NFDI4BIOIMAGE – An Initiative for a National Research Data Infrastructure for Microscopy Data

Christian Schmidt[1] and Elisa Ferrando-May[1,2]

[1]Bioimaging Center, University of Konstanz, Konstanz, Germany
[2]German BioImaging – Gesellschaft für Mikroskopie und Bildanalyse e.V.

Bioimaging and biophotonics are key enabling research technologies in the natural and biomedical sciences. They both foot to a large extent on (light) microscopy, which has transformed from a mainly qualitative observational method to a big-data quantitative approach, as exemplified by automated high-content imaging. Advancements in microscopy instrumentation are achieved at an unprecedented pace driving the production of vast amounts of bioimage data. Novel AI-based bioimage informatics tools are emerging and facilitate knowledge extraction from these highly complex data with high information density. Therefore, image processing and analysis have become an intrinsic and essential component of bioimage-based research. At present, bioimaging data is mainly stored locally and is often not systematically annotated. Proprietary software and heterogeneous file formats impede comparability. Unfolding the full potential of bioimaging and biophotonics requires a culture of image data sharing and re-use that could advance research in multiple scientific disciplines. FAIRification of bioimage data management demands the development and adoption of common standards, harmonizing data handling practices, and extensive training and user education. Leveraging our experience within German BioImaging – Society for Microscopy and Image Analysis (GerBI-GMB) in bringing together microscopy users, IT infrastructure providers, image analysts, and application specialists at core facilities in Germany, we aim to tackle these challenges. We intend to submit our proposal for a consortium within the national research data infrastructure (NFDI) in 2021 to foster a state-of-the-art, high-quality bioimage data management ecosystem in Germany's research data management landscape.

1 Introduction

The national research data infrastructure (NFDI) is currently being established in Germany as a network of closely collaborating consortia. Aims are to manage the scientific

---

Das hier beschriebene Poster ist in der Open Access-Plattform der Universität Heidelberg heiDOK unter der DOI: `https://doi.org/10.11588/heidok.00029489` veröffentlicht.

data systematically, develop standards and solutions for data handling, and provide concepts for long-term storage according to the FAIR principles, i.e., data must be findable, accessible, interoperable, and re-usable [1, 2]. Individual NFDI consortia represent a community of data users and data providers based on a scientific discipline or scientific method. Up to 30 consortia will be incorporated into the NFDI framework in a science-driven process orchestrated by the German Research Foundation, DFG, in three consecutive application calls between 2019 and 2021 [2]. By close collaboration across disciplines, all consortia contribute to research data management standards to the mutual benefit. Basic services relevant for all consortia and common aspects with shared solutions among a subgroup of the consortia will be addressed as cross-cutting topics [3, 4].

Bioimaging methods enable in-depth insight into biological and biomedical samples or materials at high spatial and temporal resolution. Method development and research in bioimage informatics facilitate the extraction of knowledge from images while promoting the advancement of both instrumentation and software. At present, data acquired in bioimaging experiments are often stored locally, and are not systematically annotated and archived according to common standards. The lack of shared research data management standards for bioimaging is also reflected in the often incomplete description of imaging methods in published research articles [5].

Recently established archives for published image data like the BioImage Archive (BIA) hosted at the European Bioinformatics Institute or the Image Data Resource (IDR) showcase the benefit of highly curated and cross-referenced image datasets to the scientific community [6, 7]. However, a large number of often proprietary file formats, insufficient metadata standards, and the need to establish new formats for cloud-based computing and storage impose considerable challenges on the FAIRification efforts for bioimage data [8, 9]. The integration of bioimage (meta)data with multimodal and multidisciplinary data sources, e.g., from established "-omics" research areas like genomics, proteomics, metabolomics, and more, needs to be addressed. Increased data accessibility and interoperability will enable the re-use of bioimage data to extract novel information from existing data of diverse sources. These developments will advance the field of bioimage informatics, including machine-learning algorithms for bioimage analysis and computational modeling.

2 Resources

One important pillar of the NFDI4BIOIMAGE initiative is German BioImaging, a well-organized network of >50 imaging core facilities spread across research institutions. Through this network, the initiative has established contacts with a large number of microscopy users [10]. A significant part of bioimaging data in Germany is acquired at or in close collaboration with core facilities, which puts NFDI4BIOIMAGE in an excellent position to survey community needs and distribute novel standards and solutions for bioimage data management. The consortium has both a methodological and disciplinary scope, the latter comprising the field of bioimage informatics. Task Areas will be devoted to, e.g., (meta)data standards and formats, including the development of next-generation file for-

mats to overcome current limitations in data handling and transformation [8]. The team includes among others leading experts in software engineering specialized in bioimage data formats and data management tools. Furthermore, NFDI4BIOIMAGE leverages the work of the Research Data Management for Microscopy (RDM4Mic) group, which is part of German BioImaging and has a long-standing collaboration with the Open Microscopy Environment (OME) consortium. OME is home to OMERO (OME Remote Objects), the most widely used management platform for microscopy data. Thus, NFDI4BIOIMAGE will promote OMERO as one possible solution for FAIR bioimage data handling. In particular, common metadata recommendations and best practices for running OMERO instances at local institutions will be addressed [11]. Members of yet another consortium, the Quality Assessment and Reproducibility for Instruments and Images in Light Microscopy (QUAREP-LiMi) group [12], will contribute to NFDI4BIOIMAGE with work on quality criteria for bioimage acquisition and their representation in metadata. This work includes collaboration with commercial vendors. With members in Euro-BioImaging and Global BioImaging, the NFDI4BIOIMAGE initiative has further well-established international connections.

