



Innovating with Machine Learning, Deep Learning, and Artificial Intelligence

February 8, 2017

19th Annual Privacy & Security Conference

Celeste Fralick, Ph.D., CQA

Chief Data Scientist and Senior Principal Engineer

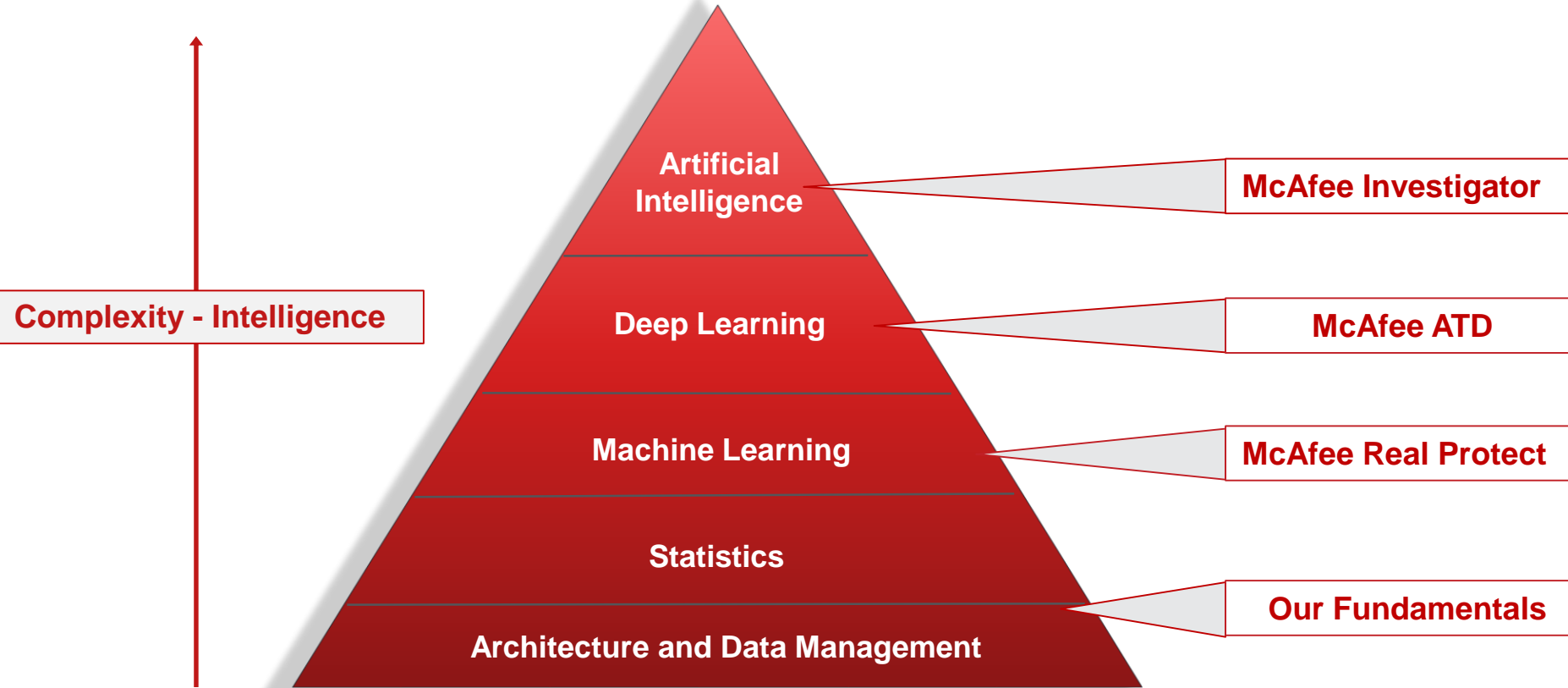
celeste_fralick@mcafee.com



