



# THE DEVELOPER'S CONFERENCE

## **Trilha – Inteligência Artificial** **AI at the Edge – Plataformas e Casos de Uso**

**Fulvio Mascara**

Cientista-chefe @ Foursys

Community Manager @ AI Brasil



THE  
DEVELOPER'S  
CONFERENCE



CARA...

- Programa desde os 11 anos
- Gosto de cozinhar e viajar
- Corredor de media e longa distância
- "Tocador" de violão e guitarra



... E CRACHÁ!

- 30 anos de carreira
- Cientista-chefe (P&D&I) da Foursys
- Community Manager – AI Brasil
  
- Tecnólogo em Proc. Dados (Mackenzie)
- Pós-Graduado em Solution Architect (FIAP)
- Pós-Graduado em Estatística Aplicada (FMU)

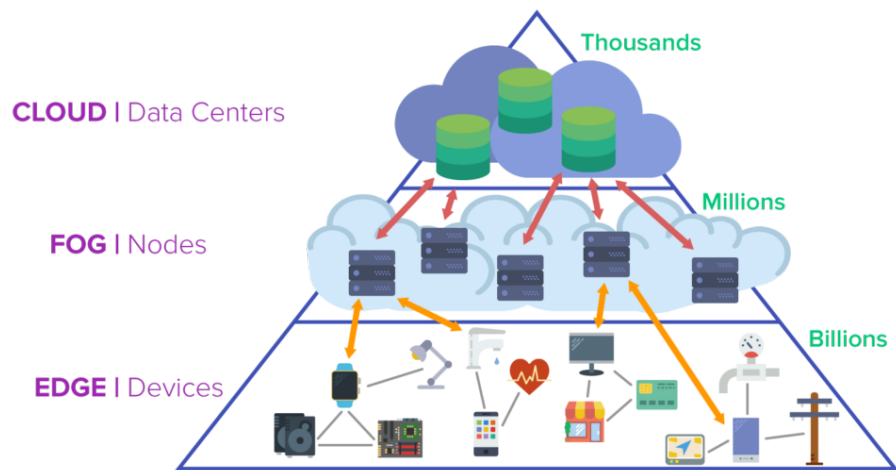
# Agenda



- Conceitos
- Principais Plataformas de Edge AI
- Exemplos de Aplicações
- Demonstração
  - Raspberry Pi
  - Jetson Nano
- Expectativas

# Conceitos

## Edge Computing X Fog Computing



- Informação é processada em uma LAN próxima aos Devices, em um Gateway IOT ou um Fog Node
- Faz a intermediação entre Edge Computing e Cloud Computing
- Informação é processada no Device (IOT)
- Computação e Storage residem no Device, na Aplicação ou no Componente que produz o dado



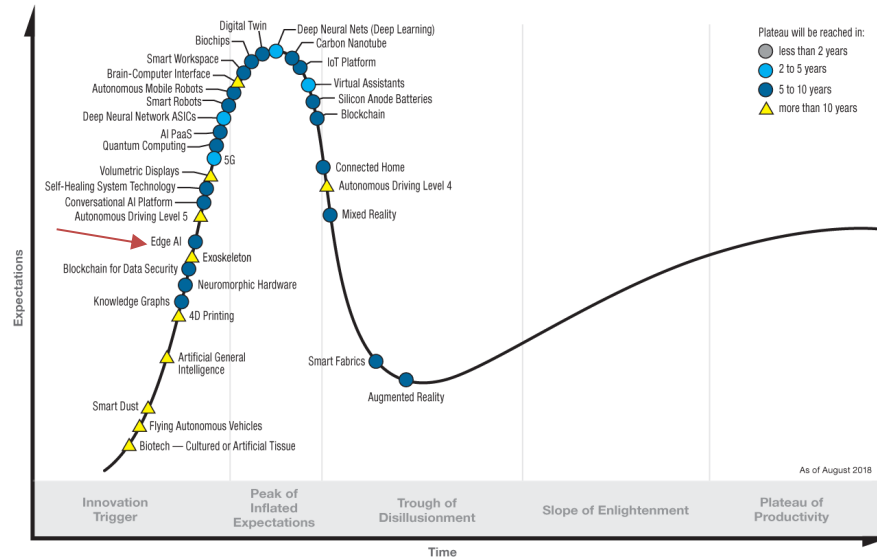
# Conceitos

## Edge AI no Radar



THE  
DEVELOPER'S  
CONFERENCE

### Hype Cycle for Emerging Technologies, 2018



[gartner.com/SmarterWithGartner](https://gartner.com/SmarterWithGartner)

Source: Gartner (August 2018)  
© 2018 Gartner, Inc. and/or its affiliates. All rights reserved.



