D. Navarro, "ascii-brain.R" (2024)

From Worm to Human: Scaling Brain Emulation

ISAAK FREEMAN

AXON@MIT.EDU

B.A. in Applied Mathematics (Neuroscience)

University of California, Berkeley (2024)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

March 2026

Authored by: Isaak Freeman

Program in Media Arts and Sciences, March 2026

Certified by: Edward S. Boyden

Y. Eva Tan Professor in Neurotechnology, *Thesis Supervisor*

Accepted by: Joseph A. Paradiso

Alexander W. Dreyfoos Professor of Media Arts and Sciences

Academic Head, Program in Media Arts and Sciences

From Worm to Human: Scaling Brain Emulation // Isaak Freeman
Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on January 23, 2026,
in partial fulfillment of the requirements for the degree of Master of Science

Abstract

Machine learning models are rapidly approaching or surpassing human performance on many metrics. In comparison, neuroscience is progressing at a slow pace. To reach the research velocity possible in software paradigms, we highlight a potential path to high-quality emulations of the brains of key model organisms such as the roundworm *Caenorhabditis elegans*, the zebrafish *Danio rerio* and the mouse *Mus musculus*.

Notably, multiple required technologies are rapidly improving. Firstly, high-resolution imaging of neuronal structure and connectivity in light microscopy or electron microscopy. Connectomics has progressed from mapping the brain of a 302-neuron nematode C. elegans to a complete adult fruit fly brain reconstruction containing approximately 140,000 neurons. Required reconstruction and proofreading cost per reconstructed neuron has fallen from roughly $16,500 to approximately $100 for zebrafish larvae.

Secondly, functional imaging has improved drastically, approaching high-resolution imaging of entire brains in the young zebrafish and cortical sections in mice containing up to 1 million neurons.

Thirdly, simulation capabilities have progressed far. Detailed emulated models of neurons and synapses are available, ranging from simple proxies to biologically accurate neurons. Under pessimistic assumptions we estimate that real-time human brain emulations require roughly 6e20 FLOP/s of compute, 700 GB memory storage per GPU, and 24 GB/s interconnect bandwidth. Mid-2020s AI clusters reach 4e20 FLOP/s, 180 GB memory per GPU, and 1.8 TB/s interconnect, with large investments into larger clusters ongoing.

Early emulation attempts are already running. Recent progress includes benchmarking-focused emulations of zebrafish brains, an entire mouse cortex, and incomplete simulations as large as 80 billion neurons, beginning to reach human-scale requirements.

Rigorous, detailed, and biologically accurate emulations of multiple model organisms are surprisingly tractable. C. elegans, larval zebrafish, and Drosophila present compelling near-term targets where structural, functional, and molecular datasets are attainable. Progress in these organisms can drive the development of technologies and structure-to-function mapping methods required to scale up brain emulation work to larger organisms, including eventually human-scale emulations.

Thesis Supervisor: Edward S. Boyden
Title: Y. Eva Tan Professor in Neurotechnology; Professor of Biological Engineering, Brain and Cognitive Sciences, and Media Arts and Sciences

Thesis Committee

Supervisor

Edward S. Boyden

Y. Eva Tan Professor in Neurotechnology
Massachusetts Institute of Technology

Reader

Kevin M. Esvelt

Associate Professor of Media Arts and Sciences
Massachusetts Institute of Technology

Reader

George M. Church

Robert Winthrop Professor of Genetics
Harvard Medical School

Preface

This thesis is my solo-authored monograph developed during my time in the Boyden Lab at MIT from 2024 to 2026.

This is a peculiar thesis: In order to argue that brain emulations are surprisingly feasible, the work underlying this piece included conducting roughly 50 in-depth interviews with researchers in connectomics, functional imaging, and neural simulation; synthesizing hundreds of papers; and drafting early material that developed the structure, estimates, and arguments presented here.

A central contribution of this thesis is the attempt to bridge fields that rarely speak to one another. Connectomics, functional imaging, neural simulation, and AI hardware each have their own literatures, metrics, and communities -- yet the feasibility of brain emulation depends on all of them simultaneously. By surveying each domain and translating their progress into common units, I aimed to make the interdependencies visible.

For instance, the Fermi estimates in this thesis reduce whole-brain simulation to concrete hardware demands: how many floating-point operations per second, how much memory per GPU, how much interconnect bandwidth between partitions, and how much raw storage for a single brain image. These order-of-magnitude calculations are not predictions; they are meant to clarify which constraints are binding today and which are likely to relax with foreseeable technology.

Similarly, the benchmarking framework developed here attempts to ground the question "how good are our simulations?" in quantitative comparisons -- not between models and biology in the abstract, but against specific experimental observables at

specific scales. The goal throughout has been to make brain emulation legible as an engineering problem with identifiable bottlenecks, rather than a speculative aspiration defined mainly by what we do not yet know.

In an extensive team effort, later collaborative outgrowth of this work was the *State of Brain Emulation Report 2025*, a multi-author companion document that extends and reworks some of the material developed here. I cite it as related work and as a later synthesis. It is available at brainemulation.mxschons.com.

> *N. Zanichelli, M. Schons, I. Freeman, P. Shiu, and A. Arkhipov, "State of Brain Emulation Report 2025," 2025. arXiv preprint: arXiv:2510.15745.*

The report was released publicly in January 2026. There are no copyright conflicts or embargo requirements for citing and discussing it here. My own contribution was to initiate the project at MIT, develop its early framing and structure, lead interviews and technical synthesis, and draft the monograph-style review work from which the later collaborative report grew.

I could not have done this alone, and I am deeply grateful to everyone who contributed.

Thanks first and foremost to Niccolò Zanichelli and Maximilian Schons, who joined in on a rainy Boston fall and a rainy Singapore winter to produce 1,000+ pages of writing between us.

Subsequently, thank you for key contributions and conversations: Adam Glaser, Adam Marblestone, Andrew Payne, Anders Sandberg, Anshul Kashyap, Anton Arkhipov, Camille Mitchell, Claire Wang, Cori Bargmann, Daniel Leible, Davi Bock, davidad, Davy Deng, Ed Boyden, Florian Engert, George Church, Glenn Clayton, Gleb, James Lin, Jianfeng Feng, Joanne Peng, Johann Danzl, Jordan Matelsky, Ken Hayworth, Kevin Esvelt, Konrad Kording, Lei Ma, Leopold Aschenbrenner, Logan Thrasher Collins, Michael Andregg, Michael Skuhersky, Michał Januszewski, Mojtaba Tavakoli, Niko McCarty, Ons M'Saad, Patrick Mineault, Phil Shiu, Quilee Simeon, Richie Kohman, Sam Altman, Srini Turagas, Tim Farkas, Tomaso Poggio, Viren Jain, Wenlian Lu, Yangning Lu, Zeguan Wang.

Lastly, my heartfelt thanks to Sarra Shubart, Kevin Esvelt, George Church, and especially Ed Boyden for the relentless support during graduate school.

You’re incomparable.

This is necessarily imperfect and incomplete. All mistakes are mine.

Welcome to the ride.

Isaak Freeman
axon@mit.edu
Cambridge, MA, 2026

Introduction

> *There lies the port; the vessel puffs her sail:*
> *There gloom the dark, broad seas. My mariners, …*
> *It may be that the gulfs will wash us down:*
> *It may be we shall touch the Happy Isles,*
> - Ulysses

Humanity is racing towards strange, unknown waters.

AI has approached, achieved or surpassed human-level performance on many conceivable benchmarks. This spans many cognitive domains: Chess (Campbell et al., 2002), Go (Silver et al., 2016), reasoning (OpenAI, 2024), parts of mathematics (Trinh et al., 2024), software development (Schluntz et al., 2024), natural language (Brown et al., 2020), reading comprehension (Rae et al., 2021), visual reasoning (OpenAI, 2024), protein structure prediction (Abramson et al., 2024), competitive programming (OpenAI, 2025), graduate-level science (Rein et al., 2023), abstract reasoning (ARC Prize, 2024), and gold-medal standard in mathematics (Google DeepMind, 2025; OpenAI, 2025). Meanwhile, investment into AI is approaching $100B yearly (Perrault & Clark, 2024). By 2030, computational power available for AI training runs is likely to scale to 2e29 FLOP or beyond — more than 10,000 times more than was used for GPT-4 (Sevilla et al., 2024). AI progress will not halt anytime soon, especially as nation states compete to avoid falling behind adversaries that possess advanced AI *and* being technologically outpaced by rivals (Aschenbrenner, 2024, Tong & Martina, 2024).

Stephen Hawking warned that AI represents either humanity's greatest achievement or an existential threat (Hawking, 2016), a sentiment resounding with Geoffrey Hinton, Dario Amodei, Yoshua

Bengio, Demis Hassabis, Sam Altman, and others signing a statement that (Hinton et al., 2023):

> *"Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."*

An Alternative Path to AI

There is an alternative path to advanced AI that is more *human*: Using detailed neuronal maps and recordings of entire human brains to emulate a human brain end-to-end.

Brain emulation.

Connectomics -- mapping the ground-truth wiring of brains down to synapses -- has been advancing nonlinearly. For decades, only the small connectome of the roundworm *C. elegans* with 300 neurons was completed. In October 2024, a team of over 200 scientists published the complete wiring diagram of an adult fruit fly brain: 139,255 neurons connected by 54.5 million synapses (Dorkenwald et al., 2024). This represents the largest complete brain ever mapped. (Murthy et al., 2024, Dorkenwald et al., 2024, Schlegel et al., 2024, Lin et al., 2024, Berg et al., 2025)

In functional imaging, the number of neurons being simultaneously recorded has doubled roughly every 7.4 years (Stevenson & Kording, 2011; Urai et al., 2022), and has since accelerated with progress toward whole-brain imaging in optically transparent larval zebrafish. For voltage imaging, whole-brain recordings

from roughly one-third of larval-zebrafish neurons have now been demonstrated (Wang et al., 2023), while the ZAPBench dataset (Lueckmann et al., 2025) demonstrated calcium imaging of more than 70,000 neurons across nearly the entire larval zebrafish brain at approximately 1 Hz.

Five months after the fruit fly connectome was completed, the Allen Institute simulated nearly 10 million mouse neurons on Supercomputer Fugaku--the closest approximation to whole-brain mammalian simulation yet achieved. (Kuriyama et al., 2025)

In December 2024, a Chinese team published that they had spun up 14,012 GPUs to run an 86-billion-neuron simulation of a human brain (Lu et al., 2024). It was too crude and oversimplified to deserve the term “emulation” -- implying capturing with high-accuracy the causal structures of a human brain -- but it was the biggest attempt at a human-scale simulation in history.

There is a massive gap between the available data and available compute. Computationally, human-scale brain simulations become feasible thanks to infrastructure built for AI. While not informed by ground-truth connectomics, simple human-scale simulations with 60-80 billion neurons already have been run on exascale compute clusters (Yamaura et al., 2020, Lu et al., 2024). This is both impressive proof-of-concept and also highlights the dire need of more ground-truth data to improve these models.

These milestones usher in a new era where mapping, simulating, and emulating brain neural circuits become engineering challenges, not speculative research.

Progressing within their rigorous subfields, these developments show that animal and human-scale simulations are increasingly plausible. This is driven by improvements in electron microscopy,

expansion microscopy, whole-brain imaging, data processing, and simulation approaches. Combined with tractable work on challenges and validation in small organisms such as the zebrafish, followed by larger-scale funding comparable to the Human Genome Project, a full reconstruction and subsequent simulation of a human brain is plausible within this century.

Why Pursue Brain Emulations?

Highly-accurate simulations offer additional value for neuroscience. Rapid, verifiable, and replicable experiments using digital reconstructions of, for instance, a simulated mouse could reveal mechanisms underlying memory, cognition, and emotion. Hypotheses about circuit function, pharmacological interventions, or lesion effects could be tested computationally before wet-lab validation--accelerating iteration, improving replicability, and enabling verification (Markram, 2006; Eliasmith et al., 2012). Computational brain models can overcome principal barriers that currently limit research: the immune system (Polikov et al., 2005), the skull (Cramer et al., 2021), lack of whole-brain access (Marblestone et al., 2013), non-reproducibility (Baker, 2016, Cobey et al., 2024), and the underlying slowness of biology and biological research (Rodriques, 2022).

Brain diseases and disorders affect more than 3.4 billion people (GBD 2021 Nervous System Disorders Collaborators, 2024). More speculatively, highly-detailed computational models would also be helpful to fulfill the promise of computational neuroscience: Better diagnostics and better treatments for brain disorders. Digital brain models could serve as testbeds for non-invasive neuromodulation (TMS, tDCS, ultrasound, deep brain stimulation), enabling systematic optimization of stimulation parameters for conditions like depression or Parkinson's disease before clinical application.

Functional imaging and basic science. Even without full emulation, partial models of neural circuits in flies, fish, and mice would constitute an unprecedented resource for basic neuroscience -- enabling in silico experiments on how specific cell populations generate and modulate behavior.

Subjective experience and consciousness research. Running digital neural systems at varying levels of fidelity could provide empirical traction on questions about consciousness and personhood that remain largely philosophical today (Tononi, 2004; Koch et al., 2016). If a digital system reproduces the functional signatures associated with conscious states, this would constitute a novel class of experimental evidence.

Longevity and digital continuity. If high-resolution brain emulations preserve the causal structure underlying personal identity, they could in principle extend conscious existence beyond biological limits -- a fundamentally different approach to longevity than molecular-level interventions (Sandberg & Bostrom, 2008). This pathway is also potentially compatible with revival from cryopreserved tissue.

Brain-inspired AI architectures. Detailed understanding of biological neural computation -- including energy efficiency (~20 W for a human brain vs. megawatts for large AI systems), existing value and alignment systems, and few-shot learning -- could inform novel hybrid architectures combining biological principles with current machine learning approaches.

Civilization resilience. Digital minds, unlike biological ones, can be backed up, copied, and distributed across physically separated compute infrastructure -- providing a form of civilizational redundancy against catastrophic risks (Bostrom, 2013).

Robotics and embodiment. Human motor planning and sensorimotor integration remain far ahead of current robotic control systems. Brain emulations could provide the complex motor and cognitive priors needed for robots operating in unstructured environments.

Lastly, the alternative path to aligned AI mentioned in the introduction. Current AI systems are trained on human-generated data but are not structurally human. A whole-brain emulation would be human by construction -- inheriting human values, motivations, and cognitive architecture -- potentially offering a qualitatively different approach to the alignment problem (Sandberg & Bostrom, 2008; Yampolskiy, 2015).

Any such effort will be enormously hard, be it a highly-detailed zebrafish emulation, mouse emulation, or even a human emulation.

From Worm to Human

Few organizations have pursued ground-truth brain simulations. The most notable effort is the Human Brain Project, proposed as early as 2009 (Markram, 2009) -- long before the advent of high-throughput electron microscopy, expansion microscopy, or large-scale single-neuron functional imaging. Without these tools, progress remained limited. HBP was marred by management controversy and by being launched before the underlying imaging and recording technologies were mature (Sample, 2014), perhaps too early for its time.

Multiple concrete and underfunded projects are out there in the open.

While optical limits make single-neuron whole-brain imaging more challenging in larger animals such as mice or monkeys, promising imaging systems are advancing in *C. elegans* and the larval stage of the zebrafish *Danio rerio*. The larval zebrafish is the largest model organism similar to mammals in which whole-brain

single-neuron imaging has been achieved and will improve drastically in the near-term -- as a vertebrate with roughly 100,000 neurons (Ahrens et al., 2013). Combined with connectomics efforts, this makes the *C. elegans* and the zebrafish promising stepping stones to validate ground-truth structural data combined with functional data leading to high-resolution simulations.

Reconstruction and proofreading remain the dominant cost of connectomics projects. Even with machine learning segmentation approaching expert-level accuracy, human proofreaders must still verify millions of putative synapses and trace ambiguous neurites -- a process that consumed over 33 person-years for the Drosophila connectome alone (Dorkenwald et al., 2024). For a mouse brain roughly 1,000 times larger, proofreading costs at current rates would be prohibitive without further automation. AI-assisted proofreading pipelines that flag only the most uncertain regions for human review are essential for scaling beyond insects.

Molecular data presents another major gap. Connectomes capture wiring but not the receptor types, ion channel densities, neuromodulator concentrations, and gene expression profiles that determine how each synapse and neuron behaves. Techniques such as spatial transcriptomics, in-situ sequencing, and multiplexed protein imaging are beginning to fill this gap for small tissue volumes, but scaling these methods to whole brains remains an open challenge. Without molecular ground truth, simulations must rely on statistical priors and literature values -- a significant source of uncertainty in current models.

State-of-the-art connectomic-driven simulation work is underfunded. For example, for simulating the worm C. elegans, we know of approximately one underfunded nonprofit driven by a few part-time volunteers (Szigeti et al., 2014), work by a graduate student at MIT on related system-identification approaches (Simeon et al., 2024), and two groups at Princeton and in Beijing attempting C. elegans simulations (Simeon et al., 2024; Zhao et al., 2024; Creamer et al., 2024).

These advances, taken together, suggest that brain emulation is transitioning from a speculative aspiration to an engineering challenge with quantifiable milestones. The question is no longer whether we can map and simulate brains, but at what scale, resolution, and cost -- and how quickly the underlying technologies will continue to improve.

By definition, large-scale *moonshot* projects will sometimes fail, sometimes succeed, and often fade into history as a mix of success and issues. For example, the Human Genome Project was repeatedly criticized as impossible and wasteful by experts in the 1980s (Collins & McKusick, 2001). Yet, after initial validation in smaller organisms, humanity decided to attempt to sequence the entire human genome and succeeded. The Human Genome Project spent ~$5B (in 2021 dollars) and took 13 years (Tripp and Grueber, 2011), leading to the discovery that most functional sequences in the human genome do not encode proteins and a drastic increase in the number of approved drugs with known protein targets (rising from less than 50% before 2001 to nearly 100% for drugs licensed in the US after the HGP) (Gates et al., 2021).

For comparison, the Manhattan Project cost over $30B (inflation adjusted) (Schwartz, 1998) and took 3 years to develop the atomic bomb (Groves, 1962). Apollo cost roughly $257B in 2020 dollars (Dreier, 2022) and took about eight years from President Kennedy's May 25, 1961 commitment to the Apollo 11 landing in July 1969 (Loff, 2015).

This thesis surveys the current state of the research that could lead to megascale emulation projects.

Scaling Structural Imaging

Starting from first principles, it is desirable to know:

1. where neurons are,
2. what kind they are,
3. what shape they are
4. what proteins and molecules are present,
5. where they form synapses,
6. what types of synapses, and
7. what their activity is.

Additional desirable information includes other cells and their functions, such as astrocytes or glial cells.

As for structural information, the most important molecule types in brains include proteins, RNA, DNA, and lipids. DNA varies minimally between cell types and DNA sequencing is easy. Lipids bound the neuron, defining morphology as well as influencing function. RNA expression is highly varied and defining for the cell's function. Proteins--for example in the form of ion channels, receptors, neuropeptides, and transcription factors--implement much of cellular activity depending on spatial context.

When considering structural imaging, the following results or axioms guide us:

Synapses are between 200 and 800 nm wide (Sheng and Kim, 2011), the synaptic cleft -- the gap between the pre- and postsynaptic cells -- is about 20-30 nm wide (Yang & Annaert, 2024), and unmyelinated axons are between 50 and 1,000 nm (Helmstaedter et al., 2013). The C. elegans brain is on the order of 10^-6 mm3. A typical Drosophila brain is up to roughly 0.1 mm3 (Rein et al., 2002). The mouse brain contains 7.5e7 neurons (Herculano-Houzel

et al., 2006) in 420 mm3 (Badea et al., 2007). The human brain is 1,400,000 mm3 containing ~8.6e10 neurons (Azevedo et al., 2009). The largest brain in a living organism is the 8,000,000 mm3 brain of a sperm whale (Kojima, 1951). Volume scales cubically with side length, making imaging a mouse or human brain non-linearly more difficult than completed efforts in C. elegans or Drosophila.

Electron Microscopy

Electron microscopy development reached resolutions less than ten nanometers in the 1930s and broke 3 nm resolution in 1944 (Haguenau et al., 2003). Thus, for most of neuroscientific history, reliably resolving synaptic structures required electron microscopes (Emmons, 2015).

Using electron microscopy, the first complete connectome of an organism was achieved in 1986. White et al. reconstructed 302 neurons of the hermaphrodite *C. elegans*, including more than 7500 chemical synapses, gap junctions, and neuromuscular junctions as a composite across multiple individual worms (White et al., 1986). Since *C. elegans* neural wiring does not differ drastically between individual worms of the same sex (Varshney et al., 2011), this initial connectome has continuously improved. Using similar electron microscopy methods, the roughly 80 additional neurons (including gap junctions) specific to the male *C. elegans* -- which are mostly present in the posterior nervous systems to execute mating functions -- were mapped in 2012 (Jarrell et al., 2012).

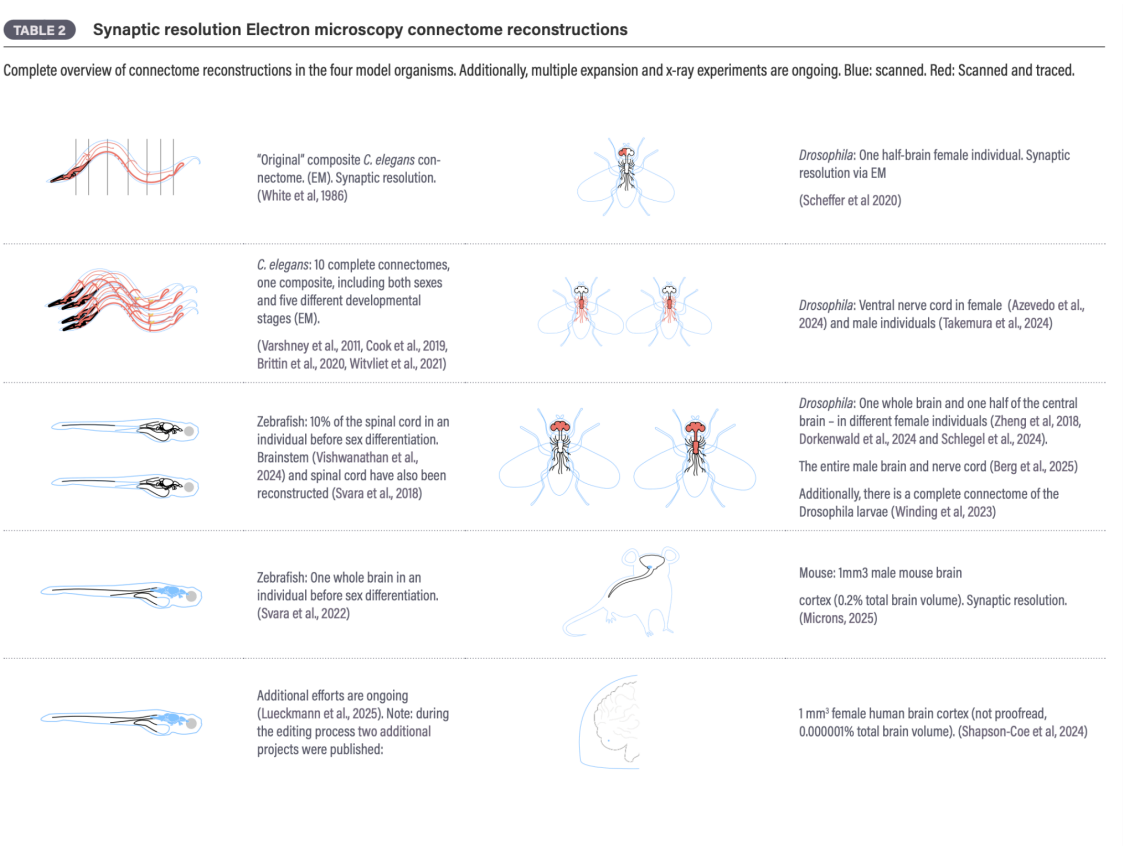

TABLE 2 **Synaptic resolution Electron microscopy connectome reconstructions**

Complete overview of connectome reconstructions in the four model organisms. Additionally, multiple expansion and x-ray experiments are ongoing. Blue: scanned. Red: Scanned and traced.

