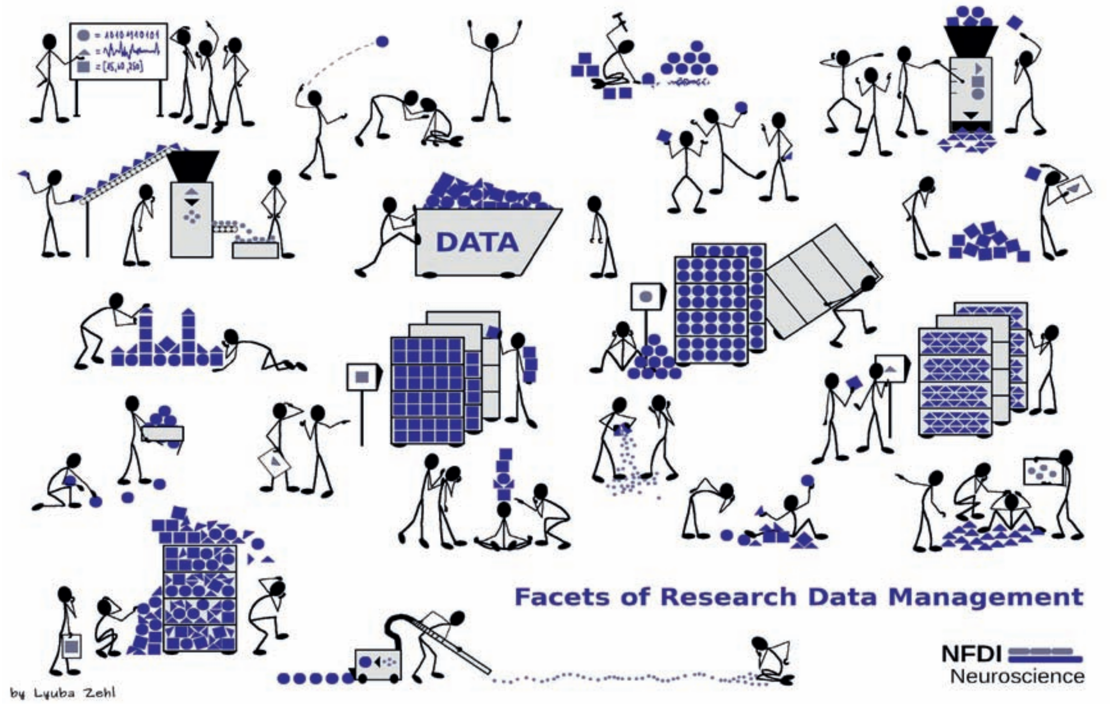
DE GRUYTER

2021 · VOLUME 27 · ISSUE 1
ISSN 0947-0875 · e-ISSN 2363-7013

NEUROFORUM

ORGAN DER NEUROWISSENSCHAFTLICHEN GESELLSCHAFT



SPECIAL ISSUE:
NFDI – NATIONALE FORSCHUNGSDATENINFRASTRUKTUR

GUEST EDITOR
Thomas Wachtler

HERAUSGEGEBEN VON
Neurowissenschaftliche Gesellschaft e.V. (NWG)

CHEFREDAKTEURIN
Petra Wahle, Bochum

NWG
NeuroWissenschaftliche Gesellschaft
German Neuroscience Society

DE G

www.degruyter.com/journals/nf

NEUROFORUM

HERAUSGEGEBEN VON
Neurowissenschaftliche Gesellschaft e.V. (NWG)

CHEFREDAKTEURIN
Petra Wahle, Bochum

REDAKTION
Susanne Hannig, Berlin

REDAKTIONSGREMIUM
Nikolai Axmacher, Bochum
Mathias Bähr, Göttingen
Niels Birbaumer, Tübingen
Tobias Böckers, Ulm
Sebastian Brandner, London, US
Katharina Braun, Magdeburg
Nils Brose, Göttingen
Ansgar Büschges, Köln
Thomas Deller, Frankfurt/M.
Ricarda Diem, Heidelberg
Veronica Egger, Regensburg
Jens Eilers, Leipzig
Eckhard Friauf, Kaiserslautern
Giovanni Galizia, Konstanz
Martin Göpfert, Göttingen
Benedikt Grothe, München
Sonja Grün, Jülich
Eckart Gundelfinger, Magdeburg
Ileana Hanganu-Opatz, Hamburg
Andreas Heinz, Berlin
Charlotte Förster, Würzburg
Michael Heneka, Bonn
Andreas Herz, München
Isabella Heuser-Collier, Berlin
Mark Hübener, Martinsried
Reinhard Jahn, Göttingen
Peter Jonas, Klosterneuburg, Österreich
Sabine Kastner, Princeton, USA
Helmut Kettenmann, Berlin
Frank Kirchhoff, Homburg
Christian Klämbt, Münster
Matthias Kneussel, Hamburg
Michael Koch, Bremen
Arthur Konnerth, München
Sigrun Korsching, Köln
Trese Leinders-Zufall, Homburg
Wolfgang Löscher, Hannover
Siegrid Löwel, Göttingen
Albert Christian Ludolph, Ulm
Heiko Luhmann, Mainz
Hanspeter A. Mallot, Tübingen
Denise Manahan-Vaughan, Bochum
Thomas Möller, Cambridge, USA
Ulrike Müller, Heidelberg
Hans-Joachim Pflüger, Berlin
Heidrun Potschka, München
Josef Rauschecker, Washington, USA
Angelika Richter, Leipzig
Andreas Ritzau-Jost, Leipzig
Christine R. Rose, Düsseldorf
Stefan Rotter, Freiburg
Isabella Sarto-Jackson, Klosterneuburg, Österreich
Constance Scharff, Berlin
Rainer Schwarting, Marburg
Sophie Seidenbecher, Aarhus, DK
Mikael Simons, Göttingen
Christian Steinhäuser, Bonn
Monika Stengl, Kassel
Christiane Thiel, Oldenburg
Stefan Treue, Göttingen
Tanja Vogel, Freiburg
Christian Wegener, Würzburg
Melanie Wilke, Göttingen
Florentin Wörgötter, Göttingen

DE GRUYTER

ABSTRACTED/INDEXED IN Baidu Scholar · Cabells Journalytics · Case · Chemical Abstracts Service (CAS): CAplus · Chemical Abstracts Service (CAS) - SciFinder · CNKI Scholar (China National Knowledge Infrastructure) · CNPIEC: cnpLINKer · Dimensions · Efield_1946CO Discovery Service · Genamics JournalSeek · Google Scholar · Japan Science and Technology Agency (JST) · J-Gate · JournalGuide · JournalTOCs · KESLI-NDSL (Korean National Discovery for Science Leaders) · Microsoft Academic · MyScienceWork · Naver Academic · Naviga (Softweco) · Primo Central (ExLibris) · Publons · Qfield_1949M (Quality Open Access Market) · ReadCube · SCImago (SJR) · SCOPUS · Semantic Scholar · Sherpa/RoMEO · Summon (ProQuest) · TDNet · Ulrich's Periodicals Directory/ulrichsweb · WanFang Data · WorldCat (OCLC) · Yewno Discover

ISSN 0947-0875 · e-ISSN 2363-7013

Alle Informationen zur Zeitschrift, wie Hinweise für Autoren, Open Access, Bezugsbedingungen und Bestellformulare, sind online zu finden unter https://www.degruyter.com/view/j/nf

HERAUSGEBER Neurowissenschaftliche Gesellschaft e.V. (NWG), Kontakt: Meino Alexandra Gibson, Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle-Straße 10, 13092 Berlin, Tel.: +49 (0)30 9406 3336, gibson@mdc-berlin.de, www.nwg-info.de

CHEFREDAKTEURIN Petra Wahle, AG Entwicklungsneurobiologie, Fakultät für Biologie & Biotechnologie, Ruhr-Universität, ND 6/72, 44780 Bochum, wahle@neurobiologie.ruhr-uni-bochum.de

REDAKTION Susanne Hannig, Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle-Str. 10, 13092 Berlin (Germany), Tel.: +49 (0)30 9406 3336, susanne.hannig@mdc-berlin.de

VERLAG Walter de Gruyter GmbH, Berlin/Boston, Genthiner Straße 13, 10785 Berlin, Germany

JOURNAL MANAGER Torsten Krüger, De Gruyter, Genthiner Straße 13, 10785 Berlin, Germany.
Tel.: +49 (0)30 260 05-173, Fax: +49 (0)30 260 05-250, E-Mail: Neuroforum.Editorial@degruyter.com

ANZEIGENVERANTWORTLICHE top-ad Bernd Beutel, Schlossergäßchen 10, 69469 Weinheim,
Tel.: +49 (0)6201 290 92-0, Fax +49 (0)6201 290 92-20 20, disposition@top-ad-online.de

COVER ILLUSTRATION ©Lyuba Zehl (https://gin.g-node.org/lzehl/ExperimenTales)

SATZ TNQ Technologies, Chennai, India

DRUCK Franz X. Stückle Druck und Verlag e.K., Ettenheim

FSC
www.fsc.org
MIX
Paper from responsible sources
FSC® C107500

VORSTAND DER AMTSPERIODE 2021–2023

Offenlegung der Inhaber und Beteiligungsverhältnisse gem. § 7a Abs. 1 Ziff. 1, Abs. 2 Ziff. 3 des Berliner Pressegesetzes: Die Gesellschafter der Walter de Gruyter GmbH sind: Cram, Gisela, Rentnerin, Berlin; Cram, Elsbeth, Pensionärin, Rosengarten-Alvesen; Cram, Dr. Georg-Martin, Unternehmens-Systemberater, Stadtbergen; Cram, Maike, Wien (Österreich); Cram, Jens, Mannheim; Cram, Ingrid, Betriebsleiterin, Tuxpan/Michoacan (Mexiko); Cram, Sabina, Mexico, DF (Mexiko); Cram, Silke, Wissenschaftlerin, Mexico DF (Mexiko); Cram, Björn, Aachen; Cram, Berit, Hamm; Cram-Gomez, Susana, Mexico DF (Mexiko); Cram-Heydrich, Walter, Mexico DF (Mexico); Cram-Heydrich, Kurt, Angestellter, Mexico DF (Mexico); Duvenbeck, Brigitta, Oberstudienrätin i.R., Bad Homburg; Gädeke, Gudula, M.A., Atemtherapeutin/Lehrerin, Tübingen; Gädeke, Martin, Student, Ingolstadt; Gomez Cram, Arturo Walter, Global Key Account Manager, Bonn, Gomez Cram, Ingrid Arlene, Studentin, Mexico, DF (Mexiko), Gomez Cram, Robert, Assistant Professor. London UK, Lubasch, Dr. Annette, Ärztin, Berlin; Schütz, Dr. Christa, Ärztin, Mannheim; Schütz, Sonja, Diplom.-Betriebswirtin (FH), Berlin; Schütz, Juliane, Berlin; Schütz, Antje, Berlin; Schütz, Valentin, Mannheim; Seils, Dorothee, Apothekerin, Stuttgart; Seils, Gabriele, Journalistin, Berlin; Seils, Christoph, Journalist, Berlin; Siebert, John-Walter, Pfarrer, Oberstenfeld; Anh Vinh Alwin Tran, Zürich (Schweiz), Tran, Renate, Mediatorin, Zürich (Schweiz).

Contents

Special Issue: NFDI – Nationale Forschungsdateninfrastruktur

Guest Editor: Thomas Wachtler

Editorial

Review articles

Nachrichten aus der Gesellschaft

Editorial

Michael Denker, Alexandra Stein and Thomas Wachtler*

Editorial

https://doi.org/10.1515/nf-2020-0042

Many of the scenarios around data management depicted on the title page of this issue of *Neuroforum* are likely to be rather familiar. Among the challenges abstracted in comic form by Lyuba Zehl are situations like transferring data between collaborating partners, integrating individual pieces of a complex, distributed data set, server shelfs overflowing from the sheer amount of data, the need to convert data between different formats, the danger of data loss, or simply the need to find an efficient, useful organization of data.

Managing research data is an everyday concern for us as scientists. The rapid developments we see in recording technology and computational methods offer enormous opportunities for neuroscience investigations – but they also bring the need to handle a massive increase in volume, diversity, and complexity of the acquired data, of data analysis workflows and of derived data. This creates difficulties for storing, accessing, and keeping track of the data. In many areas of neuroscience, it has always been a challenge for scientists to ensure reproducibility of their research and comprehensibility of the data. These problems are magnified by the continued growth of data volume and the increased need for collaboration.

Neuroscientists are not the only ones in this situation. Similar problems can be found in other fields of science. This has led to a recently established funding program for developing research data management in Germany, the *Nationale Forschungsdateninfrastruktur* (NFDI). Despite the technically sounding name, it is not the primary aim of this program to establish new data storage or computing facilities. Infrastructure is seen in a broad sense, comprising not only technical resources but also competences and structures for communication and collaboration. The NFDI is intended to be a process in the scientific community that leads to more efficient research through better use of resources, higher interoperability, and reusable data.

The NFDI consortium initiative for neuroscience – in short: NFDI-Neuro – reflects the cooperative spirit of the NFDI by uniting neuroscientists from various communities, including three large neuroscience societies: The German Neuroscience Society (NWG), the Bernstein Network Computational Neuroscience, and the Deutsche Gesellschaft für klinische Neurophysiologie. The NWG has been an important active supporter of NFDI-Neuro, and notably the jNWG, the representation of young neuroscientists within the NWG, has shown great enthusiasm for the goals of NFDI-Neuro. We are grateful to the NWG and the editorial board of Neuroforum for the invitation to this special issue, and in particular to the Editor in Chief Petra Wahle and Editorial Officer Susanne Hannig for their support.

The articles in this special issue present different perspectives on research data management and its impact on neuroscience research, including the developments envisioned for improving neuroscience research data management in the NFDI. The article by Wachtler et al. gives an overview of the NFDI-Neuro consortium initiative and the proposed approach of tackling the challenges of data management in neuroscience by making existing solutions available to neuroscientists more easily and strengthening expertise and skills in the community. Hanke et al. advocate an approach to research data management that is based on decentralized, lightweight services and thus is flexible and oriented to the needs of researchers. Denker et al. consider the practical level of the individual lab and discuss challenges, emerging solutions and conceptual considerations that arise in preparing and sharing data, based on the example of an electrophysiological study. Klingner et al. discuss challenges and possible solutions of research data management in clinical neuroscience, where translational approaches would benefit from better availability of research data. The authors consider barriers at the levels of technology, knowledge and motivation. Not the least,

---

***Corresponding author: Thomas Wachtler**, Department of Biology II, Ludwig-Maximilians-Universität München, Martinsried, Germany, E-mail: wachtler@bio.lmu.de. https://orcid.org/0000-0003-2015-6590

Michael Denker, Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany, E-mail: m.denker@fz-juelich.de

Alexandra Stein, Bernstein Coordination Site, Bernstein Network Computational Neuroscience, Freiburg und Jülich, Germany, E-mail: a.stein@fz-juelich.de

legal and ethical questions need to be addressed when dealing with data from human subjects.

In addition to these contributions, we look forward to an article about FAIR data in neuroscience by Maryann Martone in the next issue of Neuroforum. Our distributed special issue is further complemented by a symposium on “FAIR data management and data sharing in neuroscience” at the NWG’s 14th Göttingen Meeting in March 2021 (https://www.nwg-goettingen.de/2021/). We invite all neuroscientists to join the NFDI-Neuro initiative.

DE GRUYTER Neuroforum 2021; 27(1): 3–15

Review article

Thomas Wachtler*, Pavol Bauer, Michael Denker, Sonja Grün, Michael Hanke, Jan Klein, Steffen Oeltze-Jafra, Petra Ritter, Stefan Rotter, Hansjörg Scherberger, Alexandra Stein and Otto W. Witte

NFDI-Neuro: building a community for neuroscience research data management in Germany

https://doi.org/10.1515/nf-2020-0036

Abstract: Increasing complexity and volume of research data pose increasing challenges for scientists to manage their data efficiently. At the same time, availability and reuse of research data are becoming more and more important in modern science. The German government has established an initiative to develop research data management (RDM) and to increase accessibility and reusability of research data at the national level, the Nationale Forschungsdateninfrastruktur (NFDI). The NFDI Neuroscience (NFDI-Neuro) consortium aims to represent the neuroscience community in this initiative. Here, we review the needs and challenges in RDM faced by researchers as well as existing and emerging solutions and benefits, and how the NFDI in general and NFDI-Neuro specifically can support a process for making these solutions better available to researchers. To ensure development of sustainable research data management practices, both technical solutions and engagement of the scientific community are essential. NFDI-Neuro is therefore focusing on community building just as much as on improving the accessibility of technical solutions.

***Corresponding author: Thomas Wachtler**, Department Biologie II, Ludwig-Maximilians-Universität München, Grosshaderner Str. 2, 82152 Planegg-Martinsried, Germany, E-mail: wachtler@bio.lmu.de. https://orcid.org/0000-0003-2015-6590

Pavol Bauer, Department of Cellular Neuroscience, Leibniz Institute for Neurobiology, Brenneckestrasse 6, 39118 Magdeburg, Germany, E-mail: pavol.bauer@lin-magdeburg.de. https://orcid.org/0000-0003-4328-7171

Michael Denker, Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Research Center Jülich, Jülich, Germany, E-mail: m.denker@fz-juelich.de. https://orcid.org/0000-0003-1255-7300

Sonja Grün, Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Research Center Jülich, Jülich, Germany; and Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany, E-mail: s.gruen@fz-juelich.de. https://orcid.org/0000-0003-2829-2220

Michael Hanke, Institute of Neuroscience and Medicine Brain & Behaviour (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52425 Jülich, Germany; and Medical Faculty, Institute of Systems Neuroscience, Heinrich Heine University, 40225 Düsseldorf, Germany, E-mail: michael.hanke@gmail.com. https://orcid.org/0000-0001-6398-6370

Jan Klein, Fraunhofer Institute for Digital Medicine MEVIS, Am Fallturm 1, 28359 Bremen, Germany, E-mail: jan.klein@mevis.fraunhofer.de. https://orcid.org/0000-0001-9421-3549

Steffen Oeltze-Jafra, Department of Neurology, Otto von Guericke University Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany, E-mail: steffen.oeltze-jafra@med.ovgu.de. https://orcid.org/0000-0002-6962-9080

Petra Ritter, Charité – Universitätsmedizin Berlin, Freie Universität Berlin, Humboldt-Universität zu Berlin, Charitéplatz 1, 10117 Berlin, Germany; and Department of Neurology, Brain Simulation Section, Berlin Institute of Health, Charitéplatz 1, 10117 Berlin, Germany, E-mail: petra.ritter@charite.de

Stefan Rotter, Bernstein Center Freiburg & Faculty of Biology, University of Freiburg, Hansastraße 9a, 79104 Freiburg, Germany, E-mail: stefan.rotter@bio.uni-freiburg.de. https://orcid.org/0000-0003-3534-6530

Hansjörg Scherberger, Neurobiology Laboratory, Deutsches Primatenzentrum GmbH, Kellnerweg 4, 37077 Göttingen, Germany; and Department of Biology and Psychology, University of Göttingen, Göttingen, Germany, E-mail: hscherb@gwdg.de. https://orcid.org/0000-0001-6593-2800

Alexandra Stein, Institute of Neuroscience and Medicine Computational and Systems Neuroscience (INM-6), Research Center Jülich, Jülich, Germany; and Bernstein Coordination Site, Bernstein Network Computational Neuroscience, Albert-Ludwigs Universität Freiburg, Hansastraße 9a, 79104 Freiburg, Germany, E-mail: a.stein@fz-juelich.de

Otto W. Witte, Hans Berger Department of Neurology, Jena University Hospital, Am Klinikum 1, 07747 Jena, Germany; and Deutsche Gesellschaft für Klinische Neurophysiologie und Funktionelle Bildgebung (DGKN), Salvador-Allende-Platz 29, 07747 Jena, Germany, E-mail: otto.witte@med.uni-jena.de. https://orcid.org/0000-0003-2101-4105

Keywords: collaboration; data management; FAIR; research data infrastructure.

Zusammenfassung: Die kontinuierlich steigende Menge und Komplexität von Forschungsdaten stellt Wissenschaftler:innen vor besondere Herausforderungen in Bezug auf effizientes Management dieser Daten. Gleichzeitig wächst in der modernen Wissenschaft die Bedeutung von Verfügbarkeit von Forschungsdaten und deren Wiederverwendung. Aus diesem Grund hat die Bundesregierung eine nationale Initiative zur Förderung des Forschungsdatenmanagements (FDM) ins Leben gerufen: Die Nationale Forschungsdateninfrastruktur (NFDI). Das Konsortium NFDI Neuroscience (NFDI-Neuro) soll die neurowissenschaftliche Community in dieser Initiative vertreten. Wir betrachten hier die besonderen Herausforderungen und Bedarfe im Forschungsalltag sowie die vorhandenen Werkzeuge und Lösungen und stellen dar, wie die NFDI und NFDI-Neuro diese für Forschende besser verfügbar machen kann. Damit sich eine Kultur des nachhaltigen Forschungsdatenmanagements entwickeln kann, ist das Engagement der wissenschaftlichen Community unersetzlich. Daher setzt NFDI Neuroscience nicht nur auf die Verbesserung und Entwicklung von technischen Lösungen für FDM, sondern ebenso darauf, die neurowissenschaftliche Community zusammenzubringen, damit Entwickler:innen und Wissenschaftler:innen gemeinsam an nützlichen, leicht handhabbaren Werkzeugen für ein solides FDM arbeiten.

Schlüsselwörter: Forschungsdateninfrastruktur; Datenmanagement; Kollaboration; FAIR.

Introduction

Access to digital knowledge and management of data from publicly funded research are essential challenges for research and knowledge transfer. To support the digital transition of science in Germany, the federal and state governments established the Council for Information Infrastructures (RfII). The RfII comprises members representing the scientific community, providers of information infrastructures, federal and state governments and the public. Based on analysis of developments in digital science and policies, it provides advice to academia and the government supporting coordination and cooperation (https://rfii.de). In a series of discussion papers (https://rfii.de/documents/), the RfII suggested the establishment of a national initiative, the "Nationale Forschungsdateninfrastruktur" (national research data infrastructure; NFDI) to increase cooperation and efficiency of research data infrastructures.

The NFDI is envisioned to be a process spanning the entire scientific landscape, organized by consortia representing different scientific communities. Its purpose is to support scientists in efficiently managing their research data and to ensure that research data becomes findable, accessible, interoperable and reusable, according to the FAIR principles (Wilkinson et al., 2016) and in line with international standards and initiatives. The NFDI aims to build on and connect existing resources according to a comprehensive concept for research data management (RDM) that is sustainable and competitive in the international context.

