Making edge AI a reality

Products & enabling technology to accelerate edge AI adoption

How intelligent things are evolving

Smart

Autonomous decision making

Energy efficient

Personalization

Enhancing safety and security

Connected & secure

Multimodal connectivity – enabling high bandwidth to and from the cloud

Identify, authenticate & protect data - Within device, device to device, device to cloud

Transformative

Seamless user experience

Changing work & lifestyle

Enhancing health & wellness

What is artificial intelligence

The ability of **machines/objects** to perform tasks that typically require a **human intelligence.**

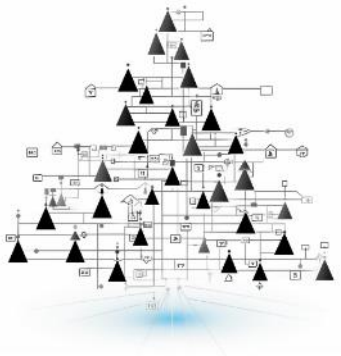
Perception

Reasoning

Learning

Decision making

From a **technical** perspective, AI is another way to program a machine, enabling it to extract meaningful information from data.

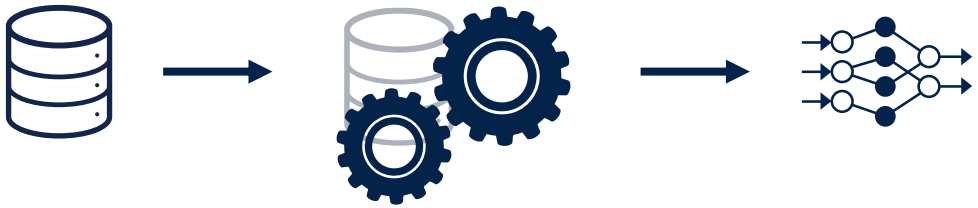


Using AI algorithms implies two phases

Training

AI models are produced using historical **datasets** and a training engine/framework.

Dataset → Training → Model



Inference

Model deployment

Inference is the process of using a trained machine learning model to make predictions or decisions based on new, unseen data.

Input data → Model & engine → Result

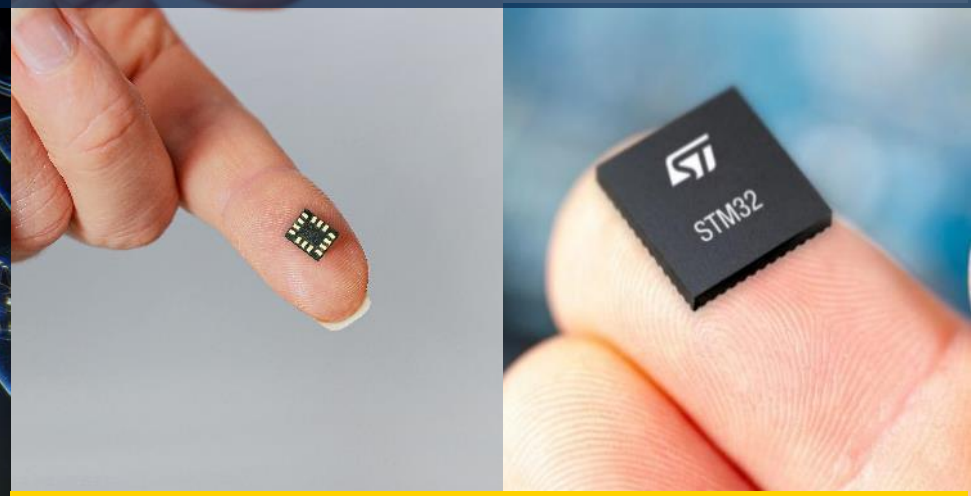
The explosion of AI-enabled devices is accelerating the inference shift from the cloud to the tiny edge

Inference **in the cloud** → Inference **on the device** → Inference **at the edge**

Cloud-centric

On device-centric

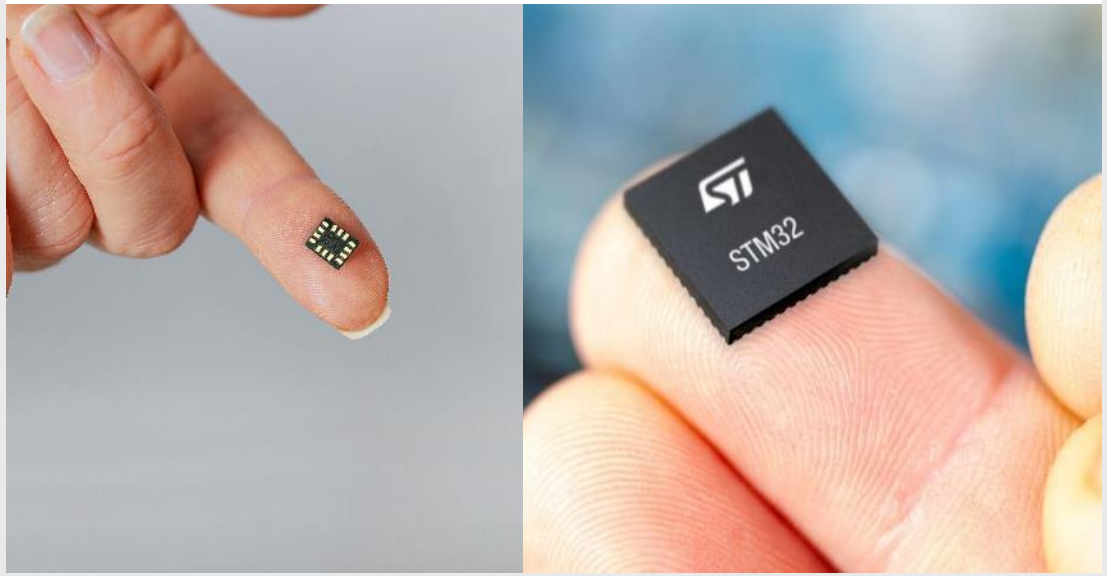
Tiny edge-centric



Enabled by a different class of hardware and software. Making AI more sustainable.

Inferring at the edge brings substantial benefits

01
10 **Reduced data transmission to generate meaningful information**



Ultralow latency
Real-time applications

Enhanced privacy and security
No data sharing in the cloud

Sustainable on energy
Low data, low power

Lower cost of inference to enable a new class of operations

Opening a new range of embedded applications

Upgrade existing devices with AI-based services

Arc fault detection

Predictive maintenance

Battery management

Reduce bom by shifting from mpus to mcus

People detection

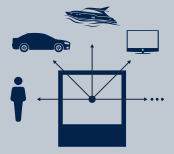
Sound analysis

Speech recognition

Offer advanced services without cloud costs

Object classification

Pose & gesture estimation



Object segmentation

A few KOPS → Tens of TOPS

Edge AI on microcontrollers and smart sensors is a key enabler for smart applications

