

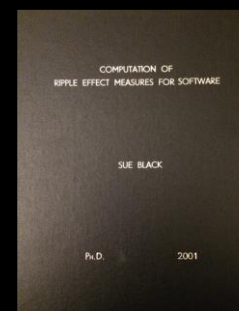
Trust Me, I'm an Algorithm: The Missing Link in AI Governance

By

Professor Sue Black OBE

Durham University, UK

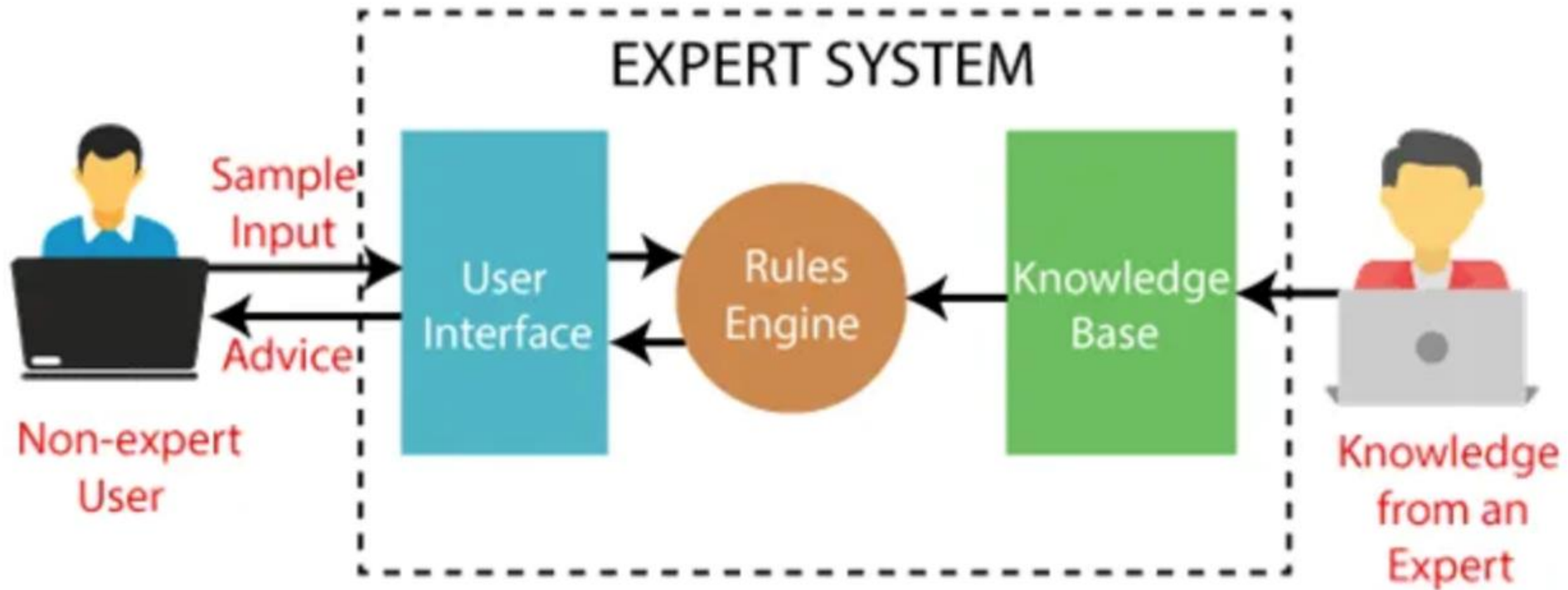




Talk to Eliza by typing your questions and answers in the input box.

> Hello, I am Eliza. I'll be your therapist today.
* Hi Eliza
> What would you like to discuss in this session?
* I'd like to discuss my relationships
> I'm not sure I understand you fully.
* my relationship with my children
> What does that suggest to you?
* That i have children
> That is an interesting response. Let me make anote of that for next time.

