“Artificial Intelligence” and Data Science

statistics and lots of data

Brian D Goodwin, PhD

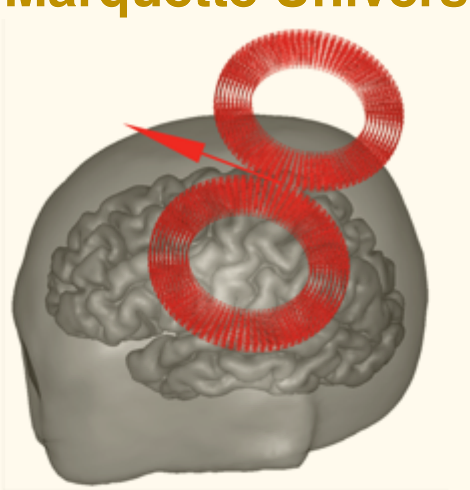
January 13, 2020

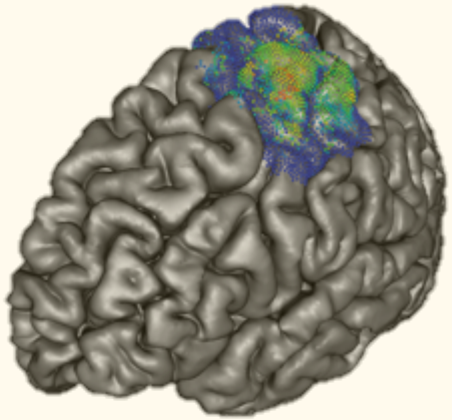
Milwaukee School of Engineering

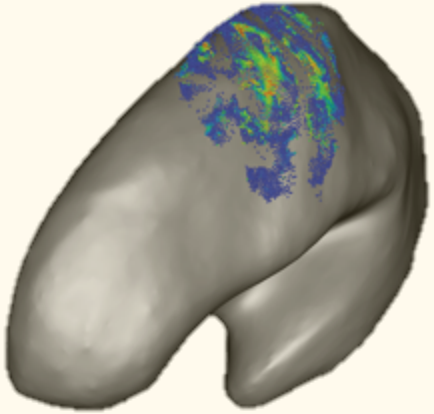
Image: https://bernardmarr.com/

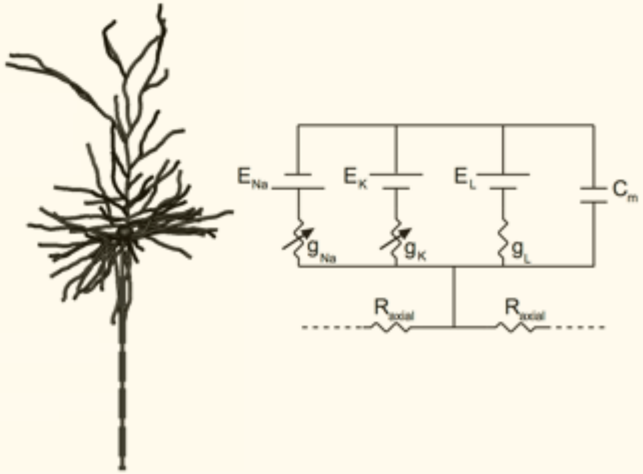
Brian D Goodwin, PhD

Marquette University



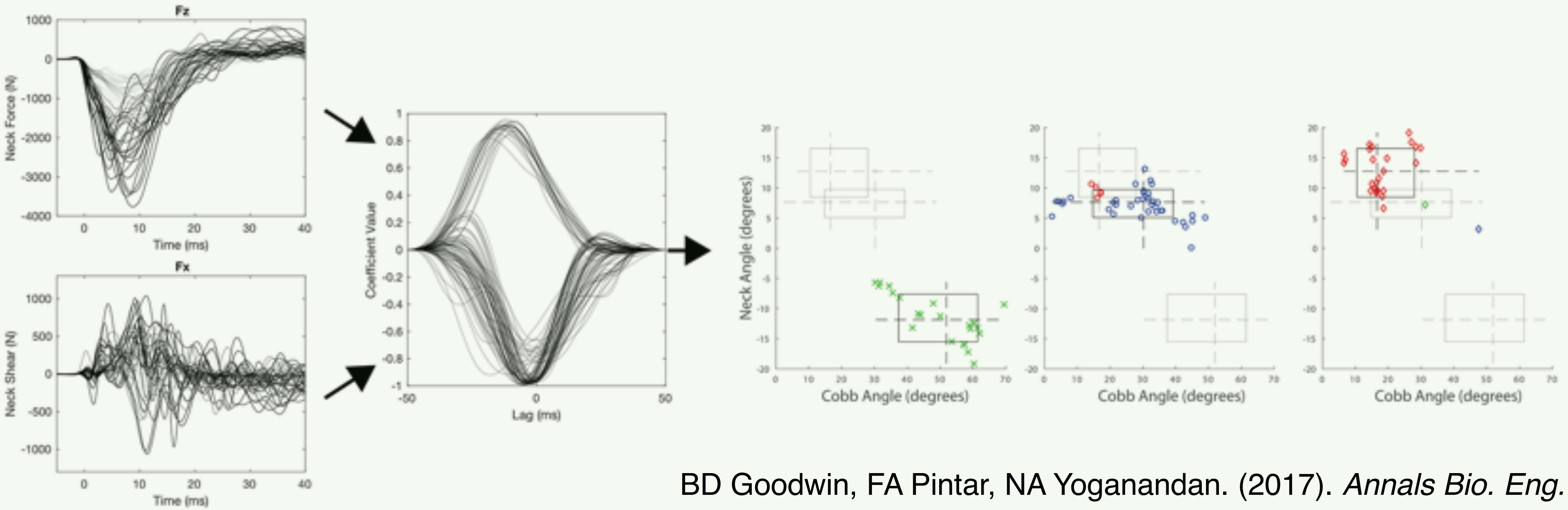






BD Goodwin, CR Butson. (2015). *J Neuromodulation*

Medical College of Wisconsin



BD Goodwin, FA Pintar, NA Yoganandan. (2017). *Annals Bio. Eng.*

There are many paths, but here is mine

- Undergrad: BS Mech. Eng. at MSOE – 2009
  - Coding/programming
  - Control systems
  - Finite element modeling (or FEM; also, FEA)
- Started a Masters in Biomedical Eng. at MU …
- Decided to just do a PhD about 6 months into my Masters
  - PhD took 5.5 years
  - Specialized in neuroscience and computational science
- Stayed in research to do a postdoctoral fellowship at MCW
  - Fellowship was about 2 year (could have been longer)
  - Specialized in biomechanics, signal processing, and machine-learning
- Transitioned into industry at IT consulting firm in the “Data & AI” space
  - 3 years
  - Specialized in cloud solution architecture, machine-learning, and optimization
- Recently accepted an offer from an AI company, Synthet^{ai}c
  - Using AI to generate synthetic datasets: large datasets that represent and characterize rare events

A few thoughts on mechanical engineering

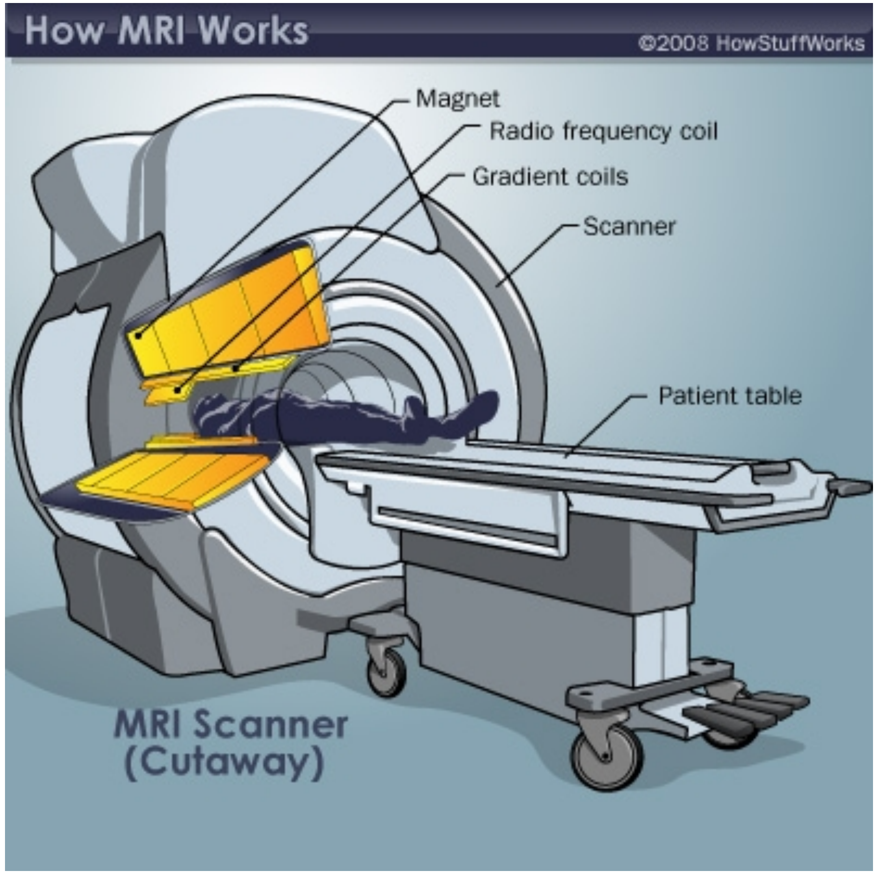
- (I'm obviously biased, but) I think a BSME sets you up for success
- Examples: the PI's I worked under and the CTO at Synthet^{ai}c
- Problems with "condensed concentrations/majors” (in my opinion)
    - Biomedical Engineering
        - Mechanical Eng.
        - Electrical Eng.
        - Systems Physiology
        - Materials Science & Biocompatilibity
    - Data Science
        - Computer Science
        - Statistics
        - Signal processing
        - Mathematics and Lin. Alg.
        - General Engineering
        - Data Engineering

What is potentially ahead for you

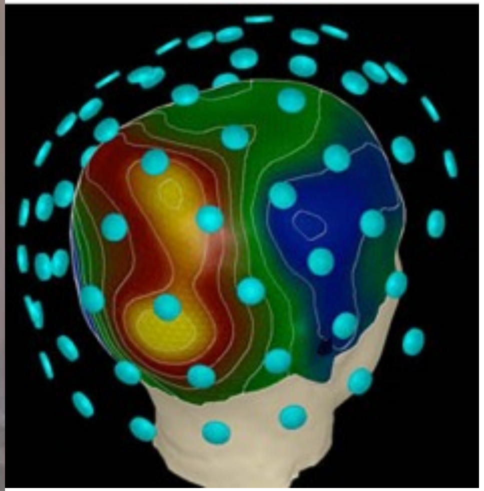
- Industry
- Academia
- How do I get the position I want?
  - Competence, but many of the competencies needed for a given job (especially your first one) are learned *on the job*.
  - At the end of the day, stories and relevant/novel contributions will land you the type of engineering job you want
  - Craft your stories accordingly
  - Question asking!