# Conceitos

## Motivações para AI at the Edge



- Reduzir latência entre a produção do dado e a decisão do que fazer com ele
  - Detectar anomalias
  - Tomar decisões em tempo real
  - Produzir mais dados
  - Melhorar continuamente os modelos
- Redução de Custos
  - Provedores de Cloud
  - API Cognitivas

# Conceitos

## Barreiras para AI at the Edge



- Segurança
  - Física: Por conta do acesso físico, muitas vezes facilitados, aos dispositivos
  - Lógica: Na integração com Nodes de Fog Computing
- Capacidade de Processamento
  - O caso de uso ou aplicação requer um modelo ou algoritmo tão complexo, incapaz de ser processado em Edge
- Capacidade de Armazenamento
  - As aplicações de IA por vezes requerem um volume de informações que não conseguem ser mantidas num dispositivo pra Edge Computing
- Portabilidade
  - Por vezes, é possível colocar o armazenamento e processamento necessário para a aplicação, mas a portabilidade pode ser comprometida

# Exemplos de Aplicações

## Agricultura – Smart Farming

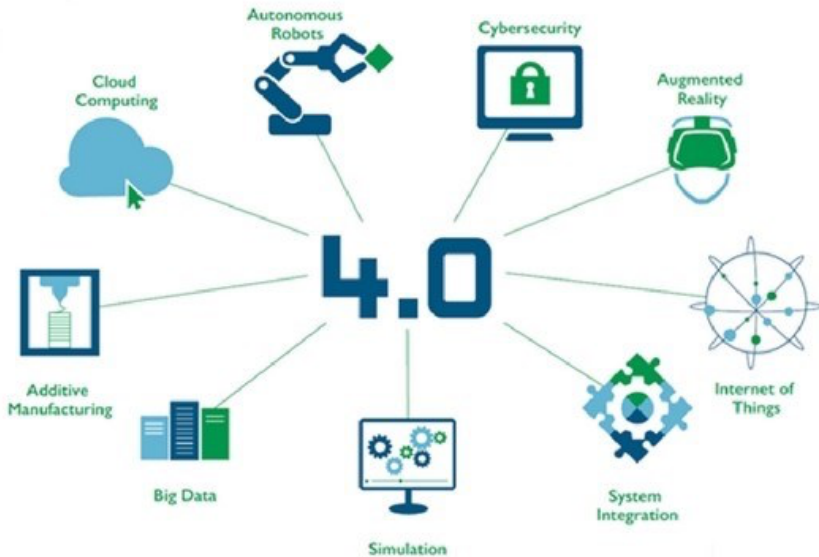


[https://s9.easternpeak.com/wp-content/uploads/2019/04/shutterstock\\_794528251-1.jpg](https://s9.easternpeak.com/wp-content/uploads/2019/04/shutterstock_794528251-1.jpg)

- **Agribots:** Através de Visão Computacional e mapeamento da área, pode realizar irrigação, plantio e colheita.
- **Farm Automation:** Sistemas autônomos com uso de IoT e modelos preditivos para gestão completa uma fazenda
- **Disaster Protection:** Através de sensores remotos, identificar anomalias em temperaturas ou demais sinais de meio ambiente e informar centros de controle.

# Exemplos de Aplicações

## Indústria – IIOT



- **Predictive Maintenance:** Com base na análise de sensores e comportamento dos mesmos na linha de produção
- **WMS:** Veículos autônomos para transportes de insumos produtivos e armazenagem
- **Disaster Protection:** Através de sensores remotos, identificar anomalias em temperaturas ou demais sinais de meio ambiente e informar centros de controle.

