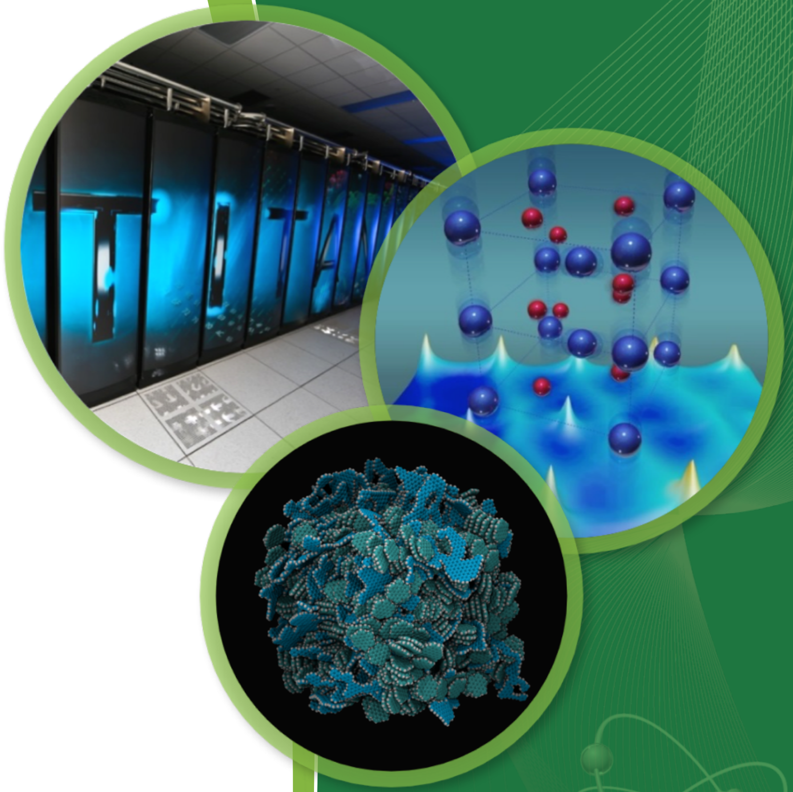
# Neuromorphic Computing: Past, Present, and Future

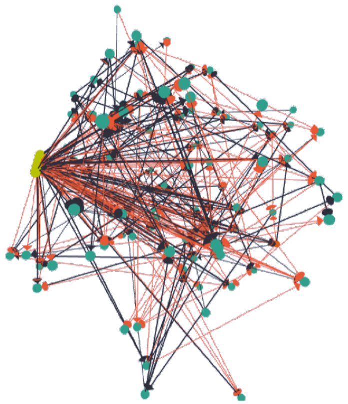
Catherine Schuman

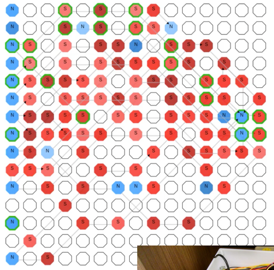
Liane Russell Fellow

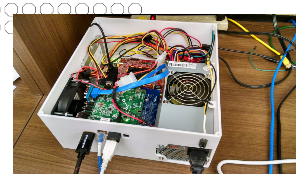
Oak Ridge National Laboratory

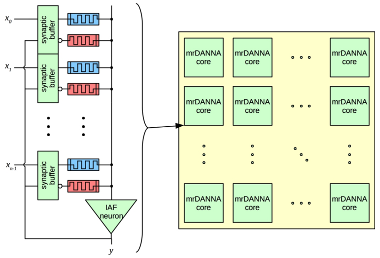


# Introduction



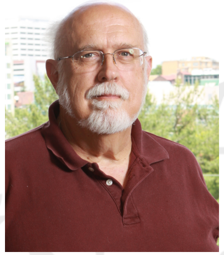




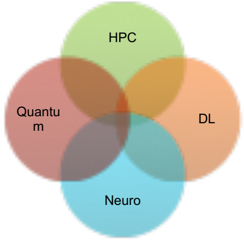


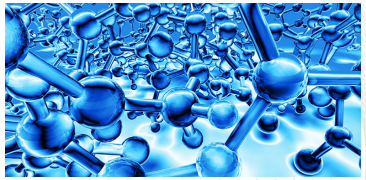

TENN LAB
NEUROMORPHIC
ARCHITECTURES. LEARNING. APPLICATIONS.

THE UNIVERSITY OF TENNESSEE KNOXVILLE



OAK RIDGE National Laboratory



OAK RIDGE National Laboratory

# Outline

- Historical Perspective
- Motivations of Neuromorphic Computing
- Key Questions:
  - Models
  - Algorithms
  - Hardware and devices
  - Applications
- What's Next?

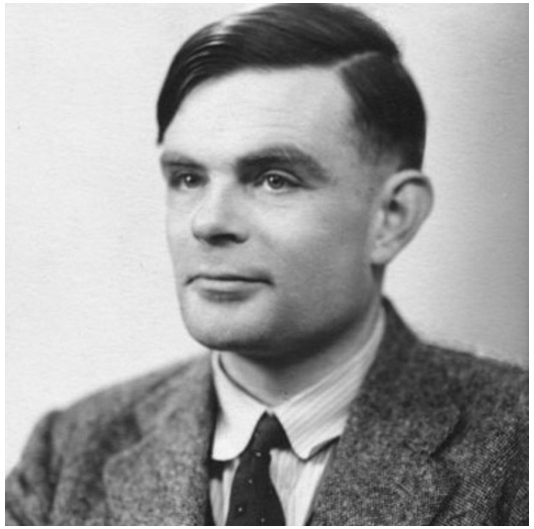
# Historical Perspective

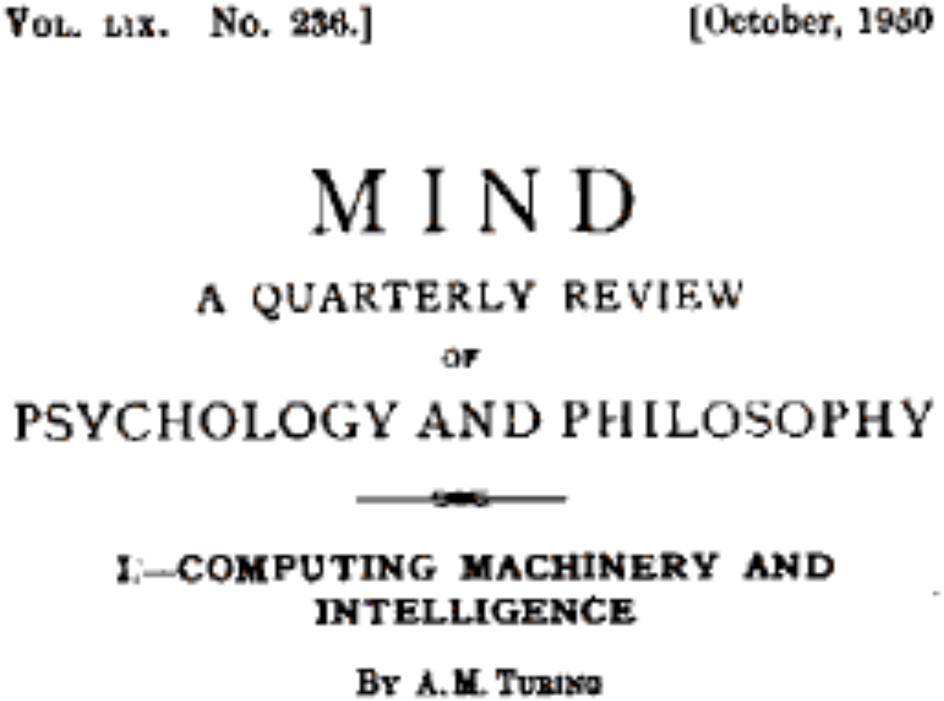
"The Analytical Engine has no pretensions whatever to originate any thing. It can do whatever we know how to order it to perform. It can follow analysis; but it has no power of anticipating any analytical relations or truths."
-- Ada Lovelace, in her notes on Charles Babbage's Analytical Engine article

Source: http://www.cs.yale.edu/homes/tap/Files/ada-lovelace-notes.html

# Historical Perspective

"Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?"
– Alan Turing, in "Computing Machinery and Intelligence"





Vol. lix. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

Image Source: https://www.biography.com/people/alan-turing-9512017

# Historical Perspective

"I will discuss the points of similarity between these two kinds of 'automata'…There are elements of dissimilarity…not only in rather obvious respects of size and speed but also in much deeper-lying areas: These involve principles of functioning and control, of over-all organization, etc."
– John von Neumann, in *The Computer and the Brain*

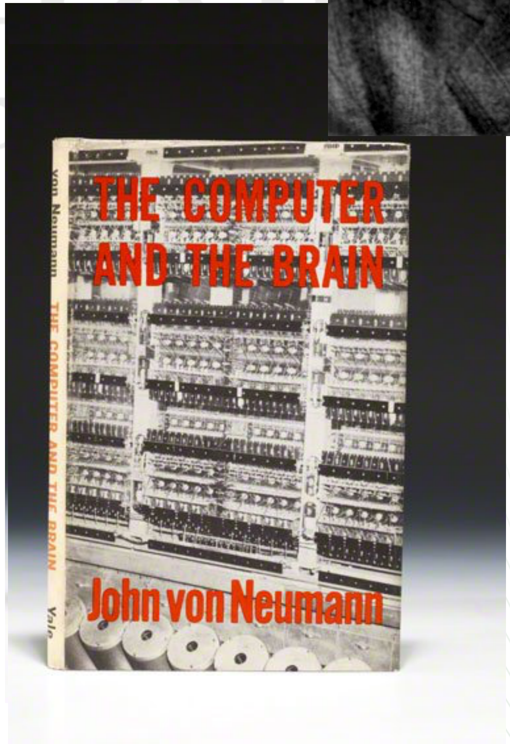



Image source: https://www.baumanrarebooks.com/rare-books/neumann-john-von/computer-and-the-brain/80575.aspx

# Neuromorphic Computing Definition Over Time

Analog electronic circuits for brain-like computation

Analog Circuits and Sensory Systems

Custom Analog, Digital, or Hybrid Circuits

Computation distributed across simple elements (neurons) with communication between elements (through synapses)

Architecture and/or the way that it performs computation is inspired by biological brains

**Carver Mead. "Neuromorphic electronic systems."** ***Proceedings of the IEEE*** **78.10 (1990): 1629-1636**

**Early-Mid 1990's – Work by Mead, Indiveri, Boahen, Douglas,…**

**Early 2000's**

**Late 2000's/Early 2010's**

# Carver Mead and Neuromorphic

“The nervous system uses, as its basic operation, a current that increases exponentially with voltage… What class of computations can be implemented efficiently using exponential functions as primitives? **Analog electronic circuits are an ideal way to explore this question.**

The fact that we can build devices that implement the same basic operations as those the nervous system uses leads to the inevitable conclusion that we should be able to build entire systems based on the organizing principles used by the nervous system. **I will refer to these systems generically as *neuromorphic systems*.”**

– Carver Mead, in ”Neuromorphic Electronic Systems,” Proceedings of the IEEE, October 1990.

Source: Mead, Carver. "Neuromorphic electronic systems." *Proceedings of the IEEE*78.10 (1990): 1629-1636.

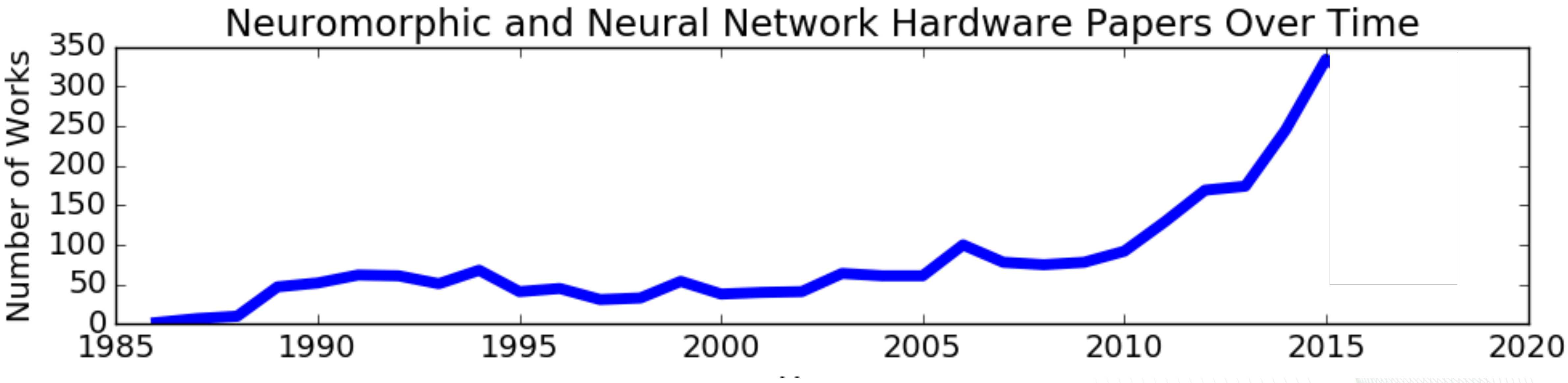
# Neuromorphic Computing Today

- ”Neuromorphic" computing distributes both computation and memory among an enormous number of relatively primitive "neurons," each communicating with hundreds or thousands of other neurons through "synapses." - Don Monroe, “Neuromorphic Computing Gets Ready for the (Really) Big Time”, CACM, June, 2014.

