

AI and Future of Workforce: A Conversation with Alistair Stephenson

Transcript

Dr Caroline Green

Hello, and welcome to Accelerating AI Ethics, the podcast of the Institute for Ethics in AI at the University of Oxford. I'm Dr. Caroline Green, and in each episode, we explore bold ideas, innovative research, and creative responses to the ethical challenges posed by artificial intelligence. Today's guest is one of our Accelerator Fellowship Program Fellows, Alistair Stephenson. Alistair is Chief Strategy and Impact Officer at the National Domestic Workers Alliance in the United States, where he works at the intersection of labor, technology and political strategy. Tell us, tell me about yourself. What's your story?

Alistair Stephenson

Yeah, thank you for having me. And I am a movement practitioner in the United States. And so my career and my life has really been oriented around questions of social justice, both abroad. I grew up in Australia and came up in traditions around gender justice, immigrant justice, LGBTQ movements. And here in the United States, I worked mostly at the intersection of labor and technology. And my work today is really housed at the National Domestic Workers Alliance, or NDWA. And our mission is really to organize and build power for nannies, house cleaners, and home care workers across the United States. And a bit of background around the domestic worker population in the US. There's 2.2 million domestic workers in the US doing the work of caring for families and homes. And these are the nannies in the playgrounds, caring for children, the house cleaners that make homes livable, and home care workers that care for older adults and people with disabilities. And in the US, this workforce is overwhelmingly women, mostly immigrants and women of color. And it's really one of the most important sectors in the US economy and one that historically has been structurally excluded from a lot of the same rights and benefits that other low-wage workers are able to benefit from. And so because of that, NDWA was founded nearly 20 years ago to really drive forward the advocacy and the rights and lifting the floor for a lot of domestic workers in the United States.

Dr Caroline Green

Incredible work you're doing there because as you say, I think here in the UK and in many other countries, it's those people, those women, most of the time. working as

caregivers and so on who are really a hidden population, but now really affected too by AI and technology. So can you tell me a little bit more about what brought you to think about AI and technology in your work?

Alistair Stephenson

Yes, absolutely. So I think this particular fellowship at Oxford came about because At a time when there are a lot of urgent questions being asked around AI and labor, obviously, a lot of questions when it comes to low-wage workers that concern displacement, surveillance, algorithmic management, what happens to low-wage work fundamentally, and how communities will bear the risk of technological change. And we know that these questions are being asked all over the world. But we noticed that there were far fewer spaces being convened specifically around the intersection of AI and care work. And that gap really matters to us because we know that care is essential social infrastructure. And so the results of these questions are quite serious. And so we wanted to really start an inquiry around really typical ethical questions that might arise as AI becomes more concrete. So how do you protect human judgment? How do you avoid turning care into a set of tasks that should be mindlessly automated? How do you prevent technology from becoming a workforce shortage or an efficiency fix that puts the risk onto workers? And that really led to a lot of our collaborations, Caroline, and a lot of the alignment between the work being done in the UK and being led over there and the US because we think that more spaces where these intersections are arising need to mature as we start to find what some of these harms are and where they're being and where they are in the workforce.

Dr Caroline Green

So when coming to the population you're working with, how are you seeing, how are people using AI on the ground and what are some of the challenges, but also benefits you are seeing?

Alistair Stephenson

Yeah, so some of the challenges starting there were actually harder to detect and harder to find in the first instance. I think a lot of our brothers and sisters in organized labor and in unions across the US were starting to document and advocate against. real displacement threats. And so we have our colleagues and allies at the dock workers union, or we have those in manufacturing plants who are immediately having to deal with automated robots, with newer types of machinery, with AI mediating a lot of human decisions. And so the task and the mission in these organizations is to really organize, mobilize, and resist. That wasn't automatically true on day one with care work, where I think a lot of people struggle to get into the imagination of, well, how can a robot start to replace like the fundamentally human labor that comes with care? So it's like looking after children. It's, you know, we're starting to see a few of these Roombas