Cities & buildings

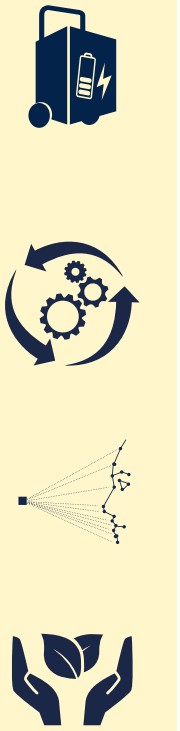
Home appliances

Industry 5.0

Healthcare & biosensing

Personal electronics

Smart mobility



- Battery management
- Arc fault detection
- Face & object recognition
- Gesture & motion recognition
- Safety | Access control
- Environmental monitoring
- Anomaly detection
- Predictive maintenance
- Energy management

and many more…

AR-Glasses
10 hours
On single-charge

Meta-Bounds 莫界

Panasonic

Virtual sensor
Tire pressure measured through the e-motor current consumption

~15-40%
Energy saving per washing cycle

Predictive maintenance
Multisensors and learning on device

oXYtronic

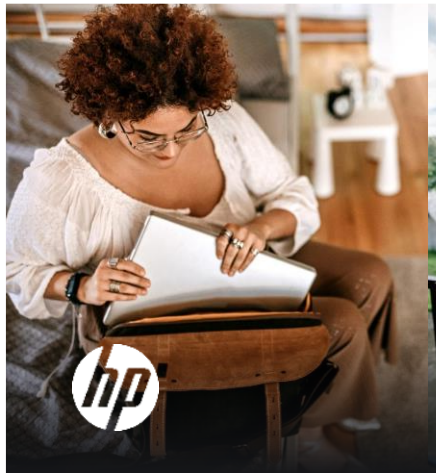
Electrical arc detection
Enhanced accuracy to avoid false detection

