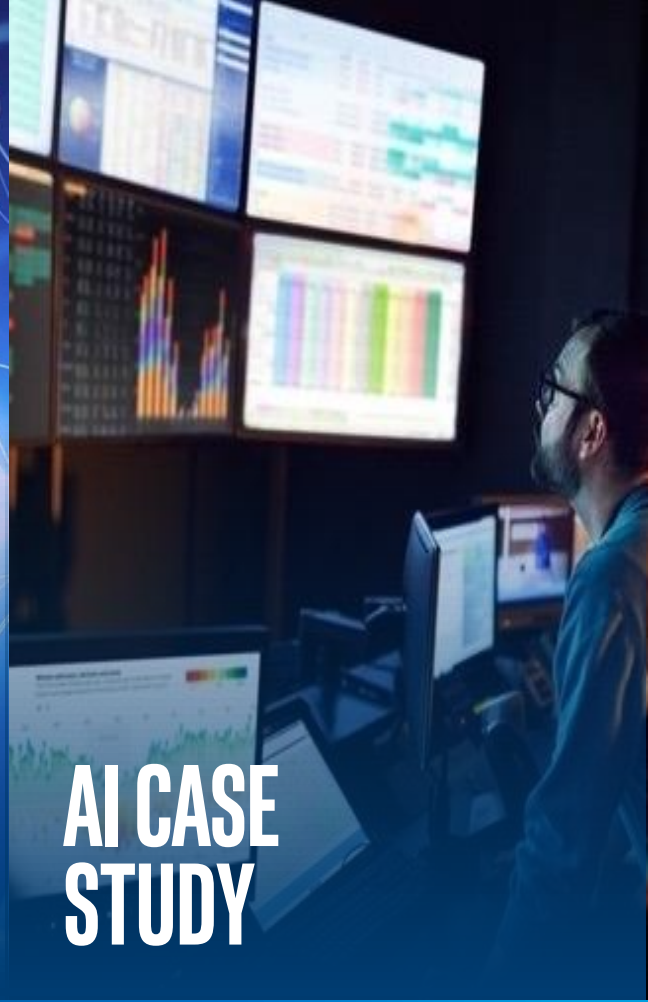
ARTIFICIAL INTELLIGENCE

BUSINESS IMPERATIVE

INTEL® AI

AI CASE STUDY



BUSINESS IMPERATIVE

INTEL® AI

AI CASE STUDY

The AI Mandate

> AI technologies are evolving fast and growing increasingly **critical** to firms' ability to win, serve, and retain customers.

FORRESTER

> ...strategic technologies for 2019 with the potential to drive significant **disruption** and deploy **opportunity** over the next five years

GARTNER

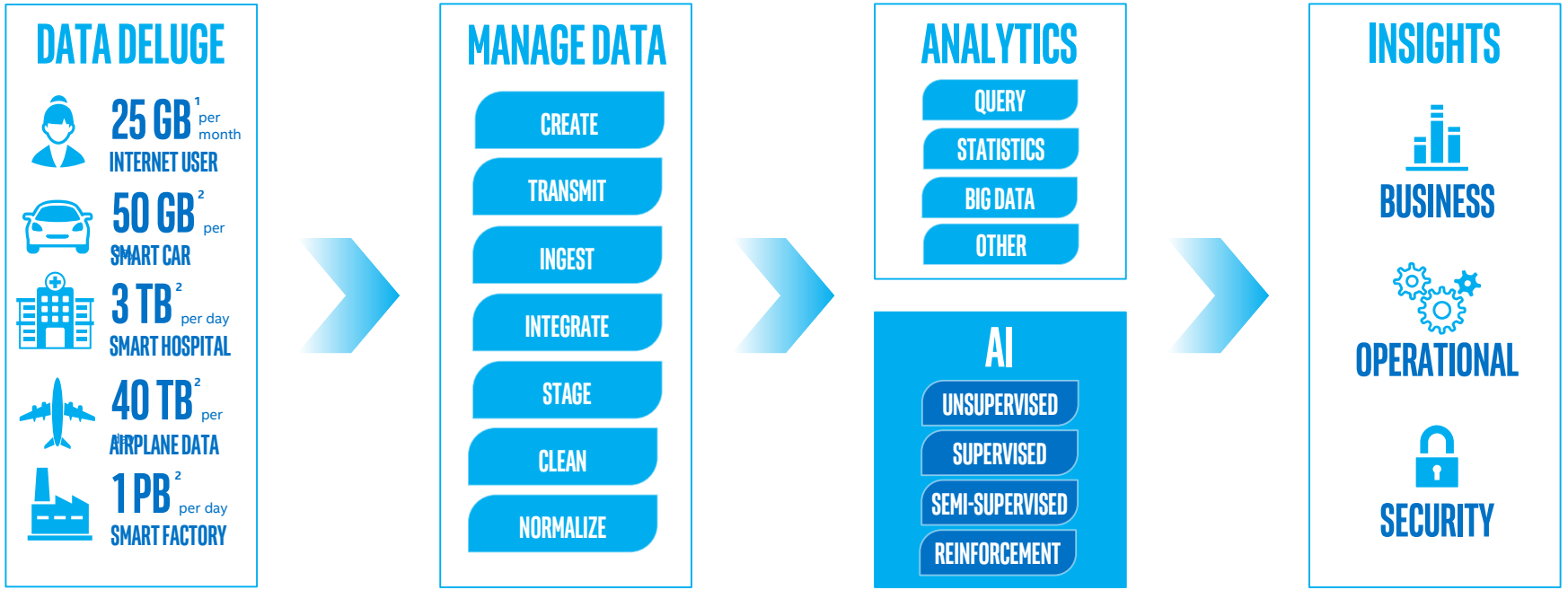
> ...**70%** of CIOs will aggressively apply data and AI to IT operations, tools, and processes by 2021.

IDC

THE TIME TO BEGIN AI ADOPTION IS NOW

Source: https://www.forrester.com/report/The+Forrester+Tech+Tide+Artificial+Intelligence+For+Business+Insights+Q3+2018/-/E-RES143252
Source: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019
Source: https://www.idc.com/getdoc.jsp?containerId=prUS44420918

Why AI?



