

Synergizing AI and CPU: Empowering Next-Generation Computing

A PROJECT REPORT

Submitted by:

21BCS6274 – Karthikeya Tallapaneni

21BCS6298 – Harshit Bharadwaj

21BCS6304 – Amar Kajla

Under The Guidance of:

Dr. Priyanka Kaushik

In Partial Fulfilment for award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE WITH SPECIALIZATION IN
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING



CHANDIGARH UNIVERSITY



ACKNOWLEDGMENT

It gives us the privilege to complete this mini-project. This is the only page where we have the opportunity to express my emotions and gratitude. It is a great pleasure in expressing sincere and deep gratitude towards my supervisor and guide Ms. Priyanka Kaushik for her valuable suggestions, guidance, and constant support throughout the completion of this project named “Neural Style Transfer”. This project, though done by us, wouldn't be possible without the support of varied people, who by their cooperation have helped us in bringing out this project successfully. We are really very thankful to Chandigarh University for providing me such a great opportunity to make such a wonderful project which can solve real-life problems and extremely valuable handsome experience along with crucial soft skills such as working in a team, communication skills, and much more. I also offer my most sincere thanks to every team member of our group who was working rigorously on this project and staff members of the Artificial Intelligence and Machine Learning department, Apex Institute of Technology, Chandigarh University for cooperation provided by them in every possible way. We thank all the faculty members and other supporting staff for the help they provided to us for the completion of our project.

Karthikeya Tallapaneni(21BCS6274)

Harshit Bharadwaj(21BCS6298)

Amar kajla(21BCS6403)

Submitted by the students of 6th Semester

Student Signatures

CONTENTS

1.	INTRODUCTION	2
2.	ARTIFICIAL NEURAL NETWORKS.....	3
3.	NEUROMORPHIC HARDWARE	6
3.1	Advantages of Spiking Neural Networks.....	9
3.2	Neuromorphic hardware implementations	11
3.3	Neuromorphic hardware leaders.....	12
3.4	Neuromorphic hardware research chips	14
3.5	Neuromorphic hardware start-ups	18
4	SOME APPLICATION AREAS	20
4.1	Mobile Applications	21
4.2	Adaptive Robotics	22
4.3	Event-Based Vision Sensors	23
5	MARKET TRENDS OF AI CHIPS	25
5.1	Cloud/Datacenter.....	27
5.2	Edge Computing	28
5.3	Future Market of Neuromorphic Chips.....	31
6	CONCLUSION	32
7	REFERENCES.....	33
8	ANNEX 1.....	41
9	ANNEX 2.....	50

1. INTRODUCTION

The Human Brain Project (HBP) (<https://www.humanbrainproject.eu/en/>) is a highly multifaceted and multidisciplinary project where various research groups – mostly neuroscientists and computing scientists - work together to understand the multi-scaled anatomy and functioning of the human brain and accordingly replicate those mechanisms *in silico*.

This report has been developed by the Innovation group of the HBP. The group, which is part of the HBP Management and Coordination subproject (SP11) and based in the Universidad Politécnica de Madrid (UPM), provides innovation management support and research commercialization advice to the HBP's partners. Among other activities, this transversal support includes the elaboration of specific technology exploitation plans, assistance on IP aspects, training on the assessment of technologies maturity, evaluation of innovation potentials, the matching of available and mature HBP results to (academic and non-academic) users' needs, and the involvement of external players, specially industrial players, in the utilization and co-creation of the developing HBP infrastructure (EBRAINS, <https://ebrains.eu/>).

Frequently, researchers and developers' cannot devote too many resources and time to devise new forms of technological innovation or plan their own protection and exploitation strategies. Their expertise and networking potential is critical, however, to increase the possibilities of an eventual and successful commercialization and utilization of their results. It seems useful therefore to provide researchers with a big picture of existing initiatives, key actors, market insights and expected trends in their area of work, thus enabling them to identify more easily strategic opportunities (e.g. those leading to obtain further research funding) and get the most out of their valuable specialized knowledge. This would contribute to equip them with enough skills and means to exploit their results, define their own roadmaps to the market and assess the possibilities of using other present and upcoming tools on their own research processes.

Based on this rationale, this report presents an updated and concise overview of a particular area of Neurocomputing hardware, the market of Neuromorphic chips¹. The document is halfway between a complete research work and a full market analysis. The study has also tried to adopt a foresight perspective that helps to draft some basic technology roadmaps and envision some future trends.

One challenging aspect of the study has been to simultaneously capture the attention of neuromorphic experts while also trying to raise the curiosity and interest of a broader audience. In this sense, initiated readers and leading computing scientists may hopefully find the report motivating enough to reflect on the pivotal parts of the Neuromorphic science around which useful, practical and innovative applications could happen.

¹ This report will be regularly updated and refined during next phase of the HBP

2. ARTIFICIAL NEURAL NETWORKS

An Artificial Neural Network (ANN) is a combination and collection of nodes that are inspired by the biological human brain. The objective of ANN is to perform cognitive functions such as problem-solving and machine learning. The mathematical models of the ANN have been started in the 1940s however, it was in silent for a long time (Maass, 1997). Nowadays, ANNs became very popular with the success of ImageNet² in 2009 (Hongming, et al., 2018). The reason behind this is the developments in ANN models and hardware systems that can handle and implement these models. (Sugiarto & Pasila, 2018)

The ANNs can be separated into three generations based on their computational units and performance (Figure 1).

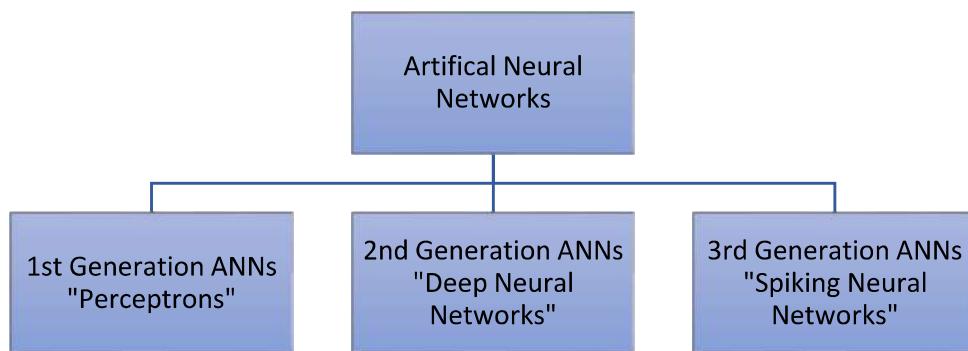


Figure 1- Generations of Artificial Neural Networks

The first generation of the ANNs has started in 1943 with the work of Mc-Culloch and Pitts (Sugiarto & Pasila, 2018). Their work was based on a computational model for neural networks where each neuron is called “perceptron”. Their model later was improved with extra hidden layers (Multi-Layer Perceptron) for better accuracy - called MADALINE - by Widrow and his students in the 1960s (Widrow & Lehr, 1990). However, the first generation ANNs were far from biological models and were just giving digital outputs. Basically, they were decision trees based on *if and else* conditions.

The Second generation of ANNs contributed to the previous generation by applying functions into the decision trees of the first-generation models. The functions work among each visible and hidden layers of perceptron and create the structure called “deep neural networks”. (Patterson, 2012; Camuñas-Mesa, et al., 2019) Therefore, second-generation models are closer to biological

² ImageNet: ImageNet is an organisation that published a “Large-Scale Hierarchical Image Database” in 2009. They conduct a large-scale visual recognition challenge since 2010. This competition gave birth to the AlexNet neural network model which achieved the challenge with only 15.3% error in 2012. (Krizhevsky, et al., 2017) Since then, AI become an area of interest and Deep learning become primary method for solving cognitive tasks; such as image, audio recognition. (Zheng & Mazumder, 2020)

neural networks. The functions of the second-generation models are still an active area of research and the existing models are in great demand from markets and science. Most of the current developments about artificial intelligence (AI) are based on these second-generation models and they have proven their accuracy in cognitive processes. (Zheng & Mazumder, 2020)

Some of the common second-generation models that are used are (Table 1) (Kennis, 2019);

Neural Network	Description	Applications	Network Image
FNNs	Each perceptron (simplest and oldest form of neurons) in one layer is connected to every perception from the next layer. Information is fed forward from one layer to the next in the forward direction only. There are no feedback loops. Thus, the data is processed, and the results are calculated on every input sequence. This network may or may not have hidden layers.	Primarily used for animal recognition, digit recognition, cheque recognition, medical diagnosis, etc.	
RNNs	Use sequential information such as timestamped data from a sensor device or a spoken sentence, composed of a sequence of terms. Unlike FNNs, inputs to RNNs are not independent of each other, and the output for each element depends on the computations of the preceding elements.	Primarily used in forecasting and time series applications, sentiment analysis and other text applications,	
Long Short-Term Memory (LSTM)	A type of RNN that is explicitly designed to hold information for long periods of time and process the incoming data, along with the previously calculated results. LSTMs contain their information in a memory and can read, write and delete information from its memory.	Primarily used for text classification, machine translation, dialog systems, speech recognition, translating videos and images to natural languages, etc.	
CNNs	Typically contain five types of layers: input, convolution, pooling, fully connected and output (more recent versions tend to be deep with more than seven or nine layers). Each layer has a specific purpose, like summarizing, connecting or activating.	Primarily used for image classification and object detection. Other applications include language processing, computer vision and video analytics.	

Source: Medium, SAS, TMT Analytics

Table 1 Most common 2nd generation ANN models (Kennis, 2019)³

The Third generation of ANN is termed as Spiking Neural Networks (SNNs). They are biologically inspired structures where information is represented as binary events (spikes). Their learning mechanism is different from previous generations and is inspired by the principles of the brain (Kasabov, 2019). SNNs are independent of the clock-cycle based fire mechanism. They do give an output (spike) if the neurons collect enough data to surpass the internal threshold. Moreover, neuron structures can work in parallel (Sugiarto & Pasila, 2018). In theory, thanks to these two features SNNs consume less energy and work faster than second-generation ANNs (Maass, 1997).

³ FNN: Feed Forward Neural Networks; RNN: Recurrent Neural Network; CNN: Convolutional Neural Network

The advantages of SNNs over ANNs are: (Kasabov, 2019)

- Efficient modeling of temporal - spatio temporal or spectro temporal - data
- Efficient modeling of processes that involve different time scales
- Bridging higher-level functions and “lower” level genetics
- Integration of modalities, such as sound and vision, in one system
- Predictive modeling and event prediction
- Fast and massively parallel information processing
- Compact information processing
- Scalable structures (from tens to billions of spiking neurons)
- Low energy consumption, if implemented on neuromorphic platforms
- Deep learning and deep knowledge representation in brain-inspired (BI) SNN
- Enabling the development of BI-AI when using brain-inspired SNN

Although there seems to be a lot of advantages of SNNs compared to ANNs (Table 2), advances in associated microchips technology, which gradually allows scientist to implement such complex structures and discover new learning algorithms (Lee, et al., 2016) (Furber, 2016), are still very recent (after the 2010s). Spiking Neural Networks technology, with only ten-year implementation in the area, is relatively young, therefore, compared to the second generation. So, it needs to be further researched and more intensively implemented to leverage more efficiently and effectively its advantages.

Category	ANN ANN (Artificial Neural Network, Deep Learning)	SNN SNN (Spiking Neural Network Algorithm)
Neuronal Activations	Multi-level (fixed or floating point)	timing domain coded spikes (binary values)
Timing expression	Recurrent connections in RNN and other networks	Membrane potential and Recurrent connections
Spatial expression	usually a more regular interconnected neuronal array. The processing of images usually adopts sliding windows in convolution operation	non-regularly interconnected neurons. Generally, there is no sliding window process (requiring parallel expanding of convolutions).
Activation function	usually nonlinear activation	No activation function
Reasoning	Convolution, pooling, multilayer perceptron model (MLP), etc.	Leaky Integrate and Fire model (LIF), etc.
Training	Back-propagation is more popular	STDP, Hebb's law, back-propagation
Normalization method	Batch normalization, etc.	Winner takes all
Represent negative neuronal values	Negative activation value	inhibitory neurons
Typical Sensor	Digital Camera, Microphone	DVS Camera
Theoretical sources	Mathematical derivation	Brain enlightenment
Common point		integration process, MLP topology

Table 2 ANN-SNN Comparison Table (Tsinghua University, 2018)⁴

⁴ Hebb's Law: Canadian Neuropsychologist Donald Hebb's rule of learning; "*In a sense, then, cells that fire together wire together*" (Zheng & Mazumder, 2020)

Large scale SNNs can be implemented both in brain simulator software like “NEST” with high-performance computing or in Neuromorphic chips which are inspired by SNNs (Knight & Nowotny, 2018). SNN simulations implemented on Central Processing Units (CPUs) or Graphics Processing Units (GPUs) are not well-suited to express the energy-efficiency and parallelism of the spike communication. (Knight & Nowotny, 2018) SNNs can fully show their competitive advantages of low energy consumption and massively parallel working when they are implemented on Neuromorphic chips (see section 3). Nowadays, the Neuromorphic chip sector has a huge interest around the world (Figure 23)(Figure 28) and the chips are gradually being achievable for scientific and industrial use. In parallel to the availability of chips, AI scientists are also improving and discovering new and more efficient SNNs learning mechanisms.

3. NEUROMORPHIC HARDWARE

The traditional von Neumann systems are multi-model systems consisting of three different units: processing unit, I/O unit, and storage unit. These modules communicate with each other through various logical units in a sequential way (Kasabov, et al., 2016). They are very powerful in bit-precise computing (Kasabov, 2019).

However, neural-networks are data-centric; most of the computations are based on the dataflow and the constant shuffling between the processing unit and the storage unit creates a bottleneck. Since data needs to be processed in sequential order, the bottleneck causes rigidity (Kasabov, 2019) (Zheng & Mazumder, 2020).

GPUs have massively parallel computing power compared to CPUs (Zheng & Mazumder, 2020). Therefore, they quickly become the dominant chips for implementing neural networks. Currently, data centers are mainly using millions of interconnected GPUs to give parallelism on processes, but this solution causes increased power consumption (Kasabov, et al., 2016).

GPUs have expedited deep learning research processes, support the development of algorithms and managed to enter the markets. However, future edge applications such as Robotics or autonomous cars will require more complex artificial networks working in real-time, low latency, and low energy consumption inference (Zheng & Mazumder, 2020).

STDP: Spike-Timing-Dependent Plasticity; It is based on the Hebb’s rule; STDP contributes to the rule by relating the impact of exact timing of pre and post synaptic neurons with the change in synaptic strength (Zheng & Mazumder, 2020)

DVS Camera: It is an event-based visual sensor that capture brightness changes asynchronously (Gallego, et al., 2019)

LIF: Leaky Integrate and Fire model; It is one of the most used Spiking Neuron model that represents the behavior of biological neurons (Zheng & Mazumder, 2020)**Error! No se encuentra el origen de la referencia.**

The requirement of energy-efficiency has oriented the industry to accelerators that are specially designed for deep learning such as the Application-Specific Integrated Circuits (ASICs) and the Field-Programmable Gate Arrays (FPGAs). ASICs are hard-wired chips designed for processing a specific type of application. FPGAs, on the other hand, are reconfigurable hardware to handle a variety of operations. Both solutions are more energy-efficient compared to GPUs (Figure 2).



Figure 2 Current AI-hardware solutions (Microsoft, 2020)

The ASICs are costly to design and not reconfigurable because they are hard-wired, but this hard-wired nature also contributes to their optimization. Throughout the data-flow optimization, they can perform better and more energy-efficiently than the FPGAs. Therefore, FPGAs serve as a prototype chip for further designing costly deep learning ASICs (Zheng & Mazumder, 2020).

Deep learning accelerators are energy-efficient and effective for current data sizes. However, they are still limited to the bottleneck of the architecture, i.e the internal data link between the processor and the global memory units (Kasabov, et al., 2016), as the load of the data size is increasing faster than the prediction of the Moore's Law⁵. It would be difficult for a built edge system that enables the process of these data (Pelé, 2019). Novel approaches beyond the von Neumann architecture are needed therefore to cope with the shuttling issue between memory/processor.