Computer vision on MCU
Cost-effective, compact, and low-power solution

AUTOTRAK™
MORE THAN JUST TRACKING

Virtual sensor
Measuring motor internal temperature based on external data

HPEGROUP

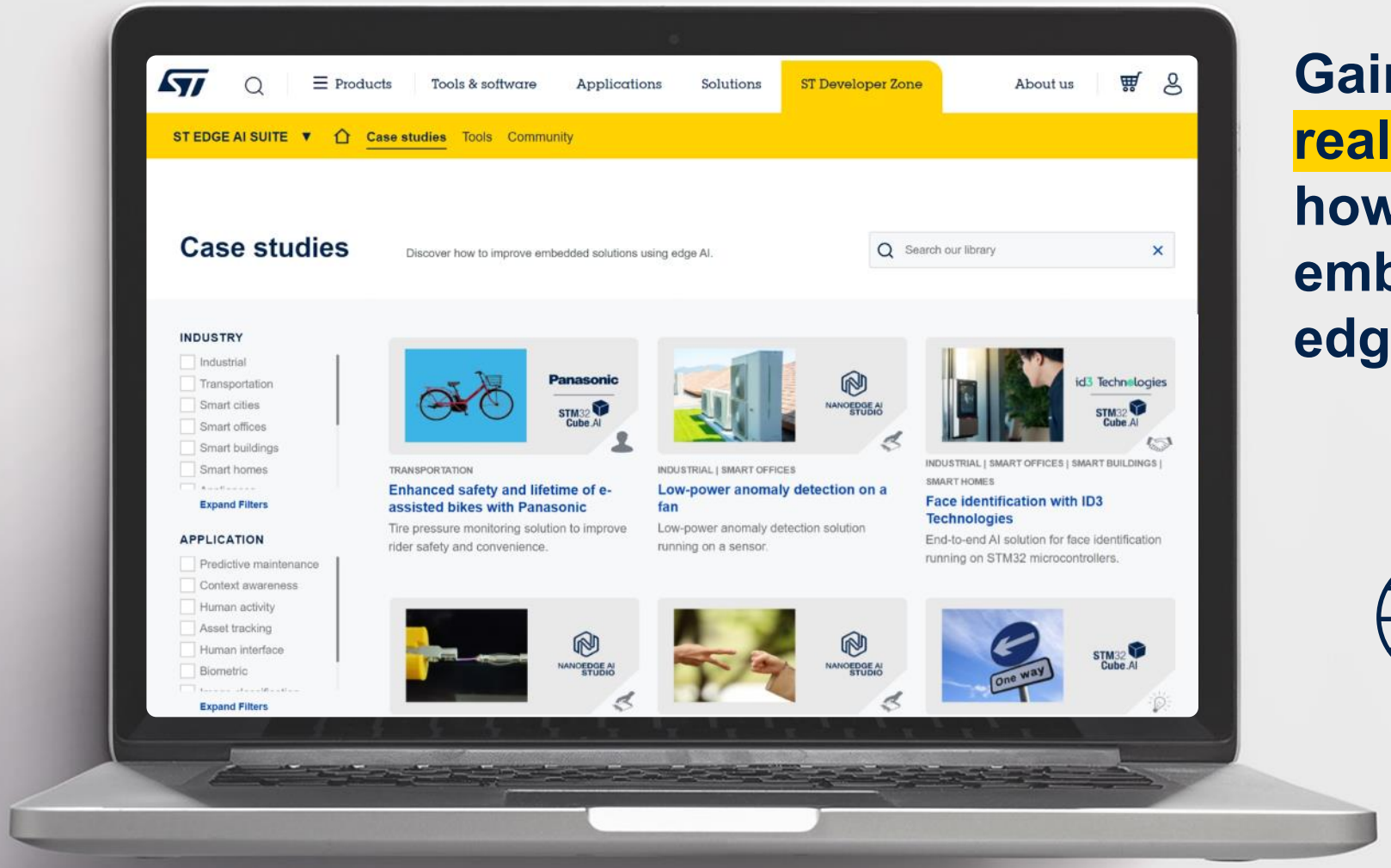


Smart context detection
In/out bag - on table

SOXAI

10+ hours
Battery-life with low-power sensor

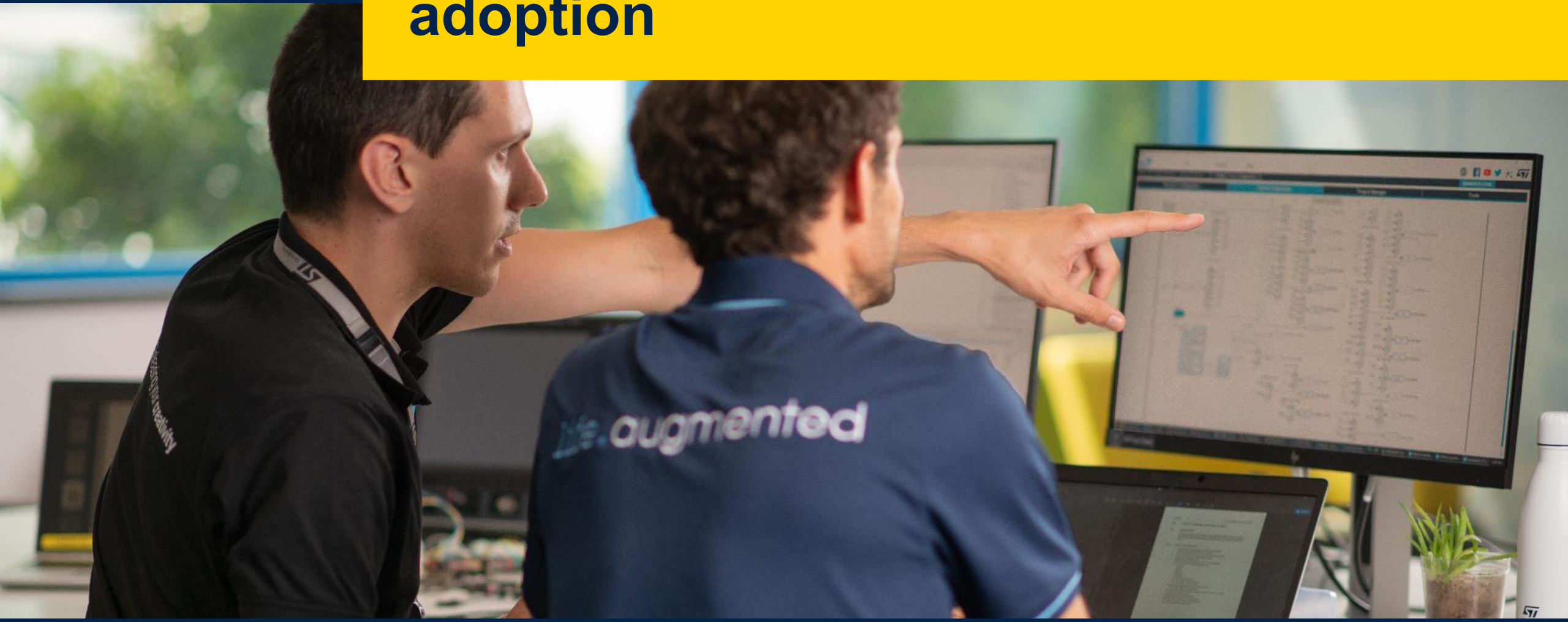
Explore inspiring case studies



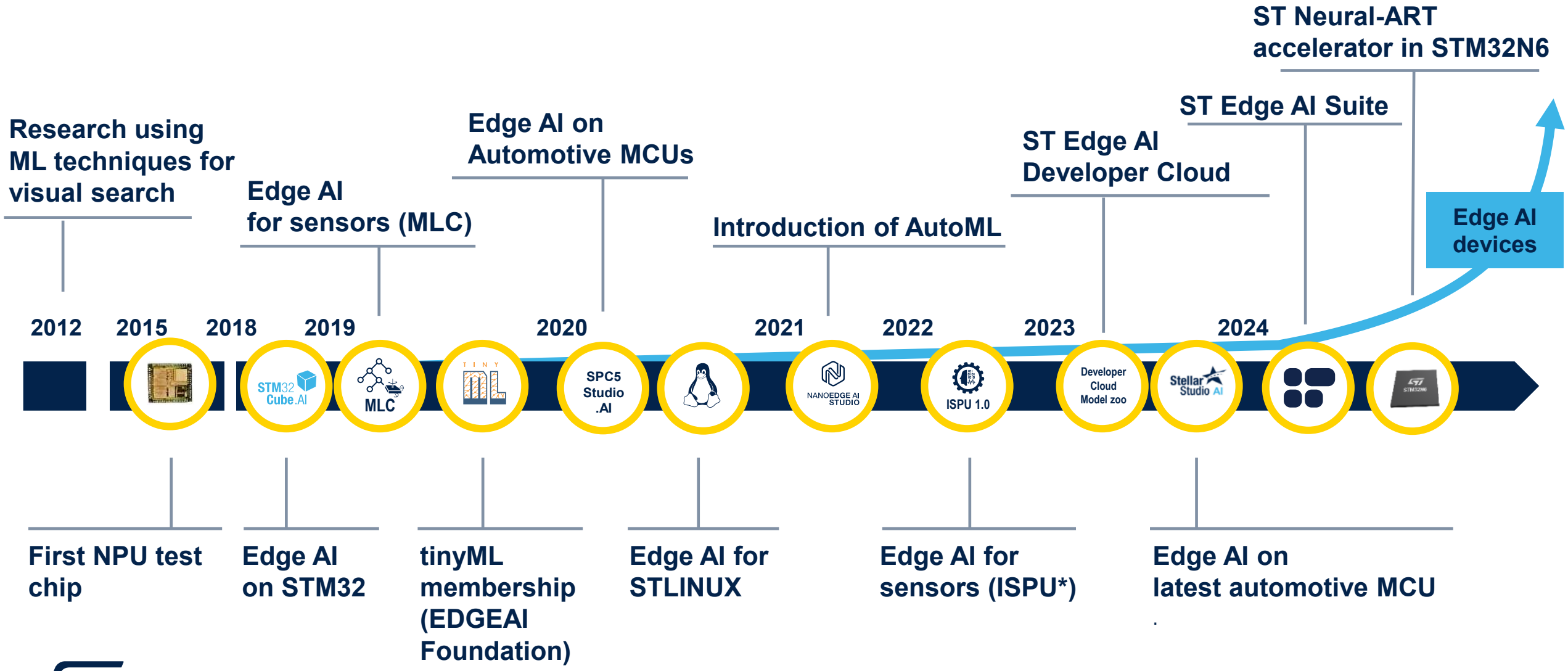
Gain valuable insights from real-world case studies on how to improve your embedded solutions using edge AI.