Transcranial Magnetic Stimulation and Magnetism in Medicine

(Nuclear) Magnetic Resonance Imaging

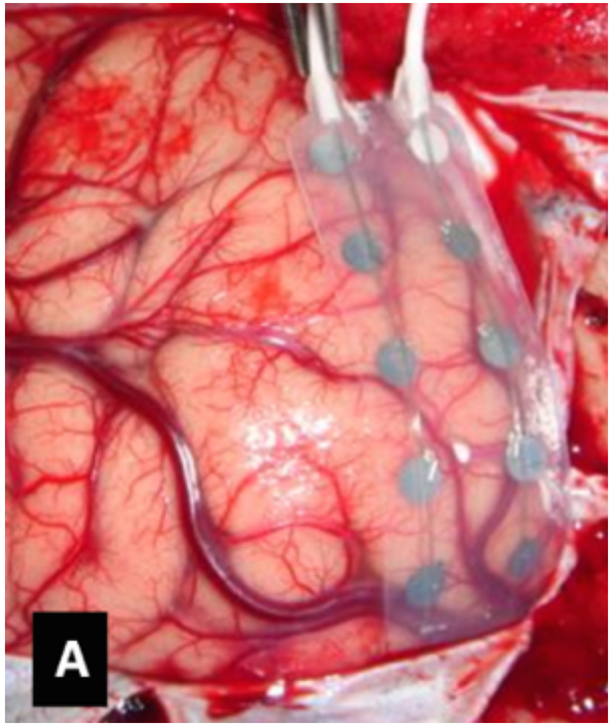


Magnetoencephalography (MEG)

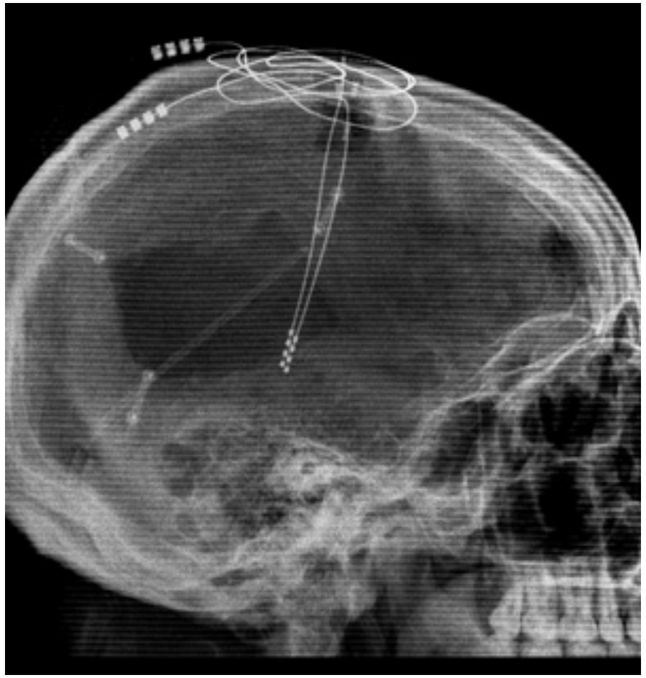


Transcranial Magnetic Stimulation

ECS

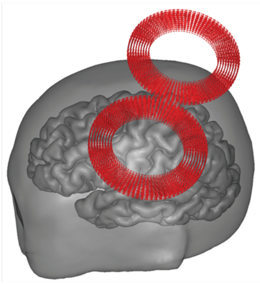


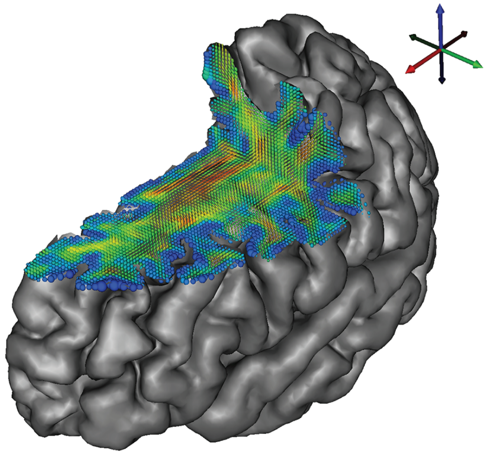
DBS

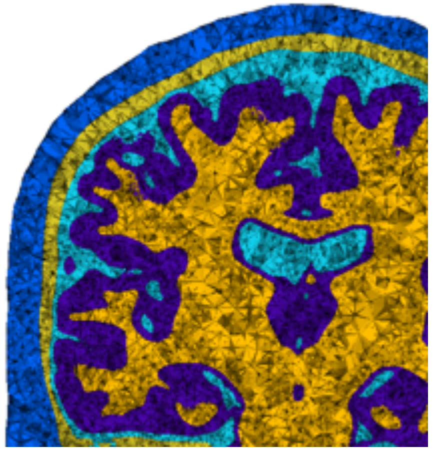


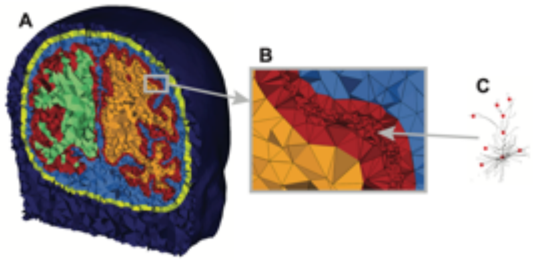
TMS

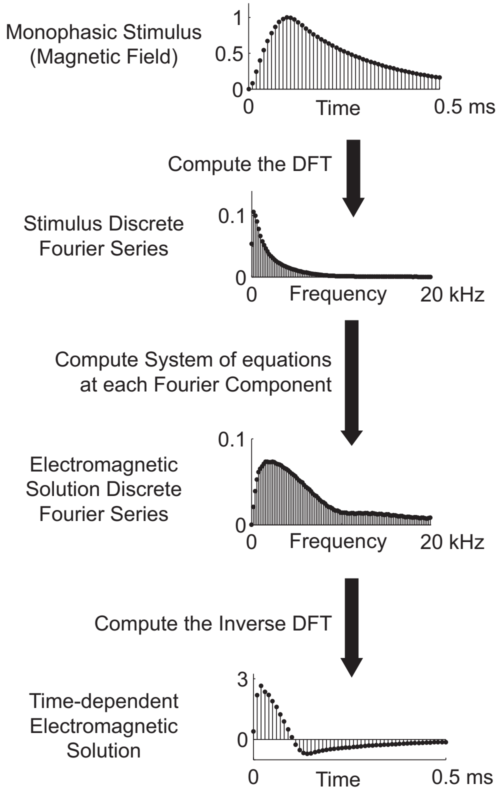
Idris Z, et al. 2013 (Chapter 2, “Clinical Management and Evolving Novel…”, Lichtor T)

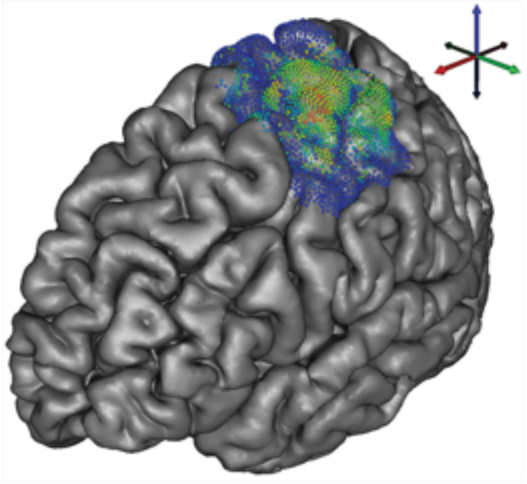








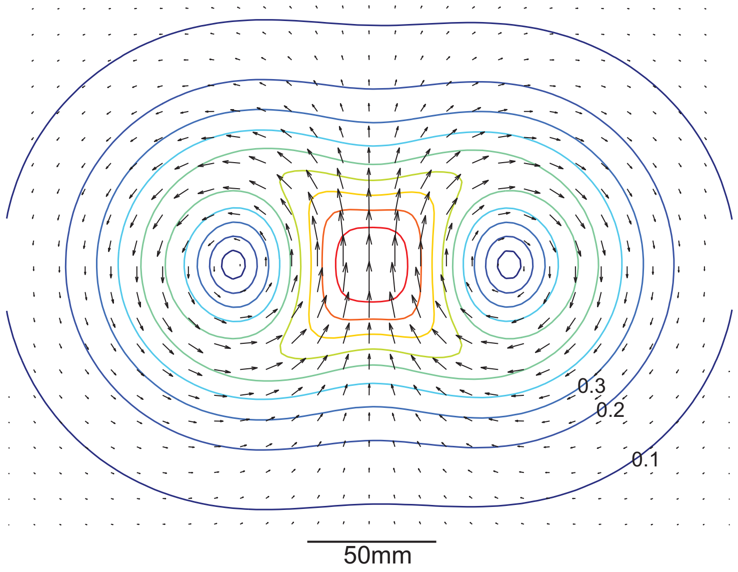


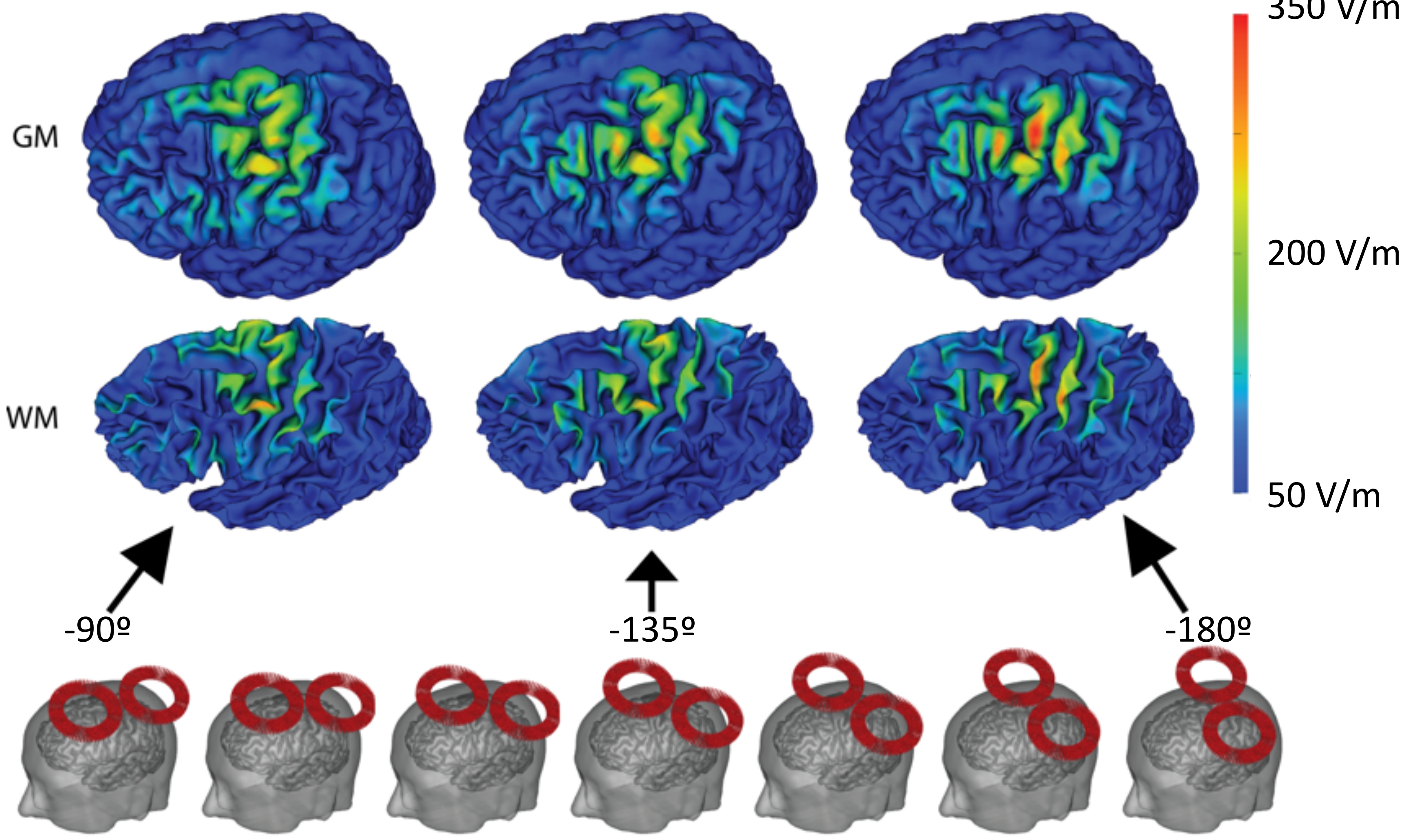


Modeling Pipeline for Prescriptive Therapy

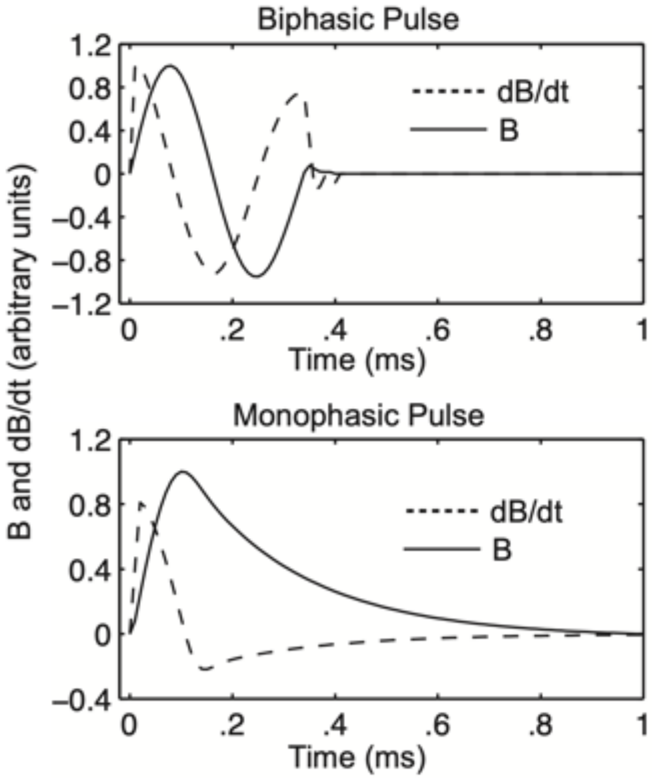
Goodwin, Butson. “Subject-Specific Multiscale Modeling...” *Neuromodulation* (2015).

