Public engagement with algorithms in public services
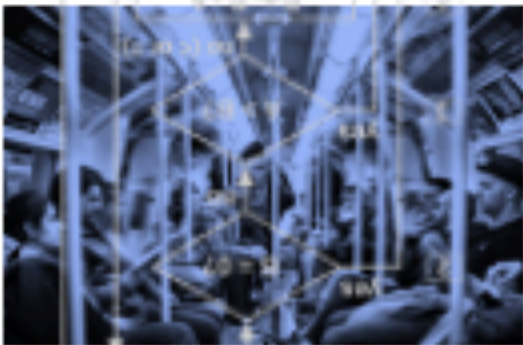
Briefing note

By Helen Pallett, Simon Burall, Jason Chilvers & Catherine Price.

University of East Anglia

3S Science, Society & Sustainability

involve

NOT-EQUAL

Just public algorithms

Algorithms are becoming more important in the provision of public services. At the same time, public debate about and resistance to their use is growing. It is in this context that this briefing note reports on a mapping of public engagement with algorithms in public services in the UK, 2013-2020, analysing data from 77 cases of public engagement.

Background and context

The events of 2020 have shown the importance, and challenge, of algorithms and associated data infrastructures in the UK. These events ranged from discussions about the data security of the NHS COVID-19 contact tracing app, to controversy about the use of an algorithm to determine A-level scores, and concerns about the unequal impacts of (app-mediated) police COVID powers on different communities[1]. Public reactions to these events demonstrate the challenges these technologies pose to public trust and social justice.

Over the last decade algorithms have become increasingly important across all areas of public life. Despite the potential for algorithms to carry out particular tasks on behalf of human actors and to improve the performance of others, scholars and activists have drawn attention to a number of social justice issues raised. These include the potential to exacerbate discrimination and inequalities[2], issues of consent and privacy, and an apparent lack of transparency or accountability[3]. Some of these issues are particularly acute in the context of public services as people may not be able to opt out, and because of the vast opportunities for data sharing across different areas of government and with the private sector.

Thus, the growth of algorithmic approaches in public services has been accompanied by recognition of the need for better regulation and governance. In particular, there are increasing calls for more public engagement in order to establish public views, build trust, and improve citizens' digital literacy[4].

There has actually been a considerable amount of public engagement with the use of algorithms in public services already. This covers a range of different technical applications and public service areas, and shows that citizens are already putting forward a multitude of relevant perspectives on the potential benefits and risks. However, existing forms of engagement are discrete, one-off and there is a dominance of institution-led or 'invited' forms of engagement. This project is one of the first to go beyond this dominant approach to map the diversity of public engagements on this topic.

The 'Just Public Algorithms' project was funded by the EPSRC through the 'Not Equal' network, and involved a mapping of cases of public engagement in this area and a workshop in February 2020 with key stakeholders

Diverse engagements and perspectives

The mapping shows the existence of many different forms of public engagement on the use of algorithms in public services through which citizens identify a wide range of potential benefits and risks.

Figure 1 summarises the sheer diversity of forms of participation the study found around the issue of the use of algorithms in public services. No single best practice model emerges from the mapping. Instead there are many different types of engagement. Often the individual cases we mapped involved more than one form of participation. There are already citizens engaged around this emerging area of technology and service provision who need to be listened to and understood. This shows that citizens do have the capacities to engage with these complex and ambiguous new technologies. They raise meaningful concerns and hopes about what they mean for the provision of public services.

The most frequently occurring forms of participation found in this mapping largely appear in the top left-hand corner of figure 1 as they tend to be issue-focussed and institution-led. Cases which were particularly significant in shaping the national conversation about algorithms in public services, and the appropriate ways of engaging citizens around this topic were generally also in these categories. Deliberative forms of engagement, especially public dialogues (n=5) and citizens juries (n=5), are the most frequently occurring among the cases mapped, with surveys almost as widely used. Both tend to be orchestrated by formal institutions – particularly government bodies or research bodies like the Nuffield Foundation or Wellcome Trust – and focus on uncovering citizen perspectives or discussing issues.

Public awareness campaigns, media campaigns and communication approaches are also significant approaches which tend to be institution-led (though some had significant involvement from civil society organisations) and focus on issues as opposed to material actions.