⁵ Moore's Law: "Gordon Moore predicted in 1965 that the number of transistors per integrated circuit chip would continue to double in each technology generation" (Lundstrom, 2003)

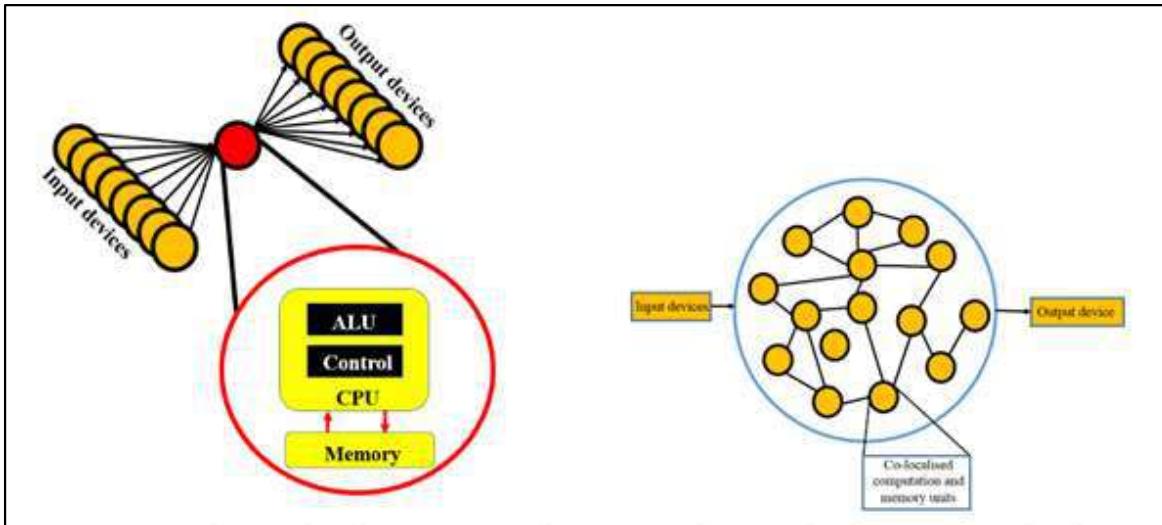


Figure 3 von Neumann vs Neuromorphic architecture

Left: The central point shows the data communication bottleneck in the von Neumann architecture (used in traditional computers).

Right: In the neuromorphic architecture the processing and memory is decentralized across a network of neuronal units (see nodes) and synapses (lines connecting the nodes), “creating a naturally parallel computing environment via the mesh-like structure” (Kasabov, et al., 2016).

As it is mentioned in the previous section, the development of SNN opens a new way of hardware architectures beyond the traditional von Neumann systems, called “Neuromorphic” (Table 3).

Neuromorphic hardware (Figure 21) is a specific brain-inspired ASIC that implements the SNNs. It has an object to reach the ability of the massively parallel brain processing in tens of watts on average. The memory and the processing units are in single abstraction (in-memory computing) (Figure 3). This leads to the advantage of dynamic, self-programmable behavior in complex environments (Kasabov, 2019). Instead of traditional bit-precise computing, neuromorphic hardware leads to “the probabilistic models of simple, reliable and power and data-efficient computing” as the “highly stochastic nature of computation of our brain” (Kasabov, 2019). Neuromorphic hardware certainly suits more for cognitive applications than precise computing.

	von Neumann Sequence	Neuromorphic
Representation of the data	Sequence of binary numbers	Spike(event) timings
Memory	1. Volatile 2. Non-volatile	1. Long term memory 2. Short term memory
Plasticity (Learning)	No	Adaptable via: 1. Long-term potentiation and depression 2. Short-term potentiation and depression
Processing	1. Deterministic 2. Centralized 3. Sequential	1. Stochastic 2. Decentralized 3. Parallel
Good At	1. High accuracy pattern recognition 2. High precision number crunching 3. Batch processing of data sets	1. Real-time processing of low-dimensional data 2. Ultra-low-power classification of sensory signals 3. Low-latency decision making

Table 3 von Neumann vs Neuromorphic difference table in a nutshell (*Kasabov, et al., 2016*) (*Indiveri, 2019*)

3.1 Advantages of Spiking Neural Networks

Learning of SNNs with neuromorphic chips can be handled both by native SNN algorithms and with the conversion of the 2nd generation algorithms of ANN into SNN.

Native SNN algorithms are theoretically promising efficient and effective, however, the practical issues with them continue. Huge efforts are made to improve these algorithms so that they can compete with 2nd generation ANN algorithms and eventually surpass them in both inference speed, accuracy and efficiency in the area of artificial intelligence. (NICE, 2019)

Another advantage of SNNs is the possibility of getting the benefits of the 2nd generation ANNs algorithms by conversion, i.e. deep learning networks (2nd generation ANN models) are mapped throughout, either empirically or mathematically, into SNN neurons. Therefore, successful deep learning operations can be converted into SNNs without any training algorithm (Zheng & Mazumder, 2020). Through this method, SNNs can reach the inference accuracy of cognitive applications with low energy consumption. (Figure 4) (Figure 5)

Applied Brain Research group, the owner of brain simulator Nengo, published a paper on 2018 to compare the Intel neuromorphic chip Loihi Wolf Mountain with conventional CPUs and GPUs (Blouw, et al., 2018). The methodology consisted in applying the 2nd generation ANN on GPUs and CPUs, and convert the 2nd generation to SNN to apply to Loihi. According to their result, for

real-time inference, (batch-size=1) neuromorphic Intel chip consumes 100x lower energy compared to GPUs which are the most common chips for implementing 2nd generation ANNs (Figure 4). Moreover, compared to Movidius, Loihi conserves the speed of the inference and the energy-consumption per inference as the number of neurons increases (Figure 5).

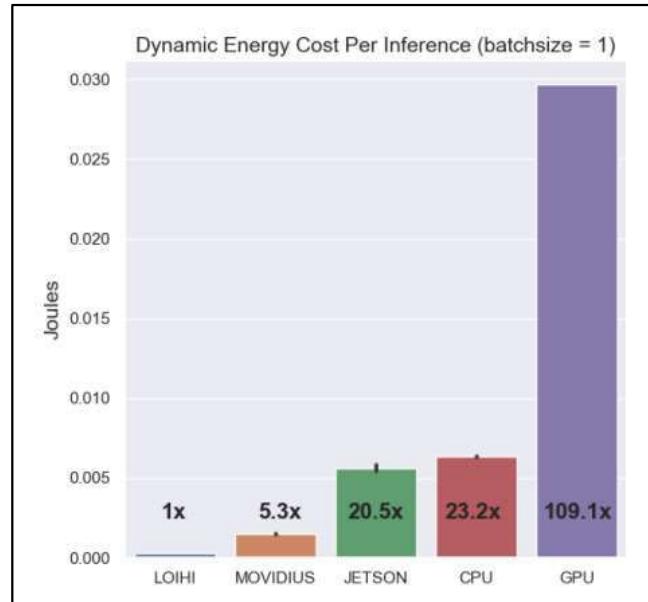


Figure 4 Loihi vs other semiconductors

Compared to Nvidia JETSON, Intel Movidius, and GPU, Loihi consumes minimum energy per inference (*Blouw, et al., 2018*)

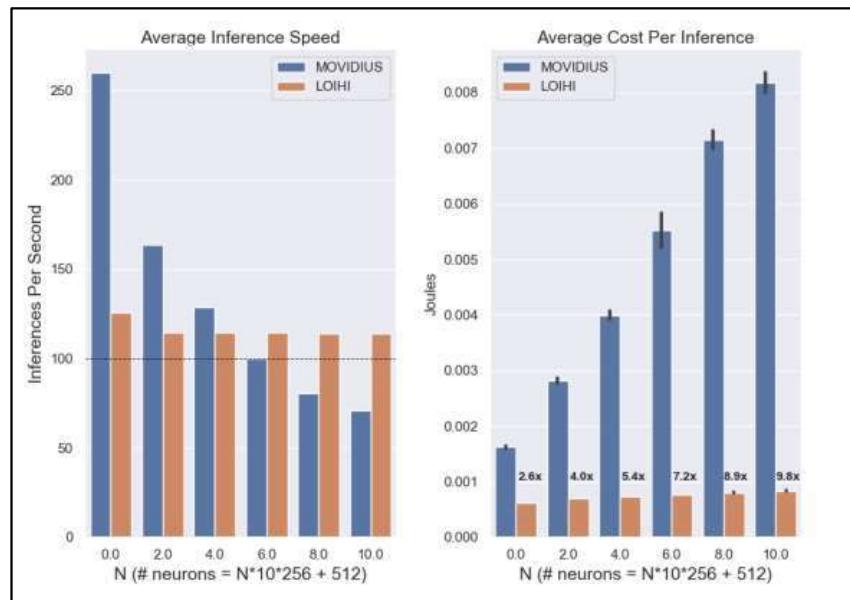


Figure 5 Performance of Loihi when implementing large scale-networks.

Loihi continues to perform in real-time even the network size increases (left). Loihi maintains the energy efficiency on the large scale networks (right) (*Blouw, et al., 2018*)

Even though the neuromorphic chips are not designed for 2nd generation of ANNs, this experiment indicates that neuromorphic chips can even outperform the traditional ones in real-time inference for large network sizes.

For the native SNN algorithms, Mike Davies -The Director of Intel's Neuromorphic Computing Lab- demonstrates that SNNs are not only energy-efficient and scalable but also more accurate and faster on specific applications (Sparse Coding, 100-1000x faster, 10000-100000x energy efficient compared to CPU). Davies also demonstrates the current achievements of SNNs with neuromorphic chips (Figure 6) and expects to see real-world applications (Davies, 2019).

- Inference and learning of sparse feature representations
- Video and speech recognition
- Event-based camera processing
- Chemosensing
- Robotics
- Adaptive dynamic control
- Anomaly detection for security and industrial monitoring
- Optimization: Constraint Satisfaction, QUBO, Convex optimization
- Autonomy: SLAM, planning, closed-loop behavior

Figure 6 Current applicable algorithms by SNNs with neuromorphic chips (Davies, 2019)

Another advantage of SNNs is their capacity to operate with event-based sensors energy-efficiently without the need for signal conversion (Gallego, et al., 2019). Event-based sensors are neuromorphic hardware that consumes less energy, while efficiently, under noise conditions.

3.2 Neuromorphic hardware implementations

Neuromorphic chips can be designed digital, analog or in a mixed way. All these designs have their pros and cons.

Analog chips resemble the characteristics of the biological properties of neural networks better than the digital ones. In the analog architecture, few transistors are used for emulating the differential equations of neurons. Therefore, theoretically, they consume lesser energy than digital neuromorphic chips (Furber, 2016) (Schemmel, et al., 2010). Besides, they can extend the processing beyond its allocated time slot. Thanks to this feature, the speed can be accelerated to process faster than in real-time. However, the analog architecture leads to higher noise, which lowers the precision. Also, the analog nature of the architecture causes leakage on signals which limits long-time learning in STDP (Indiveri, 2002).

Digital ones, on the other hand, are more precise compared to analog chips. Their digital structure enhances on-chip programming. This flexibility allows artificial intelligent researchers

to implement various kinds of an algorithm accurately with low-energy consumption compared to GPUs.

Mixed chips try to combine the advantages of the analog chips, i.e. lesser energy consumption, and the advantages of the digital ones, i.e. precision. (Milde, et al., 2017)

Yet the analog chips are more biological and promising, digital neuromorphic chips are on higher demand because they are easy to implement for real-world applications. As the learning algorithms for SNNs and hardware technology improve, analog architectures could eventually have the potential to take the position of digital.

3.3 Neuromorphic hardware leaders

3.3.1. IBM - TrueNorth

IBM with the collaboration of [the DARPA SYNAPSE program](#)⁶ built the chip “[TrueNorth](#)”. TrueNorth is a digital chip produced to speed-up the research on SNNs and commercialize it (Merolla, et al., 2014). It is not an on-chip programmable so it can be used just for inference. (Liu, et al., 2019). This is a disadvantage for the on-chip training research and at the same time limits the usage of the chip in critical applications (such as autonomous driving which needs continuous training) Efficient training - as mentioned before - is an advantage of neuromorphic hardware that unfortunately does not occur in the TrueNorth.

One IBM'S objective is to use the chip on cognitive applications such as robotics, classification, action classification, audio processing, stereo vision, etc. The chip has actually proven usefulness in relation to low energy consumption compared to GPUs (DeBole, et al., 2019). However, the TrueNorth is not yet on-sale for end-users, being only possible to request it for research reasons.

The chip is relatively old (5 years) and IBM presently seems not to be planning any new chip design but scaling it. IBM aims to invest in research focused on learning algorithms of SNNs and to take real-world applications to the market. With this goal, IBM is not only funding research (IBM labs around the world) but also sponsoring main neuromorphic hardware workshops (Neuro Inspired Computational Elements Workshop (NICE), Telluride Workshop).

IBM has also agreements with US security companies. In 2018, in partnership with The Airforce Research Laboratory, IBM unveiled the largest Neuromorphic supercomputer (Figure 22) “Blue Raven” (Figure 7). The objective is to have an intelligent, on-board, power-efficient systems on the battlefield for robust decision making (O'Brien, 2018) (Seffers, 2018).

⁶ “Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) program is to develop low-power electronic neuromorphic computers that scale to biological levels.” (DARPA, n.d.)



Figure 7 Blue Raven – A Large scale digital System (AFRL, 2018)

3.3.2. Intel - Loihi

Intel released its neuromorphic chip "[Loihi](#)" in 2018 (Davies, et al., 2018). The chip is digital and on-chip programmable. This gives flexibility to the chip so that researchers can work on a variety of learning methods, from DNN to SNN conversions, native SNN, etc. Currently, Loihi is the most effective and energy-efficient chip for cognitive applications (among the neuromorphic chips).

Like IBM, Intel is also investing in the commercialization of the neuromorphic chips and learning methodologies. Mike Davies' Lab collaborates with universities and research facilities to expand the visibility of Loihi.

Intel does not have much intention to research brain neurons, but they are more interested in cognitive applications. They expect to have a killer-app to solve real-world problems. And they believe that such an app should be related to the robotic sector, which is the one where the neuromorphic chips can more markedly express their competitive advantages, i.e. a "real-time inference with low energy consumption".

According to Mike Davies in his presentation on NICE 2019, one of the main objects of Intel is to expand the neuromorphic research society by sponsoring various workshops. The group has their special event called "Telluride Workshop" to foster their main ambition to have "real-world applications". The chips are not for sale at the moment, so only the close collaborators can benefit from the chip for research purposes. Various scale-size boards from USB to supercomputer scale (Figure 8)(Table 4) are in development for research users, while some recent agreements are progressing with Accenture, Airbus, GE, and Hitachi (Intel, 2019).