	"Original" composite *C. elegans* connectome. (EM). Synaptic resolution. (White et al, 1986)		*Drosophila*: One half-brain female individual. Synaptic resolution via EM (Scheffer et al 2020)
	C. elegans: 10 complete connectomes, one composite, including both sexes and five different developmental stages (EM). (Varshney et al., 2011, Cook et al., 2019, Brittin et al., 2020, Witvliet et al., 2021)		*Drosophila*: Ventral nerve cord in female (Azevedo et al., 2024) and male individuals (Takemura et al., 2024)
	Zebrafish: 10% of the spinal cord in an individual before sex differentiation. Brainstem (Vishwanathan et al., 2024) and spinal cord have also been reconstructed (Svara et al., 2018)		*Drosophila*: One whole brain and one half of the central brain – in different female individuals (Zheng et al, 2018, Dorkenwald et al., 2024 and Schlegel et al., 2024). The entire male brain and nerve cord (Berg et al., 2025) Additionally, there is a complete connectome of the Drosophila larvae (Winding et al, 2023)
	Zebrafish: One whole brain in an individual before sex differentiation. (Svara et al., 2022)		Mouse: 1mm3 male mouse brain cortex (0.2% total brain volume). Synaptic resolution. (Microns, 2025)
	Additional efforts are ongoing (Lueckmann et al., 2025). Note: during the editing process two additional projects were published:		1 mm^3 female human brain cortex (not proofread, 0.000001% total brain volume). (Shapson-Coe et al, 2024)

Figure 1. State of electron microscopy connectomics. Adapted from Zanichelli et al. (2025).

A comprehensive effort re-mapped the complete connectomes of both sexes, correcting errors, charting new areas, and introducing synaptic strength weighting (Cook et al., 2019). More recent work reconstructed C. elegans head connectomes across five developmental stages, useful for developmental studies (Witvliet et al., 2021; Brittin et al., 2021). C. elegans connectomes are therefore continuously increasing in quantity, accuracy, and quality as alignment, neuron identification, and reconstruction pipelines improve (Brittin et al., 2021; Witvliet et al., 2021). C. elegans is the organism where we have the most individual brains reconstructed: roughly 10 individual adult worms (Brittin et al., 2020; Cook et al., 2019; Jarrell et al., 2012; Brittin et al., 2018; Varshney et al., 2011; Witvliet et al., 2021). In C. elegans, research benefit plateaus with large populations because

the worms' brains are highly similar -- they are highly stereotyped.

In the last few years, though, *Drosophila* connectomic efforts have borne fruit. Using a custom, high-throughput, serial-section transmission electron microscopy (ssTEM), Zheng *et al.* imaged a complete fly brain in 2018 (Zheng et al., 2018). As a proof of concept, they reconstructed 120 neurons in that imaged volume.

In 2020, Scheffer *et al.* achieved a record-breaking, dense reconstruction of the central portion of a fly's brain, including more than 22,594 neurons and more than 20 million synapses (Scheffer et al., 2020). Their work represents a significant fraction of the fly brain's roughly 140,000 total neurons.

Subsequently, the FlyWire consortium completed a fruit fly connectome. (Dorkenwald et al., 2024, Schlegel et al., 2024, Lin et al., 2024) This is the largest complete brain reconstructed as of 2025.

With a full fruit fly connectome in-hand, the field is increasingly aiming higher. A multiday workshop series (National Institutes of Health, 2021) and independent investigation by the Wellcome Trust (Bosch et al., 2023) have begun roadmapping an ambitious goal: reconstructing an entire mouse brain. This mouse brain project, funded by the NIH with multiple millions USD per year, was carefully scrutinized and eventually operationalized with specific milestones, including scanning 10-15 mm^3 of mouse brain tissue.

In parallel, electron microscopy connectomics has advanced into many other organisms. Notably, the field has achieved a fully reconstructed fraction of a larval zebrafish brain with 208 neurons (Hildebrand et al., 2017), complete *imaging* of a larval zebrafish, with 2,589 axons traced (Svara et al., 2022), mapping of 15,000 neurons in the ventral nerve cord of *Drosophila* (Azevedo et al., 2024), and also a 90 x 90 x 60 micrometer cube of mouse cortex, covering 0.0005 mm^3. This block included almost

400,000 synapses and, most impressively, was completed in 4,000 work-hours. (Motta et al., 2019)

Notably, a complete cubic millimeter of human cortex was imaged and segmented, though not proofread (Shapson-Coe et al., 2024). This included 49,000 neurons and 150 million synapses, eclipsing the synapse count of an entire fly brain threefold. Ongoing projects aim to reconstruct 10-15 mm^3 of the mouse brain in the upcoming years (Manning, 2023, Januszewski, 2023), as well as a full connectome of a zebrafish (Lueckmann et al., 2025).

High-Throughput EM

Imaging an entire mouse brain will require a mega-scale project in biology.

A consortium centered around a mouse brain plans to use two 91-beam electron microscopes (Princeton Neuroscience Institute, 2023, Manning, 2023) to image a 10-15 mm^3 cube of mouse tissue -- a fiftieth of a mouse brain. According to the Wellcome Trust report on connectomics (Bosch et al., 2023), twenty electron microscopes running in parallel would require five years to image an entire mouse brain. Our current electron microscopes are not fit for the vast scale of a human brain (Collins et al., 2025), which is ~2,800 times larger than a mouse.

A quick napkin-math estimate illustrates the scale. A mouse brain volume of 500 mm3 corresponds to 5×10^20 nm3. At 10 nm isotropic resolution, that implies 5×10^17 voxels. Assuming present-day electron-microscope imaging rates of 100-200 million voxels per second (take 1.5×10^8 voxels/s as a midpoint), imaging the whole brain would require about 3.3×10^9 seconds, i.e. roughly 104 years for a single microscope running continuously.

With a deadline of 5 years, this gives (104/5 =) the estimated ~20 microscopes running in parallel.

Twenty microscopes running in parallel is a large feat, but feasible at large-project scale, especially considering that existing academic and governmental facilities already operate on the order of 5-10 microscopes (Princeton Neuroscience Institute, 2023; National Research Council Canada, 2023; School of Materials Science and Engineering, Tsinghua University, 2020). The Wellcome Trust report concludes that a mouse connectome would be an enormous feat costing billions, but still within the reach of large-scale government projects or efforts such as the consortium behind The Mind of a Mouse (Abbott et al., 2020).

However, electron microscopy has been improving rapidly even in the absence of intense demand. Peak imaging speeds of up to 1 GHz have been achieved as early as 2015 (Eberle et al., 2015) and decade-old electron microscopes are far from reaching physical or engineering limits. Public data on market size is scarce, but major producers (Zeiss, Thermo-Fisher) manufacture only hundreds to thousands of bioscience EMs annually, suggesting a relatively small market.
We are far from the engineering limit for beam lines on a microscope. As early as 2015, Zeiss produced 61-beam SEMs capable of approaching peak imaging rates of a GHz or so (Zeidler et al., 2015), with burst rates of up to 1.8 GHz and effective imaging rates of 0.3 GHz achieved in 2024 (Zheng et al., 2024).

This prototype has a 250,000x250,000 nm (hexagonal) field of view with a 5x5 nm resolution, greater than the resolution typically necessary for connectomics. It was made for the semiconductor industry, without any customizations for biology which could potentially further increase performance. This prototype is outdated by more than half a decade: Much state-of-the-art technology is proprietary and confidential. It is unclear how far this technology has advanced since, though experts we interviewed commonly think that up to 1000 beams are not implausible (Sandberg and Bostrom, 2008).

Directly extrapolating from best-case 61-beam speeds to the 2019-version with 331 beams gives:

1 GHz * 331/61 = ~5.4 GHz

However, several other factors, including data transfer, beam alignment and stage movement would likely prevent a linear scaling of throughput with the number of beams, particularly with respect to effective imaging rates achieved over months to years of operation (Rangoli, 2024).

Major EM manufacturers have not faced the demand necessary to produce such high-throughput microscopes. Responding to the demand seen by mouse connectome efforts, sufficiently large orders, or major projects towards human-scale connectomics, we would not be surprised if major innovation is latent in the electron microscope space. We estimate that improvements, most importantly in automated sample handling, reliability, and uptime could plausibly increase the speeds farther by 2-fold to 5-fold.

Assuming some improvement of traceability of larger voxels, we use 16nm as the minimum here.

Based on this, a back-of-the-envelope calculation gives:

1.4e24 nm3 / (16*16*16) nm voxel size / 1e10 (ie 10 GHz) / 0.5 (50% uptime) / 60 / 60 / 24 / 365.25 / 10 years = ~200 microscopes

I.e. 200 of such high-throughput 10-GHz mass-produced SEMs would be needed to image a human brain in roughly 10 years.

With as little as a single major private or government order incentivizing manufacturers to move from a craftsmanship approach to mass-produced economies of scale, cost could drop steeply. Given the improved performance but improved scale, assuming cost of $1M per scope, this gives 200 microscopes * $1M = $200M USD

for the cost of microscopes to image a human brain in roughly 10 years.

Slicing and Dicing

Before imaging, electron microscopy requires brain samples to be thin and stained with heavy metals (Peddie and Collinson, 2014; Bosch et al., 2023). Naturally, brains are neither.

To achieve slice thickness of ca. 50 nanometers (ssSEM) and 5 nm (FIB-SEM), the brain needs to be processed. Extensive sample preparation involves 1) fixation with aldehydes, typically a combination of glutaraldehyde and paraformaldehyde (Bosch et al., 2023), 2) staining with heavy metals to enhance contrast, typically osmium tetroxide (Tapia et al., 2012), and 3) embedding in an epoxy resin (Tegethoff and Briggman, 2024). Between these steps, the brain needs to be partitioned into small slices.

As the Wellcome Trust Report notes (Bosch et al., 2023), a vibratome could partition the brain directly after fixation, e.g. into 13 thick slices for a mouse brain. This is followed by heavy-metal staining and resin embedding. Due to the roughness of vibratome cuts leading to artefacts such as chatter and knife marks, tracing across cuts remains difficult. Yet, in our interviews, experts expressed optimism for developments in hot-knife ultra-sharp vibratome techniques (National Institutes of Health, 2021) improving cut quality sufficiently to allow traceability across thick slices.

Alternatively, one can attempt to fixate, stain, and embed an entire brain before any cutting. A 12-step protocol leveraging timelapse micro x-ray imaging to track staining progress recently achieved this for a whole mouse brain, but diffusion speed currently limits scalability to larger samples, with the protocol requiring ~6.4 weeks for the mouse brain and an extrapolated 2

years for the marmoset one (Lu et al., 2023). Other approaches have been proposed, such as whole-brain fast cryofixation to preserve structure while avoiding detrimental ice crystals. (Bosch et al., 2023)

Despite technical challenges such as sectioning artefacts, whole-brain sample preparation for EM is achievable with major investment in whole-brain staining protocols and precision sectioning methods.

Automating sample preparation would substantially reduce costs. GridTape TEM, which deposits serial sections onto TEM-compatible slot grids for parallelized multi-beam imaging, has already increased throughput for ssTEM pipelines significantly (Phelps et al., 2021). For large tissue volumes, hot-knife techniques subdivide resin-embedded samples into manageable slabs before serial sectioning (Peddie et al., 2022) -- an approach our interviewees expressed optimism about for improving cut quality across thick slices.

More broadly, our report estimates that optimizing beamlines, automated sample handling, and related infrastructure could increase effective data acquisition speeds by an order of magnitude (Zanichelli et al., 2025). The analogy to semiconductor fabs is apt: once nanometer-scale sample handling moves from artisanal to robotic, per-unit costs drop steeply and reliability improves even faster.

Neuron Reconstruction

After image data has been collected, neuron reconstruction begins. The goal of reconstruction is to recreate the 3D connectome from microscopy data. Reconstruction consists of three main steps: data preparation, tracing, and proofreading.

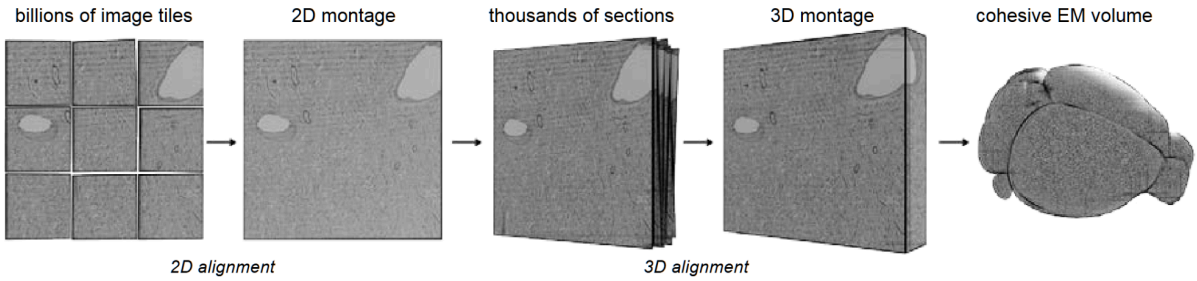

Figure 2. A visualization of the image registration step during neuron reconstruction. Source: Wellcome Trust report (Bosch et al., 2023).

Registration (aligning adjacent scans) is a critical bottleneck during data preparation, particularly for imaging approaches that involve thin-sectioning such as serial section electron microscopy. Misalignments are caused by artifacts in EM scans and slight movement/rotation of samples across imaging rounds. Robust registration ensures spatial consistency of structures, and poor alignment during the data preparation step is the dominant cause of errors in automated neuron reconstructions (Popovych et al., 2022). Recent advancements in machine learning have significantly improved alignment accuracy by addressing (non-exhaustively) artifact removal, de-warping, and noise-reduction, in some cases reducing genuine misalignments to 0.06% (Popovych et al., 2022, Scheffer et al., 2020).

Next: Segmenting cell boundaries and trace neuron skeletons. Historically, tracing was performed entirely manually, with the tracing of C. elegans connectome famously requiring over 15 years of painstaking work (Emmons, 2015). However, the prohibitive costs - over 11.2 hours per neuron on average (Zheng et al., 2018) - have driven the development of algorithmic methods designed to either assist human tracers or automate the process entirely.

Flood-filling networks (FFNs) represent a critical innovation in automated segmentation, achieving over a 10-fold accuracy

improvement versus prior methods, though at high computational cost (Januszewski et al., 2018). Reducing the compute cost of high-quality segmentation remains an active engineering target.

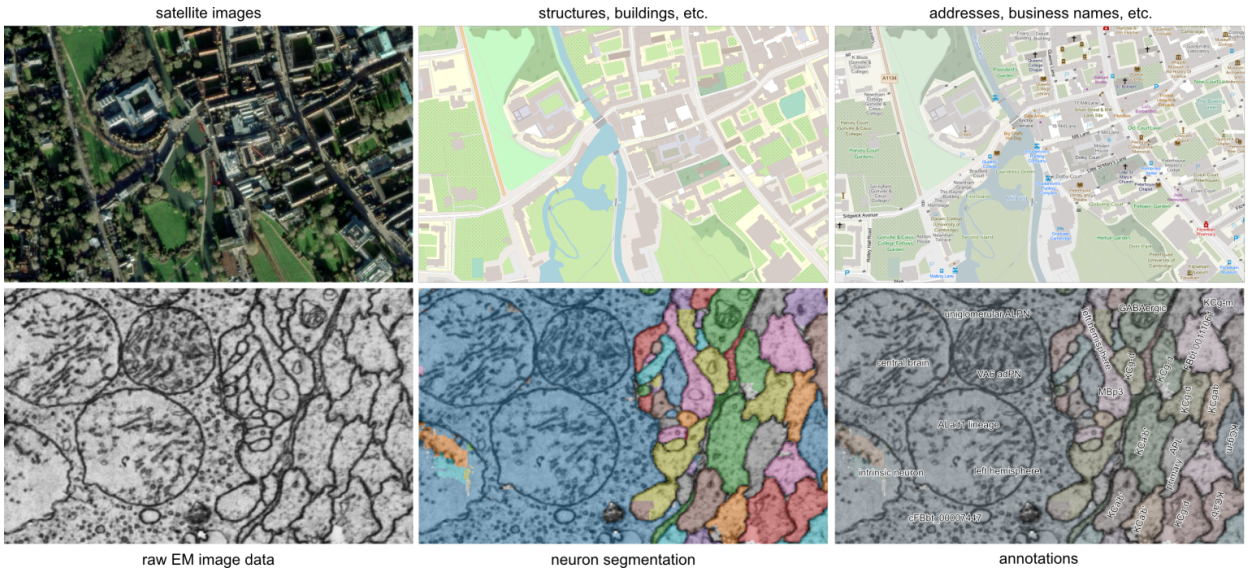

Figure 3. Brain maps at key processing stages: after registration (left), after segmentation and tracing (center), and after annotation (right). Source: Schlegel et al. (2024).

Once tracing is complete and the connectivity matrix is reconstructed and proofread, one last crucial step involves annotation, as navigating a connectome could be challenging otherwise. This process can include many steps, including synapse detection (Staffler et al., 2017), cell-type annotation (Schlegel et al., 2024), neurotransmitter classification (Eckstein et al., 2024) and more (Bazinet et al., 2023)--all of which could help downstream detailed brain emulations.

All in all, this led to the cost of neural reconstruction falling rapidly. Reconstructing a single C. elegans neuron cost roughly $16,500 in the 1980s; by 2025, a Drosophila neuron costs about $214 and a zebrafish neuron around $100 (Zanichelli et al., 2025).

Mammalian neurons remain expensive at $500-1,000 each. A billion-dollar mouse connectome -- ambitious but plausible for a

large government project -- would require getting to about $10 per neuron. A human connectome at the same budget demands roughly $0.01 per neuron, a gap of five orders of magnitude from current costs. But the trend is accelerating: proofreading still accounts for the large majority of EM connectomics costs (Bosch et al., 2023; Park et al., 2025). The structural pipeline for a mouse connectome is increasingly gated by funding and coordination rather than by physics. For a detailed breakdown of methods, costs, and projections, see our State of Brain Emulation Report (Zanichelli et al., 2025).

We note substantial optimism in our interviewees about investment into neuron reconstruction being able to substantially resolve most current bottlenecks, especially due to progress in AI and barcoding methods.

Light Microscopy

Electron microscopy (EM) is constrained by physics to a resolution limit of 0.12 nm (Penczek, 2010), with life science EMs typically achieving ~3-30 nm lateral and ~20-50 nm axial resolution (Peddie and Collinson, 2014). In contrast, traditional light microscopy is constrained to a lateral resolution of ~250 nm and an axial resolution of ~550 nm by the diffraction limit (Huang et al., 2010).

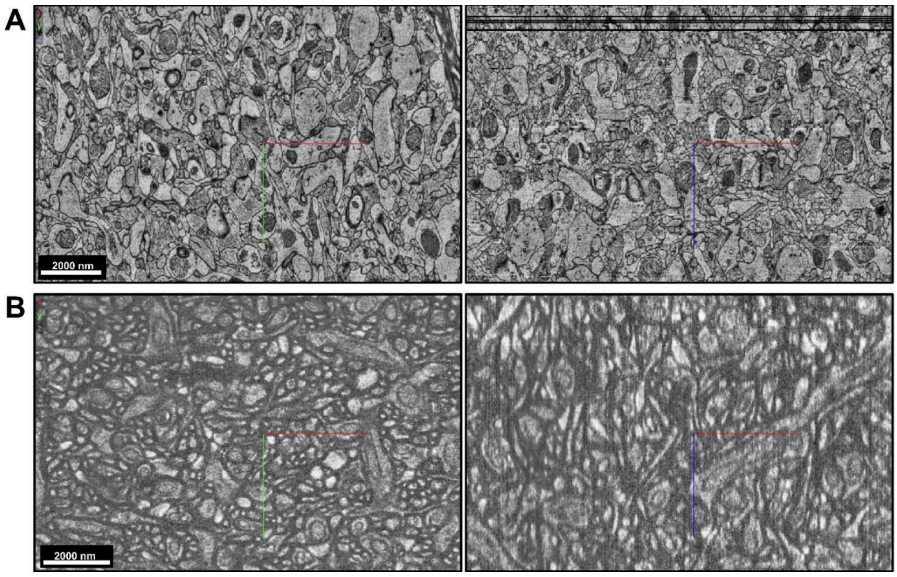

Figure 4. Comparison of raw image data in serial-section transmission electron microscopy (top) and expansion confocal light microscopy (bottom). Image from Collins et al. (2025).

Expansion Microscopy (ExM) allows for super-resolution imaging on conventional light microscopes by anchoring biomolecules to a swellable hydrogel and physically expanding the sample (Chen et al., 2015). While first implementations of this technology worked with four-fold linear expansion (Chen et al., 2015), more recent protocols can reach linear expansion factors well above 10x, bringing effective resolution toward the connectomic regime (Shaib et al., 2024; Tavakoli et al., 2024).

ExM is compatible with molecular annotation, such as protein staining (Tillberg et al., 2016). In ground-truth simulation approaches such as compartmentalized differential equations, this protein data is required to accurately model compartmental voltage (Almog and Korngreen, 2016). By merging together molecular annotations with an imaged connectome, ExM is a promising route to collect nanometer-resolution spatial information about ion channels, receptors, neuropeptide distributions, and other molecules within the brain.

Dense Labeling

Progress has been made towards dense labeling of tissue in combination with ExM to achieve EM-like contrasts. Two approaches are non-specific labeling of anchored proteins (pan-protein staining) using NHS-ester dyes (Tavakoli et al., 2024) and labeling of lipid membranes using intercalating gel-anchorable probes (Shin et al., 2024; Tavakoli et al., 2024).

Labeling of lipid membranes using ultrastructural membrane expansion microscopy (umExM) was able to demonstrate highly accurate traceability of myelinated (Rand score 0.995 +/- 0.004) and unmyelinated axons (Rand score 0.993 +/- 0.006), already at a linear expansion factor of 4x and an effective resolution of ~60 nm (Shin et al., 2024). While current approaches lack sufficient resolution, improved iterative strategies with expansion factors around 12x are in development (Shin et al., 2024).

Both dense labeling approaches would probably benefit from even higher expansion factors in the range of 20-24x, to reach the theoretically necessary effective resolution of 10-30 nm to reliably detect fine spine neck structures (Helmstaedter et al., 2013).

Subsequently, combining pan-protein and lipid staining could help drastically improve future accuracy of reconstruction.

Protein Barcoding

Tracing neurons in electron microscopy relies on the high resolution of electron microscopes and subsequent time-consuming reconstruction and error correction.

If it were possible to uniquely identify the same neuron -- at both its soma and at its far end -- that would loosen the hampering constraint of requiring virtually error-free tracing.

Even if it weren't possible to trace a neuron continuously through tissue, the expressed barcode would allow correct identification of distal synapses. In fact, a hypothetical loss of an entire section or slicing errors would be much less catastrophic, as barcoding would not rely on perfect traceability.

Brainbow and its successors are tool boxes allowing a subset of neurons of transgenic animals to randomly express different levels of fluorescent proteins. This results in neurons displaying a continuous spectrum of colors, with the initial claim of roughly a hundred distinguishable hues, allowing segmentation and tracing of individual neurons based on their differing colors. Breakthroughs in Cre/lox genetic editing and fluorescent protein engineering gave rise to the first such tools. (Livet et al., 2007, Pan et al., 2011, Cai et al., 2013, Leiwe et al., 2024)

Combining Brainbow-style barcoding with multiplexed fluorescence readouts in expansion microscopy has enabled unique identification of neurons and reconstruction of their structure, while related methods can resolve on the order of 10-15 stained targets in a single round (Linghu et al., 2020; Seo et al., 2022). Notably, super-multicolor Tetbow allows reconstruction despite two sections of a neuron being separated, beginning to leverage the core strength of barcoding (Leiwe et al., 2024). In principle, methods like Tetbow could scale to whole-brain connectomics when combined with expansion microscopy, pan-protein staining, and/or lipid staining.

