

Artificial Intelligence and a Crossroads for Humanity

Professor Giorgio Buttazzo

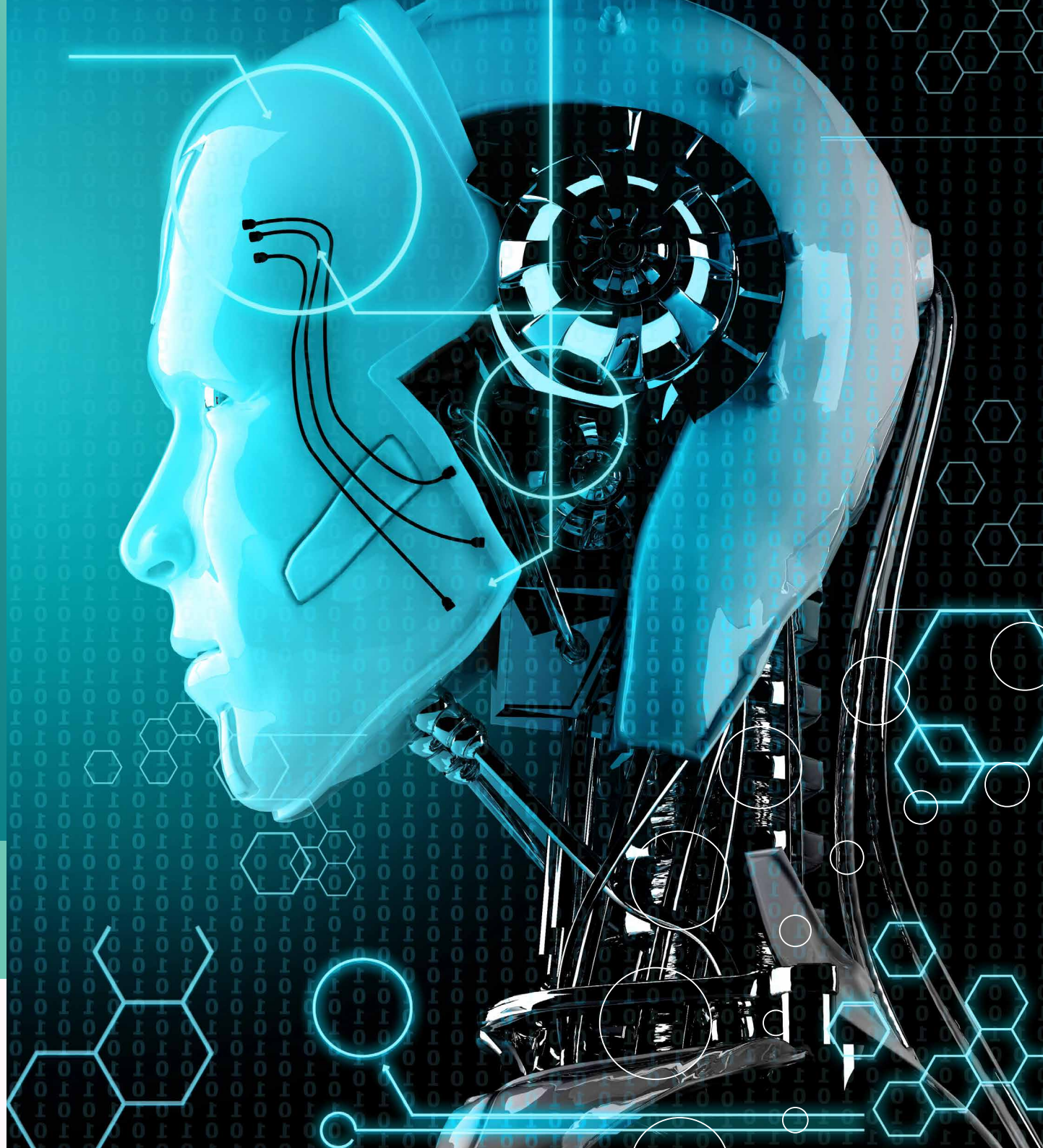
JULY 2024

doi.org/10.33548/SCIENTIA1047



ENGINEERING
& COMPUTER SCIENCE

 Scientia





Artificial Intelligence and a Crossroads for Humanity

Where do we stand with artificial intelligence? Might machines take over our jobs? Can machines become conscious? Might we be harmed by robots? What is the future of humanity? **Professor Giorgio Buttazzo** of Scuola Superiore Sant'Anna is an expert in artificial intelligence and neural networks. In a recent publication, he provides considered insights into some of the most pressing questions surrounding artificial intelligence and humanity.

