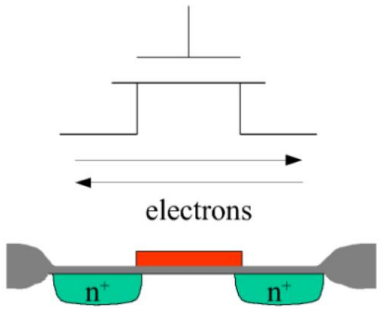
Open (Digital) Neuromorphic (Hardware)

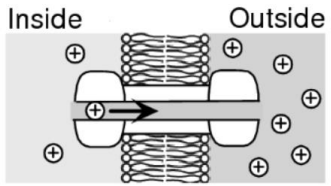
Charlotte Frenkel

Delft University of Technology, Microelectronics Department
c.frenkel@tudelft.nl

Open Neuromorphic, Politecnico di Torino (online)
13th of December 2022

Why the (digital) in the title?





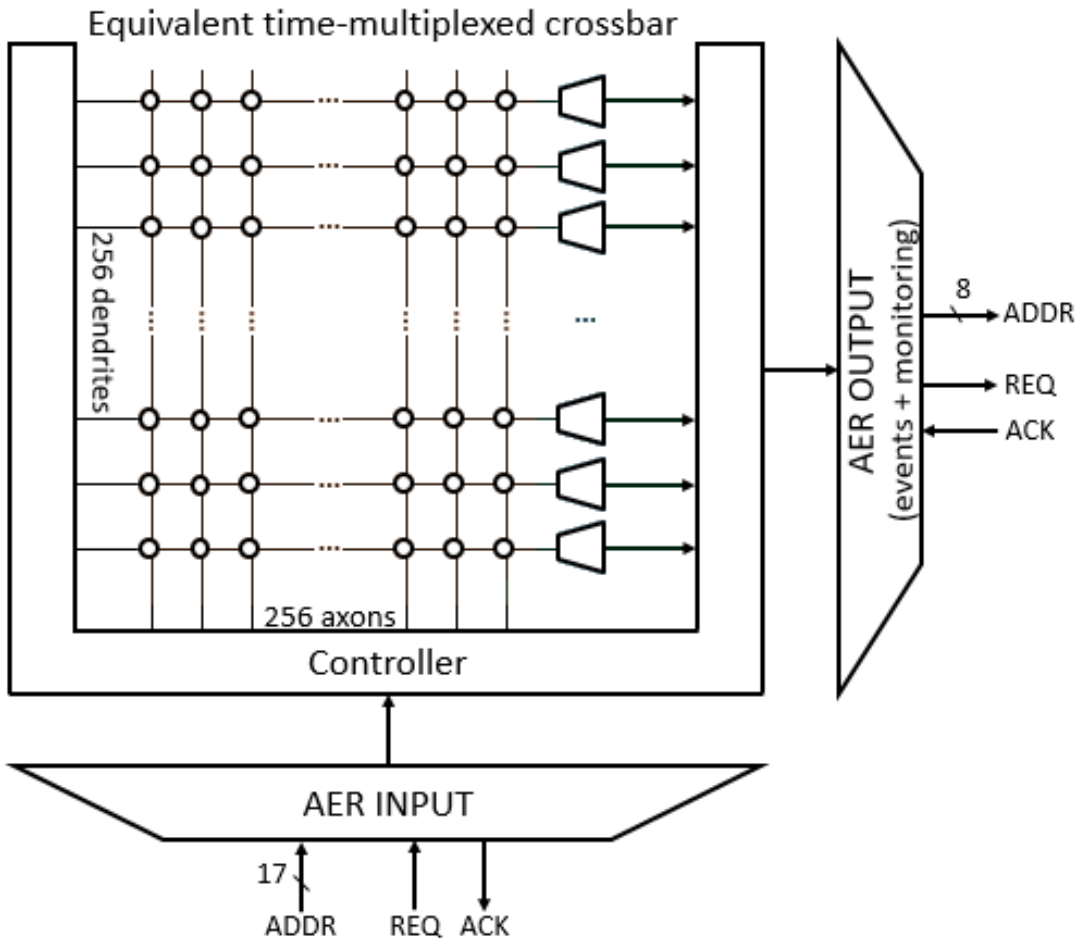
Biophysical behavior

emulation ~~simulation~~

	Analog	Digital
Biophysical versatility		
Area		
		phenomenological modeling
	Good	Bad
Design time	High	Low
Technology scaling	Low potential	High potential
Noise, mismatch, PVT	Bug or feature?	Low sensitivity

The first neuromorphic hardware?

Overview of ODIN: An experimentation platform for the edge



The basis: a nice (not-so-interesting) crossbar core

- 256 neurons
- 64k synapses

The rationale: simplicity and efficiency

- Time multiplexing
- No PDE solvers (phenomenological modelling)
- Full space and time locality

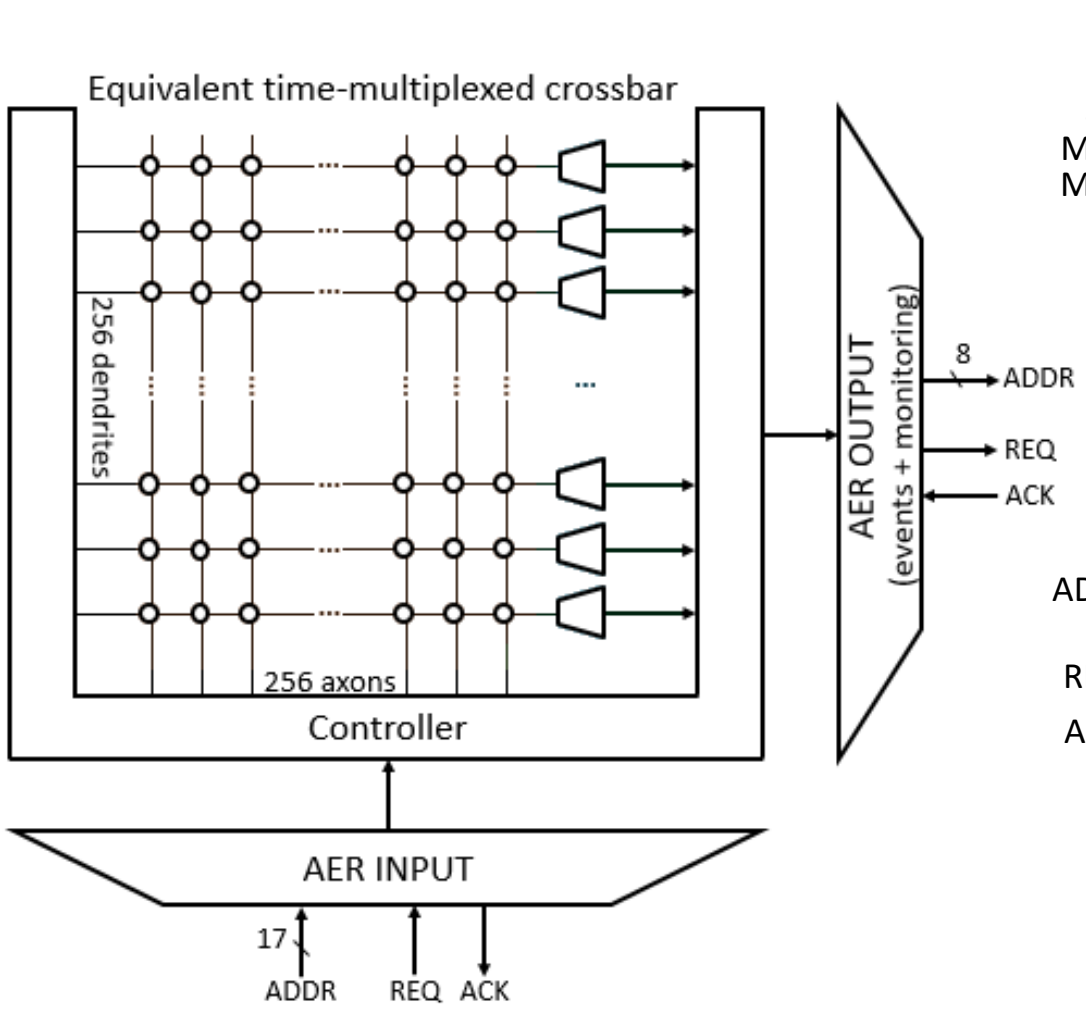
The salt and pepper: an experimentation platform

