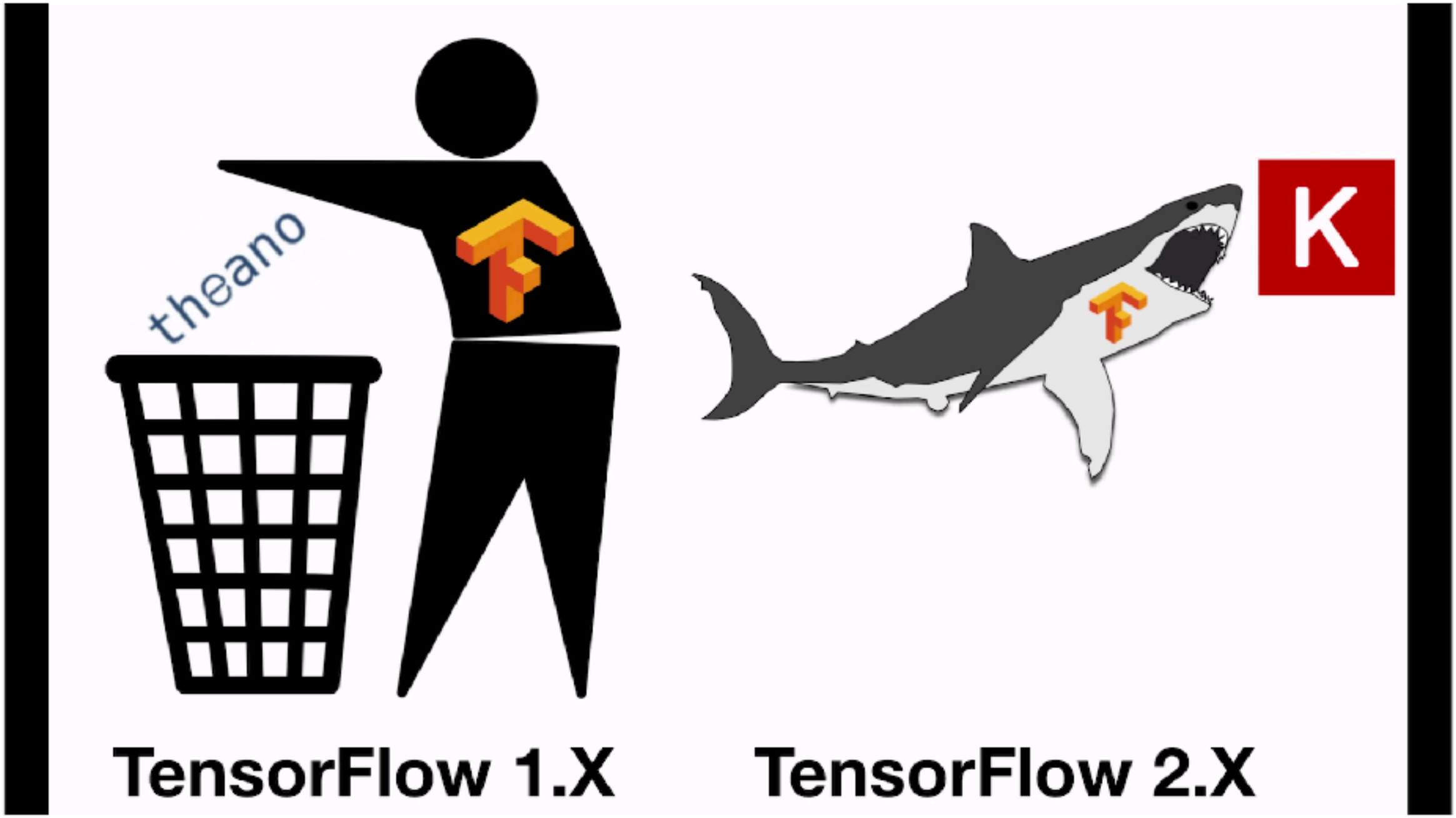
Trends in AI, ML, DL

romeo kienzler, IBM
Casablanca, 5th of March ‘19

YouTube

Search



How TensorFlow 2.0 is changing everything (eating Keras, kicking Theano)

118 views

3 3 SHARE SAVE ...

Romeo Kienzler
Published on Mar 20, 2019

SUBSCRIBE 1.2K

François Chollet
@fchollet

Follow

When there's too much hype around something, it can lead people to become cynical about it and assume it's worthless.

But cynicism is not wisdom. It's laziness. If you think you already know all the answers, there's a lot you're going to miss. Keep an open mind and keep learning

5:07 PM - 20 Mar 2019

201 Retweets **913** Likes

18 201 913

François Chollet
@fchollet

Follow

As a Keras user, what implementation of the Keras API do you use most of the time?

- Standalone Keras package
- tf.keras
- Other (e.g. MXNet-Keras)
- Show results

Vote 1,655 votes • 6 hours left

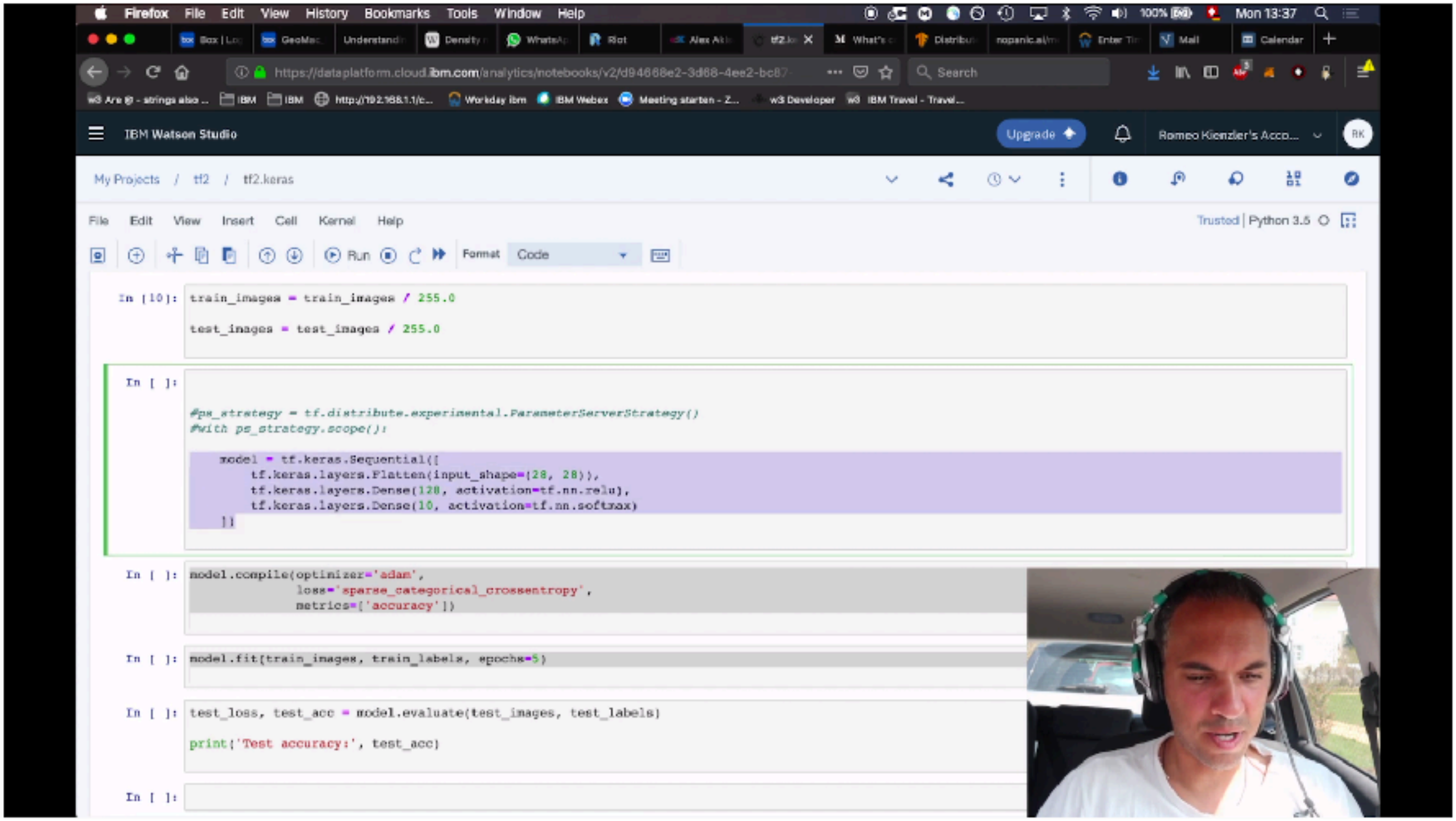
7:43 PM - 24 Mar 2019

12 Retweets **20** Likes

3 12 20

YouTube

Search



What's new in TensorFlow 2.0? Video Series - (2 of X) - Keras Distribution Strategy Integration

49 views

5 0 SHARE SAVE ...

Romeo Kienzler
Published on Apr 1, 2019

SUBSCRIBE 1.2K

romeo kienzler - ibm - 2019

tensorflow / docs

Watch 88 | Star 1,003 | Fork 963

Code | Pull requests 22 | Insights

Branch: master | docs / site / en / guide / distribute_strategy.ipynb | Find file | Copy path

romeokienzler Update distribute_strategy.ipynb e9f193d 4 days ago

5 contributors

894 lines (893 sloc) | 40.9 KB | Raw | Blame | History

Copyright 2018 The TensorFlow Authors.

In [0]:
```
#@title Licensed under the Apache License, Version 2.0 (the "License");
# you may not use this file except in compliance with the License.
# You may obtain a copy of the License at
#
# https://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
```

Distributed Training in TensorFlow

Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

Notices and disclaimers

© 2019 International Business Machines Corporation.
No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. This document is distributed “as is” without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.
IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts.

In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer’s responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer’s business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

Notices and disclaimers continued

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products.

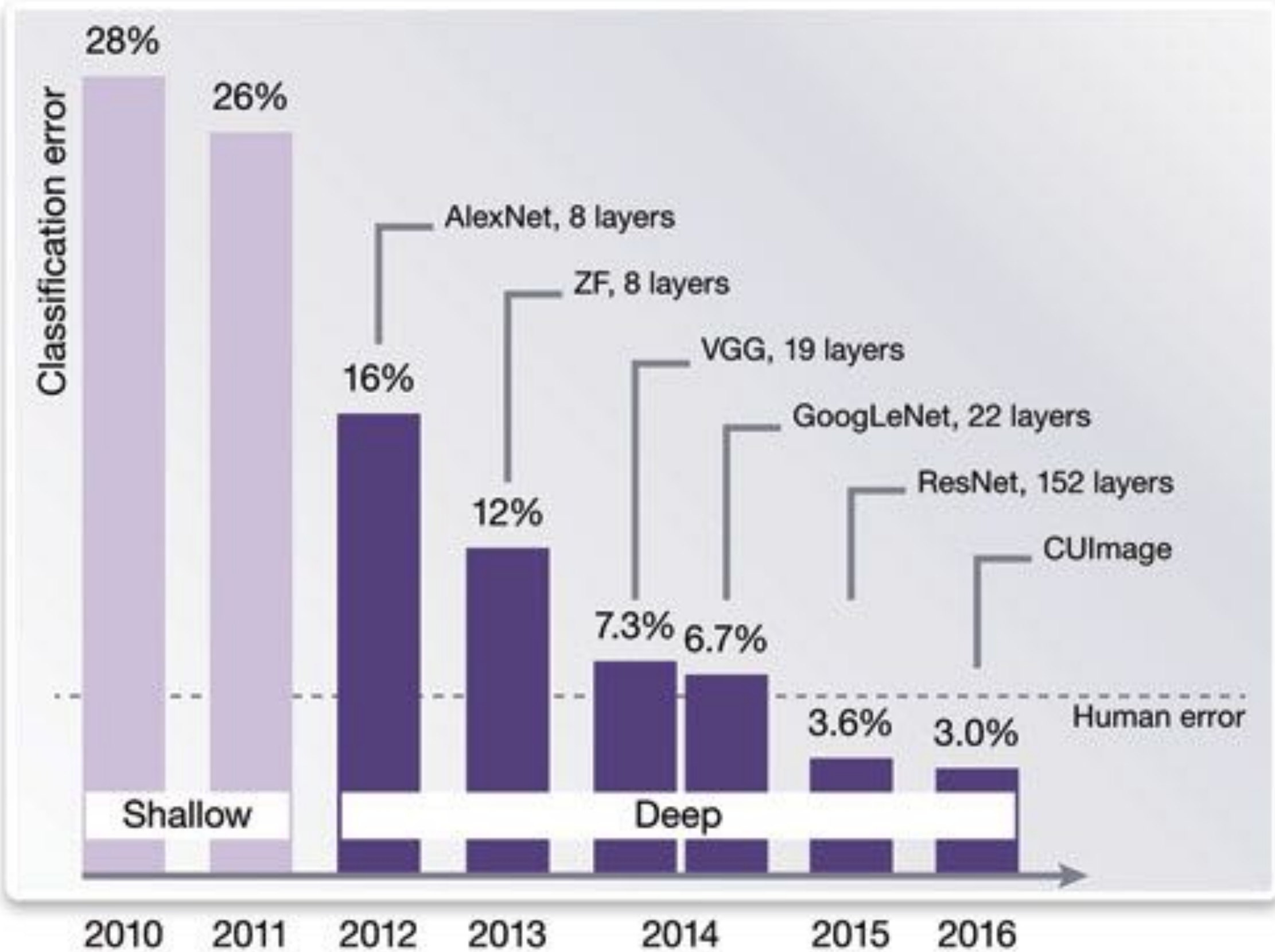
Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

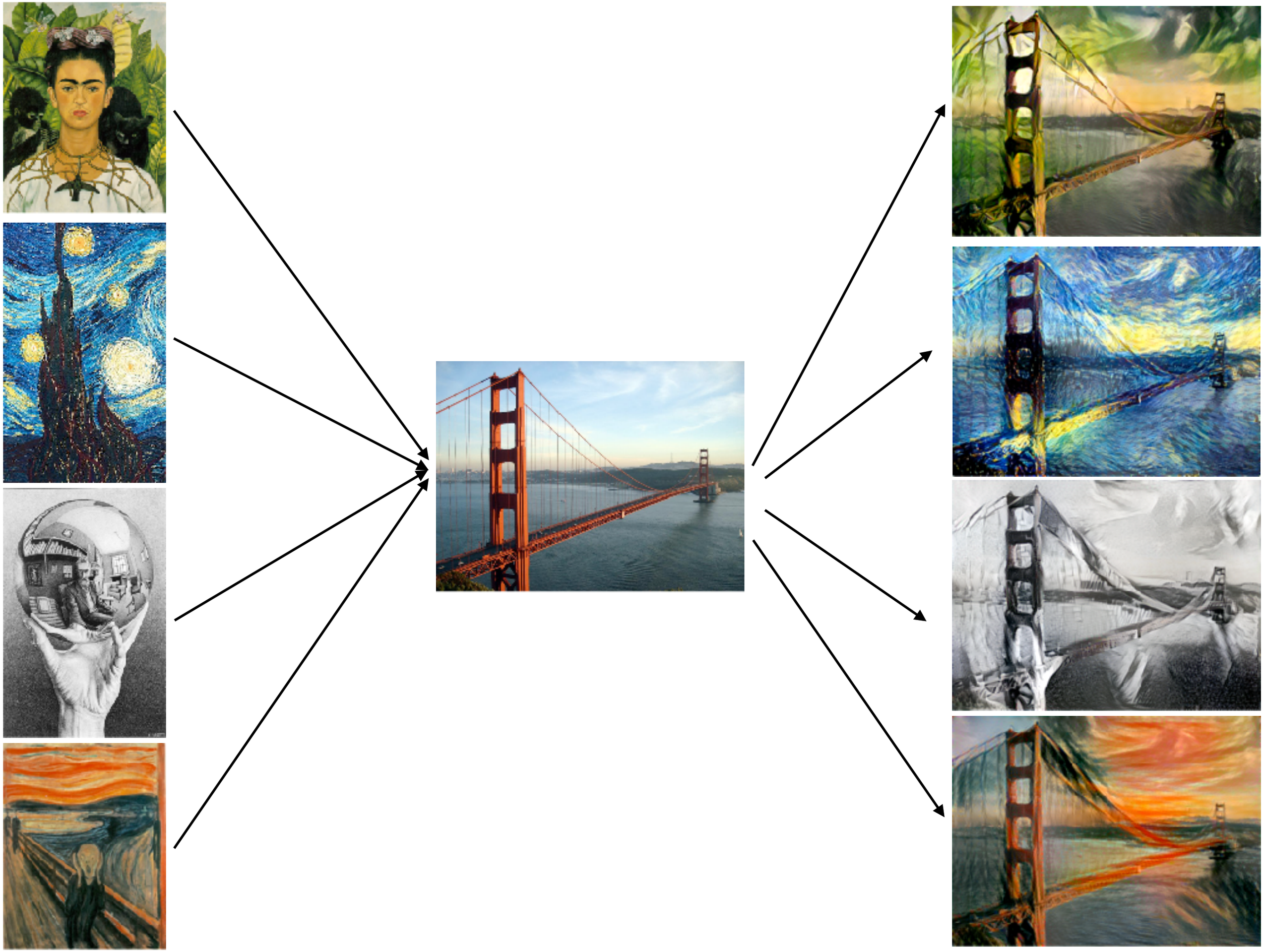
IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml.

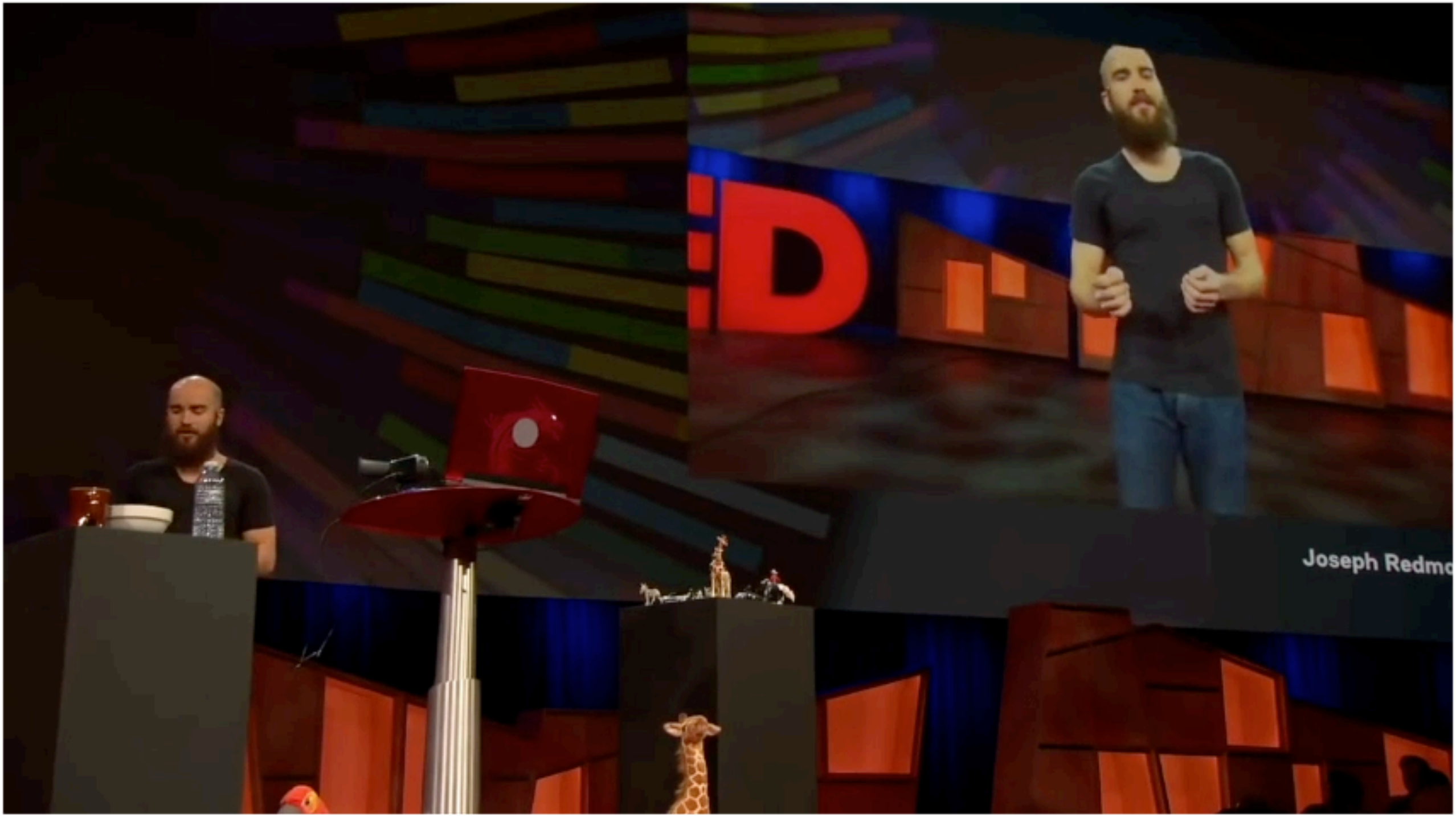
IBM’s AI journey

Life, the Universe and Everything



https://eecatalog.com/IoT/2017/10/05/as-embedded-vision-takes-off-software-frameworks-bear-watching/



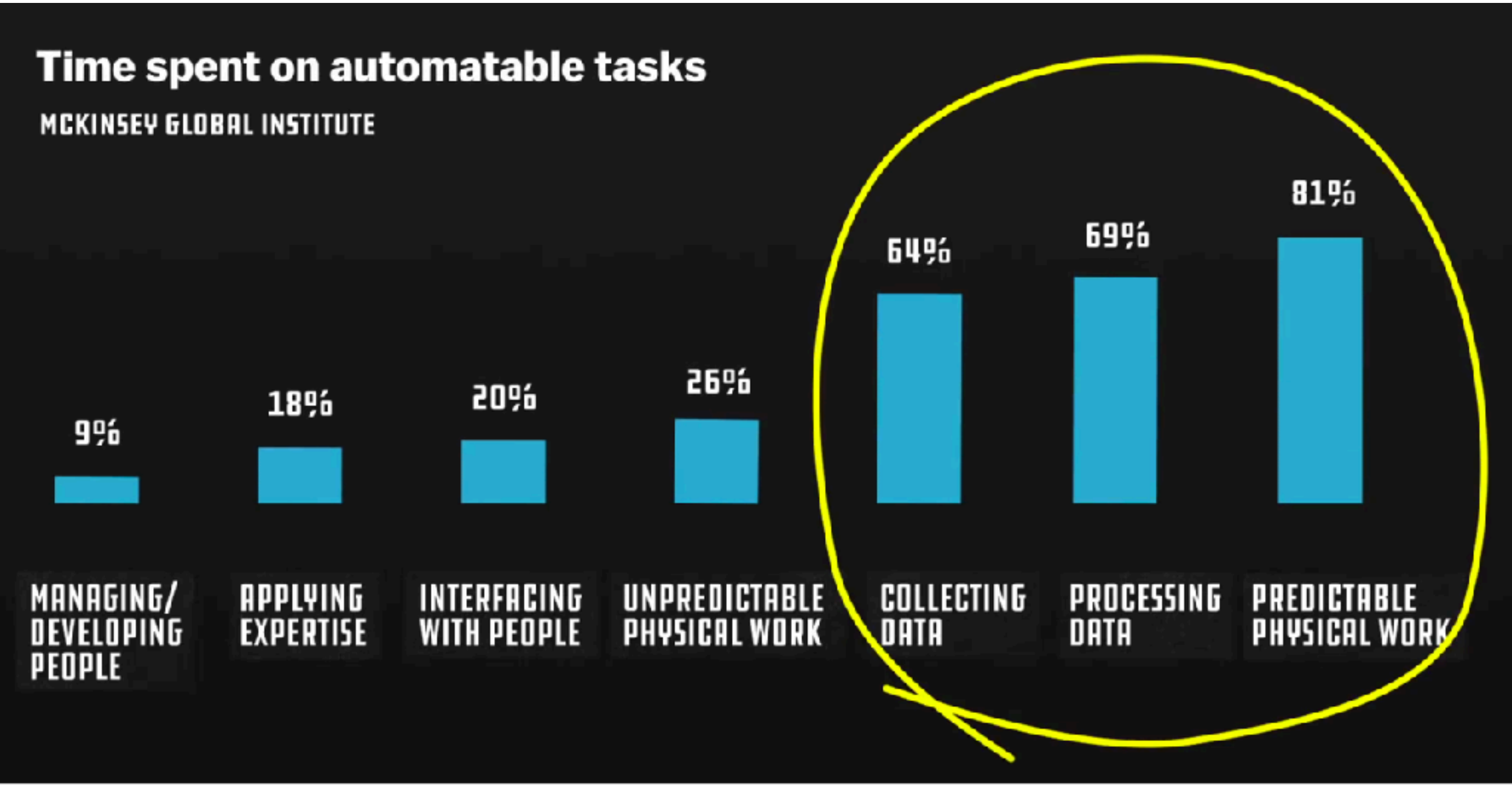

Joseph Redmo

YOLO V2
http://pjreddie.com/yolo

"construction worker in orange safety vest is working on road."

