

Artificial Intelligence Applied to Scientific Research and Ethics: How can we leverage AI algorithms to achieve better science?

Mestrado em Energias Sustentáveis | Workshop

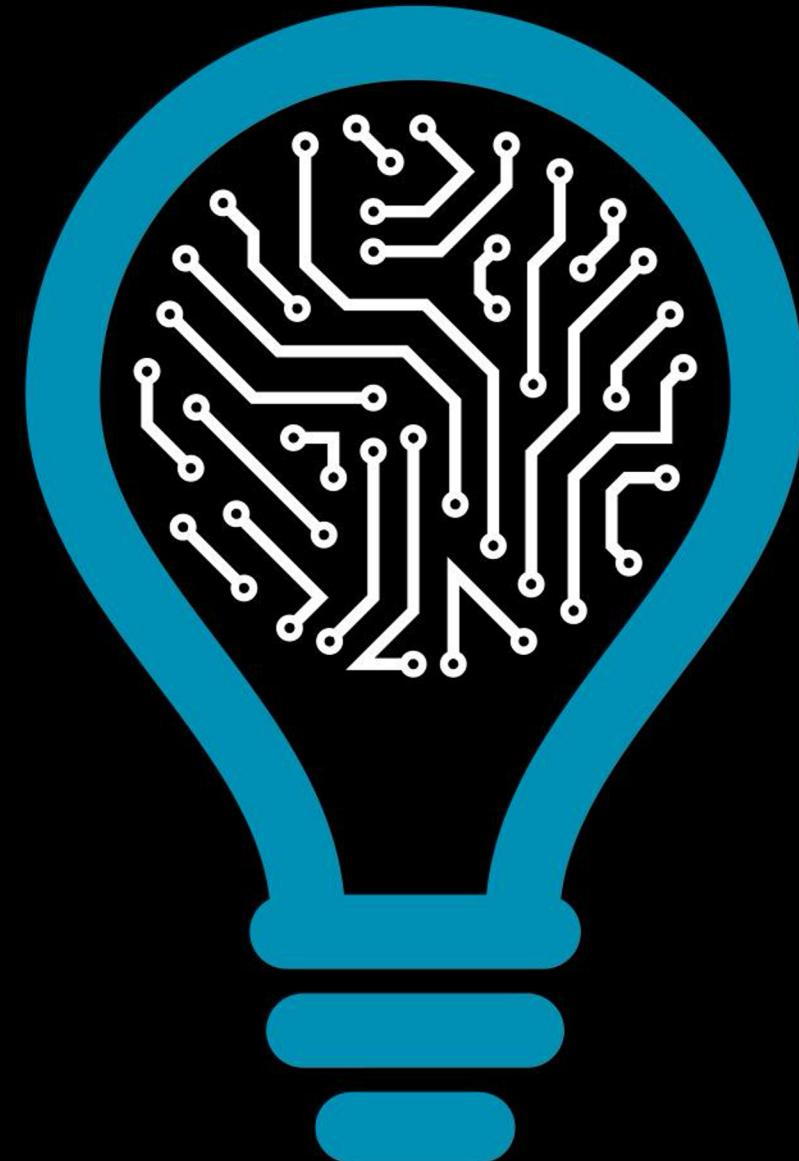
December 10, 2024 | Instituto Superior de Engenharia do Porto

Tiago Filipe Sousa Gonçalves

<https://tiagofilipesousagoncalves.github.io/>



INSTITUTE FOR SYSTEMS
AND COMPUTER ENGINEERING,
TECHNOLOGY AND SCIENCE



Outline

- 1. Modern Artificial Intelligence is Statistics on Steroids**
- 2. Regulating Science: Research Integrity and Ethics**
- 3. Artificial Intelligence meets Science**
- 4. Surfing the Waves of AI: Take-home Messages**



1. Modern Artificial Intelligence is Statistics on Steroids



In the beginning, we had rules

- In the beginning, **artificial intelligence systems were based in algorithms**:
 - An algorithm is a **set of instructions** that the system will follow to **achieve a certain goal** (direct programming)^[1]
 - These **explicit** rules were often based on **domain knowledge**
 - Hence, they were “easy” to **explain** and to **understand**
- Nowadays, we use the available data to automatically learn **programs/functions**:
 - In machine learning, we **learn from data and make predictions** (indirect programming)^[1]
 - These algorithms work by **optimising an objective function**
 - Hence, the “rules” often are **implicit** and **difficult to understand**

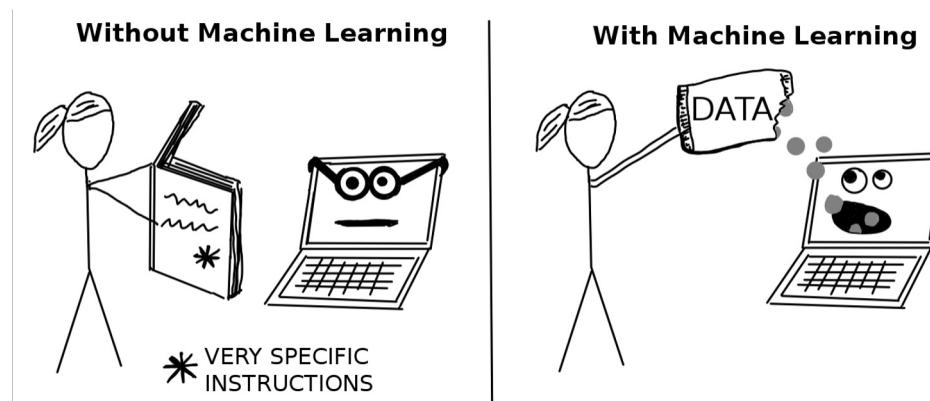


Figure - Algorithms vs Machine Learning (Image from [1])