These first-generation barcoding approaches rely on differentiating expression levels of fluorescent proteins. However, expression in different cell types is uneven, trafficking of these proteins is uneven, and fluorescent staining is uneven. Within the same neuron, a distal axon may display a slightly different barcode than the soma, complicating correct identification. Establishing a protein barcoding strategy with

better scalability powered by replacing a gradient colour with a binary system (Is the protein present or not?) may improve unambiguous barcode readout.

Combinatorics

For any protein, the presence or absence of a protein is the binary signal. Bitbow attempts to use this binary approach by using 5 binary XFPs combined with targeted localization (membrane, golgi, nucleus). (Li et al., 2021) Naively, this gives:

$2^5 * 2^5 * 2^5 - 1 = 32{,}767$

different combinations for identification of somas.

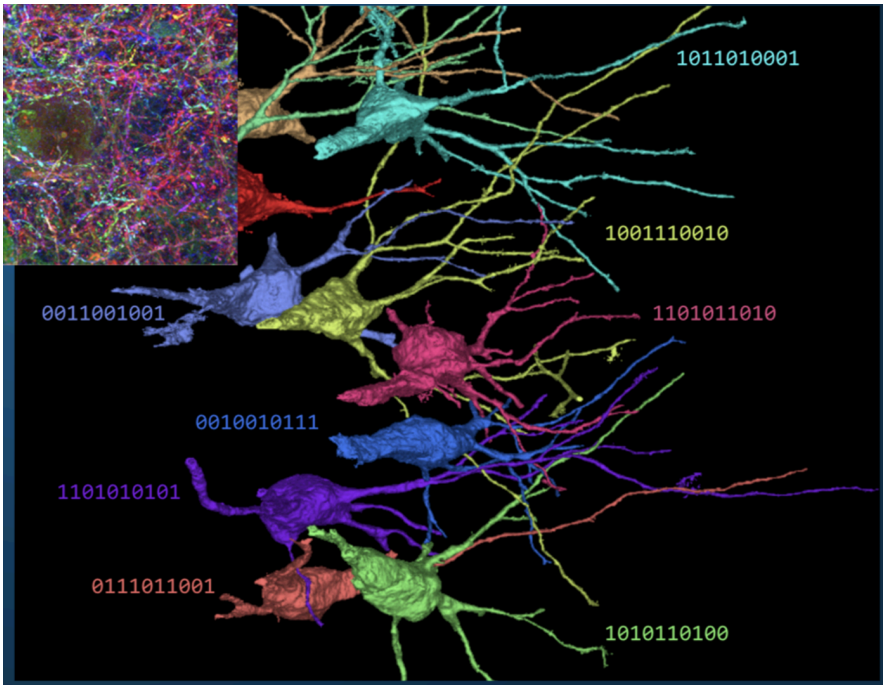

Figure 5. Binary protein barcoding illustration. Courtesy of e11 bio.

Improved barcoding approaches might leverage protein constructs such as spaghetti monsters (Viswanathan et al., 2015) to generate

high-signal binary codes for every neuron. These constructs could be delivered with AAVs. Assuming 25 randomly expressed proteins delivered via AAV, the theoretical diversity of barcodes gives:

2^{25} = ~33M different combinations.

Yet, 2^{25} assumes each protein is expressed independently with probability 0.5 -- the maximum-entropy case. AAV infection rates and expression probabilities are rarely 50%.

Assume each protein is expressed with probability p, the effective barcode space is better described by the binomial coefficient C(n, k) where k = round(n * p). For example, at a typical AAV expression probability of p = 0.3, the number of barcodes with exactly k = round(25 * 0.3) = 8 proteins expressed is C(25, 8) = 1,081,575 -- roughly 3% of the theoretical 2^{25}.

More generally, the Shannon entropy H(p) = -p*log2(p) - (1-p)*log2(1-p) gives an effective barcode space of approximately $2^{n*H(p)}$, which at p = 0.3 yields ~2^{22} = 4.3 million.

As the human brain has 8e10 neurons, most neurons do not synapse onto most neurons. Most neurons are estimated to have between 8,000 to 40,000 synapses on average in mammals (DeFelipe et al., 2002, Schüz and Palm, 1989), 4.3 million unique barcodes is a >100x margin for the number of barcodes versus average number of synapses.

As this quick combinatorics estimate highlights, barcoding need not label every neuron individually in the brain. It needs to uniquely label all the neurons in the synaptic neighborhood of a neuron. Achieving that is within the range of modern technologies.

Barcoding Issues

Such barcoding approaches face three distribution problems when delivery via AAVs:

1. Distribution to all neurons.
2. *Even* distribution among all neurons.
3. Distribution within neurons.

Delivering AAVs carrying barcodes to nearly all neurons remains challenging. This requires improved delivery methods; early intravenous PHP.eB AAV achieves 65-73% reach of cells in the mouse cortex (Chan et al., 2017), with newer variants like AAV.CAP-B10 achieving improved brain-wide transduction and reduced liver targeting (Goertsen et al., 2022). Virus-based platforms (AAV, lyssavirus, Sindbis, HSV, etc.), better intravenous delivery, high-density intra-CNS injections, and creative approaches to multi-site injection throughout the brain will be required -- or perhaps all simultaneously. Combining approaches makes 90% coverage feasible, though difficult, with extensive technology development.

Barcoding currently faces difficulties for even distribution of AAVs after injection, with high-density AAVs close to the injection site dropping off nonlinearly with increasing distance.

Once within a neuron, successful distribution of the trafficked protein epitopes to far axonal and dendritic ends -- especially into synapses -- is an open question to be evaluated. This might be alleviated by driving overexpression of the barcode proteins as far as possible without leading to neuronal dysfunction, and by targeting the expressed barcodes for synaptic trafficking (Peikon et al., 2017).

One major disadvantage of barcoding is the multiple rounds of staining, washing, and imaging required.

Whether state-of-the-art barcoding captures most neurons and especially most synapses could be verified in statistical

comparisons against electron microscope data in small, similar reconstructed samples.

PRISM

A major recent leap for barcoding technology includes Photo-connectomic Reconstruction by Iterative Staining with Molecular Annotations (PRISM) (Park et al., 2025). PRISM integrates three core advancements--binary protein cell barcoding, expansion microscopy, and self-proofreading AI segmentation models.

PRISM achieved automatic tracing accuracy 8-fold higher than conventional single-color methods. In a roughly 10 million um^3 volume of mouse hippocampus, the technique also enabled molecular mapping of synapses.

This drastically reduces the need for manual human proofreading, which currently accounts for a large fraction of connectomics costs (Bosch et al., 2023).

Protein Staining

Proteins are required for all life processes, providing cellular structure, messengers, catalysts, scaffolds, receptors, switches, motors, pumps (Alberts et al., 2002). Protein staining in connectomes will be helpful for mapping neuronal function, identifying cell types, and subsequent brain simulations.

Examples of initial targets for staining in connectomics may include, but are not limited to:

- key synaptic and axonal markers, such as PSD-95, neurofilaments such as BASOON, SHANK-2, NF-L, NF-M, NF-H, or Synapsin-1, and markers of gap junctions such as connexins,
- as well as core synaptic proteins such as AMPAR, NMDAR, GABAR, etc.,

- and key intracellular proteins mediating neuronal state, such as CaMK2.

Strategies relying on multiplexed antibody staining combined with super-resolution can now label on the order of 10-15 targets with high spatial resolution (Linghu et al., 2020; Seo et al., 2022).

In parallel, barcoding would require staining for around 20 proteins in large expanded tissue, perhaps as much as 30 proteins (Rodriques, 2009).

Both of these information-rich protein stains and barcoding protein stains require more imaging rounds.

Imaging rounds are often done overnight and take up days each. In principle, they may be done within 3-4 hours for staining and washing each. With sufficiently good protein targets, specific antibodies, and rigorous washes, some tissue degradation occurs during repeated imaging of expanded tissue, but multiple rounds of staining are fundamentally feasible in practice (Linghu et al., 2020; Seo et al., 2022).

Adding imaging rounds would be particularly time-expensive at the scale of a mouse or human brain, due to increasing the microscope time linearly. Furthermore, registration and processing of multi-round images is time-consuming. Hence breakthroughs such as multiplexing are desirable.

State-of-the-art multiplexing such like PICASSO can in principle reach up to 15 targets per round (Seo et al., 2022). With 30-bit long barcoding, 10 protein targets, and 10-fold multiplexing, that would give only 4 rounds of imaging required.

Protein staining technology may experience significant breakthroughs via e.g. approaches requiring fewer or faster washes such as adaptations of DNA-based FLASH-PAINT (Schueder et

al., 2023, Panluminate, 2022). Multiplexing of such techniques may allow dozens of stains with fewer washes.

Protein Sequencing

One of the highest-upside innovations in the space: What if we wanted to identify *all* tens of thousands of proteins in a cell?

Protein staining is limited by extensive imaging rounds. Staining for 30,000 proteins with 10-fold multiplexing per round would require 3,000 rounds, i.e. years of washing and imaging time for a single slice, not to speak of the astronomical antibody costs. Future technologies such as in-situ protein sequencing could reduce that burden dramatically, but this remains early-stage and unproven at brain scale.

In 1950, Edman et al. pioneered protein sequencing by cleaving off the n-terminus amino acid, identifying it in bulk via mass spectrometry, and repeating this process. (Edman et al., 1950) Protein sequencing has since made large strides, such as nano-pore protein sequencing (Motone et al., 2024), protein sequencing via reverse translation (Zheng et al., 2024), high-throughput identification of proteins immobilized on a glass plate (Swaminathan et al., 2018), or visualizing the shape of single proteins in expansion microscopy combined with other superresolution techniques (Shaib et al., 2024).

In principle, one could imagine anchoring all the proteins in a cell to a gel, and combining the original Edman cyclical process with expansion microscopy, and do repeated identification of amino acids.
Modern genome sequencing costs decreased super-exponentially (Wetterstrand, 2022) and all major model organisms have been sequenced (Quake, 2024). Using organism-specific protein databases, researchers can reduce the amino acid sequences needed

for reliable protein identification to just a few residues, enabling high-confidence identification.

Developing such technology faces significant difficulties in 1) binder specificity and dwell time and 2) increasing decrowding or resolution. Antibody epitopes typically identify sequences of at least ~5 amino acids with unique specificity (Huang et al., 2020), making the development of effective and long-dwelling binders against single cleaved amino acids difficult. Creative approaches such as de novo design, larger standardized antigen scaffolds, or local tethering strategies may help, but this remains speculative.

Proteins are typically ~2-10 nm large (Varongchayakul et al., 2018) and packed densely in the cell with a typical distance of 11 nm between GABA receptors at the synapse (Liu et al., 2020). 10x expansion microscopy yields an effective theoretical resolution limit of 300 / 10 = 30 nm, insufficient for decrowding tissue at typical distances of ~11 nm. This can be addressed with higher expansion factors or alternative chemistry, such as N-hydroxysuccinimide esters (Shin et al., 2024) Decrowding requirements depend on whether the goal is to localize intact proteins versus tracking proteolyzed fragments post-expansion; the latter may reach finer effective separations than the diffraction-limited microscope resolution alone would suggest.

The implications of pan-cell protein identification technology would be enormous for connectomics. With a sufficiently scalable and cheap breakthrough in-situ protein sequencing, the number of “stained” proteins might increase by orders of magnitude.

Interviewing researchers at the Boyden Lab at MIT, we remark that work on such breakthrough technologies is in need of more support and funding, especially relative to its high potential upside.

Notes On Structure

In short, imaging techniques required for connectomics are improving.

Improving: Be it the 2019 331-beam EM microscope, or breakthroughs in light microscopy such as ExA-SPIM (Glaser et al., 2024) achieving 946 megavoxels per second to image an entire 3x-expanded mouse brains in 24 hours.

While out of scope for this paper, we note that X-ray microscopy is promising in the long term as a high-throughput imaging mechanism. Under bright X-ray sources from synchrotrons, resolutions of 10nm are plausible--sufficient for connectomics. (Howells et al., 2009) Previous achievements include sub-100 nm imaging fly and mouse brain tissue using x-ray (Kuan et al., 2020) and successful imaging of individual synapses (Bosch et al., 2023).

Scaling Functional Imaging

Connectomes offer a remarkably detailed snapshot of a brain's circuitry. Yet, since data acquisition occurs post-mortem, they cannot image the dynamic changes in activity that occur in a living brain.

Hence, moving from structure to function is a challenge (Scheffer and Meinertzhagen, 2021, Bargmann, 2012, Bargmann and Marder, 2013). One way to overcome this issue is to combine reconstructed connectomes with functional information obtained in-vivo. A 2019 review attributed the lack of a successful simulation of C. elegans - the most well-characterized model organism, connectomics-wise - at least partially due to the lack of neurophysiological data (Stiefel and Brooks, 2019).

Organisms

After basic technologies for genome sequencing were established, sequencing costs started to drop at super-exponential rates in the late 2000s (Wetterstrand, 2022). Analogously, in functional imaging, the number of neurons being simultaneously recorded has roughly doubled every 7.4 years (Stevenson and Kording, 2011, Marblestone et al., 2013), but has rapidly accelerated since using *light* instead of electrodes to extract neuron activity data. Powered by genetically encoded fluorescent indicators and advanced microscopy techniques, optical imaging allows parallel recording of the neural activity of large populations of neurons at cellular resolution, and, in small transparent organisms such as *C. elegans* and larval zebrafish, can noninvasively monitor the whole-brain neural activity (Kato et al., 2015; Nguyen et al., 2016; Ahrens et al., 2012, 2013; Vladimirov et al., 2014).

C. elegans has only about half of its neurons in the frontal "brain", and the other half distributed throughout the body as ganglia, motor neurons, etc. (Arnatkevičiūtė et al., 2018). Proof-of-concept whole-brain or whole-body calcium imaging of *C. elegans* has been achieved. In 2013, Schrödel et al. combined two-photon excitation with temporal focusing to record calcium signals from 70% of all head neurons at 4-6 Hz in immobilized worms (Schrödel et al., 2013)(Prevedel et al., 2014).

More recently, single-objective oblique light-sheet microscopy (SCAPE -- Swept Confocally-Aligned Planar Excitation) accelerated cellular-resolution imaging of entire worm nervous systems to ~25.75 Hz while maintaining free animal movement (Voleti et al., 2019). Additionally, a fully automated tracking platform for freely moving imaging has further advanced C. elegans whole-body imaging (Li et al., 2021).

Despite this, large standardized whole-body activity datasets for C. elegans remain limited, and voltage indicators have been difficult to implement successfully in the worm, so work continues (Simeon et al., 2024).

Young larval zebrafish may be an excellent imaging organism. They possess a small, transparent brain. Early whole-brain calcium imaging in larval zebrafish operated around 0.8-3 Hz (Ahrens et al., 2013; Vladimirov et al., 2014). ZAPBench recently extended this to calcium imaging of more than 70,000 neurons at 1 Hz (Lueckmann et al., 2025). Whole-brain voltage imaging remains earlier-stage, but one recent preprint recorded voltage from roughly one-third of larval-zebrafish neurons simultaneously (Wang et al., 2023).

Imaging the whole adult Drosophila brain at cellular resolution is difficult due to tissue scattering in its brain. Nonetheless, two-photon microscopy has been used to monitor the calcium dynamics from nearly the entire adult fly brain at 1.95 Hz, with a sampling voxel size of 2.6×2.6×7.5 µm3 (Mann, Gallen, and Clandinin 2017).

A dramatic recent advance in mammalian functional imaging is light beads microscopy, which has enabled simultaneous two-photon calcium imaging of up to 1 million neurons across the mouse dorsal cortex at 2 Hz cellular resolution (Manley et al., 2024). Combined with 'crystal skull' curved glass windows replacing the dorsal cranium -- viable for at least 11 weeks post-surgery -- this provides optical access to an estimated 800,000-1,100,000 neurons spanning over 30 neocortical areas (Kim et al., 2016). This represents a roughly 13-fold improvement over the previous state of the art of ~75,000 simultaneously recorded mouse neurons.

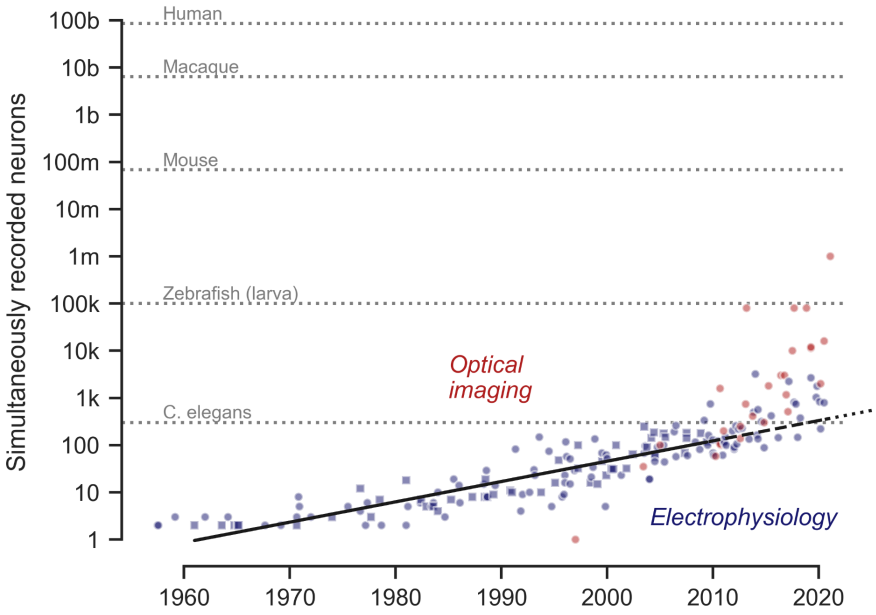

Figure 6. The count of simultaneously recorded neurons (log scale) is growing exponentially or faster. Organism sizes (in neuron count) are shown for reference; they do not imply that whole-brain imaging is currently achievable. (Urai et al., 2022).

Light-based functional imaging will likely improve in neuron count, spatial resolution, temporal resolution (voltage indicators), and quality. Other paradigms--such as magnetic resonance, electrodes, optical fibers, or molecular recording devices--face constraints but remain under development (Marblestone et al., 2013).

Human-Scale Brain Imaging

Human-scale brain imaging at non-single-neuron resolution leverages techniques like functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG), which provide valuable insights into brain activity and connectivity at the mesoscale.

fMRI measures blood oxygenation level-dependent (BOLD) signals, capturing activity changes with millimeter-scale spatial resolution and temporal resolution on the order of seconds. This allows researchers to infer broad patterns of functional connectivity between regions, although it lacks the fine-grained temporal resolution needed to resolve individual neuronal dynamics. MEG, on the other hand, records magnetic fields generated by neural activity with millisecond temporal precision, offering a complementary view of real-time activity patterns.

On the invasive recording front, Neuralink's N1 implant has demonstrated sustained human BCI performance over 670+ days post-implantation and 4,900+ hours of active brain-computer interface use. While the N1 records from far fewer neurons than light beads microscopy (hundreds vs. millions), it operates in humans and captures spiking activity at millisecond resolution from deep cortical layers that optical methods cannot reach in vivo.

The Optical Frontier

Despite these advances, single-neuron whole-brain imaging is hard. Light-based calcium or voltage imaging is constrained by tissue thickness and transparency. Small organisms such as C. elegans and larval zebrafish are often naturally transparent; humans are not. Both the skull and thick tissue prevent successful large-scale calcium imaging in mammals. Aside from the skull, scattering and absorption are the fundamental constraints on tissue penetration (Marblestone et al., 2013), often limiting penetration depth to 1-5 millimeters due to absorption and scattering.

These constraints can be partially alleviated through cranial windows, genetically transparent animals such as crystal zebrafish (Antinucci and Hindges, 2016), or tissue-clearing chemicals (Ou et al., 2024; Franzesi et al., 2024).

One way to improve optical transparency in living animals is to use absorbing molecules such as tartrazine (Ou et al., 2024), suggesting that deeper in vivo imaging without cranial windows is at least physically plausible in some regimes.

In practice, high-quality whole-brain single-neuron imaging beyond ~1-2 mm depth remains challenging with current methods (Marblestone et al., 2013). Early simulation attempts show that limited recording data may be sufficient to predict whole-brain data, i.e. that whole-brain single-neuron recordings may not always be necessary for whole-brain neural activity prediction (Beiran and Litwin-Kumar, 2024). As a first approximation, this highlights model organisms in which whole-brain single-neuron imaging is near-term feasible.

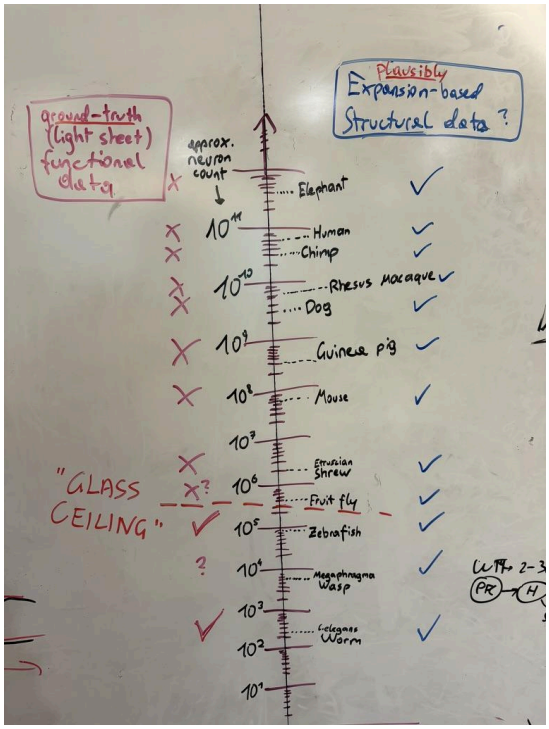

Figure 7. A 5-minute sketch on the lab whiteboard outlining how a potential "glass ceiling" might ground-truth functional data collection. (Personal drawing, 2024).

In *C. elegans and* larval zebrafish, whole-brain single-neuron imaging is now demonstrated or under development. Other understudied organisms could achieve this, such as *Xenopus laevis* tadpoles, *Hydra vulgaris* or planarian flatworms.

Adult zebrafish, mice or human brains far exceed the ~1mm scattering & absorption limits. Whole-brain single-neuron imaging therefore remains farther-out in these organisms.

We note that adult *Drosophila* occupies a borderline position between feasibility and challenge: matching zebrafish in neuron count and theoretical depth limits, yet with reduced optical transparency, presenting both promise and obstacles (Lemon et al., 2015). Whole-brain imaging has been achieved, but not at a single-neuron resolution (Aimon et al., 2019). With further development of tissue clearing techniques, whole-brain single-neuron imaging may be achievable in Drosophila.

Network analysis indicates that *Drosophila* neuron connections are fairly stereotyped, except cell bodies varying highly in location across the outer shell of the brain (Lin et al., 2024). If whole-brain single-neuron imaging in Drosophila is possible and Drosophila is sufficiently stereotyped to map the functional data of one individual fly to another individual's connectome, then that would have enormous repercussions for functionally recording and simulating complete brains. With a focused effort on garnering as much high-resolution calcium or voltage data as possible, it could be combined with the recently completed connectome of the fly.

