

# Machine Learning and AI at IBM

Lennart Frantzell PhD, Developer Advocate, IBM

4/24/2018

<http://ibm.biz/mlatibm>

# IBM and AI

- **1996, 1997. Deep Blue** The first match was played in Philadelphia in 1996 and won by Kasparov. The second was played in New York City in 1997 and won by **Deep Blue**.
- **Watson Jeopardy 2011:** The IBM Challenge aired February 14–16, 2011, and **featured IBM's Watson computer** facing off against Ken Jennings and Brad Rutter in a two-game match played over three shows. This was the first man-vs.-machine competition in Jeopardy!'s history. Watson won both the first game and the overall match

Watson uses IBM's DeepQA software and the [Apache UIMA](#), Unstructured Information Management Architecture, is an OASIS standard for content analytics.

# IBM and AI

- **1996, 1997. Deep Blue** The first match between a computer and a human in 1996 and won by Kasparov. The second match in 1997 won by **Deep Blue**.
- **Watson Jeopardy!** The world's best player of what might be humankind's most complicated board game was defeated on Tuesday by a Google computer program. Adding insult to potentially deep existential injury, he was defeated at Go — a game that claims centuries of play by humans — in China, where the game was invented.

Watson is a DeepQA software and the [Apache UIMA](#), Unstructured Information Management Architecture, is an OASIS standard for content analytics,

# From an analog to a digital world



**ASCII**





# Watson for Healthcare

## IBM Watson for Genomics helps doctors give patients new hope.

Now clinicians across the U.S. can provide precision medicine to cancer patients. See how Watson for Genomics helps enhance doctors' confidence in personalized treatment approaches.

<https://www.ibm.com/watson/health/oncology-and-genomics/genomics/>

### IBM Watson for Drug Discovery

Help researchers identify novel drug targets and new indications for existing drugs. The platform can help researchers uncover new connections and develop new treatments ahead of the competition.

[Learn more about Drug Discovery](#)

### Social Program Management

Supporting government agencies in their work to deliver health and human services that enable citizens to meet their maximum potential while protecting the most vulnerable populations.

[Learn more about Social Program Management](#)

### IBM Watson for Oncology

Spend less time searching literature and more time caring for patients. Watson can provide clinicians with evidence-based treatment options based on expert training by Memorial Sloan Kettering (MSK) physicians.

[Watson for Oncology](#)

### IBM Watson Care Manager

Use personalized care plans, automated care management workflows, and integrated patient engagement capabilities to help create more informed action plans.

<https://www.ibm.com/watson/health/>

# Data in Healthcare

- With IBM's planned [\\$2.6 billion acquisition of Truven Health](#), the company will add "200 million lives" to its data trove. "Lives" is a term typically used in the healthcare business for a data asset or record.
- And when it comes to big data analytics, the more data, the better, said IBM [\(ibm, -0.84%\)](#) Watson Health general manager Deborah DiSanzo. Truven brings still more data into IBM, which has already assembled quite the data pool, both on its own and via acquisition.

<http://fortune.com/2016/02/18/ibm-truven-health-acquisition/>

# The GPU and Machine Learning

## GPU cloud servers

GPUs harness the processing power for high performance computing, deep learning, machine learning, AI and virtual desktop infrastructure (VDI).

...The answer lies in the rise of deep learning, an advanced machine learning technique that is heavily used in AI and Cognitive Computing. Deep learning powers many scenarios including autonomous cars, cancer diagnosis, computer vision, speech recognition and many other intelligent use cases.

**Like most of the ML algorithms, deep learning relies on sophisticated mathematical and statistical computations. Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are some of the modern implementations of deep learning. .... Each type of neural net is aligned with a complex use case like classification, clustering and prediction.**

<https://www.forbes.com/sites/janakirammsv/2017/08/07/in-the-era-of-artificial-intelligence-gpus-are-the-new-cpus/#16d91afc5d16>



# Big Data, Hardware and the GPU



**Cloud TPUs**



# AI, Machine Learning and hardware

## **IBM Bets on Artificial Intelligence With Its POWER9 Chip**

The company is hoping to carve out a presence in the hyperscale data center market with its AI-optimized processor.

# **Machine Learning, the Software story**

# Harvard Business Review

HBK.ORG

OCTOBER 2012

46 **The Big Idea**  
The True Measures Of Success  
Michael J. Mauboussin

84 **International Business**  
10 Rules for Managing Global Innovation  
Kasley Wilson and Yves L. Doz

93 **Leadership**  
What Ever Happened To Accountability?  
Thomas E. Ricks

GETTING CONTROL OF **BIG DATA**

How vast new streams of information are changing the art of management  
**PAGE 59**



# The new hero of big data and analytics: The Chief Data Officer

## The CDO...

...needs a T-shaped skills set that combines business acumen with data or analytics skills



...is a business executive—not a technician, programmer or data scientist



...has technical knowledge of data and its potential as an asset



3 key steps help drive long-term value for the CDO role



Drive a C-suite mandate to take actions underpinned by data



Accelerate an enterprise-wide data strategy to enable business transformation



Reward innovation that drives corporate performance

## 5 areas in which CDOs turn data into dollars



Up to 60 percent

of CDOs surveyed could create "low-hanging fruit" successes from data that already existed within their organizations.

Learn more about why leading organizations across the globe are appointing Chief Data Officers to deliver data-driven growth and innovation that matters.

<http://ibm.com/cdo>

Reprint: The new hero of big data and analytics: The Chief Data Officer. IBM Institute for Business Value

