



The Scottish Parliament  
Pàrlamaid na h-Alba



Scotland's Futures Forum  
Fòram Alba air Thoiseach



## Artificial Intelligence and Accountability

**While AI has been developing for decades, recent years have seen increasing attention paid to its various societal impacts. These impacts range from positive and helpful to harmful and even life-threatening in some cases.**

**Parliaments, and parliamentarians, have a key role in understanding the technology and its implications for citizens.**

**How is AI being used in Scotland? What does its use mean for the Parliament? And how can we hold public bodies accountable for decisions made by algorithm?**

**WEDNESDAY 22 JUNE 2022, at the Scottish Parliament and online**

## Introduction

As part of the Futures Forum's consideration of the future of democracy in Scotland, this seminar, considered the implications of technology for democracy, focusing on the increasing use of artificial intelligence in Scotland and the ethical and governance implications.

The seminar was held in partnership with the Parliament's Public Audit Committee and chaired by the committee convener, Richard Leonard.

## Panel



**Professor Subramanian (Ram) Ramamoorthy** holds a Personal Chair of Robot Learning and Autonomy in the School of Informatics at the University of Edinburgh, where he is also the Director of the Institute of Perception, Action, and Behaviour.

Professor Ramamoorthy's research focus is on robot learning and decision-making under uncertainty, with particular emphasis on achieving safe and robust autonomy in human-centred environments.



**Professor Shannon Vallor** is the Baillie Gifford Chair in the Ethics of Data and AI at the Edinburgh Futures Institute (EFI) at the University of Edinburgh, where she is also the director of the Centre for Technomoral Futures.

Professor Vallor's research explores how new technologies, especially AI, reshape human moral character, habits, and practices.

## Chair



**Richard Leonard MSP** has served as an MSP for the Central Region since he was first elected to the Scottish Parliament in May 2016.

Richard was leader of the Scottish Labour Party from 2017 to 2021. He is a former Organiser for the GMB trade union and a former Scottish TUC economist.

Richard is convener of the Parliament's Public Audit Committee, which the partner for this event.

## Other resources

- [Read the research on Parliamentary Responses to Artificial Intelligence](#)
- [Listen to the presentations as a Scottish Parliament podcast](#)

## Presentation: Perspectives on AI and its uses

### Professor Ram Ramamoorthy

#### INTRODUCTION

Professor Ramamoorthy began by describing what constitutes AI: a collection of computational methods and tools that make various processes automated.

Although AI stands for artificial intelligence, Ram also highlighted the word “autonomy” to describe what happens when a machine starts to take over some of the decision-making capabilities that previously rested with the human participant.

# Perspectives on AI and its Uses

**Scottish Futures Forum**  
22<sup>nd</sup> June 2022

**Subramanian Ramamoorthy**  
School of Informatics, The University of Edinburgh  
Edinburgh Centre for Robotics  
The Alan Turing Institute



[View Professor Ramamoorthy's slides on the Futures Forum website](#)

As Professor Ramamoorthy pointed out, there is a very wide spectrum of AI technology currently in use. One example is the use of computer vision methods in healthcare, where automated processing of images is used to work out people's state of health or mood and what care they should receive.

Other technology is used in data mining, which is the automated processing of large amounts of the data that are being constantly collected. An example could be the use of someone's bank account data or spending patterns to feed into an automated decision on a mortgage application.

Another exciting area of development is natural language processing. In this, speech can be analysed, generated and synthesised in new ways, which lends itself to business process automation, chatbots and so on.