NFDI Neuroscience

The consortium initiative NFDI Neuroscience (NFDI-Neuro, https://nfdi-neuro.de) formed as an open community network, with the aim of acting as a platform that brings together existing solutions for RDM and assists researchers in establishing RDM as part of everyday research practice. The initiative is supported by three major neuroscience associations: Neurowissenschaftliche Gesellschaft (NWG), Bernstein Network Computational Neuroscience and Deutsche Gesellschaft für klinische Neurophysiologie (DGKN).

A key NFDI goal is that research data is handled in accordance with the FAIR principles (Wilkinson et al., 2016) to ensure findability, accessibility, interoperability and reusability, by using appropriate and interoperable solutions for data storage, data annotation, data integration and data processing. NFDI-Neuro pursues a concept where the consortium acts as a direct point of contact for researchers regarding any RDM aspects related to this goal. It will bring service providers and users together and push forward new developments based on needs identified by the neuroscientific community. As such, NFDI-Neuro will build up a competence network interwoven with the neuroscientific community. To achieve this, the NFDI-Neuro initiative has proposed a strategy that is currently under review, with a decision to be expected in the summer of 2021. The key aspects of this proposal will be elaborated in the following.

Establishing RDM infrastructure in neuroscience is a twofold challenge. First, given the complexity of neuroscientific data and workflows on one hand, and the high diversity and differences in methods and conventions in different laboratories on the other hand, creating unified and interoperable solutions is a demanding task. Second, uptake of new methods and tools is hampered by lack of resources and expertise for RDM in the neuroscience laboratories. Therefore, NFDI-Neuro takes a two-pronged approach to establishing a viable RDM infrastructure in neuroscience, combining development and provision of methods for standardized data handling with fostering collaboration and competence in RDM throughout the neuroscience community.

Researchers need practical solutions and readily available tools and services to establish efficient RDM in the daily lab routines. This implies the development of standards, tools and services that take up existing workflows and practices as they currently exist in labs, in order to minimize interference with established research structures. Progressive extension of these existing resources and building connections between resources will therefore be the focus of actions toward FAIR data management. Thus, NFDI-Neuro's strategy is based on the concepts of decentralization, use of existing infrastructure, adoption of commodity technologies and community as well as industry standards.

Community

To ensure that technical developments and solutions actually address real needs of the scientists, active participation of the neuroscience community is essential. NFDI-Neuro follows a profoundly bottom-up strategy, with utmost accessibility and openness. The scientific community is invited to participate in all activities and can actively shape the process. To support the building of the network, NFDI-Neuro proposes to employ specific instruments outlined in the following.

Transfer Teams are a central element in the NFDI-Neuro structure. These are teams of experts which combine expertise in research, as well as information technology, and RDM. They form a geographically and topically distributed network that offers ample opportunity for the community to get in contact with the initiative and to get involved, benefit and contribute. Transfer Teams will proactively seek information, initiate interactions, organize training activities, establish necessary links to inter-/national initiatives for collaboration on common solutions and drive specific developments of concepts, tools and services.

The core instrument for organizing community-driven cooperation and co-development are *Working Groups*. Here, researchers, developers and providers come together to work on specific problems, for example, metadata standards for a certain research domain, guidelines for a specific RDM task, or the definition of an interface to enhance interoperability or usability. In addition, *Dynamic Support Actions* are mechanisms to provide funding for necessary developments, which are identified as the initiative is running. This can be, for example, the implementation of an interface according to specifications defined by a Working Group, or the enhancement of a tool developed in the community to make it interoperable and more widely usable. These structural instruments are seen as a framework supporting the build-up of expertise and competence in the community in jointly addressing needs, seeking partners and sharing knowledge.

The NFDI-Neuro initiative has held a series of Community Workshops (see also Ritzau-Jost and Seidenbecher, 2019) focusing on various groups involved in the process, including individual researchers, research consortia and providers of RDM services. Across all groups, the necessity for a more coherent approach to tackle research data management was recognized. Specific implementations and guidelines were identified as potential targets for improvements. In addition, the workshops facilitated discussions in the community on overarching concerns, such as measures for ensuring the quality of data records. The process of community engagement shaped the strategy of NFDI-Neuro, and future Community Workshops will continuously drive the process and progress of the consortium, its strategy and its activities.

Research data management strategy

The conceptual and logistical challenges of integrating heterogeneous and complex high-volume data at all stages of the research lifecycle, from data acquisition to data analysis and publication, affect the individual lab but also hinder the field at large. A substantial proportion of research data gets practically lost when investigators responsible for practical data generation and acquisition—often PhD students or Postdocs—leave a lab, because metadata required for analysis and reuse are insufficiently recorded. Likewise, reusability in general as well as the usefulness of data and software repositories critically depend on sufficient data annotation and the use of interoperable formats. Various neuroinformatics initiatives have started to tackle these problems by creating a number of tools for data logistics (e.g., DataLad, https://datalad.org), metadata collection (e.g., NIDM, https://nidm.nidash.org; odML, Grewe et al., 2011), dataformats and structures (e.g., BIDS, Gorgolewski et al., 2016; NIX, http://www.g-node.org/nix; NWB, http://nwb.org), data representation (Neo, Garcia et al., 2014), data analysis (e.g., Elephant, http://python-elephant.org; FieldTrip, https://www.fieldtriptoolbox.org; Freesurfer, https://surfer.nmr.mgh.harvard.edu; MNE, https://mne.tools/) or simulation (e.g., NEST, https://www.nest-initiative.org; Neuron, https://neuron.yale.edu/). However, uptake of these solutions in the broader scientific community is limited. Utilizing new tools often requires technical skills and interoperability that are rarely found in laboratory practice. NFDI-Neuro's strategy is to leverage these existing

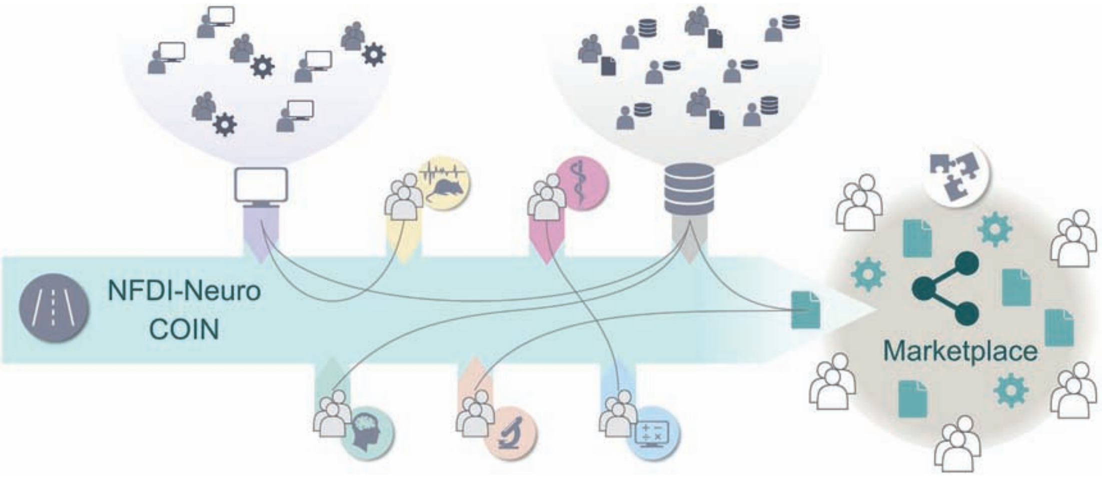


Figure 1: Illustration of the role of the NFDI-Neuro common infrastructure as a technical backbone to connect providers of tools, services and data (top), researchers in different neuroscience domains (colored icons) and the broad community of users and providers via a shared marketplace (right). The concept supports efficiency by focusing on interoperability of community resources and provides a stable foundation for collaborative work.

resources and contribute critically missing pieces to enable neuroscientists to improve their RDM throughout the data lifecycle. This strategy must aim for standardization, as is required for sharing and reuse, but must also acknowledge the inherent heterogeneity in neuroscience, which is a consequence of the diversity of approaches, levels of investigation and systems studied. NFDI-Neuro therefore takes a modular approach, establishing (a) a common infrastructure as a compatibility layer for accessing resources of the wider scientific community and (b) domain-specific, interoperable tools and interfaces to fit into the researchers' familiar research environments.

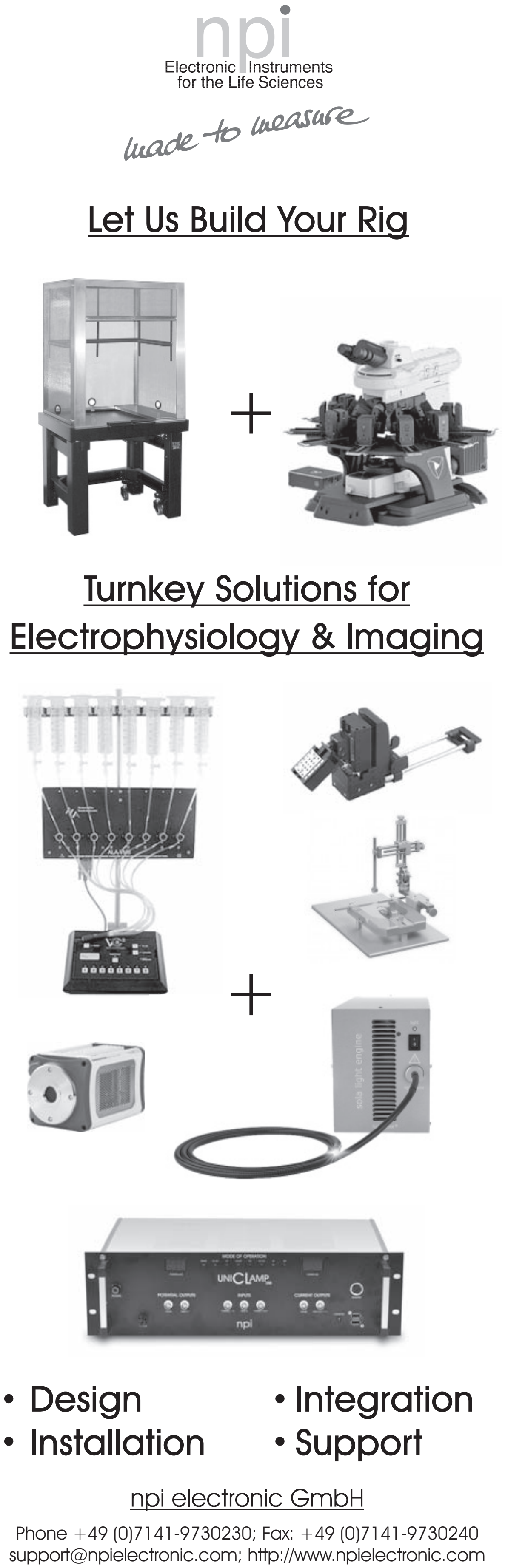
Common infrastructure

The fundamental concept of the NFDI-Neuro common infrastructure (NFDI-Neuro COIN) is to use popular commodity technologies and make them compatible with ongoing research practice to transform today's procedures for management and publication of scientific data (Figure 1). A core aim is to enable researchers to communicate their study outputs on a large number of outlets without being limited by over-simplified metadata schemata or the need to establish specific interfaces. Scientific outputs will be available via standard interfaces and open-source tools, regardless of authorship and publication venue. While this is key for making existing achievements accessible in the future, the approach of interfacing with a plurality of interoperable services (see Hanke et al., this issue) will also substantially improve the resilience of today's research infrastructure, where too often crucial activities depend on a few key pieces (e.g., GitHub or individual data hosting providers). This aims to directly improve the availability of data that are described by detailed, tightly connected metadata.

A marketplace for neuroscience research outputs

NFDI-Neuro will develop and operate a unique online marketplace for neuroscience research output. This venue will provide a unified, public entry point for discovering primary and derived data, and computational models, for potential consumers from within or outside neuroscience.

The marketplace will offer a catalog of datasets and advertise metadata of datasets. This information will be provided not only as a convenient website for human consumption, but also in machine-readable form for automated ingestion into standard industry search engines such as Google dataset search. Importantly and unlike other catalogs, a dataset need not be deposited in a public repository to be included in the marketplace. Instead, it is an expressed goal of the marketplace to also improve the



findability of datasets that have not been or cannot be published (e.g., privacy concerns prohibit the publication of high-value data).

Specific areas of addressing the needs of neuroscientists

Conceptual and practical improvements achieved by community efforts need to reach the lab workflows to be effective. Rather than trying to come up with one-size-fits-all solutions, NFDI-Neuro's strategy is to address the needs in a domain-specific way, which enables utilizing already existing building blocks in the different areas efficiently. Interoperability is achieved through the common infrastructure, which provides a unifying technical backbone on which domain-specific interfaces can be created. All developments will be directed toward connecting existing approaches and tools to the common infrastructure and establish solutions for contributing to the neuroscience data marketplace. In all domains, metadata standards, measures for quality control and training measures will be established and community support services will be provided. We briefly summarize the developments proposed to improve RDM in different areas of neuroscience in a coordinated way such that coherent and interoperable solutions arise.

Neuroimaging

In Neuroimaging, various techniques, such as CT, MRI, PET and optical imaging, are used to image structure, function and molecular architecture of the nervous system. Neuroimaging studies generate complex image data and metadata by using multiple modalities, for example, PET and MRI, or variations of the same modality, such as multiple MRI sequences yielding different contrasts, often longitudinally. Moreover, imaging is commonly paralleled by the acquisition of nonimaging data (physiological, behavioral, etc.). Subsequent processing and analysis with advanced computational methods generate a multitude of additional complex derived data. Thus, neuroscientists are confronted with massive volumes of multimodal, multidimensional, high-resolution data, derived data and metadata for which no overarching standards or RDM strategies have yet been widely established.

For neuroimaging data, two major data formats exist: DICOM (industry standard for medical imaging, dicomstandard.org) and NIfTI (community standard for data in ready-to-analyze/ exchange form, nifti.nimh.nih.gov/nifti-1). However, standardization across modalities is limited. Emerging standards for (meta)data structures beyond device/acquisition metadata (e.g., the Brain Imaging Data Structure—BIDS, Gorgolewski et al., 2016, or the Neuroimaging Data Model—NIDM, nidm.nidash.org) exist and are continuously developed. BIDS already provides a good basis for long-term data access and reuse. However, BIDS is primarily concerned with raw data description; standardization efforts for derived data are still in an early conceptual stage, and efforts to close the gap between image and other nonimage data have just started (e.g., EEG-BIDS, Pernet et al., 2019).

Measures to address these challenges will focus on standardization and integration of data from different neuroimaging modalities. This will include the development of standards for the formalization and annotation of data, metadata and derived data. Image acquisition will be automated as far as possible and will be connected with automatic quality assurance. Processing of metadata, imaging data and derived data across various modalities will be standardized to improve the accessibility of processed data for use in larger studies, as required in AI-related research questions and also to save time and resources needed for converting between the existing file formats. Standardization of processed data will also help the community in developing algorithms that can be used and adapted easily for different research questions.

Systems and behavioral neuroscience

The field of systems and behavioral neurosciences is characterized by large and valuable datasets of complex and heterogeneous data. They are elaborately evaluated for answering the original research question, but can also address other research questions, even years after the original experiment has been conducted. The development of advanced methods for data processing and analysis, in particular for the analysis of high-dimensional electrophysiological data is paramount. What is lacking are stringent and coherent RDM practices according to the FAIR principles. There is a need for new strategies and interoperable software tools to facilitate the efficient storage of primary data and metadata, as well as data provenance information. To allow the FAIR use of this data in collaborative research environments, coherent data processing and data analysis pipelines are also urgently needed. They are crucial to facilitate transfer from data acquisition to analysis, modeling and simulation and to share data with other fields of neuroscience.

Measures to address these challenges will focus on the annotation of heterogeneous data and complex data analysis workflows for reproducibility (see Denker et al., this issue). This will include developing metadata schemata for heterogeneous data acquisition workflows, standardized workflows for metadata aggregation and preprocessing and descriptions for data analysis workflows and provenance.

Cellular and molecular neuroscience

In the past years, a development toward more multiplexed and more complex data has started in the field of molecular and cellular neuroscience. With increasing availability and feasibility of *in vivo* methods, allowing for the monitoring and manipulation of neural activity on the single cell and circuit level, experimenters were facing the necessity to simultaneously monitor multiple behavioral parameters from video or other interfaced measurement probes. Moreover, the traditional *in vivo* monitoring of neural activity data with electrophysiology and light microscopy followed a trend toward high-density, high-frequency recordings from large neural populations with single-cell resolution. Representative examples for this trend were volumetric calcium imaging methods such as “Mesoscopes,” allowing for cellular resolution *in vivo* imaging within entire rodent brain hemispheres.

As a consequence of these developments, researchers were abruptly confronted with massive volumes of multimodal and highly multidimensional datasets for which currently no common management standard exists. Hence, to combine microscopic data with data of other modalities, such as behavioral measures, small animal PET, functional connectomics or electrophysiological recordings, suitable data formats will be required. Also, for efficient sharing of such complex data structures within the community and beyond, defined metadata standards are needed that include both technical information and comprehensive experiment-specific information. Moreover, analogous to challenges in preprocessing of data from high-density electrophysiological recordings, there is no standard for reproducible preprocessing of imaging data, such as deconvolution, source extraction or spike inference. As data sharing requirements are of increasing importance, the field requires a movement toward an improved data annotation and assessment of data quality.

Measures to address these challenges will focus on annotation, storage and management of molecular and cellular imaging data and preprocessing procedures.

Existing solutions for microscopy imaging data like OMERO (https://www.openmicroscopy.org) will be utilized and extended for the management of neuroscientific imaging data. Interoperability tools will be created to import preprocessing results. Data quality and reuse require high accuracy during initial upload of data into the RDM system. Inaccuracies at this early stage are hard to correct later. Therefore, not only the detailed specification of data and metadata to be documented will be developed but also a web interface that supports the experimenter during the documentation and upload process. In particular, this interface ensures the completeness and, as much as this is possible at this point, the correctness of the data. Criteria and procedures of this step are so designed as to ease the migration of data and metadata to public repositories.

Clinical neuroscience

Many of the recording techniques used in clinical neuroscience have evolved into new dimensions by the use of sophisticated analysis techniques. This allows for a semiautomatic and quantitative evaluation of MRI, such as determination of the biological age of the brain by the Brainage framework, quantification of disease progression, localization of circumscribed brain abnormalities in epilepsy, characterization of chemical properties of the brain by spectroscopy and semiautomatic detection of abnormalities in early stroke by machine learning algorithms.

Unfortunately, data formats of the techniques mentioned as well as metadata are dominated by a wide variety of proprietary solutions. Usually, the data are neither findable, nor accessible, interoperable or reusable. Similar to the neuroimaging domain, this has restrained the progress of research on many highly important topics, such as consciousness and cognition, psychiatric and neurological disorders. Only recently, a discussion on the implementation of the FAIR principles in clinical contexts has started. The implementation of these principles is hampered by the necessity to care for the security and privacy of the data and the necessity to adhere to the General Data Protection Regulations (GDPR).

Measures to address these challenges will focus on improving translation from basic to clinical research and on legal and ethical aspects in particular with regard to personal data (see Klingner et al., this issue). This will include the development of data formats and standards supporting the comparison between basic and clinical datasets and integration of clinical neuroscience data from industry healthcare systems into FAIR research infrastructure. Furthermore, certifiable standards for workflows and IT system designs are required that achieve an optimal balance between keeping data secure and remaining risk for the individual low, and simultaneously enable researchers to use personal data in compliance with GDPR.

Computational neuroscience

Computational neuroscience provides an interface between theory and most experimental approaches in neuroscience. The resulting heterogeneity is mirrored by the discipline’s own approaches, and stringent and coherent RDM is essential for sustainable progress in the field. However, RDM is currently organized on small scales, often initiated and run by individual researchers using repositories under version control. Scattered community services exist to collect data and models relevant for specific aspects, such as specific reconstructed single-neuron morphologies or network connectivity information (e.g., the Neocortical Microcircuit Collaboration Portal, Ramaswamy et al., 2015), or model simulation scripts (e.g., ModelDB, Hines et al., 2004; OpenSourceBrain, https://www.OpenSourceBrain.org; NeuroML, https://neuroml.org). New developments pursued by international initiatives are rapidly maturing to offer new technical vistas, such as the EBRAINS service catalog. However, utilizing these services to address user demands requires standardization and active involvement of the computational neuroscience community.

Measures to address these challenges will focus on three areas: First, the ability to perform robust comparisons across different approaches used in computational neuroscience by improving the description and interoperability of models and their associated metadata to enable the reproducible simulations. Second, to strengthen the ability of computational neuroscience to relate to the wealth of available experimental data, improving the design, installation and dissemination of modular, shareable, reproducible analysis pipelines across the breadth of simulated and experimental data and developing detailed generalizable schemata to describe analysis results for reuse in line with FAIR principles. Third, to align the highly heterogeneous experimental data to model outputs for rigorously performing validation testing and to increase the explanatory power of models, exploring and implementing designs for in-depth metadata of model descriptions and simulation outputs on all levels of resolution to match those of experimental data.

International initiatives

Developments toward standardization and improvement of RDM need to consider and align with activities at the international level. The NFDI-Neuro community has close connections with international initiatives.

The International Neuroinformatics Coordinating Facility (INCF, https://incf.org) was established in 2006 to coordinate development in neuroinformatics at a global level. INCF's mission is to promote the application of computational approaches in neuroscience and to provide coordination of neuroscience infrastructure through the development and endorsement of standards and best practices in support of open and FAIR neuroscience. INCF coordinates an international network of neuroinformatics initiatives, fosters collaboration and offers training resources. Recently, INCF has started to focus on coordination at the level of international organizations and initiatives and on identifying, evaluating and endorsing community standards and best practices. Given the close connections already existing between the neuroscience community in Germany and the INCF, INCF will be a strong partner for NFDI-Neuro at the international level.

The EU Flagship Human Brain Project (HBP, https://www.humanbrainproject.eu) started in 2013 with the aim to design and implement a platform for accelerating integrated, collaborative neuroscience research. The project developed six technological platforms for the fields of neuroinformatics, brain simulation, high-performance computing, medical informatics, brain-inspired computing and neurorobotics that are currently being harmonized into a cohesive offering to address cyberinfrastructure-related challenges in brain science, called EBRAINS (https://ebrains.eu). The focus of EBRAINS is the development and maintenance of a novel digital research infrastructure with tools and services for different user communities in Europe. EBRAINS operates a platform and software-as-a-service (SaaS), supporting codesign, digital workflows, open science and translating knowledge.

The Canadian Open Neuroscience Platform (CONP) provides an infrastructure for the promotion of open-science workflows and the sharing of neuroscience data. Funded by a Brain Canada grant with broad commitment across Canadian neuroscience research institutions, CONP supports basic neuroscience and clinical neuroscience research communities to share phenotypic/genotypic data and methods in an unrestricted manner, create large-scale databases, facilitate the use of advanced multivariate analytic strategies, train the next generation of computational neuroscience researchers and disseminate findings to the global community.