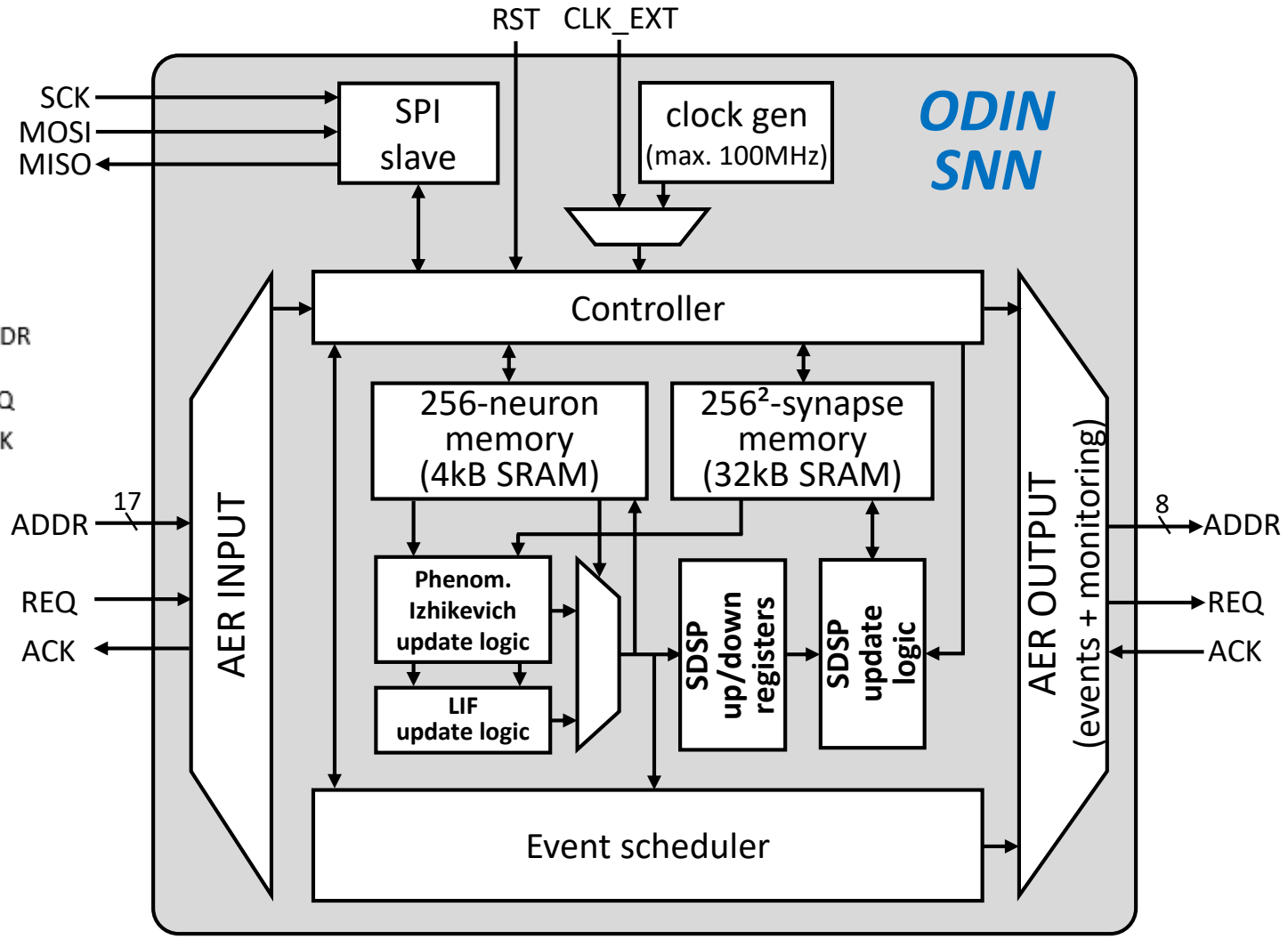
- Synaptic plasticity (SDSP)
- Large neuronal repertoire (LIF + Izhikevich)

[Frenkel, *Trans. BioCAS*, 2019a]

The first neuromorphic hardware?

Overview of ODIN: An experimentation platform for the edge

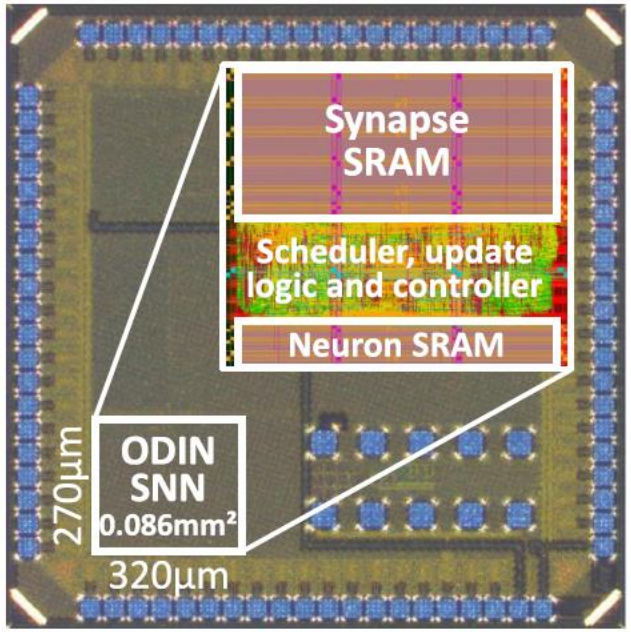




[Frenkel, *Trans. BioCAS*, 2019a]

The first neuromorphic hardware?

ODIN: The silicon demonstration



https://github.com/ChFrenkel/ODIN/

Technology	28nm FDSOI	
Implementation	Digital	
Area	0.086mm²	Record density
# neurons	256	
# synapses	64k	
# Izhikevich behav.	20	Large feature set
Online learning	SDSP, (3+1)-bit weight	
Time constant	Biological to accelerated	
Supply voltage	0.55V – 1.0V	
Leakage power (P_{leak})	27.3µW @0.55V	
Idle power (P_{idle})	1.78µW/MHz @0.55V	Lowest energy/SOP among digital designs at the time of publication, competitive with mixed-signal
Incr. energy/SOP (E_{SOP})	8.43pJ @0.55V	
Global energy/SOP ($E_{tot.SOP}$)	>12.7pJ @0.55V	
Routing flexibility/efficiency	☹ (AER)	
Fan-in	256	
Fan-out	256	

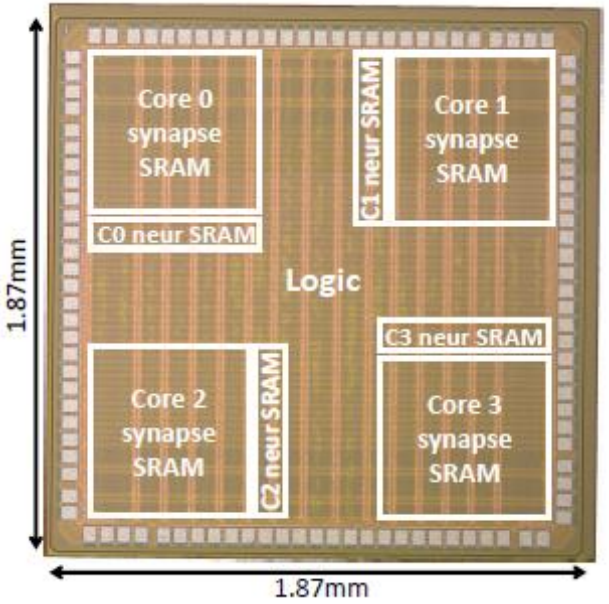
Today: 2-5pJ/SOP

[Frenkel, *Trans. BioCAS*, 2019a]

The first neuromorphic hardware?

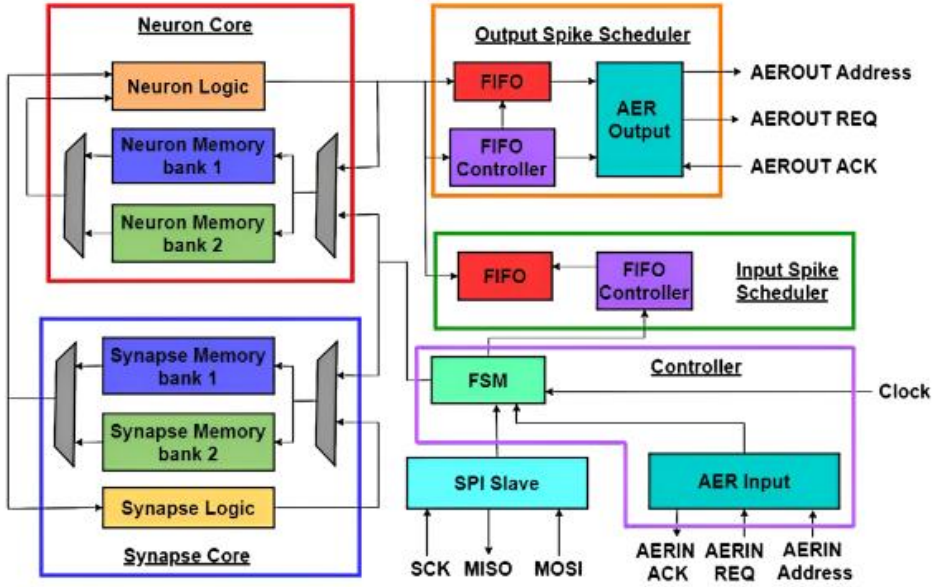
ODIN: the offspring

MorphIC



[Frenkel, *Trans. BioCAS*, 2019b]

THOR



[Senapati, arXiv:2212.01696, 2022]

tinyODIN

The basis: a nice (not-so-interesting) crossbar core

- 256 neurons
- 64k synapses

The rationale: simplicity and efficiency

- Time multiplexing
- No PDE solvers (phenomenological modelling)
- Full space and time locality

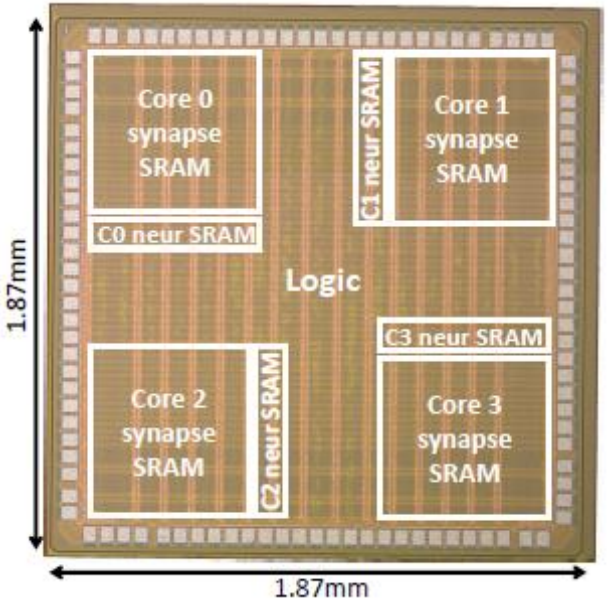
The salt and pepper: an experimentation platform

- Synaptic plasticity (SDSP)
- Large neuronal repertoire (LIF + Izhikevich)

The first neuromorphic hardware?

