



life.augmented

Deploying Artificial Intelligence at the Deep Edge

Valeria Tomaselli
STMicroelectronics

Agenda

1

Introduction to Artificial Intelligence

2

Artificial Intelligence @the Cloud: limitations

3

Artificial Intelligence @the Edge

3

IoT domain

4

Automotive domain

5

Conclusions

Artificial Intelligence

AI is a superset of all the studies where machines mimic cognitive capabilities like humans



Analysis

Where am I?

Scene classification (audio, video, environmental sensors)

Which objects are in the scene, where are they?

Video object detection/classification

What am I doing?

Activity recognition (audio, video, inertial sensors)

What's happening?

Event recognition (audio, video, inertial sensors, environmental sensors)



User Interaction

- **Command detection** (audio)
- **Speech Recognition** (audio)
- **Gesture Recognition** (inertial sensors, video)
- **User identification and mood detection** (audio, video)



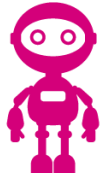
Continuous Learning

How can I detect unpredictable, unclassified events in dynamic environments?

Recurrent networks (audio, video, inertial sensors, environmental sensors)



Smart Industry



Smart Home



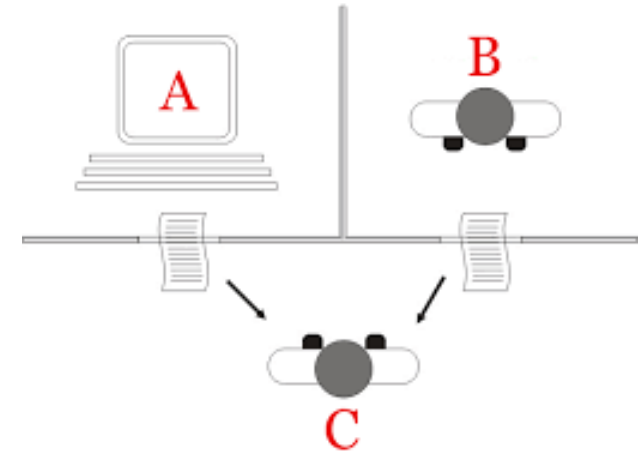
Smart City



Smart Things

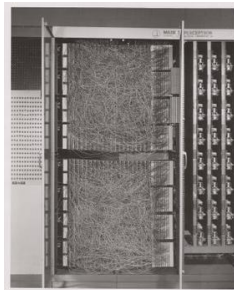
Milestones in Artificial Intelligence

- Computing Machinery and Intelligence *Mind* 49: 433–460
- “I propose to consider the question: can machines think?”
- Because thinking is difficult to define, "Are there imaginable digital computers which would do well in the imitation game?"
- “The imitation game could then be played with the machine in question... mimicking digital computer and the interrogator would be unable to distinguish them”



