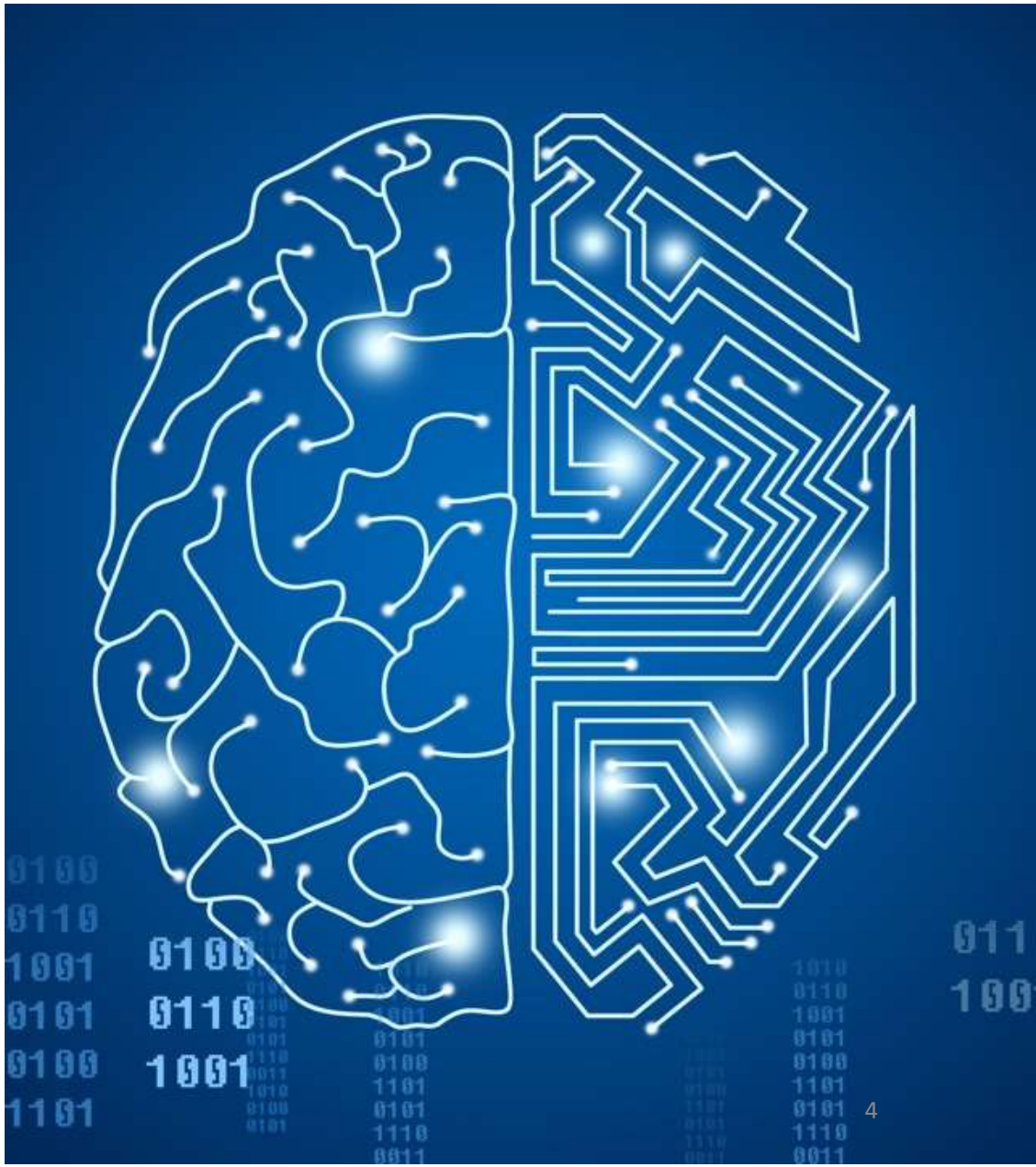
#AI4RD
Digital Health - Transformation of R&D and Pharmacology Therapy

Digital Pharmacology

Scuola di Specializzazione in Farmacologia e Tossicologia Clinica
Mar 13th 2019

What are the applications of artificial intelligence in drug discovery & development?