<https://cdn.insignia.com.au/Content/user/Insignia/img/Domino-Content-Industry-4.0-blog-image-2..jpg>

# Exemplos de Aplicações

## Saúde



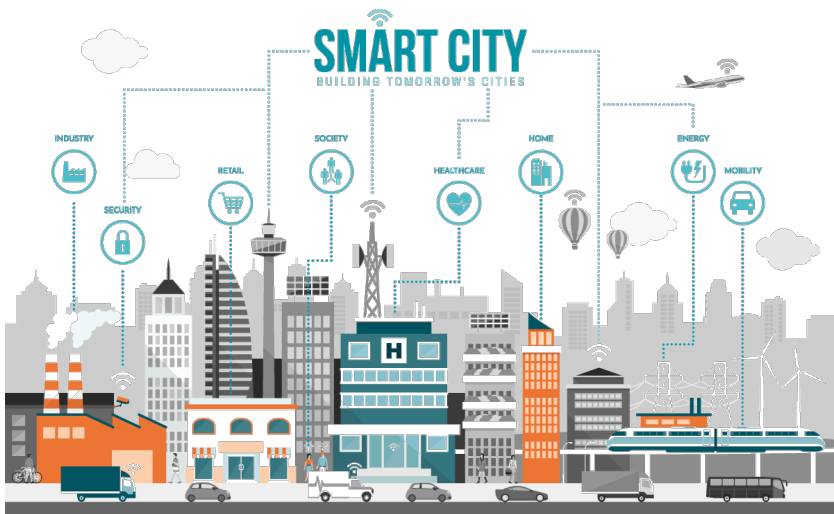
<https://www.aitrends.com/wp-content/uploads/2017/04/4-20AInHealthcare-2.jpg>

- **Wearables:** Coletando informações em real-time e indicando ações preventivas
- **Controle de Leitos:** Uso de Visão Computacional para identificar quedas de pacientes
- **Diabetes:** Dispositivos específicos de monitoramento, que utilizam machine learning para indicar tendências e sugerir ações
- **Telemedicina:** Dispositivos de coleta e auto-diagnóstico, como inputs para condutas médicas presenciais



# Exemplos de Aplicações

## Smart Cities



<http://cv.fee.unicamp.br/images/smart-city-1200px.gif>

- **Controle de Tráfego:** Com base em imagens de trânsito e controle de fluxo
- **Gestão hídrica:** Hidrômetros inteligentes, para detecção de anomalias
- **Gestão de energia:** Relógios inteligentes, com sistemas de cameras interligados

# Exemplos de Aplicações

## Segurança e Biometria



<https://www.argustrueid.com/wp-content/uploads/2017/11/3-Ways-Biometrics-LI.jpg>

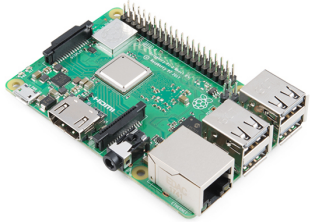
- **Controle de Acesso:** Detecção e Reconhecimento Facial, Reconhecimento por Voz
- **Biometria Comportamental:** Análise de padrões de movimento, gestos e expressões
- **Sistemas de Vigilância:** Câmeras inteligentes de segurança

