phantom have been gathered at regular intervals (once per month on average). Sociodemographic, clinic, and neuropsychological data have also been collected from over 350 patients, as well as the harmonised protocol with T13D, Flair, DTI, RS-fMRI and QSM scans. Finally, over 370 scans of murine models were collected longitudinally at 4, 8, 15, 19 months and beyond.

The platform used by RIN is a centralised architecture (Tables 1 and 2). RIN represents the so-called first component (x axis) of the original multidimensional matrix design, consisting of the Italian technology platforms. The other two axes of the matrix are described in the following two sections.

1.3.2 | IVN

There are currently five active National Virtual Institutes (IVNs). The National Virtual Institute for Dementias (Principal Investigator [PI] F. Tagliavini) was established in 2019, followed by the National Virtual Institute for Parkinson's Disease (PI P. Cortelli) and the National Virtual Institute for Multiple Sclerosis (PI G. Martino). In 2021, the National Virtual Institute for Cerebrovascular Diseases (PI P. Calabresi) and the National Virtual Institute for Rare Neurological Diseases (PI D. Pareyson) were launched.

The primary goal of the IVN is to harmonise the study protocols and procedures and standardise the diagnostic and therapeutic skills of physicians in the participating IRCCSs. The IVN also has the following secondary objectives: (i) to streamline investments and resources; (ii) to maximise the collection of electronic health records (EHR) and biological samples for research projects and advanced clinical trials; and (iii) to identify new biomarkers, therapeutic targets and innovative therapies.

As far as the National Virtual Dementia Institute is concerned, socio-demographics data, medical history, risk factors, comorbidities, objective examinations, drug therapy and neuropsychological test results are automatically extracted from various EHRs using NLP algorithms based on QABots or via classical Regular Expressions. The database is implemented using REDCap, hosted in the facilities of the IRCCS Carlo Besta in Milan.

The main platform design used by IVN is a distributed architecture (Tables 1 and 2). Each IVN has its own REDCap and XNAT platforms that follow a minimal common data schema and specific variables for each domain to structure the relationships between the data.

The IVN represents the second component (y axis) of the matrix design desired by MoH representing the clinical and therapeutic pillars.

1.3.3 | HBD

In February 2020, the MoH and MEF launched the Health Big Data (HBD) project, the only project to date involving four of the biggest Italian networks: the RIN, the Network Alliance Against Cancer, the Cardiology Network and the Pediatrics Network.

A large part of the project is based on the creation of (i) a set of local platforms in each participating IRCCS to ensure extraction, integration, and interoperability of clinical and research data adopting standards (e.g. Fast Healthcare Interoperability Resources [HL7-FHIR], Brain Imaging Data Structure [BIDS; Gorgolewski et al., 2016], Digital Imaging and COmmunications in Medicine [DICOM]) as much as possible; and (ii) a centralised High Performance Computing (HPC) e-infrastructure at the National Institute for Nuclear Physics (INFN) to ensure high throughput connectivity and AI analyses of the aggregated data. The use of 'open-source' tools (hardware or software) already in use at major international facilities will be considered for the establishment of local and central e-infrastructure platforms.

TABLE 1 Neuroinformatics initiatives at Italian, European and international level

Platforms/initiatives	Funding agency	Budget	Principal investigator	Partners involved	Funding duration
RIN	Italian Ministry of Health and Ministry of Economy and Finance	€ 5 millions	Tagliavini – Lodi	30	5 years
IVN	Italian Ministry of Health	€ 8 millions	Tagliavini – Lodi	28	2 years
HBD	Italian Ministry of Health and Ministry of Economy and Finance	€ 55 millions	Pelicci	51	10 years
neuGRID	European Commission, DG info and DG connect (7FP)	€ 7.4 millions	Frisoni – Redolfi	1	7 years
EuroPOND	European Commission (H2020)	€5 millions	Alexander	8	5 years
E-DADS	European Commission (JPND)	€1.8 millions	Alexander	6	3 years
HBP-EBRAINS	European Commission; FET flagship (H2020)	€1.2 billions	Świeboda – Jirsa – Bjiaale	100	10 years
LONI	NIH	$218 millions	Toga	63	18 years
CONP	Brain Canada platform support Grant competition award and sponsors*	$10 millions	Evans	32	3 years
GAAIN	Alzheimer's Association	$6.5 millions	Toga – Frisoni	59	5 years
US BRAIN	NIH	$5 billions	Ngai	43	15 years

Note: Table 1 lists the major neuroinformatics initiatives and their characteristics. Some initiatives last beyond the agreed project's duration because of external issues (e.g. COVID-19 and project extensions). Field-weighted citation impact is a metric from SciVal (https://www.elsevier.com/solutions/scival) that indicates the average number of citations of each platform publication compared to the average number of citations of all other similar publications in the same field in the Scopus database (https://www.scopus.com). A field-weighed citation impact above 1.00 indicates a higher number of publications than average, whereas a field-weighted citation impact below 1.00 indicates a lower number of citations than average. Acronyms: 7FP: seventh framework programme; CONP, Canadian Open Neuroscience Platform; DG: Directorate-General; EBRAINS, European Brain ReseArch INfraStructure; EC: European Commission; E-DADS: Early Detection of Alzheimer's Disease Subtypes; EuroPOND: European Progression Of Neurological Disease; FET: Future and Emerging Technologies; GAAIN: Global Alzheimer's Association Interactive Network; H2020: Horizon 2020; HBP: Human Brain Project; HBD: Health Big Data; IVN: National Virtual Institute; JPND: Joint Programme Neurodegenerative Disease research; LONI: Laboratory Of Neuro Imaging; RIN: Neuroscience and Neurorehabilitation Network; NIH: National Institutes of Health of the United States of America; BRAIN: Brain Research through Advancing Innovative Neurotechnologies; NA: not applicable; *: Irving Ludmer Family Foundation, McGill/MNI Tanenbaum Open Science Institute, Fonds de Recherche du Québec, Ontario Brain Institute/Rotman Research Institute, École Polytechnique, Institut de Cardiologie de Montréal, Western University, McLaughlin Centre/University of Toronto, Université Laval, University of Calgary, Douglas Hospital Research Foundation, Human Brain Project, University of British Columbia, Concordia University, Dell/EMC, IBM, Compute Canada and Québec Bio-Imaging Network.

TABLE 1 (Continued)

Platforms/initiatives	Real timeline extension	Infrastructure typology	FAIR compliant	Current status	Field-weighted citation impact	Maturity level
RIN	Started 2016, ongoing	Centralised	No	Active platform	.58	Basic
IVN	Started 2020, ongoing	Distributed	No	Active project	.44	Basic
HBD	Started 2020, ongoing	Federated	No	Active project/active platform	NA	Basic
neuGRID	Started 2008, ongoing	Centralised	Yes	Active platform	1.15	Intermediate
EuroPOND	2016–2020(concluded)	Distributed	NA	Ended project	2.13	Advanced
E-DADS	Started 2020, ongoing	Federated	No	Active project	NA	Basic
HBP-EBRAINS	Started 2013, ongoing	Federated	Yes	Active project/active platform	1.63	Advanced
LONI	Started 2004, ongoing	Centralised	Yes	Active platform	1.84	Advanced
CONP	Started 2017, ongoing	Distributed	Yes	Active project/active platform	1.54	Advanced
GAAIN	Started 2013, ongoing	Federated	Yes	Active platform	2.54	Advanced
US BRAIN	Started 2013, ongoing	Distributed	Yes	Active project/active platform	2.09	Advanced

Note: Table 1 lists the major neuroinformatics initiatives and their characteristics. Some initiatives last beyond the agreed project's duration because of external issues (e.g. COVID-19 and project extensions). Field-weighted citation impact is a metric from SciVal (https://www.elsevier.com/solutions/scival) that indicates the average number of citations of each platform publication compared to the average number of citations of all other similar publications in the same field in the Scopus database (https://www.scopus.com). A field-weighed citation impact above 1.00 indicates a higher number of publications than average, whereas a field-weighted citation impact below 1.00 indicates a lower number of citations than average. Acronyms: 7FP: seventh framework programme; CONP, Canadian Open Neuroscience Platform; DG: Directorate-General; EBRAINS, European Brain ReseArch INfraStructure; EC: European Commission; E-DADS: Early Detection of Alzheimer's Disease Subtypes; EuroPOND: European Progression Of Neurological Disease; FET: Future and Emerging Technologies; GAAIN: Global Alzheimer's Association Interactive Network; H2020: Horizon 2020; HBP: Human Brain Project; HBD: Health Big Data; IVN: National Virtual Institute; JPND: Joint Programme Neurodegenerative Disease research; LONI: Laboratory Of Neuro Imaging; RIN: Neuroscience and Neurorehabilitation Network; NIH: National Institutes of Health of the United States of America; BRAIN: Brain Research through Advancing Innovative Neurotechnologies; NA: not applicable; *: Irving Ludmer Family Foundation, McGill/MNI Tanenbaum Open Science Institute, Fonds de Recherche du Québec, Ontario Brain Institute/Rotman Research Institute, École Polytechnique, Institut de Cardiologie de Montréal, Western University, McLaughlin Centre/University of Toronto, Université Laval, University of Calgary, Douglas Hospital Research Foundation, Human Brain Project, University of British Columbia, Concordia University, Dell/EMC, IBM, Compute Canada and Québec Bio-Imaging Network.

