life.augmented

# Enhance your products using edge AI solutions on STM32

STM32

# A short introduction to AI concepts

life.augmented

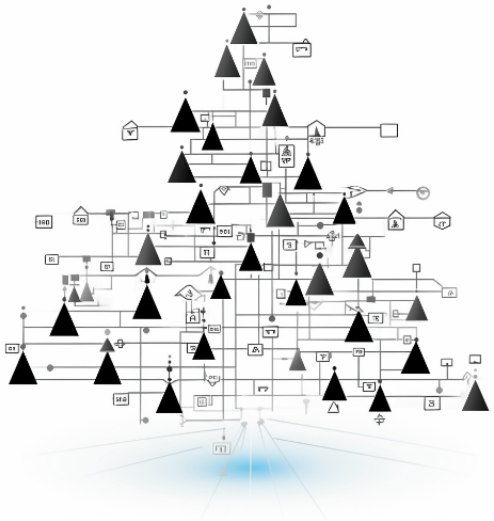
# What is artificial intelligence

**Embedded artificial intelligence**

The ability of **machines** to perform tasks that typically require **human intelligence**

From a technical perspective, AI is another way to program a machine.

Perception
Reasoning
Learning
Decision-making

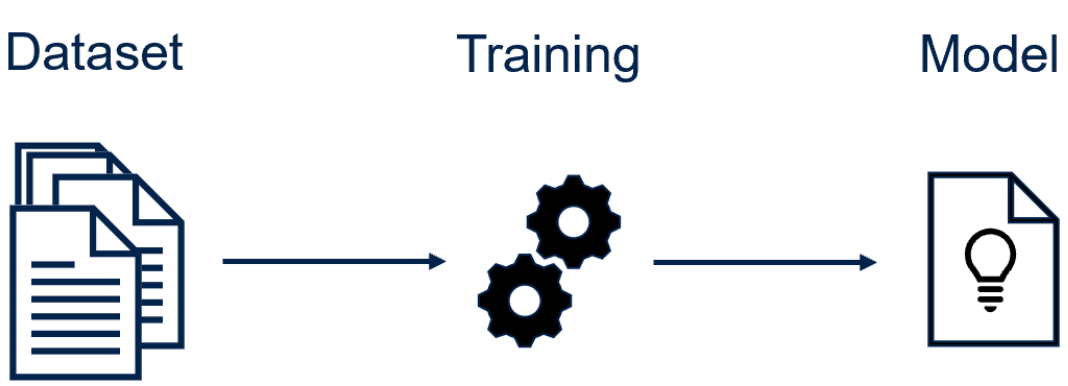


life.augmented

# Using AI algorithms implies two phases

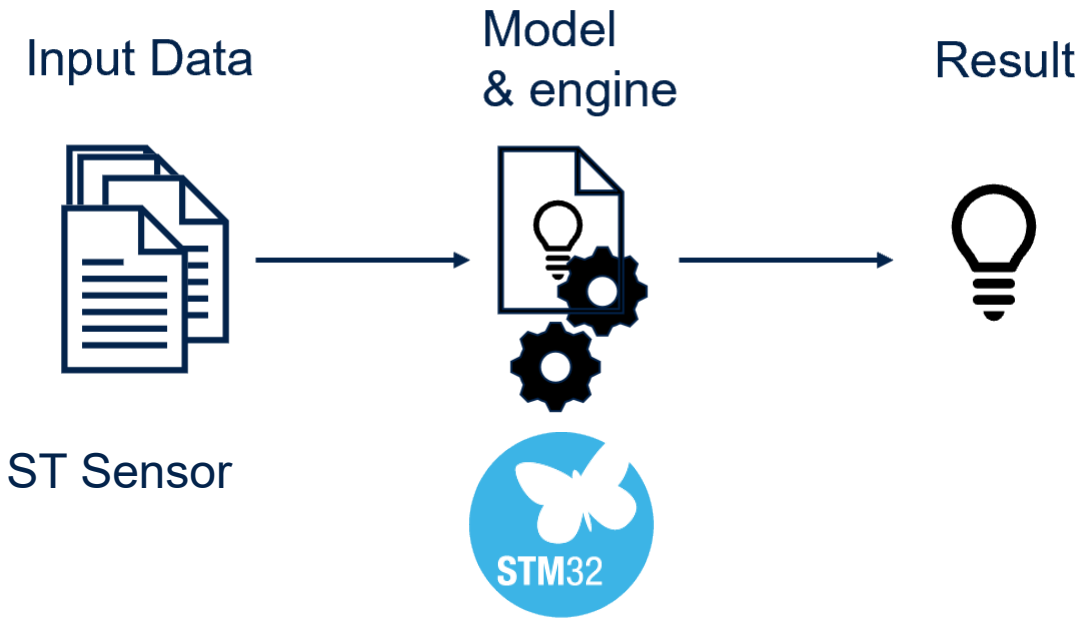
## Training

Using **datasets** (like examples)
and a training engine/framework we are
producing a **model**.



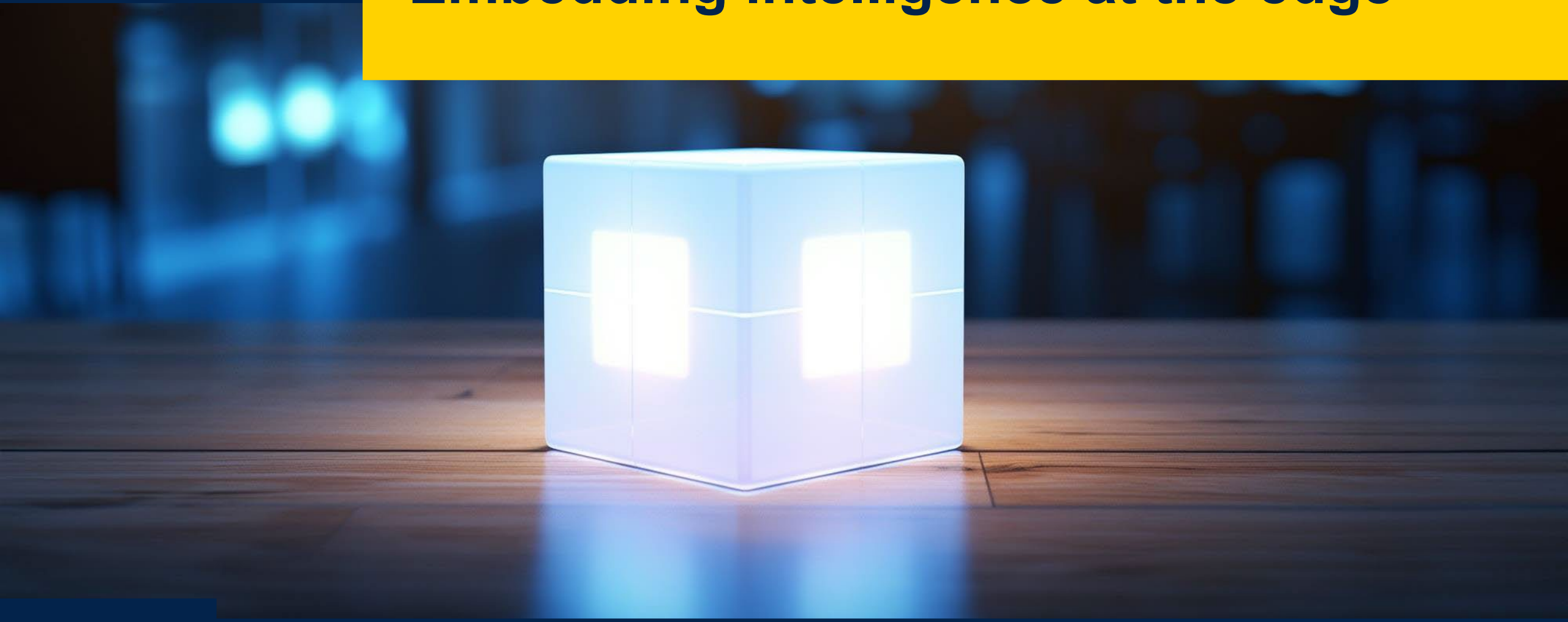

## Inference

Then during the second phase, we can use this **model**.
This is the inference phase.




life.augmented

# Embedding intelligence at the edge



life.augmented

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

Inference

## Cloud-centric

## On device-centric

## Tiny edge-centric

**Enabled by a different class of hardware and software**

life.augmented

# Inferring at the edge brings substantial benefits

**Ultra-low latency**
Real-time applications

**Reduced data transmission**
Generate meaningful information

**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

## Tiny edge-centric

life.augmented

# Opening a new range of embedded AI 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

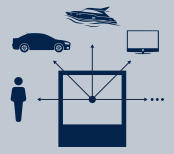
Sound Analysis

Speech Recognition

## Offer advanced services without cloud costs

Object Classification

Pose Estimation



Object Segmentation

A few KOPS → Tens of TOPS

life.augmented

# Edge AI application examples

life.augmented

# AI addresses a wide variety of projects

| Smart city | Smart buildings | Smart home | Energy | Industry 4.0 | Healthcare |
|---|---|---|---|---|---|
| Energy optimization<br>Traffic management<br>Public safety<br>Public transportation | Energy optimization<br>Access control<br>Safety<br>Predictive maintenance | Outdoor cameras<br>Babycams<br>Smart doorbells<br>Home appliances<br>Energy management<br>Lawn mowers<br>Vacuum robots | Solar panels<br>Breaker's control<br>EV chargers<br>Smart meters<br>Fraud detection | Predictive maintenance<br>Tools safety<br>Environment monitoring | Wearables<br>Patient monitoring<br>Fall detection<br>Predictive maintenance |

Battery management

Arc fault detection

Face / object recognition

Anomaly detection

life.augmented

A washing machine uses **advanced motor control algorithms** to weigh clothes and optimize water, detergent, and energy used

**~15-40%**
**Energy saving per washing cycle**

Leader in white-goods

Production starting in 2024 for **millions** units

life.augmented

# **Adding tire pressure detection capability to e-bikes** without adding any new hardware

## Virtual sensor

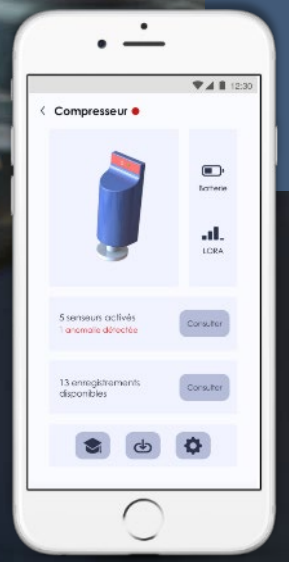
**Tire pressure measured through the e-motor current consumption**

