Transcript: Responsible AI in Research: Highlights from the NCRM Annual Lecture 2025

[0:00:00]

Gabriele Durrant: I'm Gabi, Gabi Durrant from the National Centre for Research Methods. I'm the Director of NCRM. And, yeah, we've got the panel today, and I will hand over to Professor Mark Elliot in just a few minutes, who will be chairing today's session. So, he's an expert on AI and is work leading the workstream on AI in NCRM. And he will introduce all the panel members. And I'm really delighted that all of our panel members could make it and think that this is a really important topic to discuss and, of course, I really welcome all of the participants today. And in terms of the procedures, obviously we'll be discussing with the panel a number of questions, and then towards the end, about eight o'clock, we have got the opportunity for the audience to ask questions. Just before we start with the actual panel, just a few words about NCRM, the National Centre for Research Methods. So, basically, maybe some of you, well, hopefully most of you will know, but maybe there are some that are not so familiar, so basically we are one of the biggest providers of research methods training in the UK, and we've been funded by the ESRC for, yeah, with different funding phases, five, well, four funding phases in total, and since 2004, in fact, so for quite a long time and in different ways. We really cover a wide range of research methods, quantitative, qualitative, mixed digital and so on, so really a wide range of activities.

It's not just courses; it's a whole range of activities we are offering, workshops, conferences, networks, where people get together of particular investments or with particular methodological interests, methodological special interest groups we are running and so on. So, really, there are loads of ways of getting involved or teaching or, yeah, presenting in some of our

activities. It's really learning across the life course, so CPD, continuous professional development, across all sectors. We are really reaching out to all sectors across disciplines of the ESRC in particular, the social sciences. But it's also even beyond the social sciences as well. So, obviously, anything related to methods is something that reaches out to other disciplines, maybe, for example, health and medical areas and so on. So, we've really experienced a high uptake over the last five or six years. We had more than a million website users and we've had more than 13,000 schools and event participants, so it's a really high uptake and we've really run quite a lot of training days, courses, workshops and so on. And we've published this year a report on our impact, wide-ranging impacts, so you can have a look on that on our website. We have got more than 80 impact cases studies that we got together, and quite a few of them are online. So, please have a look at that impact assessment report.

And I just wanted to say that behind all of this is a really committed team and we really focus on and we are really passionate about research methods, research methods knowledge exchange and learning about research methods, using rigorous and high-quality research methods, because I think that's really at the heart of doing good research and high-quality research. And I'm hoping that today's discussion will actually feed into this thinking about… critically thinking about what methods we are using. I'll now hand over to Mark, Mark Elliot, who will be chairing today's panel and looking at responsible use of AI in research and research methods. Thank you very much.

Mark Elliot: So, the first question, so if it's alright with you, David, we'll go to you first. Yeah? So, responsible AI can mean different things across disciplines, and Wendy mentioned the interdisciplinarity just now. From your perspective, what should we be focusing on and prioritising when we talk about responsibility in AI? Is it fairness? Transparency? Accountability? Or perhaps something else?

David De Roure: Thanks, Mark. I'd like to go back to my knowledge infrastructure perspective and actually, and also affordances. So, one of the amazing things about working with digital methods in social science, in humanities, in scholarship, is that we can build those tools, and I have… I'm proud to have a team of five research software engineers who do exactly that. They're amazing and we can build new tools. We can adapt tools. We can provide the approaches and the methods for people to do their research. And that's kind of an affordance to digital, the fact that we've put it in people's hands, researchers' hands, to create those tools, and now of course they're doing it with AI, in many different ways making use of the large language models, using AI to generate code, which is one of the things it's done surprisingly well. Lots of things we anticipated about large language models, and code generation has now become absolutely routine for programmers. It's very interesting.

[0:05:07]

Wendy Hall: Oooh, I dispute that fact.

David De Roure: Oh, okay. Well, good, yeah. What, for the software engineers…?

