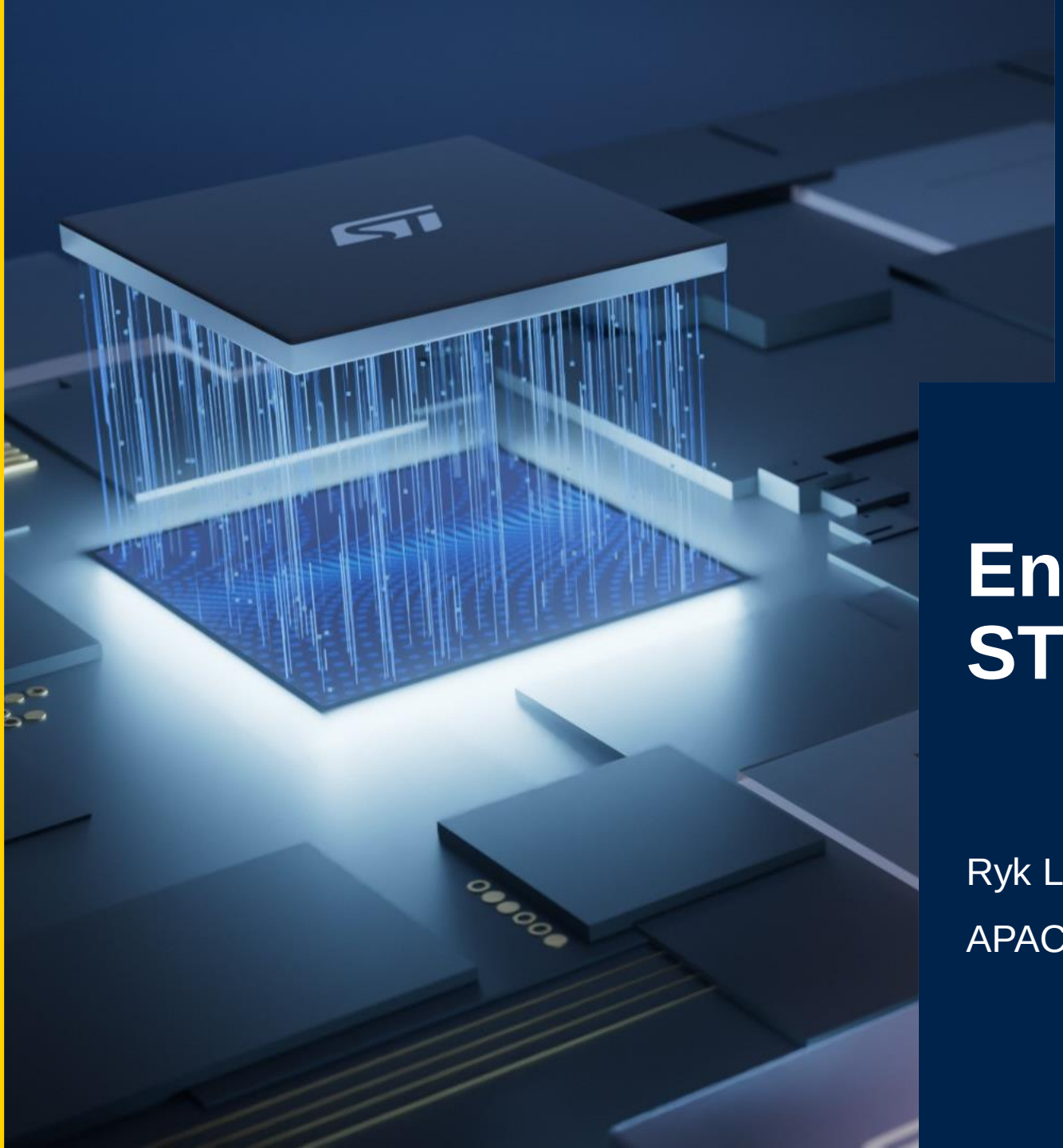


life.augmented

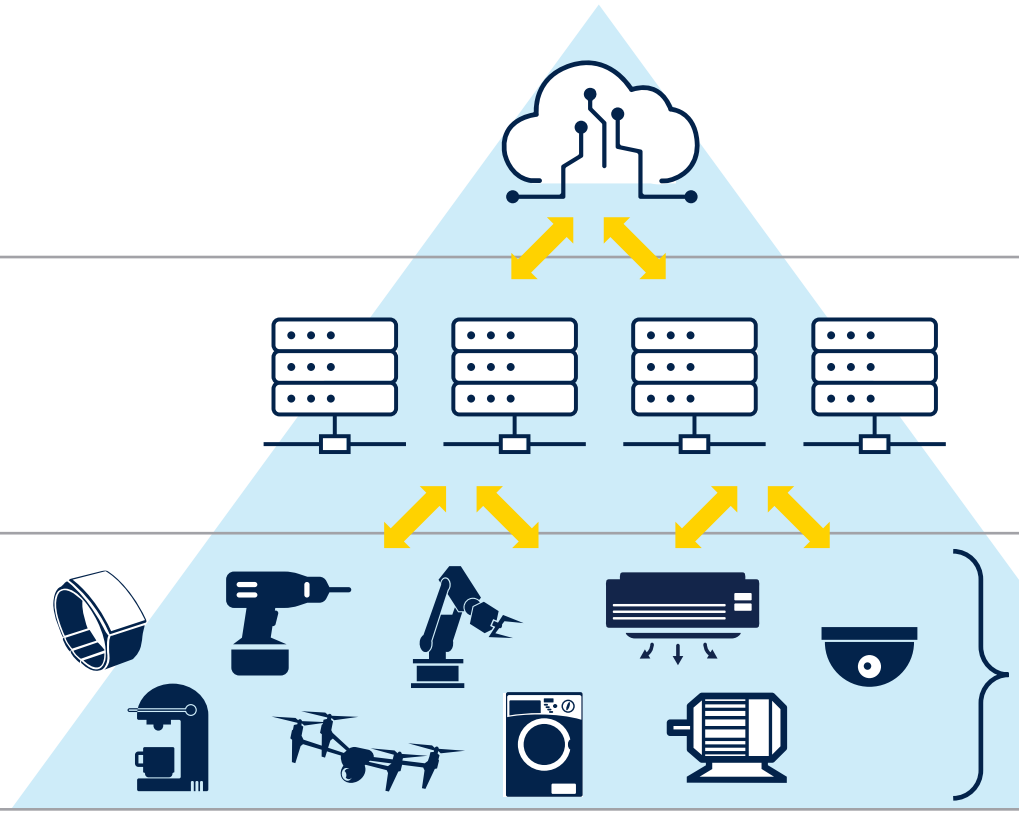
Enabling edge AI with ST Solutions

Ryk Liu,

APAC AI Competence Center

Distributed Artificial Intelligence approach

Leverage billions of devices at the edge!



Data Center Cloud

Analytics, storage, compute

Edge Nodes

IoT gateways, micro datacenters

Edge Things

Real time, local processing

life.augmented

Edge AI application processing requirements

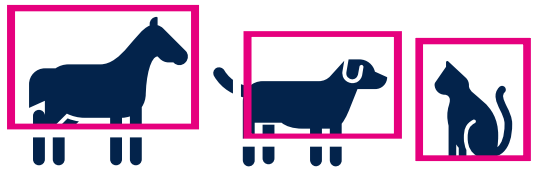
Low

- Sensor analysis
- Activity recognition (motion sensors)
- Stress analysis or attention analysis

Medium

- Multiple Objects Detection (low resolution/frame rate)
- Face/object analysis (e.g., face detection)
- Audio (keyword spotting, scene detection)

High



- Multiple Objects (high resolution/frame rate) detection/classification/tracking
- Speech (sentences) recognition
- Speech synthesis

Smart Sensors / STM32 → **next generation MCUs/MPUs**

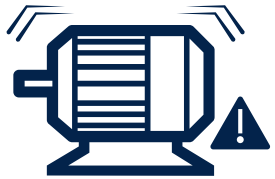
iNEMO · UNICO · ISPU · STM32 · STM32 Cube.AI · NANOEDGE AI STUDIO · X-LINUX-AI OpenSTLinux Expansion Package

life.augmented

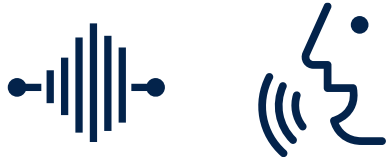
ST’s product offering

Enabling major edge AI technologies

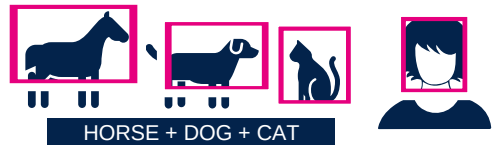
SENSING (TIME SERIES)



