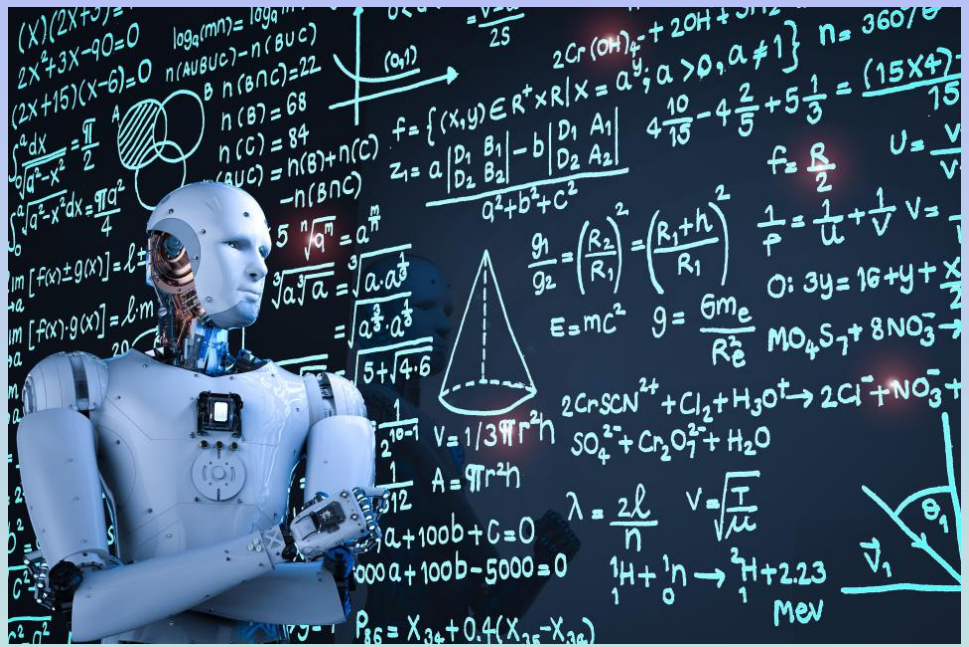
Artificial Intelligence for Health

Erwan Scornet

(Assistant Professor, Ecole Polytechnique)

1. History of AI
2. From Big Data to Deep Learning
3. AI and Health
   - Chest X-ray
   - Liver lesion segmentation
   - Genomics
   - Toxicogenetics
   - Medical
4. Perspective and issues

A first look at Artificial Intelligence



- What is Artificial Intelligence?
- What are the applications of AI?
- What are the main challenges?
- What are the issues raised by AI?

Definition of AI - Dartmouth conference

On September 1955, a project was proposed by McCarthy, Marvin Minsky, Nathaniel Rochester and Claude Shannon introducing formally for the first time the term "Artificial Intelligence".

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.

Proposal for Dartmouth conference on AI (1956)

Old applications of AI

Many tasks were achieved between 1956 and 1974.

- Computer checkers (1959, Arthur Lee Samuel) capable of challenging a respectable amateur. Interestingly, he coined the term "Machine Learning" in 1959.
- IBM Shoebox (1961) was able to recognize 16 spoken words and the digits 0 to 9.
- ELIZA (1964-1966 at MIT, Joseph Weizenbaum) was one of the first chatterbots and one of the first programs capable of attempting the Turing Test.
- WABOT-1 (Waseda University, 1972) was the world's first full-scale intelligent android.

In from three to eight years we will have a machine with the general intelligence of an average human being.

Marvin Minsky (1970, Life Magazine)

Misconception of AI

AI is about electronic device able to mimic human thinking:

- Artificial intelligence
- One famous class of AI algorithms are called neural networks.
- Android are close to humans in shape so they must think like humans.

Term invented by John McCarthy in 1956: *the science and engineering of making intelligent machines*

Most AI algorithms do **not** aim at reproducing human reasoning.

"the study and design of intelligent agents" where an intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success

Frequent definition

Artificial Intelligence is not human intelligence

What often happens is that an engineer has an idea of how the brain works (in his opinion) and then designs a machine that behaves that way. This new machine may in fact work very well. But, I must warn you that that does not tell us anything about how the brain actually works, nor is it necessary to ever really know that, in order to make a computer very capable. It is not necessary to understand the way birds flap their wings and how the feathers are designed in order to make a flying machine [...] **It is therefore not necessary to imitate the behavior of Nature in detail in order to engineer a device which can in many respects surpass Nature's abilities.**

Richard Feynman (1999)

Artificial intelligence aims at solving difficult tasks

Artificial intelligence is the science of making machines do things that would require intelligence if done by men

Marvin Minsky (1968)

2001: A Space Odyssey

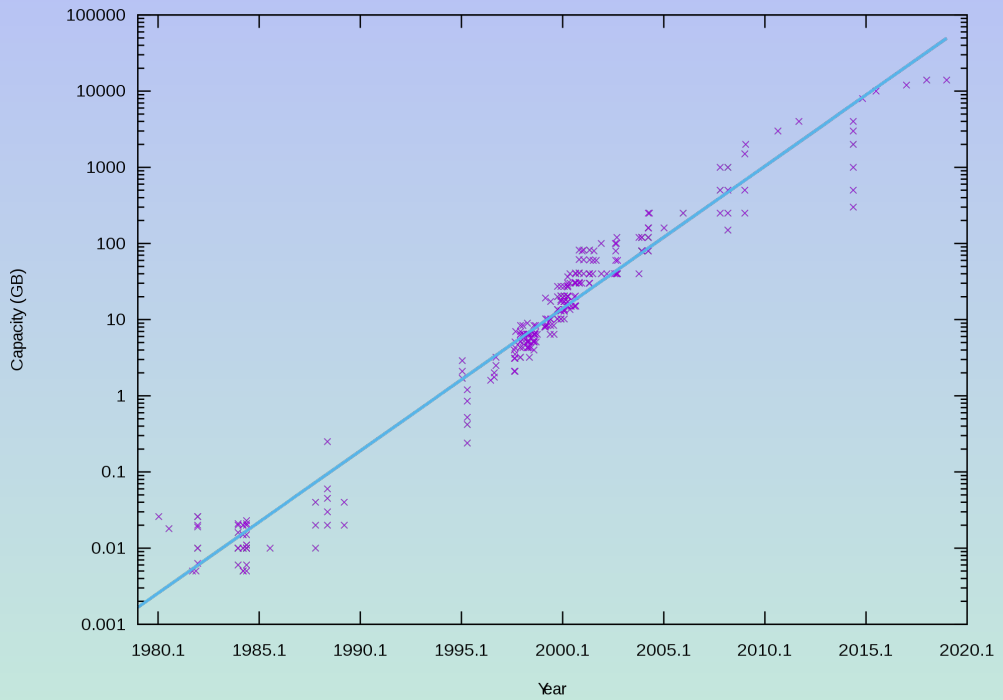
AI technology - Autonomous cars

- Originates from 1920 (NY)
- First use of neural networks to control autonomous cars (1989)
- Four US states allow self-driving cars (2013)
- First known fatal accident (May 2016)
- Singapore launched the first self-driving taxi service (Aug. 2016)
- A Arizona pedestrian was killed by an Uber self-driving car (March 2018).

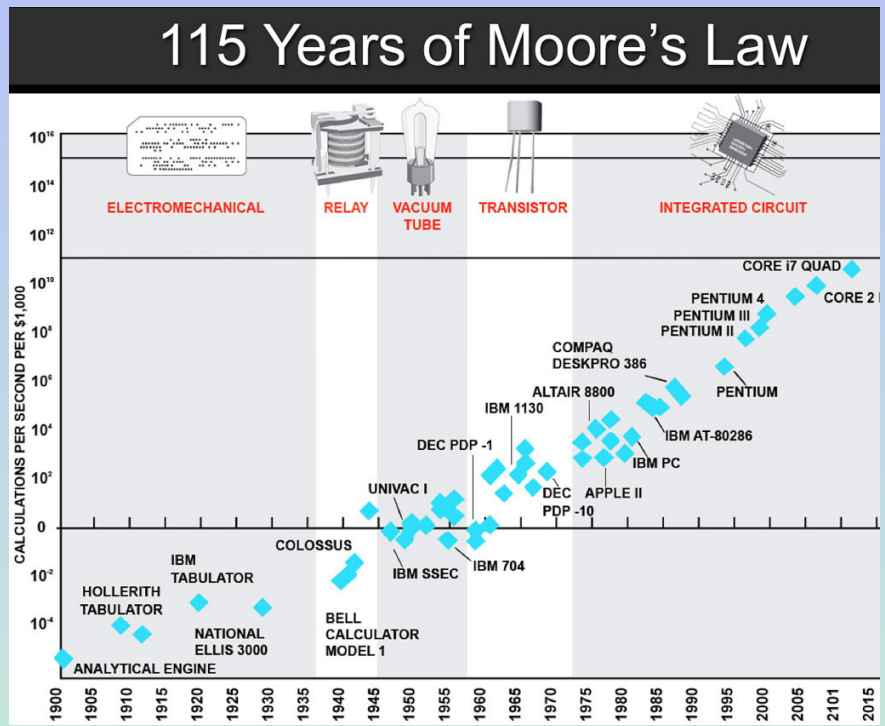
AI technology - virtual assistant / chatbot

- Voice recognition tool "Harpy" masters about 1000 words (1970s, CMU, US Defense).
- System capable of analyzing entire word sequences (1980).
- Siri was the first modern digital virtual assistant installed on a smartphone (2011).
- Watson won the TV show Jeopardy! (2011)