Wendy Hall: Yeah, but they all have to check it all and…

David De Roure: Oh, yes. But…

Wendy Hall: And do you remember the mythical man month? Really, really bad. Come on, yeah.

David De Roure: Yes, absolutely. It wasn't very… yeah. But, no, they have the skill to check it.

Wendy Hall: That's fair enough, because…

David De Roure: Yeah. I'm more worried about my students using it before they have the skills to check it.

Wendy Hall: Yeah, yeah, I agree with that. But it's not just replacing programmers yet.

David De Roure: It's not replacing them, no, but they're using it as a…

Wendy Hall: But it's changing the whole way we talk about it, yeah.

David De Roure: As a productivity tool. Yes. So, my concern is use of AI uncritically, not just at that point in the research life cycle, but more broadly, for example in scholarly communications. I gave a talk about a couple of years ago, I was talking about AI and knowledge infrastructure, and I thought I would have an academic audience that shared my critical concerns about the adoption of AI, and I discovered I had an audience who were evangelical about the adoption of AI and absolutely were coming up with new tools, new ways of publishing, so that whole process of scholarly communications about those who are writing, review of paper writing, review, all of that, with AI in the loop. And I was quite worried by that experience, and it's one of the reasons that I think the responsibility is partly with the tool builders of taking that into account. So, I haven't said that it's either fairness, transparency, accountability or something else, but I think it's responsibility for the tool builders.

Zeba Khanam: Yeah, so I think when I was thinking of this question and the approach which I have is more of a pyramid, and for me at top is why are we… the question of why tech for good? Because we live in a time where nation states do surveillance as well using AI, and that's something which we cannot run off. We have reports where AI has been used in wars as well. So, I guess it's very important why… what is the purpose? And it also helps you define my second point, which was accountability, and you touched upon that. So, I do talk to a lot of people in businesses as well, and I see two types of people. One who thinks AI is the solution, looking for the problem, so they want to use AI for everything. The answer is chatbot because they're so energised by using ChatGPT. And then there are people who are resistant. They're very happy with their spreadsheets and they don't want to use AI, so we're talking about big spectrum. Then who… but again, when we know why we are using AI, so we would know the accountability. And I think I feel like accountability is not just with tool builders, it's about… it's with people who are using AI. AI is your assistant. It's not the decision-maker because if you let AI and agents

take the decision, who is responsible for those decisions? And understanding that accountability is important.

And, again, transparency as well, but it's tough with large language models. Essentially, it's a black box that we are talking about, especially focusing on research methods of red papers, where you know the large language model was hallucinating. As such, there's no source for it. But because you let them make that decisions, you want to put that text out there. So, I guess being transparent about it as well. I feel like now ChatGPT should be one of the authors onto the paper as well, it's how… you need to let people know that it was part of the text which you have written. We also live in a world where entire population wants to write with hyphens suddenly, because that's how… and you can seriously… so I was reading a report recently, where I could even predict the prompt. So, the person said, "Here you go," and you can even… so the idea is that being transparent about it, that, "Yes, I used a large language model and it helped me make that decision, but it was not the main decision-maker". And lastly, being fair about it. So, this is how I see responsible AI. But, at the end of the day, I feel like accountability is with the individual. The tool builders essentially are big tech companies and their focus is not towards responsibility, unfortunately. It's probably these kind of conversations where we wish the focus for the big tech also becomes responsible AI. We also live in a world where there's AI psychosis, so ChatGPT has coerced people to take lives as well. So, it's dangerous times, so essentially understanding what's the role of AI, we cannot be ostrich, where we have buried our head deep down. We need to understand why we're using that kind of important… making a conscious decision in life, and I think that that's how I see responsible AI.