AUDIO



VISION



Software tools for any user profile

SENSING
AD

NANOEDGE AI STUDIO

User-friendly Auto-ML tool for STM32 MCUs

SENSING
AUDIO
VISION

STM32 Cube.AI

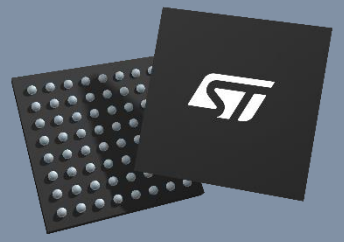
AI model optimizer and code generator for STM32 MCUs

SENSING
AUDIO
VISION

X-LINUX-AI

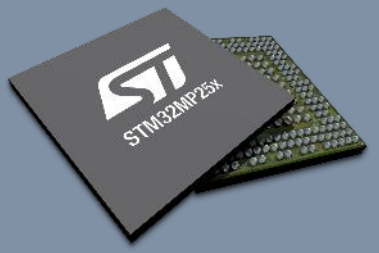
A complete AI framework for STM32 MPUs

Large choice of general purpose & accelerated hardware



STM32 MCUs

STM32N6 MCU (available soon)



STM32MP1 & STM32MP2 MPUs

NanoEdge AI Studio, the AutoML tool

840+ Active companies having completed at least one project in 2023

12,700+ Libraries compiled by business users in 2023

life.augmented

STM32Cube.AI, model optimizer and code generator

6,000+ **Unique active business users of STM32Cube.AI or STM32Cube.AI Developer Cloud in 2023**

19,000+ **STM32 projects created with STM32Cube.AI in 2023**

life.augmented

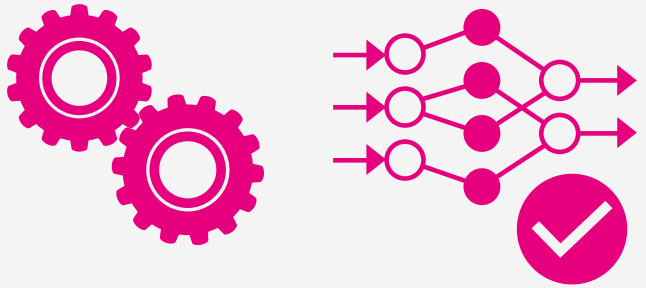
Edge AI development workflow – STM32Cube.AI