Edge AI case studies

How ST is accelerating edge AI adoption



10+ years of research, development, and deployment

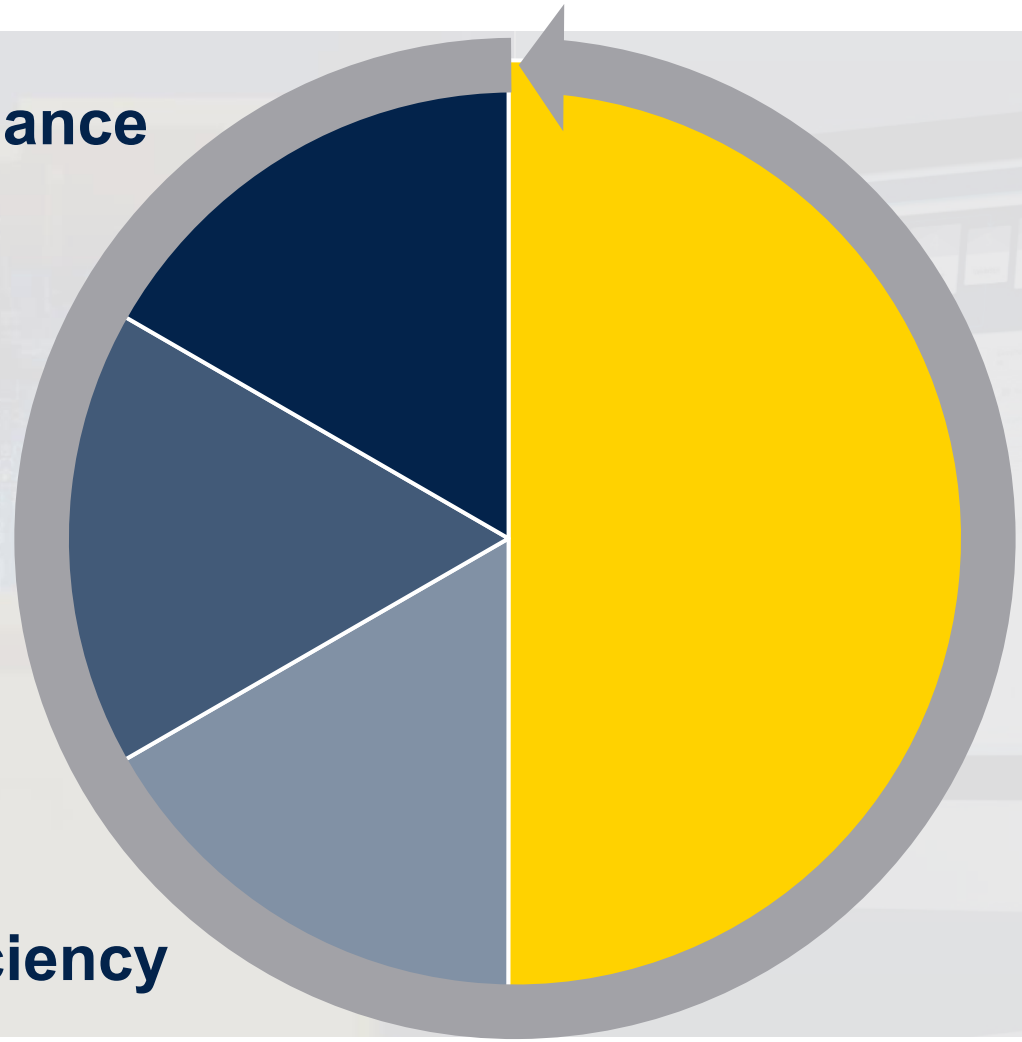


*Intelligent sensor processing unit

Taking up the edge AI developer's challenge

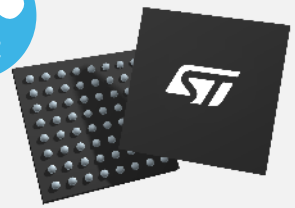
Hardware

Software



A comprehensive approach to help developers accelerate their product transformation

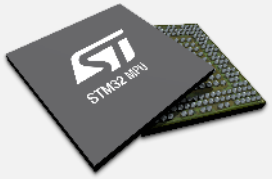
STM32



STM32 MCUs

AI

STM32N6 MCU



STM32 MPUs

AI

STM32MP2 MPU

Stellar

SPC5

Stellar MCUs

SPC5 MCUs

Sensors

AI

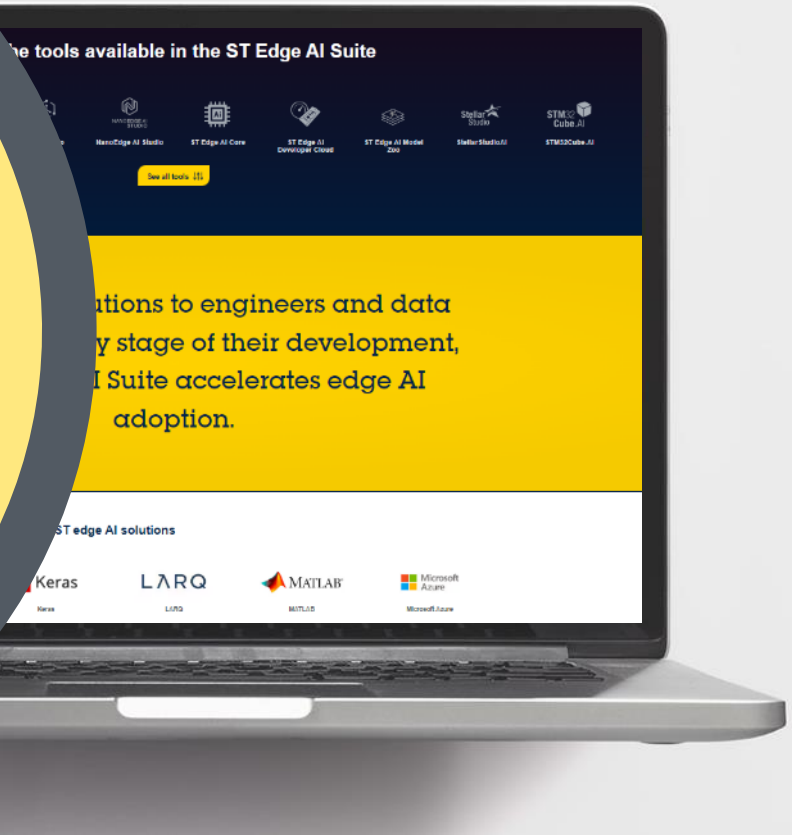
AI

MEMS with **MLC and ISPU**

Hardware platforms

Software tools

ST Edge AI Suite



Click on the products to find out more

AI With embedded AI acceleration

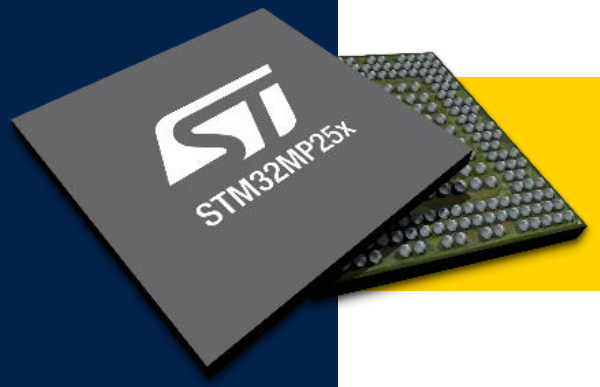
Bringing advanced compute capabilities to enable a new class of embedded machine learning applications

AI

Embedded AI hardware accelerator

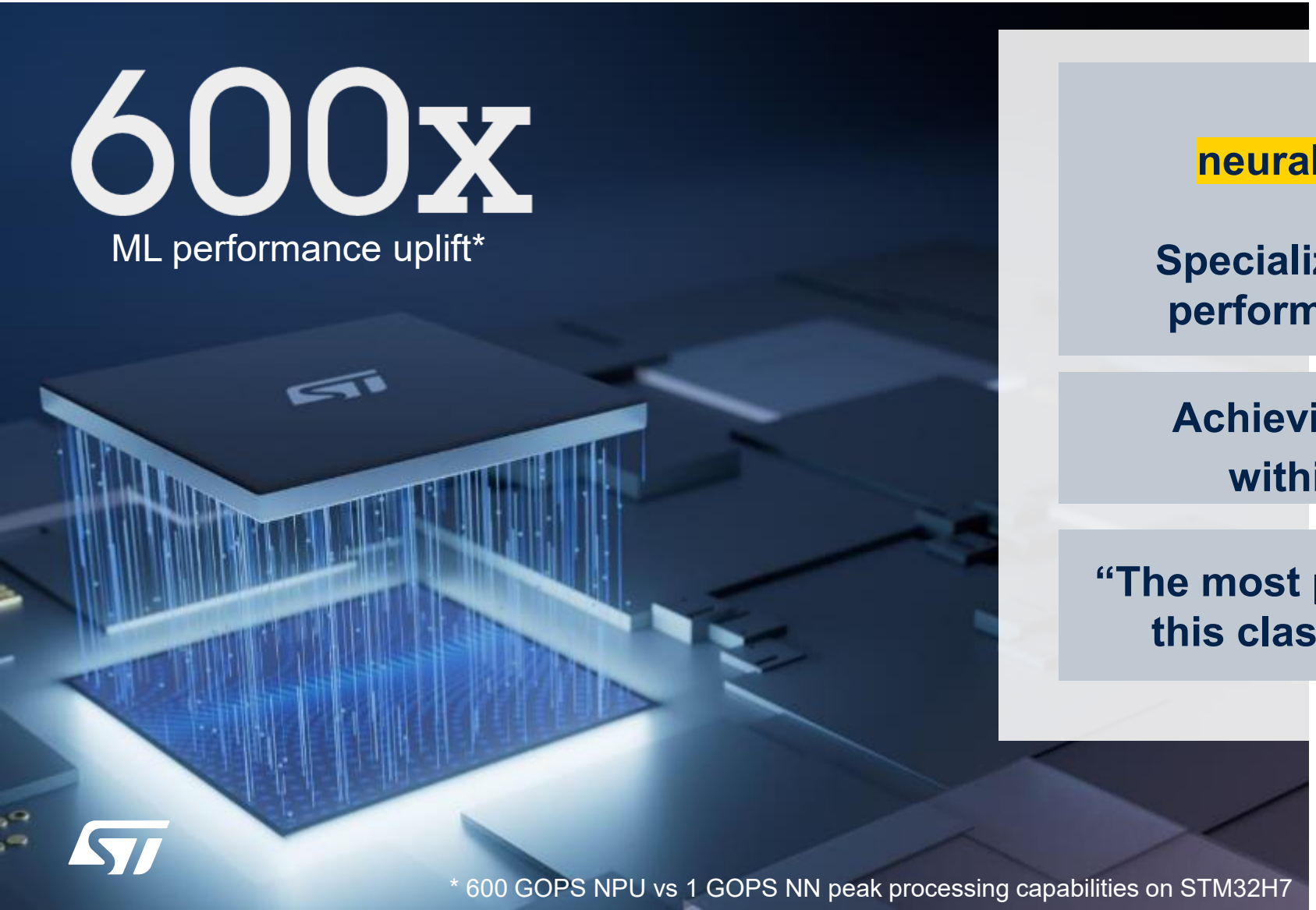
Sensors enabled with a machine learning core or an intelligent sensor processing unit (ISPU) for in-sensor AI.