Turing Machine

1936



Rosenblatt's Perceptron

1957



Cybenko's Universal Approximation Theorem

1989



Brin's Winter of AI

1990



Hinton's Turning point on Science

2006



Google AI Assistant

2018



Humanoid Robots

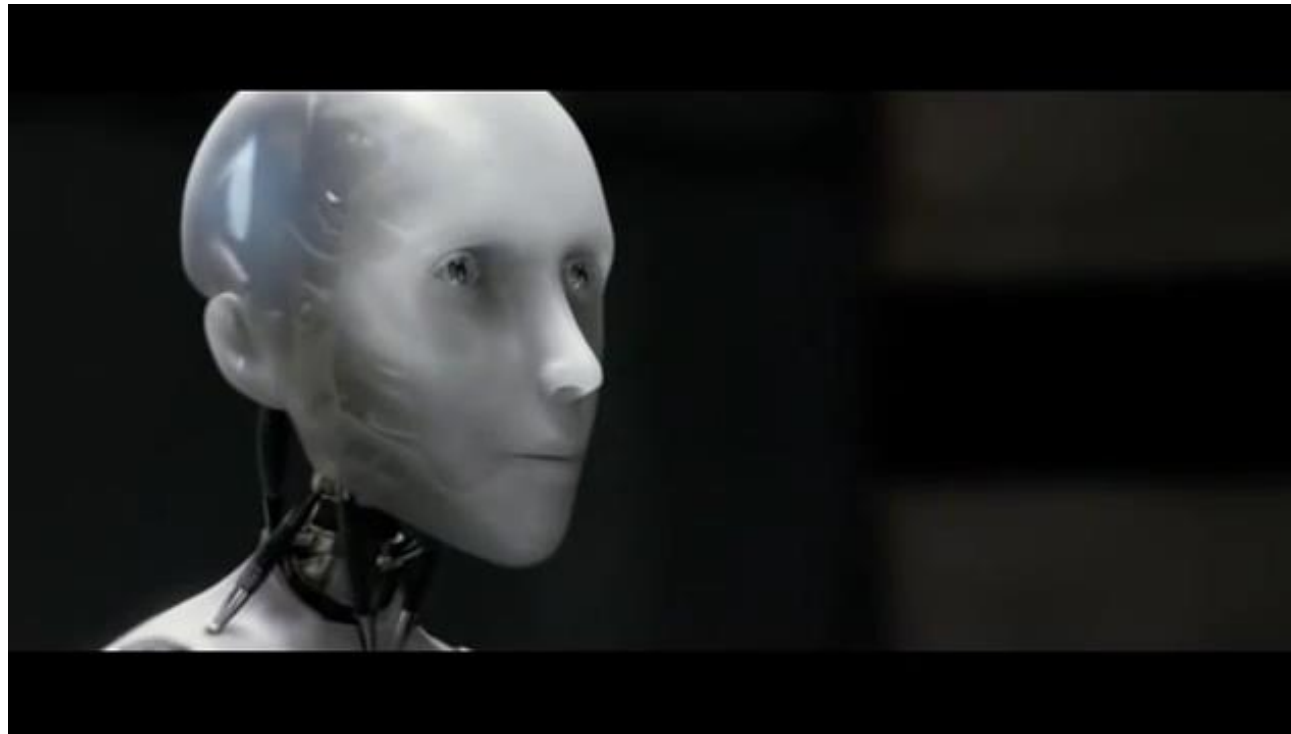
2019

Strong Artificial Intelligence

The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds

*John Searle*⁷⁻³¹⁻¹⁹³²

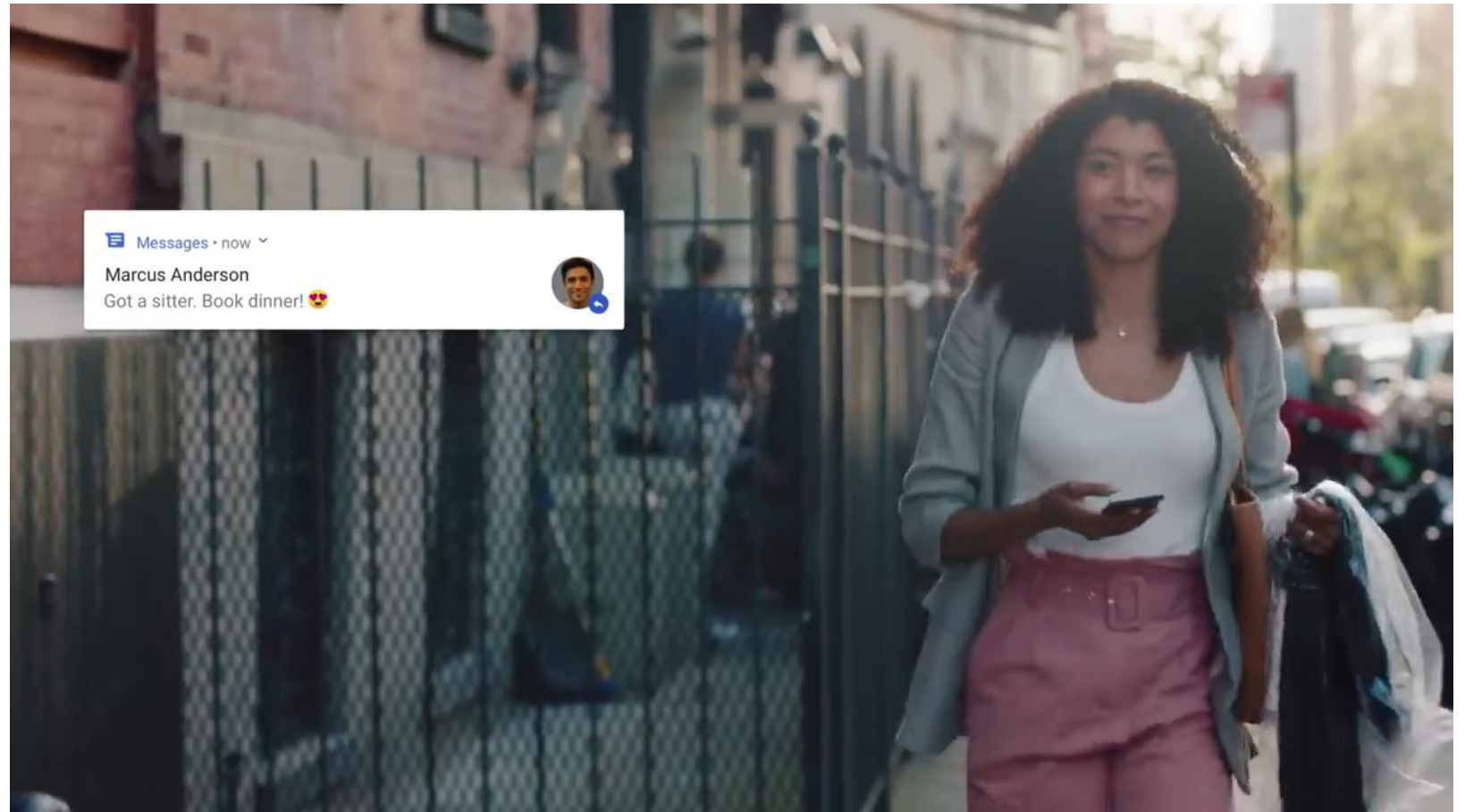