and the robots in the homes, but beyond a few examples, there hasn't really been, you know, this immediate displacement threat, at least for now. But what did start to emerge were more insidious and harder to detect interventions of AI across the workflow and workforces. And so I named surveillance being a big one and algorithmic management. So I can talk to those two briefly. On the surveillance side, we have always been an organization that has fought against surveillance in the homes. illegal surveillance with the use of ring cams or other hidden cameras and technology starting to surveil actual homes, which are domestic workers' workplaces, as well as the illegal collection of information and data, which we always want to fight against. And then with algorithmic management, We knew that increasingly with technological change, a lot of workers were starting to be connected with consumers through apps or through services that would match the preferences of a worker with the needs of a consumer or a client. except when you have an AI algorithm in the middle, sometimes they can do things like drive down wages based on invisible data that the worker hasn't actually disclosed. And so these are two examples of harms that we are actively fighting against and I think are shared concerns with a lot of other low-wage workers. On the opportunity side, though, I think these are questions that a lot of the sector and a lot of the progressive left are grappling with and are difficult questions. But where we do see potential opportunity is to potentially point the technology towards the direction of workers' rights and worker power. And so we are starting to develop our own technologies as one application of that, where a core problem that we've always struggled to solve has been how do we get core and basic information, powerful information about domestic workers' rights into the hands of workers in a way that they understand and can immediately apply? If you've worked in an organizing tradition or in a movement organization, that the bread and butter vehicle of something like this in the past has been to organize a ton of scripts so you can go knock doors, have a lot of conversations human to human, organize a lot of the information into policy documents or know your rights flyers. And this has been effective in a lot of ways, but really difficult to scale and personalize. With the opportunity around AI, we've built a coach within NDWA called Ask AIA, which allows our organizers to be able to put something and do some of this early legwork in with workers so that workers can come in and have a personalized interaction based on their sector, their background, their state, the rights available to them, their language, and not just access the information, but be able to have a conversation to make sure they really understand it and can potentially apply it.

Dr Caroline Green

Yeah, that's a really, really interesting example. Ask Aya also the way that you've gone about developing it, and we'll return to that in a minute. But first, I just want to say, and return to your point around surveillance, some of those issues we're seeing. So we work here with people, so home care workers quite a bit, who are supporting people and living in their own homes. visiting them for half an hour, an hour, a day or so, and

providing care and support. And so some of the people we've been working with here have also said to one of the issues with surveillance is that, it is a very personal individual thing caring for somebody. And some people will find it really important that there's like an element of humor or some, sort of communication that if others see and listen to that communication between the caregiver and the person who's receiving care, they would say, oh, that's a bit off. I mean, did they just say something here that I didn't find that funny? And so they would find it really difficult to be surveyed also for that reason, that they would then feel I can't actually personalize my care for somebody anymore. That it's like a real private relationship we're having here. But then it won't be anymore. It will be sent back to my employer and might misrepresent what I'm doing here, namely actually being relational with the person and adjusting my communication to the way that they need it. And so I found that really interesting. And then there's that other side where people are saying, employers are saying regulators and so on, advocacy groups who are pro surveillance, are saying, we've got a lot of abuse of vulnerable populations by those people who care for them. And so we need to safeguard them and the surveillance will help us. And we often see in the news these horrific videos of people in care homes being abused or a baby being abused by their nanny and that being caught by camera. So it sort of seems like really opposing views, but there must be a common ground here, right? I like a middle ground on how to ensure that, how to keep people safe and able to do their work. What do you think about that?

Alistair Stephenson

Yes, we've seen similar debates take place in the US and what we always return to and what our workers have really tried to imbue in our organizational mission and values is around this idea of like workers dignity. And I think when it comes to the question of AI application, we have organized around these beliefs that you cannot automate trust, you cannot automate judgment, you can't automate the relationship between a home care worker and a person they support or between a nanny and a child. And so in that way, we have tried to shift the debate from one around just displacement or replacement and towards dignity and to kind of avoid narrow frames or maybe false choices between two extremes. I think at the end of the day, there are going to be a lot of nuanced debates like this. And it's important for us that really we ground these debates in worker insight and worker stories, because at the end of the day, we know that the drift or the bias in these debates is usually against workers and to the extent that surveillance carries a real, maybe risk on both sides. It can also create fear for workers, around ordinary judgment calls. It can weaken reporting. It can make workers feel like they have to choose between keeping a job or protecting their basic privacy. And so I think, as power shifts with all of these technological changes, we really try and center around this idea of like work at dignity and making sure that they're really leading the debate.

Dr Caroline Green

Absolutely. And AI is really exacerbating that issue, right? Because people aren't just surveyed, it's also what they're saying is being recorded and then somehow being transcribed and a large language model, a language model will then create notes and send that to the employer. So how do you actually know that the notes are capturing the complexity of the context? I mean, that's a real issue, isn't it?