Storage and processing capacities

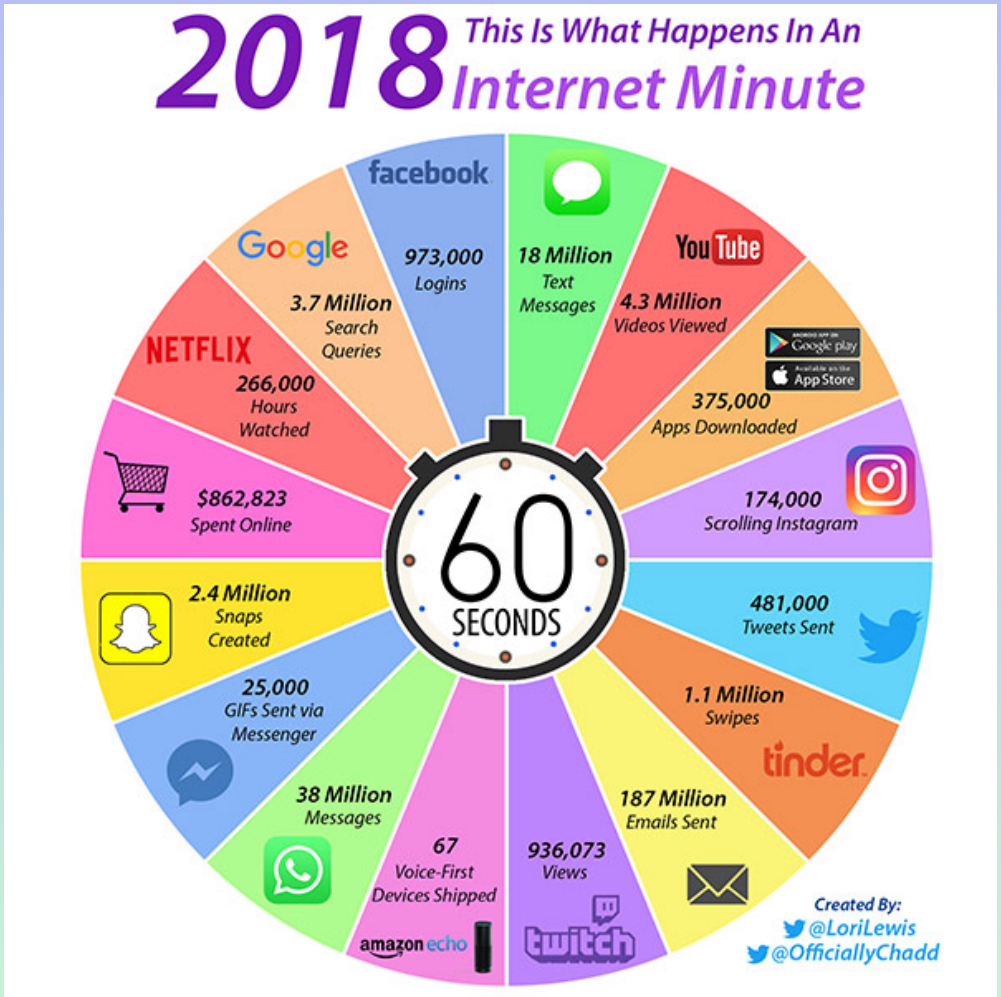


Storage capacity (Kryder's law)



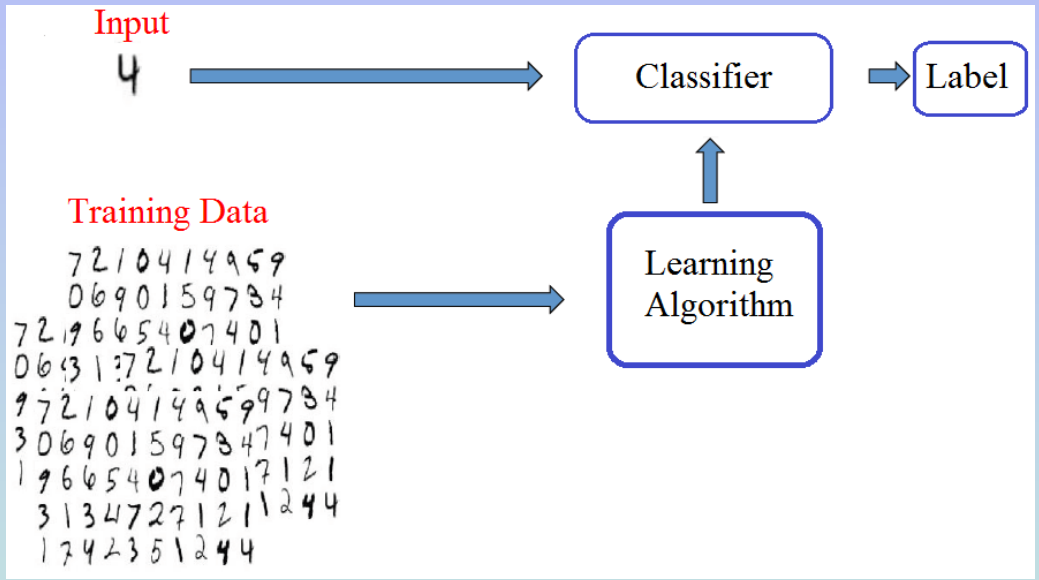
Processor capacity (Moore's law)

Role of data



Data are necessary and at the core of Machine Learning.

Machine Learning Pipeline



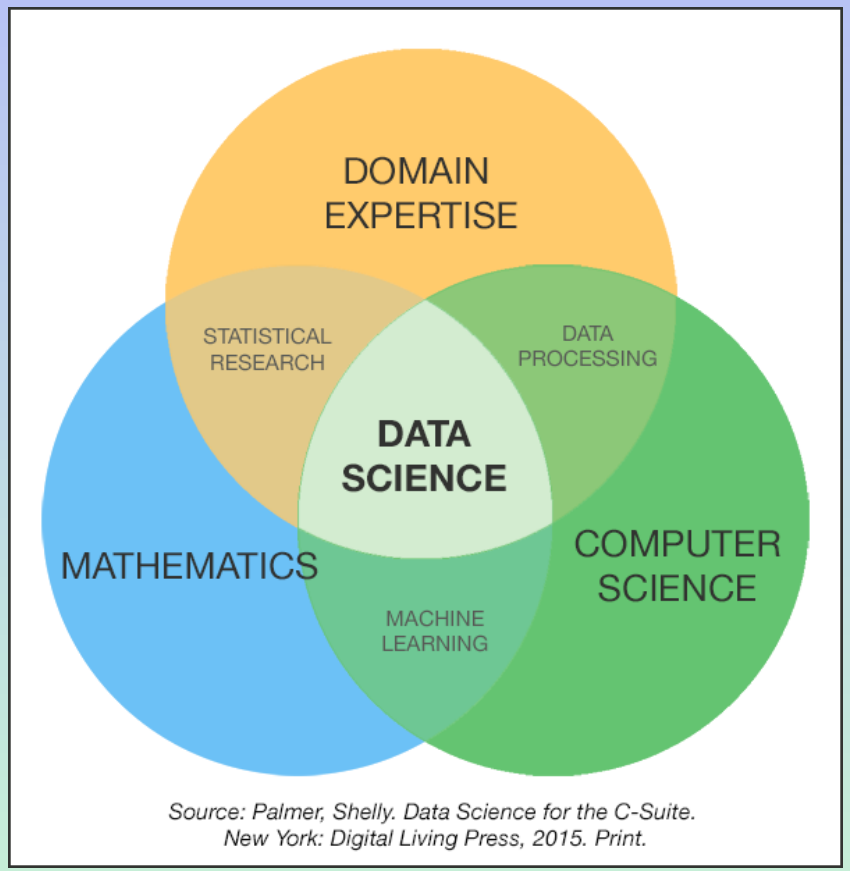
A definition by Tom Mitchell (http://www.cs.cmu.edu/~tom/)

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T , as measured by P, improves with experience E.

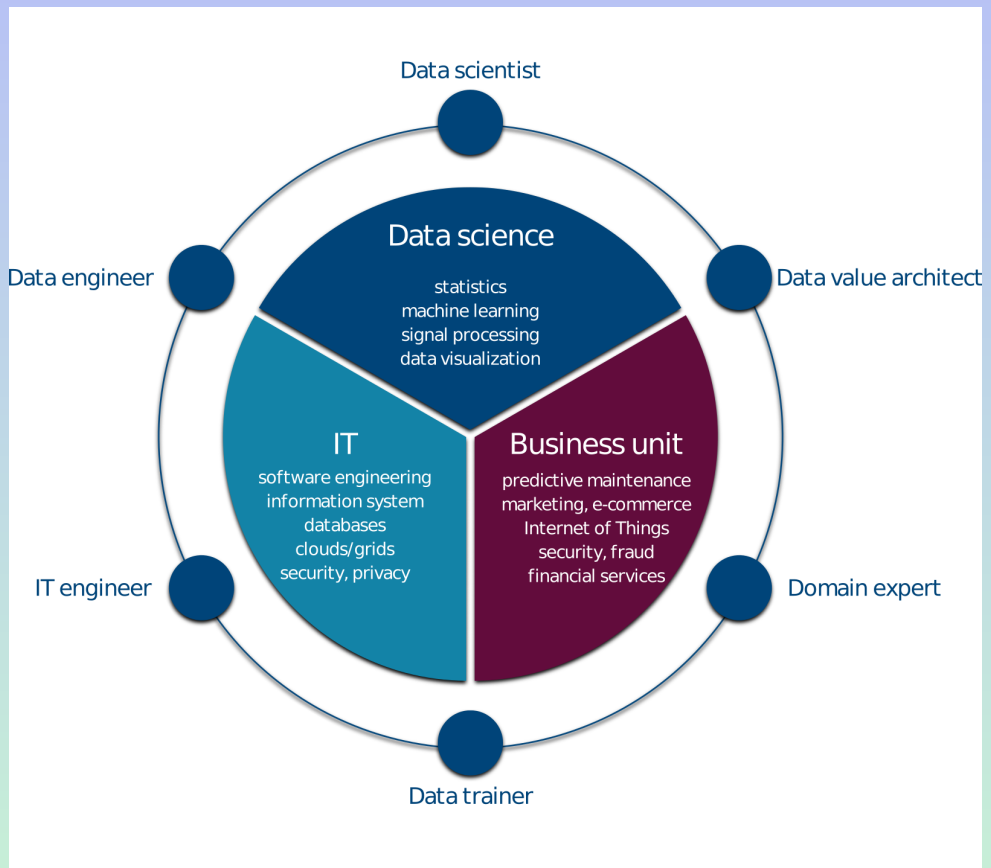
Difficulties related to (Big) data

- The prediction must be accurate: difficult for some tasks like image classification, video captioning...
- The prediction must be quick: online recommendation should not take minutes.
- Data must be stored and accessible easily.
- It may be difficult to access all data at the same time. Data may come sequentially.
- Data must be cleansed.
- Data should be relevant.

Different training and different occupations



Source: Palmer, Shelly. Data Science for the C-Suite. New York: Digital Living Press, 2015. Print.



