

The Next Era: Deep Learning for Biomedical Research



Srinivasan Parthiban

ICIC 2017, Heidelberg, Germany
October 23-24, 2017

Mastering the Game of **Go**

2017

AlphaGo Zero
Discovering new knowledge



Pharma Industry has a problem:

Cost to develop a New Drug > \$ 2.6 B and Growing



VINGYANI

**Big breakthroughs happen when
what is suddenly possible meets
what is desperately needed**

Thomas L. Friedman



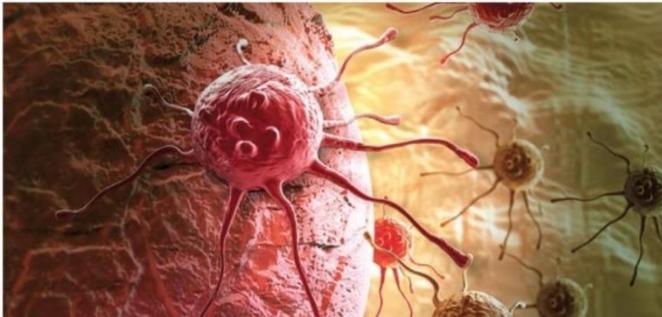
The Beginning of the AI Revolution in Pharma

MedTech

IBM unveils Watson for drug R&D, teams with Pfizer on oncology

by Amirah Al Idrus | Dec 1, 2016 8:00am

Fierce Biopharma 2016



Watson for Drug Delivery will rapidly analyze a trove of medical literature to help researchers identify new drug targets and new indications for existing drugs.

IBM is launching the latest addition to the Watson family: Watson for Drug Discovery, which will use cognitive computing to help researchers zero in on new drug targets. First up is a collaboration with Pfizer, which will use the tech to boost its immuno-oncology research.

Watson for Drug Discovery is a cloud-based platform that will use deep learning, natural language processing and machine learning to analyze medical literature and clinical trials to identify new drug targets and new indications for existing drugs.

DRUG INDUSTRY DAILY NEWS

TOPICS WATCH EVENTS SEARCH SOCIAL GUIDES LEARN

Novartis to Work with IBM Watson on Outcomes-Based Care in Advanced Breast Cancer DDD 2017

Mon, 06/05/2017 - 10:28am by Novartis



#DEALS AUGUST 28, 2017 / 4:10 PM / 2 MONTHS AGO

AstraZeneca taps AI for drug discovery in deal with Berg

Reuters Staff

August, 2017



LONDON (Reuters) - AstraZeneca (AZN.L) has forged a research collaboration with Boston-based Berg, a specialist in artificial intelligence for drug hunting, in the latest sign of big pharma's interest in using supercomputers for drug discovery.

Better drugs, faster: The potential of AI-powered humans

By Emma Woolacott
Technology of Business reporter **BBC, August 2017**

01 August 2017 | Business

f t m Share



Scientists working in tandem with artificial intelligence (AI) could slash the time it takes to develop new drugs - and, crucially, the cost - say tech companies.

NEWS

AI-powered drug discovery captures pharma interest

Nature, July 2017



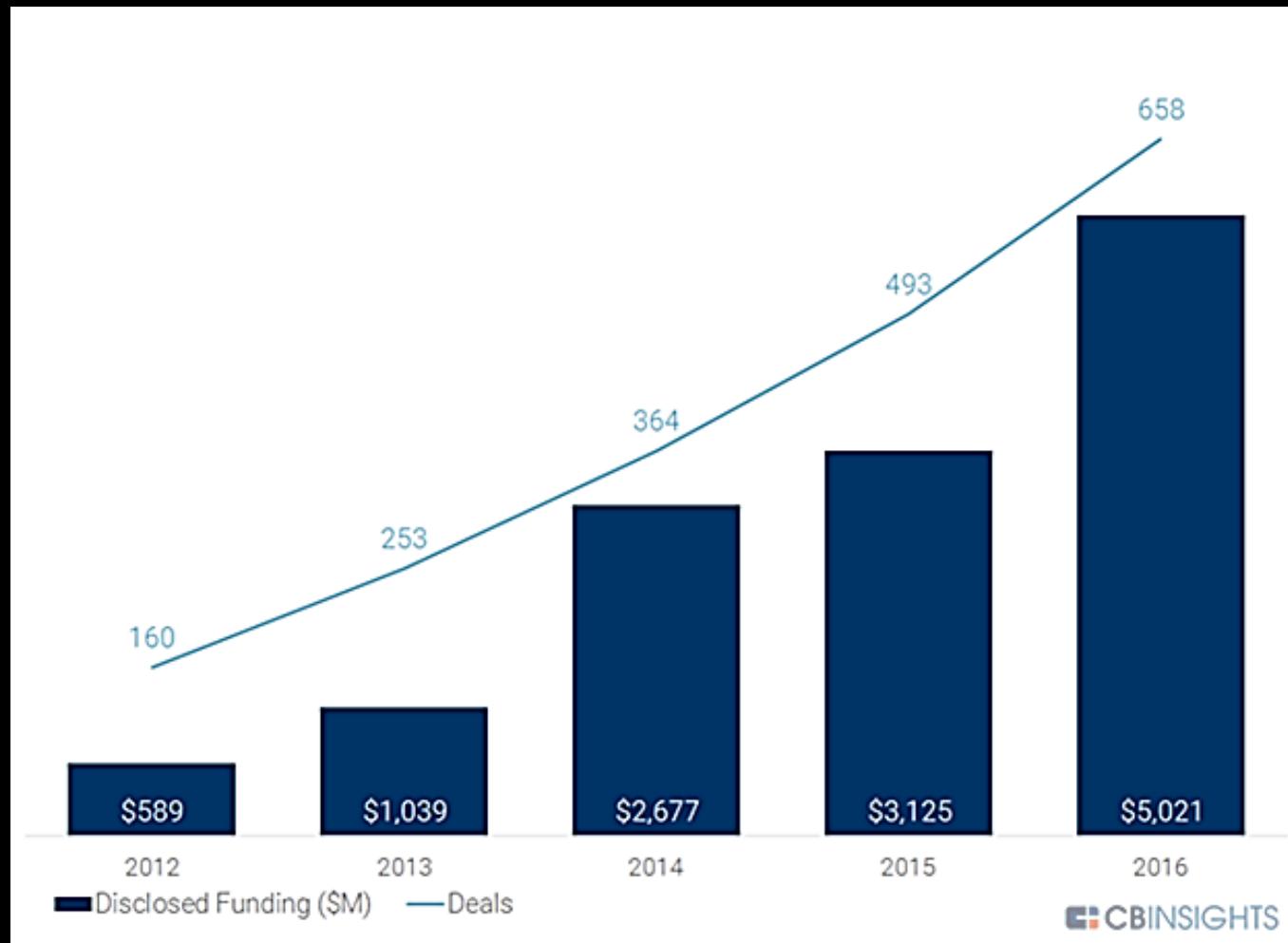
Nature Biotechnology 35, 604–605 (2017) | doi:10.1038/nbt0717-604

Published online 12 July 2017

Deep learning is starting to gain acolytes in the drug discovery space.

confluence of ever-increasing computational targets designing and optimizing compounds.

AI Annual Global **Financing History**



Healthcare emerges as hottest area of investment



Cognitive Science

**Artificial
Intelligence**

**Machine
Learning**

**Deep
Learning**

The Deep Learning Recipe



Algorithms

How many shortest-length paths are there from your house to the doughnut shop?

