# Automation and Autonomous Systems Roundtable

Final Report

**Background**

UKRI commissioned the National Coordinating Centre for Public Engagement (NCCPE) to convene a Roundtable to bring together colleagues from across UKRI and the wider research and innovation community, with interest in the areas of Automation and Autonomous systems (AAS). The work was undertaken as part of the Sciencewise[1] public dialogue programme, which is delivered by UKRI, to serve the insight and evidence needs of research funders, and policymakers.

The purpose of the roundtable was to build a shared understanding of where public perspectives may be most critically needed in AAS and identify areas where further work is necessary with the possible support of the Sciencewise team. The roundtable also explored how UKRI might best approach future cross-cutting areas like automation, which have strong public interest and significant ethical sensitivity.

In conducting this work, we worked with the UKRI Public Engagement with Research Network to identify key contacts in the research councils and Innovate, with an overview and oversight of automation and autonomous systems investments. This was followed up with a series of interviews with policy leads and researchers in advance of the roundtable and further desk research. A list of interviewees is included as an annexe to this report.

This short report summarises the discussions at the Roundtable and provides recommendations for future dialogues and capacity building in this area.

This report also includes four appendices:

A: Roundtable proceedings – notes of the discussion
B: Background reading – the paper circulated to delegates before the workshop
C: Delegate list and agenda
D: List of the experts interviewed in preparing the briefing paper

[1] Sciencewise | Supporting the commissioning of deliberative dialogue by government bodies to support socially informed and transparent policy making

## Roundtable Summary

There were many possible areas to focus future dialogues around AAS. However, a few priority areas began to emerge from discussions. These were:

- The **ethical frameworks** that should underpin AAS research.
- **Trust and trustworthiness** and the underpinning social and cultural conventions that underpin trust between humans, machines and systems.
- **Verifiability and governance, particularly the regulatory frameworks governing** AAS innovations.

These suggested dialogues might be situated in specific areas where AAS is known to have a disruptive impact, such as manufacturing, the future of work, transport and aviation.

Another potential 'hot spot' identified was:

- **Health data, and privacy** in particular around new data-driven technologies in health care.

It is notable that some participants suggested that dialogue work in any area of AAS research would be useful at the moment.

It was recommended that decisions about future dialogue areas that targeted a specific technology or application area should be prioritised by assessing: (i) Technology maturity and proximity to market/applications, (ii) Potential impact and disruption, (iii) Level of public concerns.

UKRI has funded the TAS Hub to support dialogues with a diverse set of stakeholders, including government, industry, and the public to:

- Define the research challenges.
- Inform decisions and identify the risks
- Respond to the UK's economic, environmental, and social challenges

Participants at the roundtable, particularly those attending from the TAS Hub and Nodes, were receptive to support from Sciencewise to help scope these dialogues.

Participants offered critique on the dialogue method and highlighted its limitations, and suggested additional opportunities to help embed public engagement in AAS research. It was suggested that the following could be helpful:

- a shared resource for the Research community around public engagement with AAS resources and best practices
- a shared "pool" or panel of members of the general public interested in participating in future dialogue.

**Recommendations for future work**

The discussions at the roundtable were broad in terms of possible areas for dialogues and how these might inform research pathways. Topics discussed ranged from futures work, involving users in product development, information giving, specific questions related to an

industry, and broader questions related to social desirability and the type of world we want to live in. It was clear that more time was needed to tease out these dynamics in relation to ongoing and future planned investment in AAS. It was also noted that there were specific challenges that need to be addressed beyond user engagement to thinking about non-users and how they are excluded from or may be affected by AAS.

