

SP5 Neuroinformatics Platform - Results for SGA2 Year 2: (D5.8.2 - SGA2)



Figure 1: Neuroinformatics Platform services available through the new EBRAINS web portal

The Neuroinformatics Platform provides an infrastructure for neuroscientists who wish to collaboratively curate, analyse, share, and re-use large-scale neuroscience data, ultimately contributing to making research data more Findable, Accessible, Interoperable and Reusable (FAIR). Go to <https://ebrains.eu/> to see more about data sharing, brain atlases, and services from other HBP platforms.

Project Number:	785907	Project Title:	Human Brain Project SGA2
Document Title:	SP5 Neuroinformatics Platform - Results for SGA2 Year 2		
Document Filename:	D5.8.2 (D35.2 D41) SGA2 M24 ACCEPTED 201005.docx		
Deliverable Number:	SGA2 D5.8.2 (D35.2, D41)		
Deliverable Type:	Report		
Work Packages:	WP5.1, WP5.2, WP5.3, WP5.4, WP5.5, WP5.6, WP5.7, WP5.8, WP5.9, WP5.10		
Key Result(s):	KR5.1, KR5.3, KR5.4, KR5.6, KR5.7, KR5.8, KR5.9, KR5.11, KR5.12, KR5.14		
Dissemination Level:	PU = Public		
Planned Delivery Date:	SGA2 M24 / 31 Mar 2020		
Actual Delivery Date:	SGA2 M26 / 06 May 2020; resubmitted 31 Aug 2020; accepted 5 Oct 2020		
Author(s):	Jan BJAALIE, UIO (P81)		
Compiled by:	Roman VOLCHENKOV, Martha Elisabeth BRIGG and Sofia ANDERHOLM STRAND, UIO (P81)		
Contributor(s):	<p>Ida AASEBØ, UIO (P81), KR5.1</p> <p>Mathew ABRAMS, KI (P37), KR5.1 & KR5.14</p> <p>Trygve LEERGAARD, UIO (P81), KR5.1, KR5.4 & KR5.14</p> <p>Maja PUCHADES, UIO (P81), KR5.3, KR5.4</p> <p>Lyuba ZEHL, JUELICH (P20), KR5.1Paul TIESINGA, SKU (P51), KR5.3</p> <p>Timo DICKSCHEID, Anna HILVERLING, JUELICH (P20), KR5.4, KR5.6 & KR5.7</p> <p>Felix HOFFSTAEDTER, Anna Maria GEIGER, Simon EICKHOFF, JUELICH (P20), KR5.6</p> <p>Lyuba ZEHL, JUELICH (P20), KR5.1</p> <p>Pedro LARRAÑAGA, Nikolas BERNAOLA, UPM (P68), KR5.7</p> <p>Anastasia AILAMAKI, Marc MORGAN, Oliver SCHMID, EPFL (P1), KR5.9</p> <p>Allan FRANCANI, Fabrice GAILLARD, EPFL (P1), KR5.9 & KR5.14</p> <p>Anna KRESHUK, EMBL (P123), Pascal FUA, Subeesh VASU, EPFL (P1), KR5.11</p> <p>Sonja GRÜN, JUELICH (P20), KR5.12</p> <p>Michael DENKER, JUELICH (P20), KR5.12 & KR5.14</p> <p>Andrew DAVISON, CNRS (P10), KR5.1, KR5.12, KR5.13 & KR5.14</p> <p>Axel von ARNIM, EPFL (P1), Michele MIGLIORE, Luca Leonardo BOLOGNA, CNR (P12), Boris ORTH, JUELICH (P20), Ludovic CLAUDE, CHUV (P27), Sten GRILLNER, KI (P37), Björn KINDLER, UHEI (P47), Roman VOLCHENKOV, UIO (P81), Daniele MARINAZZO, UGENT (P110), Petra RITTER, CHARITE (P122), KR5.14</p>		
SciTechCoord Review:	Martin TELEFONT, EPFL (P1)		
Editorial Review:	Annemieke MICHELS, EPFL (P1)		
Description in GA:	For consistent presentation of HBP results, SGA2 M24 Deliverables describing the accomplishments of an entire SP, WP or CDP have been prepared according to a		

	standard template, which focuses on Key Results and the outputs that contribute to them. Project management elements such as Milestones and Risks will be covered, as per normal practice, in the SGA2 Project Periodic Report.
Abstract:	The Neuroinformatics Platform contributed to the release of the new EBRAINS web portal (https://ebrains.eu/) in the fall of 2019. All HBP Platforms contribute to the new portal. The portal will be a key instrument for displaying tools and services targeting the brain research and related communities in the coming years. In the present version of the portal, the Neuroinformatics Platform has delivered services ensuring that the data needed to build brain models can be brought together, annotated, made accessible and prepared for analysis. This includes the EBRAINS Knowledge Graph with an online web interface that makes data discoverable, accessible, and possible to interpret and re-use, made possible by a dedicated team of curators. The Knowledge Graph has been opened also for submission of data from contributors external to the HBP. New versions of the atlases are released on a regular basis, together with workflows supporting feature extraction from images based on atlas regions, as well as deep analysis of functional dynamics. Overall, the Neuroinformatics Platform has created solutions for sharing and advanced analysis of data from the brain and made these solutions accessible for the neuroscience community, creating new opportunities for collaborative data analysis, re-analysis of new combinations of data and reproducible science. The Neuroinformatics Platform also hosts the High-Level Support Team, which delivers support to all users of services made available from the HBP.
Keywords:	Data sharing; data curation; data management; reference atlas; rodent; human; spatial registration; image analysis; electrophysiology; neuronal activity; prediction; infrastructure; user support; neuroinformatics
Target Users/Readers:	computational neuroscientists, neuroimaging community, HBP Platform users, Consortium members, funders, general public, policymakers, students

Table of Contents

1. Overview	7
2. Introduction	8
3. Key Result KR5.1 “A well-defined, well-documented and proven metadata curation approach which makes HBP data and models FAIR (Findable, Accessible, Interoperable, Re-usable) in a consistent and user-friendly way”	10
3.1 Outputs	10
3.1.1 Overview of Outputs	10
3.1.2 Output 1 - Optimised metadata structure	10
3.1.3 Output 2 - Curation workflow (v2.1)	11
3.1.4 Output 3 - Ontology development	12
3.1.5 Output 4 - Data and models available in the EBRAINS KG	13
3.2 Validation and Impact	13
3.2.1 Actual and Potential Use of Output(s)	13
3.2.2 Publications - not applicable to KR5.1	14
4. Key Result KR5.3 “Complete workflow going from heterogeneous experimental image data to extracted quantitative features defined in rodent atlas space, available for general use”	14
4.1 Outputs	14
4.1.1 Overview of Outputs	14
4.1.2 Output 1: QUINT workflow	14
4.1.3 Output 2: Connectomic Composition Predictor (CCP)	16
4.2 Validation and Impact	16
4.2.1 Actual and Potential Use of Output(s)	16

4.2.2	Publications	17
5.	Key Result KR5.4 “A well-defined, well-documented and proven approach for assigning spatial metadata to HBP data and models.”	17
5.1	Outputs	17
5.1.1	Overview of Outputs.....	17
5.1.1	Output 1: New versions of reference atlases	18
5.1.2	Output 2 Optimised tools and procedures for registration of brain-image data to reference atlases	19
5.2	Validation and Impact.....	20
5.2.1	Actual and Potential Use of Output(s)	20
5.2.2	Publications	21
6.	Key Result KR5.6 “Interactive exploration of function, connectivity, cytoarchitecture and receptor architecture at sub-laminar resolution in the Big Brain template space.”	21
6.1	Outputs	21
6.1.1	Overview of Outputs.....	21
6.1.2	Output 1: New release of interactive atlas viewer	22
6.1.3	Output 2: Ultrahigh-resolution 3D maps and maps of cortical layers for the BigBrain model.....	23
6.1.4	Output 3: Pre-processed data from neuroimaging cohorts available	24
6.1.5	Output 4: HPC-based workflows configured and deployed on JURECA	25
6.2	Validation and Impact.....	25
6.2.1	Actual and Potential Use of Output(s)	25
6.2.2	Publications - not available for KR5.6	26
7.	Key Result KR5.7 “Queries for structural parcellations, genetics, and activity for user-defined regions of interest”	26
7.1	Outputs	26
7.1.1	Overview of Outputs.....	26
7.1.2	Output 1: Release of JuGEx (JuBrainGeneExpression) as atlas viewer toolbox and python library	26
7.1.3	Output 2: Gene Regulatory Network Explorer - Graph visualisation and bayesian network learner tool.....	28
7.2	Validation and Impact.....	28
7.2.1	Actual and Potential Use of Output(s)	28
7.2.2	Publications	28
8.	Key Result KR5. 9 “A user-driven, data-sharing and data-management infrastructure accelerating scientific progress by providing access to large collections of curated, heterogeneous neuroscience data.”	29
8.1	Outputs	29
8.1.1	Overview of Outputs.....	29
8.1.2	Output 1: EBRAINS Knowledge Graph.....	29
8.1.3	Output 2 - Collaboratory v2.0.....	30
8.1.4	Output 3 - Image Service	31
8.2	Validation and Impact.....	32
8.2.1	Actual and Potential Use of Output(s)	32
8.2.2	Publications - not applicable for KR5.9	32
9.	Key Result KR5.11 HBP-specific workflows developed to fully leverage the current state of the art in computer vision with regards to object detection and segmentation problems. Active and transfer learning methods will be developed to lower the requirements to the necessary amount of training data	33
9.1	Outputs	33
9.1.1	Overview of Outputs.....	33
9.1.2	Output 1: Extension of the Neuroglancer viewer to brushstroke annotations (C2234).	33
9.1.3	Output 2: Web-based training of ilastik classifiers (C2235)	34
9.1.4	Output 3: Integration of ilastik with the HPC infrastructure through UNICORE (C2235)	34
9.1.5	Output 4: Transfer learning for brain vasculature segmentation (C2282), delivering pre-trained networks and code	34

9.1.6	Output 5: PyTorch back end for ilastik, new version based on GRPC, release March 2020 (C2236)	35
9.2	Validation and Impact	35
9.2.1	Actual and Potential Use of Output(s)	35
9.2.2	Publications	36
10.	Key Result KR5.12 “Comprehensive analysis toolbox to support analysis workflows for electrophysiological activity data from experiments and simulations, including support for HPC and visualisation capabilities.”	36
10.1	Outputs	36
10.1.1	Overview of Outputs	36
10.1.2	Output 1	37
10.1.3	Output 2	37
10.1.4	Output 3	37
10.2	Validation and Impact	38
10.2.1	Actual and Potential Use of Output(s)	38
10.2.2	Publications	38
11.	Key Result KR5.13 “The neural activity resource (NAR) as a central mechanism to register, annotate and browse activity data sets within the HBP”	38
11.1	Outputs	39
11.1.1	Overview of Outputs	39
11.1.2	Output 1- NAR metadata tools	39
11.1.3	Output 2	40
11.2	Validation and Impact	41
11.2.1	Actual and Potential Use of Output(s)	41
11.2.2	Publications - not applicable for KR5.13	41
12.	Key Result KR5.14 “HBP and non-HBP affiliated researchers using interoperable platforms”	41
12.1	Outputs	41
12.1.1	Overview of Outputs	41
12.1.2	Output 1 - KnowledgeSpace (KS)	42
12.1.3	Output 2 - High Level Support Team (HLST)	42
12.1.4	Output 3	43
12.2	Validation and Impact	43
12.2.1	Actual and Potential Use of Output(s)	43
12.2.2	Publications	45
13.	Conclusion and Outlook	45
14.	Index of Components	47

Table of Tables

Table 1:	Output 1 Links	11
Table 2:	Output 2 Links	12
Table 3:	Output 3 Links	13
Table 4:	Output 1 links	16
Table 5:	Output 2 Links	16
Table 6:	Output 1 Links	19
Table 7:	Output 2 Links	20
Table 8:	Output 1 Links	23
Table 9:	Output 2 Links	24
Table 10:	Output 3 Links	25
Table 11:	Output 4 Links	25
Table 12:	Output 1 Links	27
Table 13:	Output 2 Links	28
Table 14:	Output 1 Links	30
Table 15:	Output 2 Links	31
Table 16:	Output 3 Links	32

Table 17: Output 1 Links	34
Table 18: Output 2 Links	34
Table 19: Output 3 Links	34
Table 20: Output 4 Links	35
Table 21: Output 5 Links	35
Table 22: Output 1 Links	37
Table 23: Output 2 Links	37
Table 24: Output 3 Links	38
Table 25: Output 1 Links	40
Table 26: Output 1 Links	40
Table 27: Output 1 Links	42
Table 28: Output 2 Links	43
Table 29: Output 3 Links	43

Table of Figures

Figure 1: Neuroinformatics Platform services available through the new EBRAINS web portal	1
Figure 2: QUINT workflow for quantification of features using 3D reference atlases	15
Figure 3: Aligning a partial volume to BigBrain using VoluBA	19
Figure 4: Optimised user interface of interactive atlas viewer (here: BigBrain atlas)	22
Figure 5: New connectivity browser	22
Figure 6: Maximised view of the coronal plane, corresponding to Figure 4	23
Figure 7: Cytoarchitectonic maps at different scales.	23
Figure 8: Maps of cortical layers in the BigBrain, selected for same view as in Figure 6	24
Figure 9: JuGEx tool as a plugin in the interactive atlas viewer	27
Figure 10: Python code for JuGEx, imported into private collaboratory from viewer plugin	27

History of Changes made to this Deliverable (post Submission)

Date	Change Requested / Change Made / Other Action
06 May 2020	Deliverable submitted to EC
30 Jul 2020	Resubmission with specified changes requested in Review Report Main changes requested: <ul style="list-style-type: none"> • “It is also unclear why such a public document make extensive use of Components numbers which were only internal management identifiers of the project and for which an index is not even given in the document – this should be editorially cleaned before the deliverable is published.” • Although providing some useful information about the outputs per KR, the document is lacking sufficient information to validate the achievement of the KRs (see milestones section).
28 Aug 2020	Revised draft sent by SP/CDP to PCO. Main changes made, with indication where each change was made: <ul style="list-style-type: none"> • This Deliverable has been produced according to instructions, hence the use of references to Components. We added an Index of Components to make it easier for the reader to look up the information. • We consider MS5.3.2, MS5.3.3, MS5.3.5, MS5.5.13 and MS5.6.3 as validated and achieved. Extensive commentary has been submitted in the observation letter to EC on 27 Aug 2020 (HBP_SGA2_M24 Review Report Observations_200827.docx, page 9)
31 Aug 2020	Revised version resubmitted to EC by PCO via SyGMA

1. Overview

The Neuroinformatics Platform (NIP) contributed to the release of a new channel for access to the results of the Human Brain Project in the fall of 2019: the EBRAINS web portal (<https://ebrains.eu/>). Here you can find information about the effort of the HBP to build a shared digital brain research infrastructure. All HBP Platforms contribute to the EBRAINS web portal. The portal will be a key instrument for displaying tools and services targeting the brain research community and related communities in the coming years. In the present version of the portal, the Neuroinformatics Platform has delivered services ensuring that data needed to build brain models can be brought together, annotated, made accessible and prepared for analysis. Data from the HBP make up most of the content at the moment. However, the services have already started to include data from projects external to the HBP and will continue to expand in this direction. A new collaboration with journals has started. The aim is to make it possible for brain scientists to publish their data on EBRAINS at the same time as their articles are published in scientific journals. EBRAINS provides citable DOIs for the data together with data descriptors and information needed to interpret and re-use the data. The three-dimensional brain atlases developed by the HBP are tightly coupled to the data sharing service. They serve as tools for navigation and exploration, and as entry points for finding and analysing data, based on location. In close interaction with user communities, a range of new tools and workflows for analysis of structural data as well as functional dynamic data have been prepared. They will be made available through the EBRAINS web portal in 2020. Key to the successful uptake of the tools and services is the EBRAINS High-Level Support Team, which will have increased capacity for supporting external users in the next Project phase.

2. Introduction

In the second year of SGA2, the Neuroinformatics Platform (NIP), together with the HBP Communication team, paved the way for the new EBRAINS web portal at <https://ebrains.eu>. The new portal presents key services originating from the Neuroinformatics Platform, including *Find Data* (Knowledge Graph Search), *Share data* (service for curation and storage of research data and models, making data and models FAIR), and *Brain atlases* (service for online access to the HBP atlases). The portal also provides access to brain simulation and neurorobotics services and will during the next Project phase continue to be populated with more tools and services based on previous (as reported below) and ongoing developments in HBP.

This document provides an overview of Key Results (KRs), including releases, prototypes, demonstrations, major updates, and impacts of tools and services delivered by SP5 between April 2019 and March 2020. KR5.8 (Tools to enable complete workflows from metadata annotation, data curation, and spatial integration in reference atlases to viewing of multi-level data) involves tools that are described under the various KRs they support.

KR5.1: Making HBP data and models FAIR - Section 3.

