



webKnossos: efficient online 3D data annotation for connectomics

Kevin M Boergens^{1,3} , Manuel Berning^{1,3}, Tom Bocklisch², Dominic Bräunlein², Florian Drawitsch¹, Johannes Frohnhofen², Tom Herold², Philipp Otto², Norman Rzepka², Thomas Werkmeister², Daniel Werner², Georg Wiese², Heiko Wissler¹ & Moritz Helmstaedter¹ 

We report webKnossos, an in-browser annotation tool for 3D electron microscopic data. webKnossos provides flight mode, a single-view egocentric reconstruction method enabling trained annotator crowds to reconstruct at a speed of 1.5 ± 0.6 mm/h for axons and 2.1 ± 0.9 mm/h for dendrites in 3D electron microscopic data from mammalian cortex. webKnossos accelerates neurite reconstruction for connectomics by 4- to 13-fold compared with current state-of-the-art tools, thus extending the range of connectomes that can realistically be mapped in the future.

With the acceleration of 3D electron microscopic (EM) imaging of brain tissue^{1–3}, image data sets sized tens of terabytes (TB) or even petabytes (PB) are becoming available. A cubic millimeter imaged at $(15 \text{ nm})^3$ voxel (vx) size corresponds to 0.3 PB of data (Fig. 1a); a mouse brain imaged at the same resolution corresponds to 110 PB of data. Single neurons typically extend over a large fraction of the data set (Fig. 1a), making it impracticable to distribute data on hard drives to large numbers of annotators who want to follow the processes of entire neurons. At the same time, data analysis in connectomics is limited by the amount of human annotation time that can be recruited for a given analysis project⁴. Thus, enabling efficient distributed 3D data annotation in PB-sized data sets, ideally in browser, is essential.

The existing in-browser annotation tool for connectomics, CATMAID^{5,6}, uses efficient data storage and transmission in 2D image planes (comparable to Google Maps⁷), which are sequentially browsed (Fig. 1b). While this approach makes data viewing and annotation seamless in the plane of imaging, 3D neurite tracing is slowed down by the time required to progress to the subsequent image plane. Under ideal high-bandwidth, low-latency connectivity conditions (like those within research institutions), this approach yields a reconstruction speed of about 470 $\mu\text{m/h}$

when an expert follows an axon through 3D serial blockface EM (SBEM¹) data from mouse cortex (Fig. 1c). Under less optimal bandwidth and latency conditions (as are often experienced by student annotators at home and in mobile settings), however, reconstruction speed drops to 59–88 $\mu\text{m/h}$ (measured for regular 3G connectivity and for transcontinental access, Fig. 1c).

webKnossos (Fig. 1d, <https://webknossos.org>) uses 3D data storage and transmission in small cubic packages of 32^3 vx (Fig. 1b). Cubic 3D image data storage using cubes of 128^3 vx is being employed in KNOSSOS⁸ and pyKNOSSOS^{9,10}, standalone data annotation applications for connectomics. We reduced 3D cube size to 32^3 vx for lag-free in-browser data transmission and for enabling the flight-mode data visualization introduced below. webKnossos enables data interaction in 3D EM images displayed in orthogonal planes (Fig. 1d) at a speed of about 2 mm/h (Fig. 1c), which drops to 0.7–1.2 mm/h for regular 3G connectivity and transcontinental annotation (Fig. 1c). Thus, 3D data visualization when following an axon is about 4-fold faster under ideal and up to 13-fold faster under nonoptimal connectivity settings than with existing in-browser tools (Fig. 1c).

We next tested whether student annotators can be trained to interact with 3D brain image data at such speeds. Previously, annotators interacted with the image data using orthogonal image projections in the three cardinal planes for following the neuronal processes and for their annotation as ‘skeletons’ (Fig. 1d,e, ‘ortho mode’, KNOSSOS⁸). Effective tracing speed was 100–270 $\mu\text{m/h}$ ^{4,8,10–12} for reconstructions in mouse retina, zebrafish olfactory bulb and mouse cortex. We asked whether annotators can be trained to annotate faster in ortho mode, and whether a more intuitive data presentation can further accelerate human annotation. For the latter, we developed ‘flight mode’, in which the 3D image data are sampled on a hemisphere centered at the annotator’s current position (Fig. 1f). To enable such non-orthogonal data transmission and display in browser, we used (in addition to the small 3D cube size (Fig. 1b)) a simple form of path prediction (the data being loaded in a stump in direction of flight; Fig. 1g). Flight mode requires the EM image data to be sufficiently well aligned in 3D, as is routinely the case for neuronal tissue imaged using SBEM¹. In flight mode, the annotator focuses on centering the target cursor onto the axon or dendrite being followed, steering the orientation with the mouse or keyboard while moving forward (Supplementary Video 1). We suspected that this focusing on one intuitive egocentric visualization and interaction may accelerate annotation, since the user does not have to explicitly recenter the viewport and switch image plane orientation for processes running in off-axis directions (Fig. 1e).

To investigate whether flight mode in fact accelerates human 3D image data annotation, we trained 51 student annotators on

¹Department of Connectomics, Max Planck Institute for Brain Research, Frankfurt, Germany. ²Scalable minds UG (haftungsbeschränkt) & Co. KG, Potsdam, Germany.

³These authors contributed equally to this work. Correspondence should be addressed to M.H. (mh@brain.mpg.de) or K.M.B. (kevin.boergens@brain.mpg.de).

