Zuse Institute Berlin

Takustr. 7
14195 Berlin
Germany

SUMIT K. VOHRA[1], PHILIPP HARTH[2], YASUKO ISOE[3], ARMIN BAHL[4], HALEH FOTOWAT[5], FLORIAN ENGERT[6], HANS-CHRISTIAN HEGE[7], DANIEL BAUM[8]

# A Visual Interface for Exploring Hypotheses about Neural Circuits[9]

[1] 0000-0003-2443-7461
[2] 0000-0001-8831-3893
[3] 0000-0002-1391-7921
[4] 0000-0001-7591-5860
[5] 0000-0003-0372-4912
[6] 0000-0001-8169-2990
[7] 0000-0002-6574-0988
[8] 0000-0003-1550-7245
[9] Accepted on: February 04, 2023, IEEE Transactions on Visualization and Computer Graphics

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany

Telephone: +49 30 84185-0
Telefax: +49 30 84185-125

E-mail: bibliothek@zib.de
URL: http://www.zib.de

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

# A Visual Interface for Exploring Hypotheses about Neural Circuits

Sumit K. Vohra [1]*, Philipp Harth [1], Yasuko Isoe [2], Armin Bahl [3], Haleh Fotowat [2]
Florian Engert [2], Hans-Christian Hege [1] and Daniel Baum [1]

[1]Department of Visual and Data-Centric Computing, Zuse Institute Berlin (ZIB), Germany
[2]Department of Molecular and Cellular Biology, Harvard University, USA
[3]Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Germany

**Abstract**

*One of the fundamental problems in neurobiological research is to understand how neural circuits generate behaviors in response to sensory stimuli. Elucidating such neural circuits requires anatomical and functional information about the neurons that are active during the processing of the sensory information and generation of the respective response, as well as an identification of the connections between these neurons. With modern imaging techniques, both morphological properties of individual neurons as well as functional information related to sensory processing, information integration and behavior can be obtained. Given the resulting information, neurobiologists are faced with the task of identifying the anatomical structures down to individual neurons that are linked to the studied behavior and the processing of the respective sensory stimuli. Here, we present a novel interactive tool that assists neurobiologists in the aforementioned task by allowing them to extract hypothetical neural circuits constrained by anatomical and functional data. Our approach is based on two types of structural data: brain regions that are anatomically or functionally defined, and morphologies of individual neurons. Both types of structural data are interlinked and augmented with additional information. The presented tool allows the expert user to identify neurons using Boolean queries. The interactive formulation of these queries is supported by linked views, using, among other things, two novel 2D abstractions of neural circuits. The approach was validated in two case studies investigating the neural basis of vision-based behavioral responses in zebrafish larvae. Despite this particular application, we believe that the presented tool will be of general interest for exploring hypotheses about neural circuits in other species, genera and taxa.*

## 1 Introduction

A neural circuit is a network consisting of individual neurons responsible for storing and processing information, coordinating motor processes, and controlling other bodily functions. Identifying complete neural circuits related to a certain function is a complex process which requires segmentation of brain regions, reconstruction of individual neurons, ideally including synaptic connections, as well as conducting experiments that activate these circuits and record their activity to obtain functional information [1, 2]. To date, dense synaptic connectivity between neurons can only be obtained using electron microscopy (EM). This, however, requires a huge effort, having to deal with terabytes of data [3]. Recent advances in technology [4] provide an alternative approach: For individual neurons reconstructed with stochastic labeling techniques, less complete information is obtained than with EM-based reconstructions, but substantial information about connectivity is provided that allows researchers to improve their hypotheses about circuitry [5].

In order to relate all collected information to each other, including anatomical substructures and functional information, a 3D atlas [6] can be used. In this paper, we use an anatomical and functional atlas of the larval zebrafish brain [7] constructed from light microscopy data. However, our approach is neither limited to this species nor to such image data as long as all necessary data are provided. Specifically, we need an anatomically defined hierarchical partitioning of the brain into regions, a representative set of reconstructed neurons of different morphology and type connecting these brain regions, and neuron activity data obtained, e.g., in behavioral experiments with $Ca^{2+}$ imaging [8] – in addition to stimulus, gene expression and possibly electrophysiological data.

For such kind of data, we have developed a visual analysis tool that enables neurobiologists to *identify data-compatible, hypothetical neural circuits* responsible for processing specific sensory information and generating certain behavior. Data compatibility here means that the hypothetical neural circuits must be compatible with the connectivity structure of the reconstructed neurons. The typical use case is as follows: with new data at hand, knowledge from the literature and a set of conceptually possible circuit models in mind, neurobiologists use the tool to nar-

*Corresponding author: vohra@zib.de
**Keywords:** Human-centered computing, Visualization systems and tools, Web-based interaction, Scientific visualization.
**Accepted:** February 04, 2023, IEEE Transactions on Visualization and Computer Graphics.

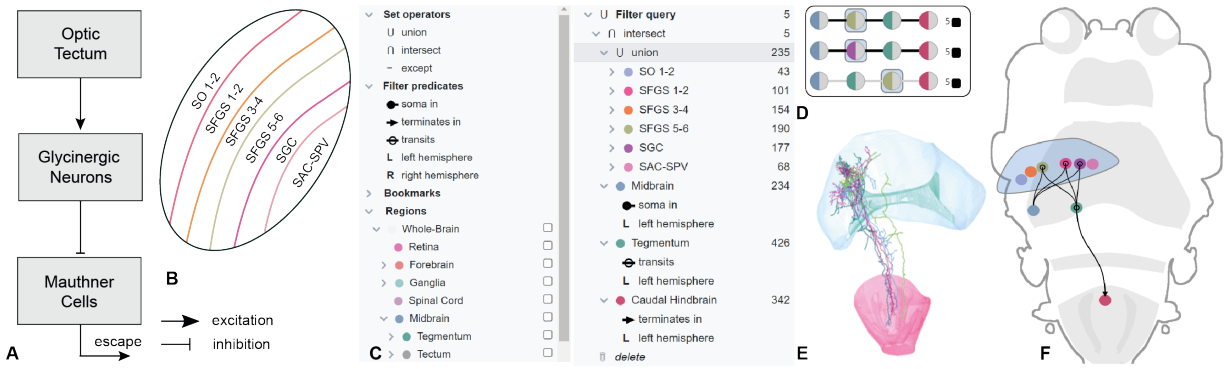


**Figure 1:** ***A****: Graphical illustration of a circuit hypothesis concerning inhibiting escape behavior in larval zebrafish, described in detail in case study 1 (Sect. 7.3).* ***B****: Layers of the optic tectum used in hypothesis 2 of case study 1 (Sect. 7.3).* ***C****: Generation of a query for searching groups of neurons, using the query builder (Sect. 6.1).* ***D****: Listing of possible pathways using the pathway browser (Sect. 6.2).* ***E****: 3D viewer showing the morphology of neurons matching the filter criteria (Sect. 6.3).* ***F****: Visualization of a neuron circuit using the circuit viewer (Sect. 6.3).*

row down the set of candidate models to a smaller subset of data-compatible models. Identifying such hypothetical neuron circuit models is a tedious task. Most of the existing tools provide minimal interactivity or require complex SQL-like query syntax for finding specific subsets of neurons. However, writing SQL queries [9] is a challenge for neurobiologists. The intuitive visualization of neural circuits corresponding to selected subsets of neurons is also very limited in existing tools (e.g., mainly the standard node-link diagram is used [10]). Therefore, in this paper, we present a novel interactive tool based on a structural-functional brain atlas that utilizes interaction and visualization techniques to facilitate interactive identification of specific neuron subsets and exploration of their complex relationships. An application of the tool is illustrated in Fig. 1.
In particular, we make the following contributions:

- A drag-and-drop-based query builder that uses Boolean operations to identify groups of neurons as possible constituents of a neural circuit of interest;
- a linked pathway browser for the visual exploration and refinement of neural pathways defined by the query builder;
- a linked 2D visualization of the constructed neural circuit using a simplified map of the anatomical brain regions facilitating the understanding of the neural circuit;
- a linked 3D anatomy viewer that allows the user to inspect the selected neurons together with the respective brain regions.

Overall, the proposed tool enables neurobiologists to interactively explore, modify, and constrain hypotheses about neural circuits based on empirical data. We demonstrate the applicability and effectiveness of the tool using two real-world case studies (Sect. 7.3 and Sect. 7.4) and a user survey.

## 2 Related work

Several tools that support the visual analysis of neural circuits have been presented. An important distinction arises from the scale at which connectomes (i.e., neural wiring diagrams) are analyzed. Noninvasive imaging techniques in human subjects, such as fMRI and DW-MRI, produce macroscale connectomes, where individual neurons and their branching structure are not resolved at the cellular level [11]. Analysis tools and visualization techniques for macroscale connectomes have been reviewed [12, 13]. At the mesoscale and microscale, where individual neurons, their branching structure, and possibly their synaptic connections are studied, a further distinction can be made on whether the tool mainly supports the reconstruction of neurons and synaptic connections from imaging data or whether the focus is on exploration of a reconstructed network, as is the case in our application. For a survey on visualization in high-resolution connectomics, see Beyer et al. [14]. Pastor et al. [15] propose a unified framework for visual exploration of the brain at all levels of organization, integrating various specific visualization tools and displaying information at different levels of abstraction; see also the cross-scale survey on visualization in connectomics [16]. In the rest of this section, we restrict ourselves to visual analysis tools for connectivity at micro- and mesoscale.

The dataset we used does not contain information on synaptic connectivity or ultrastructural morphological features as available in EM-based reconstructions of neural circuits. Examples for tools in this category are CATMAID [17] for collaborative reconstruction of neural circuits from imaging data, ConnectomeExplorer [18] for the exploration of neurites

