
aws
SUMMIT
ATLANTA | MAY 18–19, 2022

EXC102

Learning from Amazon: AI/ML at Amazon.com

Laura Squier
Specialist, AI/ML Business Development
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What we will cover today

- Why ML is widely used at Amazon.com – supply chain example
- How ML models transform the Amazon.com customer experience
- Lessons and insights from Amazon.com’s journey

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

"We intend to build the world's most customer-centric company. But there is no rest for the weary. I constantly remind our employees to be afraid, to wake up every morning terrified. Not of our competition, but of our customers.

We consider customers to be loyal to us – right up until the second that someone else offers them a better service."

Jeff Bezos (in a 1998 letter to shareholders)

Founder and Executive Chair, Amazon

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer promise: Prime One-Day | Prime Now



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What powers the customer promise



400 million products forecasted every day

300+ Fulfilment Centers, 200K+ robots, 19K trailers, and 30K+ delivery vans

Accurate delivery promise

(1 billion+ packages shipped to Prime customers last year)

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How does Amazon.com get the right items in the right place . . . even before the customer clicks the “Buy” button?

Obvious





Not so obvious





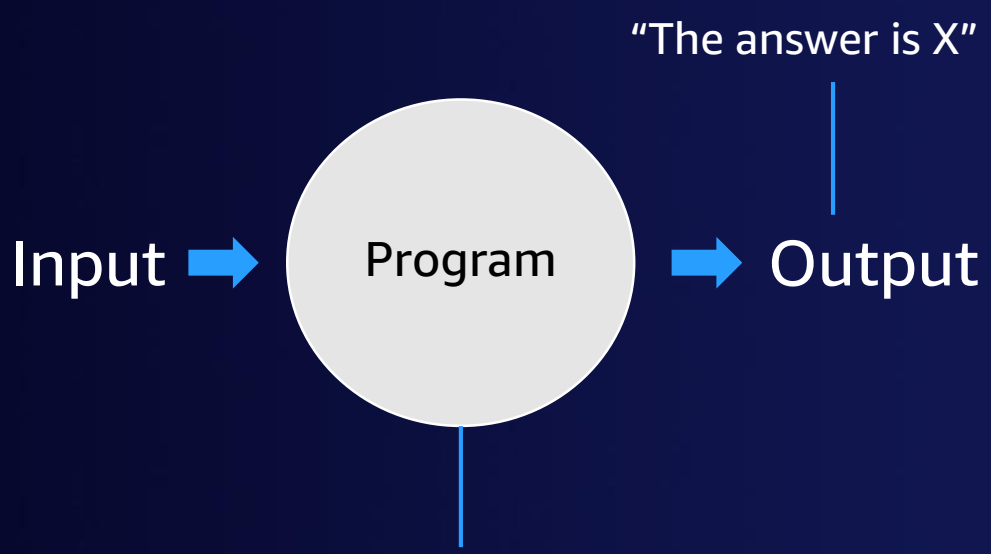


aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

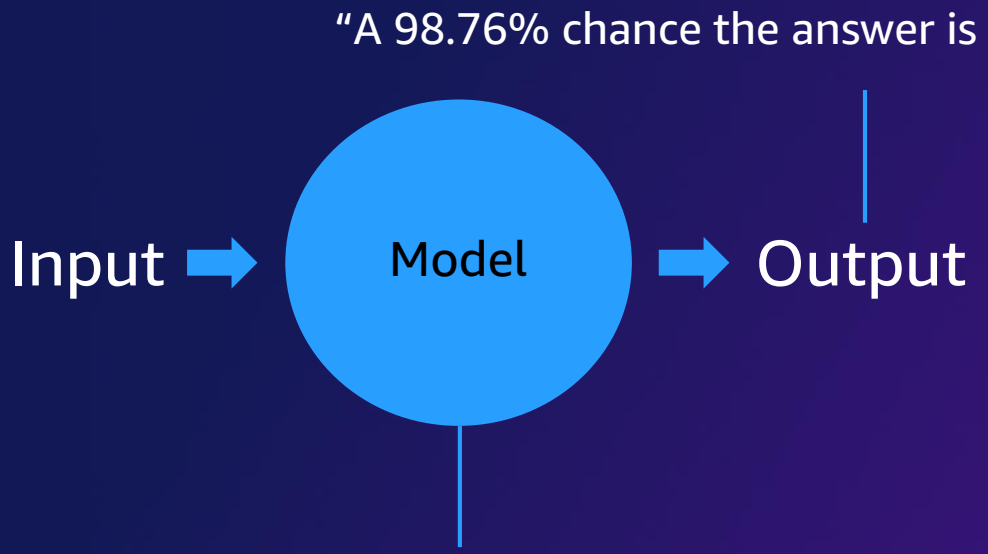
Forecasting models that learn from data

Traditional approach



A bunch of rules humans coded (e.g., IF . . . THEN . . . ELSE); limited by what humans can observe and program