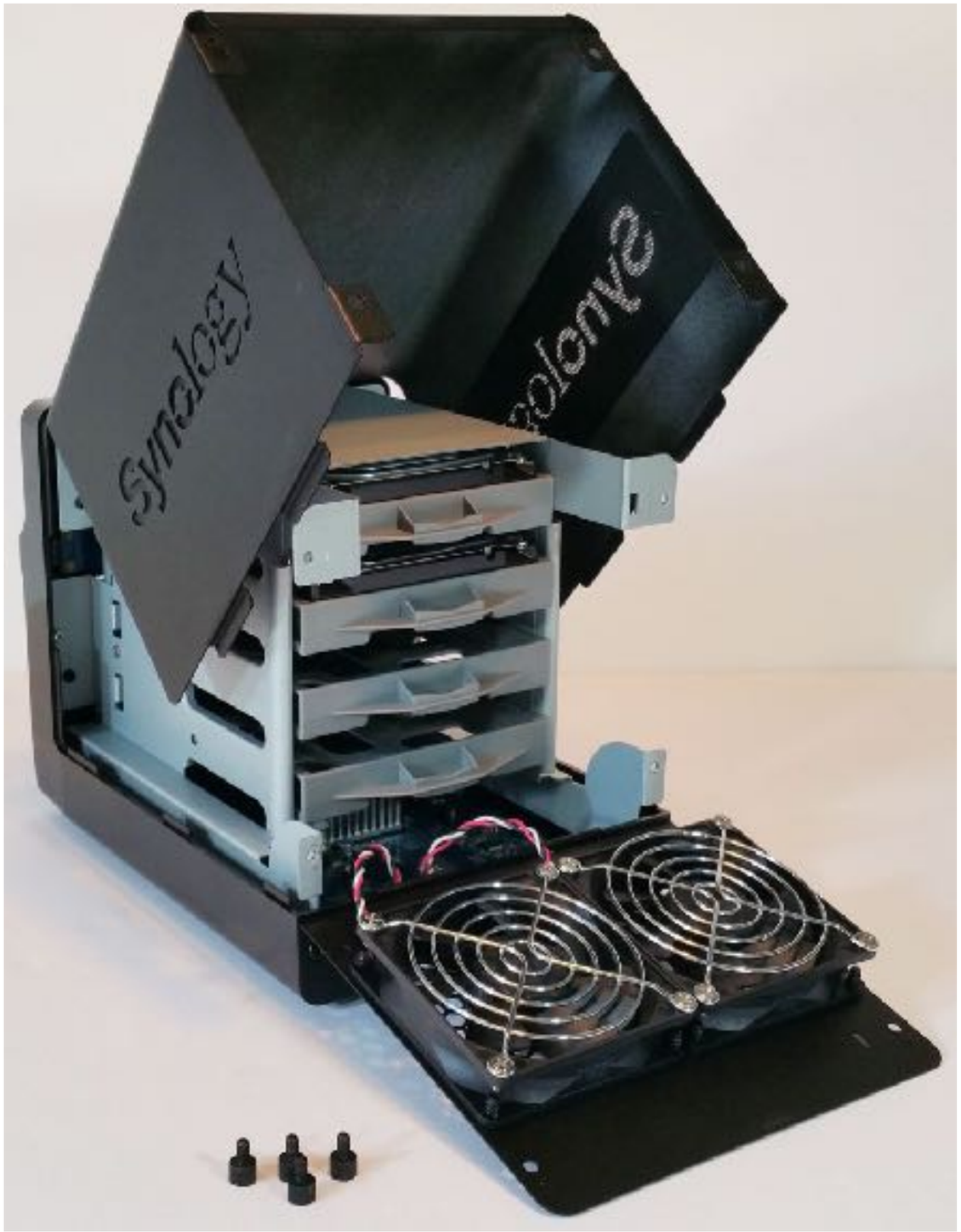
"baseball player is throwing ball in game."

"black and white dog jumps over bar."


Time spent on automatable tasks
MCKINSEY GLOBAL INSTITUTE
9%
18%
20%
26%
64%
69%
81%
MANAGING/ DEVELOPING PEOPLE
APPLYING EXPERTISE
INTERFACING WITH PEOPLE
UNPREDICTABLE PHYSICAL WORK
COLLECTING DATA
PROCESSING DATA
PREDICTABLE PHYSICAL WORK

STATUS
LAN
DSK 1
DSK 2
Synology®
DS216 SE

Openness matters…!



- Linux RAID
- OpenSSL Encryption
- Standard SATA Drives
- Linux OS
- Terminal access via SSH
- Powerful DSM Software (closed)
- AppStore like ecosystem

Hardware matters…!

Cloud Infrastructure as a Service

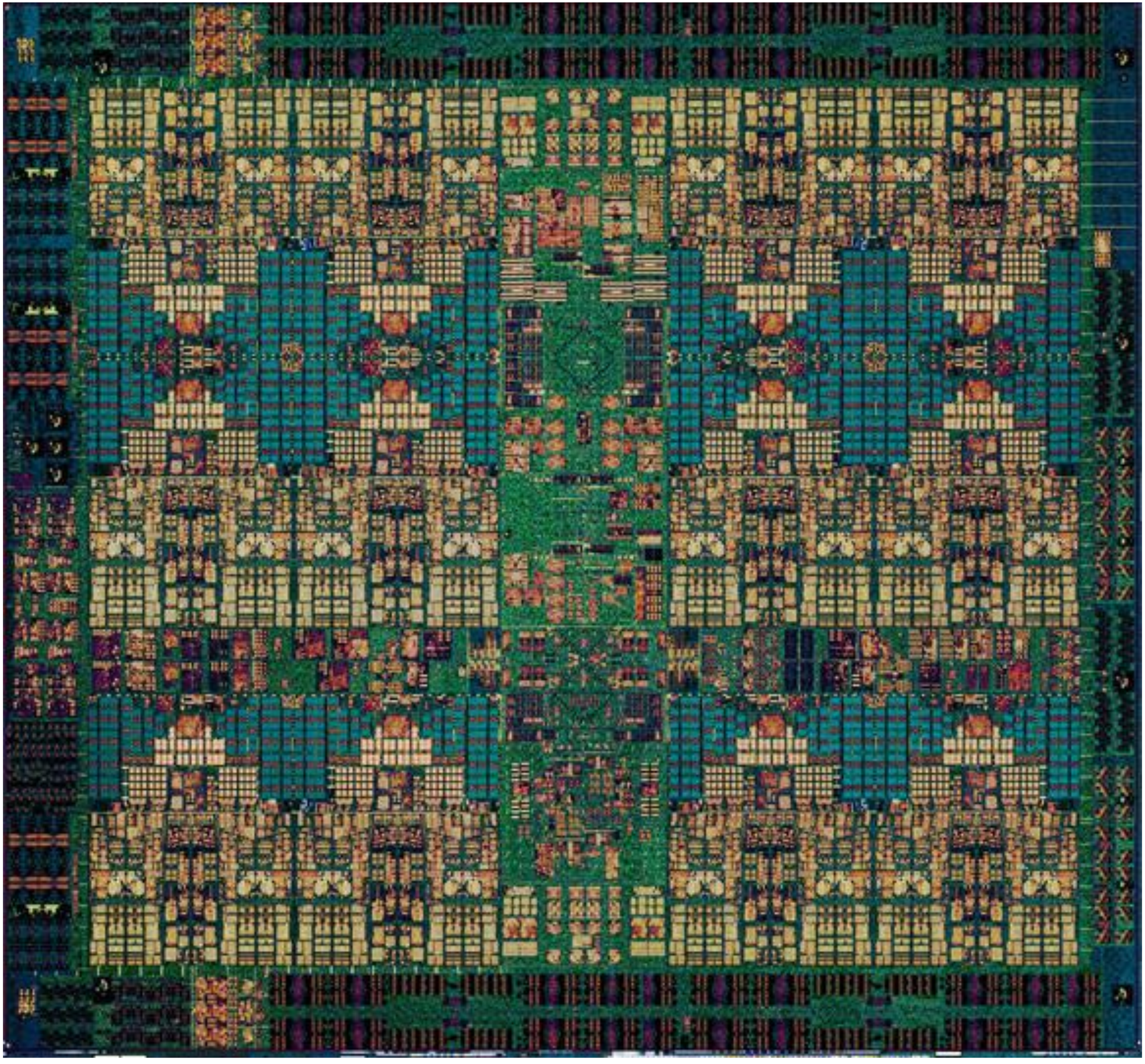
docker

kubernetes

NVIDIA

CPUs, GPUs, FPGAs, ASICs, Network and the rest…

CPU



Levels

Benefits

Current Members

How to Join

Current Members

Platinum Level

Google

IBM

Mellanox TECHNOLOGIES

Micron

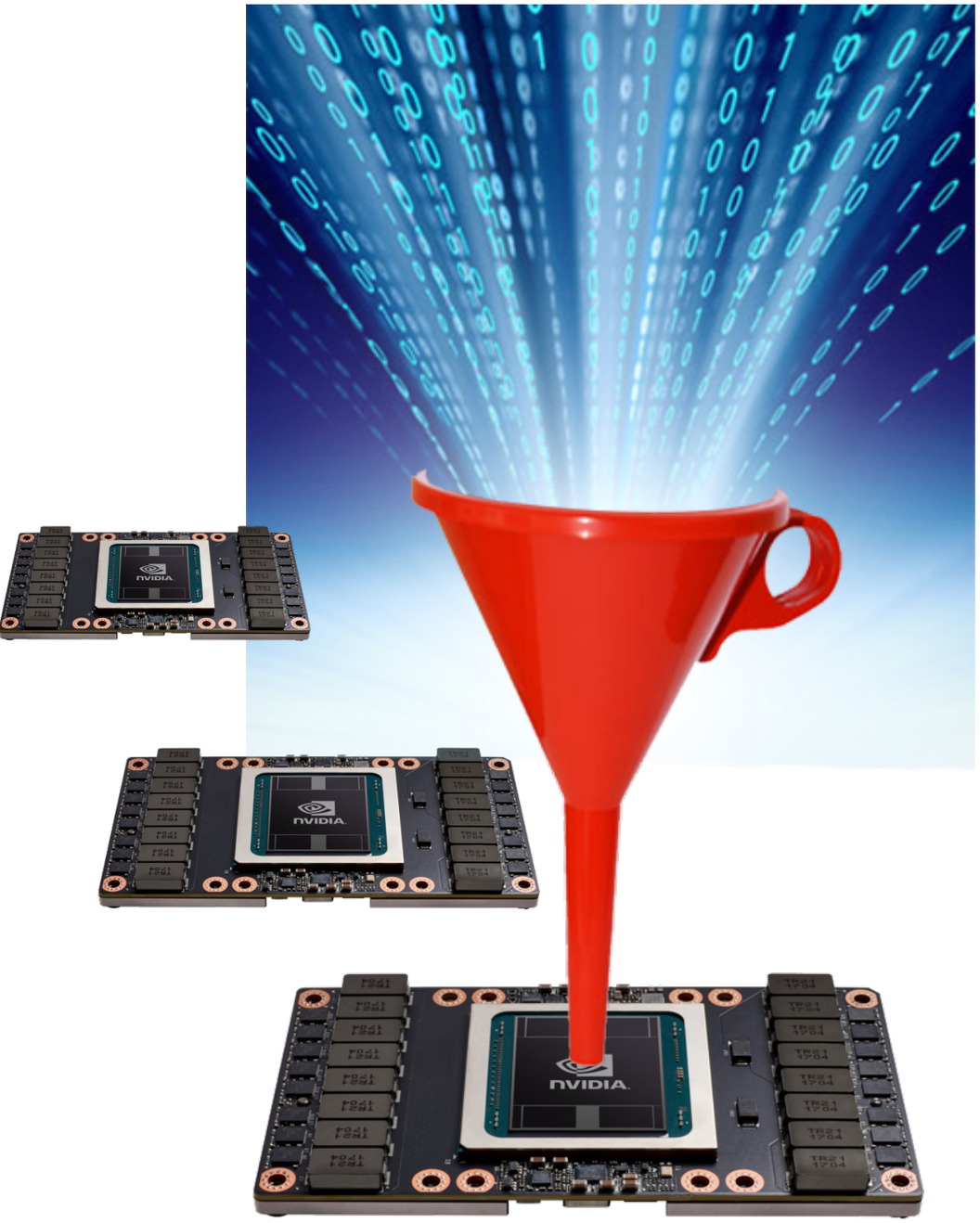
NVIDIA

redhat

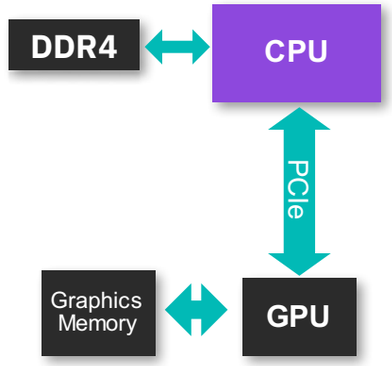
ubuntu Supported by Canonical

XILINX

GPU



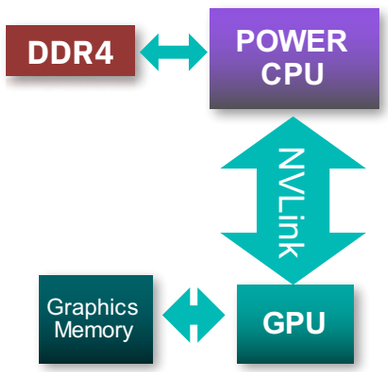
```
Time(%)   Time      Calls   Avg       Name
49.35%    29.581ms   1      29.581ms   [CUDA memcpy DtoH]
47.48%    28.462ms   1      28.462ms   [CUDA memcpy HtoD]
3.17%     1.9000ms   1      1.9000ms   naiveTransposeKernel
```



PCIe Gen3 – 32GB/sec

NVLink 1.0 – 80GB/sec

NVLink 2.0 – 150 GB/sec

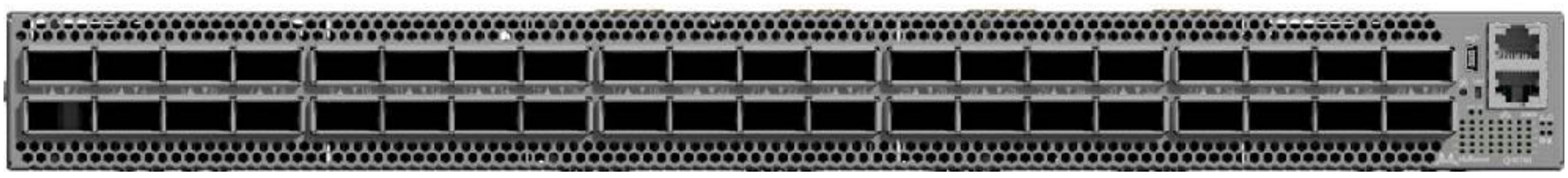


Network

QM8700 Series - Mellanox Quantum™ HDR 200Gb/s InfiniBand Smart Switches

40-port Non-blocking HDR 200Gb/s InfiniBand Smart Switch

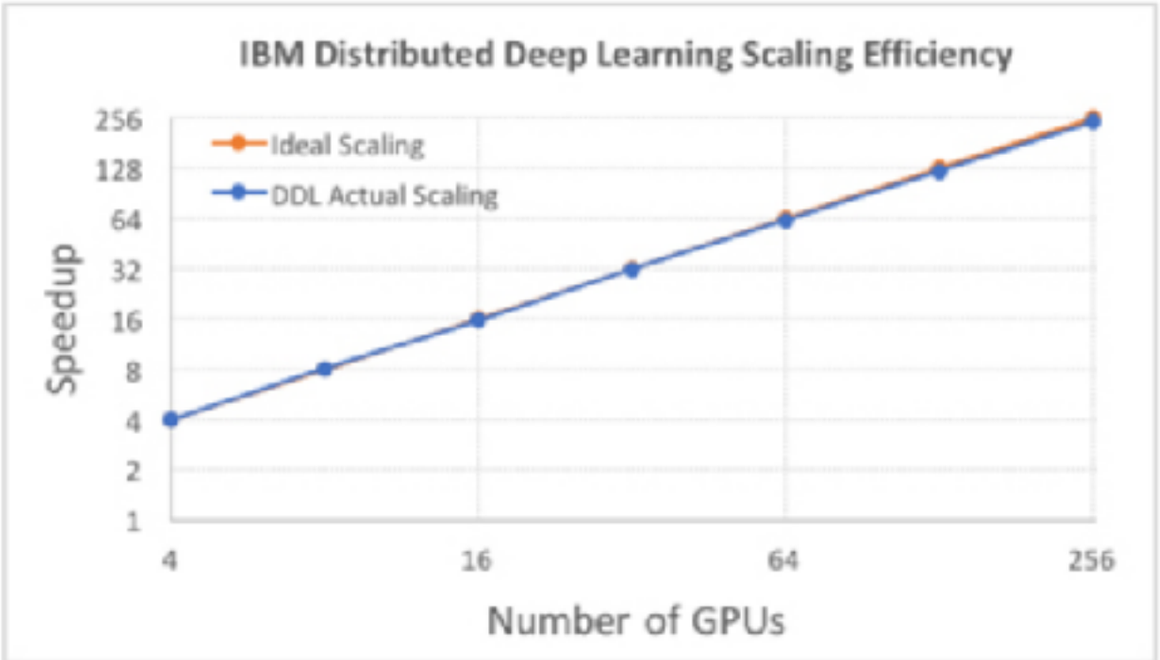
Mellanox provides the world's smartest switches, enabling in-network computing through the Co-Design Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™ technology. The QM8700 series has the highest fabric performance available in the market with up to 16Tb/s of non-blocking bandwidth with sub 90ns port-to-port latency.



August 8, 2017
Posted in: AI, Cognitive Computing

IBM Research achieves record deep learning performance with new software technology

Summary: IBM Research publishes in arXiv close to ideal scaling with new distributed deep learning software which achieved record communication overhead and 95% scaling efficiency on the Caffe deep learning framework over 256 NVIDIA GPUs in 64 IBM Power systems. Previous best scaling was demonstrated by Facebook AI Research of 89% for a training run on Caffe2, at higher communication overhead. IBM Research also beat Facebook's time by training the model in 50 minutes, versus the 1 hour Facebook took. Using this software, IBM Research achieved a new image recognition accuracy of 33.8% for a neural network trained on a very large data set (7.5M images). The previous record published by Microsoft demonstrated 29.8% accuracy.

