Trends in AI, ML, DL

romeo kienzler, IBM

The .. singularity .. is a hypothetical point in the future when technological growth becomes uncontrollable and irreversible, resulting in unfathomable changes to human civilization.

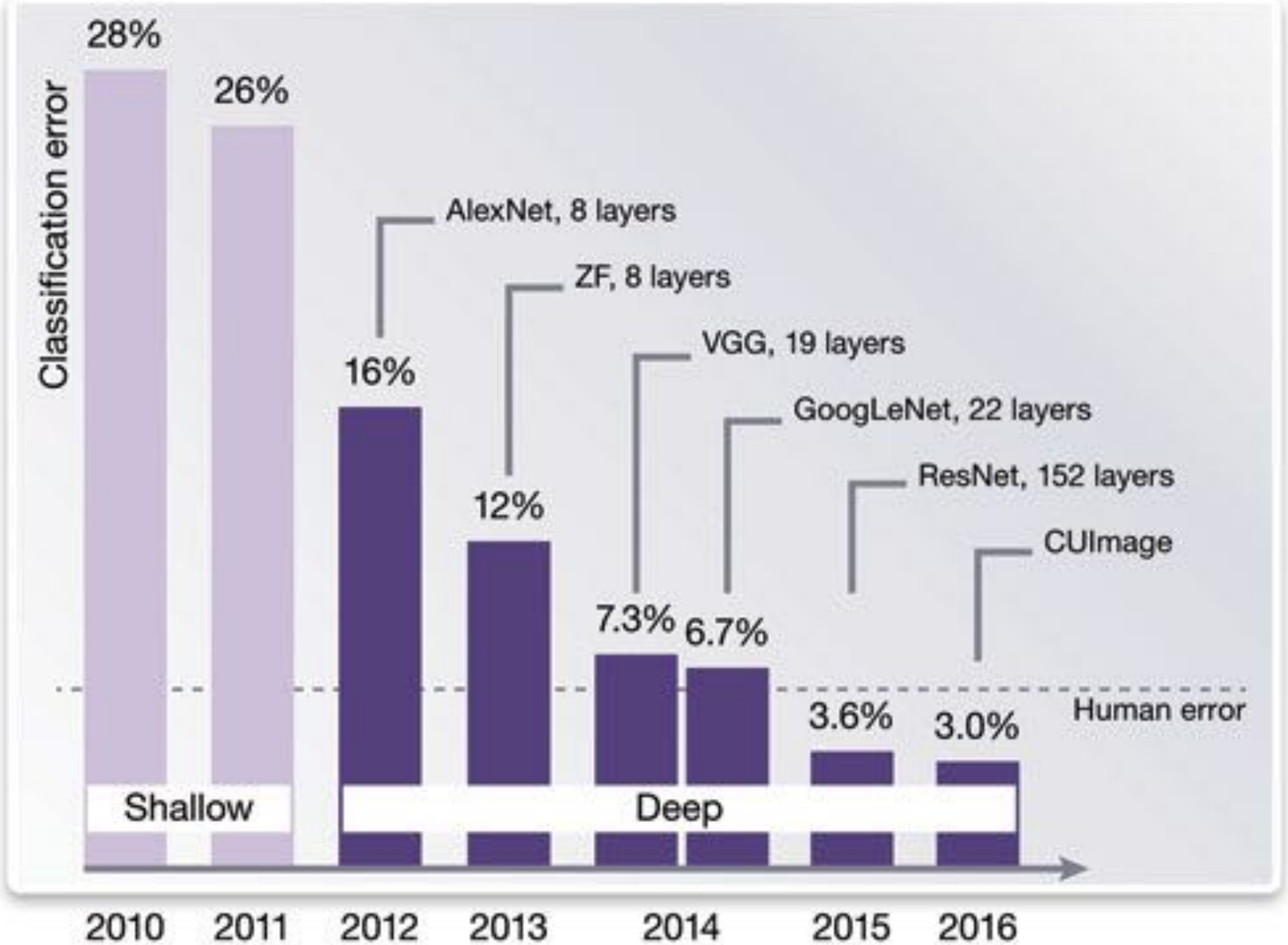
source: wikipedia

… **intelligence explosion**, an upgradable intelligent agent **..** would enter a "runaway reaction" of **self-improvement cycles**, **..** surpass all human intelligence**.**

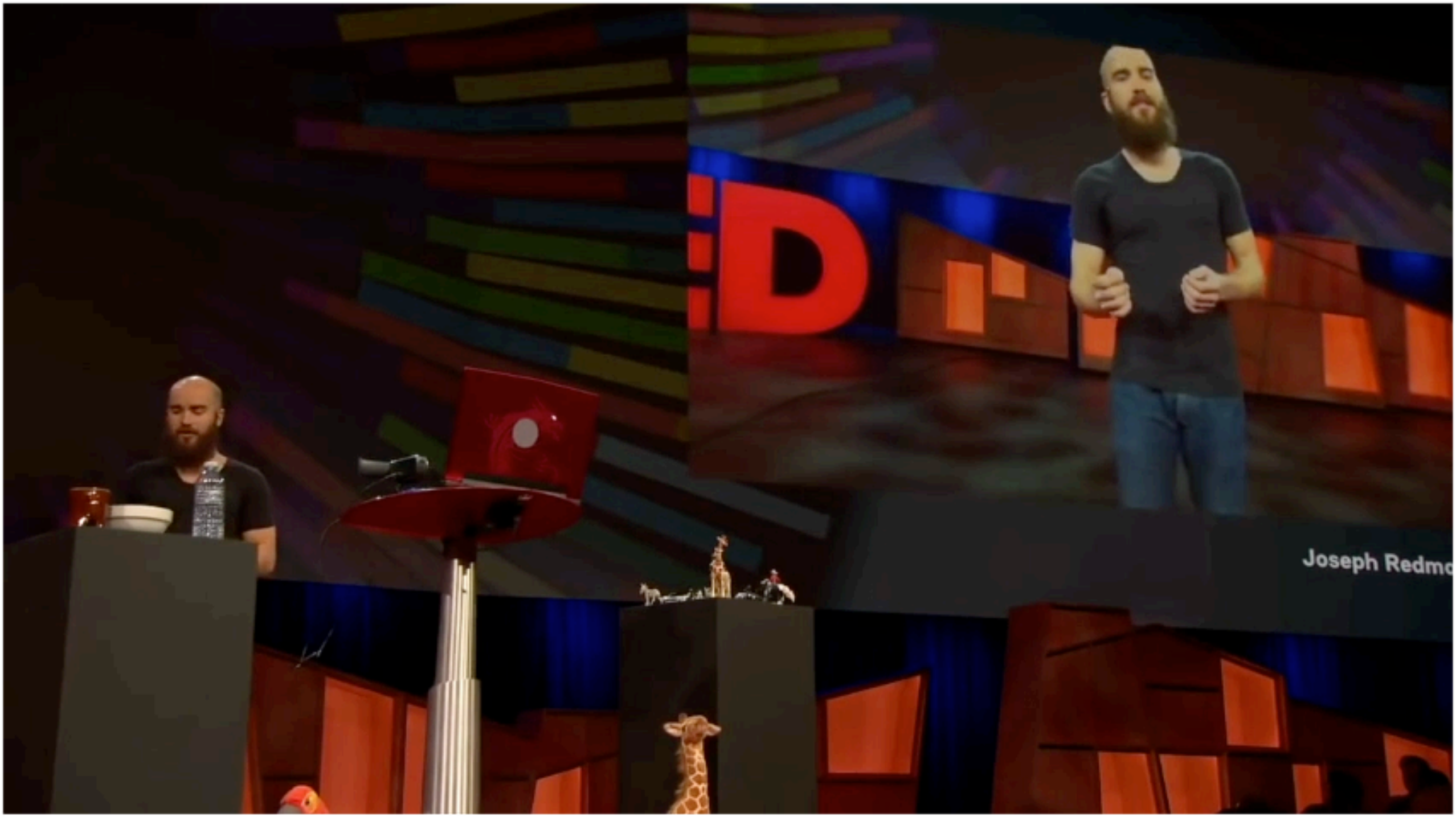
source: wikipedia

IBM’s AI journey

Life, the Universe and Everything



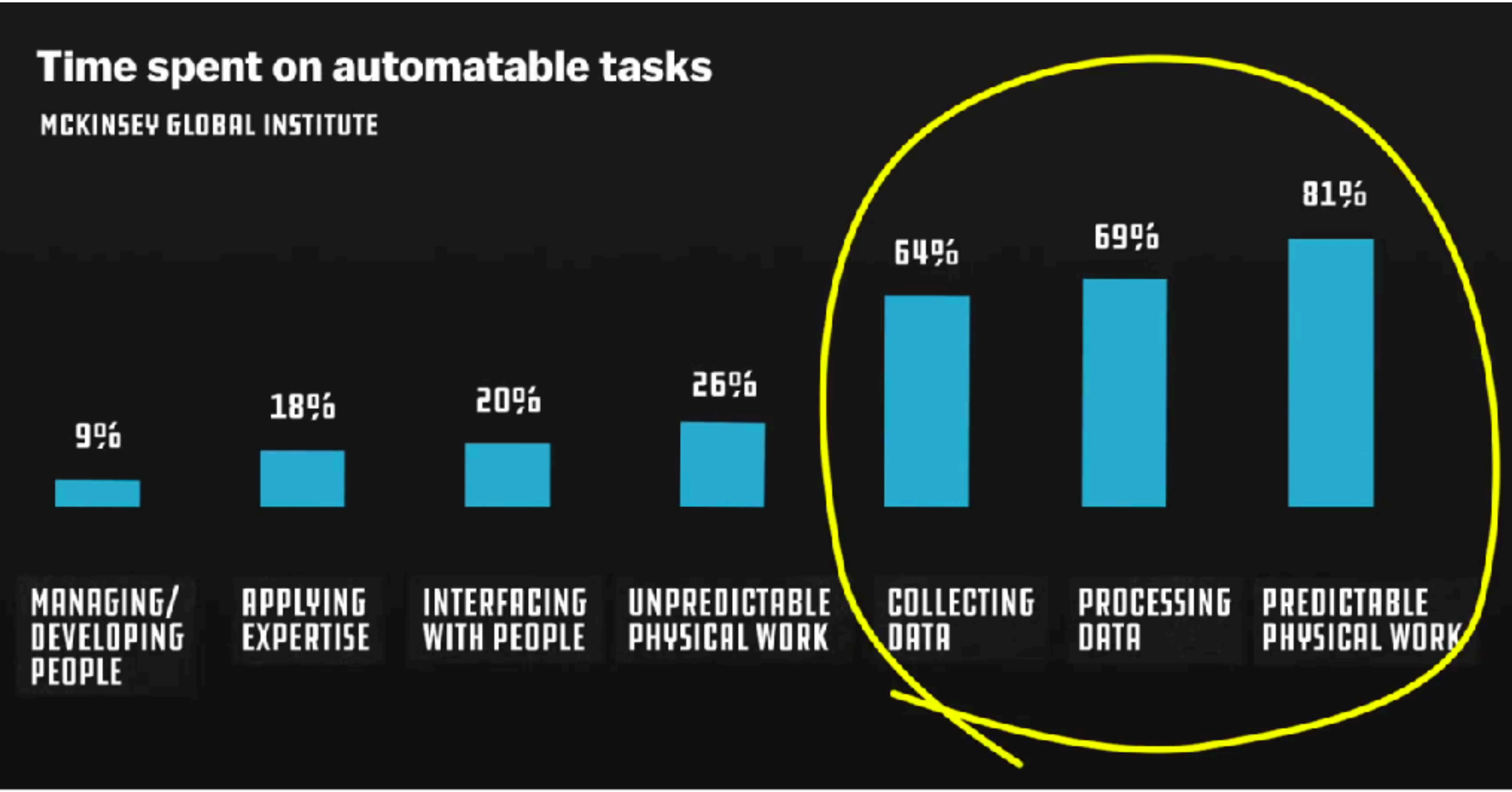
https://eecatalog.com/IoT/2017/10/05/as-embedded-vision-takes-off-software-frameworks-bear-watching/


D
Joseph Redmo


YOLO V2
http://pjreddie.com/yolo

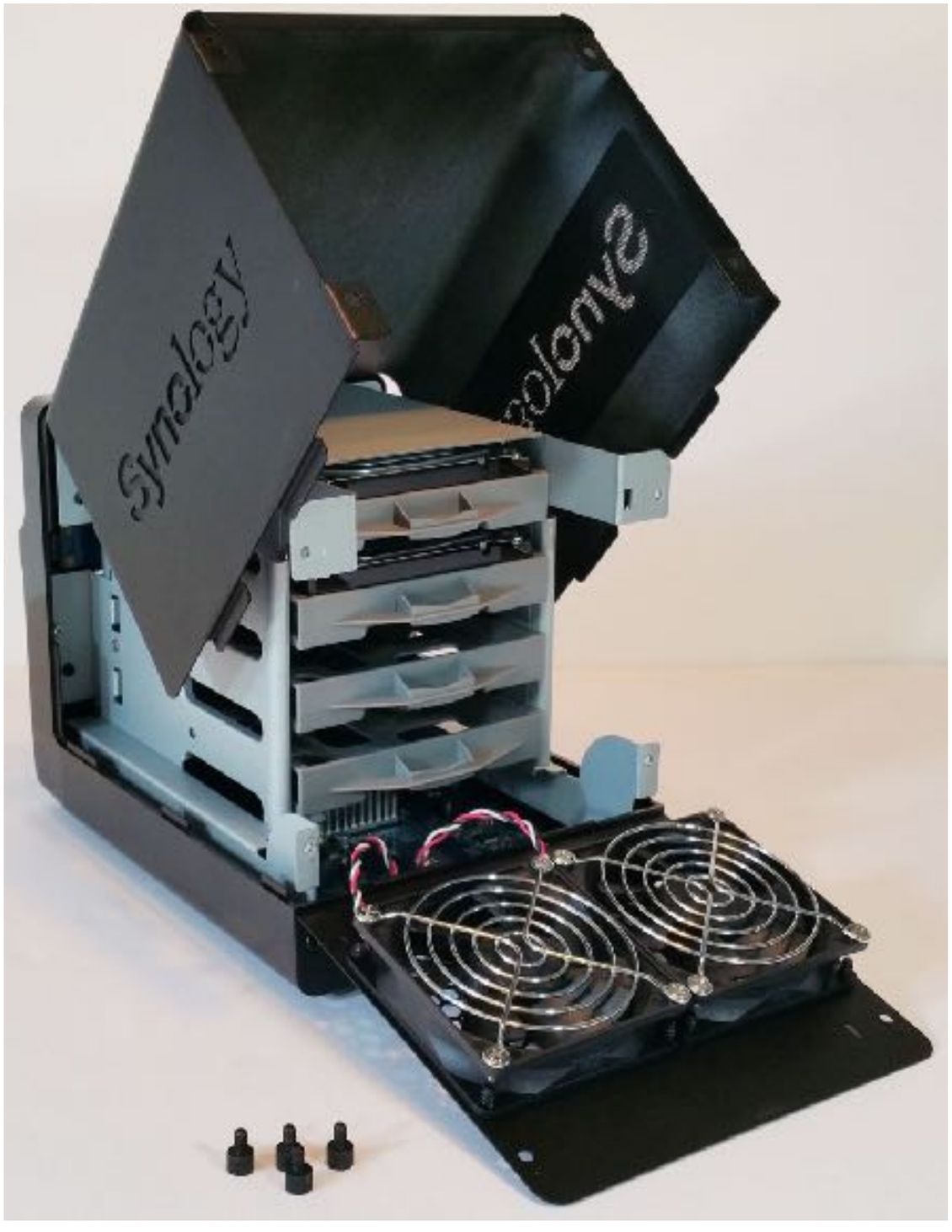
Time spent on automatable tasks

MCKINSEY GLOBAL INSTITUTE




STATUS
LAN
DISK 1
DISK 2
Synology
DS216se

Openness matters…!

