

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



Thomas RECORDING GmbH

www.ThomasRECORDING.com • www.EYE-TREC.com

224 Channel Thomas Matrix

- Up to 224 recording/stimulation channels
- Independently moveable electrode fibers
- Low noise, record while electrodes move
- Cortical and deep brain recordings
- Maximize data output by using tetrodes and/or heptodes
- Minimizes tissue damage due to conical fiber tip shapes
- **Coming soon:**
3D-Reconstruction of neuronal network & autonomous electrode movement



Visit our **websites** for more information or send an **email** to info@ThomasRECORDING.com

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>; Field-Trip, <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

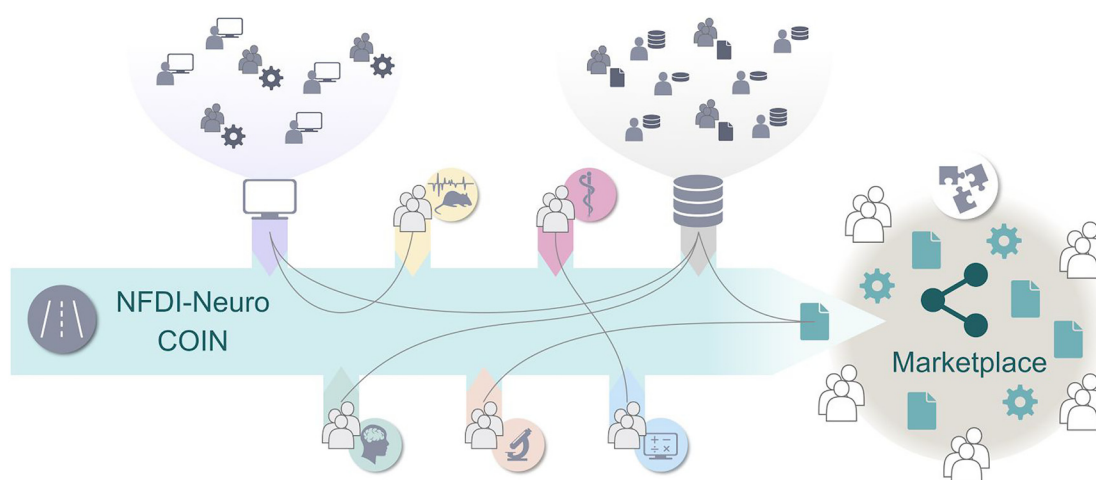


Figure 1: Illustration of the role of the NFDI-Neuro common infrastructure as a technical backbone to connect providers of tools, services and data (top), researchers in different neuroscience domains (colored icons) and the broad community of users and providers via a shared marketplace (right). The 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

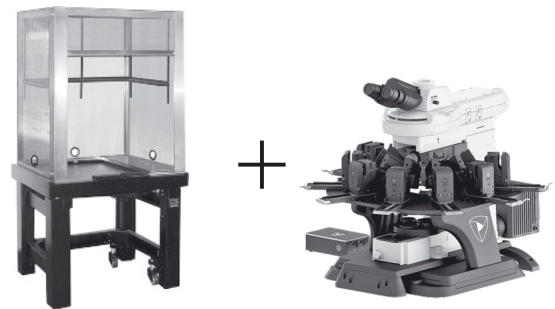
NFDI-Neuro will develop and operate a unique online marketplace for neuroscience research output. This venue will provide a unified, public entry point for discovering primary and derived data, and computational models, for potential consumers from within or outside neuroscience.