- The new metadata schemas openMINDS (refined from MINDS) and SANDS (providing formalised location metadata for anchoring of data to atlases) have been released, resulting in improved interoperability and discoverability of data via the EBRAINS Knowledge Graph.
- The curation workflow is improved to facilitate the communication with researchers and lower the threshold for publishing data on EBRAINS.
- Over 700 datasets have been released through the EBRAINS Knowledge Graph (<https://kg.ebrains.eu/search/>)

KR5.3: Workflow from image data to extracted features - Section 4.

- The QUINT analytical workflow has been updated with a new GUI and now includes non-linear atlas registration for higher precision of analyses. The workflow has been used in several large scale rodent image analysis projects, also outside of the HBP.
- The Connectomic Composition Predictor (CCP) tool has been updated to use gene modules for predictions for improved signal to noise. It can now also predict projections derived from the mouse light database maintained at Janelia farm (<http://ml-neuronbrowser.janelia.org/>).

KR5.4: Assigning spatial metadata to HBP data and models - Section 5.

- All HBP brain reference atlases have received major updates.
 - The BigBrain model has been extended with the first microscopic resolution 3D maps of cytoarchitectonic areas and cortical layers.
 - The Waxholm Space rat brain atlas has been updated with detailed and improved subdivisions of the cerebral cortex, thalamus, basal ganglia, and vestibular system.
- A new online tool for interactive alignment of partial volumes to high-resolution templates (VoluBA) was released.
- The tools for spatial registration of serial 2D images to atlas have been updated with non-linear image registration (VisuAlign, addition to QuickNII).
- Improved workflows for anchoring images to atlas and assignment of spatial location metadata (KR5.1) were released.

KR5.6: Interactive exploration of function, connectivity, cytoarchitecture and receptor architecture at sub-laminar resolution - Section 6.

- The interactive atlas viewer received a major update, with improved workflows for querying updated multimodal, region-specific brain atlas data (KR5.4), maximised views for any axial or oblique section for better appreciation of high-resolution atlases, better integration with the Knowledge Graph, and user instructions for navigating the atlas.

KR5.7: Queries for structural parcellations, genetics, and activity for user-defined regions of interest - Section 7.

- The JuGEx (JuBrainGeneExpression) algorithm for differential analysis of gene expressions by atlas regions has been implemented both as a plugin of the atlas viewer, and as a python library accessible via the Collaboratory, allowing interactive and advanced programmatic use.
- The Gene Regulatory Network explorer has reached its final software version.

KR5.9: A user-driven, data-sharing and data-management infrastructure - Section 8.

- All Knowledge Graph components have iteratively been improved based on user feedback. Various systems have been integrated with a focus on the systems' sustainability and extensibility.
- The Collaboratory has been re-engineered. It now provides fully public collabs, wiki functionality, collaborative editing of Office documents, interactive programming notebooks, an improved user interface, and powerful user & permission management.
- An Image Service was developed to provide atlas viewers and other tools fast and interactive remote access to very large images.

KR5.11: Workflows to leverage state-of-the-art computer vision to object detection and segmentation problems - Section 9.

- Using HBP compute resources, ilastik classifiers can now be trained directly in the web browser. Batch processing on supercomputer resources is possible seamlessly out of the ilastik desktop application.
- New active learning workflow and the deep learning back end allow analysis of large electron microscopy volumes of neural tissue, light microscopy volumes of cleared brain, whole slide imaging, or high-throughput screening data.

KR5.12: Toolbox to support analysis workflows for electrophysiological activity data - Section 10.

- The Elephant tool has received a refactoring of the function documentation to facilitate user accessibility, in addition to new analysis functions, such as analysis of rate vector trajectories.
- The Neo data model has received improved input handling as well as novel support for representing imaging time series.

KR5.13: The neural activity resource (NAR) to register, annotate and browse HBP activity datasets - Section 11.

- Metadata schemas for multi-electrode recordings, EEG, ECoG, MEG and two-photon calcium imaging have been deployed within the Knowledge Graph, underlying the in-depth curation of datasets using these methods (see KR5.1). Support for these experiment types, as well as for tracking provenance of data analysis and simulation workflows, has been added to the fairgraph Python library, which is now being extensively used within the Brain Simulation Platform (SP6).

KR5.14: HBP and non-HBP affiliated researchers using interoperable platforms - Section 12.

- The High Level Support Team (HLST) has successfully delivered support for users of all HBP Platforms and ingested data and models from several non-HBP projects (including five Voucher projects) into HBP/EBRAINS services.
- KnowledgeSpace has ingested released data and models from the EBRAINS Knowledge Graph and supported their integration with brain research concepts and data, models, and literature from other neuroscience data/model repositories.
- The Virtual Brain (TVB) has integrated its pipelines into EBRAINS and TVB Brain Tumor data has been included into EBRAINS.

3. Key Result KR5.1 “A well-defined, well-documented and proven metadata curation approach which makes HBP data and models FAIR (Findable, Accessible, Interoperable, Re-usable) in a consistent and user-friendly way”

The HBP curation service enables and improves the workflow for ingesting heterogeneous neuroscientific data and models into the EBRAINS Knowledge Graph (KG). Its primary aim is to provide the neuroscientific community with interoperable high-quality data and models. Along with *de facto* curation of data and models, the curation team develops technical implementations, routines and standards. These developments ensure that publication of data/models through the EBRAINS Knowledge Graph maximises each dataset/model’s Findability, Accessibility, Interoperability and Reusability (FAIR). Semi-automation of the curation process ensures that the curator’s involvement in making data FAIR is primarily quality improvement and control.

3.1 Outputs

3.1.1 *Overview of Outputs*

3.1.1.1 List of Outputs contributing to this KR

- Output 1: Optimised metadata structure
 - SANDS (Released 09.2019; C2240, C2241, C2242, C2284).
 - openMINDS v3, available since 05.2019 (C2284).
- Output 2: Curation workflow (v2.1)
 - Request and metadata forms. Released in October and November 2019 (C2284).
 - Fairgraph. Released June 2019 (C1463, C2522, C2523).
 - Preview link for data providers in the EBRAINS KG (published July 2019, C2822).
 - Updated documentation, supporting material, and tutorials for data providers. Published March 2020. (C2240, C2242, C2284, C2522, C2523).
- Output 3: Ontology development - constantly updated (C2285).
- Output 4: Data and models available in the EBRAINS KG.

3.1.1.2 How Outputs relate to each other and the Key Result

Output 1 and 3 are prerequisites of Output 2 in improving the curation service for optimising compliance with the FAIR guidelines. Output 4 is the result of Outcomes 1-3.

3.1.2 *Output 1 - Optimised metadata structure*

openMINDS and SANDS are metadata schemas/standards developed by the HBP curation team to:

- 1) Make shared neuroscientific datasets and models searchable
- 2) Enable interoperability between multimodal and heterogeneous data

- 3) Introduce flexible standards that can be continuously developed and easily adapted by other initiatives.

The schemas are developed by the curation team in close collaboration with the EBRAINS KG and ontology teams (Output 2). Both schemas are developed on GitHub and defined in JSON-schemas to facilitate their management and usage in software tools, such as the EBRAINS KG, the HBP interactive atlas viewers, or The Virtual Brain (TVB).

[openMINDS¹](https://github.com/HumanBrainProject/openMINDS) (open Metadata Initiative for Neuroscience Data Structures) will replace the currently used MINDS schemas, because it extends the data query possibilities and descriptions to individual files of registered datasets. Furthermore, openMINDS is better synchronised with corresponding ontologies (Output 2), and other schemas developed inside and outside of HBP. The previous acronym MINDS (Minimal Information for Neuroscience Data Sets) was replaced to avoid erroneous references to another metadata schema.

[SANDS²](https://github.com/HumanBrainProject/SANDS) (Spatial Anchoring of Neuroscience Data Structures) will extend the current descriptions of data in the EBRAINS KG to include user-assessed and curator-verified attribution of anatomical location to the data. SANDS will allow the HBP interactive atlas viewers to extract, transform and display the corresponding data in the HBP brain atlas (see also KR5.4)

Table 1: Output 1 Links

Component	Link to	URL
C2284	Technical and User Documentation	https://github.com/HumanBrainProject/openMINDS
C2240 C2241 C2242 C2284	Technical and User Documentation	https://github.com/HumanBrainProject/SANDS

3.1.3 *Output 2 - Curation workflow (v2.1)*

The curation services enhance, elaborate and maximise the visibility and interoperability of datasets published on the EBRAINS KG in accordance with the FAIR principles. Changes from v2.0 (M12, March 2019) are the following:

- Improved access for users via the “curation request form” October 2019 (C2284) and the “Metadata questionnaire form” which since November 2019 (C2284) is available through [HBP³](https://www.humanbrainproject.eu/en/explore-the-brain/share-data/) and [EBRAINS⁴](https://ebrains.eu/services/data-knowledge/share-data) websites.
- Model curation pipeline (<https://github.com/yzerlaut/model-curation>) released in December 2019 to improve the efficiency of curation of models in the EBRAINS KG.
- Updated documentation, supporting material, and tutorials for data providers published on the EBRAINS Share Data web page and on the new HBP Wiki Collaboratory, <https://wiki.humanbrainproject.eu/bin/view/Collabs/data-curation>, (C2240, C2242, C2284, C2522, C2523).

In addition, several tools are now in place to improve the curation workflow in SGA3:

- Fairgraph (KR5.13) enables development of more automated metadata delivery systems for the curation workflow, thereby improving quality monitoring and overview and making the curation process more efficient.

¹ <https://github.com/HumanBrainProject/openMINDS>

² <https://github.com/HumanBrainProject/SANDS>

³ <https://www.humanbrainproject.eu/en/explore-the-brain/share-data/>

⁴ <https://ebrains.eu/services/data-knowledge/share-data>

- The hbp_archive v0.9.0 (<https://pypi.org/project/hbp-archive/>) provides more efficient uploading and allows users to specify the conditions and details of their own data upload (C2522, C2523).
- Curation monitoring and management are facilitated with a ticketing system (KR5.14) and the EBRAINS KG query builder (KR5.9).

Table 2: Output 2 Links

Component	Link to	URL
C2284	Request and metadata forms	https://nettskjema.no/a/104328#/ https://nettskjema.no/a/125558/
C1463 C2284	Technical Documentation	Fairgraph (KR 5.13) https://pypi.org/project/fairgraph/
	User Documentation	https://github.com/HumanBrainProject/fairgraph
C2284	Technical and User Documentation	Model curation pipeline https://github.com/yzerlaut/model-curation
C1463 C2522 C2523	HBP archive Technical Documentation	https://pypi.org/project/hbp-archive/
C2822	Software Repository	EBRAINS KG (KR5.9) https://kg.ebrains.eu/search/
	Technical and User Documentation	https://kg.ebrains.eu/
C2822 C1474 C1477	Software Repository, Technical Documentation	EBRAINS KG Query (KR5.9) https://kg-editor.humanbrainproject.eu/query-builder Access protected (request access at kg@ebrains.eu)
C2240 C2242 C2284 C2522 C2523	Technical and User Documentation	Supporting material https://wiki.humanbrainproject.eu/bin/view/Collabs/data-curation

3.1.4 Output 3 - Ontology development

Ontologies are used to facilitate metadata standardisation in an effort to align with global neuroscientific nomenclature practices and to make the intricate and established relationships within and between neuroscientific data and metadata accessible.

The enrichment of existing ontologies is performed as a collaborative workflow between the HBP curation team and US Neuroscience Information Framework (NIF). This collaboration facilitates alignment of terminologies across projects (e.g. the US BRAIN initiative, and INCF Neuroshapes). The development of ontologies is an ongoing and continuous process. The EBRAINS KG currently pulls the latest relevant information from the NIF ontology services on a daily basis. Recent progress includes (C2285):

- Neuron Phenotype Ontology development
- Atlas ontology alignment for openMINDS and SANDS
- Methods and modalities improvements

Several improvements to the ontology engineering workflow include the automation of ontology additions, new python tools for NIF ontology work, and new ontology workflow documentation (Table 3).

Table 3: Output 3 Links

Component	Link to	URL
C2285	Technical and User Documentation	<ol style="list-style-type: none"> 1) Neuron Phenotype Ontology development and tools for working with this ontology https://github.com/SciCrunch/NIF-Ontology/blob/neurons/ttl/neuron-development.ttl https://pypi.org/project/neurondm/ 2) Atlas ontology alignment for openMINDS and SANDS https://github.com/SciCrunch/NIF-Ontology/commits/atlas 3) Methods and modalities improvements https://github.com/tgbugs/pyontutils/tree/master/nifstd/development/methods 4) Ontology workflow improvements https://github.com/SciCrunch/NIF-Ontology/releases 5) python tools for NIF ontology work https://pypi.org/project/nifstd-tools/ 6) Ontology workflow documentation https://github.com/tgbugs/pyontutils/blob/master/nifstd/scigraph/R/EADME.org

3.1.5 Output 4 - Data and models available in the EBRAINS KG

Output 4 is the result of Outputs 1, 2 and 3 in KR5.1, Output 2 of KR5.4, and Output 1 of KR5.9 (see Tables and information under each of these Outputs). The number of datasets curated and published in EBRAINS KG increased from more than 160 to over 700 during the second year of SGA2. In addition, more than 60 models and over 30 pieces of software have been curated and published through the KG. (<https://kg.ebrains.eu/search/>)

3.2 Validation and Impact

3.2.1 Actual and Potential Use of Output(s)

Output 1: Optimised metadata structure

openMINDS allows a more flexible and globally applicable metadata standard, making it easier to publish heterogeneous datasets on the EBRAINS KG. Beta testing in a replicant EBRAINS KG Space begins in April/May 2020. openMINDS is accessible on GitHub, available to external users, and already in use by the HBP curation team and external ontology developers. Similarly, the new metadata standard for spatial anchoring (SANDS) will introduce interoperable spatial standards and create a universal system for describing spatial location across all neuroscientific datasets.

Output 2: Curation workflow (v2.1):

The new curation workflow facilitates metadata validation and overall more cost-efficient curation.

- The updated request and metadata forms make it significantly easier for researchers to describe their data/metadata.
- The new curation pipeline for models provides a documented and standardised approach for modellers to maximise the impact of sharing their models.
- A preview link in EBRAINS KG shows researchers how their datasets will look when published, as a motivation for improving the dataset descriptions.
- Updated documentation, supporting materials and tutorials, based on feedback from multiple data providers over several years, to improve communication with researchers.

- The new Fairgraph and EBRAINS KG Query functionalities provide an API foundation for more automated metadata ingestion and curation progress monitoring systems in SGA3.

Output 3: Ontology development

Ontologies are continuously adapted and updated to the needs determined by the data received for curation. Through collaborative efforts involving the INCF and NIF, the HBP ontology team ensures that standards used by HBP are compatible with other databases internationally (e.g. US BRAIN initiative databases).

Output 4: Data and models available in the EBRAINS KG

The data and models available in the EBRAINS KG are produced and used in the context of research in the HBP. The sharing through the EBRAINS KG allows re-use of data in new combinations and uptake of models by other researchers, with citation of the DOIs provided.

3.2.2 Publications - not applicable to KR5.1

4. Key Result KR5.3 “Complete workflow going from heterogeneous experimental image data to extracted quantitative features defined in rodent atlas space, available for general use”

KR5.3 are pipelines for mining integrated neuroanatomical data from different sources, including those curated by the Neuroinformatics Platform. Delivering neuroscience data with suitable analytic tools is a core HBP goal.

4.1 Outputs

4.1.1 Overview of Outputs

4.1.1.1 List of Outputs contributing to this KR

- Output 1: QUINT workflow for extracting cellular features from high-resolution rodent brain images (C2275, latest release January 2020) (dependent components C2277, C3074, C2235).
- Output 2: Connectomic Composition Predictor (CCP, C935). Latest release March 2020.

4.1.1.2 How Outputs relate to each other and the Key Result

The Outputs will foster new analytical possibilities, allowing researchers to analyse spatial distributions of features in the brain, e.g. cells, connections, or proteins (QUINT workflow, Output 1) and predict the connectivity of specific brain regions with the connectivity predictor tool (CCP, Output 2).

4.1.2 Output 1: QUINT workflow

The QUINT workflow (Figure 2) is based on extensive user feedback and represents a significant improvement over the previous workflows from the first year of SGA2 (KR5.3 - Output 1; SGA2 Deliverable D5.8.1 (D35.1, D19)).

First, the new user-guided non-linear registration of the experimental images to the 3D reference atlas allows spatial anchoring of images to atlas maps with higher anatomical precision (KR5.4). The non-linear registration is performed using the VisuAlign tool v0.8 (RRID:SCR_017222), following initial linear registration obtained with QuickNII v2.2 (C2277) (RRID:SCR_016854). Output from VisuAlign is compatible with the Nutil (Neuroscience image processing and analysis utilities) tool (C3074) (RRID:SCR_017183) for quantification of the distribution of features across atlas regions.

Second, a GUI with intuitive steps and help texts was developed for Nutil v0.40 (C3074), allowing users to transform images and quantify features in atlas-defined regions-of-interest. The Nutil Quantifier function outputs an object coordinate file allowing the users to visualise their objects as point clouds using 3D viewers such as MeshView (Meshview: <https://www.nitrc.org/projects/MeshView/>; RRID:SCR_017222) or share them via EBRAINS as derived data (e.g. <https://doi.org/10.25493%2F6DYS-M3W>; <https://doi.org/10.25493/3FC2-R3R>).