Machine learning



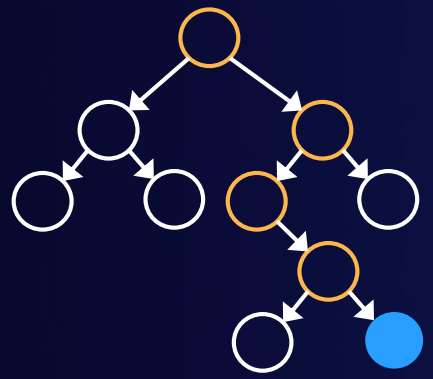
What the machine "learned" and generated from being trained by hundreds of millions of data points

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Predicting product demand at Amazon.com

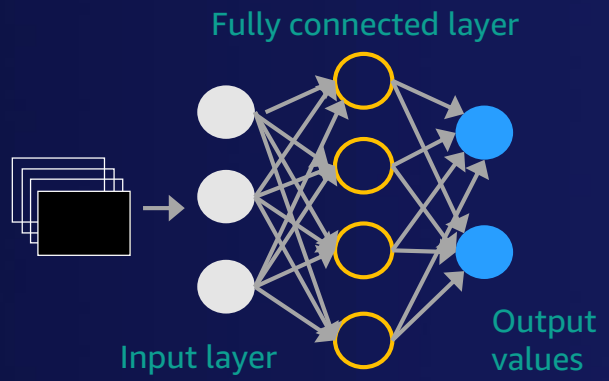
Circa 2007

Decision tree algorithm-driven model



Today

Convoluted neural networks



Billions of training data points

Powerful non-linear relationships

Automated feature engineering

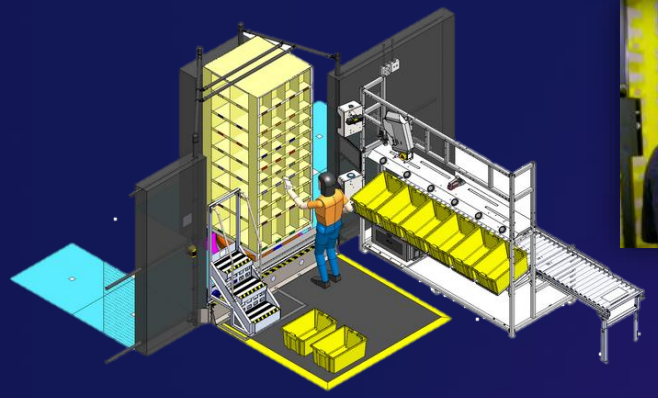
Forecast 400M+ items every day

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Package delivery: a symphony of humans and robots







- Computer tracks millions of **randomly stowed items**
- A fleet of **robots bring storage mods** to humans
- Guided stow/pick activities through the **Bin Vision System**

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What other customer experiences are transformed through ML (at scale)?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

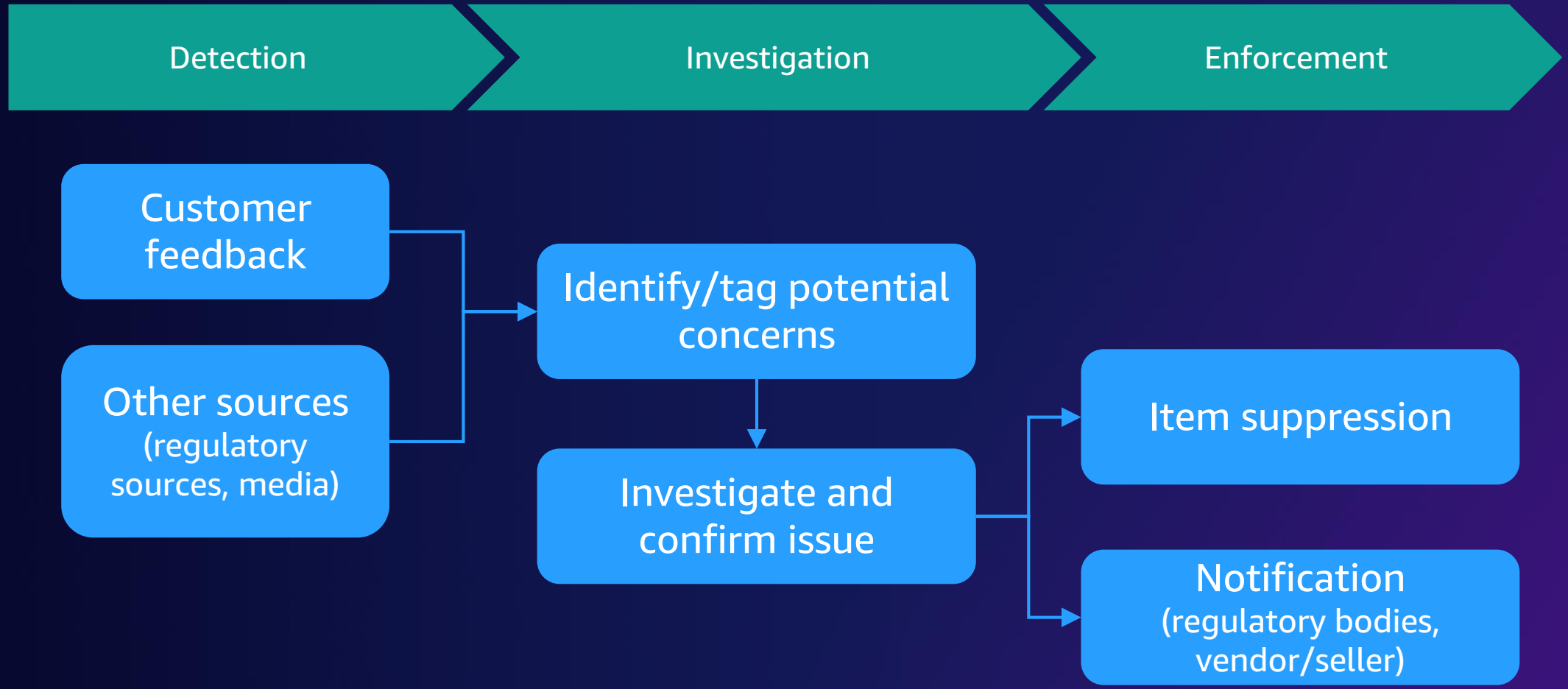
A delicious treat . . . or a medical risk?





© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What are the limitations of this approach?



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Predicting safety risk before products ship



- Billions of product and sales data updates
- Hundreds of thousands of concerns a year
- Predictive and highly accurate models

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why does shipping packaging matter?

Less of this . . .



and (a lot) more of that



75% lighter
than similarly-sized boxes

40% less space
during shipping

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML model changing the game

Basic decision-making rules

+

Visual inspection of top products

Handle small fraction of products

Untapped data

Limits to what an individual can observe

Ability to process all products Amazon sells

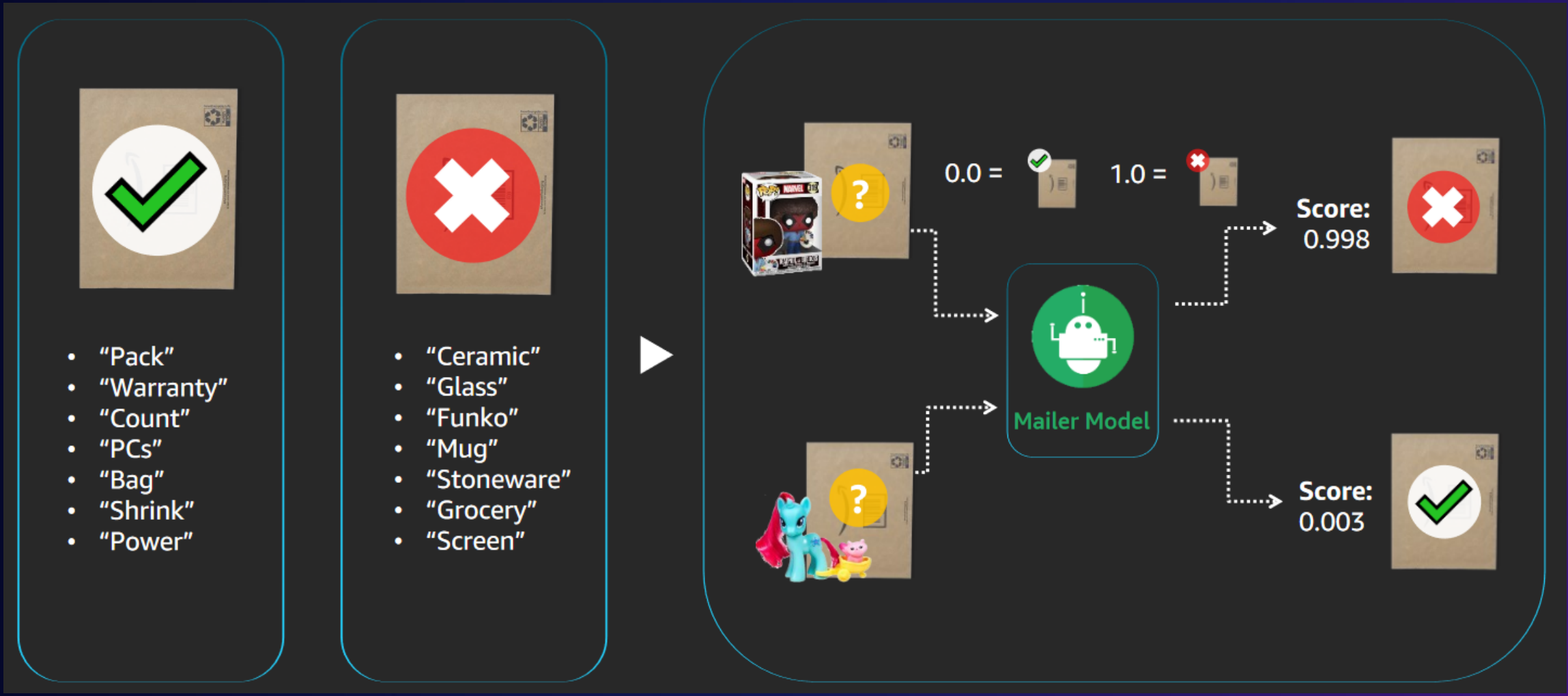
Apply complete set of intelligence generated in the delivery process