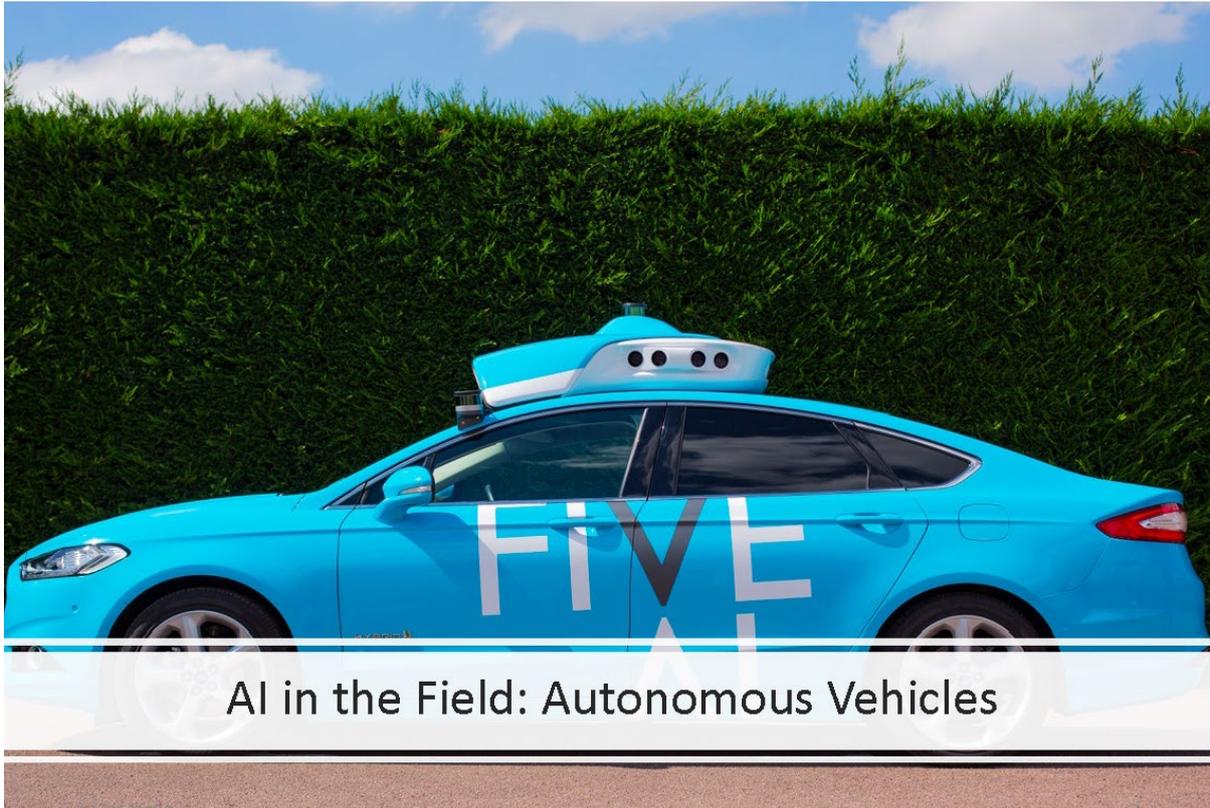
Professor Ramamoorthy emphasised the gradual nature of the trend towards the use of AI in various areas, starting with small-scale data summarisation and ending with decisions becoming fully automated.

Ram described an example from his own work at the autonomy end of the spectrum in applying machine learning to make physical devices, typically robots, more

capable. He highlighted his lab's involvement in a project with surgeons and medical professionals to use AI technology to make robots better at learning from experts. In this "learning from demonstration", an expert demonstrates a skill—such as performing an ultrasound scan or a surgical task—and the robot builds the task for itself and develops autonomy that leads to improved outcomes.

### AUTONOMOUS VEHICLES

As another practical example, Professor Ramamoorthy discussed the development of technology to deploy autonomous driving capability on the roads. Indeed, the case of autonomous cars in particular serves to raise key questions around governance, regulation and other challenges arising from the use of AI.



Professor Ramamoorthy highlighted several issues that have emerged as systems have become more automated and there are more of them. He explained how systems develop in complexity from automated parking, to lane control and, ultimately, full driving control.

He described how, historically, devices have been designed through a process in which an expert comes up with specifications that are discussed in great detail and then implemented. The certification of that is relatively straightforward.

However, the move towards data-driven technologies means that it is possible for a device to learn for itself from data that is not fully understood by the human experts. This results in the deployment of autonomous decision-making capability in some capacity.

Professor Ramamoorthy flagged up the fact that certification bodies are struggling with issues such as where in the process to intervene and how, and how guidelines should be set. Ram noted that, in essence, what begins as a technical problem becomes a broader governance challenge.

### CONTEXT

All of the technologies being discussed, Ram noted, are being developed in a fast-moving landscape. A striking example has been the adoption of telehealth technologies since the Covid pandemic.

Two years ago, people were aware of the technology but did not use it. When the pandemic hit, however, people were forced to use the technology, and this has changed the mindset. Significant numbers of the medical community, and patients, now use the technology successfully.

When technology is adopted in such a widespread way, Professor Ramamoorthy suggested that regulatory and certification bodies have to work hard to keep up.

### SOCIAL QUESTIONS

Professor Ramamoorthy also highlighted the social questions that arise from the emergence and evolution of new technologies.