EXTRACT VALUABLE INSIGHTS FROM DATA

1. Source: http://www.cisco.com/c/en/us/solutions/service-provider/vni-network-traffic-forecast/infographic.html
2. Source: https://www.cisco.com/c/dam/m/en_us/service-provider/ciscoknowledgenetwork/files/547_11_10-15-DocumentsCisco_GCI_Deck_2014-2019_for_CKN__10NOV2015_.pdf

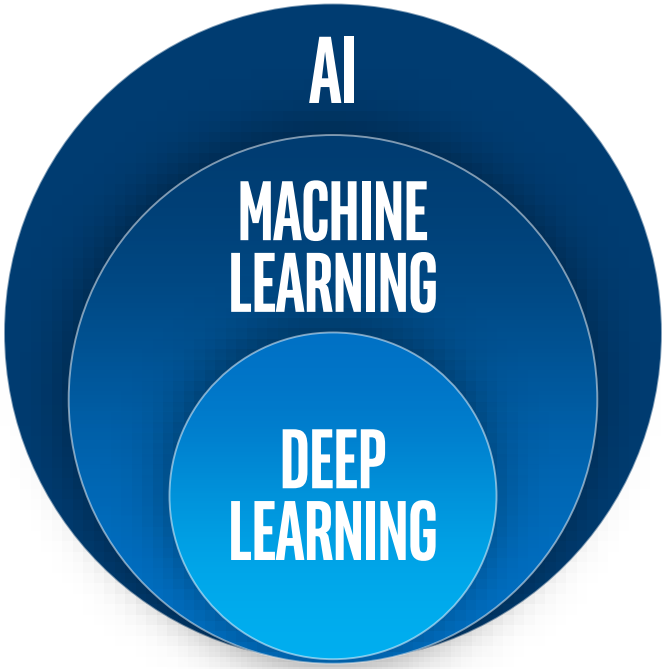
What is AI?

SUPERVISED LEARNING

UNSUPERVISED LEARNING

SEMI-SUPERVISED LEARNING

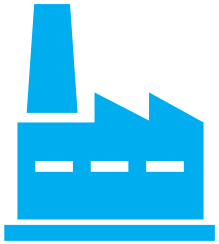
REINFORCEMENT LEARNING



- Regression
- Classification
- Clustering
- Decision Trees
- Data Generation
- Image Processing
- Speech Processing
- Natural Language Processing
- Recommender Systems
- Adversarial Networks

NO ONE-SIZE-FITS-ALL APPROACH TO AI

Which Approach is Best?

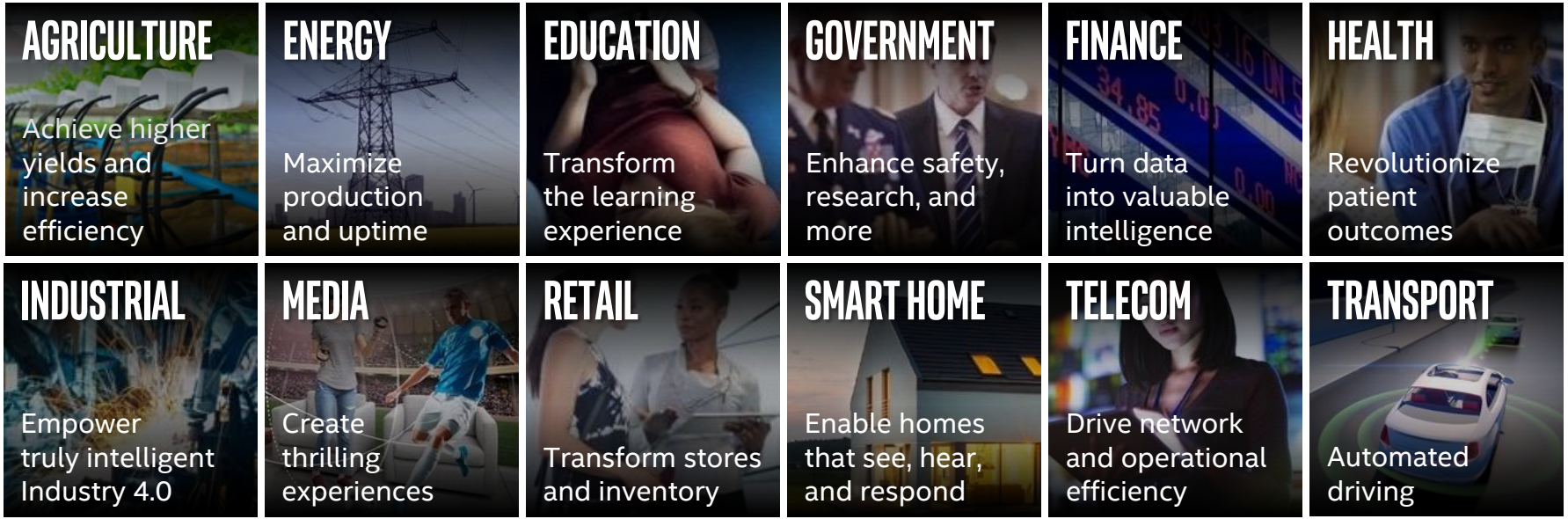


SMART FACTORY

QUESTION	METHOD	APPROACH
How many parts should we manufacture?	Historical supply and demand analysis	**Statistical Analytics**
What will our production yield be?	Algorithm learns which variables correlate to yield	**Machine Learning** ***(Unsupervised)***
Which parts have visual defects?	Algorithm learns to identify defects in images	**Deep Learning** ***(Supervised)***
Can my robotic arm learn to get better?	Algorithm that acts and adapts based on feedback	**Deep Learning** ***(Reinforcement)***

CHOOSE THE RIGHT AI APPROACH FOR YOUR CHALLENGE

AI Solutions in Every Market



OUR PARTNERS ARE DRIVING REAL-WORLD VALUE WITH INTEL® AI

Public

Solution: Ziva* Dynamics*

RESULT

INTEL® AI BUILDERS MEMBER

"The multi-core Intel® Xeon® Scalable 8164 processor allows for over 20 parallel simulations, freeing capacity while increasing overall simulation speed [to] rapidly create better, faster characters."

Client: Ziva* Dynamics*, founded by Academy Award winner and VFX pioneer James Jacobs, is changing the way computer-generated (CG) characters are made, applied, and empowered.

Challenge: Small film and game studios don't have the budget to create CG characters based on physics, anatomy and kinesiology, which makes this technology unaffordable to a wider creative audience.