TABLE 2 Neuroinformatics platform characteristics

		RIN	IVN	HBD	neuGRID	EuroPOND
Target audience		Researchers	Researchers clinicians	Researchers	Researchers clinicians	Researchers
Services offered	Data upload	Yes	Yes	Yes	Yes	No
	Data download	Yes	No*	No*	Yes	No
	Single-case analysis	No*	Yes	No*	Yes	No
	Group analysis	Yes	No*	Yes	No*	Yes
	Computations	Yes	Yes	Yes	Yes	No
	Code hosting	No	No	No*	No	Yes
	Research type	Basic/translational	Clinical	Translational/clinical	Translational/clinical	Translational
Available cohorts	Subjects number	>1000	>5000	>20,000	>5000	>10,000
	Diagnostic categories	CN DLB AD FTD PPA MM	CN MCI AD PD MS	CN MCI AD CP CVD	CN CBD LBD MDD MCI MSA PD SMI SVD AD FTD	CN MCI AD MS CJD ND
	Assessments	Cross.	Long.	Long.	Cross.	Long.
Available data types	Demographic	Yes	Yes	Yes	Yes	Yes
	Clinical	Yes	Yes	Yes	Yes	Yes
	NPSY	Yes	Yes	Yes	Yes	Yes
	MRI 1.5 T	No	Yes	Yes	Yes	Yes
	MRI 3.0 T	Yes	Yes	Yes	Yes	Yes
	FDG PET	No	Yes	No*	Yes	No
	Amyloid PET	No	Yes	No*	No*	No
	Tau PET	No	No*	No*	No*	No
	Blood	Yes	Yes	Yes	No*	No
	CSF	Yes	Yes	Yes	Yes	Yes
	Genetics	Yes	Yes	Yes	Yes	Yes
	Wearable biosensor data	No	No	Yes	No	No
	EEG	Yes	Yes	No*	No*	No

Note: Table 2 lists the target audience, provided services, cohort characteristics and main features of the data exposed by the neuroinformatics initiatives. Acronyms: AD: Alzheimer's Disease; ALS, Amyotrophic Lateral Sclerosis; ASD: Autism Spectrum Disorder; CBD: CorticoBasal Degeneration; CJD: Creutzfeldt–Jakob Disease; CN: Cognitive normal subjects; CONP, Canadian Open Neuroscience Platform; CP: Cancer Patients; Cross.: Storage of only cross-sectional data; CSF: CerebroSpinal Fluid; CVD: CardioVascular Disease; DLB: Dementia with Lewy Bodies; E-DADS: Early Detection of Alzheimer's Disease Subtypes; EuroPOND: European Progression Of Neurological Disease; FDG: 18F-FluoroDeoxyGlucose; FTD: FrontoTemporal Dementia; GAAIN: Global Alzheimer's Association Interactive Network; HBD: Health Big Data; HBP: Human Brain Project; IVN: National Virtual Institute; LBD: Lewy Body Dementia; Long.: Storage of both cross-sectional and longitudinal data; LONI: Laboratory Of Neuro Imaging; MCI: Mild Cognitive Impairment; MDD: Major Depressive Disorder; MM: Murine Models; MRI: Magnetic Resonance Imaging; MS: Multiple Sclerosis; MSA: Multiple System Atrophy; ND: Neurodevelopmental Disorders; Npsy: Neuropsychological; PD: Parkinson's Disease; PET: Positron Emission Tomography; PPA: Primary Progressive Aphasia; RIN: Neuroscience and Neurorehabilitation Network; SMC: Subjective Memory Complaints; SMI: Severe Mental Illness; SVD: Small Vessel Disease; TBI: traumatic brain injury; *: Not available at the time of writing, but with the possibility of adding the biomarker or service in the future.

TABLE 2 (Continued)

	E-DADS	HBP-EBRAINS	LONI	CONP	GAAIN	US BRAIN
Target audience	Researchers	Researchers	Researchers	Researchers	Researchers	Researchers
Services offered	No	Yes	No	Yes	Yes	Yes
	No	Yes	Yes	Yes	Yes	Yes
	No*	No*	No	No	No	No*
	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	No	Yes	No	Yes
	Translational/clinical	Basic/translational	Basic/translational	Translational	Translational	Basic/translational
Available cohorts	>50,000	>250,000	>80,000	>100,000	>500,000	>150,000
	CN MCI AD	CN MCI AD TBI MM	CN SMC MCI AD PD FTD ASD MM	CN MCI AD PD FTD ASL MS MM ND	CN MCI AD	CN MCI AD PD TBI MM
	Long.	Long.	Long.	Long.	Long.	Long.
Available data types	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	No*	Yes	Yes	No*	Yes	Yes
	No*	No*	Yes	No*	Yes	Yes
	No*	No*	Yes	No*	No*	Yes
	No*	No*	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	Yes	Yes	Yes	Yes	Yes	Yes
	No	Yes	No	Yes	No	Yes
	No	Yes	No	Yes	No	Yes

Note: Table 2 lists the target audience, provided services, cohort characteristics and main features of the data exposed by the neuroinformatics initiatives. Acronyms: AD: Alzheimer's Disease; ALS, Amyotrophic Lateral Sclerosis; ASD: Autism Spectrum Disorder; CBD: CorticoBasal Degeneration; CJD: Creutzfeldt–Jakob Disease; CN: Cognitive normal subjects; CONP, Canadian Open Neuroscience Platform; CP: Cancer Patients; Cross.: Storage of only cross-sectional data; CSF: CerebroSpinal Fluid; CVD: CardioVascular Disease; DLB: Dementia with Lewy Bodies; E-DADS: Early Detection of Alzheimer's Disease Subtypes; EuroPOND: European Progression Of Neurological Disease; FDG: 18F-FluoroDeoxyGlucose; FTD: FrontoTemporal Dementia; GAAIN: Global Alzheimer's Association Interactive Network; HBD: Health Big Data; HBP: Human Brain Project; IVN: National Virtual Institute; LBD: Lewy Body Dementia; Long.: Storage of both cross-sectional and longitudinal data; LONI: Laboratory Of Neuro Imaging; MCI: Mild Cognitive Impairment; MDD: Major Depressive Disorder; MM: Murine Models; MRI: Magnetic Resonance Imaging; MS: Multiple Sclerosis; MSA: Multiple System Atrophy; ND: Neurodevelopmental Disorders; Npsy: Neuropsychological; PD: Parkinson's Disease; PET: Positron Emission Tomography; PPA: Primary Progressive Aphasia; RIN: Neuroscience and Neurorehabilitation Network; SMC: Subjective Memory Complaints; SMI: Severe Mental Illness; SVD: Small Vessel Disease; TBI: traumatic brain injury; *: Not available at the time of writing, but with the possibility of adding the biomarker or service in the future.

The type of data to be collected and subsequently shared is heterogeneous and includes clinical data (EHRs and patient follow-up data), '-omics' data (genomics, transcriptomics, proteomics and metabolomics), imaging and radiomics data. In the medium term, data from biosensors, environmental, and social activities will also be included (Tables 1 and 2). Great emphasis will be placed on innovative ways to integrate unstructured data (Wieder & Nolte, 2022), especially with advanced data lake technologies (Google Big Lake, Amazon Web Services [AWS] Lake formation). The goal of HBD is to develop predictive and prescriptive computational analysis models capable of generating knowledge in real time and opening new possibilities for prevention and therapy. Special attention will be placed on improving the Information Technology **(**IT) tools already available at participating IRCCSs. The HBD is currently working on the development of NLP tools with the goal of developing a reliable pipeline for the extraction of medical data from EHR for the creation of high-quality curated research databases. The main architecture used by the HBD is the federated one.

The HBD represents the third component (z axis) of the matrix originally desired by MoH. It is the abstraction layer of the RIN/IVN framework from the brain to other organs (i.e. from brain to heart) or other clinical domains (i.e. from neuro-pathologies to onco-pathologies).

1.4 | European level

1.4.1 | NeuGRID

The neuGRID platform (https://neugrid2.eu) (Redolfi et al., 2009), originally founded by EC in two successive waves of the 7FP, is a web-based HPC e-infrastructure designed to help neuroscientists to perform high-throughput analyses.

It provides an array of atomic algorithms and libraries (Freesurfer, FSL, ANTS, SPM and MINC tools) for quantification of surrogate MRI/PET imaging biomarkers of neurodegeneration (De Francesco et al., 2021; Redolfi et al., 2015); ML tools for single case analysis (Archetti et al., 2021; Redolfi et al., 2020); and secure archiving repositories based on a MySQL relational database.

NeuGRID has been identified by the Re3data initiative (Pampel et al., 2013) as an official research data repository compatible with the FAIR principle (Martone, 2022). NeuGRID contains data from more than 5000 subjects (Tables 1 and 2) from seven open-access multicentre studies (I-ADNI, PharmaCOG, VITA, OASIS, ARWIBO, WMH-AD and EDSD) organised in the BIDS format.

In neuGRID, the collection of large amounts of image data is associated with computationally intensive data analysis. The allocated computational resources consist of 350 Central Processing Units (CPU) cores, 2000 GigaBytes (GB) of Random Access Memory (RAM), and 150 TeraBytes (TB) of physical memory providing users with coarse-grained parallelisation via a Sun Grid Engine scheduler that enables faster data processing compared to a traditional scenario whilst preventing users from wasting time installing and learning complex software.

These advanced computing resources and the user-friendly environment, coupled with specific sections for each pipeline, have made neuGRID the core platform for multiple Italian (RIN, IVN and HBD), European (European Progression of Neurological Disease [EuroPOND] and E-DADS) and international (GAAIN) projects. NeuGRID employs mainly a centralised architecture (Figure 1a).

1.4.2 | EuroPOND

At the European level, the growing interest in neuroinformatics under Horizon 2020 has opened up the possibility of funding specific projects. EuroPOND has been exemplary in the development and validation of a number of data-driven algorithms that have successfully modelled a wide range of neurodegenerative diseases such as dementia (Archetti et al., 2019; Oxtoby et al., 2017; Young et al., 2014), multiple sclerosis (Dekker et al., 2021; Eshaghi et al., 2018), Parkinson's disease (Oxtoby et al., 2021), prion disease (Pascuzzo et al., 2020), Huntington's disease (Wijeratne et al., 2018, 2021) and amyotrophic lateral sclerosis (Gabel et al., 2020)) as well as normal ageing (Vinke et al., 2018) and neurodevelopment (Gui et al., 2019; Sa de Almeida et al., 2021).

EuroPOND led the way in extending computational models from pure and well-collected research data to clinical not fully curated and harmonised data for testing disease models. In addition to demographic, clinical, cognitive, and biological features, special emphasis was placed on neuroimaging data following the holistic approach typical of neuroinformatics (Table 2).

Most of the algorithms developed in EuroPOND are based on the fundamental idea that a disease follows an average trajectory with acceptable variability across patients, and that the trajectory can be inferred based on historical patients' data (Oxtoby & Alexander, 2017).

The goal of a common disease trajectory has been achieved using different approaches: (i) discrete models, in which disease is represented by a sequence of discrete

time steps; and (ii) spatiotemporal models, in which disease is modelled as a continuous set of biomarker trajectories over time.