STM32N6 with proprietary ST Neural-ART Accelerator for unmatched levels of performance in an MCU.



STM32MP2 bringing neural acceleration to the embedded Linux world.

First version of Neural-ART Accelerator embedded inside STM32N6



ST proprietary neural processing unit (NPU)

Specialized hardware accelerator performs advanced AI inference

Achieving edge AI performance within MCU power budget

“The most powerful accelerator seen in this class of device so far.” *EETimes*

Read more in the blog

Read our whitepaper

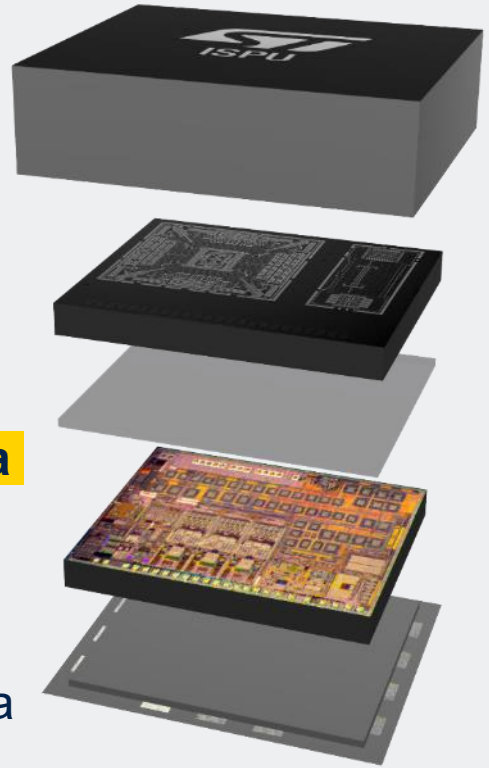
* 600 GOPS NPU vs 1 GOPS NN peak processing capabilities on STM32H7

In-sensor AI for real-time processing
The intelligent sensor processing unit

Raw data from MEMS sensor

ISPU processes data locally (a few µW)

Only meaningful data



MEMS sensor
3-axis accelerometer & 3-axis gyroscope

Sensor hub
to collect data from additional external sensors (up to 4)

ISPU core

Standard package: 2.5 x 3 mm^2

Read our whitepaper

In-sensor AI for real-time processing
The machine learning core (MLC)

MLC is an in-sensor classification engine based on decision tree logic

Motion sensing

External sensors

Sensor with machine learning core

Sensor data
- Accelerometer
- Gyroscope

Computation block
- Filters
- Digital predefined features

Decision tree
- Metaclassifier
- Results

Main processor

MLC is able to **increase accuracy** with a **better context detectability**, **offloading the main processor** while the built-in sensors identify motion data

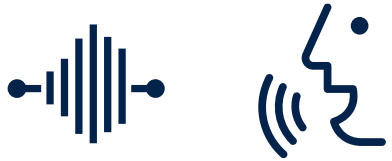
Enabling products and software

STM32 MCU offering for running AI

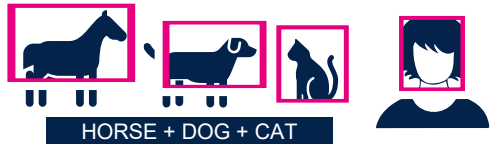
Enabling major edge AI technologies

SENSING (TIME SERIES)