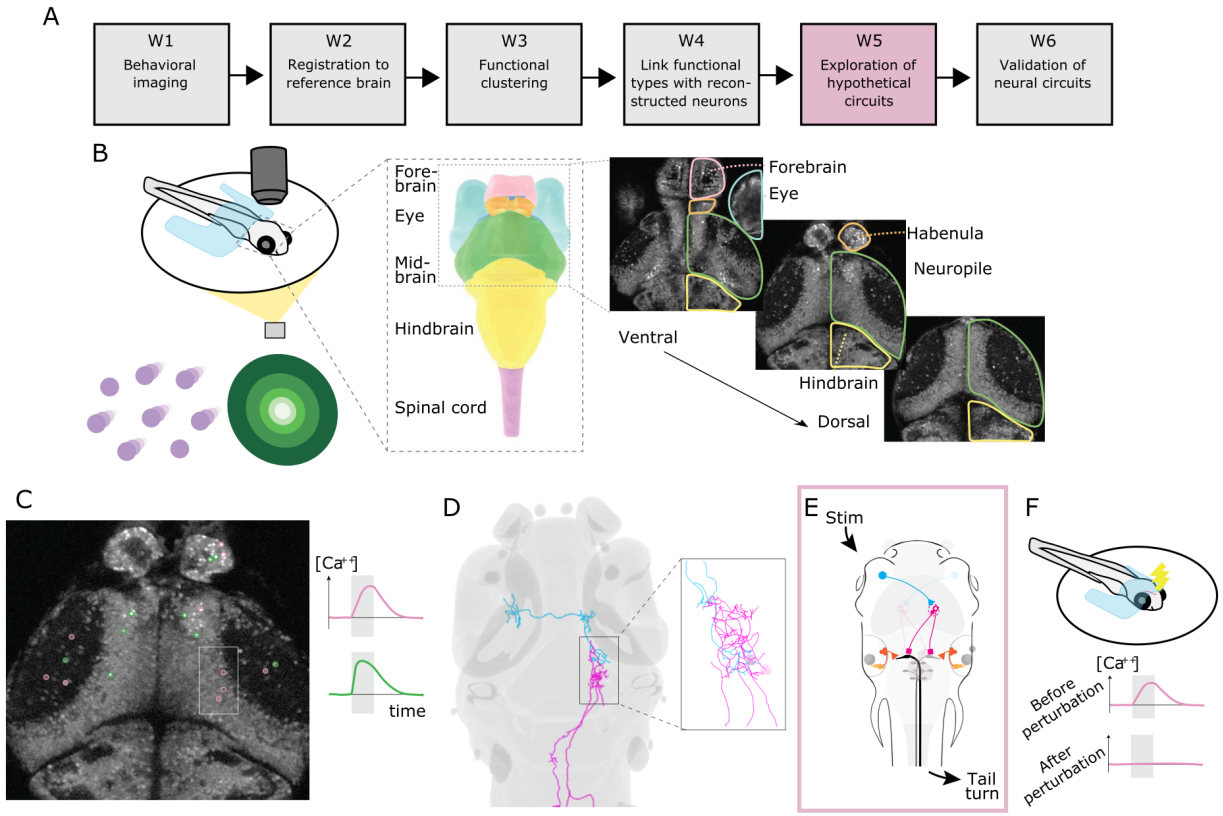


**Figure 2:** ***Overall workflow*** *for uncovering neural circuits – exemplified by circuits controlling movement behavior of larval zebrafish following visual stimuli. For a complete description of the steps, see Sect. 3; note that for step W2 there is no subfigure here.* ***A****: All major working steps.* ***B****: First, a behavior imaging experiment is performed: (left) simplified view of the experimental setup and two types of stimuli: coherent dot motion stimuli (left purple) and looming stimuli (right, green); (mid) schematic drawing of relevant brain regions; (right) the activity of neural cells is recorded with* $Ca^{2+}$ *imaging in multiple z planes.* ***C****: The bodies of active neurons are segmented and functionally characterized based on their activity curves.* ***D****: From a large set of reconstructed neurons, those spatially closest to the active cells are determined – assuming that with some probability that they belong to the same functional type.* ***E****: Based on all the available information and pre-existing hypotheses about neural circuits, data-compatible circuits are constructed and explored.* ***F****: For the most likely hypothetical circuit, validation steps are performed, such as experimental perturbation.*

with their connectional and morphological features using visual queries, and NeuroLines [19], which introduced abstracted representations for the branching topology of neurons. Neuroglancer [20] is a browser-based application that uses WebGL to display 3D imaging volumes, meshes, segmentations, and annotations. Other tools for visualizing connectomes at different scales are the Connectome Viewer Toolkit [21] and brainrender [22]. In our application context, connectivity is defined in terms of projection patterns of neurons (i.e., where their branches terminate and overlap) and which regions are traversed. To retrieve such information, BrainGazer [23] supports interactive path queries by drawing freehand sketches of a suspected path, which is then matched against the dataset. NeuroMAP [24] utilizes abstract circuit diagrams to represent connectivity in the fly brain. Efficient computation and visualization of higher-order overlap beyond pairs of neurons has also been addressed [25, 26]. A tool for visually analyzing the structural organization principles of neural circuits at several levels was presented by Dercksen et al. [27]. Brain network analysis is not restricted to structural connectivity but may involve multimodal data, such as gene co-expression patterns in conjunction with different anatomical parcellations [28], which can be interactively explored and compared in tools like BrainTrawler [10] or NeuroVIISAS [29].

Connectivity analyses are frequently performed in a standardized anatomical model of the species [30], such as the Allen mouse brain atlas [31]. Examples of atlas-based visual analysis systems include the Virtual fly brain browser [32] and NeuronNavigator [33], which allow filtering of neurons by specifying terminal and transit regions. In our case, the underlying anatomical model is a zebrafish brain atlas [7] that offers a hybrid 2D/3D approach for visualization of brain regions, neuron morphologies and neuron somata. A list of comparable resources for zebrafish, such as the Larval Standard Brain [34], is provided by the ZFIN network [35]. Navigation in atlases is often facilitated by means of hierarchical and ontology-

based navigation [36]. A recent development is the integration of neuroinformatic resources and visual analysis tools in web-based platforms, such as Open Source Brain [37], EBRAINS [38], and Gepetto [39]. Specific visualization techniques have also been developed, including the multi-channel Maximum Intensity Difference Accumulation (MIDA) technique in BrainGazer [23], or perceptually effective techniques for visualizing filamentous structures traversing volumetric regions [40]. Finally, Ganglberger et al. [41] presented a Voronoi tessellation and force directed-based approach for 2D layout of 3D brain graphs.

Important components for data analysis are query mechanisms and information visualization techniques for representing abstract information, e.g., about the hierarchical organization of structural units and their connectivity. A good overview of hierarchical visualization techniques is provided by Schulz et al. [42]. More recent surveys have summarized existing techniques for the visualization of multivariate networks [43] and group structures in graphs [44]. Set-based filter queries for neuronal entities have previously been used in tools like ConnectomeExplorer [18]. However, the intuitive visual representation of nested filter queries remains subject of ongoing research (see, e.g., QueryVis [45] or DataPlay [46]). In this respect, an important novelty of the work presented here is the query builder (Sect. 6.1), which provides a tree-like representation of the user-defined filter query, linked to an abstract representation of the filtered neuronal pathways (Sect. 6.2) and semi-abstract anatomical views (Sect. 6.3).

# 3 Overall workflow

Mapping neuronal dynamics and circuit structure across acquisition modalities remains a challenging task in systems neuroscience. The presented workflow, consisting of six major steps (cf. Fig. 2A), must therefore generally be further tailored to the specific problem at hand [47]. We briefly describe all six workflow steps to enable a comprehensive understanding. The focus of our work, however, is on step W5.

**W1: Behavioral experiment and $Ca^{2+}$ imaging.** For zebrafish larvae, imaging of the entire brain is possible. While the animal is exposed to a stimulus, the activity of its neurons is recorded with $Ca^{2+}$ imaging and its behavior, such as the position of its tail, is recorded (Fig. 2B). The result are time series reflecting the stimulus information, neural activities, and behavioral response. Of interest are neurons whose activity correlates with the stimulus or a behavioral output. These neurons are referred to as "active".

**W2: Registration of activity maps to the reference brain.** For each individual, all time steps of the segmented time-dependent activity maps are registered to the brain atlas to relate it to other information.

**W3: Functional clustering and activity types.** For each individual and each time step, the bodies of the active neurons are segmented in the activity maps (Fig. 2C) and activity curves are derived. Then the active cells are functionally characterized and clustered, based on correlations between the activities, the stimulus signals and the output signals. Thereby the main "functional cell types" are identified. Using fluorescence imaging on double transgenic fish lines, the activity type (excitatory vs. inhibitory) can be inferred from gene expression patterns. The derived activity type is then associated with spatially nearby neurons.

**W4: Link functional types with reconstructed neurons.** In the $Ca^{2+}$ images, only the cell bodies of the active neurons are visible; their connectivity is unknown. In the following, the analysis is therefore continued with *proxies*, which originate from a set of reconstructed neurons that is as representative as possible. Their morphology and connectivity is known and stored in a database; and their probable functional type is inferred from spatially close active cells that have been functionally classified.

**W5: Exploration of hypothetical circuits.** In the preceding steps, the following information has been gained: A *set of neurons* that are spatially close to "active cells" that contribute to the observed behavior. About each of these reconstructed neurons, we know (1) its 3D position in the reference atlas, (2) its morphology and putative connectivity, (3) the following brain regions: the one in which its cell body is located, as well as those through which its neurites pass and in which they terminate, (4) its presumed functional type , and (5) (sometimes) also its presumed type of activity (excitatory or inhibitory). The presumed functional type of a neuron gives an indication of the specific subfunction of the neural circuit to which it (more precisely: the nearby active cell for which it is a proxy) is likely to contribute.

Information (4) and (5), together with the restrictions from (2) and (3), provide constraints for circuit models that may underlie the generation of the behavior of interest. This information is *incomplete* (e.g., important neurons may be missing among the active cells) and it is *uncertain* (e.g., the connectivity information and the functional type). This means that a particularly extensive and multi-faceted validation is necessary, after designing a hypothetical circuit. The subworkflow of this working step is depicted in Fig. 4: First, the neurobiologist identifies brain regions of interest w.r.t. functional information. Then she/he generates queries to find in the database of reconstructed neurons all those that satisfy prede-

termined conditions (regions for nucleus, terminals and transit; more complex pathways through multiple regions). Then the neurobiologist selects specific subsets from all neurons contained in the hypothetical, preliminary circuit. She/he also visualizes the selected neurons in the 3D brain atlas. In this way, the biologist can test, modify, and refine his/her hypotheses to finally arrive at the few alternatives that best fit the data.

**W6: Validation of neural circuits.** Circuit validation is an iterative process that requires a series of experiments and analyses in which modified and refined hypotheses are tested and increasingly constrained by data. For this purpose, numerical simulations of the hypothetical circuits as well as further physiological experiments (patch clamping, EM, further $Ca^{2+}$ imaging, modified behavioral experiments and perturbation) are used. In this way, neuroscientists arrive at ever better validated models.

# 4 Biological data

The model organism we are investigating here is larval zebrafish (6 days post fertilization) – a small vertebrate, comprising around 100,000 neurons. It is capable of performing complex behavior like visually guided optomotor and escape maneuvers, hunting, etc. The biological entities (**B1-3**) used in this paper are obtained from several imaging modalities: Confocal microscopy was used to define the brain parcellation (**B1**); whole-brain fluorescence confocal imaging [7] was used to obtain gene expression information (**B1**); $Ca^{2+}$ imaging was used [48] for determining neuronal excitability (**B2**); and stochastic labeling techniques [49] were used to label single neuron morphologies, but without distinguishing axons and dendrites (**B3**).

**B1 Parcellation of the brain and related ontology.** The zebrafish brain is divided into structurally and genetically defined 3D regions resulting from spatial heterogeneity in microscopy images and gene expression maps. Based on the structural information, the zebrafish brain is commonly divided into six main regions: Forebrain, Midbrain, Hindbrain, Retina, Ganglia and Spinal Cord. These major regions were further subdivided into 129 brain subregions [7]. In addition, brain regions were derived from 11 gene expression datasets. They represent regional clusters where specific genes are expressed, e.g., gad1b. For each gene, we obtained multiple clusters. In total, we have 189 such regions, which we call "gene cluster regions". We extracted 3D meshes of the boundaries of these brain regions and stored them in binary format.

**B2 Cell bodies or functional cell bodies.** This data comprises a set of neurons for which information about neurites, i.e. dendrites and axons, are missing. These cell bodies can then be functionally classified as explained in the workflow (Sect. 3, **W3**). We stored the cell body information in a CSV file. Users can also upload their custom CSV files for analysis.