This leaves the zebrafish as a common vertebrate model organism in which we plausibly image all of its neurons with near-term technology. It is the largest model organism in this thesis for

which near-whole-brain, single-neuron functional recording appears plausible. In comparison to C. elegans, the zebrafish has a more vertebrate-like brain structure, extensive neural plasticity, complex behaviours, and complex learning mechanisms (Koch and Reid, 1999; Herculano-Houzel, 2009). We note that many other animal models are either too large for ground-truth whole-brain single-neuron imaging or exhibit less mammal-like neural dynamics, making the zebrafish a promising platform where high-resolution functional imaging combined with subsequent connectomics is plausibly achievable.

With extensive ground-truth structural, functional, and behaviour data collection, it could serve as a thorough testbed for the first end-to-end simulated animal. In parallel, as technology progresses, we are optimistic more promising model organisms will open up across species.

.| / ... / .

Simulations

We focus on neuron-level simulations. Lower-resolution approaches exist, such as next-voxel prediction from fMRI data (Tang et al., 2023). Here we cannot do justice to the large and complex field of computational neuroscience at all levels of resolutions, and consider only cellular-resolution simulations.

Computational neuron models can be of varying degrees of complexity: from extremely simplified models such as the binary McCulloch-Pitts model (McCulloch and Pitts, 1943; Izhikevich, 2003; Brette and Gerstner, 2005) to highly biophysically accurate models such as multicompartmental Hodgkin-Huxley models (Hodgkin and Huxley, 1952; Haufler et al., 2023; Petousakis & Apostolopoulou, 2023). Similarly, models simulating synaptic transmission can be more or less biophysically accurate, ranging from simplified models such as current-based exponential synapses

to highly biophysically accurate models (Ecker et al., 2024) such as multi-state kinetic receptor models Ecker et al., 2024.

Importantly, questions remain about what neural properties must be directly measured versus inferred, what can be learned from measurements, and what additional constraints are needed for accurate simulation. In general, strategic simplification can likely maintain high levels of biological realism; for review, see Einevoll et al., 2019.

Efforts have been made to simulate brains in detail, including the Blue Brain Project (BBP) (Markram 2006) that was initiated with the ambitious goal of creating a comprehensive digital simulation of a human brain. The BBP first computationally modeled ~31,000 rat cortical neurons (Markram et al., 2015). At its completion in 2024, the BBP released models and simulations of larger portions of the rat brain: a 4.2 million-neuron model of non-barrel somatosensory cortex (Isbister et al., 2026) and a 456,000-neuron model of the hippocampal region CA1 (Romani et al., 2024). However, substantial gaps in knowledge of the overall circuit architecture--especially the lack of a synapse-level connectome--led to an approach relying heavily on computationally filling-in missing information (Hill et al., 2012), resulting in significant challenges.

The Allen Institute is similarly simulating mouse brain circuits, beginning with primary visual cortex (V1). Using both multicompartmental biophysically detailed and point-neuron models (Haufler et al., 2023; Billeh et al., 2020; de Vries et al., 2020), and integrating a vast array of experimental data on cortical neuron types and their morpho-electric properties (Gouwens et al., 2018; Gouwens et al., 2019; Gouwens et al., 2020), connectivity (Gouwens et al., 2020; Campagnola et al., 2022), and activity in vivo (de Vries et al., 2020; Siegle et al., 2021), they ran a 230,000-neuron model of mouse V1 with realistic visual inputs (Arkhipov et al., 2018; Billeh et al., 2020). This model has been used to investigate structure-function

relationships in the cortical circuits (Billeh et al., 2020; Galván Fraile et al., 2024), study the biophysical mechanisms sculpting electric fields recorded in the cortex (the local field potential, or LFP) (Rimehaug et al., 2023; Rimehaug et al., 2024; Meneghetti et al., 2024), and explore and compare task training and learning in biological and artificial neural nets (Chen et al., 2022). Subsequently, in a collaboration with the Japanese supercomputer Fugaku, they scaled their simulation efforts to ~10 million neurons with 26 billion synapses across 86 brain regions. (Kuriyama et al., 2025)

Other groups have also released models of mammalian cortical circuits at the ~10,000-100,000-neuron scale using biophysically detailed, point-neuron, and hybrid approaches; for a recent review, see Dura-Bernal et al. (2024). Recent applications include studying slow waves in a human thalamocortical network model (Marsh et al., 2024) and testing new pharmacology in a model of human cortical function in depression (Guet-McCreight et al., 2024).

OpenWorm is developing a similar representation of the 302 neurons in C. elegans, attempting to build a comprehensive simulation environment of both functional prediction and a simulated body of the worm (Szigeti et al., 2014). C. elegans efforts have been hampered by graded, non-spiking dynamics and, until recently, a relative lack of direct neurophysiological data (Randi et al., 2023; Stiefel and Brooks, 2019). Critically, cell capacitance and conductance values are often assumed rather than empirically derived, as are key details such as axonal lengths and propagation delays (Szigeti et al., 2014; Sarma et al., 2018). A recent successor to OpenWorm called BAAIWorm has emerged as an integrative data-driven model simulating brain-body-environment interaction (Zhao et al., 2024).

Indeed, C. elegans has a small brain, but it is far from a simple one. Constrained by a limited capacity for complex interneuronal communication, the worm’s brain is forced to maximize

computational processing within each neuron, with signaling proteins occupying over 20% of the worm's genome (Sterling and Laughlin, 2015).

Attempting to partially account for such factors, Creamer et al., 2024 fit a connectome-constrained linear model to whole brain *C. elegans* recordings of neural activity during optogenetic perturbation of single neurons and achieved relative correlations for held-out neurons up to 0.92 (Creamer et al., 2024). Further progress is possible, especially with more data from extensive light-based microscopy work combined with protein staining (Tillberg et al., 2016), a peptidergic atlas (Ripoll-Sánchez et al., 2023, Wang et al., 2024) and in-situ mRNA sequencing (Alon et al., 2021).

Shiu et al. present a *Drosophila* simulation model that applies a leaky integrate-and-fire framework to simulate circuit-level behaviors like feeding and grooming (Shiu et al., 2024). They base it on ground-truth neural connections from the complete fly connectome (Dorkenwald et al., 2024) and neurotransmitter data. The model's predictions, validated through experiments, are a promising proof-of-concept on how connectome-based models can effectively replicate behavioral circuits. Further, Lappalainen et al., 2024 combine machine learning and the connectome to create a connectome-constrained model of the fly visual system trained on the computation of visual motion from naturalistic visual stimuli. This modeling accurately predicts known parameters of the Drosophila visual system, for example, the contrast selectivity of all 31 cell types for which this has been previously determined. Likewise, Cowley et al., 2024 created a deep neural network corresponding to the Drosophila visual system in which the response of artificial neurons accurately matches their experimentally determined response properties on data the network was not trained on.

Current efforts include the call for an inter-laboratory project to comprehensively characterize neuronal input-output functions

for the 300 neurons of *C. elegans* to its neuronal output, with the goal of predicting behavior and responses to novel perturbations. (Haspel et al., 2023).

Structure & Function

As shown in the optical frontier section, single-neuron whole-brain functional data is currently limited to a few small organisms, and a connectome, no matter how detailed, is a static snapshot of a brain's structure. This opens up a gap: The gap of moving from structure to function. How could we go from a static snapshot of brain structure to a simulation capable of driving an accurate brain function?

Note we focus on one scenario here, where the field is in a structure-dominant paradigm. This serves to highlight one plausible path to whole-brain emulation; alternative paradigms are plausible and likely, such as function-to-function prediction. Due to this focus, we cannot cover function-to-function predictions. Using functional data to predict structural connectivity via e.g. co-correlation has also been demonstrated (Creamer et al., 2024), but is not discussed here.

In principle, the single-neuron function-to-function simulations seem feasible in organisms below the optical frontier, i.e. in small brains amenable to whole-brain voltage imaging. Training function-to-function machine learning models has been proposed in *C. elegans*, as roughly 500 hours of calcium functional data are available. (Simeon et al., 2024a, Simeon et al., 2024b) Yet, current modelling attempts lack sufficient data to accurately predict functional activity, suggesting the need for large-scale data collection in *C. elegans* (Simeon et al., 2024a, Stiefel and Brooks, 2019). Current functional data is limited to at best ~200 of the 300 neurons, emphasizing the need for improved microscopes capable of whole nervous-system recordings in conditions allowing the worms to move as freely as possible. Further improvements

could include voltage imaging instead of calcium imaging, and datasets including out-of-distribution activity with robust optogenetic stimulation of as many neuron subsets as feasible, ultimately leading to deeper understanding of neuron activity causality. (Ahrens et al., 2013; Kim et al., 2017)

Early attempts at whole-brain structure-to-function prediction Shiu et al. (2024) have been demonstrated. They present a *Drosophila* simulation model based on ground-truth neural connections from the complete fly connectome (Shiu et al., 2024). To instantiate this connectome, leaky integrate-and-fire (LIF) neurons were used -- a system of simple differential equations developed as early as 1907 (Abbott, 1999). The model attempts to simulate circuit-level behaviors, like feeding and grooming. Validated by in-vivo experiments, this is a promising proof-of-concept on how connectome-based models can effectively replicate behavioral circuits. (Shiu et al., 2024)

Notably, the fly connectome was produced via electron microscopy, unable to directly capture membrane proteins such as neurotransmitter receptors or other molecular data. Also note that this was done with simple differential equations and no compartments, i.e. as an extremely simplified model of a brain directly based on the fly connectome. Additionally, almost no fruit-fly functional data was used in tuning that model. Yet, it accurately reproduced some fly behaviours. The simplest connectome with no molecular information and century-old integrate-and-fire-style differential equations (Abbott, 1999) lacking compartments or biophysical detail were sufficient to produce limited behavior. A thoroughly impressive empirical result.

Yet, it remains to be shown that this paradigm can generalize to mammal-scale brains. The critical empirical question is how much structural and functional data is required. In our interviews, researchers were tentatively optimistic that detailed structure-to-function models with little functional data will be

imperfect but may still replicate surprising amounts of behaviour and consistent neuronal activity, as recently demonstrated in whole-brain Drosophila simulations (Shiu et al., 2024).

This would potentially scale as a key approach to mammalian simulations, such as shown in the Allen Institute V1 model (Billeh et al., 2020; Rimehaug et al., 2023). In some interviews, optimism is informed by how robust stable brain states are as attractor spaces in functional activity: Evidence includes returns to stable brain states from diverse conditions such as seizures, anaesthesia, coma, blunt trauma, sleep, hydrocephalus, and stroke recovery.

Integrating Functional Data

Of course, if available, tuning on functional data would be of obvious benefit.

Assuming a mouse or human connectome is achieved, future structure-to-function work may require more complex fine-tuning, e.g. using different parameters in extended Hodgkin-Huxley or Izhikevich neurons to simulate different cell types accurately.

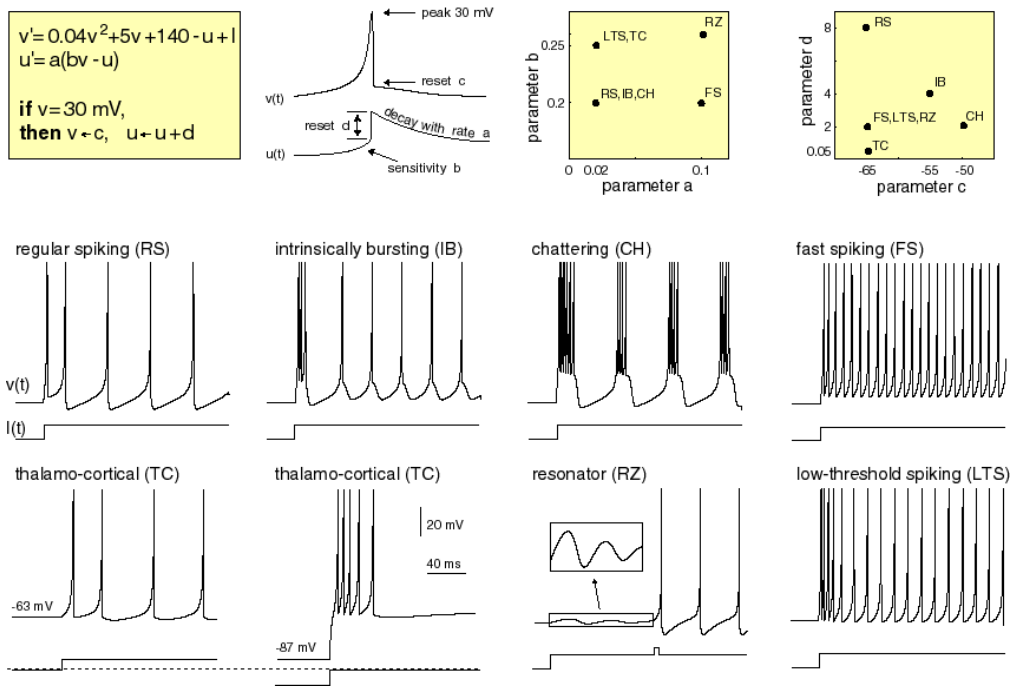

Figure 8. Spiking behaviors of known cortical neuron types. Reproduced from [Izhikevich (2003)](#).

One proposal would be to extend these equations with more voltage-dependent terms, especially at synaptic compartments. Each term would represent a class of receptors identified at that synapse and requires tuning through in-vivo trials of receptor dynamics. In the model, parameters of those voltage-dependent terms may be constrained by the differing levels of receptors found at that synapse in expansion-based light microscopy. This constraint also serves to constrain the possible solution space. This is a fully differentiable system, as the prediction as well as the parameters for each voltage-dependent term and its probability curves are smooth and continuous.

With improving structural and functional data, such modelling approaches should improve systematically. Notably, the free variables may also implicitly account for influences not explicitly modelled, ranging from plausibly relevant factors like ephaptic coupling (Anastassiou et al., 2011) to speculative or implausible factors, such as cosmic rays.

Additionally, neuroscience offers an extensive corpus of behaviour and parameters of various cell types; while not ideally standardized or replicable, scraping that corpus could serve as important anchors for parameters across neuron types.

This suggests using a major tool which has proven itself useful for finding optima in differentiable landscapes: Gradient descent (Ruder, 2016). This approach has been demonstrated for both point-neuron (Stanojevic et al., 2024) and biophysically detailed network models (Chen et al., 2022; Oláh et al., 2022; Zhang et al., 2023; Deistler et al., 2024).

The predictions of these "extended" differential equations would benefit from tuning on ground-truth functional data to produce accurate typical activity. The ideal whole-brain single-neuron functional dataset may only be available in a few limited organisms, such as *C. elegans* and the larval zebrafish. This again highlights especially the latter as a promising validation platform for complete end-to-end integration of structural connectomic data, functional whole-brain data, and a differentiable simulation system to achieve a high-resolution simulation of a whole brain.

It may also be possible to inform a primarily structure-to-function model with more limited functional data, such as partial brain, slice or organoid functional recordings, or fMRI, MEG, or EEG data, for example by using spatial or temporal upsampling.

Scale & Compute

Total compute in supercomputers has been rising exponentially for decades and systems in the 2020s have reached exaflop territory, for example El Capitan at more than 1.5 exaflops (Thomas, 2024), with typical capital costs on the order of hundreds of millions of dollars (Joseph, 2023). It is also worth noting that many frontier AI clusters do not appear in TOP500, since companies

usually do not disclose full details about their infrastructure. As a result, the computational capacity of state-of-the-art systems likely exceeds published supercomputer rankings. Publicly disclosed AI clusters now operate at the scale of tens of thousands to more than 100,000 accelerators; xAI's Colossus cluster, for example, has scaled beyond 100,000 H100 GPUs (Musk, 2024).

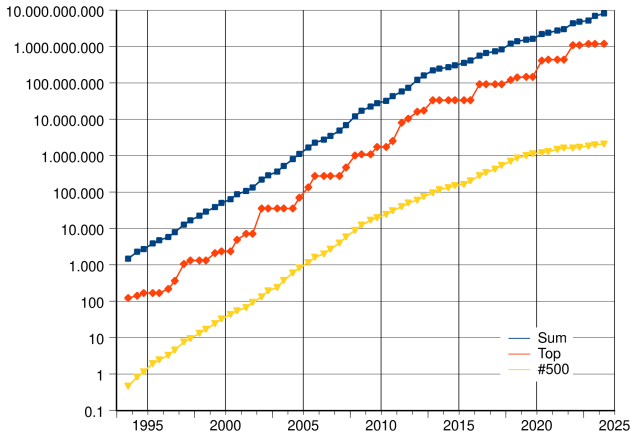

Figure 9. Rise of supercomputing performance. Source: AI.Graphic, CC BY-SA 3.0, via Wikimedia Commons.

Using point-neuron models with different levels of detail, the computational feasibility of large-scale brain simulations has already been demonstrated, including up to human-scale. Yamazaki et al. (2021) ran a model with the number of neurons and connections consistent with those in the human cerebellum (~68 billion neurons, comprising ~80% of total brain neurons) using LIF neurons on a supercomputer capable of up to 11 petaflops, at 600 times slower than real time.

As the reader may note, there currently is no human synapse-level connectome available. Given the predictable and repetitive structure of the human cerebellum, they statistically and algorithmically designed a cerebellum based on histological studies and known distributions of neurons synapsing onto each other.

Similarly, Lu et al., 2024 instantiated a human-scale simulation with 86 billion neurons and 47.8 trillion synapses as a proof-of-concept. On their GPU cluster, this simulation ran between 60 to 120 times slower than real time.

Yet, lacking a ground-truth connectome, these primarily serve as a proof-of-concept that extremely simplified human-scale simulations with simple neuron models are already feasible. As AI data center build-out is becoming one of the largest infrastructure projects in history, we are optimistic about surpassing compute requirements of more complex human brain emulations in the upcoming decades.

To bound discussion, a few brief Fermi estimates for compute requirements reveal substantial but tractable demands:

A lower bound: For leaky integrate-and-fire (LIF) neurons we assume approximately 40 floating-point operations per millisecond each (Izhikevich, 2004; Brette et al., 2007). For 8.6×10^10 neurons at real-time simulation speed, this yields 40 × 8.6×10^10 × 10^3 ≈ 3.4 petaFLOP/s -- roughly the tensor throughput of a single H100-class GPU at FP16 precision (~1 petaFLOP/s dense, or nearly 2 petaFLOP/s with structured sparsity; NVIDIA Corporation, 2023).

An upper bound: We assume Hodgkin-Huxley neurons with 1,000 compartments and 7,000 synapses and approximately 690 floating-point operations per compartment per millisecond (using Euler-exponential solvers; Hines & Carnevale, 2001). Synaptic integration contributes approximately 10^4 operations per spike with 10 Hz average firing rates (Koch, 1999; Brette et al., 2007). This scales to approximately 10^20 FLOP/s--comparable to next-generation 100,000-GPU AI clusters such as xAI's Colossus (~10^20 FLOP/s). (Patel, 2024)

We highlight these serve as order-of-magnitude ballparks, not as thorough estimates.

Critically, the majority of compute cost derives from synaptic modeling rather than neuronal dynamics; the per-neuron gap between LIF and Hodgkin-Huxley may be small relative to synaptic detail costs. Given ongoing hardware improvements driven by artificial intelligence investment, raw computational power is unlikely to emerge as the primary bottleneck in the near term.

Spanning from a single GPU for simplified models to large AI clusters for biophysically detailed simulations under pessimistic assumptions, human-scale brain simulations are surprisingly computationally accessible today in terms of raw computational power.

Memory & Interconnect

The widening gap between processing power and memory/interconnect capabilities has created bottlenecks in large-scale neural simulations that will likely require more thorough integration of memory and processing.

Over the past 30 years, the rate of improvement in processing power has far exceeded the rate of improvement in memory and interconnect development. Peak hardware FLOPS have improved by roughly 3x every 2 years, compared to only 1.6x and 1.4x for memory and interconnect bandwidth respectively, leading to the so-called memory wall (Gholami et al., 2024, An et al., 2024). Memory capacity has also undergone a similar trend: today's El Capitan boasts 1.74 exaFLOPS and over 5.4 petabytes of HBM3 memory (Thomas, 2024); NEC's Earth Simulator boasted 41 teraFLOPS and 10 terabytes of DRAM memory in 2004 (Sato, 2004); representing an approximately 42,400-fold increase in processing power and just a 540-fold increase in memory capacity. Like many

other workloads, computational simulations of brain tissue have been affected by this trend: a 2014 study identified interconnect bandwidth as a bottleneck for spiking neural network simulations (Kunkel et al., 2014) and a recent in-depth analysis found memory bandwidth and interconnect latency to represent key bottlenecks for all types of neuron simulations studied - from point neurons with current-based synapses to multicompartmental models with conductance-based synapses - at high fan-ins and neuron counts (Cremonesi et al., 2020).

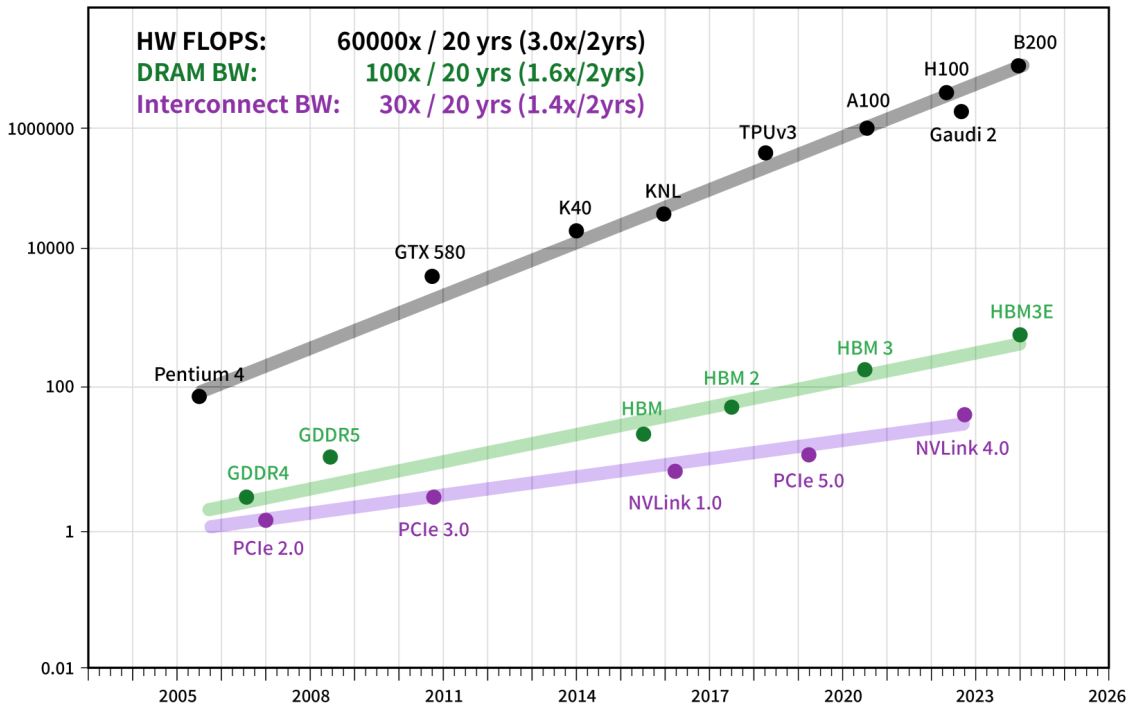

Figure 10. Memory and interconnect have not kept pace with processing power. Y-axis is normalized peak hardware FLOPS relative to the R10000 processor. Compiled from Sato (2004), NVIDIA Corporation (2023), and Thomas (2024).

Napkin math suggests: Memory requirements: with 64 bytes per synapse (for IDs, type, and receptor states) and 64 bytes per neuron, plus 2x overhead for high-performance computing bookkeeping, we obtain approximately 7×10^16 bytes (~70

petabytes) total. Distributed across 100,000 GPUs, this yields approximately 700 gigabytes per GPU. Current top-tier GPUs possess 80-192 gigabytes of HBM, placing this estimate less than an order of magnitude beyond current hardware capacity. While tight, this is not intractable, especially given ongoing memory scaling in the GPUs.