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

+ up's
↑
+ right's
→

$$\binom{11}{5} = \binom{11}{6} = 330 \text{ paths}$$

$\binom{11}{5} = \binom{11}{6} = 330 \text{ paths}$

$P \models Q \vdash R \models P \vee Q \models P \vee R \models (P \vee Q) \wedge (P \vee R)$

T	T	T	T	T	T
T	T	F	T	T	T
T	F	T	T	T	T
F	T	F	T	T	T
F	F	T	F	T	T
F	F	F	T	F	T
F	F	F	F	T	T

$\sum_{i=1}^n i = 1 + 2 + 3 + \dots + n$

$A_1 - A_2 = 4$
 $A_2 - A_3 = 4$
 $A_3 - A_4 = 4$
 \vdots
 $+ A_n - A_1 = 4n$
 $A_n = A_1 + 4n$
 $K_{3,3}$

Onto

One-to-one

$S_n = 11866$

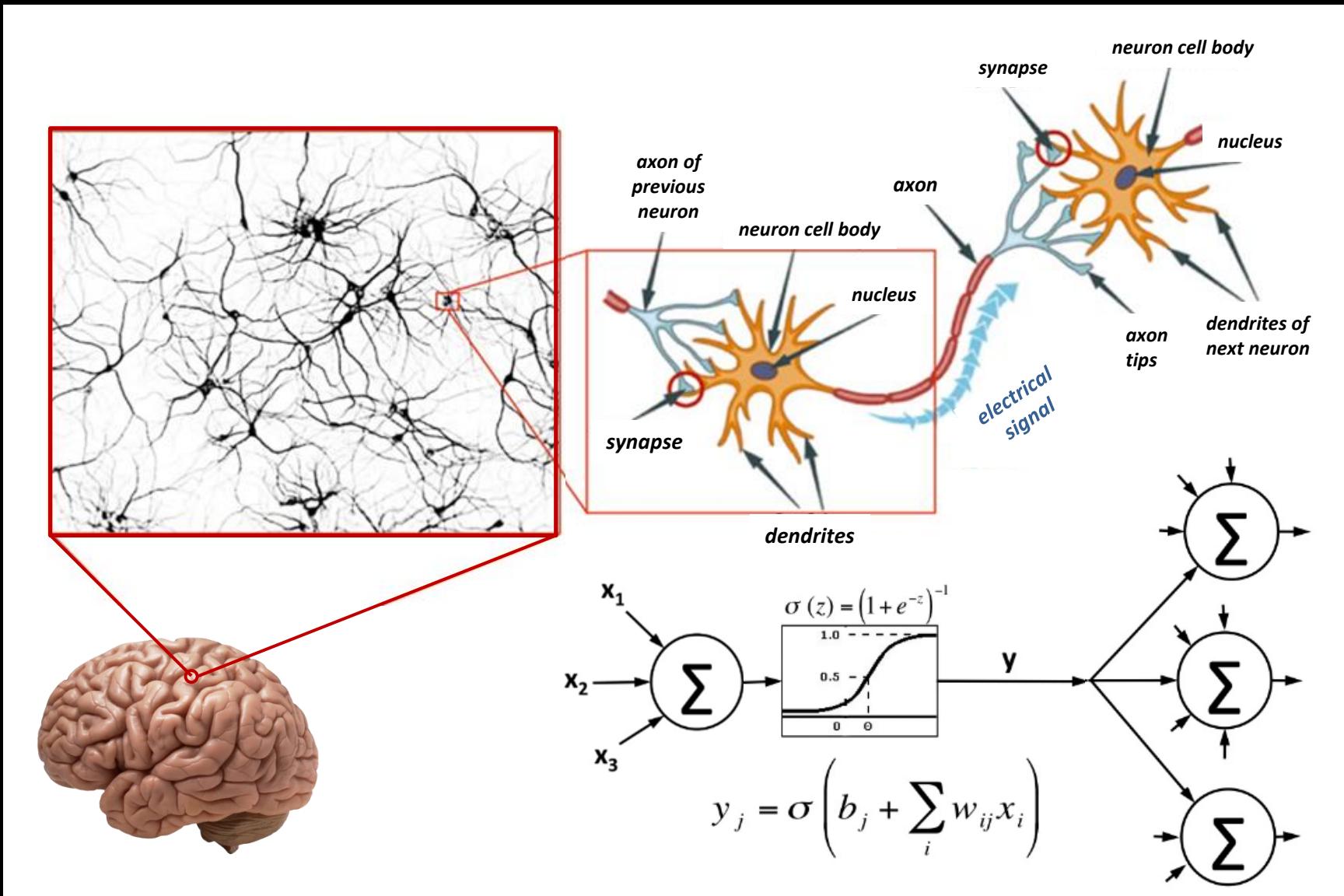
$\exists x \forall y (x \geq 2y \rightarrow x > y+1)$
 converse: $\exists x \forall y (x > y+1 \rightarrow x \geq 2y)$
 Negation: $\neg (\exists x \forall y (x \geq 2y \vee x > y+1))$
 $\forall x \exists y (x \geq 2y \wedge x \leq y+1)$
 Contrapositive: $\exists x \forall y (x \leq y+1 \rightarrow x < 2y)$

$P.I.E. \text{ Example:}$

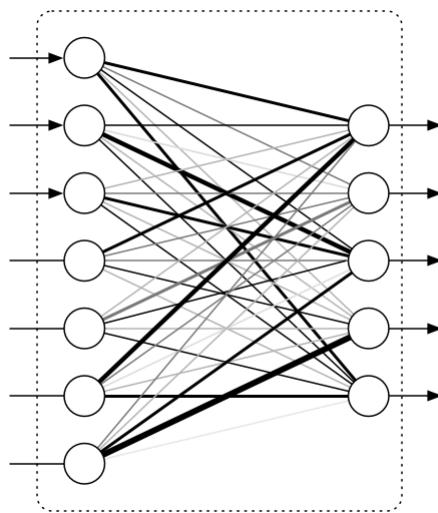
$A = \{2, 4, 10, 7\}$

$\binom{6}{1} - [\binom{6}{1} \cdot \binom{6}{2} + \binom{6}{3} \cdot \binom{6}{3} + \binom{6}{4} \cdot \binom{6}{4} + \binom{6}{5} \cdot \binom{6}{1}]$