1 Data preparation

Data acquisition | Data processing

Data logging and curation tools

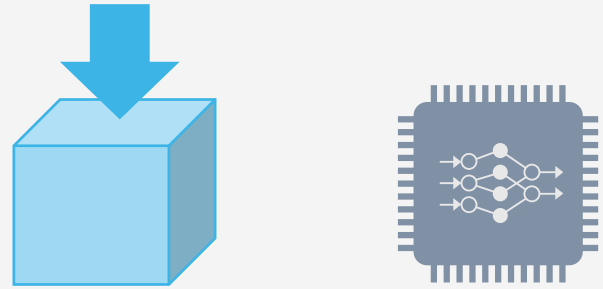
2 Model development



Model selection and training | Model validation

scikit learn | Keras | TensorFlow Lite | PyTorch | ONNX

3 Model implementation



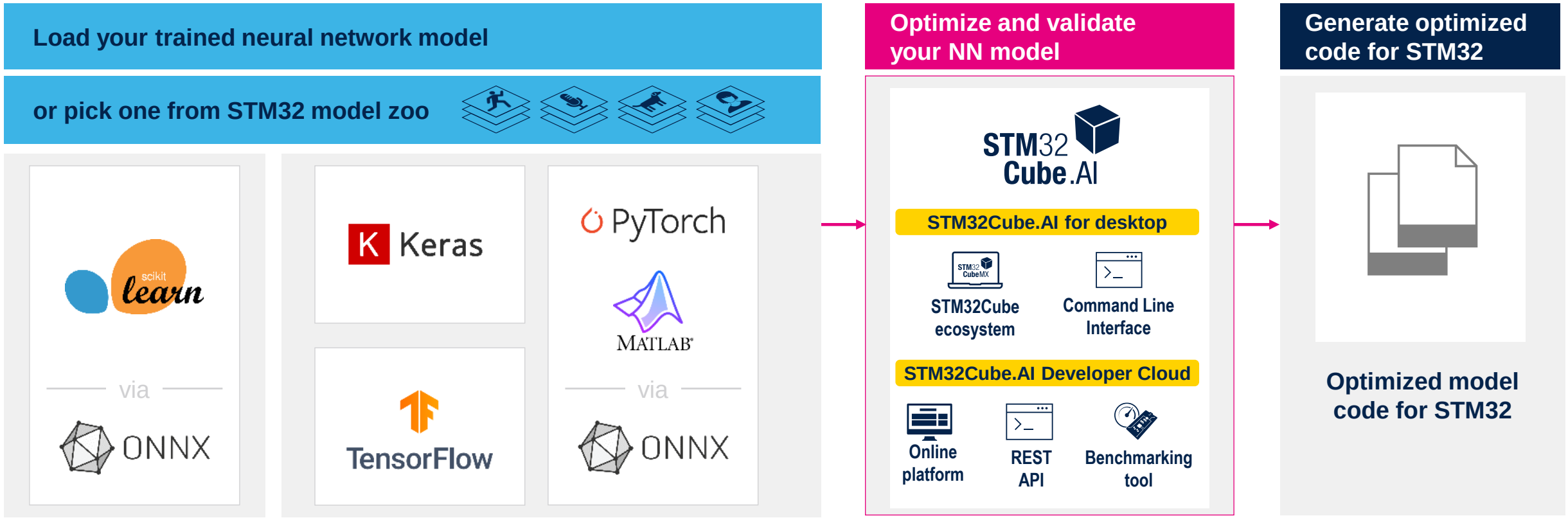
Model Optimization | Deployment

Edge AI toolkit

STM32 Cube.AI | STM32

life.augmented

One tool – two versions to deploy AI on STM32



STM32 Cube.AI

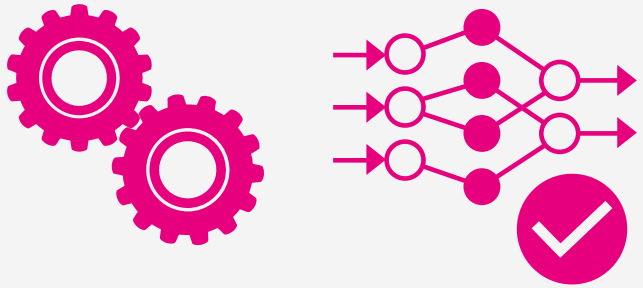
AI development workflow – NanoEdgeAI Studio

