Machine Learning in Engineering: Panacea or Deep Trouble ?

Kostas Plataniotis

ECE Department - University of Toronto

kostas@ece.utoronto.ca

www.dsp.utoronto.ca

Distinguished Lecturer Series "Leo the Mathematician"
School of Informatics
Aristotle University of Thessaloniki

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

What this presentation is all about ?

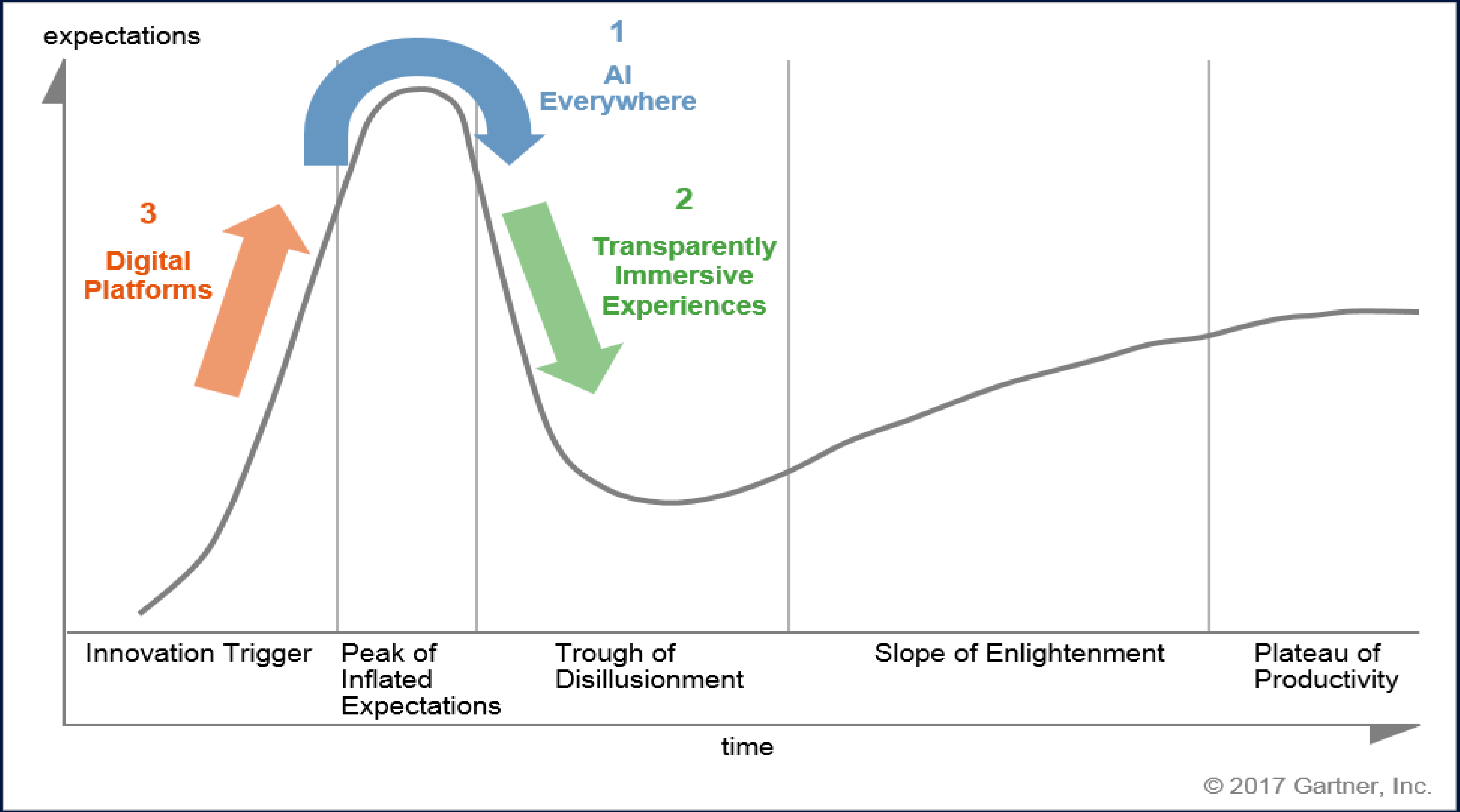
A personal account of (some) key issues in the emerging field of machine learning

(relevant to the engineering practice)

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering UNIVERSITY OF TORONTO

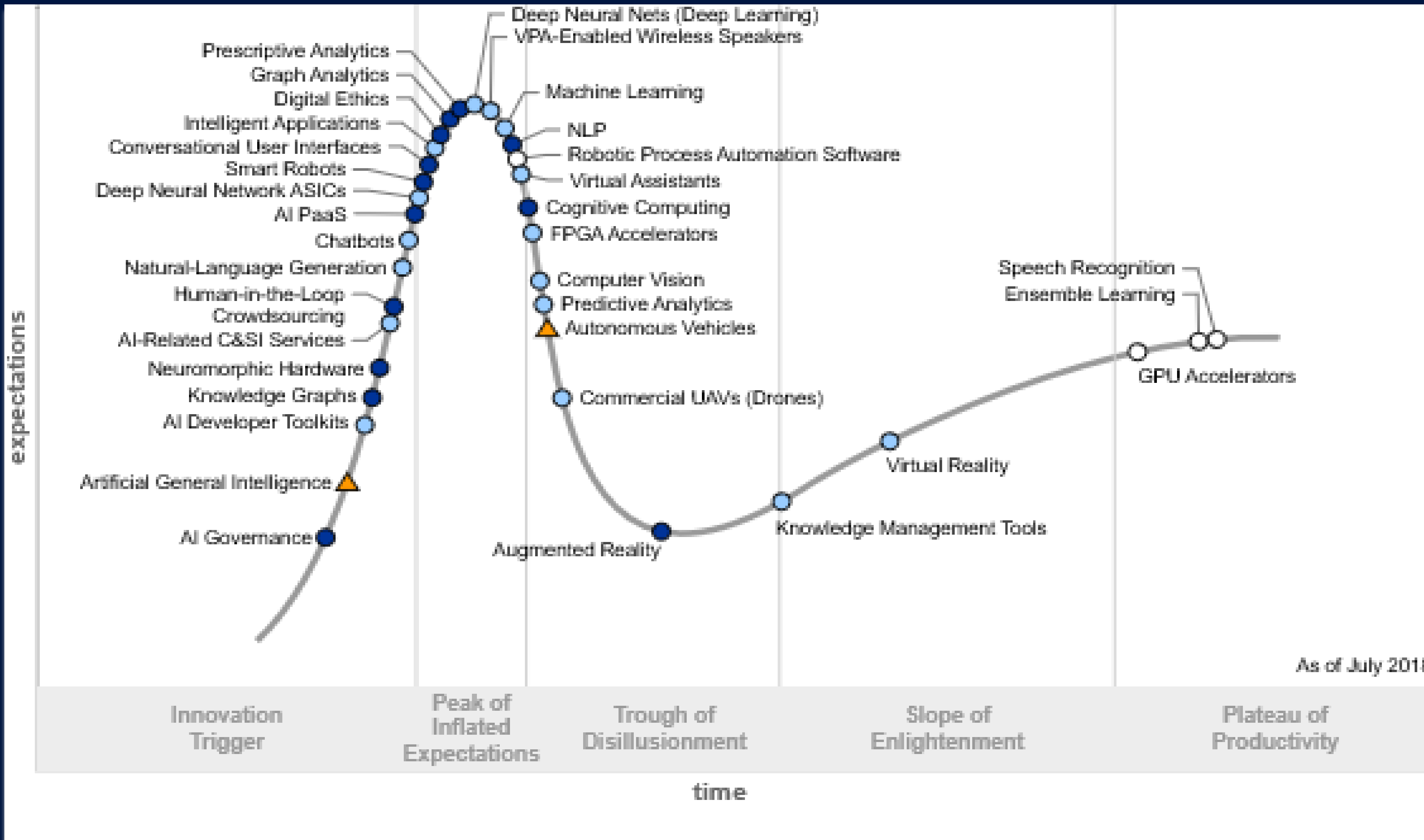
FACULTY OF APPLIED SCIENCE & ENGINEERING

Why a presentation on Machine Learning ?
The "hype cycle" (2017-Gartner)



The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY
OF APPLIED
SCIENCE &
ENGINEERING

The "hype cycle" (2018-Gartner) (in data science and machine learning)



The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

“Priority Matrix” for Artificial Intelligence (2018 Gartner)

Priority Matrix for Artificial Intelligence, 2018

benefit	years to mainstream adoption: less than 2 years	2 to 5 years	5 to 10 years	more than 10 years
transformational	**Speech Recognition**	**AI-Related C&SI Services** **Chatbots** **Deep Neural Nets (Deep Learning)** **Intelligent Applications** **Machine Learning** **Virtual Assistants** **VPA-Enabled Wireless Speakers**	Cognitive Computing Conversational User Interfaces Neuromorphic Hardware NLP	Artificial General Intelligence Autonomous Vehicles
high	**Ensemble Learning** **GPU Accelerators** **Robotic Process Automation Software**	AI Developer Toolkits Commercial UAVs (Drones) Computer Vision Deep Neural Network ASICs Natural-Language Generation Predictive Analytics	AI Governance AI PaaS Augmented Reality Digital Ethics Graph Analytics Human-in-the-Loop Crowdsourcing Knowledge Graphs Prescriptive Analytics Smart Robots	
moderate		FPGA Accelerators Knowledge Management Tools Virtual Reality		
low				

As of July 2018

ID: 357478

Outline

- **A definition (or two)**
- AI/ML: The big picture
- Altum Visum on deep learning networks
- Machine Learning: Myths & Realities
- Machine Learning as a process
- Long Term View: Explainable Artificial Intelligence (XAI)
- Long Term View: Human-like intelligence
- Epilogue

It’s all Greek to me

How we learn / know something:

- **Techné** (skill) - Knowing by doing. A carpenter learns to build by building, a potter by making pots.
- **Epistemé** (science) - Knowing by demonstration. Scientific facts are capable of being repeatedly demonstrated.
- **Nous** (intuition) - Knowing without the demonstration of invariable facts.

Nicomachean Ethics - Aristotle

It’s still Greek to me

The pertinent questions :

what are we learning and why?

The Aristotelian answer:

The goal of **episteme´** is to know truth from falsehood. The goal of **phronesis (nous)** is to know good from bad, and the goal of **techné** is to know how to express and appreciate beauty.

The Aristotelian view:

Each of these kinds of knowledge is a uniquely human capacity, thus the aim of learning is to help human beings become more fully human.

Nicomachean Ethics - Aristotle

(Lay) Definitions – I

Learning: The activity or process of gaining knowledge or skill by studying, practicing, being taught, or experiencing something. (Merriam Webster Dictionary).

Machine: a mechanically, electrically, or electronically operated device for performing a task. Archaic : a constructed thing whether material or immaterial. (Merriam Webster Dictionary).

(Lay) Definitions - II

- **Artificial Intelligence (AI)**: the broader concept of machines being able to carry out tasks in a way that we would consider "smart". [1]
- Machine Learning (ML): a current application of AI based around the idea that we should really just be able to give machines access to data and let them learn for themselves. [1]

[1] Bernard Marr, What Is The Difference Between Artificial Intelligence And Machine Learning?, Forbes Magazine, accessed online, December 6, 2016.