The current state of the art makes use of custom codes for analysis of whole brain volumes by registration to reference atlases and often relies on 3D reconstruction (Kim *et al.* 2017⁵; Vandenberghe *et al.* 2016⁶; Furth *et al.* 2018⁷; Xiong *et al.*, 2018⁸). The QUINT workflow goes beyond the state of art by contributing to discoveries based on histological experimental data, enabling precise quantifications, visual output of the data, without the need for coding knowledge nor usage of 3D reconstruction. In addition, this workflow can be used with different rat or mouse reference atlases: Allen Mouse atlas⁹ CCF v3 (2015 and 2017) and the Waxholm Space (WHS) atlas of the Sprague Dawley rat brain v2 and v3 (Papp *et al.* 2014¹⁰; Kjonigsen *et al.* 2015¹¹; Osen *et al.* 2019¹²).

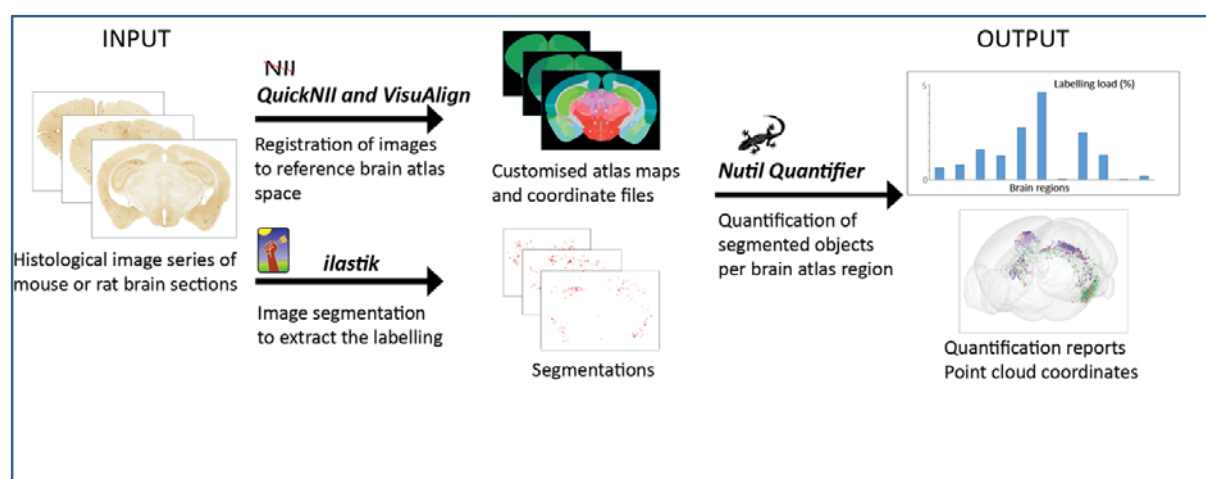


Figure 2: QUINT workflow for quantification of features using 3D reference atlases

⁵ Kim, Y., Yang, G. R., Pradhan, K., Venkataraju, K. U., Bota, M., Garcia Del Molino, L. C., *et al.* (2017). Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* 171, 456.e422-469.e422. doi: 10.1016/j.cell.2017.09.020

⁶ Vandenberghe, M. E., Herard, A. S., Souedet, N., Sadouni, E., Santin, M. D., Briet, D., *et al.* (2016). High-throughput 3D whole-brain quantitative histopathology in rodents. *Sci. Rep.* 6:20958. doi: 10.1038/srep20958

⁷ Fürth, D., Vaissière, T., Tzortzi, O., Xuan, Y., Martin, A., Lazaridis, I., *et al.* (2018). An interactive framework for whole-brain maps at cellular resolution. *Nat. Neurosci.* 21, 139-149. doi: 10.1038/s41593-017-0027-7

⁸ Xiong, J., Ren, J., Luo, L., and Horowitz, M. (2018). Mapping histological slice sequences to the allen mouse brain atlas without 3D reconstruction. *Front. Neuroinform.* 12:93. doi: 10.3389/fninf.2018.00093

⁹ © 2004 Allen Institute for Brain Science. Available from: mouse.brain-map.org

¹⁰ Papp, E. A., Leergaard, T. B., Calabrese, E., Johnson, G. A., and Bjaalie, J. G. (2014). Waxholm space atlas of the Sprague-Dawley rat brain. *Neuroimage* 97, 374-386. doi: 10.1016/j.neuroimage.2014.04.001

¹¹ Kjonigsen, L. J., Lillehaug, S., Bjaalie, J. G., Witter, M. P., and Leergaard, T. B. (2015). Waxholm space atlas of the rat brain hippocampal region: three-dimensional delineations based on magnetic resonance and diffusion tensor imaging. *Neuroimage* 108, 441-449. doi: 10.1016/j.neuroimage.2014.12.080

¹² Osen KK, Imad J, Wennberg AE, Papp EA, Leergaard TB. Waxholm Space atlas of the rat brain auditory system: Three-dimensional delineations based on structural and diffusion tensor magnetic resonance imaging. *Neuroimage*. 2019 Oct 1;199:38-56. doi: 10.1016/j.neuroimage.2019.05.016.

Table 4: Output 1 links

Component	Link to	URL
C2275	Technical Documentation	https://wiki.ebrains.eu/bin/view/Collabs/quint/
	User Documentation	Publication P2245: https://doi.org/10.3389/fninf.2019.00075
C2277	Software Repository & Technical Documentation	https://www.nitrc.org/projects/quicknii https://kg.ebrains.eu/search/instances/Software/e5395520-4a17-4be9-97da-e12a663fcd81
	User Documentation	https://www.nitrc.org/projects/quicknii/ ; P2182: https://doi.org/10.1371/journal.pone.0216796
C3074	Software Repository & Documentation	https://www.nitrc.org/projects/nutil/ https://kg.ebrains.eu/search/instances/Software/95053862-450a-45f7-8f16-eb3f781376b4

4.1.3 Output 2: Connectomic Composition Predictor (CCP)

The CCP is a Jupyter notebook tool that delivers a predicted mesoconnectome, based on tract tracing data from the [Allen Institute](#)¹³. Version 2.1, released March 2020, provides the ability to use gene modules for predictions. The gene modules are generated using dictionary learning to sparsify the representation, and thus reduce noise, as well as based on cell-type specific gene profiles derived from single-cell RNAseq. The predictions are validated by a number of quantitative measures. We find that dictionary based modules yield similar predictive accuracy compared to using the full gene expression data set, whereas the RNA-seq derived modules did not perform as well. The [mouse light database](#)¹⁴ provides axonal morphologies, which we have used to generate cell-type specific region-based connectivity. We have added code to the CCP to use gene expression patterns to predict these mouse-light mesoconnectomes as well. The current state of the art is the prediction of binary connectivity for an older version of the Allen Brain Atlas using wild-type mouse data (e.g. Ji *et al.* 2014¹⁵). Here, we go beyond the state of the art by making non-binary predictions of the connection strength, using the latest version of the Common Coordinate Framework (CCF 3.0), which includes cortical layers, and including the results from *Cre*-mouse lines to make the predictions cell-type specific.

Table 5: Output 2 Links

Component	Link to	URL
C0935	Technical and User Documentation	https://github.com/ntimonid/Connectomic-Composition-Predictor-CCP-

4.2 Validation and Impact

4.2.1 Actual and Potential Use of Output(s)

Output 1: The updated QUINT workflow allows researchers to achieve precise atlas maps for their experimental histological data and perform quantitative analyses of labelled objects in whole brains and specific regions. The tools are intuitive and do not require programming skills or specialist

¹³ <https://connectivity.brain-map.org/projection>

¹⁴ <http://ml-neuronbrowser.janelia.org/>

¹⁵ Ji, S., Fakhry, A. & Deng, H. (2014) Integrative analysis of the connectivity and gene expression atlases in the mouse brain. *Neuroimage*, **84**, 245-253. doi: 10.1016/j.neuroimage.2013.08.049

knowledge. The Jackson Laboratory (Bar Harbor, USA) has used the QUINT workflow extensively and contributed valuable feedback. Several laboratories have expressed interest after our presentations at several conferences and workshops (INCF 2019 Warsaw, Poland; Cajal school of Neurosciences, Bordeaux, France, Sep. 2019; SfN 2019, Chicago, USA; HBP student conference 2020, Pisa, Italy; HBP Summit, Feb 2020, Athens, Greece). Download rates of the tools are increasing.

Output 2: The CCP will be useful for students and neuroscientists studying mouse brain mesoscopic connectivity, computational neuroscientists incorporating whole-brain connectivity patterns in their simulations, and translational neuroscientists linking gene expression patterns to distortions in connectivity. For the release of CCP version 2.1, we included three use cases that illustrate the potential use of the output. CCP was presented at “BrainsUnited” workshop at INCF 2019 Warsaw, for representatives from HBP, Partnering Projects and other brain initiatives. Participants expressed interest in applying the techniques to the marmoset brain, as the Japan brain initiative is collecting both gene expression data and tract-tracing connectivity data.

4.2.2 Publications

Software used in the QUINT (Output 1) workflow has been shared and described in publications.

- P2182: Puchades *et al.* 2019, Spatial registration of serial microscopic brain images to three-dimensional reference atlases with the QuickNII tool. doi: [10.1371/journal.pone.0216796](https://doi.org/10.1371/journal.pone.0216796).
- P2245: Yates *et al.* 2019, QUINT: Workflow for Quantification and Spatial Analysis of Features in Histological Images From Rodent Brain. doi: [10.3389/fninf.2019.00075](https://doi.org/10.3389/fninf.2019.00075).

For Output 2 (CCP) a manuscript that demonstrates a prediction pipeline for perform multimodal data integration to improve the accuracy of the predicted mesoconnectome and support other neuroscience use cases has been made available.

- P2499: Temonidis *et al.* 2019, Prediction of a cell-type specific mouse mesoconnectome using gene expression data <https://doi.org/10.1101/736520>

5. Key Result KR5.4 “A well-defined, well-documented and proven approach for assigning spatial metadata to HBP data and models.”

A key goal of the HBP’s Neuroinformatics Platform is to organise and share heterogeneous, multi-level and multi-modal neuroscientific data from human, rat, and mouse brains by their anatomical location defined in brain reference atlases, thereby providing a topographically organised information system. To achieve this goal, curation of datasets addresses not only structured general metadata (Tier 1 curation), but also proper definition of location in the brain (Tier 2 curation). Such location assignment can be specified at the level of brain regions (reference parcellation of the brain) or in terms of spatial coordinates (reference coordinate space). This KR establishes the standards, brain reference atlases, and methods to assign such location information to data. The resulting tools are used in the Curation Workflow (KR5.1), included in analytic workflows (KR5.3) and accessible to external users of the HBP RI.

5.1 Outputs

5.1.1 Overview of Outputs

5.1.1.1 List of Outputs contributing to this KR

- Output 1: Improved reference atlases

- JuBrain probabilistic cytoarchitectonic atlas (C2319, C2264)
- Waxholm Space rat brain atlas with additional delineations (C2169)
- Allen Mouse CCF version 3 and Waxholm Space rat brain atlas v3 integrated in HBP's interactive atlas viewer (C2167)
- Output 2: Optimised tools and procedures for registration of brain-image data to reference atlases
 - Release of VoluBA - interactive online service for anchoring of partial volumetric data to high-resolution reference models (May and July 2019; C2267, C2269, C2372)
 - New release of QuickNII tool (May 2019) with improved procedures for anchoring 2D images to reference atlases, nonlinear adjustment with the VisuAlign tool, and inspection and analysis of 2D image data with superimposed custom atlas images with the LocaliZoom tool (C2277, C2244, C2272)

5.1.1.2 How Outputs relate to each other and the Key Result

Output 1 provides better coverage and higher detail of the reference brain parcellations, key for improving the assignment of atlas region names to data (see KR5.1). Tier 2 curation also includes coordinate-based assignment of data. The new SANDS metadata scheme (KR5.1) serves as a basis for location-specific interactions between the Interactive atlas viewer, EBRAINS KG and spatial search. Such coordinate-based location assignment is particularly challenging for high-resolution partial volumetric data in human brains, due to the limited field of view and high variability of brain morphology. With the release of VoluBA (Output 2), we have now provided an interactive online tool to address this problem in a simple and intuitive fashion. In a future release, VoluBA will also employ the new SANDS scheme to provide spatial metadata directly to the Knowledge Graph. More accurate and efficient registration of 2D serial images to atlas is achieved with the improved tools and workflows (Output 2).

5.1.1 *Output 1: New versions of reference atlases*

The JuBrain probabilistic cytoarchitectonic atlas version 18 was released in fall 2019. Besides optimised delineations of many existing areas, this version includes new maps for areas 6d1, 6d2, 6d3, Id7, OP8, OP9, STS1, STS2, 6mp, 6ma, p24ab, p24c, p32, hIP4, hIP5, hIP6, hIP7, hIP8, hOP1, hOc6, Ia, Id2, Id3, Id4, Id5, Id6, FO4, FO5, FO6, FO7, Ig3, CA, IF, MF, VTM, DG, TrS, hPO1, hPO, TE 2.2, TE 2.1, TI, and Tel. With this version, the JuBrain atlas is defined in both the MNI Colin 27 single-subject and the ICBM152 asymmetric multi-subject average reference spaces, and can be explored in both of them in the Interactive atlas viewer. This makes the atlas accessible for a wider spectrum of use cases, and allows for a better integration with other atlases, datasets and tools, that are typically only available in one of the spaces. For example, via the ICBM 152 asymmetric space, the JuBrain atlas now allows for interactive differential gene expression analysis using the JuGEx tool (C2254), which interfaces directly to the Allen brain atlas.¹⁶

The HBP human brain atlases now also include ultrahigh-resolution 3D cytoarchitectonic maps defined in the BigBrain model, each of which corresponds directly to a probabilistic map, hereby providing links between the macroscale, where we capture brain variability, and the microscale, where we provide high anatomical detail (see KR5.6).

The Waxholm Space atlas of the rat brain (C2169), was released as version 3 in May 2019 with several new and improved delineations. Version 4 with detailed subdivisions of the cerebral cortex, thalamus, basal ganglia and vestibular system has been validated by HBP users and will be released via nitrc.org in early SGA3.

¹⁶ © 2010 Allen Institute for Brain Science. Allen Human Brain Atlas. Available from: <http://human.brain-map.org/>

The latest versions of the Allen Mouse CCF version 3 and Waxholm Space rat brain atlas (v3) were integrated into the Interactive atlas viewer and the QuickNII tool.

Table 6: Output 1 Links

Component	Link to	URL
C2319, C2264	Online atlas	https://jubrain.humanbrainproject.eu
	User Documentation	https://kg.ebrains.eu/search/instances/Project/a684ecccc76fc522773286a895bc8436
C2169	Software Repository and documentation	https://www.nitrc.org/projects/whs-sd-atlas

5.1.2 Output 2 Optimised tools and procedures for registration of brain-image data to reference atlases

A common problem in high-resolution brain atlasing is spatial anchoring of high-resolution volumes of interest (VOIs) from specific imaging experiments into the detailed anatomical context of a high-resolution reference model like BigBrain¹⁷, to allow integration of partial volumetric data at the level of cortical areas, individual cortical layers, or subcortical structures. Downloading and interacting with reference models of whole human brains at microscopic resolution is usually out of reach for neuroscientists due to the sheer size of the data and the lack of practical tools. EBRAINS atlas services now provide a first version of a novel interactive volumetric alignment tool - VoluBA - which allows anchoring of such VOIs to very large reference volumes. VoluBA is implemented as a web service that allows users to upload a VOI from their own imaging experiment and perform interactive anchoring by direct manipulation of its position and rotation using the mouse pointer, flipping of 3D coordinate axes, as well as entering of point landmarks for refined alignment with affine components. The resulting transformation parameters can be downloaded for reference and for submission to HBP data curation. The aligned image volume can be directly opened in the Interactive atlas viewer to view the VOI in full anatomical context, compare it to atlas regions, and access data linked to the region of interest. As an experimental feature, we also developed functionality for nonlinear alignment of cortical VOIs, which exploits the equivolumetric model of cortical depth to optimize alignment of cortical layers across subjects.