DANDI is a platform for publishing, sharing and processing cellular neurophysiology data, funded by the US BRAIN Initiative. It aims to enable reproducible practices, publications and reuse of data, reduce the need to contact data producers by enriching the data with comprehensive metadata, with the goal to build a living repository that enables collaboration within and across labs, and for others, the entry point for research.

Training

Educating researchers is a key element for improving RDM. Thus, a major task in promoting knowledge and competence in RDM throughout the neuroscience community will be to develop and implement a coordinated training concept. Currently, training in RDM is either not provided at all or carried out in an ad hoc manner. What is lacking is a well-structured solution, taking into consideration all stages of the research education and research career.

Measures to address these challenges will include the coordination of courses relevant to the different subdomains in neuroscience, developing RDM curricula which can be integrated in Bachelor and Master or doctoral programs. These shall be designed in a modular fashion containing modules covering basic and more general aspects of RDM as well as advanced courses for specific RDM aspects concerning the various neuroscience subdomains and various data types or analyses. Train-the-trainer activities and networking for colleagues who teach RDM will complement these measures, enhancing the overall quality of RDM training.

Training measures will be coordinated in collaboration with the Graduate Schools for Neuroscience in Germany (www.neuroschools-germany.com) and the jungeNWG on a national level. On a European level, NFDI-Neuro will cooperate with the Federation of European Neuroscience Societies (FENS) and the Network or European Neuroscience Schools (NENS).

Conclusions

We all know how challenging it is to make our data findable, understandable and reusable. The NFDI as a whole and NFDI-Neuro in particular are aimed at building a network for jointly addressing RDM needs and developing solutions that will make these tasks easier. NFDI-Neuro will be a forum for exchange and building of expertise supporting neuroscientists in their data management efforts. In the long run, the benefits will outweigh the investment we have to make now. Let us together tackle this challenging task—become part of NFDI-Neuro!

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This research was funded by LMU excellent as part of LMU Munich's funding as University of Excellence in the German Excellence Strategy, Helmholtz Metadata Collaboration (HMC), European Union's Horizon 2020 research and innovation programme under grant agreement no. 826421 (VirtualBrainCloud) and 945539 (HBP SGA3), Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE), German Federal Ministry of Education and Research (BMBF 01GQ1905), ERC 683049; German Research Foundation CRC 1315, CRC 936, CRC-TRR 295 and RI 2073/6-1; Berlin Institute of Health & Foundation Charité, Johanna Quandt Excellence Initiative.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

Denker M., Grün S., Wachtler T., Scherberger H.. Reproducibility and efficiency in handling complex neurophysiological data. Neuroforum 2021. *27*;25–25.

Garcia, S., Guarino, D., Jaillet, F., Jennings, T.R., Pröpper, R., Rautenberg, P.L., Rodgers, C., Sobolev, A., Wachtler, T., Yger, P., and Davison, A.P. (2014). Neo: An object model for handling electrophysiology data in multiple formats. Front. Neuroinf. *8*, 10.

Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci. Data *3*, 1–9.

Grewe, J., Wachtler, T., and Benda, J. (2011). A bottom-up approach to data annotation in neurophysiology. Front. Neuroinf. *5*, 16.

Hanke M., Pestilli F., Wagner A.S., Markiewicz C.J., Poline J.-B., Halchenko Y.O.. In defense of decentralized research data management. Neuroforum 2021. *27*;15–15.

Hines, M.L., Morse, T., Migliore, M., Carnevale, N.T., and Hines, M.L. (2004). ModelDB: A database to support computational neuroscience. J. Comput. Neurosci. *17*, 7–11.

Klingner C.M., Ritter P., Brodoehl S., Gaser C., Scherag A., Güllmar D., Rosenow F., Ziemann U., Witte O.W.. Research data management in clinical neuroscience – the NFDI initiative. Neuroforum 2021. In this issue.

Pernet, C.R., Appelhoff, S., Gorgolewski, K.J., Flandin, G., Phillips, C., Delorme, A., and Oostenveld, R. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci. Data *6*, 103.

Ramaswamy, S., Courcol, J.D., Abdellah, M., Adaszewski, S.R., Antille, N., Arsever, S., Ateneken, G., Bilgili, A., Brukau, Y., Chalimourda, A., et al. (2015). The neocortical microcircuit collaboration portal: A resource for rat somatosensory cortex. Front. Neural Circ. *9*, 44.

Ritzau-Jost, A. and Seidenbecher, S. (2019). NFDI Neuroscience: Advocating cross-community data management in neuroscience. Neuroforum *25*, 279–280.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data *3*, 1–9.

Bionotes

Thomas Wachtler
Department Biologie II, Ludwig-Maximilians-Universität München, Grosshaderner Str. 2, 82152 Planegg-Martinsried, Germany
wachtler@bio.lmu.de
https://orcid.org/0000-0003-2015-6590

Thomas Wachtler has a background in physics and received his diploma and doctoral degree from the University of Tübingen. He was postdoctoral researcher at the Salk Institute for biological Studies and at the Universities of Freiburg and Marburg. His research interests are in the neural mechanisms of sensory processing with focus on vision. In his research he combines experimental and computational approaches, including electrophysiology, psychophysics and computational modeling, to study the neural principles of processing and coding in the visual system and how they relate to the properties of the sensory environment and to perceptual phenomena. He is also working on neuroinformatics developments in the context of the International Neuroinformatics Coordinating Facility. Since 2009, he has been Scientific Director of the German Neuroinformatics Node at LMU Munich, leading developments of tools and services for research data management in neuroscience.

Pavol Bauer
Department of Cellular Neuroscience, Leibniz Institute for Neurobiology, Brenneckestrasse 6, 39118 Magdeburg, Germany
pavol.bauer@lin-magdeburg.de
https://orcid.org/0000-0003-4328-7171

Pavol Bauer studied Computer Science at the Technical University of Vienna, Austria and received a doctoral degree in Scientific Computing from the Uppsala University, Sweden. In 2018, he pursued his postdoctoral research at the Neural Networks lab at the German Center for Neurodegenerative Diseases in Bonn. In 2020, he joined the Department of Cellular Neuroscience at the Leibniz Institute of Neurobiology, where he heads the Neural Data Science working group. His current research is focused on cell-type specific hippocampal neural activity and its correlations to the animal behavioral patterns. To improve the national standards of RDM in bioimaging, he is currently active within the Information Infrastructure for BioImage Data (I3D:bio) project as well as the NFDI4BIOIMAGE consortium.

Michael Denker
Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Research Center Jülich, Jülich, Germany
m.denker@fz-juelich.de
https://orcid.org/0000-0003-1255-7300

Michael Denker received his diploma in physics from the University of Göttingen, Germany, in 2002. In 2004, he started his doctoral studies in the lab of Sonja Grün at the Free University Berlin. In 2006, he became a researcher at the RIKEN Brain Science Institute, Japan. He was awarded his PhD in 2009 at the Free University Berlin, Germany. In 2011, he joined the Institute of Neuroscience and Medicine (Research Centre Jülich, Germany) and now leads the group Data Science in Electro- and Optophysiology Behavioral Neuroscience. His research interests are the analysis of the correlation structure of neural activity and its relationship to signals that express population activity, and the establishment of workflows that improve the reproducibility of data analysis in neurophysiology.

Sonja Grün
Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Research Center Jülich, Jülich, Germany
Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany
s.gruen@fz-juelich.de
https://orcid.org/0000-0003-2829-2220

Sonja Grün received her diploma in physics from the Eberhard Karls University in Tübingen (1991) and her PhD in physics from Ruhr University Bochum (1996). After being a postdoc at the Hebrew University (Jerusalem) and at Max-Planck Institute in Frankfurt (M), she became a junior professor at Freie University Berlin in 2002 and unit/team leader at RIKEN Brain Science Institute (Tokyo), in 2006. Since 2011, she is a full professor at RWTH Aachen University and leads the group Statistical Neuroscience (INM-6, Research Center Jülich), and was appointed director of INM-6/INM-10 in 2018. Her work focuses on the development of analysis strategies and tools that uncover concerted activity in massively parallel electrophysiological recordings from the cortex, which led to the additional focus on research data management.

Michael Hanke
Institute of Neuroscience and Medicine Brain & Behaviour (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52425 Jülich, Germany
Medical Faculty, Institute of Systems Neuroscience, Heinrich Heine University, 40225 Düsseldorf, Germany
michael.hanke@gmail.com
https://orcid.org/0000-0001-6398-6370

Michael Hanke studied psychology in Halle (Saale) and Magdeburg, Germany. After being a postdoc in the lab of James Haxby at Dartmouth College, together with Yaroslav O.Halchenko working on NeuroDebian and PyMVPA, he moved back to Magdeburg as a Juniorprofessor. Since 2019, he is a professor at Heinrich Heine University Düsseldorf and head of the Psychoinformatics lab at the Institute of Neuroscience and Medicine (Brain and Behavior) of the

Research Center Jülich. His group contributes to the development of the DataLad software, and develops workflows and training materials for research data management in neuroscience.

Jan Klein
Fraunhofer Institute for Digital Medicine MEVIS, Am Fallturm 1, 28359 Bremen, Germany
jan.klein@mevis.fraunhofer.de
https://orcid.org/0000-0001-9421-3549

Jan Klein is a member of the management board of the Fraunhofer Institute for Digital Medicine MEVIS (FME) and heads the neuroimaging activities at FME. Klein has led numerous industry and EU projects on multiple sclerosis, stroke imaging and neurosurgery and got first prizes at last year's MICCAI registration challenge (CuRIOUS 2019) and at the Eurographics conference for his application on neurosurgical planning.

Steffen Oeltze-Jafra
Department of Neurology, Otto von Guericke University Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany
steffen.oeltze-jafra@med.ovgu.de
https://orcid.org/0000-0002-6962-9080

Steffen Oeltze-Jafra heads the working group Medicine and Digitalization at the Department of Neurology, Otto von Guericke University (OVGU) Magdeburg, Germany. He is also a Privatdozent at the Faculty of Computer Science, OVGU. In 2016, he received a habilitation in Computer Science; in 2010, a Ph.D. in Computer Science and in 2004, a diploma in Computational Visualistics from the OVGU. From 2016 to 2018, Steffen was Scientific Director "Digital Patient- and Process Model" and Junior Research Group Leader at the Innovation Center Computer Assisted Surgery (ICCAS), Medical Faculty, Leipzig University, Germany. His research interests are in the quantitative analysis of clinical routine data, the visual analysis of biomedical data and in model-based clinical decision support. He is currently working on an imaging-based continuous registration and quantitative analysis of brain structure and function of all patients with neurological and neuropsychiatric disorders in Saxony-Anhalt.

Petra Ritter
Charité – Universitätsmedizin Berlin, Freie Universität Berlin, Humboldt-Universität zu Berlin, Charitéplatz 1, 10117 Berlin, Germany
Department of Neurology, Brain Simulation Section, Berlin Institute of Health, Charitéplatz 1, 10117 Berlin, Germany
petra.ritter@charite.de

Petra Ritter studied medicine at the Charité University Medicine Berlin. She spent a large part of her clinical traineeships and practical year abroad: at the universities UCLA and UCSD in Los Angeles and San Diego, the Mount Sinai School of Medicine in New York and the Harvard Medical School in Boston. In 2002, she received her license to practice medicine. In 2004, she completed her doctoral thesis at the Charité, and in 2010, she received habilitation in Experimental Neurology. After being Max Planck Minerva research group leader from 2011 to 2015, she assumed the lifetime position of BIH Johanna Quandt Professor for Brain Simulation at Berlin Institute of Health (BIH) and Charité Universitätsmedizin Berlin, one of Europe's largest university hospitals. Since 2017, she is director of the Brain Simulation Section at Charité Universitätsmedizin Berlin. Ritter holds an ERC Consolidator grant and serves in the leadership of several national and international neuroinformatics consortia.

Stefan Rotter
Bernstein Center Freiburg & Faculty of Biology, University of Freiburg, Hansastraße 9a, 79104 Freiburg, Germany
stefan.rotter@bio.uni-freiburg.de
https://orcid.org/0000-0003-3534-6530

Stefan Rotter is the Managing Director of the Bernstein Center Freiburg, an interdisciplinary research facility for Computational Neuroscience and Neurotechnology at the University of Freiburg. With a strong background in Mathematics and Physics, he is now Professor of Computational Neuroscience at the Faculty of Biology. His research is focused on functional and plastic networks in the brain. Large-scale numerical simulations are an important research tool for him, but he also regards them as a suitable vehicle for teaching students from various disciplines the concepts and results of modern Computational Neuroscience. Currently, he is spokesperson of the User Committee for large-scale IT infrastructures in Baden-Württemberg.

Hansjörg Scherberger
Neurobiology Laboratory, Deutsches Primatenzentrum GmbH, Kellnerweg 4, 37077 Göttingen, Germany
Department of Biology and Psychology, University of Göttingen, Göttingen, Germany
hscherb@gwdg.de
https://orcid.org/0000-0001-6593-2800

Hansjörg Scherberger heads the Neurobiology Laboratory at the German Primate Center and is Professor for Primate Neurobiology at Göttingen University (since 2008). He received his Master in Mathematics (1993) and his Medical Doctor (1996) from Freiburg University, Germany, and subsequently was trained in systems electrophysiology at the University of Zurich (1995–1998) and the California Institute of Technology (1998–2003) before leading a research group at the Institute of Neuroinformatics at Zurich University and ETH (2004–2009). His research focuses on neural coding and decoding of hand movements, their interactions with sensory systems, and he develops brain-machine interfaces to read out movement intentions for the development of neural prosthetics to restore hand function in paralyzed patients.

Alexandra Stein
Institute of Neuroscience and Medicine Computational and Systems Neuroscience (INM-6), Research Center Jülich, Jülich, Germany
Bernstein Coordination Site, Bernstein Network Computational Neuroscience, Albert-Ludwigs Universität Freiburg, Hansastraße 9a, 79104 Freiburg, Germany
a.stein@fz-juelich.de

Alexandra Stein heads the Bernstein Coordination Site (BCOS), the central coordination of the Bernstein Network Computational Neuroscience. She studied Biology at LMU Munich and continued with her doctoral research in the field of sensory neuroscience. After a postdoc at the Bernstein Center Computational Neuroscience in Munich, she continued a career in science management under the upwind of the first Excellence Initiative. After coordinating the Graduate School of Systemic Neuroscience (GSN-LMU) for several years she turned to the Graduate Center, the central coordination, advice and service unit for doctoral studies at LMU Munich. Since 2017, she heads the Bernstein Coordination Site in Freiburg.

Otto W. Witte
Hans Berger Department of Neurology, Jena University Hospital, Am Klinikum 1, 07747 Jena, Germany
Deutsche Gesellschaft für Klinische Neurophysiologie und Funktionelle Bildgebung (DGKN), Salvador-Allende-Platz 29, 07747 Jena, Germany
otto.witte@med.uni-jena.de
https://orcid.org/0000-0003-2101-4105

Otto W. Witte studied medicine, psychology and mathematics in Münster and London. He worked as a postdoc in Neurophysiology in Münster with E.-J. Speckmann before he moved to Düsseldorf where he received his neurology education with H.-J. Freund. Since 2001, he is director of the Hans Berger Department of Neurology in Jena. His scientific interests include brain plasticity and brain aging, as well as experimental and clinical brain imaging. As the secretary of the DGKN, he heads the office of the society which supports research and innovation in the exciting field of clinical neurophysiology and functional brain imaging, and is engaged in establishing standard procedures and quality control measures.

DE GRUYTER Neuroforum 2021; 27(1): 17–25

Review article

Michael Hanke*, Franco Pestilli, Adina S. Wagner, Christopher J. Markiewicz, Jean-Baptiste Poline and Yaroslav O. Halchenko

In defense of decentralized research data management

https://doi.org/10.1515/nf-2020-0037

Abstract: Decentralized research data management (dRDM) systems handle digital research objects across participating nodes without critically relying on central services. We present four perspectives in defense of dRDM, illustrating that, in contrast to centralized or federated research data management solutions, a dRDM system based on heterogeneous but interoperable components can offer a sustainable, resilient, inclusive, and adaptive infrastructure for scientific stakeholders: An individual scientist or laboratory, a research institute, a domain data archive or cloud computing platform, and a collaborative multisite consortium. All perspectives share the use of a common, self-contained, portable data structure as an abstraction from current technology and service choices. In conjunction, the four perspectives review how varying requirements of independent scientific stakeholders can be addressed by a scalable, uniform dRDM solution and present a working system as an exemplary implementation.

Keywords: BrainLife; Canadian Open Neuroscience Platform; DataLad; Interoperability; OpenNeuro.

Zusammenfassung: Dezentrale Forschungsdatenmanagement (dFDM) Systeme verwalten digitale Forschungsdaten mit mehreren Teilnehmern, ohne dabei von einem zentralen Service abhängig zu sein. Zur Verteidigung von dFDM präsentieren wir vier Perspektiven: Einzelne Wissenschaftler, Institutionen, Datenarchive, Analyse-Plattformen und Konsortien, die zeigen, dass heterogene, aber auf interoperablen Komponenten basierende dFDM Systeme, im Gegensatz zu zentralisierten oder föderierten Lösungen, eine nachhaltige, resiliente, offene und anpassungsfähige Infrastruktur für wissenschaftliche Interessensgruppen sein können. Allen ist die Verwendung einer einheitlichen, portablen Datenstruktur gemein, die als Abstraktion von aktuell verwendeten Technologien zum Einsatz kommt. Zusammengenommen zeigen diese Perspektiven beispielhaft anhand eines in der Praxis verwendeten Systems, wie vielfältige Anforderungen unterschiedlicher Interessengruppen durch eine skalierbare dFDM Lösung adressiert werden können.

Schlüsselwörter: BrainLife; Canadian Open Neuroscience Platform; DataLad; Interoperabilität; OpenNeuro.

Introduction

Research data management (RDM) is an increasingly important topic for individual scientists, institutions, infrastructure providers, and large-scale research collaborators. This shift in attention is driven by ethical considerations, threats to the trustworthiness of research outputs, and the desire to maximize the impact of publicly funded research. Generic, large-scale storage and computing infrastructure has existed internationally for a considerable time. Yet, the

---

***Corresponding author: Michael Hanke,** Institute of Neuroscience and Medicine Brain & Behavior (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52425 Jülich, Germany; and Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University, 40225 Düsseldorf, Germany, E-mail: michael.hanke@gmail.com. https://orcid.org/0000-0001-6398-6370
Franco Pestilli, Department of Psychology, The University of Texas at Austin, 108 E Dean Keeton St, Austin, TX 78712, TX, USA, E-mail: pestilli@utexas.edu. https://orcid.org/0000-0002-2469-0494
Adina S. Wagner, Institute of Neuroscience and Medicine Brain & Behavior (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52425 Jülich, Germany, E-mail: adina.wagner@t-online.de. https://orcid.org/0000-0003-2917-3450
Christopher J. Markiewicz, Department of Psychology, Stanford University, 450 Jane Stanford Way, Building 420, Stanford, CA 94305, CA, USA, E-mail: markiewicz@stanford.edu. https://orcid.org/0000-0002-6533-164X
Jean-Baptiste Poline, McConnell Brain Imaging Centre, Faculty of Medicine, McGill University, 3801 University Street, Montreal, Quebec, H3A 2B4, Canada, E-mail: jean-baptiste.poline@mcgill.ca. https://orcid.org/0000-0002-9794-749X
Yaroslav O. Halchenko, Department of Psychological and Brain Sciences, Dartmouth College, 419 Moore Hall, Hinman Box 6207, Hanover, NH 03755, NH, USA, E-mail: yaroslav.o.halchenko@onerussian.com. https://orcid.org/0000-0003-3456-2493

apparent lack of fit for domain-specific or regionalized data exchange and publication use cases has motivated a large number of localized, domain-specific developments or deployments of RDM solutions. These emerging solutions address some of the immediate needs, in part motivated by the increasing enforcement of minimum RDM standards by funding agencies. Yet as of today, the lack of infrastructure allowing interoperability across RDM systems still limits the potential impact that the research data can make to science and society.

This problem can be addressed by establishing a network of interoperable but independently governed and funded services that jointly form a decentralized research data management system (dRDM). Such a system makes digital research objects available across a network of participating institutions and investigators for publication, query, retrieval, backup or archive, and collaborative evolution. Importantly, this is achieved without critically relying on central services, thereby offering a high level of resilience against any failure of individual network components, including technical errors, but also institutional failure like discontinued funding.

Two primary models of decentralization can be distinguished: (1) A *federation*, where a single technology is utilized across partner sites, to provide a homogeneous solution, and (2) *interoperability*, where multiple technologies are used across partner sites but integrated into a single but heterogeneous set of components. On the one hand, the federation model dramatically simplifies the technical challenges. Simplicity comes at a cost though, as it constrains all partner sites to the deployment and maintenance of a single (homogeneous) software solution that might be suboptimal for many partners; a "one-size-must-fit-all" problem that can limit the type of partners involved in the federation. On the other hand, the interoperability model allows decentralization based on a network of heterogeneous software solutions. Each participant site is free to employ the optimal, site-specific solution avoiding the challenges and limitations of a "one-size-must-fit-all" approach. Though in such a system the challenge is shifted to establishing effective interoperability between the different technologies employed.

Arguably, the interoperability model is more flexible and inclusive as it allows a more diverse set of partner sites to participate. More importantly, the interoperability model can improve the widespread application and resilience of dRDM. For example, established analysis and deployment workflows at each site can stay working, while interoperability with other sites can be established in parallel, for those projects requiring it, rather than requiring disruptive infrastructural changes that can simultaneously impact multiple laboratories or researchers. In the following, we present four perspectives on the utility of this type of dRDM. All four share a common principle: the use of a uniform data structure as a common denominator that facilitates independent development of software adapters to instruments and services that enable interoperability and data flow between all relevant infrastructure components and participants. While various standards and implementations of such data structures exist (e.g. BagIt, Kunze et al., 2018; Frictionless Data Package, Walsh et al., 2017; or Dat, McKelvey et al., 2020), all presented perspectives share the use of DataLad's datasets (Hanke et al., 2020) as key technology choices. This particular implementation is a domain-agnostic lightweight data structure that offers joint version control capabilities for code and data (based on the industry standard Git, git-scm.com), supports arbitrarily structured metadata, and is capable of tracking the identity and availability of dataset components via the git-annex software (Hess, 2020) without requiring universal data access or actually containing the file content. This makes it possible to construct a dataset as a standardized overlay data structure which references content in heterogeneously organized data portals or databases. Moreover, it does not hide or bypass existing institutional access protection mechanisms and leaves authorization procedures in the responsibility of the data owners (see Figure 1).

dRDM perspective: one laboratory or researcher

From the perspective of individual researchers, their laboratories, and collaborators, dRDM can improve day-to-day operations and make them robust against disruptive infrastructural changes. If data are uniformly accessible regardless of their storage location, scientists can orchestrate collaborative workflows and access not only to the data collected locally but also from external (public) resources in a streamlined fashion. Moreover, researchers utilizing a dRDM model can ensure consistent and robust data management across local and institutional information technology (IT) environments. For example dRDM makes it trivial to deploy a processing script from a local copy of data within the laboratory to a larger scale version of the data hosted in a datacenter. And as most researchers, in particular at early career stages, frequently move their workplace to different institutions (Guthrie et al., 2017), the benefits of this feature extend beyond a single workplace. When research agendas comprise a longer time frame, such that an employment change does not necessarily imply a fresh start and the discontinuation of previous projects, the potentially substantial and disruptive transition to a new institution and IT environment can be alleviated or prevented by a dRDM-based system.