American philosopher. He is currently Willis S. and Marion Slusser Professor Emeritus of the Philosophy of Mind and Language and Professor of the Graduate School at the University of California, Berkeley.



*I, robot*²⁰⁰⁴

Weak Artificial Intelligence

- Focused on one narrow task
- Google Assistant



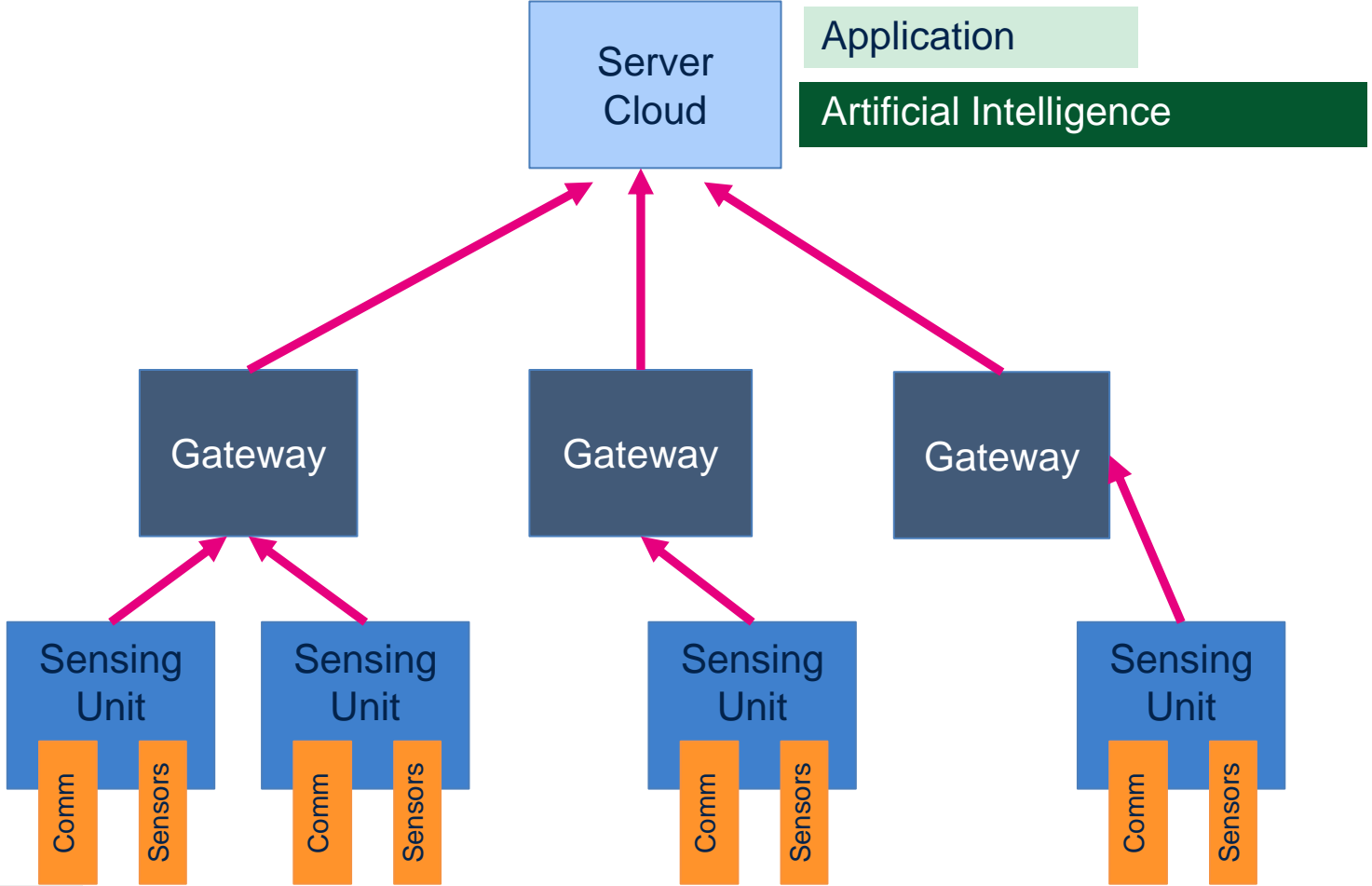
<https://www.youtube.com/watch?v=-qCanuYrR0g>