# Principais Plataformas

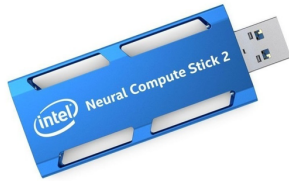
Edge Computing para AI – SBC and Sticks



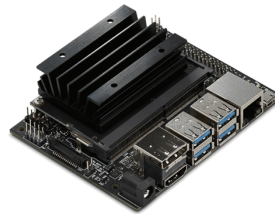
THE  
DEVELOPER'S  
CONFERENCE



Raspberry Pi



Intel Movidius



Jetson Nano



Google Coral



AWS DeepLens

# Principais Plataformas

## Raspberry Pi



# Caffe

 Keras

 TensorFlow

# PYTORCH

  
OpenCV

 ONNX

# mxnet

 Microsoft  
CNTK

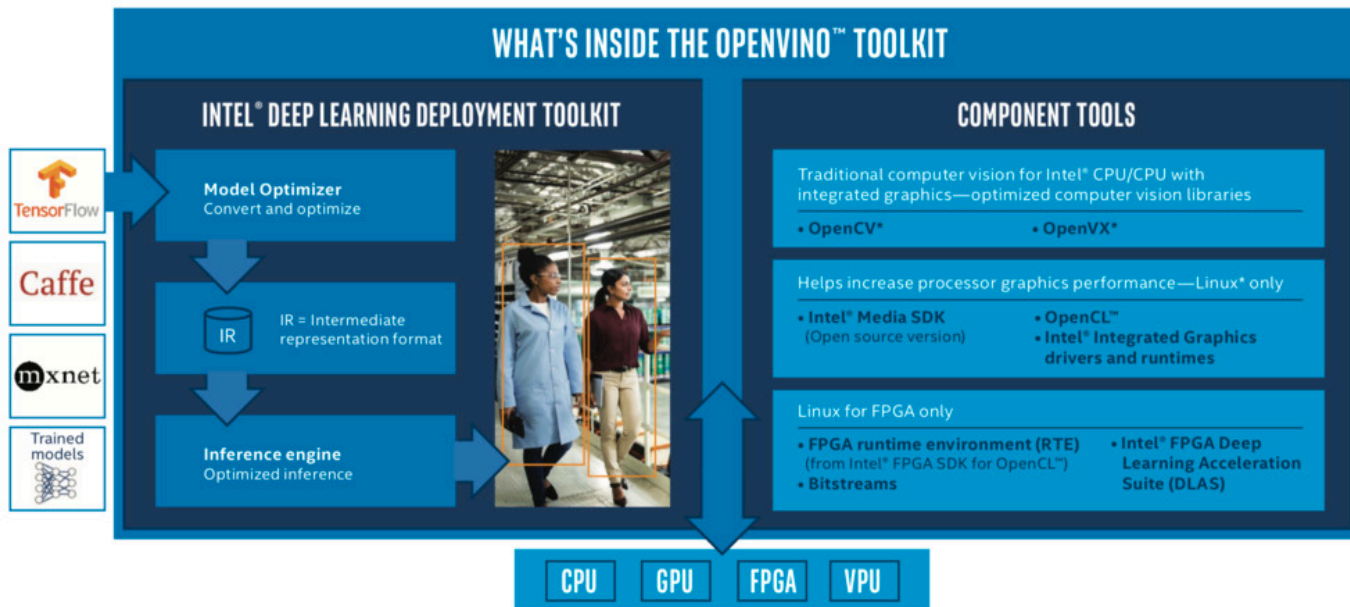
Mais infos em: <https://www.raspberrypi.org>

# Principais Plataformas

## Intel Movidius – Neural Compute Stick



THE  
DEVELOPER'S  
CONFERENCE



Mais infos em: <https://software.intel.com/en-us/articles/transitioning-from-intel-movidius-neural-compute-sdk-to-openvino-toolkit>

