Transcript: NCRM Annual Lecture 2024

[0:00:00]

Gabriele Durrant: Good evening everybody. I would like to welcome everyone here tonight to the NCRM Annual Lecture. So, it's an annual event and we have got a really exciting one this year.

I am Gabi Durrant. I'm the director of the National Centre for Research Methods, NCRM. And this very short video about NCRM and then a little bit of the background is actually linking to an extremely significant year for our centre, we are celebrating our 20th year anniversary in fact.

Obviously NCRM has changed a lot over these years and different remits, different teams and so on but really the focus is on research methods.

And this lecture effectively forms part of this celebration and it's absolutely fantastic to see, you know, so many different people here from various different backgrounds, academia, particularly from government, from industry, from charity sectors and so on. We've had more than 600 registrations in fact and obviously this is online as well so everything will be recorded, and also later on the videos and the event will be available from our website after a few days of processing and so on.

I'm really delighted to have, yeah, so many people here from these different sectors and disciplines and so on and it's really sort of a telling reflection on the importance of the topic first of all that we are going to listen to in just a moment but also of course of the reach of the centre.

And just before we move to the actual lecture, I just want to briefly say about the schedule for this event tonight and just a couple of words about NCRM. So I will in a little bit introduce Noortje Marres, our key speaker for tonight, and she will speak for about 45 minutes and then we will have two

discussions for about 15 minutes and then we have about 30 minutes for sort of questions and open discussion and people that are online can also use the YouTube livestream to input into that .

We may not be able to answer all of the questions given the time available but we'll see.

Just very briefly, a little bit of background about NCRM, so NCRM is one of the key UK providers of research methods, training and capacity building, so we are really focusing on high quality training, that is across sectors. And we really have the mission to sort of advance methodological expertise across the social sciences and beyond, so particularly also this interdisciplinary, and we are really focusing with our lecture today on that.

We are funded by the ESRC and at the moment we are a network of 12 partners led by the University of Southampton all across the UK and we are running very many courses and events across the year and it's really on a vast array of topics, research methods related of course. We have several hundred researchers that we are training every year, in fact over the last four years we trained more than 4,000 people just on courses alone, obviously in addition we have got loads of events where sometimes several thousand people sign up for and some things are obviously online.

We are also supporting the research community via for example, online learning resources and everything is available free on our website so please have a look on our website to visit that and see a little bit what we are offering. And we have also developed strategic partnerships with, for example, ESRC investments and other key non-academic stakeholders across the UK over the last few years.

And we have run, or we are running, a number of strategic initiatives which really aim to sort of nurture innovation and research methods. We are bringing together researchers from diverse fields, really different ways of life and research environments and we are really aiming to respond to the very

many challenges facing us today and into the future, particularly obviously from a methods perspective.

[0:04:09]

And that really brings me very neatly to tonight's lecture and the sort of topic. We will explore one of the most compelling new issues facing researchers and indeed more broadly society and the sort of focus on the rise of technological changes including artificial intelligence.

And basically throughout this year, 2024, we in NCRM we are trying to explore the role of generative AI on social research and the challenges and also opportunities that actually will bring is just one of the topics we focus on.

So I'm really thrilled that this evening we have a highly distinguished leader in this critical area delivering our annual lecture and I would like to introduce Professor Noortje Marres tonight.

I will just say a few words about you before I hand over. She is a really leading scholar in the field of science, technology in society. She is based at the University of Warwick and she has conducted extensive research on participation in technological societies and made hugely influential contributions to interdisciplinary methods and methods development so that's really something we are keen on.

Her recent research focuses on experiments for example examining diverse forms of testing in societal settings from street trials and we hear a little bit about that in your talk of intelligent vehicles to fact checks and media environments and so on. And her lecture today is entitled, 'After the Automation of Methods the Case for Situational Analytics,' and we will explore here the new sort of challenges that AI and other technological changes pose to the sciences of society

And drawing on your recent research, Professor Marres will present a new set of methods that social researchers across disciplines have developed to

address these different challenges in society, situational mapping, situational analytics, so we will hear a lot more about that.

And just very briefly I would also like to introduce the two discussions afterwards, Carrie Friese, Associate Professor of Sociology at LSE and you've been a co-developer of situation analysis, and also Rachel Coldicutt, researcher and strategist specialising in the social impact of new and emerging technologies and you are founder and executive director of research consultancy, for example Careful Industries, and in 2019 Rachel was awarded an OBE for services for The Digital Society.