Antonio Pelliccia

IBM Business Development
Life Science

Artificial Intelligence

is NOT generalist

Artificial Intelligence has to be trained

by **expert people** with tons of **data**


Artificial Intelligence
is NOT a software to be installed

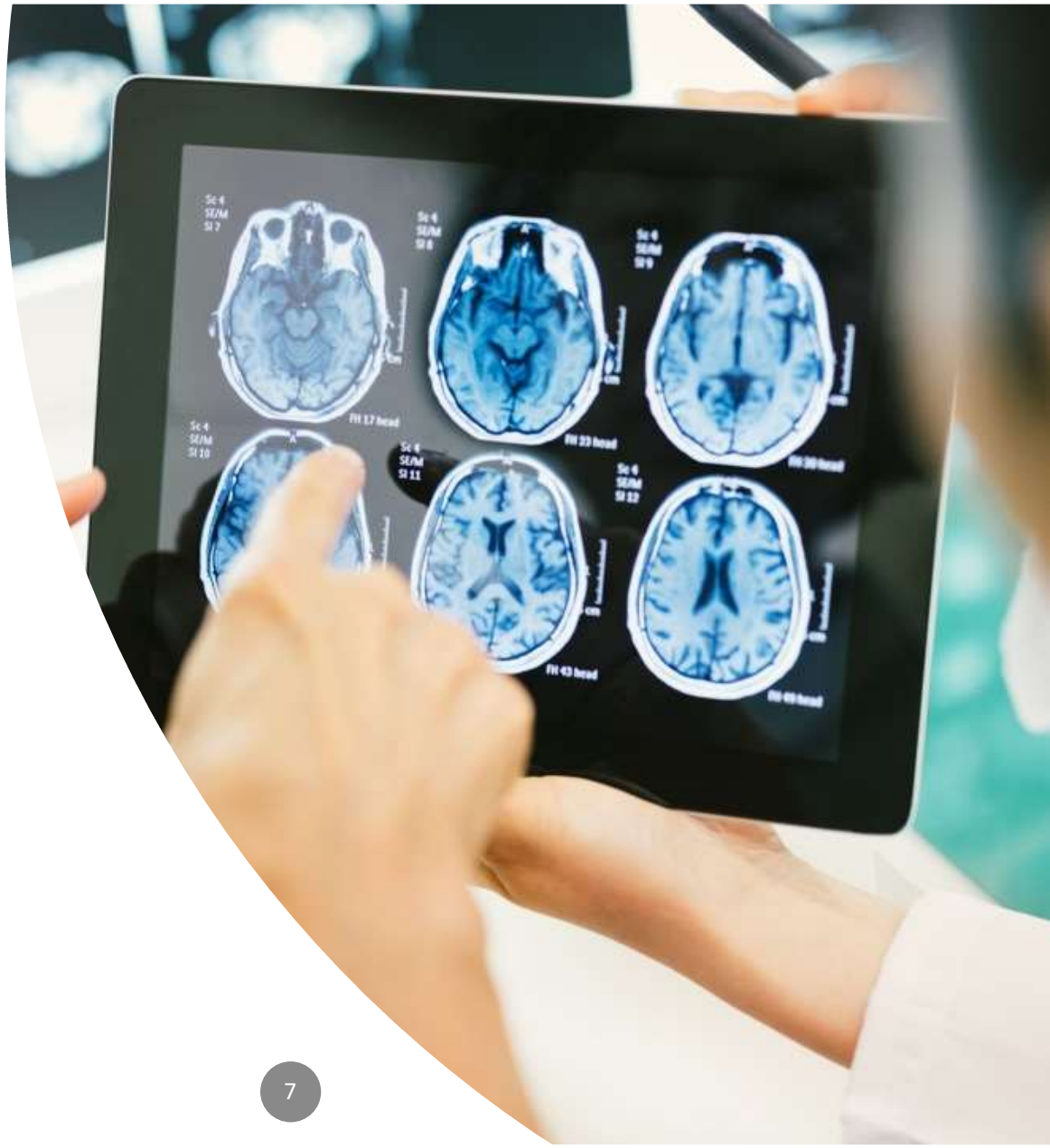
Artificial Intelligence needs

a new technology infrastructure and a new ecosystem


Artificial Intelligence
does not find something that does not exist