[0:10:24]

Mark Carrigan: I mean, the obvious sociological question is why people are using it in irresponsible ways, and there are issues of design and there are issues of context. I was relatively relaxed with the first generation of consumer-facing large language models compared to my critical sociology friends, who were

very pessimistic, and that was because it seemed to me that as much as the discourse of prompt engineering is misleading and unhelpful in some ways, it did point to the intellectual discipline involved in using them effectively. In order to get an output that served a particular purpose, you have to explain what it is you wanted it to do. And to explain what it is you wanted it to do, you had to be clear yourself. And so, for academics and students, I was, for a while, somewhat reassured that that would be a check on irresponsible use. Not a firm check, but a check. And something interesting has happened since then with successive generations of these models, particularly through the innovation in post-training rather than the actual models themselves and their underlying capabilities, they've become much more adept at inferring expectations from fragmented, unclear or otherwise inchoate statements by a user. And so you're now able to produce good-enough outputs without being clear about what you're asking. And in fact, if you try something like GPT 5 Pro, the £200-per-month Open AI subscription, what that can do with just a sentence is really quite remarkable. And that worries me because that makes irresponsible use a lot easier.

And we're beginning to see a similar process to the one that took place with social media, sometimes talked about as enshitification, Open AI announced today that they're introducing their shopping feature into ChatGPT. Meta this week announced that from December onwards they'll be using chat data to personalise adverts. And we're seeing a trend now where the LLMs are increasingly guiding the user. It's still in its very early stages, but I find this worrying. Meta are experimenting with a system where the Llama model will proactively reach out and nudge users. At the moment they're talking about a safeguard, so if it does it once and the user doesn't respond, it won't do it again. But I do think when you look at the political economy and the need to recruit the huge investment and the capital build-out that's taken place, we're going to see a commercial playbook quite like what happened with social media. Consider the difference between Twitter in 2010 and Elon Musk's X in 2025, and I worry we're going to see a similar trajectory for these commercial-facing large language models, with disastrous results potentially.

And so sociotechnically, I think responsibility is becoming inherently more difficult, but there's also the context in which this has emerged because at least in higher education, the sector is in crisis. There's a crisis of funding, there's a crisis of morale, there's a crisis of workload. And I think this is the most salient feature for understanding why academics would misuse these systems, because they… we feel pressured and stressed, and there's all sorts of competitive demands that are leading people to try and get a competitive advantage. And so, in a sense, I think there is a problem here, but it shines a mirror, it holds a mirror up to a system that was already struggling and cracking and riven with internal contradictions. Scholarly publishing was in a terrible state before the introduction of large language models, but I think the proliferation of AI slop in papers is possibly going to tip the publishing system into a kind of critical point from which there's no return.

[0:14:12]

And just finally, if we think about students, I've become increasingly frustrated with the tendency to penalise students and pathologise them as irresponsible, because when we think about the conditions many of our students are now facing, the data here is really clear. Compared to even just five years ago, the proportion of our students who are doing large amounts of paid work by necessity to survive through their degree. If we've designed degree programmes that presuppose students who have no other commitments, then yeah, of course, a student who's working 30 hours a week is going to seize on this to make that degree viable to them. You know, the number of students who are commuting to campuses because they can't afford to live near them. The number of students who have caring responsibilities is much higher than was the case five years ago. And so there are real problems. There are sociotechnical problems and there are sociological problems, but I think we have to be careful to see this context. And the language of responsibility makes me nervous. I can see the

importance of having this conversation, but it makes me nervous as a sociologist. There's a risk it can be individualising.

