



A Comprehensive, FAIR File Format for Neuroanatomical Structure Modeling

A. E. Sullivan¹ · S. J. Tappan¹ · P. J. Angstman¹ · A. Rodriguez¹ · G. C. Thomas¹ · D. M. Hoppes¹ · M. A. Abdul-Karim¹ · M. L. Heal¹ · Jack R. Glaser¹

Accepted: 1 June 2021 / Published online: 2 October 2021
© The Author(s) 2021

Abstract

With advances in microscopy and computer science, the technique of digitally reconstructing, modeling, and quantifying microscopic anatomies has become central to many fields of biological research. MBF Bioscience has chosen to openly document their digital reconstruction file format, the Neuromorphological File Specification, available at www.mbfbioscience.com/filespecification (Angstman et al., 2020). The format, created and maintained by MBF Bioscience, is broadly utilized by the neuroscience community. The data format's structure and capabilities have evolved since its inception, with modifications made to keep pace with advancements in microscopy and the scientific questions raised by worldwide experts in the field. More recent modifications to the neuromorphological file format ensure it abides by the Findable, Accessible, Interoperable, and Reusable (FAIR) data principles promoted by the International Neuroinformatics Coordinating Facility (INCF; Wilkinson et al., *Scientific Data*, 3, 160018, 2016). The incorporated metadata make it easy to identify and repurpose these data types for downstream applications and investigation. This publication describes key elements of the file format and details their relevant structural advantages in an effort to encourage the reuse of these rich data files for alternative analysis or reproduction of derived conclusions.

Keywords Neuromorphology · Neuroimaging · Morphological modeling · Neuron reconstruction · FAIR data · Vasculature reconstruction

Background

Digitally reconstructing and modeling microscopic anatomical structures has become important in many fields of research, none more so than the field of neuroscience. A principal use of this technique is 3D neuronal reconstruction, which allows researchers to create and analyze accurate and quantifiable neuronal models derived from microscopic specimens (Meijering, 2010; Ascoli 2007). By geometrically representing structures contained in the image data, detailed morphometric analyses, simulations, and electrotonic modeling of the neurons can be performed. Unlike raw microscopic image data, the reconstruction data specifies the individual neuronal components and denotes the x, y, and z position of every point of the modeled structures.

The field of digital morphological reconstruction has evolved and expanded for more than 50 years (Fig. 1). The origin of this technique dates back to the seminal work of Edmund Glaser and Hendrik Van der Loos who first described it in the paper published in 1965, “A semi quantitative computer-microscope for the analysis of neuronal morphometry”. This paper describes a system for attaching x-y-z transducers to a microscope stage, tracing the branches of a Golgi-stained neuron, and outputting the result to a plotter (Fig. 1a; Glaser & Van der Loos, 1965). This work was continued by the father and son team of Edmund Glaser and Jack Glaser, who founded MBF Bioscience (at that time known as MicroBrightField) in 1988. MBF Bioscience expanded on the original invention of Glaser and Van der Loos to develop NeuroLucida (Fig. 1b; Glaser & Glaser, 1990), which has become a widely used microscope system for neuron reconstructions in humans and other species (Halavi et al., 2012; Parekh & Ascoli, 2013; Blackman et al., 2014; Usher et al., 2018), with more than 6,500 citations (Rance et al., 1990; Schiller et al., 1997; Wong et al., 2002; Henriksen et al., 2010; Földy et al., 2013; Lázaro et al., 2018; Ullah et al., 2020). Its

✉ Jack R. Glaser
jglaser@mbfbioscience.com

¹ MBF Bioscience, Williston, VT, USA

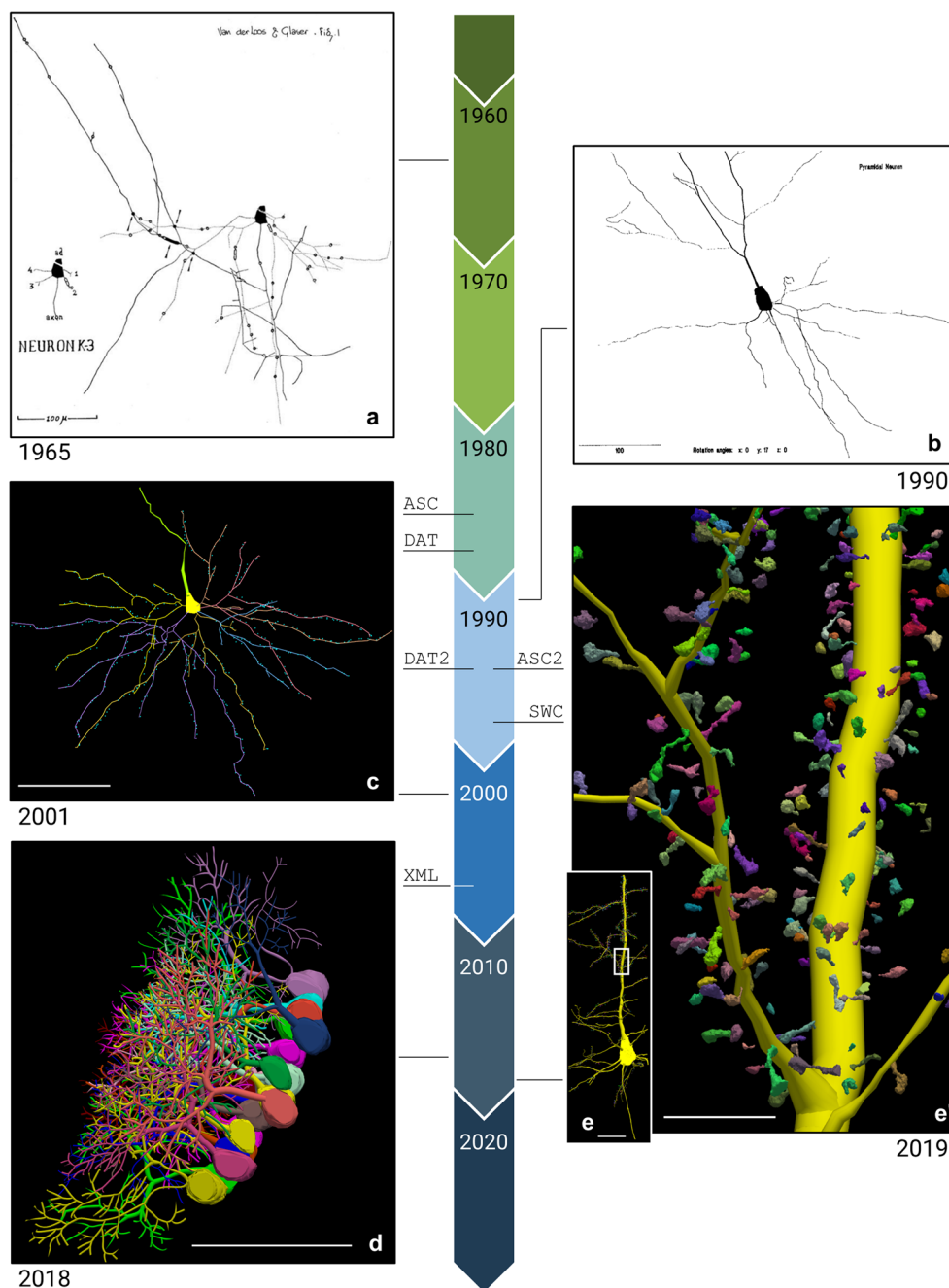


Fig. 1 A timeline depicting the evolution of MBF Bioscience’s digital neuron reconstruction between 1960 and 2020. The lines connecting each image to the timeline indicate when in time the tracing was generated. The birth of each data file format is indicated with the line under each file name: ASC (1986), DAT (1988), DAT2 and ASC2 (1995), SWC (1998), XML (2007). The timeline’s colored arrows each represent one decade and are labeled with the first year of that decade. Each neuronal reconstruction includes the publication date below the image. **(a)** A neuronal reconstruction “produced by the first computer-assisted neuron tracing system, NeuroLucida’s ancestor” (Glaser & Vanderloos, 1965). The scale bar equals 100 micrometers. **(b)** “The hard copy, monochrome output of the neuron of Fig. 3...” in the paper, “Neuron imaging with NeuroLucida—a PC-based system for image combining microscopy.”

(Glaser & Glaser, 1990). The scale bar equals 100 micrometers. **(c)** A reconstruction of a human supragranular pyramidal cell from the Brodmann area (BA), superior frontopolar zone (BA10). Spines have been mapped using point markers (blue) along the cell’s dendrites (Jacobs et al., 2001). The scale bar equals 100 micrometers. **(d)** Purkinje cells (PCs) reconstructed using NeuroLucida 360 from the cerebellar vermis of male mice (Nedelescu et al., 2018). The scale bar equals 100 micrometers. **(e)** A NeuroLucida 360 reconstruction of a *Drosophila* pyramidal neuron, showing the soma, apical/ basal dendrites, and axon segments (Gao et al., 2019). The white box demonstrates the approximate location of **e’**. The scale bar equals 100 micrometers. **(e’)** A zoomed in look at the 3D spines reconstructions of the cell tracing shown in **e**. The scale bar equals 25 micrometers

popularity has overtaken other integrated microscopy systems such as the Eutectic Neuron Tracing System (NTS; Parekh & Ascoli, 2013).

From its initial release in 1988, NeuroLucida made use of two file formats for storing and archiving tracing data: a portable ASCII text format, and a binary platform-dependent format. The ASCII text format, normally referred to as the NeuroLucida ASC format, was conceived as a human-readable and portable format that could be easily imported into other software, regardless of underlying computing hardware or operating system. The binary format, referred to as the NeuroLucida DAT format, was intended as a more computationally efficient alternative for use within the NeuroLucida environment. Both of these formats have continued to evolve over the past 30 years to accommodate new research requirements and as of this writing are still fully supported by MBF products.

By the mid-1990 s, with the increase in memory and computing power of personal computers along with the connectivity afforded by the now burgeoning internet, full-cell neuronal tracings created with NeuroLucida were being created and shared (Turner et al., 1995; Mpodozis et al., 1996; Jackson & Cauller, 1997; Wu & Karten, 1998; Prusky & Arjannikova, 1999; Gabriele et al., 2000; Ghosh et al., 2011). In 1998, a paper presented an online archive system for neuronal reconstructions (Cannon et al., 1998). The archive (www.neuro.soton.ac.uk), initially populated by a set of 87 neurons from the hippocampus reconstructed using NeuroLucida, offered an editor and a format converter. The system's native archival file format, SWC, was a simple and portable text-based format that could represent arbitrarily-complex neuronal geometries by providing a list of positions and radius in a hierarchical parent/child arrangement that form a collection of minimally connected cylindrical segments. Due to its simple, portable, and topologically constrained format, SWC became a preferred file format for archival and sharing of neuronal tracings for use in compartment modeling and computational neuroanatomy applications. One notable example of SWC's success has been its adoption as the standard format used by NeuroMorpho.Org (Ascoli et al., 2007), to date the largest collection of publicly accessible 3D neuronal reconstructions. Though the database converts all data to the SWC format for continuity, the database currently contains 65,898 neuronal reconstructions generated from MBF Bioscience software in formats including DAT, ASC, and XML (NeuroMorpho.Org, n.d.). While the SWC format is more compact and sufficient for compartment modeling and basic morphometry applications, it offers minimal or no support to detailed subcellular structures such as spines, somas, varicosities, and puncta, as well as support for other biologically relevant structures such as blood vessels or integrated annotations constructs such as markers, text, regions, and contours. It also lacks the means to preserve information about any original source data. Evolution of the SWC format to

support 4D, time-varying data has occurred (e.g. He & Cline, 2011; Nanda et al., 2018), however, the above limitations still apply.

Over the last twenty years, a shift from analyzing slides at the microscope to capturing images and image volumes for post-hoc analysis has steadily occurred. This change has expanded annotation potential for microscopy modalities (fMOST, microCT, EM, multiphoton, confocal, light sheet) and image data sizes. The advancement of microscope technologies with increasing resolution and imaging capabilities together with exquisitely specific labeling methods (pseudorabies transsynaptic labeling and transgenic techniques to name a couple), has enabled the visualization of smaller and subtler structures. This has driven the need for representing the relationship between neuronal morphology and smaller, subcellular structures such as spines, synapses, and varicosities or boutons (Jacobs et al., 2001; Le Bé et al., 2007; Arellano et al., 2007).

Starting in 2007, MBF created a new file format based on the Extensible Markup Language (XML). XML, a World Wide Web Consortium (W3C; Bray, Paoli, Sperberg-McQueen, Maler, & Yergeau, 2008) standard, is well accepted and recognized as having a number of important attributes for data sharing: simplicity, extensibility, and self-description. The new format, often referred to as the NeuroLucida XML format, is the subject of this article. It recreates the same data elements already present in the ASC and DAT files, while greatly improving data accessibility and extensibility. Additionally, XML's hierarchical structure allows the contained reconstruction data to unambiguously denote the relationship of described elements in the context of their anatomical region at the tissue or organ level. The NeuroML format has a similar file structure and is currently recognized as an INCF standard for modeling neuronal systems. During development, the NeuroML format took into account NeuroLucida's cellular data structure, which is likely why structural congruencies with the neuromorphological file are seen (Crook et al., 2007). Though the format structures are similar, the neuromorphological file format is a more broadly based format for neuroanatomical modeling than NeuroML which lacks some morphological detail.