For example, as part of a US infrastructure project, it was suggested that sensors on cyclists or pedestrians could be connected with cars to solve technical challenges relating to self-driving vehicles. However, as Ram pointed out, the involvement of other people changes the dynamic, as regulation no longer concerns only the device and its user but passers-by.

Ram noted that similar questions arise in healthcare. With the increasing use of AI in medicine and diagnostics, there are questions around the balance of decision making between the human expert and the automated machine.

As Professor Ramamoorthy outlined, there are many initiatives worldwide to look at these issues. There is a lot of interest from Governments and industry bodies in how testing, verification and certification can be improved.

In closing, Ram discussed his part, along with Professor Vallor, in the [Trustworthy Autonomous Systems Programme](#), which brings together a multidisciplinary group of experts to explore what the issues are, how software engineering and development paradigms can be changed and how to influence the broader conversation in society around ethics, accountability and so on.

## Presentation: AI and Accountability in the Public Sector

### Professor Shannon Vallor

Professor Vallor spoke about the vast scope of AI applications in the public sector and outlined the profound accountability issues and challenges which they raise. She pointed out that even systems that are not technically defined as AI can be deployed in ways that present substantial issues.

# AI and Accountability in the Public Sector

Scotland's Futures Forum, Scottish Parliament, 22 June 2022

**Shannon Vallor**

**Baillie Gifford Professor of Ethics of Data and AI**

The University of Edinburgh

Director, Centre for Technomoral Futures

Turing Fellow, The Alan Turing Institute



THE UNIVERSITY of EDINBURGH  
Edinburgh Futures Institute

The  
Alan Turing  
Institute

[View Professor Vallor's slides on the Futures Forum's website](#)

### AI IN THE PUBLIC SECTOR WORLDWIDE

To provide a context to her presentation, Professor Vallor outlined where AI applications are used throughout the world:

**Policing:** face recognition, crime prediction, criminal profiling, lie detection, licence plate reading, image analysis and reconstruction

**Judicial:** bail/pretrial release decisions, recidivism risk assessment

**Health Care:** diagnostic image reading, medication monitoring and delivery, robotic surgery, triaging, risk flagging, organ and bed allocation, personalised treatment selection

**Social Care:** fall/activity monitoring, location tracking, medication monitoring, biometric monitoring (breathing, sleep patterns, pulse), behavioural analysis and prediction, social care support matching

**Education:** student risk assessment, exam proctoring/cheating detection, classroom attention/gaze tracking, behavioural monitoring, student sentiment analysis, automated marking

**Immigration:** border security, detention surveillance, claim approval, risk assessment, identity verification, fraud detection, lie detection

**Public Welfare:** benefits fraud and abuse detection, automated application review, benefits determination and automatic adjustments

**National Security/Defence:** autonomous weapons and vehicles, cybercrime and cyberattack detection/prevention, encryption/decryption, surveillance, behavioural analytics and profiling, suspect or target identification, suspect or target tracking, risk and strategy assessment

**Transportation and Infrastructure:** smart city sensors, traffic management, autonomous vehicles, road/sewer/bridge defect detection, maintenance prioritisation, water and air quality monitoring, weather/flood prediction, emergency services dispatching

As Professor Vallor highlighted, these emerging use cases raise issues about the lack of the knowledge required for oversight.

For example, which of the above technologies are being deployed in the UK or Scotland, and which of these are even scientifically legitimate? She pointed out that there is not necessarily consensus among the AI community on the safe and legitimate use of certain tools, and yet they are being deployed by Governments in a way that impacts citizens.

One example is sentiment analysis in the classroom, which is being used in China and is available elsewhere. Shannon noted that Microsoft had recently announced that it would be retiring the emotion recognition capabilities in its facial recognition software. It did that because it did not think that the technology was scientifically or ethically robust enough to sell.

However, while some large tech companies might withdraw from sale technology that they consider is not scientifically or ethically robust, third-party software vendors may still make a pitch to Government agencies with these same technologies, which could affect vulnerable citizens.

### MANAGEMENT AND MITIGATION

Professor Vallor highlighted other questions to consider.

- **Which of Scotland's public agencies are adequately resourced to deploy these tools safely and reliably?**
- **Which technologies have worked well in other jurisdictions, and which have not?**
- **Which of them present currently unmanageable ethical risks, and what are those risks?**
- **Who is endangered by these risks, and what is their path of redress?**
- **What is required to mitigate the manageable risks?**

Shannon cited a key issue: public agencies are often not well funded or staffed enough to be able to answer those questions themselves. This presents a substantial risk to accountability and good governance.