Handle non-obvious use cases (e.g., collectables)

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

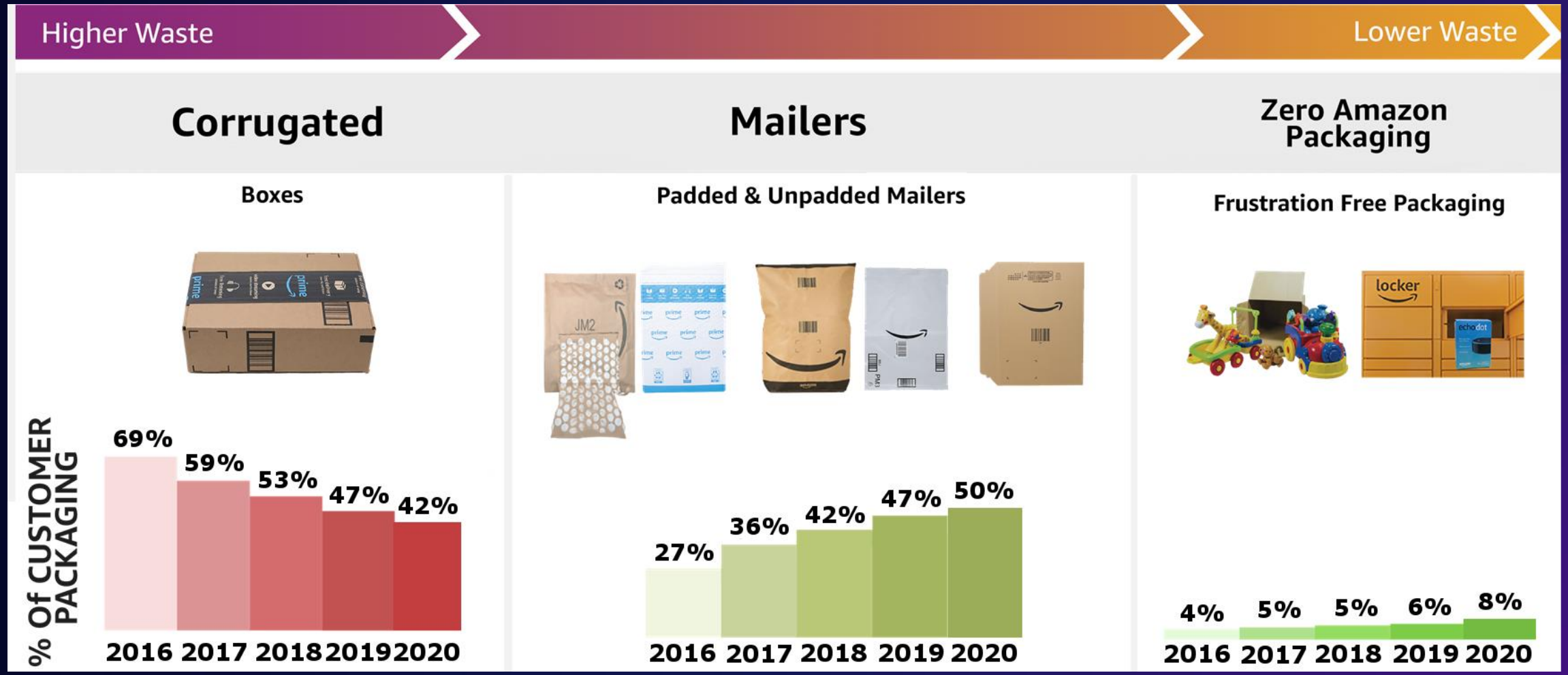
What is a collectable?



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

>50% of Amazon's shipments in mailers in 2020



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

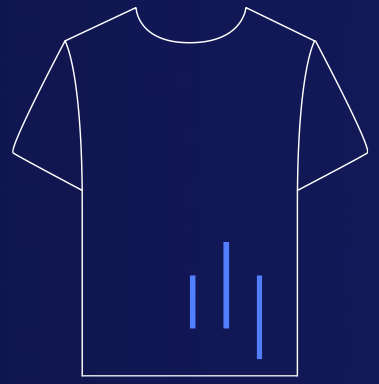
Team can launch new models in weeks



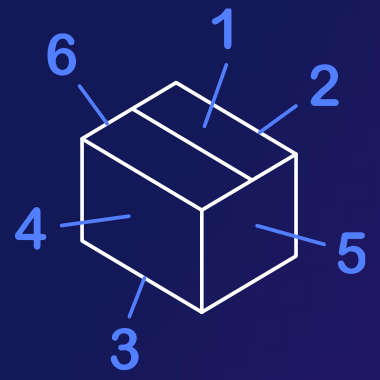
Suitable for paper bag?