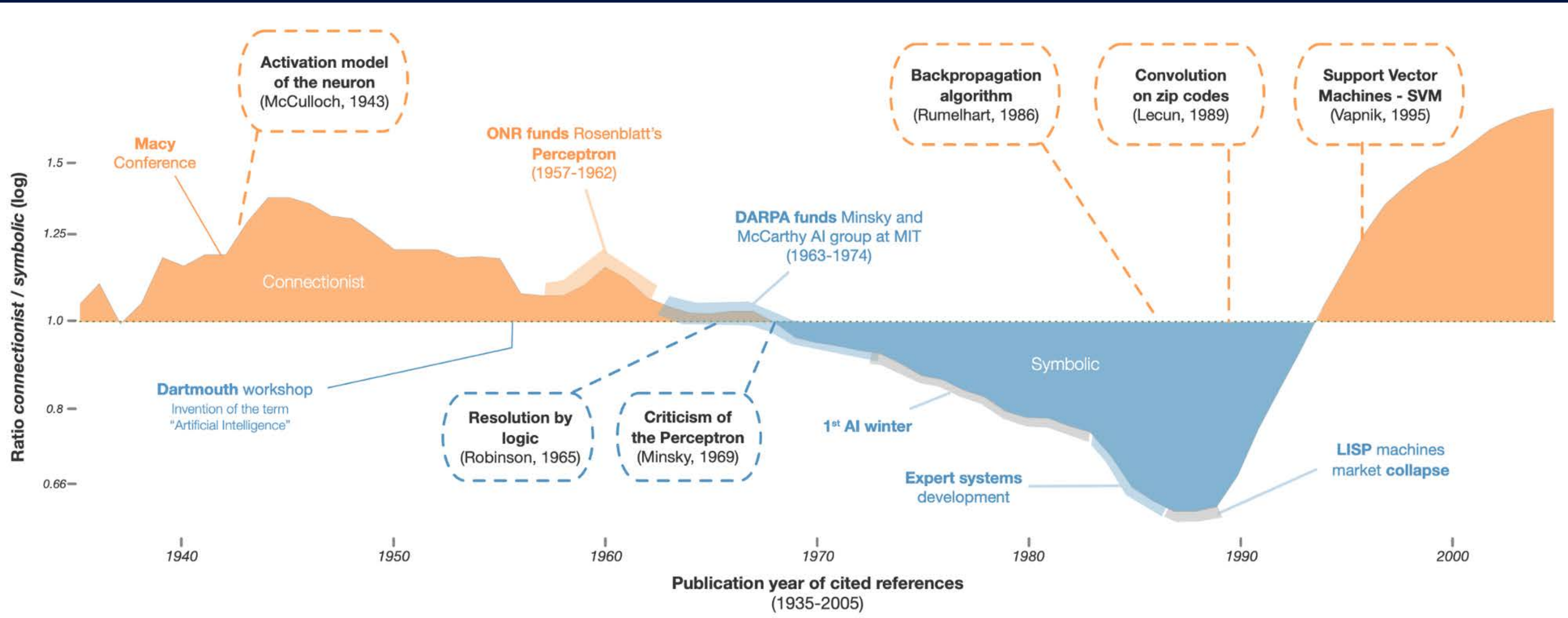
Outline

- A definition (or two)
- **AI/ML: The big picture**
- Altum Visum on deep learning networks
- Machine Learning: Myths & Realities
- Machine Learning as a process
- Long Term View: Explainable Artificial Intelligence (XAI)
- Long Term View: Human-like intelligence
- Epilogue

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Machine Learning: Big Picture



Credit: Carlos E Perez

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Some fundamentals

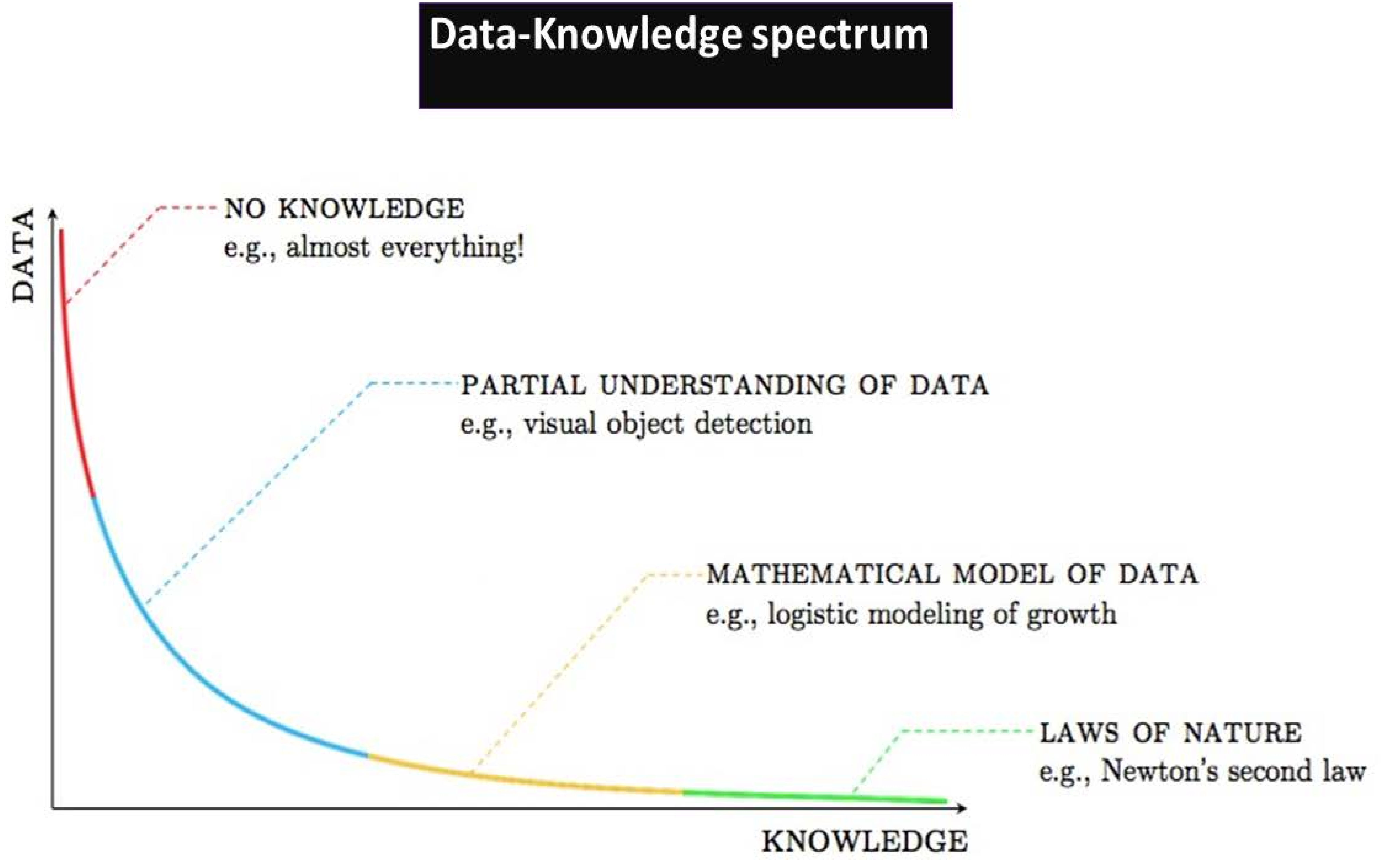
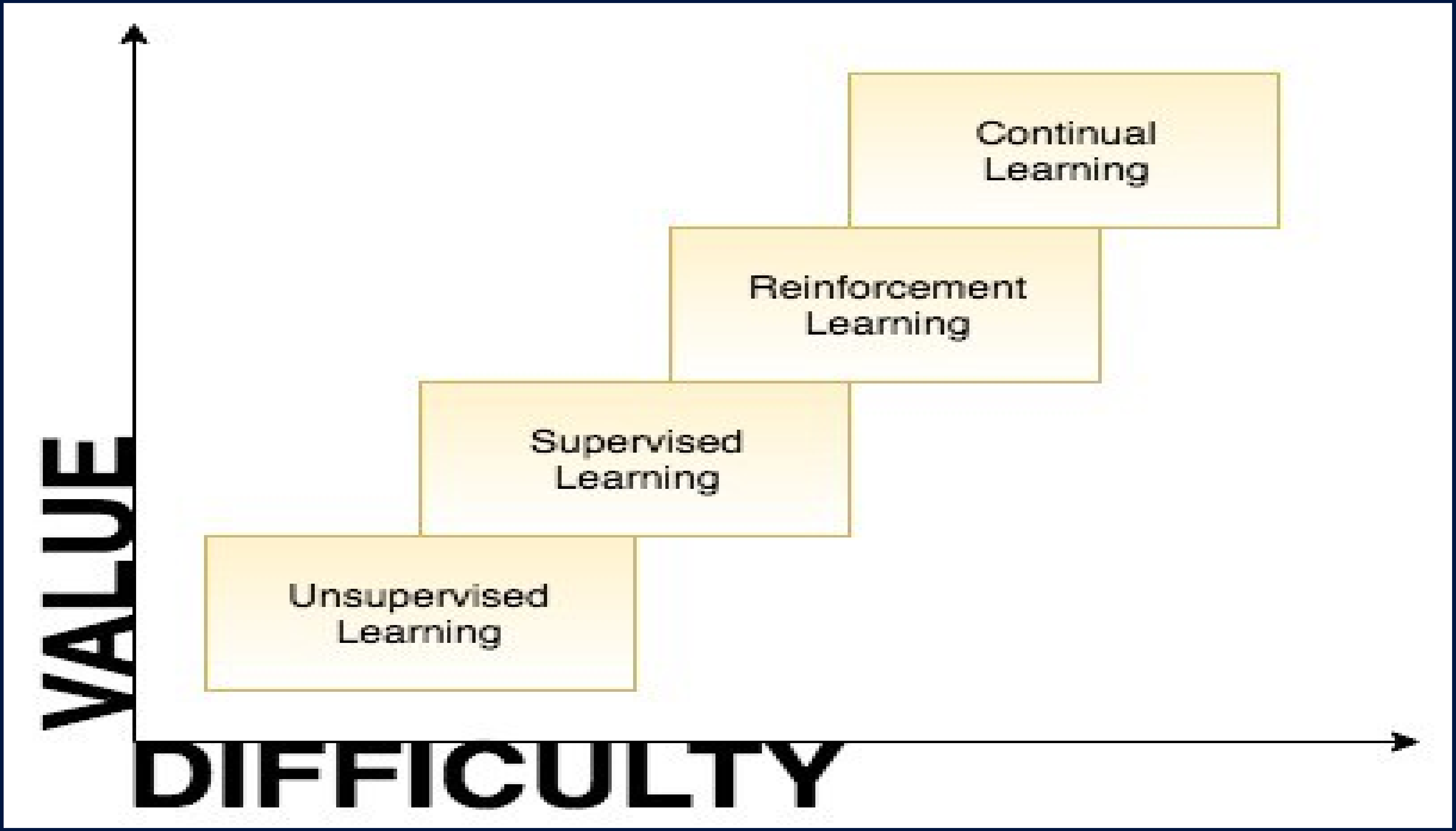


Image taken from Machine Learning Refined, by Watt – Borhani - Katsaggelos

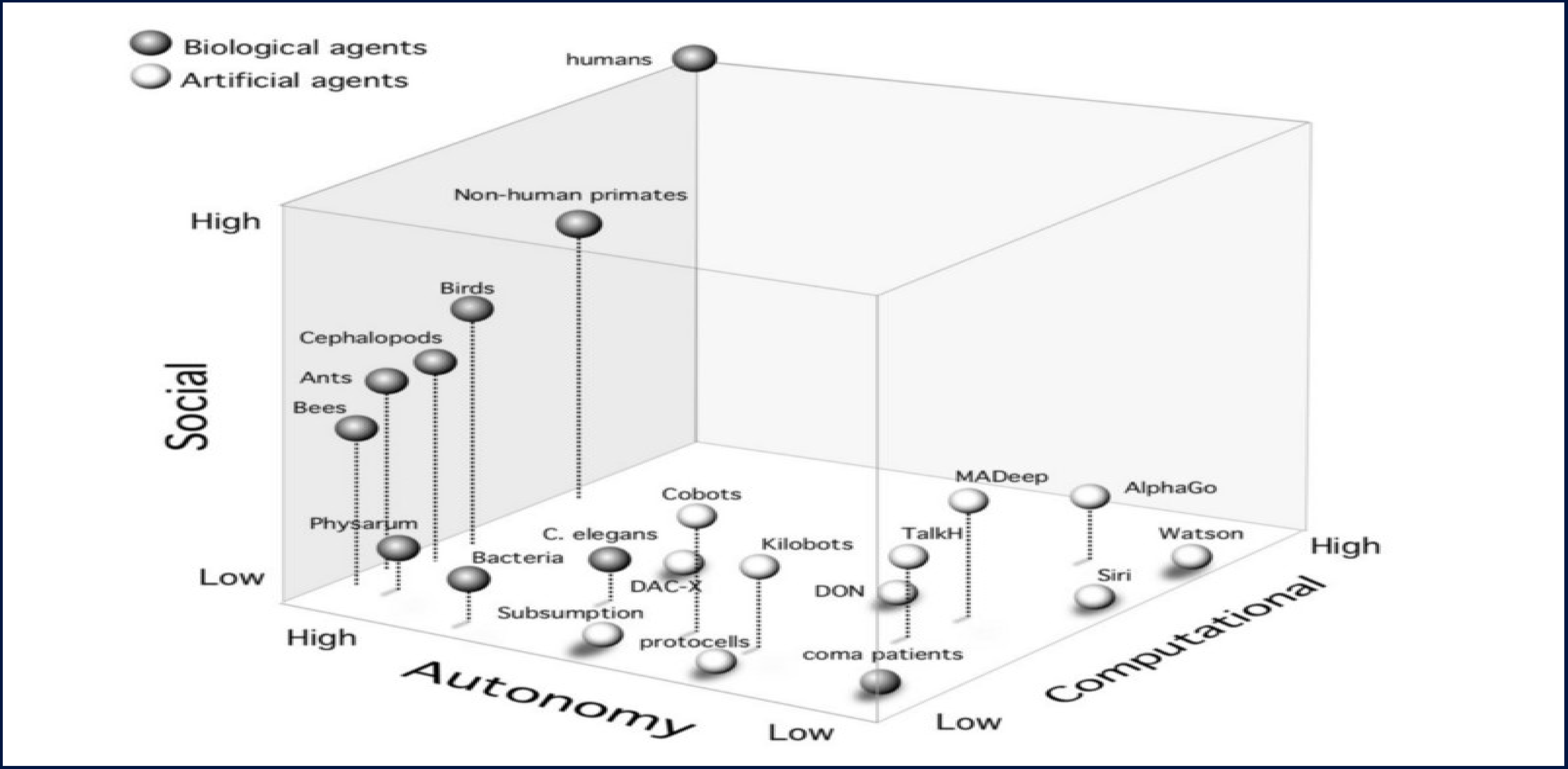
Types of Learning



Credit: Carlos E Perez

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Problems to be solved w/t AI



Credit: https://arxiv.org/pdf/1705.11190.pdf

Example: Computational Pathology (CP)

Definition

"Computational Pathology investigates a **complete probabilistic treatment** of scientific and clinical workflows in general pathology, i.e. it combines experimental design, statistical pattern recognition and survival analysis within an **unified framework** to answer scientific and clinical questions in pathology."