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

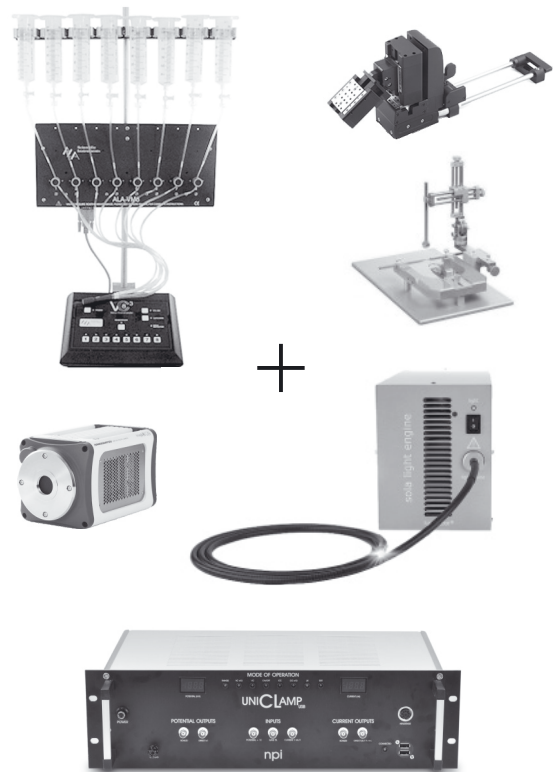
npi
Electronic Instruments
for the Life Sciences

made to measure

Let Us Build Your Rig



Turnkey Solutions for Electrophysiology & Imaging



- Design
- Installation
- Integration
- Support

npi electronic GmbH

Phone +49 (0)7141-9730230; Fax: +49 (0)7141-9730240
support@npielectronic.com; <http://www.npielectronic.com>

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.



WORLD
PRECISION
INSTRUMENTS
Instrumenting scientific ideas

Discover our new Microinjectors

Designed to simplify intracellular injection and a variety of other microinjection tasks, WPI introduces three new Microinjectors for all your Microinjection Applications.

- Convenient internal pressure source (MICRO-ePUMP and uPUMP only)
- Intuitive user-interface
- Small footprint takes up very little bench space
- Easy to navigate with touch screen and control knob
- High precision for reproducible injections – microliters to picoliter



MICRO-ePUMP



PV850



uPUMP

Applications:

- Microinjection of diverse compounds and biomolecules – DNA, RNA, proteins
- Pre- and post-implantation in embryos of various species – mice, rats, monkeys, bovine, pigs, zebrafish, etc.

For more information please visit us at **wpi-europe.com**

World Precision Instruments Germany GmbH Tel +49 (0)6031 67708-0 wpide@wpi-europe.com

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

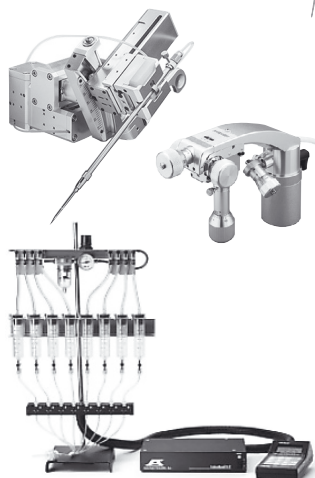
and neurorobotics that are currently being harmonized into a cohesive offering to address cyberinfrastructure-related challenges in brain science, called EBRAINS (<https://ebrains.eu>). The focus of EBRAINS is the development and maintenance of a novel digital research infrastructure with tools and services 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.

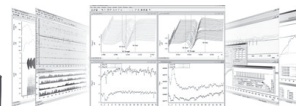
DANDI is a platform for publishing, sharing and processing cellular neurophysiology data, funded by the US BRAIN Initiative. It aims to enable reproducible practices, publications and reuse of data, reduce the need to contact data producers by enriching the data with comprehensive metadata, with the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

SCIENCE PRODUCTS GmbH

for Research in Life Sciences



2-Photon Microscopy, Amplifiers, Micromanipulators, Stimulators & Stimulus Isolators, Electrodes, Data Acquisition & Data Analysis Systems, Microinjection & Perfusion Systems, Temperature Controlling Systems, Stereotaxic Instruments, Micropipette Pullers, Microforges & Bevelers, Glass Capillaries, Wires, Tables & Faraday Cages ...and more!



SCIENCE PRODUCTS GmbH
Hofheimer Str. 63 · D-65719 Hofheim · Germany
Tel.: +49 6192 / 90 13 96 · Fax: +49 6192 / 90 13 98
info@science-products.com · www.science-products.com

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 of 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, Brennekestrasse 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.