Two key focal points for future work have emerged for Sciencewise and UKRI following these conversations:

**1. Developing capacity for public engagement with AAS**

1. The development of a "public panel" that can inform future research priorities in this area. The panel could be supported over a more extended time period and engage participants in the design and delivery of AAS research. The panel should be co-developed with participants, and we should also track changes in attitudes, behaviours and beliefs of panel members over time.
2. A programme of support for senior researchers and research leaders of large scale, socially sensitive programmes in AAS. The aim would be to better equip them with the skills to work across disciplines and embed public and stakeholder engagement. This could be delivered via:
    - A programme of events
    - Development of a network
    - Development of resources and tools
    - A consultancy model linking up with social science/dialogue expertise etc.
3. Facilitate cross council networking on engagement and social intelligence needs to identify priority areas so as not to duplicate effort and join resources. We assume that this would encompass both insights into how publics understand emerging technologies (frames and meanings that attach to them) and also the concerns / ethical concerns etc. that they have

**2. Possible future dialogues (Focus Areas)**

No single priority area emerged from the Roundtable. However, several themes appeared which could inform future dialogue activity:

1. Public dialogue on the development of ethical principles that could inform AAS research. It is anticipated that these could then act as a reference point for researchers, policy and industry stakeholders.
2. Further dialogue work on notions of trust and what this means in the context of different AAS technologies (this may be addressed in work on the ethical principles).
3. Targeted work with specific individuals, groups or communities about future impact areas (i.e., work, transport, health and health care etc.) For instance, what are the specific requirements of niche road users, hazards & concerns? E.g. cyclists, horse-riders, disabled pedestrians, workers in a particular industry, animal herders.
4. Engagement work to deliver understanding of how to align different and sometimes conflicting interests between stakeholders (i.e., researchers, industry, policy experts and the public)

We recommend that Sciencewise and UKRI follow up with participants from the roundtable to help further scope a future dialogue in AAS and address (i) how the people involved in the dialogue could be engaged with AAS research over an extended period, and to (ii) consult people about how they might want to be engaged with this process.

The Roundtable notes are included in Appendix A.

# Appendix A: Roundtable Proceedings

**1. Specific challenges in integrating public engagement in Automation and Autonomous Systems research?**

Participants were invited to reflect on the role of public engagement (PE) in their work and the specific challenges they experience embedding public perspectives in developing Automation and Autonomous Systems research.

The complete participant list is included in Appendix B for reference.

Feedback from participants:

- UKRI has recently invested in this area via the TAS Hub. This multi-disciplinary hub brings together the £33M Trustworthy Autonomous Systems Programme, funded by the UKRI Strategic Priorities Fund. They currently support more than 25 different projects spanning different domains, including CAVs, health and wellbeing, disaster response, agriculture, robotics etc. Public engagement is a core part of delivering this kind of research. Co-creation is a core principle in the programme.
- We often talk in public participation about user involvement, users of technology as being core. There are a range of methods to support this; however, with AAS, we are not just looking at users (as highlighted in the discussion paper). We are looking at non-users (i.e. facial recognition in public spaces).
- We have seen how the lack of diversity of the teams designing these systems has had terrible consequences leading to data not being representative, embedding faulty or poor decision making within algorithms and processes that harm disadvantaged groups (for instance, Ofqual).
- Over the past two years, EPSRC has promoted good news stories from their research investments. For instance, the CASCADE programme (referenced in the discussion paper) utilised drones to deliver medical supplies. These examples just share information with the public, but they are an essential part of the landscape.
- Robotics is a very 'public friendly' topic. But researchers can experience public engagement as simply talking to the public, and most of the time, it is a one-way conversation. Occasionally we can get ideas from those conversations. The Royal Society Summer Science Exhibition was better and had successfully begun to open up a two-way dialogue with research over a week. We feel that the public does not generally have enough information about what an Autonomous System is to be able to tell us what the key issues are. Therefore we spend a lot of time explaining what AAS is and debunking myths, by which time the event has ended. We really need that co-creation element for Public Engagement to work well, but this takes commitment over time. It's clear that public engagement is important, and technology development should be led by public opinion. We're open to support for this from wherever we can get it.
- One of the challenges we all face is trust in AAS. However, there are two opposing trends here. The first is that AAS are becoming more complex in nature. The opposite pressure is that to gain trust; we need more technical transparency. With increasing complexity, it becomes more challenging to explain how these systems work and unpack aspects of safety.