**B3 Neuron morphologies.** A set of reconstructed neuron morphologies, including cell body and neurites, that are required (distinction of axons and dendrites is not necessary) to infer, with some probability, the morphology and connectivity of nearby neurons. In this paper, we use approximately 3,000 neurons reconstructed by Baier and colleagues [4] using light microscopy with fluorescent reporter transgene BGUG (Brn3c:Gal4, UAS:GAP43-GFP) labelling. The morphologies of these reconstructed neurons consist of the cell body and the neurites (dendrites and axons are not distinguished). Bahl et al. [47] registered all 3,000 neurons to the zebrafish reference brain. The neuron morphologies are stored in standard SWC format consisting of a set of connected points. Neurons are extracted on the fly from these files and can be accessed via hierarchical trees.

The minimal data requirements for using the presented tool are a list of brain regions with or without hierarchy and a set of neuron morphologies. Although the circuit generated using minimal data may be incomplete, it already allows researchers to gain information about potential circuits. These in turn can be used to conduct more targeted behavioral experiments and collect further data that provide new hypotheses. The building of hypothetical neural circuits is an iterative refinement process which is supported by the presented tool.

# 5 Visual design

## 5.1 Design considerations

The idea of a tool supporting working step **W5** originated from meetings with neurobiologists who expressed their concerns about lack of analysis tools for exploring neural circuits. As result, the domain experts, consisting of three post-docs and the group leader, all of whom are co-authors of this work, met regularly with the visual researchers over a period of 24 weeks. The domain experts provided the necessary domain knowledge and a list of common analysis tasks that neurobiologists would like to perform semi-automatically. Based on this input, requirements were identified (Sect. 5.2) and a prototype tool was implemented. In the development process, we used the multilevel-typology-of-visual-tasks approach [50] to qualitatively analyze the design process.

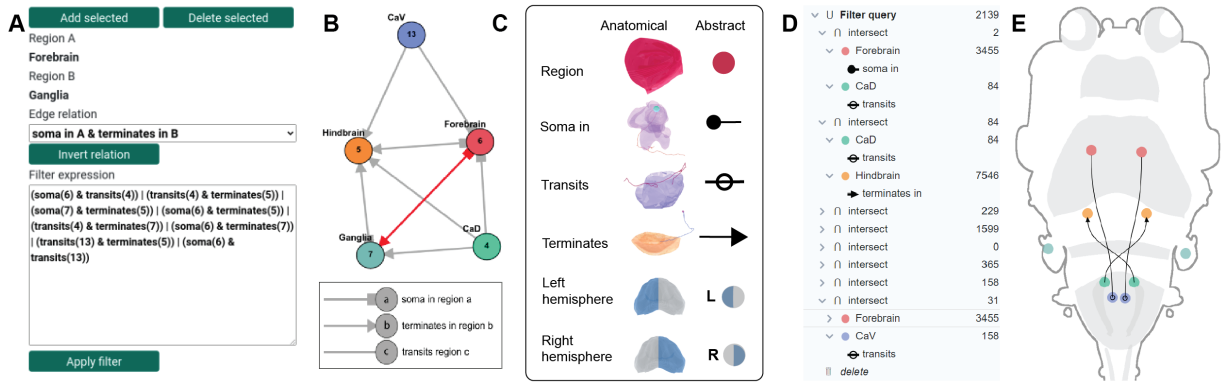


**Figure 3:** ***Design prototype and proposed solution.*** ***A:*** *Initial visual interface for specifying query expressions.* ***B:*** *First prototype for displaying the constructed neural circuit as abstract 2D graph.* ***C:*** *Anatomical units and their abstract representation.* ***D:*** *New query builder (Sect. 6.1) for interactive query building.* ***E:*** *Circuit viewer (Sect. 6.3) for intuitive visualization of neural circuits.*

Our first aim was to map functional information about cell bodies and related data onto hierarchical brain regions and to enable a quick perception of the data. Such a mapping allows neurobiologists to find interesting brain regions which may play a role in a neural circuit. Trees and tree maps represent one common way to support the quick perception of cell body distributions across brain regions. We decided to use a tree map [51] because of its compactness and its ability to scale well with hierarchical regions.

Neurobiologists often use node-link diagrams to represent neural circuits. Here, nodes are generally brain regions or functional units which are specialized in performing certain tasks or portions of tasks. Therefore, our first prototype for exploring neural circuits was a query editor for writing expressions (Fig. 3A) along with an abstract 2d node-link diagram (Fig. 3B). In this prototype, nodes represent brain regions and are visually encoded as circles. An edge between two nodes represents a relation between two regions. However, this approach provided no spatial information to the user and could lead to visual clutter when many pathways exist. Neurobiologists also had difficulties in writing query expressions as this can easily get complex. These limitations motivated us to develop a novel 2d visualization and intuitive ways for specifying queries graphically.

In particular, we needed a semi-abstract anatomical embedding of the node-link diagram, similar to what our domain experts would use when sketching on paper. The resulting circuit viewer (Fig. 6) shows in its default configuration an outline of the zebrafish brain obtained by projecting brain region boundaries and anatomical landmarks (shaded areas) from different imaging layers into a common plane [41]. In addition, we add further visual queues to the background and use spline curves for connecting brain regions. To cope with the large number of possible pathways that can result even from a neuron query with few brain regions, we designed the pathway browser (Fig. 6C,F), which uses simple aggregation glyphs to reduce the number of pathways and contour overlays to highlight selected items.

## 5.2 Task analysis

Prior to developing the tool, our team of visual data analysts and neurobiologists identified the following domain goals, requirements and corresponding tasks.

**Domain goals** The main goal of neurobiologists is to explore hypotheses about neural circuits, in particular their connectivity structure. For workflow step **W5**, which is the focus of this work, two subgoals can be recognized: (1) Identification of subsets of the neurons building hypothetical neural circuits that are in accordance with biological hypotheses; (2) Exploration of selected neurons from (1) and their projection preferences.

An example of (1) is the analysis of neurons involved in inhibiting escape response (Sect. 7.3). An example of (2) is identifying where the axon of a specific neuron connects to the other neurons.

**Domain-driven tasks** To support the above mentioned subgoals, the following tasks were identified:

**T1 Select brain regions.** To identify brain regions of interest, users need to get a quick overview of specific neurons per brain region, especially functional information derived from activity maps of active neurons.

**T2 Select neuron subsets.** Given several selected brain regions, domain users want to identify neurons that are located in certain regions, connect regions, or traverse regions. The task is therefore to develop a simple, intuitive tool that allows users to focus on the data and removes the burden to write complicated query expressions.

**T3 Pathway analysis.** Identifying neuron connec-

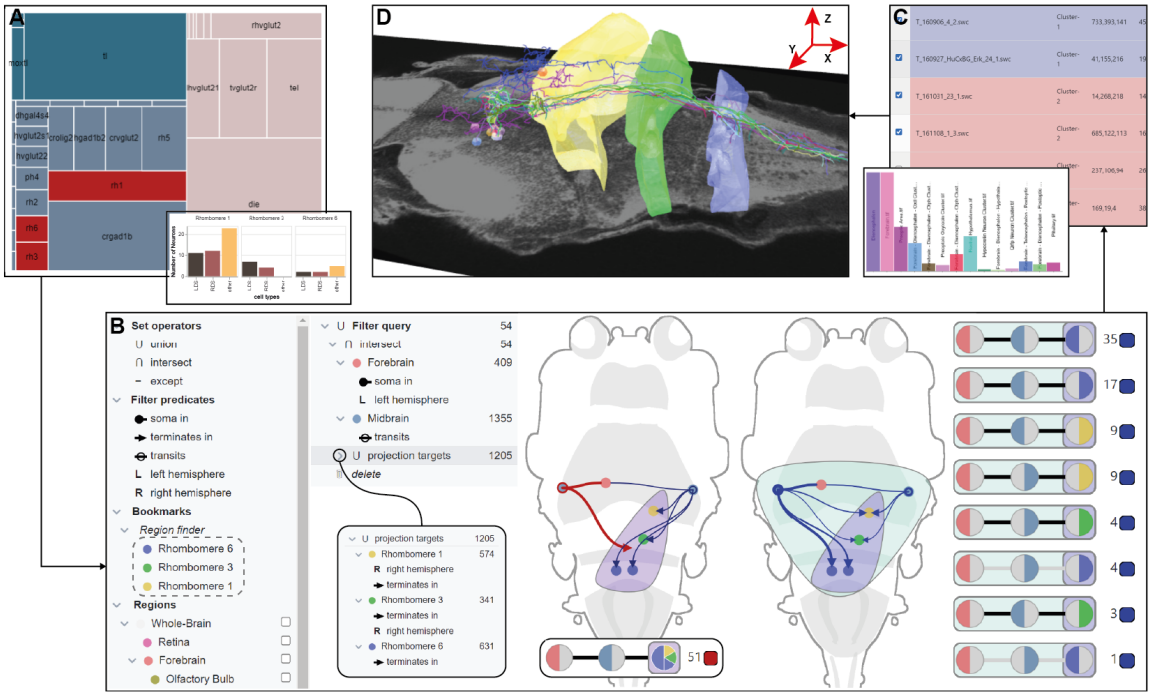


**Figure 4:** ***Exploration of hypothetical circuits*** *(corresponding to working step W5 in the overall workflow, Fig. 2A):* ***A****: Functionally interesting brain regions are identified using the region finder (cf. Sect. 6.1).* ***B*** *left: A hypothetical circuit is generated using the query builder (cf. Sect. 6.1) and (****B*** *mid and right) visualized using the circuit viewer (cf. Sect. 6.3).* ***C****: All neurons contained in the circuit are then passed to the neuron browser (cf. Sect. 6.2).* ***D****: Selected neurons and anatomical context are visualized in the anatomical viewer (cf. Sect. 6.3).*

tions between brain regions requires visualizations that give users a quick overview of the pathways between brain regions. Multiple pathways or parallel neural circuit models might exist that are responsible for different behaviors. Thus, a tool is needed that allows users to explore all possible pathways and narrow them down based on further information including connection strength and expertise.

**T4 Connectivity analysis.** Once brain regions and interesting neurons between these have been identified, it is important to build neural circuits that might be responsible for transforming sensory information into behavioral responses. For example, what connections does the axon of a specific neuron have with dendrites of other neurons? In our case, users will use presumed connectivity information since synapse information is missing. Key for all these tasks is interactive exploration, enabling users to combine the domain expert knowledge with the observed data.

## 5.3 Scalability challenges

Our aim is to build a scalable platform that allows users to explore and navigate reconstructed neuron morphologies, perspectively even for large EM datasets with many neurons. With this aim, we face the following scalability challenges:

**S1 Number of neuron morphologies.** Our current dataset is small, consisting of only 3,000 morphologies reconstructed from light microscopy data. This is about 3% of the total neurons present. Our collaborators are constantly working on segmenting neurons using EM datasets. In the next few years, this data will become available and we will have to deal with approximately 100,000 neurons.

**S2 Number of pathways.** The number of combinatorially possible pathways increases exponentially with the number of brain regions, so that it would quickly become impossible for users to navigate through a list enumerating all such pathways to test their hypotheses.

Our main approach to addressing these scalability challenges, which also guides the design of the visual components discussed in the next section, is as follows:

To ensure that our platform can handle large numbers of neuron morphologies (**S1**) we created synthetic test data and used technical means such as server-side pagination of query results (Sect. 7.1) to verify that our system meets this requirement.