1 Data preparation

Data acquisition

Data processing

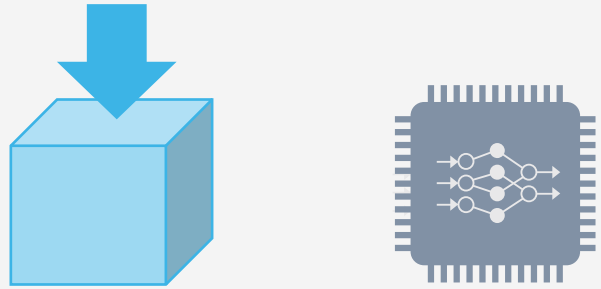
2 Model development



Model selection and training

Model testing

3 Model implementation



Model library creation

Model inference

Data logging tools

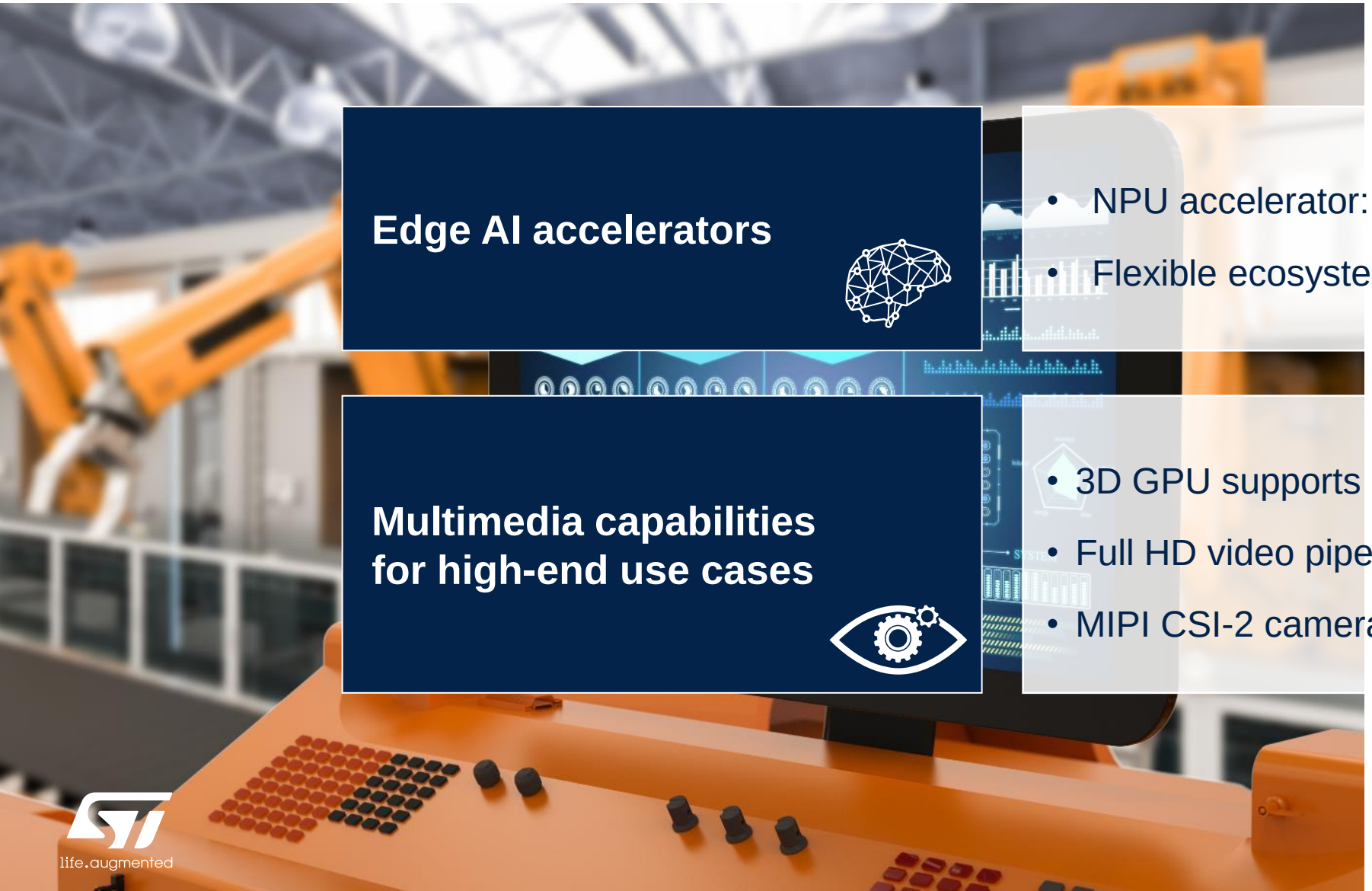
Automated Edge AI Software

NANOEDGE AI STUDIO

STM32

life.augmented

64-bit MPU with advanced edge AI capabilities



Edge AI accelerators

- NPU accelerator: **up to 1.35 TOPS**
- Flexible ecosystem to run AI on CPU, GPU, or NPU