TYPE HERE



Source: [Expert Systems in Artificial Intelligence — Javatpoint](#)

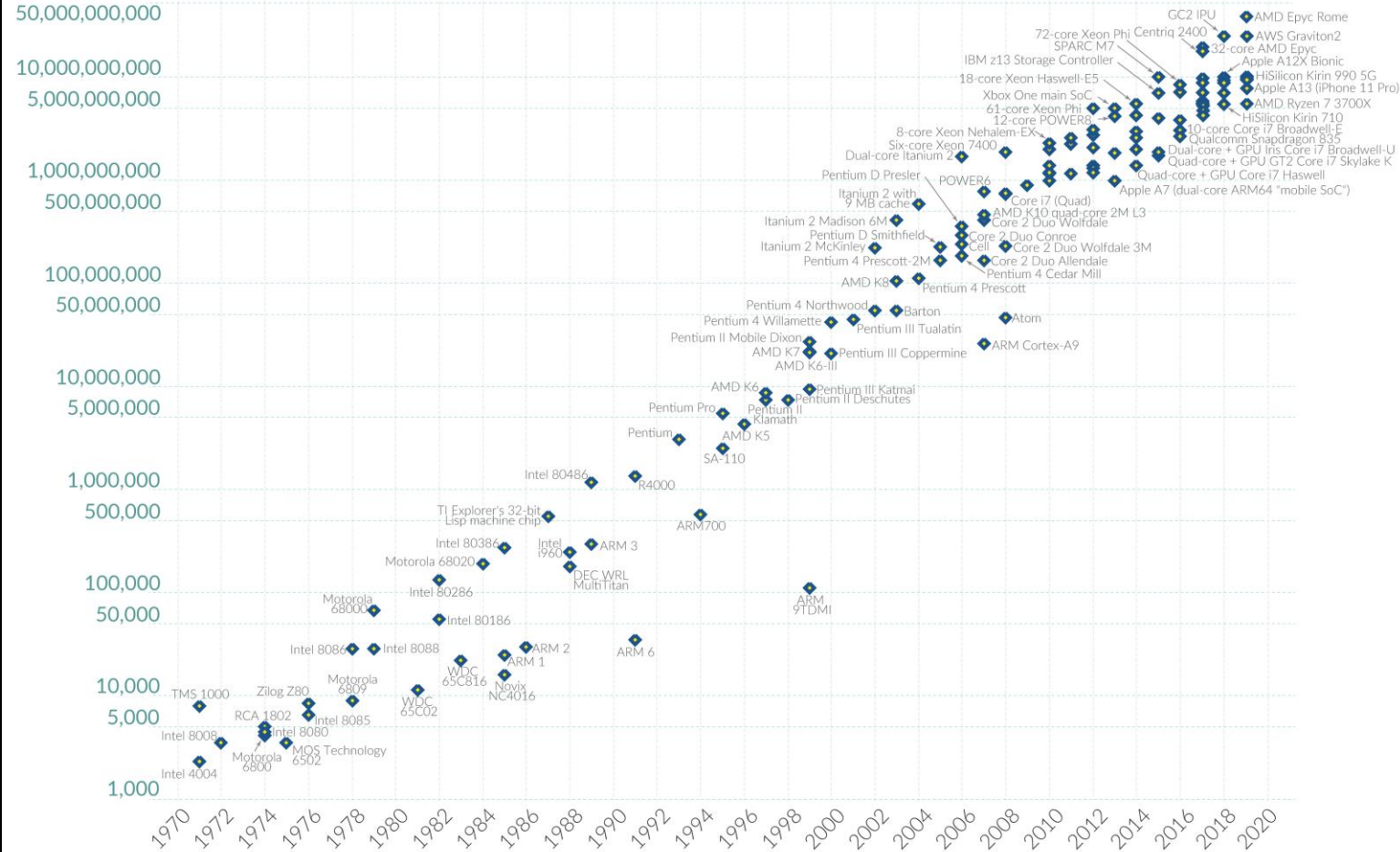


Moore's Law: The number of transistors on microchips doubles every two years



Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

Transistor count



Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)
 OurWorldinData.org – Research and data to make progress against the world's largest problems. Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

DALL·E

Some of the most popular tools based on Large Language Models:



- Chatbots and Virtual Assistants
- Content Creation
- Language Translation
- Text Analysis and Summarisation
- Personalised Recommendation
- Education
- Research
- Information Retrieval



**Governance/Regulation +
?**

= Responsible AI

Home / Business / 'Qatar leads in AI regulation implementation'

'Qatar leads in AI regulation implementation'

25/05/2023 / Business



BREAKING | Bashar al-Assad's whereabouts remain unknown after reports he fled the capital

Economy | Technology

AI needs 'global governance', cannot be left to market, UN panel says

Expert body makes seven recommendations on regulating AI, including establishment of fund to help developing countries.