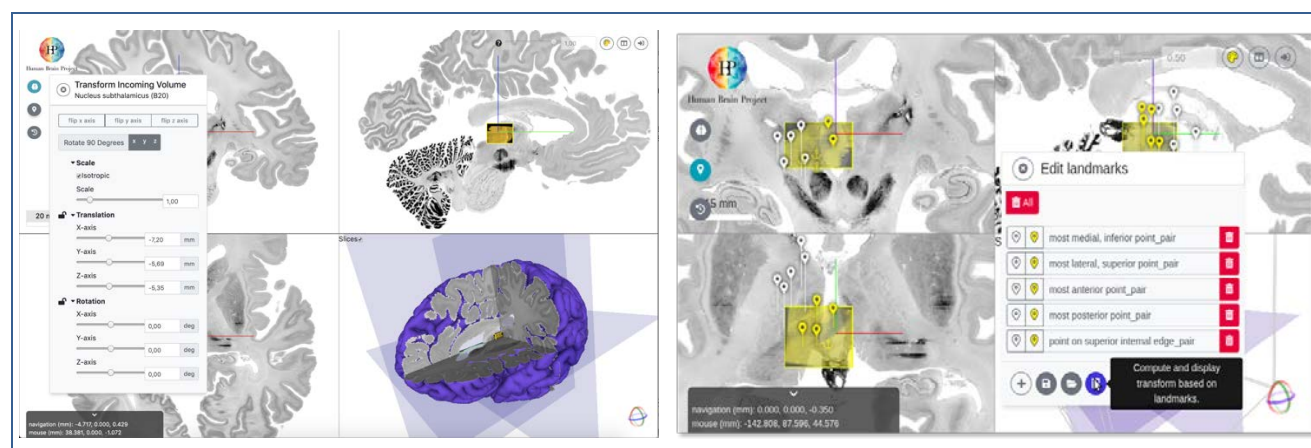


Figure 3: Aligning a partial volume to BigBrain using VoluBA

Left: basic positioning, right: refinement with landmarks.

The affine registration of experimental 2D rodent images to the 3D reference atlases with the QuickNII tool (C2277) (RRID:SCR_016854) is now complemented with a non-linear registration step, using the VisuAlign tool v0.8 (RRID:SCR_017222) released in January 2020. Output from VisuAlign is

¹⁷ Amunts et al., 2013; DOI: 10.1126/science.1235381

compatible with the Nutil tool (KR5.3) (RRID:SCR_017183) for quantification of distributions across brain regions of features extracted with use of the QUINT workflow (KR5.3).

Table 7: Output 2 Links

Component	Link to	URL
C2267	Software repository	https://github.com/HumanBrainProject/cortical-voluba
	User Documentation	https://voluba-user-doc.apps-dev.hbp.eu/
C2269	Online service	https://voluba.apps.hbp.eu/#/
C2372	User Documentation	https://voluba-user-doc.apps-dev.hbp.eu/
C2277	Software Repository	https://www.nitrc.org/projects/quicknii
	Technical Documentation	https://kg.ebrains.eu/search/?facet_type[0]=Software#Software/e5395520-4a17-4be9-97da-e12a663fcd81 and https://www.nitrc.org/projects/quicknii
	User Documentation	https://www.nitrc.org/projects/quicknii/ (P2182 doi:10.1371/journal.pone.0216796)
C2244	User Documentation	https://www.frontiersin.org/articles/10.3389/fnana.2018.00082/full (P1530)
C2272	Data Repository	https://kg.ebrains.eu/search/instances/Project/af8d3519-9561-4060-8da9-2de1bb966a81

5.2 Validation and Impact

5.2.1 Actual and Potential Use of Output(s)

Output 1: The updated reference atlases are available to all online users of the HBP/EBRAINS atlas services, in particular via the Interactive atlas viewer. The reference atlases are potentially interesting to the entire neuroscience community, the size of which we estimate to 70,000, with slightly varying needs of specific target groups, which include neuroanatomists (estimated to 26,000), experimental and computational neuroscientists (estimated to 30,000), the neuroimaging community (estimated to 6,000), neurosurgeons (estimated to 2,500), and brain-inspired AI researchers (estimated to 500).

Output 2: The new VoluBA web application fills a gap among the spatial anchoring tools, since, to our knowledge, there is no other tool which allows interactive positioning of partial volumes in Terabyte-sized reference volumes. VoluBA has been used by the curation teams to anchor a range of human datasets to BigBrain, including highfield fMRI, PLI and cell-stained histological data. The public usage numbers of the VoluBA application, however, are still small, as it is only available to the public since a short time period.

The QuickNII tool (also part of the QUINT workflow, KR5.3) is a much used tool for registration of histological images to rodent brain atlases. The tool is used, e.g. at The Jackson laboratory (Bar Harbor, USA); Norwegian University of Science and Technology (Trondheim, Norway), UPenn (Philadelphia, USA), Korea Brain Research Institute (Daegu, Korea), and Macquarie University (Sydney, Australia). The number of downloads of the tool on the Nitrc.org page has increased during SGA2 year 2 and has reached 1,300. Within HBP, QuickNII has been used for spatial location assignment of 47 datasets, released through the Knowledge Graph during SGA2 year 2.

5.2.2 Publications

Output 2:

- P2182: Puchades *et al.* 2019, Spatial registration of serial microscopic brain images to three-dimensional reference atlases with the QuickNII tool. doi: 10.1371/journal.pone.0216796.
Significance: The QUICKNII tool has a key role in registration of 2D image data to rodent atlases and is a cornerstone of the QUINT analytical workflow (KR5.3)
- P1879: Osen *et al.* 2019, [Waxholm Space atlas of the rat brain auditory system: Three-dimensional delineations based on structural and diffusion tensor magnetic resonance imaging](https://doi.org/10.1016/j.neuroimage.2019.05.016). doi: 10.1016/j.neuroimage.2019.05.016
Significance: Significant update of the reference WHS rat atlas with detailed subdivisions in many anatomical regions enabling precise anatomical location and analysis of experimental results in rats.

6. Key Result KR5.6 “Interactive exploration of function, connectivity, cytoarchitecture and receptor architecture at sub-laminar resolution in the Big Brain template space.”

A unique aspect of HBP’s multilevel Brain Atlas is that it describes brain anatomy at different spatial resolutions, going from the millimetre to the micrometre scale, and provides access to regional characterisations of brain structure, function and connectivity. A particularly challenging problem is the integration of different data modalities at the level of different cortical layers, which are only resolved at the resolution of tens to few hundreds of micrometres. To date, there is no framework available that would allow interactive exploration of brain anatomy at this level. This KR addresses that problem, aiming to provide explorative access to microscopic resolution multimodal data through visual interaction with the high-resolution Big Brain template.

6.1 Outputs

6.1.1 Overview of Outputs

6.1.1.1 List of Outputs contributing to this KR

- Output 1: New releases of interactive atlas viewer (C2802, C2370, C2371)
- Output 2: Ultrahigh-resolution 3D maps and maps of cortical layers for the BigBrain model (C2272, C2264)
- Output 3: Pre-processed data from neuroimaging cohorts available (C2820, C2821)
- Output 4: HPC-based workflows configured and deployed on JURECA (C2820, C2821)

6.1.1.2 How Outputs relate to each other and the Key Result

The Interactive viewer (Output 1) allows 3D browsing of high-resolution brain models like the BigBrain, together with overlays of maps and additional imagery. The recently developed and integrated maps of BigBrain areas and cortical layers (Output 2) allow to identify different cytoarchitectural regions, explore their boundaries, and link the BigBrain model to region-specific multimodal data. In combination with the new VoluBA tool for spatial anchoring (KR5.4), this finally

provides the possibility to aggregate high-resolution multimodal data in the BigBrain space for investigation in detailed, yet whole-brain anatomical context. Outputs 3 and 4 are directly related to the JuBrain probabilistic cytoarchitectonic atlas using this regional delineation to compute connectivity, receptor density and provide high-resolution cell density.

6.1.2 Output 1: New release of interactive atlas viewer

The Interactive atlas viewer, accessible at <https://interactive-viewer.apps.hbp.eu>, has received several updates during the last year (C2371). Overall, it now provides optimised user navigation for selecting different reference atlases, navigating and selecting brain regions, and searching for region specific data (Figure 4). The behaviour of the application now more closely resembles the interaction patterns of popular online services like google maps, and is therefore more intuitive to most users. The Interactive atlas viewer now includes a connectivity browser (Figure 5; release scheduled March 2020), which allows to retrieve a connectivity matrix for the given parcellation if available, and browse or select regions according to their connection strengths (C2260). As of now, average structural connectivity data of the “1,000 brains” cohort is supported. Further improvements include the possibility to enter a maximised view into any axial or oblique section (Figure 6), which allows to better appreciate high resolution atlases, and performance improvements by more efficient data handling.

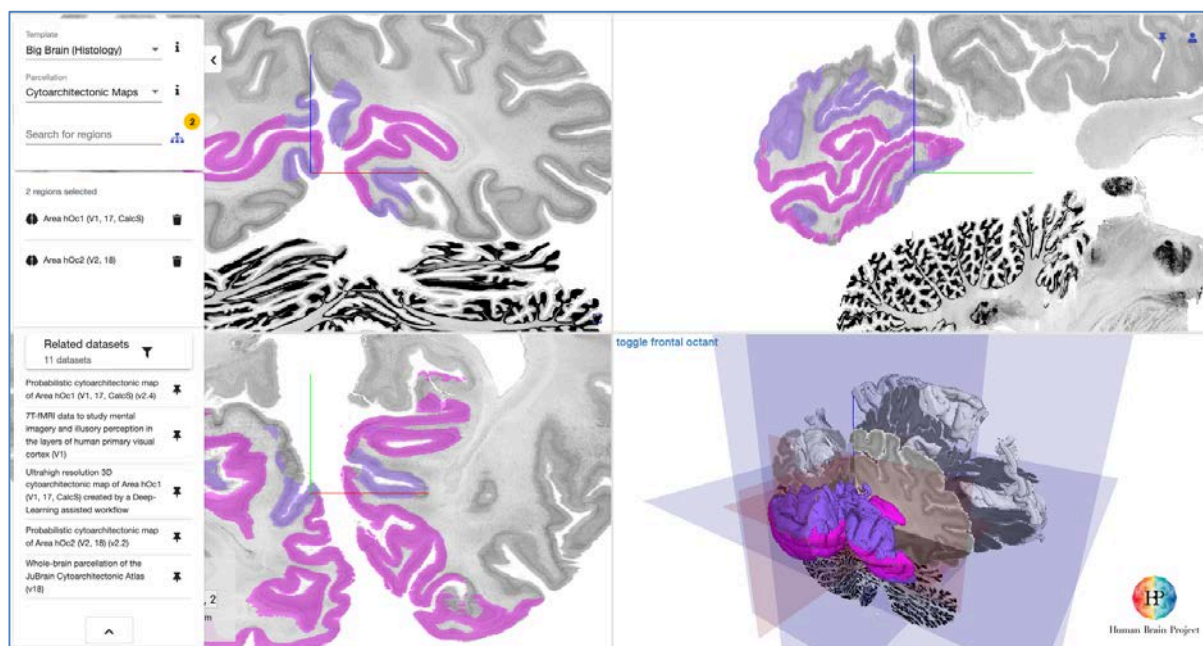


Figure 4: Optimised user interface of interactive atlas viewer (here: BigBrain atlas)

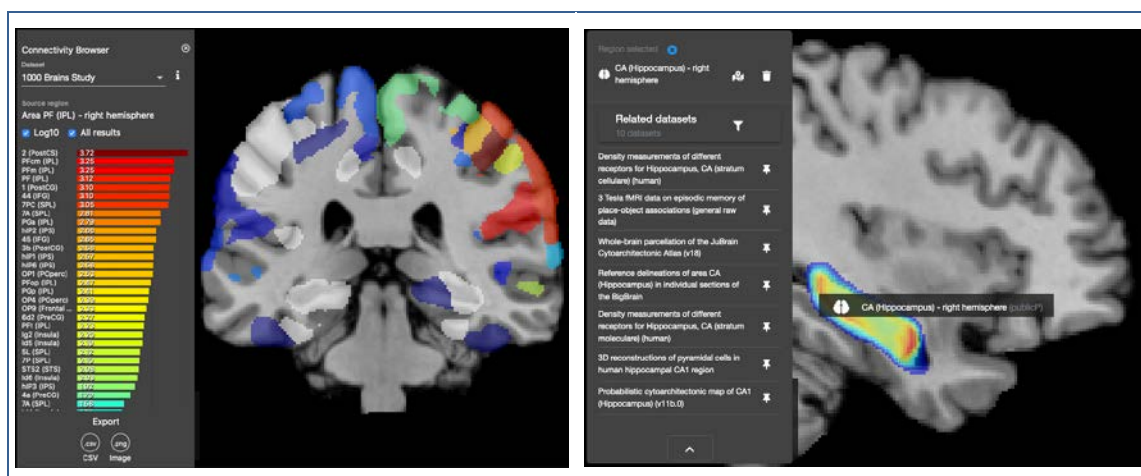


Figure 5: New connectivity browser

New connectivity browser functionality (left) and streamlined multimodal data search (right)

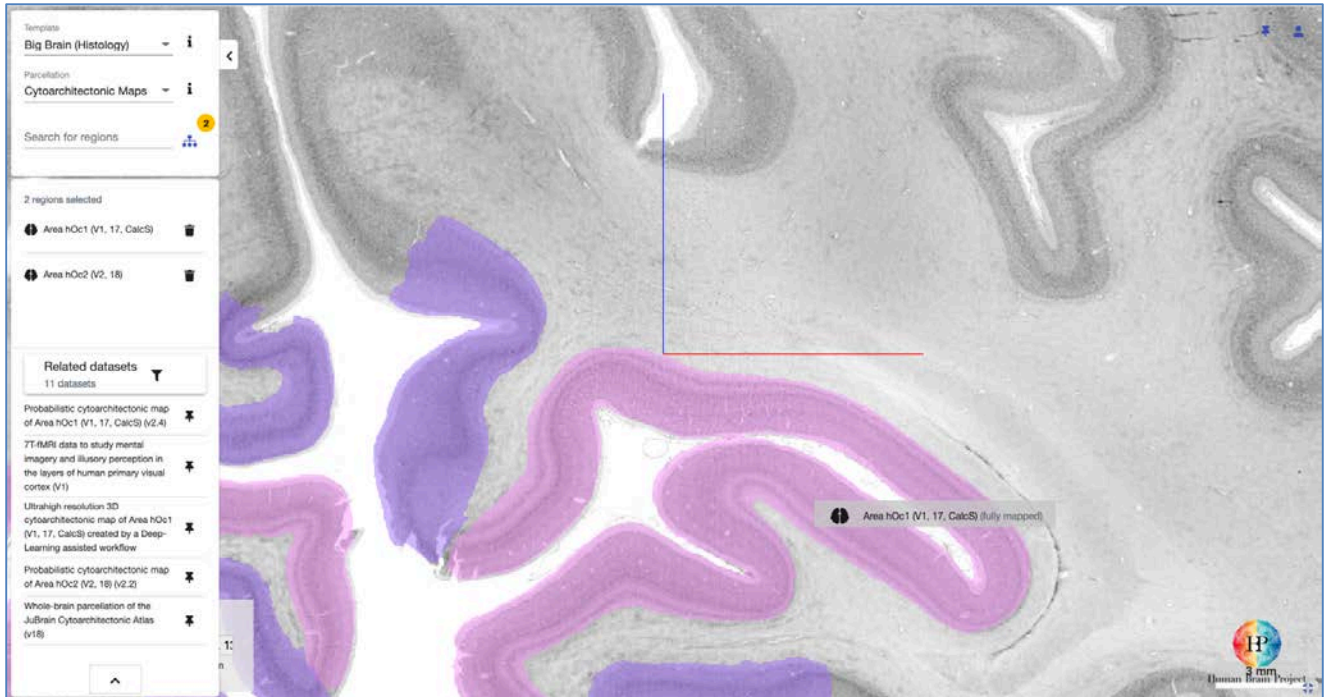


Figure 6: Maximised view of the coronal plane, corresponding to Figure 4

Table 8: Output 1 Links

Component	Link to	URL
C2371	Online Service	https://interactive-viewer.apps.hbp.eu
C2260	User Documentation	https://interactive-viewer-user-documentation.apps-dev.hbp.eu/
C2370		
C2802		

6.1.3 Output 2: Ultrahigh-resolution 3D maps and maps of cortical layers for the BigBrain model

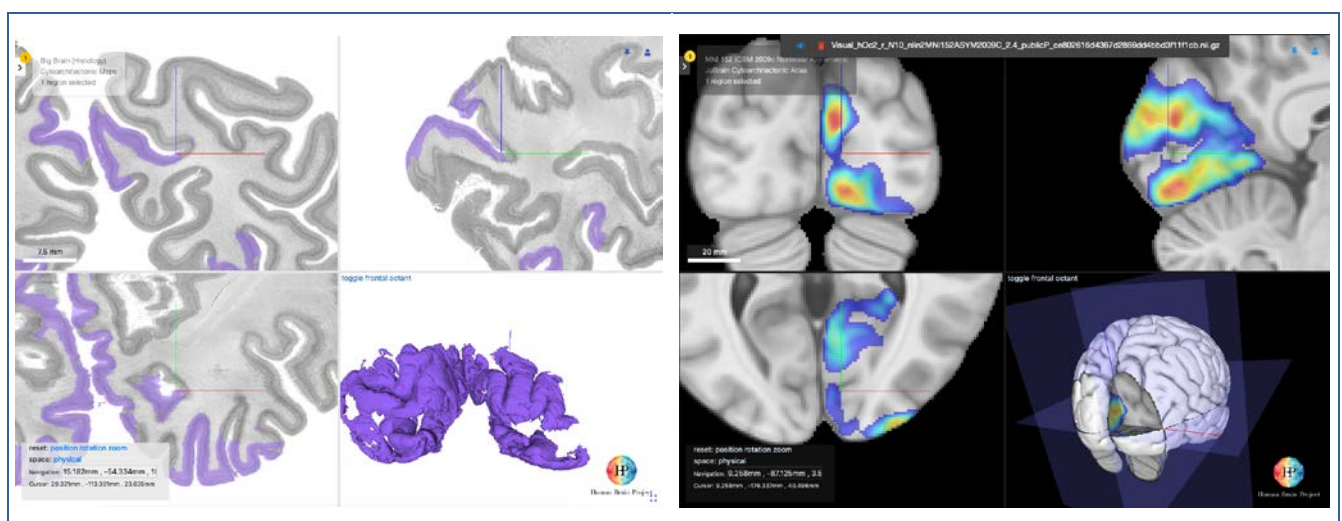


Figure 7: Cytoarchitectonic maps at different scales.