IBM

Just in case you didn't know

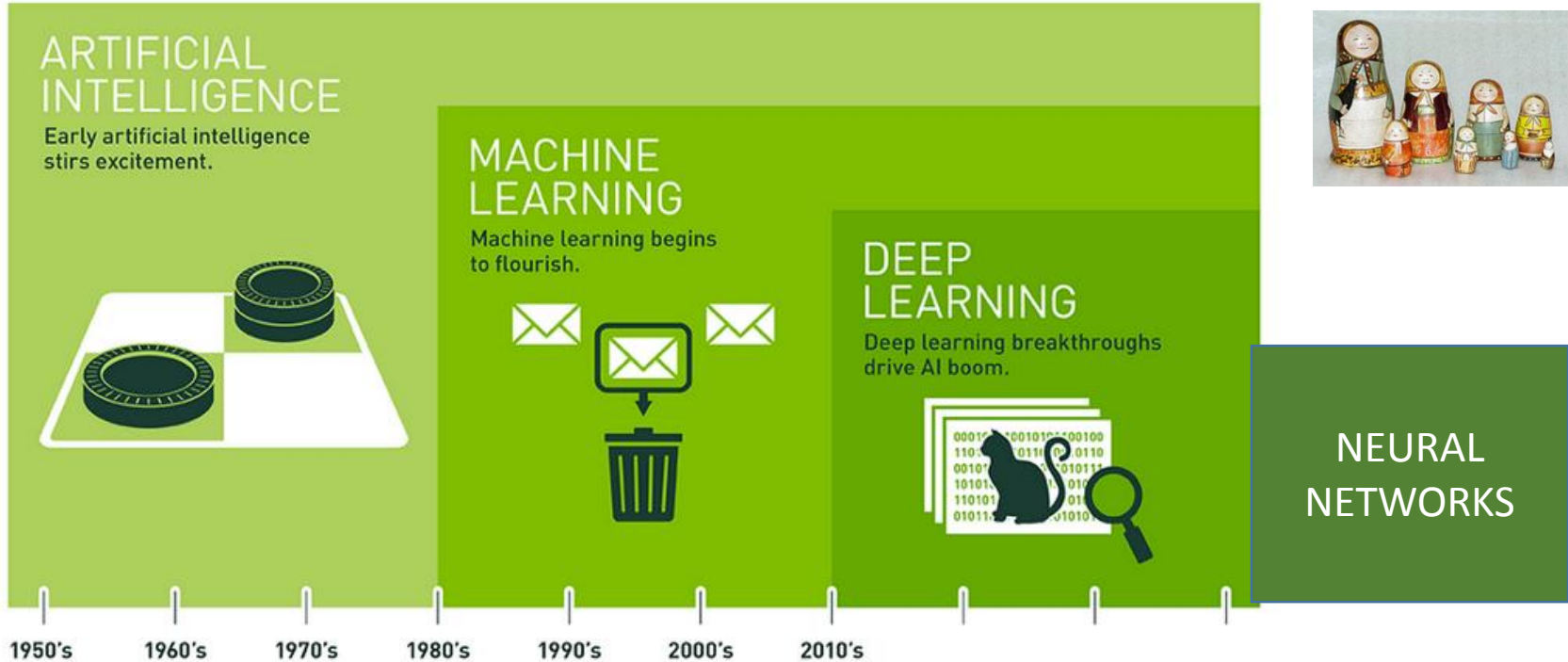




**But, ~ 80% of all ML projects are taken up with finding, transferring, washing and structuring the data.**



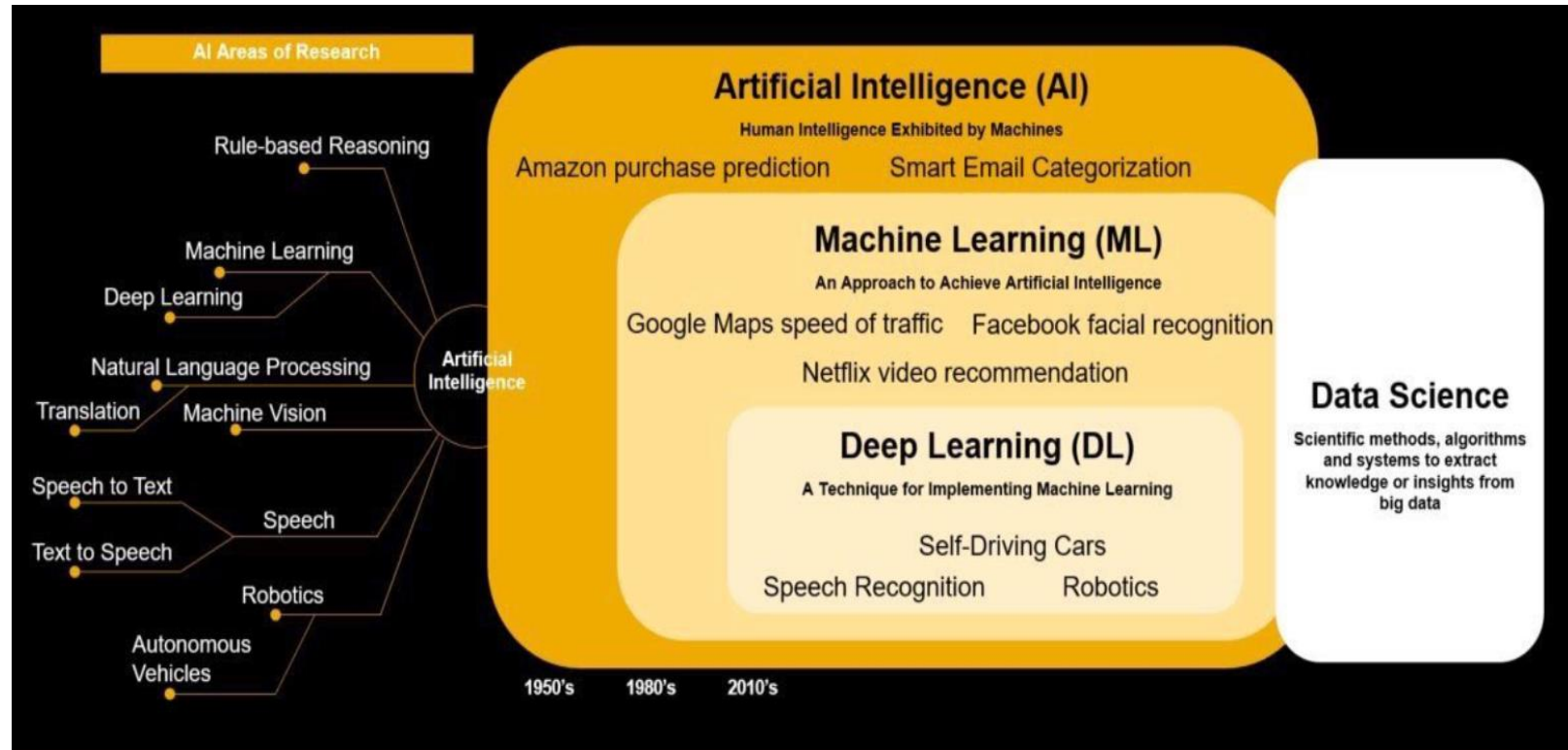
# Terminology: From AI to ML to DL



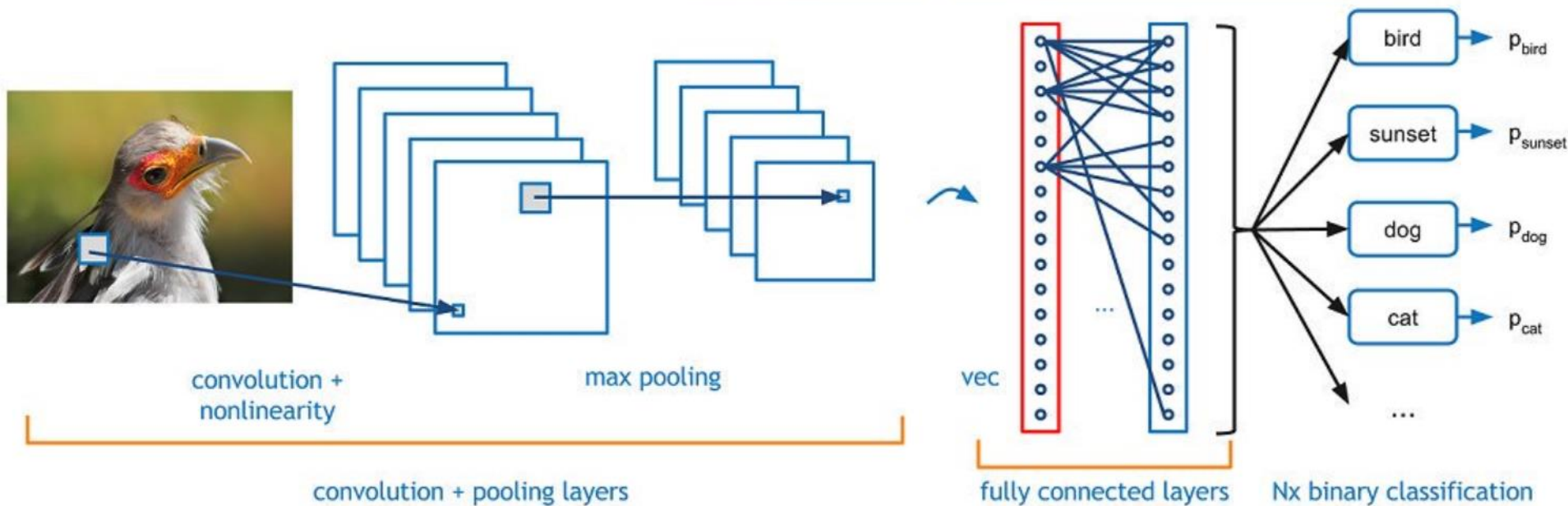
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.