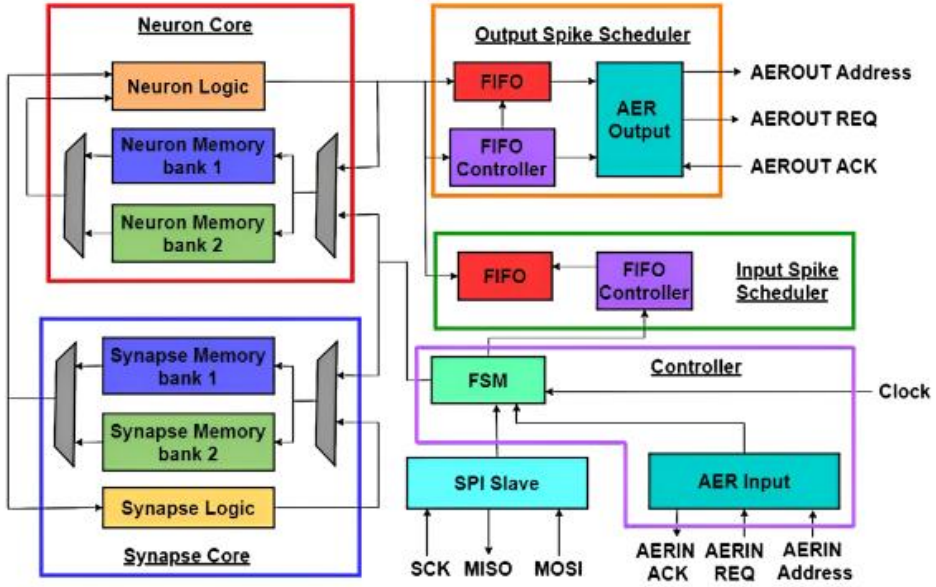
ODIN: the offspring

MorphIC



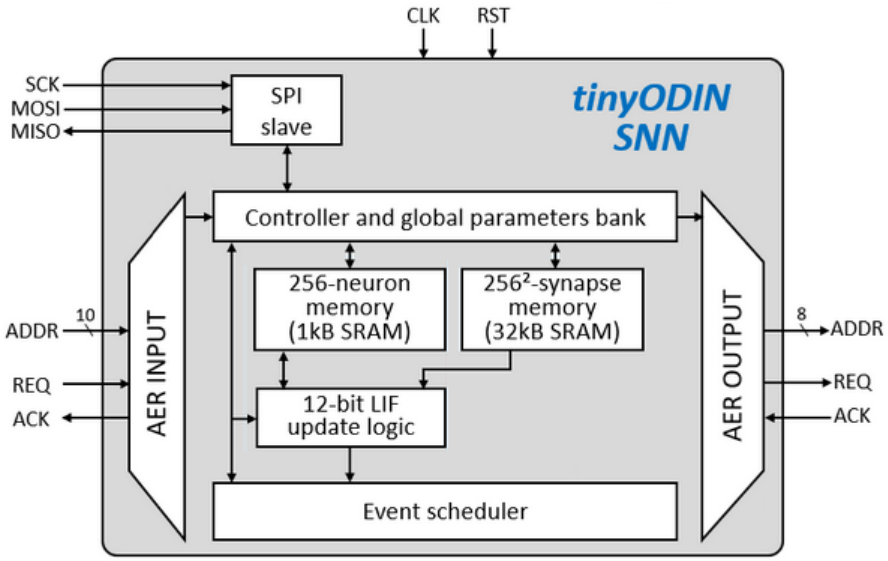
[Frenkel, *Trans. BioCAS*, 2019b]

THOR



[Senapati, arXiv:2212.01696, 2022]

tinyODIN



To be released soon!

Why should we open neuromorphic hardware?

(Just in case you still have doubts!)

Neuromorphic hardware is not (yet) as available as the Raspberry Pi

There's a strong need for flexible testbeds for new experiments!

Opening neuromorphic hardware is good for the community

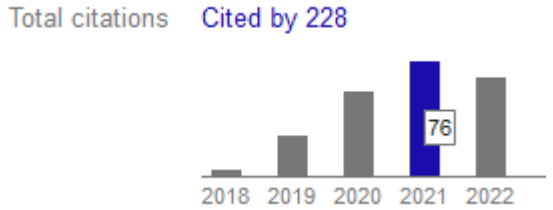
Remember this time you had to redesign published work from scratch?

Open results = reproducible results.

You can accelerate new progress in the field!

Opening neuromorphic hardware is good for you!

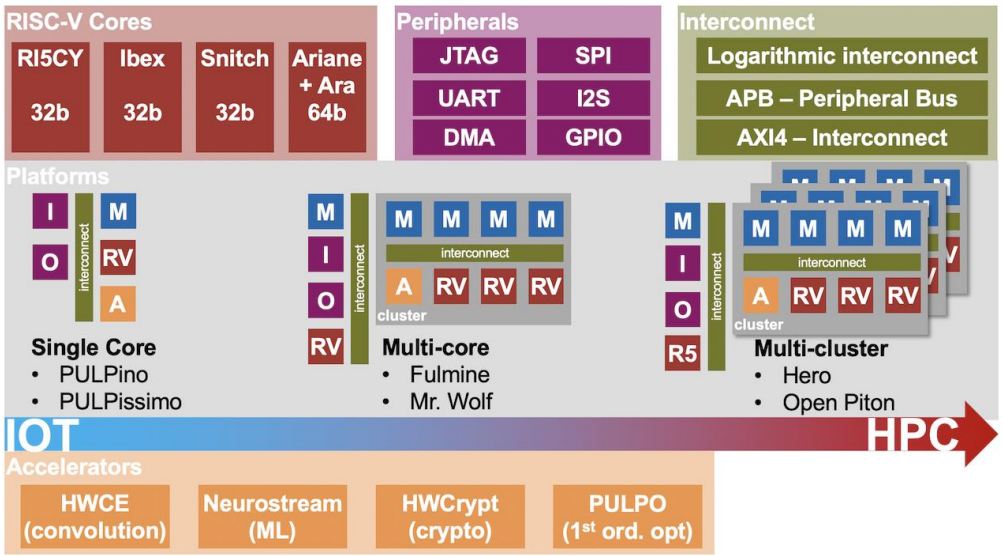
Opening your developments is good for your citations. ☺



How should we open the hardware as well?

From full open-source ecosystems to licenses

Take a (humbling) inspiration from the champions: the PULP platform



https://pulp-platform.org/

Which license?

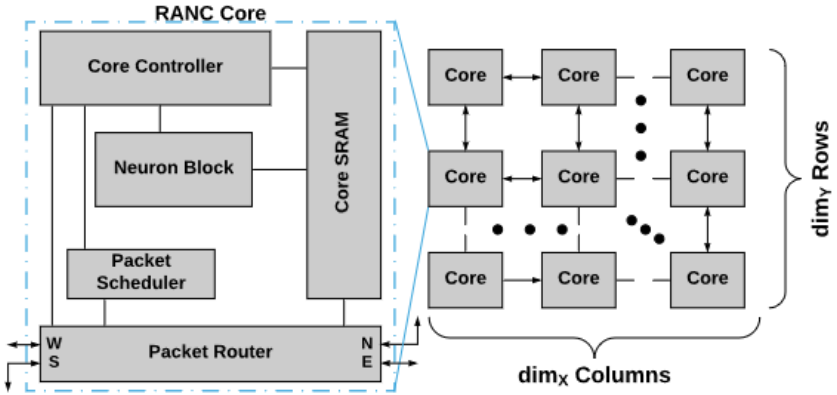
For a permissive license, the *Solderpad Hardware License* is the way to go!
(Although in theory, the *Apache 2.0* license should be applicable to hardware too)
https://solderpad.org/

For a copyleft use, see also the *CERN Open Hardware License (OHL)*.
https://ohwr.org/cernohl/

Well, this is not all about ODIN either!

Large-scale, flexible and scalable, or tiny and adaptive?