- Public engagement is central in two very crucial ways. First, it is integral to our research methods that inform our policy positions. We use public engagement strategies like citizen juries / citizen councils to consult on particular issues. The second way it is central is that we integrate recommendations about public engagement in our policy recommendations (i.e. recommendations around participatory algorithmic assessments).
- We need to think more about long-term engagement with research participants. What are our commitments to the people who contribute to our work? We are trying to move beyond one-off short-term interventions towards keeping people involved over a long period of time.
- We need to dig more deeply and meaningfully into the perspectives of marginalised communities. These can be difficult to surface using deliberation methods because you do not have the depth and context in which you can understand the root of people's perspectives (i.e. the social and cultural context).
- The way in whch we tend to think about dialogue is embedded within a systemic social science approach. We are keen to move beyond notions of trust and public readiness towards understanding the social need and social desirability of emerging technologies. This cannot be done without looking at the richer deeper issues such as social inequity and workplace displacement. This is not something you can do with public dialogue alone – through public dialogue forms an important part of the process and helps us understand the framing of the research.

**2. Can we identify 'hot spots' which should be prioritised for more in-depth dialogue/engagement with the public?**

Our mapping prior to the roundtable made it clear that there are several areas where there is scope for future dialogue work. We put forward four possible 'hot spots' at the roundtable to stimulate discussion. These were:

- Future of work
- Health and health care
- Future flight
- Verifiability and governance

Participants were invited to reflect on their priorities and pressing decisions where public engagement with AAS would be valuable (for instance, developing a new funding programme or developing a public engagement strategy for a research programme).

***Feedback from participants:***

*Group 1: Paul Manners, Ram Ramamoorthy, Radu Călinescu, Mike Collins, Fern Elsdon-Baker Triantafyllos Gkikopoulos, Philippa Lang*

- There is an important democratic question about how much power and autonomy should be given to autonomous systems. This decision should be led by legislators and the public, not by technology.

- We're keen to move beyond questions of public trust and readiness to a process of understanding social needs. Technologies need to be framed by non-vested interests.
- The TAS Node for Verifiability is developing a process of engaging the public on the legal and ethical aspects to validate what the technology does. The validation aspect needs public engagement for developing specific requirements or principles to inform policies and regulations.
- Many organisations worldwide are working on ethics and high-level principles in this area, but there is a need to avoid duplication.
- Humanities and social science research can help us avoid the deficit model of engagement that assumes public dialogue is a form of representation and treats it as a consultation. Engagement and dialogue should be part of a broader programme of work because conversations are very complex, and those social norms are themselves subjective. We need to understand their context.
- It was suggested that part of the solution is to establish panels of lay expertise and take a more longitudinal approach to engagement, but it must be recognised that this is not 'The Answer'.
- We need to be careful about defining diversity when considering public engagement and ensure that we think beyond characteristics towards lived experience.

*Group 2: Ed Manley, Mohammad Reza Mousavi, Yan Yip, Deborah Kroll, Richard Morris, Suzannah Lansdell, Clio Heslop*

- Currently working in MRC on digital health and AI. Public engagement needs to focus on privacy and encourage people to allow the use of their data.
- Engagement is being developed by Ada Lovelace Institute for the AHRC. The current focus is – a) increasing public understanding, trust, acceptance b) consumer acceptance and adoption. AHRC also has a small portfolio on space technology and earth observation, which has a lot of potential for public engagement.
- Innovate the UK has funded around 94 projects in this area, with more recent ones focusing on demonstrations of the technology and hence are more public-facing. We have been thinking about engaging people in the future; our big challenge is normalising the technology.
- We're interested in understanding the demands and barriers for automated vehicles, especially among older people in the UK and Canada. Thinking about where we will see changes in travel demand and secondary impacts on the public, in particular different parts of the population. We're also interested in health data behaviour and privacy.
- Participants who were working more closely with business and business-led innovation urged for future dialogues that moved beyond 'the general public's' attitude to AAS. We need to hear from "niche groups" –the specific requirements of niche road users, hazards & concerns: E.g. cyclists, horse-riders, disabled pedestrians, animal herders. We have a reasonably good idea of general public attitudes, behaviours and what they want.
- We're interested in the balance people want between safety, usefulness and comfort for autonomous systems. How much risk are people willing to take in order to get the usefulness? Also, what ethical frameworks do people want to see implemented? We hope that the verification frameworks should be informed by

what the public. Fairness, lack of bias, transparency are generic and transferrable across different domains. But there are also very specific frameworks needed, e.g. medical systems have very different ethical requirements than a passenger car.