Might leak?



Foldable?



No Amazon package needed?



Prone to damage?

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon + AI/ML = better customer experience

Automation

Better, cheaper, and faster

Augmentation

Smarter and more convenient

Enrichment

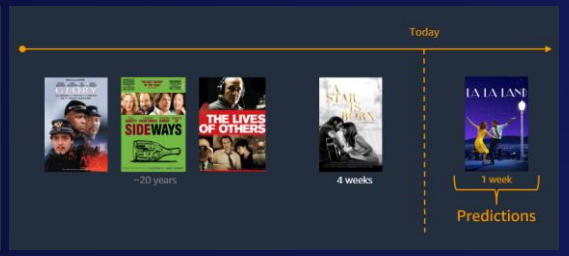
New features/services

Invention

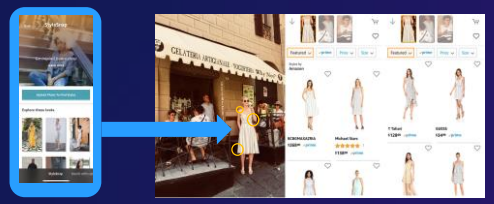
New products and categories



Supply chain automation and operations



Personalized Amazon Video recommendation



StyleSnap – find that style in the picture



No check out lines – ever

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What did Amazon learn was most important to get right in the ML journey?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

It’s (always) the basics

? What are the most common challenges with AI/ML adoption?

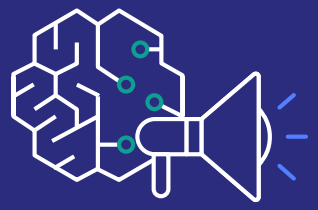
Challenge	% of respondents
Skills of staff	56%
Understanding AI benefits and uses	42%
Data scope and quality	34%
Finding use cases	26%
Integration complexity	26%
Defining the strategy	25%
Security or privacy concerns	20%
Measuring the value	17%
Governance issues or concerns	13%
Finding funding	12%

Source: Gartner Research Circle members survey, July 2019

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Solve for a system of enablers



ML-first mindset and culture



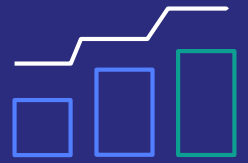
Empowering teams for the mission



Powerful data platform



Choosing the right first project



Scaling

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Start with one question



ML-first mindset and culture

How will you use machine learning?

(“We won’t” is not an acceptable answer)

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Focus team on delivering value



ML-first mindset and culture

Our tenets (unless you know better ones)

1. We are successful when our partners succeed
2. We are teachers
3. **We innovate to achieve breakthroughs**
4. Customer trust is paramount
5. **We are builders**
6. **We enable automated data-driven decision-making**
7. We build for internet scale
8. We enable ML application development at speed
9. We are talent magnets

3 Many of the opportunities at Amazon have no known scientific solutions. **We take risks** and go beyond the frontier of known methods.

5 Our expertise makes us good consultants, but we are only **delivering results** when platforms, tools and solutions make it to production.

6 We replace “magic numbers” and rule-based logic with models that learn from data. We help **optimize every system** at Amazon using ML.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Break down the silos



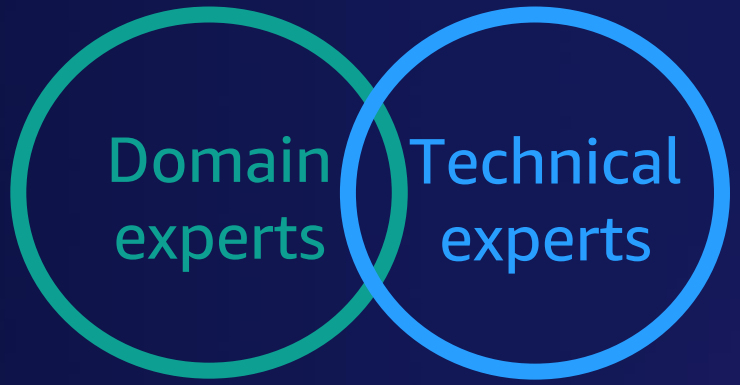
Empowering teams for the mission



Product manager
Requirements/validations

Data scientist
Model development and testing







Data engineer
ML data foundation and performance



Research scientist
Cutting-edge theories and applications



ML platform engineer
Model production and optimization

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Evolve the teams as you grow and scale



Empowering teams for the mission

	How teams might start	Scaling within teams	Scaling across teams
Team level	 Product manager	 Product manager  Data scientist	 Product manager  Data scientist
Function/LoB level	 Data engineer	 Data engineer	 Research scientist  Platform engineer  Data engineers
Enterprise level/ consultants	 Data scientist	 Platform engineers	 Other specialist