Artificial Intelligence

helps people to perform repetitive tasks **faster and easier**



What is AI?

"**Machine learning** is a field of computer science that gives computer systems the ability to "learn" from data, **without being explicitly programmed**"

Samuel, Arthur L. (1959). "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*

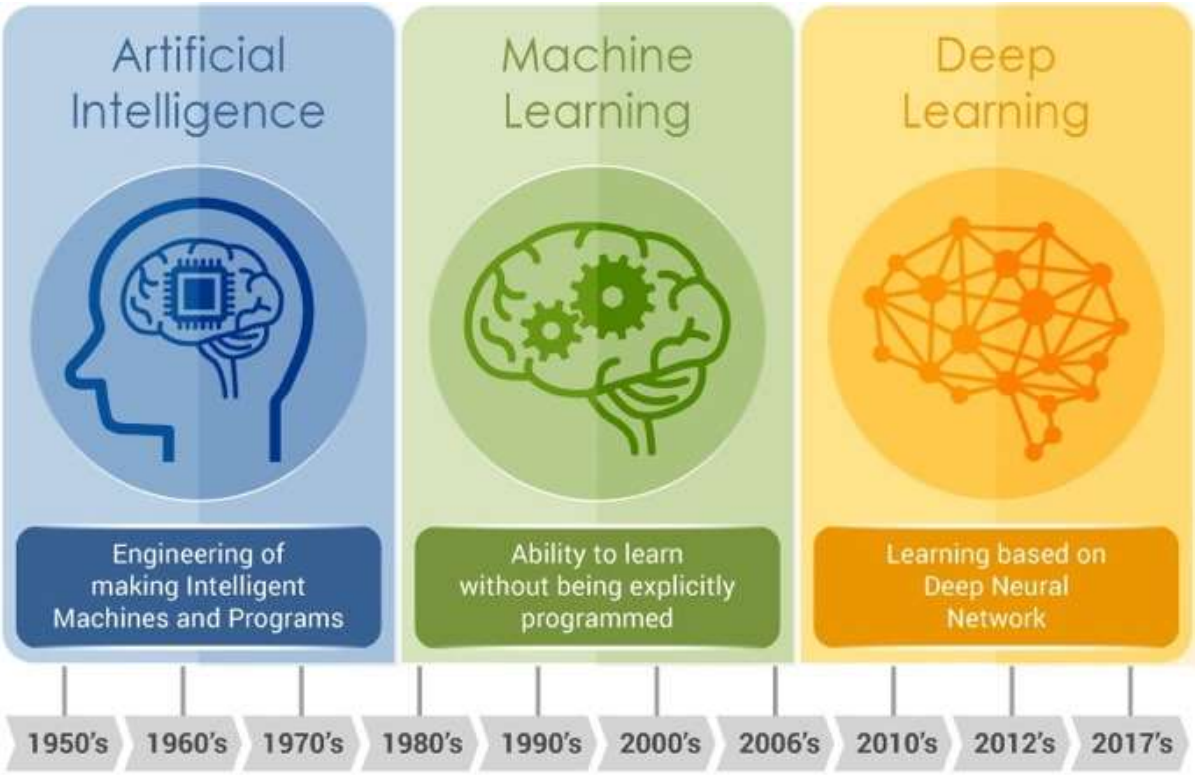
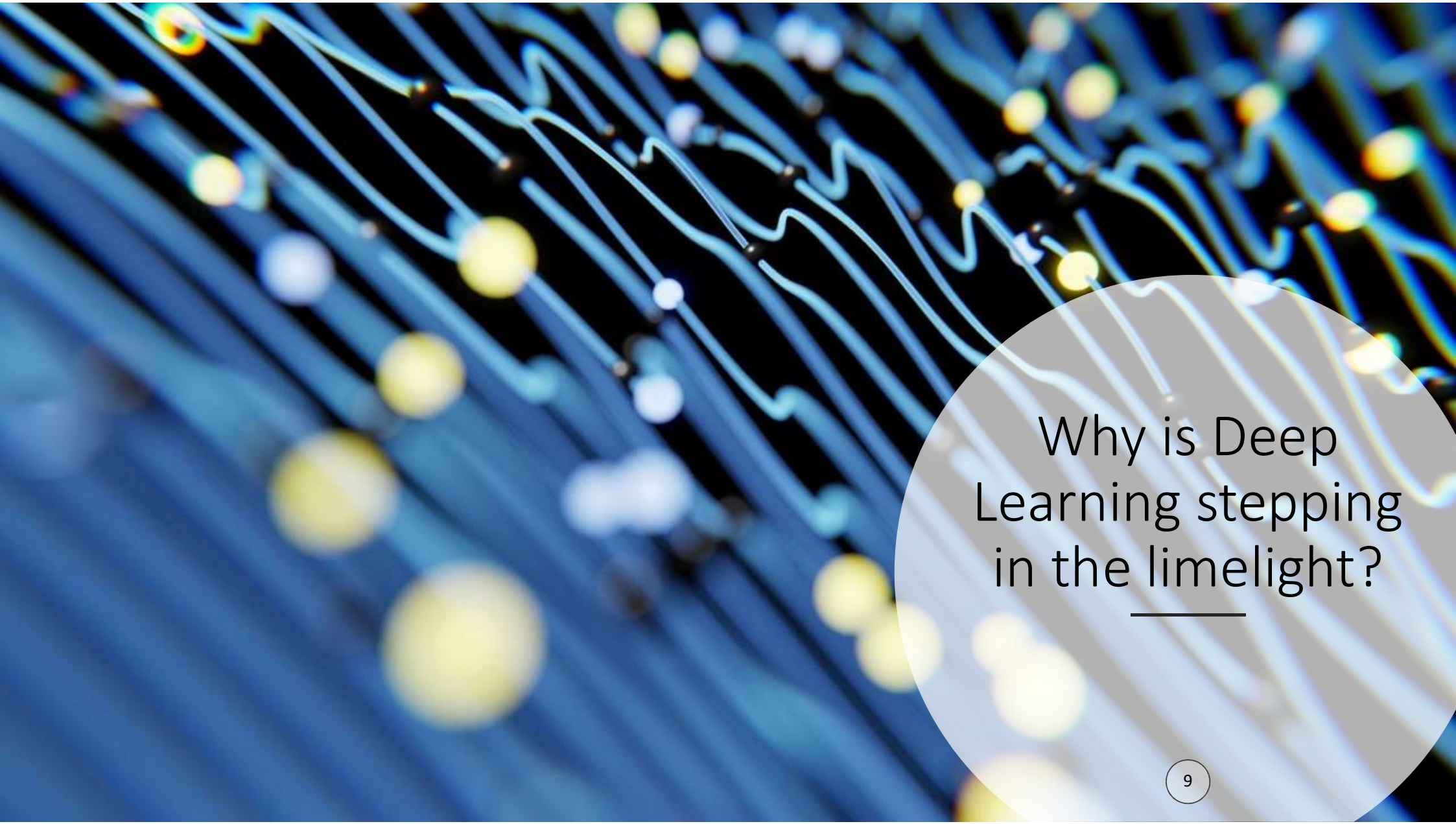