[Fuchs 2011]

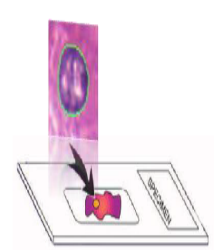
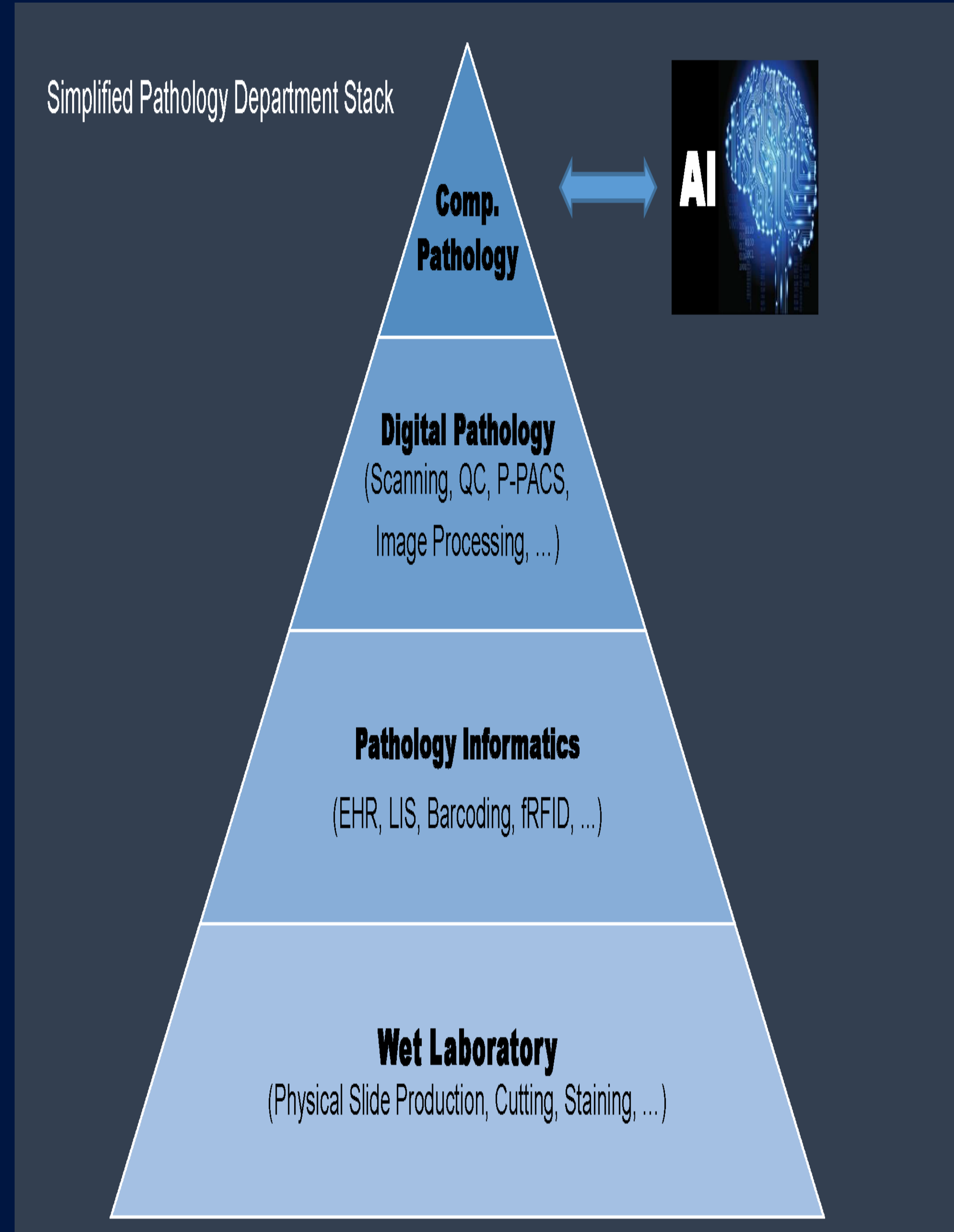


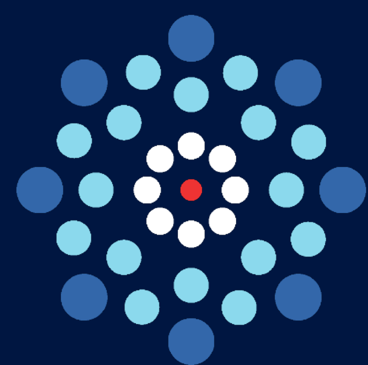
Image Credit : Thomas J Fuchs & PAIGE.AI



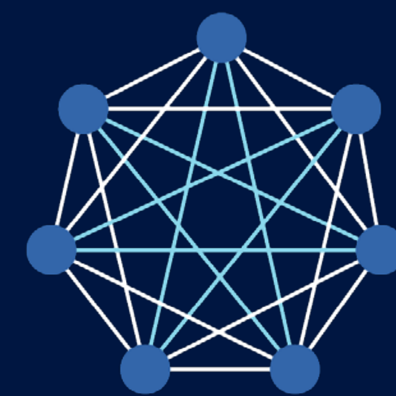
Outline

- A definition (or two)
- AI/ML: The big picture
- **Altum Visum on deep learning networks**
- Machine Learning: Myths & Realities
- Machine Learning as a process
- Long Term View: Explainable Artificial Intelligence (XAI)
- Long Term View: Human-like intelligence
- Epilogue

Deep Neural Networks – Where we are

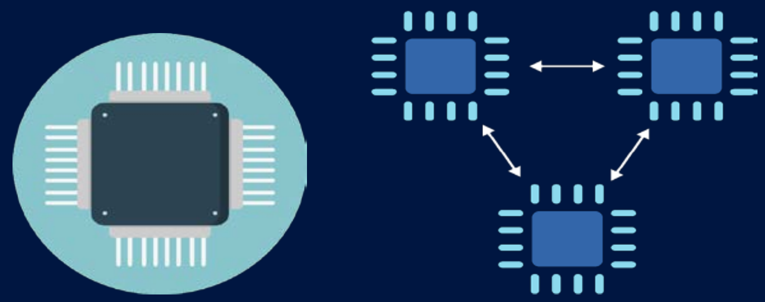


Large data



Large and complex models

Frameworks & Libraries



Training Hardware

The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY
OF APPLIED
SCIENCE &
ENGINEERING

The “Data Sets” (laboratory style)

Image credit: Xuedong Huang

The *ImageNet* Large Scale Visual Recognition Challenge (ILSVRC) evaluates algorithms for object detection and image classification at large scale

Image Credit : Ferenc Huszar

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

The “Data Sets” (laboratory style)

Dataset Sizes: Computer Vision vs. Computational Pathology

1 Whole Slide
= 100,000 x 60,000
= **6 billion pixels**

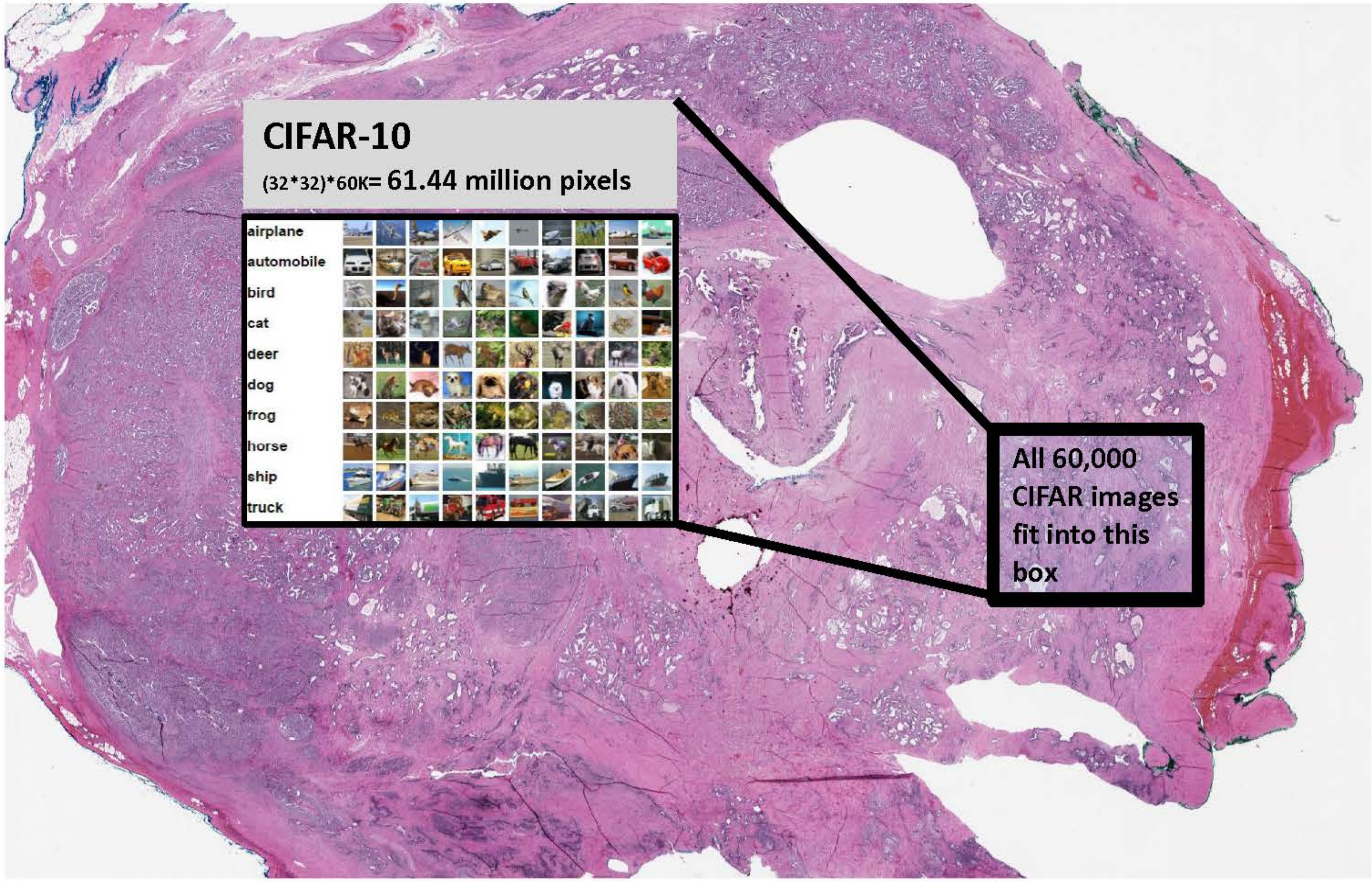
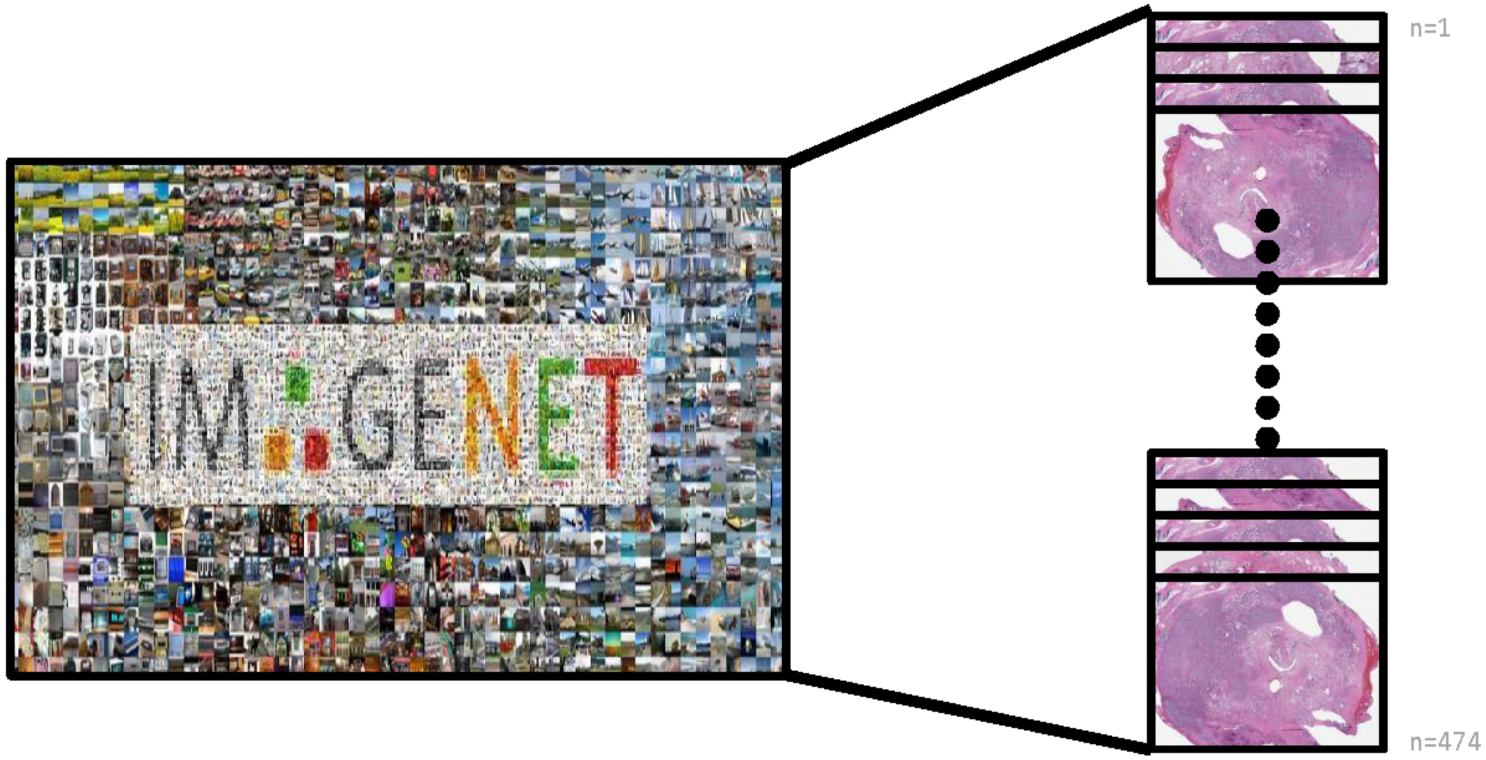