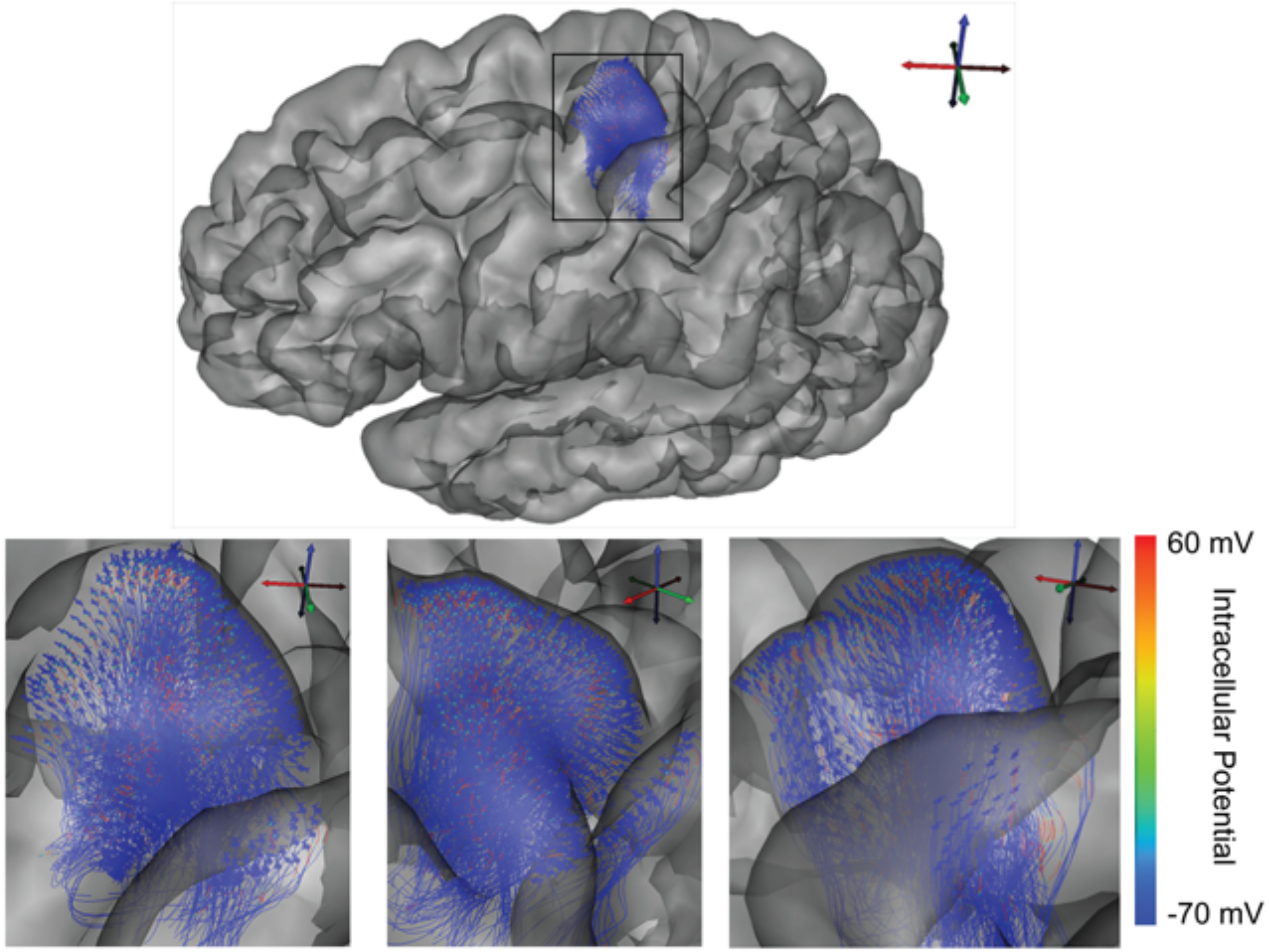
E-field and Brain Stimulation



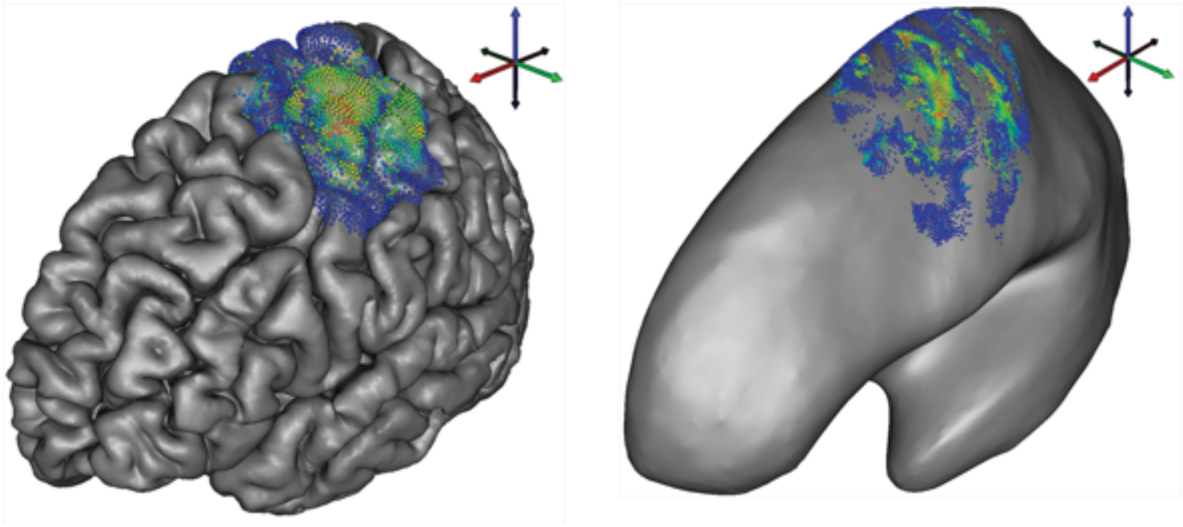


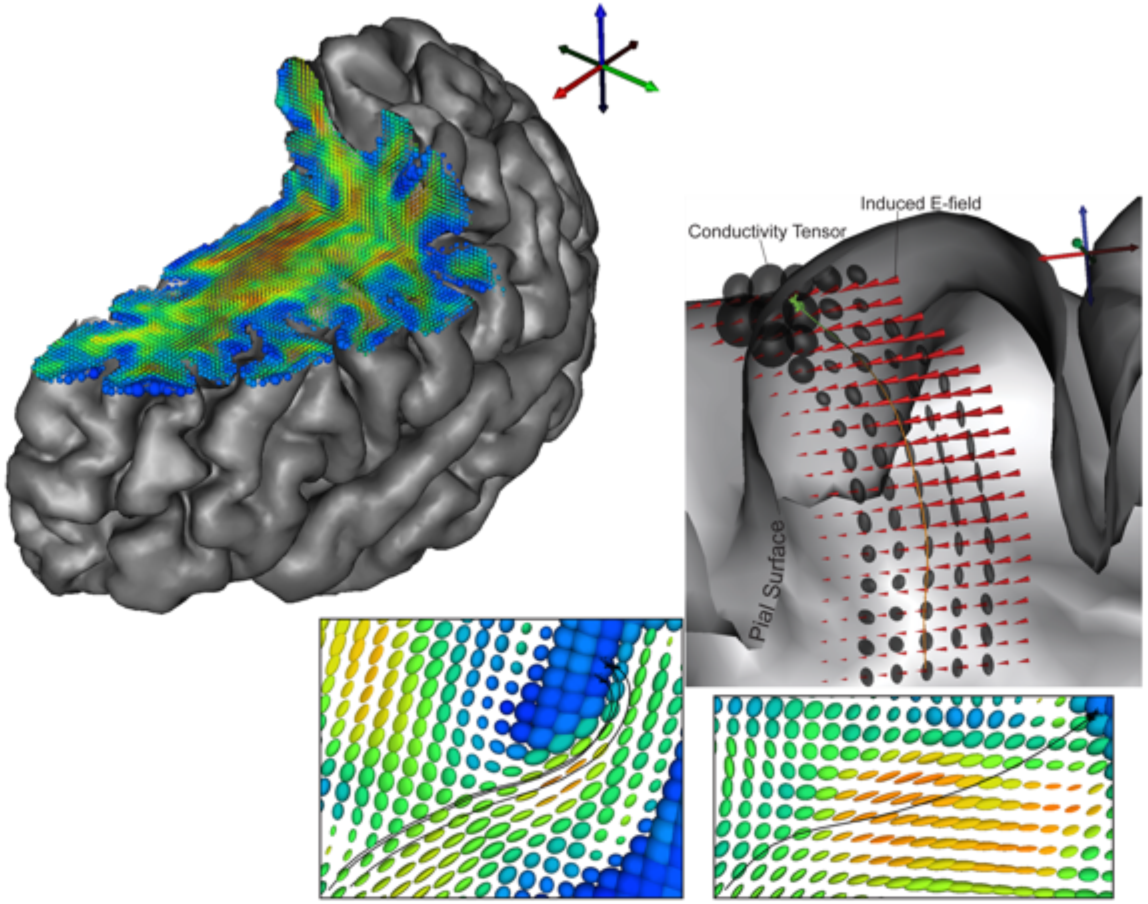
Computational Neuroscience and HPC



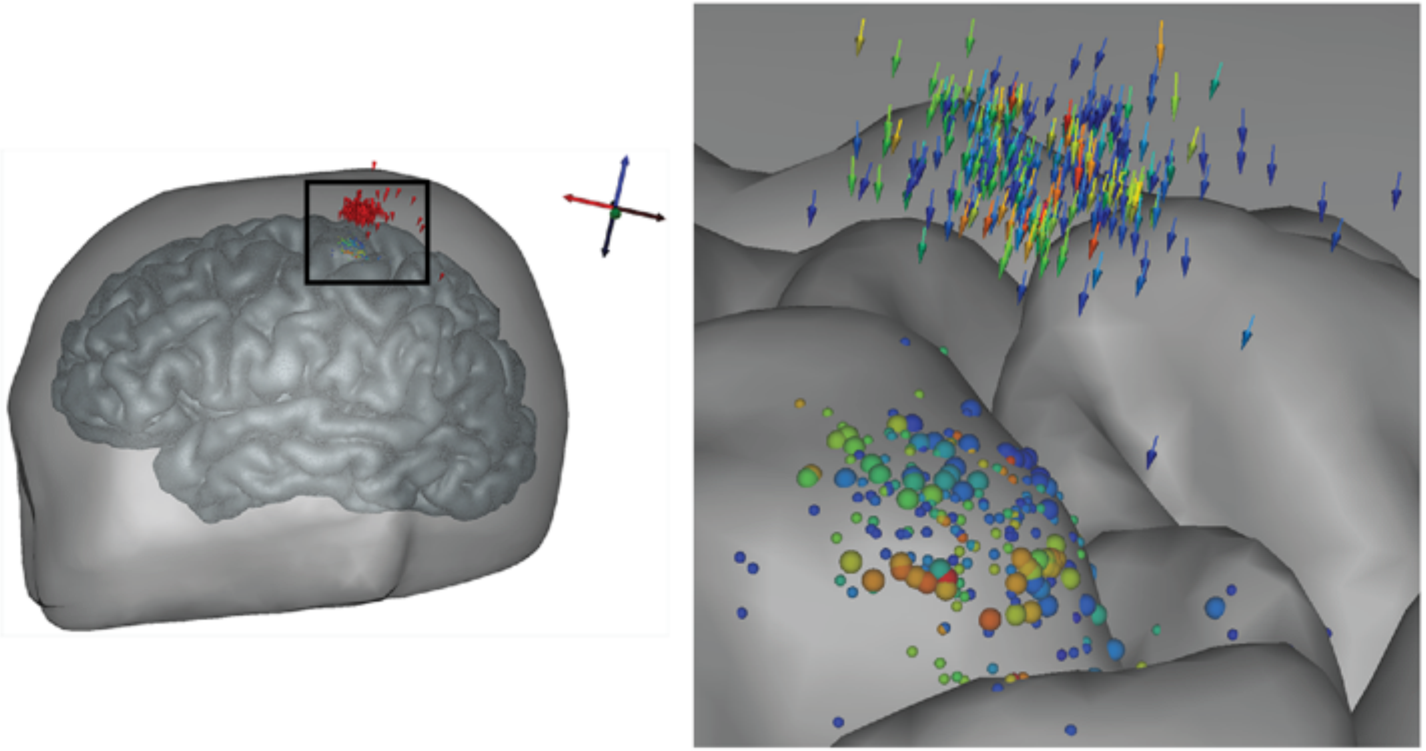


Patient-specific Modeling Pipeline Results