RECEIVED 23 MARCH 2016; ACCEPTED 17 MAY 2017; PUBLISHED ONLINE 12 JUNE 2017; DOI:10.1038/NMETH.4331

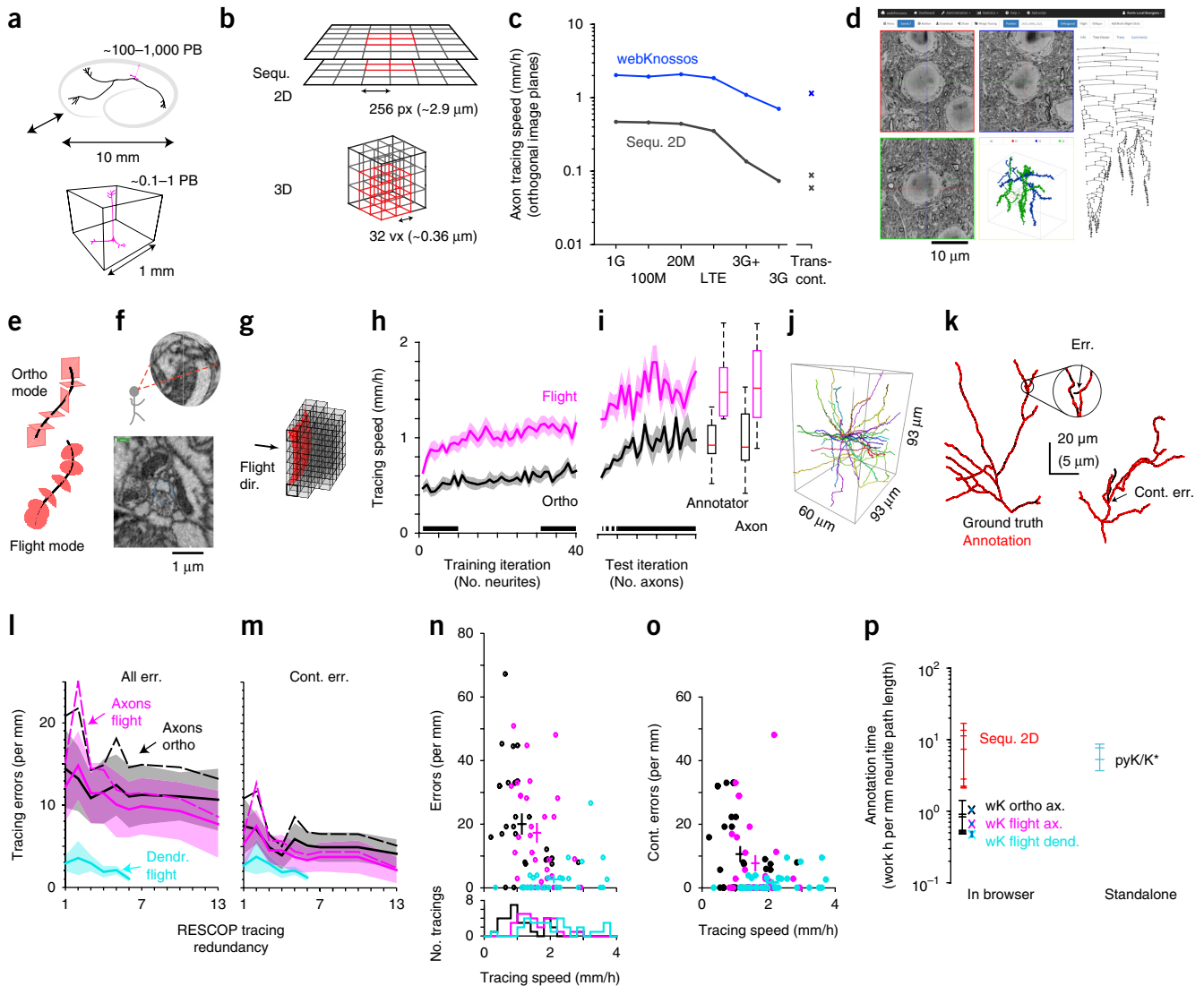


Figure 1 | In-browser 3D annotation of axons and dendrites for connectomics. (a) Sketch of mouse whole-brain 3D EM data set size (top) and 1 mm³ of cerebral cortex (bottom) compared with the extent of a typical single pyramidal neuron (dendrites, magenta; axon, black). (b) Sketch of online data delivery modes using lateral prefetching in 2D (top, Google maps, CATMAID) and in 3D (bottom, webKnossos). (c) Comparison of data transmission when following a neurite using 2D image prefetching (black, in CATMAID) and 3D data prefetching in webKnossos (blue, ortho mode) under different bandwidth and latency conditions. Transcont, transcontinental access (<https://www.openconnectomeproject.org> from Europe, custom CATMAID and webKnossos from South America, bottom to top). (d) webKnossos in-browser user interface with orthogonal viewports (*xy*, *yz*, *xz*), one 3D skeleton viewport, and the abstract tree viewer (right). (e) Sketch of viewing surface orientation (red) in orthogonal mode (top) and flight mode (bottom). (f) Flight-mode egocentric 3D image sampling on a hemisphere (top), yielding a single flight-mode data view (bottom). (g) Example of 3D prefetching in flight mode given flight direction and current position (gray, prefetched webKnossos cubes; red, flight-mode image surface). (h) Annotator speed training in ortho mode (black, $n = 25$ annotators) and flight mode (magenta, $n = 26$ annotators) on 40 neurites in cortex (randomly ordered per annotator). (i) Tracing speed test on 20 randomly selected cortical axons (including branches) performed by 26 annotators 8 weeks after training. Dashed line, first five neurites presented again from training. Solid line, 20 test axons. Box plots report tracing speed for these 20 test axons in ortho (black) and flight mode (magenta), reported over $n = 26$ annotators (left) and $n = 20$ axons (right); flight 1.51 ± 0.04 mm/h, ortho 0.96 ± 0.03 mm/h ($n = 520$, mean \pm s.e.m.). (j) Display of 20 test axons within data set boundary. (k) Illustration of tracing error measurement in one of the 20 axons—local errors (inset, less than 10- μ m path length) and continuation errors (right). Black, ground truth; red, six-fold consolidated flight-mode tracing. (l) Tracing errors reported over tracing redundancy (using RESCOP⁸) for 10 axons in ortho mode, flight mode and for 10 dendrites in flight mode (cyan), respectively. (m) same as l but only continuation errors (see k). Dashed lines in l and m, path length corrected for each RESCOPed skeleton (see Online Methods). (n) Relation between tracing speed and error rates for single-annotator reconstructions ($n = 30$) of axons and dendrites (colors as in l and m). Crosses indicate mean \pm s.e.m. (o) same as n for continuation errors only. (p) Summary comparison of annotation time requirements for neurite reconstruction. Data from c, (black, webKnossos, <https://webknossos.org>; red, CATMAID) and from i and n (crosses). Asterisk, annotation consumption documented in published work; pyK, consumption by experienced annotators¹⁰; K, KNOSSOS annotation consumption from refs. 8, 11 and 12.

