



Machine Learning on AWS

Julien Simon

Principal Technical Evangelist, AI and Machine Learning

@julsimon



Welcome to Amazon.com Books!

*One million titles,
consistently low prices.*

(If you explore just one thing, make it our personal notification service. We think it's very cool!)

SPOTLIGHT! -- AUGUST 16TH

These are the books we love, offered at Amazon.com low prices. The spotlight moves **EVERY** day so please come often.

ONE MILLION TITLES

Search Amazon.com's million title catalog by author, subject, title, keyword, and more... Or take a look at the books we recommend in over 20 categories... Check out our customer reviews and the award winners from the Hugo and Nebula to the Pulitzer and Nobel... and bestsellers are 30% off the publishers list...

EYES & EDITORS. A PERSONAL NOTIFICATION SERVICE

Like to know when that book you want comes out in paperback or when your favorite author releases a new title? Eyes, our tireless, automated search agent, will send you mail. Meanwhile, our human editors are busy previewing galleys and reading advance reviews. They can let you know when especially wonderful works are published in particular genres or subject areas. Come in, meet Eyes, and have it all explained.

YOUR ACCOUNT

Check the status of your orders or change the email address and password you have on file with us. Please note that you **do not** need an account to use the store. The first time you place an order, you will be given the opportunity to create an account.