Panasonic e-bike

**Hundreds of thousands** of units annually, starting in 2024

life.augmented | Panasonic

# **After-market predictive maintenance**
intelligent sensor with wireless connectivity



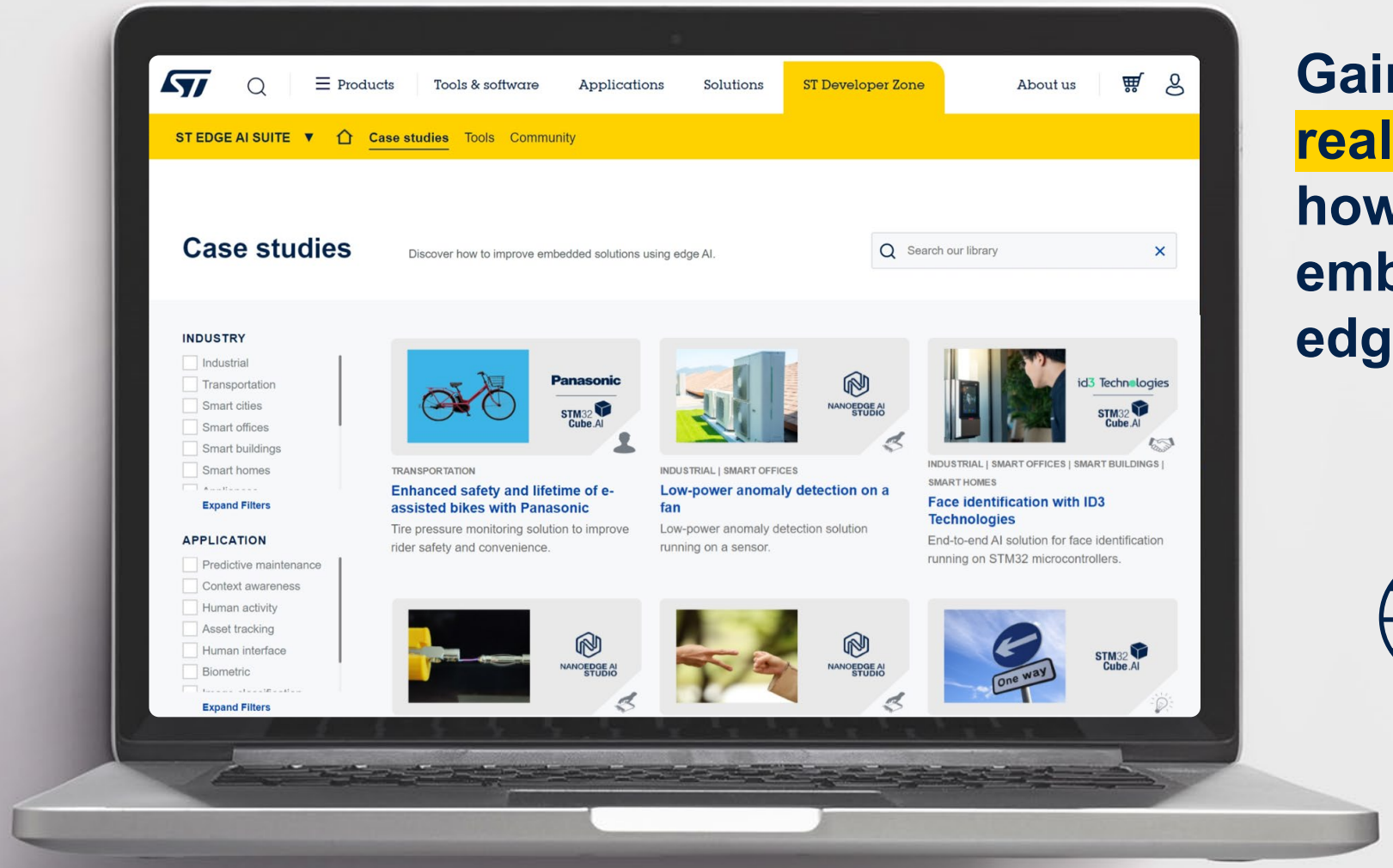

## **Predictive maintenance**
**Multisensors and learning on device**

EMEA company

Deployed at **Volvo Trucks** manufacturing plant

life.augmented | OXYtronic

# Explore inspiring case studies



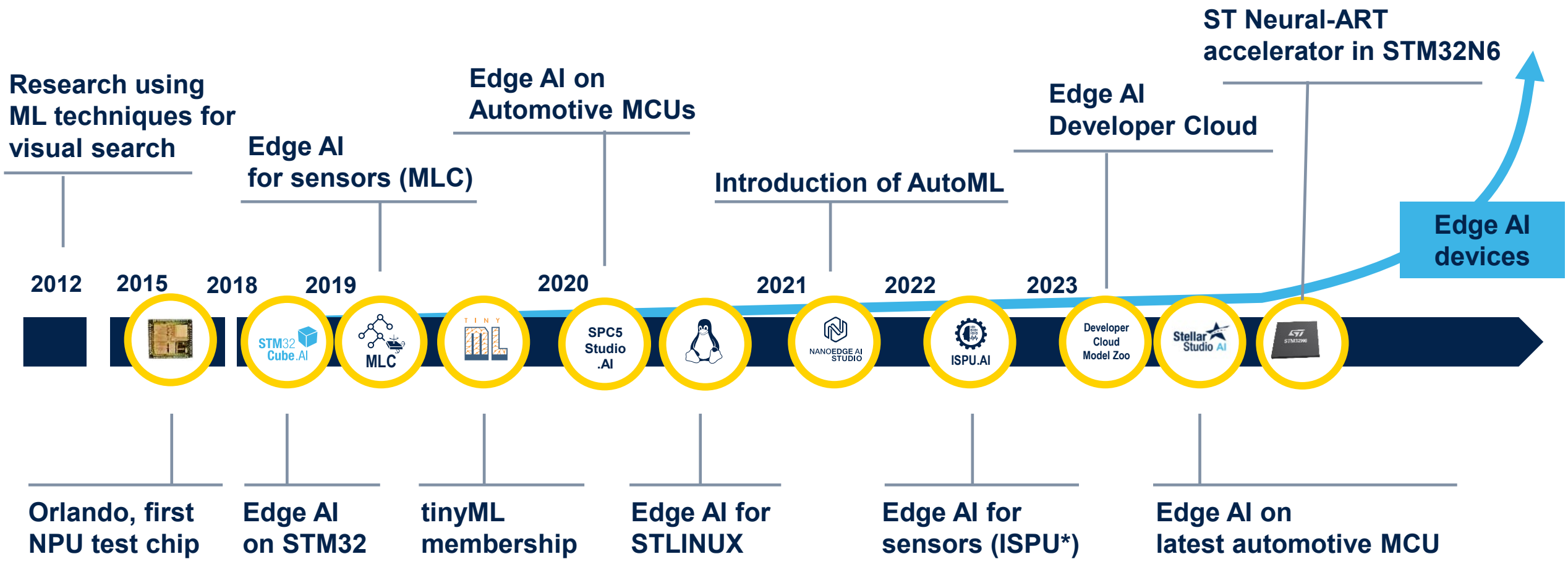

**Gain valuable insights from real-world case studies on how to improve your embedded solutions using edge AI.**

Edge AI case studies

# ST’s investment in edge AI enabling broad adoption

life.augmented

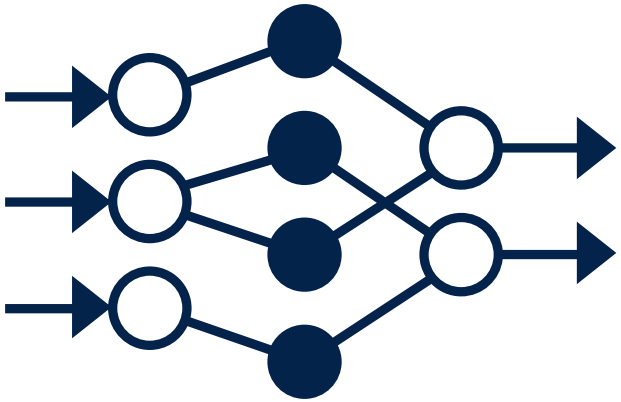
# More than a decade of research, development, and deployment