Very recently, a collaboration with the FAIR Data Informatics (FDI) lab through participation in the NIH Common Fund Initiative, Stimulating Peripheral Activity to Relieve Conditions (SPARC), has prompted MBF Bioscience to adapt the XML file format once again to embrace Findable, Accessible, Interoperable, and Reusable (FAIR) data standards (Wilkinson et al., 2016). The project highlights the importance of generating FAIR digital reconstruction and modeling data for anatomical and neuronal structures by incorporating the necessary metadata at the file level. MBF Bioscience has taken steps to ensure the image segmentation data produced by its products abides by the FAIR data standards for neuroscience, defined and promoted by the International Neuroinformatics Coordinating Facility (INCF;

Wilkinson et al., 2016). Following these modifications, the neuromorphological file specification was released (Angstman et al., 2020). As demonstrated by implemented data standards such as the Neurodata Without Borders neurophysiology data standard (Rübel et al., 2019), defined data format standards help to ensure generated resources are reusable and reproducible by the scientific community (Abrams et al., 2021). We believe that embracing the transparency and systematic organization of the XML data file format will permit for further accommodation of the scientific community's advocacy for open and accessible research.

Purpose

The MBF Bioscience neuromorphological segmentation file structure has been driven for over 30 years by the ever-advancing science, technology, and input of neuroscientists throughout the world. The purpose of this paper is to document the relevant, systematic, and flexible nature of the file structure and demonstrates its effectiveness as a solution for the microscopic anatomy morphology data standard. Through these definitions, the neuromorphological file specification serves as an important format for the exchange of neuromorphological data for archival and exploratory research. We hope this will facilitate the development of other software and tools for downstream applications and investigation of the increasing scope of data that is stored in this file format.

File Structure Summary

A brief description of selected data elements of the neuromorphological file structure are detailed in this section. These elements were chosen to highlight their unique data structure's direct impact on microanatomical models, analytics, and data reusability. The neuromorphological file structure is fully defined in the Neuromorphological File Specification available at <http://www.mbfbioscience.com/filespecification> (Angstman et al., 2020). The file specification will continue to be updated as needed to define added and/or modified data elements in use with related MBF Bioscience software.

The neuromorphological file format is an Extensible Markup Language (XML) 1.0 (Fifth Edition) format and includes two organizational aspects, elements, and attributes. These aspects are defined in the XML file specification provided by W3C (Bray, Paoli, Sperberg-McQueen, Maler, & Yergeau, 2008). All neuromorphological data files include header elements, introducing the file via essential metadata, followed by morphological structure modeling elements.

File-Level Metadata

In the header section, metadata attributes define the expected file structure, software application, and version number of the neuromorphological file structure. By establishing the expected data structure for the document, the ability to interpret

Figure 2(a) shows a software window titled "Welcome! To begin annotation, specify the following:". It is divided into two main sections: "Subject information:" and "Select criteria for anatomic terms:".

The "Subject information:" section contains four fields: "Species:" with a dropdown menu showing "Rattus norvegicus"; "Subject ID:" with a text input field containing "001"; "Sex:" with a dropdown menu showing "Male"; and "Age:" with a text input field containing "12" and a dropdown menu showing "Weeks".

The "Select criteria for anatomic terms:" section contains three fields: "Organ:" with a dropdown menu showing "Kidney"; "Species:" with a dropdown menu showing "Rattus norvegicus"; and "Parcellation:" with a dropdown menu showing "Species Independent".

Figure 2(b) shows the XML code generated from the form in (a):

```
<sparcdata>
  <subject species="http://purl.obolibrary.org/obo/NCBITaxon_10116"
  subjectid="001" sex="Male" age="12 Weeks"/>
  <atlas organ="Kidney" label="Species Independent"
  rootid="http://purl.org/sig/ont/fma/fma7203"/>
</sparcdata>
```

Fig. 2 (a) The subject and annotation term list selection window within MBF software. The fields in the left section detail the subject information of the sample origin. The fields in the right section determine the anatomical term list provided to the user for annotation. The selected values indicate this sample originated from a 12-week-old male rat with the subject identifier 001. The anatomical term list selected for annotation was the rat kidney term list. The parcellation indicates Species Independent, meaning there currently is no parsed term list for the rat

kidney. Instead, a generic term list of all kidney anatomies, independent of species, is provided to the user. (b) A neuromorphological data file related to a microscopy sample from a 12-week old male rat kidney delineated using anatomical terminology from the Foundational Model of Anatomy (FMA) ontology database. The species and <atlas> rootid store IRIs that are linked to the species and the term list origin selected in a. The IRI includes the unique identifier for the species or parcellation

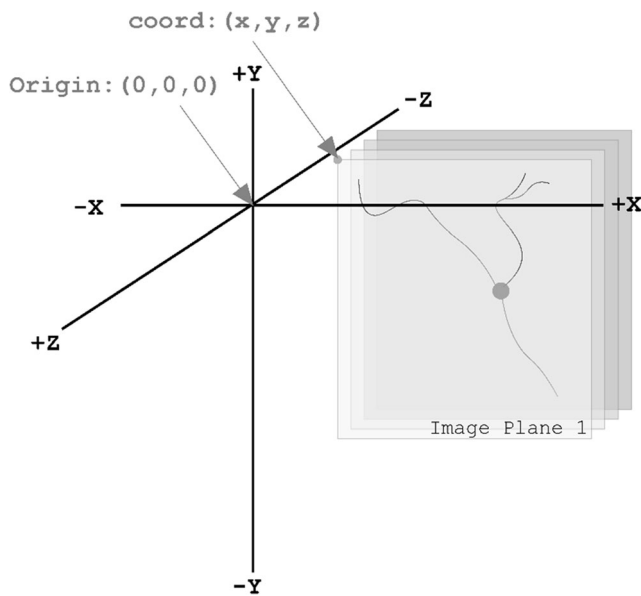


Fig. 3 Demonstrates the 3D coordinate space with an origin point of (0, 0, 0). The gray planes represent a 3D image volume with an image location coordinate, coord: (x, y, and z). Note the direction of the Z-axis. The most positive image plane of the 3D volume is the first image plane. The following image planes are in the same X and Y location, but their z location descends incrementally based on the z scaling. The units of this coordinate space are in micrometers (μm)

provisional file formats is preserved. Header elements are not used to model morphological structures and therefore differ from traced data elements.

Essential information regarding the software application includes the application name, application version, application Research Resource Identifier (RRID), and institution RRID. The institution RRID specifies the company, organization, or institution that produced the software which generated the neuromorphological file. Both application RRID and institution RRID are globally unique and persistent identifiers registered in the SciCrunch knowledge base. Reporting metadata regarding the software application used to generate the data

file ensures the data generated is reproducible, reusable, and citable. This metadata will never get separated from the traced data because it is all saved at the file-level, unlike a sidecar metadata file.

SciCrunch InterLex Terminology Linkage

Another header section of the neuromorphological file format stores critical subject and annotation metadata for each data file. The subject metadata is user-defined (left panel of Fig. 2a) and includes fields for subject species, identifier, sex, and age of the sample. The anatomical terminology list used for annotating neuromorphological or fiducial structures are selected using an API connection with the SciCrunch InterLex Terminology Portal (right panel of Fig. 2a). This information is recorded in the < atlas > child element of the < sparadata > section of the neuromorphological file format (Fig. 2b). The metadata in this section accommodates organ, species, parcellation, and International Resource Identifier (IRI) for the atlas or ontology database if no atlas is available.

The unique identifiers for the subject species, the anatomical term list, and delineated anatomical regions are recorded in the data file. This aids in repurposing and reusing any anatomically relevant data. A list of terms separated by organ, species, and atlas/parcellation scheme is provided to the user for annotation. Terms in the SciCrunch database are a collation of all recognized anatomical databases (e.g., NCBI, FMA, UBERON). Additions and edits are managed by SciCrunch (Grethe et al., 2014). This infrastructure allows investigators to use a consistent and structured lexicon when referring to multi-species and multi-scale anatomies. With the API connection to the database, up-to-date term lists can be provided at the time of annotation, ensuring the data file produced supplies a robust understanding of how the traced data was derived. Reducing the barrier to include relevant metadata about the experiment is key to obtaining more data that is interoperable.

Table 1 Neuron summary analysis. The neuron summary analysis produced from the tracing of the cell in Fig. 4b. The quantity of the cell bodies contours, axons, and dendrites that were analyzed is reported in column 2. The field describing the quantity of the cell body indicates the number of 2D contours that construct the single 3D cell body. The 3D measurements such as surface area and volume apply to the 3D cell body equivalent to the shell of the 2D contours. The 2D measurements such as length, area, and mean area were obtained using the 2D contours. The

length reported for the cell body indicates the perimeter of all contours that make up the individual cell body (MBF Bioscience, 2020). The area and mean area values are reported as N/A (not applicable) for the 3D Axon and Dendrite trees. The complexity refers to the normalization and comparison of trees among fundamentally different neurons (Pillai et al., 2012). This analysis does not apply to cell bodies, therefore, N/A is the value reported

Name	Quantity	Length (μm)	Mean Length	Area (μm^2)	Mean Area	Surface (μm^2)	Mean Surface	Volume (μm^3)	Mean Volume	Complexity
Cell Body	20	1440.628	72.031	6135.247	306.762	14407.086	14407.086	1992.254	1992.254	N/A
Axon	1	1013.302	1013.302	N/A	N/A	3180.255	3180.255	897.261	897.261	1013.302
Dendrite	6	1395.549	232.591	N/A	N/A	4659.897	776.649	1450.534	241.756	43262.016

SciCrunch services link recognized synonyms to the term identifier guaranteeing a database query of a term abbreviation or synonym will still return all applicable results by pulling any data tagged with the term. The clear file-level subject and annotation term list metadata within the neuromorphological data files can be queried and allows the files to be sorted by species, subject ID, sex, age, and organ of the sample origin.

Coordinate Space

The schematic in Fig. 3 demonstrates the coordinate space for all neuromorphological data files is three-dimensional. All coordinates and measurements are reported in micrometer units (μm). The origin point of the coordinate system is ($x = 0, y = 0, z = 0$).

Using the image scaling defined as micrometers/pixel, the anatomical structures are readjusted into a real-world coordinate space enabling robust qualitative and quantitative analysis to be performed on the data elements. The example in Fig. 4b demonstrates the digital reconstruction of a 3D microscopy image (Fig. 4a). The morphometric measurements in Table 1 were obtained using the reconstruction in Fig. 4b and NeuroLucida Explorer's neuronal summary analysis (MBF Bioscience, 2020).

As seen in Table 1, the individual data elements (cell body, dendrite, and axon) can be analyzed as unique structures though they are all part of the same neuron. The importance of storing detailed coordinates for each individual element within the neuromorphological data file lies in the analysis flexibility. Not only can these structures be analyzed individually, but they can be analyzed in relation to one another. It is possible to determine the length and volume of each dendritic tree. Using a similar analysis method, the volume of dendritic trees that fall within an anatomical volume such as an airway or a colonic layer can be determined, providing the neuronal density within that region (Table 2).

Microscopy Image Association

Raw image data is not saved within neuromorphological data files, rather they are linked with a file location, name, and other information about the image. Because the image data is not saved to file, the tracing data file is “lightweight”, easy to store, transfer, and read. However, the file path and name of the image(s) are conserved, conveying the provenance of the data as it relates to the images in which it was derived.

The neuromorphological data file can be associated with any number of source images. The images can be either 2D (a single image plane) or 3D (multiple image planes from a single file or multiple files). Image data can be combined in several ways inside the neuromorphological data file. The simplest is the single 3D or 2D image file. The lineage of the derivative data is recorded in the file, regardless of

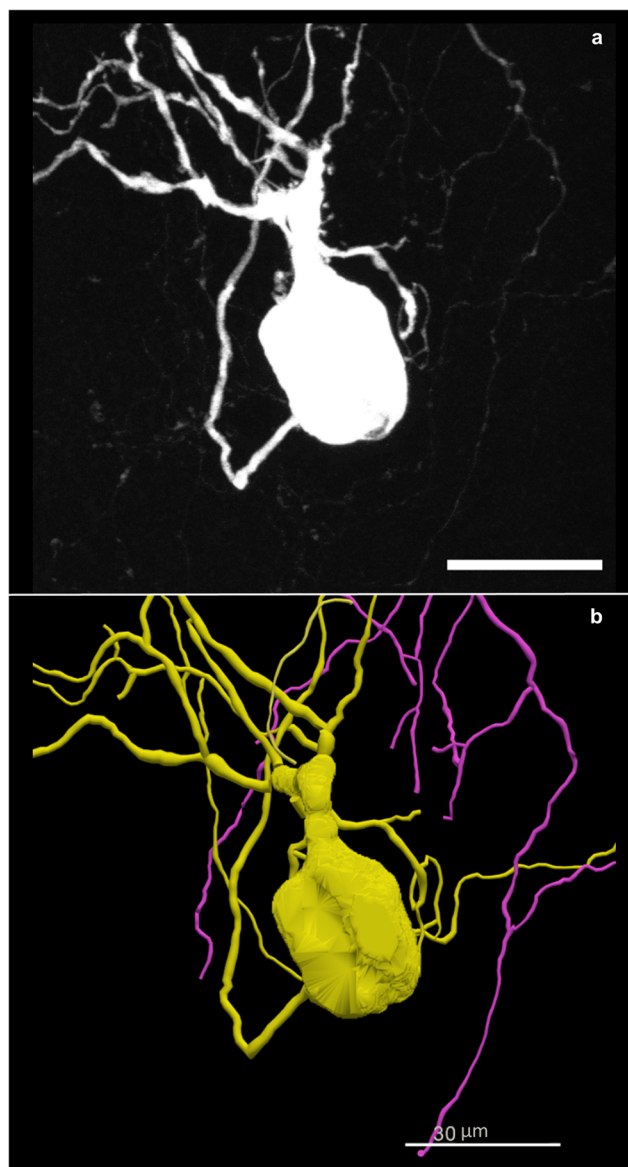


Fig. 4 (a) A neuron from the mouse stellate ganglion backfilled with Neurobiotin. A 3D image was acquired on a Leica confocal microscope with a 40x objective lens. The scale bar equals 30 micrometers. The image scaling is: $X = 0.2228 \mu\text{m}/\text{pixel}$, $Y = 0.2228 \mu\text{m}/\text{pixel}$, $Z = -0.5 \mu\text{m}$. (b) A tracing of the neuron in a created using NeuroLucida 360. The data is stored in the neuromorphological file format. The trace data elements include contours that make up the cell body and trees that reconstruct the neuron's dendrites and axon. The scale bar equals 30 micrometers (Cho et al., 2020)

complexity. This type of record keeping is rarely cataloged in other digital reconstruction data formats file formats.