Left: Microscopic resolution 3D map of area hOc2 in BigBrain right: corresponding probabilistic map in ICBM 152 asymmetric Space.

The HBP atlases now include ultrahigh-resolution 3D cytoarchitectonic maps defined in the BigBrain model (C2272), which have been created within SP2 using a Deep Learning algorithm (C2375), and

integrated to the atlas by the human data curation team (C2264). Each of these novel maps corresponds directly to a probabilistic map in the JuBrain cytoarchitectonic atlas. In this way, we have implemented links between the macroscale, where we capture brain variability using probabilistic maps, and the microscale, where we provide remarkably high anatomical detail in a single subject (Figure 7).

Further, we curated and deeply integrated new whole-brain maps of all cortical layers¹⁸ as an external dataset, contributed through a collaboration. These cortical layer maps are available as an additional parcellation of the BigBrain model, and can be selected along with the cytoarchitectonic maps. This allows for the first time interactive 3D exploration of cortical areas and layers at the microscopic scale through a public online service.

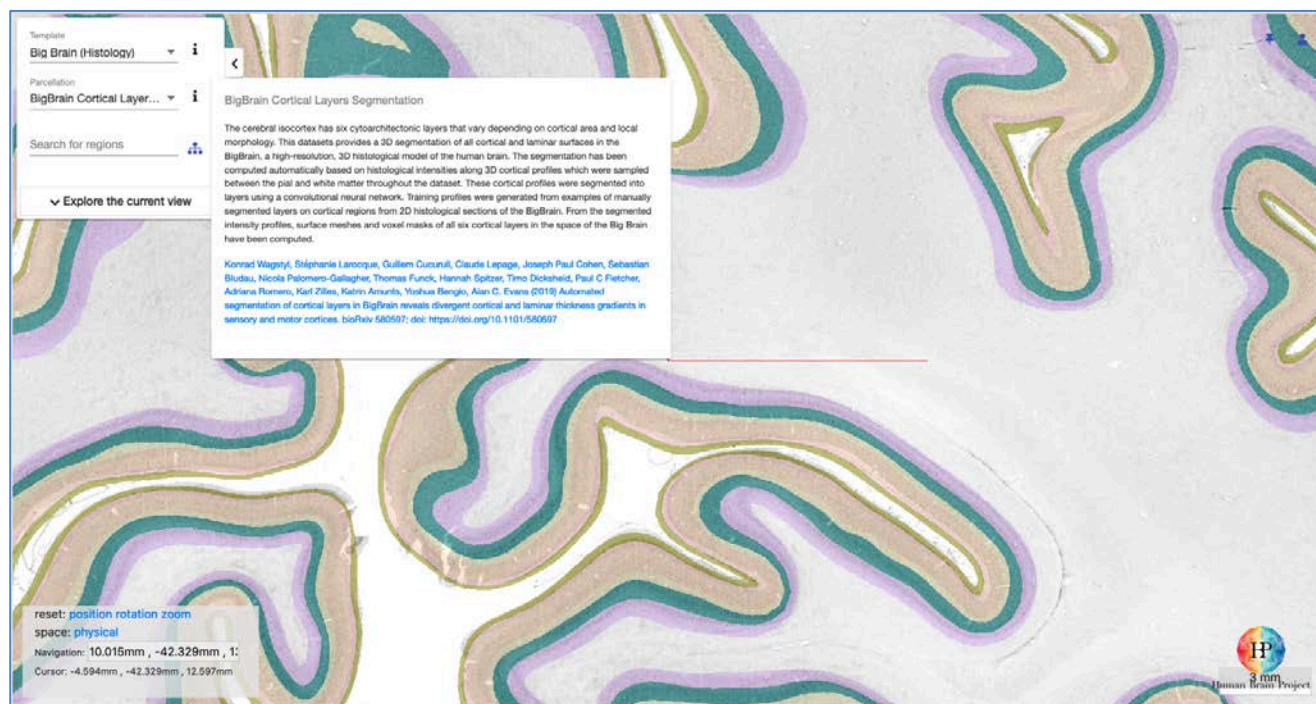


Figure 8: Maps of cortical layers in the BigBrain, selected for same view as in Figure 6

Table 9: Output 2 Links

Component	Link to	URL
C2272	Data Repository and User Documentation	https://kg.ebrains.eu/search/instances/Project/af8d3519-9561-4060-8da9-2de1bb966a81
C2264	Online Atlas	https://bigbrain.humanbrainproject.eu
	User Documentation	https://kg.ebrains.eu/search/instances/Project/af8d3519-9561-4060-8da9-2de1bb966a81

6.1.4 Output 3: Pre-processed data from neuroimaging cohorts available

Structural connectivity (SC) and resting-state functional connectivity (FC) have been calculated for and averaged over several hundreds of healthy subjects from the Human Connectome Project (HCP) and the enhanced Nathan Kline Institute-Rockland Sample (eNKI-RS) data sets. Several brain atlases based on different parcellation approaches have been used including functional Schaefer atlas with 100, 200 and 400 cortical regions as well as anatomical Harvard-Oxford atlas and cytoarchitectonic JuBrain atlas.

¹⁸ Wagstyl K et al. (2019) Automated segmentation of cortical layers in BigBrain reveals divergent cortical and laminar thickness gradients in sensory and motor cortices <https://doi.org/10.1101/580597>

Table 10: Output 3 Links

Component	Link to	URL
C2820	Data Repository	https://kg.ebrains.eu/search/instances/Dataset/50c215bc-4c65-4f11-a4cd-98cc92750977
	Software Repository	https://gitlab.humanbrainproject.org/felixh/hcp_vbm
C2821	Software Repository	https://gitlab.humanbrainproject.org/felixh/hcp_fs/ (access inquiries through support@ebrains.eu)

6.1.5 *Output 4: HPC-based workflows configured and deployed on JURECA*

A pipeline for extraction of structural connectivity (SC) has been developed. It can be applied to raw DW-MRI data and used to calculate the whole-brain tractography and extract SC including the matrices of streamline counts and average streamline path lengths for a given brain atlas in volumetric format and standard MNI space. The pipeline is utilising the functions from the well-known (free, open-source) software packages such as Freesurfer, FSL, MRtrix, and ANTs and has a flexible modular architecture. It can be executed on a personal computer and has also been optimised for parallel processing on high-performance computational clusters (JURECA in FZJ) for efficient usage of computation nodes. The DataLad software has been extended and validated for use on HPC systems. All necessary software components are released as free and open-source software with DataLad 0.12.3, available for download from <http://datalad.org> with extensive user documentation provided at <http://handbook.datalad.org>. It enables reproducible execution of containerised analysis pipelines which encapsulate specialised neuroimaging data analysis tools, like Freesurfer, and facilitates the comparative inspection of their output. Moreover, specific enhancements have been implemented to improve scalability for large-scale neuroimaging datasets.

Table 11: Output 4 Links

Component	Link to	URL
C2820	Software Repository	https://gitlab.humanbrainproject.org/felixh/hcp_vbm
C2821	Software Repository	https://gitlab.humanbrainproject.org/felixh/hcp_fs/ (access inquiries through support@ebrains.eu)

6.2 Validation and Impact

6.2.1 *Actual and Potential Use of Output(s)*

Outputs 1, 2 and 3

The new high-resolution maps and connectivity data extracted from cohorts are accessible via the different functions of the interactive atlas viewer. The viewer has been presented at many of the public HBP events, and also in dedicated hands-on sessions and trainings at international meetings which include the INCF 2019 (Warsaw, Poland) and SfN 2019 (Chicago, USA). The reference atlases are potentially interesting to the entire neuroscience community, as described in Section 5.2.1.

The usage numbers of the interactive viewer are continuously growing since its first release. Between October 2019 and February 2020, the number of uniquely identified users per month doubled from around 500 to more than 1,000.

Output 4

As a practical demonstration of the capabilities, a DataLad dataset for the open-access components of the Human Connectome Dataset (15 million files, 80 TB total size) has been published on GitHub for public use at <https://github.com/datalad-datasets/human-connectome-project-openaccess> (actual data access requires credentials that can be obtained from the consortium, instructions are

included). In order to facilitate the registration of generated pre-processed data in the Knowledge Graph, an initial version of a DataLad extension to report captured computational data provenance as JSON-LD compliant linked-data has been published at <https://github.com/datalad/datalad-metalad>.

6.2.2 Publications - not available for KR5.6

7. Key Result KR5.7 “Queries for structural parcellations, genetics, and activity for user-defined regions of interest”

Some data modalities, like local recordings of brain function over time or expression of genes in different parts of the brain, cannot be adequately explored by superimposing their images with a reference. Instead, they often require very specific analytical tools that would make it possible, for example, to understand whether their signals in different regions of the brain differ significantly or not. This KR implements software interfaces that allow a first set of such specific analysis tools to be connected to brain atlases.

7.1 Outputs

7.1.1 Overview of Outputs

7.1.1.1 List of Outputs contributing to this KR

- Output 1: Release of JuGEx (JuBrainGeneExpression) as atlas viewer toolbox and python library. (since 04.2019; C2254, C2265)
- Output 2: Gene Regulatory Network Explorer - Graph visualisation and bayesian network learner tool. Released November 2019 (C2380, 2379)

7.1.1.2 How Outputs relate to each other and the Key Result

Output 1 is a direct implementation of a tool for exploring gene expression by atlas regions. Output 2 is the visualisation tool for the models.

7.1.2 Output 1: Release of JuGEx (JuBrainGeneExpression) as atlas viewer toolbox and python library

The JuGEx algorithm for differential analysis of gene expressions by atlas regions (Method developed by SP2) has been implemented both as a plugin (C2254) of the atlas viewer (C2371), as well as a python library accessible via the Collaboratory, thus allowing for both interactive and advanced programmatic use. For the plugin, the python code was wrapped into a back end web service that is interactively controlled via the plugin, providing an intuitive GUI that allows visual selection of brain areas and auto-completion for selecting candidate genes. Once an experiment is configured, it can be executed directly from the viewer. The resulting probe locations can be visualised in the viewer as landmarks (C2265), and all results can be downloaded as a structured .csv file. On request, the plugin can generate the explicit Python code that corresponds to this computation, and directly import it into a private user Collab for advanced programmatic use in a Jupyter notebook, allowing users to transition seamlessly from exploratory analysis to sustainable, larger scale data

experiments. Here, code modifications and batch experiments can be easily carried out based on the initial experiment in the atlas viewer.

Table 12: Output 1 Links

Component	Link to	URL
C2254	Online Service	https://interactive-viewer.apps.hbp.eu/
	Plugin Repository	https://github.com/FZJ-INM1-BDA/webjugex-iav-plugin
	Technical and User Documentation	https://www.fz-juelich.de/SharedDocs/Downloads/INM/INM-1/DE/jugex.html?nn=2163780
C2265	Online Service	https://interactive-viewer.apps.hbp.eu/

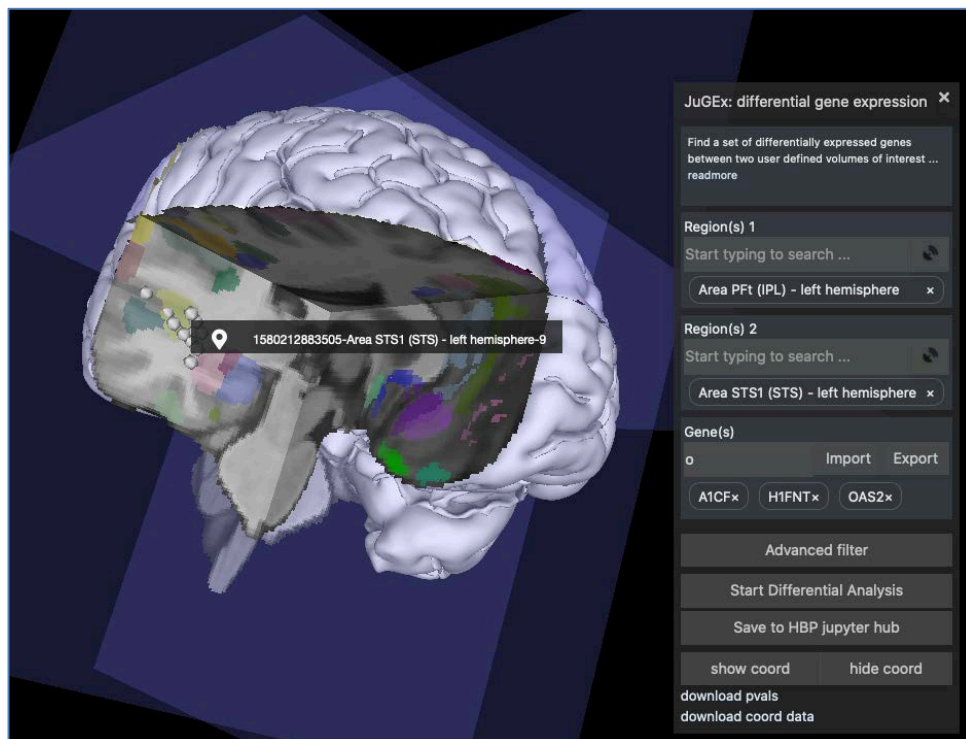


Figure 9: JuGEx tool as a plugin in the interactive atlas viewer

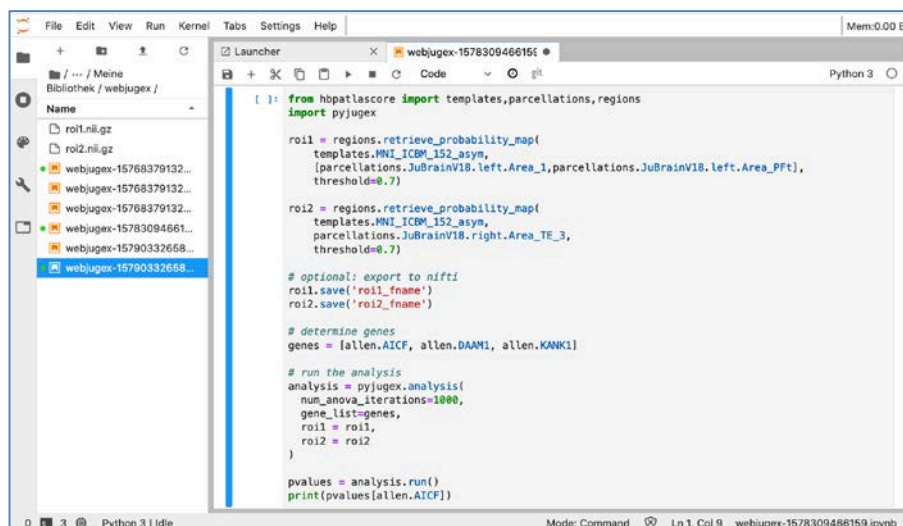


Figure 10: Python code for JuGEx, imported into private collaboratory from viewer plugin

7.1.3 *Output 2: Gene Regulatory Network Explorer - Graph visualisation and bayesian network learner tool*

We have updated the Gene Regulatory Network explorer to its final version, available (C2380) as a standalone tool. Since Y1 we have finished implementing visualisation features that help with the interpretability of massive networks, such as adding metadata to the nodes finding communities or exporting subnetworks for easier analysis, and allow for very fast inference. Further, this tool has been used to visualise Gene regulatory network models for the full human genome using brain gene expression data (C2379)

Table 13: Output 2 Links

Component	Link to	URL
C2380	Online Service	https://neurosuites.com/morpho/ml_bayesian_networks
	Technical and User Documentation	https://www.biorxiv.org/content/10.1101/2020.02.04.934174v1 (P2330)
C2379	Data Repository	https://gitlab.com/mmichiels/fges_parallel_production/tree/master/BNs_results_paper
	Technical and User Documentation	https://www.biorxiv.org/content/10.1101/2020.02.05.935007v1 (P2329)

7.2 Validation and Impact

7.2.1 *Actual and Potential Use of Output(s)*

The JuGEx method is of potential interest for most experimental neuroscientists, the total number of which we estimate to 30,000. Current user numbers of these tools is not yet known, as the release was late in SGA2. However, the interactive atlas viewer, which hosts the plugin implementation, currently counts more than 1,000 unique visitors per month.