Wendy Hall: The trouble with AI is we've seen, and, God, at Southampton we study this stuff, all social media stuff, with… we've seen what's happened with social media and we're not learning those lessons now, and it's all moving too fast. And the companies have got to make a return on their investment. There is a lot of talk about a dotcom crash equivalent with AI, because it's very understood that it… it's very clear that companies are overvalued, they're not making money, and they've got to find ways to make money. And they are not regulated in any shape, sense or form. So, when, for example, OpenAI put out GPT 5, they can say about it whatever they want. I think of them as snake oil salesmen. Somebody called Nigel Farage that, didn't they? Was it the Prime Minister? Anyway, I think of Sam Altman, oh, brain fog, I think of Sam Altman as a snake oil salesman. And there is nothing regulated that's stopping them saying whatever they want to say and releasing whatever product they want to release, whether it's been tested or not, whether anyone's thought about the things. It's irrelevant to them. It's all about how am I going to make more money than the company down the road or whatever.

In China they are taking it more sensibly, I think. The way they're handling this has much more regard for the safety of society than we have. I'm too tired tonight to explain much about that, but I'm trying to write a… with a Chinese colleague, a paper about this at the moment. And it's quite a surprise. I mean, if you just look at the… what Trump said with the US AI action plan, which basically says no regulation. There's some sensible stuff when you get down to recommendation 20, it looks quite sensible, but at the top it's all MAGA and no regulation. That came out. A week later, the Chinese released their action plan, this was all back in July, it's all moving so fast, where they were talking about making things much more open, more transparent, being clearer about how the products are tested and evaluated, and actually working with the United Nations, which of course because I was

part of that, I'm supporting what the United Nations are trying to do, flawed organisation as it is, there's no other global convener in this space. But… so you had a complete flip, complete flip geopolitically from the US being the good guys and the Chinese being the bad guys. This is very detached, of course, from the sort of thing you're talking about, which is what's happening in the classrooms, in our front rooms, you know, with the kids on their smartphones. But it sort of sums up what… the context in which all this is being rolled out.

And the last thing that is just last… this week, last week, no, I think it was just yesterday, sorry, a senator in… of course, the government's just shut down in the US, I don't know how they do this stuff, but anyway, and a senator has put in something to the senate bill to say, "You can think of AI as a product, and if it harms you, sue people". It's all about take that approach. If it does… it's not about actually saying this is… I think we have to go to a… sorry, at a meta level, I don't mean Meta the company, I mean meta, we have to go to a much more assurance model, which is actually what our government's talking about. Peter Kyle's speech at Mansion House was brilliant on this. And we actually have to get the companies, when they produce a new product, to say how they've tested it, how they've evaluated it, and what… to declare the algorithms.

[0:19:48]

I just want to finish with one other thing and then… you talked about this term AI slop as just… it's only in the last few weeks, isn't it, people have started talking about that? Maybe it's longer, but it's a great… but my worry is a bit more than that you were describing. I think we are risking the integrity of the internet because it's not just scholarly publications. Any time anybody takes something from ChatGPT without checking it, thinking about it, puts it up on the internet, that's being used to train the future… so you are… we're completely polluting the internet with rubbish or slop. That's what we'll call it. And I really worry about this for future generations and, you know, we're

polluting the digital world in the same way as we've polluted the physical world, but it's happening much faster. And I don't see how you un-pollute it.

Mark Elliot: So, large language models have shown impressive capabilities for generating human-like text, but they raise concerns about misinformation, could also call that slop in the aggregate, bias, lack of transparency. How can we mitigate these risks without compromising their potential for innovation? So, I guess this is kind of, yeah, okay, that's a problem. How do we go about solving that problem? Do you want to start again, David?

David De Roure: Sure. This has segued really well and I want to pick up on a couple of the comments there, and I'm going to start very positively because I think if anyone can do this, this community can. So, we're really good with metadata, with provenance. We're really good at sharing models. We're really good at interpreting data. And in fact, this contrasts between my humanities world and my social sciences world because… and social media is an exception we might want to come back to, social sciences traditionally has been really good at collecting data and curating that data. We all know this well. And it isn't just accidents of preservation and discovery that you have in archaeology, it's something else. We go out and get the data. What I think we have to do, and this might be touching on technical solutions, Mark, we need to do lots of testing. I think what we need is some benchmarks actually, so when a new version of something comes out, we can run our benchmarks against it and see how it's going, see whether they still work, see what the biases are in this new model. So, that's something we can do constructively.