EuroPOND validated several discrete disease models: the event-based model (EBM) (Fonteijn et al., 2012; Young et al., 2014), where disease evolution is represented by a sequence of events corresponding to biomarkers becoming abnormal. Variants of EBM were developed (Venkatraghavan et al., 2019, 2021), differing primarily in how normal and abnormal biomarkers are distinguished, as no a priori labelling of data was assumed. EBMs are powerful tools because they could be trained on cross-sectional data. EBMs, however, had significant limitations, mainly because the event sequences do not provide information about the timing and spacing between events, and the notion that biomarkers suddenly step from normality to abnormality may be biologically inaccurate.

These issues were overcome by the spatiotemporal modelling approaches. Two spatiotemporal models have been developed in EuroPOND, namely LEASPy (LEArning Spatiotemporal Patterns in Python) (Koval et al., 2018, 2021) and the Gaussian Process Progression Model (GPPM) (Abi Nader et al., 2020; Lorenzi et al., 2019). Both approaches rely on longitudinal data for training, and in both cases, individual trajectories are usually translated and warped to build a continuous trajectory in the biomarker space–time. The main difference between these two approaches is that LEASPy assumes a fixed shape for the biomarker trajectory (typically sigmoid), whereas GPPM makes no assumptions about the shape of each biomarker trajectory.

Indeed, the a priori assumption of a unique trajectory common to all patients may not be biologically accurate, so EuroPOND also developed a subtype model called SuStaIn (subtype and stage inference) (Young et al., 2018) that combines subtype modelling and disease progression modelling, allowing the discovery of disease variants that correspond to different disease progression patterns. SuStaIn has been used to identify neurodegeneration patterns in AD (Archetti et al., 2021; Young et al., 2018), fronto-temporal dementia (Young et al., 2018) and multiple sclerosis (Eshaghi et al., 2021). The main architecture used by EuroPOND was primarily a distributed architecture (Table 1), and the software and models developed within the consortium are publicly available (https://github.com/EuroPOND).

1.4.3 | E-DADS

E-DADS (Early Detection of Alzheimer's Disease Subtypes) is a Joint Programme Neurodegenerative Disease research (JPND) initiative launched in December 2020. E-DADS (https://e-dads.github.io) aims to unify the models previously developed in EuroPOND (SuStain and LEASPy) with the pipeline of Nonnegative Matrix Factorisation (NMF) to create a single hypermodel for AD. NMF (Ten Kate et al., 2018) is a ML based approach for detecting AD subtypes in imaging data that allows identification of atrophy features that are correlated in one subset of individuals but not in other subsets. Such nonlinear associations cannot generally be detected using classical methods.

The hypermodel of E-DADS is based on neuroimaging, clinical, cognitive, biological, genetic data and polygenic risk scores that can quantify risk and predict disease progression years in advance. This may trigger interventions (lifestyle changes and drug treatments) that can delay or slow down the onset of dementia (Ten Kate et al., 2018).

E-DADS data include more than 50,000 subjects in the AD spectrum from multiple data sets (Alzheimer's Disease Neuroimaging Initiative [ADNI], UK Biobank, Australian Imaging Biomarkers and Lifestyle Study of Ageing [AIBL], Amsterdam Data Cohort, Insight46, NeuGRID data, GAAIN data [see below]), creating a representative cohort that links early-life predictors to AD progression, and BIDS standard is employed in all centres to better handle the data.

An important part of the E-DADS project is based on the development of a neuroinformatics prototype, which will be made available through an online platform, to enable the use of the disease hypermodel by the general public, with dedicated areas for clinicians, pharmaceutical companies, patients and lay people.

E-DADS initiative is structured as a federated architecture as shown in Figure 1c.

1.4.4 | HBP-EBRAINS

Started in 2013, the Human Brain Project (HBP) is the largest and, with a duration of 10 years, also the longest brain science project ever conducted in Europe. The HBP is part of the Future Emerging Technologies (FET) Flagship initiative, a new funding programme launched by the EC that aims to achieve visionary goals such as understanding the multiscale organisation of the brain (Amunts et al., 2016, 2019).

HBP pries on advanced data processing methods, neuroinformatics tools and artificial intelligence computational models. The community can access the cloud-based services of HBP platforms through EBRAINS, the European Brain ReseArch INfraStructure (https://ebrains.eu).

The main EBRAINS services (Amunts, Axer, et al., 2022; Amunts, DeFelipe, et al., 2022) are (i) 'knowledge graph', a system to make data integrated and searchable, accessible, interoperable and reusable; (ii) 'atlases,' to make visualisation of brains comparable across species by highlighting similarities and differences, zooming in and out of each region of interest (up to 20 μm at cell resolution), extracting data, and using these data as input for simulations; (iii) 'simulation platforms' (D'Angelo & Jirsa, 2022) for modelling the brain such as the Brain Scaffold Builder (BSB), The Virtual Brain (TVB) and the Medical Informatics Platform (MIP). BSB developed a new framework for cerebral cortex reconstruction via morphologically realistic single-neuron models. TVB developed hundreds of interconnected circuits and networks of the brain, enabling automatic conversion of user-specific model equations into fast simulation code (Matzke et al., 2015; Schirner et al., 2022). MIP developed a federated imaging platform that enables the analysis of MRI/PET scans in European hospitals and research centres (Redolfi et al., 2020); and (iv) 'brain-inspired services and tools' to develop neuromorphic computing and neurorobotic platforms. These tools occupy a special position as they enable new insights into the brain or inspire the development of more efficient Information and Communication Technologies (ICT) for AI (e.g. innovative learning gradient approach that mimics neuron behaviour for DL applications (Göltz et al., 2021); co-design of robots with enhanced decision-making capabilities and contextual awareness for use in autonomous exploration in situations such as deep sea search and rescue in disaster areas; and high performance supercomputing).

EBRAINS neuroinformatics services are computationally intensive and require access to the latest supercomputing resources on the Fenix infrastructure, which brings together scalable storage and computing resources at several leading HPC sites, namely CINECA (Italy), CEA (France), BSC (Spain), JSC (Germany) and CSCS (Switzerland). The backend architecture envisioned by the HBP-EBRAINS is the federated one (Figure 1c).

1.5 | International level

1.5.1 | LONI

The Laboratory of Neuro Imaging (LONI), at the University of Southern California (USC) (https://www.loni.usc.edu/), is the most advanced neuroinformatics platform ever developed (Tables 1 and 2).

LONI provides several services to the community, mainly the provision of algorithms and data storage. As far as data analysis is concerned, LONI algorithms can be accessed both independently or through a graphical system (LONI pipeline). The LONI pipeline is a visual programming interface for the creation, execution and distribution of advanced neuroimaging analysis pipelines created by the user (Dinov et al., 2009; Rex et al., 2003) or directly by LONI developers. There are hundreds of workflows ranging from bioinformatics (BLAST or PLINK for genotype/phenotype data analysis) to imaging analysis (Freesurfer, FSL, ITK libraries for morphometric or functional analysis) specific to clinical or preclinical models.

The LONI Image and Data Archive (IDA), a large HPC infrastructure, provides tools and resources for de-identification, integration and sharing of neuroscience data. External researchers are granted access to LONI HPC resources on the basis of ad hoc scientific collaboration agreements. IDA is a repository for long-term and permanent storage of neuroimaging, clinical, biospecimen and genetic research data (Crawford et al., 2016) in accordance with the FAIR principles. Among IDA datasets, the ADNI has collected the most widely used cohort of data in AD research (Aisen et al., 2010). Since its inception in 2004, the ADNI has gone through several phases: ADNI-1 (2004–2009) collected demographic, clinical, biological and cognitive data from 800 subjects across the AD spectrum (Petersen et al., 2010), followed by ADNI-GO (2009–2011, 200 subjects) (Jack et al., 2010), ADNI-2 (2011–2016, 700 subjects) (Aisen et al., 2015) and ADNI-3 (2016–2021, 133 subjects) (Weber et al., 2021).

During all ADNI phases, subjects were followed longitudinally, and high-resolution neuroimaging data (T13D, T2/PD, Flair, fMRI, DTI and FDG-PET, amyloid PET and tau PET) were also collected. ADNI data have been published in nearly 4000 scientific articles, and ADNI has made an extraordinary contribution to paving the way for big data in neuroscience. LONI hosts many other public and private datasets not only on neurological but also on psychiatric disorders. The backend used by LONI is the centralised architecture (Figure 1a).

1.5.2 | Canadian neuroinformatics platforms

Canada plays a central role in the development of the neuroinformatics platforms (Das et al., 2017). The Brain Institute Centre for Ontario Data Exploration (BrainCODE) electronic infrastructure (Rotenberg et al., 2018; Vaccarino et al., 2018) represents one of the most advanced examples of neuroinformatics to date. BrainCODE provides specialised relational databases for data acquisition depending on the records to be archived, that is, (i) REDCap for clinical and neuropsychological

information; (ii) XNAT for imaging and electroencephalogram (EEG) data; and (iii) LabKey for the -omics data. All archived information complies with privacy and security regulations and specific format standards (DICOM, Neuroimaging Informatics Technology Initiative – NIfTI, European Data Format – EDF). To efficiently query, manage and analyse the large amount of electronic health records, Brain-CODE also combines innovative 'data lake' strategies with traditional databases. In this way, Brain-CODE can provide its 600 active users with efficient data services through web-based ad hoc dashboards and the ability to post-process brain data via R, Python or Jupyter notebooks. Brain-CODE relies on a computing cluster with 1000 CPU cores, a Graphics Processing Unit (GPU) node (dedicated to ML analytics) and 9000 GB RAM memory. Brain-CODE relies on the CANAIRE high bandwidth backbone (https://www.canarie.ca).

In 2017, the Brain Canada Foundation launched the Canadian Open Neuroscience Platform (CONP) bringing together many of the country's leading scientists in basic and clinical neuroscience from McGill University, Laval University and the Ontario Brain Institute to form a critical mass for Open Science. The goal of the CONP is to lead Canadian neuroscience into a new era of commonly shared, digitally integrated, data and algorithmic-rich research. To accomplish these tasks, the CONP has decided to invest in several components of neuroinformatics: (i) the technical infrastructure that facilitates the storage and processing of data via HPC or cloud ecosystems (CBRAIN [Sherif et al., 2014], Virtual Imaging Platform [VIP, Glatard et al., 2013], Compute-Canada [https://computecanada.ca] and AWS) for a total of 100,000 CPU cores; (ii) advance workflow systems necessary to extract from raw data meaningful disease biomarkers for AD or other brain pathologies; and (iii) web-based relational database (LORIS [Das et al., 2012]) to facilitate data query and retrieval (Tables 1 and 2). To date, the CONP portal (https://portal.conp.ca/) provides 90 pipelines (primarily algorithms for imaging, neuroinformatics and bioinformatics analyses) and 70 datasets (primarily neuroimaging, EEG, transcriptomics and genomics data) sorted with a variety of filters for the benefit of the end user (Duchesne et al., 2019). Detailed descriptions are provided for each pipeline and dataset. Both computational pipelines and data cohorts can be explored, searched and selected by specifying keywords. The majority of the datasets exposed follow BIDS standards. CONP analysis tools can be executed via the Boutiques system (Glatard et al., 2018) to be run locally or on HPC/cloud platforms through ad hoc containers (Docker and Singularity). Boutiques are innovative and interoperable framework, supported by multiple virtual platforms, that enables the creation of neuroinformatics applications via structured descriptors. The backend architecture adopted by CONP is shown in Figure 1b.