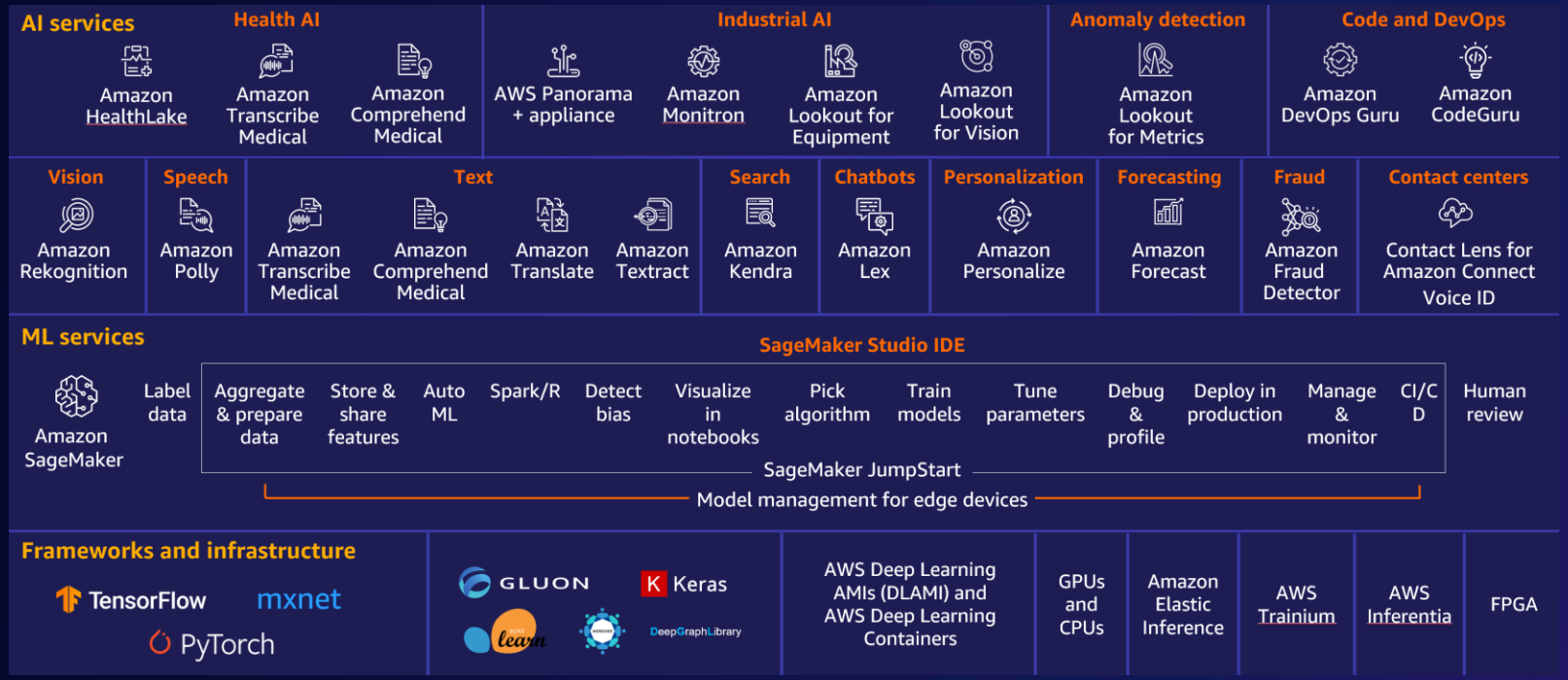
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Apply the right tool for the job

Empowering teams for the mission

Example: AWS AI/ML service stacks



Simple, powerful and ready to use

Control and flexibility at the most granular levels

TensorFlow, the TensorFlow logo and any related marks are trademarks of Google Inc.
PyTorch, the PyTorch logo and any related marks are trademarks of Facebook, Inc.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data fuels ML-driven innovations



Amazon.com's big data marketplace

- Innovation happens at the edge
- Self-service discovery and subscription
- Bring your own clusters/queries

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Speed and substance is critical



