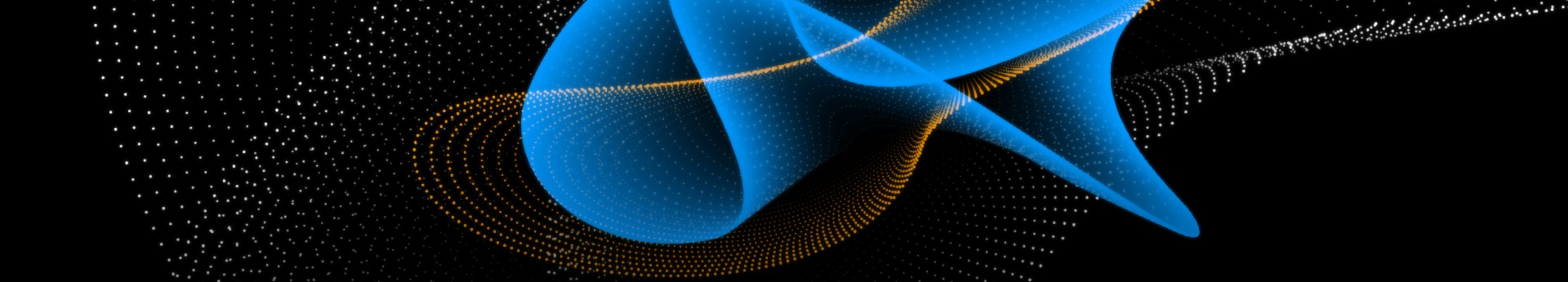


Amazon AI for every developer

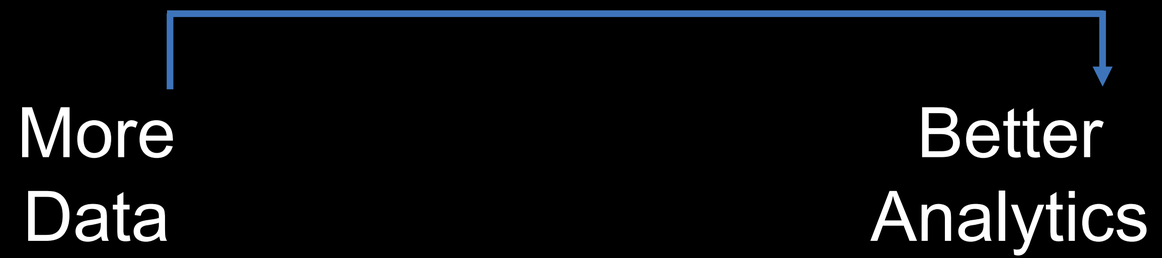
Julien Simon

AI Evangelist, EMEA

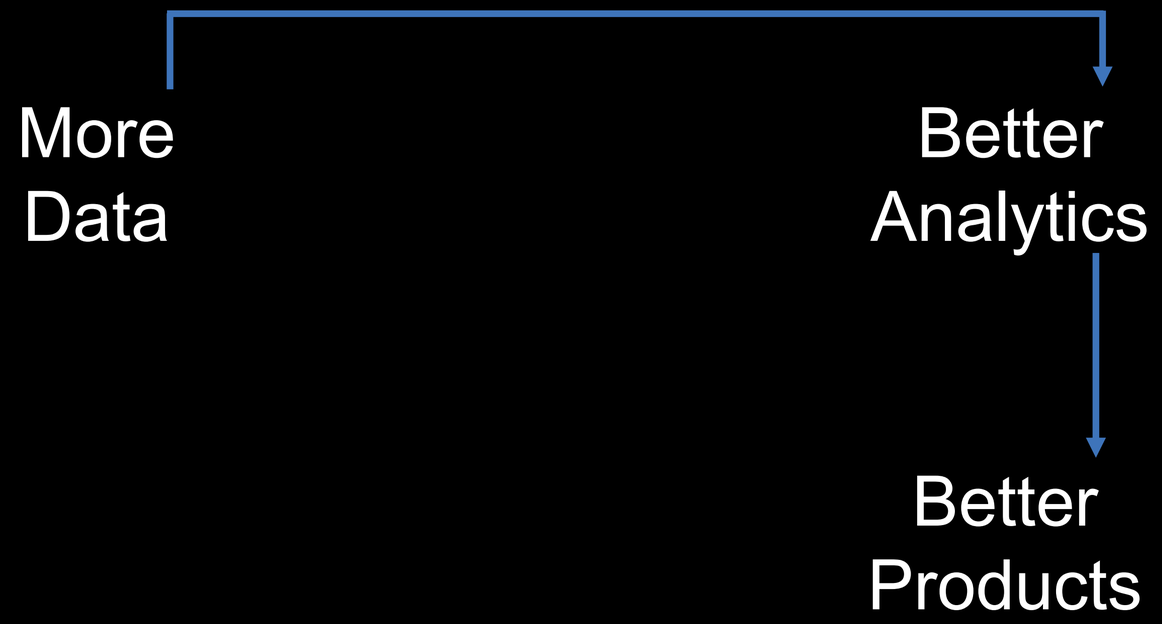
@julsimon

aws

A Flywheel For Data

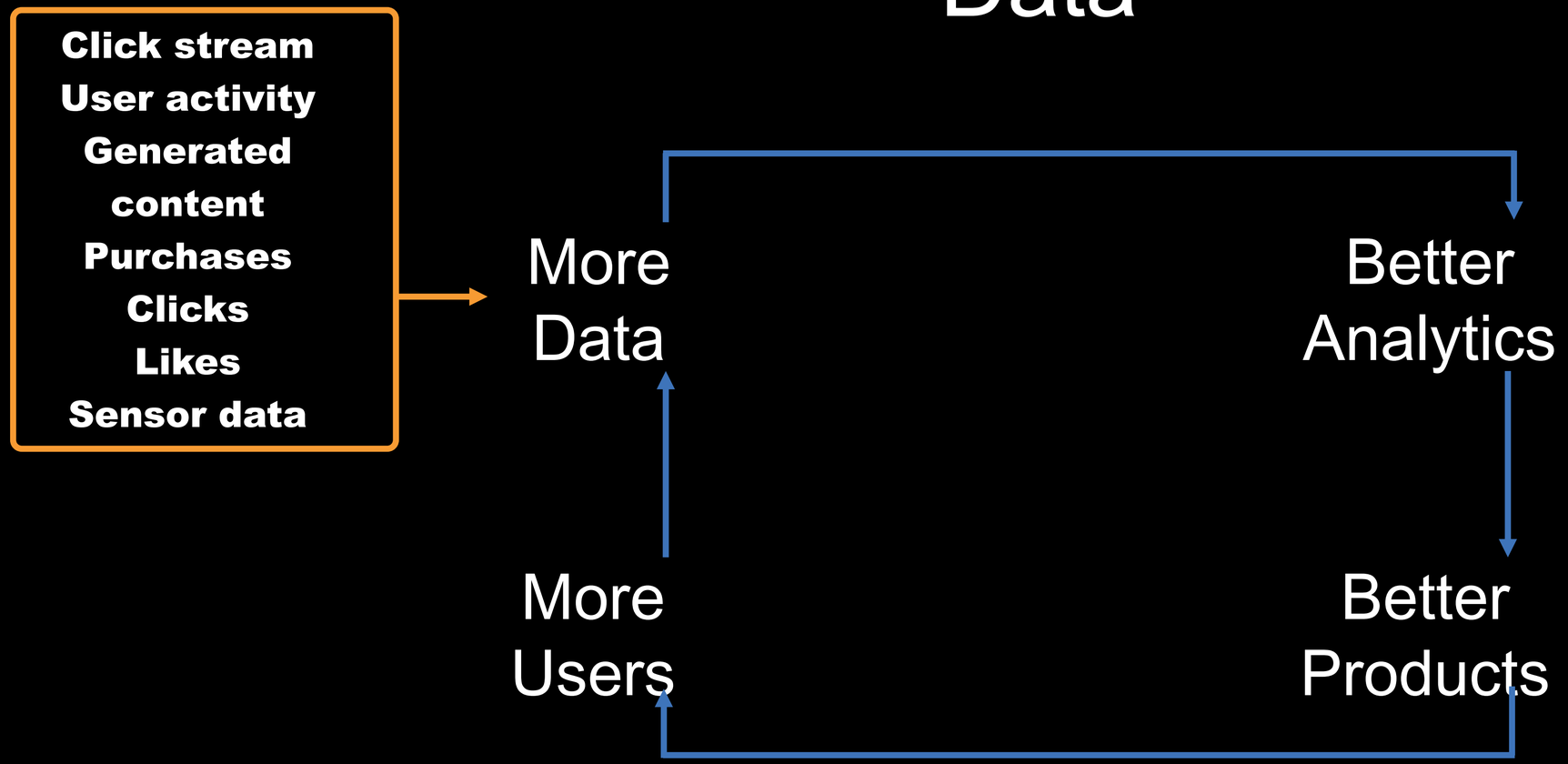


A Flywheel For Data

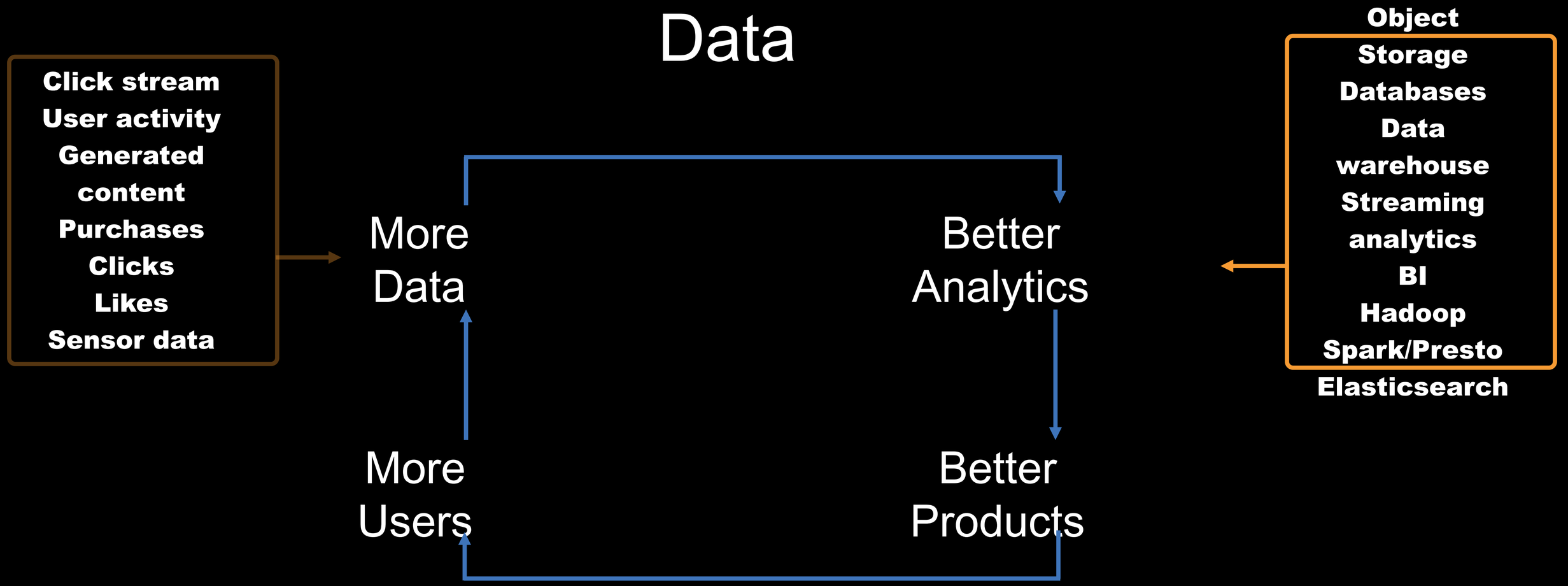


A Flywheel For Data

A Flywheel For Data



A Flywheel For Data



A Flywheel For Data

A Flywheel For Data

- Artificial Intelligence: design software applications which exhibit human-like behavior, e.g. speech, natural language processing, reasoning or intuition
- Machine Learning: teach machines to learn without being explicitly programmed
- Deep Learning: using neural networks, teach machines to learn from complex data where features cannot be explicitly expressed

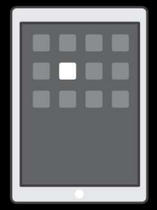
Amazon.com,1995

Artificial Intelligence At Amazon

Thousands Of Employees Across The Company Focused on AI

Discovery & Search

Fulfilment & Logistics



Enhance Existing Products

Define New Categories Of Products

Bring Machine Learning To All


amazon
echo
amazon
25,000 skills

Machine Learning for customer support