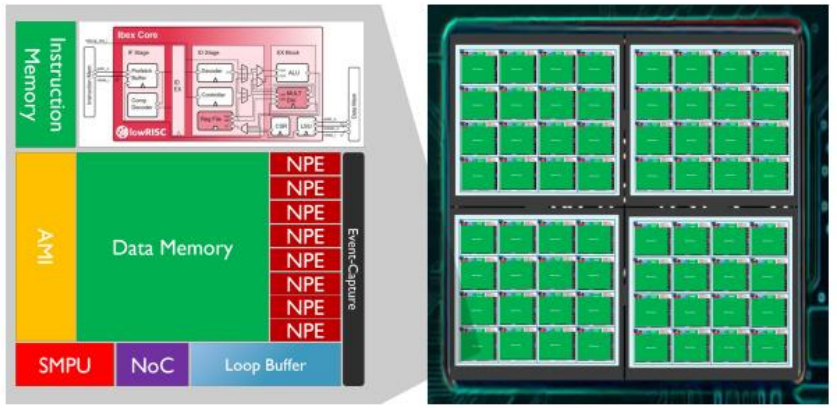
RANC



[Mack, *Trans. CAD*, 2020]

https://github.com/UA-RCL/RANC

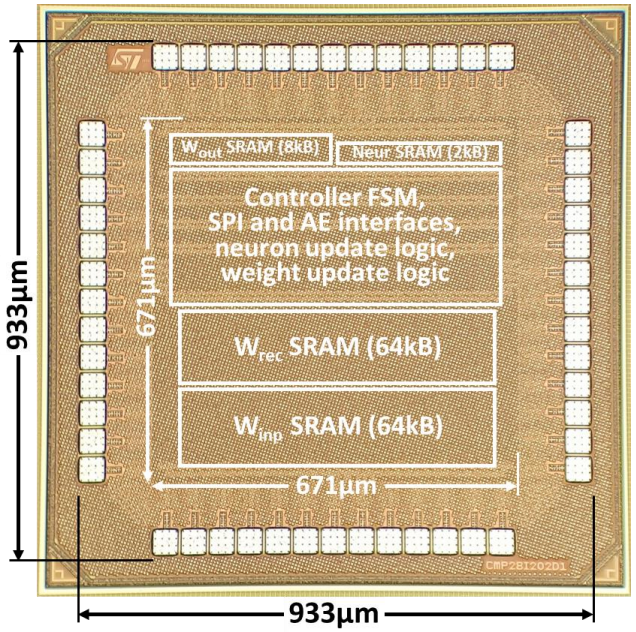
SENeCA



[Yousefzadeh, *AICAS*, 2022]

…Soon-to-be open-source. ☺

ReckOn

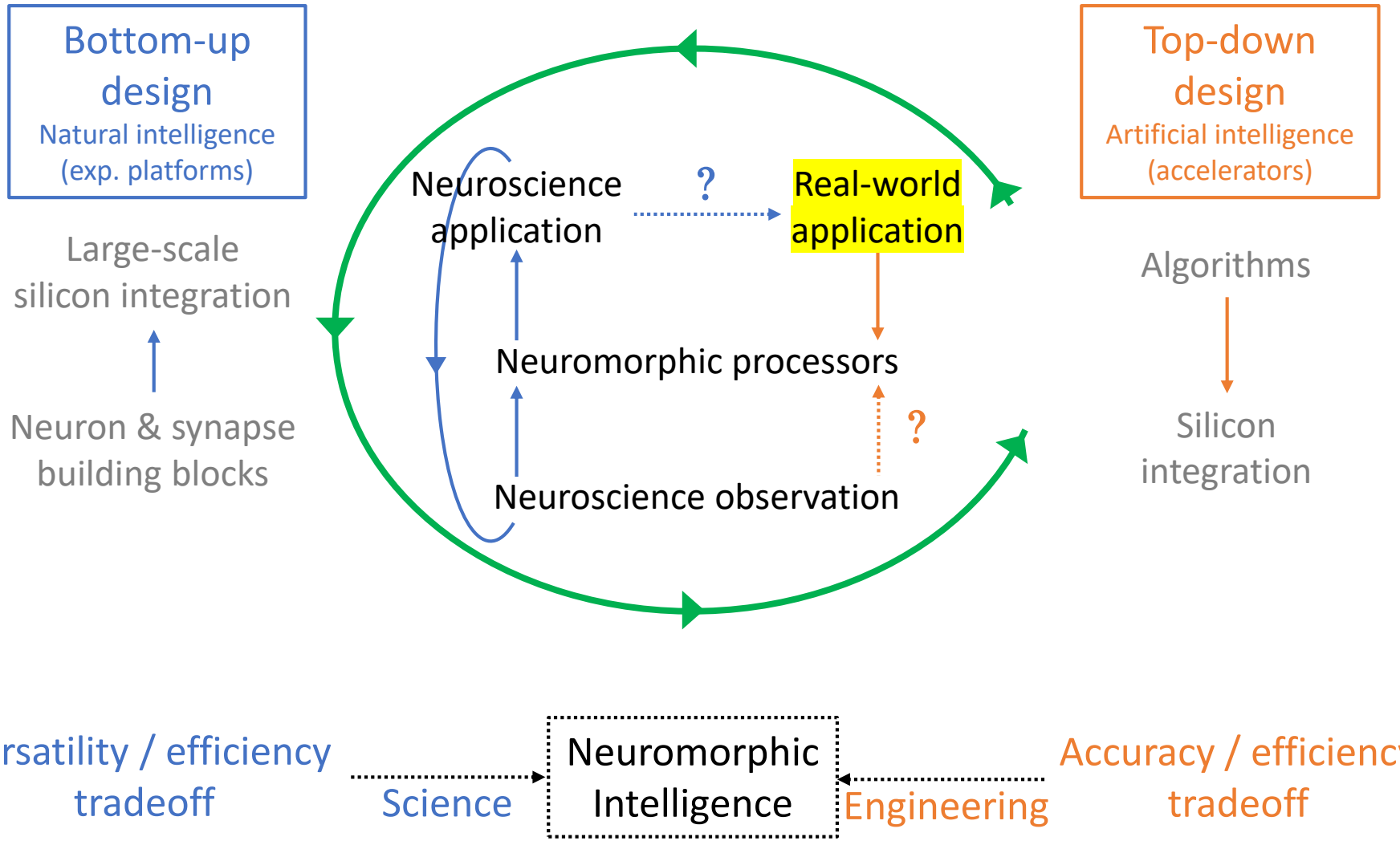


[Frenkel, *ISSCC*, 2022]

https://github.com/ChFrenkel/ReckON

Neuromorphic Engineering – How?

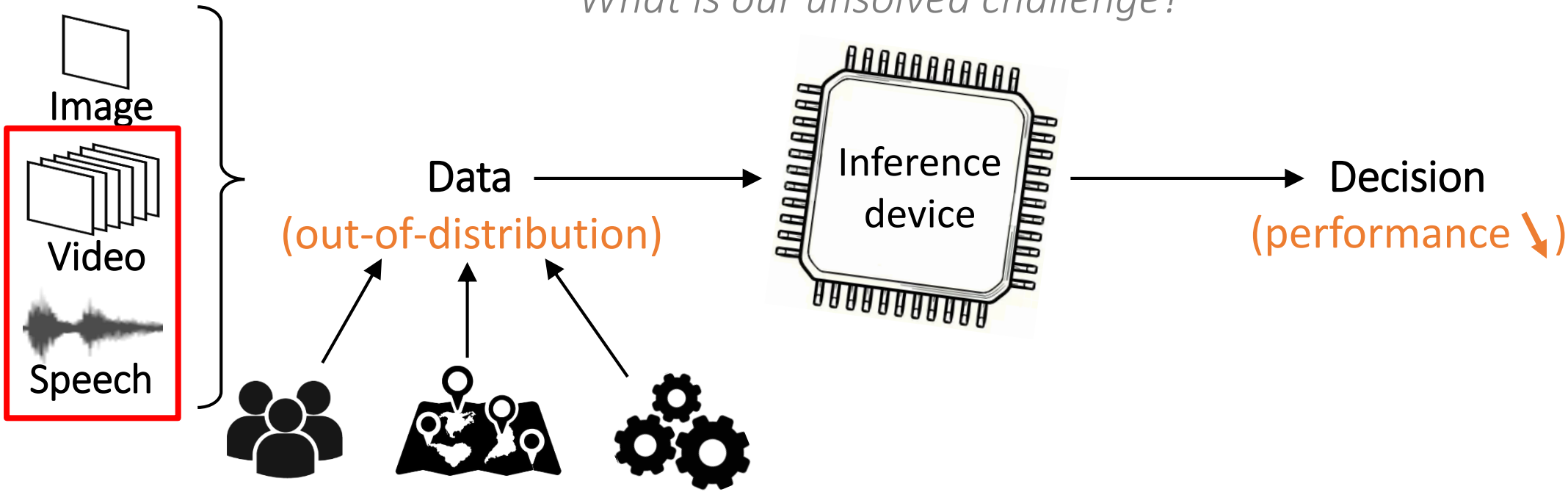
Which design strategy toward neuromorphic intelligence?



[Frenkel, *arXiv*, 2021]

The real world application

What is our unsolved challenge?



Different users, environments, task requirements

More training data before deployment?