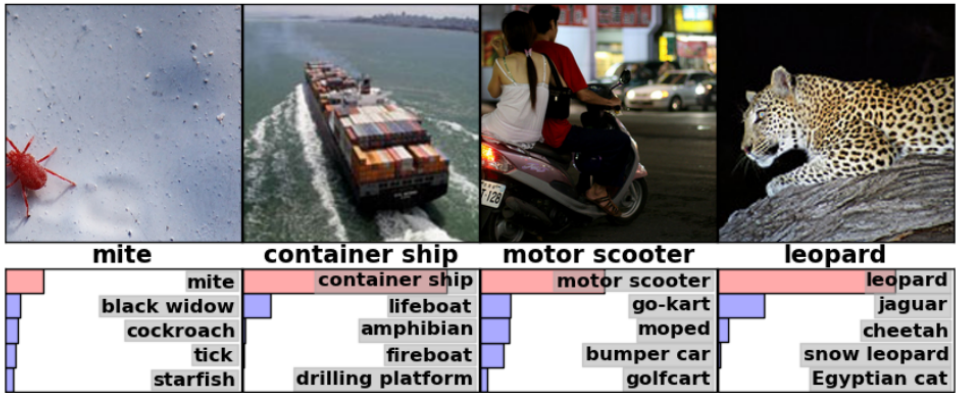
Buzzwords

- **Big data** is an all-encompassing term for any collection of data sets so large and complex that it becomes difficult to process using traditional data processing applications.
- **Statistics** is the study of the collection, analysis, interpretation, presentation and organization of data.
- **Machine learning** is the subfield of computer science that gives computers the ability to learn without being explicitly programmed.
- **Artificial Intelligence** research is defined as the study of *intelligent agents*: any device that perceives its environment and takes actions that maximize its chance of success at some goal.
- **Data science** is the study of the generalizable extraction of knowledge from data, yet the key word is science.

Baron Schwartz, Twitter, Nov 2017

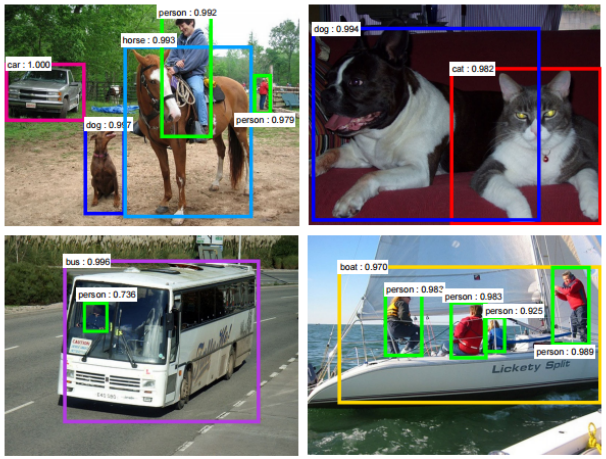
When you're fundraising, it's AI / When you're hiring, it's ML / When you're implementing, it's linear regression / When you're debugging, it's printf()

Deep Learning: Vision



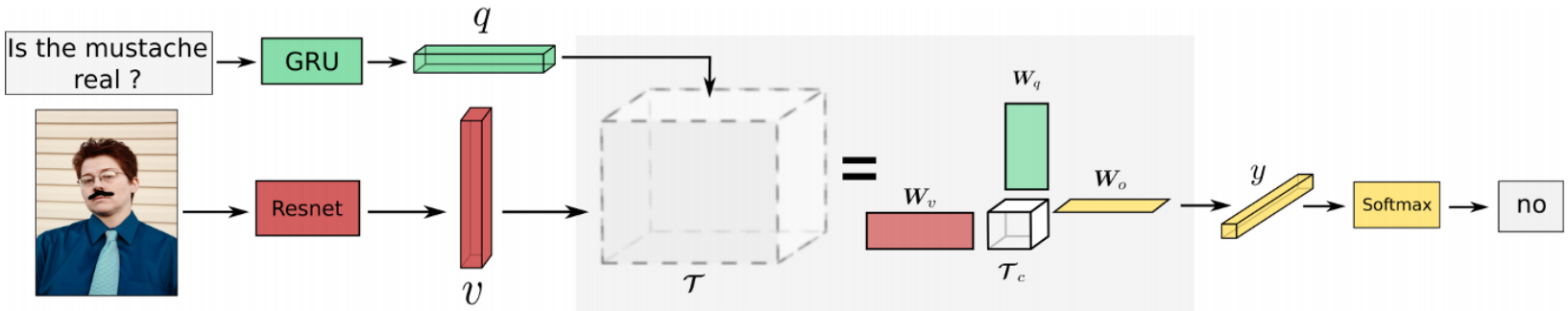
[Krizhevsky 2012]

[Ciresan et al. 2013]



[Faster R-CNN - Ren 2015]

[NVIDIA dev blog]



[VQA - Mutan 2017]

"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

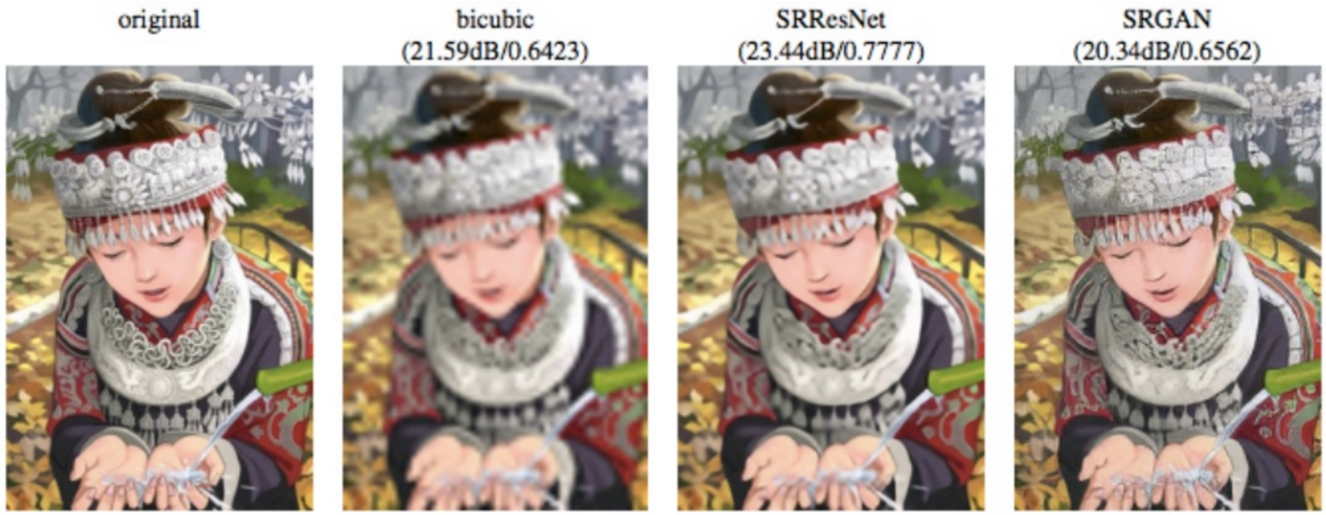
"two young girls are playing with lego toy."

"boy is doing backflip on wakeboard."

[Karpathy 2015]

[DeepDream 2015]

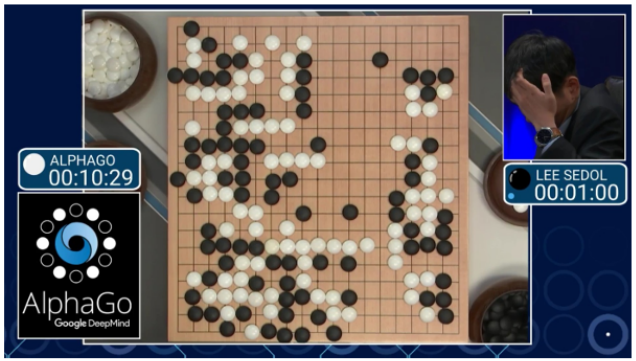
[Gatys 2015]

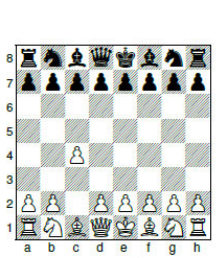


[Ledig 2016]

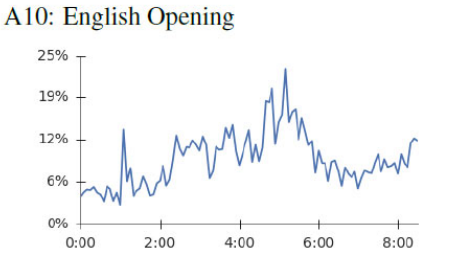
Deep Learning: Generative Models

Deep Learning: Generative Models



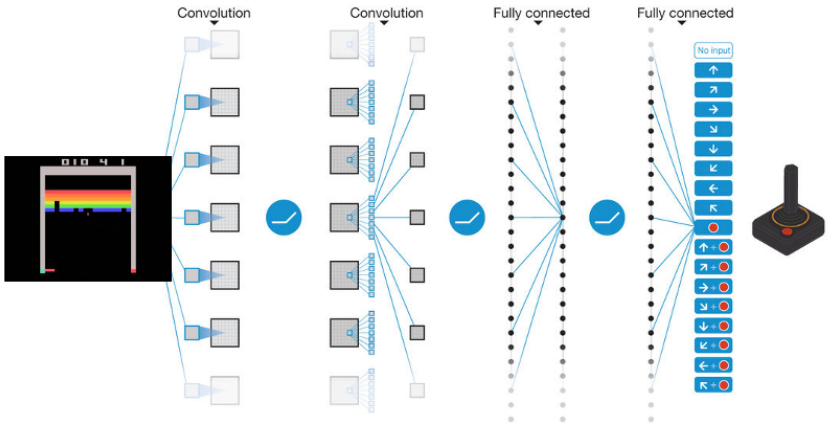


w 20/30/0, b 8/40/2

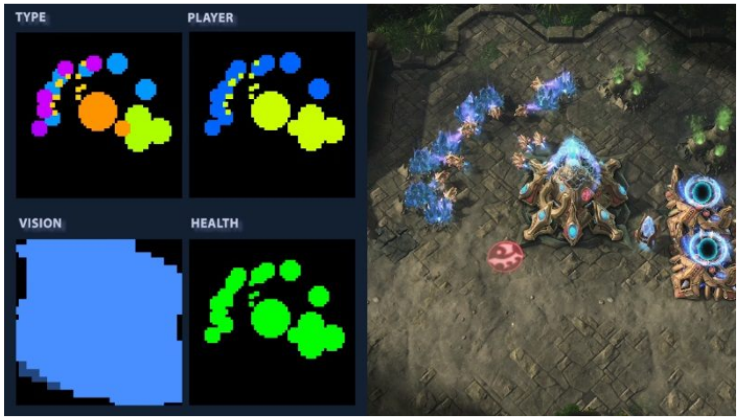


1...e5 g3 d5 cxd5 ♘f6 ♗g2 ♘xd5 ♘f3

[Deepmind AlphaGo / Zero 2017]