- “Although in the original definition, the term neuromorphic was restricted to the set of analog VLSI circuits that operate using the same physics of computation used by the nervous system (e.g., silicon neuron circuits that exploit the physics of the silicon medium to directly reproduce the bio-physics of nervous cells), **the definition has now been broadened to include analog/digital hardware implementations of neural processing systems, as well as spike-based sensory processing systems**.” – Indiveri, et al., “Neuromorphic silicon neuron circuits,” Frontiers in Neuroscience, May 2011.

# Neuromorphic Computing Today

**A neuromorphic computer is a computer whose underlying architecture and the way that it performs computation is inspired by biological brains.**



# Why neuromorphic computing?

# Why Neuromorphic Computing?

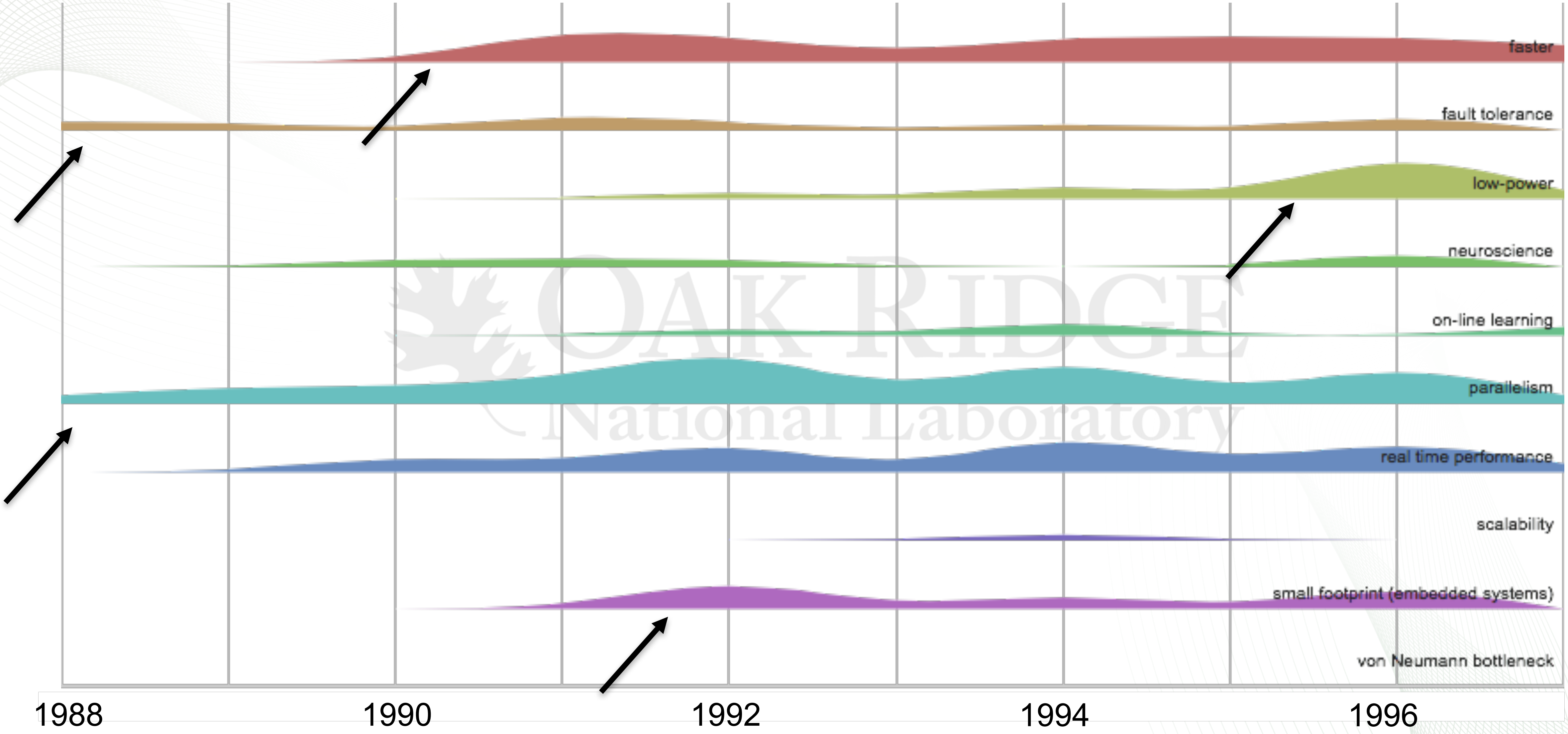
- Carver Mead's reasons for analog neuromorphic systems:
  - **Power efficiency and size:** "Perhaps the most intriguing result of these experiments has been the suggestion that adaptive analog systems are 100 times more efficient in their use of silicon, and they use 10,000 times less power than comparable digital systems."
  - **Robustness**: "It is also clear that these systems are more robust to component degradation and failure than are more conventional systems."
  - **Beyond silicon**: "I have also argued that the basic two-dimensional limitation of silicon technology is not a serious limitation in exploiting the potential of neuromorphic systems."

Source: Mead, Carver. "Neuromorphic electronic systems." *Proceedings of the IEEE* 78.10 (1990): 1629-1636.

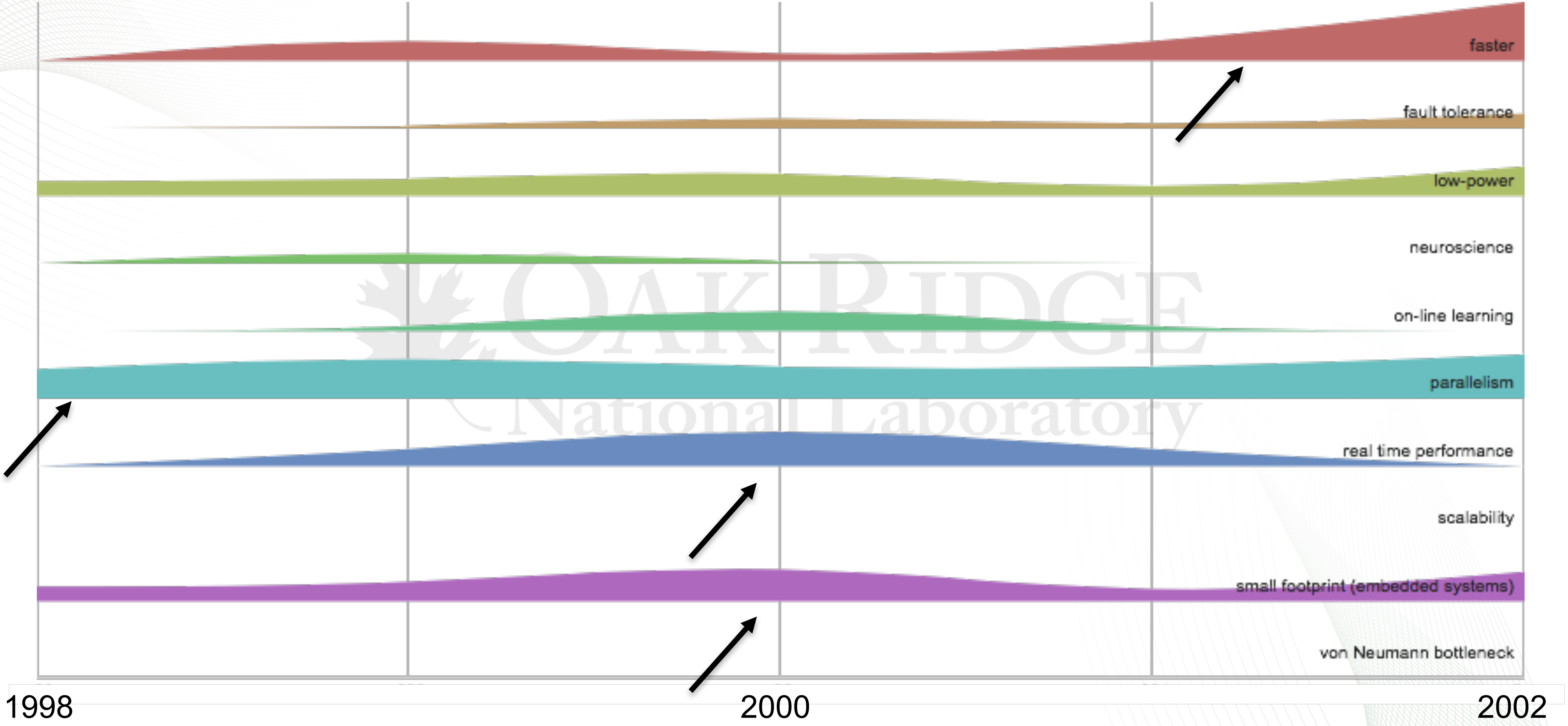
# Motivations for Neuromorphic and ANNs in Hardware

1. Low power
2. Parallelism
3. Faster/Speed
4. Real-time performance
5. Small footprint
6. Fault tolerance/Robustness
7. Neuroscience
8. On-line learning
9. Scalability
10. von Neumann bottleneck