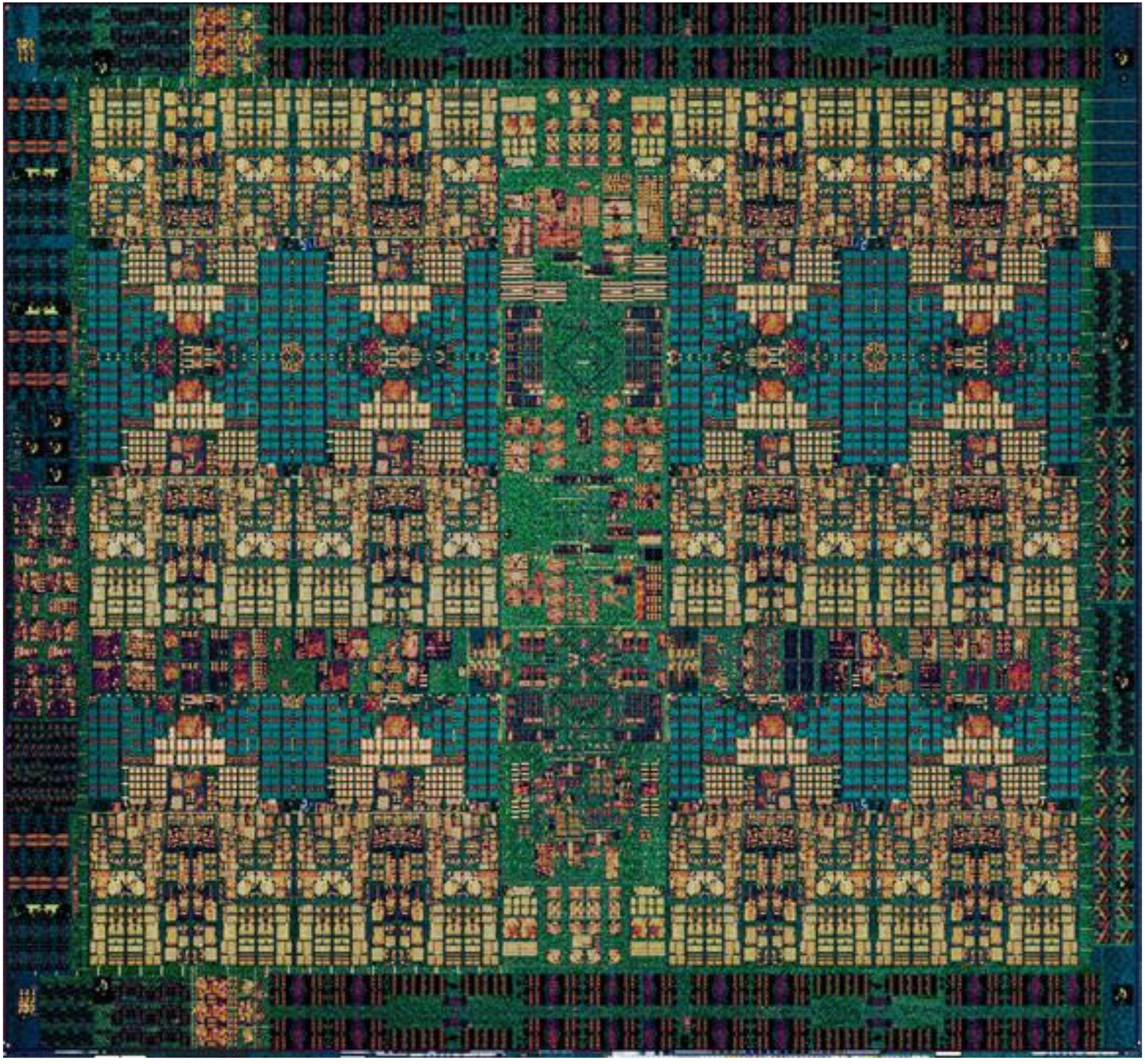


- Linux RAID
- OpenSSL Encryption
- Standard SATA Drives
- Linux OS
- Terminal access via SSH
- Powerful DSM Software (closed)
- AppStore like ecosystem

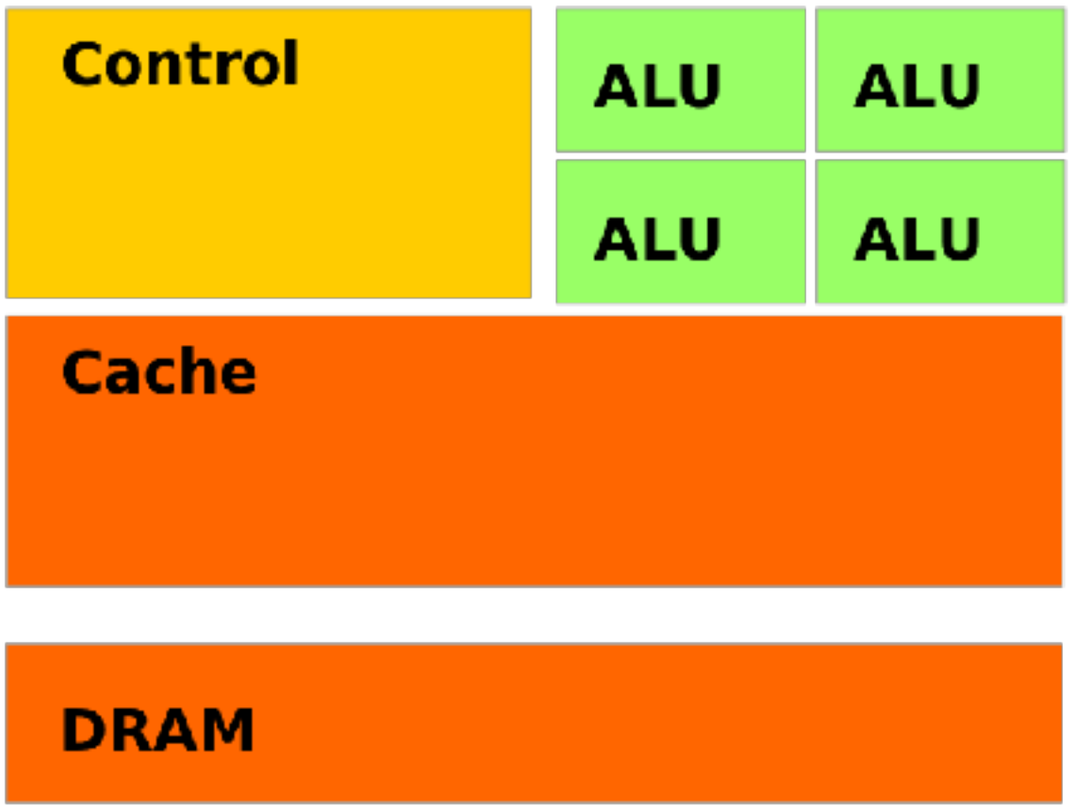
Hardware matters…?!

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

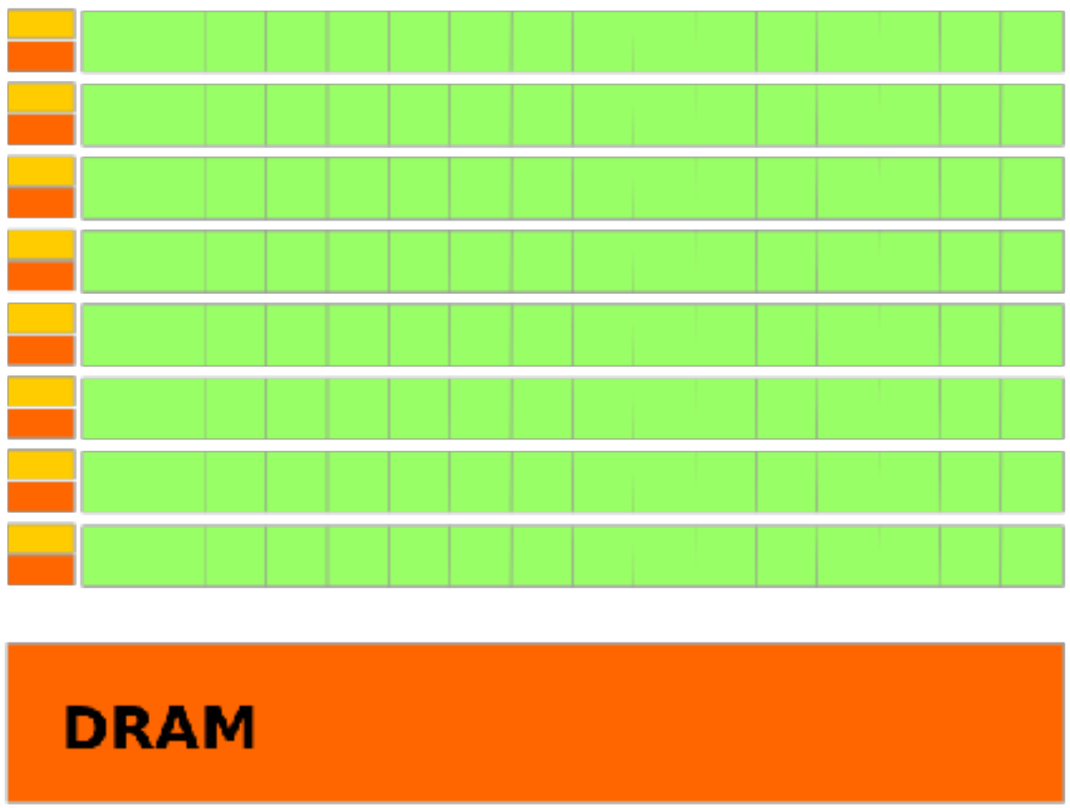
CPU



GPU

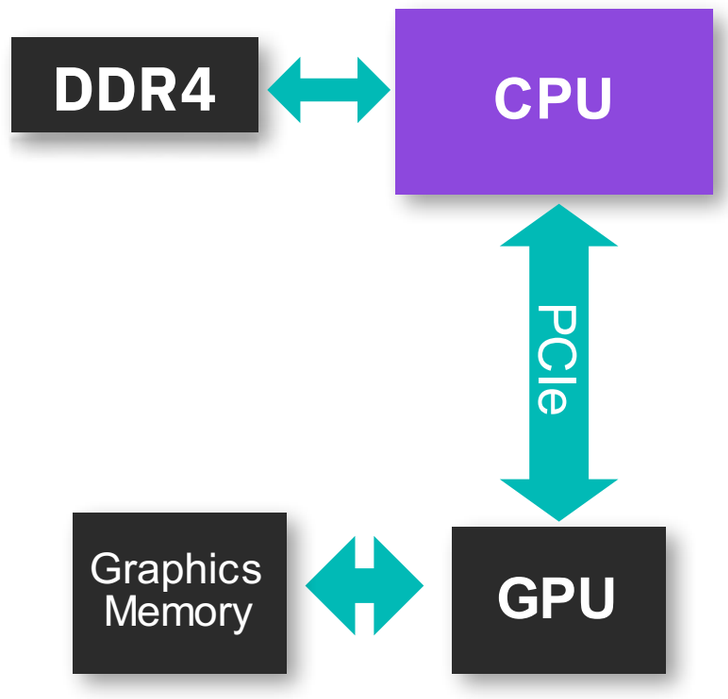


CPU



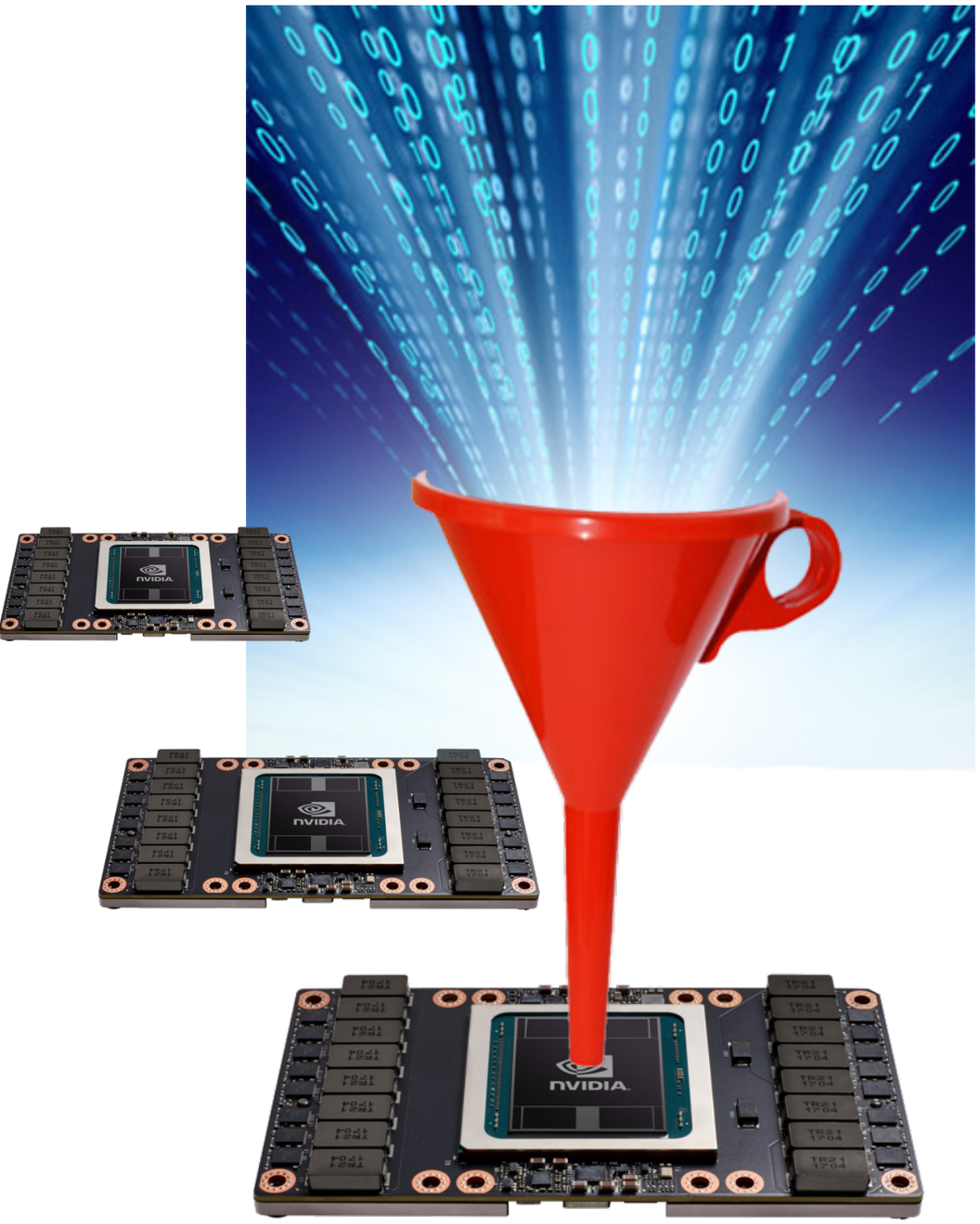
GPU

```
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
```