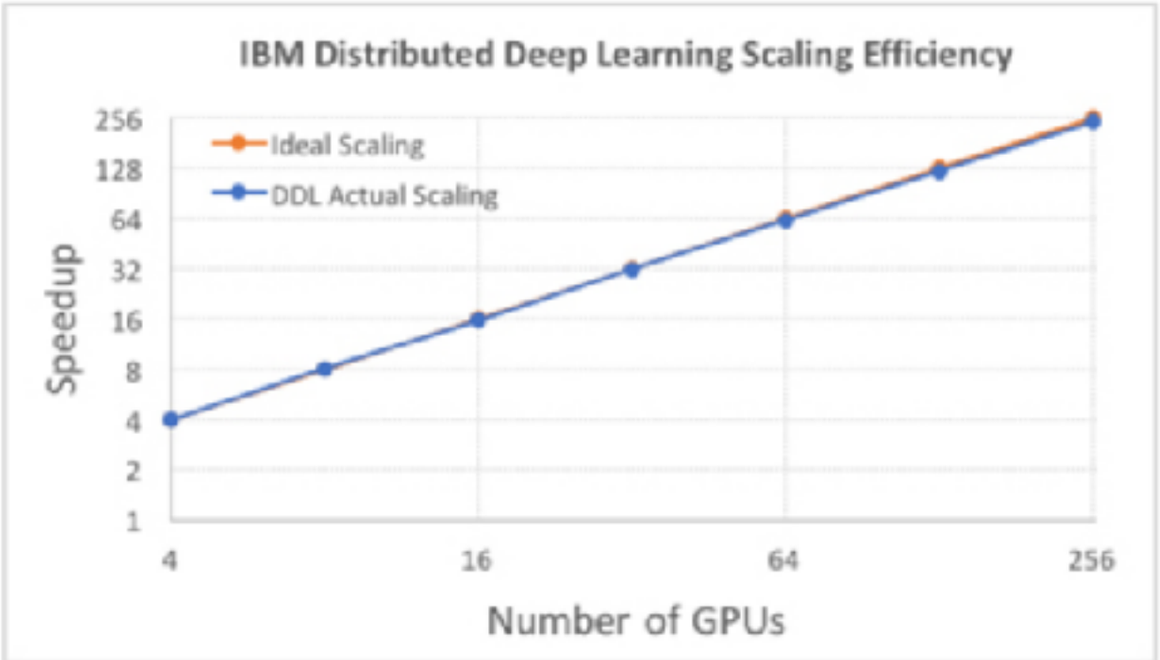


August 8, 2017
Posted in: AI, Cognitive Computing

IBM Research achieves record deep learning performance with new software technology

Scaling Efficiency
Facebook: 89%
IBM: 95%

Summary: IBM Research publishes in arXiv close to ideal scaling with new distributed deep learning software which achieved record communication overhead and 95% scaling efficiency on the Caffe deep learning framework over 256 NVIDIA GPUs in 64 IBM Power systems. Previous best scaling was demonstrated by Facebook AI Research of 89% for a training run on Caffe2, at higher communication overhead. IBM Research also beat Facebook's time by training the model in 50 minutes, versus the 1 hour Facebook took. Using this software, IBM Research achieved a new image recognition accuracy of 33.8% for a neural network trained on a very large data set (7.5M images). The previous record published by Microsoft demonstrated 29.8% accuracy.



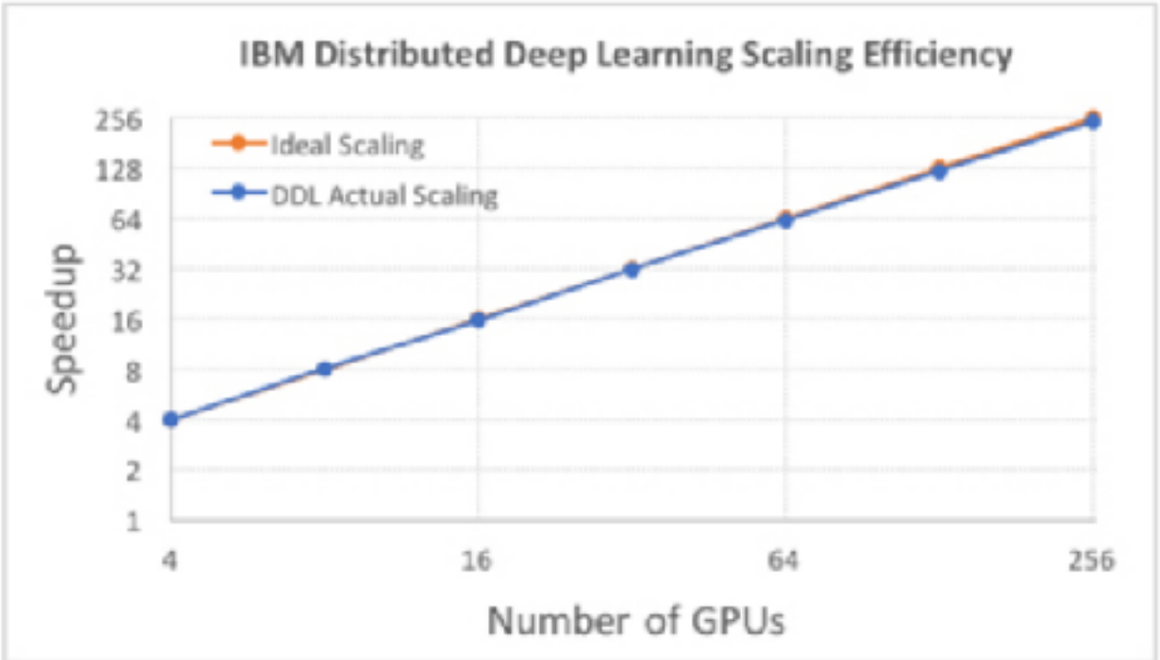
August 8, 2017
Posted in: AI, Cognitive Computing

IBM Research achieves record deep learning performance with new software technology

Scaling Efficiency
Facebook: 89%
IBM: 95%

Runtime
Facebook: 1h
IBM: 50 minutes

Summary: IBM Research publishes in arXiv close to ideal scaling with new distributed deep learning software which achieved record communication overhead and 95% scaling efficiency on the Caffe deep learning framework over 256 NVIDIA GPUs in 64 IBM Power systems. Previous best scaling was demonstrated by Facebook AI Research of 89% for a training run on Caffe2, at higher communication overhead. IBM Research also beat Facebook's time by training the model in 50 minutes, versus the 1 hour Facebook took. Using this software, IBM Research achieved a new image recognition accuracy of 33.8% for a neural network trained on a very large data set (7.5M images). The previous record published by Microsoft demonstrated 29.8% accuracy.



August 8, 2017
Posted in: AI, Cognitive Computing

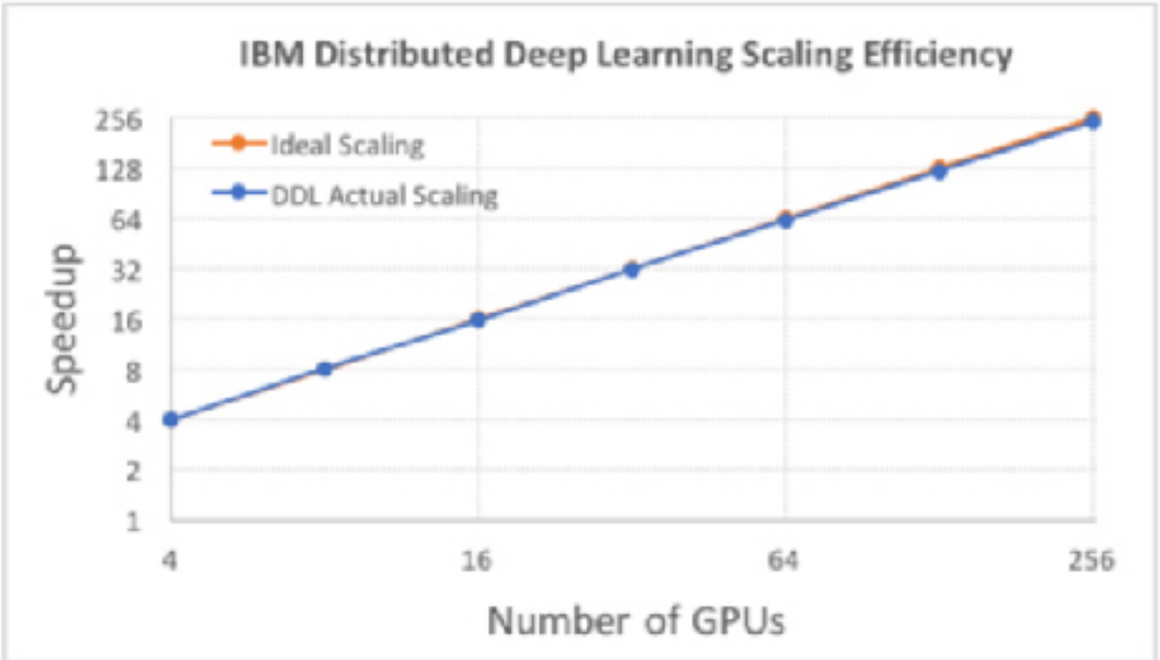
IBM Research achieves record deep learning performance with new software technology

Scaling Efficiency
Facebook: 89%
IBM: 95%

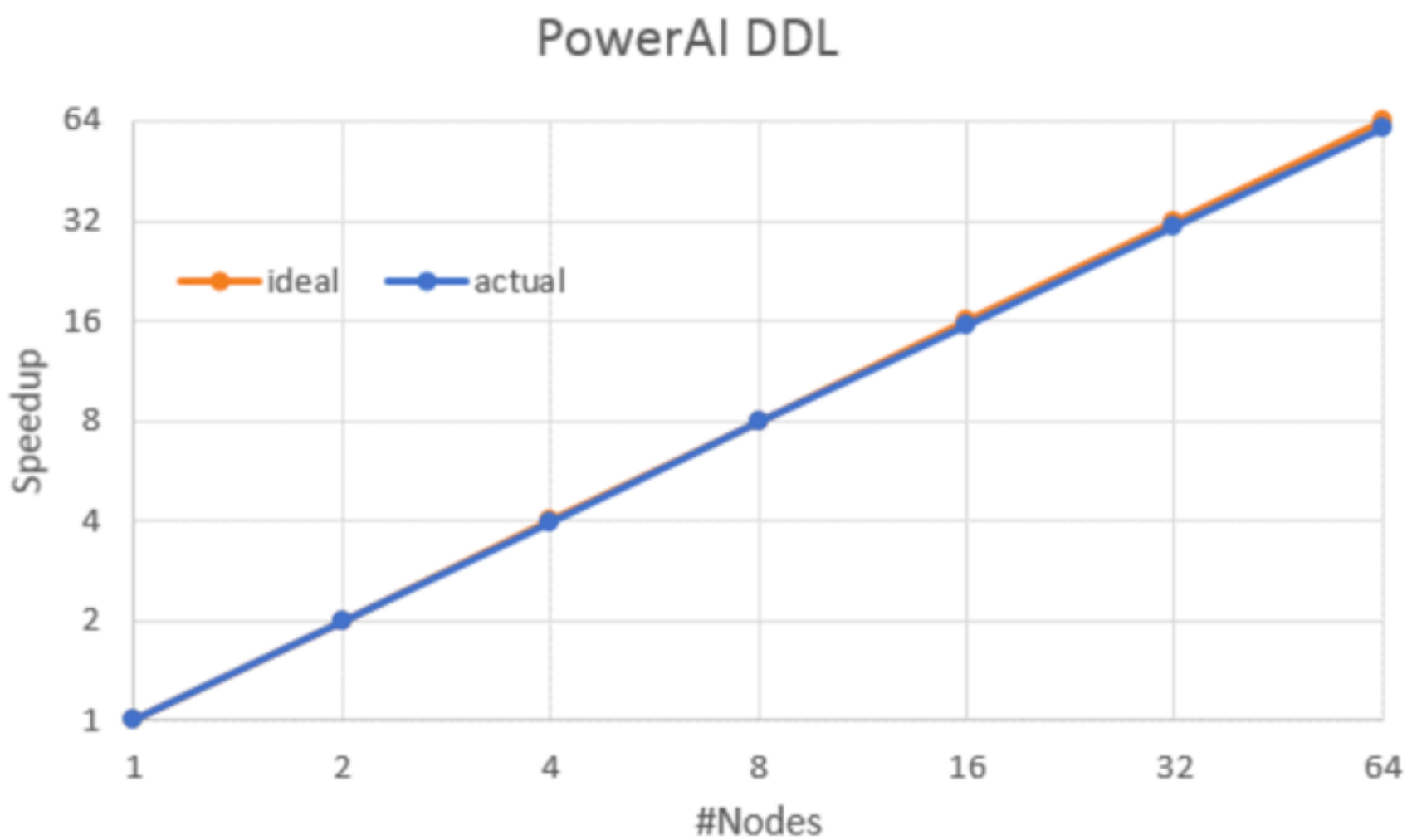
Runtime
Facebook: 1h
IBM: 50 minutes

Summary: IBM Research publishes in arXiv close to ideal scaling with new distributed deep learning software which achieved record communication overhead and 95% scaling efficiency on the Caffe deep learning framework over 256 NVIDIA GPUs in 64 IBM Power systems. Previous best scaling was demonstrated by Facebook AI Research of 89% for a training run on Caffe2, at higher communication overhead. IBM Research also beat Facebook's time by training the model in 50 minutes, versus the 1 hour Facebook took. Using this software, IBM Research achieved a new image recognition accuracy of 33.8% for a neural network trained on a very large data set (7.5M images). The previous record published by Microsoft demonstrated 29.8% accuracy.

Accuracy
Microsoft: 29.8%
IBM: 33.8%



Disclaimer: This is a beta talk - don't kill me - Disclaimer: This is a beta talk - don't kill me - Disclaimer: This is a beta talk - don't kill me



#GPUs	4	8	16	32	64	128	256
#Nodes	1	2	4	8	16	32	64
Speedup	1.0	2.0	3.9	7.9	15.5	30.5	60.6
Scaling efficiency	1.00	1.00	0.98	0.99	0.97	0.95	0.95

Figure 2: Resnet-50 for 1K classes using up to 256 GPUs with Caffe.

© 2017 IBM Corporation

U.S. DEPARTMENT OF ENERGY

Office of Science

NNSA

National Nuclear Security Administration

- 200 petaflops
- 4600 nodes
- 9200 Power9 CPU's
- 220800 cores
- 1766400 hyper threads
- 27600 nVidia V100 GPUs
- 10 PB memory
- Power 15 MW
- 1% of Ethereum's hash rate

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,282,544	122,300.0	187,659.3	8,806
2	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
3	**Sierra** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/NNSA/LLNL United States	1,572,480	71,610.0	119,193.6	

IBM PowerAI Platform

PowerAI Software Distribution

Deep Learning Frameworks

- Caffe
- NVIDIA Caffe
- IBM Caffe
- torch
- TensorFlow
- theano
- Chainer

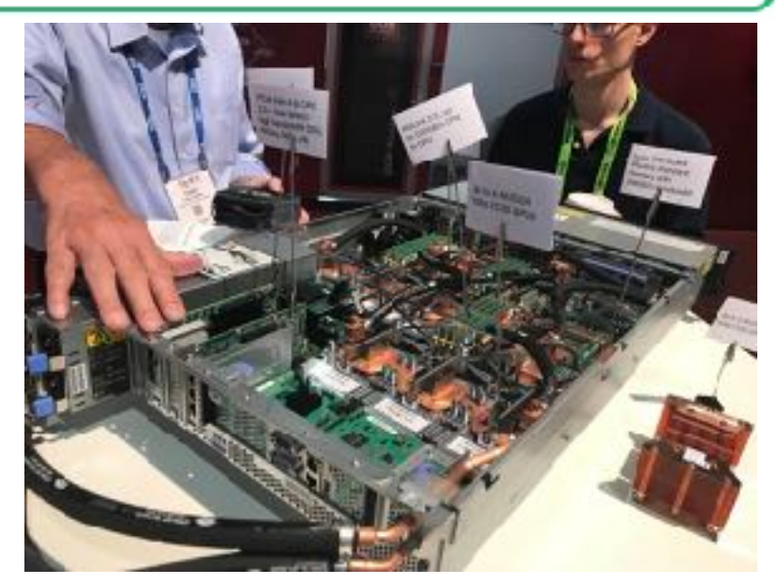
Supporting Libraries

- DIGITS
- OpenBLAS
- Distributed Frameworks
- Bazel
- NCCL

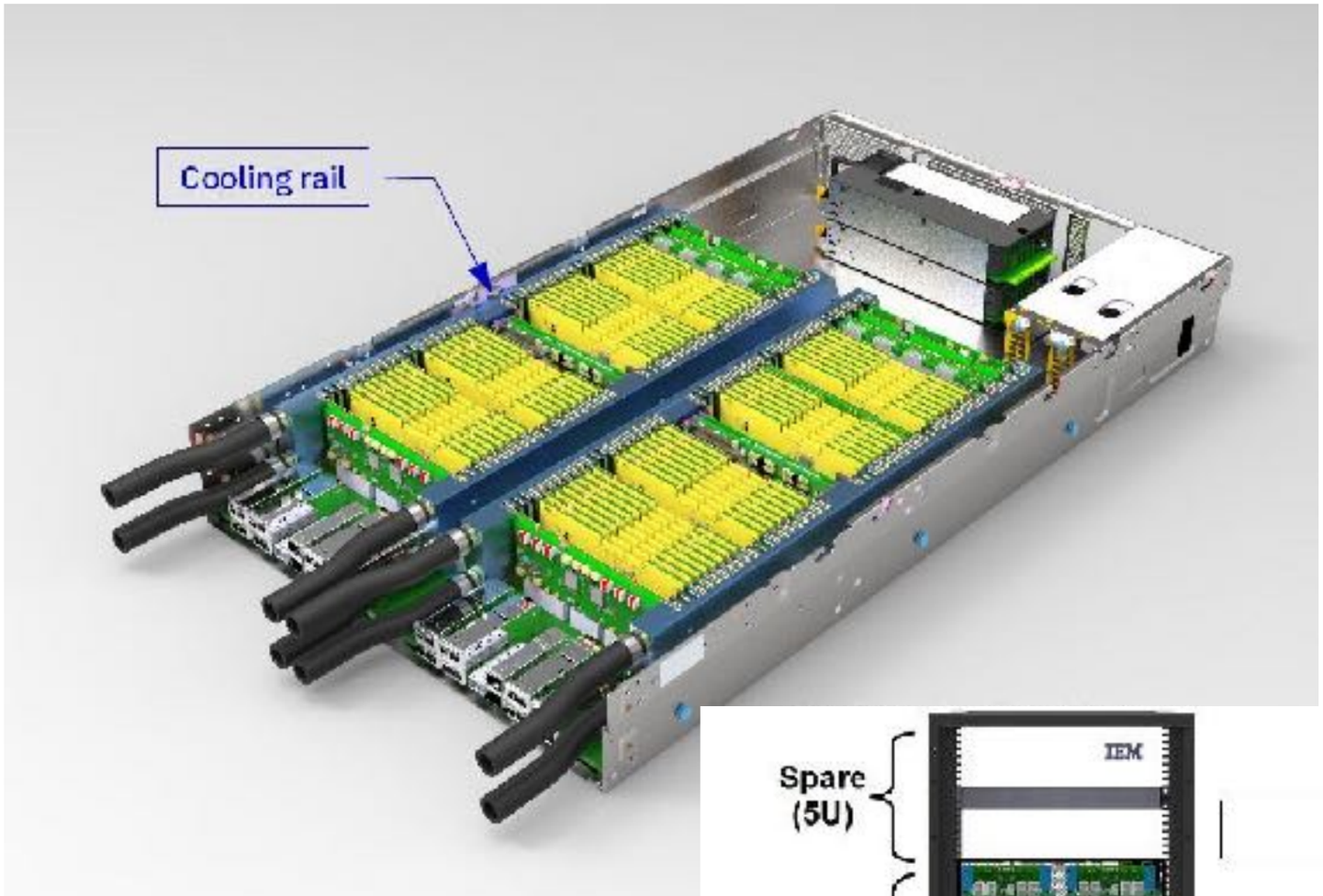
IBM Power System for HPC, with NVLink

Breakthrough performance for GPU accelerated applications, including Deep Learning and Machine Learning.