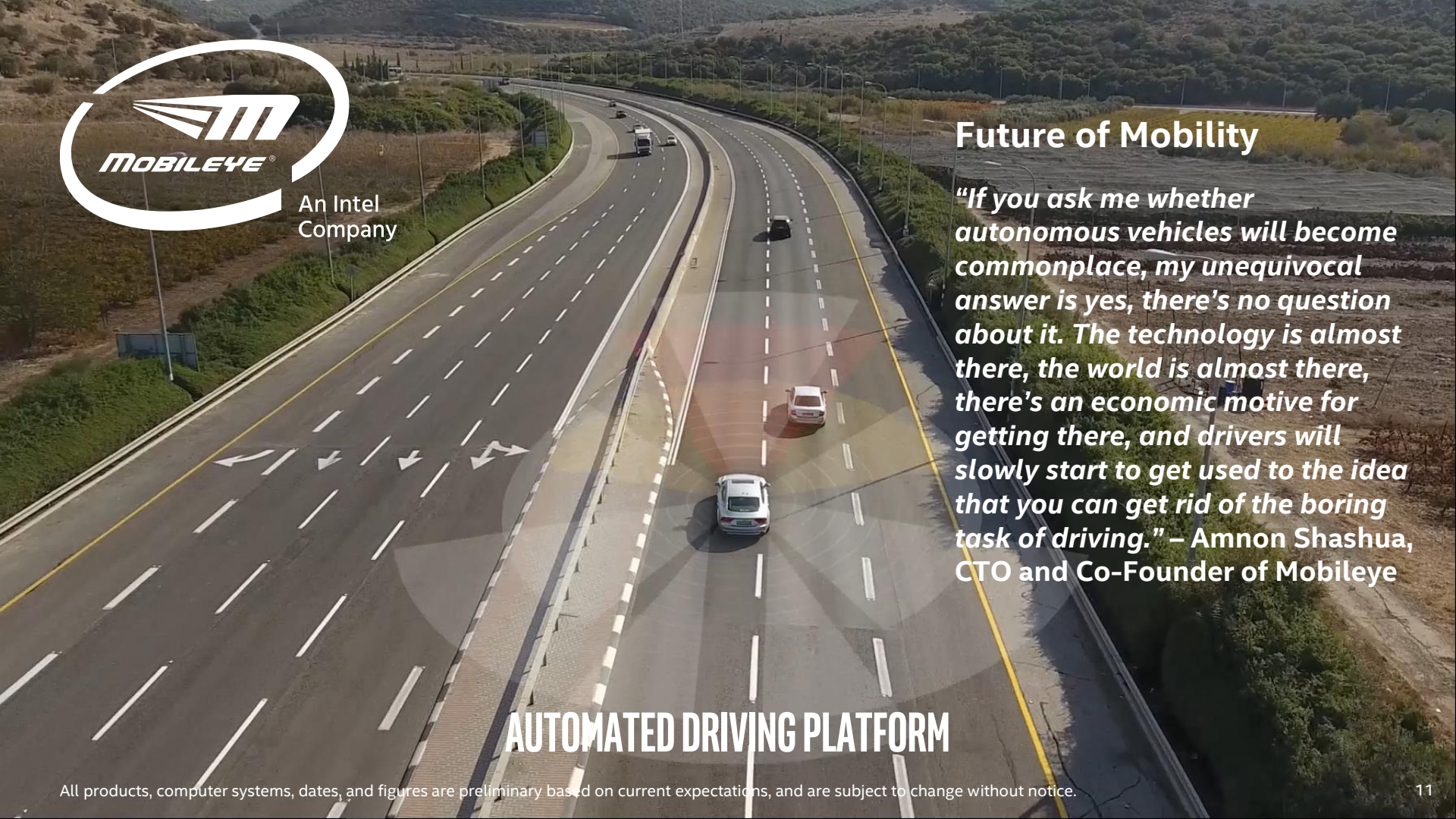
Solution: Ziva leverages standard models to create data sets of movement and mass that studios can adapt without starting from scratch. By mirroring the elements of locomotion, Ziva's ML technology enables studios, brands, and people to replicate the real world in films, games, and VR. For example, skin, muscle, fat can all be modified from base models to change a four-legged base cat model into an elephant.

Video & Solution Brief: link

https://zivadynamics.com/sundance2018

*Other names and brands may be claimed as the property of others.
Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

THIS IS AI
ENGINEERING
EXCITEMENT
ON INTEL
THE
MEG
ONLY IN THEATERS



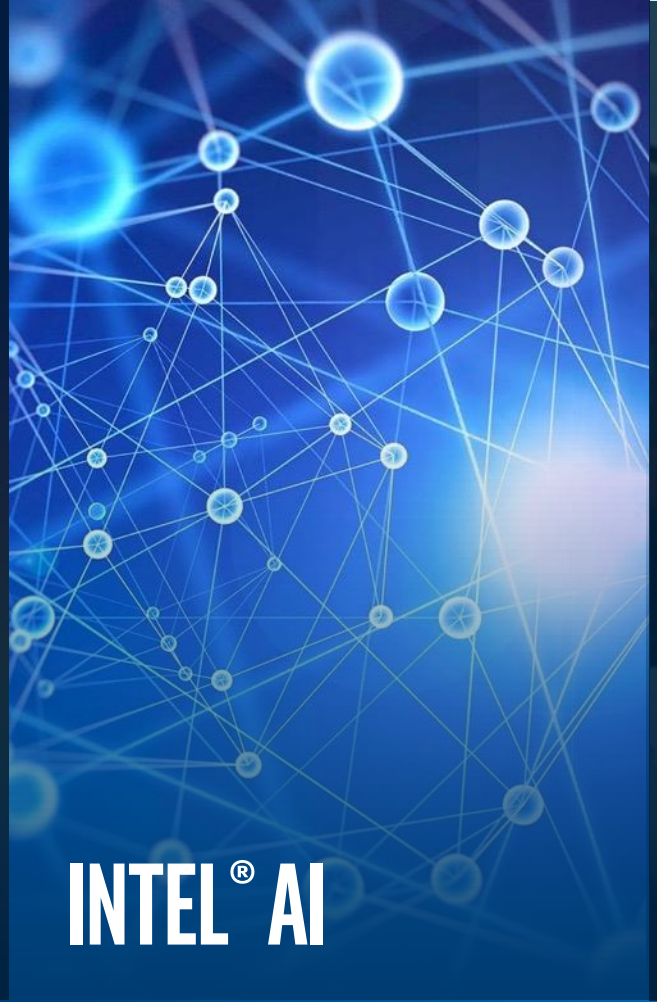
Future of Mobility

"If you ask me whether autonomous vehicles will become commonplace, my unequivocal answer is yes, there's no question about it. The technology is almost there, the world is almost there, there's an economic motive for getting there, and drivers will slowly start to get used to the idea that you can get rid of the boring task of driving." **– Amnon Shashua, CTO and Co-Founder of Mobileye**

AUTOMATED DRIVING PLATFORM

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.