Most brain connections are local, lending themselves to parallelization, but cross-partition communication remains non-trivial. With clever partitioning, perhaps 10% of synapses cross GPU boundaries. This yields roughly 6×10^13 cross-partition synapses (from 6×10^14 total with 10% cross-partition traffic). Spike events can be encoded compactly at approximately 4 bytes each, including timing and neuron ID. At 10 Hz firing rates across 10^5 GPUs, this translates to approximately 24 GB/s per GPU.

Many interventions -- such as moving to lower-precision arithmetic (Micikevicius et al., 2018) -- could help decrease memory and interconnect requirements and boost processing power. The H100 GPU, for example, sees a boost from 67 teraFLOP/s (FP64 Tensor Core) to roughly 1 petaFLOP/s at dense FP16 tensor throughput, or nearly 2 petaFLOP/s with structured sparsity (NVIDIA Corporation, 2023). Indeed, the effective hardware-utilization rate would remain bottlenecked by memory and interconnect bandwidth, much like how data-movement and latency bottlenecks are expected to limit the training of frontier AI systems (Erdil and Schneider-Joseph, 2024).

Data Storage

Mapping mammalian brains will generate enormous amounts of data. At 10 nm isotropic resolution, 1 mm³ of brain tissue corresponds to 10^15 voxels; assuming 1 byte per voxel, that is roughly 1 petabyte of raw data. For a human brain volume of about 1.4 million mm³, this implies on the order of 1.4 zettabytes of raw

image data before compression. Compression can reduce this substantially, with state-of-the-art approaches demonstrating ~128× reduction without compromising reconstruction (Li et al., 2024). In other words, the storage required for a human connectome could plausibly fall into the 10-100 petabyte range after compression and downstream processing. Datacenters such as those operated by Backblaze already store data at exabyte scale (Backblaze, 2022), which is ten to a hundred times larger than that compressed requirement.

A brief comparison to other large-scale scientific efforts such as CERN comes to mind. In CERN's case, their facility has storage capacities exceeding 1 exabyte and throughput on the order of 100-200 GB/s, which is more than enough to keep up with data generated from our microscope paths (20GB/s) ("CERN Data Centre", n.d., Bosch et al., 2023). CERN's datacenter cost 83m CHF ($93m USD) to build, and on the order of $100-200m USD to run per year ("Facts and Figures", n.d.). Assuming we build a data center with ~10 petabytes of storage, 100x smaller than CERN's, a rough cost estimate gets us to anywhere from $10-50 million for the initial build and compute costs ("Data Centre Cost", n.d.).

Concretely, Collins et al. estimate roughly $2B for storage costs of raw imaging data of a human brain. (Collins et al., 2025) Hence, a 10x-compressed image of a complete human brain would cost approximately $200 million at 2023 prices. Storage costs decline roughly 10-fold per decade (Collins et al., 2025). The final simulation data would not be the raw imaging archive but a highly annotated connectome graph, which is orders of magnitude more storage-efficient.

Therefore, while initial imaging data will be expensive to store, the actual connectome representation needed for simulation is manageable with current datacenter technology.

Lastly, we note that neuromorphic systems can implement asynchronous, event-driven computation, significantly reducing

both latency requirements and energy consumption in large-scale neural simulations. The field has seen steady progress, from IBM's TrueNorth in 2014 (implementing 1 million neurons and 256 million synapses per chip) (Akopyan et al., 2015) to Intel's Loihi platforms (Davies et al., 2018, Orchard et al., 2021) and more recent developments like the Hala Point system at Sandia National Laboratories, with over 1.15 billion neurons and 128 billion synapses (Intel, 2024).

While massive investment in AI accelerators -- orders of magnitude larger than in neuromorphic computing -- has led to a research landscape where repurposing AI hardware for biological neural-network simulation is often the preferred approach due to its widespread availability, mature software ecosystems, and sheer raw computational capability (Landsmeer et al., 2024; Deistler et al., 2024), neuromorphic and other non-von Neumann architectures remain promising paths to address compute, storage, memory, and energy demands.

Benchmarking

Metrics of Success

What constitutes success?

Dedicated benchmarking work for brain emulations is still sparse, though recent efforts such as ZAPBench and broader synthesis work have begun to articulate the problem (Lueckmann et al., 2025; Zanichelli et al., 2025).

The space of possible metrics is effectively unbounded, and we do not yet know which metrics best capture the functional essence of brain computation--a situation unlike protein structure prediction or weather forecasting, where success criteria are well defined. Ideal benchmarks will differ between possible goals: Sub-neuron level accuracy, neuron-level accuracy,

behavioural goals, accurate causal models, usefulness on tasks, etc.

We first and foremost emphasize the need for a thorough benchmarking sub-field, developing various benchmarks for various purposes. Below we briefly touch on the state of the field and a few early key ideas to initiate discussion.

Deterministic Metrics

One could take a recording of C seconds of whole-brain neural activity and try to predict the next T seconds, forming a metric of accuracy by comparing future activity vectors neuron by neuron and averaging the deviation over time. This is the spirit of ZAPBench (Lueckmann et al., 2025), which proposes, for example, T=30 seconds and evaluates mean absolute error frame by frame and neuron by neuron before averaging to a final score. The score is repeated across tasks in the zebrafish recordings, such as exposure to flashing light versus sinusoid drifting gratings. Related approaches have been applied to C. elegans whole-brain data (Simeon et al., 2024; Creamer et al., 2024) and to human non-invasive brain decoding (Tang et al., 2023).

Stochastic Issues

A key issue with these approaches is that they do not account for the stochastic and chaotic nature of much neural activity. In physical brains, stochasticity is inherent, e.g., due to thermal noise affecting ion channel dynamics (Faisal et al., 2008). As previously noted (Sandberg and Bostrom, 2008), deterministic metrics such as mean absolute error (MAE) do not capture the chaotic dynamics. For example, in a free-running simulation two zebrafish models might start in similar initial conditions, but subsequent brain states are likely to diverge, due to the property of chaotic systems that they are highly sensitive to

small deviation in initial conditions, as in the “butterfly effect” (Lorenz, 1963). How can we capture this?

Our goal needn’t necessarily be to carry forward the dynamical activity trajectory of a set of neurons deterministically for a long time, as in the ZapBench benchmark. Instead, our high-level goal is to “accurately capture the learned behaviors, memories and computations of individual animals”. Thus, we must accurately model the *underlying dynamics* of the nonlinear, stochastic system - including all its attractors and other key types of memory states or patterns - but not necessarily predict a given trajectory of that dynamical system perfectly.

Similarly, climate models can predict complex climate effects such as El Nino, even though weather models cannot predict exact weather more than a few weeks ahead.

More generally, ideas from chaotic dynamical systems theory apply here. In that field, it has been observed that one can reconstruct the attractor structure of a system even from limited and noisy observations (Takens, 1981; Sauer et al., 1991). Here, we briefly sketch several approaches relevant to metrics given the potentially chaotic and stochastic nature of nervous systems.

Stochastic Distribution Matching

How would one measure brain state typicality? A naive way to measure brain state typicality is to compare activity distributions of neurons, circuits, or regions. For example, a specific cortical neuron might be active with an average 10 Hz spike rate with a ±5 SD, whereas the simulated neuron is erroneously active with a typical 20 Hz ± 10 spike rate. A naive error metric is the difference between spike rates. Other distributional metrics could be Total Variation Distance, KL Divergence, or Jensen-Shannon Divergence, or, as was used in a large-scale simulation of the mouse primary visual cortex (Billeh et al., 2020), a KS test-based similarity between distributions

over multiple neurons. Analogous to ZAPBench, averaging this error across all neurons in a population yields a provisional total score.

Of course, distributions can also be computed across all neurons in a brain region of interest and compared on a population-level instead.

For example, Rimehaug et al. (2023) demonstrated quantitative comparison of simulated and experimentally recorded local field potentials (LFP) in mouse primary visual cortex, using a biophysical model of over 50,000 neurons to reproduce current source density patterns--an approach that validates simulations at a mesoscale level inaccessible to single-neuron metrics alone.

Finally, neuronal activity varies substantially across brain states. Thus, a distributional metric might be most useful when conditioned on brain state.

Distribution metrics can be chosen based on empirical usefulness and robustness, such as total variation distance, KL divergence, Jensen-Shannon divergence, or another suitable measure of distributional difference.

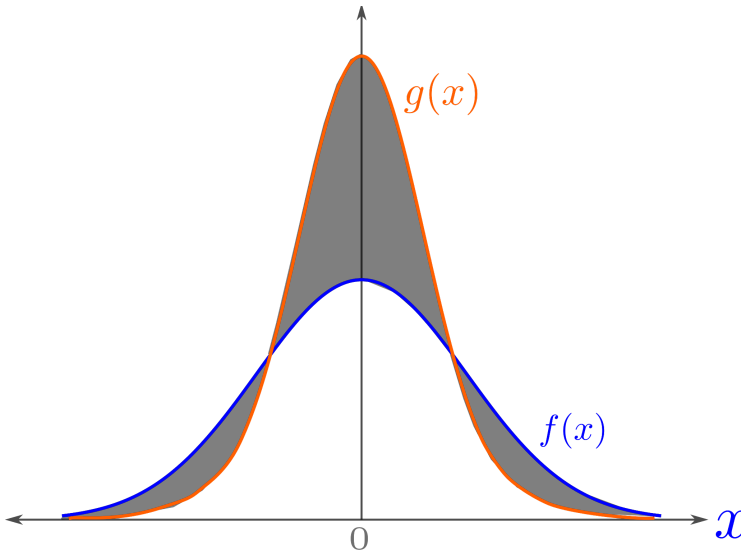

Figure 11. Schematic illustration of total variation distance between two distributions. Adapted from [Wikipedia: Total variation distance of probability measures](#).

If ground truth neuronal activity is not given, such as in calcium imaging, deconvolution methods can estimate spike rates (Vogelstein et al., 2009). Direct electrophysiological or voltage-imaging recordings can bypass this limitation.

If brain state is not explicitly given by e.g. a performed task, one could incorporate a classification step to determine which "brain state" is active under a particular stimulus. Different stimuli might produce distinct patterns of neural activity, allowing a classifier to differentiate responses to, for example, flashing light versus drifting gratings (Chen et al., 2018).

If conditional on the task at hand, a model produces statistically indistinguishable firing rates, this may be an impressive feat. However, this measure of typicality is not sufficient to guarantee causal dynamics, as models may learn per-neuron activity distributions directly rather than the causal rules governing inter-neuron influence.

Behavioural Metrics

Ultimately, effective “systems-level” verification occurs when a simulated organism exhibits real-like behavior in a virtual environment. For instance, a *C. elegans* model may be tested on chemotaxis or thermotaxis tasks, while a larval zebrafish model can be validated on optomotor or phototactic tasks. Benchmarks might include reflex latencies or typical locomotor patterns--each confirming whether the underlying neural computations match biological norms.

A complementary benchmark is stimulus-evoked response fidelity: When the organism receives known stimuli, does the simulation generate comparable neuronal responses and behavioral outputs? This bridges neural and behavioral levels and can be quantified via response latencies, tuning curves, and behavioral choice probabilities.

The embodied Turing test is holistic and intuitive: If a simulated C. elegans navigates chemical gradients, responds to touch, and forages like a real worm, something important has been captured. But behavioral equivalence is neither necessary nor sufficient for mechanistic accuracy--a lookup table could pass behavioral tests without representing any neural mechanism. Behavioral benchmarks are therefore best combined with neural-level metrics, not used as standalone criteria. (Zador et al., 2023)

One emerging framework for validating brain emulations is the 'embodied Turing test' proposed by Zador et al. (2023), which evaluates whether a simulated nervous system can replicate the behavioral repertoire of the original organism when embedded in an equivalent body (physical or simulated). Unlike traditional Turing tests focused on language, this framework evaluates sensorimotor integration, learning, and adaptive behavior -- capacities that depend on the full neural circuit rather than a single cognitive module. Initial practical benchmarks are

beginning to emerge for C. elegans and Drosophila, where both connectomes and behavioral libraries are available.

Benchmark Suites

Given the many dimensions among which to measure model performance, we propose to focus on benchmark suites rather than single benchmarks. We expect no absolute threshold to define a "successful" simulation. Instead, there are successively better approximations that capture more and more. We hope for the field to develop harder benchmark suites guiding ever-better simulations of brains (Einevoll et al., 2019). This mirrors current AI practice, where no single benchmark defines a model's strength, but rather a suite of benchmarks across various tasks. (Srivastava et al., 2022, Liang et al., 2022) (Chiang et al., 2024; Kiela et al., 2021)

Critically, the appropriate benchmarks depend on the goals of the simulation. A simulation aimed at behavioral fidelity (reproducing learned behaviors in a virtual environment) demands different metrics than one aimed at mechanistic accuracy (capturing causal dynamics under perturbation) or applied utility (building better sensory processing models for robotics). We organize metrics below by type rather than importance, acknowledging that different research programs will weigh them differently.

An early example of a benchmarking suite is Brain-Score (Schrimpf et al., 2020; brain-score.org), which evaluates how well artificial neural networks predict neural responses in the primate visual system across multiple benchmarks simultaneously. Brain-Score demonstrates that no single metric suffices and that composite benchmark suites -- combining neural predictivity, behavioral alignment, and anatomical correspondence -- are needed. A similar multi-metric approach has been advocated for

brain simulations (Arkhipov et al., 2018), and the same logic should apply to whole-brain models.

A practical approach is to focus on metrics that are experimentally accessible and scientifically informative, accepting that any simulation will be trained on limited data and must generalize beyond it. For context, Billeh et al. (2020) trained a V1 model on one second of spontaneous activity and a single drifting grating, then validated on over 100 seconds of diverse stimuli. Achieving robust generalization will require organized community efforts: standardized competitions (akin to CASP for protein folding), regular workshops, and evolving benchmark suites that prevent saturation on any single metric.

To illustrate how these goals translate to concrete criteria, consider what "successful mouse brain emulation" might mean. A behavioral benchmark suite might include: Navigation in novel mazes, fear conditioning and extinction, social interaction patterns, and circadian activity rhythms. Neural correspondence benchmarks would require that simulated neurons in the hippocampus exhibit place fields, that visual cortex neurons show orientation tuning, and that population dynamics during sleep replay match recorded patterns.

A particularly stringent test would be personality matching--do individual simulated mice, initialized from different biological individuals' connectomes, show the behavioral variability observed in the source animals? And crucially, the simulation must make novel predictions: if it predicts a specific behavioral deficit from a specific perturbation, and that prediction is confirmed experimentally, this provides strong evidence for mechanistic accuracy.

The Pressing Need for Benchmarks

In short, like in recent AI progress, benchmarks will be necessary to accurately guide development, quantification, and comparison of a diversity of brain emulations.

Given the size of the possible space of benchmarks, our notes here are necessarily incomplete. Among other approaches, we do not cover:

- Computational mechanics (Shalizi and Crutchfield, 1999), which automatically discovers the temporal and spatial scales at which a system's dynamics become predictable
- Machine learning using automated discovery of state variables: Chen et al. (2022)
- Metrics of causality at various scales, Granger causality, transfer entropy.
- The use of perturbation-based evidence for genuine causality (Haspel et al., 2023) and privileging perturbation data (optogenetic, pharmacological, lesion) over purely observational metrics.
- Metrics that enable success for out-of-distribution generalization of neuron activity and behaviour
- Benchmarks across spatial scales: Measurable neuronal activity spans from individual ion channels opening to whole-brain activity.
- Benchmarks across timescales: Neuronal activity ranges from sub-millisecond dynamics of ion channels to decades.
- Benchmarks for adaptive behaviours such as short-term synaptic depression, long-term potentiation and neuroplasticity

We look forward to an extensive sub-field of neuroscience asking this important question of: How do we measure success?

The Mega-Project

It took decades to go from sequencing the bacteriophage φX174, then *C. elegans*, *Drosophila*, and *Mus musculus* to reach a complete human genome (Sanger et al., 1977, C. elegans Sequencing Consortium, 1998, Adams et al., 2000, Mouse Genome Sequencing Consortium, 2002, International Human Genome Sequencing Consortium, 2001).

As noted before, the Manhattan Project cost over $30B (inflation adjusted) (Schwartz, 1998) and took 3 years to develop the atomic bomb (Groves, 1962). Apollo cost roughly $257B in 2020 dollars (Dreier, 2022) and took about eight years from President Kennedy's May 25, 1961 commitment to the Apollo 11 landing in July 1969 (Loff, 2015).

These approximate bounds suggest that human-scale connectomics and simulation might be on the scale of investment between the Human Genome Project and the Manhattan Project, while costing less than the Apollo Program. A rough range based on our estimates and interviews is $5-50B over 10-25 years.

In the absence of required data, we restrain from forecasting more detailed numbers. Estimates around such ambitious projects are especially sensitive to how much the underlying technology improves before and during the project. We are broadly optimistic about the improvement in light microscopy, molecular identification, electron microscopy, reconstruction and proofreading, and simulation techniques.

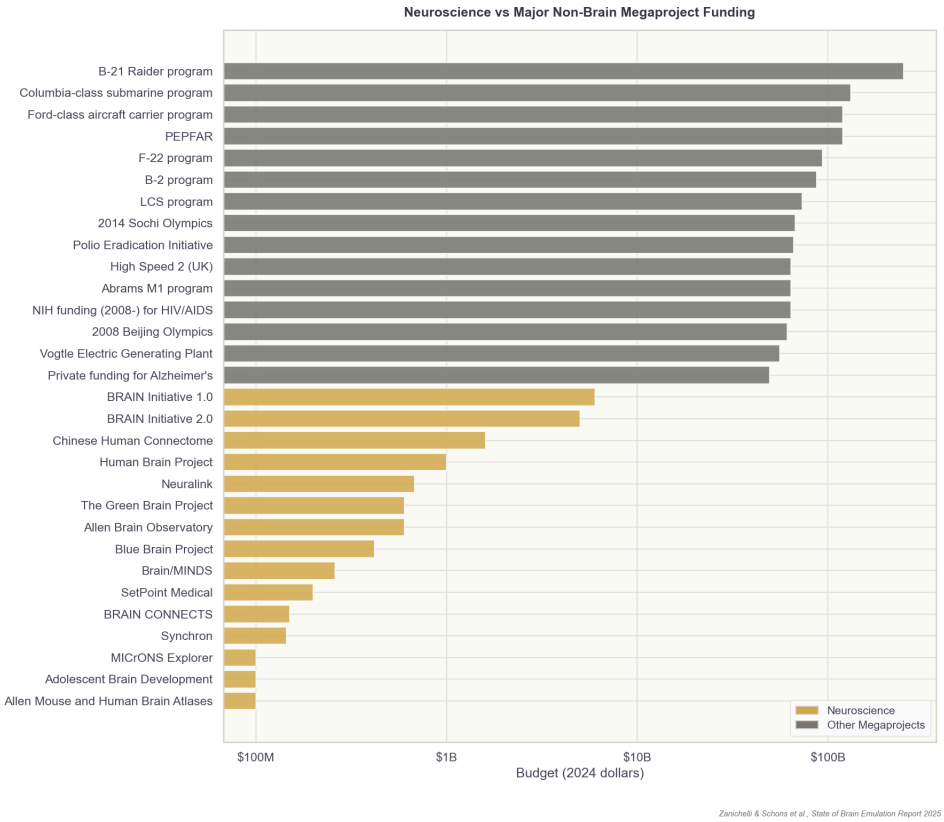

Figure 12. Megaproject examples. Adapted from Zanichelli et al. (2025).

Modern machine learning may be the single largest driver of cost reduction.

Modern machine learning has generated highly accurate predictive models across weather, biology, and natural language processing, among other domains. These same systems are increasingly being adapted into scientific assistants and tool-using workflows (Bran et al., 2024). They will likely improve further as AI companies continue investing heavily in larger underlying models (Perrault & Clark, 2024).

Machine learning could reduce the amount of resources needed. As a concrete example, SmartEM, a single beam electron microscopy

approach that demonstrated a 7-fold decrease in acquisition time by adjusting the microscope dwell time based on the difficulty of a given area would subsequently be to segment (Meirovitch et al., 2024). Another example, in-detector ML-based data compression could help sidestep data transfer limits for imaging approaches currently bottlenecked by detector capabilities such as X-ray ptychography (Du et al., 2021, Valentin et al., 2023).

Machine learning has already transformed connectomics proofreading, such as Google's flood-filling networks or segmentation-guided contrastive learning (SegCLR). As these models improve, the dominant cost of connectomics -- human proofreading, which currently accounts for over 90% of project budgets -- should fall dramatically.

Machine learning also offers paths to faster simulation. Neural network surrogate models can approximate the dynamics or required parameters of biophysically detailed neuron models faster than numerical integration. More ambitiously, connectome-constrained neural networks have begun to predict neural activity directly from wiring diagrams: Lappalainen et al. (2024) used the Drosophila connectome to predict responses across 64 neuron types in the visual system, reproducing findings from over 20 prior experimental studies. If such structure-to-function methods generalize to mammalian circuits, they could partially substitute for the functional recording data that remains hardest to collect at scale.

AI scaling laws -- the empirical finding that model performance improves as a predictable power-law function of compute, data, and parameters (Kaplan et al., 2020) -- might be a useful analogy for brain simulation. If simulation accuracy scales similarly with the quantity and resolution of input data, then each incremental improvement in connectomics throughput, functional recording density, or molecular profiling would yield predictable gains in simulation fidelity, even before fundamental algorithmic breakthroughs.

As AI companies continue to build ever-larger clusters, the marginal cost of allocating a fraction of that capacity to neuroscience simulation shrinks.

Storage is a parallel bottleneck. A single cubic millimeter of brain tissue imaged at EM resolution generates roughly a petabyte of raw data; a full human brain would produce exabytes. Cloud storage costs have fallen roughly 10-fold per decade, but the scale of whole-brain datasets will still require purpose-built archival infrastructure -- likely co-located with compute to minimize data transfer costs, analogous to how genomics centers co-locate sequencing instruments and analysis clusters.

Any serious large-scale connectomics effort would require automation at every stage of the pipeline. The semiconductor industry offers a useful analogy: chip fabrication involves cleaning, etching, doping, and layering at nanometer precision with extraordinarily low error rates, all inside billion-dollar fabs that run continuously with minimal human intervention. Brain sample preparation -- fixation, slicing, staining, embedding -- demands comparable precision at comparable scales, yet remains largely artisanal. Transitioning from hand-operated microtomes and manual staining protocols to automated, high-throughput sample preparation lines is arguably as important as any improvement in microscopy or reconstruction software.

The same applies downstream: automated quality control, real-time segmentation during acquisition, and closed-loop imaging systems that re-scan ambiguous regions could dramatically reduce both cost and error accumulation in large-scale projects.

.||. / |.|.

Closing

This thesis is necessarily incomplete. We have not covered: non-connectomics approaches to emulation, including functional paradigms that bypass detailed structural reconstruction; structure-to-function inference methods that remain in their infancy; X-ray microscopy and other emerging imaging modalities; the ethical, legal, and philosophical dimensions of creating digital minds. We have focused on what we believe to be one possible central pipeline -- connectomics, functional recording, and simulation.

What would it take to emulate a human brain? This thesis has surveyed the three core pillars -- structural mapping, functional recording, and simulation -- and found that each has advanced substantially since the Sandberg and Bostrom roadmap of 2008, yet none is close to the scale required for whole human brain emulation.

In connectomics, the completion of the Drosophila whole-brain connectome (Dorkenwald et al., 2024; Schlegel et al., 2024) and the emergence of petascale mouse-cortex connectomics programs mark genuine inflection points. On the recording front, calcium imaging and multi-electrode arrays have scaled from tens to tens of thousands of simultaneous neurons, but whole-brain coverage at single-neuron, single-spike resolution in mammals remains beyond current reach. On the simulation side, GPU-accelerated simulators and neuromorphic hardware have made real-time simulation of cortical-column-scale circuits feasible, but the memory wall and interconnect bottleneck remain formidable at whole-brain scale.