AUDIO



VISION



Software tools for any user profile

SENSING

NANOEDGE AI STUDIO

User-friendly AutoML tool for STM32 MCUs

SENSING

AUDIO

VISION

ST Edge AI Developer Cloud

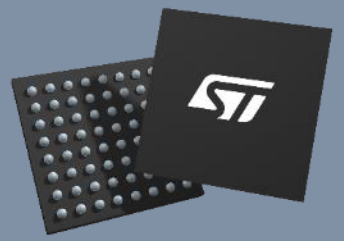
STM32 Cube.AI

ST Edge AI Core

AI model benchmark, optimizer and code generation for STM32 MCUs

ST Edge AI Suite

Large choice of general purpose & accelerated hardware

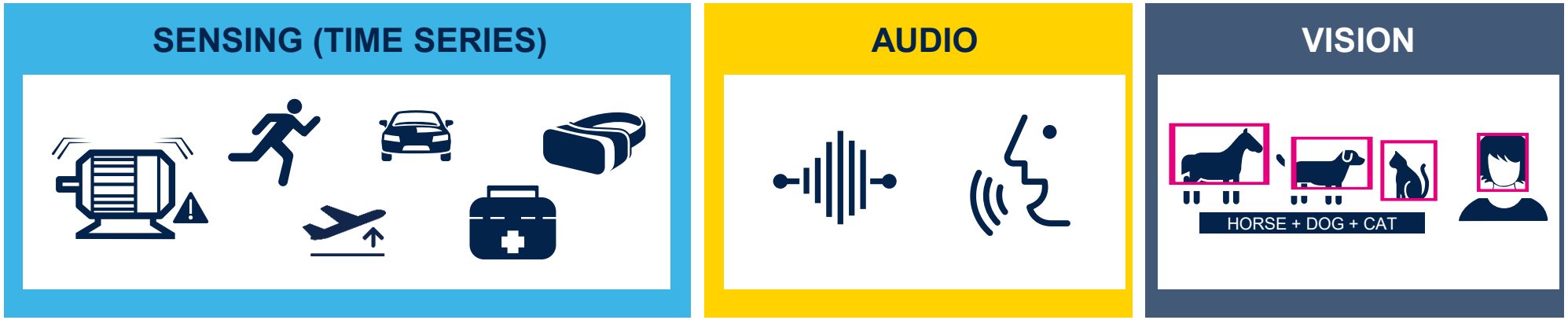


STM32 MCUs

STM32N6 MCU with AI acceleration

STM32 MPU offering for running AI

Enabling major edge AI technologies



Software tools for any user profile

SENSING
AUDIO
VISION

ST Edge AI Developer Cloud

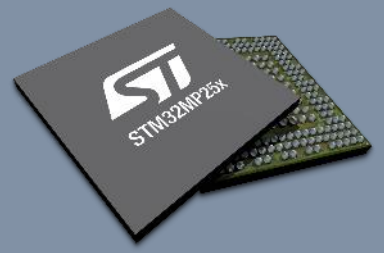
X-LINUX-AI

ST Edge AI Core

A complete AI framework for STM32 MPUs

ST Edge AI Suite

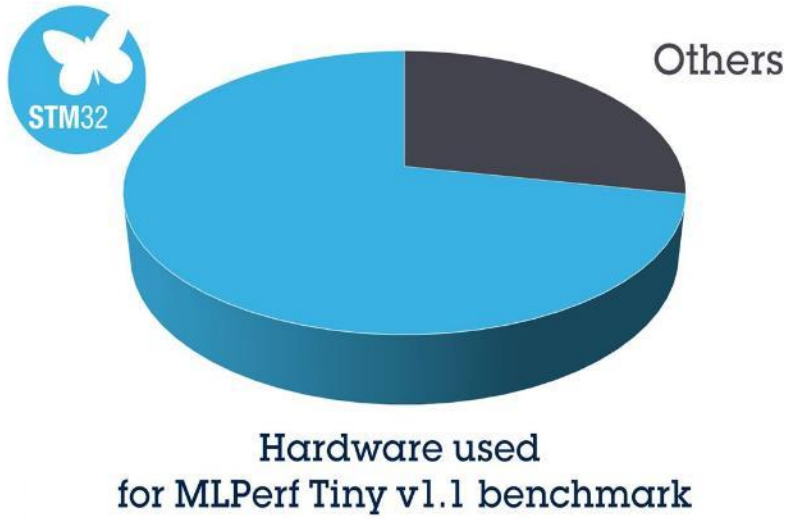
Large choice of general purpose & accelerated hardware



STM32MP1 & STM32MP2 MPUs

STM32 is the de facto platform for embedded AI

ML Commons
ML Perf Tiny benchmark



#1 General purpose MCU provider

Multiyear top contributor in edge AI benchmarks

Only online platform to offer MCU benchmarking

Sensor for running edge AI

Software tools to develop **in-sensor AI** features

MEMS Studio

- Profile and optimize NN and ML models
- Configure the MEMS machine learning core and ISPU

ST AIoT Craft

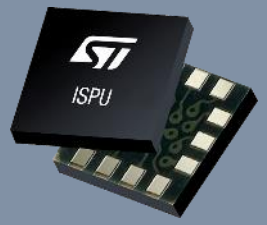
Create end-to-end sensor-to-cloud solutions

ST Edge AI Suite

A broad range of **evaluation** & **prototyping** tools




20+ smart sensors released



6-axis IMUs with ISPU

6-axis IMUs with high-g axl

6-axis IMUs for industrial

6-axis IMUs for automotive

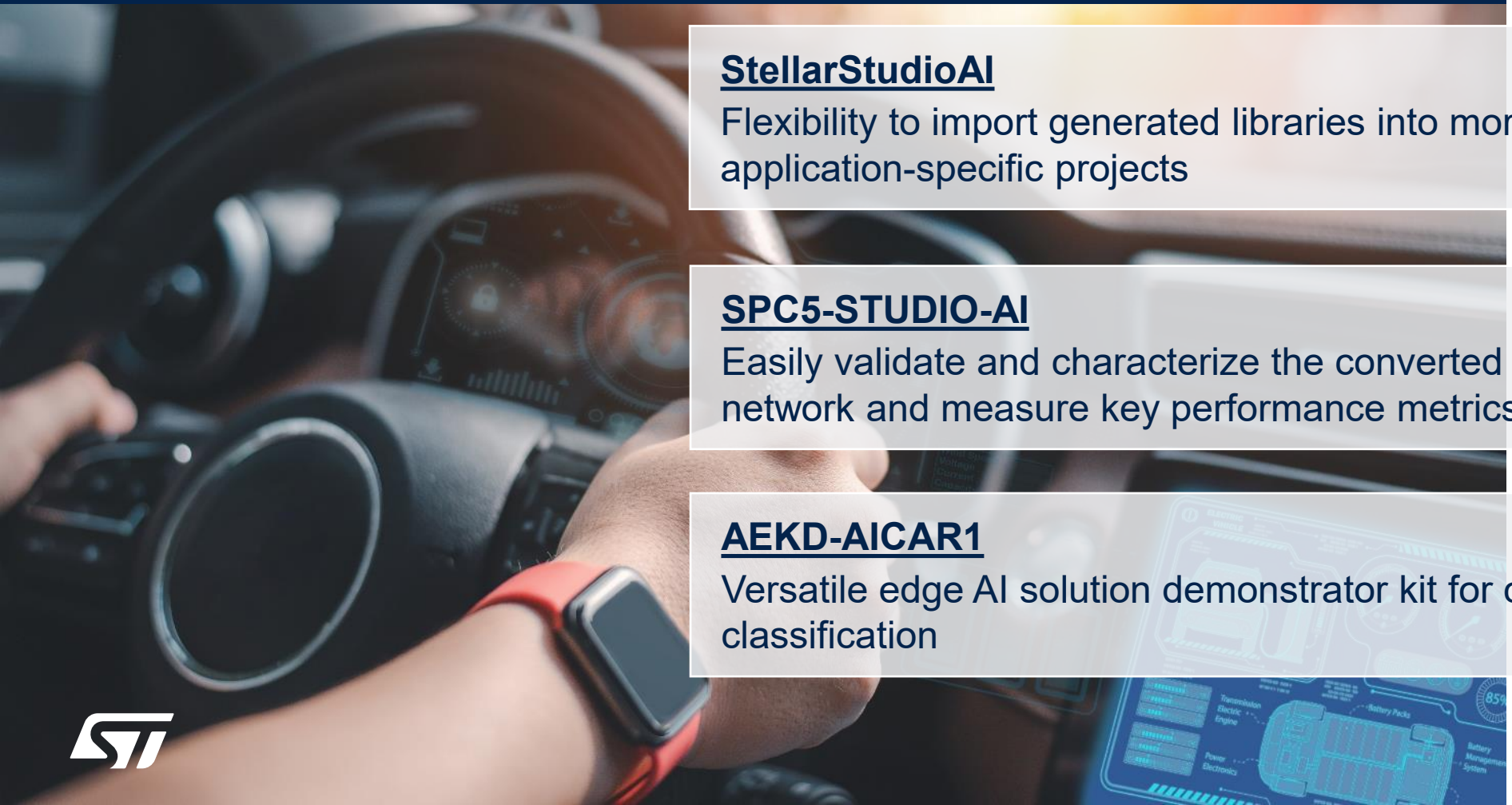
3-axis smart accelerometers

Industrial accelerometers

Industrial inclinometers

Automotive MCUs running AI

Edge AI provides more safety, better efficiency and car maintenance, and a personalized driving experience.