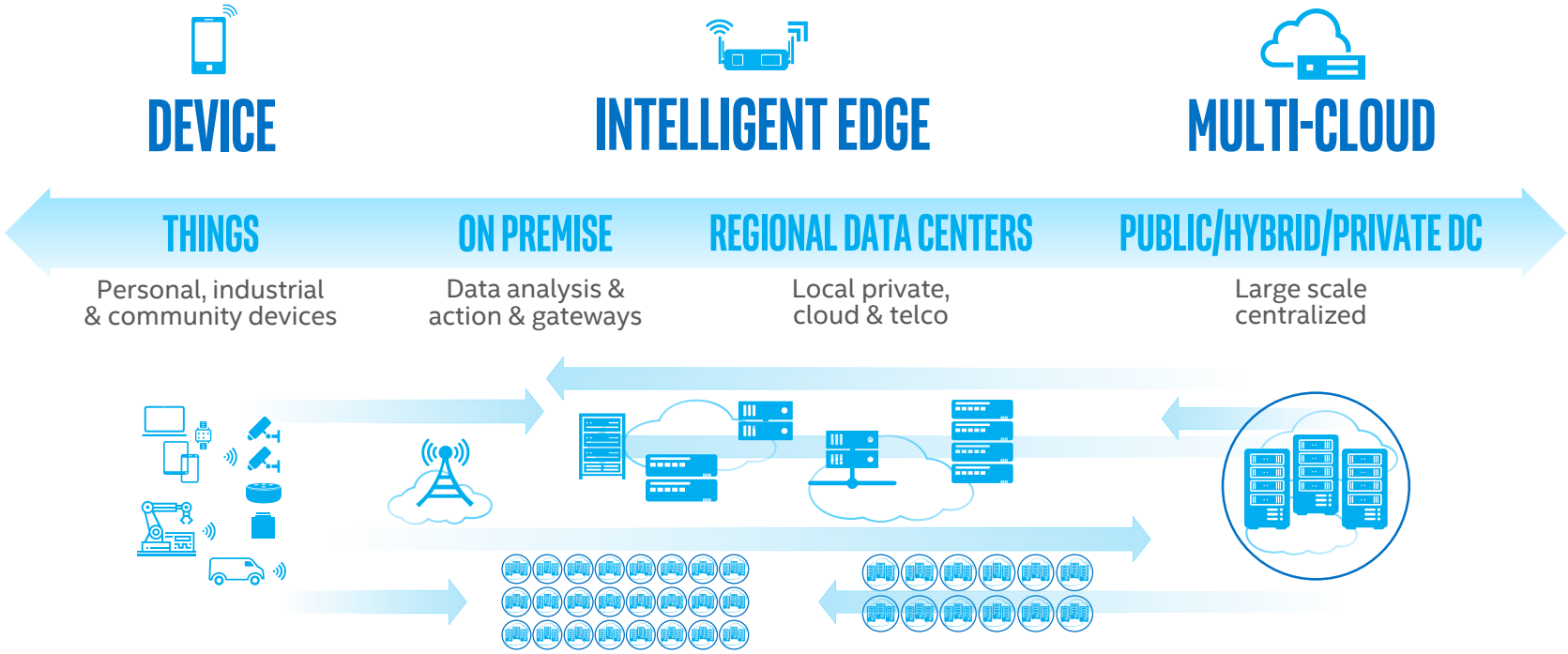
BUSINESS IMPERATIVE



INTEL® AI

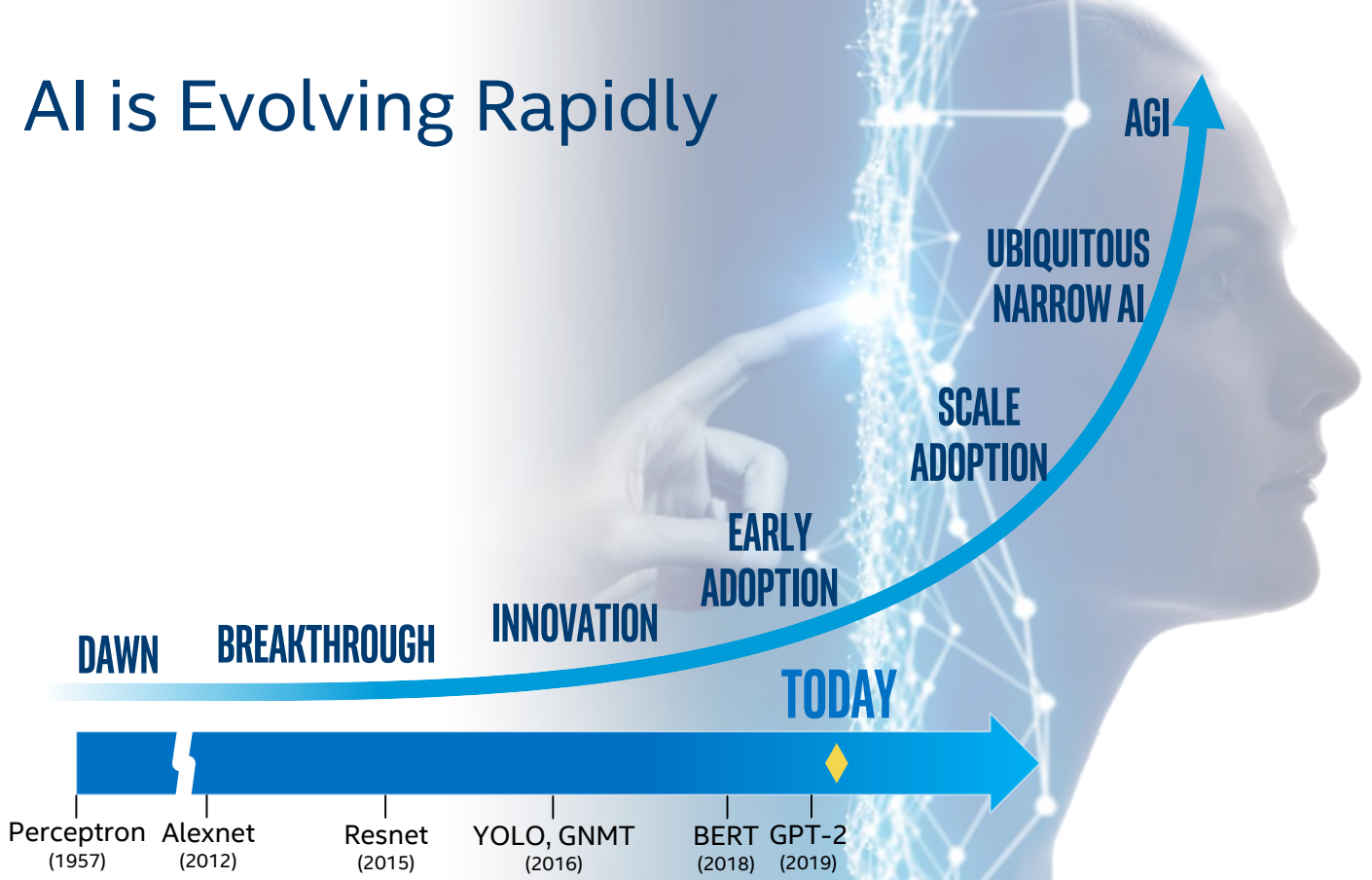
AI CASE STUDY

AI Opportunities are Diverse



All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

AI is Evolving Rapidly



MASSIVE MODELS & DATA SETS

◆

GROWING SCALE & UBIQUITY

◆

2X AI COMPUTE NEED EVERY ~3 MONTHS

◆

SYSTEM, SW & HW EVOLVING RAPIDLY

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

Intel® AI Strategy

VIBRANT COMMUNITY

Drive innovative use cases

Pioneer leading-edge AI

Fuel the ecosystem

INDUSTRY & OPEN SOFTWARE

Optimize customer software

Unify APIs across Intel

Empower developers

PLATFORM WITH BEST HARDWARE

Extend the CPU

Lead in acceleration

Build a common platform

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

Accelerate Your AI Journey with Intel

DISCOVER

Get started faster with community support

DATA

Tame the deluge with a modern data layer

MODEL

Speed up development with open AI software

DEPLOY

Deliver on the best AI hardware for your needs

COMMUNITY

CONSULT Intel® AI

PARTNER AI Builders
AI In Production

LEARN AI Developer Program

SOFTWARE

DATA MANAGEMENT Choice of 50+ Optimized Tools for Data Preparation

MACHINE LEARNING Intel® DAAL · Intel® Distribution for Python* · Analytics Zoo

DEEP LEARNING TensorFlow · BigDL · ONNX · OpenVINO · mxnet · PyTorch · Model Zoo · & More

HARDWARE

MOVE Silicon Photonics · Omni-Path Fabric · Ethernet

STORE OPTANE DC PERSISTENT MEMORY · OPTANE DC SOLID STATE DRIVE

PROCESS Xeon · FPGA · GPU · Nervana · Movidius

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

Get Started Faster

with community support

CONSULT

intel AI

Consult with your Intel and supplier representative(s) to learn more

Visit: plan.seek.intel.com/SMARTForm_ICS

PARTNER

AI BUILDERS
(Cloud to Device)

AI IN PRODUCTION
(IOT Edge/Device)

Partner with an Intel® AI provider and/or access a catalog with >100 solutions

Visit: builders.intel.com/ai
software.intel.com/ai-in-production

LEARN

AI DEVELOPER PROGRAM

Learn AI skills with the FREE¥ Intel® AI Developer Program, including cloud access

Visit: software.intel.com/ai

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

Tame the Deluge

with a modern data layer

See also:
Analytics Gold Deck

DATA MANAGEMENT

CREATE | TRANSMIT | INGEST | INTEGRATE | STAGE | CLEAN | NORMALIZE

END-TO-END: SAP*, Microsoft*, Oracle*, SAS*, Cloudera*, IBM*...

TRENDING:

C3IoT*
Thoughtspot*
Streamsets*
Confluent* (Kafka*)...

BlueData*
MemSQL*
RedisLabs*
Cassandra*
Pandas*...

Aerospike*
MarkLogic*
Splunk*
Spark*
SKLearn*...

SOLUTIONS: 30+ AAI & HPC Solutions (Genomics, ICPD, Splunk...)

