Data integration through brain atlas ing: Human Brain Project tools and strategies


Published in:
European psychiatry

DOI:
10.1016/j.eurpsy.2018.02.004

Citation for published version (APA):
Review

Data integration through brain atlasing: Human Brain Project tools and strategies

Ingvild E. Bjerke, Martin Øvsthus, Eszter A. Papp, Sharon C. Yates, Ludovico Silvestri, Julien Fiorilli, Cyriel M.A. Pennartz, Francesco S. Pavone, Maja A. Puchades, Trygve B. Leergaard, Jan G. Bjaalie

*Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, Norway
European Laboratory for Non-linear Spectroscopy, Sesto Fiorentino, Italy
*Cognitive and Systems Neuroscience Group, SILS Center for Neuroscience, University of Amsterdam, The Netherlands

ABSTRACT

The Human Brain Project (HBP), an EU Flagship Initiative, is currently building an infrastructure that will allow integration of large amounts of heterogeneous neuroscience data. The ultimate goal of the project is to develop a unified multi-level understanding of the brain and its diseases, and beyond this to emulate the computational capabilities of the brain. Reference atlases of the brain are one of the key components in this infrastructure. Based on a new generation of three-dimensional (3D) reference atlases, new solutions for analyzing and integrating brain data are being developed. HBP will build services for spatial query and analysis of brain data comparable to current online services for geospatial data. The services will provide interactive access to a wide range of data types that have information about anatomical location tied to them. The 3D volumetric nature of the brain, however, introduces a new level of complexity that requires a range of tools for making use of and interacting with the atlases. With such new tools, neuroscience research groups will be able to connect their data to atlas space, share their data through online data systems, and search and find other relevant data through the same systems. This new approach partly replaces earlier attempts to organize research data based only on a set of semantic terminologies describing the brain and its subdivisions.

1. Introduction

Knowledge about brain functions is needed not only to help alleviate mental illness, neurological diseases, memory loss and aging, but also to learn more about the normal organization of the brain and the massive information arrays that are handled by natural neural networks. Converting data from the brain to knowledge about the healthy and diseased brain requires comprehensive data integration, since the brain can be fully characterized and understood only by combining a wide range of data acquired at different spatial and temporal scales and with manifold techniques. As in many other fields of science, however, the capacity to generate massive amounts of new data has by far surpassed the ability to integrate and make sense of the data collected. The challenge of combining and unifying disparate multimodal and multilevel data collected from the brain into meaningful and valuable information, has triggered developments in the field of neuroinformatics, which combines neuroscience with information technology and deals with the creation of data systems that are required to achieve advanced integration of data needed to understand the nervous system [1–3].

The Human Brain Project (HBP), a multinational European brain research initiative, was formed to advance neuroscience and medicine and to create brain-inspired information technology [4]. The project comprises multiple subprojects, including experimental neuroscience research and infrastructure platforms contributing tools and workflows for data analysis, modeling and simulation, robotics, and ethics frameworks. A key concept is the integration of heterogeneous data from human and rodent brain in a common data system, and the use of such data for analysis and computational modeling and simulation of brain function. The project’s Neuroinformatics Platform, in close interaction with other parts of the project, covers the preparatory steps for data integration: making data discoverable, accessible, interpretable, and reusable, as captured in the recently introduced FAIR Guiding Principles (Findability, Accessibility, Interoperability, and Reusability; [5]). A prerequisite for data integration is that the nature and relationships of data parameters are well defined, and possible to compare. As a start, basic features of data to be used in data integration

http://dx.doi.org/10.1016/j.eurpsy.2018.02.004
0924-9338/© 2018 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
efforts must be consistently and reliably described. Metadata is the term used to describe and characterize features of data (“data about data”), and the use of controlled vocabularies and ontologies is key to making metadata consistent [6–9]. A special category of metadata related to the brain is spatial metadata, describing the anatomical location from which the data originates. Location is a critical descriptor for brain data, since the brain is organized into numerous structures with different subparts, cell groups, or layers, and with function closely coupled to location. Location thus provides a natural basis for organizing data across spatial scales, and reference atlases of the brain and neuroinformatics tools have emerged as relevant technologies in support of such efforts [10–12].

In this paper, we illustrate key principles of the HBP workflows used for registration of a range of data categories to reference atlases of the brain. Neuroscientists upload experimental data and associated metadata to the HBP data systems. A curation service delivered by the Neuroinformatics platform organizes the data and helps standardize the metadata. The curation process includes registration of experimental data to common frameworks represented by the reference atlases. We exemplify the workflows using rodent brain data and demonstrate how data from different experimental modalities, registered to reference atlases, can be combined and unified with other data categories. Finally, we discuss the application of these approaches for characterization and analysis of rodent models of neurological diseases.