40 neurites in a 3D EM data set from mouse cortex sized $93 \times 60 \times 93 \mu\text{m}^3$, imaged at $11.24 \times 11.24 \times 28 \text{ nm}^3$ using SBEM (data not shown; **Fig. 1h**). The sequence of neurites to trace was shuffled for each student (total path length of all 40 neurites, 6.11 mm; number of branch points, 258; overall branch point rate, 42 per mm). The student annotators came from our pool of trained annotators and were thus experienced in connectomic data reconstruction (**Supplementary Fig. 1**). They were randomly assigned to two groups of 25 and 26 students, respectively. One group was asked to trace in ortho mode, the other in flight mode. We supplied both groups with an 8.5-min tutorial movie (a separate movie for each group), which encouraged them to increase their movement speed whenever they felt they could go faster during reconstructions (**Supplementary Fig. 1b,d**). To enable constantly maximized tracing speed for each annotator, we automatically tracked the fraction of time during which the annotators proceeded at the preset movement speed (i.e., when holding the forward key pressed while navigating) and reminded them to increase their preset movement speed if they spent more than 75% of their tracing time constantly pressing the forward key. In ortho mode (**Fig. 1h**), annotators initially traced at $0.49 \pm 0.04 \text{ mm/h}$ (first ten processes traced, mean \pm s.e.m.) and accelerated to $0.63 \pm 0.05 \text{ mm/h}$ (last ten processes traced, $P < 10^{-3}$, Wilcoxon signed-rank test). In flight mode (**Fig. 1h**), annotators started at a speed of $0.84 \pm 0.05 \text{ mm/h}$ (faster than ortho mode, $P < 10^{-4}$, Wilcoxon rank-sum test) and were able to increase their speed to an average of $1.11 \pm 0.07 \text{ mm/h}$ ($P < 10^{-4}$, Wilcoxon signed-rank test), 1.8 times faster than ortho mode tracing ($P < 10^{-5}$, Wilcoxon rank-sum test). These data indicate that a substantial reconstruction speed gain comes from per-user speed maximization, enabled by efficient 3D data handling, and an additional gain from the egocentric flight mode data interaction.

To test whether this tracing speed can be routinely achieved for axons in mammalian cerebral cortex (**Fig. 1i,j**), we next randomly selected 20 axons from a $(2.5 \mu\text{m})^3$ region in the same data set (**Fig. 1j**) and asked students trained in ortho mode to again use ortho mode and students trained in flight mode to again use flight mode. In each group, 13 of the trained students participated in this second experiment 8 weeks after the training (their initial training performance had been indistinguishable from the whole group, $P > 0.24$, Wilcoxon rank-sum test). We first presented five of the processes reconstructed during training to calibrate the persistence of the training effects, and then we presented the 20 new axon seeds in random order to all tracers (**Fig. 1i**; note that the five neurites from the training session were not included in the final speed measurement). Tracers resumed annotation at the speed attained during training and were able to further accelerate, yielding a reconstruction speed of $0.96 \pm 0.03 \text{ mm/h}$ in ortho and $1.51 \pm 0.04 \text{ mm/h}$ in flight mode (**Fig. 1i**, mean \pm s.e.m., $n = 20$ previously unseen randomly selected cortical axons; total path length, 2.53 mm; overall branch-point rate, 39 per mm).

But were faster tracings more error prone? We next quantified the rate of errors for 10 randomly drawn axons out of the 20 test axons. For each axon, we manually counted the number of incorrect stops and incorrect continuations (**Fig. 1k**, performed by two expert annotators blinded to tracing mode, see Online Methods) and distinguished between errors yielding continuation mistakes (i.e., a premature stop or missed branch of a major part of the axon) or local errors (yielding less than $10 \mu\text{m}$ neurite

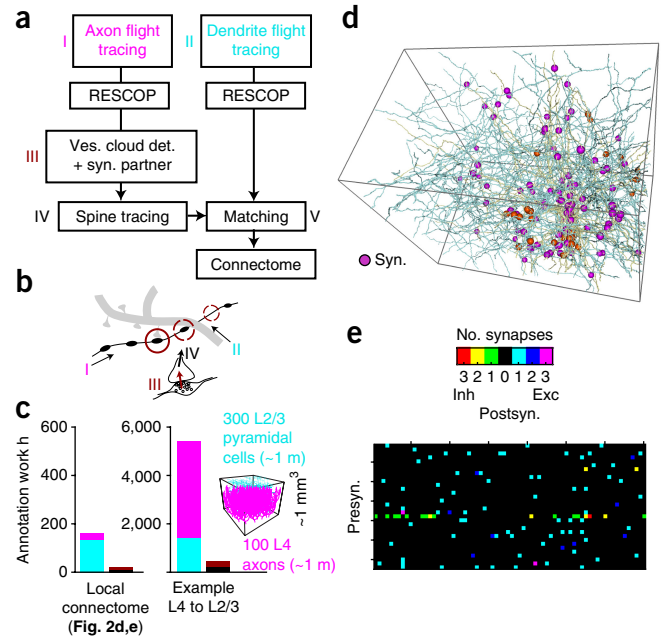


Figure 2 | Connectome reconstruction using webKnossos. **(a)** Flow chart and illustration **(b)** of connectome reconstruction steps. Ves. cloud det., vesicle cloud detection; syn. partner, identification of postsynaptic partner. **(c)** Comparison of annotation times for axons, dendrites, spines and synapses including the speed gains for neurite reconstruction in webKnossos (**Fig. 1**) for the local connectome shown in panels **d** and **e** and for a sparse example reconstruction (estimate, see inset, 100 layer-4 cortical axons innervating 300 layer-2/3 pyramidal cells within a cubic millimeter of cortex tissue, amounting to 2 m total path length). **(d)** Reconstruction of 497 dendrites and 32 axons in local SBEM data set from mouse cortex following the workflow in **a**. Colored spheres indicate excitatory (Exc., violet) and inhibitory (Inh., red) synaptic contacts (syn.). **(e)** Resulting connectome between 32 presynaptic axons (presyn.) and 497 postsynaptic (postsyn.) dendrites (only the 70 innervated dendrites are shown).