# AI, Machine Learning, Deep Learning



# Convolutional Neural Networks



- <https://dzone.com/articles/a-beginners-guide-to-understanding-convolutional-n>

# CheXNet: Deep Learning in Radiology

Our algorithm, CheXNet, is a 121-layer **Convolutional neural network** trained on ChestX-ray14, currently the largest publicly available chest X-ray dataset, containing over 100,000 Frontal view X-ray images with 14 diseases.

<https://stanfordmlgroup.github.io/projects/chexnet/>

Stanford ML Group

## CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning

Pranav Rajpurkar\*, Jeremy Irvin\*, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. Ng

1. **Four practicing academic radiologists annotate a test set, on which we compare the performance of CheXNet to that of radiologists.**
2. **We find that CheXNet exceeds average radiologist performance** on the F1 metric. (In statistical analysis of binary classification, the F1 score (also F-score or F-measure) is a measure of a test's accuracy.)

# CheXNet: Deep Learning in Radiology

Our algorithm, CheXNet, is a 121-layer **Convolutional neural network** trained on ChestX-ray14, currently the largest publicly available chest X-ray dataset containing over 112,120 images. Frontal view only. 14 disease categories.

Test Set with 420 images



1. **Four practicing academic radiologists annotate a test set, on which we compare the performance of CheXNet to that of radiologists.**
2. **We find that CheXNet exceeds average radiologist performance** on the F1 metric. (In statistical analysis of binary classification, the F1 score (also F-score or F-measure) is a measure of a test's accuracy.)

# But wait, is it the whole story?

Medium



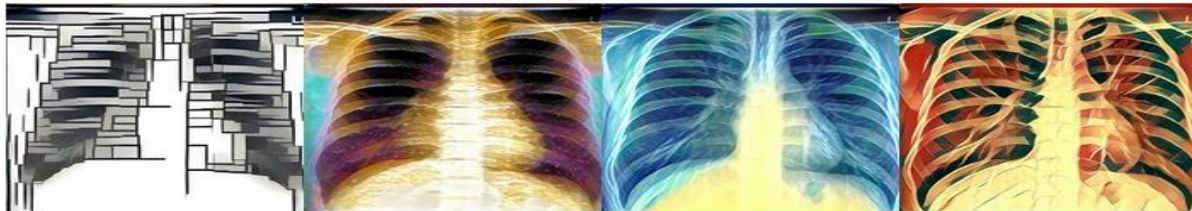
Bálint Botz

Follow

MD/PhD, Diagnostic radiology resident, Researcher mainly focusing on neuroinflammation

Nov 20, 2017 · 4 min read

## A Few Thoughts About CheXNet—And The Way Human Performance Should (And Should Not) Be Measured



<https://medium.com/@BalintBotz/a-few-thoughts-about-chexnet-and-the-way-human-performance-should-and-should-not-be-measured-68031dca7bf>

<https://lukeoakdenrayner.wordpress.com/2018/01/24/chexnet-an-in-depth-review/>

# But wait, is it the whole story?

Medium



Bálint Botz

Follow

MD/PhD, Diagnostic radiology resident, Researcher mainly focusing on neuroinflammation

Nov 20, 2017 · 4 min read

## A Few Thoughts About ChexNet Way Human Performance Should Not

You have to understand ML, DL and the domain you apply it to



<https://medium.com/@BalintBotz/a-few-thoughts-about-chexnet-and-the-way-human-performance-should-and-should-not-be-measured-68031dca7bf>

<https://lukeoakdenrayner.wordpress.com/2018/01/24/chexnet-an-in-depth-review/>

# The work horses: Deep (Machine) Learning Frameworks

Caffe theano



Keras

- [http://cs231n.stanford.edu/slides/2017/cs231n\\_2017\\_lecture8.pdf](http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture8.pdf)



# The Open Source Caffe Deep Learning Framework

## Caffe

- **Developed at Berkeley**
- **Expressive architecture** encourages application and innovation. Models and optimization are defined by configuration without hard-coding. Switch between CPU and GPU by setting a single flag to train on a GPU machine then deploy to commodity clusters or mobile devices.
- **Extensible code** fosters active development. **In Caffe's first year, it has been forked by over 1,000 developers and had many significant changes contributed back.** Thanks to these contributors the framework tracks the state-of-the-art in both code and models.
- **Speed** makes Caffe perfect for research experiments and industry deployment. **Caffe can process over 60M images per day with a single NVIDIA K40 GPU\*.** That's 1 ms/image for inference and 4 ms/image for learning and more recent library versions and hardware are faster still. We believe that Caffe is among the fastest convnet implementations available.