Issues: cost, robustness, flexibility

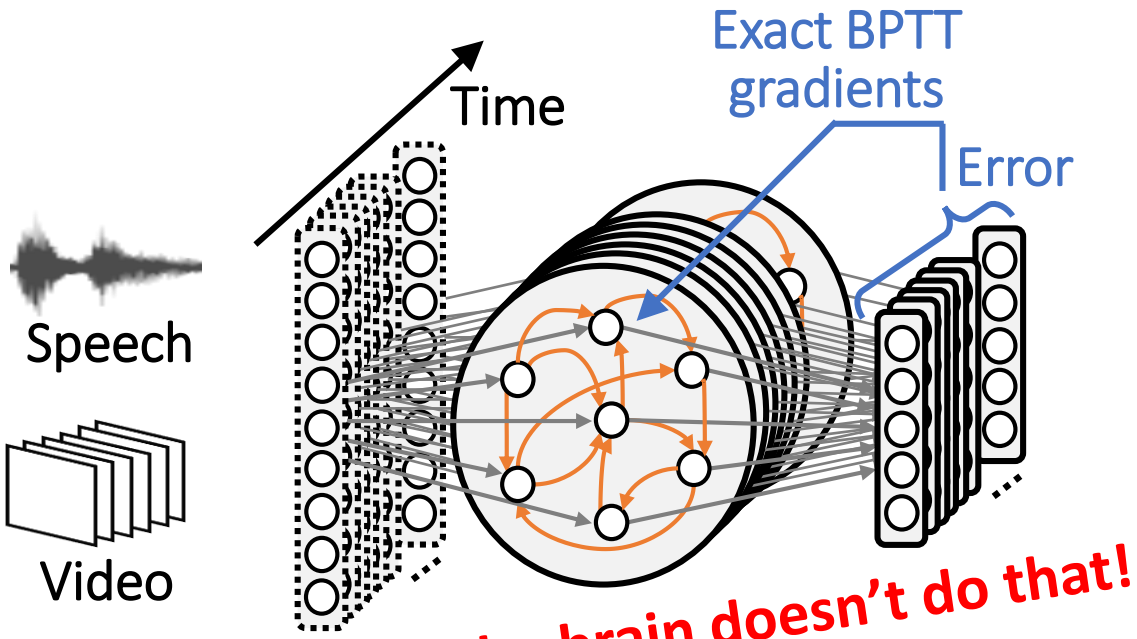
Data exchange with the cloud?

Issues: power budget, privacy

On-chip training (end-to-end)

The real world application

What is our unsolved challenge?



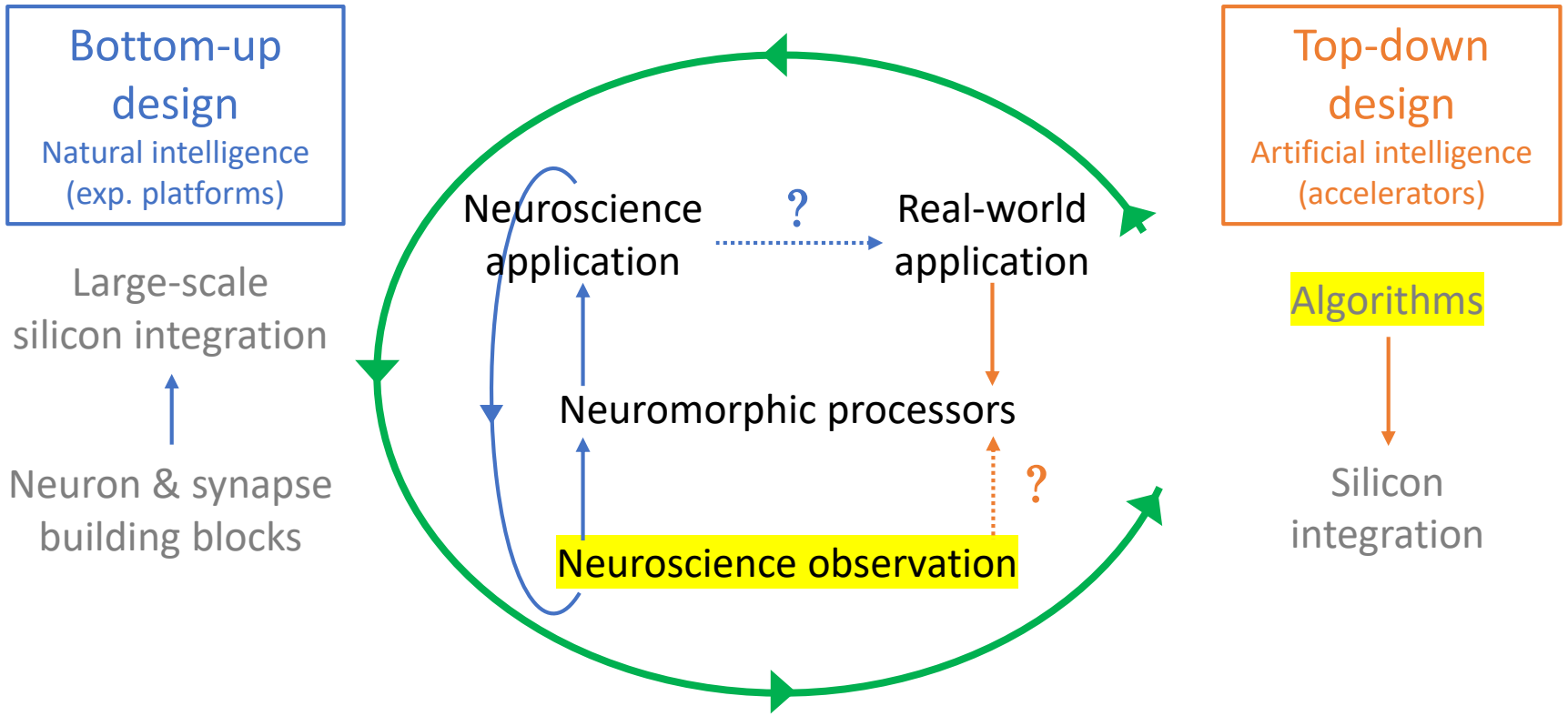
- Unrolling in time: very deep network
  (current learning ICs for static stimuli: ≤3 layers)
- Intractable memory/latency requirements
- No end-to-end on-chip solution to date

Hint: Obviously, the brain doesn't do that!

Key challenge: On-chip learning over long timescales while keeping a fine-grained temporal resolution

Neuromorphic Engineering – How?

Which design strategy toward neuromorphic intelligence?



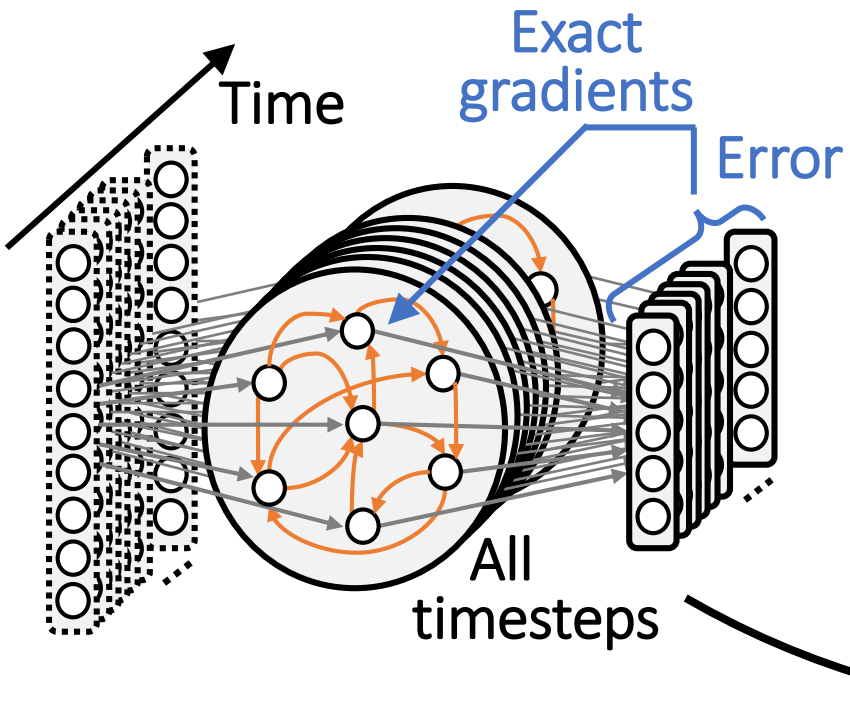
Versatility / efficiency tradeoff — Science → Neuromorphic Intelligence ← Engineering — Accuracy / efficiency tradeoff

[Frenkel, *arXiv*, 2021]

What does neuroscience tell us?

