EBRAINS FAIR data services SC1 - status at the end of SGA3 (D4.12- SGA3)

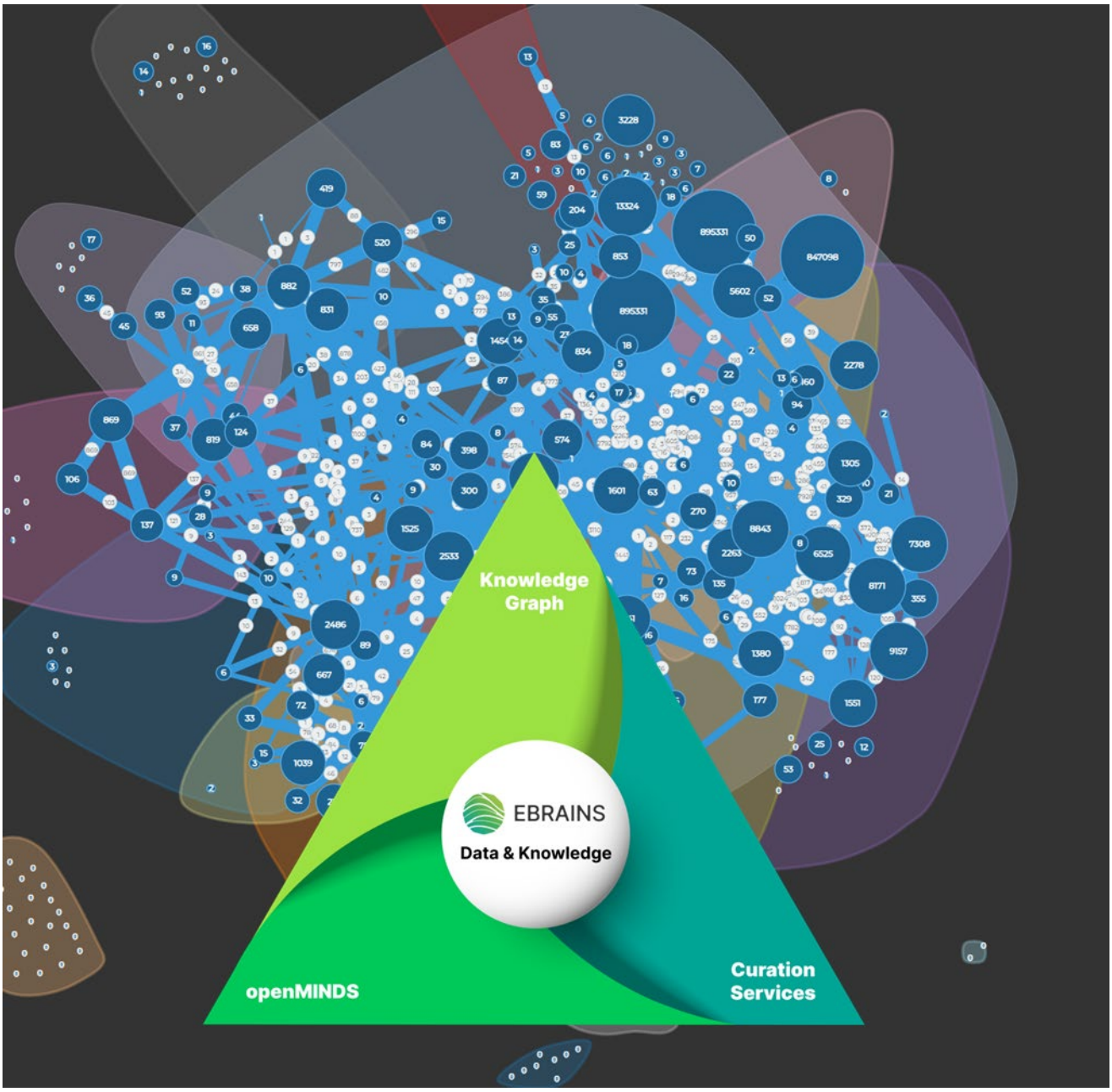

Figure 1: EBRAINS Data and Knowledge Service

The EBRAINS Data and Knowledge Services integrate essential components behind the EBRAINS data sharing and discovery platform. The services enable the broader neuroscience community to find, share, interpret and reuse neuroscience data, models, and software. The three main pillars that constitute the EBRAINS Data and Knowledge Services, as elaborated on in this report, are represented by the Knowledge Graph, the openMINDS metadata framework, and the Curation Services.

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	D4.12 EBRAINS FAIR data services (SC1) status at the end of SGA3
Document Filename:	D4.12 (D43) SGA3 M42 SUBMITTED 231002.docx
Deliverable Number:	SGA3 D4.12 (D43)
Deliverable Type:	Other
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M42 / 30 SEP 2023
Actual Delivery Date:	SGA3 M43 / 02 OCT 2023
Author(s):	Jan G. BJAALIE, UIO (P81), Oliver SCHMID, EBRAINS (P1), Archana GOLLA, UIO (P81), Eszter A. PAPP, UIO (P81), Ingrid REITEN, UIO (P81), Lyuba ZEHL, JUELICH (P20) Mathew ABRAMS, KI (P37), Simisola AKINTOYE, DMU (P16), Damian EKE, DMU (P16), William KNIGHT, DMU (P16), Andrew DAVISON, CNRS (P10)
Compiled by:	Martha Elisabeth BRIGG, UIO (P81)
Contributor(s):	
WP QC Review:	Timo DICKSCHEID, JUELICH (P20)
WP Leader / Deputy Leader Sign Off:	Jan G. BJAALIE, UIO (P81)
T7.4 QC Review:	Formatted by Annemieke MICHELS, EBRAINS (P1)
Description in GA:	Knowledge Graph release with improved functionality and extended content, and updated inventory of related FAIR data, tools and services released through the EBRAINS portal.
Abstract:	The EBRAINS Data and Knowledge services are inclusive, accommodating data from all areas of neuroscience, thus serving the broader neuroscience community interested in adhering to the FAIR principles. The services provide robust support for data curation, comprehensive domain-specific metadata, and user-friendly interfaces and APIs that enhance discoverability. Simultaneously, they uphold responsible data compliance, data protection protocols, and governance mechanisms. The core components — the Knowledge Graph, the openMINDS metadata framework, and the curation workflow — have reached a high level of maturity and usability. As of September 2023, over 2000 researchers have contributed research data, and a continuous influx of research data is being received, including contributions from sources beyond the Human Brain Project. Some of this is driven by journal recommendations. Integration with other EBRAINS services presents novel opportunities for data reuse and advanced analysis, offering unique possibilities for exploring new combinations of data. This document is a follow up from Deliverable D4.1(D32), outlining the status of the EBRAINS Data and Knowledge services before October 2020, and D4.7 (D38), outlining progress between October 2020 and October 2021.
Keywords:	Data sharing; data curation; data management; infrastructure; data compliance, ethics, ontology, EBRAINS.

Target Users/Readers:	Computational neuroscientists, HBP/EBRAINS users, Consortium members, funders, general public, policymakers, students.

Table of Contents

1. Introduction 4
2. Improvement and growth of services since December 2021 (M19) 4
2.1 EBRAINS Knowledge Graph 5
2.1.1 Key components 5
2.1.2 Associated Tools 7
2.1.3 User benefit 7
2.2 EBRAINS Curation 7
2.2.1 Migration to openMINDS 8
2.2.2 openMINDS 8
2.2.3 Ontologies 9
2.2.4 Curation workflow 10
2.2.5 User benefits 11
2.3 Human Data Gateway 12
2.3.1 User benefits 13
2.4 Live papers 13
2.4.1 User benefits 15
2.5 EBRAINS Compliance Management, Data Management, and Data Governance 15
2.5.1 User benefits 16
2.6 KnowledgeSpace 16
2.6.1 User benefits 18
2.6.2 Work in progress 18
3. Education, Outreach, and other interaction with users of the services 19
3.1 EBRAINS Knowledge Graph 19
3.2 EBRAINS curation 19
3.3 EBRAINS Compliance Management, Data Protection, and Data Governance 21
3.4 EBRAINS KnowledgeSpace 21
4. Future Perspectives 21

Table of Tables

Table 1: Development between October 2021 (M19) and September 2023 (M42) of key components of the EBRAINS Knowledge Graph 5

Table of Figures

Figure 1: EBRAINS Data and Knowledge Service 1
Figure 2: Increase in number of contributors, datasets, models, and software 8
Figure 3: Data Sharing Collab 10
Figure 4: Distribution of experimental modalities across datasets shared via the HDG 12
Figure 5: A dialog box from the Live Paper Builder app. 13
Figure 6: Lists of resources in a published live paper, with buttons to open interactive viewers. 14
Figure 7: KnowledgeSpace metadata schema. 17
Figure 8: KnowledgeSpace (KS) features. 18

1. Introduction

The EBRAINS FAIR data services, referred to as the EBRAINS Data and Knowledge services, were initially introduced on the Human Brain Project website in April 2018. Subsequently, in October 2019, the EBRAINS web portal was launched with these services incorporated into the new portal.

This report is the third and final in a series serving to provide an overview of the status of the EBRAINS Data and Knowledge services. The second report (SGA3 D4.7) included a progress report for the period from November 2020 to October 2021, whereas the present report summarizes progress since October 2021.