Alistair Stephenson

That's right, yeah. And we're not just a believer in I guess like the goodness and the power of workers, part of the care ecosystem is making sure that we really support good employers, And I think when you take the employer and domestic worker relationship together, it is really about the primacy of the trust between them and making sure that the relationship comes first and that the rights are the floor, but there are all types of different axes for how to manage a relationship responsibly and to have something that's like long-term and trusting. And I think this question goes to that.

Dr Caroline Green

Yeah, so before we go on to the project that you'll be working on as part of your fellowship here, because that directly links to what we've just been talking about, I'd like to hear more about Ask Aya because so in March we had the AI and Social Care Summit here in Oxford and we had your colleagues come and attend that conference and present how Ask Aya came into being. And people here in the UK, those people who were part of that conference were incredibly inspired by how that worked. So please explain a little bit and how did Ask Aya come about? What was some of the kind of process and ideas behind it?

Alistair Stephenson

Absolutely. So yeah, I mentioned before that the conditions around AI and impacting domestic work weren't immediately obvious, you know, three or four years ago when organized labor were starting to ask these questions. And so we actually started this process with a learning tour. We took 10 or 12 of our domestic worker leaders that were elected in our councils to the Silicon Valley. And with this cohort, we met with technologists at Google and Anthropic. We met with ethicists. We met with academics and policymakers. And together started to ask basic questions around where are these harms taking place? where are the opportunities, what's on the horizon? And we're able to leave the journey with an agenda that was like our first AI workers agenda that we were able to then ratify with our big national assembly of 10s of thousands of domestic workers. And also take forward a few territories for exploration. And what we wanted to do was make sure that we were fighting and advocating on multiple fronts and not just assuming that one lever would be the best lever for power. And so to that end, we

ended up with a framework that was a four-part framework. On the one hand, we were talking about bargaining first. And so domestic workers don't get too belong to a lot of collective bargaining or union contexts. But we thought it was really important that we joined in our colleagues with the rest of labor in collective bargaining arrangements. And a lot of the early work to protect workers' rights was in union context. And so one pillar for us was just supporting bargaining efforts that were happening to the left and right of us. Downstream of that, we had a pillar around advocate. And so we needed to make sure that we were being organized around the basic research of what was happening and what can be resolved in policy. And so some of the harms I've spoken about earlier, the risks around surveillance, the risks around algorithms, how can we support and help to design basic legislation across the states tackling preemption and all of these big issues with a lot of our colleagues and in coalition with a lot of other bodies. The third pillar was really called, it was really around redistribution. And so this is really a pillar that starts to ask questions around the wealth being generated from AI and the fact that we don't yet have big tax programs, schemes or instruments to direct that wealth back into the economy. We know that there is a lot of concentration of wealth and a lot of concentration of power happening in Silicon Valley. And what would it mean if some of that started to funnel and translate back into the social programs that will really support a care economy? There are a lot of folks in the US who are asking big questions around tax and redistribution, and we wanted to be part of that to make sure that the basic social safety net remained in view and a priority for those who are doing this work. And then finally, we had a pillar called Build. And Build is really where a lot of Ask AIA has come to fruition. We are trying to ask a question, what if AI is here to stay, how could we make sure that it is pointed in the direction of workers and not in the direction of further extraction or exploitation of workers as a result of AI? And this has led us into a lot of important discussions around ethical AI and what have you. And so With all of this, we embarked on a design project and a massive consultation project with our workers, engaging over 1000 of them to design and develop the first AI tool for domestic workers. And what resulted was Ask Aya. And it really helps workers prepare for common but high stakes situations. And this was the goal. So how can we use AI to help workers prepare for asking for a raise, setting boundaries with employers, documenting an agreement, understanding local labor standards or rights, drafting messages to employers or having difficult conversations. And so in that way, we have used the the application of a chatbot and what people are used to in an information exchange and actually customized and personalized a domestic worker experience on top of it. is built on top of Anthropic's Claude and we have put guardrails and a lot of knowledge into this product to make sure that we can provide workers like a really personalized and rich interactive experience that supports them.

Dr Caroline Green

Incredible. So it's really something that very much was developed by the workers for the workers. And how do you continue to develop it now? Does it, do you have feedback loops and do you continue your consultation or how does that work?

Alistair Stephenson

Yeah, so it's been so important for us to have designed the original conception of Aya with all of these workers. And from the beginning, because This is kind of the air we breathe at NDWA. We were so intentional about setting up governance structures at all levels as we got into this technological development. And so we have a working group at the board level. Our board is comprised almost entirely of domestic worker leaders from across the country. They really steward our AI governance, our principles and our policy. We have working groups at the product level. And so they are really informing the design and some of the core decisions around how IR looks and feels and reacts. And then we have a working group of practitioners that include domestic workers, that include our staff who are like product and knowledge and policy experts that are doing human in the loop review on a regular basis. and that go over all of the anonymized data to make sure that it is really accurate and supportive and centering the dignity of workers. And so these are a lot of structures we set up from the beginning for ongoing monitoring and design.