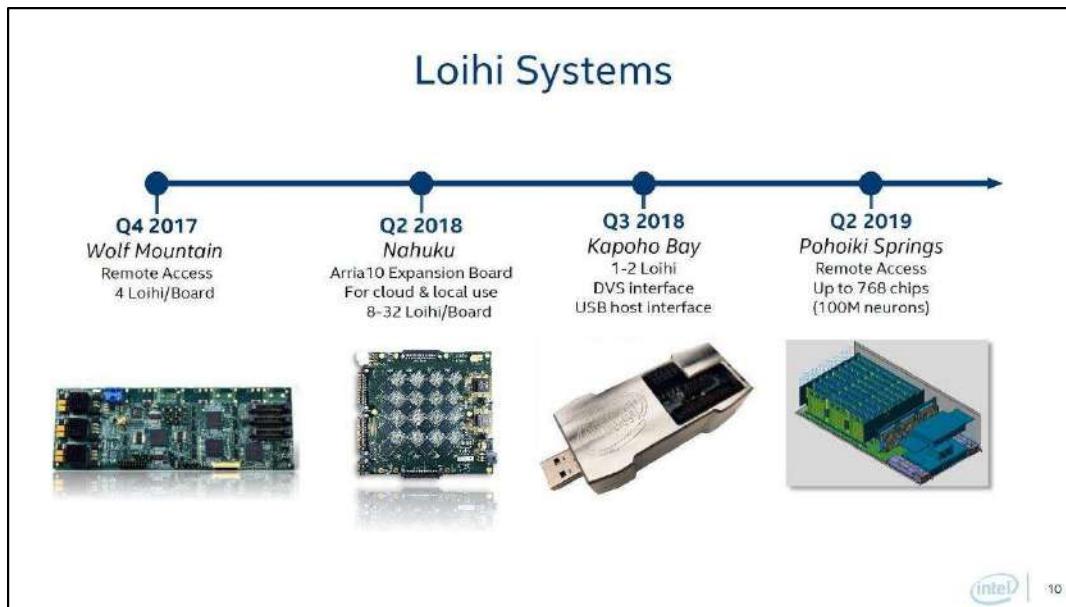


Figure 8 Intel roadmap for Loihi systems (NICE, 2019)

3.4 Neuromorphic hardware research chips

There are mutual relationships and interdependency between neuroscience and neuromorphic hardware. As neuroscience researchers discover the physical brain communications, mapping and learning mechanisms of the neural networks, these findings are designed and implemented in neuromorphic hardware. In turn, neuromorphic chips contribute to the effectiveness of neuroscience by emulating the brain models and allowing neuroscientists to make more complex and efficient experiments.

Different visualizations on the contribution of neuromorphic chips, key people and actors, integrated circuits in the market, and interconnections in the area have been included in the annex (elaborated with the road mapping software tool, 'Sharpcloud') to facilitate an overview of the area.

3.4.1. The University of Manchester - *SpiNNaker*

[SpiNNaker](#) (Figure 9)(Figure 25)(Figure 26) is a digital, on-chip programmable hardware designed by the University of Manchester, under the supervision of Steve Furber (Figure 24) (Furber, et al., 2013). SpiNNaker was the first on-chip programmable digital chip, so a large variety of research has been conducted around it.



Figure 9 SpiNNaker - Large Scale digital Neuromorphic System (Jordan, 2018)

Within the Human Brain Project, the focus of the University of Manchester is the study of the brain neurons rather than cognitive applications. However, as it is flexible, it can serve as a chip for cognitive applications too. Therefore, the Neurorobotics Platform of HBP is taking advantage of the SpiNNaker chip as hardware for robotic applications. The first generation SpiNNaker can be used on-cloud through HBP collaboration portal and the physical boards can be sold for research aims. Currently, SpiNNaker-1 is the world's largest neuromorphic computing platform and will assist EBRAINS. It has around 100 external users who can access the machine through the HBP Collaboratory. And there are around 100 SpiNNaker-1 systems in use by research groups around the world.

Even though the CMOS technology of the first SpiNNaker chip dates from the last decade, the recent study of Steve Furber and his team reveals that it manages to simulate Julich cortical unit⁷ in real-time and in an energy-efficient way. SpiNNaker-1 surpasses the HPC (runs in 3 times slower than real-time) and GPU (runs 2 times slower than real-time) in terms of processing speed and consumes significantly lesser energy compared to HPC and GPU (Oliver, et al., 2019). This study expresses the high flexibility and the long-term potential of the SpiNNaker architecture, and contributes to the competitive advantage of neuromorphic technology.

⁷ Julich cortical unit is the Cell-Type Specific Cortical Microcircuit founded by Tobias Potjans and Markus Diesmann. (Tobias & Markus, 2014)

Steve Furber expects that the neuromorphic technology will take the place of current solutions once they significantly show their commercial advantages in a specific application area ("killer app") "There are many promising potential applications, but none has yet met this critical

threshold", he expresses in our internal communication. His second generation of the SpiNNaker chip has been produced in HBP as a prototype and it is 10 times more energy-efficient and more effective in cognitive processes (Mayr, et al., 2019). The large scale version of this second generation of SpiNNaker will be deployed in TU Dresden with funding from the Saxon Science ministry (HBP Press Release, 2019).

3.4.2. Heidelberg University - *BrainScaleS*

BrainScaleS (Figure 10) (Figure 27) is an analog, on-chip STDP programmable hardware designed by the Heidelberg University currently under the supervision of Johannes Schemmel (Schemmel, et al., 2010). BrainScaleS is suited for brain-research applications that serve to discover and emulate the neuronal activities of biologic structures, as its architecture is the closest to the biological one. It is an accelerated system that runs 10000 times faster than biological speed, thus enabling the discovering of the various parts of the brain quicker (Schemmel, et al., 2017). However, the analog structure and acceleration (signal loss during such a rapid transmission) causes noise in signals and decreases the accuracy compared to other neuromorphic digital chips.

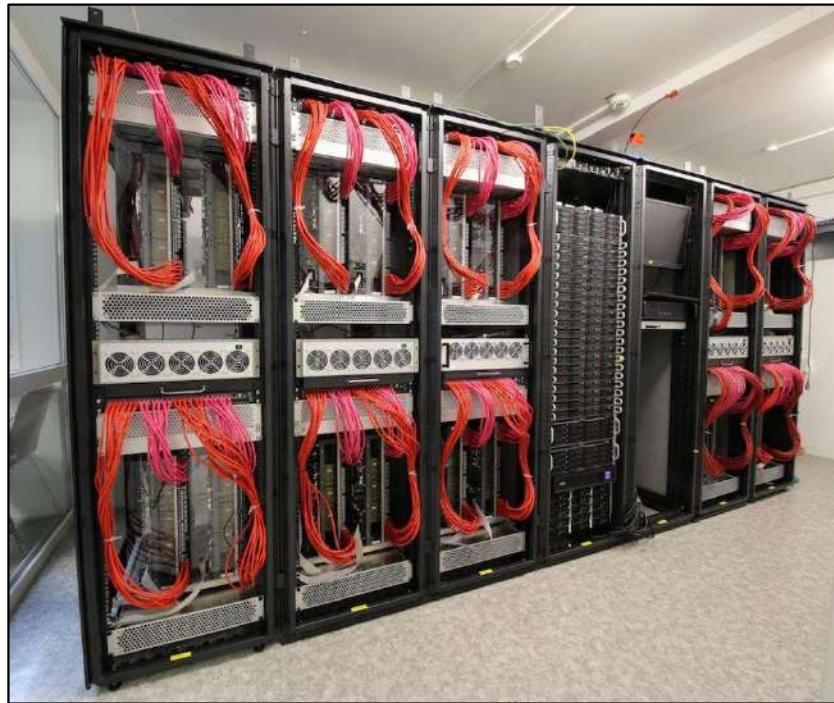


Figure 10 BrainScales – A Large Scale Analog Neuromorphic System

(HBP Neuromorphic Computing Platform, n.d.)

In the Human Brain Project, the main objective of the Heidelberg University neuromorphic research group is to contribute to the study of the neural networks of the human brain instead of focusing on commercial cognitive applications (Schemmel, et al., 2010). Due to its mentioned accuracy issues, it is not well-suited for real-word commercial applications yet. However, as the STDP learning methodologies and hardware improves, the evolution of the chip is promising in terms of both accuracy and energy consumption.

BrainScales 2 is on the prototype stage of maturity, and closer to its objective of supporting brain emulation. The chip will gradually help to capture more essential parts of the brain complex bio-structures in parallel with the advancements in neuroscience (Schemmel, et al., 2017) (Schemmel, 2019).

3.4.3. Stanford University – *Braindrop & Neurogrid*

[Neurogrid](#) (Benjamin, et al., 2014) and [Braindrop](#) (Neckar, et al., 2019) are subthreshold analog mixed-signal neuromorphic hardware built by the Stanford University under the supervision of Kwabena Boahen. Neurogrid dates from 2014 and Brarindrop has been recently launched in 2019. Stanford University designs chips for the research purposes, rather than commercial.

The object of Standford University was to capture the attention of the research community in the area of neuromorphic hardware with the NeuroGrid. The new object is to expand it to a wider community with its user-friendly programmability. Their designs on neuromorphic hardware stand out for their lower energy consumption compared to other chips. Especially, the new Braindrop chip claims to consume lower energy than the Intel Loihi and the “energy-efficient” AI accelerator of the Tesla chip.

3.4.4. Zhejiang & Hangzhou Dianzi University – *Darwin*

Darwin is the first neuromorphic chip from China. However, the chip remains in the prototype stage. (SHEN, et al., 2016) A recent post from the university reveals the fact that they are working on the 2nd version of this chip. (Zhejiang University, 2019)

3.4.5. CEA-Leti

[CEA-Leti](#) is the research institute of The French Alternative Energies and Atomic Energy Commission. They claim that their “SPIRIT” is the first non-volatile resistive memory neuromorphic chip in the world. The resistive memories are analog memories that surprisingly resemble the plasticity of brain synapses. Therefore, it can boost the potential of neuromorphic chips in the future. The prototypes will be available in 2021 (CEA-Leti Press Release , 2019).

Neuromorphic Systems	Chip	Number of Chips and Cores
BrainDrop Board	BrainDrop	1 X 4096core
BrainScaleS System	Hicann	1 Wafer Scale (352 chip can fit)
BrainScaleS 2 System	Hicann 2	1 Wafer Scale (under development)
Kapaho Bay	Loihi	2 x 128core
Nahuku	Loihi	32 x 128core
Pohoiki Springs	Loihi	768 x 128core
Wolf Mountain	Loihi	4 x 128core
NeuroGrid Board	NeuroGrid	16 x65536core
102 Machine	SpiNNaker 1	4 x 18core
103 Machine	SpiNNaker 1	48 x 18core
104 Machine	SpiNNaker 1	576 x 18core
105 Machine	SpiNNaker 1	5760 x 18core
106 Machine	SpiNNaker 1	57600 x 18core
SpiNNaker 2 System	SpiNNaker 2	70000 x 144core
NS16e	TrueNorth	16 x 4096core
NS16e-4	TrueNorth	64 x 4096core
NS1e Board	TrueNorth	1 x 4096core
NS1e-16	TrueNorth	16 x 4096core
SyNAPSE Board	TrueNorth	16 x 4096core

Table 4 Neuromorphic Systems and their origin chip.

The detailed version can be found in the UPM SharpCloud (the performance comparison is not available since there is no universal benchmarking for neuromorphic chips)

3.5 Neuromorphic hardware start-ups

Market-Ready, end-user programmable chips are an essential need for neuromorphic computing to expand its visibility and to achieve a variety of “real-world applications” with an increasing number of users. As large technology companies are waiting for the technology to become more mature, some start-ups are planning to release their chips in 2020 to fill the gap and to have a competitive advantage against those tech-giants. Three start-ups below are very active in the AI fairs and workshops and presenting the capabilities of their chip transparently.

3.5.1. Brainchip - Akida

[Brainchip](#) is an Australian company with the ambition of producing its own neuromorphic chip Akida. The chip enables to implement native SNN based algorithms or CNNs through their CNN2SNN converter software. The main advantage of Akida is that they can be on the market very soon, thus giving them the advantage of spreading out the technology among the early adopters of the SNN. They expect to begin sampling processors in early 2020 (Demler, 2019).

The company has started to sell Akida IP in mid-2019 and targeted applications in vision and acoustic systems. They are evaluating to open two innovation centers in one Australia (due to historical roots) and the second one in China (due to China's high potential in edge computing) (BrainChip Holdings Ltd., 2020). The Japanese manufacturer SocioNext is investing in Akida to manufacture the chip and plan to offer the chip to the customers through their platform (Smith, 2020).

3.5.2. aiCTX

[aiCTX](#) is a Zurich based start-up founded by Giacomo Indiveri who is one of the top researchers in the area. They have a variety of products in their catalog based on the DYNAPs (Moradi, et al., 2018) chip researched at the University of Zurich. They have modified and improved this original chip into different purposes. The chip DYNAP-SE2 is suitable for real-time applications in the area of robotics and medical health applications. DYNAP-SEL enables on-line learning and real-time implementation of large-scale models with its large fan-in and fan-out network connectivity. The latest available chip DYNAP-CNN is adapted for Spiking Convolutional Neural Networks which makes it the best candidate for visual processing applications via input from event-based vision sensors. (aiCTX, 2020)

Development kits for all these three chips are already available with the names of "Rockpool", "CorTeX ConTrol" and "Sinabs" respectively. Therefore, we can expect to see them in the market very soon.

aiCTX, with its innovative implementations, is on the radar of the investors, e.g. Pre Angel Capital (2018) and Baidu Ventures (2017). Moreover, the CES Asia Innovation award winner DYNAPs chip provided them with a very competitive EU grant (aiCTX, 2020).

Recently, the company researches the product "SPECK" with the collaboration of [iniVation](#) (an event-based camera producer). "SPECK" is an event-based mobile camera with an embedded neuromorphic chip. This allows the utilization of low latency and energy-efficient visual applications.

3.5.4. GrAI

[GrAI](#) is founded in 2016 within the iBionext Start-up Studio in Paris by Ryad Benosman, Bernard Gilly, Giacomo Indiveri, Xavier Lagorce, Sio-Hoi Leng, Bernabe Linares-Barranco and Atul Sinha. “GrAI One” is a hybrid neuromorphic chip that supports both SNN and ANN models, based on their technology “NeuronFlow” (Jonathan, 2020). The chip was produced in late 2019 and they have recently presented their hardware development kit at the tinyML Summit 2020. iBionext had funded 15 million dollars in 2017 (Nieke, 2019). They claim to consume lesser energy per inference than Intel Loihi and the IBM TrueNorth (Jonathan, 2019).

Besides the above mentioned start-ups that are orienting their chip design efforts in the spiking neuromorphic ASICs, other ones are producing ASICs and adopting novel approaches. These systems are hardwired in relation to the way they represent neural networks. Some initiatives are based on the optimization of the dataflow between the processing unit and the memory unit, e.g. “near data processing”, “High Bandwidth Memory”, “Hybrid Memory Cubes” (Ganguly, et al., 2019). Other ones are using novel memory architectures such as memristors (Ganguly, et al., 2019). For example, SCAiLE (SCalable AI for Learning at the Edge) is a consortium of three startups (Gyrfalcon Technology Inc., mtes Neural Networks Corporation (mNN) and Robosensing Inc.) whose objective is to combine advanced acceleration hardware, resistive memory (ReRAM), optimized neural networks *“to create ready-made, power-efficient solutions with unsupervised learning and event recognition capability”* (businesswire, 2019). Another start-up (Groq) has an on-chip memory for providing high-bandwidth of data source and eliminate the need for external memory (Gwennap, 2020). The outcomes from these non-Neuromorphic but innovative start-ups will certainly contribute as well to the Neuromorphic area since the visibility of chips and achievability of novel memory solutions will increase.

4 SOME APPLICATION AREAS

During the next decade, we will see how Neuromorphic computing gradually transforms the nature and functionalities of a wide range of scientific and non-scientific applications. In this report, we will briefly describe three specific but very large areas on which this emerging field of computing science is likely to impact more rapidly and intensively: 1) mobile applications, which are dramatically affecting our daily lives, are increasingly demanding more powerful processing capacities and abilities, 2) adaptive robotics, whose technological advance runs in parallel and is intimately linked to the progress of AI, needs to draw on the ‘human thinking’ mechanisms provided by neuromorphic chips to offer solutions more closely and effectively matched to the domestic and/or industrial users’ necessities, and 3) event-based vision sensors, that although may look, in principle, a less impactful area of application than the previous ones, certainly allow adaptive robotics to be fed with reliable visual signals and react accordingly with precise human-like responses.

4.1 Mobile Applications

Intelligent software is essential for the current usage of mobile applications. They cover from image processing to text processing, audio processing, etc. (Table 5).