Artificial Intelligence @the Cloud limitations



Intelligence must not be Centralized

- ↓ Scalability
- ↑ Latency
- ↓ Units
- ↑ Energy
- ↑ Intrusions
- ↑ Bandwidth



Cyber

Object

Physical

CPS Domains



Voice Recognition choking the Cloud

- Average Person**
- 16,000 utterances ¹/day @ 163 words/minute ²
 - ≈ 98 minutes speech/day



Audio coding @ 128Kbps ≈ 94 MB/day



1 Million People ≈ 94 TB, 98 minutes in a day



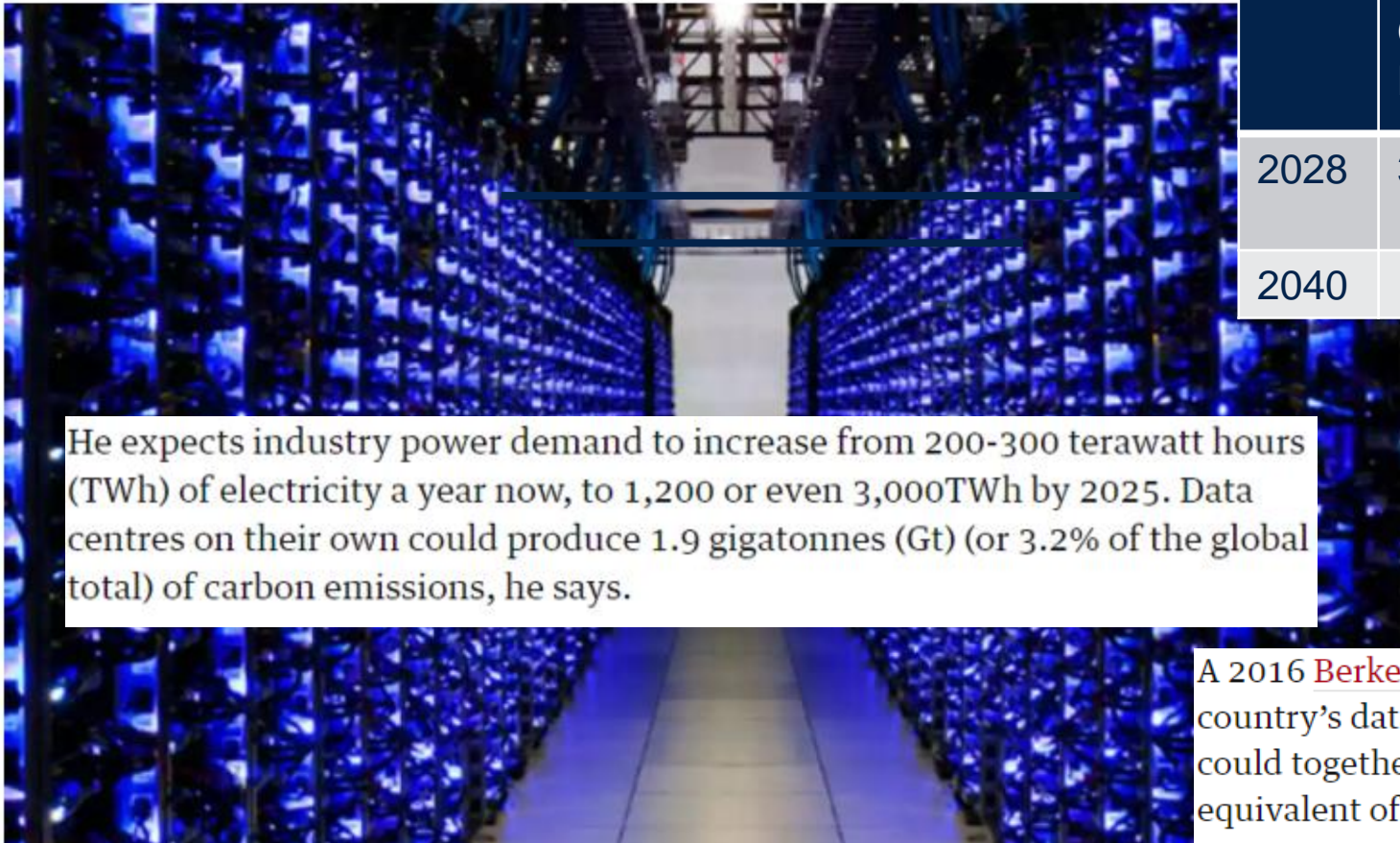
3.8 Billion
People by 2021

≈ 11 PB
3 minutes in a day

'Tsunami of data' could consume one fifth of global electricity by 2025

<https://www.theguardian.com/environment/2017/dec/11/tsunami-of-data-could-consume-fifth-global-electricity-by-2025>

Billions of internet-connected devices could produce 3.5% of global emissions within 10 years and 14% by 2040, according to new research, reports [Climate Home News](#)



Year	IoT Global Carbon Emissions
2028	3.5%
2040	14%

Year	Power demand datacenters
2020	Up to 100 TWh (only USA)
2025	600 TWh

He expects industry power demand to increase from 200-300 terawatt hours (TWh) of electricity a year now, to 1,200 or even 3,000TWh by 2025. Data centres on their own could produce 1.9 gigatonnes (Gt) (or 3.2% of the global total) of carbon emissions, he says.