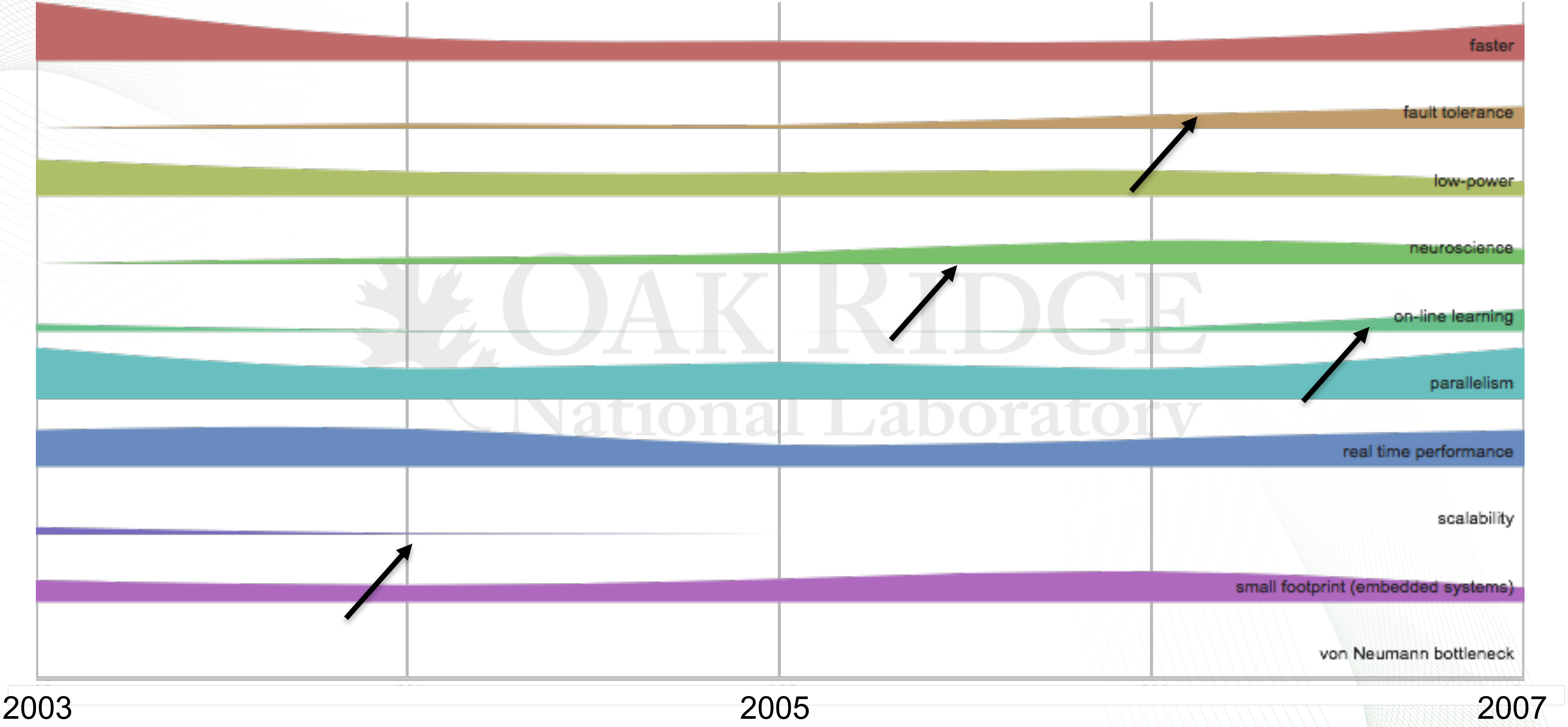
# Motivations: 1988-1997



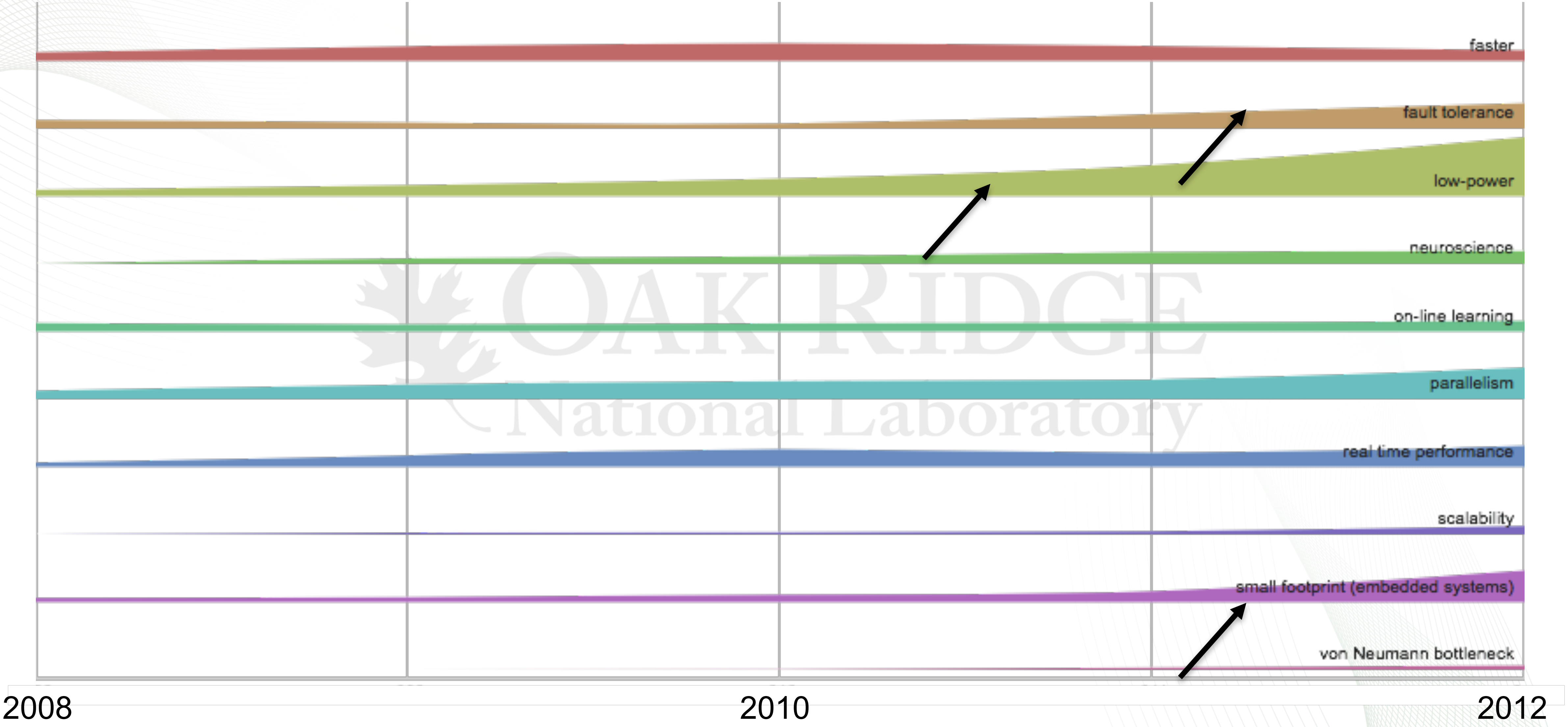
# Motivations: 1998-2002



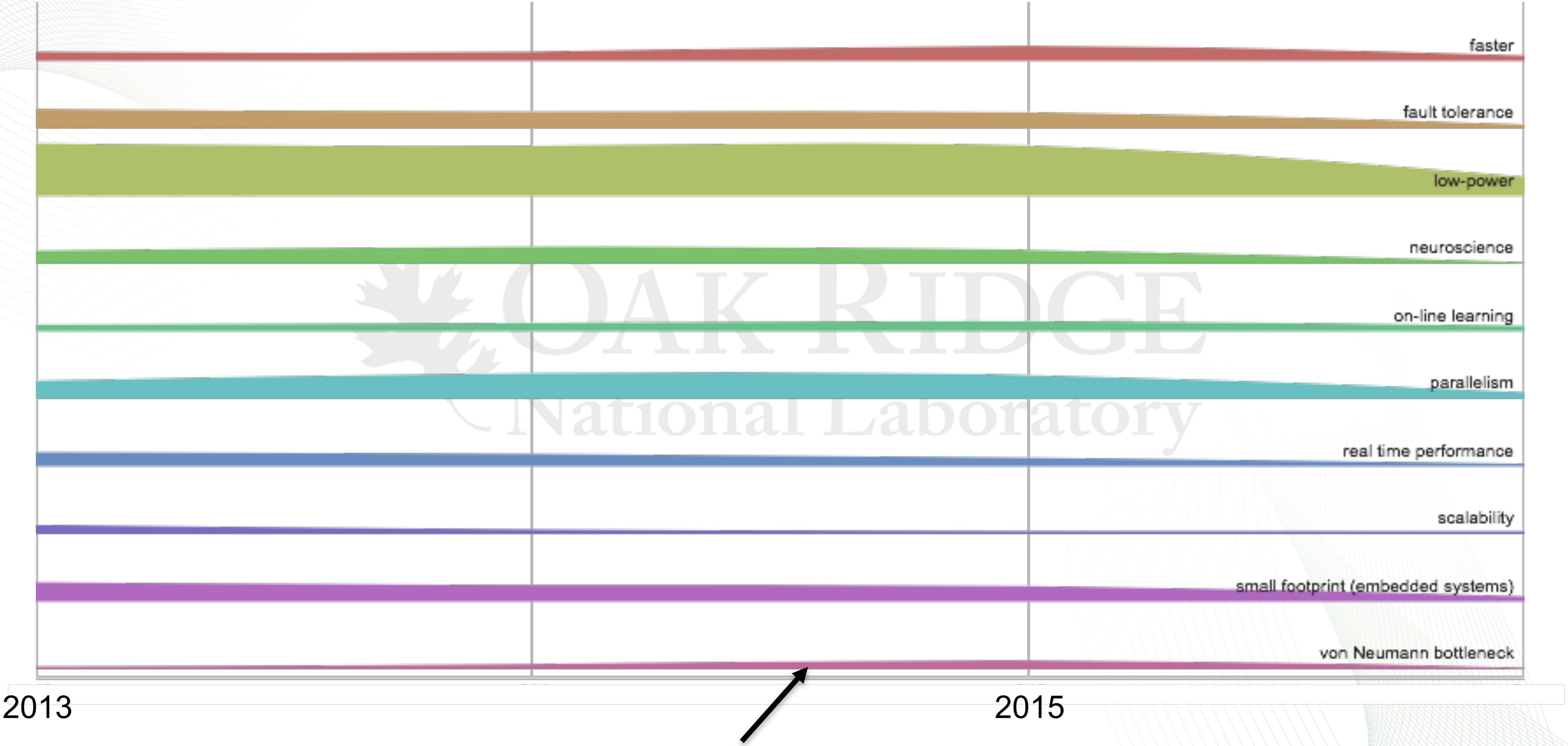
# Motivations: 2003-2007



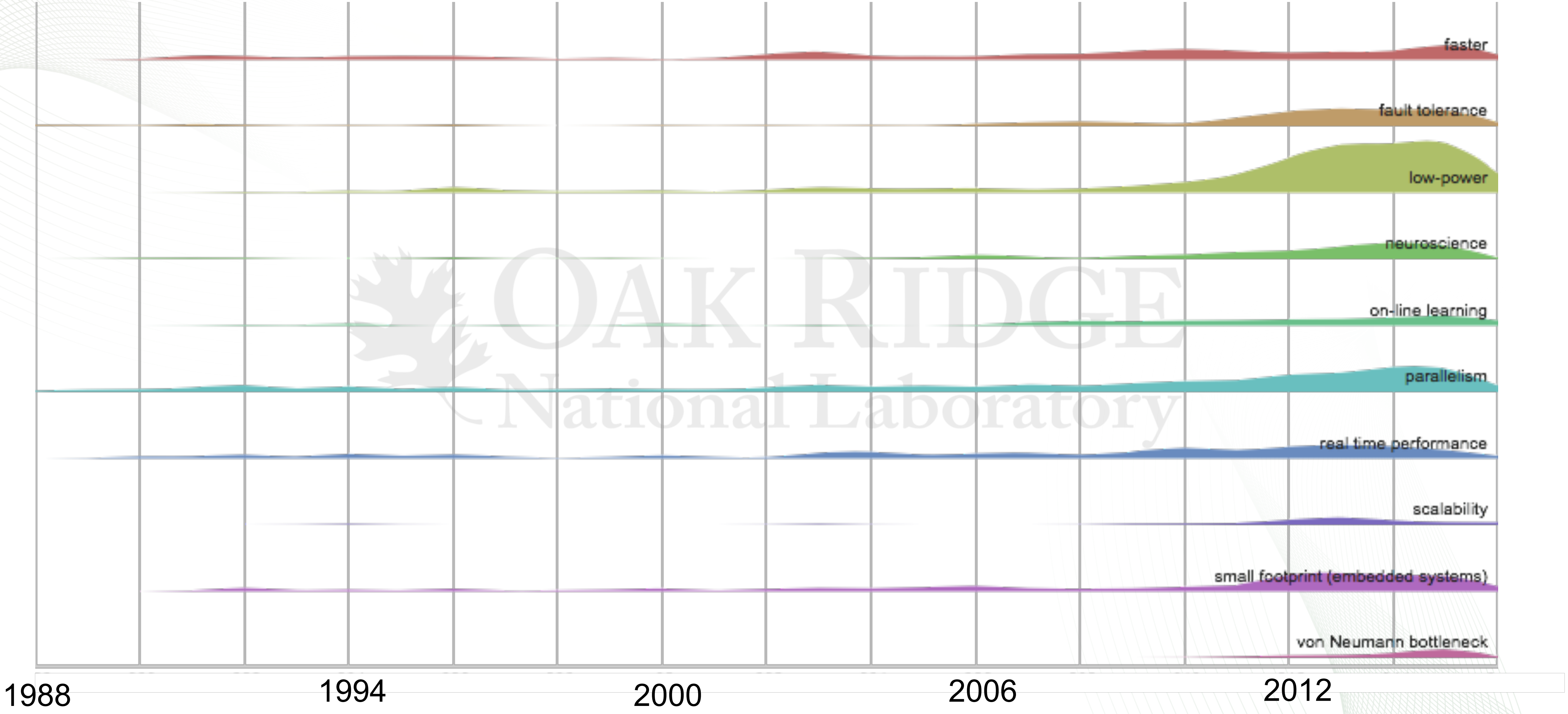
# Motivations: 2008-2012



# Motivations: 2013-2016



# Overall View: 1988-2016



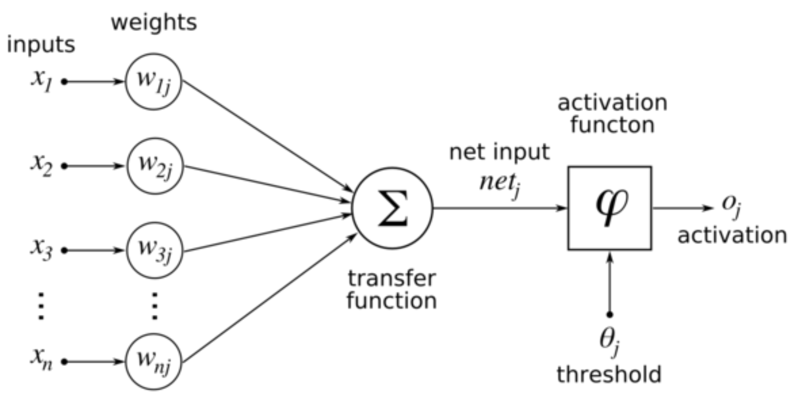
# Key Questions in Neuromorphic Computing

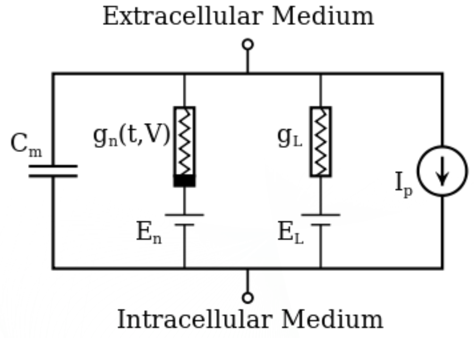
- Models
- Training and Learning
- Hardware, Devices, and Materials
- Applications

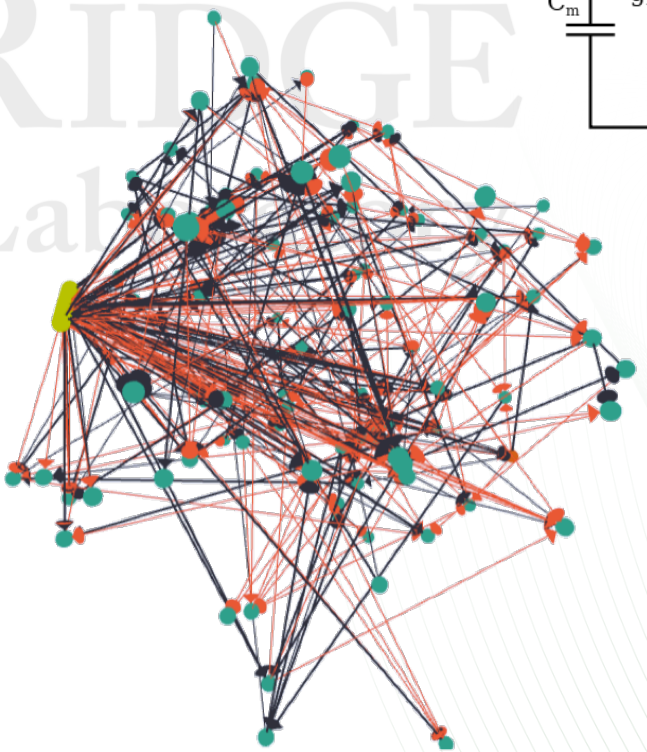
# What model should be implemented?