[Atari Games - DeepMind 2016]



[Starcraft 2 for AI research]

Chest pathologies [1]

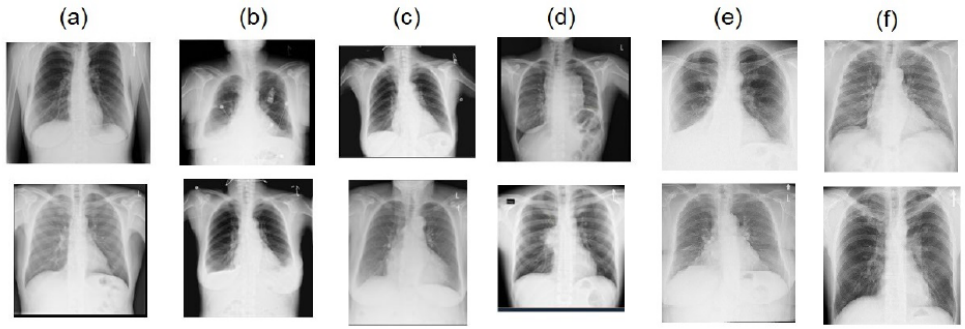
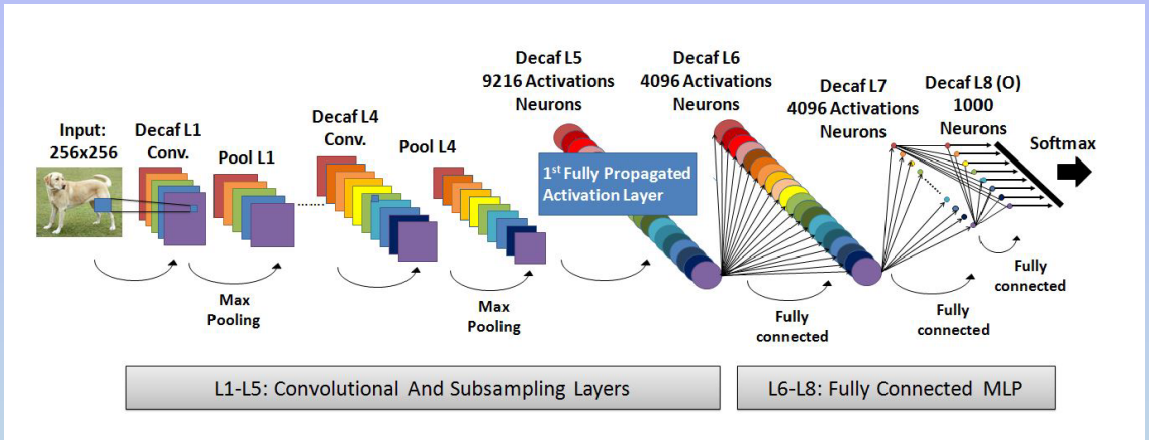


Fig. 1. Chest X-rays categories examples. (a) healthy, (b) left or right effusion, (c) enlarged heart (cardiomegaly), (d) enlarged mediastinum, (e) left or right consolidation, (f) multiple pathologies: enlarged heart, mediastinum, left and right effusion and left or right consolidation (Source: Diagnostic Imaging Department, Sheba Medical Center, Tel Hashomer, Israel)

637 X-ray images (6 chest pathologies):

- Right Pleural Effusion (73 images)
- Left Pleural Effusion (74 images)
- Right Consolidation (58 images)
- Left Consolidation (45 images)
- Cardiomegaly (154 images)
- Abnormal Me-diastinum (145 images)

Convolutional Neural Network



Existing CNN architecture, pretrained on ImageNet (large image data base)

Use SVM to predict a class based on the extracted features

Additional features:

- Bag-of-Visual-Words (BoVW), particularly useful to categorize X-rays on organ level (ImageClefcompetitions, `http://www.imageclef.org`)
- GIST descriptors [4]

Results I

Table 1. AUC accuracy metric classification performance. We use the following abbreviation: Ln for $Decafn$, + for fusion, FS for feature selection method

Descriptor	**Right Pleural Effusion**	**Left Pleural Effusion**	**Right Consolidation**	**Left Consolidation**	**Cardiomegaly**	**Abnormal Mediastinum**	**Healthy vs. Pathology**
GIST	0.85	0.79	0.77	0.41	0.96	0.73	0.88
BoVW	0.89	0.87	0.78	0.65	0.94	0.74	0.85
L5	0.91	0.81	0.80	0.75	0.95	0.79	0.90
L6	0.91	0.82	0.85	0.76	0.94	0.80	0.90
L7	0.90	0.79	0.75	0.79	0.93	0.79	0.89
L5+L6+GIST	0.92	0.82	0.83	0.68	0.96	0.79	0.91
L5+L6+L7	0.92	0.82	0.83	0.78	0.94	0.80	0.91
FS (5000)	0.93	0.82	0.84	0.78	0.95	0.80	0.92

Results II

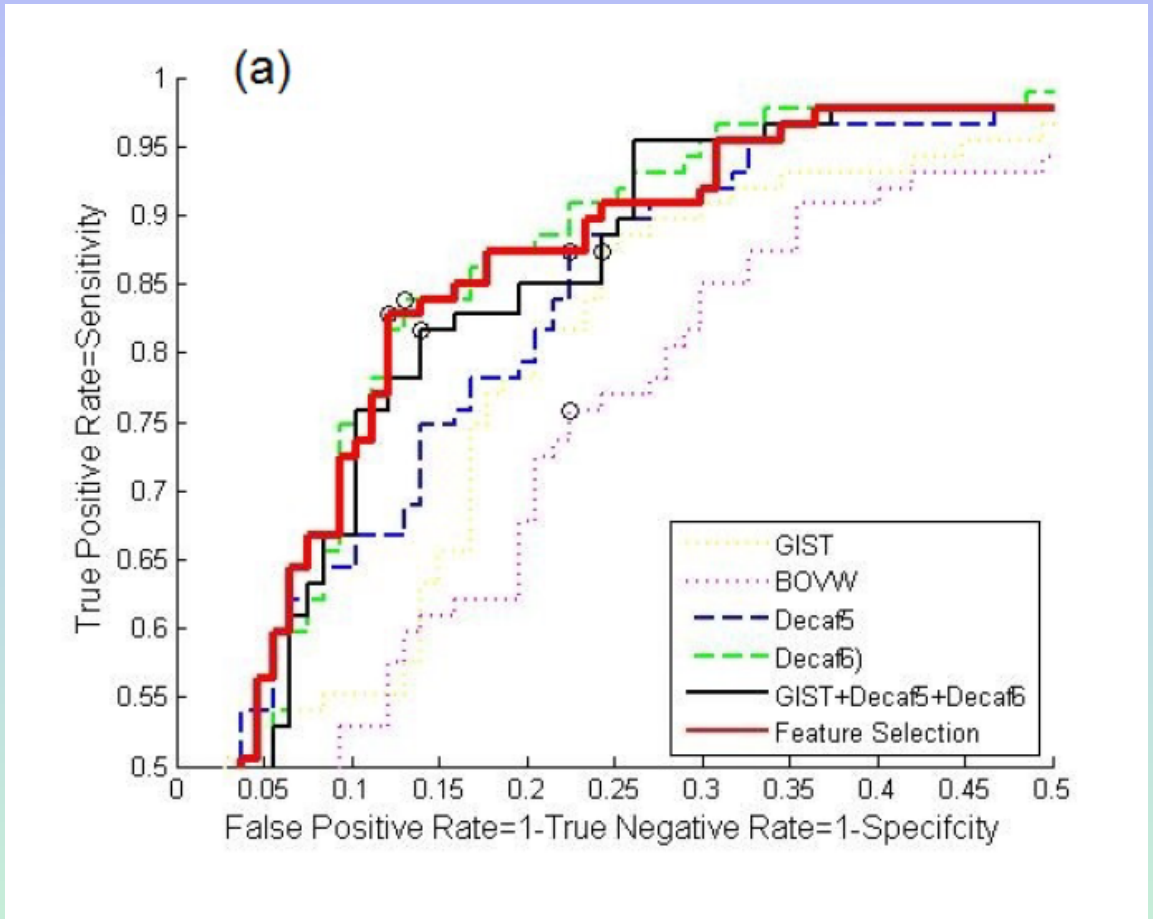
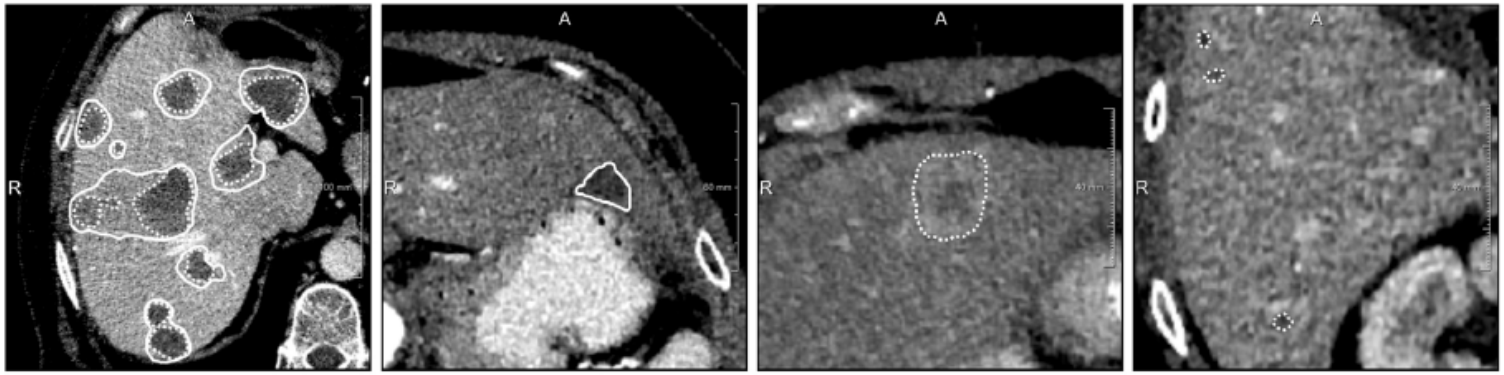


Figure: Healthy vs. Pathology ROC;

Liver lesion segmentation [2]



a b c d

Figure 3. MTRA (*dashed*) vs. LiTS (*solid*) annotations. (**a**) Case with low dice/correspondence (**b**) Case where a LiTS reference tumor was missed (**c**) Case where MTRA found a lesion in a case with no tumors according to LiTS reference (**d**) Case where small additional tumors were found by the MTRA.

