

tinyML[®] EMEA

Enabling Ultra-low Power Machine Learning at the Edge

tinyML EMEA Technical Forum 2021 Proceedings

June 7 – 10, 2021

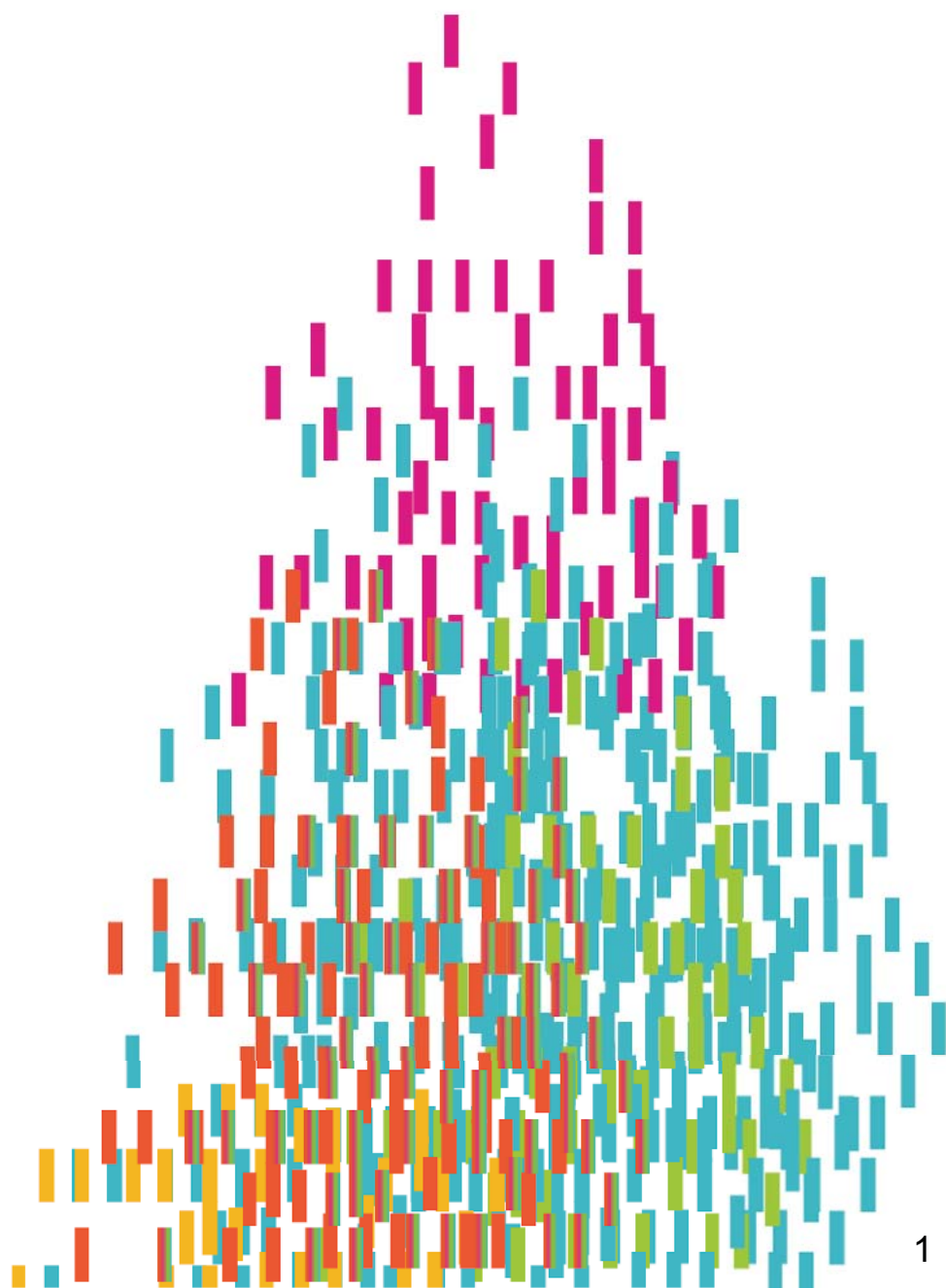
Virtual Event



www.tinyML.org

EON Tuner

AutoML for real-world
embedded devices



Hi I'm Jan!

