The National Academies of
Sciences, Engineering, and Medicine

Roundtable on:

Accelerating Climate Progress with AI: From Science to Action Workshop

January 13, 2026

Proceedings by:

CASET Associates, Ltd.
caset@caset.net

TABLE OF CONTENTS

P R O C E E D I N G S

DR. MELVIN: Thank you so much for joining us today for this workshop, our workshop, Accelerating Climate Progress with AI: From Science to Action. I am April Melvin. I am the staff lead with the National Academies that have been working with the Planning Committee to develop this workshop. I am thrilled to see you in person and online today.

Just a few housekeeping items before we get going. We ask that you are partners with us in fostering a positive and constructive environment for this meeting. The institution and the Planning Committee are committed to fostering a professional, respectful, and inclusive environment where all participants feel safe and welcome to participate in an atmosphere that is free of harassment and discrimination. If you see or experience something that makes you uncomfortable, if you are in person, please feel free to approach me or you can reach directly to our HR Department to report any concerns you may have.

To engage in the conversation today, we have folks that will be participating remotely as well as a number of folks that are in the room here with us. We encourage you to log into Slido following this QR code and use this as a platform for asking questions, leaving comments, up-voting questions or comments that others make, and also responding to a variety of interactive polls that we will be sharing throughout the next two days.

For those that are in-person participants, we will also be providing a microphone for you to ask questions in real time during the discussion question. You can feel free to participate in either way. And if you are in person, we also ask that you share your name and affiliation when you speak.

With that, I would like to hand it over to Steve Sain, who is the chair of the Planning Committee for this activity.

(Applause)

Agenda Item: Welcome and Setting the Stage: AI in Climate Research and Action

DR. SAIN: Welcome. Thank you for being here today, both those that are here in person and those of you that are online.

This workshop is one of many different work streams of the roundtable for AI and climate change that the National Academies has. There are a number of members of the roundtable that are here in attendance. If you are looking for more information, there is a QR code, the website, or you can grab some information from the front desk.

I would like to first off by thanking both the staff who helped to put this together. There is absolutely no way we could have done this without their incredible ability to manage change and put everything together. But I would like to thank the committee for all their hard work, lots of meetings over the last couple of months, to put together what I think is going to be a really outstanding workshop.

This workshop came out of conversations within the roundtable, thinking about how AI is really starting to have a dramatic impact on climate science and climate modeling. But what we wanted to do is take a step further and see how AI and AI-derived products are coming out of climate modeling and having impacts beyond, whether we are actually talking about impacts research or we are moving onto decision making and policy decisions. And that really has been the start of the focus of driving the workshop. But as the workshop even evolved, we started thinking more about things like a broader view of what climate action actually is.

With that in mind, we put together hopefully a workshop with lots of different people from lots of different backgrounds, from the technical side to people who are actually using AI as part of their day job.

Day 1 of today’s workshop will focus mostly on trying to get into the details, into the weeds, if you will, on a number of different key areas. This is not by any stretch all possible applications of AI and climate action but at least a handful that are going to hopefully set up some really nice conversations about what is going on with AI and climate actions and also setting up Day 2 where we will likely move forward into thinking about how these things can impact more broadly and how we move forward from here.

We are going to try out these Slido questions. We are going to hold up for a second. The online folks I guess are not online yet.

This is the first question in the Day 1 poll. I loved the last answer, my boss made me about why you are attending today. Let us go on.

Part of these questions were, one, to get us used to using Slido and using these polls. We are going to have a number of these polls over the course of the next couple of days as another mode of audience participation.

This question, the one that is one the screen, is about your primary affiliation. We just want to know the make-up of the audience. The first question on primary affiliation -- probably not too much of a surprise, academia, faculty and staff. It comes in at about a little over 40 percent, students at about 10 or 11 percent, private, 18, public, 14, nonprofit, 12. It is a solid mix across a number of different sectors.

The second question has to do with how familiar you are with the use of artificial intelligence in climate science or decision-making tools for climate action. Anybody want to guess? No. Roughly half answered I have some basic knowledge. About 30 percent, I have a solid understanding on one or both of these topics. And then the rest are split between an expert who uses the tools regularly and new topic. Again, a nice mix across there as well.

The last question was about one to two words -- when you hear the terms AI and climate together, what comes to mind? I am actually really happy with what is the dominate word there. Again, audience participation. Anybody want to guess? Opportunity. Nice. Do you see this? Yes, but I am really please to see that opportunity comes up. There are some -- it is a great mix. There are some like scary change, negative impacts, but there is also a lot of positivity there. That is what we are hoping to try to get through today.

Anybody have any questions about the process here or the roundtable or the committee or the organization of the workshop? We can at least fill up some time that way. The roundtable is a standing function within the National Academies. It has a number of different workstreams. There have been other workshops. There is this workshop and there are a number of different other activities. Ann, over here, is the head of the roundtable, chair of the roundtable. Chair.

Without further ado, I would like to introduce our first keynote address, David Rolnick, who is the assistant professor and a Canada CIFAR AI Chair from McGill University. David.

Agenda Item: Keynote Presentation: Artificial Intelligence and Climate Change: Opportunities, Challenges, and Dangers

DR. ROLNICK: Hello everybody. Can you hear me okay? Great. I am hearing some yesses. Can you see my slides? Great. I cannot see anybody or really hear anybody so I am just going to keep on talking into this void. It is great to be there with all of you virtually. It would not be a conference on tech if it did not have some tech issues right at the start.

I am going to be talking overview about the space of AI and climate change, which is the focus of my own work as a computer science professor. To start out, I want to give a really big picture sense of where AI can be thought of, dialing into the space of climate action. The space of climate action is really big and touched on so many different areas of climate science but also climate change adaptation and mitigation from areas involving earth sciences, involving heavy industry, buildings and cities, agriculture, forestry, and other land use, transportation, electricity systems, and a whole range of different are as involved in resilience, in ecosystem modeling, and ecosystem services. Lots of different areas. And AI can be relevant in all of these cases. It is not the answer in any one of them.

It can play specific roles in helping facilitate existing strategy for climate action across all of these different areas and different sectors. I am not going to go through all of those. But if you are interested in checking out a full overview of lots of different areas, definitely all, but many within this rich panoply of different applications, you can check out our overview paper that we wrote looking into those sectors and more.

But I do want to touch on some overarching opportunities for AI that we see across all of these different areas because this can help us think about where AI can be relevant, the kind of role that AI can play in many different contexts.

One theme that we see again and again is distilling raw data into actionable information. Taking information that is coming from some big and unstructured data gathering procedure and turning that into something that can then be usable by decision makers or by scientists, by other people.

For example, this is data coming from Climate TRACE, which is a coalition of for profit and nonprofit entities that monitor the world's greenhouse gas emissions and make that data actionable for policymaking and for meaningful change. Some of that is using AI machine learning with remote sensing data in order to get real-time info on where missions are happening, when, and what we can do about it. I encourage you to check out data from Climate TRACE. Similarly, disaster response, understanding where is flooding happening in real time so we can respond to it. This is something that is already being done by the UN and other agencies and organizations.

The second theme I want to touch on is improving operational efficiency. This is often relevant in the context of mitigation where we can think in terms of complicated, automated systems. They are already automated. But we can make them run more efficiently, use less energy, use less raw materials. Heating and cooling systems used in buildings. HVAC systems are a great example.

Also, industrial processes and factories that are manufacturing, for example, steel and cement, which together contribute about 15 percent of global greenhouse gas emissions. AI is already starting to be used in all these different contexts, everywhere from the smart thermostat in your home to the much more challenging problem of optimizing factories to use less energy and raw materials.

Forecasting. Predictions of what is going to happen. Now, that can be really useful across lots of different settings but I particularly want to touch on the examples of nowcastings or predicting in very short timeframes. Electricity supply and demand so how much electricity can be produced if you are looking at solar and wind power. These things vary from moment to moment. You need to understand how much is available to meet demand. You need to understand demand as well. This requires often very accurate and very fine-grained weather forecasts and then also predictions of how much electricity people are going to use. You do not end up not meeting demand or overproducing, which can often mean that you have wasted often fossil fuels in, for example, spinning reserve and thereby produce more greenhouse gas emissions.

And then the final theme I want to touch on is in speeding up simulations. Oftentimes we have simulations that we need to run for various different climate relevant problems such as extreme weather modeling or long-term climate modeling. We would like to have these processes run faster because actually modeling climate and weather takes a really long time. Climate models can take months or even a year to run because there is just so much physics involved in a simulation. That is accurate but sometimes AI can help speed these processes up. It can help step in and approximate some pieces of that model or even the entire model in certain cases, especially for shorter-term predictions like for weather rather than for long-term climate.

We are seeing increasingly in AI stepping in and emulating the physics-based simulations or downscaling, which is taking data that comes out of a simulation or comes out of real time sensors and increasing the resolution of that data so making a grainy picture suddenly high resolution, very sharp. That can be really useful in improving the accessibility of use of time-intensive simulations for meaningful action based upon climate and weather data.

These are some opportunities from AI at a really high level, thinking about some of the ways in which AI can be relevant. These are just some examples.

I want to dive into a couple of more examples from my own group’s work in more detail. But I also want to touch on something that was hinted at I think in some of those responses to the initial survey, which is another side. I often hear and I am very glad to be hearing healthy skepticism about AI. AI can help solve climate change but it uses a lot of energy. I have added a question mark here because I do not think this is wrong but I think it is missing part of a point here.

AI is using vast amounts of energy in lots of ways now. This is a major problem. We are seeing the growth of these exponential curves. This is a prediction. Taken with a grain of salt and there are huge variants between the different predictions. But it is expected that AI is going to mammothly consume lots of energy moving into the future and maybe more importantly, it is already leading to more energy consumption right now. It is leading to a vast production of data centers. It is leading to coal-fired power plants being built or not closed down to meet energy demand right now. This is a huge problem. The amount of AI energy usage forecasted in the future is in the order of large countries. This is a major problem.

But it is not the flipside of the coin that we were looking at in the previous slides. It is not like AI is helping climate change but it is also hurting because fundamentally when we say AI, we mean actually a lot of different things. The stuff that is helping is not necessarily the same as the stuff that is hurting. And the motivations for these things can be very different.

Think of AI like a menagerie. If you can think of a menagerie of different animals, you can think of a menagerie of different AI algorithms. You can think of, not to push my metaphor too much here -- I am going to see how far it goes. You can think of elephants. You can think of bees. You can think of a lot of other animals obviously. But let us imagine a menagerie of different AI algorithms like some elephants and some bees. On the elephant side, let us think about algorithms like Gemini, ChatGPT, DeepSeek, and Midjourney. A lot of the algorithms within the past few years have been very much in the news.

On the bees’ side, we are going to be looking at algorithms that are also actually in the news but not necessarily as flashy, remote sensing, AI for optimization and control like those factories, sensing the floods, time series analysis like those predictions of electricity consumption, physics-informed machine learning. These are kinds of the AI that are being widely reported on and integrated across society, just not necessarily as much in the public mind.

Fundamentally, there are some differences here. First of all, you can think about the elephants obviously being bigger. Massive energy, often consolidated infrastructure. Those big data centers. Whereas the bees might be scalable, distributed use. They might be done on laptops or phones. They might not be requiring the same level of infrastructure or even the same kind of infrastructure.

But then maybe more fundamentally -- again, I am a computer scientist. These are technically different. Now, they are both AI. They are both machine learning. They are both even deep learning. They are neural networks. They have all these commonalities but fundamentally they are trying to do different things and they are achieving the goals differently. Different goals achieve differently. These are large-scale generative AI algorithms whereas these are classification optimization and prediction algorithms. You might say if they are both deep learning, if they are both neural networks, why does it matter? Well, because they are trying to do different things. If you are trying to generate text, that is different, or generate images, that is different than if you are trying to determine whether this image has a flaw. It turns out the algorithms you need are very different in those different circumstances.

I want to also highlight a couple of other things here. Firstly, the way in which one designs these massive scale algorithms is often different from the way when designs the distributive scalable algorithms. These are often top-down designs because they are designed to do one thing -- sorry. They are designed to be one algorithm that by virtually it being big and consolidating, it is trying to do lots of things. But those things are all generative AI things. They are all like generating the answer to a question or generating images or movies whereas these are much more trying to do this specific thing particularly well. They are also trying to do different things. They are trying to answer somebody's problem when it comes to the problem they are facing with how do we forecast the availability of electricity or how do we optimize our factory. This is a very user-centered problem. If it is done right, it is collaborative. The design really focuses on the problem.

The kicker is that these kinds of generative AI algorithms are actually only seldom relevant in those climate-relevant areas that we mentioned before, that entire big map. There are some uses for these elephants. We are going to see some, including in the next talk, that are great uses for elephants, large-scale generative AI algorithms. But most of the uses that you will find for AI in climate action are these bees. They are widely needed. They are used across sectors and they are just a different kind of AI fundamentally.

This is the contrast I want to draw. I want to see by looking at the menagerie of the algorithms, AI use cases across society. What we would like to see is some elephants and lots of bees. But oftentimes nowadays because the elephants have become really flashy, we are seeing everybody assuming that we need large-scale generative AI, LLMs for everything. Everybody needs a chatbot. And people are thinking that the hard thing is you really want to have chatbots for everything. Chatbots will solve all of our problems. The chatbot that answers how we should optimize our factory(?), the chatbot that answers where the floods are going to be. That is not right. The hard problem is actually flood predictions or forecasting how you should control this piece of machinery in your factory. The elephants are not just overkill for such problems. They are the wrong tool. They will not give you the right answer. This is like a farmer trying to pollinate their crops with elephants. You need fundamentally bees.

This is why I feel like the problem of such a lot of energy consumption, which is all driven by the elephants now. We have seen in the past couple of years is everybody wants to build chatbots for everything. That is focusing, to some extent, on the wrong thing. We do not need elephants for everything.

We also do not need to think about just size by itself. People often say we will build smaller LLMs. We will use smaller LLMs for everything. This is also a problem. Smaller elephants. Again, they are still elephants. Fundamentally, not the same thing as bees.

People are increasingly suggesting we can have elephants but we can build problem-specific elephants. No. We do not want generative AI for everything. We want generative AI for some things. This is where, I think, we should be moving towards a few elephants and lots and lots of bees.

And then also, a lot of situations where we do not need AI at all. Fundamentally, sometimes when people turn to AI now, they just need search or they just need a good user interface or they just need some software engineering. Really thinking in terms of a menagerie of AI and other engineering tools is really useful.

Let us see some examples. This is some work being done in my own group, motivated by the problem of agricultural remote sensing where we are trying to take images like this and create maps to look like this for governments around the world that are trying to map agriculture and understand how to step in as climate change affects productivity.

Let us compare different approaches. ChatGPT, DeepSeek, and so on, standard vision transformers, which are a classic approach in computer vision and image processing, and then our approach, Galileo. Let us look at a number of parameters, which is a proxy for the size of the model here. ChatGPT and all the others are about a trillion parameters. Mammoth. These are elephants because they are large-scale generative AI. Both vision transformers and Galileo are not this kind of large-scale generative AI so they are both bees. Standard vision transformers are about 100 million parameters. Galileo can be effective. They are as small as a million parameters. This is again both bees, a million times smaller than

ChatGPT.

And the kicker here is ChatGPT and DeepSeek are not going to help. They are not going to do the job here. Standard vision transformers will help somewhat. But what you really want is an algorithm that was designed for the problem at hand. Galileo is a self-supervised transformer that is designed specifically for remote sensing. I will get to some technical details very quickly on the next slide. And actually, in addition to higher accuracy and being smaller, it also needs much less labeled data. Really by thinking about a specialized tool and by not making it be a large-scale generative AI tool, which is not the kind of problem to be solved here, we really end up with something that is very useful.

What is going on here is self-supervised. We are trying to deal with the fact that there is not very much labeled data available for crops around the world and often it is biased in terms of different geographies. We are trying to build in all of these different remote-sensing sensors coming from different satellites. We are trying to use them as labels. We basically remove part of the data and then ask a transformer-based neural network to reconstruct the rest of the data. And this allows us to get a meaningful data representation that can then be used for different downstream tasks with very minimal labeled data actually being used by the user.

This final classifier could just be a random forest. They might take this final vector, which might even just be 100 or 1000 dimensional vector and throw it into their existing pipeline. Random forest, linear aggression. It can really help because all of the processing has already been done under the hood.

And we find that our algorithms actually outperform lots of all of the prior algorithms that we have seen while being also smaller here. We can see that the accuracy is higher than everything else and the computational cost is lower.

I should note that everything we are comparing against is still orders of magnitude lower than something like DeepSeek or ChatGPT, which is all the way over here off the slide. It is so big.

This is an example. But I really want to touch on a couple of things that we are seeing here. First, the frontiers of innovation that are needed for making this kind of progress. I want to think about these different kinds of ways that we need AI to be built. These are the same kinds of things that you sometimes see in machine learning and AI conferences but really brought the fore by the particular needs of users, things like out-of-distribution generalization, uncertainty quantification, explainability and interpretability, physical constraint satisfaction, multi-modal data, limited label availability, and causality. These are all problems that we tackle on a day-to-day basis in building machine learning algorithms, driven by the needs of users and really thinking about impactful innovations means, thinking about user needs, and specific applications even though sometimes when you build for a single user needs, you end up with something that is usable for lots of different users, which is what we had with Galileo. It was originally intended for agriculture and now we are using it to monitor the water stored in snowpacks in the Alps and in the Rocky Mountains because it ended up being really designed for challenges across remote sensing.

Thinking about innovations driven by user needs and specific applications. I want to highlight. This is cutting-edge AI innovation for both elephants and bees. It was actually more relevant often for bees because sometimes the elephants -- they are too big and unwieldy. They are designed by a few companies. Often it is tough to build really for the needs of particular users. Often the bees end up really driving a lot of this innovation.

But state-of-the-art AI is in both elephants and bees. State-of-the-art AI does not mean bigger AI. It means smarter AI. There are a lot of different ways to build smarter AI.

I also want to highlight that AI is not just about the newest, flashiest AI. There are a lot of uses for older methods like simple neural networks and random forests. Here is this fossilized bee. There is still use for the older algorithms. They are widely used in many cases.

I want to touch on one more example quickly before closing, which is some of our projects on gathering data on biodiversity. Here, we are motivated by the fact that there are a lot of species out there. And actually, half of them are insects. But if you look at the data that we have on biodiversity, this is the IUCN Redlist, looking at what species have even been assessed for risk of extinction. You see that the number of species of insects that have been assessed for risk of extinction is half the number of birds. We have only tried to get data on half the number of insects as birds despite the fact that there are literally 100 times more species of insects than birds.

How do we get this data? We are working as part of a global coalition of organizations that build hardware and software and algorithmic systems to monitor insect biodiversity scalably, using technologies that look like this. These are automated camera traps that can actually attract and photograph insects. And then our AI algorithms can process this data and turn it into meaningful species identification like saying these are the particular species with these confidences and you can get data at scale when previously you would need very small number of experts in the world to go out to a vast number of locations. Clearly, impossible to do.

Now, I want to touch on a few things here and just wrap up by identifying why this is relevant to the previous discussion. First of all, I want to highlight that the core algorithm here is actually a ResNet50. It is a standard AI approach, not really a fossil bee. But they are definitely not the start-of-the-art flashiest models because that is actually all you need in these cases. But there is a lot of innovation needed around that.

In particular, one of the things that we find is this is one of our deployments in the cloud forest in Panama. We went down there and we found 2000 species of moths at this one location and 1000 of them were new to science. They were undescribed. And that is a huge challenge for standard AI approaches. Standard AI approaches cannot deal with data that is totally different from what they have seen before.

The frontier for us is often undescribed species. We are dealing with a problem called open-set recognition where we need to build computer vision recognition -- computer vision algorithms that can deal with undescribed species. These are on top of the core AI algorithms, which are a little bit more simpler and old school. This is the mixture of different things that we are seeing here. I guess these are old bees from that standpoint. Some of them are older, classical bees and some of them are the latest, flashiest, cutting edge bees.

And then the final piece I want to highlight here is that it is not enough just to build the algorithms. You also have to build the software and the human capacity around that. There is this global network of organizations deploying these systems. A lot of the work has been building the software stack to allow ecologists to use these tools. They do not necessarily know how to run AI machine learning code themselves. We built the Antenna platform to enable them to do that and make it accessible.

Takeaways overall. I want to highlight that there are opportunities for AI to advance action across many different sectors. There is a need for cutting-edge AI. But it is mostly not large-scale generative AI. Impactful innovations are driven by user needs by really thinking deeply about what that is.

If you are interesting in finding out more, I want to point to resources at Climate Change AI. We are a nonprofit that catalyzes impactful work at the intersection of climate change and AI. It is the hat that I have in addition to my main professor's hat. We have reports like that tackling climate change and machine learning report for lots of different audiences, including for policymakers and other different stakeholders.

We have events, most recently at NeurlPS 2025. You can check out our thousand plus peer reviewed papers online. Our summer schools, which will be happening again this summer and there are materials for it online. We have taught over 15,000 students now, people coming from a whole range of backgrounds. They do not have to be students, people from companies, public sector about the intersection of AI and climate change.

We also have a range of different funding opportunities that we run or make available. If you sign up for our newsletter, you can stay informed about all of these opportunities and more across the space of AI and climate change.

Thank you so much for listening. There is some time for questions.

(Applause)

DR. SAIN: Thanks, David. Can you hear us okay?

DR. ROLNICK: I can hear you great.

DR. SAIN: Fantastic. We will start. Does anybody have any questions for David in the audience? I think we have a microphone here.

DR. WOOTTEN: Hi David. I am Adrienne Wootten. I am a research scientist at the University of Oklahoma and on the program committee. Let us say you have all these different bees, representing all these different tools and things and what have you. They are all different, ever so slightly different. And with everything becoming more open source also on top of it and ease of access, what is to stop someone who is unfamiliar from grabbing the wrong tool for their needs and getting a weird -- to try and make the metaphor work across pollination and things --

DR. ROLNICK: It is a great question. I think that it is a big issue. I will say that I am actually less worried about it with specialized tools and with general ones because I think that the big problem that we often have with general tools is that people do not know when they are getting the wrong tool. It is still the wrong tool under the hood. But if it is user friendly enough or if it seems general then you do not know when you are getting the wrong answer.

But in terms of specialized tools, you need specialized gateways sometimes or you need single gateways that provide avenues to lots of other tools. That could be websites that say use this tool if you have this problem. It could be word of mouth sometimes if the space is educated enough about the challenges involved. This is one of the reasons why Climate Change AI and other organizations exist to build capacity within the space so that people understand what AI can and cannot do and how to find the right AI tools and use those right AI tools.

I think there is some role that may be played by the elephants just in routing to the bees like sometimes there is a need for interfacing with something via chatbot. But the chatbot is not answering the question. The chatbot is saying you should use this tool. That could be one direction not to go.

But I think that by and large, there is not a replacement for expertise. These tools are not going to replace human experts. They are going to enable them to do their work much more scalably. I do not think we should think in terms of AI enabling people who know nothing about something to suddenly do it.

We have seen it again and again, for example, in programming. I think that programming is a great use of generative AI in certain cases but only for experienced programmers because we find the non-experienced programmers are actually getting really bad results. They do not even know what is going on. Across the board whether you are working with elephants like large-scale generative AI or bees, you need to have some baseline knowledge of what you are doing, not necessarily from the AI perspective, but from the domain perspective.

DR. SAIN: I just want to remind you, both in room and online, that you can submit questions via Slido. But we do have time for one more question.

PARTICIPANT: Hi David. Great presentation. Two quick questions. One is where do you see the convergence between the elephant and bee. I am not talking about immediate ones but looking forward to maybe a few years. Where do you see things potentially converge?

The second question is about decision making. A lot of the bees and elephants are used for research and modeling. But when it goes to the decision making, do you have some more insights about the leverage of AI and then potential tools that may help?

DR. ROLNICK: As I mentioned, I think that the best intersection between elephants and bees is elephants routing to bees. People, again, confuse having an interface with solving a problem. A chatbot is not going to solve your problem but it might route you to tools that can solve your problem. The hard problems in AI are often being solved by the bees but then the chatbots provide a nice shiny interface. That could be a way forward. But that means we need mostly investment in bees and some investment in elephants. That is one to address your first question.

Can you remind me of your second question?

DR. SAIN: It was about decision making.

DR. ROLNICK: Decision making. I think that fundamentally we need translators. We need more people from a science background who are actually interfacing with decision making whether that is in public policy or in other kinds of decision making. I do not think there is a replacement for that fundamentally.

Angel is going to be talking about some tools that make it easier for people to again do their jobs of communication better. But fundamentally, we are not replacing communicators. I do not think we should be replacing communicators any more than we should be replacing scientists.