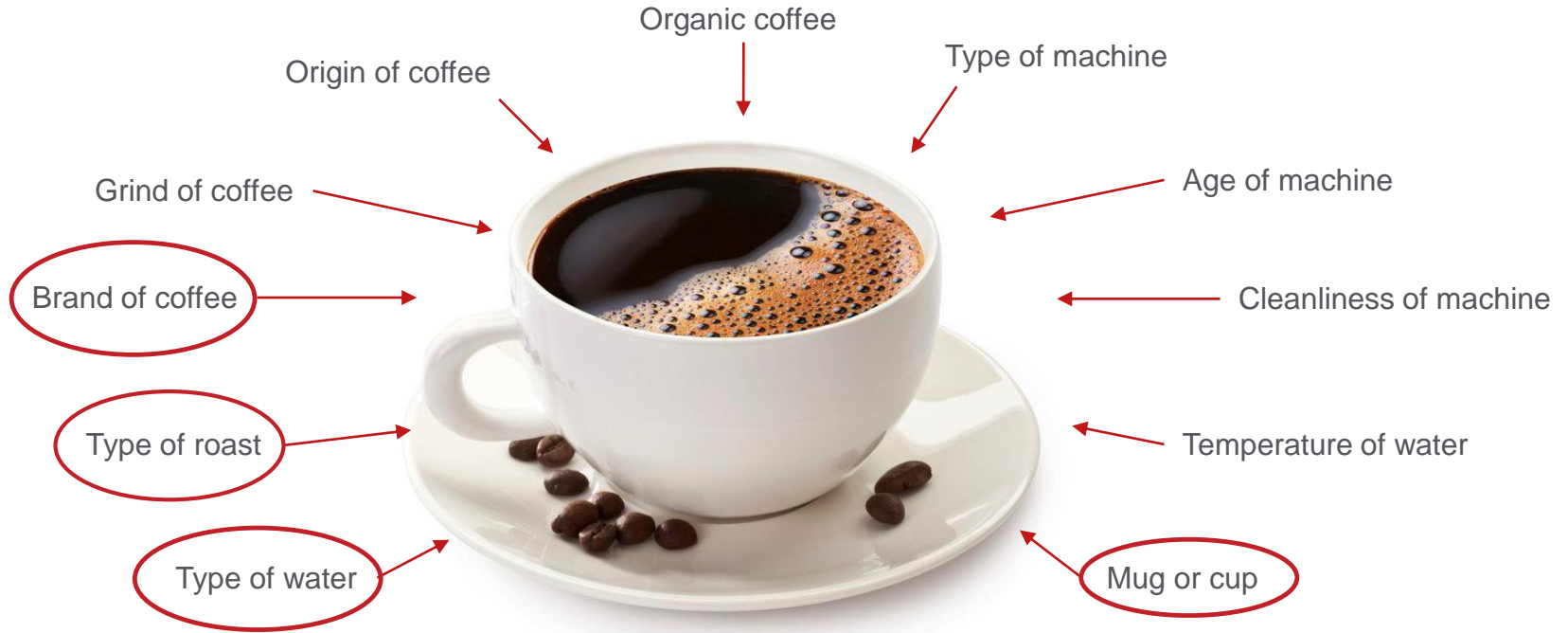
Agenda

- Distinguishing Between Machine Learning, Deep Learning, and AI
- Myths of Analytics and Machine Learning
- Growing threat: Adversarial ML
- Other Areas of Innovation
- Separating the Hype From Reality
- Applying ML, DL, and AI to McAfee Products
- Avecto and McAfee
- Final Comments