The first report (SGA3 D.4.1, October 2020) emphasized how the EBRAINS Data and Knowledge services had been constructed around the EBRAINS Knowledge Graph, how the FAIR (Findable, Accessible, Interoperable, and Reusable) principles for data sharing had been implemented, and how scientists could easily share, find, and access data through intuitive interfaces and APIs. The curation process in 2020 was primarily manual but well developed, delivering metadata enrichment and quality assurance. The metadata schemas were referred to as HBP-MINDS, later to be considerably updated and expanded as the openMINDS framework. The services reported on were:

- EBRAINS Knowledge Graph
- EBRAINS Curation
- EBRAINS Compliance Management, Data Protection, and Data Governance
- KnowledgeSpace

The second report (SG3 D4.7, October 2021) provided information on two new services:

- Human Data Gateway
- Live papers

The Human Data Gateway was developed to facilitate safeguarded sharing of highly pseudonymised human data. While metadata are openly shared in the Knowledge Graph to make the data easily discoverable, this new service offered controlled access to the data, subject to specific requirements that data consumers had to adhere to. The Live Papers were introduced as a measure to facilitate the interactive exploration and re-use of datasets, models, and code in the context of published articles with which they are associated.

In addition, the report emphasised the large improvements of the previously established services, the user benefits resulting from these improvements, the large growth of content, and the extensive community engagement activities serving to promote Open Science in general and the EBRAINS Data and Knowledge services in particular.

Building on the two previous reports, the present report describes the enhancements implemented in the above-mentioned services since October 2021, including advantages experienced by users, the continuing substantial expansion of content and growth in number of contributors, and the broad range of community engagement initiatives.

2. Improvement and growth of services since December 2021 (M19)

The following sections describe the changes implemented in each of the services since October 2021 (M19), as well as the user benefits resulting from these changes.

2.1 EBRAINS Knowledge Graph

The EBRAINS Knowledge Graph (KG)[1] is the metadata management system of the EBRAINS Data and Knowledge services. It provides fundamental services for making neuroscientific data, models, and related software FAIR and is the central storage system of the collected and curated metadata in EBRAINS. Being a multicomponent solution, the Knowledge Graph allows the data-consuming end user to access the registered meta-data via an intuitive user interface (Knowledge Graph Search[2] or programmatically via a REST API[3] supported by additional convenience UIs (Knowledge Graph Query Builder[4]).

For (meta-)data producers and curators, the KG provides (besides the API for automation and mass-manipulation of meta-data) a management user interface allowing users without programming skills to directly manage their metadata on the graph with the Knowledge Graph Editor[5]. This includes the support for publication workflows and fine-granular permission management.

Although driven by the challenging metadata management needs of the Human Brain Project because of its heterogeneity of data, the technical solution of the various components of the EBRAINS Knowledge Graph has been designed from the beginning to be generic and flexible enough to be applied in any scientific or non-scientific metadata management.

Being a central underlying component for various of EBRAINS services, the focus of the Knowledge Graph development since the end of 2021 has been set on sustainability and therefore to further enforce the reliability of the service (improvements of the codebase, deployment pipelines, application monitoring, error reporting, documentation, and performance / scalability). Additionally, several steps have been taken to improve the usability for end users and to onboard new user groups with the existing services (e.g. data providers adding their metadata directly in KG Editor). The KG Search has been further extended to expose the rich metadata of openMINDS in appealing and informative representations and extended by additional types and structures.

2.1.1 Key components

Table 1 shows key components of the EBRAINS Knowledge Graph, with a description of changes and improvements since 2021.

Table 1: Development between October 2021 (M19) and September 2023 (M42) of key components of the EBRAINS Knowledge Graph

Key Component	Description of changes / improvements
Documentation (T4.1 & T4.2) https://docs.kg.ebrains.eu	An automated pipeline aggregates the documentation from the various software components to a central documentation platform. Another automated pipeline allows to automatically compose video tutorials out of individual snippets for being able to quickly update to minor UI changes.
Operations (T4.2)	The previous version of the KG (v2) has been successfully deprecated and shutdown. All dependencies in the KG components have been cleared from the code base. The maintenance pipelines have been optimized and stabilized. Additional monitoring, load testing, error reporting and code analytics have been integrated to further improve the quality of services.

[1] https://kg.ebrains.eu/
[2] https://search.kg.ebrains.eu/
[3] https://core.kg.ebrains.eu/
[4] https://query.kg.ebrains.eu/
[5] https://editor.kg.ebrains.eu/

Core (T4.1) https://core.kg.ebrains.eu	The KG API has been stabilized, finalised, and released in v3. The old v2 was deprecated and eventually - after a successful client migration - shut down. Several performance issues, especially for complex metadata structures, have been resolved and the spatial search capabilities have been extended. A new, virtual space called "review" allows end users to validate complex metadata structures in the KG Editor. The documentation[6] of the KG Core has been rewritten and its docker images – next to the already public source code[7] – have been publicly exposed at https://docker-registry.ebrains.eu/ New, autogenerated Python and TypeScript based SDKs[8] (API wrappers with convenience functionalities and authentication processes) have been published and released in PyPi[9] and NPM[10]
Search UI (T4.2) https://search.kg.ebrains.eu	In combination with the openMINDS migration, the KG Search has been equipped with improved and extended visualizations for content types, dataset versions (e.g. filter mechanisms and grouping by file bundles, display of cross-links between resources, hierarchical representation of specimen), citation (e.g. download of bibtex) and more. Also, new representations have been added such as "Behavioural protocol," "Workflow," "Brain atlas," "Webservice." User feedback and other feature requests have been actively user tested and resulted – amongst others - in the replacement of the fuzzy full-text search with a "did you mean" suggestion mechanism. Highlighting functionalities such as "Top trending" and "New" have been introduced and performance optimizations have been implemented.
Editor (T4.1) https://editor.kg.ebrains.eu	The usability of the KG Editor has been improved by various mechanisms (additional information for suggestions, moving instances between spaces, validation mechanisms, etc.). Adaptations to achieve an integration into the collab-based curation workflow have been made. Amongst the changes, the KG Editor has been integrated as a Collab community app. Additionally, the KG Editor is now capable of providing minimal support instances of unknown types making it a complete viewer of the overall graph structures.
Query Builder (T4.1) https://query.kg.ebrains.eu	A complete redesign of the UI layout and its flow has improved usability and made the query builder accessible for even more users without a lot of technical background. A comprehensive video tutorial to introduce the query builder has been produced and linked in the UI to make the usage of the application even more comprehensible.
Statistics (T4.1) https://stats.kg.ebrains.eu	No further extensions have been applied for KG Statistics except for maintenance work (keeping dependencies up to date).
Automation (T4.1)	No major extensions alongside maintenance work (keeping dependencies up to date), bug fixing and adaptation to new KG Core Python SDK (see above).
Metadata monitoring (T4.1)	The metadata monitoring has been extended to support the new extended curation workflow based on collabs. Additionally, workflows such as for the approval of ethics and the internal reporting of views and downloads have been integrated making the metadata monitoring board a central information place for the curation process.

[6] https://docs.kg.ebrains.eu/
[7] https://github.com/HumanBrainProject/kg-core
[8] https://github.com/HumanBrainProject/kg-core-sdks
[9] https://pypi.org/project/ebrains-kg-core/
[10] https://www.npmjs.com/package/@ebrains/kg-core

2.1.2 Associated Tools

The integration with various existing EBRAINS services such as EBRAINS atlas viewers and Jupyter Notebooks, with the EBRAINS KG have been maintained, partially enriched, and successfully migrated to the new KG version. Thanks to the stabilized API and the corresponding improved toolset with the language specific SDKs, the use by the community has been extended: more and more tools and adapters have been built to interact/rely on the information either directly or indirectly via siibra-python[11] and fairgraph[12]. One example of such a tool is DataLad[13]. Additionally, new services were integrating with the EBRAINS KG (e.g. TVB[14], NeuroCONNECT[15]) in a server-to-server manner.

2.1.3 User benefit

The developments since October 2021 have resulted in several improvements for the end users.

A key achievement is the finalisation and release of the Knowledge Graph API v3 in combination with the availability of rich metadata thanks to the openMINDS structures. Thanks to the more detailed and precise information provided by the new metadata structures, we were able to significantly improve the representations in the KG Search and to expose the power of the underlying API for the end users. Integrations into workflows such as the collab-based curation workflow can profit from the dynamicity of the tools built around the core functionality of the KG and expose it to even more users – and potentially non-expert users - whilst still providing a rich set of capabilities.

Last but not least, the users profit from a system which has been proven to be production-stable for several years and acts as a reliable centralised metadata system for neuroscientific needs.

2.2 EBRAINS Curation

The role of the EBRAINS curation service is to facilitate and guide the publication of research data, models, and software in the EBRAINS Knowledge Graph according to the FAIR guiding principles (Wilkinson et al. 2016[16]). In particular, curators ensure that the submitted research products are accurately annotated with metadata in line with the openMINDS standard, which makes the data discoverable via the EBRAINS KG search, and in turn also indexed by public search engines. Data files registered in the EBRAINS Knowledge Graph are integrated with relevant analysis software according to their content type, enhancing their reusability and interoperability. The curation team also helps data providers create human-readable summaries of their data, so-called Data Descriptors, to give a good starting point for researchers aiming to reuse the data.

In the course of HBP SGA3, the overall number of datasets, models, and software curated and integrated in the EBRAINS Knowledge Graph has continued to increase (Figure 2). In addition to research products generated by the Human Brain Project, we have received a steady stream of external research products curated and published on EBRAINS. We have also seen a substantial increase in the number of researchers choosing to share data on EBRAINS. Almost half of the researchers who submit their data to EBRAINS curation have been referred to EBRAINS by a colleague. We attribute this growth to the increased visibility of the EBRAINS website, public dissemination, and outreach activities by the curation team and the HBP communications team, as well as EBRAINS being a recommended data sharing repository by publishers such as Springer Nature and Frontiers.