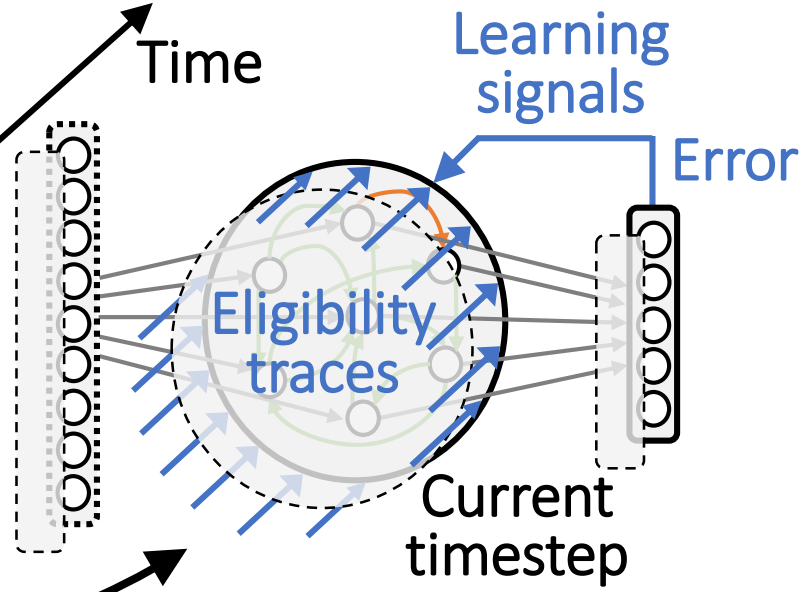
Synaptic eligibility traces are worth looking into!

Backprop through time (BPTT, backward)



Eligibility propagation (e-prop, forward)

[Bellec, *Nat. Comms*'20]



$$\frac{dE}{dW_{ij}} \simeq \sum_t L_j^t e_{ji}^t$$

e-prop

HW/algo co-design: Local decoupling of space and time

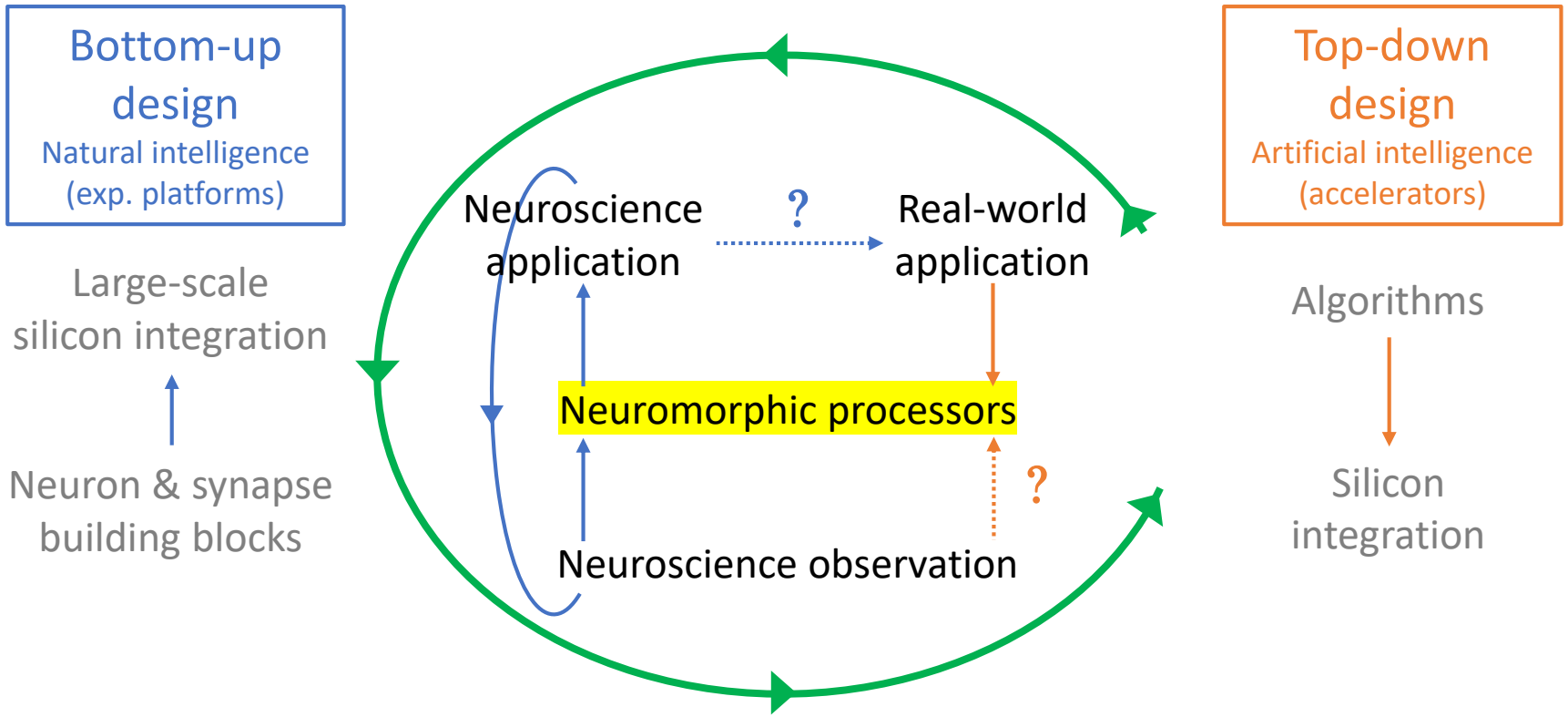
Biological plausibility ↗

Space and time locality ↗

On-chip memory requirements ↘

Neuromorphic Engineering – How?

Which design strategy toward neuromorphic intelligence?

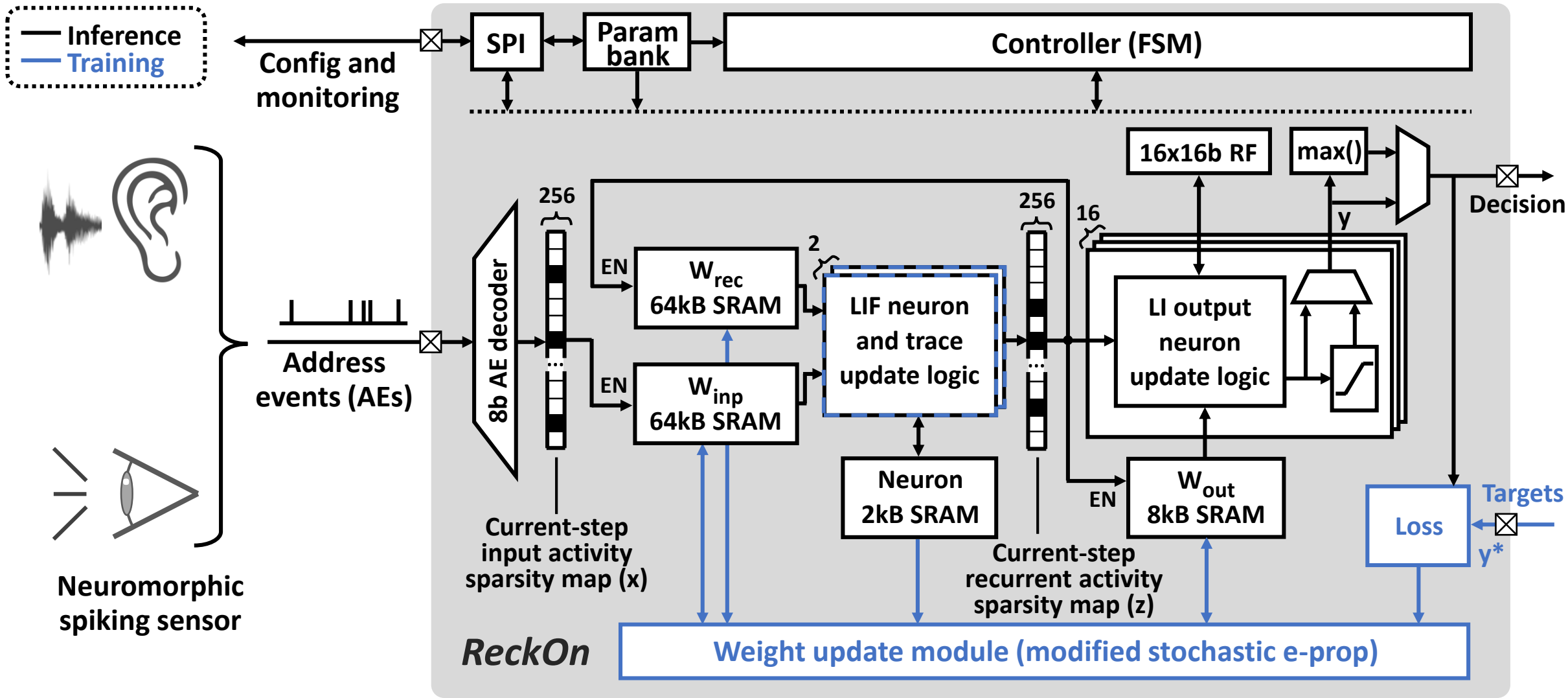


Versatility / efficiency tradeoff — Science → Neuromorphic Intelligence ← Engineering — Accuracy / efficiency tradeoff

[Frenkel, *arXiv*, 2021]