The path forward is not a single moonshot but a series of interlocking achievements: complete molecularly-annotated connectomes, paired with whole-brain functional recording at cellular resolution, feeding validated biophysically detailed simulations -- each organism a proving ground for the next. Whether the result constitutes emulation in the strong sense -- a system that reproduces not only connectivity but functional dynamics and emergent properties -- remains an open empirical question.

The field's trajectory suggests that the question is shifting from whether to when and how. The answer will depend not only on technological advances but on sustained, large-scale, coordinated effort that has historically delivered humanity's most consequential achievements. I hope this thesis serves as a useful map of the terrain -- inevitably incomplete, but grounded in the engineering realities of 2025 and in the work that I carried out at MIT.

For readers interested in the collaborative synthesis that grew out of this work, see the *State of Brain Emulation Report* (Zanichelli et al., 2025) at brainemulation.mxschons.com. Follow the researchers cited throughout -- they are at the frontier of neuroscience and brain emulation in real time. If this field interests you, get involved. If you would like to support the field, many of the mentioned research efforts are funding-constrained. If I can help you get involved, reach out anytime at axon@mit.edu and i@isaak.net.

The challenges are vast, the community is welcoming, and the potential impact is extraordinary: This may be one of humanity's most pressing quests to embark on.

> *To strive, to seek, to find, and not to yield.*
> - Ulysses

References

1. Abbott, L. F. (1999). Lapicque's introduction of the integrate-and-fire model neuron (1907). Brain Research Bulletin, 50(5-6), 303-304. https://doi.org/10.1016/S0361-9230(99)00161-6
2. Abi Akar, N., Cumming, B., Karakasis, V., Küsters, A., Klijn, W., Peyser, A., & Yates, S. (2019). Arbor -- a morphologically-detailed neural network simulation library for contemporary high-performance computing architectures. arXiv preprint arXiv:1901.07454. https://arxiv.org/abs/1901.07454
3. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Willmore, L. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 630(8025), 493-500. https://doi.org/10.1038/s41586-024-07487-w
4. Adam, Y., Kim, J. J., Lou, S., Zhao, Y., Xie, M. E., Brinks, D., ... & Cohen, A. E. (2019). Voltage imaging and optogenetics reveal behaviour-dependent changes in hippocampal dynamics. Nature, 569(7756), 413-417. https://doi.org/10.1038/s41586-019-1166-7
5. Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., ... & Venter, J. C. (2000). The genome sequence of Drosophila melanogaster. Science, 287(5461), 2185-2195. https://doi.org/10.1126/science.287.5461.2185
6. Ahrens, M. B., Li, J. M., Orger, M. B., Robson, D. N., Schier, A. F., Engert, F., & Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. Nature, 485(7399), 471-477. https://doi.org/10.1038/nature11057
7. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., & Keller, P. J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nature Methods, 10(5), 413-420. https://doi.org/10.1038/nmeth.2434
8. Aimon, S., Katsuki, T., Jia, T., Grosenick, L., Broxton, M., Deisseroth, K., ... & Cohen, A. E. (2019). Fast near-whole-brain imaging in adult Drosophila during responses to stimuli and behavior. PLOS Biology, 17(2), e2006732. https://doi.org/10.1371/journal.pbio.2006732
9. Akopyan, F., Sawada, J., Cassidy, A., Alvarez-Icaza, R., Arthur, J., Merolla, P., ... & Modha, D. S. (2015). TrueNorth: Design and tool flow of a 65 mW 1 million neuron programmable neurosynaptic chip. IEEE

Transactions on Computer-Aided Design of Integrated Circuits and Systems, 34(10), 1537-1558. https://doi.org/10.1109/TCAD.2015.2474396

10. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Molecular Biology of the Cell (4th ed.). Garland Science, New York.

11. Almog, M., & Korngreen, A. (2016). Is realistic neuronal modeling realistic? Journal of Neurophysiology, 116(5), 2180-2209. https://doi.org/10.1152/jn.00360.2016

12. Alon, S., Goodwin, D. R., Sinha, A., Wassie, A. T., Chen, F., ... & Boyden, E. S. (2021). Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. Science, 371(6528), eaax2656. https://doi.org/10.1126/science.aax2656

13. Amodei, D. (2024). Machines of loving grace: How AI could transform the world for the better. Dario Amodei Blog. https://darioadmodei.com/machines-of-loving-grace

14. Amsalem, O., Inagaki, H. K., Yu, J., Svoboda, K., & Darshan, R. (2024). Sub-threshold neuronal activity and the dynamical regime of cerebral cortex. Nature Communications, 15, 7958. https://doi.org/10.1038/s41467-024-51390-x

15. An, W., Xie, Z., Chen, H., Wang, Q., Yang, S., Wang, H., Yang, G., & Tian, Y. (2024). A survey on the GPU memory wall. arXiv preprint arXiv:2408.14158. https://arxiv.org/abs/2408.14158

16. Anastassiou, C. A., Perin, R., Markram, H., & Koch, C. (2011). Ephaptic coupling of cortical neurons. Nature Neuroscience, 14, 217-223. https://doi.org/10.1038/nn.2727

17. Antinucci, P., & Hindges, R. (2016). A crystal-clear zebrafish for in vivo imaging. Scientific Reports, 6, 29490. https://doi.org/10.1038/srep29490

18. ARC Prize. (2024). ARC Prize 2024: Abstraction and reasoning corpus benchmark. https://arcprize.org/

19. Arkhipov, A., Gouwens, N. W., Billeh, Y. N., Gratiy, S. L., Iyer, R., Wei, Z., Xu, Z., Abbasi-Asl, R., ... & Koch, C. (2018). Visual physiology of the layer 4 cortical circuit in silico. PLOS Computational Biology, 14(11), e1006535. https://doi.org/10.1371/journal.pcbi.1006535

20. Arnatkeviciute, A., Fulcher, B. D., & Fornito, A. (2018). A practical guide to linking brain-wide gene expression and neuroimaging data. NeuroImage, 189, 353-367. https://doi.org/10.1016/j.neuroimage.2019.01.011

21. Aschenbrenner, L. (2024). Situational awareness: The decade ahead. https://situational-awareness.ai/

22. Azevedo, A. W., Lesser, E., Phelps, J. S., Mark, B., Elabbady, L., ... & Tuthill, J. C. (2024). Connectomic reconstruction of a female Drosophila ventral nerve cord. Nature, 632(8024), 352-359. https://doi.org/10.1038/s41586-024-07389-x

23. Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., ... & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology, 513(5), 532-541. https://doi.org/10.1002/cne.21974

24. Backblaze. (2022). Hard drive stats for 2022. Backblaze Blog. https://www.backblaze.com/blog/backblaze-drive-stats-for-2022/

25. Badea, A., Ali-Sharief, A., & Johnson, G. A. (2007). Morphometric analysis of the C57BL/6J mouse brain. NeuroImage, 37(3), 683-693. https://doi.org/10.1016/j.neuroimage.2007.06.029

26. Bahri, Y., Dyer, E., Kaplan, J., Lee, J., & Sharma, U. (2024). Explaining neural scaling laws. Proceedings of the National Academy of Sciences, 121(27), e2311878121. https://doi.org/10.1073/pnas.2311878121

27. Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature, 533(7604), 452-454. https://doi.org/10.1038/533452a

28. Bargmann, C. I. (2012). Beyond the connectome: How neuromodulators shape neural circuits. BioEssays, 34(6), 458-465. https://doi.org/10.1002/bies.201100185

29. Bargmann, C. I., & Marder, E. (2013). From the connectome to brain function. Nature Methods, 10(6), 483-490. https://doi.org/10.1038/nmeth.2451

30. Bazinet, V., Hansen, J. Y., & Misic, B. (2023). Towards a biologically annotated brain connectome. Nature Reviews Neuroscience, 24(9), 555-566. https://doi.org/10.1038/s41583-023-00752-3

31. Beiran, M., & Litwin-Kumar, A. (2024). Prediction of neural activity in connectome-constrained recurrent networks. bioRxiv. https://doi.org/10.1101/2024.02.22.581667

32. Beniaguev, D., Segev, I., & London, M. (2021). Single cortical neurons as deep artificial neural networks. Neuron, 109(17), 2727-2739. https://doi.org/10.1016/j.neuron.2021.07.002

33. Berg, S., Nern, A., Schlegel, P., Matsliah, A., Dorkenwald, S., et al. (2025). Sexual dimorphism in the complete connectome of the

Drosophila male central nervous system. bioRxiv. https://doi.org/10.1101/2025.10.09.680999

34. Billeh, Y. N., Cai, B., Gratiy, S. L., Dai, K., Iyer, R., Gouwens, N. W., ... & Arkhipov, A. (2020). Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. Neuron, 106(3), 388-403.e18. https://doi.org/10.1016/j.neuron.2020.01.040

35. Bommasani, R., Hudson, D. A., Adeli, E., ... & Liang, P. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258. https://arxiv.org/abs/2108.07258

36. Bosch, C., Ackels, T., Pacureanu, A., Zhang, Y., Peddie, C. J., ... & Schaefer, A. T. (2022). Functional and multiscale 3D structural investigation of brain tissue through correlative in vivo physiology, synchrotron microtomography and volume electron microscopy. Nature Communications, 13, 2923. https://doi.org/10.1038/s41467-022-30199-6

37. Bosch, C., Collinson, L., Costa, A., Jefferis, G., & Schlegel, P. (2023). Wellcome Trust report: Scaling up connectomics. Wellcome Trust. https://wellcome.org/reports/scaling-connectomics

38. Bostrom, N. (2013). Existential risk prevention as global priority. Global Policy, 4(1), 15-31. https://doi.org/10.1111/1758-5899.12002

39. Bran, A. M., Cox, S., Schilber, O., Baldassari, C., White, A. D., & Schwaller, P. (2024). Augmenting large language models with chemistry tools. Nature Machine Intelligence, 6(5), 525-535. https://doi.org/10.1038/s42256-024-00832-8

40. Breakspear, M. (2017). Dynamic models of large-scale brain activity. Nature Neuroscience, 20, 340-352. https://doi.org/10.1038/nn.4497

41. Brette, R., & Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. Journal of Neurophysiology, 94(5), 3637-3642. https://doi.org/10.1152/jn.00686.2005

42. Brittin, C. A., Cook, S. J., Hall, D. H., Emmons, S. W., & Cohen, N. (2018). Volumetric reconstruction of main neurons in the brain of the nematode C. elegans. Neural Computation, 30(2), 443-460.

43. Brittin, C. A., Cook, S. J., Hall, D. H., Emmons, S. W., & Cohen, N. (2021). A multi-scale brain map derived from whole-brain volumetric reconstructions. Nature, 591(7848), 105-110. https://doi.org/10.1038/s41586-021-03284-x

44. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., ... & Amodei, D. (2020). Language models are few-shot learners. In Advances

in Neural Information Processing Systems, 33, 1877-1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

45. C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode C. elegans: A platform for investigating biology. Science, 282(5396), 2012-2018. https://doi.org/10.1126/science.282.5396.2012

46. Cai, D., Cohen, K. B., Luo, T., Lichtman, J. W., & Sanes, J. R. (2013). Improved tools for the Brainbow toolbox. Nature Methods, 10(6), 540-547. https://doi.org/10.1038/nmeth.2450

47. Campagnola, L., Seeman, S. C., Chartrand, T., Kim, L., Hoggarth, A., Gamlin, C., ... & Ko, A. L. (2022). Local connectivity and synaptic dynamics in mouse and human neocortex. Science, 375(6585), eabj5861. https://doi.org/10.1126/science.abj5861

48. Campbell, M., Hoane, A. J., & Hsu, F. (2002). Deep Blue. Artificial Intelligence, 134(1-2), 57-83. https://doi.org/10.1016/S0004-3702(01)00129-1

49. Caron, S. J. C., Ruta, V., Abbott, L. F., & Axel, R. (2013). Random convergence of olfactory inputs in the Drosophila mushroom body. Nature, 497(7447), 113-117. https://doi.org/10.1038/nature12063

50. CERN. (n.d.-a). CERN Data Centre. https://home.web.cern.ch/science/computing/data-centre

51. CERN. (n.d.-b). Facts and figures about the LHC. https://home.cern/resources/faqs/facts-and-figures-about-lhc

52. Chan, K. Y., Jang, M. J., Yoo, B. B., Greenbaum, A., Ravi, N., ... & Gradinaru, V. (2017). Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. Nature Neuroscience, 20(8), 1172-1179. https://doi.org/10.1038/nn.4593

53. Chen, F., Tillberg, P. W., & Boyden, E. S. (2015). Expansion microscopy. Science, 347(6221), 543-548. https://doi.org/10.1126/science.1260088

54. Chen, G., Scherr, F., & Maass, W. (2022). A data-based large-scale model for primary visual cortex enables brain-like robust and versatile visual processing. Science Advances, 8(44), eabq7592. https://doi.org/10.1126/sciadv.abq7592

55. Chen, X., Mu, Y., Hu, Y., Kuan, A. T., Nikitchenko, M., Randlett, O., Chen, A., Gavornik, J. P., Sompolinsky, H., Engert, F., & Ahrens, M. B. (2018). Brain-wide organization of neuronal activity and convergent sensorimotor transformations in larval zebrafish. Neuron, 100(4), 876-890.e5. https://doi.org/10.1016/j.neuron.2018.09.042

56. Chiang, W.-L., Zheng, L., Sheng, Y., Angelopoulos, A. N., Li, T., Li, D., Zhang, H., Zhu, B., Jordan, M. I., Gonzalez, J. E., & Stoica, I. (2024). Chatbot Arena: An open platform for evaluating LLMs by human preference. Proceedings of the 41st International Conference on Machine Learning (ICML 2024), 235, 8359-8388. https://proceedings.mlr.press/v235/chiang24b.html

57. Clarke, S. (2003). Henry Ford and mass production. In The Oxford Encyclopedia of Economic History. Oxford University Press.

58. Cobey, K. D., Fehlmann, C. A., Franco, M. C., Ayala, A. P., ... & Moher, D. (2024). Reproducibility challenges in neuroscience research. eLife, 13, e86365.

59. Collins, F. S., & McKusick, V. A. (2001). Implications of the Human Genome Project for medical science. JAMA, 285(5), 540-544. https://doi.org/10.1001/jama.285.5.540

60. Collins, L. T., Huffman, T., & Koene, R. A. (2025). Comparative prospects of imaging methods for whole-brain mammalian connectomics. Cell Reports Methods, 5(2), 100988. https://doi.org/10.1016/j.crmeth.2025.100988

61. Cook, S. J., Jarrell, T. A., Brittin, C. A., Wang, Y., ... & Emmons, S. W. (2019). Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature, 571(7763), 63-71. https://doi.org/10.1038/s41586-019-1352-7

62. Cowley, B. R., Minhas, J. S., Deguchi, S., Tse, J., Wang, S. S., et al. (2024). Neural activity prediction for populations of unknown size. Nature. https://doi.org/10.1038/s41586-024-07451-8

63. Cramer, S. W., Carter, R. E., Aronson, J. D., Kodandaramaiah, S. B., Ebner, T. J., & Chen, C. C. (2021). Through the looking glass: A review of cranial window technology for optical access to the brain. Journal of Neuroscience Methods, 354, 109100. https://doi.org/10.1016/j.jneumeth.2021.109100

64. Creamer, M. S., Leifer, A. M., & Pillow, J. W. (2024). Bridging the gap between the connectome and whole-brain activity in C. elegans. bioRxiv. https://doi.org/10.1101/2024.09.22.614271

65. Cremonesi, F., Hager, G., Wellein, G., & Schurmann, F. (2020). Analytic performance modeling and analysis of detailed neuron simulations. International Journal of High Performance Computing Applications, 34(4), 428-449. https://doi.org/10.1177/1094342020912528

66. Davies, M., Srinivasa, N., Lin, T.-H., Chinya, G., Cao, Y., ... & Wang, J. (2018). Loihi: A neuromorphic manycore processor with on-chip

learning. IEEE Micro, 38(1), 82-99. https://doi.org/10.1109/MM.2018.112130359

67. Davies, M., Wild, A., Orber, G., Cavigelli, L., Plank, B., Risbud, S., ... & Loihi 2 Team. (2021). Advancing neuromorphic computing with Loihi 2. IEEE Micro, 41(5), 14-23. https://doi.org/10.1109/JPROC.2021.3067593

68. de Vries, S. E. J., Lecoq, J. A., Buice, M. A., Groblewski, P. A., Ocker, G. K., Oliver, M., ... & Koch, C. (2020). A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. Nature Neuroscience, 23(1), 138-151. https://doi.org/10.1038/s41593-019-0550-9

69. DeFelipe, J., Alonso-Nanclares, L., & Arellano, J. I. (2002). Microstructure of the neocortex: Comparative aspects. Journal of Neurocytology, 31(3-5), 299-316. https://doi.org/10.1023/A:1024130211265

70. Deistler, M., Kadhim, K. L., Pals, M., Beck, J., Huang, Z., Gloeckler, M., ... & Macke, J. H. (2024). Differentiable simulation enables large-scale training of detailed biophysical models of neural dynamics. bioRxiv. https://doi.org/10.1101/2024.08.21.608979

71. Dorkenwald, S., Matsliah, A., Sterling, A., Schlegel, P., ... & Seung, H. S. (2024). Neuronal wiring diagram of an adult brain. Nature, 634(8028), 124-138. https://doi.org/10.1038/s41586-024-07558-0

72. Dreier, C. (2022). An improved cost analysis of the Apollo program. Space Policy, 60, 101476. https://doi.org/10.1016/j.spacepol.2022.101476

73. Du, M., Di, Z. W., Gursoy, D., Xian, R. P., Kozorovitskiy, Y., & Jacobsen, C. (2021). Upscaling X-ray nanoimaging to macroscopic specimens. Journal of Applied Crystallography, 54, 386-394. https://doi.org/10.1107/S1600576721002314

74. Dura-Bernal, S., Herrera, B., Lupascu, C., Marsh, B. M., Gandolfi, D., Marasco, A., ... & Arkhipov, A. (2024). Large-scale mechanistic models of brain circuits with biophysically and morphologically detailed neurons. Journal of Neuroscience, 44(40), e1236242024. https://doi.org/10.1523/JNEUROSCI.1236-24.2024

75. Eberle, A. L., Mikula, S., Schalek, R., Lichtman, J. W., & Denk, W. (2015). High-resolution, high-throughput imaging with a multibeam scanning electron microscope. Journal of Microscopy, 259(2), 114-120. https://doi.org/10.1111/jmi.12224

76. Ecker, A., Isbister, J. B., Pokorny, C., Bolaños-Puchet, S., Egas Santander, D., ... & Reimann, M. W. (2024). Modeling and simulation of neocortical micro- and mesocircuitry. Part II: Physiology and experimentation. bioRxiv. https://doi.org/10.1101/2023.05.17.541168

77. Eckstein, N., Bates, A. S., Hartenstein, V., ... & Funke, J. (2024). Neurotransmitter classification from electron microscopy images at synaptic sites in Drosophila melanogaster. Cell, 187(6), 1548-1563. https://doi.org/10.1016/j.cell.2024.03.034

78. Edman, P. (1950). Method for determination of the amino acid sequence in peptides. Acta Chemica Scandinavica, 4, 283-293.

79. Einevoll, G. T., Destexhe, A., Diesmann, M., Grün, S., Jirsa, V., de Kamps, M., ... & Schürmann, F. (2019). The scientific case for brain simulations. Neuron, 102(4), 735-744. https://doi.org/10.1016/j.neuron.2019.03.027

80. Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012). A large-scale model of the functioning brain. Science, 338(6111), 1202-1205. https://doi.org/10.1126/science.1225266

81. Emes, R. D., Pocklington, A. J., Anderson, C. N. G., ... & Grant, S. G. N. (2008). Evolutionary expansion and anatomical specialization of synapse proteome complexity. Nature Neuroscience, 11(7), 799-806. https://doi.org/10.1038/nn.2135

82. Emmons, S. W. (2015). The beginning of connectomics: A commentary on White et al. (1986). Philosophical Transactions of the Royal Society B, 370(1666), 20140309. https://doi.org/10.1098/rstb.2014.0309

83. Erdil, E., & Schneider-Joseph, H. (2024). Compute requirements for algorithmic innovation in frontier AI systems. Epoch AI Report. https://epochai.org/

84. Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. Nature Reviews Neuroscience, 9(4), 292-303. https://doi.org/10.1038/nrn2258

85. Fitch, W. T. (2023). Editorial: Cellular computation and cognition. Frontiers in Computational Neuroscience, 17, 1107876. https://doi.org/10.3389/fncom.2023.1107876

86. Friedrich, R. W., Jacobson, G. A., & Zhu, P. (2010). Circuit neuroscience in zebrafish. Current Biology, 20(8), R371-R381. https://doi.org/10.1016/j.cub.2010.02.039

87. Fuller-Wright, L. (2024). Princeton connectomics project update: Proofreading progress. Princeton University News. https://www.princeton.edu/news/

88. FutureHouse. (2025). Robin: Demonstrating end-to-end scientific discovery with a multi-agent system. https://www.futurehouse.org/research-announcements/demonstrating-end-to-end-scientific-discovery-with-robin-a-multi-agent-system

89. Galvan Fraile, J., et al. (2023). Systematic comparison of expansion microscopy protocols for DNA and RNA detection. Methods, 218, 1-13. https://doi.org/10.1016/j.ymeth.2023.08.015

90. Galván Fraile, J., Scherr, F., Ramasco, J. J., Arkhipov, A., Maass, W., & Mirasso, C. R. (2024). Modeling circuit mechanisms of opposing cortical responses to visual flow perturbations. PLOS Computational Biology, 20(3), e1011921. https://doi.org/10.1371/journal.pcbi.1011921

91. Gates, A. J., Gysi, D. M., Kellis, M., & Barabasi, A.-L. (2021). A wealth of discovery built on the Human Genome Project - by the numbers. Nature, 590(7845), 212-215. https://doi.org/10.1038/d41586-021-00314-6

92. GBD 2021 Nervous System Disorders Collaborators. (2024). Global, regional, and national burden of disorders affecting the nervous system, 1990-2021. The Lancet Neurology, 23(4), 344-381. https://doi.org/10.1016/S1474-4422(24)00038-3

93. Gholami, A., Yao, Z., Kim, S., Hooper, C., Mahoney, M. W., & Keutzer, K. (2024). AI and memory wall. In Proceedings of the IEEE International Symposium on High-Performance Computer Architecture. IEEE. https://doi.org/10.1109/MM.2024.3373763

94. Glaser, A. K., et al. (2024). Expansion-assisted selective plane illumination microscopy for nanoscale imaging of centimeter-scale tissues. eLife, 13, e91979. https://doi.org/10.7554/eLife.91979

95. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association, 102(477), 359-378. https://doi.org/10.1198/016214506000001437

96. Goertsen, D., Flytzanis, N. C., Goeden, N., Chuapoco, M. R., Cummins, A., Chen, Y., ... & Gradinaru, V. (2022). AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. Nature Neuroscience, 25(1), 106-115. https://doi.org/10.1038/s41593-021-00969-4

97. Gottweis, J., et al. (2025). Towards AI co-scientists. Google DeepMind Technical Report. https://storage.googleapis.com/coscientist_paper/ai_coscientist.pdf

98. Gouwens, N. W., Sorensen, S. A., Baftizadeh, F., Buice, M. A., Lee, B. R., Jarsky, T., ... & Murphy, G. J. (2020). Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells. Cell, 183(4), 935-953. https://doi.org/10.1016/j.cell.2020.09.057

99. Gouwens, N. W., Sorensen, S. A., Berg, J., Lee, C., Jarsky, T., Ting, J., et al. (2018). Systematic generation of biophysically detailed models for diverse cortical neuron types. Nature Communications, 9, 710. https://doi.org/10.1038/s41467-017-02718-3

100. Gouwens, N. W., Sorensen, S. A., Berg, J., Lee, C., Jarsky, T., Ting, J., et al. (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. Nature Neuroscience, 22(7), 1182-1195. https://doi.org/10.1038/s41593-019-0417-0

101. Groves, L.R. (1962). Now It Can Be Told: The Story of the Manhattan Project. Harper & Row, New York.

102. Guet-McCreight, A., Moradi Chameh, H., Mazza, F., Prevot, T. D., Valiante, T. A., Sibille, E., & Hay, E. (2024). In-silico testing of new pharmacology for restoring inhibition and human cortical function in depression. Communications Biology, 7, 245. https://doi.org/10.1038/s42003-024-05907-1

103. Gutierrez, G. J., & Wang, S. (2023). Gap junctions and neural circuits. Current Biology, 33(12), R549-R551. https://doi.org/10.1016/j.cub.2023.05.025

104. Hadjiabadi, D., Lovett-Barron, M., Raikov, I. G., Sparks, F. T., Liao, Z., Baraban, S. C., ... & Soltesz, I. (2021). Maximally selective single-cell target for circuit control in epilepsy models. Neuron, 109(16), 2556-2572. https://doi.org/10.1016/j.neuron.2021.06.007

105. Haguenau, F., Hawkes, P. W., Hutchison, J. L., ... & Williams, D. B. (2003). Key events in the history of electron microscopy. Microscopy and Microanalysis, 9(2), 96-138. https://doi.org/10.1017/S1431927603030101

106. Haspel, G., Baker, B., Beets, I., ... & Kording, K. P. (2023). The time is ripe to reverse engineer an entire nervous system. arXiv preprint arXiv:2308.06578. https://arxiv.org/abs/2308.06578

107. Hassenstein, B., & Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungswahrnehmung des Russelkafers Chlorophanus. Zeitschrift fur Naturforschung, 11b, 513-524.