[11] https://github.com/FZJ-INM1-BDA/siibra-python
[12] https://github.com/HumanBrainProject/fairgraph
[13] https://www.datalad.org/
[14] https://thevirtualbrain.org/
[15] https://www.biomax.com/research/brain-science/
[16] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

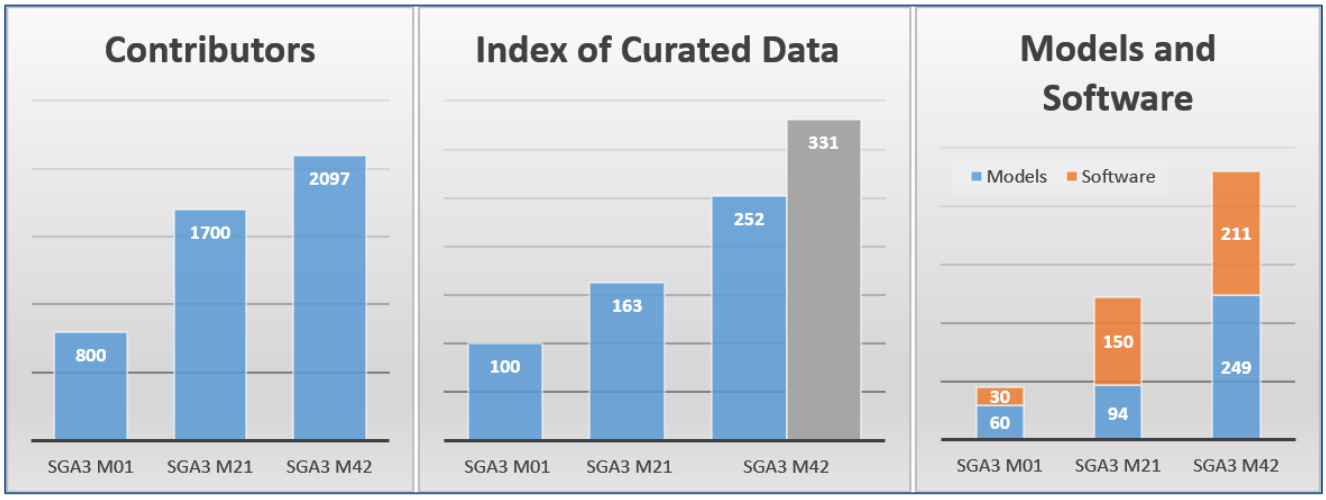

Figure 2: Increase in number of contributors, datasets, models, and software

The graphs reflect the increasing number of contributors and research products published in the EBRAINS Knowledge Graph from M01 (April 2020) to M42 (September 2023). The Index of Curated data represents the relative percent-wise increase in the number of subjects and samples from M01. The grey column at SGA3 M42 shows the increase when also including derived data (reuse of subjects and samples in new datasets).

The routines and workflows employed in the EBRAINS curation service are continuously adjusted to accommodate user feedback. Besides extensive communication with data providers who use the data sharing service, a user satisfaction survey is sent out to all researchers upon the successful publication of their data on EBRAINS. New developments and optimisations of the service are based on these interactions, as well as input from the broader research community acquired through extensive communication and dissemination activities (see Section 3). Where applicable, community feedback is discussed with the EBRAINS Knowledge Graph team and the openMINDS development team to further improve the overall user experience of the EBRAINS infrastructure.

2.2.1 Migration to openMINDS

The metadata representations of all research products (datasets, models, software) shared through the EBRAINS Knowledge Graph (KG) prior to April 2021 were successfully migrated to openMINDS (section 2.2.2) in the past reporting period.

The migration included a re-curation of the research products to harmonise their representation and increase their integration in the EBRAINS KG. The increase in consistency of metadata registrations in the KG led to improved visualisations in the EBRAINS KG Search and facilitated the interoperability to other EBRAINS services (e.g. the EBRAINS Atlases).

Moreover, the usage of the well-defined, ontology-driven openMINDS terms (e.g., age categories, species) and constructs (e.g. brain atlases, common coordinate spaces) strengthened the alignment of EBRAINS to other data sharing and metadata standardisation efforts (section 2.2.3).

The migration to openMINDS also allowed the registration of new research products in the EBRAINS KG, including now data/metadata models, brain atlases, common coordinate spaces, web services, live papers, and workflow recipes. Moreover, openMINDS allows now to provide in parallel general as well as specialist (in-depth) metadata descriptions of these research products within the same graph database allowing for an integrative registration process to stepwise increase the FAIRness of shared data.

2.2.2 openMINDS

openMINDS is an open-source metadata framework for graph database management systems, such as the EBRAINS Knowledge Graph (KG). The framework consists of (i) metadata schema specifications for an extendible set of metadata models, (ii) libraries of well-defined, serviceable metadata instances, and (iii) supportive tooling for handling openMINDS in Python or MATLAB.

(i) The openMINDS schema specifications are developed in a customized syntax which is easier to read than most formal schema languages and allows the inclusion of machine-readable configuration hints for data management systems. Schema specifications are organized into interlinked metadata models to facilitate domain-specific development and maintenance, as well as modular extensions. openMINDS v3.0 includes the following metadata models (essentials in bold; extensions in italic):

+ *chemicals* (focus: description of chemical compounds and mixtures)

+ **controlledTerms** (focus: description of common terminologies)

+ *computation* (focus: description of computational workflows)

+ **core** (focus: description of general origin, location, and content of data)

+ *ephys* (focus: description of electrophysiology experiments)

+ *publication* (focus: description of publication resources and live papers)

+ *SANDS* (focus: description of anatomical location of data, and brain atlases)

+ *specimenPrep* (focus: description of specimen preparation procedures)

+ *stimulation* (focus: description of stimulation processes)

(ii) The provision of serviceable metadata instances that are well-defined and linked to other standardization efforts (e.g. ontologies) increases the overall integration of data within and beyond the data management system. openMINDS maintains extendible libraries of such instances for various common terminologies (e.g., age categories, species, etc.), licenses, content types (software specific file formats), or complex constructs, such as brain atlases or common coordinate spaces.

(iii) To facilitate the preparation of openMINDS compliant metadata collections independent of a data management system, openMINDS offers small, management-systems independent Python and MATLAB packages to handle schema specifications and instances. Within EBRAINS, management-system specific tooling for openMINDS was developed by the EBRAINS KG and Curation team (e.g. the EBRAINS KG Editor, the Python package fairgraph).

Besides being adopted by EBRAINS, the openMINDS framework has been implemented by The Virtual Brain (TVB) and NeuroCONNECT by the Labvantage-Biomax GmbH. Furthermore, openMINDS is under consideration to serve as the metadata framework for the Karolinska Institute Centre for Imaging Research, The BrainMINDS large-scale brain project in Japan, and the Duchenne Data Foundation.

To foster international awareness and adoptions, all openMINDS developments has recently been moved to the newly founded GitHub organization "openMetadataInitiative[17]." As part of the moving, the openMINDS build pipeline was revised to better accommodate the flexible modularity of the framework. The main building process of the openMINDS schemas is now reintegrated into the main repository[18] and separated from continuative building processes of other schema formats (e.g., JSON-Schema[19]). In addition, the development and provision of instances is separated from the schema specifications[20] to facilitate their openMINDS version-specific maintenance. Finally, the documentation was revised and centralized to be available on "Read the Docs[21]".

2.2.3 Ontologies

The ontologies reported in Deliverable D4.1 of the HBP SGA3 Grant Agreement continue to be maintained in a collaboration between the EBRAINS teams and US Neuroscience Information Framework (NIF), thereby facilitating improvements for use by EBRAINS as well as alignment of terminologies with other projects (e.g., the US BRAIN initiative). The Atlas Ontology Model (AtOM) has been published and released (Kleven et al, 2023, P3987). Current efforts focus on integrating

[17] https://github.com/openMetadataInitiative
[18] https://github.com/openMetadataInitiative/openMINDS
[19] https://github.com/openMetadataInitiative/openMINDS_json-schema
[20] https://github.com/openMetadataInitiative/openMINDS_instances
[21] https://openminds-documentation.readthedocs.io/en/latest

various brain atlases into the EBRAINS KG through the openMINDS metadata model interpretation of AtOM. As a new project, the ontology engineering team is currently developing an ontology for behavioural tasks.

2.2.4 Curation workflow

The procedures and workflows for sharing of research products have been further standardised and improved, both from the user's and the curator's perspective. Below, we outline developments that have enhanced the user experience when it comes to sharing data on EBRAINS.

We have automated parts of the submission process by generating dataset-specific collabs (Figure 3) for sharing data and metadata with the curation team. The collab gives the data provider guidance for each step in the data curation process, thus reducing the need for e-mail contact with the curator. A progress indicator in the collab informs the data provider about which steps are completed and approved and which remain open. Upon the acceptance of the curation request, the data provider receives an email from the curator with the link to the data sharing collab for the specific dataset. From that point, the data provider can manage the submission of metadata, data descriptor and data in the order they prefer. The curator is available for support throughout the process.