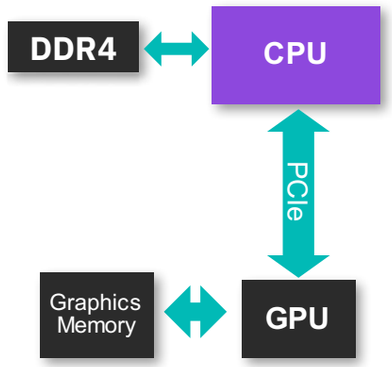


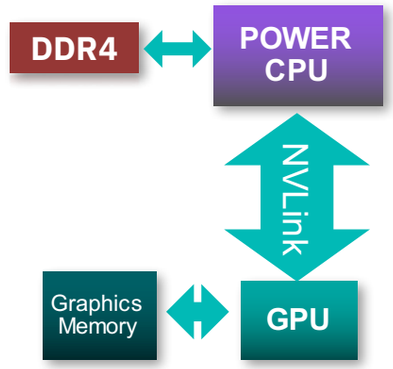
PCIe Gen3 – 32GB/sec

PCIe Gen4 – 64GB/sec



```
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

PCIe Gen4 – 64GB/sec

NVLink 1.0 – 80GB/sec

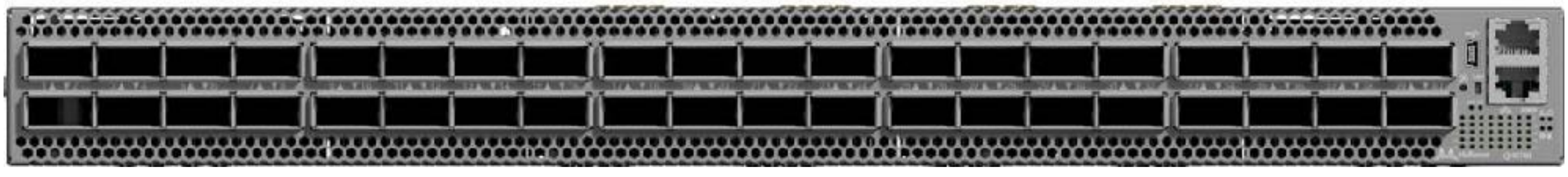
NVLink 2.0 – 300GB/sec

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

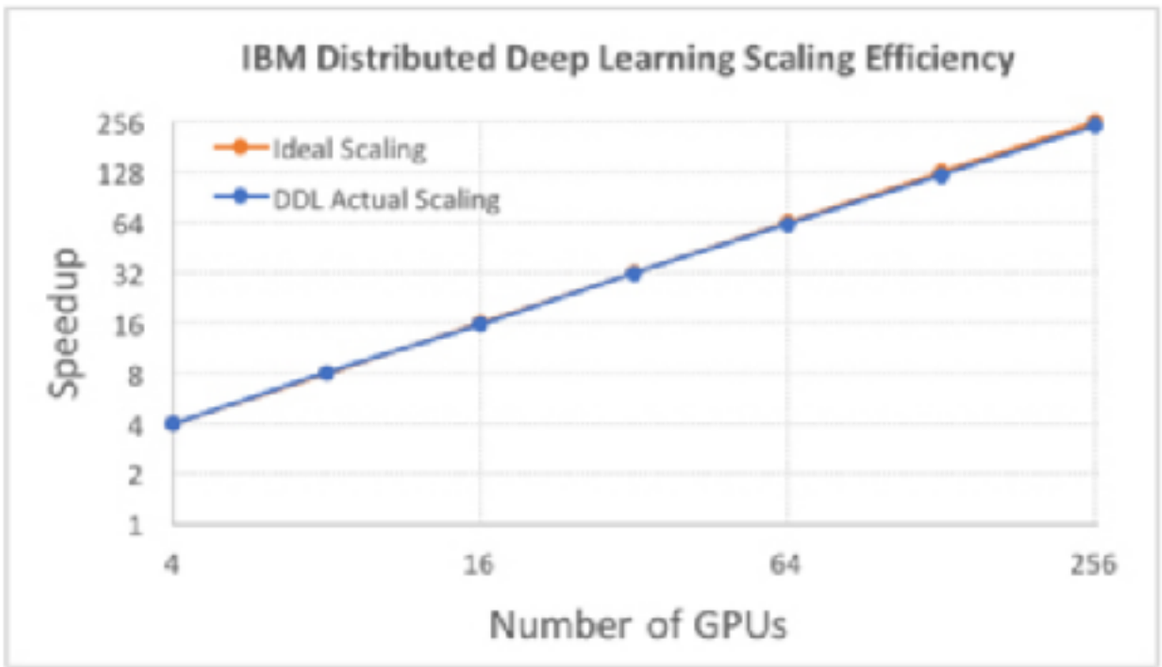
IBM Research achieves record deep learning performance with new software technology

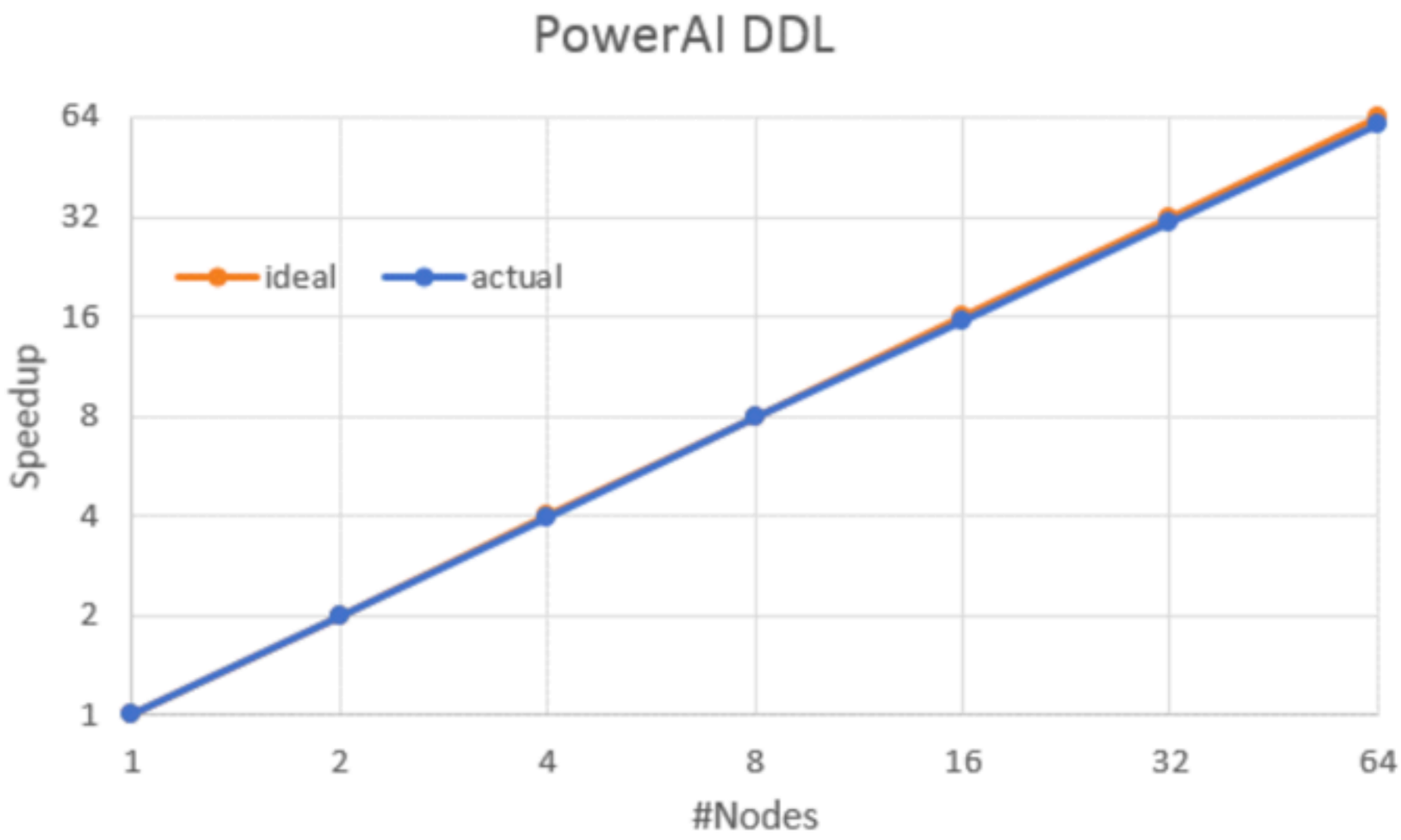
Scaling Efficiency
Facebook: 89%
IBM: 95%

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%





#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.

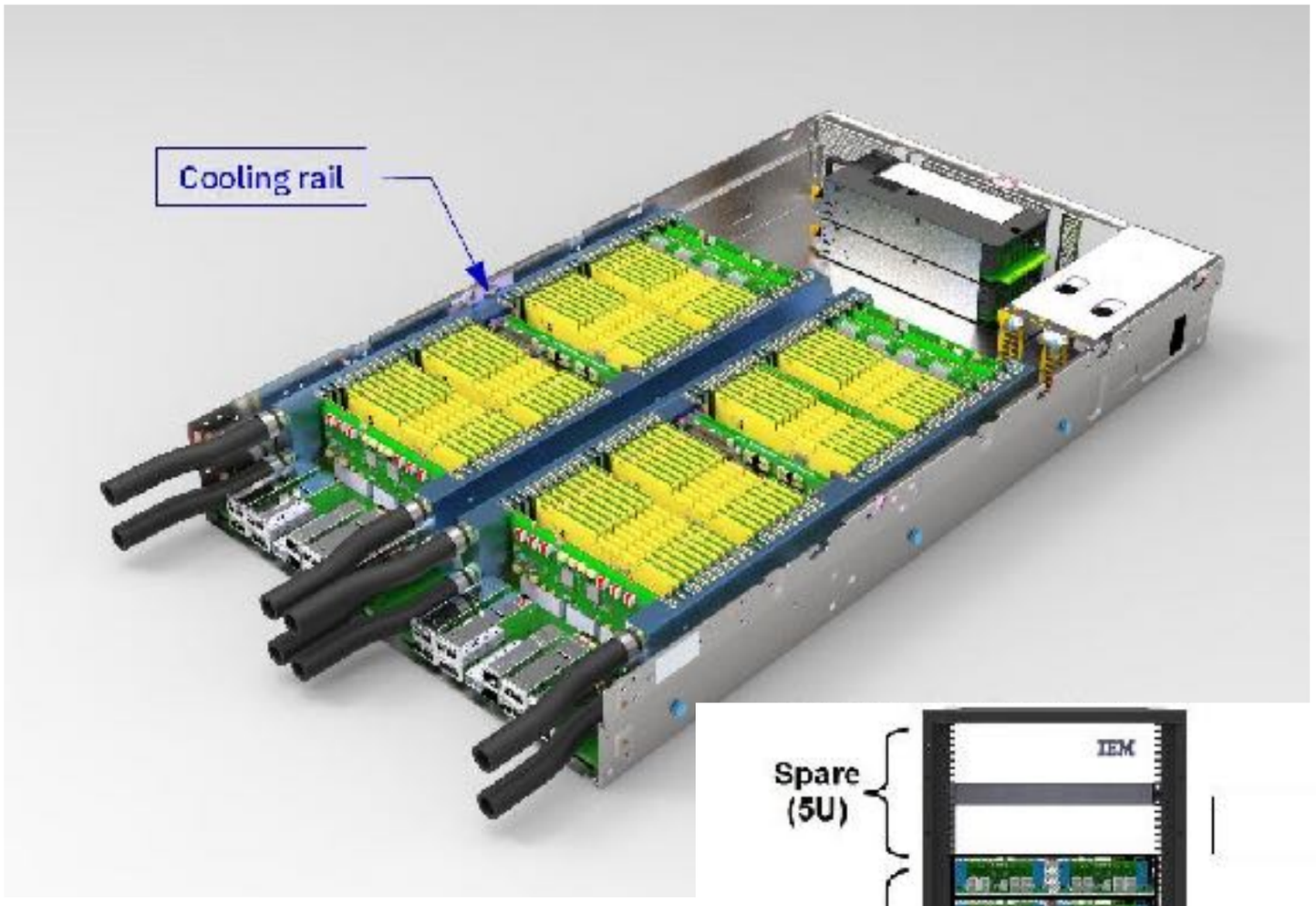
U.S. DEPARTMENT OF ENERGY
Office of Science

National Nuclear Security Administration

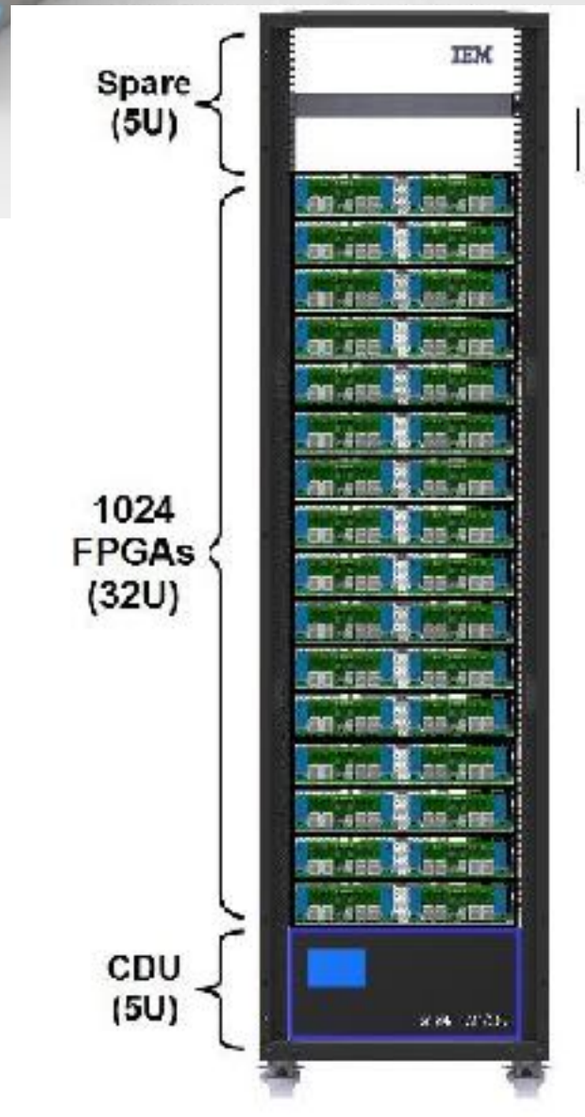
- 200 petaflops
- 4600 nodes
- 9200 Power9 CPU’s
- 220800 cores
- 1766400 hyper threads
- 27600 nVidia V100 GPUs
- 10 PB memory
- Power 15 MW
- 1% of Ethereum’s hash rate

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	

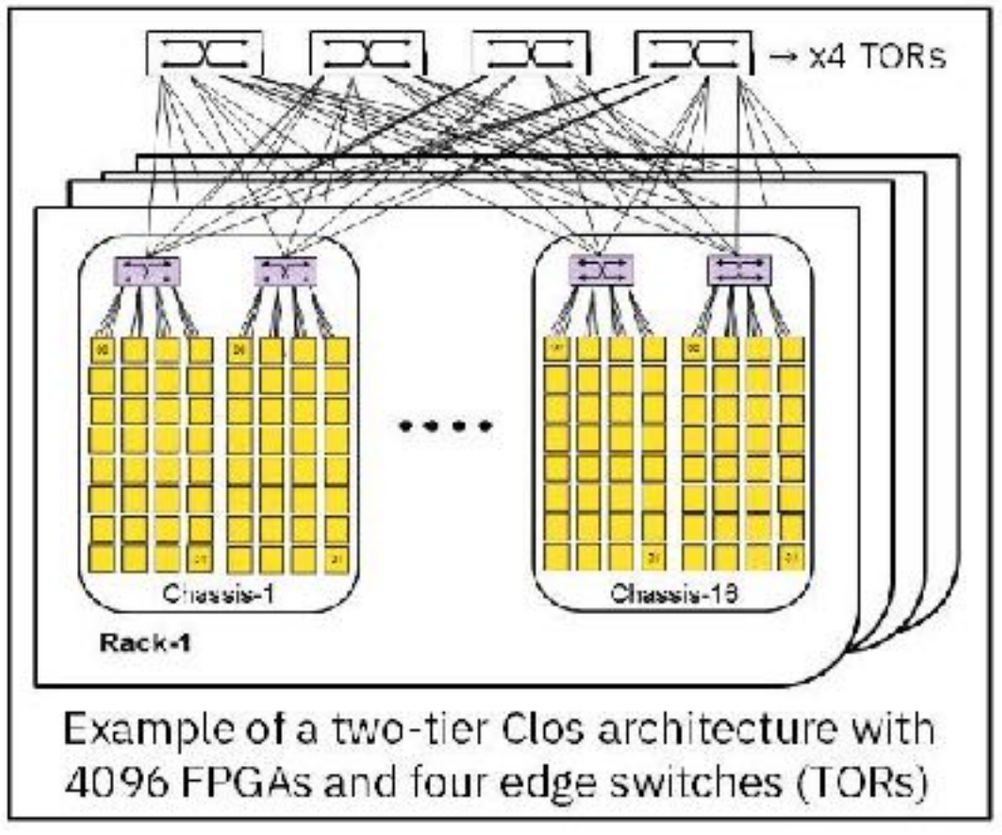
FPGAs



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



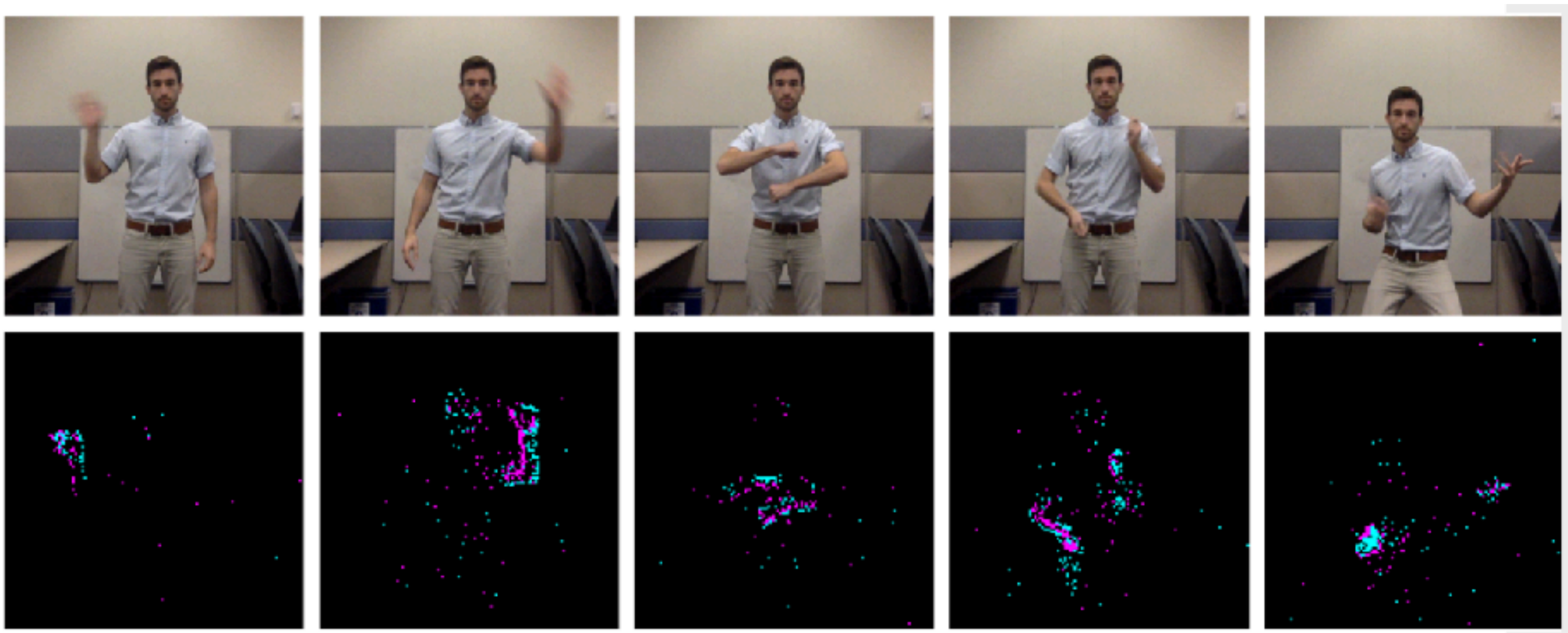
1024 FPGAs → 2.8M DSPs,
$2x10^{15}$ Fixed-Point Multiply-Accumulates/s
10 Tb/s bi-sec. Bw – 16 TB DDR4 – 40 kW max.



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

ASICs

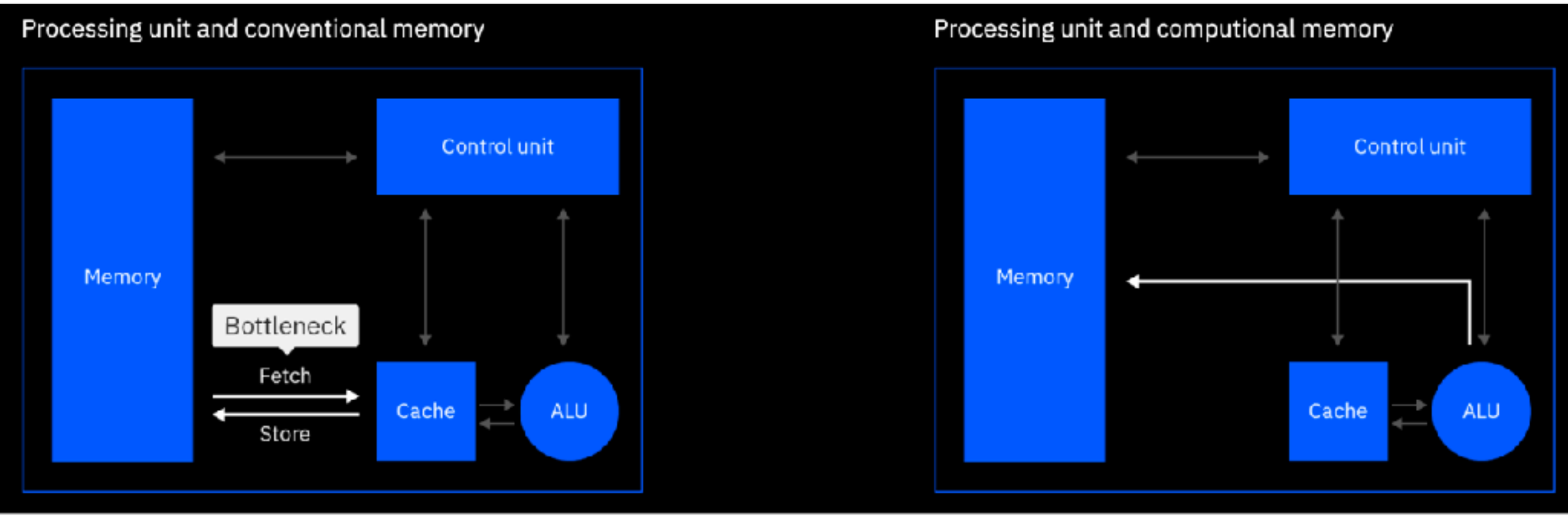