Image Credit : Thomas J Fuchs & PAIGE.AI

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

The “Data Sets” (laboratory style)

Dataset Sizes: Computer Vision vs. Computational Pathology

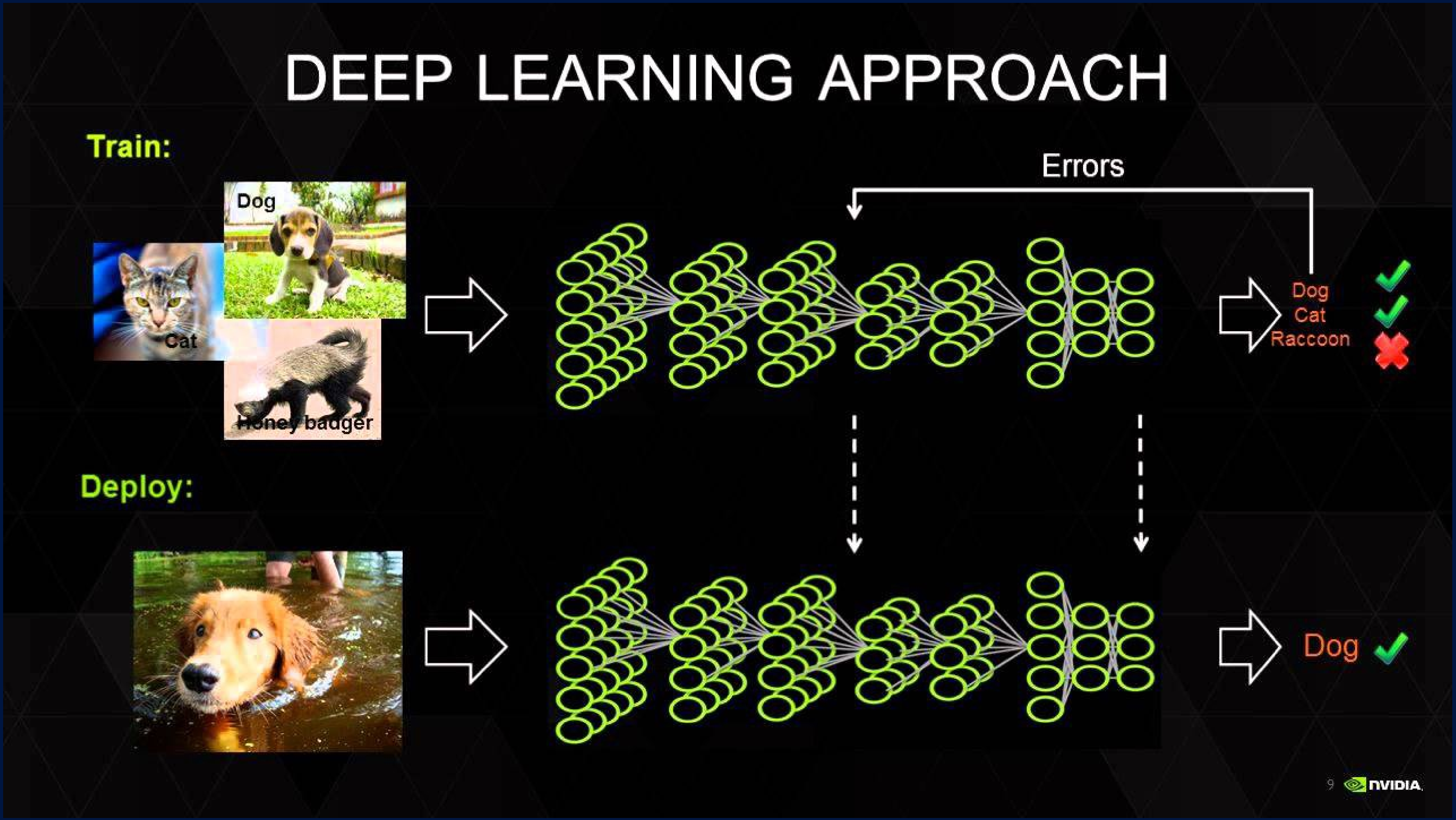


All of ImageNet
482 x 415 * 14,197,122
= 2.8 trillion pixels

474 Whole Slides
100,000 x 60,000 *474
= 2.8 trillion pixels

Image Credit : Thomas J Fuchs & PAIGE.AI

Modern Deep Neural Networks (DNN)



Convolutional NN (DNN) in popular blogs - I

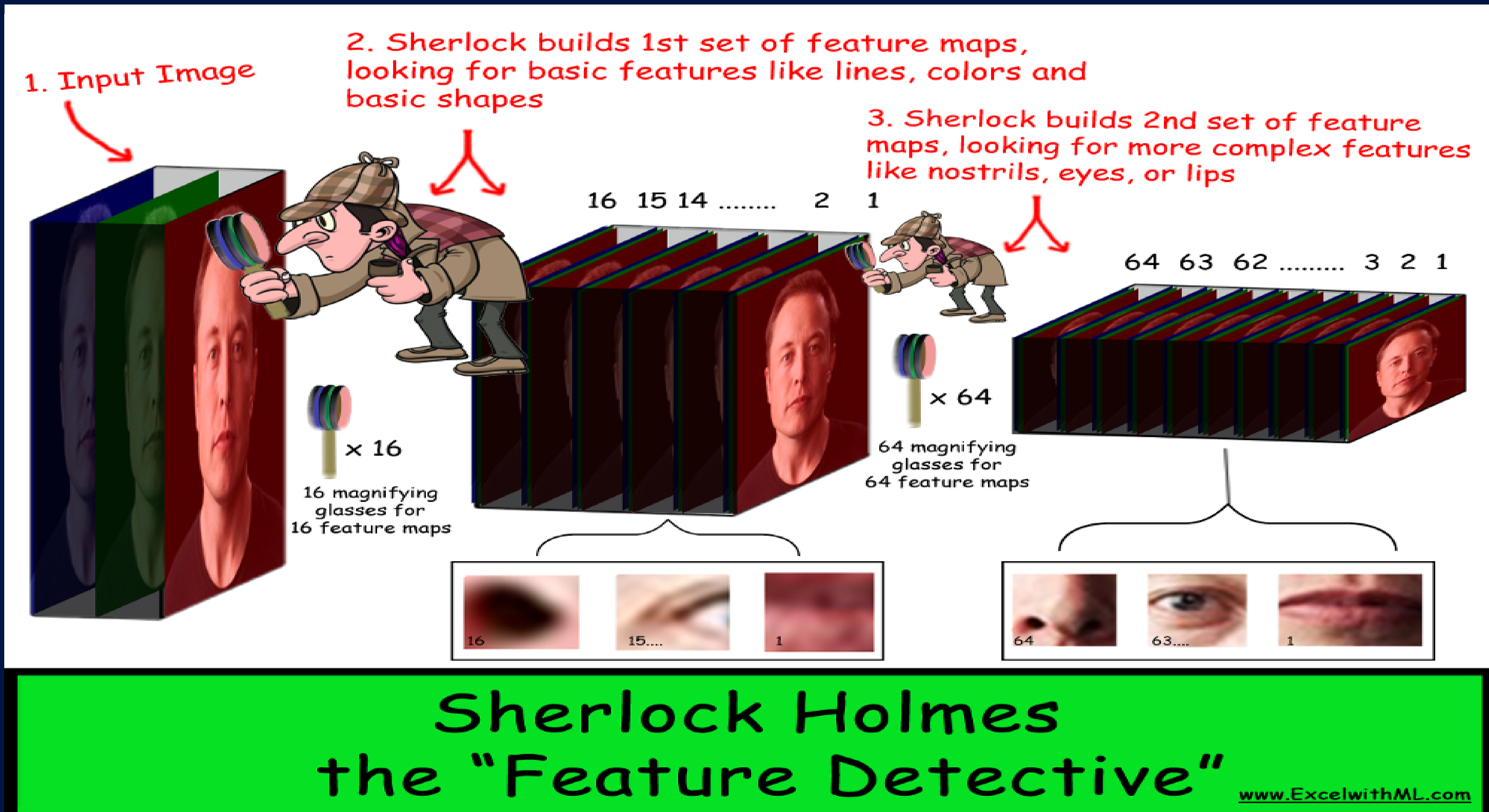


Image Credit: Dave Smith;
https://towardsdatascience.com/cutting-edge-face-recognition-is-complicated-these-spreadsheets-make-it-easier-e7864dbf0e1a
Accessed; August 7, 2018