B Election '24 U.S. Economy International Affairs

COMMENTARY

Reforming data regulation to advance AI governance in Africa

Chinasa T. Okolo
March 15, 2024



News Sport Business Money Opinion Ukraine Travel Health

See all Business

China and US sign Bletchley declaration warning of 'catastrophic' risks from AI

The EU and 27 countries agree to work together to address dangers from the technology



Participants at the Bletchley Park summit expressed concerns about risks posed by AI in cybersecurity and biotechnology Credit: Chris J. Ratcliffe/Bloomberg

News provider of the year

The Guardian

UK ▾

News

Opinion

Sport

Culture

Lifestyle



UK US politics World Climate crisis Middle East Ukraine Football Newsletters Business Environment



Universal credit

Revealed: bias found in AI system used to detect UK benefits fraud

Exclusive: Age, disability, marital status and nationality influence decisions to investigate claims, prompting fears of 'hurt first, fix later' approach





Bias in Large Language Models

Sarah Wyr (PhD Student)

Department of Computer Science
Durham University, UK





What is the UK's Post Office IT scandal about and who is involved?

Prosecution of thousands of people has been called one of the greatest miscarriages of justice in British history



AI's mysterious 'black box' problem, explained



Graphic by Violet Dashi. Illustrations by Nadia and Simple Line via Adobe Stock

QATAR / QATAR

Kudos to Afghan Girls Robotics Team

DOHA

PUBLISHED ON MARCH 27, 2022 | 12:12 AM



“If all students in low-income countries had just basic reading skills (and nothing else), an estimated 171 million people could escape extreme poverty.”

United Nations Educational, Scientific, and Cultural Organization UNESCO

“A well-educated workforce is more productive, innovative and adaptable to changing market conditions.”

<https://www.timeshighereducation.com/campus/how-convey-benefits-higher-education-government-and-policymakers>

“A country’s economy becomes more productive as the proportion of educated workers increases since educated workers can more efficiently carry out tasks that require literacy and critical thinking”

<https://www.investopedia.com/articles/economics/09/education-training-advantages.asp#:~:text=A%20country's%20economy%20becomes%20more,education%20also%20carries%20a%20cost.>



Regulation



29% reduction in
fatal injuries

+

Education



250,000 children's lives saved

=

60% reduction
in fatalities

<https://www.sernis.com/news/being-the-green-cross-man-beats-being-darth-vader-any-day>
<https://www.pacts.org.uk/40-years-of-uk-seatbelt-wearing-saves-thousands-of-lives-2/>



📷 Tobacco accounts for more than 1 million deaths a year and 10% of all deaths in Europe, the WHO estimates. Photograph: Roberto La Rosa/Alamy

Smoking tobacco

Regulation - bans in public places+ tax increases

+

Education – Campaigns about the negative effects of smoking

UK 2011 19.8%

UK 2023 11.6%

One in three adults in UK and Ireland eat five or more daily portions of fruit and veg

England bans single-use plastic: What is and isn't included in the new rules?



Copyright Unsplash

Regulation – “Sugar tax”

+

Education – “5 a day”

Regulation – Single use plastic ban

+

Education – “Bag for Life”

**Governance/Regulation +
Education**

= Responsible AI



Create an image of me taking a selfie from the stage at the AI summit in Qatar with the audience in the background



Professor Sue Black OBE
Durham University, UK





Professor Sue Black OBE • You



Professor of Computer Science and Technology Evangelist, Forbes World To...

1d •

🤔 Is AI good or evil? Or both? I would love to know what you think.