Combining the [IBM TrueNorth neurosynaptic processor] with an [iniLabs Dynamic Vision Sensor] (DVS)



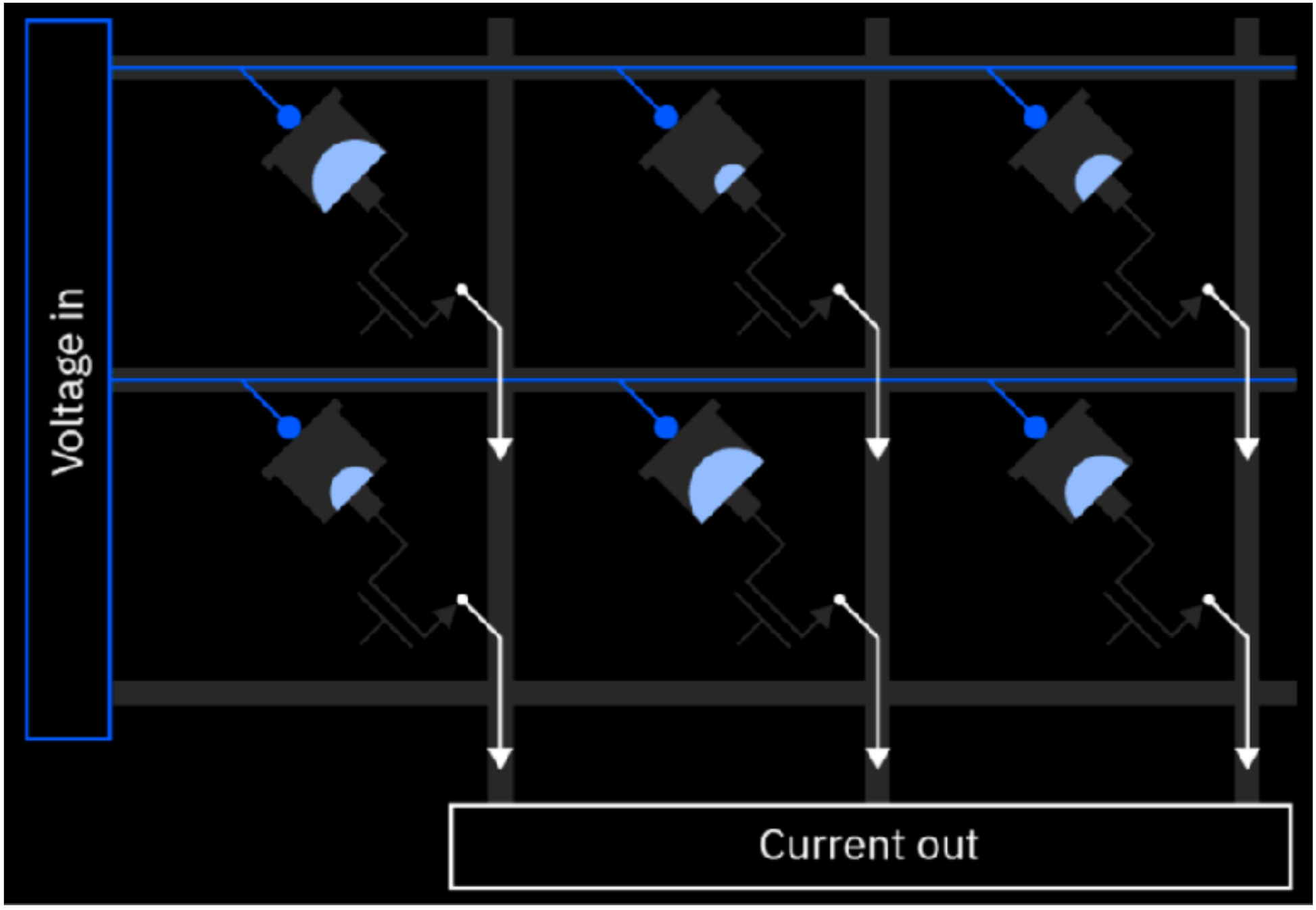
96.5 percent accuracy within a tenth of a second consuming under 200 mW

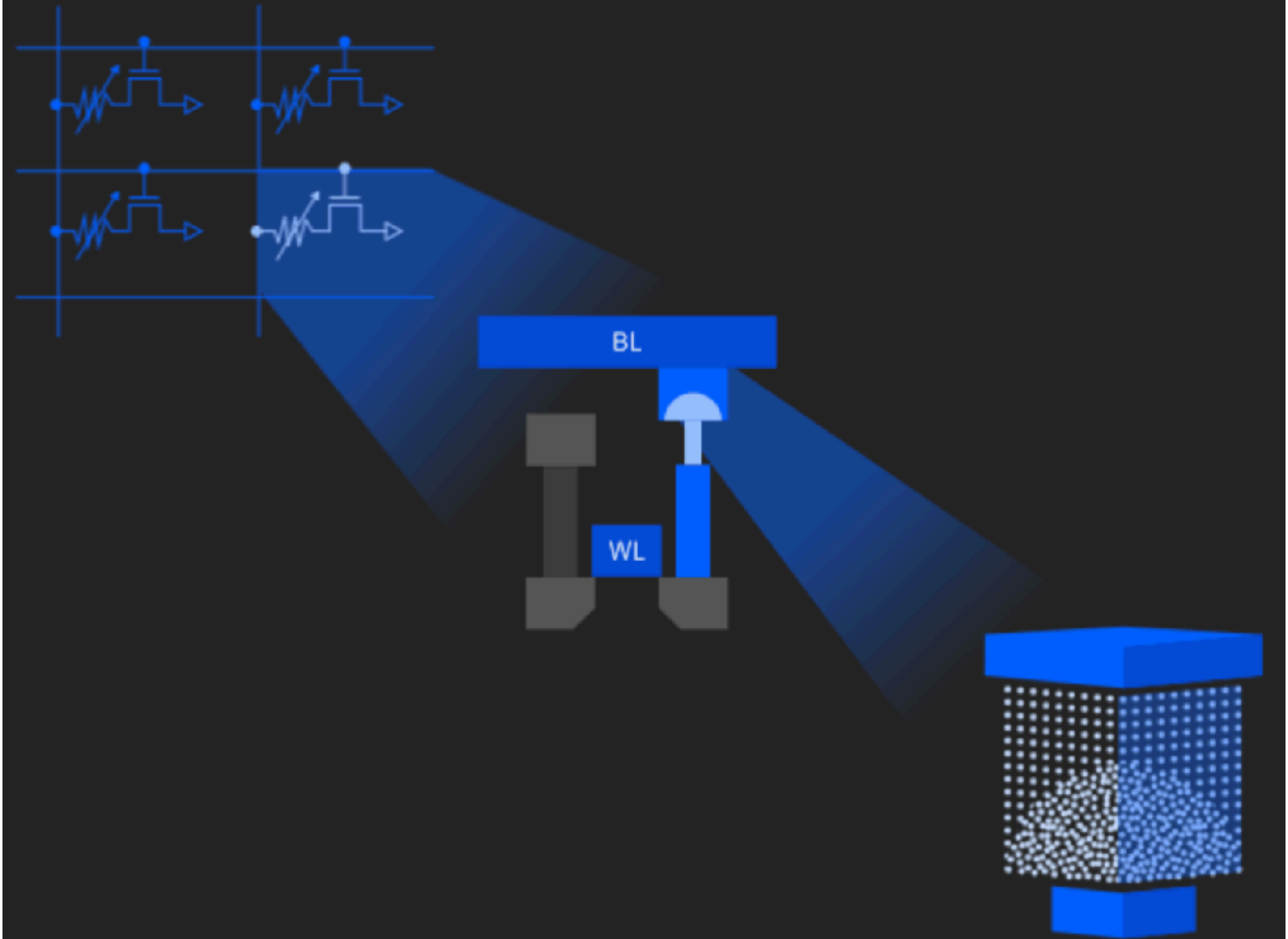
https://www.youtube.com/watch?v=g08IW-qRomM

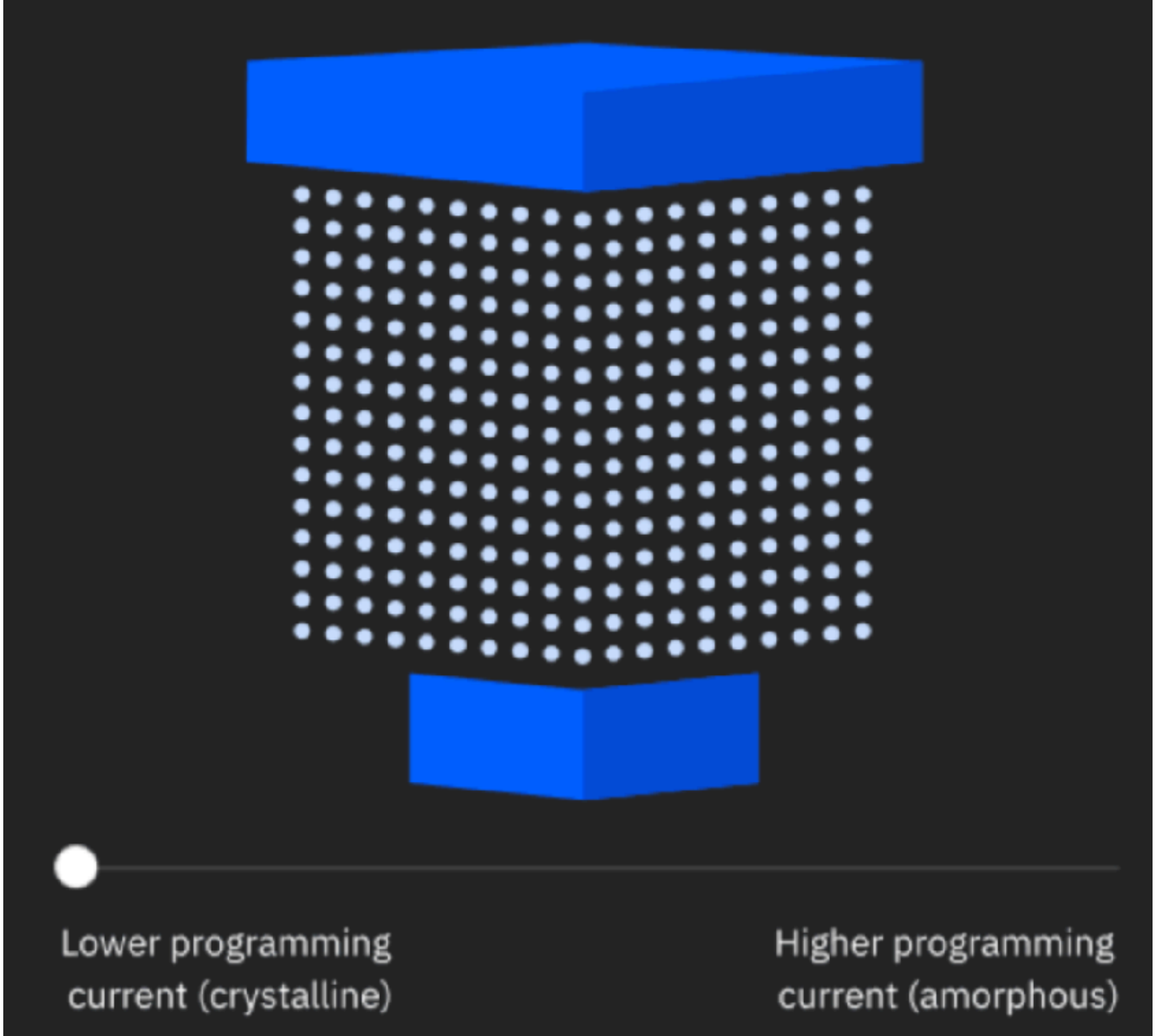
IBM Fusion Chip

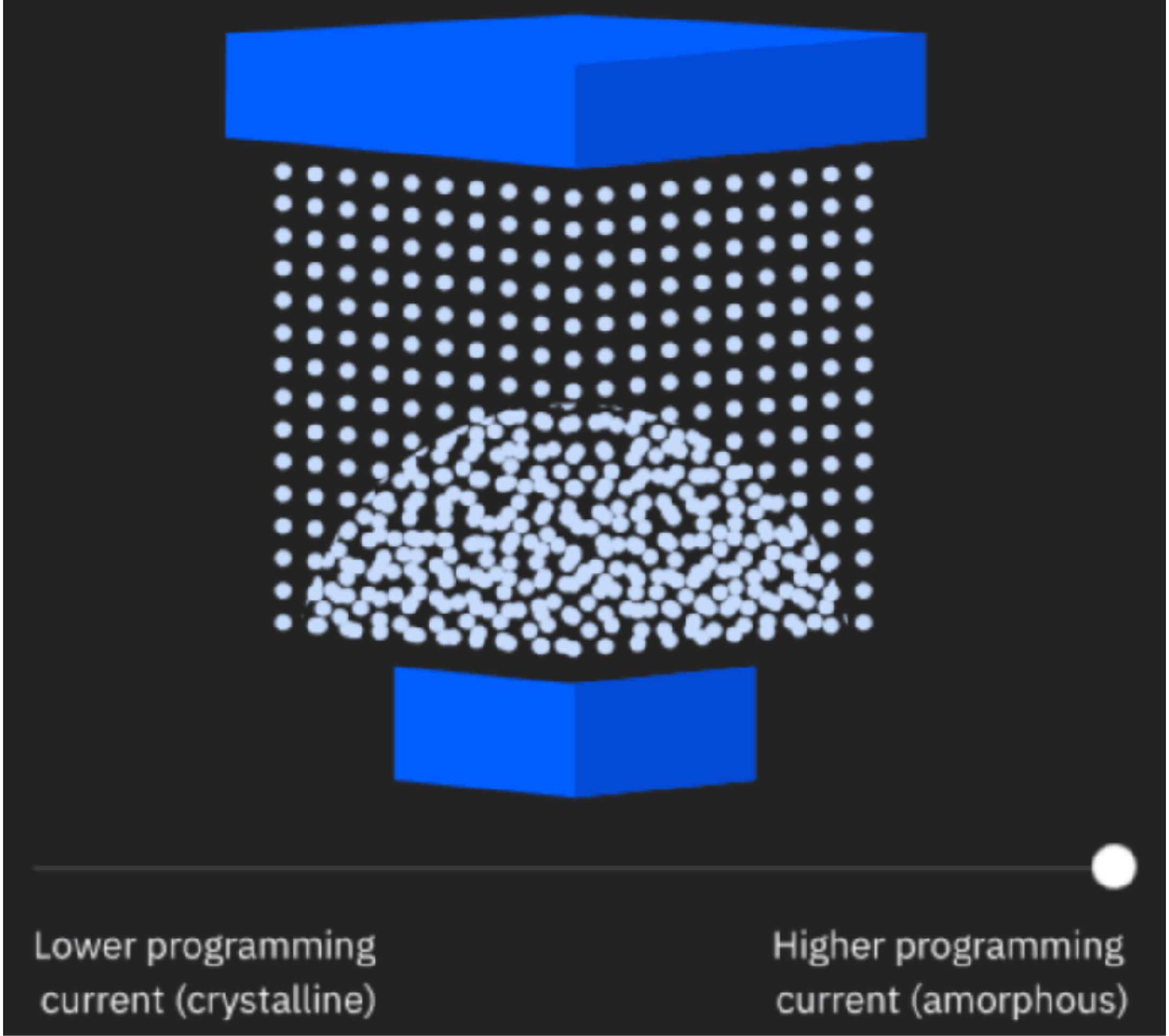


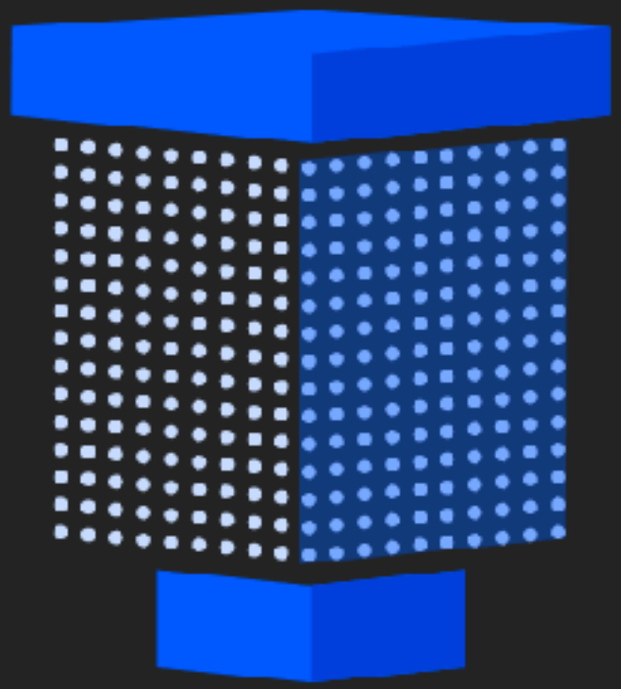
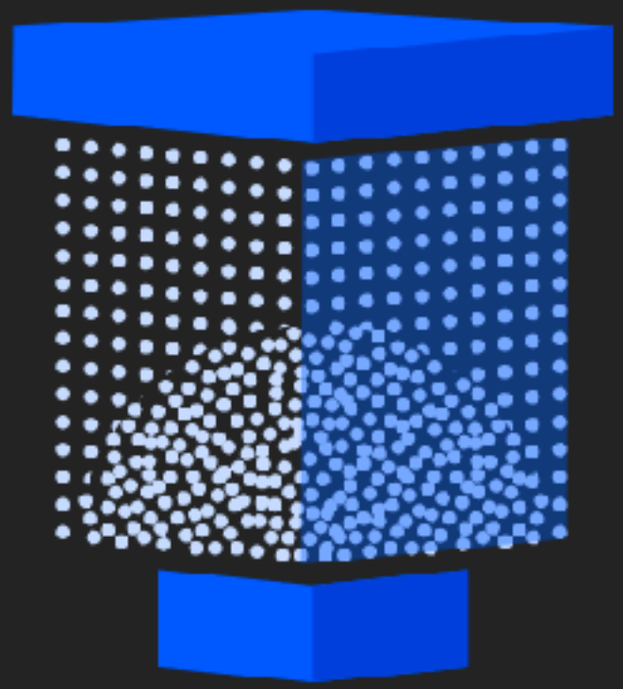
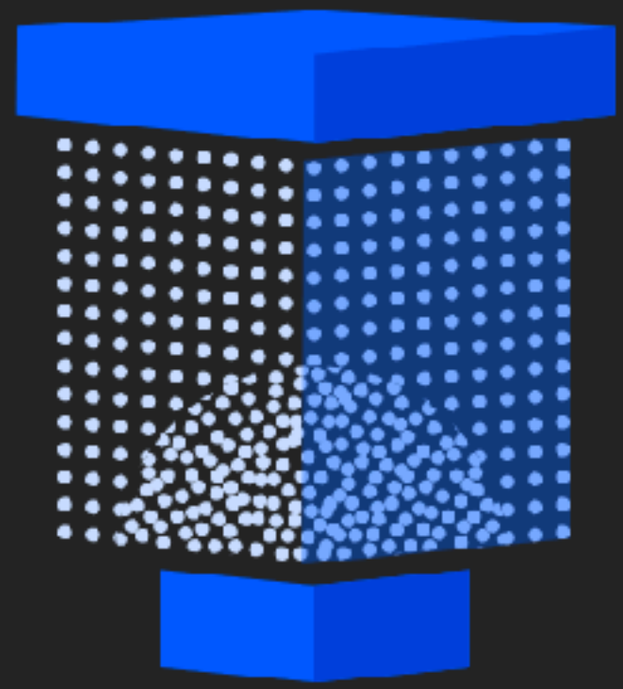
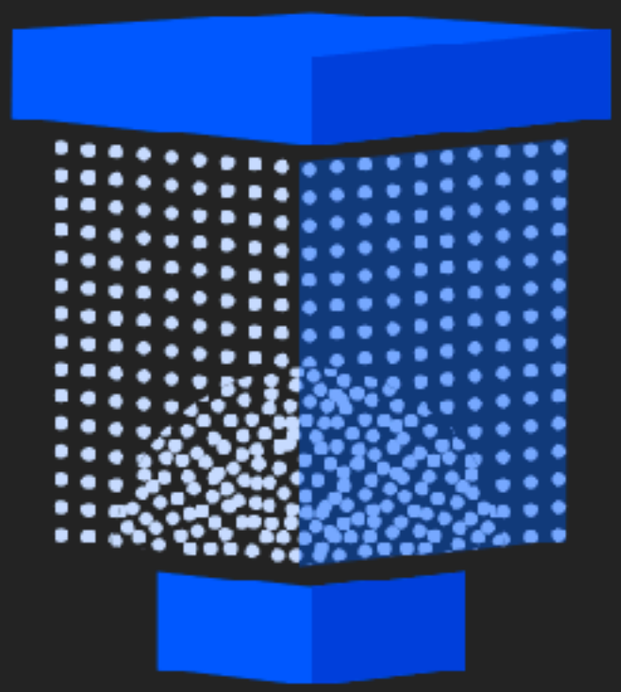
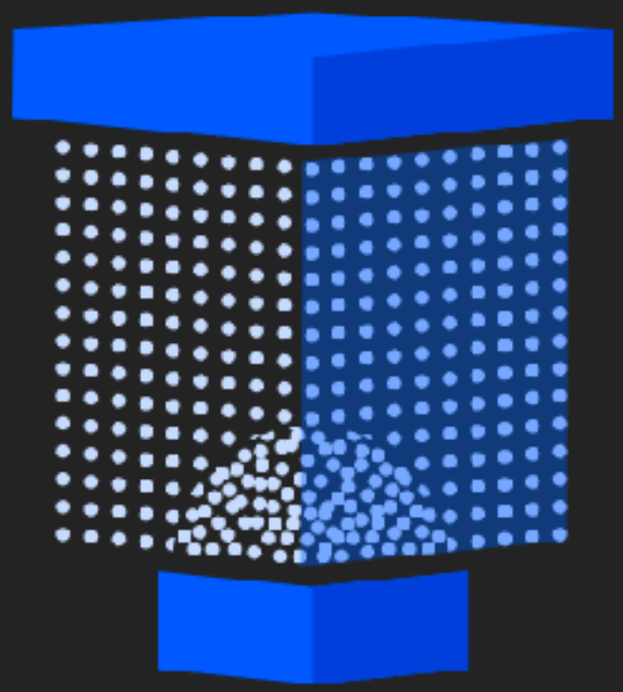
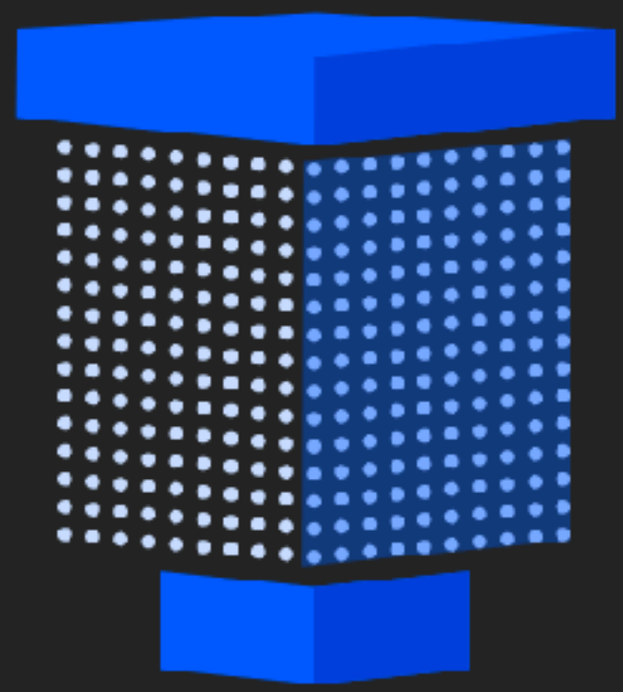
IBM Fusion Chip




BL
WL


Lower programming
current (crystalline)
Higher programming
current (amorphous)


Lower programming
current (crystalline)
Higher programming
current (amorphous)

https://en.wikipedia.org/wiki/GeSbTe

Fusion chip

First layer

784

250

Neural network

First layer

784

250

Fusion chip

Second layer

250

10

Neural network

Second layer

250

10

Fusion chip

Neural network

+

-

…resulting in **unfathomable** changes to human civilization.

…resulting in **unfavourable** changes to human civilization.

So what does it take to trust a decision made by a machine?

(Other than that it is 99% accurate)?

Did anyone tamper with it?

Is it fair?

Is it easy to understand?

Is it accountable?

Did anyone tamper with it?