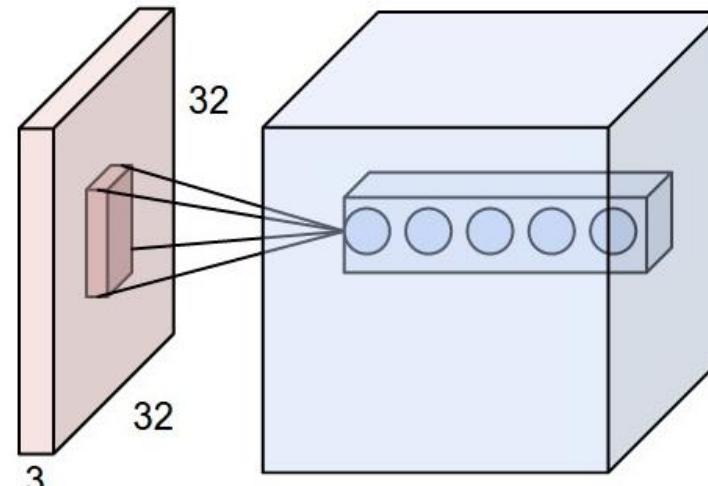
Neurons and the Brain



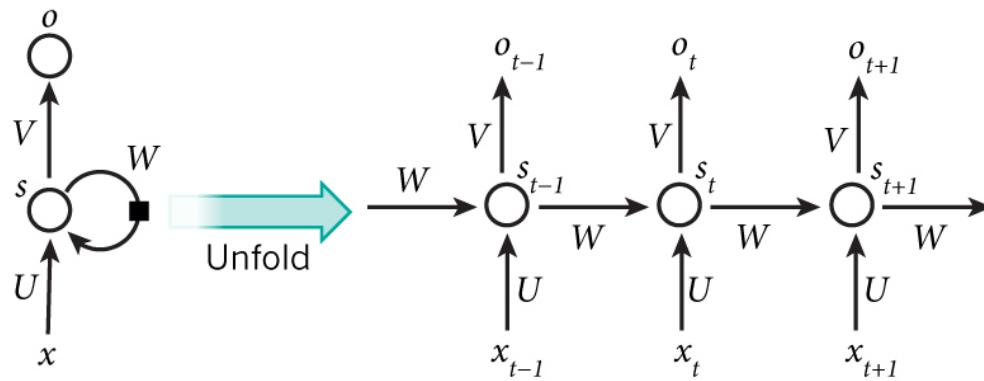
Deep Learning Primitives



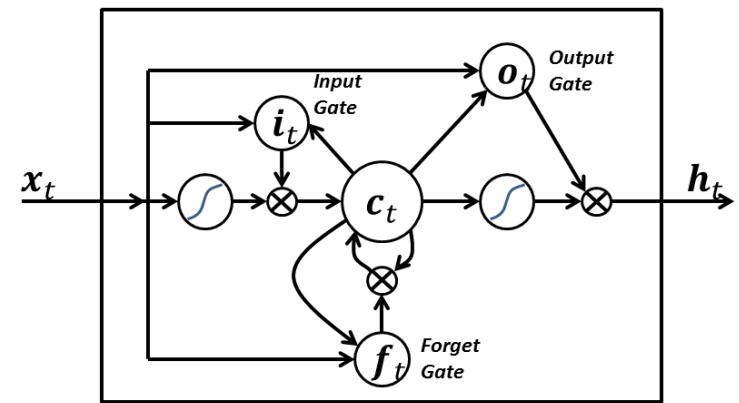
Fully Connected Layer



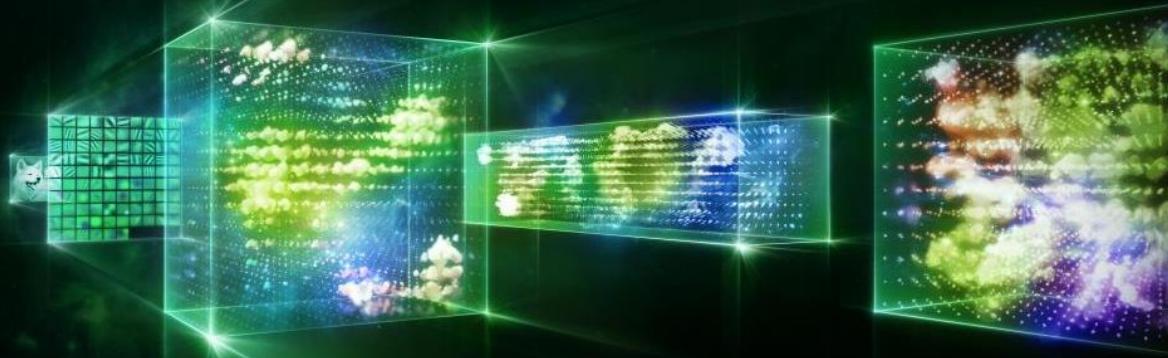
Convolutional Layer



Recurrent Neural Network Layers



Long Short Term Memory (LSTM) Cells



Dog



Welcome to the First International Beauty Contest
Judged by Artificial Intelligence
Beauty.AI 2.0

Be the First Beauty Queen or King
Judged by Robots

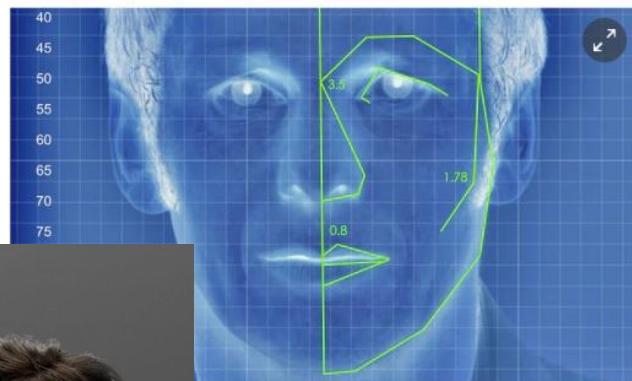
Watch our video

GET IT ON Google play App Store

A robotic head and a woman's profile with a wireframe overlay, illustrating the AI judging process.

New AI can guess whether you're gay or straight from a photograph

An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions



The Algorithm Will See You Now



Radiology



Pathology

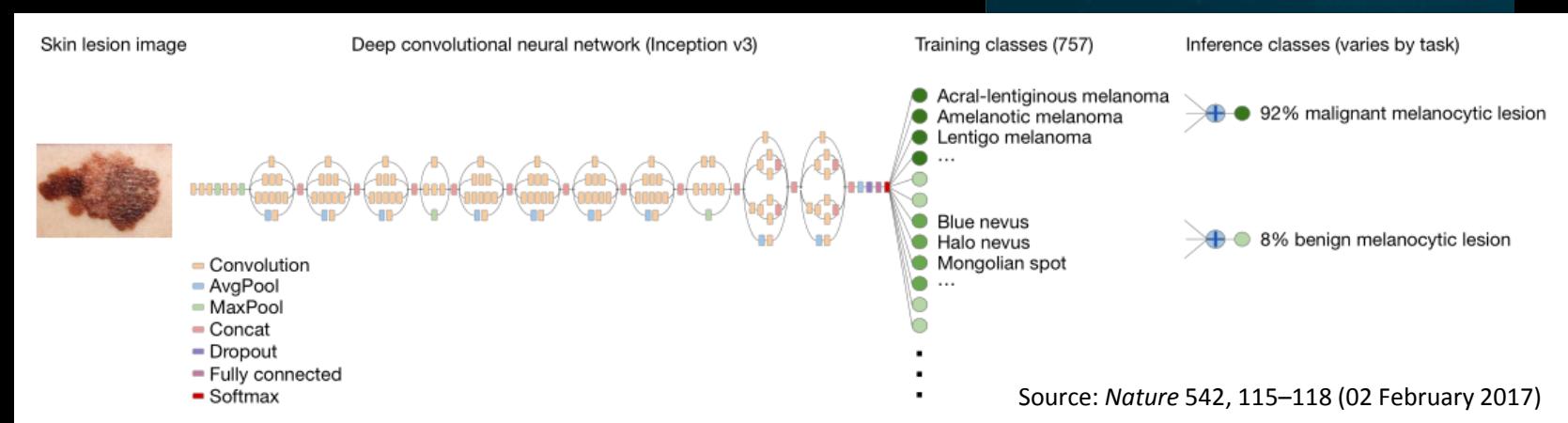
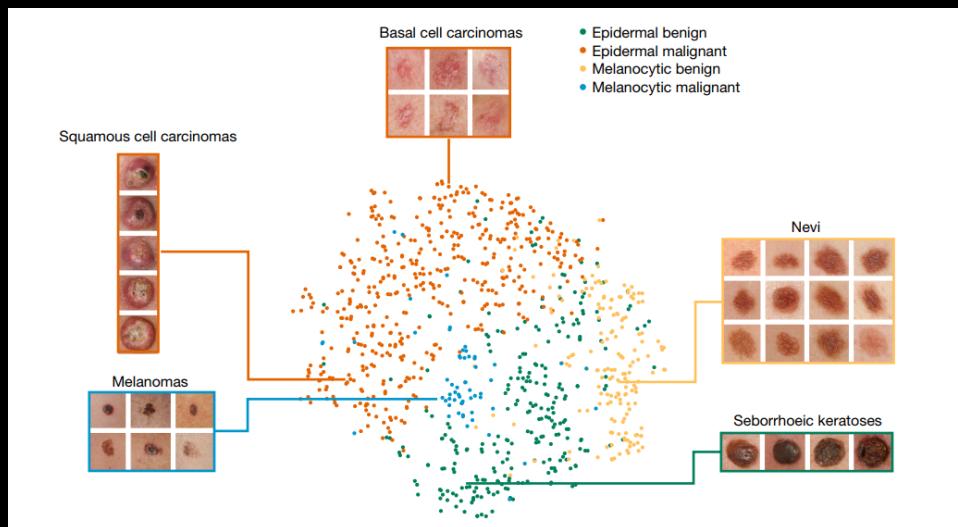


