Data and AI – What's Happening

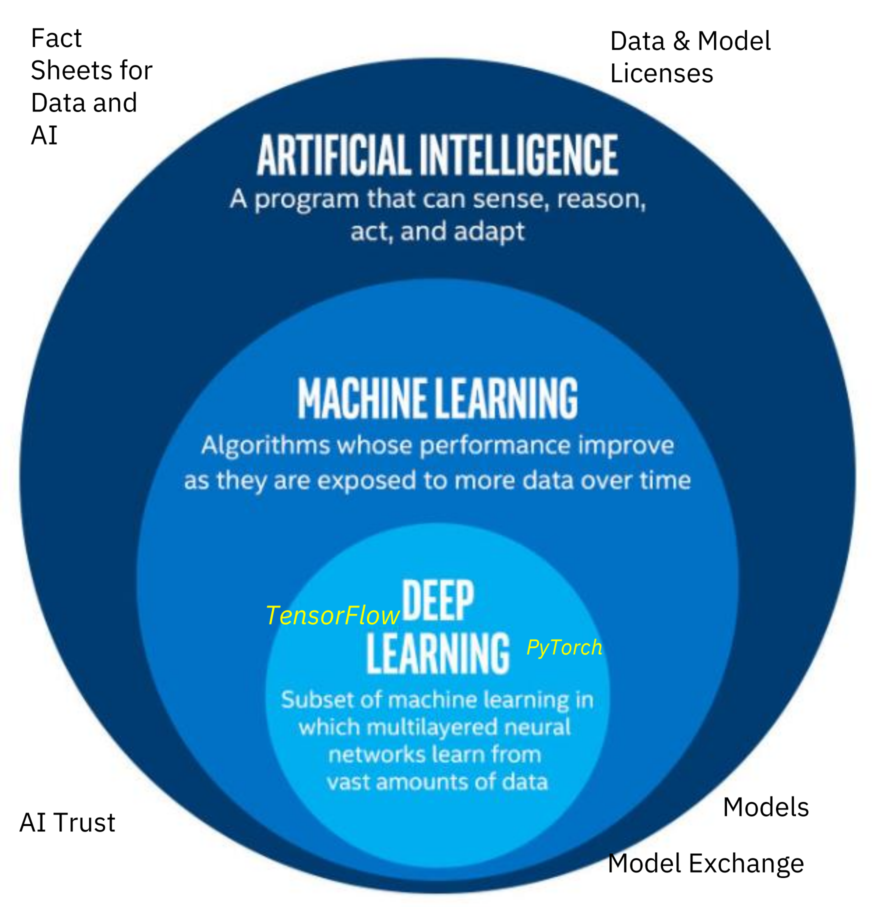
Susan Malaika
Senior Technical Staff

@sumalaika
malaika@us.ibm.com
https://developer.ibm.com/opentech/category/susan-malaika/

IBM **Developer**

IBM

Artificial Intelligence, Machine Learning & Deep Learning



Fact Sheets for Data and AI

Data & Model Licenses

ARTIFICIAL INTELLIGENCE

A program that can sense, reason, act, and adapt

MACHINE LEARNING

Algorithms whose performance improve as they are exposed to more data over time

TensorFlow

DEEP LEARNING

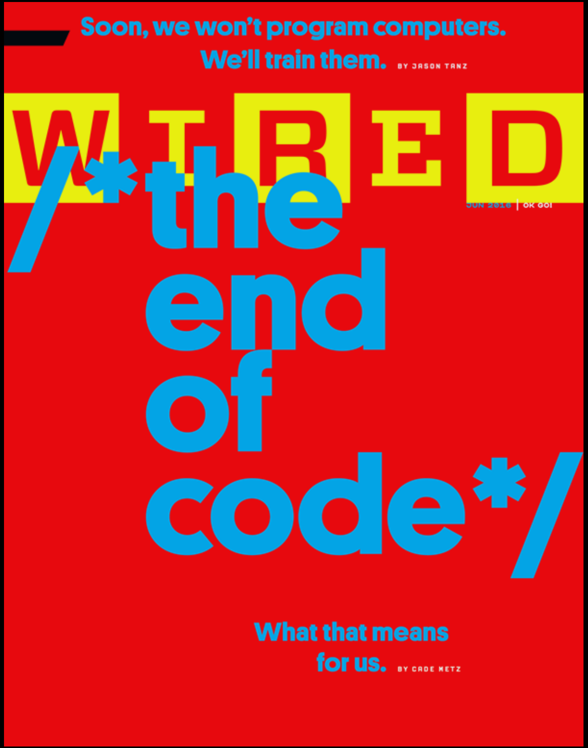
PyTorch

Subset of machine learning in which multilayered neural networks learn from vast amounts of data

AI Trust

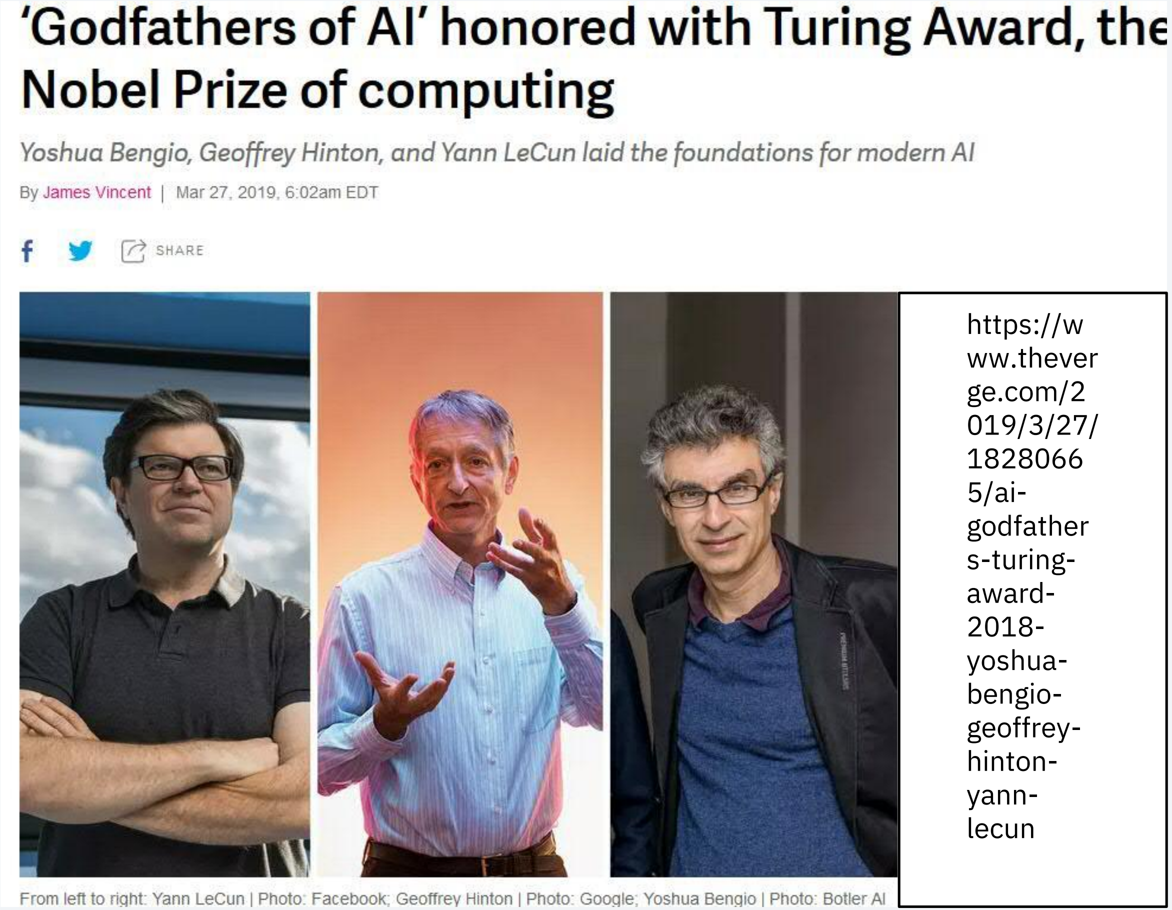
Models

Model Exchange


Soon, we won't program computers. We'll train them. BY JASON TANZ
WIRED
JUN 2016 | OK GO!
/*the end of code*/
What that means for us. BY CADE METZ

Training Computers instead of Programming Them

2019 Turing Award



‘Godfathers of AI’ honored with Turing Award, the Nobel Prize of computing

Yoshua Bengio, Geoffrey Hinton, and Yann LeCun laid the foundations for modern AI

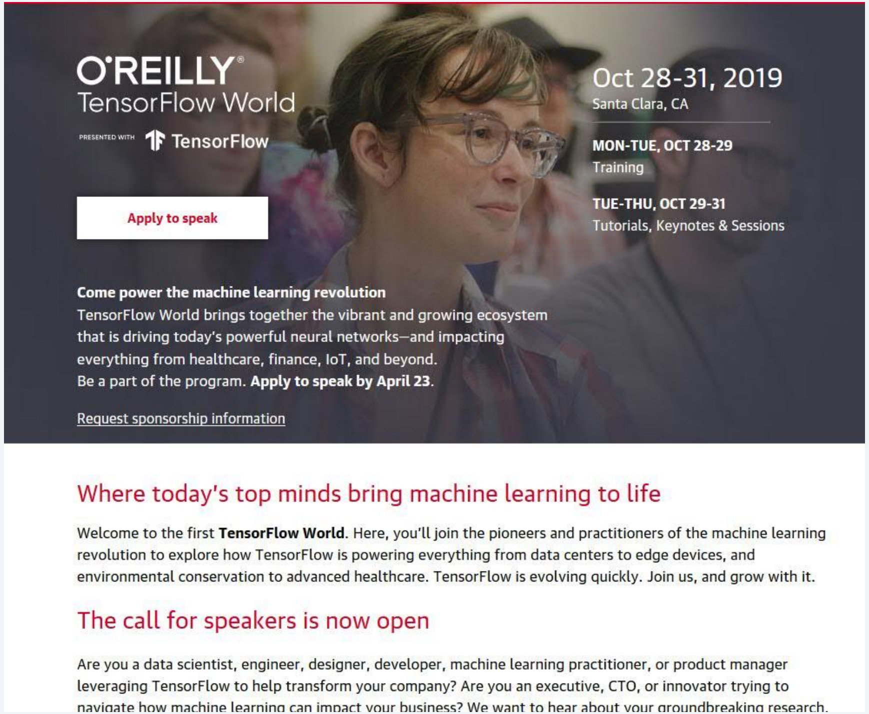
By James Vincent | Mar 27, 2019, 6:02am EDT

From left to right: Yann LeCun | Photo: Facebook; Geoffrey Hinton | Photo: Google; Yoshua Bengio | Photo: Botler AI

https://www.theverge.com/2019/3/27/18280665/ai-godfathers-turing-award-2018-yoshua-bengio-geoffrey-hinton-yann-lecun

First Tensorflow Conference

2019 - Conference



O'REILLY®
TensorFlow World
PRESENTED WITH TensorFlow

Oct 28-31, 2019
Santa Clara, CA

MON-TUE, OCT 28-29
Training

TUE-THU, OCT 29-31
Tutorials, Keynotes & Sessions

Apply to speak

Come power the machine learning revolution
TensorFlow World brings together the vibrant and growing ecosystem that is driving today's powerful neural networks—and impacting everything from healthcare, finance, IoT, and beyond.
Be a part of the program. **Apply to speak by April 23.**

Request sponsorship information

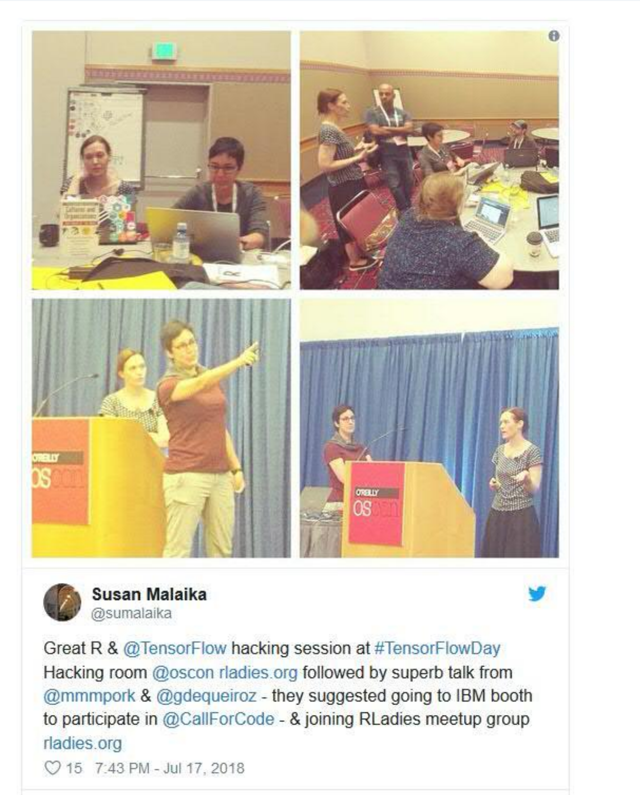
Where today's top minds bring machine learning to life

Welcome to the first **TensorFlow World**. Here, you'll join the pioneers and practitioners of the machine learning revolution to explore how TensorFlow is powering everything from data centers to edge devices, and environmental conservation to advanced healthcare. TensorFlow is evolving quickly. Join us, and grow with it.

The call for speakers is now open

Are you a data scientist, engineer, designer, developer, machine learning practitioner, or product manager leveraging TensorFlow to help transform your company? Are you an executive, CTO, or innovator trying to navigate how machine learning can impact your business? We want to hear about your groundbreaking research.