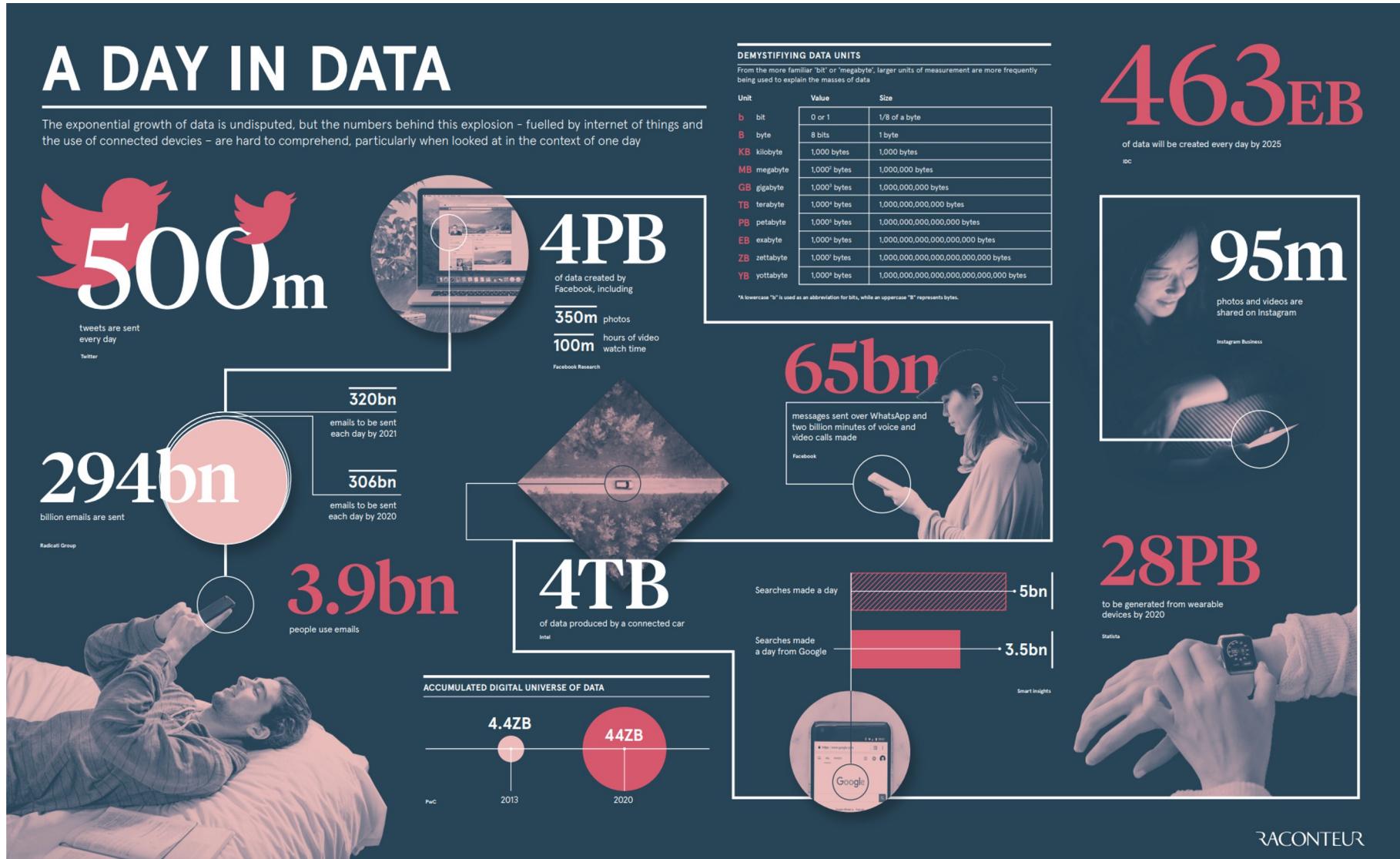


Today, we have data^[1]

- **The paradigm is changing:** most of the daily tasks and services can now be performed with the aid of **digital applications** or **gadgets**
- High-tech companies such as Google, Facebook, Netflix or Amazon **have access to huge amounts of data from several data sources and users:**
 - This phenomenon suggests that the *business of data* will become a **significant sector of the global economy**^[2]
 - There are several **open-source data sets with millions of entries** (e.g., ImageNet^[3])
- Data is referred as **the new oil**^[4]
 - The main impact on humanity is related to **the way data can improve our lives**
 - **A proper management process of the “dark side” of data must be implemented**, but the **advances in data fuels are worth the effort**



Yes, lots of data^[1, 2, 3]





We have more computational power than ever

- The fundamental concepts of artificial intelligence and deep neural networks have been around since 1940^[1]
 - Frank Rosenblatt proposed one of the first approaches to the design and training of artificial neural networks: **the Perceptron**^[2]
- The development of **powerful computer processing units (CPUs)** and the leveraging of the **graphical processing units (GPUs)**^[3] for computation allowed the training of deep and complex algorithms in “human time”

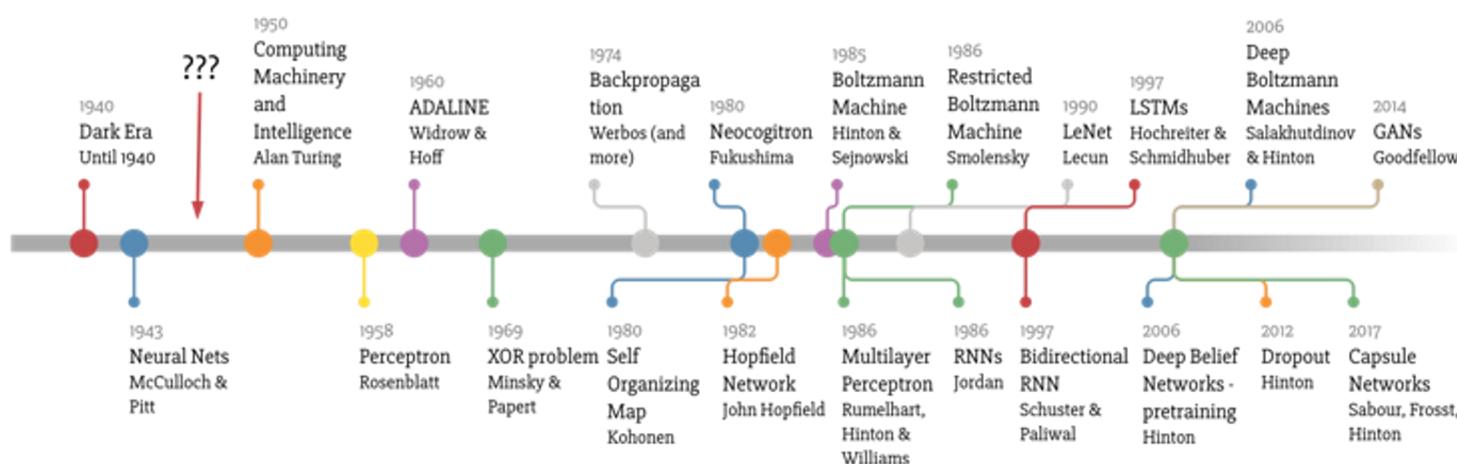


Figure - A (tentative) deep learning timeline (Image from [1])



Deep learning *versus* traditional machine learning^[1]

- Traditional machine learning required **experts to extract meaningful features** (i.e., domain-specific features) from raw data and feed them into machine learning algorithms to obtain classification/regression models:
- Deep learning “only” requires **raw data and labels** to achieve high-performing models, since it **automatically extracts the patterns**
 - Deep learning algorithms are suitable for **representation learning**, i.e., finding the **best representation of the data** that optimises a given optimisation objective

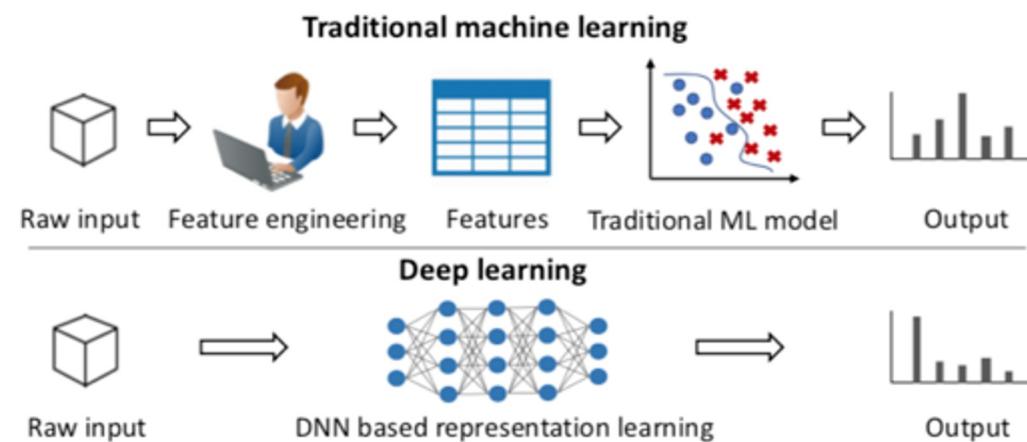


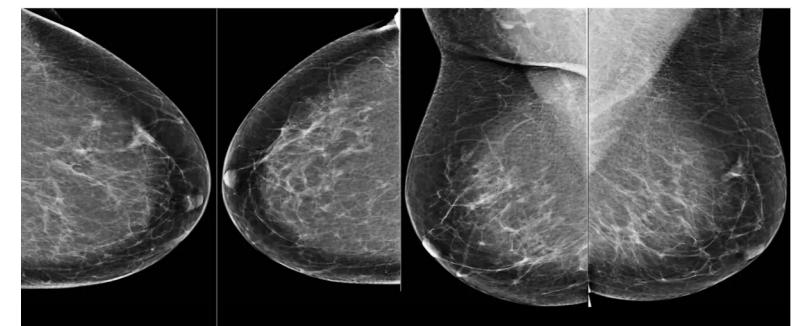
Figure - Deep learning vs traditional machine learning (Image from [2])



Technology has been *challenging* human performance...

