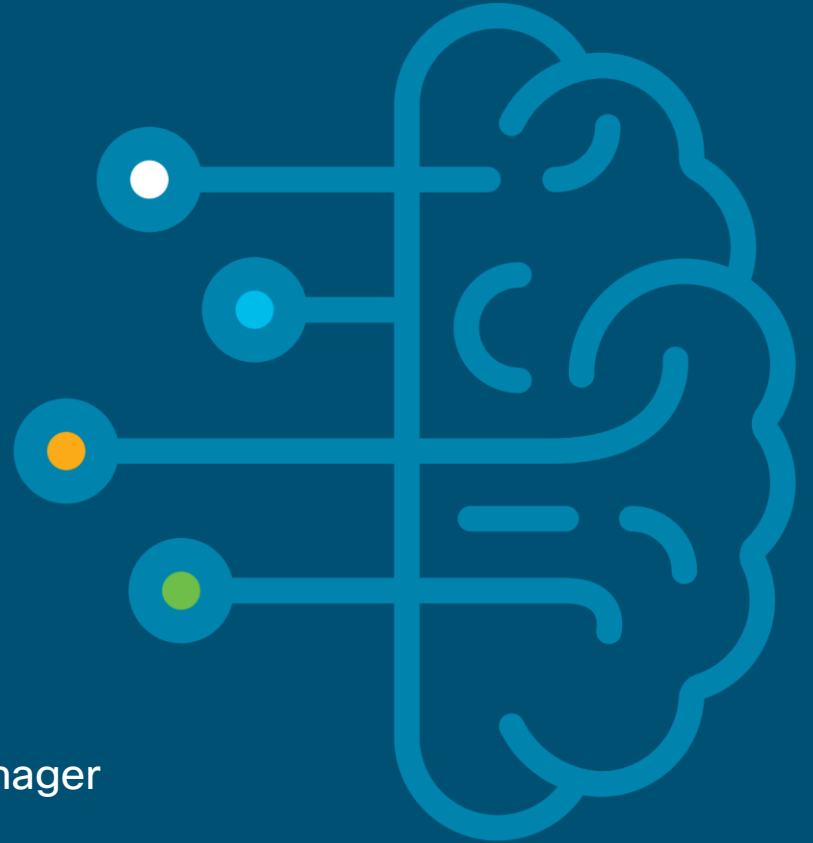




# Bridging the Gaps in AI/ML

Francoise Rees, AI/ML Solution Marketing Manager  
September, 2019



# A Top CXO Priority

**2.5 quintillion<sup>1</sup>**

bytes of data generated per  
day in 2017

**\$3.9 trillion<sup>2</sup>**

AI derived business value by  
2022

**85%<sup>3</sup>**

of CIOs will be piloting AI  
programs by 2020

Sources: 1. IBM

Sources: 2. Smarter with Gartner, "2018 Will Mark the Beginning of AI Democratization", Dec 2017, <https://www.gartner.com/smarterwithgartner/2018-will-mark-the-beginning-of-ai-democratization/>

Sources: 3. Gartner, "Forecast: The Business Value of Artificial Intelligence, Worldwide, 2017-2025," Published 12 March 2018 - ID G00348137

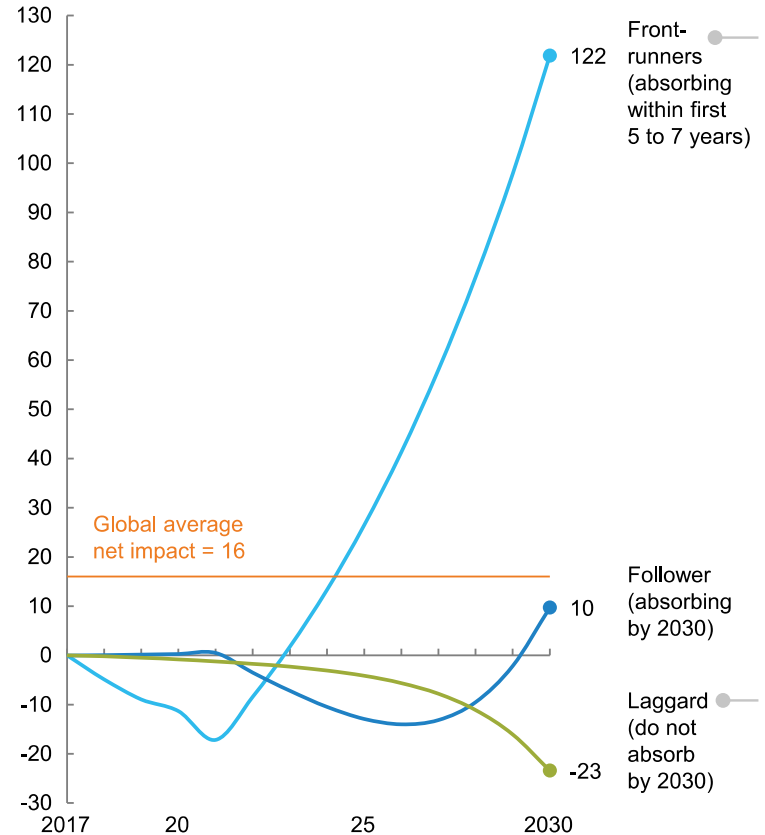
# Relative Change in Cash Flow by AI Adoption Cohort