2018 – Community Days



Susan Malaika
@sumalaika

Great R & @TensorFlow hacking session at #TensorFlowDay Hacking room @oscon rladies.org followed by superb talk from @mmmpork & @gdequeiroz - they suggested going to IBM booth to participate in @CallForCode - & joining RLadies meetup group rladies.org

♡ 15 7:43 PM - Jul 17, 2018

Center for Open Source Data and AI Technologies

Code - Build and improve practical frameworks to enable more developers to realize immediate value (e.g. FfDL, MAX, Tensorflow, Jupyter, Spark)

Content – Showcase solutions to complex and real world AI problems

Community – Bring developers and data scientists to engage with IBM (e.g. MAX)

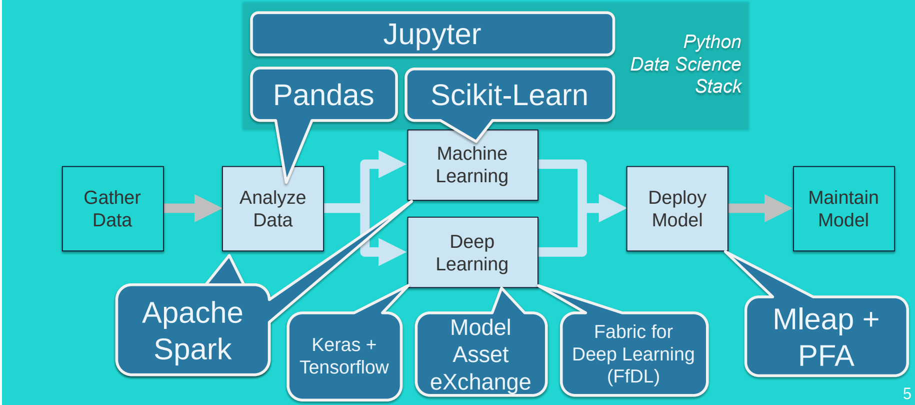
CODAIT

codait.org

codait (French)
= coder/coded

https://m.interglot.com/fr/en/codait

Improving Enterprise AI lifecycle in Open Source



Jupyter

Pandas

Scikit-Learn

Python Data Science Stack

Gather Data

Analyze Data

Machine Learning

Deep Learning

Deploy Model

Maintain Model

Apache Spark

Keras + Tensorflow

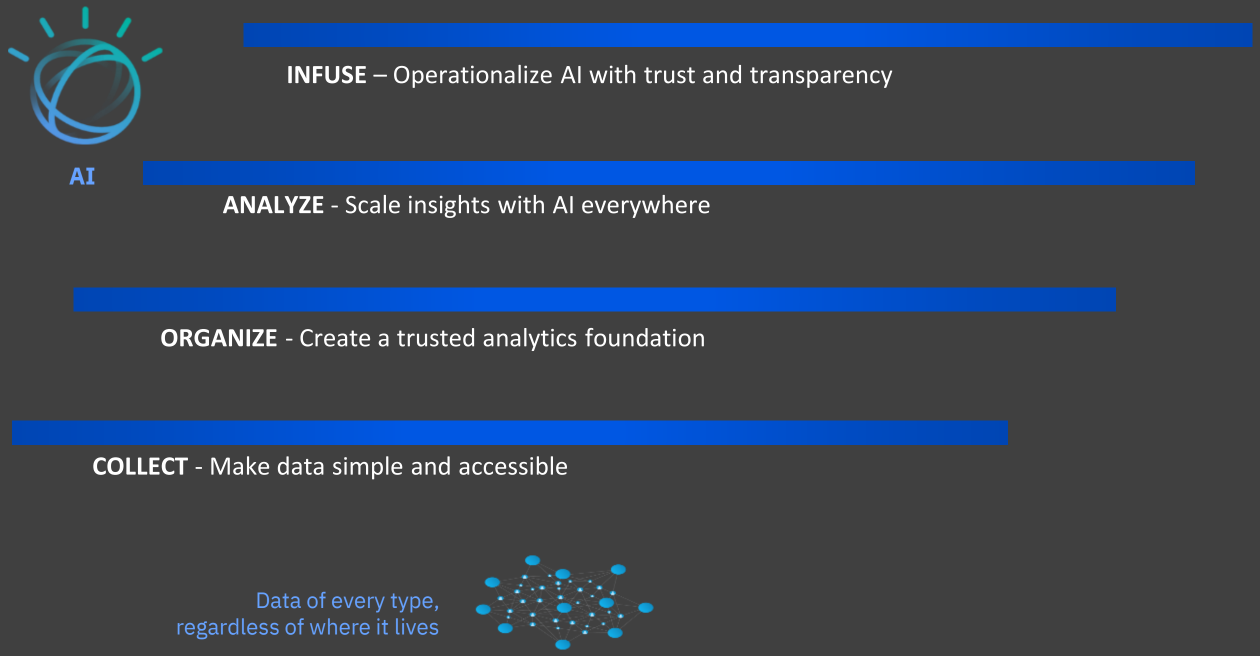
Model Asset eXchange

Fabric for Deep Learning (FfDL)

Mleap + PFA

The AI Ladder

A prescriptive approach to accelerating the journey to AI



INFUSE – Operationalize AI with trust and transparency

ANALYZE - Scale insights with AI everywhere

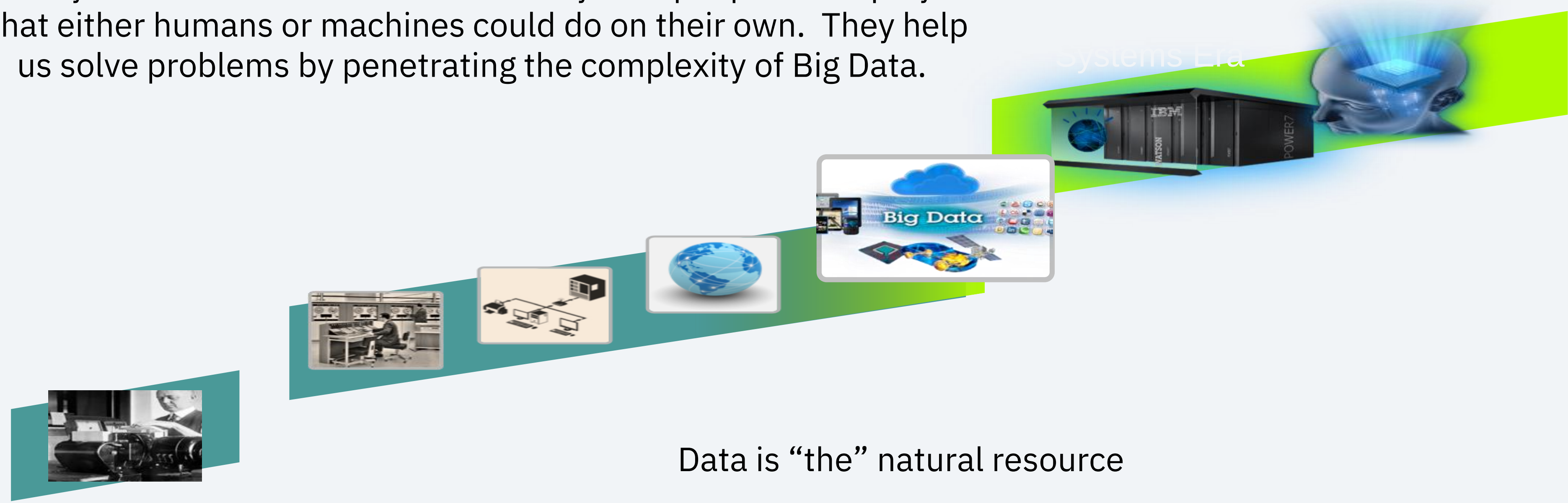
ORGANIZE - Create a trusted analytics foundation

COLLECT - Make data simple and accessible

Data of every type,
regardless of where it lives

Eras of Computing

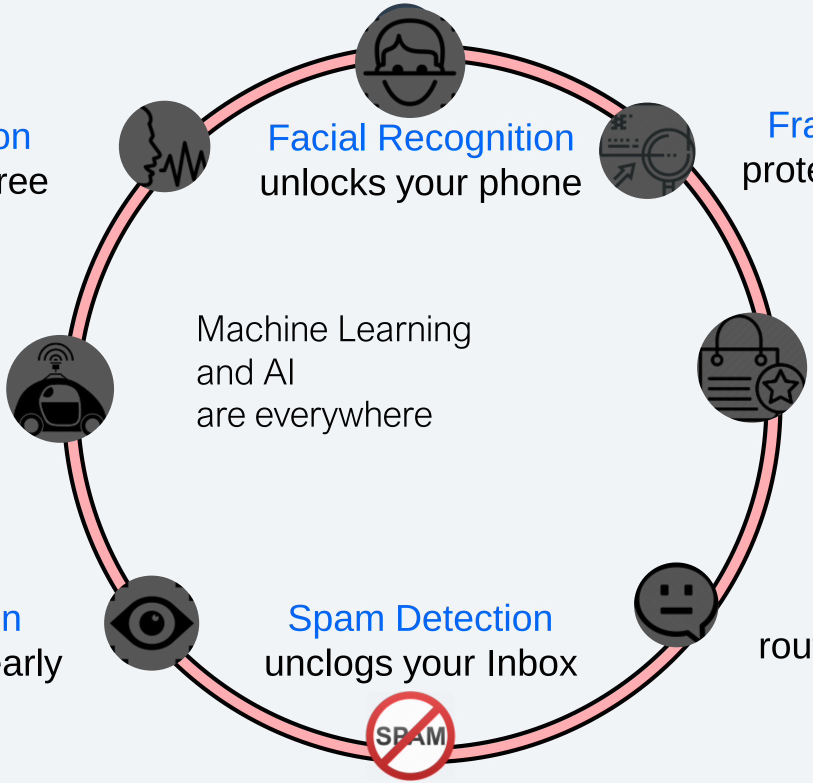
AI Systems learn and interact naturally with people to amplify what either humans or machines could do on their own. They help us solve problems by penetrating the complexity of Big Data.



Data is “the” natural resource

AI is “the” approach to exploit that resource

AI Applications



Machine Learning
and AI
are everywhere

Speech Recognition
lets you go hands-free

Facial Recognition
unlocks your phone

Fraud Detection
protects your credit

Autonomous Vehicles
detect pedestrians

Recommendations
help you shop faster

Machine Vision
detects cancer early

Spam Detection
unclogs your Inbox

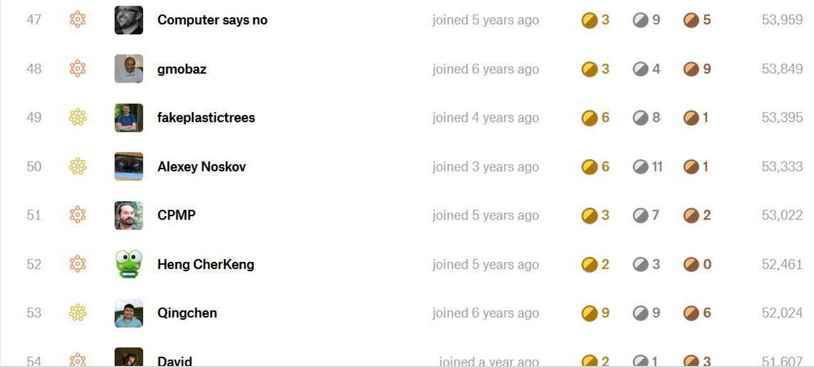
Chat Bots
route calls quicker

The Emergence of LeaderBoards in AI

Leaderboards – such as those in Kaggle - the home of data science contests that utilize open tech & datasets for predictive modeling – resulting in:

- Ranking of data scientists world-wide
- Fresh datasets, data science models, methods, and education - all in open source
- Coursera class https://www.coursera.org/learn/competitive-data-science/home/welcome
- Companies (that sponsor datasets and contests) who benefit through:
    - Recruitment of great employees ; Eminence of own employees; Excellent publicity ; Better understanding of what can be done with their data, Being part of a global AI conversation around open technologies

Kaggle: "a way to organize the brainpower of the world's most talented data scientists and make it accessible to organizations of every size" - Hal Varian, Google



47	Computer says no	joined 5 years ago	● 3	● 9	● 5	53,959
48	gmobaz	joined 6 years ago	● 3	● 4	● 9	53,849
49	fakeplastictrees	joined 4 years ago	● 6	● 8	● 1	53,395
50	Alexey Noskov	joined 3 years ago	● 6	● 11	● 1	53,333
51	CPMP	joined 5 years ago	● 3	● 7	● 2	53,022
52	Heng CherKeng	joined 5 years ago	● 2	● 3	● 0	52,461
53	Qingchen	joined 6 years ago	● 9	● 9	● 6	52,024
54	David	joined a year ago	● 2	● 1	● 3	51,607

Sample contest:

Women in Data Science Datathon Feb 2018

Xi Lui and Ye Wang, Worcester Polytechnic Institute's data science graduate program, beat 230 teams composed of students, faculty, and professional data scientists from 26 countries.

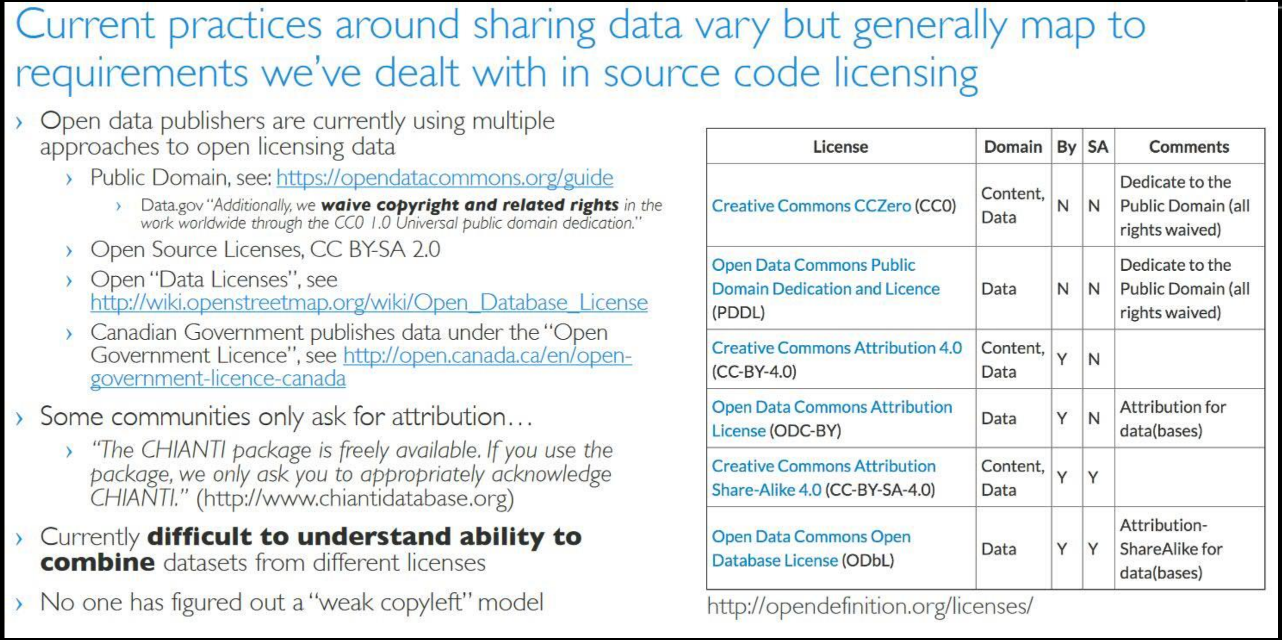
IBM had 12 wonderful teams in the contest (more than any other institution) - the highest ranked was at 7

The contest goal was to predict if a person is male or female by examining the responses the people gave to some questions.