2. Assigning location to data from the brain: reference atlases for spatial data integration

Assigning anatomical location to data acquired from the brain can be done at the level of structures of the brain, or through the use of coordinate systems. Reference atlases offer a standardized representation of anatomical location and are commonly used tools to assign location to data. They are thus highly relevant tools for spatial data integration.

Traditional atlases of the brain are books or articles containing series of two-dimensional (2D) diagrams of cross-sections through a brain, showing delineations and names of the structures present in each section (see, e.g. [13,14]). The structural delineations are typically derived from interpretations of features observed in histological sections. The limitations of such 2D reference atlases are primarily the missing information between each diagram (due to use of incomplete histological material) and the single plane of orientation of the diagrams. For future data integration efforts, a new generation of three-dimensional (3D) reference atlases is used ([15–19]; see also www.brain-map.org; www.nitrc.org; http://scalablebrainatlas.incf.org). These atlases are based on either 3D reconstruction of images taken from complete series of histological sections, or on tomographic imaging data acquired with, e.g., magnetic resonance imaging (MRI) and diffusion tensor imaging (DTI). The volumetric data in these reference atlases allows the image volumes to be re-sliced and viewed in arbitrary angles with minimal loss of image quality due to the isotropic acquisition of the images. For rodent experimental data, the Human Brain Project uses two volumetric atlases (Fig. 1): the Allen Mouse Brain Atlas [20] and the Waxholm Space rat brain atlas [16,21].

The first version of the Allen Mouse Brain Atlas was reconstructed from multiple series of Nissl stained histological sections [22]. Building on the high level of architectonical details available in the histological material, more than a thousand anatomical regions were delineated. The latest version of the atlas (v3) contains a high resolution population average image volume from serial two-photon tomography, integrating shape and signal intensity over 1675 specimens (mouse.brain-map.org). The Allen

![Image](image-url)

**Fig. 1.** Three-dimensional rodent brain reference atlases used by the HBP. (A) The Allen Mouse Brain Atlas (v3; 13; see also [22]) and (B) the Waxholm Space atlas of the Sprague Dawley rat brain [16,21]. The atlases have been sliced coronally, giving a view of deep brain structures. The volumetric reference atlas delineations have isotropic voxels and can be arbitrarily sliced in any angle of orientation.
Mouse Brain Atlas employs a Cartesian coordinate system, referred to as the Common Coordinate Framework, complemented by a curved coordinate system in the cerebral cortex to enable integration of information from different cortical depths.

The Waxholm Space rat brain atlas provides volumetric anatomical delineations of the rat brain based on high resolution isotropic MRI and DTI of a single animal. The atlas is navigated using the Waxholm Space coordinate system with its origin set to the midpoint of the anterior commissure (Fig. 2). The primary entry point for experimental data into Waxholm Space is through image alignment to the MRI template. Regions of interest are then assigned coordinates, allowing comparison to other datasets that have been mapped to the same atlas. Coordinates can also be transformed to other coordinate systems for access to information available on the basis of location [23].

All datasets, regardless of spatial scale or experimental modality, are associated with an anatomical location in the brain. To bring diverse types of experimental data into the same anatomical space, 2D or 3D images are aligned to a reference atlas (for an overview of alignment approaches, see e.g. [24,25]). Once the anatomical location of the data is established, regions or features of interest are assigned coordinates, allowing for comparison to other mapped datasets and queries for additional information [26]. The HBP Neuroinformatics Platform develops software tools and services that facilitate the process from data acquisition and annotation, through assignment of anatomical location to data, to assembly of and access to spatially indexed information about the brain. The examples below serve to illustrate data integration through the use of rodent brain reference atlases and HBP tools.

3. Examples of spatial data integration

Registration of experimental data to a reference atlas requires informatics tools and well defined procedures. Such tools and procedures are currently being established in the HBP Neuroinformatics Platform. We here provide examples illustrating different aspects of the challenges involved and the opportunities emerging, following registration of data to the common frameworks represented by the reference atlases (Figs. 3–5; for overview of ongoing rodent atlasing developments in HBP, see [27–32]). The challenges involved are primarily related to the ability to efficiently and accurately position large series of images in reference atlas space. The opportunities are manifold and include: (1) the creation of cut planes of the reference atlas that can be shown as overlays on 2D experimental images at any angle of orientation [28]; (2) the assignation of reference atlas coordinates to features observed or extracted using image processing (Figs. 3 and 4); (3) the ability to analyze the distribution of features in atlas space or according to atlas structures (Figs. 3–5); and (4) the ability to search for other available information based on location of interest in atlas space or according to atlas structures (Fig. 5).