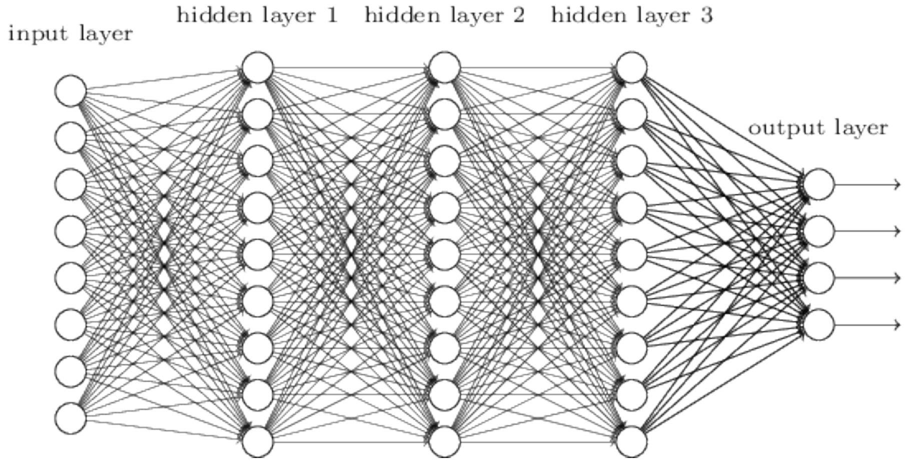
# Models

- Neuron models:
  – Works that are specifically focused on building a particular neuron model in hardware.
- Network models:
  – Works that implement full network models in hardware.