The second challenge (**S2**) is more principled, as the exponential growth in the number of potential pathways cannot be compensated by performance improvements. We address this with a two-staged filter process realized through the *Query builder* (Sect. 6.1) and the *Pathway browser* (Sect. 6.2). The *Query builder* enables the user to define a coarse selection of poten-

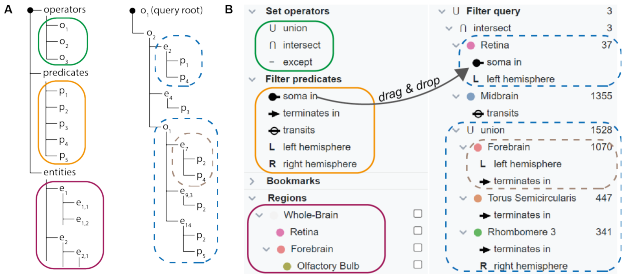


**Figure 5:** ***Query builder.*** ***A****: Set operators, filter predicates, and biological entities (left tree) are used to construct a filter query (right tree).* ***B****: In our application, filter predicates describe spatial attributes of neurons in relation to regions. The user builds a query by dragging operators, spatial predicates and regions into the query tree.*

tial pathways based on regions of interest pertaining to the hypothesis about the neural circuit and various filter predicates. The resulting pathways are then displayed in the *Pathway browser*, sorted by their actual support in the dataset (i.e., the number of neurons compatible with the pathway). The *Pathway browser* enables selection of single pathways for display and features aggregation symbols so that families of similar pathways can be collapsed (Fig. 6E).

# 6 Visual components

The development of the neural circuits explorer and its visual components was motivated by the need to link the given anatomical and morphological data with functional data to support the identification of data-compatible neural circuits. As result, a three-step process was implemented, reflected by the three tool panels: (1) *Query* panel; (2) *Selection* panel; and (3) *Viewer* panel. Each of the panels comprises several tools. The *Query* panel (Sect. 6.1) supports the intuitive, interactive building of a user query that selects a set of neurons and a corresponding set of pathways. This pre-selection of neurons and pathways can be further refined in the *Selection* panel (Sect. 6.2), resulting in a final set of neurons that build or are part of the hypothetical neural circuit we are looking for. This neural circuit can then be visualized in the *Viewer* panel (Sect. 6.3) using different visualization tools.

## 6.1 Query panel

**Query builder** The main component of the *Query* panel is the *Query builder* (**B1-3, T2**). It allows the user to extract subsets of neurons with certain attributes from a neuron database. Technically speaking, it enables the user to create complex database queries interactively by simply dragging items and selecting operators. It consists of two juxtaposed trees as shown in Fig. 4B (left) and in Fig. 5. The left tree contains all components that are needed to build a circuit query, including *Set operators*, *Filter predicates*, a list of *Bookmarks* containing special brain regions, and a subtree with all brain *Regions* in the predefined parcellation. Fig. 3C shows the filter predicates along with an anatomical depiction and the visual abstractions used in the proposed tool. The right tree is a representation of the nested *Filter query* which produces a filtered set of neurons that realize the neural circuit.

The user builds the query by dragging & dropping elements from the left tree into the query tree (Fig. 5B). Leaf nodes in the query tree are regions or filter predicates attached to regions. A region node without predicates returns all neurons that can be associated with the region in both hemispheres. Predicates (e.g., 'soma in' and 'left hemisphere', c.f. Fig. 6A,D) filter the set of neurons returned by a region node. Set operators (e.g., union, intersect) are intermediate nodes (except for the query root node), processing the results of their respective child nodes (Fig. 5). On user changes, the query tree is reevaluated recursively and the numbers of neurons that match the filter query are displayed at each node/subquery. This gives an immediate feedback as to how useful the query is, i.e., whether it can lead to the desired neural circuit and which subqueries return no results. Any modifications of the filter query will be immediately passed on to the other panels, where the possible neuronal pathways and the resulting circuit will be visualized; c.f. Fig. 6B,C,E-G.

Subqueries in the query tree can be collapsed and expanded as needed, i.e. the corresponding visual elements are hidden or visible to the user (**S2**). This is indicated by the highlighted red and green boxes of the filter query shown in Fig. 6A. While the subqueries are expanded in Fig. 6A, they are collapsed in Fig. 6D. Selecting as well as collapsing/expanding of subqueries will also influence the visualization of the pathways and the circuit, as will be described in later.

**Region finder** The *Region finder* can be used to identify brain regions of interest in terms of functional information (**B1-2, T1**). The component consists of several sub-components (Fig. 4A) including a hierarchical treemap and bar/pie chart views. The hierarchical tree map provides an overview of the spatial distribution of how neuron populations are spatially distributed across brain regions defined in the ontological hierarchy. To avoid visual clutter, abbreviated region names are displayed by default; the full name becomes visible when the user hovers over a particular cell of the tree map. Selections can be done on different levels of the anatomical hierarchy. The

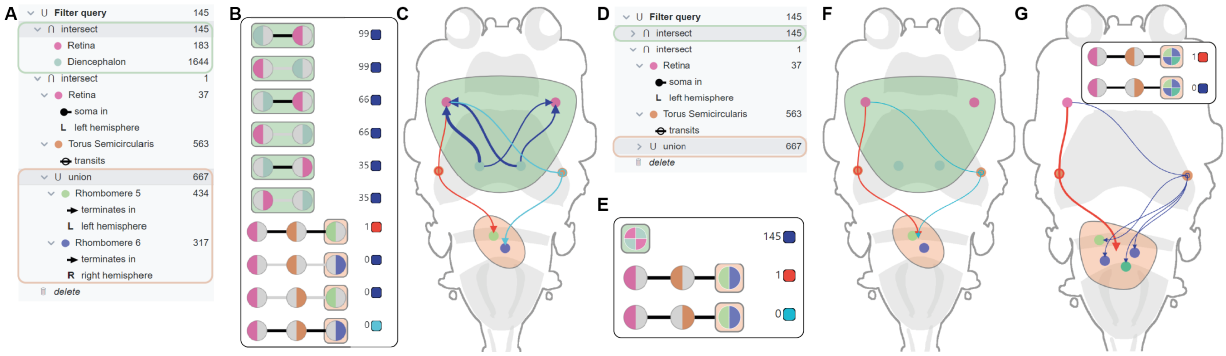


**Figure 6:** *Visual components (Sect. 6). **A**: Query builder (Sect. 6.1), showing the example query; the selected sub-queries are highlighted by green and orange bounding boxes, respectively. **B**: Pathway browser (Sect. 6.2), displaying the list of possible connections realizing the example query. **C**: Circuit viewer (Sect. 6.3), showing the realized neural circuits and regions corresponding to selected sub-queries enclosed by an envelope. The thickness of the blue connections encodes to number of neurons. **D**: Query builder (Sect. 6.1), showing the collapsed sub-queries. **E**: Pathway browser (Sect. 6.2), displaying the list of partially aggregated pathways. **F**: Circuit viewer (Sect. 6.3), showing the pathways as partially aggregated edges. **G**: Circuit viewer (Sect. 6.3), depicting two different selection modes for partially aggregated pathways. In the default mode, the links to all regions in the corresponding set are shown, e.g., pathways in the right hemisphere. Alternatively, an aggregated representation of pathways, e.g., the red line leading to the orange region, is displayed as an aggregated edge.*

selected brain regions can be added as bookmarks to the *Query builder* (Sect.6.1) for further processing. The bar/pie chart views allow the user to visualize and compare the distributions of cell types across selected regions. On hovering over the bars or pies, the exact proportion of the cell type in the brain region is displayed. Bar charts are the default; alternatively, pie charts are available for visualizing a large number of cell types.

## 6.2 Selection panel

**Pathway browser** The neuron *Pathway browser* (Fig. 6C,F,G) shows the list of all possible pathways related to the current query sorted by the number of neurons in the pathways. Examples of pathways together with their generating queries are shown in Fig. 6B,C. Each region is represented as a circle. Left and right portion of the circle are colored based on its anatomical location, i.e., left or right hemisphere. The users can select/deselect all paths, or single paths by clicking on them. The selected edges will be drawn as black lines. Only the selected paths will be shown in the *Circuit viewer* (Sect. 6.3, Fig. 6B).

If a set operation is selected in the hierarchical *Query builder*, paths will be enclosed by an envelope in the *Pathway browser* as well as in the *Circuit viewer*, see, e.g., the green region in Fig. 6B,C. The background color of the collapsed paths has the same color as the polygon in the *Circuit viewer*. If a set operation is collapsed, paths will also be collapsed in both views, see, e.g., green region in Fig. 6B,E. Aggregated regions in the pathways will be shown as colored pies in the pathway browser, preserving the location in the left and right hemisphere (Fig. 6E,G). For example, the first four pathways in Fig. 6C are partially based on the subquery highlighted in green (Fig. 6A). In Fig. 6D, the subquery is collapsed and, hence, the corresponding pathways at the bottom are also partially aggregated (Fig. 6E). When the user selects an aggregated node, by default, the links to all regions in the corresponding set are shown; however, it can also be displayed as an aggregated edge, see, for example, the red line leading to the orange region in Fig. 6H.

**Neuron browser** The *Neuron browser* (Fig. 4C) allows the user to select individual neurons to be displayed in the *2D* and *3D viewers* (Sect.6.3). It contains the list of neurons that match the filter criteria of active edges in the *Pathway browser* and the *Circuit viewer*. In addition, it shows meta-information about each reconstructed neuron, e.g., soma location, morphological type, and spatial extent across brain regions.

## 6.3 Viewer panel

**Circuit viewer** The *Circuit viewer* is central for understanding the generated queries and supports the interactive exploration as well as the refinement of the query in many different ways (**B1-3, T3**). It provides an intuitive visualization of the built neural circuit, see, e.g., Fig 6C,F,G.