life.augmented

*Intelligent sensor processing unit

# The challenge of deploying embedded AI



AI expertise

Data

Memory footprint

Inference time

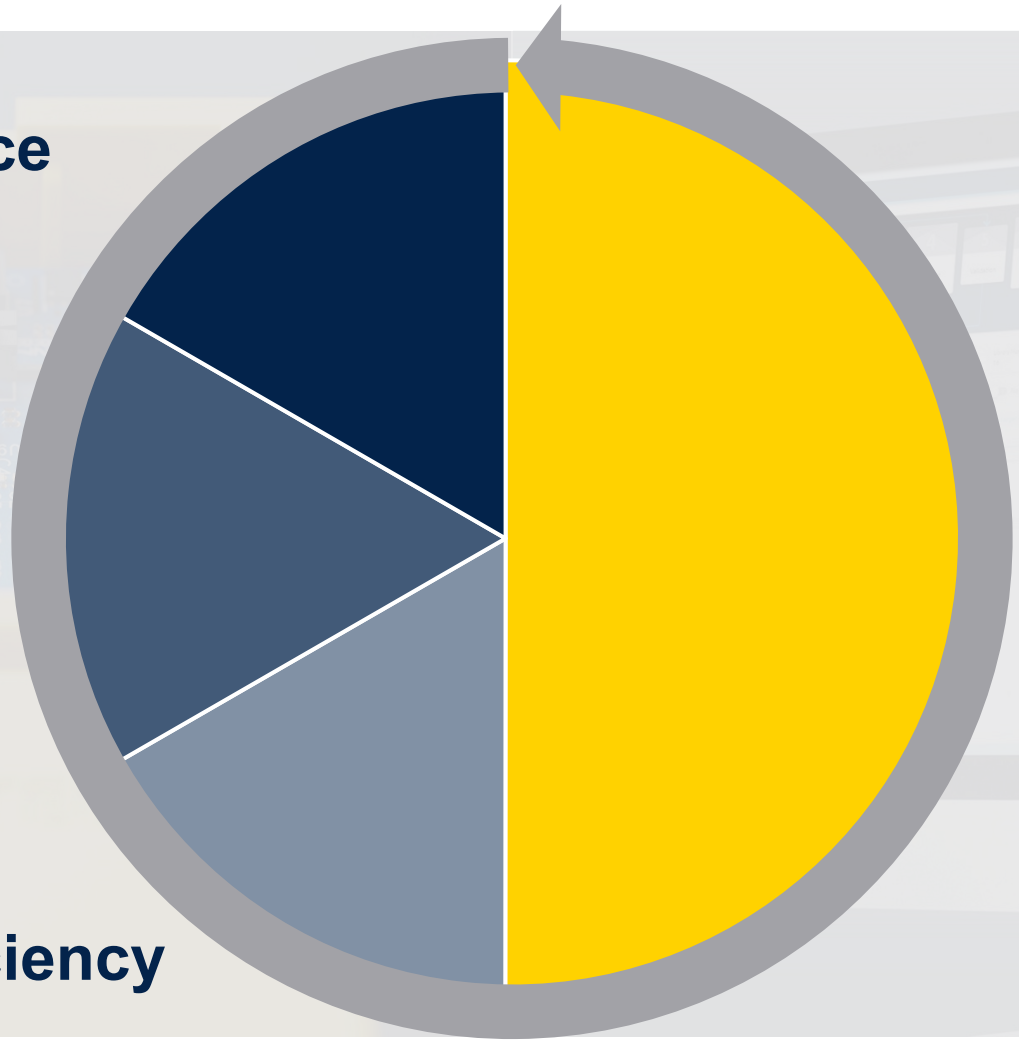
Power consumption

eSW development

**Deploying embedding AI on MCUs is NOT a walk in the park**

life.augmented

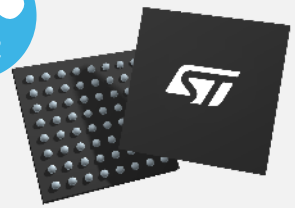
# Taking up the edge AI developer’s challenge




life.augmented

# A comprehensive approach to help developers accelerate their product transformation

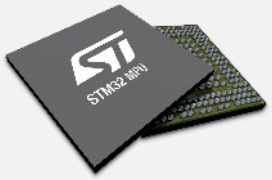
STM32



**STM32** MCUs

NEW

**STM32N6** MCU



**STM32** MPUs

NEW

**STM32MP2** MPU

Stellar

**Stellar** MCUs

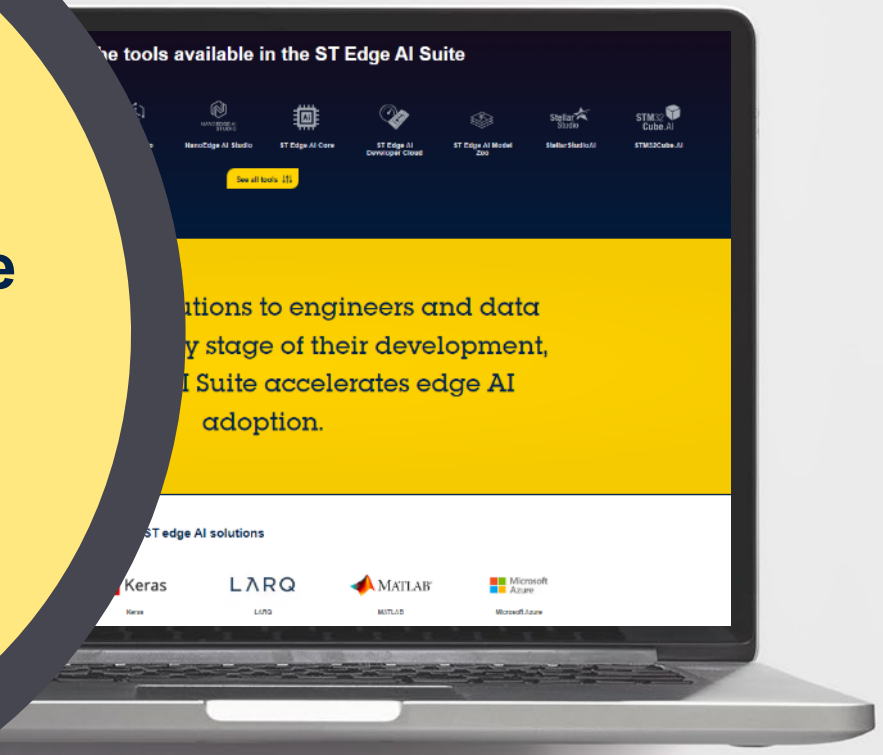
Sensors

**MEMS** MLC/ISPU

**Hardware platforms**

**Software tools**

## ST Edge AI Suite




life.augmented

Click on the products to find out more

# STM32 product offering

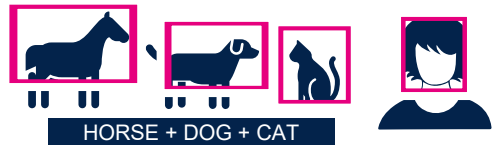
## Enabling major edge AI technologies

### SENSING (TIME SERIES)

ANOMALY DETECTION

### AUDIO

### VISION




## Software tools for any user profile

SENSING
AD

NANOEDGE AI STUDIO

User-friendly AutoML tool for STM32 MCUs

SENSING
AUDIO
VISION

STM32 Cube.AI

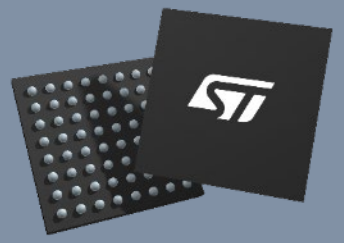
AI model optimizer and code generator for STM32 MCUs