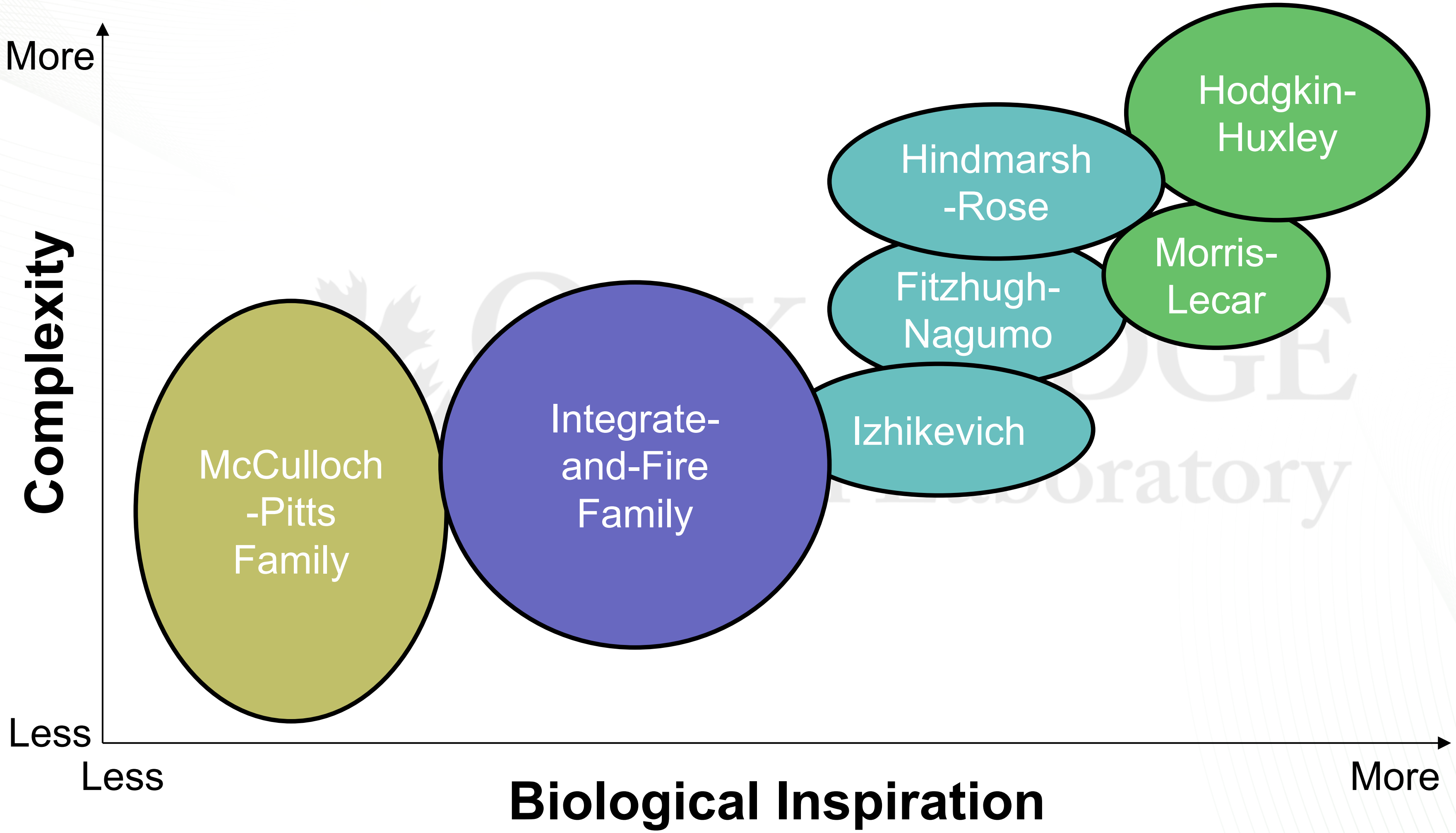


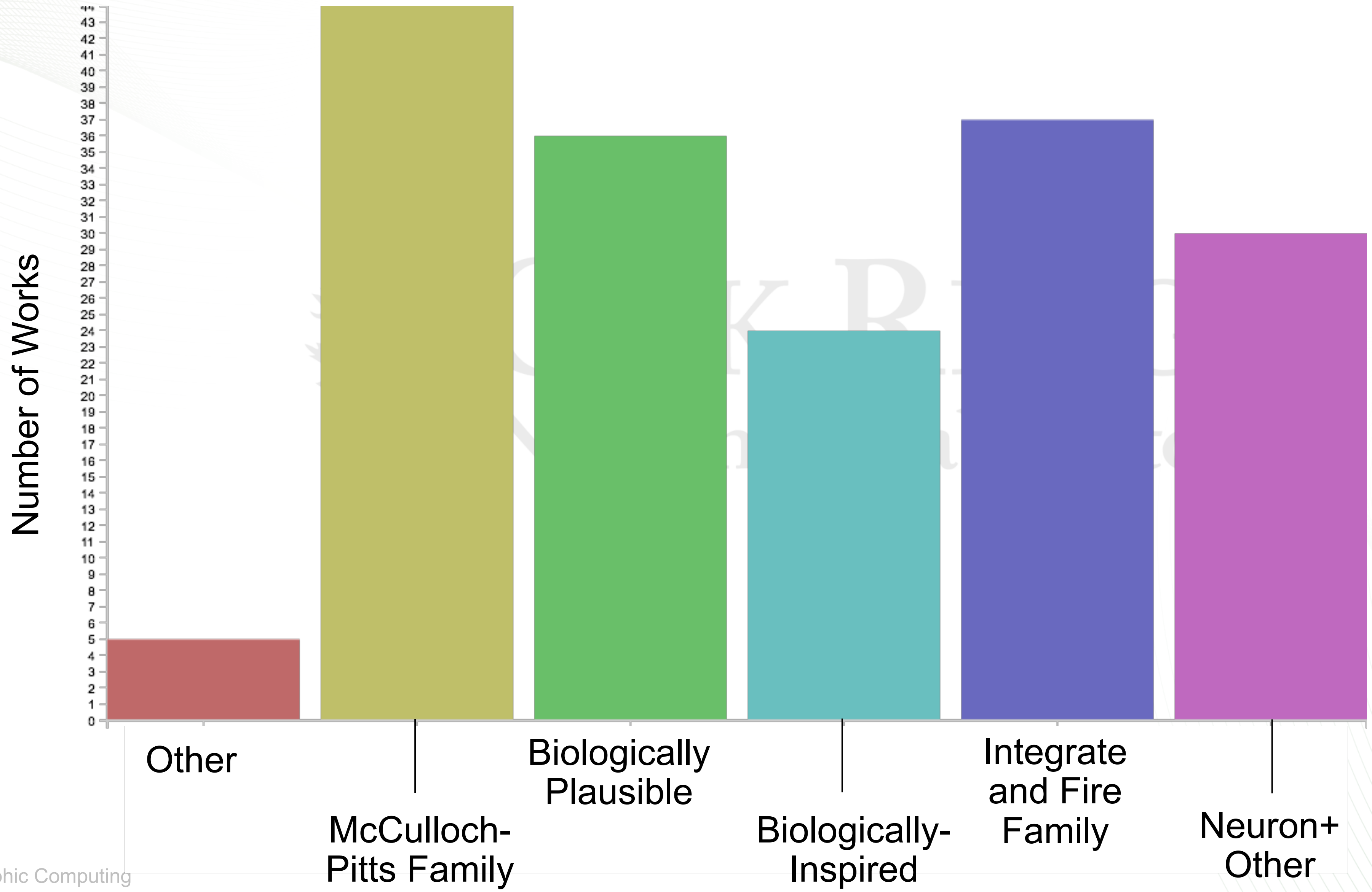





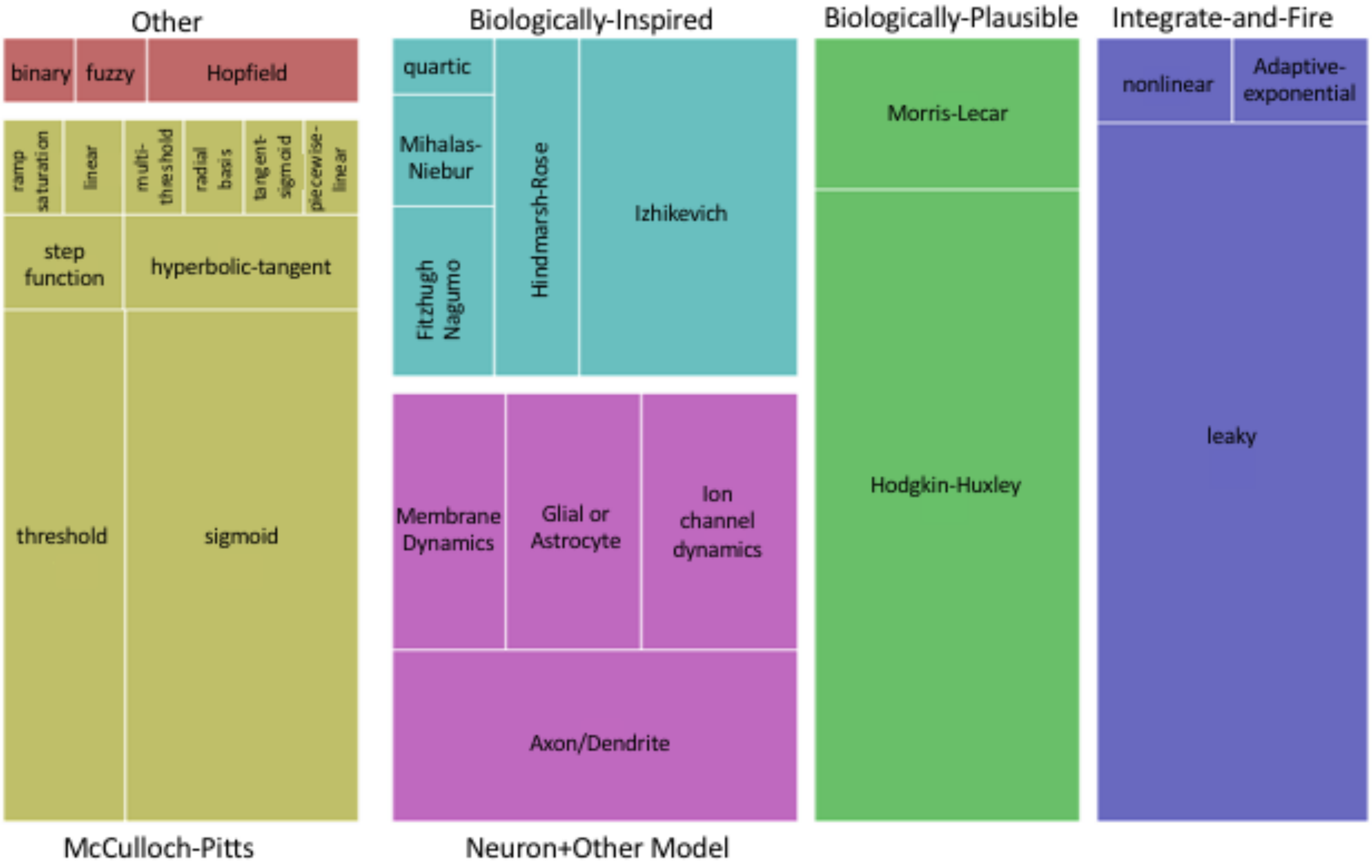
# Neuron Models



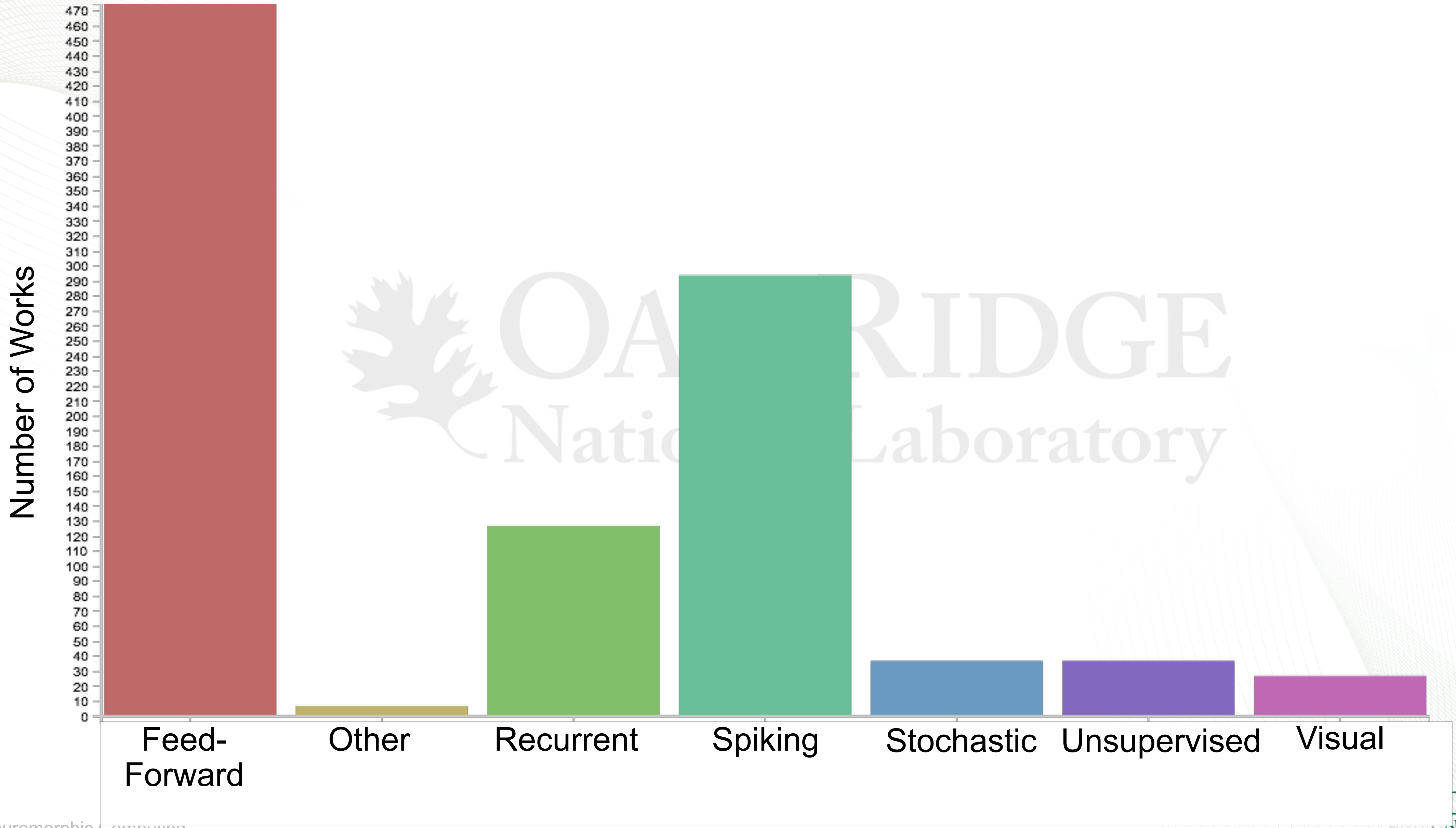
# Neuron Model Families



# Neuron Models

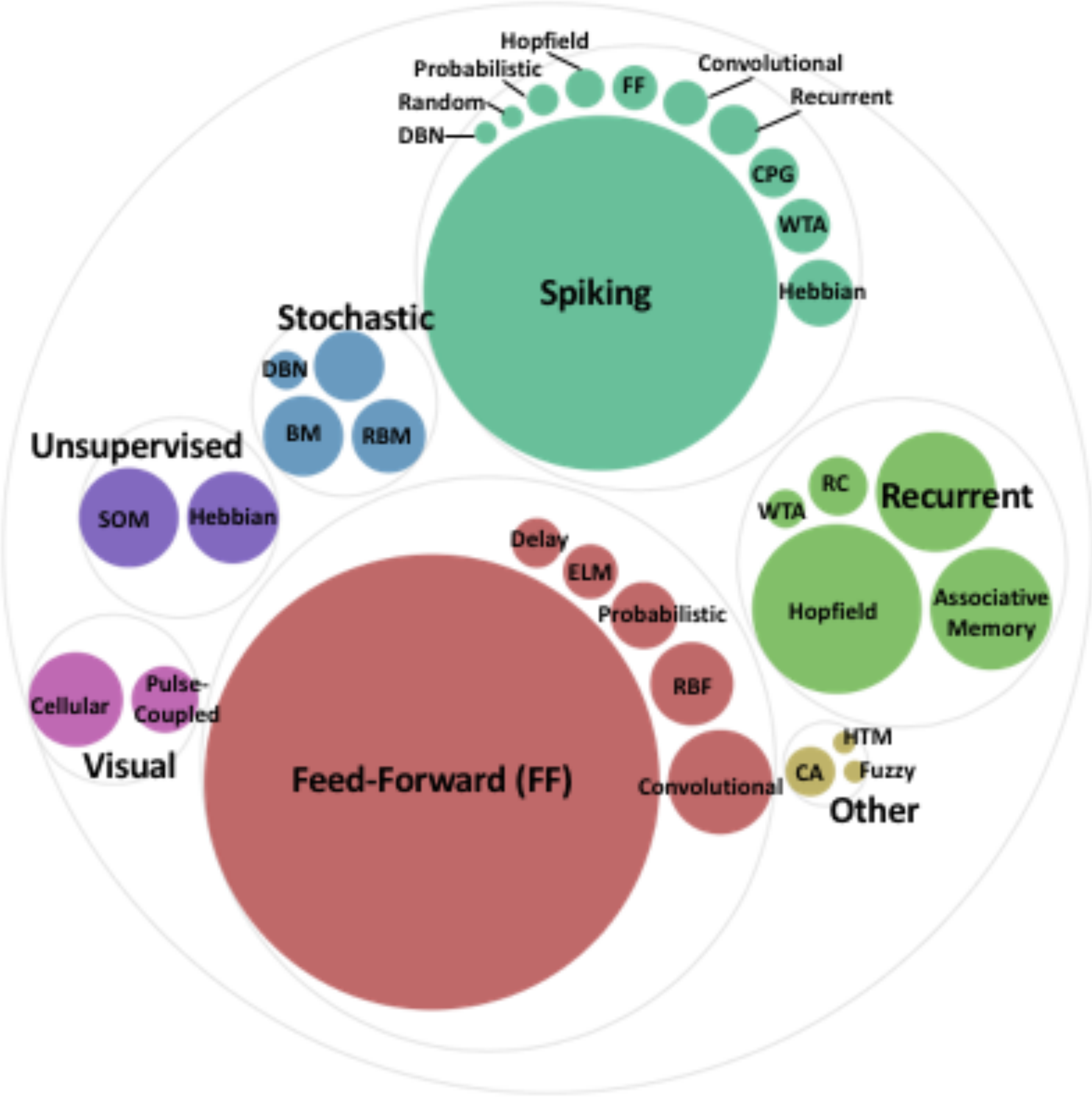


# Network Model Families

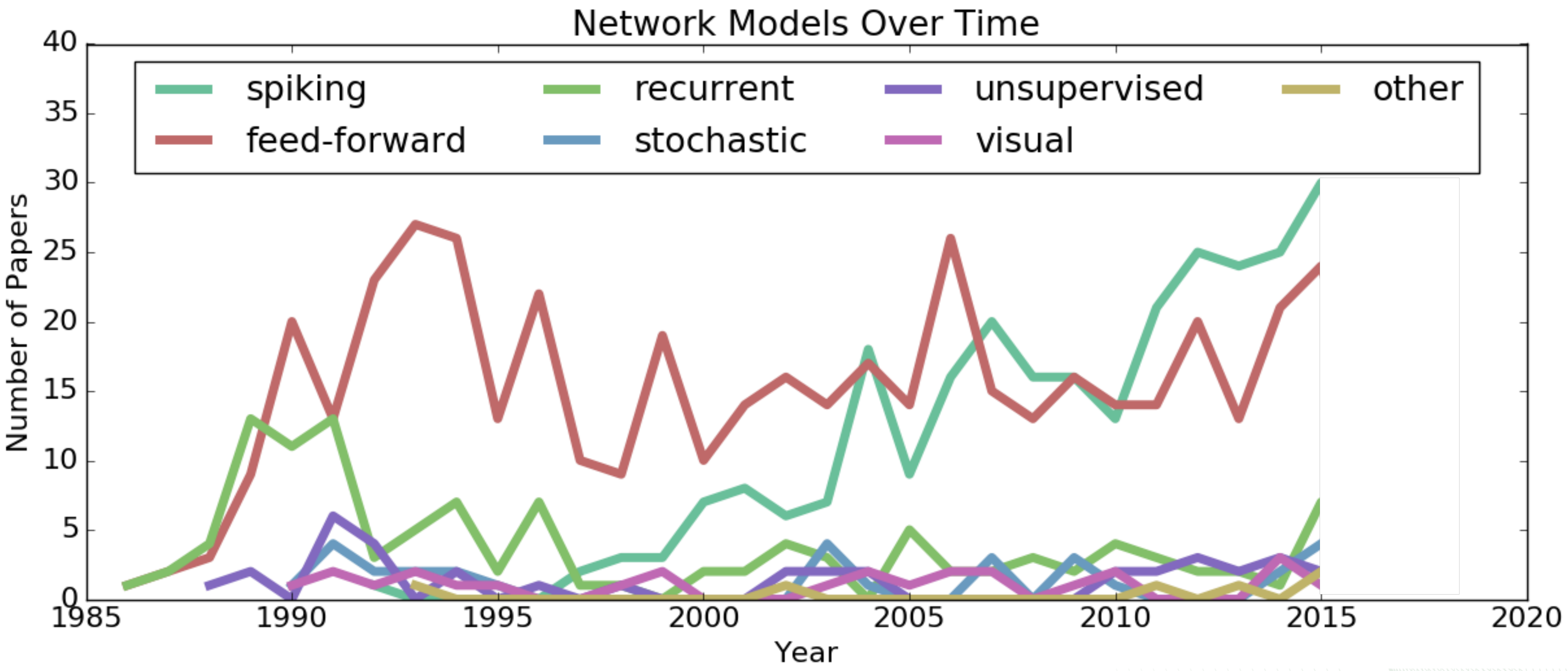


# Network Models

- A wide variety of network models have been implemented in hardware or neuromorphic systems.
- Spiking neuromorphic systems have been used to implement a variety of other types of network models.



# Network Models



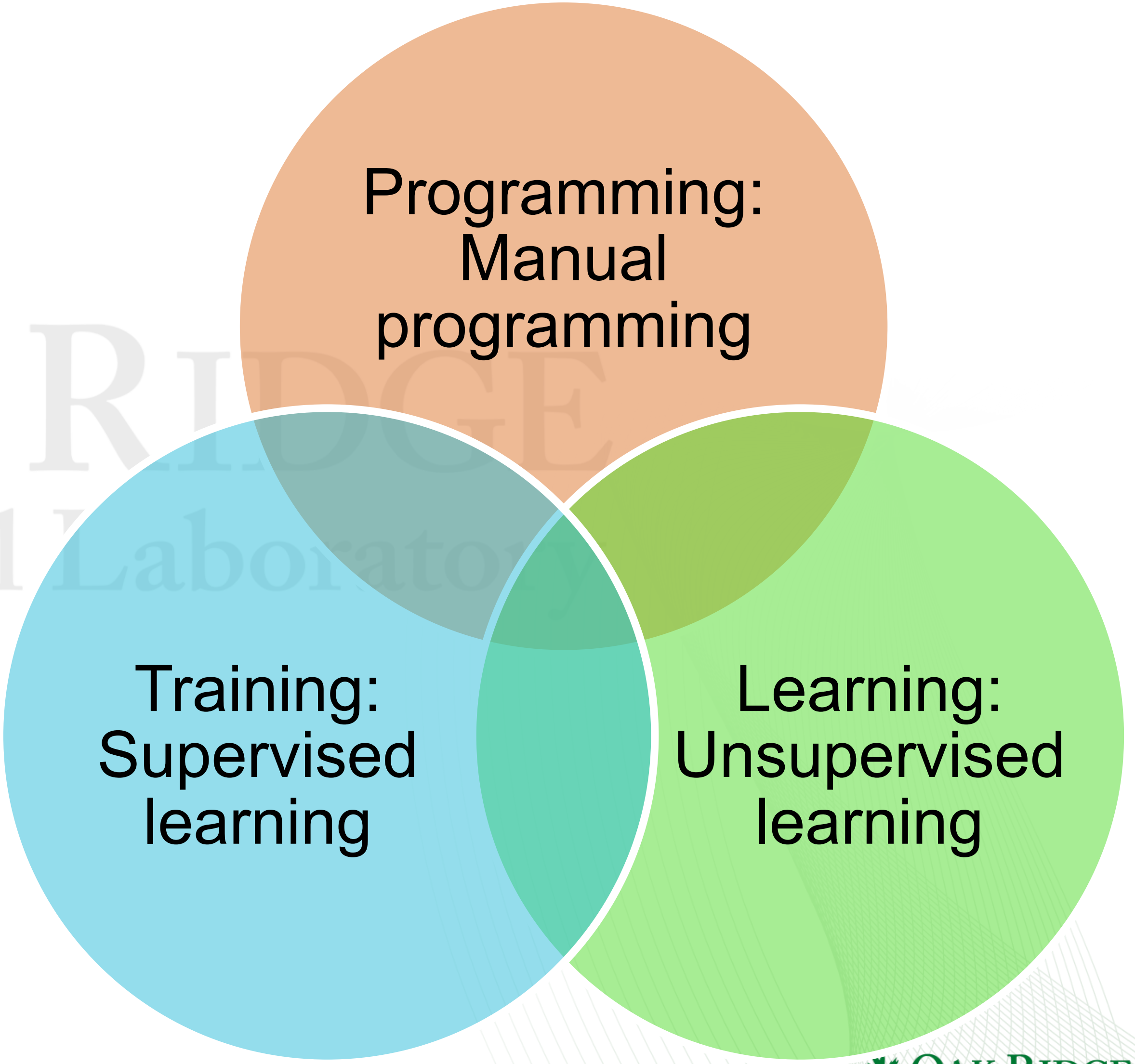
# General Model Concerns and Open Questions