1.5.3 | GAAIN

In 2013, the Alzheimer's Association established the Global Alzheimer's Association Interactive Network (GAAIN) initiative (https://gaain.org), a federated platform that makes more than 500,000 patient records with more than 35,000 clinical, neuropsychological, genetic and imaging data (Tables 1 and 2) from 60 partners around the world accessible in aggregated form. GAAIN allows all neuroscientists interested in studying AD to query and analyse the data using filtering criteria.

GAAIN is suitable for conducting meta-analyses (Neu et al., 2017). From the GAAIN dashboard (Ashish et al., 2016; Toga et al., 2016), users can run various models commonly used in epidemiology, such as (i) linear regression: to measure how well two variables are correlated each other; (ii) logistic regression: to model the role variables may play between a control and case group; (iii) Cox regression: to model how variables may affect subject survival times; and (iv) Mantel–Haenzel meta-analysis: to quantify a pooled odds ratio by combining independent case–control studies together. The GAAIN architecture consists of a central server at USC and a client server running at each partner's site. To cope with the different ontologies of the datasets, GAAIN has developed a semi-automatic mapping tool called Entity Mapping capable of mapping and encode different naming conventions (Ashish et al., 2016). Data partners have full control over the shared data as they can stop communication at any time, and the shared data are immediately deleted from the cache memory of the GAAIN Central server (Neu et al., 2016). The backend architecture adopted by GAAIN is shown in Figure 1c.

1.5.4 | BRAIN research initiative(s)

In 2013, U.S. President Barack Obama's 15-year plan was funded by the NIH to accelerate the development of innovative neurotechnologies aimed at understanding how the healthy brain works and how neural circuits function. The 'Brain Research through Advancing Innovative Neurotechnologies' programme (BRAIN - https://braininitiative.nih.gov) also aims to find new ways to predict, treat, cure and prevent brain disorders. To achieve these goals, BRAIN's strategy was first to empower the scientists with data standards and then to develop the necessary archives and e-infrastructures. The BRAIN

neuroinformatics programme decided that, instead of an all-encompassing e-infrastructure, the project could create a few cloud infrastructures for individual scientific areas. Specifically, the programme consists of three main components: (i) data archives, such as BossDB, DANDI, BIL and NeMO tools that allow users to access data without moving petabytes of information to local computers; (ii) data standards, such as NWB for electrophysiology data and DICOM for imaging data; and (iii) software tools for data integration, interoperability and analysis (Hsu et al., 2020).

Following the establishment of the HBP-EBRAINS and US BRAIN initiatives, additional five similar BRAIN initiatives emerged in China (China Brain Project), Japan (Japan Brain-MINDS project), Korea (Korea Brain Initiative), Canada (Canadian Brain Research strategy) and Australia (Australian Brain Alliance) culminating in the establishment of the International Brain Initiative (IBI) in 2017. The IBI is neither a funding body nor a research project, but has a coordinating role in promoting interoperability among the seven 'BRAINS initiatives' worldwide (Quaglio et al., 2021).

1.6 | Standard operative procedure and best practice for the neuroinformatics

Neuroscience has entered a 'golden age' with many large-scale initiatives attempting to understand brain function and dysfunction by combining theoretical, computational and technological approaches.

Although new data in neuroscience can be obtained rather easily, there is one major problem that neuroinformatics is trying to solve: the poor reproducibility of results (Ioannidis et al., 2014). This is particularly true for neuroimaging data which, even using the best practice for standardisation, show a remarkable variability because of the samples, the scanners and the analysis pipelines (Schilling et al., 2021). Nature surveyed 1500 scientists and 90% agreed that there was a crisis; 70% said they had tried to replicate another group's experiment but failed (Baker, 2016). Indeed, this represents a major lack of efficiency and is particularly striking in this day when results are constantly published under more stringent requirements imposed by scientific journals in the interest of rigour and transparency.

To overcome this 'impasse', significant efforts have been made to standardise and establish best practices. In 2005, INCF was established as an independent international organisation to promote the standardisation of data and e-infrastructures in neuroscience.

In 2016, the INCF played a prominent role in defining and disseminating the FAIR principles (Poline et al., 2022). These best practices reinforced many other historical de facto standards established in neuroimaging practice, such as DICOM or NIfTI, and in clinical practice with the implementation of the HL7-FHIR standards.

Currently, the INCF has endorsed eight standards and best practices to support neuroscience (Abrams et al., 2021). A notable example is the BIDS for imaging data. Since then, a significant community of researchers has begun to take interest in the BIDS standard, leading to subsequent expansion to include the organisation of EEG and magnetoencephalography data. Moreover, to increase the momentum around the FAIR open science, other institutes such as the IBI (Eke et al., 2022) and the Institute of Electrical and Electronics Engineers (IEEE) started to work closely with INCF.

The INCF has developed a robust governance process to discuss, support and endorse new standards. Table 3 provides a complete list of current standards developed for the neuroscience community. Clearly, more standards will need to be developed in the coming years to achieve syntactic interoperability, which allows two or more systems to communicate and exchange data even if they use different programming languages, and semantic interoperability, where data are preserved and fully understood by each system, between neuroinformatics platforms. The development and adoption of such standards with the aim of improving the replicability of results through data interoperability is an advantage for robust open neuroscience.

1.7 | The next 10 years of neuroinformatics platforms

The neuroimaging-focussed neuroinformatics platforms presented in this review might undergo substantial reshaping in the near future. Nevertheless, it is important for neuroscientists to know what is available today and how these e-infrastructures might evolve in the future.

Neuroinformatics platforms will continue to promote frameworks capable of integrating multimodal and multi-scale data holistically to enable the use of analytics such as ML, AI and federated learning approaches (TensorFlow and PyTorch). In addition, these platforms will increasingly seek to (i) leverage innovative technologies for data connectivity and exchange (MapReduce programs developed by Google and Hadoop Java technology); (ii) consolidate Single Sign-On technology to ensure user friendly access to data and networked services; (iii) improve platforms connectivity by a multiple of 100 GB per second to pave the way for efficient terabit connectivity; and (iv) improve real-time collection and processing of biosensor technologies.

TABLE 3 INCF standards and best practice

Standard/best practices	Description	Developed by	Status
FAIR	A set of best practices to help researchers to develop software of excellent quality, reproducible, and widely reusable by other researchers	INCF	Adopted
BIDS	BIDS is a standard that prescribes a formal method for naming and organising MRI data and metadata. BIDS speeds up the curation process for databases (e.g. neuGRID, LORIS and XNAT) and the data export procedures	INCF	Adopted
PyNN	PyNN is a simulator-independent language for building neuronal network models. The PyNN API aims to model neuron populations, layers, columns, and the connections between them, while providing access to the details of individual neurons and synapses as needed	INCF	Adopted
NeuroML	NeuroML is an XML-based standardised model description language for computational neuroscience that provides a common data format for defining and exchanging descriptions of neuronal cell and network models. NeuroML focusses on models based on the biophysical and anatomical properties of real neurons, including details of neuronal morphologies and membrane conductance	INCF	Adopted
MBF v4.0	MBF v4.0 provides an openly documented and widely used digital reconstruction and modeling structure for microscopic anatomies. The XML based file structure has been used by neuroscientists around the world for over 30 years. File metadata provides detailed information about the origin of the sample, ensuring that provenance of derived data can be traced and that important source information is not separated from the data	INCF	Adopted
NIX	NIX electrophysiology file format enables storage of fully annotated scientific datasets along with the metadata and their relationships in a unified comprehensive format	INCF	Adopted
NWB	NWB is a data standard for neurophysiology. NWB is designed to store a variety of neurophysiological data, intracellular or extracellular electrophysiological experiments, data from optical physiological experiments, and stimulus data	INCF	Adopted
DAQCORD	DAQCORD is a framework for designing eCRF to achieve high data quality. These best practices represent a comprehensive set of data quality indicators for large observational clinical studies	INCF	Adopted
NIDM	NIDM provides a representation of mass univariate GLM results. In neuroimaging, 'mass univariate' analyses are applied to a wide range of data: functional MRI, structural	INCF	Under development

(Continues)

TABLE 3 (Continued)

Standard/best practices	Description	Developed by	Status
	MRI, PET, EEG and MEG data. The mass univariate model is fitted to each voxel independently, in parallel, and then unified across the different analysis software packages. The implementation of NIDM within FSL and SPM, two of the major neuroimaging libraries, provides an automated solution for sharing maps generated by different studies		
BMI/BCI	BMI/BCI develop and provide standards for systems that enable closed-loop interaction with artificial devices (e.g. miniaturised EEG) based on information obtained from measurements of nervous system activity	IEEE	Under development
IDG tool	IDG aims to define a standard that addresses the different international constraints on data sharing (i.e. ethical principles, national and international laws, regulations and policies) while taking into account privacy requirements and the need for open science in neuroscience	IBI	Under development

Note: Table 3 lists the standard and best practices available to the neuroscience community at the time of writing this manuscript. Acronyms: API: Application Programming Interface; BCI: Brain Computer Interface; BIDS: Brain Imaging Data Structure; BMI: Brain-Machine Interfacing; DAQCORD: Data Acquisition, Quality and Curation for Observational Research Designs; eCRF: electronic Case Report Form; FAIR: Findable, Accessible, Interoperable and Reusable; FSL: FMRIB Software Library; GLM: General Linear Model; IBI: International Brain Initiative; IDG: International Data Governance; IEEE: Institute of Electrical and Electronics Engineers; INCF: International Neuroinformatics Coordinating Facilities; MBF: Neuromorphological File Format; NIDM: NeuroImaging Data Model; NIX: Neuroscience Information eXchange; NWB: Neurodata Without Borders; pyNN: Python Neuron Network models; SPM: Statistical Parametric Mapping; XML: Extensible Markup Language.