Image scaling in x and y (micrometers per pixel) and the z scaling (micrometers between image planes) for the image(s) is reported alongside the related image name and file path. Additionally, the total number of image planes is included to further convey the associated image(s) structure. The x, y, and z coordinates of the upper, left-hand corner of an image are also reported. These values provide the location of the image

Table 2 A description of child elements and attributes of the neuromorphological file's edgelist element (Angstman et al., 2020). A description for the edgelist's child element of a vessel, its child elements,

attributes, and values. An indentation before the element name and a dashed border are used to indicate a child element in the table

Element	Attribute	Description
<edgelist>		A set of edgelist child elements whose attributes define the relationship between an edge and its two nodes.
<edgelist>		Defines the relationship between edge and node elements, indicating the branches and loops that form the vessel.
	id	The unique identifier generated for every edgelist.
	edge	Indicates the ID of the edge element that the edgelist relates to a start and end node.
	sourcenode	The ID of the node that the edge starts from.
	targetnode	The ID of the node that the edge ends at.

within the data file coordinate space. To reuse multi-image neuromorphological files, it is necessary to know the image scaling, number of image planes, image order, and image location within the 3D coordinate space. These elements are stored alongside the corresponding image file path(s) and name(s). The file-level metadata indicating all source image(s) associated with the morphological data promotes the reusability of the data captured within the neuromorphological file format.

The ability to relate neuronal morphology reconstructions to multiple source images is a valuable aspect of the neuromorphological file format's versatility. Applications of this format include the generation of morphological and anatomical reconstructions on 2D and 3D image montages as well as image data comprised of many individual images of physically serially sectioned tissue. In the case that multiple 2D image planes are combined to construct a 3D image volume, all image paths and file names are recorded in the neuromorphological file. The image order and z location, key metadata for repurposing this data, is noted alongside the image element for these image types. The same is true for a singular image generated by merging source images that make up color channels to construct a multi-channel microscopy image. Lastly, registration of neuronal reconstruction from a high-resolution image to an annotated low-resolution image is possible due to this structure. This multi-resolution image segmentation process can provide anatomical context, especially in large organs or tissue samples that are difficult to image comprehensively at high magnification.

This case study depicted in Fig. 5 demonstrates how Cho et al. utilized multi-resolution image segmentation strategies to understand the structure-function relationships of cells within the stellate ganglion (Cho et al. Data set in progress). After performing electrophysiological recordings, the cells were labeled with a fluorescent dye and processed for imaging. Using Neurolucida 360, Cho et al. reconstructed select neurons from images taken with a 40x objective lens (Fig. 5b). The entire stellate ganglion was also imaged with a 10x objective lens (Fig. 5c) and contoured using integrated FAIR anatomy terminology lists. The neuron reconstruction of the

40x, high-resolution image (Fig. 5b) was registered to the appropriate location in the 10x image (Fig. 5f demonstrates the x, y, and z location of the backfilled cell bodies can be discerned), providing anatomical context within the stellate ganglion. The registration of cellular reconstructions to the whole stellate ganglion allowed the researchers to bring the physiological and morphological data into context with the entire ganglion (Cho et al. Data set in progress).

Incorporation of file-level metadata including microscopy image association is an essential component of the data provenance of the morphology file. The neuromorphological file format takes the approach of maintaining the original source image as its original format, ensuring all source image metadata is conserved (e.g. objective magnification, channel IDs, modality). By incorporating access to both the source image and its derived digital reconstruction files, data repositories can enable users to re-use, repurpose, or reproduce this data by providing all necessary components. By only saving the file location(s) and name(s), factors like image dimensions, number of image planes, number of channels, and bit depth do not affect the reconstruction data file size. Storing this data can increase the size of the reconstruction data file by over 99%. File size can bog down downloading, 3D visualization, and web rendering speeds. For those only interested in repurposing the reconstruction data, this is an exceptional benefit as it notably decreases load times. These factors illustrate the value of the neuromorphological file format's microscopy image association scheme.

Morphological Structure Modeling

The traced data elements include all data models of neuromorphological structures and additional annotations. A data file will not necessarily contain all types of traced data elements and typically will include more than one of a single traced element in a data file. For example, one file could have cell body contours and neuronal trees where another could incorporate cell body contours, neuronal trees, and marker elements.

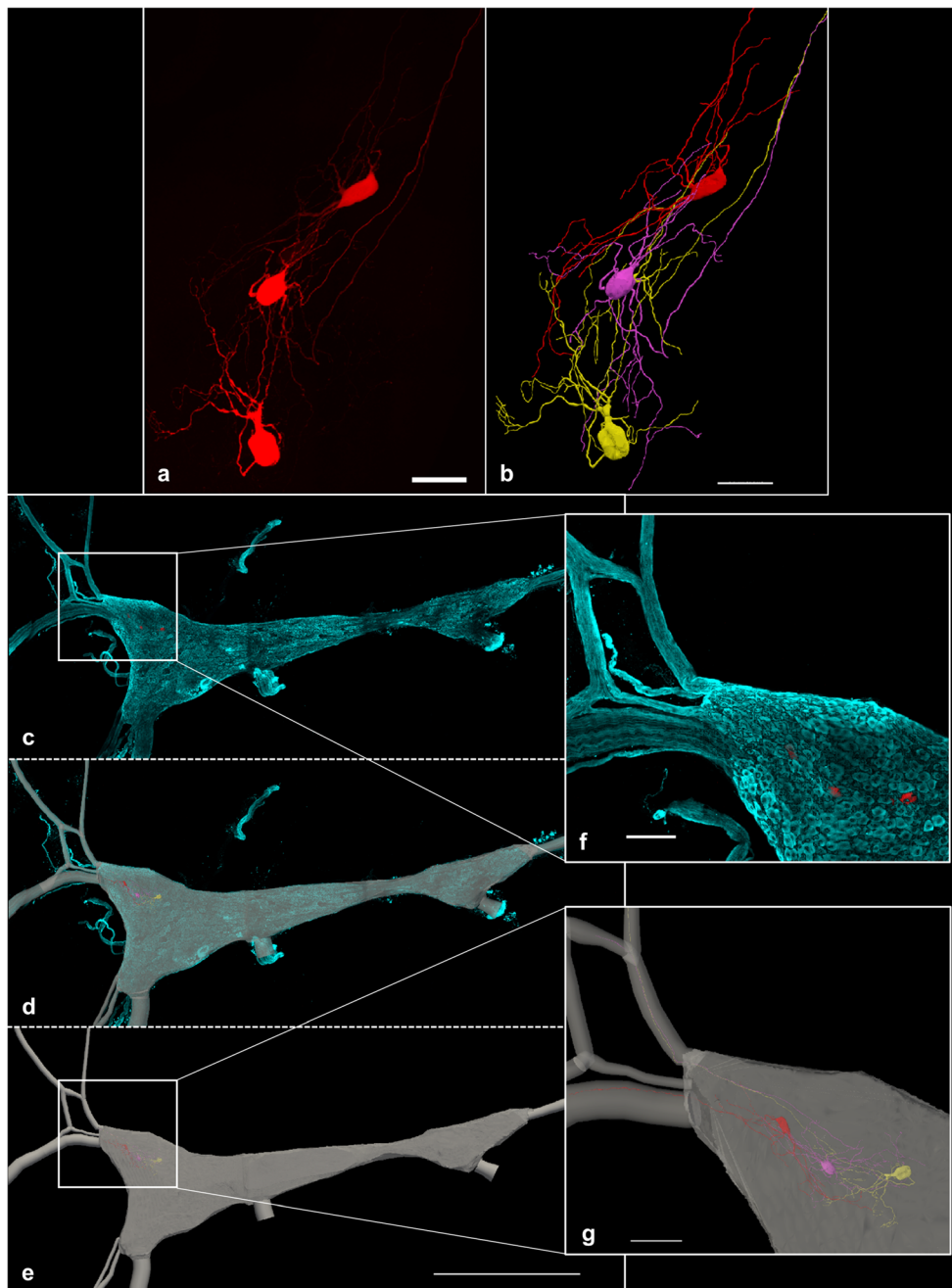


Fig. 5 This figure illustrates the application of the multi-resolution image segmentation that is possible with the neuromorphological file format. **(a)** The 3D image, acquired on a Leica confocal microscope using a 40x objective lens, shows neurons from the stellate ganglion backfilled with Neurobiotin. The scale bar equals 50 micrometers. **(b)** A neuronal reconstruction obtained from the 40x, high-resolution image in **a**. Tree elements were used to represent the neuronal dendrites and axons of the cells. The cell bodies are represented using serial z contours, shelled into a three-dimensional volume. **(c)** A 10x, low-resolution tile scan image was acquired using Leica confocal microscope. The images include the entire stellate ganglion labeled with tyrosine hydroxylase (TH) (cyan) and the neurons backfilled with Neurobiotin (red). This same group of backfilled

A variety of trace data elements have been added to the neuromorphological file, their format augmented for performance, morphometric modeling accuracy, and analytical

neurons was imaged at 40x **(a)**. **(d)** A 3D reconstruction of the 10x, low resolution whole stellate ganglion image **(c)** overlaid with that image. **(e)** 2D contours of the ganglia's area were delineated at serial z image planes. They were shelled into a 3D volume to represent the stellate ganglion (gray). Tree elements were used to represent the path of the nerve fibers stemming from the ganglion. These structures were segmented using the 10x image in **c**. The 1000 micrometer scale bar shown in **e** is applicable **c**, **d** and **e**. **(f)** A zoomed in snapshot of the boxed location displayed over **c**. Scale bar equals 100 micrometers. **(g)** A zoomed in snapshot of the boxed location displayed over **e**. Axon innervation to the Inferior cardiac nerve (top) and Ventral ansa subclavia (bottom) can be mapped and visualized. Scale bar equals 100 micrometers (Cho et al. 2020)

potential based on feedback from top neuroscientists in the field. Below, the structure of relevant neuromorphological data elements are detailed providing the background for discussing

the FAIR aspects of the data and the structure's relevance to representing and analyzing the neuromorphology.

Trees

In the neuromorphological file structure, the tree element is used to represent non-looping branching structures within microscopy data such as axons, dendrites, and airways. Trees consist of an origin, branches, nodes, and endings. The one neuronal dendrite within the 3D confocal image shown in Fig. 6a was reconstructed (Fig. 6b and c). The reconstructed dendrite is represented in the schematic in Fig. 6d to demonstrate the neuromorphological file's tree data structure. The starting point of a tree element is referred to as the origin (O) and the points that follow make up the root branch of the tree. All trees must have at least an origin and root branch, but typically have branching points called nodes. Nodes are where a segment of the tree splits into multiple branch child elements. The branch elements are made up of an ordered list of points that connect nodes to nodes, and nodes to endings. Endings are the last point of a branch or tree where the segment terminates.

The neuromorphological file format stores each point location (x, y, z position, and diameter) for a tree and its branches. The tree element is formatted as described above, to clearly relate relevant anatomical features such as the nodes indicating origin, ending, and bifurcation points. High-level and detailed analysis can be performed on the computer-readable morphometric reconstruction data of trees and branches. This includes quantifying the number of branches or terminals within an anatomical region or determining the proximity of two branches such as a neuronal fiber and a blood vessel. By classifying the neuronal fiber types, for example, axons, dendrites, and apical dendrites, the total number of each can be determined and the details of the trees can be compared. Tree details including length, surface area, volume, total terminations, branch angles, complexity (Table 1, column 11), etc., and even the extension of neurons can be determined. Analyzing realistic, meaningful, and quantifiable neuron reconstructions can help researchers to conclude the structure-function relationship that defines a specific subset of neurons.