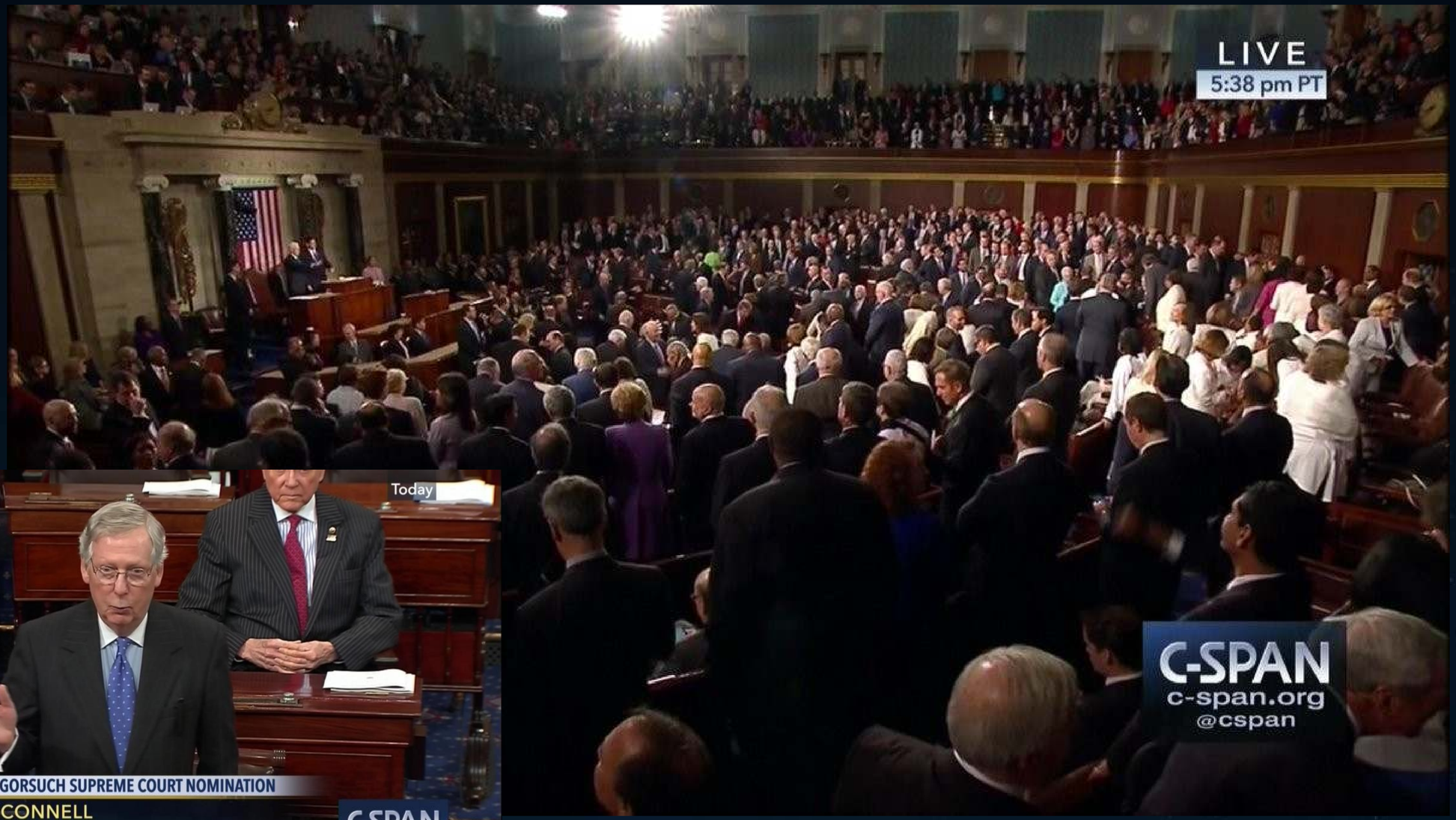
Not all customer interactions can be solved in a self-service mode. Therefore, Amazon operates large customer support centers where Customer Service Representatives (CSR) handle customer requests.

The machine learning models described above are used to optimize the human interactions of these requests.

For example, they are used to route the customer call to the best CSR before the customer has even started to speak! They are also used again during the call.

Selected customers running AI on AWS

C-SPAN



https://aws.amazon.com/solutions/case-studies/cspan/

- Expedia have over 10M images from 300,000 hotels
- Using great images boosts conversion
- Using Keras and AWS GPU instances, they fine-tuned a pre-trained Convolutional Neural Network using 100,000 images
- Hotel descriptions now automatically feature the best available images

https://news.developer.nvidia.com/expedia-ranking-hotel-images-with-deep-learning/

Liberty Mutual.
INSURANCE

https://www.libertymutual.com/liberty-mutual-mobile/amazon-alexa

https://techcrunch.com/2017/09/27/bmw-to-bring-alexa-to-its-cars-starting-in-2018/