So it's going to be an absolutely fascinating lecture I believe and I imagine everyone is really thoroughly looking forward to hearing much more about the topic and us discussing afterwards, so without any further delay I'd like to hand over to our key speaker for this evening Professor Noortje Marres, thank you.

(Applause)

Noortje Marres: Hello everybody, good evening. It's a real pleasure and also an honour to give this lecture this evening at the invitation of the National Centre for Research Methods.

I will be looking forward to the responses of these two esteemed discussions and also to then discuss with all of you afterwards.

To start I would like to introduce you to Yugasa Bot, an AI enabled focus group moderator. This bot can run focus groups in 100 languages. It can structure your data analysis for you and it will document everything. Now expected by the end of the evening we will be able to decide together whether we believe this claim, but this is how we're often introduced to AI and the application of AI in social science.

[0:08:37]

Since the release of powerful new large language models in 2018, the fully automated generation of natural language has been hailed as the next frontier in the seemingly inexhaustible digital revolution. Computational systems like Open AI's ChatGPT that needs no introduction, a system like that has not only had a massive uptake around the world but it is also deemed capable by many of humanlike reasoning and expression.

And the proliferation of those kinds of claims has also helped in spawning a wide range of additional applications built on ChatGPT to create real world applications of chatbots and large language models.

Now it appears that social research is an especially fertile ground for the creation of these customised applications of large language models, so alongside focus group bots like this dedicated Chatbot, there is an application of ChatGPT for focus group moderation which has been developed by a company called Speak Ai, there are AI enabled surveys that we might discuss later and applications like BERTopic which adapt this large language model for the sematic analysis of textual data.

Now these kinds of applications of generative AI for social research are often pitched as a way of scaling up research designs, of making resource savings and optimising research management. As Speak Ai, the creator of this ChatGPT application for focus group research puts it, "ChatGPT for focus groups will save you 80% of your time and 80% of your costs", that's the way in which the sort of, yeah, the application is pitched.

Crucially however the use of generative AI in social research is not just framed as an operational matter, the promotional material of the same company lists various ways in which chatbot moderated focus groups are better than their human led variant, noting that GPT can be used to ask questions in a more natural way and will allow for more meaningful conversation and insight.

Now I am then particularly struck by the ways in which advocates of AI based social science and science make the case for the uptake of AI and social science on substantive grounds. So here is an example of a similar claim by organisational scientists Koehler and Sauermann, who state in a recent article, 'AI can perform core research tasks such as generating research questions, processing data and solving problems'. Generative AI in other words, is being framed today as a way of delegating knowledge production to machines, and I believe that that is what's most distinctive in some ways.

In this lecture I would therefore like to ask on what grounds is this justified and with what consequences? I will begin by offering some wider reflections on the uptake of large language models in social research and I will identify a number of more general challenges that it poses for the creation of knowledge about society.

I will then introduce an interdisciplinary method called Situational Mapping which I believe can help equip us to address these challenges.

Next I will present some findings and some insights from recent collaborative research projects in the area of AI in society that I have conducted together with others, some of whom are in this room, at the University of Warwick.

And in this research on AI in society we used AI based methods alongside other methods and I will, yeah, try and introduce that range but often, and indeed perhaps always within a framework of situational mapping.

[0:13:46]

Our use of AI in these projects was relatively basic but I think it can offer a good ground for exploring some of the implications of AI for social science from a practitioner perspective. In my conclusion I will summarise what I believe will be required of us to address these implications.

I will make the case, some advanced warning that we will need to develop new and possibly quite different evaluative frameworks for assessing

robustness of our methods in social research, but I will also make a wider claim that will try and sort of make it necessary for social researchers from many different backgrounds to become much more curious than we sometimes are today about the way in which automation unsettles our methods and disrupts more established frameworks in social science research.

Okay, so we'll start with an overview of challenges. Machine learning based applications for the generation of natural language are today widely used for a variety of tasks in science and in social science. Automated transcription of speech has become quite common for online interviews which involves the creation of more or less instant textual transcripts, so is the use of automated language translation and the use of language generators in peer review and scientific writing.

The latter was recently brought to my attention when someone mentioned a study which showed a significant increase in the use of the word delve in the titles and abstracts of scientific articles, a word for which apparently the chatbot, ChatGPT, had a fondness. It was also pointed out to me that this does not necessarily mean that scientists now have ChatGPT write their articles for them, it rather indicates a growing reliance on large language models to improve clarity of expression and to speed up the writing process.