Convolutional NN (DNN) in popular blogs - II

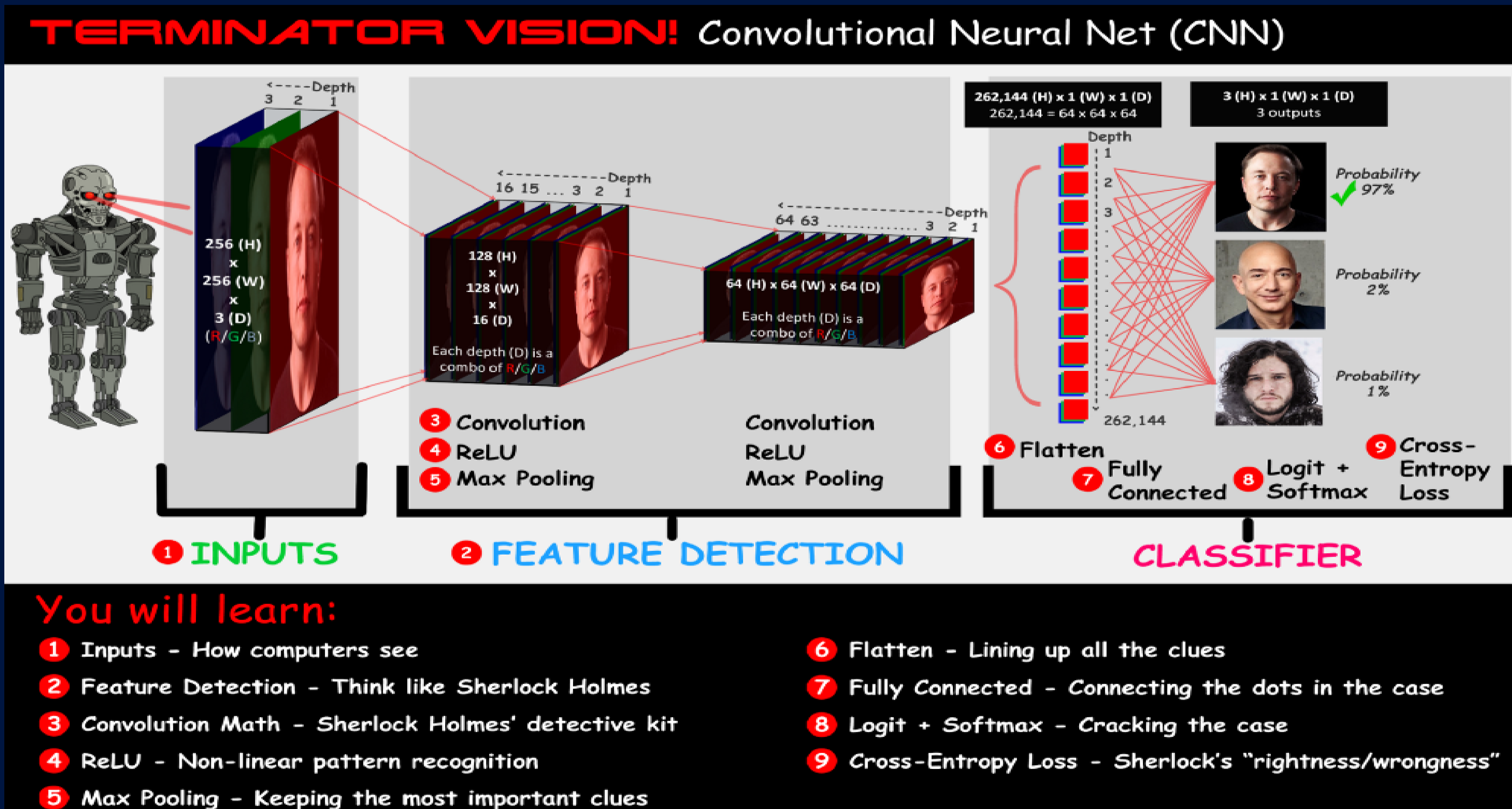
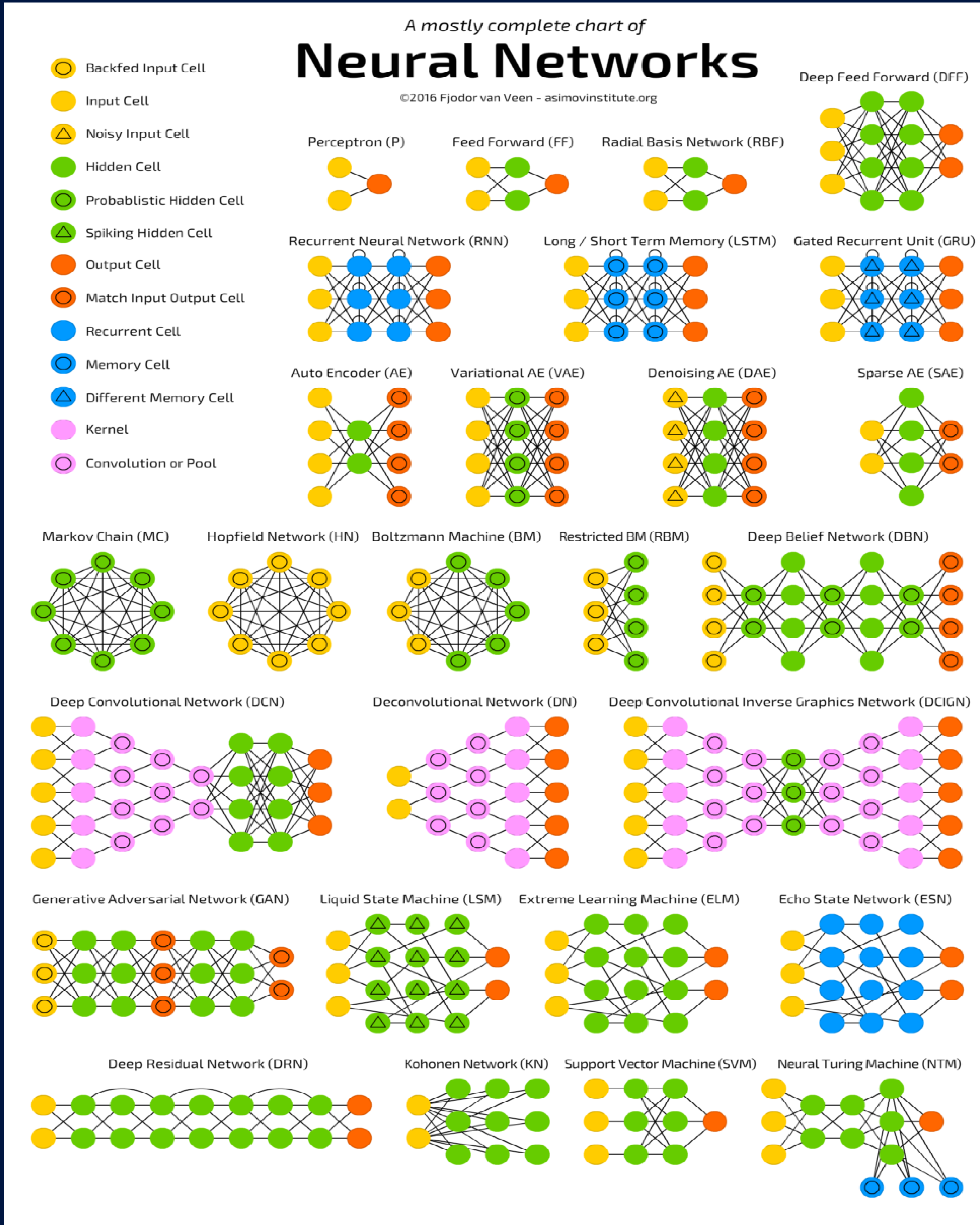


Image Credit: Dave Smith; https://towardsdatascience.com/cutting-edge-face-recognition-is-complicated-these-spreadsheets-make-it-easier-e7864dbf0e1a Accessed; August 7, 2018

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING



Taxonomy

Source:

http://www.asimovinstitute.org/neural-network-zoo/

Training Hardware

Deep Learning Hardware (2016)

GPUs: Nvidia is dominating

One of the first GPU neural nets was on a NVIDIA GTX 280 up to 9 layers neural network. (2010 Ciresan and Schmidhuber)

- Nvidia chips tend to outperform AMD
- More importantly, all the major frameworks use CUDA as first-class citizen. Poor support for AMD's OpenCL

NVIDIA®

Libraries – A ‘revolution’ in the making ?

Python For Data Science *Cheat Sheet*
SciPy - Linear Algebra
Learn More Python for Data Science Interactively at www.datacamp.com

SciPy
The SciPy library is one of the core packages for scientific computing that provides mathematical algorithms and convenience functions built on the NumPy extension of Python.

Interacting With NumPy
Also see NumPy

```
>>> import numpy as np
>>> a = np.array([1,2,3])
>>> b = np.array([(1+5j,2j,3j), (4j,5j,6j)])
>>> c = np.array([[(1.5,2,3), (4,5,6)], [(3,2,1), (4,5,6)]])
```

Index Tricks

`>>> np.mgrid[0:5,0:5]`	Create a dense meshgrid
`>>> np.ogrid[0:2,0:2]`	Create an open meshgrid
`>>> np.r_[3,[0]*5,-1:1:10j]`	Stack arrays vertically (row-wise)
`>>> np.c_[b,c]`	Create stacked column-wise arrays

Shape Manipulation

`>>> np.transpose(b)`	Permute array dimensions
`>>> b.flatten()`	Flatten the array
`>>> np.hstack((b,c))`	Stack arrays horizontally (column-wise)
`>>> np.vstack((a,b))`	Stack arrays vertically (row-wise)
`>>> np.hsplit(c,2)`	Split the array horizontally at the 2nd index
`>>> np.vpslit(d,2)`	Split the array vertically at the 2nd index

Polynomials

`>>> from numpy import poly1d`	
`>>> p = poly1d([3,4,5])`	Create a polynomial object

Vectorizing Functions

```
>>> def myfunc(a):
        if a < 0:
          return a*2
        else:
          return a/2
>>> np.vectorize(myfunc)
```
Vectorize functions

Type Handling

`>>> np.real(b)`	Return the real part of the array elements
`>>> np.imag(b)`	Return the imaginary part of the array elements
`>>> np.real_if_close(c,tol=1000)`	Return a real array if complex parts close to 0
`>>> np.cast['f'](np.pi)`	Cast object to a data type

Other Useful Functions

`>>> np.angle(b,deg=True)`	Return the angle of the complex argument
`>>> g = np.linspace(0,np.pi,num=5)`	Create an array of evenly spaced values (number of samples)
`>>> g [3:] += np.pi`	
`>>> np.unwrap(g)`	Unwrap
`>>> np.logspace(0,10,3)`	Create an array of evenly spaced values (log scale)
`>>> np.select([c<4],[c*2])`	Return values from a list of arrays depending on conditions
`>>> misc.factorial(a)`	Factorial
`>>> misc.comb(10,3,exact=True)`	Combine N things taken at k time
`>>> misc.central_diff_weights(3)`	Weights for Np-point central derivative
`>>> misc.derivative(myfunc,1.0)`	Find the n-th derivative of a function at a point

Linear Algebra
Also see NumPy

You'll use the `linalg` and `sparse` modules. Note that `scipy.linalg` contains and expands on `numpy.linalg`.

```
>>> from scipy import linalg, sparse
```

Creating Matrices

```
>>> A = np.matrix(np.random.random((2,2)))
>>> B = np.asmatrix(b)
>>> C = np.mat(np.random.random((10,5)))
>>> D = np.mat([[3,4], [5,6]])
```

Basic Matrix Routines

Inverse	
`>>> A.I`	Inverse
`>>> linalg.inv(A)`	Inverse
Transposition	
`>>> A.T`	Tranpose matrix
`>>> A.H`	Conjugate transposition
Trace	
`>>> np.trace(A)`	Trace
Norm	
`>>> linalg.norm(A)`	Frobenius norm
`>>> linalg.norm(A,1)`	L1 norm (max column sum)
`>>> linalg.norm(A,np.inf)`	L inf norm (max row sum)
Rank	
`>>> np.linalg.matrix_rank(C)`	Matrix rank
Determinant	
`>>> linalg.det(A)`	Determinant
Solving linear problems	
`>>> linalg.solve(A,b)`	Solver for dense matrices
`>>> E = np.mat(a).T`	Solver for dense matrices
`>>> linalg.lstsq(F,E)`	Least-squares solution to linear matrix equation
Generalized inverse	
`>>> linalg.pinv(C)`	Compute the pseudo-inverse of a matrix (least-squares solver)
`>>> linalg.pinv2(C)`	Compute the pseudo-inverse of a matrix (SVD)

Creating Sparse Matrices

`>>> F = np.eye(3, k=1)`	Create a 2X2 identity matrix
`>>> G = np.mat(np.identity(2))`	Create a 2x2 identity matrix
`>>> C[C > 0.5] = 0`	
`>>> H = sparse.csr_matrix(C)`	Compressed Sparse Row matrix
`>>> I = sparse.csc_matrix(D)`	Compressed Sparse Column matrix
`>>> J = sparse.dok_matrix(A)`	Dictionary Of Keys matrix
`>>> E.todense()`	Sparse matrix to full matrix
`>>> sparse.isspmatrix_csc(A)`	Identify sparse matrix

Sparse Matrix Routines

Inverse	
`>>> sparse.linalg.inv(I)`	Inverse
Norm	
`>>> sparse.linalg.norm(I)`	Norm
Solving linear problems	
`>>> sparse.linalg.spsolve(H,I)`	Solver for sparse matrices