Task Areas on technical infrastructures and on data linking & multimodal data integration will guide the choice and configuration of local and decentralized hardware. They will work towards increased interoperability of microscopy data with other data types and platforms. Integration of these solutions with laboratory information management systems and electronic lab notebooks is another topic dealt with by the consortium. We also build on established connections to the IDR and BIA and will promote user-friendly workflows and guidelines to connect these repositories to national storage and archiving solutions. A Task Area on bioimage informatics aims to establish user-friendly tools and interfaces for FAIR bioimage analysis, enabling researchers from all disciplines to use state-of-the-art tools and software.

The resources mentioned above enable the initiative to focus on all aspects of the bioimaging data life cycle from experiment planning to data acquisition, annotation and storage, image analysis, data processing, and finally to publication and archiving. Ultimately, the re-use and mining of existing data should be routinely considered when planning to conduct a bioimaging experiment and should belong to the standard skills of researchers in the life sciences. NFDI4BIOMAGE aims to become a reliable resource for bioimage data users, data generators, and data stewards, and will do so by making services and solutions sustainable and openly available to the whole community. Notably, the consortium will generate benefits for the NFDI as a whole by collaborating with discipline-specific NFDI consortia and contributing to cross-cutting topics.

3 Support

German BioImaging has been funded by DFG from 2012 to 2017 as a scientific network. In 2017, it was transformed into a scientific society with the legal form of a non-profit association. A project for bioimage data management within the DFG funding line In-

formation Infrastructures for Research Data initiated by members of German BioImaging has been recently granted funding. It represents a first step for developing FAIR bioimage data management practices based on the platform OMERO. The project, termed I3D:bio (Information Infrastructure for BioImage Data) will start at the beginning of 2022. Fostering good practices of research data management and open science is a shared vision of the participating institutions.

4 Conclusions

The National Research Data Infrastructure paves the way for a community-driven, collaborative effort to facilitate research data management in Germany and network data internationally. We aim to make a substantial contribution within the framework of the NFDI with our focus on microscopy, biophotonics, and bioimage analysis as a consortial initiative applying in the third call for proposals in 2021. We welcome interested researchers, IT professionals, and, in general, people from the community and NFDI consortia to get in contact and collaborate on the aims of NFDI4BIOIMAGE, provide input and feedback. Please visit https://nfdi4biomage.de for further information.

Acknowledgements

Bibliography

[1] Wilkinson, M.D., et al.: The FAIR Guiding Principles for scientific data management and stewardship. Sci Data, 2016. 3: p. 160018. `https://doi.org/10.1038/sdata.2016.18`

[2] Bund-Länder-Vereinbarung zu Aufbau und Förderung einer Nationalen Forschungsdateninfrastruktur (NFDI) vom 26. November 2018, in BAnz AT 21.12.2018 B10. 2018: Germany.

[3] The development of the National Research Data Infrastructure (NFDI). Second statement of the NFDI Expert Committee. 2020. URL: `https://www.dfg.de/download/pdf/foerderung/programme/nfdi/stellungnahme_nfdi_201112_en.pdf` (last visit: May 10th, 2021)

[4] Bierwirth, M., et al.: Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung. Zenodo, 2020. `http://doi.org/10.5281/zenodo.3895209`

[5] Marques, G., T. Pengo, and M.A. Sanders, Imaging methods are vastly underreported in biomedical research. Elife, 2020. 9.e55133 `https://doi.org/10.7554/eLife.55133`

[6] Ellenberg, J., et al.: A call for public archives for biological image data. Nat Methods, 2018. 15(11): p. 849-854. `https://doi.org/10.1038/s41592-018-0195-8`

[7] Williams, E., et al.: The Image Data Resource: A Bioimage Data Integration and Publication Platform. Nat Methods, 2017. 14(8): p. 775-781. `https://doi.org/10.1038/nmeth.4326`

[8] Moore, J., et al.: OME-NGFF: scalable format strategies for interoperable bioimaging data. bioRxiv, 2021. `https://doi.org/10.1101/2021.03.31.437929`

[9] Swedlow, J.R., et al.,: A Global View of Standards for Open Image Data Formats and Repositories. Nature Methods, 2021, `https://doi.org/10.1038/s41592-021-01113-7`

[10] Ferrando-May, E., et al.: Advanced light microscopy core facilities: Balancing service, science and career. Microsc Res Tech, 2016. 79(6): p. 463-79. `https://doi.org/10.1002/jemt.22648`

[11] RDM4Mic. 2021; URL: `https://german-bioimaging.github.io/RDM4mic.github.io/` (last visit: Aug 9th, 2021)

[12] Nelson, G., et al.: QUAREP-LiMi: A community-driven initiative to establish guidelines for quality assessment and reproducibility for instruments and images in light microscopy, J Microsc. 2021 Jul 2. doi: 10.1111/jmi.13041. Online ahead of print