SENSING
AUDIO
VISION

X-LINUX-AI

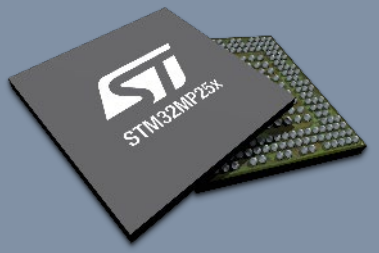
A complete AI framework for STM32 MPUs

## Large choice of general purpose & accelerated hardware



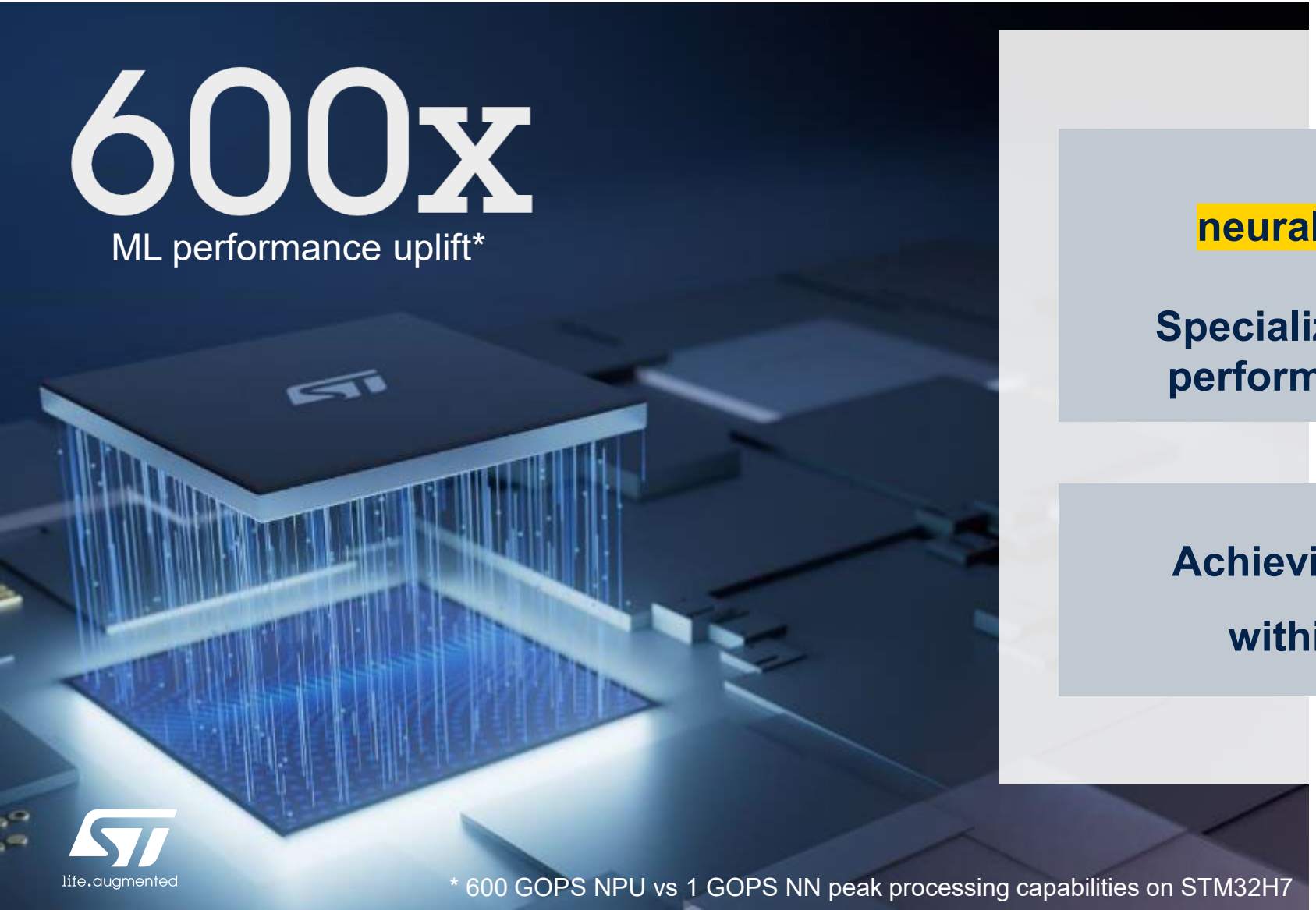
STM32 MCUs

STM32N6 MCU with AI acceleration




STM32MP1 & STM32MP2 MPUs

# ST Neural-ART Accelerator in STM32N6




**ST proprietary**
**neural processing unit (NPU)**

**Specialized hardware accelerator performs advanced AI inference**

**Achieving edge AI performance within MCU power budget**

Read more in the blog

Read our whitepaper

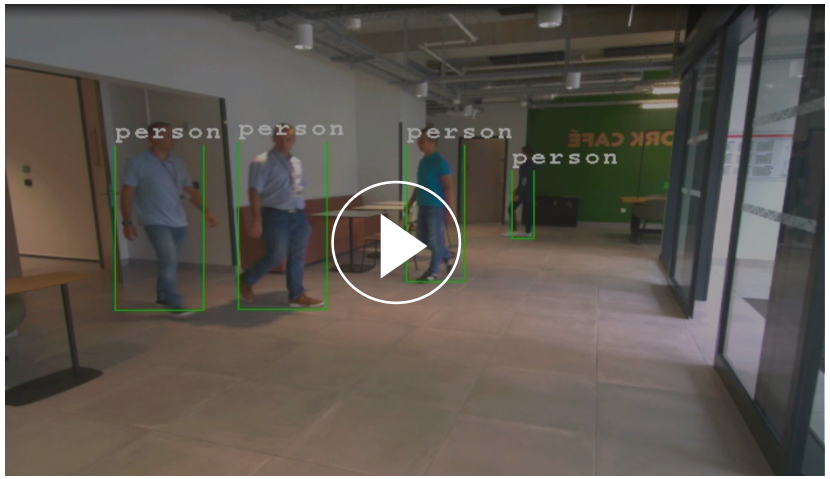
life.augmented

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

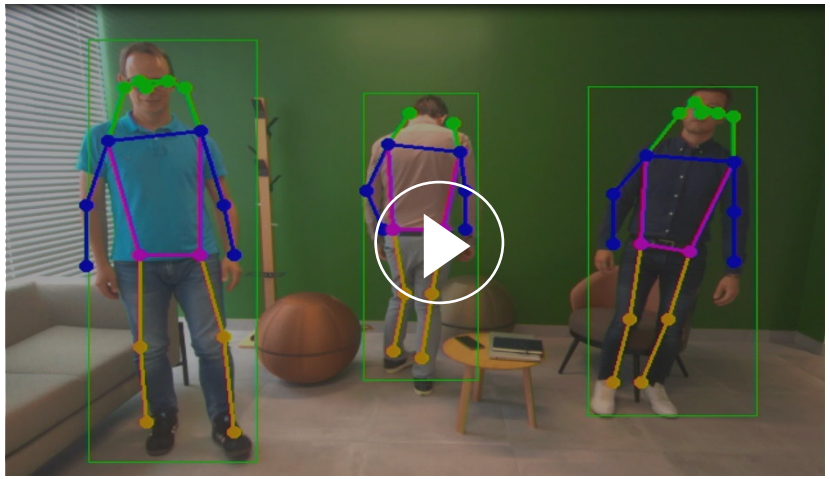
# How the STM32N6 changes the game

**An MPU-like end-user experience. Available on an MCU.**

*People detection*



