Release of updated atlas curation supporting material (D5.2.1 – SGA2)

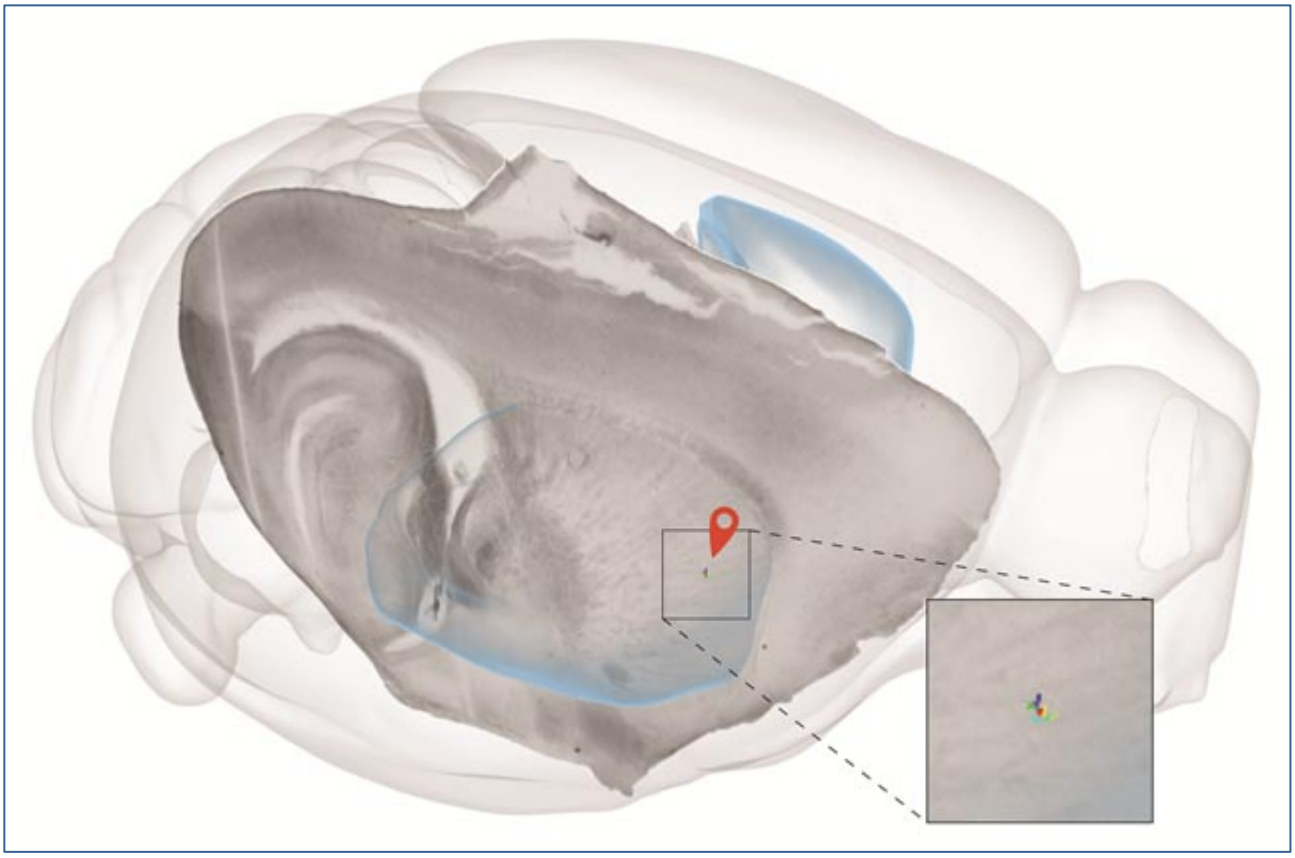

Figure 1: The anatomical location of a 3D reconstructed neuron in the mouse striatum

To determine the anatomical location of a 3D reconstructed neuron in the mouse striatum, a histological image of a sagittal mouse brain section in which the cell body of a 3D reconstructed neuron is visible, was registered to the Allen Common Coordinate framework using the HBP QuickNII tool. The spatial coordinates for the anatomical position of the neuron were extracted and registered as location metadata. The position of the image in atlas space is here visualised using the Scalable Brain Atlas Composer. For further details about the data set and best practices for determining and documenting anatomical location, see Bjerke *et al*. (Front Neuroanat, 12:82, 2018).