❤️ Please take 2 minutes to answer my quick questions and share

🙌 Thank you so much, I will send you the results

#AI #poll #questions

The screenshot shows a Google Docs poll form with the following sections:

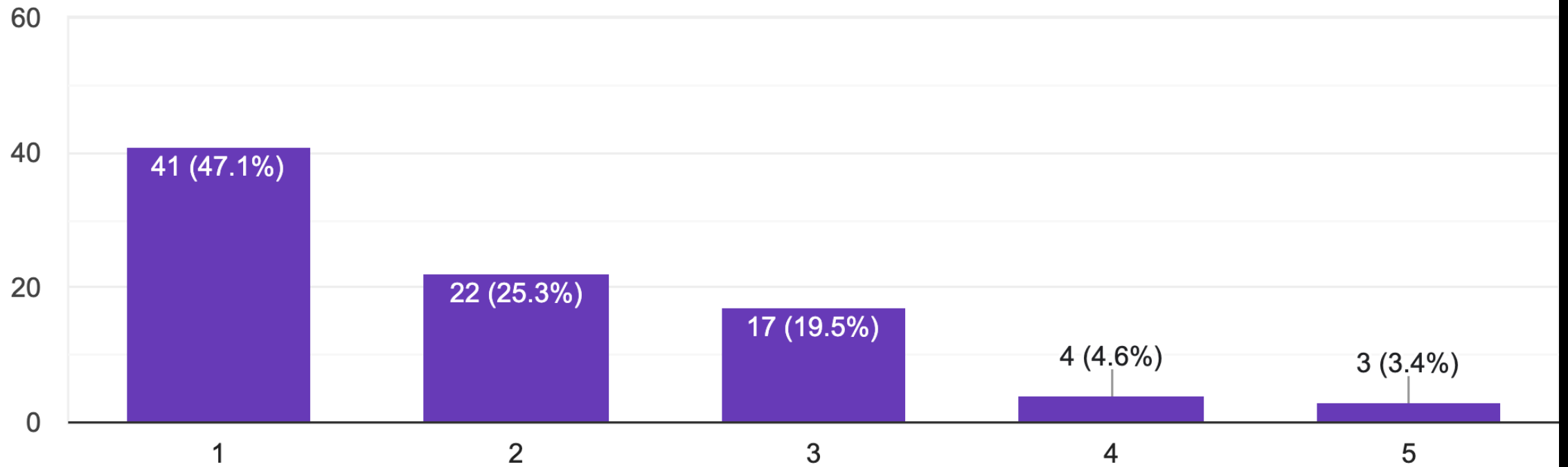
- Title:** Is AI good or evil?
- Intro:** Please answer these short questions to help us understand what people think you of AI. There are no right or wrong answers, thank you very much. I will email you the results.
- Question 1:** Email? (with a text input field for "Your email")
- Question 2:** Are you worried about AI? (with radio buttons for "Yes" and "No")
- Question 3:** If you are worried about AI, what concerns you? (with a text input field)

Is AI good or evil?