As Professor Vallor noted, this in turn raises a big question:

**Who is responsible for ensuring that public sector use of AI in Scotland is scientifically legitimate, safely and reliably implemented, ethically deployed, and accountable to the public and those at risk?**

### ACCOUNTABILITY AND TRUST GAPS

Professor Vallor noted that the UK has in place robust legal frameworks for data protection, intellectual property and copyright, and information governance. However, she flagged up an issue: public agencies may conflate adherence to these standards with meeting ethical requirements. It is manifestly not.

As Shannon stated, ethics often enters gaps where legal accountability is, or is perceived as, porous, weak and inadequate. As technology moves fast, there are many such gaps, and there are ethical expectations arising from various sources, such as professional societies, the public and whistle-blowers.

Professor Vallor highlighted as a key concern the need to address those gaps and build public trust in the way that those technologies are used in the public sector.

Shannon moved on to link the question of trust with the requirement for accountability for power, in particular where that power may endanger specific vulnerabilities and interests. She distinguished between three types of accountability:

- **Retrospective accountability:** whether the person or agency deploying the power will answer for any unjust harm it causes.
- **Prospective accountability:** where the technology has not been deployed yet, who will be accountable when it is?
- **Character accountability:** whether the person or agency has thus far been trustworthy with the interests of people and the community.

She identified that, where there is a trust gap in technology, there are three ways to restore trust or bridge that gap. Those are:

- **Hard constraints**—local or global prohibition or restriction of the technology.
- **Robust duties**—creating duties of care assigned to specific parties that allow the technology to be deployed safely, reliably and accountably.
- **Strict liability for harm**—imposing specific and appropriate sanctions.

### ETHICAL RISK AND ACCOUNTABILITY IN AI

Professor Vallor cited the specific ethical risks and vulnerabilities that need to be addressed in respect of AI:

- Unpredictable and brittleness performance of AI systems.
- Unjust bias arising from historical data or inappropriate design decisions.
- Opacity of AI or machine learning decisions, known as the “black box problem”.
- Deployment at speed and scale, which can impede meaningful human control and lead to automation bias.
- Distinctive vulnerabilities of groups targeted for public sector use cases, whose autonomy, dignity, rights and wellbeing may be disregarded in order to attain key efficiencies or satisfy political aims.

Shannon also highlighted some barriers to accountability in AI:

- Lack of resource in public agencies to identify or manage risks appropriately.
- Optimism bias, with overreliance on legal compliance and risk left unaddressed.
- Techno-solutionist imperatives: not everything needs AI attached to it.
- Lack of technical skills to create appropriate, robust models and safeguards, which makes public agencies vulnerable to exploitation by unscrupulous vendors.
- Fears of over-regulation stifling innovation and adoption; whereas failure to appropriately regulate also stifles innovation in the long run because people will be reluctant to adopt risky and unaccountable AI tools.
- Inadequate channels for identifying, reporting and contesting harms.

### EXPERTISE AND RESPONSIBILITY

In concluding, Professor Vallor highlighted several opportunities regarding accountable AI in the public sector.

She noted that, in the UK, there are growing resources, such as those provided by the [Alan Turing Institute](#), to guide public sector agencies. In addition, there is an opportunity for investment in producing the knowledge to help them bridge the gaps, along with new training pipelines to develop AI ethics expertise that public agencies can use.

She highlighted Scotland's strong commitment to responsible AI and [strong AI strategy](#), and noted that Scotland's other advantage is public trust.

Finally, Professor Vallor noted that devolved agencies can create new cultures of accountability and care in deploying AI that provide a sound model for others to follow.

## Q&A session

### Struggling to keep pace

Participants began by addressing the key theme of the seminar: what can the public sector, the Scottish Parliament and its committees do to ensure accountability with regard to AI? It was noted that the public sector in general struggles to keep pace with the speed of change. The need for oversight of data collection by certain bodies, such as the police, was highlighted.

Professor Vallor stressed that the first step for Government agencies is to ensure there is a high-level overview of where AI is being deployed in Scotland's public sector, and to what end. She also noted that Governments need to learn from each other and share information, in particular where public sector deployments have not gone well.

Shannon flagged up an interesting dynamic that Governments can use as a lever for change. The big technology companies have signalled an openness to limited regulation, because they have the resources to invest in mitigating risks and avoiding harm and they do not want to be undercut by other companies. She pointed out that Governments can therefore work together with companies on things like managing bias to achieve mutually beneficial goals.