Recent developments have advanced the tree classification to further permit branch specific annotations via the API connection with the SciCrunch InterLex Terminology Portal detailed above. As seen in Fig. 7c, the rootclass defines the anatomical classification of the tree or branch. The TraceAssociation property of the tree or branch elements stores an IRI that is linked to the anatomy term used to annotate the branch, including the globally unique and persistent identifier (PID) for that term. The data file communicates the term list that the root class was selected from via the recorded anatomical term list in the <atlas> child element of the sparcd data section.

Tree structures such as neuronal branches or bronchial trees (Fig. 7a) can be modeled and classified via SciCrunch curated ontologies. Figure 7b demonstrates the classification of two branches of the airway with the appropriate anatomical names found in the SciCrunch term list. Note the data file (Fig. 7c) includes both anatomy terms (Trachea and Left main bronchus) and two separate term identifiers, one for each segment of the tree. Other anatomical structures that follow a non-looping branch pattern can use a similar modeling structure.

Reporting each point of the tree enables child elements to be associated at unique locations along the branch. This is valuable for representing neuronal morphologies such as spines, synapses, and varicosities. By representing the relationship of structures like the tree and spine in the data file, joint analysis can be performed to determine the tree's spine density, spine class densities, or average spine diameter. Even the distance from an individual spine's base to the tree origin can be calculated. The spine-tree association is further detailed in the section below.

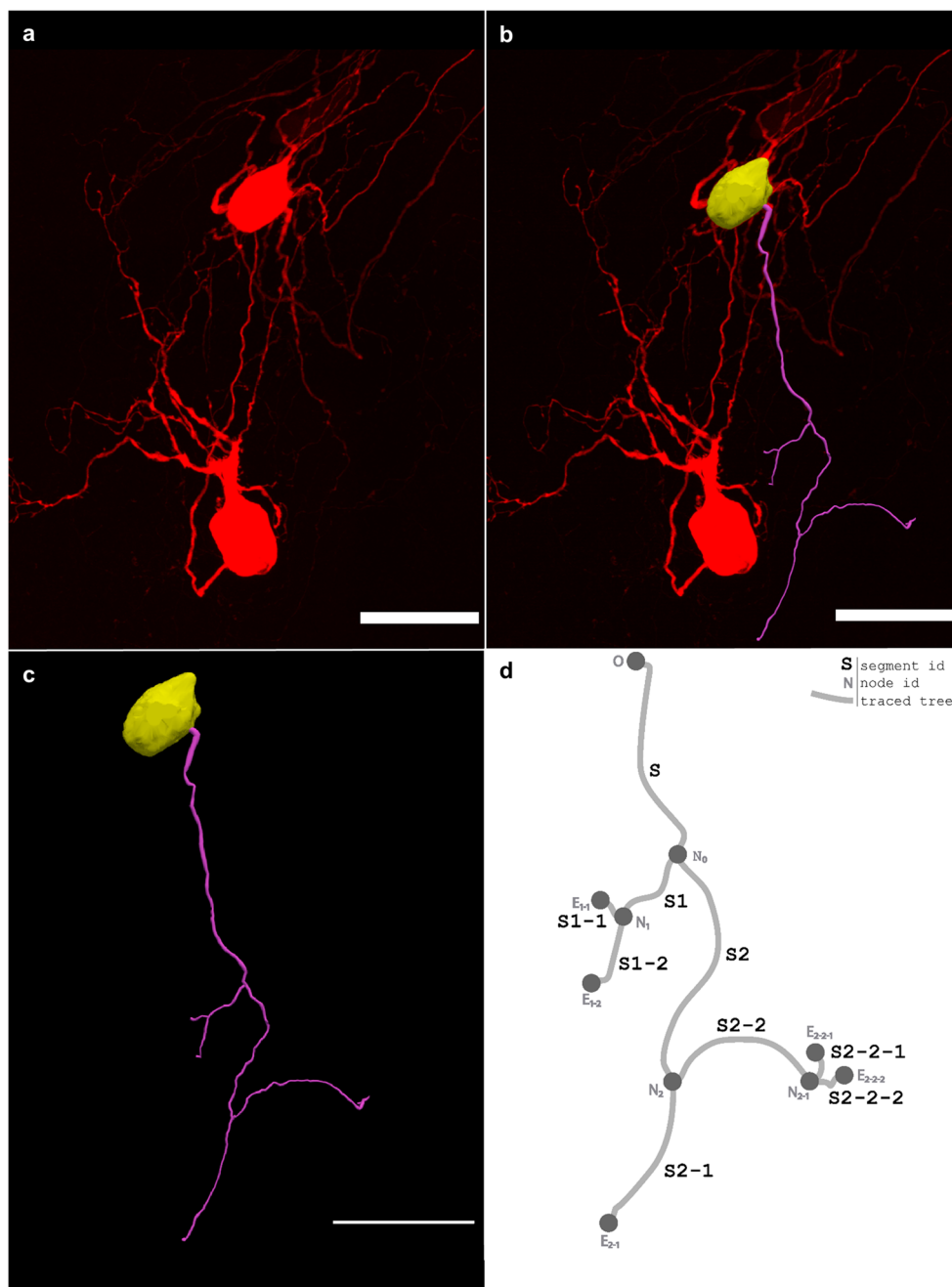
Dendritic Spines

Spines are small protrusions off of dendritic branches. The neuromorphological data format embeds the spine element in the associated branch at the tree points that the spine occurs.

The Backbone property of the spine describes the points that construct the spine volume (Fig. 8a). The number of total points is listed first. Following this, the x, y, and z coordinates along with the diameter of each coordinate are listed in order of proximity to the branch. The diameter thickness of each point determines the three-dimensional shape of the modeled spine. The first point is the branch insertion point along the branch centerline. The next two points model the base and neck of the spine, with the second point element of the spine on the surface of the dendrite. The fourth and fifth points model the head and tip location of the spine.

By modeling the key dynamic connection component of brain cells, dendritic spines, we can analyze how these dendritic spines grow, disappear, and change shape over time. Researchers are able to extract accurate 3D measurements of these dendritic spines to validate novel techniques testifying to the data's reliability. By analyzing 1,500 dendritic spine reconstructions, Gao et al. were able to confirm that physically increasing tissue size through expansion microscopy did not cause damage to the structural components. This helped to demonstrate the utility of the novel expansion microscopy (ExM) technique and validate its ability to obtain comprehensive morphometrics of delicate dendritic spines in combination with lattice light sheet microscopy (LLSM) and digital reconstructions. Due to the data's format, the range of spine head diameters, neck diameters, backbone lengths, and neck backbone lengths could be extracted from the morphometric models. The team found the

Fig. 6 (a) The 3D image, acquired on a Leica confocal microscope with a 40x objective lens, shows neurons from the stellate ganglion backfilled with Neurobiotin. (b) A 3D reconstruction of one dendrite and one cell body of the backfilled neuron from the stellate ganglion (a) overlaid with that image. (c) The same reconstruction shown in b with no image data. A model of one cell body (yellow) and one neuronal tree (pink) was produced using NeuroLucida 360 (Cho et al., 2020). (d) An unscaled diagram demonstrating the structure of a tree with each segment shown as a line and labeled with the segment name (ex. S2-2-2). The origin (O), nodes (N), and endings (E) of the tree are marked with a circle. The root segment (S) begins with the origin (O) point and terminates with the node (N₀). The child segments of N₀, S1, and S2 terminate with nodes N₁ and N₂. The child segments of N₁, S1-1, and S1-2 terminate with endings E₁₋₁ and E₁₋₂. N₂ has two child segments, S2-1 and S2-2. Segment S2-1 has no bifurcations, so it terminates with ending E₂₋₁. Segment S2-2 bifurcates at node N₂₋₁. Lastly, the branches S2-2-1 and S2-2-2 terminate with endings E₂₋₂₋₁ and E₂₋₂₋₂. The scale bars in (a)–(c) are equal to 50 micrometers



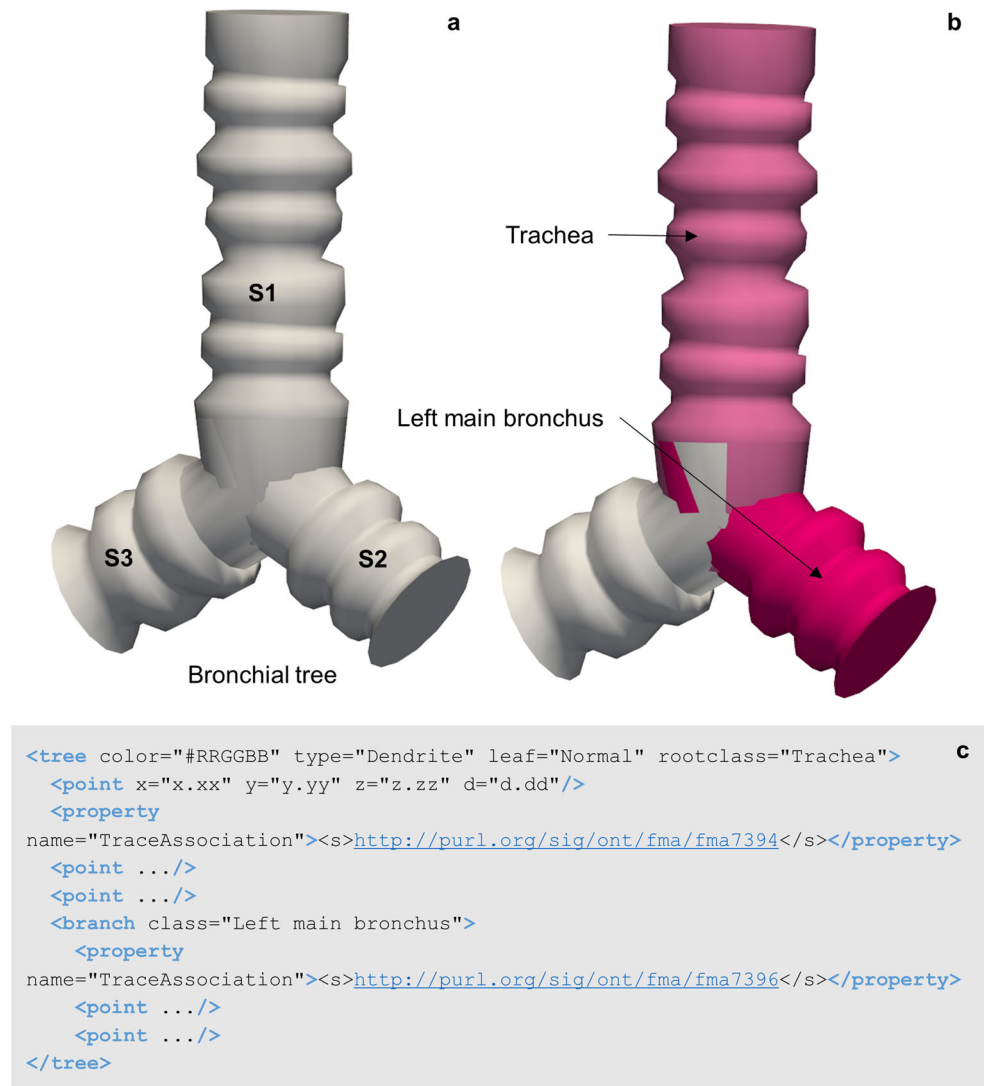
spine metrics collected from the expansion lattice light sheet microscopy (ExLLSM) samples proved consistent with a relevant electron microscopy (EM) study (Gao et al., 2019). As mentioned above, the relationship between spines and dendritic trees can be analyzed to provide a contextual understanding of the nanostructure's arrangement on a unique grouping of neuronal fibers. Individual spine metrics written to the data file include but are not limited to volume, classification (Rodriguez et al., 2008), and backbone lengths. Even subunits of the spine can be compared such as spine head position, head length, and head diameter vs. neck length, and neck diameter. The level of detail

available for analysis is due to the file format's detail of the element used for modeling spines.

Anastomoses

Structures like vasculature, nerves, or fascicles, can be represented with a different branching structure than neuronal trees. In the neuromorphological file format, vessel elements are made up of points called nodes that connect branches called edges (Fig. 9a). An edge is a collection of connected points. The connection relationships are described with the edgelist element (Fig. 9b).

Fig. 7 (a) A 3D reconstruction of a bronchial tree utilizing the tree elements of the neuromorphological file format. (b) The Trachea and left main bronchus segments of a lung airway named using the SciCrunch terminology link through MBF Bioscience software. (c) The data representation of the bronchial tree. The point elements that fall directly within the tree element represent the first segment (S1). Segment 1 (S1) of the bronchial tree is classified as the Trachea. The branch element and the points that are enclosed make up segment 2 (S2) of the bronchial tree, classified as the Left main bronchus. Unique identifiers for each term are indicated in the segment's Trace Association property



The vessel elements can have edges that loop, whereas tree elements can only branch. The looping capability of the vessel elements is modeled with what is known as a graph structure. The graph structure has a wide range of applications and can be used to model anything from hyperlinks in webpages to transportation of goods (Siek, Lee, & Lumsdaine, 2001). What makes it so powerful for modeling anatomical morphologies is its ability to represent biological structures such as anastomoses in vascular or nerve networks.

In addition, the same branch specific system for anatomical classification of trees has been implemented for vessel segments. Segment names and PIDs are stored using the rootclass and TraceAssociation. The corresponding atlas or parcellation scheme IRI is stored in the sparcddata element. This allows scientists to generate FAIR data files with specific, globally unique anatomical terms for structures modeled using the vessel elements at the time of segmentation.