- On AI technology, we're interested in the contrast between different cultural groups and communities. What good looks like to one group won't carry across; what is trustworthy for one country might look different in another. E.g. there are very different cultures around driving in different countries! There are almost endless layers of perspective that would add nuance.
- From a health point of view, this conversation shouldn't just be about including the public, but the clinicians and whether they have the trust and skills to use the technology. Otherwise, it won't get adopted into practice.
- Finding a way to get through to the people who reject technology and don't want to be part of the discussion is critical.
- There are conflicting interests between stakeholders that haven't been resolved – push from technology manufacturers, whilst not ready to offer technical transparency that the public want. We need a diverse representation of stakeholders and a democratic dialogue approach to bring people together. This has been missing to date.
- Clear communication is still essential. Many people have concerns about new technology, but when it's explained, then they're more accepting.
- It was helpful in my research to facilitate a broad conversation about the implications of the technology. Many areas came up that took us by surprise. For example, we thought the main concern would be safety, but actually, it was job losses. We can't make assumptions about what people will be concerned about and need to keep the conversation open to allow those views to be surfaced.

*Group 3: Joel Fischer, Kira Allmann, Anne Toft, Andrew Tyrer, Diane Beddoes, David Owen*

- We are not just looking at engagement between researchers and users but also non-users. We need to facilitate broad conversations about the implications of AAS and put co-creation at the heart of what we do. We shouldn't take it for granted that this is generally the case across the landscape.
- At Innovate UK, we've done work around the public acceptance of drones with NESTA. This space has limited technical blockers, but public approval and regulation remain the two key blockers.
- We are interested in how do we bring in those excluded and marginalised. What methods can help us engage with people who are not involved: Participatory Design, PPI, Citizen Juries. It would be helpful to have case studies and examples of practice with AAS.
- It's easy to do dialogue as a one-off activity, but everything is dropped once the research is over. We need to 'sure up' a more long term engagement with the public and build capacity for engagement with AAS.
- There are a whole host of topics where dialogue work would be helpful. Research tends to have an outlook of 5 to 10 years. What kind of future do we want to look at and design for. CAVs, the relationship between AAS and net-zero, how AAS can be part of the solution. These are all areas we would be interested in.
- In robotics, there are many arguments that still need to be won. There is a fear about robots replacing jobs, yet there is also a dichotomy here about people not

being available to do those jobs or longer-term arguments about AAS creating higher paid positions within the economy.
- Engagement needs to involve policymakers and broader departments. CAV (Gov) what does this mean for the policy and legislation phase.
- At the Ada Lovelace Foundation, we are thinking more about the systems and processes that sit behind the technology. Automated decision-making systems in Health and a wide range of other domains – it's exciting to hear people's perspectives in these areas.
- There is always more work needed on unpacking differences between Acceptance, Trust and Transparency.

**3. Stepping back and looking forward: the role of Sciencewise supporting UKRI**

In the final part of the Roundtable, we explored how Sciencewise might contribute to UKRI's future ambitions to increase the engagement and involvement of the public with its programmes.

Participants were invited to reflect on the host of developments across UKRI, which foreground the need to embed sensitivity to public concerns and attitudes within UKRI's systems and in UKRI's expectations of the researchers it funds.