DR. SAIN: Thank you. Thank you, David. I appreciate you bearing with us with some technical issues earlier. Let us thank David for a fantastic presentation.

(Applause)

DR. SAIN: Our next keynote this morning is Angel Hsu, who is the associate professor of Public Policy and the Environment at the University of North Carolina.

Agenda Item: Keynote Presentation

DR. HSU: Hi. Good morning, everyone. Let us hope that my screen share works. Thanks so much to the National Academies for the invitation to address you remotely from freezing Oslo. I am here participating in the IPCC's Special Report for Cities and Climate Change lead author meeting. And trust me. I would much rather be in sunny California. It is freezing here. It has been snowing all day. It is really gross.

But I am also really glad to be part of this conversation because the intersection of AI and decision making -- that was a great final question to David -- is really moving incredibly fast. This was something that people saw as experimental to something that is now shaping how information is being produced, discovered, and acted on.

Today in my short time, I am going to focus on one specific slice of climate AI. David did a great job at illustrating a lot of the use cases particularly for machine learning and for very specific prediction problems. But I am going to tackle the elephant head on and I am going to talk about generative AI, particularly large language models and chatbots and how we are seeing them being adapted for climate decision making.

I am asking the question, what happens when these tools actually do become a frontline for climate information and decision making. David and I have known each other for a long time. I completely agree with everything that he said. But I think what we are seeing is we are seeing a lot of non-expert users that are turning to these tools and trying to understand where should I be investing in resources or, in particular, interventions to address climate mitigation and adaptation. I think it can be potentially dangerous when you have non-expert users who are turning to these tools as their frontlines of accessing those insights.

As someone who teaches data science in a public policy department, I can tell you that the students are using AI 100 percent of the time for all other coding. I agree with David. They do not necessarily get the highest quality answers. But it is really challenging for them to decipher what is actually highly accurate information versus what requires a lot more scrutiny.

To extend David's metaphor, hopefully, you are not sick of bees and elephants. I had no idea what he was going to say. But I want to ask in this presentation what happens when the bees are actually trying to become elephants or when everyone is actually trying to use these elephants to solve what should be bee problems. I completely agree. We need to be designing smarter AI systems for climate decision making and ask ourselves when is AI absolutely necessary.

There was a great piece that Sir Nicholas Stern from London School of Economics and colleagues published last year that categorizes the opportunity. That was great to see in the word cloud that opportunity was front and center. I think this is a really great piece that talks through the range of dimensions in climate adaptation and mitigation and financing where they see a lot of potential for AI to accelerate climate solutions.

They include things, as David mentioned, so transforming complex systems, things like energy systems and transportation networks, urban ecosystems where decisions are often interdependent and things do not happen linearly. It might be really helpful to try to optimize a solution to find where the innovations lie by looking at nonlinear and non-expected relationships that may not be easily predicted by traditional statistical models.

Another is innovating technology and resource efficiency, as David mentioned, design optimization, predictive questions, smarter infrastructure, and how to optimize some of those areas.

Another is modeling climate systems and policy. My group right now for the special report on cities for the IPCC were trying to ask the question, can we actually use machine learning to take advantage of all the large-scale data sets on building height in transportation network and mobility and waste emissions to actually get more detailed climate scenarios that speak specifically to cities because most of the climate models and integrated assessment models are so coarsely resolved, they can really only give you detail at the country level at best.

And of course, there is a lot of potential for AI to assist in adaptation resilience so early warning for extreme events and disasters, climate risk management and strategic planning when there is a lot of uncertainty. There is a lot of AI being applied in that domain.

And then last and this is where I see a lot of potential for these generative AI chatbots is for AI to nudge behavioral change where there can be pattern recognition and personalization that can help people and institutions reduce emissions or manage the risk. Let us face it. And I have been dealing with this because we have been going through a lot of the scientific literature. Climate science is incredibly dense. It is really complicated. And even doing these assessments, we try to make it accessible. It really is very challenging.

That is where GenAI can have a lot of potential to make information a lot more accessible to the everyday user to make decisions and policies about climate. David talked about the different types of AI. I have illustrated it here in this type of nested diagram where we can see that AI applications are really varied. There is a lot of confusion about what AI really means. Whenever people say there is a lot of potential for AI and climate, I think that is the first question we need to be asking. What exactly do you mean by AI?

Because on its broadest level, we are just talking about computers and algorithms that can replicate human reasoning tests. But then there are more specific approaches. These are where the bees might be located like machine learning. That is looking at pattern recognition for prediction purposes.

Within that bubble there, you also have deep learning, which uses more complex neural networks for more complicated reasoning tasks. A lot of the applications that David mentioned so integrating satellite remote sensing data, doing this kind of forecasting, risk modeling, classification techniques. There is actually a fairly well-established set of use cases that he went over. We know how to validate them. We know how to quantify the error and uncertainty and how to integrate them into workflows.

But generative AI -- that is the core of this diagram. These are where area models are being used to generate new text, new images, audio and other content. I argue that this is really the new frontier. Because large language models or LLMs are powering chatbots that are making it easier for people to interact and obtain information much more easily than before, we are facing a totally new paradigm where it is not about data scarcity. My whole dissertation was written about data scarcity and how we can innovate to help better inform and drive more evidence-based decision making. We are confronted with a completely opposite problem where we have information deluge and so much information being generated by these GenAI tools that it has become really difficult for decision makers to decipher what is actually trusted, credible information.

You can see on this plot here -- this is just from mid-2025. I am sure that the more recent statistics are even greater than what is here. But now there are nearly 600 million users a month that are using popular chatbots and GenAI tools like Claude, ChatGPT, Gemini, DeepSeek, and Perplexity. It is making AI a household name whereas before -- a couple of years ago in my data science class, I would ask students, have you heard of AI or machine learning. They would glaze over. Now I would confidently say every single one of my students is using these GenAI chatbots all the time every day.

Globally, we are seeing some of the same patterns as well. It is no longer a niche technology. They are becoming a near universal interface for AI. On the left, you can see this map from the World Bank. It shows how ChatGPT traffic is distributed globally. It is not confined to a single region to the Global North. This is now being used virtually everywhere.

And then in the next context, the Pew did a really interesting survey, asking about Americans usage of AI tools. You can see many Americans report interacting with AI several times a week. It is not just that they log in, they said what is this, and they quickly turn it off.

They are having a substantial share of the responses saying that they use AI daily or even multiple times a day.

For climate change and climate decision making, these numbers really matter because these chatbots are quickly becoming the way people are obtaining information on a daily basis. I was actually shocked in my class how students were pulling out ChatGPT as a first line of information acquisition as opposed to a traditional search or asking me who is standing in the classroom.

Just as conversations are taking place, discussing what the implications are for AI usage across a variety of different climate domains, adaptation, mitigation, et cetera, we have to also understand what the implications are for every day user that are seeking information or guidance on climate science and action particularly when these chatbots -- if you look on ChatGPT and all of these tools, they clearly have disclaimer that they are not necessarily meant to be purveyors of accuracy or truth. They actually say the opposite that there is a high possibility for false misinformation or hallucination. Even though these tools were not necessarily originally designed as climate chatbots, we have to assume that climate questions are being asked and they are increasingly being routed through these general-purpose assistants.

What kind of implication is that going to have more broadly for how we think about these chatbots and this technology in informing decision making? To start we have to think about AI and try to unpack the black box of AI. The first challenge is that these systems are trained on massive, messy, and unevenly documented data. We have to first remember that training data is not neutral. It is often geographically skewed towards the Global North, which has more English language content, more digitized institutions, more media coverage, more research. We can see in the IPCC alone, many of the citations, the vast majority are coming from the Global North. That is research. That is easier to scrape and index. The model's background knowledge can end up reflecting a very small part of the world.

Here is a study that illustrates the skewness of underlying training data in a lot of these LLMs in much more detail. You can see in this one study, they looked across thousands of public data sets' use in AI training. And the geographic skew is enormous. Ninety percent of the training data come from North America and fewer than 4 percent from Africa, which is disproportionately impacted by climate impacts.

When you combine this imbalance with the lack of training data coming from these very vulnerable areas, you get a system that may be excellent at producing fluent responses and point you to all these types of innovative climate solutions. But they may not be actually attuned to a local context.

When it comes to informing and guiding climate adaptation, this is a huge deal because adaptation is highly local and content specific. Governance capacity informality. We have been tackling with that all week here. Infrastructure fragility, public health baselines, cultural constraints and available resources all impact the relevance of a particular climate solution.

If the model that is trying to inform that decision making does not resemble those on-the-ground conditions, it can still give an answer that sounds really great but it is biased and likely drifting towards the context that it knows better from the training data but may not at all be appropriate for the local condition.

Here are some other challenges with applying GenAI for climate decision making and information seeking on this slide. This is just the tip of the iceberg. There are a lot of challenges. But I wanted to just point on a few.

Another issue is the fact that when we use these types of large language models in these AI chatbots, we have no idea how the model is prioritizing certain evidence and that has been a challenge that the domain experts in the IPCC -- we have been dealing with that in the assessment is how do we weigh great literature versus a paper that has a thousand citations versus another paper in another journal that may only have a few other citations. But for these AI models, we have no idea how it is weighing those different types of evidence and that can be really important in building trusted insights.

Another well-known issue is hallucination, which I alluded to in the introduction of this presentation. Climate policy and decision making questions are full of uncertainty, caveats, and conditional findings. GenAI can fill in some of the gaps with plausible sounding detail. But they may not be fully supported.

A third is confirmation bias or sycophancy. If a user asks a leading question or comes in with a strong prior, the model may optimize for being agreeable or helpful rather than being careful and corrective. I cannot tell you the number of times I have played around with these chatbots, ChatGPT or Gemini, and I ask it a question and I know that the response is wrong. I say no, ChatGPT, you are not correct. And immediately, it flips and gives me a totally different answer and tells me that I am right. What happens when we start to narrow our worldviews and the options of information and the viewpoints that we receive because we are only getting our primary source of information from chatbots. To me, that is potentially really dangerous.

And then that has to also deal with my fourth point that is about summary creep. Climate evidence often depends on scope, a particular region, a scenario, a certain level of confidence. But with GenAI summaries, it produces over generalizations. It has a tendency to turn a particular example into generic and overly generalized statements. That can be really challenging because as a decision maker, you want to know. That is what we are doing here in the IPCC is we are evaluating the literature and providing certain confidence statements and confident levels to different pieces of scientific evidence. If you do not have a human doing that, then it all kind of sounds like it could be the same and it is really difficult to understand where is there more or less certainty to motivate a particular decision.

Finally, what we are finding is that LLMs have computational challenges. A lot of the questions that you want to answer about climate data or climate scenarios that involve units or baselines or time horizons and quantitative comparisons -- that is not yet possible to do with LLMs. Although I may say this and then three months later, there could be a new innovation that makes us point out a date.

For example, if you were to ask one of these chatbots, how many companies in Germany have set a net-zero target. It is not able to actually do that kind of computation unless there is a report that says Germany has X number of companies that have set a net-zero target.

If we are applying GenAI to climate information and decision making, especially for policy, adaptation planning, risk and communication, then we have to engineer systems that address these weaknesses.

One of the ways that is being discussed to help solve this type of black box problem with AI and to make AI more trustworthy in decision making is pushing explainable AI or explain ability. Transparency and explainability are often described as this missing link between model outputs that may be really nice on the surface but you have no idea how it is being generated.

I found this paper that was published in Nature Geoscience last year. I like this Venn diagram because it reflects how climate risk questions actually work. We care about where something happens, when it happens, what futures strive it, whether it be temperature, precipitation deficits, land cover infrastructure, underlying vulnerability.

In Geoscience, the right model is not the one that necessarily predicts the best. It is the one that helps you understand the mechanisms to understand what can I actually act on and where should resources be placed.

But in reality, as the authors note that you can see in this bar chart, this type of explainable AI is still really limited in the climate in the Geoscience's literature. You can see that across millions of archived abstracts, only a small share is mentioning explainable AI compared to AI overall. There still needs to be a lot more work to make this a reality. That is where climate-specific chatbots enter.

I have shown a couple here. In full disclaimer, the two chatbots on the left, ChatNetZero and ChatNDC, are chatbots in my group that has produced. And the core idea is simple, recognizing that there are a lot of these elephants that are totally bloated that are over-parameterized with billions of parameters and often highly inaccurate. When ChatGPT came out, I asked it, can companies still continue to burn fossil fuels and set a credible net zero target. It told me yes. And the scientific literature had already made very clear statements and a lot of these emerging regulations saying actually you cannot. You cannot have a credible net zero target and not have a plan to phase out fossil fuels. But ChatGPT in these big models did not necessarily reflect that.

What we did is we used a RAG process or a retrieval augmented generation process to use the elephants but then to try to make it more bee like by saying to the OpenAI APR or whatever model we are using, only pull your answers from the select group of documents that are verified. I know that they are scientifically robust. And that is similar to this chat climate chatbot here, which is taking all of the thousands of pages of IPCC synthesis and trying to make it more discoverable and interactive with a chatbot interface. That is one way to address this problem and a blow to these elephants.

And then you have other approaches like climateGPT, which is actually trying to build a foundational model. It is using Llama, which is Meta’s LLM, to then endogenize it to some extent with more domain-specific knowledge to make the terminology about climate sharper.

And then you have this agentic model here, this climate site tool, which is basically a chatbot that can act as an agent to help you plan climate decisions and to help execute a multi-step climate action workflow based on a set of data or scenarios. It might decide for you and it might say in the back end, based on this question that you are asking for a particular climate decision-making problem, I will look up a relevant data set, pull out the correct indicators, run a small analysis, summarize the result, and cite the sources. Instead of you prompting the chatbot independently through a series of prompts, it just makes those kinds of decisions for you. And that can be really powerful because, as David said, a lot of people who may not have domain knowledge, they might be tasked, particularly in Global South or capacity-strained context -- they might not have the training to go in and understand what information do I need to actually develop my climate adaptation and climate mitigation plan. They may be turning to an agentic chatbot that says I need to figure out what I need to do on climate adaptation and come up with a risk plan. They might rely on one of these chatbots.

This slide here illustrates why they can be really appealing. We conducted a study where we asked these eight questions to the climate domain-specific chatbots as well as to the generic ones like Gemini and ChatGPT and Coral. When we compared the factual accuracy of the climate domain-specific chatbots, we were really encouraged. We said if we compare the original source documents and the most authoritative climate science, we are getting it right with the ChatNetZero. It has a much higher factual accuracy than these generic models.

We also found that the generic models had a tendency to embellish. They would say really nice things like Walmart and Amazon are increasingly vocal about sustainability. That sounds really nice but what does that actually mean. You could be vocal in a negative way. You could be lobbying against stricter climate regulation or you could be actually proactively trying to articulate a decarbonization plan. It could be both ways. We found that these generic models had a tendency to actually inflate and embellish what they were saying.

But the catch and I did not include that on this slide but you can look at the study. When we surveyed 50 climate policy experts and we blinded all the responses, we found that people actually preferred the longer responses, the embellished responses of Gemini and ChatGPT because they were longer even if they were inaccurate and also highly embellished. That really, I think, impresses the point that we have to think about and David said this, the user in mind. We have to think about how the user prefers to receive information and how they are going to interpret the outputs of these responses. They may not prefer the short and pithier responses that we constrain ChatNetZero with.

But I also want to talk about another elephant and that is AI's footprint. David talked a lot about this. I will not spend too much time on this particular side but just to say that data centers consume a small proportion of overall US energy consumption but that is expected to grow quite a lot. As David mentioned, a lot of these elephants that are over-parameterized actually have a huge energy footprint.

But I want to get a little bit more specific and tell you about how we are actually trying to get more specific numbers. First off about the design of climate-specific chatbots. And when we try to control for things like hallucination and try to improve the credibility of the responses, what implication that has for energy consumption, and also to test some of the prompting behaviors and the types of questions that users may be asking of chatbots to understand if you are using a chatbot to ask information about climate science or climate facts, instead of using a traditional Internet search, does that have a higher energy implication?

These are never before seen results. I am sharing them with you now. We are working on publishing this in a paper that will be available soon. You are getting a sneak peek.

But on the left in the chatbot, these are two different prompts of ChatGPT where we are using ChatGPT 4.0. And on the left, this is the energy consumption of a prompt. We fed the same prompt to all these different models. For ChatGPT 4.0, when you said limit your response to 200 characters, the energy impact is actually much lower. That was exciting because last year Sam Altman made the statement and he said stop being polite to ChatGPT. It is costing us millions of dollars every time we say please and thank you.

When I heard that, I thought does that actually mean that it is consuming more energy and having to generate more tokens to then generate a response. What we actually see and I will show this on the next slide is actually that is not accurate. But you can see here that with ChatNetZero with the domain specific LLM and to try to design for a bee, we can actually then reduce the overall energy consumption compared to the generic model.

The ChatNDC, because it has an agentic approach so we are asking the agent itself to do the hallucination check whereas ChatNetZero does that hallucination check outside of the model itself, it is actually a lot more energy intensive.

Even though the agentic check for ChatNDC is 61 percent faster, it uses 25 times more energy. People could be designing for the bee but then unintentionally creating an elephant depending on what design choice they make.

Going back to that Sam Altman question, we tested this and we found that the same prompt -- I do not know if you can see my cursor, but on the left, GPT 4.0-mini and then this short when we constrain the output has a huge impact in terms of the energy consumption.

But what we found is that whether or not you input a haiku or like paragraphs of an IPCC report in the input, it is actually the output that really matters and has the more energy-intensive piece. That was really surprising to us. Even if we put an input question that was 600 times longer in ChatNetZero, then GPT 4.0-mini still shows lower average energy use because the response that is generated is actually much shorter. Do not worry. You can go and be as polite as you want to ChatGPT. It actually is not going to have an impact on the overall energy consumption of the output. That was one thing that we found.

I know I am running out of time. I will just maybe go passed this here and then just talk about -- I can just briefly mention Jevon's Paradox. Also, we have to think about the number of bees that are being built because I think that is also a question when it comes to Jevon's Paradox. This is the idea that as efficiency approves and cost drops, the demand for these kinds of tools often rises. And total consumption can actually increase rather than decrease. We may not be capitalizing on the opportunity for AI if everyone decides that they need to have their own climate chatbot for different purposes. I can tell you that it is actually happening. I feel like everyday people are contacting me in different agencies and different institutions are saying I want my own chatbot to do essentially the same thing. We are starting to think about can we design climate-specific infrastructure and software to help people and lower the overall energy and carbon impacts when everyone is trying to essentially do the same thing and just try to make their own data or their own processes more streamlined and efficient so they can actually capitalize on those opportunities for AI.

And then as David mentioned and in order to do this and to ensure that we can have credible climate AI applications, we have to continue to ensure that humans are in the loop at every process, at every point. We can never replace that domain and expert and that community lived experience.

I think this bottom example shows that when we do not design AI systems that co-create solutions with real people and communities, we may end up reinforcing existing biases or injustices in the current systems. There is a real emergence of the potential for algorithmic redlining that actually reproduces historical redlining where there were these maps that were used to make decisions about how to actually lend and prevent certain communities of color and disenfranchise communities' money so that they could live in more desirable areas.

On the right, you see -- this is a machine learning, unsupervised, statistical classification where using existing data, it ends up producing a map that looks very similar to this redline map. This is a warning of why we need to continue to incorporate human perspectives and humans in the loop when we design these climate AI systems.

And last, I just wanted to talk about governance because I think that we have to really think about governance and guardrails right out of the gate because this technology is evolving so quickly. I serve on the North Carolina Governor's AI Leadership Council. He appointed 25 representatives from different aspects of workforce development, education, health care, and I am the energy and environmental person on the committee. We are helping the State of North Carolina develop a roadmap for responsible AI usage. I think this could be a really interesting model for other states to also replicate. We are thinking through. What are the guardrails? What are the concerns? Who are the stakeholders that we need to engage? How should we design education pipelines to make sure that when we educate kindergarteners all the way to college students that we are helping to inform the possibilities of AI and prepare them to handle a world that is going to inundated with AI tools but to recognize that you still need to have that domain expertise.

There are different examples out there. I put a couple of examples. ISO and the EU are developing these types of frameworks for responsible AI. I love this example of the Dutch government that has this algorithmic register. Any time you use AI to inform public decisions or any type of policy or regulation, you have to also log it into this register. I think that there are ways in which we can actually have responsible climate AI usage but it certainly starts exactly with what David says and evaluating do I really need AI for this purpose. What is the use case for it? How can it be done responsibly? What underlying data do I have to train those models? I think importantly what we have shown in some of our work is how do we benchmark the output and ensure that it is accurate and that it actually reflects the inputs and the on-the-ground conditions and does not end up reinforcing existing biases.

I know I went over but hopefully we still have some time for questions. This is a QR code to sign up for our newsletter and to stay up to date. We just launched a new center called CLAIM, the Center for Climate Leadership and AI-Driven Integrity in Mitigation to look at these societal intersections between AI technologies and decision making. If you work in this space or you are keen to get involved, please get in touch. Thanks so much. I really appreciate this opportunity.

(Applause)

DR. SAIN: Thanks, Angel. A lot of great work there. I think we have time for one question. I think we are going to go to the online audience or the Slido for this one.

PARTICIPANT: This is a question. There has been some discussion around it and I think you started to touch on it a little bit more at the end, Angel. How is the accuracy quality of the climate data used for training these AIs assured? Trust or more bees to QC data or multiple cross-references.

And related, is the data used to train AI curated and 100 percent accurate or trained on bulk data or a mix?

DR. HSU: That is a really great question. And I was actually hoping David would talk about this because in the prep call that we had for this last week, he had some really great insights. And maybe, David, you could share the paper where there is really a challenge.

What we are seeing is that a lot of the AI models that are being used for climate applications are optimized and developed for computer science. Then it is looking at how well I have this training data set and I pull out some of the data and then you hold it out and then you use that to test how well the model then reflects the held-out testing data. But that does not necessarily reflect on-the-ground conditions or in particular with GenAI and text generation whether or not that response is actually accurate because it is very easy. We show this in that ChatNetZero paper where you can have very high similarity between different texts and similar patterns. And those are the metrics that traditionally have been used to know whether or not this algorithm performed in a high-functioning way or in a highly accurate way. But it could give you a totally different insight or be interpreted totally different by an end user.

Particularly in climate policy, what we are finding is that there are very few benchmark data sets that can actually help us test the accuracy of that generated output. I think that is a huge problem. I think that that is something that we have to also invest in if we want AI to become a reality and a credible tool for climate action. We have to actually improve the overall accuracy and the quality of the training data and also get more representative.

That slide about the training data being heavily skewed toward the Global North, there is a lot about climate impacts and about climate policy and climate action in the Global South that we simply do not have enough information on. We need to make sure that we are not chasing the shiny object of AI and then forgetting that we still need to build a lot in just basic data infrastructure to make sure that the AI is actually credible.

I think there is a lot more that needs to be done in this aspect. In my CLAIM project, that is one of the things that we want to do. We actually want to have benchmark data sets where people have said this AI input uses these kinds of things. But then this is the actual output that we are trying to impact or effect. We have collected on this and let us see how well those two align because that is simply just not enough.

In the work that I am doing on net-zero climate policy and decision making, it is very challenging. We try to trust experts and we try to bring experts into that and help us evaluate. But that study also showed that experts also have a very hard time distinguishing what is hallucinated or what is embellished versus what is actual fact and what matches what companies say that they are doing or what the policies actually say. Several of the respondents in our survey said I cannot tell the difference between what was ChatGPT and what was actually ChatNetZero and what I should be evaluating as the most accurate. It can be very confusing. I think a lot more needs to be done in this area. We need a lot of people with a lot of different types of expertise contributing.

DR. SAIN: Thank you, Angel. I think we are going to wrap up this session now. Let us thank Angel again.

We had scheduled a short break. We are going to take a quick break please very quickly. Restrooms are outside. I think there is still coffee and everything out there. We will see you back here in five minutes. Can we do it in five minutes? Thank you.

(Break)

Agenda Item: Using AI to Advance Climate Science to Meet User Needs

DR. SAIN: We are going to move on to the next phase of the agenda for today. This is where we are going to focus a little bit, get into a little bit of the weeds on a couple of very important areas, again, not inclusive or trying to be inclusive but a number of key areas to get us continuing to think about some of the issues at the intersection of AI and climate action.

The first panel is on wildfire. I will turn it over to Hugo. I will also say one more thing. For questions from the audience, we now have two microphones set up at the top of each stairwell. If you want to ask a question, just try to make your way up to the microphone. Thank you.

Agenda Item: Living with Wildland Fire: AI to Inform Adaptation

DR. LEE: Welcome. My name is Hugo Lee and I am a data scientist from NASA's Jet Propulsion Laboratory. I will be moderating this session, Living with Wildland Fire: AI to Inform Adaptation.