Image source: https://www.slideshare.net/linagora/deep-learning-in-practice-speech-recognition-and-beyond-meetup


Why is Deep Learning stepping in the limelight?

Healthcare Data are exploding
(from 153 Exabytes in 2013 to 2.314 Exabytes in 2020 - IDC)

Exogenous data
(Behaviour, Socio-economic, Environmental)

60% of determinants of health

Volume, Variety, Velocity
(1.100 TB per lifetime)

Genomics and Biologic data
30% of determinants of health
Volume (6TB per lifetime)

Clinical data
10% of determinants of health
Variety (0.4TB per lifetime)

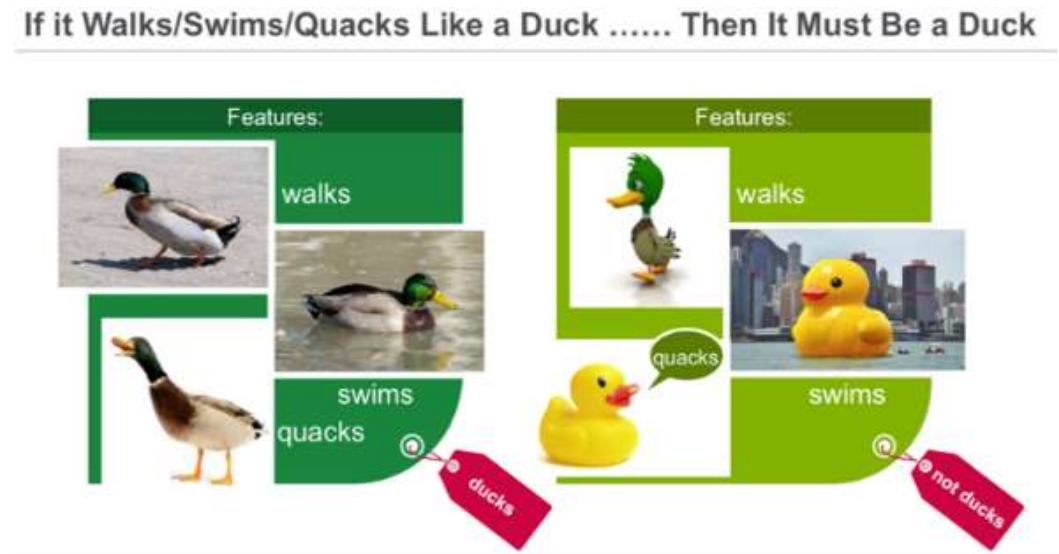
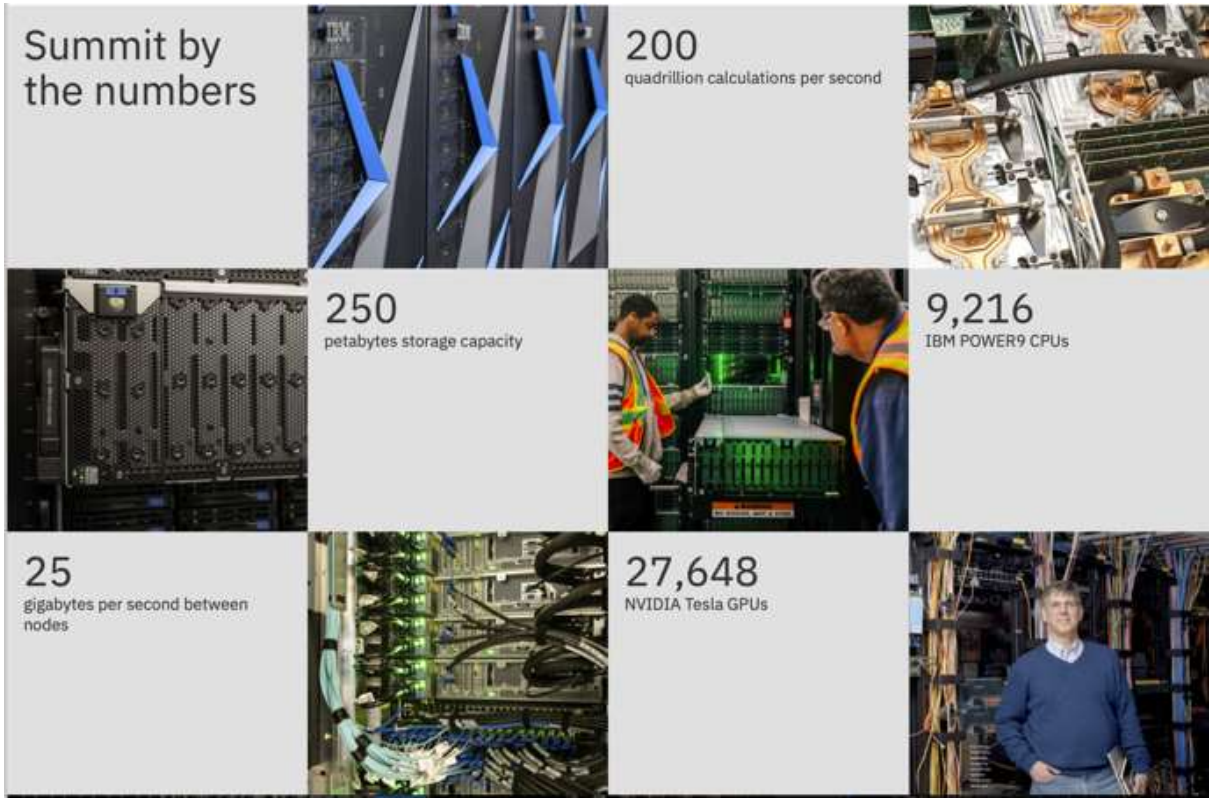


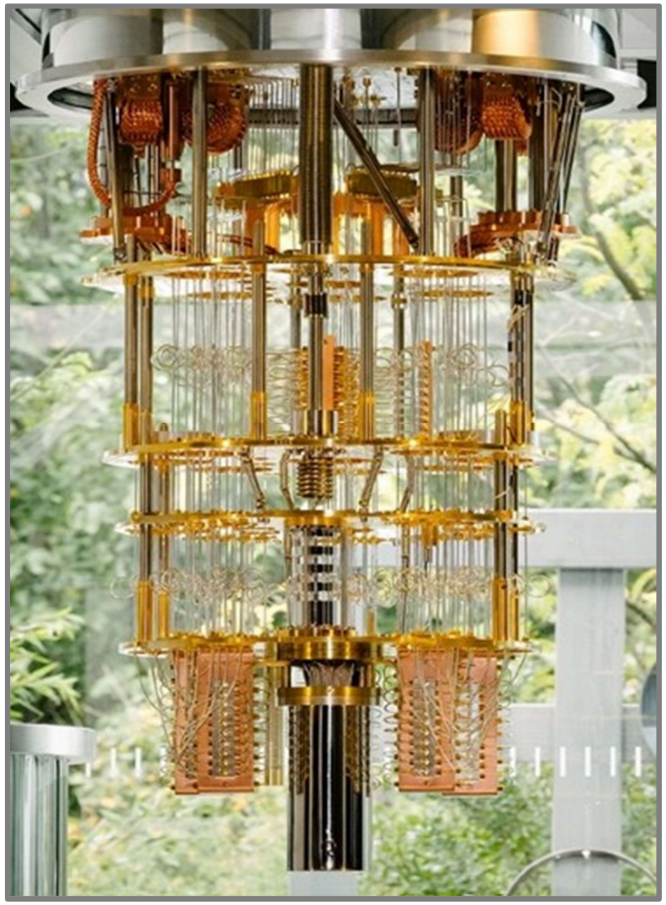
Image Source: https://goo.gl/B8Aaex

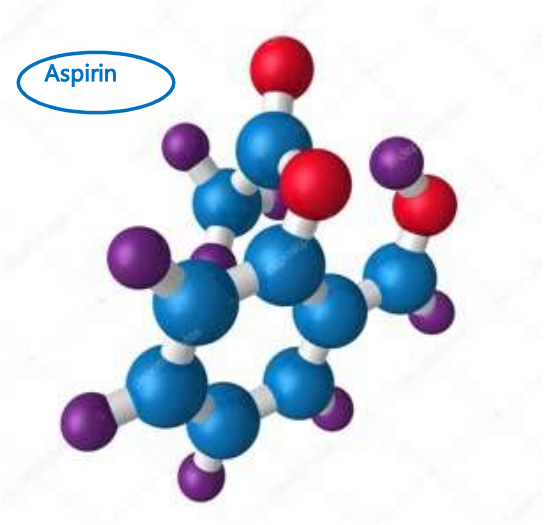
IBM builds the world's fastest supercomputers
What will we do with 200 petaflops?