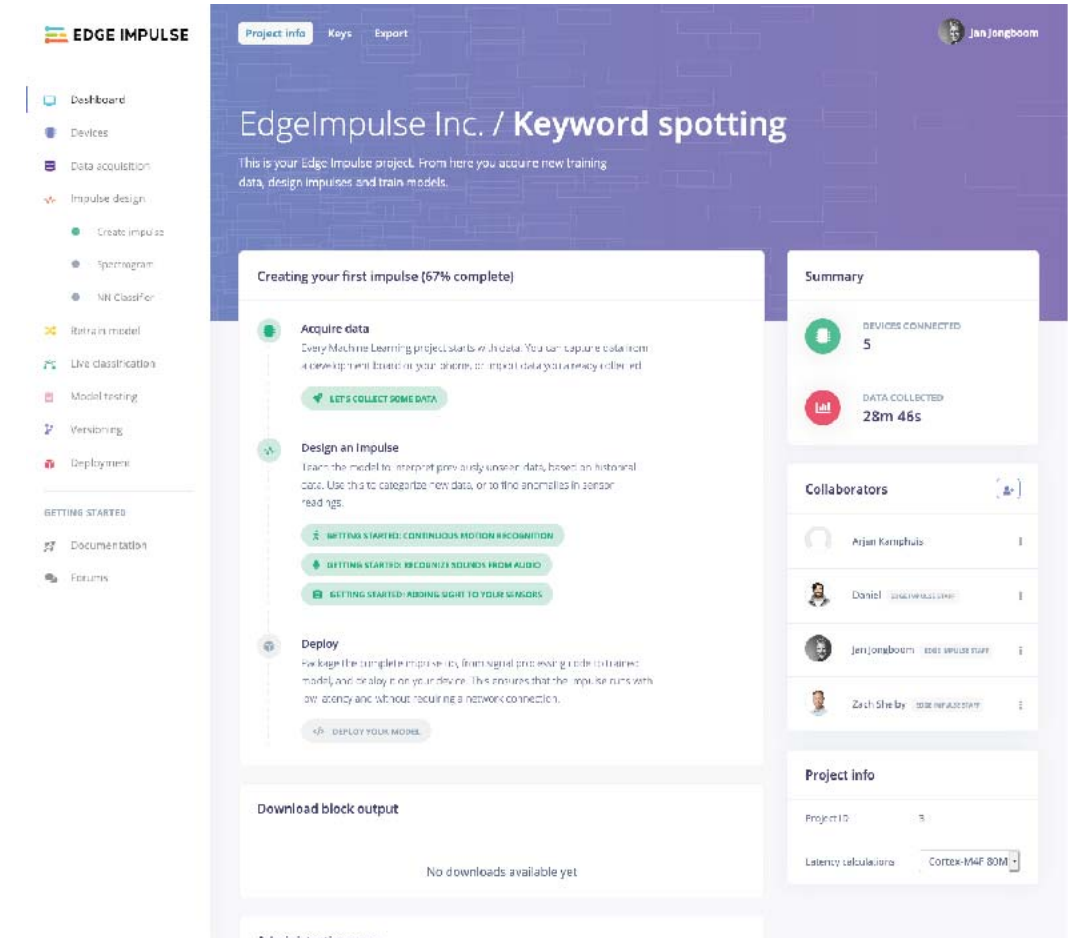
I cofounded **Edge Impulse** in 2019

The place to build embedded Machine Learning models

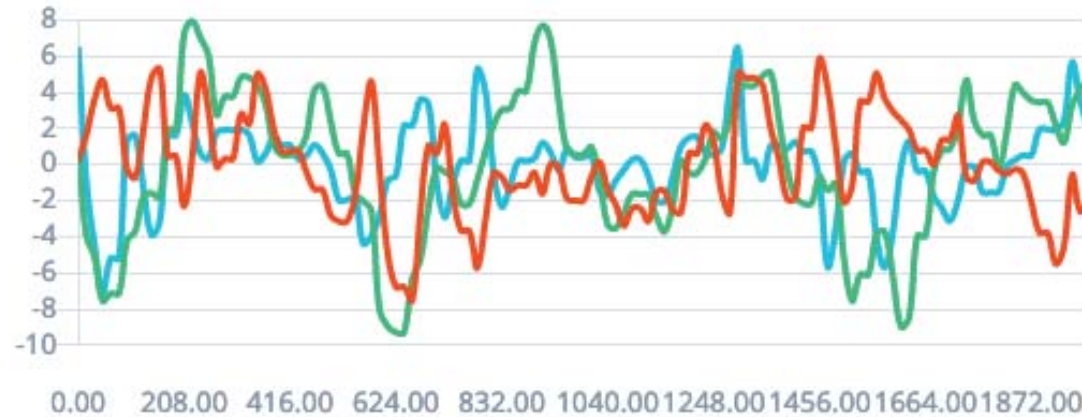
From engineers (not data scientists!) for engineers

Every step of the way, from data collection to deployment

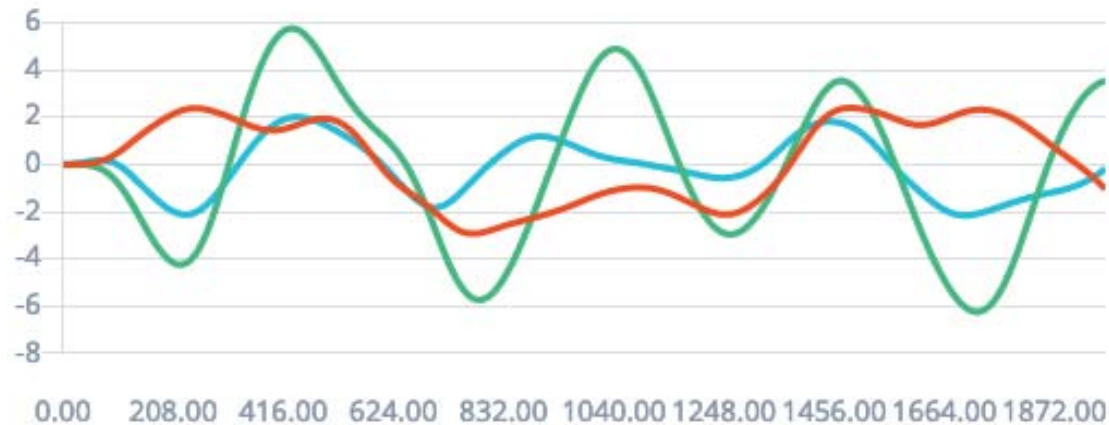
Already >26,000 real ML projects created!



Signal processing + ML = ❤️



Apply low-pass filter...



= much easier job
for the ML algo



Leveraging signal processing

On-device intelligence is not new

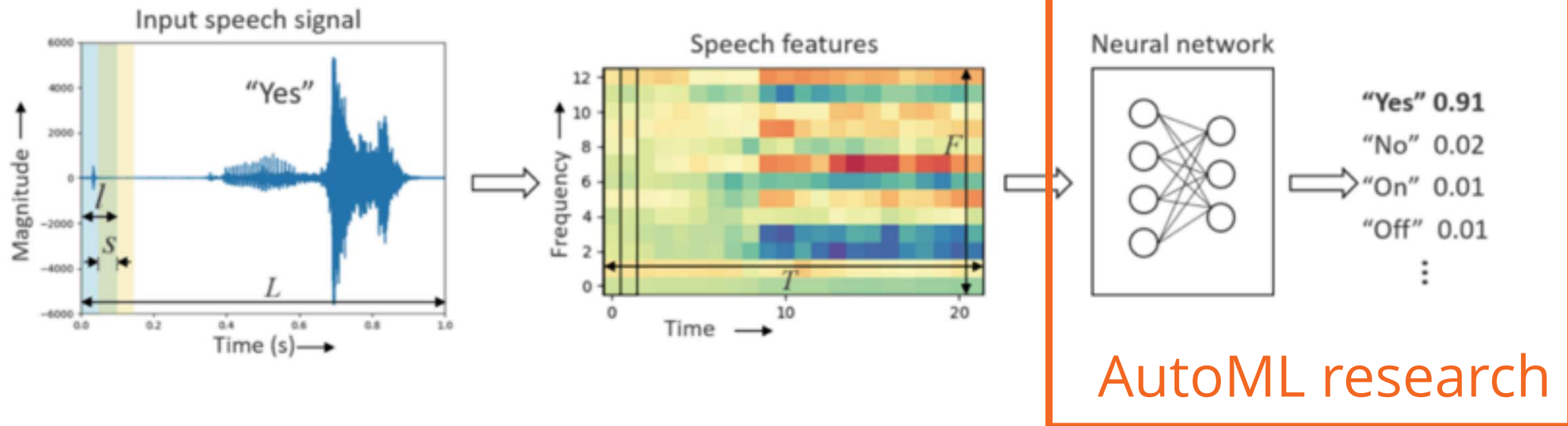
Neural networks are inefficient, if you can preprocess? Do so!