Community Data License Agreement

There are two CDLA license agreements:

- “Sharing” based on a form of copyleft designed to encourage recipients to participate in reciprocal sharing of data
- “Permissive” an approach similar to permissive open source licenses (e.g. Apache, BSD or MIT) where recipients are not required to share any changes



Current practices around sharing data vary but generally map to requirements we’ve dealt with in source code licensing

- Open data publishers are currently using multiple approaches to open licensing data
  - Public Domain, see: https://opendatacommons.org/guide
    - Data.gov “Additionally, we **waive copyright and related rights** in the work worldwide through the CC0 1.0 Universal public domain dedication.”
  - Open Source Licenses, CC BY-SA 2.0
  - Open “Data Licenses”, see http://wiki.openstreetmap.org/wiki/Open_Database_License
  - Canadian Government publishes data under the “Open Government Licence”, see http://open.canada.ca/en/open-government-licence-canada
- Some communities only ask for attribution…
  - *“The CHIANTI package is freely available. If you use the package, we only ask you to appropriately acknowledge CHIANTI.”* (http://www.chiantidatabase.org)
- Currently **difficult to understand ability to combine** datasets from different licenses
- No one has figured out a “weak copyleft” model

License	Domain	By	SA	Comments
Creative Commons CCZero (CC0)	Content, Data	N	N	Dedicate to the Public Domain (all rights waived)
Open Data Commons Public Domain Dedication and Licence (PDDL)	Data	N	N	Dedicate to the Public Domain (all rights waived)
Creative Commons Attribution 4.0 (CC-BY-4.0)	Content, Data	Y	N	
Open Data Commons Attribution License (ODC-BY)	Data	Y	N	Attribution for data(bases)
Creative Commons Attribution Share-Alike 4.0 (CC-BY-SA-4.0)	Content, Data	Y	Y	
Open Data Commons Open Database License (ODbL)	Data	Y	Y	Attribution-ShareAlike for data(bases)

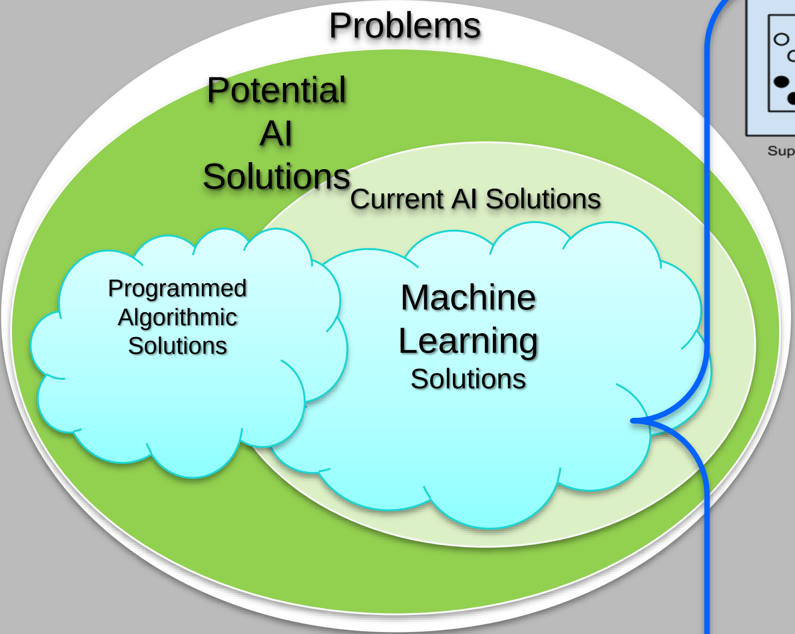
http://opendefinition.org/licenses/

Candidate users of CDLA:

Communities training AI and ML systems
Public-private infrastructure (e.g. data on traffic)
Researchers
Companies with mutual interests in sharing data

License Announced in November 2017 by Linux Foundation
https://www.linuxfoundation.org/press-release/2017/10/linux-foundation-debuts-community-data-license-agreement/

More on AI



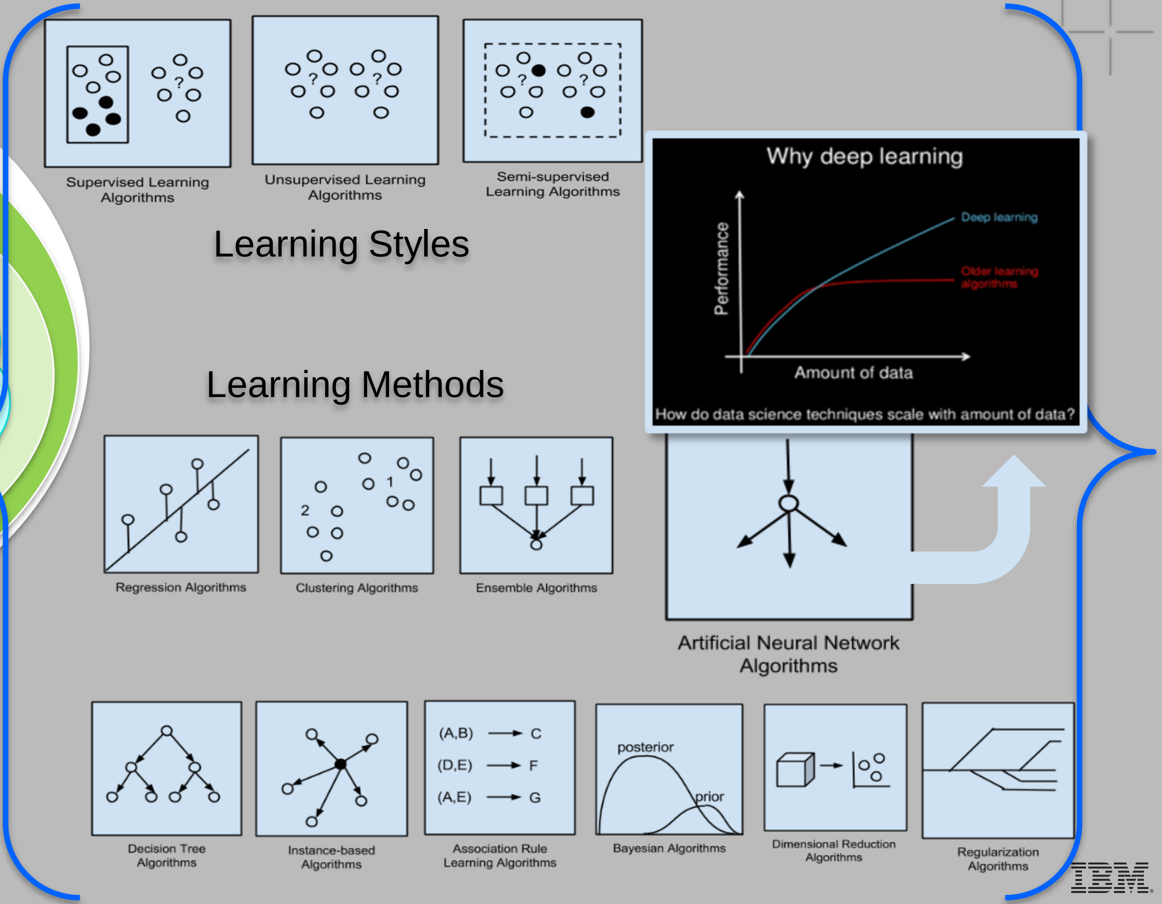
Problems

Potential AI Solutions

Current AI Solutions

Programmed Algorithmic Solutions

Machine Learning Solutions



Learning Styles

Supervised Learning Algorithms

Unsupervised Learning Algorithms

Semi-supervised Learning Algorithms

Learning Methods

Regression Algorithms

Clustering Algorithms

Ensemble Algorithms

Artificial Neural Network Algorithms

Why deep learning

How do data science techniques scale with amount of data?

Decision Tree Algorithms

Instance-based Algorithms

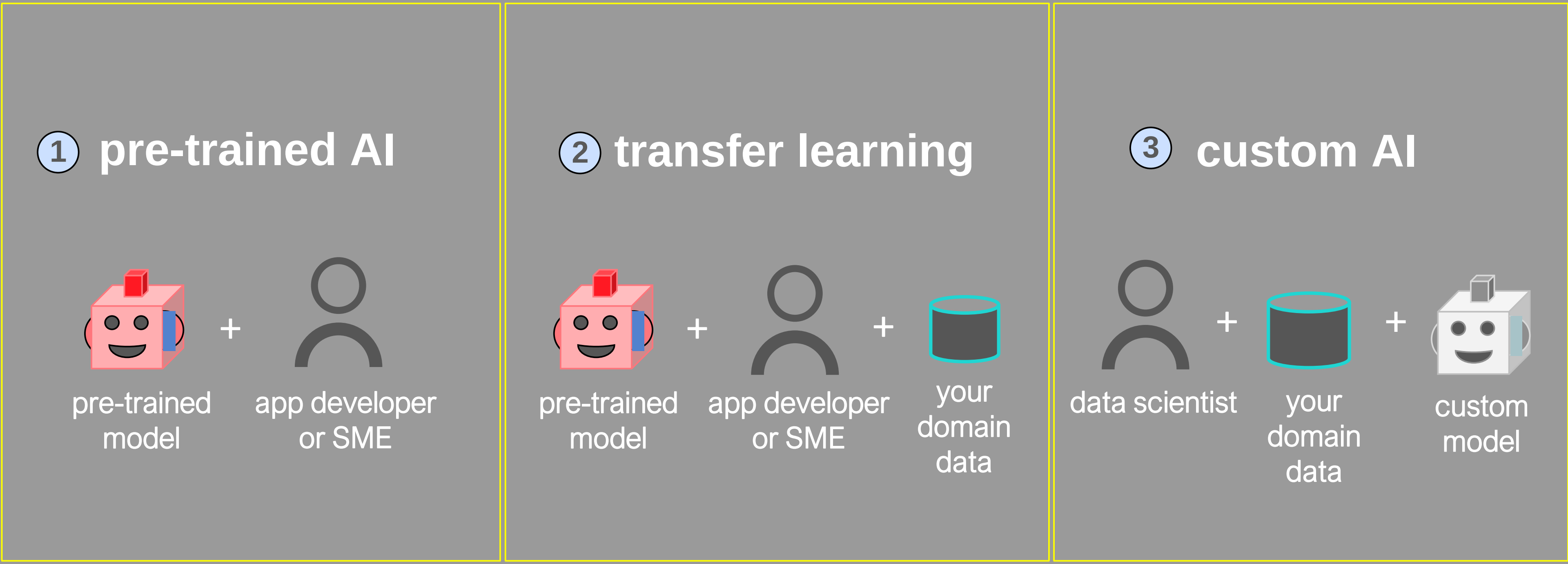
Association Rule Learning Algorithms

Bayesian Algorithms

Dimensional Reduction Algorithms

Regularization Algorithms

Three approaches for building AI Models



1 pre-trained AI

pre-trained model + app developer or SME

2 transfer learning

pre-trained model + app developer or SME + your domain data

3 custom AI

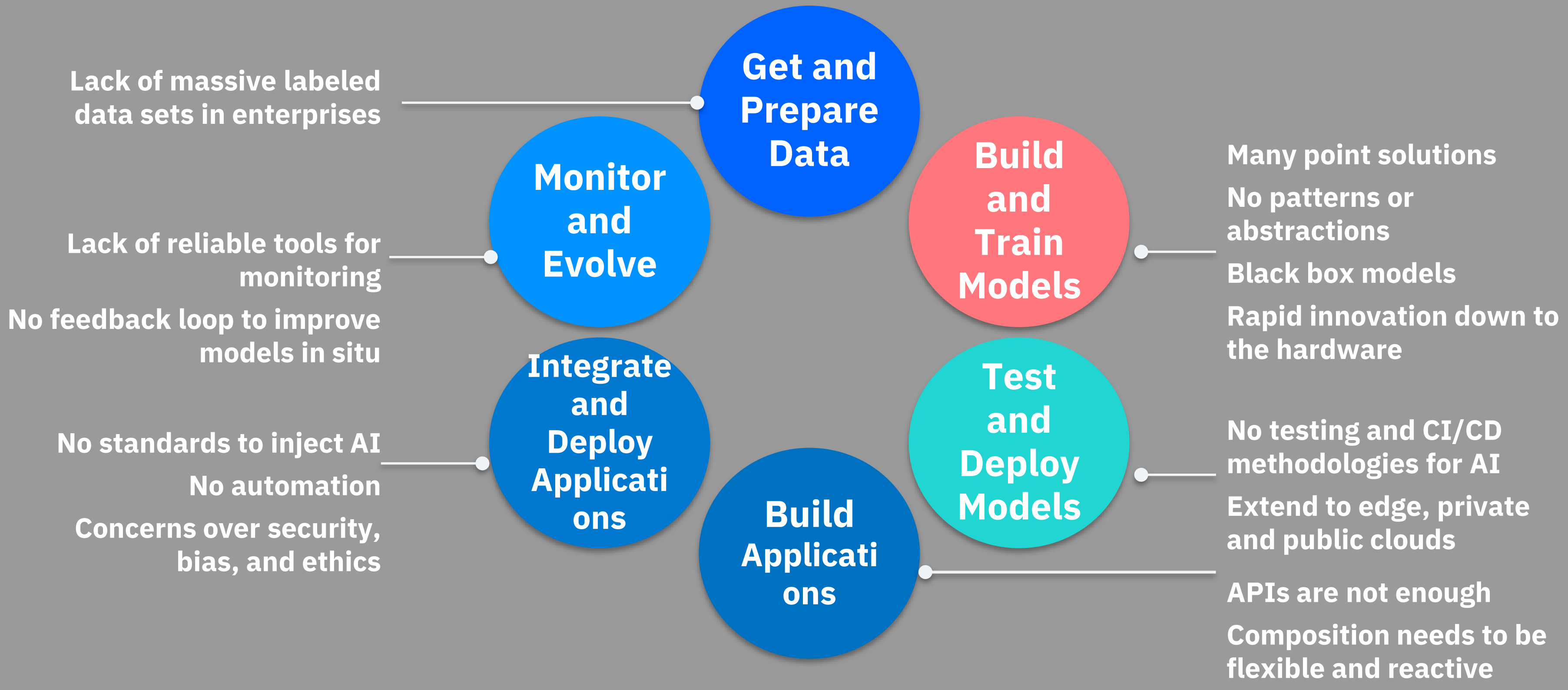
data scientist + your domain data + custom model

Watson Visual Recognition
Natural Language Understanding
Watson Speech to Text
Watson Text to Speech
…

Watson Visual Recognition
Natural Language Classifier
Watson Speech to Text
…

Watson Studio (Mostly open source)
Watson Machine Learning
Deep Learning as a Service
 (FfDL is open source project)
…

Building AI is still hard



Get and Prepare Data

Lack of massive labeled data sets in enterprises

Build and Train Models

Many point solutions

No patterns or abstractions

Black box models

Rapid innovation down to the hardware

Test and Deploy Models

No testing and CI/CD methodologies for AI

Extend to edge, private and public clouds

Build Applications

APIs are not enough

Composition needs to be flexible and reactive

Integrate and Deploy Applications

No standards to inject AI

No automation

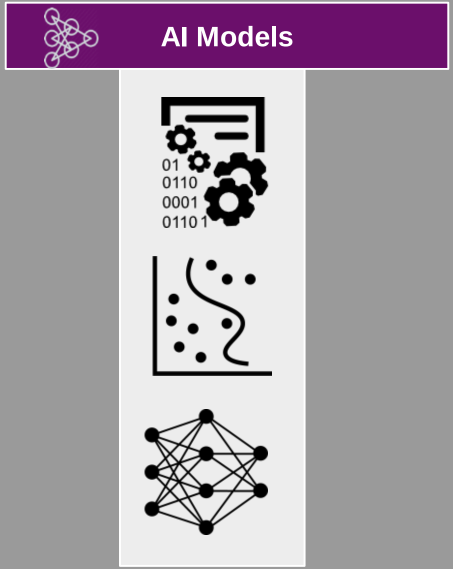
Concerns over security, bias, and ethics