The Gene Regulatory Network Explorer (Output 2) is used internally by the Computational Intelligence Group at UPM to validate algorithms for learning gene regulatory networks, by comparing structure in examples with known real structures, and as a quick way for testing various algorithms for learning Bayesian networks on available data. It is also being used as a learning tool for students of the Masters in Computational Biology and Masters in Artificial Intelligence in UPM. The Gene Regulatory Network Explorer (C2379) provides a graphical interface for common algorithms to learn Bayesian networks and visualise graphs, making it well suited for non-experts to rapidly learn and explore rough Maximum Likelihood models. It has shown promise as an educational tool and its use in this area could be scaled.

7.2.2 *Publications*

These publications are showcasing some example uses of Output 2.

- P2330: Michiels *et al.* 2020. NeuroSuites-BNs: An open web framework for massive Bayesian networks focused on neuroscience. <http://dx.doi.org/10.1101/2020.02.04.934174>
Significance: Learning massive interpretable gene regulatory networks of the human brain by merging Bayesian Networks
- P2329: Bernaola *et al.* 2020. Learning massive interpretable gene regulatory networks of the human brain by merging Bayesian Networks <http://dx.doi.org/10.1101/2020.02.05.935007>
Significance: NeuroSuites-BNs: An open web framework for massive Bayesian networks focused on neuroscience for Output 2

8. Key Result KR5. 9 “A user-driven, data-sharing and data-management infrastructure accelerating scientific progress by providing access to large collections of curated, heterogeneous neuroscience data.”

The HBP has developed the tools for a data-sharing and data-management infrastructure. The EBRAINS Knowledge Graph ecosystem provides a complete toolset for metadata management and dissemination. The system supports (semi-)manual and automated workflows and provides multi-datasource integration. The HBP Collaboratory offers protected workspaces for teams to develop reproducible workflows, and share their work publicly through wiki content, documents and notebooks. It provides also a more general solution for collaborative edition of documentation (rich text, wiki or MS Office-type documents) and a user-friendly file repository. An image service offers fast interactive remote access to very large images for atlas viewers and other tools.

8.1 Outputs

8.1.1 *Overview of Outputs*

8.1.1.1 List of Outputs contributing to this KR

- Output 1: EBRAINS Knowledge Graph. Progressively released between March 2019 and March 2020 (C2822, C2619, C1469, C1474, C2239).
- Output 2: Collaboratory v2.0. Progressively released between March 2019 and February 2020 (C358, C372, C373, C374, C540, C532, C2482 and C2617).
- Output 3: Image service. Progressively released between March 2019 and March 2020 (C1461).

8.1.1.2 How Outputs relate to each other and the Key Result

The HBP has developed the essential building blocks of a data-sharing and data-management infrastructure. EBRAINS Knowledge Graph Editor is a key tool used by the curation team to efficiently organise the metadata (KR5.1). Once validated, the data become publicly searchable and receives a DOI. The goal is for the EBRAINS KG to become the “go-to” reference for neuroscience FAIR datasets. The infrastructure also offers optimised access to large image datasets stored centrally in the HBP research infrastructure and referenced by the EBRAINS KG, to avoid having to download GB/TB-sized images for interactive browsing.

The HBP Collaboratory offers protected workspaces for users and teams to develop reproducible workflows, and potentially share their work publicly. It is also a more general solution for collaborative edition of documentation (rich text, wiki or MS Office-type documents) and a user-friendly file repository. The new version adds versioning, a key for user adoption of collaborative tools.

8.1.2 *Output 1: EBRAINS Knowledge Graph*

The EBRAINS KG has been migrated as one of the first services to the new EBRAINS Platform. This metadata management solution consists of multiple components and allows different user groups to find, share, and access high-quality data. The already existing components have been improved and released to the users in regular intervals:

- EBRAINS [KG Search](#)¹⁹ (C2619) has been redesigned for better usability and extensibility. It now supports additional data structures (models, software, etc.) and allows to provide web-browser compatible previews for files for faster user-orientation and to improve the visual appearance. Additionally, actions have been taken for better integration in other search engines (e.g. Google) and the internal search logic has been optimised.
- [EBRAINS KG Editor](#)²⁰ (C1469) has been extended to support additional metadata structures such as software and to simplify the curation workflow by the possibility to invite specific users (typically data providers) to review individual datasets before the release to the public.
- As part of the EBRAINS KG Service (C1474), several other systems (e.g. Knowledge Space, 3D viewers) have been integrated by using the Query API which is supported by the KG Query Builder - a user interface to simplify the programmatic consumption of metadata from the KG. Additionally, several automation processes have been implemented to execute regular probability checks, analyse inconsistent states and to auto-generate derived data (e.g. preview videos of Nifti files for the Search UI).
- New implementation for the Spatial Search (C2239) allows cross-space (subjects) queries, an extended query language with complex shapes support and performance optimisations.

Additionally, a new version of the EBRAINS KG core services has been prepared and released as v3.0-beta: it addresses the requirements of a more fine-grained permission control, improved identifier management, native inference mechanisms, live reflection of the underlying data-structure, simplified client management, etc. This development therefore ensures the continuity and sustainability of the system.

The EBRAINS KG as an ecosystem seems to be unique in its combination of tools, services and applied conventions to support the collaborative and iteratively extensible scientific metadata annotation process. Due to its close integration with semi-automatic processes, such as the curation workflows and the automated ingestion and consolidation of external data sources, the EBRAINS KG builds an infrastructure to collect extensive amounts of metadata while supporting the establishment of quality assurance mechanisms. It also tries to reduce the complexity of metadata consumption by the heterogeneous clients connecting to the EBRAINS KG by the customisation of scopes and reflection on its internal structure.

Table 14: Output 1 Links

Component	Link to	URL
C2822 (incl. related components)	Software Repository	https://github.com/HumanBrainProject/kg-search-ui Access protected (request access at kg@ebrains.eu): https://github.com/HumanBrainProject/kg-editor https://github.com/HumanBrainProject/kg-query-api https://github.com/HumanBrainProject/kg-service https://github.com/HumanBrainProject/kg-core
	Technical and User Documentation	https://kg.ebrains.eu
C2239	Software Repository	https://github.com/epfl-dias/mercator
	Technical and User Documentation	https://epfl-dias.github.io/mercator/

8.1.3 Output 2 - Collaboratory v2.0

The Collaboratory has been re-engineered to replace homemade features with standard open-source solutions. Multiple releases during 2019 and 2020 incorporated user feedback:

¹⁹ <https://kg.ebrains.eu/search>

²⁰ <https://kg-editor.humanbrainproject.eu/>

- GUI and core engine (C540) have been replaced with [Xwiki²¹](#), an extensible wiki platform that brings full-text search and versioning.
- Collaboratory Storage (C373, C532 & C2617) became [Collaboratory Drive²²](#) and is now based on Seafile, a state-of-the-art document repository with much more user friendly interface and file versioning.
- HBP Identity Service (C374) based on MITREid Connect has been replaced by KeyCloak, an OpenID Connect and SAML compatible authentication and authorisation service.
- Document App (C2617) based on Etherpad has been replaced by OnlyOffice, a collaborative document editing solution, compatible with Microsoft Office documents: <https://office.humanbrainproject.eu>
- Jupyter notebook server (C358) has been upgraded to JupyterLab with a simplified access to the Collaboratory drive.

Table 15: Output 2 Links

Component	Link to	URL
C540	Software Repository	https://github.com/xwiki/xwiki-platform/
	Technical Documentation	https://www.xwiki.org/xwiki/bin/view/Documentation/AdminGuide/
	User Documentation	https://www.xwiki.org/xwiki/bin/view/Documentation/UserGuide/
C373, C532 & C2617	Technical and User Documentation	https://download.seafile.com/published/seafile-manual/home.md
C374	Technical Documentation	https://www.keycloak.org/documentation.html
	User Documentation	https://wiki.ebrains.eu/bin/view/Collabs/collaboratory-community-apps
C2617	Software Repository	https://office.humanbrainproject.eu
	Technical and User Documentation	https://helpcenter.onlyoffice.com/server/linux/document/index.aspx
C358	Technical and User Documentation	https://wiki.ebrains.eu/bin/view/Collabs/jupyter-notebooks-collaboratory https://jupyter.readthedocs.io/en/latest/architecture/how_jupyter_ipython_work.html

8.1.4 Output 3 - Image Service

An Image Service was developed and deployed. It supports the ingestion of various image formats used for neuroscience datasets. Images are ingested to generate image chunks at various resolutions to optimise remote access time to very large images. Clients of the service are typically used by brain atlas viewers and neuroscience image analysis tools. The architecture of the service is voluntarily kept simple to support very large amounts of simultaneous accesses, which can arise from public access to public shared datasets. Support for the image service is already built into the NeHuBa image viewer and the EBRAINS KG allowing for the service to be increasingly accessible directly from the KG transparently for end users. The current release of the image service was deployed in February/March 2020.

²¹ <https://wiki.humanbrainproject.eu/>

²² <https://drive.humanbrainproject.eu/>

Table 16: Output 3 Links

Component	Link to	URL
C1461	Online Service	https://img-svc.apps.hbp.eu/
	User Documentation	https://wiki.ebrains.eu/bin/view/Collabs/hbp-image-service-user-guide/

8.2 Validation and Impact

8.2.1 *Actual and Potential Use of Output(s)*

Output 1 is the central registry for quality-assured metadata about scientific artefacts (such as datasets, models and software) produced within and outside of the HBP. As such, it provides a highly visible front end application as part of its publicly and anonymously accessible EBRAINS service next to tooling for the metadata management and APIs.

The different tools of the EBRAINS KG have continued to be in active use by HBP internal and external users during Y2 of SGA2. Between M13-M24, the Search UI has been visited >10,000 times, >7,000 search queries have been executed and almost 2,000 downloads (either full datasets or individual files) have been triggered. EBRAINS KG Editor has been in daily active use by the curation team (>38,000 manipulation/releasing operations) and internal and external systems interact programmatically via the ~650 stored queries.

The components of the EBRAINS KG have been running stable in a productive environment since several months. Nevertheless, since the numbers of users/integrated systems are expected to continue to grow, one key focus of M13-M24 has been to continuously improve the systems' maintainability and scalability.

Output 2: The Collaboratory is used by multiple stakeholders: project managers for their daily needs, scientists use their notebooks to run jobs, analysis and as "live papers", students participate in MOOCs. Currently there are 900 HBP and 3,000 external users registered. While Collaboratory v1.0 has been in production for years, Collaboratory 2.0 first release occurred in March 2019. Its functionality and user base will grow with the community of users progressively migrating their content and usage to the new platform.

Output 3: The image service was deployed in M23/M24. The identified users of the service are the brain image and atlas viewers developed in HBP (LocaliZoom and NeHuBa) and neuroscience image analysis tools such as ilastik. The dataset curation service will also use the service and link available images directly to the appropriate viewing tool in the EBRAINS KG providing end users with an Open With functionality for large neuroscience datasets. The service will also be available to neuroscientists also for datasets not yet curated and for derived image data.

8.2.2 *Publications - not applicable for KR5.9*

9. Key Result KR5.11 HBP-specific workflows developed to fully leverage the current state of the art in computer vision with regards to object detection and segmentation problems. Active and transfer learning methods will be developed to lower the requirements to the necessary amount of training data

Computer vision is developing faster than ever before, constantly pushing the state of the art in object detection and segmentation. These methods are, however, not directly accessible to neuroscientists who might wish to apply them for brain image analysis. The two main barriers are the expertise required and the amount of training data they need. This KR aims to alleviate both.

9.1 Outputs

9.1.1 *Overview of Outputs*

9.1.1.1 List of Outputs contributing to this KR

M13-M24 Outputs list:

- Output 1: Extension of the Neuroglancer viewer to brushstroke annotations (C2234)
- Output 2: Web-based training of ilastik classifiers (C2235)
- Output 3: Integration of ilastik with the HPC infrastructure through UNICORE (C2235)
- Output 4: Transfer learning for brain vasculature segmentation (C2282), delivering pretrained networks and code
- Output 5: PyTorch back end for ilastik, release March 2020 (C2236)

9.1.1.2 How Outputs relate to each other and the Key Result

Output 1 allows users to introduce annotations directly into the Neuroglancer viewer used in NeHuBa. These annotations enable Output 2, where they are used to train an ilastik classifier directly in the browser. Output 3 is stand-alone, but essential for this KR as it allows HBP users to use pre-trained ilastik classifiers for large-scale prediction on the HPC resources. Output 5 provides the means to execute the methods developed in Output 4 for users without extensive deep learning expertise.

9.1.2 *Output 1: Extension of the Neuroglancer viewer to brushstroke annotations (C2234).*

We developed an extension to the popular Neuroglancer viewer which serves as the base viewer for the NeHuBa and VoLuBa HBP products. The annotation tool is implemented as an extension of the native Neuroglancer annotation functionality. The tool creates 3D vertices in the WebGL context used by Neuroglancer, mimicking ilastik brushstroke annotations. The widgets are implemented in TypeScript directly in the Neuroglancer codebase. Brushstrokes can be painted and deleted, history of annotations is available along with bookmarks, arbitrary colours are supported which are also used to display predictions produced in Output 2.

Table 17: Output 1 Links

Component	Link to	URL
C2234	Technical Documentation	https://github.com/ilastik/neuroglancer/tree/web_predictions

9.1.3 *Output 2: Web-based training of ilastik classifiers (C2235)*

Using the annotation tool from Output 1, we reproduced the functionality of the ilastik Pixel Classification workflow with a web-based training interface. A user can select pixel features, introduce labels for various classes, interactively train the Random Forest classifier and, in a close feedback loop, correct the classifier predictions. The training routine is executed remotely on a CSCS Virtual Machine. The trained classifier can be used to launch large-scale batch predictions as described in Output 3.

Table 18: Output 2 Links

Component	Link to	URL
C2235	Technical Documentation	https://github.com/ilastik/ilastik/tree/API_benchmarks

9.1.4 *Output 3: Integration of ilastik with the HPC infrastructure through UNICORE (C2235)*

We developed a Django-based web service for seamless launch of ilastik batch jobs at the supercomputer of CSCS via UNICORE. The desktop application of ilastik now supports upload of the trained ilastik project directly to this new web service running on CSCS infrastructure. Users log in using the OIDC infrastructure of the EBRAINS Platform, which makes the HPC resources readily available for prediction by all HBP/EBRAINS users. The consumed resources are assigned to the ilastik service account; user management is performed by the web service, logging the consumed compute resources.

Table 19: Output 3 Links

Component	Link to	URL
C2235	Technical Documentation	https://github.com/ilastik/cloud_ilastik/

9.1.5 *Output 4: Transfer learning for brain vasculature segmentation (C2282), delivering pre-trained networks and code*

We developed a transfer learning segmentation approach based on deep learning framework. Our solution is an unsupervised learning approach that can work under the scenario wherein the labels are available for a source domain and one would like to transfer that knowledge to generate good results on an unlabelled target domain. An example application scenario would be when we have labelled data for one part of the brain and we want to get results on another part of brain (for which we do not have labels). We have experimentally verified the effectiveness of this approach for the task of brain vasculature segmentation. The target images used in this experiment were captured under a different condition as compared to the source data. The pre-trained networks and the inference code corresponding to brain vascular segmentation will be released to run through the Pytorch back end of ilastik as described in Output 5.

Table 20: Output 4 Links

Component	Link to	URL
C2282	Technical and User Documentation	https://www.ilastik.org/documentation/voxelsegmentation/voxelsegmentation https://github.com/subeeshvasu/hbp-DL-seg-codes/tree/0.1.0/documentation

9.1.6 *Output 5: PyTorch back end for ilastik, new version based on GRPC, release March 2020 (C2236)*

PyTorch back end for ilastik (tiktorch) has been significantly reworked, aiming for more stability. The previous version used ZeroMQ, we have now switched to GRPC to make it faster and more reliable. The new back end conforms to other implementations of model runners and allows ilastik team members to focus development on core functions of delivering interactive training instead of supporting RPC. New session manager has been introduced to allow for reconnections to a running tiktorch server instance. The front end (an ilastik applet) has also been simplified to improve user friendliness. Neural network training is now provided as experimental functionality, accessible in the debug mode. In order to simplify sharing of pretrained networks, we have further extended the previously introduced model configuration template, which now includes a more complete and better structured description of the inference and training process. We have used this configuration to enable running inference for models developed in Output 4.

Table 21: Output 5 Links

Component	Link to	URL
C2236	Technical Documentation	https://github.com/ilastik/ilastik

9.2 Validation and Impact

9.2.1 *Actual and Potential Use of Output(s)*

The intended users of our Outputs are scientists in the HBP and beyond who are already acquiring imaging data or would be interested to exploit the existing imaging data available through the EBRAINS KG. The integration of ilastik with the EBRAINS infrastructure has been tested by selected beta users. The new active learning workflow of ilastik is released for the benefit of all ilastik users, including many thousands outside HBP. The deep learning back end is also targeting all ilastik users who need neural network-based segmentation, such as neuroscientists who want to analyse large electron microscopy volumes of neural tissue, light microscopy volumes of cleared brain, whole slide imaging or high-throughput screening data.