Dermatology

AI and Your Skin

Dermatologist-level classification of skin cancer with deep neural networks

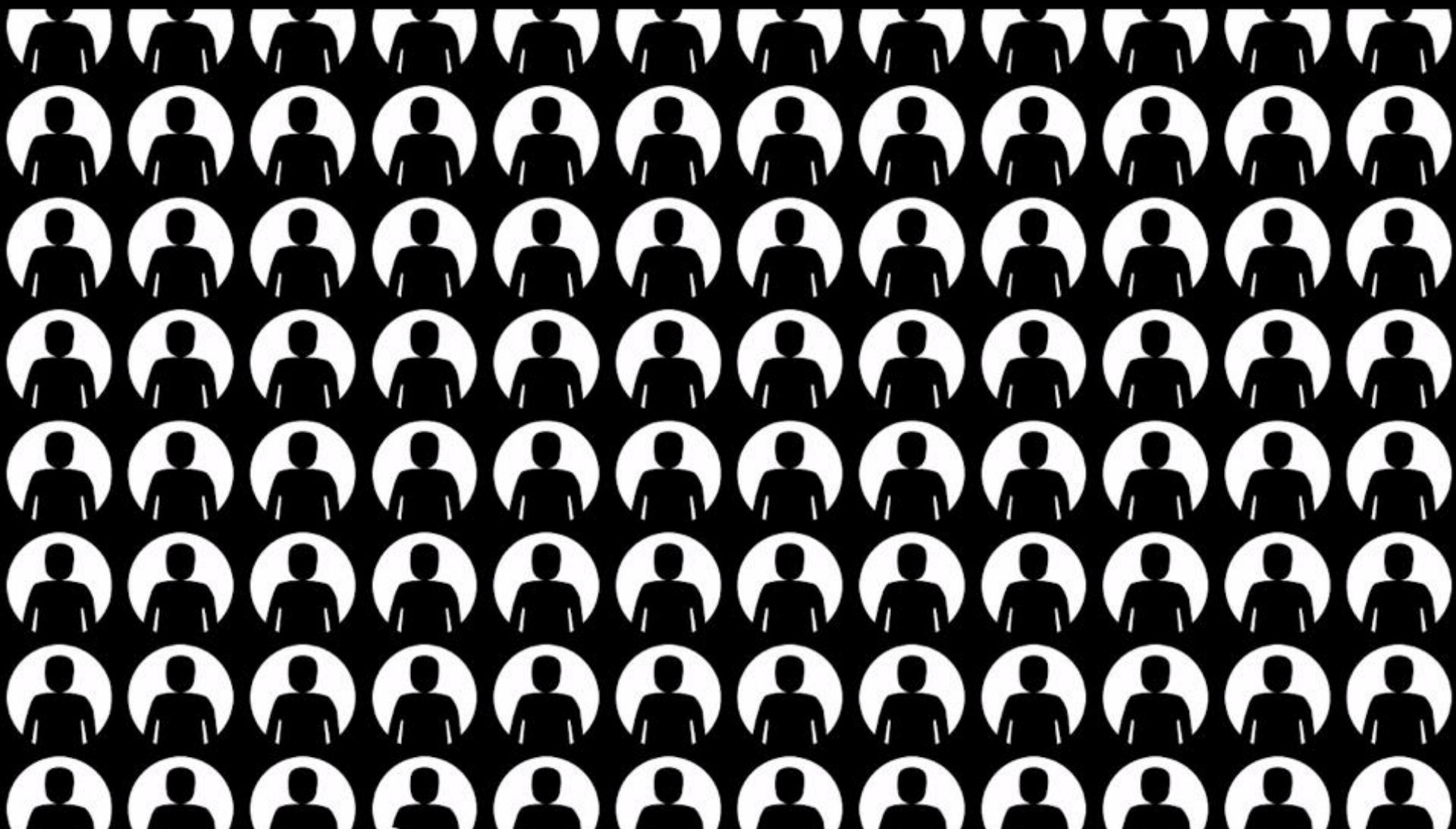
Andre Esteva^{1*}, Brett Kuprel^{1*}, Roberto A. Novoa^{2,3}, Justin Ko², Susan M. Swetter^{2,4}, Helen M. Blau⁵ & Sebastian Thrun⁶