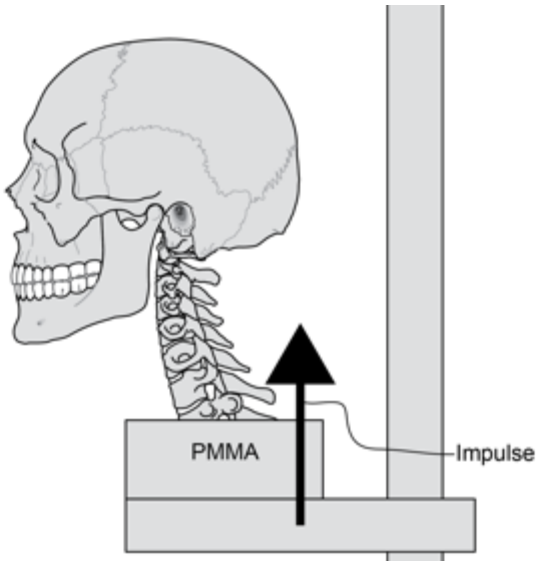


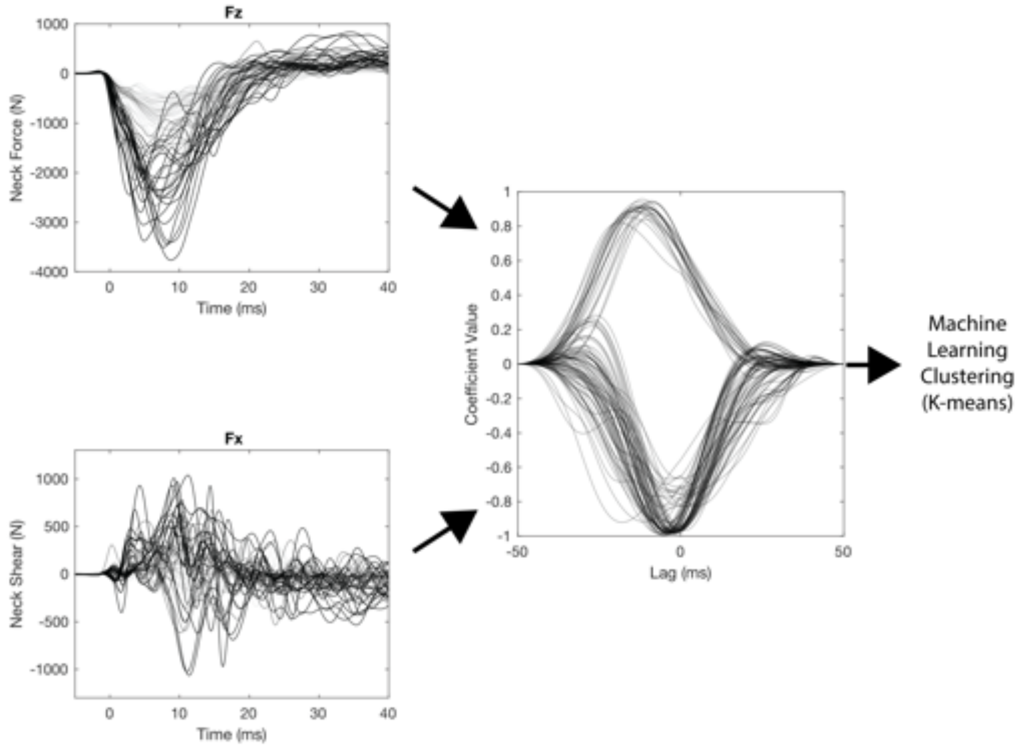
Human Subject Comparison



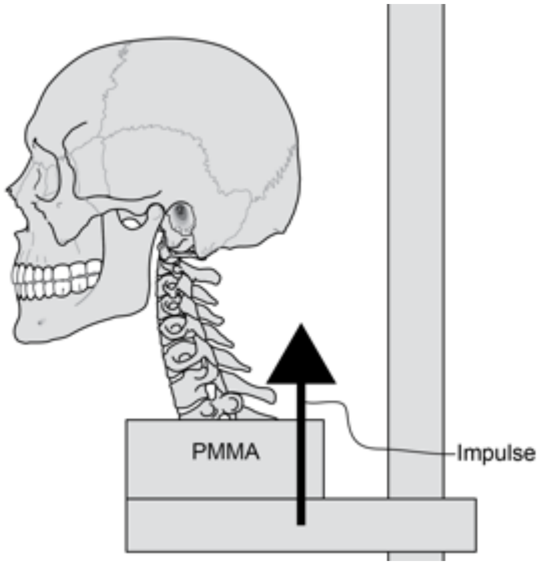
Goodwin, Marquette University (2015)

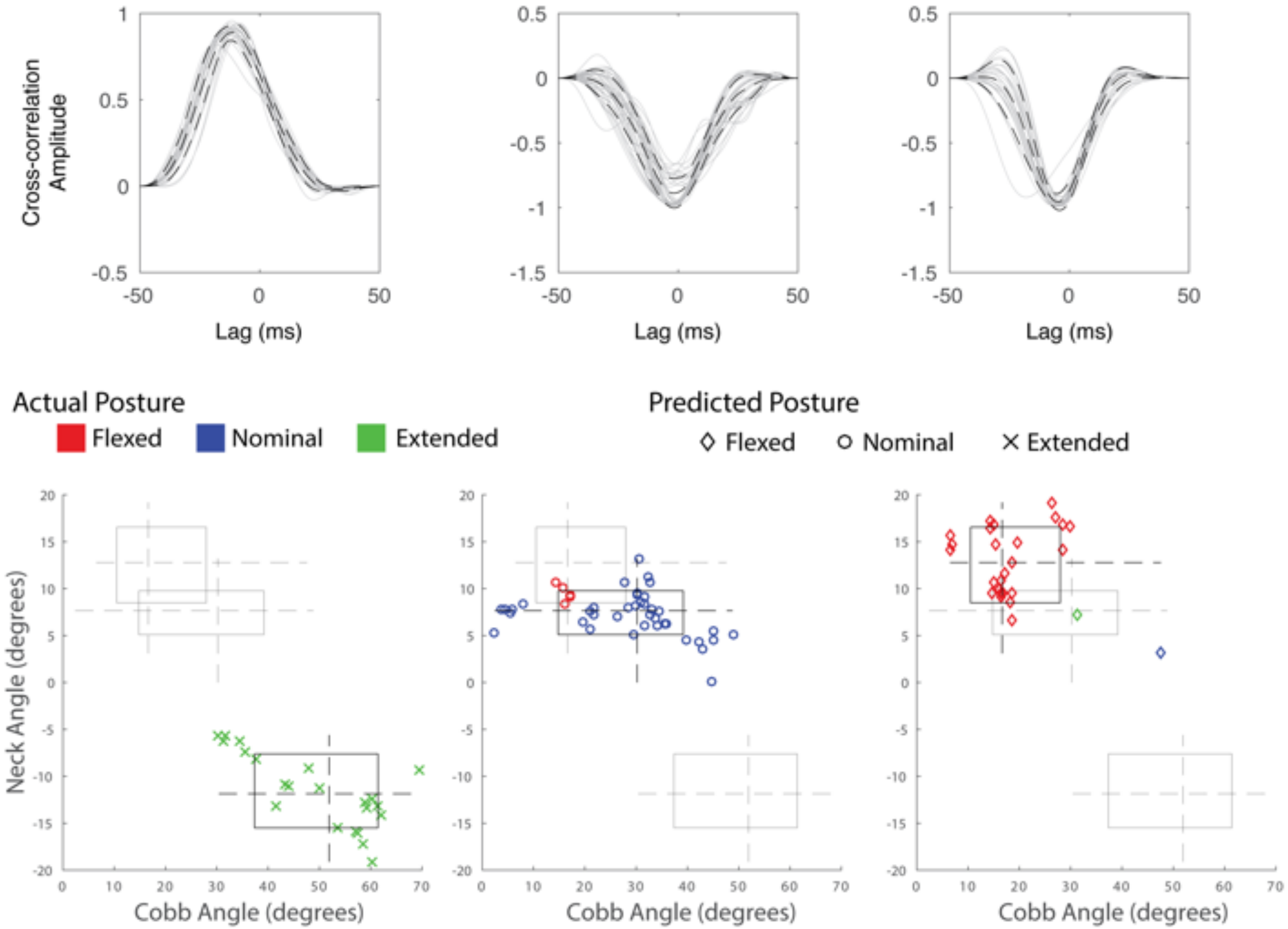
Biomechanics and Multi-channel HF Signals