Significantly reduce input features, leading to smaller networks.

Cleans up input



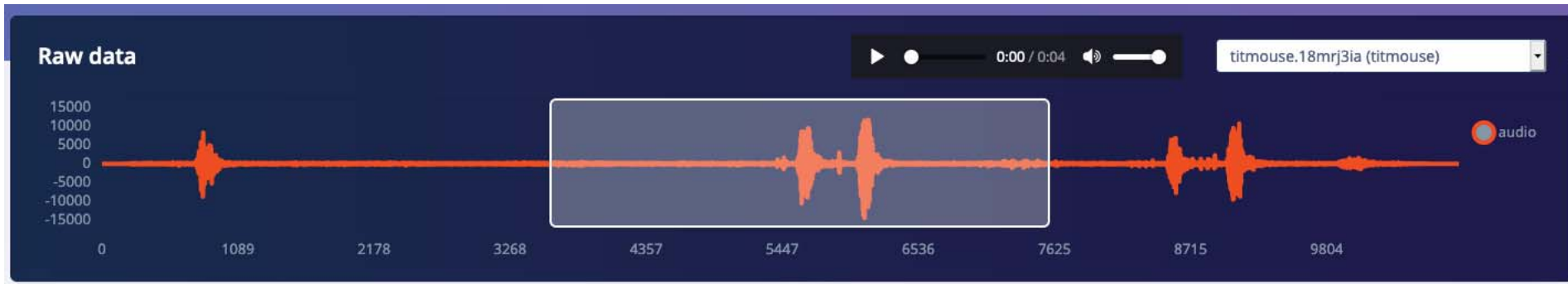
ML Sensor pipeline is often combination



ML Sensor pipeline is often combination



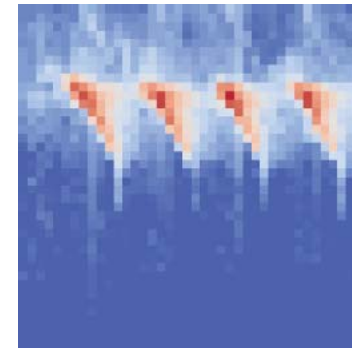
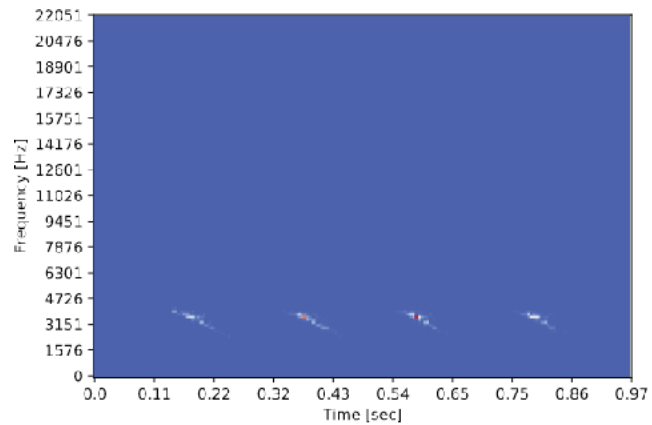
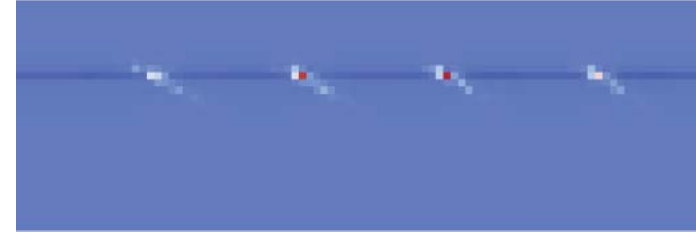
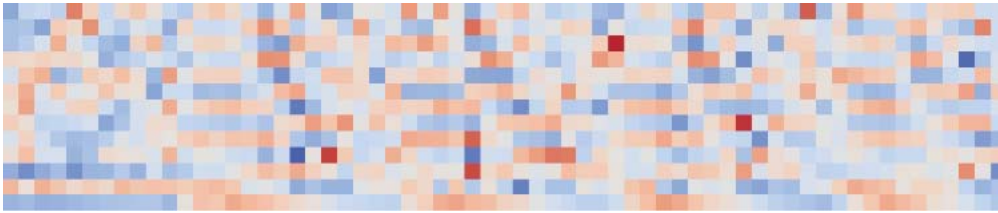
Wide range of parameters



Window length, window step, downsample?