- There are, at least, two popular events that created a revolution in the History of AI:
 - In 1997, IBM's Deep Blue beat the Chess World Champion Garry Kasparov^[1]
 - In 2016, Google's DeepMind AlphaGo learn to play Go alone (i.e., through reinforcement learning policies) and beat the Go World Champion Lee Sedol^[2]
- The two events above are examples of the **(virtually) unlimited boundaries of the application of artificial intelligence** to our daily lives
 - In 2020, Google's DeepMind published a paper in *Nature* suggesting that “its model was able to spot cancer in de-identified screening mammograms with fewer false positives and false negatives than experts”^[3, 4]

Figure - Medical Image Analysis: Mammograms (Image from [4])

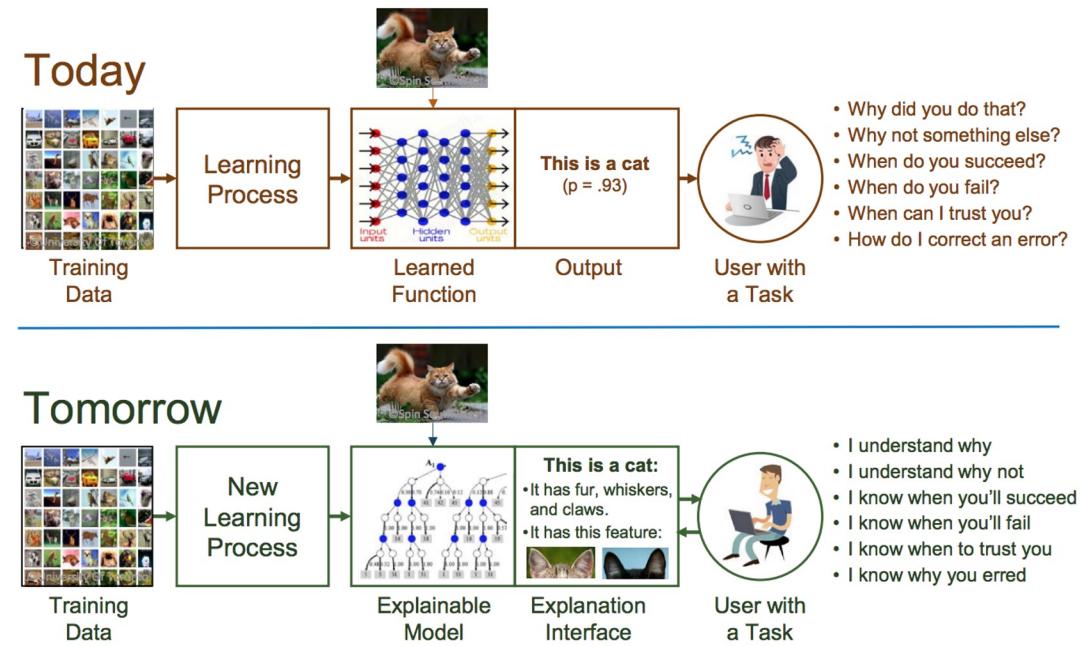




Do we understand the features learned by these models?

- Even if the models achieve high performances, **it is not trivial to assure that they are learning features that are relevant for that domain** (i.e., black box behaviour)
 - Machine learning models **are good at extracting correlations**
- While this **may not be an issue in several domains** (e.g., recommendation systems), in others, it is of utmost importance that the **system is capable of transparently showing** the reasons behind its decisions (e.g., healthcare)

Figure - The future of machine learning algorithms
(Image from [1])





Responsible AI: Modern problems require modern solutions

- **Responsible AI** is a framework that guides how we should address the challenges around artificial intelligence from both an **ethical, technical and legal** point of view^[1]
 - We must resolve ambiguity for where responsibility lies if something goes wrong!
- This framework relies on fundamental principles^[2]:
 - Accountability
 - Interpretability
 - Fairness
 - Safety
 - Privacy

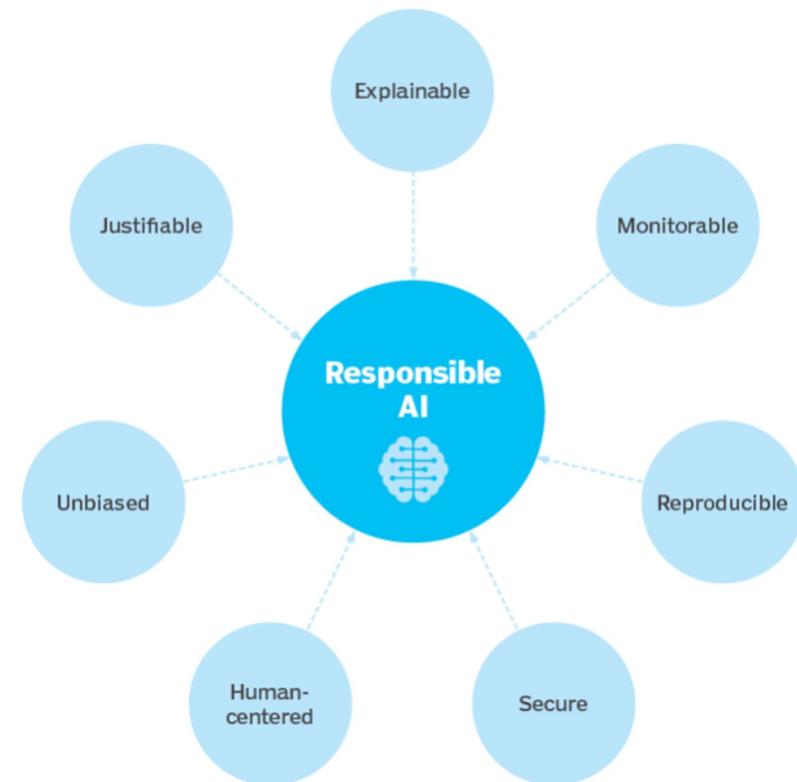


Figure - Responsible AI (Image from [1])



2. Regulating Science: Research Integrity and Ethics



My first contact with Research Integrity and Ethics happened during my PhD with a course on the topic

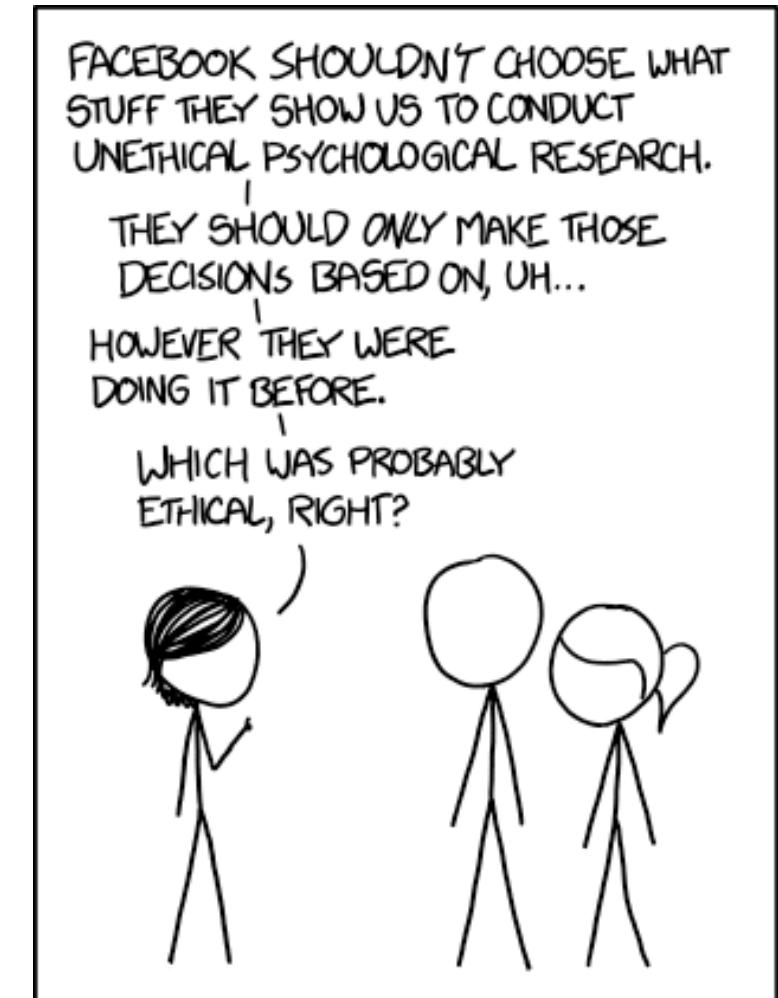
- I had the privilege of participating in the **3rd Edition of the Training Course in Research Ethics/ Integrity**, organised by Susana Magalhães (**Unit for Responsible Conduct in Research, i3S**)