In this regard we might call a lot of the applications of AI and social science as involving the automation of operational tasks. However, crucially the use of AI in social science is not just limited to these kinds of tasks, meaning tasks that do not involve the creation of new knowledge necessarily, operational tasks, I think it's important to be precise in how I define them, are then tasks that are about conserving and disclosing meaning during the processing of data and the right of results, they are not about creating new knowledge or new meaning.

However, over the last years computational social scientists have advocated what I think of as the delegation of knowledge acts to AI such as the annotation of interview and conversation data most notably.

Gilardi and colleagues have reviewed the use of ChatGPT for the analysis of content, discourse and conversation positing that the chatbot now outperforms human coders for these types of tasks and claiming that they do better than their human counterparts for annotation tasks including relevance detection, stance, topic analysis and frame detection.

While machine learning models have been used for several decades in computational social science, what is significant about this kind of claim is that here general purpose pre-trained models, GPT, this is what GPT stands for, general purpose pre-trained I came to realise, that these general generic models can now be used without any custom training by social scientists on their data and this without even providing the model with examples of their data, so with no custom training or examples at all.

Now this sort of very ambitious claim that now we can perform social science with general purpose AI has been questioned by other colleagues including my colleague Michael Costello who's here and I hope will join us in the discussion later. But even those who are more cautious in their application of large language models in social science still make very ambitious claims to the effect that for instance, the claim by Ziems and all that large language models can reliably classify and explain social phenomena like political ideology. These same authors also propose that social scientists can use these applications to implement social theory in their writing and using ChatGPT and other models to stylistically restructure utterances so that they are brought in line with a social theoretical framework, so these are examples of this kind of delegation of knowledge acts to AI.

Now scholars for some years now have worked hard to demonstrate that there are significant problems with these proposed uses of generative AI in social science, first and foremost they have shown that this use of large language models for social research is marred by significant biases. These models generally speaking perform much better in the analysis of data that they have been trained on which means they tend to do much better at analysing the English language internet, that's sort of the brief summary.

Ashwin and colleagues found that large language models perform less well in the analysis of non-English interview data and they also found that the results of such analysis are less accurate for distinctive social economic attributes, so the speech of some subjects can be more reliably analysed using AI than the speech of other subjects and speech in other contexts.

I find it important to note that such biases introduced by AI in social science do not only concern the quality of AI based social data analysis itself but in the longer term may well affect the wider social science knowledge landscape.

As free to use models like ChatGPT perform especially well on easily available English language data, their popularity will quite probably make growth of this type of research more likely while the analysis of minority phenomena, if I can use such a term, will continue to require custom configured models and custom configured research design, that is to say the uptake of general purpose AI to conduct social research may result in biases and growing biases in the overall empirical coverage of social science.

The push towards the automation of social research may also contribute to growing research asymmetries in social research but automated social research tools, which will be easy to use, will work especially well for mainstream English language phenomena and are as such, also likely to be available to many researchers. Iterative custom configured research is, and will then become more likely, to be framed by comparison as a kind of luxury research.

The uptake of AI in social science is then likely to have important implications for the political economy of knowing society in an age of AI.

But I would like to direct your attention now also to something else, namely the question of whether and how AI enabled methods will help to advance social science.

I have been fond of noticing the use in articles by computational social scientists of the work of the philosopher of science, Thomas Kuhn. The term 'paradigm change' is very often mentioned in computational social science articles about the use of AI for knowing society, and there is one quote in particular that keeps returning in these articles, namely Kuhn's claim that paradigm change and advances in science do not just come about through the formulation of new ideas but through the invention of new tools and new methodologies.

There is something odd, however, about the use of this quote in these articles and the use of these quotes in relation to AI based social science because it seems that to date the most successful examples of the use of large language models in social science operate within strikingly conventional methodological frameworks, focus group research, content analysis, conversation analysis. Yes, the tools are certainly new but the methodologies perhaps not so much.

[0:24:48]

I would like to discuss what I'd like to provocatively name this relative lack of methodological innovation in AI based social science in more detail in what follows but to sum up I can then say that I see two important challenges that the use of large language models poses for social science alongside the major problem of bias.

So this is first these growing asymmetries between instrumental research, a sort of fully automated research that works with free and easy to use tools and then on the other side, custom made, custom configured social research

projects, so the asymmetry between those types of social research and second, this paradoxical failure to realise the opportunities that new tools open up for the creation of new methodology in social science.

I will go on in a moment to show how these challenges of AI how they can be addressed in social research that works with situational mapping but before I do that, I would just like to give one more lens on this problematic of how AI is introduced in social science because I think it will, yeah, it will help us, it will put us in a better position to evaluate what AI can do for social science and what it cannot do for social science.