These applications mostly require more processing power than mobile phones can currently handle. Therefore, they are using supercomputers via service calls to offer AI services. Although this way of service call works, it also has some critical issues such as (Pathak, 2017)

- They are limited by the speed of the internet connection.
- Their responsiveness depends on the service speed.
- They are causing privacy concerns (Howard, 2019).

On-device AI, therefore, is essential to solve these problems and enable leading-edge technologies.

Since the Snapdragon 820/836 (2016) mobile phones include processors with AI accelerators (Ignatov, et al., 2019). The most advanced chips at the time (2020 January) are Apple A13, Qualcomm Snapdragon 865 and Huawei Kirin 990. These chips can handle some edge AI applications such as face recognition, real-time translation, photo segmentation, voice recognition. However, their processing capacity is still limited for large-scale, complex or parallel-tasking services. Embedding extra processing power is not sustainable with the limitations of battery technology. Therefore, neuromorphic chips can have a big impact in this sector with their efficiency in real-time AI services.

usage	detailed usage	as core feature
image: 149	photo beauty: 97	94 (96.9%)
	face detection: 52	44 (84.6%)
	augmented reality: 19	5 (26.3%)
	face identification: 8	7 (87.5%)
	image classification: 11	6 (54.5%)
	object recognition: 10	9 (90%)
	text recognition: 11	4 (36.3%)
text: 26	word&emoji prediction: 15	15 (100%)
	auto-correct: 10	10 (100%)
	translation: 7	3 (42.8%)
	text classification: 4	2 (50%)
	smart reply: 2	0 (0%)
audio: 24	speech recognition: 18	7 (38.9%)
	sound recognition: 8	8 (100%)
other: 19	recommendation: 11	2 (18.1%)
	movement tracking: 9	4 (44.4%)
	simulation: 4	4 (100%)
	abnormal detection: 4	4 (100%)
	video segment: 2	1 (50%)
	action detection: 2	0 (0%)
total: 211		171 (81.0%)

Table 5 The detailed usage areas of AI processing on Smartphones (Xu, et al., 2019)

Neuromorphic devices are just pointing out the state of art issues in terms of energy efficiency. New architectures, learning algorithms, memory advancements, chip materials are challenges under intensive research in both universities and technology companies. According to the founder of the RAIN Neuromorphics, if current trends on neuromorphic chips continue, the mobile-edge chips could be on the market by mid-2020s (Kendall & Kumar, 2020).

4.2 Adaptive Robotics

Adaptive Robotics is a game-changer, an area with increasing demand and interest from enterprises throughout the world. Their smart sensing and autonomous decision capabilities are key features in the fourth industrial revolution. To increase autonomy and functioning in real-world conditions, the following aspects should be addressed:

- Sensors are needed that can capture high-dimensional data from the environment
- Fast (low-latency response) and energy-efficient processing power
- Self-learning capacity is needed to ensure precise behavior under dynamic and changing environments (Bing, et al., 2018)

Nowadays, artificial neural networks are being used to handle intelligent motor controls which include perception and learning. ANNs often require GPUs to process real-time tasks (Glatz, et al., 2019). However, these systems have difficulties to process these high-demand operations (Bing, et al., 2018). Firstly, their training functionality requires expensive computational power to eliminate latency and sustain responsiveness. Secondly, they do consume a lot of energy to handle large-scale operations. In fact, this power-latency trade-off is one of the main disadvantages of ANNs that are working with GPUs (Bing, et al., 2018) (Glatz, et al., 2019).

One of the solutions for addressing this issue is the cloud processing through service calls. The high processing capacity of the supercomputers allows them to handle learning and perception capabilities and enable them to control robots wirelessly. This option, while being functional in some cases, is not actually sustainable:

- *Communication Delays over the Cloud*; a huge amount of data can arise from the robotic applications from perception to SLAM (Simultaneous Localization and Mapping) and navigation. This can cause considerable delays in Cloud-Robot communication. Moreover, if there is no connection available, this will cause inoperability (e.g. space mission robots, miner robots, underwater operation robots, etc.)
- *Privacy and Security in Cloud–Robot Systems*; as the data is stored and processed over the cloud remotely, the robots are vulnerable to hacker attacks. This could lead to critical security issues such as changing the behavior of the robot or data-theft (Saha & Dasgupta, 2018).

Edge computing, which is taking the processing capacity closer to the device, is therefore required. The power-latency trade-off issue of current von Neumann based processing units should be replaced with energy-efficient, parallel functioning, scalable ones.

Neuromorphic chips are one of the best candidates to handle edge-robotic applications with their brain-like energy efficiency and scalable neuron systems. There is a huge amount of research on this area and Robotics is expected to be the killer-app for Neuromorphic chips.

Neuromorphic chips need to work with traditional CPUs for communicating with robot actuators and sensors. This may diminish the leading advantage of the energy efficiency of these chips (Knight & Nowotny, 2018). However, recent researches from Yulia Sandemirkaya (Glatz, et al., 2019) and hybrid Tianjic chip (Pei, et al., 2019) have already explored the state of the art of energy efficiency. In particular, Yulia Sandemirkaya and her team have managed to operate a robot with complete neuromorphic control eliminating the interference of the CPU. Therefore, they have shown the possibility of energy-efficient adaptive robotics with a visual sensor (Glatz, et al., 2019) (Pei, et al., 2019). On the other hand, Tsinghua University from China has presented the first neuromorphic hybrid chip “Tianjic” (Figure 11).

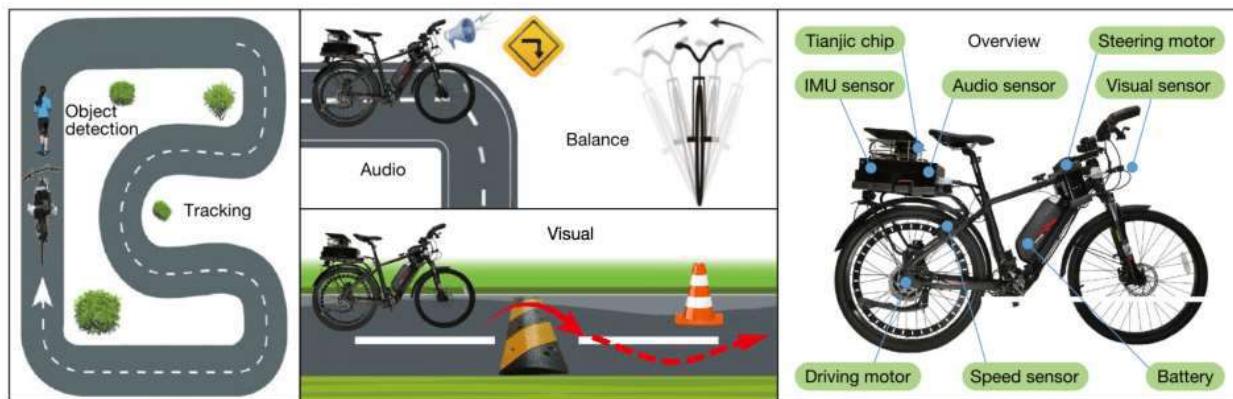


Figure 11 Autonomous driving bicycle with hybrid Tianjic chip (Pei, et al., 2019)

This programmable chip supports both ANN, SNN operations simultaneously by communicating with robotic actuators in an energy-efficient way. Their real-world application - which is controlled by the Tianjic chip - “smart bicycle” (Tsinghua University, 2019) can already compete with current adaptive robots in terms of intelligence. It is expected to have these hybrid chips in the market sooner than neuromorphic-only chips.

4.3 Event-Based Vision Sensors

Event-Based visual sensors are also bio-inspired – like neuromorphic chips - and they do also benefit from the low power consumption advantage. They measure the per-pixel brightness changes asynchronously, instead of capturing images with a fixed rate (Gallego, et al., 2019) (Unitectra, 2016). They do have critical advantages over conventional cameras, e.g. low latency,

less power consumption, high-temporal resolution, and high-dynamic-range. Their real-time working capability is essential for the artificial intelligence sector such as Robotics, autonomous vehicles, etc. Currently, most of the visual AI-applications have been researched and developed with these sensors: (Gallego, et al., 2019)

- Feature Detection and Tracking
- Optical Flow Estimation
- 3D reconstruction
- Pose estimation and SLAM
- Visual-Inertial Odometry
- Image Reconstruction
- Motion Segmentation
- Recognition
- Auditory Sensors

Two of these applications “Feature Detection” and “Motion Segmentation” are ready and available for customers’ use. The rest is still in the research phase because they cannot fetch the static object and “colors” by themselves. The processing unit is necessary for these applications, so traditional chips cannot be the long-term solution for event-based sensors. Real-time control of sensors with neuromorphic chips is an active area of research that aims to tackle this bottleneck. Yulia Sandemirskaya has recently worked on an event-based visual sensor with control of Intel Loihi chip and her team managed to achieve “Object tracking” (Renner, et al., 2019).

It is expected to have market-ready neuromorphic event-based systems in the horizon of two years. One example is the “SPECK” event-based vision sensor with DYNAPs neuromorphic control (Figure 12) (aiCTX, 2019).



Figure 12 SPECK - Event-based camera with neuromorphic chip control (Speck.ai, 2019)

Once successfully implemented with SLAM applications(Figure 29), these cameras can become the prior choice over traditional cameras used in the area of Robotics or autonomous driving.

5 MARKET TRENDS OF AI CHIPS

AI is very popular today and the market of chips is receiving an increasing interest and attention from the markets. A list of companies can be found in tables 6 and 7 of annex 2. Many applications are already adopted by end-users and numerous emerging applications are expected to happen in the short term. This rising demand will affect the plans of semiconductor companies. According to McKinsey's report "Artificial-intelligence hardware: New opportunities for semiconductor companies", estimated CAGR (Compound annual growth rate) of AI-supported semiconductors will be 18-19% between 2017-2025 compared to 3-4% of non-AI semiconductors (Figure 13) (Batra, et al., 2018). Report from TMT Analytics also correlated with McKinsey's and expects the market of AI-supported semiconductors to achieve 66 billion dollars by 2025 (Figure 14) (Kennis, 2019).

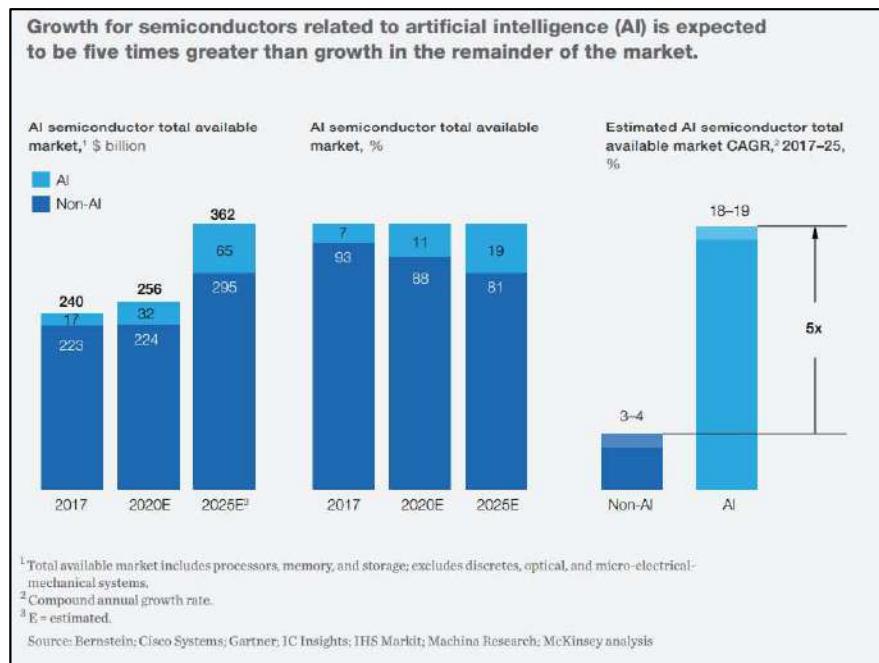
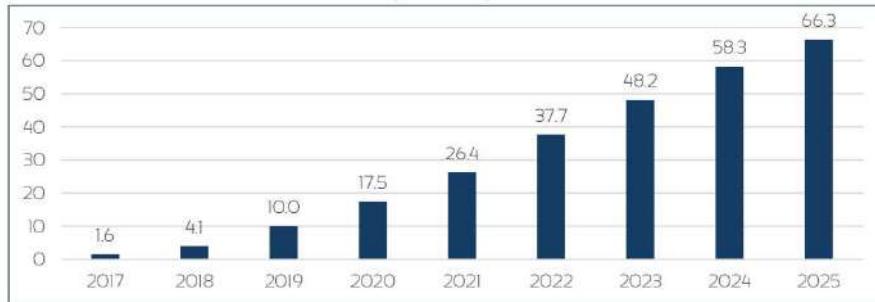


Figure 13 Market of AI & Non-AI semi-conductors between 2017-2025

By 2025 AI-supported semiconductors are estimated to take %19 of the share (Batra, et al., 2018)



Source: Tractica, TMT Analytics

Figure 14 AI-supported Semiconductors *market volume* (Kennis, 2019)

Currently available semiconductors for AI applications are the CPUs and the AI accelerators. The AI accelerators are leading the market because of the computing limitations of CPUs. Available AI accelerators are the GPUs, ASICs, and FPGAs, as mentioned in chapter 3.

GPUs have a lot of the parallel processing cores which give them a significant advantage for processing AI training and the inference. However, they do have a high power consumption cost which is not sustainable for future applications. Nvidia is at present the sector leader in the GPU area, especially for training.

Emerging FPGAs, on the other hand, can have 10 times more power-efficiency than GPUs but have lower performance. In applications where energy-efficiency is the top priority FPGAs can be the alternative solution. Some key players in the FPGA semiconductors are Intel - with its chip series Agilex -, Baidu - with Kunlun - and Xilinx – which has a quite large offer of chips.

Among AI Accelerators, ASICs shows the best performance, lesser power consumption, and efficiency. However, designing special functioning ASIC is highly costly and is not reconfigurable. Therefore, ASICs should be used when the market of specific AI applications is adequate for the design investment (Du & Du, 2018). Google is leading the sector with its chip Tensorflow Processing Unit (TPU). In parallel, INTEL is investing in ASICs with the collaboration of startups, “Nervana” and “Habana” (Intel decided to continue with “Habana” by January 2020) (Synced, 2020).

AI usage can be classified into four quadrants. The quadrants are the result of combining “Cloud/HPC/Data-Center” and “Edge/Embedded Computing”, which refer to the implementation place, and the “Training” or “Inference”⁸, which are the purpose of AI utilization (Figure 15). There is no viable solution for edge training since the data-size is huge compared to processing power for edge devices. Neuromorphic chips can fill this gap by its novel solutions.

⁸ Training: “Training refers to the process of creating a machine learning algorithm.”

Inference: “Inference refers to the process of using a trained machine learning algorithm to make a prediction.” (Paul, 2019)

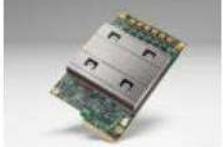
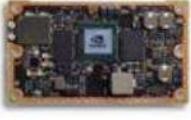
	Cloud/HPC/Data Center	Edge/Embedded
Training	 <ul style="list-style-type: none"> • High Performance • High Precision • High flexibility • Distributed • Need for Scale 	
Inference	 <ul style="list-style-type: none"> • High Throughput • Low Latency • Power Efficiency • Distributed • Scalable <p>CPU/GPU/ASIC/FPGA</p>	 <ul style="list-style-type: none"> • Wide range of Requirements (from ADAS to wearable) • Low-moderate Throughput • Low Latency • Power Efficiency (5-10 TOPS/W) • Low Cost <p>CPU/GPU/FPGA/ASIC</p>

Figure 15 Current semi-conductor and their location of use

There is still a gap in the edge-training part (Tsinghua University, 2018)

According to IBM supported research by “TIRIAS Research Group”, 95% of the AI processing is the “inference” one. This can clearly show that the inference AI chip market is dominating over training purposes (IBM PartnerWorld, n.d.).