**Speed of AI/ML Deployment is Important**

<https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy>

Relative changes in cash flow by AI adoption cohort

% change per cohort, cumulative

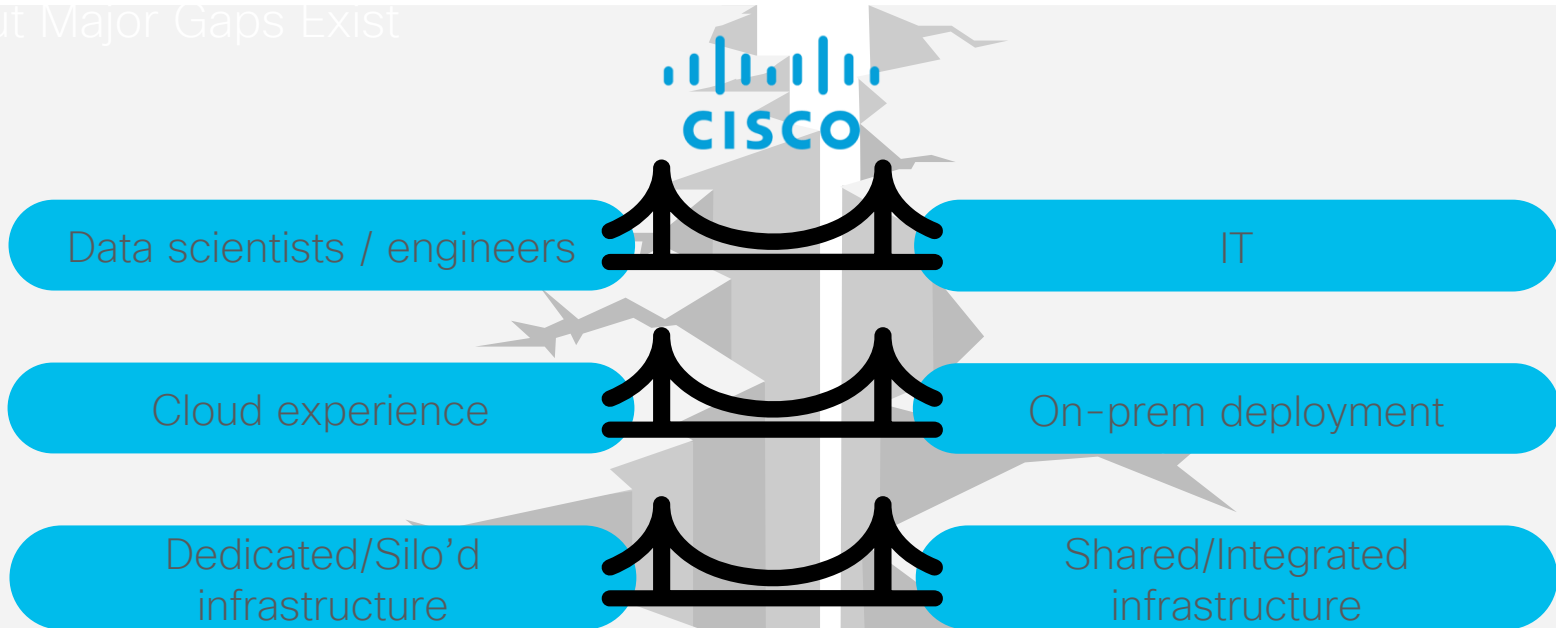


NOTE: Numbers are simulated figures to provide directional perspectives rather than forecasts.