A 2016 [Berkeley laboratory report](#) for the US government estimated the country's data centres, which held about 350m terabytes of data in 2015, could together need over 100TWh of electricity a year by 2020. This is the equivalent of about 10 large nuclear power stations.

Artificial Intelligence @the Edge



Artificial Intelligence at the Edge

Moving part of intelligence closer to the data acquisition



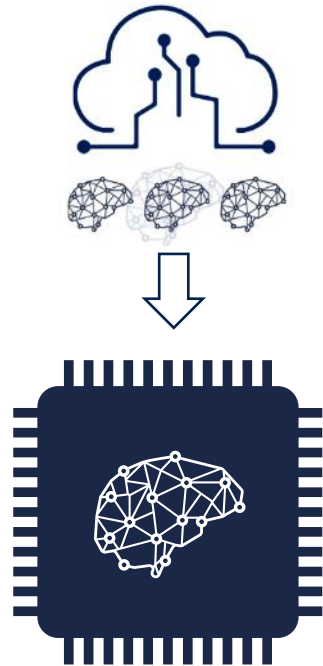
Better user experience



Realtime, no latency



Reliable



Add **new functions** and **services** with **Embedded AI**

Optimized Cloud usage








Privacy by design (GDPR compliant)

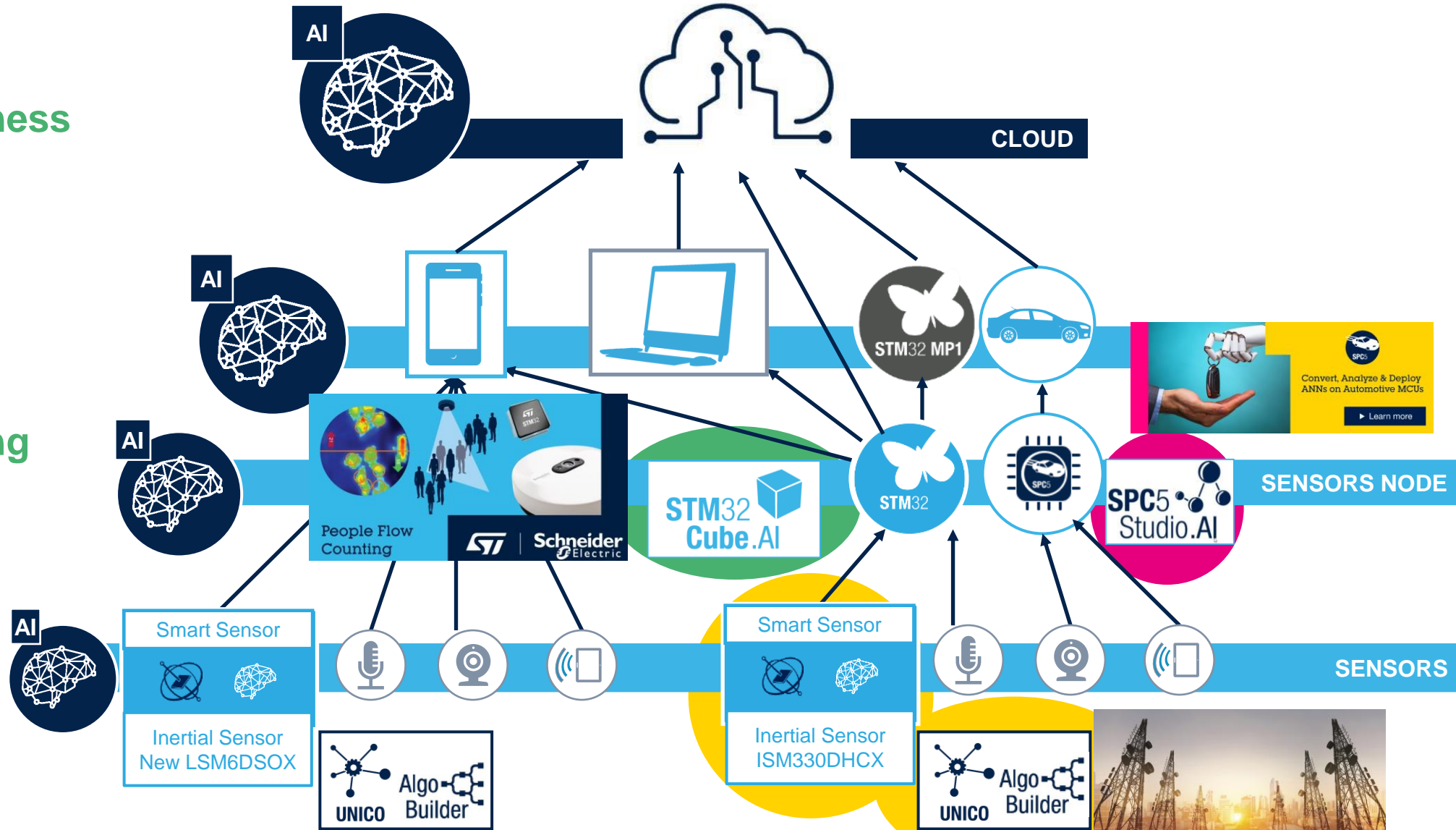


Sustainable on energy



eAI@ST: Distributed Intelligence & ST Ecosystem

-  Responsiveness
-  Bandwidth
-  Privacy
-  Security
-  Energy Saving

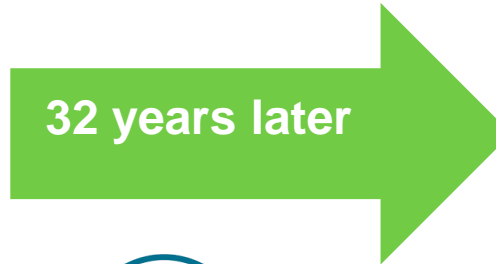


Computers vs embedded μ controllers

Olivetti M24¹⁹⁸³



- Intel 8086
- 8 MHz
- 128 KB RAM
- 16 KB ROM
- 1.84 W
- 360 \$



10x

more computational
power

100x

more efficient and
cheap

STM32 L4²⁰¹⁵



- STM32 MCU L4
- 80 MHz
- 128 KB RAM
- 1 MB Flash
- < 20mW
- < 4 €



IoT Domain

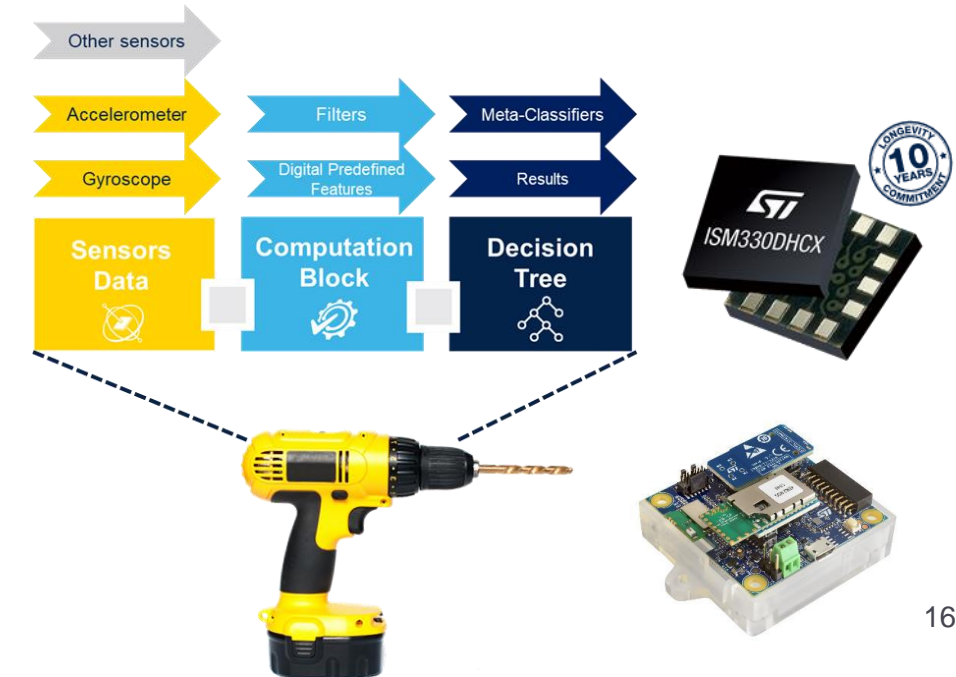
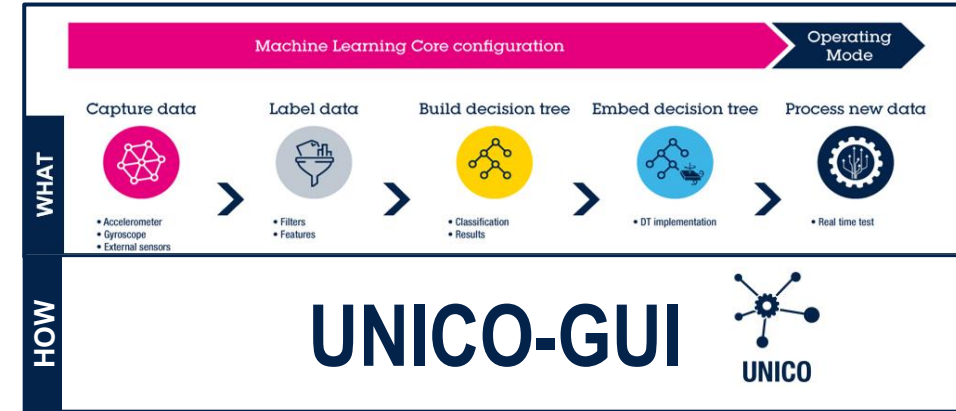


Drilling Machine classification using MLC ISM330DHCX @ STWIN

- 1 Idle (1)
- 2 Drilling (5)
- 3 Screw Driving (9)