CPU

ANALYTICS? Analytics Gold Deck

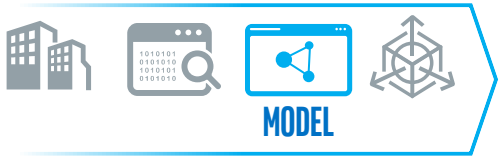
AI? See Next Slide>>

Visit: *www.intel.com/analytics*

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

Speed Up Development

with open AI software



TOOLKITS
App Developers

MACHINE LEARNING

ANALYTICS ZOO

DEEP LEARNING

MODEL ZOO

OpenVINO™

LIBRARIES
Data Scientists

Intel® Data Analytics Acceleration Library (DAAL)

Intel® Distribution for Python* (Sklearn*, Pandas*)

R (Cart, Random Forest, e1071)

Distributed (MlLib on Spark, Mahout)

Frameworks

TensorFlow, BigDL, Caffe, ONNX, mxnet, PyTorch

More framework optimizations in progress...

Intel Tools

NAUTA

RL Coach

NLP Architect

NN Distiller

KERNELS
Library Developers

Intel® Math Kernel Library (Intel® MKL)

Intel® Machine Learning Scaling Library (Intel® MLSL)

Intel® Math Kernel Library for Deep Neural Networks (Intel® MKL-DNN)

nGraph

CPU

CPU ▪ GPU ▪ FPGA ▪ ACCELERATOR

Visit: *www.intel.ai/technology*

1 An open source version is available at: 01.org/openvinotoolkit

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

Developer personas show above represent the primary user base for each row, but are not mutually-exclusive

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.

Deploy with Unprecedented

AI hardware choice

DEPLOY

MOVE FASTER

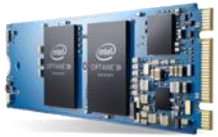
INTEL® SILICON PHOTONICS

INTEL® ETHERNET

INTEL® OMNI-PATH FABRIC

STORE MORE

INTEL® OPTANE™ SSD



INTEL® OPTANE™ DC PERSISTENT MEMORY

PROCESS EVERYTHING

CPU

GPU, FPGA

ACCELERATORS

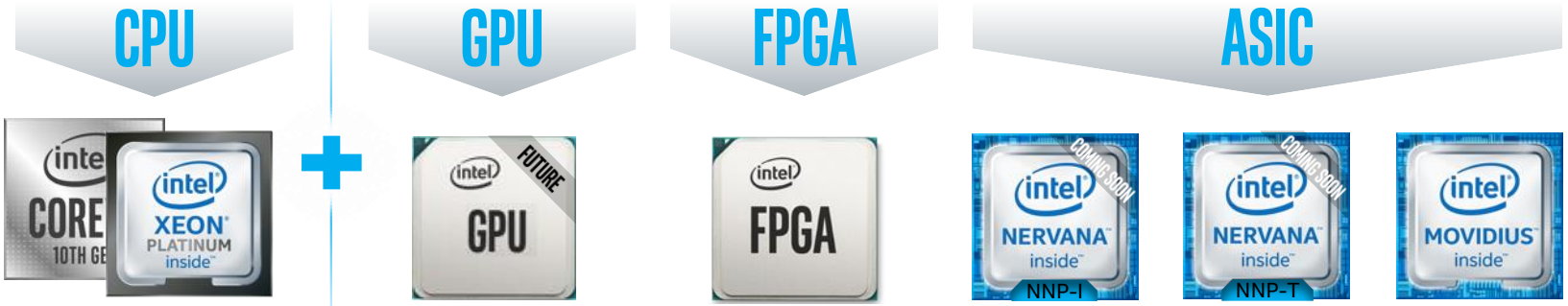
Visit: www.intel.ai/technology

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others

Intel® AI Hardware

DEVICE | INTELLIGENT EDGE | MULTI-CLOUD

OPTIMIZED FRAMEWORKS & SOFTWARE



Visit:

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
1Unified software stack development in progress DL=Deep Learning

Intel® AI Use Cases

CPU

Intel® Xeon® Scalable Processors

MULTI-CLOUD

JD.com*	HYBRID ANALYTICS + AI	Fast time-to-solution on Spark* with MlLib & BigDL
CERN*	HPC AND AI	Fast time-to-solution for deep learning in classic workflows
Novartis*	LARGE DL TRAINING	Fast DL training for large image recognition in drug discovery
Taboola*	DEEP LEARNING INFERENCE	High throughput real-time recommendation (billion items)
Ziva*	MACHINE LEARNING	Animating movie creatures using machine learning techniques

CPU

Intel® Xeon® Scalable Processors

INTELLIGENT EDGE

GE Health*	DEEP LEARNING INFERENCE	Low TCO for image recognition in CT scanner for radiology

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others.