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.

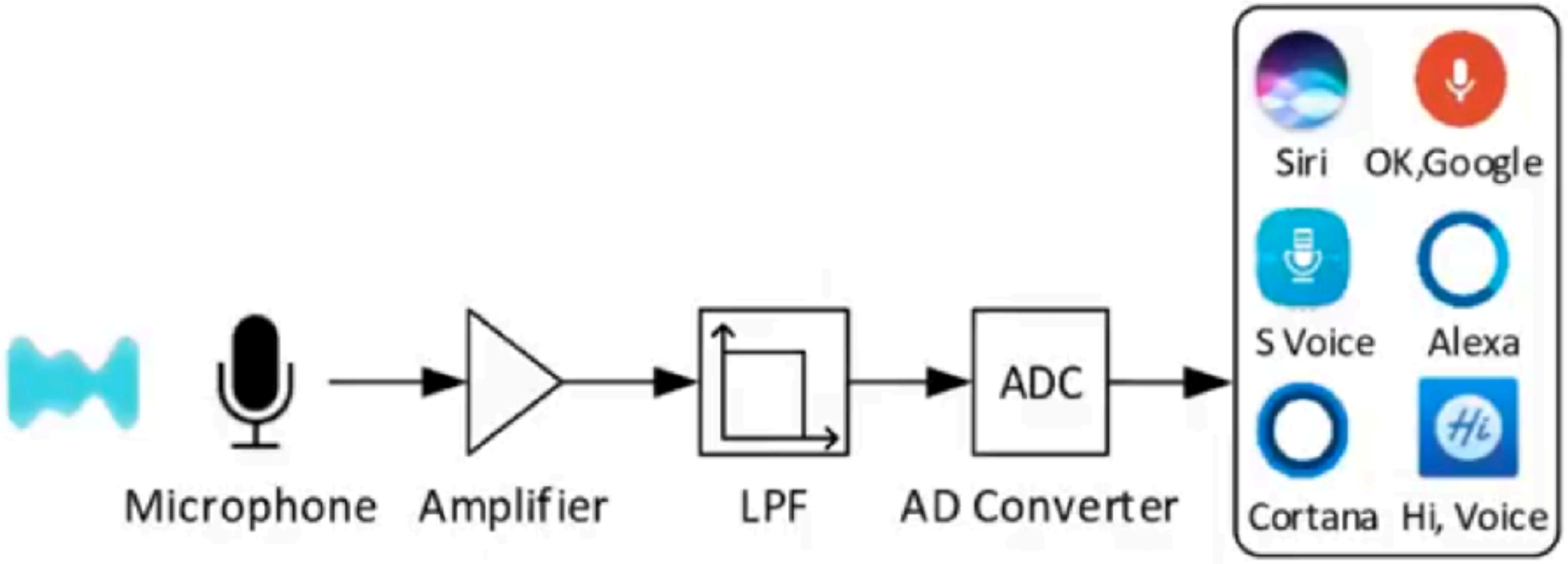
Attack noise hides pedestrians from the detection system.

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


STOP
© Washington University


STOP
HATE
© Washington University

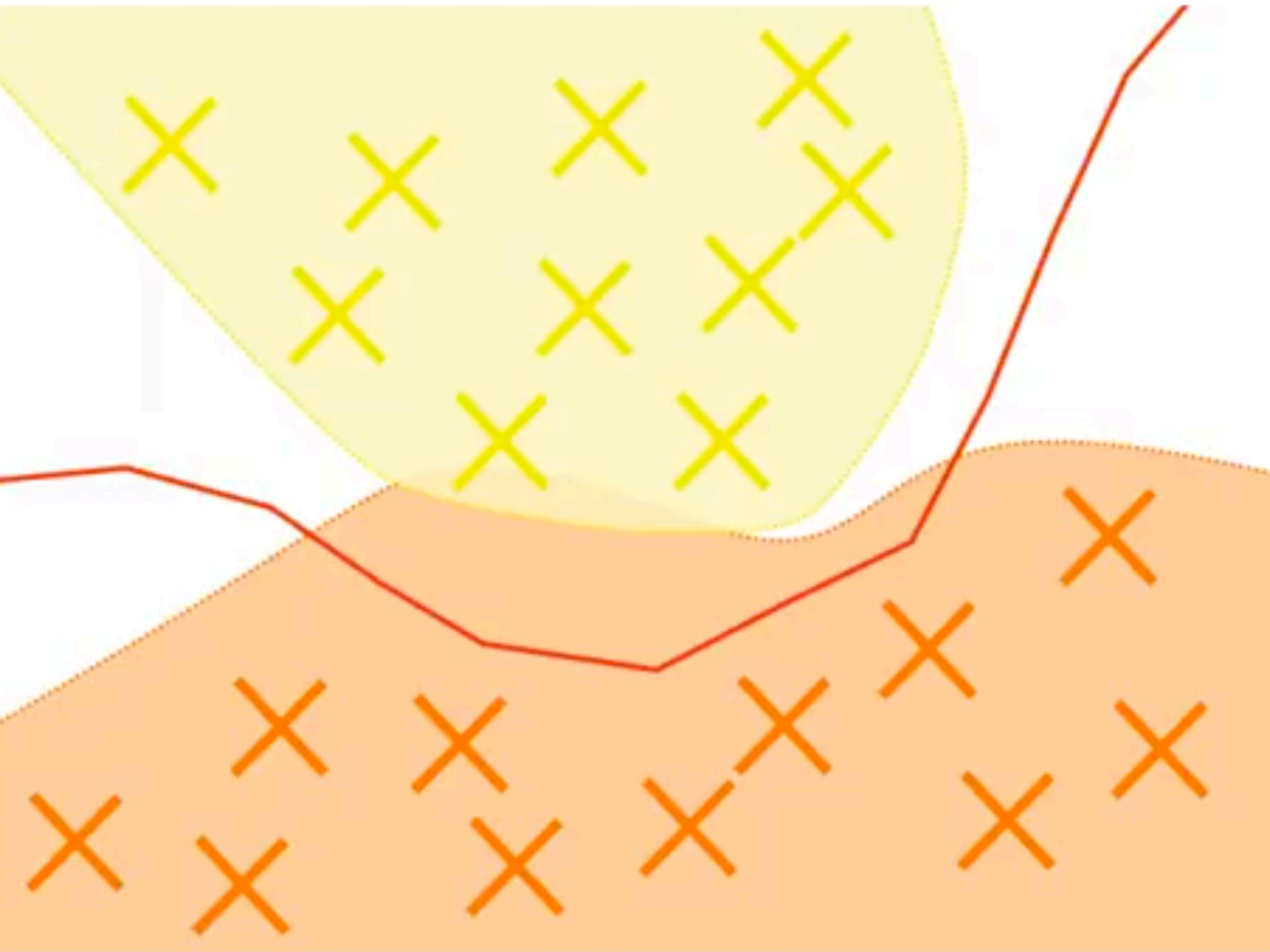

SPEED
LIMIT
45

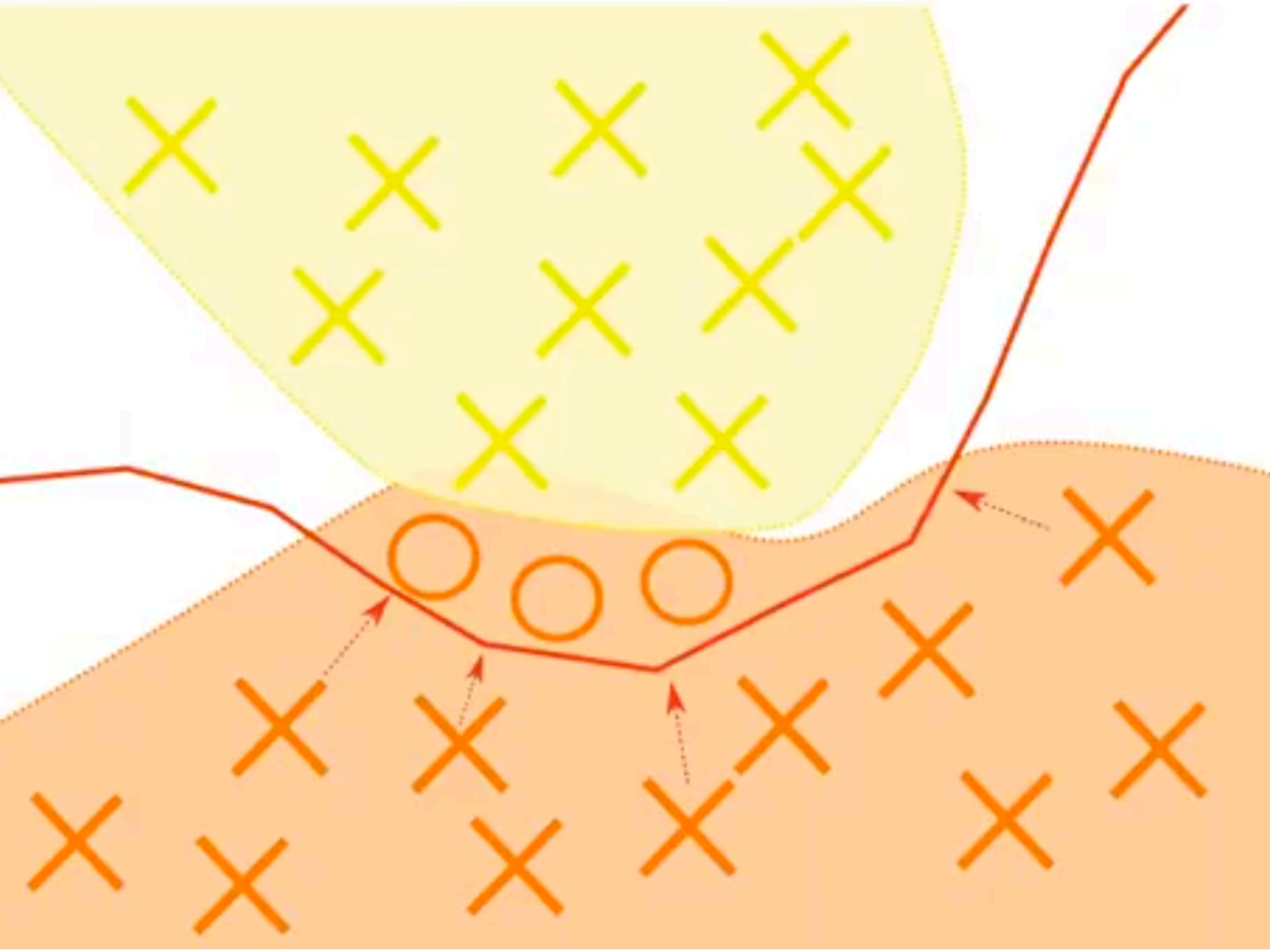


Okay Google, text John!

- 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

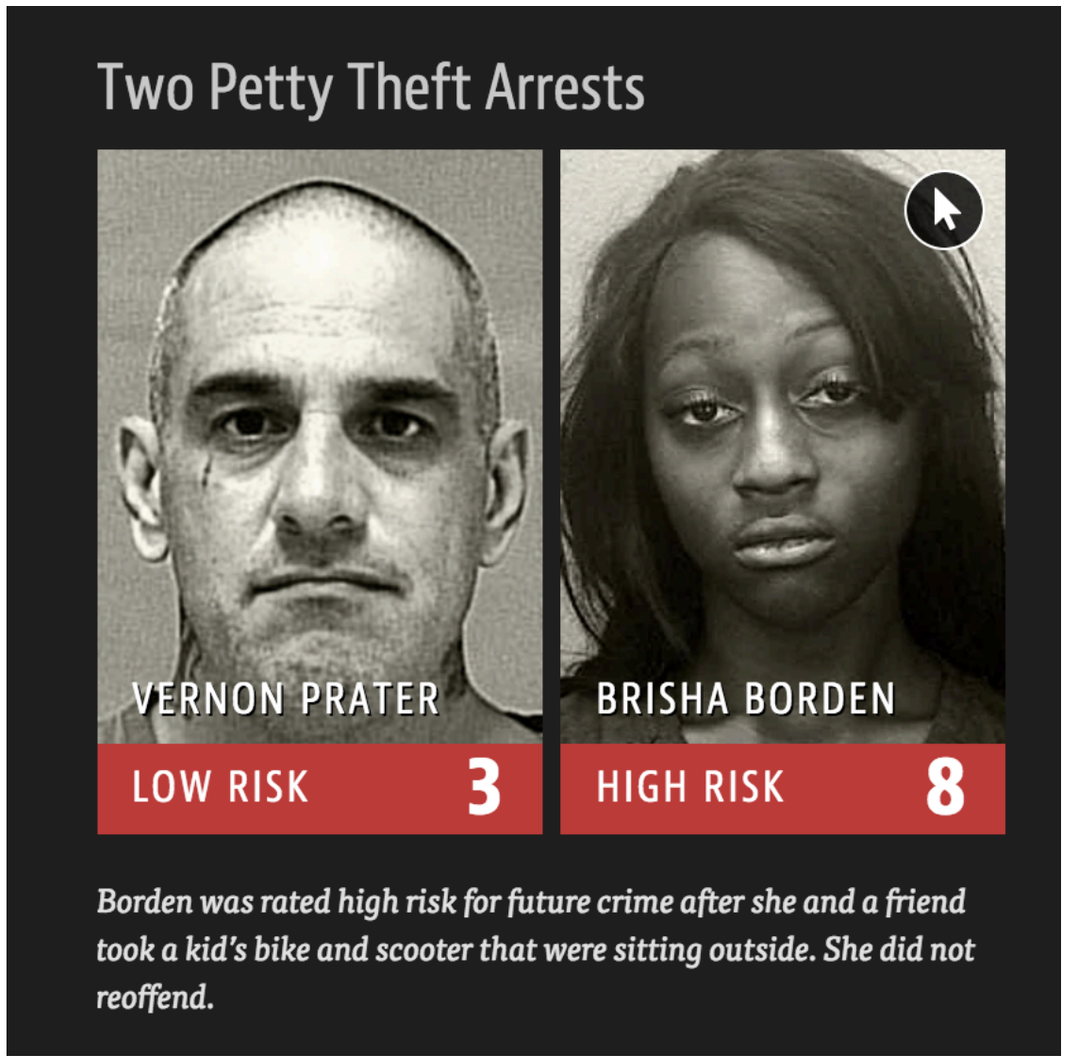
Attacks	Defenses
DeepFool	Feature Squeezing
Fast Gradient Method	Spatial Smoothing
Jacobian Saliency Map	Label Smoothing
NewtonFool	Adversarial Training
Universal Perturbation	Virtual Adversarial Training
C&W Attack	Gaussian Augmentation
Virtual Adversarial Method	

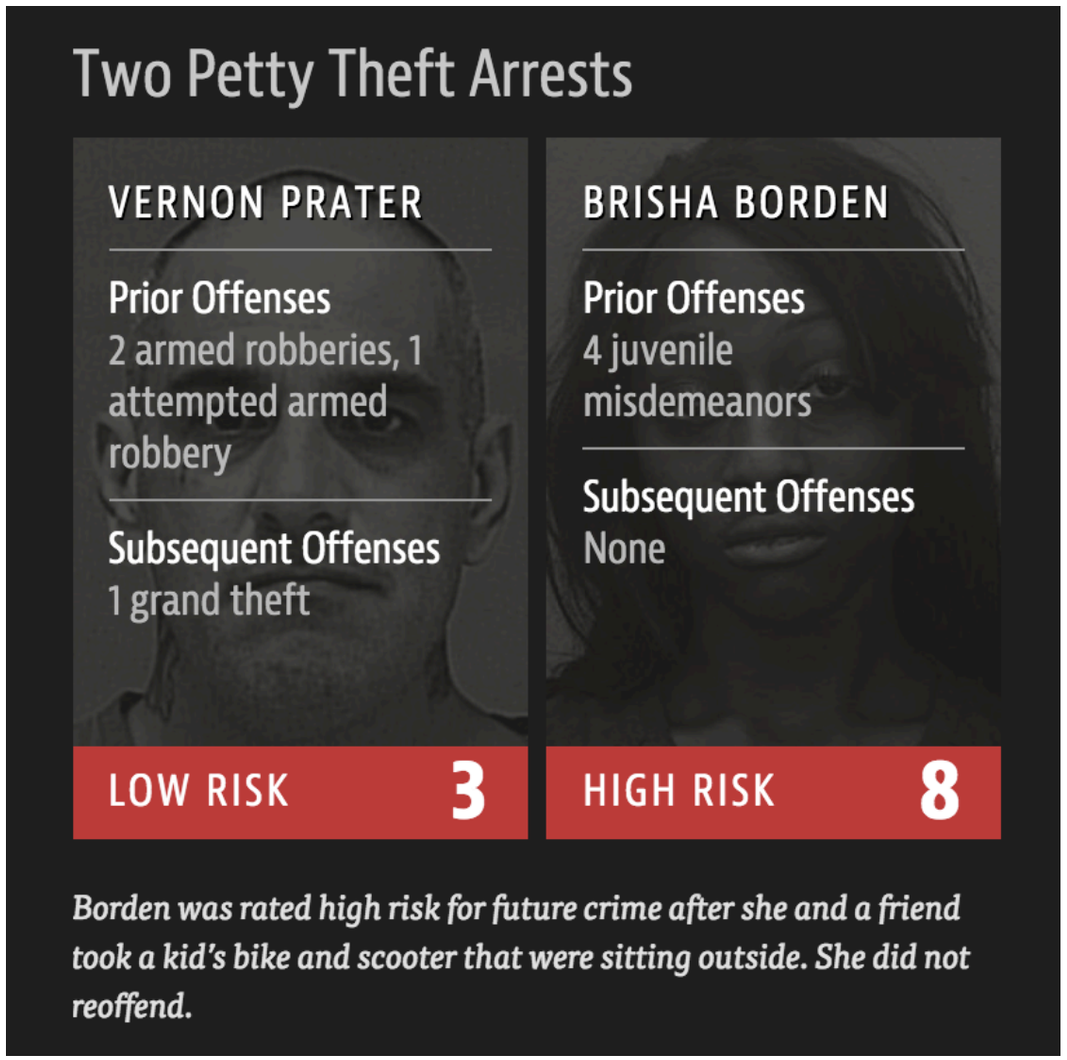
Frameworks	Metrics
TensorFlow	Loss sensitivity
Keras	Empirical robustness
PyTorch (soon)	CLEVER
MXNet (soon)	

Is it fair?

Bias

Northpointe's COMPAS algorithm widely used since 2008 in Broward County, Florida is racially biased





flagging black people 45% vs. white people 24% for risk for future 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)

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

AI Fairness 360 - Demo

Data — Check — Mitigate — Compare

Back Next

2. Check bias metrics

Dataset: German credit scoring
Mitigation: none

Protected Attribute: Sex

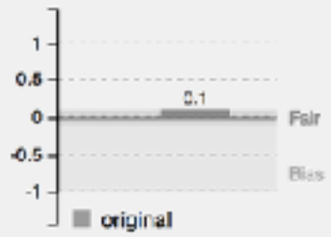