- Machine learning algorithms scaled on Summit will help medical researchers with a comprehensive view of the cancer population
- Using a mix of AI techniques, researchers will be able to identify patterns in the function, cooperation, and evolution of human proteins and cellular systems



A future quantum processor could simulate a drug molecule – this would require a conventional computer larger than 10 percent of the size of the earth





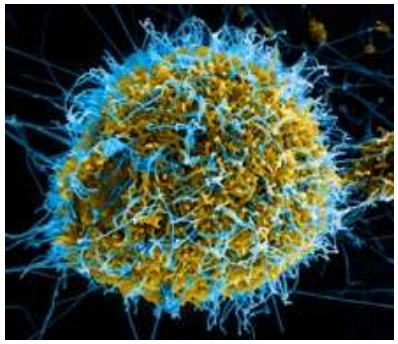
Type of Scaling	Time to Solve Problem				
Classical algorithm with exponential runtime	10 secs	2 mins	330 years	3300 years	Age of the Universe
Quantum algorithm with polynomial runtime	1 min	2 mins	10 mins	11 mins	~24 mins

Why should pharmacology use AI?

To help accelerate drug discovery, leading organisations and researchers are turning to AI because its ability to reveal hidden patterns and predict novel connections in biomedical data at a scale no human or traditional computing methods could possibly achieve
(IBM IBV study)

New applications of AI in drug discovery are emerging to transform the development process

Generating novel drug candidate

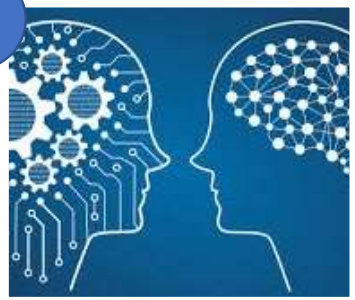


Understanding disease mechanism

Aggregating and synthesizing information

- AI for drug target identification and validation (Genentech/GNS, GSK/Insilico Medicine)
- AI for target based phenotypic drug discovery (GSK/Exscientia, Takeda/Numerate, Atomwise/IBM)
- AI for drug repurposing programs (Sanofi/Recursion, Astellas/Numedii)
- AI for biomarkers development (Sanofi/Berg Health)
- AI for analysing research literature, publications, patents (IBM Watson for Drug Discovery)