# Principais Plataformas

## Intel Movidius - Inferência



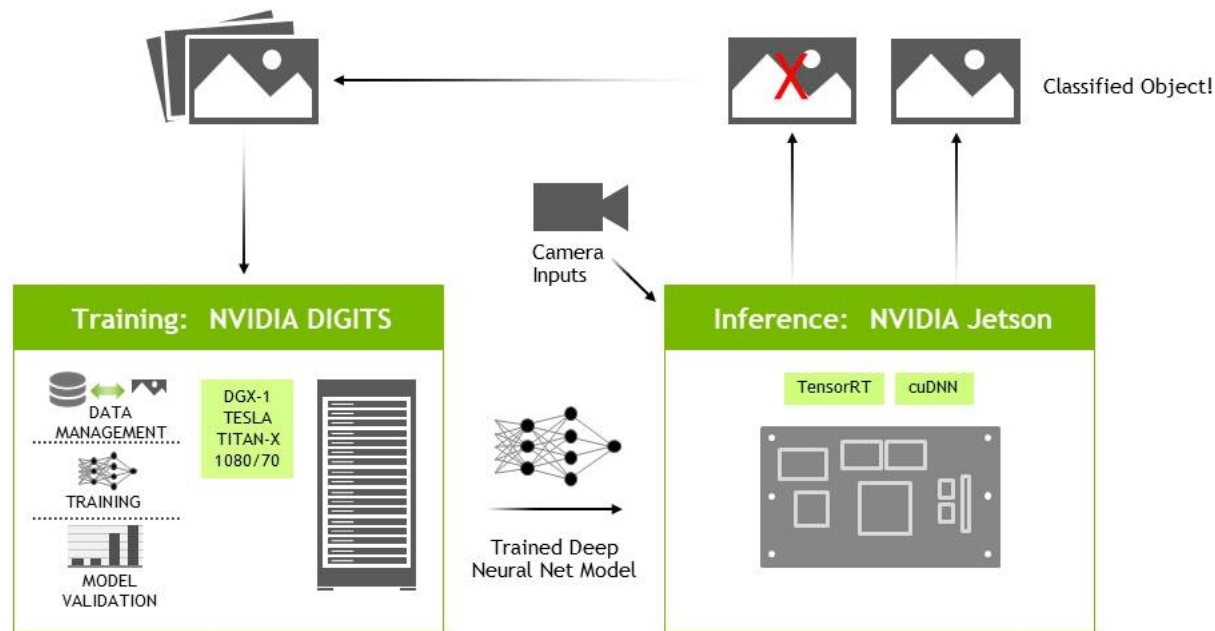
THE  
DEVELOPER'S  
CONFERENCE





# Principais Plataformas

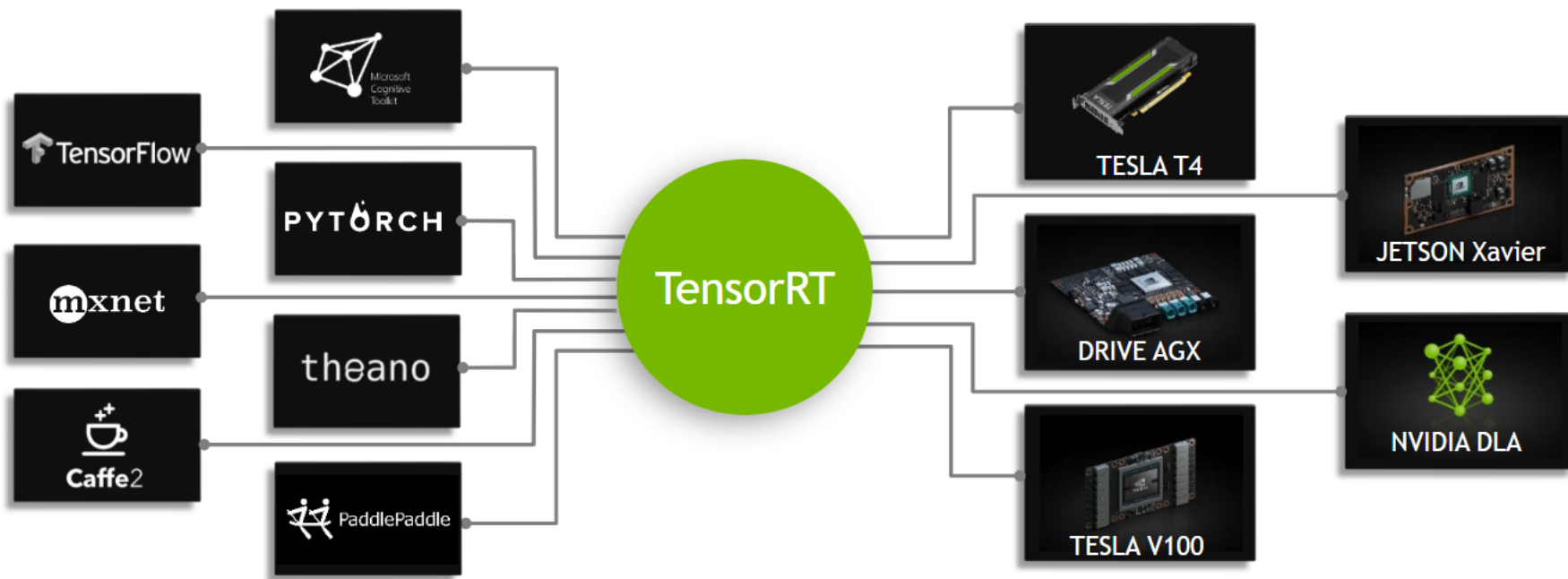
## NVIDIA - Jetson Nano (DIGITS)



Mais infos em: <https://github.com/dusty-nv/jetson-inference/blob/master/docs/digits-workflow.md>