5.1 Cloud/Datacenter

GPUs are currently dominating the cloud market for training services. Nvidia is the main actor in the field and this trend will likely continue during a couple of years. On the other hand, ASICs are emerging in the market and Google is already using their specified ASIC system “Tensor Processing Units” on its data centers. It is expected that ASICs will reach a 28 billion dollars chip market by 2026, which will be almost half of the total AI chip market (Bloomberg Business, 2019).

FPGAs research and development are basically supported by Intel and Microsoft. However, their lower performance compared to ASIC or GPU will limit opportunities in a market that will only demand efficient AI solutions. Their expected market share will be around 9 billion by 2024 (Bloomberg Business, 2018).

The inference is presently dominated by the traditional CPU datacenters (Figure 16). The CPU dominance will be gradually replaced by ASICs as soon as the utilization of the latter will become widespread. As the complexity of the tasks increase and datasets becomes larger, the inference cost of CPUs will be much higher. ASICs can provide the solution to this dilemma by the “increased parallelism, memory transfer reductions, and workload reductions” (Wang, et al., 2019).

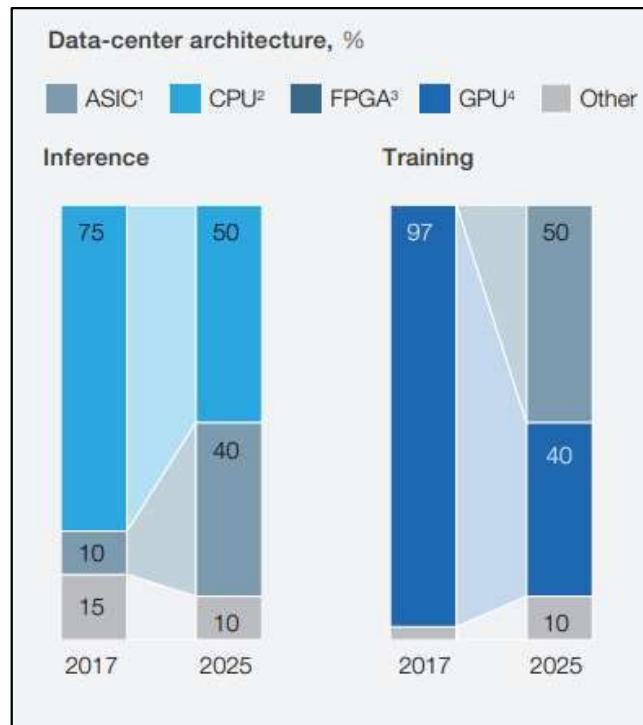


Figure 16 Data-Center semiconductors between 2017-2025

Data-centers is expected to replace CPUs over ASICs by late 2020s in the area of inference, GPUs are losing their dominance over ASICs in training by mid-2020s (Batra, et al., 2018)

5.2 Edge Computing

Edge computing represents the future of AI however, the amount of data transactions is increasing tremendously. Most of these data are just unnecessary bulky data which could make data-centers inadequate in the near future. Moreover, latency and real-time processing are crucial in some applications (health, space, robotics, etc.). Edge inference is inevitable to solve these issues.

According to Chetan Sharma Consulting, the Edge Market size is expected to reach 4.1 Trillion dollars by 2030. And half-trillion of this size will be located in the edge hardware market, which also includes the chip sector (Figure 17) (Chetan Sharma Consulting, 2019).

By 2018 the AI edge chip market was less than 100 million dollars; however, the demand will be huge. The top mobile chip producers, e.g. ARM architecture on Apple, Qualcomm, Huawei, have already edge inference and will certainly continue to invest. McKinsey believes that the edge inference chip market will get around 5 billion by 2025 and might surpass the data-center inference market by 2030 (Batra, et al., 2018).

At present, dominant processors in the edge market are CPUs. However, for large scale, real-time applications, CPUs will not be enough, and they will replace with ASICs by 2025 (Figure 18). On the other hand, edge training – even though being a very important area - is not efficient yet. There are some methodologies such as federated learning which boosts privacy and limits the data size. Unfortunately, this solution does not cover yet the latency related issues (Figure 19).

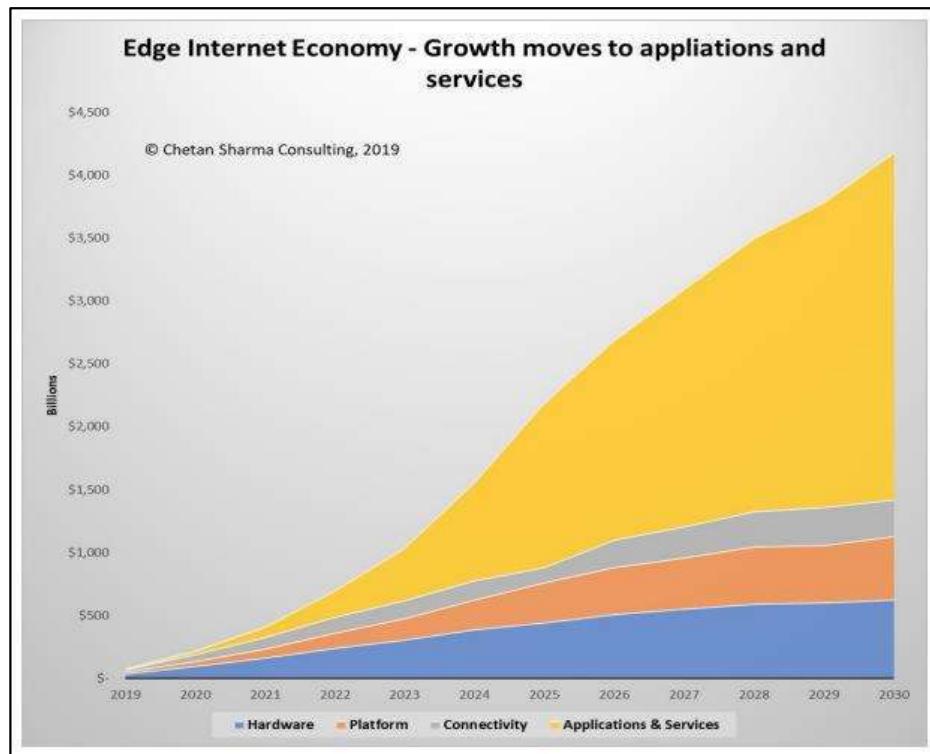


Figure 17 Edge Internet Economy,

The share of Hardware devices have an incrementing graph; around 500 billion dollars by 2030 (Chetan Sharma Consulting, 2019)

At both data centers and the edge, demand for training and inference hardware is growing.

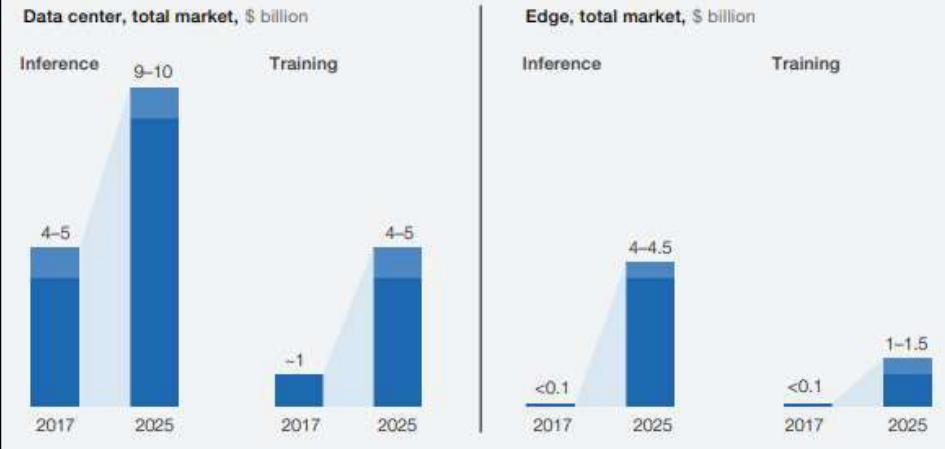


Figure 18 Market of Data Center & Edge-Computing semiconductors between 2017-2025

Huge Interest in Data Center training and Edge inference mid-2020s; Development in Edge training can be also seen in (Batra, et al., 2018)



Figure 19 Edge-Computing semiconductors between 2017-2025

ASICs are expected to take the place of CPUs in inference by mid-2020s. In edge training ASICs are again will rise, however, the total market is still low due to the technology inadequacy (Batra, et al., 2018)

5.3 Future Market of Neuromorphic Chips

Compared to AI Accelerators, Neuromorphic chips seems to be the best option in relation to “parallelism”, “energy efficiency” and “performance”. They can handle both AI inference and training in real-time. Moreover, edge training is possible through neuromorphic chips (Kendall & Kumar, 2020). However, learning methodologies should be improved their accuracy. In addition, there is no yet market-ready neuromorphic chips to widespread and cope with the potential user size. The start-ups mentioned in chapter 3.5 are expected to release their chip in the market in 2020. The success of aiCTX and BrainChip can be determinant for the future of neuromorphic computing. The hybrid research chip “Tianjic” has also tested on real-world applications and it could be a good sample for the transition period.

Yole and TMT Analytics expect that the market size of neuromorphic chips can reach billion-dollar by mid-20s with a growth of 51% between 2017-2023 (Yole Development, 2019) (Kennis, 2019). If they can manage to get ahead and demonstrate their potential under the pressure of the currently-successful AI accelerators, Neuromorphic chips are expected to take a solid place in the market by mid-20s and possibly achieve market domination by 2030 (Figure 20).

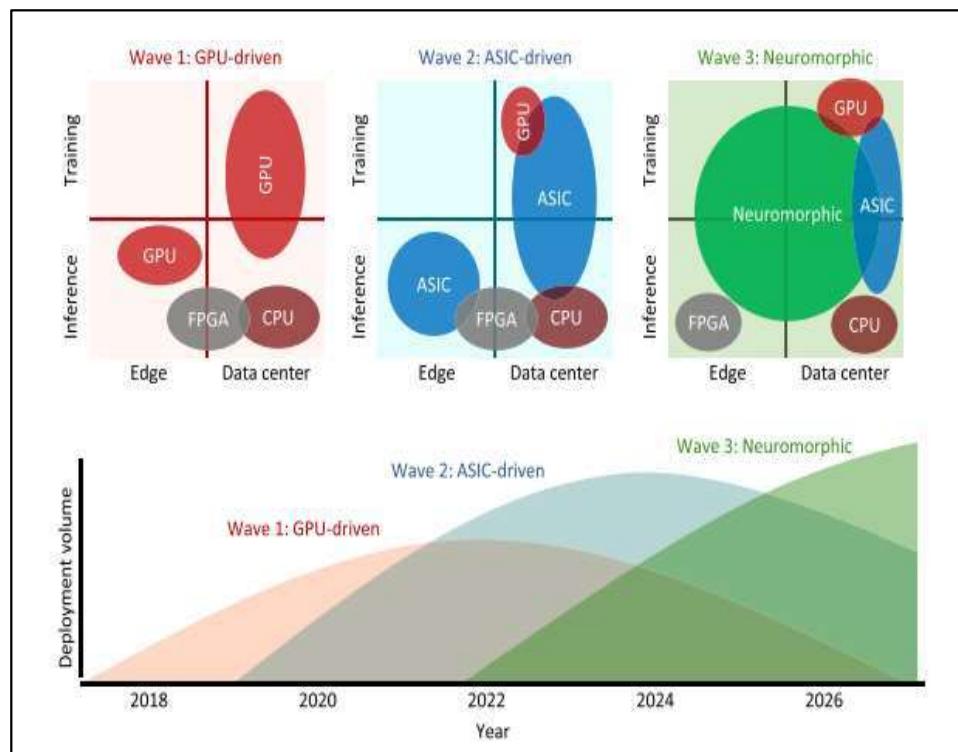


Figure 20 Expected waves for the semi-conductor dominance in the area of AI (Kendall & Kumar, 2020)

6 CONCLUSION

In our emerging and dynamic AI-based society, research and development on AI is to a large extent focused on the improvement and utilisation of deep neural networks and AI accelerators. However, there is a limit in the architecture of traditional von Neumann systems, and the exponential increasing of data-size and processing requires more innovative and powerful solutions. Spiking Neural Networks and Neuromorphic computing, which are well-developed and known areas among neuroscientist and neuro-computing researchers, are part of a trend of very recent and novel technologies that already contribute to enable the exploration and simulation of the learning structures of the human brain.

This report has explained the evolution of the artificial neuronal networks, the emergence of SNNs and their impact on the discovery of neuromorphic chips. It has been discussed the limitations of the traditional chips and the eventual influence of neuromorphic chips on demanding AI applications. The main players have been identified in the area, and have been related to current and future applications. The study has also described the market advantages of neuromorphic chips when comparing with other AI semiconductors. Neuromorphic chips are compatible with event-based sensors applications and emerging technologies such as photonic, graphene or non-volatile memories. They have a huge potential in the development of AI and can certainly be a dominant technology in the next decade.

Hopefully, the report has served to briefly give some light on the complexity of this challenging computing area. While staying loyal to our objective of offering a practical description of the most recent advances, we have also tried to be instructive enough so that to increase the interest and visibility of the topic to the non-specialised audience. For other readers the study may represent a promising and challenging step towards a more profound understanding of the area that could eventually support the creation of roadmaps, the exploration of new industrial applications, or the analysis of synergies between these novel chips and other related emerging trends.

7 REFERENCES

AFRL, 2018. *AFRL, IBM UNVEIL WORLD'S LARGEST NEUROMORPHIC DIGITAL SYNAPTIC SUPER COMPUTER*. [Online]

Available at: <https://afresearchlab.com/news/afrl-ibm-unveil-worlds-largest-neuromorphic-digital-synaptic-super-computer/>

[Accessed March 2020].

aiCTX, 2019. *aiCTX and iniVation announce Speck: a micropower smart vision sensor for mobile and IoT devices*. [Online]

Available at: [aiCTX and iniVation announce Speck: a micropower smart vision sensor for mobile and IoT devices](#)

[Accessed March 2020].

aiCTX, 2020. *aiCTX*. [Online]

Available at: <https://aictx.ai/>

[Accessed March 2020].

aiCTX, 2020. *DYNAP technology*. [Online]

Available at: <https://aictx.ai/technology/>

[Accessed March 2020].

Batra, G. et al., 2018. *Artificial-intelligence hardware: New opportunities for semiconductor companies*, s.l.: McKinsey&Company.

Benjamin, B. V. et al., 2014. Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations. *Proceedings of the IEEE*, 102(5), pp. 699-716.

Bing, Z. et al., 2018. A Survey of Robotics Control Based on Learning-Inspired Spiking Neural Networks. *Frontiers in Neurorobotics*, Volume 12.

Bloomberg Business, 2018. *Embedded FPGA Market Expected to Reach \$8,981 Million, Globally by 2024*. [Online]

Available at: <https://www.bloomberg.com/press-releases/2018-06-14/embedded-fpga-market-expected-to-reach-8-981-million-globally-by-2024>

[Accessed Mach 2020].

Bloomberg Business, 2019. *ASIC Chip Market to Reach \$28.04 Bn, Globally, by 2026 at 8.6% CAGR: Allied Market Research*. [Online]

Available at: <https://www.bloomberg.com/press-releases/2019-12-13/asic-chip-market-to-reach-28-04-bn-globally-by-2026-at-8-6-cagr-allied-market-research>

[Accessed March 2020].

Blouw, P., Xuan, C., Hunsberger, E. & Eliasmith, C., 2018. *Benchmarking Keyword Spotting Efficiency on Neuromorphic Hardware*. s.l., ICPS Proceedings, NICE 19, pp. 1-8.

BrainChip Holdings Ltd., 2020. *BrainChip Holdings (ASX:BRN) Year-End Update 2019*, Australia: s.n.

businesswire, 2019. *AI Innovators Join Forces in Consortium for Development and Commercialization of Best-in-Class AI Computing Platform*. [Online]
Available at: <https://www.businesswire.com/news/home/20190219005358/en/AI-Innovators-Join-Forces-Consortium-Development-Commercialization>
[Accessed March 2020].