Distinguishing Between ML, DL, and AI



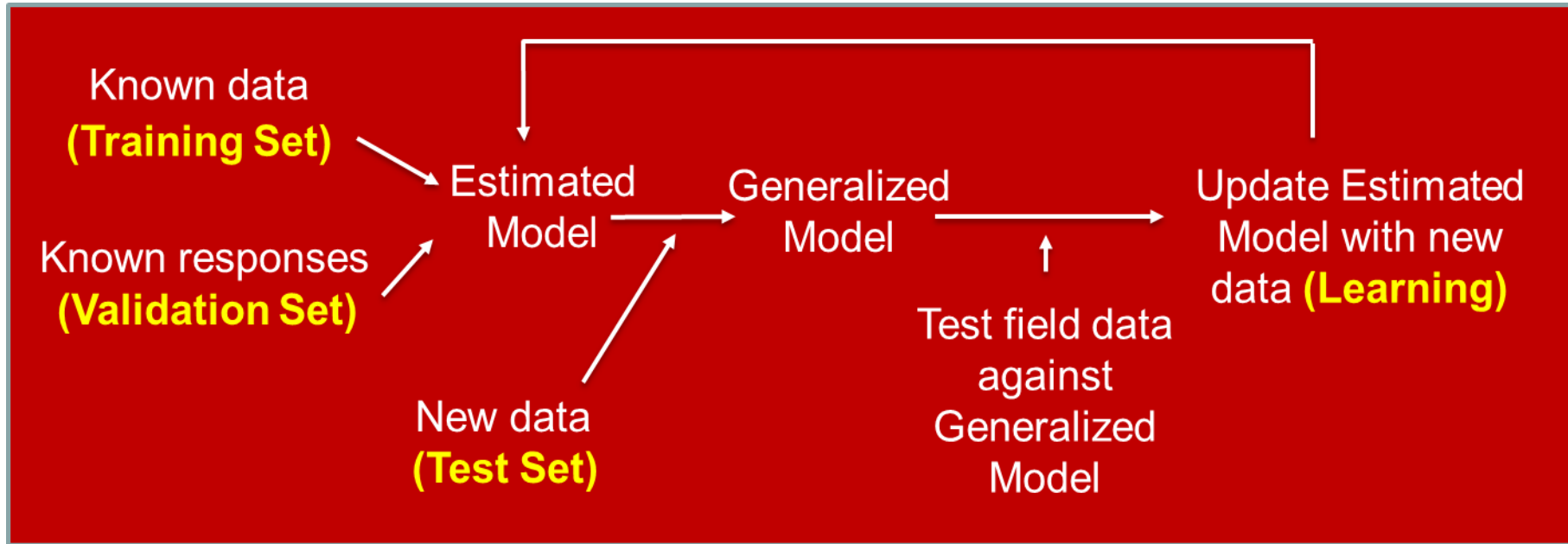
What Is A “Feature”?



A **Feature** is an individual measurable property or characteristic of something being observed.

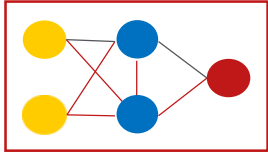
Distinguishing Between ML, DL, and AI

Machine Learning: Automated analytics that learn over time and recognizes patterns. Often applied to more complex (predictive and prescriptive) algorithms.

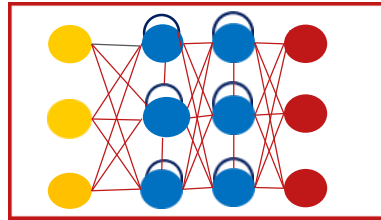


Distinguishing Between ML, DL, and AI

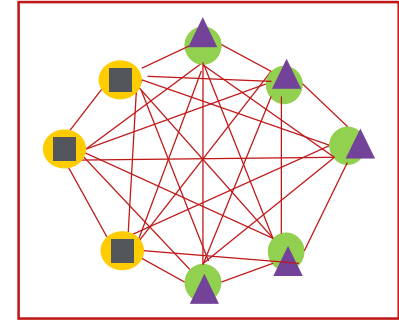
Neural Networks: Loosely based on neuronal structure of brain, uses layers with mathematical transformations and previous data to learn good vs bad data. If it is many layers, it is often referred to as “**Deep Learning**”. There are many different types of neural network algorithms, including “cyclic” (FF), “acyclic” (RNN) and many more iterations.



Feed Forward (FF)



Recurrent Neural Network (RNN)



Boltzmann Machine (BM)

Distinguishing Between ML, DL, and AI

Artificial Intelligence: Typically self-learning systems that apply an ensemble of complex algorithms to mimic human-brain processes, including decision making, reasoning, and problem solving. Cognitive Computing is an example of AI.

Natural Language
Processing &
Understanding

+

Deep Learning
Machine Learning



Value judgments

Prediction Perception

Comprehension

Interaction Sensory

Reasoning Logic

Memory Attention

Image Recognition

Consciousness of Being

Myths of Machine Learning

ML is devoid of human intervention!

- **WRONG!** Humans must still prepare, clean, model and assess data sets long term

ML can produce results from any data in any situation!

- **WRONG!** Unstructured data is notoriously challenging and can lead to inaccuracies

ML is scalable in all cases!

- **WRONG!** Some ML algorithms are better suited for larger data sets.

ML is plug-n-play!