Sparse Matrix Functions

`>>> sparse.linalg.expm(I)`	Sparse matrix exponential

Matrix Functions

Addition	
`>>> np.add(A,D)`	Addition
Subtraction	
`>>> np.subtract(A,D)`	Subtraction
Division	
`>>> np.divide(A,D)`	Division
Multiplication	
`>>> A @ D`	Multiplication operator (Python 3)
`>>> np.multiply(D,A)`	Multiplication
`>>> np.dot(A,D)`	Dot product
`>>> np.vdot(A,D)`	Vector dot product
`>>> np.inner(A,D)`	Inner product
`>>> np.outer(A,D)`	Outer product
`>>> np.tensordot(A,D)`	Tensor dot product
`>>> np.kron(A,D)`	Kronecker product
Exponential Functions	
`>>> linalg.expm(A)`	Matrix exponential
`>>> linalg.expm2(A)`	Matrix exponential (Taylor Series)
`>>> linalg.expm3(D)`	Matrix exponential (eigenvalue decomposition)
Logarithm Function	
`>>> linalg.logm(A)`	Matrix logarithm
Trigonometric Functions	
`>>> linalg.sinm(D)`	Matrix sine
`>>> linalg.cosm(D)`	Matrix cosine
`>>> linalg.tanm(A)`	Matrix tangent
Hyperbolic Trigonometric Functions	
`>>> linalg.sinhm(D)`	Hypberbolic matrix sine
`>>> linalg.coshm(D)`	Hyperbolic matrix cosine
`>>> linalg.tanhm(A)`	Hyperbolic matrix tangent
Matrix Sign Function	
`>>> np.signm(A)`	Matrix sign function
Matrix Square Root	
`>>> linalg.sqrtm(A)`	Matrix square root
Arbitrary Functions	
`>>> linalg.funm(A, lambda x: x*x)`	Evaluate matrix function

Decompositions

Eigenvalues and Eigenvectors	
`>>> la, v = linalg.eig(A)`	Solve ordinary or generalized eigenvalue problem for square matrix
`>>> l1, l2 = la`	Unpack eigenvalues
`>>> v[:,0]`	First eigenvector
`>>> v[:,1]`	Second eigenvector
`>>> linalg.eigvals(A)`	Unpack eigenvalues
Singular Value Decomposition	
`>>> U,s,Vh = linalg.svd(B)`	Singular Value Decomposition (SVD)
`>>> M,N = B.shape`	
`>>> Sig = linalg.diagsvd(s,M,N)`	Construct sigma matrix in SVD
LU Decomposition	
`>>> P,L,U = linalg.lu(C)`	LU Decomposition

Sparse Matrix Decompositions

`>>> la, v = sparse.linalg.eigs(F,1)`	Eigenvalues and eigenvectors
`>>> sparse.linalg.svds(H, 2)`	SVD

Asking For Help

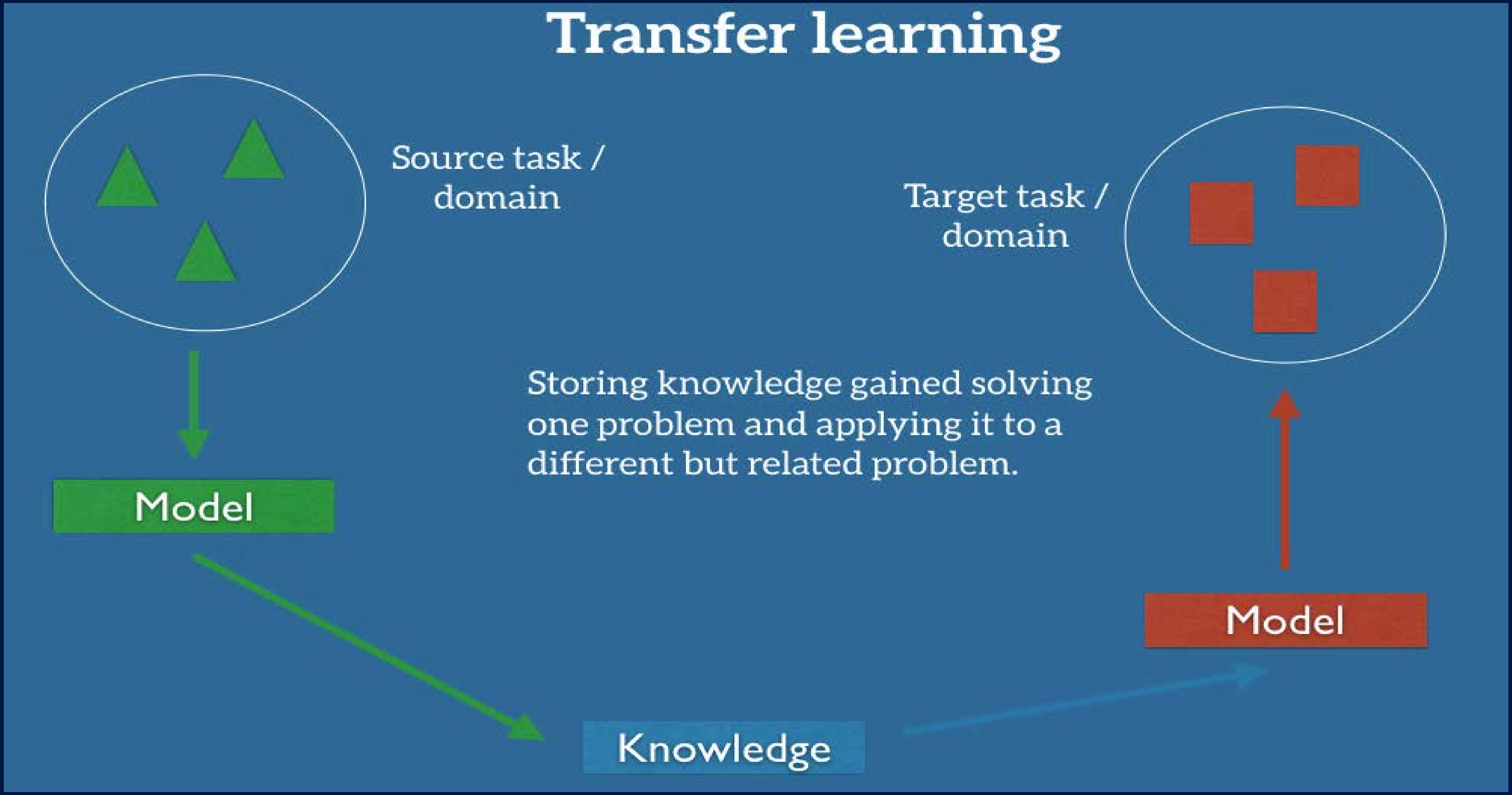
```
>>> help(scipy.linalg.diagsvd)
>>> np.info(np.matrix)
```

DataCamp
Learn Python for Data Science Interactively

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Deep Learning → Real World Problems



Neural Nets - Challenges

a young boy is holding a baseball bat

Statistically impressive, but individually unreliable

"Deep Visual-Semantic Alignments for Generating Image Descriptions" by Andrej Karpathy, Li Fei-Fei (CVPR 2015).

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Neural Nets - Challenges

Conclusion: Inherent flaws can be exploited

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Neural Nets - Challenges

Internet trolls cause the AI bot, Tay,
to act offensively

Skewed training data creates maladaptation

On Deep Neural Networks

Bored Yann LeCun
@boredyannlecun

Following

Spend the holidays with your family, not reading arXiv papers. You'd be wasting your time anyway because—2019 spoiler alert—*CONVOLUTION IS ALL YOU NEED!* #torched #feelthelearn #Christvolution

7:10 PM - 23 Dec 2018

Outline

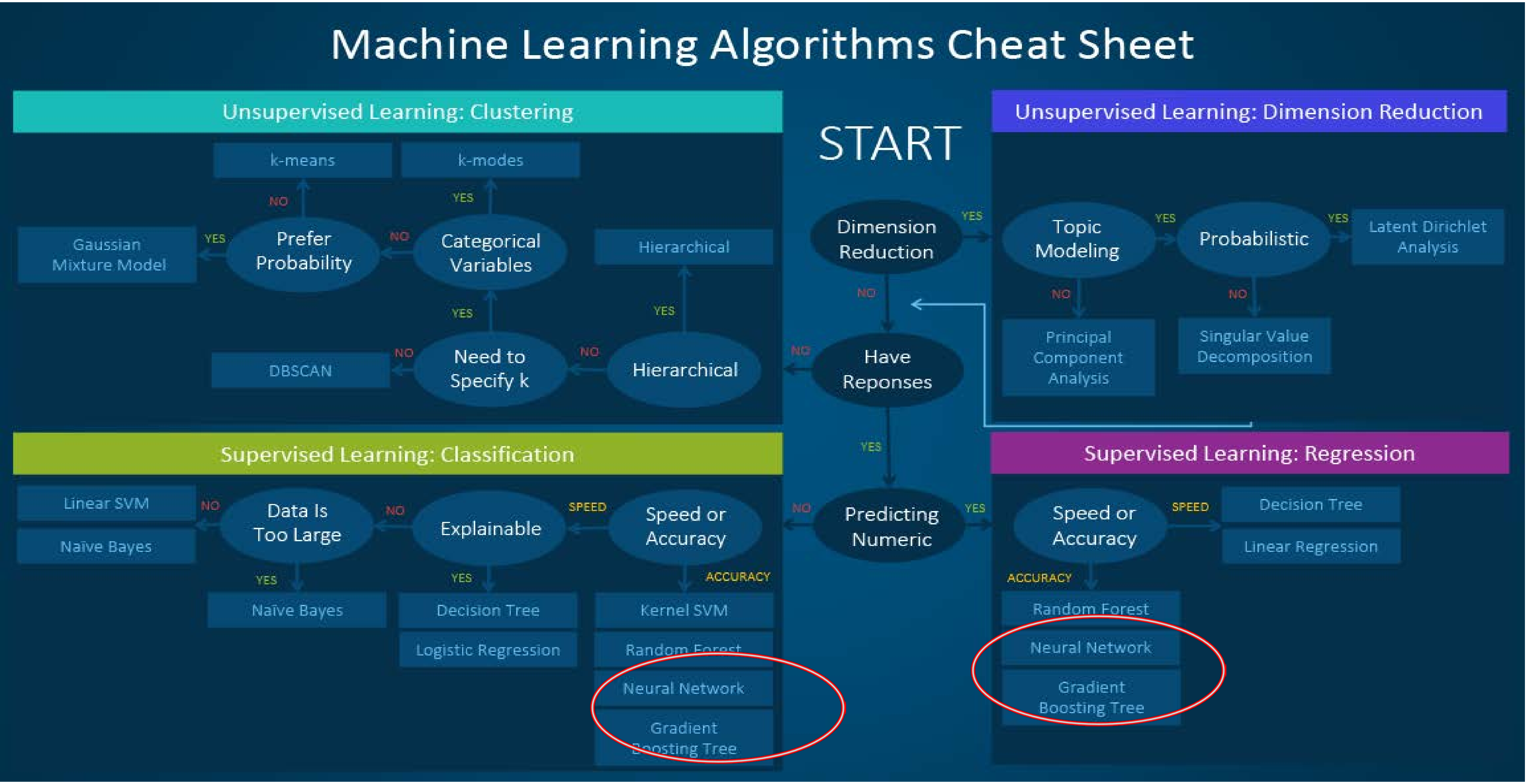
- A definition (or two)
- AI/ML: The big picture
- Altum Visum on deep learning networks
- **Machine Learning: Myths & Realities**
- Machine Learning as a process
- Long Term View: Explainable Artificial Intelligence (XAI)
- Long Term View: Human-like intelligence
- Epilogue

Myth

Machine Learning = Deep Neural Networks

Quiz question: When the term “A.I. Winter” was invented and why ?