Due to the data structure that stores each unique vessel point and its diameter, detailed analysis of length, surface area, volume, average thickness, and even tortuosity can be obtained through the vascular models. In a recent study, vascular reconstructions were generated of micro-CT rat brains on control and blast-exposed rats to examine the effects of traumatic brain injury on vascular networks (Gama Sosa et al., 2019). Though the micro-CT images showed a clear decrease in vasculature structures, the reconstruction of both control and blast-induced rat brains provided meaningful quantitative results to back this observation. The results showed that the blast-exposed rat brains decreased in length by about 50 %, in surface area by about 50 %, and in volume by about 60 % (Gama Sosa et al., 2019).

The structure also permits anastomose analysis including loop length, surface area, volume, and average diameter, which can aid in the classification of looping structures. The

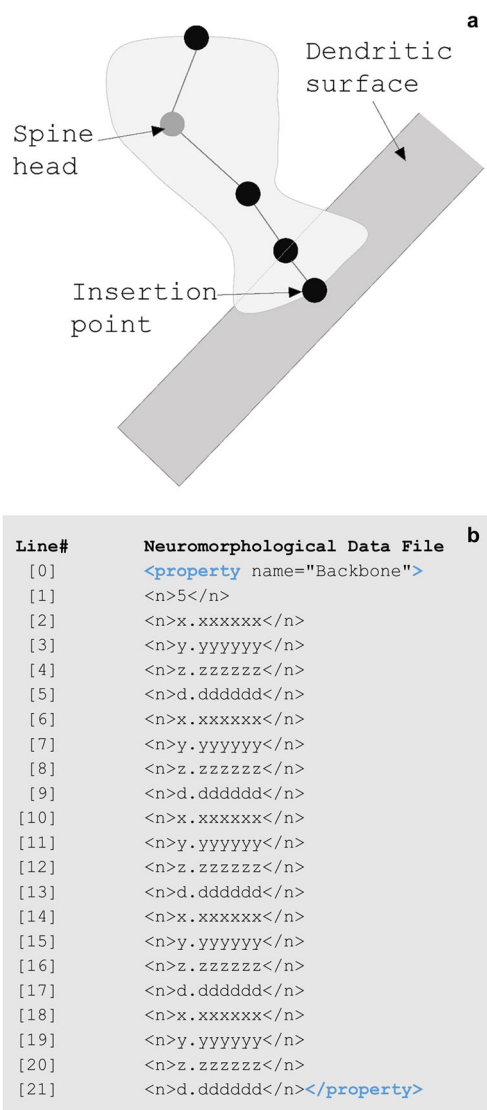


Fig. 8 (a) A diagram of a dendritic spine along a neuronal tree. The five points of the spine are represented with circles. The coordinates of these points are reported in the `<property name="Backbone">` number string including an x, y, and z location along with a thickness, d. The spine head is marked with a gray circle. (b) The Backbone of a spine includes a string of numbers. The line numbers and return spaces present in **b** were added for clarity and do not exist in the data file structure. Line [1]’s value reports the total number of points that make up the spine. The values from line [2] through [21] make up each of the four spine coordinates (x, y, z, and d). The first point (x = line [2], y = line [3], z = line [4], and d = line [5]) listed is the insertion point where the spine is located along the tree

data can guide researchers in predicting the function of each loop class based on its morphometric traits.

Like other data elements of the neuromorphological file, vessels can be analyzed at a high level, reporting the network summary, loop count, or the branching within an anatomical structure. By coupling two data types, vessels and contours, we are able to further our understanding of vasculature abnormalities that may occur. For example, reconstructing anatomical regions alongside vasculature networks can allow for the

exploration of network loop density within specific regions of control and treatment samples. More applications of utilizing contours for delineating anatomical regions in two or three dimensions are further detailed in the following section.

Anatomical Structure Delineation

A contour element is a named list of sequentially connected points. Contours are often used to delineate anatomical regions within image data. The TraceAssociation property of the contour stores an IRI that is linked to the anatomy term used for a contour and includes the PID for that term. If this child element is present in the contour, then the contour name and TraceAssociation are both saved to the data file via the API connection with the SciCrunch InterLex Terminology Portal discussed in the above section. The contour name is selected from the anatomical term list recorded in the sparcrunch’s `<atlas>` child element. The contour element in Fig. 10b has the name “Glomerulus” which was selected from the species independent kidney terminology list. The corresponding metadata associating this term to its term list is exemplified in the `<atlas>` element of Fig. 10b.

Contours can be used to delineate 2D structures or they can be grouped to represent a 3D surface or volume. 2D or 3D cell bodies are represented using the contour element or groups of contour elements. All cell body contours have names containing “soma”. 3D cell bodies are traced using multiple contours of the same name at different z positions, outlining the entire z space of the cell body region. These structures can provide anatomical context or they can be used in data analysis. The density of a reconstructed neuronal network within an anatomical region can be analyzed based on the volume of the region’s contours.

In a recent study by Achanta et al., researchers utilized the neuromorphological file format to delineate regions of the rat heart and mark the location of neurons of the intrinsic cardiac nervous system (ICN). The contour element was used extensively to map the cardiac structures in two dimensions (Fig. 11a) – such as the endocardium of left and right atria, auricles, and ventricles – within the histological sample to create a 3D reconstruction of the heart (Fig. 11b; Achanta et al., 2020). The annotation methodology was used again by Leung et al. to compare ICN neuron location and distribution across multiple male and female rat hearts (Leung et al., 2020).

Contour-style segmentation can be performed in other software applications and widgets, such as ImageJ’s ROI Manager Tool, but the output file format (e.g., ROI file) often lacks data provenance, structural organization, and is unreadable to humans outside of the ImageJ application. The contour element within the neuromorphological file format addresses these concerns with clear data lineage, concise and human-readable detail regarding the point makeup of each contour as

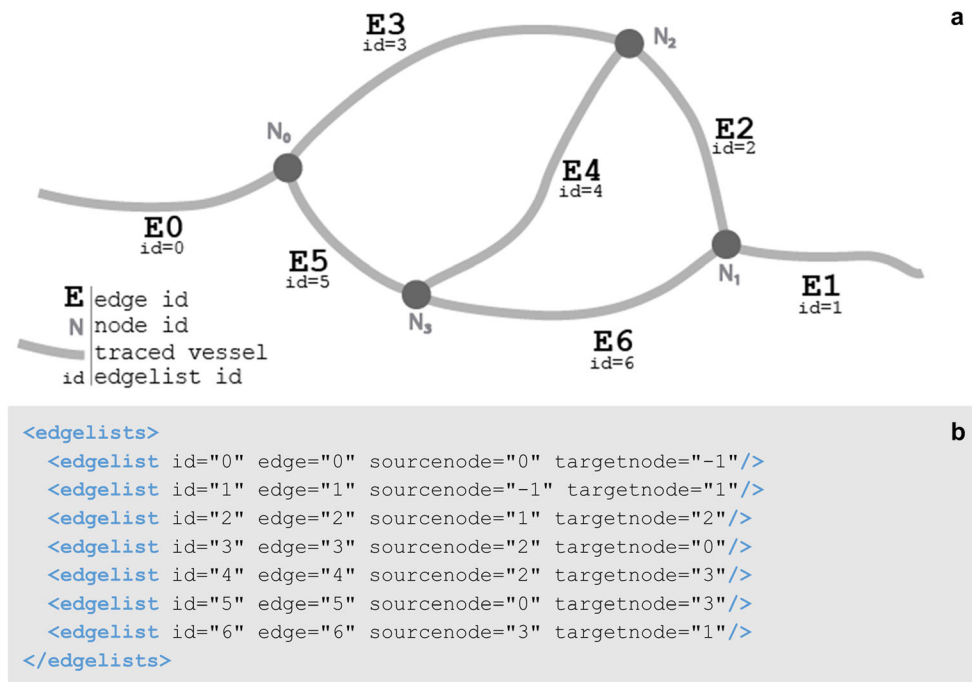


Fig. 9 (a) A diagram of the edgelist element of **b**. Each edgelist and edge id correspond to one of the vessel branches. These are labeled appropriately. The edgelist sourcenode and targetnode inform the start and endpoint of the vessel branch or edge. For example, edge="4" (E4) begins at node 2 (N_2) and ends at node 3 (N_3). This connection is indicated in edgelist id=4 (see **b**). This connection of the vessel back onto itself creates a loop structure. (b) The data structure for the edgelist child

element of a vessel. Each edgelist id attribute corresponds to the edgelist ids in **a**, informing how the vessel edge elements connect to the node elements. The edge attributes correspond to the edge ids in **a**. The sourcenode and targetnode values refer to a node id in **a**. If either the sourcenode or target node values equal -1, this means that there is no starting or ending node

it relates to the structure within the microscopy image(s), and is packaged in a single file for further ease of reuse and interoperability.

Point Markers

The marker element is used to represent single point locations in the data file. This is useful for mapping and/or counting cell types (Zaborszky et al., 2015), synapses (Le Bé et al., 2007), and other punctuate objects. Additionally, they can be used to mark fiducial points or anatomically significant single coordinate locations, as seen in Fig. 12, to assist the mathematical registration of segmentation data to a generic 3D scaffold model. The most recent addition to the data file permits marker naming via SciCrunch terminology lists, including the corresponding PID in the TraceAssociation element.

Since the marker element can be used to represent a variety of biological and non-biological structures, there are a myriad of ways to classify the markers within the data file, allowing others to properly decipher a data file with little inquisition. These attributes include name, color, and shape, all of which can be applied to multiple or individual marker elements.

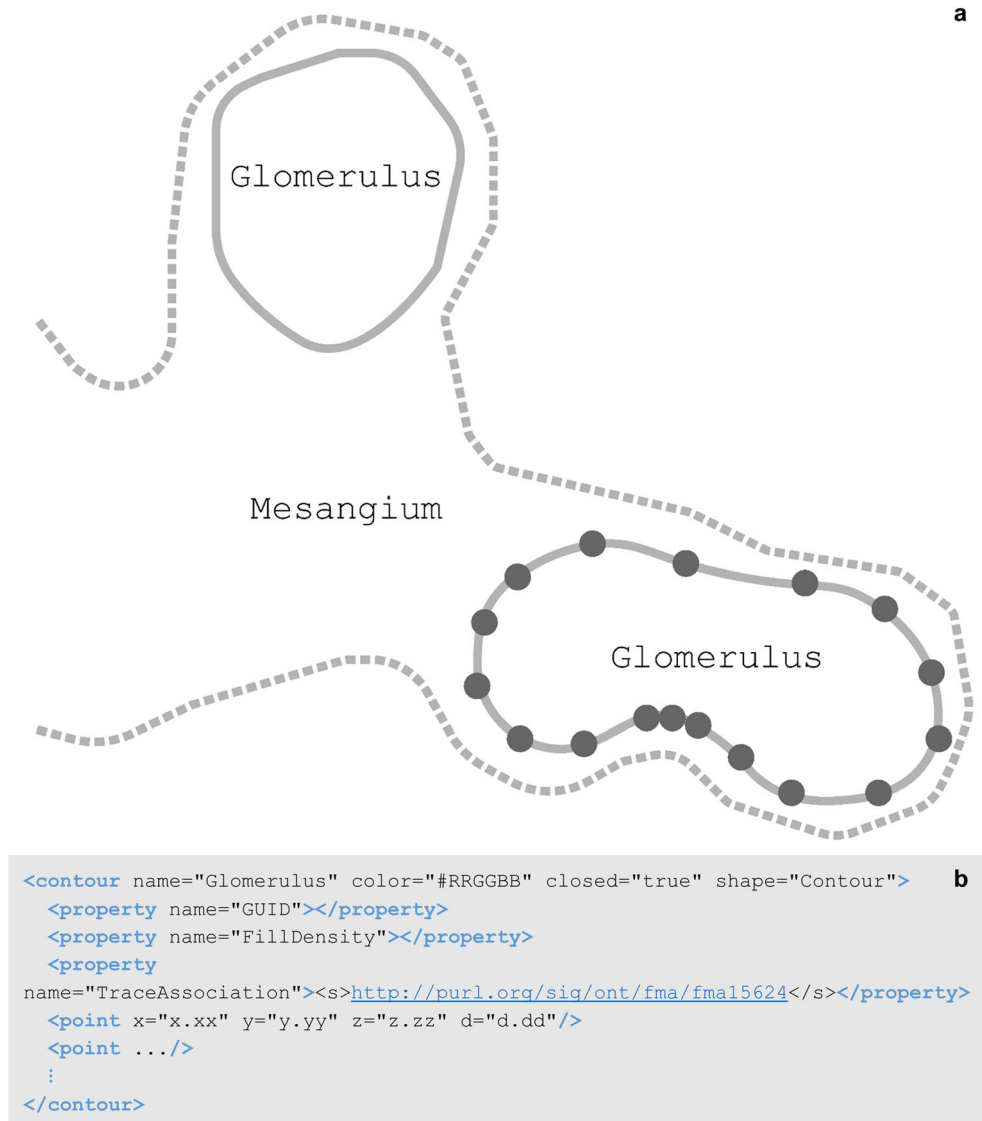
One application of marking fiducials for registration to the 3D organ scaffolds engineered by the Auckland Bioengineering Institute (ABI; Leung et al., 2020). These

scaffolds allow researchers to compare segmentation data from multiple subjects by matching fiducial points in each segmentation file to those same fiducial points within a generic organ scaffold. Mapping segmented data to a common space can help correct for sample deformation due to experimental imaging protocols. By fitting the samples to the generic organ scaffold via fiduciary point matching, objects of interest can be compared to each other with true anatomical context.