Wildland fire is a place where AI is already showing promise and potential. AI supports detection over active fires, situational awareness, and potentially informing land management, and adaptation decisions.

At the same time, we are hearing weird challenges such as data access and quality and downscaling to local decision needs and standardizing trustworthy approaches when AI is in the loop.

In the next hour, we will explore what is working, where gaps remain, and what enabling conditions are needed for AI to support fire-related climate action in a responsible and practical way. We will start with brief opening remarks from each panelist. Then I will facilitate discussion across a few themes with Q&A. As mentioned, for Q&A, we will alternate between in-room questions at the microphones behind the lecture room and virtual questions from the Slido.

First, we will hear from Andre Perkins from the Allen Institute for AI. Andre, could you briefly introduce your role and how AI is being used at your institute related to the session topic and what you are excited about and what is a challenge.

DR. PERKINS: I did have a few slides. I am Andre Perkins. I am a senior research engineer at the Allen Institute for artificial intelligence where I work on the

Climate Modeling Team. Essentially, we are a small team of domain scientists who are working to try and build machine learning tools that can help make climate modeling more accessible, more easily usable to the community. The Allen Institute is a nonprofit. Our tagline is building breakthrough AI to tackle the world's largest problems. But essentially, everything we do is in the open sphere for the public good.

Our team has actually partnered with large modeling institutions that we do our work with, primarily the Geophysical Fluid Dynamics Laboratory, our GFDL at Lawrence Livermore National Labs, who both build and maintain large climate modeling systems. And we are also partnered with other organizations of course for their machine learning expertise such as NVIDIA Square Lines Project.

At Ai2, we have actually done a few different strategies to try and build machine learning actually into climate models. But the most recent work -- we have tried a few different strategies so far. But the most recent one that we have been seeing a lot of success in -- I will not go into all the details here -- is something we call emulation that has been mentioned a few times.

Essentially, we take an ML model and output from one of these physical models and train the ML to do the same thing. We train it to mimic the step-by-step prediction process of the whole atmosphere.

When you train one of these models, the nice thing is then you can do the very same predictions. You have a very similar variable set so you can do the same sorts of analyses that you would typically do on these models. But of course, you can do them much more quickly.

Our first foray into this a couple of years ago was an atmosphere-only model, which we referred to as ACE. We trained it to do six-hourly predictions of the whole atmospheric state. The figure on the left is one outcome of that and showcasing that. Even though we are only training these models to do six-hour predictions, they get many of the emergent processes such as tropical cyclone formations. The plot on the bottom is our model run over some historical time period and all the hurricanes that is generated compared to one of the training data sets above there.

And then in the last year, we have also branched out into actually what we are vying for in the climate modeling space, which are these coupled model predictions. We have combined our atmosphere model with an ocean model emulator from the M-square-lines project, put them together and started running them in tandem. This other figure is just an example of that outcome. We are showing how it is able to replicate the seasonal behaviors of sea ice in the northern hemisphere. We have looked at this.

We have also looked at its ability to spontaneously and realistically simulate the El Nino-Southern Oscillation, which is of course responsible for a large part of the year-to-year variability, especially here on the West Coast.

But the nice part of about these emulators is you build them so that they can replicate some sort of process you are interested in representing and researching but they are extremely fast. We can run about 1500 simulated years per day on a single GPU. Very accessible simulation platform for domain scientists to start to prototype and build from. We are, of course, very invested in making sure these are skillful and useful for long-term simulations.

More relevant for this panel here, which is of course more concerned with local impacts and local extreme weather. We have been exploring how we can also take these climate model emulators, which are typically to very coarse resolution say you are only representing 100 square kilometers as a single point of information in the model itself, how we can take that, and then map back into something that is usable for impacts. This is typically done with downscaling, which has been also mentioned previously but we are essentially training a downscaling model, which we can then pair it with our climate model emulator that still runs extremely fast to say downscale decades of information overreach and on a single GPU in about 45 minutes.

Just for comparison, if you were to run a very high-resolution simulation on a supercomputer for the whole globe, it would take about three months of work. Often, these are very expensive.

On the upper right panel, I am just showing an example of wind speeds for a wind event over California. On the left is what you might see coming out of a climate model. And in the middle is what we can generate with our climate emulation and then downscaling process. And on the right is an example from the original 3-kilometer model for an event like this. I am showing the same thing for other large-scale impacts below showing what it can do for a hurricane wherein the climate model, you can make arguments about whether or not it is actually represented there but we can recover those details nicely.

I will just close. Again, our organization, our team is very concerned with doing this for public good. We are fully open data, open models and open science as kind of our strategy for this. We are heavily involved with our partners in investigating in how we can successfully leverage these models and leverage their speed but also more importantly where they fail. And our strategy of actually partnering with these institutions is with the hope that this is a proof of concept so that other organizations who build these very expensive, very large models will take up similar processes and then provide that information so that we can have a multitude of these things being produced in tandem and then start to be used by scientists and in the community. I am very excited to talk about this on the panel and looking forward to the conversation. Thanks.

(Applause)

DR. LEE: Thank you, Andre. And next joining us remotely is Ilkay Altintas from the University of California San Diego. Ilkay, same prompts, your role and responsibilities relating to wildfire and what kind of AI tools you are using and what is promising and what fields are risky or limiting.

DR. ALTINTAS: I hope you are able to see my slides. I am at the University of California at San Diego. I have a lab called the Societal Computing and Innovation

Lab at UC San Diego and the San Diego Supercomputer Center. Over the last decade in health, I have been leading a program called WIFIRE. WIFIRE was one of the early fire tech research and development programs. We have both operational programs and also research programs under the scale related to wildfires and also climate in general.

The panel generally talks about AI and trustworthiness in climate. Initially, I would like to say when we talk about trustworthy climate information, we often focus on the science itself where we talk about observations, models, peer review. But increasingly, trust is also shaped just like we heard on the open data and open science -- trust is shaped by these: openness, reproducibility of the systems in AI that are produced at science, how data are shared, how models in AI are trained, and how transparent they are, how fair they are. And others can independently inspect rerun or validate these results.

Of course, another part of it is that AI does not inherently make climate information more or less trustworthy. It amplifies whatever conditions we create around data and workflows and original physics-based science it builds on. If climate data is siloed or poorly documented or in accessible or models are not available to reproduce, it really leads to those weaknesses to scale AI or scale (indiscernible).

And the other part of it is if data are open/value curated by clear provenance and context, AI, I think, can accelerate insights, evolving, forcing that trust and confidence. These things go hand in hand.

That is why, as we also just heard, we are talking about similar things that open data and reproducible open science are not optional anymore and they are prerequisites for AI, use of AI and trusting AI-enabled climate information. Findability as in fair principles. Also, findability, accessibility, usability are important but also added to it executable data and models and accessibility as much as usable. These are important. Results can also be examined to enable collaborations and challenged through these. And of course, a big part of it is improvement of these things by the broader community.

I would like to frame this as a question of AI readiness in that sense. It is not just about the algorithms or computing or data itself. It is about climate ecosystems can support open access to high-value data sets and those platforms' availability to be the foundational platform so we can have transparent meta-data, lineage, provenance. We can also interpret those outputs. This is really related to today's panel's questions. How does the use of AI change what we mean by trust and climate information? Risks emerge when AI systems are built on closed data, whether opportunities open up when AI is paired with reproducible workflows or shared infrastructure because again ultimately the challenge is whether AI will shape the climate side. It already is as we just heard. But we build open reproducible foundations needed so AI can accelerate climate action and trust while strengthening rather than eroding trust, which is another part of this.

This is where I would like to mention the National Data Platform that we have been working on. It is designed as an enabling infrastructure for open reproducible science, bringing together siloed data and other digital assets like executable workflows and shared computing environments under one standardized umbrella. Its purpose, as it relates to climate science, is to make climate-relevant assets or the digital assets AI ready by design while preserving their openness and transparency needed to earn that trust across research and enable collaborations within policy and practice as well as research.

And NDP is not just one big platform and interface that is usable. That is one part of it. But the platform aspects of it is it lends itself to other, more specialized platforms and commons-like things, enabling structures to build on it.

One of those things that was built on is the Wildfire Commons so it is a concrete example to translate research into impact and through democratizing digital access for wildfire and wildland fire science and enable them to computing and the community of practice needed to use that computing.

It is at the Wildfire Commons, which focuses in that sense overall, wildfire to start with but also climate-driven disasters where trust and timeliness are essential equally.

And how it works is Wildfire Commons builds three things. One of them is an expert network. Another one is a non-monetary marketplace to actually build fire tech and innovations around AI and models and cutting-edge things. And the fire forest platform, the third component, is built on top of the National Data Platform. It is using NDP services. It is building its catalog, but it is also building the specialized user interfaces and models for wildland fires.

I hope we can talk more about these things on the panel. Thank you for this part of my remarks.

(Applause)

DR. LEE: Next is James Randerson from the University of California-Irvine. James, you have been seen like an AI influence for wildfire science and data. What do you find most useful and what are you concerned about?

DR. RANDERSON: I do have slides somewhere. I am a professor in the Department of Earth System Science here at UC Irvine. I also have a joint appointment in civil and environmental engineering and ecology and evolutionary biology.

I have had the privilege of being able to work on a series of machine learning related fire problems over the last 15 years or so. One of the ones that I was going to highlight was trying to predict the final fire size at the time of ignition. It is kind of trying to predict which of your social media posts will go viral before you push the button on it for good or for bad. And it is really critical for wildfires because there are a lot of situations, including some work that we just published about a week ago in Science Advances where multi-ignition fires are quite common actually from these storm systems that Andre was describing. Some of them are wrapping around and coming up to California. They can induce hundreds of lightening ignitions within a day. There is a need to prioritize where you triage and place your resources in a limited environment.

And another class or problem that I have worked on a lot that has been exciting but highly data limited is S2S prediction of both burned area and fire emissions. On seasonal time scales, how can you improve fire outlooks? For that, we have taken advantage, I think. It has been really a fun stepping back and forth between basic science and predictability. There is ecological memory in soil moisture reservoirs. There are different types of ecological memory that give you the ability to predict wildfires up to six or eight months in advance.

One example that has been really exiting is in the Amazon. If you look at the preceding wet season, it is too wet during the wet season for there to be any fires. But depending on how the soil moisture reserves build up, if there is an inadequate build up of the soil moisture, in the following dry season, a lot of the trees will become very limited in their evapotranspiration, the humidity drops. This is just a summary of some of these problems. I will wrap this up soon.

In the Amazon, six months before the fire season, if you do not get enough rain, then the trees run out during the following dry season. You have this predictability because the rainfall during the wet season is really set by El Nino, the state of the North Atlantic. There are these predictable ocean indices around the world. We have developed an early warning system there.

And then more recently, we have been working on problems related. For example, in real time, how do you identify? If you see a thermal anomaly from a satellite, how can you identify whether that is, for example, a wildfire that is starting or a solar facility or, for example, a gas flare that is less likely to take off and do damages? We are trying to classify using different approaches.

And also, another direction of my lab is trying to predict the risk associated with individual Santa Ana events. Today, for those who are visiting, it is a beautiful set of days in Southern California. We are in a Santa Ana event. And Andre, you showed a really nice example of some of the winds that can happen during those events. And those winds -- they are intense. They drive the most devastating fires, Palisades and the Eaton fire. But not all Santa Ana events generate these wildfire risks. It is really tied to the rainfall and the status of the live fuel moisture in the months before you get to a Santa Ana event. We are trying to use machine learning to look for this and in collaboration with Alex Hall at UCLA. We have been working on different aspects of that.

I think the holy grail in this space is to build a new generation of fire spread models. There are fire spread models that work on empirical-physical relationships. Some have coupling with the atmosphere like more fire. We are poised to generate a new class of predictive models potentially that inject different smoke and aerosols into the atmosphere and at the same time enable and improve predictions. I am working with a team right now to try to basically use machine learning together with improved parameter ops, optimization of physics-based models to develop a new class of models and then also to use that to ideally identify where in the landscape should you place fuel treatments to have the biggest effect. That is the diagram on the left.

And on the right, one direction that my lab is going in is to develop -- it is kind of like a reanalysis for fire. The top panel is an example of every 12 hours trying to track how a fire is growing and then fusing that with information about climate. We are trying to build these data cubes of how the fires are spreading and then overlaying stacked information about the ecology, the topography, the climate, ideally the winds like from your effort, and get all this information together for thousands and thousands of fires and tens of thousands to hundreds of thousands of spread time steps so that we have the data to be able to develop machine learning approaches.

And the right bottom is a panel from Rebecca Scholten that shows the duration of wildfires and how long they last across the world from harvesting information from viewers' observations.

These are some of the barriers, I think, for AI model use in wildland fire research. I think there are two elements. One is that compared to the data available even though it is highly biased as we heard this morning from Angel, there is an acute shortage of high-quality wildland fire data for model development.

For S2S prediction for these outlooks, we have a very short time series. At the same time, for fire events, we really need high-quality -- basically, it is like a re-analysis for weather that includes information about ignition timing, spread, extinction, fire suppression, but also -- the fire suppression is really important and the impacts. Once we have this kind of re-analysis, I think it will open up the ability of computer scientists to really engage more deeply in the problem.

And then relating to what Ilkay said, barriers to the trustworthiness of AI-derived wildfire predictions.

This is a really critical challenge for wildfire operations. It is so essential that you are placing firefighting crews in different positions around the fire. In no foreseeable future do I see a rapid shift to AI systems. I think the way this may happen potentially is that there is trickle that is increasing in its flow of potentially like AI ensemble members for fire weather and then potentially also ensemble members for fire models. I see a future where maybe fire managers on the GIS spatial component of an operational team might be drawing from AI but there is not going to be a whole scale replacement because of this critical need to understand and interpret known physical models.

For integration in operational systems, I think clear identification of AI-derived ensemble members is really important. That is not happening right now in weather. When you get your weather on your phone, you cannot tell if it is coming from a physics-based model or if it is coming from an AI model. I think that is a real challenge for us.

And then I also am highlighting what Ilkay said and my last point. I think in this field, simplicity and reproducibility should be a priority, even if there are tradeoffs in terms of the model performance that you are developing.

DR. LEE: Thank you, James.

And finally, I am inviting Alan Talhelm from CAL FIRE to the podium. From the management and the operational perspective, where can AI actually help and what do you need from the research and the tool development community to make those models more useful?

DR. TALHELM: Great. Thank you, everybody, for having me and good morning to you all. I am Alan Talhelm. I am the assistant deputy director for Climate and Energy at CAL FIRE. For those of you who are familiar with CAL FIRE, you may have a question of why we have somebody who is devoted to climate energy at CAL FIRE. You probably think of us as a firefighting organization and we are very much a firefighting organization. But we are also the state’s forestry department.

My role is to interface with the rest of the state that is making climate policy. I also oversee portfolio grant programs that are funded through the state’s cap and trade program. Those grant programs are focused on forest health, doing resilience treatments like prescribed fire at scale on our landscapes, building tribal wildfire resilience, and then investing in business and workforce development, including research and development grants to help us with the next set of tools that we can use to help build a more resilient landscape and more resilient communities in California.

My background is an ecologist. I come from the forest ecology carbon cycle world but now very much I am in the policy space as well as doing the science. I am also collaborating with Ilkay and others to help build new science and data tools to help the scientific community, help land managers, help first responders better respond to wildfires and help make better investments.

Just level setting on the California wildfire crisis. We are almost exactly a year passed the LA wildfires, the Palisades and the Eaton fires, where there were 30 fatalities and more than 16,000 structures destroyed in Los Angeles. I also want to say that, yes, we are a Santa Ana wind event. But to calm your nerves, we have also had a very wet winter so far in the last weeks. If anyone in the room from out of the area has been given a little anxiety, I want to soothe you a little bit.

But I also want to point out that as deadly and as destructive as their LA fires were, that was not even the most destructive fire over the last ten years in California. In 2018, we had the Camp Fire, which destroyed the town of Paradise, 85 fatalities, nearly 19,000 structures destroyed basically in the course of a day.

I also want to give you some context for where California has come in terms of our wildfire experience. For many decades, the largest fire in recorded California history, going back about 100 years, was the Matilija Fire, down here in Southern California in 1932 of 220,000 acres. That was the state's largest fire until 2003 when the Cedar Fire erupted in San Diego County. That was 273,000 acres.

The Cedar Fire is now, I believe, the tenth largest fire in California history. The Matilija Fire is now the 17th largest fire in California's history. We now have had a fire over a million acres within the State of California. Nearly all these fires have happened in the last decade, many of them in the last five years.

Something has changed in California as it relates to our relationship with fire. There are a lot of factors in that, some of which are our past land management, some of which is climate change, some of which is where humans are on the landscape and how we are interacting with that landscape.

What has been the policy response to this? Well, the state has invested in more firefighters. In a couple of years, CAL FIRE will be about 15,000 people. This has more than doubling of our strength from a decade ago.

We have an updated and expanded aircraft fleet. We now have C-130s that we have purchased with the help of the federal government from federal agencies. We have night flying, Black Hawk helicopters, which are purpose built for wildfire fighting. We have new technology like the ALERTCalifornia Network, which I will talk about later.

We have an Interagency Action Plan which comes to the realization that this is not just a firefighting and land management issue. This is a water quality issue. This is a wildlife habitat problem. This is a transportation. This cuts across all angles of government. It is a land use planning problem. It is such a huge societal problem that we all need to come together to effectively resolve it.

And then the state, through my programs and other programs, have invested in proactive land management, including pathways for wood utilization so that maybe some of the vegetation management that we can do can pay its way out of the forest and then also workforce to actually get all this stuff done in California.

What are the roles for AI that I am seeing that we are already seeing? One is early detection of wildfires. We have this ALERTCalifornia system. We have more than a thousand cameras distributed on mountain tops and towers across the state. With the help of AI, these cameras are constantly scanning for wildfire ignitions and detections with the hope of detecting these ignitions before anyone makes a 911 call. This has been very successful. It was named Innovation of the Year in Time Magazine a couple of years ago.

We need improved fire behavior modeling for sure, as other panelists have mentioned. This is a frontier. We particularly have a weakness in modeling the behavior of fire and the spread of fire coming from wildlands into the built environment. We have had a lot of practice at modeling wildfires in the landscape and forests for 30 years or more, 40 years. But this is a frontier of figuring out how the transmission of wildfire happens from the wildland to our urban environment and then where it goes once it is in the built environment.

But also even on the landscape scale, we need to get improved predictions about where fires are going and particularly how do you integrate the role of fire suppression in these predictions of fire behavior.

Enhanced community planning is another way we can use AI to help us figure out how to evacuate communities quickly, when to evacuate communities, how to design communities more effectively so they are more fire safe, more resilient to fire.

Optimize land management. We are at the point now where we can have model predictions of what are the carbon tradeoffs of different treatments going to be. What are the wildlife tradeoffs going to be? What are the water quality tradeoffs going to be? What are the changes in fire behavior going to be? How do we optimize these treatments to get the best value out of the investments that we are making on the landscape?

Advanced firefighting technology. We are starting to see this already happening. Automated helicopters, drones, other things like that that we can use as tools to help put out fires when we have those early detections.

And then another space that I have not heard or I have not seen as much investment as I would like to coming from the forest ecology forestry side is new technologies for vegetation management. One of the things that my program has invested in is new tools, things like a BurnBot, which is basically a tool to do a machine that does prescribe fire for us. Automated masticators which go out and do vegetation thinning, tree thinning out on landscapes so we do not have to have a human there. Those are the things we need to invest in because frankly the cost of doing all the vegetation management work that we need to do, all the restoration work that needs to happen on the landscape is tremendous. Unless we bring down that cost, we are not going to be able to scale it to do the work that we need to be able to do out on the landscape. That is a place where I think technology is really needed to help.

What are the barriers to doing this? I think trustworthiness is going to be very important. I think firefighters and the land management community are very much willing and interested in new tools. But they are also somewhat of a conservative group. If we go the AI route and get it wrong too many times, there are going to be people who will want to pull back and not go down that road and are not going to trust AI and these types of tools for the future because they see the risk and the flipside of getting it wrong, the lives that are lost, the property that is lost if we use one of these tools and it does not turn out the way the model says that it is going to. I think that means we need to be very open and honest when we are developing these AI tools about their strengths and weaknesses as we deploy them.

DR. LEE: Thank you, all.

(Applause)

DR. LEE: Now, I am moving into an interactive discussion. I will pose a few questions. And panelists, please -- I will invite multiple perspectives from all of you and also feel free to ask questions. We will alternate between microphone questions and online questions via Slido.

Let us start with trustworthiness. In your view, what makes wildfire-related climate data or information trustworthy? And how does the use of AI change that trust calculus? Is it validation, provenance, as Ilkay mentioned, or some uncertainty or communication or something else?

PARTICIPANT: I guess to start on that, it feels to me that building on some of the things that Ilkay said that there needs to be a really clear history like a rigorous evaluation of the information. It needs to be published in open access literature. It needs to be reproducible. And again, I feel like probably in wildfire research but also in other aspects of weather, you cannot have a blending. We need to keep a clear separation and understanding of where the streams are coming in and where you are using more traditional sources.

When we are doing reanalysis, we have to know exactly what the climate data streams are. We are trying to pick the highest quality ones based on the performance of the re-analysis. I feel like there are a lot of important analogs for fire in that space.

DR. LEE: Thank you for mentioning the need for wildfire re-analysis data. I would love to hear from someone like Andre or Ilkay about the provenance and the importance of the open data and models.

DR. ALTINTAS: I think important -- I mentioned a little bit of course in my opening remarks. AI does not make it more or less trustworthy but I think we can increase the trust in the process of provenance and governance of data in AI-enabled pipelines. One way to achieve this is going from these national static products to dynamic pipelines if data is open ended enabling platforms like the ones I mentioned today were the NDP and the Wildfire Commons. Trust shifts from that snapshot. Is this correct to a workflow question? It is a pipeline question. Can I understand reproduce monitor?

One example I can give is Alan and I worked on the Wildfire and Landscape Resilience Data Hub. We have a scientific committee deciding on some of the things that depend on the open data. Through this, the state actually opened up data for resilience metrics and also interagency tracking of vegetation treatments. We have scientific consensus on how some numbers are created and generated algorithms, not just AI.

But then we also -- while we do that, we make the pipelines that led to creation of, for instance, footprint calculations to the public. Some of these are challenged, of course, and they should be challenged when they are being used.

Then we offer -- these things can actually be reproduced by others who might want to bring in their own algorithms or some AIs on top of this. They can also generate their versions of it and compare.

This is, in a sense, a workflow question. It raises the bar for transparency a little bit on about how it was produced to can I reproduce and monitor the outputs generated over time. Is it auditable? Is it versioned?

The second part of this is data application. AI actually helps boost data to increase trust in terms of its -- maybe making -- filling gaps, satellites, sensors and re-analysis. We actually heard all of these in the products. Can we fill the gaps and create more spatially and temporally available data sets through AI and also decrease those biases and things like that through AI for sparse, uneven, historical data sets? These are actually some of the things that we need to mitigate from.

And the third one is very important. With AI, uncertainty can be very explicit or dangerously implicit. We need to really in essence balance this. Physics-based climate models well-understood explainable answers of these frameworks. They improve over time.

Then we can quantify this and create ensembles or confidence intervals are more statistically covariate-type approaches. There are many other things that can be done here. They increase trust. But when we treat AI outputs to be very precise, that is the expectation from, by the way, the users when they are making decisions if something is true or false. How do we actually not turn this into a black box uncertainty and also find ways to communicate that uncertainty for decision makers?

One thing we have done in our fire risk program for initial attack, for instance, with -- we have an initial attack program led by (name) in partnership with CAL FIRE, a Fire-Integrated Real-Time Intelligence Information System. I think that is what it stands for. We lead a fusion center at UC San Diego for it. And the models of fire behavior models there -- we can actually come up with the uncertainty of those, of course. But then you communicate it as some percentage or things like that. It is sort of decreases, we noticed, the ability to take decisions from it, also ability to trust it.

But if we can turn our confidence in the inputs and the outputs into low, medium, and high confidence and explain that a little bit, it helps the decision maker. How do we make these explicit/implicit balancing and communicate the right thing to the right audience and really explain the results?

DR. LEE: Thank you. Do you have any additional sentences?

DR. PERKINS: I will not expand. I think everything they have said is really great. In our space, in particular, we are at the state where we are pretty much only a research-grade product. We do not make any guarantees and we are explicit in the literature as such. But it does bring up, of course, the danger in that when these are so easily accessible and able to run. Anybody can take them and use them to a certain degree and what those use those for could be anything. That is a very active area of what we are trying to think about and how we build language around these products and more specifically how we make sure our partners who are helping us build and release products that nominally are attached to their name to product, how we can be careful about that.