Future platform interoperability should also rely on defining and executing pipelines through simple interfaces based on natural language understanding, removing any implementation details and difficulties. Interoperability will be facilitated by the introduction of Web 4.0 technologies (Almeida, 2017) (definition of intelligent software agents and new model of machine-to-machine communication) and applications that facilitate information exchange and collaboration between different infrastructures. In the future, neuroscientists will no longer face a steep learning curve to understand computing ecosystems (scripting, programming interfaces, data transfer protocols and shells for remote data retrieval).

Implementation of the aforementioned improvements, along with rigorous clinical validation of the tools provided by neuroinformatics platforms, might envision a futuristic use case that a physician will face in the next 10 years, likely the following: 'I need to quantify MRI biomarkers for a group of subjects within the next 2 hours.' This analysis, triggered by a voice message through the physician's smartphone, will be performed in a neuroinformatics platform that is automatically configured *ex-novo* in a cloud environment within minutes. Thanks to intelligent software agents, which are able to access and anonymise the hospital raw data and homomorphically encrypt the personal data using blockchain techniques to maximise security, the requested biomarkers will be calculated. The AI models will also provide the physician with an estimate of the biomarkers' evolutions based on the patient's genetic characteristics, comorbidities and lifestyle habits. Finally, the neuroinformatics agents will display the fresh results on the physician's tablet prior to his/her clinical appointments.

2 | CONCLUSIONS

Neuroinformatics has attracted much attention since its inception and has fostered the development of several platforms for integrating, analysing, and sharing data and theories across scales and neuroscience sectors/disciplines, favouring standardisation. Despite the progress made in the last decade, efforts towards deeper and more significant harmonisation are needed to enable the maturation of current platforms towards cutting-edge scientific e-infrastructures.

Indeed, opportunities for interoperability are being developed. The Italian MoH and MEF have pursued this

vision with a multidimensional 'matrix' approach. In addition, the Italian 'National Recovery and Resilience Plan' (PNRR), supported by the NextGenerationEU programme, is expected to promote large-scale investments in the digitalisation of health technology to further advance the interoperability of biomedical research platforms.

In Europe, through the 'European Strategy Forum on Research Infrastructures' (ESFRI) (https://www.esfri.eu), opportunities to promote infrastructure interoperability are being developed in member states and associated countries over the last three framework programmes. ESFRI will continue to play a strategic role in supporting European e-infrastructures at national and global levels in the coming years. For example, the ESFRI roadmap has allowed the launch of the EBRAINS-Italy project based on PNRR funding.

At the international level, the Coordination and Support Actions (CSA) for Research Infrastructures will enable the harmonisation of available health infrastructures from 2021 to 2027. In particular, this call has the following objectives: (i) consolidate and expand the international dimension of research infrastructures; (ii) promote scalability, security and interoperability of platforms; and (iii) support open science.

All these concrete efforts represent the driving force for the near future to create a global and interoperable neuroinformatics platform for neurology without inefficiencies or overlaps. Neuroinformatics represents the bridge for transnational collaborations, and the emergence of common standards such as BIDS and best practices such as FAIR culture will be the key to global interoperability.

ACKNOWLEDGEMENTS

The present work was partially funded by

1- the National funding of Italian Ministry of Health in the framework of the grant RETE IRCCS DI NEUROSCIENZE E NEURORIABILITAZIONE (RRC-2016-2361095; RRC-2017-2364915; RRC-2018-2365796; RRC-2019-23669119_001) and in the framework of the grant ISTITUTI NAZIONALI VIRTUALI (RCR 2020–23670067 and RCR-2021-23671214);
2- the Ministry of Economy and Finance (CCR-2017-23669078 and CCR-2018-23669122);
3- the Italian Ministry of Health in the framework - Ricerca corrente 2019 ERANET - JPND JPco-fuND 2 – 'Early Detection of Alzheimer's Disease Subtypes' E-DADS and European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 825664 – JPco-fuND 2;
4- European Union's Horizon 2020 research and innovation programme under grant agreement No. 666992; and
5- European Union's Horizon 2020 Framework Program for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3) to ED and from Centro Fermi project 'Local Neuronal Microcircuits' to ED. Special acknowledgement to EBRAINS and FENIX for informatic support and infrastructure.
6- Open access funding provided by BIBLIOSAN.

CONFLICTS OF INTEREST

The authors declared that they have no competing interests.

AUTHOR CONTRIBUTIONS

AR: conceptualisation, methodology, resources, formal analysis, data curation, writing, supervision.
DA: software, data curation, writing.
SDF: software, data curation, writing.
CC: software, writing.
FT: project administration, review – editing.
RL: project administration, review – editing.
JCL: review – editing.
RG: review – editing.
CAMGWH: review – editing.
DCA: project administration, review – editing.
EDA: review – editing.

All authors contributed to the article and approved the submitted version.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15854.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in

- RIN (https://www.reteneuroscienze.it/en/istitutinazionali-virtuali/);
- IVN (https://www.reteneuroscienze.it/en);
- HBD (https://www.alleanzacontroilcancro.it/en/progetti/health-big-data/);
- neuGRID (https://neugrid2.eu);
- EuroPOND (http://europond.eu/software/);
- E-DADS (https://e-dads.github.io/);
- HBP-EBRAINS (https://www.ebrains.eu);
- LONI (https://ida.loni.usc.edu/login.jsp);
- CONP (https://conp.ca/);

- US BRAIN (https://braininitiative.nih.gov/);
- GAAIN (https://gaain.org/)

ORCID

Alberto Redolfi https://orcid.org/0000-0002-4145-9059
Damiano Archetti https://orcid.org/0000-0003-0818-7685
Silvia De Francesco https://orcid.org/0000-0002-9277-9087

REFERENCES

Abdulrahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C., & Guizani, M. A. (2021). Survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet of Things Journal*, *8*(7), 5476–5497. https://doi.org/10.1109/JIOT.2020.3030072

Abi Nader, C., Ayache, N., Robert, P., & Lorenzi, M. (2020). Monotonic Gaussian process for spatio-temporal disease progression modeling in brain imaging data. *NeuroImage*, *205*, 116266. https://doi.org/10.1016/j.neuroimage.2019.116266

Abrams, M. B., Bjaalie, J. G., Das, S., Egan, G. F., Ghosh, S. S., Goscinski, W. J., Grethe, J. S., Hellgren Kotaleski, J., Tatt Wei Ho, E., Kennedy, D. N., Lanyon, L. J., Leergaard, T. B., Mayberg, H. S., Milanesi, L., Mouček, R., Poline, J. B., Roy, P. K., Strother, S. C., Boon Tang, T., ... Martone, M. E. (2021). A standards organization for open and FAIR neuroscience: The international neuroinformatics coordinating facility. *Neuroinformatics*, *1*, 12. https://doi.org/10.1007/s12021-020-09509-0

Aisen, P. S., Petersen, R. C., Donohue, M. C., Gamst, A., Raman, R., Thomas, R. G., Walter, S., Trojanowski, J. Q., Shaw, L. M., Beckett, L. A., Jack, C. R., Jagust, W., Toga, W. W., Saykin, A. J., Morris, J. C., Green, R. C., & Weiner, M. W. (2010). Clinical core of the Alzheimer's disease neuroimaging initiative: Progress and plans. *Alzheimer's & Dementia*, *6*(3), 239–246. https://doi.org/10.1016/j.jalz.2010.03.006

Aisen, P. S., Petersen, R. C., Donohue, M., & Weiner, M. W. (2015). Alzheimer's disease neuroimaging initiative. Alzheimer's disease neuroimaging initiative 2 clinical core: Progress and plans. *Alzheimer's & Dementia*, *11*(7), 734–739. https://doi.org/10.1016/j.jalz.2015.05.005

Almeida, F. (2017). Concept and dimensions of web 4.0. *International Journal of Computers and Technology*, *16*(7), 7040–7046. https://doi.org/10.24297/ijct.v16i7.6446

American College of Radiology. (2018). *Phantom test guidance for use of the small MRI phantom for the ACR MRI accreditation program*. American College of Radiology.

Amunts, K., Axer, M., Bitsch, L., Bjaalie, J., Brovelli, A., Caspers, S., Costantini, I., D'Angelo, E., De Bonis, G., DeFelipe, J., Destexhe, A., Dickscheid, T., Diesmann, M., Eickhoff, S. B., Engel, A., Fousek, J., Furber, S., Goebel, R., Günterkün, O., ... Vanduffel, W. (2022). The coming decade of digital brain research—A vision for neuroscience at the intersection of technology and computing. *Zenodo*. https://doi.org/10.5281/zenodo.6345821

Amunts, K., DeFelipe, J., Pennartz, C., Destexhe, A., Migliore, M., Ryvlin, P., Furber, S., Knoll, A., Bitsch, L., Bjaalie, J.G., Ioannidis, Y., Lippert, T., Sanchez-Vives, M.V., Goebel, R., & Jirsa, V. (2022). Linking brain structure, activity, and cognitive function through computation, *eNeuro*, *9*(2), ENEURO.0316-21.2022. https://doi.org/10.1523/ENEURO.0316-21.2022

Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., & Lippert, T. (2016). The human brain project: Creating a European research infrastructure to decode the human brain. *Neuron*, *92*(3), 574–581. https://doi.org/10.1016/j.neuron.2016.10.046

Amunts, K., Knoll, A. C., Lippert, T., Pennartz, C. M. A., Ryvlin, P., Destexhe, A., Jirsa, V. K., D'Angelo, E., & Bjaalie, J. G. (2019). The human brain project—Synergy between neuroscience, computing, informatics, and brain-inspired technologies. *PLoS Biology*, *17*, e3000344. https://doi.org/10.1371/journal.pbio.3000344

Archetti, D., Ingala, S., Venkatraghavan, V., Wottschel, V., Young, A. L., Bellio, M., Bron, E. E., Klein, S., Barkhof, F., Alexander, D. C., Oxtoby, N. P., Grisoni, G. B., & Redolfi, A. (2019). Multi-study validation of data-driven disease progression models to characterize evolution of biomarkers in Alzheimer's disease. *NeuroImage: Clinical*, *24*, 101954. https://doi.org/10.1016/j.nicl.2019.101954