Camuñas-Mesa, L., Linares-Barranco, B. & Serrano-Gotarredona, T., 2019. Neuromorphic Spiking Neural Networks and Their Memristor-CMOS Hardware Implementations. *Materials*, 12(17).

CEA-Leti Press Release , 2019. *CEA-Leti Builds Fully Integrated Bio-Inspired Neural Network with RRAM-Based Synapses & Analogue Spiking Neurons*. [Online]
Available at: <http://www.leti-cea.com/cea-tech/leti/english/Pages/What's-On/Press%20release/CEA-Leti-Builds-Fully-Integrated-Bio-Inspired-Neural-Network-with-RRAM-Based-Synapses-Analogue.aspx>
[Accessed March 2020].

Chetan Sharma Consulting, 2019. *Edge Internet Economy*, s.l.: Chetan Sharma Consulting.

DARPA, n.d. *Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) (Archived)*. [Online]
Available at: <https://www.darpa.mil/program/systems-of-neuromorphic-adaptive-plastic-scalable-electronics>
[Accessed March 2020].

Davies, M., 2019. *Realizing the Promise of Spiking Neuromorphic Hardware*, s.l.: RPG Workshops.

Davies, M. et al., 2018. Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. *IEEE Micro*, 33(1), pp. 82-99.

DeBole, M. V. et al., 2019. TrueNorth: Accelerating From Zero to 64 Million Neurons in 10 Years. *Computer*, 52(5), pp. 20-29.

Demler, M., 2019. *BRAINCHIP AKIDA IS A FAST LEARNER*, s.l.: The Linley Gorup.

Du, L. & Du, Y., 2018. Hardware Accelerator Design for Machine Learning. *Machine Learning - Advanced Techniques and Emerging Applications*.

Furber, S., 2016. Large-scale neuromorphic computing systems. *Journal of Neural Engineering*, 13(5).

Furber, S. et al., 2013. Overview of the SpiNNaker System Architecture. *IEEE TRANSACTIONS ON COMPUTERS*, 62(12).

Gallego, G. et al., 2019. Event-based Vision: A Survey. *arXiv.org*.

Ganguly, A., Muralidhar, R. & Singh, V., 2019. Towards Energy Efficient non-von Neumann Architectures for Deep Learning. *20th International Symposium on Quality Electronic Design (ISQED)*.

Glatz, S. et al., 2019. Adaptive motor control and learning in a spiking neural network realised on a mixed-signal neuromorphic processor. *2019 International Conference on Robotics and Automation (ICRA)*.

Gwennap, L., 2020. *GROQ ROCKS NEURAL NETWORKS*. [Online]

Available at: <https://groq.com/wp-content/uploads/2020/01/Groq-Rocks-NNs-Linley-Group-MPR-2020Jan06.pdf>

[Accessed March 2020].

HBP Neuromorphic Computing Platform, n.d. *The BrainScaleS “physical model” system*. [Online]

Available at: <https://electronicvisions.github.io/hbp-sp9-guidebook/pm/pm.html>

[Accessed March 2020].

HBP Press Release, 2019. *Saxon Science Ministry delivers 8 Mio Euro to TU Dresden for second generation SpiNNaker machine*. [Online]

Available at: <https://www.humanbrainproject.eu/en/follow-hbp/news/second-generation-spinnaker-neuromorphic-supercomputer-to-be-built-at-tu-dresden/>

[Accessed March 2020].

Hongming, C. et al., 2018. The rise of deep learning in drug discovery. *Drug Discovery Today*, 23(6).

Howard, A., 2019. *Introducing the Next Generation of On-Device Vision Models: MobileNetV3 and MobileNetEdgeTPU*, s.l.: s.n.

IBM PartnerWorld, n.d. *Why Your AI infrastructure Needs Both Training and Inference*, s.l.: s.n.

Ignatov, A. et al., 2019. AI Benchmark: All About Deep Learning on Smartphones in 2019. *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*.

Indiveri, G., 2002. Neuromorphic Bisable VLSI Synapses with Spike-Timing-Dependent Plasticity. *Advances in Neural Information Processing Systems*.

Indiveri, G., 2019. *DYNAP-SEL: An ultra-low power mixed signal Dynamic Neuromorphic Asynchronous Processor with Self Learning abilities*, NICE Workshop: s.n.

Intel, 2019. *Accenture, Airbus, GE and Hitachi Join Intel Neuromorphic Research Community*. [Online]

Available at: <https://newsroom.intel.com/news/accenture-airbus-ge-hitachi-join-intel-neuromorphic-research-community/>

[Accessed March 2020].

Jonathan, T., 2019. *GRAI ONE NEURONFLOW PROCESSING FOR THE EDGE*. [Online]

Available at: <https://www.graimatterlabs.ai/files/GML-Linley-FPC-OCT-24-FINAL-new.pdfv>

[Accessed March 2020].

Jonathan, T., 2020. *GrAI One – A Hybrid Neuromorphic and Dataflow Processor*. [Online]

Available at: https://www.tinyml.org/summit/abstracts/GrAI_Matters_Lab_demo_abstract.pdf

[Accessed March 2020].

Jordan, K., 2018. 'Human brain' supercomputer with 1 million processors switched on for first time. [Online]

Available at: <https://techxplore.com/news/2018-11-human-brain-supercomputer-million-processors.html>

[Accessed March 2020].

Kasabov, K. N., 2019. *Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence*. Berlin, Germany: Springer.

Kasabov, N., Sengupta, N. & Scott, N., 2016. From von neumann, John Atanasoff and ABC to Neuromorphic computation and the NeuCube spatio-temporal data machine. *2016 IEEE 8th International Conference on Intelligent Systems (IS)*.

Kendall, D. J. & Kumar, S., 2020. The building blocks of a brain-inspired computer. *Applied Physics Reviews*, 7(1).

Kennis, M., 2019. *BrainChip Holdings Limited: Taking Deep Learning to higher levels*, s.l.: TMT Analytics.

Knight, J. & Nowotny, T., 2018. GPUs Outperform Current HPC and Neuromorphic Solutions in Terms of Speed and Energy When Simulating a Highly-Connected Cortical Model. *Frontiers in Neuroscience*.

Krizhevsky, A., Sutskever, I. & Hinton, G., 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), pp. 84-90.

Lee, J., Delbruck, T. & Pfeiffer, M., 2016. Training Deep Spiking Neural Networks Using Backpropagation. *Frontiers in Neuroscience*, Volume 10.

Liu, Y., Zhang, W. & Li, P., 2019. Enabling Non-Hebbian Learning in Recurrent Spiking Neural Processors With Hardware-Friendly On-Chip Intrinsic Plasticity. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 9(3), pp. 465-474.

Lundstrom, M., 2003. Moore's Law Forever?. *Science*, 299(210).

Maass, W., 1997. Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, 10(9), pp. 1659-1671.

Mayr, C., Sebastian, H. & Steve, F., 2019. SpiNNaker 2: A 10 Million Core Processor System for Brain Simulation and Machine Learning. *arXiv:1911.02385*.

Merolla, P. et al., 2014. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345(6197).

Microsoft, 2020. *What are field-programmable gate arrays (FPGA) and how to deploy*. [Online] Available at: <https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-fpga-web-service> [Accessed March 2020].

Milde, M. et al., 2017. Obstacle Avoidance and Target Acquisition for Robot Navigation Using a Mixed Signal Analog/Digital Neuromorphic Processing System. *Frontiers in Neurorobotics*.

Moradi, S., Qiao, N., Stefanini, F. & Indiveri, G., 2018. A Scalable Multicore Architecture With Heterogeneous Memory Structures for Dynamic Neuromorphic Asynchronous Processors (DYNAPs). *IEEE Transactions on Biomedical Circuits and Systems*, 12(1), pp. 106-122.

Neckar, A. et al., 2019. Braindrop: A Mixed-Signal Neuromorphic Architecture With a Dynamical Systems-Based Programming Model. *Proceedings of the IEEE*, 107(1), pp. 144-164.

NICE, 2019. *NICE 2019 - Day 1n Panel Discussion*. [Online]

Available at:

https://www.youtube.com/watch?v=SyDISx5zzRA&list=PLJ506hQ4g3Th3sDNaHigmK6Pr0YTH_aZo&index=14&t=0s

[Accessed March 2019].

Nieke, R., 2019. *Eindhoven-connected Grai Matter Labs edges to the AI market*. [Online]

Available at: <https://bits-chips.nl/artikel/eindhoven-connected-grai-matter-labs-edges-to-the-ai-market/>

[Accessed March 2020].

O'Brien, J. W., 2018. *AFRL, IBM unveil world's largest neuromorphic digital synaptic super computer*. [Online]

Available at: <https://www.wpafb.af.mil/News/Article-Display/Article/1582310/afrl-ibm-unveil-worlds-largest-neuromorphic-digital-synaptic-super-computer/>

[Accessed 2020].

Oliver, R. et al., 2019. Real-time cortical simulation on neuromorphic hardware. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 378(2164).

Pathak, N., 2017. Artificial Intelligence for .NET: Speech, Language, and Search.

Patterson, J. C., 2012. *MANAGING A REAL-TIME MASSIVELY-PARALLEL NEURAL ARCHITECTURE*. Phd Thesis ed. Manchester: UNIVERSITY OF MANCHESTER.

Paul, D., 2019. *Training versus Inference*. [Online]

Available at: <https://blogs.gartner.com/paul-debeasi/2019/02/14/training-versus-inference/> [Accessed March 2020].

Pei, J. et al., 2019. Towards Artificial General Intelligence with Hybrid Tianjic Chip Architecture. *Nature*, 572(7767), pp. 106-111.

Pelé, A.-F., 2019. *Neuromorphic Promises Better AI*. [Online]

Available at: <https://www.eetimes.com/neuromorphic-promises-better-ai/#> [Accessed March 2020].

Renner, A., Evanusa, M. & Sandamirskaya, Y., 2019. Event-based attention and tracking on neuromorphic hardware. *arXiv*.

Saha, O. & Dasgupta, P., 2018. A Comprehensive Survey of Recent Trends in Cloud Robotics Architectures and Applications. *Robotics*, 7(3).

Schemmel, J., 2019. *Turing, or Non-Turing? That is the question*, s.l.: NICE Workshop.

Schemmel, J. et al., 2010. A wafer-scale neuromorphic hardware system for large-scale neural modeling. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*.

Schemmel, J., Kriener, L., Muller, P. & Meier, K., 2017. An accelerated analog neuromorphic hardware system emulating NMDA- and calcium-based non-linear dendrites. *2017 International Joint Conference on Neural Networks (IJCNN)*.

Seffers, G., 2018. *AFRL Anticipates Arrival of Neuromorphic Supercomputer*. [Online]

Available at: <https://www.afcea.org/content/afrl-anticipates-arrival-neuromorphic-supercomputer> [Accessed March 2020].

SHEN, J. et al., 2016. Darwin: a neuromorphic hardware co-processor based on Spiking Neural Networks. *SCIENCE CHINA*, Volume 59.

Smith, J., 2020. *BrainChip's (ASX:BRN) Akida tech to be sold through Socionext*. [Online]

Available at: <https://themarketherald.com.au/brainchips-asxbrn-akida-tech-to-be-sold-through-scionext-2020-03-23/> [Accessed March 2020].

Speck.ai, 2019. *SPECK*. [Online]

Available at: <https://www.speck.ai/> [Accessed March 2020].

Sugiarto, I. & Pasila, F., 2018. Understanding a Deep Learning Technique through a Neuromorphic System a Case Study with SpiNNaker Neuromorphic Platform. *MATEC Web of Conferences*.

Synced, 2020. *Intel AI GM Naveen Rao Leaves After Habana Chosen Over Nervana*. [Online] Available at: <https://syncedreview.com/2020/02/26/intel-ai-gm-naveen-rao-leaves-after-habana-chosen-over-nervana/> [Accessed March 2020].

Tobias, C. P. & Markus, D., 2014. The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. *Cerebral Cortex*, 24(3), pp. 785-806.

Tsinghua University, 2018. *White Paper on AI Chip Technologies*, Beijing: Tsinghua University.

Tsinghua University, 2019. *Computer science, neuroscience, and self-driving bicycle*. [Online] Available at: <https://www.youtube.com/watch?v=eTJOcBCFmDc>

Unitectra, 2016. *Dynamic Vision Sensor (DVS)*. [Online] Available at: <https://www.youtube.com/watch?v=0ZEM57DZJes> [Accessed March 2020].

Wang, E. D. J., Zhao, R. & Ng, H.-C., 2019. Deep Neural Network Approximation for Custom Hardware: Where We've Been, Where We're Going. *ACM Computing Surveys*.

Widrow, B. & Lehr, M., 1990. 30 years of adaptive neural networks: perceptron, Madaline, and backpropagation. *Proceedings of the IEEE*, 78(9), pp. 1415-1442.

Xu, M. et al., 2019. A First Look at Deep Learning Apps on Smartphones. *The World Wide Web Conference on - WWW 19*.

Yole Development, 2019. *Neuromorphic Sensing and Computing 2019; Market and Technology Report*. [Online] Available at: https://yole-i-micronews-com.osu.eu-west-2.outscale.com/uploads/2019/09/YD19039_Neuromorphic_Sensing_Computing2019_sample.pdf [Accessed March 2020].

Zhejiang University, 2019. ‘*Darwin NPU 2*’ developed to process information faster. [Online] Available at: <https://www.zju.edu.cn/english/2019/0902/c19573a1614979/page.htm> [Accessed March 2020].

Zheng, N. & Mazumder, P., 2020. *Learning in energy-efficient neuromorphic computing: algorithm and architecture co-design*. s.l.:Wiley-IEEE Press.

List of Figures and Tables

Table 1 Most common 2 nd generation ANN models (Kennis, 2019)	4
Table 2 ANN-SNN Comparison Table (Tsinghua University, 2018)	5
Table 3 von Neumann vs Neuromorphic difference table in a nutshell.....	9
Table 4 Neuromorphic Systems and their origin chip.	18
Table 5 The detailed usage areas of AI processing on Smartphones	21
Table 6 Basic info on potentially interested Companies	51
Table 7 Assessment of exploitation potential	55

Figure 1- Generations of Artificial Neural Networks	3
Figure 2 Current AI-hardware solutions	7
Figure 3 von Neumann vs Neuromorphic architecture	8
Figure 4 Loihi vs other semiconductors.....	10
Figure 5 Performance of Loihi when implementing large scale-networks.....	10
Figure 6 Current applicable algorithms by SNNs with neuromorphic chips.....	11
Figure 7 Blue Raven – A Large scale digital System	13
Figure 8 Intel roadmap for Loihi systems	14
Figure 9 SpiNNaker - Large Scale digital Neuromorphic System	15
Figure 10 BrainScales – A Large Scale Analog Neuromorphic System.....	16
Figure 11 Autonomous driving bicycle with hybrid Tianjic chip	23
Figure 12 SPECK - Event-based camera with neuromorphic chip control.....	24
Figure 13 Market of AI & Non-AI semi-conductors between 2017-2025.....	25
Figure 14 AI-supported Semiconductors market volume.....	26
Figure 15 Current semi-conductor and their location of use	27
Figure 16 Data-Center semiconductors between 2017-2025	28
Figure 17 Edge Internet Economy,.....	29
Figure 18 Market of Data Center & Edge-Computing semiconductors between 2017-2025.....	30
Figure 19 Edge-Computing semiconductors between 2017-2025.....	30
Figure 20 Expected waves for the semi-conductor dominance in the area of AI	31
Figure 21 Neuromorphic Chips.	41
Figure 22 Neuromorphic Scale-Systems.	42
Figure 23 UPM relationship diagram of Neuromorphic Computing.	43
Figure 24 Neuromophic Computing Family, Sample. "Steve Furber" and his connections.	44
Figure 25 Neuromophic Computing Family, Sample B "SpiNNaker"	45
Figure 26 Neuromophic Computing Family, Sample C "SpiNNaker" in detail.....	46
Figure 27 Neuromophic Computing Family, Sample D. "Schemel J." and his connections.....	47
Figure 28 Neuromophic Computing Family, Sample E "Wold-Map"	48
Figure 29 UPM roadmap for Event-Based Sensors.....	49

Neuromorphic Chips

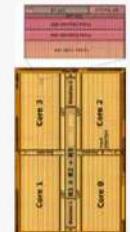
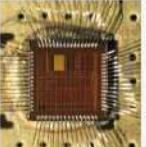
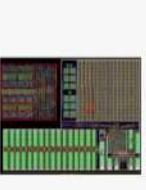
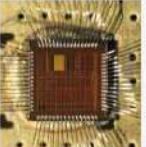
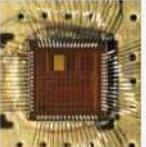
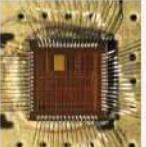
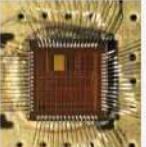
Akida			BrainDrop			Darwin 2			DYNAPS			HICANN
HICANN 2			Loihi			Neurocore			ODIN			SpiNNaker
SpINNaker 2			Tianjic			TrueNorth						

Figure 21 Neuromorphic Chips.