A Brief History of Neural Networks and Deep Learning

In artificial intelligence (AI), computers can be taught to process data using neuron-like computing systems inspired by the mechanisms used by the human brain. These so-called neural networks represent a type of machine learning ('deep learning') in which interconnected nodes or neurons are able to adapt and learn from data to recognise patterns and solve complex problems.

After a somewhat lukewarm reception in the earlier part of the 1900s, interest in neural networks was reignited in the early 1980s when scientists developed new approaches that more closely resembled the functioning of the human brain and were capable of working as associative memories and self-organised cortical maps. The upshot of these advances was that neural networks could be effectively applied across a broad spectrum of disciplines, including physics, robotics, agriculture, and economics (to name but a few). But by the end of the century, interest had yet again stalled, and research advances weren't able to offer any meaningful theoretical or applied benefits.

Professor Giorgio Buttazzo of Scuola Superiore Sant'Anna is an expert in safety-critical software, neural networks, and the future application of AI, particularly the risks and benefits of superintelligent systems. He explains that in the 2000s, developments in both theory and technology converged to offer huge new potential for neural networks and deep learning. At the same time, computers advanced considerably, and cloud data storage and high-speed networks became more readily accessible, opening up this emerging and exciting field to increasing numbers of researchers.

Finally, according to Professor Buttazzo, the ImageNet Large-Scale Visual Recognition Challenge, a sort of annual computer vision Olympics organised to stimulate the development of innovative algorithms for solving complex problems in computer vision, also contributed to the explosion of deep learning. In 2012, for the first time, the winner for image classification was a neural network known as AlexNet, which considerably outperformed the established algorithmic solutions. By 2014, a neural network developed by Google – GoogLeNet – reached a human-level performance. So, from 2018, the organising committee ramped up the task to include more complex tasks. And year by year, the number of tasks in which machines outperform humans continues to increase.

Then and Now

In 2001, Professor Buttazzo published an article outlining his perspective on the likely, perhaps inevitable, development of consciousness in machines. Here, he argued that by 2029, conscious machines could be a reality. More than 20 years later, Professor Buttazzo revisited his predictions in light of recent developments in artificial intelligence, and took stock of the possible risks and opportunities of this for humankind.

In his 2023 publication, Professor Buttazzo observed that artificial intelligence surpassed human performance in several tasks. In industrial production, robots have been used in place of humans for many years due to their superior precision and speed. In medical science, neural networks have been shown to be more efficient and effective than humans in diagnosing heart disease, lung cancer, skin cancer, and even neurodegenerative disease. Neural networks are even providing stiff competition in creative



fields (such as classical music composition) and language comprehension (as evidenced by the many translation and transcription programmes now available and, of course, chatbots, which simulate human conversation).

Are Jobs at Risk?

Given the clear utility of neural networks in routine tasks, it is not surprising that concerns surround the replacement of humans in different occupations. For example, since June 2022, driverless taxis have carried passengers in California (negating the need for human taxi drivers), and transportation systems continue to undergo adaptations based on technological advances. The general consensus is that in the coming years, AI will undertake many jobs, including those related to legal practice, medicine, and even software programming.

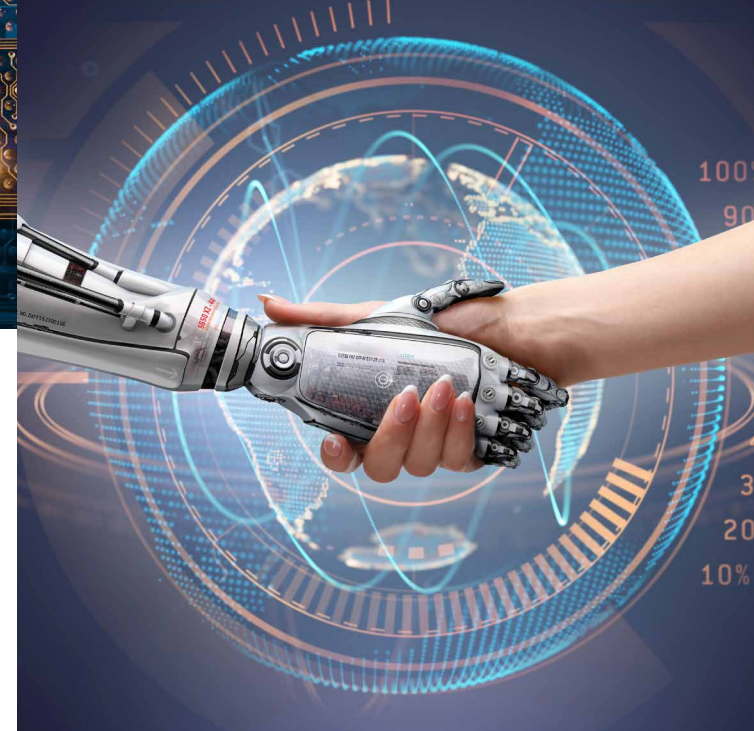
Professor Buttazzo takes a reasoned stance on this. He proposes that just as AI will replace human jobs, it will create new jobs, and if we look back through history, the nature of 'work' has always been ever-changing. Instead, he argues that it is not the replacement of jobs that is a concern but the speed at which this might occur, and he emphasises the importance of thinking ahead and developing appropriate strategies to manage the transition.