Archetti, D., Young, A. L., Oxtoby, N. P., Ferreira, D., Mårtensson, G., Westman, E., Alexander, D. C., Frisoni, G. B., & Redolfi, A. (2021). Inter-cohort validation of SuStaIn model for Alzheimer's disease. *Frontiers in Big Data*, *4*, 661110. https://doi.org/10.3389/fdata.2021.661110

Ashish, N., Bhatt, P., & Toga, A. W. (2016). Global data sharing in Alzheimer disease research. *Alzheimer's Disease & Associated Disorders*, *30*(2), 160–168. https://doi.org/10.1097/WAD.0000000000000121

Ashish, N., Dewan, P., & Toga, A. W. (2016). The GAAIN entity mapper: An active-learning system for medical data mapping. *Frontiers in Neuroinformatics*, *9*, 30. https://doi.org/10.3389/fninf.2015.00030

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Braga, F., Pasqualetti, S., & Panteghini, M. (2018). The role of external quality assessment in the verification of in vitro medical diagnostics in the traceability era. *Clinical Biochemistry*, *57*, 23–28. https://doi.org/10.1016/j.clinbiochem.2018.02.004

Calhoun, V. D. (2015). A spectrum of sharing: Maximization of information content for brain imaging data. *Gigascience*, *4*(1), s13742–014–0042–5. https://doi.org/10.1186/s13742-014-0042-5

Campos, E. M., Saura, P. F., González-Vidal, A., Ramos, J. L., Bernabé, J. B., Baldini, G., & Gómez-Skarmeta, A. F. (2022). Evaluating federated learning for intrusion detection in internet of things: Review and challenges. *Computer Networks*, *203*, 108661. https://doi.org/10.1016/j.comnet.2021.108661

Chen, C. C., Wan, Y. L., Wai, Y. Y., & Liu, H. L. (2004). Quality assurance of clinical MRI scanners using ACR MRI phantom: Preliminary results. *Journal of Digital Imaging*, *17*, 279–284. https://doi.org/10.1007/s10278-004-1023-5

Clark, D., McLaughlin, P., Woo, E., Hwang, K., Hurtz, S., Ramirez, L., Eastman, J., Dukes, R., Kapur, P., DeRamus, T., & Apostolova, L. (2016). Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. *Alzheimer's and*

Dementia: Diagnosis, Assessment and Disease Monitoring, *2*(1), 113–122. https://doi.org/10.1016/j.dadm.2016.02.001

Crawford, K. L., Neu, S. C., & Toga, A. W. (2016). The image and data archive at the Laboratory of Neuro Imaging. *NeuroImage*, *124*(Part B), 1080–1083. https://doi.org/10.1016/j.neuroimage.2015.04.067

D'Angelo, E., & Jirsa, V. (2022). The quest for multiscale brain modelling. *Trends in Neuroscience*, *45*(10), 777–790. https://doi.org/10.1016/j.tins.2022.06.007

Darwish, A. (2018). Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications. *Future Computing and Informatics Journal*, *3*(2), 231–246. https://doi.org/10.1016/j.fcij.2018.06.001

Das, S., Glatard, T., Rogers, C., Saigle, J., Paiva, S., MacIntyre, L., Safi-Harab, M., Rousseau, M. E., Stirling, J., Khalili-Mahani, N., MacFarlane, D., Kostopoulos, P., Rioux, P., Madjar, C., Lecours-Boucher, X., Vanamala, S., Adalat, R., Mohaddes, Z., Fonov, V. S., … Evans, A. C. (2017). Cyberinfrastructure for Open Science at the Montreal neurological institute. *Frontiers in Neuroinformatics*, *10*, 53. https://doi.org/10.3389/fninf.2016.00053

Das, S., Zijdenbos, A. P., Harlap, J., Vins, D., & Evans, A. C. (2012). LORIS: A web-based data management system for multi-center studies. *Frontiers in Neuroinformatics*, *5*, 37. https://doi.org/10.3389/fninf.2011.00037

De Francesco, S., Galluzzi, S., Vanacore, N., Festari, C., Rossini, P. M., Cappa, S. F., Frisoni, G. B., & Redolfi, A. (2021). Norms for automatic estimation of hippocampal atrophy and a step forward for applicability to the Italian population. *Frontiers in Neuroscience*, *15*, 656808. https://doi.org/10.3389/fnins.2021.656808

Dekker, I., Schoonheim, M. M., Venkatraghavan, V., Eijlers, A. J. C., Brouwer, I., Bron, E. E., Klein, S., Wattjes, M. P., Wink, A. M., Geurts, J. J. G., Uitdehaag, B. M. J., Oxtoby, N. P., Alexander, D. C., Vrenken, H., Killestein, J., Barkhof, F., & Wottschel, V. (2021). The sequence of structural, functional and cognitive changes in multiple sclerosis. *NeuroImage: Clinical*, *29*, 102550. https://doi.org/10.1016/j.nicl.2020.102550

Dinov, I. D., Van Horn, J. D., Lozev, K. M., Magsipoc, R., Petrosyan, P., Liu, Z., Mackenzie-Graham, A., Eggert, P., Parker, D. S., & Toga, A. W. (2009). Efficient, distributed and interactive neuroimaging data analysis using the LONI pipeline. *Frontiers in Neuroinformatics*, *3*, 22. https://doi.org/10.3389/neuro.11.022.2009

Duchesne, S., Chouinard, I., Potvin, O., Fonov, V. S., Khademi, A., Bartha, R., Bellec, P., Collins, D. L., Descoteaux, M., Hoge, R., McCreary, C. R., Ramirez, J., Scott, C. J. M., Smith, E. E., Strother, S. C., & Black, S. E. (2019). The Canadian dementia imaging protocol: Harmonizing national cohorts. *Journal of Magnetic Resonance Imaging*, *49*(2), 456–465. https://doi.org/10.1002/jmri.26197

Eke, D. O., Bernard, A., Bjaalie, J. G., Chavarriaga, R., Hanakawa, T., Hannan, A. J., Hill, S. L., Martone, M. E., McMahon, A., Ruebel, O., Crook, S., Thiels, E., & Pestilli, F. (2022). International data governance for neuroscience. *Neuron*, *110*(4), 600–612. https://doi.org/10.1016/j.neuron.2021.11.017

Eshaghi, A., Marinescu, R. V., Young, A. L., Firth, N. C., Prados, F., Cardoso, M. J., Tur, F., De Angelis, F., Cawley, N., Brownlee, W. J., De Stefano, N., Stromillo, M. L., Battaglini, M., Ruggieri, S., Gasperini, C., Filippi, M., Rocca, M. A., Rovira, A., Sastre-Garriga, J., … Ciccarelli, O. (2018). Progression of regional grey matter atrophy in multiple sclerosis. *Brain*, *141*(6), 1665–1677. https://doi.org/10.1093/brain/awy088

Eshaghi, A., Young, A. L., Wijeratne, P. A., Prados, F., Arnold, D. L., Narayanan, S., Guttman, C. R. G., Barkhof, F., Alexander, D. C., Thompson, A. J., Chard, D., & Ciccarelli, O. (2021). Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nature Communications*, *12*, 2078. https://doi.org/10.1038/s41467-021-22265-2

Fonteijn, H. M., Modat, M., Clarkson, M. J., Barnes, J., Lehmann, M., Hobbs, N. Z., Scahill, R. I., Tabrizi, S. J., Ourselin, S., Fox, N. C., & Alexander, D. C. (2012). An event-based model for disease progression in Alzheimer's disease and Huntington's disease. *NeuroImage*, *60*(3), 1880–1889. https://doi.org/10.1016/j.neuroimage.2012.01.062

Gabel, M. C., Broad, R. J., Young, A. L., Abrahams, S., Bastin, M. E., Menke, R. A. L., Al-Chalabi, A., Goldstein, L. H., Tsermentseli, S., Alexander, D. C., Turner, M. R., Leigh, P. N., & Cercignani, M. (2020). Evolution of white matter damage in amyotrophic lateral sclerosis. *Annals of Clinical and Translational Neurology*, *7*(5), 722–732. https://doi.org/10.1002/acn3.51035

Glatard, T., Kiar, G., Aumentado-Armstrong, T., Beck, N., Bellec, P., Bernard, R., Bonnet, A., Brown, S. T., Camarasu-Pop, S., Cervenansky, F., Das, S., Ferreira da Silva, R., Flandin, G., Girard, P., Gorgolewski, K. J., Guttmann, C. R. G., Hayot-Sasson, V., Quirion, P. O., Rioux, P., … Evans, A. C. (2018). Boutiques: A flexible framework to integrate command-line applications in computing platforms. *Gigascience*, *7*(5), giy016. https://doi.org/10.1093/gigascience/giy016

Glatard, T., Lartizien, C., Gibaud, B., da Silva, R. F., Forestier, G., Cervenansky, F., Alessandrini, M., Benoit-Cattin, H., Bernard, O., Camarasu-Pop, S., Cerezo, N., Clarysse, P., Gaignard, A., Hugonnard, P., Liebgott, H., Marache, S., Marion, A., Montagnat, J., Tabary, J., & Friboulet, D. (2013). A virtual imaging platform for multi-modality medical image simulation. *IEEE Transactions on Medical Imaging*, *32*(1), 110–118. https://doi.org/10.1109/TMI.2012.2220154

Göltz, J., Kriener, L., Baumbach, A., Billaudelle, S., Breitwieser, O., Cramer, B., Dold, D., Kungl, A. F., Senn, W., Schemmel, J., Meier, K., & Petrovici, M. A. (2021). Fast and energy-efficient neuromorphic deep learning with first-spike times. *Nature Machine Intelligence*, *3*, 823–835. https://doi.org/10.1038/s42256-021-00388-x

Gorgolewski, K., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*, 160044. https://doi.org/10.1038/sdata.2016.44

Gui, L., Loukas, S., Lazeyras, F., Hüppi, P. S., Meskaldji, D. E., & Borradori Tolsa, C. (2019). Longitudinal study of

neonatal brain tissue volumes in preterm infants and their ability to predict neurodevelopmental outcome. *NeuroImage*, *185*, 728–741. https://doi.org/10.1016/j.neuroimage.2018.06.034

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Heo, J., Yoon, J. G., Park, H., Kim, Y. D., Nam, H. S., & Heo, J. H. (2019). Machine learning-based model for prediction of outcomes in acute stroke. *Stroke*, *50*(5), 1263–1265. https://doi.org/10.1161/STROKEAHA.118.024293