Last June, tuSimple drove an autonomous truck

for 200 miles from Yuma, AZ to San Diego, CA

https://www.oreilly.com/ideas/self-driving-trucks-enter-the-fast-lane-using-deep-learning

julia

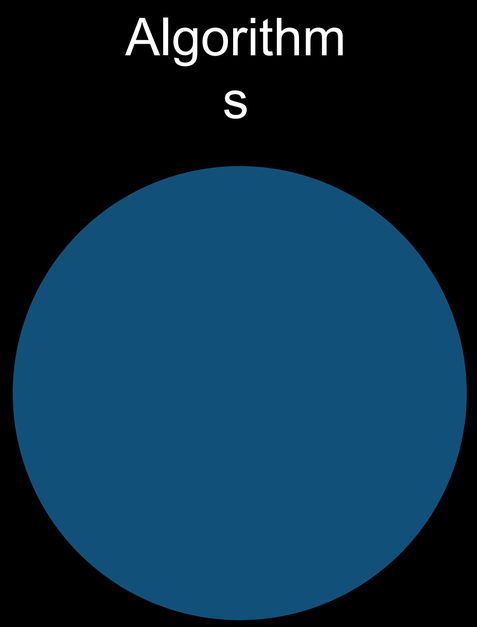
Detecting retinopathy with Deep Learning

https://juliacomputing.com/blog/2016/11/16/deep-eyes.html

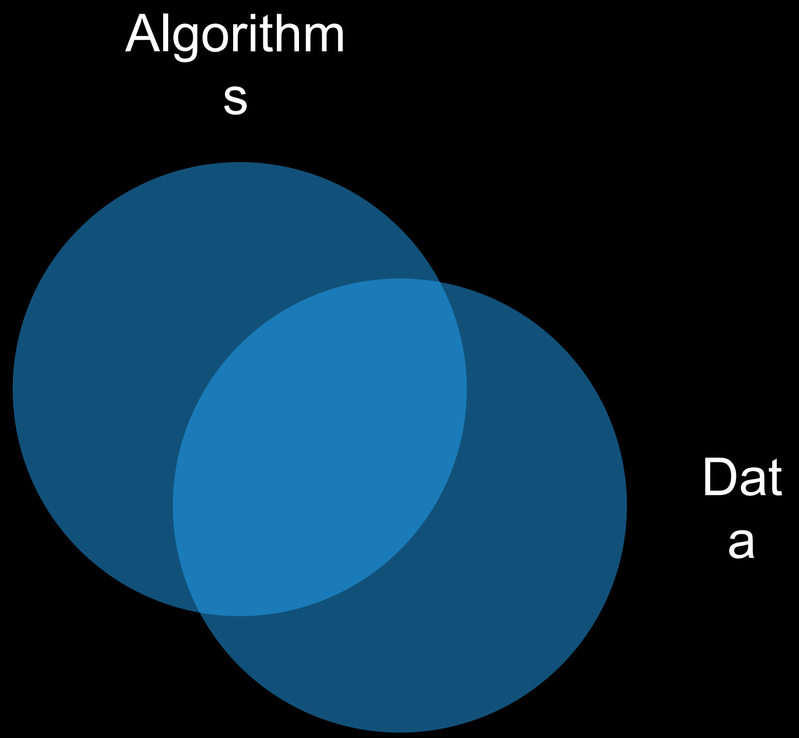
Amazon AI

Intelligent Services Powered By Deep Learning

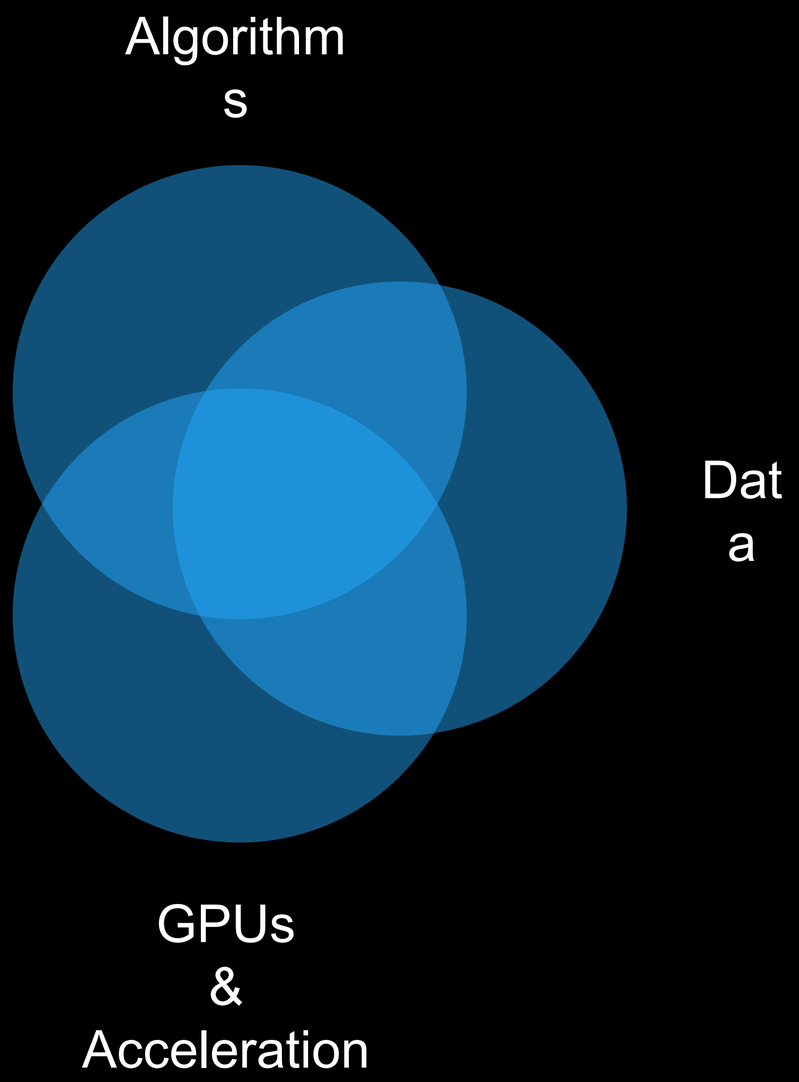
The Advent Of Deep Learning



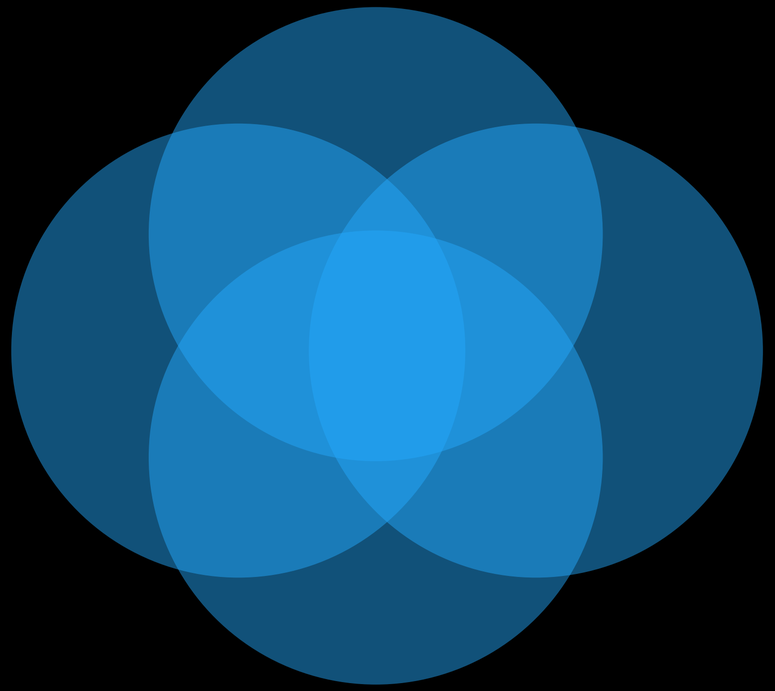
The Advent Of Deep Learning



The Advent Of Deep Learning



The Advent Of Deep Learning



Can We Help Customers
Put Intelligence At The Heart Of
Every Application & Business?

Questions, questions…

What's the business problem my company has failed to solve?