Sets

Set properties can be found in any trace data element. It is utilized for naming and grouping data of one or many trace data elements. These elements can either be the same type (e.g., all tree elements) or different types (e.g., spine, tree, and contour elements). By placing traced elements into a set, researchers can embed their knowledge and expertise on the sample into the traced data file, bridging gaps between those reusing and repurposing the data. Applications for the set property include: defining relationships to anatomies, annotating anatomies consisting of more than one traced data element, associating traced elements and supplemental data, and grouping objects to select and edit all at once. For example, neurons can be grouped together with a set name "intraganglionic laminar endings" to represent sensory

Fig. 10 (a) A schematic of contoured regions of a renal corpuscle. The marked point locations (represented as circles on one glomerulus contour) are connected with a line (solid, dashed) to generate an area that represents an anatomical region in two-dimensions. The dashed line represents the mesangium region where the solid lines represent glomeruli. The glomeruli contours are closed where the mesangium is an open contour indicating the structure continues and the contour represents the layer of the mesangium that falls within the renal corpuscle. **(b)** A Glomerulus contour element, child elements, attributes, and values as they appear in the segmentation data file. This contour is a closed contour indicating the first and last point elements are connected. In this **b**, the `<property>` child elements exclude all values for concision with the exception of the TraceAssociation property. The value of the TraceAssociation property is the IRI to the Glomerulus term in the FMA kidney ontology term lists. The point elements have been abbreviated in this **b**. A contour usually contains a list of many point elements, connected in the order they are listed in the contour

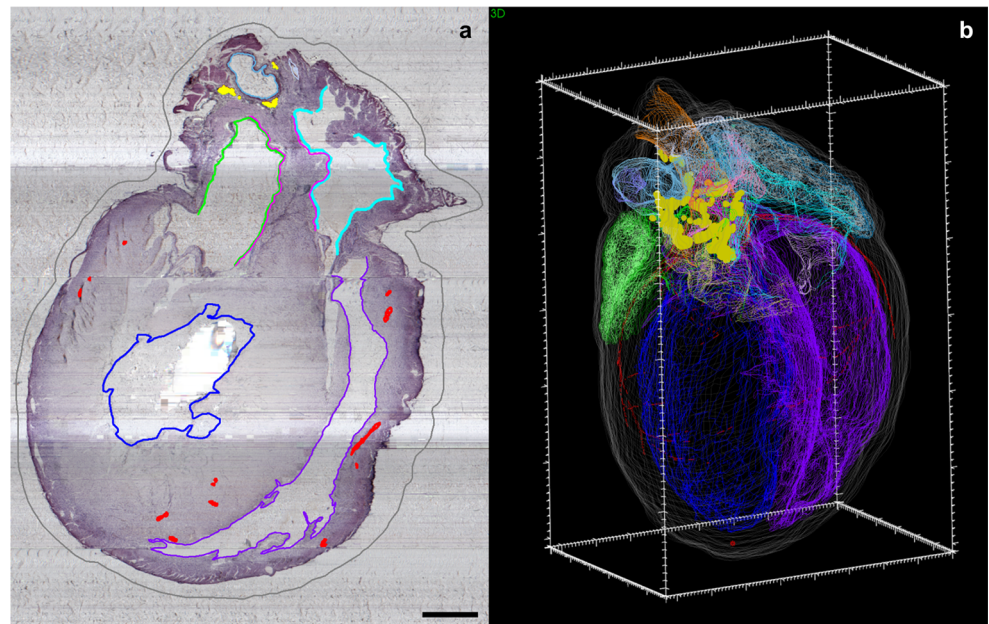


endings within the colon, a key neuronal structure for sensing muscle stretching. This annotation is carried with the segmentation data and can be incorporated in downstream applications and investigations. Elements can also be associated with multiple set properties. This is one of the set property's most useful abilities because annotations about an overarching structure or relationships to other data modalities can be made while still including anatomical context for those structures. Figure 13c demonstrates one use case for associating multiple sets with one data element. Note the multiple set properties indicating this axon is associated with the electrophysiology readings number 18105039-091 and the axon innervation. Those looking to repurpose this data file can identify which of the dataset's electrophysiology files pair with this axon. The anatomical context of the axon within the greater structure of the stellate ganglion is apparent through the set property denoting the path of the fiber.

The data collected by Cho et al. demonstrates the importance of the set property in communicating additional relationships through the neuromorphological file format. Each backfilled cell from the stellate ganglion was reconstructed (Fig. 13a; Cho et al. Data set in progress) and placed into a set named with the electrophysiology identifier (Fig. 13b). The axon innervation was labeled for each cell indicating the nerve fiber that the axon passes through as it exits the stellate ganglion. Encapsulating this otherwise autonomous anatomical context known only by original researchers promotes easy and constructive reusability of this file format's modeled morphologies.

Not only can sets be used to link alternative data types to morphological reconstructions, they can also be used to associate anatomical terminologies to structures that include a variety of data types. For example, a set grouping cell body contours, dendrites, trees, and spines could be named as a specific cell type. The flexibility of the set data type permits multiples sets

Fig. 11 The anatomy of the male Fischer rat heart, including the intrinsic cardiac nervous system neurons (yellow), were mapped using MBF Bioscience’s TissueMapper application **(a)** to create a comprehensive 3D reconstruction **(b)**. **(a)** The scale bar is equal to 1000 micrometers. **(b)** The 3D scale bar’s minor ticks are equal to 100 micrometers on the x-axis, 200 micrometers on the y-axis, and 100 micrometers on the z-axis



for any given element, meaning functional data types as described above can be coupled with globally unique PIDs. This file-level metadata makes it easy for the reconstructions to be

repurposed and opens the doors for quantitative data collation of cellular morphologies that can help further scientific discovery.

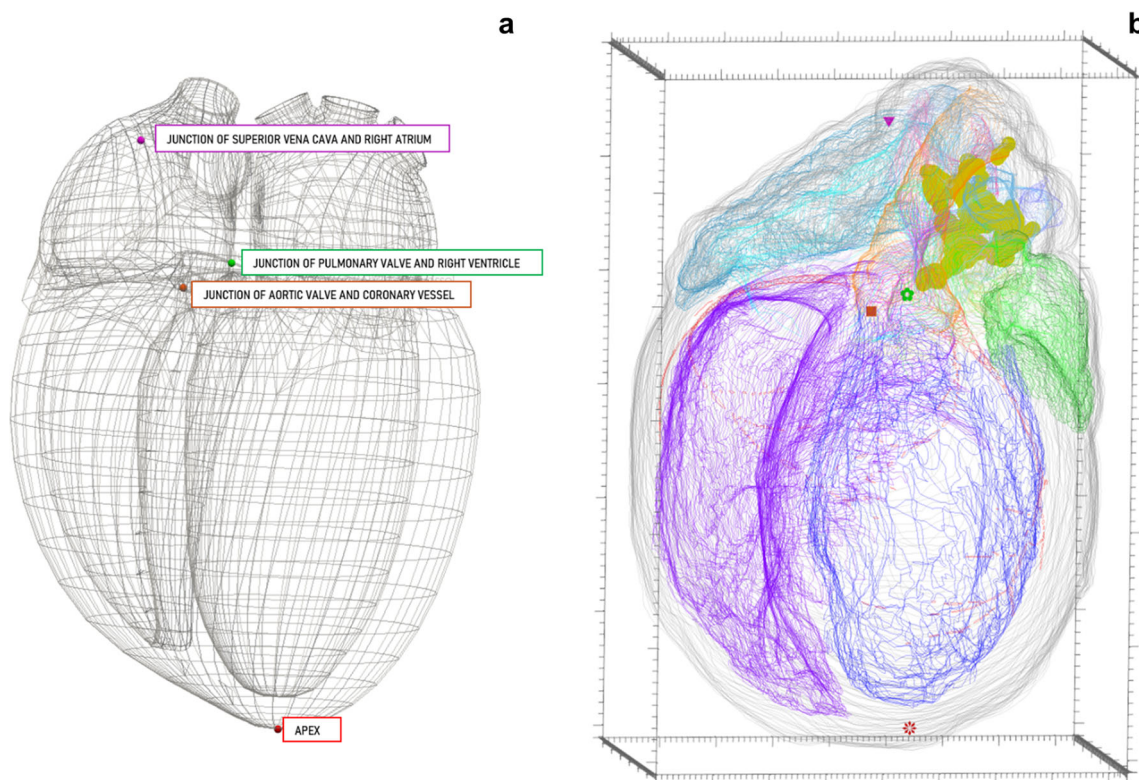


Fig. 12 A generic heart scaffold **(a)** beside segmentation **(b)** in a multi-viewport 3D environment to identify and mark concordant fiducial points essential for registration with the common coordinate scaffold. The triangle marker in **b** represents the discrete location of the junction of the superior vena cava and the right atrium. This location is also marked in **a** in the matching marker color and alongside the associated marker name.

The following marker pairs follow the same format as described above for the triangle marker: flower marker **(b)**= junction of the pulmonary valve and the right ventricle **(a)**, square marker **(b)**= Junction of aortic valve and coronary vessel **(a)**, star marker **(b)**= Apex **(a)**. **(b)** The 3D scale bar’s minor ticks are equal to 100 micrometers on the x-axis, 200 micrometers on the y-axis, and 100 micrometers on the z-axis

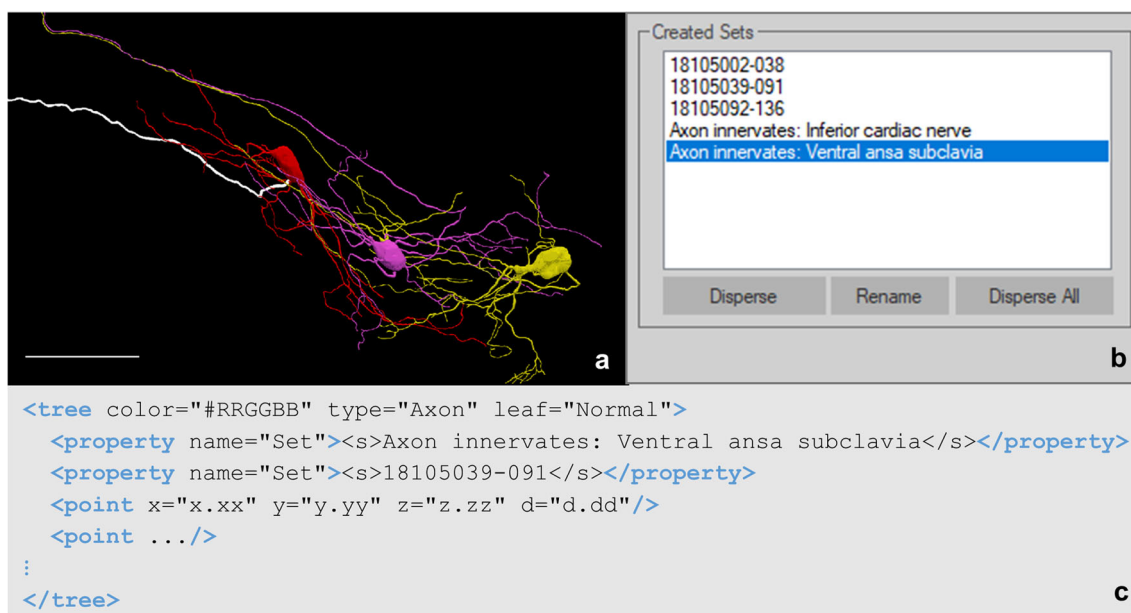


Fig. 13 (a) Neuron reconstructions generated with NeuroLucida 360 based on the image described in Fig. 5a. Highlighted in white is the axon for cell 18105039-091. The cell ID corresponds to the electrophysiology readings taken for each cell backfilled with Neurobiotin. The scale bar is equal to 100 micrometers (Cho et al. Data set in progress). (b) The names of all created set in the tracing shown in a.

The highlighted set, Axon innervates: Ventral ansa subclavia, describes which nerve of the stellate ganglion cell 18105039-091 innervates. (c) The tree element of the axon for cell 18105039-091. The point elements in this tree have been abbreviated using an ellipsis to draw focus to the structure of all created sets for this axon

Discussion

The neuromorphological segmentation file structure has evolved for over 30 years. It is used in software applications such as NeuroLucida and NeuroLucida 360, which has become prevalent systems for neuron reconstruction, with more than 6,500 citations. The file format is widely utilized in projects including the Human Brain Project, the Blue Brain Project, and the NIH SPARC program. Currently, the Blue Brain Portal hosts over 1000 NeuroLucida neuron reconstruction data files, and the Human Brain Project's EBRAINS Knowledge Graph has accumulated around 100 NeuroLucida files to date (EPLF Blue Brain Project, n.d.; Human Brain Project, n.d.). Additionally, other software e.g., NEURON, a simulation environment for neurons and networks can utilize NeuroLucida data files for computational modeling applications. Publication of the file format specification will increase the utility and sustainability of neuromorphological data across research fields and continue to support data-driven science while aiding in data management.