StellarStudioAI

Flexibility to import generated libraries into more complex application-specific projects

SPC5-STUDIO-AI

Easily validate and characterize the converted neural network and measure key performance metrics

AEKD-AICAR1

Versatile edge AI solution demonstrator kit for car state classification

A comprehensive approach to help developers accelerate their product transformation

STM32

STM32 MCUs

STM32N6 MCU

STM32 MPUs

STM32MP2 MPU

Stellar

SPC5

Stellar MCUs
SPC5 MCUs

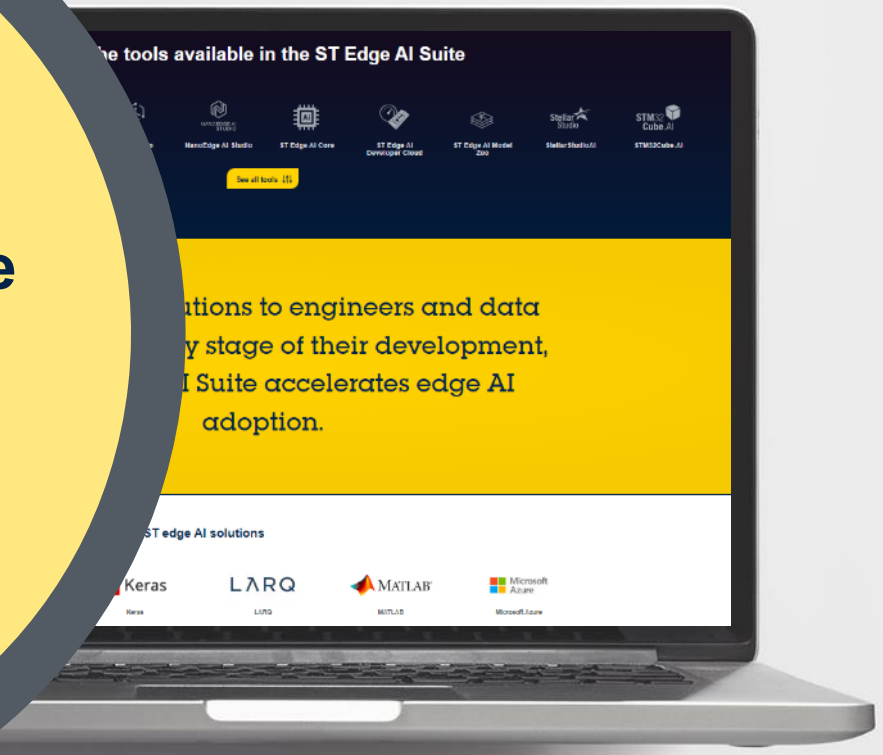
Sensors

MEMS with MLC and ISPU

Hardware platforms

Software tools

ST Edge AI Suite



Click on the products to find out more

Creating an edge AI solution requires multiple skills and tasks

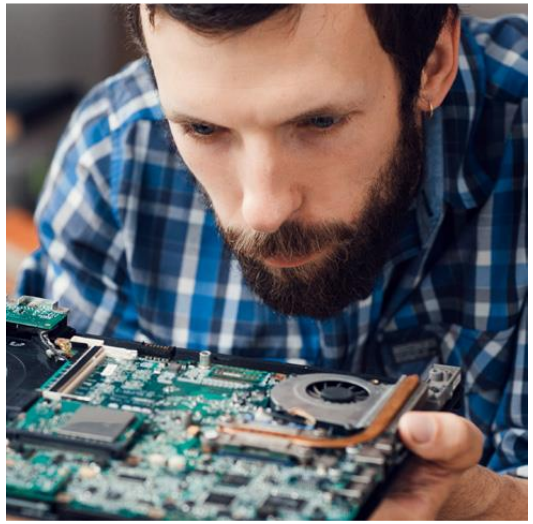
Embedded Software Engineer

"I need to focus on essentials. I prefer to start from existing examples to cut down the time to develop AI solutions."

ML Engineer Data scientist

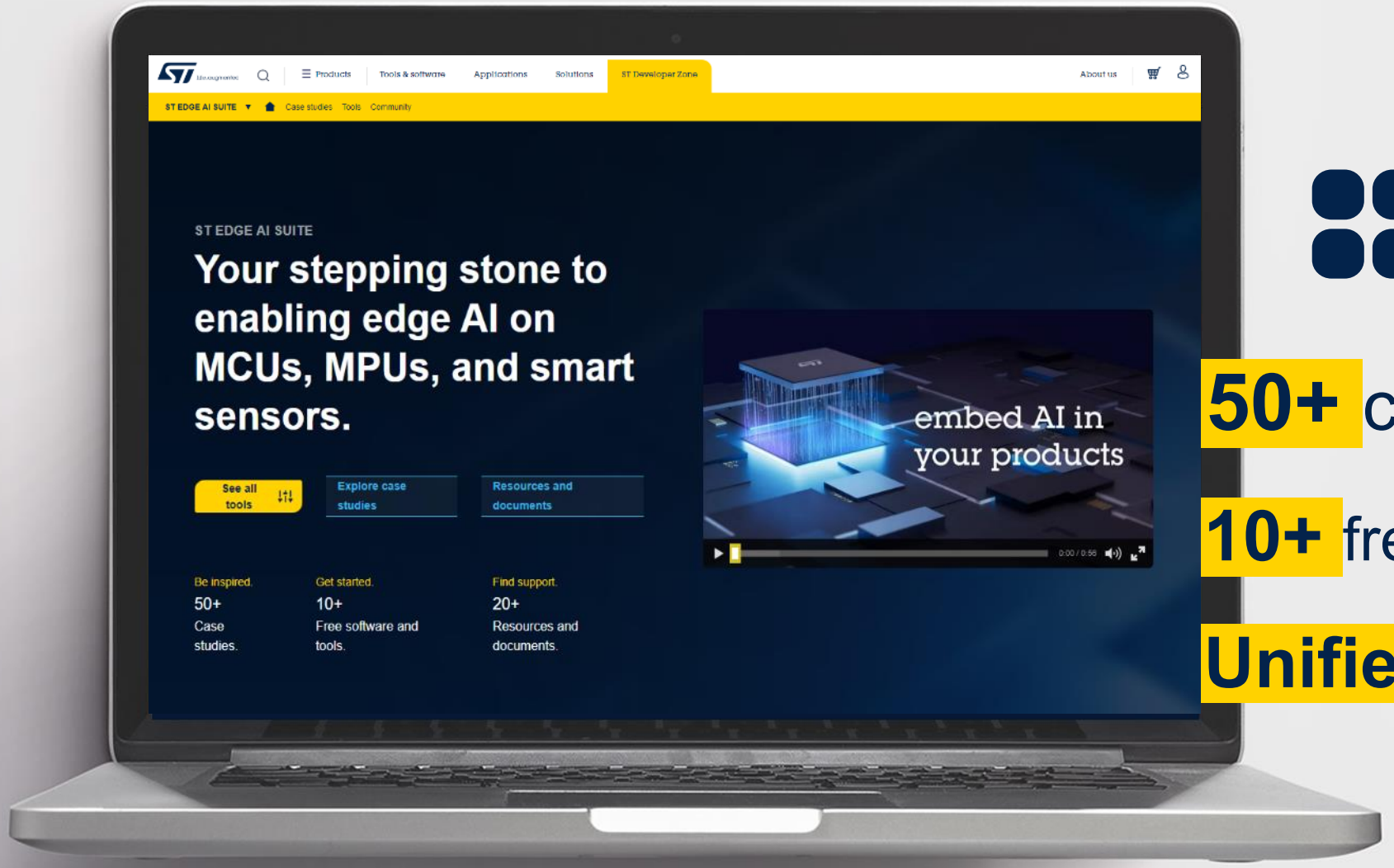
"I mainly work in a scripting environment for dataset management, training and optimizations."