# Principais Plataformas

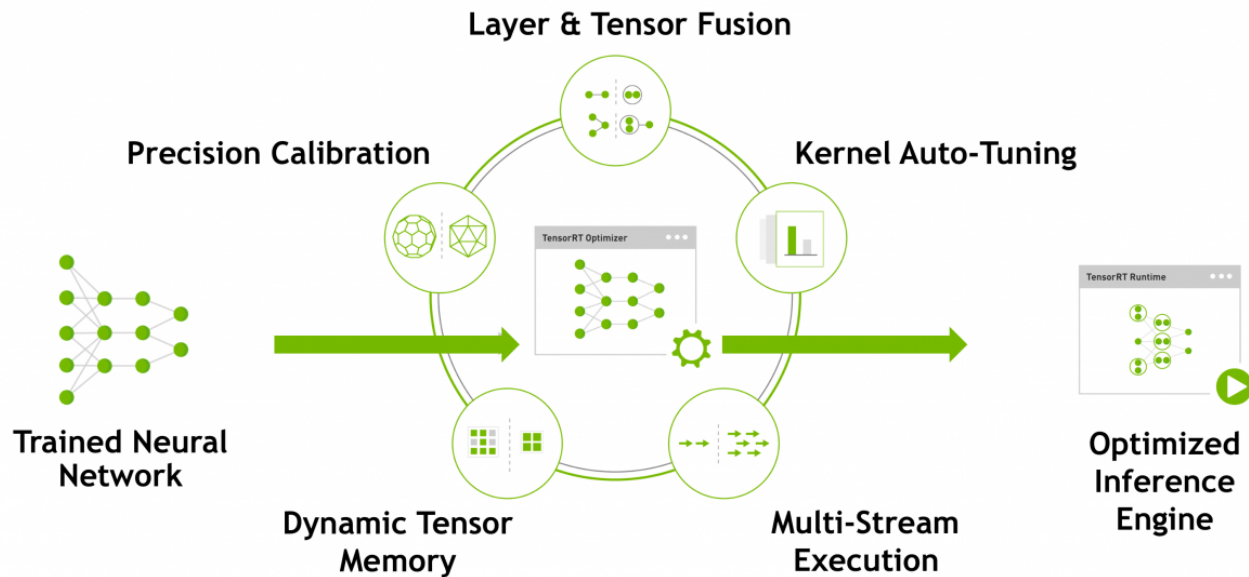
## NVIDIA - Jetson Nano (TensorRT)



Mais infos em: <https://docs.nvidia.com/deeplearning/sdk/tensorrt-developer-guide/index.html>

# Principais Plataformas

## Jetson Nano – Inferência (TensorRT)



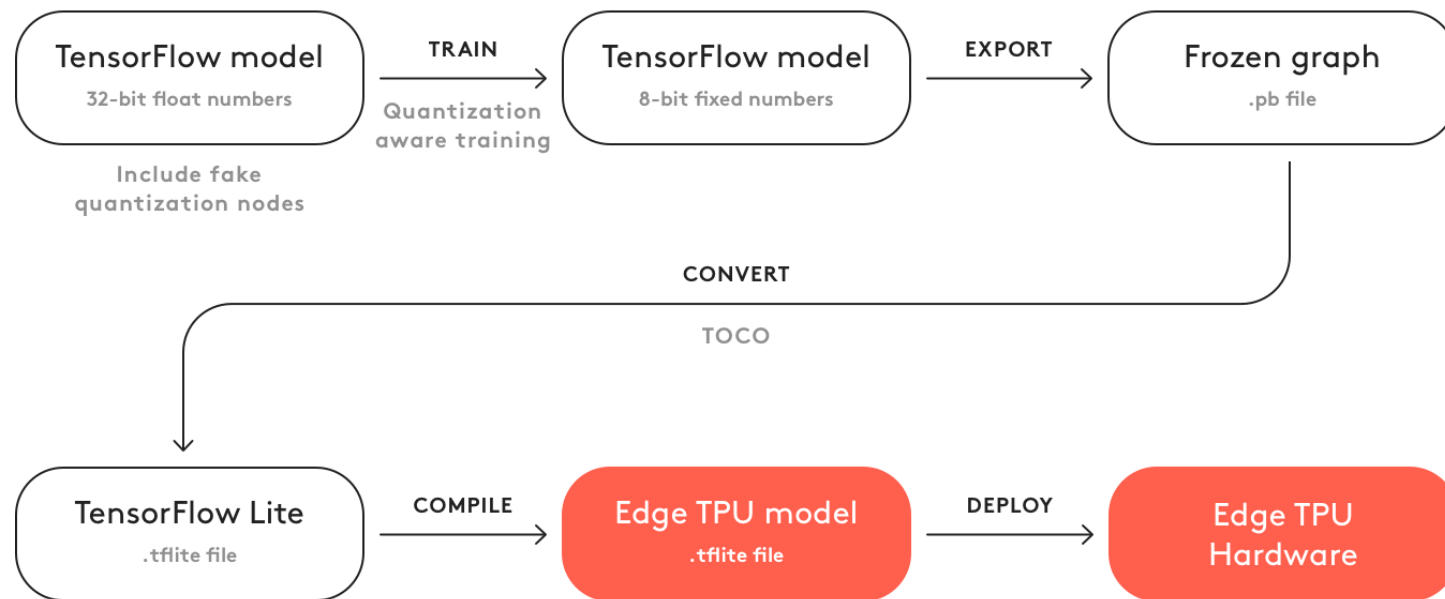
Mais infos em: <https://docs.nvidia.com/deeplearning/sdk/tensorrt-developer-guide/index.html>