The map of this view is based on the real anatomy of a zebrafish larva. In addition to the outlines of the fish larva, some higher-level regions are highlighted in gray to give the user more precise positional information. The component is generic and any other

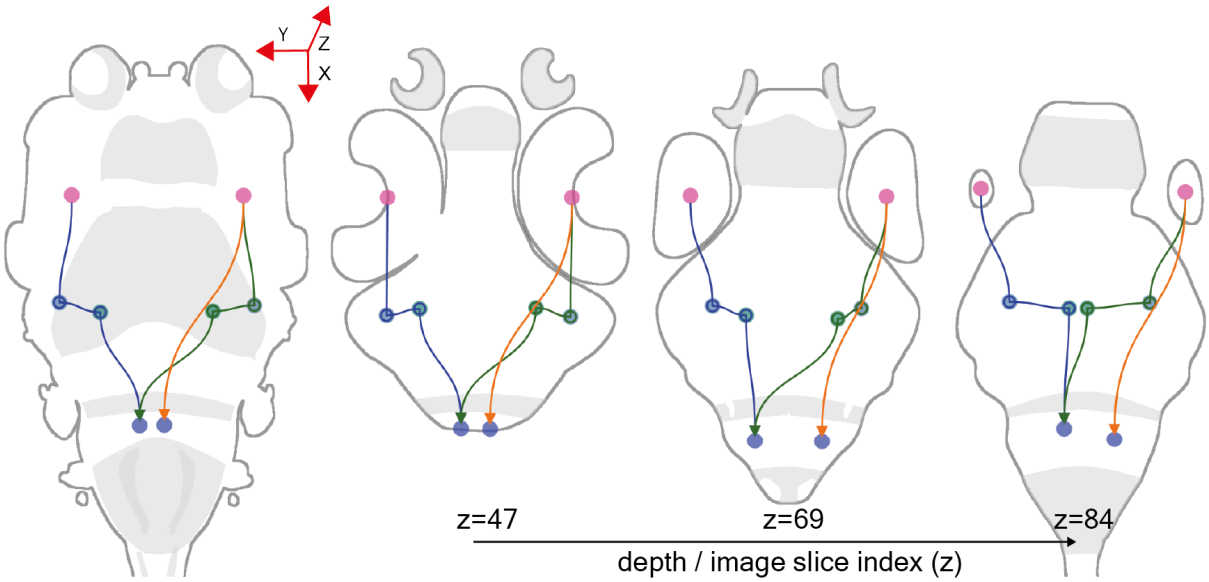


**Figure 7:** ***Slice view:*** *Functionality of the slice view (Sect. 6.3) is to give neurobiologists an idea how the locations of the circuit nodes change along the z-direction with changing anatomy.*

base anatomy can be loaded. In the *Circuit viewer*, each brain region is represented by a filled circle that has the same (user-definable) color as in the *Query builder*. A list of possible anatomical locations, where a node representing the region could be placed, is precomputed. To distinguish between left and right hemisphere, we need two locations per brain region. By default, nodes will be placed at the centroid of the region. If the centroid is already in use, another nearby location is selected from the precomputed list. Connections between nodes are drawn using spline curves. An arrow at the end of the curve indicates the 'terminates in' relation. The 'soma in' relation is not visualized but the 'transits' relation is, using a dark ring around the node. Width and color of connections can be changed interactively. The width can also be scaled according to the number of neurons, if desired by the user (blue connections in Fig. 6C). Nodes can also be moved interactively to any preferable location; corresponding links will automatically follow.

The viewer comes in two flavors. In the default, the anatomical outlines are generated by projecting the zebrafish larva into the plane. Alternatively (see Fig. 7), a slice-based representation of the zebrafish and the precomputed locations for each region can be used. The main purpose of this view is to give neurobiologists an idea how the locations of the circuit nodes change along the z-direction with changing anatomy. Note that this option shows cross-sections of the anatomy while the default option shows its maximum projection.

**2D and 3D viewer** The *2D* and *3D viewers* allow the user to display selected brain regions, cell bodies and cell morphologies in 2D and 3D to support the visual exploration of a neural circuit (**B1, B3, T4**; Fig. 1F; Fig. 4D). Visualization of multi-channel data is also supported, allowing users to render multiple image datasets in the viewers as separate color channels, e.g., gad1b expression (red), vglut2a expression (green), and the reference brain anatomy (blue), respectively.

## 6.4 Implementation

The visual analysis tool has been integrated into a web-based zebrafish atlas [52]. The implementation of the *Query builder* is based on react-complex-tree [53]. Bar and pie graphs make use of Vega and Vega-Lite [54], a declarative grammar to generate interactive charts. The *Circuit viewer* and the *Pathway browser* are implemented as custom modules using functionality from D3.js [55]. The 2D and 3D anatomical viewers are based on Babylon.js [56]. The backend routines are implemented in Python using gunicorn [57] as web server. The filter query is also implemented in Python, supporting all valid set operations (union: $\cup$, intersection: $\cap$, difference: $\setminus$, symmetric difference: $\wedge$). Offline precomputation and caching mechanisms are applied to handle spatial relations, e.g., 'transits', 'terminates in', for large neuron networks.The data component for storing the anatomical image stacks, brain region meshes, and cell morphology data is based on the MinIO file server, and on MongoDB.

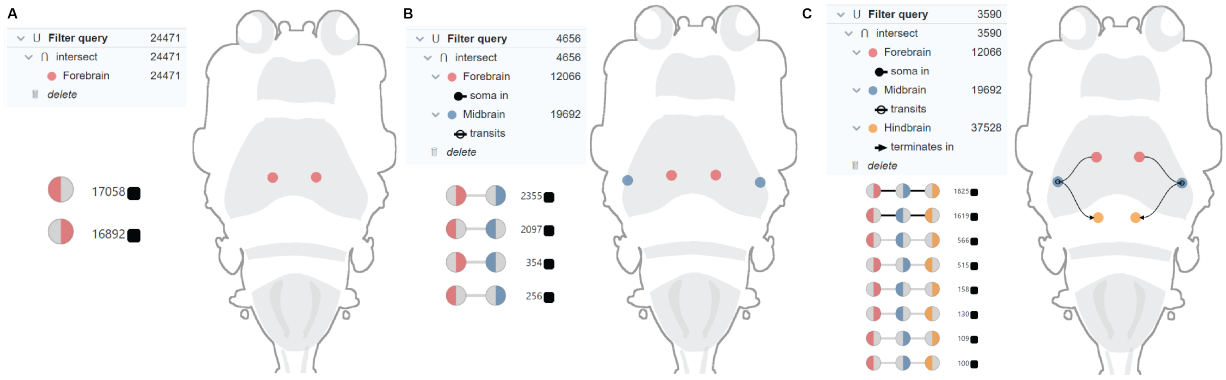


**Figure 8:** *Elementary case study (Sect. 7.2).* ***A****: Circuit viewer (Sect. 6.3) showing the forebrain nodes in the left and right hemisphere.* ***B****: Circuit viewer (Sect. 6.3) showing the forebrain and midbrain nodes in the left and right hemisphere; pathway browser (Sect. 6.2) showing the list of possible connections between forebrain and midbrain.* ***C****: Circuit viewer (Sect. 6.3) displaying the forebrain, midbrain and hindbrain nodes in the left and right hemisphere; listing of possible and selected pathways using the pathway browser (Sect. 6.2).*

# 7 Evaluation and case studies

We tested the visual interface using two case studies that seek to experimentally link brain-wide functional imaging data and behavior in a six-days post-fertilization larval zebrafish. The first study identified populations of neurons that responded during visual looming stimulation. In the second case study, the neural bases of sensorimotor decision making were investigated in the context of a movement integration task. We performed case studies in a complementary fashion. In the first use case (Sect. 7.3), we started with the existing circuits and found an interesting set of neurons and regions. In the second use-case (Sect. 7.4), we assume that the circuit is unknown and show how neurobiologists propose a circuit, given functional cell bodies, by obtaining a basic idea of connections between regions. The case studies were conducted jointly by three neurobiologists (all postdocs) and a developer of the tool via screen sharing.

## 7.1 Scalability evaluation

In order to evaluate the scalability of our platform, we developed a simple neuron simulator using existing brain parcellations, cell types and neuron morphologies. The simulator has two modes for creating neurons, one using brain parcellations and the other using cell types. In the first mode, the simulator selects a random brain region and then randomly selects a neuron from the set of neurons that have their soma in the selected brain region. The selected neuron is then spatially translated to a new neighborhood. This is repeated until the desired number of neurons has been created. In the second mode, the neuron simulator repeatedly selects a random cell type and randomly picks a neuron from the set of neurons of that cell type. Using this simulator, we generated two synthetic datasets with 10,000 and 50,000 neurons for each mode.

Applying queries to these data sets was instantaneous and smooth. The rendering of pathways and neuron morphologies was also smooth and interactive. The neuron browser automatically switches to server-side pagination mode if the number of neurons matching the query exceeds more than 1,000 and renders only the current page with a maximum of 500 neurons. A hypothetical neural circuit related to certain behavior generally consists of fewer than 100 neurons; the rendering capacity of our 3D viewer is much higher. The 3D viewer interactively rendered 1,500 neurons, consisting of more than 32,000 neurites.

The time for preprocessing the model with 50,000 neurons was less than 120 minutes on a standard PC with 32 GB memory and an Intel I7 processor. The preprocessed data can be fully loaded into the main memory (ca. 250 MB for the dataset with 50,000 neurons).

## 7.2 Elementary case study

This case study demonstrates a simple query operation and how the corresponding neuron pathways are generated. The case study was performed using a simulated dataset with 50,000 neurons with three large regions, i.e. forebrain, midbrain and hindbrain, for performing query operations. The exact query is to "filter subsets of neurons having their somas in the Forebrain, transit through the Midbrain, and terminate in Hindbrain". We started by dragging the 'intersect' operator and the 'Forebrain' region into the filter query component (Fig. 8A). By default, it shows the number of neurons for both the left and right hemisphere. Next, we dragged the Midbrain and applied the filter predicate 'soma in' to Forebrain and 'transits' to Midbrain, respectively (Fig. 8B). In the following, we use abbreviations for neuronal con-

nections between brain regions, where F, M, and H stand for forebrain, midbrain, and hindbrain, and l and r stand for left and right, respectively. The pathway browser (Fig. 8B) shows four possible connections, i.e. Fl $\rightarrow$ Ml, Fl $\rightarrow$ Mr, Fr $\rightarrow$ Ml, and Fr $\rightarrow$ Mr. Finally, we dragged 'Hindbrain' into the query component and applied the 'terminates in' predicate (Fig. 8C). The recomputed neuron pathways show eight possible connections (Fig. 8C) sorted by the number of neurons, i.e. Fl $\rightarrow$ Ml $\rightarrow$ Hl, Fl $\rightarrow$ Ml $\rightarrow$ Hr, Fl $\rightarrow$ Mr $\rightarrow$ Hl, Fl $\rightarrow$ Mr $\rightarrow$ Hr, Fr $\rightarrow$ Ml $\rightarrow$ Hl, Fr $\rightarrow$ Ml $\rightarrow$ Hr, Fr $\rightarrow$ Mr $\rightarrow$ Hl, and Fr $\rightarrow$ Mr $\rightarrow$ Hr. The drag-and-drop interaction to create this query is also shown in the accompanying video.

### 7.3 Case study 1: Analysis of neuron populations involved in inhibiting escape response

This case study utilizes data from experiments designed to find network elements that inhibit escape to non-threatening, i.e., dimming stimuli. An object approaching on a collision course was simulated by a dimming disk on a light background and presented to the animal from the side, while the behavior of the animal was monitored. For a more detailed understanding, the experts also performed 2-photon whole-brain functional imaging during the experiment. This identified distinct populations of neurons tuned to different aspects of the stimulus, e.g., expansion or luminance, as well as to motor behavior [59].

Mauthner neurons (also known as motor command neurons), located in rhombomere 4 of the hindbrain, are known for triggering escape response [58]. The question is, whether the information about the dimming aspect of stimulus is in fact relayed to the downstream motor networks, and whether the dimming pathway is part of an inhibitory network that stops Mauthner cells from firing. In fact, both dimming and looming pathways have been identified to leave the tectum and project to downstream networks [49]. More precisely, what is the neuronal circuitry responsible for relaying this information and which brain regions are involved in this? There are two competing hypotheses about the circuit. **Hypothesis 1:** There are neurons that go from the optic tectum to the caudal hypothalamus and connect with dopaminergic neurons, which then project to and excite glycinergic clusters [58] that in turn connect to the Mauthner cells (Fig. 9A). **Hypothesis 2:** There are neurons that go from the optic tectum directly to glycinergic neurons, which in turn connect to the Mauthner cells (Fig. 1A).