SOURCE: McKinsey Global Institute analysis

# Operationalizing Machine Learning

But Major Gaps Exist



## Gap 1:

# Data Scientist and IT



Data  
Engineers  
Data  
Scientists

+



IT Team

- IT not familiar with AI/ML SW stack
- SW: Open source vs. Enterprise grade
- Desire for simplicity and flexibility of HW and SW deployment at scale

IT needs to be part of the design team for AI/ML from the beginning

## Cisco Solutions

NGC Containers



NVIDIA NGC-Ready



ANAconda



HX Turnkey Starter Kit  
for dev and  
deployment



Inference

UCS/HX C220  
Servers



Test/Dev

UCS C240 Servers  
HyperFlex GPU  
Nodes



Deep Learning

UCS C480 ML  
HX C480 ML



UCSD Workflow

## Gap 2:

# Consistent Cloud and On-Premise Machine Learning Experience



+



Cloud

On  
Prem

- Fast deployment
- Test-dev
- Simplicity
- APIs
- One time
- Data gravity and integration
- TCO and Performance
- Governance
- Remodel, retraining
- At scale production

## Cisco Hybrid Cloud



Data



Collect



Clean



Correlate



Train



Kubeflow

UCS



Kubeflow



Anthos

HyperFlex



Kubeflow



Anthos

Google Cloud



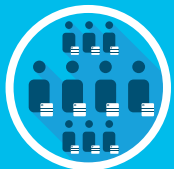
Partner  
of the Year

Google Cloud

Cisco is #2 contributor to  
Kubeflow 2.5M lines of  
code

### Gap 3:

# Dedicated vs. Shared Infrastructure



Dedicated



Shared

### Data Scientists Experience

- Dedicated workstation vs. Share ML infrastructure

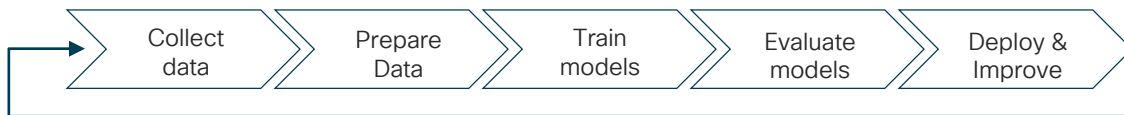
### Data Center Architecture

- Silo ML infra vs. Integrated Data Center Architecture

### Networking Architecture

- InfiniBand vs. Scalable Ethernet

## Cisco: Unified Architecture



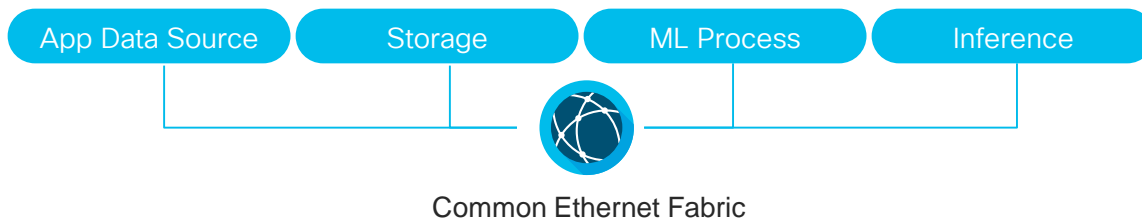
CLUDERA



FlexPod  
A Cisco and NetApp Solution

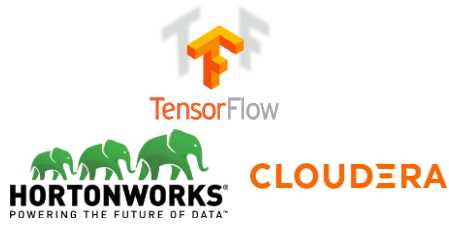
FLASH STACK CI

SwiftStack



# UCS and HyperFlex AI / ML Solutions

NGC on Cludera & Hortonworks



UCS

NGC TensorFlow on OpenShift



UCS

Kubeflow on UCS and GCP



UCS

Google Cloud

 **FlexPod**  
A Cisco and NetApp Solution



 **FLASH**  **STACK CI**

C480 ML



**NVIDIA.**



Customer Sentiment  
Inference on HyperFlex 4.0



# Activating Data with the Power of UCS

## Cisco Computing Solutions for AI/ML



### Unified Architecture

---

Adaptable Cloud-Managed  
Systems for Distributed IT



### Demystifying AI/ML Stacks

---

Validated Solutions with  
Industry Leaders



### Powering the Full AI Data Lifecycle

---

Accelerating Insight  
and Action

