

# The AI revolution in scientific research

---

Source: The Royal Society and The Alan Turing Institute Report

# From the t-test to the frontiers of AI

- Scientists aspire to understand the workings of nature, people, and society by:
  - Formulating hypotheses,
  - Designing experiments, and
  - Collecting data

The aim is analyzing and better understanding natural, physical, and social phenomena

# From the t-test to the frontiers of AI

- Data collection and analysis is a core element of the scientific method, and scientists have long used **statistical techniques** to aid their work
  - **t-test (Student test)** gave researchers a new tool to extract insights from data in order to test the veracity of their hypotheses.
  - Such mathematical frameworks were vital in extracting as much information as possible from data that had often taken significant time and money to generate and collect

# From the t-test to the frontiers of AI

- The application of statistical methods to scientific challenges can be seen throughout history:
  - The analysis by Johannes Kepler of the astronomic measurements of Tycho Brahe in the early seventeenth century led to his formulation of the laws of planetary motion, which subsequently enabled Isaac Newton (and others) to formulate the law of universal gravitation.

# From the t-test to the frontiers of AI

- The application of statistical methods to scientific challenges can be seen throughout history:
  - The laboratory at Rothamsted was established as a centre for agricultural research, running continuously monitored experiments from 1856 which are still running to this day. Ronald Fisher – a prominent statistician – was hired to work there in 1919 to direct analysis of these experiments. His work went on to develop the theory of experimental design and lay the groundwork for many fundamental statistical methods that are still in use today.

# From the t-test to the frontiers of AI

- The development of artificial intelligence (AI) techniques offered additional tools for extracting insights from data.
  - Alan Turing grappled with the idea of machine intelligence.
  - In 1950, he posed the question “can machines think?”, and suggested a test for machine intelligence – subsequently known as the Turing Test – in which a machine might be called intelligent, if its responses to questions could convince a person that it was human



# AI as an enabler of scientific discovery

- Using genomic data to predict protein structures:
  - By predicting the protein's shape, scientists can identify proteins that play a role in diseases, improving diagnosis and helping develop new treatments.
  - Determining the shape of a protein from its corresponding genetic sequence – the protein-folding challenge – is a complex task.
- The AlphaFold project at DeepMind has created a deep neural network that predicts the distances between pairs of amino acids and the angles between their bonds, and in so doing produces a highly-accurate prediction of an overall protein structure.



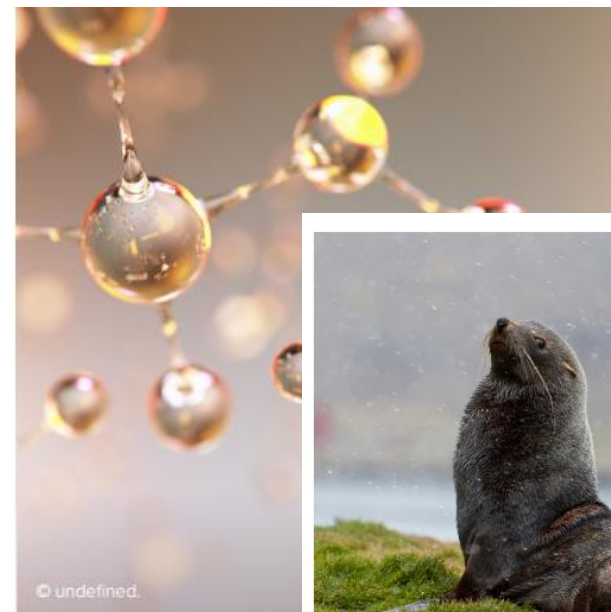


# AI as an enabler of scientific discovery



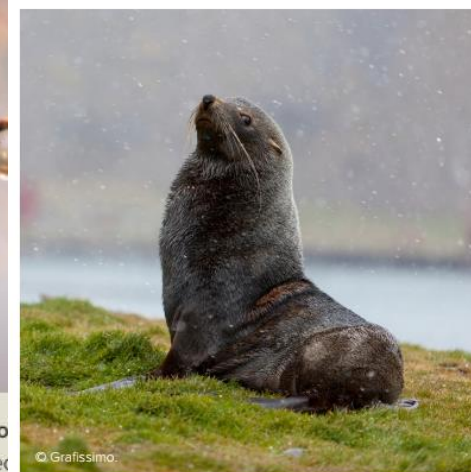
**Materials characterisation using high-resolution imaging**  
Materials behave differently depending on their internal

l structure is often extracted by  
gh them and studying the resulting  
contemporary approaches for  
ering patterns are iterative and  
ntion of scientists. The scope of this  
he options of using machine learning  
rring the structural information of  
g the scattering patterns<sup>15</sup>.