DR. TALHELM: I just want to build on one of the examples that Ilkay mentioned which is this data commons that we have around wildfire. To date, a lot of the work that the state has done in that space has not involved AI.

But I think the process is a good example of trust and maybe a metaphor for what we need to do.

We have built this database of land management, of vegetation management activities related to wildfire. It has been a really complex process. It has been complex from a technical side, yes, for sure, integrating all these data sources and validating and things like that. But the challenge has been a much more political one than a technical one. The use and communication of these data from different sources. We have a couple of dozen different land management agencies that are feeding data into the system is very sensitive because there are dollars tied to these things. There are politics tied to all of these things. We had to do a lot of relationship building with these different data sources and have built a lot of deep trust. And it took us a long time to do that. There was not really a way to make that go fast. This is very much of advances moving at the speed of relationship-type of situation.

The output that we have had, I think, has been very valuable and we have seen people use this database that we have created in ways that we did not necessarily imagine or anticipate when we built it. And people are using it for incident responses but they are using it for planning investments. They are using it for scientific analyses. They are using it for tracking whether the state is going to be hitting its land management goals related to carbon neutrality. Uses have been really diverse. It has taken a lot of leg work to build that trust in the relationship. And then we have tired to maintain that, as Ilkay mentioned, by making as much open source and open access as we can so that someone who has questions about all of this stuff can investigate as more deeply and see the homework that we did to build this process.

DR. LEE: Thank you. Could you read out one question from Slido related to trustworthiness? Can we get a room question? Please go up to the microphone and ask a question. Please tell me your name and affiliation.

PARTICIPANT: Just a quick question. It might be more for Alan. I am not sure -- as part of the industries that are associated or tied in with -- by the way, my name is (name). I am with the Hydrologic Research Center in San Diego.

My question is about your interactions or engagement with industry stakeholders. In the scope of your work, have you engaged with the insurance industry? Because one of the things that comes up in properties that are destroyed is, of course, is California becoming less and less insurable against fire hazards.

PARTICIPANT: Yes, we are interacting with the insurance industry. The state has its own Department of Insurance. One of the parts of this that makes that a little bit more complicated is we have an elected insurance commissioner here in California. They are independent from the rest of the administration. They act in their own interest. That does sometimes create some political challenges. Yes, we are actively working with the insurance industry.

We work very much with the -- IBHS, I think, is their acronym, which is the industry group around insurance risk modeling and best practices and things like that to help us develop our own standards for things like home hardening and community planning to help us communicate to the public to policymakers of what types of interventions should be made at that scale in order to increase wildfire resilience and decrease the risk to property.

DR. LEE: One more question and then we will move on.

PARTICIPANT: This is a question actually for Andre. I actually asked it on Slido but I was not sure it was being read. It is about your modeling. I think in one part of your presentation, you mentioned how the data that you generate in your emulator -- I was not sure exactly how that is done. Is there any kind of embellishment, as we heard another speaker say, about emulated data that could be embellished by any kind of machine learning? Is there anything to calibrate that with the actual physics model?

DR. PERKINS: I think maybe you are getting at the fact that machine learning when it goes out of bounds, there are no guarantees or constraints on the outputs of that prediction itself. For the models we train, we train them in such a way that we can run them very similarly to how the original physics-based simulation was generated so we can compare long-term statistics and extremes and ensemble bounds against the original model insomuch is that we actually have data from the original model. It is hard to run a three-kilometer model more than a few times over long term. But we do very traditional, I would say, atmospheric science or climate science focus assessments of the outputs in that way.

We do notice things in these models that are incorrect. For instance, if you look in the upper atmosphere, the stratosphere kind of behaves very slowly on timescales compared to what you are considering for weather. But it is also very heavily impacted by climate change.

We first were training our models so they were only sensitive to sea surface temperature. We were using that as our analog for forcing this over time. And then you would see the stratosphere flipping through different climate regimes very quickly where it should not be happening.

We are involved in looking into where these models go wrong and how badly they go wrong and perhaps why and if we can fix it. It is a long process in that. But I will leave it there.

DR. LEE: Thank you.

We will move on to the next question. As panelists have already mentioned some challenges and concerns, we will move on to the next two questions. What opportunities do you see and what conditions do you think are needed to facilitate AI-accelerating climate action?

DR. PERKINS: Maybe I will just start from that one. We are involved in building these models because we have a long-term vision of that this information in the climate modeling ecosystem has largely been funded by the public for the public good. We would like for climate modeling and simulation to operate in that space as well. We are building these models with the idea that the communities/stakeholders will be able to create information that is relevant to themselves.

In that, we still believe we will need people who are educated in the domain to use and analyze and also convey some of this information but having models where you do not need access to a super computer to run is a big part of what we think we can provide in that space.

DR. TALHELM: As I mentioned earlier in my presentation, I think there is a big opportunity to use AI and help us develop new land management tools to actually put treatments on the ground.

I think there is a risk there in that the dollars that we have for these projects are limited. I think there is going to be some reluctance to investing in a new tool if you are not certain it is going to work because you are making that investment in that space because you are trying to protect a community, protect a sensitive habitat. If your investment goes away, you lose those dollars because you invested in a piece of technology that you thought was going to work but was not successful. You have now lost that opportunity to protect whatever resource you are trying to protect.

To get over this barrier, we need more pilot projects. And this is the thing that government is not necessarily well equipped to do. This is not a space where we have a lot of history of success in this. This is a place where I think philanthropy can come in, where venture capital can come in, to do these types of pilots and demonstrate the success. We have seen some of this already happening with a partner of CAL FIRE, the Moore Foundation, who has piloted some new technology in the Lake Tahoe Basin, which I think has been a tremendous demonstration. We need more of those types of things if we are going to be able to successfully integrate AI and new technologies into the wildfire resilience space.

DR. RANDERSON: I think for both weather and for wildfire AI success, on the fire weather side, there is a critical infrastructure of observation both from surface and from satellites that is central. It enables re-analysis.

I think on the weather side, some of the training and the evaluation relies on these major center re-analysis products like the ones from European Union, ERA5, NCEP, HRRR in the US. Those systems are powerful but I think it is easy for the community to not recognize that without a sustained investment in those basic observations that we are really shutting off our ability to use these new tools.

I think right now NOAA’s budget is under threat. These data are critical. It is an opportunity and it is a challenge at the same time but we have to have these operating systems.

The same thing -- there are system that NASA uses, as Hugo well knows, that provide the land surface conditions. Not all the products that we want for specifying the land surface state. We cannot take those for granted and it is really a critical thing for building the new generation of AI.

DR. LEE: -- conditions you think that are needed to facilitate AI to accelerate climate action.

DR. ALTINTAS: The good part of going last is I can agree with everyone. I think something new I can add to it is AI can accelerate climate action if you do some things right. The first thing is first. It needs to be grounded in the right action. We need to solve the right problem to take action and collect and assess our needs in a multisector way so we can create the right innovation platform for a science-driven action and an AI-driven action.

It needs to be grounded in open, high-quality data and embedded in those decision support workflows through those partnerships and also supported by and enabled by platforms and governance. I think we need to create an approach and framework around it.

In the Societal Computing and Innovation Lab, we actually created an innovation approach around actionable science, which starts by getting together and coming out with the needs assessment to finding usable or useful things to solve, first of all, with the state if possible, through cutting-edge science and AI, and then creating an innovation pathway that integrates AI and science into those processes. Since we start with partners in both multisector industry and government entities together with research, those partnerships typically lead to then scalable and sustained use of those things that come out of the innovation pathways as usable, actionable tools.

DR. LEE: April, I think we have time for two questions. Panelists, please keep the answers to less than one minute.

DR. MELVIN: I am going to actually try to combine two questions here, read two out together. This one, I think, was targeted at Andre but if others want to chime in, please do so. How does the emulator uncertainty or the ensemble spread change with forecast length in comparison to traditional numerical models and also which specific tools do you use to train your models and did any tool work better, for instance, more efficient, faster, more accurate than others?

DR. PERKINS: Maybe I will address the second part first. Most of our model training is done using typical machine learning tools so Pi 2, AArch, and Python and other things. We are not really doing anything special in that arena. In fact, a lot of the things we borrow from frontier research like NVIDIA -- but it is mostly a byproduct of us as domain scientists, knowing what is important for our problem and building the implementation around what we are interested in for our simulation platform.

As far as the uncertainty for forecast timescales, we have not actually invested a lot of time in weather forecasting, in particular, mostly because we are focused on longer-time horizons so building stable simulations. A lot of the statistics and extremes we look at are generated -- are gathered statistics over multi-year/multi-decade rollouts to look at that.

I think there is, of course, a lot of interesting work being done, especially in a lot of labs like ECMWF. They have invested heavily since they have the premier ensemble forecasting sort of facility for the globe right now. They have also invested in AI on this. I would point to their work, looking at ensemble spread and calibration and other more short-term forecasting statistics.

But of course, we are interested in it. We just do not necessarily have the capacity to focus on everything that could be done with this model at that time.

DR. LEE: We will move to the next question.

DR. MELVIN: Sure. This is another insurance question. How interested are insurance providers in longer-term fire projections? What are the opportunities to use AI to improve them?

DR. TALHELM: I really do not want to speak on behalf of the insurance industry. I do not know if you have insight there. I definitely would say there is some interest in the insurance industry. I can tell you that that is, in fact, occurring. There was a hand in the audience that shot up when I said that. Maybe we have somebody who has more insight.

Yes, there is interest in the longer-term risk. But I should also say that insurance policies are written one year at a time largely. There is an opportunity for the insurance industry to react on a shorter timescale and they do not necessarily need to be looking three decades out into the future when they are making some of their decisions.

DR. RANDERSON: -- interesting to me. It feels to me like there are dozens now of risk maps that have been developed by re-insurance insurance companies and also public ones like the State of California has a really nice map.

I find really curious right now in this space is that a lot of that data is being held very closely internally and it is very difficult to assess the uncertainty in the risk like in a 1-year time horizon or a 30-year time horizon. It is very difficult to get quantitative estimates of the uncertainty. And I think that is actually a challenge for the market that AI can help. There is machine learning. There are ways you can combine physics-based models with ignition probability distributions to come up with these risks for properties or for buildings. Right now, we do not really have a framework for doing quantitative uncertainty quantification on it.

DR. ALTINTAS: I can add something because some interesting things are happening actually right now for the future. There is fear in this moment and time. California Department of Insurance is actually spearheading the creation of the first publicly available wildfire catastrophe model in the United States. This is to stabilize, bring some more writ reforms, I think broader reforms to stabilize the insurance market and also make these risk assessments more open. It is looking to inform regulatory decisions and support risk reduction and bring more consortium models between academia and government to support risk model generations and their explainability. I hope these things will, in time, improve also the proper use of AIs in this area.

DR. LEE: Thanks a lot to all the panelists and to all of you. Let us do the quick poll and then we will transition to the next discussion session.

Please respond to the question on the screen. What do you think should be the top priority, near-term focus area to improve the accuracy and useability of wildland fire information developed using AI?

We are seeing results together. The majority of people chose more high-quality observational data as the top priority, followed by the improved communication and transparency about how AI is used in wildland fire science and/or development of applications that can inform decision making. It is not a coincidence that the third one is use of AI to improve model downscaling and fine-scale spatial resolution data. About 9 percent think that addressing workforce limitations is the top priority.

Thank you, everyone.

(Applause)

Agenda Item: AI for Water Resource Management

DR. WOOTTEN: Now we are transitioning into our next panel. I am mindful that I am between everybody in the room and lunch at this point. I will dive right into it. It is a pleasure and an honor to be moderating this session. I am Adrienne Wootten. I am with the University of Oklahoma as a research scientist. I am one of the workshop Planning Committee members.

When I think about AI and climate in hydrology and the other connections of water quality, water resources, and water resource management, I look at it as probably some of the most fascinating and complex challenges because of the nature of all the different folks who are water resource managers in different parts of the country, be it as using my own region where I work in the South Central US as an example in Louisiana where they have the consistent challenge of too much water and getting a lot of their water from the Mississippi River Basin and other rivers in the area, going all the way West to New Mexico where they have the consistent problem of just not having enough water and getting much of their water from the North American Monsoon through the Rio Grande Basin as well as the snowpack in the Rocky Mountains and Southern Colorado. All of these are fascinating challenges as we move forward in a changing climate, affecting all of these.

This panel is going to get into those questions about how AI can be used with water resource management, questions in climate science, but also there is also a lot of uncertainty as we have already hit on. We will get into that also and some of the other similar questions, as you have heard in the last panel.

Each of our panelists here on stage will have eight to ten minutes to talk on a couple of things, first their brief description of their position and responsibilities related to the panel topic and a brief overview of the AI involvement in that particular topic, what are they excited about, perhaps, more importantly, what are you nervous about, and what specific kinds of tools do you like to utilize.

Before we begin, I will remind folks asking questions in the room, please go to the microphones in the back if you want to ask a question during the question-and-answer session. And for folks on Slido, please feel free to ask your questions. Since I know folks were asking questions, you can also up-vote questions on Slido if there is a particular question you would like to hear answered. Please do that.

Without further ado, I am going to turn this over to our first panelists. I think you have slides, Debaditya. To our first panelist, this is Debaditya Chakraborty.

DR. CHAKRABORTY: Thanks, Adrienne, and thanks to the National Academy of Science for this timely workshop and the invitation. I was asked to talk a little bit about trustworthy AI and how we can effectively apply AI in water resource management.

My name is Debaditya Chakraborty. I am an associate professor at the University of Texas at San Antonio. My background is mostly in applied AI in both physical and biological domains so ranging from water resource management to climate change to cancer studies. We work on a particular type of AI called explainable and counterfactual AI that helps us answer questions like why is the AI predicting something and are there any hidden patterns in the data that we can reveal and in the process of doing so, discover things that are useful, in this case, useful to water resource managers.

What are the different promises that I see with AI? That is the most exciting part. The first thing that we are really interested in is understanding why the AI is predicting something and what are the different impacts of interventions like the climate interventions, any kind of adaptation and mitigation strategies that are being implemented? What are the impacts of that? Can we quantify those impacts using artificial intelligence so moving certainly beyond correlation to more looking at the causation like action A results in this particular outcome, uncovering hidden patterns in the data sets, using explainable AI, trying to find out if you have a particular basin, what are the different regions within that basin that you can apply conservation on to enhance water recharge in the aquifer, for example? That is another promise.

But there are some challenges as well like sometimes black box models are preferred because of the hype in certain situations like people would want to apply transformers instead of extra boost, let us say, for example. Does it lead to the best outcome possible based on what we are doing when you have structured tabular assets, region-specific, where you need precise solutions? Maybe not. Maybe it is still prudent to go with an algorithm like Random Forest or extra boost over transformers if you want to get the best results that are useful to water resource managers. Kind of asking ChatGPT to give me a solution for a problem and blindly taking that answer is probably not a good approach. That is one of the nervous outcomes of using AI. The danger of using opaque models versus more of the grey box models or white box models.

And also, the threat of premature deployments before doing proper generalization tests and understanding the explanations, understanding why the model is predicting something. If you deploy certain models that might lead to negative outcomes.

This is one of the researches that we did where we looked at a basin. We looked at several basins but this is just one snapshot from a basin to understand how AI can help us understand which parts of that basin of land conservations would enhance recharge so using things like (indiscernible) explanations like SHAP line, integrated gradience, trying to understand this is where land conservation might enhance recharge to the best possible ways. These are the kinds of questions if you can answer.

That leads to user crust and also usefulness of the results rather than just giving a point prediction and telling where the recharge is going to be in the future.

Another thing is model selection. As I was saying, sometimes it is easy to get carried away and using a very complicated architecture versus a simpler architecture. But what we saw in our research is that when it comes to skewed data sets, applying slightly simpler models rather than deep learning convolution of neural networks or transformers can lead to better predictions and also more meaningful physical outcomes.

We did this research where we looked at emulating behavior of recharge from physical models. We had a USGS-based physical model and we were trying to emulate recharge. We saw some differences between the AI-based emulations and the underlying physical models that were used to train the AI. For example, in this situation, the points marked 1, 2, and 3, we saw significant discrepancies where USGS physical models said zero charge whereas our AI was predicting super high recharge during those spaces.

In those situations, maybe using some kind of commonsense and trying to overlay real world analysis like observed analysis like groundwater levels on top of it to see which ones are actually making more sense. By the way, the groundwater levels, J-17s, are the green lines. They were not used as features in the model. But when we overlaid that on top of the predictions, we saw that there is consistency between the AI-based emulations and the observed data versus the physical models that were used to train the AI.

Even though AI was trained on the physical models, it kind of overcame the problems that we see with the parameterizations of phys8ical models.

And then again, several other verifications like from different simulations, not just one, but looking at satellite data sets or other kinds of simulations to see whether AI makes sense. Not just one layer of tests but when you apply several layers of tests on top of the AI-based results, that, in my opinion, leads to more trustworthiness and ultimately better adoption of the technology.

This is an interesting problem that we also looked at like comparing simpler tree-based models versus deep learning when it came to short-term predictions of water levels or spring flows. We found that again tree-based models like our ensembles outperformed when we were predicting one to two weeks ahead of time in comparison to other models like LSTM or transformers or CNN.

This is interesting because there was one instance in the CNN when it actually gave up like flat lined. In the training period, it was producing amazing results like .98 or .99 R-squared power used but -- because during the inference phase, it gave up. Sometimes when you ask ChatGPT and it gives you a very silly answer and we catch it red handed. Here, we caught this one doing the same thing.

It is also not just important to look at point predictions but also look at the usefulness. Is it predicting the extreme values as well like the drought period like the critical stages when the resources are in the critical danger phases? Can AI actually identify those instances? Again, we saw that tree-based ensembles outperformed deep learning, another thing that we need to keep in mind.

Complexity does not equal accuracy. What we saw and -- over smoothing is a critical problem with deep learning. We have to be careful about it. And data scarcity says sometimes -- for example, the recharge problem we saw. The data was very scarce and skewed. In those cases, we have to be really careful, using very complicated models, complicated black box models.

The next thing is counterfactual AIs, not just understanding why AI is predicting something but answering questions like what would happen if we take certain interventions or what would have happened if this intervention was not taken. In those cases, we use counterfactual AI and see the impacts of applying a certain mitigation or adaptation techniques in terms of addressing certain climate change impact.

It is important to go beyond the hype, especially in hydro-informatics, to not just use black box complicated models, but try to use simpler models with post-hoc explanation techniques and counterfactuals to understand why the model is predicting something and what would happen if certain actions are taken, which will enhance the usefulness of the models and also increase the trustworthiness and enhance adoption.

Thank you very much and I will look forward to the discussions.

(Applause)

DR. WOOTTEN: Thank you, Debaditya.

With that, we will move on to our second panelist before we open it up for discussion and that is Kathleen Boomer. I have two different affiliations for you so I may let you introduce yourself.

DR. BOOMER: So grateful for this opportunity to contribute to an important conversation. My name is Kathy Boomer. I am with the Foundation for Food and Agricultural Research. We are a grantmaking organization, funded through the Farm Bill and as such, we are charged with supporting applied research.

I work in the water space and working at FFAR provides an amazing opportunity to engage and interact with technical experts along the entire food supply chain, decision makers in industry, farmers, in the board rooms of big ag, food and beverage companies, as well as policymakers at local, state, and federal scales and nonprofits.

I, myself, am looking at that work, looking at how AI is being used, what are the concerns and the constraints around AI, and what is the excitement around AI in those communities through the lens of a critical zone scientist and also a landscape modeler.

I have also had tremendous opportunities to work with experts in decision science in indigenous knowledge systems and surely, I am standing on their shoulders in my contribution here today.

Before I dive into trust building through AI, I would just like to share how I am seeing the challenge in front of us. Really, when we are thinking about climate action, it is to ensure water and food security for future generations.

We have a lot of technology, a lot of know-how. But in the water space, let me share those advanced technologies such as precision irrigation and conservation drainage that are implemented on well less than 5 percent of our national and global crop lands. What is explaining this gap between technology and implementation?

I would share that it comes down to trust and I am going to start with what is most intuitive to many of us in the room and that is a weak system framework to transfer information from our labs in the field or in the computer lab to real-world conditions. Stakeholders, land managers, farmers do not trust or do not see the relevance of that information to making decisions in the field.

Similarly, there is a concern about the integrity of the accounting systems that we are using to inform management to evaluate our progress toward our shared goals. A lot of uncertainty about green washing, for example, by Walmart and other companies like that.

What I would say, however, as the most important reason, is that many of us generating that information do not understand the decision context or the process for making decisions.

What is really exciting to see is that we are increasingly realizing that trust building does not come from disseminating information out to end users. It really comes through the integration of local knowledge with academic knowledge so really shifting from a competitive mindset to a collaborative mindset. This is increasingly being bared out by cutting edge social, economic research and decision science. And it is really fun to see how it is dovetailing with indigenous knowledge systems.

Furthermore, there is a very well-established process for facilitating that collaboration. Managing a system requires a community effort, space to reflect on the challenge, thinking about what are the potential solutions to mitigating that challenge, assessing the tradeoffs in a formal way that drives consensus and eventually consent to take action. And then of course, it is critically important to have evaluation, monitoring, and research to evaluate the efficacies of the actions and to reset the next reflection, if you will, in confronting those global challenges.

Importantly, this process requires trust, space for reciprocity, valuing everybody in the conversation, finding ways to incorporate their knowledge. And that means creating space for feedback and having willingness to adjust times and modeling approaches to think about those tradeoffs and best solutions.

I will share -- I do not think of AI -- I will put on the table that I am not very excited about AI as a predictive tool. However, I think AI can be valuable to fast tracking this trust-building process.

To date, there has been a lot of resistance to that process because of the time and the complexity of engaging in this kind of conversation. But AI really can empower us to move this much faster and much more effectively.

For example, in the reflection phase in meetings like this, AI could be valuable to cataloging our different ideas, our conceptual models, cataloging our ideas around mitigation strategies, what are the concerns and constraints around this mitigation strategies, surfacing conceptual models, and then synthesizing those diverse perspectives in an equitable way.

A lot of the talks we have heard today really speak to the assessment stage, improving the models, improving our capacity to estimate and evaluate those tradeoffs.

Yes, AI will be invaluable to improving our model performance. But I think even more exciting is the promise of AI to help with the pre-processing and the post-processing of those model data and even to do that multi-modeling approach so that we can begin to not only develop more relevant models to the decision making, but we can also better understand the uncertainties, start to decompose and understand what is the structural uncertainty versus the parameter uncertainty versus the uncertainty in the data that we are collecting. In this way, it really helps to guide us as to where, when, and what data we need to monitor and collect in a way to strengthen our decisions over time.

Two other quick points about this slide is first

I hope you will see where I need to emphasize that this is a value-driven process. It has to be led by people. We can use models to fast track these conversations but they cannot replace the work and the synthesis and the evaluation of those data.

The other point is to realize this framing really lays out a roadmap for the kind of workforce development we need. We need colleagues who are comfortable with using AI to facilitate and elicit these ideas. We need colleagues who have a holistic understanding of modeling who appreciate the breadth and the diversity of models and, in particular, appreciate the value and power of conceptual modeling and how to translate that to our more mechanistic and statistical models.

We need colleagues who recognize the pitfalls of AI, ensuring transparency, look out for spurious results, and direct and context shifts. And then of course, being able to translate those model results back to, not only our colleagues on the ground, implementing those actions, but our colleagues who are collecting data and thinking about how they can help facilitate decision making into the future.

There is one other thread of indigenous knowledge systems that I cannot resist putting on the table with this group as a critical zone scientist. When indigenous communities come together to start this process of collaboration, they acknowledge that you cannot start that work without acknowledging a common belief, the common set of beliefs, and that common set of beliefs is based on water is life. They understand that the ebbs and flows of water across time, daily, seasonally, annually across millennia, shape the distribution of plants and animals across the landscape, the land forms, as well as where, when, and what crops we can grow or where fire is more likely to pose a risk.

It is really interesting to contrast that with the AI models that we have been using and to question are they really capturing this conceptual model. Most of the time we are driving these models by precipitation data, which do not capture the more complex water system.

I share this as I think it is a wonderful example of an alternative conceptual model that is maybe not captured when we try to feed a lot of data into the system. Importantly, if we do not capture those alternative models that greatly enhances the risk of making a bad decision because we are not capturing the hidden costs of how we are managing our land and water resources. This ultimately could really undermine technology development as well as water and food security.

In closing, a plea to center on water. But then when we are thinking about trust building through collaboration, what we need is a commitment to knowledge co-development, elevating conceptual modeling, also thinking about promoting the idea of multi-models instead of the best model. This mindset could be invaluable to helping us shift from does a practice work or not to when and where a practice can provide the targeted benefits if properly designed and managed. Multi-models could help with the shifts over time and space. Nested modeling, in particular, could be really helpful to understanding why drift occurs, for example. I touched on the value of multiple models for understanding uncertainty.