Project Number:	785907	Project Title:	Human Brain Project SGA2

Document Title:	Release of updated atlas curation supporting material
Document Filename:	D5.2.1 (D29.1 D73) SGA2 M12 ACCEPTED 190722
Deliverable Number:	SGA2 D5.2.1 (D29.1, D73)
Deliverable Type:	Report
Work Package(s):	WP5.2 (lead), WP5.1 and WP5.4
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA2 M12/ 31 MAR 2019
Actual Delivery Date:	SGA2 M12 / 18 MAR 2019; ACCEPTED 22 Jul 2019

Authors:	Trygve LEERGAARD, Camilla BLIXHAVN, and Maja PUCHADES, UIO (P81)
Compiling Editors:	Trygve LEERGAARD, Jan BJAALIE, and Sofia ANDERHOLM STRAND, UIO (P81)
Contributors:	Krister ANDERSSON, UIO (P81), Task Leader T5.1.1 Camilla BLIXHAVN and Ulrike SCHLEGEL, UIO (P81), Rodent curation team Anna HILVERLING, Lyuba ZEHL, Sara ZAFARNIA, and Stefan KÖHNEN, JUELICH (P20), Human curation team Maja PUCHADES, Ingvild BJERKE, Martin Øvsthus, Sharon YATES, and Heidi KLEVEN UIO (P81), Brain atlasing and analysis; curation support Milica MARKOVICH, EPFL (P1), Curation team liaison Mathew Abrams, KI (P37), Ontology development Elodie LEGOUEE, CNRS (P10), Neural activity curation team Andrew DAVISON, CNRS (P10), Task Leader T5.7.2 Timo DICKSCHEID, JUELICH (P20), WP Leader WP5.3, WP5.4 Trygve LEERGAARD, UIO (P81), Task Leader T5.2.1, T5.2.2 Jan BJAALIE, UIO (P81), WP leader WP5.1, WP5.2
SciTechCoord Review:	Martin TELEFONT, EPFL (P1)
Editorial Review:	Annemieke MICHELS, EPFL (P1)

Description in GA	Users can access tutorials and SOPs from HBP workbench or request support from UiO atlas curation team (T5.2.2)
Abstract:	To facilitate ingestion of experimental brain data in the HBP infrastructure, the HBP curation team has created a HBP Data Integration Guide, containing updated information packages for researchers wishing to share their data via HBP. The Data Integration guide includes workflow descriptions with detailed instructions and tutorials for all steps in the process. The documents are released from a public HBP collaboratory, as support for data providers throughout the entire process of project planning, data collection, and data release.
Keywords:	Data sharing, Data integration, Brain atlas, Location metadata, Curation
Target Users/Readers:	Neuroscientists wanting to share experimental data via the HBP infrastructure

Table of Contents

Table of Figures

1. HBP curation support: Data Integration Guide

A key objective for the HBP data curation team is to organise and validate experimental brain data and metadata before the data are made publicly discoverable and accessible through the HBP Knowledge Graph (KG) (https://www.humanbrainproject.eu/en/explore-the-brain/search/) in accordance with the F.A.I.R. principles (Findable, Accessible, Interoperable and Re-usable), with a license defining the conditions for usage and Data Cite DOIs (Digital Object Identifiers) for each data set.

A particular focus in SGA2 has been to improve the curation process established in SGA1, aiming at increased throughput. This has been achieved by establishing a revised data / model curation workflow (reported in SGA2 Deliverable D5.1.1 "Completion of SGA1 data/model curation into NIP metadata database"), and by preparing updated documentation and supporting material for data providers who will share their data through the HBP infrastructure (present Deliverable). The present Deliverable describes the release of the updated atlas curation supporting material, together with the complete data curation support material which is bundled in the present release of the Data Integration Guide.

The sharing of data in the HBP infrastructure is based on metadata handled in a three-tiered process: 1) basic metadata, 2) location metadata, and 3) method-specific in-depth metadata. The online curation supporting material reflects this organisation, see below. A major improvement in the curation support is the establishment of a single contact point with a central ticketing system for both HBP and non-HBP users requesting data curation.