life.augmented

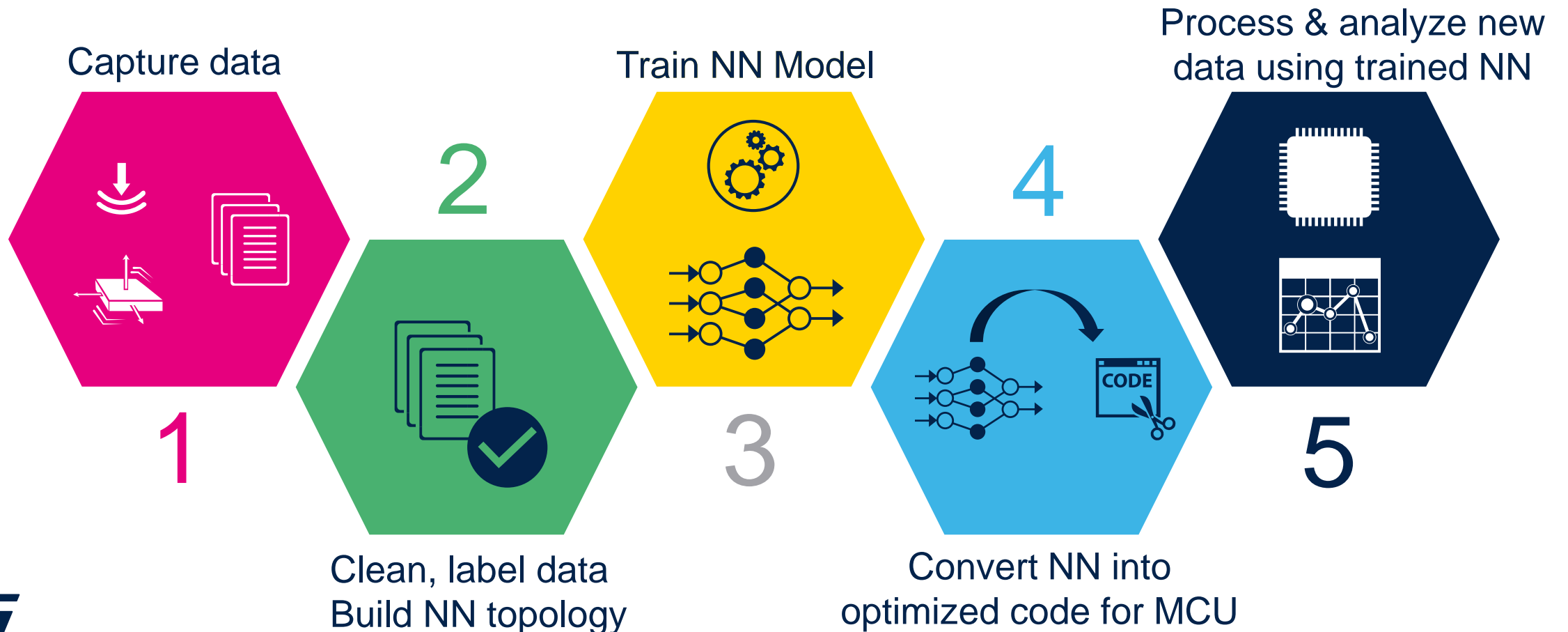


Short Command recognition on STM32L4

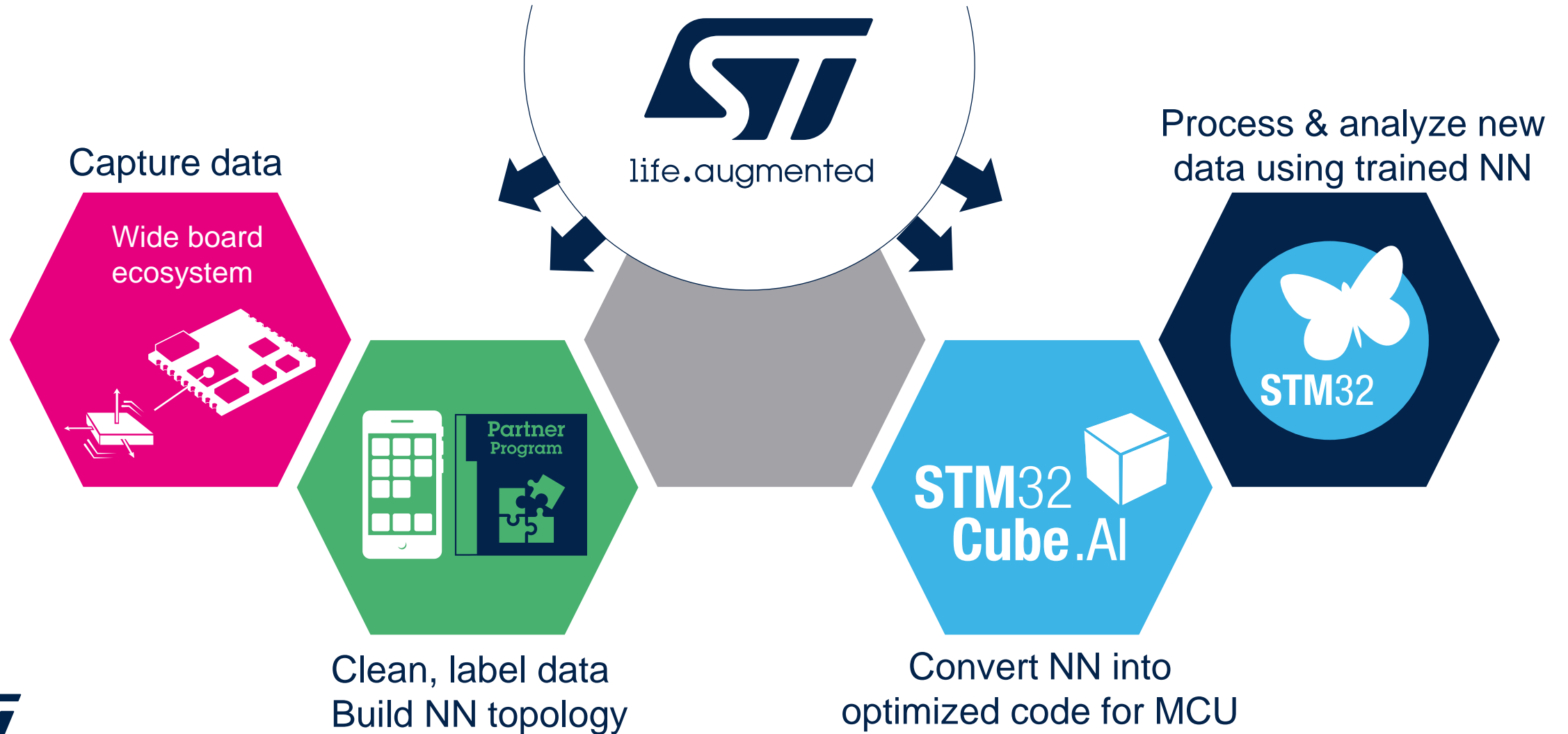
- **AI LOCAL Processing ONLY** thanks to:
 - **Audio Activation Word** for system wake-up followed by **Short Command Recognition on STM32L476JG@80MHz** with 1MB Flash, 128KB RAM present in
 - **ST SensorTile**, including also 1 microphone, inertial and environmental sensors, and BTLE module
- The recognized command is sent to the **BlueCoin Actuation board on Robot** via BTLE to enable the Robot action.