The structural elements of the neuromorphological file format were influenced by input and feedback from leading neuroscientists to propel scientific discovery using quantifiable digital reconstructions of biological structures. The structural elements were also designed to generate Findable, Accessible, Interoperable, and Reusable (FAIR) data and metadata. The scientific community continues to move toward an open and collaborative climate. FAIR data standards are being developed

and adopted by funding institutions, further influencing the shift. Another factor driving FAIR data is the growing importance of bioinformatics and database queries due to the surplus of scientific data of all kinds. Due to the growing emphasis on producing FAIR data, we plan to submit the neuromorphological file format as a standard for digitally reconstructed microscopic anatomies referred to as the neuromorphological file standard. The elements of the neuromorphological file were designed to be Findable by humans and computers to allow the file type to be queried by online databases via globally unique, persistent, and up-to-date identifiers. These identifiers are searchable via the SciCrunch InterLex Terminology Portal. With the most recent enhancements to the file format, allowing for anatomical term association for every data element of the neuromorphological data file, researchers can readily conform to FAIR data principles. The file-level information within the neuromorphological files can be leveraged by common fund research projects or initiatives (i.e., HuBMAP) to enhance searching functions within online collectives of experimental data. The human-readable and published file format is Accessible to anyone looking to repurpose this data type (Angstman et al., 2020). Encoded in the well-recognized XML format, the metadata and tracing data's Interoperability is one of the neuromorphological file format's vital attributes (World Wide Web Consortium, 2008). The detailed specification of the file format is open to the public. Its elements have been decided to enhance interoperability with limited human interaction. Because the format is XML and

human-readable, it can be easily viewed and parsed with a variety of software, e.g., MATLAB and Python, extending its outreach to drive scientific research forwards. Tools such as NEURON, the neuronal simulation modeling software, already support NeuroLucida reconstruction data and their development team is currently working to support the neuromorphological data format. By publishing the neuromorphological file specification we hope to expand the utility of this data type and encourage re-use and repurposing of the reconstructions. The metadata conserved in the header elements of the neuromorphological file provides a robust understanding of how the traced data was derived such as the generating software, sample origin, and image scaling. The association of the morphometric modeling data with this metadata provides file-level information on the data's origin. Because this metadata is stored in the same location as the morphology models, the two are never separated ensuring the Reusability of these data types. These characteristics demonstrate that the neuromorphological file format meets the FAIR data principles developed and endorsed by the INCF (Wilkinson et al., 2016). However, there are many additional and unique benefits to each element of the neuromorphological file. These include the thousands of morphometric analyses available for each distinct element or for elements in combination. These relevant advantages of the neuromorphological file format validate the file format as a rich digital reconstruction format for microscopic neuronal anatomies. A broadly accepted, neuronal tracing format, such as the neuromorphological file format, that is open and FAIR can catalyze scientific research (Meijering, 2010; Halavi et al., 2012; Parekh & Ascoli, 2013; Abrams et al., 2021).

Interoperability

As described above, one of the major aspects that makes this data format interoperable is the human and machine-readable XML format and the open file specification. The format allows for the inclusion of broadly accepted, searchable, and trackable terminology from the SciCrunch Registry to describe the morphological features, following FAIR data principles by specifically reporting the URI of the term or resource within the data file (Grethe et al., 2014). It provides knowledge information regarding the morphological reconstruction in an accessible and broadly accepted format. By providing fields that indicate the software used to generate the morphological reconstructions and reporting the appropriate RRID for the software, the neuromorphological file format appropriately cross-references resources that aid in indicating the provenance of the data using unique and globally persistent identifiers. The format of these URIs and RRIIDs are described in detail along with their relationship to the SciCrunch database in the neuromorphological file specification. The SciCrunch API that is used by MBF Bioscience software to add the terms and unique identifiers to

the data file is open to all tool developers and is provided by the FAIR Data Informatics (FDI) Lab (FDI Lab, n.d.).

One proposed application of the neuromorphological data file's file-level metadata is to incorporate it with existing data repositories, e.g., NeuroMorpho.Org. With proper interpretation of the reconstruction data files, metadata fields could be populated automatically, reducing researcher efforts and the potential for human error. Some of the metadata recorded in neuromorphological reconstruction data files align with the existing NeuroMorpho.Org metadata fields including the following: subject, species, gender, age, tissue thickness, reconstruction software (including the application RRID), anatomical region (i.e., organ), anatomy sub-region, cell type, sub-cell type, structural domains (e.g., axons, dendrites, and somas), and attributes (e.g., branch angles, 3D point space, and diameter resolution) (Parekh et al., 2015). Another example includes leveraging the microscopy image metadata associated with the reconstruction to pair these source data elements within a database, promoting the reuse of these rich data files. In Parekh, Armañanzas, and Ascoli's paper, they describe the variation of neuronal morphology reconstruction formats collated on the NeuroMorpho.Org database. Some reconstructions lack completeness in three categories: structural domains, physical integrity, and morphological attributes (Parekh et al., 2015). The reconstruction data format does not impact physical integrity, as no automated curation currently exists to ensure all morphological structures have been reconstructed to completeness. However, all structural domains and morphological attributes discussed can be modeled and stored appropriately within the neuromorphological data file to represent and analyze these microscopic structures (e.g., 3D data that includes contours, cell bodies, spines, varicosities, diameter resolution, and branch angle information; 2015). To ensure the neuromorphological data file is interoperable, it is essential to providing the means for databases like NeuroMorpho.Org to ingest the file-level metadata captured by this format.

Development efforts are ongoing to improve the interoperable nature of the file format, including the creation of an open, accessible converter for the neuromorphological file format. This is further detailed in the [Future Directions](#) section of this manuscript.

Governance

The neuromorphological file format specification is open to the community and can be accessed at www.mbfbioscience.com/filespecification. The file specification will continue to be updated as needed to define added or modified data elements. To facilitate the integration of community contributions, we plan to generate a standard mechanism for enhancements such as documentation corrections, data storage, and feature requests. The basis of the community

contributions will be facilitated through the publicly available web forum generated to provide support for tool developers implementing the file format. We refer interested participants to <https://forums.mbfbioscience.com/>.

Code Availability

Due to intellectual property right restrictions, we cannot provide source code or its documentation for the commercially available NeuroLucida, NeuroLucida 360, Vesselucida 360, and Tissue Mapper software. Free trials are available at <http://www.mbfbioscience.com>.

Future Directions

Currently, reconstructions in the neuromorphological file format can only be generated within MBF Bioscience proprietary software. In an effort to broaden the use of the neuromorphological file format and integrate other reconstruction formats into the open ecosystem, MBF Bioscience plans to develop a tool for reading and writing the neuromorphological format. This tool will be developed in Python and made available to the community via GitHub to allow users to adapt it for their specific needs. We predict that this tool will help to make the neuromorphological file format more accessible to the neuroscience community easing integration with software tools. Tool builders can use the neuromorphological reader/writer to develop tools that: convert alternative digital neuroanatomical reconstruction file formats to the open, and FAIR neuromorphological file format, convert the neuromorphological data files to a format that can be read by their unique software, and/or extract valuable metadata stored within neuromorphological data files. Simple examples of neuromorphological data files will be hosted on the GitHub repository. We plan to support the code for the neuromorphological reader/writer through adequate documentation hosted on the file specification webpage (www.mbfbioscience.com/filespecification) while addressing any user feedback submitted through the file format web forum (<https://forums.mbfbioscience.com/>).

Another goal is to track the number of software platforms that accept the neuromorphological file format along with the number of users of this data type. The SPARC Portal (<https://sparc.sicence>) will be a good starting point for collecting these metrics because of the required level of curation for all morphological reconstructions. The project intends to track dataset downloads and page visits, making the desired use metrics accessible.

As previously mentioned, the development of the file structures modeled within the neuromorphological format evolves alongside the advances of microscopes, computer systems, and experimental design. It is realistic to predict that this process will continue, as adapting to these changes is something MBF

Bioscience is well versed in. We plan to continue to modify the neuromorphological file format as needed and release new versions of the related resources in a coordinated fashion. We also recognize the importance of community contributions to further enhance the neuromorphological file format. A plan is in place to encourage open community input and process submissions for additions and modifications. Contact through email and forums will help provide the necessary checks and balances for the advancement and longevity of this open and FAIR file format. Lastly, MBF Bioscience welcomes collaborations with tool developers to help incorporate related data types into the neuromorphological file format. We hope these additional resources will help enhance the FAIRness of the data type and put digital reconstructions of microscopy anatomies into context, so they can tell their own story.

Information Sharing Statement Due to intellectual property right restrictions, we cannot provide source code or its documentation for the commercially available NeuroLucida, NeuroLucida 360, Vesselucida 360, and Tissue Mapper software. Free trials are available at <http://www.mbfbioscience.com>. Additional resource related to the Neuromorphological File Format that currently exist or that are developed in the future will be made available for non-commercial use on the file specification webpage: www.mbfbioscience.com/filespecification.

Acknowledgements We would like to dedicate this manuscript to the memory of Dr. Edmund Glaser in appreciation for devoting his career of more than four decades to the field of neuroscience. In 1963, he co-invented and patented the Image-Combining Computer Microscope and pioneered the method of quantifying 3D neuronal morphometry. This technology applied computer techniques to neuroanatomical research, permitting scientists to precisely quantify the brain's three-dimensional structure. In 1987, Dr. Glaser co-founded the company MicroBrightField (now MBF Bioscience) with his son, Jacob (Jack) Glaser. Over the ensuing decades, reconstructing neuronal structures was reduced from hours to minutes and measurement precision was able to achieve sub-micron accuracy. Large assemblies of neuronal networks can now be examined in quantitative detail in three dimensions. Dr. Edmund Glaser's accomplishments in the development of computer microscopy and his contributions to neuromorphological reconstruction were unparalleled and essential to the formulation of this file format.

We would also like to recognize Maryann Martone and the SPARC project for accelerating the publication of the neuromorphological file format.

Funding NIH HHS 1OT3OD025349-01

Data Availability The Neuromorphological File Specification (4.0) can be found at www.mbfbioscience.com/filespecification.