Choosing the right first project

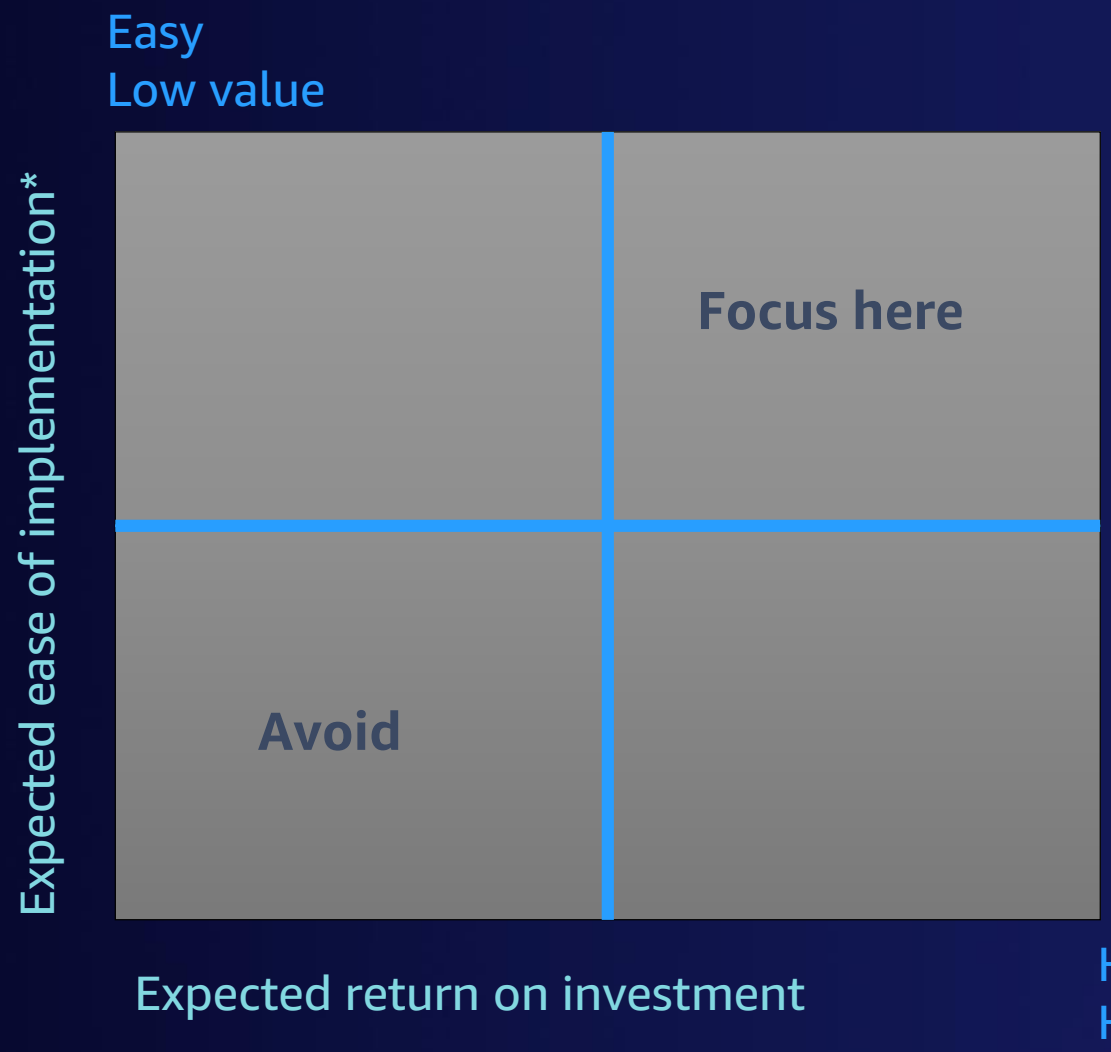
1. Does it solve a real and significant problem for your customer?
2. Where/how does machine learning unlock new capabilities?
3. Are there places where we already have a lot of untapped data?
4. Can we achieve success in 6–10 months?
5. Is it important enough to get sustained attention and support?

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Easy prioritization matrix

Who Should Participate?
Leadership, line of business, domain experts, data science team



Ease of implementation:
Includes data availability and ability to execute; how simple it is to assemble data, develop models, and deploy models to include any cultural and process changes that must occur to use the results

ROI:
Includes both hard and soft ROI, where the value is measured as reduced costs, improved profits, improved satisfaction, enhanced productivity, etc.