Hsu, N. S., Fang, H. Y., David, K. K., Gnadt, J. W., Peng, G. C., Talley, E. M., Ward, J. M., Ngai, J., & Koroshetz, W. J. (2020). The promise of the BRAIN initiative: NIH strategies for understanding neural circuit function. *Current Opinion in Neurobiology*, *65*, 162–166. https://doi.org/10.1016/j.conb.2020.10.008

Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., Schulz, K. F., & Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, *383*(9912), 166–175. https://doi.org/10.1016/S0140-6736(13)62227-8

Jack, C. R., Bernstein, M. A., Borowski, B. J., Gunter, J. L., Fox, N. C., Thompson, P. M., Schuff, N., Krueger, G., Killiany, R. J., Decarli, C. S., Dale, A. M., Carmichael, O. W., Tosun, D., & Weiner, M. W. (2010). Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. *Alzheimer's & Dementia*, *6*(3), 212–220. https://doi.org/10.1016/j.jalz.2010.03.004

Koval, I., Bône, A., Louis, M., Lartigue, T., Bottani, S., Marcoux, A., Samper-González, J., Burgos, N., Charlier, B., Bertrand, A., Epelbaum, S., Colliot, O., Allassonnière, S., & Durrleman, S. (2021). AD course map charts Alzheimer's disease progression. *Scientific Reports*, *11*(1), 8020. https://doi.org/10.1038/s41598-021-87434-1

Koval, I., Schiratti, J. B., Routier, A., Bacci, M., Colliot, O., Allassonniere, S., & Durrleman, S. (2018). Spatiotemporal propagation of the cortical atrophy: Population and individual patterns. *Frontiers in Neurology*, *9*, 235. https://doi.org/10.3389/fneur.2018.00235

Kruse, C. S., & Beane, A. (2018). Health information technology continues to show positive effect on medical outcomes: Systematic review. *Journal of Medical Internet Research*, *20*(2), e41. https://doi.org/10.2196/jmir.8793

LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, *2010*, 253–256. https://doi.org/10.1109/ISCAS.2010.5537907

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, *36*(4), 1234–1240. https://doi.org/10.1093/bioinformatics/btz682

Lorenzi, M., Filippone, M., Frisoni, G. B., Alexander, D. C., & Ourselin, S. (2019). Probabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in Alzheimer's disease. *NeuroImage*, *190*, 56–68. https://doi.org/10.1016/j.neuroimage.2017.08.059

Manenti, R., Gobbi, E., Baglio, F., Macis, A., Ferrari, C., Pagnoni, I., Rossetto, F., Di Tella, S., Alemanno, F., Cimino, V., Binetti, G., Iannaccone, S., Bramanti, P., Cappa, S. F., & Cotelli, M. (2020). Effectiveness of an innovative cognitive treatment and telerehabilitation on subjects with mild cognitive impairment: A multicenter, randomized, active-controlled study. *Frontiers in Aging Neuroscience*, *12*, 585988. https://doi.org/10.3389/fnagi.2020.585988

Marcus, D. S., Olsen, T. R., Ramaratnam, M., & Buckner, R. L. (2007). The extensible neuroimaging archive toolkit: An informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics*, *5*(1), 11–34. https://doi.org/10.1385/ni:5:1:11

Martone, M. E. (2022). A decade of GigaScience: The importance of community organizations for open and FAIR efforts in neuroinformatics. *Gigascience*, *11*, giac060. https://doi.org/10.1093/gigascience/giac060

Matzke, H., Schirner, M., Vollbrecht, D., Rothmeier, S., Llarena, A., Rojas, R., Triebkorn, P., Domide, L., Mersmann, J., Solodkin, A., Jirsa, V. K., McIntosh, A. R., & Ritter, P. (2015). TVB-EduPack-an interactive learning and scripting platform for the virtual brain. *Frontiers in Neuroinformatics*, *9*, 27. https://doi.org/10.3389/fninf.2015.00027

McMahan, B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, in Proceedings of Machine Learning Research*, *54*, 1273–1282. Available from https://proceedings.mlr.press/v54/mcmahan17a.html

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., Trojanowski, J. Q., Toga, A. W., & Beckett, L. (2005). The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America*, *15*(4), 869–877. https://doi.org/10.1016/j.nic.2005.09.008

Nayak, L., Dasgupta, A., Das, R., Ghosh, K., & De, R. K. (2018). Computational neuroscience and neuroinformatics: Recent progress and resources. *Journal of Biosciences*, *43*(5), 1037–1054. https://doi.org/10.1007/s12038-018-9813-y

Nelson, E. K., Piehler, B., Eckels, J., Rauch, A., Bellew, M., Hussey, P., Ramsay, S., Nathe, C., Lum, K., Krouse, K., Stearns, D., Connolly, B., Skillman, T., & Igra, M. (2011). LabKey server: An open source platform for scientific data integration, analysis and collaboration. *BMC Bioinformatics*, *12*, 71. https://doi.org/10.1186/1471-2105-12-71

Neu, S. C., Crawford, K. L., & Toga, A. W. (2016). Sharing data in the global alzheimer's association interactive network. *NeuroImage*, *124*(Part B), 1168–1174. https://doi.org/10.1016/j.neuroimage.2015.05.082

Neu, S. C., Pa, J., Kukull, W., Beekly, D., Kuzma, A., Gangadharan, P., Wang, L. S., Romero, K., Arneric, S. P., Redolfi, A., Orlandi, D., Frisoni, G. B., Au, R., Devine, S., Auerbach, S., Espinosa, A., Boada, M., Ruiz, A., Johnson, S. C., … Toga, A. W. (2017). Apolipoprotein E genotype and sex risk factors for Alzheimer disease: A meta-analysis. *JAMA Neurology*, *74*(10), 1178–1189. https://doi.org/10.1001/jamaneurol.2017.2188

Nigri, A., Ferraro, S., Gandini Wheeler-Kingshott, C. A. M., Tosetti, M., Redolfi, A., Forloni, G., D'Angelo, E., Aquino, D., Biagi, L., Bosco, P., Carne, I., De Francesco, S., Demichelis, G., Gianeri, R., Lagana, M. M., Micotti, E., Napolitano, A., Palesi, F., Pirastru, A., … Bruzzone, M. G. (2022). Quantitative MRI harmonization to maximize clinical impact: The RIN-neuroimaging network. *Frontiers in Neurology*, *13*, 855125. https://doi.org/10.3389/fneur.2022.855125

Olivieri, I., Chiappedi, M., Meriggi, P., Mazzola, M., Grandi, A., & Angelini, L. (2013). Rehabilitation of children with hemiparesis: A pilot study on the use of virtual reality. *BioMed Research International*, *2013*, 695935. https://doi.org/10.1155/2013/695935

Oxtoby, N. P., & Alexander, D. C. (2017). Imaging Plus X: Multimodal models of neurodegenerative disease. *Current Opinion in Neurology*, *30*(4), 371–379. https://doi.org/10.1097/WCO.0000000000000460

Oxtoby, N. P., Garbarino, S., Firth, N. C., Warren, J. D., Schott, J. M., & Alexander, D. C. (2017). Data-driven sequence of changes to anatomical brain connectivity in sporadic Alzheimer's disease. *Frontiers in Neurology*, *8*, 580. https://doi.org/10.3389/fneur.2017.00580

Oxtoby, N. P., Leyland, L. A., Aksman, L. M., Thomas, G. E. C., Bunting, E. L., Wijeratne, P. A., Young, A. L., Zarkali, A., Tan, M. M. X., Bremner, F. D., Keane, P. A., Morris, H. R., Schrag, A. E., Alexander, D. C., & Weil, R. S. (2021). Sequence of clinical and neurodegeneration events in Parkinson's disease progression. *Brain*, *144*(3), 975–988. https://doi.org/10.1093/brain/awaa461

Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., Goebelbecker, H. J., Gundlach, J., Schirmbacher, P., & Dierolf, U. (2013). Making research data repositories visible: The re3data.org registry. *PLoS ONE*, *8*(11), e78080. https://doi.org/10.1371/journal.pone.0078080

Pascuzzo, R., Oxtoby, N. P., Young, A. L., Blevins, J., Castelli, G., Garbarino, S., Cohen, M. L., Schonberger, L. B., Gambetti, P., Appleby, B. S., Alexander, D. C., & Bizzi, A. (2020). Prion propagation estimated from brain diffusion MRI is subtype dependent in sporadic Creutzfeldt–Jakob disease. *Acta Neuropathologica*, *140*, 169–181. https://doi.org/10.1007/s00401-020-02168-0

Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., Jack, C. R., Jagust, W. J., Shaw, L. M., Toga, A. W., Trojanowski, J. Q., & Weiner, M. W. (2010). Alzheimer's disease neuroimaging initiative (ADNI): Clinical characterization. *Neurology*, *74*(3), 201–209. https://doi.org/10.1212/WNL.0b013e3181cb3e25

Poline, J. B., Kennedy, D. N., Sommer, F. T., Ascoli, G. A., Van Essen, D. C., Ferguson, A. R., Grethe, J. S., Hawrylycz, M. J., Thompson, P. M., Poldrack, R. A., Ghosh, S. S., Keator, D. B., Athey, T. L., Vogelstein, J. T., Mayberg, H. S., & Martone, M. E. (2022). Is neuroscience FAIR? A call for collaborative standardisation of neuroscience data. *Neuroinformatics*, *10*, 1007/s12021-021-09557-0. https://doi.org/10.1007/s12021-021-09557-0

Quaglio, G., Toia, P., Moser, E. I., Karapiperis, T., Amunts, K., Okabe, S., Poo, M. M., Rah, J. C., Koninck, Y., Ngai, J., Richards, L., & Bjaalie, J. G. (2021). The international brain initiative: Enabling collaborative science. *The Lancet Neurology*, *20*(12), 985–986. https://doi.org/10.1016/S1474-4422(21)00389-6

Redolfi, A., De Francesco, S., Palesi, F., Galluzzi, S., Muscio, C., Castellazzi, G., Tiraboschi, P., Savini, G., Nigri, A., Bottini, G., Bruzzone, M. G., Ramusino, M. C., Ferraro, S., Gandini Wheeler-Kingshott, C. A. M., Tagliavini, F., Frisoni, G. B., Ryvlin, P., Demonet, J. F., Kherif, F., … D'Angelo, E. (2020). Medical informatics platform (MIP): A pilot study across clinical Italian cohorts. *Frontiers in Neurology*, *11*, 1021. https://doi.org/10.3389/fneur.2020.01021