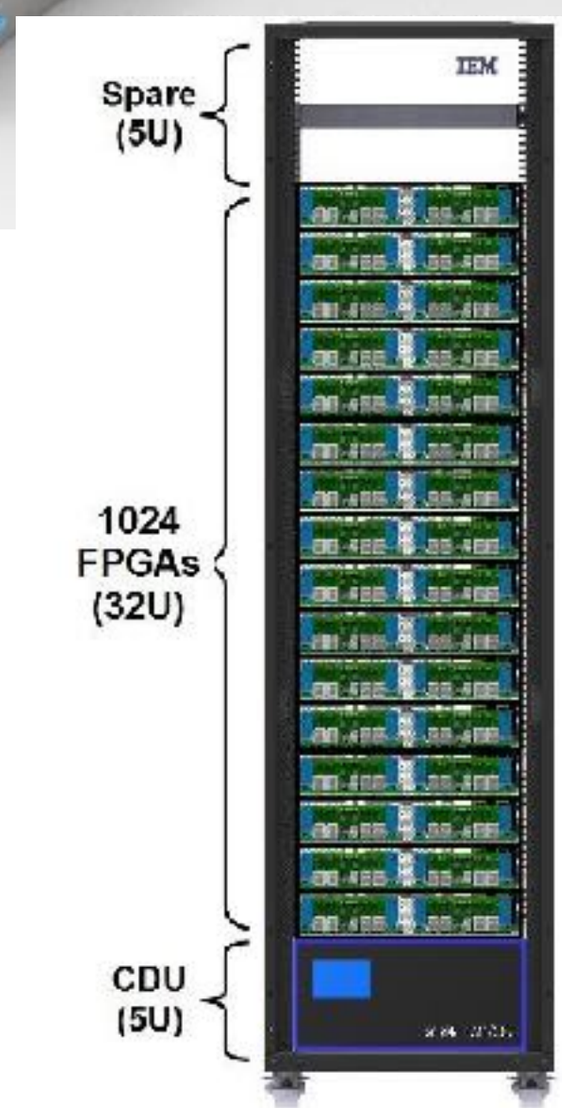
- **Ubuntu Linux**
- **SSH Terminal Access**
- **CUDA**
- **Keras**
- **Scikit-Learn**
- **Caffe2**
- **PyTorch**
- **R**
- **Scikit-Learn**



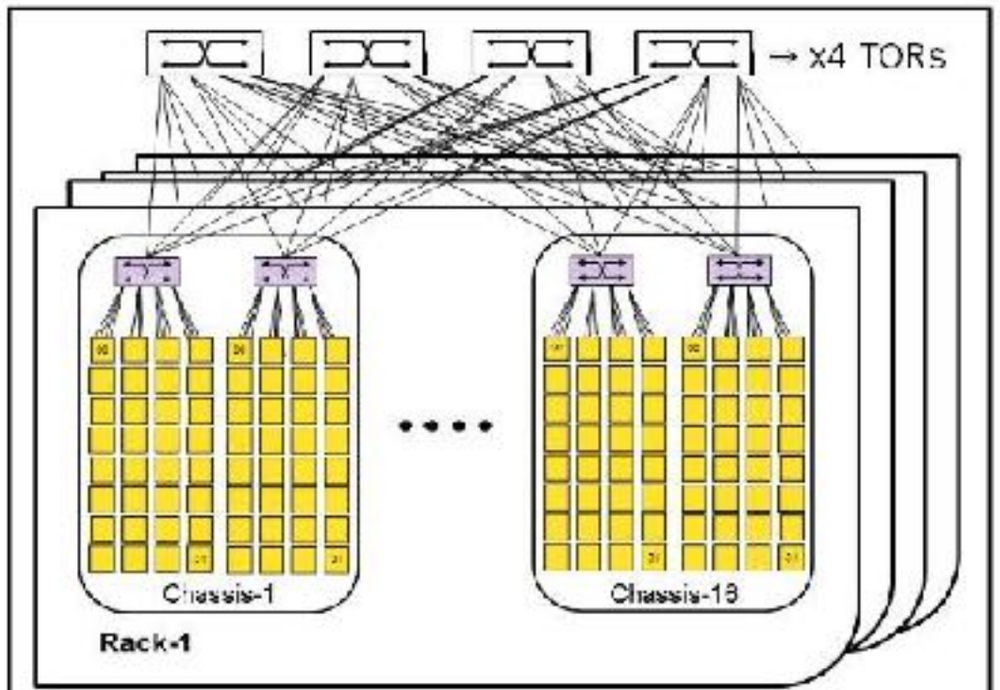
FPGAs



https://www.linkedin.com/pulse/how-do-you-squeeze-1000-fpgas-dc-rack-francois-abel/



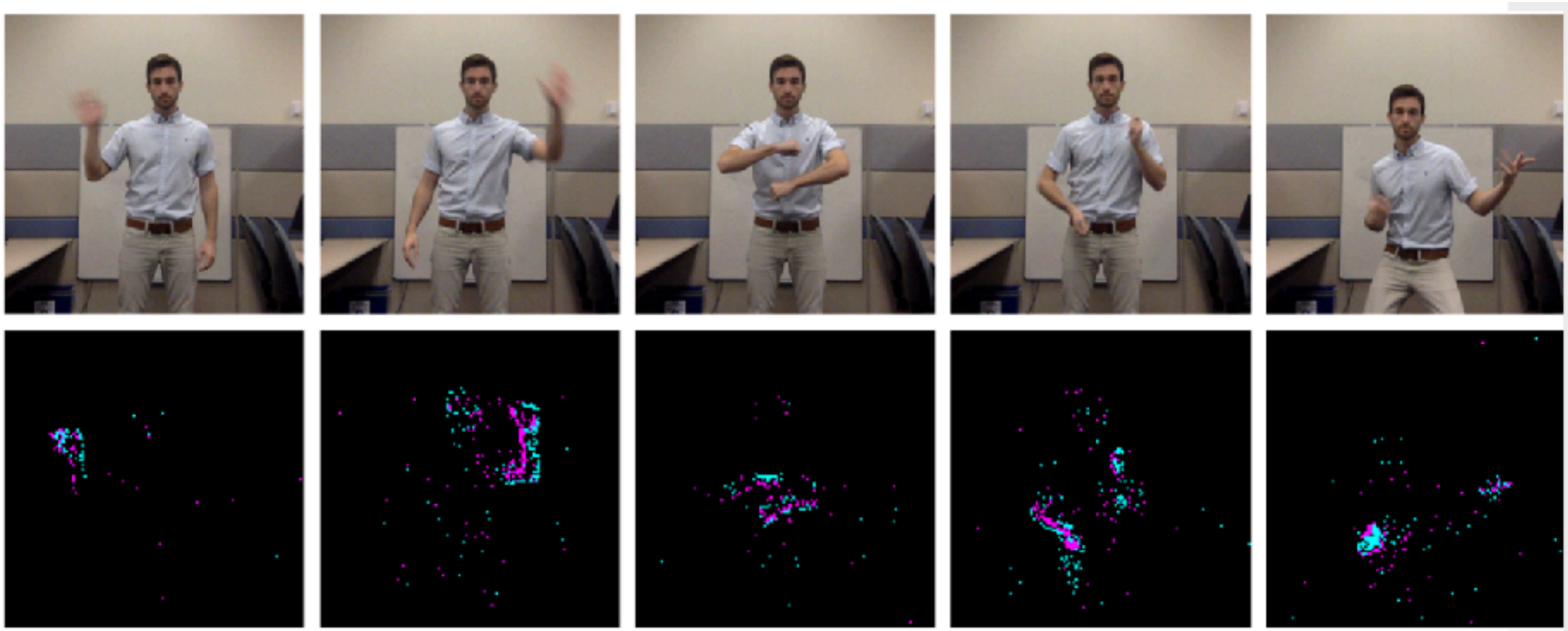
1024 FPGAs → 2.8M DSPs,
$2x10^{15}$ Fixed-Point Multiply-Accumulates/s
10 Tb/s bi-sec. Bw – 16 TB DDR4 – 40 kW max.



Example of a two-tier Clos architecture with 4096 FPGAs and four edge switches (TORs)

ASICs

Combining the IBM TrueNorth neurosynaptic processor with an iniLabs Dynamic Vision Sensor (DVS)



96.5 percent accuracy within a tenth of a second consuming under 200 mW

https://www.youtube.com/watch?v=g08IW-qRomM

Robustness…

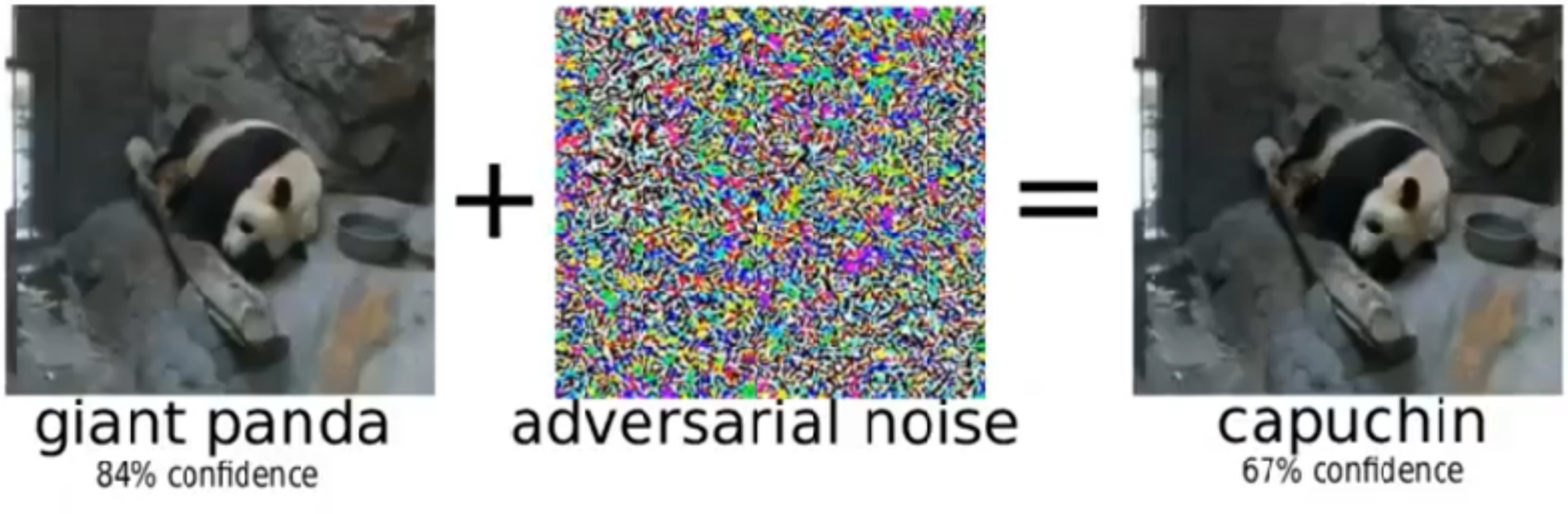
(a) Husky classified as wolf (b) Explanation

Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.

	Before	After
Trusted the bad model	10 out of 27	3 out of 27
Snow as a potential feature	12 out of 27	25 out of 27

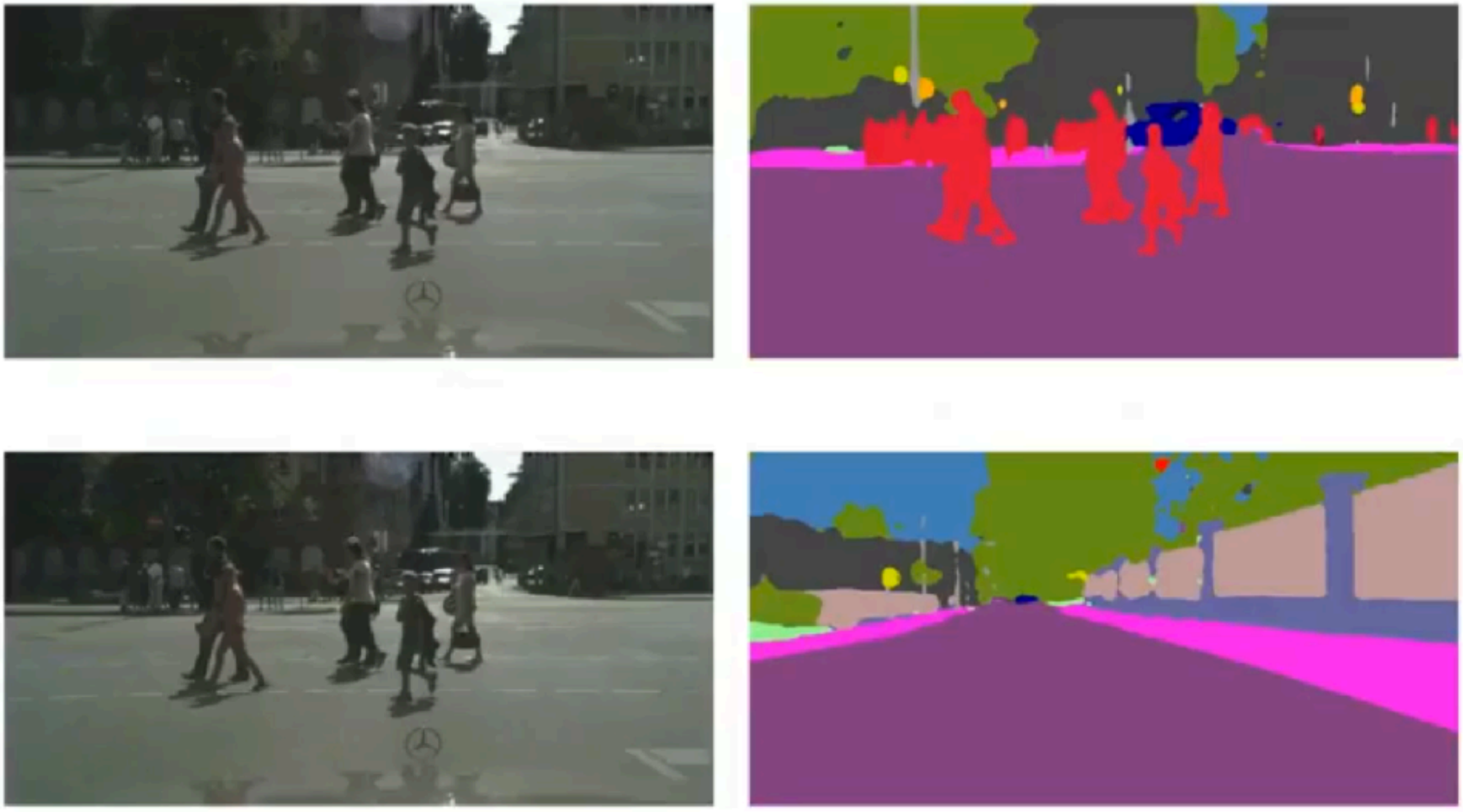
https://hackernoon.com/dogs-wolves-data-science-and-why-machines-must-learn-like-humans-do-41c43bc7f982

Adversarial Examples



- Perturb model inputs with crafted noise
- Model fails to recognize input correctly
- Attack undetectable by humans
- Random noise does not work.

Attack noise hides pedestrians from the detection system.

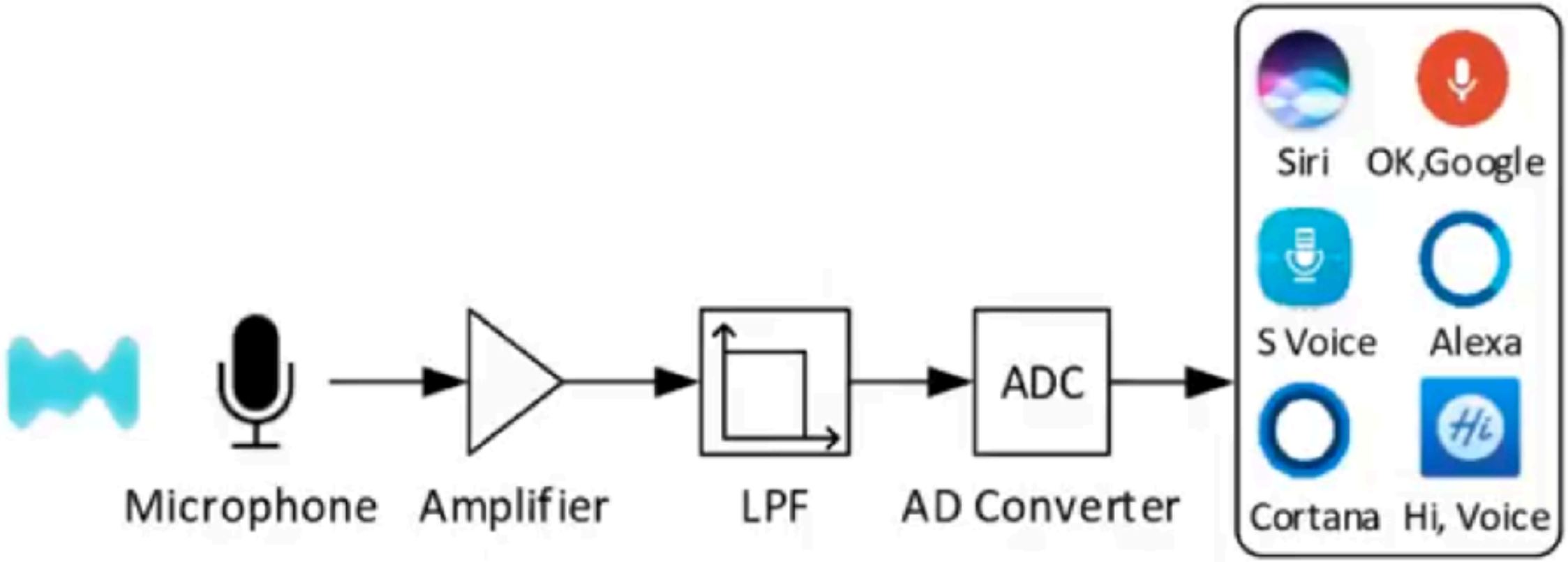


Metzen et al., Universal Adversarial Perturbations against Semantic Image Segmentation.
https://arxiv.org/abs/1704.05712

STOP
© Washington University

LOVE
STOP
HATE
© Washington University

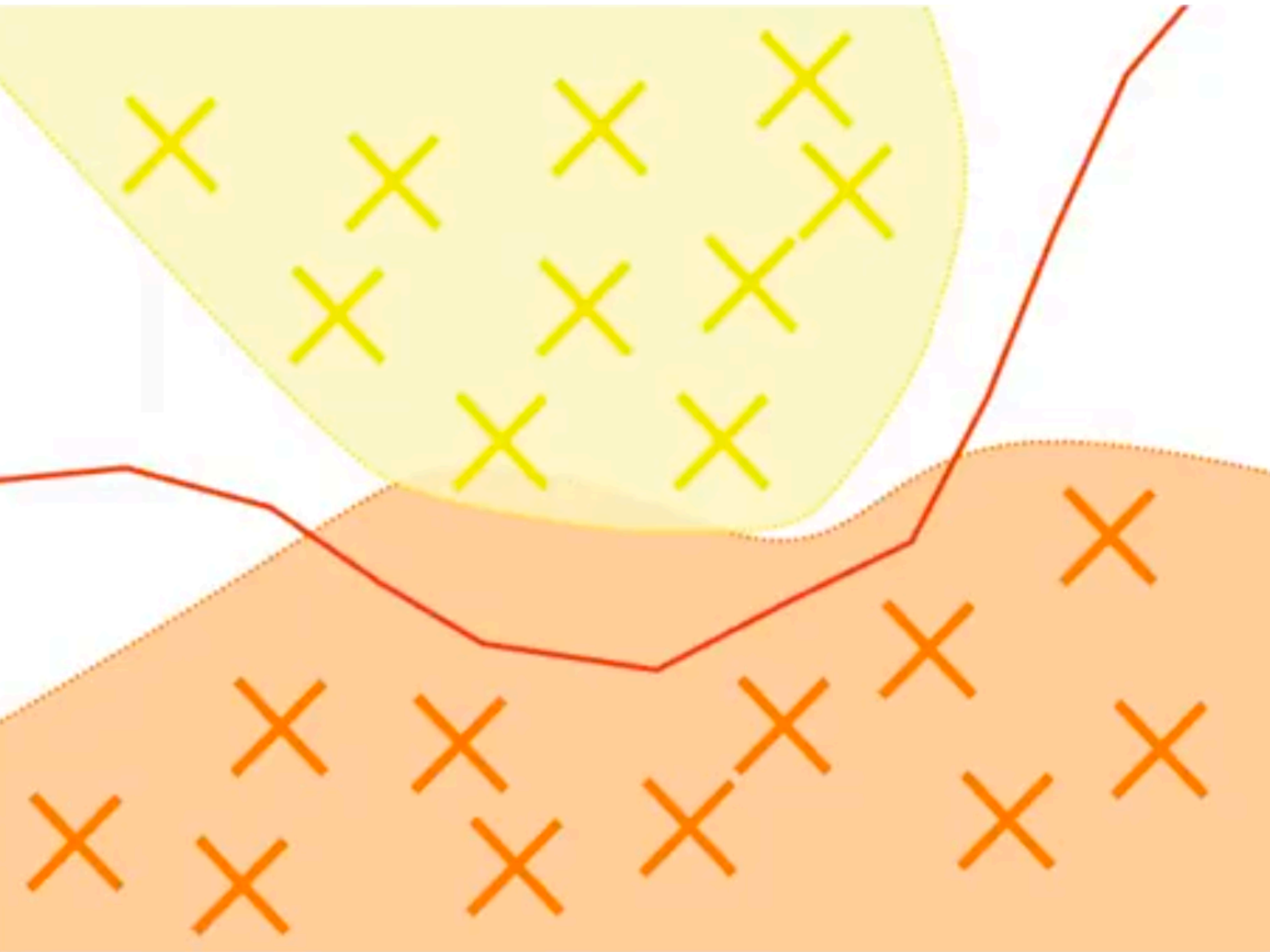
SPEED
LIMIT
45

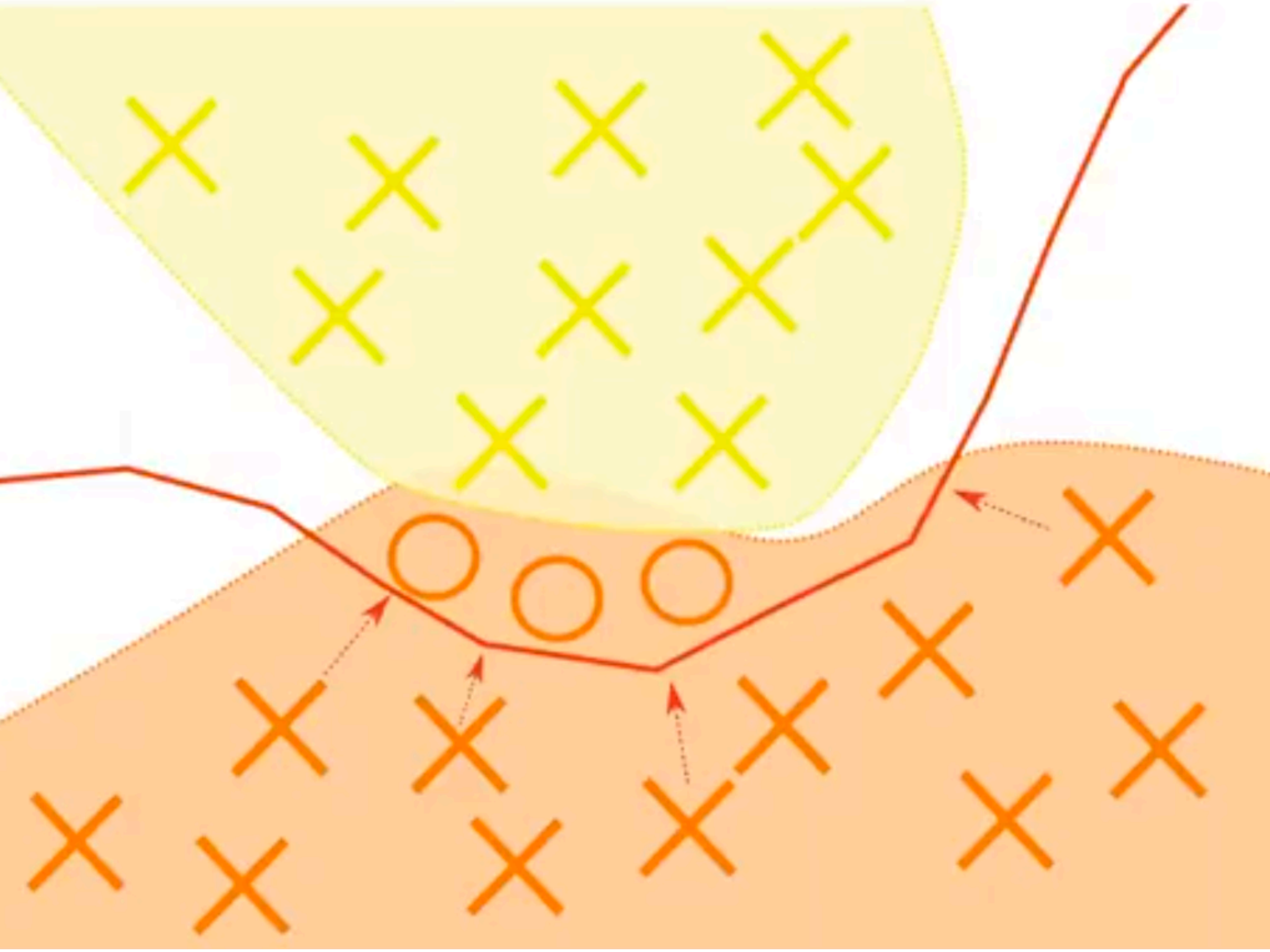


Okay Google, text John!