Intel® AI Use Cases

FPGA

Intel® FPGA

MULTI-CLOUD	Microsoft*	**REAL-TIME REC. ENGINE** Real-time recommendations and more workload acceleration
	Manjeera*	**REAL-TIME TRANSCRIPTS** Real-time transcription acceleration
	JD.com*	**TEXT RECOGNITION** Faster time-to-market for custom CNN & LSTM for end-to-end text recognition
INTELLIGENT EDGE	QNAP*	**VISION INFERENCE** Faster time-to-market for custom CNN workload with OpenVINO™ toolkit
	NEC*	**FACE RECOGNITION** Faster time-to-market for custom CNN workload for surveillance and retail
	Alibaba*	**REAL-TIME VISION** Real-time video encoding and decoding for smart city project

ASIC

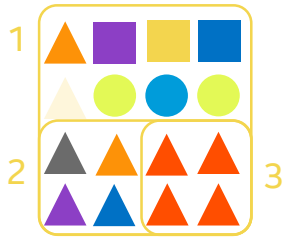
Intel® Movidius™ Myriad™ X VPU

INTELLIGENT EDGE	HPE*	**VISION AT THE EDGE** Video analytics and DL inference in an edge server blade
DEVICE	Hikvision*	**VISION IN THE DEVICE** Deep learning-based computer vision at low power

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others.

AI Compute Considerations

WORKLOADS

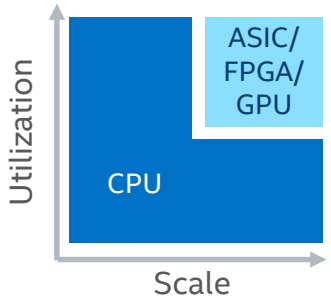


What is my workload profile?

REQUIREMENTS

What are my use case requirements?

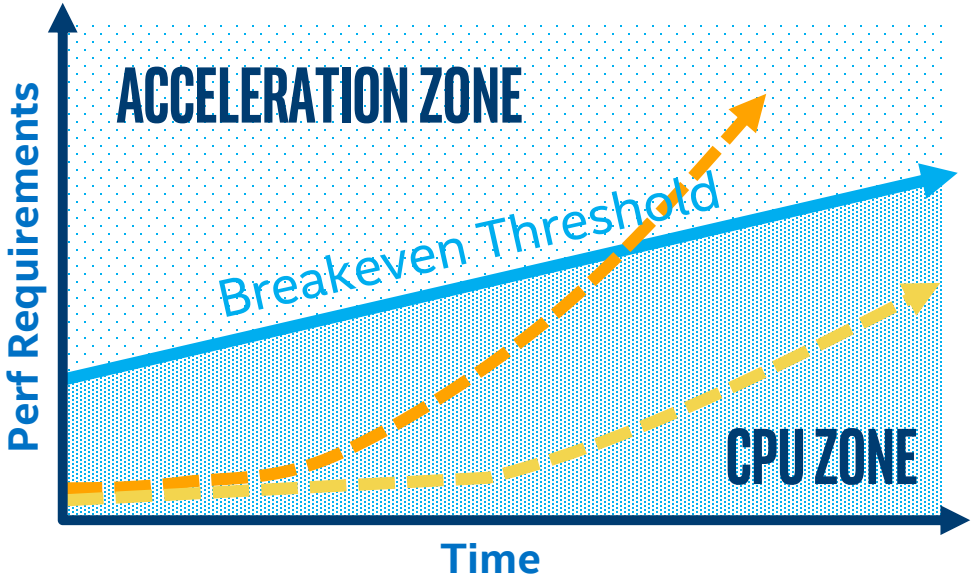
DEMAND



How prevalent is AI in my environment?

Note: word cloud source is www.wordart.com
¥Free = available to download/access at no cost to qualified developers who are enrolled in the program
*Other names and brands may be claimed as the property of others.

Bust the Deep Learning Myth



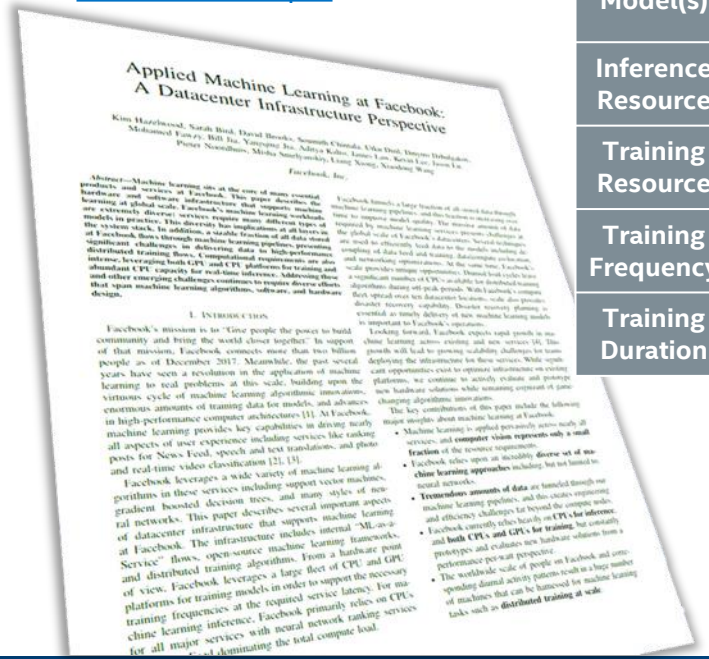
"A GPU is required for deep learning..." **FALSE**

- Most enterprises (---) use CPU for machine and deep learning needs
- Some early adopters (---) may reach a deep learning tipping point when acceleration is needed[1]

1"Most" of enterprise customers based on survey of Intel direct engagements and internal market segment analysis