And the last point that I want to make is that kind of smart modeling is essential for improving models over time. We all know that more parsimonious models are better models. This really speaks not only from a model accuracy-precision-credibility standpoint but also from a resource intensification need. For example, we know that more and more data is not necessarily going to come up with a better answer. We cannot afford the more we have, the more data centers we need. We cannot keep going down that path. We really need to think about how we can leverage this technology more effectively. And AI is essential for that to fast track facilitation and also pre-post-model data synthesis. Thank you.

(Applause)

DR. WOOTTEN: Fantastic. Thank you, both, so much. We will move right ahead into questions and answers. Again, a reminder for those in the room. You can go to the microphones and please use your name and affiliation before you ask your question. And for folks on Slido, of course, type in your questions and vote for ones you also want to hear about. We will go back and forth.

Seeing as you both got into my first two questions already, I am going to modify this a little bit on the fly because both of you approached the trust question from extremely different angles with it. That does lead to an interesting thing that I know we are probably getting at a little bit in this workshop in that you have a water manager and you have a hydrology modeler per se. Let us just use those two as an example here. They have two different definitions of trust essentially. How do you bridge that divide between the two of them to use AI in water resource management for climate change actions?

DR. BOOMER: I would say trust can be thought of as a need to surface different ideas about how the system is behaving and how we can manage the system. Having comfort with finding those uncertainties and confronting those uncertainties together is a really exciting opportunity, not only to foster collaboration on the ground, but to advance our research.

DR. CHAKRABORTY: These questions are not just coming up. It has been there for a while. When I first started working in AI back in 2013, we were faced with the same challenges, same questions like some kind of healthy speculation about can we trust these models. From a data science point of view, the answer is always going to come from what kind of generalization test have you done to ensure that the AI is not just memorizing but it is actually learning the patterns and giving you an answer that is based on rational interpretation of the data set that it has been trained on.

Newer questions are related to how holistic your data set is and can you expand the horizon of the data set. There is certainly scoped to improve that and enhance the outlook of AI whether it is going with the analogy in the morning, whether it is a bee or an elephant. Is it giving you precise answers? The way to do that would be through as many generalization tests as you can possibly do.

DR. BOOMER: There might be a complementary need, which is field testing those data so we can do a lot of model performance estimates. But at the end of the day, we need to take the information that you are generating and test it in the decision context. Is it useful? Does it make sense from the end user's perspective? Does it make sense in terms of their conceptual model of how the system is behaving and might respond to management practices?

DR. CHAKRABORTY: Absolutely. 100 percent agree with that. Sometimes AI goes against -- and that creates a major challenge like we faced in the recent past. We went to domain experts to generate completely unbiased models on their own and to make sure that AI is not doing something out of the ordinary so yes, to go back to your point -- expertise needed. AI --

DR. WOOTTEN: Excellent. Do you have a question? Go right ahead.

DR. MORRISON: Thank you for those talks. I am Monica Ainhorn Morrison. I am at the National Center for Atmospheric Research in Boulder. There is an element that came up in both of your talks although, I think, from very different perspectives that is interesting and I wanted you to elaborate on. A lot of the data that we are using is not adequately capturing phenomena that communities that might be downstream users of these tools care about. I think that

that is a data scarcity issue but also it might be a consequence of the fact that climate data and earth system data is relatively unique in that the data does not explicitly describe things. We define them and then we extract the information in order to do that sort of interpretation. I am just wondering how especially with hydroclimate, hydrological phenomena how we can do a better job of interrogating our data to make sure that it is representing things in the way that communities are experiencing them and what that might mean for our data processes when it comes to training and evaluating our models.

DR. BOOMER: To answer your question, I would want to start with sharing my definition for applied research and that would be informing a decision. A cutoff line for me is if I am at FFAR and we fund applied research, is that modeling work being developed to inform a decision and whose decision?

An important process of strengthening that bridge is not only sharing the needs from the decision maker's point of view but also engaging those decision makers in developing the scenario assessments that you might use or that you apply your model to inform.

I will pause there and --

DR. CHAKRABORTY: To answer your data scarcity question, it is a challenging aspect and data scarcity (indiscernible) in the data set. One particular type of AI like generative AI, different models like diffusion models, auto encoders, latent space perturbations -- these are right now giving us good, early results in terms of overcoming the data scarcity problems. But that again leads us to a challenge like can we trust the data sets that we are generating using those diffusion models.

Right now, we are working on trying to control the generation of synthetic data in a way that makes physical sense. We are utilizing some Bayesian statistics to make sure that it is well controlled within the means, not outside the boundary. That is showing some early potential but let’s not get ahead of ourselves. But it is a good solution probably to overcome data scarcity.

DR. BOOMER: I think we will always be dealing with data scarcity that our conditions are changing over time and we will never have enough data to capture what has happened in the past in a way to predict new conditions. But we can think about, for example, in your work, I know a big challenge from a producer’s perspective, for example, is even thinking about whether or not to invest in capital infrastructure for water management, whether or not to shift in terms of which crops to grow. These are decisions that are closer in scale to where you work and examples of how your work might more effectively inform practitioners on the ground.

DR. WOOTTEN: With that, I think we have a Slido question, April.

DR. MELVIN: This is a question about drought. How can AI be used to improve drought monitoring, drought assessment, and drought forecasting when a spell of dry weather transitions to a drought?

DR. CHAKRABORTY: Can you repeat that question?

DR. MELVIN: Sure. How can AI be used to improve drought monitoring assessment and forecasting when a spell of dry weather transitions to a drought?

DR. CHAKRABORTY: That is a particularly challenging aspect because again when we look at historical data sets, trying to just predict droughts will result in a skewed classification problem. How can we address that challenge? Something we have done in the past is minority oversampling and majority under sampling. That particular approach works. As I said, to answer the previous question, diffusion models are showing some potential in terms of generating these under-sampled data points especially like drought. If you have a hundred-year time period and have 10 or 15 instances of droughts like major droughts that creates a particular challenge in predicting it because you are trying to predict a skewed tail of a distribution. How can we oversample that skewed tail? Diffusion models right now are showing some potential that we can do.

And then the next step is the prediction step is kind of simple. I would probably fit an extra boost to Random Forest to once the data set is ready, we use that to model the droughts and use that for future protections, using climate forecasts from RCPs or SSP scenarios and trying to predict decades forward what the drought scenario would look like.

Also, the scale is a factor so looking at regional scales rather than global scales would make more sense and give us precise answers. It is something we could also look at.

DR. BOOMER: This is a question that starts to make me uncomfortable because I want to know who is asking the question and what do they think of when they are saying AI can provide us with the answer. Even if we were all using the same AI model, if you asked geomorphologists and hydrologist to answer that question versus crop modelers, I bet you would come up with very different answers because of how they implemented the model, what data they chose to feed into the model. What a great opportunity just to underscore a need to think creatively about uncertainty and being careful about answering these kinds of questions.

DR. WOOTTEN: Indeed and I would add to that actually that building off the thing with drought, the flipside of that, the flooding and extreme rains, we have run into very much the same problem where you are looking at the far tail of a different -- the opposite tail of the distribution and how much of a challenge that can be.

I see we have one up here. Go ahead.

PARTICIPANT: (Name) NC State University. The thing I want to talk about -- ask a question about robustness of those tools especially in terms of extreme events especially with the data that can be either intentionally with low quality or because of the natural conditions that, for example, Hurricane Helene knocked out a lot of the river gauges that you do not have information of real time of those high-quality data as input to a lot of those AI tools for water resource management. How are you thinking about the robustness of those tools to support the decision making, especially the data as input? It can be either intentionally or unintentionally with low quality. Sometimes people just like to mess with data --

DR. BOOMER: I would agree. Further, I would share

-- have concerns about the water models but because water is so fundamental to driving all other agroecosystem processes, I worry about the robustness of carbon market models, of TMDLs. Again, having a more robust approach to understanding uncertainty and to valuing -- thinking about valuing characterizing uncertainty in a way where we can understand where more information, more focus is needed. I am thinking of value of information analyses that could really help over time to improve the robustness of our models would be invaluable to our efforts.

DR. CHAKRABORTY: The particular types of explanations like local explanations are especially valuable to look at different inflection points in the model's predictions so trying to answer the question of why a model is predicting a drought and whether it is the right prediction or not. What we could do is we could look at the local interpretations, assuming that it is an explainable model, AI model. And then trying to take those inflection point explanations to a domain expert and ask them questions from a domain point of view does not make sense. Based on your experience, do you think that this is physically possible or biologically possible depending on what kind of topic you are discussing? That can help enhance the robustness of the model as well as the validity of the model.

DR. WOOTTEN: Awesome. I think we have time for maybe one more quick question but I will ask the panelists here to keep it brief because we are between them and lunch.

DR. MELVIN: Another online question. One of the big challenges in co-production convergence is communicating across disciplinary boundaries and perspectives. Any pitfalls or benefits of incorporating AI into that process?

DR. BOOMER: I am seeing AI as being a solution to that reality. Of course, you would need to have experts in applying that AI tool. But I think what is really exciting about AI is its capacity to help us understand each other's vocabularies, to help us see where there are alignments and how we are thinking about the system and its response. Maybe even more exciting is the different ideas, the differences. Again, hoping we can compel a paradigm -- I would call it a paradigm shift of instead of trying to ignore uncertainty, to surface those uncertainties and look at it as a really exiting opportunity to advance our work together and foster collaboration.

DR. CHAKRABORTY: I think AI is definitely useful in communication but it does not eradicate the need for human expertise. We need more AI experts. We need more domain experts. We need more educators when it comes to properly communicating the usefulness of AI and how to use AI in the proper way.

DR. WOOTTEN: Excellent. Thank you, both, to our lovely, fantastic panelists. As you head off to lunch, I can see people are already responding in one to two words. What is the primary challenge to incorporate AI in climate planning for water source management and/or water quality? There are a lot of them of the same size, which is interesting. There is not a whole lot of convergence there.

I believe we now have lunch for about an hour and then it is off to our next sessions in the afternoon. Am I correct? Fantastic.

(Lunch Break)

AFTERNOON SESSION

Agenda Item: Use of AI in Agriculture and Land Management

DR. GUAN: Hello everybody. Welcome back. I am Kaiyu Guan. I am a professor from the University of Illinois. I am running a particular interesting session that I feel is related to using AI for agriculture and land management. For this particular session, we are very honored to have three speakers actually coming from a quite diverse area of the agriculture research or industry. And then I am going to briefly introduce them. We have Catherine Nakalembe from the University of Maryland, who is also part of NASA Harvest. And then we have David Lobell from Stanford University and then we also have Emma Bassein from John Deere.

And then just a high-level layout of this particular session. We will give every speaker about five to six minutes to talk about high-level work that they have been working on and some of the relevant topics related to this. And after that, we will get into the discussion. I really want this to be a very engaging discussion that everybody -- that we all participate.

And then the design of this particular panel that I am envisioning is we have the researchers and the industry people that come from the large agricultural system that studies US agriculture, US farmers versus we have the (indiscernible) farmers from Africa that have quite different needs. And then both of these agricultural types of systems require experience to climate change and then thinking about AI will help. And then we also have the industry experience as well as the academia and practitioner's ideas.

With that, I will first welcome Catherine to give a presentation about her work.

DR. NAKALEMBE: Hello. I am Catherine Nakalembe. I am an assistant professor at the University of Maryland. I chose this title from model to multipliers and the word translational here is critical in the applications of geographical information, science data, using AI to scale and implement projects and activities related to agriculture.

The next slide summarizes a lot of what I do. If I could run an organization, it would not be one that works on one specific component here where we have data collection, that field agent who might be using their phone. It might be an instrument (indiscernible), et cetera. We have ground sensors. And then one of the backbones of my work is the vast amounts of satellite data that we have access to, going back all the way to '70s although this is not consistent in places.

And the other component is trying to get value out of that so that it is relevant for agriculture. If you want to do crop-type mapping, you need crop-type labels from the field and then you can train your machine learning model to create your crop-type map. But then you can input and do crop-specific, crop condition assessments, which to do that you need rainfall, temperature, data, et cetera. But you not only need it for now and in the future to forecast, but you also need to be able to go back.

And the other thing that is important is around scale. As Kaiyu was just mentioning, what might work really well for the US so looking at the Midwest for mapping, maize, wheat, et cetera, is a lot easier to do in the Midwest than it is to do, for example, in Western Kenya. I sometimes say that I can close my eyes and map US agriculture but I cannot do that when I am trying to do, for example, indiscernible), because of the complexity of the system.

But ultimately, one of the things that Kaiyu was asking is to make a connection as to how is this useful in decision making, et cetera. You have to repeat this process continuously in order to be able to inform risk assessment, disaster assessments, policies for improving agriculture be it irrigation, where you put infrastructure, what crops are doing well where, and how can you improve them. This is an overview of the majority of the things that I do.

On the next slide what I was trying to show here is that when you are actually on the ground where the farmer is working, be it in Illinois or in Kenya, it is a lot messier than we think. We, I think, sometimes assume that when you go to a machine learning conference or the AGU, somebody will show you my model has an F1 Score of something. It is better than everybody else's. It can help a community wide do X. We could save this amount of water if we implemented this method. It does not actually work like that in reality because on the ground, it is as if all these things that we talk about and conferences completely disappear.

I like to call this the messy middle because one of the reasons they disappear is that to get them to be relevant for finance, investment, and infrastructure, optimizing at the farm requires working with people more than working with motors and machines. This is very complicated. We are not very skilled at that particularly if we are working from -- you are used to working on a computer with your data. You can wrangle it and improve it.

You can come up with all sorts of methods for making it better. But when it comes to translating it -- this was also mentioned, I think, at the very beginning. Making sense of you might have a black box model that predicts that if you added an extra kilogram of NPK to this field, you get this much out of it. Convincing the farmer to believe you and do the scheduling that you decided will be the best is not as easy as it might sound when you write it in a paper.

I like to criticize abstracts because one of the things that they say -- the last sentence is how is this useful and you always have to have that sentence. But then when you read deeper at the work, it is much more often irrelevant than what is needed on the ground.

When you look on the ground -- I like to show this because I think it represents not only publications. There are lots of those around methods and machine learning, AI, remote sensing. We are really changing the world in terms of publications and data sets, et cetera. We have data sets, models, papers, reports, publications, et cetera, that tell us about what is going on about agriculture. If you search any place on the planet, you will probably find something that says this.

But these do not typically translate into delivery of fertilizer, the right fertilizer. They do not translate into an insurance policy for the farmer who needs it the most. They do not translate into an early warning for a person who might be affected by a flood in the next three minutes, et cetera.

The reason why in my title that I had the word translation is that we need to be able to come from my fancy graphic at the beginning into something that is actionable and useful. One of the examples I used to use that is really critical is if you get a warning on your phone when the phone buzzes for a flood warning that is action. It tells you, do not drive in that direction. That does not happen for many of the things that we want to improve.

In order to move forward going towards what are the things that we could do, over on my next slide, I summarize what are five fundamental things that might be useful that could make early warning, remote sensing, et cetera, and AI data sets and methods be useful for a farmer, Mary, who is working in her field in Tanzania in this particular case. What is it that Mary would have needed to receive so that she does not lose her harvest? It is much more complex than the messy part that I showed in the beginning.

One of the first things that I think about a lot is while I really like working and analyzing data, labeling and creating and using new models, some of the things that are most useful is the fact that it is available on time. It might not be 100 percent accurate. It might not have the best F1 Score. It might not have used the most recent network, what was mentioned before. A simple Random Forest model or a regression model that gives me information to make a decision. The flood warning that is available immediately is much better than something that -- Catherine spent six months developing the best method using all the data that exists in the world. It applies the same to farmers in this case.

And then the next one is -- this has also been mentioned in many different ways where if we are going to build even more sophisticated models, we have to recognize the importance of data, ground truth data, what data are being fed into the models that we end up using. In the beginning, one of the speakers mentioned the geographic bias enabled data. When it comes to agriculture, this bias is even more phenomenal in the sense that while I can close my eyes and map corn, wheat, soy, et cetera, in Illinois again, for me to be able to create a really good accurate map or a map at all of where maize, wheat might be growing in some way is not that straightforward because the data are not available. To get to it, you need to invest in this type of data, huge investments, to collect the diverse combinations of data sets required.

To do crop type mapping, one of my projects that I worked on -- I call it helmets labeling crops. This is where AI is helping us address a critical problem. What we do is take pictures as you drive with GoPros and then we are using a computer vision pipeline to basically detect crops and then translate those into crop-type labels. This is a drop in the ocean because we are looking at very huge, very diverse agricultural systems.

On my next slide, I wanted to show another example of why ground truth is really important. Last year in July, I believe the satellite was launched for soil moisture, which is critical for predicting agriculture. However, in terms of calibration plans for the mission, there is a huge gap in where the data are available. There is not a single station for soil moisture that is part of the carve-out work that is being done by NASA in the continent of Africa, which has the biggest burden in terms of crop loss and impacts of drought in communities.

Trying to address this problem goes back to if you want to have the best algorithms, the best method, it is not only the NASA data that has the input that is going to make it work. We need to address the foundations of what does soil moisture of X really -- that is measured from NASA translated into soil moisture in Ethiopia where the soil type might be different from Malawi, for example. Another example of something that I am trying to work on.

And then on the next slide, trying to build towards making AI data sets and products useful is ensuring that there is financing. And the financing here I mean as if I can predict and show that an investment in irrigation infrastructure in this area can increase yield by X amount, that investment in the irrigation infrastructure needs to be available, otherwise, it is just creating anxiety, knowing where the possibilities are and where the capabilities are so trying to understand what are the limitations. We can model those but addressing the need, the gap, to be able to deliver the solution requires financing beyond just the research itself.

And then number four is policy. We were talking about this at lunch where fundamentally some things do not have to be fancy. They just have to be done. Let us say we know trains are better than cars. If you put in a rail line, you would save so much money. You would reduce pollution, et cetera. This applies towards policy and agriculture. If we know that irrigation infrastructure works, then we put in policy that enables financing to put the infrastructure where it is most required. If we know that data are important, we put in policy to collect data just like the USDA collects data through the extension network. It is the same thing that applies even in the context for Africa to sustain it over long term.

And then number five is around working with people, people to bring not only context. The people on the ground and the places where we think our models would be useful and relevant have different interpretations. I think Kathy from the last panel was saying who is asking that question because the response would be different. If I am a policymaker and I am trying to improve yield, my question will be very different than if I am farmer who is trying to improve yield. Where is the shop where I can buy the fertilizer, et cetera? Trying to work with different people from different contexts to understand and be able to build that workflow that completes the cycle is really important.

One other thing that I do not include in this is number six is how we measure impact has to be very different. It is not the accuracy of the model that is most important. It is the effect that will help a community or a particular place get better from what your prediction was.

This is my next to slide, which is recognizing that more data sets, more models, more compute, better estimates and forecasts do not directly translate into impact. There is still a huge gap between platforms, data sets, reports, and models and the impact that we hope to realize that scale.

One of the biggest gap fillers is policy that is institutional to invest in not only data systems and infrastructure that make it possible for farmers to succeed. Thank you.

On my last slide, it is just saying thank you and some of the work that I am going to be doing trying to institutionalize and implementing these types of things. Thank you.

(Applause)

DR. GUAN: Thanks, Catherine. We will have David. You will be the next.

DR. LOBELL: Okay. Thanks, Kaiyu. Hi everyone. I am sorry I am not there in person.

(Video 1 ends)

DR. LOBELL: -- that I've found often take a while for people who are used to thinking about AI, but not you sit thinking about agriculture, it would take a while for them to sort of understand or appreciate. So I think the first is simply how much patience agriculture requires. In the context of AI, for example, you can produce with these new AI-based climate forecasts, as I'm sure you've already talked about, really exciting forecast accuracies, and you can produce data that are better than anything that has existed before.

But then if you want to see it have impact, you have to get a few things to happen. One is you have to have the trust and the confidence of the users to actually use that, which takes a while to build that trust. And then once you make the decision, you actually have to get lucky, in a sense, that you actually come across a year where say that forecast is actually beneficial to farmers and they see the benefits, and then that kind of reinforces the uptake. So any sort of agriculture technology, if we think of AI-based forecasts as one new type of agricultural technology, it could easily take a decade before we're sure that it's really impactful and that it's working, and that can be frustrating, I think, to people in the AI space, but it's I think an inevitable lesson for most of agriculture. It's just things take time, both because of trust and because of just the variability of the system.

The second point, I think, is very related to what Catherine was saying, was how AI is often thought to be kind of a substitute at this point for collection of data, but in most places it's really so reliant on good data that it's really best thought of as a complement or a multiplier, and it really increases the value of more data collection. And I would say that the biggest successes we've seen in terms of AI in agriculture to date are in cases where we have a lot of high-quality labels, and these are -- I think Catherine's example was good, where people are on the ground with helmet cams going through and acquiring lots and lots of georeference data. There's another great example, which is Plantix, which is a company using phones to diagnose diseases, and they've had a bunch of experts label many thousands of images to generate the types of accuracies that then have shown to be of use to farmers, and Plantix is now used by I think millions of farmers to do diagnostics with AI.

We have a recent example looking at using phones to do yield estimation in the field, and that only works because the partners we're working with, in this case Pula, have done thousands of crop cuts on fields and have actually good measurements in these locations of what the yields are.

Another good example actually is when we can annotate images with field boundaries. That's something that you can get people to do pretty rapidly. It's easy to create a lot of labels, and then AI can really take off, and we're seeing now the production of AI-based maps of where fields are around the world because of that.

But there are so many other variables, as Catherine alluded to, where it's really quite early stages yet in terms of having actual good training data for these models, and I think the investment in that hasn't sort of matched the potential of what's there.

Finally, I'll say I think it's useful to distinguish between AI for generating data or forecasts that are more accurate and cheaper than before and using AI to sort of optimize or analyze data to inform decisions. I think the second is where I get very excited in the long term about our ability at the policy level or at the field level to really optimize management much more than we can today. But again, I think that is going to require a much firmer base of data. We're seeing examples of that in many cases. I won't go through them, but you can imagine as you piece together lots of data on how farms are being managed, what policies are in place, and what the outcomes are that you can start to generate insights into what sort of a better system would look like. But without that base of data that we have to train on, we're not going to get there. So again I think it points back to the importance of investing in traditional data sources, which is I think an ironic point that I think a lot of people want to skip over.

So I'm going to keep it short, Kaiyu, and we then we can get into it in Q&As.

DR. GUAN: Thanks a lot, David.

Last but not least, we will welcome Emma to provide her inputs.

MS. BASSEIN: Thank you, everyone. I'm Emma Bassein, I work at John Deere, which I know is a little bit of a different vibe than some of the people in the room here. I manage a team called Sustainability Solutions, which is working with our growers on adopting sustainable practices and engaging with the ecosystems services markets. So things like carbon credits or other programs along those lines.

I'm here in part because we've worked pretty closely with Kaiyu on a few things, the University of Illinois, and others. But a lot of people think of John Deere as a machinery company, and we certainly are and have been for nearly 200 years. But we also are surprisingly a technology company. Almost all of our machines report back data in near real-time, so we have hundreds of thousands of machines reporting back agronomic and machine data to a centralized database, where growers can use that information to optimize their operations.

With hundreds of thousands of machines covering 350 million acres of data, we have what is likely the largest agronomic database in the world. So what do we actually do with that, and that's a great question and part of what we're here to talk about.

AI at Deere is one tool of many to help our growers solve specific problems. My team in particular is helping growers access the sustainability markets. It's understanding their own data, optimizing the practices that they're using in their field and then connecting with programs that want to compensate them for making those changes, taking that risk in their operation. And there's a lot of places where AI can play into that, including modeling the actual carbon outcomes, backfilling data that growers may not have complete sets, doing data quality analysis for the information coming out of the machine.

But the other larger area where AI plays at John Deere is around what I would call mitigation and adaptation to climate change, if we think of broadly mitigation as reducing the impact of a particular sector and adaptation as adjusting to the variability that we're going to see in an increasingly changing world.

On the mitigation side, there are a few industries where the actual business interest of the industry is so well aligned with reducing inputs and use of products. The economic benefit of doing more with less is very apparent in agronomy, using less fertilizer, using less fuel, margins are pretty small, so all of the technologies that we have in the field are generally around how do we get the machines to produce more of the product, how do we get the land to produce more of the product, with fewer inputs, with fewer passes, that sort of thing.

So some of the AI technologies that we have out on the market include one called a predictive feed rate control, that's actually using remote sensing to see the condition of the field and adjust the speed of a harvester so it's going the correct speed to have the least losses as possible while it's harvesting. That actually really changes the outcome of that production.