Might We Be Harmed By Robots?

Ethics are used to define and regulate human behaviour. But what about intelligent robots? In the 1940s, Isaac Asimov formulated three ethical laws for robots, including the First Law ('A robot may not injure a human being...'). Professor Buttazzo notes that the practical application of these laws would be difficult due to ambiguity in the language and complexity of the situations that may arise. But would algorithmic approaches be any more useful? Decisions are often not easily or readily made. A clear challenge for the future is to find a way to 'teach' robots common sense in the application of ethical rules and behaviour in different contexts.

A Crossroads: Future Outcomes for Humanity

Professor Buttazzo is clear that the trajectory of the increasing presence of intelligent machines is set to continue. Indeed, he notes that machines are likely to achieve human intelligence around 2030 and surpass human capabilities very shortly after. He warns that there are two opposing outcomes from this. The first is that the uncontrolled evolution of machines could possibly lead to a direct conflict with humans, resulting in a battle that humans would be unlikely to win – extinction. The second is that humans embrace the opportunities offered by AI to control their own evolution in order to adapt, change, and merge with these technologies – and, in doing so, achieve immortality.



Professor Giorgio Buttazzo of Scuola Superiore Sant'Anna is an expert in safety-critical software, neural networks, and the future application of AI, particularly the risks and benefits of superintelligent systems. He explains that in the 2000s, developments in both theory and technology converged to offer huge new potential for neural networks and deep learning.



MEET THE RESEARCHER



Professor Giorgio Buttazzo, Department of Excellence on Robotics and AI, Scuola Superiore Sant'Anna, Pisa, Italy

Professor Giorgio Buttazzo achieved his PhD in Computer Engineering at the Scuola Superiore Sant'Anna of Pisa in 1991. After a research period at the University of Pennsylvania, he took up the position of Assistant Professor at the Scuola Superiore Sant'Anna of Pisa, where he founded and directed the Real-Time Systems Laboratory (ReTiS) – a world-leading research group. From 1998 to 2005, he held the position of Associate Professor at the University of Pavia, where he directed the Robotics Laboratory of the Computer Science Department. In 2003, he co-founded Evidence s.r.l., a spin-off company of the Scuola Superiore Sant'Anna, providing software solutions for real-time embedded systems. He has been Editor-in-Chief of the Journal of Real-Time Systems and Associate Editor of IEEE Transactions on Industrial Informatics and ACM Transactions on Cyber-Physical Systems. He has authored 10 books on real-time systems and over 300 papers in the field of real-time systems, robotics, and neural networks, and has received 15 Best Paper Awards.



CONTACT

Giorgio.Buttazzo@santannapisa.it
<https://retis.santannapisa.it/~giorgio/>



FURTHER READING

G. Buttazzo, Rise of artificial general intelligence: risks and opportunities, *Frontiers in Artificial Intelligence*, 2023, 6, 1226990. DOI: <https://doi.org/10.3389/frai.2023.1226990>

G. Buttazzo, Artificial Consciousness: Hazardous Questions (and Answers), *Journal of Artificial Intelligence in Medicine*, 2008, 44(2), 139–146, 2008. DOI: <https://doi.org/10.1016/j.artmed.2008.07.004>

G. Buttazzo, Artificial Consciousness: Utopia or Real Possibility?, *Computer*, 2001, 34(7), 24–30. DOI: <https://doi.org/10.1109/2.933500>.



Find out more at [scientia.global](https://www.scientia.global)