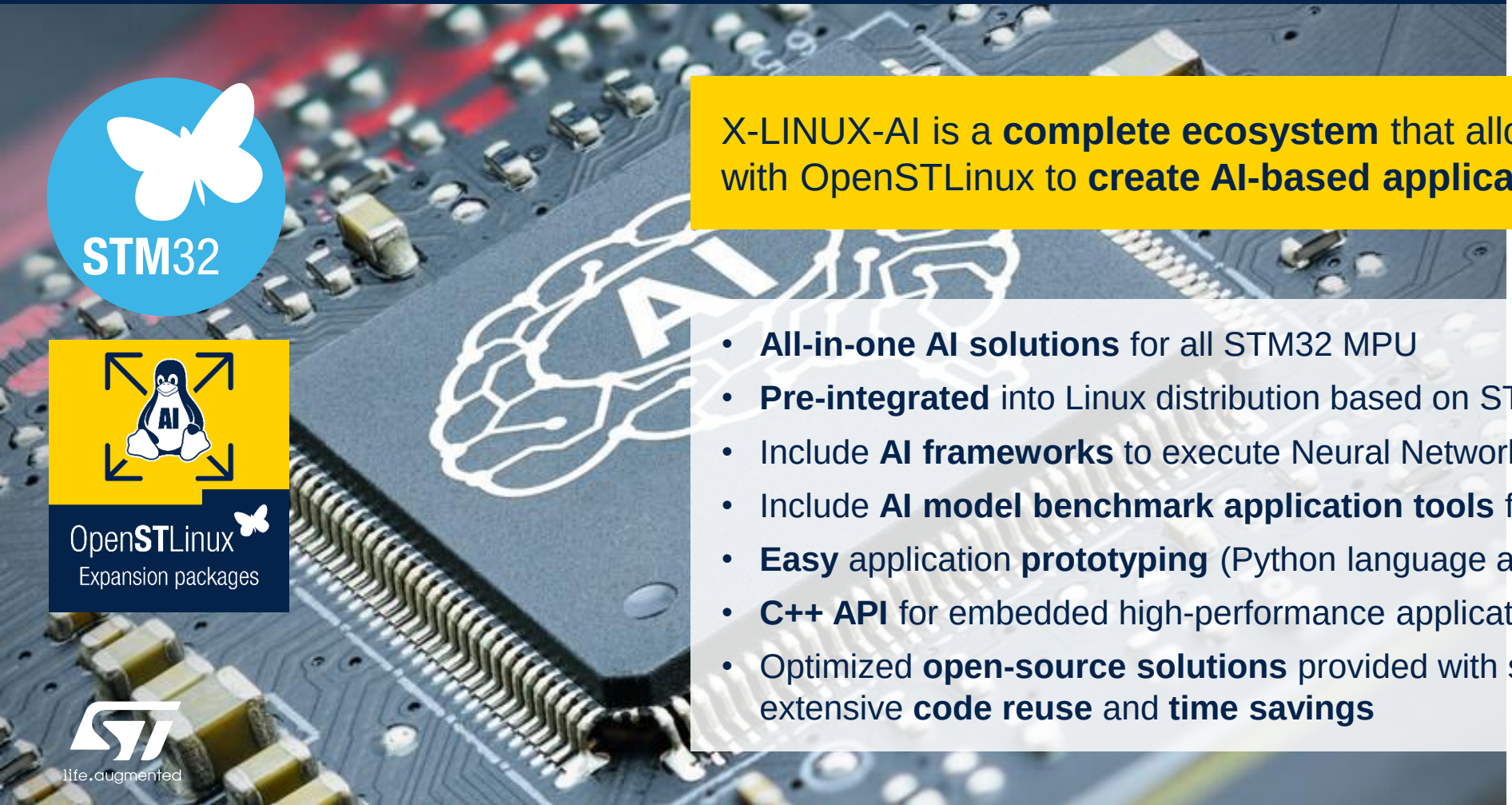
Multimedia capabilities for high-end use cases

- 3D GPU supports up to 1080p resolution
- Full HD video pipe with RGB, LVDS & DSI outputs
- MIPI CSI-2 camera interface with ISP

life.augmented

X-LINUX-AI: enabling edge AI on STM32 MPUs

A Free open-source software package for AI



STM32

AI

OpenSTLinux
Expansion packages

X-LINUX-AI is a **complete ecosystem** that allow developers working with OpenSTLinux to **create AI-based applications very easily**

- **All-in-one AI solutions** for all STM32 MPU
- **Pre-integrated** into Linux distribution based on ST environment
- Include **AI frameworks** to execute Neural Network models
- Include **AI model benchmark application tools** for MPU
- **Easy** application **prototyping** (Python language and AI frameworks Python API)
- **C++ API** for embedded high-performance applications
- Optimized **open-source solutions** provided with source codes that allow for extensive **code reuse** and **time savings**