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Remove that bottleneck

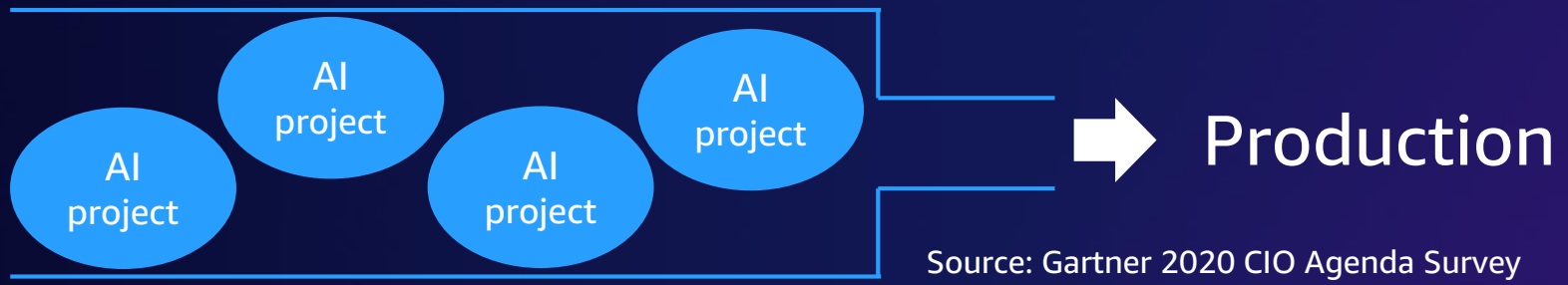
Scaling

Number of AI projects estimated for 2021 compared to 2020

2x

Planned to actually deploy solutions

40%



Source: Gartner 2020 CIO Agenda Survey

aws

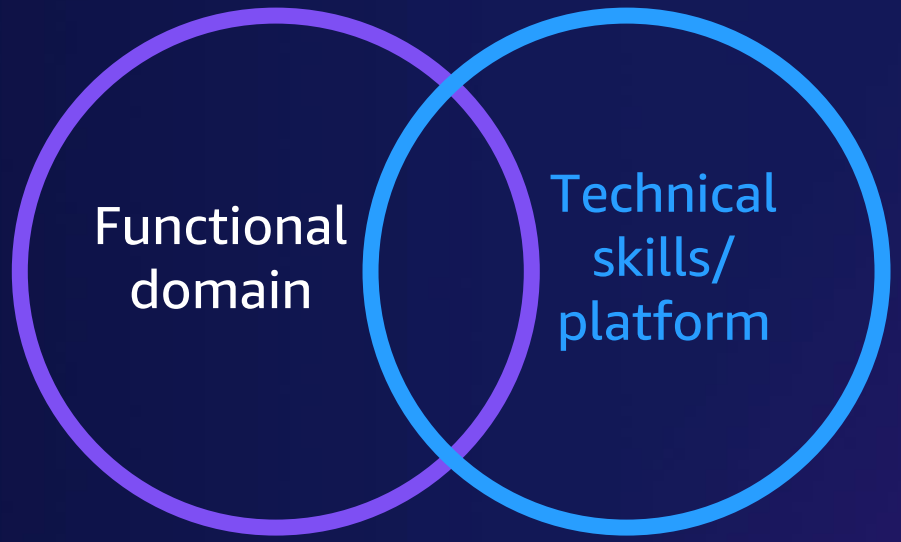
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Disconnected tools and processes block scaling

Scaling

1. Lack of common mentality/priorities
2. Teams in silos, not cross-functional

3. Inability to define and measure actual/ differentiated business outcome



6. Lack of integrated end-to-end data pipeline and production platform

4. Unique aspects of ML delivery methods
5. Governance, compliance, and security

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Eliminate undifferentiated heavy lifting from teams



Legacy ML platforms

- Homegrown platform, tools, and assets
- Multiple different ML platforms and pervasive one-time ML practices
- Perception of custom platforms are cheaper to run

"Hidden costs" blocking speed and scale

- Significant ongoing manual effort to maintain, retrain, and optimize models
- Additional IT/engineering expertise needed to set up and deploy ML solutions
- Business impact from increasing cost of poor quality

Increasing use of AWS ML toolset

- Standardizing model development lifecycle and data pipeline infrastructure on Amazon SageMaker
- Tools are automating model tuning, retraining, and optimization – traditionally hard-to-find skills
- Deployment and maintenance done with "clicks"

Nonlinear growth in speed and scale examples

- Significantly improve the developer-scientist ratio (going from 4–5:1 to 1:1) in months
- Retraining time decreasing from days to just hours: better models, free up team
- More than 10x model coverage without incremental headcount increase

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Summary

1. Amazon sees AI/ML as **an investment** toward **delivering better services**
2. ML delivers a **differentiated impact** and **nonlinear scaling** for Amazon
3. Building and scaling ML usage at enterprise need requires clear strategies across the "**system of enablers**"
4. Creating the **right way of working** and providing teams with the **right tools for the job** is critical toward building and scaling ML-driven innovation engine
5. Identify and remove the blockers to **path to production**

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Common blockers and opportunities

Vision and strategy

- AI/ML ambition and vision clearly defined
- AI/ML initiatives clearly aligned to desired business outcome, and holding teams accountable
- The problem space to address with ML clearly understood and defined
- Clarity to where we see biggest impact overall, near-term impact, and what should be next

Culture, skills, and organization

- Clearly defined roles and interaction models, but with room to adapt and grow
- Talent development/ acquisition strategy is not an afterthought
- Applying and amplifying the patterns of successful cross-functional teaming
- “Hitch the wagon” to the AI/ML-enabled innovations to generate excitement and momentum

Data and technology foundations

- Address the biggest blockers/enablers from data and tooling perspectives but know where to start
- “Dogs not barking” and “paper cuts” to pressure-test where we are over/under investing
- Vocally self-critical – what do we need to try and learn more/faster to evolve our data and technology roadmap?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Learn in-demand AWS Cloud skills

AWS Skill Builder

Access **500+ free** digital courses and Learning Plans

Explore resources with a variety of skill levels and **16+** languages to meet your learning needs

Deepen your skills with digital learning on demand

Train now

AWS Certifications

Earn an industry-recognized credential

Receive Foundational, Associate, Professional, and Specialty certifications

Join the **AWS Certified community** and get exclusive benefits

Access **new** exam guides

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.