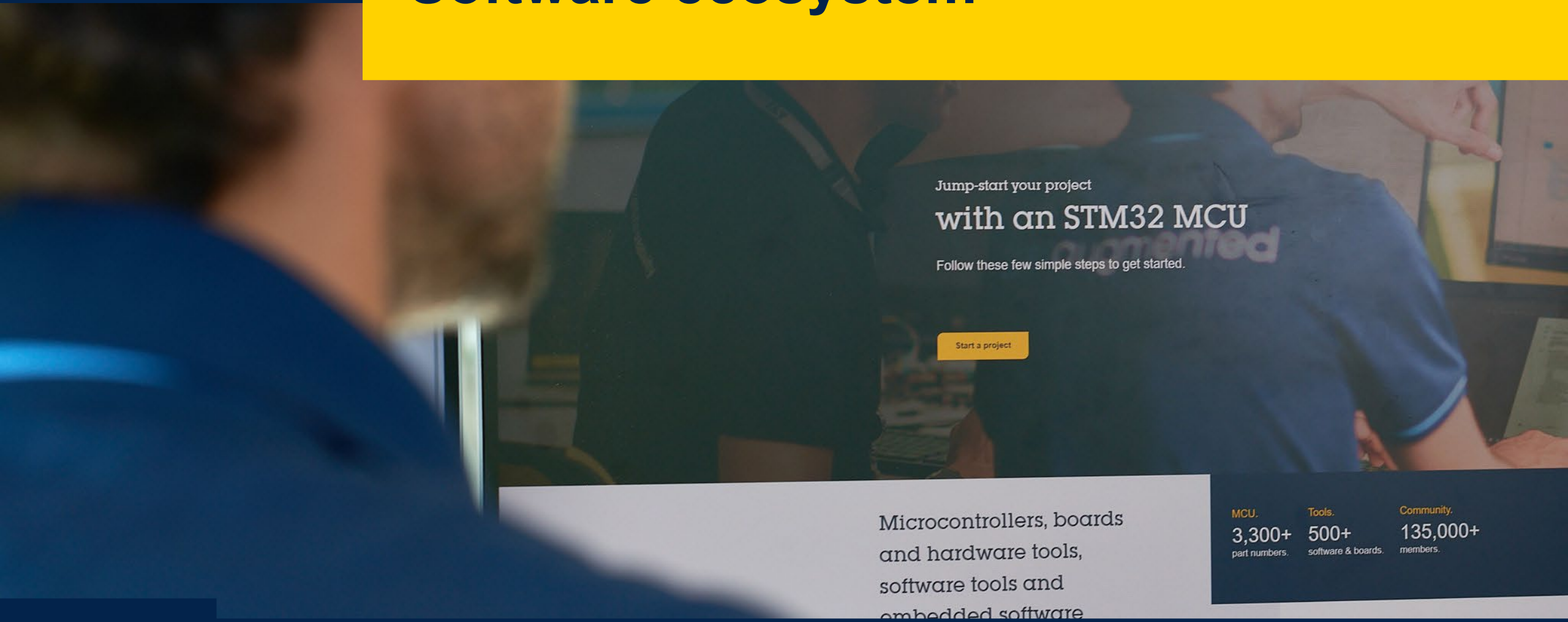

*Pose estimation*



*Hand landmark*

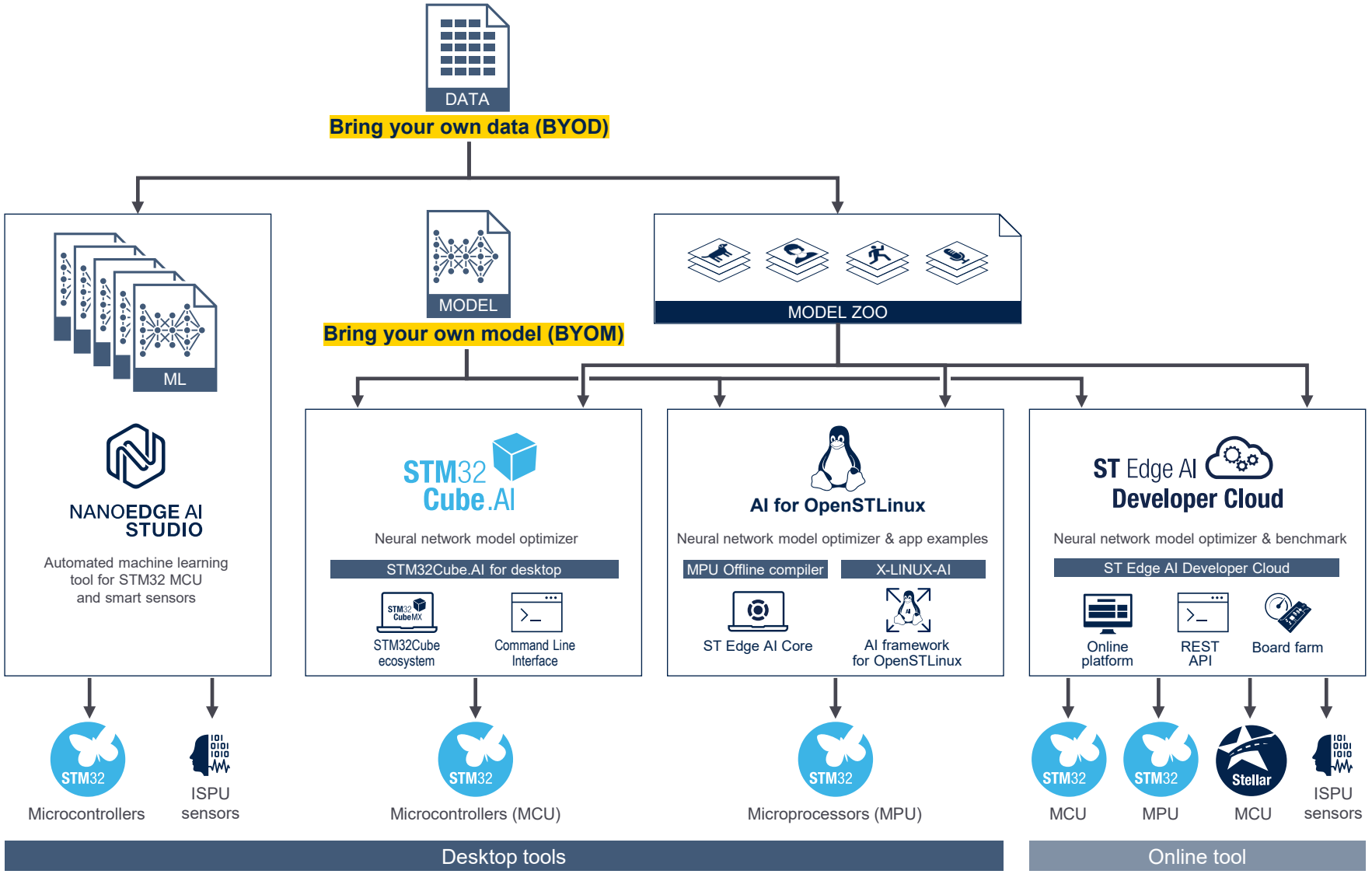
Watch all videos on YouTube

# Software ecosystem




life.augmented

# A broad collection of free tools

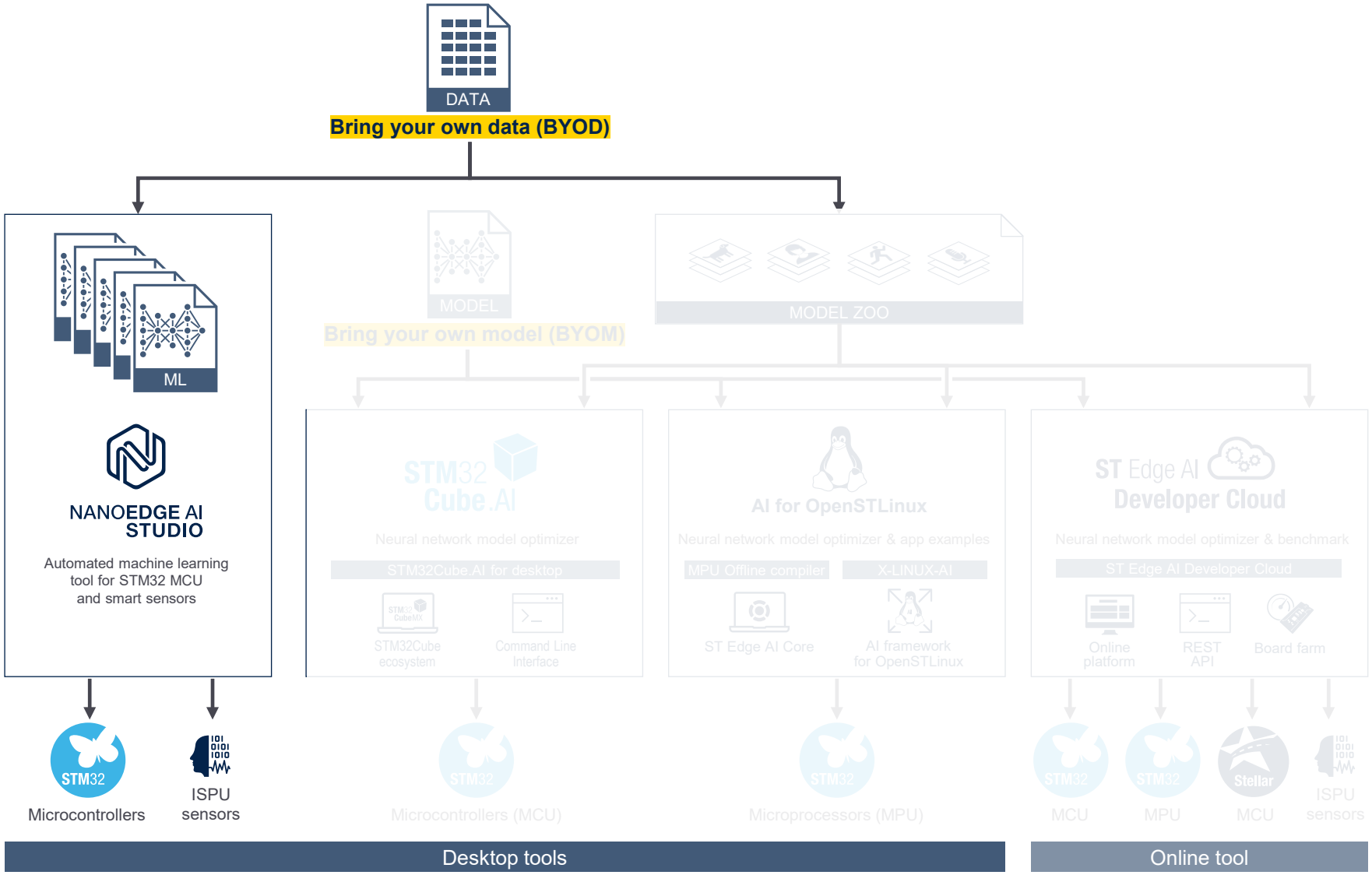



life.augmented

# NanoEdge AI Studio, the AutoML tool

life.augmented

# Software tool: NanoEdge AI Studio




life.augmented

# Simplified edge AI development workflow

**Deployment of NanoEdge AI Studio libraries, the market reference AutoML tool, is completely free for unlimited quantities on any STM32**

**… and available on any Arm® Cortex® -M MCU***

NANOEDGE AI STUDIO

The **best combination** for given data:
ML model, hyperparameters, and preprocessing

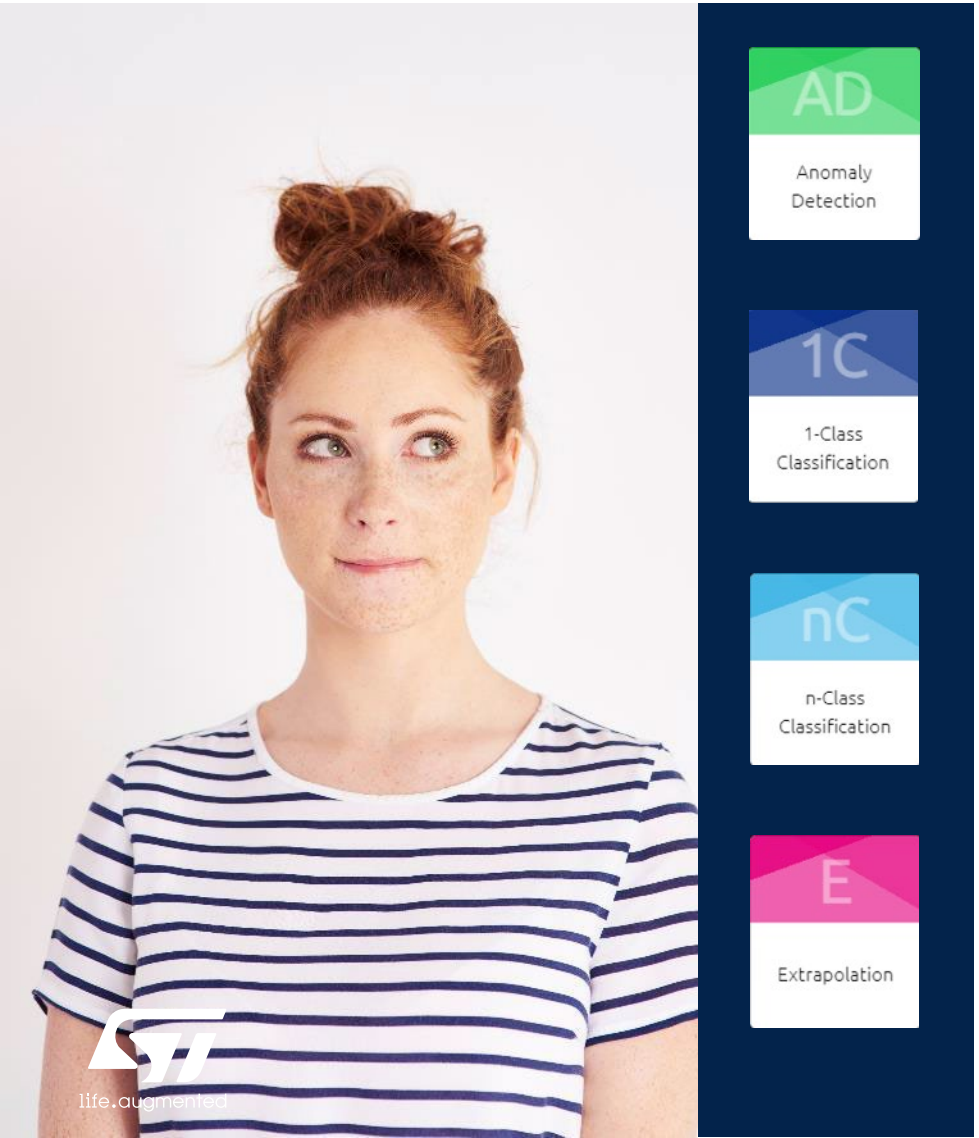
**On-device learning capability** to fine-tune a deployed solution without retraining