So the three challenges that I just presented that AI poses for social science in my view have much to do with what I think of as the delegation model for the use of AI in social research, this model sets up the wholesale transfer of knowledge acts from human researchers to automated systems as the normative frame and the ideal to strive for in using AI for social science.

This kind of approach is latent, it's implicit in the very definition of AI which is often defined in terms of the capacity of machines to perform tasks that would have otherwise required human intelligence. So this kind of definition of AI implies already an evaluative principle, namely one that sets up an equivalence between human and machine performance as the standard by which to assess AI.

This principle is still with us today in the shape of methods of accuracy testing of large language models, an approach whereby indeed both human and machine performance in certain knowledge text tasks is judged against the same standard, so there's a kind of an equivalence and idealised equivalence between human and machine performance in knowledge tasks that is implied by these kind of evaluative frameworks like accuracy testing.

Computational social science apply a similar principle when they use criteria like intercoder agreement to evaluate how well machines and humans are at the coding of content, the annotation of data. Now the problem with this

approach is that in order to evaluate the quality of research performed by machines and humans, it posits their sameness, it posits equivalence between humans and machines.

The limitations of such an approach are now being discussed including by Tornberg who argues in favour of what we could call a human in the loop approach or a coordination model for the use of AI. He proposes that the aim should be to integrate large language models into research design but to make this integration conditional not on whether the alignment between machine and human can be assumed but whether it can be achieved, so there's a kind of a commitment to undo this assumption that for machines to do well in social science they have to be equivalent in their performance to humans and the other way around.

Now I think this is one of the most promising ways forward and I will explain why I think so but it's also really important to note that it's very difficult today to pursue any alternative to this kind of delegation model because these ideas of equivalence between human and machine tasks are baked into AI based methods as we will see.

How then to address these in some ways quite general and deep seated challenges that AI poses for knowing society.

[0:30:18]

Now it is here that a distinctive set of interdisciplinary methods; digital, visual and participatory methods of situational mapping can help us to advance in how we understand the role of AI in knowing society.

The approach situational mapping has routes in sociology going back to the early 20th Century and has been adopted across several disciplines over the last decades. Situational mapping can be defined in terms of following three commitments, first it makes the situation the unit of analysis. Situations have been defined very differently by sociologists but it almost always involves a grounding of our understanding of society in a specific place in society, say,

in a hospital or a city or the commitment to ground social analysis in a particular social grouping, the perspective of, say, antinuclear activists or nurses.

The second feature of situational mapping is that it takes up data mapping as a qualitative method, it uses these techniques of mapping entities to surface and qualify what composes situations. So in the words of Adele Clarke, the aim of situational mapping is to specify what entities of varying scale and type compose the situation? So it's a compositional method you could say in that regard.

Historically this approach has been developed, as I already mentioned, in sociology and it has historically also been primarily used for the analysis of interview data and fieldwork data but it has since been taken up in many different fields including in design research where as part of participatory design situational mapping is used to plot future scenarios in this case of mobility practices of how we travel and the plotting of these future scenarios in a social space composed of governance, ecology, culture, infrastructure, economy.

And also the approach has been used in digital media studies as here in early work where we relied on hyperlink analysis to map conflicts that were emerging on the Web in the early 2000s around the construction of the Narmada dams in Gujarat in India.

Situational mapping then what I'm trying to say is an avowedly interdisciplinary method.

To clarify what is involved in the use of situational mapping to analyse online data I have previously offered the term situational analytics. Situational analytics combines fieldwork based situational analysis with digital methods which these days often rely on automated data capture via so called APIs for application programming interface, so capturing data through semiautomated means from platforms like Twitter or YouTube. And here you see an example

of a mapping of the introduction of automated vehicles in different cities and different places in the West Midlands, and the analysis of issues this raised on Twitter, yeah, in that early period in 2016.

I included this example to make clear that when you use situational mapping to analyse online phenomena you inevitably end up mapping situations across different settings, in this case across Twitter and the streets and places where automated vehicles were being introduced, so there's a kind of a heterogenous kind of, yeah, space in which situations are mapped in that case.

It's also important to note that once you work with this kind of online data that situational analytics end up often scaling up the situation and the analysis of it because it is now possible to analyse the situation within a whole platform setting, a context, so how does the phenomenon unfold on Twitter? And we often then use semiautomated methods to analyse and create maps of situations at scale.