These include:

- Heightened expectations around EDI for all programmes
- An increasing focus on 'responsible' research, in terms of ethics, trust, integrity etc
- Individual councils developing their responsible innovation frameworks (e.g. EPSRC's Responsible Research and Innovation; BBSRC's Wider Perspectives work)
- The contribution of arts, humanities and social sciences research to an increasingly interdisciplinary approach to the social contexts for science

***Feedback from participants:***

- We need to generate a more sophisticated understanding of the different types of stakeholders involved. This will help us have more precise conversations about future dialogues. For example, we mentioned that there are non-users; what are the different types of non-user that we would like to engage? How do we conceive of these stakeholders in ways that are not extractive and simplistic?
- Participants would find it incredibly helpful to have resources and support for public engagement with AAS. There was an authentic call for help integrating the perspectives of different stakeholders, users, non-users etc, in future research priorities. The TAS Hub asked for a 'pooled resource' that they could share with the research community that shares best practices and support.
- Re-emphasising the discussions in the breakout group. We could recruit a panel of public representatives that want to engage with AAS research. It is incredibly challenging for researchers to build these panels from scratch, particularly reaching communities that go beyond our networks. There is potentially a lot of time being invested in this across the research system and much duplication of effort.

- Is there an opportunity to do a public dialogue about research engagement and pathways to engage with research before we begin implementing those?
- As part of the Future Flights initiative, we have been thinking about ways to pool social science expertise and offer this out in a consultancy model for AAS research. We're interested in embedding social science research within the processes, building networks of researchers across the country who can work across disciplines.

## Appendix B: Discussion paper

The following report was provided as pre-reading material for participants at the roundtable.

**1. What is Sciencewise?**

The Sciencewise programme is delivered by UKRI. It is managed by a consortium led by the participation experts, Involve together with NCCPE and the BSA. The Sciencewise team provide a portfolio of services to UKRI and to policymakers across government to support bringing public perspectives into policy and research.

The team has identified the area of Automaton and Autonomous systems as an area where Sciencewise has already undertaken some relevant dialogue but where there may be scope for additional activity. This roundtable is intended to identify opportunities Sciencewise could further support in collaboration with UKRI colleagues.

The table below describes the services currently offered by the Sciencewise programme:

**Delivering effective deliberative dialogue**

- Provision of expertise to support the: development of public dialogue framing and questions, commissioning, and evaluating
- Provision of expertise to support the integration of the deliberation with other engagement processes
- Up to 50% of the funding for deliberative dialogue

**Delivering faster engagement where policy time horizons are short**

- Provision of expertise to support the identification of effective engagement methodologies
- Connection to UKRI's suite of evidence and insight gathering tools

**Supporting identification of opportunities for early engagement to avoid conflict**

- Delivery of anticipatory roundtables focused on policy or technology areas in order to identify medium-term policy requirements and technological advances which will have a significant societal impact.

**Tapping into existing knowledge about public perspectives on scientific and technological innovation**

- A bank of social intelligence reports on key Sciencewise themes which can be quickly updated

**Providing access to wider innovation and experimentation in public participation**

- Through strong networks and access to UKRI's dialogue experimentation fund

## 2. Why Automation and Autonomous Systems?

AAS was selected by the UKRI Public Engagement and Sciencewise teams as it is a major priority for investment with the promise of rapid application but risks of low public acceptability. It is an area of research that may be subject to substantial public interest, and it has the potential to bring significant potential public benefits and harms alongside raising substantial risks and ethical concerns[2]. It is also a topic with great potential for cross-council interdisciplinary working.

## 3. Mapping the Research and Innovation Landscape

Automation and Autonomous Systems research does not sit neatly into a defined box. It is a multi-disciplinary area of emerging science with tributaries from several domains, including but not limited to (i) Robotics, (ii) Digital technologies and (iii) Human–technology interaction.

There are a range of investments across UKRI Themes/Research Areas which include work on autonomation and autonomous systems. Perhaps the most significant programme UKRI has invested in is the Trusted Autonomous Systems Hub, which represents a significant cross-council investment. The Hub sits at the centre of the £33M Trustworthy Autonomous Systems Programme, funded by the UKRI Strategic Priorities Fund. The programme aims to build capability in the area of autonomous systems, bringing together a number of diverse communities working to understand and enable trustworthiness.

Alongside the Trustworthy Autonomous Systems Programme automation and autonomous systems is embedded as an objective or work package within several small, medium and large grants, network funding and centres for doctoral training.

We have included a separate paper with a summary of these investments, for information.