Whilst the majority of cases of engagement were institution-led and issue focussed, there were
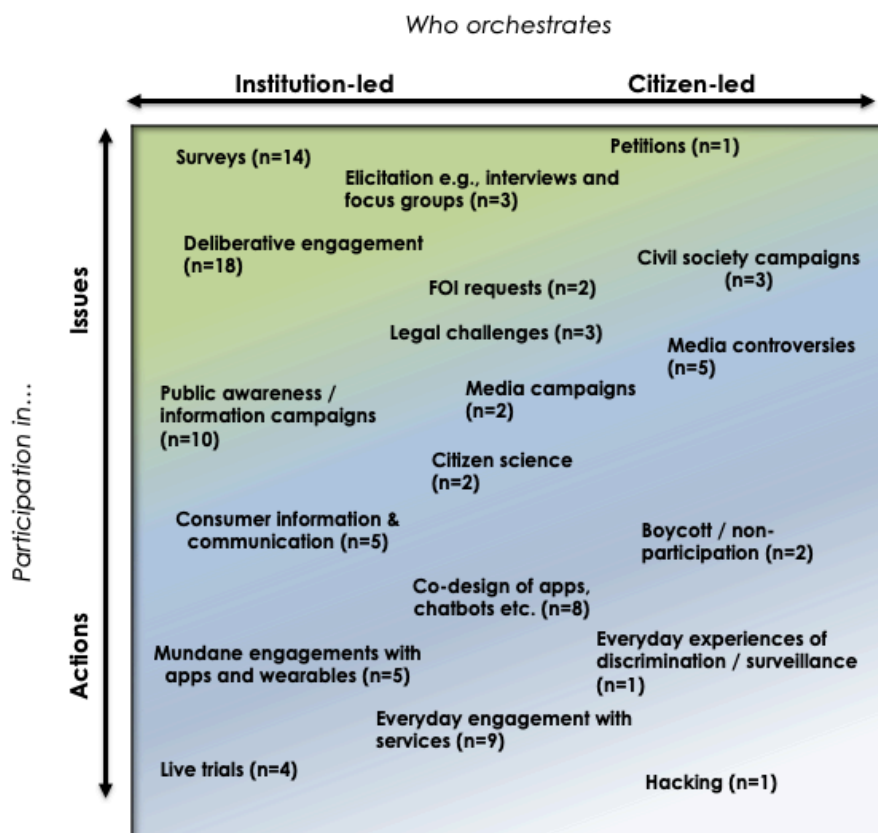

Figure 1 Models of engagement visualised according to who is orchestrating and what people are participating in

many examples found which took different forms and were more action-focussed and / or citizen-led. Emergent examples of more citizen-led forms of engagement around this topic, include civil society campaigns and petitions.

Some of the more action-oriented forms of engagement were quite mundane, including: engagement with chatbots and apps designed for service users; and unwitting participation in trials of technologies like facial recognition technology in policing. More emergent citizen-led and action-oriented forms of engagement included: the co-design of new apps and chatbots with users; deliberate non-participation in apparently compulsory algorithmic systems such as databases for aggregating education records; and the use of face paints by activists to 'hack' facial recognition systems.

Risks and benefits of algorithms in public services

Benefits	Risks
Better services	Lack of privacy
Greater efficiency	Discrimination
Better allocation of resources	Mistakes and inaccuracies
	Data sharing

Figure 2 Summary of main potential benefits and risks of the use of algorithms in public services identified in cases mapped

Benefits

It was frequently pronounced that algorithms had the potential to generate significant improvements in public services; from improved diagnoses of medical conditions, better care of patients, better informed public health decisions, and improved capacities to catch criminals and safeguard communities.

The automation of key processes, especially those related to information provision, was expected to make service provision more efficient and timelier, as well as potentially freeing up time and resources for other activities.

However, it was noted by participants in the stakeholder workshop that many of these benefits may primarily be felt by the service providers rather than service users. Other cases from the mapping also illustrate that the discriminatory potential of these approaches means that benefits are likely to be disproportionately enjoyed by more privileged groups. This includes those living in cities with easy access to well-funded teaching hospitals, those without existing police or social service records, and those whose immigration status is not under question.

Risks

Some of the risks identified in the cases mapped were mainly concerned with *data collection, storage and sharing*. The most commonly identified were:

- privacy
- the sharing or storage of data without a person's knowledge
- data security
- informed consent and confidentiality

Another category of risks identified in the cases mapped are more concerned with the *applications of algorithms and associated technologies*. The most commonly identified were:

- discrimination or bias,
- inaccuracies and mistakes
- exacerbating existing inequalities
- abuses of human rights or other harms.

A further category of concerns identified were more broadly about the *governance of these approaches and their potential to provide good public services*. These concerns include:

- transparency and accountability
- governance and oversight
- the potential for new approaches to be substituted for other more beneficial forms of service provision.

These findings show that it is not only that the risks and benefits of newly adopted algorithmic approaches in public services need to be considered, but that they also need to be compared against alternative solutions to the same problems.