**Bring your own data** approach:
no need to create edge AI models

**under a special license agreement*

life.augmented

# State-of-the-art machine learning for smarter products



AD
Anomaly Detection

1C
1-Class Classification

nC
n-Class Classification

E
Extrapolation

> **I want to anticipate product failures**

> **I need to detect any outliers**

> **I want to identify the activity, the environment, the usage**

> **I need to predict future states**

life.augmented

# NanoEdge AI Studio workflow

**Create your dataset And set up your project**

**Project parameters**
max ram, max flash, microcontroller ref.

011001
100111
110100
111000

**Dataset**
Collection of data from sensors

**Analyze, choose, and validate the best ML processing**

NANOEDGE AI STUDIO

Automated Machine Learning tool
**for STM32 MCU and smart sensors**

STM32

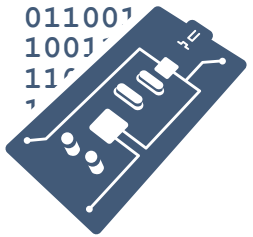
ISPU Sensors

**Generate an optimized library**

.a Library

**Precompiled library (.a) to link to your main code**

**Train your model directly on the device**



**On-device learning**
(for anomaly detection)

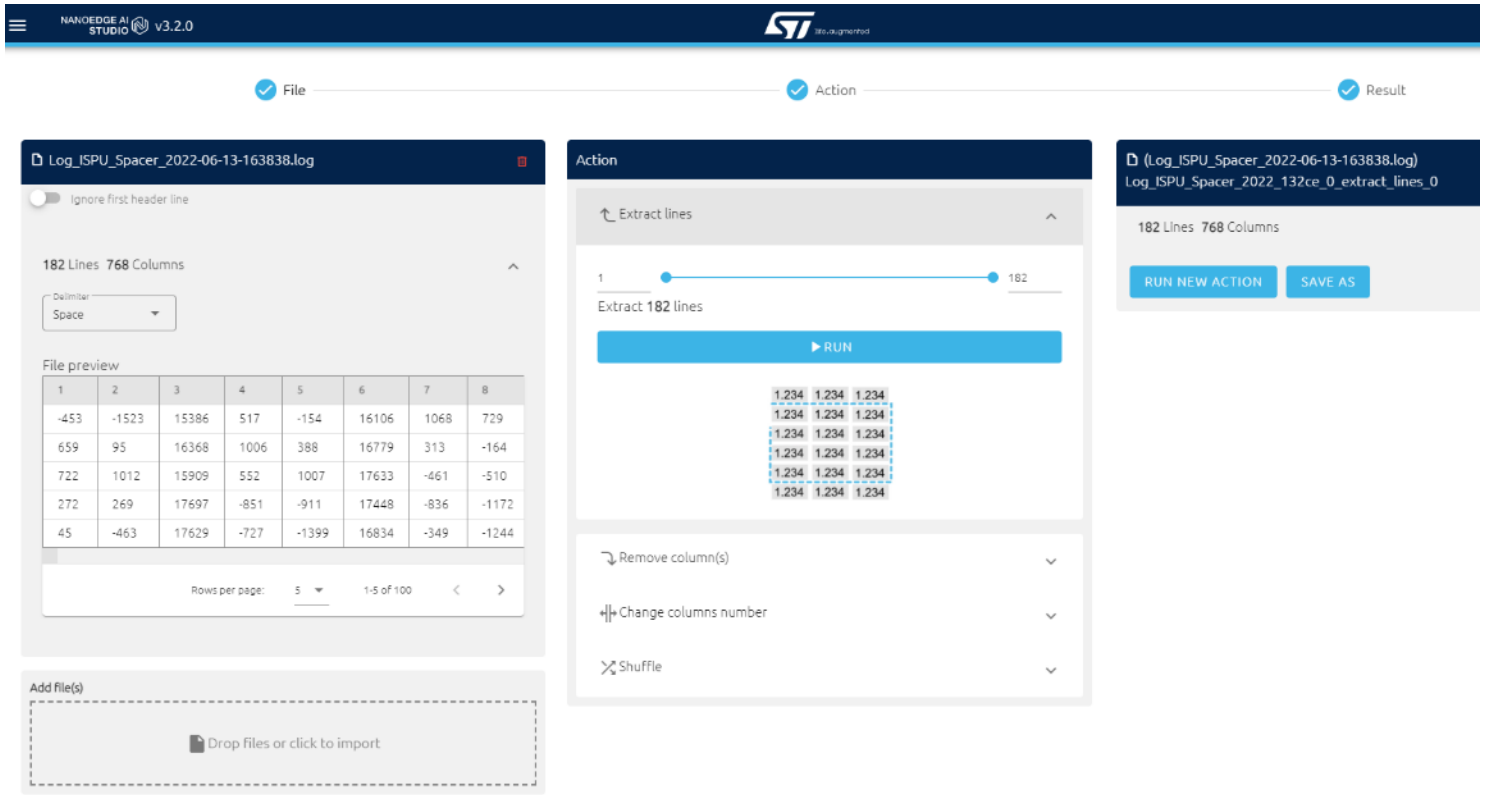
life.augmented

# All the tools needed for data preparation

DM

Data Manipulation

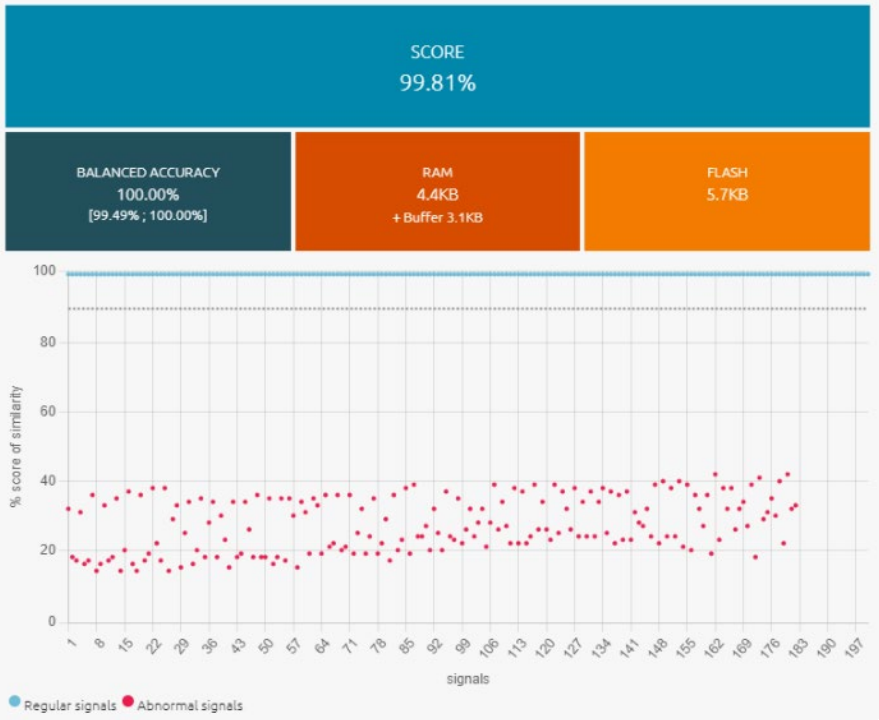
**“I want to clean my data quickly and efficiently”**




life.augmented

# Get the best ML algorithm for your application

## Internal benchmarking tool picks the best algorithm for your data

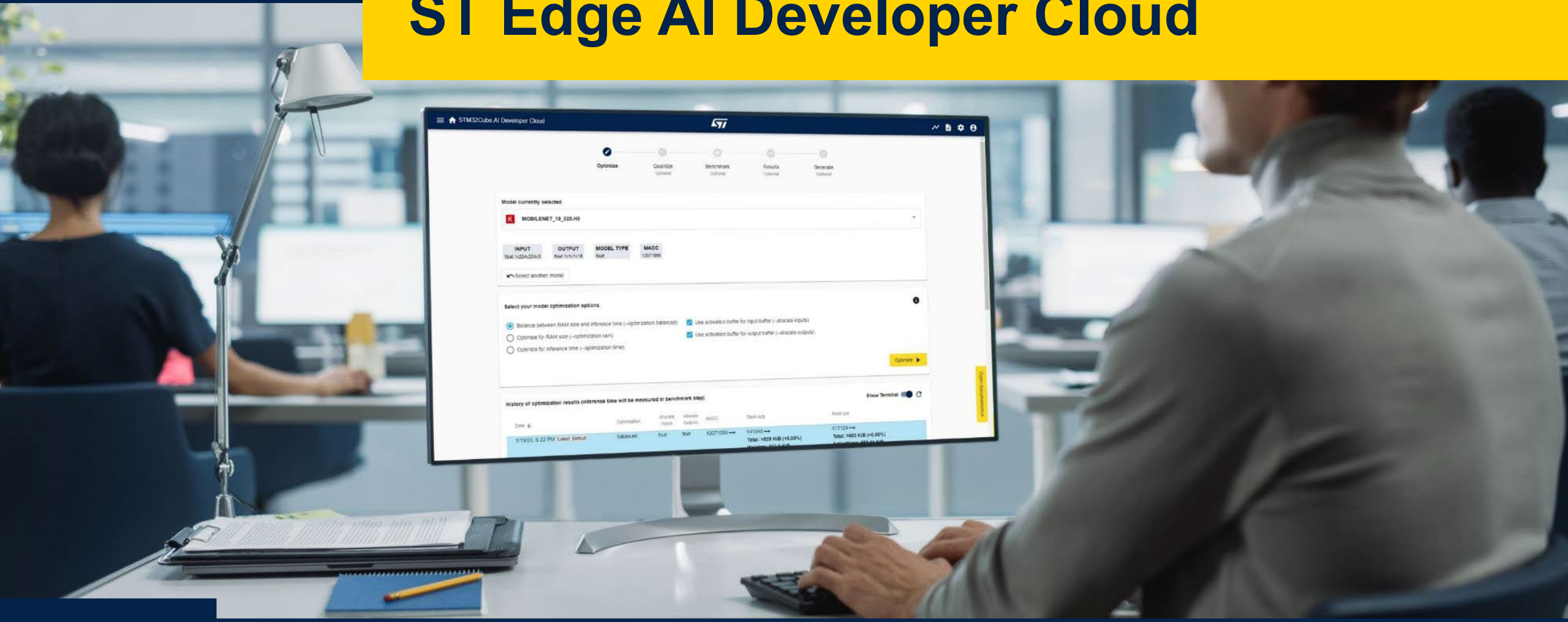


NanoEdge AI Studio improves the performance of the model.