## 4. What public engagement has already been undertaken relevant to AAS?

The application of automation and autonomous systems raises very significant social and ethical considerations, for instance: in the labour market (with impacts on employment), in healthcare (for instance in using robots for personal care), in agriculture, defence and surveillance, and in transport (where there is already significant debate about driverless cars).

We found that dialogues and social intelligence had been undertaken on autonomous vehicles, weapons, robots and automation in the workplace. Although the overall numbers are small, there appears to have been more dialogue on Autonomous Vehicles than in any other area.

[2] See for example: RAENG (2009) Autonomous Systems: Social, Legal and Ethical Issues, Royal Society (2017) Machine Learning: The power and promise of computers that learn by example, BEIS (2021) The Economic Impact of Robotics and Autonomous Systems Across UK Sectors

Experience from previous dialogues suggests that if the purpose of the dialogue is around engaging people with areas of governance and leadership over future research, then discussing Automation and Autonomous Systems too broadly can result in very diffuse and unfocused conversation[3]. Future dialogue work could sensibly be situated within specifics, for example specific industries or trades, potential technologies and specific policy options.

Interviewees noted how challenging it can be to balance breadth (opening up big questions like "what the world should look like" can be productive) but drilling down into specific sectors and application spaces is important to get the nub and nuance of public concerns and interests.

A summary of the existing social intelligence and dialogue work his included in a separate paper, for reference.

**5. What are the gaps and opportunities for future public engagement in this area?**

It is clear from our mapping that there are several areas, where there is scope for future dialogue work. A number of possible areas for future dialogue have emerged from the interviews and desk research to date. We've highlighted four below. We expect others to emerge from discussions at the roundtable and offer these as 'starters for ten'.

Each of the suggested dialogues could help:

- Investigate the existing social narratives surrounding AAS (i.e., what assumptions and stories do people carry about AAS?)
- Look at the perceived social harms/goods risks and concerns across users and non-user groups or communities, including relationships to broader social concerns (e.g., future of employment and broader economic concerns, climate crisis, conservation, the impact of COVID-19 etc.)?
- Help build in public perspectives to future funding calls and programmes.

[3] See for example: https://www.involve.org.uk/our-work/our-projects/practice/what-does-meaningful-public-engagement-look-ai-and-ethics

**Four possible focal points for future dialogue**

1. Future of Work

The recent Government white paper on levelling up places automation in the same sentence as climate change as a potential risk to the Government's attempt to tackle the growing regional inequality in the country:

> "the UK's economic geography starts from a particularly vulnerable position. Existing stocks of capital are depleted in large parts of the UK, as earlier sections described. And the future shocks hitting some of these places, such as automation and climate change, are potentially large and long-lasting. This generates a lack of spatial resilience across the UK and calls for a public policy response".

It is well known that the transition to increased automation, if left unmanaged, could negatively affect certain sectors and places. This brings into question whether a place-based dialogue on the future of work in specific industry areas may be of value. There have been comparatively few dialogues exploring the future of work, particularly in the context of place and/or a specific industry.

Key notes:

- Dialogue and social intelligence work in this space is limited.
- It has been identified by the UK government as an area of significant harm.
- We do not yet know the potential for cross-council collaboration around this, or the potential to inform future research agendas.

2. Autonomous Systems and Automation in Health Care

AAS in the health field is transforming how health and social care is delivered. For example, there is a growing focus on robots in the operating room, in clinical settings to support health workers and to provide care at home, enabling vulnerable patients to stay out of the hospital, but raising concerns about 'depersonalising' care. During the COVID-19 pandemic, hospitals and clinics began deploying robots for a much wider range of tasks to help reduce exposure to pathogens. It's become clear that the operational efficiencies and risk reduction provided by health robotics offer value in many areas.

Furthermore, recent work by the Ada Lovelace Foundation is highlighting how data-driven technologies and the systems within which they operate have are becoming a central part of the health infrastructure – a trend accelerated by the pandemic. Tools, such as symptom tracking and digital contact tracing apps, are being mobilised at pace and their use during the pandemic may well become the norm for the future.