Overall, ilastik was downloaded 60,000 times in 2019. Usage tracking is not possible due to privacy reasons and thus we do not have an exact breakdown of what parts of ilastik are used the most. From user communication at ilastik trainings, image analysis conferences and our support forum, we conclude that the newly introduced ilastik capabilities (active and transfer learning, deep learning back end) are of great interest to the complete ilastik user base (Outputs 4 and 5). The annotation capabilities of Output 1 will also be attractive to other users of Neuroglancer, in the HBP and beyond. Outputs 2 and 3 provide simple and user-friendly access to the HPC computing resources, addressing the challenges of remote large data processing for all EBRAINS users interested in imaging data.

For validation, all functionality has been released in ilastik, accessible with the “—hbp” flag. The web training interface can be explored at web.ilastik.org. Outputs 1, 2, 3 and 5 are at TRL 6, Output 4 is at TRL 4.

9.2.2 Publications

- P2348 (*in validation*); Berg *et al.* 2019; ilastik: interactive machine learning for (bio)image analysis <http://dx.doi.org/10.1038/s41592-019-0582-9>

Significance: Presentation of ilastik software and features as per outputs in KR5.11 in a high impact journal. Publication validates and showcases the outputs in KR5.11

10. Key Result KR5.12 “Comprehensive analysis toolbox to support analysis workflows for electrophysiological activity data from experiments and simulations, including support for HPC and visualisation capabilities.”

The primary output of neuronal network simulations performed in the HBP is dynamic activity that reflects the concerted activity of vast numbers of brain cells on multiple scales. Such simulations are complemented by experimental data recorded using a broad range of measurement techniques. KR5.12 provides software to represent, analyse, and compare such data, in particular to make quantified statements about the scope of validity of the developed models with respect to real brains. The result are generic data structures and analysis methods that enable scientists to perform the exact same analyses protocols for network simulations and biological data. The analysis methods are the basis of model validation at the neuron and network activity levels. KR5.12 provides in particular state-of-the-art measures to exploit the availability of large parallel data streams in order to extract the statistical and correlative structure hidden within large samples of activity data.

10.1 Outputs

10.1.1 Overview of Outputs

10.1.1.1 List of Outputs contributing to this KR

M13-M24 Outputs list:

- Output 1: Electrophysiological Analysis Toolkit – Elephant v0.7, released in March 2020 (C348)
- Output 2: Software library Elephant Visualisation, released in March 2020 (C2122)
- Output 3: Software library Neo v0.8.0, released in September 2019 (C361)

10.1.1.2 How Outputs relate to each other and the Key Result

Output 3 represents the foundation library for Outputs 1 and 2, and is used extensively by the neural activity resource (KR5.13), as well as other tools and services outside this KR (e.g. in SP6). It provides the underlying unifying data model for simulated and experimental activity data, as well as interfaces to a large number of file formats and software back end. Output 2 is a new library to extend Output 1 by visualisation capabilities to facilitate both interactive analysis by users, but also service-based automation of visualising analysis results. Output 1 is also the basis for multiple library components of the HBP validation services.

10.1.2 *Output 1*

Elephant (C348) is an open-source software library that provides analysis methods for activity data from experiment and simulation, in particular parallel spike trains and population activity time series data. It is a community open-source project, with contributors from both inside and outside the HBP. Elephant releases leading up to v0.6.4 (Nov 5, 2019) contain new functionality, in particular novel analysis of rate trajectories in large-scale spike data using Gaussian-process factor analysis or the enhanced SPADE method for revealing temporally extended patterns of activity. These methods complement a heterogeneous array of tools in Elephant for the comprehensive investigation of concerted activity from different angles within a single framework that is not available in other tools. Additional work was invested in starting a major refactor of the documentation system and documentation consistency, as well as several performance improvements on methods to investigate correlative structures in spiking data and parallelisation of the ASSET method (v0.7, March 2020)

Table 22: Output 1 Links

Component	Link to	URL
C348	Software Repository	https://github.com/NeuralEnsemble/elephant
	Technical and User Documentation	https://elephant.readthedocs.io

10.1.3 *Output 2*

Elephant Visualisation (C2122) represents a complementary library for Elephant to provide functionality related to visualisation of data and analysis results. A first version of the library provides functions to visualise Neo data objects, e.g. raster plots for spike trains. In addition, the library provides prototypical visualisations of simple and advanced analysis methods, thereby simplifying the process of browsing analysis outcomes. Further prototype work on the interactive viewer for higher-order correlation results was performed, combining layouting of time-variant graph structures using force-directed graph layouting, combined with hypergraph visualisations.

Table 23: Output 2 Links

Component	Link to	URL
C2122	Software Repository	https://github.com/INM-6/viziphant
	Technical and User Documentation	https://elephant.readthedocs.io/

10.1.4 *Output 3*

Neo (C361) is a package for representing electrophysiology data in Python, together with support for reading a wide range of neurophysiology file formats. It is essential to guarantee interoperability of HBP tools independent of the data sources. It is a community open-source project, with contributors from both inside and outside the HBP. Version 0.8.0 has the following new features:

- “lazy loading” -a feature introduced for raw file back ends that enables the selective loading of partial data during runtime.
- ImageSequence object was added to support the representation of image time series, such as calcium imaging data.
- new functionality to manipulate Neo objects, e.g. to easy allow handling of trial definitions on the data.

Table 24: Output 3 Links

Component	Link to	URL
C361	Software Repository	https://github.com/NeuralEnsemble/python-neo
	Technical and User Documentation	https://neo.readthedocs.io/

10.2 Validation and Impact

10.2.1 Actual and Potential Use of Output(s)

Elephant (Output 1 and 2) is the central component to perform analysis on activity data within the HBP infrastructure. In particular, it is successively validated through HLST-supported interactions, such as the analysis of spatial activity in the full development of the initial version of the slow wave analysis pipeline SWAP (in particular, C2051, C2053), and interactions for analysing correlations in rodent electrophysiology data within SP3. In addition, Elephant constitutes the back end to the “NetworkUnit” validation library, which was employed in publications performing model validation (SP6). Elephant contains contributions from 33 developers, has 71 stars and 59 forks on Github, and is a dependency for 49 other repositories. At the time of writing, the project homepage had received 58 unique visitors in the preceding 14 days. The test suite currently features 515 tests.

Neo (Output 3) is used by HBP and wider neuroscience community, directly and as a dependency of tools such as PyNN and Elephant. The Neo project has 148 stars and 156 forks on Github, and is a dependency for 136 other repositories. At the time of writing, the project homepage had received 161 unique visitors in the preceding 14 days. The test suite was increased from 992 tests (release 0.6.0) to 1,107 tests.

10.2.2 Publications

P2084: Stella *et al.* (2019) 3d-SPADE: Significance evaluation of spatio-temporal patterns of various temporal extents. *Biosystems*, 185:104022. - Output 1

Significance: This publication contributes a major update to the SPADE method, made available by Elephant, to include the search for higher-order patterns with delays. Work on this paper also led to significant improvements in speed of the implemented method.

11. Key Result KR5.13 “The neural activity resource (NAR) as a central mechanism to register, annotate and browse activity data sets within the HBP”

Recordings of neural activity (NA) are typically performed to answer specific hypotheses, but such data may in addition be reused by other scientists, potentially in combination with other data from other labs, to answer different questions that were not considered by the original lab. Such reuse requires that the data are thoroughly annotated with detailed information (“metadata”) about the experimental conditions (the same applies to details of the model and simulation, for simulated data). To ensure that the annotations (KR5.1) are consistent across different data sets, the HBP develops metadata schemas (used to structure metadata within the KG, KR 5.9), together with tools to make it easy to find, understand, and reuse NA data (whether for further analysis, for model building or for model validation). The set of metadata schemas, software and web services related specifically to NA recordings collectively forms the Neural Activity Resource (NAR).

11.1 Outputs

11.1.1 Overview of Outputs

11.1.1.1 List of Outputs contributing to this KR

M13-M24 Outputs list:

- Output 1: NAR metadata tools (C1463):
 - Web app (releases August, September 2019 and March 2020)
 - fairgraph Python library (releases June, November 2019, March 2020)
 - EBRAINS KG schemas (continuous updates in KG)
- Output 2: Neo Viewer (C2124; releases May, August 2019, February 2020).

11.1.1.2 How Outputs relate to each other and the Key Result

Output 1 allows curators, and users with the necessary access permissions, to register, annotate and browse NA datasets, with more fine-grained, in-depth metadata. Output 2 depends on the Neo library (KR 5.12), and allows users to visualise data files that use any of the file formats supported. Output 2 is embedded within the NAR Web app (Output 1).

11.1.2 Output 1- NAR metadata tools

NAR provides tools and documents to support the creation, curation and use of in-depth metadata about NA datasets (from experiments and simulations). The available tools are:

- web app for viewing in-depth metadata (stored in the EBRAINS KG);
- Fairgraph Python library for searching, creating, downloading such metadata for use in modelling, simulation and data analysis;
- metadata schemas in SHACL format, for validating metadata.

The web app is available as a Collaboratory app ("NAR") and as a stand-alone service at <https://nar-browser-app.brainsimulation.eu/app/>. In both cases, an HBP identity account is required for access, and users must agree to the data usage agreement. The app was extended to support viewing detailed metadata for EEG, multi-electrode array, and two-photon calcium imaging experiments.

The NAR Python library was renamed "*fairgraph*", and released on the [Python Package Index²³](#). There have been several releases throughout the year with new features:

- support for provenance tracking of simulation and data analysis workflows
- support for additional experiment types (two-photon calcium imaging, ECoG, etc.)
- support for the EBRAINS KG Query API (C2822), enabling workflows that combine programmatic interaction with the EBRAINS KG via *fairgraph* and manual interaction via the EBRAINS KG Editor (C1469)
- greatly improved test suite, with continuous integration

Metadata schemas for EEG, ECoG, extracellular multi-electrode recordings, and two-photon calcium imaging have been published in the EBRAINS KG. Discussions on adopting these schemas as community standards are under way in the context of the INCF

²³ <https://pypi.org/project/fairgraph/>

(<https://incf.github.io/neuroshapes/>). Development of schemas for fMRI metadata, making use of the existing standards BIDS and NIDM, is under way.

In its current state, the NAR meets or exceeds "beyond state of the art criteria in the SGA2 grant agreement: the tools are in daily use for Tier 3 (in-depth) curation of datasets; all metadata are stored in the EBRAINS KG.

Table 25: Output 1 Links

Component	Link to	URL
C1463	Software Repositories	https://github.com/HumanBrainProject/fairgraph https://github.com/HumanBrainProject/neural-activity-resource https://github.com/INCF/neuroshapes https://pypi.org/project/fairgraph/
	Technical Documentation	https://fairgraph.readthedocs.io/
	User Documentation	https://fairgraph.readthedocs.io/ https://incf.github.io/neuroshapes/

11.1.3 Output 2

Neo Viewer is a web service and accompanying Javascript package for visualising NA data in a web browser. It supports a wide range of file formats (supported by the Neo library); can be easily deployed by data providers or consumers (available as a Docker container), and matches server resources to user demand.

The Neo Viewer has been enhanced with support for visualisation of spike trains, and can now display multiple traces on the same axes. Web-developers embedding the viewer in their pages now have a greater range of options to customize the viewer. The performance of the server component was also greatly improved, making use of the "lazy-loading" features introduced in Neo v0.8. The demonstration server for the [Neo Viewer service](#)²⁴, which also supports the Brain Simulation Platform (BSP) Live Papers, was migrated from a commercial cloud provider to the ICEI/Fenix VM service at CSCS.

Related visualisation tools are available from the Geppetto project (<https://geppetto.readthedocs.io/>) and the Allen Institute Cell Types database (<http://celltypes.brain-map.org/>). The main advantages of Neo Viewer over Geppetto are that it is much lighter and easier to use, and that it supports a very wide range of data formats. The Allen Institute database supports interactive visualisation, but it is customised for the specific protocols used in the Allen Institute and it is not distributed for use with other data outside the Allen Institute database.

Table 26: Output 1 Links

Component	Link to	URL
C2124	Software Repository	https://github.com/NeuralEnsemble/neo-viewer
	Technical Documentation	https://neo-viewer.brainsimulation.eu
	User Documentation	https://neo-viewer.brainsimulation.eu

²⁴ <https://neo-viewer.brainsimulation.eu/>

11.2 Validation and Impact

11.2.1 *Actual and Potential Use of Output(s)*

Output 1: The primary direct users at present are data curators; tutorials on the use of fairgraph were also given to computational and experimental neuroscientists at the 1st Elephant User Workshop and CodeJam #10 workshop. The primary indirect users are modellers using the Model Validation Framework of SP6. Estimation of user numbers is difficult because the tools are typically used as components of other services. There have been 380 direct downloads of fairgraph in the last six months.

The NAR tools are potentially of use to anyone wishing to reuse NA datasets shared through EBRAINS, either directly or through the BSP online use cases. We expect that, once the API of the fairgraph Python library has stabilised, it will be the primary tool for interacting with the EBRAINS KG for data-driven model building, simulation, and validation, or for advanced data analysis.

Output 2: The Neo Viewer is being used as part of some of the BSP “live papers” that accompany traditional journal articles (<https://humanbrainproject.github.io/hbp-bsp-live-papers/>) to enable interactive visualisation of all the electrophysiology recordings. It is also embedded within the NAR web app, allowing interactive previews of individual time series recordings alongside the associated metadata. We are not currently aware of any direct users outside HBP (i.e. people embedding the Javascript component in their own web sites or running their own servers). However, it could potentially be used by any neuroscientist wishing to share electrophysiology data, whether from experiment or simulation, and by anyone wishing to inspect such shared data, and could be adopted by other data sharing initiatives outside the HBP. Work is underway to improve performance with large datasets.

11.2.2 *Publications - not applicable for KR5.13*

12. Key Result KR5.14 “HBP and non-HBP affiliated researchers using interoperable platforms”

Here, we report on the recent achievements in:

- 1) Building a community-based, data-driven encyclopaedia for neuroscience that provides an interface between current brain research concepts and the data, models, and literature
- 2) Delivering user support to users of all HBP Platforms through the High-Level Support Team
- 3) Building the links to The Virtual Brain, the platform allowing for personalised brain simulations by linking computational neuroscience with individual brain imaging data

12.1 Outputs

12.1.1 *Overview of Outputs*

12.1.1.1 List of Outputs contributing to this KR

- Output 1: KnowledgeSpace (KS; C2186, latest release 02.2020)
- Output 2: High-level support team (HLST; C3085, C3084; HBP vouchers - C3059, C3089, C3091, C3104).

- Output 3: The Virtual Brain: (BSP-TVb, released 07.2019, C3046; TVb-NIP, C3047 latest release 03.2019; TVb-HPAC, C3048, latest release November 2019; TVb EduPack, C3049 and 3061, latest update February 2019; HPAC-TVb Pipeline, C3048 released November 2019)

12.1.1.2 How Outputs relate to each other and the Key Result

Output 1 integrates released data and models from EBRAINS KG (KR5.9) with brain research concepts and data, models, and literature from other neuroscience data/model repositories.

Outputs 2 deals with support issues related to all aspects of HBP Platforms from basic inquiries regarding Platform access to more complex solutions that require modification of existing workflows or require considerable contribution of HLST teams, e.g. external dataset curation, or development/integration of new models (HBP vouchers).

Output 3 bridges HBP Platforms (BSP, NIP, HPAC) through workflows developed in The Virtual Brain (KRc8.1-3; for more details, see SGA2 CDP8 Deliverable D5.10.1 (D80.1 D123)).

12.1.2 Output 1 - KnowledgeSpace (KS)

Latest release of [KS²⁵](#) (February 2020) was achieved through the ingestion of released data and models from EBRAINS KG and their integration with brain research concepts and data, models, and literature from other neuroscience data/model repositories. The back end was upgraded to a graph-based data system with Elastic search used for improved text search and graph for semantic search. All code was refactored to provide API endpoints to unify all queries from the front end that was handled by direct Elastic search queries from the UI, and an API endpoint is available for other platforms to fetch descriptions, data, and models from KS.

Table 27: Output 1 Links

Component	Link to	URL
C2186	Data Repository	https://knowledge-space.org
	Technical Documentation	https://github.com/OpenKnowledgeSpace/KnowledgeSpace
	User Documentation	https://knowledge-space.org/documentation

12.1.3 Output 2 - High Level Support Team (HLST)

HLST is HBP's and EBRAINS single entry point for research infrastructure user support. Anyone can submit a request by means of email to support@humanbrainproject.eu or support@ebrains.eu. Further, requests can be submitted through the HBP or EBRAINS web pages (Table 28). All of the above arrive at the front line ticketing system where HLST staff route and address support tickets. Apart from general type of support HLST provides, specialised support in various areas have been covered in SGA2: Data Curation support, Functional Data Analysis support, BSP support, MIP support, Neuromorphic Computing support, Neurorobotics support and Simulation and Data Analytics Workflow support

Further, five external Vouchers were anchored to HLST teams in order to implement these projects on the HBP infrastructure. The projects incorporated are HIPPOCAMPOME (NTNU, NO); MAPS (Uni Roma, IT); HippoPlasticity (LSMUNI, LT); ADNeuronMode (IPMC CNRS, FR); and neurodynlib (Uni Helsinki, FI).