And optimizes it to reduce footprint and latency

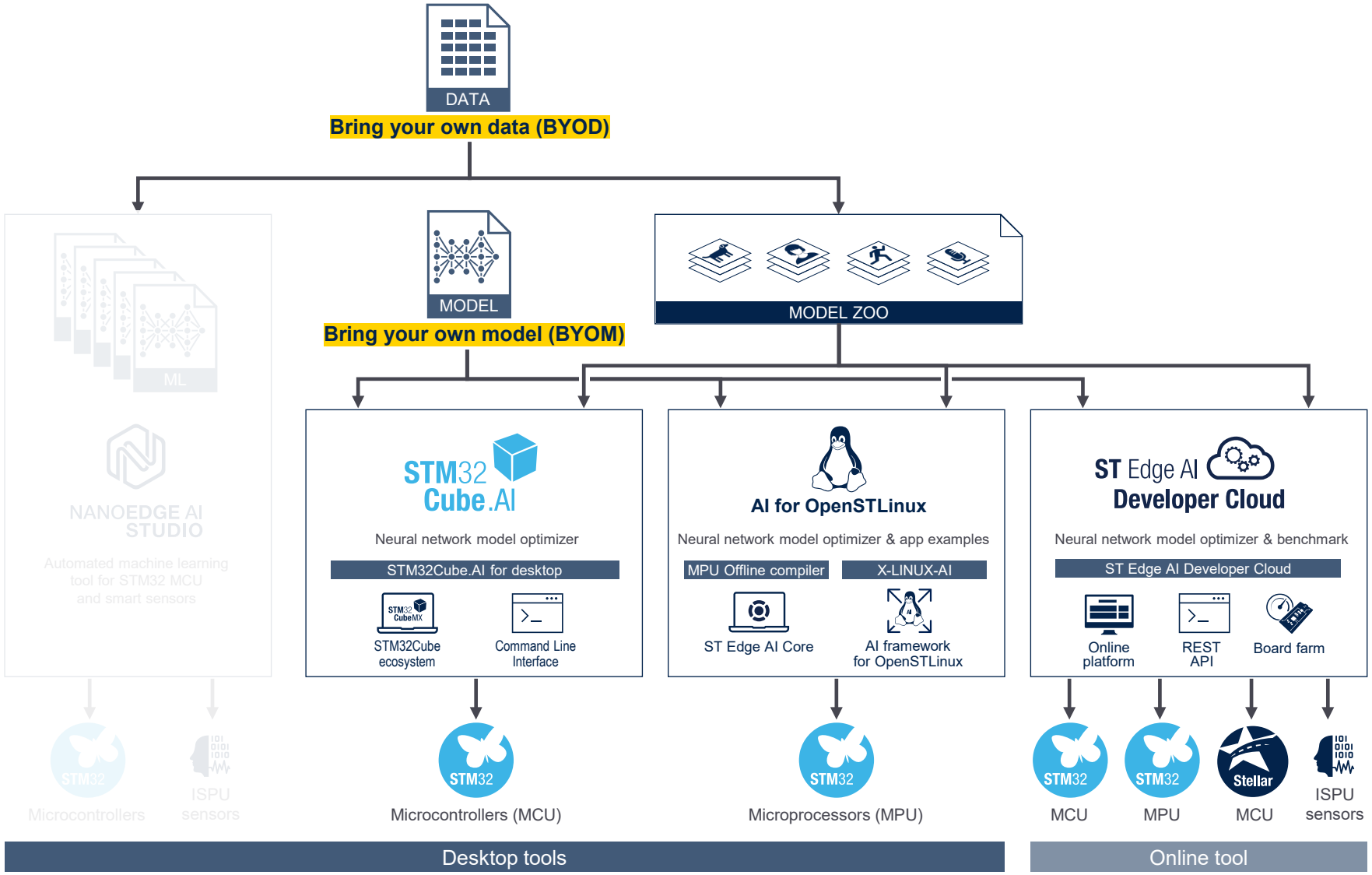
life.augmented

# STM32Cube.AI &
# ST Edge AI Developer Cloud



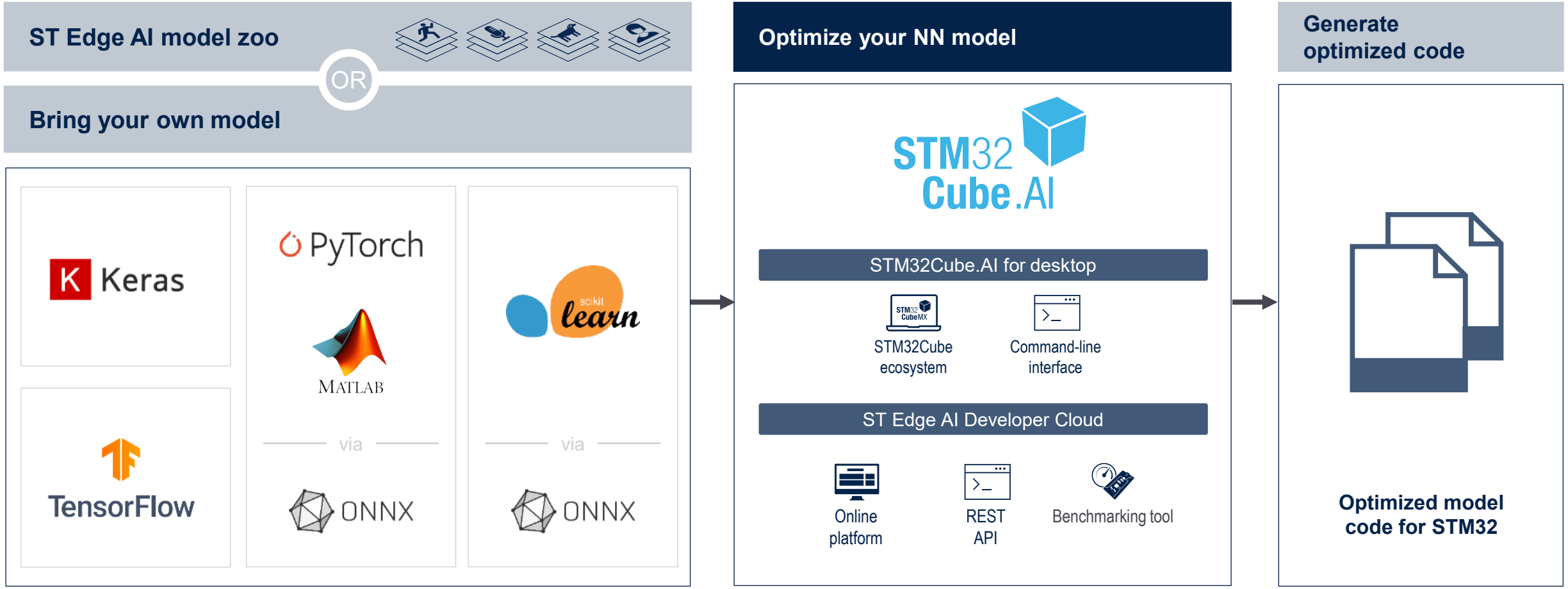

life.augmented

# Software tool: STM32Cube.AI




life.augmented

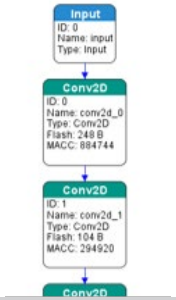
# One tool – two versions to deploy AI on STM32



# The 3 pillars of STM32Cube.AI

## Graph optimizer

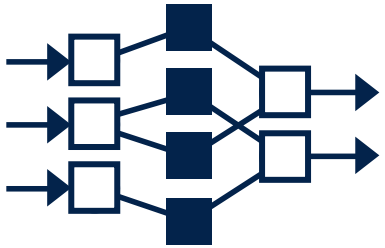
Automatically improve performance through graph simplifications & optimizations that benefit STM32 target hardware architectures




- Auto graphs rewrite
- Node/operator fusion
- Layout optimization
- Constant-folding…
- Operator-level info to fine-tune memory footprint and computation

## Quantized model support

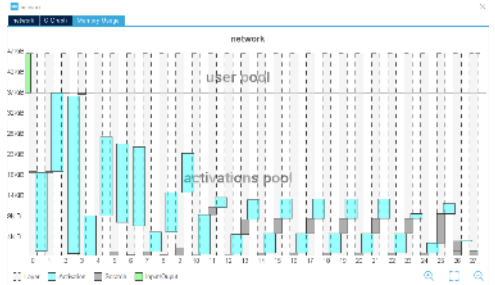
Import your quantized ANN to be compatible with STM32 embedded architectures while keeping their performance



- From FP32 to Int8 or mixed-precision
- Minimum loss of accuracy
- Code validation on target
  - Latency
  - Accuracy
  - Memory footprint

## Memory optimizer

Optimize memory allocation to get the best performance while respecting the constraints of your embedded design




- Memory allocation
- Internal/external memory repartition
- Model-only update option

STM32Cube.AI is **free of charge**, available both in graphical interface and in command line.

life.augmented

# Start with edge AI optimized models STM32 model zoo

**A collection of application-oriented models optimized for STM32**

Human activity

Motion Sensing

Image classification

Computer vision

Audio event detection

Audio classification

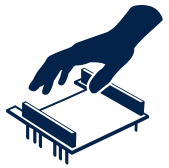
Object detection

Computer vision

**Hosted on GitHub**

**< | >** **Model training scripts**

- Scripts to generate and validate



**Application code example**

- Designed to host optimized NN models
- Automatically generated from the trained models
- Easy to deploy for end-to-end evaluation