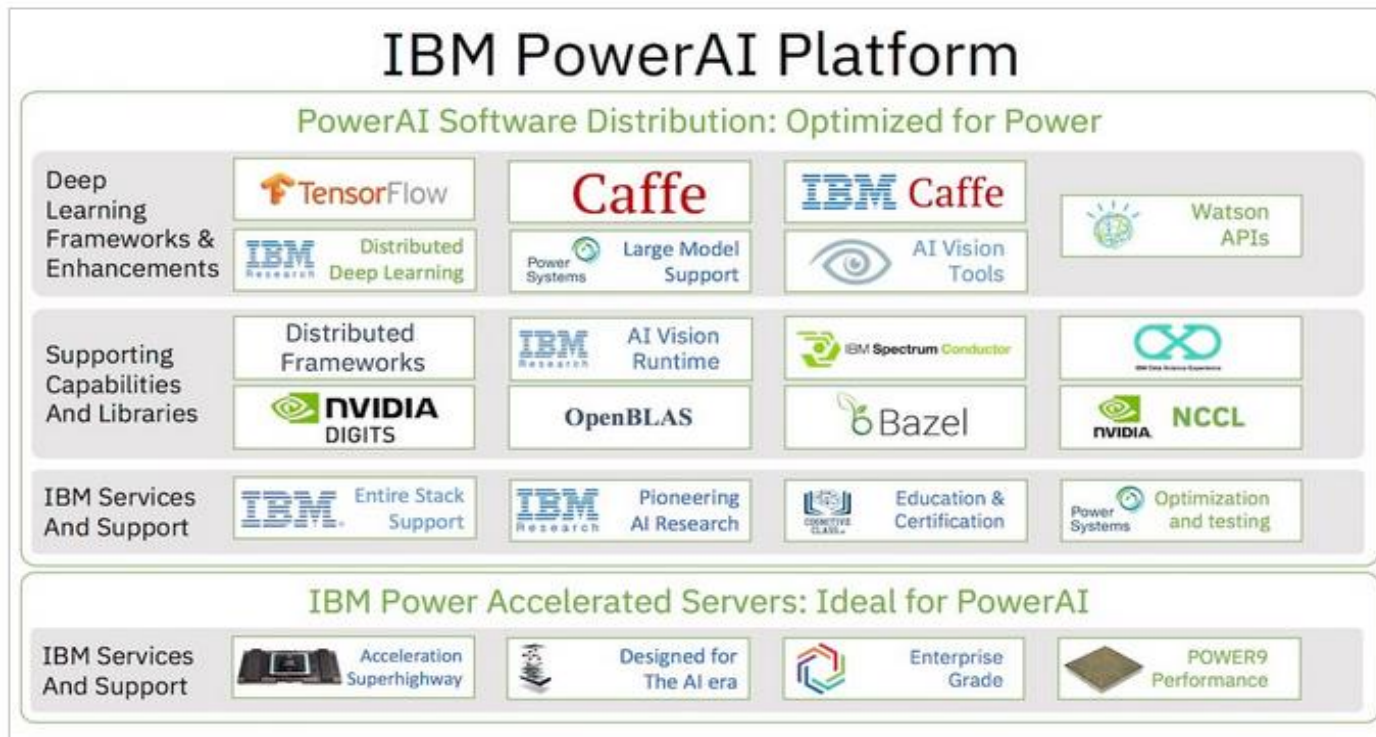


# Note on Keras



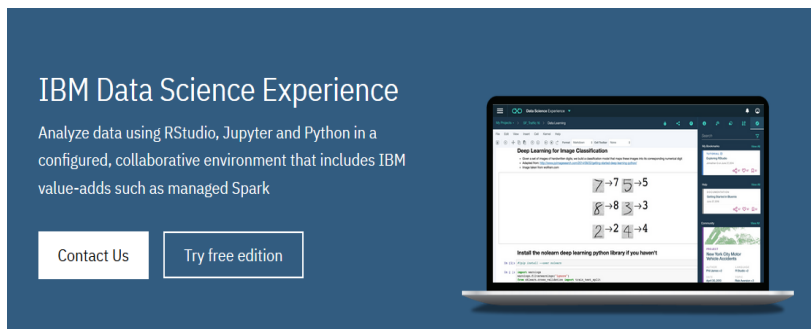
Keras is a high-level neural networks API, written in Python and capable of running on top of [TensorFlow](#), [CNTK](#), or [Theano](#). It was developed with a focus on enabling fast experimentation.

# IBM PowerAI Platform



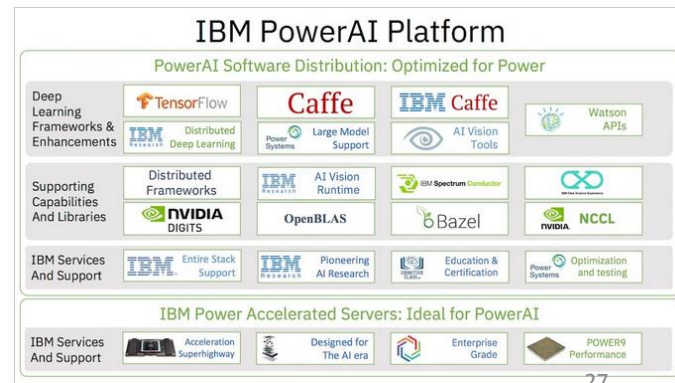
<https://www.ibm.com/us-en/marketplace/deep-learning-platform>

# Watson Studio and PowerAI



The Data Science Experience is an interactive and collaborate cloud-based environment designed to be **a place where data scientists can use such tools as RStudio, Jupyter, Python, Scala, Spark and IBM's Watson Machine Learning technology** to drive insights into their data and derive information useful to their businesses. It was rolled out last year first for the public cloud, and later was optimized for private clouds.

IBM is bringing the two together by integrating the PowerAI deep learning enterprise software distribution into the Data Science Experience.



# Watson in the IBM Public Cloud

# Watson Studio in the IBM Cloud

The screenshot shows the IBM Cloud console interface for the Watson Studio service. At the top is a dark navigation bar with the IBM Cloud logo and links for Catalog, Docs, Support, and Manage. Below the navigation bar, the page title 'Watson Studio' is displayed. The main content area is divided into several sections. On the left, there is a description of Watson Studio, a 'View all' link, and a metadata table. The description states that Watson Studio democratizes machine learning and deep learning to accelerate the infusion of AI in business. The metadata table lists the author as IBM, the publication date as 03/20/2018, the type as a service, and the location as United Kingdom, US South. In the center, there are configuration options for the service name (Watson Studio-gm), region (US South), organization (alf@us.ibm.com), and space (dev). Below these options is a 'Features' section with four bullet points: 'Use what you know, learn what you don't', 'Power on demand', 'Be a founding member', and 'Collaborate for better outcomes'. Each bullet point has a brief description of the feature. At the bottom of the page, there are links for 'Need Help?' and 'Contact IBM Cloud Sales', an 'Estimate Monthly Cost' link, a 'Cost Calculator' link, and a prominent blue 'Create' button.

IBM Cloud

Catalog Docs Support Manage

View all

## Watson Studio

Watson Studio democratizes machine learning and deep learning to accelerate infusion of AI in your business to drive innovation. Watson Studio provides a suite of tools and a collaborative environment for data scientists, developers and domain experts.

Service name:

Watson Studio-gm

Choose a region/location to deploy in:

US South

Choose an organization:

alf@us.ibm.com

Choose a space:

dev

[View Docs](#) [Terms](#)

**AUTHOR** IBM

**PUBLISHED** 03/20/2018

**TYPE** Service

**LOCATION** United Kingdom, US South

### Features

- **Use what you know, learn what you don't**  
Start from a tutorial, start from a sample, or start from scratch. Tap into the power of the best of open source (RStudio, Jupyter Notebooks) and Watson services for flexible model creation. Use Python, R, or Scala. Stop downloading and configuring analysis environments and start getting insights.
- **Power on demand**  
Enterprise-scale features on demand. From data exploration and preparation, to enterprise-scale performance. Manage your data, your analytical assets, and your projects in a secured cloud environment.
- **Be a founding member**  
Take advantage of shared data sets, notebooks, models, and tutorials. Share your work with your team and your peers across job roles. Join a vibrant community of data scientists, developers, and domain experts across industries, functions, and organization types.
- **Collaborate for better outcomes**  
Work with your peers on projects to find better solutions together. Share your knowledge and your work easily with visualizations and code — and help fuel the advancement of data science and AI for all.

Need Help? [Contact IBM Cloud Sales](#)

Estimate Monthly Cost [Cost Calculator](#)

Create

<https://console.bluemix.net/catalog/services/watson-studio>

# Watson Studio



Welcome Lennart!

Watson Studio is part of IBM Watson.

[Try out](#) other IBM Watson apps.

Get started with key tasks



Refine data



New notebook



Deep learning



New Modeler flow



New model



# Deep Learning in Watson Studio

## What is Deep Learning?

IBM's experiment-centric deep learning service within Watson Studio allows data scientists to visually design their neural networks and scale out their training runs while auto-allocation means paying only for the resources utilized. Optimized for production environments, scale up your training using the NVIDIA® Tesla® V100 GPU with your preferred deep learning framework then easily deploy to the cloud or at the edge.

Deep Learning features	<b>Experiment Assistant</b>  Initiate and monitor batch training experiments then compare cross-model performance in real-time without worrying about log transfers and scripts to visualize results. You focus on designing your neural networks. We'll manage and track your assets.	<b>Open and flexible</b>  Use your preferred deep learning framework: Tensorflow, Keras, PyTorch, Caffe and more. Manage your deep learning experiments with the tools you prefer: command-line interface (CLI), Python library or an interactive user interface.	<b>Elastic GPU compute</b>  Train neural networks in parallel using market-leading NVIDIA® Tesla® GPUs - K80, P100, and V100. Pay only for what you use. Auto-allocation means no more remembering to shutdown your cloud training instances. No clusters or containers to manage.
	<b>Hyperparameter optimization</b>  Efficiently automate searching your network's hyperparameter space to ensure the best model performance with the fewest training runs.	<b>Neural Network Modeler (beta)</b>  Visually design your neural networks. Drag-and-drop layers of your neural architecture then configure and deploy using the most popular deep learning frameworks.	<a href="#">Let's talk</a>