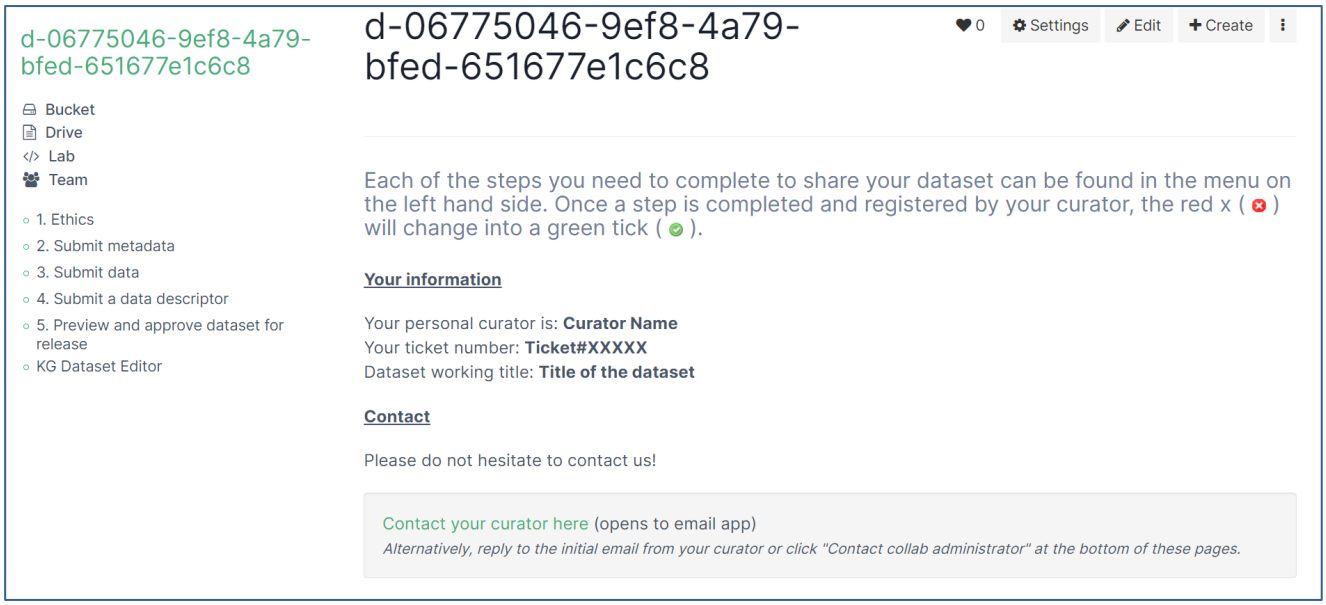

Figure 3: Data Sharing Collab

Users are provided with guidance on individual curation steps and can monitor the progress of the curation process using the data sharing collab.

The process of submitting metadata has been streamlined through the creation of the EBRAINS Metadata Wizard (https://github.com/HumanBrainProject/ebrains_wizard), a JavaScript tool designed for the submission of openMINDS-compatible metadata to the EBRAINS platform. This tool simplifies the metadata submission process by guiding users through the inclusion of all essential information required for publishing a dataset on EBRAINS. Additionally, it encourages the inclusion of recommended metadata elements to enhance data discoverability and adherence to FAIR principles. The Metadata Wizard clearly distinguishes between mandatory and optional fields, reducing the effort required for data providers who wish to provide only the essential metadata. Users can conveniently download or upload previously submitted metadata and save partially completed submissions, provided they are using the same computer and browser. Upon submission, the completed form is automatically transmitted to the curation team for further processing.

Metadata can also be submitted through interaction with the EBRAINS Knowledge Graph (KG) via a private workspace. Access to this space is provided on an individual basis by a curator via the dataset-specific collab (mentioned above). Within this private workspace, users can interact with uploaded metadata and establish connections between the metadata and existing publicly accessible entries in the Knowledge Graph.

The EBRAINS data descriptor template underwent modifications with the primary goal of providing improved guidance to data providers in creating more informative dataset descriptions. The updated template (version 3) was taken into use in November 2022. Notable changes include a reduction in the number of sections and a reorganization to emphasize four main thematic areas: i) summary (concise overview of the dataset), ii) materials and methods (details about the materials used and the methods employed), iii) usage notes (suggestions and recommendations for utilizing the dataset effectively) and iv) data records (description of the dataset contents and their organization within the repository is outlined).

The data proxy is continuously improved, resulting in a more stable service for researchers uploading data to EBRAINS. We have also developed an interactive python script which facilitates programmatic upload of files.

The curation workflow has been optimised to ensure a closer alignment with community efforts. We have assessed how well our data sharing services follow the FAIR guiding principles, available via https://wiki.ebrains.eu/bin/view/Collabs/data-curation/#HEBRAINScommitstotheFAIRprinciples. Also, to ensure better coupling to the BIDS standard, we have introduced the requirement to complete a BIDS validation prior to registering the dataset as a dataset following BIDS standards.

Documentation Resources for disseminating information about the data sharing process to data providers and data users have improved. The public data curation collab (https://wiki.ebrains.eu/bin/view/Collabs/data-curation) was updated and holds now an up-to-date and more in-depth description of the data sharing process, including a step-by-step procedure for sharing of datasets, models and software, elaboration of additional options for research products shared in the KG, a contact section telling more about the team behind the curation, and a FAQ section. Additionally, the ebrains.eu entry for the Share data service (https://ebrains.eu/share-data) is revised with the launch of the new web portal.

The (meta)data migration to openMINDS improved data integration within the EBRAINS KG, facilitating the discovery of cross-relations between different research products and their versions. Moreover, the usage of well-defined instances of the openMINDS libraries harmonised the metadata description across research products within the EBRAINS KG, and established connections to metadata/data of other standardisation (e.g. ontologies) and data sharing efforts (cf. Knowledge Space). Through the flexibility of openMINDS, general and specialist (in-depth) data registration can now be conducted in parallel in the EBRAINS Knowledge Graph facilitating the modularisation of the registration process and the curation workflow.

2.2.5 User benefits

Easier to find information: With the recent improvements, it is easier for the data provider to find information about what is required for sharing research products via EBRAINS, and the process of sharing is simplified. For sharing of datasets, the full process can now be managed and tracked by the data provider in the dataset specific data sharing collab. This gives freedom for the data provider to choose which steps to focus on when, and to complete the data sharing at his/her own pace. It also reduced the need for email communication with the curator, making the process more efficient.

Enhanced Metadata with openMINDS: The updated openMINDS framework, alongside new ontology models, has significantly enriched the metadata associated with the research products. This expansion means that users now have access to a more comprehensive set of information about the data, models, and software shared through the EBRAINS Data and Knowledge services. This detailed metadata provides valuable context and improves the usability and interpretability of the shared resources. Researchers can better understand the nature of the data, its source, and its relevance to their work, making it easier to locate and utilize relevant information.

Interoperability Across Repositories: The preparations made for adoption of openMINDS by other brain initiatives and data repositories fosters greater interoperability across different platforms. This means that data shared in diverse repositories becomes more comparable and interpretable for users. This interoperability promotes collaboration, reduces data fragmentation, and streamlines the research process by eliminating the need for researchers to navigate multiple data formats and structures.

Efficiency through more Automated Curation: The more automated and well-documented curation workflow represents a significant efficiency gain for data providers. By automating certain aspects of the curation process, data providers can save time and effort in preparing and submitting their research data. This streamlining of the curation workflow ensures that data is processed and made available to users more quickly and efficiently. It also reduces the potential for errors, ensuring the data's quality and reliability meet the highest standards.

Overall, these advancements contribute to a more user-friendly and efficient environment for users who can more easily discover, access, and leverage neuroscience data, ultimately accelerating progress in the field.

2.3 Human Data Gateway

The Human Data Gateway (HDG) is an extension of the EBRAINS data sharing services for making strongly pseudonymised human data available to registered users from recognised academic institutions upon acceptance of the EBRAINS Data Use Agreement (https://ebrains.eu/terms#data-use-agreement). While the data themselves are protected by the HDG, where user access is time-restricted and tracked, metadata such as experimental methods, study target, and affected brain regions are publicly available on EBRAINS and help make the data discoverable to the public.

The Human Data Gateway was introduced in 2021. The service has been in regular use since then, resulting in a total of 37 datasets curated and published (Fig. 4). The list of datasets shared via the HDG is available under the data accessibility filter named controlled access (https://search.kg.ebrains.eu/?category=Dataset&dataAccessibility[0]=controlled%20access) in the Knowledge Graph Search. HDG data represent a variety of modalities, dominated by neuroimaging, behavioural and electrophysiology data (Figure 4). The design of GDPR compliant workflows for responsible sharing of data via the Human Data Gateway has been described in more detail in the related publication by Karakasidis and Vassalos, 2023 (DOI: 10.5220/0011743000003405).

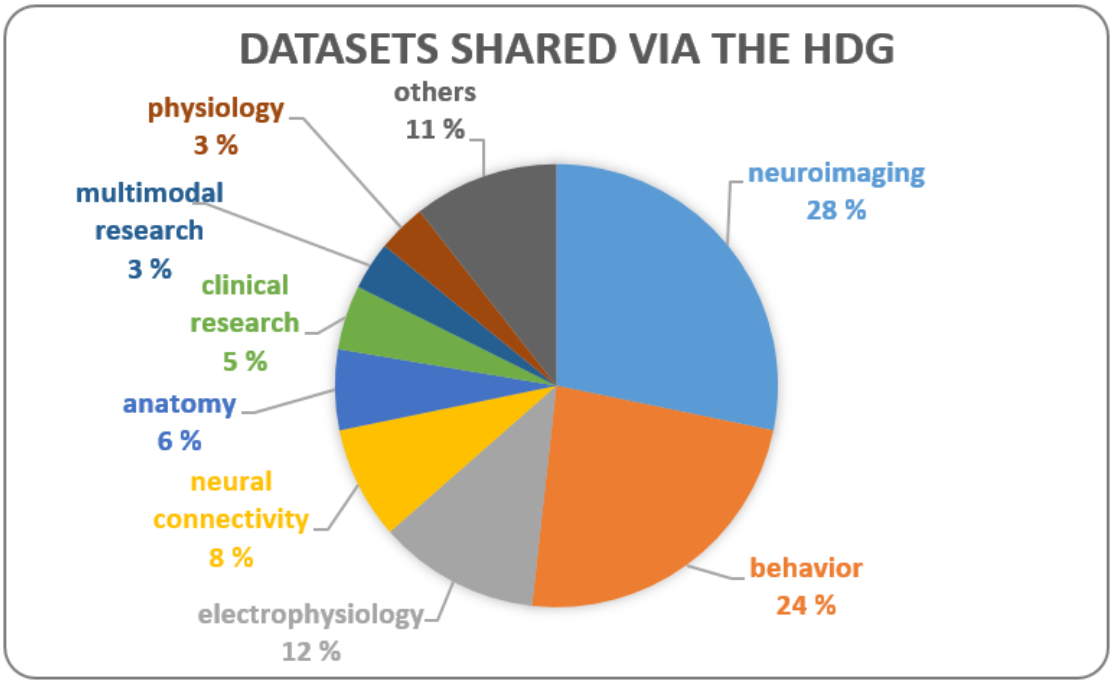

Figure 4: Distribution of experimental modalities across datasets shared via the HDG