- **WRONG!** There are many ML algorithms to train and each model must be validated. Selecting the right data set & model takes insight and time

ML is always predictive!

- **WRONG!** There are ML algorithms that classify only and do not predict.

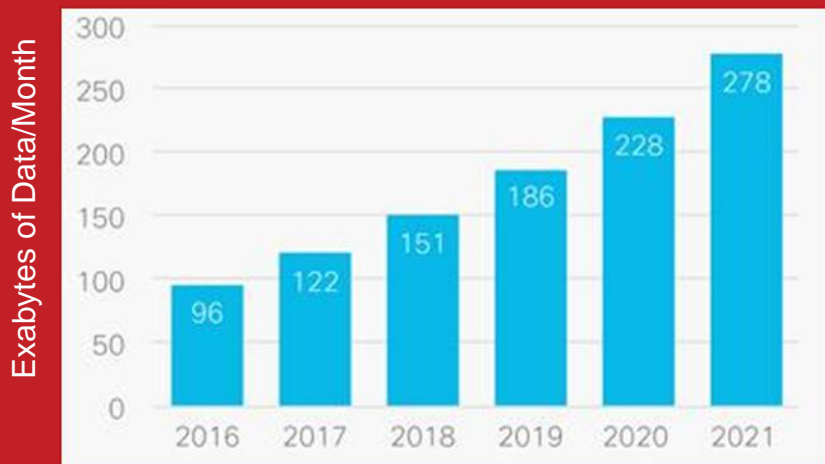
ML is hack proof!

- **WRONG!** If we can build it, hackers can build something better!

Adversarial Machine Learning

The study and design of machine learning algorithms that can resist attacks

Cisco VNI forecasts 278 EB per month of IP traffic by 2021



Source: Cisco VNI, 2017.



Taxonomy of Attacks



Influence > Causative & Exploratory

- Causative – impacts Training
- Exploratory - impacts Learning



Specificity > Targeted & Indiscriminate

- Targeted – Specific features
- Indiscriminate - Any classifier



Security > Integrity & Availability

- Integrity – Treats data, sample
- Availability – Overwhelms FP



Evasion and Poisoning

- Evasion: Increases FN w/perturbations
- Poisoning: Impacts Training data



Privacy and Availability

- Balance required

To ensure we understand adversarial attack vectors

Demonstration of Adversarial Attack Using Digits

Original Training



Following adversarial attack



Demonstration of Adversarial Attack

- **DREBIN Android malware dataset**
 - 625 Malware samples of *Fake Installer*
 - 700 Benign samples
- **A Deep Learning Neural Net was developed (4 layer)**
 - ~700 Features
 - 99.47% accuracy with Validation Set
 - 97.67% accuracy with Test Set with 1 False Negative
- **Developed evasion attack with only modifying Feature values**
 - **99.67%** malware evaded detection by Neural Net
 - Average number of Features modified: **11**



Example of Additional Innovation Areas

- Human Machine Teaming
- Artificial Intelligence
- File-less and server-less
- IoT
- Secure Home Gateway





Separating the Hype from Reality

- How often does the ML model actually “learn” ?
- How accurate is the ML model?
- Is the ML model diagnostic or predictive?
- How does your “data decay” impact the model?

Avecto + McAfee joint value proposition

Together is power

defendpoint

Privilege Management

+

 McAfee™

ePO and TIE/DXL

Client deployment | Policy management | Centralized auditing and reporting | Actionable intelligence

Dan Deganutti
Canadian Director

Dan.Deganutti@Avecto.com

McAfee Data Exchange Layer

Standardized integration and communication to break down operational silos

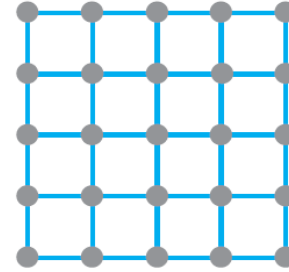
Disjointed API-Based Integrations



Result

- Slow, heavy, and burdensome
- Complex and expensive to maintain
- Limited vendor participation
- Fragmented visibility

Collaborative Fabric-Based Ecosystem (DXL)