Now the possibility that I'm interested in here is how can these methods of situational mapping and situational analytics provide an alternative framework and a different orientation for addressing the challenges that AI poses for knowing society? Rather than answering this question in the abstract, what I would like to do is to show you how we apply situational mapping and the principles of situational mapping in recent research on AI in society and in that way hopefully make clear to you how, yeah, we offer a different framework in doing so.

[0:36:24]

Shaping AI is a three year international research project in which we mapped public discourse about AI in four countries for a 10-year period between 2012 and 2022. In this project we applied the principles of situational mapping in the following way; to start with we grounded our analysis of UK controversies about AI in a specific community. Rather than identifying controversies about

AI by querying Twitter or querying the media, we started by consulting with AI and society experts in the UK which we defined broadly in terms of all those with a stake in the issue and committed to genuine debate.

In our online experts consultation which we conducted in the Autumn of 2021, we asked what according to you, so according to these experts, has been most controversial about AI in the last ten years just grounding our analysis in the perspective and the standpoint of this community.

We sent out our consultation to 250 experts and received 53 responses. Now it took us some time to interpret these responses as it was not immediately obvious what type of controversies they had identified. In the sociology of science a controversy is usually defined in terms of the staging of disagreement between experts but we found relatively few examples of this among the responses, so what had we found instead? Through close reading of the consultation responses we identified three different types of what you could call forms of controversiality of AI, different ways in which AI was becoming or had become controversial and we called these topics frictions and problematisations.

Now focusing on the middle category, that of AI friction, these were especially prominent among the results to our consultation. Frictions are not disputes or controversies per se but they refer to technological systems, sites or incidents, in which AI gives rise to demonstrable trouble, harm or contestation in specific environments in society, and here you see a range of these frictions that were mentioned in the consultation, and here you find them listed in the middle category, so these are responses to the question what makes AI controversial?

We also indicate here the topic in relation to which these frictions, AI frictions, were mentioned by our respondents. Now you can see that the topic of facial recognition has an especially broad number of frictions associated with it and so do tracking and targeting and corporate research culture. You also see the GPT3 in Autumn 2021 already figured quite prominently among the AI

frictions and so did the firing or resignation of Timnit Gebru, a Google researcher who had written a controversial paper about large language models.

Now is this then how AI became controversial in the UK in the relevant period 2012 to 2022? Did it become controversial through the topicalization of these frictions? I'm afraid it is more complicated than that because we continued conducting these consultations in the different countries that participated in the projects and found again quite a different distribution of AI friction among the responses to our consultations in four countries. The consultations in France, Canada and Germany were conducted in a later period in 2022 and '23 and here you see large language models as a category emerging as really a dominant sort of site of this friction.

One reason I'm showing you this is to show that when we ground the mapping of AI controversies in different communities, which is what we did through this method, we become confronted by the variability and you could even say the instability of our empirical object. We started by looking for controversies but we found frictions instead and these frictions did not have stable properties. The frictions that were surfaced in the 2021 consultation in the UK were very different from those surfaced in the other consultations.

[0:41:47]

Once you adopt a situational mapping approach to AI in society we find ourselves frequently shifting our frames and we find ourselves frequently reframing our empirical object, and this is something that can seem incompatible with the stable frame of reference that seems required for robust data analysis.

Now in the next phase of our research we worked with AI methods to further develop our situational mapping of AI frictions. To analyse debates sparked by AI we turned to Twitter as a setting where at that time one could still find exchanges between academics, journalists, activists and industry

perspectives on AI. We curated Twitter datasets for a selected set of our AI frictions by querying the Twitter API for publications that had been mentioned as relevant to these specific topics in the consultation.

We used large language models in different ways in this work including to determine which tweets are in scope and out of scope of the debates that we were interested in but we were introduced in particular by shaping AI team members Michael Castelle and James Tripp, to BERTopic for bidirectional and coder representations for transformers, a term that might need further unpacking that I can give here.

What we learned about this model is that it enables the analysis of so called word embeddings to determine what are representative topics for a given text. It does this by using a pre-trained large language model to convert text to a large number of what are called vector representations in high dimensional space, these embeddings. And these are then reduced through cluster techniques and a measure called term frequency inverse document frequency to determine a relevance of topics and which makes it possible to create weightings for different topics and then generate succinct topic representations.

Now what we found fascinating and I can only say exciting in applying BERTopic to our Twitter data is that it seemed to us that BERT was capable of capturing contextual expression, so here you see the topics that BERT detected in the Twitter conversations for our four objects, Gaydar is an application for the prediction of sexual orientation based on facial analysis using machine learning, NHS DeepMind involves the controversy around the use of an app by DeepMind in the Royal Free Hospital and which involves a large amount of data being transferred to DeepMind.