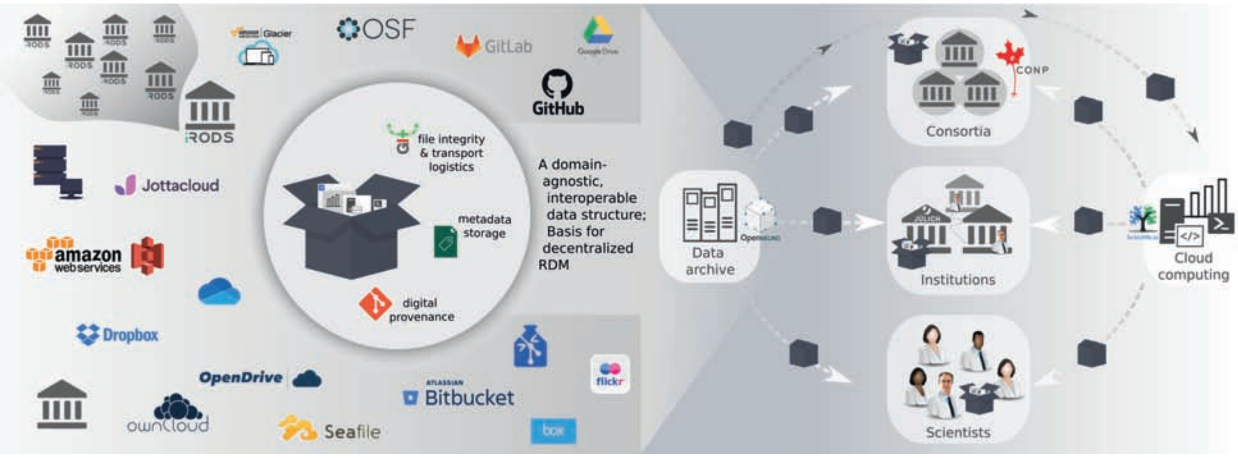


Figure 1: A common, portable data structure allows establishing interoperability between diverse participant sites.
Left: A common data structure can serve as a uniform abstraction layer to interface any number of commercial or institutional storage services, which may be centralized or federated systems. *Right:* The portable nature of the data structure facilitates data exchange between archive and compute services, as well as collaboration among individual researchers or formal consortia. Moreover, it provides institutions with the flexibility to evolve their infrastructure without needlessly impacting scientific workflows.

Without dRDM, and depending on the magnitude of the differences between IT systems and policies, the necessary changes can be severe. Consider, for example, a transition from an environment with ample storage and shared computing resources, to a workplace with minimal local resources, but an institutional cloud storage service account. Before, all data holdings were accessible with low latency as if stored on a single big hard drive. Computing resources had direct data access, and analysis scripts could reference the desired data by (hardcoded) paths. After the transition, scripts cease to work because there is no local storage resource large enough to hold all data for analysis. Instead, additional, service-specific software has to be used to pull required data from the cloud and deposit results into the cloud. Essentially all analysis implementations of the past have to be manually adjusted to work in the new environment, an error-prone process that in itself is a threat to the reproducibility of results.

Using a common data structure as an abstraction of an analysis environment has the potential to substantially ease such transitions. In the case of a DataLad dataset, it is possible to comprehensively include all components of a compute- or data-intensive analysis in a single, version-controlled unit. This includes input data of any number and size, analysis code in any programming language, and even complete computational environments in the form of software container images. The dataset offers an intuitive application programming interface (API) for data access that hides the peculiarities of a particular IT environment and enables the development of analysis codes with improved portability properties. For example, a particular input file for an analysis can be referenced using a simple local path, relative to the root path of the analysis dataset: *input/datasetA/file1.dat*. An analysis script that requires this file can ensure this by executing the shell command *datalad get input/datasetA/file1.dat*. Importantly, the analysis script does not need to reflect that datasetA, which contains the file of interest, is a different modular data unit that is presently hosted on a particular storage service. Consequently, the analysis script does not need to be adjusted whenever the availability of datasetA changes because it has been transferred to a different institution. Instead, the DataLad software can be centrally configured to look for datasets, identified by a globally unique identifier and a precise version, at a different or additional location. Given that the data structure also allows for change tracking, it is possible to retrospectively discover how data were manipulated, improving the transparency and reproducibility of conducted projects.

For an individual researcher or laboratory, the barrier of entry into such a system is low. With no confinement to external services or file types, a scientist can transition new or existing projects into a common data structure independently and can typically achieve this without assistance, additional infrastructure, or project structure change. Nevertheless, the adoption of a common data structure such as DataLad's datasets implies the necessity to acquire additional expertise, e.g. from documentation, user training, or tutorials, and also an individual's interest in doing so. Efforts such as ReproNim's (repronim.org) webinars, teaching resource collections, and teaching fellowships, or in-depth, user-focused documentation formats such as the DataLad Handbook (Wagner et al., 2020) facilitate this.

dRDM perspective: a research institute

Like individual laboratories or researchers, research institutes also exist in a volatile environment. It is in their best interest to provide their scientists with the latest technologies to maximize their competitive advantage, boost research efficiency, and consequently increase the attractiveness and reputation of a research environment. However, the desire to quickly adopt new technologies has to be counterbalanced with the need to keep the cumulative cost of legacy infrastructure and procedures at a manageable level. This is compounded by the fact that institutions are generally responsible for guaranteeing a certain level of longevity for all research outputs, for example, the retention of research data, typically for at least a decade.

For the same reason as for individual researchers or laboratories, readiness for future infrastructure transitions, it makes sense for research institutions to utilize a portable, common data structure as an abstraction layer for RDM operations. The key feature of data structures, like DataLad's datasets, is that they present researchers with a familiar view, a project directory on a filesystem, and internally translate *requests for data by location* (i.e. a file path) into *requests for data by identity* (i.e. a UUID or a checksum). This represents a powerful paradigm shift, as it enables future modifications of the content lookup and retrieval without changing the user/research-facing data representation.

The Institute of Neuroscience and Medicine Brain & Behaviour (INM-7) of the Research Center Jülich uses DataLad datasets not only to manage access to large-scale neuroimaging datasets, like the UKBiobank (Miller et al., 2016), or the Human Connectome Project (HCP, van Essen et al., 2013), but also as a system to archive completed projects. Institute members can discover all managed datasets via a collection that is maintained as a DataLad superdataset (a dataset comprising a versioned collection of datasets) hosted on a local GitLab (gitlab.com) instance. Independent of the hosting choice of the original data provider, institute members can access any data file by requesting it through the institute's dataset collection, as described above. File access permissions are managed either directly by the respective data owners (e.g. each HCP user obtains their own credentials from the HCP consortium) or by controlled access to local downloads of restricted datasets (e.g. dedicated access group for signatories of the UKBiobank data usage agreement). Importantly, data access procedures remain uniform and fine-grained, regardless of whether an analysis is developed on a student's laptop or is computed on the institute's cluster system. This RDM setup also facilitates the *ad hoc* usage of resources at the Jülich Supercomputing Center (JSC). Institute staff can stage individual data resources on the JSC storage systems, and the DataLad software can transparently obtain dataset content on this independently operated resource without requiring individual adjustments of datasets, or analysis scripts. When a study is completed and archived, its DataLad dataset, including the incorporated study metadata, remains fully discoverable and accessible through the institute's dataset collection. However, file content can be administratively moved from fast and expensive "hot" storage to higher latency bulk storage, and eventually onto tape backup systems, all without structurally impacting dataset access for institute members. Combined with data access statistics, this flexibility allows institute staff to maintain an optimal compromise of data access latency and storage demands without individual user negotiations.

dRDM perspective: a domain data archive or computing platform

Domain data archives seek to provide high-reliability datasets access to all authorized researchers, with a secondary mandate to ensure that publicly funded data are findable via internal search or external indexing. Archives treat datasets as a natural unit of organization, and the necessary considerations are ingress, validation and metadata extraction, storage, publication, and egress. By adopting common data standards coupled with ingress and egress validation mechanisms, an archive team can focus development efforts on the key tasks of ensuring data access, availability, and findability.

For example, OpenNeuro (Gorgolewski et al., 2017) is a public neuroimaging data repository. Rather than imposing its own schema to which submitters must adapt their data, the archive adopted the community-developed Brain Imaging Data Structure (BIDS) standard for data organization and metadata (Gorgolewski et al., 2016). To assure reliable data access, and to serve the wide community of users, the archive relies on commercial infrastructure and uses Amazon Web Services to host the web interface and the Simple Storage Service (S3) to host the data. However, to ensure the long-term availability of the data, it requires a data model that is not tied to any specific vendor, hosting platform, or technology. In addition to the data model, OpenNeuro also desired making data available through generalized, stable interfaces independent of a particular storage platform or vendor. Consequently, the archive adopted DataLad to represent datasets internally (within the archive). This choice enables data change tracking and a common protocol for data egress (i.e. Git combined with git-annex). Data ingestion is also facilitated by DataLad. When a dataset is submitted to the archive, a DataLad dataset is created and binary files with imaging data are annexed. The dataset owner makes at least one “snapshot” to mark the dataset as complete and then publishes it in the archive. When the dataset is published, all files are uploaded to S3, and the URLs provided by S3 are associated with the annexed files. Finally, the DataLad dataset is published to a GitHub repository, to allow findability by other researchers even beyond the OpenNeuro Archive. The use of high-availability, permissive, third-party services ensures data are accessible even if the primary website suffers from downtime. At the same time, the data model does not depend on either service and can be ported to other services as new technologies emerge.

Version control and persistent identifiers are central features of the OpenNeuro data model. Datasets may change over time as new data are added or metadata is updated, and analyses of a dataset depend critically on the state of the dataset at the time of analysis. Dataset snapshots are represented as Git tags, allowing analyses to refer to the version of the dataset used via its version number (as opposed to by checksum). In addition, data object identifiers (DOIs) are issued for each snapshot of the dataset, ensuring that the particular version of the dataset may be cited in publications and facilitate the reproduction of analyses.

The use of DataLad and the published datasets on GitHub allows OpenNeuro datasets to be available beyond the archive. A variety of computational systems even without direct interaction with OpenNeuro can reference and access the datasets. For example, a researcher interested in developing a new analysis method might test the code during development on their personal computer by fetching an OpenNeuro dataset for testing or validation. The same researcher can then run a scaled-up version of the analysis on a high-performance computing cluster, which may host OpenNeuro datasets in a centralized location within a datacenter with minimal effort, simply reusing the data model and DataLad version tracking mechanisms. Finally, a cloud-based computational platform may expose OpenNeuro datasets to its users to increase data availability and enhance the general utility of the services offered.

As datasets are published and accumulate in one or several accessible repositories, new opportunities emerge for data aggregation and reuse across datasets (Avesani et al., 2019). Common metadata standards are essential to effectively harmonize data from multiple sources and enable research questions at scales previously impracticable. Furthermore, a common data standard can facilitate the aggregation of data from multiple sources. The effective separation of metadata (Git) and data (git-annex) is a key feature of the DataLad model that ensures that the metadata can be made accessible even when there are legal and ethical barriers to openly sharing data. It is thus becoming possible to develop tools to aggregate data from multiple providers without requiring an explicit effort from those providers. The dRDM model breaks some of the barriers and facilitates aggregation, curation, and upcycling data, allowing central archives such as OpenNeuro to act as stewards rather than gatekeepers.

Key partners that can be effectively served by the proposed dRDM model are cloud computing platforms. BrainLife (brainlife.io) is one of the most recent open and publicly funded platforms developed with the goal to serve researchers facilitating access, sharing, or reuse of data processing methods. The code implementing the data processing method can be submitted to BrainLife and registered as a web service (an App). The BrainLife platform allows automated tracking of the analyses execution and orchestrates data processing on diverse compute resources via a convenient graphical web interface or command line interfaces. BrainLife is not meant to be a data archive but a registry for reusable processing methods used in published scientific articles. The computational platform is compliant with the BIDS data standard so as to facilitate users’ data ingress and egress. Recently, the BrainLife team has used DataLad to connect the platform users with hundreds of BIDS-compliant datasets that are made publicly available as DataLad datasets. BrainLife uses DataLad to offer automated import “with the push of a button” of datasets that users have published on a variety of public archives. BrainLife benefits from the dRDM standardization in two ways: (1) Metadata standardization enables automatic identification of relevant dataset components, extraction of key data properties, and match-making of applicable analysis implementation against available data types, and (2) the abstraction of data transport logistics provided by DataLad’s datasets enables BrainLife to automatically obtain (pull) data files from the original providers, for example, from OpenNeuro, avoiding manual access to each data archive. Taken together, BrainLife is an example of a highly accessible computing platform that translates the potential of a dRDM system to the immediate computing needs of researchers, by connecting to independent standardization efforts without suffering from the need to continuously adjust to implementation changes in a large number of data portal and metadata access APIs.

dRDM perspective: a collaborative multisite consortium, the Canadian Open Neuroscience Platform

The need for data sharing across institutions and states is fueled by the requirement of large sample sizes to enable well-powered and generalizable studies and for distributing the cost of data acquisition across sites. These large consortia generally opt for centralized data hosting, which simplifies data harmonization and management. However, large numbers can also be achieved through many independently acquired datasets that have the potential to better represent a more diverse population, an important factor for the construction of biomarkers. The Canadian Open Neuroscience Platform (CONP) is a consortium aiming for this goal and was funded in part to share neuroscience datasets across Canada within a comprehensive ethical and legal framework, establishing a repository of data implementing the Findable, Accessible, Interoperable, Reusable (FAIR) principles (Wilkinson et al., 2016).

While the central CONP data portal (portal.conp.ca) could have been only a set of links pointing to original infrastructures, this would not have given direct data access across datasets and would have been of limited utility for information aggregation. On the other extreme, centralizing data would have been infeasible. Critically, ethical or institutional policy requirements would have prevented transferring data to a central data storage for a number of datasets that are presently accessible on the platform. To keep the governance of datasets local, the CONP needed to adopt a distributed solution, while still making the data accessible directly through a single portal.

Adopting a portable, common data structure, like DataLad’s dataset, as an abstraction provided the CONP a shared and centralized space for distributing the metadata, while keeping the links to the original data locations. Metadata descriptors implemented using the DAta Tag Suite (DATS) model (Sansone et al., 2017) are incorporated in the centrally hosted dataset Git repositories, while original raw data are hosted on diverse platforms (OSF.io, Zenodo.org, Loris.ca, Braincode.ca, and others). The CONP uses a crawler to discover datasets on external services, like OSF or Zenodo, and builds a minimal DATS model for each dataset to make these data findable and accessible through the CONP portal. This offers a simple procedure for researchers who both want to share data in a general repository but also make these data discoverable in a neuroscience specialized portal.

Presently, CONP users must access datasets exclusively using the DataLad software. This imposes requirements, such as the necessity to deploy the software for any consumer. However, not all data consumption scenarios require that each participant operates a full-featured node of the dRDM system. Consequently, the CONP is working on convenient export functionality, such as an in-browser dataset downloader, to lower the technical threshold for interaction with its users. Because such a solution relies on standardized data access records, it can also be used by any other project using the data structure for dRDM.

Conclusions

As illustrated by the four perspectives presented here, dRDM, built on a common, portable data structure that enables uniform access to all relevant commercial and institutional data services, is a flexible model that can scale from personal computing environments to individual institutions, all the way to large-scale collaborations in multisite consortia. The inclusive nature of this RDM approach that avoids one-size-must-fit-all prescription of centralized or federated services is suitable for introducing RDM standards and procedures in heterogeneous fields of endeavor. Consequently, it has also been selected as a strategic component of the NFDI Neuroscience initiative, a consortium that aims to consolidate neuroscience RDM in Germany along these lines.

Using the DataLad software and its datasets as an exemplary implementation of a common portable data structure, it is possible to curate and maintain unified data distributions collating data from the wide range of data providers. One such distribution is datasets.datalad.org, which currently provides a single point of entry for public or authenticated access to over 5,000 DataLad datasets covering over 200 TBs of neuroscience research data from hundreds of archives, initiatives, or individual laboratories. Among others, this collection also includes the superdatasets for CONP and OpenNeuro and through them provides access to all datasets managed by the respective entities. In turn, this collection is used by BrainLife to automatically discover datasets that can be processed on its platform.

Standardizing on a technology implies a substantial risk and installs a single point of failure in a complex system. However, standardization of core components also limits the variability that subsequent developments need to consider and ultimately enables more progress to be made with the same finite resources. In the case of DataLad, risks are introduced by three components: two small-scale developments (DataLad, git-annex) and the version control system Git. Git is a globally adopted industry standard. The chance of a technology failure without an adequate mitigation opportunity can be considered minimal. Both DataLad and git-annex build on Git, adding only documented, plain-text data structures to the content managed by Git. In the case of catastrophic failure (discontinuation of the development), the interpretability of data contained in these structures is unimpaired. Moreover, both software components are openly developed (public code history, issue tracker, support channels) and are available under recognized free software licenses (MIT, Affero GPL), such that continued maintenance by a third party can be considered feasible. This use of general-purpose protocols and technologies makes it possible to present scientific data in a readily usable form on platforms and forums, such as GitHub, that are used by a large audience of nonresearchers, thereby dramatically increasing the exposure of publicly funded research output, and successfully utilizes them for improving the capabilities and resilience of global dRDM.

Acknowledgements: A.S.W., F.P., M.H., and Y.O.H. were, in part, supported by a US-German CRCNS project, co-funded by the BMBF and the US National Science Foundation (BMBF 01GQ1905; NSF 1912266). A.S.W. and M.H. were supported by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826421 (VirtualBrainCloud) and 945539 (HBP SGA3). A.S.W. is supported by a ReproNim/INCF Training Fellowship. NSF OAC-1916518, NSF IIS-1912270, NSF IIS-1636893, NSF BCS-1734853, Microsoft Faculty Fellowship supported F.P. C.J.M. is supported by NIH R24-MH117179. Y.O.H. was also, in part, supported by NIH 1P41EB019936-01A1. J-.B.P. was partially funded by the National Institutes of Health (NIH) NIH-NIBIB P41 EB019936 (ReproNim), NIH-NIMH R01 MH083320 (CANDIShare), and NIH RF1 MH120021 (NIDM), the National Institute Of Mental Health under Award Number R01MH096906 (Neurosynth), as well as the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives initiative and the Brain Canada Foundation with support from Health Canada.

Glossary

API: An application programming interface (API) defines interactions between multiple software intermediaries. An API can be entirely custom, specific to a component, or it can be designed based on an industry standard to ensure interoperability.

Checksum: A checksum is a small-sized datum derived from a block of digital data for the purpose of detecting errors that may have been introduced during its transmission or storage.

UUID: A universally unique identifier (UUID) is a 128-bit number used to identify information in computer systems.

Version control: Version control (also known as revision control) is a class of systems responsible for managing changes to computer programs, documents, or other collections of information.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This article was funded by National Science Foundation (BMBF 01GQ1905; NSF 1912266), European Union's Horizon 2020 under grant agreement no. 826421 (VirtualBrainCloud), and 945539 (HBP SGA3), Microsoft Faculty Fellowship, National Institutes of Health, National Institute of Mental Health under Award Number R01MH096906, Canada First Research Excellence Fund, Health Canada, Bundesministerium für Bildung und Forschung, and National Institute of Biomedical Imaging and Bioengineering.

Conflict of interest statement: The authors declare no conflicts of interest.

References

Avesani, P., McPherson, B., Hayashi, S., Caiafa, C.F., Henschel, R., Garyfallidis, E., Kitchell, L., Bullock, D., Patterson, A., Olivetti, E., et al. (2019). The open diffusion data derivatives, brain data upcycling via integrated publishing of derivatives and reproducible open cloud services. Sci. Data *6*, 1–13.

Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci. Data, *3*, 1–9.

Gorgolewski, K., Esteban, O., Schaefer, G., Wandell, B., and Poldrack, R. (2017). OpenNeuro—A free online platform for sharing and analysis of neuroimaging data. Organ. Hum. Brain Mapp. Vancouver, Canada *1677*.

Guthrie S., Lichten C., Harte E., Parks S., Wooding S. 2017. International Mobility of Researchers (Cambridge, UK: RAND EUROPE).

Hanke, M., Halchenko, Y.O., Poldrack, B., Meyer, K., Solanky, D.S., Alteva, G., Gors, J., MacFarlane, D., Häusler, C.O., Olson, T., et al. (2020). DataLad/DataLad: 0.13.5 (October 30, 2020, Version 0.13.5). Zenodo, https://doi.org/10.5281/zenodo.4161813.

Hess, J. (2020). Git-annex, https://git-annex.branchable.com.

Kunze, J., Littman, J., Madden, E., Scancella, J., and Adams, C. (2018). The bagIt file packaging format (V1.0). RFC 8493 RFC Editor, https://doi.org/10.17487/rfc8493.

McKelvey, K. and Buus, M. (2020). Dat protocol foundation, https://dat.foundation.

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., et al. (2016). Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. *19*, 1523–1536.

Sansone, S.A., Gonzalez-Beltran, A., Rocca-Serra, P., Alter, G., Grethe, J.S., Xu, H., Fore, I.M., Lyle, J., Gururaj, A.E., Chen, X., et al. (2017). DATS, the data tag suite to enable discoverability of datasets. Sci. Data *4*, 170059.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., and Wu-Minn HCP Consortium. (2013). The Wu-Minn human connectome project: An overview. Neuroimage *80*, 62–79.

Wagner, A.S., Waite, L.K., Meyer, K., Heckner, M.K., Kadelka, T., Reuter, N., Waite, A.Q., Poldrack, B., Markiewicz, C.J., Halchenko, Y.O., et al. (2020). The DataLad Handbook (Version v0.13). Zenodo, https://doi.org/10.5281/zenodo.3905791.

Walsh, P. and Pollock, R. (2017). Frictionless data package, https://specs.frictionlessdata.io/data-package.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data *3*, 1–9.