loss or neurite addition). **Figure 1l** reports the rate of errors for flight-mode and ortho-mode tracings for single-annotator reconstructions and for consolidations of multiple reconstructions of the same axon (consolidated using RESCOP⁸). We found that, first, the rate of tracing errors was not distinguishable between ortho and flight mode tracings ($P > 0.34$ for all errors, $P > 0.28$ for continuation errors, Wilcoxon rank-sum test). Second, the average error rate for single-annotator reconstructions obtained at the achieved speed in webKnossos was not worse than the error rates reported previously in mouse retina^{8,11} and cortex¹³. The rate of continuation errors (**Fig. 1m**) was 7.5 ± 3.4 per mm in ortho mode and 5.3 ± 3.0 per mm in flight mode for single-annotator tracings. We finally asked whether a speed-accuracy tradeoff could be observed in either of the tracing modes. For this we correlated the rate of errors with tracing speed in single-annotator tracings (**Fig. 1n**). No positive correlation could be found for ortho mode nor for flight mode (ortho $r = -0.5$, $P = 0.007$, flight $r = -0.4$, $P > 0.05$, Pearson's correlation). This also was true when only analyzing the continuation errors (**Fig. 1o**, ortho $r = -0.34$, $P > 0.05$, flight $r = -0.20$, $P > 0.28$, Pearson's correlation).

We thus conclude that annotators can be trained to trace cortical axons at $1.51 \pm 0.04 \text{ mm/h}$ in flight mode without a reduction in accuracy. webKnossos can support this speed online and

in browser, and it provides a 6- to 15-fold improvement over published tracing speeds, depending on the reconstructed data sets (summarized in **Fig. 1p**).

To determine tracing speeds for dendrites, we reconstructed the shafts of ten randomly chosen dendrites (1.8-mm total path length, overall branch point rate 16 per mm) and measured tracing errors as described above. Since dendrites are about three times wider in diameter¹⁴, annotators could zoom out further and fly along dendrites faster than they do along axons. We found that tracing speed was 2.11 ± 0.16 mm/h including branch-point reconstruction, and single-annotator error rates were 2.7 ± 0.69 errors per mm dendrite (**Fig. 1l–o**).

In mammalian brains, which constitute a main challenge of connectomics, about 90% of the neuronal processes are axons^{14,15}. The speed gain for representative axon reconstruction in flight mode was therefore critical for the acceleration of connectomic reconstruction in mammalian cortex. However, the reconstruction of connectomes additionally requires the identification of synapses and the assignment of postsynaptic partners to the respective neuronal cell bodies and dendrites. **Figure 2a,b** illustrates a workflow for such connectome reconstruction. In this workflow, axons and dendritic shafts are reconstructed first (including branch points). Then a synapse movie mode is activated in webKnossos; in this mode, the user can fly along the pre-traced axon and click into the postsynaptic process whenever a synapse is encountered. In the final step, the postsynaptic partner (in about 90% a spine head, **Supplementary Fig. 2a**) is seeded for annotators to trace back to the main shaft of dendrites. In this workflow, the fraction of time spent on synapse annotation is small for sparse reconstructions (7–11% for typical network reconstructions, **Fig. 2c**, **Supplementary Fig. 2b**) but is becoming more substantial for dense reconstructions, approaching about 50% of reconstruction time (**Supplementary Fig. 2b**). Since synapse detection requires only a local image classification (unlike neurite tracing), automated synapse detection is likely to soon replace manual synapse detection in dense connectome reconstructions (e.g., refs. 16–19). To exemplify a full connectomic reconstruction using webKnossos, we finally reconstructed 497 dendrites (total path length of 93.6 mm, tracing redundancy 3), and determined all synapses with 32 axons from the training set (**Fig. 1j**), 4.55 mm path length of axons, tracing redundancy 6). We detected 104 synapses in this local connectome (**Fig. 2d,e**) (total annotation time was 27.3 h for axons, 133 h for dendrites, and 19.2 h for synapses).

In summary, webKnossos accelerates human 3D data interaction for EM-based connectomics in browser by about 4- to 13-fold, which likely saturates human interaction speed with 3D EM data of nervous tissue using flight mode. While tested on well-aligned 3D SBEM data from mammalian cortex, these results are expected to be comparable for other neuropil with comparable neurite morphology (especially branch-point rates, e.g., in subcortical structures and ganglion cells in mammalian retina, see <http://www.neuromorpho.org>). Reconstructions in highly anisotropic and potentially less well-aligned image data can still profit from the speedup because of faster display rates in ortho mode tracings (see **Fig. 1i**). Thus webKnossos can serve as a versatile high-efficiency tool for 3D image data annotation in various 3D image analysis settings in connectomics and other fields.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank D.Y. Buckley, E. Klinger and P. Laserstein for helpful discussions and comments on the manuscript; C. Roome, F. Kaiser and C. Guggenberger for outstanding IT support; and J. Striebel, S. Mischkewitz, N. Ring, A. Motta, B. Staffler and M. Zauser for contributions to the code. We thank E. Arzt and A.M. Burgin for hosting KMB at the Instituto de Investigación en Biomedicina de Buenos Aires, Argentina. We thank J. Abramovich, D. Acay, M. Ali, S. Babl, A. Bansbach, N. Berghaus, D. Beyer, L. Bezenberger, F. Bock, D. Boehm, A. Budco, L. Buxmann, D. Celik, H. Charif, N. Cipta, T. Decker, J. Depnering, V. Dienst, S. Dittmer, L. Dufter, J.E. Martinez, H. Feddersen, A. Feiler, L. Feist, L. Frey, K. Friedl, A. Gaebelein, L. Geiser, D. Greco, M. Groothuis, S. Gross, A. Haeusel, S. Hain, J. Hartel, M.-L. Harwardt, B. Heftrich, M. Helbig, J. Heller, S. Hermann, J. Hoeltke, T. Hoermann, M. Hofweber, R. Huelse, R. Jakob, R. Jakoby, R. Janssen, M. Karabel, S. Kempf, L. Kirchner, J. Knauer, R. Kneissl, T. Koecke, P. Koening, M. Kolodziej, F. Kraemer, E. Kuendiger, D. Kurt, M. Kuschnerreit, E. Laube, D. Laubender, M. Lehnardt, Irene Meindl, Iris Meindl, J. Meyer, P. Mueller, N. Neupaertl, M. Perschke, M. Poltermann, M. Praeve, K. Puechler, M. Reiner, T. Reimann, V. Robl, F. Ruecker, T. Ruff, F. Sahin, D. Schelieu, V. Schuhbeck, S. Seiler, I. Stasiuk, B. Stiehl, A. Strubel, R. Thieleking, K. Trares, P. Wagner, Y. Wang, J. Wastl, A. Weber, S.S. Wehrheim, C. Weiss, M. Werr, A. Weyh, and J. Wiederspohn for annotation work.