The scope of the HDG is restricted to human data without direct identifiers and with very few indirect identifiers (strongly pseudonymised, de-identified data). Depending on the nature of the data, it may not always be possible to remove all identifying elements from e.g. imaging data without limiting the reusability potential and re-analysis value of the data. EBRAINS is therefore developing a Service for Sensitive Data (https://www.healthdatacloud.eu/) for sharing clinical data

of high sensitivity that requires further safeguards in place. These developments are independent of the Human Data Gateway, but in a close dialogue with the EBRAINS Data and Knowledge Services.

2.3.1 User benefits

The Human Data Gateway represents a low-threshold entry point to the EBRAINS data sharing service for researchers working with strongly pseudonymised human data. Users of the HDG benefit from full access to the EBRAINS data curation service, including guidance on data organization and metadata management. HDG datasets are offered free storage and streamlined data upload directly from the browser via the easy-to-use EBRAINS data proxy graphical user interface. Large data can also be uploaded using Python. Data re-use is facilitated by a fast and automated access request process and one-click data download, while security measures are in place to track access to each file and protect the data from unauthorised use.

2.4 Live papers

EBRAINS Live Papers are structured and interactive online documents that complement published scientific articles. Live papers include text, figures, links to publicly available research products (datasets, models, or code), and integrated interactive tools that allow users to explore and visualise models and datasets and run simulations and Jupyter notebooks.

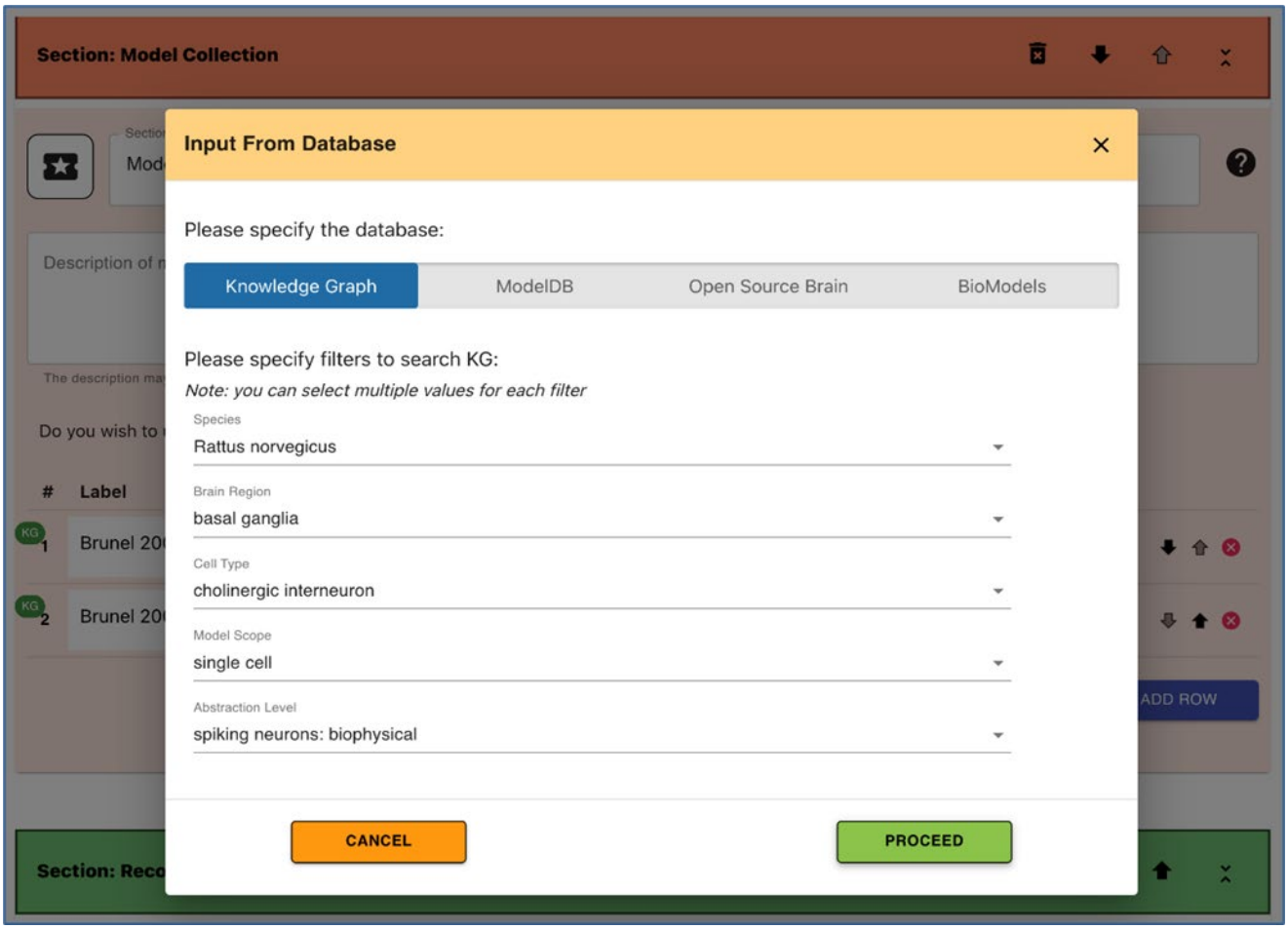

Figure 5: A dialog box from the Live Paper Builder app.

The dialog box allows the user to link to models in public repositories. Similar dialogs are available for electrophysiology data and for morphological reconstructions.

The service includes a "Live Paper builder" app accessible with an EBRAINS account through the "Publish your own Live Paper" button at https://www.ebrains.eu/data/live-papers/live-papers, and

a live paper viewer which is publicly accessible through the "Browse and search Live Papers" button on the same page (all live papers are published with open access licences).

The Live Paper builder allows users to add links to research products from several public repositories, currently EBRAINS Knowledge Graph, the Allen Brain Atlas, NeuroMorpho.org, ModelDB, OpenSourceBrain, and BioModels (Figure 5). For the Knowledge Graph, a search interface is available; for the other repositories, a dataset or model identifier must be entered. Links to Jupyter notebooks in the EBRAINS Lab may also be added, and formatted text and figures can be added using Markdown syntax to provide context and explanations. In-development live papers may be downloaded as a JSON document and/or saved to the Knowledge Graph into the private "myspace" or a "collab space". The latter option allows in-development live papers to be shared with/edited by collaborators using the standard EBRAINS Collaboratory access controls.