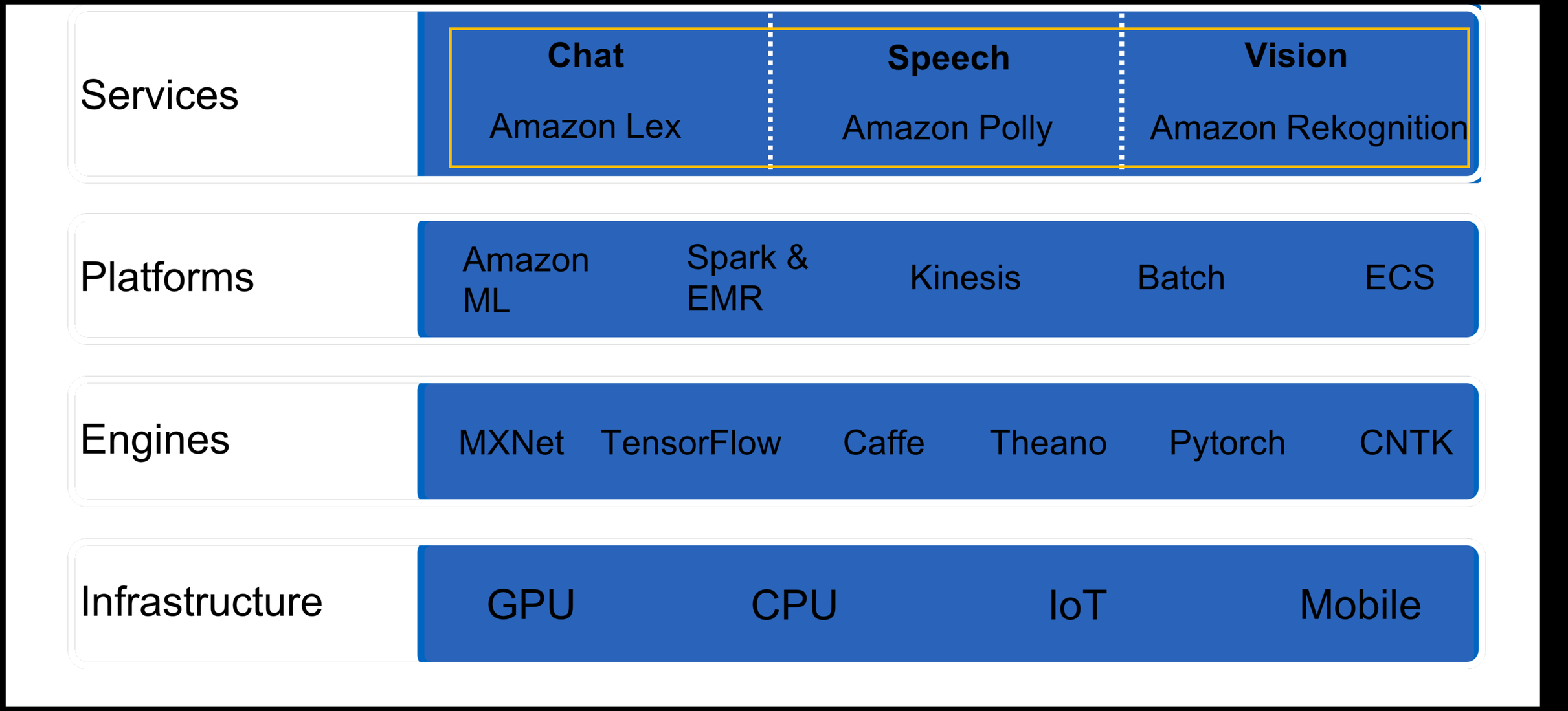
Should I design and train my own Deep Learning model?

Should I use a pre-trained model?

Should I use a SaaS solution?

Same questions as "Big Data" years ago

Amazon AI for every developer



Amazon AI: Three New Deep Learning Services

Pol
ly

Life-like
Speech

Amazon AI: Three New Deep Learning Services

Polly

Life-like Speech

Rekognition

Image Analysis

Amazon AI: Three New Deep Learning Services

Polly

Life-like Speech

Rekognition

Image Analysis

Lex

Conversational Engine

Amazon AI: Three New Deep Learning Services

Polly

Life-like Speech

Rekognition

Image Analysis

Lex

Conversational Engine

Polly: Life-like Speech Service

Converts text
to life-like speech

Fully
managed

50
voices

24
languages

Low
latency,
real time

Polly: A Focus On Voice Quality & Pronunciation

1. Automatic, Accurate Text Processing

-  “Today in Seattle, WA, it’s 11°F”
-  ‘"We live for the music" live from the Madison Square Garden.’

Polly: A Focus On Voice Quality & Pronunciation

1. Automatic, Accurate Text Processing

2. Intelligible and Easy to Understand

Polly: A Focus On Voice Quality & Pronunciation

1. Automatic, Accurate Text Processing
2. Intelligible and Easy to Understand
3. **Add Semantic Meaning to Text**


“Richard’s number is 2122341237“


“Richard’s number is 2122341237“

Telephone Number

Polly: A Focus On Voice Quality & Pronunciation

1. Automatic, Accurate Text Processing
2. Intelligible and Easy to Understand
3. Add Semantic Meaning to Text
4. **Customized Pronunciation**

“My daughter’s name is Kaja.”

“My daughter’s name is Kaja.”

Polly: Life-like Speech Service

High quality, through best-in-class deep learning

Deep functionality

Easy to use & thoughtfully integrated

Built for production

Low cost

Amazon AI: Three New Deep Learning Services

Polly

Life-like Speech

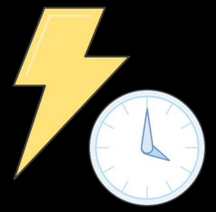
Rekognition

Image Analysis

Lex

Conversational Engine

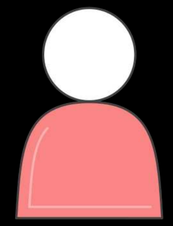
Rekognition: Search & Understand Visual Content



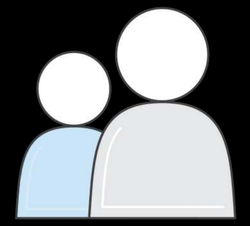
Real-time & batch image analysis

Object & Scene Detection

Facial Detection



Facial Analysis

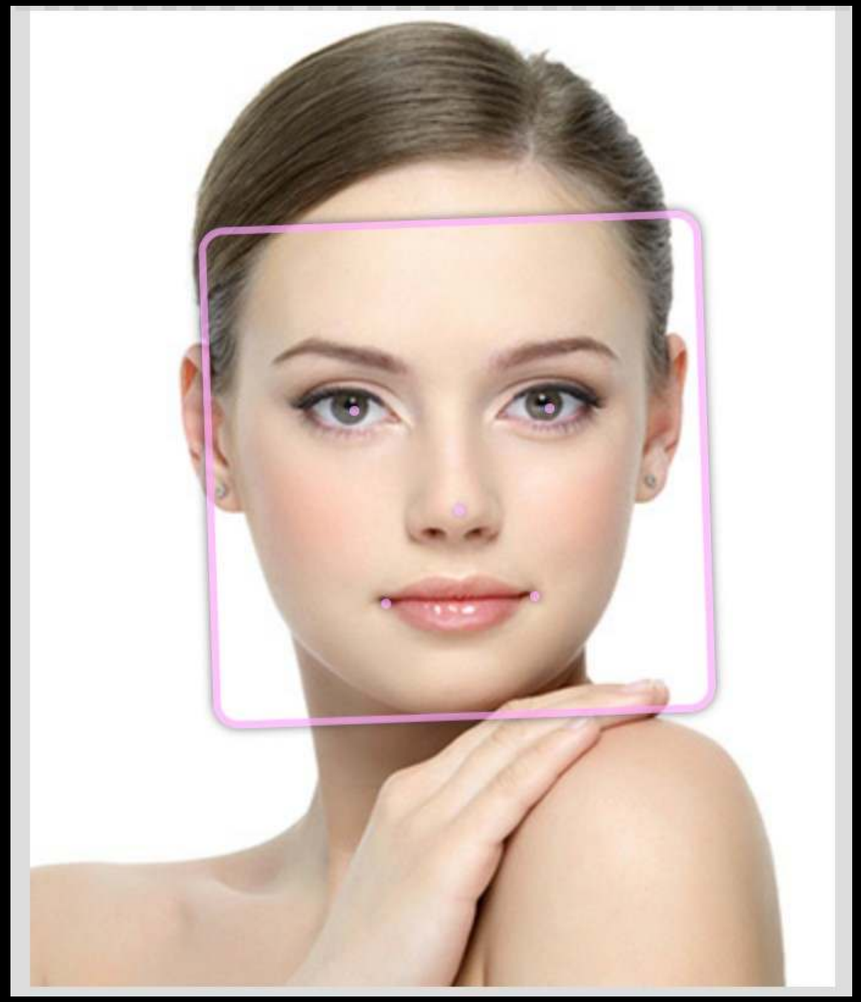


Face Search

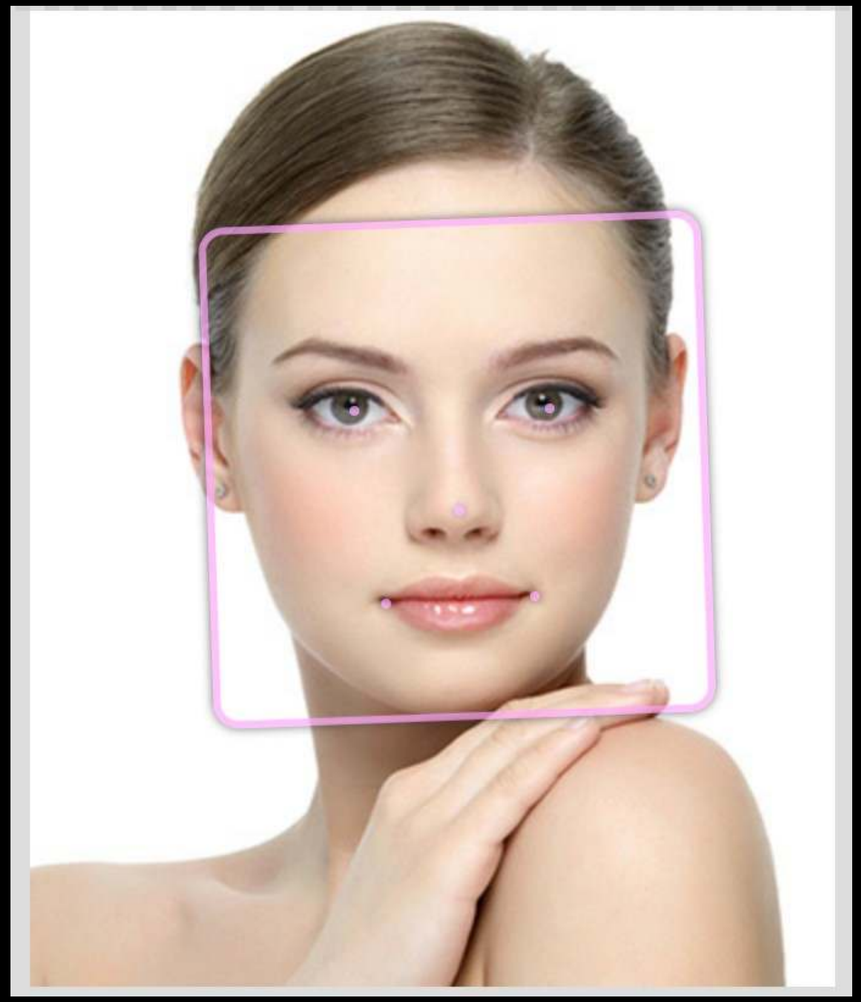
Rekognition: Object & Scene Detection

Category	Confidence
Bay	99.18%
Beach	99.18%
Coast	99.18%
Outdoors	99.18%
Sea	99.18%
Water	99.18%
Palm_tree	99.21%
Plant	99.21%
Tree	99.21%
Summer	58.3%
Landscape	51.84%
Nature	51.84%
Hotel	51.24%