- Many spiking implementations are translations of other types of neural networks onto spiking networks.
  - Pro: Leverage previous research
  - Con: Inhibits development on spiking neural networks themselves
- How restricted should the neuromorphic system be?
  - Topology – Fixed or programmable?
  - Neuron/synapse model – Capabilities that can be tuned or turned on and off?
- Can we leverage neuromorphic systems to accelerate network model study?
  - E.g., as GPUs did for convolutional neural networks and deep learning

# What training/learning method should be used?

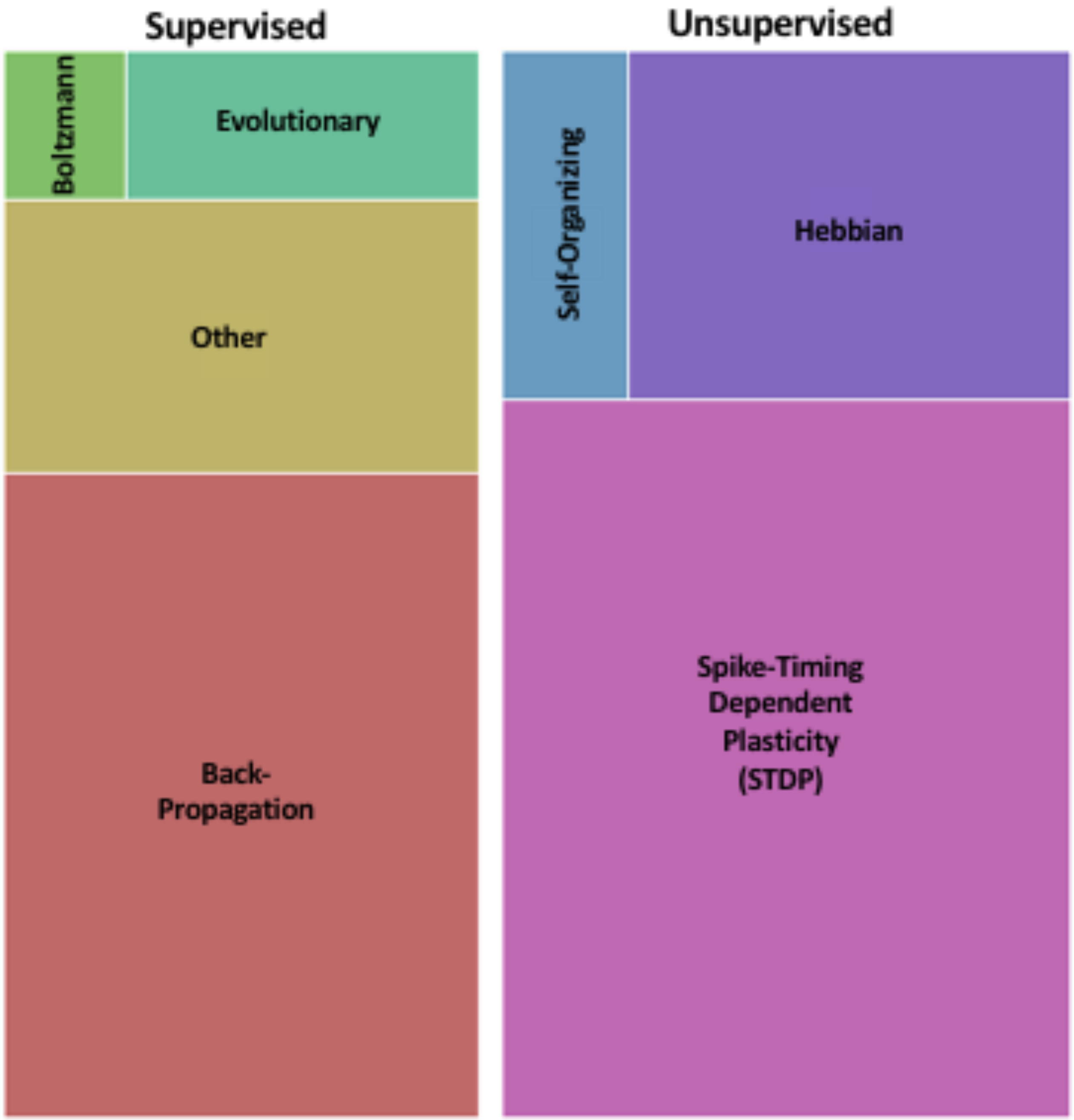
# Choosing an Appropriate Training/Learning Algorithm

- Programming is primarily used for systems that are neuroscience-driven.
  - Setting parameters based on those observed in biological brains.
- Training algorithms are the most well-understood and have had significant demonstrated success.
- Learning algorithms are the “holy grail.”



# Training/Learning

- Training and learning algorithms here are those that have been developed specifically for neuromorphic systems, and one of the following is true:
  - Customized in some way to deal with restrictions
  - Implemented on-chip
  - Chip-in-the-loop



# Training/Learning Mechanisms

| Algorithm | Any Model | Device Quirks | Complex to Implement | On-Line | Fast Time to Solution | Demonstrated Broad Applicability | Biologically-Inspired or Plausible |
|---|---|---|---|---|---|---|---|
| **Back-Propagation** | No | Maybe | Yes | No | Yes | Yes | Maybe |
| **Evolutionary** | Yes | Yes | No | No | No | Yes | Maybe |
| **Hebbian** | No | Yes | No | Yes | Maybe | No | Yes |
| **STDP** | No | Yes | Maybe | Yes | Maybe | No | Yes |

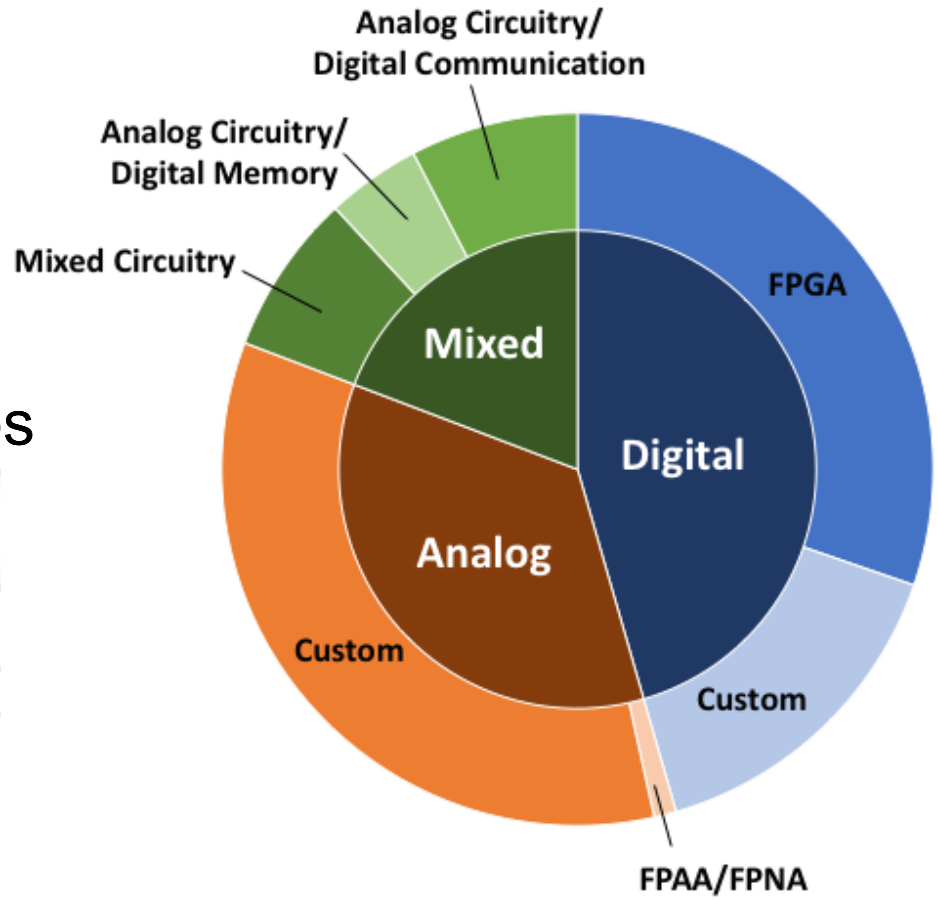
# Training/Learning – Concerns and Open Questions

- STDP is frequently implemented, but its capabilities are not frequently explored.
  - Though it's been shown to be useful in several cases, the specific implementation of STDP and the context of the greater model can have a significant effect on performance.
- It's not enough to train off-line and off-chip.
  - Training/learning is clearly an important component of neuromorphic systems use.
  - We must consider how the neuromorphic systems themselves can be used during part or all of the training process.

# What hardware/devices/materials should be used to implement neuromorphic systems?

# Hardware Implementations

- Different hardware classification criteria:
  - Digital/Analog/Mixed/Other
  - Programmable Architectures vs. Custom Chips
  - Degree of parallelism
  - “General purpose” vs. application specific
  - On-chip or off-chip learning
  - Input types
  - Communication networks

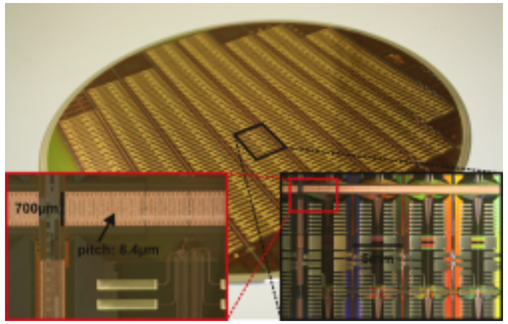



N. Izeboudjen, C. Larbes, and A. Farah, “A new classification approach for neural networks hardware: from standards chips to embedded systems on chip,” *Artificial Intelligence Review*, vol. 41, no. 4, pp. 491–534, 2014.

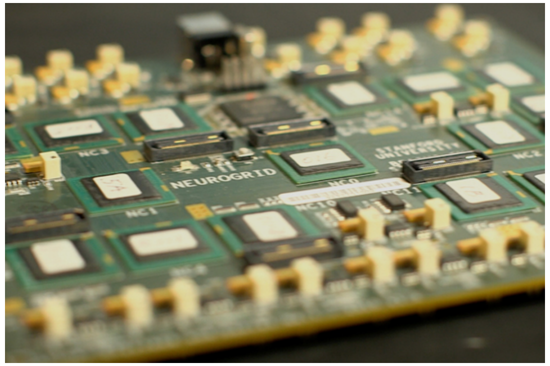
# Major Neuromorphic Projects

## SpiNNaker:
- Fully digital
- Many small integer cores
- Custom interconnect
- Flexible model and topology


## TrueNorth:
- Fully digital
- Custom ASIC
- Fixed model (LIF)
- Highly optimized



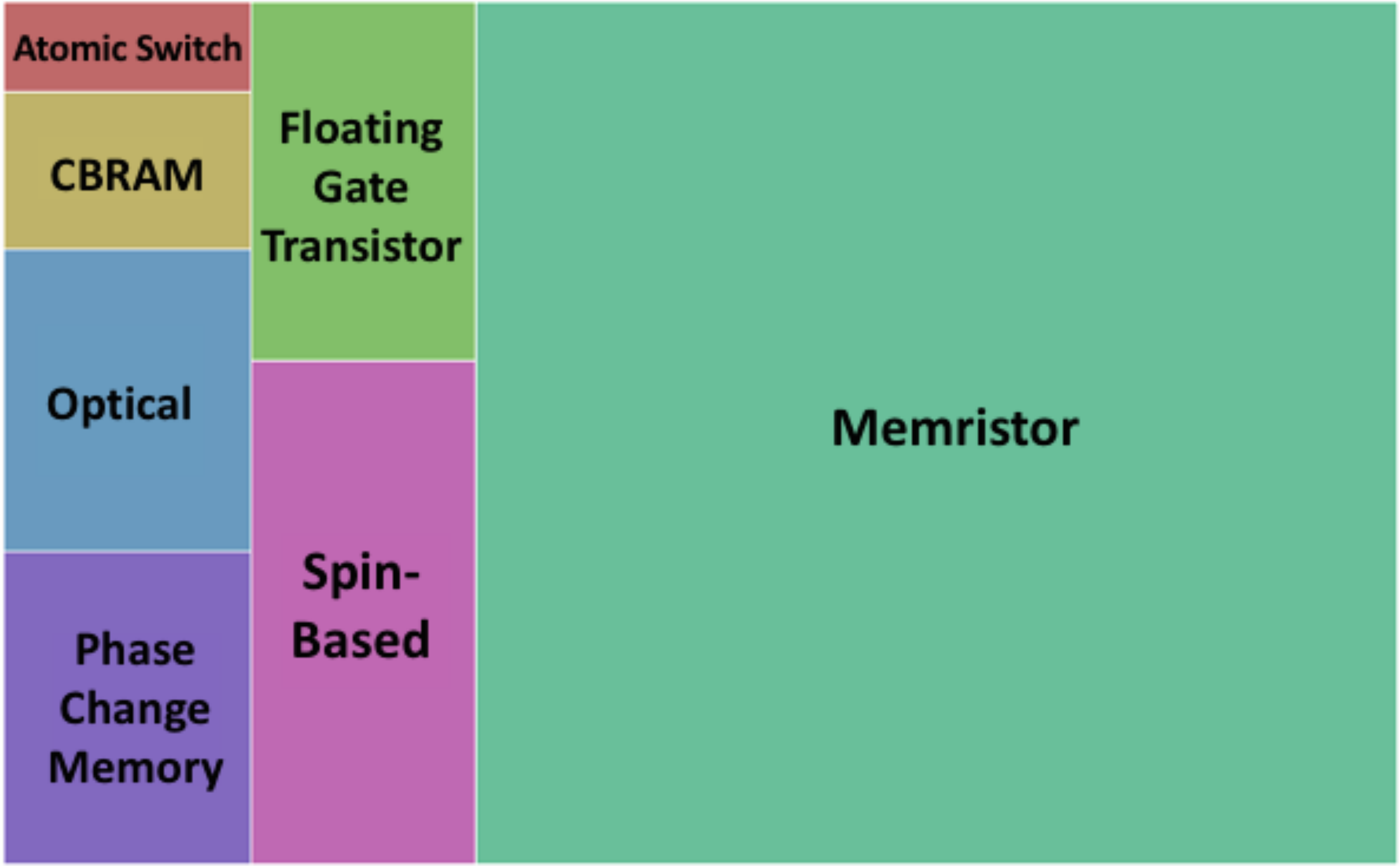

## BrainScaleS:
- Hybrid analog/digital
- Wafer-scale
- Super-threshold operation
- Relatively high clock rate




## Neurogrid
- Hybrid analog/digital
- Sub-threshold operation
- Relatively low clock rate

SpiNNaker Image: http://wp.doc.ic.ac.uk/hipeds/wp-content/uploads/sites/78/2016/01/The-SpiNNaker-Project-Seminar-Slides.pdf
IBM TrueNorth: http://www.techrepublic.com/article/ibms-brain-inspired-chip-truenorth-changes-how-computers-think-but-experts-question-its-purpose/
BrainScaleS: http://www.artificialbrains.com/brainscales
Neurogrid: http://news.stanford.edu/pr/2014/pr-neurogrid-boahen-engineering-042814.html