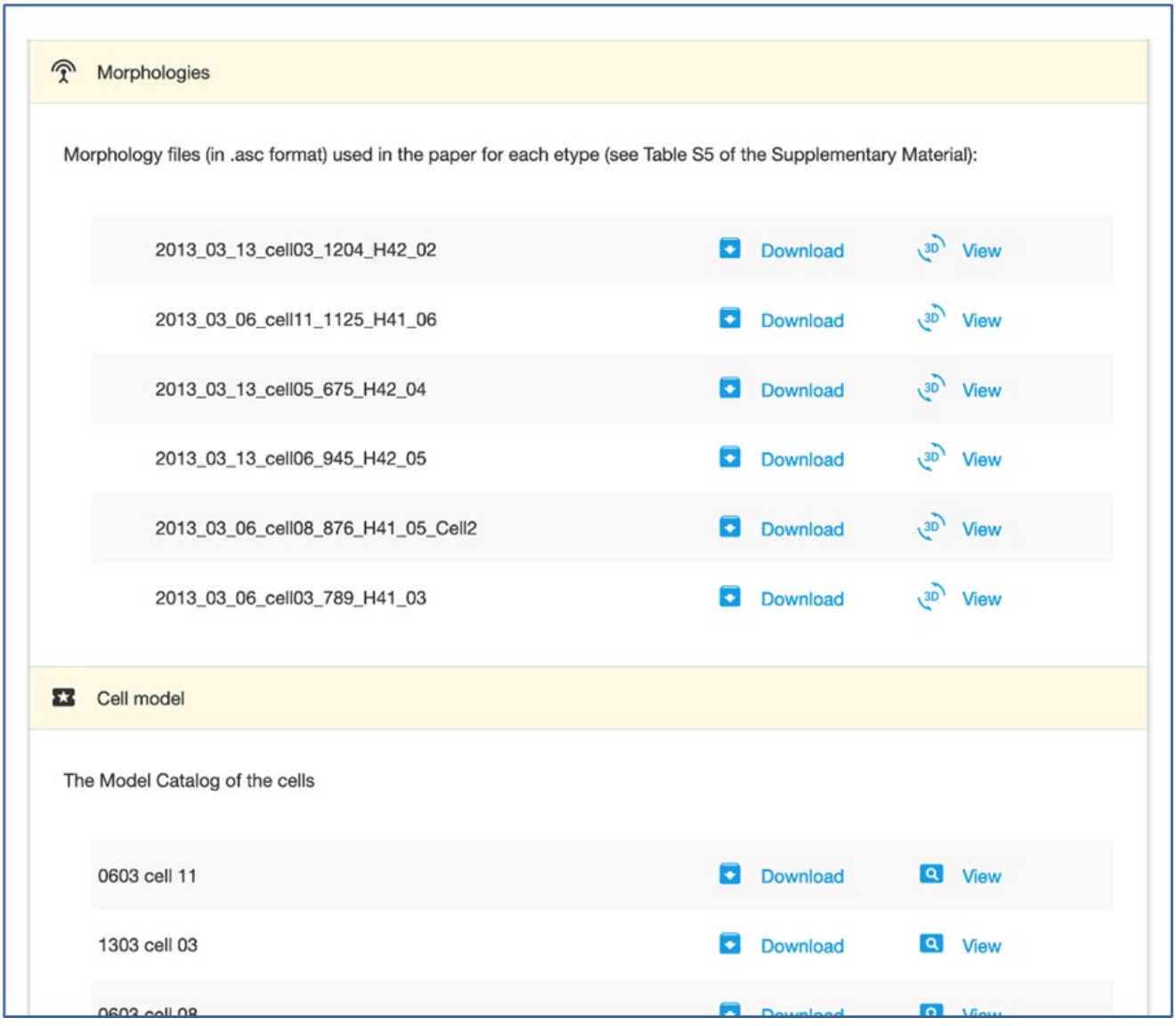

Figure 6: Lists of resources in a published live paper, with buttons to open interactive viewers.

When a live paper is ready for publication the authors may request curation through EBRAINS support. The curation process checks that all linked resources are publicly available in suitable repositories with compatible licences. If any resources are not stored in suitable repositories, the authors are invited to request curation of these resources through the EBRAINS Curation service. Once curation is complete, a DOI is attributed, and the live paper is made available online.

Where the live paper serves as supplementary material to a traditional manuscript, the authors may share it anonymously with reviewers prior to publication by setting a password that can be communicated to the journal editor.

The Live Papers viewer automatically detects resources for which an associated EBRAINS tool is available and either embeds the tool in page or provides a button to launch it in a new browser tab (Figure 6). Currently supported tools include the NeoViewer widget for electrophysiology data, the HBP Neuron Morphology Viewer for neuronal reconstructions, and the BlueNAAS in-browser simulation tool for models of neurons.

Since the Live Papers service was first introduced, a number of features have been added to the builder app: (1) metadata for an associated publication can now be fetched by entering its DOI, as an additional option alongside uploading a PDF or entering the metadata by hand; (2) authors can clone an existing live paper as a starting point for a new one, making it easier to keep the same structure across related papers; (3) large numbers of data files or other research products can be added in an easier way, with an optional spreadsheet-like interface.

As well as new functionality, the visual appearance has been updated to better match the EBRAINS colour scheme; the documentation has been improved and a video tutorial released on YouTube, https://www.youtube.com/watch?v=h6BxdWN1Jj4.

A paper describing the Live Papers service was published (Appukuttan et al. 2022; doi:10.1007/s12021-022-09598-z; P3328).

The most extensive changes to the service are invisible to the end user: the service and its data were migrated from version 2 to version 3 of the Knowledge Graph. This was a challenging process that involved developing a new openMINDS extension for scientific publications (see above). The benefits are greater interoperability for live paper metadata and the possibility for automatic enrichment of live papers through following links in the KG. Finally, continued maintenance and improvement of the underlying code has resulted in performance improvements, greater reliability, and better maintainability.

2.4.1 User benefits

By sharing their data, models, and code through EBRAINS, scientists satisfy data-sharing mandates according to FAIR principles, and in a way that maximises the potential impact of their work. Since many traditional publications are based on multiple datasets and/or multiple models, Live Papers provide an additional mechanism that groups and contextualises related datasets, code, and models, and enables interactive exploration and re-use of these research products.

Live Papers also allow the grouping of resources across different public repositories, not only EBRAINS but also the Allen Brain Atlas, NeuroMorpho.org, ModelDB and others.

Individual published live papers are citable and can be accessed directly through their DOI.

In most cases, Live Papers functions as interactive supplementary material to a published manuscript. While a small number of publishers such as eLife have started publishing their own live papers, the EBRAINS Live Papers service allows authors to associate interactive, reproducible documents with publications in any journal, in a publisher-agnostic way.

2.5 EBRAINS Compliance Management, Data Management, and Data Governance

EBRAINS compliance management, data management and data governance has collaborated with several key stakeholders to ensure that data integrated into EBRAINS met ethical, social, and legal requirements. We have coordinated several pivotal pieces of work towards this end:

1) The completion of a suite of Data Protection Impact assessments for all applicable EBRAINS services to ensure that any personal data processing occurring in EBRAINS meets GDPR requirements.
2) The management of the EBRAINS ethics compliance process, as part of the EBRAINS data curation pathway, where data intended to be shared through the EBRAINS Knowledge Graph are reviewed for compliance with applicable legislation.

3) Updating and improving several critical EBRAINS policy documents.
   a) EBRAINS Access Policy[22]
   b) EBRAINS Data Provision Protocol[23]
   c) EBRAINS Data Use Agreement[24]
   d) EBRAINS Cookie Statement[25]
   e) EBRAINS General Terms of Use[26]
   f) EBRAINS Privacy Statement[27]

We have collaborated with colleagues in the Human Brain Project to run a set of Capacity Development workshops for EBRAINS users: https://www.humanbrainproject.eu/en/science-development/ethics-and-society/ethics-society-training-resources/human-animal-data-ebrains/. These workshops focused on the ethical and data governance challenges associated with sharing brain data on the scale of EBRAINS and garnered positive feedback. These capacity development resources will continue to be available for EBRAINS users via the HBP website.

2.5.1 User benefits

EBRAINS compliance management, data management and data governance provide expert support to EBRAINS users as they seek to share data on EBRAINS. The ethics compliance check is a key step in the data curation pathway, and it ensures that EBRAINS only shares data which meets ethical and legal requirements. Similarly, we provide support to EBRAINS data and knowledge personnel – who will often be the first port-of-call for users with questions relating to the ethics and data governance challenges of sharing data through EBRAINS.

EBRAINS users can, therefore, be assured that data they access on EBRAINS platforms is collected and shared responsibly. Data providers, similarly, can be assured that the data they share through EBRAINS is managed appropriately, that the rights of their data subjects are respected and any queries they might have on towards ethics compliance, data protection or data governance are addressed by experts in those fields.

2.6 KnowledgeSpace

The KnowledgeSpace (KS; http://knowledge-space.org) is a community-based encyclopaedia and data catalogue for neuroscience that provides a unique, global interface between brain research concepts and the data, models, and literature that support or weaken their definition, as well as educational resources associated with those concepts. Today, KS contains over 90,000 ontology terms and over 700,000 publicly available datasets (excluding literature results from PubMed) from 17 of the world's leading neuroscience repositories covering a wide range of species (human, macaque, mouse, and rat) and data types (anatomical, connectivity, expression, microscopy, models, morphology, neuroimaging, and physiology) in both the central and peripheral nervous systems. In addition, KS provides users with access to over 700 neuroscience courses, lectures, and tutorials through indexing the INCF TrainingSpace.

As a global service supported by EBRAINS, KS provides EBRAINS with a catalogue of data and models generated outside of the HBP— thus increasing the volume and heterogeneity of the data and models available through the EBRAINS Research Infrastructure (RI), as well as providing a much easier mechanism for non-HBP affiliated data repositories to connect and make their datasets discoverable

[22] https://files.ebrains.eu/file/e4b05476-d2f0-49c2-8b45-41f9c317892e/EBRAINS_AISBL_Access_Policy_04_2022_85c8e61216.pdf
[23] https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_Data_Provision_Protocol_dfe0dcb104.pdf
[24] https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_Data_Use_Agreement_90858e7836_ef3ee29d50.pdf
[25] https://files.ebrains.eu/file/e4b05476-d2f0-49c2-8b45-41f9c317892e/EBRAINS_Website_Cookie_Statement_2022_3119732739.pdf
[26] https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_General_Terms_of_use_e457353c1a_d2122f84c2.pdf
[27] https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_Privacy_Statement_2022_80958229c5.pdf

through the Knowledge Graph since the KS schema is not limited to the openMINDS schema required by Knowledge Graph. KS has resulted in the development of Google Dataset Search-like search engine for neuroscience. Unlike Google Dataset Search which is high level (i.e. search only at the dataset level), KS provides a similar search capability together with extensions for improved neuroscience dataset discovery (i.e. search down to the data file level). To our knowledge, KS is the only FAIR data indexing service that provides users with a single point of entry to access the datasets generated by four of the world's large-scale brain projects: the HBP, the US BRAIN Initiative, Japan Brain/MINDS, and the Canadian Open Neuroscience Platform (Figure 7).