The updated supporting material is now released in the Collab "Preparing and integrating data for the Human Brain Project". This open Collab is available for users with an HBP account: (https://collab.humanbrainproject.eu/#/collab/7574/nav/281006)

The Collab front page shows infographics of all topics covered by the Collab.

The detailed supporting material found in the Collab includes:

1) The HBP Data Integration Guide
2) Tier 1: Basic metadata, making data findable in the HBP Knowledge Graph
3) Tier 2: Location metadata, increasing data visibility in HBP atlas viewers and exploitation through HBP analytic workflows
4) Tier 3: Method specific, deep metadata, optimising data accessibility and reusability
5) Overview of ontologies used in HBP curation

1.1 HBP Data Integration Guide

The HBP Data Integration Guide is available as a short version in the Collab. The full version can be downloaded from the Collab Storage as a PDF file.

The HBP curation team is available to provide support to data providers with the organisation and uploading of data and metadata, per request to: curation-support@humanbrainproject.eu.

1.2 Tier 1: Basic metadata making data findable in the HBP Knowledge Graph

The Collab guides the data providers through the mandatory steps of the Tier 1 data curation process. A series of guideline questions are highlighted to elucidate the process and assist users in delivering the requested metadata, before they upload their data to HBP storage and get feedback from the curators on the organization of the data and the completeness of the metadata. At this level important formalities regarding ethics and access to HBP data storage facilities for uploading of data

(also for researchers outside the HBP consortium) are explained. A list of metadata keys with definitions is provided.

1.3 Tier 2: Atlas integration by assignment of anatomical location

The Collab provides guidance on the assignment of anatomical location to experimental data. This step is of key importance to enable spatial data queries, visualisation of data in HBP viewers, and for utilising the different analytic workflows established by HBP. For the present release, information is provided to explain the:

- Anatomical references atlas resources provided for the human, rat, and mouse brain by HBP
- Different anatomical parcellation schemata supported by HBP
- Tutorials presenting the different workflows available for spatial registration of data to the reference atlases used by HBP, including use of semantic links, spatial registration of images to reference atlases, and use of spatial coordinates defining points of interest.

The different tools used for registration of 2D and 3D images in the atlas are introduced, together with links to tool-specific Collabs:

https://collab.humanbrainproject.eu/#/collab/5484/nav/42798

https://collab.humanbrainproject.eu/#/collab/5401/nav/42067

1.4 Tier 3: Optimising data accessibility and reusability

The Collab shows the present offer for adding in-depth metadata beyond what is captured by the basic metadata schemas used in Tiers 1 and 2. This step adds detailed metadata describing a range of methodological details, facilitating the re-use of the data shared through the Knowledge Graph. The nature of these metadata varies considerable with the methodology used. Through SGA1 and SGA2 the Neural Activity Data resource of HBP has developed metadata schemas of in-depth metadata for neural activity data. The collection of tools and documents can be accessed from the "Neural Activity Resource" Collab:

https://collab.humanbrainproject.eu/#/collab/1635/nav/14491

More detailed method-related metadata will also be captured for the remaining data categories registered to the HBP infrastructure (work in progress).

1.5 Ontologies

To enable precise and efficient queries of the Knowledge Graph, HBP uses and contributed to establishing a set of internationally accepted **controlled vocabularies** and **ontologies** that have been approved as part of the basic metadata schema used in Tier 1. These controlled vocabularies define the name, definition, type, and other properties of a scientific term. In addition, ontologies provide formal definitions for relationships between terms. Beyond providing a common language for metadata keys and values, another benefit of using ontologies in KG is that it can use these relationships to improve query results. The online support documents provide an overview of the ontologies and terminologies used for HBP metadata. This information is available in the collab: https://collab.humanbrainproject.eu/#/collab/7574/nav/119329.

1.6 Ongoing work towards next release

The current release of Curation support material has incorporated feedback and experiences hitherto accumulated through the SGA1 and SGA2 phases of the project. The resources will be regularly updated as workflows and tools are further developed. For the next release of the Curation support material, web design will be optimised with more use of infographics, figures and diagrams. The HBP curation team will work together with the High Level Support Team (HLST) to further improve the data registration and curation process, aiming to make it as automated and time efficient as possible.