Data:

- 131 abdominal CT scans
- The CT scans come with reference annotations of the liver and tumors done by trained radiologists (LiTS reference)
- Another set of annotations given by a medical-technical radiology assistant (MTRA)

Fully Convolutional Neural Networks

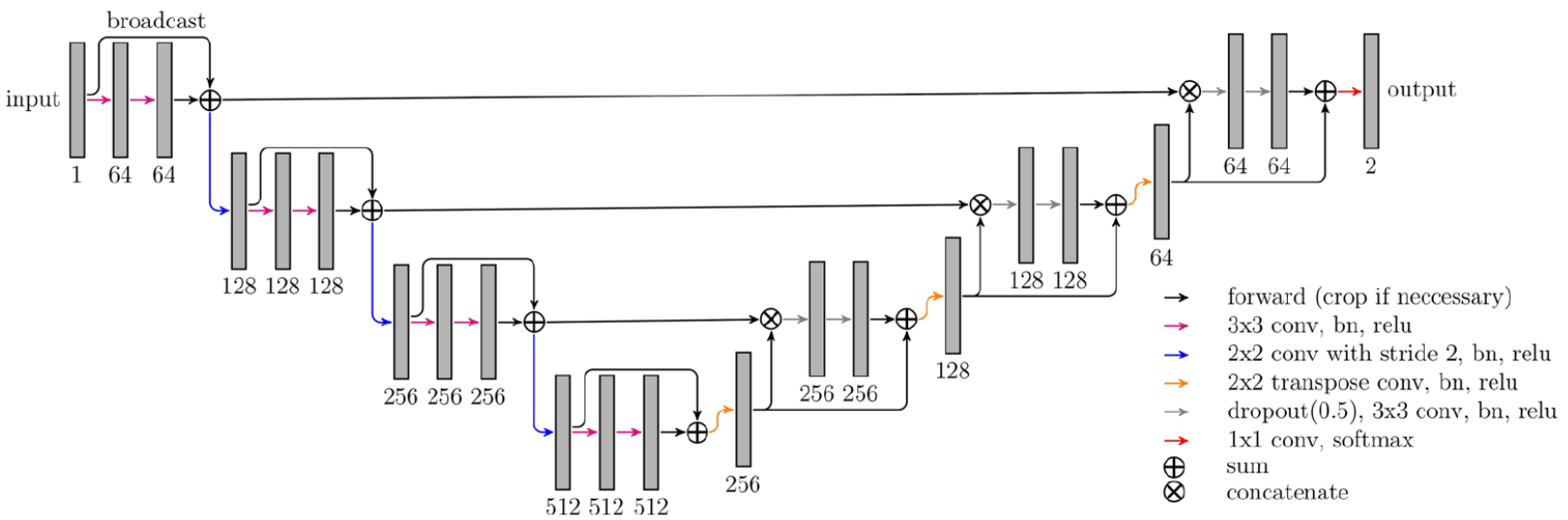


Figure 1. Overview of the neural network architecture. The numbers denote the feature map count.

Based on U-net [6] designed specifically for Biomedical Image Segmentation.

Neural Network Output compared to annotated data

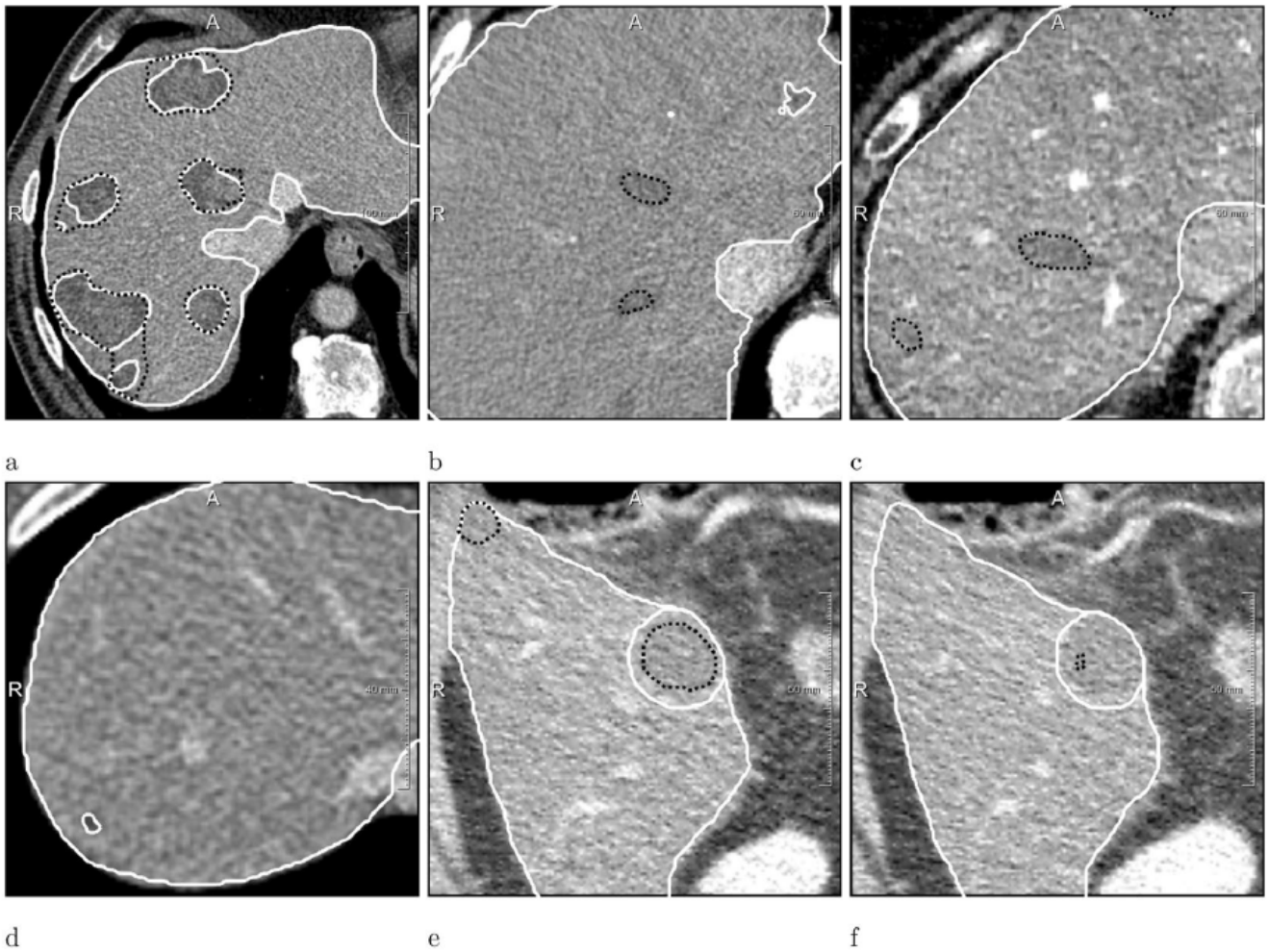


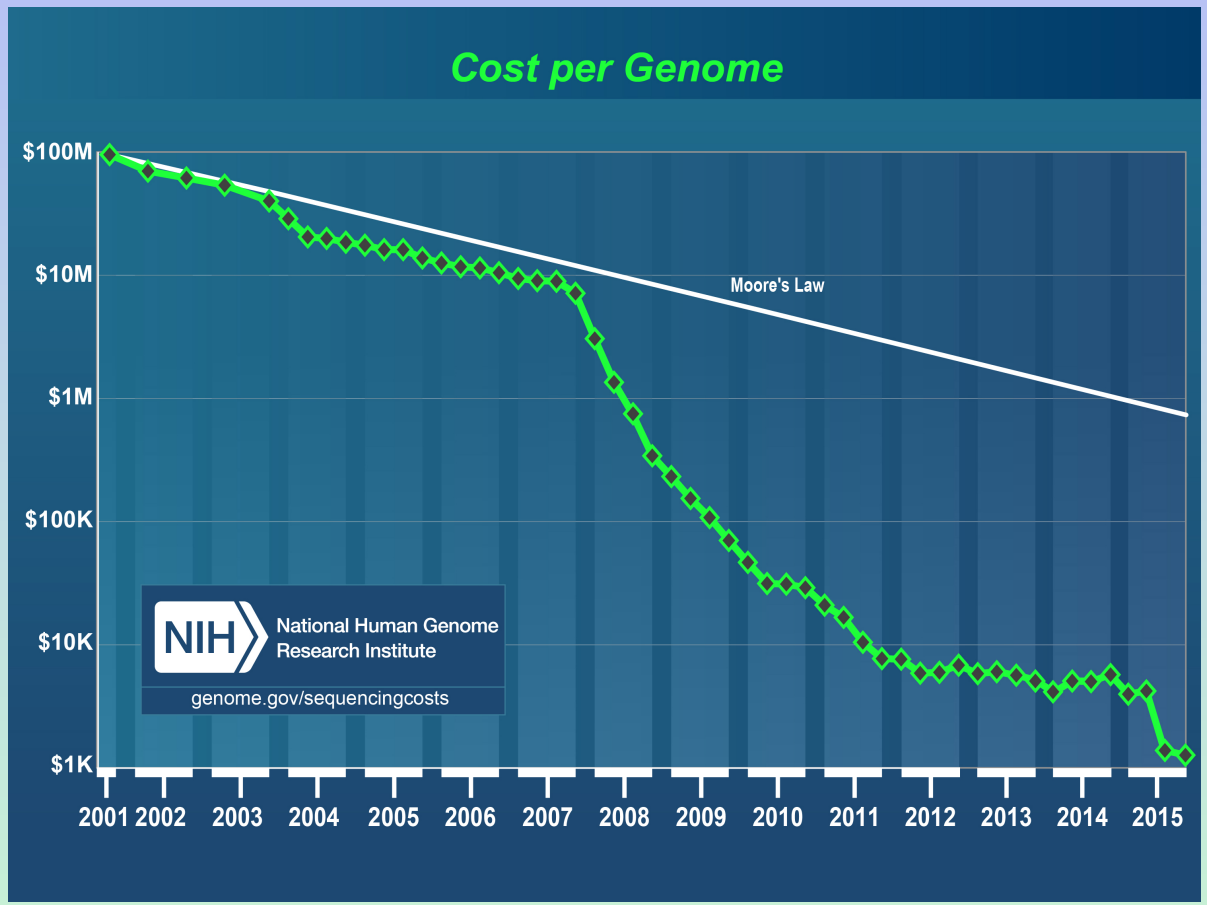
Figure 5. Neural network (*black*) compared with the LiTS (*white*) annotations. (**a**) Case with 0.85 dice/case (**b**,**c**) Cases with 19 and 16 FPs (**d**) Case where a small tumor was not detected (**e**,**f**) Case where tumor segmentation strongly differed on consecutive slices.