Reproducibility, at two levels. Computational reproducibility, can two people get the same results from doing the same experiment, is a really good way of sharing practice, sharing our code, sharing our prompts in order to do that. But there's that sort of deeper scientific method type of reproducibility, which is I do a thing here, you try a similar experiment there with your other infrastructure. Does it have the same result? We can do these things. The culture isn't always there to do them, which is interesting. Perhaps we can develop that. Provenance is something that technically we're quite good at.

Every time we have a chain of processing of data, we could be recording exactly where that's come from. Every time we use an AI model, there's a supply chain, there are standards, SBOM is used in the US if it's a software supply chain, there's an AI version of that, there's TAIBOM which is being developed here. Some of those technologies have been around for a while. They're not widely deployed. I think this goes to some of the points you were making. So, we've been able to do provenance with W3C prov standards for years. Do we do it? Is it in our tools? Not really. So, how do we get that adoption? So, I think we know the answer to this. Also, I think we should get on with it. There's nothing to stop us doing some of these things now. Build the benchmarks, let's have a project to build some benchmarks. We're spending a lot of time talking about it, I think we could do some doing.

[0:23:50]

Zeba Khanam: I'll pick up from some of the points which were discussed earlier, and it's essentially where do you draw the line? And I feel like especially in academic research, we publish for the sake of publishing, because we want to meet that KPI requirement, whereas innovation is something which requires time, so the idea of benchmarking. The other thing about large language model is that it's very non-deterministic in nature, so about reproducibility, that's a question. Essentially, understanding that if you're innovating it, why are we doing it? We're not running a rat race like big tech companies. Academic research has to be a space of creativity, right? And that comes with time. And I feel like it's different from industrial research because you don't have that pressure. So, building that culture where innovation is done, where you're kind of producing something or building something which is essentially required, you're adding something new, and that's where use of large language model would be more judicious in nature. Not just using it to write a paper, saying something which is required and is not a slop. I think that the idea is that, yes, we have 30 hours a week, it can help us augment it, but how can we reduce that pressure? How can we build that culture where we are innovating, which is required, essentially. And I think that's where I feel like I

hope academics evolves in a direction where it plays a role of innovating, which would have impact, essentially. And interestingly, it would not come from big tech. Big tech would not. Their focus is completely different so that's part of my thoughts. These are my thoughts and where can we draw the line and how we can build the culture for benchmarking, reproducibility and all, but essentially large language models, how can we use them in our day-to-day lives?

Mark Carrigan: Yeah, I'm in complete agreement with that. I mean, I've often thought about that as the kind of artisanal culture. So, when we can produce this material in…

Wendy Hall: Artisanal culture?

Mark Carrigan: Craft culture. Caring for what you build. I will self-censor, if you want me to.

Wendy Hall: No, that's alright. If I don't know, I'll ask.

Mark Carrigan: So, that culture of caring for what we're producing, so not just churning them out because we have a sense of that's what we need to do to remain competitive or that's what one does, but to invest and reflect on why we're doing what we're doing. And I think that kind of professional ethic is crucial. I don't think it's sufficient, but I think it's necessary. And how we could create spaces in which we can inculcate that, I mean there's existing professional resources and idea of methodological reflexivity, but again, in professional reflection on how we approach the decisions we make as researchers, better?

Wendy Hall: Yeah. Thank you.

Mark Carrigan: Great. And so going forward, I mean, we can create spaces in which those discussions are more likely. We can think about things like graduate education. But we clearly need more than that.

Mark Elliot: Right. We have one more question. I'll start with you, Mark, if you will, because it's a social science question. So, how is AI reshaping the methodological landscape in the social sciences, both in terms of capabilities and new risks, like black box models, data opacity, reproducibility, which Zeba has already mentioned. Have you got any thoughts about this?