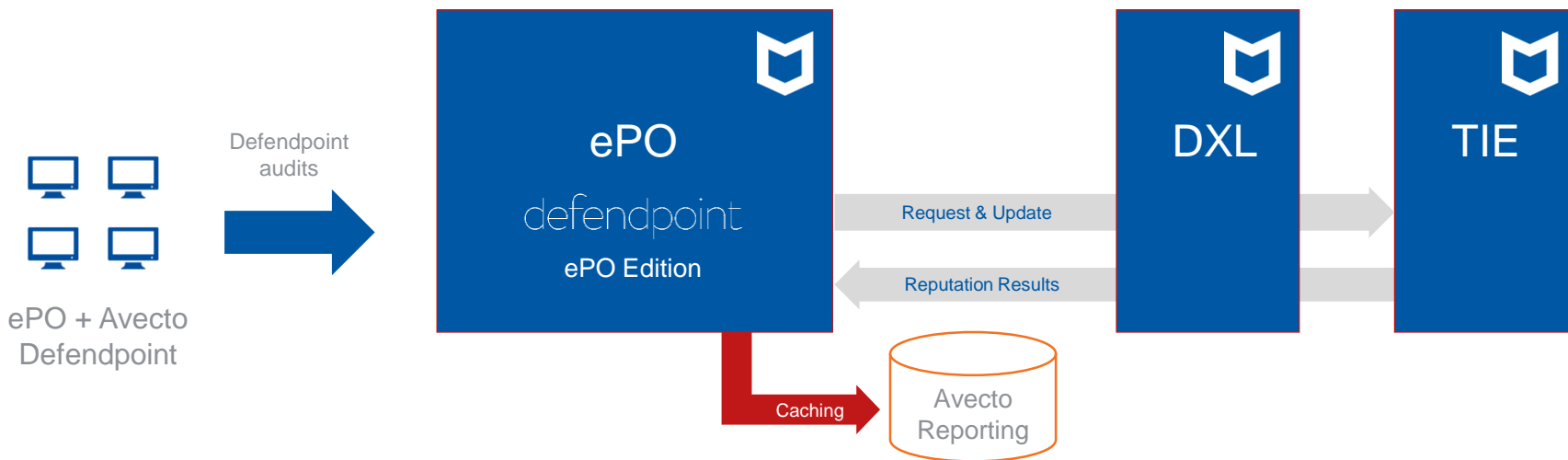
Result

- Fast, lightweight, and streamlined
- Simplified and reduced TCO
- Open vendor participation
- Simplicity- one time integration

Defendpoint ePO Edition + McAfee DXL

One of the first SIA partners to integrate with DXL/TIE

- Avecto Defendpoint is DXL ready and integrated with Threat Intelligence Exchange (TIE)
- TIE application reputation information used to drive Defendpoint configuration changes
- Defendpoint reporting enables risk-based policy adjustments



SIA MOST VALUABLE PARTNER OF THE YEAR, 2017
McAfee

Why Privilege Management?

Typical use cases

- **Reduce the attack surface** exponentially
- Over-locked or under-locked users
- Achieve compliance
- Security audits—internal and external
- Insider and outsider threats
- Legacy or custom-developed apps
- Remote access to printer drivers/network settings
- Limit third-party vendors
- Help desk calls and limited visibility of the IT estate
- Happy, productive, safe users



Threat landscape

Reasons to remove admin rights



94%

of **critical vulnerabilities** reported by Microsoft in 2016 would be mitigated by removing admin rights

Microsoft Vulnerabilities Report 2016



100%

of **Internet Explorer vulnerabilities** in 2016 would be mitigated by admin rights

Microsoft Vulnerabilities Report 2016



85%

of **intrusions mitigated** by implementing the **Top 4 Mitigation Strategies:** Whitelisting, patching, and least privilege