- Stealthy voice commands recognized by devices
- Humans cannot detect it.

Zang et al., Dolphin Attack: Inaudible Voice Commands, ACM C 2017





Adversarial Robustness Toolbox

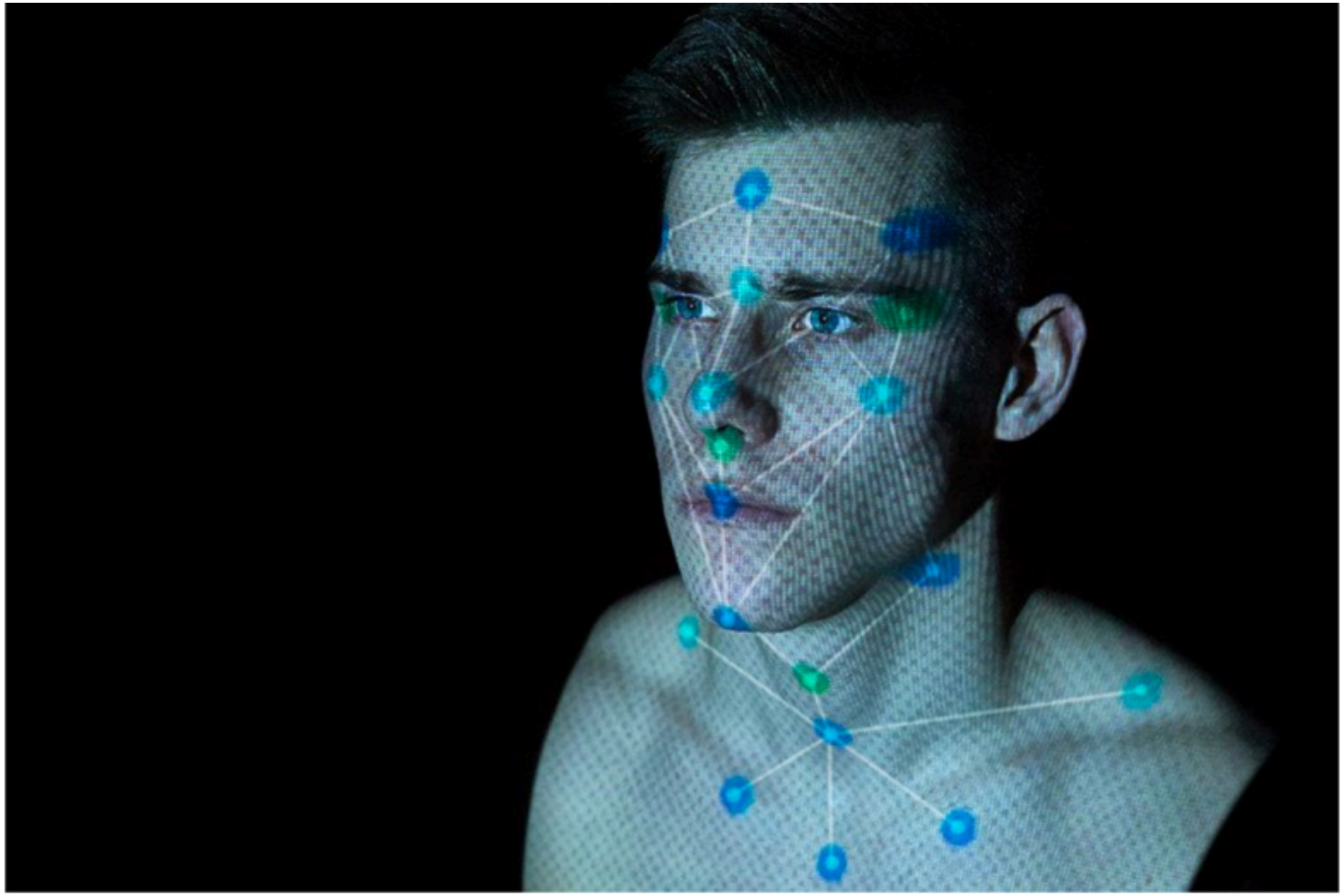
https://github.com/IBM/adversarial-robustness-toolbox

Attacks	Defenses
DeepFool	Feature Squeezing
Fast Gradient Method	Spatial Smoothing
Jacobian Saliency Map	Label Smoothing
NewtonFool	Adversarial Training
Universal Perturbation	Virtual Adversarial Training
C&W Attack	Gaussian Augmentation
Virtual Adversarial Method	

Frameworks	Metrics
TensorFlow	Loss sensitivity
Keras	Empirical robustness
PyTorch (soon)	CLEVER
MXNet (soon)	

Bias

COMPAS algorithm widely used in US is racially biased



John Lamb/Getty

flagging black people 45% vs. white people 24% for risk of recommitting crime

The problem of racist AI is not always a problem of the AI. It is the problem of a racist world (moral-robots.com)

AI is just a mirror that refuses to be sugar-coated (moral-robots.com)

82 percent of enterprises are considering AI deployments, but 60 percent fear liability issues. (IBMs Institute for Business Value)

IBM launches tools to detect AI fairness, bias and open sources some code

Bias detection and black box transparency is limiting AI deployments at scale, argues IBM.

By Larry Dignan for Between the Lines | September 19, 2018 -- 04:01 GMT (12:01 GMT+08:00) | Topic: Artificial Intelligence

https://www.zdnet.com/article/ibm-launches-tools-to-detect-ai-fairness-bias-and-open-sources-some-code/

IBM Research Trusted AI

Home Demo Resources Community

AI Fairness 360 - Demo

Data Check Mitigate Compare

Back Next

2. Check bias metrics

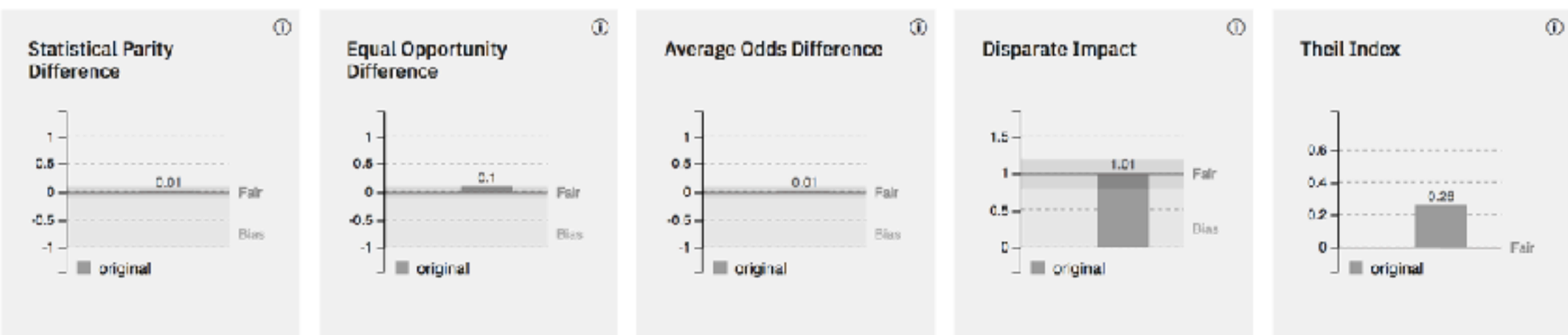
Dataset: German credit scoring
Mitigation: none

Protected Attribute: Sex

Privileged Group: ***Male***, Unprivileged Group: ***Female***

Accuracy with no mitigation applied is 76%

With default thresholds, bias against unprivileged group detected in 0 out of 5 metrics



○ **Reweighing**

Weights the examples in each (group, label) combination differently to ensure fairness before classification.

◉ **Adversarial Debiasing**

Learns a classifier that maximizes prediction accuracy and simultaneously reduces an adversary's ability to determine the protected attribute from the predictions. This approach leads to a fair classifier as the predictions cannot carry any group discrimination information that the adversary can exploit.

○ **Reject Option Based Classification**

Changes predictions from a classifier to make them fairer. Provides favorable outcomes to unprivileged groups and unfavorable outcomes to privileged groups in a confidence band around the decision boundary with the highest uncertainty.

4. Compare original vs. mitigated results

Dataset: German credit scoring

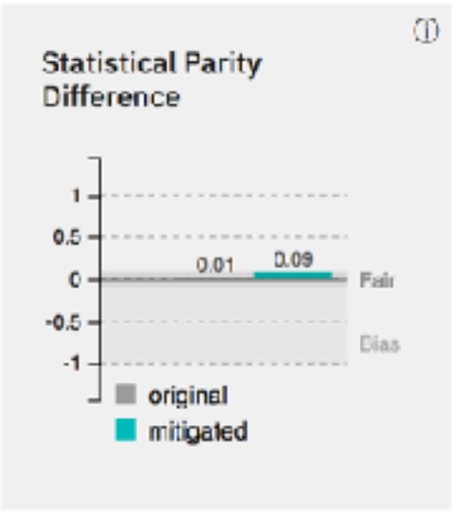
Mitigation: Adversarial Debiasing algorithm applied

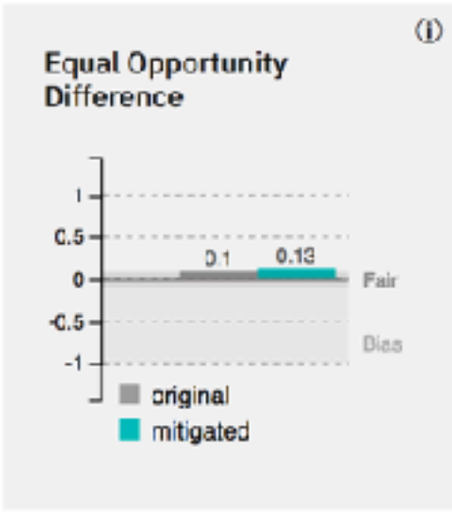
Protected Attribute: Sex

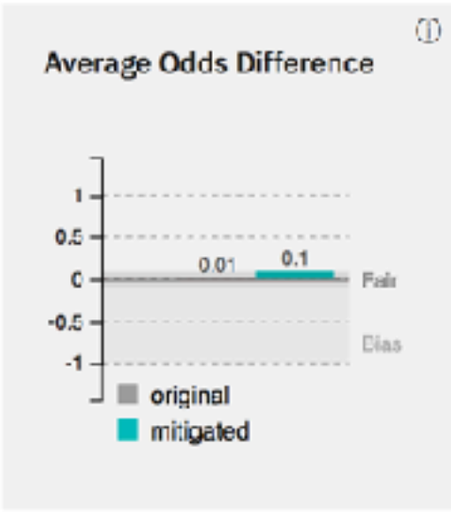
Privileged Group: *Male*, Unprivileged Group: *Female*

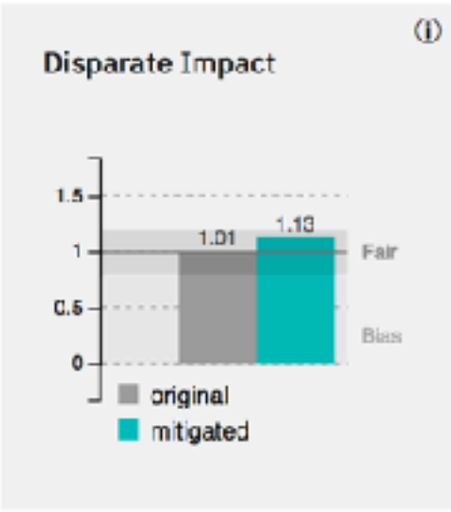
Accuracy after mitigation changed from 76% to 62%

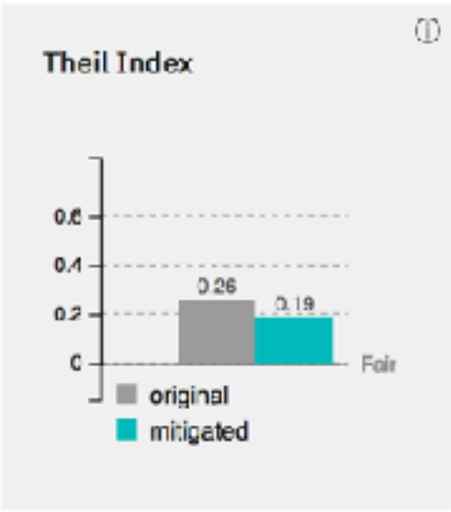
Bias against unprivileged group unchanged after mitigation (0 of 5 metrics indicate bias)







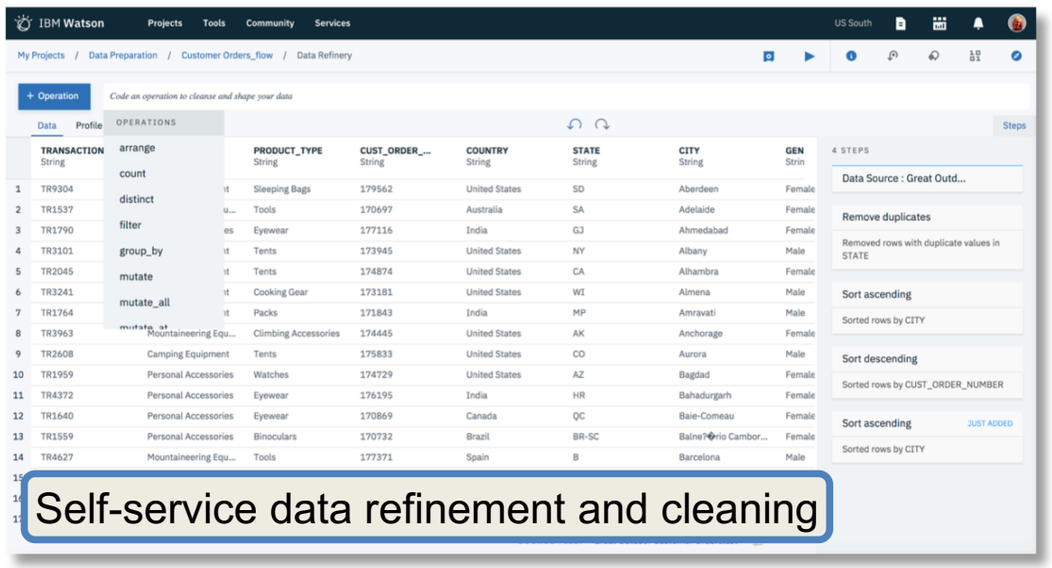




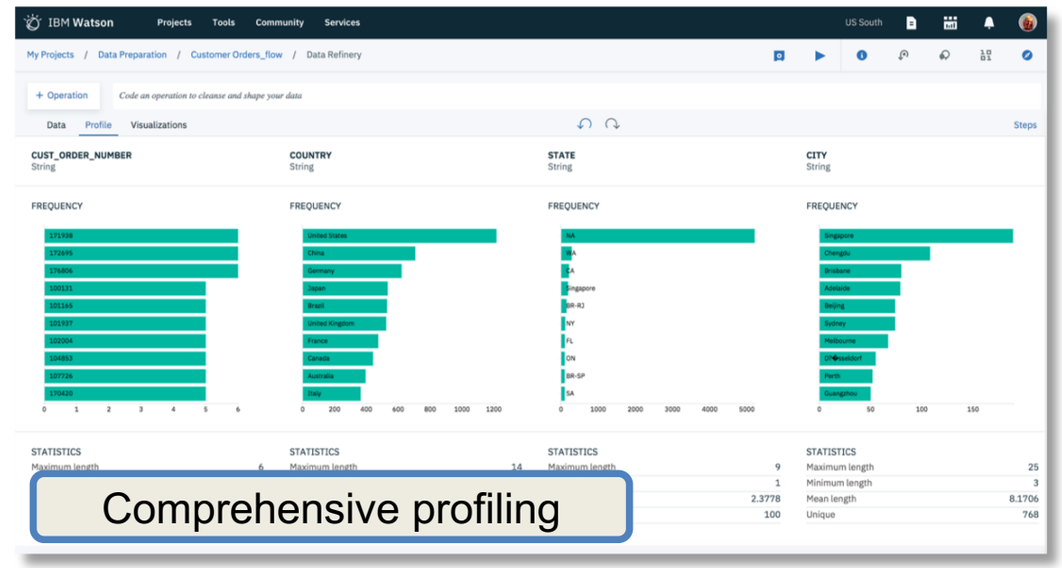
IBM Watson Studio

Data Refinery

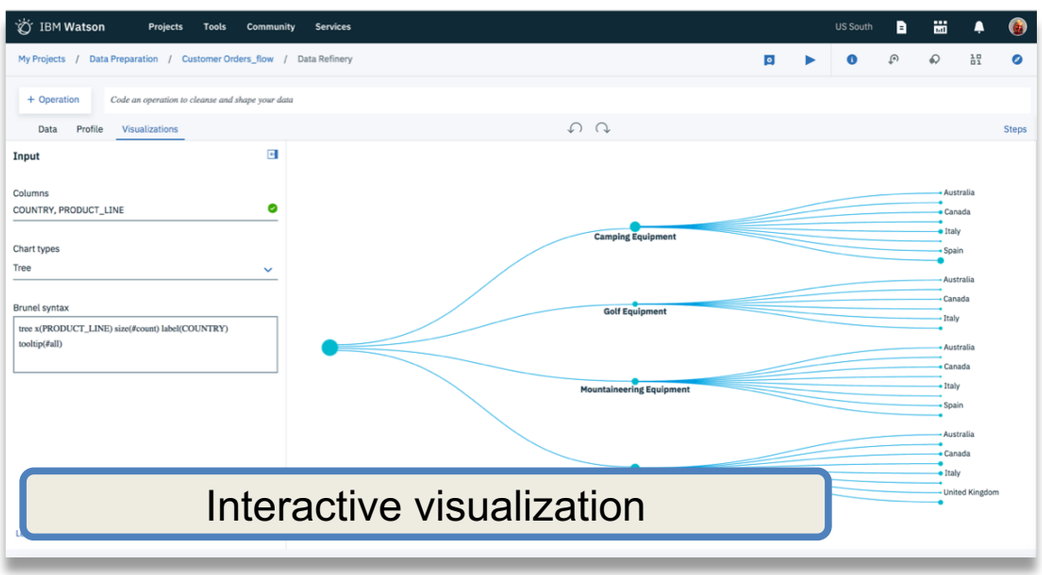
Making data fit for use



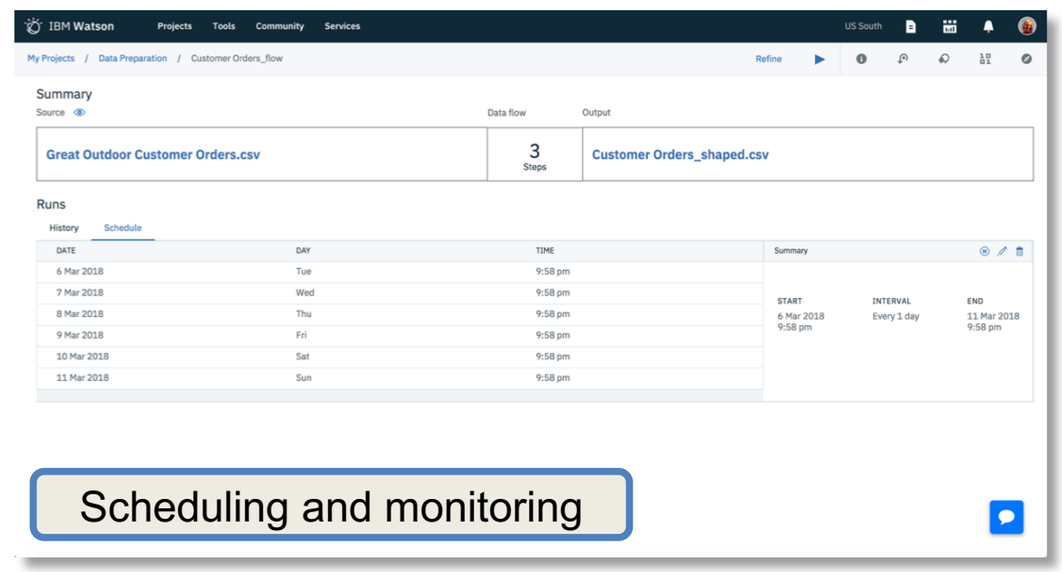
Self-service data refinement and cleaning