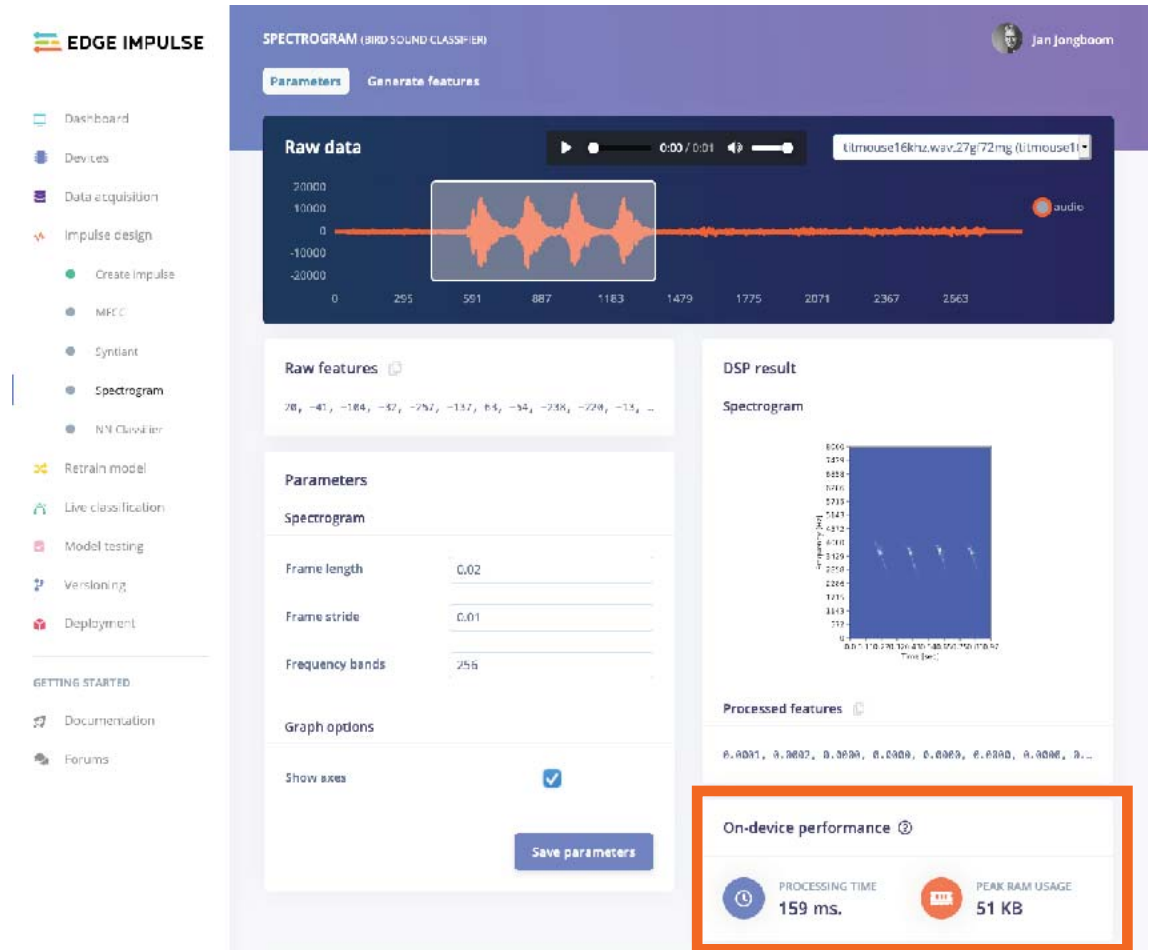
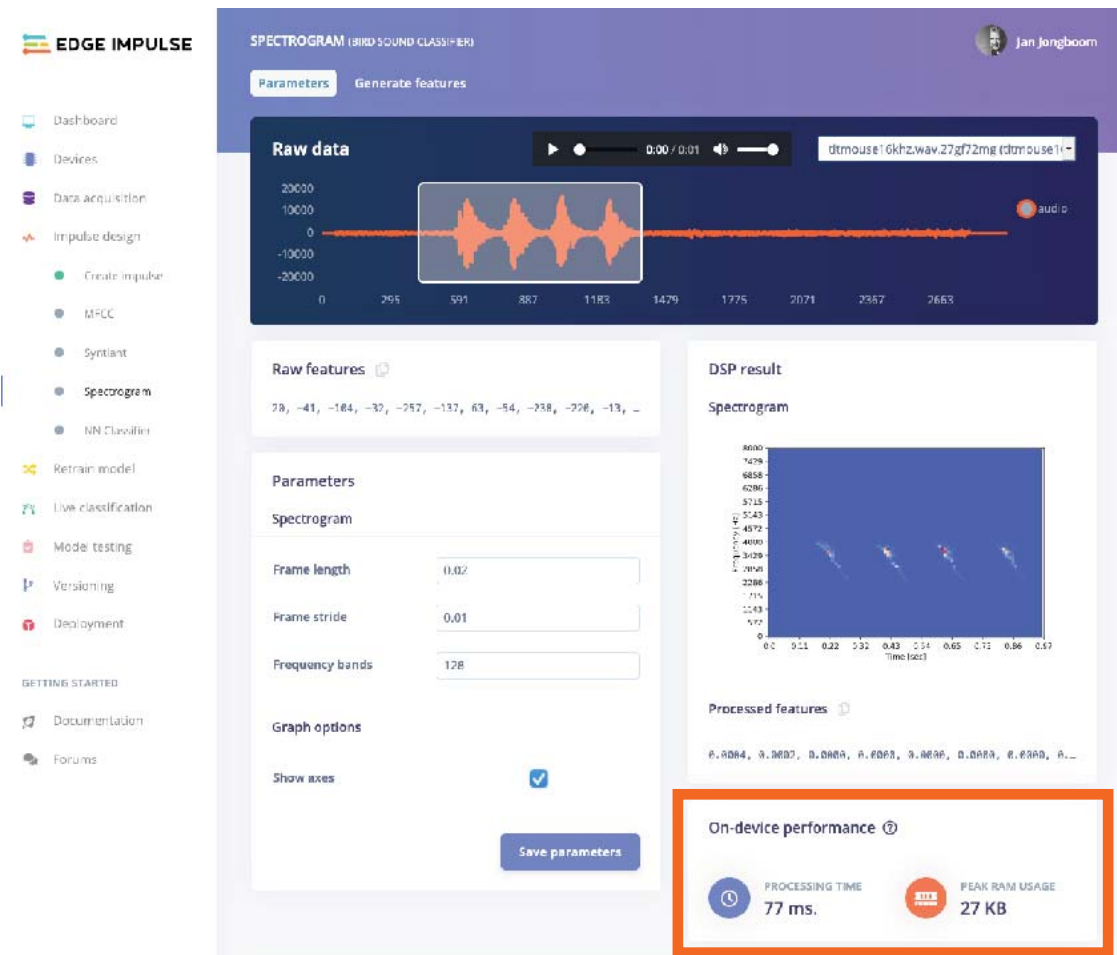
Wide range of parameters



+ endless configuration options



Constrained targets - what's worth it?