108. Haufler, D., Ito, S., Koch, C., & Arkhipov, A. (2023). Simulations of cortical networks using spatially extended conductance-based neuronal models. The Journal of Physiology, 601(2), 461-485. https://doi.org/10.1113/JP284030

109. Hawking, S. (2016). This is the most dangerous time for our planet. The Guardian, December 1, 2016. https://www.theguardian.com/commentisfree/2016/dec/01/stephen-hawking-dangerous-time-planet-inequality

110. Helmstaedter, M. (2013). Cellular-resolution connectomics: Challenges of dense neural circuit reconstruction. Nature Methods, 10(6), 501-507. https://doi.org/10.1038/nmeth.2476

111. Henke, B. L., Gullikson, E. M., & Davis, J. C. (1993). X-ray interactions: Photoabsorption, scattering, transmission, and reflection at E = 50-30,000 eV, Z = 1-92. Atomic Data and Nuclear Data Tables, 54(2), 181-342. https://doi.org/10.1006/adnd.1993.1013

112. Herculano-Houzel, S., Mota, B., & Lent, R. (2006). Cellular scaling rules for rodent brains. Proceedings of the National Academy of Sciences, 103(32), 12138-12143. https://doi.org/10.1073/pnas.0604911103

113. Hildebrand, D. G. C., Cicconet, M., Torres, R. M., ... & Lee, W.-C. A. (2017). Whole-brain serial-section electron microscopy in larval zebrafish. Nature, 545(7654), 345-349. https://doi.org/10.1038/nature22356

114. Hill, S.L., Wang, Y., Riachi, I., Schürmann, F., & Markram, H. (2012). Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. Proceedings of the National Academy of Sciences, 109(42), E2885-E2894. https://doi.org/10.1073/pnas.1202128109

115. Hines, M. L., & Carnevale, N. T. (2001). NEURON: a tool for neuroscientists. The Neuroscientist, 7(2), 123-135. https://doi.org/10.1177/107385840100700207

116. Hinton, G., Bengio, Y., et al. (2023). Statement on AI risk. Center for AI Safety. https://www.safe.ai/statement-on-ai-risk

117. Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in

nerve. Journal of Physiology, 117(4), 500-544. https://doi.org/10.1113/jphysiol.1952.sp004764

118. Howells, M. R., Beetz, T., Chapman, H. N., ... & Starodub, D. (2009). An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. Journal of Electron Spectroscopy and Related Phenomena, 170(1-3), 4-12. https://doi.org/10.1016/j.elspec.2008.10.008

119. Huang, B., Babcock, H., & Zhuang, X. (2010). Breaking the diffraction barrier: Super-resolution imaging of cells. Cell, 143(7), 1198-1213. https://doi.org/10.1016/j.cell.2010.12.002

120. Hubbell, J. H., Veigele, W. J., Briggs, E. A., Brown, R. T., Cromer, D. T., & Howerton, R. J. (1975). Atomic form factors, incoherent scattering functions, and photon scattering cross sections. Journal of Physical and Chemical Reference Data, 4(3), 471-538. https://doi.org/10.1063/1.555523

121. Intel Corporation. (2024). Intel builds world's largest neuromorphic system: Hala Point. Intel Newsroom. https://www.intel.com/content/www/us/en/newsroom/news/intel-builds-worlds-largest-neuromorphic-system.html

122. International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. Nature, 409(6822), 860-921. https://doi.org/10.1038/35057062

123. Isbister, J. B., Ecker, A., Pokorny, C., Bolaños-Puchet, S., Egas Santander, D., ... & Reimann, M. W. (2026). Modeling and simulation of neocortical micro- and mesocircuitry. Part I: Anatomy. eLife, 13, e99688. https://doi.org/10.7554/eLife.99688

124. Izhikevich, E. M. (2003). Simple model of spiking neurons. IEEE Transactions on Neural Networks, 14(6), 1569-1572. https://doi.org/10.1109/TNN.2003.820440

125. Jacques, S. L. (2013). Optical properties of biological tissues: a review. Physics in Medicine & Biology, 58(11), R37-R61. https://doi.org/10.1088/0031-9155/58/11/R37

126. Januszewski, M. (2023). Google Research embarks on effort to map a mouse brain. Google Research. https://research.google/blog/google-research-embarks-on-effort-to-map-a-mouse-brain/

127. Januszewski, M., & Jain, V. (2024). Foundation models for connectomics. bioRxiv. https://doi.org/10.1101/2024.11.24.625067

128. Januszewski, M., Kornfeld, J., Li, P. H., Pope, A., Blakely, T., ... & Jain, V. (2018). High-precision automated reconstruction of neurons with flood-filling networks. Nature Methods, 15(8), 605-610. https://doi.org/10.1038/s41592-018-0049-4

129. Jarrell, T. A., Wang, Y., Bloniarz, A. E., ... & Emmons, S. W. (2012). The connectome of a decision-making neural network. Science, 337(6093), 437-444. https://doi.org/10.1126/science.1221762

130. Jirsa, V. K., Stacey, W. C., Quilichini, P. P., Ivanov, A. I., & Bernard, C. (2014). On the nature of seizure dynamics. Brain, 137(8), 2210-2230. https://doi.org/10.1093/brain/awu133

131. Joseph, J. (2023). The $600M cost of frontier AI training runs. Epoch AI Analysis. https://epochai.org/

132. Jouppi, N. P., Yoon, D. H., Kurian, G., Li, S., Patel, N., ... & Young, C. (2021). Ten lessons from three generations of tensor processing units. In Proceedings of the 48th International Symposium on Computer Architecture (pp. 1-14). https://doi.org/10.1145/3470496.3533033

133. Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). Scaling laws for neural language models. arXiv preprint arXiv:2001.08361. https://arxiv.org/abs/2001.08361

134. Kato, S., Kaplan, H. S., Schrödel, T., Skora, S., Lindsay, T. H., Yemini, E., Lockery, S., & Zimmer, M. (2015). Global brain dynamics embed the motor command sequence of Caenorhabditis elegans. Cell, 163(3), 656-669. https://doi.org/10.1016/j.cell.2015.09.034

135. Khanna, P., et al. (2024). Large-scale neural recordings with single-neuron resolution in human cortex using high-density CMOS microelectrode arrays. Nature Neuroscience, 27, 579-589. https://doi.org/10.1038/s41593-023-01520-x

136. Kiela, D., Bartolo, M., Nie, Y., Kaushik, D., Geiger, A., Wu, Z., Vidgen, B., Prasad, G., ... & Williams, A. (2021). Dynabench: Rethinking benchmarking in NLP. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 4110-4124. https://doi.org/10.18653/v1/2021.naacl-main.324

137. Kim, D. H., Kim, J., Marques, J. C., Grama, A., Hildebrand, D. G. C., Gu, W., Li, J. M., & Robson, D. N. (2017). Pan-neuronal calcium imaging with cellular resolution in freely swimming zebrafish. Nature Methods, 14(11), 1107-1114. https://doi.org/10.1038/nmeth.4429

138. Kim, T.H., Zhang, Y., Lecoq, J., Jung, J.C., Li, J., Zeng, H., Niell, C.M., & Bhatt, G. (2016). Long-term optical access to an estimated one million neurons in the live mouse cortex. Cell Reports, 17(12), 3385-3394. https://doi.org/10.1016/j.celrep.2016.12.004

139. Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. Nature Reviews Neuroscience, 17(5), 307-321. https://doi.org/10.1038/nrn.2016.22

140. Kojima, T. (1951). On the brain of the sperm whale (Physeter catodon L.). Scientific Report of the Whales Research Institute, 6, 49-72.

141. Kortemme, T. (2024). De novo protein design--From new structures to programmable functions. Cell, 187(3), 526-544. https://doi.org/10.1016/j.cell.2023.12.028

142. Kuan, A. T., Phelps, J. S., Thomas, L. A., ... & Lee, W.-C. A. (2020). Dense neuronal reconstruction through X-ray holographic nano-tomography. Nature Neuroscience, 23(12), 1637-1643. https://doi.org/10.1038/s41593-020-0704-9

143. Kunkel, S., Schmidt, M., Eppler, J. M., ... & Helias, M. (2014). Spiking network simulation code for petascale computers. Frontiers in Neuroinformatics, 8, 78. https://doi.org/10.3389/fninf.2014.00078

144. Kuriyama, R., Akira, K., Yamazaki, T., Arkhipov, A., Green, L., Herrera, B., Dai, K., et al. (2025). Microscopic-level mouse whole cortex simulation composed of 9 million biophysical neurons and 26 billion synapses on the supercomputer Fugaku. Proceedings of SC25. https://doi.org/10.1145/3712285.3759819

145. Landsmeer, L. P. L., De Schutter, E., & van Elburg, R. A. J. (2024). Efficient simulation of 3D reaction-diffusion in models of neurons and networks. Neurocomputing, 598, 127953. https://doi.org/10.1016/j.neucom.2024.127953

146. Lappalainen, J. K., Tschopp, F. D., Prakhya, S., McGill, M., Nern, A., Shinomiya, K., ... & Turaga, S. C. (2024). Connectome-constrained networks predict neural activity across the fly visual system. Nature, 634, 1132-1140. https://doi.org/10.1038/s41586-024-07939-3

147. Leiwe, M.N., Fujimoto, S., Baba, T., Moriyasu, D., Saha, B., Sakaguchi, R., Inagaki, S., & Imai, T. (2024). Automated neuronal reconstruction with super-multicolour Tetbow labelling and threshold-based clustering of colour hues. Nature Communications, 15, 5210. https://doi.org/10.1038/s41467-024-49455-y

148. Lemon, W. C., Pulver, S. R., Hockendorf, B., ... & Keller, P. J. (2015). Whole-central nervous system functional imaging in larval

Drosophila. Nature Communications, 6, 7924. https://doi.org/10.1038/ncomms8924

149. Leonard, M.K., et al. (2024). Large-scale single-neuron speech sound encoding across the depth of human cortex. Nature, 626, 593-602. https://doi.org/10.1038/s41586-023-06839-2

150. Li, P. H., Januszewski, M., Tyka, M. D., Maitin-Shepard, J., Blakely, T., Berger, D. R., ... & Jain, V. (2024). EM-Compressor: Electron microscopy image compression in connectomics with variational autoencoders. bioRxiv, 2024.07.07.601368. https://doi.org/10.1101/2024.07.07.601368

151. Li, Y., Lu, Z., Otecko, N. O., et al. (2021). Bitbow: a digital format of Brainbow enables highly efficient neuronal lineage tracing and morphology reconstruction in single brains. Frontiers in Neural Circuits, 15, 732183. https://doi.org/10.3389/fncir.2021.732183

152. Liang, P., Bommasani, R., et al. (2022). Holistic evaluation of language models. arXiv preprint arXiv:2211.09110. https://arxiv.org/abs/2211.09110

153. Lichtman, J. W. (2014). Connectome data storage: Low petabytes expected. Keynote at BRAIN Initiative Conference.

154. Lin, A., Yang, R., Dorkenwald, S., Matsliah, A., Sterling, A. R., ... & Murthy, M. (2024). Network statistics of the whole-brain connectome of Drosophila. Nature, 634(8028), 153-169. https://doi.org/10.1038/s41586-024-07615-6

155. Linghu, C., Johnson, S. L., Valdes, P. A., Shemesh, O. A., Park, W. M., ... & Boyden, E. S. (2020). Spatial multiplexing of fluorescent reporters. Cell, 183(6), 1479-1494. https://doi.org/10.1016/j.cell.2020.10.001

156. Liu, Y.T., Tao, C.L., Zhang, X., Xia, W., Shi, K., Bhatt, D.G., et al. (2020). Mesophasic organization of GABAA receptors in hippocampal inhibitory synapses. Nature Neuroscience, 23(3), 359-368. https://doi.org/10.1038/s41593-020-00729-w

157. Livet, J., Weissman, T. A., Kang, H., ... & Sanes, J. R. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature, 450(7166), 56-62. https://doi.org/10.1038/nature06293

158. Loff, S. (2015). Apollo program history and achievements. NASA History Office. https://www.nasa.gov/mission_pages/apollo/

159. Lorenz, E. N. (1963). Deterministic nonperiodic flow. Journal of the Atmospheric Sciences, 20(2), 130-141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

160. Lu, Y., Wang, H., Ko, H. R., et al. (2023). A 12-week protocol of EMBED-SEM for whole mouse brain connectomics. bioRxiv. https://doi.org/10.1101/2023.09.26.559613

161. Lu, Y., Wang, Y., Zhang, J., Liang, Y., Li, X., ... & Feng, J. (2024). Simulation and assimilation of the digital human brain. Nature Computational Science, 4(12), 890-898. https://doi.org/10.1038/s43588-024-00731-3

162. Lueckmann, J. M., Boelts, J., Gao, R., Nonnenmacher, M., & Macke, J. H. (2025). ZAPBench: A benchmark for whole-brain activity prediction from neural calcium imaging. arXiv preprint arXiv:2503.02618. https://doi.org/10.48550/arXiv.2503.02618

163. Manley, J., Lu, S., Barber, K., Demas, J., Kim, H., Meyer, D., Martinez Traub, F., & Vaziri, A. (2024). Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number. Neuron, 112(6), 885-898. https://doi.org/10.1016/j.neuron.2024.02.011

164. Mann, K., Gallen, C. L., & Clandinin, T. R. (2017). Whole-brain calcium imaging reveals an intrinsic functional network in Drosophila. Current Biology, 27(15), 2389-2396.e4. https://doi.org/10.1016/j.cub.2017.06.076

165. Manning, C. (2023). Princeton large-scale connectomics project: Two 91-beam electron microscopes. Princeton Neuroscience Institute.

166. Marblestone, A. H., Zamft, B. M., Maguire, Y. G., ... & Boyden, E. S. (2013). Physical principles for scalable neural recording. Frontiers in Computational Neuroscience, 7, 137. https://doi.org/10.3389/fncom.2013.00137

167. Markram, H. (2006). The Blue Brain Project. Nature Reviews Neuroscience, 7(2), 153-160. https://doi.org/10.1038/nrn1848

168. Markram, H. (2009). A Brain in a Supercomputer. Talk at TED Global 2009. Available at: https://www.ted.com/talks/henry_markram_supercomputing_the_brain_s_secrets

169. Markram, H., Muller, E., Ramaswamy, S., ... & Schurmann, F. (2015). Reconstruction and simulation of neocortical microcircuitry. Cell, 163(2), 456-492. https://doi.org/10.1016/j.cell.2015.09.029

170. Marsh, B., Navas-Zuloaga, M. G., Rosen, B. Q., Sokolov, Y., Delanois, J. E., Gonzalez, O. C., Krishnan, G. P., Halgren, E., & Bazhenov, M. (2024). Emergent effects of synaptic connectivity on the dynamics of global and local slow waves in a large-scale thalamocortical network model of the human brain. PLOS Computational Biology, 20(7), e1012245. https://doi.org/10.1371/journal.pcbi.1012245

171. McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5(4), 115-133. https://doi.org/10.1007/BF02478259

172. Meirovitch, Y., et al. (2024). SmartEM: Machine-learning guided electron microscopy. bioRxiv. https://doi.org/10.1101/2023.10.05.561103

173. Meneghetti, N., Rimehaug, A. E., Einevoll, G. T., Mazzoni, A., & Ness, T. V. (2024). Kernel-based LFP estimation in detailed large-scale spiking network model of mouse visual cortex. bioRxiv. https://doi.org/10.1101/2024.11.29.626029

174. Micikevicius, P., Narang, S., Alben, J., Diamos, G., Elsen, E., Garcia, D., ... & Wu, H. (2018). Mixed precision training. In Proceedings of the International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1710.03740

175. Motone, K., Kontogiorgos-Heintz, D., Wee, J., Kurihara, K., Yang, S., ... & Nivala, J. (2024). A strategy to load, rethread and read protein sequences through a nanopore. Nature, 633, 662-669. https://doi.org/10.1038/s41586-024-07935-7

176. Motta, A., Berning, M., Boergens, K. M., ... & Helmstaedter, M. (2019). Dense connectomic reconstruction in layer 4 of the somatosensory cortex. Science, 366(6469), eaay3134. https://doi.org/10.1126/science.aay3134

177. Mouse Genome Sequencing Consortium. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature, 420(6915), 520-562. https://doi.org/10.1038/nature01262

178. Mu, S., Dorkenwald, S., Li, P. H., Januszewski, M., Berger, D. R., Maitin-Shepard, J., ... & Jain, V. (2023). Multi-layered maps of neuropil with segmentation-guided contrastive learning. Nature Methods, 20(12), 2011-2020. https://doi.org/10.1038/s41592-023-02059-8

179. Murthy, M., Dorkenwald, S., Schlegel, P., Matsliah, A., Sterling, A. R., ... & Lin, A. (2024). The whole-brain connectome of Drosophila melanogaster. Nature, 634(8028), 110-123. https://doi.org/10.1038/s41586-024-07557-1

180. Musk, E. [@elaboratkingfish]. (2024, September 4). Doubling the Colossus supercluster to 200K H100 GPUs. X (formerly Twitter). https://x.com/elonmusk/status/1831089398592344424

181. National Institutes of Health. (2021). BRAIN Initiative: Multiday workshop series on brain connectivity. NIH Report. https://braininitiative.nih.gov/

182. Nguyen, J. P., Shipley, F. B., Linder, A. N., Plummer, G. S., Liu, M., Setru, S. U., Shaevitz, J. W., & Leifer, A. M. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving Caenorhabditis elegans. Proceedings of the National Academy of Sciences, 113(8), E1074-E1081. https://doi.org/10.1073/pnas.1507110112

183. Nowinski, W. L. (2024). Data challenges for whole-brain connectomics: Storage, compression, and analysis at the zettabyte scale. Frontiers in Neuroinformatics, 18, 1386001.

184. NVIDIA Corporation. (2023). NVIDIA H100 Tensor Core GPU architecture whitepaper. NVIDIA. https://resources.nvidia.com/en-us-tensor-core

185. Oláh, V. J., Pedersen, N. P., & Rowan, M. J. M. (2022). Ultrafast simulation of large-scale neocortical microcircuitry with biophysically realistic neurons. eLife, 11, e79535. https://doi.org/10.7554/eLife.79535

186. OpenAI. (2024). GPT-4o and reasoning capabilities. OpenAI Technical Report. https://openai.com/

187. OpenAI. (2025). Competitive programming with OpenAI o3. arXiv preprint arXiv:2502.06807. https://arxiv.org/abs/2502.06807

188. Orchard, G., Frady, E. P., Rubin, D. B. D., Sanborn, S., Shrestha, S. B., Sommer, F. T., et al. (2021). Efficient neuromorphic signal processing with Loihi 2. 2021 IEEE Workshop on Signal Processing Systems (SiPS). https://ieeexplore.ieee.org/document/9605018

189. Ou, J., Duh, Y.-S., Rommelfanger, N. J., Keck, C. H. C., Jiang, S., Brinson, K., ... & Hong, G. (2024). Achieving optical transparency in live animals with absorbing molecules. Science, 385(6713), eadm6869. https://doi.org/10.1126/science.adm6869

190. Pan, Y. A., Livet, J., Sanes, J. R., Lichtman, J. W., & Schier, A. F. (2011). Multicolor Brainbow Imaging in Zebrafish. Cold Spring Harbor Protocols, 2011(1), pdb.prot5546. https://doi.org/10.1101/pdb.prot5546

191. Panluminate. (2022). Panluminate. https://www.panluminate.com/

192. Park, S. Y., Sheridan, A., An, B., Jarvis, E., Lyudchik, J., ... & Boyden, E. S. (2025). Combinatorial protein barcodes enable

self-correcting neuron tracing with nanoscale molecular context. bioRxiv, 2025.09.26.678648. https://doi.org/10.1101/2025.09.26.678648

193. Patel, D. (2024). 100,000 H100 clusters: Power, network topology, Ethernet vs InfiniBand, reliability. SemiAnalysis Newsletter. https://newsletter.semianalysis.com/p/100000-h100-clusters-power-network

194. Peddie, C. J., & Collinson, L. M. (2014). Exploring the third dimension: Volume electron microscopy comes of age. Micron, 61, 9-19. https://doi.org/10.1016/j.micron.2014.01.009

195. Peddie, C. J., Liv, N., Hoogenboom, J. P., & Collinson, L. M. (2022). Volume electron microscopy of cells and tissues: imaging biological ultrastructure in three dimensions. Emerging Topics in Life Sciences, 6(5), 575-590. https://doi.org/10.1042/ETLS20220039

196. Peikon, I. D., Kebschull, J. M., Vagin, V. V., Ravens, D. I., Sun, Y.-C., ... & Zador, A. M. (2017). Using high-throughput barcode sequencing to efficiently map connectomes. Nucleic Acids Research, 45(12), e115. https://doi.org/10.1093/nar/gkx462

197. Penczek, P. A. (2010). Fundamentals of three-dimensional reconstruction from projections. Methods in Enzymology, 482, 1-33.

198. Penczek, P. A. (2011). Resolution measures in molecular electron microscopy. Methods in Enzymology, 482, 73-100.

199. Perrault, R., & Clark, J. (2024). The 2024 AI Index Report. Stanford Institute for Human-Centered AI. https://aiindex.stanford.edu/report/

200. Petousakis, K.-E., Apostolopoulou, A. A., & Poirazi, P. (2022). The impact of Hodgkin-Huxley models on dendritic research. The Journal of Physiology, 601(15), 3091-3102. https://doi.org/10.1113/JP282756

201. Phelps, J. S., Hildebrand, D. G. C., Graham, B. J., Kuan, A. T., Thomas, L. A., Nguyen, T. M., ... & Lee, W.-C. A. (2021). Reconstruction of motor control circuits in adult Drosophila using automated transmission electron microscopy. Cell, 184(3), 759-774.e18. https://doi.org/10.1016/j.cell.2020.12.013

202. Poggio, T., & Reichardt, W. (1973). Considerations on models of movement detection. Kybernetik, 13(4), 223-227. https://doi.org/10.1007/BF00274887

203. Polikov, V. S., Tresco, P. A., & Reichert, W. M. (2005). Response of brain tissue to chronically implanted neural electrodes. Journal of Neuroscience Methods, 148(1), 1-18. https://doi.org/10.1016/j.jneumeth.2005.08.015

204. Popovych, S., Macrina, T., Kemnitz, N., ... & Seung, H. S. (2022). PetaByte-scale multi-beam electron microscopy image alignment for connectomics. bioRxiv. https://doi.org/10.1101/2022.05.16.492103

205. Pospisil, D. A., Aragon, M. J., Dorkenwald, S., Matsliah, A., Sterling, A. R., Schlegel, P., ... & Pillow, J. W. (2024). The role of invisible factors in neural simulation accuracy. Nature Neuroscience, 27, 1320-1330.