Redolfi, A., Manset, D., Barkhof, F., Wahlund, L. O., Glatard, T., Mangin, J. F., & Frisoni, G. B. (2015). Head-to-head comparison of two popular cortical thickness extraction algorithms: A cross-sectional and longitudinal study. *PLoS ONE*, *10*(3), e0117692. https://doi.org/10.1371/journal.pone.0117692

Redolfi, A., McClatchey, R., Anjum, A., Zijdenbos, A., Manset, D., Barkhof, F., Spenger, C., Legré, Y., Wahlund, L. O., Barattieri di San Pietro, C., & Frisoni, G. B. (2009). Grid infrastructures for computational neuroscience: The neuGRID example. *Future Neurology*, *4*(6), 703–722. https://doi.org/10.2217/fnl.09.53

Rex, D. E., Ma, J. Q., & Toga, A. W. (2003). The LONI pipeline processing environment. *NeuroImage*, *19*(3), 1033–1048. https://doi.org/10.1016/s1053-8119(03)00185-x

Rotenberg, D. J., Chang, Q., Potapova, N., Wang, A., Hon, M., Sanches, M., Bogetic, N., Frias, N., Liu, T., Behan, B., El-Badrawi, R., Strother, S. C., Evans, S. G., Mikkelsen, J., Gee, T., Dong, F., Arnott, S. R., Laing, S., Dharsee, M., … Jankowicz, D. (2018). The CAMH neuroinformatics platform: A hospital-focused brain-CODE implementation. *Frontiers in Neuroinformatics*, *12*, 77. https://doi.org/10.3389/fninf.2018.00077

Sa de Almeida, J., Meskaldji, D. E., Loukas, S., Lordier, L., Gui, L., Lazeyras, F., & Hüppi, P. S. (2021). Preterm birth leads to impaired rich-club organization and fronto-paralimbic/limbic structural connectivity in newborns. *NeuroImage*, *225*, 117440. https://doi.org/10.1016/j.neuroimage.2020.117440

Schilling, K. G., Rheault, F., Petit, L., Hansen, C. B., Nath, V., Yeh, F. C., Girard, G., Barakovic, M., Rafael-Patino, J., Yu, T., Fischi-Gomez, E., Pizzolato, M., Ocampo-Pineda, M., Schiavi, S., Canales-Rodríguez, E. J., Daducci, A., Granziera, C., Innocenti, G., Thiran, J. P., … Descoteaux, M. (2021). Tractography dissection variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset? *NeuroImage*, *243*, 118502. https://doi.org/10.1016/j.neuroimage.2021.118502

Schirner, M., Domide, L., Perdikis, D., Triebkorn, P., Stefanovski, L., Pai, R., Prodan, P., Valean, B., Palmer, J., Langford, C., Blickensdörfer, A., van der Vlag, M., Diaz-Pier, S., Peyser, A., Klijn, W., Pleiter, D., Nahm, A., Schmid, O., Woodman, M., … Ritter, P. (2022). Brain simulation as a cloud service: The virtual brain on EBRAINS. *NeuroImage*, *251*, 118973. https://doi.org/10.1016/j.neuroimage.2022.118973

Sherif, T., Rioux, P., Rousseau, M. E., Kassis, N., Beck, N., Adalat, R., Das, S., Glatard, T., & Evans, A. C. (2014). CBRAIN: A web-based, distributed computing platform for collaborative neuroimaging research. *Frontiers in Neuroinformatics*, *8*, 54. https://doi.org/10.3389/fninf.2014.00054

Ten Kate, M., Dicks, E., Visser, P. J., van der Flier, W. M., Teunissen, C. E., Barkhof, F., Scheltens, P., & Tijms, B. M. (2018). Atrophy subtypes in prodromal Alzheimer's disease are associated with cognitive decline. *Brain*, *141*(12), 3443–3456. https://doi.org/10.1093/brain/awy264

Ten Kate, M., Redolfi, A., Peira, E., Bos, I., Vos, S. J., Vandenberghe, R., Gabel, S., Schaeverbeke, J., Scheltens, P., Blin, O., Richardson, J. C., Bordet, R., Wallin, A., Eckerstrom, C., Molinuevo, J. L., Engelborghs, S., van Broeckhoven, C., Martinez-Lage, P., Popp, J., … Barkhof, F. (2018). MRI predictors of amyloid pathology: Results from the EMIF-AD multimodal biomarker discovery study. *Alzheimer's Research & Therapy*, *10*, 100. https://doi.org/10.1186/s13195-018-0428-1

Toga, A. W., Neu, S. C., Bhatt, P., Crawford, K. L., & Ashish, N. (2016). The global Alzheimer's association interactive network. *Alzheimer's & Dementia*, *12*(1), 49–54. https://doi.org/10.1016/j.jalz.2015.06.1896

Ulfenborg, B., Karlsson, A., Riveiro, M., Andersson, C. X., Sartipy, P., & Synnergren, J. (2021). Multi-assignment clustering: Machine learning from a biological perspective. *Journal of Biotechnology*, *326*, 1–10. https://doi.org/10.1016/j.jbiotec.2020.12.002

Vaccarino, A. L., Dharsee, M., Strother, S., Aldridge, D., Arnott, S. R., Behan, B., Dafnas, C., Dong, F., Edgecombe, K., El-Badrawi, R., El-Emam, K., Gee, T., Evans, S. G., Javadi, M., Jeanson, F., Lefaivre, S., Lutz, K., MacPhee, F. C., Mikkelsen, J., … Evans, K. R. (2018). Brain-CODE: A secure neuroinformatics platform for management, federation, sharing and analysis of multi-dimensional neuroscience data. *Frontiers in Neuroinformatics*, *12*, 28. https://doi.org/10.3389/fninf.2018.00028

Van Horn, J. D. (2021). Bridging the brain and data sciences. *Big Data*, *9*(3), 153–187. https://doi.org/10.1089/big.2020.0065

Venkatraghavan, V., Bron, E. E., Niessen, W. J., & Klein, S. (2019). Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. *NeuroImage*, *186*, 518–532. https://doi.org/10.1016/j.neuroimage.2018.11.024

Venkatraghavan, V., Klein, S., Fani, L., Ham, L. S., Vrooman, H., Ikram, M. K., Niessen, W. J., & Bron, E. E. (2021). Analyzing the effect of APOE on Alzheimer's disease progression using an event-based model for stratified populations. *NeuroImage*, *227*, 117646. https://doi.org/10.1016/j.neuroimage.2020.117646

Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, *88*, 428–436. https://doi.org/10.1016/j.jbusres.2017.12.043

Vinke, E. J., de Groot, M., Venkatraghavan, V., Klein, S., Niessen, W. J., Ikram, M. A., & Vernooij, M. W. (2018). Trajectories of imaging markers in brain aging: The Rotterdam study. *Neurobiology of Aging*, *71*, 32–40. https://doi.org/10.1016/j.neurobiolaging.2018.07.001

Vu, M. T., Adalı, T., Ba, D., Buzsáki, G., Carlson, D., Heller, K., Liston, C., Rudin, C., Sohal, V. S., Widge, A. S., Mayberg, H. S., Sapiro, G., & Dzirasa, K. (2018). A shared vision for machine learning in neuroscience. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(7), 1601–1607. https://doi.org/10.1523/JNEUROSCI.0508-17.2018

Weber, C. J., Carrillo, M. C., Jagust, W., Jack, C. R., Shaw, L. M., Trojanowski, J. Q., Saykin, A. J., Beckett, L. A., Sur, C., Rao, N. P., Mendez, P. C., Black, S. E., Li, K., Iwatsubo, T., Chang, C. C., Sosa, A. L., Rowe, C. C., Perrin, R. J., Morris, J. C., … Weiner, M. W. (2021). The worldwide Alzheimer's disease neuroimaging initiative: ADNI-3 updates and global perspectives. *Alzheimer's & Dementia*, *7*(1), e12226. https://doi.org/10.1002/trc2.12226

Wieder, P., & Nolte, H. (2022). Toward data lakes as central building blocks for data management and analysis. *Frontiers in Big Data*, *5*, 945720. https://doi.org/10.3389/fdata.2022.945720

Wijeratne, P. A., Johnson, E. B., Gregory, S., Georgiou-Karistianis, N., Paulsen, J. S., Scahill, R. I., Tabrizi, S. J., & Alexander, D. C. (2021). A multi-study model-based evaluation of the sequence of imaging and clinical biomarker changes in Huntington's disease. *Frontiers in Big Data*, *4*, 662200. https://doi.org/10.3389/fdata.2021.662200

Wijeratne, P. A., Young, A. L., Oxtoby, N. P., Marinescu, R. V., Firth, N. C., Johnson, E. B., Mohan, A., Sampaio, C., Scahill, R. I., Tabrizi, S. J., & Alexander, D. C. (2018). An image-based model of brain volume biomarker changes in Huntington's disease. *Annals of Clinical and Translational Neurology*, *5*(5), 570–582. https://doi.org/10.1002/acn3.558

Young, A. L., Marinescu, R. V., Oxtoby, N. P., Bocchetta, M., Yong, K., Firth, N. C., Cash, D. M., Thomas, D. L., Dick, K. M., Cardoso, J., van Swieten, J., Borroni, B., Galimberti, D., Masellis, M., Tartaglia, M. C., Rowe, J. B., Graff, C., Tagliavini, F., Frisoni, G. B., … Alexander, D. C. (2018). Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference. *Nature Communications*, *9*, 4273. https://doi.org/10.1038/s41467-018-05892-0

Young, A. L., Oxtoby, N. P., Daga, P., Cash, D. M., Fox, N. C., Ourselin, S., Schott, J. M., & Alexander, D. C. (2014). A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain*, *137*(9), 2564–2577. https://doi.org/10.1093/brain/awu176

How to cite this article: Redolfi, A., Archetti, D., De Francesco, S., Crema, C., Tagliavini, F., Lodi, R., Ghidoni, R., Gandini Wheeler-Kingshott, C. A. M., Alexander, D. C., & D'Angelo, E. (2022). Italian, European, and international neuroinformatics efforts: An overview. *European Journal of Neuroscience*, 1–23. https://doi.org/10.1111/ejn.15854