Introducing the EON Tuner!

 **EDGE IMPULSE**

 Dashboard

 Devices

 Data acquisition

 Impulse design

 Create impulse

 MFCC

 NN Classifier

 Retrain model

 Live classification

 Model testing

 Versioning

 Deployment

 EON Tuner

EON TUNER (DCASE2018)

 mathijs

EON Tuner

 Config

 Logs

Finding the most optimal architecture for your model (17 model variants evaluated / 50 variants total)



 cortex-m7-216mhz  1000 ms  RAM: 128kB • ROM: 1024kB  

97% 10x10 | SYNTIA | Clone | a65 

PERFORMANCE



ACCURACY

too	97	3	0	0	0	0
dis	2	94	4	0	0	0
eat	0	0	100	0	0	0
soc	4	0	2	93	0	0
vac	0	0	0	0	100	0

97% 10x10 | SYNTIA | Clone | 133 

PERFORMANCE



ACCURACY

too	98	2	0	0	0	0
dis	4	88	8	0	0	0
eat	0	0	100	0	0	0
soc	2	2	0	96	0	0
vac	0	0	0	0	100	0

96% 10x10 | SYNTIA | Clone | 2f1 

PERFORMANCE



ACCURACY

too	95	5	0	0	0	0
dis	2	92	4	2	0	0
eat	0	3	98	0	0	0
soc	2	0	2	96	0	0
vac	0	0	0	0	100	0

 Filters

Status

- ☒ Pending
- ☒ Running
- ☒ Completed
- ☒ Failed

 View

Data set

- ☐ Validation
- ☐ Train

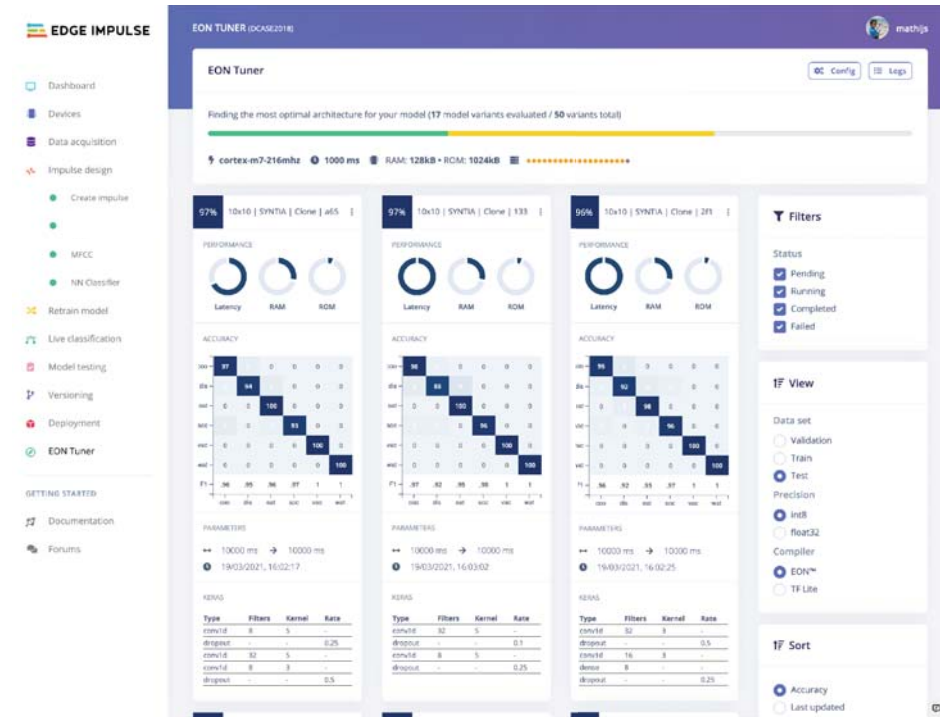
EON Tuner

Find best model for sensor data over mix of input blocks, DSP blocks and ML blocks

Specify device constraints

Extensible with your own DSP and ML blocks

Biggest win: "we found DSP configuration that works so well, we no longer need ML"



Still an engineering tool!

Model search completed (100 model variants evaluated)

EON Tuner settings

Target Advanced

Find the optimal architecture for your machine learning model

The EON™ Tuner will evaluate many candidate model architectures (selected based on your target device and latency requirements) concurrently to help you find the best performing architecture for your application.

The search process can take up to 24 hours to complete. We will notify you by e-mail on completion of the search process. While the search is running you can view the progress on this page at any time.

Dataset category

Speech (Keywords)

- Speech (Keywords)
 - Yes/no
 - Hello world
- Speech (Continuous)
 - Emotions
 - Speaker gender
- Audio (Events)
 - Breaking glass
 - Gunshot
- Audio (Continuous)
 - Household activities
 - Traffic level

? Help me choose the right dataset category

0.95 mfcc-keras-9ae ☆

PERFORMANCE ⓘ

250 ms 128 kB 1024 kB

DSP NN Unused

ACCURACY ⓘ

hel	98	2	0
noi	2	92	6
unk	7	5	88
F1	.95	.93	.91

hel noi unk

INPUT ⓘ

↔ 1000 ms → 1000 ms

MFCC ⓘ

↔ 0.02 → 0.01 | 13 | 32

KERAS ⓘ

Type	Filters	Kernel	Rate
conv1d	8	5	-
dropout	-	-	0.25
conv1d	8	5	-
dropout	-	-	0.25