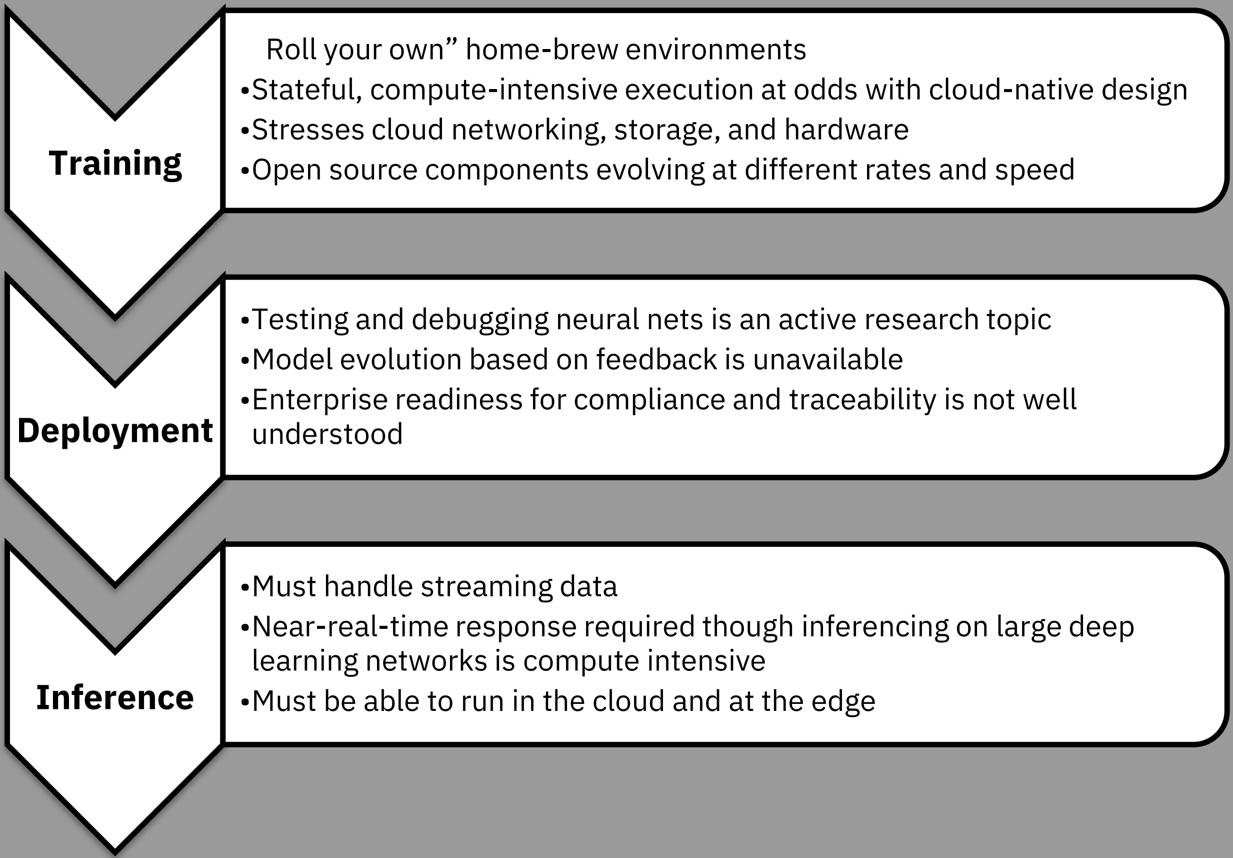
Monitor and Evolve

Lack of reliable tools for monitoring

No feedback loop to improve models in situ

Challenges of Building and deploying AI Models Today





Training

Roll your own” home-brew environments

- Stateful, compute-intensive execution at odds with cloud-native design
- Stresses cloud networking, storage, and hardware
- Open source components evolving at different rates and speed

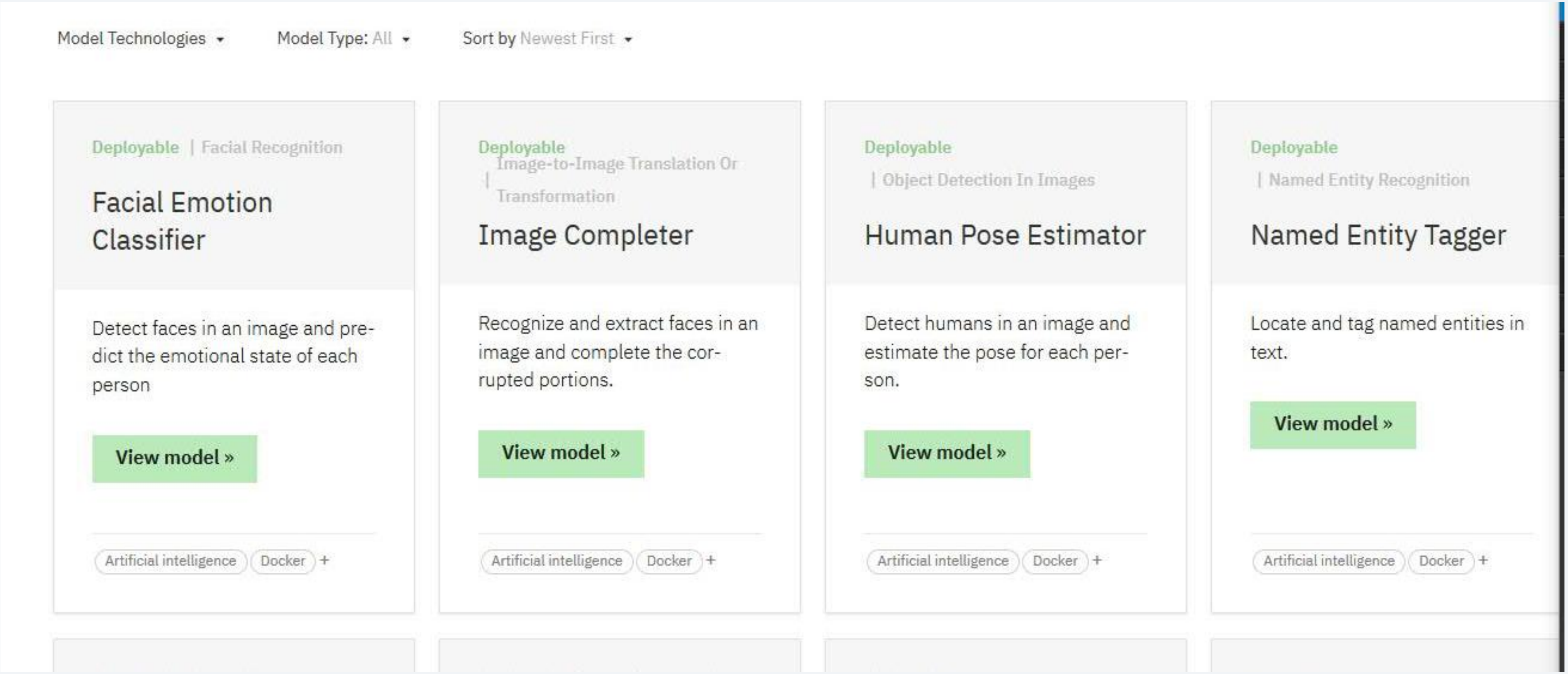
Deployment

- Testing and debugging neural nets is an active research topic
- Model evolution based on feedback is unavailable
- Enterprise readiness for compliance and traceability is not well understood

Inference

- Must handle streaming data
- Near-real-time response required though inferencing on large deep learning networks is compute intensive
- Must be able to run in the cloud and at the edge

Model Asset Exchange
https://developer.ibm.com/exchanges/models/all/



Deployable | Facial Recognition

Facial Emotion Classifier

Detect faces in an image and predict the emotional state of each person

Deployable | Image-to-Image Translation Or Transformation

Image Completer

Recognize and extract faces in an image and complete the corrupted portions.

Deployable | Object Detection In Images

Human Pose Estimator

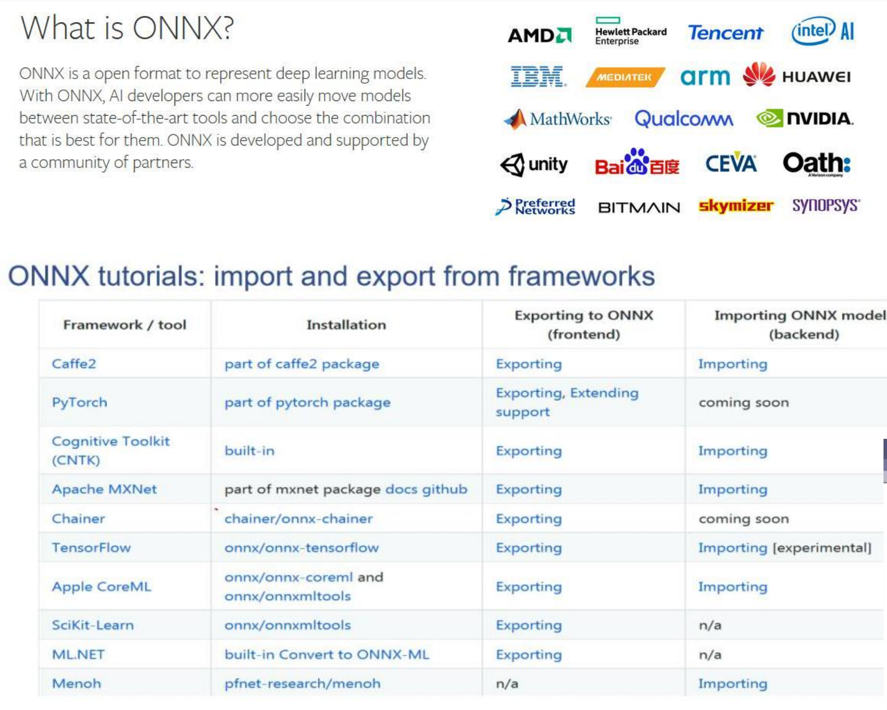
Detect humans in an image and estimate the pose for each person.

Deployable | Named Entity Recognition

Named Entity Tagger

Locate and tag named entities in text.

Storing, Sharing, Reusing, Composing and Deploying Machine Learning Models

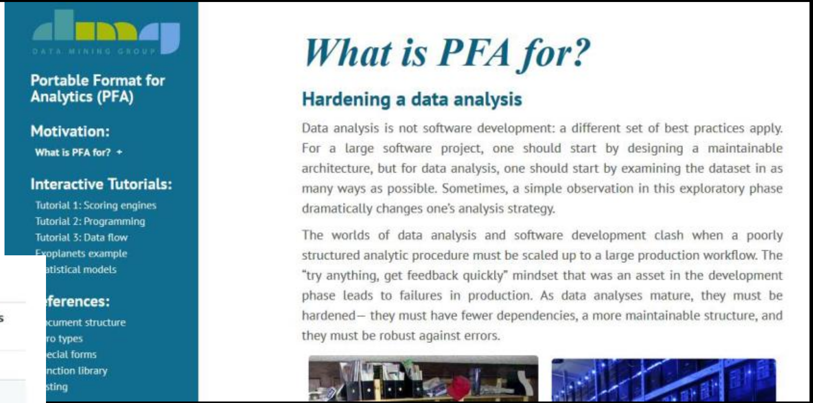


What is ONNX?

ONNX is a open format to represent deep learning models. With ONNX, AI developers can more easily move models between state-of-the-art tools and choose the combination that is best for them. ONNX is developed and supported by a community of partners.

ONNX tutorials: import and export from frameworks

Framework / tool	Installation	Exporting to ONNX (frontend)	Importing ONNX models (backend)
Caffe2	part of caffe2 package	Exporting	Importing
PyTorch	part of pytorch package	Exporting, Extending support	coming soon
Cognitive Toolkit (CNTK)	built-in	Exporting	Importing
Apache MXNet	part of mxnet package docs github	Exporting	Importing
Chainer	chainer/onnx-chainer	Exporting	coming soon
TensorFlow	onnx/onnx-tensorflow	Exporting	Importing [experimental]
Apple CoreML	onnx/onnx-coreml and onnx/onnxmltools	Exporting	Importing
SciKit-Learn	onnx/onnxmltools	Exporting	n/a
ML.NET	built-in Convert to ONNX-ML	Exporting	n/a
Menoh	pfnet-research/menoh	n/a	Importing

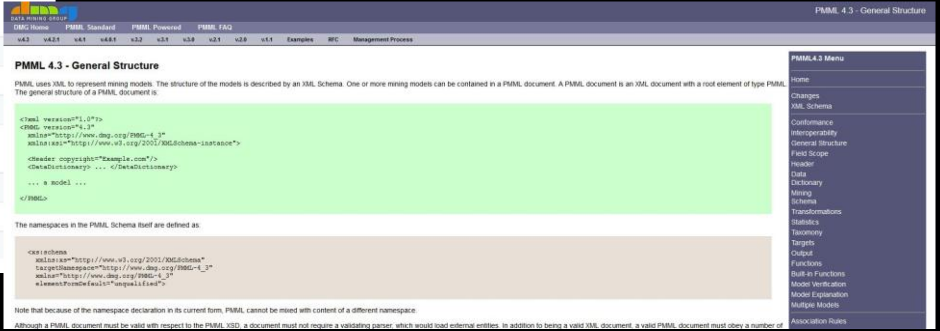


What is PFA for?

Hardening a data analysis

Data analysis is not software development: a different set of best practices apply. For a large software project, one should start by designing a maintainable architecture, but for data analysis, one should start by examining the dataset in as many ways as possible. Sometimes, a simple observation in this exploratory phase dramatically changes one's analysis strategy.

The worlds of data analysis and software development clash when a poorly structured analytic procedure must be scaled up to a large production workflow. The "try anything, get feedback quickly" mindset that was an asset in the development phase leads to failures in production. As data analyses mature, they must be hardened— they must have fewer dependencies, a more maintainable structure, and they must be robust against errors.



PMML 4.3 - General Structure

PMML uses XML to represent mining models. The structure of the models is described by an XML Schema. One or more mining models can be contained in a PMML document. A PMML document is an XML document with a root element of type PMML. The general structure of a PMML document is:

```
<?xml version="1.0"?>
<PMML version="4.3"
  xmlns="http://www.dmg.org/PMML-4_3"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">

  <Header copyright="Example.com"/>
  <DataDictionary> ... </DataDictionary>

  ... a model ...

</PMML>
```

The namespaces in the PMML Schema itself are defined as:

```
  <xs:schema
    xmlns:xs="http://www.w3.org/2001/XMLSchema"
    targetNamespace="http://www.dmg.org/PMML-4_3"
    xmlns="http://www.dmg.org/PMML-4_3"
    elementFormDefault="unqualified">
```

Note that because of the namespace declaration in its current form, PMML cannot be mixed with content of a different namespace.

Although a PMML document must be valid with respect to the PMML XSD, a document must not require a validating parser, which would load external entities. In addition to being a valid XML document, a valid PMML document must obey a number of

AI Engineering: An emerging discipline

Data handling tools

Image/ Video

Audio

Text

Language

DLaaS Cloud Platform & Access to Frameworks

TensorFlow

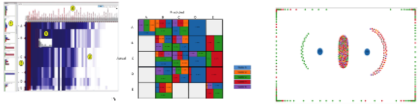
Caffe

theano

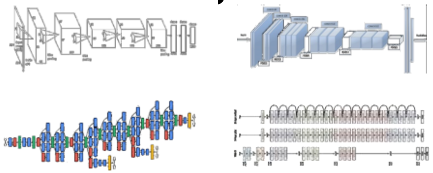
OpenCV

torch

Visualization & Human-Computer interaction



Network optimization tools



Computation & Distributed Learning

Power Systems

NVIDIA

cuDNN

AI Model Lifecycle Management

DEV
OPS
CREATE
PLAN
RELEASE
CONFIGURE
MONITOR
PACKAGE
VERIFY

AI Open Scale

AI Benchmarks

DeepBench : This benchmark targets low-level operations that are fundamental to deep learning, such as matrix-multiplication, convolutions, and communications, and aims to identify the most appropriate hardware but the benchmark does not consider time-to-accuracy.

TensorFlow : The TensorFlow performance benchmarks are similar to DeepBench, in that they identify the most appropriate hardware, but not time-to-accuracy currently. They are also tied to the TensorFlow Framework.

