https://doi.org/10.1038/s41583-025-00954-x

The emergence of NeuroAI: bridging neuroscience and artificial intelligence

Sadra Sadeh & Claudia Clopath

Neuroscience has inspired artificial intelligence (AI) for decades but, in recent years, AI tools have begun to revolutionize neuroscience research. The emerging field of NeuroAI has the potential to transform large-scale neural modelling and data-driven neuroscience discovery. The field must balance exploiting AI's power while maintaining interpretability and biological insight.

Early artificial intelligence (AI) models in the form of neural networks drew directly from neurobiological inspiration: perceptron mimicked sensory neurons, Hopfield networks modelled associative memory, and convolutional neural networks borrowed the columnar organization of the visual cortex. The principles of recently emerged AI models have been inspired by neurobiology too: researchers developed reinforcement learning to explain reward-modulated learning in the brain, deep neural networks imitated the multilayer organization of neuronal networks, and attention mechanisms in the brain motivated transformers. These AI models succeeded because they abstracted key principles from neural computation.

Emergence of NeuroAI

NeuroAI has historically referred to this direction of interaction between neuroscience and AI: how principles of neural architecture and function can inspire AI models. However, the dynamics of the neuroscience–AI relationship seems to have been reversed recently, thanks to the rapid progress in the development of powerful machine learning (ML) tools, as exemplified in the recent achievements of large language models (LLMs). The potential yields of applying similar AI models to large-scale and complex datasets in neuroscience are currently so huge that NeuroAI seems to be in its AI-exploitation phase. In this Comment, we focus on this recent direction of NeuroAI, namely AI exploitation in neuroscience, and outline its key promises and challenges, noting that the balance between the two directions NeuroAI takes is key for its long-term success as a field.

Promises of NeuroAI: AI for neuroscience

Advances in ML and AI models can revolutionize neuroscience by contributing to large-scale data analysis, building foundation models of neural activity, and helping to generate and test hypotheses.

Data organization and data analysis. The most immediate and transformative promise of NeuroAI lies in tackling neuroscience's big data challenge. Modern neuroscience generates terabytes and petabytes of data daily across multiple scales, from dendritic activity to whole-brain imaging. This wealth of information could reveal unprecedented insights into brain function, yet it can also create as much confusion as clarity. The curse of dimensionality threatens to bury signals beneath noise, making traditional analytical approaches inadequate.

ML and AI tools can provide essential support for complex data organization and analysis. Much of the labour-intensive curation and preprocessing of large-scale neuroscience data can be automated. For instance, LLMs can be used to transform metadata annotation to curate complex datasets more efficiently and data analysis pipelines can be streamlined by agentic AI solutions. The analysis process itself has been revolutionized through modern ML and AI tools, such as in image and video processing[1] and cellular segmentation[2].

A central question in neuroscience is how neural structure gives rise to emergent neural activity and ultimately drives animal behaviour. NeuroAI can help with data processing across all these scales. Two prominent examples demonstrate this impact at the two ends: connectomics and behavioural analysis. The remarkable achievement of mapping the complete fruit fly connectome[3] would have been impossible using traditional manual neuron tracing methods. These automated AI tools[4] now enable scaling connectomics to larger networks like mouse and human brains — an endeavour nearly impossible just years ago. Similarly, recent ML and AI tools for automated behavioural analysis[5] have revolutionized the field, enabling quantification of animal behaviour with unprecedented speed and precision. With these advances, NeuroAI promises to help to tackle fundamental neuroscience challenges like linking precise neural connectivity and activity patterns to complex behavioural motifs.

Modelling and predicting neural activity. The potential of NeuroAI can, however, extend beyond data analysis to help with modelling and predicting neural activity as a bridge between structure and function. First, it can help in quantifying and assessing the content of neural representations. For instance, meaningful low-dimensional patterns can be extracted from complex datasets through dimensionality reduction in the latent space of AI models[6]. Second, NeuroAI can go beyond representational analysis, and offer foundation models of neural activity: models that are trained on large-scale datasets to capture complex, high-dimensional statistical relationships in the data. Advanced computational resources now enable training foundation models on datasets from specific brain regions, such as by stimulating the visual cortex with a large set of natural stimuli[7].

Traditional neuroscience modelling has used two main approaches: top-down models grounded in normative theories like Bayesian learning, and bottom-up models constrained by established neural circuit properties. NeuroAI introduces a third paradigm without explicit theoretical assumptions: data-driven foundation models. Learning statistical regularities directly from massive neural datasets offers the potential to discover emergent principles of brain function through data alone.

NeuroAI models that capture statistical relationships within neural data can then provide strong predictive capabilities for neural activity patterns. Recent NeuroAI tools have shown remarkable potential in predicting human brain function, including in decoding speech[8]. The next frontier involves developing more generic NeuroAI models that can predict multimodal stimuli and can maintain effectiveness for out-of-distribution predictions when encountering novel stimulus categories. These types of tools can be used to advance brain–machine interfaces for enhancing or restoring brain function[9]. They also hold notable promise for predicting brain dysfunction, including early detection of neurodegenerative diseases like Alzheimer disease and identifying neural activity patterns that precede epileptic seizures.