COMPAS relates to the use of recommenders or predictive algorithms in the courts in the US to assess likelihood of recidivism, which sparked a lot of controversy, and finally stochastic parrots is the debates pertaining to the

dismissal of Google researcher Timnit Gebru and her paper that she coauthored with others on the dangers of stochastic parrots.

Now the word clustering's that BERT generated, and this is after URLs and Twitter handles and emojis and ad mentions were all removed from the data, so it's from clean data, but what this kind of word clustering's for us seem to capture really well was the unresolvedness of conversations like people don't paper, it's not quite a topic, it's more an issue of debate.

Now in this respect BERTopic for us aligned really well with the experience of coding Twitter conversations manually. So we had undertaken manual coding of Twitter conversations as part of this study and we were struck by the ways in which often meaning was unstable or meaning not applied in the conversations that we coded which made it very difficult to assign stable topics to these conversations.

So here you see an example of this, it's a conversation which starts with a defence of the decision in which Timnit Gebru ended up leaving Google, but then the conversation turns into a debate about the sources of bias in large language models whether these derive from data or whether there are also other sources of bias involved in the application of large language models.

[0:47:26]

Now in a conversation like this you see both on the level of the tweets as on the level of the reply chain that meaning shifts constantly and often indeed multiple frames may be applicable to utterances such as in this you miss the word 'only', does that pertain to the debate about the dismissal or about the bias of the models?

Now when one does topic labelling one is forced to resolve that kind of ambiguity and we duly obliged by attributing topics to conversations and solving this by attributing multiple topics in some cases, but we felt that this kind of stabilising of the meaning of the conversations through this kind of

labelling in a way meant that we analysed an object that wasn't actually the lively conversation that we wanted to analyse.

Now when comparing the codes of BERT and the codes that we created manually for these topics, there is again a sense where the sort of reliance on a fixed frame to evaluate our research for us did not get us to the interesting insight, so you could try and establish an equivalence when evaluating how the codes by BERT, which you see in brown and the codes in purple which are the codes that we created according to our codebook, how they are equivalent, you know, is resigned, fired, didn't resign possibly a debate about corporate research culture? We can know that while the human coding and these are the codes of that friction, that debate around stochastic parrot on the level of the whole dataset, so this is not a Tweet or conversation specific.

You can say, "Oh BERT found climate environmental car and environmental green centres, human coding gave us environmental impact" so we can try and look for the degree of equivalence but it seemed much more important to us to look instead for where BERT and us deviated interestingly.

So while our labels identified what you could call problems on a structural level, our research culture, BERTopic are located on a you could say empirical level resigned, didn't resign. Where we coded from male white privilege in AI BERT codes for antiracism, racism. And so it is this kind of exploring of the way in which there is a contestation and a negotiation of the meanings of this conversation that actually that is what you can begin to explore when you put the human next to the machine code.

So not a fixed frame of is it the same or is it not the same but how is meaning being negotiated and contested within this space opened up by machine interpretation?

So I will return to this question of evaluation in the conclusion but I briefly want to show you research from a different project called AI industries because I think it's very important that even though we are now exploring how we can make AI work for social research, we should not necessarily conclude that if it can work for social research, that from the standpoint of situational analysis this kind of AI based data analysis is the way forward because I think it very much is not.

So in a different project called AI industries we work with a very different set of mapping methods to locate AI in society and to do an analysis of how AI manifests in specific social environments as here in the intelligent mobility testbed for the testing of automated vehicles that has been created in Coventry.

Now we are currently creating and preparing a set of listening walks through the street where we invite residents and others who live in Coventry to walk with us and to map how does AI manifest in the street?

Now when you conduct these walks this is what AI data capture looks like, the form of these little brown boxes that are hanging in the lampposts and also in the traffic lights. At the same time this kind of infrastructure has been framed online as providing excess to testbed data, so it turns out that this Holyhead Road has an API through which users can access scene specific data assets and gain access to real world live data feeds.

[0:53:32]

Now there is a disjuncture in this regard which I think is really important, between the way in which AI manifests as a method that we can deploy within the space of data analysis and the way in which AI manifests as hardware in actual environments in society. And it is to explore and really thematise these disjunctures that again situational mapping becomes an important method, so we're working with also using digital technologies like the Unheard City app which detects the Wi-Fi and Bluetooth signals that the testbed equipment

is emitting to create mappings of the sort of AI space as it is seen from the street.