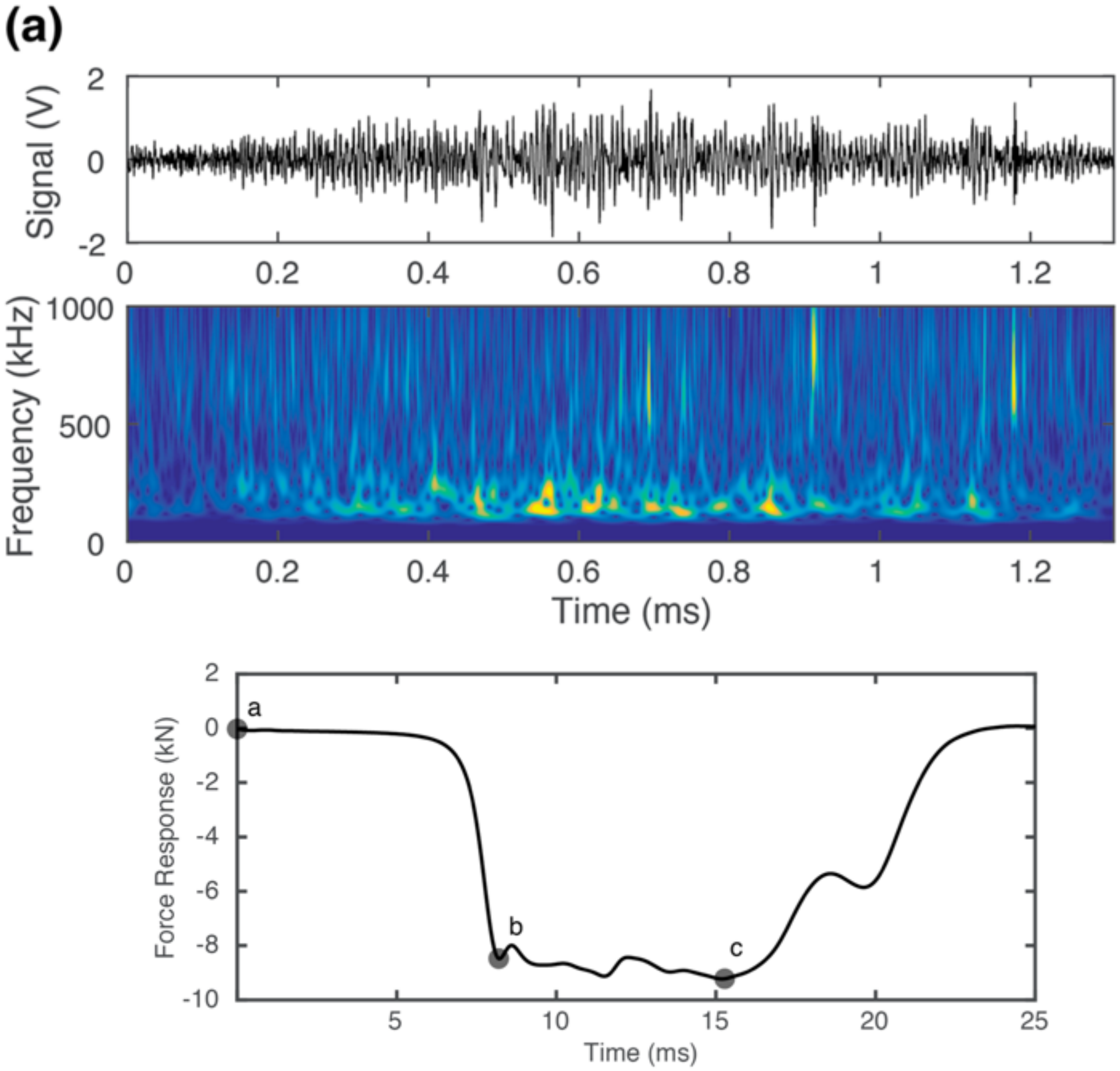


Biomechanics and Multi-channel HF Signals





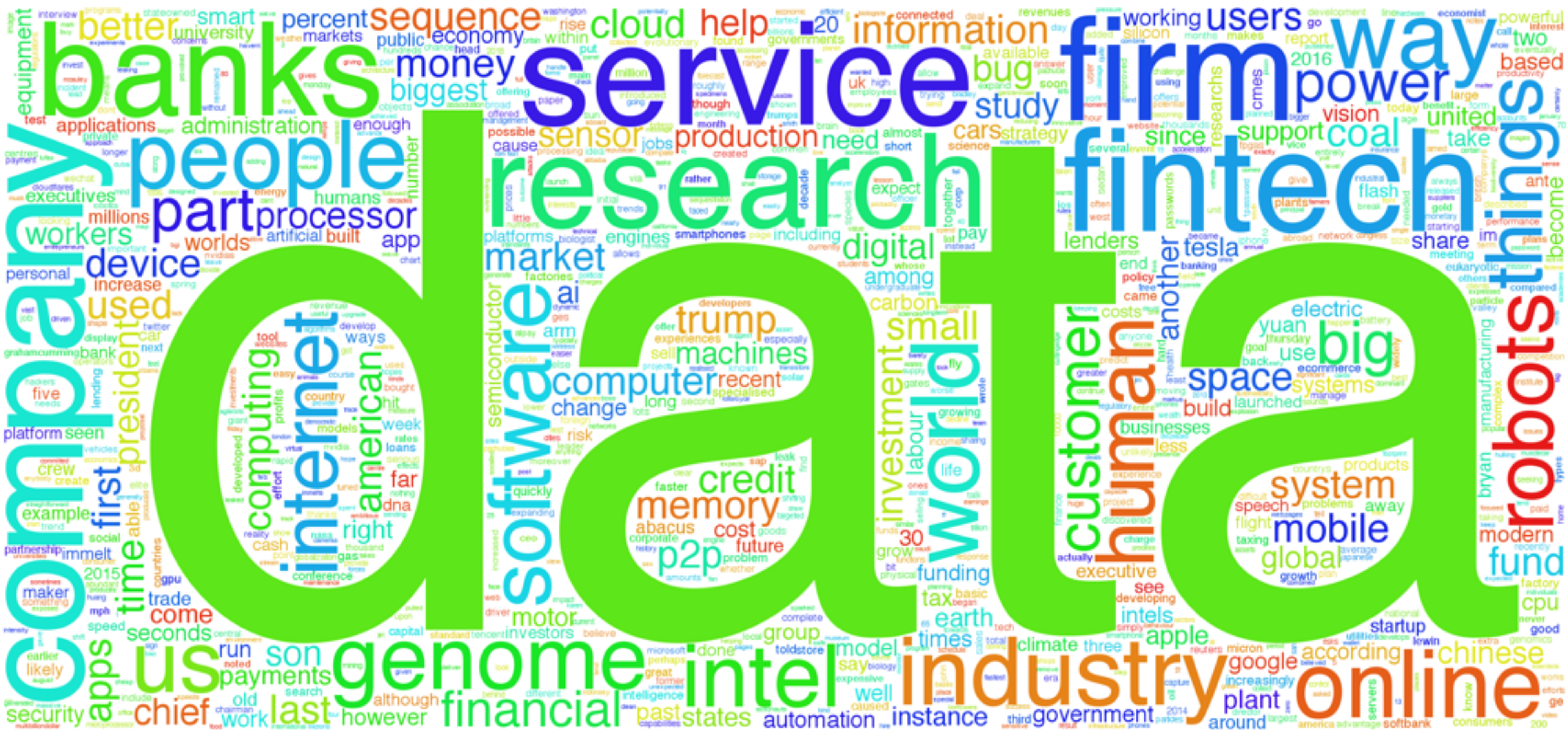
Acoustic Emissions from Bone Fracture Events



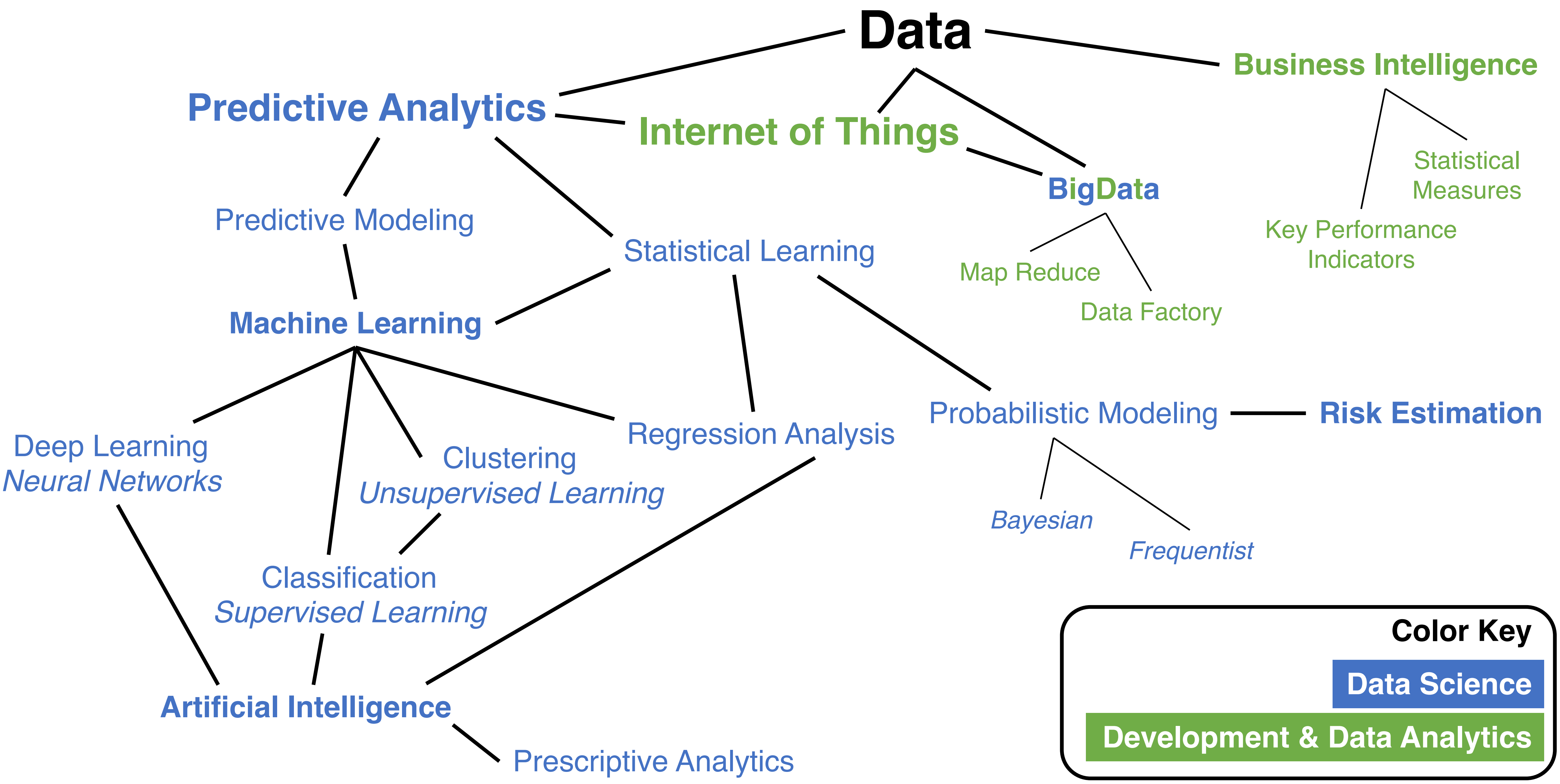
Goodwin, et al. “Acoustic Emission Signatures During Failure of Vertebra and Long Bone.” *Annals of Biomedical Engineering*, 2017.

Business, **Tech**, and **Science** News Text Sources:

The Economist; Wall Street Journal; Science (AAAS); Reuters; BBC

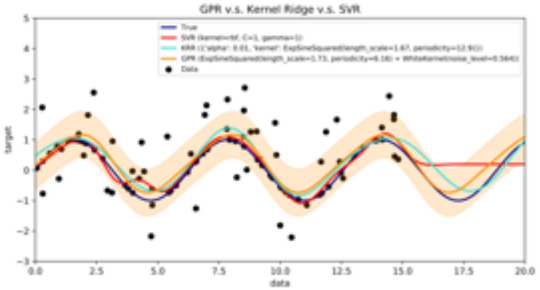


What industry thinks about Data Science

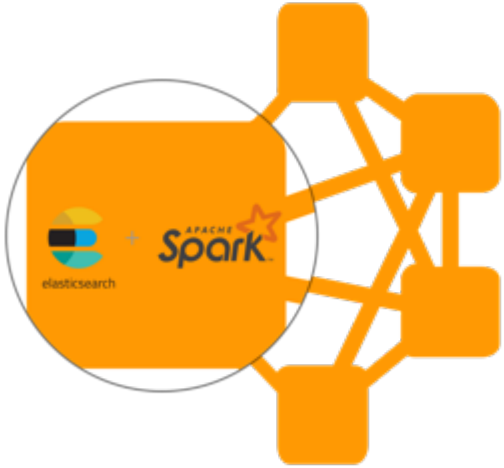


Data science is perceived as many things…

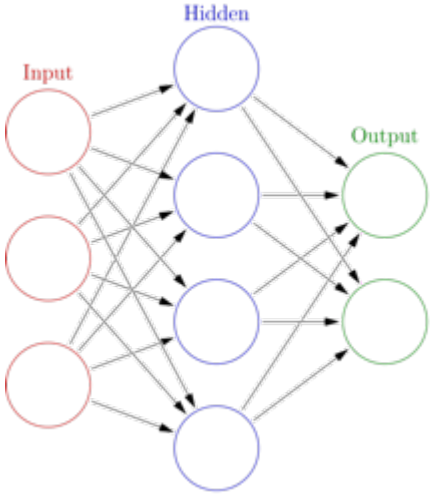
Statistics



"Big Data"

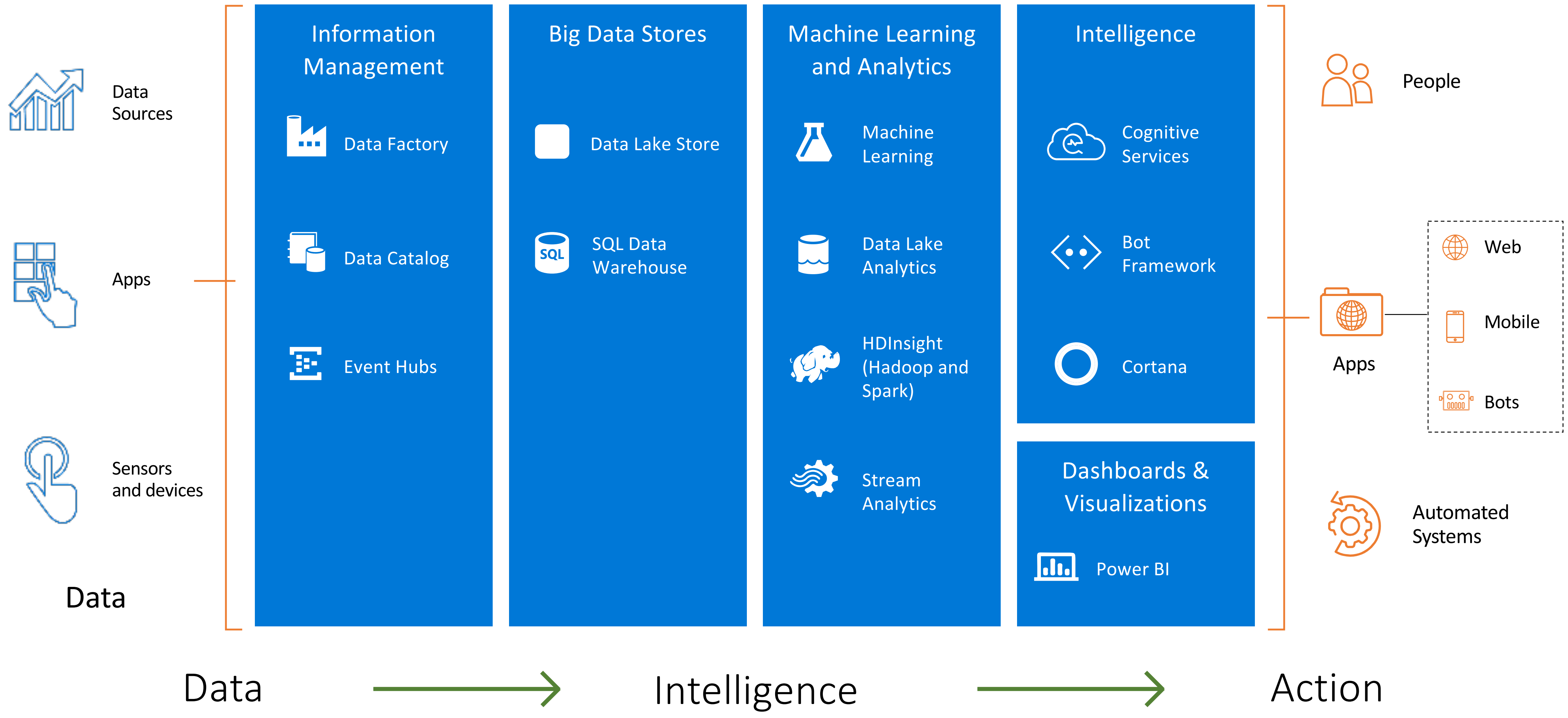


Machine-learning



Business Intelligence and Cloud

Cloud Solution Architecture



Statistics! Question...