Public concerns around discrimination come out strongly from the examples of public engagement around policing, immigration and social care. Concerns about surveillance and human rights were mainly related to policing and immigration. Citizens expressed worries that already marginalised communities face the greatest potential harms from the adoption of these approaches because of algorithmic biases built into assumptions of particular software, disproportionate amounts of existing data available about members of these communities (due to, for example, contact with the Home Office or social services), and discriminatory assumptions and actions built into earlier datasets used to train algorithms.

Concerns about inaccuracies and mistakes emerged from cases across a range of different public service areas, including mistaken identities in facial recognition technology used in policing, examples of overly long sentences given through justice system software, or misdiagnoses of illnesses in healthcare.

Mapping participation with algorithms in public services

The mapping shows evidence of a range of public engagement activity around the use of algorithms in public services across different parts of the UK, different public service areas, different institutional settings, and different technical applications
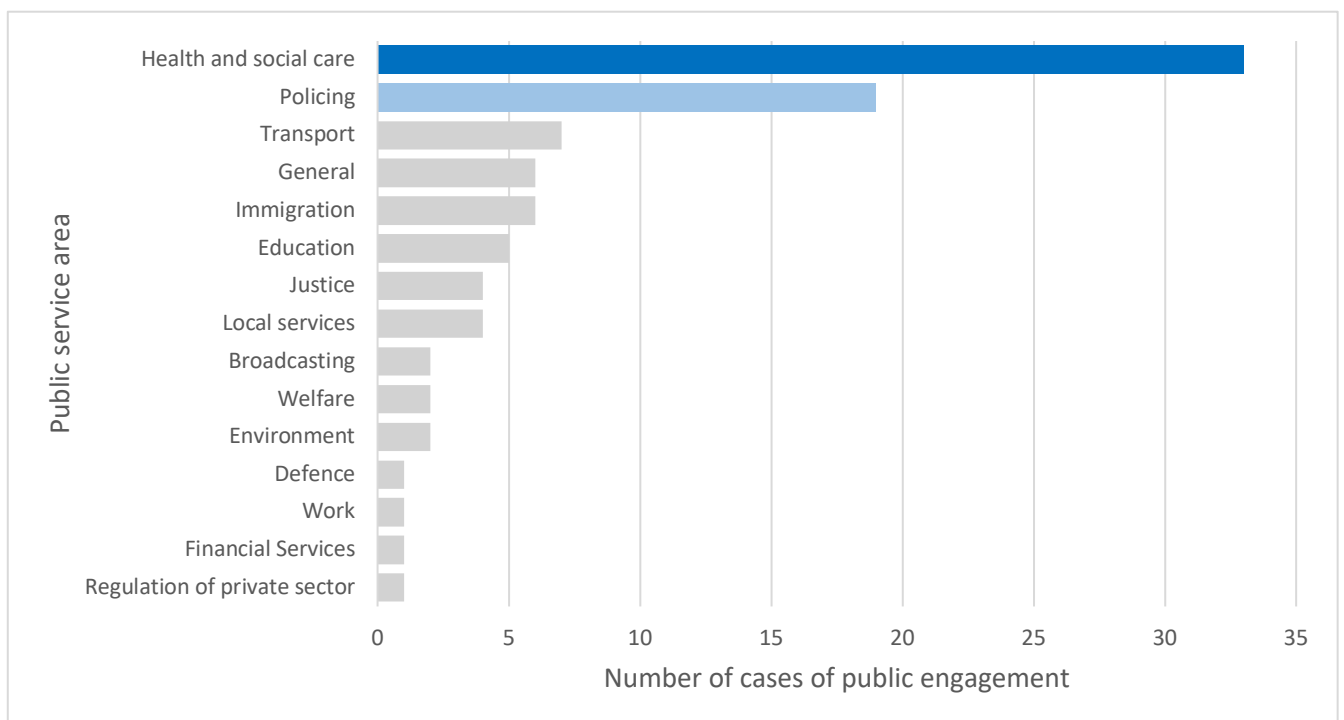

Figure 3 Number of cases of public engagement by public service area

Geographical distribution

The cases mapped were mainly nationwide (n=35) or focussed on England (n=25) and London (n=13), though there were also some cases specifically concerned with Scotland (n=6) and Wales (n=7).

Public service areas

As figure 3 illustrates, most of the cases of engagement mapped were around health and social care or policing. These are both areas where these approaches have been controversial and attracted media attention, such as around the sharing of NHS data with private companies, concerns about discriminatory potential of predictive policing, and the use of biometrics and facial recognition technologies by police.