A number of other places in picking the best route through the field, knowing which fields to go to. Many large operations in the United States have fields, hundreds of fields, distributed across multiple counties, sometimes multiple states, they have to move equipment to get to the right place at the right time. If you put that equipment where the fields are soggy and wet, you can't get into them, you've now lost really important days in your planting season.

So all of that is stuff that I consider to be part of this mitigation and adaptation strategy of how do we help farmers who are already dealing with a variable system control and work on that system better.

We rely on our own data but also a lot of public data for this, so I want to thank everyone who's worked on weather modeling. We use soil maps that are really important, we use a lot of remote sensing data, mostly Landsat and Sentinel-2, because they're available and have long histories.

We really appreciate the NAS surveys. That's not an AI product but it's something that is really important for continuing to understand how farmers are farming.

One of the things that I'm hearing a lot in these is talking about modeling risk and uncertainty, and it's such a different thing when you're dealing with customers. At the end of the day, our models are successful if a farmer looks at it and says, yup, that's better than I could have done. Which is both a very high and a very low bar, depending on the day. I think David said something along the lines, you have to get lucky. You could have a model that's like 85 percent accurate but if the first time they use it it's going to tell them something in that 15 percent where it was wrong, they're going to be, nope, we're done.

So that's part of it. And as a business we're able to transfer, get some of that trust and that buy-in, by setting up business models that actually help adoption. With some of our more advanced AI products, we actually have a you only pay us if it actually got you savings structure, so that we can build that trust and adoption over time.

So those are options we have as a company, that are not necessarily available to science, but I think it's an interesting thing of how do you build risk tolerance over time, how do you work with the end user to communicate that risk, and move through that?

The elephant in the room, any time I say we have the largest agronomic database in the world, everyone's like great, how do we get access to it? And it's tough because the data is individually controlled by those growers. So this is not John Deere's private data stash that we can hand out any way we want. This is data that's coming off actual operators, and they have to give us permission to use it in any way that's appropriate for that.

That being said, we do have a number of research partnerships with a number of universities where people are able to have PhD students that become parttime employees and actually get to work with that dataset to do research, but one of the things that I would love to bring up in this community is thinking about now to use farm managements systems like Operations Center, but there are other ones, as a research tool. If you think about it, we have done this problem of solving how to stream data from the field, into a database and structure it in a way that makes sense for agronomic understanding, and that's a thing, Ops Center is free to use, you can log in and use it. We also would be happy to introduce you to our university partnerships person.

But thinking about how do we use the framework that growers are already using for their data, how they understand their data, to also do the research? Because one of the things that I'm noticing is we try and get farmers engaged in these ecosystem services markets, it's like the language that's used in the models, that's modeling carbon, is entirely different than the actual implement that was attached to the tractor. If you're talking about tillage, what percentage of the field was still residue, that sort of thing, it's not directly translatable to what the farmer actually has data on.

So one of the benefits that could come from doing more research inside the actual farm management platforms is you have this data that's automatically labeled in a lot of cases, in the language that farmers use. And if we start building that into how we think about these products, it might be more translatable to the growers themselves.

So I want to throw that out there, and I think we've talked, those of us who are nerds in the company have talked a lot about how do we potentially eventually produce anonymized useful clean datasets, and I don't have an answer to that yet. I just want to say it out loud. But there are some of us who are talking about it. And so it's a question of how are we serving our customers for that, how are we serving the industry with that, and how can we make that justification that it's worth the risk that there is of this data privacy concern from growers.

But one way that you as researchers could do it is you can actually ask for direct permission from those growers to get access to their data in the system. But if we wanted to produce a broader anonymized dataset, it would take a lot of work, so it's something that I'm open to hearing what would be most valuable there, how could we do that in a way that is respectful to farmer data privacy, but also helps move things like this forward.

So those are some of the questions that are on my mind as we go into this discussion.

(Applause)

DR. GUAN: Thank to all the speakers. Let's get started. I prepared these three questions, but I also encourage all the audience here, if you have questions, to please go to the microphone at the end of the stadium, but also online, for the online audience, if you have questions please post it, and we should have time to entertain questions from online.

To start with, I really appreciated the conversation, but trustworthy is one of the topics that we today discussed. What is your trustworthy checklist, if you actually plan to use any AI tool or algorithm in your work? And maybe another similar question is what if you wanted to use these tools with your -- ask the tools to be used by the stakeholders, maybe the answer is slightly different.

DR. NAKALEMBE: For one of the things that's happened in our field is the fact that every day there's a new map, every day there's a new dataset, there's a new foundational dataset, a new crop type map, a new global cropland map, et cetera. And at some point I was like I'm done making these maps, but one of the things that I ended up doing, which I also do a lot of labeling like David was talking about, we used to have a label café in our group -- was to create what I would consider as our benchmark label dataset, and so if you produce a new map, we run it through, and if it's better than the other one then we're like, okay, now we're going to switch and use this as input for our crop conditions assessment.

So having an independent way of evaluating the dataset, not as it was evaluated by the people who produced it, going back to the point of geographic biases -- if you look at a lot of these maps they do really well for the United States, for France, we have like 99 percent accuracies. But then start to look at Malawi at like 35 percent, et cetera.

One was trying to reduce the work that I do if somebody else is doing it, but at the same time pointing out where things are falling short in order to build on onto other things. Then going to stakeholders and users like ministries of agriculture, I work a lot with ministries of agriculture like in Kenya for example, to improve maybe the yield modeling workflow or their crop conditions assessment work. I walk them through these steps of evaluation, so that they understand it, but that they're also able to reproduce it. And then we make the decision with them together. This enables access and trust as well as they understand the limitations, because also the ministry will be like, well, we want our own map that we created ourselves, and then you explain the cost, and when you weigh the cost they understand that it might be better to use the next best.

But it's not always that straightforward. For something like yield, to collect a benchmark dataset for yield, you'd have to access the kind of data that Emma has, for example. It's also something you need to do seasonally, which is why NAS data become really important for evaluation. So it goes back to investment. Trustworthiness requires that you invest in infrastructure and processes that can make it accessible, understandable, and verifiable. Otherwise it's like haiku of my model is better than yours, because I have .5 percent better than yours.

MS. BASSEIN: Yeah, I think it's very similar. We do a similar thing. If there's someone claiming that they have a new model for detecting agronomic practices, because we have a large dataset we can go an create a test set for ourselves for that, we run it against that, we say how well this is doing, is this going to help us fill in the data gaps that we have?

My particular team, again, we're trying to connect farmers to ecosystem services market, which means mostly we're trying to fill in field level historical practice data, and then verifying and going forward, and so one of the big questions we have is we're asking growers to go through and fill out this data so we're like if we fill in this data, does it actually save you time? Is it close enough to correct that it's going to actually make the process better, or is it going to frustrate you? So it's actually more of a user experience question in a lot of the cases that we're doing, for things like cover crop detection or tillage detection, if we're going to surface that information. It's a design question more than it is a science one.

DR. LOBELL: I don't have much to add. I think I'm closer to Catherine in terms of the type of user I am, but I'm not really on the front lines. Having benchmark datasets is really critical for all of us, and it's getting harder and harder to be sure that the benchmarks weren't used to train the models that are being put forward, but I think there's still plenty of tasks where these models just fall really short on any sort of local benchmark.

DR. GUAN: My next question, that is where do you see the biggest bottleneck to scale AI in agriculture, but also potentially what would be the opportunities to actually advance AI use in agriculture, or advance climate outcomes. Bottleneck, opportunity, whatever you want to talk about.

DR. NAKALEMBE: Mine, I showed it in my slide, one of the -- David has talked about it, too -- in the geographies that I'm interested in where you can learn a lot of things about adaptation, where the huge gaps in adaptation, mitigation, huge investments that translate into enormous human suffering that could be avoided with better predictions, et cetera, the biggest missing thing is label data, where I could pretrain my model on data, let's say, in India, or geography that's similar to the place that I'm studying, let's say somewhere in western Kenya But in order for me to evaluate it and know that it's actually performing well for that region, I would need to collect the data. So I think one of the biggest bottlenecks is just having data to be able to run any of the, even the most basic model. I've been trying to do rice mapping, for example, in Ghana, and I spent hours labeling -- rice is easy to interpret so I can create the data, but it wouldn't be that difficult if I was trying to do it somewhere else where it exists, or if there's some method and process for sustaining it.

And in terms of opportunities, I think another dimension, something that's been happening more and more, there's a huge push of chatbots in agriculture. Like advisory, they call them e-extension or something like that. These really, really worry me, particularly in the places where I'm looking, not that I'm against being able to type when should I plant and you get a response. And the reason is, because I fundamentally understand that the underlying data for informing that response that you get from the prompt, is highly questionable. Some of the most basic things we're doing really bad at mapping, but then the chatbot gives you some -- it always gives you a response, right? So that, you know, some models give up when they can't. But these still give you a response. They'll tell you to plant on Tuesday, and I hope you pay me back if I lose my money.

But that creates a very big problem, because this ease of getting a response, which is like getting some sort of direction, puts people in a very risky situation where the consequences are really huge. A forecast for a flood, you could say there's going to be a massive flood in the next 24 hours, and then it doesn't happen, but then somebody rushed their harvest and everything, or the opposite. They didn't do anything because they got the wrong response.

So underlying, whatever the layers that go into building the model that translates into what the chatbot will produce, is very questionable, and you could use ChatGPT and you will still get a response, and this is very problematic.

I'm trying to say something that's hopeful. Opportunity, right? I think the opportunity also lies in us being truthful and honest about what it is that we're able to measure with certainty, or giving uncertainty around model performance, et cetera. And explaining the limitations. I'm very proud to say that my model is not the best, but I developed it for this region, it works well because of this and this, and it doesn't do well at this and this. And then trying to maybe institute those types of things in how things are shared and communicated I think is much more important because there's such huge risks associated with giving people some information that completely is misleading.

MS. BASSEIN: It's wild how different the scenario is in largescale U.S. agriculture. Because a lot of our framers have tons and tons of land, so if they're going to try something new they're going to try it on one field this year and they're going to see how it goes, and they're going to test it out, and they're going to get to see that and the risk of that failing is pretty low. And then additionally, if they try it across more, there's crop insurance, like there are so many things that our farmers in the United States have that make adopting technology so much easier, and yet they still are pretty skeptical about it and don't always like to adopt it. I think of that trust thing that David was talking about in his opening comments.

I wonder how there is that ability for collective learning across smaller holders. Is there a way to be we're going to try this on a couple farms, but we're going to help each other out if -- and better, actual funding to support those farmers?

On my side, when we talk about, again, I'm looking at very different scale, carbon market type of situation. One of the things that we need is just we actually don't really have great measurements for soil carbon, and I think that this is like one of those fundamental things that you say over and over again. Your model's only as good as your input data, and our soil carbon measurement has such high error bars that creating a model that actually represents what's happening in an agricultural system -- it's not meaningless, Kaiyu would come over and strangle me if I said that, because that's an area that he spends a lot of time on, but it's hard when there's so much error there.

But, yes, I think there's a ton of opportunity. I know we're only scratching the surface on what we have. Our database is large but it's also messy, so there's a lot of opportunity for us to continue to do that and get better insights out of it.

If you look at how technology adoption in the United States has changed agricultural yields, we hit record agricultural yields for corn and soy this last year at close to almost 50 percent higher than 30 years ago in corn, and 40 percent higher in soy, per acre. So when you think about the opportunity for technology adoption, AI is just part of that. There's just so much there, still. There's still more in the United States, and if you think about how that's true in other systems.

DR. NAKALEMBE: I was just going to add that I have a chatbot for soil carbon.

(Laughter)

DR. LOBELL: I have an unfair advantage here, I guess, going last. But I would say one of the bottlenecks I think related to Emma's last point is there is a lot of overpromising going on right now in terms of the AI community, in terms of just generally the carbon markets and other things, where people are setting expectations that are dangerously high, and I think when the science is not there, that runs a real risk of burning trust over time. So that's the negative.

I think there are a few big opportunities that didn't come up yet. I think one is, like I've kind of worked both in domestic and other developed countries, and then developing countries, I think it is the case that in the United States it is also very hard to get farmers to use these things, partly because farmers are already pretty sophisticated and optimized, and providing value to them is quite difficult. But I do think the policy, seen as a lot more rudimentary, there's a lot of opportunity to use AI to improve and design better incentives that farmers face, that they can then reoptimize.

So in addition to the type of precision agriculture that Emma's talking about, I think there's a space for precision policy, or at least something less crude than what we do now, that could really improve in the United States and Europe and other similar geographies. There's also that opportunity in developing countries, but I think there the opportunity increase on farm management is also bigger.

In the developing country context, I think one of the big opportunities is an ability to understand better what farmers actually want to know, because of all these chatbots which may be giving them imperfect answers and that may be a risk, but the upside is you get actually maybe better than ever before an understanding of what the research priorities should be, or what the modeling priorities should be, and I think Plantix, I'll go back to that example, I have collaborated with them for a long time, and I think it's an example first off of how good labels can lead to a good product, and they did that. But also now because this product is very well used, there's a pretty good understanding of which diseases farmers are dealing with in which years and which locations. And all sorts of things like that could I think feed back into research in a way that historically we were sort of -- there's more guesswork involved than what would actually be useful.

A final point on opportunities I would make is that Catherine and I have been saying over and over that we need better labels. I think there is a big opportunity to improve labels quite rapidly with a lot of other technologies that are coming along, so it's not just about doing the very sort of traditional measures, but trying to do things that can really accelerate that. I think there's big opportunities there overall.

DR. GUAN: I want to follow up David and especially this discussion, you particularly mentioned about the policy. If you work with individual farmers, it's very much bottom-up, and then if you work on policy it's very much top-down. Probably changing one thing will quickly have much bigger impact. Could you please expand that, elaborate on that a little bit more, and give some ideas like how AI actually has been used and what do you envision these type of application of AI use for national policy of agriculture, relative to environmental or even climate, should be pursued.

DR. LOBELL: I think actually, I don't have any great examples of things that have already been done. But I think the types of things that -- I don't know if you've presented anyway -- but for example, in the United States, you could imagine policies that are trying to incentivize cover crops, butt just in the locations where they actually will help improve water quality and improve farm productivity, but not in the locations where they won't. And we would use AI to sort of understand exactly when and where those practices should be done.

Similarly with an example from the EU, there's a lot of the new rules in the EU, they're trying to promote more sustainable agriculture, but what they have is very coarse rules like every farmer should have a rotation once every three years for every field. And the farmers push back on that because they don't want that kind of constraint, but the reality is probably there's only about 30 or 40 percent of the cases where that actually is really important, and the others you probably are best to do something else and focus on other things.

So historically, policy has not been very informed by our scientific understanding, other than kind of what on average might work. Even then it's not perfect. But I think if Kaiyu hasn't sort of talked about his own work, I think that kind of thing where you're getting down to very detailed understanding of where and when things should happen, then you provide the incentives to do it, and then that works much better, and taxpayer dollars are going much more efficiently towards the things that we want.

DR. GUAN: We can entertain a few questions on the internet.

PARTICIPANT: I'll give you two questions. The first is specifically for David, and the other is more general. David, you mentioned that agricultural models benefit for decision-making might not bear fruit until years later. Are there ways of collecting historical data, back-casting kinds of modeling, to increase the trustworthiness of AI?

And then the second more general question is has agricultural AI data shown any new or unique perspectives on how to improve crop yields? For example, does the data show crop rotation is more important than fertilization, or vice versa?

DR. LOBELL: I can start, I guess, before I forget the two questions. Actually I've already sort of forgotten the first question. It was about the hind-casting.

I think the short answer is those are useful exercises for scientists to do to understand kind of how long you should expect it to take and what the risks of things not working out, and those sorts of things. But from a user perspective, they're going to suspect that things have been sort of tuned to historical stuff and is kind of looking better than it is, and that's fine, I think what I would do, too. It's like any sort of investment strategy that people show, like if you had invested this way the last 20 years it would have done great. So it only goes so far. But I think from a scientific perspective it's important to do those things, to really understand the risks and the timescales involved.

I do think, to the second question, that's where my research group is doing a lot of work trying to understand exactly which practices help where and when. Rotation is something that seems to be underappreciated in a lot of contexts. Even in Africa we can see early signs that the places that are rotating are doing, are benefitting more. We see that in other regions as well. Cover cropping has like a mixed record, I would say. I think we're starting to get those kind of insights, and the more that the data come along, the more we'll have confidence in those insights.

MS. BASSEIN: For the yield question, like I just said, yield for row crops in the United States have come up incredibly. Obviously, this is not true for everything, it's not true for everywhere. But one of the questions I'd love us to start asking is how do we improve resilience rather than yields. So it's not is your peak yield this year higher than it was last year, but in years where there's intense rain and there's years where there's intense drought, how are you doing relative to the other folks around you? And there is pretty strong evidence for things like reduced tillage and cover crops and those sorts of things to help resilience.

And the way commodity markets work, that's actually a really big economic benefit for farmers as well, because if you do well in the year that everyone else does well, prices collapse, like you're seeing right now, where there were record yields across the world. That's actually not great for farmers, but if you can do well in years where others are struggling you actually get a better economic reward for that.

So looking at resilience rather than total yield. And then to the back-casting, we hear growers say I need to have seen someone in my county do it, before I will. It's not a modeling exercise at all; it's I need to see someone else in the ground that's within 10 miles of me. It's really specific that growers want to see that example. They want to know that it applies to them, they want to know it applies to their soil type, to their terrain, to their climate, everything.

DR. MCKINNON: Karen McKinnon at UCLA. The first question was just about AI for high-value crops, fruit, nuts vegetables, things like that. And then the second as a Californian, AI for improving water efficiency, reducing water use for agriculture if there's advances either happening there or even if there's possibilities.

DR. LOBELL: I think for specialty crops in general, it is an area where you'd see a lot of adoption of AI just because of the economics making more sense, I would say. But my guess is that most of the useful stuff early on will be ground-based sensors, and it's very difficult -- Catherine can maybe speak to this, but from the type of stuff we do with remote sensing, it's been harder. We have work going on. But I think for specialty crops, things like rapid photo-based assessments or soil probes and things like that feeding into management is definitely an area where there's a lot of interest and activity going on.

In terms of the California water efficiency, those are kind of also very much based on perennial crops, I would say there's opportunity, but again I think farmers are already pretty sophisticated in what they're doing, and a lot of it may come more down to policy incentives and policy being smarter through AI of understanding exactly where water should be used and for what.

DR. NAKALEMBE: I can add a little bit. I think the thing that comes to my mind is some of the additional opportunities is around being able to create maps and products for places where we weren't able to do before. And this is not just AI, it is availability of remote sensing datasets, compute, which allows us to process more and more, and then storage. Because before we were unable to do a lot of this. You'd have a very good computer with your own cluster, et cetera, so being able to do things in the cloud allows me on my computer here to be able to keep labeling and mapping an entire country. So that is a great opportunity.

But AI methods around mapping and modeling, they allow us to create more products along the chain that could feed into the management decisions. So it's not like -- I think sometimes some might think of it like you type into ChatGPT, your end result is you want an email. So you do it, you get your email draft, and then you send the email if you're happy with it. But for agriculture, in order to predict yield for farmer X in Illinois, you need a crop type map, soil moisture, all of these different variables that need to come together into some -- it could also be different models, and they're aggregated together into another model, that then gives a prediction. So being able to do and combine these very different datasets to have an improved estimate of something is a great opportunity space, but there's still so much more to do towards making it 100 percent.

And I loved David's point saying precision policy. We were talking at lunch at one of our tables, if you don't measure it, you don't know it doesn't exist, but the idea, I think David worked on yield gaps, for example -- having a yield gap analysis, and how that has changed, as have some places stagnated, have they gone back, gives a policymaker a direction and a location for where certain investments could make a big difference. So being able to produce better methods, better products, to inform those policies, make them more precise, I think is a huge opportunity space.

And I really like working -- I don't like the word top-down, but it's 100 percent true, but it helps you see more broadly where the potential is, and it can guide where you can have the most impact.

DR. GUAN: As we close out this session, I want to end with some aspirational high notes. I want all three of you to think about, we are in the early stage of AI, we all know this, and thinking about what's going to happen in the next 10 years, what would be one thing that you think would be very exciting, you're actually looking forward to seeing it happen in the agricultural space that actually benefits the climate. So maybe name one thing, just to be inspirational.

MS. BASSEIN: When I look at this from the perspective of what we're doing at Deere, one of the things that I'm really excited about is the potential for autonomy, and I want to explain why. One of the reasons growers cite for doing large monoculture operations is it takes a lot of people to do diversity. It takes a lot more management. If you're managing different parts of your field differently, that's really complicated, it's really hard to do.

The thing that I hope that autonomous farm equipment as well as automation in the settings that are adapting to different conditions in the field enables over time is the ability to have more precise planting, more precise management, and just better overall opportunities for creating diversity in our agriculture system. So that's something that I am aspirationally hopeful for.

DR. GUAN: So automation to facilitate diversity, diversification of the system?

MS. BASSEIN: Yes.

DR. NAKALEMBE: Mine is going back to my chatbots. And my hope is that, let's say 2030, 2050, that my farmer Mary that I shared in that photo at some point would get a text message that tells her that the forecast says there'll be a flood. You've been approved for financing, you can go pick it up here, and do nothing. That's my ultimate chatbot. Not one that tells Mary flood coming, good luck to you. Something around that. But to get to the point where financing approved, you can pick it up here, means that somebody has invested or there is a system where there's some verification, it is worth putting money in this area, and that would require a lot of improvement going back into the back end of this process.

DR. GUAN: Last but not least, David.

DR. LOBELL: Good question, Kaiyu. I would say very briefly, the aspiration would be that historically a lot of the technologies we've had have sort of been favoring largescale grain production, which is not a bad thing, per se, but I think AI has the chance to both skew things back towards smallholder farmers and towards diversified crops. We know that a lot of the reason healthy crops are more expensive is because it has huge labor requirements. We know one of the reasons that smallholders struggle is because of the difficulty of making capital investments.

And I think maybe to build on Catherine's example, I would say like an app where you would maybe not even have to do it yourself, but you would have a service provider come in and do all these activities without you having to have large capital investments, and being able to say, like in the California case, having sort of lower cost nutritious foods -- not just in California but everywhere -- I think those are both possible in ways they weren't before.

DR. GUAN: Thank you so much, and thanks everybody for the engagement. Let's give a round of applause.

(Applause)

Thanks, Emma, thanks, Catherine, thanks, David.

It's the time for doing the poll, so please log into the system again. It's the time to answer the question for this particular panel. We're going to show the screen, and then we will also talk about the results.

The question is based on the discussion in this panel, where do you think the largest opportunities lie for advancing AI in agriculture in the near term?

We have three options: option one, expanding collection of high-quality data; second, using AI to provide more localized, granular information to decision-making; and the third, but not least, enhance communication to stakeholders.

Please vote, and then we can live look at the result. We have a quite balanced one. The first one seems to be leading the effort, the data, everybody agrees the data is important. Seems the second and the third are very similar. We all recognize we're going to continue collecting the data, training the model, and then communication is very critical.

Again, thanks, appreciate everybody's participation.

(Applause)

Agenda Item: AI in Urban Planning for Climate Change Impacts & Adaptation

DR. MENDEZ: Good afternoon, everyone. It's a pleasure to be here, and again, welcome to the Beckman Center here in beautiful Irvine, California, for those that are just joining us or joining us online.

My name is Michael Mendez, I'm an associate professor of environmental planning and public policy and a Chancellor's Fellow here at the University of California, Irvine, just down the street. And you're going to be joining us today for an innovative panel on looking at how

AI is used in urban planning for climate change impacts and adaptation.

So we're going to be covering a variety of issues from different perspectives, particularly highlighting the challenges and opportunities for AI-enhanced tool support, urban design, and also the translation of data related to flooding, urban heat islands, and to understanding of risks for various sectors, including city planning, reinsurance, and the general public.

I think this is a growing area, particularly for scholars and practitioners of urban planning. You don't automatically think this being used. I can tell, before we go onto our expert panel, tell a bit of story from my perspective as a scholar and a practitioner.

I currently sit as a gubernatorial appointee to the state's regional water quality control board, and we're divided between nine regions based on watersheds, and I'm part of region 4, which represents Los Angeles and Ventura County watersheds and represents over 11 million people in that area.

Recently, late last year, we toured one of the recent fire zone areas, Altadena. As the regional water quality control board, we help implement the federal and state Clean Water Acts, so we were touring with the county to see the rebuilding process of the greater Los Angeles fires in that particular place, Los Angeles, Altadena, and see the rebuilding process and how county governments were responding to that rebuilding process and the needs of residents and businesses.

Part of that was creating a one-stop shop, bringing all agencies within the county, the massive county of Los Angeles, together to be attentive to various rebuilding processes and the bottlenecks. During that conversation, we learned of a new tool that the city planning department in particular was going to roll out. It was an AI-derived product for plan check.