Bionotes

Michael Hanke
Institute of Neuroscience and Medicine Brain & Behavior (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52425 Jülich, Germany
Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University, 40225 Düsseldorf, Germany
michael.hanke@gmail.com
https://orcid.org/0000-0001-6398-6370

Michael Hanke studied psychology in Halle (Saale) and Magdeburg, Germany. After being a postdoc in the laboratory of James Haxby at Dartmouth College, together with Yaroslav O.Halchenko working on NeuroDebian and PyMVPA, he moved back to Magdeburg as a junior professor. Since 2019, he is a professor at Heinrich Heine University Düsseldorf and head of the Psychoinformatics Lab at the Institute of Neuroscience and Medicine (Brain and Behavior) of the Research Center Jülich. His group contributes to the development of the DataLad software and develops workflows and training materials for research data management in neuroscience.

Franco Pestilli
Department of Psychology, The University of Texas at Austin, 108 E Dean Keeton St, Austin, TX 78712, TX, USA
pestilli@utexas.edu
https://orcid.org/0000-0002-2469-0494

Franco Pestilli is an associate professor in the Department of Psychology at the University of Texas, Austin. Prior to that, he was an associate professor at Indiana University. Dr. Pestilli holds a Ph.D. from New York University and a B.A. from the University of Rome La Sapienza and received Postdoctoral Training at Stanford University and Columbia University. Dr. Pestilli is the author of over 60 publications spanning multiple fields of science, such as cognitive and computational neuroscience, vision science, neuroanatomy, computer science, and neuroinformatics. Dr. Pestilli's scientific projects have been funded by the National Science Foundation, the National Institute of Health, Department of Defense, Association for Psychological Science, and Microsoft Research. Dr. Pestilli is an elected Fellow of the Association for Psychological Science and Psychonomics Society. He was awarded a Microsoft Faculty Fellowship in 2019 and the Janet Taylor Spence Award for Transformative Early Career Achievements by the Association for Psychological Science in 2016. He is an editorial board member for Scientific Data, Scientific Reports, and Brain Structure and Function. Dr. Pestilli is a principal investigator for the Midwest Big Data Hub, director of the Advanced Computational Neuroscience Network, and founder of the open science platform brainlife.io.

Adina S. Wagner
Institute of Neuroscience and Medicine Brain & Behavior (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52425 Jülich, Germany
adina.wagner@t-online.de
https://orcid.org/0000-0003-2917-3450

Adina S. Wagner is a doctoral researcher in the Psychoinformatics Lab at Research Center Jülich and lead of the DataLad Handbook documentation project. She holds a ReproNim/INCF fellowship and teaches workshops on research data management and reproducible open science.

Christopher J. Markiewicz
Department of Psychology, Stanford University, 450 Jane Stanford Way, Building 420, Stanford, CA 94305, CA, USA
markiewicz@stanford.edu
https://orcid.org/0000-0002-6533-164X

Christopher J. Markiewicz is a software developer at the Center for Reproducible Neuroscience at Stanford University and the technical lead for the OpenNeuro database. He holds a Ph.D. in Cognitive and Neural Systems from Boston University and is a core developer for open source neuroimaging tools including NiBabel, Nipype, PyBIDS, fMRIPrep, and FitLins. He is also a maintainer for the Brain Imaging Data Structure (BIDS) standard.

Jean-Baptiste Poline
McConnell Brain Imaging Centre, Faculty of Medicine, McGill University, 3801 University Street, Montreal, Quebec, H3A 2B4, Canada
jean-baptiste.poline@mcgill.ca
https://orcid.org/0000-0002-9794-749X

Since 1990, Jean-Baptiste Poline has worked on the development of methods for the analysis of functional imaging data (mostly fMRI) and more specifically in the statistical modeling and inference aspects. He was a postdoctoral fellow in London (1994–1997) working on SPM and data analysis. He was twice elected as the educational chair for the Organization for Human Brain Mapping (OHBM). He joined the SHFJ-Orsay and then NeuroSpin in 1998 where he has developed an interest in imaging genetics and was responsible for a large multicentric neuroimaging genetic database. He moved to UC Berkeley in 2011 and then to McGill as an associate professor in 2017. He chairs the Council for Training, Science and Infrastructure at the International Neuroinformatics Coordinating Facility. He co-founded Frontiers in Brain Imaging Methods in 2012 partly to address the need for reproducibility and the development of methods in the brain imaging field and recently developed the OHBM Aperture publishing project.

Yaroslav O. Halchenko
Department of Psychological and Brain Sciences, Dartmouth College, 419 Moore Hall, Hinman Box 6207, Hanover, NH 03755, NH, USA
yaroslav.o.halchenko@onerussian.com
https://orcid.org/0000-0003-3456-2493

Yaroslav O. Halchenko studied optoelectronic engineering at VSTU (Vinnitsa, Ukraine) and then obtained Masters in Computer Science at UNM (Albuquerque, NM, USA) and Ph.D. at NJIT (Newark, NJ, USA). In 2009, he joined PBS Department at Dartmouth College as a postdoc in the laboratory of James Haxby at Dartmouth College, to work together with Michael Hanke on NeuroDebian and PyMVPA. After the expected but still unfortunate departure of Dr.Hanke back to Germany, Yaroslav continued his work at Dartmouth. In 2018, he accepted a research associate professor position and established the Center for Open Neuroscience (CON) at PBS Department. CON develops new and contributes to many ongoing projects and initiatives to make neuroscience research more open, robust, efficient, and trustworthy. The CON contributes to the development and dissemination of the DataLad and many related software, data sharing, and scholastic resources.

DE GRUYTER Neuroforum 2021; 27(1): 27–34

Review article

Michael Denker, Sonja Grün, Thomas Wachtler and Hansjörg Scherberger*

Reproducibility and efficiency in handling complex neurophysiological data

https://doi.org/10.1515/nf-2020-0041

Abstract: Preparing a neurophysiological data set with the aim of sharing and publishing is hard. Many of the available tools and services to provide a smooth workflow for data publication are still in their maturing stages and not well integrated. Also, best practices and concrete examples of how to create a rigorous and complete package of an electrophysiology experiment are still lacking. Given the heterogeneity of the field, such unifying guidelines and processes can only be formulated together as a community effort. One of the goals of the NFDI-Neuro consortium initiative is to build such a community for systems and behavioral neuroscience. NFDI-Neuro aims to address the needs of the community to make data management easier and to tackle these challenges in collaboration with various international initiatives (e.g., INCF, EBRAINS). This will give scientists the opportunity to spend more time analyzing the wealth of electrophysiological data they leverage, rather than dealing with data formats and data integrity.

Keywords: FAIR; NFDI; open data; research data management; systems neuroscience.

Zusammenfassung: Die Aufbereitung eines neurophysiologischen Datensatzes mit dem Ziel, ihn zu teilen und zu veröffentlichen, ist schwierig. Viele der verfügbaren Werkzeuge und Dienste für einen reibungslosen Ablauf einer Datenpublikation sind noch im Entstehen und nicht gut integriert. Außerdem fehlen Handlungsempfehlungen und konkrete Beispiele für die Publikation eines vollständigen Datensatzes aus elektrophysiologischen Experimenten. Angesichts der Heterogenität des Feldes können solche einheitlichen Richtlinien und Prozesse nur gemeinschaftlich formuliert werden. Eines der Ziele der Konsortiumsinitiative NFDI-Neuro ist es, für die System- und Verhaltensneurowissenschaften eine solche Gemeinschaft aufzubauen. NFDI-Neuro will die Bedürfnisse dieser Community für ein verbessertes Datenmanagement aufgreifen und in Zusammenarbeit mit verschiedenen internationalen Initiativen (z.B. INCF, EBRAINS) angehen und lösen. Hierdurch bleibt den Wissenschaftlern in Zukunft mehr Zeit zur Analyze ihrer reichhaltigen elektrophysiologischen Daten, anstatt sich mit Datenformaten und Datenintegrität befassen zu müssen.

Schlüsselwörter: FAIR; NFDI; offene Daten; Forschungsdatenmanagement; Systemneurowissenschaften.

Introduction

Neurophysiology may be considered one of the most common approaches in neuroscience to gain an understanding of the internal processes that underlie neuronal information processing by examining neural activity at various scales of observation. The technique profits from constant technological evolution of recording techniques that enable ever more intricate experimental designs to investigate neuroscientific questions. Although our growing insight into neuronal computation is manifested in the development of increasingly realistic models of brain dynamics through simulation and theory, electrophysiology arguably continues to provide the most valuable experimental counterpart upon which the process of cross-validation is based.

***Corresponding author: Hansjörg Scherberger,** Neurobiology Laboratory, Deutsches Primatenzentrum GmbH, Kellnerweg 4, 37077 Göttingen, Germany; and Department of Biology and Psychology, University of Goettingen, Goettingen, Germany, E-mail: hscherb@gwdg.de. https://orcid.org/0000-0001-6593-2800
Michael Denker, Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany, E-mail: m.denker@fz-juelich.de. https://orcid.org/0000-0003-1255-7300
Sonja Grün, Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany; Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany, E-mail: s.gruen@fz-juelich.de. https://orcid.org/0000-0003-2829-2220
Thomas Wachtler, Department Biologie II, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany, E-mail: wachtler@bio.lmu.de. https://orcid.org/0000-0003-2015-6590

However, the new opportunities offered by the rapid technological and conceptual developments during the last decades do not come for free. The increase in complexity of modern-day experimentation is mirrored in an intricate pool of data, resulting from various hardware and software components built by industrial manufacturers or in-house workshops, and combined in versatile ways to enable novel experimental designs. The amount of experimental skill, time, and creativity that enters such experiments leads to a situation where the protocols conducted in individual labs are to a large extent unique to the respective research group. In order to describe such experiments at a level of being reproducible (Denker and Grün, 2016; Plesser, 2018) and the resulting data becoming practically findable and reusable by other researchers, as stipulated by the FAIR principles (Wilkinson et al., 2016; see also Wachtler et al., this issue), there is a substantial conceptual difficulty in documenting data acquisition and postprocessing at minute levels of detail due the inherent heterogeneity and complexity. Moreover, in the absence of automatization, as typically associated with more standardized processes in science, the costs for a thorough description of the recorded data seem prohibitive. Therefore, the intricate nature of this heterogeneous data bundle, together with the need to integrate all data into a form that is suitable for the anticipated analysis, leads to situations where scientists come up with ad-hoc and highly customized data analysis solutions. This is not only time consuming and inefficient but also error prone and hardly reproducible.

The increased complexity of experimentation is a challenge for organizing the data, and it also amplifies the urgency to conceptualize and implement better solutions of data management. Firstly, while data from electrophysiology have always been considered particularly precious, given the amount of invested resources (time, money, and animal life), the richness and complexity of modern experiments further increase their value by expanding the number of possible scientific questions that can be addressed by a single data set. Indeed, the primary use of the original experiment often leverages only a small part of the full potential of the data. In consequence, data sets may be of relevance for research questions a long time after the original recording and for a large audience, which makes it infeasible for the original experimenters to guide and check that data are handled appropriately. Such scenarios include the notorious case of the PhD student leaving a research group without providing sufficient information about the recorded data and the preprocessing and postprocessing steps applied to them during the PhD project. Another case is the situation where data are analyzed in parallel by a number of laboratories applying different methods and having different research questions with the aim to synergize their findings, like in a collaboration between an experimental and a theoretical lab. For the exchange of results and findings, these partners need to have a consistent description of the data. Even beyond the initial use of a data set within a laboratory and among their collaborators, more and more scientists embrace the idea of making their data publically available, recognizing not only the added attention and appreciation that well-curated data are generating for the experimentalist, but also the increased efficiency for progressing science and the potential to create new research questions.

Despite all of these suggested merits of striving to handle and manage electrophysiology data in a more optimal fashion, we perceive that reality is far from the ideal situation where all of the essential, cumbersome housekeeping of acquired data is automated and the description of the experiment can be saved with the proverbial click of a button. Yet, the topic of data handling is currently erupting in a burst of activity in the field of (neuro-)informatics. Nevertheless, we still observe a prevailing gap between the design of emerging tools, services and processes, and their implementation in concrete experimental settings that are helpful for experimentalists.

The anticipated NFDI-Neuro consortium therefore considers its task to communicate between the world of computer science and neuroscience and to enable the neuroscience community to make better use of the existing tools and services. On the other hand, NFDI-Neuro also supports the neuroscience community to link its established data acquisition and postprocessing workflows to existing tools and to identify lacking tools, processes, and guidelines. The heterogeneity of the data and experimental approaches in the neuroscience community demands to base discussions on concrete examples and build on experiences. NFDI-Neuro considers the establishment of such an exchange to be a primary goal for its activities surrounding electrophysiology.

To stimulate such a discussion, we report here on current challenges, solutions, and shortcomings as we encountered them in our research routines (e.g., Zehl et al., 2016) and during efforts to publish an experimental data set consisting of spiking activity and local field potentials of a macaque monkey performing a motor task (Brochier et al., 2018). We will outline the experiences we had in curating such a data set, ranging from dealing with a collaborative environment consisting of multiple labs, up to an ongoing and dynamic data acquisition process established over the course of years, and the need to publish and easily maintain data in an accessible format.

Considerations when curating electrophysiological data sets

Adoption of any type of data management workflow should ideally disrupt the research workflow only minimally and require only little attention of the researcher. One of the most rewarding strategies is therefore to select data curation procedures that remain constant, consistent, and mostly automated. Given the unpredictable nature of the research process, it may, however, be tempting to design the data curation process "on the fly" as the experiment is being set up. However, anticipating the future use of the data can help guide design decisions early on, which expedite the establishment of a stable data curation pipeline. This includes the consideration of the perhaps most challenging scenario from the start: sharing data with strangers. By design, this approach will help to perceive the process of data acquisition from the perspective of a data consumer (e.g., the remotely analyzing scientist) not the data provider (i.e., the experimenter). Such considerations will address a diversity of issues such as making the data and corresponding metadata available in formats independent of specific programming languages, avoidance of idiosyncratic software codes, and favoring easily comprehensible data descriptions over those present in the original hardware and control software implementations.

For example, if data are provided as "raw" (or primary) data, representing directly the output of the recording setup, users accessing these data need to understand the specific data structure or, equivalently, require specific software for reading the data. While this is conceptually a feasible approach, in practice, it may often fail even at the level of reading the raw binary data since codes for reading the corresponding file formats often differ in the various programming languages and rarely receive professional maintenance and thorough testing. When it comes to interpreting the data contained in these raw files, matters tend to become even more difficult. For example, in typical recordings, raw data contain only certain marked events in time, such as events indicating the start of a certain experimental trial or the time point of a stimulus presentation. Uniquely identifying and describing trials, however, is a long way from these marked events. For example, it may involve the need to interpret the type of a given trial based on subsequent events, in case alternative stimuli, manipulations, or behaviors are possible. It may also involve the interpretation of the performance in a trial based on separate behavioral measurements, e.g., reaction time. Ultimately, each trial must be labeled by an informative identifier that is based on this information and that supports the implementation of the planned analysis of the data. Finally, the outcome of these preparatory steps performed on the primary data (e.g., the resulting trial identification) needs to be completely consistent for any user of the data; otherwise, the comparison of the results of the different variations of data analysis is not reliable and will be reduced to an act of belief.

A second aspect to consider in developing the data curation workflow is that data are ideally made available for initial inspection soon after the first recordings are performed and should then already resemble the anticipated final output structure. Failure to analyze the data set early on bears the danger that potential shortcomings in the data are not noticed at an early stage. However, starting the data analysis using ad-hoc and makeshift solutions may prevent the later adoption of a more rigorous data management concept for those projects, for example, because they might not be not backward compatible. In such a situation, scientists who rely on such initial solutions might therefore not profit from future adjustments in data acquisition or postprocessing workflows, for example, to account for adaptations in postprocessing parameters or to incorporate additional descriptive metadata that were previously not considered.

Adopting the view of the naïve data consumer and the resulting need of a rigorous, comprehensible, and unambiguous data output, it became clear that both a defined process for data acquisition and postprocessing, as well as stable tools to implement corresponding standards in support of this process, are required (see Figure 1). The need for early access to the processed data, while allowing adjustments to the process as the experiment progresses, further indicated that going from the raw recorded data to the resulting processed data package must be reproducible

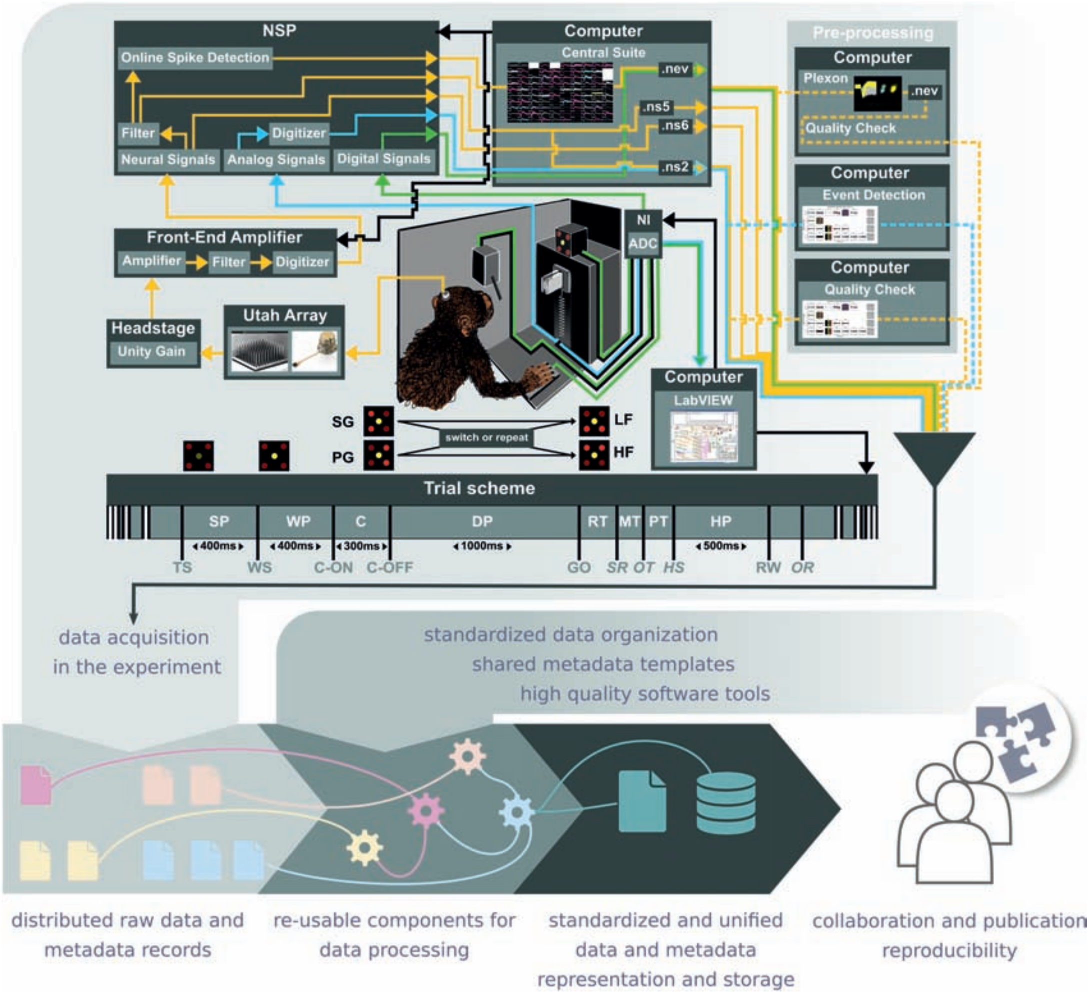


Figure 1: Data acquisition and postprocessing workflows simplify the process of preparing data for analysis, collaboration, and sharing. Data and metadata recorded in electrophysiological experiments are typically distributed across multiple disconnected files and stored in various proprietary and custom file formats. Development of standardized components for processing the recorded data based on common standards for data organization, shared ontologies and terminologies for metadata, and high-quality community tools leads to a comprehensive data representation that facilitates collaboration and fast publication via repositories. Top image modified from Zehl et al., 2016 (licensed under CC-BY).

at any time. As described in Zehl et al. (2016), such an approach requires – as much as possible – an automated way of building both the skeletal structure of data and metadata for the complete experiment and filling this structure with a particular data set. Such a generating process is able to produce a complete and consistent data package that is robust with respect to subsequent structural changes during the lifetime of the experiment, i.e., in case the data structure or specific parts of the metadata descriptions need to be adjusted.

While the implementation of such a process for data curation ultimately leads to a well-documented data set, designing this workflow is laborious when performed from scratch, since it requires the researcher to consider the design decisions of the process in minute detail. Yet, while experiments differ, individual parts of this curation workflow can most likely be singled out and used for multiple experiments as easily adaptable building blocks. Sharing and reuse of such components of data curation workflows are therefore key elements to facilitate researchers to produce comparable, complete, and versatile workflows in reasonable time before starting the experiment and thus enable a more structured way to implement the data curation process. The work programme of NFDI-Neuro supports this process by establishing mechanisms for researchers to share their data acquisition, postprocessing and analysis pipelines, to identify commonalities, and to produce common components for building individual data acquisition pipelines.

Metadata

When it comes to the underlying tools and services that enable such a data curation process, luckily the field is in a more advanced position. A first challenge we faced was the high degree of data fragmentation and descriptive metadata in the form of different files and file formats that need to be bundled for proper data curation. Here, the open metadata markup language (odML) offers an easy approach to adopt a machine readable and fully flexible data model that supports structuring and storing such metadata (Grewe et al., 2011). In this way, the wealth of details related to the experiment and each data set could be exposed to collaborators in an easily understandable manner. Still, the design of the actual metadata hierarchy for a particular experiment is a challenge, in particular in absence of standardized vocabularies or ontologies that suggest, based on prior experience, what metadata to record and how to label them.

More and more efforts are being drawn up to alleviate the problem of metadata organization and storage. These are in part proposed by data repository providers where metadata help in making data sets discoverable and interpretable, e.g., the CRCNS database (Teeters et al., 2008; http://crcns.org), EEGBase (Papez and Moucek, 2013), GIN (https://gin.g-node.org), or detailed metadata schemas that are developed as part of the EBRAINS curation service (https://ebrains.eu/service/share-data), such as openMINDS to describe high-level minimal metadata (cf., https://github.com/HumanBrainProject/openMINDS). Also, for more in-depth metadata describing further experimental details, efforts have started to pool and harmonize metadata templates for different experimental aspects like hardware components, experimental paradigms, and measurement techniques (cf., e.g., Bower, 2009). The emerging terminologies are commonly based on community contributions and published work, such as the G-Node terminologies (https://terminologies.g-node.org/), NeuroElectro (Tripathy et al., 2014; https://neuroelectro.org/), ontologies, and terminologies provided as part of the NIF information framework (Imam et al., 2012; https://neuinfo.org/). In part, metadata schemas inspired by computational neuroscience are equally relevant for neurophysiology, e.g., NeuronDB and ModelDB (Hines et al., 2004; https://senselab.med.yale.edu/neurondb).