# Emerging Components

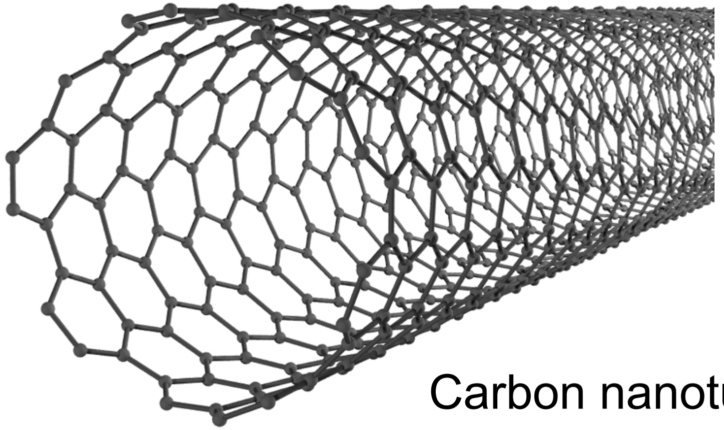




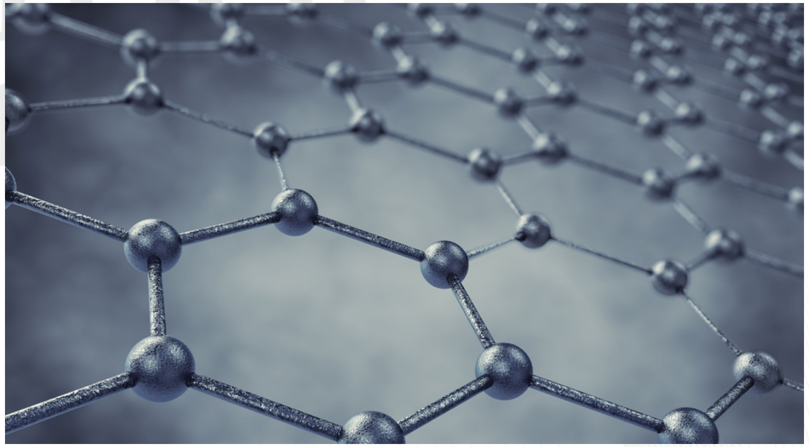

OAK RIDGE National Laboratory

# Materials

- Materials:
  – A variety of metal-oxides
    - $HfO_x$, $TiO_x$, $WO_x$, $TaO_x$, etc.
  – Carbon nanotubes
  – Graphene
  – Ferroelectric materials
  – Polymer and organic-based memristors and transistors



Carbon nanotubes



Graphene

Carbon nanotube: https://www.digitaltrends.com/computing/ibm-carbon-nanotubes-moores-law/
Graphene: https://www.extremetech.com/extreme/211437-extremetech-explains-what-is-graphene

# Materials

- Circuit components fabricated with different materials can have different behaviors:
  - Number and type of resistance states
  - Endurance
  - Stability
  - Reliability
  - Switching speeds
  - Cost
  - Tunability

**Table 1:** Examples of bipolar metal oxide memristors and their operational characteristics.

| Material | TE/BE | $V_{SET}/V_{RESET}$ | $\Delta R=R_{Off}/R_{On}$ | Switching speed | Retention time | Endurance | Fabrication process | References |
|---|---|---|---|---|---|---|---|---|
| ZnO | Ag/Cu | 1.2 V/-1.25 V | 1000 | – | – | >500 cycles | Electrohydrodynamic printing | [110] |
| | Pt/Pt | 1 V/-0.5 V | 100 | 10 ms | – | $10^6$ cycles | RF-magnetron sputtering | [111] |
| $TiO_2$ | Pt/Pt | – | 100 | – | – | – | ALD | [112] |
| | TiN/Pt | +1 V/-1.5 V | 10 | 1 μs | $10^4$ s | $10^4$ cycles | RF-reactive sputtering | [97] |
| | TaN-TiN/TiN-TaN | 1.5 V/-1.5 V | 1.5 | – | – | $10^2$–$10^3$ cycles | Sputtering | [113] |
| | Al/Al | 3 V/-2 V | ~50 | – | $10^4$ s | 100 cycles | Plasma-enhanced ALD | [114] |
| LaO | ITO/$SrTiO_3$ | 5 V/-1.6 V | 200 | – | >$4\times10^4$ s | 2000 cycles | Pulsed laser deposition | [115] |
| $TaO_x$ | Pt/Pt | – | – | | 10 years at 85°C | $10^9$ cycles | Sputtering | [116] |
| | W/Pt | – | – | – | >10 years | $10^4$ cycles | RF-magnetron sputtering | [117] |
| NiO | Pt/Pt | >10 V/<-10 V | – | – | >$10^4$ s | – | Pulsed laser deposition | [118] |
| | Au/Au | +5.2 V/-6 V | – | – | – | – | Electrochemical plating | [96] |
| $HfO_2$ | TiN/TiN | 1.5 V/-1.4 V | 100 | <10 ns | >500 min at 200°C | >$10^6$ cycles | ALD | [93] |
| | TiN/TiN | – | >50 | 5 ns | $10^5$ s at 200°C | $5\times10^7$ cycles | ALD | [61] |
| $ZrO_2$ | ITO/Ag | 1 V/-1 V | >10 | – | $10^6$ s at 27°C | >50 cycles | Electrohydrodynamic printing | [119] |
| | Ag/Ag | 3 V/-3 V | ~100 | – | – | – | Electrohydrodynamic printing | [120] |
| | TiN/Pt | 0.8 V/-0.5 V | – | – | $10^4$ s at 27°C | $10^3$ cycles | RF-magnetron sputtering | [121] |
| $CeO_2$ | Au/Au | 2.4 V/-3 V | $10^4$ | – | – | – | Sol-gel (drop-coating) | [122] |
| $AlO_x$ | Al or CNT/CNT | – | – | – | $10^5$ s | $10^4$ cycles | ALD | [123] |
| | Cu/W | 1.3 V/-0.05 V | 500 | – | $10^3$ s | – | E-beam evaporation | [124] |
| $Al_2O_3$ | Ti/Pt | 1.4 V/-1.7 V | <1000 | 10 ns | $10^4$ s | – | RF-magnetron sputtering | [59] |
| $Cu_2O$/CuO | Pt//Nb-STO | – | $10^5$ | – | – | – | Plasma assisted molecular beam epitaxy | [125] |
| $Gd_2O_3$ | ITO/ITO | +2 V/-2 V | – | – | – | $10^3$ cycles | Pulsed laser deposition | [126] |
| $GdO_x$ | Cr/TiN | <+4 V/-4 V | >70 | – | $3\times10^4$ s | $10^5$ cycles | E-beam evaporation | [127] |
| MnO | Ti/Pt | 0.7 V/-1.1 V | – | – | $10^4$ s at 85°C | $10^5$ cycles | RF-reactive sputtering | [128] |

TE, top electrode; BE, bottom electrode; "–", data not found in the associated reference paper.

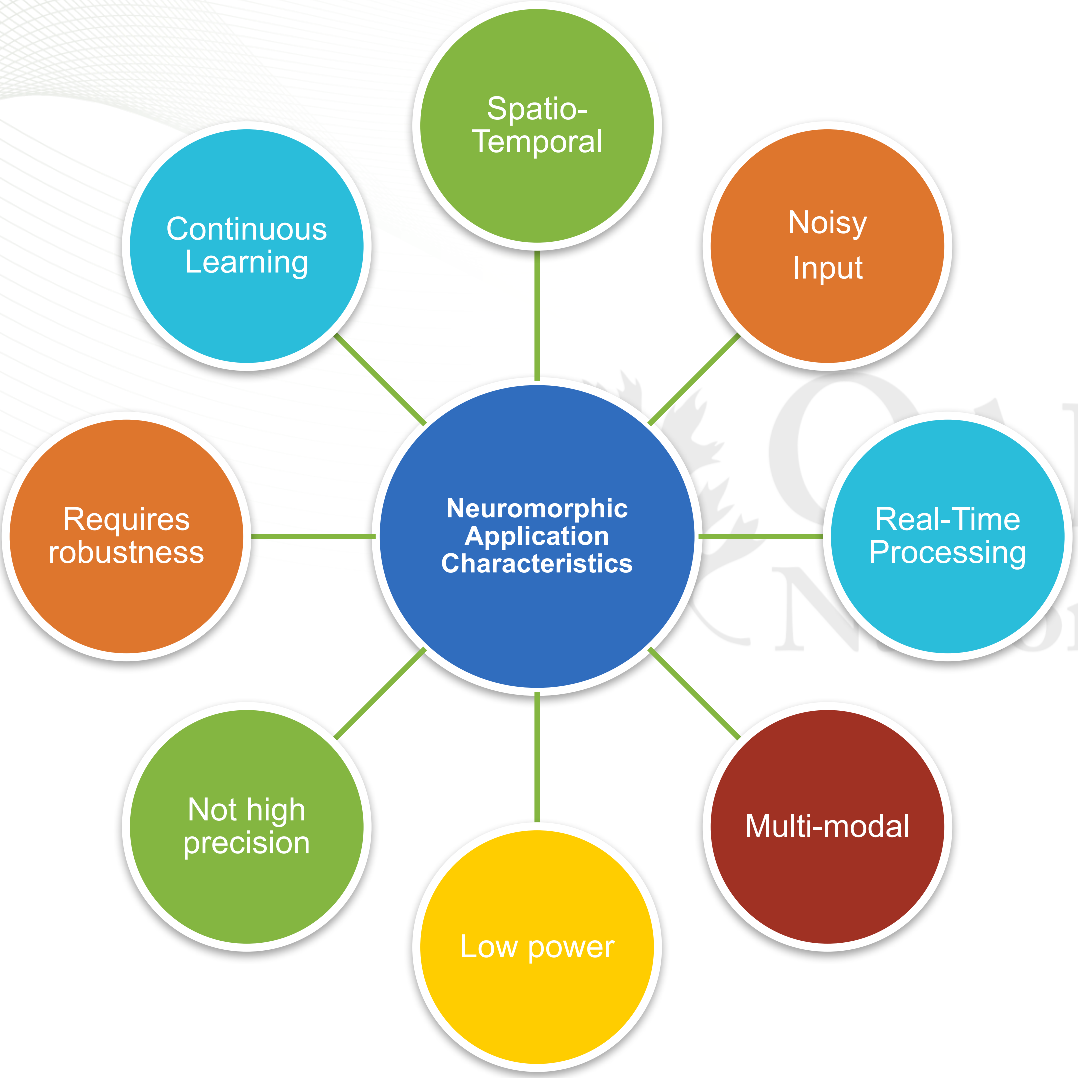
Table: Mohammad, Baker, et al. "State of the art of metal oxide memristor devices." *Nanotechnology Reviews* 5.3 (2016): 311-329.

# Hardware/Devices/Materials Open Questions and Concerns

- High-level chip design questions: programmability, general vs. application-specific, digital/analog/hybrid, etc.
- What are the most appropriate emerging components to use to build neuromorphic systems?
- Can we customize materials/selection of materials specifically for neuromorphic use?
- Most literature in the materials science and low-level circuits does not tie to real usage:
  - Co-design will be extremely important moving forward!