The key steps behind Neural Networks



ST toolbox for Neural Networks



AI Solutions on STM32

A full development ecosystem to create AI applications



AI extension for STM32CubeMX to map pre-trained Neural Networks



STM32 Community with dedicated Neural Networks topic and AI expert partners



Trainings, hands on, MOOCs and partners videos



FP-AI-VISION1



Person presence detection
Food classification



FP-AI-SENSING1



People activity recognition
Audio scene classification



FP-AI-NANOEDG1



Condition-based monitoring



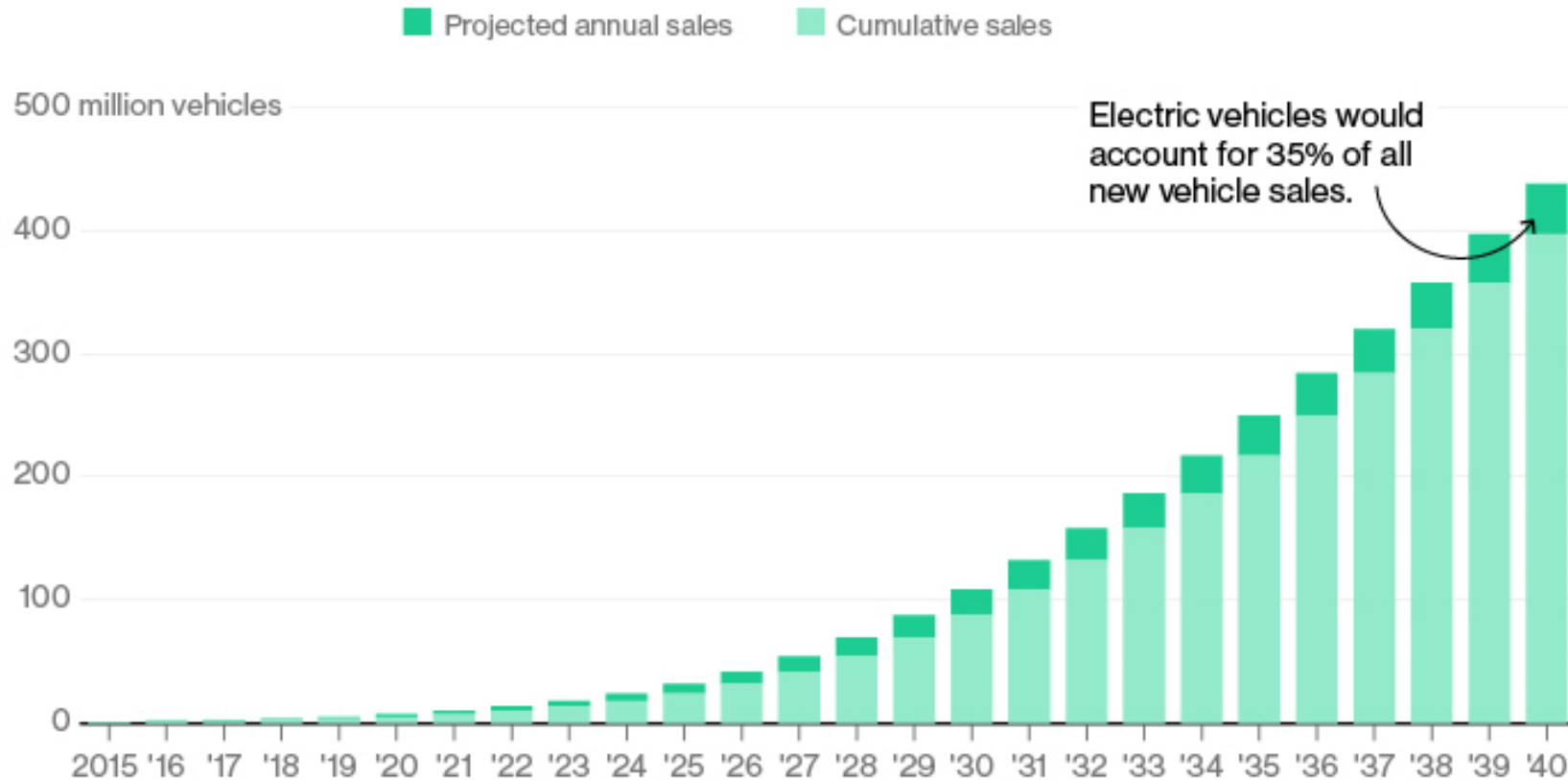
Automotive Domain



life.augmented

The rise of electric cars

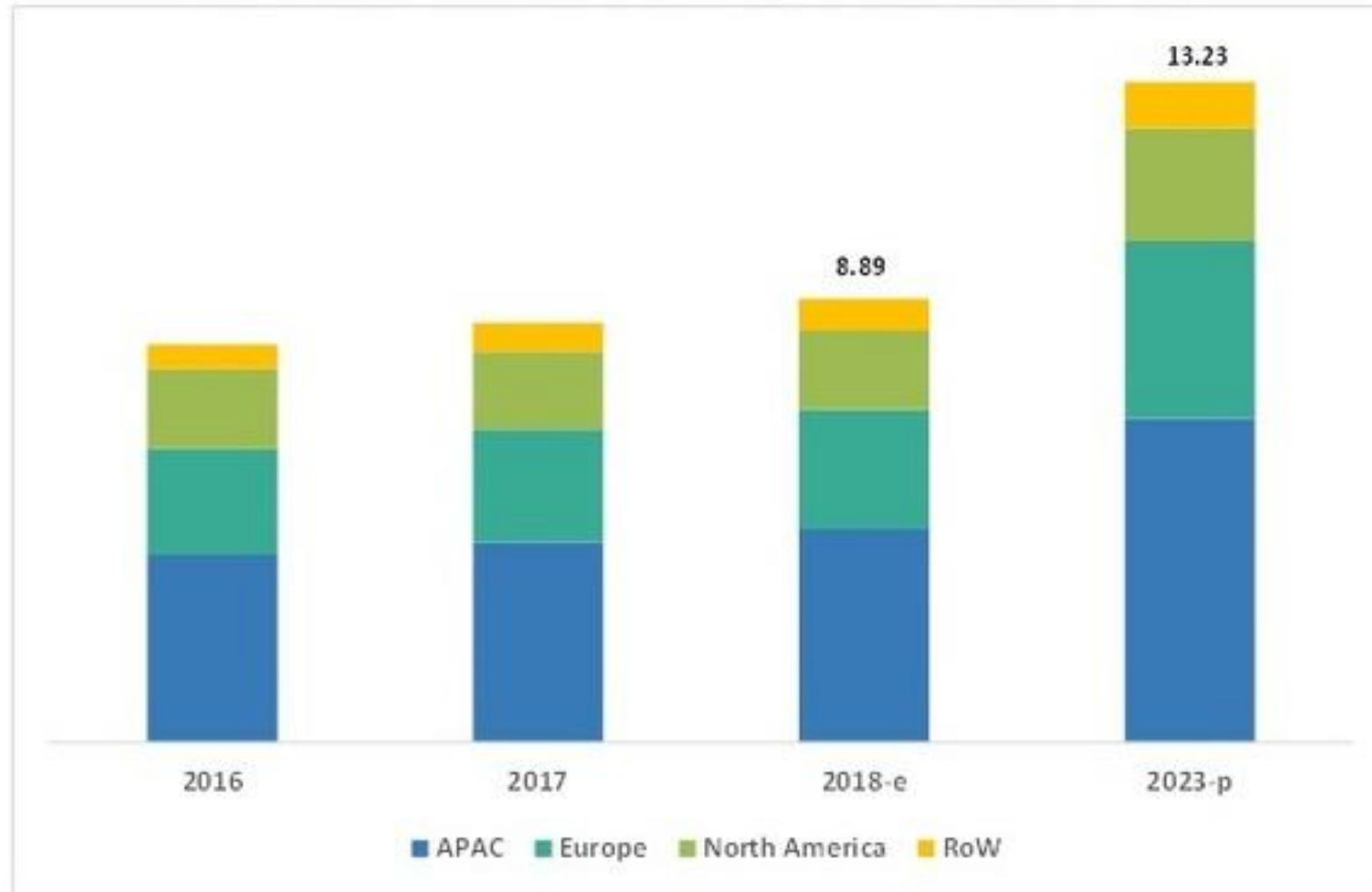
By 2022 electric vehicles will cost the same as their internal-combustion counterparts. That's the point of liftoff for sales.