One of the main goals of NFDI-Neuro's task area for electrophysiological data will be to work toward making these resources interoperable and easy to integrate into a detailed data acquisition and postprocessing workflow already at the planning stage of the experiment. For this, we envision that components of the workflow provide automatic metadata for stereotypical processing steps and assist in finding appropriate metadata descriptions for those parts of the curation workflow that require customization with respect to the specific experiment.

Data formats

The next question we encountered was that of choosing a data format in which the final data packet would be available. Typically, recorded data are stored in files, often using a file format specified by the manufacturer of the recording system. The researcher is then presented with two possible scenarios:

(1) The shared data files are left untouched and are accompanied by a piece of code that loads these data and metadata in accordance with the experiment.
(2) Alternatively, a new data file is created that contains the annotated and curated data and metadata in a standardized format.

Either option has advantages and disadvantages. In the first scenario, data duplication is minimal, an important factor for experiments generating large quantities of data. Moreover, keeping the original data minimizes the risk of potential errors in moving data from one representation to another. On the downside, the recipient of the shared data will be presented with a proprietary data format that requires a highly customized loading routine. Such loading routines are in danger of becoming outdated over time and rarely receive testing by a larger community to prevent errors. In the second scenario, this danger can be prevented by supplying the data set in a standardized format that is read by well-tested and maintained loading code that tends to be more stable over time.

More importantly, the second scenario has two further advantages. First, a common standard data file format will simplify the use of curated data in multiple programming languages. For example, at the time of publication of our data set (Brochier et al., 2018), these file formats were not yet sufficiently mature; consequently, a second set of data files (mat-format for Matlab users) had to be supplied in addition to the original data files and the Python code, thus duplicating storage space. Second, many vendor-specific formats are designed from the perspective of the recording system, i.e., data packets from multiple channels and are written progressively to file. From a consumer perspective, however, one of the most common scenarios is to read recording traces of one or several selected channels. The corresponding data samples are distributed across the raw data file, causing suboptimal performance in loading and processing. In contrast, standardized data formats provide more efficient data and metadata storage for the end user.

Both scenarios require a well-defined access to data stored in the various file formats. Efforts to form an alliance with manufacturers to provide a common, platform-independent, and well-tested basis for data access is still far from reality, despite early efforts by the Neuroshare initiative (http://neuroshare.sourceforge.net/index.shtml) aiming to unify access to various file formats from different vendors. In the Python world, the Neo data object model (Garcia et al., 2014) currently hosts the most comprehensive set of loading routines resulting from a community effort, which anticipates synergy with the SpikeInterface project aimed at evaluating spike sorter performance (Buccino et al., 2020; https://spikeinterface.readthedocs.io). In this design concept, data are represented in a common, generic structure independent of the source, which provides easy data access in a generic fashion from applications, analysis scripts, or other components of the data processing workflow.

When pursuing the first scenario of data publication (as we did for the data set described in Brochier et al., 2018), the most efficient and robust approach to construct the accompanying code for data access was to rely on a public community library such as Neo to handle the actual data loading and then annotate and reshape the resulting data object in a second step to optimally present its structure to the user. For the second scenario, data can be saved in a common, generic file format. Perhaps due to the high diversity of vendors of electrophysiological recording systems, such a common file format had not been available. However, to close this gap, two promising and complementary efforts have recently started. The first, Neurodata Without Borders (Teeters et al., 2015; https://www.nwb.org), offers a highly structured, HDF5-based format (Hierarchical Data Format version 5) to hold neurophysiological data sets based on a defined, optimized scheme. The second, NIX (Stoewer et al., 2014; http://www.g-node.org/nix), is a file format more customizable and suitable to combine structured data and arbitrary metadata records and fully compatible with odML-based metadata descriptions. Support to connect the Neo object model is continuously improved for both formats. With respect to the organization of data files at the file system level, ongoing efforts exist to extend structures such as the BIDS schema (Gorgolewski et al., 2016) to electrophysiology (cf., e.g., Pernet et al., 2019 for EEG, or discussions of the newly formed INCF special interest group on standardized data organization for electrophysiology), as are initiatives to establish interfaces with databases (e.g., Reimer et al., 2020).

NFDI-Neuro places a main focus on fostering these efforts toward common data models and file formats, and on making them interoperable with existing programming languages and storage solutions in the laboratories. Eventually, robust backing of data descriptions is key to ensure a smooth transition of electrophysiological data between any kind of data producer and data consumer.

Research data repositories

An important decision that must be made when deciding to share data is the physical storage location to use. Indeed, a number of data repositories exist to choose from, ranging from discipline-agnostic solutions such as Figshare (https://figshare.com) or Zenodo (https://zenodo.org), to generic institutional repositories, and to services catering specifically to the neuroscience community. Besides exposing the data set to a more targeted audience, the advantage when choosing one of the latter solutions is that these repositories are often able to interpret the contained data files as long as community standards are being adhered to. For example, the G-Node Infrastructure service used to store Brochier et al. (2018) is able to parse and display the odML encoded metadata schemes (see, e.g., https://gin.g-node.org/INT/multielectrode_grasp/src/master/datasets/i140703-001.odml), and the EBRAINS Knowledge Graph can link data sets to a corresponding view of its anatomical location in a brain atlas viewer.

Given the diversity of solutions that are available for sharing electrophysiological data, NFDI-Neuro's approach is to build a common infrastructure as a connecting layer, which will make access to data independent of specific storage solutions. In this way, researchers are able to choose the best repository for their data based on considerations of formal requirements, computational demands, and capabilities of the repository store, while being able to simplify discoverability and access to the data.

Looking back, when we started to develop strategies to best share and publish data sets, we soon realized that a more efficient and robust process for data curation is essential in our field of science. However, it was only by way of experience that the prevailing gaps in our workflows became apparent. For this reason, we are confident that establishing a process within NFDI-Neuro to foster the interaction between the experimental realities in the laboratories and the development of sophisticated tools will lay the foundation that in the future scientists need to worry less about the technicalities of managing their data, but instead can appreciate the creativity sparked by analyzing the richness of state-of-the-art neural recordings.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This study is supported by LMUexcellent, Helmholtz Association, European Union's Horizon 2020 Framework Programme under grant no. 945539 (Human Brain Project SGA3), Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE), the German Federal Ministry of Education and Research (BMBF 01GQ1903), and the German Research Foundation (CRC889, FOR1847).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

Bower, M.R., Stead, M., Brinkmann, B.H., Dufendach, K., and Worrell, G.A. (2009). Metadata and annotations for multi-scale electrophysiological data. 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Minneapolis, MN: IEEE), pp. 2811–2814.

Brochier, T., Zehl, L., Hao, Y., Duret, M., Sprenger, J., Denker, M., Grün, S., and Riehle, A. (2018). Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task. Sci. Data *5*, 180055.

Buccino, A.P., Hurwitz, C.L., Garcia, S., Magland, J., Siegle, J.H., Hurwitz, R., and Hennig, M.H. (2020). SpikeInterface, a unified framework for spike sorting. eLife *9*, e61834.

Denker, M. and Grün, S. (2016). Designing workflows for the reproducible analysis of electrophysiological data. Brain-Inspired Computing. K. Amunts, L. Grandinetti, T. Lippert, and N. Petkov, eds. (Cham: Springer International Publishing), pp. 58–72.

Garcia, S., Guarino, D., Jaillet, F., Jennings, T., Pröpper, R., Rautenberg, P.L., Rodgers, C.C., Sobolev, A., Wachtler, T., Yger, P., et al. (2014). Neo: An object model for handling electrophysiology data in multiple formats. Front. Neuroinf. *8*, 10.

Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci. Data *3*, 160044.

Grewe, J., Wachtler, T., and Benda, J. (2011). A bottom-up approach to data annotation in neurophysiology. Front. Neuroinf. *5*, 16.

Hines, M.L., Morse, T., Migliore, M., Carnevale, N.T., and Hines, M.L. (2004). ModelDB: A database to support computational neuroscience. J. Comput. Neurosci. *17*, 7–11.

Imam, F., Larson, S., Grethe, J., Gupta, A., Bandrowski, A., and Martone, M. (2012). Development and use of ontologies inside the neuroscience information framework: A practical approach. Front. Genet. *3*, 111.

Papez, V. and Moucek, R. (2013). Data and metadata models in electrophysiology domain: Separation of data models into semantic hierarchy and its integration into EEGBase. 2013 IEEE International Conference on Bioinformatics and Biomedicine (Shanghai, China: IEEE), pp. 539–543.

Pernet, C.R., Appelhoff, S., Gorgolewski, K.J., Flandin, G., Phillips, C., Delorme, A., and Oostenveld, R. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci. Data *6*, 103.

Plesser, H.E. (2018). Reproducibility vs. replicability: A brief history of a confused terminology. Front. Neuroinf. *11*, 76.

Reimer, M.L., Bangalore, L., Waxman, S.G., and Tan, A.M. (2020). Core principles for the implementation of the neurodata without borders data standard. J. Neurosci. Methods, 108972, https://doi.org/10.1016/j.jneumeth.2020.108972.

Stoewer, A., Kellner, C.J., Benda, J., Wachtler, T., and Grewe, J. (2014). File format and library for neuroscience data and metadata. Front. Neuroinform. Conference Abstract: Neuroinformatics 2014, https://doi.org/10.3389/conf.fninf.2014.18.00027.

Teeters, J.L., Godfrey, K., Young, R., Dang, C., Friedsam, C., Wark, B., Asari, H., Peron, S., Li, N., Peyrache, A., et al. (2015). Neurodata without borders: Creating a common data format for neurophysiology. Neuron *88*, 629–634.

Teeters, J.L., Harris, K.D., Millman, K.J., Olshausen, B.A., Sommer, F.T. (2008). Data sharing for computational neuroscience. Neuroinformatics *6*, 47–55.

Tripathy, S.J., Savitskaya, J., Burton, S.D., Urban, N.N., and Gerkin, R.C. (2014). NeuroElectro: A window to the world's neuron electrophysiology data. Front. Neuroinf. *8*, 40.

Wachtler, T., Bauer, P., Denker, M., Grün, S., Hanke, M., Klein, J., Oeltze-Jafra, S., Ritter, P., Rotter, S., Scherberger, H., et al. (2021). NFDI-Neuro: Building a community for neuroscience research data management in Germany. Neuroforum, (this issue).

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data *3*, 160018.

Zehl, L., Jaillet, F., Stoewer, A., Grewe, J., Sobolev, A., Wachtler, T., Brochier, T.G., Riehle, A., Denker, M., and Grün, S. (2016). Handling metadata in a neurophysiology laboratory. Front. Neuroinf. *10*, 26.

Bionotes

Michael Denker
Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
m.denker@fz-juelich.de
https://orcid.org/0000-0003-1255-7300

Michael Denker received his diploma in physics from the University of Göttingen, Germany, in 2002. In 2004, he started his doctoral studies in the lab of Sonja Grün at the Free University Berlin. In 2006, he became a researcher at the RIKEN Brain Science Institute, Japan. He was awarded his PhD in 2009 at the Free University Berlin, Germany. In 2011, he joined the Institute of Neuroscience and Medicine (Research Center Jülich, Germany) and now leads the group Data Science in Electro- and Optophysiology Behavioral Neuroscience. His research interests are the analysis of the correlation structure of neural activity and its relationship to signals that express population activity and the establishment of workflows that improve the reproducibility of data analysis in neurophysiology.

Sonja Grün
Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany
s.gruen@fz-juelich.de
https://orcid.org/0000-0003-2829-2220

Sonja Grün received her diploma in physics from the Eberhard Karls University in Tübingen (1991) and her PhD in physics from Ruhr University Bochum (1996). After being a postdoc at the Hebrew University (Jerusalem) and at Max-Planck Institute in Frankfurt (M), she became a junior professor at Freie University Berlin in 2002, and unit/team leader at RIKEN Brain Science Institute (Tokyo) in 2006. Since 2011, she is a full professor at RWTH Aachen University and leads the group Statistical Neuroscience (INM-6, Research Center Jülich), and was appointed director of INM-6/INM-10 in 2018. Her work focuses on the development of analysis strategies and tools that uncover concerted activity in massively parallel electrophysiological recordings from the cortex, which led to the additional focus on research data management.

Thomas Wachtler
Department Biologie II, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany
wachtler@bio.lmu.de
https://orcid.org/0000-0003-2015-6590

Thomas Wachtler has a background in physics and received his diploma and doctoral degree from the University of Tübingen. He was a postdoctoral researcher at the Salk Institute for biological Studies and at the universities of Freiburg and Marburg. His research interests are in the neural mechanisms of sensory processing with focus on vision. In his research, he combines experimental and computational approaches, including electrophysiology, psychophysics, and computational modeling to study the neural principles of processing and coding in the visual system and how they relate to the properties of the sensory environment and to perceptual phenomena. He is also working on neuroinformatics developments in the context of the International Neuroinformatics Coordinating Facility. Since 2009, he has been the Scientific Director of the German Neuroinformatics Node at LMU Munich, leading developments of tools and services for research data management in neuroscience.

Hansjörg Scherberger
Neurobiology Laboratory, Deutsches Primatenzentrum GmbH, Kellnerweg 4, 37077 Göttingen, Germany
Department of Biology and Psychology, University of Goettingen, Goettingen, Germany
hscherb@gwdg.de
https://orcid.org/0000-0001-6593-2800

Hansjörg Scherberger heads the Neurobiology Laboratory at the German Primate Center and is a professor for primate neurobiology at Göttingen University (since 2008). He received his diploma in mathematics (1993) and his medical doctor degree (1996) from Freiburg University, Germany, and subsequently was trained in systems electrophysiology at the University of Zurich (1995–1998) and the California Institute of Technology (1998–2003) before leading a research group at the Institute of Neuroinformatics at Zurich University and ETH (2004–2009). His research focuses on neural coding and decoding of hand movements, their interactions with sensory systems, and he develops brain-machine interfaces to read out movement intentions for the development of neural prosthetics to restore hand function in paralyzed patients.

DE GRUYTER Neuroforum 2021; 27(1): 35–43

Review article

Carsten M. Klingner*, Petra Ritter, Stefan Brodoehl, Christian Gaser, André Scherag, Daniel Güllmar, Felix Rosenow, Ulf Ziemann and Otto W. Witte

Research data management in clinical neuroscience: the national research data infrastructure initiative

https://doi.org/10.1515/nf-2020-0039

Abstract: In clinical neuroscience, there are considerable difficulties in translating basic research into clinical applications such as diagnostic tools or therapeutic interventions. This gap, known as the "valley of death," was mainly attributed to the problem of "small numbers" in clinical neuroscience research, i.e. sample sizes that are too small (Hutson et al., 2017). As a possible solution, it has been repeatedly suggested to systematically manage research data to provide long-term storage, accessibility, and federate data. This goal is supported by a current call of the DFG for a national research data infrastructure (NFDI). This article will review current challenges and possible solutions specific to clinical neuroscience and discuss them in the context of other national and international health data initiatives. A successful NFDI consortium will help to overcome not only the "valley of death" but also promises a path to individualized medicine by enabling big data to produce generalizable results based on artificial intelligence and other methods.

Keywords: NFDI; research data management; FAIR principles.

Zusammenfassung: In den klinischen Neurowissenschaften gibt es erhebliche Schwierigkeiten, Erkenntnisse aus der Grundlagenforschung in therapeutische klinische Strategien umzusetzen. Diese Lücke wurde als „Tal des Todes" (Hutson et al., 2017) bezeichnet und hat zu der Ansicht geführt, dass die klinisch-neurowissenschaftliche Forschung nicht optimal aufgestellt ist. Als mögliche Lösung wurde vorgeschlagen, Forschungsdaten systematisch zu verwalten, um eine langfristige Speicherung, Zugänglichkeit und Vernetzung der Daten bereitzustellen. Dieses Ziel wird durch einen aktuellen Aufruf der DFG für eine nationale Forschungsdateninfrastruktur (NFDI) unterstützt. In diesem Artikel werden aktuelle Probleme und mögliche Lösungen der NFDI für die klinische Neurowissenschaft beschrieben. Ein erfolgreiches NFDI-Konsortium wird dazu beitragen, nicht nur das „Tal des Todes" zu überwinden, sondern verspricht auch einen Weg zur individualisierten Medizin, indem die daraus resultierenden „Big Data" zusammen mit Methoden des maschinellen Lernens genutzt werden können.

***Corresponding author: Carsten M. Klingner**, Hans Berger Department of Neurology, University Hospital Jena, Erlanger Allee 101, 07747 Jena, Germany; and Biomagnetic Center, Jena University Hospital, Jena, Germany, E-mail: carsten.klingner@med.uni-jena.de. https://orcid.org/0000-0001-8815-3391
Petra Ritter, Brain Simulation Section, Department of Neurology, Berlin Institute of Health, Charité – Universitätsmedizin Berlin, Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany; Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany, E-mail: petra.ritter@charite.de
Stefan Brodoehl, Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany; Biomagnetic Center, Jena University Hospital, Jena, Germany, E-mail: stefan.brodoehl@med.uni-jena.de. https://orcid.org/0000-0002-6096-8684
Christian Gaser, Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany; Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany, E-mail: christian.gaser@uni-jena.de
André Scherag, Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital, Jena, Germany, E-mail: andre.scherag@med.uni-jena.de. https://orcid.org/0000-0002-9406-4704
Daniel Güllmar, Medical Physics Group, Institute for Diagnostic and Interventional Radiology, Jena University Hospital, Jena, Germany, E-mail: daniel.guellmar@med.uni-jena.de. https://orcid.org/0000-0001-8693-8099
Felix Rosenow, Department of Neurology, Epilepsy Center Frankfurt Rhine-Main, University Hospital Frankfurt, Frankfurt am Main, Germany; Center for Personalized Translational Epilepsy Research (CePTER), Goethe-University Frankfurt, Frankfurt am Main, Germany, E-mail: rosenow@med.uni-frankfurt.de. https://orcid.org/0000-0002-3989-7471
Ulf Ziemann, Department of Neurology and Stroke, University of Tübingen, Tübingen, Germany; Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany, E-mail: ulf.ziemann@uni-tuebingen.de
Otto W. Witte, Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany, E-mail: otto.witte@med.uni-jena.de. https://orcid.org/0000-0003-2101-4105

Schlüsselwörter: NFDI; Management von Forschungsdaten; FAIR Prinzipien.

Introduction

Clinical neuroscience (CNS) investigates the function and dysfunction of the human nervous system. It heavily relies on data from patients and healthy subjects. CNS employs functional measurements to characterize the brain and the peripheral nervous system: these comprise multichannel electroencephalography (EEG) and magnetoencephalography (MEG) and the associated—rather complex—evaluation techniques (e.g., inverse solutions, dynamic causal modeling, connectivity analyses). These techniques are used to quantify and localize signals and analyze interactions between brain regions and brain states, such as consciousness and sleep. Further functional measurements comprise, but are not limited to, visual, acoustic, somatosensory, and motor evoked potentials (EVPs), or invasive and noninvasive electromyography (EMG). Other techniques are used to directly influence the brain's function, such as electrical (usually with direct current) and magnetic brain stimulation or deep brain stimulation. The techniques are often combined, for example, in sleep polygraphy, and setups for computer-brain interfaces and closed-loop stimulation paradigms. An essential set of techniques was added in recent years by the availability of functional magnetic brain resonance imaging, which allows a direct correlation between structure and function. Many of these techniques have evolved into new dimensions using sophisticated analysis techniques. This allows for a semiautomatic and quantitative evaluation of, for example, disease progression, localization of circumscribed brain abnormalities in epilepsy, characterization of the brain's chemical properties by spectroscopy, and semiautomatic detection of abnormalities in early stroke by machine learning algorithms.

Unfortunately, data formats of these techniques and associated metadata are dominated by a wide variety of proprietary industry solutions. In this regard, CNS differs considerably from neuroimaging, where Digital Imaging and Communications in Medicine (DICOM) serves as a standard, promoting sharing and interoperability. In CNS—generally speaking—data are neither findable nor accessible, interoperable, or reusable. This has restrained the progress of research on many highly essential topics, such as consciousness and cognition, and psychiatric and neurological disorders.

Moreover, the availability of research data for the community remains low (Bryan Heidorn, 2008). This problem has led to considerable difficulties in translating basic research into clinical applications such as diagnostic tools or therapeutic interventions. This gap, known as the "valley of death," (Hutson et al., 2017) has been attributed mainly to the problem of "small numbers" in CNS research, that is too small sample sizes. Besides, data availability declines rapidly over time, with an estimated 17% increase in data loss probability per year after the publication of the article (Vines et al., 2014). Only recently, a discussion on the implementation of the FAIR principles (findability, accessibility, interoperability, and reusability) (Wilkinson et al., 2016) was started.

Key barriers

In the following, we identify significant barriers for the FAIRification of CNS data. Those comprise technical, knowledge, motivational, resource, legal, and ethical barriers. The goal of national research data infrastructure (NFDI) neuroscience is to develop solutions to overcome these barriers.

Technical barriers

Research data sharing in compliance with the FAIR principles place high demands on technical implementations. The resulting technical barriers have long been known, and there has been a lot of research and publications that have investigated possible solutions. Most obstacles have been solved theoretically and can also be solved in practice, but sustainable working implementations and political and financial commitment remained limited so far. Therefore, technical barriers remain significant challenges for the availability and use of data in all medical research areas and the public healthcare system (Packer, 2018; van Panhuis et al., 2014).

Challenges (I)

Research infrastructure for a lawful (that is General Data Protection Regulation (GDPR) compliant—see below) long-term storage of personal data in a data warehouse with archiving and backup solutions is often not available. Also, standardized annotation/metadata models are lacking. Only with proper annotation, contextual information is provided that allows meaningful data use in the long term.

Objective

Research data should be easy to find, and the availability should be guaranteed.

Description

The retention or archiving of data is often not prioritized, especially if research is performed with limited capacities and resources (Terry et al., 2018). In most cases, the data are “available” for the prescribed or recommended retention period. Usually, the data are collected for a specific project or to answer a specific scientific hypothesis. The storage of the data is typically limited to this conceptual framework. Therefore, this conceptual framework has to be changed to a framework for long-term data storage. In that case, there will be little effort needed to choose, for example, a data format designed for long-term data storage and to move data in repositories that can be easily accessed. To improve data findability, the data should be published along with metadata in data portals and warehouses. The number of sustainable data retrieval systems should be increased. Moreover, the archiving of data must be an independent process that runs in parallel to research; it should be independent of relocation and fluctuation of people, computers, offices, physical damage to paper or electronic files, computer viruses, etc.