Privileged Group: ***Male***, Unprivileged Group: ***Female***

Accuracy with no mitigation applied is 76%

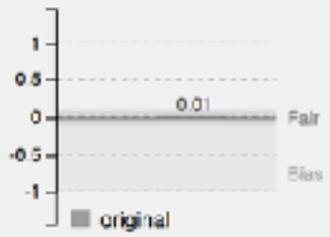
With default thresholds, bias against unprivileged group detected in 0 out of 5 metrics

Statistical Parity Difference

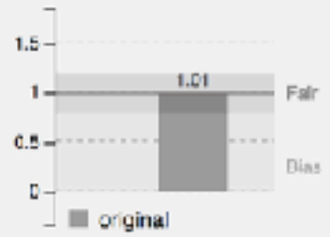
Equal Opportunity Difference



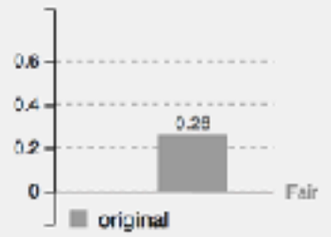
Average Odds Difference



Disparate Impact



Theil Index



○ **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.

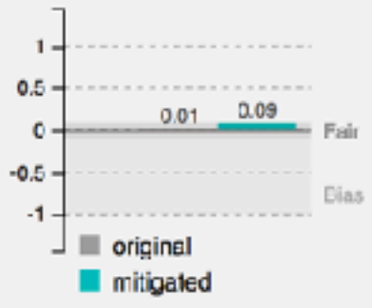
4. Compare original vs. mitigated results

Dataset: German credit scoring
Mitigation: Adversarial Debiasing algorithm applied

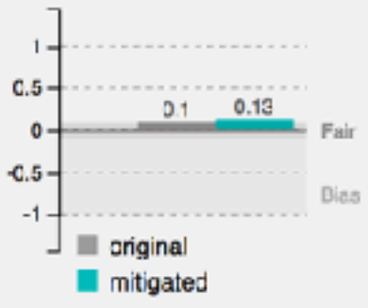
Protected Attribute: Sex

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)

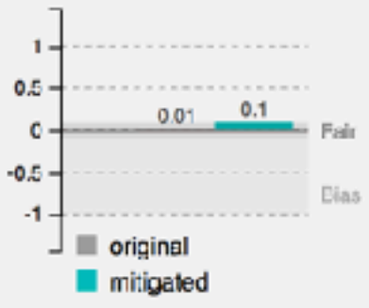
Statistical Parity Difference



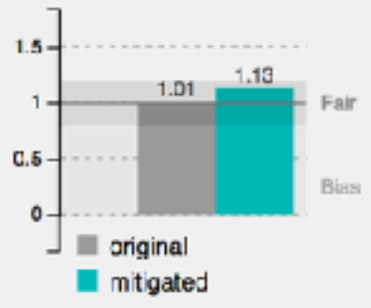
Equal Opportunity Difference



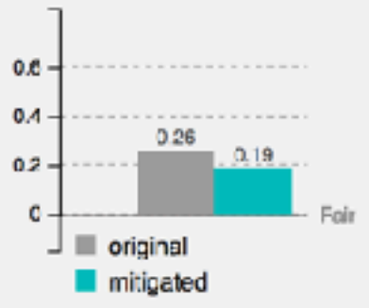
Average Odds Difference



Disparate Impact



Theil Index



IBM Watson Studio

Romeo Kienzler's Account

RK

My Projects / ··· / hello_fairness

File Edit View Insert Cell Kernel Help

Trusted | Python 3.6

Format

Run

Code

In [10]:
```
classicaltion_metric = \
    ClassificationMetric(
        dataset_ground_truth,
        dataset_classifier,
        unprivileged_groups=unprivileged_groups,
        privileged_groups=privileged_groups)