Rekognition: Facial Detection

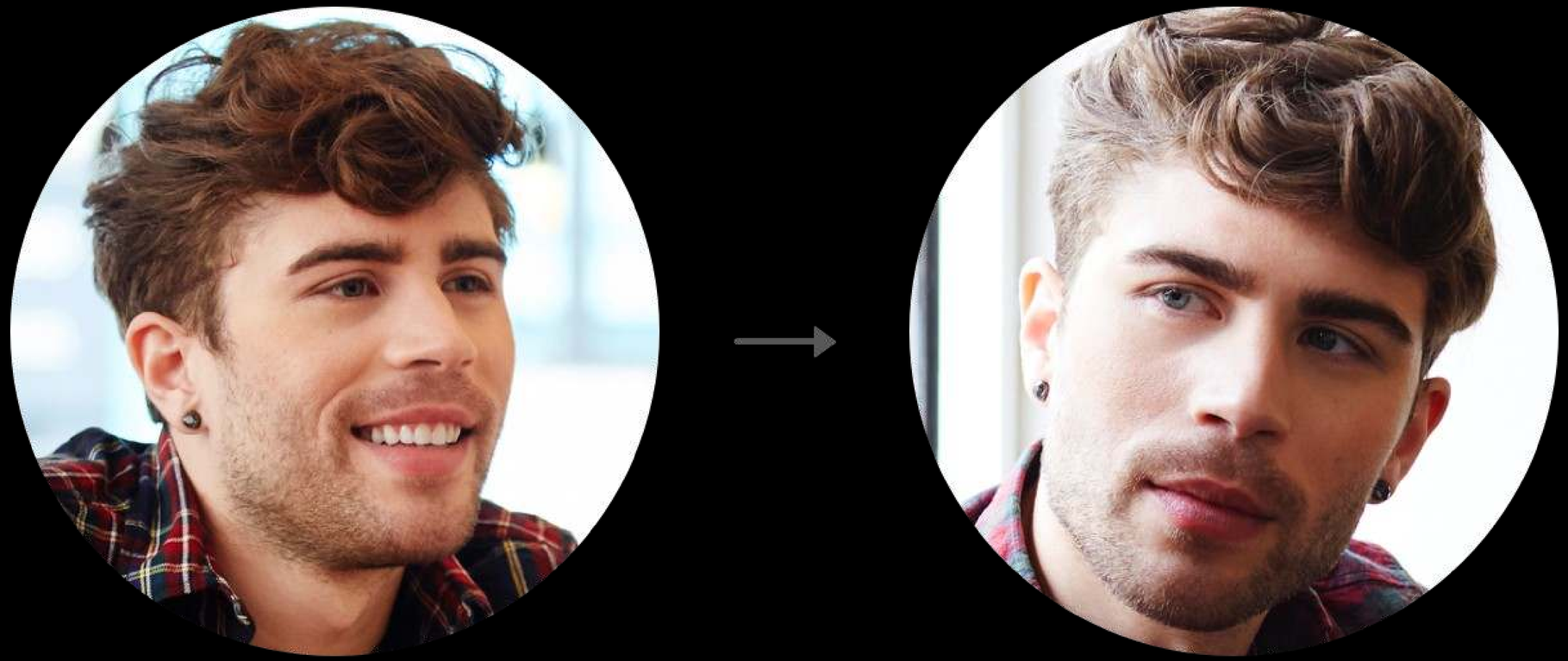


Rekognition: Facial Analysis



Emotion: calm: 73%
Sunglasses: false (value: 0)
Gender: female (value: 0)
Mouth open wide: 0% (value: 0)
Eye closed: open (value: 0)
Glasses: no glass (value: 0)
Mustache: false (value: 0)
Beard: no (value: 0)

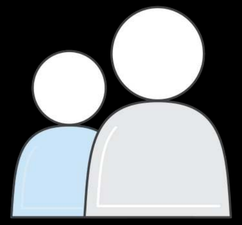
Rekognition: Compare Faces



Similarity: **97.0%**

Rekognition: Facial Search

Facial verification
(compare two faces)



Face Search
(compare many faces)

Visual Similarity Search
(find similar faces)

Rekognition: Search & Understand Visual Content

High quality, through best-in-class deep learning

Deep functionality

Easy to use & thoughtfully integrated

Built for production

Low cost

Amazon AI: Three New Deep Learning Services

Polly

Life-like Speech

Rekognition

Image Analysis

Lex

Conversational Engine

Lex: Build Natural, Conversational Interactions In Voice & Text

Voice & Text
“Chatbots”

Power
s
Alexa

Voice
interactions
on mobile, web
& devices

Text interaction
with Slack &
Messenger
(with more
coming)

Enterprise
Connector
s

Salesforce
Microsoft
Dynamics
Marketo
Zendesk
Quickbooks
Hubspot


Flight Booking
Origin
Destination
Departure Date

“Book a flight to London”

Flight Booking	
Origin	
Destination	
Departure Date	

“Book a flight to London”

Book Flight London

Automatic Speech Recognition

Flight Booking	
Origin	
Destination	
Departure Date	

"Book a flight to London"

Book

Flight

London

Automatic Speech Recognition

Flight booking

London Heathrow

Knowledge Graph

Natural Language Understanding

Grammar Graph

Utterances

Flight Booking	
Origin	
Destination	
Departure Date	

"Book a flight to London"

Book

Flight

London

Automatic Speech Recognition

Flight booking

London Heathrow

Natural Language Understanding

Knowledge Graph

Grammar Graph

Utterances

Flight Booking	
Origin	
Destination	London Heathrow
Departure Date	

"Book a flight to London"

Location

Book

Flight

London

Automatic Speech Recognition

Location

Flight booking

London Heathrow

Knowledge Graph

Natural Language Understanding

Grammar Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	

"Book a flight to London"

Location

Automatic Speech Recognition

Book | Flight | London

Location

Natural Language Understanding

Flight booking

London Heathrow

Knowledge Graph

Grammar Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	

"When would you like to fly?"

Prompt

"Book a flight to London"

Location

Book

Flight

London

Automatic Speech Recognition

Location

Flight booking

London Heathrow

Knowledge Graph

Natural Language Understanding

Grammar Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	

"When would you like to fly?"

Prompt

Polly

"When would you like to fly?"

“Next Friday”

“When would you like to fly?”

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	

Next Friday

“Next Friday”

Automatic Speech Recognition

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	

"Next Friday"

Next

Friday

Automatic
Speech Recognition

Flight booking

11/18/2016

Knowledge
Graph

Natural Language
Understanding

Grammar
Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	11/18/2016

"Next Friday"

Next

Friday

Automatic
Speech Recognition

Flight booking

11/18/2016

Natural Language
Understanding

Knowledge
Graph

Grammar
Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	11/18/2016

λ

"Next Friday"

Next Friday

Automatic
Speech Recognition

Flight booking

11/18/2016

Knowledge
Graph

Natural Language
Understanding

Grammar
Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	11/18/2016

λ

"Your flight is booked for next Friday"

Confirmation

"Next Friday"

Next | Friday

Automatic Speech Recognition

Flight booking

11/18/2016

Knowledge Graph

Natural Language Understanding

Grammar Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	11/18/2016

λ

"Your flight is booked for next Friday"

Confirmation

Polly

"Your flight is booked for next Friday"

"Next Friday"

Next

Friday

Automatic
Speech Recognition

Flight booking

11/18/2016

Knowledge
Graph

Natural Language
Understanding

Grammar
Graph

Utterances

Flight Booking	
Origin	Seattle
Destination	London Heathrow
Departure Date	11/18/2016