Responding to what is increasingly being taken to be the near future of health care, there may be benefits to public dialogue in this area.

Key notes:

- Dialogue and social intelligence work in this space are limited. Though the focus needs to be further refined (i.e. Robots,data-driven technologies, which aspects of health and healthcare etc.)
- We do not yet know the potential for cross-council collaboration around this, or the potential to inform future research agendas.

3. Future Flight

The Future Flight Challenge has been allocated £125M under the Industrial Strategy Challenge Fund (ISCF) and aims to stimulate the development and application of new aviation technologies in the UK. ESRC has invested in a Future Flights Research Director and there is currently a working group considering the key research questions that could be addressed by the social science community. There is currently a small mini dialogue underway exploring Future Flight which is being supported and co-funded by Sciencewise. The expectation is that there will be future potential dialogue in this area.

Currently, the benefits and risks in this area are generally described in the abstract or at a fairly high level - for example, NESTA (2018)[2] and PWC (2018)[3] including potential use cases[4].

Key notes:

- Dialogue and social intelligence work in this space is limited.
- It has the potential to inform policy via the ISCF and future research activity helping to point out issues that could arise in the implementation and allowing in some cases for mitigation during the design process.
- It has good cross-council potential.

4. Verifiability and governance

Currently, autonomous system developers are not required to submit evidence to prove that the components of a design will result in entirely trustworthy functioning. Furthermore, governance and regulation of automated systems is still based on the idea that a person is always in charge of the system, so can step in if need be. For example, most planes have an autopilot feature that facilitates limited autonomy in flight, however much of the regulatory framework depends on the pilot being there to anticipate all eventualities and take control at any time. What criteria and values should underpin a future kitemark or regulatory frameworks for autonomous systems?

As Autonomous Systems become more capable and incorporate increasingly more automated decision-making mechanisms, typically AI-based, the existing frameworks for certification and assurance become strained: issues of liability, accountability and responsibility become harder to pin down. Currently the stakeholders involved in addressing these issues can be broadly defined as experts in various domains (i.e. legal,

policy, industry, technical and research), there has been little or no attempt to engage the public in this level of detail and to involve them in debates.

Key notes:

- Dialogue and social intelligence work in this space is limited.
- Verifiability and governance are two key nodes within the UKRI TAS programme, suggesting high likelihood of cross-council impact.
- The dialogue would need to be focused on specific application and technical areas; or it could take a range of applications and draw out overarching recommendations.

**6. Stepping back and looking forward: the role of Sciencewise supporting UKRI**

This review has seen us trying to 'retrospectively' identify possible starting points for public engagement across the topic of Automation and Autonomous systems. It has revealed a patchwork of activity to capture social intelligence and potential gaps for future cross-council commissioning. Our interviews have indicated an appetite for more proactive, cross-council collaboration for future programmes.

This reflects UKRI's new strategy, which commits the organisation to increasing public involvement in setting future priorities and increasing the inclusivity of the research and innovation it commissions. The new Public Engagement Strategy has a key goal of 'listening and acting on diverse voices on key research and innovation priorities'.

There are a host of developments across UKRI which are foregrounding the need to embed sensitivity to public concerns and attitudes within UKRI's systems and in UKRI's expectations of the researchers it funds. These include:

- Heightened expectations around EDI for all programmes
- An increasing focus on 'responsible' research, in terms of ethics, trust, integrity etc
- Individual councils developing their own responsible innovation frameworks (e.g. EPSRC's Responsible Research and Innovation; BBSRC's Wider Perspectives work)
- The contribution of arts, humanities and social sciences research to an increasingly interdisciplinary approach to the social contexts for science

We are interested in participants' views of how Sciencewise can complement and add value to these developments, and how it might best contribute to the development and scoping of new cross council programmes.