- 1 in 1000 products are defective
- QC performs a defect test that has a 95% accuracy;
  - i.e., a 95% true positive rate and 5% false positive rate
- The test never fails to identify a defective product (false negatives are impossible)

- QC tests a product at random, and the test is positive (indicating defective product)

What is the probability that the product is *actually* defective?

We are poor intuitive statisticians

An answer in plain words:

- If 1000 randomly selected billets were tested...
- Then 50 parts will test positive (5% false positive rate)
- But only 1 of them is *actually* defective.
- Therefore, $\frac{1}{51} \approx 0.02$ or $\mathbf{2}\%$

Answer...

Human intuition is really expensive, yet it's error prone in 2 delicate environments:
1) one requiring use of statistics and
2) one where many variables are interacting with each other.

D: positive defect test
defect: part is actually defective

$$p(defect) = 0.001$$
$$p(normal) = 0.999$$
$$p(D|defect) = 1.00$$
$$p(D|normal) = 0.05$$

$$p(D) = p(D|defect)p(defect) + p(D|{\sim}defect)p({\sim}defect)$$
$$p(D) = 1.00 * 0.001 + 0.05 * 0.999 = 0.05095$$

$$p(defect|D) = \frac{p(D|defect)p(defect)}{p(D)} = \frac{1.00 * 0.001}{0.05095} = 0.0196 \approx \mathbf{2\%}$$

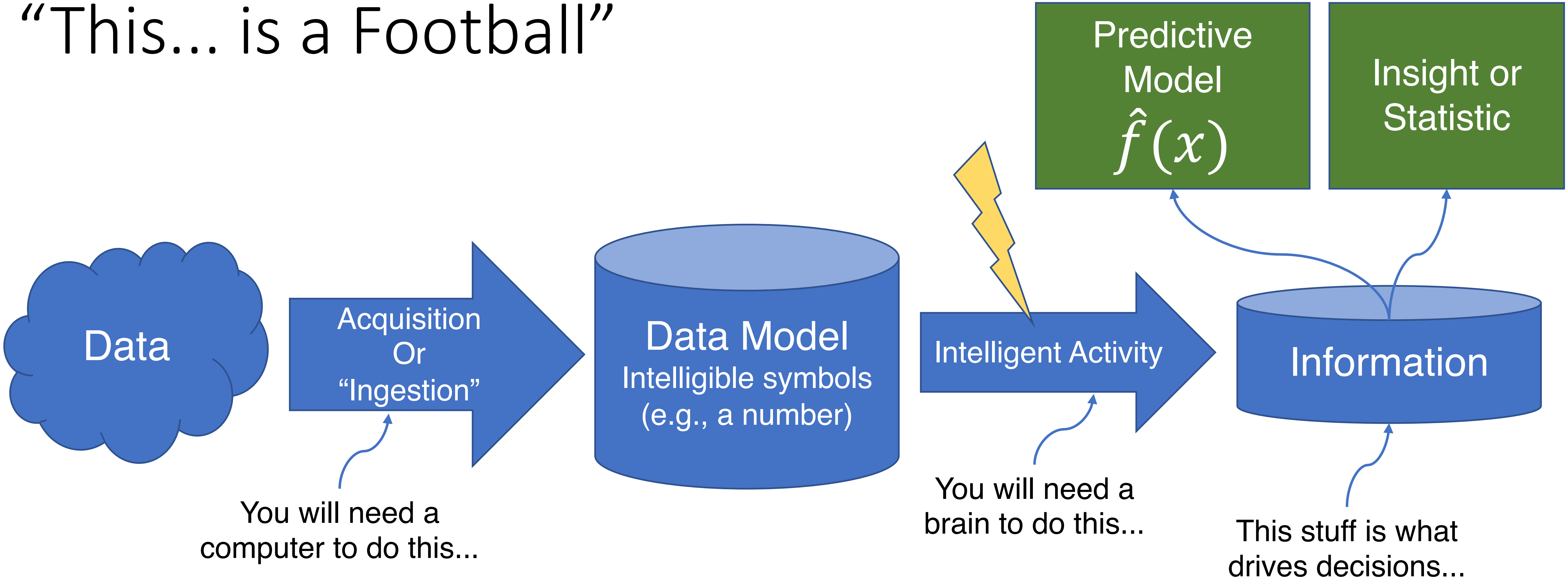
A Caveat...

”But the cleverest of algorithms are no substitute for human intelligence and knowledge of the data in the problem.”

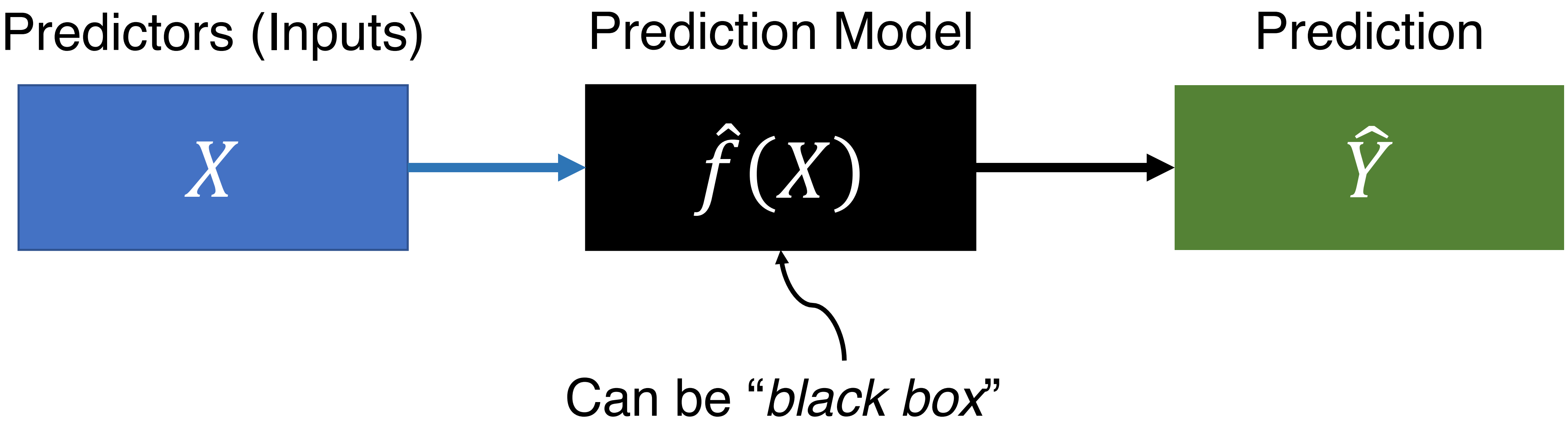
- Brieman and Cutler, UC Berkley, inventors of the Random Forest ML algorithm.

https://www.stat.berkeley.edu/~breiman/RandomForests/cc_philosophy.htm

“This... is a Football”



At the end of the day: it's a function



Why Estimate f ?

The system, f, has **predictive power** (prediction).
The system, f, has **explanatory power** (inference).

Say you have some system that produces an output

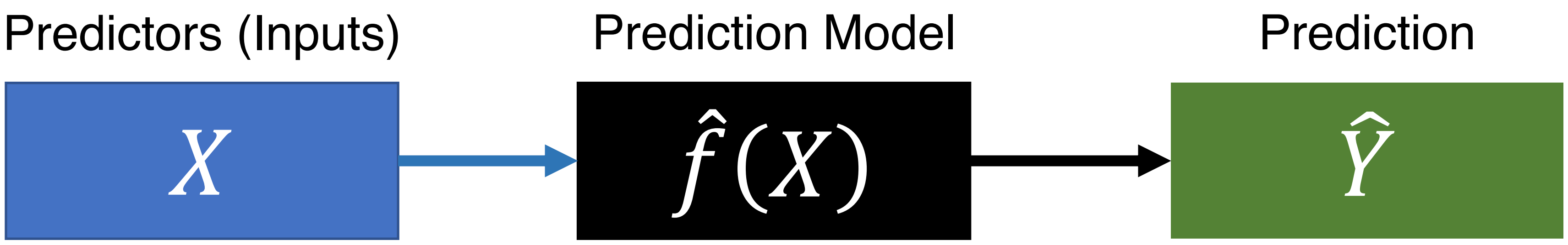
$$Y = f(X) + \epsilon$$

But it's impossible to know all of the factors (X features) that influence Y, let alone the exact function, f. So, we try to estimate this:

$$\hat{Y} = \hat{f}(X)$$

A more tangible (supervised learning) example…

	X		Y
	variable_a	variable_b	labels
1	0.54444248	-0.3646484	0
2	1.10292398	3.0249860	0
3	0.99040419	-1.2835897	0
4	-0.75436252	2.5110819	0
5	1.56692947	3.2313980	0
6	-0.24228152	4.0610301	1
7	0.24556087	1.9519334	1
8	0.52413679	3.1674849	1
9	-0.02794355	0.5639772	1
10	-0.72091824	3.5863003	1



See: https://playground.tensorflow.org/

A more tangible (supervised learning) example…

	variable_a	variable_b	labels
1	0.54444248	-0.3646484	0
2	1.10292398	3.0249860	0
3	0.99040419	-1.2835897	0
4	-0.75436252	2.5110819	0
5	1.56692947	3.2313980	0
6	-0.24228152	4.0610301	1
7	0.24556087	1.9519334	1
8	0.52413679	3.1674849	1
9	-0.02794355	0.5639772	1
10	-0.72091824	3.5863003	1

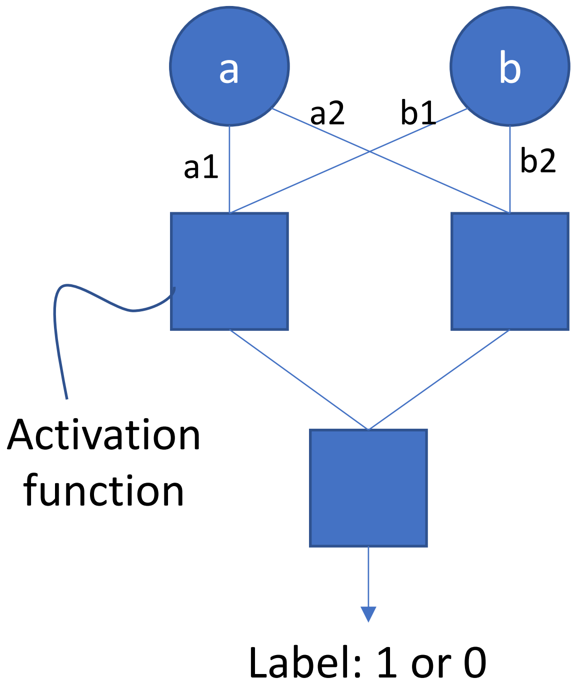
Learning or "Training" set

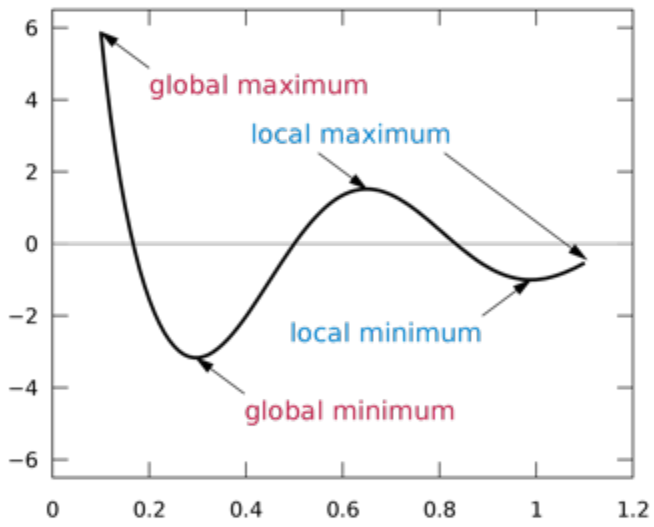
	variable_a	variable_b	labels
1	0.54444248	-0.3646484	0
2	1.10292398	3.0249860	0
3	0.99040419	-1.2835897	0
6	-0.24228152	4.0610301	1
7	0.24556087	1.9519334	1
8	0.52413679	3.1674849	1