01/05/2021, 02:33:56

0.95 mfcc-keras-e8e ☆

PERFORMANCE ⓘ

250 ms 128 kB 1024 kB

DSP NN Unused

ACCURACY ⓘ

hel	97	0	3
noi	2	91	7
unk	6	3	91
F1	.95	.94	.9

hel noi unk

INPUT ⓘ

↔ 1000 ms → 1000 ms

MFCC ⓘ

↔ 0.04 → 0.02 | 13 | 32

KERAS ⓘ

Type	Filters	Kernel	Rate
conv1d	8	3	-
dropout	-	-	0.1
conv1d	16	5	-
dense	8	-	-
dropout	-	-	0.25

01/05/2021, 02:35:14



Eta: soon! 🚀

www.edgeimpulse.com



Premier Sponsor



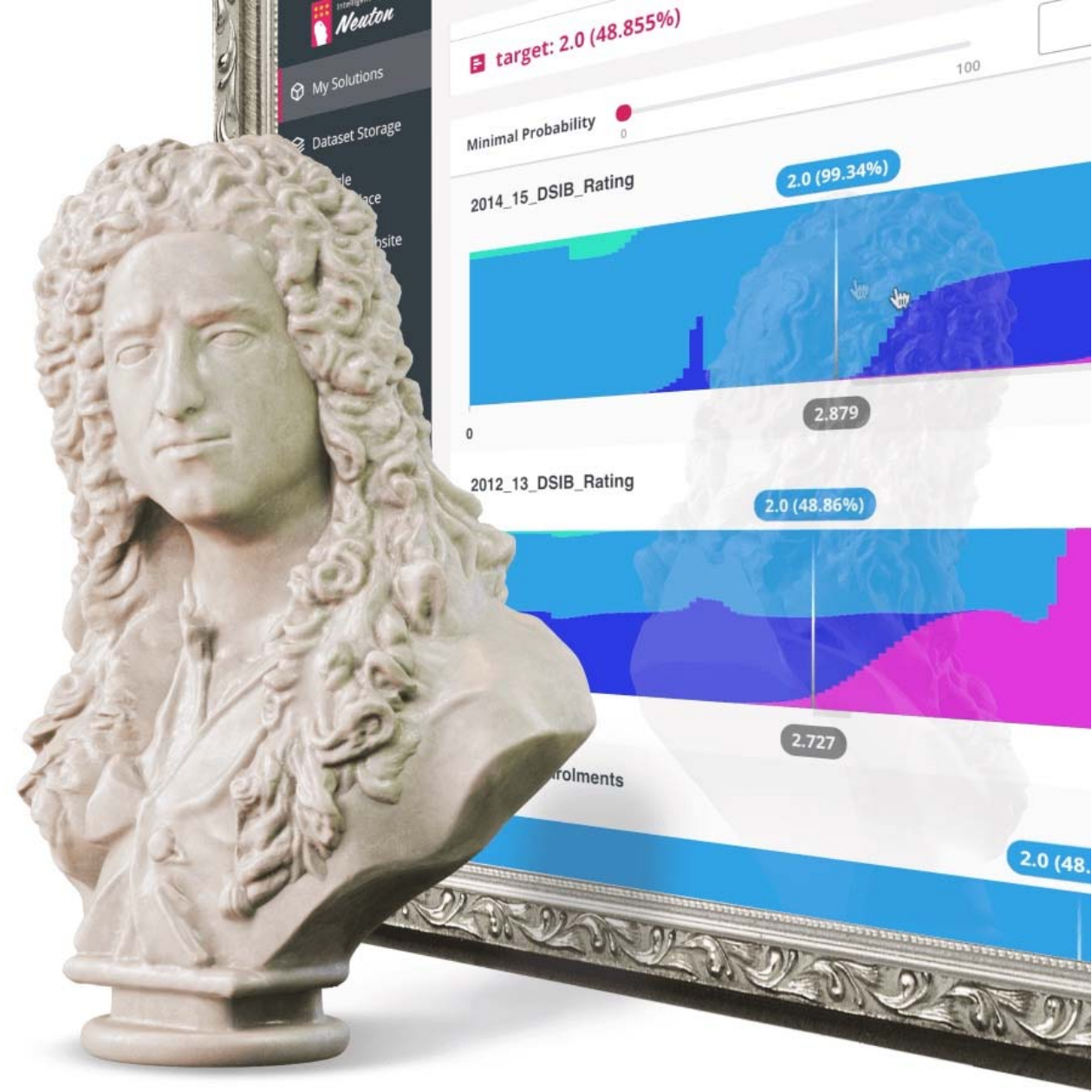
Automated TinyML

Zero-code SaaS solution

**Create tiny models, ready for embedding,
in just a few clicks!**

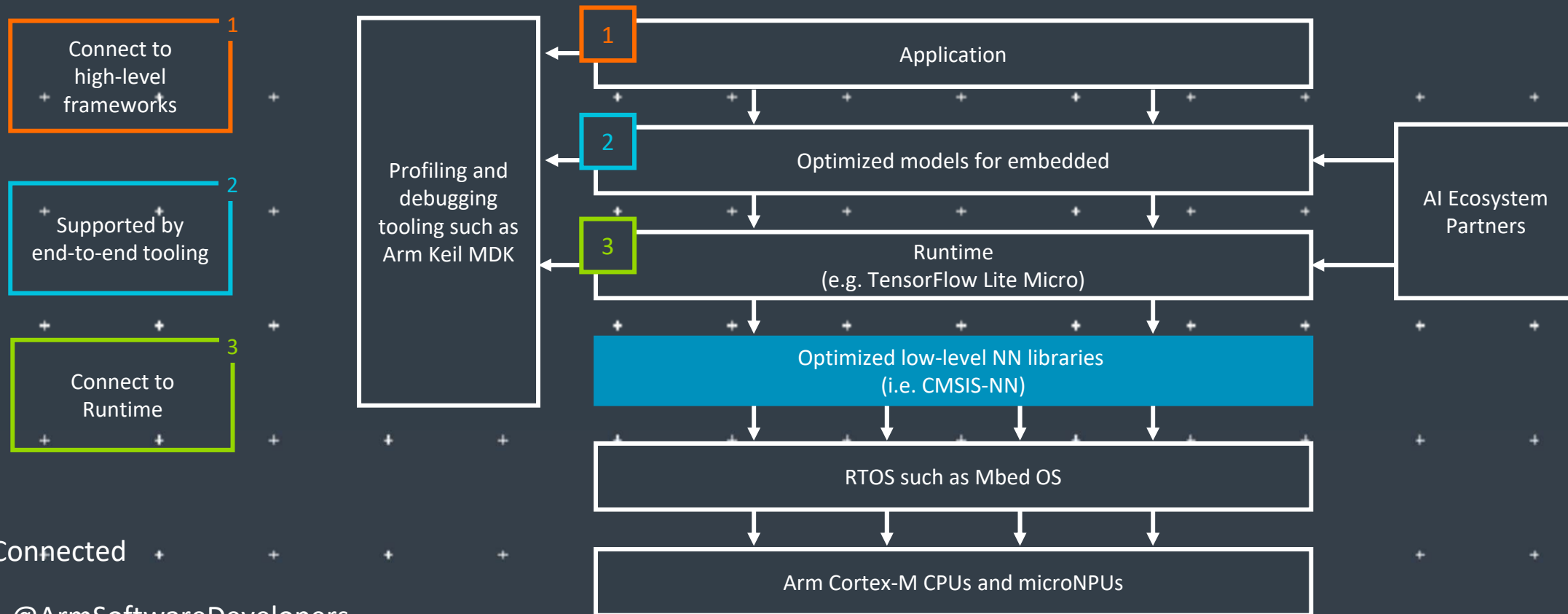
Compare the benchmarks of our compact models to those of TensorFlow and other leading neural network frameworks.