NeuroAI should not, however, remain limited to learning statistical relationships, but should also help in building mechanistic and causal models of neural activity. These models will incorporate detailed biological properties of neural circuits, including cellular characteristics (neuroanatomy and electrophysiology) and network properties (network connectivity and plasticity). Although ambitious, NeuroAI's ultimate contribution to understanding brain function and dysfunction depends on developing such biologically informed mechanistic models. This approach promises to tackle longstanding neuroscience questions such as how neural dynamics emerge from different architectures; which plasticity rules enable efficient learning across circuits; how the brain selectively stores and forgets information to avoid saturation and catastrophic forgetting; how sleep benefits brain function; and how the brain achieves efficient information processing with minimal energy compared to current AI models.

Hypothesis generation and hypothesis testing. It is tempting to envisage that NeuroAI may advance beyond data analysis and modelling to assist neuroscientists with something more fundamental: generating novel hypotheses for scientific discovery. This capability operates at multiple stages, from synthesizing existing knowledge and generating new hypotheses, to evaluating competing hypotheses and identifying the most promising candidates for experimental testing.

LLMs excel at integrating knowledge across the entire neuroscience literature, facilitating initial literature screening before hypothesis generation. They can also combine existing hypotheses in conceptual and innovative ways, suggesting novel research directions for investigators[10]. Moreover, NeuroAI tools can narrow the scope of potential neuroscience experiments by identifying optimal hypotheses or models to test. Given the vast space of possible experiments following hypothesis generation, AI tools can conduct 'thought experiments' to recommend the most efficient experimental approaches.

Challenges and future directions in NeuroAI

The rapid and powerful advances in NeuroAI demand consideration of emerging challenges. Interpretability remains the most pressing concern. Whereas AI models may accurately capture neural data patterns, understanding the underlying mechanisms proves difficult. This applies both to the mechanism of operation of AI models as well as the neural mechanisms generating the patterns captured. A model that perfectly predicts neural activity without providing mechanistic insight offers limited scientific value.

Equally concerning is the risk of over-reliance on computational tools at the expense of fundamental analytical skills or biological insights. As analysis pipelines grow increasingly complex, reproducibility may become an issue, cross-validation and assessing the statistical power may be more difficult, and the researchers may lose connection with their data. The field risks creating a generation of neuroscientists capable of executing sophisticated analyses but unable to critically evaluate their results.

Addressing these challenges requires a balanced approach to NeuroAI. Educating the next generation of NeuroAI researchers with both strong analytical skills and solid biological foundations would be crucial for the field's success. A balanced approach should also strive to make the neuroscience–AI interaction more bidirectional: while current NeuroAI focuses primarily on applying AI tools to neuroscience, the reverse direction can be equally transformative, leading to the development of biologically informed AI models.

Conclusions

The ultimate success of NeuroAI will be measured not by the sophistication of its computational tools, but by its ability to advance our understanding of the most complex object in the known universe: the brain. Achieving this goal requires maintaining the balance between respecting biological complexity and harnessing computational power, ensuring that as AI is used to understand the brain, researchers do not lose sight of what makes biological intelligence so remarkable in the first place.

Sadra Sadeh[1,2] & Claudia Clopath[3] ✉

[1]Centre for Developmental Neurobiology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [2]The Francis Crick Institute, London, UK. [3]Bioengineering Department, Imperial College London, London, UK.

✉e-mail: c.clopath@imperial.ac.uk

Published online: 07 August 2025

References

1. Ravi, N. et al. SAM 2: segment anything in images and videos. Preprint at https://arxiv.org/abs/2408.00714 (2024).
2. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
3. Schlegel, P. et al. Whole-brain annotation and multi-connectome cell typing of *Drosophila*. *Nature* **634**, 139–152 (2024).
4. Schmidt, M., Motta, A., Sievers, M. & Helmstaedter, M. RoboEM: automated 3D flight tracing for synaptic-resolution connectomics. *Nat. Methods* **21**, 908–913 (2024).
5. Vogt, N. Automated behavioral analysis. *Nat. Methods* **18**, 29 (2021).
6. Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360–368 (2023).
7. Wang, E. Y. et al. Foundation model of neural activity predicts response to new stimulus types. *Nature* **640**, 470–477 (2025).
8. Metzger, S. L. et al. A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* **620**, 1037–1046 (2023).
9. Wairagkar, M. et al. An instantaneous voice-synthesis neuroprosthesis. *Nature* https://doi.org/10.1038/s41586-025-09127-3 (2025).
10. Gottweis, J. et al. Towards an AI co-scientist. Preprint at https://arxiv.org/abs/2502.18864 (2025).

Acknowledgements

The authors thank A. Schaefer, P. Gleeson, F. Iacaruso and N. Skatchkovsky for reading an earlier version of this Comment and providing their insightful comments on it.

Competing interests

The authors declare no competing interests.