# IBM Watson Studio and Neural Network Modeler

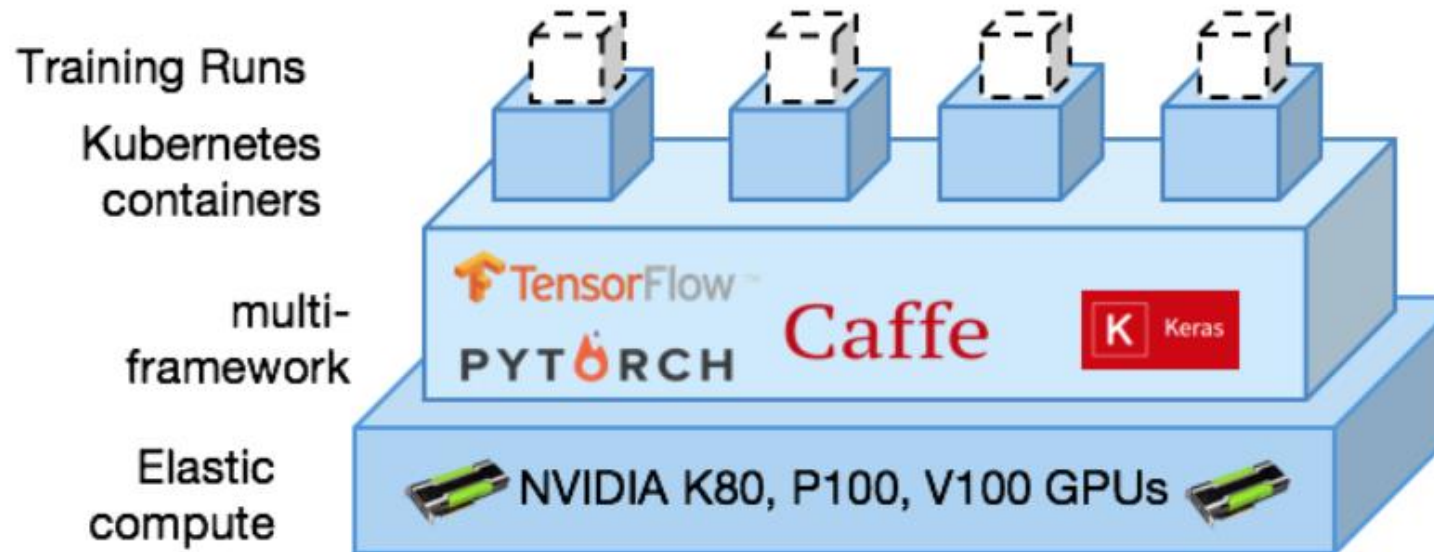
The screenshot displays the IBM Watson Studio Neural Network Modeler interface. The top navigation bar includes 'IBM Watson', 'Projects', 'Tools', 'Catalog', 'Community', and 'Services'. The breadcrumb trail shows 'My Projects / Deep Learning Experiments / CNN on CIFAR10'. The left sidebar, titled 'Search Nodes', lists categories: Input, Activation, Convolution, and Core. The 'Core' category is expanded, showing nodes like Concat, Dense, Dropout, Elementwise Operations, Flatten, Reshape, and Split. The main workspace shows a flowchart of a CNN architecture. The flow starts with 'Image Data', followed by two parallel paths of 'Conv 2d' and 'ReLU' layers, then a 'Pooling 2d' layer, and a 'Dropout' layer. These paths merge into a 'Flatten' layer, followed by a 'Dense' layer, another 'ReLU' layer, another 'Dropout' layer, and a final 'Dense' layer. The output goes to a 'Softmax' layer, which branches into 'Accuracy' and 'Sigmoid Cross-Entropy' loss, with the latter connected to an 'RMSE' metric. The right sidebar shows the 'Conv 2d' node settings:

- Number of filters\***: 32
- Kernel row\***: 3
- Kernel col\***: 3
- Stride row**: 1
- Stride col**: 1
- Border mode**: ☐ VALID ☒ SAME
- Initialization**: glorot\_normal
- Rise**: (slider)

At the bottom right of the settings panel are 'Cancel' and 'Save' buttons.

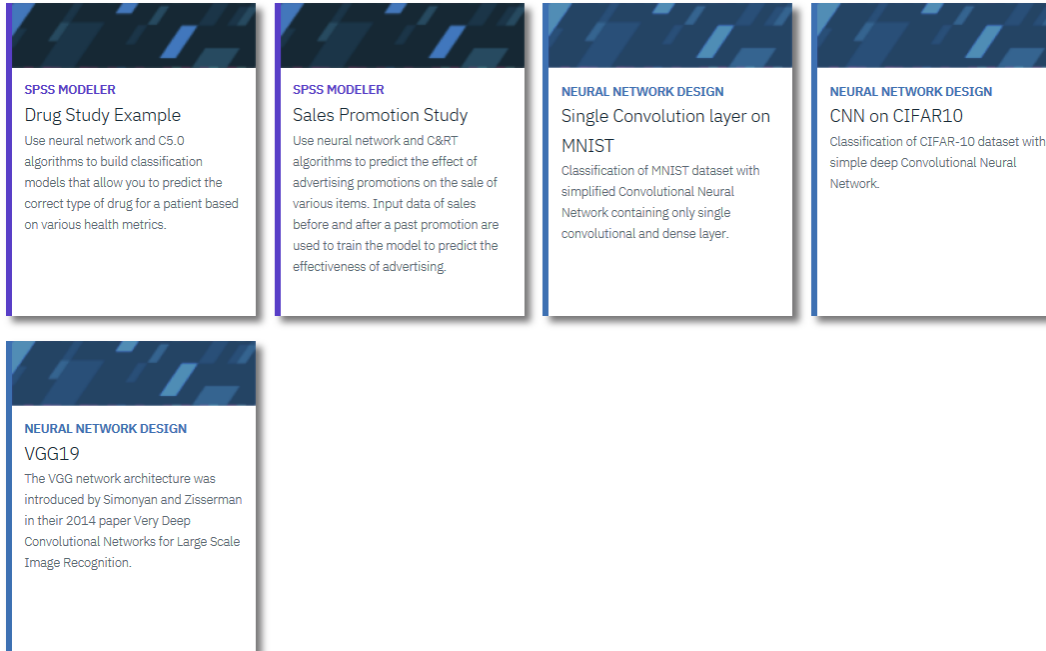


# Deep Learning as a Service within Watson Studio



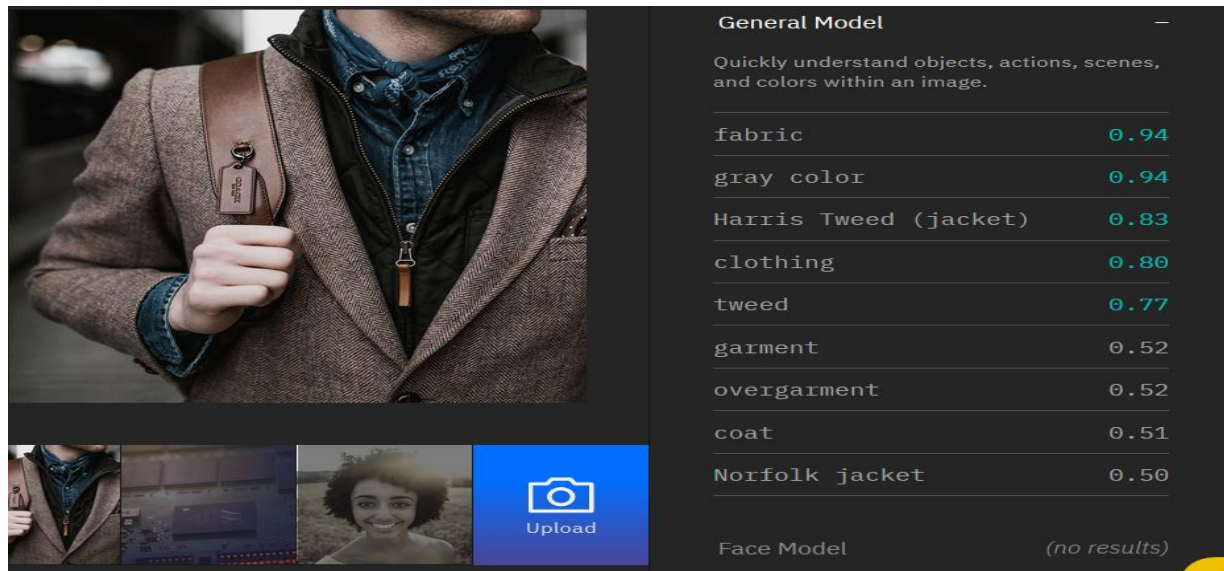
<https://www.ibm.com/blogs/research/2018/03/deep-learning-advances/>

# Watson Studio Modeler, From Example



You can't create a Modeler yet.