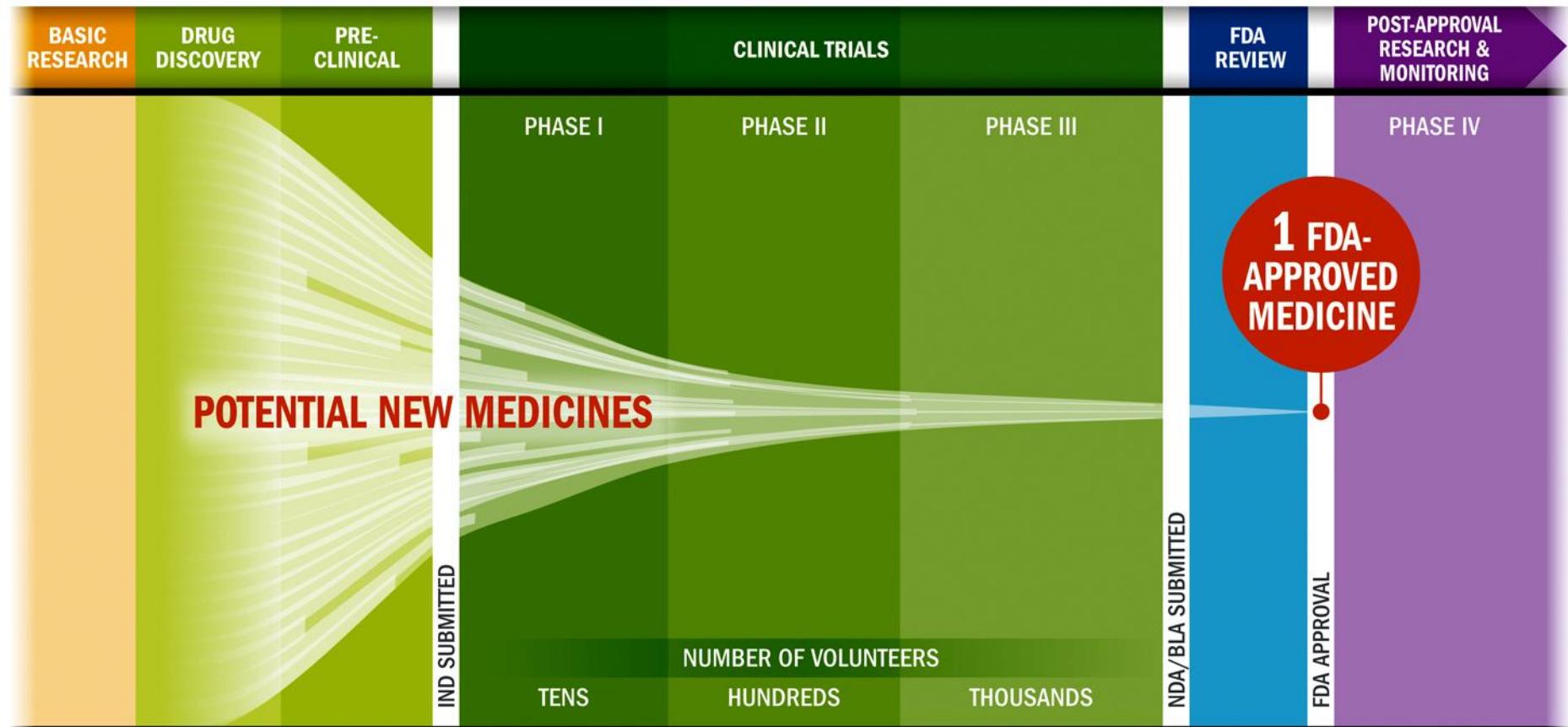
Electronic Health Records



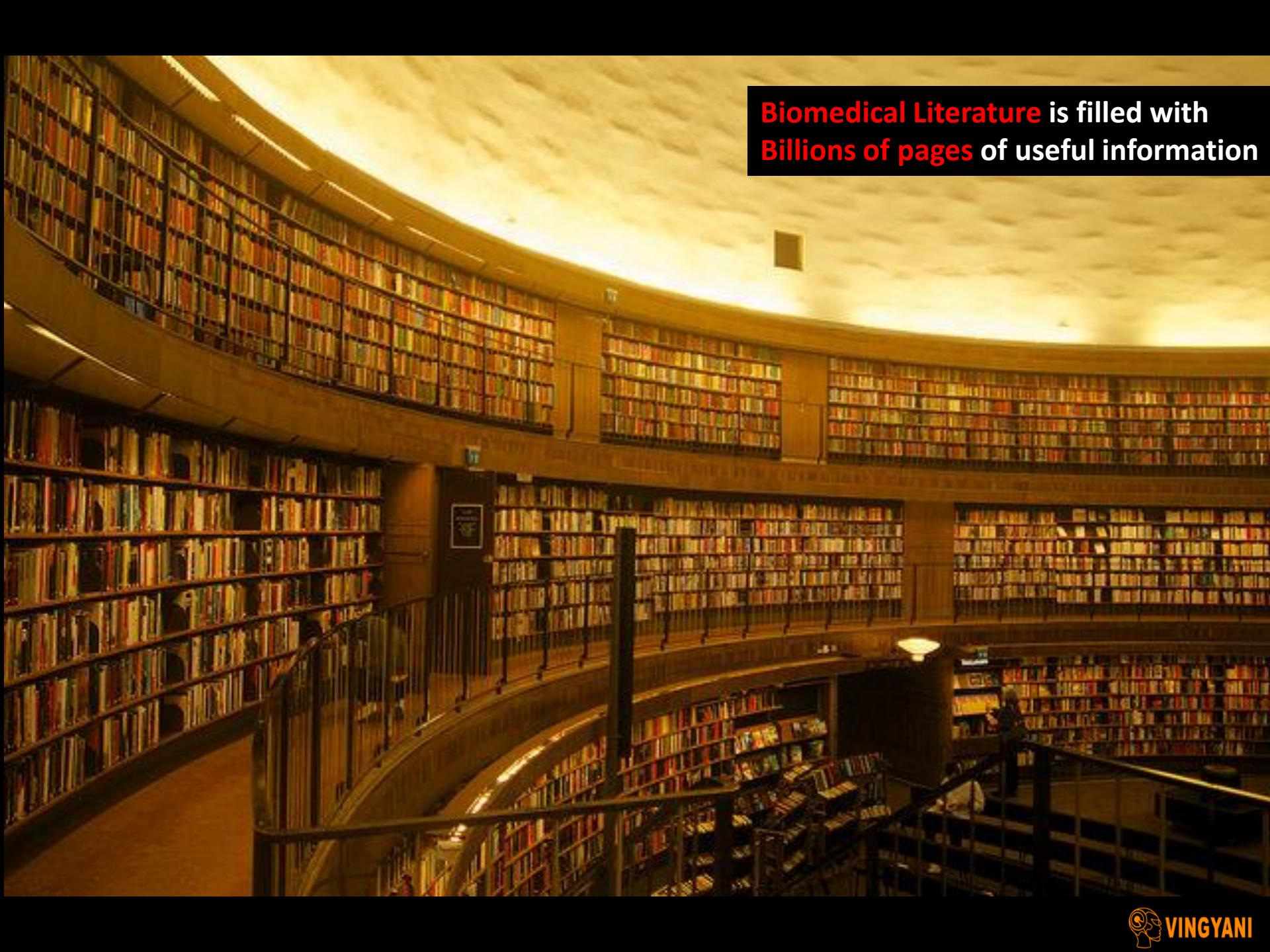
Electronic Health Records



Biopharmaceutical Research and Development Process



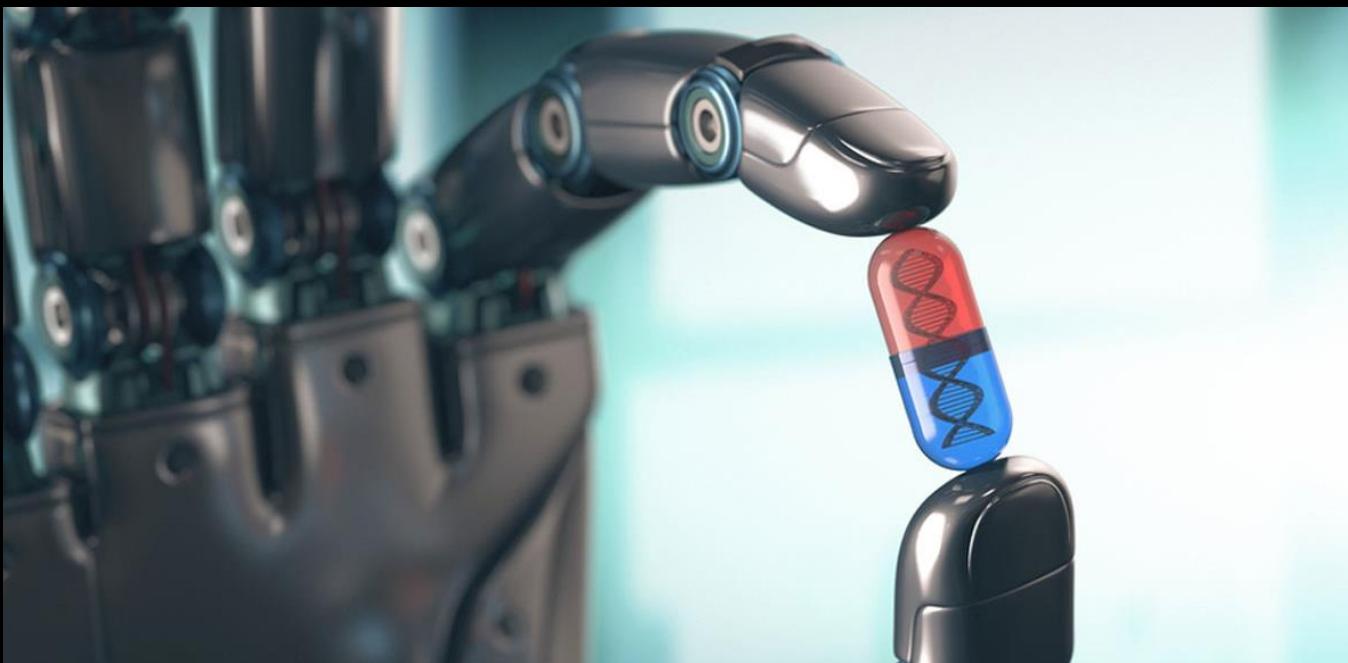
Key: IND: Investigational New Drug Application, NDA: New Drug Application, BLA: Biologics License Application

A photograph of a large, modern library. The space is circular and features multiple levels of curved bookshelves filled with books. The shelves are made of wood and are illuminated by warm, recessed lighting in the ceiling. The floor is a light-colored wood or laminate. A small staircase leads up to the second level. In the top right corner of the image, there is a black rectangular box containing red text.

Biomedical Literature is filled with
Billions of pages of useful information

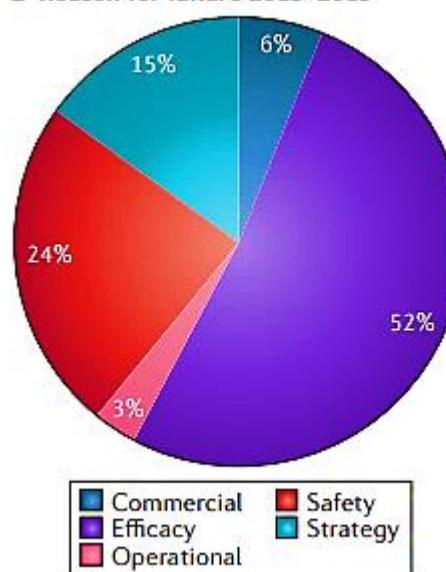
Chemists pin hopes on deep learning for drug discovery

**Pharma scientists are exploring what
the algorithm behind voice and face
recognition can do for chemistry**

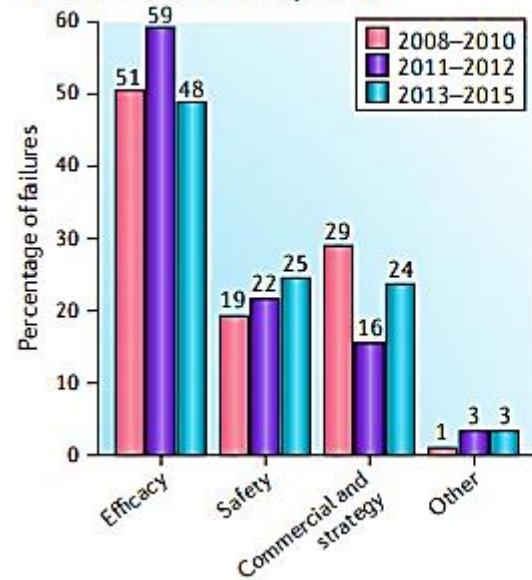


Causes of Clinical Failure

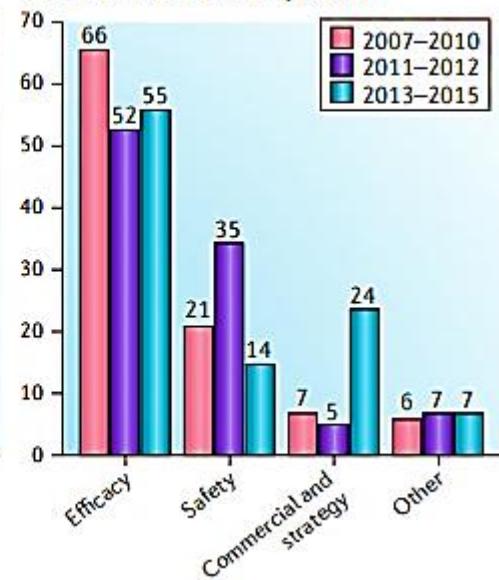
a Reason for failure 2013–2015



b Reason for failure in phase II



c Reason for failure in phase III



Data Repositories

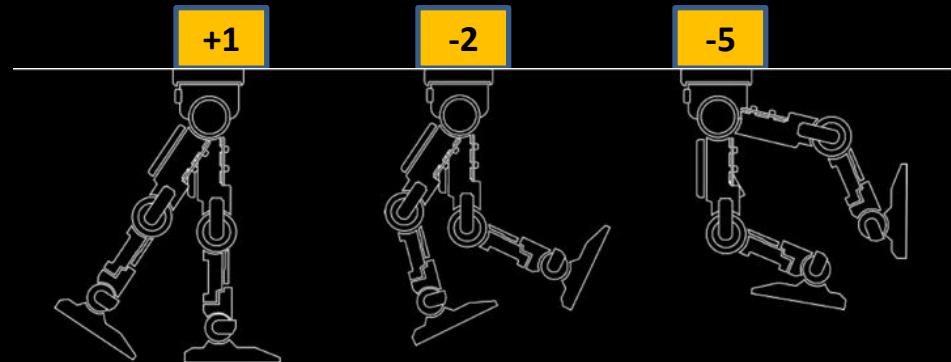
Database	Unique Compounds	Experimental facts	Main data types
ChEMBL v.21	1,592,191	13,968,617	PubChem HTS assays and data mined from literature
BindingDB	529,618	1,207,821	Experimental protein-small molecule interaction data
PubChem	>60M	>157M	Bioactivity data from HTS assays
Reaxys	>74M	>500M	Literature mined property, activity and reaction data
SciFinder (CAS)	>111M	>80M	Experimental properties ^{13}C and ^1H NMR spectra, reaction data
GOSTAR	>3M	>24M	Target-linked data from parents and articles
AZ IBIS	-	>150M	AZ in-house SAR data points
OCHEM	>600k	>1.2M	Mainly ADMET data collected from literature