classicaltion_metric.theil_index()
```

Out[10]: 0.2772588722239781

In [ ]:

AI Fairness 360 Toolbox

https://github.com/IBM/AIF360

Is it easy to understand?

Explainability...

Hi can you please tell my why my number 004179 is blocked?

request #38815

Question about WhatsApp for Android Inbox ×

support@support.whatsapp.com Tue, Jun 25, 9:18 AM

to me

##- WhatsApp Support -##

Hi,

Thanks for your message.

We understand you're currently unable to access WhatsApp and are working diligently to answer your request. We appreciate your patience and will get back to you as soon as possible. For more information, please read this article.

support@support.whatsapp.com Jun 28, 2019, 7:06 PM

to me

##- WhatsApp Support -##

Hi,

Thanks for your message.

Your WhatsApp account has been banned because your activity violated our Terms of Service.

Be aware that we ban accounts if we believe the account activity is in violation of our Terms of Service. Please review the "Acceptable use of our services" section in our Terms of Service carefully to learn more about the appropriate uses of WhatsApp and the activities that violate our Terms of Service.

We might not issue a warning before banning your account. If you think your account was banned by mistake, please respond to this email and we'll look into your case.

Note: WhatsApp reserves the right to modify, suspend or terminate service for any reason without prior notice, at our sole discretion.

WhatsApp Support Team

Sun, Jun 30, 10:39 AM

to support

Hi

I can't find a reason why my number has been barned regarding your terms and services

Please explain

Thanks a lot!

support@support.whatsapp.com Jun 30, 2019, 10:52 AM

to me

##- WhatsApp Support -##

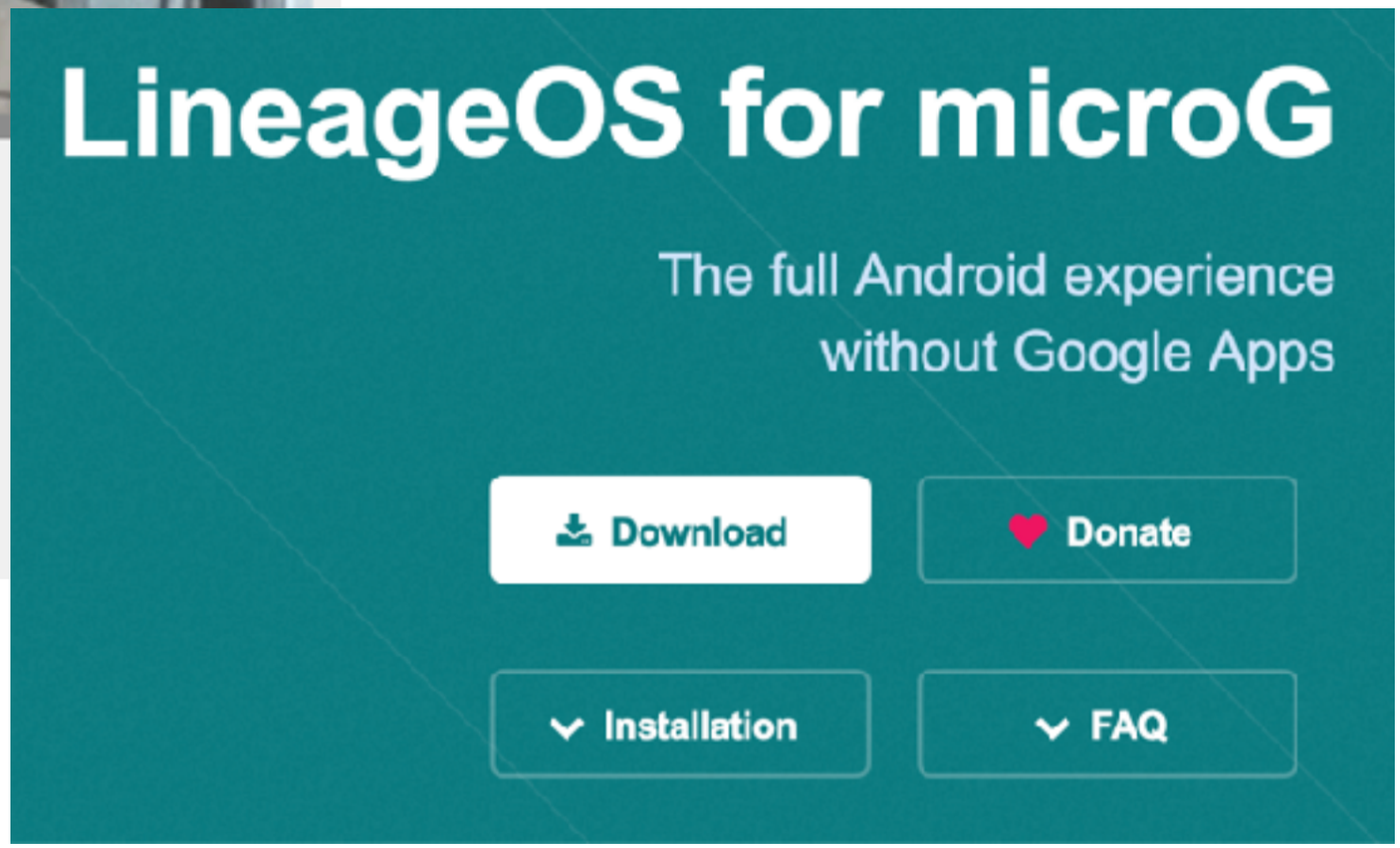
Hi,

We have reason to believe your account activity has violated our Terms of Service and decided to keep your account banned. We received a large number of complaints about your account and in order to protect our users' privacy, we won't disclose the nature of the complaints.

Responses to this email thread won't be read.

FairPhone2 Adventures: Replacing the internal...

48 views • 6 months ago



Signal

SUPPORT BLOG DEVELOPERS CAREERS DONATE EN

Fast, simple, secure.

Privacy that fits in your pocket.

Android

iPhone

Desktop

 romeo kienzler

Why Riot? Features Free! Help Open Source Get Started

RIOT.IM

Liberate your communication

Communicate the way you want with Riot - a universal secure chat app entirely under your control.

Get started

LAUNCH NOW ON YOUR BROWSER

DOWNLOAD APP FOR DESKTOP

Download on the App Store

GET IT ON Google Play

GET IT ON F-Droid

Riot is for everyone, from casual chat to high powered collaboration

IBM CODAIT

 romeo kienzler

Keybase

Login Install

The Big Stellar Space Drop

2 Billion Lumens for Everyone

Just about $122 million USD

98,595 registered so far. The next round is **Nov 15, 2019.**

You'll get **1,014** Lumens (XLM) each month if no one else registers.

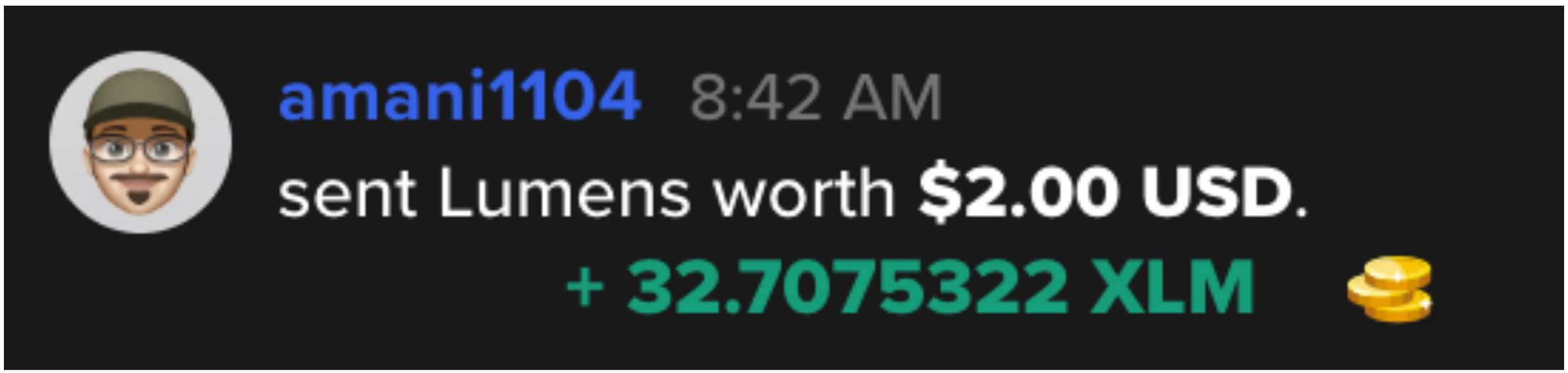
September 9, 2019

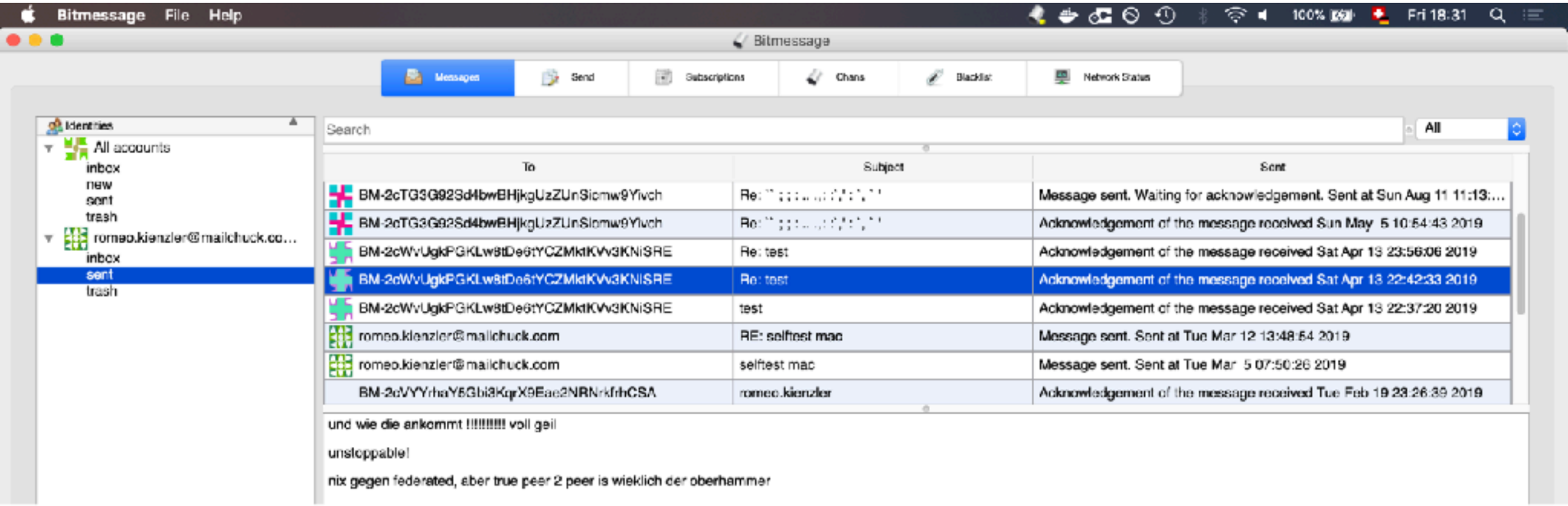
UPDATE #3 October 8. 💖 Good news, everyone! The Space Drop is back on, and the requirements to register are now looser.

At Keybase, we've **blogged before** why the Stellar network is our favorite cryptocurrency technology.

- Transactions take only 5 seconds and cost under 1¢
- Stellar does not burn mountains of fossil fuels
- Stellar has a "path payments" system for magical currency and token conversions.

The last point is special. I've got US dollars, and you want Japanese yen? No problem - 5 seconds later I've sent you yen, around the world.


amani1104 8:42 AM
sent Lumens worth $2.00 USD.
+ 32.7075322 XLM

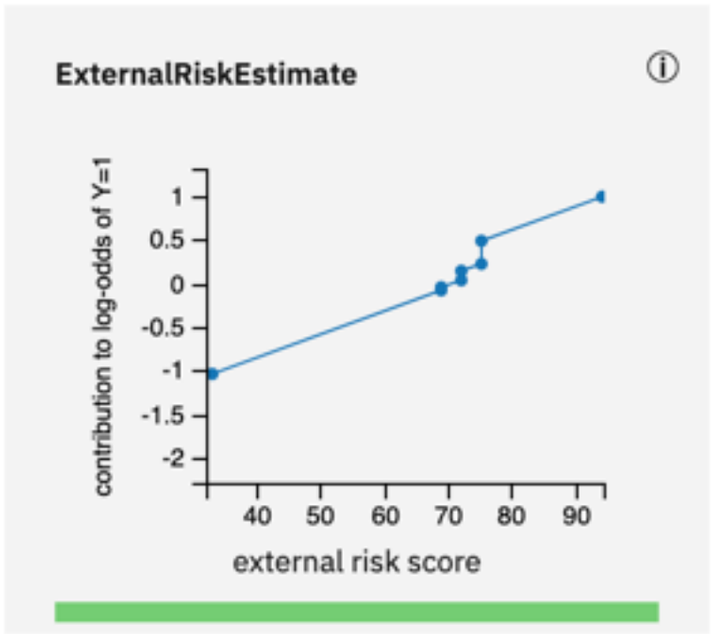


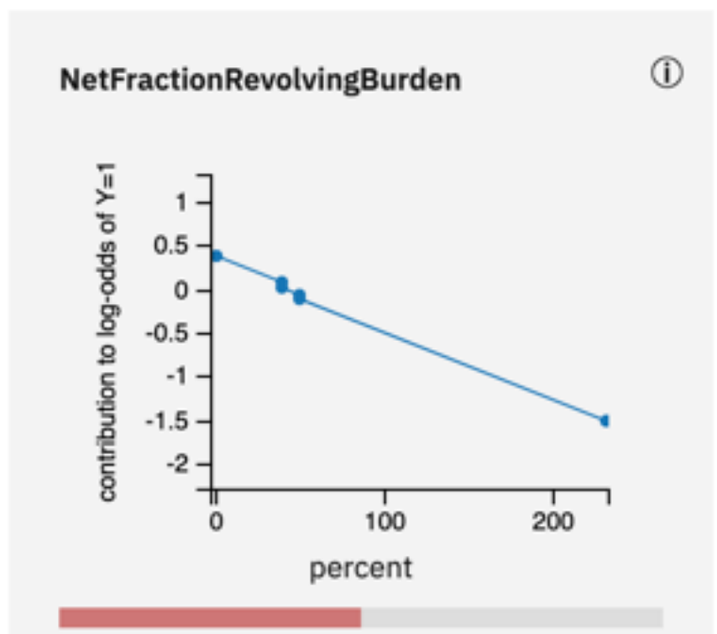
AI Explainability 360 - Demo

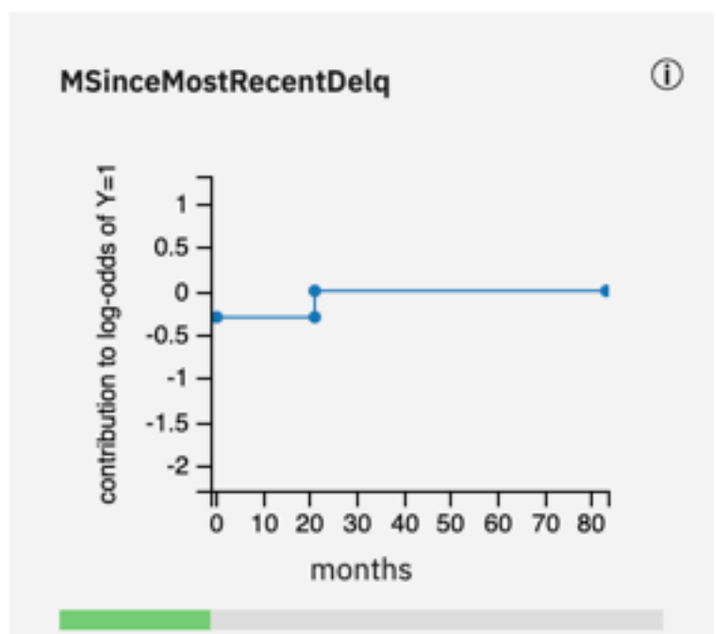
Data — Consumer — Explanation

A Data Scientist wants to understand:

What is the overall logic of the model in making decisions?
Is the logic reasonable, so that we can deploy the model with confidence?