To test **hypothesis 1** (Fig. 9), the experts first searched the subset of neurons projecting from the optic tectum (with subregions Stratum Periventriculare and Neuropil, located in the midbrain) to the caudal hypothalamus by interactively selecting the regions with the query builder and combining them with appropriate union and intersection operators, followed by applying filter predicates 'soma in' to 'Optic Tectum' and 'transits' to 'Caudal Hypothalamus', respectively (Fig. 9B). In this way, the experts were able to find three such neurons (Fig. 9B,F). Next, the experts performed a query operation (Fig. 9E) searching for neurons that project from the caudal hypothalamus to any of 10 glycinergic interneuron clusters that are located next to the Mauthner cell in the hindbrain (Fig. 9D) and inhibit escape response. No neuron matching this query could be found, see red box in Fig. 9E. Then, they looked for neurons that make connections from any of these 10 glycinergic interneuron clusters to the Mauthner cell region. In total, 47 such neurons were found (see supp. video) and the most interesting glycinergic clusters, i.e., with neurons transiting through these regions and projecting to the Mauthner region, are clusters 1, 3 and 6 (Fig. 9G). Overall, this first hypothesis looks promising and several connections are narrowed down using our tool. However, further experiments are required to find connections from caudal hypothalamus to potentially involved glycinergic interneuron clusters.

To test **hypothesis 2** (Fig. 1), the experts tried to find neurons that pass through any of the tectal layers, which cover large portions of the Stratum Periventriculare and Neuropil [60], e.g., SO 1-2, SFGS 1-2 (Fig. 1B), have their soma in the midbrain, transit through the medial tegmentum and terminate in the caudal hindbrain (Fig. 1C-F). The experts were able to find five such neurons as shown in the anatomical viewer (Fig. 1E). Next, they searched for neurons that transit through or terminate in glycinergic interneuron clusters or the caudal hypothalamus, located in the forebrain. There is one such neuron in the left hemisphere that passes through the caudal hindbrain but does not make connections with glycinergic interneuron cluster sites. While performing the analysis on the right hemisphere, we found that there are two neurons that make connections with glycinergic interneuron cluster 9 but do not transit through the caudal hypothalamus. Furthermore, there are neurons from the glycinergic interneuron cluster 9 that make connections with the Mauthner cells.

As result of the above analyses, we obtained a list of potential glycinergic interneuron clusters being involved in the inhibition of escape response. These findings including the identified functional cell bodies will be reported in more detail in a neurobiology journal. Regarding the tested hypotheses, hypothesis

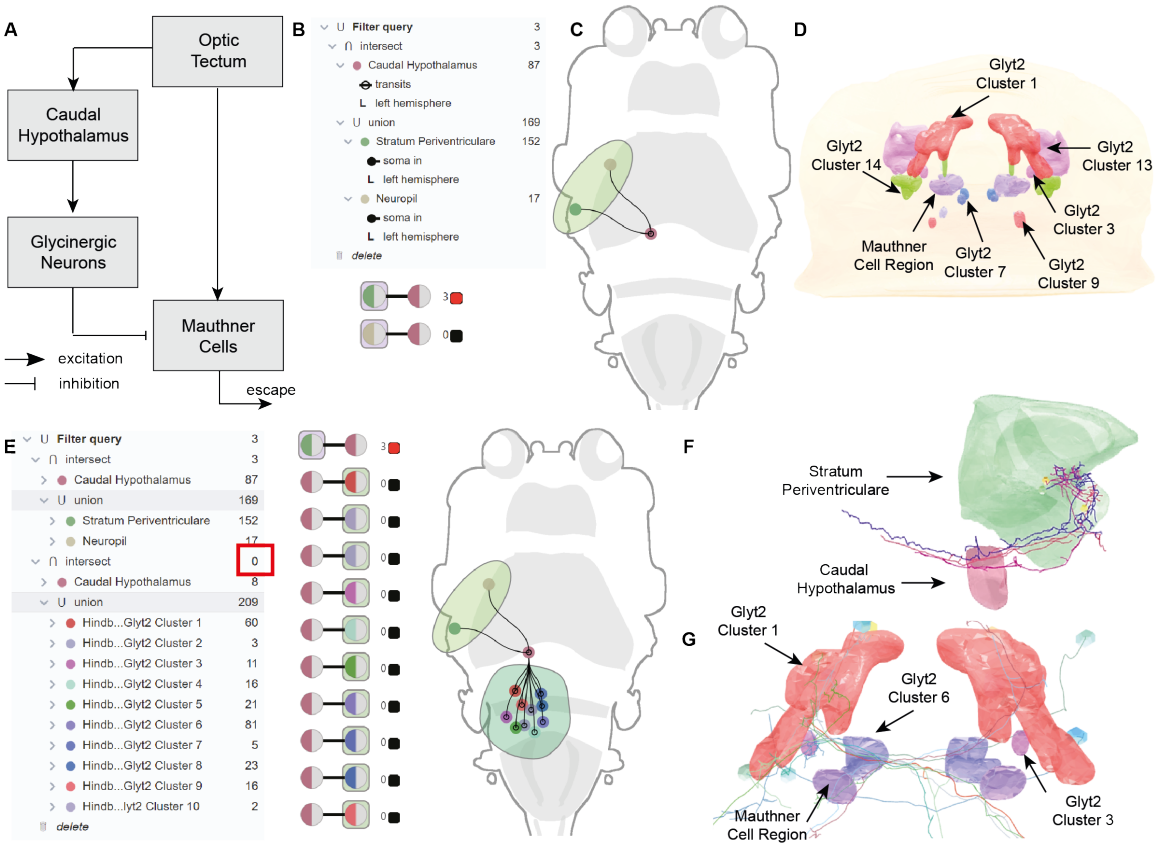


**Figure 9:** ***A****: Graphical illustration of the circuit hypothesis regarding the inhibiting escape behavior, proposed by Yao et al. [58], described in more detail in case study 2 (Sect. 7.3).* ***B****: Query builder (Sect. 6.1);* ***C****: Circuit viewer (Sect. 6.3);* ***D****: 3D visualization of glycinergic interneuron clusters that are located next to Mauthner cell region.* ***E****: Example of a query, described in Sect. 7.3. The number of neurons satisfying each subquery are displayed with a subquery returning no results (red box);* ***F****: 3D view revealing the connectivity of neurons that have their soma in the stratum periventriculare and pass through the caudal hypothalamus;* ***G****: 3D view revealing the connectivity of neurons transiting through glycinergic interneuron clusters and projecting to the Mauthner region.*

2 seems more likely since when testing hypothesis 1, we could not find any connections from caudal hypothalamus to potentially involved glycinergic interneuron clusters. However, as this might be due to missing data, it will be subject to further investigation.

## 7.4 Case study 2: Analysis of neuron populations involved in the integration of visual motion cues

This case study examines the optomotor reflex, which causes many animals to orient toward a visual motion stimulus and is thought to be important for compensating the involuntary drift in flowing streams. In this behavioral experiment, the animal was presented with a coherent dot stimulus of approx. 1000 dots projected from below in a circular arena. The dots are presented at 4 different coherence levels, i.e. 0%, 25%, 50%, and 100%, either moving left- or rightwards relative to the body orientation of the fish. The coherence level means that a certain fraction of dots moves coherently in one direction, while the remaining dots stochastically disappear and reappear at random locations in the arena. This type of stimulus makes it more difficult for the animal to determine the correct motion direction and requires the brain to integrate information over time. To link these behavioral observations to neural dynamics, the experts used 2-photon whole-brain functional imaging to search for direction-selective neurons with slow time constants. They found such neurons in several parts of the brain, with the highest density in the anterior hindbrain. Using these data, the experts then created data-compatible, hypothetical neural circuits that could mimic the observed dynamics within – and across – the identified brain regions. Many open questions remain, such as what the neurotransmitter identity of the identified cells is, how their precise anatomy looks like, or which neurons may be connected.

Using the proposed framework, the experts started by analyzing the spatial distribution of cell bodies across brain regions by selecting regions in the hierarchical tree map view of the region finder, as shown in Fig. 10A and B. Several regions of interest were found

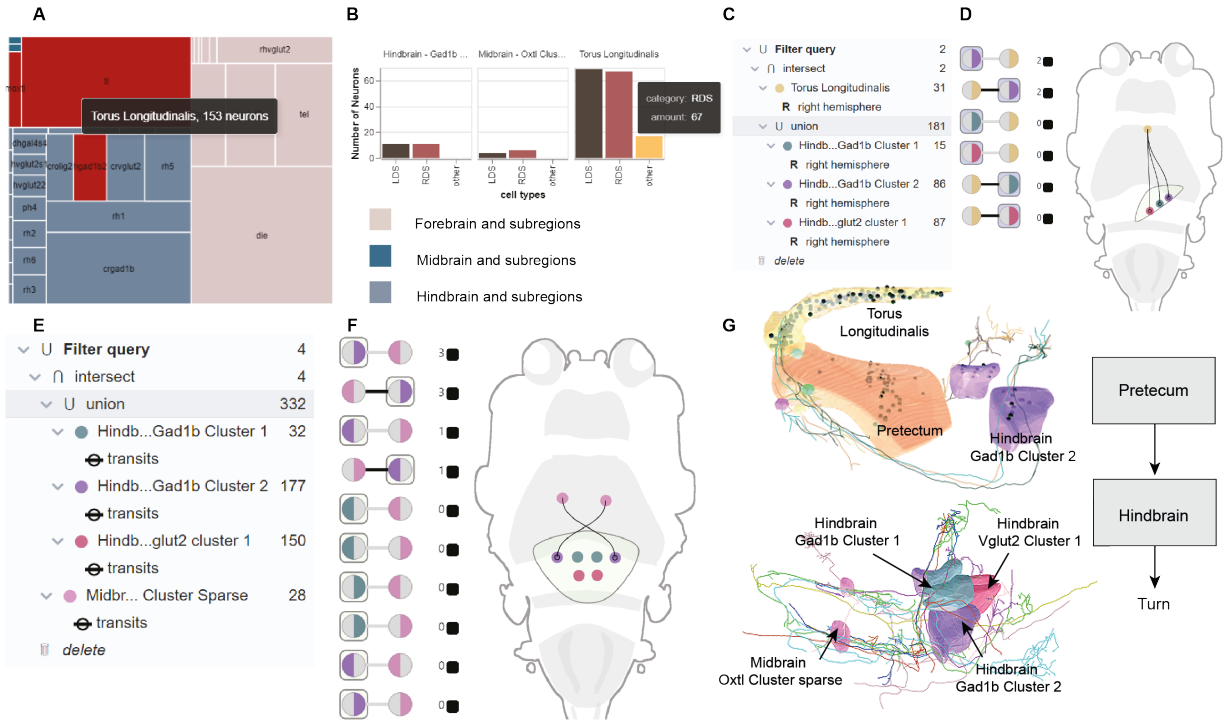


**Figure 10:** ***A****: As first step, the user identifies brain regions of interest using the tree map view. The selected cells are highlighted in red and the name of a brain region appears when hovering over a cell.* ***B****: The functional cell type compositions of regions selected in the tree map are shown in bar charts.* ***C****: User query, showing the selection of anterior hindbrain as group entity.* ***D****: Circuit viewer (Sect. 6.3) showing the anterior hindbrain region nodes in the right hemisphere and the list of possible paths connecting with torus longitudinalis.* ***E****: User query for finding the contralateral projecting neurons from midbrain oxytocin cluster to anterior hindbrain;* ***F****: Display of all possible contralateral projections from the midbrain oxytocin cluster; the widths of the projected links are automatically adjusted according to neuron distributions; the path with the highest number of projection is shown in green.* ***G****: 3D view showing neurons passing through the anterior hindbrain cluster and torus longitudinalis. Somata of these neurons are in the pretectum. Next, the neurons passing through the midbrain oxytocin cluster and making contralateral projections in anterior hindbrain regions are shown; graphical depiction of visual motion cues circuit.*

by exploring the hierarchy of regions, cell types and spatial information: telencephalon vglut2 rind, diencephalon left and right habenula vglut2 cluster (in the forebrain), torus longitudinalis, pretectum, oxytocin cluster (in the midbrain) and gad1b cluster1-2, vglut cluster1, gad1b enriched area and olig2 enriched area (in the hindbrain). This confirms results that the expert had found before [47].