1



Machine Learning

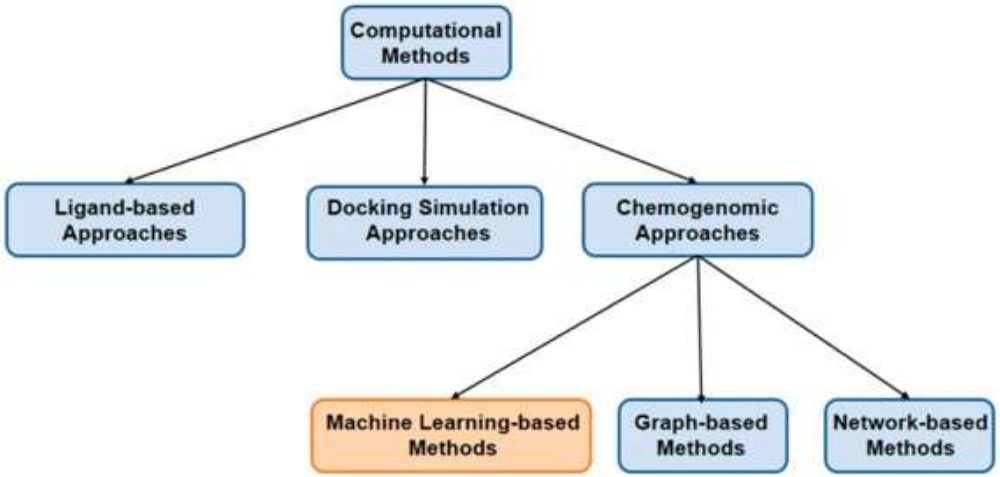
Text and unstructured data

2

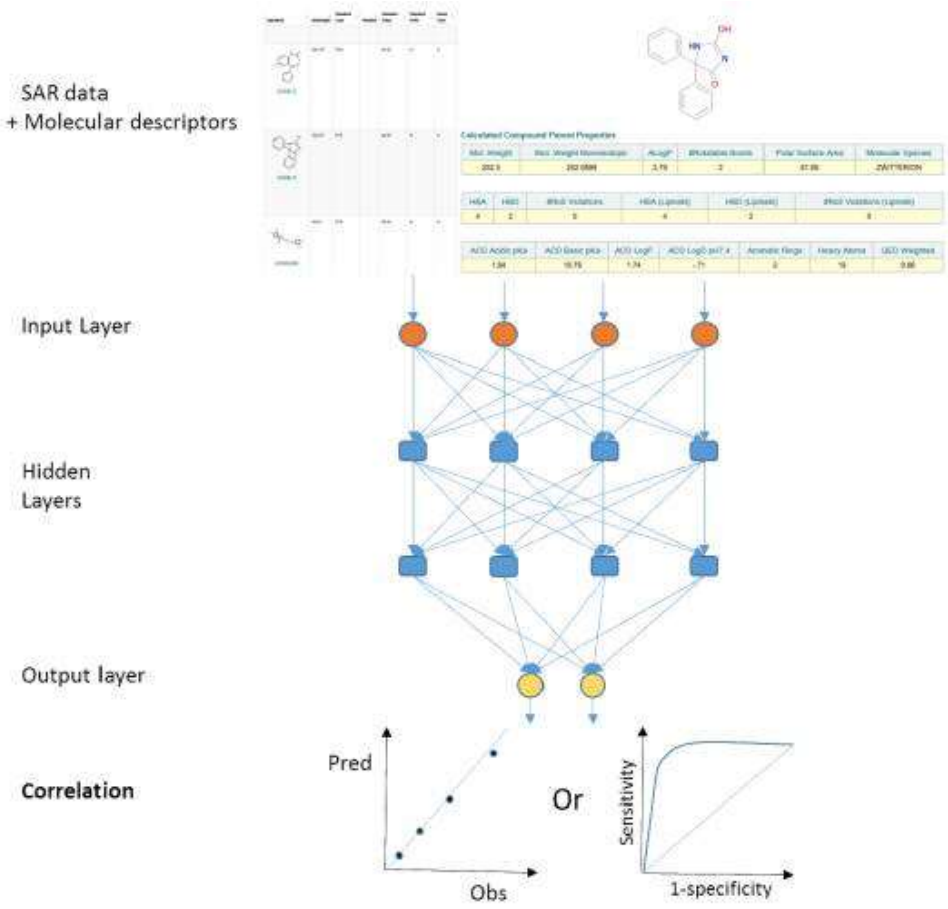
Deep Learning

Text, Images, DB

New methods are emerging for pharmacology research and development

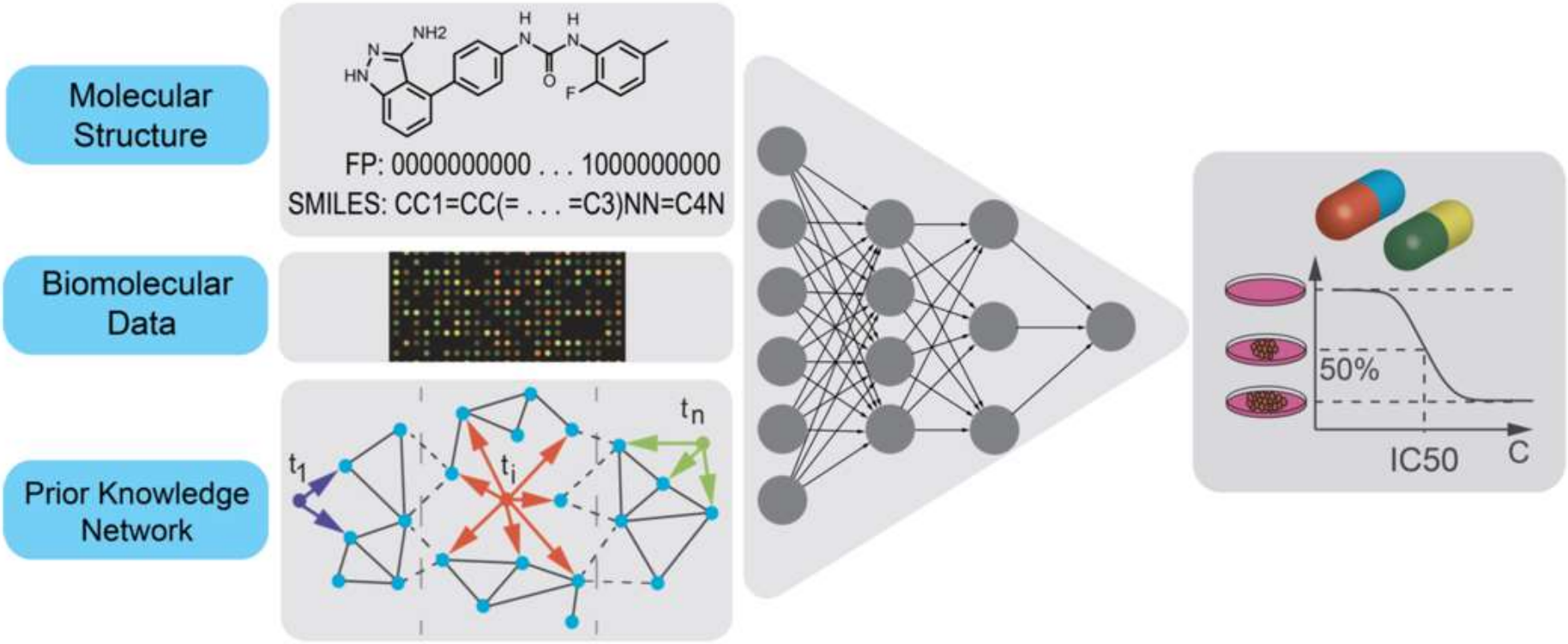


Source: Machine Learning for Drug-Target Interaction Prediction, Ruolan Chen, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6225477/



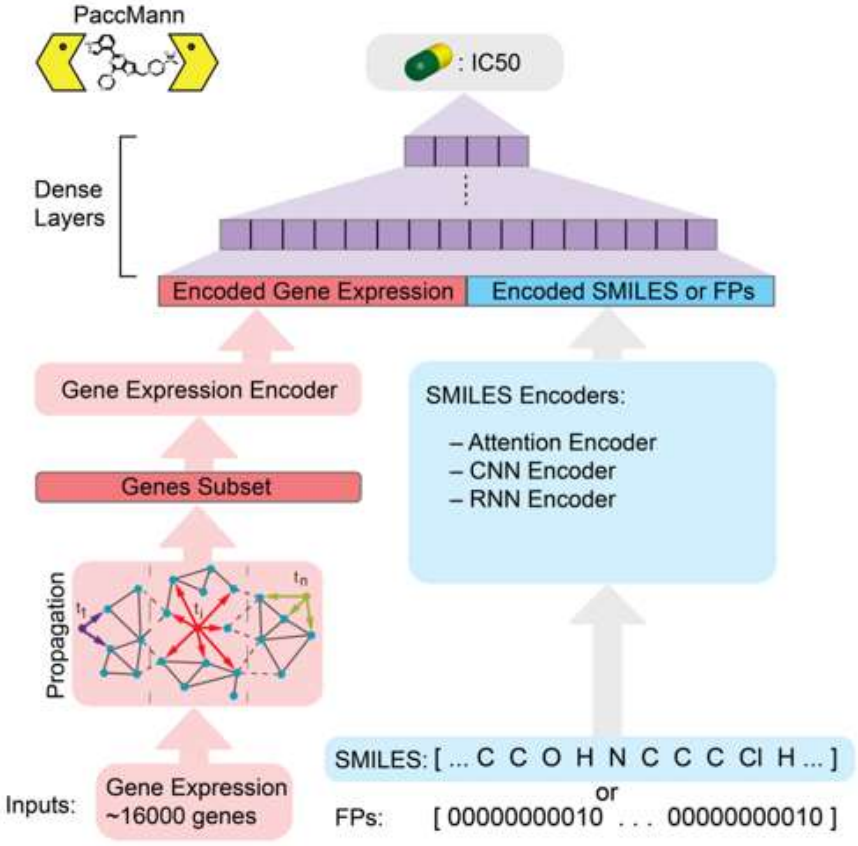
Source: The Next Era: Deep Learning in Pharmaceutical Research, Sean Ekins, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5042864/