Declarations

Conflicts of Interest/Competing Interests MBF Bioscience is a commercial entity, and the authors affiliated with them are company employees.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abrams, M. B., Bjaalie, J. G., Das, S., Egan, G. F., Ghosh, S. S., Goscinski, W. J., Grethe, J. S., Kotaleski, J. H., Wei Ho, E. T., Kennedy, D. N., Lanyon, L. J., Leergaard, T. B., Mayberg, H. S., Milanesi, L., Mouček, R., Poline, J. B., Roy, P. K., Strother, S. C., Tang, T. B., Tiesinga, P., Wachtler, T., Wójcik, D. K., & Martone, M. E. (2021). A standards organization for open and FAIR neuroscience: the International Neuroinformatics Coordinating Facility. *Neuroinform.* <https://doi.org/10.1007/s12021-020-09509-0>.
- Achanta, S., Gorky, J., Leung, C., Moss, A., Robbins, S., Eisenman, L., Chen, J., Tappan, S., Heal, M., Farahani, N., Huffman, T., England, S., Cheng, Z. J., Vadigepalli, R., & Schwaber, J. S. (2020). A comprehensive integrated anatomical and molecular atlas of rat intrinsic cardiac nervous system. *iScience*, *23*(6), 101140. <https://doi.org/10.1016/j.isci.2020.101140>.
- Angstman, P. J., Tappan, S. J., Sullivan, A. E., Thomas, G. C., Rodriguez, A., Hoppes, D. M., Abdul-Karim, M. A., Heal, M. L., Glaser, J.R. (2020). *Neuromorphological File Specification (4.0)*. Neuromorphological File Specification. Retrieved March 23, 2020, from www.mbfbioscience.com/filespecification
- Arellano, J. I., Benavides-Piccione, R., DeFelipe, J., & Yuste, R. (2007). Ultrastructure of dendritic spines: correlation between synaptic and spine morphologies. *Frontiers in Neuroscience*, *1*(1), 131–131. <https://doi.org/10.3389/neuro.01.1.1.010.2007>.
- Ascoli, G. A., Donohue, D. E., & Halavi, M. (2007). NeuroMorpho.Org: a central resource for neuronal morphologies. *Journal of Neuroscience*, *27*(35), 9247–9251. <https://doi.org/10.1523/jneurosci.2055-07.2007>.
- Blackman, A. V., Grabuschni, S., Legenstein, R., & Sjöström, P. J. (2014). A comparison of manual neuronal reconstruction from biocytin histology or 2-photon imaging: morphometry and computer modeling. *Frontiers in Neuroanatomy*, *8*, 65. <https://doi.org/10.3389/finana.2014.00065>.
- Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., & Yergeau, F. (2008). *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. W3C. Retrieved August 8, 2020, from <https://www.w3.org/TR/2008/REC-xml-20081126/>
- Cannon, R. C., Turner, D. A., Pyapali, G. K., & Wheal, H. V. (1998). An on-line archive of reconstructed hippocampal neurons. *Journal of Neuroscience Methods*, *84*(1–2), 49–54. [https://doi.org/10.1016/s0165-0270\(98\)00091-0](https://doi.org/10.1016/s0165-0270(98)00091-0).
- Cho, Y., Tshanani, A., Sullivan, A. E., Tappan, S. J., Ardell, J. L., Shivkumar, K., & Tompkins, J. D. (2020). Atlas of mouse stellate ganglion neurons with axonal projections to the heart. (Version 1) [Data set in progress]. Blackfynn. <https://doi.org/10.26275/atzo-uhlm>.
- Crook, S., Gleeson, P., Howell, F., Svitak, J., & Silver, R. A. (2007). MorphML: level 1 of the NeuroML standards for neuronal morphology data and model specification. *Neuroinformatics*, *5*(2), 96–104. <https://doi.org/10.1007/s12021-007-0003-6>.
- EPLF Blue Brain Project. (n.d.). *Neurons*. EPLF Blue Brain Portal. Retrieved August 6th, 2020, from <https://portal.bluebrain.epfl.ch/resources/models/neurons-2/>
- FDI Lab. (n.d.). *Our Services*. FDI Lab. Retrieved March 24, 2021, from <https://www.fdilab.org/services>
- Földy, C., Malenka, R. C., & Südhof, T. C. (2013). Autism-associated neuropilin-3 mutations commonly disrupt tonic endocannabinoid signaling. *Neuron*, *78*(3), 498–509. <https://doi.org/10.1016/j.neuron.2013.02.036>.
- Gabriele, M. L., Brunso-Bechtold, J. K., & Henkel, C. K. (2000). Plasticity in the development of afferent patterns in the inferior colliculus of the rat after unilateral cochlear ablation. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, *20*(18), 6939–6949. <https://doi.org/10.1523/JNEUROSCI.20-18-06939.2000>.
- Gama Sosa, M. A., De Gasperi, R., Perez Garcia, G. S., Perez, G. M., Searcy, C., Vargas, D., Spencer, A., Janssen, P. L., Tschiffely, A. E., McCarron, R. M., Ache, B., Manoharan, R., Janssen, W. G., Tappan, S. J., Hanson, R. W., Gandy, S., Hof, P. R., Ahlers, S. T., & Elder, G. A. (2019). Low-level blast exposure disrupts gliovascular and neurovascular connections and induces a chronic vascular pathology in rat brain. *Acta Neuropathologica Communications*, *7*(1), 6. <https://doi.org/10.1186/s40478-018-0647-5>.
- Gao, R., Asano, S. M., Upadhyayula, S., Pisarev, I., Milkie, D. E., Liu, T. L., Singh, V., Graves, A., Huynh, G. H., Zhao, Y., Bogovic, J., Colonell, J., Ott, C. M., Zugates, C., Tappan, S., Rodriguez, A., Mosaliganti, K. R., Sheu, S. H., Pasolli, H. A., Pang, S., ... Betzig, E. (2019). Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution. *Science*, *363*(6424), eaau8302. <https://doi.org/10.1126/science.aau8302>.
- Ghosh, S., Larson, S. D., Hefzi, H., Marnoy, Z., Cutforth, T., Dokka, K., & Baldwin, K. K. (2011). Sensory maps in the olfactory cortex defined by long-range viral tracing of single neurons. *Nature*, *472*(7342), 217–220. <https://doi.org/10.1038/nature09945>.
- Glaser, E. M., & Van der loos, H. (1965). A semi-automatic computer-microscope for the analysis of neuronal morphology. *IEEE Transactions on Bio-medical Engineering*, *12*, 22–31.
- Glaser, J. R., & Glaser, E. M. (1990). Neuron imaging with NeuroLucida—a PC-based system for image combining microscopy. *Computerized Medical Imaging and Graphics: the Official Journal of the Computerized Medical Imaging Society*, *14*(5), 307–317. [https://doi.org/10.1016/0895-6111\(90\)90105-k](https://doi.org/10.1016/0895-6111(90)90105-k).
- Grethe, J. S., Bandrowski, A., Banks, D. E., Condit, C., Gupta, A., Larson, S. D., Li, Y., Ozyurt, I. B., Stagg, A. M., Whetzel, P. L., Marenco, L., Miller, P., Wang, R., Shepherd, G. M., & Martone, M. E. (2014). SciCrunch: A cooperative and collaborative data and resource discovery platform for scientific communities. *Frontiers in Neuroinformatics. Conference Abstract: Neuroinformatics 2014*. <https://doi.org/10.3389/conf.fninf.2014.18.00069>.
- Halavi, M., Hamilton, K. A., Parekh, R., & Ascoli, G. A. (2012). Digital reconstructions of neuronal morphology: three decades of research trends. *Frontiers in Neuroscience*, *6*, 49. <https://doi.org/10.3389/finns.2012.00049>.
- He, H. Y., & Cline, H. T. (2011). Diadem X: automated 4 dimensional analysis of morphological data. *Neuroinformatics*, *9*(2–3), 107–112. <https://doi.org/10.1007/s12021-011-9098-x>.
- Henriksen, E. J., Colgin, L. L., Barnes, C. A., Witter, M. P., Moser, M. B., & Moser, E. I. (2010). Spatial representation along the proximodistal axis of CA1. *Neuron*, *68*(1), 127–137. .
- Human Brain Project. (n.d.). *EBRAINS Find Data*. EBRAINS. Retrieved August 6th, 2020, from <https://kg.ebrains.eu/search/>
- Jackson, M. E., & Cauter, L. J. (1997). Evaluation of simplified compartmental models of reconstructed neocortical neurons for use in

- large-scale simulations of biological neural networks. *Brain Research Bulletin*, 44(1), 7–17. [https://doi.org/10.1016/s0361-9230\(96\)00380-2](https://doi.org/10.1016/s0361-9230(96)00380-2).
- Jacobs, B., Schall, M., Prather, M., Kapler, E., Driscoll, L., Baca, S., Jacobs, J., Ford, K., Wainwright, M., & Treml, M. (2001). Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cerebral Cortex*, 11(6), 558–571. <https://doi.org/10.1093/cercor/11.6.558>.
- Lázaro, J., Hertel, M., Sherwood, C. C., Muturi, M., & Dechmann, D. K. N. (2018). Profound seasonal changes in brain size and architecture in the common shrew. *Brain Structure & Function*, 223(6), 3823–2840. <https://doi.org/10.1007/s00429-018-1666-5>.
- Le Bé, J. V., Silberberg, G., Wang, Y., & Markram, H. (2007). Morphological, electrophysiological, and synaptic properties of corticocortical pyramidal cells in the neonatal rat neocortex. *Cerebral Cortex*, 17(9), 2204–2213. <https://doi.org/10.1093/cercor/bhl127>.
- Leung, C., Robbins, S., Vadigepalli, R., Schwaber, J., Heal, M., Tappan, S., Huffman, T., Farahani, N., & Cheng, Z. (2020). Distribution of ICN Neurons in Male and Female 3D Reconstructed Rat Hearts (Version 1) [Data set]. *Blackfynn*. <https://doi.org/10.26275/IVO4-0RZY>.
- MBF Bioscience. (2020). Neuron Summary (branched structure). Analysis results. Neuron Summary and Cell Body Details. https://www.mbfioscience.com/help/neurolucida_explorer/Content/Analyze/BranchedStructure/neuronSumm.htm.
- Meijering, E. (2010). Neuron tracing in perspective. *Cytometry. Part A: the Journal of the International Society for Analytical Cytology*, 77(7), 693–704. <https://doi.org/10.1002/cyto.a.20895>.
- Mpodozis, J., Cox, K., Shimizu, T., Bischof, H. J., Woodson, W., & Karten, H. J. (1996). GABAergic inputs to the nucleus rotundus (pulvinar inferior) of the pigeon (*Columba livia*). *The Journal of Comparative Neurology*, 374(2), 204–222. [https://doi.org/10.1002/\(SICI\)1096-9861\(19961014\)374:2<204::AID-CNE4>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1096-9861(19961014)374:2<204::AID-CNE4>3.0.CO;2-6).
- Nanda, S., Chen, H., Das, R., Bhattacharjee, S., Cuntz, H., Torben-Nielsen, B., Peng, H., Cox, D. N., De Schutter, E., & Ascoli, G. A. (2018). Design and implementation of multi-signal and time-varying neural reconstructions. *Scientific Data*, 5, 170207. <https://doi.org/10.1038/sdata.2017.207>.
- Nedelescu, H., Abdelhack, M., & Pritchard, A. T. (2018). Regional differences in Purkinje cell morphology in the cerebellar vermis of male mice. *Journal of Neuroscience Research*, 96(9), 1476–1489. <https://doi.org/10.1002/jnr.24206>.
- NeuroMorpho.Org. (n.d.). *Metadata*. NeuroMorpho.Org. Retrieved March 9th, 2021, from <http://neuromorpho.org/MetaData.jsp>
- Parekh, R., Armañanzas, R., & Ascoli, G. A. (2015). The importance of metadata to assess information content in digital reconstructions of neuronal morphology. *Cell and Tissue Research*, 360(1), 121–127. <https://doi.org/10.1007/s00441-014-2103-6>.
- Parekh, R., & Ascoli, G. A. (2013). Neuronal morphology goes digital: a research hub for cellular and system neuroscience. *Neuron*, 77(6), 1017–1038. <https://doi.org/10.1016/j.neuron.2013.03.008>.
- Pillai, A., de Jong, D., Kanatsu, S., Krugers, H., Knapman, A., Heinzmann, J., Holsboer, F., Landgraf, R., Joëls, M., & Touma, C. (2012). Dendritic morphology of hippocampal and amygdalar neurons in adolescent mice is resilient to genetic differences in stress reactivity. *PLoS One*, 7(6), e38971. <https://doi.org/10.1371/journal.pone.0038971>.
- Prusky, G., & Arjannikova, T. (1999). Intracellular filling and reconstruction of identified neurons in fixed rat brain slices. *Brain Research. Brain Research Protocols*, 3(3), 313–319. [https://doi.org/10.1016/s1385-299x\(98\)00055-5](https://doi.org/10.1016/s1385-299x(98)00055-5).
- Rance, N. E., McMullen, N. T., Smialek, J. E., Price, D. L., & Young, W. S. III. (1990). Postmenopausal hypertrophy of neurons expressing the estrogen receptor gene in the human hypothalamus. *The Journal of Clinical Endocrinology and Metabolism*, 71(1), 79–85. <https://doi.org/10.1210/jcem-71-1-79>.
- Rodriguez, A., Ehlenberger, D. B., Dickstein, D. L., Hof, P. R., & Wearne, S. L. (2008). Automated three-dimensional detection and shape classification of dendritic spines from fluorescence microscopy images. *PLoS One*, 3(4), e1997. <https://doi.org/10.1371/journal.pone.0001997>.
- Rübel, O., Tritt, A., Dichter, B., Braun, T., Cain, N., Clack, N., Davidson, T. J., Dougherty, M., Fillion-Robin, J., Graddis, N., Grauer, M., Kiggins, T. J., Niu, L., Ozturk, Schroeder, W., Soltész, I., Sommer, F. T., Svoboda, K., Lydia, N., Frank, L. M., & Bouchard, K. (2019). NWB:N 2.0: An accessible data standard for neurophysiology. *bioRxiv*, 523035. <https://doi.org/10.1101/523035>.
- Schiller, J., Schiller, Y., Stuart, G., & Sakmann, B. (1997). Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. *Journal of Physiology*, 505(3), 605–616.
- Siek, J., Lee, L., & Lumsdaine, A. (2001). *Review of Elementary Graph Theory*. Boost C++ Libraries. Retrieved July 10, 2020, from https://www.boost.org/doc/libs/1_73_0/libs/graph/doc/graph_theory_review.html
- Turner, D. A., Li, X. G., Pyapali, G. K., Ylinen, A., & Buzsáki, G. (1995). Morphometric and electrical properties of reconstructed hippocampal CA3 neurons recorded in vivo. *The Journal of Comparative Neurology*, 356(4), 580–594. <https://doi.org/10.1002/cne.903560408>.
- Ullah, F., Asgarov, R., Venigalla, M., Liang, H., Niedermayer, G., Münch, G., & Gyengesi, E. (2020). Effects of a solid lipid curcumin particle formulation on chronic activation of microglia and astroglia in the GFAP-IL6 mouse model. *Scientific Reports*, 10(1), 2365. <https://doi.org/10.1038/s41598-020-58838-2>.
- Usher, W., Klacansky, P., Federer, F., Bremer, P. T., Knoll, A., Yarch, J., Angelucci, A., & Pascucci, V. (2018). A virtual reality visualization tool for neuron tracing. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), 994–1003. <https://doi.org/10.1109/TVCG.2017.2744079>.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., Hoen, P. A. C., 't, Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, P. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>.
- Wong, A. M., Wang, J. W., & Axel, R. (2002). Spatial representation of the glomerular map in the *Drosophila* protocerebrum. *Cell*, 109(2), 229–241. [https://doi.org/10.1016/s0092-8674\(02\)00707-9](https://doi.org/10.1016/s0092-8674(02)00707-9).
- Wu, C. C., & Karten, H. J. (1998). The thalamo-hyperstriatal system is established by the time of hatching in chicks (*Gallus gallus*): a cholera toxin B subunit study. *Visual Neuroscience*, 15(2), 349–358. <https://doi.org/10.1017/s0952523898152148>.
- Zaborszky, L., Csordas, A., Mosca, K., Kim, J., Gielow, M. R., Vadasz, C., & Nadasdy, Z. (2015). Neurons in the basal forebrain project to the cortex in a complex topographic organization that reflects corticocortical connectivity patterns: an experimental study based on retrograde tracing and 3D reconstruction. *Cerebral Cortex*, 25(1), 118–137. <https://doi.org/10.1093/cercor/bht210>.