Mark Carrigan: Yeah. I mean, I've seen some great examples, including at an event a few people here were at, at LSE in June, and some really creative work that's been done of building LLM-based analytics as part of funded research projects. And I came away from that event aware that they were all… they all had research software engineers working on the project and they were all using API access to design their own bespoke solutions, and that, I think, is really crucial to maintain control over the process because if we are going to see commercial software go through this kind of enshitification phase, I think the idea that we would see people using commercial large language models for research methods is terrifying. And anecdotally, people are doing this and they should be told to stop urgently for all sorts of reasons, not least of all ethics and data governance. But I think even off-the-shelf solutions really worry me, because the risk is that unless control is exercised over the process, qualitative research in particular, the kind of analytic encounter at the heart of it is going to be lost in a black box of prepackaged software. And this concern is not a new one. I remember when I was taught research for quantitative methods as an MA student and the tutor insisted on writing equations on a transparency because we weren't allowed to use SBSS until we'd learnt the basics. I never learnt the basics; I just learnt to use SBSS, so maybe he was right.

[0:29:11]

So, it's a continuation of a general theme that the more we use software in research, there's always a concern that we're being deskilled in the process. But this seems like a much more worrisome and more radical potential deskilling. And while there are creative projects that are using things like API access effectively, I worry what happens when these diffuse into real-world

context, because talking to colleagues and acquaintances in government and commercial social research, LLM-driven analytics have been taken up much more quickly in those sectors because there's more of an immediate need to demonstrate value for money effectively. And I think we have to be aware that that pressure is coming to higher education because once it becomes normalised that we can do qualitative research in particular a lot more quickly, funders, universities, colleagues are going to start expecting that we do that. And I find that really, really worrying because things will get lost. So, I think there are some… there are ways of doing this responsibly. But I think even if we are taking those responsible strategies, we have to imagine, if we're building something new, how might it be used three or four years down the line in a different context?

David De Roure: I think I'll just briefly go back to my music example. So, there are sort of… we've used AI to generate music that's indistinguishable from humans. That's like a mimic. We could have a debate about whether that's stochastic parrots, which is a phrase that people may be aware of. But then the way I'm using it now is very much to support creativity and innovation and disruption and experimentation. And, in a way, I think there's another sort of affordance in the digital world of play. So, it sounds like it runs contrary to some of the things we've said, but I want to encourage the freedom for people to experiment and to play with these things, and that would help us create new methods. And then we can't do that uncritically for all the reasons I've said. But, on the one hand, I think that what we're seeing is potential routine automation can make our research more robust, for example, and we can scale things up. But on the other hand, an affordance of AI is to facilitate some plays and experimentation, some new ideas, and then do it properly. But that's one of the affordances.

Zeba Khanam: I want to talk about the fact that when you talk to a technical audience whose research is in building AI, their focus is on accuracy and fine tuning layers and weights and parameters. But if you talk to a social science audience, that's where you can discuss policy, you can discuss about interaction, and I

think that's why I feel like it's so important for social science researchers to get involved in AI because of what we were alluding to about governance and citizens. And that kind of conversation can come only from social science background, so I feel like it's a good opportunity. Also, the idea of doing multidisciplinary research because, in the end, the silos I feel like can be broken by AI if we go in that direction. But we're in nascent stage, but hopeful that we can come together from a technical viewpoint and social science viewpoint.

