

# Roundtable 1: AI and Machine Learning in Public Opinion Research

*Dr Zach Dickson (LSE), Dr Tom Robinson (LSE), Dr Ben Tappin (LSE), Chaired by Dr Miriam Sorace (UoR)*

The discussion focused on the benefits and challenges of using Machine Learning (ML) and AI, primarily Large Language Models (LLMs), in public opinion research.

We talked about the **opportunities that LLMs presented** to public opinion research, focusing on three areas: administering dynamic surveys, conducting large scale interviews, and generating synthetic data.

- An LLM running in the backend could, for instance, rephrase a question on the go while administering a survey, or dynamically assign bespoke experimental treatments.
- LLMs perform very well as interviewers, enabling the depth of qualitative research but conducted on a much larger scale than normal.
- We can also use LLMs for generating synthetic data where they are great at predicting the mean, but currently their variance is far too low.

We gauged the **overall sentiment in the room toward LLMs**. Most were not entirely opposed to the use of LLMs but there were concerns about bias and the lack of reproducibility.

- LLMs have potential uses such as classifying and coding open text responses in surveys, and for fact checking information that it would be difficult for humans to do.
- But there are worries about everyone focusing on LLM compatible research which might detract from other useful research, as well as researchers not reporting which models they are using.

We also discussed the applications and limitations of **ML in social science research**.

- Often causal questions in social science can be effectively approached as predictive tasks, which aligns with where ML is most useful.
- While ML can be helpful for predicting heterogeneous treatment effects, it is not necessarily doing anything that a regression model couldn't do, but the benefit is that it is doing it in an inductive way.
- But the benefits of ML can be overstated, and we should question how much we actually benefit from the marginal improvements in accuracy it provides, given the added compute and complexity.
- The black box nature of ML can obscure the "why", which is something we are interested in as social scientists.

Overall, we acknowledged the transformative potential of AI in public opinion research while emphasizing the need for careful consideration of its limitations and ethical implications. We agreed on some **best practices** when it comes to using AI for social science research including using multiple LLMs for the same task, defining your ground truth before you apply AI to your research, and being realistic about what it can actually help with.

## Roundtable 2: Identity Politics and Public Opinion Research

*Professor Rob Johns (SOTON), Dr Katharina Lawall (UoR), Dr Stuart Turnbull-Dugarte (SOTON), Professor Jason Reifler (SOTON), Chaired by Dr Marta Antonetti (LSE)*

The roundtable explored various aspects of identity politics and how they shape public opinion and how we approach public opinion research.

We discussed **negative partisanship** and **polarisation**, and their effects on inclusion in political spaces.

- Negative partisanship — identifying strongly against a political party — can be a more powerful motivator than partisan identity.
- This negative partisanship is contributing to political polarisation which in turn can negatively affect women in politics, who are much more likely than men to drop out.
- This is partly a result of very gendered differences on how political spaces, like a council meeting, are experienced in terms of their hostility. Women are significantly more likely to perceive them as hostile than men are.

We also talked about the changing attitudes towards **queer individuals**.

- There had been a dramatic shift over time away from reported homophobic attitudes to them now being largely socially unacceptable.
- But despite that we still see evidence of non-expressed homonegative preferences in terms of how LGBTQ+ individuals navigate through some aspects of life — potentially also indicated by the use of social liberal language as a cover for illiberal views.
- Among younger generations, there is a big gender divide in queer identification with many more women than men identifying as bisexual or gay — men are probably underreporting due to perceived (and real) penalisation for admitting same-sex activity.

We touched upon how research has changed with regards to how people reacted to **corrective information** about misperceptions they believed.

- While earlier research showed that people did not respond well to corrective information and that it in fact reinforced misperceptions, experiments now show that corrective information actually does have an effect on their attitudes.
- This ought to change how we think about the limits of political identity, and what the effects of that are.

We also explored **national identity** and how our understanding of it has changed in relation to public opinion.

- National identity used to be a fixed part of people's identity, but what we have observed in UK politics in recent years is national identity becoming endogenous to other things, including ideology.
- This change is motivated by political events like the Scottish independence referendum and the EU referendum — pro-independence Scots have become less likely to identify as British, and liberal Britons more likely to identify as European. This suggests that changes in political opinions are driving changes in identity, not the other way around.

Together these insights point to different aspects of identity politics (gender, sexuality, partisan identity, national identity) and how they are interacting with public opinion. We concluded that identities are endogenous to policy and to an extent all politics is identity politics. The link between certain identities and politics can shift over time and so some identities can get depoliticised.

## Roundtable 3: Advances and Challenges in Survey Research

*Professor Will Jennings (SOTON), Luke Taylor (Verian), Joe Twyman (Deltapoll), Anthony Wells (YouGov), Chaired by Professor Sara Hobolt (LSE)*

This discussion focused on the state of survey research, particularly on the challenges that public opinion firms are facing.

There was much focus on what was seen as the biggest problem for the polling industry right now which is the increasing problems with **panel quality** in surveys.

- There has been a sharp rise in the last few years of fraudulent responses in surveys driven by the lower barriers to generating fake survey responses due to the expansion of AI.
- Often around 20% of responses typically have to be thrown out but it can even be many more than that — a lot of AI response at present are very poor and easy to spot but the issue is what happens when they get better and it gets more difficult to spot them.
- Firms will need to be doing more to validate that responses come from human beings.
- Another problem with panel quality is that surveys frequently source from a blend of panels, which can include duplicates as a result — but this can be minimised using postcode data.

We evaluated the benefits and drawbacks of “gold standard” **random probability samples**.

- Given the time and budget constraints for political polling work – this approach only tends to be possible using random probability samples that have been pre-recruited to participate in other surveys using very short surveys that are conducted online only.
- The main advantage is in how it tackles the panel quality problem by requiring the use of a physical address.
- But it does not ensure a completely representative sample and weighting is required to compensate for observable non-response bias, it's expensive to maintain, and the sample sizes means it only really works at a national level and less for subsamples.

We also highlighted that despite major methodological changes over time, polling actually broadly remains quite accurate globally, though while noting concern that we don't know how **information dissemination** is influencing people's political views as much nowadays.

- Polling has evolved and become more complex than it used to be many decades ago when you could just broadly base it on the class structure, but despite that, UK excluded, recent elections around the world have been good for the polling industry.
- There is more obscurity than there used to be on how information sources are affecting people's political views as they increasingly get their news from WhatsApp or YouTube which doesn't say much about what content they are being exposed to — making it much harder to design survey questions around media diet and information environment.

Finally, we reflected on what has **changed over time** and what hasn't in the polling industry.

- Many things have improved such as the speed at which we can get 1,000+ responses and the variety of things you can do as part of the survey response experience thanks to the online survey methods, compared to older door-to-door methods.
- But some of the same problems still exist to do with getting people to actually take part in surveys, and with people who do take part in surveys being more engaged than the average person — and there are also still a lot of bad survey questions out there.

While polling and surveys have advanced a lot over the years, they continue to face challenges around panel quality, non-response bias, and an increasingly opaque media environment.