Multi-modal prediction of IC50 drug sensitivity (PaccMann)



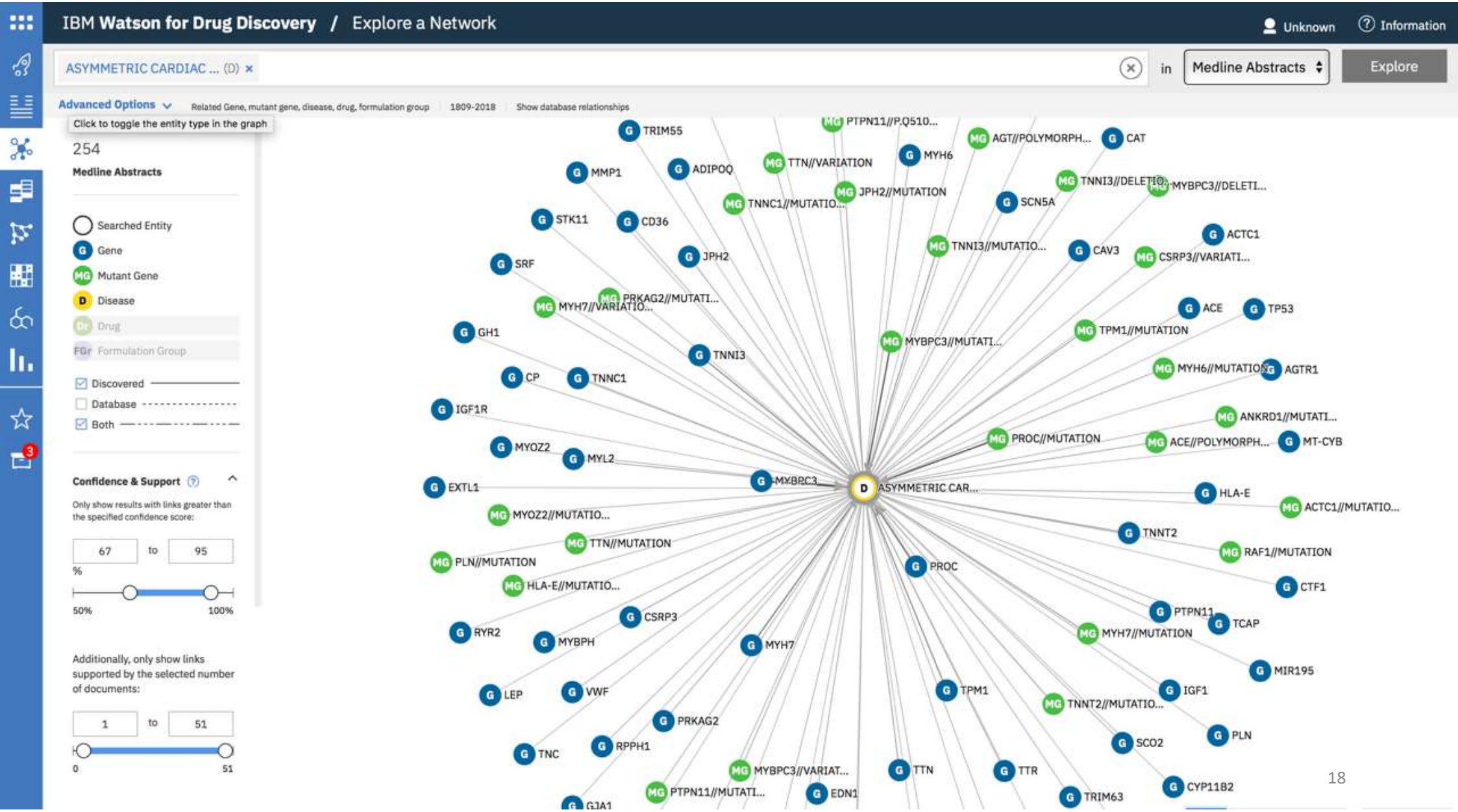
PaccMann: Prediction of anticancer compound sensitivity with multi-modal attention-based neural networks
IBM Research https://arxiv.org/pdf/1811.06802.pdf

The PaccMann algorithm allows to predict drug effectiveness starting from a specific molecular profile



Encoder type	Drug structure	RMSE Best	RMSE Average
Deep baseline (DNN)	Fingerprints	0.114	0.123 ± 0.008
Bidirectional recurrent (bRNN)	SMILES	0.106	0.118 ± 0.007
Stacked convolutional (SCNN)	SMILES	0.120	0.133 ± 0.012
Self-attention (SA)	SMILES	**0.089**	0.112* ± 0.007
Contextual attention (CA)	SMILES	0.095	**0.110* ± 0.008**
Multichannel convolutional attentive (MCA)	SMILES	0.106	0.120 ± 0.001

Drug	Cell line	Cancer type	Top-5 attended genes	IC50 Predicted	IC50 Measured
Afatinib	UMC-11	lung (NSCLC)	F13A1, MYH4, ATOH8, SEMA4A, NES	0.505	0.493
BX-912	YH-13	glioma	RNASE2, HOXA13, CBR3, FABP1, HDC	0.532	0.5
GSK319347A	EW-12	bone	CD300A, RHBDL2, NES, TFF3, SOCS1	0.597	0.7
JW-7-24-1	OVTOKO	ovary	HDC, EIF2A, RNASE2, ANGPTL6, CBR3	0.502	0.49
PI-103	MV-4-11	leukemia	TFF3, ATOH8, RBP2, ITIH3, GRIP1	0.362	0.33
TGX221	SW962	urogenital system	CBR3, RNASE2, FABP1, HDC, SH3D21	0.621	0.66
S-Trityl-L-cysteine	NCI-H187	lung (SCLC)	RHBDL2, NR1H4, MYH4, NES, APCS	0.535	0.502
Fedratinib	BL-41	lymphoma	TFF3, ATOH8, RBP2, MAPK7, ARHGEF33	0.382	0.428
Tipifarnib	RCC10RGB	kidney	EIF2A, HDC, CBR3, PIK3R5, HOXA13	0.542	0.544
Midostaurin	GAK	skin	SVOP, FABP1, HDC, F13A1, FGFR3	0.507	0.477