Wendy Hall: It's a three-way stakeholdership. You've got the companies that are producing the technology, you've got the governments that have somehow got to tackle supporting the innovation, which the EU AI Act doesn't do so much, although well done for them for trying. But the supporting innovation whilst keeping citizens safe, and they're in a total dilemma about how to do that. We haven't sorted out with social media, and this is social media on steroids. So, to me, we are going to be in the same position with the climate change issues, right? Our future generations are going to be stuck with what we're creating today and they're going to have to sort it out in some way. And the best thing we can do, I think, is educate them, we've been talking a lot about that, educate them first of all not so much in being responsible, I think we should move to a not being irresponsible. I think we should talk about being irresponsible with AI and making sure everybody understands what it is. Sue Black, great friend of mine, great advocate of women in computing, she gives a super talk on how responsible AI is AI as we know it today plus education. I honestly believe that we have to have a global dialogue on this and we have to have, again, this is coming out of the UN, but it's very similar to what's happening in other areas, scientific panels, which will include social scientists, not just the physical scientists, that are discussing the science of AI.

[0:34:40]

Mark Carrigan: I was doing an internal presentation recently about exactly this and saying that no-one can adequately answer that question and anyone who currently

says that they can I think shouldn't be trusted because there's an assumption that there's a straightforward notion of their literacy that prepares students for a workplace saturated with AI systems, but that workplace even in five years' time I think will look very different from the contemporary workplace. And I think there's a lot of educational work that needs to be done to think about what that preparation is. The element that I think has to be a part of that is learning how to learn. I mean, it's a cliché, but the capacity to work with new systems, to understand and reflect on how to adopt them, use them individually and collectively, but that's a tiny part of the answer. I worry that there's a tendency to see preparing students as being a case of teaching them to use ChatGPT. I don't think that's the answer, not least of all because ChatGPT in five years' time will be very different to ChatGPT now.

I'd be very interested in finding out more about the Chinese answer to this. I mentioned earlier, I started this year's teaching earlier on this week and found a roomful of new students, many of whom are from China, who had all used LLMs throughout their undergraduate degree. And I was really struck when I asked how many of them had been formally taught this use as part of their degree. And so I think in the short terms…

Wendy Hall: What was the answer?

Mark Carrigan: Oh most of them.

Wendy Hall: Yeah.

Mark Carrigan: Yeah, most of them had been formally taught. And so on my to-do list for when I get back to Manchester is to try and find out about these curricula, because I really want to see what is being taught. And I think we can do things that prepare students in the short term, but it's very hard to change things rapidly in UK universities, and it's a curriculum that will have to change every year. I teach a postgraduate ed-tech programme and we have a constant struggle of persuading our faculty that actually we have to update everything every year. You know, we can't continue year in, year out. And so I think the

structures of the university make it hard to be sufficiently adaptive. The structure of academic workloads makes it hard to be sufficiently adaptive, but we can do short and medium-term preparation if we continually update it. But the long term? I have no idea and I'm very sceptical of people who say that they do have an idea of that long-term answer.

Wendy Hall: I wanted to share my experiences. I've just come back, I came back from China via Rome, okay? I went straight from Hong Kong to Rome for a two-day meeting of the ACM publications board. I'm co-chair of the publication… and ACM publishes the big… most of the computer science… a lot of the computer science papers, journals and conferences. And we have huge policies about what people… and we have to keep it updated on what people can and can't do in scholarly publications with AI. It is going to… I think you said it earlier, Mark, it's going to change the nature, or was it you, Zeba? It's going to change the nature of scholarly publications generally, and we have to prepare our students for that too. It is going to… but I think we're getting very pessimistic, which is what tends to happen. Actually, AI is going to change a lot of things for the better because the slog of research that… I mean, I was telling one of my PhD students this week about when I was doing my PhD and I had to go to the library to see anything, and we had to… if our library didn't have that journal, we had to send off a postcard to interlibrary loan and wait for weeks to be told whether that journal was available or not. And then if it was, it came in the post and you had it for three weeks or whatever. And that seems like lightyears ago that we were doing our research that way. And AI, here we are today, where AI is going to change everything about the way we do research. And in some ways it's going to be fantastic, because we will be able to trawl all the… do the… be much more thorough in the way that we do it. But what the AI can't do at the moment is the synthesis, and that's what you… at the moment, it is a data analysing predictive technology and we need to get our students to understand that that's what it is at the moment and how it works and how they can use it responsibly.