Comparison between Human and Computer recognition performances

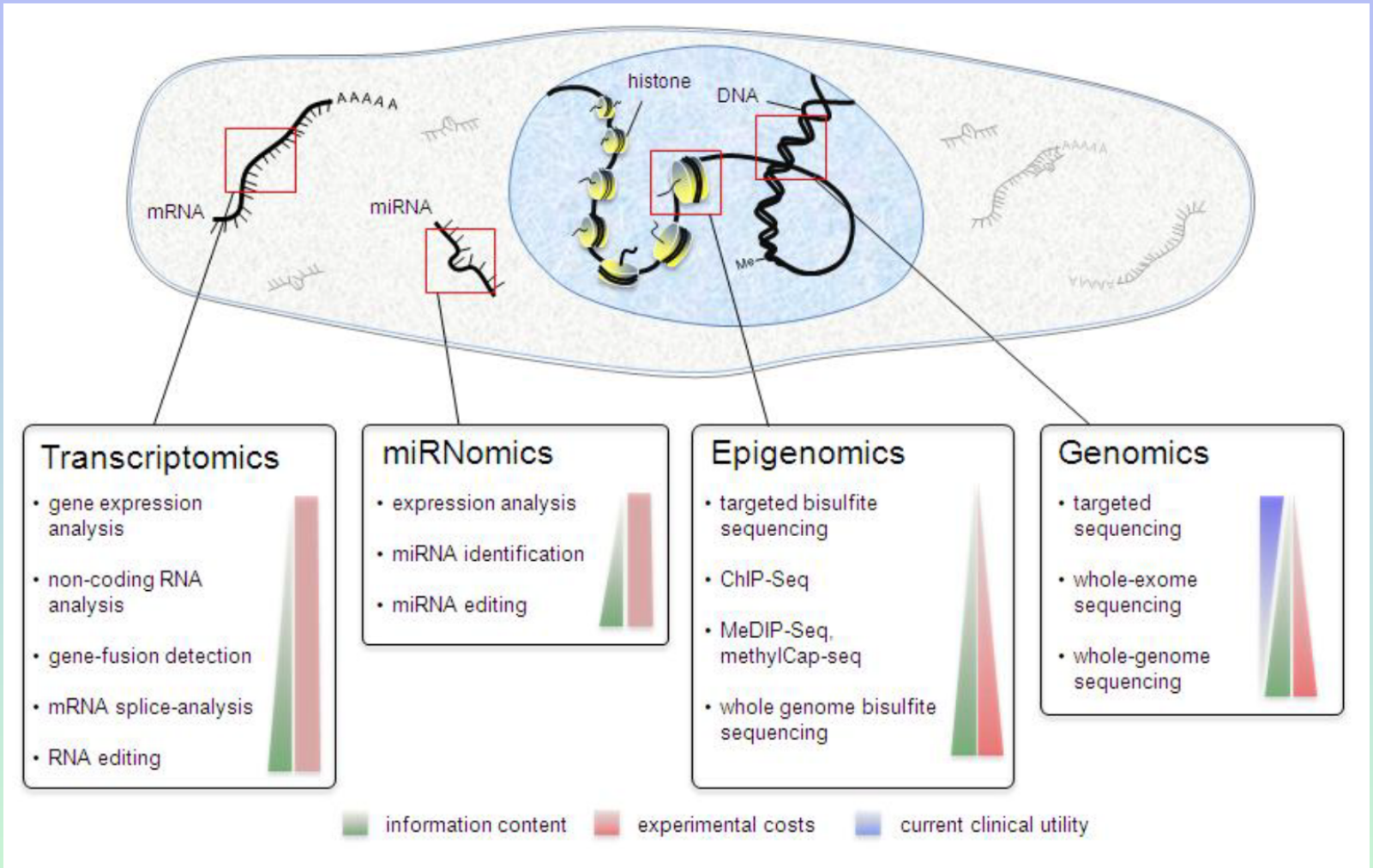
	Recall	Recall $\geq$ 10 mm	FP per case	Dice per case	Dice per correspondence	Merge error	Split error
Human vs. Human							
MTRA (LiTS)	0.92	0.94	2.6	0.70 ± 0.27	0.72 ± 0.11	11	5
LiTS (MTRA)	0.62	0.85	0.3	0.70 ± 0.27	0.72 ± 0.11	5	12
Computer vs. Human							
FCN (MTRA)	0.47	0.75	4.7	0.53 ± 0.37	0.72 ± 0.11	7	13
FCN (LiTS)	0.72	0.86	4.6	0.51 ± 0.37	0.65 ± 0.16	12	14
FCN + RF (LiTS)	0.63	0.77	0.7	0.58 ± 0.36	0.69 ± 0.18	11	10

Table 1. Mean metric values for human vs. human and computer vs. human comparisons. The parentheses denote the dataset used as a reference for the computation of evaluation metrics.

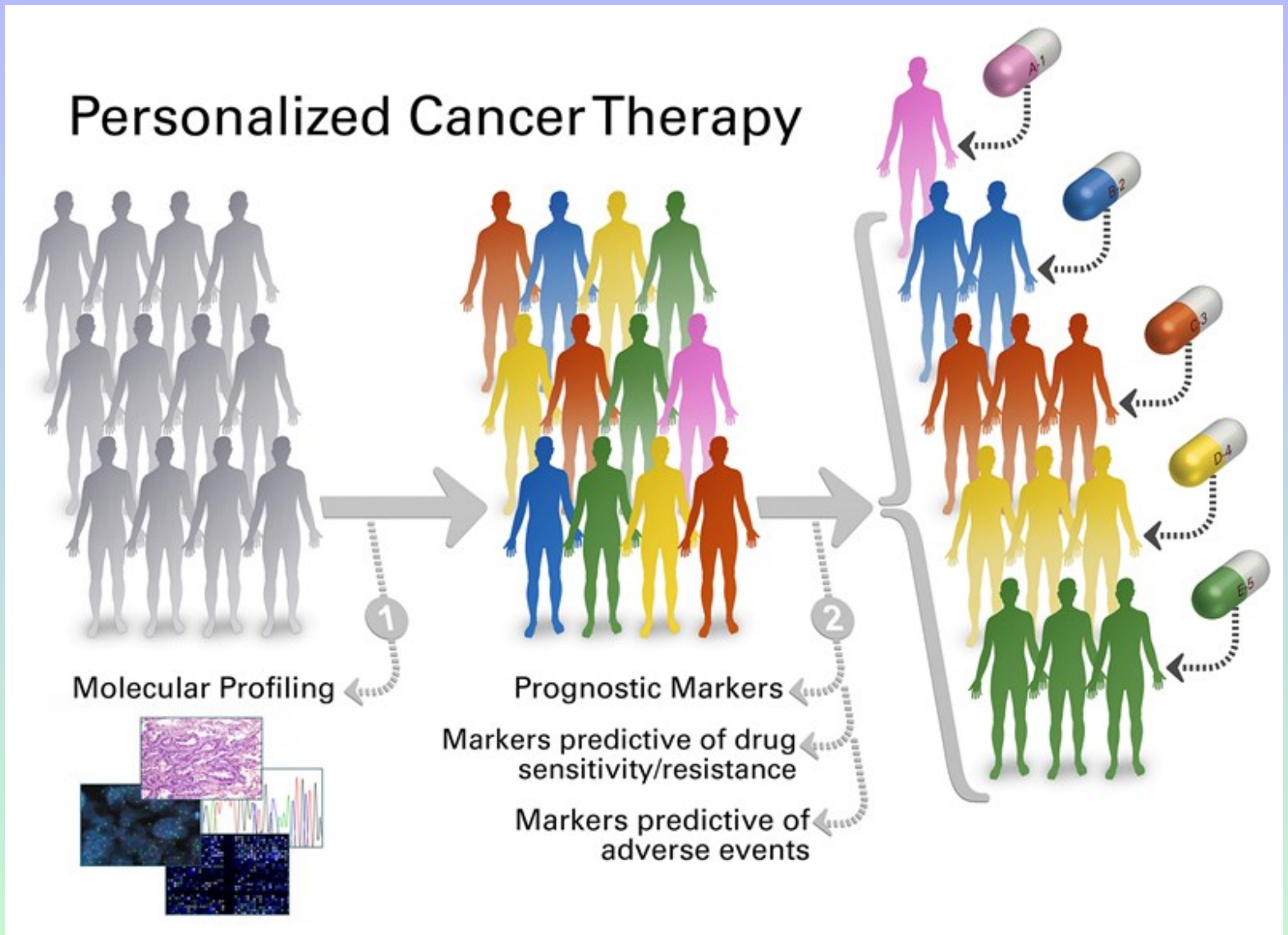
Cost per genome



Different fields in Omics [3]