DAWNBench : DAWNBench allows different deep learning methods to be compared by running a number of competitions. It was the first major benchmark suite to examine end-to-end deep learning training and inference. It does not address data preparation and hyper-parameter optimization work.

MLPerf : MLPerf defines the primary metric as the wall clock time to train a model to a target quality, often hours or days. The target quality is based on the current state of the art publication results, less a small delta to allow for run-to-run variance. MLPerf does not address hyper-parameter optimization nor data preparation.

- The **MLPerf Closed Model Division** specifies the model to be used and restricts the values of the hyper parameters (batch size, learning rate, etc.) which can be tuned in an attempt to create a fair and balance comparison of the hardware and software systems.
- The **MLPerf Open Model Division**, only requires that same task must be achieved using the same data, but provides fewer restrictions

AI is now used in many high-stakes decision making applications

Credit

Employment

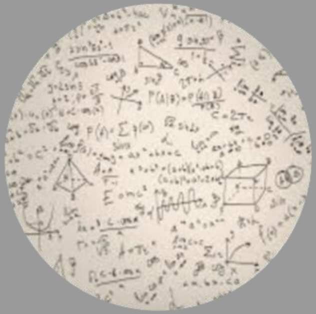
Admission

Sentencing

What does it take to trust a decision made by a machine?

(Other than that it is 99% accurate)

Is it fair?



Is it easy to understand?

Did anyone tamper with it?

1866.
Buckmaster. 13.13.0.
Shed lent S. Hand. Ret.
27. Stidolph. 3.3.0
28. Inge. 3.0.0.
28. Case. 6.6.0.
29. Eastwood. hire S. Hand.
29. Farnham. 1.5.0. S. Hand.
January. 1867.
Duck. 4.10.0. S. Hand
Brickell 6.6.0.

Is it accountable?

IBM Vision for Trusted AI

Pillars of trust, woven into the lifecycle of an AI application

FAIRNESS

EXPLAINABILITY

ROBUSTNESS

ASSURANCE

supported by an instrumented platform

AI OpenScale

AI learns whatever the data teaches it

Image Search

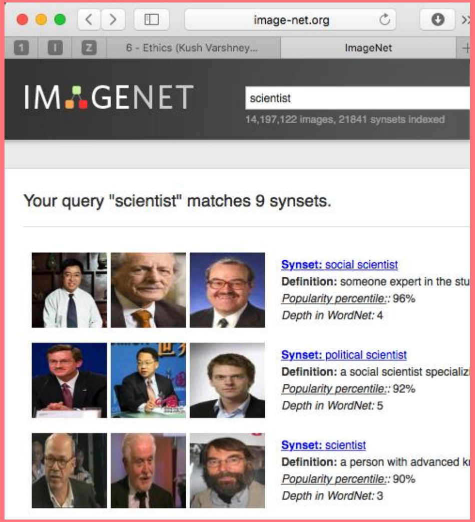


image-net.org

6 - Ethics (Kush Varshney...

ImageNet

IMAGENET

scientist

14,197,122 images, 21841 synsets indexed

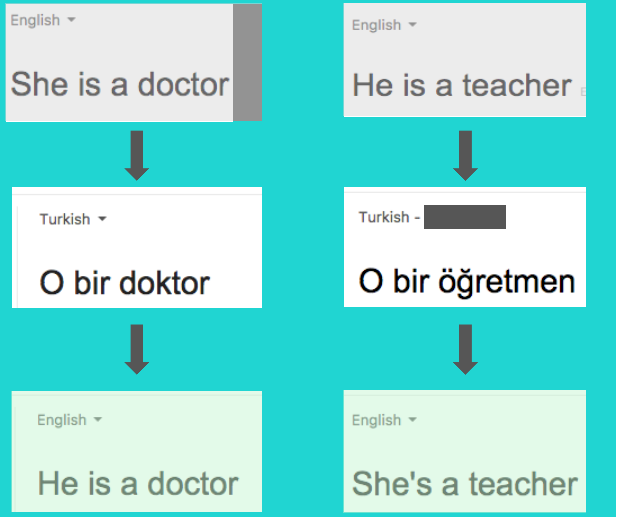
Your query "scientist" matches 9 synsets.

Synset: social scientist
Definition: someone expert in the stu
Popularity percentile:: 96%
Depth in WordNet: 4

Synset: political scientist
Definition: a social scientist speciali
Popularity percentile:: 92%
Depth in WordNet: 5

Synset: scientist
Definition: a person with advanced k
Popularity percentile:: 90%
Depth in WordNet: 3

Language Translation



English

She is a doctor

Turkish

O bir doktor

English

He is a doctor

English

He is a teacher

Turkish - [redacted]

O bir öğretmen

English

She's a teacher

Google Translate: English / Turkish

Chatbot Interactions

The Twitter profile picture of Tay

Microsoft Tay chatted racist and xenophobic epithets learned from interacting with people

Unwanted bias and algorithmic fairness

Machine learning, by its very nature, is always a form of statistical discrimination

Discrimination becomes objectionable when it places certain privileged groups at systematic advantage and certain unprivileged groups at systematic disadvantage.

Illegal in certain contexts:

> e.g. Equal Credit Opportunity, The Equal Pay Act, The Americans With Disabilities Act, …

… but not well understood in others.

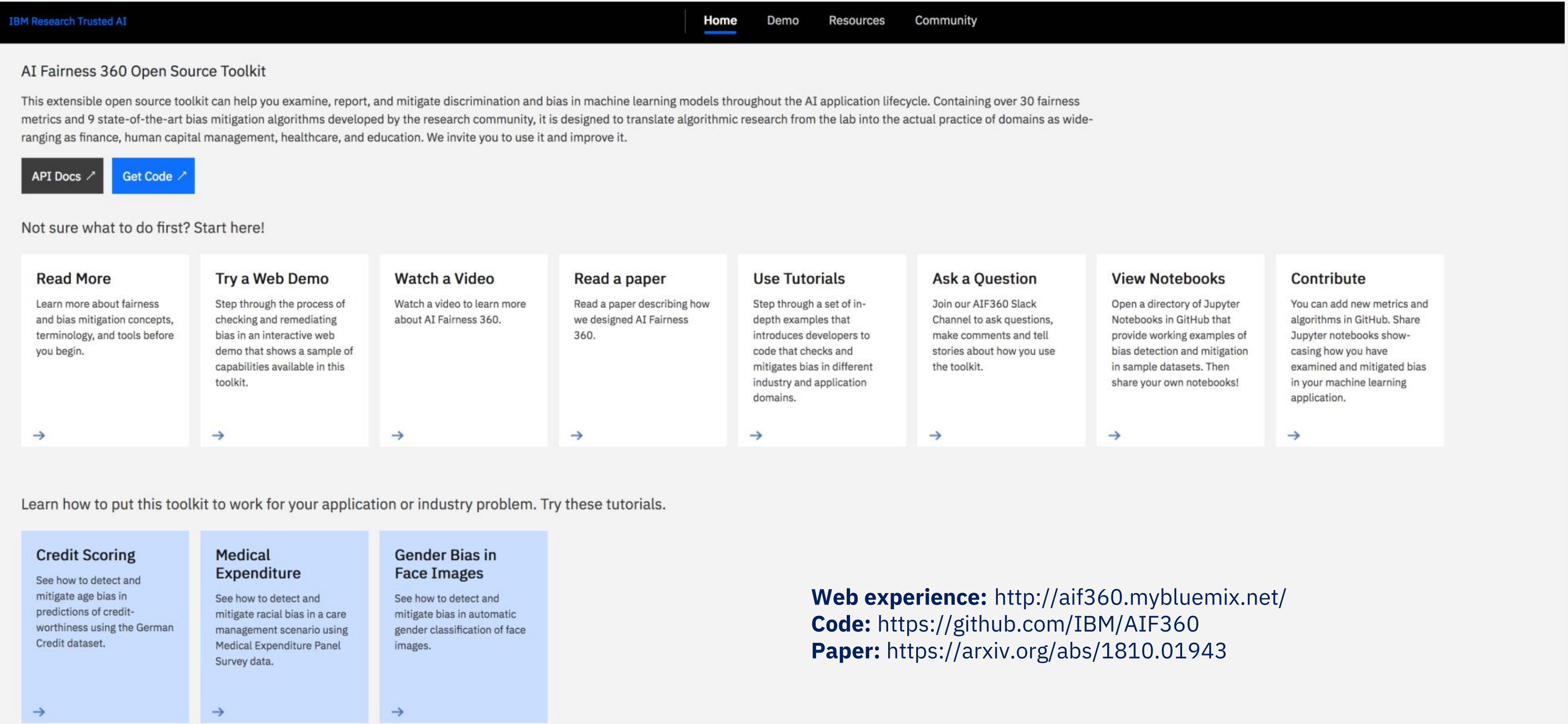
Unwanted bias in training data yields models that scale the bias out

- Prejudice in labels,
- undersampling or oversampling,

- … but bias can creep in due to incorrect model build, selection or deployment.

AI Fairness 360

An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias



AI Fairness 360 Open Source Toolkit

This extensible open source toolkit can help you examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. Containing over 30 fairness metrics and 9 state-of-the-art bias mitigation algorithms developed by the research community, it is designed to translate algorithmic research from the lab into the actual practice of domains as wide-ranging as finance, human capital management, healthcare, and education. We invite you to use it and improve it.

Not sure what to do first? Start here!

Read More Learn more about fairness and bias mitigation concepts, terminology, and tools before you begin. →

Try a Web Demo Step through the process of checking and remediating bias in an interactive web demo that shows a sample of capabilities available in this toolkit. →

Watch a Video Watch a video to learn more about AI Fairness 360. →

Read a paper Read a paper describing how we designed AI Fairness 360. →

Use Tutorials Step through a set of in-depth examples that introduces developers to code that checks and mitigates bias in different industry and application domains. →

Ask a Question Join our AIF360 Slack Channel to ask questions, make comments and tell stories about how you use the toolkit. →

View Notebooks Open a directory of Jupyter Notebooks in GitHub that provide working examples of bias detection and mitigation in sample datasets. Then share your own notebooks! →

Contribute You can add new metrics and algorithms in GitHub. Share Jupyter notebooks showcasing how you have examined and mitigated bias in your machine learning application. →

Learn how to put this toolkit to work for your application or industry problem. Try these tutorials.

Credit Scoring See how to detect and mitigate age bias in predictions of credit-worthiness using the German Credit dataset. →

Medical Expenditure See how to detect and mitigate racial bias in a care management scenario using Medical Expenditure Panel Survey data. →

Gender Bias in Face Images See how to detect and mitigate bias in automatic gender classification of face images. →

Web experience: http://aif360.mybluemix.net/
Code: https://github.com/IBM/AIF360
Paper: https://arxiv.org/abs/1810.01943

Trusted AI

AI Fairness 360:

- **30+ fairness metrics/checkers**
- **10 bias "mitigators"**
- **industry tutorials**

aif360.mybluemix.net/

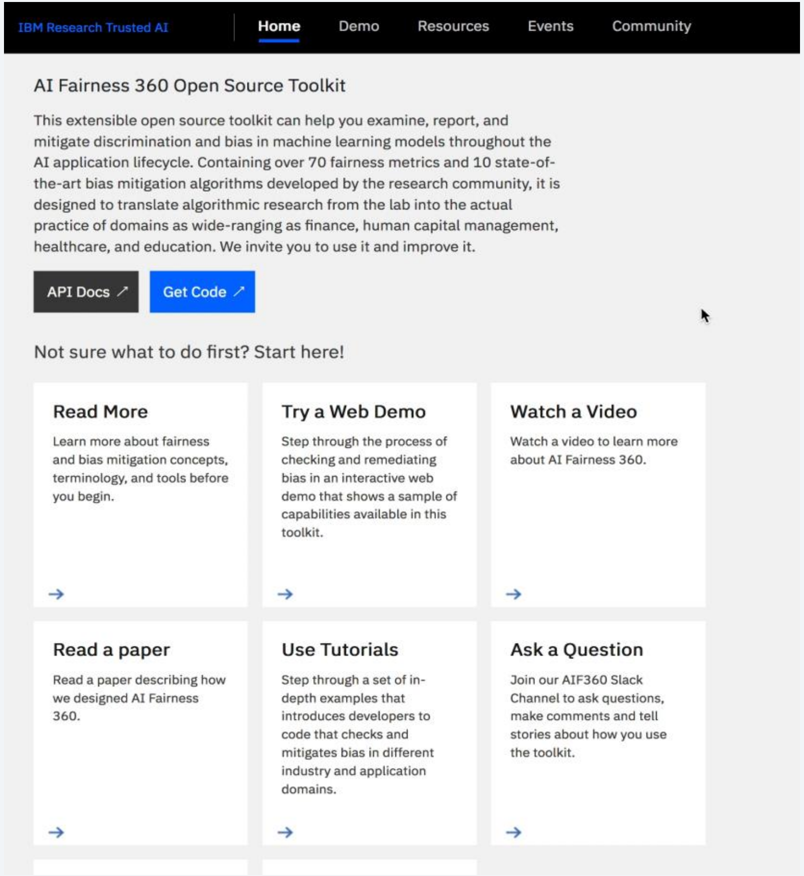
Differentiation

Comprehensive bias mitigation toolbox (including unique algorithms from IBM Research)

Several metrics and algorithms that have no available implementations elsewhere

Extensible, (e.g. scikit-learn's fit/predict paradigm)

Designed to translate new research from the lab to industry practitioners



AI Fairness 360 Open Source Toolkit

This extensible open source toolkit can help you examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. Containing over 70 fairness metrics and 10 state-of-the-art bias mitigation algorithms developed by the research community, it is designed to translate algorithmic research from the lab into the actual practice of domains as wide-ranging as finance, human capital management, healthcare, and education. We invite you to use it and improve it.

Not sure what to do first? Start here!