life.augmented

Evolution of system-based AI

Till today

Sensor + GP MCU

AI running on general purpose MCU

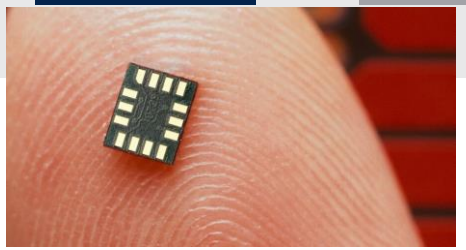
Sensor — MCU

Today

ISPU + MCU

ISPU integrated in the sensor ASIC

ISPU + Sensor — MCU



Tomorrow

ISPU 2.0 + AI MCU

AI distributed among ISPU and AI MCU

ISPU + Sensor — AI MCU

ISPU

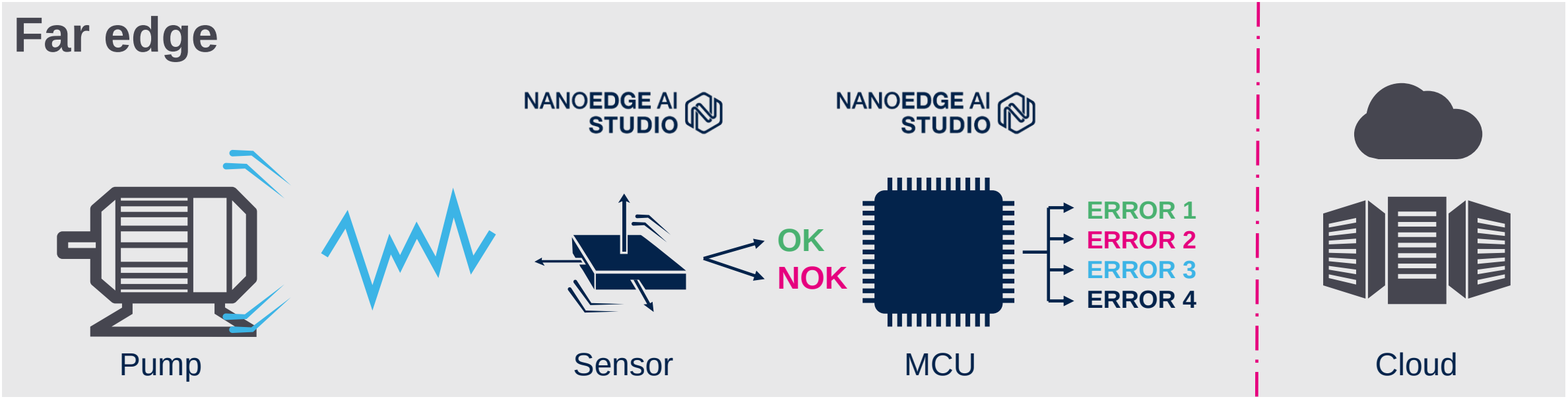
Intelligent Sensor Process Unit

Small area down to 8 Kgates	**Standard package** 3 x 2.5 x 0.83 mm	**Optimization** Power consumption vs. performance
Full precision Floating Point Unit	**Binary Neural Network** convolution acceleration	**RAM based** 40 kB (program + execution)

life.augmented

NanoEdge AI running on sensors
One step closer to the signal

Possibility to distribute the tasks between the sensor and the MCU



life.augmented

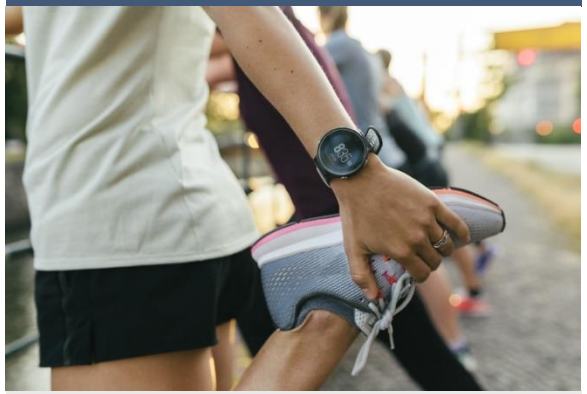
Integrate your ML models more easily with our application-oriented code examples

Time series-based monitoring

FP-AI-MONITOR1

- Predictive maintenance and much more sensor-monitoring apps
- Runs Libraries from NanoEdge AI Studio