docs.google.com

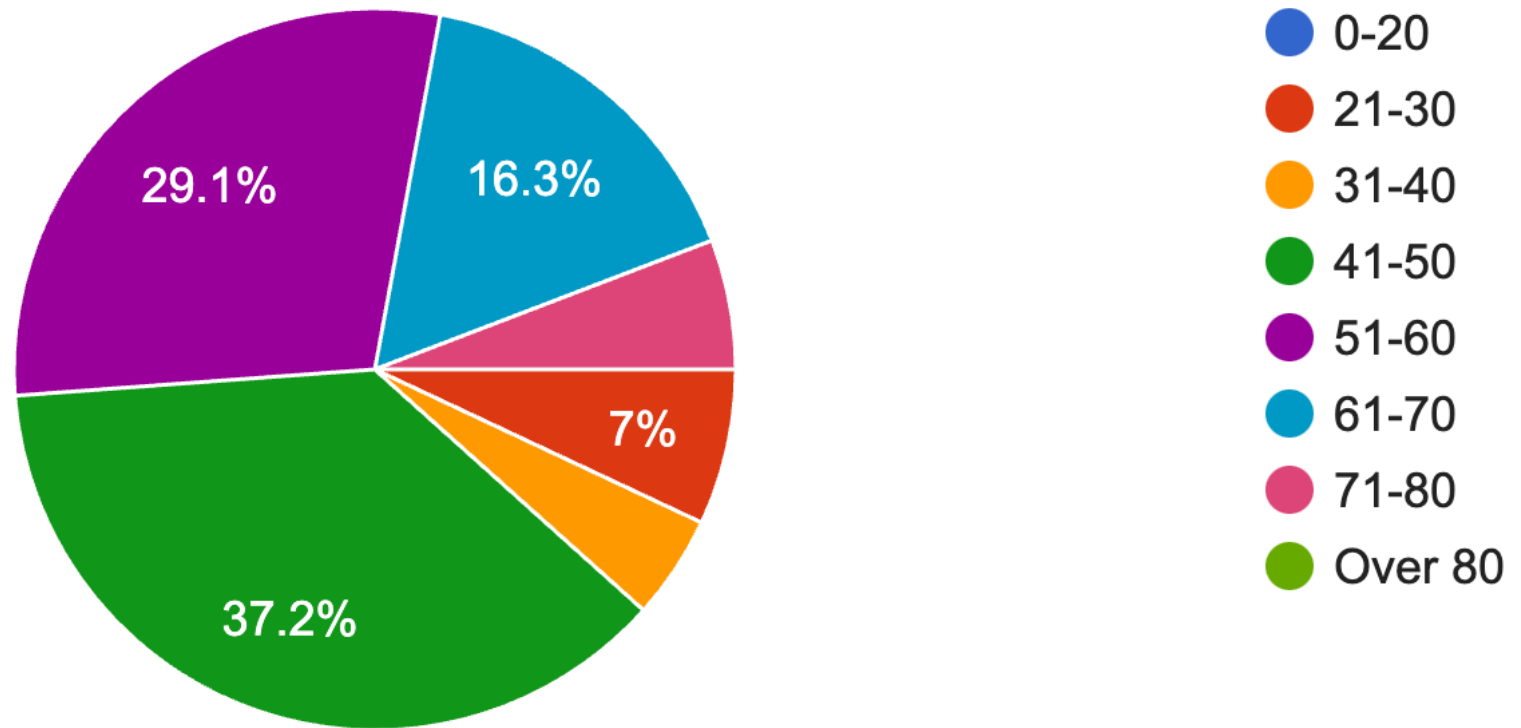
How confident are you using technology?

87 responses



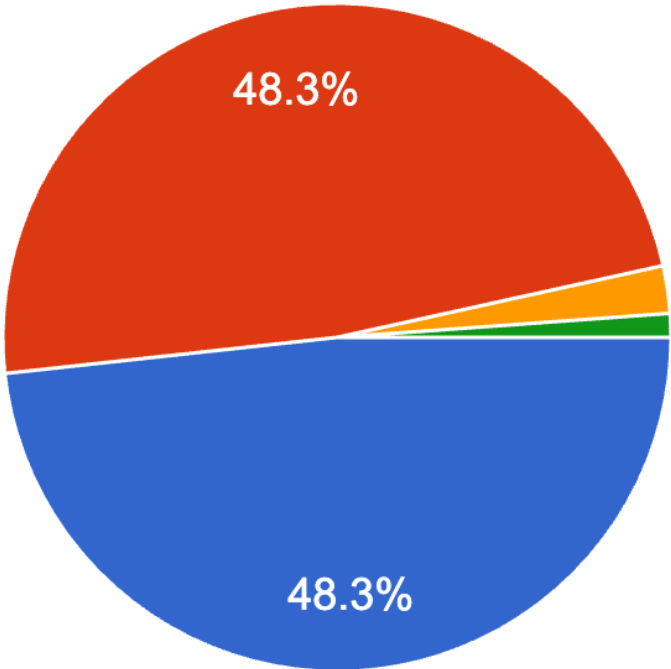
How old are you?

86 responses



What gender are you?