That's when individuals, homeowners or businesses, wanted to rebuild a building, they could submit it with their architect into an AI-enhanced tool that they solicited through a consulting firm, and it's proprietary to the county, that would check the plan, the architectural plans, for compliance with city planning laws, building and safety, to cut down the time it would take to get approval, which could take anywhere in the normal process months to years, to their goal of 30 days.

That's currently being unfolded, but I thought was a real-world example of how various forms of AI is being used in the urban planning field.

But today we have a stellar panel of experts that span academia, local government, and NGO and the consulting world. They're going to be introducing themselves for five minutes, so I'm just going to give you their general description, and they're going to talk a little bit more about the work that they do, their position, and how it relates to AI and how they're currently using it in their work.

The first that will go up is Adam Nayak, who's a PhD candidate at Columbia University. Next we have Chris Belasco, who's the City of Pittsburgh, one of the chief information officers, data officers. And Mariela Alfonzo, who is the CEO and founder of State of Place, which is an NGO and consulting firm.

I'll hand it over next to Adam, if you can take no more than five minutes to briefly introduce yourself, how do you use AI in the work and research that you do, and how is it informing the field in general?

MR. NAYAK: Thanks, everyone, for being here. It's really exciting to be embarking on such an important topic. I'm going to talk a little bit about futures in AI for climate risk, and particularly what we're focused on is this idea of spatiotemporal planning for an evolving climate.

This is an introduction, a little bit, to me. I am a PhD candidate from Columbia, and my background specifically is spanning engineering, policy, and climate science settings, which really informs a lot of the questions that I've been curious about and pursuing, particularly in the context of insurance and thinking about some problems that we've been facing in the insurance sector in the United States.

As most folks have probably experienced, and as seen in the news, we have a new sort of climate crisis emerging that has manifested within our insurance systems. The headlines really reveal this, more specifically, of how our insurance is helping to actually prepare and finance projects more specifically that are proactive and support urban planning efforts and resilience, to climate-driven disasters.

The National Flood Insurance Program has accumulated billions of dollars in debt for the U.S. government. We also have largescale insurers withdrawing from markets in Florida, and more recently in California, and when we want to zoom in on this problem as planners and decisionmakers, we need to think about the different timescales in terms of how we can plan for future disasters.

So we have potentially the adaptive spectrum in which insurance systems are a subset -- we plan for in largescale infrastructure projects, more like a 100-year type flood event, or 100-year return on a disaster. We have home mortgages that are typically around 30 years in length. And then we have at the short term, insurance policies which are renewed annually.

At the same time, we're working across the hydroclimatic timescales. We have climate change, which is over 100 years in uncertainty for our projection horizons. We have multidecadal impacts and oscillations, such as the Atlantic multidecadal oscillation that impact our systems. We also have sub-decadal oscillations -- El Nino, more commonly. And then we have things like seasonality and sub-annual oscillations that also affect the ways in which our extremes propagate through this system.

More traditionally, if we think about how insurers quantify risk, this is usually region- and hazard-specific. Insurers will look at a given area, they will assess most specifically the vulnerability of that given area to a given hazard, and they'll use a tool called a catastrophe model to specifically look at the loss as a function of both the exposure to that area and that hazard, and the vulnerability that's expected given the return probability of event. So then the risk becomes the probability distribution over these individual loss events that we then assess more independently.

In sum, insurance is mostly priced by this idea of the return probability of the hazard, plus some buffer for uncertainty. And our question is, well, how can we map this back to our climate systems? Because from a hydroclimatic perspective, floods are extreme realizations of dynamical climate systems. And how can we use AI to more specifically consider the space and time dynamics that shape these clustered risks more naturally?

Our ocean and atmospheric processes are going to shape climate circulation patterns that then drive extremes that manifest through both space and time and are often clustered. What's scary about this process for insurers is that these space-time risks can lead to basically simultaneous losses that are unaccounted for in traditional actuarial modeling. This can lead to risk balance issues when we're balancing checkbooks.

So my work is specifically focused on thinking about placing individual loss events within their hydrometeorological contexts and using AI and machine learning tools to do this.

So there's three main areas that I want to talk about in relation to climate risk and AI, that can be really future opportunities for us to basically redesign our insurance systems, and also think about preparedness across asset portfolios. These are across understanding, projection, as well as adaptation and system redesign.

One example specifically is that we use unsupervised learning for understanding by using basically both data from historical reanalysis of our natural environment, as well as claims information and disaster aid disbursements to map the space and time dynamics associated with given disasters. Here you see basically disaster in which it's Hurricane Ike, that low pressure system moves north towards the midwest, and in the subsequent weeks we see a series of severe convective storms and tornadoes that cause subsequent damages. We can then map this back to the actuarial losses that are associated with this synoptic pattern.

We can also use the machine learning algorithms to more generally classify the different types of extremes that we're seeing from a space and time perspective. We see sequential tropical cyclones, recurrent riverine floods, severe convective storm cycles, as well as inland storms that follow hurricanes, and these systems are naturally interdependent and are shaped by atmospheric dynamics such as moisture recycling patterns, thermal convection, as well as atmospheric blocking patterns.

So our question forward is how can we use AI for our future projection and prediction? How can we consider these uncertainties in planning and decision-making?

We use a three-step process where we first look at combining signal-processing tools with dynamical statistics for heavy tails and integrating this with a real strength that we see in machine learning, which is in pattern extraction and looking at how we can forecast a low-frequency signal forward, and then condition our heavy tail simulations on such signal.

In our model, we use deep learning specifically paired with explainable AI, such as integrated gradients, to basically pair wavelet spectral analysis and coherence with interpretable attribution specifically to teleconnections for global climate variability. This helps us to look at specifically, given a set of extreme events, spatially dispersed, what sort of patterns do we see as explainable attributions that could be driving this sort of behavior?

And then we want to map it back to our systems. So we work specifically with also using qualitative interviews and working with stakeholders in the insurance sector to think about ways in which risk pooling can be informed by the ways that we consider our climate systems. So we look at risk quantification and catastrophe modeling techniques that try to improve the space and time representation of these hazards.

This is important for considering adaptive planning, as well as policy and regulation around adoption of AI tools within the insurance sector and thinking about preparedness for communities more generally. Most folks can't necessarily self-insure or buffer their individual losses with their own savings. So it's very important looking forward for us to consider ways we can redesign insurance systems to be more accessible and climate-sensitive in order to protect more communities.

Thank you.

(Applause)

DR. MENDEZ: Thank you, Adam. A great overview. Look forward to being in more in conversation with you on understanding the spatiotemporal risk and projection models.

Next up, we have Chris, and we look forward to your introduction.

DR. BELASCO: Thank you, and thanks for having me. My name is Chris Belasco. I'm a city government employee. I'm with the Department of Innovation and Performance at the City of Pittsburgh. I'm the chief data officer.

My relationship to climate resilience planning is part of the work we do, to work very closely with our division of sustainability and resilience in our Department of City Planning.

A couple of years ago, we worked together on a project that was related to doing tree planting, which you know is a climate adaptation solution, kind of a really great opportunity to help forestall predicted effects of temperature increase in climate change, and for our region, we're expecting around a 3 degree Fahrenheit increase in temperature, but that varies, different places. We'll show you in a second what that means.

We did this project with a couple of other external partners. Resilient Cities Catalyst, and they're a climate-adaptation resilience-planning company, nonprofit, and a couple of local partners including Tree Pittsburgh and UrbanKind Institute. We worked together, received a grant from ICLEI, which was formally known as International Cities for Local Environmental Initiatives -- they shortened it -- and Google.org to do this project.

There were folks from the Department of City Planning, folks from the Division of Forestry who were involved, and in addition to that, our GIS team. So I want to acknowledge all those folks for working on this.

Our effort was to try to update the city of Pittsburgh's 2030 climate action plan by having some ideas about the placement of trees as it related to both policy, community organizing, and then in data. So the data space is a little bit about what we're talking about here. I'm going to present one of the maps that we assembled by working with Google.org, where they used a computer vision model, downward on Google's satellite imagery, to detect unique tree crown in the city.

Why were we interested in that? The city collects and is almost completely ready to release a new street tree inventory, so we have a handle on the trees that we maintain in the streets, but you can see in those purple dots, there are lots of places where there's incredible tree density. The city is very hilly. There's an ordinance that prevents the development of many very steep hills and cliffs throughout the city of Pittsburgh. We call those greenways, so we own all those trees in our greenways, and in our parks. So we wanted to get a handle on the scope of trees that we actually own.

We leveraged that AI model and some climate risk data from First Street to try to understand where there would be ideal tree placements for the future. This gave us sort of an operative understanding of places to try to plant trees and forestall coming change.

Just briefly, the folks at Google.org who helped us with their model were really transparent about where it worked and where it didn't. If you look at the lower right, you can see some red dots aligned with some green ones. The risk of using an overhead satellite model is that if there are shadows, the satellite's not going to photograph a tree. So if we're using a computer vision model to detect that unique tree crown, we're not going to see them.

You can see that it did pretty well outside of that area, but what we knew from that is that a lot of those places were living in the urban canyon, so that gave us the ability to snap our existing street tree inventory to that, to help better identify places where the model missed.

So if we can get to extra time, I'm going to try to show you one slide that also helps us to think a little bit about the risk of just kind of relying on a computer vision model to help identify where unique tree crown is, because since the project, the forestry service has released some canopy estimates on counts of trees in the city, so this round number of 500,000 or so that we were able to achieve from this, and the forestry service's estimates are quite different.

What this project ended up allowing us to do was to identify places, particularly in neighborhoods that --

(Audio drops 1:05:27-1:09:55 in video)

DR. ALFONZO: Hi, everyone. So great to be here with you today. My name is Mariela Alfonzo. Just one clarification. State of Place is not an NGO. We are a software technology and advisory company. Maybe one day we'll be an NGO.

So another clarification is that I'm also not a climate scientist, and I'm technically not a technologist. But I run a technology company. My background is actually in urban design and behavior. I come from sort of the spatiosocioecology. I actually got my PhD here at UCI.

And this kind of slide encapsulates like what I do and why I do, essentially like where we live, built environment, predetermines how well we live, and this goes from poor schools, poor mobility, poor infrastructure, poor health, and of course, worse climate outcomes. It's all interrelated.

And I wanted to kind of guide you through a little bit of like why. So if you just kind of sit for a second with these two pictures, how do they make you feel? What do you prefer? What place might you avoid?

It turns out that these kinds of places impact more than just our feelings. They're going to impact our choices, our behaviors, our perceptions, and then in turn, they're going to impact the outcomes that we are all trying to change and optimize and improve, whether that's climate change or health or otherwise.

And it turns out that all of these different built environment factors also have a huge exponential cost when they are designed poorly. So what my work has really been trying to do for 25 years is quantify all this. So how does the built environment actually influence our behaviors and perceptions, and in turn, how does that translate into different aspects of value, and not just economic value, but also social, health and environmental value.

What I want to bring home today is that the built environment factors that are impacting climate change are the same built environment factors that are also impacting health outcomes, real estate values, property taxes, different aspects of crime, and all of these things are interconnected and it's really important to kind of think about things from a more systems perspective, more holistically, for a couple of reasons.

One, it helps you kind of understand interventions and developments that are going to give you the biggest bang for the buck, which is increasingly important, of course. So what's the most effective decision that can give us value across all these different dimensions? But also, too, it helps from a communications standpoint. We're all in the same space here, same minds. We're not afraid of climate change. We're not afraid of saying that.

But there are some communities in which that's not going to be the case, some administrations that we're working within that that might not be the best sort of foot forward that you want to put, but also just from a resident perspective, that might not be the thing that's most salient for that.

So the communication might be more, oh, well, I want to be able to walk my child safely to school. It turns out that the same built environment features that can facilitate that can also facilitate improvements on climate change, and that's really what my work is and I want to tell you a little bit more about how we do that.

Before I do that, I wanted to quickly set the stage. I know we talked a lot about AI. We've kind of talked about this in all these different frameworks. This is just a simple visual to say there's multiple different aspects of AI, and we use three out of the four, although we're aspiring to use four, and I'll tell you more how in a second.

So first we're using computer vision. So I mentioned we collect data on the built environment. So obviously we need to -- we are using computer vision and large language models, which I'll talk about in a second, to quantify that data from images. So these are using convolutional neural networks, which I used the analogy of somebody being blindfolded, an elephant, everybody is touching different things, well, all the neural networks are getting different layers and together then they can piece together what an image is, and that's kind of what we use.

And then on the large language model side, you guys all know this, this is based off of transformers, kind of like understanding all the words all at once so it can then predict the next word.

So CLIP is a model that we use that's produced by OpenAI that combines these two approaches where you're basically able to use a natural language prompt and it's able to understand whether an image that you feed it actually has that feature.

So we have developed our own little, I don't know, we call it IRIS. It's not really a robot, but it has different sensors and images or cameras, rather, and we drive that around. It's our own version of Google Street View. We collect the images, multiple images per block, and then we use CLIP to be able to understand whether or not a certain feature exists. We collect 150 features: sidewalks, trees, benches, crosswalk markings, curb cuts, all these different things, and then what CLIP does is give you a confidence, how confident it is that it actually is there. Very important thing. It doesn't actually tell you it's there or not.

So like I said, we have 150 features. We put them together in a score that measures the built environment. We divide that into 10 different aspects of the built environment. We show that in a software. It kind of helps you understand why you got the score you got. We're showing it spatially as well and heat maps.

We then have a sort of a SimCity tool where you can play around with the data and say, well, what happens if I add trees, sidewalks, benches, what happens to your score? And more importantly then, well, this is just future aspirational. We can use generative AI to kind of visualize this as opposed to just show you 50 versus 70.

And then we use machine learning to basically understand how these different aspects of the built environment influence different outcomes so that we can help identify what specific changes and where to make to optimize any of the outcomes that you want, whether it's climate change or health or economics or all of the above.

And then what we do is forecast based off of the scenarios that the users use what is going to be the actual impact of those changes in real time, whether that's economic or how much it's going to decrease collision rates, things like that.

Quickly, because I'm somehow very much out of time, we did this in Philly. We really wanted to kind of create a narrative of the built environment influences all these things, especially from an equity perspective. This was during COVID. So what we found was that for places below versus above average state of place, the index of our built environment, we saw huge differences in COVID rates across positivity, hospitalization, and fatalities. We saw differences in terms of what percentage were in the floodplain versus not. We saw differences in the heat differential, specifically within the parks and public spaces dimension within the index.

We saw differences in crime, in this case violent crime, huge differences again, below average, and then we also have -- that was just all for Philly, and they were basically able to weed this narrative, hey, the built environment crosses all these silos in our departments; maybe we should work together, and then these are some other models that we've done where we've looked at differences in property tax, the diabetes rate, vehicle collision risk, vehicle miles traveled.

Again, the whole point of what I wanted to tell you here is that the whole is bigger than the sum of its parts, and you need to know your audience and this helps you kind of tailor that message to your audience. You can get more effective decisions and get buy-in, trust, approvals and funding for those projects. Thank you.

(Applause)

DR. MENDEZ: Thank you for that, and apologies for calling State of Place an NGO, but I can definitely say your company is a technology company focused on the social good. So thank you for that presentation.

I was excited of the panel presentations that you all did, and you know, oftentimes we think of AI, generally as a negative narrative that it's replacing people, it's not human-centered, it's not realistic of marginalization or different equity measures, or conditions. I was surprised and more happy to hear of that human-centered or people-centered approach you're doing to ground truth your models, your projections, your products across communities, particularly -- which is a particularly important aspect of urban planning.

So I want to hear from each of you the first question is how are people contributing to the refining the revisions and updates of these models or products that you're containing, where do people fit in into that process?

Adam, if you'd like to start.

MR. NAYAK: Yeah, that's a great question, super-important. I think generally speaking even just entering into this work, the motivation around even looking at insurance systems in the first place was this idea of people and how we finance and provide basically the necessary capital to support communities proactively.

Our process specifically actually started with just a series of conversations first with folks in the reinsurance sector as well as folks across settings in both governance and in other parts of the industry that really informed the ways that we were approaching critical failure points, and one thing that we're really interested in with our analysis is this sort of systems perspective of how can we can take data and information that is widely available and translate that into sort of actionable science in a way that gets at the root sources of these sort of problems that we're seeing and what's causing accumulations of debt, what's causing failures, and what sort of adaptive mechanisms can we move towards in the reinsurance sector, in city planning, in infrastructure and asset management that are considering the ways in which our system isn't adapting those systems proactively, and that all involves thinking about people and communities.

DR. MENDEZ: And maybe a quick follow-up question; as a scientist and the work that you do, we don't really think of scientists or engineers doing these qualitative interviews. How has that process been for you in terms of the data you're gathering and incorporating into your projections?

MR. NAYAK: That's a great question. I think at least how I was thinking about these problems in the very beginning was motivated by the fact that I did some work in climate policy in D.C. and realized a lot of the optimization algorithms, a lot of the problem-solving techniques and decision-making frameworks that we were using it as engineers don't map back to how people actually make decisions. So I think the idea of integrated and mixed methods frameworks that sort of help us to better inform our science are really critical, and the way that we integrate those pieces involves both systematic detailing of qualitative methods in both across policy documents, across literature that's already been published, and important memos that have been released, but then also like engaging with those folks in ways that we can actually design our engineering and our algorithms and the ways that we're approaching data analysis and geospatial analysis in a way that is more holistic and gets at those specific problems.

And so I think that that is something that I hope to see a lot more of, especially with a lot of scientists moving in this direction with a lot of interest in community-based practice and community-based building of new insurance systems has been something on the horizon. So I think it's really exciting.

DR. MENDEZ: Thank you. What a great perspective that you bring into that field.

And then for Chris and Mariela, again, you know focusing on that ground truth being questioned and understanding that these models, these products are not being developed in a vacuum, and you both mentioned some aspect of livability, community wellbeing, or equity. So how do you translate that into data and metrics of your products?

DR. BALASCO: Yeah, sure. The translation of the quality and the livability back into the work that we're doing kind of starts both organizationally; I think that doing this project allowed several departments that had not kind of previously coordinated to do some renewed emphasis on coordinating, deploying a scarce resource, which a tree is, not because of planning it, the tree is expensive, but the work to plant a tree is expensive.

So figuring out ways to make sure that we were doing it in the right place, which actually just does involve ground truth and going to the residents and making sure that they are willing to take a tree in front of their house, and if they're not, then we find another place to put a tree, and that's kind of their right.

In one instance, the most recent planting involved coordinating with our Department of Mobility and Infrastructure, they manage our right of way, they manage our sidewalks, to also put sidewalks in in places that the sidewalks had degraded. So the tree, thinking of some of the different built environment aspects that Mariela is talking about, the tree was part of a more holistic strategy with another department that was heretofore not really coordinated.

There was one other example in a separate data project that I am doing in another neighborhood that also has a high socioeconomic need, and we engaged with the residents. We brought the IR heat guns, the laser heat guns, and we talked about trees and what shade can do. Then we introduced the heat guns and we invited them outside. It was actually a cloudy day in Pittsburgh, surprise, I don't know if you know this, but Pittsburgh is like the second cloudiest city next to Seattle.

And so we had, it was a cloudy day, and we still had them kind of, invited them to look at, you know, the temperature with the IR gun in the shade of a tree and where the sun would have been peeking through on a sunny day, we're still seeing 10 degree heat differences there, Fahrenheit, and also taking some measurements on a sunny day, and we were seeing 25 degree temperature differences, and I showed them that, too, and said this isn't the same weather that we had.

But it can get really visceral very quickly, which helps to leverage someone's interest and activate a willingness to be part of the planning process then.

DR. ALFONZO: You saw the first slide that I put, right? So spatial justice is kind of core to my work, as is kind of a human-centered approach. So we do this in like basically these mixed methods, so on the spatial justice side, we're trying to quantify that we're trying to make that very visible, both like spatially but also from our modeling and being able to say like, oh, these are the implications of divestment in the built environment that is structurally inequitable. It's not just from a morality perspective. There's real costs to it, so that you can then speak to that in a way that it makes it easier for decisionmakers and policymakers to justify that expense.

But the other thing that we do is we work directly with multiple stakeholders. So like when we work with our cities, we're trying to understand from them what is the problem that we can help them solve. That's why I say I'm not a technologist. I created State of Place as a like academic then turned entrepreneur, because I saw a problem that I couldn't address within the confines of traditional academia, and so I -- like you made your solution based on a problem, and yet, I've never been an urban planner. I have never worked in the city.

So that means I'm always constantly trying to understand from the city's perspective like how can I actually help you not create a solution that's going to create more work for you? So that's an important piece, but the other thing that we do more recently that I've been wanting to do this for a very long time is try to work directly with the stakeholders themselves, residents.

So we recently did a workshop in Sacramento that was a community-based participatory action research framework, which basically meant like the stakeholders were sort of cocreating the research or at least the next phase of research. So what we did is we deployed our little robot IRIS, collected data for the built environment for 2,000 blocks within this city council district, and then we presented them the data.

I showed them the software. I showed them their scores and like, okay, what do you think about these scores, because while we're collecting objective data, A, we know that these are not telling us what's there, they're predicting, they're probabilistic models, so they might be wrong, but 2, even if they're right, perceptions don't always mirror reality. So that in and of itself is a datapoint that our customer, which is a community-based organization, needed to understand, too. Because if there's a disconnect there, why is that and then how do we work to address that disconnect?

Then we also did essentially like a capacity-building workshop where I did a much longer version of what I presented here on AI, on research 101, like how do you connect the dots between the built environment and your lived experience, and then after that, I was like, okay, now, what are the outcomes that you actually want us to predict? So we went in there, we hadn't done any of the forecasting, and now the next phase is, okay, now these are the quality of life outcomes that we can go, go model, bring it back, and then we can figure out, well, what are the specific built environment changes within these blocks that can help them achieve those outcomes.

It's a very different approach than typical, like, technologists, I would say.

DR. MENDEZ: That's great, and I really appreciate everyone's approach to understanding the different perspectives. The contextual issues going on for whatever stakeholder or community that you're working with, understanding that if we look at technologies as pictography and mapmaking, those are abstractions, and obviously AI is another generation of an abstraction and the need to loop back in and have feedback loops and, as you mentioned, have the systems thinking approach to understanding the contextual environment in which you're dealing with.

With that in mind, what opportunities do you see for AI for climate action and for urban planning in general? We'll start with Chris.

DR. BELASCO: One of the opportunities I see is when you think of AI based tools as intermediate inputs. You know, we heard early on in the talks today about like putting a human in the loop. One of the ideas that we were able to kind of bring about here was to make -- to obtain an estimate of a resource that we have that we are not otherwise able to get. We can't send people into the steep hillside to find out how many trees are there.

So what we wanted to do was to find another way to rely on machines to do so. Another really interesting model that Mariela talked about would be able to help us identify the number of retaining walls that were built in the 1960s that are going to start to age in the city where the recordkeeping wasn't so great.

So the concept here is, okay, let's bring that into a decision-making tool that then allows to prioritize or inspect different assets that we would want to better manage. So I see these as pieces of a whole that are not in our case, you know, that are part of kind of like descriptive inference that would allow us to go out and do interventions that would help make sure that people are safe or that we're improving climate resilience.

MR. NAYAK: Yeah, I think there's just inordinate potential I think across the board both in one area that is particularly interesting, I think, is thinking about ways we can better integrate ML and AI into current catastrophe models and the ways that we are modeling hazards in a way that is also reflective of how first how regulation and policy really shaped the ways that folks in the reinsurance and insurance industry are making decisions. There's a lot of importance around validation, a lot of importance around explainability, and so making sure that our methods match the needs of the sectors is really important.

The potential also for climate driven and climate informed mechanisms for reinsurance, for parametric insurance, I think are really massive and could provide a lot of support earlier on to folks particularly for even forecast-based insurance, getting payments to folks on the ground faster in a way that enables emergency management practices much quicker, and I think generally speaking, we have seen a lot of improvement in our forecasting models in weather, as well as in emulators as folks mentioned earlier, and I think that this sort of speed and enabling that speed can really improve our systems practices from a modeling perspective.

DR. MENDEZ: If I could follow up with that, I know a good majority of the data they use comes from the insurance industry. They have their own datasets. But also some of it's from the federal government, from places like NCAR, the National Center for Atmospheric Research. How are cuts or blocking of data affecting the work that you do and your colleagues?

MR. NAYAK: Yeah, it's definitely I guess -- first of all a lot of publicly available datasets, particularly reanalysis datasets and climate projections, are public datasets that are crucial to the work that is being done, and so cuts in this sector and also insecurity for folks who like are doing the critical work to maintain these systems continue to improve model projections, continue to get the forecast out. If folks are facing job insecurity and they aren't able to perform those functions at the best of their ability or even feel that they need to switch jobs, that creates a slowdown that really will not enable progress in a way that is proactive and really will potentially enable a lot of tools for a lot of people.