Amazon.com, 1995

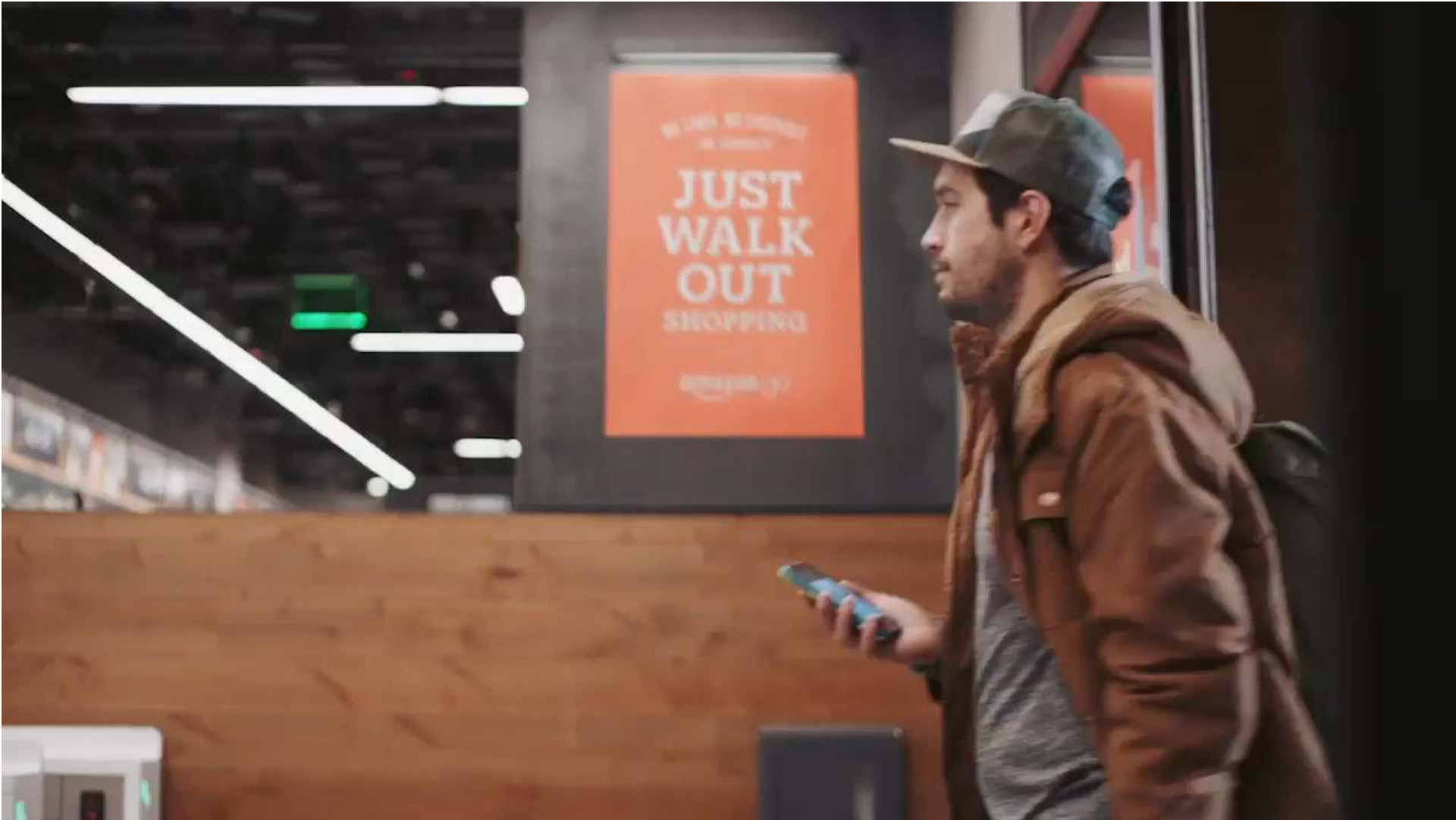


amazon
fulfillment

amazonrobotics



amazon



NO CASHIERS.
NO CHECKOUTS.
NO BAGGERS.
**JUST
WALK
OUT**
SHOPPING

amazon

amazon echo



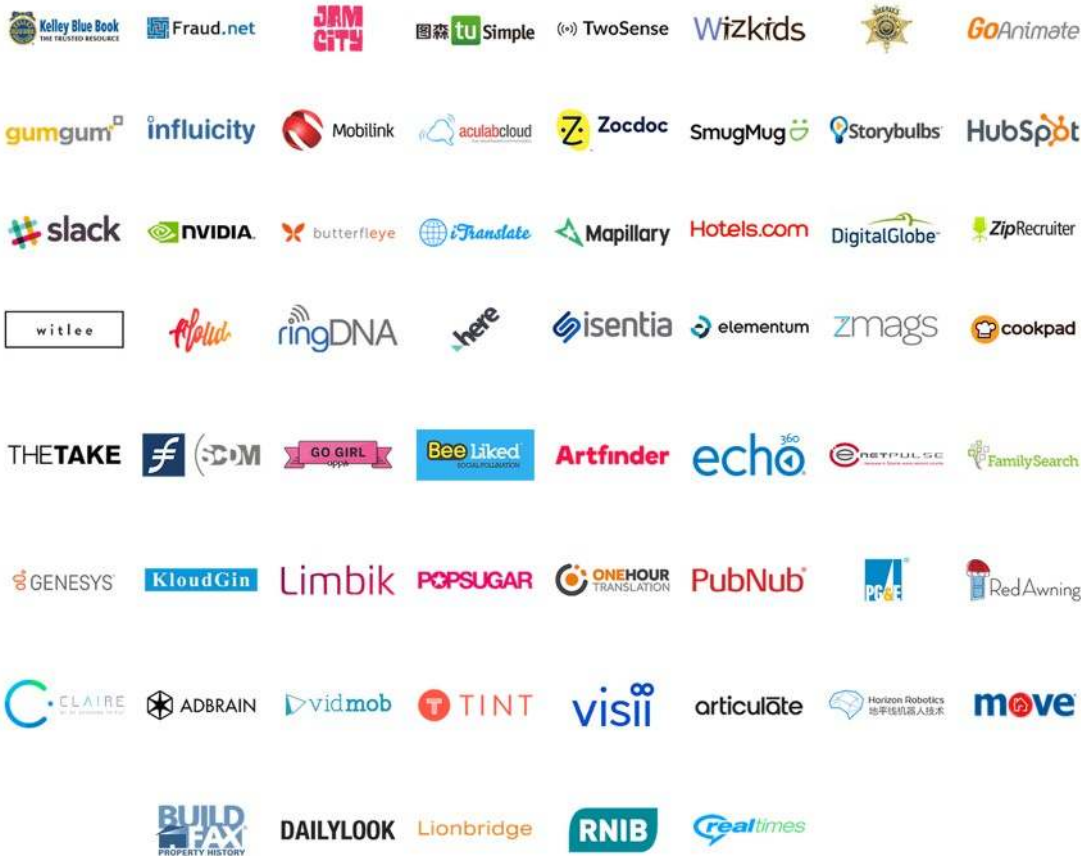
Our mission

Put Machine Learning in the hands
of every developer and data scientist

More ML is built on AWS than anywhere else



More ML is built on AWS than anywhere else



AWS ML Stack

Application Services

API-driven services: Vision, Language & Speech Services, Chatbots

Platform Services

Deploy machine learning models with high-performance machine learning algorithms, broad framework support, and one-click training, tuning, and inference.

Frameworks & Infrastructure

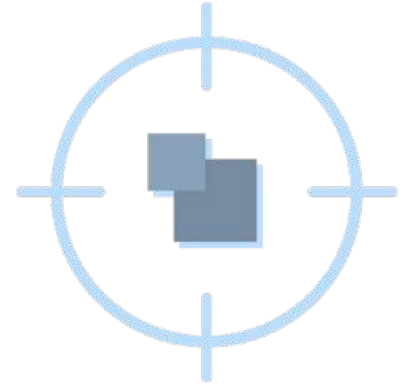
Develop sophisticated models with any framework, create managed, auto-scaling clusters of GPUs for large scale training, or run prediction on trained models.



Application Services

Amazon Rekognition

Deep Learning-based image analysis service



Object & Scene Detection




City	91.4%
Downtown	91.4%
Metropolis	91.4%
Urban	91.4%
Building	56.2%
High Rise	56.2%
Dock	55.6%
Pier	55.6%
Dawn	52.4%
Dusk	52.4%

Facial Analysis





looks like a face	99.8%
appears to be female	100%
age range	23 - 38 years old
smiling	99.4%
appears to be happy	93.2%
wearing eyeglasses	99.9%
wearing sunglasses	97.6%

Crowd Detection – up to 100 faces




Choose a sample image



Use your own image


Image must be .jpeg or .png format and no larger than 5MB. Your image isn't stored.

 **Upload** or drag and drop

Go

Done with the demo?
[Learn more](#)

▼ Results



looks like a face	99.9 %
appears to be male	99.9 %
age range	26 - 43 years old
smiling	98.5 %
appears to be happy	99.7 %
not wearing glasses	99.9 %

[Show more](#)

Facial Search



=



Similarity

97%



≠



≠



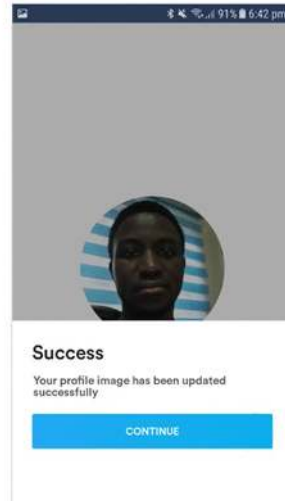
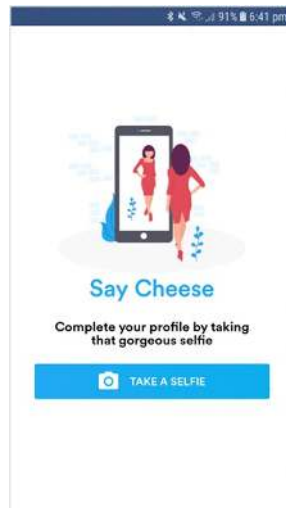
Aella Credit



Aella Credit is a financial services company based in West Africa that provides **instant loans to individuals** by using biometric and employer data.

In emerging markets, identity verification and validation is a major challenge for people who don't have easy access to retail banking services.

Aella Credit uses **Amazon Rekognition**, for biometric identity verification on their mobile application.



<https://aws.amazon.com/blogs/machine-learning/aella-credit-empowers-underbanked-individuals-by-using-amazon-rekognition-for-identity-verification/>

Marinus Analytics

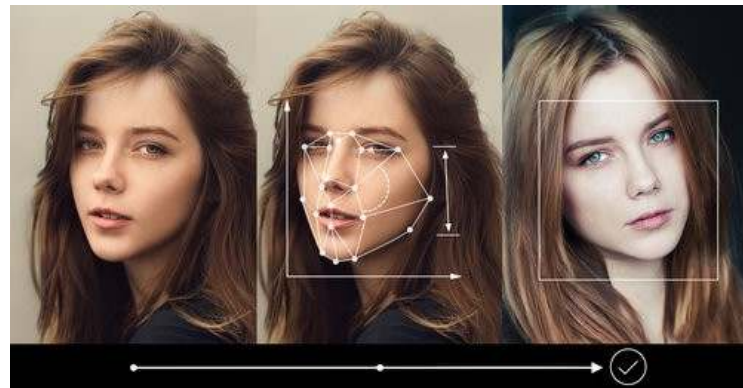


Marinus Analytics provides law enforcement with tools founded in artificial intelligence.