# Watson Visual Recognition



The screenshot displays the Watson Visual Recognition interface. On the left, a large image of a person wearing a brown Harris Tweed jacket is shown. Below this image are three smaller thumbnails: the same jacket, a close-up of the jacket's texture, and a portrait of a woman. To the right of the main image is a panel titled 'General Model' with the description 'Quickly understand objects, actions, scenes, and colors within an image.' Below this is a list of recognized items with their confidence scores:

Item	Score
fabric	0.94
gray color	0.94
Harris Tweed (jacket)	0.83
clothing	0.80
tweed	0.77
garment	0.52
overgarment	0.52
coat	0.51
Norfolk jacket	0.50


At the bottom of the interface, there is a blue 'Upload' button with a camera icon. Below the 'General Model' section, the 'Face Model' section is visible with the text '(no results)'.

Quickly and accurately tag, classify and train visual content using machine learning.

Train models effortlessly with Watson Studio

<https://www.ibm.com/watson/services/visual-recognition/demo/index.html#watson-demo>

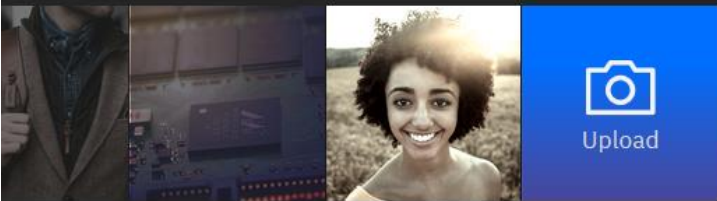
# How about ML and some Chow Mien?



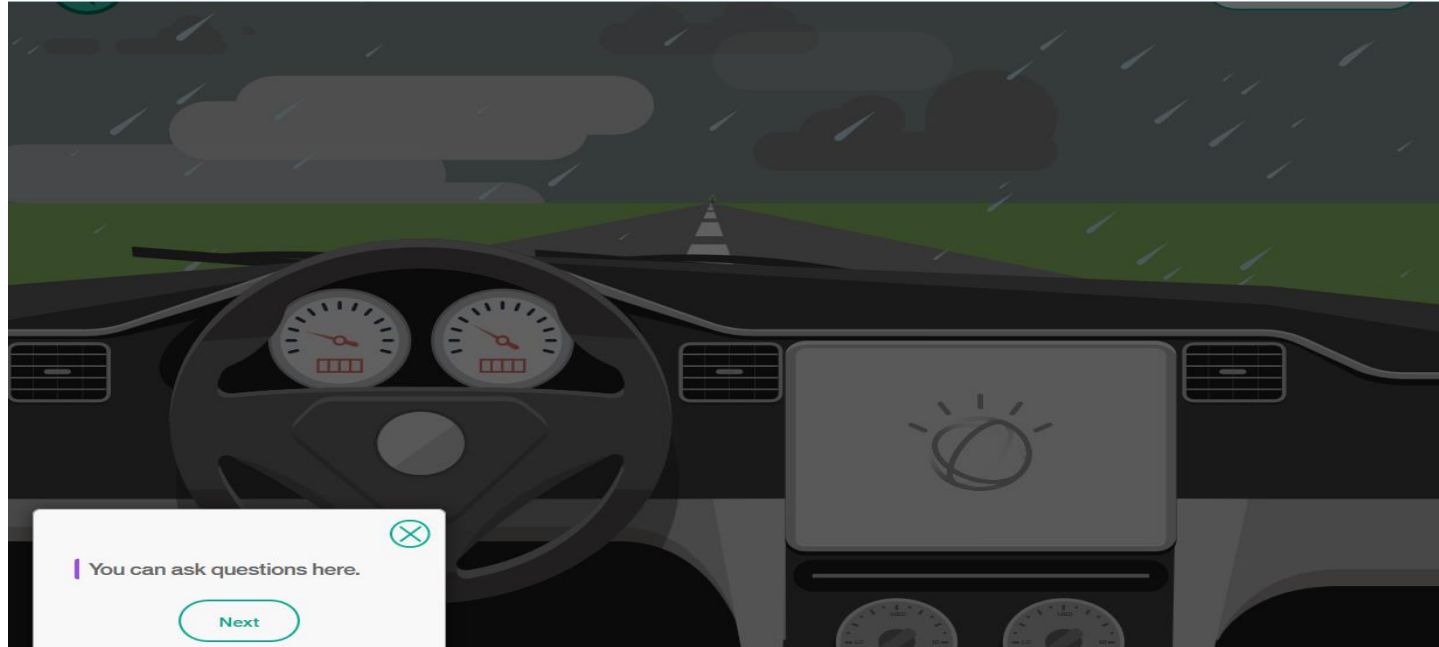
General Model

Quickly understand objects, actions, scenes, and colors within an image.

dish	0.85
nutrition	0.85
food	0.85
food product	0.79
light brown color	0.75
chow mein	0.71
vegetable	0.66
plant	0.66
bean sprout	0.66