# Principais Plataformas

## Google Coral



THE  
DEVELOPER'S  
CONFERENCE



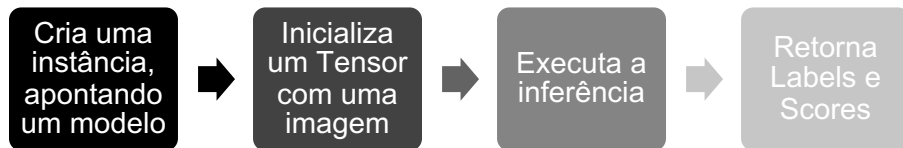
Mais infos em: <https://coral.withgoogle.com/docs/edgetpu/models-intro/>

# Principais Plataformas

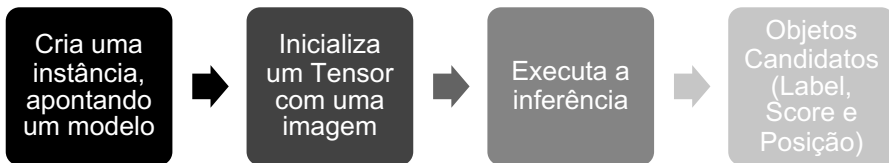
## Google Coral - Inferência



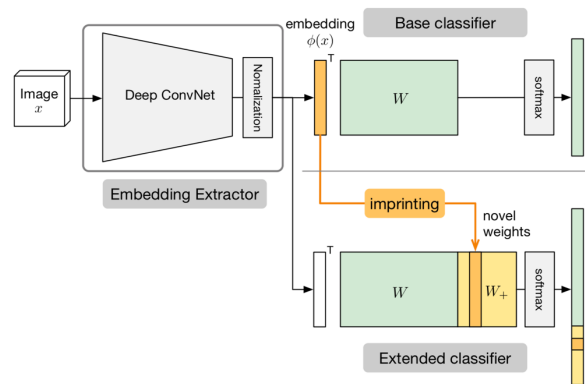
### Classification Engine



### Detection Engine



### Imprinting Engine (Transfer Learning)



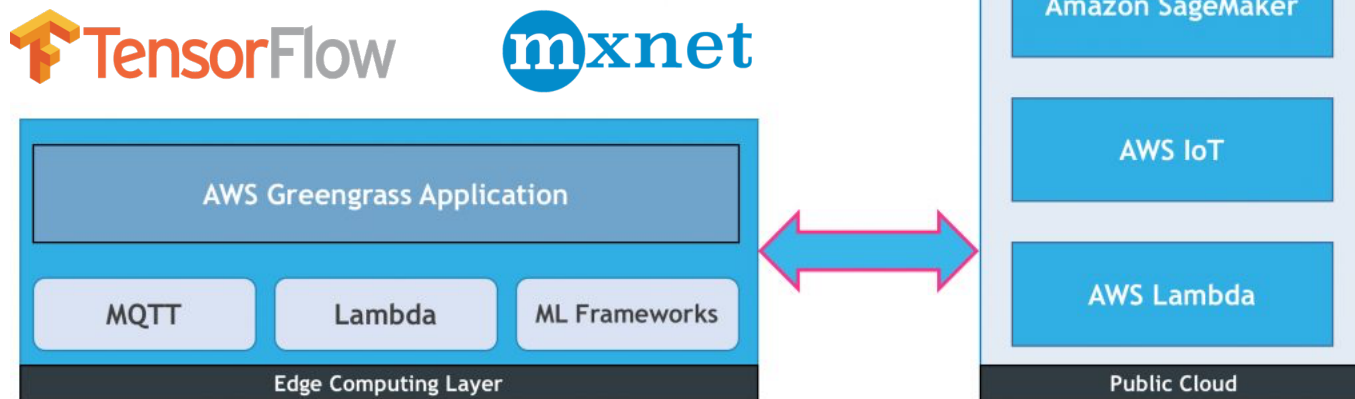
### Low-Shot Learning with Imprinted Weights