This course presented 8 fundamental topics

1. Fundamentals of Bioethics and Research Ethics
2. Good Researcher what it is and why it matters
3. Preventing and managing conflicts in research
4. Authorship and Publication Ethics
5. Misconduct and Unacceptable Practices
6. Data Protection and Intellectual Property
7. Open Science
8. Science Communication and Citizen Engagement and Vulnerability and Equity in Research





Ethics in Science: Lessons learned from World War II

1. Scientific progress does not always result in greater good for humanity (it can also lead to suffering and death)
2. Science is not a value in itself but must remain an instrument for realising human goals (the goals do not justify all means)
3. All knowledge has a practical application, so scientists must take responsibility for foreseeing the possible consequences of the knowledge they build (and preventing its harmful uses)
4. While Science answers for what we can do, Ethics will state what we should do, taking as a criterion the human value
5. Science's self-regulation is not enough to guarantee the goodness of its ends; we need Ethical scrutiny to achieve this
6. **New technologies have lost their status as inert and passive to become dynamic and active, no longer depending solely on the designs of the user**



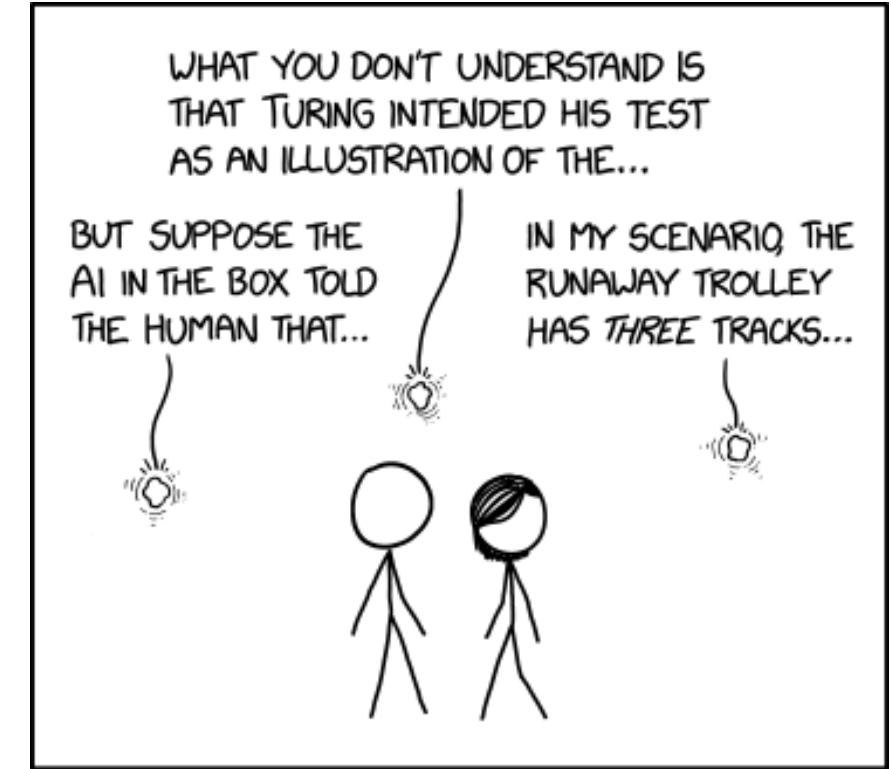
Research Integrity in Science: Why it matters

Scientific Research Integrity sets out the structuring of Ethical principles for responsible scientific research and innovation:

- Truth, rigour and objectivity
- Independence, impartiality and impartiality
- Co-operation and honesty
- Transparency and justice
- Commitment and social responsibility

We still need an Ethical reflection applied to scientific research to:

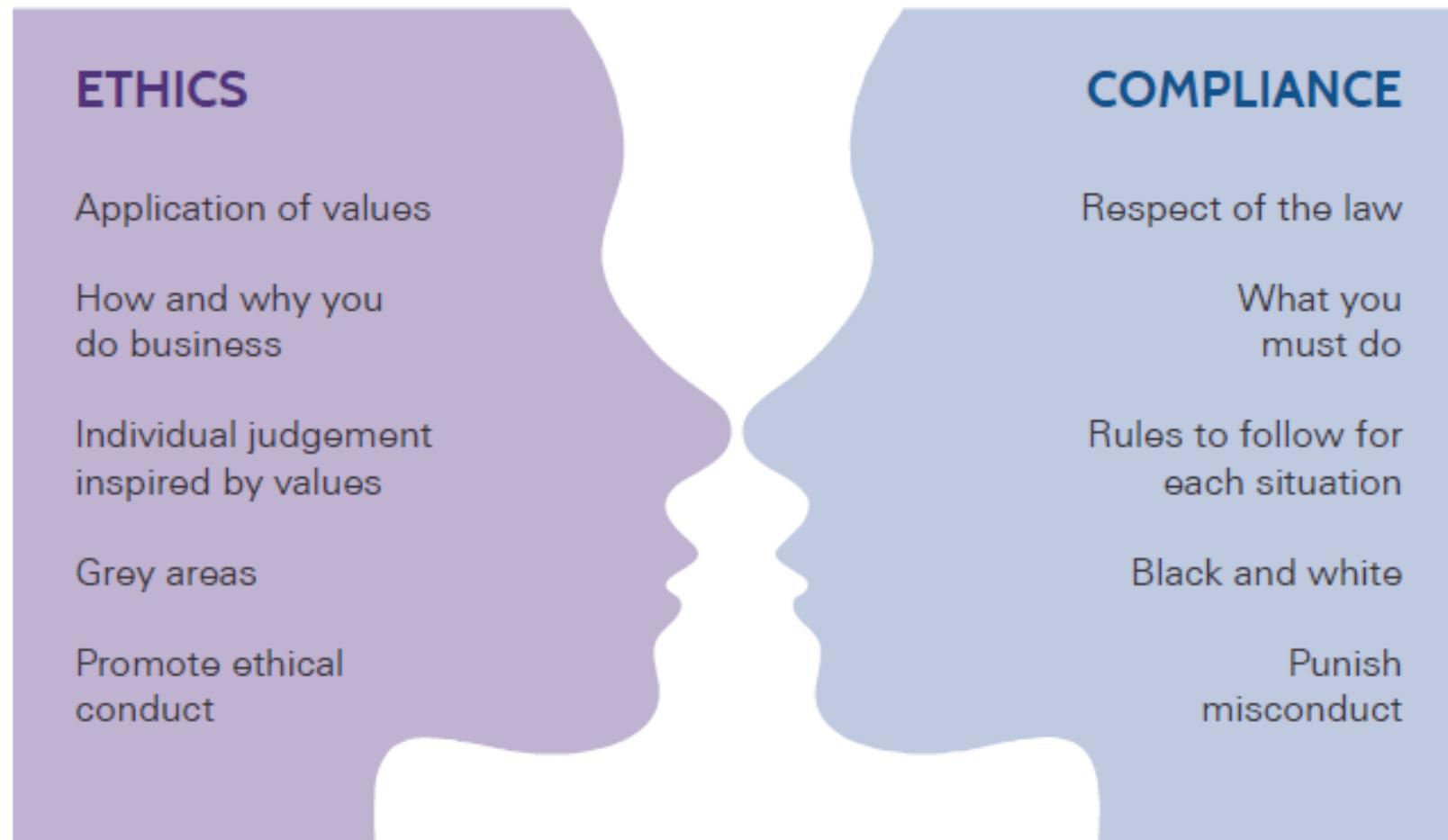
1. Ensure that scientific progress and technological innovation conform to common morality
2. Guarantee the goodness of its aims and impacts
3. Maintain the human as the purpose of human creation



IN RETROSPECT, GIVEN THAT THE SUPERINTELLIGENT AIs WERE ALL CREATED BY AI RESEARCHERS, WHAT HAPPENED SHOULDN'T HAVE BEEN A SURPRISE.