Read More

Learn more about fairness and bias mitigation concepts, terminology, and tools before you begin. →

Try a Web Demo

Step through the process of checking and remediating bias in an interactive web demo that shows a sample of capabilities available in this toolkit. →

Watch a Video

Watch a video to learn more about AI Fairness 360. →

Read a paper

Read a paper describing how we designed AI Fairness 360. →

Use Tutorials

Step through a set of in-depth examples that introduces developers to code that checks and mitigates bias in different industry and application domains. →

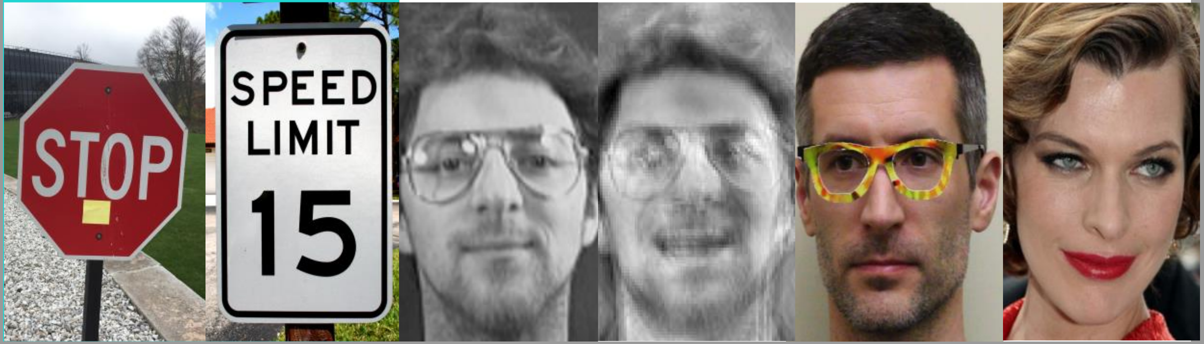
Ask a Question

Join our AIF360 Slack Channel to ask questions, make comments and tell stories about how you use the toolkit. →

Trusted AI: Making AI Secure

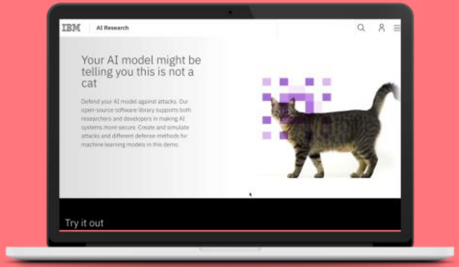
AI Models are vulnerable

Poison training data and corrupt models



Adversarial Robustness Toolkit

The most comprehensive toolkit for “attacking” and defending AI

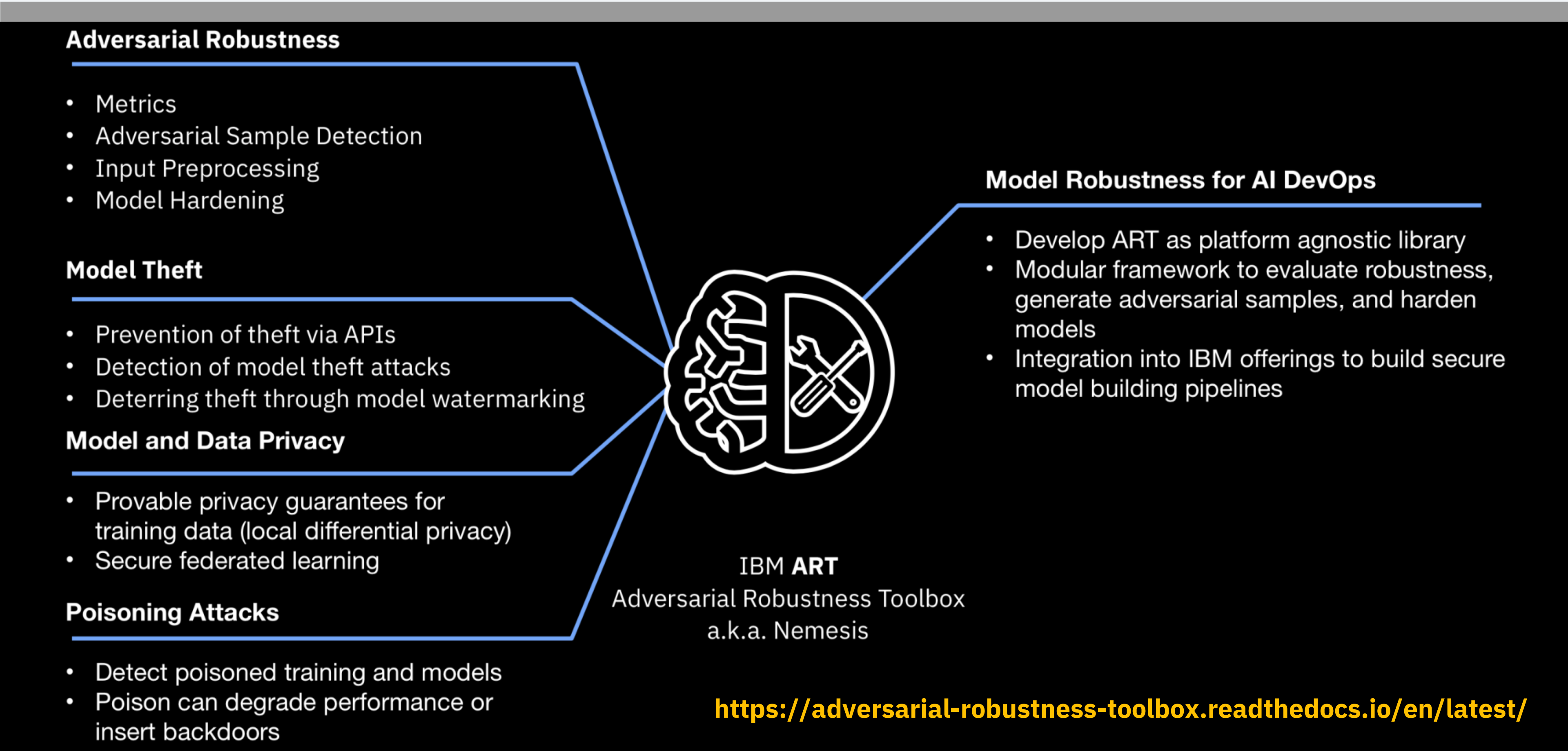


Your AI model might be telling you this is not a cat

Defend your AI model against attacks. Our open-source software library supports both researchers and developers in making AI systems more secure. Create and simulate attacks and different defense methods for machine learning models in this demo.

Try it out

IBM created ART, an open-source adversarial robustness toolkit



Adversarial Robustness

- Metrics
- Adversarial Sample Detection
- Input Preprocessing
- Model Hardening

Model Theft

- Prevention of theft via APIs
- Detection of model theft attacks
- Deterring theft through model watermarking

Model and Data Privacy

- Provable privacy guarantees for training data (local differential privacy)
- Secure federated learning

Poisoning Attacks

- Detect poisoned training and models
- Poison can degrade performance or insert backdoors

Model Robustness for AI DevOps

- Develop ART as platform agnostic library
- Modular framework to evaluate robustness, generate adversarial samples, and harden models
- Integration into IBM offerings to build secure model building pipelines

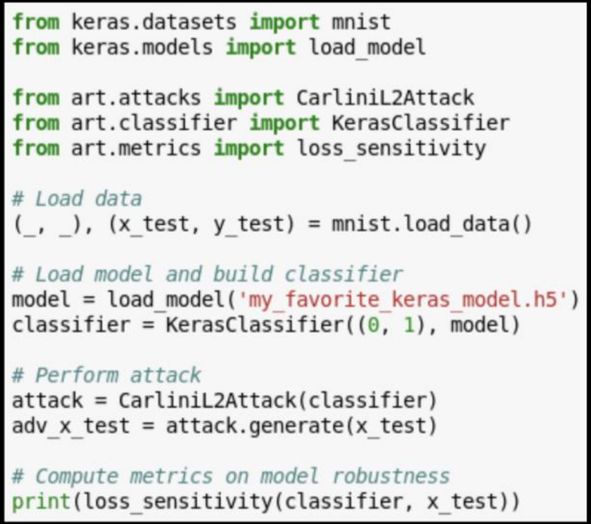
IBM **ART**
Adversarial Robustness Toolbox
a.k.a. Nemesis

https://adversarial-robustness-toolbox.readthedocs.io/en/latest/

Adversarial Robustness Toolbox (ART)

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

- Python library, 7K lines of code
- State-of-the-art attacks, defences and robustness metrics



Load ART modules →

Load classifier model (Keras, TF, PyTorch etc) →

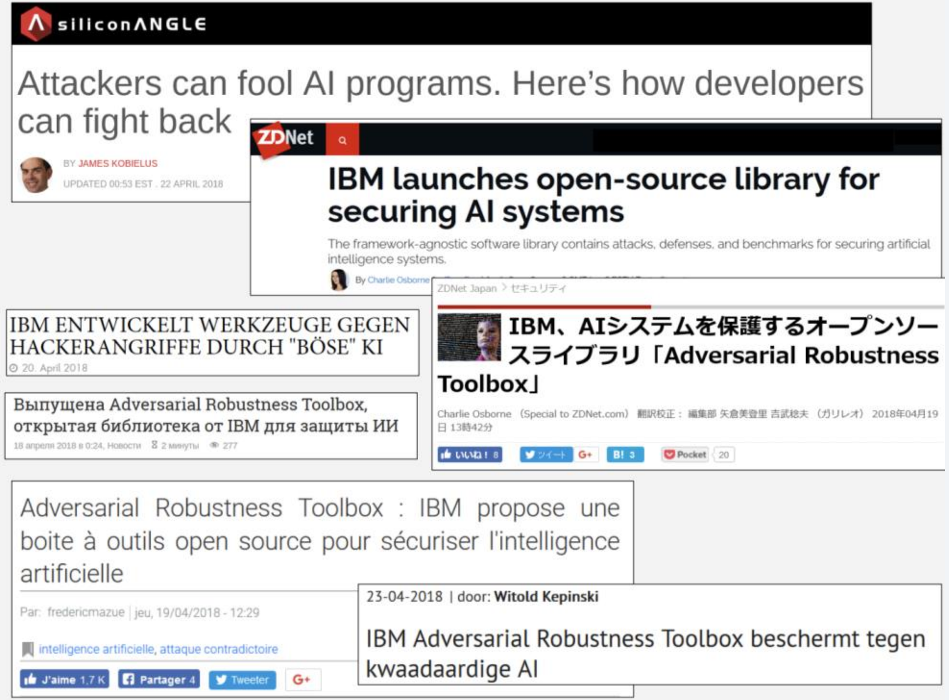
Perform attack →

Evaluate robustness →

```
from keras.datasets import mnist
from keras.models import load_model

from art.attacks import CarliniL2Attack
from art.classifier import KerasClassifier
from art.metrics import loss_sensitivity

# Load data
(_, _), (x_test, y_test) = mnist.load_data()

# Load model and build classifier
model = load_model('my_favorite_keras_model.h5')
classifier = KerasClassifier((0, 1), model)

# Perform attack
attack = CarliniL2Attack(classifier)
adv_x_test = attack.generate(x_test)

# Compute metrics on model robustness
print(loss_sensitivity(classifier, x_test))
```

Open-source release @ RSA 2018:

- ~ 3.5K+ views of IBM blogs
- ~ 100+ news outlets covering release
- ~ 1.3M+ Social Media potential impressions
- ~ 5K+ views of GitHub repo



siliconANGLE

Attackers can fool AI programs. Here's how developers can fight back

BY JAMES KOBIELUS

UPDATED 00:53 EST . 22 APRIL 2018

ZDNet

IBM launches open-source library for securing AI systems

The framework-agnostic software library contains attacks, defenses, and benchmarks for securing artificial intelligence systems.

By Charlie Osborne

IBM ENTWICKELT WERKZEUGE GEGEN HACKERANGRIFFE DURCH "BÖSE" KI

20. April 2018

ZDNet Japan > セキュリティ

IBM、AIシステムを保護するオープンソースライブラリ「Adversarial Robustness Toolbox」

Charlie Osborne （Special to ZDNet.com） 翻訳校正： 編集部 矢倉美登里 吉武稔夫 （ガリレオ） 2018年04月19日 13時42分

いいね！ 8 ツイート G+ B! 3 Pocket 20

Выпущена Adversarial Robustness Toolbox, открытая библиотека от IBM для защиты ИИ

18 апреля 2018 в 0:24, Новости 2 минуты 277

Adversarial Robustness Toolbox : IBM propose une boite à outils open source pour sécuriser l'intelligence artificielle

Par: fredericmazue | jeu, 19/04/2018 - 12:29

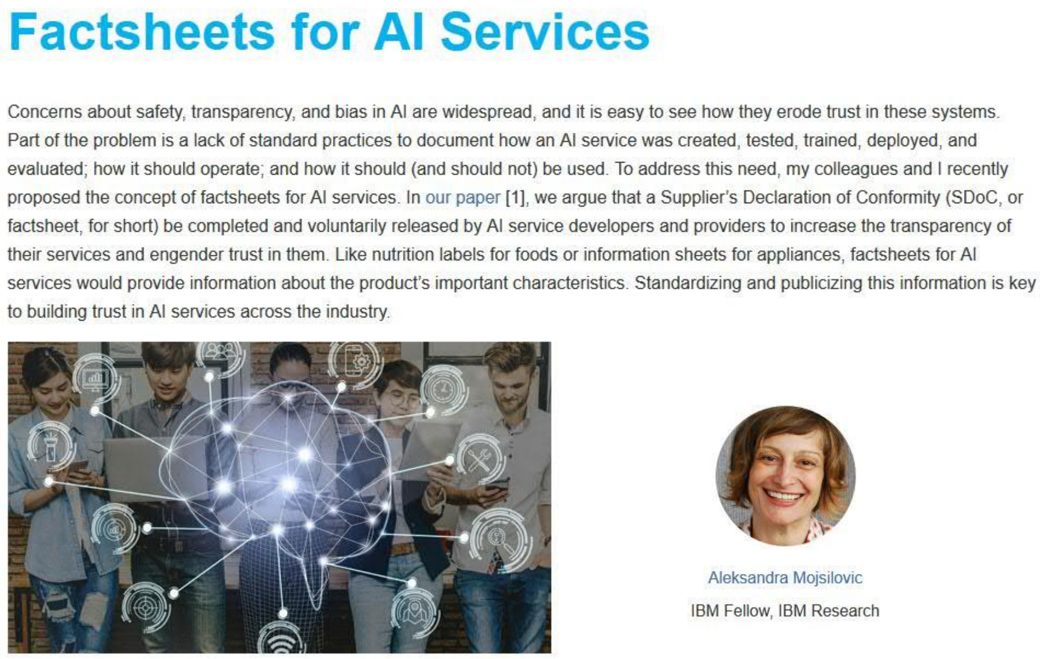
intelligence artificielle, attaque contradictoire

J'aime 1,7 K Partager 4 Tweeter G+

23-04-2018 | door: Witold Kepinski

IBM Adversarial Robustness Toolbox beschermt tegen kwaadaardige AI

Trust in AI Systems

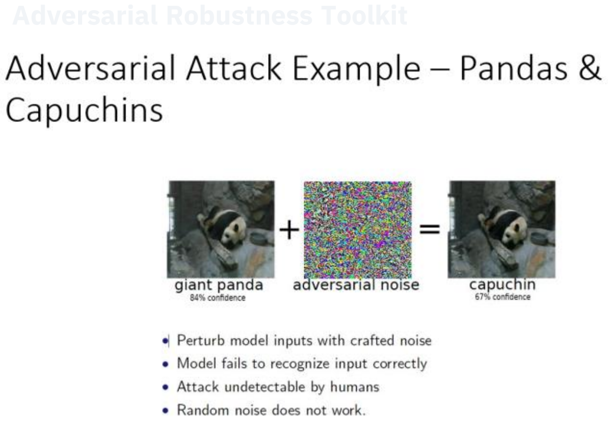


Factsheets for AI Services

Concerns about safety, transparency, and bias in AI are widespread, and it is easy to see how they erode trust in these systems. Part of the problem is a lack of standard practices to document how an AI service was created, tested, trained, deployed, and evaluated; how it should operate; and how it should (and should not) be used. To address this need, my colleagues and I recently proposed the concept of factsheets for AI services. In our paper [1], we argue that a Supplier's Declaration of Conformity (SDoC, or factsheet, for short) be completed and voluntarily released by AI service developers and providers to increase the transparency of their services and engender trust in them. Like nutrition labels for foods or information sheets for appliances, factsheets for AI services would provide information about the product's important characteristics. Standardizing and publicizing this information is key to building trust in AI services across the industry.