The expert started with a query operation for which he had a hypothesis in mind but had never seen such connections: select all neurons that transit through torus longitudinalis, that make connections in anterior hindbrain areas (i.e. gad1b cluster1-2, vglut cluster1), as shown in Fig. 10C,D. Closer inspection of these neurons revealed that their somata are located in the Pretectum (Fig. 10G). This query filtered out two neurons, the morphology of which is shown in Fig. 10G. They reveal an interesting connection pattern between these three regions. Next, the expert performed a query operation to explore neurons that branch out in a midbrain oxytocin-like cluster and connect to the anterior hindbrain. The query shown in Fig. 10E filters out these neurons. Finally, the expert performed a query for all neurons with somata in anterior hindbrain. These neurons branch out in several other regions, e.g. hindbrain gad1b cluster1 and have contralateral projections. Based on the above analysis, the expert narrowed down aspects of the hypothesis about the neural circuit responsible for decision making in the larval zebrafish brain, gaining the basic idea about involved input and output connections between neuronal populations. Specifically, the information obtained about potential local and long-range connections could be used to postulate a densely interconnected circuit motif with interhemispheric interactions. Such a network model could implement the empirically observed slow circuit dynamics and also predict behavioral decisions.

## 7.5 User survey

In order to obtain qualitative feedback and further improve the tool, we conducted a user study. We invited participants from the fields of neuroscience, computational neuroscience, and data science/visualization.

In the survey, we asked the participants to watch

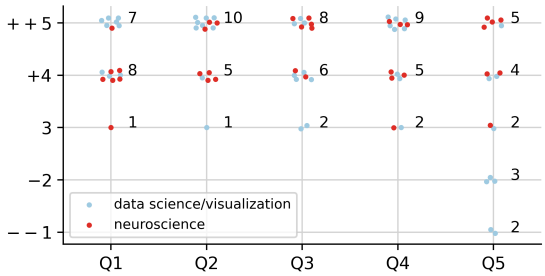


**Figure 11:** *Survey responses of the participants (9 from data science/visualization, 7 from neuroscience).*

a 9:16 min long video introducing the tool (see Supplementary Material) and encouraged them to read the questionnaire before watching the video. Additionally, we provided an online link to interactively explore the tool. The questionnaire's first part concentrated on a few generic questions regarding gender, what best describes their field of work (neuroscience, visualization, computer/data science, other), their lab affiliation (Engert lab, Baum lab, other) and their current role (principal investigator, postdoc, PhD candidate, other). Next, we asked the participants to rate the following questions using a 5-item Likert scale [61] (from (1) strongly disagree to (5) strongly agree, or from (1) very unlikely to (5) very likely). The questions underlying the ratings were: *Q1: Overall, the tool was easy to use*; *Q2: The query panel functionality (i.e., drag and drop) was easy to use and the query builder components (i.e., set operators, filter predicates, and biological entities) provided in the panel are appropriate*; *Q3: The pathway browser would make it easier to explore pathways and filter neurons of interest*; *Q4: The circuit viewer provides an intuitive visualization of the built neural circuit*; *Q5: The tool itself and/or the proposed interactive way to create queries could be of interest in my own ongoing research, if the necessary changes were made*. Finally, we asked the participants two optional questions with free textual response: *Q6: What changes would be required to make the tool more appealing for your own work?*; *Q7: What new features in the tool would you like to see in the future?*

In total, 16 users participated in the survey of whom 9 were data science/visual researchers and 7 potential users from neuroscience. 11 participants were male, 4 were female, and 1 preferred not to say. *Q1*, *Q2*, *Q3*, and *Q4* were rated positive by all participants (Fig. 11). The response for question *Q5* was mixed: the tool was rated positively by all neurobiologists (Fig. 11), but the data scientists were (unsurprisingly) not sure whether it could help them in their research. Nevertheless, one data scientist, who works on semiconductor wafers, wrote: "The tool would be helpful for my work if the query panel would allow me to filter wafers of interest based on defects and metrological properties". One of the neurobiologists suggested changing the "filter predicates" heading with a heading based on biological meaning. Another neurobiologist suggested using biological meanings for the "union and intersect" operators. The response for *Q7* was also positive and several features were requested by the participants, e.g., "Orthogonal views in the circuit viewer as multiple regions are overlapping when the fish is seen from the top.", "Drag and Drop is great but I absolutely like keyboard shortcuts and I think at least deleting items in the query should be there." Other requested features included help pages and support for all browsers.

# 8 Conclusion

Neuroscience is faced with a vast and ever-growing amount of information about individual neurons, their morphological, genetic and functional types, and their connections. We presented a visual analysis tool that allows neuroscientists to quickly extract those pieces of information that are relevant to their research questions. In particular, the tool supports the *advancement of hypotheses about neural circuits*: it enables neuroscientists to evaluate existing hypotheses, expand incomplete ones, refine coarse ones, constrain and narrow down vague ones and, based on previously unknown information, develop entirely new ones.

The web-based tool is built into a 3D brain atlas that provides all relevant spatial information and additional information related to spatial entities to interactively generate data-compatible neural circuits. It is unique in combining a fully interactive hierarchical filter query editor with the real-time visualization of all resulting pathways, which can be interactively explored in a novel 2D visualization that is based on the anatomical map of the zebrafish, which greatly supports the understanding of the built neural circuit. The additional pathway view allows an easy selection and highlighting of important pathways which helps the user to keep an overview of the neuronal circuit that can become quite complicated. The filtered neurons can be further analyzed by linking them to additional views, including a 3D anatomical viewer that shows selected brain regions and neurons. Scalability was demonstrated using a large synthetic dataset that approximates the properties of large experimental datasets that will become available soon. However, transitioning to dense EM connectomes with synaptic connectivity information would likely necessitate the integration of fast subgraph enumeration algorithms [62]. As this scalability requirement

was not immediate, we leave it to future work.

The practical applicability of the tool was confirmed in a qualitative user survey and two case studies, elucidating neural circuits of the zebrafish larva that are responsible for the transformation of (visual) sensory information into a (motor) response. It is now being used in current brain research to generate, test and constrain ideas about how circuits are implemented in the fish brain. Future work will include generalizations that increase the range of applications in neuroscience and, for example, require support of further filter predicates. Currently, filter predicates are pre-defined. Support for user-defined predicates, based on, e.g., functional data, would expand the functionality considerably. Although the presented 2D visualization is favorable for interactive exploration, an additional representation of (possibly simplified) neural circuits in 3D could further improve understanding.

## 9 Acknowledgment

This work was supported by the National Institute of Neurological Disorders and Stroke (NINDS) of the National Institutes of Health (NIH, 1U19NS104653) and the Deutsche Forschungsgemeinschaft (DFG, SPP 2041 Computational Connectomics, 347210657). F. E. received funding from the NIH (U19NS104653, 1R01NS124017), the National Science Foundation (IIS-1912293) and the Simons Foundation (SCGB 542973). A. B. was supported by the German Emmy Noether Program (DFG 429442687) and the Zukunftskolleg Konstanz. We thank the Herwig Baier lab for sharing neuron morphologies.

## References

[1] Shin-ya Takemura et al. “A visual motion detection circuit suggested by Drosophila connectomics”. In: *Nature* 500.7461 (2013), pp. 175–181. DOI: 10.1038/nature12450.

[2] Tomoko Ohyama, Casey M Schneider-Mizell, Richard D Fetter, et al. “A multilevel multimodal circuit enhances action selection in Drosophila”. In: *Nature* 520.7549 (2015), pp. 633–639. DOI: 10.1038/nature14297.

[3] David Grant Colburn Hildebrand, Marcelo Cicconet, Russel Miguel Torres, et al. “Whole-brain serial-section electron microscopy in larval zebrafish”. In: *Nature* 545.7654 (2017), pp. 345–349. DOI: 10.1038/nature22356.

[4] Michael Kunst et al. “A Cellular-Resolution Atlas of the Larval Zebrafish Brain”. In: *Neuron* 103.1 (2019), pp. 21–38. DOI: 10.1016/j.neuron.2019.04.034.

[5] Rainer W Friedrich and Adrian A Wanner. “Dense Circuit Reconstruction to Understand Neuronal Computation: Focus on Zebrafish”. In: *Annu. Rev. Neurosci.* 44 (2021), pp. 275–293. DOI: 10.1146/annurev-neuro-110220-013050.

[6] Robert Brandt et al. “Three-dimensional average-shape atlas of the honeybee brain and its applications”. In: *J. Comp. Neurol.* 492.1 (2005), pp. 1–19. DOI: 10.1002/cne.20644.

[7] Owen Randlett et al. “Whole-brain activity mapping onto a zebrafish brain atlas”. In: *Nat. Methods* 12.11 (2015), pp. 1039–1046. DOI: 10.1038/nmeth.3581.

[8] Werner Gobel and Fritjof Helmchen. “In vivo calcium imaging of neural network function”. In: *Physiology* 22.6 (2007), pp. 358–365. DOI: 10.1152/physiol.00032.2007.

[9] Stephen M. Plaza et al. “neuPrint: An open access tool for EM connectomics”. In: *Front. Neuroinform.* 16 (2022). ISSN: 1662-5196. DOI: 10.3389/fninf.2022.896292.

[10] Florian Ganglberger et al. “BrainTrawler: A visual analytics framework for iterative exploration of heterogeneous big brain data”. In: *Comput. Graph.* 82 (2019), pp. 304–320. ISSN: 0097-8493. DOI: 10.1016/j.cag.2019.05.032.

[11] Alex Fornito, Andrew Zalesky, and Edward T. Bullmore. *Fundamentals of Brain Network Analysis*. English. Elsevier Academic Press, 2016, xvii, 476 pages. ISBN: 9780124079083. DOI: 10.1016/C2012-0-06036-X.

[12] Daniel S. Margulies et al. “Visualizing the human connectome”. In: *NeuroImage* 80 (2013), pp. 445–461. DOI: 10.1016/j.neuroimage.2013.04.111.

[13] Weifeng Chen, Lei Shi, and Wei Chen. “A survey of macroscopic brain network visualization technology”. In: *Chin. J. Electron.* 27.5 (2018), pp. 889–899. DOI: 10.1049/cje.2018.04.007.

[14] Johanna Beyer et al. “A Survey of Visualization and Analysis in High-Resolution Connectomics”. In: *Comput. Graph. Forum* 41.3 (2022), pp. 573–607. ISSN: 1467-8659. DOI: 10.1111/cgf.14574.

[15] Luis Pastor et al. “A Unified Framework for Neuroscience Morphological Data Visualization”. In: *Appl. Sci.* 11.10:4652 (2021). DOI: 10.3390/app11104652.