## **The balance of power**

Participants highlighted issues with the balance of power regarding AI, not only between Governments and the private sector but at a micro level, for example between doctors and patients.

It was noted that, while the seminar was largely about accountability within the public sector, there were concerns around accountability and governance around AI data and tech issues in the private sector.

Why should we have to rely on companies themselves withdrawing products they consider to have risks? Where can we take the conversation around understanding public control of private sector activity?

It was argued that there appear to be huge problems with the regulation of AI, and yet, collectively, around the world, we are not trying hard enough to address them.

The possibility of greater citizen involvement in testing AI was raised, with a need for a more democratic understanding of the possibilities regarding the technology.

Procurement was mentioned as a useful area of interaction between the public and private sectors; participants raised the question of how much discussion there is internationally regarding how technology can be ethically sourced. How can the Parliament's Public Audit Committee, for example, benchmark what oversight bodies are doing about the procurement of software by the police and other bodies?

Professor Vallor emphasised that setting standards for procurement is another lever that Government can use to raise the floor for such technologies across the board.

Professor Ramamoorthy discussed how we might take lessons from approaches to governance and accountability in fields such as engineering. He observed that a lot of AI came out of the development of the internet, in an environment where there was no real cost for errors.

However, companies are now looking at it differently, and building models from the ground up in the hope that some of the obvious harms and biases can be mitigated or controlled, or at least understood.

Ram emphasised that the internet approach is not the only way to think about commerce—highlighting, for example, how we think about safety and reliability in aviation. He stressed that if the whole regulatory infrastructure around AI moves towards that type of approach, the right questions will be raised, stemming from how technology is thought about in the context of the ecosystem, whether that is public or private.

## **Learning from our mistakes**

Picking up on one of the themes that Professor Vallor raised, participants considered the extent to which AI can embed structural bias in a way that could affect risk profiling, women in business and so on. It was noted that people around the world are grappling with that challenge, and participants considered how legislatures might address it.

Professor Vallor emphasised that bias within AI is not a symptom of defective technology; rather, it is inherent in the way that the technologies work. She pointed out that, as we train AI systems on historical data, we are asking them to learn from

our own mistakes, bias and social failures. She noted that the biggest tech companies spend a great deal of energy and expertise on managing the risk of bias, and she stressed that it is a risk that has to be managed rather than eliminated, as there is no way to build AI without worrying about that aspect.

Participants raised the question of whether AI itself could be used to identify and tackle systemic bias. Professor Vallor described that as an on-going challenge for Governments, and she stressed that there are ways to use these tools as mirrors to highlight the flaws in our own processes. She urged Governments, rather than running away from challenges such as bias, to run towards them and embrace them.

She noted that, in some cases, that may mean that an AI application might not be fit for purpose because there is no way to make the risk of bias manageable or acceptable, whereas in other cases, there are things we can do.

The key lesson Professor Vallor flagged up is that we should expect to find bias and plan for how to manage the risks. She emphasised that being well-informed can help Government agencies and others to make decisions about the risk-benefit balance with AI and how to manage it through auditing and benchmarking.

### **Transcending human capability?**

The recent case of a Google AI specialist who raised concerns that, in his view, one of the company's AI programs had achieved consciousness was discussed.

Beside the bigger question of whether that is possible and the broader ramifications it could have for society's use of AI, the associated question of whether whistle-blowers working with AI would have their concerns taken seriously was highlighted.

Professor Ramamoorthy argued that, while this concern has been a long-standing philosophical problem, large corporations may see such stories as a useful PR strategy.

Professor Vallor explained that the current technology is not the sort of thing that could achieve sentience. The tools predict strings of data from past patterns of data—in a sense, a sophisticated version of autocomplete—whereas humans can reason and comprehend.

She described the recent Google incident succinctly as an unfortunate case of someone falling victim to illusion, and stated that we do not have to worry about the current generation of technologies sneaking up on us and becoming conscious without our realising it.

Professor Ramamoorthy highlighted that, while we might see the potential for AI to transcend human capability as a danger, the real issues that we need to grapple with are much simpler.

As with all technologies, the important thing is to keep an eye on the misuse or ill-considered use of AI, rather than worrying about whether technology has transcended limits in a way that we might not expect.

### **Attendees**

The seminar was open to Members of the Scottish Parliament, their staff and Parliamentary officials, as well as others in the Futures Forum network. Among these, the following elected Members attended: Maggie Chapman MSP, Richard Leonard MSP, and Michelle Thompson MSP.