Reinforcement Learning

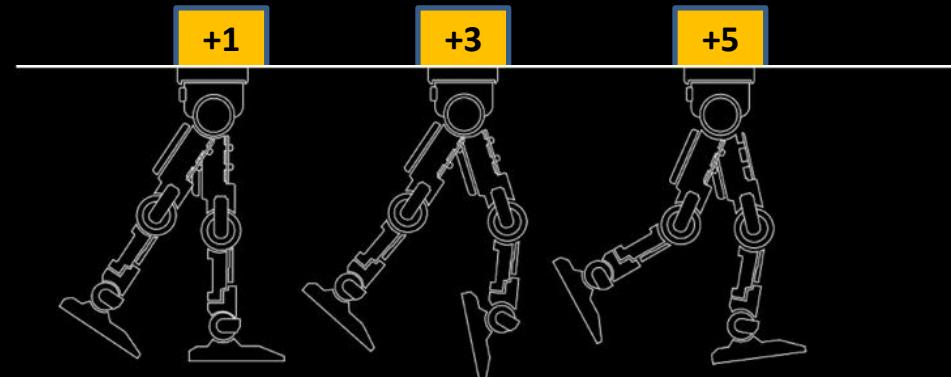
$$R = \sum_{t=0}^{\infty} \gamma^t r_{t+1} \quad f(\beta) := \frac{1}{N} \sum_{i=1}^N L(\beta; (x_i, y_i)) + R(\beta)$$