The ReckOn open-source neuromorphic processor

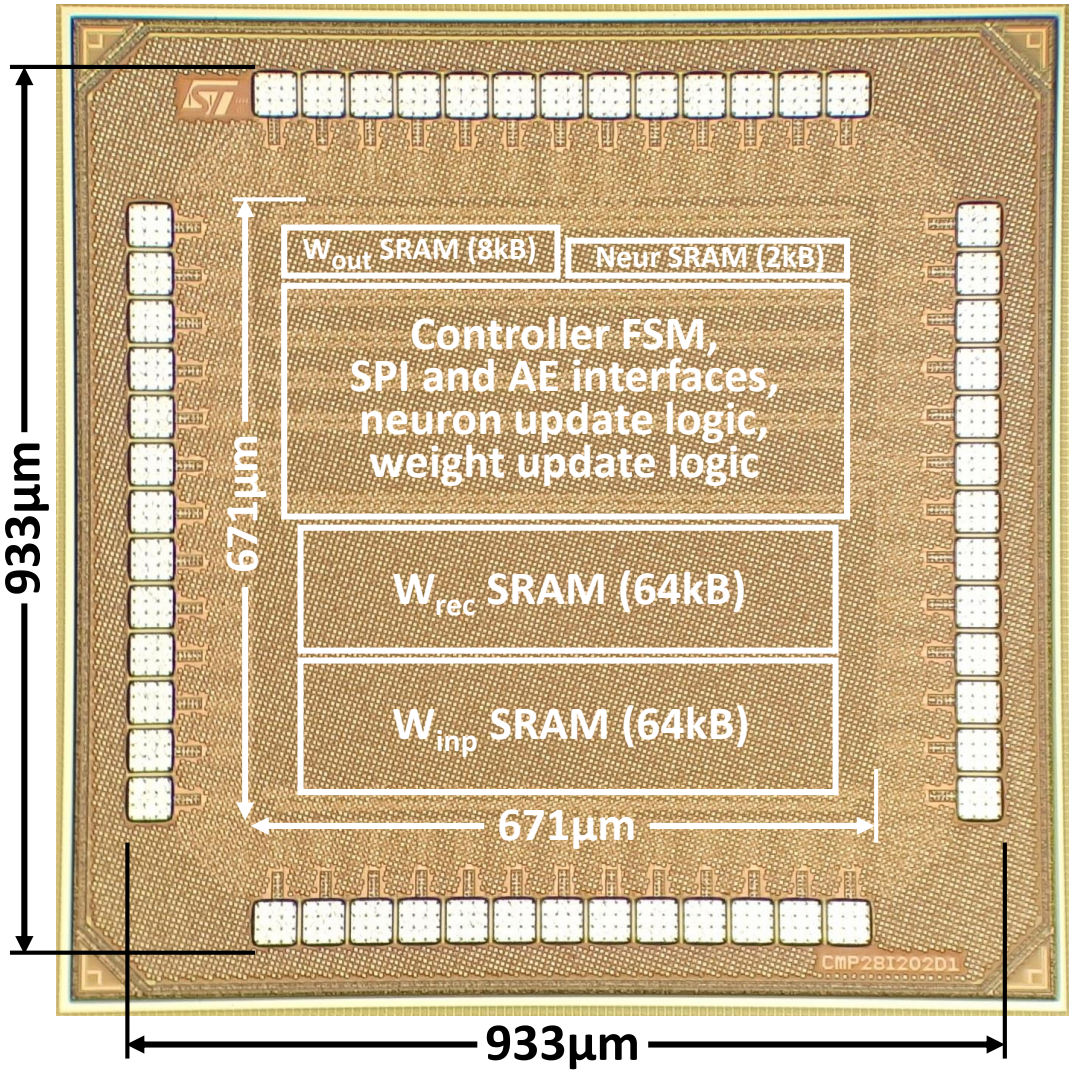
System diagram



[Frenkel, *ISSCC*, 2022]

The ReckOn open-source neuromorphic processor

Chip microphotograph and summary

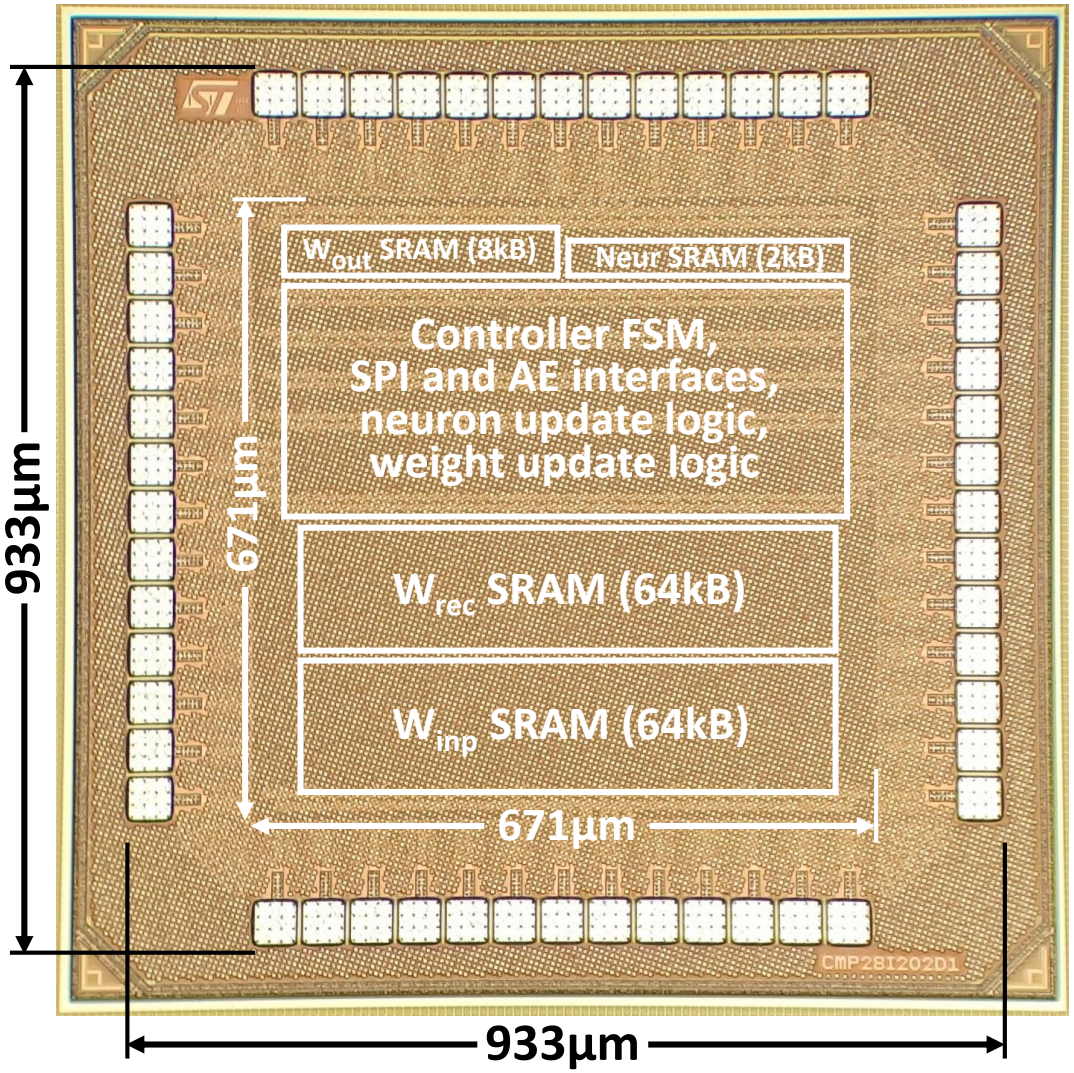


Technology	28nm FDSOI CMOS
Core size	0.67 x 0.67 mm² **0.45mm²**
Die size	0.93 x 0.93 mm²
SRAM	138kB **+ 0kB ext. DRAM!**
Network	Spiking RNN
Training timespan	Max. 32k steps

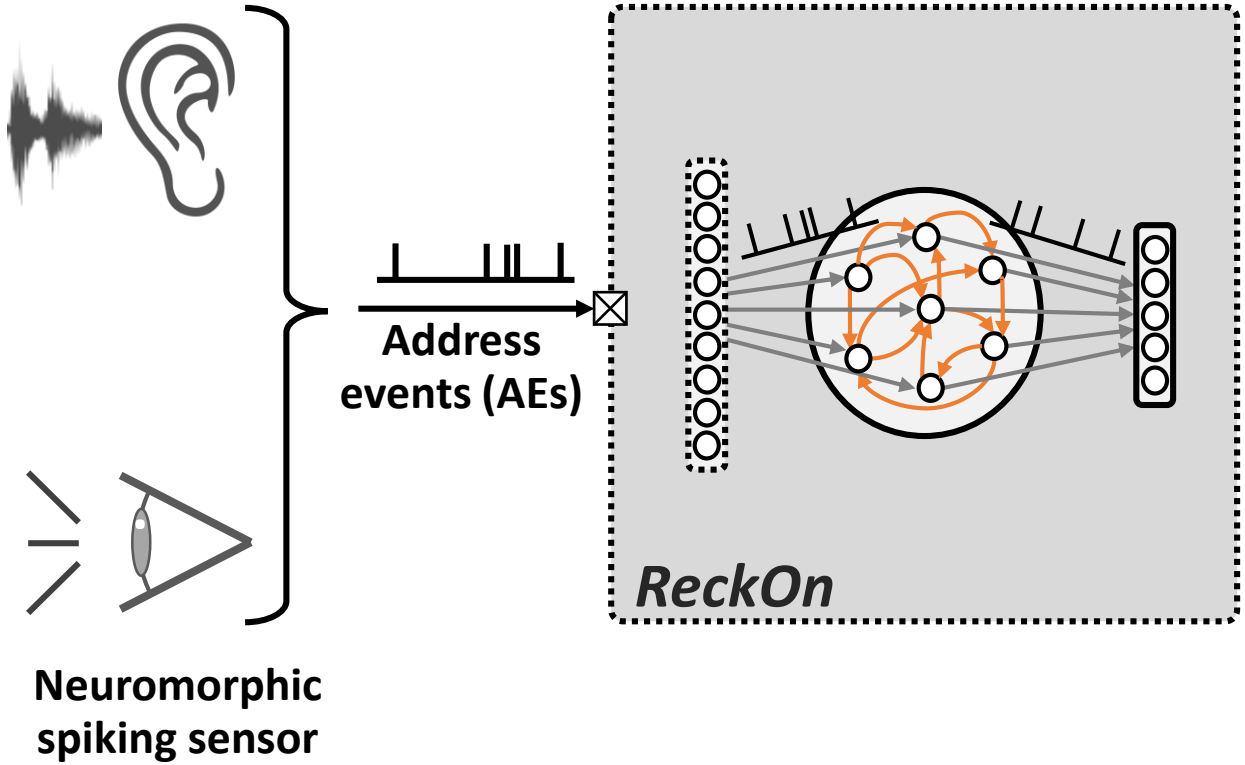
[Frenkel, *ISSCC*, 2022]