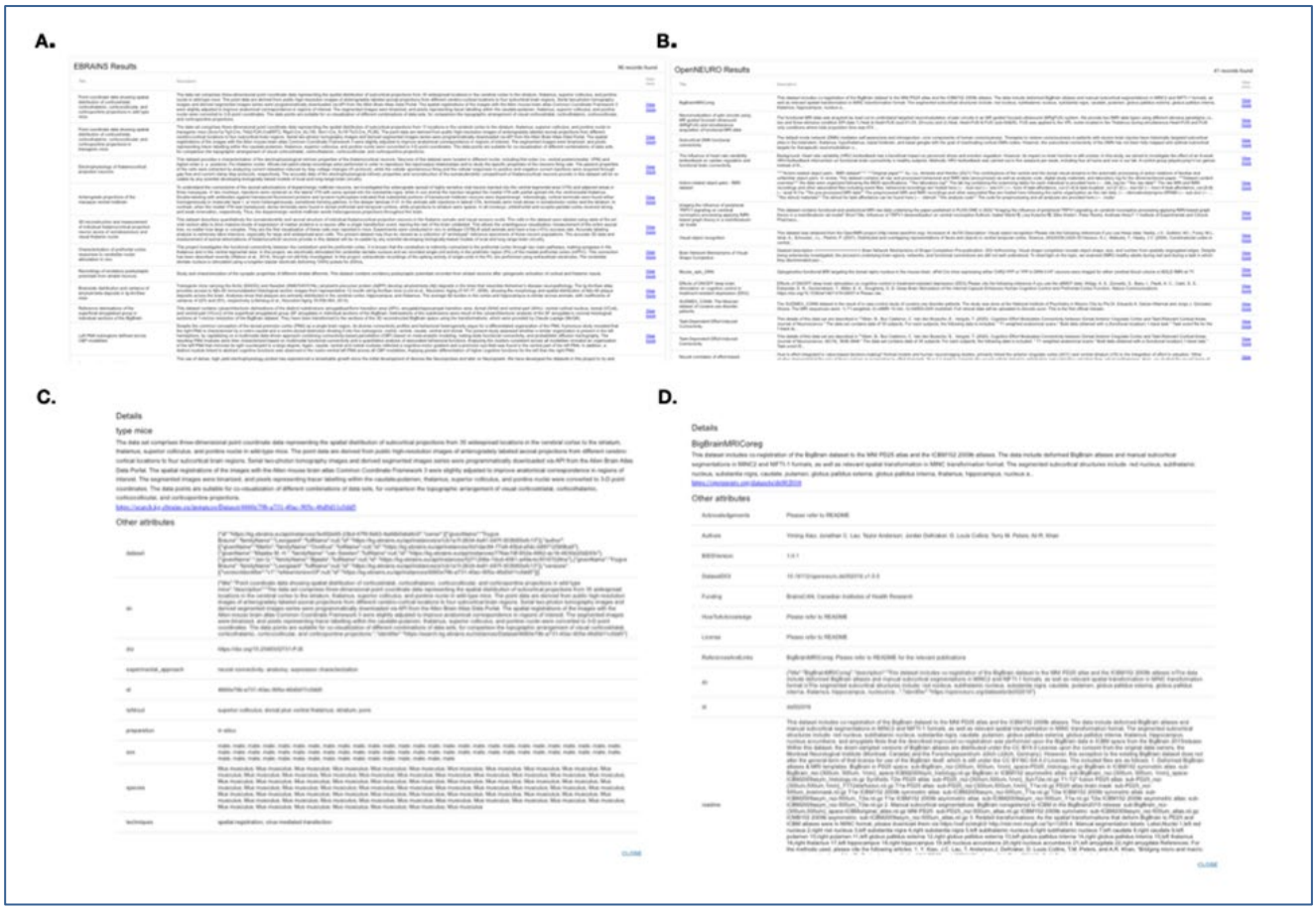

Figure 7: KnowledgeSpace metadata schema.

A. Display of KS search results for EBRAINS and B. for openNeuro shows the results using the KS schema view which provides users with a brief description of the data with the option to expand to view the results according to C. the EBRAINS metadata schema and D. the openNeuro metadata schema for more detailed information about the data.

An important feature of the KS user interface is that it adapts to the schema of each repository indexed. That is, when a search is conducted within a single repository, the results are based on its schema; and when a federated search is conducted across multiple repositories, the search results are based on a minimal common schema. KS users have several options for searching: dataset-only, literature-only, courses/tutorials-only, or the encyclopaedia view which combines the dataset and literature with a Wikipedia-like description with the ability to receive federated results from across data/model sources (Figure 8). Users can refine their searches by either filtering by data source or by adding additional text in the search bar. Of special note, the dataset-only search returns data cards that include summaries of the data files with the option to expand for more detailed metadata.

There are APIs to enable 3rd party resources to programmatically access KS. For example, the Neurobot data sharing tool currently uses KS APIs to programmatically access the National Institute for Neurological Disorders and Stroke's (NINDS) common data elements (CDEs), a standardised and precisely defined question paired with a set of allowable responses that are used systematically across different sites and studies to ensure consistent data collection, for traumatic brain injury which enables users of the tool to annotate data with the NINDS CDEs for traumatic brain injury. It is important to note that all NINDS CDEs have been integrated into KS.

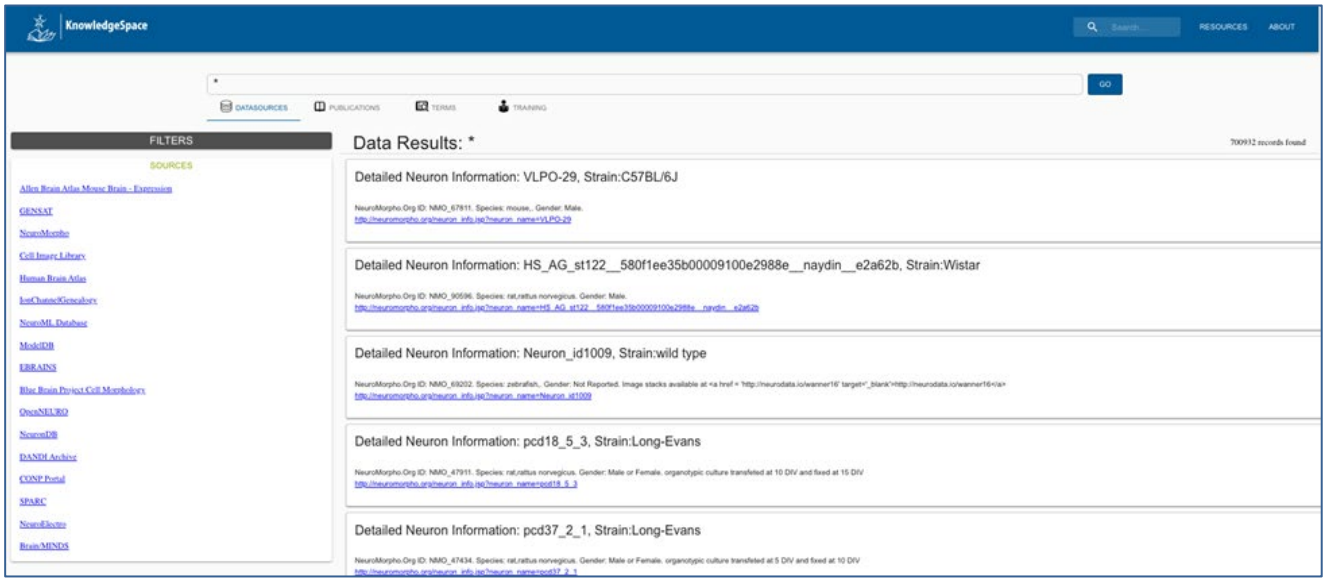

Figure 8: KnowledgeSpace (KS) features.

The search interface enables users to conduct 4 major classes of query: dataset-only, literature-only, encyclopaedia (combines dataset and literature with Wikipedia), and for training resources. Users can further filter search based on the data source. Note: KS integrates data generated by the world's large-scale brain projects.

2.6.1 User benefits

KS provides users with a single point of entry catalogue for some of the world's leading neuroscience data repositories. Today, KS contains over 90,000 ontology terms and over 700,000 datasets (excluding literature results from PubMed) covering both the central and peripheral nervous system in mouse, rat, human, and macaque over a wide range of modalities (gene expression, morphology, models, physiology, connectivity, histology, microscopy, and MRI). In addition to providing users with a one-stop shop catalogue of datasets from many sources, KS also enables repositories to demonstrate that they are adhering to the FAIR Principles since only repositories adhering to the FAIR principles are indexed in KS.

2.6.2 Work in progress

We will continue to increase the range and types of data and models, improve the user experience, and achieve a tighter integration with the EBRAINS Knowledge Graph, as outlined below.

Expanding the range of data and model types:

As new data and models are added to the existing data sources, they will be indexed in KS.

Improving the user experience:

1) Enhanced search capabilities: The data and model panels in the encyclopedia will be updated so that users can benefit from the data set search functionality from the encyclopaedia (expanded modalities, inter/intra source search).
2) Further curation of terms in the encyclopedia.
3) Tutorials: New multimedia content (including the updated how to guide and video tutorial) will be embedded into the KS website.

3. Education, Outreach, and other interaction with users of the services

The period since October 2021 has been highly active with respect to events targeting both data providers and data consumers. The next sections provide concrete examples of the Data and Knowledge team´s engagement. Summarising, the events have focused on 1) increasing awareness and overall visibility of services, 2) enabling the integration and collaboration with our services and 3) enhancing user engagement and enabling use of our services within the neuroscience community.

3.1 EBRAINS Knowledge Graph

Below are examples of outreach opportunities where we have promoted the EBRAINS Knowledge Graph services.