Traffic Jam, is a suite of tools for use by law enforcement agencies on sex trafficking investigations.

Before using **Amazon Rekognition**, their only recourse was manual processing; this was time-intensive or not possible.

Now, investigators are able to take effective action by searching through **millions of records in seconds** to find victims.



<http://www.marinusanalytics.com/articles/2017/10/17/amazon-rekognition-helps-marinus-analytics-fight-human-trafficking>

Image Moderation



▼ Results

Suggestive	83.5%
Female Swimwear Or Underwear	83.5%

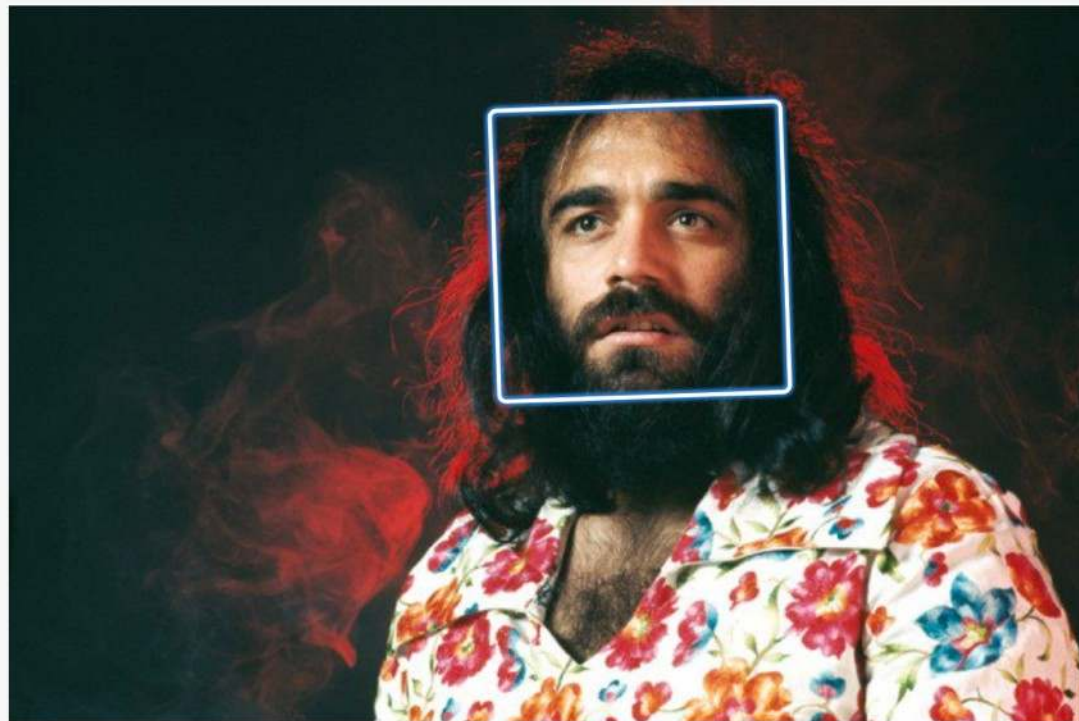
Explicit Nudity

- Nudity
- Graphic Male Nudity
- Graphic Female Nudity
- Sexual Activity
- Partial Nudity

Suggestive

- Female Swimwear or Underwear
- Male Swimwear or Underwear
- Revealing Clothes

Celebrity Recognition



Done with the demo?

[Learn more](#)

▼ Results



Demis Roussos

[Learn More](#)

Match confidence

99 %

► Request

► Response

Text in Image



▼ Results

US English only

| 5T6E652 |

► Request

▼ Response

```
{
  "TextDetections": [
    {
      "Confidence": 96.38984680175781,
      "DetectedText": "5T6E652",
      "Geometry": {
        "BoundingBox": {
          "Height": 0.05606506019830704,
          "Left": 0.38905566930770874,
          "Top": 0.7383729219436646,
          "Width": 0.10844030976295471
        },
        "Polygon": [
          {
            "X": 0.38905566930770874,
            "Y": 0.7383729219436646
          },
          {
            "X": 0.49749597907066345,
            "Y": 0.7726403474807739
          },
          {
            "X": 0.4896218776702881,
            "Y": 0.8287054300308228
          },
          {
            "X": 0.38905566930770874,
            "Y": 0.7383729219436646
          }
        ]
      }
    }
  ]
}
```

Amazon Rekognition – latest features

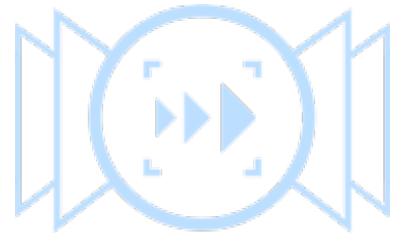
August 6 – Text in Image: **improved accuracy** and support for **rotated text** (from -90° to +90°)

August 23 – Face collections: new **DescribeCollection** API to simplify management of collections

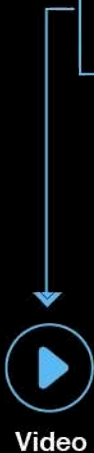
September 18 – Face collections: **filter the quality and quantity of faces** indexed for face recognition (manual threshold or automatic)