Embedded Hardware Engineer



"I need simple tools to test my AI solutions on different targets. I need to assess and benchmark before real hardware development."

By bringing solutions to engineers and data scientists at every stage of their development, the ST Edge AI Suite accelerates edge AI adoption.



ST Edge AI Suite

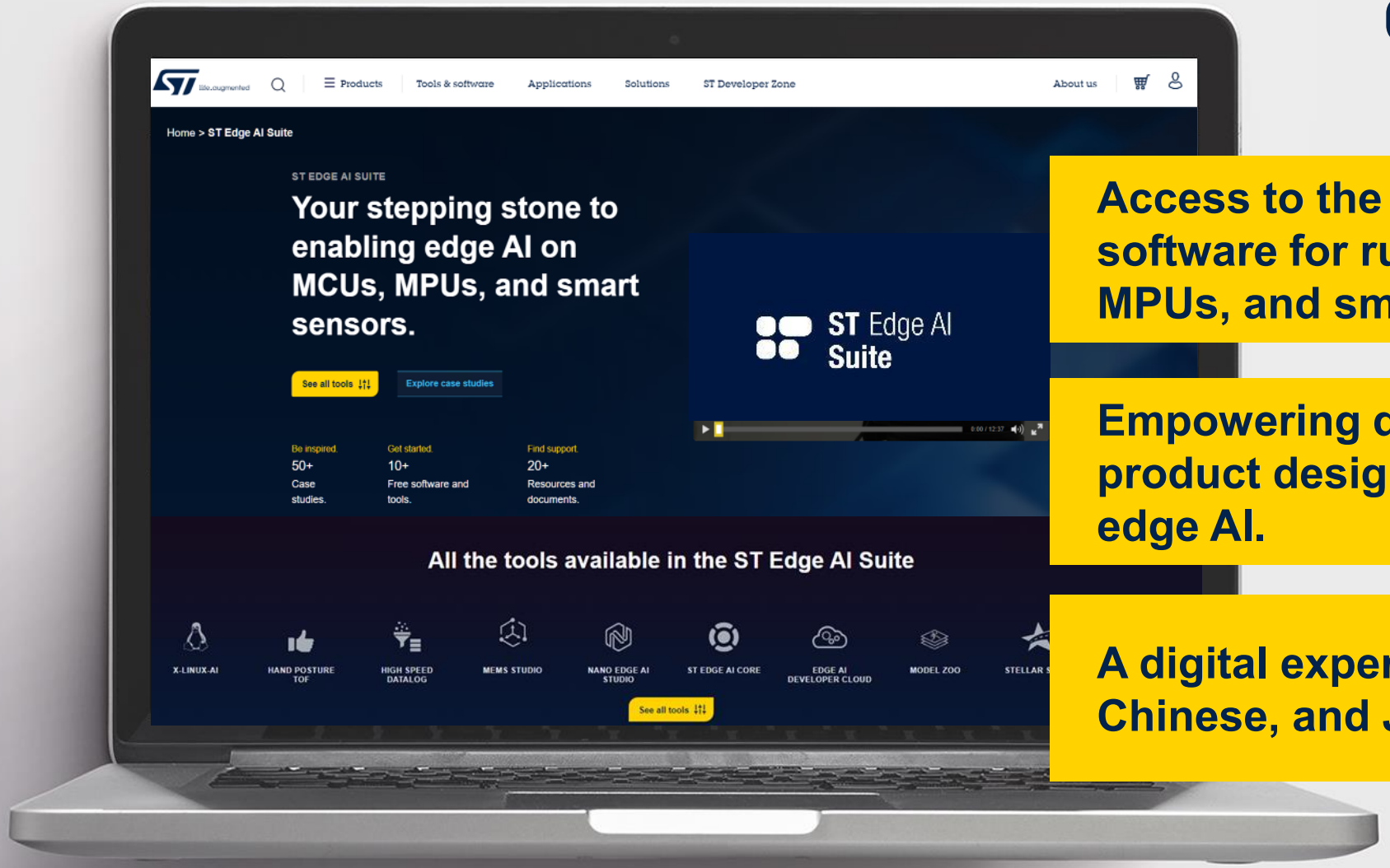
50+ case studies

10+ free software tools

Unified AI core technology

st.com/st-edge-ai-suite

ST Edge AI Suite



Access to the enabling hardware and software for running edge AI on MCUs, MPUs, and smart sensors.

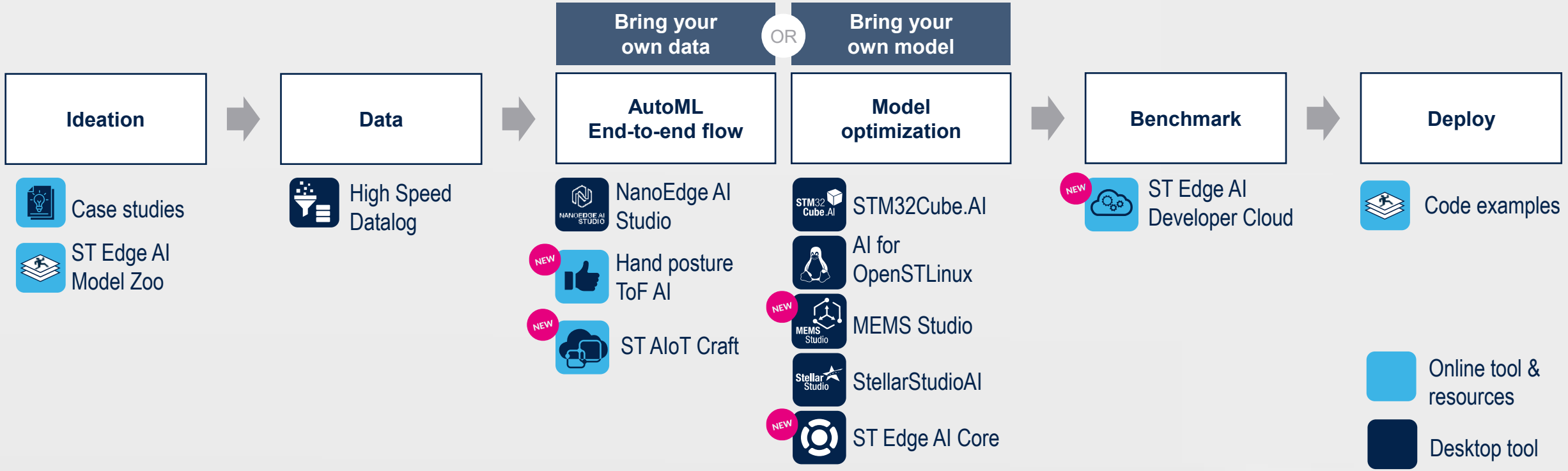
Empowering developers, data scientists, and product designers to harness the power of edge AI.

A digital experience available in English, Chinese, and Japanese.

ST Edge AI Suite

Free tools to run edge AI on MCUs, MPUs, and smart sensors

Find the tools you need to optimize and deploy machine learning algorithms, from data collection to final deployment on hardware.



Our technology starts with You

Find out more at st.com/edge-ai