

How should we think smarter about AI?

Report by Belinda Schwehr

Farnham Humanists hosted a public meeting on the 13th November, about 'Artificial Intelligence - the Risks, Benefits and the Challenges'.

Humanists share a non-religious ethical outlook. Belinda Schwehr, Chair, told the 80-strong audience that Farnham Humanists had picked this topic for a public interest debate, because of the need to avoid the extremes of complacency or fear that are common currency at present.



David Wood (Chair, London Futurists) referenced the Bletchley Park Summit earlier this month and said it had taken sensible steps towards international safety standards for the development of generative AI ('frontier AI') as opposed to the current deep-learning large language models in use.

He called for people, business and governments to be techno-agile in the pursuit of checks and balances, using not just an accelerator but also the steering and brakes that could be available. He gave a rousing explanation of why techno-scepticism, techno-suppression or techno-promiscuity make no sense for anyone seriously interested in managing the risks presented, if the undoubted opportunities for good are to be harnessed.

The meeting also heard from Daniel Dancey (software engineer and AI speaker, and Treasurer for Dorset Humanists) about well-documented 'AI fails' and unavoidable disruption in the labour market. He explained how the commercial advantages of early



release of any new generation of AI program are so extensive that big business does not have the rational incentive to iron out all the bugs.

Cyber-attacks, fake audio and image generation and the degradation of the entire information environment were identified as major risks, and an educator in the audience added her own concerns that AI will be encouraging a generation of people who only know (at best) how to use the tool (and at worst, to cheat) when learning, without absorbing the actual knowledge.

The audience also raised questions about the meaning of intelligence, whether doing anything more than simulating emotional understanding or a theory of mind would ever be possible for AI, and how it happens that current AI sometimes 'hallucinates' and gives an authoritative but wrong answer, instead of saying it just doesn't know.

The report prepared for the Bletchley Park summit emphasised that the overarching risk is a loss of trust in the technology itself, which would permanently deny humanity its transformative positive benefits.

The Chair concluded by endorsing the heartfelt call from Dr Fei-Fei Li, the Co-Director of Stanford University's AI Institute -

"So if we want it to play a positive role in tomorrow's world, it must be guided by human concerns."

From Farnham Humanists' perspective, those concerns need to be the best of humanity's values, rather than the base, exploitative or the harmful.

The voluntary collection taken at the end was for ComputerAid International which has provided access to technology to over 14.5 million people.