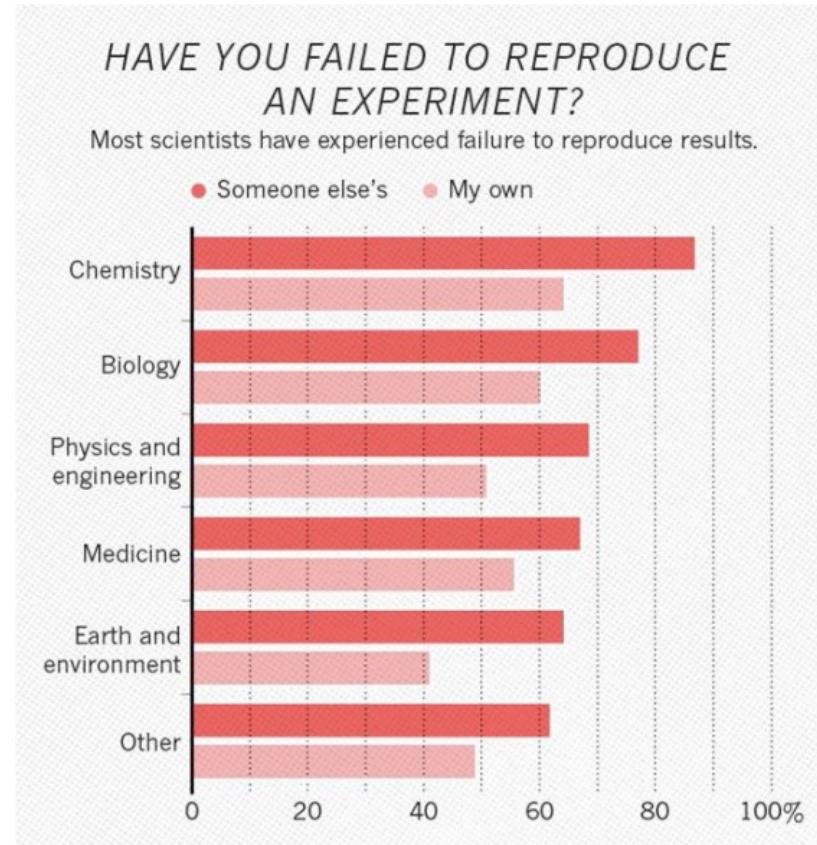
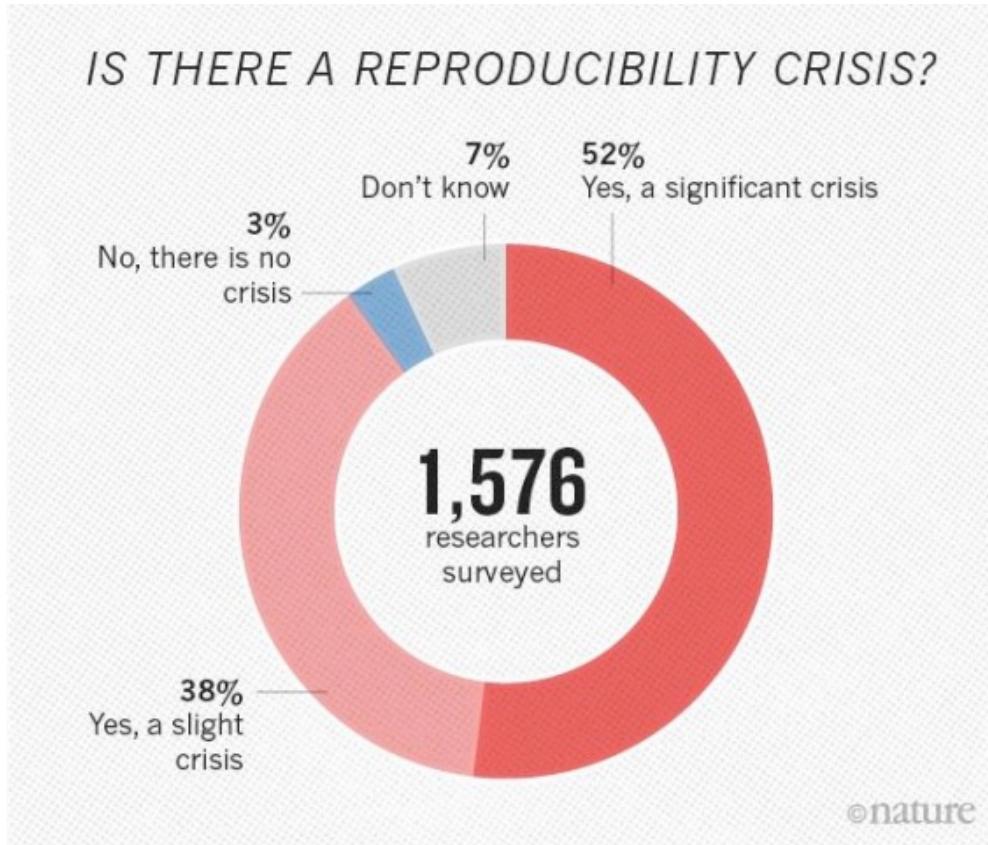


The duality between Ethics and Integrity (or Compliance)





Hot Topic: Ensuring that Science is Reproducible^[1]





All you need is “The Embassy of Good Science”[1]

The screenshot shows the homepage of The Embassy of Good Science. The header features a yellow navigation bar with the site's name on the left and links for Home, Themes, Reports, Resources, Community, Training, and About. A search bar and a user icon are on the right. The main content area has a large yellow background with the text "Your platform for research integrity and ethics" in bold black font. Below this is a video thumbnail for "The Embassy story" (2:07 min). At the bottom left is a link to "What's here for me?", and at the bottom right is a yellow circular icon with a question mark.

the embassy of good science

home themes reports resources community training about

search the embassy Q 👤

Your platform for research integrity and ethics

▶ The Embassy story 2:07 min

[What's here for me?](#)

?

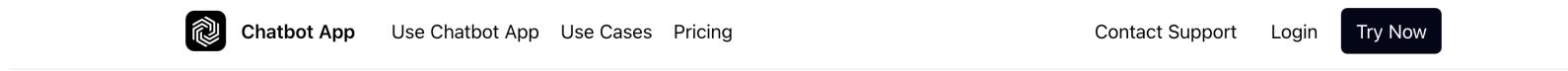


3. Artificial Intelligence meets Science



Should we ChatGPT all over the way?

- Nowadays, my perception is that everyone uses ChatGPT for everything, but are we using it correctly?



Chatbot App: All-in-One AI Chatbot

Chatbot App offers an intuitive user interface for accessing large language models, including GPT-4o, Claude 3.5 Sonnet, and Google Gemini, all at an affordable price with a single membership.

To start using Chatbot App on the web, mobile, and desktop, simply click or tap the 'Get Started' button.

Get Started

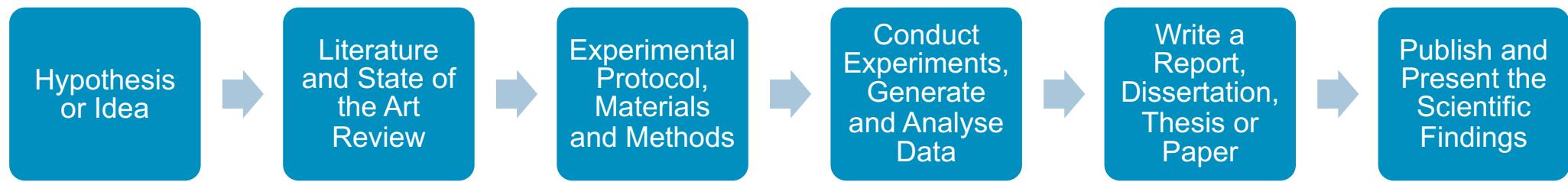
Accessible at:



Sources: <https://chatbotapp.ai/>



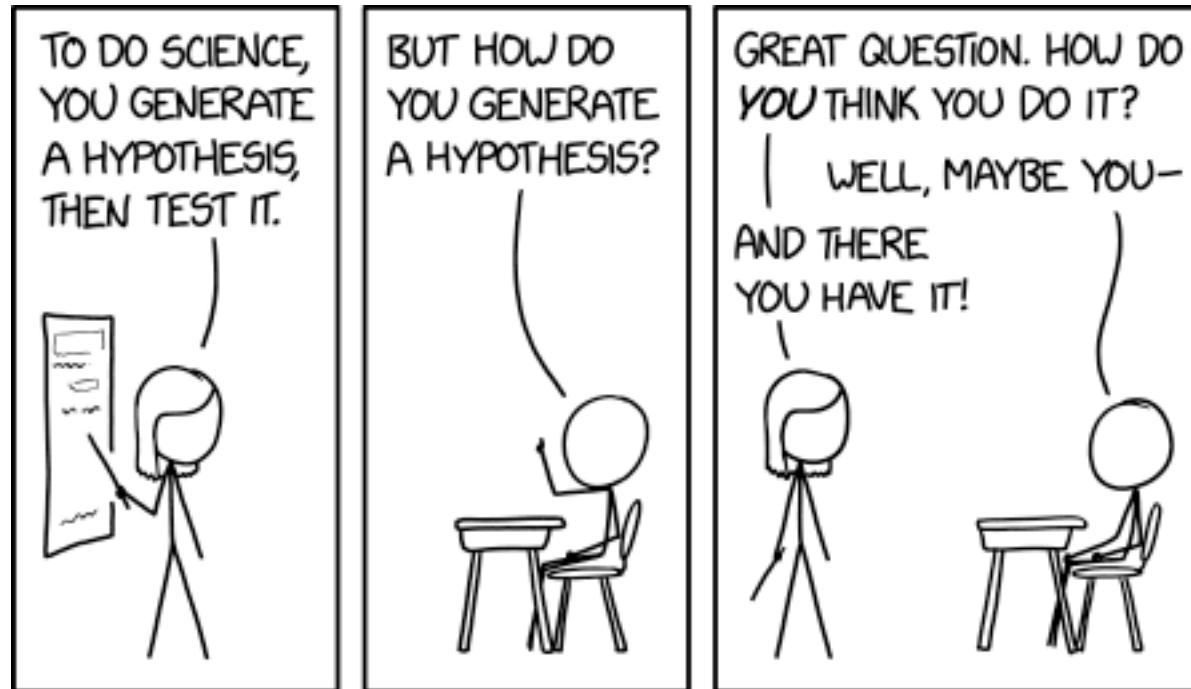
From research to publication: the typical pipeline





Hypothesis or Idea

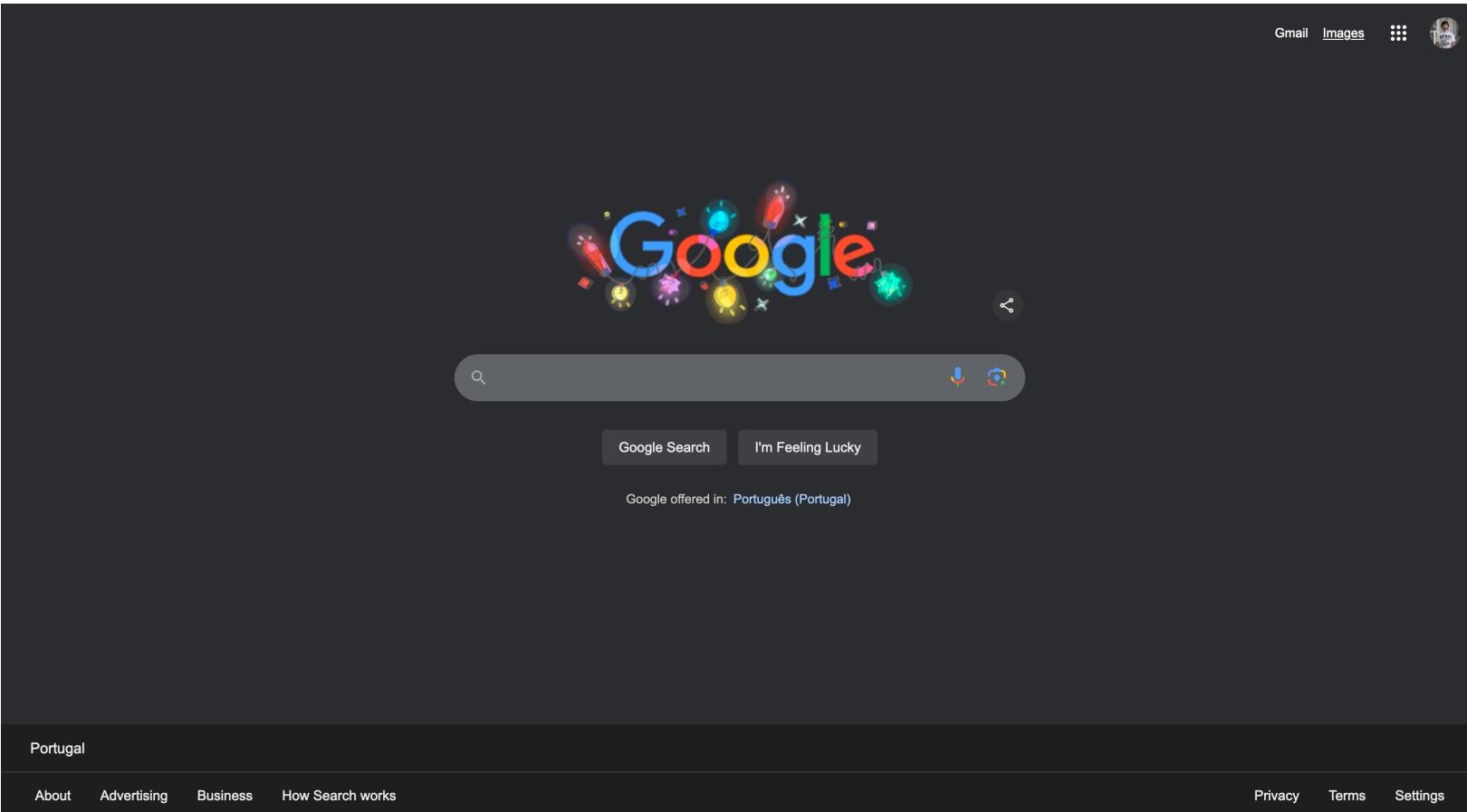
- Sometimes, you don't need AI...





Hypothesis or Idea

- Google^[1] is still a nice and (fairly) good research engine





Hypothesis or Idea

- **Perplexity^[1]** might be a good starting point to search your hypotheses or ideas

The screenshot shows the Perplexity AI interface. At the top left is the logo and navigation links: 'perplexity' (with a magnifying glass icon), 'New Thread', and a search bar with a 'K' icon. On the left is a sidebar with 'Home', 'Discover', 'Spaces', and 'Library'. The main area has a dark background with white text. It features a large input field with the placeholder 'Ask anything...'. Below it are buttons for 'Focus' and 'Attach', and a 'Pro' toggle switch. A question 'What do you want to know?' is centered above the input field. At the bottom of the input field are four recent search queries: 'Recent research on cold plunge therapy', 'Preventing spam phone calls', 'Global university rankings 2024', and 'Ski resort with most snow this season'. At the bottom left are 'Sign Up' and 'Log in' buttons. At the bottom right are links for 'Pro', 'Enterprise', 'Store', 'Blog', 'Careers', 'English (English) ▾', and a help icon.



Literature and State of the Art Review

- **Elicit^[1]** allows researchers to get a list of relevant references/articles/papers related to the question they ask in the platform

The screenshot shows the homepage of the Elicit website. At the top, there is a navigation bar with links for 'Elicit', 'Features', 'Testimonials', 'Pricing', 'FAQ', 'Careers', 'Sign In', and a prominent 'Sign Up' button. The main visual is a large green gradient background with the text 'Analyze research papers at superhuman speed' in white. Below this, a subtext reads: 'Automate time-consuming research tasks like summarizing papers, extracting data, and synthesizing your findings.' There are two 'Sign Up' buttons, one above the other, and a link 'Learn More'. At the bottom, a section titled 'TRUSTED BY RESEARCHERS AT' lists several logos: GOV.UK, Google, Stanford, THE WORLD BANK, and NASA.

Analyze research papers at superhuman speed

Automate time-consuming research tasks like summarizing papers, extracting data, and synthesizing your findings.