λ

Hotel Booking

Lex: Build Natural, Conversational Interactions In Voice & Text

High quality, through best-in-class deep learning

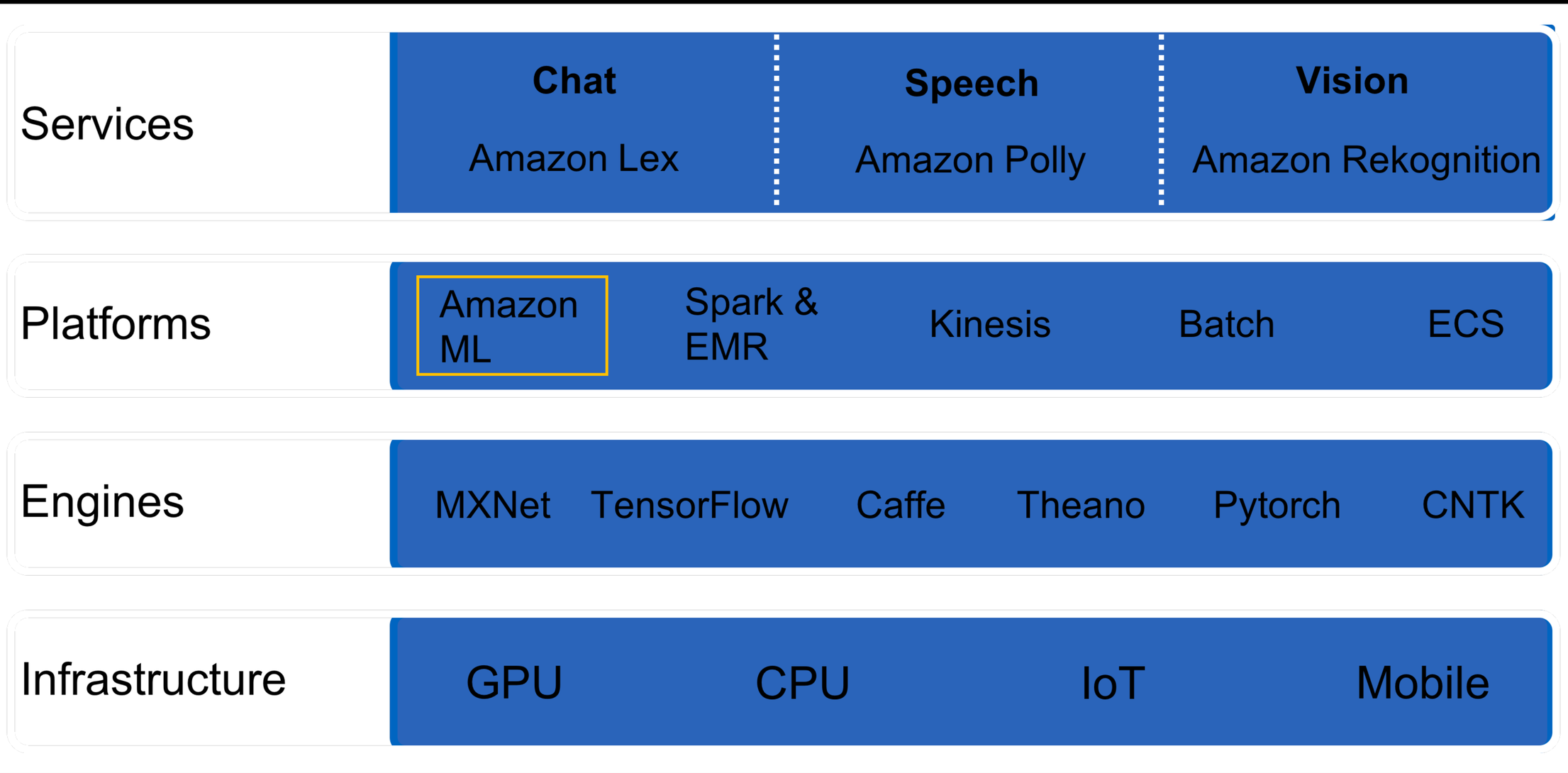
Deep functionality

Easy to use & thoughtfully integrated

Built for production

Low cost

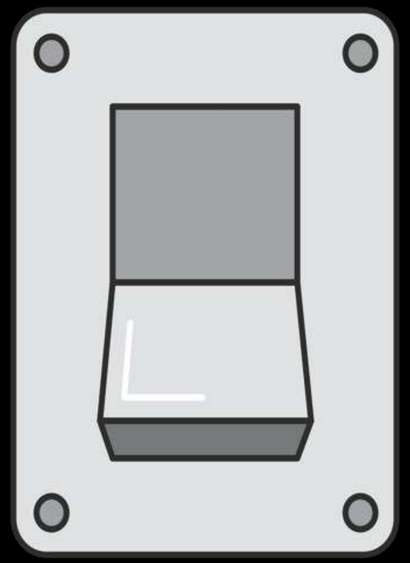
Amazon AI for every developer



Amazon Machine Learning

- Easy-to-use, managed machine learning service built for developers
- Robust, powerful machine learning technology based on Amazon's internal systems
- Create prediction and classification models using your data already stored in the AWS Cloud
- Deploy models to production in seconds

Fully managed model and prediction services



End-to-end service, with no servers to provision and manage

One-click production model deployment

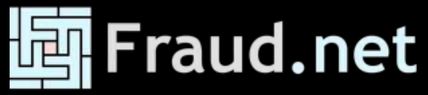
Programmatically query model metadata to enable automatic retraining workflows

Monitor prediction usage patterns with Amazon CloudWatch metrics

Fraud.net Uses AWS to Quickly, Easily Detect Online Fraud

> Amazon Machine Learning helps us reduce complexity and make sense of emerging fraud patterns.

Oliver Clark
CTO,
Fraud.net



Fraud.net is the world's leading crowdsourced fraud prevention platform.

- Needed to build and train a larger number of more targeted machine-learning models
- Uses Amazon Machine Learning to provide more than 20 models
- Easily builds and trains models to effectively detect online payment fraud
- Reduces complexity and makes sense of emerging fraud patterns
- Saves clients $1 million weekly by helping them detect and prevent fraud

Upserve Uses AWS to Help Restaurants Predict Business

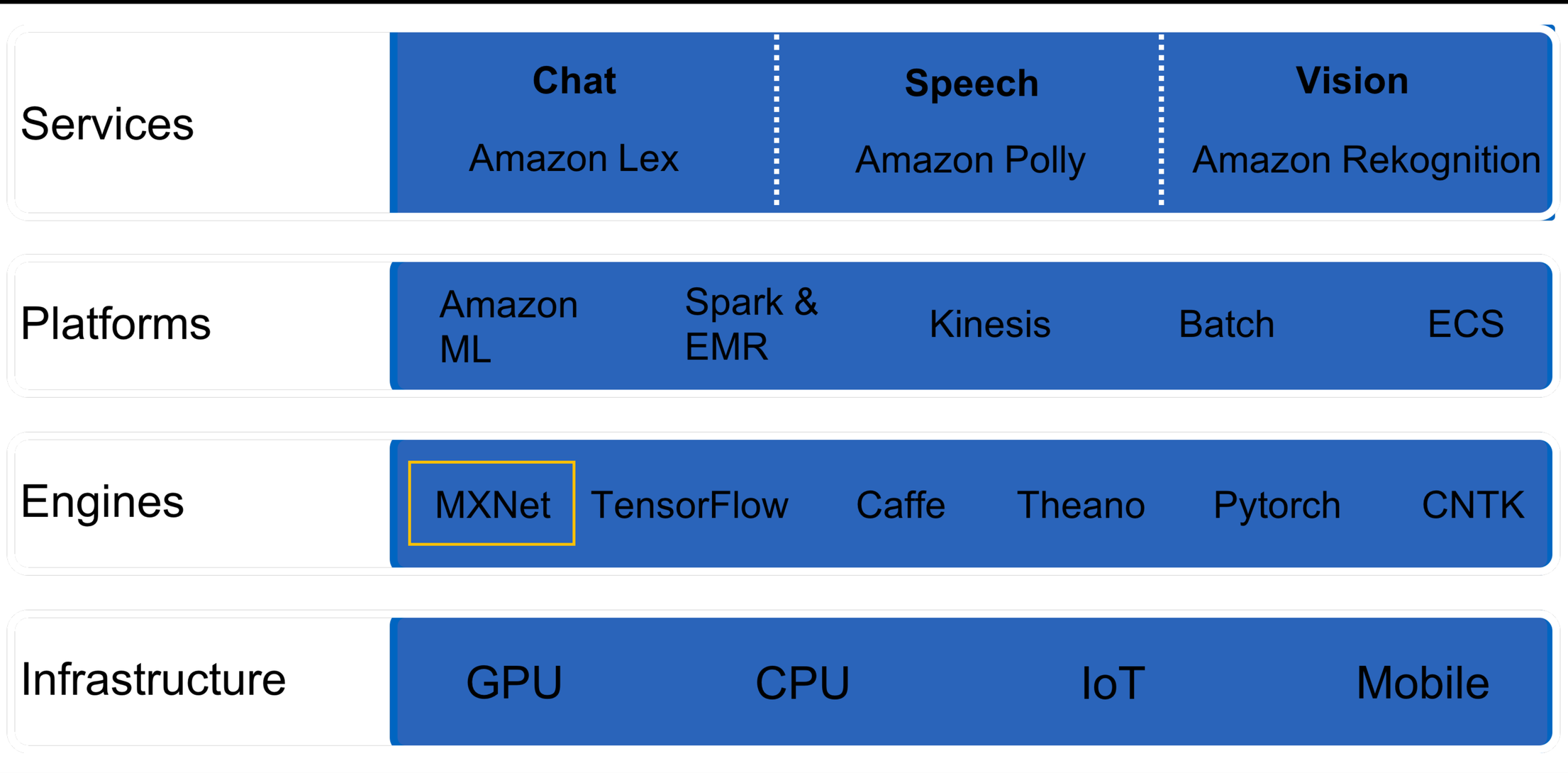
“Using Amazon Machine Learning, we can predict the total number of customers who will walk through a restaurant’s doors in a night.”