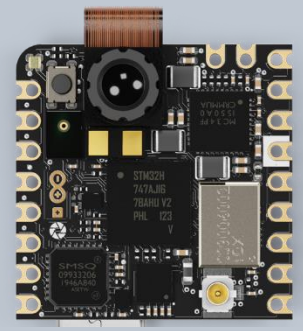
Audio and sensing



FP-AI-SENSING1

- Human activity recognition
- Acoustic scene classification
- Data logging, labeling, and result on Bluetooth® Low Energy applications

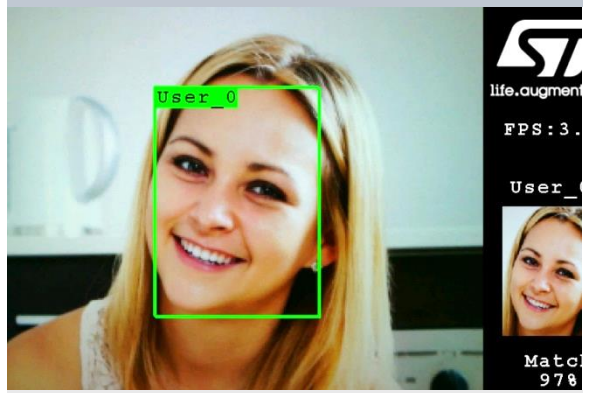
Computer vision



FP-AI-VISION1

- Food recognition (CNN)
- Person presence detection (CNN)
- People counting (object detection NN)
- Image processing library

Face recognition



FP-AI-FACEREC1

- Face detection and recognition
- Fully functional without cloud connection

life.augmented

A proven technology already adopted by multiple clients

NANOEDGE AI STUDIO

INDUSTRIAL | DEMO

Fan anomaly detection based on vibrations

Learn to detect abnormal behavior at the edge on a vibrating machine.

nke WATTECO

NANOEDGE AI STUDIO

INDUSTRIAL | CUSTOMER

AI solution for industrial predictive maintenance with NKE Watteco

Predictive maintenance solution for industrial equipment.

CROUZET

NANOEDGE AI STUDIO

TRANSPORTATION | CUSTOMER

AI solution for monitoring automatic doors with Crouzet

Predictive maintenance on motors for automatic door motors.

NANOEDGE AI STUDIO

INDUSTRIAL | DEMO

Anomaly detection in an electric motor

Current sensing to detect abnormal behaviors in motors.

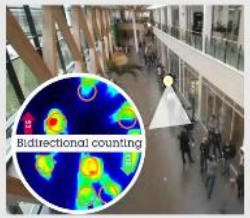
OXYtronic

NANOEDGE AI STUDIO

INDUSTRIAL | CUSTOMER

AI solution for industrial predictive maintenance with Oxytronic

Predictive maintenance solution for industrial equipment.



Schneider Electric

STM32 Cube.AI

SMART OFFICE | CUSTOMER

People flow counting Sensor with Schneider Electric

An innovative approach to measure people flows using an in-house thermal sensor.

STM32 Cube.AI

SMART CITY | DEMO

Acoustic scene classification

Identify different environments (indoor, outdoor, in-car) using a simple microphone.

STM32 Cube.AI

WEARABLES | DEMO

Human Activity Recognition

Easily identify 5 different activities with a 3D accelerometer.

STM32 Cube.AI

INDUSTRIAL | DEMO

People presence detection (visual wake word)

Human detection on high-performance MCU.

STM32 Cube.AI

INDUSTRIAL | DEMO

Aftermarket wireless digit reader

Equip meters with aftermarket wireless & low-power readers.

life.augmented

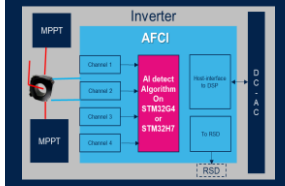
Find out more use cases: https://stm32ai.st.com/use-cases

Reference designs developed at the AI competence center

AI Competence Center

Material recognition based on ToF

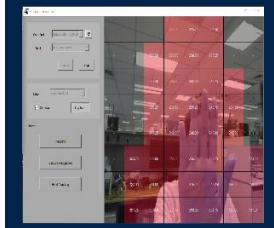
Face detection for thermal camera



Solar Arc Fault detection

WM clothes weight estimation

(*) WM drum anti collision system



3D ToF Gesture recognition

(*) Industrial radar people detection

Water meter digit recognition

And more available or under development…

(*) POC Demo developed under close collaboration with customer (proprietary dataset).

life.augmented

life.augmented