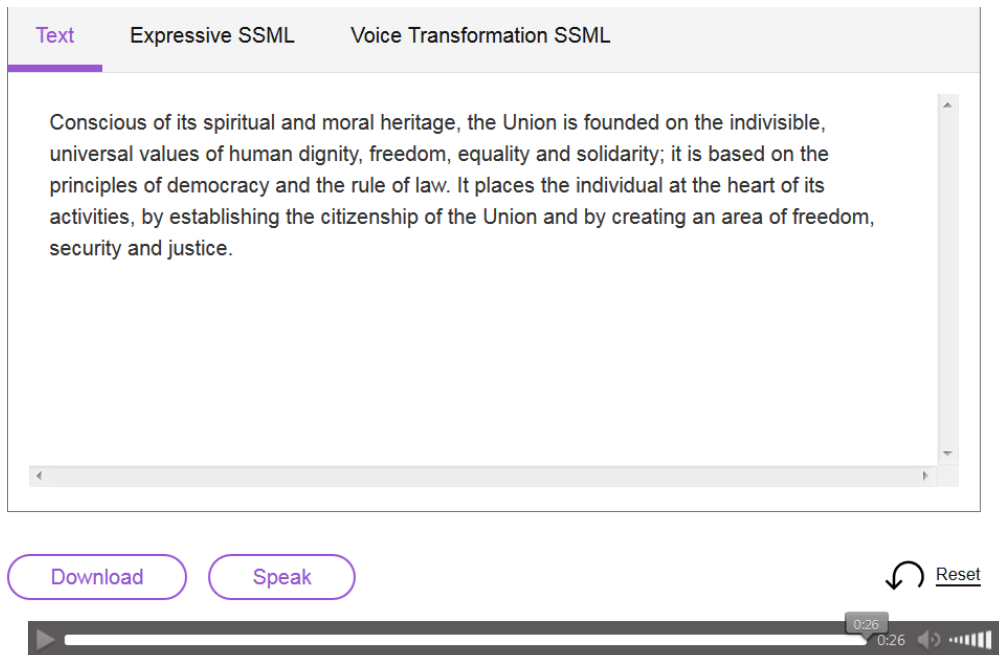
# Let's go for a drive with a Chatbot



- <https://conversation-demo.ng.bluemix.net/>

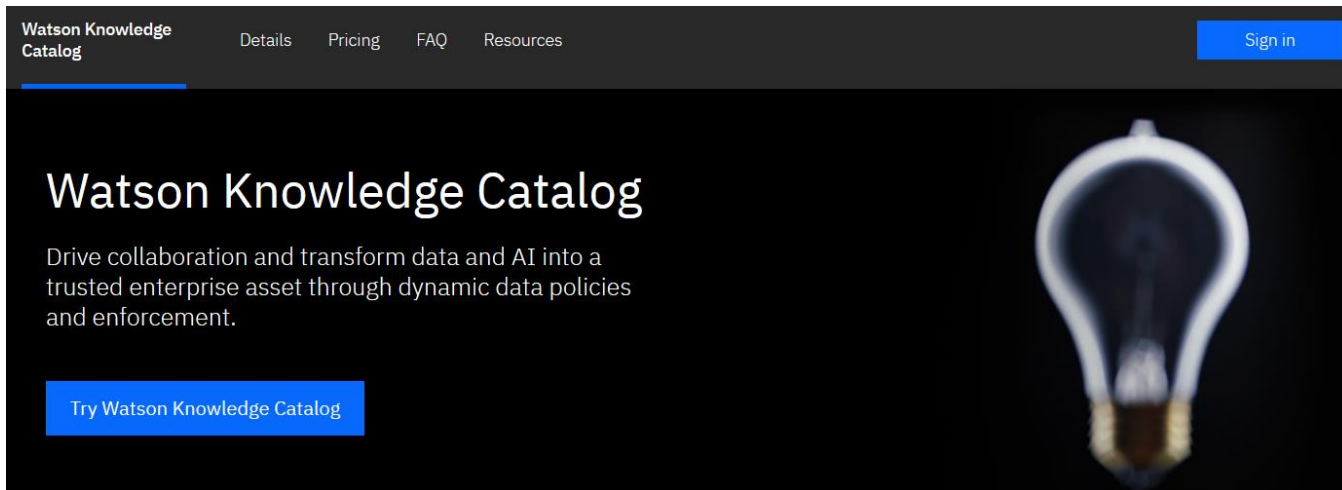
# Watson Text to Speech

- <https://text-to-speech-demo.ng.bluemix.net/>



The screenshot shows the Watson Text to Speech demo interface. At the top, there are three tabs: "Text" (selected), "Expressive SSML", and "Voice Transformation SSML". The "Text" tab is active, displaying a text input area with the following text: "Conscious of its spiritual and moral heritage, the Union is founded on the indivisible, universal values of human dignity, freedom, equality and solidarity; it is based on the principles of democracy and the rule of law. It places the individual at the heart of its activities, by establishing the citizenship of the Union and by creating an area of freedom, security and justice." Below the text input area, there are two buttons: "Download" and "Speak". To the right of these buttons is a "Reset" button with a circular arrow icon. At the bottom of the interface, there is a progress bar and a volume control icon. The progress bar shows a duration of 0:26.

# Watson Knowledge Catalog



<https://www.ibm.com/blogs/bluemix/2018/03/ibm-watson-knowledge-catalog-find-prepare-data-analytic-assets-power-ai/>

# Watson Machine Learning

Artificial intelligence and machine learning

Home

Why Watson?

Build with

Explore machine learning and cognitive computing resources, and build artificial intelligence functions into your app.

Read the O'Reilly book on getting started with AI

<https://developer.ibm.com/ai/>



# IBM's CodaIT

# IBM CodaIT

## Center for Open-Source Data & AI Technologies

Improving the Enterprise AI Lifecycle in Open  
Source



<https://developer.ibm.com/code/open/centers/codait/>

# 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
- provides a standardized approach to classify, annotate, and deploy these models for prediction and inferencing,
- including models that can be deployed and customized in IBM's AI application development platform, **Watson Studio**.

# IBM Code Model Asset Exchange

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

<https://developer.ibm.com/code/exchanges/models/>

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

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



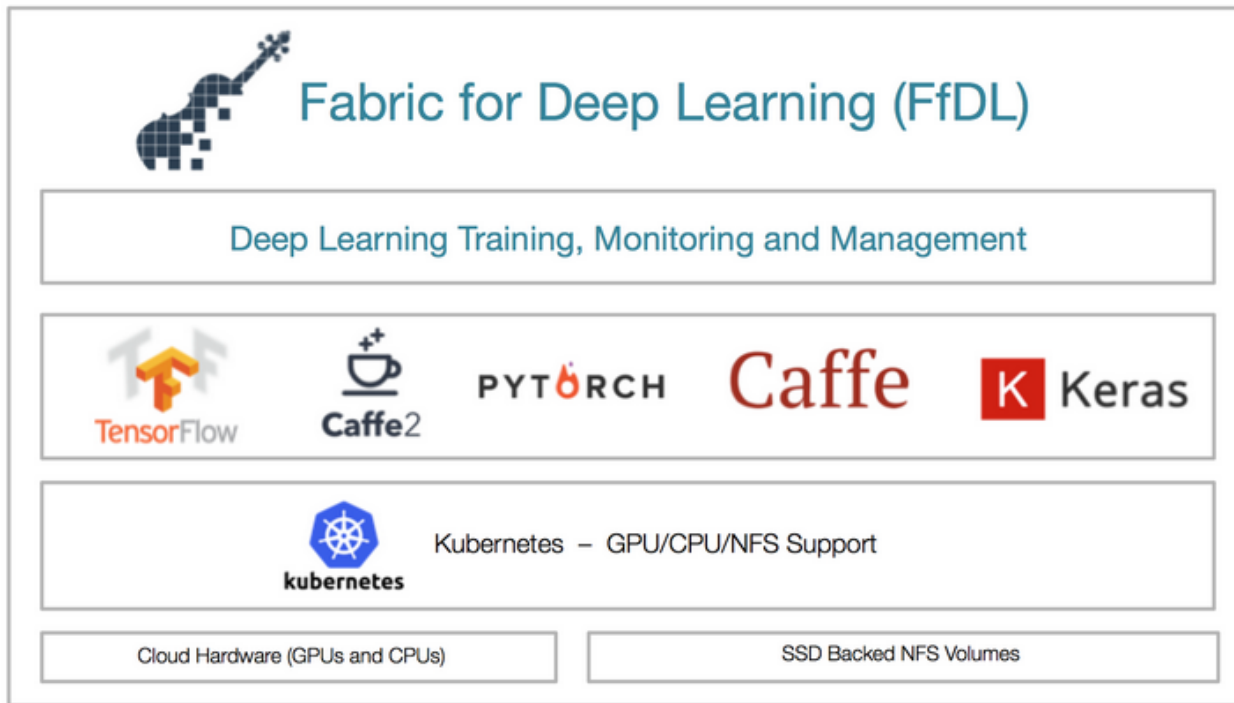
## Image Caption Generator

Generate captions that describe the contents of images.