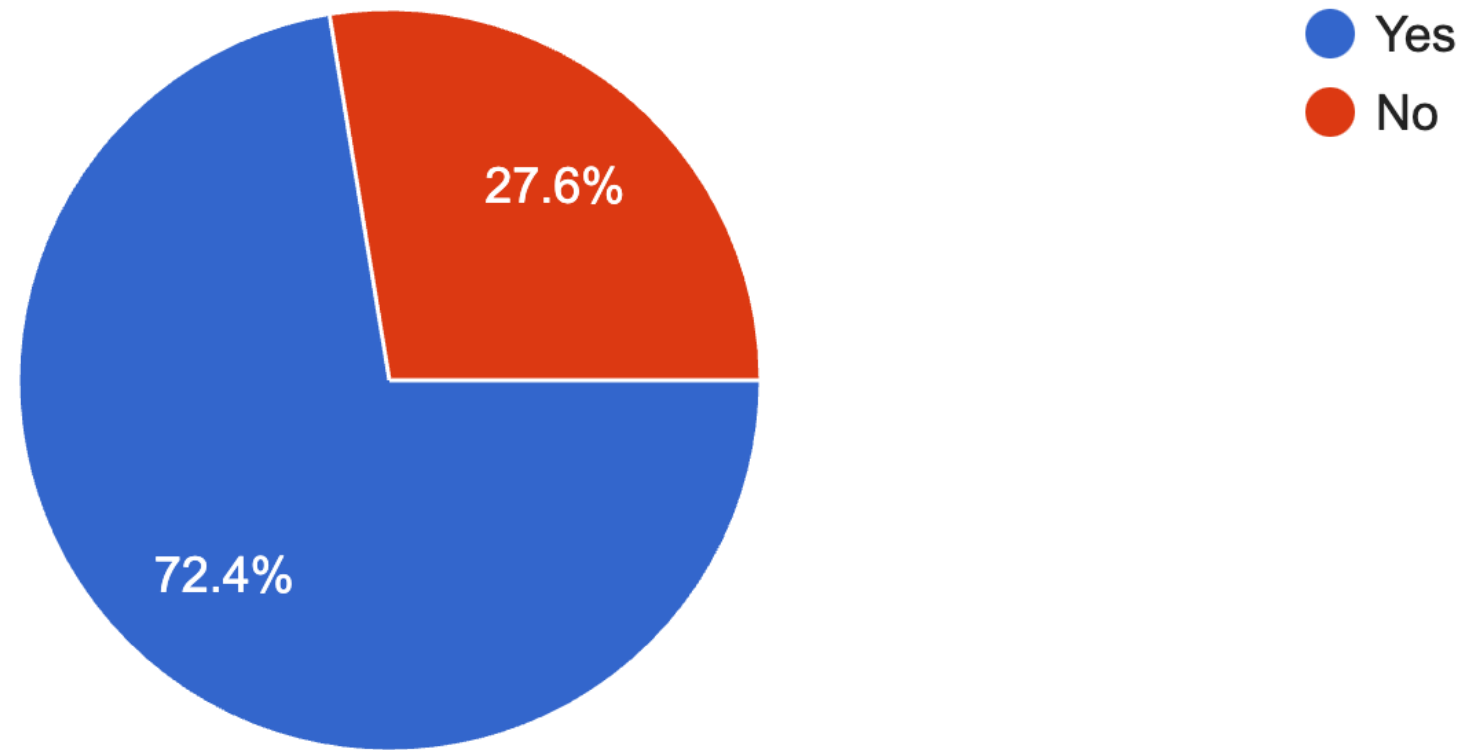
87 responses



- Female
- Male
- Non-binary
- Other/prefer not to say

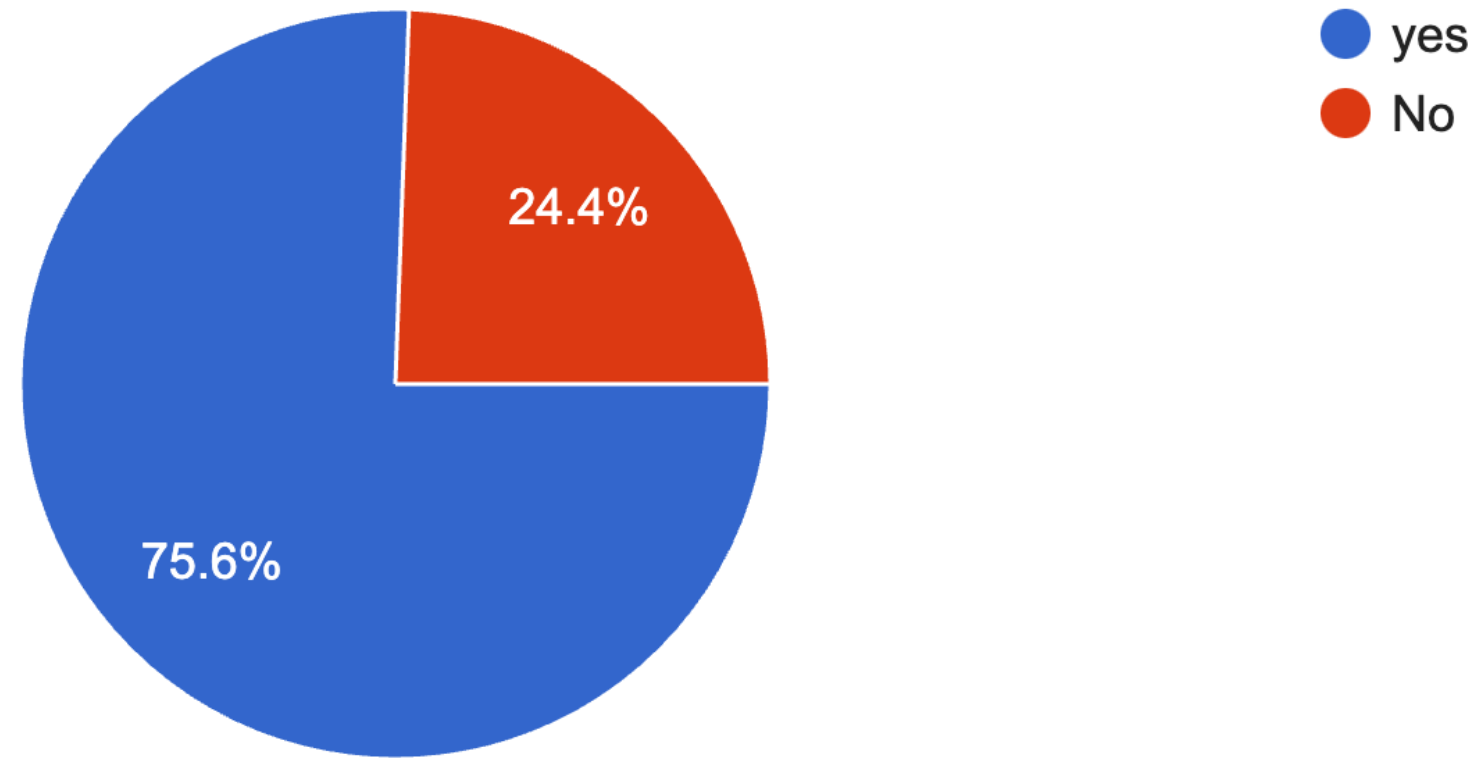
Are you worried about AI?

87 responses



Are you excited by AI?

86 responses



What concerns you about AI?

- That it is changing too fast for us to keep up
- **I'm not worried about software, I'm worried about the people creating the software**
- **I'm not worried about the technology, I'm more worried about who controls what, and the potential to manipulate people**
- Inaccuracies leading to harm, loss of human creativity, deepfakes used for harassment, scam/crime campaigns, misinformation
- AI created misinformation and the impact on society. Failure of systems and laws to adapt to AI quickly enough. Lack of training and support meaning that people are left behind or scared of AI
- Harms caused by biases in the data AI is trained on

What excites you about AI?

- The future
- The opportunities to drive efficiency
- The endless possibilities to change the world
- **Its potential to solve problems currently deemed to be too complex**
- A great aggregator to summarize knowledge made by others
- Breakthroughs in science (e.g. on antimicrobial resistance, protein structures)
- Mostly just incredible potential
- It can help with many accessibility needs
- The ability to spot patterns that humans can't see
- **The potential to do boring jobs leaving people to do the more interesting things**

Anything else you would like to tell me?

Don't think AI itself is the issue. It's who has access to its information and how that it is used

I think of AI a bit like a knife. Used correctly and safely, it's a tool that can help with a variety of tasks. In the wrong hands it can become a weapon. And it can be dangerous to use if you're not aware of the risks

It's not the AI I am worried about, it's the people's response that fills me with fear

Terrifying possibility and excitement about a new technology all rolled into one

My sons use it in their music making in an experimental way, sampling it and making new work. I think it's a great tool for artists to push for new ideas



+



⇒



Durham University

Now streaming on all major platforms...

100 MOMENTS THAT ROCKED COMPUTER SCIENCE

With Professor Sue Black OBE

 Alan Emtage	 Anne-Marie Imafidon MBE	 Bobby Seagull	 Dame Wendy Hall	 Sir Dermot Turing	 Megan Smith
--	--	--	--	--	--