There has long been a public conversation around the uses of data and associated technologies in **healthcare** – these tended to be the topics of the earliest examples of public engagement from the mapping– and this is a public service area which the vast majority of the population has regular engagement with. There are a number of key specialist organisations around healthcare – namely the NHS and the Wellcome Trust – which have been involved in orchestrating mainly deliberative or survey-style public engagement on this topic.

Around **policing** there has been a large range of very different kinds of engagement, from protests, to live trials of new technologies or court cases. Furthermore, these cases are characterised by the involvement of very different kinds of organisations including public sector agencies, campaign groups and research bodies.

Significant organisations

The Nuffield Foundation and the RSA have also been significant bodies calling for and orchestrating public engagement around algorithms in public services. Campaign groups like Big Brother Watch and Liberty have been significant in stimulating engagement around policing and justice, whilst 'digitaldefendme' has been significant in shaping engagement with education issues. Market research companies such as Ipsos MORI and YouGov were involved in a number of cases from the mapping, as were some of the technology companies themselves like Amazon and Google.

Technologies

The cases mapped cover a range of different technical applications. **Data collection, sharing and use** were the most common technical focus. Among the most recent cases there is a strong interest in **facial recognition technology**. These cases covered a range of different engagements from more activist style engagement to more formal deliberative processes.

Predictive analytics and risk assessment analytics were also a significant technical focus of cases and one which has been similarly controversial. This technical application has been the focus of a significant number of activist engagements, information campaigns and academic projects.

Apps, chatbots and wearables were both the subject and means of many cases of citizen engagement, mainly encompassing more mundane or everyday engagements.

Next Steps

This project illustrates the value of mapping public engagement with the use of algorithms in public services in order to identify emerging concerns and hopes about these approaches, and to observe new areas of interest and forms of engagement. It also demonstrates that whether or not the government formally engages citizens, they will engage with the issues themselves, leading the debate and the reaction to emerging technologies.

This is particularly true in the context of the current COVID-19 pandemic which is bringing to the fore a range of issues related to sharing of health data with private companies, surveillance potential of tracking contact, movements and infection risk, as well as broader issues related to the treatment of immigrants. This is likely to lead to an explosion in new cases of engagement around the use of algorithms in public services which raise new challenges and concerns for policy actors to pay attention to. The mapping approach adopted here provides a more comprehensive evidence base around issues, concerns and hopes raised through citizen engagement than reliance on one singular public engagement process – however high profile and well-run it is. Our approach encompasses multiple different framings of the issues and objects at hand, and goes beyond a focus on only the institutionally orchestrated forms of participation which are usually highlighted.

Policy makers concerned about public trust in the use of algorithms and AI in the delivery of public services need to urgently consider how to engage more proactively, and in a sustained way with the multiplicity of public engagement going on in this area.

The field of energy policy offers one potential way forward, where an Observatory for Societal Engagement with Energy (SEE) has been established by the UK Energy Research Centre to carry out these mappings of public engagement with energy systems and net zero on an ongoing basis[5].

A dedicated body to take forward this mapping and monitoring of public engagement around the use of algorithms in public services would contribute better synthetic understandings of societal values and public conversations, and also allow the exploration and shaping of broader narratives around the use of algorithms in public services.

Key functions for this body identified from our stakeholder workshop include to:

1. curate and maintain an open repository of data on cases of citizen engagement in this area to be shared and used by actors within and outside of government.
2. identify forms of engagement, communities, public service areas, and perspectives which may currently be missing from the national conversation.
3. give broader guidance and support to those working in this field, for example through reports and toolkits.
4. coordinate between diverse actors, avoiding both the replication of effort across different projects and organisations, and to connect together discussions which are currently occurring in different silos.
5. provide a space and focus within this emergent field for reflection and discussion, allowing the challenging of dominant narratives and assumptions, and critically reflecting on the potential futures of such approaches.
6. draw on lessons and intelligence from beyond the UK.

References

1. Centre for Data Ethics and Innovation. (2021). *COVID-19 Repository & Public Attitudes: 2020 in Review*.
2. Noble, S. (2018). *Algorithms of oppression*. New York: NYU Press.
3. Donovan, J., Caplan, R., Matthews, J., & Hanson, L. (2018). *Algorithmic Accountability: A Primer*.
4. Balaram, B., Greenham, T., & Leonard, J. (2018). *Engaging citizens in the ethical use of AI for automated decision-making*. RSA: London; Adams, L., & Burall, S. (2019). *How to stimulate effective public engagement on the ethics of artificial intelligence*. Involve: London; Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). *Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research*.
5. https://ukerc.ac.uk/research/see/

Acknowledgements

Suggested citation: Pallett, H., Burall, S., Chilvers, J. & Price, C. (2021) 'Public Engagement with Algorithms in Public Services' 3S Research Group Briefing Note.