# What are the appropriate applications for neuromorphic computers?

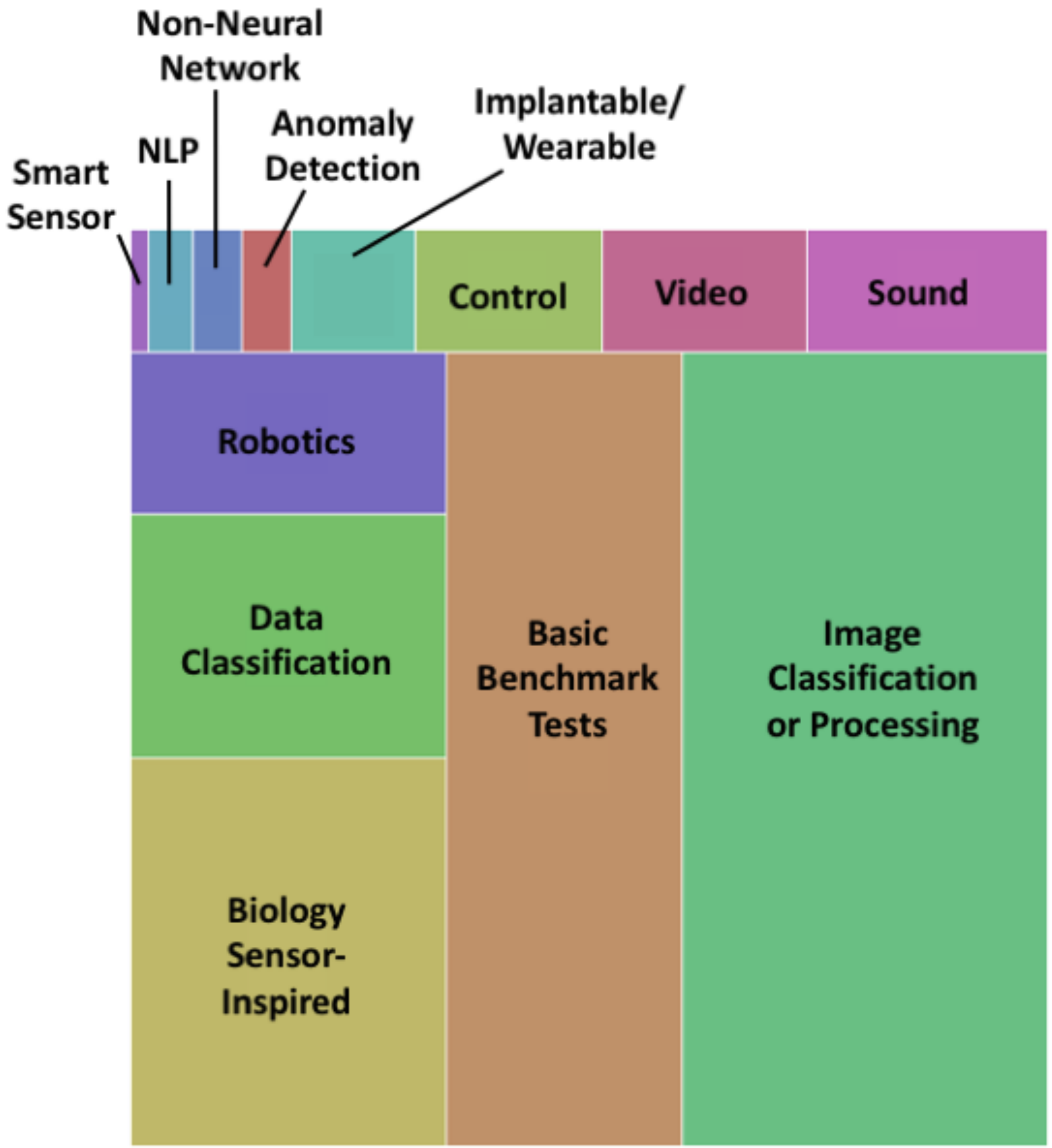
# Applications of Neuromorphic Computing



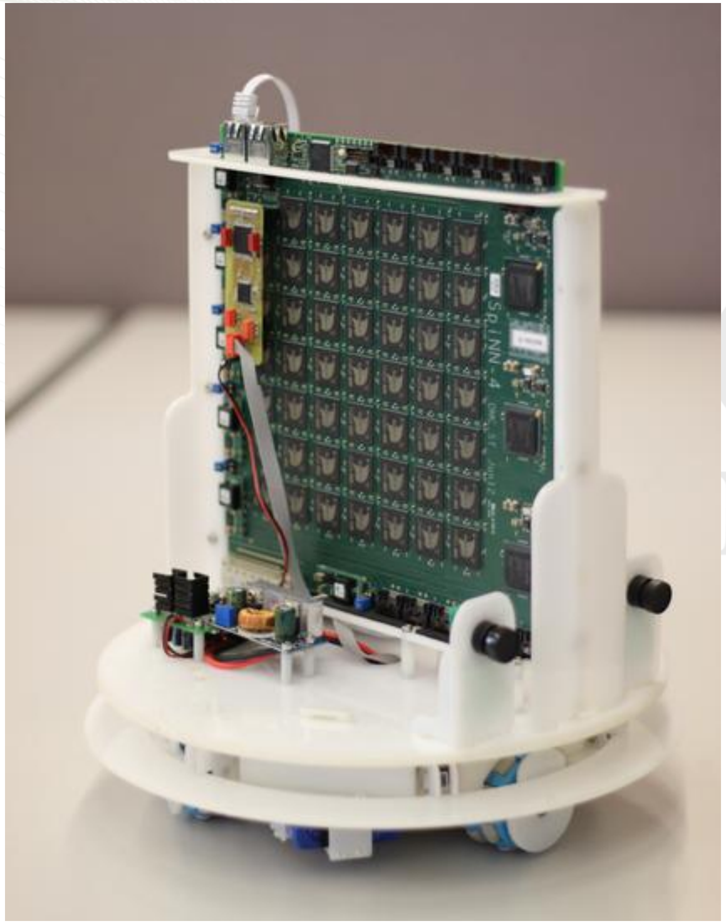

- Co-processor
- Large-scale data analytics
- Cyber security
- Autonomous vehicles
- Robotics
- Internet of things
- Smart sensors

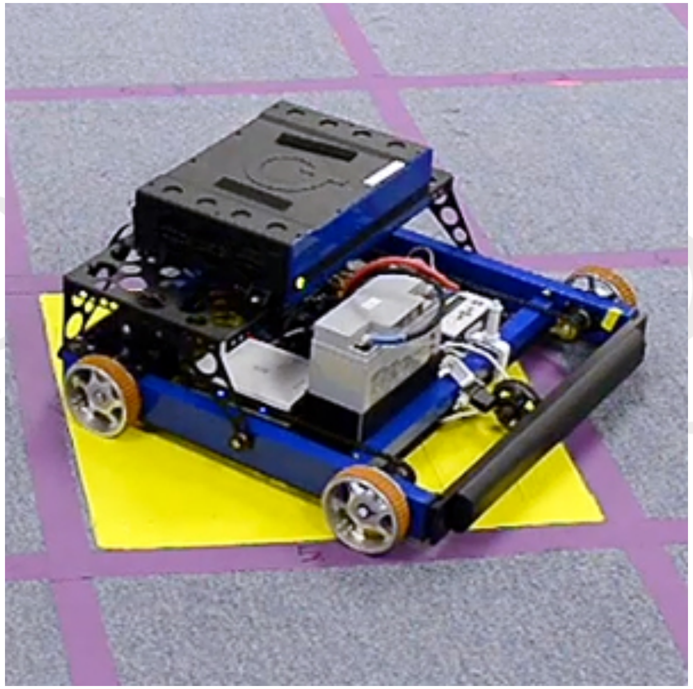
# What's Been Done?

- Tremendous focus on image classification and processing.
- Opportunity for big impacts in:
  - Implantables/wearables
  - Internet of things
  - Smart sensors
  - Robotics
  - Control
  - Anomaly Detection



# Example Application: Robotics



SpiNNaker on a robot: http://www.neuromorphs.net/nm/wiki/2013/uns13
Qualcomm: https://www.technologyreview.com/s/520211/qualcomm-to-build-neuro-inspired-chips/

# Summary

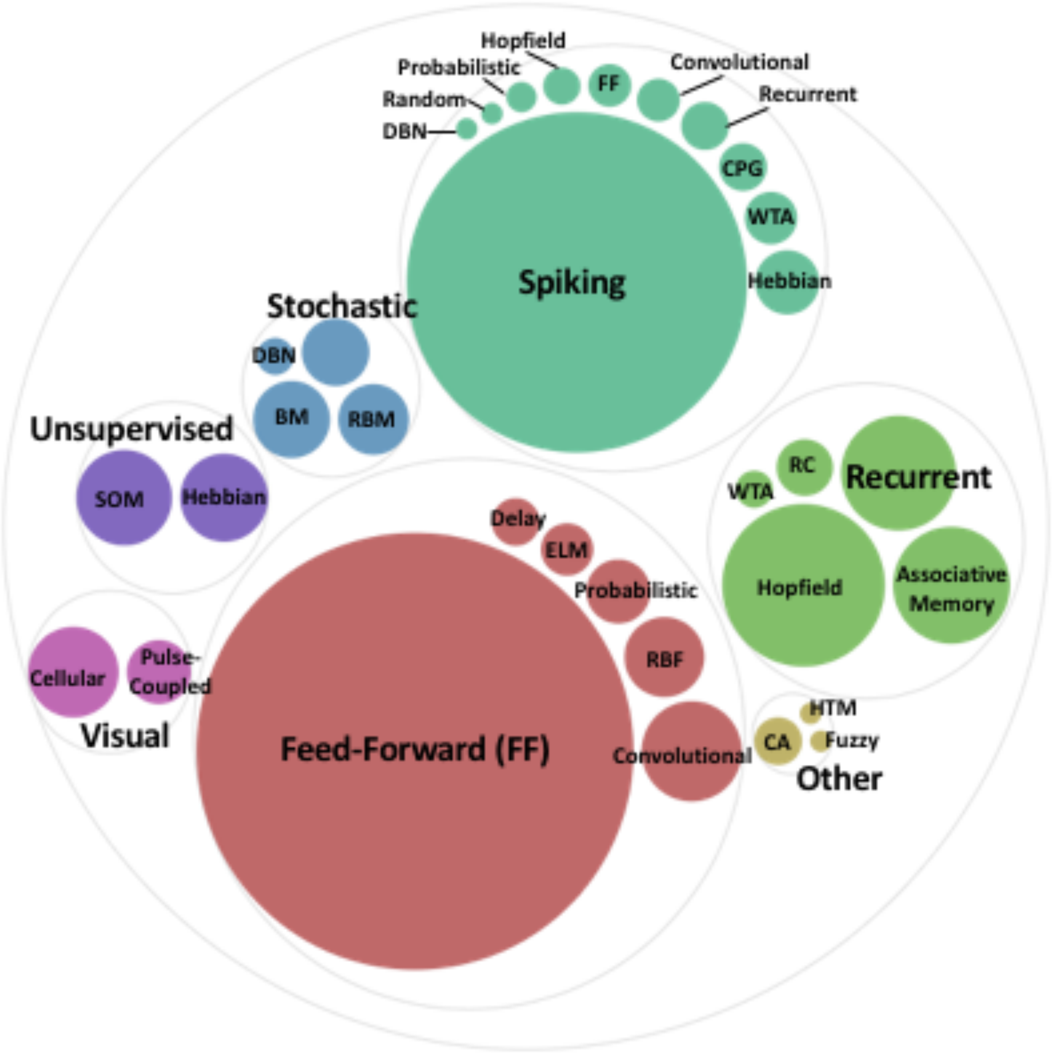
- There has been a significant amount of work in building neuromorphic systems and neural network hardware over the last few decades.
  - Wide variety of models, algorithms, hardware/devices, and applications have been explored.
- However, we have barely scratched the surface of what is possible to do within the field of neuromorphic computing.
  - Significant work to do in algorithm development, custom and emerging hardware and materials for hardware, and applications.

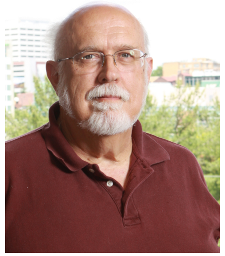
# What's Next?

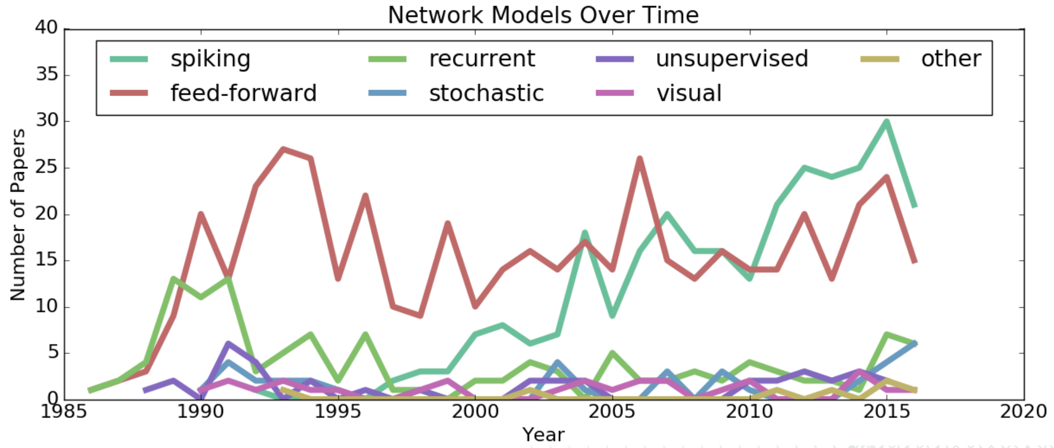
- Model and algorithms research:
  - Wider availability of spiking neuromorphic systems → More research on spiking neural networks and their capabilities and learning algorithms
  - Wider adoption of neuromorphic systems → Increased need for supporting software
- Hardware/devices:
  - Advances in materials fabrication and characterization → Advanced co-design of neuromorphic systems and materials
- Applications:
  - Explosion of use-cases for low power, small footprint smart embedded systems → Opportunity for embedded neuromorphic systems to shine
  - Spatiotemporal scientific data explosion → Opportunities for neuromorphic co-processors or smart in situ analyzers

# Neuromorphic Computing Survey Paper

"A Survey of Neuromorphic Computing and Neural Networks in Hardware"
https://arxiv.org/abs/1705.06963





# Thank You!

Email: schumancd@ornl.gov
Website: CatherineSchuman.com