Amazon Rekognition Video

Deep Learning-based video analysis service



Video Analysis



Rekognition

Context

Spatial

Temporal

Auditory

Features

Objects, Scenes and Concepts



Facial Analysis



Female: 99.9%
Eyes are open: 98.4%
Calm: 86.0%

Face Recognition



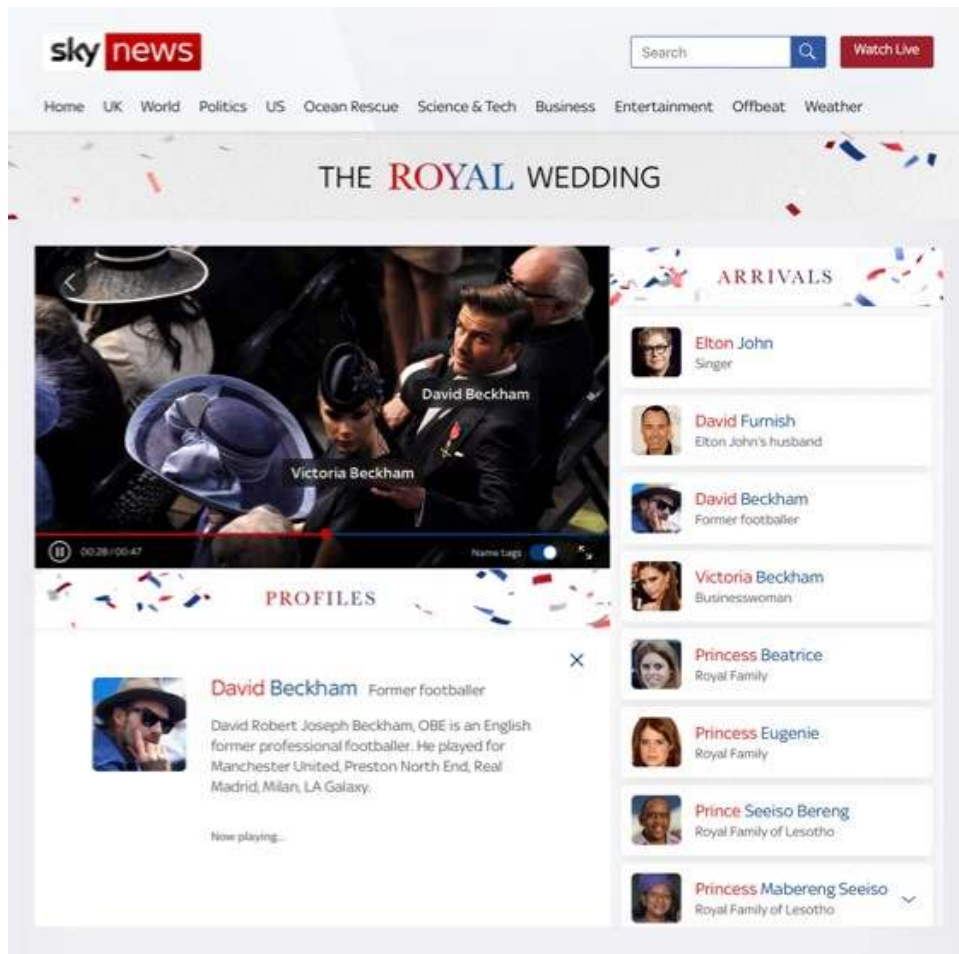
Person 1

Activity Detection

Sports, Running

Object Tracking





Sky News used **Amazon Rekognition** to perform **real-time identification of guests** as they enter St. George's Chapel

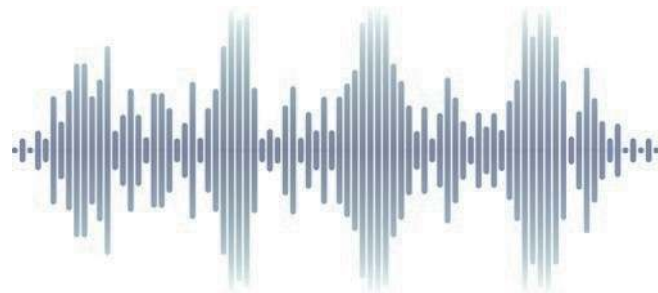
Amazon Polly

Deep Learning-based text-to-speech service

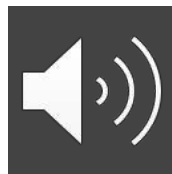


Amazon Polly: Text In, Life-like Speech Out

“Salut, je
m’appelle Léa. Je
suis la nouvelle
voix française de
Polly.”



Amazon Polly



A Focus On Voice Quality & Pronunciation

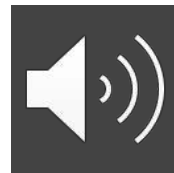
Support for Speech Synthesis Markup Language (SSML) Version 1.0

<https://www.w3.org/TR/speech-synthesis>

<speaking xml:lang="en-US">

The price of this book is <prosody rate="60%">45€</prosody>

</speaking>



Breathing



<https://aws.amazon.com/blogs/machine-learning/amazon-polly-releases-new-ssml-breath-feature/>

<amazon:auto-breaths>

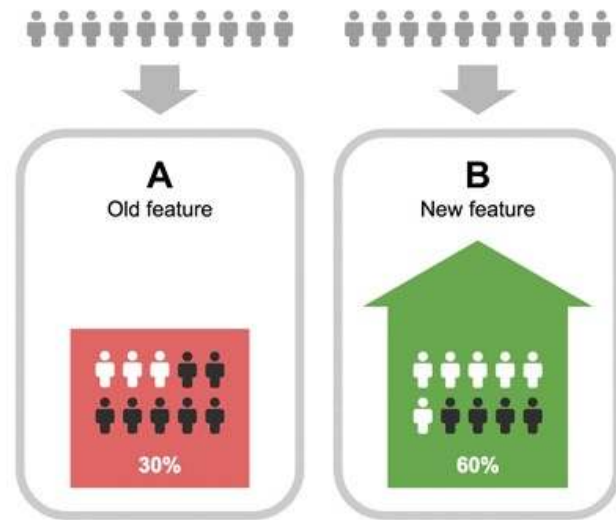
Duolingo



Duolingo is the most popular language-learning platform and the most downloaded education app in the world, with more than **170 million users**.

They have run six A/B tests, testing an **Amazon Polly** voice against a voice from other TTS providers.

For all of these experiments, the winning condition was the Amazon Polly voice



<https://aws.amazon.com/blogs/machine-learning/powering-language-learning-on-duolingo-with-amazon-polly/>

Amazon Polly – latest features

July 17 – **Input limit** now 100K, store **output files** in S3

August 2 – New voice: Aditi, **bilingual Indian/English**

September 14 – New voice: Zhiyu, **Mandarin Chinese**

54 voices, 27 languages

Amazon Translate

Neural Machine Translation Service

Now generally available



FREE TIER

Amazon Translate

Natural and fluent language translation



Real-time
translation



Batch analysis



Automatic language
recognition



Low cost

Languages

English to/from:

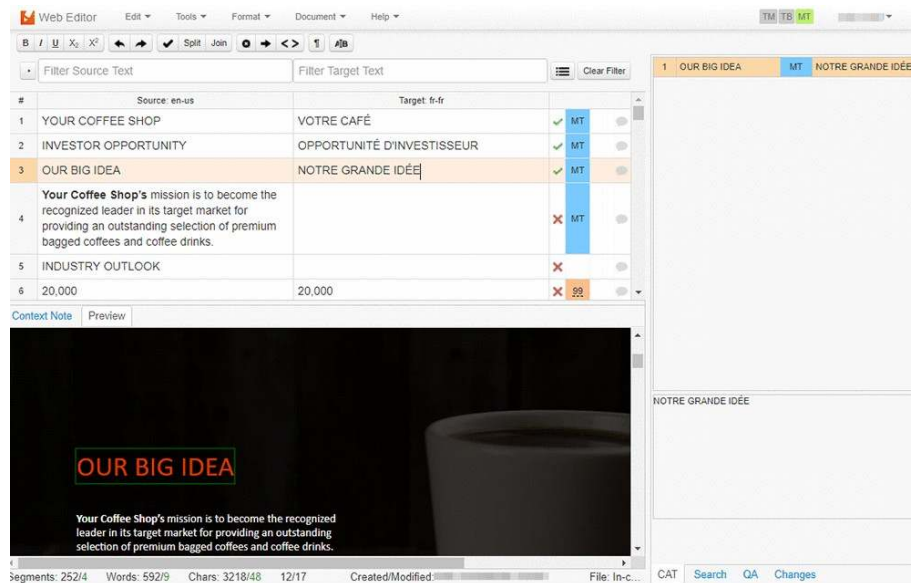
- Spanish
- Portuguese
- German
- French
- Arabic
- Simplified Chinese
- Japanese
- Russian
- Italian
- Traditional Chinese
- Turkish
- Czech

Coming soon : Danish, Dutch, Finnish, Hebrew, Polish, and Swedish

Memsourse was founded in 2010 and has become a leader in **cloud-based translation technology**.

Memsourse and **Amazon Translate** are connected via the Amazon Translate API.

Results are available in **real-time** from within the Memsourse translation workbenches, Memsourse Web Editor, and Desktop Editor.



<https://aws.amazon.com/blogs/machine-learning/amazon-translate-now-available-in-the-memsourse-translation-management-system/>

Amazon Transcribe

Automatic speech recognition service

Now generally available



FREE TIER

Amazon Transcribe

Automatic conversion of speech into accurate, grammatically correct text



Multiple
languages



Intelligent punctuation
and formatting



Timestamp
generation



Support for
telephony audio



Recognize multiple
speakers



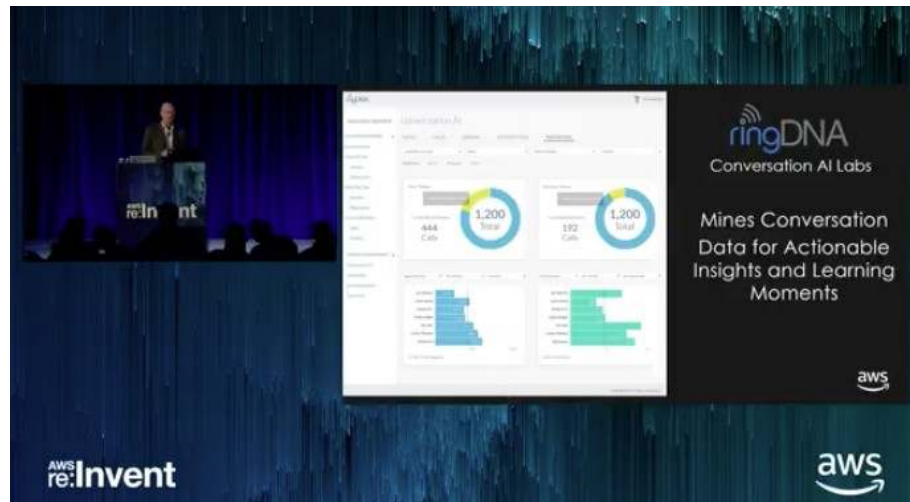
Custom
vocabulary

RingDNA is an end-to-end communications platform for sales teams. Hundreds of enterprise organizations use RingDNA to increase productivity, engage in smarter sales conversations, gain predictive insights and improve their win rate.

Speech to Text

*"A critical component of RingDNA's Conversation AI requires best of breed speech-to-text to deliver transcriptions of every phone call. RingDNA is excited about **Amazon Transcribe** since it provides high-quality speech recognition at scale, helping us to better transcribe every call to text"*