Build Fast. Build Once. Never Compromise.



Executive Sponsors

Arm: The Software and Hardware Foundation for tinyML



Stay Connected



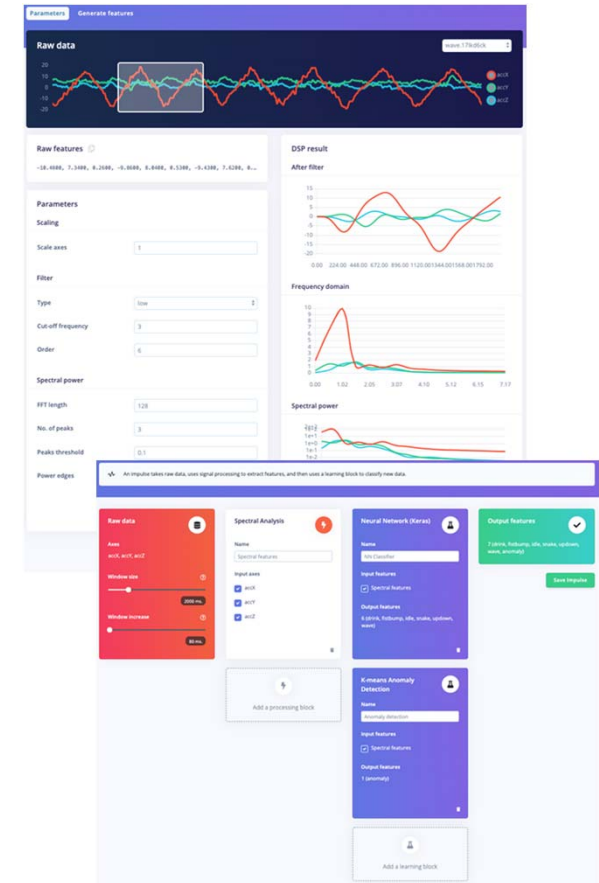
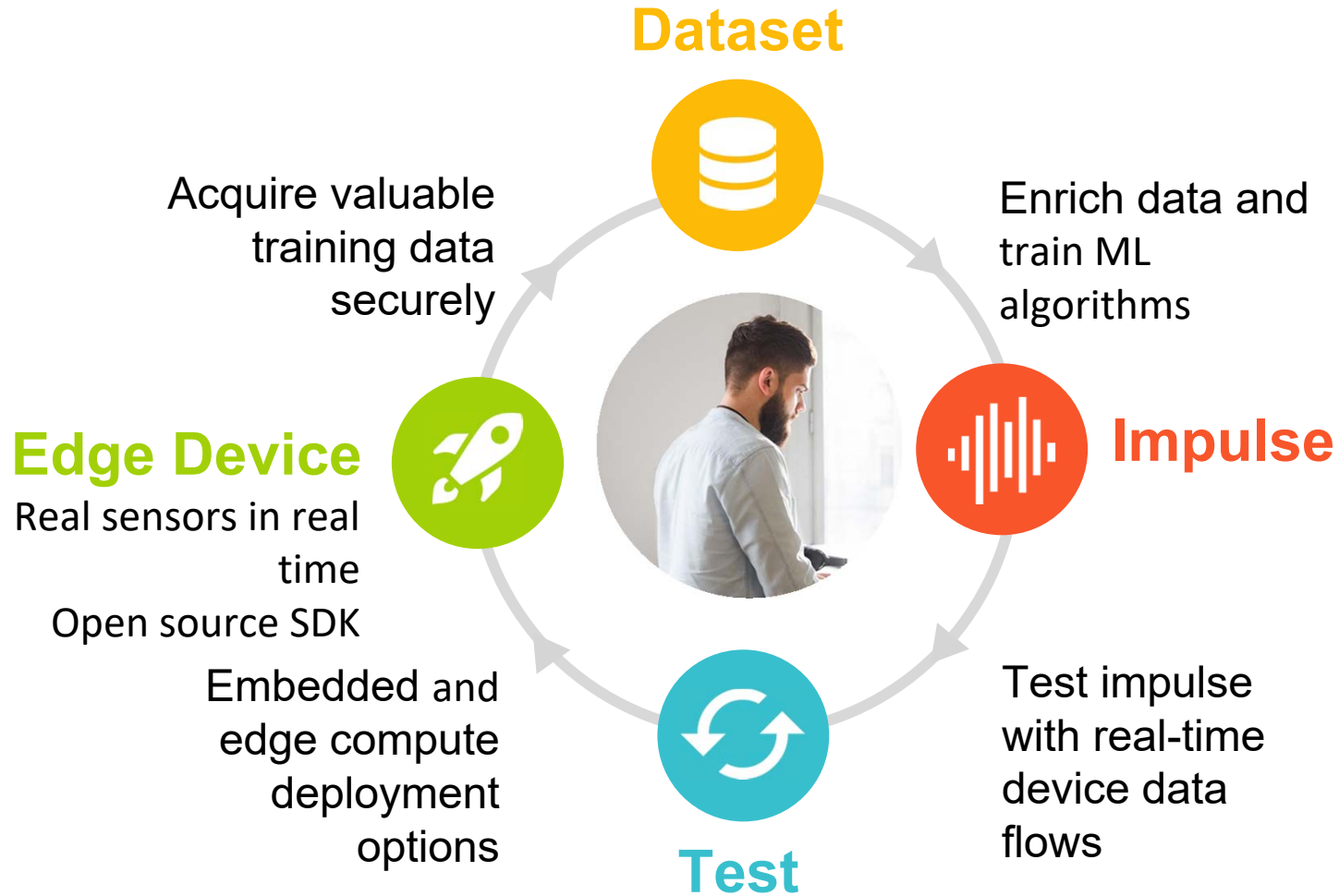
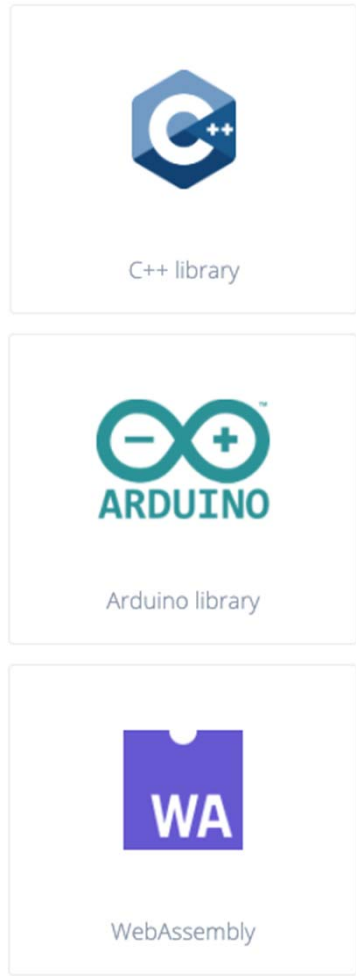
@ArmSoftwareDevelopers