Use Math to Represent the Goal of Walking

Receive a Score for Each Move

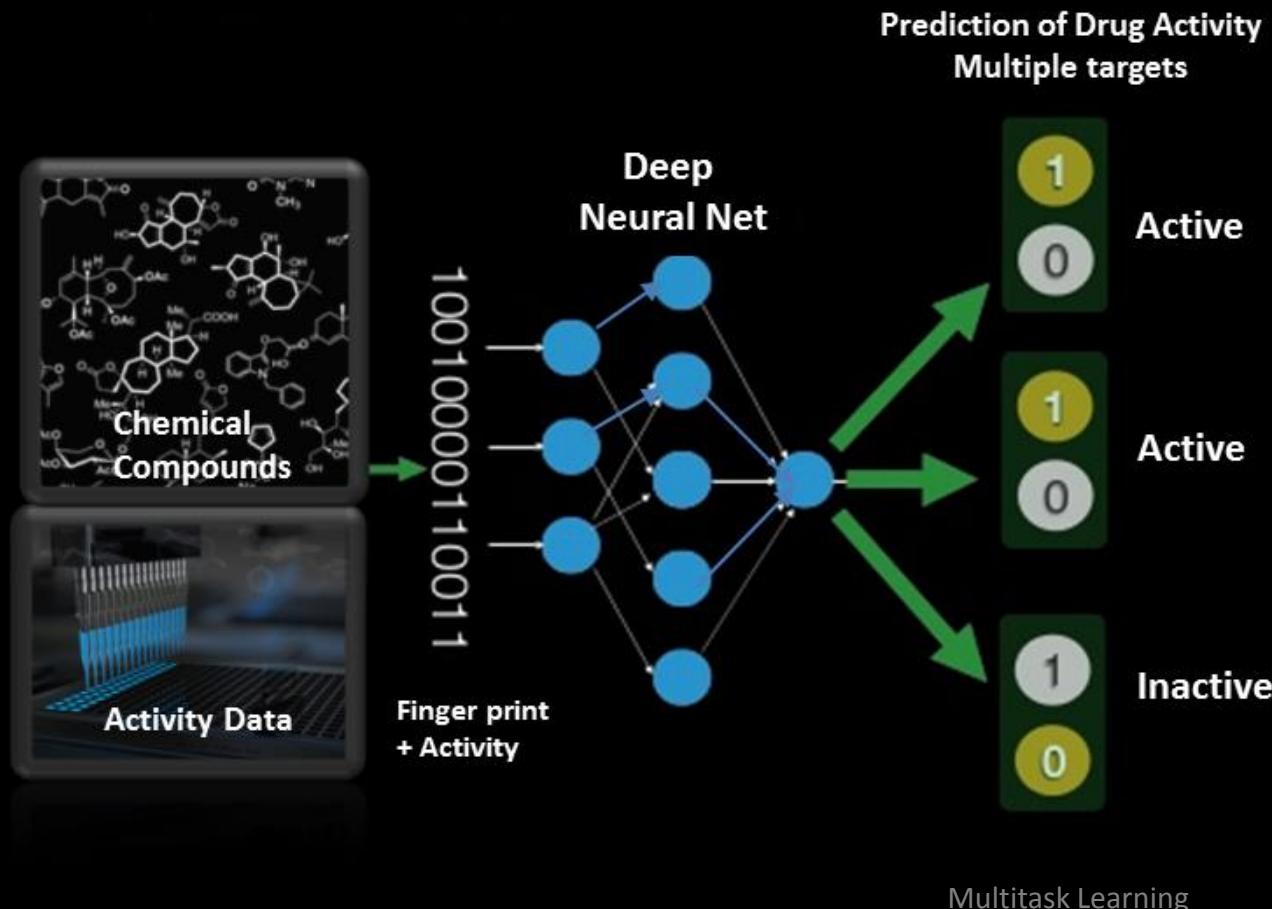


Experiment with Different Moves

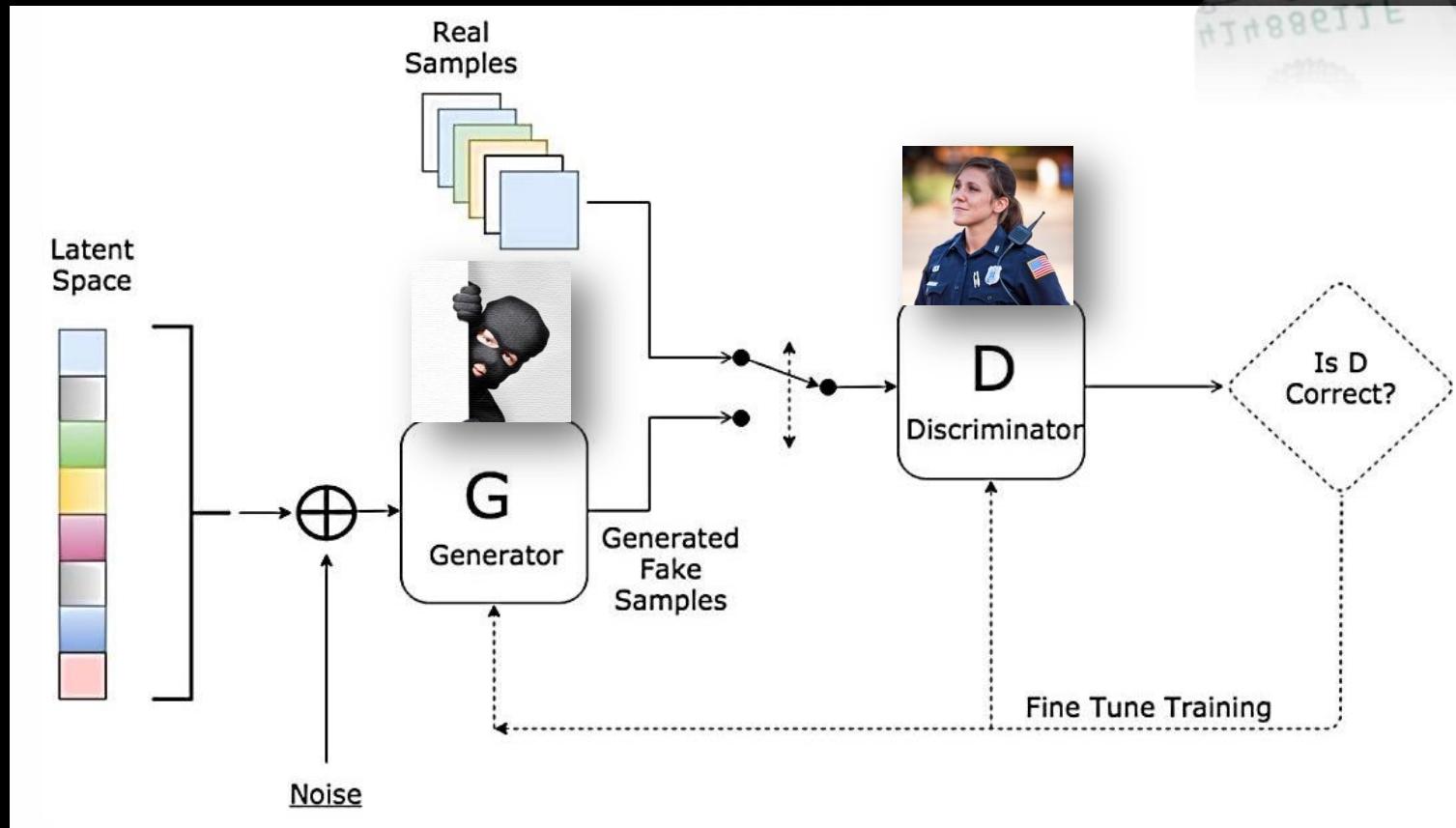


Interacting with the Environment

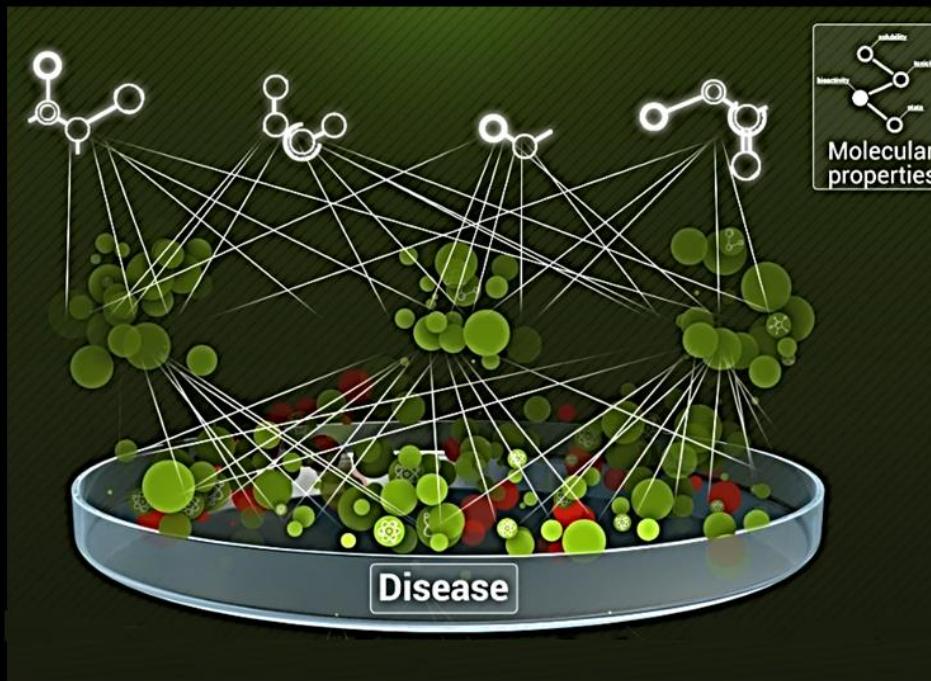
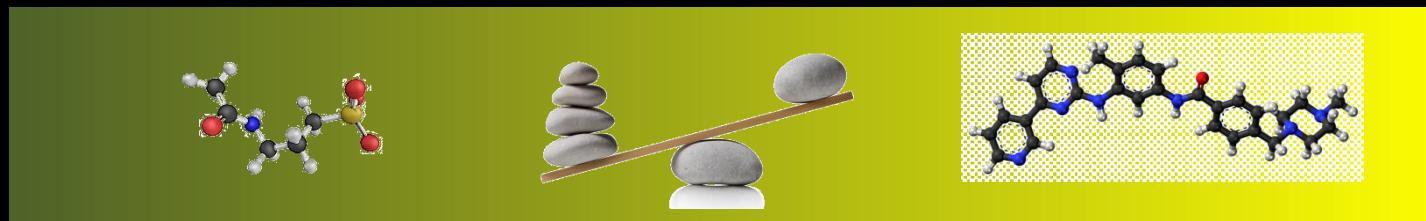
The Dawn of a New Era for Data Driven Medicine?



Generative Adversarial Networks



Drug Database



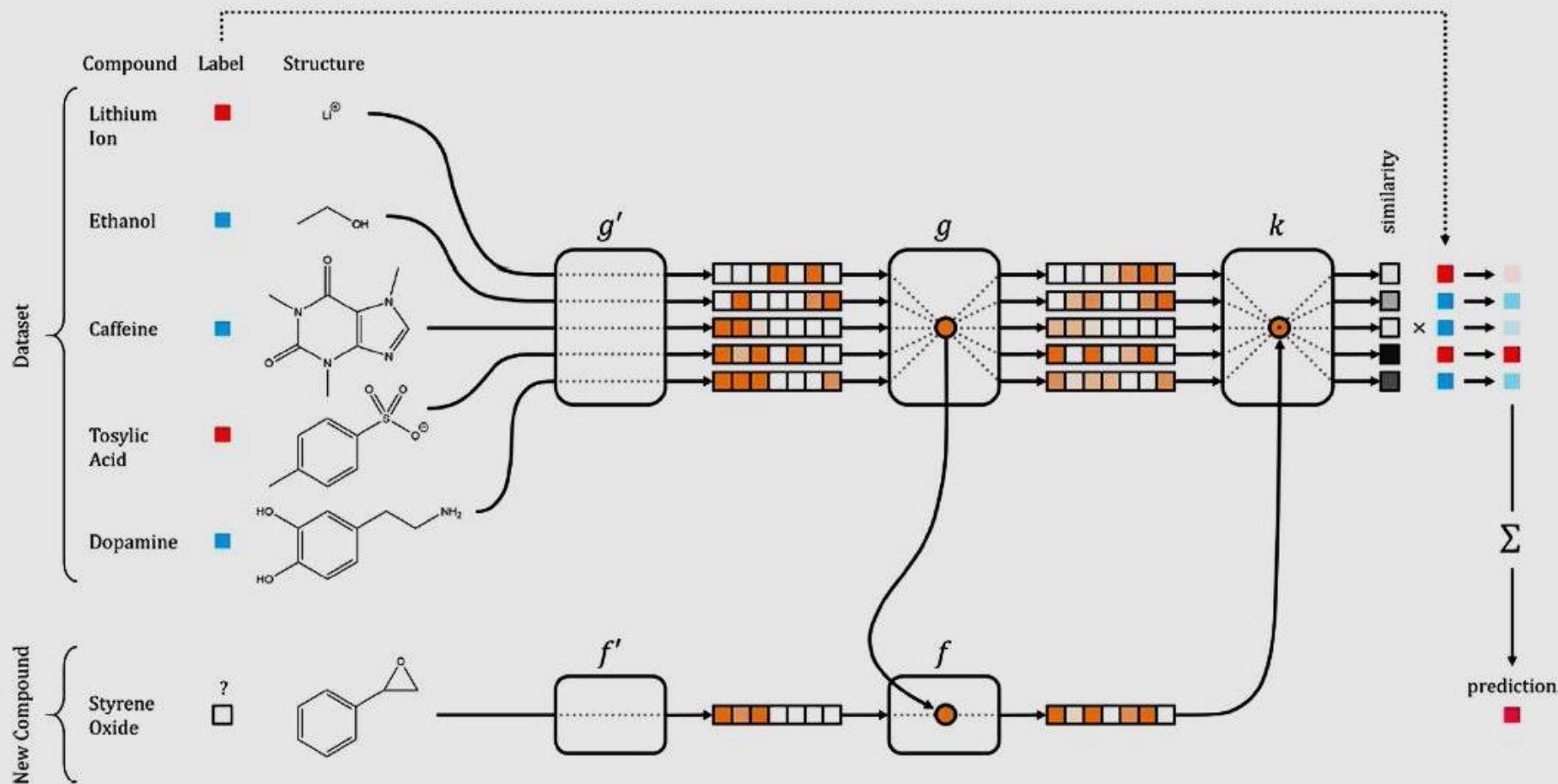
**Neural Networks for
Meaningful Leads**