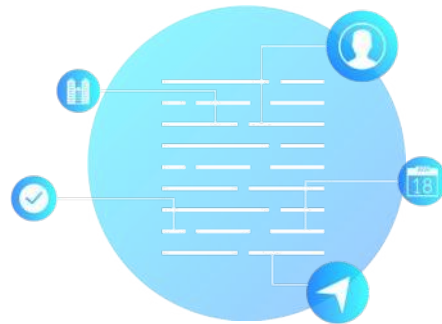
Howard Brown, CEO & Founder, RingDNA



https://www.youtube.com/watch?v=1ZJ_f1bDdog

Amazon Comprehend

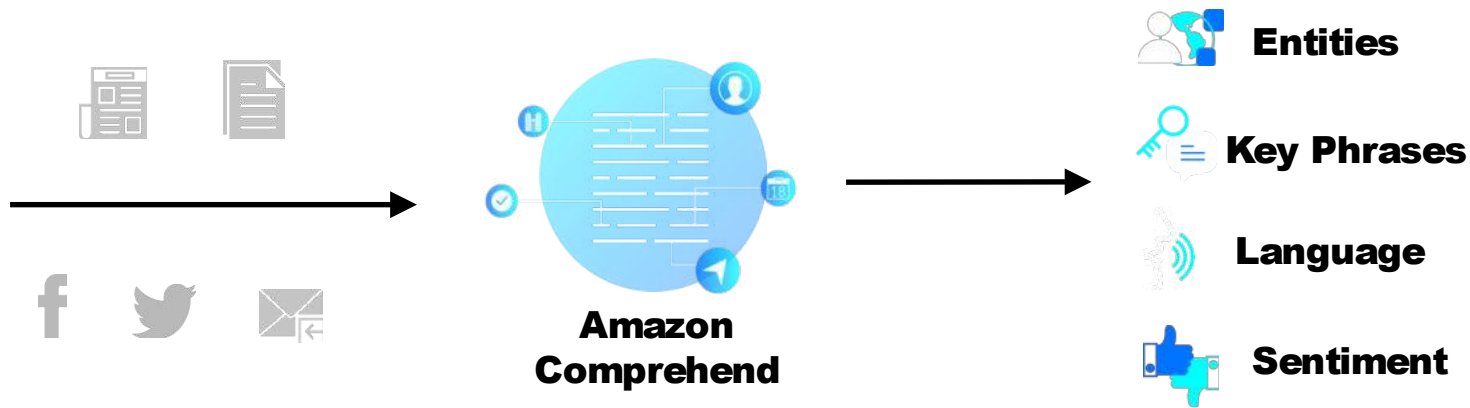
Natural Language Processing



FREE TIER

Fully managed natural language processing

Discover valuable insights from text



ClearView Social



**CLEAR
VIEW
SOCIAL**

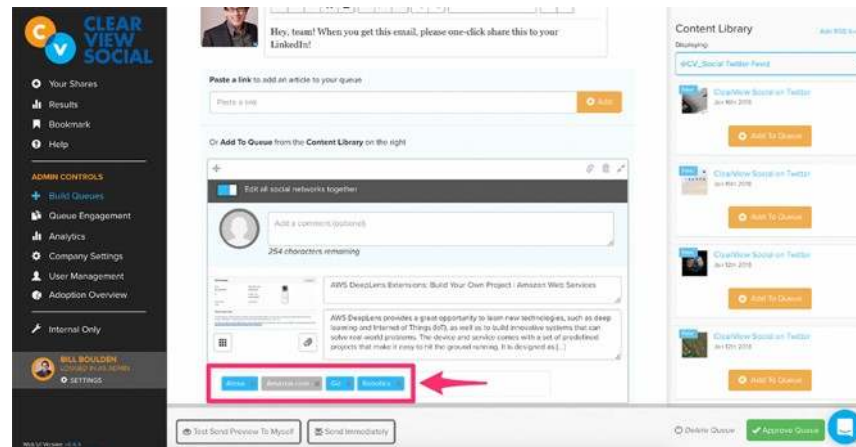
ClearView Social enables a company's employees to share approved content on LinkedIn, Twitter, and other social networks.

To eliminate the reliance on manual tagging, ClearView Social turned to **Amazon Comprehend** to find **insights** and **relationships** in text.

Natural Language Processing

"We use **Amazon Comprehend** to read an article and extract topics, which are automatically tagged using machine learning. This automatic tagging helps customers easily estimate the market value of their engagement according to the current bid prices from the Google AdWords API"

Bill Boulden, CTO, ClearView Social



<https://aws.amazon.com/blogs/machine-learning/clearview-social-uses-amazon-comprehend-to-measure-the-impact-of-social-sharing/>

Topic modeling



**Amazon
Comprehend**

STORM



WORLD
SERIES



STOCK
MARKET



WASHINGTON



AUSTRALIA HEALTH

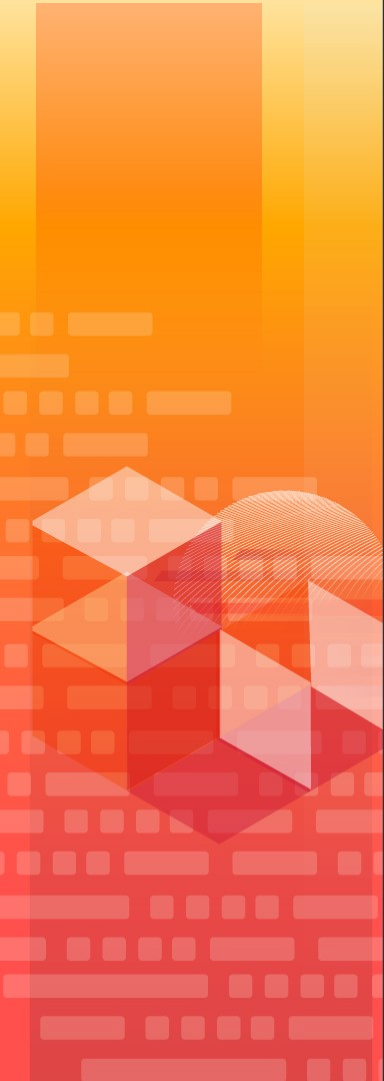


CRISIS



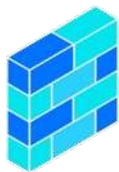
MACHINE
LEARNING





Platform Services

ML is still too complicated for everyday developers



Collect and prepare
training data



Choose and
optimize your ML
algorithm



Set up and manage
environments for
training



Train and tune model
(trial and error)



Deploy model
in production



Scale and manage
the production
environment

Amazon SageMaker

Easily build, train, and deploy Machine Learning models



Collect and prepare
training data



Choose and
optimize your ML
algorithm



Set up and manage
environments for
training



Train and tune model
(trial and error)



Deploy model
in production



Scale and manage
the production
environment

Amazon SageMaker



Pre-built
notebooks for
common
problems



Built-in, high-
performance
algorithms

ALGORITHMS

K-Means Clustering
Principal Component Analysis
Neural Topic Modelling
Factorization Machines
Linear Learner

XGBoost
Latent Dirichlet Allocation
Image Classification
Seq2Seq,
And more!

FRAMEWORKS

Apache MXNet,
Chainer
TensorFlow, PyTorch

Caffe2, CNTK,
Torch



Set up and manage
environments for
training



Train and tune
model (trial and
error)



Deploy model
in production



Scale and manage the
production environment

Build

Amazon SageMaker



Pre-built
notebooks for
common
problems



Built-in, high-
performance
algorithms

Build



One-click
training



Hyperparameter
optimization

Train



Deploy model
in production



Scale and manage
the production
environment

Amazon SageMaker



Pre-built
notebooks for
common
problems



Built-in, high-
performance
algorithms



One-click
training



Hyperparameter
optimization



One-click
deployment



Fully managed
hosting with auto-
scaling

Build

Train

Deploy

FREE TIER

Selected Amazon SageMaker customers



THOMSON REUTERS



DOW JONES



Digital Globe

<https://aws.amazon.com/solutions/case-studies/digitalglobe-machine-learning/>



- Operating **Earth imaging satellites** and providing **image analysis** services.
- Over **100 PB of imagery**.
- Extensive use of Machine Learning on **SageMaker** to extract information from images.
- Working with the AWS ML Lab, built a predictive model reducing cloud storage costs by **50%**.