The examples shown in Figs. 3 and 4 are centered on whole brain mapping of categories of neurons or other labeling features in mouse brain. Fig. 3 shows a dataset acquired with high-resolution light sheet microscopy following tissue clearing [33]. The resulting 3D dataset was initially globally registered to the Allen Mouse Brain Atlas, before a local registration procedure centered on a region of the cerebral cortex to increase the precision of the positioning. For both steps, affine transformations were used. Registration could be further improved by introducing non-linear transformations. Coupled with automatic cell localization algorithms [34], the approach allows efficient whole-brain mapping of neuronal distributions, yielding quantitative information about neuronal densities, numbers, and spatial distributions, normalized to reference atlas space. Such detailed quantitative data allow comparisons among parts of the brain, and between brains from different subjects, at a high level of granularity, and may in combination with other data be used to make inferences about functional properties of brain regions, to feed computational models of brain systems, or to monitor changes occurring in models of aging and disease.

The two datasets shown in Fig. 4 are series of 2D microscopic images from histological sections processed to visualize different features in the mouse brain. The first represents a sagittally oriented series of sections from a normal brain (Fig. 4A–F), whereas the other represents a near coronal series of sections from a disease model (Fig. 4G–I). Both datasets are registered to the Allen Mouse Brain Atlas using affine transformations. The angle of sectioning of each of the series was precisely defined as part of the registration process, resulting in a good fit to the atlas. By

![Fig. 2. External and internal reference points and coordinate systems in a rat brain reference atlas. (A) Cranial landmarks bregma (br), lambda (la), and the interaural line (IAL) are marked on a rat skull (volumetric rendering from μCT; DigiMorph Library, University of Texas at Austin). The position of the brain within the skull is indicated by the yellow outline. (B) The axes of the Waxholm Space coordinate system (blue lines) are shown relative to the anterior commissure (structure in color). The position of bregma and lambda are indicated above the brain surface (derived from the Waxholm Space atlas of the Sprague Dawley rat brain; [46]).](image-url)
registering the images to a common reference atlas, the anatomical location of features of interest were derived from the atlas overlay images and described either semantically (referring to names of structures), or defined by spatial atlas coordinates. Following image processing applied to the images in a batch process, labeled objects were extracted from image background and quantified [23,31,35,36]. By automatic assignation of an atlas coordinate to each centroid point, the labeled features were co-visualized and quantified in context of atlas regions (Fig. 4). The quantitative distributions observed in this example are well suited for comparing densities and number of identified features in each atlas-defined region of the brain (plaques in the disease model as shown in Fig. 4G–L, and labeled neurons as shown in Fig. 4A–F). The data obtained are suitable for integration with other categories of data available in the same common atlas space, aiming at improved understanding of organizing principles in the brain or changes occurring in disease.

Fig. 5 is an example showing the determination of anatomical position of electrodes (i.e., tetrodes) in the rat brain by registration of histological sections to a reference atlas. In a series of experimental animals, craniotomy was performed and recording tetrodes were lowered into the brain (Fig. 5A) at locations stereotaxically measured from the bregma landmark of the skull (Fig. 2; [37,38]). The tetrodes were targeted at the same area. Following implantation, rats performed a behavioral task taxing their perceptual and memory capacities of the neural system under scrutiny. The locations of the tetrode tips were verified by registration of section images from the brains of the animals to the Waxholm Space rat brain atlas. Tetrode tip locations could be identified in the section images and could thereby be located in atlas space (Fig. 5D). Thereby, the exact anatomical locations for each of the recording sites were determined as defined by anatomical area and atlas coordinates. The results show that the recording sites are positioned in or close to different subregions of the perirhinal cortex (Fig. 5E), which is relevant for interpretation of the electrophysiological results. This example also highlights the opportunity for accessing other types of information available for the areas in which the recording sites were located. Thus, an online database linked to the different atlas regions was searched to extract information about the known anatomical connections of the regions of interest. In our example we consulted the Temporal-lobe.com database, which holds a comprehensive collection of literature-derived data about the neural connections of the rat hippocampal system [39,40]. The query revealed that the two subregions (areas) of the perirhinal cortex, in which the recording sites were located, have different sets of afferent and efferent connections (Fig. 5F). This allows us to associate the electrophysiological characteristics and behavioral correlates of cells with their anatomic locations in two perirhinal areas and the networks in which they are embedded. The example thus illustrates how mapping of data to a common reference space facilitates data interpretation. Many data categories from the brain, including neural connectivity data, can only be interpreted correctly if the spatial location is known. A further approach, which is also pursued in the Human Brain Project because it cannot be realized in humans, is to combine intracellular patch clamp recordings in vivo with single-cell staining and 2-photon targeted recordings from identified cellular phenotypes — enabling the coupling of single-cell morphology, functional electrophysiology and brain atlasing (see, e.g., [41]). Finally, these functional-anatomic data are being used in computational modelling and simulation projects of HBP. In the example shown in Fig. 5 the data serve to specify
models of multisensory integration and memory processing in sensory cortices and the medial temporal lobe system.