Reality



Credit: Hui Li, SAS Analytics, April 2017

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Outline

- A definition (or two)
- AI/ML: The big picture
- Altum Visum on deep learning networks
- Machine Learning: Myths & Realities
- **Machine Learning as a process**
- Long Term View: Explainable Artificial Intelligence (XAI)
- Long Term View: Human-like intelligence
- Epilogue

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

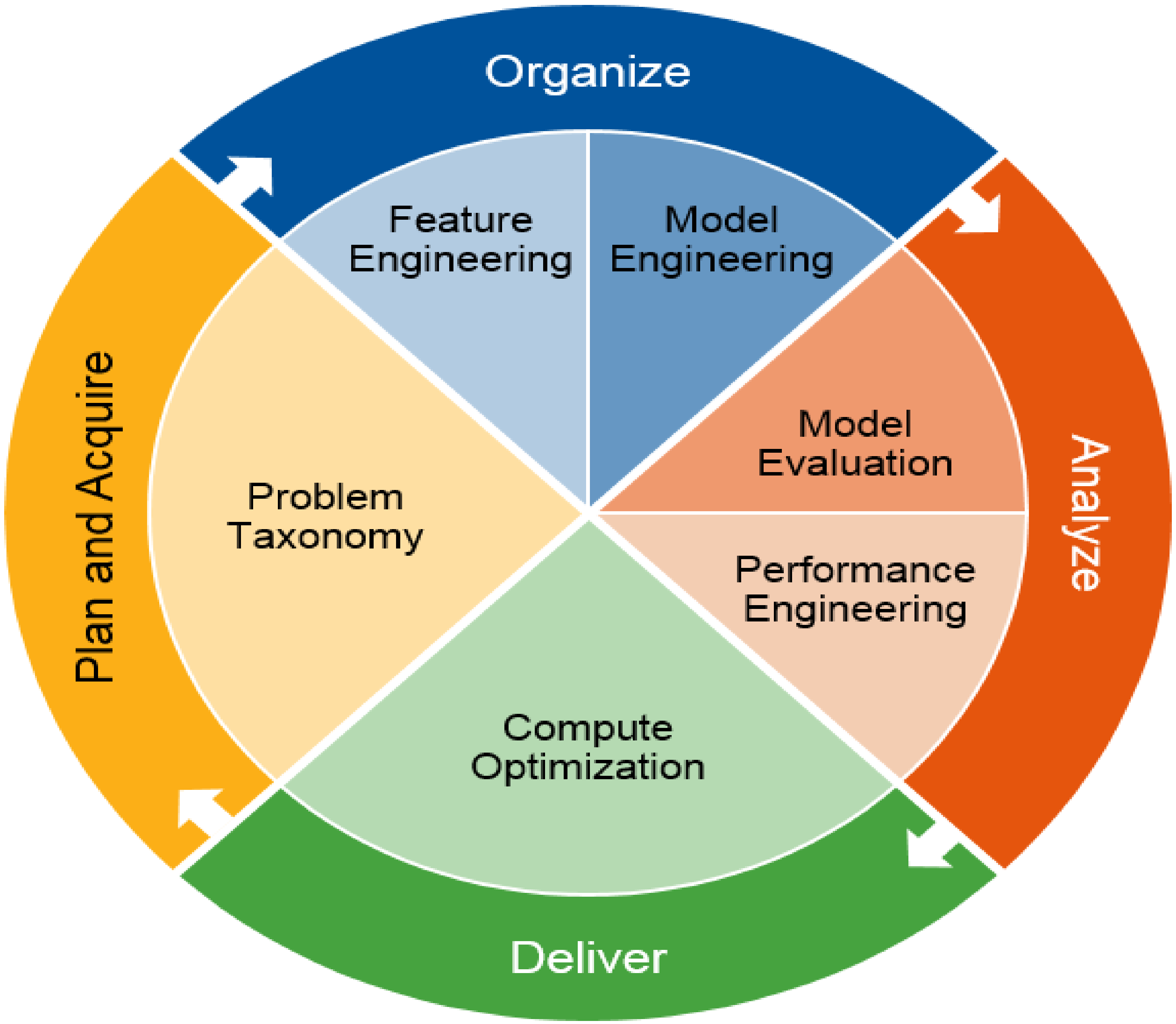
A definition revised:

"Machine learning (ML): a subset of artificial intelligence (AI) is more than a technique for analyzing data. It's a system that is fueled by data, with the ability to learn and improve by using algorithms that provide new insights without being explicitly programmed to do so."

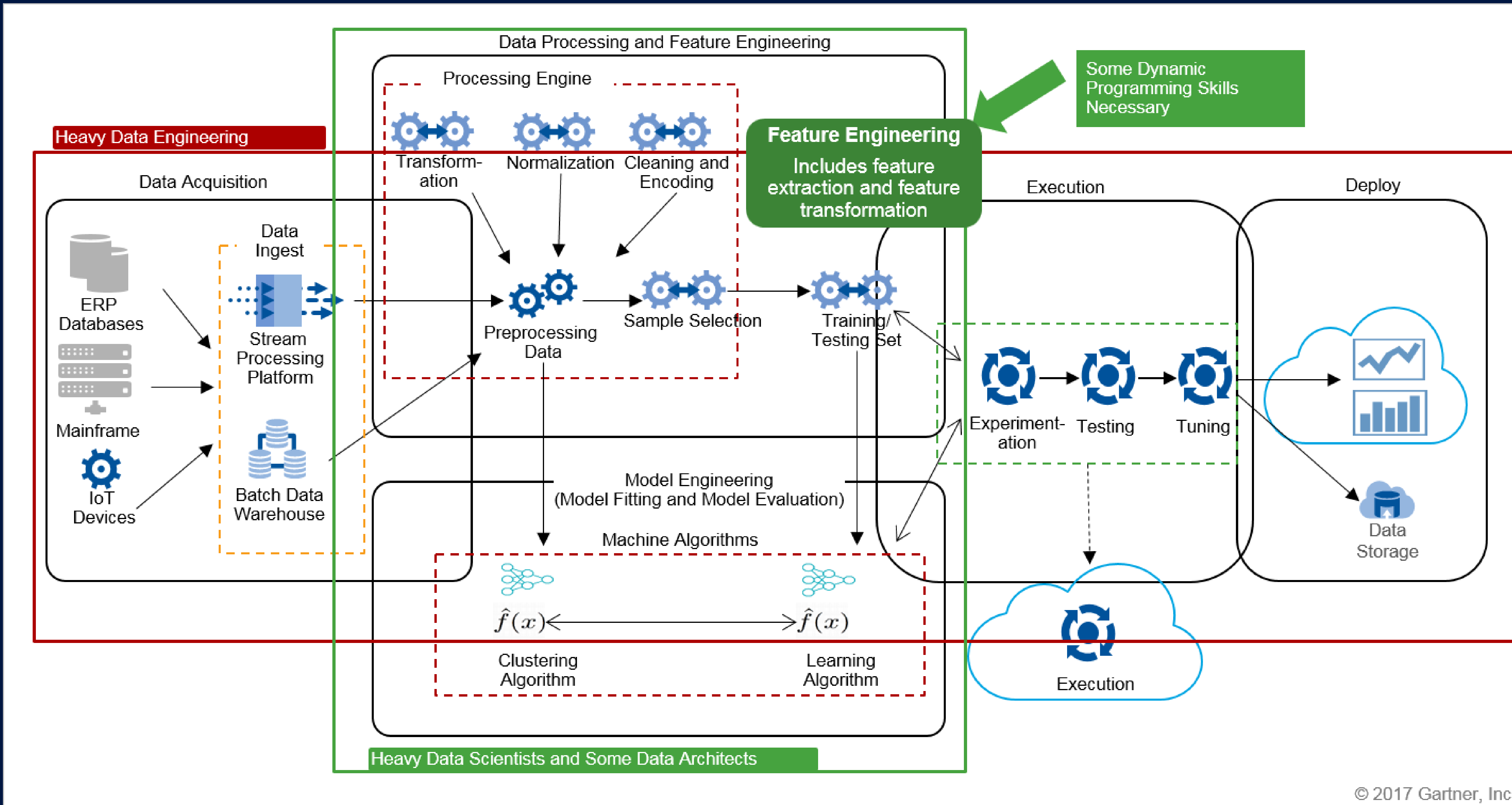
Gartner, "Preparing and Architecting for Machine Learning", Technical Professional Advice, published January 17, 2017.

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

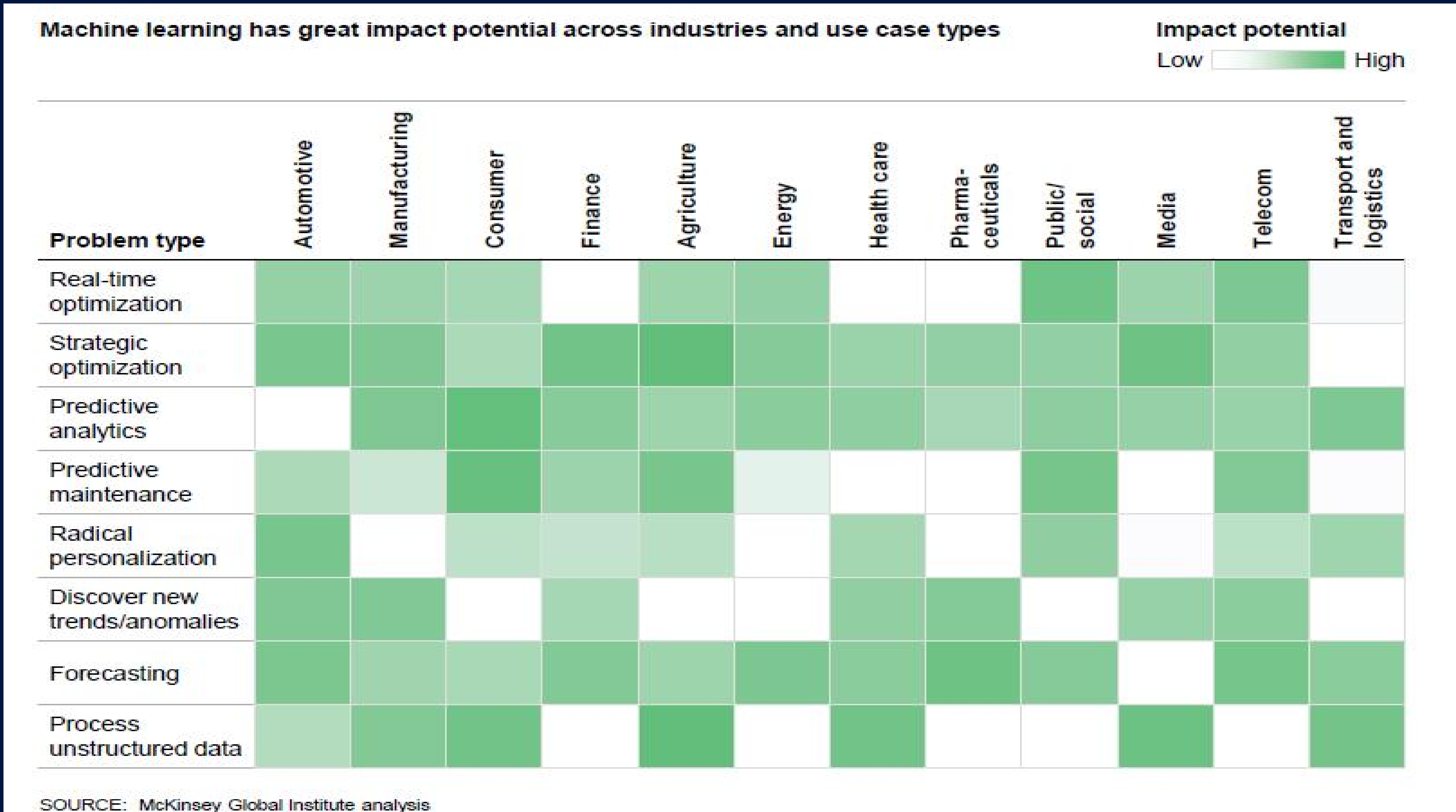
Engineering Cycle of Machine Learning



Machine Learning – Skills Set Requirements



What is the expected impact & where



Outline

- A definition (or two)
- AI/ML: The big picture
- Altum Visum on deep learning networks
- Machine Learning: Myths & Realities
- Machine Learning as a process
- **Long Term View: Explainable Artificial Intelligence (XAI)**
- Long Term View: Human-like intelligence
- Epilogue