Sign Up

Or

Learn More

TRUSTED BY RESEARCHERS AT

GOV.UK Google Stanford THE WORLD BANK NASA



Literature and State of the Art Review

- **Consensus^[1]** helps academics to get references and appropriate literature related to their questions

The screenshot shows the homepage of the Consensus AI Search Engine. At the top, there is a navigation bar with the Consensus logo, Product, Resources, Pricing, About, Contact, Login, and Try for free buttons. Below the navigation bar, the main title "AI Search Engine for Research" is displayed in large, dark font. A subtitle "Find & understand the best science, faster." is shown below the title. A search bar with the placeholder "Ask the research..." and a magnifying glass icon is centered. Below the search bar, there are four example search queries: "Does exercise improve cognition?", "Can cash transfers reduce poverty?", "Are statins effective in the elderly?", and "Can mindfulness help with sleep?". Each query has a small icon and a question mark icon. At the bottom of the page, there is a link "Try an example search".



Literature and State of the Art Review

- **Scispace^[1]** aims to help researchers to understand research papers better

The screenshot shows the Scispace homepage. At the top, there is a navigation bar with the SCISPACE logo, Pricing, Chat with PDF, Login, and Sign up buttons. On the left, a vertical sidebar contains icons for Home, Chat, Search, Write, Share, Lists, Quotes, Flashcards, AI, Document, Megaphone, Google Chrome, and a neural network icon. A circular profile picture placeholder is at the bottom. The main content area features the heading "The Fastest Research Platform Ever" and the subtext "All-in-one AI tools for students and researchers." Below this is a search bar with the placeholder "Get insights from top papers directly" and a magnifying glass icon. A section titled "Try asking or searching for:" lists several AI-generated questions: "How does climate change impact biodiversity?", "Why are aging Covid patients more susceptible to severe complications?", "How does social media affect the college selection process?", "What are the interesting theories about dark matter and dark energy?", and "What is the significance of higher-dimensional algebra?". At the bottom, there is a "Popular Tools" section with two cards: "Chat with PDF" (with a speech bubble icon) and "AI Writer" (with a pen icon). The "Chat with PDF" card includes the subtext "Get all answers backed by citations." and the "AI Writer" card includes the subtext "Write new research papers. Assisted by AI."

Sources: [1] <https://typeset.io/>



Literature and State of the Art Review

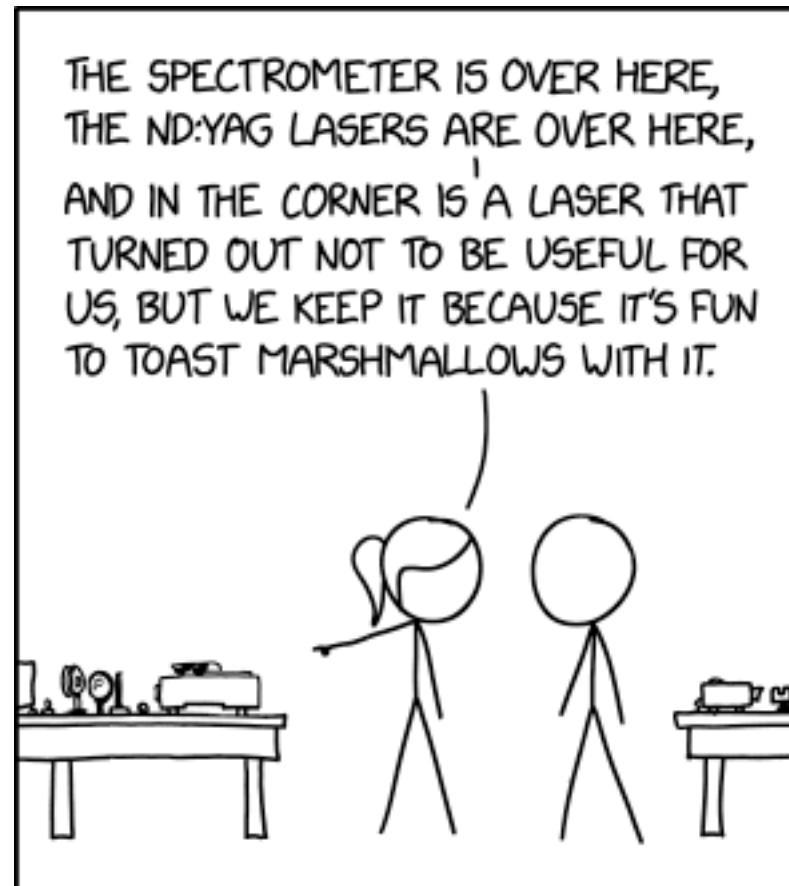
- ChatPDF^[1] lets you chat with any PDF

The image shows the homepage of ChatPDF. At the top, there are three buttons: '+ New Chat', '+ New Folder', and a 'Sign in' button. Below these, there are two small thumbnail images of previous chats. A text prompt encourages users to 'Sign in for free to save your chat history'. On the right side of the main content area, there are three accolades: '#1 PDF Chat AI Original', 'Q's answered every day 1,000,000+', and 'Gen AI apps of 2024 Top 50'. The main heading 'Chat with any PDF' is displayed prominently. Below it, a subtext reads: 'Join millions of [students, researchers and professionals](#) to instantly answer questions and understand research with AI'. A large dashed-dotted rectangular area is provided for uploading a PDF file, with the text 'Click to upload, or drag PDF here' and a purple 'Upload PDF' button. A purple arrow points from the text 'DRAG + DROP YOUR PDF FILE HERE' to the top right corner of the dashed-dotted area. The bottom left corner of the image shows a dark mode sidebar with language selection ('EN') and other settings.



Experimental Protocol, Materials and Methods

- Sometimes, you can't use AI...



EVERY LAB IN EVERY FIELD HAS
SOME PIECE OF EQUIPMENT LIKE THIS.



Conduct Experiments, Generate and Analyse Data

- Julius^[1] intends to help you with data analysis

[Use Cases ▾](#)**Julius**[Docs](#)[Community](#)[Sign Up](#)[Sign In](#)

Analyze your data with computational AI.

Chat with your files and get expert-level insights in seconds.

[Try Julius AI >](#)

Loved by over 1,200,000 users worldwide



Charts & Graphs

Create sleek looking data visualizations.

[Try now →](#)

Insights

Ask anything to your data, and get answers.

Advanced Analysis

Perform modeling and predictive forecasting.

Problem Solving

Scan-and-solve math, physics, and chemistry.

Reports

Generate polished analyses and summaries.

[Show Code ▾](#)

Here's the pairplot visualizing the relationships between the features of the Iris dataset, colored by species:





Write a Report, Dissertation, Thesis or Paper

- Overleaf^[1] is not about AI (although it has some AI running in the background), but it will (probably) be your best friend during the writing process



Features & Benefits ▾ Templates Plans & Pricing Help ▾ Sign up Log in

\begin{anything}

Write like a rocket scientist with Overleaf
—the collaborative, online LaTeX editor that *anyone* can use.

OR

Enter your email Enter your password

Sign up for free

By registering, you agree to our [terms of service](#) and [privacy notice](#).

The screenshot shows the Overleaf web interface. At the top, there's a navigation bar with links for Features & Benefits, Templates, Plans & Pricing, Help, Sign up (in a green button), and Log in. Below the navigation is a large purple header with the text '\begin{anything}' in a stylized font. Underneath the header, there's a sub-headline: 'Write like a rocket scientist with Overleaf —the collaborative, online LaTeX editor that *anyone* can use.' Below this are two sign-up buttons: 'Sign up with Google' and 'Sign up with ORCID'. A 'Sign up for free' button is prominently displayed in a green rounded rectangle. At the bottom of the sign-up section, there's a small note: 'By registering, you agree to our [terms of service](#) and [privacy notice](#)'. The main area of the interface shows two open LaTeX documents. The left document is titled 'My Paper on Astronomy and Computing' and contains a mathematical equation:
$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0$$
. The right document is titled 'Exploring the Nexus of Astronomy and Computing' and includes a small image of a galaxy.