And, yeah, just to refer to this image on the right, what you see on the right is the way in which the testbed is described as part of the mandatory public sort of consultation for the testbed which, you know, means that for a lot of residents and people who pass through the street, the testbed is completely invisible so it's kind of an act of making visible how AI is creating the street as data space but to do that within the street.

I am going to conclude and sum up. So I've been taking you on some detours through research on AI in society which also in some cases works with AI methods. I've taken you on this kind of meandering tour to make clear that when it comes to evaluating what is the impact on AI for social science and what is the right way for social science to engage with AI, it is really important I think that we keep a very open mind.

So the models that prescribe performance testing or accuracy testing as the way to evaluate the use of AI in social science posit a very restrictive, homogenising framework where it is really the equivalence between how humans code data and how machines code data that serves as the kind of evaluative framework.

I've been trying to argue that AI based methods also have very different possibilities where they can make us attentive to the way in which meaning proliferates is negotiated is contested and is fundamentally unstable in social life, and so the question is can we develop evaluative frameworks for the use of AI in social science that affirm that kind of instability of meaning in social life? And it's what I would like to kind of, yeah, really put in the centre as a bigger challenge for which there isn't necessarily a quick solution.

Also wanted to just emphasise that staying within the frameworks of social data analysis can have really quite blinding effects. I've shown it in these last slides for automated mobility testing infrastructure which is of course a

different data capture infrastructure than the ones that we maybe use as social scientists, but I think there is an important similarity in that AI based methods often create very little opportunity for the subjects whose data provides the raw material for analysis, to actually speak back to the analysis.

So this is why I think as a second challenge or way of responding to the challenge of AI, developing frames for participatory method like listening walks, is actually central, should be central to how we respond to AI in social science.

And finally and this is also a moment I think to hand over to our respondents, is that in approaching these challenges in this way what I'm doing is I'm really centring sort of the situational qualities of social life, the messiness, the negotiation of meaning without which things don't happen, the need to always coordinate action for it to even be interpretable and that to centre these qualities of social life and of social phenomena we will I think find ourselves working sometimes with and sometimes against the automation of method and in doing so I hope that what we can really put centre stage is not necessarily the drive to automate social science but the question of where this meets the requirements of knowing society.

(Applause)

[0:59:33]

Gabriele Durrant: Thank you very much Noortje. I would like to come forward our first discussion, Carrie, thank you.

Carrie Friese: So first I just want to say a big thank you to NCRM for inviting me to come at Noortje's key note address this evening and I also want to thank Noortje for giving us such a timely, rich and important talk on AI and social science research methods.

So I should start by saying that I come to this event, and I'm assuming I was asked to do this as an expert in situational analysis. I was one of Adele

Clarke's PhD students when she was writing the first edition of this book, thereafter Rachel Washburn and I coauthored with Adele the second edition of Situational Analysis as well as two edited collections on Situational Analysis in Practice.

So in this context I've given many workshops on the method and it is a largely qualitative method, and these workshops often focus on making situational social worlds arenas and positional maps based on as Noortje said, largely interview material, field notes from ethnography but also documents of various kinds.

These workshops are very hands on and analogue if you like, we use paper and pens, we work around a table and we really focus in on different people's projects and the situations that they're interested in, except that these aren't always analogue of course because the workshops are often on Zoom and maybe that's another story but maybe that's part of the story.

It was in one of these workshops that I was asked if and how digital methods could be included in situational analysis and I have to say I was incredibly relieved that Noortje had already answered that question for me in developing situational analytics. And I've since argued that situational analysis needs situational analytics not only to study the digital, not only if it's being used to study a typically digital situation, but also to do reflexive research whenever digital infrastructures appear in our situational maps and whether or not those situations are of interest to us as a site of the digital.

I'm really convinced by Noortje's argument that we need to use digital methods in order to surface their infrastructural effects including on our research but I have also argued that this cannot come at the expense of the very real embodied people sitting at the end of the computer screen or not within our situations as well.

And so I think this key note is really important because we see Noortje explicitly bringing together her longstanding work on material participation on the one hand with situational analytics on the other and I think this is a really exciting way to foreground both people and things.

So I have to admit that I have never used ChatGPT or any other site of generative AI because I would personally rather not feed these neural networks at the current moment when the only consensus seems to be that there are serious and uneven distributed dangers. And I should say that Noortje herself has given me the confidence in using my outside position in relation to the digital as an analytic device of sorts.