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.

Neuromorphic Scale-Systems

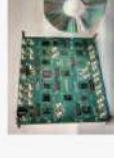
	102 board		104 board		105 board		BrainDrop Board		NS16e-4		NSTe Board		NS1e-16
	Nahuku		NeuroGrid		Poholki Springs		SpiNNaker 2 System		SyNAPSE Board		Wolf Mountain		
	Kapaho Bay		BrainScales 2 System		BrainScales System		NSTe Board		NS1e-16				

Figure 22 Neuromorphic Scale-Systems.

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.

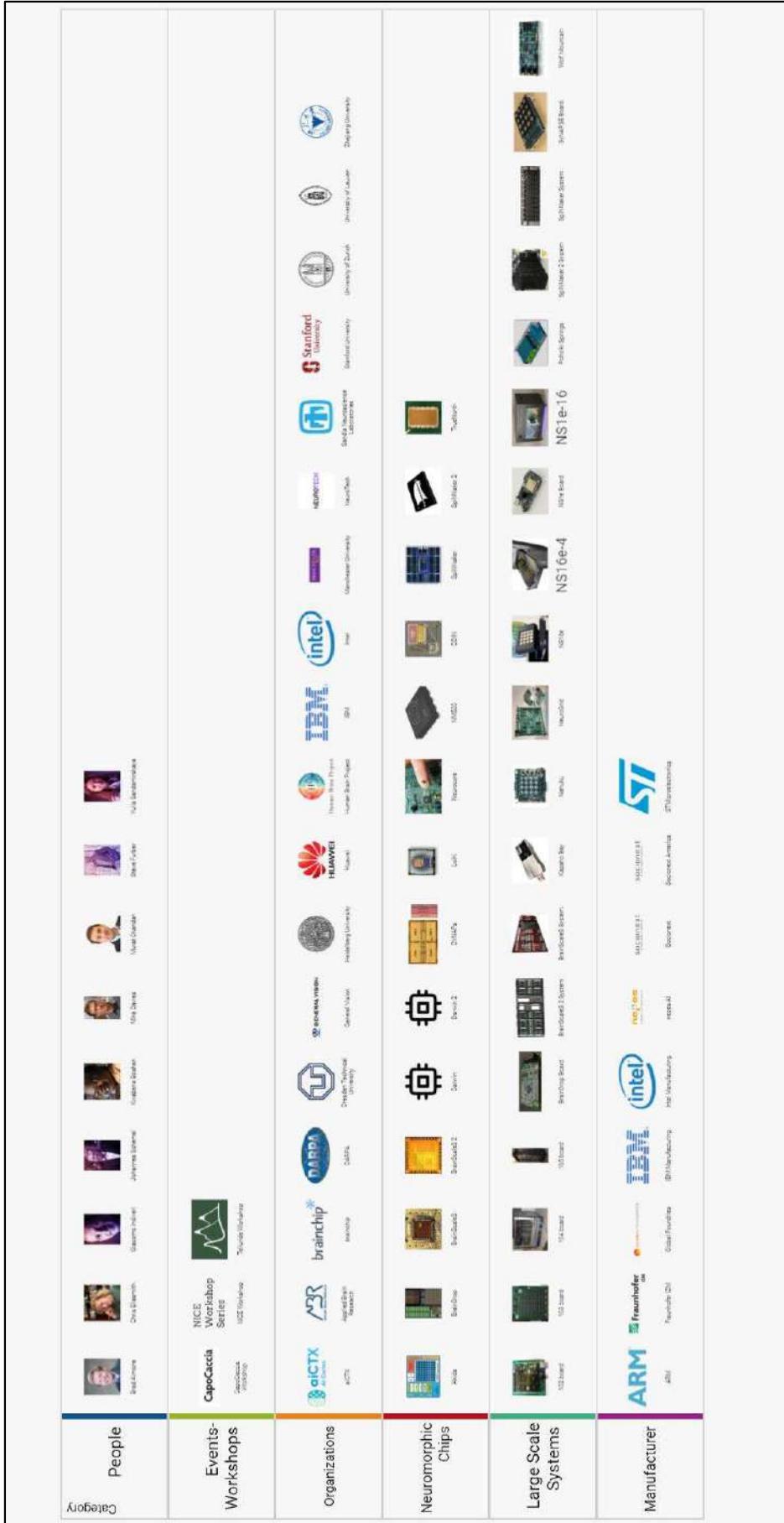


Figure 23 UPN1 relationship diagram of Neuromorphic Computing.

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.

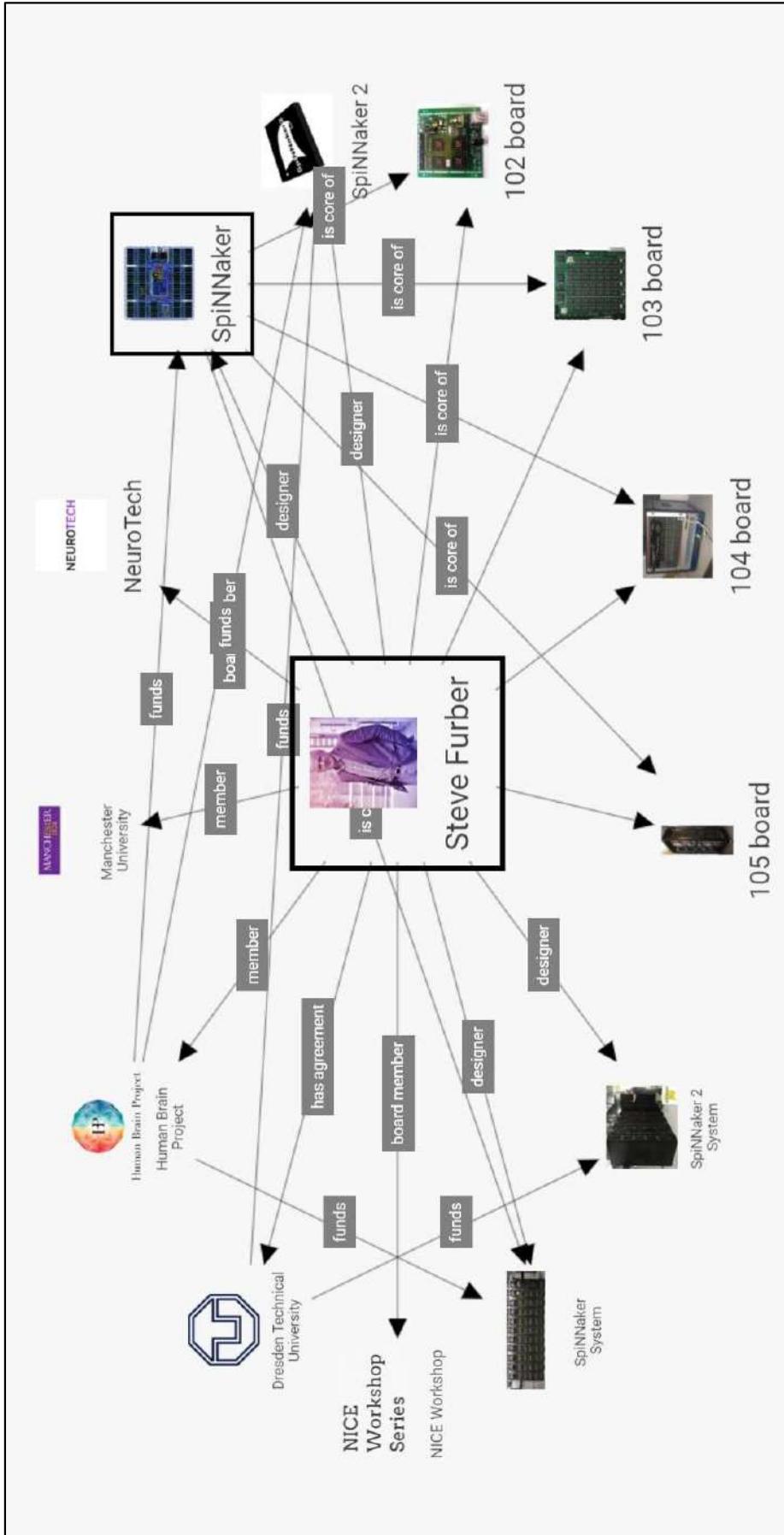


Figure 24 Neuromorphic Computing Family. Sample. "Steve Furber" and his connections.

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.

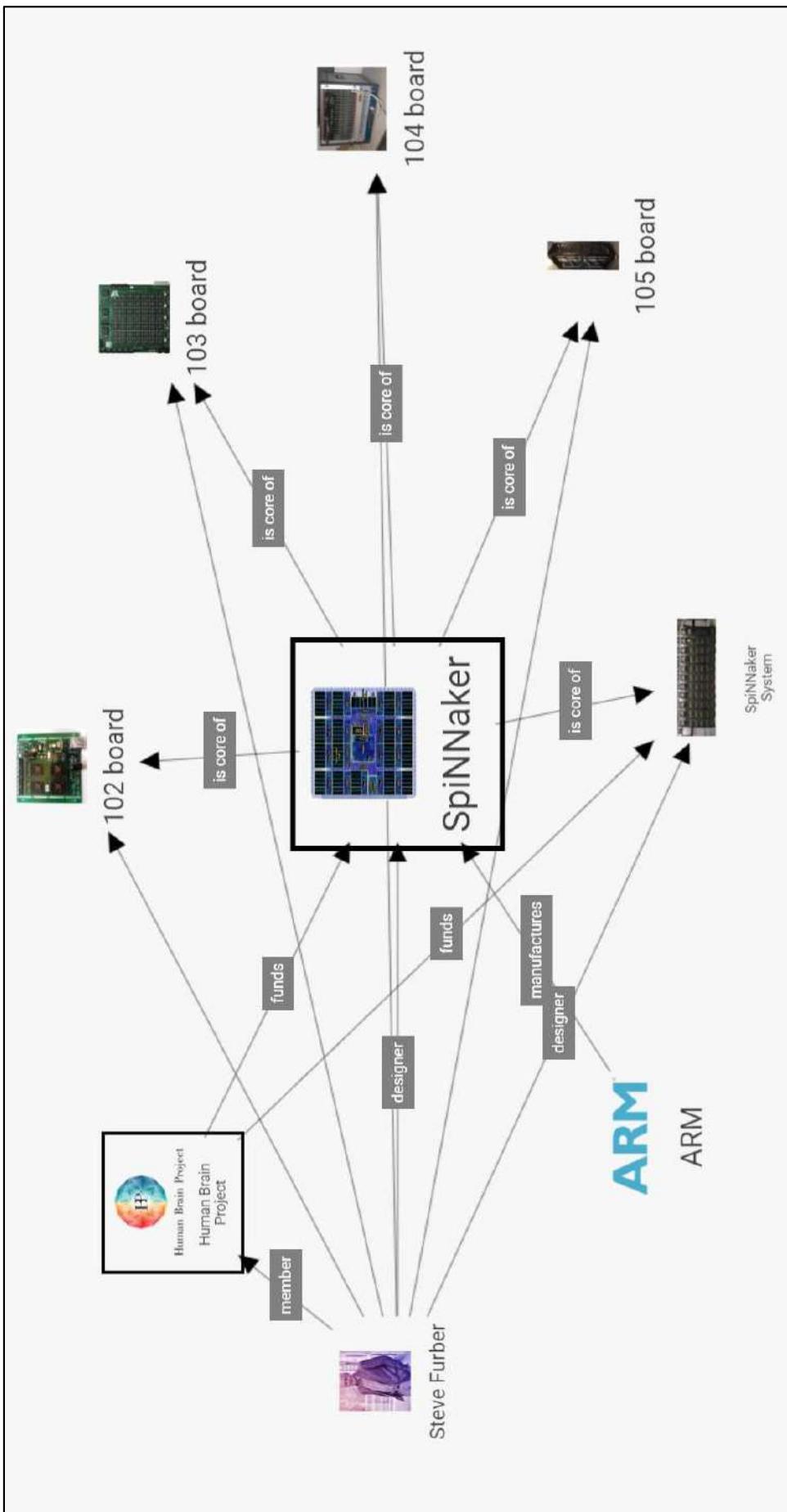


Figure 25 Neuromorphic Computing Family, Sample B "SpiNNaker"

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.

Figure 26 Neuromorphic Computing Family, Sample C "SpiNNaker" in detail.

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.

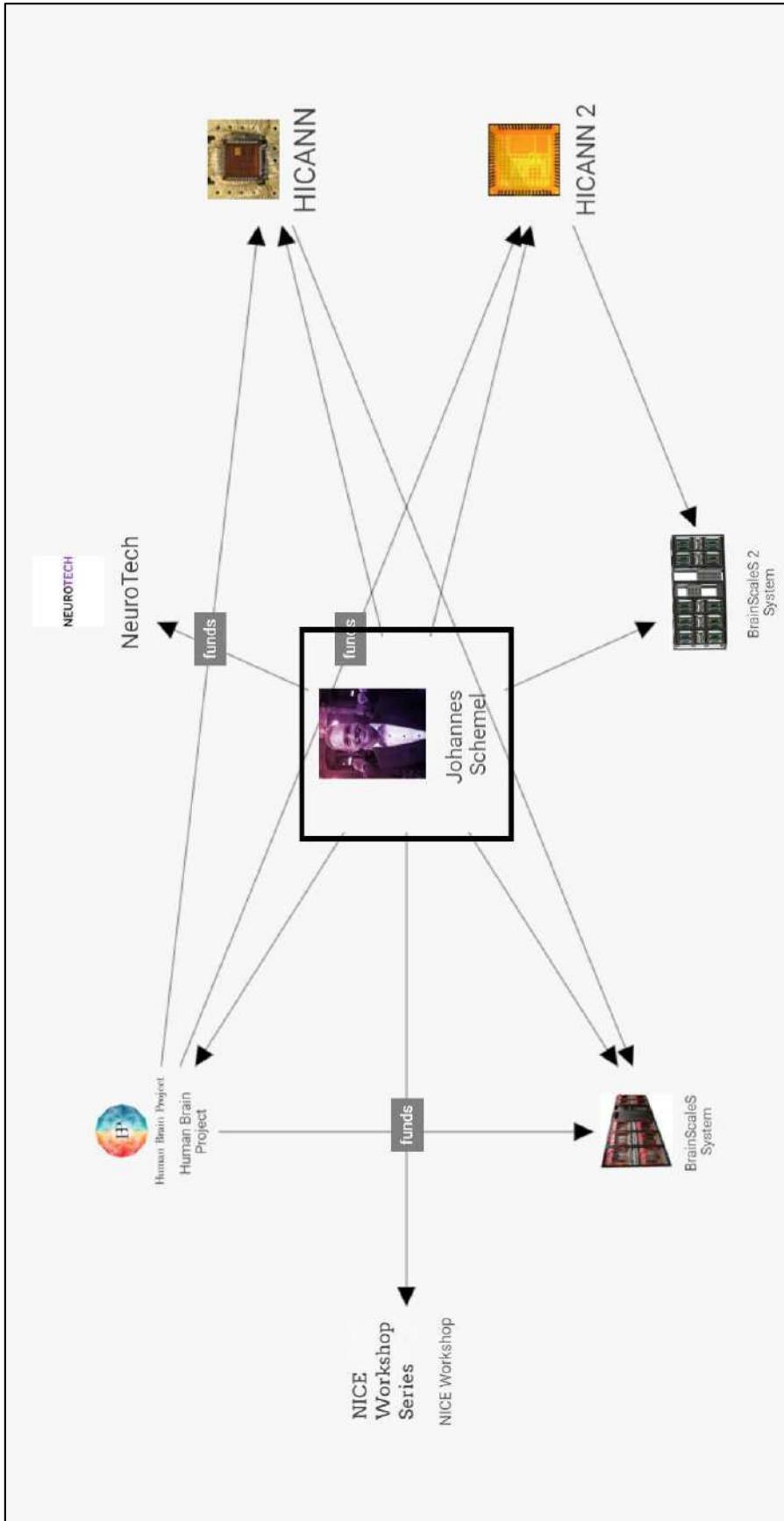


Figure 27 Neuromorphic Computing Family, Sample D. "Schiemel J." and his connections.
The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them