Get this model

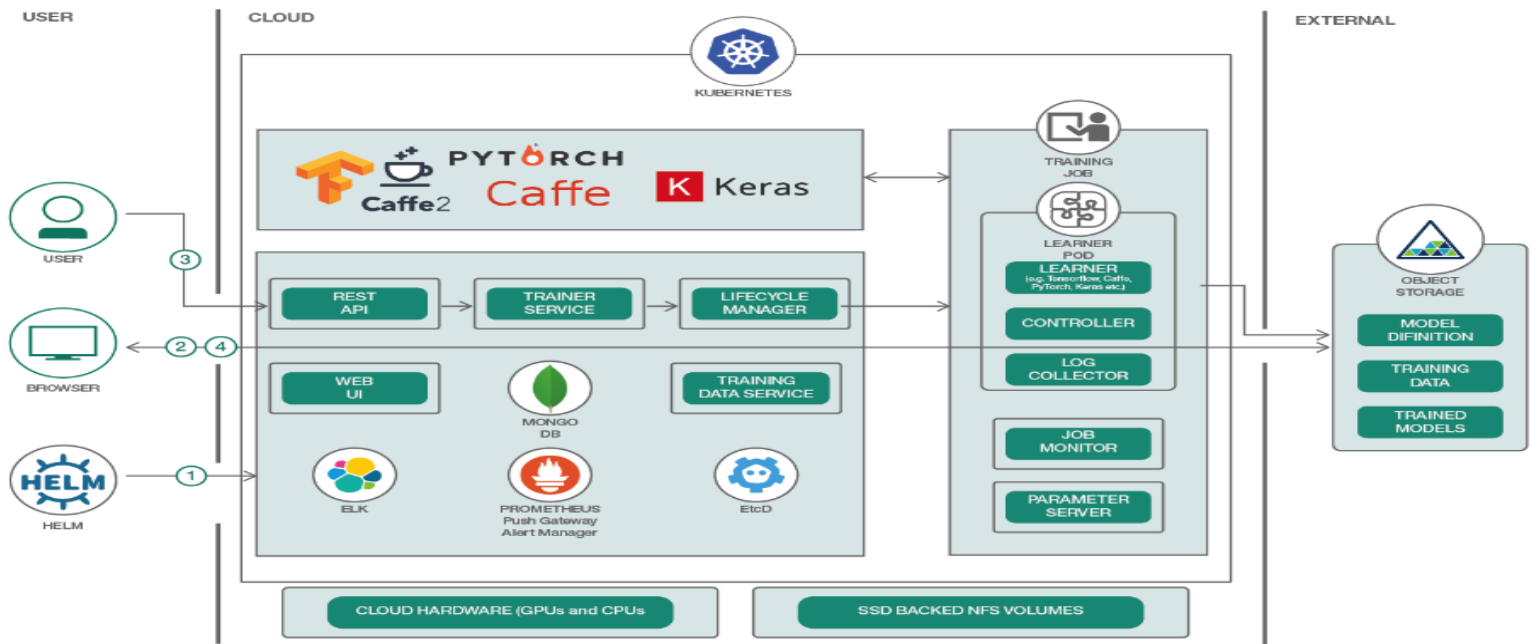
including models that can be deployed and customized in IBM's AI application development platform, **Watson Studio**.

# Fabric for Deep Learning(FfDL)



<https://developer.ibm.com/code/2018/03/20/fabric-for-deep-learning/>

# Install and consume a deep learning platform on Kubernetes with TensorFlow, Caffe, PyTorch, and more



<https://developer.ibm.com/code/patterns/deploy-and-use-a-multi-framework-deep-learning-platform-on-kubernetes/>

# Adversarial Robustness Toolbox (ART v0.1)

---

docs **passing**

This is a library dedicated to **adversarial machine learning**. Its purpose is to allow rapid crafting and analysis of attacks and defense methods for machine learning models. The Adversarial Robustness Toolbox provides an implementation for many state-of-the-art methods for attacking and defending classifiers.

The library is still under development. Feedback, bug reports and extension requests are highly appreciated.

## Supported attack and defense methods

---

The Adversarial Robustness Toolbox contains implementations of the following attacks:

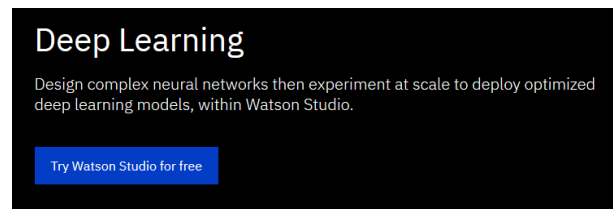
- Deep Fool ([Moosavi-Dezfooli et al., 2015](#))
- Fast Gradient Method ([Goodfellow et al., 2014](#))
- Jacobian Saliency Map ([Papernot et al., 2016](#))
- Universal Perturbation ([Moosavi-Dezfooli et al., 2016](#))
- Virtual Adversarial Method ([Moosavi-Dezfooli et al., 2015](#))
- C&W Attack ([Carlini and Wagner, 2016](#))
- NewtonFool ([Jang et al., 2017](#))

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

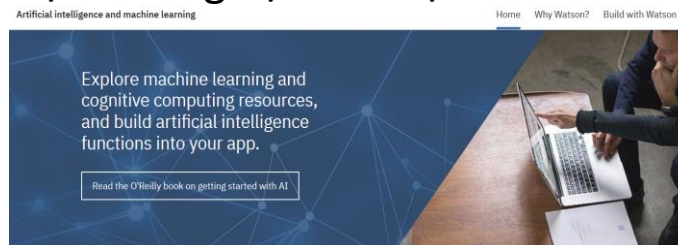
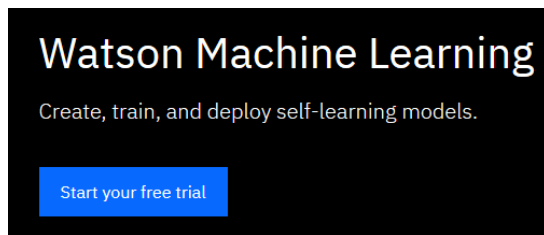
# Where do we go from here?

- kaggle.com Open Source Datasets

kaggle™



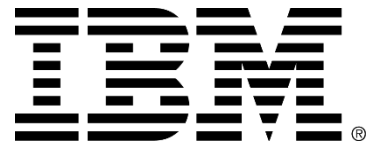
<https://developer.ibm.com/code/exchanges/models/> <https://www.ibm.com/cloud/machine-learning>



<https://www.ibm.com/cloud/machine-learning>

<https://developer.ibm.com/ai/>





# IBM Machine Learning for z/OS

*A private cloud (on-premises) deployment of IBM Machine Learning*

**Cost effective, low latency,  
high security**



Benefit from your private cloud  
(on-premises) z Systems  
investments

- Gain advantage from z Systems infrastructure, people and processes
- Leverage z Systems data in place while combining structured and unstructured data from z and non-z data sources
- Access live transactional data

# Tell me about the terminology


SAN FRANCISCO — Apple has hired Google’s chief of search and artificial intelligence, John Giannandrea, a major coup in its bid to catch up to the artificial intelligence technology of its rivals.

Apple said on Tuesday that Mr. Giannandrea will run Apple’s “**machine learning and A.I. strategy**,” and become one of 16 executives who report directly to Apple’s chief executive, Timothy D. Cook.

<https://www.nytimes.com/2018/04/03/business/apple-hires-googles-ai-chief.html>

[illegible]

# Data science and machine learning at IBM



Analytics > Data science >

## Data science and machine learning

IBM has one of the leading data science platforms, which enables you to easily collaborate across teams, use the top open-source tools and scale at the speed your business requires

[Learn more](#)

[Let's talk](#)

[↓ Overview](#)   [↓ Capabilities](#)   [↓ Featured product](#)   [↓ Products](#)

# Runtimes Supported by Watson Machine Learning\*



**More frameworks including deep learning are on the Roadmap**

\*IBM ML in DSX Local is currently in Beta and only supports Spark ML models (scikit-learn and other deep learning frameworks are on the roadmap).