USING AMAZON SAGEMAKER TO CUT CLOUD STORAGE COSTS IN HALF



Amazon SageMaker – latest algo features

June 7 – Hyper parameter optimization generally available

July 12 – two new algos: k-Nearest-Neighbor, object detection (SSD)

July 13 – improvements to DeepAR, Word2Vec and Linear Learner (multi-class classification)

Amazon SageMaker – latest infra features

July 17 – **Pipe mode** supported for Tensorflow training jobs

July 17 – **Batch transform** for non-real time prediction

August 14 – SageMaker APIs now fully supported on **AWS PrivateLink**

September 4 – New **custom header** for the InvokeEndPoint API



Frameworks & Infrastructure


Amazon EC2 C5 / C5d instances

C5: Next Generation
Compute-Optimized
Instances with
Intel® Xeon® Scalable
Processor


AWS Compute optimized instances support the new Intel® AVX-512 advanced instruction set, enabling you to more efficiently run vector processing workloads with single and double floating point precision, such as AI/machine learning or video processing.

*25% improvement in
price/performance over C4*


C4

36 vCPUs
 "Haswell"
4 Gbps to EBS
60 GiB memory

C5

72 vCPUs 2X vCPUs
 "Skylake"
12 Gbps to EBS 3X throughput
144 GiB memory 2.4X memory

AVX 512
2X performance



Amazon EC2 P3 Instances

The fastest, most powerful GPU instances in the cloud

- P3.2xlarge, P3.8xlarge, P3.16xlarge
- Up to eight NVIDIA Tesla V100 GPUs in a single instance
 - 40,960 CUDA cores, 5120 Tensor cores
 - 128GB of GPU memory
- 1 PetaFLOPs of computational performance – *14x better than P2*
- 300 GB/s GPU-to-GPU communication (NVLink) – *9x better than P2*

AWS Deep Learning AMIs

Preconfigured environments to quickly build Deep Learning applications

Conda AMI

For developers who want pre-installed pip packages of DL frameworks in separate virtual environments.

Base AMI

For developers who want a clean slate to set up private DL engine repositories or custom builds of DL engines.

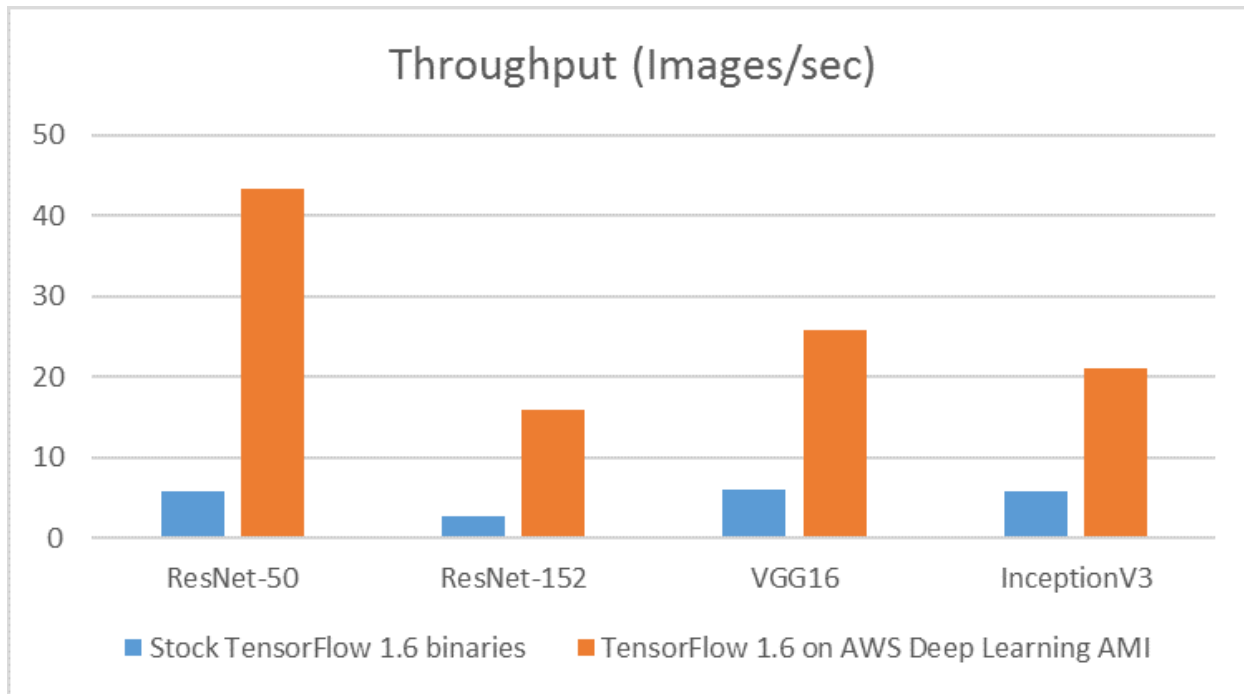
AMI with source code

For developers who want preinstalled DL frameworks and their source code in a shared Python environment.



<https://aws.amazon.com/machine-learning/amis/>

Faster TensorFlow training on C5



<https://aws.amazon.com/blogs/machine-learning/faster-training-with-optimized-tensorflow-1-6-on-amazon-ec2-c5-and-p3-instances/>

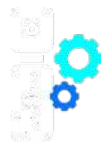
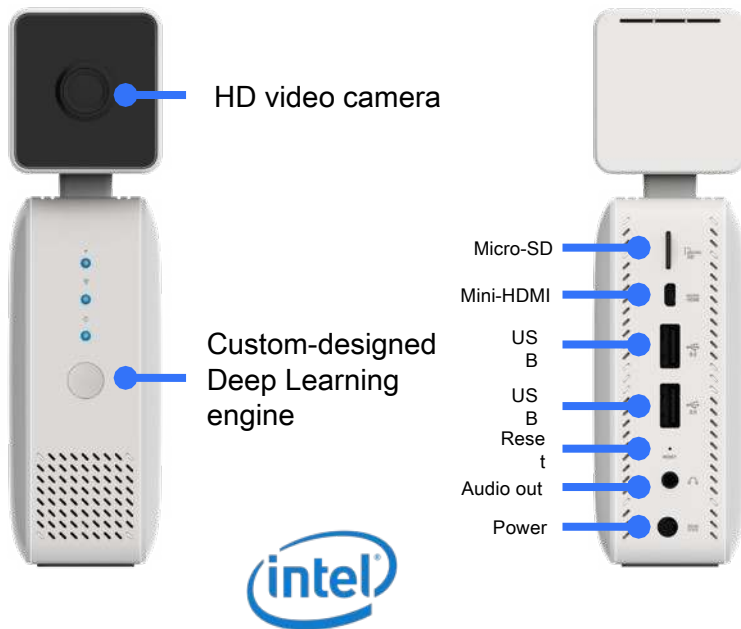
Autonomous driving systems with Apache MXNet



Can we do more to put ML in the
hands of all developers (literally)?