AUTHOR CONTRIBUTIONS

M.H. initiated and supervised the project; K.M.B., M.B., T.B., N.R., T.W. and M.H. developed specifications and conceptual design with contributions by H.W.; T.B., D.B., J.F., T.H., P.O., N.R., T.W., D.W., G.W. and K.M.B. implemented the software; H.W., M.B., K.M.B. and F.D. provided data; K.M.B., M.H., H.W. and M.B. analyzed the data; M.H., K.M.B. and M.B. wrote the manuscript with contributions by all authors.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

- Denk, W. & Horstmann, H. *PLoS Biol.* **2**, e329 (2004).
- Eberle, A.L. *et al. J. Microsc.* **259**, 114–120 (2015).
- Kasthuri, N. *et al. Cell* **162**, 648–661 (2015).
- Helmstaedter, M. *Nat. Methods* **10**, 501–507 (2013).
- Saalfeld, S., Cardona, A., Hartenstein, V. & Tomancak, P. *Bioinformatics* **25**, 1984–1986 (2009).
- Schneider-Mizell, C.M. *et al. eLife* **5**, e12059 (2016).
- Rasmussen, L. *Lect. Notes Comput. Sci.* **3579**, 7 (2005).
- Helmstaedter, M., Briggman, K.L. & Denk, W. *Nat. Neurosci.* **14**, 1081–1088 (2011).
- Wanner, A.A., Genoud, C. & Friedrich, R.W. *Sci. Data* **3**, 160100 (2016).
- Wanner, A.A., Genoud, C., Masudi, T., Siksou, L. & Friedrich, R.W. *Nat. Neurosci.* **19**, 816–825 (2016).
- Helmstaedter, M. *et al. Nature* **500**, 168–174 (2013).
- Berning, M., Boergens, K.M. & Helmstaedter, M. *Neuron* **87**, 1193–1206 (2015).
- Hua, Y., Laserstein, P. & Helmstaedter, M. *Nat. Commun.* **6**, 7923 (2015).
- Braitenberg, V. & Schüz, A. *Anatomy of the cortex: statistics and geometry* Vol. 18 (Springer Science & Business Media, 2013).
- Narayanan, R.T. *et al. Cereb. Cortex* **25**, 4450–4468 (2015).
- Becker, C., Ali, K., Knott, G. & Fua, P. *IEEE Trans. Med. Imaging* **32**, 1864–1877 (2013).
- Kreshuk, A., Koethe, U., Pax, E., Bock, D.D. & Hamprecht, F.A. *PLoS One* **9**, e87351 (2014).
- Staffler, B. *et al.* Preprint at <http://biorxiv.org/content/early/2017/01/22/099994> (2017).
- Dorkenwald, S. *et al. Nat. Methods* **14**, 435–442 (2017).

ONLINE METHODS

3D SBEM data, animal experiments. The 3D EM image data was acquired using serial blockface electron microscopy (SBEM¹) from primary somatosensory cortex layer 4 of a P28 C57BL/6 male mouse, same data set as in ref. 12 ($93 \times 60 \times 93 \mu\text{m}^3$, $11.24 \times 11.24 \times 28 \text{ nm}^3$ resolution, data set 2012-09-28_ex145_07x2_new2). All animal experiments were carried out with approval of the local animal research authorities (Regierungspräsidium Oberbayern, Germany) and in accordance with the German Animal Welfare Act.

Image data transmission. For efficient volume data transmission, data is requested in small cubic packages 32^3 voxel in size ('buckets') stored at original 8-bit depth. For bandwidth-limited settings, each voxel in a bucket is trimmed to the 4 most significant bits for transmission ('4-bit mode', user activatable). Buckets are requested along a priority ranking based on the current view point and direction of movement. In ortho mode, all buckets in the plane of the active viewport are loaded first, prioritized by the Manhattan distance to the viewport center. The buckets of the next two bucket layers in direction of movement are prioritized next, with two- and four-fold-reduced priority, respectively. All buckets are loaded at the user-specified magnification level.

In flight mode the preview volume is a square frustum with basis sized 5×5 buckets, height of 2.5 buckets, oriented along the current movement direction, and top sized 4×4 buckets. All buckets fully or partially contained in this volume are requested at original magnification, prioritized by the Manhattan distance between the respective bucket center and the current viewpoint. The prioritized bucket request queue is updated on each user movement.

Measurement of 3D data transmission speed: webKnossos.

The speed of data transmission for 3D navigation in webKnossos (Fig. 1d) was measured as follows. webKnossos was run at a Hetzner (Gunzenhausen, Germany) data center on a server with the following specifications: Intel(R) Xeon(R) CPU E3-1245 V2 ($4 \times 3.4 \text{ GHz}$); 32 GB RAM; $15 \times 3 \text{ TB}$ HWRaid HDD. The EM data set was viewed in 4-bit mode in webKnossos run in Google Chrome (version 56) on a computer in the MPI for Brain Research. A neurite was picked and followed using the forward and arrow keys, keeping the forward key pressed where possible (which resulted in continuous image stream in webKnossos). To emulate reduced connectivity settings, the developer tools function of Google Chrome was used. The 'transcontinental' experiment (Fig. 1c) was performed on a computer connected to the network of the Instituto de Investigación en Biomedicina de Buenos Aires, Argentina, accessing the webKnossos instance running on the Hetzner server in Germany (see above).