IBM Watson for Drug Discovery / Explore a Network
Unknown
Information
ASYMMETRIC CARDIAC ... (D) ×
in
Medline Abstracts
Explore
Advanced Options
Related Gene, mutant gene, disease, drug, formulation group
1809-2018
Show database relationships
Click to toggle the entity type in the graph
254
Medline Abstracts
Searched Entity
Gene
Mutant Gene
Disease
Drug
Formulation Group
Discovered
Database
Both
Confidence & Support
Only show results with links greater than the specified confidence score:
67
to
95
%
50%
100%
Additionally, only show links supported by the selected number of documents:
1
to
51
0
51
TRIM55
PTPN11//P.Q510...
AGT//POLYMORPH...
CAT
MMP1
ADIPOQ
TTN//VARIATION
MYH6
TNNI3//DELETIO...
MYBPC3//DELETI...
JPH2//MUTATION
TNNC1//MUTATIO...
SCN5A
STK11
CD36
ACTC1
TNNI3//MUTATIO...
CAV3
CSRP3//VARIATI...
JPH2
SRF
PRKAG2//MUTATI...
MYH7//VARIATIO...
ACE
TP53
GH1
TPM1//MUTATION
MYBPC3//MUTATI...
TNNI3
MYH6//MUTATION
AGTR1
CP
TNNC1
IGF1R
ANKRD1//MUTATI...
PROC//MUTATION
ACE//POLYMORPH...
MT-CYB
MYOZ2
MYL2
EXTL1
MYBPC3
D
ASYMMETRIC CAR...
HLA-E
ACTC1//MUTATIO...
MYOZ2//MUTATIO...
TNNT2
TTN//MUTATION
RAF1//MUTATION
PLN//MUTATION
PROC
HLA-E//MUTATIO...
CTF1
PTPN11
CSRP3
TCAP
RYR2
MYBPH
MYH7
MYH7//MUTATION
MIR195
LEP
VWF
TPM1
IGF1
TNNT2//MUTATIO...
PRKAG2
PLN
RPPH1
SCO2
TNC
MYBPC3//VARIAT...
TTN
TTR
CYP11B2
PTPN11//MUTATI...
EDN1
TRIM63
GJA1

tions of key driver genes in colorectal cancer ession and metastasis.

DLINE®/PubMed®, a database of the U.S. National Library of Medicine.

t publication name:	Cancer metastasis reviews
t publication date:	Mar 1, 2018
le issue:	1
MID:	29322354
ocument identifier:	10.1007/s10555-017-9726-5
d common main author list:	Huang,Dongdong; Sun,Wenjie; Zhou,Yuwei; Li,Peiwei; Chen,Fang; Chen,Hanwen; Xia,Dajing; Xu,Enping; Lai,Maode; Wu,Yihua; Zhang,Honghe

dence for **BRAF//MUTATION**

nd in **Abstract**

group analysis stratified by ethnic populations indicated that the BRAF mutation was related to CRC astasis (combined OR 1.42, 95% CI 1.18-1.71) and distant metastasis (combined OR 1.51, 95% CI 1.20-) in an Asian population.

to snippet

ıct

ciation between mutations of key driver genes and colorectal cancer (CRC) metastasis has estigated by many studies. However, the results of these studies have been contradictory. Here, rm a comprehensive analysis to screen key driver genes from the TCGA database and validate the hese mutations in CRC metastasis. Using bioinformatics analysis, we identified six key driver amely APC, KRAS, BRAF, PIK3CA, SMAD4 and p53. Through a systematic search, 120 ublished by November 30, 2017, were included, which all showed roles for these gene mutations metastasis. A meta-analysis showed that KRAS mutations (combined OR 1.18, 95% CI 1.05- d p53 mutations (combined OR 1.49, 95% CI 1.23-1.80) were associated with CRC asis, including lymphatic and distant metastases. Moreover, CRC patients with a KRAS (combined OR 1.29, 95% CI 1.13-1.47), p53 mutation (combined OR 1.35, 95% CI 1.06- SMAD4 mutation (combined OR 2.04, 95% CI 1.41-2.95) were at a higher risk of distant asis. **Subgroup analysis stratified by ethnic populations indicated that the BRAF mutation ated to CRC metastasis (combined OR 1.42, 95% CI 1.18-1.71) and distant metastasis ed OR 1.51, 95% CI 1.20-1.91) in an Asian population.** No significant association was found mutations of APC or PIK3CA and CRC metastasis. In conclusion, mutations of KRAS, MAD4 and BRAF play significant roles in CRC metastasis and may be both potential ers of CRC metastasis as well as therapeutic targets.

s Found Within This Document

s)	Found in text as
	KRAS
	BRAF
	PIK3CA
	SMAD4
	p53

TECHNOLOGY HELPS SCIENTISTS DISCOVER POTENTIAL NEW TREATMENTS FOR ALS

BARROW Neurological Institute

The challenge

Millions of pages of research, nearly 1,500 possible target proteins and wildly disparate clinical data made progress extremely slow for scientists seeking new drug therapies for ALS.

The transformation

The IBM Watson™ for Drug Discovery platform is helping Barrow Neurological Institute narrow research scope and uncover new pathways of interest for drug therapies in the fight against ALS.

The results

5 new proteins identified in months rather than years by analysing large amounts of disparate data more quickly than traditional methods

80% of top-ranked targets were proven to be linked to ALS

Identifies new pathways of interest for drug therapies that scientists may not have considered otherwise

Thanks

- Antonio Pelliccia