The ReckOn open-source neuromorphic processor

Chip microphotograph and summary



- Event-driven / sparsity-aware computation
- Sensor-agnostic raw-data processing
- Task-agnostic processing and learning



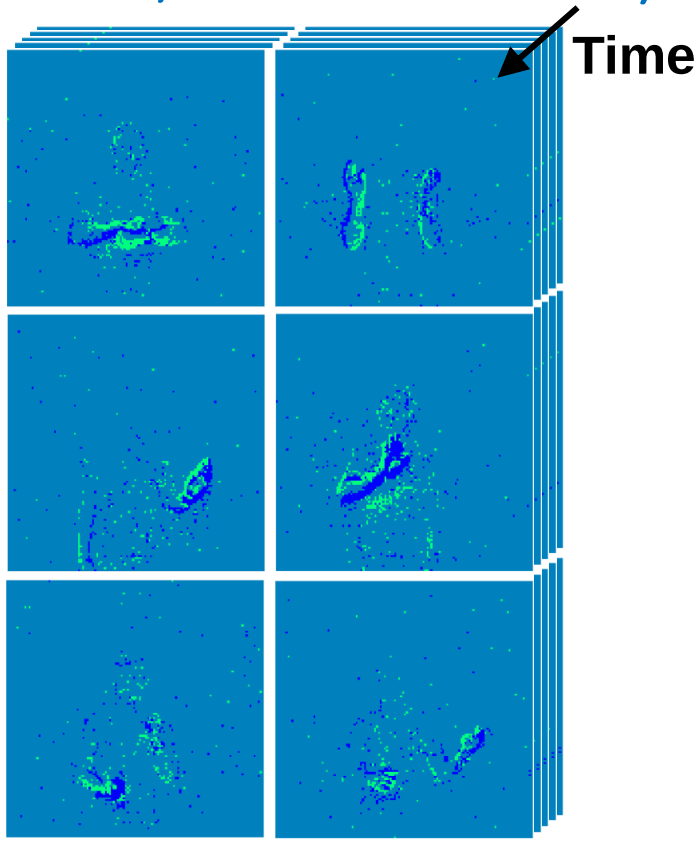
[Frenkel, *ISSCC*, 2022]

The ReckOn open-source neuromorphic processor

Benchmarking for task-agnostic processing

Vision

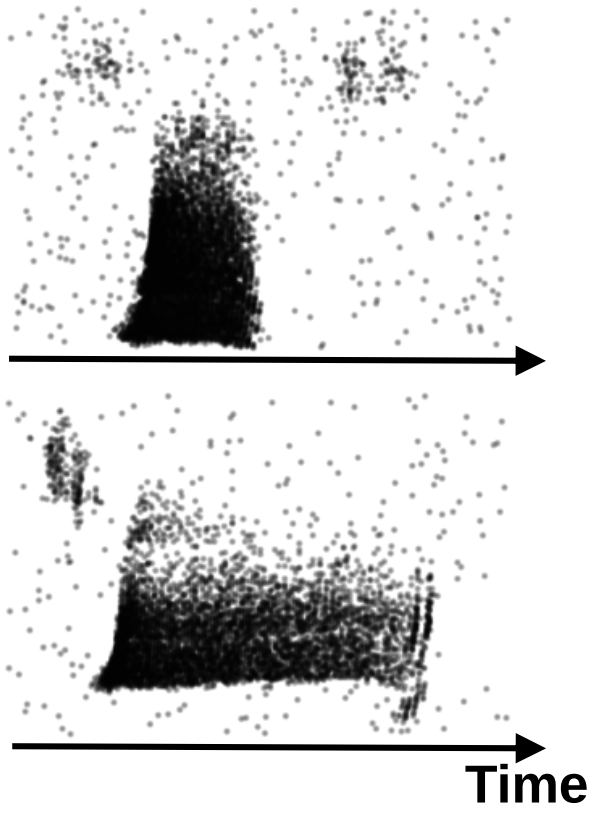
IBM DVS Gestures dataset
(10 classes, shrinked to 16x16)



Accuracy: 87.3% (28µW @0.5V)

Audition

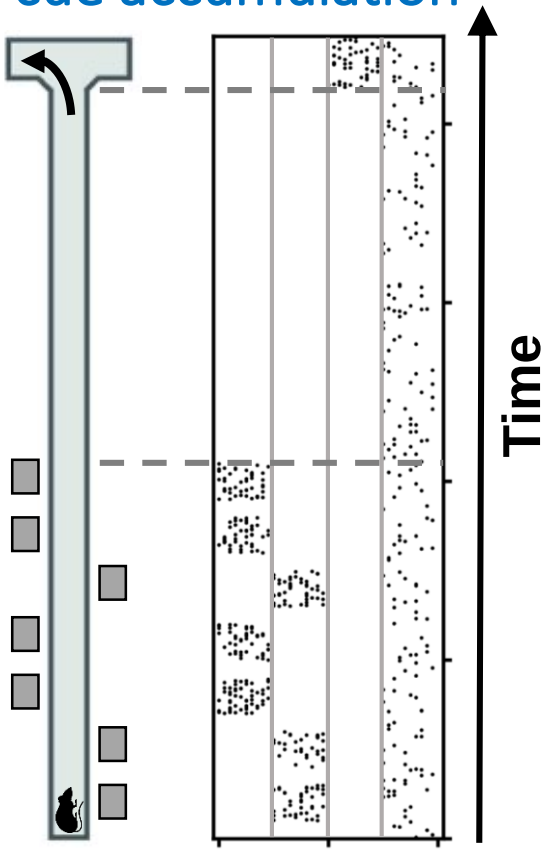
Spiking Heidelberg Digits (EN) dataset
(1-word KWS, 1:1 target vs. filler, 1:3 sub)



Accuracy: 90.7% (46µW @0.5V)

Navigation

Delayed-supervision
cue accumulation



Accuracy: 96.4% (14µW @0.5V)

[Frenkel, *ISSCC*, 2022]

Conclusions and Perspectives

- The neuromorphic community is enthusiastic and young: open-source developments are especially welcome!
- Open neuromorphic HW today is still mostly digital. Open PDKs now outline open-sourced mixed-signal neuromorphic HW as well.
- We have most pieces of the puzzle. How/when will we have a full neuromorphic ecosystem (like the PULP platform)?

Questions?

@C_Frenkel
Charlotte-Frenkel
cfrenkel
c.frenkel@tudelft.nl
ChFrenkel
chfrenkel.github.io

- (tiny)ODIN:	C. Frenkel et al., "A 0.086-mm² 12.7-pJ/SOP 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28nm CMOS," *IEEE Trans. BioCAS*, 2019.	*Open-sourced!* github.com/ChFrenkel/ODIN, *with tinyODIN to be released soon.*
- MorphIC:	C. Frenkel et al., "MorphIC: A 65-nm 738k-synapse/mm² quad-core binary-weight digital neuromorphic processor with stochastic spike-driven online learning," *IEEE Trans. BioCAS*, 2019.	
- THOR:	M. Senapati et al., "Thor – A neuromorphic processor with 7.29 G TSOP²/mm²Js energy-throughput efficiency," *arXiv preprint arXiv:2212.01696*, 2022.	*Soon-to-be open-sourced*
- RANC:	J. Mack et al., "Ranc: Reconfigurable architecture for neuromorphic computing," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 2020.	*Open-sourced!* github.com/UA-RCL/RANC
- SENeCA:	A. Yousefzadeh et al., "SENeCA: Scalable energy-efficient neuromorphic computer architecture," *IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2022.	*Soon-to-be open-sourced*
- ReckOn:	C. Frenkel and G. Indiveri, "ReckOn: A 28-nm sub-mm² task-agnostic spiking recurrent neural network processor enabling on-chip learning over second-long timescales", *IEEE International Solid-State Circuits Conference (ISSCC), 2022.*	*Open-sourced!* github.com/ChFrenkel/ReckOn
- Review:	C. Frenkel, D. Bol and G. Indiveri, "Bottom-up and top-down neural processing systems design: Neuromorphic intelligence as the convergence of natural and artificial intelligence", *arXiv preprint arXiv:2106.01288, 2021.*	