Australian Government Defence Signals Directorate



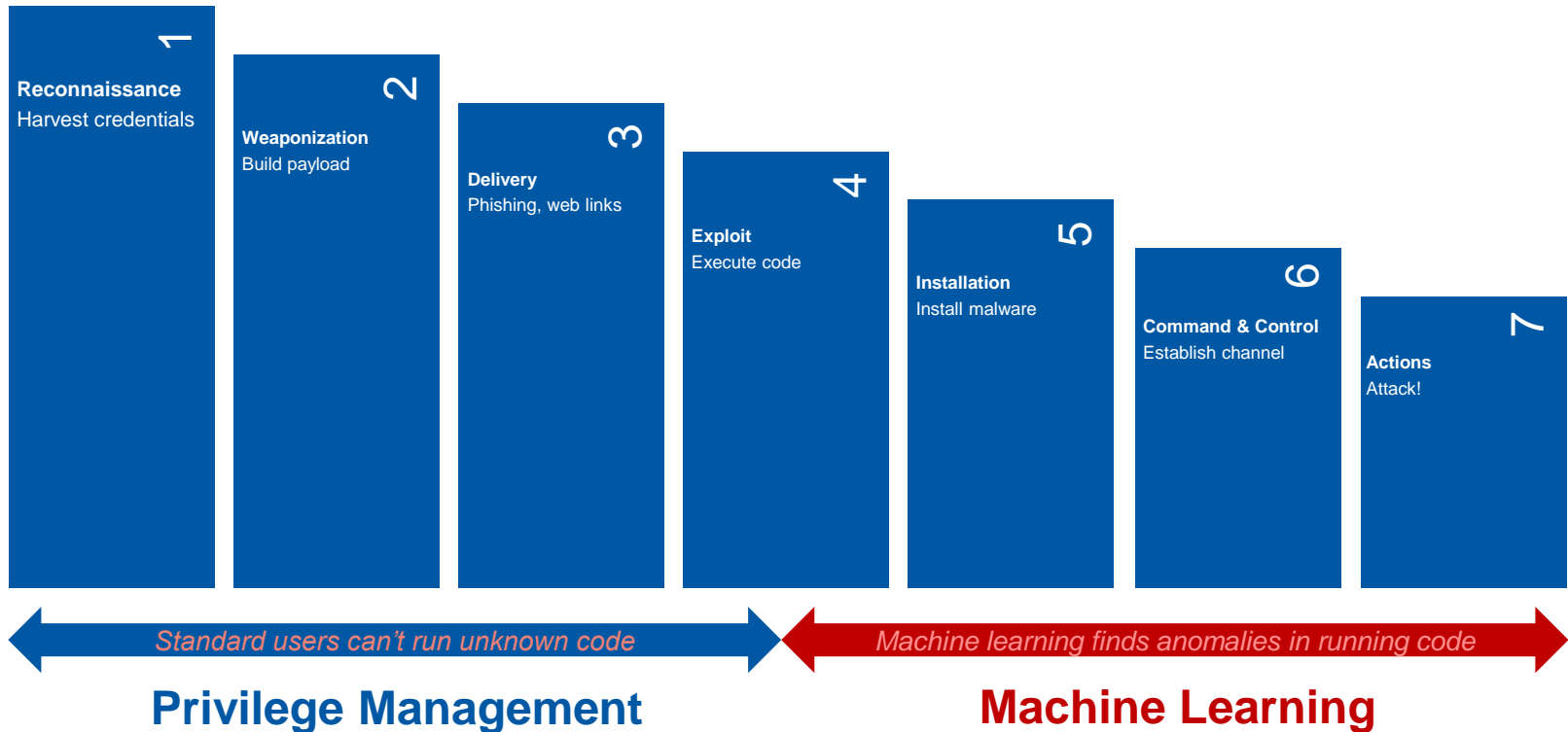
80%

of security breaches involve **privileged credentials***

Forrester Wave Privilege Identity Management 2016

Cyber Kill Chain

Privilege management and machine learning: better together



Avecto Defendpoint ePO edition

A **fully integrated solution** to quickly and successfully remove admin rights

Centralized management via McAfee ePolicy Orchestrator[®]

Technology integration and real time intelligence with McAfee Threat Intelligence Exchange and Data Exchange Layer (TIE/DXL)

A **proactive approach** to endpoint security

Reduce your attack surface **overnight**





Summary

- Adversarial ML is a viable threat and one of many innovation areas
- Myths and misleading statements abound regarding analytics
- We are continuing to integrate ML, DL, and AI into our solutions and ecosystem



McAfee, the McAfee logo and [insert <other relevant McAfee Names>] are trademarks or registered trademarks of McAfee, LLC or its subsidiaries in the U.S. and/or other countries.

Other names and brands may be claimed as the property of others.
Copyright © 2017 McAfee, LLC.