[16] Hanspeter Pfister et al. "Visualization in Connectomics". In: *Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization*. Ed. by Charles D. Hansen et al. Springer, 2014, pp. 221–245. DOI: 10.1007/978-1-4471-6497-5_21.

[17] Stephan Saalfeld et al. "CATMAID: Collaborative annotation toolkit for massive amounts of image data". In: *Bioinformatics* 25.15 (2009), pp. 1984–1986. ISSN: 13674803. DOI: 10.1093/bioinformatics/btp266.

[18] Johanna Beyer et al. "ConnectomeExplorer: Query-guided visual analysis of large volumetric neuroscience data". In: *IEEE Trans. Vis. Comput. Graph.* 19.12 (2013), pp. 2868–2877. ISSN: 10772626. DOI: 10.1109/TVCG.2013.142.

[19] A. K. Al-Awami et al. "NeuroLines: A Subway Map Metaphor for Visualizing Nanoscale Neuronal Connectivity". In: *IEEE Trans. Vis. Comput. Graph.* 20.12 (2014), pp. 2369–2378. DOI: 10.1109/TVCG.2014.2346312.

[20] *Neuroglancer: WebGL-based viewer for volumetric data*. URL: https://github.com/google/neuroglancer.

[21] Stephan Gerhard et al. "The Connectome Viewer Toolkit: An Open Source Framework to Manage, Analyze, and Visualize Connectomes". In: *Front. Neuroinform.* 5 (2011), p. 3. DOI: 10.3389/fninf.2011.00003.

[22] Federico Claudi et al. "Visualizing anatomically registered data with brainrender". In: *Elife* 10 (2021), e65751. DOI: 10.7554/eLife.65751.

[23] Stefan Bruckner et al. "BrainGazer - Visual Queries for Neurobiology Research". In: *IEEE Trans. Vis. Comput. Graph.* 15.6 (2009), pp. 1497–1504. DOI: 10.1109/TVCG.2009.121.

[24] Johannes Sorger et al. "neuroMAP - Interactive graph-visualization of the fruit fly's neural circuit". In: *IEEE Symposium on Biological Data Visualization, BioVis 2013, Atlanta, GA, USA, Oct 13-14, 2013*, pp. 73–80. DOI: 10.1109/BioVis.2013.6664349.

[25] Nicolas Swoboda et al. "Visual and Quantitative Analysis of Higher Order Arborization Overlaps for Neural Circuit Research". In: *Eurographics Workshop on Visual Computing for Biology and Medicine, VCBM 2014, Vienna, Austria*. Eurographics, 2014, pp. 107–116. DOI: 10.2312/vcbm.20141189.

[26] Nicolas Swoboda et al. "Visualization and Quantification for Interactive Analysis of Neural Connectivity in Drosophila". In: *Comput. Graph. Forum* 36.1 (2017), pp. 160–171. DOI: 10.1111/cgf.12792.

[27] Vincent J. Dercksen et al. "Synaptic Connectivity in Anatomically Realistic Neural Networks: Modeling and Visual Analysis". In: *Eurographics VCBM*. 2012, pp. 17–24. DOI: 10.2312/VCBM/VCBM12/017-024.

[28] Florian Ganglberger, Nicolas Swoboda, Lisa Frauenstein, et al. "Iterative Exploration of Big Brain Network Data". In: *Eurographics Workshop on Visual Computing for Biology and Medicine*. Eurographics Association, 2018, pp. 77–87. ISBN: 978-3-03868-056-7. DOI: 10.2312/vcbm.20181231.

[29] Sebastian Schwanke et al. "Towards Differential Connectomics with NeuroVIISAS". In: *Neuroinformatics* 17.1 (2019), pp. 163–179. ISSN: 15392791. DOI: 10.1007/s12021-018-9389-6.

[30] A. Maye, T.H. Wenckebach, and H.-C. Hege. "Visualization, Reconstruction, and Integration of Neuronal Structures in Digital Brain Atlases". In: *Int. J. Neurosci.* 116.4 (2006), pp. 431–459. DOI: 10.1080/00207450500505860.

[31] Allen Institute. *Documentation:Allen Mouse Brain Connectivity Atlas*. URL: http://help.brain-map.org/display/mouseconnectivity/Documentation.

[32] N. Milyaev et al. "The Virtual Fly Brain browser and query interface". In: *Bioinformatics* 28.3 (2012), pp. 411–415. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btr677.

[33] Ching-Yao Lin et al. "The Neuron Navigator: Exploring the information pathway through the neural maze". In: *IEEE Pacific Visualization Symposium, PacificVis 2011, Hong Kong, China, 1-4 March, 2011*. 2011, pp. 35–42. DOI: 10.1109/PACIFICVIS.2011.5742370.

[34] *The Larval Standard Brain*. URL: https://zebrafish.tefor.net.

[35] *ZFIN - Anatomy Atlases and Resources*. URL: https://zfin.atlassian.net/wiki/spaces/general/pages/1892876451.

[36] Anja Kuß et al. "Ontology-Based Visualization of Hierarchical Neuroanatomical Structures". In: *Proc. Eurographics Workshop on Visual Computing for Biomedicine, VCBM 2008, Delft, Netherlands*. Eurographics, 2008, pp. 177–184. DOI: 10.2312/VCBM/VCBM08/177-184.

[37] Padraig Gleeson, Matteo Cantarelli, B. Marin, et al. "Open Source Brain: A Collaborative Resource for Visualizing, Analyzing, Simulating, and Developing Standardized Models of Neurons and Circuits". In: *Neuron* 103.3 (2019), 395–411.e5. ISSN: 10974199. DOI: 10.1016/j.neuron.2019.05.019.

[38] Michael Schirner, Lia Domide, Dionysios Perdikis, et al. "Brain simulation as a cloud service: The Virtual Brain on EBRAINS". In: *NeuroImage* 251 (2022), p. 118973. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2022.118973.

[39] Matteo Cantarelli, Boris Marin, Adrian Quintana, et al. "Geppetto: A reusable modular open platform for exploring neuroscience data and models". In: *Philos. Trans. R. Soc. B* 373.1758 (2018), p. 20170380. ISSN: 14712970. DOI: 10.1098/rstb.2017.0380.

[40] A. Kuß et al. "Effective Techniques to Visualize Filament-Surface Relationships". In: *Comput. Graph. Forum* 29.3 (2010), pp. 1003–1012. DOI: 10.1111/j.1467-8659.2009.01703.x.

[41] Florian Ganglberger, Monika Wißmann, Hsiang-Yun Wu, et al. "Spatial-data-driven layouting for brain network visualization". In: *Computers & Graphics* 105 (2022), pp. 12–24. ISSN: 0097-8493. DOI: 10.1016/j.cag.2022.04.014.

[42] H-J Schulz, S Hadlak, and H Schumann. "The Design Space of Implicit Hierarchy Visualization: A Survey". In: *IEEE Trans. Vis. Comput. Graph.* 17.4 (2011), pp. 393–411. DOI: 10.1109/tvcg.2010.79.

[43] C. Nobre et al. "The state of the art in visualizing multivariate networks". In: *Comput. Graph. Forum* 38.3 (2019), pp. 807–832. ISSN: 14678659. DOI: 10.1111/cgf.13728.

[44] Corinna Vehlow, Fabian Beck, and Daniel Weiskopf. "Visualizing Group Structures in Graphs: A Survey". In: *Comput. Graph. Forum* 36.6 (2017), pp. 201–225. ISSN: 14678659. DOI: 10.1111/cgf.12872.

[45] Aristotelis Leventidis et al. "QueryVis: Logic-based Diagrams help Users Understand Complicated SQL Queries Faster". In: *Proc. ACM SIGMOD Int. Conf. Manag. Data*. ACM, 2020, pp. 2303–2318. DOI: 10.1145/3318464.3389767.

[46] Azza Abouzied, Joseph Hellerstein, and Avi Silberschatz. "DataPlay". In: *Proc. 25th Annu. ACM Symp. User Interface Softw. Technol. - UIST '12*. ACM Press, 2012, pp. 207–218. DOI: 10.1145/2380116.2380144.

[47] Armin Bahl and Florian Engert. "Neural circuits for evidence accumulation and decision making in larval zebrafish". In: *Nat. Neurosci.* 23.1 (2020), pp. 94–102. DOI: 10.1038/s41593-019-0534-9.

[48] Miranda Robbins et al. "Calcium imaging analysis – how far have we come?" In: *F1000Research* 10 (2021), p. 258. DOI: 10.12688/f1000research.51755.2.

[49] Thomas O. Helmbrecht et al. "Topography of a Visuomotor Transformation". In: *Neuron* 100.6 (2018), 1429–1445.e4. DOI: 10.1016/j.neuron.2018.10.021.

[50] Matthew Brehmer and Tamara Munzner. "A Multi-Level Typology of Abstract Visualization Tasks". In: *IEEE Trans. Vis. Comput. Graph.* 19.12 (2013), pp. 2376–2385. DOI: 10.1109/tvcg.2013.124.

[51] Willy Scheibel et al. "A taxonomy of treemap visualization techniques". In: *VISIGRAPP 2020 - Proc. 15th Int. Joint Conf. Comput. Vis. Imaging and Comput. Graph. Theory and Appl.* Vol. 3. 2020, pp. 273–280. ISBN: 9789897584022. DOI: 10.5220/0009153902730280.

[52] Sumit Kumar Vohra et al. "A multiscale, explorative atlas of the larval zebrafish's brain". in preparation.

[53] *React Complex Tree*. URL: https://github.com/lukasbach/react-complex-tree.

[54] Arvind Satyanarayan et al. "Vega-Lite: A Grammar of Interactive Graphics". In: *IEEE Trans. Vis. Comput. Graph.* 23.1 (2017), pp. 341–350. DOI: 10.1109/tvcg.2016.2599030.

[55] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. "D3 data-driven documents". In: *IEEE Trans. Vis. Comput. Graph.* 17.12 (2011), pp. 2301–2309. DOI: 10.1109/TVCG.2011.185.

[56] *BabylonJs: The WebGL-based rendering engine*. URL: https://www.babylonjs.com.

[57] *Gunicorn – Python WSGI HTTP Server for UNIX*. URL: https://gunicorn.org/.

[58] Yuanyuan Yao et al. "Visual Cue-Discriminative Dopaminergic Control of Visuomotor Transformation and Behavior Selection". In: *Neuron* 89.3 (2016), pp. 598–612. DOI: 10.1016/j.neuron.2015.12.036.

[59] Haleh Fotowat and Florian Engert. *Neural circuits underlying habituation of visually evoked escape behaviors in larval zebrafish*. preprint, bioRxiv, doi: 10.1101/2022.06.23.497366. June 2022.

[60] Timothy W. Dunn et al. “Neural Circuits Underlying Visually Evoked Escapes in Larval Zebrafish”. In: *Neuron* 89.3 (2016), pp. 613–628. DOI: 10.1016/j.neuron.2015.12.021.

[61] Rensis Likert. “A technique for the measurement of attitudes”. In: *Arch. Psychol.* 22 (1932), pp. 5–55.

[62] Brian Matejek et al. “Edge-colored directed subgraph enumeration on the connectome”. In: *Sci. Rep.* 12 (1 2022), p. 11349. DOI: 10.1038/s41598-022-15027-7.