Measurement of 3D data transmission speed: sequential 2D.

The speed of data transmission for sequential 2D image navigation (Fig. 1c) was measured as follows. We followed the instructions by the CATMAID authors to optimize server performance (published under https://groups.google.com/forum/#!topic/catmaid/vE_4iLrPv4). A CATMAID instance (version 2016.12.16) was installed on a server in the compute center of the Max Planck Society (Garching, Germany) with the following specifications: Xeon E5-2630 12 cores, 128 GB RAM, 10 Gb network, JBOD of $4 \times$ Intel DC S3500 240 GB SSDs, Ubuntu 14.04. Postgres and

data partitions resided on SSDs; XFS was used with noatime. The 3D image data set 2012-09-28_ex145_07x2_new2 (s. above) was converted to a series of $256 \times 256 \text{ px}$ jpg images, compressed by 75% with jpg headers removed to further reduce file size as suggested by the CATMAID authors. The data were resliced into three image series along the three cardinal directions.

The 'transcontinental' experiment was performed by accessing CATMAID on <https://www.openconnectomeproject.org> from a desktop computer in the MPI for Brain Research and by accessing the custom CATMAID instance at the Max Planck datacenter in Germany from a computer connected to the network of the Instituto de Investigación en Biomedicina de Buenos Aires, Argentina. All tests were performed in Google Chrome (version 56). The viewports in CATMAID and webKnossos were set to similar size.

Annotator training. For training annotators at high-speed annotation (Fig. 1h), 40 training neurites were selected from the cortex data set (see above). For this, a bounding box sized $4.5 \times 4.5 \times 4.2 \mu\text{m}^3$ was chosen. Then, two annotators were asked to reconstruct all processes within this bounding box. Next, each of the reconstructed processes was classified as axon, dendrite or glia. Finally, 40 of the 68 processes classified as axons were randomly selected. For each process, an expert annotator defined a starting position and a starting direction (required for flight-mode annotation).

51 annotators were trained. These annotators were randomly assigned to two groups (flight ($n = 26$) and ortho mode ($n = 25$)). Annotators were asked to watch an introductory video, which instructed them to increase their maximum velocity setting (the speed at which the annotator progresses through the data when the space key is held down continuously). Each annotator was presented with the 40 training processes in random order. The annotators' preset maximum velocities were monitored during the annotation process. If the ratio of tracing speed and preset maximum velocity was higher than 0.75 for an entire tracing, the annotator was notified via e-mail and asked to increase the maximum velocity setting for the next annotation.

Axon test reconstruction. For the test of axon reconstruction speed (Fig. 1i), the 51 previously trained annotators were asked to reconstruct 20 randomly selected axons 8 weeks after the initial training. 26 annotators signed up for this experiment (13 that had been trained on flight mode and 13 that had been trained on ortho mode). These annotators had not been faster in the final ten training iterations than the whole group of annotators ($P = 0.246$ (Wilcoxon rank-sum test) and $P = 0.250$ (t -test)) and had not been faster in the final training iteration ($P = 0.699$ (Wilcoxon rank-sum test) and $P = 0.649$ t -test).

To select a set of representative axons, a $(2.5 \mu\text{m})^3$ bounding box (located randomly within the cuboid of $15 \mu\text{m}$ edge length centered to the data set center) was chosen that did not contain a soma, and all neuronal and glial processes within this bounding box were reconstructed by one expert annotator. Then all processes were classified as axonal, dendritic or glial. Three additional expert annotators proofread the annotation. Then, 20 of the 41 processes labeled as axonal were randomly selected, and for each axon a seed position and initial orientation were defined within the bounding box.

The 26 annotators were first asked to again reconstruct five neurites from the training experiment (these five neurites were randomly chosen from the 40 training seeds and were the same

for all annotators; the sequence in which these were presented was randomized for each annotator). Then each annotator was asked to reconstruct the 20 previously unseen test axons (in a sequence randomized per annotator) in flight mode or ortho mode. After all annotators had finished, all annotations were automatically scanned for open branch points (i.e., positions at which the annotator had set a branch point flag but had forgotten to jump back to for inspection) and seed nodes with a degree of 1 (i.e., starting points which had only been traced in one direction). 17 open branch points (12 at first node) and 22 unidirectional seeds were detected (5 in flight, 17 in ortho) of 2,170 fully annotated branch points, total. In these cases the annotators were asked to go back to the task and continue their annotations. The code for this automated annotation checking is provided in **Supplementary Software 1**.

Dendrite reconstruction. 497 dendrites were reconstructed in flight mode by 47 of the annotators previously trained in ortho or flight mode (see above). Those annotators that had worked in ortho mode before were asked to watch the instruction movie for flight mode before performing the dendrite reconstructions. Annotators were instructed to set the data set quality setting in webKnossos to medium (which means that image data is displayed at lower resolution) and not to reconstruct spines.

The 497 dendrite seeds for the connectome reconstruction (Fig. 2) were drawn randomly from a set of over 2,000 dendrites that had previously been reconstructed using webKnossos. For all dendrites, the z-axis pointing toward the data set center was used as initial flight orientation.

Measurement of annotation speed. For measurement of annotation speed, the path length of a given neurite and the time it took to annotate that neurite were determined. To measure neurite path length from a skeleton annotation, the lengths of all edges within a skeleton were summed (as in refs. 8 and 11). However, this method has two caveats. First, noise in the placement of skeleton nodes will be biased to only increase apparent skeleton length, not decrease it, which could potentially lead to an overestimation of annotation speed. Second, this effect will depend on the density of placed skeleton nodes. Since in flight mode the skeleton nodes are placed automatically, the density of skeleton nodes is substantially higher in flight mode than in ortho mode tracings (flight, 6.3 ± 0.68 nodes/ μm ; ortho, 1.79 ± 0.67 nodes/ μm , measured on the ten axons used for Fig. 11–o). To account for these potential biases, we first used nonuniform rational b-spline (NURBS²⁰)-based skeleton smoothing to calibrate the effect of node placement noise on skeleton path length (**Supplementary Fig. 1f**); using the skeleton nodes as support knots, NURBS spline order (i.e., the degree of smoothing) $NO = 4$ and clamping the first and last node. We found that post-NURBS path length measurements of flight tracings are still on average $14.93 \pm 1.08\%$ (mean \pm s.e.m.) longer than ortho tracings (**Supplementary Fig. 1g**). To correct for this and for the difference in node densities between tracing modes, we scaled NO in dependence of skeleton node density D_s (in number of nodes per μm edge-based skeleton length),