So in this context as an outsider I had some questions for Noortje. One thing that I think is powerful analytically about situational analysis is in highlighting the invisible and this comes up particularly in your research within Warwick. Who or what is not present? What is unsaid? Who is represented by others but unable to represent themselves? These are the questions that situational analysis asks in order to make power and power relations palpable. I wondered can AI surface the invisible or the unsaid or is this the limit of AI? And what are the implications of how this question is answered?

[1:04:07]

Now while I have refrained from using things like ChatGPT, I am also interested in and curious about the possibilities of AI for qualitative research and so I found myself listening to Noortje's talk with some very practical questions in mind as well. I am part of a consortium of researchers from several countries who are currently seeking funding to conduct qualitative interviews with very large samples that aim to be as nationally representative as possible and with this we're building on Professor David Grusky's American Voices project.

I was drawn to this project because we can confirm that the interviews were indeed with real people and not an AI generated interview of what a person

with sociodemographic x might say, in that sense we aim to target people who don't leave as many digital traces and for us the question of bias is therefore the question.

We would still have people do the qualitative interviews and I have to say that I am unconvinced by the proponents of AI that they could do a better job, I just can't wrap my head around that belief but one of our aims is to test large language models to see if they could be used to help us with the coding of what would be a huge dataset.

So we would need to use AI to identify sections of text on both specifically tangible but also less tangible topics. Noortje's talk really made me question the extent to which we are assuming the delegation of tasks approach to AI in this coding process, and I have to say I had a bit of an aha moment where I realised just how much I have assumed this delegation model whenever I think about AI, and so this shift that things could be otherwise for me was quite powerful.

To think instead about alignments and participatory models I think is a really productive way for us to possibly do this coding quite differently. If funded we would work collaboratively with computational social scientists and data scientists to see how humans and AI can work together on different types of qualitative coding, and so I do see the potential for an intervention with the kinds of alignments that Noortje suggests.

And so I am wondering what the varieties of participation might look like going forward as I do assume that AI will become an increasing part of qualitative research, so thank you very much.

(Applause)

Rachel Coldicutt: Hello everyone. So I find myself in an unusual moment here because I'm a practitioner and the research that I do is really a means to an end, and I think what is very interesting about what Noortje has outlined this evening is something that is very troubling when we're looking at particularly policy and

regulating technology, particularly new technologies that are rolling out because one of the things that particularly troubles me is what do we do before participatory models? How do we know what is really happening as it is happening?

One of the things we've seen over the last I suppose maybe 20 years is the extent to which the regulatory adaptations that happen happen really often based on prevalence, so to give you an example that is not AI but thinking about online safety.

We're in a world in which we've had message boards and forums for over 30 years now and just at the end of last year in the UK we had the first piece of legislation which is the beginning of trying to get our arms around that. It will probably take at least another four years for the regulatory environment around that to be actually working and stood up which means that by the time we're really regulating this thing that has emerged and been emerging, we will have 30 years at least of human activity to think about.

[1:09:13]

And there's a real opportunity and a moment I think to use mapping and understanding of what is happening now as a lever to what comes prior. So the things that I'm particularly interested in I think are how do we know when an outcome, whether it is a good or a harm, is just beginning to emerge? How do we see it happening very often to the most marginalised communities?

And, you know, the burden of evidence that is asked for is very often more than anecdotes or heresy, which means that we're often left with polling. So polling tells us about a thing that has emerged, it's already happened. By the time an issue can be captured in a yes, no, don't know form, everybody knows about it and I think what's incredibly valuable and interesting is to think how mapping of the sort we've seen here might be a thing the community

owns and does. Rather than mapping happening to us it might happen by us or with us and become a piece of evidence that we're all able to use.

And I think overall it will take us a long time to rollout properly anticipatory methods and models into regulatory ways of working.

I spent a little bit of time working with regulators in the UK and the burden of evidence that is needed is very, very high but I think there's something fascinating and interesting, particularly as we're thinking about how automated technologies will be rolled out in ways that most of us do and don't notice/understand how we might all have the ability to map differently and show the impacts.

And just to conclude, you know, I would say that lots of the work that I do is about helping people to work out what's likely to happen next and normally the people who are given the power to think about what is likely to happen next they might be enormous companies, they might be defence specialists. Most of us have to rely on what we sense in the moment or what we talk about with our friends, and that actually now that the prevalence of the sorts of tools that we could all use for mapping is greater than ever before, perhaps it's an opportunity, perhaps there's like a covert thing that could happen here where we could all be telling other kinds of our story. Thank you.

(Applause)