²⁵ https://knowledge-space.org/wiki/thalamus/dataspace/scr_017612_ebrains

Table 28: Output 2 Links

Component	Link to	URL
C3084 C3085	Web site	https://www.humanbrainproject.eu/en/hbp-platforms/support/ https://support.humanbrainproject.eu/ https://ebrains.eu/support
	Technical and User Documentation	https://zammad-user-documentation.readthedocs.io/en/latest/
C3059	Model Repository and User Documentation	https://senselab.med.yale.edu/ModelDB/showmodel?model=244412#tabs-1 https://humanbrainproject.github.io/hbp-bsp-live-papers/index.html
C3091	Model Repository and User Documentation	https://senselab.med.yale.edu/ModelDB/showmodel?model=256388#tabs-1 https://humanbrainproject.github.io/hbp-bsp-live-papers/index.html
C3089	Technical Documentation	https://cxsystem2.readthedocs.io/en/latest/index.html
	User Documentation	https://cxsystem2.readthedocs.io/en/latest/tutorials.html
C3104	Data repository	https://kg.ebrains.eu/search/?q=witter&facet_type[0]=Contributor#Project/1198dafe-12fe-469b-9263-17dd0b46aa3e https://humanbrainproject.github.io/hbp-bsp-live-papers/index.html

12.1.4 Output 3

TVB has been integrated into various HBP Platforms (NIP, BSP, HPAC). EduCase has been included in the TVB GitHub repository linked in EBRAINS, as well as in the HBP Collab. Further, TVB brain tumour data were published on EBRAINS KG. Our effort has treasured the efforts made by the neuroinformatics community towards standards for data sharing, resulting in derived data being included in the BIDS derivatives. (For more details, see SGA2 CDP8 Deliverable D5.10.1 (D80.1 D123)).

Table 29: Output 3 Links

Component	Link to	URL
C3049, C3061	Data Repository	https://kg.ebrains.eu/search/instances/Dataset/a696ccc7-e742-4301-8b43-d6814f3e5a44
	Technical Documentation	https://github.com/the-virtual-brain/tvb-educase-braintumor
	User Documentation	https://training.incf.org/studytrack/virtual-brain-simulation-platform
C3046	Software Repository	https://tvb-nest.apps.hbp.eu/ https://hub.docker.com/r/thevirtualbrain/tvb-nest
C3047	Software Repository	https://hub.docker.com/r/thevirtualbrain/tvb_converter
C3048	Software Repository	https://github.com/the-virtual-brain/tvb-hpc

12.2 Validation and Impact

12.2.1 Actual and Potential Use of Output(s)

Output 1: KS currently contains 1.678.580 units of data and models from 14 different sources. During M13 - M24, KS welcomed over 3.300 users that are mainly from two groups: modellers looking for data to support their models and instructors looking for sample datasets to use for training and

education. The majority of users are from outside Europe (60.4%); the number of European users has increased during this period from 34% to 39.6%. KS could potentially be used by researchers seeking to contextualize their data and large-scale brain projects looking to integrate their data with other large-scale brain projects; as an indicator of FAIR for repositories—if the data and models are discoverable through KS, they meet a certain level of FAIRness. Moreover, potential users could use KS as a template for how to model data and access underlying APIs and data models/schemas for data and terminologies.

Output 2: In total HLST have served around 1600 tickets in SGA2 Y2 for users from all over the world, receiving high (4,7 of 5) score from over 400 customer satisfaction surveys. Requests deal with everything from getting access to platforms to more specific questions about Collaboratory, updating the documentation or tutorials, software deployment (including deployment of MIP), support with EBRAINS KG, JupyterHub, PLUS, EMDESK, Gitlab and data containers.

During SGA2, the HLST service team received 28 requests from 23 different researchers for curation of 75 non-HBP datasets, including data from Voucher projects Hippocampome, MAPS, and ADNeuron. At M24, 65 of these datasets were findable via the EBRAINS Knowledge Graph (25 open, 40 under embargo), while 10 datasets were under curation. Voucher projects (MAPS, HippoPlasticity, ADNeuronMode, neurodynlib) result in new, previously non-existing models which both validate the data provided by the external scientists, and bridge HBP researchers with external community. In addition to research data, MIP HLST provided support to curate clinical datasets and define in collaboration with clinicians a set of harmonized variables (CDEs) for new pathologies.

Other examples of advanced projects with HLST staff are implementation of brain activity data use cases from SP3, support of neuromorphic applications (SpiNNCer, SuperSpike) and implementation of various workflows on HPC resources (TVB image processing, Brayns visualisation software, ilastik, and [learning-to-learn](#)²⁶ (L2L) workflow using Fenix/ICEI resources at ETHZ-CSCS).

Further, the HLST teams are involved in support of users (external) in writing voucher and partnering project applications, applications for resource allocations at HPC centers (both internal and external, e.g. implementation support for GLIF models (Mihalas, Allen Institute), very-large-scale models of the primate brain (Doya, OIST / Igarashi, RIKEN), Astrocyte models (Linne, Tampere, Finland)), as well as partnering and voucher projects (CerebNEST (Pedrocchi, Milano), NEST-Desktop (Rotter, Freiburg)).

Finally, HLST worked on increase of user base of all HBP platforms by presenting their work and services through various events: Elephant User Workshop (Paris, 11.2019), NRP user workshop (Munich, 01.2020), SpiNNaker Workshop, the HBP Student Conference, HBP brain days in Heidelberg, NEST 2019 conference, CodeJAM#10 and HBP summit in Athens.

Output 3: TVB has been downloaded over 7000 during SGA2 (mostly external users) and is currently used in a clinical trial in France to guide surgical planning in Epilepsy. More specific TVB-HBP pipelines are mostly used by HBP researchers (measured by container downloads): TVB-NEST - 129 downloads; TVB-Pipeline: 25 downloads; Fast_TV- 19 downloads. The brain tumor data derivatives showcased at EBRAINS, provide a clear example in a completely accessible and reproducible clinical case. This clinical case study was advertised in TVB nodes and conferences, as well as in the dedicated HBP meetings and open days, raising interest and attracting potential users.

Potential use of TVB outputs: biomarker development, surgical planning (e.g. tumour, epilepsy), therapy development, diagnosis, simulation of person-specific brain models to reveal inter-individual differences in brain dynamics, e.g., related to pathologies or behavioural features.

²⁶<https://wiki.humanbrainproject.eu/bin/view/Collabs/I2I-icei-demonstrator/L2L%20ICEI%20demonstrator%20documentation/>

12.2.2 Publications

Output 2

- P2296: Stimberg *et al.* 2020: Brian2GeNN: accelerating spiking neural network simulations with graphics hardware <http://dx.doi.org/10.1038/s41598-019-54957-7>

Significance: This research exploits the HPC infrastructure provided through the HPAC Platform and bridges between the artificial and biological foundations of learning for applications in research, education and industry.

- P1866: Solinas *et al.* 2019: A kinetic model for Brain-Derived Neurotrophic Factor mediated spike timing-dependent LTP <http://dx.doi.org/10.1371/journal.pcbi.1006975>

Significance: The model explains the magnitude and time course of both transient forms of LTP (t-LTP) and allows predicting t-LTP properties that result from an altered BDNF turnover.

- P2112: Pousinha *et al.* 2019: Amyloid Precursor Protein C-Terminal Domain Alters CA1 Neuron Firing, Modifying Hippocampus Oscillations and Impairing Spatial Memory Encoding. <http://dx.doi.org/10.1016/j.celrep.2019.08.103>

Significance: The model supports the experimental findings suggesting that AICD pathological levels, observed in AD mouse models and in human patients, might contribute to progressive neuron homeostatic failure, driving the shift from normal aging to AD.

Output 3:

Both publications validate components relevant for TVB by using software products developed during SGA2. P2145 further explores the potential of the “virtual surgery” performed by intervening on the structural connectivity pre-surgery

- P1826: Stefanovski *et al.* 2019, Linking molecular pathways and large-scale computational modeling to assess candidate disease mechanisms and pharmacodynamics in Alzheimer's disease <https://doi.org/10.3389/fncom.2019.00054>
- P2145: Aerts *et al.* 2019. Modeling brain dynamics after tumor resection using The Virtual Brain. <http://dx.doi.org/10.1016/j.neuroimage.2020.116738>

13. Conclusion and Outlook

In October 2019, the Neuroinformatics Platform initiated the migration of its tools and services from the project centric HBP website to the new EBRAINS web portal. The launching of the EBRAINS web portal represented the most significant change in the work plan during the reporting period. The new portal, developed in close collaboration with the HBP Communications team, targets the broader community of brain researchers and is populated with content from all HBP Platforms. At the end of SGA2, it brings to light a range of publically available services developed by the Neuroinformatics Platform. It shows how to find data and models (*Find data*, the EBRAINS Knowledge Graph Search UI), how to share and publish data (*Share data*, request curation aimed at making data and models FAIR), and how to navigate the brain with use of the HBP *Brain atlases*. The next step will be to migrate remaining Neuroinformatics Platform tools and services to the EBRAINS web portal, including the tools and workflows for registration of data to atlases, image segmentation and analysis of data in an atlas context, and analysis of activity data. This will be a priority in the beginning of the next Project phase, SGA3.

The HBP Knowledge Graph, now referred to as the EBRAINS Knowledge Graph, was introduced in the spring of 2018 and has been redesigned and technologically improved during SGA2. The Knowledge Graph provides a complete toolset for metadata management and dissemination with datasets stored in the first instance of FENIX. All data and models in the Knowledge Graph are curated. The curation team has developed efficient workflows to facilitate the process of assigning general metadata to the datasets (Tier 1 curation), information about location in the brain (Tier 2 curation), and in several instances in-depth more specialised metadata (Tier 3 curation). Half way into SGA2, the

Knowledge Graph contained over 160 datasets from 40 projects (e.g. clusters of outputs from HBP research), contributed by more than 200 researchers. At the end of SGA2, it had grown to include more than 700 datasets from 80 projects, contributed by more than 700 researchers. With a large number of datasets in the curation pipeline at the end of SGA2, a further substantial increase can be projected for the early phase of SGA3.

Based on the lessons learnt while establishing the Knowledge Graph service, the data and models curation team has developed suitable metadata schemas, employed available ontologies, and contributed to improving and expanding the relevant ontologies. This effort, critical for making data findable, interpretable, and re-usable, has been performed in collaboration with the US Neuroscience Information Framework and the International Neuroinformatics Coordinating Facility, thereby facilitating alignment across international brain projects. The uptake of the Knowledge Graph in the research community is already quite significant, reflected in the large number of searches performed through the Search UI. In the next Project phase, emphasis will be on establishing the EBRAINS Knowledge Graph as the leading European solution for making brain research data FAIR, with uptake across all international brain initiatives. In this regard, a route for EBRAINS data publishing in parallel with release of journal articles has been established during SGA2 and will be announced and brought into action early in SGA3.

Closely associated with the Knowledge Graph service, the Collaboratory continues to function as a unifying service in support of all HBP Platforms. With the release of Collaboratory 2, users have access to several new and improved functionalities such as fully public collabs, wiki functionality, interactive programming notebooks, improved user interface, and powerful user and permission management. The majority of the services presented through the EBRAINS web portal in SGA3 will use Collaboratory 2.

The HBP Brain atlases has received major updates and continues to be one of major services delivered by the HBP. The tip of the iceberg of the brain atlas capabilities of HBP has been available on the EBRAINS web portal since October 2019, with the Interactive Atlas Viewer being the central tool for displaying and navigating the atlases. This online tool was improved with workflows for querying updated multimodal, region-specific brain atlas data, maximised views for any axial or oblique section for better orientation and appreciation of the available data, better integration with the Knowledge Graph, and user instructions for navigating the atlas.

The Neuroinformatics Platform has continued to deliver improved analytical tools and workflows for structural and functional data. The widely used ilastik tool for machine learning based segmentation, classification and tracking of 2D and 3D images, has been further modified and adapted to the needs of brain research. Using HBP compute resources, ilastik classifiers can now be trained directly in the web browser. Batch processing on supercomputer resources is possible seamlessly out of the ilastik desktop application. New active learning workflow and the deep learning back end allows analysis of large electron microscopy volumes of neural tissue, light microscopy volumes of cleared brain, whole slide imaging, or high-throughput screening data. A new workflow combining multiple Neuroinformatics Platform tools now allows users to register data to an atlas, perform image segmentation, and sort the extracted features into atlas defined brain regions at a level of granularity determined by the user. Analysis tools for dynamic functional data from experiments and simulations has been further updated with a focus on facilitating user accessibility and the adding of new functionality to cover needs of more users. Extensive testing of the Neuroinformatics Platform analytical workflows has taken place in SGA2. The next step will be to migrate the tools and workflows, with accompanying tutorials and user manuals, to the EBRAINS web portal in order to facilitate the further uptake of the data sharing, atlases, and other related services.

Finally, the Neuroinformatics Platform has been responsible for launching the High-Level Support Team (HLST), which was migrated to the EBRAINS web portal this year. At the end of SGA2, this service has successfully delivered support for users of all HBP Platforms and ingested data and models from several non-HBP projects into the HBP/EBRAINS services. HLST will be ramped up further in the next project phase to ensure that user errors, bugs, or technical issues are resolved, to receive feedback guiding further development of the services, and to provide advanced support enabling new future capabilities to be added to the EBRAINS services.

14. Index of Components

Component ID	Component name	More information
C1461	Image Service	table 16
C1463	Neural Activity Resource metadata tools (web app, fairgraph python library, metadata schemas for EEG)	table 2, table 25
C1469	EBRAINS KG editor	table 14
C1474	EBRAINS KG service	table 14
C2122	Visiphant (Elephant visualization library)	table 23
C2124	Neo viewer for neural activity data	table 26
C2167	Spatial coordinate system (Waxholm Space and stereotaxic references) defined in Allen Mouse Brain Atlas template	table 7
C2169	Waxholm Space rat brain reference atlas	table 6
C2186	Knowledgespace	table 27
C2234	Extension of the Neuroglancer viewer to brushstroke annotations	table 17
C2235	ilastik software	table 18
C2236	ilastic voxel segmentation and ilastic back end	table 20, table 21
C2239	Spatial search (Mercator)	table 14
C2240, C2241, C2242	SANDS metadata schema,	table 1
C2244	Best practice recommendations for assignment of anatomical location to experimental rodent brain data	table 7
C2254	JuGEx plug in for atlas viewer	table 12
C2260	Connectivity matrix export tool of the human atlas viewer	table 8
C2264	Jubrain human atlas	table 6, table 9
C2265	Support for visual landmarks in 3D atlas viewer	table 12
C2267	Voluba (Service for computing a map of cortical depth from a cortical segmentation)	table 7
C2269	Live warping for interactive affine volume registration (Voluba)	table 7
C2272	Cytoarchitectonic mappings in Big Brain visual areas	table 7, table 9
C2275	Quint automated workflow for analysis of serial microscopic image data registered in the Knowledge Graph	table 4
C2277	QuickNII software	table 4, table 7
C2282	ilastik cloud (integration with HPC)	table 19
C2284	openMINDS metadata schema and curation request form	table 1
C2285	Ontology for data/model discoverability annotation	table 3
C2319	JuBrain probabilistic cytoarchitectonic maps	table 6
C2370	Display of area-specific textual descriptions and images in the 3D atlas viewer	table 8
C2371	Web-based multi-resolution three-planar viewer for large image volumes	table 8
C2372	Iterative workflow loop for landmark adjustment (Voluba)	table 7
C2379	Gene regulatory network models	table 13
C2380	Gene Regulatory Network explorer software	table 13
C2522	Curation workflow: Support for data upload and download	table 2
C2523	Curation workflow: Support for data transfer	table 2
C2617	Only Office for Collaboratory	table 15

C2619	EBRAINS KG search	table 14
C2802	Adoption of 3D atlas viewer to a unified viewer architecture	table 8
C2820	Volume-based pipelines for structural connectivity	table 10, table 11
C2821	Surface-based processing pipelines for structural connectivity	table 10, table 11
C2822	EBRAINS KG	table 2, table 14
C3049	TVB EduPack (INCF)	table 29
C3046	TVB-NEST pipeline	table 29
C3047	TVB converter	table 29
C3048	TVB-HPC pipeline	table 29
C3059, C3091, C3089, C3104	HBP voucher projects (data and models from external partners integrated by HLST)	table 28
C3061	TVB epilepsy dataset	table 29
C3074	Nutil (Neuroscience image processing and analysis utilities)	table 4
C3084, C3085	High Level Support Team	table 28
C348	Electrophysiological Analysis Toolkit - Elephant	table 22
C358	JupyterLab (Jupyter notebooks for Collaboratory)	table 15
C361	Neo software package	table 24
C373, C532, C2617	Collaboratory Drive	table 15
C374	HBP identity service	table 15
C540	Xwiki (collaboratory GUI and core)	table 15
C935	Connectomic composition predictor (rodent)	table 5