$$NO = \min \left(\text{ceil} \left(c_1 \left(\frac{D_s}{c_2} \right)^{c_3} \right), N_n \right)$$

with N_n (the total number of skeleton nodes per tracing) and parameters c_1 , c_2 and c_3 . The parameters were adjusted such that the average path lengths of neurites from the training set were similar for flight- and ortho-mode tracings (resulting in $c_1 = 50$, $c_2 = 5$, $c_3 = 4$; as a result, average ortho and flight path length for a given axon agreed within $1.79 \pm 1.16\%$). See **Supplementary Figure 1g** and **Supplementary Software 1–3** for comparisons of skeleton path length measurements based on edge length and NURBS smoothing with fixed NO and variable NO , respectively. The length measurements involving NURBS smoothing with fixed NO and variable NO reduced the path length obtained from the simple edge length addition method by less than 20% (**Supplementary Fig. 1g**). The variable NO path length measurement method was used for speed measurements in the axon test set and the dendrite tracings.

To determine the annotation time of a given tracing, the administrative API of webKnossos was used (**Supplementary Software 4**) to log autosave events. Autosave events are triggered when the annotator is actively tracing within the last 30 s and the last autosave was more than 30 s ago. Therefore, during annotation work, an autosave is submitted every 30 s (but not during pauses the annotator chooses to take). Annotation time was measured as the number of autosave events times 30 s. This is also the time used for determining annotator payment.

Annotation redundancy: RESCOP. For determining the dependence of annotation error rates on annotation redundancy (Fig. 11–o), multiple annotations of the same neurite from different annotators were consolidated using RESCOP⁸. Briefly, the priors and decision boundaries were fitted separately for axons traced in ortho and flight mode (**Supplementary Fig. 1h,i**). The priors were fitted using 20 randomly selected annotations for each neurite from the training annotations (i.e., 800 annotations for ortho and flight mode, respectively, total of 865,121 edges, 866,721 nodes for computing the vote histogram, **Supplementary Fig. 1h**). The resulting decision boundaries are shown in **Supplementary Figure 1i**.

Measurement of annotation error rates. For the measurement of annotation errors, 10 of the 20 test axons (Fig. 1j) were randomly selected. For these ten axons, a ground truth annotation was generated. To do this, the axon was first traced by one expert annotator. Then, all annotations of this axon from all tracers and tracing modes were superimposed; and all locations of discrepancy between the experts' annotation and all other annotations were inspected. Remaining errors in the expert annotation were corrected. Finally, two additional experts verified the ground truth annotation independently.

Then, for each of the ten axons, three ortho-mode and three flight-mode annotations were randomly selected and their discrepancies to the ground truth annotation counted as in ref. 8 (Fig. 5c in this reference). Similarly, consensus skeletons at redundancies 2, 3, 4, 5, 6, 7, 10 and 13 were computed using RESCOP⁸ (see above) for each of the ten axons and the two tracing modes, respectively. For each redundancy, 3 sets of tracings were randomly drawn from the available 13 tracings per axon and tracing mode. Thus, together, 540 reconstructions were error analyzed.

Error analysis was done by one expert annotator and proofread by a second expert annotator. Both experts were blinded to the tracing mode in which the reconstructions were performed.

For error analysis, the reconstructions were plotted in the three cardinal projections, overlaid with the ground truth reconstruction. Errors were classified into missing branches (false negatives) and wrongly added branches (false positives). Jumps from one process into another were counted twice, once as FP and once as FN. Errors were further classified according to the length of added or omitted neurite pieces ($>10\ \mu\text{m}$, $5\text{--}10\ \mu\text{m}$, $3\text{--}5\ \mu\text{m}$, $1\text{--}3\ \mu\text{m}$; discrepancies smaller than $1\ \mu\text{m}$ were not counted, as in ref. 8). Error segments larger than $10\ \mu\text{m}$ were classified as continuation errors (Fig. 1k).

For the measurement of errors in dendrite reconstructions, 10 of the 497 reconstructed dendrites were randomly selected and error-annotated as for the axon reconstructions. Since error rates were substantially lower for dendrites than for axons (Fig. 1l–o), only redundancies 1–6 were evaluated for dendrites.

Path length of consolidated reconstructions. To make our results comparable to error rates reported in refs. 8, 11 and 13, we normalized the number of errors to the path length of the ground truth skeleton for each axon (Fig. 1l,m). However, since some annotations were shorter (due to missed neurite pieces) and others longer (due to added neurite pieces), we wanted to assure that our conclusions about error rates (Fig. 1l–o) were still correct when instead the neurite path length of the actual tracing or consolidation was used for error rate computation. To do so, we determined the path length for each RESCOP-consolidated reconstruction by generating a version of the ground truth reconstruction that matched the respective RESCOP-consolidated reconstruction (including its possible false negative errors), and we measured that skeleton's path length as described above.

Synapse annotation and connectome reconstruction. To exemplify the full analysis workflow for reconstructing connectomes using webKnossos, we used all axons from the training reconstructions (32 axons, at RESCOP redundancy 6, step I in Fig. 2a–c) and 497 dendrite reconstructions (Fig. 2d, at redundancy 3, step II in Fig. 2a–c). For synapse annotation (step III in Fig. 2a–c), a synapse movie mode in webKnossos was used (this mode is automatically activated for webKnossos tasks of type 'synapseannotation'). This was built as an extension of flight mode in which the previously reconstructed skeleton was displayed. The annotator was asked to mark synapses by setting a single node into the postsynaptic process while navigating along the axon. For the synapse movie mode, the (consolidated) reconstruction was first cut into unbranched parts, and each of these parts was presented to the annotators (see Supplementary Software 1 for the corresponding MATLAB code).