Validation or "Testing" set

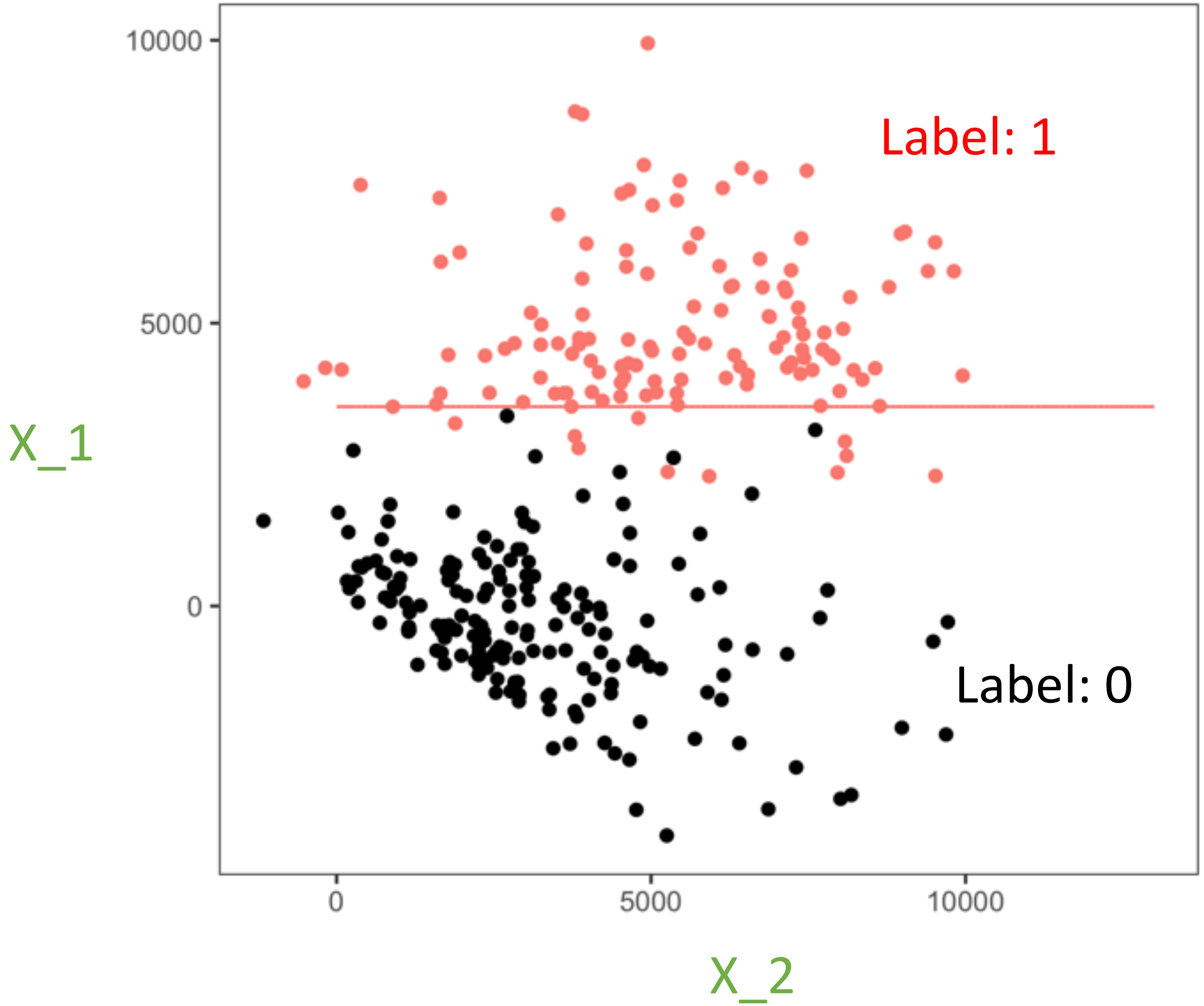
4	-0.75436252	2.5110819	0
5	1.56692947	3.2313980	0
9	-0.02794355	0.5639772	1
10	-0.72091824	3.5863003	1

Neural Network



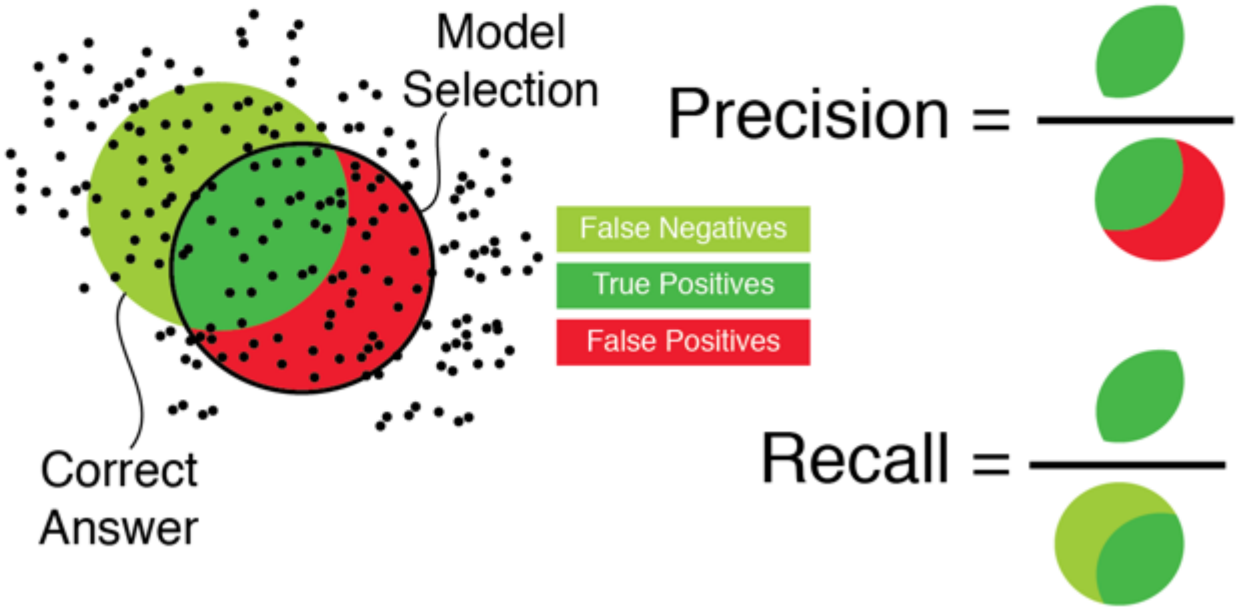


See: https://playground.tensorflow.org/



Label: 1
Label: 0
X_1
X_2
10000
5000
0
0
5000
10000

Model Performance



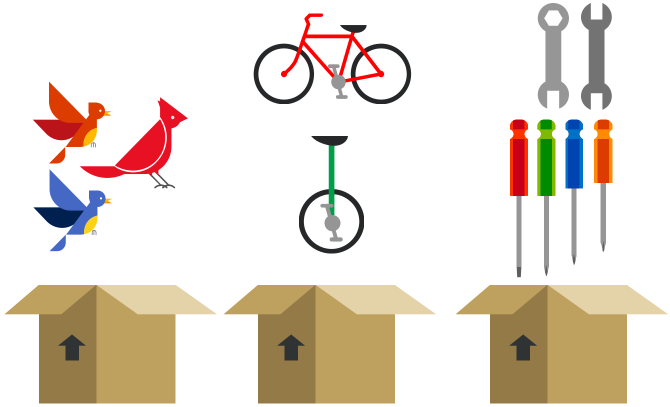
Predicted \ Actual		
SVM	FALSE	TRUE
FALSE	216	1
TRUE	23	9

Precision: 9 / (9+23) = 28.1%

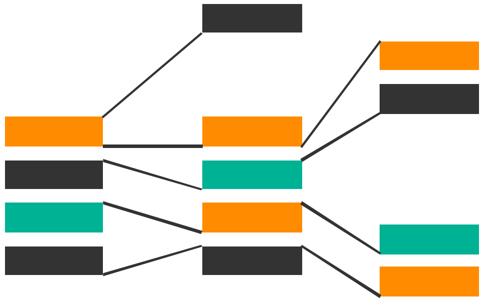
Recall: 9 / (9+1) = **90%**

Machine Learning & Estimating a $f(x)$

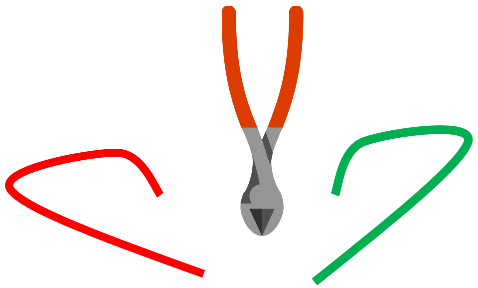
- Supervised Learning
  - Classification problems; e.g., classifying images, anomaly detection, etc.
  - Features have a known class and you want to predict the class for a new set of features.

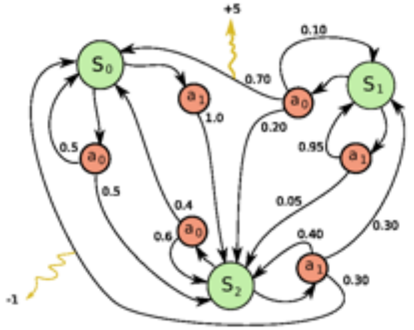


- Unsupervised Learning
  - Data clustering
  - No known labels
  - Grouping data points with similarities among their features.

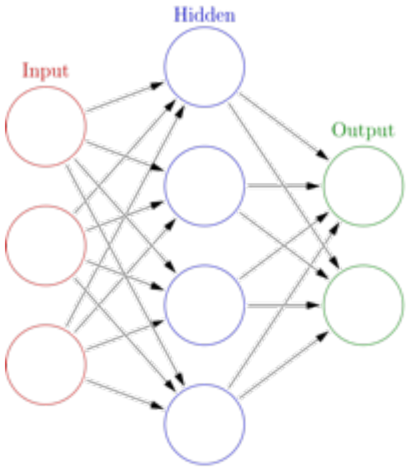


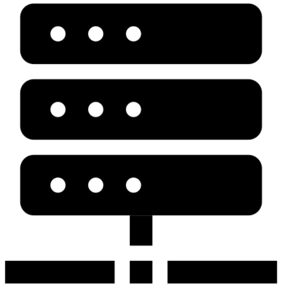
- Reinforcement Learning
  - Neural Network or Deep Learning
  - Optimizing the performance of some action
  - Common in Robotics