# Appendix C: Participant List and Agenda

## Roundtable: Public Engagement with Automation and Autonomous systems

**March 2nd, 11.00 – 1.00**

## Participants

UKRI

- Deborah Kroll, Senior Investment Manager for AI & Digital, AHRC
- Richard Morris, Innovation Lead: Autonomous & Connected Vehicles, Innovate UK
- Andrew Tyrer, Challenge Director: ISCF Robots for a safer world, Innovate UK
- Triantafyllos (Trias) Gkikopoulos, Innovation Lead - Robotics & Artificial Intelligence, Innovate UK
- Philippa Lang, Programme Manager PER, UKRI
- Anne Toft, Senior Portfolio Manager, EPSRC
- Yan Yip, Programme Manager, MRC

Academic / Research Leads

- Fern Elsdon-Baker, Professor of Science, Knowledge and Belief in Society, University of Birmingham
- Ed Manley, Professor of Urban Analytics, University of Leeds
- Professor Ram Ramamoorthy, Chair of Robot Learning and Autonomy, University of Edinburgh
- Mohammad Reza Mousavi, Professor of Software Engineering, KCL
- Kira Allmann, Researcher, Ada Lovelace Institute
- Joel Fischer, Research and Engagement Director for TAS Hub, University of Nottingham
- Radu Calinescu, Department of Computer Science, University of York

Sciencewise

- Suzannah Lansdell, Project Identification Lead and DES, Sciencewise
- Diane Beddoes, Senior Dialogue and Engagement Specialist and Lead Evaluator, Sciencewise
- Paul Manners, Co-Director, NCCPE
- David Owen, Associate, NCCPE
- Clio Heslop, Head of Policy, Partnerships and Impact, British Science Association

## Apologies

Neeraj Suri, Distinguished Professor & Chair in Cyber Security, Lancaster University
Joseph Ellery, Senior Research Portfolio Manager: Technology, ESRC
Sarah Church, Head of Technology, Work, Education and Skills, ESRC
Paul Colville-Nash, Programme Manager, MRC
Kate Delvin, Co-chair and Director of Engagement for TAS Hub, KCL
Elisabetta Cherchi, Professor of Transport, Newcastle University
Samantha McGregor, Head of AI & Design, AHRC
Bruce Etherington, Strategic Lead: Industrial Challenge Funds, ESRC
Sophia Abbasi, Head of Policy; Strategic Planning, Evidence & Engagement, BBSRC
Richard Kino, Policy Manager, BBSRC
Sara El-Hanfy, Head of AI and Machine Learning, Innovate UK
Graham Bukowski, Public Engagement Lead, UKRI
Simon Burall, Programme Director and Involve Lead, Sciencewise
Mike Collins, Head of Public and Stakeholder Engagement, AHRC
Victoria Mico Egea, Senior Portfolio Manager, Artificial Intelligence & Robotics, EPSRC

## Agenda

| | |
|---|---|
| **11.00** | **Welcome** |
| **11.05** | **Framing the roundtable**<br>• UKRI's ambitions for public engagement<br>• The role of Sciencewise |
| **11.15** | **Public engagement with Autonomous systems**<br>• The role of public engagement in Autonomous Systems<br>• Opportunities and priority areas for future public engagement<br>• Next steps |
| **12.40** | **How do we embed social intelligence and public engagement into the research ecosystem?** |
| **1.00** | **Close** |

## Appendix D: Interview List

Mike Collins, Head of Public and Stakeholder Engagement, AHRC
Victoria Mico Egea, Senior Portfolio Manager, Artificial Intelligence & Robotics, EPSRC
Yan Yip, Programme Manager, MRC
Kira Allmann, Researcher, Ada Lovelace Institute
Clio Heslop, Head of Policy, Partnerships and Impact, British Science Association
Sarah Church, Head of Technology, Work, Education and Skills, ESRC
Paul Colville-Nash, Programme Manager, MRC
Kate Delvin, Co-chair and Director of Engagement for TAS Hub, KCL
Elisabetta Cherchi, Professor of Transport, Newcastle University
Sophia Abbasi, Head of Policy; Strategic Planning, Evidence & Engagement, BBSRC
Bruce Etherington, Strategic Lead: Industrial Challenge Funds, ESRC
Samantha McGregor (she/her), Head of AI & Design, AHRC
Andrew Stafford, Head of the Society, Governance and Security Portfolio, ESRC