206. Potjans, T. C., & Diesmann, M. (2014). The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. Cerebral Cortex, 24(3), 785-806. https://doi.org/10.1093/cercor/bhs358

207. Prevedel, R., Yoon, Y.-G., Hoffmann, M., Pak, N., Wetzstein, G., Kato, S., ... & Vaziri, A. (2014). Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. Nature Methods, 11(7), 727-730. https://doi.org/10.1038/nmeth.2964

208. Princeton Neuroscience Institute. (2023). Princeton connectomics: Large-scale serial EM imaging project. Princeton University. https://pni.princeton.edu/

209. Quake, S. R. (2024). The era of complete genomics. Nature, 627, 504-510.

210. Rae, J. W., Borgeaud, S., Cai, T., ... & Irving, G. (2021). Scaling language models: Methods, analysis, and insights from training Gopher. arXiv preprint arXiv:2112.11446. https://arxiv.org/abs/2112.11446

211. Randi, F., Sharma, A. K., Dvali, S., & Leifer, A. M. (2023). Neural signal propagation atlas of Caenorhabditis elegans. Nature, 623(7986), 406-414. https://doi.org/10.1038/s41586-023-06683-4

212. Rangoli, S. (2024). Neuromorphic computing: Long-term operational perspectives. Nature Electronics, 7, 680-690.

213. Rein, D., Hou, B. L., Stickland, A. C., Petty, J., Pang, R. Y., Dirani, J., ... & Bowman, S. R. (2023). GPQA: A graduate-level Google-proof Q&A benchmark. arXiv preprint arXiv:2311.12022. https://arxiv.org/abs/2311.12022

214. Rein, K., Zöckler, M., Mader, M. T., Grübel, C., & Heisenberg, M. (2002). The Drosophila standard brain. Current Biology, 12(3), 227-231. https://doi.org/10.1016/S0960-9822(02)00656-5

215. Rimehaug, A. E., Dale, A. M., Arkhipov, A., & Einevoll, G. T. (2024). Uncovering population contributions to the extracellular potential in the mouse visual system using Laminar Population Analysis. bioRxiv. https://doi.org/10.1101/2024.01.15.575805

216. Rimehaug, A. E., Stasik, A. J., Hagen, E., Billeh, Y. N., Siegle, J. H., Dai, K., ... & Arkhipov, A. (2023). Uncovering circuit mechanisms of current sinks and sources with biophysical simulations of primary visual cortex. eLife, 12, e87169. https://doi.org/10.7554/eLife.87169

217. Ripoll-Sanchez, L., Watteyne, J., Sun, H., Fernandez, R., Taylor, S. R., ... & Schafer, W. R. (2023). The neuropeptidergic connectome of C. elegans. Neuron, 111(22), 3570-3589.e5. https://doi.org/10.1016/j.neuron.2023.09.010

218. Rodriques, S. G. (2009). Barcoding-based approaches for mapping brain connectivity. PhD Thesis, Harvard University.

219. Rodriques, S. G. (2022). Why is progress in biology so slow? Sam Rodriques. https://www.sam-rodriques.com/post/why-is-progress-in-biology-so-slow

220. Romani, A., Antonietti, A., Bella, D., Budd, J., Giacalone, E., Kurban, K., ... & Markram, H. (2024). Community-based reconstruction and simulation of a full-scale model of the rat hippocampus CA1 region. PLOS Biology, 22(11), e3002861. https://doi.org/10.1371/journal.pbio.3002861

221. Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747. https://arxiv.org/abs/1609.04747

222. Sample, I. (2014, October 8). Henry Markram's Blue Brain Project: The quest to simulate the brain. The Guardian.

223. Sandberg, A., & Bostrom, N. (2008). Whole brain emulation: A roadmap (Technical Report #2008-3). Future of Humanity Institute, University of Oxford. https://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf

224. Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, 74(12), 5463-5467. https://doi.org/10.1073/pnas.74.12.5463

225. Sarma, G. P., Lee, C. W., Portegys, T., Ghayoomie, V., Jacobs, T., Alicea, B., ... & Larson, S. D. (2018). OpenWorm: overview and recent advances in integrative biological simulation of Caenorhabditis elegans. Philosophical Transactions of the Royal Society B, 373(1758), 20170382. https://doi.org/10.1098/rstb.2017.0382

226. Sato, H. (2004). DRAM memory scaling: A historical perspective. IEEE Solid-State Circuits Magazine, 6(4), 22-30.

227. Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. Journal of Statistical Physics, 65(3-4), 579-616. https://doi.org/10.1007/BF01053745

228. Scheffer, L. K., & Meinertzhagen, I. A. (2021). A connectome is not enough - what is still needed to understand the brain of Drosophila? Journal of Experimental Biology, 224(13), jeb242740. https://doi.org/10.1242/jeb.242740

229. Scheffer, L. K., Xu, C. S., Januszewski, M., Lu, Z., Takemura, S.-y., ... & Plaza, S. M. (2020). A connectome and analysis of the adult Drosophila central brain. eLife, 9, e57443. https://doi.org/10.7554/eLife.57443

230. Schlegel, P., Yin, Y., Bates, A. S., ... & Jefferis, G. S. X. E. (2024). Whole-brain annotation and multi-connectome cell typing of Drosophila. Nature, 634(8032), 139-152. https://doi.org/10.1038/s41586-024-07686-5

231. Schrimpf, M., Kubilius, J., Hong, H., Majaj, N. J., Rajalingham, R., Issa, E. B., Kar, K., Bashivan, P., Prescott-Roy, J., Geiger, F., Schmidt, K., Yamins, D. L. K., & DiCarlo, J. J. (2020). Brain-Score: Which artificial neural network for object recognition is most brain-like? bioRxiv. https://doi.org/10.1101/407007

232. Schrödel, T., Prevedel, R., Aumayr, K., Zimmer, M., & Vaziri, A. (2013). Brain-wide 3D imaging of neuronal activity in Caenorhabditis elegans with sculpted light. Nature Methods, 10(10), 1013-1020. https://doi.org/10.1038/nmeth.2637

233. Schueder, F., Rivera-Molina, F., Su, M., Marin, Z., Kidd, P., Rothman, J. E., ... & Bewersdorf, J. (2023). DNA-based FLASH-PAINT super-resolution microscopy. Nature Methods, 20, 1310-1320.

234. Schuz, A., & Palm, G. (1989). Density of neurons and synapses in the cerebral cortex of the mouse. Journal of Comparative Neurology, 286(4), 442-455. https://doi.org/10.1002/cne.902860404

235. Schwartz, S. I. (1998). Atomic Audit: The Costs and Consequences of U.S. Nuclear Weapons Since 1940. Brookings Institution Press. https://www.brookings.edu/books/atomic-audit/

236. Secure I.T. Environments Ltd. (n.d.). What does it cost to build a data centre? https://siteltd.co.uk/blog/what-does-it-cost-to-build-a-data-centre/

237. Seo, J., et al. (2022). PICASSO allows ultra-multiplexed fluorescence imaging of spatially overlapping proteins without

reference spectra measurements. Nature Communications, 13, 2475. https://doi.org/10.1038/s41467-022-30168-z

238. Sevilla, J., Heim, L., Ho, A., Besiroglu, T., Hobbhahn, M., & Villalobos, P. (2024). Can AI scaling continue through 2030? Epoch AI. https://epoch.ai/blog/can-ai-scaling-continue-through-2030

239. Shaib, A. H., Chouaib, A. A., Chowdhury, R., Altendorf, J., Mihaylov, D., ... & Rizzoli, S. O. (2024). One-step nanoscale expansion microscopy reveals individual protein shapes. Nature Biotechnology, 42, 1320-1330. https://doi.org/10.1038/s41587-024-02431-9

240. Shalizi, C. R., & Crutchfield, J. P. (1999). Computational mechanics: Pattern and prediction, structure and simplicity. arXiv preprint cond-mat/9907176. https://arxiv.org/abs/cond-mat/9907176

241. Shapson-Coe, A., Januszewski, M., Berger, D. R., ... & Lichtman, J. W. (2024). A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. Science, 384(6695), eadk4858. https://doi.org/10.1126/science.adk4858

242. Sheng, M., & Kim, E. (2011). The postsynaptic organization of synapses. Cold Spring Harbor Perspectives in Biology, 3(12), a005587. https://doi.org/10.1101/cshperspect.a005587

243. Shin, T. W., Wang, H., Zhang, C., Ou, J., Chen, Z., Jiang, S., ... & Boyden, E. S. (2024). Dense, continuous membrane labeling and expansion microscopy visualization of ultrastructure in tissues. bioRxiv, 2024.03.07.583776. https://doi.org/10.1101/2024.03.07.583776

244. Shiu, P. K., Sterne, G. R., Spiller, N., Franconville, R., Sandoval, A., ... & Scott, K. (2024). A Drosophila computational brain model reveals sensorimotor processing. Nature, 634(8033), 210-219. https://doi.org/10.1038/s41586-024-07763-9

245. Siegle, J. H., Jia, X., Durand, S., Gale, S., Bennett, C., Graddis, N., ... & Koch, C. (2021). Survey of spiking in the mouse visual system reveals functional hierarchy. Nature, 592, 86-92. https://doi.org/10.1038/s41586-020-03171-x

246. Silver, D., Huang, A., Maddison, C. J., ... & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484-489. https://doi.org/10.1038/nature16961

247. Simeon, Q., Venâncio, L., Skuhersky, M. A., Nayebi, A., Boyden, E. S., & Yang, G. R. (2024). Scaling properties for artificial neural network models of a small nervous system. bioRxiv. https://doi.org/10.1101/2024.02.13.580186

248. Srivastava, A., Rastogi, A., Rao, A., Shoeb, A. A. M., Abid, A., Fisch, A., ... & Wu, J. (2022). Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models. arXiv preprint arXiv:2206.04615. https://arxiv.org/abs/2206.04615

249. Staffler, B., Berning, M., Boergens, K. M., ... & Helmstaedter, M. (2017). SynEM: Automated synapse detection for connectomics. eLife, 6, e26414. https://doi.org/10.7554/eLife.26414

250. Stanojevic, A., Woźniak, S., Bellec, G., Cherubini, G., Pantazi, A., & Gerstner, W. (2024). High-performance deep spiking neural networks with 0.3 spikes per neuron. Nature Communications, 15, 6793. https://doi.org/10.1038/s41467-024-51110-5

251. Sterling, P., & Laughlin, S. B. (2015). Principles of neural design. MIT Press. https://mitpress.mit.edu/9780262327329/principles-of-neural-design

252. Stevenson, I. H., & Kording, K. P. (2011). How advances in neural recording affect data analysis. Nature Neuroscience, 14(2), 139-142. https://doi.org/10.1038/nn.2731

253. Stiefel, K. M., & Brooks, D. S. (2019). Why is there no successful whole brain simulation (yet)? Biological Cybernetics, 113(1-2), 155-167. https://doi.org/10.1007/s00422-018-0787-1

254. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science, 364(6437), 255. https://doi.org/10.1126/science.aav7893

255. Strotton, M., Hosogane, T., di Michiel, M., Moch, H., Varga, Z., & Bodenmiller, B. (2023). Multielement Z-tag imaging by X-ray fluorescence microscopy. Nature Methods, 20, 1310-1322. https://doi.org/10.1038/s41592-023-01977-x

256. Sui, X., Wang, X., et al. (2024). Deep-STARmap: Scalable spatial transcriptomics. bioRxiv. https://doi.org/10.1101/2024.08.05.606553

257. Svara, F., Förster, D., Kubo, F., Januszewski, M., dal Maschio, M., ... & Baier, H. (2022). Dense connectomics of larval zebrafish with 2,589 axons traced. Current Biology, 32(18), 3924-3934.

258. Swaminathan, J., Boulgakov, A. A., Hernandez, E. T., Bardo, A. M., Bachman, J. L., ... & Marcotte, E. M. (2018). Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. Nature Biotechnology, 36(11), 1076-1082. https://doi.org/10.1038/nbt.4278

259. Szigeti, B., Gleeson, P., Vella, M., Khayrulin, S., Palyanov, A., ... & Larson, S. (2014). OpenWorm: An open-science approach to modeling Caenorhabditis elegans. Frontiers in Computational Neuroscience, 8, 137. https://doi.org/10.3389/fncom.2014.00137

260. Takahashi, T., Zhang, H., Agetsuma, M., Nabekura, J., Otomo, K., Okamura, Y., & Nemoto, T. (2024). Two-photon microscopy for cortical tissue access. Nature Methods, 21, 1000-1010.

261. Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L.-S. Young (Eds.), Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898 (pp. 366-381). Springer. https://doi.org/10.1007/BFb0091924

262. Tang, J., LeBel, A., Jain, S., & Huth, A. G. (2023). Semantic reconstruction of continuous language from non-invasive brain recordings. Nature Neuroscience, 26(5), 858-866. https://doi.org/10.1038/s41593-023-01304-9

263. Tapia, J.C., Kasthuri, N., Hayworth, K.J., Schalek, R., Lichtman, J.W., Smith, S.J., & Bhatt, J. (2012). High-contrast en bloc staining of neuronal tissue for field emission scanning electron microscopy. Nature Protocols, 7(2), 193-206. https://doi.org/10.1038/nprot.2011.439

264. Tavakoli, M. R., Lyudchik, J., Januszewski, M., Vistunou, V., Agudelo Duenas, N., ... & Danzl, J. G. (2024). Light-microscopy-based connectomic reconstruction of mammalian brain tissue. Nature, 634, 580-590. https://doi.org/10.1038/s41586-024-08968-9

265. Tegethoff, T., & Briggman, K. L. (2024). Quantitative evaluation of embedding resins for volume electron microscopy. Frontiers in Neuroscience, 18, 1286991. https://doi.org/10.3389/fnins.2024.1286991

266. Tennyson, Alfred, Lord. (1842). Ulysses. Poetry Foundation. https://www.poetryfoundation.org/poems/45392/ulysses

267. Thomas, D. (2024). The computing costs of large-scale brain simulation: A 1.5 exaflop analysis. In Proceedings of the International Conference on High Performance Computing.

268. Tillberg, P. W., Chen, F., Piatkevich, K. D., Zhao, Y., Yu, C. C. J., ... & Boyden, E. S. (2016). Protein-retention expansion microscopy of cells and tissues. Nature Biotechnology, 34(9), 987-992. https://doi.org/10.1038/nbt.3625

269. Tong, A., & Martina, M. (2024). U.S. government commission pushes Manhattan Project-style AI initiative. Reuters.

https://www.reuters.com/world/us/us-government-commission-pushes-manhattan-project-style-ai-initiative-2024-11-19/

270. Tononi, G. (2004). An information integration theory of consciousness. BMC Neuroscience, 5, 42. https://doi.org/10.1186/1471-2202-5-42

271. Trinh, T. H., Wu, Y., Le, Q. V., He, H., & Luong, T. (2024). Solving olympiad geometry without human demonstrations. Nature, 625(7995), 476-482. https://doi.org/10.1038/s41586-023-06747-5

272. Tripp, B. C., & Grueber, W. B. (2011). The Human Genome Project: Timeline and legacy. Annual Review of Genomics and Human Genetics, 12, 83-110.

273. Urai, A. E., Doiron, B., Leifer, A. M., & Churchland, A. K. (2022). Large-scale neural recordings call for new insights to link brain and behavior. Nature Neuroscience, 25(1), 11-19. https://doi.org/10.1038/s41593-021-00980-9

274. Vahdat, A. (2024). Jupiter data center networking now scales to 13 petabits per second. Google Cloud Blog. https://cloud.google.com/blog/products/networking/

275. Valentin, J., Rodenburg, J. M., Juschkin, L., et al. (2023). Electronic ptychography for high-throughput microscopy. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1055, 168665. https://doi.org/10.1016/j.nima.2023.168665

276. Varongchayakul, N., Huttner, D., Grinstaff, M. W., & Meller, A. (2018). Sensing native protein solution structures using a solid-state nanopore. Scientific Reports, 8, 1017. https://doi.org/10.1038/s41598-018-19332-y

277. Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., & Chklovskii, D. B. (2011). Structural properties of the Caenorhabditis elegans neuronal network. PLoS Computational Biology, 7(2), e1001066. https://doi.org/10.1371/journal.pcbi.1001066

278. Villette, V., Chavarha, M., Dimov, I. K., Bradley, J., Pradhan, L., ... & Lin, M. Z. (2019). Ultrafast two-photon imaging of a high-gain voltage indicator. Cell, 179(7), 1590-1608. https://doi.org/10.1016/j.cell.2019.11.004

279. Viswanathan, S., Williams, M. E., Bloss, E. B., Stasevich, T. J., Speer, C. M., ... & Looger, L. L. (2015). High-performance probes for light and electron microscopy. Nature Methods, 12(6), 568-576. https://doi.org/10.1038/nmeth.3365

280. Vladimirov, N., Mu, Y., Kawashima, T., Bennett, D. V., Yang, C.-T., Looger, L. L., ... & Ahrens, M. B. (2014). Light-sheet functional imaging in fictively behaving zebrafish. Nature Methods, 11(8), 883-884. https://doi.org/10.1038/nmeth.3040

281. Vogelstein, J. T., Watson, B. O., Packer, A. M., Yuste, R., Jedynak, B., & Paninski, L. (2009). Spike inference from calcium imaging using sequential Monte Carlo methods. Biophysical Journal, 97(2), 636-655. https://doi.org/10.1016/j.bpj.2008.08.005

282. Voleti, V., Patel, K. B., Li, W., Perez Campos, C., Bharadwaj, S., Yu, H., ... & Hillman, E. M. C. (2019). Real-time volumetric microscopy of in vivo dynamics and large-scale samples with SCAPE 2.0. Nature Methods, 16(10), 1054-1062. https://doi.org/10.1038/s41592-019-0579-4

283. Vorwerk, J., Clerc, M., Burger, M., & Wolters, C. H. (2024). Comparison of boundary element and finite element approaches to the EEG forward problem. Frontiers in Human Neuroscience, 18, 1335212. https://doi.org/10.3389/fnhum.2024.1335212

284. Wang, C., Vidal, B., Sural, S., Loer, C., Aguilar, G. R., Merritt, D. M., Toker, I. A., Vogt, M. C., Cros, C. C., & Hobert, O. (2024). A neurotransmitter atlas of C. elegans males and hermaphrodites. eLife, 13, RP95402. https://doi.org/10.7554/eLife.95402.3

285. Wang, Z., Zhang, J., Symvoulidis, P., Guo, W., Zhang, L., Wilson, M. A., & Boyden, E. S. (2023). Imaging the voltage of neurons distributed across entire brains of larval zebrafish. bioRxiv, 2023.12.15.571964. https://doi.org/10.1101/2023.12.15.571964

286. Wetterstrand, K. A. (2022). DNA sequencing costs: Data from the NHGRI Large-Scale Genome Sequencing Program. National Human Genome Research Institute. https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

287. White, J. G., Southgate, E., Thomson, J. N., & Brenner, S. (1986). The structure of the nervous system of the nematode Caenorhabditis elegans. Philosophical Transactions of the Royal Society B, 314(1165), 1-340. https://doi.org/10.1098/rstb.1986.0056

288. Witvliet, D., Mulcahy, B., Mitchell, J. K., Meirovitch, Y., Berger, D. R., ... & Zhen, M. (2021). Connectomes across development reveal principles of brain maturation. Nature, 596(7871), 257-261. https://doi.org/10.1038/s41586-021-03778-8

289. Yamaura, H., Igarashi, J., & Yamazaki, T. (2020). Neural simulation on exascale compute clusters. In Proceedings of SC20: International Conference for High Performance Computing.

290. Yamazaki, T., Igarashi, J., Makino, J., & Ebisuzaki, T. (2021). Real-time simulation of a cat-scale artificial cerebellum on PEZY-SC processors. International Journal of High Performance Computing Applications, 35(2), 204-212. https://doi.org/10.1177/1094342021998540

291. Yampolskiy, R. V. (2015). The space of possible mind designs. In Artificial General Intelligence (pp. 218-227). Springer. https://doi.org/10.1007/978-3-319-21365-1_23

292. Yang, J., & Annaert, W. (2024). The synaptic cleft: Molecular architecture and function. Nature Reviews Neuroscience, 25, 120-135.

293. Yin, W., Brittain, D., Borseth, J., Scott, M. E., Williams, D., Perkins, J., ... & da Costa, N. M. (2020). A petascale automated imaging pipeline for mapping neuronal circuits with high-throughput transmission electron microscopy. Nature Communications, 11(1), 4949. https://doi.org/10.1038/s41467-020-18659-3

294. Yoon, Y.-G., Dai, P., Wohlwend, J., Chang, J.-B., Marblestone, A. H., & Boyden, E. S. (2017). Feasibility of 3D reconstruction of neural morphology using expansion microscopy and barcode-guided agglomeration. Frontiers in Computational Neuroscience, 11, 97. https://doi.org/10.3389/fncom.2017.00097

295. Zador, A. M., Escola, S., Richards, B., Ölveczky, B., Bengio, Y., Boahen, K., Botvinick, M., Chklovskii, D., ... & Churchland, A. (2023). Catalyzing next-generation Artificial Intelligence through NeuroAI. Nature Communications, 14, 1597. https://doi.org/10.1038/s41467-023-37180-x

296. Zanichelli, N., Schons, M., Freeman, I., Shiu, P., & Arkhipov, A. (2025). State of Brain Emulation Report 2025. arXiv preprint arXiv:2510.15745. https://doi.org/10.48550/arXiv.2510.15745

297. Zeidler, D., et al. (2015). Multi-beam scanning electron microscopy for high-throughput imaging. Microscopy and Microanalysis, 21(S3), 1249-1250. https://doi.org/10.1017/S1431927615007060

298. Zhang, Y., He, G., Ma, L., Liu, X., Hjorth, J. J. J., Kozlov, A., He, Y., Zhang, S., Kotaleski, J. H., Tian, Y., Grillner, S., Du, K., & Huang, T. (2023). A GPU-based computational framework that bridges neuron simulation and artificial intelligence. Nature Communications, 14, 5798. https://doi.org/10.1038/s41467-023-41553-7

299. Zhao, M., Huang, T., Zheng, Y., Wang, W., Li, M., Yang, K., ... & Shan, S. (2024). An integrative data-driven model simulating C. elegans brain, body and environment interactions. Nature Computational Science, 4(12), 978-990. https://doi.org/10.1038/s43588-024-00738-w

300. Zheng, Y., et al. (2024). Peptide sequencing via reverse translation of peptides into DNA. bioRxiv. https://www.biorxiv.org/content/10.1101/2024.05.31.596913v1

301. Zheng, Z., Lauritzen, J. S., Perlman, E., Robinson, C. G., Nichols, M., ... & Bock, D. D. (2018). A complete electron microscopy volume of the brain of adult Drosophila melanogaster. Cell, 174(3), 730-743. https://doi.org/10.1016/j.cell.2018.06.019

AI Use Acknowledgement: Portions of this thesis were drafted, edited, and fact-checked with the assistance of Claude (Anthropic, 2025), and GPT5 (OpenAI, 2025), both large language models.

Made on Earth by humans, with AIs.