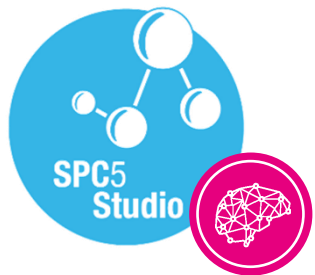
Sources: Data compiled by Bloomberg New Energy Finance, Marklines



Automotive microcontrollers market by region 2023, USD billion



Source: MarketsandMarkets Analysis

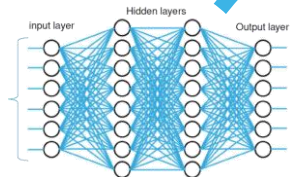


SPC5 STUDIO.AI

Artificial Intelligence toolchain for SPC5

Design & pre-train

your neural network



Analyze & optimize!



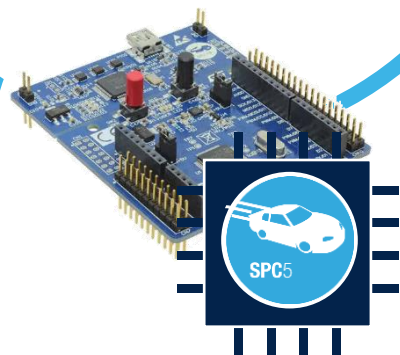
Convert

automatically to C code

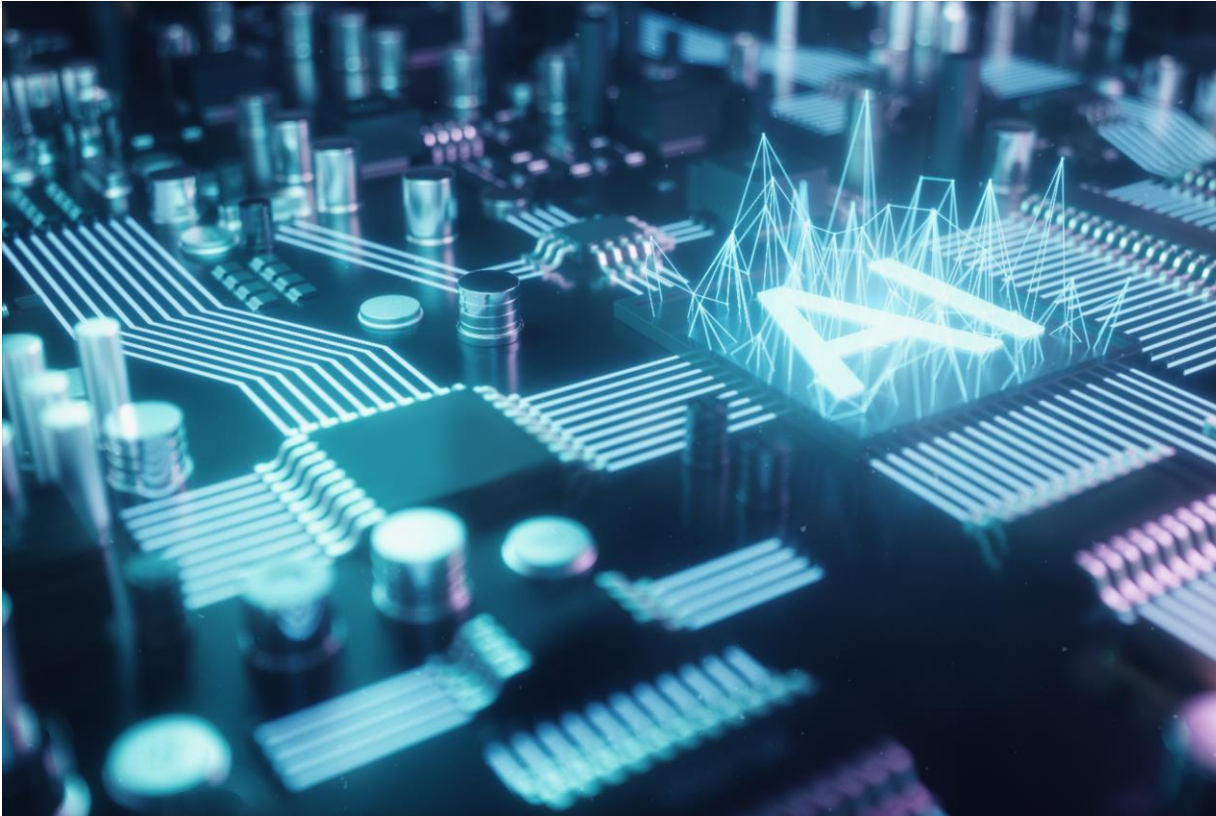
Analytics included!

Process

Execute on
SPC5 Chorus MCU



Conclusions



- Artificial intelligence is **highly pervasive** in all market segments
- The implementation of AI solutions **leads to improved quality and increased efficiency**
- **A distributed approach** to Artificial Intelligence **leads to more secure, private, energy-efficient, prompt solutions** close to where events are generated
- Artificial Intelligence solutions **can be already enabled in the Sensors and in Microcontrollers efficiently and easily** thanks to specific tools and development boards.

Thank you

© STMicroelectronics - All rights reserved.

ST logo is a trademark or a registered trademark of STMicroelectronics International NV or its affiliates in the EU and/or other countries.

For additional information about ST trademarks, please refer to www.st.com/trademarks.

All other product or service names are the property of their respective owners.



life.augmented