Precision Medicine

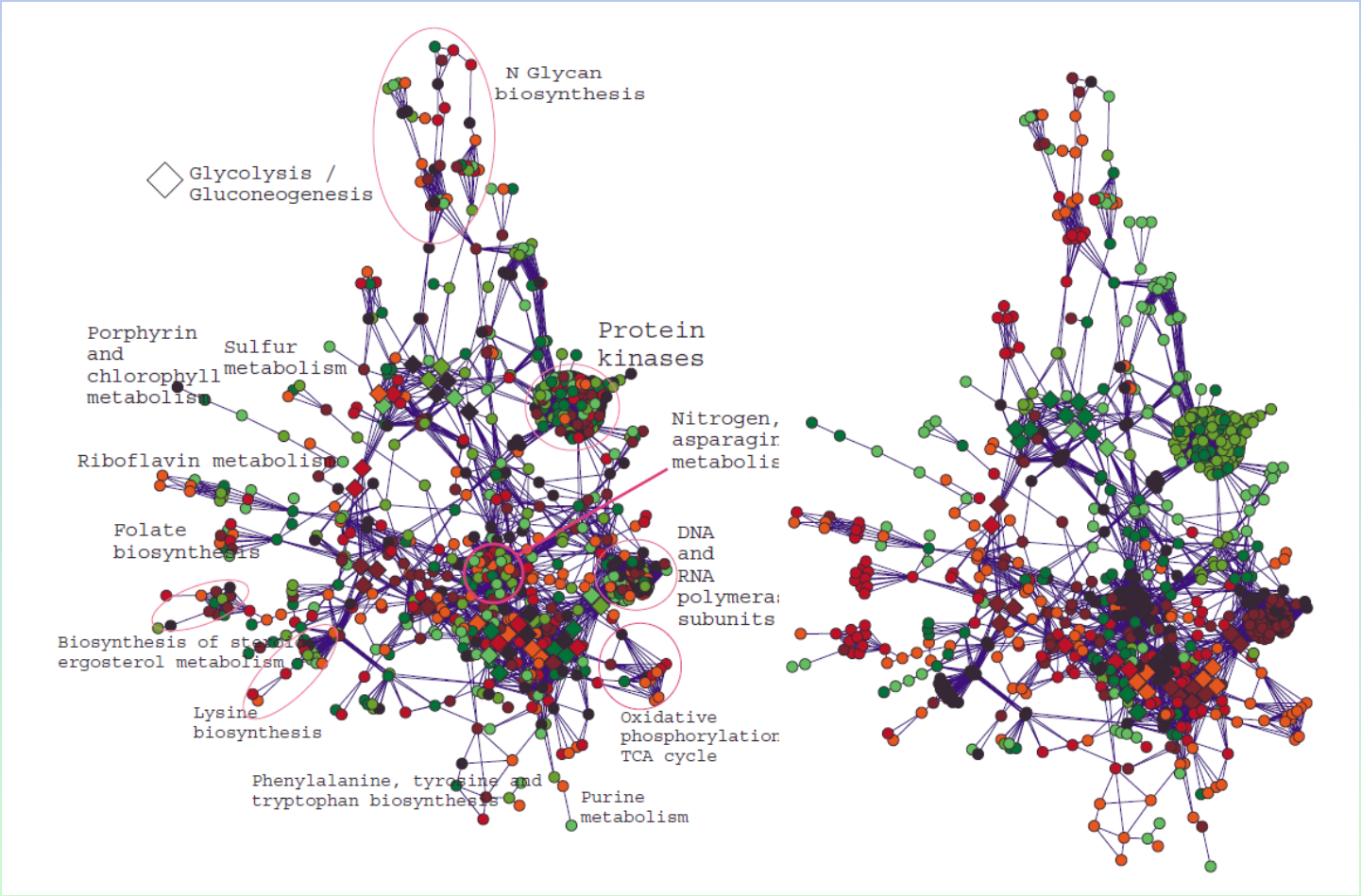


https://pct.mdanderson.org

Need for robustness - incorporating knowledge in the model

[5] Effect of low irradiation on Saccharomyces cerevisiae strains: 6 irradiated groups vs 11 non-irradiated groups.

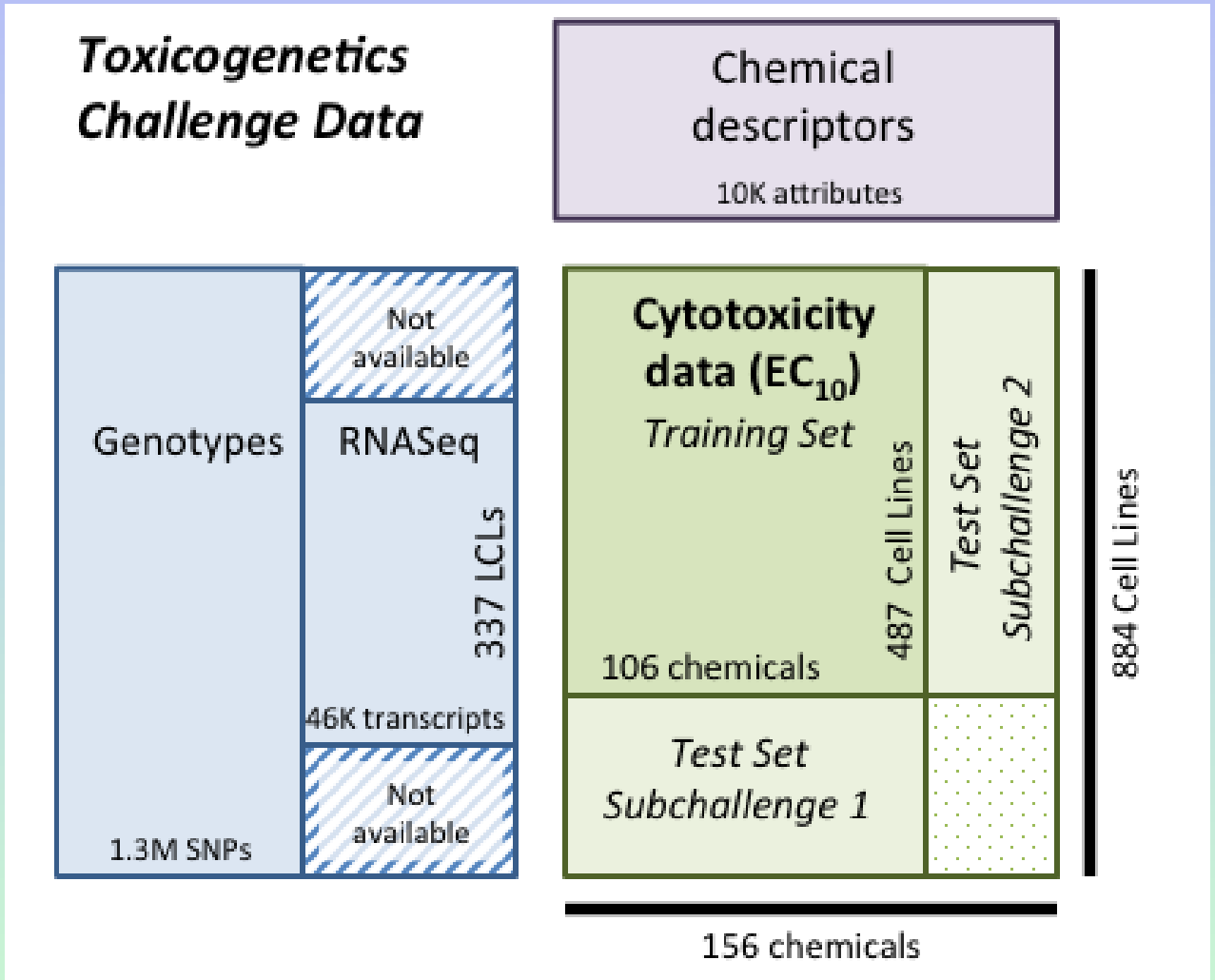
Aim: Detect the irradiation looking only at the transcriptional changes.



What is toxicogenetics ?

- Different responses to drugs or environmental chemicals according to genotypes
- Aim : provide a personalized treatment for patients

DREAM8 Toxicogenetics challenge



Traumabase

15000 patients/ 250 variables/ 11 hospitals, from 2011 (4000 new patients/ year)

Center	Accident	Age	Sex	Weight	Height	BMI	BP	SBP	
1	Beaujon	Fall	54	m	85	NR	NR	180	110
2	Lille	Other	33	m	80	1.8	24.69	130	62
3	Pitie Salpetriere	Gun	26	m	NR	NR	NR	131	62
4	Beaujon	AVP moto	63	m	80	1.8	24.69	145	89
6	Pitie Salpetriere	AVP bicycle	33	m	75	NR	NR	104	86
7	Pitie Salpetriere	AVP pedestrian	30	w	NR	NR	NR	107	66
9	HEGP	White weapon	16	m	98	1.92	26.58	118	54
10	Toulon	White weapon	20	m	NR	NR	NR	124	73

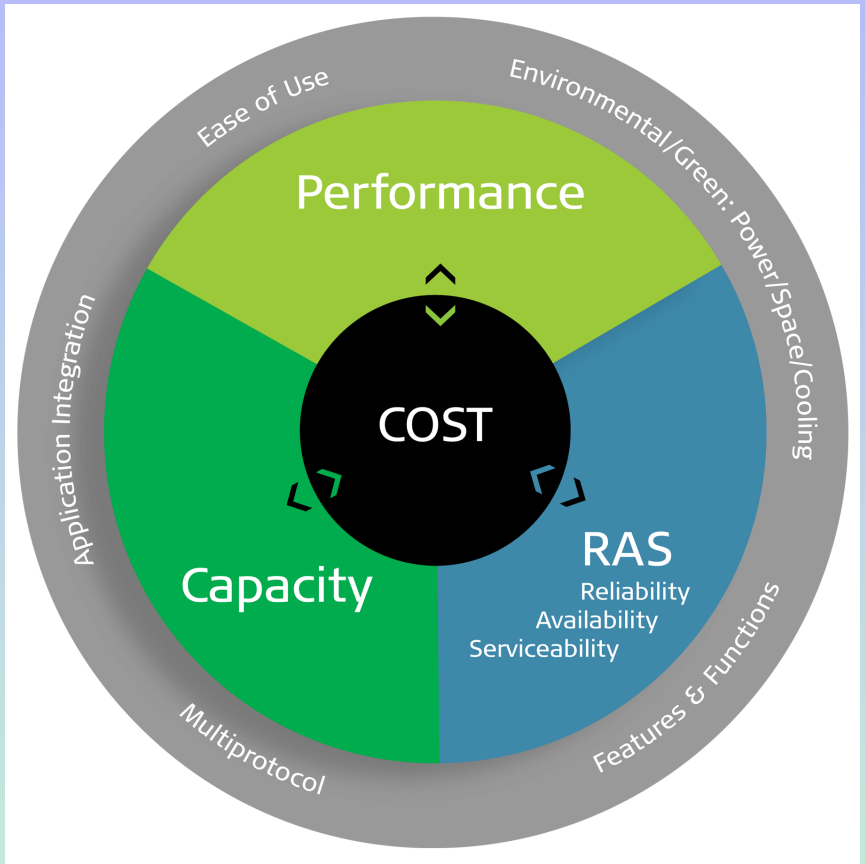
.....

	SpO2	Temperature	Lactates	Hb	Glasgow	Transfusion	.....
1	97	35.6	<NA>	12.7	12	yes	
2	100	36.5	4.8	11.1	15	no	
3	100	36	3.9	11.4	3	no	
4	100	36.7	1.66	13	15	yes	
6	100	36	NM	14.4	15	no	
7	100	36.6	NM	14.3	15	yes	
9	100	37.5	13	15.9	15	yes	
10	100	36.9	NM	13.7	15	no	

$\Rightarrow$ **Predict** whether to start a blood transfusion, to administer fresh frozen plasma, etc...