Write a Report, Dissertation, Thesis or Paper

- Grammarly^[1] uses AI to help you write better



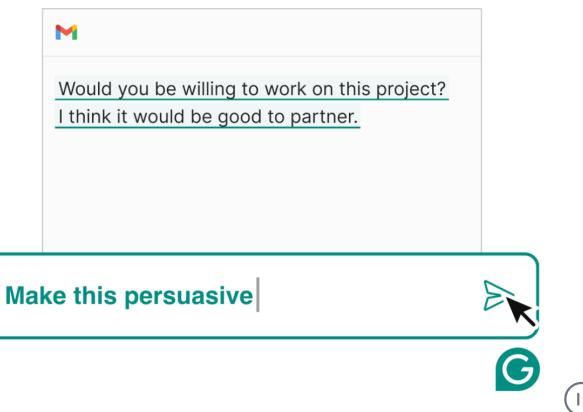
Responsible AI that ensures your writing and reputation shine

Work with an AI writing partner that helps you find the words you need—to write that tricky email, to get your point across, to keep your work moving.

[Sign up It's free →](#)

[Sign up with Google](#)

By signing up, you agree to the [Terms and Conditions](#) and [Privacy Policy](#). California residents, see our [CA Notice at Collection](#).



Trusted by 70,000+ teams and 30 million people



EVERLANE

BlackRock

zoom

Expedia



ATLASSIAN

align





Write a Report, Dissertation, Thesis or Paper

- DeepL^[1] will help you to translate the difficult stuff

The screenshot shows the DeepL Translator homepage. At the top, there is a navigation bar with the DeepL logo, 'Products' (New), 'Solutions', 'Pricing', 'Apps', and a 'Start free trial' button. Below the navigation bar are three main translation options: 'Translate text' (33 languages), 'Translate files' (.pdf, .docx, .pptx), and 'DeepL Write' (AI-powered edits). The main workspace has two dropdown menus: 'Detect language' and 'English (British)'. A large text input area is labeled 'Type to translate.' with instructions to drag and drop files or click the microphone to translate speech. Below this is a 'Dictionary' section with the instruction 'Click on a word to look it up.' On the left side of the workspace, there are two small icons: a blue square with a white 'A' and a white square with a blue 'A'.



Publish and Present the Scientific Findings

- **Gamma^[1]** might give you some inspiration when creating your presentations

The screenshot shows the homepage of the Gamma AI website. At the top left is the Gamma logo (a stylized 'G'). At the top right are links for English, Pricing, Careers, Login, and a 'Try for free' button. The main visual features a large, colorful illustration of a person's head and shoulders. Inside the head, a smaller figure wearing a space helmet stands on a staircase made of floating rectangular blocks, looking up at a starry galaxy. The text 'A new medium for presenting ideas.' is displayed in a large, multi-colored font (purple, red, orange) across the center of the head. Below this, the text 'Powered by AI.' is shown in a dark font. Underneath the main title, there are two descriptive lines: 'Beautiful presentations, documents, and websites.' and 'No design or coding skills required.' A blue 'Sign up for free' button is located below these lines. At the bottom of the page is a small thumbnail image showing a person's hands working on a laptop.



4. Surfing the Waves of AI: Take-home Messages



Can we trust AI? Towards “Trustworthy AI” in the EU

- The European Commission appointed a group of experts to provide advice on its artificial intelligence strategy: **High-Level Expert Group on AI**^[1]
- **According to the Guidelines, trustworthy AI should be:**
 - **Lawful:** respecting all applicable laws and regulations
 - **Ethical:** respecting ethical principles and values
 - **Robust:** both from a technical perspective while taking into account its social environment
- Several important **guidelines were proposed**^[2]:
 - **Human agency and oversight:** AI systems should empower human beings
 - **Technical Robustness and safety:** AI systems need to be resilient and secure
 - **Privacy and data governance:** data governance mechanisms must be ensured
 - **Transparency:** the data, system and AI business models should be transparent
 - **Diversity, non-discrimination and fairness:** AI systems should be accessible to all
 - **Societal and environmental well-being:** AI systems should benefit all human beings
 - **Accountability:** ensure responsibility and accountability for AI systems and their outcomes



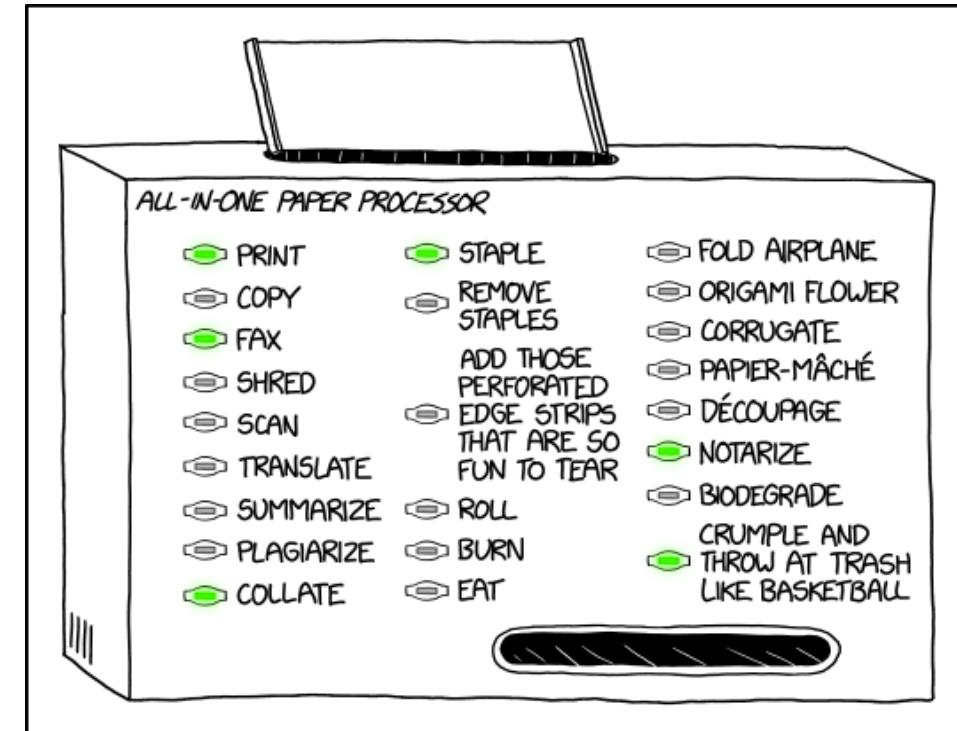
The AI Act and how it will impact our lives

- The AI Act^[1] is a **document proposed by the European Commission** that contains several **harmonised rules^[2]** regarding **AI applications**, emphasising that its approach is shaped by EU values and **risk-based**, ensuring both **safety** and **fundamental rights protection**
- What does the AI Act propose?^[2]
 - **Prohibition of unacceptable AI practices** (e.g., social scoring)
 - **Regulation of high-risk AI systems** (e.g., AI used in the context of recruitment)
 - **Conformity assessment** (i.e., under the EU product safety framework)
 - **Transparency obligations for potentially deceptive AI systems**
 - **Ex post market surveillance** (i.e., post-market monitoring system)
 - **Governance** (i.e., authorities must be appointed for the application and implementation)
 - **Pre-emption of national AI regulatory frameworks** (i.e., regulated by the EU)
 - **Monitoring and enforcement** (i.e., done by the Member States)
 - **Compliance with the prohibitions and regulatory requirements**



An Accurate and Honest Summary of this Session

- The development of data-driven artificial intelligence applications is impacting our lives, motivating the need for ethical, legal and technical regulatory frameworks based on specific principles: accountability, interpretability, fairness, safety, privacy
- Think about AI software and applications as tools: worry about knowing how these algorithms work and how you can leverage their power to improve the quality of your research and work
- Always bear in mind the principles of Ethics and Research Integrity, and think of the impacts of using AI-driven applications in your research and work
- Multidisciplinary work is, more than ever, of utmost importance and useful



Artificial Intelligence Applied to Scientific Research and Ethics: How can we leverage AI algorithms to achieve better science?

Mestrado em Energias Sustentáveis | Workshop

December 10, 2024 | Instituto Superior de Engenharia do Porto

Tiago Filipe Sousa Gonçalves

<https://tiagofilipesousagoncalves.github.io/>



INSTITUTE FOR SYSTEMS
AND COMPUTER ENGINEERING,
TECHNOLOGY AND SCIENCE