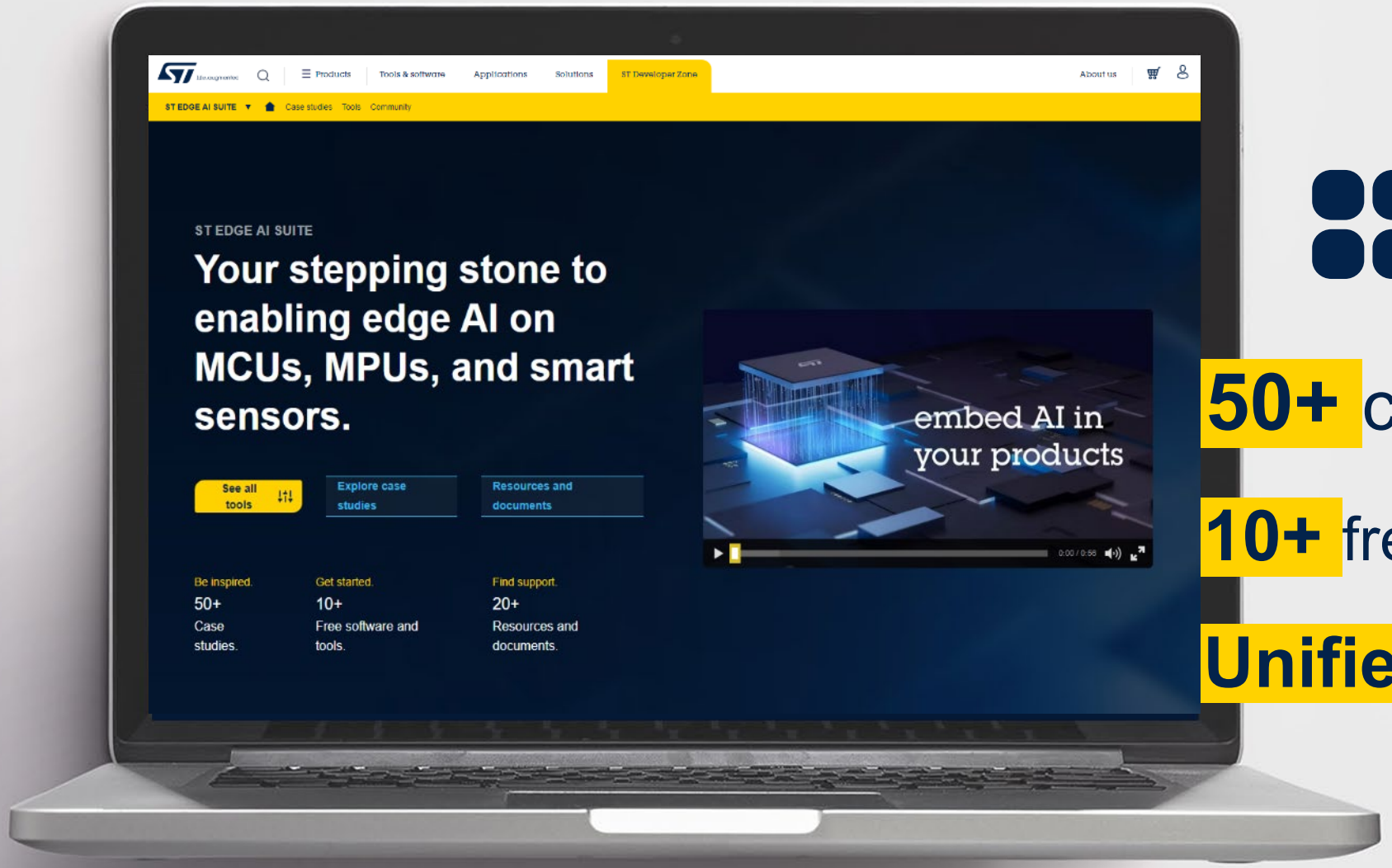
life.augmented

# Discover the ST Edge AI Developer Cloud

# ST Edge AI Suite

life.augmented

# 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**

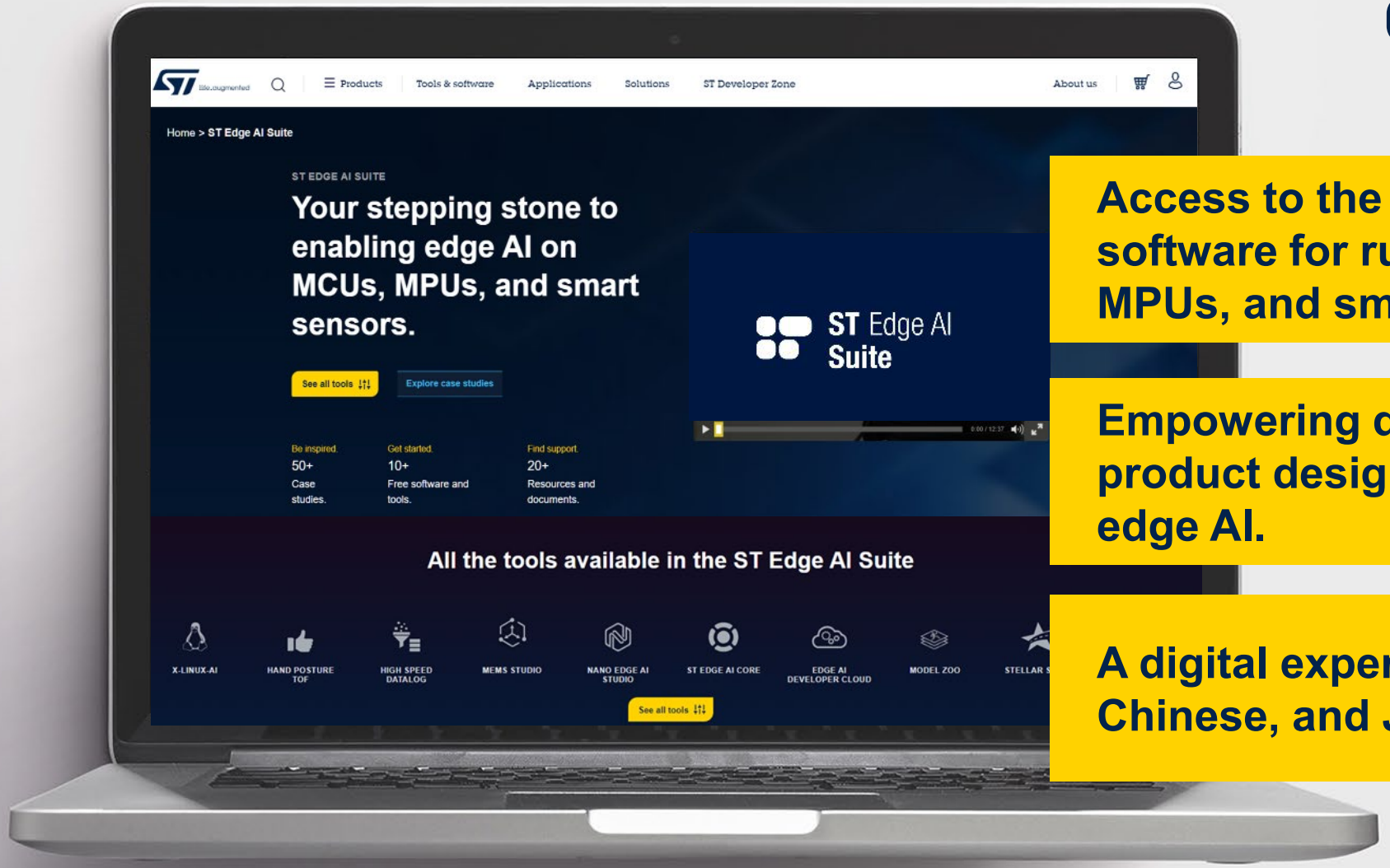
**50+** case studies

**10+** free software tools

**Unified** AI core technology

life.augmented

# ST Edge AI Suite



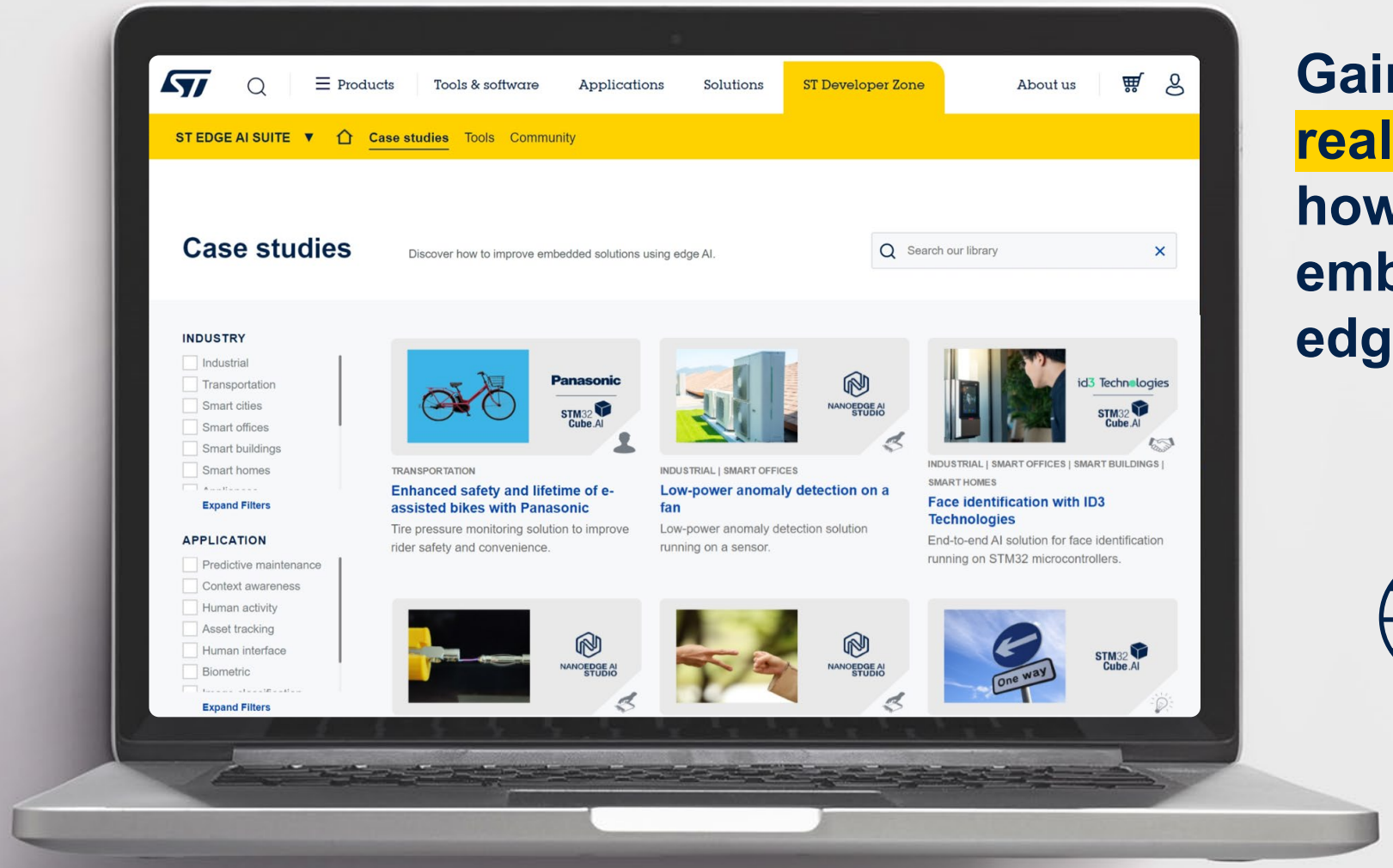

**Access to the enabling hardware and software for running edge AI on MCUs, MPUs, and smart sensors.**

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

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

life.augmented

# Explore inspiring case studies



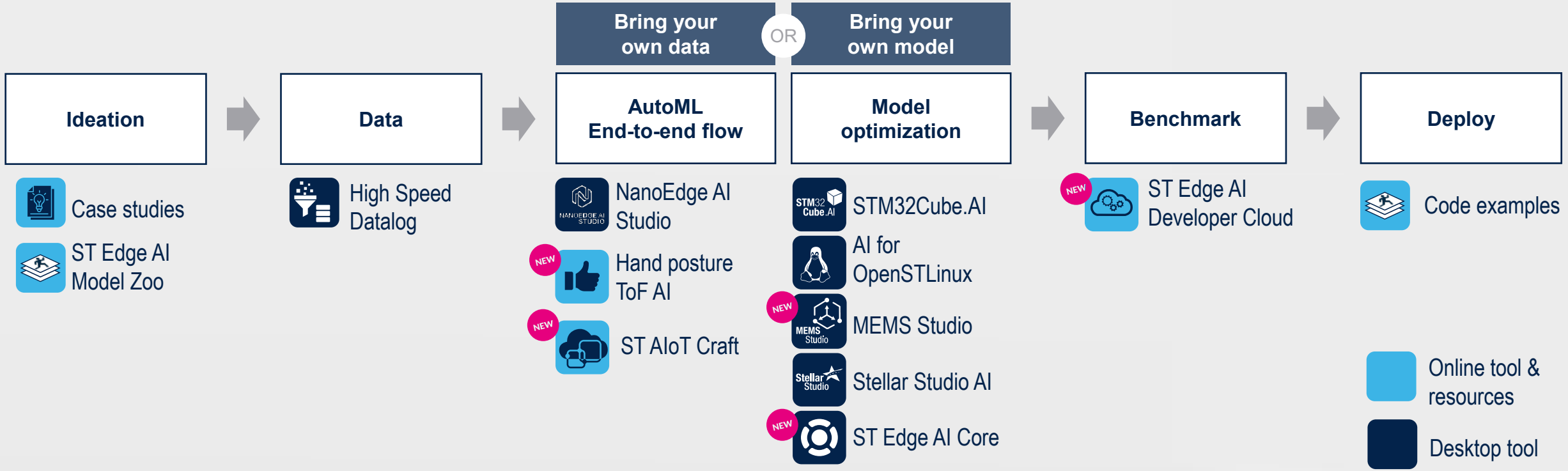

**Gain valuable insights from real-world case studies on how to improve your embedded solutions using edge AI.**

Edge AI case studies

# 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.**




life.augmented

# Our technology starts with You

Find out more at www.st.com/stm32ai

© 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