Bright Fulton
Director of Infrastructure Engineering,

Upserve provides online payment and analytical software to thousands of restaurant owners throughout the U.S.

- Needed its restaurant management platform to provide more predictive analytics
- Builds and trains more than 100 machine learning models weekly
- Streams restaurant sales and menu item data in real time
- Helps restaurateurs predict nightly business

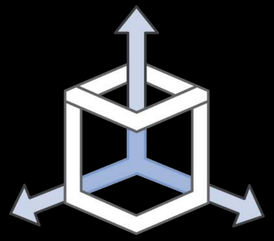
Amazon AI for every developer



Apache MXNet: Open Source library for Deep Learning

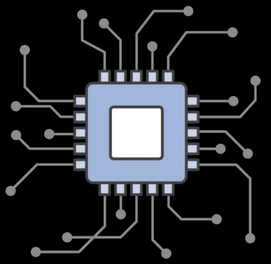
Programmable

Simple syntax,
multiple languages



Portable

Highly efficient
models for mobile
and IoT

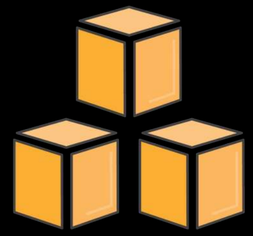


High Performance

Near linear scaling
across hundreds of GPUs

Most Open

Accepted into the
Apache Incubator



Best On AWS

Optimized for
Deep Learning on AWS

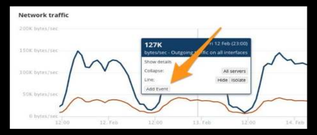
More information at
mxnet.io

mx.model.FeedForward

model.fit

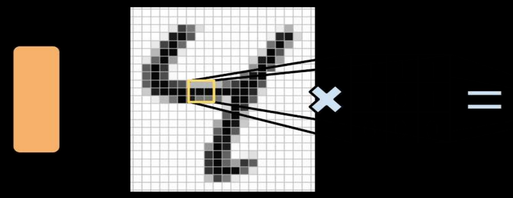
Input

Output



mx.sym.SoftmaxOutput

mx.sym.**FullyConnected**(data, num_hidden=128)



mx.sym.**Convolution**(data, kernel=(5,5), num_filter=20)

mx.sym.**Pooling**(data, pool_type="max", kernel=(2,2),

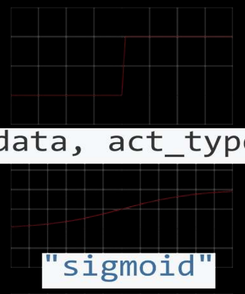
stride=(2,2)

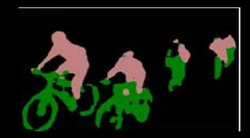
lstm.lstm_unroll(num_lstm_layer, seq_len, len, num_hidden, num_embed)

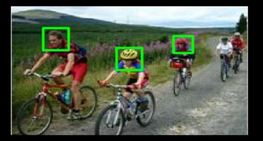
*cos(w, **queen**) = cos(w, **king**) - cos(w, **man**) + cos(w, **woman**)*

mx.symbol.**Embedding**(data, input_dim, output_dim = k)

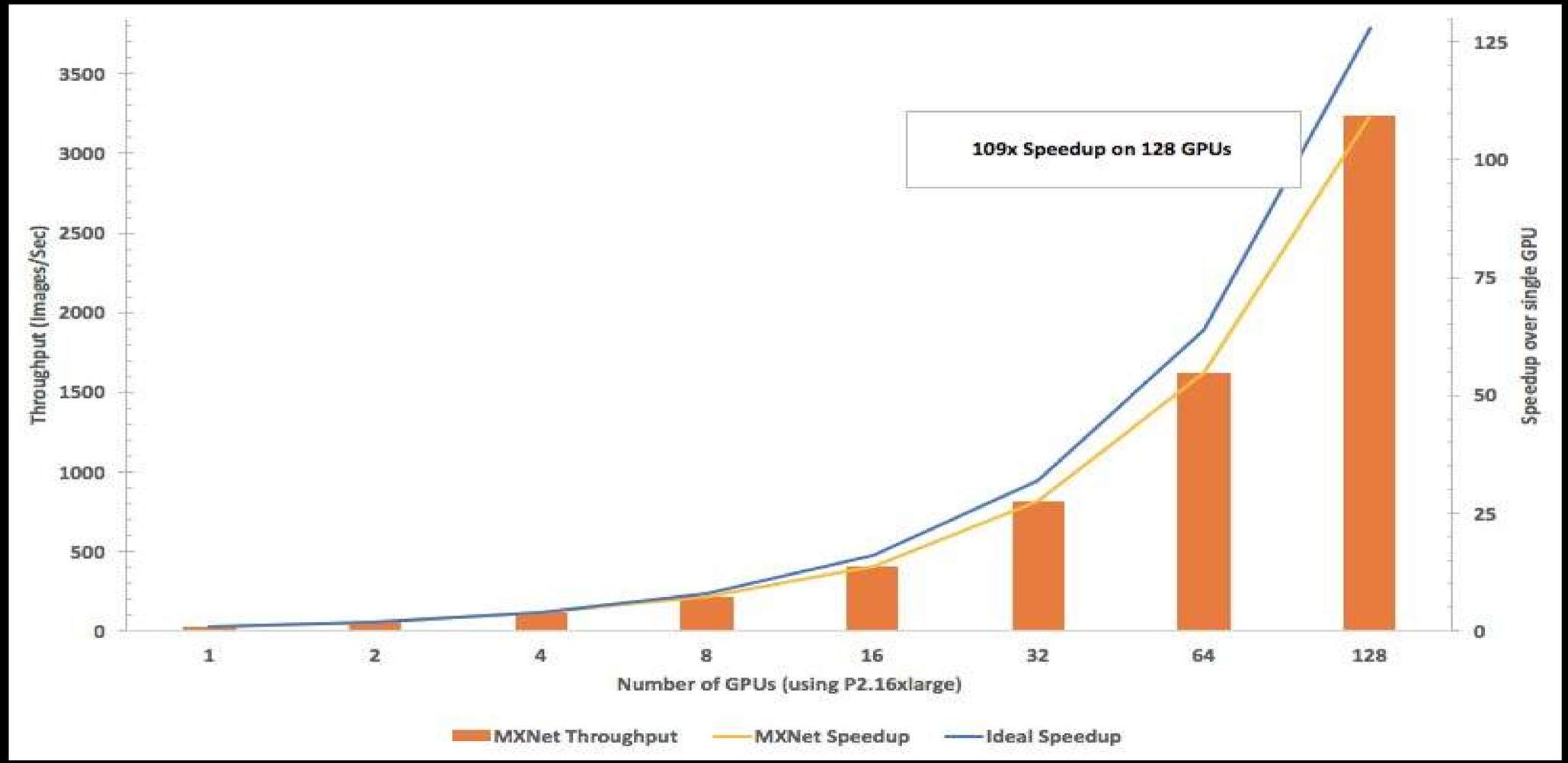
mx.sym.**Activation**(data, act_type="xxxx")







MXNet: Scalable Deep Learning Framework



AWS Deep Learning AMI

Up to~40k CUDA cores

Apache MXNet

TensorFlow

Theano

Caffe

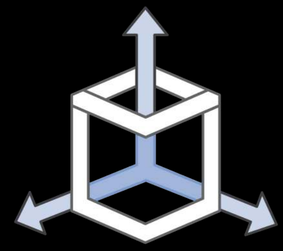
Torch

Pre-configured CUDA drivers

Anaconda, Python3

+ CloudFormation template

+ Container Image



One-Click GPU Deep Learning

Gluon: Deep Learning gets even easier

- Announced October 11th
- Available now in MXNet, soon in Microsoft Cognitive Toolkit
- Developer-friendly high-level API
- Dynamic networks can be modified during training
- No compromise on performance
- Extensive model zoo

https://github.com/gluon-api/
https://aws.amazon.com/blogs/aws/introducing-gluon-a-new-library-for-machine-learning-from-aws-and-microsoft/

Gluon Model Zoo

vgg11	VGG-11 model from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg13	VGG-13 model from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg16	VGG-16 model from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg19	VGG-19 model from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg11_bn	VGG-11 model with batch normalization from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg13_bn	VGG-13 model with batch normalization from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg16_bn	VGG-16 model with batch normalization from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
vgg19_bn	VGG-19 model with batch normalization from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.