<https://arxiv.org/pdf/1712.07136.pdf>

Mais infos em: <https://coral.withgoogle.com/docs/edgetpu/api-intro/>

# Principais Plataformas

## AWS Deeplens



Mais infos em: <https://aws.amazon.com/pt/deeplens/>

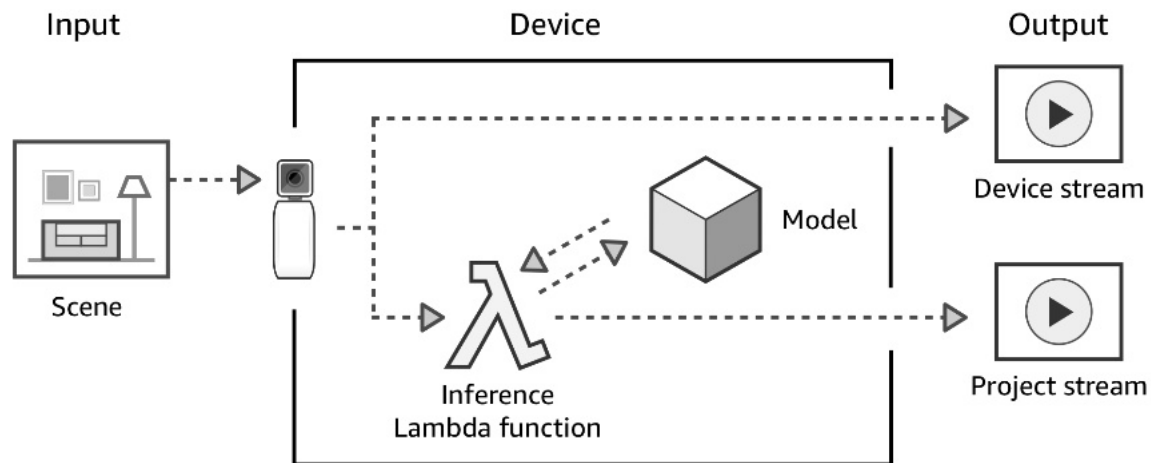


# Principais Plataformas

## AWS DeepLens - Inferência



THE  
DEVELOPER'S  
CONFERENCE



Mais infos em: <https://aws.amazon.com/pt/deeplens/>

**ENOUGH TALK**

**SHOW ME A DEMO!**



# Expectativas ... e sonhos

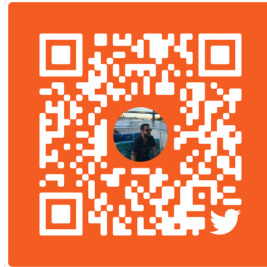


- Mais otimizadores para inferências
- Algoritmos otimizados para treinamento de redes “at the Edge”
- Mais plataformas e frameworks para aplicações de NLP
- Raspberry Pi com GPU que permita treinar redes

# Contatos



[in/fulviomascara](https://www.linkedin.com/in/fulviomascara)



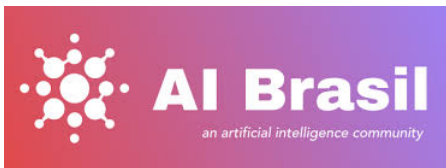
[@fulviomascara](https://twitter.com/fulviomascara)



[fulvio.mascara](mailto:fulvio.mascara)



[fulviomascara](https://github.com/fulviomascara)



**LinkedIn**



*meetup*



 **WhatsApp**



**facebook**

**Muito obrigado !**



**THE DEVELOPER'S  
CONFERENCE**