Dr Caroline Green

Yeah, that sounds like, so do you think that could be a model you would also want to see translated more widely into AI development of products?

Alistair Stephenson

Absolutely, yes. I think what we're seeing with all of this upfront rigor when it comes to not just product, but the governance and the structures itself is actually netting out in the impact. We have run a lot of tests that put IR side by side with a traditional out-of-the-box AI model. And we're seeing some really striking early pilot results. There's one particular set that I wanted to share with you, which is across a six week beta test with a variety of our domestic workers, 93% of them reported applying Aya's advice in real workplace situations, and 25% of them reported negotiating pay increases as a result of their interaction. And those numbers were markedly higher than the ordinary interaction or a conversation with a ChatGPT or a Claude. And we think that this comes down to a lot of the responsible design and decisions and knowledge that we've imbued into AIA. But really what was striking and what took us by surprise was less the reports of the accuracy and helpfulness of the rational information. So like I'm entitled to \$15 an hour in New Jersey, for example. and more that people felt encouraged and validated in the dignity of their work. And these are insights about the tone of voice of Aya. These are insights about like the way that Aya responds or reacts and validates in a conversation. And that might feel like second nature to a lot of us in an affirmative sort

of AI conversation. But there are subtle things that we're doing in the language to make sure that we're using the same verbiage as a worker, that we are making sure that we are matching the conditions and really understanding all sides of the work equation. And that's really important because in an environment where domestic workers just don't get validated for anything at all and are so often like invisibilized and marginalized in this economy. These interactions are very meaningful.

Dr Caroline Green

So it strikes me, you know, what's really interesting is that you're not only supporting your workers there. through that information, empowering them, but also it's about connection. It's like a community of people really behind this. Not only have you built it in community, but it's also people drawing on community through that AI system, which I think is really powerful. And what also strikes me about it is that, obviously the purpose of it is to support people in their rights, in their dignity, in their working conditions, rather than it being profit driven. That's what most of the AI tools and companies we have out there, it's about scaling, it's about making profit. no matter what, really. And so this is such a different approach. But I'd still love to ask you how, in terms of financial sustainability, because it does cost something, right? Also to develop it to everything, how do you do that? How do you keep it financially sustainable?

Alistair Stephenson

Yeah, so I mean, NDWA is a nonprofit to begin with. And so we make sure that we source revenue in all sorts of ways that are diversified. We have support from institutions, from individuals, from grassroots fundraising. And that mix has always supported the different technological experiments or innovations that we've had. But you're right that scaling it is a different issue and it takes real resources to house a product and technology team within an organizing institution. Our hope though is that IR will always be free for domestic workers and come with membership and membership is also free. And so we just want to make sure that we only scale to the degree that we can do it responsibly in line with our values, in line with our guardrails. and just stay focused on this North Star, which is can tools like this really work to shift power and less with questions around, as you've said, will this eventually be the thing that turns us a profit or helps us with revenue? Really, we're so focused on the domestic worker experience.

Dr Caroline Green

Yeah, it's really nice. It's a really amazing example of how you're using the power of AI to benefit this group of people. which takes me now on to your Accelerator Fellowship project. Let's talk about that a little bit. What are you hoping to achieve while you're part of the fellowship? And by an active part, you'll always be a part of the fellowship now. But yeah, what's your project all about?