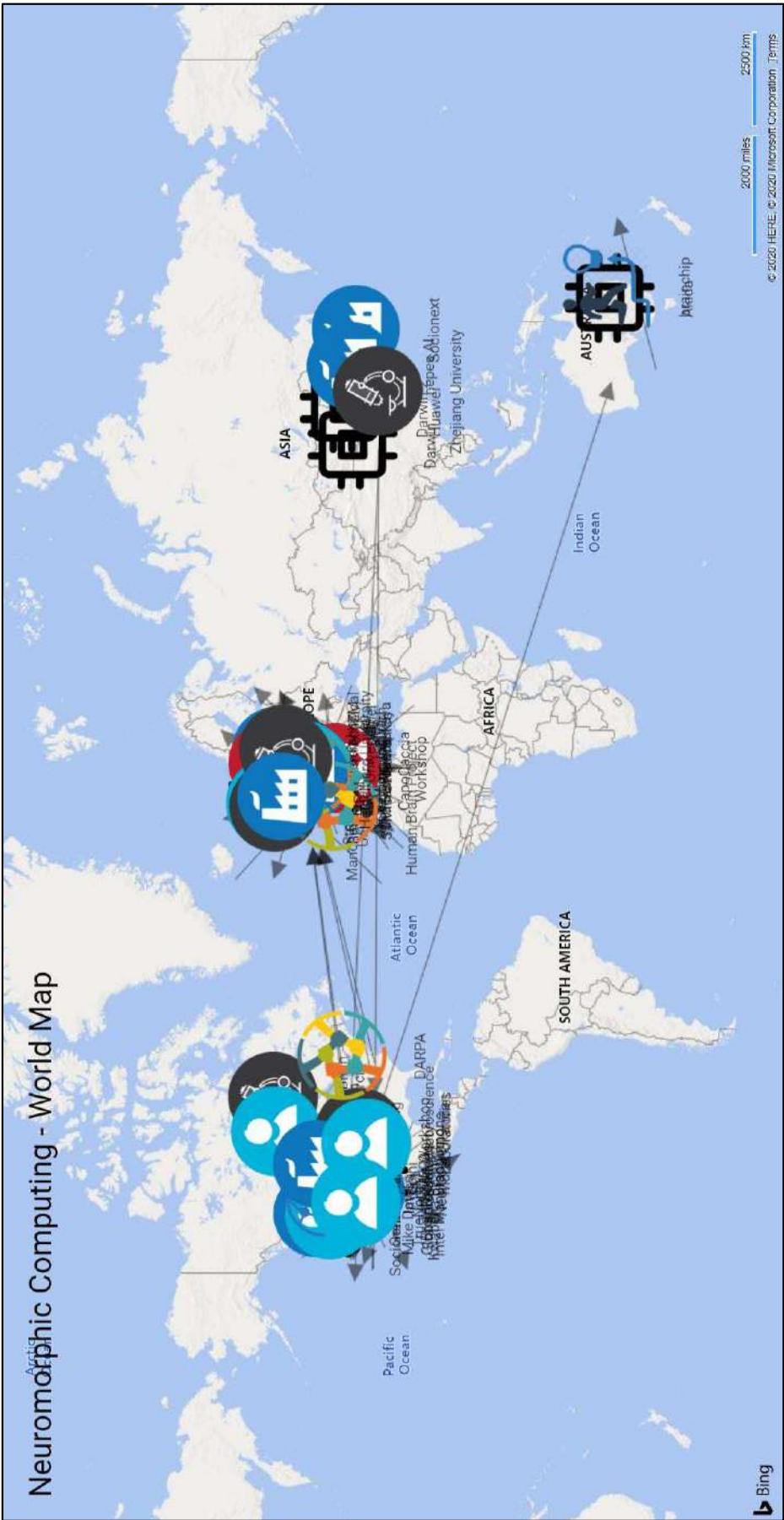


Figure 28 Neuromorphic Computing Family, Sample E "World-Map"

The detailed version can be found through interactive UPM SharpCloud. The interactive diagrams give information about each actor (can be the chip, person, event, etc.) and their connection among them.



Figure 29 UPM roadmap for Event-Based Sensors
The application availability diagram for both control with traditional chips (light-green) and control with neuromorphic chips (pink). The detailed version can be found through interactive UPM SharpCloud.

9 ANNEX 2

Tables 6 and 7 present some private companies that could be potentially interested in collaborating with HBP or exploiting neuromorphic chips.

In particular, table 6 provides basic info about the companies. Note that the table includes companies of different sizes (micro for 1-10 employees, small 10-50, medium 50-250, large more than 250)

The table 7 shows a basic assessment of the companies' exploitation potential, based on the following criteria:
(scale: 1 - minimum - to 5 – maximum)

- “Innovativeness”: novelty of the product (s) portfolio or area of work of the company.
- “Knowledge”: expertise and specialisation of the company about neuromorphic technology.
- “Economic Interest”: economic impact of the potential product for the HBP partners.
- “Effective Integration”: feasibility of the integration of neuromorphic technology into an available product (s) of the company.
- “Capacity to Reach Market”: potential of the company to connect with relevant market actors and materialize exploitation.
- “Accessibility”: commercial/collaboration dialogue with the relevant company can be easily established.

Table 6 Basic info on potentially interested Companies

Id	Company	Website	Location	Sector	Size	Area	Relation to Neuromorphics	Products
1	Neurons Inc	https://neuronsinc.com/	Denmark	Biotechnology	Medium	Neurosciences devices	Interested in robotics but want to set up partnerships with industrials. Their Machine Learning Applications can be implemented through neuromorphic chips	Neurons Inc applies neuroscience tools and insights to help business to better understand unconscious and conscious responses.
2	GoodAI	https://www.goodai-solutions.com/	Czech Republic	Industry AI Solutions	Small	Artificial Intelligence, Machine Learning, Software	The company can use our neuromorphic chips for its AI solutions.	AI solutions such as "Predictive Maintenance", "Visual Inspection", "Asset Management" or "custom" for industries (manufacturing, gas, electricity etc.)
3	SPIXII	https://www.spixii.com/	United Kingdom	AI Customer Service	Small	Chatbot, machine learning, exploring voice	Cognitive Virtual Assistant can be implemented through Neuromorphic Chips	SPIXII virtual chatbot for data analysis and task automation
4	ec2ce easy to see	https://www.ec2ce.com/	Spain	Agricultural AI Solutions	Small	Agriculture, Artificial Intelligence, Big Data, Machine Learning	The company can use our neuromorphic chips for its AI solutions.	AI solutions such as "Prediction of weekly production", "Prediction of productivity", "Pest Management" or "Optimization of fertilization", "Crop Quality", "commodity price" for agriculture
5	Think Silicon	https://think-silicon.com/	Greece	GPU producer	Small	Artificial Intelligence, Computer Vision, Internet of Things, Semiconductor	They are producing AI accelerator for end users. spINNaker might be marketed through them to increase user-size.	They sell Low-Power AI accelerator based on GPU
6	Deep learning partnership	http://www.deeplp.com/	United Kingdom	Consultancy about AI	Micro	AI, Blockchain, deep learning	They do have training of neuromorphic programming (theory to practice) As they are based in London, SpINNaker can be	Consulting Company that gives training and workshop about AI

7	CENTURY Tech	https://www.century.tech/	United Kingdom	Education	Small	Machine learning	used in training activities.
8	MBDA	https://www.mbdasystems.com/	France	Defence Systems	Large	Artificial intelligence for aeronautics, spatial and armaments industries	CENTURY is a personalised learning platform combining A.I., big data and behavioural science to provide real-time insights to educators. They are producing air missiles and recently acquired Numalis for AI solutions for smart missile systems (detect, track, intercept and destroy aerial targets.)
9	Cognicor Technologies	https://www.cognicor.com/	India	AI Customer Service	Medium	Analytics, Artificial Intelligence, Banking, Big Data, FinTech, InsurTech, Natural Language Processing	Neuromorphic chips can be used in their AI solutions (data analyses and insight from platform users) Very interested by AI and by using innovation platform as a watch tool
10	Sum&Substance	https://sumsub.com/	United Kingdom	IT Services	Medium	Artificial Intelligence, Computer Vision, FinTech	Sum&Substance is an independent developer of solutions for remote identification and verification of customers, partners and employees.
11	Cybernano	https://www.cybernano.eu/	France	Medical Data Science	Micro	in vitro assays, patients following physiologic signals interpretation (cardiomyocytes and potentially soon neurons)	Data Science for Health Innovation (the analysis of cardiac signals) (Quality-by-Design for Medical Device & Drug Development) (Biomedical Signal Processing)(automate the analysis of your Biological Data)
12	BASF	https://www.bASF.com/	Germany	Chemistry	Large	AI, robotic, system biology	Their recent agreement with TUM for integration AI solutions. The joint research work will investigate issues such as the solubility of complex mixtures or dyes as well as predicting the aging process of catalysts. Joint agreement can be done for research

13	Sundance	https://www.sundance.com/	United Kingdom	Hardware Retail	Medium	Artificial Intelligence, Wireless	They are selling embedded chips and boards for end users. SpiNNaker might be marketed through them to increase user-size.	embedded processors and boards
14	AURA Innovative Robotics	https://aurarobotix.com/	Spain	Medical devices	Small	BioTechnology, Health Care, Science and Engineering (Neurological diseases, exoskeleton)	Currently, they are engaged with OSCANN & ORTE. They will incorporate AI and OSCANN for diagnose. Neuromorphic chips can be a solution with visual detection capabilities.	OSCANN & ORTE, OSCANN is a diagnosis platform that combines the analysis of ocular and oculocephalic movements, ORTE is an innovative medical robotic solution for the diagnosis, treatment and rehabilitation of the human upper-limb, that combines a sophisticated musculoskeletal computer model of the upper-limb with a sensorized exoskeleton that helps the patient to perform the treatment prescribed by the therapist and monitors progress.
15	SimplicityBio	https://www.quartz.bio/	Switzerland	Bioinformatic	Micro	Machine learning for pharmaceutical industry	Machine learning is used for data analysis therefore neuromorphic chips can be used.	QuartzBio- Data integration, Analyse and Visualisation tool
16	Sensimed	https://www.sensimed.ch/	Switzerland	Smart Lenses	Small	Lenses Medical devices for data collection	Collected data can be used for detection of Glaucoma. Therefore, neuromorphic chips can be useful	SENSIMED Triggerfish is a smart lens device that provides insights into the ocular volume changes throughout the day and night for glaucoma treatment
17	Green Running Ltd (creators of Very)	https://very.energy/	United Kingdom	Smart Wwhite Goods	Small	Artificial Intelligence, Energy, Energy Efficiency, Energy Management, Machine Learning, Peer to Peer, Smart Building, Smart Home	Very specialises in high-speed data acquisition and AI, providing cutting-edge fault-finding technology to white goods manufacturers to make their appliances smarter and more sustainable	
18	GreenSoft	http://www.greensoft.com.ro/	Romania	Software Service Provider	Small	Data management	Their monitoring, data analyse solutions can be	GreenSoft is a software service provider that incorporates innovation and technology from

							AI to improve client competitiveness on market.
19	Guerbet	https://www.guerbet.com/	France	Pharmaceutical	Large	DM, AI for radiology	Recent Agreement with IBM Watson for medical imagining. Can be a partner in neuromorphic implementation.
20	ImaBiotech	https://www.imabiotech.com/	France	Pharmaceutical	Small	Molecular imaging, big data software	Machine learning solutions for element detection , localization and quantification can be done by neuromorphic chips
21	Seldon	https://www.seldon.io/	United Kingdom	Software Library	Micro	Machine learning	Their library can be researched on neuromorphic chips but not a potential market opportunity.
22	Bioptimize	http://www.biooptimize.com/	France	Bio Technology	Micro	AI	Developed a AI algorithm that allow to treat data set with a weak number of data. This can be implemented on neuromorphic chips
23	ASPerTice	https://www.asperdice.net/	France	Software Consultancy	Small	AI and data management	They might increase their consultancy portfolio with neuromorphic
24	Heuro Labs	http://www.heurolabs.com/	Germany	Data Science	Small	Data management	No news since 2017. A recent company, their object to have cognitive applications, robotics through data analyse. These applications can be implemented on neuromorphic chips.

Table 7 Assessment of exploitation potential

Id	Company	Usage	Current Situation	Innovativeness	Knowledge	Economic Interest	Effective Integration	Capacity to Reach Market	Accessability	Exploitation Potential
1	Neurons Inc	Use chip for AI solutions	Active	5	3	4	4	4	4	24
2	GoodAI	Use chip for AI solutions	Active	4	3	4	4	4	4	23
3	SPIXII	Use chip for AI solutions	Active	3	3	4	4	4	5	23
4	ec2ce easy to see	Use chip for AI solutions	Active	4	3	4	4	3	4	22
5	Think Silicon	Distribution	Active	4	3	4	4	3	4	22
6	Deep learning partnership	Workshops, Training	Active	2	5	2	5	2	5	21
7	CENTURY Tech	Use chip for AI solutions	Active	4	2	3	4	4	4	21
8	MBDA	Use chip for AI solutions	Active	4	1	5	3	4	4	21
9	CogniCor Technologies	Use chip for AI solutions	Active	3	3	4	4	3	4	21
10	Sum&Substance	Use chip for AI solutions	Active	4	2	4	3	4	4	21
11	Cybernano	Use chip for AI solutions	Active	4	2	3	3	4	4	20
12	BASF	Use chip for AI solutions	Active	3	2	4	2	5	4	20
13	Sundance	Distribution	Active	2	2	4	4	3	4	19
14	AURA Innovative Robotics	Use chip for AI solutions	Active	4	2	3	3	3	4	19
15	SimplicityBio	Use chip for AI solutions	Active	3	2	3	3	4	4	19
16	Sensimed	Use chip for AI solutions	Active	4	2	3	2	3	4	18
17	Green Running Ltd (creators of Verv)	Use chip for AI solutions	Active	4	2	4	2	3	3	18
18	GreenSoft	Use chip for AI solutions	Active	3	2	3	2	4	4	18
19	Guerbet	Use chip for AI solutions	Active	3	1	3	2	4	4	17
20	ImabioTech	Use chip for AI solutions	Active	3	2	4	3	3	2	17
21	Seldon	Use chip for AI solutions	Active	3	2	2	2	4	4	17
22	Bioptimize	Use chip for AI solutions	Active	3	2	3	2	3	4	17
23	ASPeritise	Use chip for AI solutions, Workshops, Training	Active	2	2	2	2	2	5	15
24	Heuro Labs	Use chip for AI solutions	Active	2	2	2	2	2	3	13