AWS DeepLens

The world's first Deep Learning-enabled video camera for developers



HD video camera
with on-board
compute optimized
for Deep Learning



Integrates with Amazon
SageMaker and AWS
Lambda



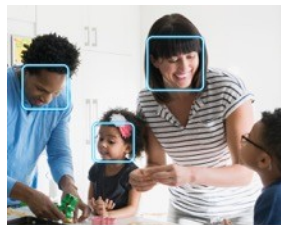
From unboxing
to prediction in
<10 minutes



Tutorials, examples,
demos, and pre-built
models

Get Started with Deep Learning

It takes less than 10 minutes with AWS DeepLens



FACIAL DETECTION
Detect faces of people.



ACTIVITY RECOGNITION
Recognize more than 50 kinds of actions such as brushing teeth, applying lipstick, and playing guitar.



HOT DOG NOT HOT DOG
A hot dog or not a hot dog, that is the question.



OBJECT DETECTION
Accurately detect and recognize objects.



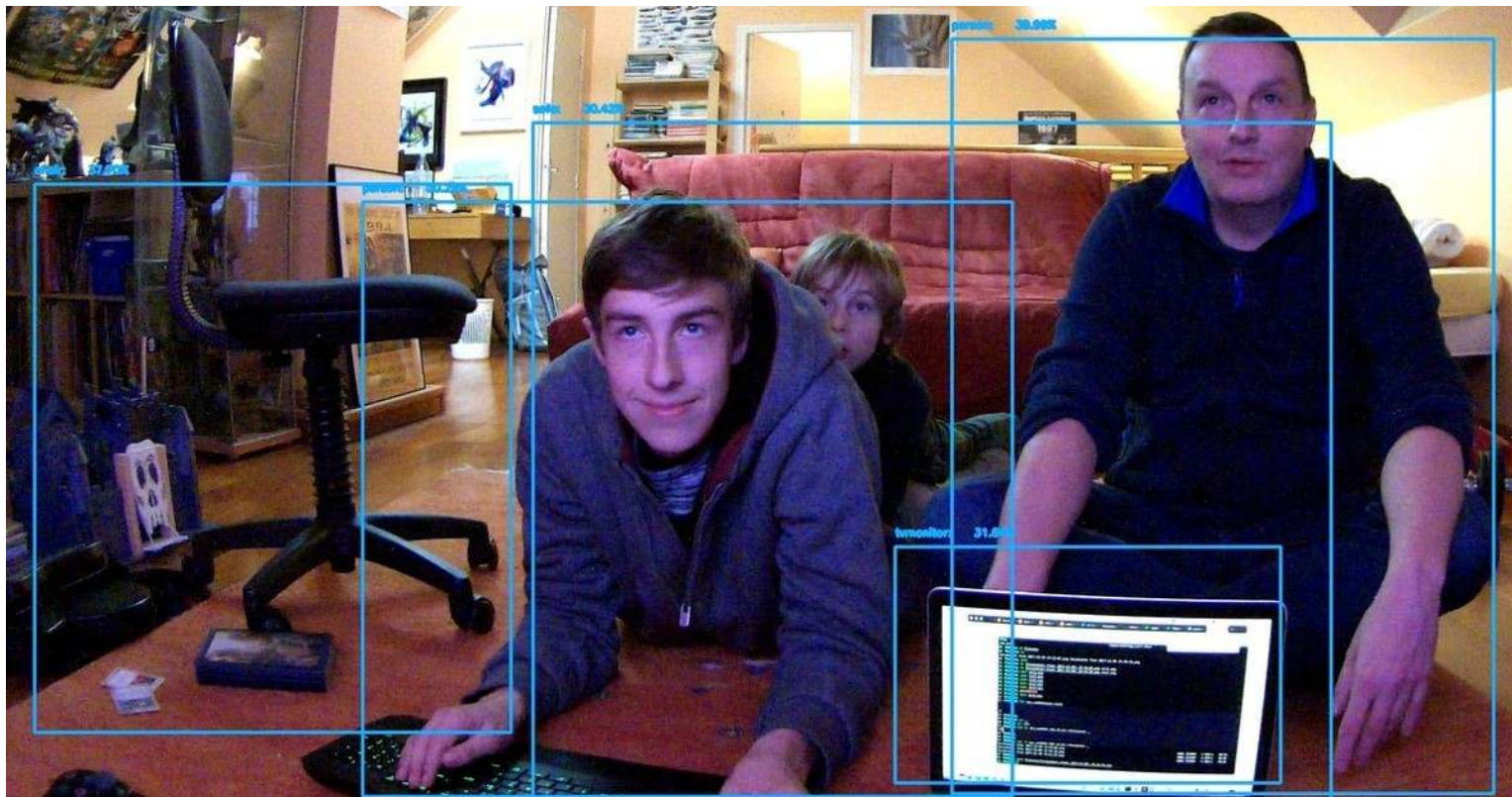
ARTISTIC STYLE TRANSFER
Transfer the style from one image, such as a painting, to an entire video sequence captured by your AWS DeepLens in real-time.



CAT AND DOG RECOGNITION
Detect whether it's a cat or dog.

Build custom Deep Learning models in the cloud using Amazon SageMaker, or use the collection of pretrained models included with AWS DeepLens

Object detection with AWS DeepLens



Democratization of ML

APPLICATION SERVICES



**Amazon
Rekognition**



**Amazon
Polly**



**Amazon
Lex**

**Amazon
Comprehend**



**Amazon
Rekognition
Video**



**Amazon
Transcribe**



**Amazon
Translate**

PLATFORM SERVICES

**Amazon
SageMaker**

AWS DeepLens

FRAMEWORKS AND INTERFACES

**Deep
Learning AMI**

Apache
MXNet

Caffe2

CNTK

PyTorch

TensorFlow

Theano

Torch

Keras

Gluon



Thank you!

<https://ml.aws>

<https://aws.amazon.com/blogs/machine-learning>

<https://medium.com/@julsimon>

<https://youtube.com/juliensimonfr>

Julien Simon

Principal Technical Evangelist, AI and Machine Learning

@ **julsimon**