1. History of AI
2. From Big Data to Deep Learning
3. AI and Health
   - Chest X-ray
   - Liver lesion segmentation
   - Genomics
   - Toxicogenetics
   - Medical
4. Perspective and issues

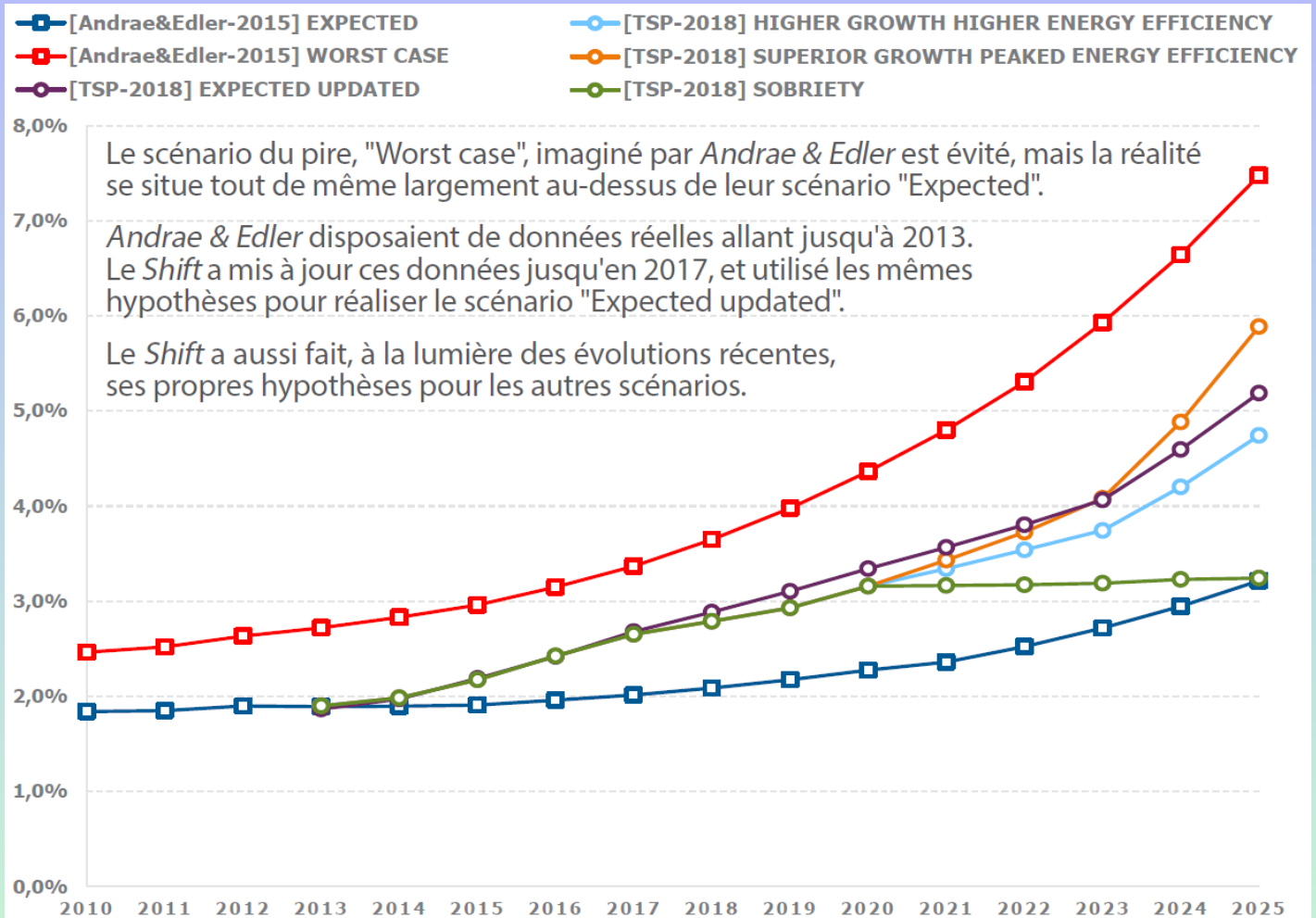
Environmental issues



Data centers:

- 2% of the total electricity consumption in the US.
- 626 billion liters of water.
- 2% of total global greenhouse emissions.

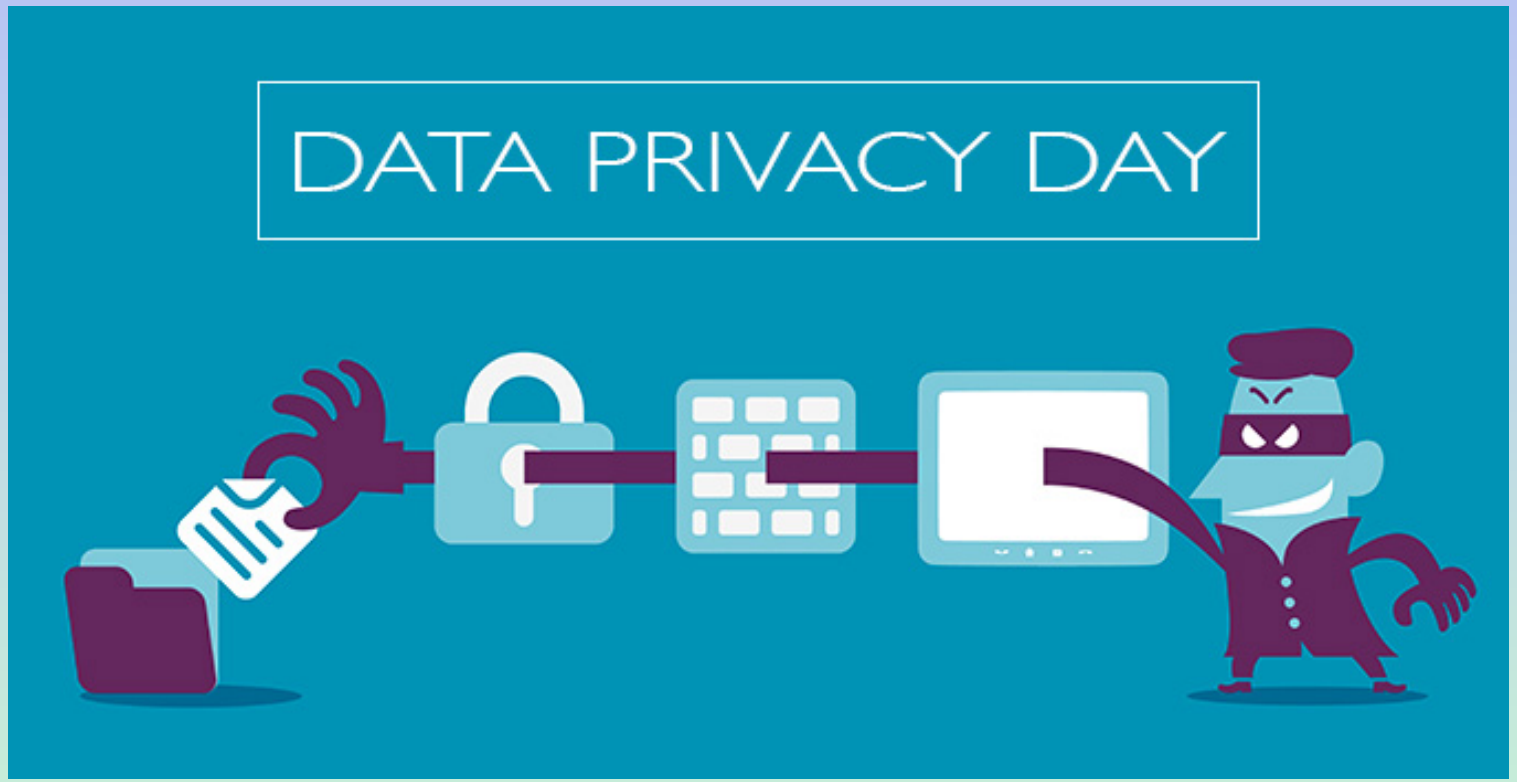
Environmental issues



Évolution de la consommation énergétique mondiale du numérique entre 2010 et 2025, rapportée à la consommation énergétique mondiale totale

[Source : scénarios et calculs *The Shift Project* 2018, à partir de Andrae & Edler 2015]

Data privacy



Ethical issues - Bias and use of AI

- Tay (”thinking about you”) was an AI released by Microsoft via Twitter on March 2016. It was shut down when the bot began to post inflammatory and offensive tweets, only 16 hours after its launch.
- Correctional Offender Management Profiling for Alternative Sanctions (Compas) used in the US.
- Social Credit System
- John James Vlahos died of cancer in february 2017. His son, James, implemented an IA on Facebook Messenger named “dadbot”.

What kind of society do we want?

Far from terminator

- Stephen Hawking BBC, Dec 2 2014

  The development of full artificial intelligence could spell the end of the human race. We cannot quite know what will happen if a machine exceeds our own intelligence, so we can't know if we'll be infinitely helped by it, or ignored by it and sidelined, or conceivably destroyed by it.

Take-home messages

- AI originates from 1950s but became more prevalent in the last seven years due to computational power, and data explosion.
- AI is not about reproducing human brain or human cognition.
- AI is about solving task that are thought to require human intelligence.
- Wherever you can collect data, you can use artificial intelligence, keeping in mind that specific questions require specific data.
- Many challenges in AI developing the algorithmic tools (computer science, mathematics), and avoiding major pitfalls both technical (bias in data) and ethical.

Take-home messages II

- AI is particularly useful for unstructured data like images, audio signal, texts and videos.
- Achieving good performances with AI often requires a lot of data, which can be cumbersome in particular fields such as medicine (rare pathology).
- You do not need to make assumptions on your data in order to implement AI. But you also don't have a precise understanding of what the model does.

Thank you!

References I

[1] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan. Chest pathology identification using deep feature selection with non-medical training. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 6(3):259–263, 2018.

[2] G. Chlebus, A. Schenk, J. H. Moltz, B. van Ginneken, H. K. Hahn, and H. Meine. Automatic liver tumor segmentation in ct with fully convolutional neural networks and object-based postprocessing. *Scientific reports*, 8(1):15497, 2018.

[3] K. Frese, H. Katus, and B. Meder. Next-generation sequencing: from understanding biology to personalized medicine. *Biology*, 2(1):378–398, 2013.

[4] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3):145–175, 2001.

[5] F. Rapaport, A. Zinovyev, M. Dutreix, E. Barillot, and J.-P. Vert. Classification of microarray data using gene networks. *BMC bioinformatics*, 8(1):35, 2007.

[6] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.