Deep Learning Use Case

Source Paper:
research.fb.com/wpcontent/uploads/2017/12/hpca-2018-facebook.pdf


Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective

Services	Ranking Algorithm	Photo Tagging	Photo Text Generation	Search	Language Translation	Spam Flagging	Speech
Model(s)	MLP	SVM,CNN	CNN	MLP	RNN	GBDT	RNN
Inference Resource	CPU	CPU	CPU	CPU	CPU	CPU	CPU
Training Resource	CPU	GPU & CPU	GPU	Depends	GPU	CPU	GPU
Training Frequency	Daily	Every N photos	Multi-Monthly	Hourly	Weekly	Sub-Daily	Weekly
Training Duration	Many Hours	Few Seconds	Many Hours	Few Hours	Days	Few Hours	Many Hours

LARGE CLOUD USERS EMPLOY CPU EXTENSIVELY FOR DEEP LEARNING

Intel® Distribution of OpenVINO™ Toolkit

DEEP LEARNING

Caffe TensorFlow ONNX mxnet KALDI

- Model Optimizer
- Inference Engine
- Supports 100+ public models, incl. 30+ pretrained models

COMPUTER VISION

OpenCV OpenVX OpenCL™

- Computer vision library (kernel & graphic APIs)
- Optimized media encode/decode functions

SUPPORTS MAJOR AI FRAMEWORKS

Rapid adoption by developers

CROSS-PLATFORM FLEXIBILITY

Multiple products launched based on this toolkit

HIGH PERFORMANCE, HIGH EFFICIENCY

Breadth of product portfolio

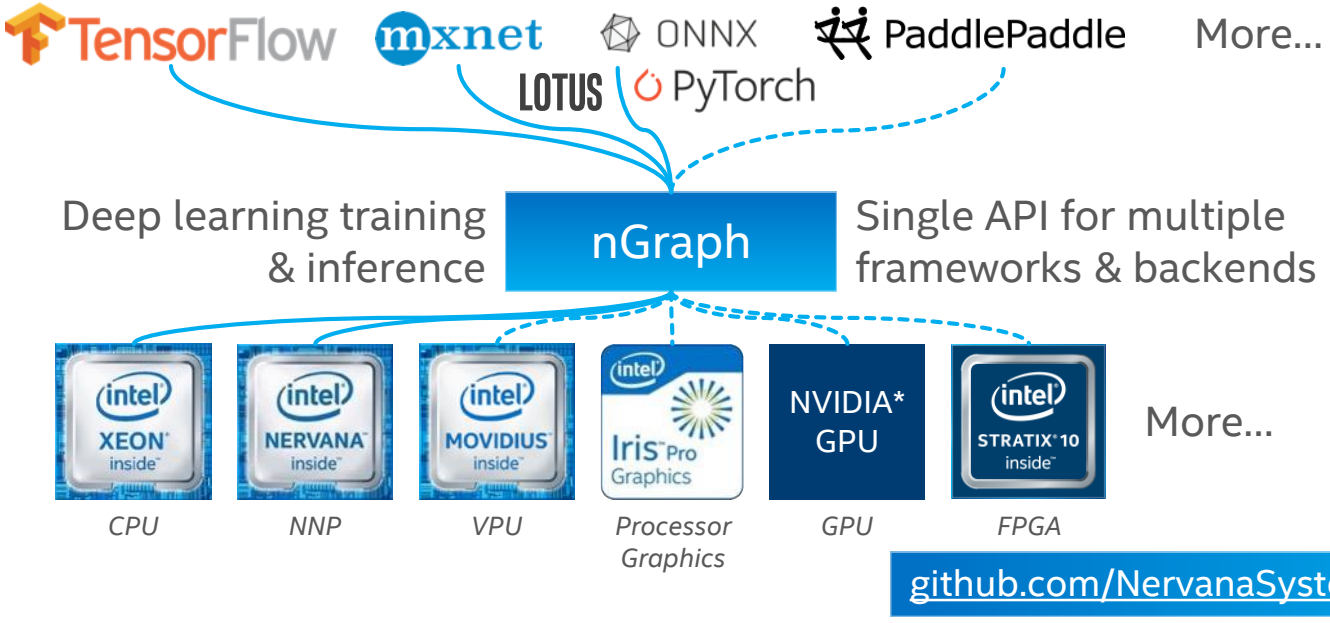
STRONG ADOPTION + RAPIDLY EXPANDING CAPABILITY

SOFTWARE.INTEL.COM/OPENVINO-TOOLKIT

Optimization Notice
Obtain open source version at 01.org/openvinotoolkit
Other names and brands may be claimed as the property of others.

Intel® nGRAPH™ Compiler

Work in progress



github.com/NervanaSystems/ngraph

OPEN-SOURCE C++ LIBRARY, COMPILER & RUNTIME FOR DEEP LEARNING ENABLING FLEXIBILITY TO RUN MODELS ACROSS A VARIETY OF FRAMEWORKS AND HARDWARE

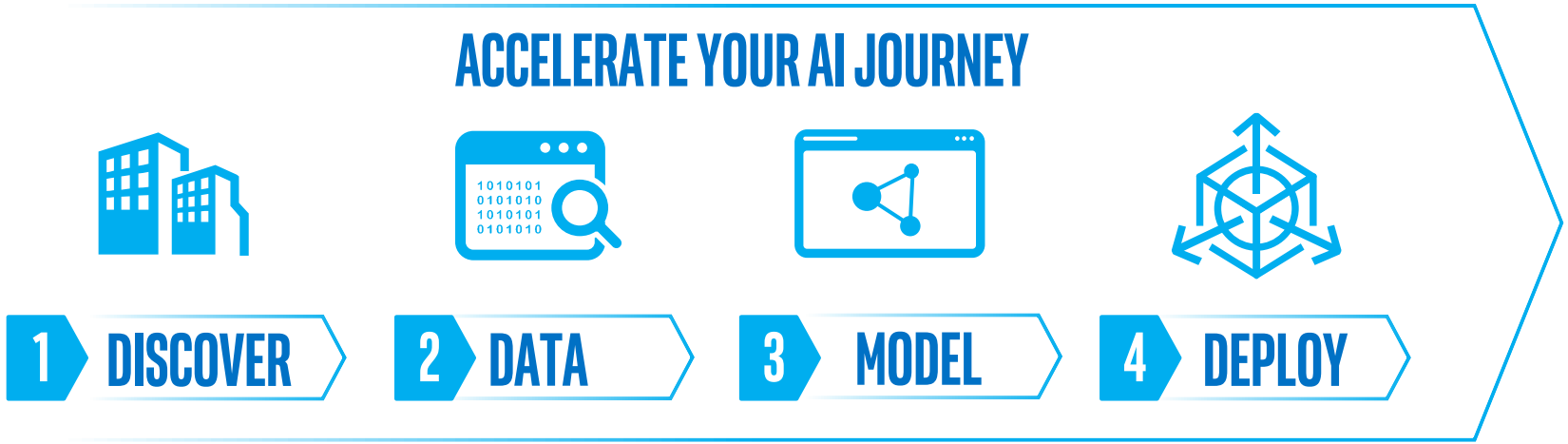
All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others
Optimization Notice