## Understanding complex o

The goal of this pilot project  
Centre and The Alan Turing  
possibilities for machine le  
predicting the process of t  
plants. Triterpenes are com  
a large and important class  
with diverse commercial ap  
agriculture and industrial s  
all synthesized from a sing  
can then be further modifi  
give over 20,000 structurally diverse triterpenes. Recent  
machine learning models have shown promise at  
predicting the outcomes of organic chemical reactions.  
Successful prediction based on sequence will require



## Satellite imaging to support conservation

Many species of seal in the Antarctic are extremely  
difficult to monitor as they live exclusively in the sea-ice  
zone, a region that is particularly difficult to survey. The  
use of very high-resolution satellites enables researchers  
to identify these seals in imagery at greatly reduced cost  
and effort. However, manually counting the seals over the  
vast expanse of ice that they inhabit is time consuming,  
and individual analysts produce a large variation in count  
numbers. An automated solution, through machine  
learning methods, could solve this problem. divina quick.



## Understanding social history from archive material

Researchers are collaborating with curators to build  
new software to analyse data drawn initially from millions  
of pages of out-of-copyright newspaper collections  
from within the British Library's National Newspaper  
archive. They will also draw on other digitised historical  
collections, most notably government-collected data,  
such as the Census and registration of births, marriages  
and deaths. The resulting new research methods will  
allow computational linguists and historians to track  
societal and cultural change in new ways during the



• Finding patterns in astronomical data: Research in  
astronomy generates large amounts of data and a key  
challenge is to detect interesting features or signals from  
the noise, and to assign these to the correct category  
or phenomenon. For example, the Kepler mission is

in stellar activity, or other systematic trends. Before the  
data can be analysed, these so-called instrumental  
artefacts need to be removed from the system. To help  
with this, researchers have developed a machine learning  
system that can identify these artefacts and remove them



**Driving scientific discovery from particle physics  
experiments and large scale astronomical data**  
Researchers are developing new software tools



# Are things going to get better or worse?



Next life: Metaverse

## Neurotechnology, Elon Musk and the goal of human enhancement

Brain-computer interfaces could change the way people think, soldiers fight and Alzheimer's is treated. But are we in control of the ethical ramifications?



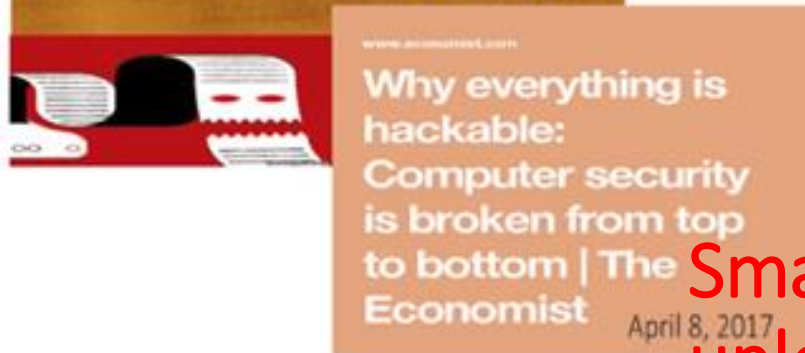
Neuralink



## Immersive experience



# Are things going to get better or worse?



## Hackers Can Access Elon Musk's Neuralink to Steal Memories, Skills and Thoughts, Warn Experts

■ Elon Musk, who has been working on his startup Neuralink since 2016, is set to release news of a possible working prototype of the brain chip on August 28

By Bhaswati Guha Majumder

August 6, 2020 19:45 +08



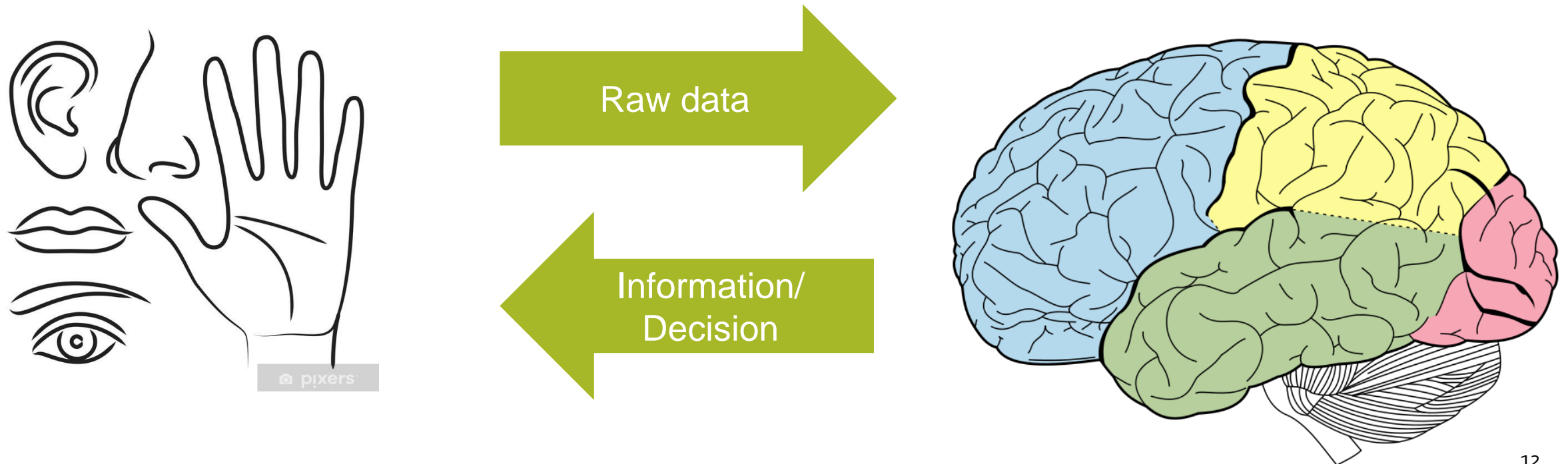
Smart Environment could become horror environment unless we design security (and privacy) correctly