Big Picture

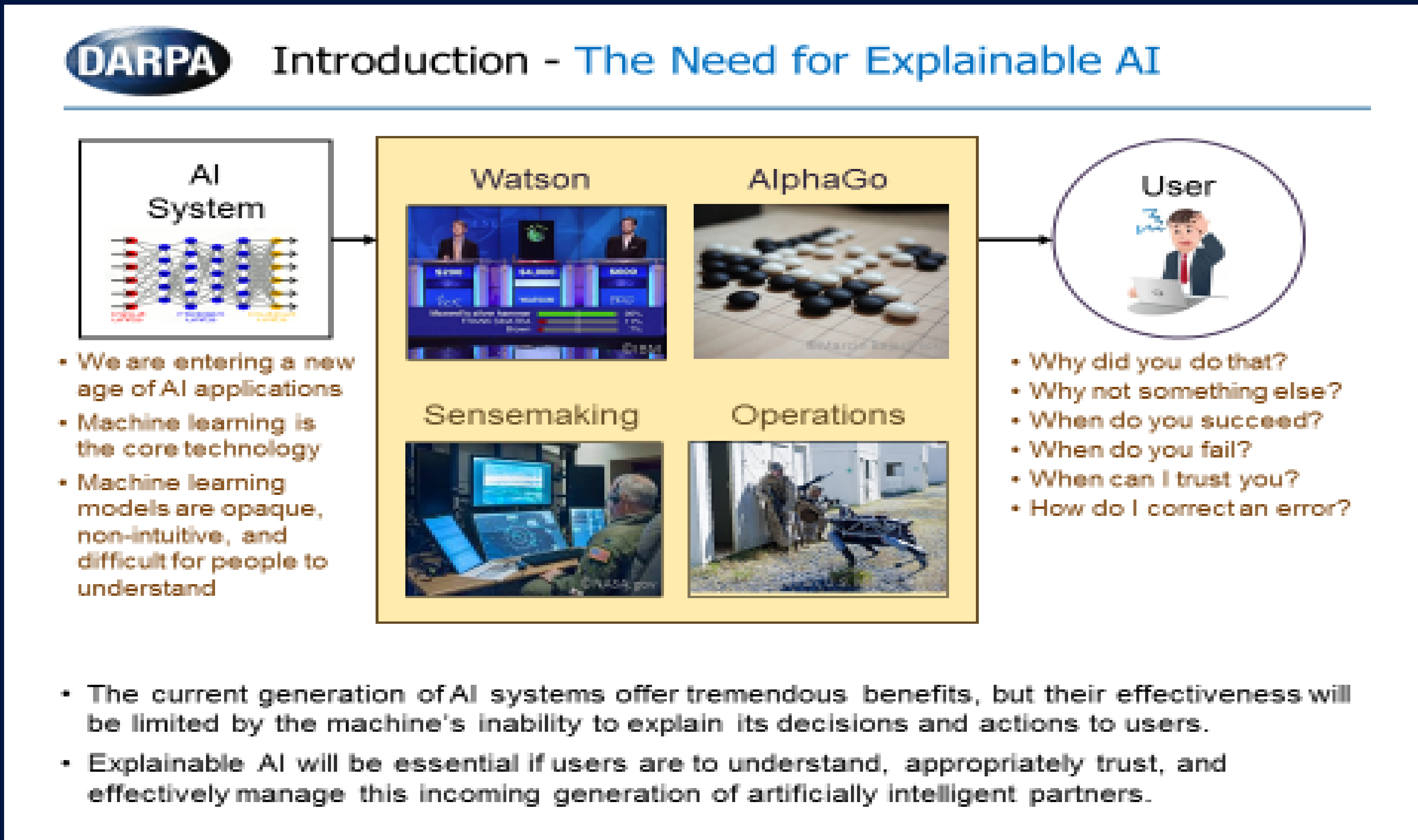
Long Term Research In Machine Learning [1]

- Understanding theoretical capabilities & limitations
- Developing scalable systems
- Pursuing research on general-purpose artificial intelligence
- Developing more capable and reliable robots
- Advance hardware for improved AI/Creating AI for improved hardware
- Fostering research on human-like AI
- Improving fairness, transparency, and accountability by design

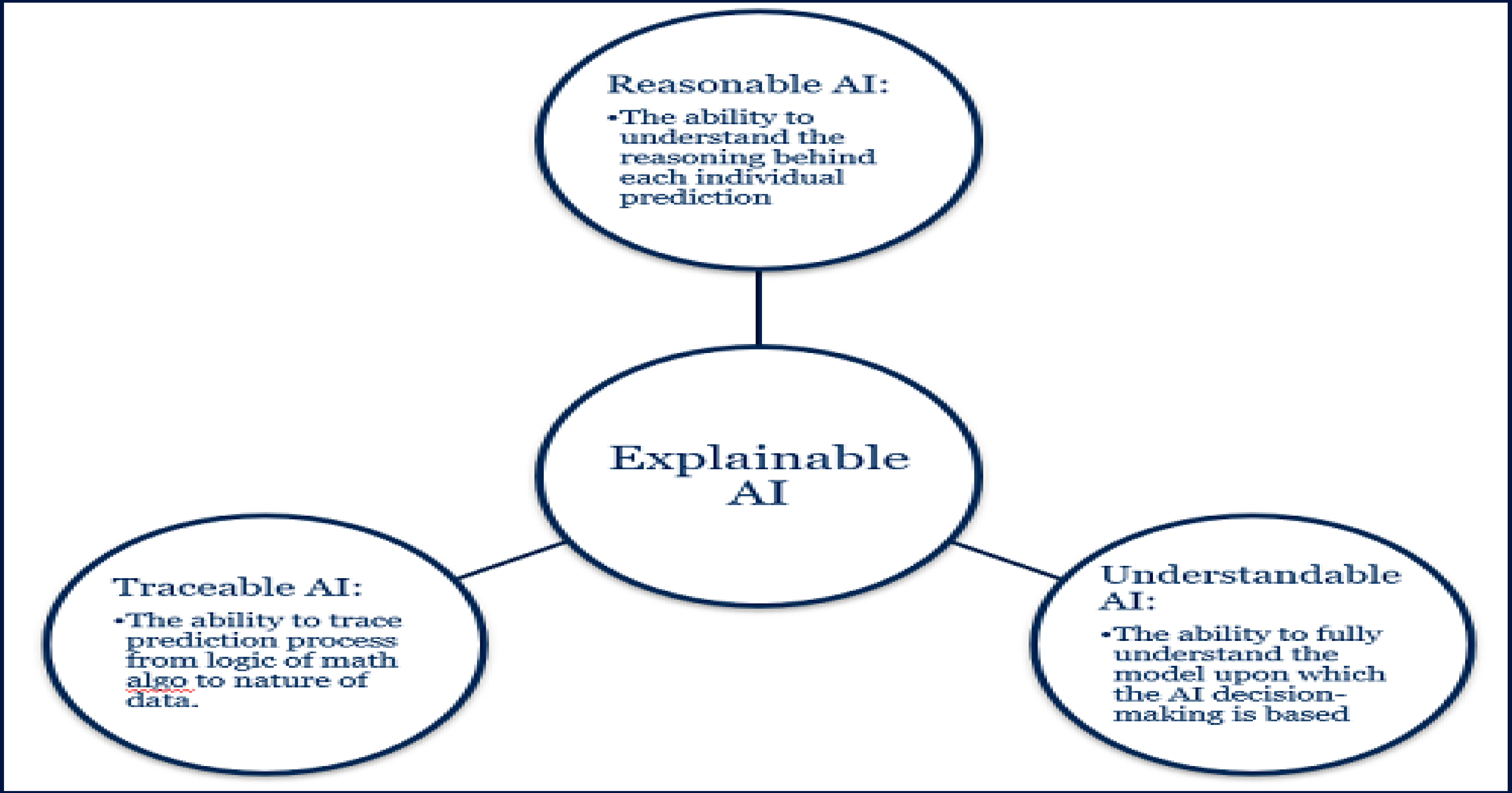
[1] The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update, A report by the Selected Committee on Artificial Intelligence of the National Science & Technology Council, Michael Katsios, June 2019.

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Explainable Artificial Intelligence



Explainable AI in Medical Applications

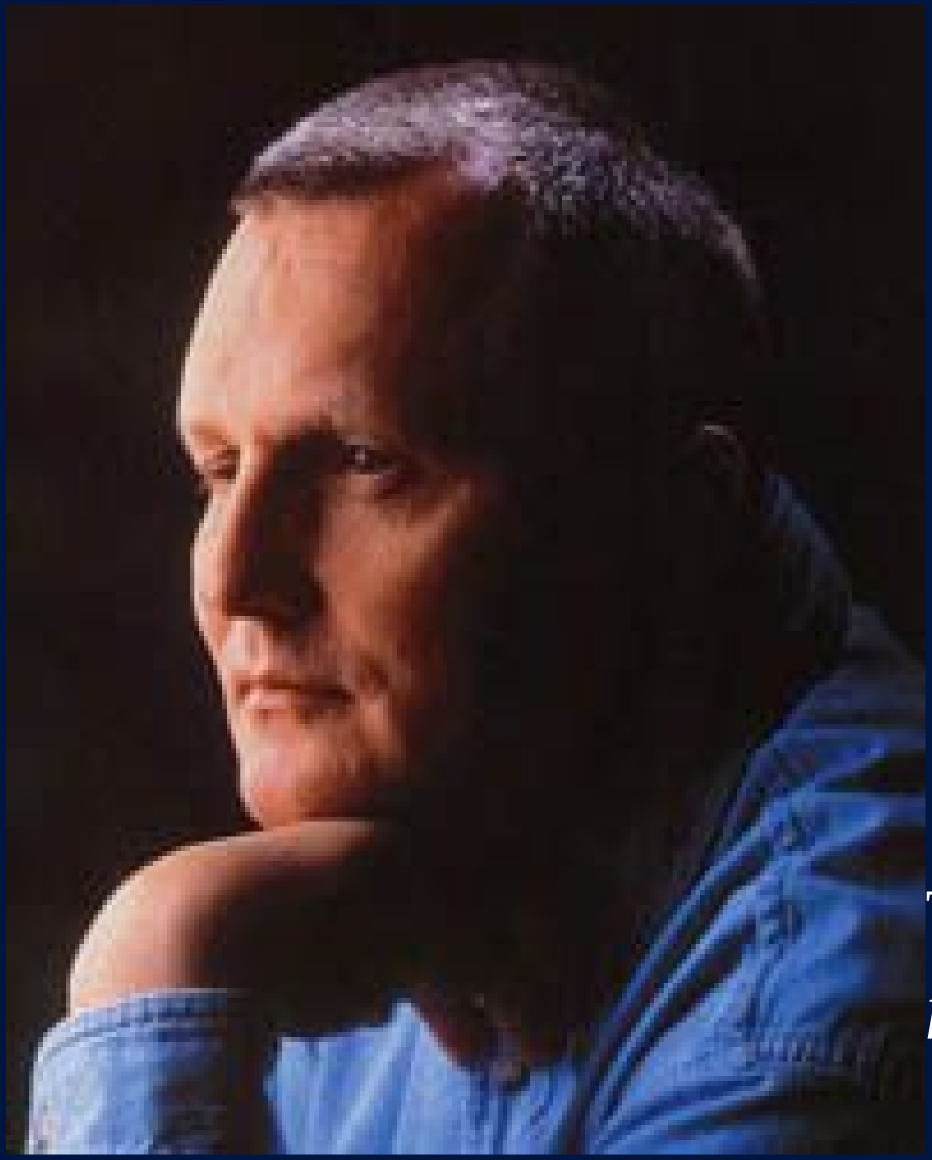


Credit: Saurabh Kaushik

Outline

- A definition (or two)
- AI/ML: The big picture
- Altum Visum on deep learning networks
- Machine Learning: Myths & Realities
- Machine Learning as a process
- Long Term View: Explainable Artificial Intelligence (XAI)
- **Long Term View: Human-like intelligence**
- Epilogue

An old (?) paradox



The Moravec's Paradox (1988): "it is comparatively easy to make computers exhibit adult level performance on intelligent tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility" . [1]

The paradox is sometimes simplified by the phrase: *Robots find the difficult things easy and the easy things difficult.*

[1] Hans Moravec, ***Mind Children:*** **The Future of Robot and Human Intelligence, Harvard University Press, 1988,** (**ISBN 0674576187**).

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Human Intelligence Characteristics

Fostering research on human-like AI [1]

- Efficient use of resources: meta-reasoning
- Efficient use of data: meta-learning

Meta-reasoning: human meta-cognition/active learning: awareness of one's own internal states; accuracy of memory; confidence in judgment; reasoning intelligently on how to collect information; intelligently re-use elements of cognitive and motor skills.

Meta-learning: efficient use of data; leverage commonalities across tasks that all have a similar character.

T.L. Griffiths et al, Doing more with less: meta-reasoning and meta-learning in humans and machines, Current Opinion in Behavioral Sciences, vil. 29, pp. 24-30, 2019.

Outline

- A definition (or two)
- AI/ML: The big picture
- Altum Visum on deep learning networks
- Machine Learning: Myths & Realities
- Machine Learning as a process
- Long Term View: Explainable Artificial Intelligence (XAI)
- Long Term View: Human-like intelligence
- **Epilogue**

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Big Picture

Is DNN (or ML in general) a "Deus ex Machina Moment" ?

Epilogue

- Machine learning is best-suited for dealing with big, **albeit curated,** data.
- Supervised networks (DNN) can learn semantically relevant representations useful in areas such as (image) classification, content-aware advertising, content filtering, social networks.
- Preparing data for Machine Learning pipelines is challenging.
- Machine Learning implies “learning” – the ability to generalize from experience – not yet there.

Thank you!

kostas@ece.utoronto.ca

www.dsp.utoronto.ca