 romeo kienzler

AI Explainability 360 - Demo

Data — Consumer — Explanation

A Loan Officer wants to understand:

Why is the model recommending this person's credit be approved or denied?
How can I inform my decision to accept or reject a line of credit by looking at similar individuals?

Using Similar Examples to Inform a Loan Decision

A Loan Officer typically makes the final decision when accepting or rejecting a customer's loan request. When using a predictive model, a Loan Officer wants to understand how and why the model came to that prediction in order to make an informed and trusted decision. One algorithm within AI Explainability 360—ProtoDash—works with an existing predictive model to show how the customer compares to others who have similar profiles and had similar repayment records to the model's prediction for the current customer, which helps to evaluate and predict the applicant's risk. Based on the model's prediction and the explanation for how it came to that recommendation, the Loan Officer can make a more informed decision.

Select a customer the Loan Officer wants to understand

Alice	Robert
Approved	Denied

 romeo kienzler

AI Explainability 360 - Demo

Data — Consumer — Explanation

A Bank Customer wants to understand:

Why was my application rejected?
What can I improve to increase the likelihood my application is accepted?

Providing Contrastive Explanations for Insight into Loan Application Outcomes

The Bank Customer wants to know how and why the decision was made to accept or reject their loan application. The explanation given will help them understand if they've been treated fairly, and also provide insight into what – if their application was rejected – they can improve in order to increase the likelihood it will be accepted in the future. To help provide that insight and suggest avenues for improvement, we will use the Contrastive Explanations Method (CEM) algorithm available in AI Explainability 360. This algorithm sits on top of an existing predictive model and helps detect both the features that a bank customer could improve (e.g., amount of time since last credit inquiry, average age of accounts), and also further detects the features that will increase the likelihood of approval and those that are within reach for the customer. See examples below.

Select a customer asking for explanations

Jason	Ann	Julia
Denied	**Denied**	**Denied**

 romeo kienzler

AI Explainability 360

https://github.com/IBM/AIX360

 romeo kienzler

Is it accountable?

Data Lineage...

Kubeflow

Pipelines

Experiments

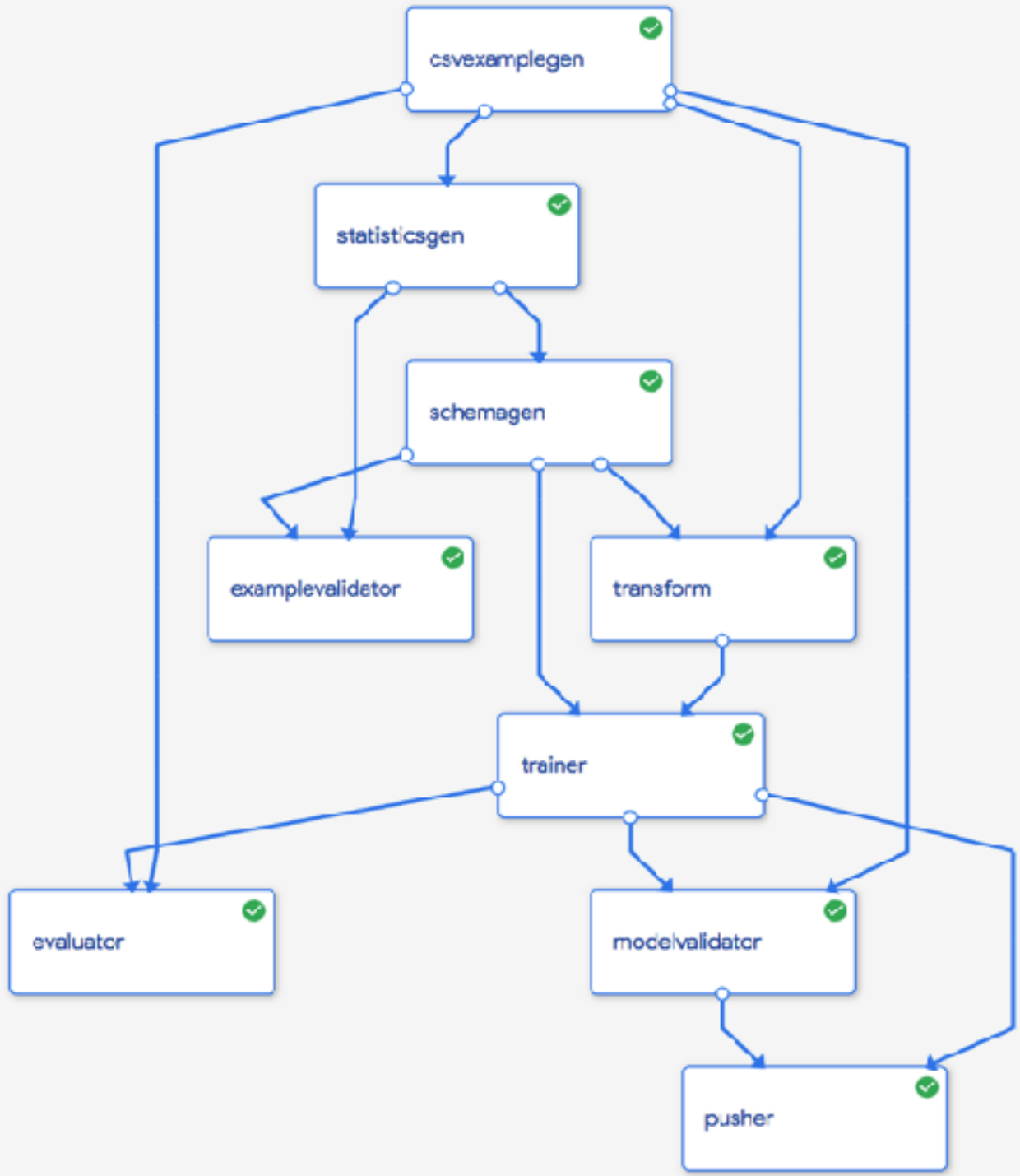
Notebooks

All runs

Chicago Taxi Pipeline

Clone run Refresh

Graph Run output Config

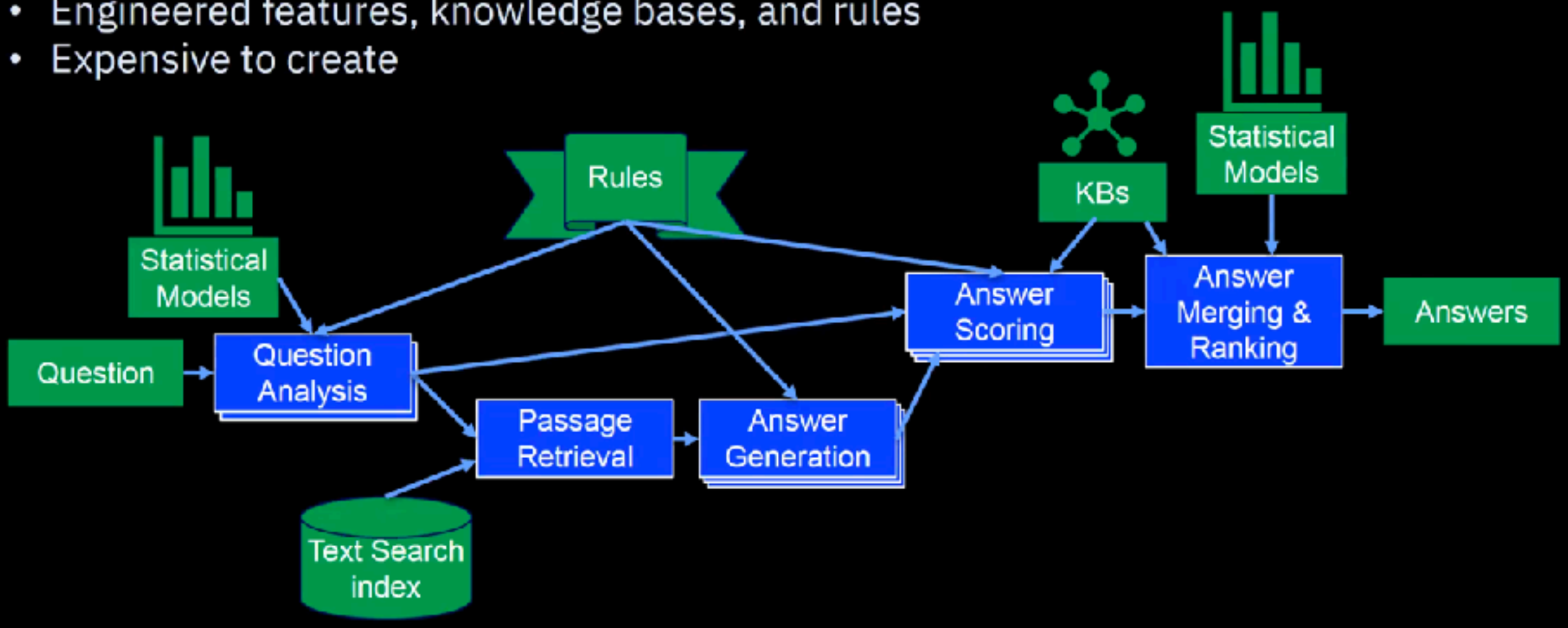


When will we have a generic and universal question answering machine?

Is DeepLearning only sufficient for Question Answering in 2019?

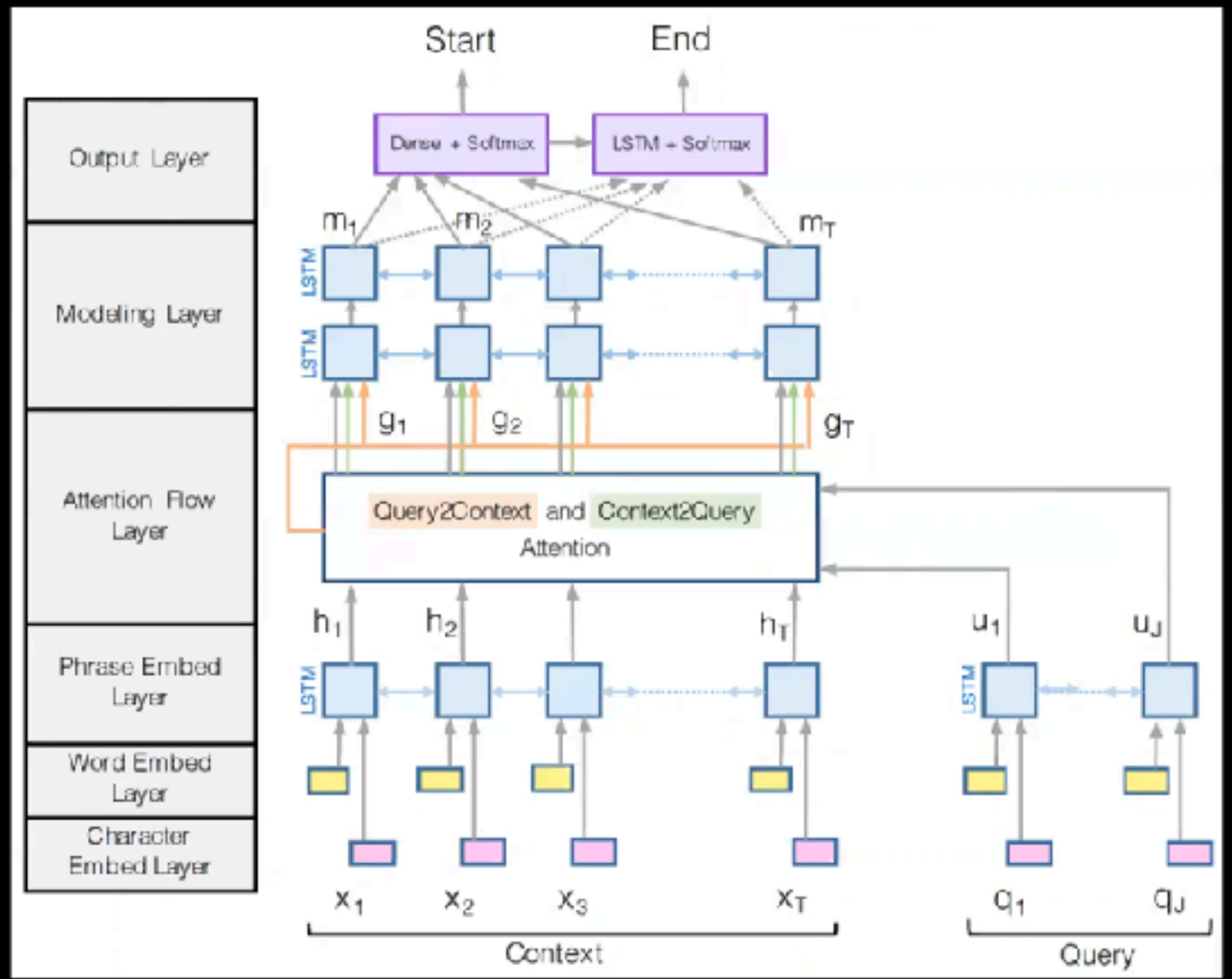
DDQA: Multi-Strategy Factoid Question Answering

- 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

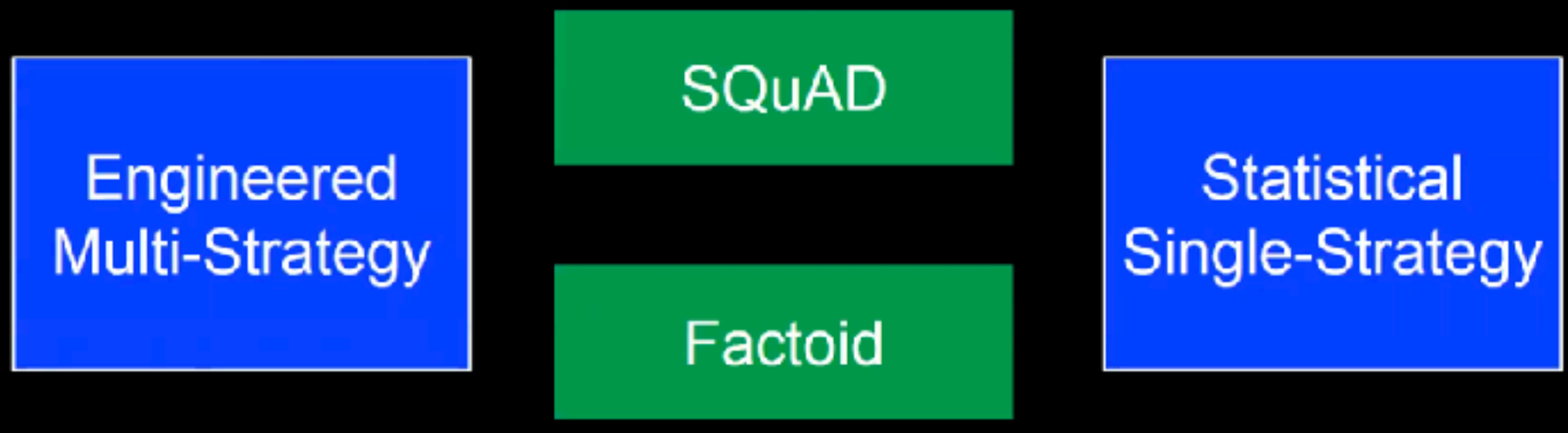
- Off-the-shelf Deep Neural Net
- Some engineering on structure
- No manually engineered features



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



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

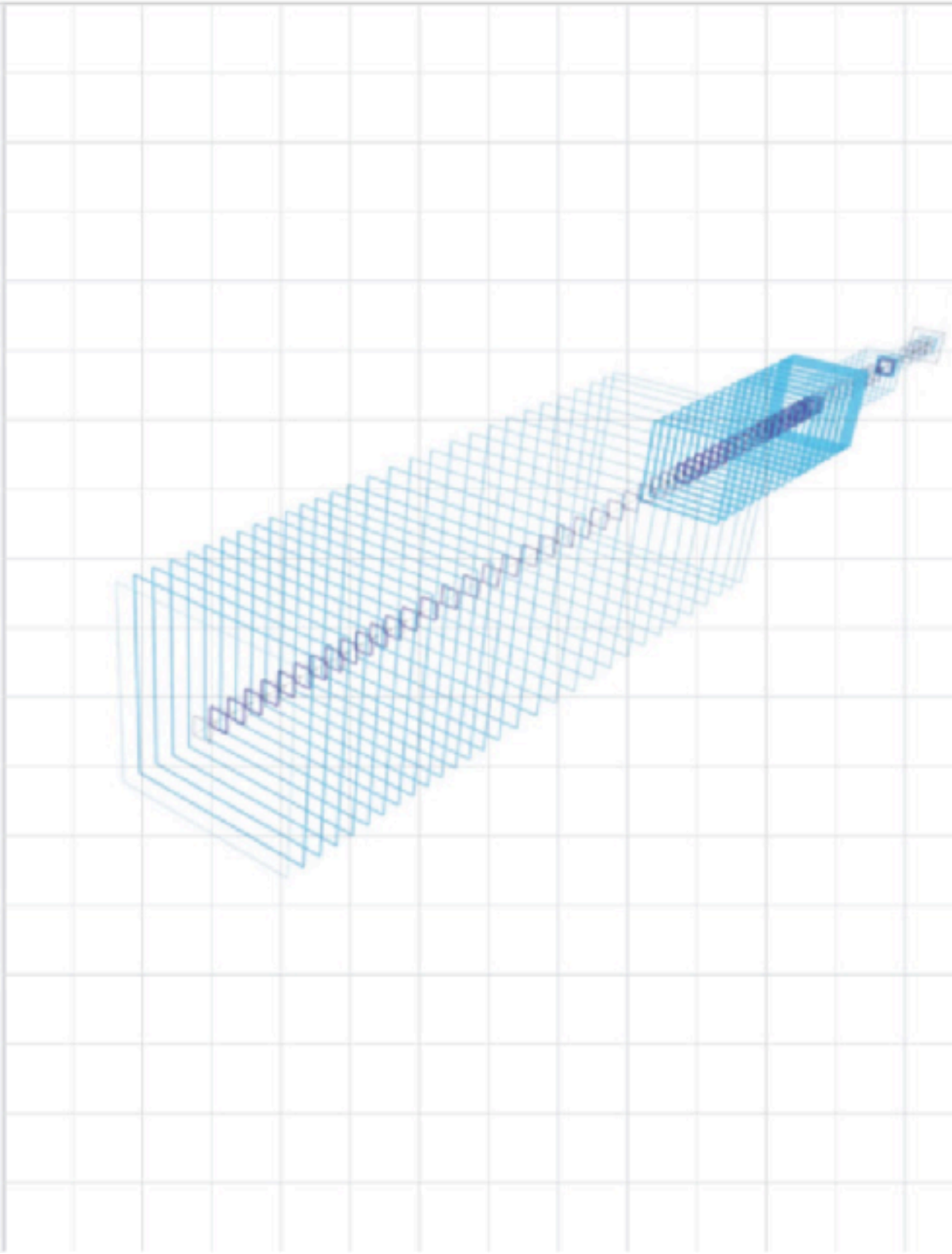
TRAINING DATA

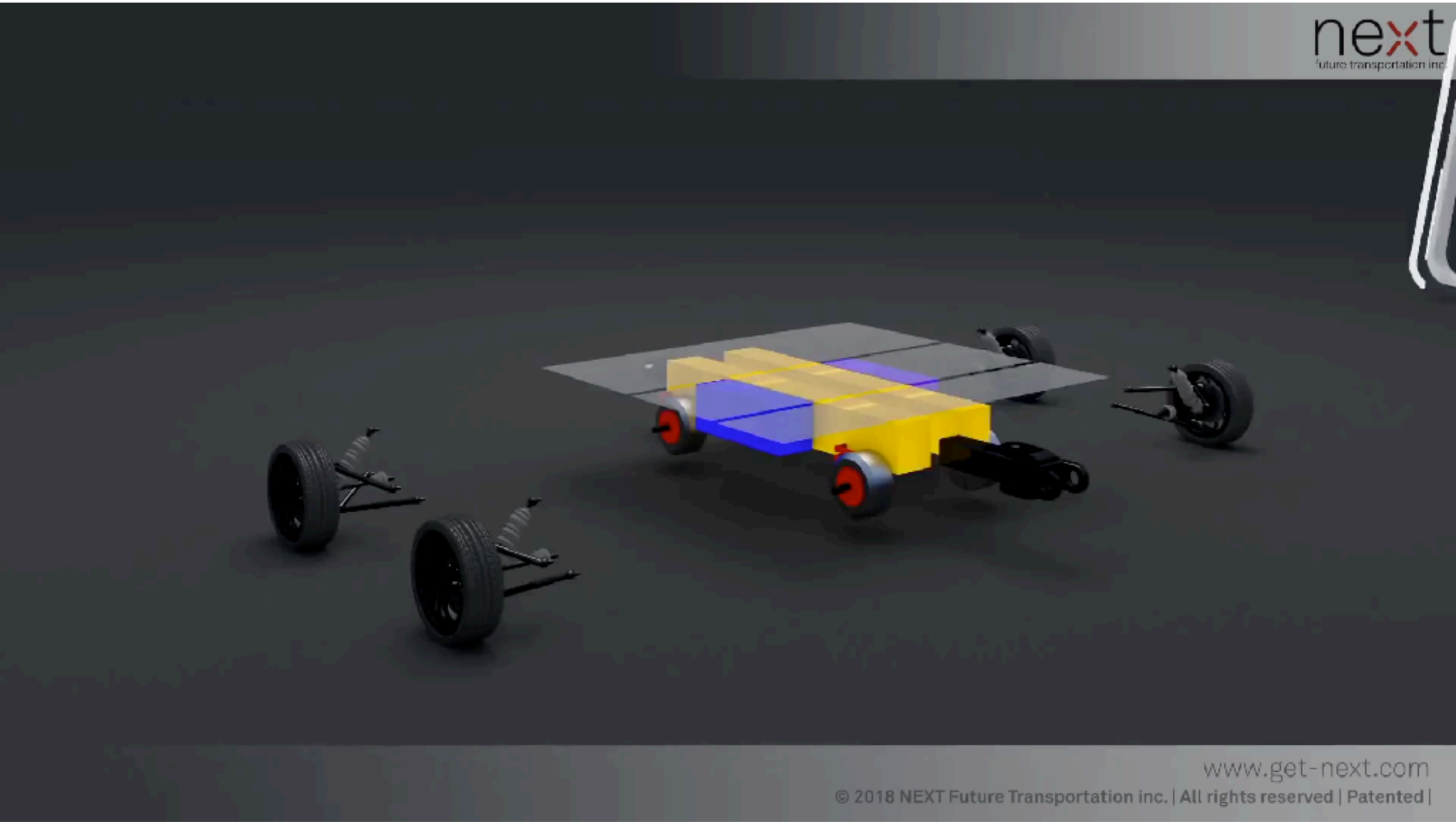
Source	Test Sample
TrainingSet.zip	25%

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




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…

OpenData

Speech Recognition:
Proliferation

Cortana

Mycroft

Echo

Siri

Google Home

Apple HomePod


3
teaspoons in
1 tablespoon

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

your TODOs

- Take my course courses => http://nopanic.ai/datascience
- Contribute to Mozilla Common Voice => @mikehenrty
- Encrypt your emails using PGP (or S/MIME)
- Use a privacy preserving messenger
- Use open Android alternative
- Choose your apps and permissions carefully

If data is oil for machine learning, data is plutonium for DeepLearning