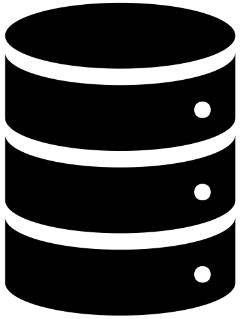




Pillars of Data Science



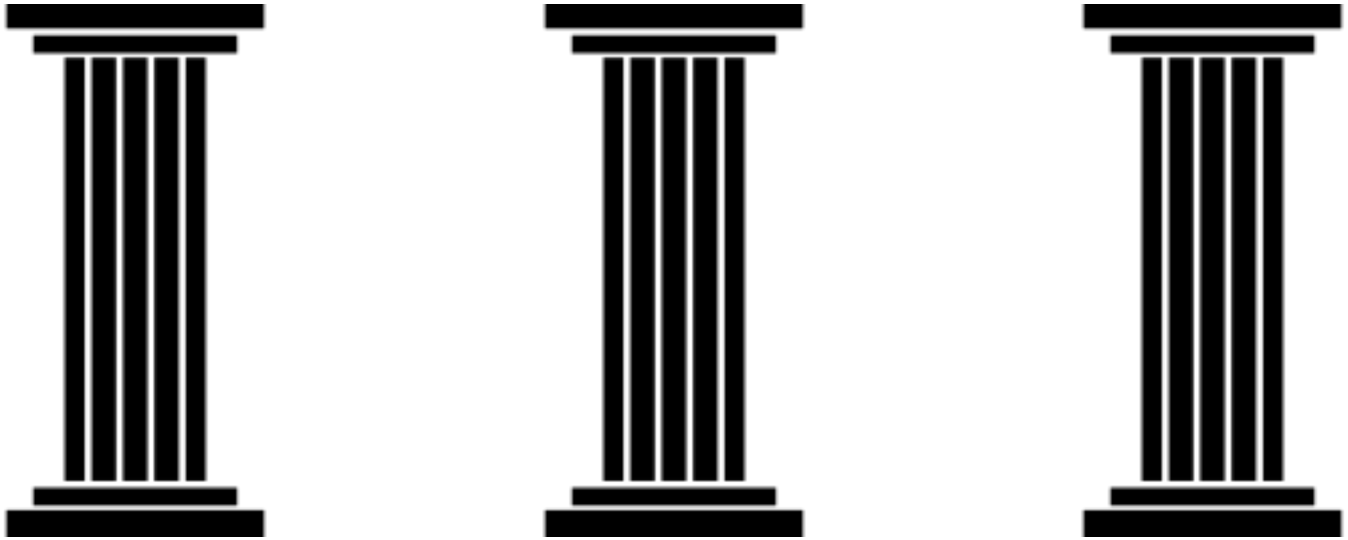




Models

Compute

Data



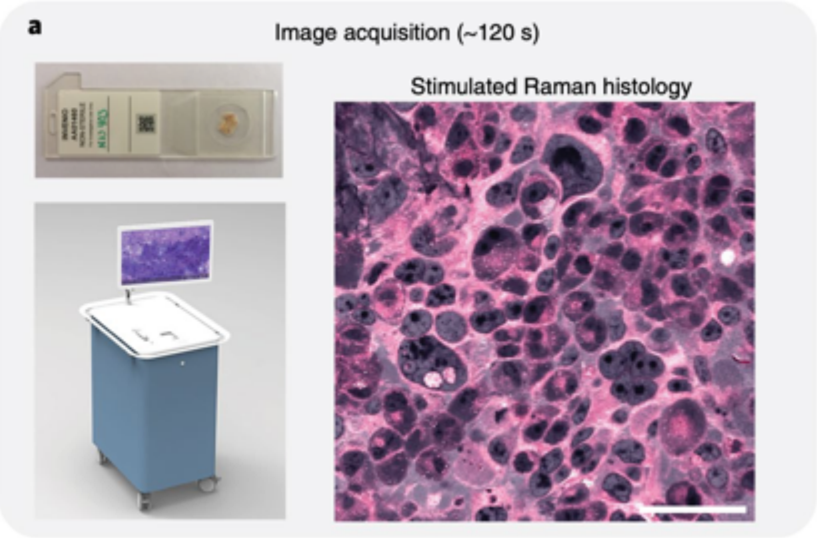
PROGRESSIVE GROWING OF GANS FOR IMPROVED QUALITY, STABILITY, AND VARIATION

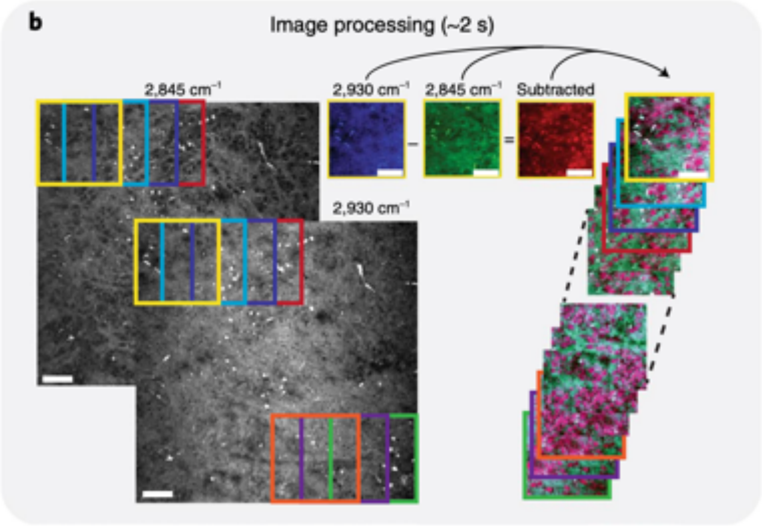
Tero Karras NVIDIA
Timo Aila NVIDIA
Samuli Laine NVIDIA
Jaakko Lehtinen NVIDIA and Aalto University
{tkarras,taila,slaine,jlehtinen}@nvidia.com

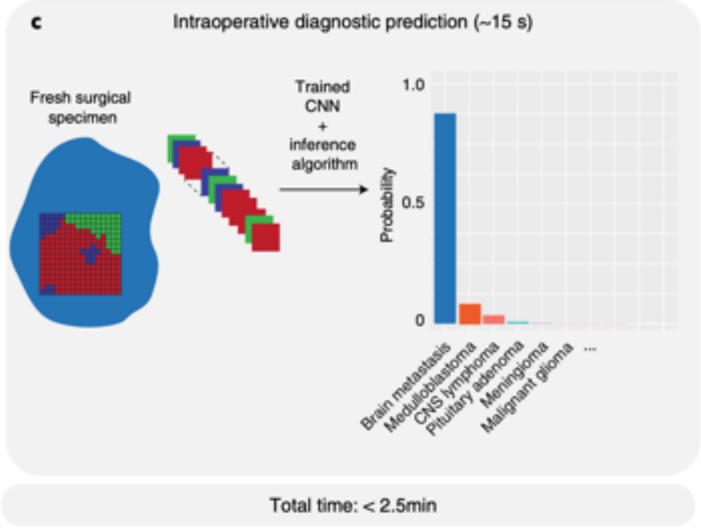
Karras, T. et al., 2017. Progressive Growing of GANs for Improved Quality, Stability, and Variation. In *ICLR 2018*. pp. 1–26.

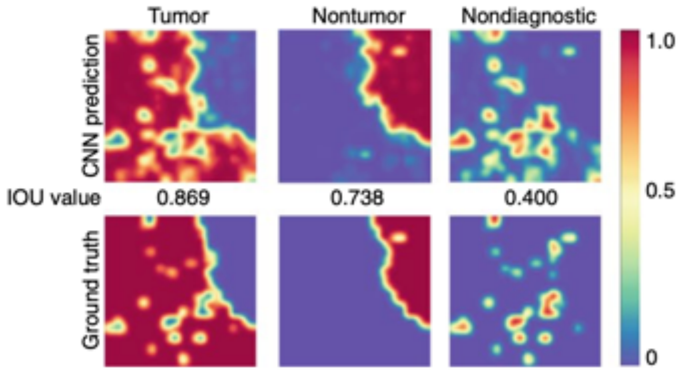
Improving intra-operative process in neurosurgery

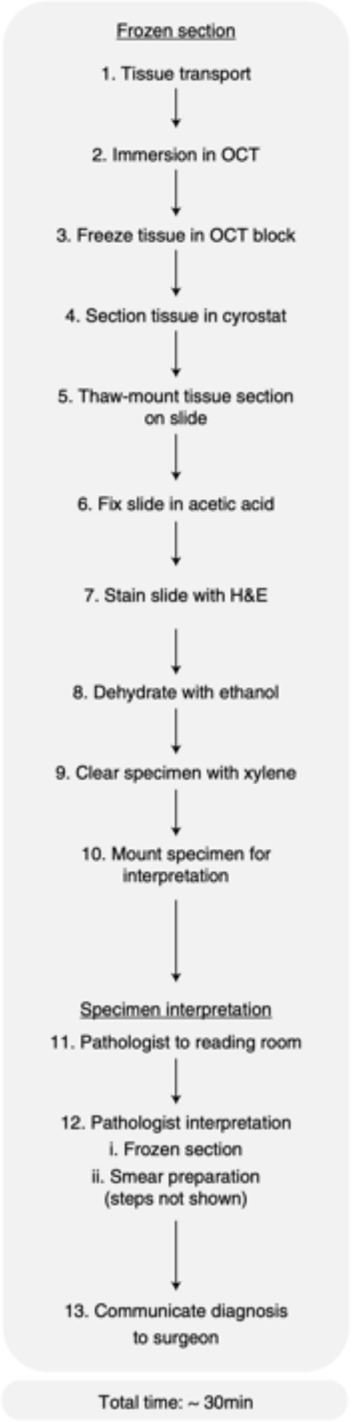
Hollon, et al. 2020, Nature.











Mastering the game of Go without human knowledge

David Silver[1]*, Julian Schrittwieser[1]*, Karen Simonyan[1]*, Ioannis Antonoglou[1], Aja Huang[1], Arthur Guez[1], Thomas Hubert[1], Lucas Baker[1], Matthew Lai[1], Adrian Bolton[1], Yutian Chen[1], Timothy Lillicrap[1], Fan Hui[1], Laurent Sifre[1], George van den Driessche[1], Thore Graepel[1] & Demis Hassabis[1]

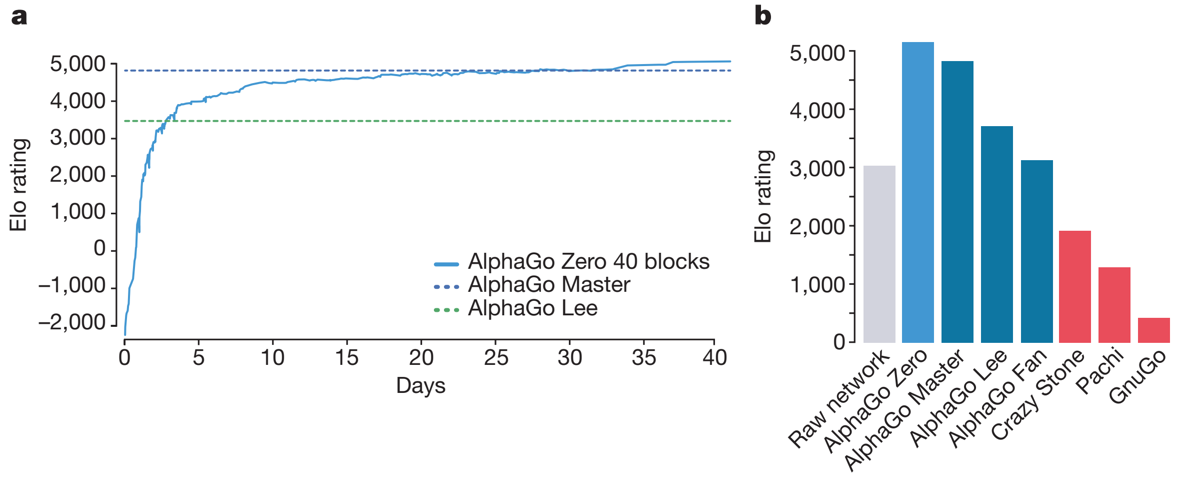
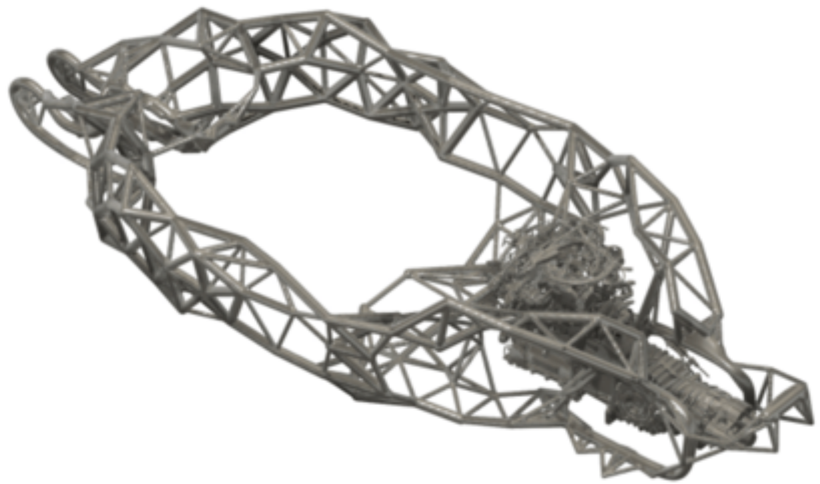


Figure 6 | Performance of AlphaGo Zero. a, Learning curve for AlphaGo Zero using a larger 40-block residual network over 40 days. The plot shows the performance of each player α_{θ_i} from each iteration i of our reinforcement learning algorithm. Elo ratings were computed from evaluation games between different players, using 0.4 s per search (see Methods). **b**, Final performance of AlphaGo Zero. AlphaGo Zero was trained for 40 days using a 40-block residual neural network. The plot shows the results of a tournament between: AlphaGo Zero, AlphaGo Master (defeated top human professionals 60–0 in online games), AlphaGo Lee (defeated Lee Sedol), AlphaGo Fan (defeated Fan Hui), as well as previous Go programs Crazy Stone, Pachi and GnuGo. Each program was given 5 s of thinking time per move. AlphaGo Zero and AlphaGo Master played on a single machine on the Google Cloud; AlphaGo Fan and AlphaGo Lee were distributed over many machines. The raw neural network from AlphaGo Zero is also included, which directly selects the move a with maximum probability p_a, without using MCTS. Programs were evaluated on an Elo scale[25]: a 200-point gap corresponds to a 75% probability of winning.

Silver, D. et al., 2017. Mastering the game of Go without human knowledge. *Nature*, 550, p.354.

Design Optimization



- Constrained optimization
  - Design constraints
  - Optimizing for specific circumstance (e.g., lap speed on a specific course)
  - Structural engineering and optimization
- Reinforcement learning
  - Massive parameter space
  - Simulation environment for quantifying outcomes

https://www.fastcompany.com/3054028/inside-the-hack-rod-the-worlds-first-ai-designed-car