4. Conclusion

The Human Brain Project introduces a systematic use of reference atlases as one of several ways to integrate neuroscience research data. A suite of neuroinformatics tools are currently being developed and shared with the research community, aimed at facilitating the multiple steps involved in data integration through the use of the reference atlases. In addition to the efforts here described for the rodent brain, the project uses and extends reference atlases for the human brain, including the BigBrain [42]. Furthermore, advanced tools are developed to support the spatially more complicated task of organizing research data for the human brain (see, e.g., [43]).

Brain atlasing is equally relevant for organizing research data from the healthy and the diseased brain. By registering large amounts of heterogeneous neuroscience data to a common brain reference atlas, and making the images available for inspection and analysis in web-based viewer tools, the Human Brain Project will provide new opportunities for the use of data from normal specimens as benchmarks for interpretation of disease related changes in patients or animal models. One of these opportunities is to pinpoint electrophysiologically characterized neuronal populations to sub-regions with unprecedented anatomic resolution. Secondly, having characterized the location of neuropathology in a disease specimen, the atlas based infrastructure will allow researchers to look up information that is spatially relevant to the regions of interest, which may facilitate the interpretation of findings and assist in the planning of new experiments. The new infrastructure and procedures will likely also contribute to increased precision of the phenotypic characterization of rodent disease models (see, e.g., [44,45]) and enable comparisons across models.

Brain atlasing is one of several components of the HBP Neuroinformatics Platform (www.humanbrainproject.eu; see, Explore the brain). The platform organizes data from the HBP, in addition to incorporating data from other sources. The tools and services currently being launched will be openly available to the research community. The platform will provide new opportunities for discovery based research through the use of spatial data.

Fig. 4. From brain images to spatially integrated numbers.
Illustration showing spatial integration of data extracted from different serial images cut at different planes of orientation. (A–F) Series of microscopic images showing distribution of parvalbumin mRNA using in-situ hybridization [15,47]. (G–L) Series of microscopic images showing distribution of amyloid plaque primarily in cerebral cortex and hippocampus, in the Tg2576 model for Alzheimer disease, visualized by immunohistochemistry [28]. The images show serial histological sections cut in sagittal (A and B) and coronal (G–H) planes, that are spatially registered to the Allen Mouse Brain Atlas (C and D, and I and J). Labeled features of interest are extracted using the machine learning software tool ilastik [35,36] (E and K). The extracted features can then be represented by centroid point coordinates that can be displayed in 3-D atlas space (F and L). These spatially integrated data points can be co-displayed in specific atlas regions of interest (M and N), and numbers of objects can be extracted and compared at the level of atlas regions (O), in this example the left piriform cortex. This approach allows for the combined analysis of data that would otherwise be difficult to compare due to different cutting planes. Scale bars: 1 mm. (Material in C–L provided by M. Hartlage-Rübsamen and S. Rossner; animal experiments approved by Landesdirektion Sachsen, license no T28/16.) Abbreviations: CB, cerebellum; CP, caudoputamen; CTX, cortex; HC, hippocampus; HYP, hypothalamus; PIR, piriform cortex; TH, thalamus.
integration as well as other neuroinformatics approaches tailored for different categories of data.

We envision that this brain atlasing-based approach to data integration will facilitate a new direction for neuroscience research, where data to a larger degree are shared and re-used to generate new knowledge and improve our understanding of the brain.

Acknowledgements

We thank Christopher Coello, Gergely Csucs, Dmitri Darine, and Hong Qu for expert technical assistance, Antonino Paolo Di Giovanna and Giacomo Mazzamuto for help in preparing experimental data used in Fig. 3, and Martin Vinck, Jeroen Bos, and Sven van der Burg for help in providing the experimental data used in Fig. 5. Material from the transgenic Alzheimer disease mouse model (Fig. 4) was contributed by Maite Hartlage-Rübsamen and Steffen Rossner (Paul-Flechsig-Institute for Brain Research, University of Leipzig, Germany). Funded by the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 720270 (HBP SGA1) and The Research Council of Norway under Grant Agreement No. 269774 (INCF).

References