@ArmSoftwareDev

Resources: developer.arm.com/solutions/machine-learning-on-arm

TinyML for all developers



www.edgeimpulse.com



Advancing AI research to make efficient AI ubiquitous

Power efficiency

Model design, compression, quantization, algorithms, efficient hardware, software tool

Personalization

Continuous learning, contextual, always-on, privacy-preserved, distributed learning

Efficient learning

Robust learning through minimal data, unsupervised learning, on-device learning

A platform to scale AI across the industry



Perception

Object detection, speech recognition, contextual fusion



Reasoning

Scene understanding, language understanding, behavior prediction



Action

Reinforcement learning for decision making



Edge cloud



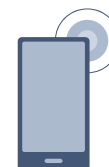
Cloud



IoT/IIoT



Automotive



Mobile

SYNTIANT

[Syntiant Corp.](#) is moving artificial intelligence and machine learning from the cloud to edge devices. Syntiant's chip solutions merge deep learning with semiconductor design to produce ultra-low-power, high performance, deep neural network processors. These network processors enable always-on applications in battery-powered devices, such as smartphones, smart speakers, earbuds, hearing aids, and laptops. Syntiant's Neural Decision Processors™ offer wake word, command word, and event detection in a chip for always-on voice and sensor applications.

Founded in 2017 and headquartered in Irvine, California, the company is backed by Amazon, Applied Materials, Atlantic Bridge Capital, Bosch, Intel Capital, Microsoft, Motorola, and others. Syntiant was recently named a [CES® 2021 Best of Innovation Awards Honoree](#), [shipped over 10M units worldwide](#), and [unveiled the NDP120](#) part of the NDP10x family of inference engines for low-power applications.

www.syntiant.com



@Syntiantcorp

Platinum Sponsors



Part of your life. Part of tomorrow.

www.infineon.com



Reality AI[®]

Add Advanced Sensing to your Product with Edge AI / TinyML

<https://reality.ai>



info@reality.ai



[@SensorAI](https://twitter.com/SensorAI)



[Reality AI](#)

Pre-built Edge AI sensing modules, plus tools to build your own

Reality AI solutions

Prebuilt sound recognition models for
indoor and outdoor use cases

Solution for industrial anomaly detection

Pre-built automotive solution that lets cars
“see with sound”

Reality AI Tools[®] software

Build prototypes, then turn them into
real products

Explain ML models and relate the function
to the physics

Optimize the hardware, including
sensor selection and placement

Gold Sponsors



LatentAI

Adaptive AI for the Intelligent Edge

latent.ai



Build Smart IoT Sensor Devices From Data

SensiML pioneered TinyML software tools that auto generate AI code for the intelligent edge.

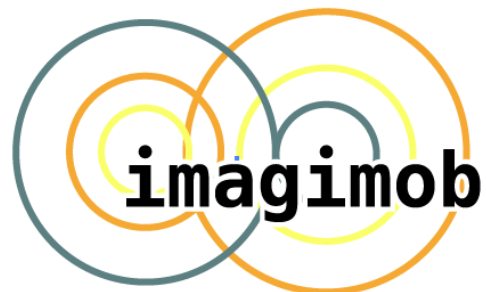
- End-to-end AI workflow
- Multi-user auto-labeling of time-series data
- Code transparency and customization at each step in the pipeline

We enable the creation of production-grade smart sensor devices.



sensiml.com

Silver Sponsors



Copyright Notice

The presentation(s) in this publication comprise the proceedings of tinyML® EMEA Technical Forum 2021. The content reflects the opinion of the authors and their respective companies. This version of the presentation may differ from the version that was presented at tinyML EMEA. The inclusion of presentations in this publication does not constitute an endorsement by tinyML Foundation or the sponsors.

There is no copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies and may contain copyrighted material. As such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author(s) or their companies.

tinyML is a registered trademark of the tinyML Foundation.

www.tinyML.org