[0:39:20]

David De Roure: There's something we can do now, but it's going to be… it's not going to be the answer in the future. And it goes to the question about teaching. So, next week I'll be doing the first classes on our digital scholarship master's, and I will tell the students to use AI and I will tell them to reflect on the use of AI and to put a section at the end of the dissertation explaining how they've used AI, and we can do that now rather than saying, "Do not use AI," because I want to prepare them for a world with AI. Now I can be explicit. It can be very deliberate and we can, as scholars, reason about our use of AI, reflect on it and discuss it. In a certain period of time, and maybe it's only a year, then it's implicit use of AI. People are coming in already in something that's, quotes, AI natives, and all the tools have it in or not, or who knows, and then what do we do? But at the moment we can be explicit and reflective and use our critical skills.

Mark Carrigan: I completely agree. And that building into the software I think is really dangerous. Our university did the Copilot 365 trial recently and…

Wendy Hall: Yeah, I just got the email about that today.

Mark Carrigan: It's terrible software. I tried to find a single use and I couldn't. But the idea that this is being built into everything, so the burden becomes on you to avoid outsourcing cognitively to it is really dangerous. So, I mean, communicating that danger to students can be part of it. But very practically, I'm doing a lot of work at the moment about disciplinary norms around students using AI, because we have decades of educational theory about different sorts of learning activities, and when you look at the evidence and when you talk to students, the ways in which students use generative AI is really multifaceted. There's a kind of academic assumption that students are predominantly using it to completely outsource writing. And actually, the evidence is that that's a growing portion of use, but that's far from majority use. You know, the first example that I saw a student using in 2023 was asking what ethnographic meant. And I spoke to her and she said, "Yeah, I recognised that from

research methods, but I didn't understand how you used it there so I was asking it to contextualise it". And so I think we need to be better at distinguishing between the fine-grained practices that students are engaging within and to understand what we should steer them towards and what we should steer them away from. But the further I've got into that topic, the more I've realised that that looks very different in difference disciplines. For engineers, for computer scientists, in some ways there's a lot less flexibility about the kind of learning that has to take place, and I find that really challenging to get my head around as someone whose background is philosophy and social science.

Wendy Hall: So, in computer science it's going to change beyond belief as a discipline. Social science may stay the same. AI is going to completely, as Dave was alluding to earlier, we will have to completely change what we teach and how we teach it.

[0:42:24]

David De Roure: I think we will get to that point. There's some really interesting experiments already in using AI for every stage of the scientific process and in publishing that through the existing structure successfully. Equally, they're experimenting for using pieces of music like that, but if the definition is contribution to knowledge then sure, yeah, that's going to happen.

Wendy Hall: Because none of us have talked about… we haven't really tonight talked about sustainability and the energy use, and I think we should touch on that because it is part of responsibility. And yesterday, someone was doing a simple sum, something you could do on a calculator, on ChatGPT. The amount of electricity that uses. This is part of what we have to say to people, it's like… I mean even Sam Altman said, "Stop saying thanks to ChatGPT," because it just… it takes that as a prompt and everything whirrs up. I do think down the line actually, I think the datacentres could be a good thing. I think they… I mean just take the investment in the UK, I think they will lead to economic growth just because people building them will mean there's jobs,

but I do think we can think of them as big heat pumps. I think we are working on the technology where we can invert. They'll take a lot… they'll need electricity to gear up, but they'll create a lot of heat, and we can reverse that. So, you might want a datacentre in your backyard because you'll get free heat and light from it down the line. I think there are wonderful things. You know, we're very good at doing this as the human race of making the best of it. But I do really think at the moment we should be telling people not to use them to do simple things. It's just, A, lazy, and B, is helping destroy the planet.

[End of Transcript]