# Research questions to advance the application of AI in science

- Let start with the Human Brain: Information processing
- AI Hypothesis: the human brain as a biological computer System
  - Store the information
  - Process and analyze the information using complex algorithms

# Research questions to advance the application of AI in science

- Let start with the Human Brain: Information processing





# Research questions to advance the application of AI in science

- DATA IS IMPORTANT

## البيانات.. نفط الغد

ما هي البيانات التي يجري التنافس عليها ويستخدم الصراع حولها؟



# Research questions to advance the application of AI in science

- DATA MANAGEMENT

- Is there a principled method to decide what data to keep and what to discard, when an experiment or observation produces too much data to store?
- How will this affect the ability to re-use the data to test alternative theories to the one that informed the filtering decision?



# Research questions to advance the application of AI in science

- DATA MANAGEMENT

- What does 'open data' mean in practice where the data sets are just too large, complex and heterogenous for anyone to actually access and understand them in their entirety



# Research questions to advance the application of AI in science

- DATA MANAGEMENT

- How can scientists search efficiently for rare or unusual events and objects in large and noisy data sets?

# Research questions to advance the application of AI in science

- DATA MANAGEMENT

The Data Preparation Pipeline



# Research questions to advance the application of AI in science

- AI METHODS AND CAPABILITIES
  - How can machine learning help integrate data from different sources collected under different conditions and for different purposes, in a way that is scientifically valid (for example, integrate observations of the same system taken at different scales)

# Research questions to advance the application of AI in science

- AI METHODS AND CAPABILITIES
  - How can researchers re-use data which they have already used to inform theory development, while maintaining the rigor of their work?

# Research questions to advance the application of AI in science

- AI METHODS AND CAPABILITIES
  - How can AI methods produce results which are transparent as to how they were obtained, and interpretable within the disciplinary context?



# Research questions to advance the application of AI in science

- AI METHODS AND CAPABILITIES
  - How can research help create more advanced, and more accurate, methods of verifying machine learning systems to increase confidence in their deployment?

# Research questions to advance the application of AI in science

- INTEGRATING SCIENTIFIC KNOWLEDGE

- Is there a rigorous way to incorporate existing theory/ knowledge into a machine learning algorithm, to constrain the outcomes to scientifically plausible solutions?
- How can AI be used to actually discover and create new scientific knowledge and understanding, and not just the classification and detection of statistical patterns?

# What is Artificial Intelligence?



# What is Artificial Intelligence?

- The expression 'artificial intelligence' today is therefore an umbrella term.
- It refers to:
  - a suite of technologies that can perform complex tasks when acting in conditions of uncertainty, including visual perception, speech recognition, natural language processing, reasoning, learning from data, and a range of optimization problems

# What is Artificial Intelligence?

Artificial



Not natural

Intelligence



Ability to understand, think, learn and act

# What is Artificial Intelligence?

- Artificial Intelligence (AI)
  - AI is intelligence demonstrated by machines, as opposed to natural intelligence displayed by ~~animals including humans~~.
  - Leading AI textbooks define the field as the study of "intelligent agents": any system that perceives its environment and takes actions that maximize its chance of achieving its goals.
  - Some popular accounts use the term "AI" to describe machines that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving", however, this definition is rejected by major AI researchers.



# What is Artificial Intelligence?

- Artificial General Intelligence (AGI)
  - AGI is the hypothetical ability of an intelligent agent to understand or learn any intellectual task that a human being can.
  - It is a primary goal of some artificial intelligence research and a common topic in science fiction and futures studies.
  - AGI can also be referred to as strong AI, full AI, or general intelligent action. (Although academic sources reserve the term "strong AI" for computer programs that experience sentience or consciousness.)

# What is Artificial Intelligence?

- Individual Artificial Intelligence: neurocomputer interface
  - Completely new type of artificial intelligence: technological symbiosis against traditional electronic systems
    - Biotechnological neurocomputer symbiosis
- Human brain and a machine will work together in a dual complementary system
  - Both components will complement and reinforce each other