Challenges (II)

Interoperability is limited by a lack of metadata describing data content, origin, methods, etc.

Objective

Generation of a published standard for the formalization and annotation of data in CNS at the individual level includes different modalities as subdomains.

Description

CNS brings together many submodalities and supplements them with clinical data. To this end, we need to define minimum requirements for data annotation that complement the metadata standards of the included data (Khvastova et al., 2019). Very likely, these metadata will be domain-specific and require a definition of domain-specific metadata. A consensus about minimum requirements for data annotation should be reached within the community, for example, through community discussions, publications, and workshops. Based on these results, a virtual environment of federated data repositories can be created. To improve the understanding, usability, and spreading of the standard, guidelines for the acquisition, annotation, and curation of multimodal data and a best-practice implementation should be provided.

Challenges (III)

At present, data exchange between IT systems in CNS and IT systems in the public healthcare system is limited (mostly non-existing) (Winter et al., 2018). Data in the domain of CNS and the public healthcare system are mainly stored in proprietary formats from a wide variety of manufacturers without intramodal or intermodal compatibility.

Objective

CNS data are available in a FAIR research infrastructure and can be exchanged between research and healthcare infrastructures by specified standards. Standards for data exchange and archiving should ensure interoperability with the existing research data infrastructure.

Description

Unlike in the medical imaging domain, where data are typically fully compliant with the DICOM standard, the situation differs in CNS. Device and software manufacturers are often small- to medium-sized companies that employ their own proprietary data formats. Very recently, a DICOM standard for clinical neurophysiological data has been proposed, which facilitates the integration of electrophysiological data into common and interoperable picture archiving and communication systems (PACS). This proposed standard does not cover processed data: in particular, annotations and clinical findings are typically not structured and stored in syntactically or semantically standardized ways, while different complementary initiatives exist in clinical and the research communities (e.g., the Brain Imaging Data Structure [BIDS]) (Gorgolewski et al., 2016). To date, this approach covers only a small part of the data, which is mainly raw data and some derivative data. At least in disease description, there has been an improvement in recent years through the general use of the ICD system (Office of the Secretary, HHS, 2014; The Lancet, 2019).

However, data formats' heterogeneity and the lack of a uniform standard represent a central obstacle in creating interoperability (technological, semantic, and organizational). An essential aspect of new data standards is the provision of reference implementations. This software must validate data conformity as well as to collect broader user feedback. The software should be capable of automatic metadata extraction, creation, validation, and exchange of standardized data formats. It should be available for several major programming languages that are widely used in CNS (e.g., Python, Java, Matlab). This software aims to offer well-documented and easily understandable best-practice implementations used as templates in the community.

Knowledge barriers

Challenges

Researchers might have limited knowledge about the existing solutions in the theoretical, software, and hardware domain. A lack of knowledge can prevent the use of existing and available solutions and thus represent a significant barrier to all goals of a sustainable research data management.

Objective

The objective is to create awareness and offer training for the use of new technologies and tools for data sharing

Description

The acceptance and a wide range of applications are of crucial importance. Therefore, continuous feedback from and to the neuroscience community and education and teaching is of utmost importance. Support should be organized centrally, for example, by a domain-specific help desk. Additionally, the neuroscience community should more closely cooperate with standardization organizations such as "Health Level Seven International" (HL7), which have solved many of the interoperability questions across medical domains already.

Motivational and resource barriers

The process of data sharing requires human and technical resources for data preparation, annotation, communication with recipients, computer equipment, and, above all, time. Time and resources are chronically lacking in CNS as in every other scientific research branch. What are the reasons that data providers motivate to share data—extrinsic and intrinsic?

Challenges

The challenges are too little extrinsic and intrinsic incentives, the fear of researchers that mistakes or quality issues are identified in their data/results with a high degree of transparency, and no reward system for data sharing.

Objectives

The objective is to create incentives for researchers to care about data sharing.

Description

There is some incentive needed (personal and institutional) to prioritize data sharing over other pressing duties. Currently, data providers for secondary analyses receive little attention and credit. In CNS, most data are collected to test scientific hypotheses. This process of testing and thereby falsifying or corroborating hypotheses motivates scientists and delivers the feeling of extending human knowledge. To endure the efforts of preparing, annotating, and sharing of data currently has little clear focus. The potential profit, even for the greater good, is not nameable. The risks and benefits of exchanging data are often perceived as unfair. Data producers often feel exploited when they benefit little from their work, whereas data users who analyze the data quickly and publish results can significantly benefit and may be characterized as "research parasites" (Longo and Drazen, 2016). The metrics and perception currently relevant in Germany—but also internationally—clearly favor the publishing researchers with little regard to who is the data producer.

Data providers could be discredited by errors found during secondary use of their data. Particularly, scientists who have invested time and effort in data collection might fear that scientific credit may be lost if data recipients with greater capacity or more sophisticated analysis methods could gain the majority of credit (Pisani and AbouZahr, 2010). The possibility that data providers may disagree with the intended secondary use of their data can also reduce the motivation for data sharing (van Panhuis et al., 2014). All these motivational issues can only be overcome with a cultural shift that values data sharing and quality assurance through the detection of errors by the community and through a shift toward the notion that mistakes happen and that their professional management can increase recognition and trustworthiness of the concerned scientists rather leading to their disqualification.

Legal and ethical aspects in CNS

Lawful processing of personal data is a cross-cutting topic relevant beyond neurosciences and regulated through the EU General Data Protection Regulations in place since 2018. Anyone who collects data from people is responsible for obtaining a lawful basis for doing so and for protecting the rights and freedom of "data subjects," that is, the identifiable natural persons from whom the data have been collected. This applies equally to researchers and institutions. This responsibility also has implications for data sharing. Researchers often do not know the law in all its complexity with its implications regarding who is the data controller, what are the technical and organizational measures to keep the data safe during storage and processing, and what are the rules for data sharing.

Challenges

Identifiable health data is, according to GDPR, a special category of data under high protection. Only under certain conditions—that guarantee the data subject's rights and freedoms—their processing is legal. Processing of health data—according to GDPR—required a data protection impact assessment that is a detailed analysis of the risks for the data subjects and implemented measures to mitigate those. Additionally, it requires evaluating the data protection concept by a data protection officer. Due to the complex legal aspects and a lack of expert knowledge and support in this domain, personal health data processing is often hindered.

Objectives

The objective is to establish a framework for the legal and ethical handling of personalized data and appropriate support mechanisms for researchers.

Description

In a patient-centered science branch like CNS, most of the data come from an individual. A clear distinction between data with personal identifiers and completely anonymous data are not always possible, which leads to restrictive guidelines for all types of data for data protection reasons. Aggregated data without personal identifiers are often not detailed enough for certain research questions or applications. Existing tools and standards for deidentifying personal identifiers such as statistical data masking may not be known or available in many contexts (Wartenberg and Thompson, 2010). Some records consist primarily of personal identifiers and may become useless if identification precludes the possibility of reidentification—since it is precisely the personal information relevant for research. We need certifiable standards or blueprint solutions for organizational and technical measures for the data protection of workflows and IT systems. These standards must find the optimal balance between keeping data secure and simultaneously enable researchers to use personal data in compliance with GDPR. The data protection concept shall use all practically implementable protection measures that can help prevent unauthorized access and abuse of personal data and enable the data subjects' rights, for example, for information, data deletion, and transparency.

Best practices must be developed and published to describe how to guarantee lawful neuroscience health data processing. The development process should involve collaboration with related international (e.g., Open Brain Consent (Pernet et al., 2020), European Open Science Cloud [EOSC], the Virtual Brain Cloud, the Human Brain Project, or the European Health Data Space) and national efforts, like the German group "data security" of the TMF—Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. or the Medical Informatics Initiative. The best practices should address the separation of data availability (others can see which data are available) and accessibility (others can also access the data), the separation and handling of identifying data, the provisioning of metadata while guaranteeing data privacy, and linking institutional to the NFDI-Neuro Common Infrastructure (COIN). One important step might be the development of an online helpdesk for the generation of data protection concepts based on successful concepts collected from the community.

Conclusion

The development of a national research data infrastructure in CNS is a colossal undertaking that requires

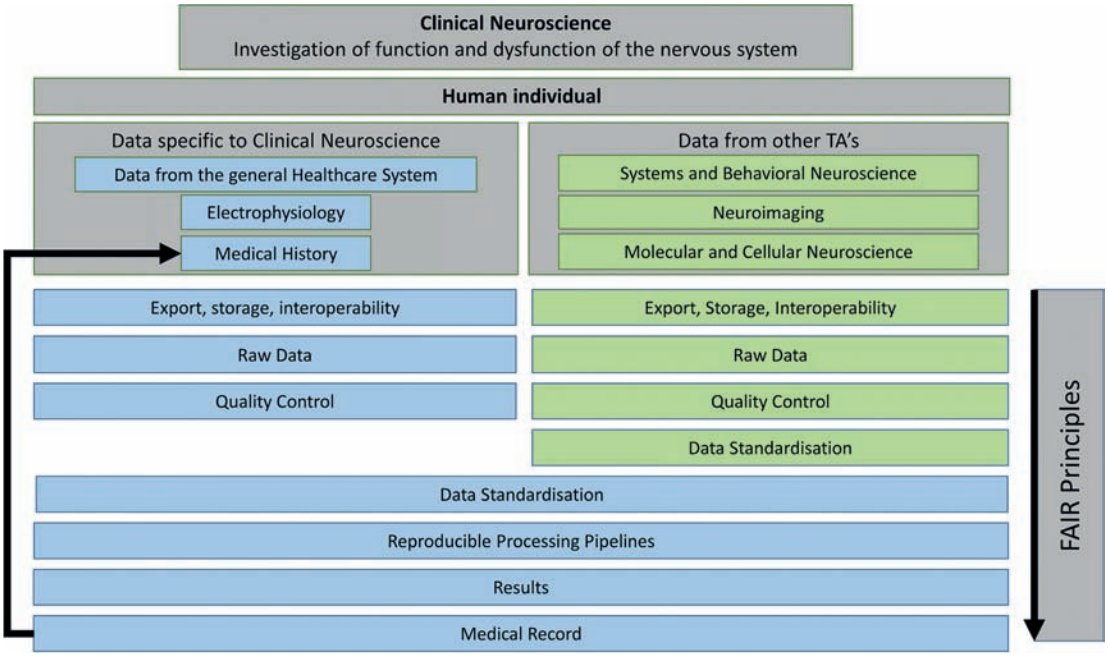


Figure 1: Main data sources and knowledge extraction process in clinical neuroscience.

considerable effort from the entire research community. The goal is the establishment of FAIR principles for multimodal neuroscience health data that comprise neurophysiological, neuroimaging, behavioral, and clinical data of patients and healthy volunteers (Figure 1). This endeavor will be a lengthy process made up of many different measures and requires structured cooperation with standardization organizations and neighboring consortia and disciplines.

Competences and expertise

The scientists working on this area have a long-standing interest and expertise in research data (Khvastova et al., 2019; Klingner et al., 2016; Rauch et al., 2019; Ritter et al., 2013; Schirner et al., 2018; Sonntag et al. 2018; Winter et al., 2018). They cooperate with the data integration center of the UKJ as part of the German medical informatics initiative within the consortium "Smart Medical Information Technology for Healthcare" (SMITH; www.smith.care), the Virtual Research Environment and Health Data Platform of the Charité and Berlin Institute of Health (BIH), with the Medical Informatics Initiative, the national research data infrastructure initiative NFDI4-Health, the European Open Science Cloud, the Virtual Brain Cloud, the International Neuroinformatics Coordination Facility, the Human Brain Project, and several other projects with related topics. Owing to the interaction with the scientific community through the German Society for Clinical Neurophysiology and Functional Imaging (DGKN), interactions with many user sites and the industry can be integrated. The DGKN also has established national expert commissions which will contribute their expertise in working groups.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: The authors received support from German Research Foundation (DFG) for 1738 B2; Wi 830 12-1; CRC 1315; CRC 936, CRC-TRR 295 and RI 2073/6-1; BMBF Gerontosys JenAge (FKZ 031 5581B); BMBF SMITH (FKZ 01ZZ1803C); EU EpimiRNA (FP7 GA No. 602130), EU H2020 826421, 785907; 945539, ERC 683049; Berlin Institute of Health & Foundation Charité, Johanna Quandt Excellence Initiative, LOEWE Hessen (Center for Personalized Translational Epilepsy Research – CePTER).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

Gorgolewski, K. J., 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., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data *3*, 160044.

Heidorn, P. B. (2008). Shedding light on the dark data in the long tail of science. Libr. Trends *57*, 280–299.

Hutson, P.H., Clark, J.A., and Cross, A.J. (2017). CNS target identification and validation: Avoiding the valley of death or naive optimism? Annu. Rev. Pharmacol. Toxicol. *57*, 171–187.

Khvastova, M., Witt, M., and Krefting, D. (2019). Towards interoperability in clinical research: enabling FHIR on the open source research Platform XNAT. Stud. Health Technol. Inf. *258*, 3–5.

Klingner, C.M., Brodoehl, S., Huonker, R., and Witte, O.W. (2016). The processing of somatosensory information shifts from an early parallel into a serial processing mode: a combined fMRI/MEG study. Front. Syst. Neurosci. *10*, 103.

Longo, D. L. and Drazen, J. M. (2016). Data sharing. N. Engl. J. Med. *374*, 276–277.

Office of the Secretary, HHS (2014). Administrative simplification: change to the compliance date for the International Classification of Diseases, 10th Revision (ICD-10-CM and ICD-10-PCS) medical data code sets. Final rule. Fed. Regist. *79*, 45128–45134.

Packer, M. (2018). Data sharing in medical research. BMJ *360*, k510.

Pernet, C.R., Heunis, S., Herholz, P., and Halchenko, Y.O. (2020). The open brain consent: Informing research participants and obtaining consent to share brain imaging data, https://doi.org/10.31234/osf.io/f6mnp.

Pisani, E. and AbouZahr, C. (2010). Sharing health data: good intentions are not enough. Bull. World Health Organ. *88*, 462–466.

Rauch, G., Röhmel, J., Gerß, J., Scherag, A., and Hofner, B. (2019). Current challenges in the assessment of ethical proposals-aspects of digitalization and personalization in the healthcare system. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz *62*, 758–764.

Ritter, P., Schirner, M., McIntosh, A.R., and Jirsa, V.K. (2013). The virtual brain Integrates computational modeling and multimodal neuroimaging. Brain Connect. *3*, 121–145.

Schirner, M., McIntosh, A.R., Jirsa, V., Deco, G., and Ritter, P. (2018). Inferring multi-scale neural mechanisms with brain network modelling. Elife *7*, e28927.

Sonntag, H., Haueisen, J., and Maess, B. (2018). Quality assessment of MEG-to-MRI coregistrations. Phys. Med. Biol. *63*, 075003.

Terry, R.F., Littler, K., and Olliaro, P.L. (2018). Sharing health research data - the role of funders in improving the impact. F1000Res *7*, 1641.

The Lancet (2019). ICD-11. Lancet *393*, 2275.

van Panhuis, W.G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A.J., Heymann, D., and Burke, D.S. (2014). A systematic review of barriers to data sharing in public health. BMC Public Health *14*, 1144.

Vines, T. H., Albert, A.Y., Andrew, R.L., Débarre, F., Bock, D.G., Franklin, M.T., Gilbert, K.J., Moore, J.S., Renaut, S., and RennisonD.J. (2014). The availability of research data declines rapidly with article age. Curr. Biol. *24*, 94–97.

Wartenberg, D. and Thompson, W.D. (2010). Privacy versus public health: the impact of current confidentiality rules. Am. J. Public Health *100*, 407–412.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data *3*, 160018.

Winter, A., Stäubert, S., Ammon, D., Aiche, S., Beyan, O., Bischoff, V., Daumke, P., Decker, S., Funkat, G., Gewehr, J.E., et al. (2018). Smart medical information technology for healthcare (SMITH): data integration based on interoperability standards. Methods Inf. Med. *57*, e92–e105.

Bionotes

Carsten M. Klingner
Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany
Biomagnetic Center, Jena University Hospital, Jena, Germany
carsten.klingner@med.uni-jena.de
https://orcid.org/0000-0001-8815-3391

Carsten M. Klingner studied computer science and medicine at the Technical University Berlin and Charité Berlin. He received a diploma in computer science and a doctorate after his state examination in medicine. Then he moved to Jena where he received his neurology education with Prof. O.W. Witte. Since 2015, he is the head of the Biomagnetic Center at the Hans Berger Department of Neurology in Jena. His scientific interests include brain plasticity and the interaction between sensory and motor signals in the human brain. He performed multiple studies in the field of CNS that combined functional imaging and behavioral experiments. In the course of these studies, he developed a strong interest for research data management in CNS.

Petra Ritter
Brain Simulation Section, Department of Neurology, Berlin Institute of Health, Charité – Universitätsmedizin Berlin, Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany
Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany
petra.ritter@charite.de

Petra Ritter studied medicine at the Charité University Medicine Berlin. She spent a large part of her clinical traineeships and practical year abroad: at the universities UCLA and UCSD in Los Angeles and San Diego, the Mount Sinai School of Medicine in New York, and the Harvard Medical School in Boston. In 2002, she received her license to practice medicine. In 2004, she completed her doctoral thesis at the Charité, and in 2010, she received habilitation in experimental neurology. After being Max Planck Minerva research group leader from 2011 to 2015, she assumed the lifetime position of BIH Johanna Quandt Professor for Brain Simulation at Berlin Institute of Health (BIH) and Charité Universitätsmedizin Berlin, one of Europe's largest university hospitals. Since 2017, she is the Director of the Brain Simulation Section at Charité Universitätsmedizin Berlin. Ritter holds an ERC Consolidator grant and serves in the leadership of several national and international neuroinformatics consortia.

Stefan Brodoehl
Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany
Biomagnetic Center, Jena University Hospital, Jena, Germany
stefan.brodoehl@med.uni-jena.de
https://orcid.org/0000-0002-6096-8684

Stefan Brodoehl studied medicine and mathematics in Jena. In 2006, he received his license to practice medicine. In 2007, he completed her doctoral thesis at the Charité, and in 2017, he received habilitation in experimental neurology. Since 2007, he worked in the Hans Berger Department of Neurology in Jena where he received his neurology education. Since 2019, he is the head of the dementia center of the university hospital in Jena. His scientific interests include brain plasticity and brain aging, as well as experimental and clinical brain imaging. He has applied these methodological interests in many studies to investigate healthy and nonhealthy cerebral aging processes.

Christian Gaser
Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany
Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany
christian.gaser@uni-jena.de

Christian Gaser studied electrical engineering and technical acoustics in Chemnitz and Dresden. He received his PhD degree in neuroscience from the University of Magdeburg, Germany, in 2001. He was a postdoctoral fellow at Harvard Medical School in Boston in 2001 and was a visiting postdoctoral fellow at UCLA in Los Angeles, Mount Sinai School of Medicine in New York, Australian National University in Canberra, and Auckland University. Currently he is an associate Professor of Computational Neuroscience and Neuroimaging and the Head of the Structural Brain Mapping Group at the University of Jena. His research program is directed toward the development of advanced computation tools for the analysis of structural brain data. More specifically, he is heavily involved in the development of algorithms and tools for processing of voxel- and surface-based imaging data which encompasses segmentation, surface reconstruction, and disease prediction. He has developed several software tools, including the Computational Anatomy Toolbox for performing voxel- and surface-based morphometry, which is widely used by the scientific community (>40000 downloads).

André Scherag
Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital, Jena, Germany
andre.scherag@med.uni-jena.de
https://orcid.org/0000-0002-9406-4704

André Scherag is the director of the Institute for Medical Statistics, Informatics and Data Sciences and deputy scientific director of the Center for Clinical Studies Jena. He studied psychology and biostatistics in Marburg and Heidelberg. After obtaining his doctorate in Marburg in biostatistics, he moved to the University Hospital Essen in 2007. In Essen, he was the head of Biometry at the Center for Clinical Trials Essen and Head of the Research Group Biometry and Bioinformatics. After his habilitation and *Venia legendi* for “Medical Informatics, Biometry and Epidemiology,” he moved to Jena in 2013 as a full Professor of Clinical Epidemiology. His research interests are diverse and cover topics from clinical, genetic, and translational epidemiology, as well as clinical trial statistics, biometry, bioinformatics, biostatistics, and digital medicine. He is particularly interested in methodological challenges connected to translating and communicating results from basis and translational research to clinical research and even evidence-based clinical practice (and thus to put some of the promises made by precision medicine into effect). In addition, he is the deputy speaker of consortium SMITH (“Smart Medical Information Technology for Healthcare”; www.smith.care) which is funded within the German medical informatics funding scheme. Interoperable solutions for (research) data management in medicine and the healthcare system are currently implemented as part of the data integration centers within SMITH.

Daniel Güllmar
Medical Physics Group, Institute for Diagnostic and Interventional Radiology, Jena University Hospital, Jena, Germany
daniel.guellmar@med.uni-jena.de
https://orcid.org/0000-0001-8693-8099

Daniel Güllmar received a diploma in media technology and a doctoral degree in biomedical engineering at the Ilmenau University of Technology. He has started his scientific work in the field of EEG/MEG source localization in combination with advanced volume conductor models at the Biomagnetic Center Jena. During this research phase, he worked also as a visiting scientist at the A. Martinos Center for Biomedical Imaging in Boston, MA, USA, in 2005/2006. After his PhD phase, he shifted his focus on diffusion MRI methods to study neurological and psychiatric disease, as well as psychological mechanisms in interdisciplinary collaborations. Currently, he analyzes the strengths and weaknesses of deep learning methods in biomedical imaging.

Felix Rosenow
Department of Neurology, Epilepsy Center Frankfurt Rhine-Main, University Hospital Frankfurt, Frankfurt am Main, Germany
Center for Personalized Translational Epilepsy Research (CePTER), Goethe-University Frankfurt, Frankfurt am Main, Germany
rosenow@med.uni-frankfurt.de
https://orcid.org/0000-0002-3989-7471

Felix Rosenow studied medicine at the Free University of Berlin and received a research grant from the Max-Planck Institute for

Neurological Research Cologne and subsequently an MD degree at the University of Cologne, where he also was trained as a board-certified neurologist. He received postdoctoral training at the Cleveland Clinic Foundation (Department of Neurology, Section of Epilepsy and Sleep). He was Professor for Neurology/Epileptology at the Philipps University Marburg and is now Professor and Director of the Epilepsy Center Frankfurt Rhine Main at the University Hospital Frankfurt. His research interests include clinical and translational epilepsy research, clinical neurophysiology, and genetics. He currently serves as a member of the Research Strategy Committee of the Medial Faculty and as a speaker of the LOEWE Center for Personalized Translational Epilepsy Research (CePTER) of the Goethe-University Frankfurt and as the President of the German Society of Clinical Neurophysiology and Functional Imaging (www.DGKN.de).

