



# SUMMARY

# Legal Notices and Disclaimers

This presentation is for informational purposes only. INTEL MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS SUMMARY.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. Check with your system manufacturer or retailer or learn more at [intel.com](https://www.intel.com).

This sample source code is released under the [Intel Sample Source Code License Agreement](#).

Intel, the Intel logo, neon, and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.

\*Other names and brands may be claimed as the property of others.

Copyright © 2018, Intel Corporation. All rights reserved.

# AI on the PC

- **AI on PC** represents a large addressable market for computation at the edge that involves AI inferencing.
- AI on PC is about being able to exploit the compute capabilities of CPUs to deliver richer human-computer interactions.
- The rate at which data is created will increase and the competition for brokering data between PCs at the edge and cloud services will be a limitation. As such, in-place computation of inputs from cameras/sensors will take center-stage to ensure timely and secure processing of those data streams.

# Summary

- Intel has technology in Data Center, Cloud, and PC to support AI.

## 1. Hardware

Capabilities highlighted:

- HW CPU – i7 Core 8<sup>th</sup> Gen, i7-8550U (CFL)
- HW CPU – i7 Core 8<sup>th</sup> Gen i7-8565U (WKL)
- Intel Integrated GPU – HD graphics
- Camera, Movidius Neural Compute Stick
- WinML via Microsoft Windows 10 – 64 bit
- OpenVINO™ Toolkit, Intel Dist. of Python
- IDE: Visual Studio Community Ed. 2017
- TensorFlow with Tutorials, Labs

## 2. Tools

## 3. Solutions

- Intel is positioning the PC as a critical AI engine at the edge.

# TensorFlow

- **TensorFlow** is an open-source software library for dataflow programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks.
- **TensorFlow on the PC** enables some training and inference capabilities at the Edge.
- Training for models in TensorFlow can take hours, days, weeks. TensorFlow can further run inference/evaluation on the trained models – but these can take time. As such, to elicit real-time uses of the models they must be “frozen” and exported, and eventually converted for use in real-time systems such as OpenVINO.

# End-to-End TensorFlow Workflow

- End-to-End TensorFlow workflow presents in steps how all of the pets imaging data that is to be used in the final custom inception\_V1 model is curated – all the way from data aggregation, clean-up and preparation for consumption by TensorFlow.
- The next steps in the workflow is about preparing for training using the TensorFlow framework and an GoogLeNet Inception v1 topology. It initiates training and view a completed graph, and learn about the relationship between accuracy and loss.
- The final steps in the workflow perform a test evaluation of the trained graph and prepares it for exporting out of TensorFlow for eventual runtime usage by either the OpenVINO toolkit or Windows Machine Learning.

# Inference at the Edge / AI on PC

- **Inference** refers to the process of inferring things about the world by applying your model to new data. In the context of machine learning refers to the process of taking a model that's already been trained and using that trained model to make useful predictions.
- **Edge computing** devices include PCs, etc.
- **Inference at the Edge** refers to the process of pushing inference models to the edge devices and perform such computations locally, timely and independent of access to network or cloud resources.
- Form more information on Intel AI on PC, check the links below:

<https://software.intel.com/en-us/ai-academy/ai-on-pc> and

<https://devmesh.intel.com/>

# OpenVINO toolkit

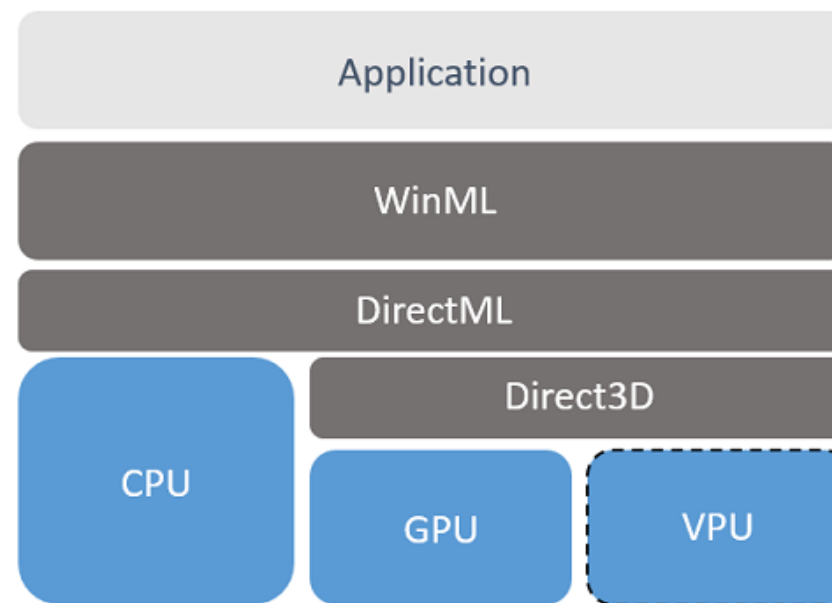
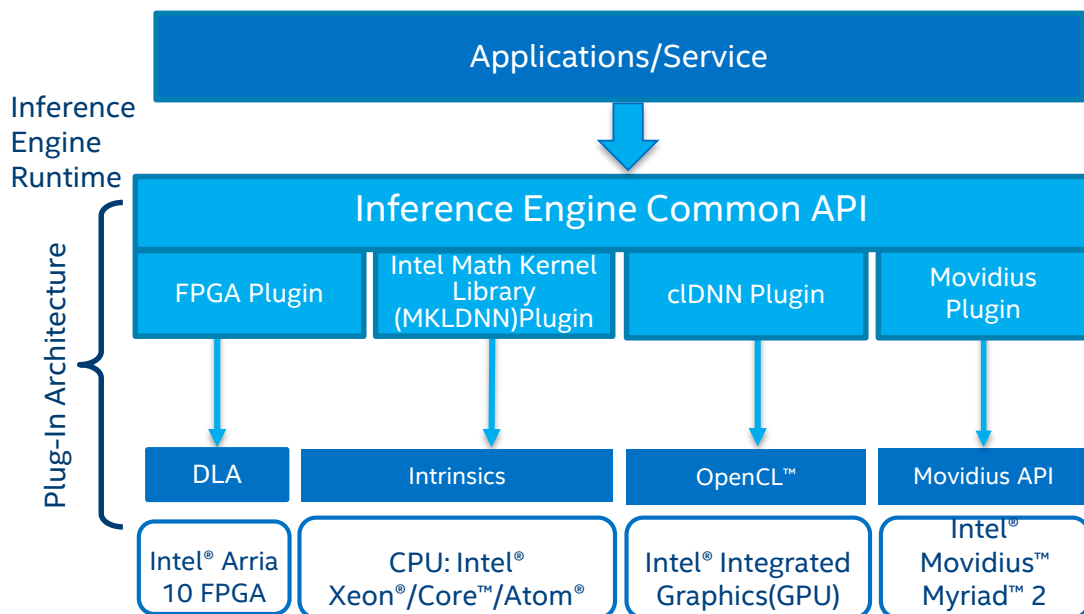
- **The OpenVINO toolkit** enables bringing to PCs advanced inference models that developers can craft into rich and compelling AI applications doing classification, object detection and object recognition workflows.
- Models developed for use with OpenVINO Model Optimizer and Inference Engine are the result and output of model training sessions such as those depicted in “*End-to-end TensorFlow Workflow*” module.



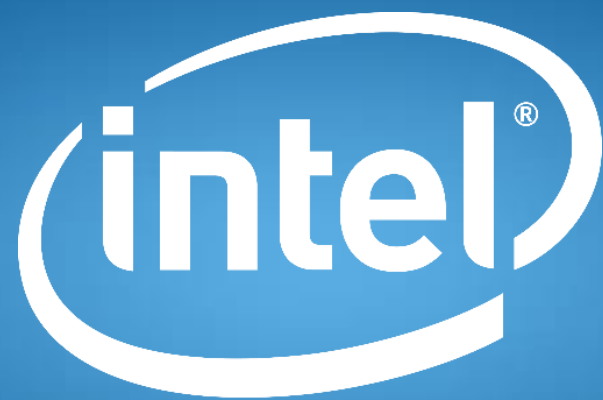
# Windows Machine Learning (WinML)

- **Windows ML** allows you to use trained machine learning models in your Windows apps.
- The trained machine learning models can be derived from frameworks such as TensorFlow, and into the ONNX model intermediate representation.
- **ONNX** is the standard model format for WinML and it can be used to automatically generate (through Visual Studio tools for AI) the interfaces that represent the inputs, outputs, and model itself for an application to bind with.

# Conclusion



- Intel has developed the 3 pillars of hardware, toolkits, and solutions to enable AI on the PC.
- With the OpenVINO toolkit, Intel is delivering a world-class experience in developing use-cases of vision AI applications on the PC.
- Intel extends AI on PC capabilities by supporting Microsoft Machine Learning solutions.



experience  
what's inside™