- A workshop to actively push the integration with the KG by other services: TVB-EBRAINS Integrated Workflows 7th Workshop: Updates Knowledge Graph – 2021.12.17
- A presentation of the KG from a user's perspective: FENS 2022 (External Event) - A Practical Guide to Using the EBRAINS Knowledge Graph in (your) Research – 2022.07.13
- A social media campaign to make the broad audience aware of the available tools: EBRAINS Data and Knowledge Social Media Campaign – 2022.04.23 to 2022.07.23
- A workshop for students in neuroscience to make them aware of the importance of metadata management in general and to show them the possibilities with the power of the developed toolset around the EBRAINS KG: EBRAINS training on Brain Atlases and Simulation Services (for members of Austrian Neuroscience Association - ANA (Internal Event) - Hands on workshop on EBRAINS Knowledge Graph and metadata management – 2022.09.13

3.2 EBRAINS curation

Below are examples of events where EBRAINS Curation services have been promoted. The examples are classified thematically.

Events promoting Open MINDS including specialised courses with close interaction with users, training event, and talks at major conferences:

- A course promoting user interaction: CodeJam#12. Fairgraph; General overview to openMINDS; openMINDS Python API - 2021.11.23
- A workshop promoting EBRAINS and openMINDS and allowing for interaction with users: "In2PrimateBrain Skills" Workshop - 2021.12.02
- TVB-EBRAINS Integrated Workflows 7th Workshop: Update on openMINDS - 2021.12.17
- TVB-Cloud & Health Data Cloud Technical Coordination: Talk on openMINDS - 2022.08.26
- 6th BigBrain Workshop: Talk titled 'openMINDS meets BigBrain - capturing complex histology data in a graph database' - 2022.10.27
- Satellite Symposium at SfN 2022 by IBI Data Standards and Sharing Working Group: Talk on open Metadata Initiative for Neuroscience Data Structures (openMINDS) - 2022.11.14
- Ontology Engineering Workshop - 2022.11.17
- NeuroPSI ICN Weekly Seminar: Talk on openMINDS – 2022.12.12
- BICAN Metadata Specification Interchange Workshop: Talk titled 'openMINDS - A metadata framework for graph database systems." - 2023.01.24
- TVB-Cloud & Health Data Cloud Technical Coordination: Talk on openMINDS - 2023.03.03

- HBP Summit 2023: Posters titled ‘openMINDS SANDS’, ‘MATLAB user interface for openMINDS metadata registration’, ‘openMINDS extensions for in-depth graph database descriptions of electrophysiology experiments’, ‘Pioneering open neuroscience - the role of the curation services’ - 2023.03.28
- NeuroPSI-Chen Institute Joint Conference on Brain, Behaviour & Beyond: Poster titled ‘openMINDS - Flexible metadata models for neuroscience’ - 2023.05.11.
- openMINDS hackathon - 2023.06.30
- OHBM 2023, Montréal, Canada: Poster titled ‘openMINDS - SANDS’ - 2023.07.22-26
- HIBALL Hackathon, Montréal, Canada: Hands on Session for openMINDS - 2023.07.27

Events promoting the curation workflow and related components of the EBRAINS Data and Knowledge services, including specialised courses with close interaction with users, training event, and talks at major conferences:

- EAN-EBRAINS Joint Workshop: The Future of Medical Data Sharing in Clinical Neurosciences: Talk titled ‘How to make data public: ethics, regulations, solutions’ - 2021.12.09
- HBP UK seminar - 2022.03.22
- FLAG-ERA 2022 Project Workshop: Data Management Open Discussion: ‘EBRAINS Data and Knowledge Services’ - 2022.03.22
- Young Researchers Event Denmark: EBRAINS - a digital European Infrastructure for next-generation basic & clinical neuroscience: Talk titled ‘Managing, sharing, and publishing research data through EBRAINS,’ Open discussion about data sharing: opportunities and challenges, rewards, and pitfalls - 2022.05.16
- EBRAINS Workshop: Brain Activity across Scales and Species: Analysis of experiments and simulations (BASSES): Hands-on Session 1: Handling EBRAINS Data, Hands-on Session 4: Validating models against data in EBRAINS - 2022.06.13-15
- Organisation & Participation in Nature Portfolio Webcast: "Making small data big: Building a graph database from experimental brain research" - 2022.06.27
- FENS 2022: posters titled ‘A Practical Guide to Using the EBRAINS Knowledge Graph in (your) Research’, presentation titled ‘Making the most of precious data: challenges and opportunities in mining open data’, ‘openMINDS - Flexible metadata models for neuroscience’ - 2022.07.08-09
- EBRAINS satellite event at the FENS Forum 2022.: Hands-on session on ‘EBRAINS Data Services’ - 2022.07.08
- Interaction at bigger event. HBP booth at FENS Forum 2022 - 2022.07.09
- HBP Partnering Projects Meeting: ‘Status quo and outlook: EBRAINS Data & knowledge service.’ Disseminating information about EBRAINS Data & Knowledge services to partnering projects members. Some were familiar with our services and wanted to learn more and others were not familiar and learned about new opportunities within EBRAINS - 2022.09.05
- INCF Neuroinformatics Assembly 2022: Talk titled ‘EBRAINS Data & Knowledge - exploring data through metadata’ - 2022.09.12
- HBP booth at SfN Neuroscience 2022 - 2022.11.12
- HIBALL Winter School 2023: Tutorial on EBRAINS Data & Knowledge - How to share own data and explore the shared data from others - 2023.02.09
- EBRAINS Brain Simulation School 2023 “Training on modelling neurons, circuits, and cognitive functions”: Talks titled ‘The Knowledge Graph and data sharing’, ‘Hands-on: Knowledge Graph database’- 2023.05.19-23
- EBRAINS – Euro-BioImaging workshop. Liaising with another large Research Infrastructure and exploring possible means of collaboration - 2023.05.31

3.3 EBRAINS Compliance Management, Data Protection, and Data Governance

EBRAINS Compliance Management, Data Protection, and Data Governance has been involved with a number of workshops, outreach events and dissemination activities, such as:

- Key contributions to the Human Brain Project Capacity Development Plan, specifically relating to the handling of human and animal data on EBRAINS platforms. https://www.humanbrainproject.eu/en/science-development/ethics-and-society/ethics-society-training-resources/human-animal-data-ebrains/. This set of tailored training resources aims to ensure that EBRAINS users have the knowledge and skills needed to responsibly handle data on EBRAINS platforms.
- Collaborating with the Danish Board of Technology, members of staff from De Montfort University organised and ran the successful "European Health Data Space: Navigating the Challenges of Responsible Data Governance" workshop held in Brussels on the 22nd of June 2023. The event included 22 experts from across a broad spectrum of expertise who came together to brainstorm and articulate the potential data governance challenges associated with the forthcoming European Health Data Space. At this single day event, participants identified challenges across three themes: technical, legal, and ethical/social challenges, participants then prioritised and subsequently articulated the most pressing or important challenges under those themes. A short report on the workshop outcomes is under development and is expected to be produced in August 2023. EBRAINS will be a key contributor to the European Health Data Space, and so it is important that EBRAINS is at the forefront of discussions around the usage of health data, and particularly brain data, in such spaces.

  EBRAINS Compliance Management, Data Protection and Data governance presented a panel at the 7th HBP Student Conference on Interdisciplinary Brain Research titled "Responsible Brain Data Governance: Ethical and Legal Considerations", which discussed challenges associated with the handling of brain data, and suggested potential solutions and resources for students. https://www.humanbrainproject.eu/en/follow-hbp/events/hbp-student-conference-7.

3.4 EBRAINS KnowledgeSpace

We conducted a pilot project with an undergraduate neuroscience laboratory at the University of California, Los Angeles to determine whether the user interface (UI) improved the ability of users to find appropriate datasets through KS to conduct analyses. Feedback on the UI was positive, and we were provided with several feature requests which we plan to implement in future developments. The results of the pilot were presented as a poster at Society for Neuroscience 2022.

In addition, efforts are underway to refine KS integration in the INCF TrainingSpace[28], an open access hub of multimedia educational resources and tool tutorials.

4. Future Perspectives

The future of the EBRAINS Data and Knowledge services involves a commitment to serving the global neuroscience community. The services are positioned to play a pivotal role in advancing neuroscience research and fostering collaboration on a global scale. Key aspects include:

- Continued Growth: The services are expected to experience continued growth in terms of the volume and diversity of data, models, and software shared by researchers in the neuroscience community. As more researchers recognize the benefits of data sharing and compliance with FAIR principles, and the value added by using EBRAINS, the repository is likely to continue to expand.

[28] https://training.incf.org/

- Enhanced Usability: The services will likely continue to improve interfaces and APIs to ensure that researchers can easily discover, access, and utilize the wealth of data and resources available. Efforts to improve search capabilities, data visualization, and data analysis tools will further enhance the user experience.
- Advanced Metadata and Ontologies: The openMINDS framework and metadata standards are expected to evolve to accommodate an even broader range of research products and domains within neuroscience. This will result in more detailed and structured metadata, enhancing the contextual information available to users.
- Interoperability: The expected adoption of openMINDS by other brain initiatives and data repositories will contribute to greater interoperability, making it easier for users to access and compare data from various sources. This cross-repository compatibility will foster collaboration and data integration.
- Incorporating New Technologies: The services are likely to incorporate new technologies and data formats as they emerge in the field of neuroscience. This adaptability will ensure that EBRAINS remains at the forefront of data sharing and management.
- Global Impact: EBRAINS Data and Knowledge services aim to have a global impact by serving the international neuroscience community. As more researchers and organizations worldwide recognize the value of FAIR data sharing, EBRAINS is poised to become a global hub for neuroscience data.
- Advanced Analysis and Workflows: Integration with other EBRAINS services will offer advanced data analysis and workflow capabilities, enabling researchers to perform sophisticated analyses and gain deeper insights from the data. This will further facilitate discoveries and innovations.
- Ethical and Regulatory Compliance: EBRAINS will continue to prioritize ethical considerations and regulatory compliance, ensuring that data shared through its services adhere to applicable EU and national regulations and ethical principles.