VGG	VGG model from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.
get_vgg	VGG model from the "Very Deep Convolutional Networks for Large-Scale Image Recognition" paper.

resnet18_v1	ResNet-18 V1 model from "Deep Residual Learning for Image Recognition" paper.
resnet34_v1	ResNet-34 V1 model from "Deep Residual Learning for Image Recognition" paper.
resnet50_v1	ResNet-50 V1 model from "Deep Residual Learning for Image Recognition" paper.
resnet101_v1	ResNet-101 V1 model from "Deep Residual Learning for Image Recognition" paper.
resnet152_v1	ResNet-152 V1 model from "Deep Residual Learning for Image Recognition" paper.
resnet18_v2	ResNet-18 V2 model from "Identity Mappings in Deep Residual Networks" paper.
resnet34_v2	ResNet-34 V2 model from "Identity Mappings in Deep Residual Networks" paper.
resnet50_v2	ResNet-50 V2 model from "Identity Mappings in Deep Residual Networks" paper.
resnet101_v2	ResNet-101 V2 model from "Identity Mappings in Deep Residual Networks" paper.
resnet152_v2	ResNet-152 V2 model from "Identity Mappings in Deep Residual Networks" paper.

ResNetV1	ResNet V1 model from "Deep Residual Learning for Image Recognition" paper.
ResNetV2	ResNet V2 model from "Identity Mappings in Deep Residual Networks" paper.
BasicBlockV1	BasicBlock V1 from "Deep Residual Learning for Image Recognition" paper.
BasicBlockV2	BasicBlock V2 from "Identity Mappings in Deep Residual Networks" paper.
BottleneckV1	Bottleneck V1 from "Deep Residual Learning for Image Recognition" paper.
BottleneckV2	Bottleneck V2 from "Identity Mappings in Deep Residual Networks" paper.
get_resnet	ResNet V1 model from "Deep Residual Learning for Image Recognition" paper.

mobilenet1_0	MobileNet model from the "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" paper, with width multiplier 1.0.
mobilenet0_75	MobileNet model from the "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" paper, with width multiplier 0.75.
mobilenet0_5	MobileNet model from the "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" paper, with width multiplier 0.5.
mobilenet0_25	MobileNet model from the "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" paper, with width multiplier 0.25.

MobileNet	MobileNet model from the "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" paper.

densenet121	Densenet-BC 121-layer model from the "Densely Connected Convolutional Networks" paper.
densenet161	Densenet-BC 161-layer model from the "Densely Connected Convolutional Networks" paper.
densenet169	Densenet-BC 169-layer model from the "Densely Connected Convolutional Networks" paper.
densenet201	Densenet-BC 201-layer model from the "Densely Connected Convolutional Networks" paper.

DenseNet	Densenet-BC model from the "Densely Connected Convolutional Networks" paper.

inception_v3	Inception v3 model from "Rethinking the Inception Architecture for Computer Vision" paper.
Inception3	Inception v3 model from "Rethinking the Inception Architecture for Computer Vision" paper.

alexnet	AlexNet model from the "One weird trick..." paper.
AlexNet	AlexNet model from the "One weird trick..." paper.

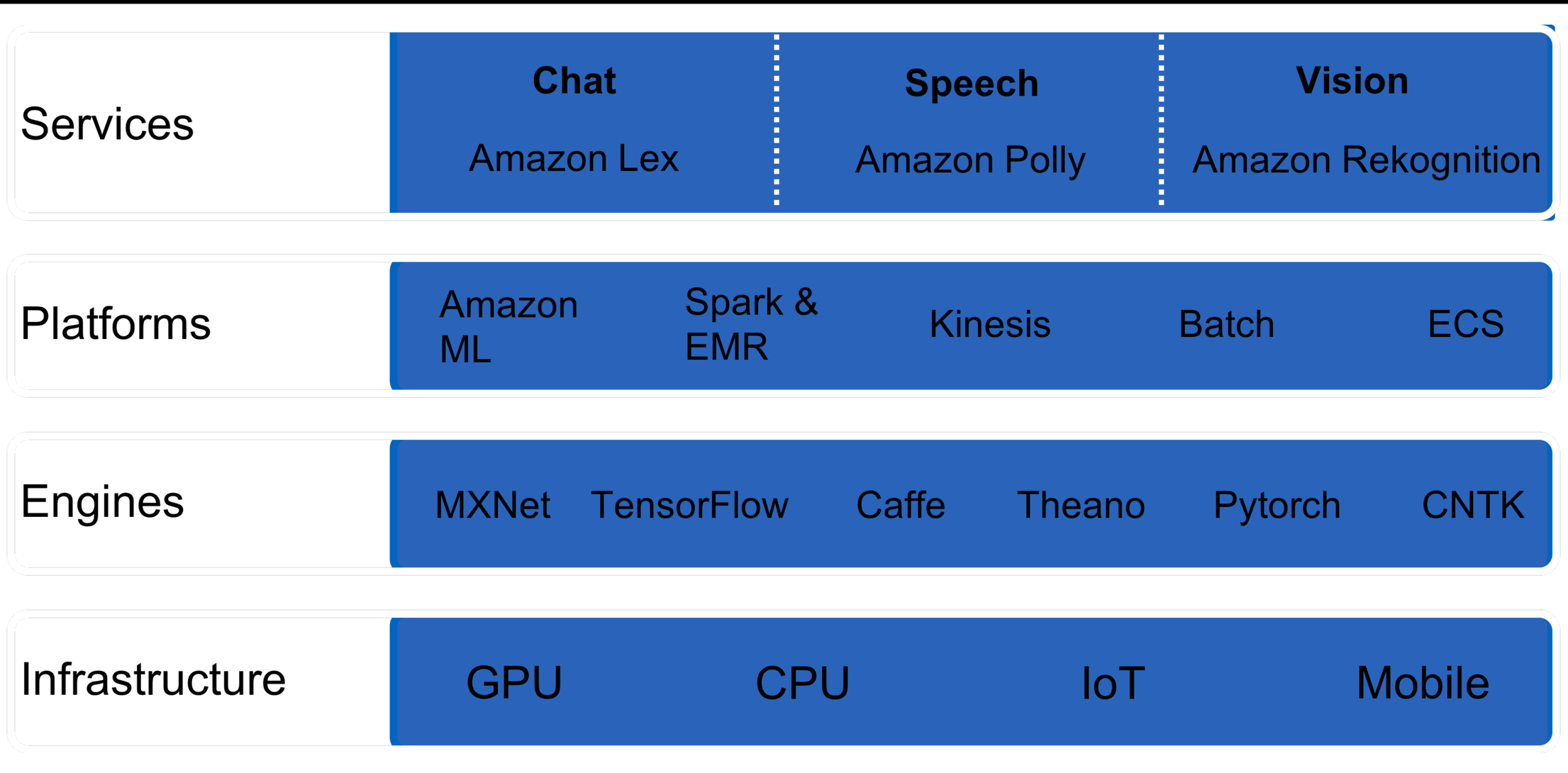
squeezenet1_0	SqueezeNet 1.0 model from the "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size" paper.
squeezenet1_1	SqueezeNet 1.1 model from the official SqueezeNet repo.

SqueezeNet	SqueezeNet model from the "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size" paper.

VGG
ResNet
AlexNet
DenseNet
SqueezeNet
Inception
MobileNet

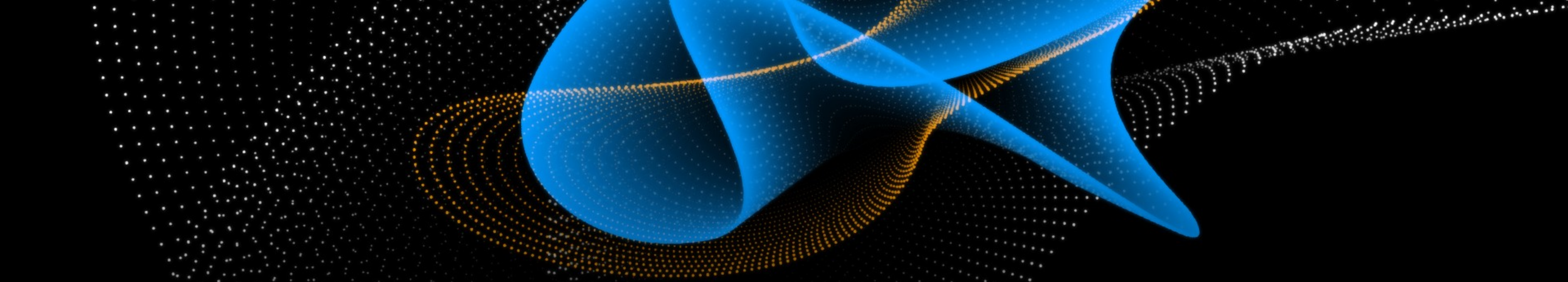
https://mxnet.incubator.apache.org/versions/master/api/python/gluon/model_zoo.html

Amazon AI for every developer



AWS is the Center of Gravity for Artificial Intelligence

http://aws.amazon.com/ai
https://reinvent.awsevents.com/



Thank you!

Julien Simon

AI Evangelist, EMEA

@julsimon
http://medium.com/@julsimon