Ten annotators were trained for synapse annotation in an introductory 1-h seminar followed by two training axons for which they received immediate feedback. Then all annotators were asked to determine the output synapses of all 32 axons. To measure the precision and recall of synapse detection by student annotators, four of the axons were randomly selected, and synapse detection errors were determined by expert annotators. The student annotator with optimal precision and recall of synapse detection (precision 96%, recall 89%) was selected for the generation of the output connectome. In addition, annotators were instructed to mark axons as putative inhibitory axons if the majority of output synapses were made onto shafts.

For the axons that the best annotator marked as inhibitory, a second annotator was asked to annotate the synapses of that axon. For the annotation of inhibitory synapses, the annotator was instructed not to focus on speed of synapse annotation. The resulting synapse annotations were reviewed by an expert annotator to establish error rates for inhibitory synapse annotation (no error in 20 reviewed synapse annotations).

This procedure operated at $1.2 \pm 0.5\ \text{h per mm}$ candidate axon segment length ($n = 151$, excitatory axons; $1.8 \pm 1.0\ \text{h per mm}$ for all axons, $n = 178$).

To determine whether the postsynaptic targets of the reconstructed axons matched any of the 497 dendrites in the connectome, the annotation of the postsynaptic partner in synapse mode was used as a new seed for an annotation task (step IV in Fig. 2a–c). The annotators for these tasks were asked to only reconstruct the postsynaptic structure (in about 90% of cases a spine) in ortho mode until it entered a dendritic shaft and to then place three additional skeleton nodes in the shaft center to simplify the matching to dendrite reconstructions. This annotation had a consumption of $31.1 \pm 28.0\ \text{s}$ annotation time per spine, mean \pm s.d., $n = 975$; i.e. $2.3 \pm 1.3\ \text{h per mm}$ axon path length. Error rates of this postsynaptic process annotation were established by inspection of 30 randomly selected postsynaptic structures by an expert annotator (one wrong annotation).

To match the postsynaptic partner reconstructions (step V in Fig. 2a) with the 497 dendrite reconstructions, we finally measured the average distance d_{pd} between all dendrites and the three shaft nodes of each postsynaptic partner reconstruction, and we detected the dendrite with the smallest average distance. To determine an attachment threshold—i.e. a maximum average distance d_{pd^*} up to which a postsynaptic partner reconstruction was considered to match a dendrite reconstruction—we used a randomly chosen set of 200 partner reconstructions. In these, the distribution of d_{pd} (Supplementary Fig. 2a) indicated a threshold distance d_{pd^*} of 250 nm. To determine the error rate of postsynaptic partner matching, we evaluated the matching in an additional set of 200 randomly chosen spines and their closest dendrite (21 true positives, 1 false positive, 178 true negatives, no false negatives). All code for these procedures is available in Supplementary Software 1.

Connectome annotation time estimates. For the annotation time approximation of an example L2/3–L4 cortical connectome (Supplementary Fig. 2c), we used $1.5\ \text{mm/h}$ reconstruction speed and six-fold redundancy for axons, and $2.1\ \text{mm/h}$ reconstruction speed at three-fold redundancy for dendrites. For estimating the reconstruction time spent on synapse annotation (Supplementary Fig. 2b), two approaches for synapse annotation were considered. One, axon-based synapse annotation (Fig. 2a,b), proceeds along axons, marking synapses and identifying postsynaptic partners, which are then matched to dendrite reconstructions (see "Results"). The other, dendrite-based synapse annotation, proceeds along dendrites, reconstructing all spines along dendritic shafts. Spine annotation along dendrites proceeds at about 40 s per spine (time taken to reconstruct a spine and mark its presynaptic partner) at a spine density of about 1 per μm dendrite length. In both strategies, we assumed that only proximities of axons and dendrites at less than $5\ \mu\text{m}$ distance need to be investigated for synapses. Therefore, depending on the density of axons and

dendrites in a given reconstruction task, the length of axons and dendrites that need to be synapse searched varies, which yields an optimal strategy for any given volume density of axons and dendrites. Dendrite and axon reconstruction speed used for these estimates was 2.1 mm/h and 1.5 mm/h (at three- and six-fold redundancy, respectively).

Statistical tests. The speed and the error comparison between ortho and flight tracers and the speed comparison between the subset of annotators for the second experiment and all annotators used Wilcoxon rank-sum test. The speed comparison for annotators between beginning and end of training used Wilcoxon signed-rank test. The correlation between tracing error rate and tracing speed was computed using Pearson's correlation.

Software availability, code availability and licensing. webKnossos is available for testing at <https://demo.webknossos.brain.mpg.de> together with example data sets: the published retina data sets e2198 (ref. 21), k0563 (refs. 8, 11 and 21) and e2006 (ref. 11), a $20 \times 20 \times 20 \mu\text{m}^3$ sized subvolume of the data set 2012-09-28_ex145_07x2_new2 used for webKnossos testing (see above), and an example fluorescence data set (FD0149-2, data not shown). See **Supplementary Video 2** for an introductory video.

The webKnossos source code is provided as **Supplementary Software 5** and is also available at <https://github.com/scalableminds/webKnossos>. webKnossos is licensed under the AGPLv3 license (this applies to all source code files in **Supplementary Software 1–5** and GitHub repository). webKnossos uses the following software packages and technologies: Scala, JDK 8, Play, mongoDB, WebGL, ThreeJS, Backbone, sbt.

Data availability statement. webKnossos is openly accessible at <https://demo.webknossos.brain.mpg.de>, where data sets from retina and cortex can be browsed and annotated. The entire SBEM data set of the mouse cortex that support the findings in this study are available from the corresponding author upon reasonable request. webKnossos is open source, source code is available as **Supplementary Software 5** and at <https://github.com/scalableminds/webknossos>. All reconstructions used in this study are available in **Supplementary Software 1, 2 and 3**. Source data for **Figures 1 and 2** are available online.

- Piegl, L. & Tiller, W. *The NURBS book* (Springer Science & Business Media, 2012).
- Briggman, K.L., Helmstaedter, M. & Denk, W. *Nature* **471**, 183–188 (2011).