Comprehensive profiling

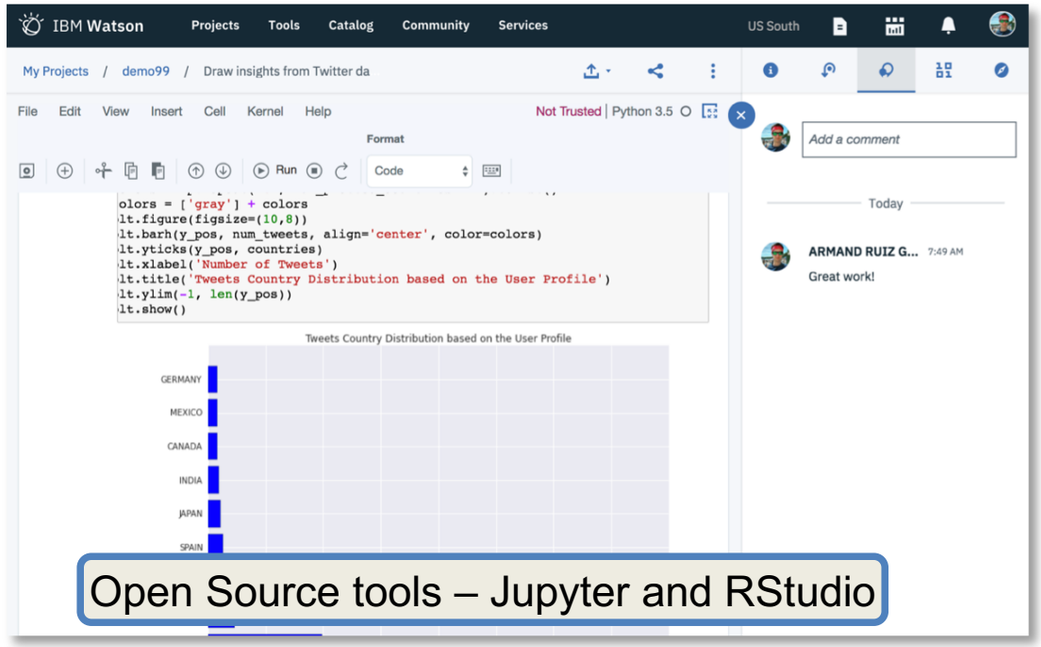


Interactive visualization

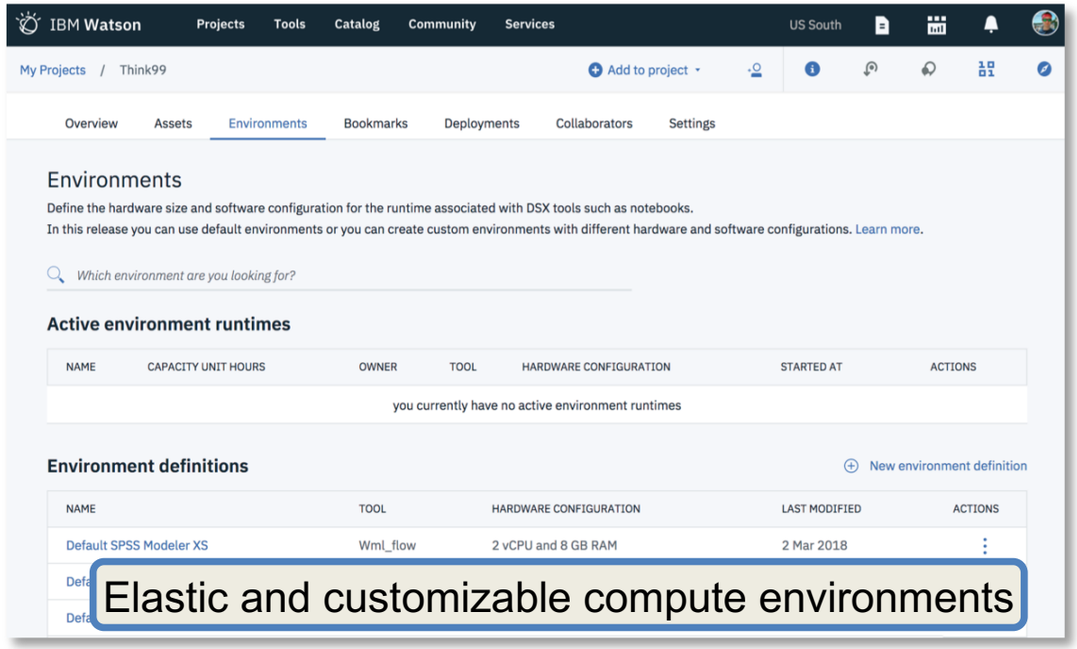


Scheduling and monitoring

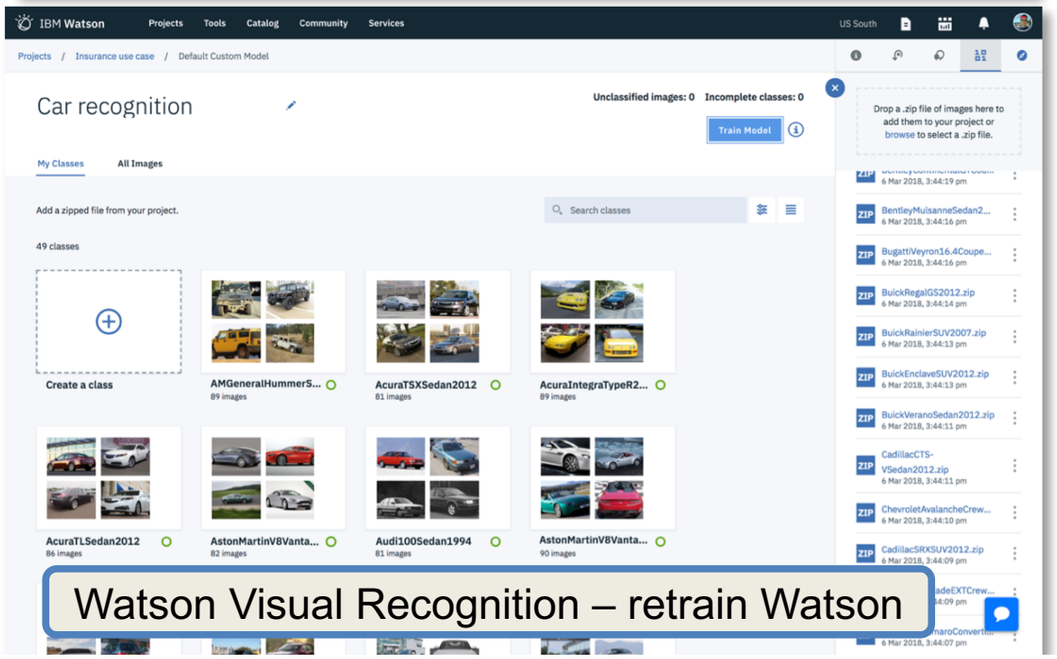
Watson Studio



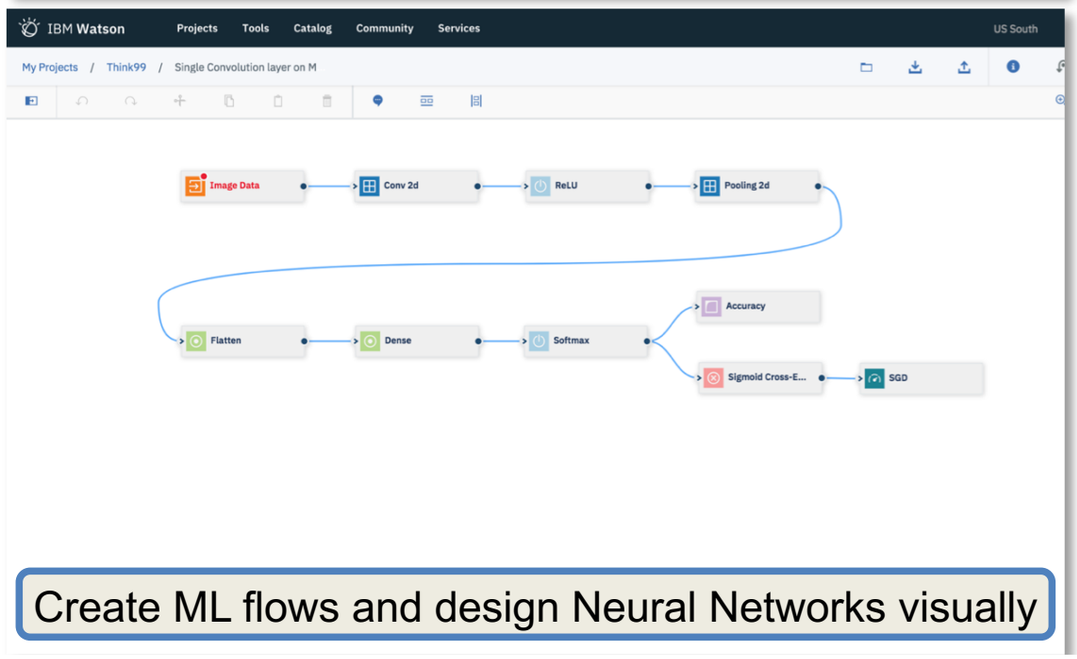
Open Source tools – Jupyter and RStudio



Elastic and customizable compute environments



Watson Visual Recognition – retrain Watson

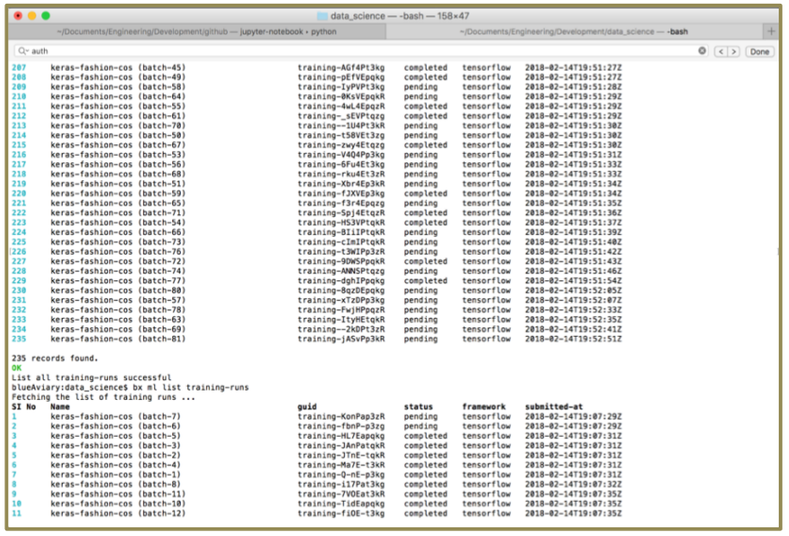


Create ML flows and design Neural Networks visually

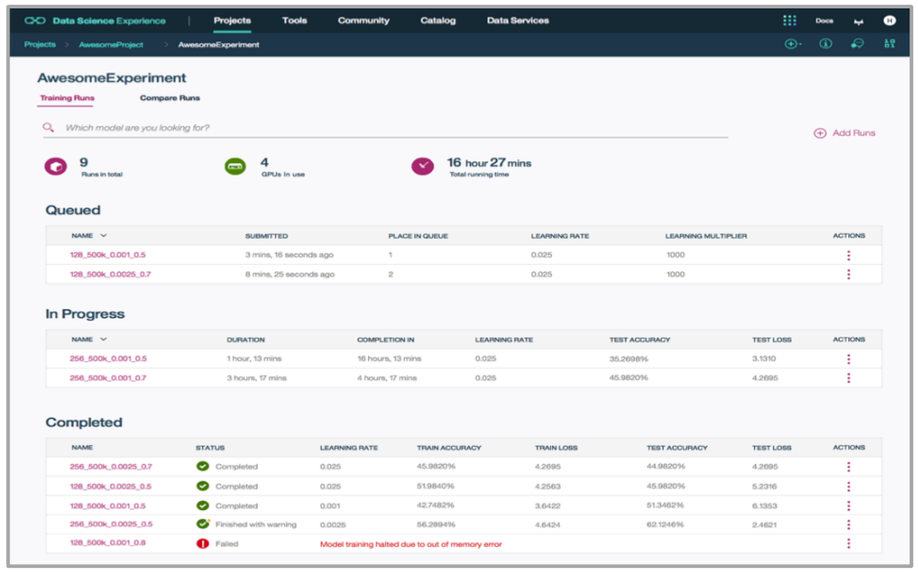
Watson Machine Learning

Simplifying deployment & management of ML models in production apps

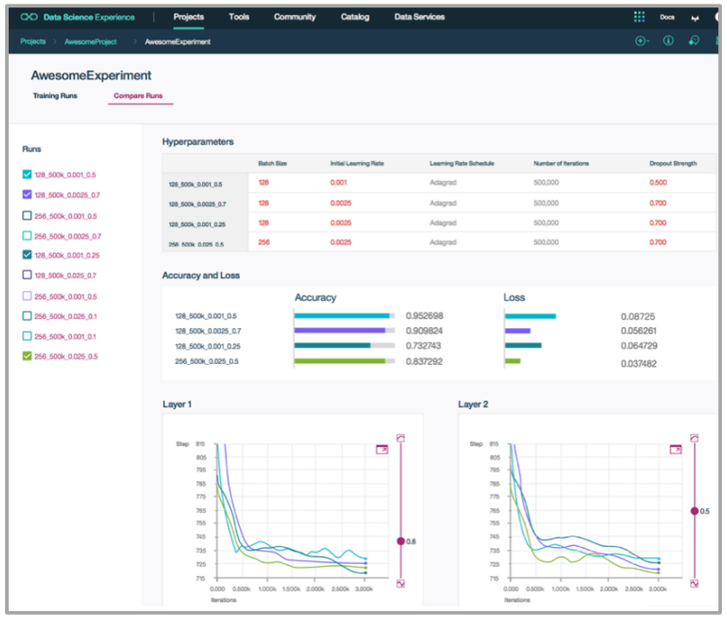
- Python client
- command line interface (CLI)
- UI



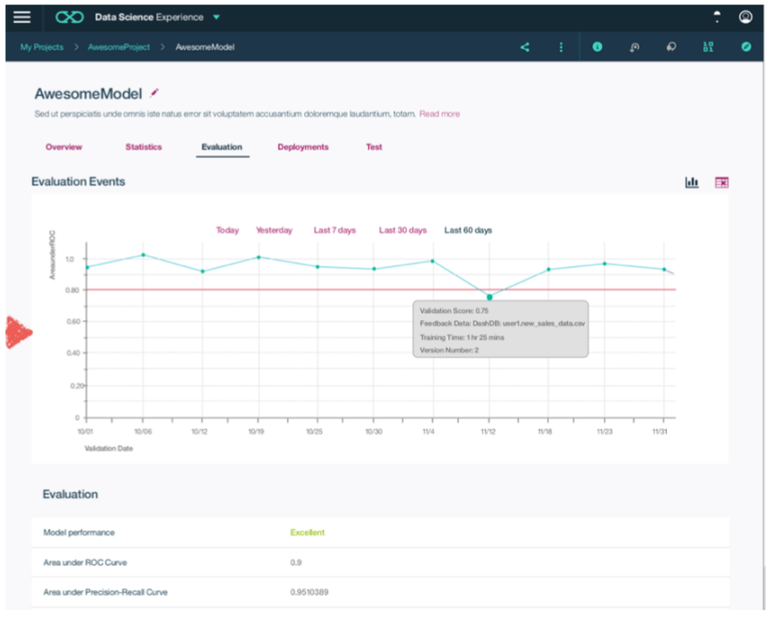
- parallel neural networks training across NVIDIA GPUs
- Pay per use
- Auto-deallocation



- Monitor batch training experiments
- compare cross-model performance

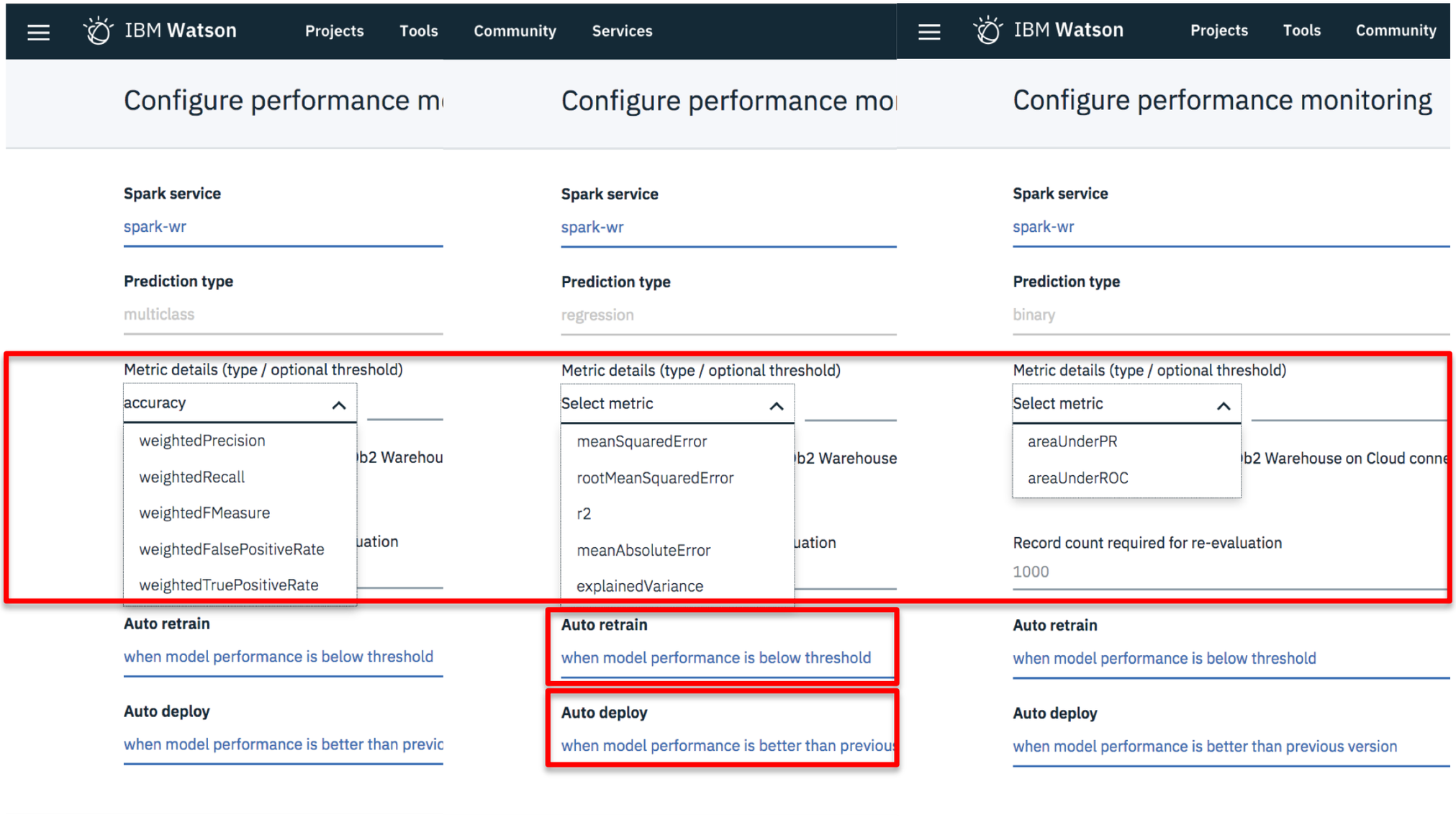


- Model deployment
- Model monitoring and performance evaluation
- continuous learning and retraining



Model Performance Evaluation

Supports Spark 2.1



24
Runs in total

1 hr, 34 min, 33 sec
Total running time

Queued

NAME	SUBMITTED
No training runs found.	

In progress

NAME	DURATION	LOSS	VAL_ACC	VAL_LOSS	ACC
training-VcXNl-RmR	0 hr, 5 min, 46 sec	1.707	0.714	4.115	0.358
training-sHzDl-gmR	0 hr, 5 min, 6 sec	2.322	0.439	4.172	0.137
training-3oQH_-Rmg	0 hr, 5 min, 13 sec	1.401	0.678	4.822	0.463

Completed

NAME	STATUS	DURATION	LOSS	VAL_ACC	VAL_LOSS	ACC	ACTIONS
training-c-_Ilagig	completed	0 hr, 6 min, 46 sec	0.820	0.814	2.908	0.692	
training-YEvSL-gmR	completed	0 hr, 6 min, 20 sec	0.446	0.860	2.208	0.837	

Overview Data Kernels Discussion Leaderboard Rules Team My Submissions **Late Submission**

3098	▼ 343	Bowman	0.821683	47	3y
3099	▲ 53	a2b	0.821677	1	3y
3100	▲ 2	Jelmer Krauwer	0.821676	4	3y
3101	▼ 227	Cory Lesmeister	0.821674	6	3y
3102	▲ 66	A.M.A	0.821672	3	3y
3103	▲ 1	Xianen	0.821670	4	3y
3104	▲ 19	lo Ko	0.821654	8	3y
3105	▲ 124	Pakoun	0.821637	24	3y
3106	▼ 543	Romeo Kienzler	0.821620	41	3y
3107	▼ 166	Jacob Zweig	0.821620	2	3y
3108	▲ 131	RobertSalzwedel	0.821620	7	3y
3109	▲ 21	data-science	0.821583	20	3y
3110	▼ 129	Sébastien Romano	0.821579	25	3y
3111	▼ 389	Beeshere	0.821575	28	3y
3112	▲ 251	EdwardLandesberg	0.821573	2	3y
3113	▲ 84	Donlo Wong	0.821559	7	3y
3114	▼ 3	ShawnPoindexter	0.821556	3	3y

kaggle Search Competitions Datasets Kernels Discussion Learn •••

Santander Customer Satisfaction

Which customers are happy customers?