BUSINESS IMPERATIVE

INTEL® AI

AI CASE STUDY

Intel® AI Case Study



Intel® AI Case Study

DISCOVER

IDENTIFY

Identify prospects internally and using the 70+ AI solutions in Intel's portfolio; then assess business value of each one

PRIORITIZE

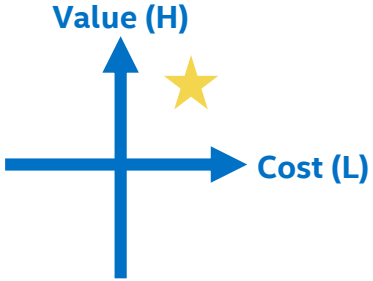
Prioritize projects based on business value and cost to solve with Intel guidance; choose industrial defect detection via DL[1]

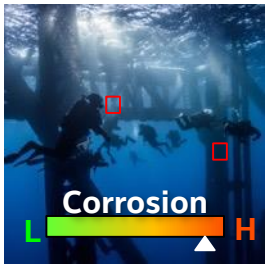
CONSIDER

Consider ethical, social, legal, security and other risks and mitigation plans with Intel advisors prior to kickoff

ORGANIZE

Organize internally to get buy-in, support new development philosophy and grow developer talent via Intel® AI





intel AI DEVELOPER PROGRAM

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others Optimization Notice

Intel® AI Case Study

DATA

011010110110
110101101011
001011010100
011010110110
110101101011
001011010100
011010110110
110101101011
001011010100
011010110110
110101101011
001011010100

INGEST

Ingest streaming data from drones using a popular software tool among the many that run on the CPU

STORE

Store data in block storage (for high-performance) in a data lake with guidance from an Intel storage partner

PREPARE

Prepare data by performing cleanup and integration using popular software tools that run on the CPU

ACT

Act on the data using one of the many popular CPU tools for data analytics and visualization

12 weeks

Source Data	Transmit Data	Ingest Data	Clean up Data	Integrate Data	Stage Data
10%	*5%*	*5%*	*60%*	*10%*	*10%*

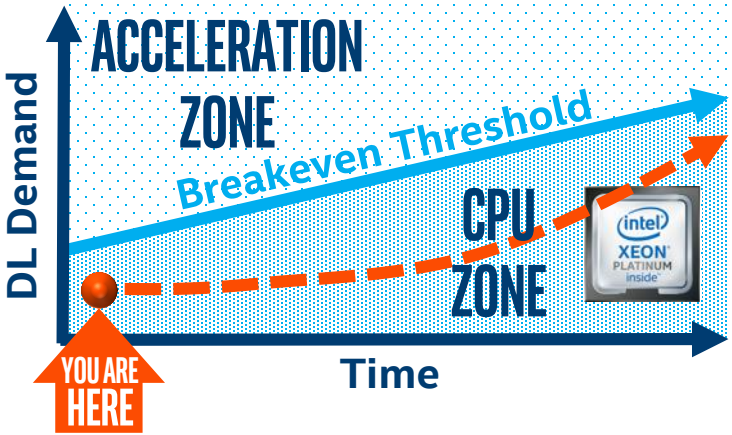
All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others Optimization Notice

Intel® AI Case Study

MODEL

SETUP

Set up compute environment; DL training (~7% of journey) acceleration NOT worthwhile due to high setup time & cost



MODEL

Model development through training a deep neural network using an Intel-optimized DL framework

TEST

Test the deep learning model using a control data set to determine if accuracy meets requirements

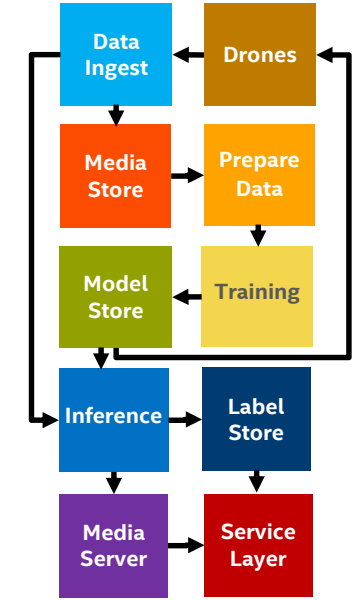
DOCUMENT

Document the code, process, and key learnings for future reference

12 weeks

Train *(Topology Experiments)*	Train *(Tune Hyper-parameters)*	Test Inference	Document Results
30%	*30%*	*20%*	*10%*

Intel® AI Case Study



DEPLOY

ARCHITECT

Architect AI deployment with Intel® AI Builders

IMPLEMENT

Implement AI in production environment

SCALE

Scale to more sites and users as demand grows

ITERATE

Iterate on the models with new data over time

REMOTE DEVICES

Drone
Drone
Drone

10 Drones

Real-time object detection and data collection

Drone
Drone
Drone

Per Drone

1x Intel® Core™ processor
1x Intel® Movidius™ VPU

MEDIA SERVER

Media Store
Media Store
Media Store

110 Nodes

8 TB/day per camera

10 cameras

3x replication

1-year retention

4 mgmt nodes

Media Store
Media Store
Media Store

Per Node

1x 2S 61xx
20x 4TB SSD

MULTI-USE CLUSTER

Data Ingestion
Data Ingestion
Data Ingestion
Data Ingestion
Inference
Inference
Inference
Inference
Prepare Data
Prepare Data
Service Layer
Service Layer
Service Layer
Media Server
Media Server
Media Server

4 Nodes
One ingestion per day, one-day retention

4 Nodes
20M frames per day

2 Nodes
Infrequent op

3 Nodes
Simultaneous users

3 Nodes
10k clips stored

DATA STORE

Model Store
Model Store
Model Store
Model Store
Label Store
Label Store
Label Store
Label Store
Label Store

Per Node

1x 2S 81xx
5x 4TB SSD

4 Nodes
1-year of history

4 Nodes
Labels for 20M frames /day

ADV. ANALYTICS

Training

16 Nodes

Intermittent use
1 training/month for <10 hours

Training

Per Node

1x 2S 81xx
1x 4TB SSD

SOFTWARE

- *OpenVino™ Toolkit*
- *Intel® MKL-DNN*
- *TensorFlow**
- *Intel® Movidius™ Software Development Toolkit*

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
*Other names and brands may be claimed as the property of others Optimization Notice

THANK YOU