One-Shot Learning

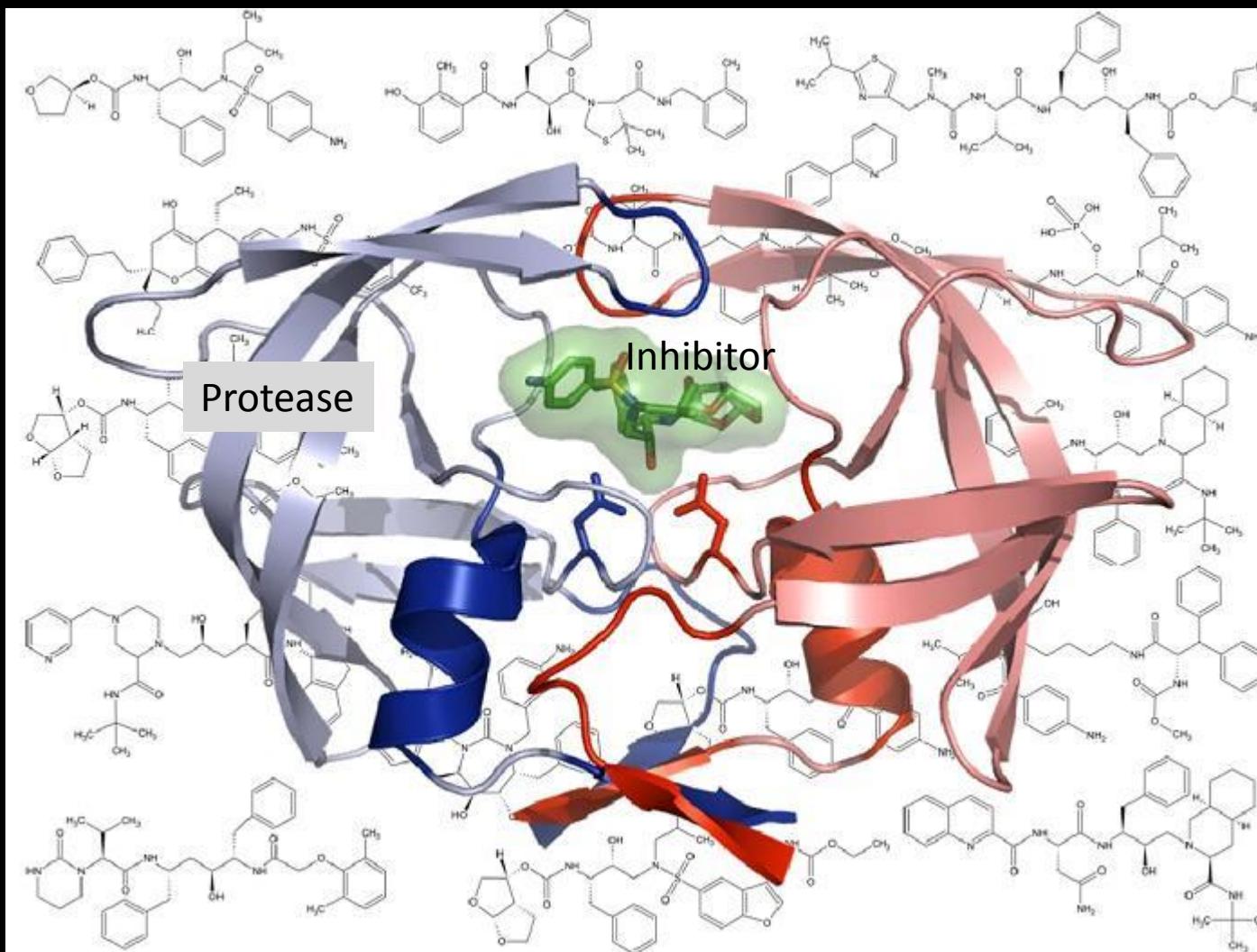


You show Giraffe once to a 3 year old kid, you do not need to teach the kid again

Deep Learning in Drug Discovery



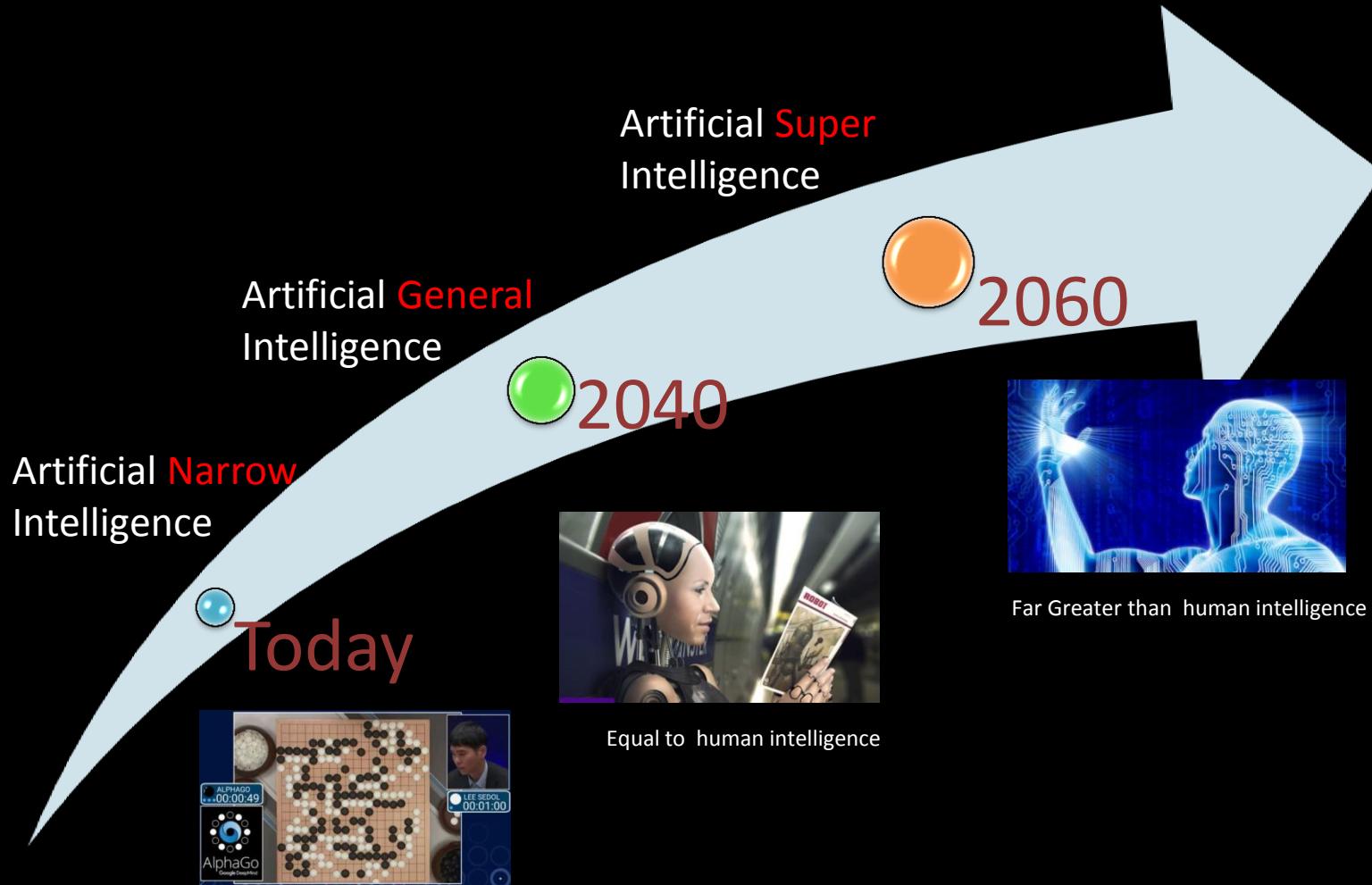
HIV Protease Inhibitors are Powerful **anti-AIDS** drugs



New Players are Entering the Healthcare Space



How to think about AI...





Thank You

The World is
Going to Change
With or Without
You...
Get Ready!

parthi@**vingyani**.com

[www.vingyani](http://www.vingyani.com).com

Chennai, Tamil Nadu, India