Ulf Ziemann
Department of Neurology and Stroke,
University of Tübingen, Tübingen, Germany
Hertie Institute for Clinical Brain Research,
University of Tübingen, Tübingen, Germany
ulf.ziemann@uni-tuebingen.de

Ulf Ziemann has research expertise in human motor cortex, excitability, plasticity, motor learning, TMS, brain state–dependent stimulation, neuropharmacology, and TMS-EEG and clinical expertise in stroke, neurointensive care, neuroimmunology, and clinical neurophysiology; holds current positions as the Editor-in-Chief of "Clinical Neurophysiology," ExCo member of the International Federation of Clinical Neurophysiology, Deputy Editor of "Brain Stimulation," Associate Editor of "Journal of Neuroscience," and President of the German Society for Clinical Neurophysiology and Functional Imaging (DGKN); has received awards including Richard-Jung Prize of the German Society of Clinical Neurophysiology and Functional Neuroimaging (DGKN), National Institutes of Health (NIH, Bethesda, USA) Merit Award, and NIH Fellowship Award for Research Excellence; and has published 350 peer-reviewed publications, 40 book chapters, and 7 books, with cumulative IF of 2.070, ISI citations of 32.000, ISI h-index of 90, and Google Scholar h-index of 103.

Otto W. Witte
Hans Berger Department of Neurology, Jena
University Hospital, Jena, Germany
otto.witte@med.uni-jena.de
https://orcid.org/0000-0003-2101-4105

Otto W. Witte studied medicine, psychology, and mathematics in Münster and London. He worked as a postdoc in Neurophysiology in Münster with E.-J. Speckmann before he moved to Düsseldorf where he received his neurology education with H.-J. Freund. Since 2001, he is the director of the Hans Berger Department of Neurology in Jena. His scientific interests include brain plasticity and brain aging, as well as experimental and clinical brain imaging. As the secretary of the DGKN, he heads the office of the society which supports research and innovation in the exciting field of clinical neurophysiology and functional brain imaging, and is engaged in establishing standard procedures and quality control measures.

DE GRUYTER Neuroforum 2021; 27(1): 45–49

Nachrichten aus der Gesellschaft

https://doi.org/10.1515/nf-2021-0002

Ergebnis der Wahl zum Vorstand der Neurowissenschaftlichen Gesellschaft e.V. für die Amtsperiode 2021–2023

Zum Stichtag 31. Januar 2021 wurden 537 Wahlzettel eingesandt. Das entspricht einer Wahlbeteiligung von 25,6 %. Davon waren 494 Wahlzettel gültig, 43 mussten als ungültig gewertet werden und sind nicht in das Abstimmungsergebnis eingegangen. Die ordnungsgemäße Durchführung der Wahl wird vom Wahlleiter, Prof. Dr. Michael Synowitz, Kiel, bestätigt.

Präsidentin	Prof. Dr. Christine R. Rose (Düsseldorf)		
	Ja: 469	Nein: 13	Enthaltung: 12
Vizepräsident	Prof. Dr. Frank Kirchhoff (Homburg): 248 Prof. Dr. Christian Wegener (Würzburg): 226		
Generalsekretär	Prof. Dr. Christian Steinhäuser (Bonn)		
	Ja: 469	Nein: 8	Enthaltung: 17
Schatzmeister	Prof. Dr. Ansgar Büschges (Köln)		
	Ja: 478	Nein: 7	Enthaltung: 9

Sektionssprecher

Computational Neuroscience	Prof. Dr. Uli Egert (Freiburg): 32 Prof. Dr. Sonja Grün (Jülich): 86
Entwicklungsneurobiologie/ Neurogenetik	Prof. Dr. Constance Scharff (Berlin): 60 Prof. Dr. Tanja Vogel (Freiburg): 33
Junge NWG (jNWG)	Dr. Sophie Seidenbecher (München): 16
Klinische Neurowissenschaften	Prof. Dr. Mathias Bähr (Göttingen): 77
Kognitive Neurowissenschaften	Prof. Dr. Christiane Thiel (Oldenburg): 53 Prof. Dr. Melanie Wilke (Göttingen): 31
Molekulare Neurobiologie	Prof. Dr. Dirk Dietrich (Bonn): 57 Prof. Dr. Tobias Böckers (Ulm): 83
Neuropharmakologie/ -toxikologie	Prof. Dr. Heidrun Potschka (München): 35 Prof. Dr. Markus Schwaninger (Lübeck): 29
Systemneurobiologie	Prof. Dr. Ileana Hanganu-Opatz (Hamburg): 69
Verhaltensneurowissenschaften	Prof. Dr. Martin Göpfert (Göttingen): 72 Prof. Dr. Gary Lewin (Berlin): 60
Zelluläre Neurowissenschaften	Prof. Dr. Veronica Egger (Regensburg): 122 Prof. Dr. Birgit Liss (Ulm): 53

Der Vorstand der Amtsperiode 2021–2023 setzt sich somit wie folgt zusammen:
Präsidentin: *Prof. Dr. Christine R. Rose (Düsseldorf)*
Vizepräsident: *Prof. Dr. Frank Kirchhoff (Homburg)*
Generalsekretär: *Prof. Dr. Christian Steinhäuser (Bonn)*
Schatzmeister: *Prof. Dr. Ansgar Büschges (Köln)*

Sektionssprecher
Computational Neuroscience: *Prof. Dr. Sonja Grün (Jülich)*
Entwicklungsneurobiologie/Neurogenetik: *Prof. Dr. Constance Scharff (Berlin)*
Junge NWG (jNWG): *Dr. Sophie Seidenbecher (München)*
Klinische Neurowissenschaften: *Prof. Dr. Mathias Bähr (Göttingen)*
Kognitive Neurowissenschaften: *Prof. Dr. Christiane Thiel (Oldenburg)*
Molekulare Neurobiologie: *Prof. Dr. Tobias Böckers (Ulm)*
Neuropharmakologie/-toxikologie: *Prof. Dr. Heidrun Potschka (München)*
Systemneurobiologie: *Prof. Dr. Ileana Hanganu-Opatz (Hamburg)*
Verhaltensneurowissenschaften: *Prof. Dr. Martin Göpfert (Göttingen)*
Zelluläre Neurowissenschaften: *Prof. Dr. Veronica Egger (Regensburg)*

Ehrenpräsident: *Prof. Dr. Albert C. Ludolph (Ulm)*

Der neue Vorstand tritt sein Amt mit dem Ende der Göttinger Tagung der NWG am 30. März 2021 an.

Stipendien für virtuelle Göttinger Tagung 2021 vergeben

Die folgenden Bewerber wurden für ein Stipendium für die Teilnahme an der virtuellen Göttinger Tagung 2021 (22.–30. März) ausgewählt. Den erfolgreichen Stipendienbewerbern wurde die Registrierungsgebühr für die Teilnahme an der Tagung erlassen:

(1) Aksan, Bahar (Heidelberg, Germany)
(2) Anisimova, Margarita (Hamburg, Germany)
(3) Armasescu, Florian-Vintila (Bucharest, Romania)
(4) Bicakci, Ahmet (Magdeburg, Germany)
(5) Bica-Popi, Melania (Bucharest, Romania)
(6) Constantin, Oana M. (Hamburg, Germany)
(7) de Tredern, Eloïse (Gif-sur-Yvette, France)
(8) DeMaegd, Margaret Louise (Normal, USA)
(9) Folschweiller, Shani (Freiburg, Germany)
(10) Ghenghea, Mihail-Sebastian (Bucharest, Romania)
(11) Grosu, Andreea-Violeta (Bucharest, Romania)
(12) Kleis, Piret (Freiburg, Germany)
(13) Kocovic, Dušica M. (Belgrade, Serbia)
(14) Lange, Sven (Bonn, Germany)
(15) Lehman, Maxime (Gif-sur-Yvette, France)
(16) Liedtke, Maik (Rostock, Germany)
(17) Mahishi, Deepthi (Leipzig, Germany)
(18) Menne, Laura (Hannover, Germany)
(19) Milicevic, Katarina (Beograd, Serbia)
(20) Pierzchlinska, Anna (Szczecin, Poland)
(21) Rodrigues Neves, Ana Catarina (Coimbra, Portugal)
(22) Scharr, Melanie (Tübingen, Germany)
(23) Schmaul, Samantha (Mainz, Germany)
(24) Senn, Lara (Modena, Italy)
(25) Vestring, Stefan (Freiburg, Germany)
(26) Vitale, Maria Rosaria (Würzburg, Germany)
(27) Völkner, Christin (Rostock, Germany)
(28) Wolf, Katharina (Erlangen, Germany)

Bewerben konnten sich Studenten, Doktoranden und Postdocs, die zum Zeitpunkt der Bewerbung maximal 35 Jahre alt waren und an der Göttinger Tagung mit einem eigenen Beitrag als Erstautor teilnahmen. Als Bewerbungsunterlagen waren ein kurzer Lebenslauf, eine Publikationsliste (falls vorhanden), eine Kopie des Abstracts sowie ein kurzes Empfehlungsschreiben gefordert.
Wir gratulieren den erfolgreichen Kandidaten!

Schilling Forschungspreis der Neurowissenschaftlichen Gesellschaft 2021

Der Schilling Forschungspreis der Neurowissenschaftlichen Gesellschaft wurde 2021 an Katrin Franke, die eine Juniorgruppe am Department für Augenheilkunde in Tübingen leitet, verliehen.

Wie wir die Welt sehen, hängt davon ab, welche Informationen unsere Augen an das Gehirn senden. Bereits im Auge - der ersten Verarbeitungsstufe des visuellen Systems - zerlegt die Netzhaut das eingehende visuelle Signal in parallele Bilddarstellungen, die das Gehirn über wichtige Aspekte der Umgebung wie Kontrast, Bewegung oder Farbe informieren. Der volle Umfang der visuellen Informationen, die das Auge an das Gehirn sendet, war jedoch für lange Zeit weitgehend ungelöst.

Katrin Franke erhält den Schilling-Preis der Neurowissenschaftlichen Gesellschaft 2021 für ihre Arbeit zur funktionalen Charakterisierung von Informationskanälen der Netzhaut. Durch die Kombination von funktionellen Populationsmessungen mit maschinellem Lernen zeigte ihre Arbeit, dass das Auge ein viel komplexeres Signal an das Gehirn sendet als bisher angenommen. Weiterhin konnte sie darlegen, wie inhibierende neuronale Schaltkreise für die Erzeugung dieser funktionellen Vielfalt von entscheidender Bedeutung sind. In jüngerer Zeit konzentrierte sich ihre Forschung auf die Frage, wie sich einzelne visuelle Merkmale wie Farbe in den ersten Verarbeitungsstufen des visuellen Systems der Maus entwickeln. Ihre Arbeit hat dazu beigetragen besser zu verstehen, wie die Netzhaut das eingehende Bild verarbeitet, um relevante Aspekte aus der Umwelt zu extrahieren. Ein derart detailliertes Verständnis der gesunden Netzhautfunktion ist der Schlüssel zur Identifizierung von Funktionsdefiziten, welche verschiedenen degenerativen Netzhauterkrankungen zugrunde liegen, und ist damit Voraussetzung für die Entwicklung neuer Behandlungsstrategien.

Katrin Franke promovierte an der International Max Planck Research School for Neural and Behavioral Sciences am Graduiertenkolleg Tübingen. Seitdem ist sie Juniorgruppenleiterin am Bernstein Center for Computational Neuroscience und dem Department für Augenheilkunde der Universität Tübingen, mit längeren Forschungsperioden am Baylor College of Medicine in Houston, USA.

Dieser Preis wird alle zwei Jahre in den ungeraden Jahren durch die Neurowissenschaftliche Gesellschaft e.V. für herausragende Leistungen auf dem Gebiet der Hirnforschung verliehen. Der von der Hermann und Lilly Schilling-Stiftung für medizinische Forschung finanzierte Förderpreis in Höhe von 20.000 Euro soll junge Wissenschaftler*innen bis zu einem Alter von 35 Jahren unterstützen. Voraussetzung für eine erfolgreiche Bewerbung ist eine durch Publikationen dokumentierte hervorragende Forschungsarbeit. Der/die Bewerber*in sollte in einem deutschen Labor arbeiten oder als Deutsche*r im Ausland tätig sein. Die Bewerbung kann entweder direkt oder durch Vorschlag erfolgen. Bewerbungen aus allen Gebieten der Neurowissenschaften sind willkommen. Mitgliedschaft in der Neurowissenschaftlichen Gesellschaft ist nicht Voraussetzung.

Die Preisverleihung erfolgte auf der virtuellen Göttinger Tagung der Neurowissenschaftlichen Gesellschaft am 22. März 2021, die Preisträgerin hielt dort einen Vortrag.

ALBA Declaration für Inklusion und Gleichstellung

ALBA NETWORK
Towards diversity and equality in brain sciences

Die Neurowissenschaftliche Gesellschaft e.V. gehört zu den Unterstützern der ALBA Declaration, mit der sich 143 neurowissenschaftliche Institutionen zu Inklusion und Gleichstellung verpflichten.

Das Alba-Netzwerk wurde 2018 von der ehemaligen FENS-Präsidentin Carmen Sandi, die auch Vorsitzende des Netzwerks ist, und einer internationale Gruppe von Neurowissenschaftlern ins Leben gerufen. Es hat sich zur Aufgabe gemacht, Fairness und Diversität in der Hirnforschung zu fördern. Es soll vermieden werden, dass sich talentierte Forscher*innen in den Neurowissenschaften nicht vollständig entfalten können, weil nicht überall die Prinzipien einer fairen, vorurteilslosen, diversen Arbeitsumgebung geschaffen sind. Dafür möchte das ALBA Netwerk konkrete, evidenz-basierte Handlungsempfehlungen geben. Dazu gehören unter anderem klar definierte Kriterien für Rekrutierung, Anstellung oder Kandidatenauswahl, um versteckte, etablierte Mechanismen von Ungleichheit zu erkennen und zu beseitigen, ein Verhaltenskodex im Umgang mit Mitarbeiter*innen und Kolleg*innen, transparente Karrierestrukturen und die Möglichkeiten für ein ausgewogenes Verhältnis von Arbeit und Familie.

Eine Registrierung als ALBA member ist über die Website des Netzwerks möglich.
Zur Website: http://www.alba.network/network
Kontakt: E-mail: info@alba.network

NEU auf dasGehirn.info

Im Monat Dezember stand erneut der Themenschwerpunkt **Struktur und Funktion neuronaler Netzwerke** im Mittelpunkt. Der zweite Teil der Themenpartnerschaft mit dem SFB 870 steht unter der Überschrift „*Nur eine Zelle*“ und vertieft das Thema mit den folgenden Artikeln:

Erfolgsgeheimnis Kommunikation
Tausende Antennen, ebenso viele Sender und ein superschnelles Datenkabel erlauben Neuronen den gleichzeitigen Informationsaustausch mit vielen anderen Zellen. Dazwischen werden die Signale räumlich und zeitlich präzise verrechnet.

Kommunikationslöcher: Ionenkanäle
In Ionenkanälen verbirgt sich der Schlüssel zum Verständnis der elektrischen Aktivität von Nervenzellen. Ohne sie läuft buchstäblich: Nichts.

Highspeed dank Myelin
Oligodendrozyten umwickeln die Axone von Nervenzellen mit einer isolierenden Myelinschicht, damit elektrische Signale schneller reisen können. Dabei gilt es, die genaue Beschaffenheit der Straßenabschnitte sorgfältig zu planen.

Was Nerven verbindet
Eine Sache der richtigen Vernetzung: Synapsen ermöglichen es uns, flexibel auf Situationen zu reagieren und zu lernen.

Der Animationsfilm **Neuron ≠ Neuron** schaut auf die Biophysik hinter den Neuronen: Um die 86 Milliarden Neurone bevölkern das Gehirn. Sie unterscheiden sich sehr in Form und Funktion – wobei die eine die andere bedingt: Wer, wie die corticale Pyramidenzelle, primär mit Lernen befasst ist, darf sich Zeit lassen. Ganz im Gegensatz zu einer Zelle der oberen Olive, die beim Richtungshören gegen die Zeit läuft.

Das Schwerpunktthema **Sucht** in Themenpartnerschaft mit dem TRR 265, mit dem dasGehirn.info im Mai 2020 online gegangen ist, wurde um ein Video-Interview ergänzt:
Sucht hat viele Facetten. Wir sprechen mit Prof. Andreas Heinz von der Berliner Charité, Sprecher des TRR 265, über verschiedene Arten, die Rolle von Stress, aktuelle Therapien und künftige Möglichkeiten.

In der Rubrik **Neues aus der Wissenschaft** macht dasGehirn.info im Dezember 2020 und Januar 2021 auf die folgenden **Pressemeldungen** aus den Instituten aufmerksam:

- Kartierung der Netzwerkentwicklung im Gehirn | Max-Planck-Institut für Hirnforschung (03.12.2020)
- Wie sich das Gehirn im Raum orientiert | Heinrich-Heine Universität Düsseldorf (08.12.2020)
- Visuelles Kurzzeitgedächtnis ist vielschichtiger als angenommen | Ruhr-Universität Bochum (08.12.2020)
- Menschen hören das, was sie zu hören erwarten | Technische Universität Dresden (09.01.2021)

DasGehirn.info gibt künftig neurowissenschaftlichen Lehr- und Forschungseinrichtungen Gelegenheit, sich und ihre Arbeit in einem Videoportrait vorzustellen. Wie das aussehen kann, zeigt das Portrait von The Center for Integrative Physiology and Molecular Medicine (CIPMM) in Homburg.

Möchten Sie eine Pressemeldung an “dasGehirn.info” weitergeben oder Ihr Institut vorstellen, wenden Sie sich bitte an Arvid Leyh (E-mail: a.leyh@dasgehirn.info).

Neueintritte

Folgende Kolleginnen und Kollegen dürfen wir als Mitglieder der Neurowissenschaftlichen Gesellschaft begrüßen:

Ebru Acun (Magdeburg)
Nikolaos Aggelopoulos, Dr. (Göttingen)
Emad Amini (Würzburg)
Anna Antoniou, Dr. (Bonn)
Gregor Bucher (Göttingen)
Anna-Sophia Hartke (Hannover)
Daniel Hillier (Göttingen)
Hanna Hörnberg, PhD (Berlin)
Alexander Jais (Leipzig)
Lakshay Khurana (Göttingen)
Pia Kruse (Freiburg)
Mirjam Montag (Kaiserslautern)
Sabina Nowakowska (Magdeburg)
Enya Paschen (Freiburg)
Anna Pierzchlinska (Szcecin, Polen)
Nicole Richter (Marburg)
Parthiban Saravanakumar (Magdeburg)

Der Mitgliedsstand zum 20. Januar 2021 beträgt 2.114 Mitglieder.

Ausblick

Martin Korte et al.
Respiratory viral infections and associated neurological manifestations

Valentin Nägerl et al.
Nanoscale imaging of the functional anatomy of the brain

Thomas Bosch et al.
Neurons interact with the microbiome: an evolutionary-informed perspective

Martin Brüne
Mental Health and Biological Evolution: Implications for Psychiatry and Psychosomatic Medicine

NeuroWissenschaftliche Gesellschaft
German Neuroscience Society

Neurowissenschaftliche Gesellschaft e.V. (NWG)

- Beitrittserklärung -

Hiermit erkläre ich meinen Beitritt zur Neurowissenschaftlichen Gesellschaft e.V. (NWG).

Eintrag in das Mitgliederverzeichnis:

Name

Vorname

Titel

Dienstadresse

Universität/Institut/Firma

Straße

PLZ/Ort Land

Telefon/Email

Privatadresse

Straße

PLZ/Ort

Telefon

Rechte und Pflichten der Mitgliedschaft siehe Satzung (nwg-info.de/de/ueber_uns/satzung). Mit meiner Unterschrift bestätige ich, dass ich die Satzung sowie die Datenschutzrichtlinie (nwg-info.de/de/datenschutz) zur Kenntnis genommen habe und diese anerkenne.

Datum/Unterschrift

Ich unterstütze den Antrag auf Beitritt zur NWG e.V.

Datum/Unterschrift des Mitglieds Datum/Unterschrift des Mitglieds

Bitte senden Sie Ihren Antrag an die Geschäftsstelle der NWG:

Stefanie Korthals
Neurowissenschaftliche Gesellschaft e.V.
MDC
Robert-Rössle-Str. 10
13092 Berlin

Email: korthals@mdc-berlin.de
Tel.: +49 30 9406 3127

Ich optiere für folgende 2 Sektionen:

- ☐ Computational Neuroscience
- ☐ Entwicklungsneurobiologie/Neurogenetik
- ☐ junge NWG (jNWG)
- ☐ Klinische Neurowisschenschaften
- ☐ Kognitive Neurowissenschaften
- ☐ Molekulare Neurobiologie
- ☐ Neuropharmakologie und -toxikologie
- ☐ Systemneurobiologie
- ☐ Verhaltensneurowissenschaften
- ☐ Zelluläre Neurobiologie

Ich bin Student ☐ ja ☐ nein
(Bescheinigung anbei)
Ich bin ☐ weiblich ☐ männlich ☐ divers

☐ **Ich erkläre mich einverstanden**, dass meine Daten zum Zwecke wissenschaftlicher Informationsvermittlung (z.B. **FENS-Mitgliedschaft**) weitergegeben werden.
Diese Entscheidung kann jederzeit über die Geschäftsstelle oder das Mitgliederportal auf der Website widerrufen werden.

Jahresbeitrag (bitte ankreuzen):

- ☐ 100,– €/Jahr Seniors (Prof.)
- ☐ 80,– €/Jahr Postdocs (PhD, Dr., etc.)
- ☐ 40,– €/Jahr Studenten, Doktoranden, Mitglieder in Elternzeit oder im Ruhestand, Arbeitslose

Überweisung:
Bankverbindung: Berliner Bank AG
IBAN: DE55 1007 0848 0463 8664 05
BIC: DEUTDEDB110

Einzug über Kreditkarte (VISA/Mastercard):

Kartennr.: ____________________

gültig bis: __________ Betrag: __________

Dreistellige Sicherheitsnr.: __________

Karteninhaber: ____________________

Unterschrift: ____________________

SEPA-Lastschriftmandat:
(Gläubiger-IdentNr: DE64NWG00001110437)

Ich ermächtige die Neurowissenschaftliche Gesellschaft e.V. von meinem Konto

bei der Bank: ____________________

IBAN: ____________________

BIC: ____________________

einmal jährlich den Mitgliedsbeitrag in Höhe von € ______ einzuziehen und weise mein Kreditinstitut an, die von der NWG auf mein Konto gezogenen Lastschriften einzulösen.

Ort, Datum: ____________________

Unterschrift: ____________________

Kontoinhaber: ____________________

Anschrift: ____________________

____