Aleksandra Mojsilovic

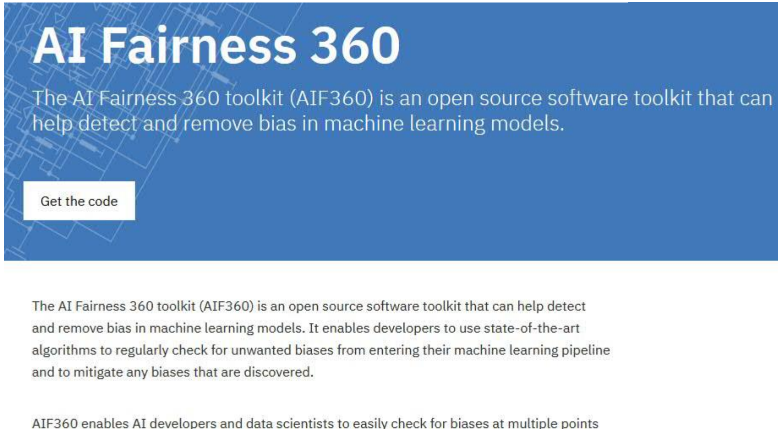
IBM Fellow, IBM Research



Adversarial Robustness Toolkit

Adversarial Attack Example – Pandas & Capuchins

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



AI Fairness 360

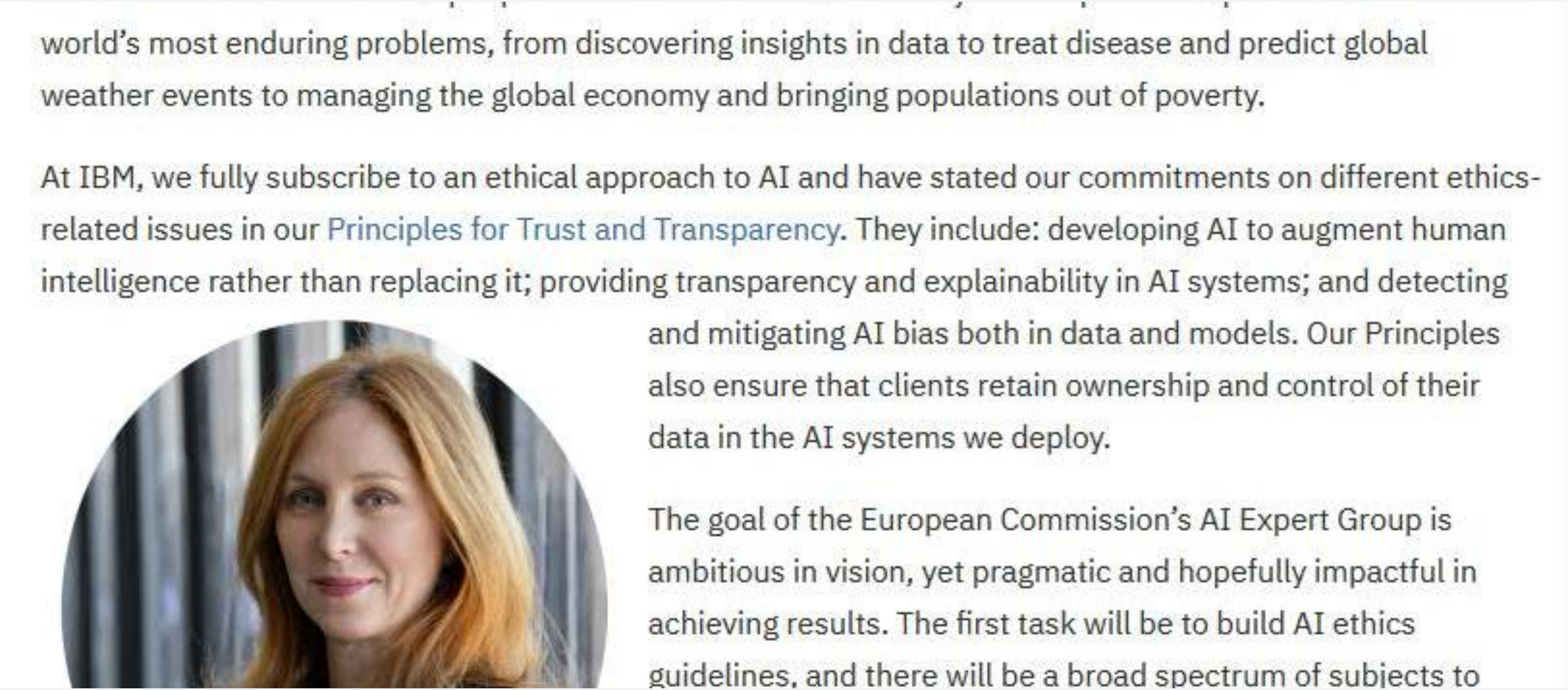
The AI Fairness 360 toolkit (AIF360) is an open source software toolkit that can help detect and remove bias in machine learning models.

The AI Fairness 360 toolkit (AIF360) is an open source software toolkit that can help detect and remove bias in machine learning models. It enables developers to use state-of-the-art algorithms to regularly check for unwanted biases from entering their machine learning pipeline and to mitigate any biases that are discovered.

AIF360 enables AI developers and data scientists to easily check for biases at multiple points

AI Ethics https://www.ibm.com/blogs/policy/francesca-rossi-ai/

https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf

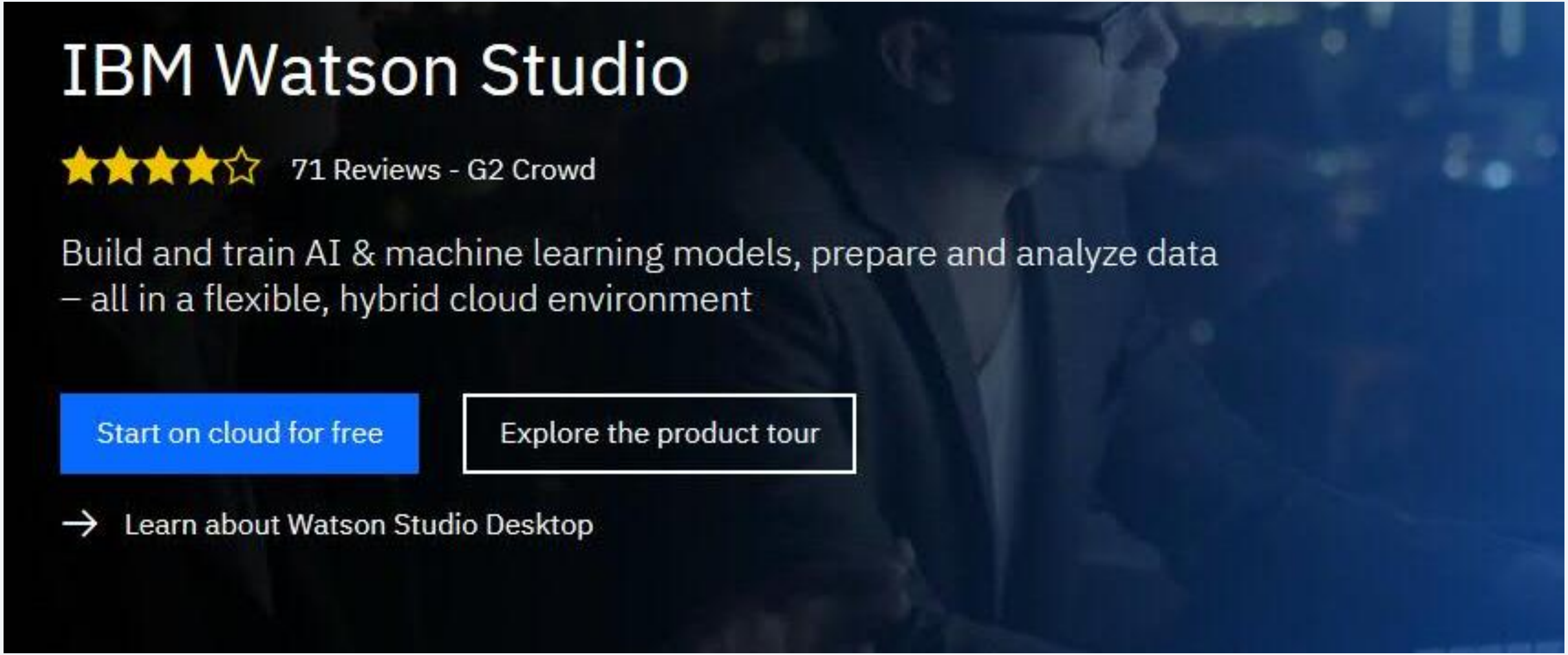


world’s most enduring problems, from discovering insights in data to treat disease and predict global weather events to managing the global economy and bringing populations out of poverty.

At IBM, we fully subscribe to an ethical approach to AI and have stated our commitments on different ethics-related issues in our Principles for Trust and Transparency. They include: developing AI to augment human intelligence rather than replacing it; providing transparency and explainability in AI systems; and detecting and mitigating AI bias both in data and models. Our Principles also ensure that clients retain ownership and control of their data in the AI systems we deploy.

The goal of the European Commission’s AI Expert Group is ambitious in vision, yet pragmatic and hopefully impactful in achieving results. The first task will be to build AI ethics guidelines, and there will be a broad spectrum of subjects to

Watson Studio



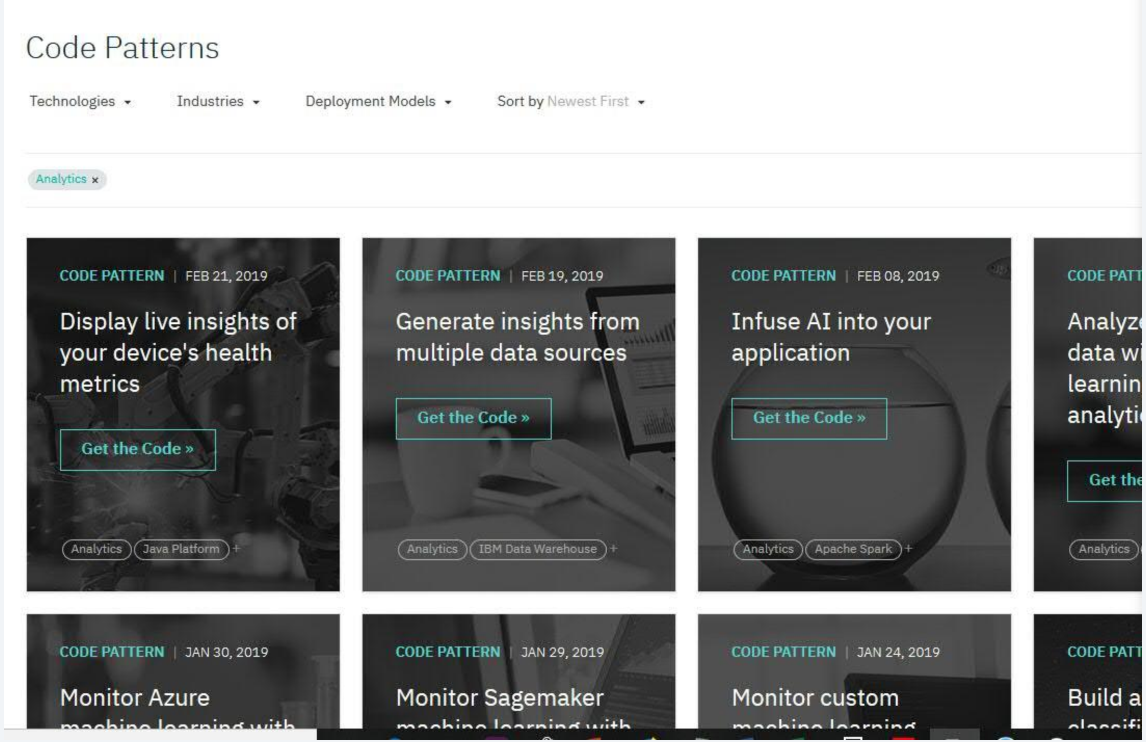
IBM Watson Studio

★★★★☆ 71 Reviews - G2 Crowd

Build and train AI & machine learning models, prepare and analyze data – all in a flexible, hybrid cloud environment

Code Patterns

https://developer.ibm.com/patterns/



Code Patterns

CODE PATTERN | FEB 21, 2019
Display live insights of your device's health metrics

CODE PATTERN | FEB 19, 2019
Generate insights from multiple data sources

CODE PATTERN | FEB 08, 2019
Infuse AI into your application

CODE PATT
Analyz
data wi
learnin
analyti

CODE PATTERN | JAN 30, 2019
Monitor Azure machine learning with

CODE PATTERN | JAN 29, 2019
Monitor Sagemaker machine learning with

CODE PATTERN | JAN 24, 2019
Monitor custom machine learning

CODE PATT
Build a
classifi

IBM AI Learning & Certification: AI Literacy for ALL

Sharing our deep AI knowledge & experience from working with hundreds of enterprise clients.

Now Live:

Online learning with our AI Learning Catalog

Custom, in-lab AI Learning Experiences

Coming Soon:

IBM AI Certification for the end-to-end enterprise AI workflow.

Invest in your future. Start your AI Learning now.

http://community.ibm.com/aiskills

Get involved

Call for Code 2019

How can you participate?

Developers register for the challenge, get started building applications that will save lives.
www.developer.ibm.com/callforcode

Support Call for Code:

- Host a day for your organization
- Provide promotional support for the initiative
- Donate in-kind: charitable donations, offer a VC pitch to the winning team or donate your technology

https://callforcode.org/become-a-supporter/

Sponsor, show your full support with a sponsorship.
https://callforcode.org/become-a-sponsor/

Visit www.developer.ibm.com/callforcode

Backup

Data & AI : What's Happening

Talk Summary : This session will provide a brief overview of what's happening with AI with a particular emphasis on data - and then provide a summary of the IBM offerings and products that support AI and Data Science

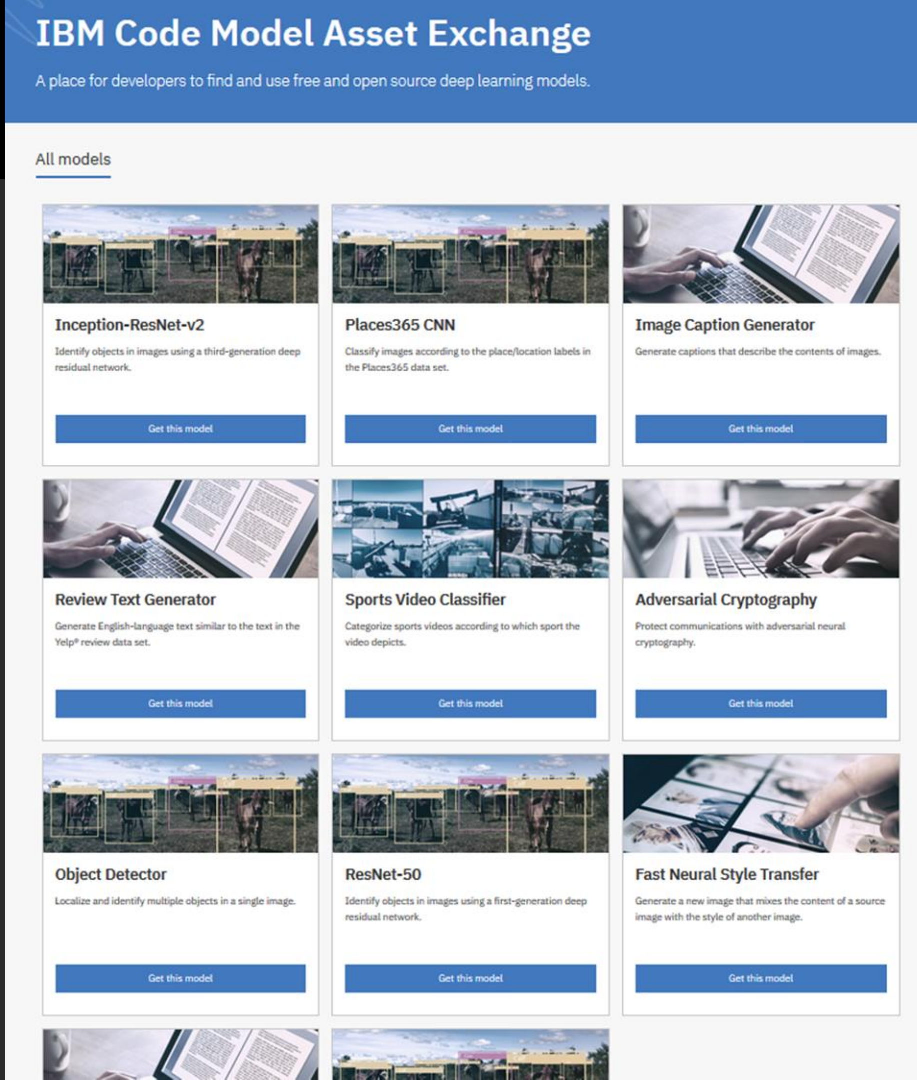
Bio : Susan Malaika is Senior Technical Staff in the Cognitive Applications group in IBM focusing on open source for Data & AI, Susan also leads a tech community of a few hundred volunteers in the New York area & she loves hackathons.