$60,000 · 5,123 teams · 3 years ago

Overview Data Kernels Discussion Leaderboard Rules Team My Submissions Late Submission

Public Leaderboard Private Leaderboard

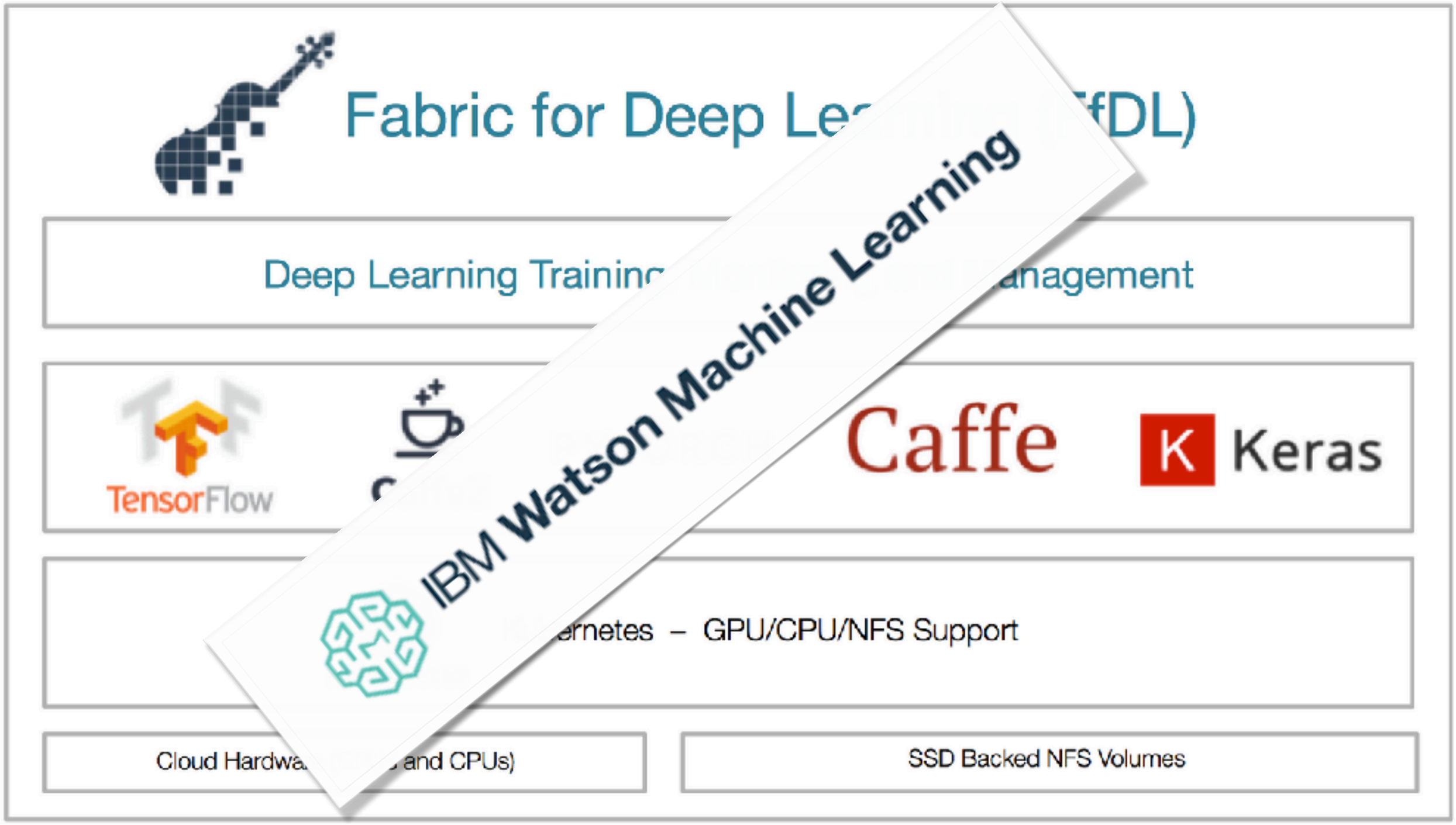
The private leaderboard is calculated with approximately 50% of the test data.

This competition has completed. This leaderboard reflects the final standings.

Refresh

In the money Gold Silver Bronze

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲3	Shize & Nir			0.829072	298	3y
2	▲848	kg_joi			0.829062	35	3y
3	▲6	#1 Leustagos		+8	0.828530	227	3y
4	▲1	Why so noise?			0.828466	195	3y


Fabric for Deep Le fDL)
Deep Learning Trainin anagement
TensorFlow
Caffe
Keras
IBM Watson Machine Learning
ernetes – GPU/CPU/NFS Support
Cloud Hardwa and CPUs)
SSD Backed NFS Volumes

IBM Code Model Asset Exchange

A place for developers to find and use free and open source deep learning models.

All models

Inception-ResNet-v2

Identify objects in images using a third-generation deep residual network.

Get this model

Places365 CNN

Classify images according to the place/location labels in the Places365 data set.

Get this model

Image Caption Generator

Generate captions that describe the contents of images.

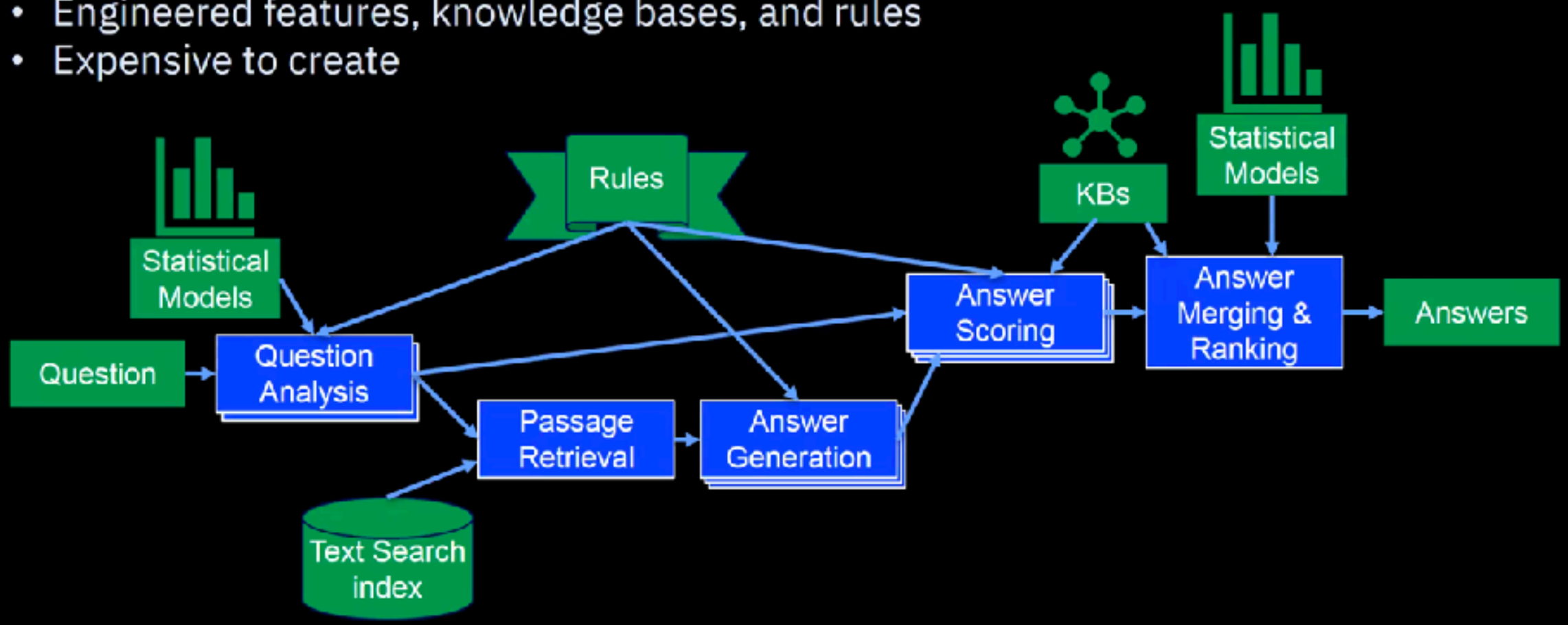
Get this model

When will we have a generic and universal question answering machine?

Is DeepLearning only sufficient for Question Answering in 2019?

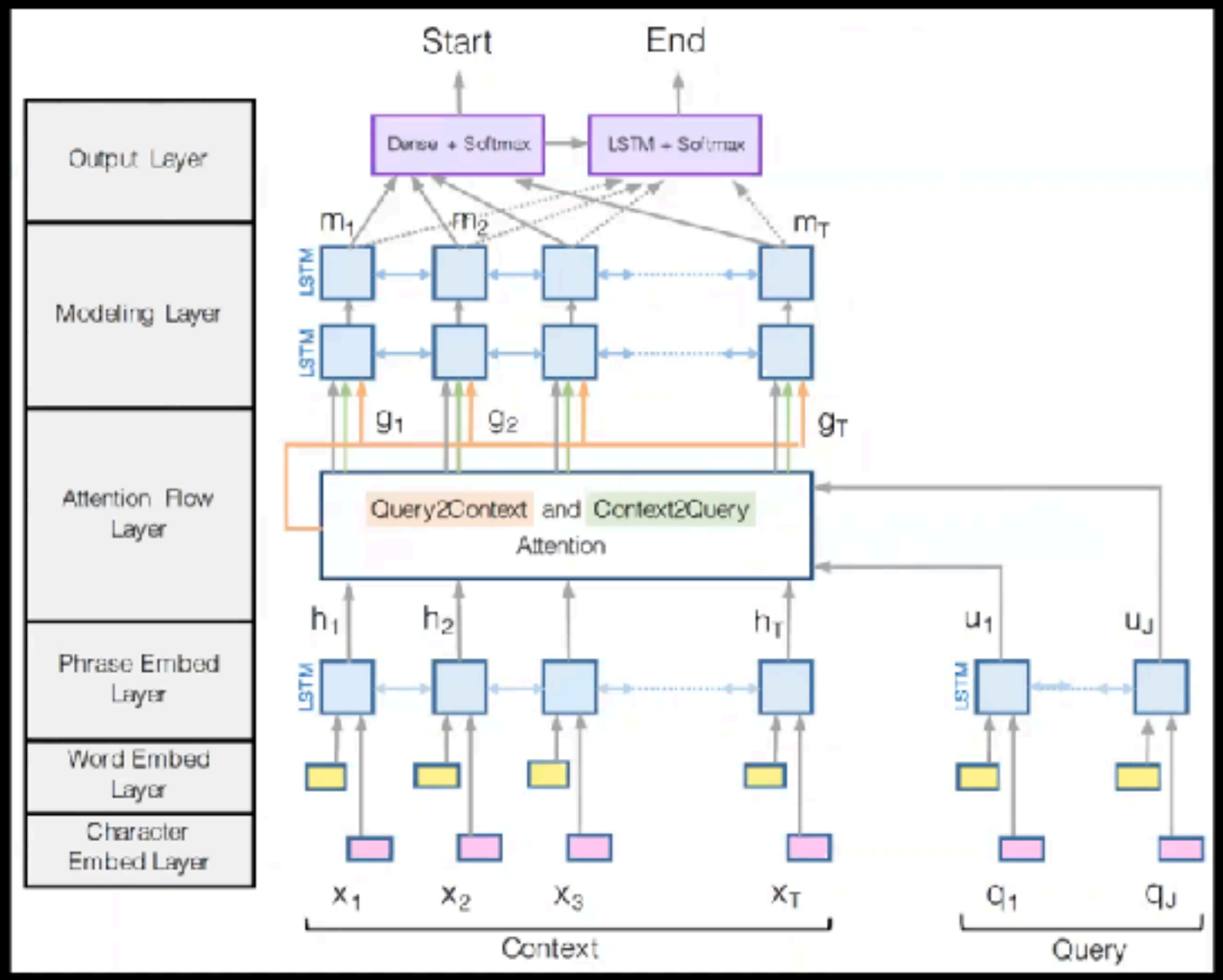
DDQA: Multi-Strategy Factoid Question Answering

- DDQA = Discovery DeepQA
- Simpler version of IBM Watson 1.0, designed for cloud
- Engineered features, knowledge bases, and rules
- Expensive to create



Bidirectional Attention Flow

- Off-the-shelf Deep Neural Net
- Some engineering on structure
- No manually engineered features



Seo, M., et al. Bidirectional attention flow for machine comprehension. *ICLR 2017*.

Hypothesis

A system that excels at SQuAD will also excel at factoid question answering

Engineered Multi-Strategy

SQuAD

Factoid

Statistical Single-Strategy

SQuAD Results

	Exact Match	Mean Rec. Rank
Statistical	66%	71%
DDQA + Statistical	67%	73%

- The statistical system alone provides nearly all of the power.
- Adding DDQA provides very little benefit despite all of its great cos

Factoid-1527 Results

	Exact Match	Mean Rec. Rank
Statistical	15%	21%
DDQA + Statistical	47%	56%

- The statistical system alone provides very little power.
- Adding DDQA provides enormous benefit.

the future

AI that creates AI

Vehicle Repair Estimator

Neural Network Synthesis

Deploy | Download Model

STATUS

IBM AI OpenScale has successfully created a synthetic deep learning model with **NeuNetS**. Your NeuNetS model was trained with your data set and tested for accuracy, precision, and recall.