Alistair Stephenson

Yeah, so the kind of backdrop for a lot of these questions for us was How do we enter into this space around ethical AI, but really make sure that we are setting up the right structures and the right vehicles to retain and strengthen worker voice throughout? There are not a lot of examples yet in a sort of like low maturity field of worker AI being used to support workers inside or outside of the workplace. And that's true across the world. There are more things happening. There are more experiments that are growing. But a lot of this work is being developed for the first time in pilots that are happening right now. And I think what's such a great opportunity with some of this inquiry is not just making sure that we ship these products and get them out the door, but that in doing so, we create models that can be then shared across the sector, shared with colleagues in different countries and in different markets, and therefore can provide comparison and experiment for people around how much and to what degree worker voice was preserved or got us to a tipping point around worker rights. And to that end, the fellowship is really a way for us to ground some of these questions in a live example, which is Ask Aya. and then to be able to develop really strong frameworks around it that can be used by others. So we're doing a lot of this work to surround Aya with a lot of important work perspective. And I think there is a lot of like important research already existing that says, you know, it's really important to do that. It's really important to give workers the seat at the table when it comes to technology development and to, I guess, enshrine work of voice in a lot of the advances happening around things like AI. But I think what's not clear is beyond the level of principles and beyond like these first efforts to write down values alongside an AI, how can we actually govern a tool like this ongoing and day-to-day? I think there is a level of, I don't want to be too cynical when I say this, but like symbolic principle setting when it comes to workers and the ethics around AI. But I think what is more difficult is as a product matures and as more and more people use it and those incentives maybe come into competition with each other, how do we make sure we're actually preserving power among workers who are not necessarily in the day-to-day of the product development and aren't the driving force between the AI like core technological advances itself? And so for us, this is really the center of the project. How do we create, how do we build on the level of these principles that go towards fairness, transparency, accountability, and human rights, and then start to tackle these harder questions around who decides how systems change on an ongoing basis? Who sees the harms? Who has authority to intervene? Who controls the data? And how can we try and answer these questions in structure and in framework and in policy so that we can have a roadmap or a model to share with others as this work develops. And a final thing I'll just say here is that we're not just doing this because as a worker organization, it's morally incumbent on us to figure out these questions. We also think it is strategic. If we can figure out how to answer some of these questions and create like true like worker stacks of governance product

frameworks and put it all together and we do it enough times in the US context or abroad, there are enough then signals in the market that this is like a validated approach to building AI that maybe sits differently from the way corporations are doing it and those signals in the market are so important because that is what policymakers look to, say there's a different way of doing things and this is how we're going to regulate and this is how you have to have worker voice in the mix. And so we're trying to rush these questions towards us now as regulators kind of catch up in the United States and start to set new norms and standards.

Dr Caroline Green

And I think also as a tool of empowerment of workers, right, to say, well, actually, look, we've got a way here that we want things to work, how we can bring our voice in here. Because I really, I think that point you made, we can start. co-designing or being some kind of co-work on how we integrate people into building AI systems. So we can start with good intentions and we can have lots of voices in the beginning, maybe even in a little bit further down the road too, we're still including people, we're still listening to them and so on. But as then these systems scale, I think that point you made, how do we continue to ensure that there is that real co-production, I would say, in the workplace. And I think that that's really, really important. And so we're so excited to work with you on this as we are grappling with these questions too. And I'm also really excited to be working with you for another reason, namely to shed some light on this group, incredibly large group of people who are doing such important work, but who are often so hidden. and so affected though by how AI and tech is influencing their work life. People are often so concerned that AI is actually going to take over their jobs. And it's something that I must admit, I was first surprised when I was listening to the care workers we're working with now who are working in care homes and so on, that they said, oh, well, I'm worried that I'm not going to have a job in five years down the line. But we've got a huge scarcity of care workers here in this country, as in many other countries. So I said, oh, I was wondering about that, of why people would be worried that they not have a job with so much demand. But I can see now where they're coming from because people are designing for picking over. And so what would you say as sort of like a roundup of this podcast, what would you say to people who are working in care when it comes to being concerned, but also hopeful towards AI and tech?

Alistair Stephenson

I would say that these contradictions are being felt among domestic workers all across the world, amongst all workers. And I think that it's okay for us to like sit in both hope and ambition as well as rightful cynicism and asking challenging questions. I think what we know to be true is that people have this real fear and workers are already using AI as well as employers. You know, there are workers who are using it in their daily lives. And often, you know, there is a lot of emerging research around like what the impact is of

the daily use of AI among women. And this is like understudied, like how women are using AI to manage work, family life, medical questions, and certainly their own personal care questions around the daily administration of life under austerity, you know, in conditions where it just isn't a lot of help to be accessed. And I think that as we grapple with some of these questions, we can't just bifurcate this debate into threat opportunity, good and bad. There are a lot of very important nuances. And so it's incumbent on us as researchers, activists, organizers to view AI in both ways and potentially to use the best of it to build worker power while we figure out the rest.

Dr Caroline Green

Wonderful. Thank you so much, Alistair. It was a really great conversation and inspiring conversation. So thanks for that.

Alistair Stephenson

Yes, thank you so much for having me. It's been an honor.

Dr Caroline Green

This has been the Accelerating AI Ethics podcast at the Institute for Ethics and AI at the University of Oxford. If you have enjoyed this episode, please subscribe and share. Until next time, thank you for listening.