For more information about Susan please see https://developer.ibm.com/opentech/category/susan-malaika/

MAX - Model Asset Exchange

- MAX is a one-stop exchange for data scientists and AI developers to consume models created using their favorite machine learning engines like TensorFlow, PyTorch, and Caffe2, and provides a standardized approach to classify, annotate, and deploy these models for prediction and inferencing.

- Visit the Model Asset Exchange at: https://developer.ibm.com/code/exchanges/models/



IBM Code Model Asset Exchange

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

Inception-ResNet-v2

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

Places365 CNN

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

Image Caption Generator

Generate captions that describe the contents of images.

Review Text Generator

Generate English-language text similar to the text in the Yelp® review data set.

Sports Video Classifier

Categorize sports videos according to which sport the video depicts.

Adversarial Cryptography

Protect communications with adversarial neural cryptography.

Object Detector

Localize and identify multiple objects in a single image.

ResNet-50

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

Fast Neural Style Transfer

Generate a new image that mixes the content of a source image with the style of another image.

AI is Not Magic: It’s Time to Demystify and Apply

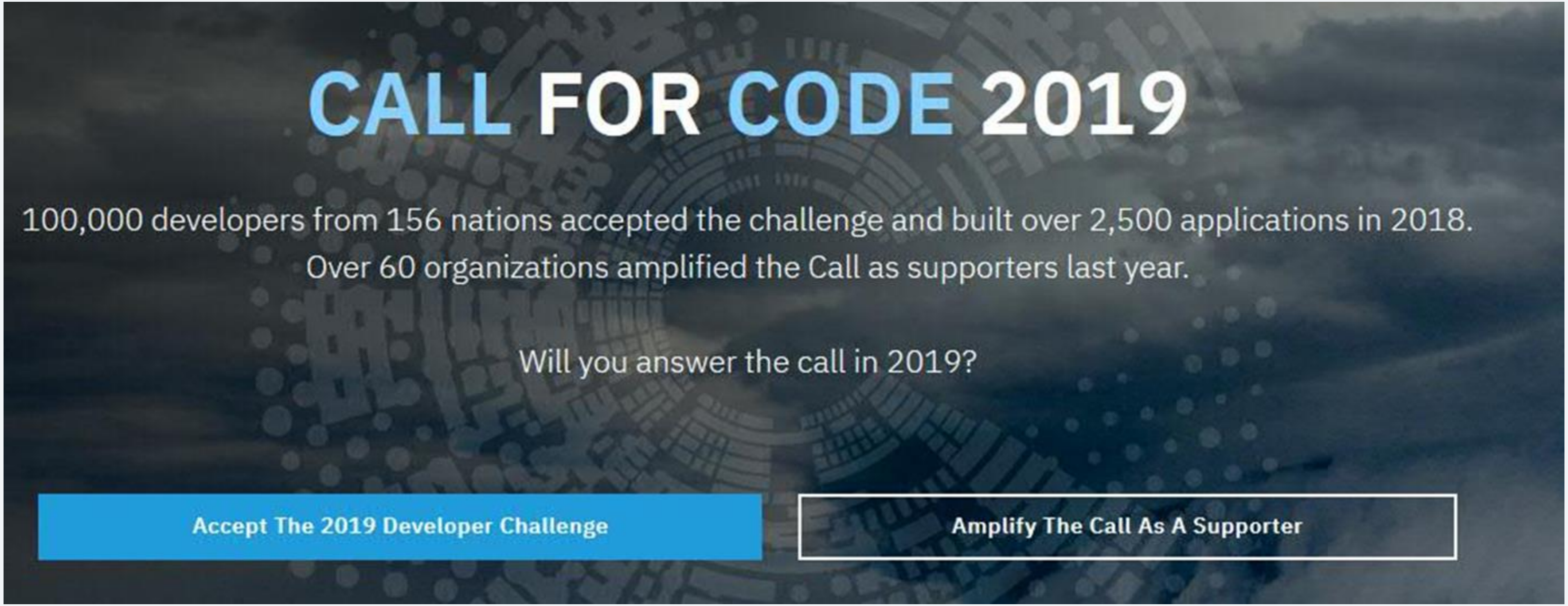
A unified, modern data fabric.

A development environment and engine.

Human features.

AI management and exploitation.

Participate : Registrations Opened 2019-03-25 - https://callforcode.org/



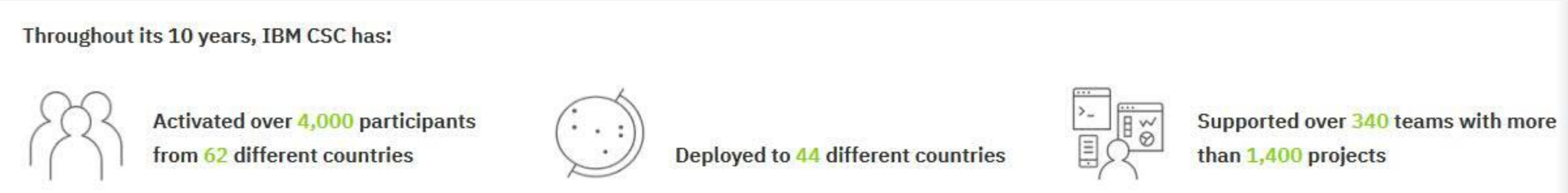
CALL FOR CODE 2019

100,000 developers from 156 nations accepted the challenge and built over 2,500 applications in 2018.
Over 60 organizations amplified the Call as supporters last year.

Will you answer the call in 2019?

Accept The 2019 Developer Challenge
Amplify The Call As A Supporter

IBM Corporate Service Corps

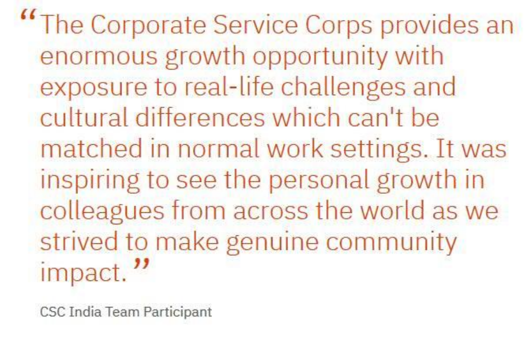


Throughout its 10 years, IBM CSC has:

Activated over 4,000 participants from 62 different countries

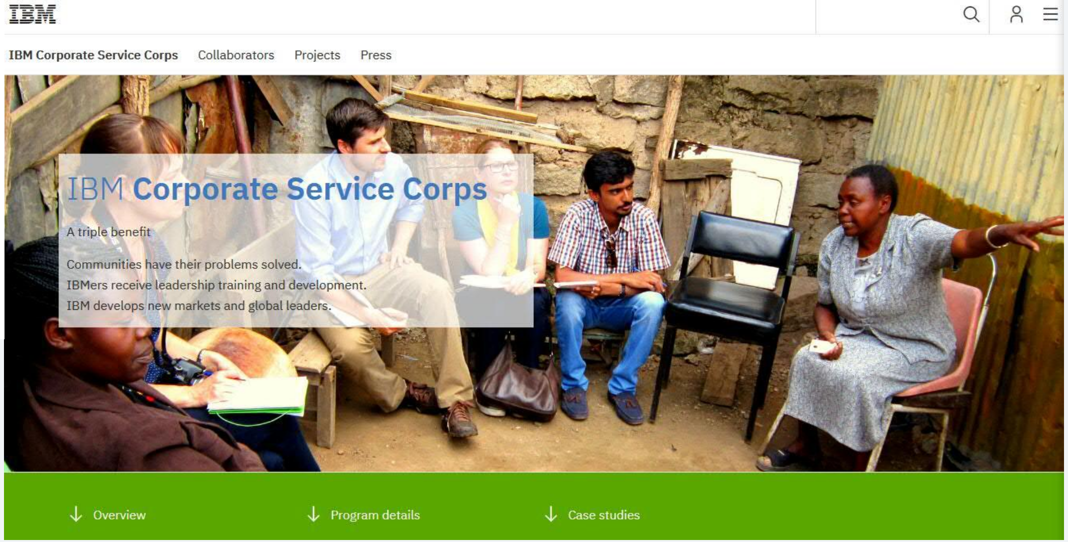
Deployed to 44 different countries

Supported over 340 teams with more than 1,400 projects



> “The Corporate Service Corps provides an enormous growth opportunity with exposure to real-life challenges and cultural differences which can't be matched in normal work settings. It was inspiring to see the personal growth in colleagues from across the world as we strived to make genuine community impact.”

CSC India Team Participant



IBM **Corporate Service Corps**

A triple benefit

Communities have their problems solved.
IBMers receive leadership training and development.
IBM develops new markets and global leaders.

Natural disasters are among the world's greatest challenges…

800,000+

worldwide deaths attributed to earthquakes since 2010

25%

coastline areas that met or surpassed record number of flood days

800+

confirmed tornadoes touched down in 2018

17 million

acres lost to wildfire in the United States in the last 2 years

22

named storms in the Eastern Pacific region this year – a record

18

volcanos considered a "very high threat" in the U.S. alone

Agenda

- Data
- ML – Machine Learning,
- DL – Deep Learning
- AI Trust
- Join the Call for Code

Code Response

Code and Response is an IBM initiative which provides a platform to create and deploy open source technologies to tackle some of the world's biggest challenges.

Coding challenges includes Call for Code, CGIU student codeathons

Solution deployment starting with Call for Code 2018 winner Project OWL

Volunteer in disaster relief efforts with the American Red Cross & more

Code and Response™ is supported by NGOs, governments, global technologists, as well as the IBM Corporate Service Corps.

www.developer.ibm.com/code-and-response

CALL FOR CODE®

100k Developers

156 Nations

2,500+ Applications

Part of Code and Response™, this annual global developer challenge is a great way to get involved. It inspires developers to create sustainable software solutions to prepare for, respond to and recover from natural disasters. www.developer.ibm.com/callforcode

The winning team receives:

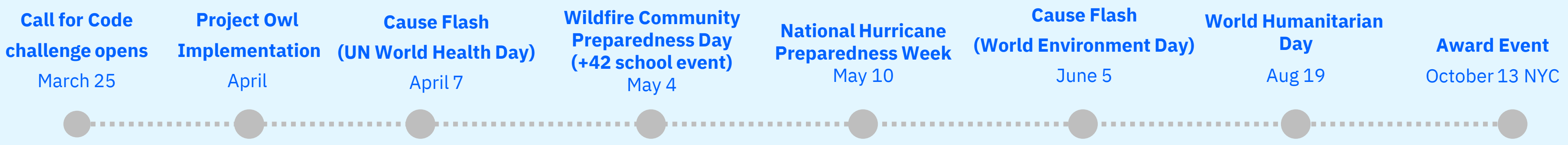
- A **$200K** cash prize
- Open Source Support from The Linux Foundation
- Meetings with mentors and potential investors
- Solution implementation through Code and Response™

Get involved

Support Call for Code, and host a day

Become an affiliate, donate in-kind

Sponsor, show your full support with a sponsorship



Call for Code challenge opens
March 25

Project Owl Implementation
April

Cause Flash (UN World Health Day)
April 7

Wildfire Community Preparedness Day (+42 school event)
May 4

National Hurricane Preparedness Week
May 10

Cause Flash (World Environment Day)
June 5

World Humanitarian Day
Aug 19

Award Event
October 13 NYC

The IBM open source way
https://developer.ibm.com/open/culture/

Training

Open Source @ IBM Program touches

78,000

IBMers annually

Recognition

We recognize our open source leaders with

300+

cash awards annually

Tooling

Our open source management tool suite is used over

30,000+

times per month

Organization

Our Open Source Core Team includes

~12 FTEs

supporting all of IBM

Consuming

Virtually all of our products contain open source

3000+

packages reviewed every month

Jeffrey Borek

Great turn out for #OpenSource @IBM Lab Tour event today. w Jochen talking about when #Standards & OS collide! 😃 IBM SW Labs, Germany #Cloud

Michael Fork

Excited to talk to a full house about The Next Chapter of Open Source at IBM with @Kevin_Nowka, @timmoore_1, and @bradtopol

Contributing

We invest in community code & innovation

1500+

GitHub repos

Datasheets Proposal

- The machine learning community has no standardized way to document how and why a dataset was created, what information it contains, what tasks it should and should not be used for, and whether it might raise any ethical or legal concerns. To address this gap, we propose the concept of datasheets for datasets.
- In the electronics industry, it is standard to accompany every component with a datasheet providing standard operating characteristics, test results, recommended usage, and other information. Similarly, we recommend that every dataset be accompanied with a datasheet documenting its creation, composition,intended uses, maintenance, and other properties.
- Datasheets for datasets will facilitate better communication between dataset creators and users, and encourage the machine learning communityto prioritize transparency and accountability.

Sample questions:

- Why was the dataset created? (e.g., was there a specific intended task gap that needed to be filled?)
- Who funded the creation of the dataset?
- What preprocessing/cleaning was done? (e.g., discretization or bucketing, tokenization, part-of-speechtagging, SIFT feature extraction, removal of instances)
- If it relates to people, were they told what the dataset would be used for and did they consent?If so, how? Were they provided with any mechanism to revoke their consent in the future or for certain uses?
- Will the dataset be updated? How often, by whom?

Datasheets for Datasets https://arxiv.org/pdf/1803.09010.pdf

Summary

This talk reviews the challenges and metrics for enterprise workloads, the benchmark tests that are available, and the gaps which need to be filled.

The paper, that this talk is based on, identifies the following areas as important to enterprises concerned about performance:

- **1. Model training performance**
  - data labeling / preparation
  - time-to-accuracy
  - computational time / cycles
  - throughput-to-accuracy
- **2. Hyper-parameter optimization performance**
- **3. Inference runtime performance**

The talk offers a summary table of the main three AI areas important to enterprises, alongside:

- Workload profile
- Important performance indicators to assess the task's efficiency
- Potential technical bottlenecks to look out for that could limit the AI tasks performance delivered by a given solution.