Prepare → Synthesize → Test → Deploy

ACCURACY	PRECISION	RECALL
93% .46	94% .47	92% .48

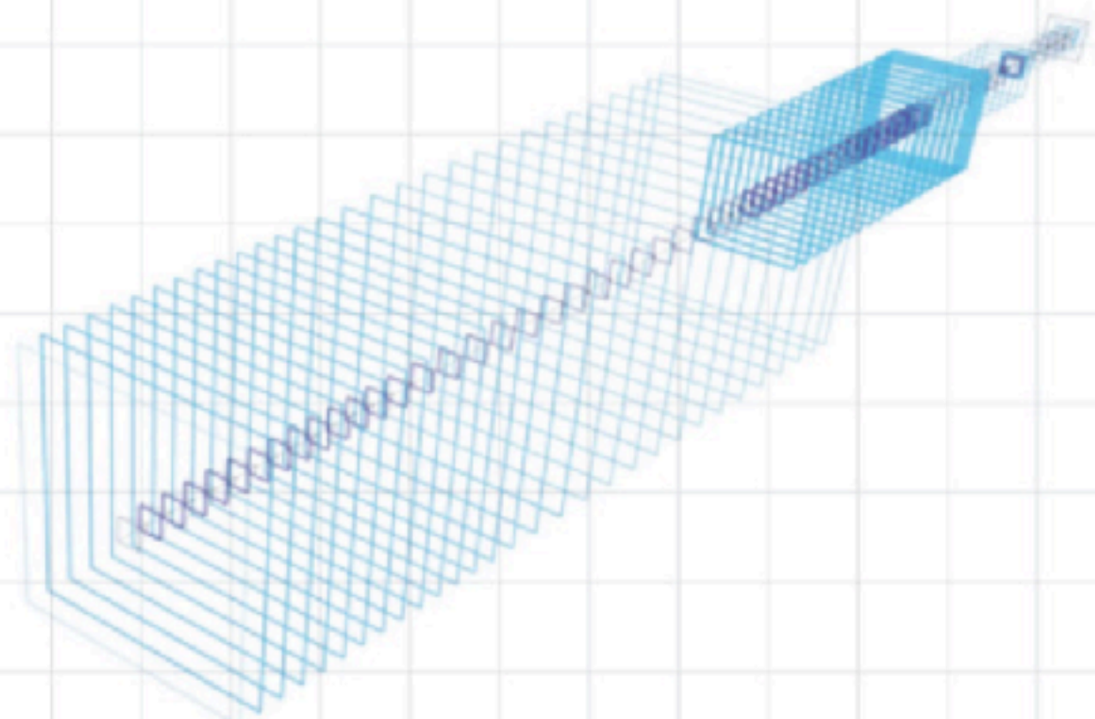
TRAINING DATA

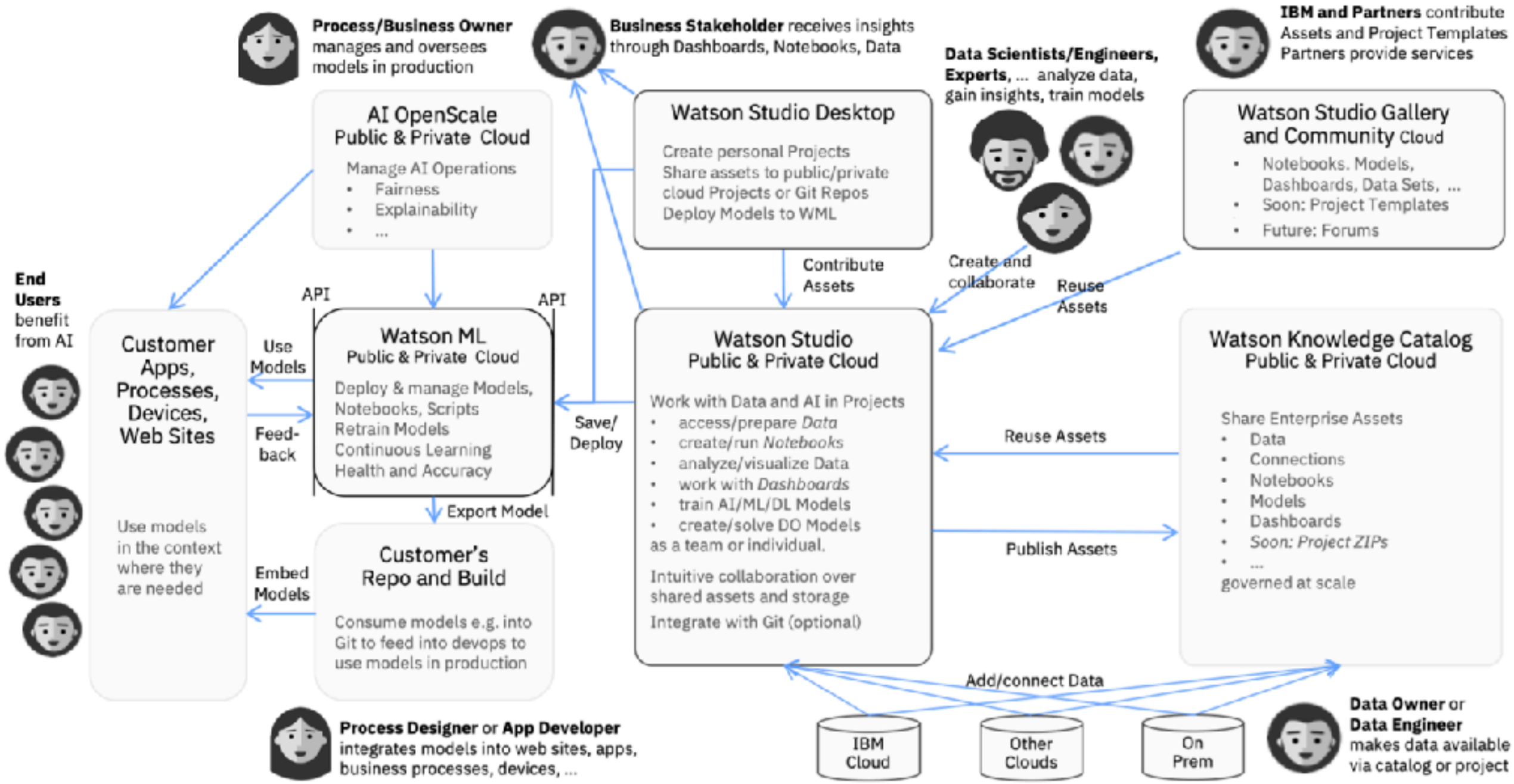
Source	Test Sample
TrainingSet.zip	25%

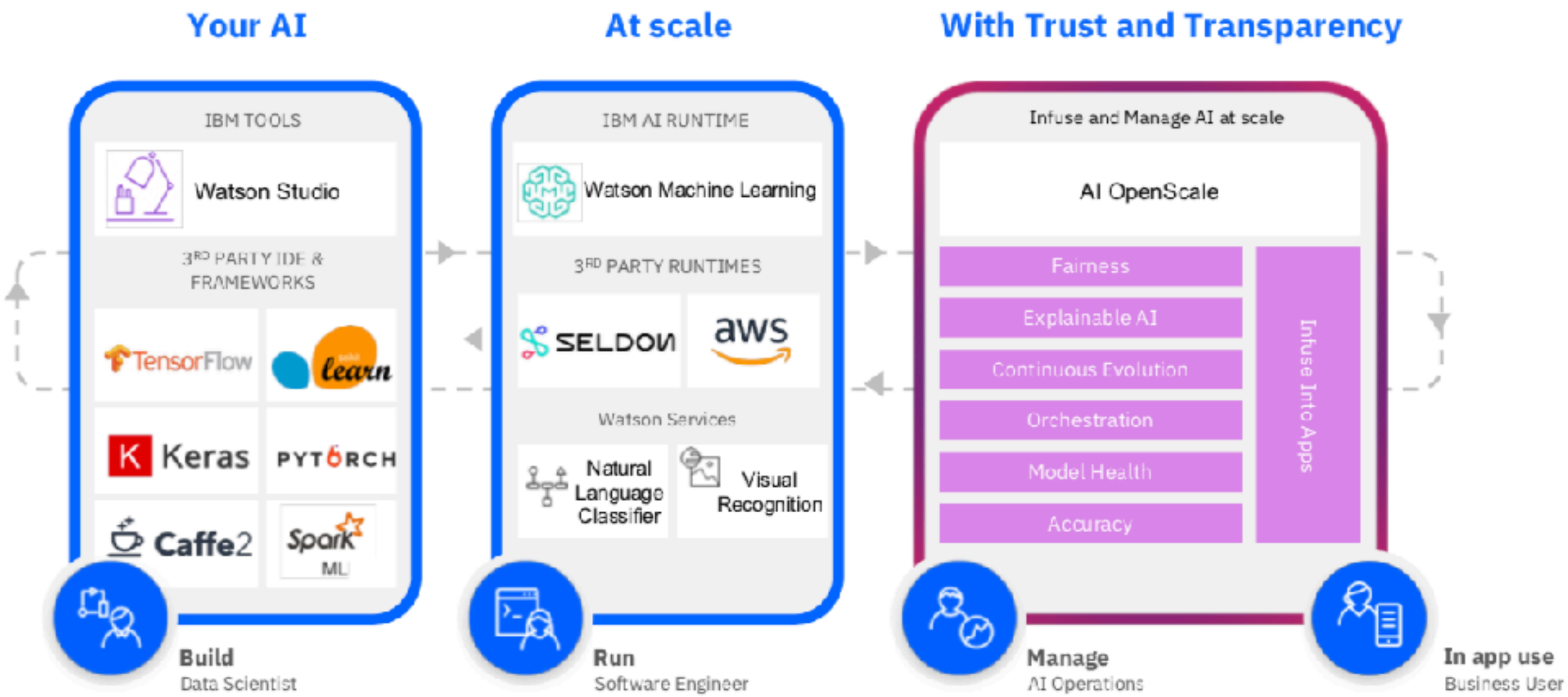
Size	Features	Labels
10,245	256	765

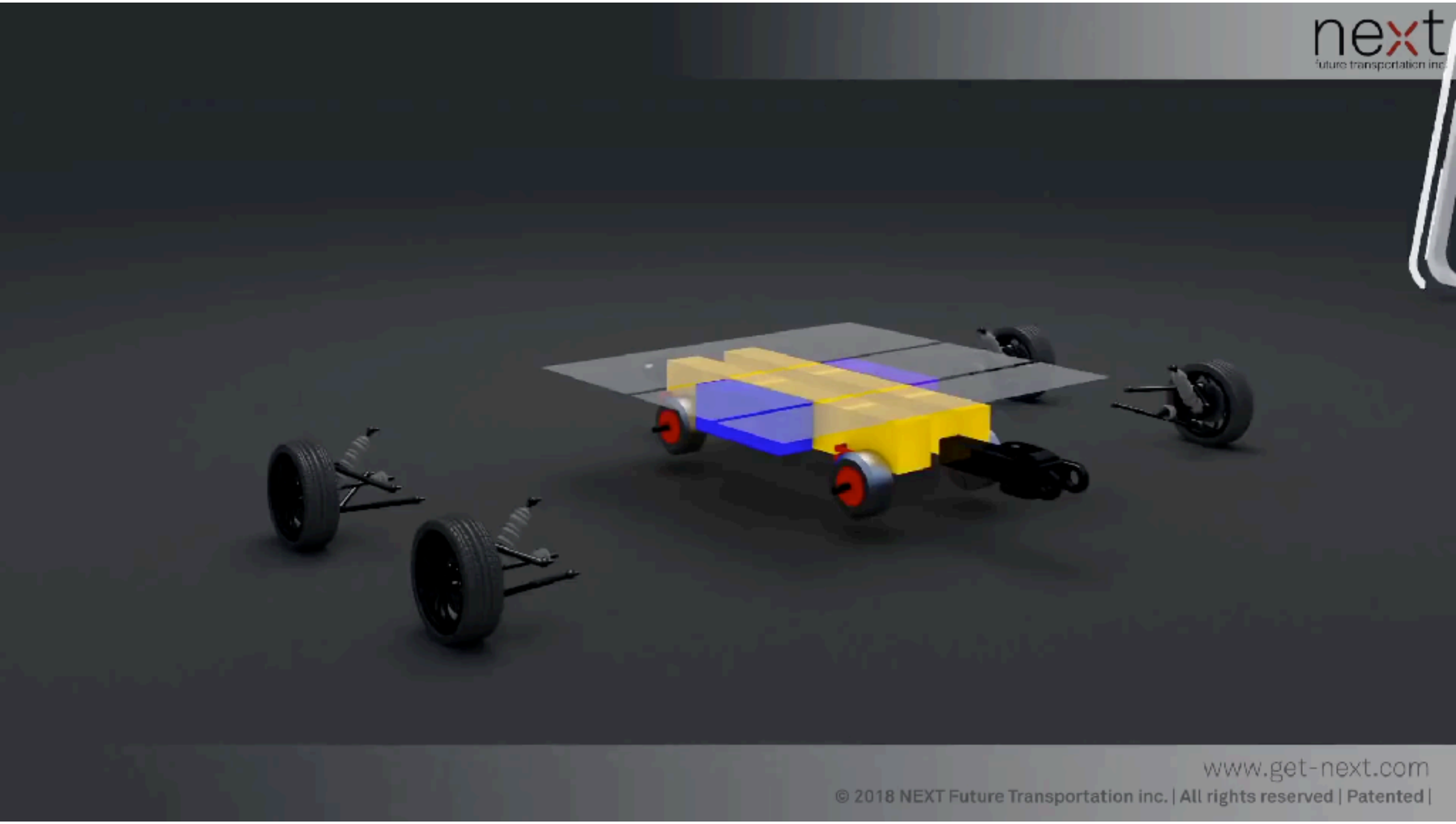
DEPLOYMENT

AIOS Instance	WML Instance	Model Type	Framework
aios-tx	wml-ws	wml-1.1	Tensorflow 1.5




Process/Business Owner manages and oversees models in production
Business Stakeholder receives insights through Dashboards, Notebooks, Data
Data Scientists/Engineers, Experts, ... analyze data, gain insights, train models
IBM and Partners contribute Assets and Project Templates
Partners provide services
AI OpenScale
Public & Private Cloud
Manage AI Operations
Fairness
Explainability
...
Watson Studio Desktop
Create personal Projects
Share assets to public/private cloud Projects or Git Repos
Deploy Models to WML
Watson Studio Gallery and Community Cloud
Notebooks, Models, Dashboards, Data Sets, ...
Soon: Project Templates
Future: Forums
Contribute Assets
Create and collaborate
Reuse Assets
End Users benefit from AI
Customer Apps, Processes, Devices, Web Sites
Use models in the context where they are needed
Use Models
Feed-back
API
API
Watson ML
Public & Private Cloud
Deploy & manage Models, Notebooks, Scripts
Retrain Models
Continuous Learning
Health and Accuracy
Save/ Deploy
Export Model
Watson Studio
Public & Private Cloud
Work with Data and AI in Projects
access/prepare Data
create/run Notebooks
analyze/visualize Data
work with Dashboards
train AI/ML/DL Models
create/solve DO Models
as a team or individual.
Intuitive collaboration over shared assets and storage
Integrate with Git (optional)
Reuse Assets
Publish Assets
Watson Knowledge Catalog
Public & Private Cloud
Share Enterprise Assets
Data
Connections
Notebooks
Models
Dashboards
Soon: Project ZIPs
...
governed at scale
Embed Models
Customer's Repo and Build
Consume models e.g. into Git to feed into devops to use models in production
Process Designer or App Developer integrates models into web sites, apps, business processes, devices, ...
Add/connect Data
IBM Cloud
Other Clouds
On Prem
Data Owner or Data Engineer makes data available via catalog or project


Your AI
At scale
With Trust and Transparency
IBM TOOLS
Watson Studio
3RD PARTY IDE & FRAMEWORKS
TensorFlow
learn
Keras
PYTORCH
Caffe2
Spark ML
IBM AI RUNTIME
Watson Machine Learning
3RD PARTY RUNTIMES
SELDON
aws
Watson Services
Natural Language Classifier
Visual Recognition
Infuse and Manage AI at scale
AI OpenScale
Fairness
Explainable AI
Continuous Evolution
Orchestration
Model Health
Accuracy
Infuse Into Apps
Build
Data Scientist
Run
Software Engineer
Manage
AI Operations
In app use
Business User


next
future transportation inc.
www.get-next.com
© 2018 NEXT Future Transportation inc. | All rights reserved | Patented |

THE Late Late SHOW
WITH JAMES CORDEN
olli
olli

PM

Boston Dynamics

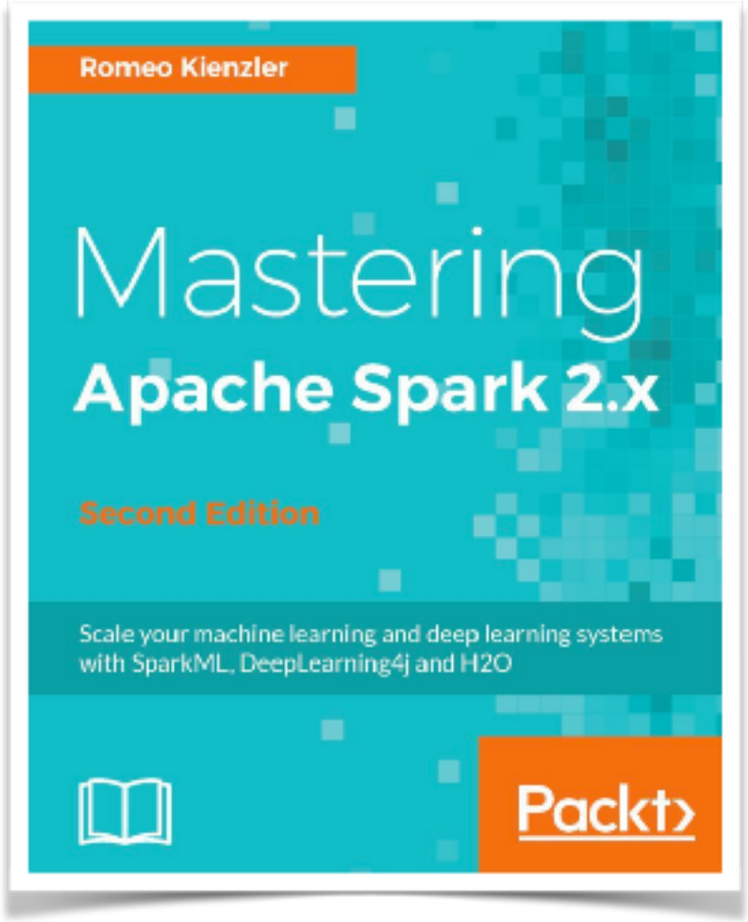
UBER
HACKED

Openness matters…

Center for Open-Source Data & AI Technologies

Improving the Enterprise AI Lifecycle in Open Source

- **Apache Spark (50,000 lines of code)**
- **TensorFlow**
- **Keras**
- **Apache SystemML (65,000 lines of code)**
- **Apache Arrow**
- **Apache Bahir**
- **Apache Toree**
- **Apache Zeppelin**
- **Apache Livy**
- **Fabric for DeepLearning (FfDL)**



OpenData

Speech Recognition: **Proliferation**

Echo

Cortana

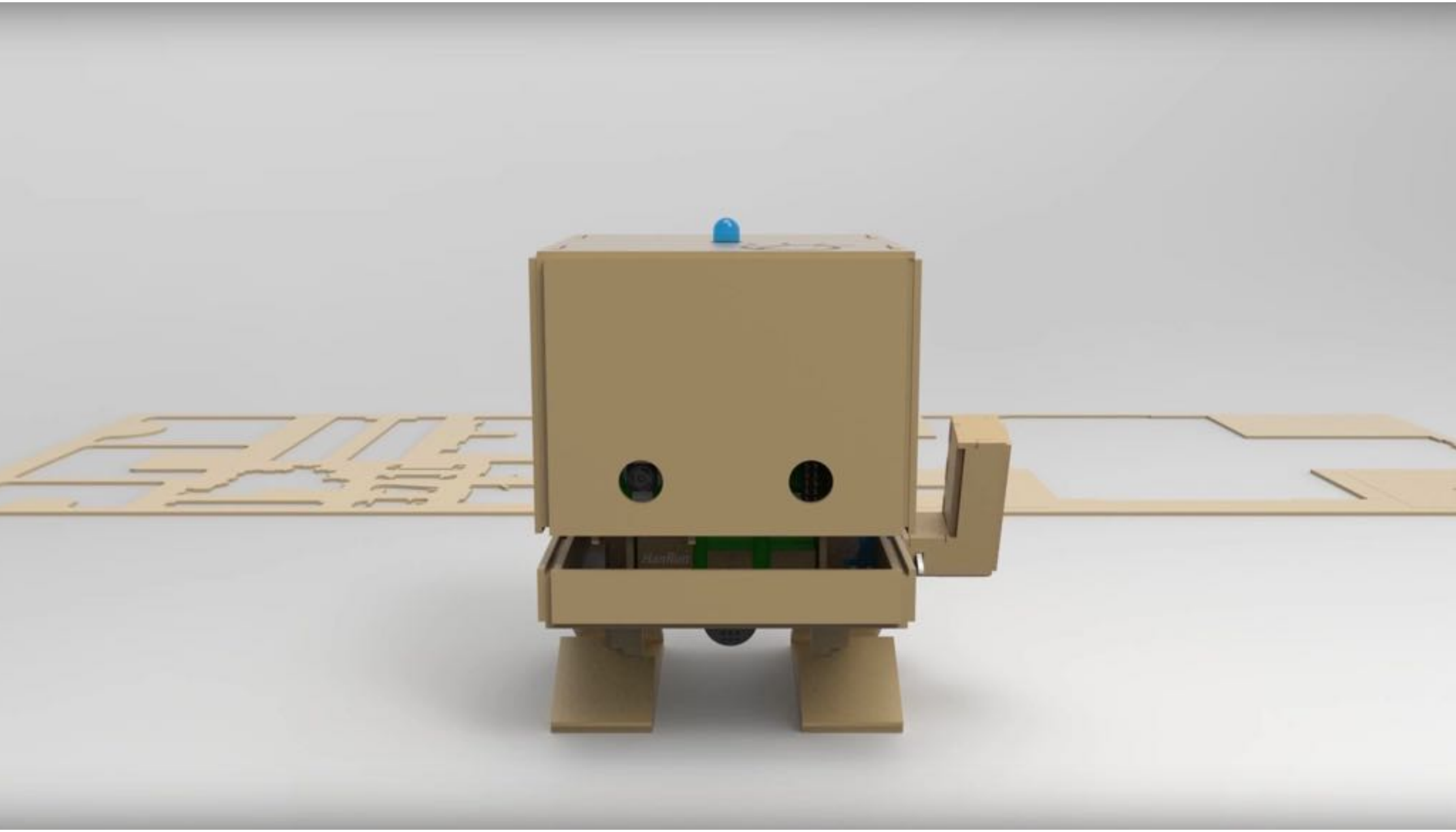
Mycroft

Siri

Google Home

Apple HomePod

3 teaspoons in
1 tablespoon



ibmtjbot / tjbot

Watch 61 | Star 386 | Fork 224

Code | Issues 9 | Pull requests 3 | Projects 0 | Wiki | Insights

IBM TJBot https://ibmtjbot.github.io

tjbot raspberry-pi watson prototyping cognitive

262 commits | 3 branches | 5 releases | 10 contributors | Apache-2.0

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

jweisz updating `apikey` to be consistent with tjbotlib 1.5.x — Latest commit 2dabcb9 on Oct 17

bootstrap	updating `apikey` to be consistent with tjbotlib 1.5.x	a month ago
featured	Update README.md	4 months ago
images	updated the wiring diagram	2 years ago
recipes	make sure tjbot always speaks a response	5 months ago
CONDUCT.md	cleaning up permission bits	2 years ago
CONTRIBUTING.md	Broken Links	a year ago
DCO1.1.txt	first push .. go love	2 years ago
LICENSE	Update LICENSE	a year ago
MAINTAINERS.md	revert to initial attribution method for tjbot. maintainers listed as...	2 years ago
README.md	Update README.md	a month ago

Common Voice
moz://a

CONTRIBUTE DATASETS LANGUAGES PROFILE

0 1 ENGLISH

Don't see your language on Common Voice yet?

Request a Language

Launched

Search

For these launched languages the website has been successfully localized, and has enough sentences collected, to allow for ongoing *Speak* and *Listen* contribution.

English	*Speakers* 32,631	*Validated Hrs* 522h / 1.2k
German	*Speakers* 2,034	*Validated Hrs* 141h / 1.2k
Kabyle	*Speakers* 337	*Validated Hrs* 92h / 1.2k

CONTRIBUTE MITMACHEN TTEKKI

mozilla / DeepSpeech

Watch 559 | Star 8,457 | Fork 1,467

Code | Issues 119 | Pull requests 3 | Projects 3 | Wiki | Insights

A TensorFlow implementation of Baidu's DeepSpeech architecture

deep-learning machine-learning neural-networks tensorflow speech-recognition speech-to-text

1,316 commits | 14 branches | 20 releases | 55 contributors | MPL-2.0

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

kdavis-mozilla Merge pull request #1736 from mozilla/issue1735 ... Latest commit f5d3dcd 2 days ago

bin	Addressed review comment	2 days ago
data	Remove old versions of decoder binary files	13 days ago
doc	Adding streaming API Support to the GUI Tool	5 days ago
examples/vad_transcriber	Adding streaming API Support to the GUI Tool	5 days ago
images	Compressed gif	a year ago
native_client	Remove NodeJS v11 hack	5 days ago
taskcluster	Run CTC Decoder build on merge and expose artifacts	8 days ago
util	Preprocessing: use all available threads	3 days ago

To all our readers in Germany.

It's a little awkward, so we'll get straight to the point: on this Thursday we humbly ask you to protect Wikipedia's independence. We depend on donations averaging about € 21.48, but 99% of our readers don't give. If everyone reading this gave a small amount, we could keep Wikipedia thriving for years to come. The price of your Thursday coffee is all we need. When we made Wikipedia a non-profit, people warned us we'd regret it. But if Wikipedia became commercial, it would be a great loss to the world. Wikipedia is a place to learn, not a place for advertising. It unites all of us who love knowledge: contributors, readers and the donors who keep us thriving. The heart and soul of Wikipedia is a community of people working to bring you unlimited access to reliable, neutral information. Please take a minute to help us keep Wikipedia growing. Thank you! — *Jimmy Wales, Wikipedia Founder*

€ 2.0 Mio. Still missing: € 6.1 Mio.

DONATION ACCOUNT Wikimedia Foerdergesellschaft BIC BFSWDE33BER IBAN DE33 1002 0500 0001 1947 00 Where does my donation go?

one-time | monthly | yearly

5 € | 10 € | 20 € | 25 €

50 € | 100 € | Other amount

Direct Debit | Bank Transfer

Credit Card | PayPal

Proceed with the donation

IBM Advanced Data Science Specialization Certificate on Coursera

Fundamentals of Scalable Data Science
www.coursera.org/learn/ds

Advanced Machine Learning and Signal Processing
www.coursera.org/learn/advanced-machine-learning-signal-processing

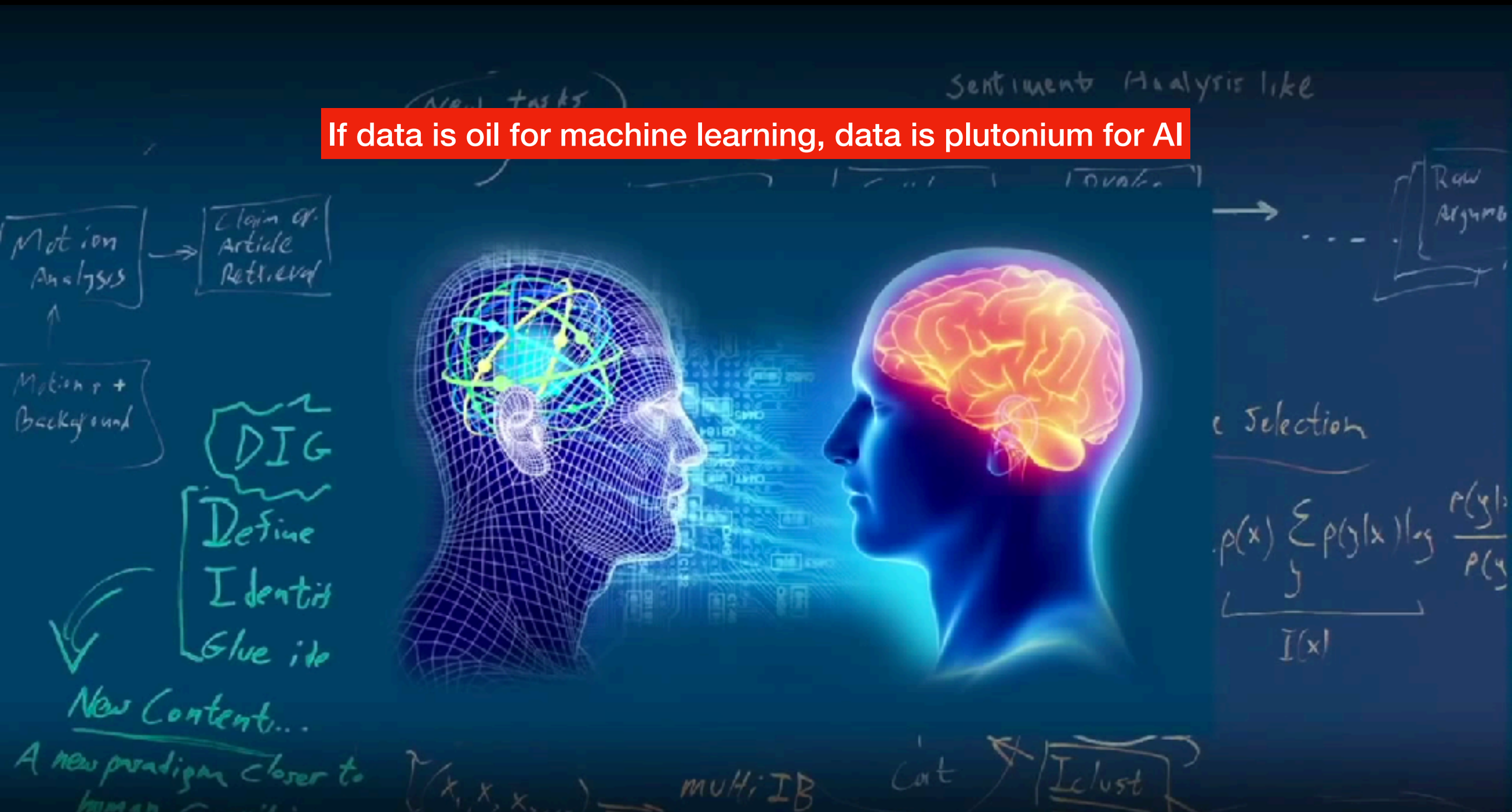
Applied AI with DeepLearning
www.coursera.ai/learn/ai

Advanced Data Science Capstone Project
https://www.coursera.org/learn/advanced-data-science-capstone

nopanic.ai/datascience

IBM Cloud Free Tier which enables all the open source presented here:

nopanic.ai/ibmcloud

your TODOs

- Take my courser courses => http://nopanic.ai/datascience
- Contribute to Mozilla Common Voice => @mikehenrty
- Encrypt your emails using PGP (or S/MIME)
- Use Signal or TOX for messaging
- Use open Android alternative
- Choose your apps and permissions carefully