The models rely on public data. There's also proprietary data for sure. But those public datasets are widely used across the board, and I think understanding them, understanding their various limitations and uncertainties, this all comes from science that is really important to be open access, publicly available. That's also true of these insurance datasets that are proprietary to in order to best enable future systems to better protect our communities, we do need to understand how the current systems are working, and I think that's a huge limitation as well that I think there's a lot of folks really concerned about these issues and ways we can better promote open science.

DR. MENDEZ: Thank you for that, appreciate it. And Mariela, the opportunities you see for the next generation of AI urban planning, urban design.

DR. ALFONZO: Yeah, some of it I spoke to, but I think for me what's critical is models that connect the dots, and some other folks have mentioned this kind of not just predicting the bad outcome, but telling you what to do to prevent it or what to do to mitigate it, and that's one of the things that we've always tried to do, because in urban planning, that's the point. You're like how can you make it better, not just where the traffic is. How can we reduce it, right?

And I think that with especially because of the data that we're collecting, we're starting to get more longitudinal data, more time series data, both on the built environment side and on the outcomes side. On the outcomes side, that's been there, but on the built environment side, the type of data we're collecting historically, you just don't have at scale, because it was impossible to do with humans. I did it on the ground 20 years ago.

So being able to have that, which we're now building, we now have D.C. metro over the course of 10 years, now we can start to get really into just more robust models that help us understand now just how much of the built environment is influencing this place versus that place, but how much of the change in the built environment that we saw over a certain amount of time, how much is that influencing the change in whatever the outcome is that we're modeling, and that just statistically speaking just helps us do more and understand that problem more.

And then on the generative AI side, and I don't mean like LLM generative, I mean like the GANs generative, that's -- I mentioned that earlier, like being able to do these SimCity type models that are not just predicting here's a 30 or a 60 statistically like as a nerd, that's exciting for me, but most communities don't care. They don't understand. I mean, they care, but they don't fully get it.

So being able to show visualize their block and how that's going to change, both in terms of the built environment and the outcomes, I really see a huge opportunity with that, although I have to think about how much water we're going to be using to do it. But I do see that that can really help move the needle from a policymaker perspective, from a resident perspective, and start getting people to not fear change as much.

DR. MENDEZ: And in our prep call, we did talk about the issues of water and energy use as urban environmental planners, environmental engineers. That's a key aspect. So as the field of urban planning moves forward, we hope that there can be stronger equity considerations in creating these data centers and particularly for the communities around them that are suffering from some of the water and energy and polluting uses.

But I would like -- I have one more question, and then we're going to open it up to the audience and then

Slido, questions from online as well, and in terms of we talked about the human perspective about it being a feedback loop of changing your models, but beyond that, what would you say makes trustworthy climate data, environmental data, information trustworthy?

DR. BELASCO: Sure, thanks, Michael.

From the local government official perspective, I suppose, I think that residents are kind of -- a lot of what we're talking about here as you pointed out, you know, maps to like they kind of reflect these abstractions of reality. You know, trustworthiness is kind of explainable and relatable.

So the closer we get to the resident, the more that kind of what would they understand about what matters to them, and in the case of the shade of the tree I mentioned earlier, you know, the more that the outcome that we're trying to understand is something that kind of suits their frame of reference, the more kind of useful that modeling decision is.

Now, we're talking about all throughout the day, all manner of kinds of disaster modeling and agriculture and water and fire, so the different things mean different things to different people in the country, but how we get to the resident if we're thinking about that, trustworthiness is really about what are the people that are helping to use those models, thinking with them in mind, can that be shown, and how does that explainability kind of come back to what they know or what they can refer to?

DR. ALFONZO: I think a huge part of it is transparency. I think many of us have been talking about that throughout the day. I mean, if I'm going to tell my city customers that it's like, oh, yeah, there's going to be a tree, there's going to be a sidewalk, and 100 percent -- like, no, they're going to be, oh, this block, you told me there was a light and there's no light there, and then you just completely -- we've said this -- you completely lose trust in them.

So being able to have them understand there's a confidence that we're predicting this. This isn't detecting it, right? And I think that people don't fully get that yet and just being able like in our software, we're updating it and we're going to put the confidence, because before, we used to do this data manually. So there was error, but people just understood people make errors.

Now with AI, they just don't, they don't fully grasp that, so that's really, really important, to just be transparent. The other thing is that when I first was talking to a data scientist like 15 years ago, they were like, oh, you don't, you can't, I was telling them about my conceptual frameworks, like, you can't do that, you're going to bias the model. I was like what are you talking about bias the model? Like we need to understand what are the factors that are likely going to be influencing this outcome and then go grab that data and then feed it into these models to better understand, and dah-dah-dah.

I understand where they were coming from, but that creates, much more likely to create a black box, and also that means that not only do we not understand what the relationships are that predicted something, it goes back to what I said before; how do you then figure out, well, what levers do I pull if it's a black box?

So bias in the sense of like theoretical like frameworks I think are really important, and that way you can help a person understand how you came to those decisions, why you chose the data that you chose, and they can poke holes. They can say, oh, well, you didn't consider X, Y, Z, and that's fine.

But I think that like not being afraid of the bias without a theoretical model is doing to influence or sort of frame a problem I think is really important. I don't know, that sounds very basic, but like as a social scientist, like that's how we start, and that's not always the case in data science and AI computer science.

DR. MENDEZ: This idea that this data process is apolitical, neutral, and the metrics that you pick have no bias in them in the first place, but just by the very metrics that you pick has some level of preference, a level of likeability, one of my favorite quotes from sustainability science at Donella Meadows is we measure what we value. I think that goes to the heart of what you're saying is like who gets to choose what we value and what we measure, as well?

MR. NAYAK: I think it's also just considering data as a way in which society operates from an institutional perspective and trust operates institutionally through data in the United States is really crucial not only to help folks understand specifically how decisions are being made, but also understanding why folks might be really skeptical of the work you're doing, because there's histories there and systems which didn't necessarily support all people.

And data has been used to do that historically, and so I think considering that context is really important particularly for engineers who are data scientists, like myself, and even in the context of AI, this becomes a much larger conversation, because of just this nature of how these models operate and how we don't fully understand those inner workings because of the black box architecture, but also just thinking about that relative to widely trusted data sources that we already use.

Some of what we do in our work is really dependent on tails and specifically meaning extremes that are unexpected. We have a really small sample of those events, and so even using standard reanalysis datasets for some of those activities can sort of spatially and temporally disaggregate and smooth out those tails in a way that is really not advantageous for our modeling purposes.

So we often turn to things like gauge data and ground truth data, as has been discussed in other panels, for this reason. There's different understandings of even where that source material comes from that is widely trusted today but can be obscured because it is sort of a widely accepted standard.

I think that's going to be super important as we move towards the future of AI and adopting these technologies more widely is where did that information actually come from and was that ground truth information? Was that remotely sensed information? Was that assimilated information from multiple sources? How did we get that?

And that ties back into a discussion on how we really promote open access science.

DR. MENDEZ: Thank you for that. Now we have about 10 minutes or so, and I want to open it up to questions from the audience and also online. Maybe I'll take one from online first.

PARTICIPANT: Sure. The first question is directed towards Mariela. Could you use Google Street View open source imagery versus your own proprietary camera systems for largescale applications?

DR. ALFONZO: That's where we started. So, yes, we used Google Street View for many years. Last year, actually now October 2024, we got a grant from USDOT to kind of scale up what we had been doing and part of what we wanted to do was actually divorce ourselves from third party imagery, for a variety of reasons. One is we don't have any control over how often Google Street View collects their data. There's sometimes gaps in the data. Sometimes they redo an area, but they only do sort of high traffic areas, and that can create bias and also inequities.

And just even from our own sort of commercialization standpoint, when we came to the city, we were like, oh, we're going to analyze the Google Street View. Okay, cool. Just that's one timepoint data, and it took us two years to get the contract, and they're like, okay, just come back to us in another five years. It just doesn't work from a scalability perspective from the company, but also just from a scalability perspective in terms of our actual mission.

So, yes, we can use it, we have used it, but we purposely have revamped our entire pipeline so that we can be more flexible and actually give cities much more current data. Also this allows us to do seasonal data, because we can deploy our own capture system multiple times throughout the year, and we have phase 2 of that USDOT grant that's starting in hopefully a couple of months where the goal is to actually design a system that cities can deploy on their own fleet and then they can collect real-time data, and we can also collect other sensor data that we can use to combine with the image data to get features that computer vision is just never going to get because it's a little bit stupid for certain things.

So like pavement quality, it just won't understand that. So we combined that with like accelerometer data, then you can start to really understand pavement quality, but also feedback into the image data and make it hopefully understand that eventually. So it's just much more robust in terms of what we can do by having our own system.

DR. MENDEZ: Okay, we have some great questions. One for Adam and the next to the entire panel, and then I'll take in-person questions.

The next is from Arthur Lee to Adam. Ultimately, will the insurance industry need to change its actuarial tables for risk calculations and increase its premium? And to have the government taxpayers share in the risk payments and paying out damages?

MR. NAYAK: That's an important question. I think generally speaking, there's a lot of great ideas right now in early stages to how we can better consider the ways in which our insurance systems right now are having some struggles with our natural disasters more generally. One of the ways in which is being more proactive and also the market is doing really, really well, is in insurance-linked securities and reinsurance catastrophe bonds, for example, that are being issued in different sectors. For example, like the New York City subway system, MTA, has a catastrophe bond to reinsure their subway system in the case of hurricanes. The ways that these payments are disbursed, though, from reinsurance and specifically triggered, are not always necessarily reflective of these spacetime realities.

For example, sequential hurricanes would count as separate storms and might not hit those same indemnity triggers that would allow for disbursement or, for example, a series of severe convective storms across the midwest would not necessarily be linked to the tropical cyclone that brought in that low pressure system and resulted in those subsequent storms if it had an appropriate time lag. So how do we actually get those systems to adapt to our climate systems is really important. How we price these things is going to be dependent on the ways in which our system, which includes both reinsurance, insurance, and policy and regulation, adapts to these new realities.

So I think there's creative ways to address these problems, but they will require innovative thinking, and it doesn't necessarily just have to do with premium raises. I think that's something that is reactionary today, because also like we talked about earlier with annual policy renewals, sometimes these decisions can be reactionary. So how can we be more proactive with our design?

DR. MENDEZ: Very complicated issue. Just for the audience members that may not know, reinsurance is the insurance companies for insurance companies, correct? And parametric is a cash buyout for a catastrophic event?

MR. NAYAK: Parametric insurance is basically triggered by a threshold that relates to, like, meteorological variable. So, oh, our windspeeds reached this threshold. So then that will result in a payout without necessarily having to have auditors go and check the indemnities of all these properties that storm resulted in this threshold of an event, hence we disbursed payment.

DR. MENDEZ: Okay, great. Thank you. I think I'll go to in-person. There's a great question online, but I want to make sure in-person has a chance. So do we have any questions in person? Up there, yes.

DR. ALESSI: My name is Marc Alessi from the Union of Concerned Scientists. I just wanted to say this was a really excellent panel, very diverse in your views. My question is something that kind of affects all of us in the room, but I think it especially affects you three, and that is the existence of like bad actors or bad systems. For example, Adam, I feel like your research kind of supports restructuring insurance, but I also know that some insurance companies are reporting record profits because they're using the climate crisis as an excuse to raise premiums.

And then in the case of Mariela and Chris, I think your research is interesting because it exists in a system of systemic injustice, like obviously wealthy, white neighborhoods are more likely to have a better built environment. So I'm just curious how you guys kind of deal with the existence of these like bad systems in your research.

DR. ALFONZO: Well, some people ask me: aren't you building a gentrification tool? No. Because part of what we do is, first of all, when you invest in a place and make it better, we're in a capitalist society, it will be more valuable. So by quantifying what that value is, we're giving information to cities, policymakers, to begin to actually create anti-displacement mechanisms. We can predict, for example, like what the rent will go up, both on a commercial and residential side, because commercial displacement is also a really important problem that isn't talked about enough, and that information is gold for our customers, because that helps them create a pathway to not have these continued systemic injustices and still be able to reinvest in these areas.

There's probably more to that question, but I'm going to let Chris.

DR. BELASCO: Thanks, and thanks, it's a great question. One of the things that we've tried to do to help frame our work around equity and justice is to look at aspects where communities are lacking sort of an outcome that they may seek to improve so that they're more in charge of kind of trying to activate some of these instruments of the city government to be more responsive to them, and our central framing has been social determinants of health, of which one pillar is the built environment. So we can't affect all elements of the built environment. We don't like manage the plumbing permits, for example. That's at a different government level.

But our efforts to try to unlock the various residents' interests in improving their own health outcomes relates to which aspects of the built environment they'd be most interested in trying to work with the front doors of city government to bring about.

So that's like a shared value between the folks who are working in city government who want to make their residents' lives better, and then the residents themselves, which increases trust to help kind of mitigate some of that longstanding concerns related to neglect.

MR. NAYAK: This is a super important question, and I think one of the things I can start with is that --

so most of the research that I've been doing more specifically is actually on the National Flood Insurance Program which is publicly managed through FEMA. And the National Flood Insurance Program came to be out of the fact that in the 20th century we had insurers withdraw in a widespread fashion from flood markets because of unprecedented amounts of risk. So sounds pretty similar, right, to what we're seeing now with new markets and homeowners, specifically homeowners and property insurance in Florida and California, and yet with this system of insurance that is federally centralized, we are seeing problems with massive amounts of debt accumulation.

There's not necessarily a profit incentive here for the federal government, because it's publicly managed. However, the management system and the pricing systems are still failing. So there is this question of like, okay, well, if we have a publicly managed system and we are facing massive challenges there, what does that mean also for private companies or even more local and regionally-based preparedness hubs in ways that we think about financial risk and conceptualize financial risk.

Even largescale infrastructure projects are going to require massive amounts of financing. So how do we think about ways in which we integrate those financial tools into our engineering pipelines? That's a much larger, broader question. So the privilege of being I guess in an academic setting is that you can kind of reimagine these boundaries, right?

Does it mean that the system inherently is a bad actor, or does that mean that the ways in which we have thought about the system as it currently is manifested, is that system functional? It is inherent to this system that there is exploitative elements. So you have got to be cognizant of that and think about ways we can be proactive, but even if we think about like disaster aid, disaster aid disbursements only amount to about $5,000 on average to individual households. That's not going to help folks actually recover from a massive disaster.

And so insurance mechanisms, because you pay a little to get a lot more, can be really helpful. So that mechanism might not be the broken part. It's maybe the system failure at a higher level that we need to rethink.

DR. MENDEZ: Thank you. With that, unfortunately we're running out of time, and I've been asked to have you join us on Slido and answer a poll question.

In one or two words, what is your primary takeaway from the AI in Urban Planning for Climate Change

Impacts and Adaptation panel? One or two words? What is your -- what was this panel about? What did it underscore?

Complicated, innovation, high risk, data, injustice, uncertain, building trust, data integration, need trust, banned words. I don't know what that -- resilient, innovation, complicated, creative. Ethical risk, odd, weird. Resilience.

That's good to see weird and surprise creative, because as we know, urban planning engineering are often considered boring. So weird is better than boring.

All right, well, thank you. It's been really a great pleasure to speak with everyone here. I didn't know what to expect, but I thought it was an innovative approach showing how the value of AI could have a people-centered or human or community-centered approach to addressing climate adaptation and action. Thank you.

(Applause)

(Break)

Agenda Item: AI: Solution or Obstacle for Climate Action?

DR. SAIN: All right. We have one more speaker today, and I think we're bringing her up. I'm very pleased to have Francesca Dominici, who will be talking about AI as a solution or obstacle for climate action. Francesca is the director of the Data Science Initiative and a professor at Harvard University.

Over to you, Francesca.

DR. DOMINICI: Thank you so much for giving me this opportunity and I'm really sorry I'm not in person, because this is a topic, the topic of this workshop is really front and center of the work that I am doing. So I hope this will be helpful for your conversation and I'll be looking forward to catching up what the other speakers have been discussing.

I think what I'm going to do, and I know I only have 15, 20 minutes. I'm going to be very brief and succinct. But I want to give you an overview of things that are front and center in my research. I think, first of all, just mention this trillion-dollar highly political scientific question that we are dealing in terms of health impact of climate change related exposures, the health impact of air pollution, unfortunately I am sure you have read that yesterday the EPA is thinking not to use health as base for quality regulation.

So I think that the first part of what I'm going to talk about is the fact that AI can be a promising solution to really understand what are the most effective interventions to protect public health in the presence of weather-related event and climate exposure and I'm going to tell you what we are doing in terms of developing the first foundation model for climate adaptation that really takes account all of the data from the healthcare system.

And then I'm going to turn the table on the other side where the degree to which AI is actually harming public health and I'll tell you a little what we were doing in terms of environmental impact of the AI infrastructure.

So I think in the context of climate adaptation, we are, I would say, really at the point where we don't know what it's going to look like, whether AI is going to make our public health better or AI is going to make public health worse. It's actually been interesting for me to work simultaneously both in the utopia and the dystopia part of the equation.

The type of question I have spent my career addressing, which are highly contentious, even more contentious starting yesterday, is really to see as whether or not even low level of exposure to air pollution including wildfires increase the risk of adverse health outcomes. What is the cost, whether or not coal-fired pollution is more dangerous than all sources pollution, whether wildfires, what type of health outcome wildfire impact, and more importantly, what to do about it. These are the type of really politically and really important public health questions that have been addressing.

One element, and that's why AI and the AI Foundation model has been attracting my attention is that to another part that I've been doing in the last now 20 years is to really build the largest database where we are gathering data and harmonizing data from the entire U.S. healthcare system. So all of the Medicare data, Medicaid data, in all the United States. So 67 million individuals older than 65 in Medicare, several millions in Medicaid. Medicaid is all ages, but for lower socioeconomic status, or for people that are eligible in Medicaid. These are individual level data. They are linked for over 20 years. We know every single hospitalization and for what cause. Basic demographic and place of residence.

And one thing to just to help you to navigate probably some of the conversation that you have had, I'm anticipating before, during the day, was a lot of AI and foundation model to digest information on climate and climate simulation and weather data and satellite data. So just think for a moment that what we're doing is some sort of equivalent, but instead of doing on hurt data, on health data, right, on all the U.S. healthcare system, because we have all the ICD-9 code and we have the 95 percent population of elderly, and 90 percent of the population in Medicaid.

So I think one element is really building foundation model that not only ingests data on whether in climate simulation, but ingests data both on what I call the exposure, which could be air pollution exposure and weather-related exposure, but also information on the entire U.S. healthcare system.

All of these individual level trajectory of cause-specific hospitalization and also ingest what we think about confounders, which will be variable that it could be both affected by air pollution and could affect health outcome in our, on the causal pathway, and so you can imagine everything related to health to socioeconomic status.

So this is just basically -- and I use the box doing an imagining because what we're doing, we're both doing exploratory analysis, but we're also trying to estimate association and so we're trying with this data by doing what I mean is like looking at effectiveness of intervention, but also thinking about causal interpretation, and so what AI can do is into the imagining world, where try to figure out what would have happened, let's say, to different health outcomes or health responses under a constellation of different action.

So one example in terms of policy-related where we are -- this was specific, not in the context of AI, but when we published this paper in the New England Journal of Medicine, this was the first analysis where we really tried to assess the causal impact of even low levels of air pollution on mortality for all of the elderly in the United States, and so you can see that the data science, the sample size, was 623 million observation, and then we did also an analysis stratified by age and gender and socioeconomic status.

This analysis had tremendous impact under the Biden administration in passing more stringent national ambient air quality standards. So there was a report in the New York Times, and here you see that on the graphic what we are doing is we are estimating the decrease in mortality by gender and by socioeconomic status and race, under different air quality intervention. So basically we are looking at how much reduction you see in the risk if we were going to revise the national ambient air quality standard for PM2.5 from 12 to 11 or from 12 to 10, 12 to 9, and 12 to 8, and you see that the benefit would be mostly for Black Americans and white Americans who are also enrolled in Medicaid.

By the way, this was the paramount study that led to the revision of the National Ambient Air Quality Standard from 12 micrograms per cubic meter to 9.

Now, this was, I would say, traditional analysis, but now what we are doing, we are having AI learning from this 9 terabytes of data and be able to do this type of analysis in a much more automatized way, but also that can learn from the potential high dimensional correlation and association from the different ICD-9 code.

So there is a lot that we have been doing in terms of using -- and you know, we also did a very extensive analysis published in Science that looked at the fact that air pollution is even more damaging to human health when it comes from coal-fired powered plants, that is really a topical situation right now, because what's happening with the Trump administration, we are reopening coal-fired powered plants to be able to provide more energy to AI.

So tremendous amount of impact. This work was just the basis, right, to building the foundation model for AI that I'm going to talk about. What we are doing now is building on this foundation of data, policy relevant question. There basically what we are doing is instead of looking at one exposure at a time and health outcome at a time, basically taking all together.

And so we are -- we all know that unfortunately we have been exposed to more extreme weather, more wildfires, and so I think that what we want to do is to have this foundation model learning from all of the climate stressors, simultaneously, and then link them to the entire U.S. healthcare system.

So the goal is really to find, basically allowing AI to learn all this multimodal data, because of course we also have satellite data and other types of data, and then be able to really assess in a more comprehensive way effectiveness of intervention on climate adaptation, both at the local level and at the national level.

So to just give you like just a little bit, we hope to finalize this ClimaCare, which is the first foundation model for healthy climate adaptation so they learn from the U.S. healthcare system, the leader is a postdoctoral fellow in my lab that is on the job market in computer science, Claudio Battiloro, and basically the idea is that the foundation model would be pretrained on the entire U.S. healthcare system times all of the environmental data and potentially be augmented with some of the climate simulation that people are developing, and societal data.

The goal is to produce unified embeddings that capture the complex spatiotemporal relationship between climate stressors, socioeconomic variables, and health outcomes. We are now deploying to where we have trained the model and we are evaluating the model and so many benchmarked downstream tasks, and so for example, can reproduce what we have published in the New Journal of Medicine, right?

And then the most important thing is that after you have trained the model and you have evaluated with different benchmarks, the great thing is now you can interrogate and to identify and be able to tell you, localize effective intervention that could be at the same time acting on different climate stressors or different health outcomes.

So I think the general idea, as you can see, there is a massive work behind the scene of data harmonization, misalignment of data, spatiotemporal data, multimodal data, and again, there is no way that in a few minutes I can give you all of the details, but bottom line, the architecture of a foundation model are pretty much similar across different areas. It's just that this one is further complicated by the fact that we are dealing with not only multimodal data, different spatiotemporal scale, but also with ingesting both climate stressor and health data and that the other complication is that, which honestly we are still working on, to figure out the degree to which we can really talk about causality in terms of effectiveness of intervention versus just correlation or prediction.

So this is all the type of downstream task I mentioned, spatial interpretation, spatial interpolation, extrapolation, forecasting, causal inference, all of that.

Now, this is one piece. The other piece is now let's say we have figured it out, okay? So we now are in a world where we have this agent, this agent can tell us at a localized level what would be the most effective intervention for public health to respond to climate crisis and to pollution. Now the question is are we better off or not?

So at the same as literally simultaneously we are working on developing this foundation model, massive computation, ingestion of terabytes of data, we're also start asking ourselves, is this explosion of AI and infrastructure of AI good for public health?

So at the same time I'm sure that this is a topic that you guys have touched as well is that you know that is all over the news, we are now have proliferation of these data centers. There are AI infrastructures that are requiring and they're basically being built in many parts of the United States, and these are very energy hungry.

So the question is in one hand we want to develop AI potentially to make us better, but then by developing a lot of AI, we need of lot of data centers that power the computation, and would that make that worse?

So we are hopefully at the final stage of this work that has been peer reviewed and hopefully publishing soon where we are doing a rigorous and replicable analysis where we have built a data science pipeline that linked all of the largest data centers that provide energy of AI with the electricity where the electricity is coming from, from which power plants, and the degree to which the source of energy is a fossil fuel or not.

This map shows in black dots, you don't see very well, the location of the data centers, with the smaller dots, these are the power plants that provide electricity to data centers. Most of these dots are in red, because most of the power plants that provide electricity to data center are fossil fuel.

And so actually there is a trend where we're reopening as I mentioned before new coal-fired power plants that were previously decommissioned to provide electricity to data center.

So basically we are now building I would say a data platform to be able to start interrogating and asking questions like what's the electricity consumption, what are the CO2 emission, what is the fuel mix of the power plants providing electricity, which states have the highest CO2 emission, and I always think that if I build all of the data, I can then start interrogating in terms of decision-making. Right now, to just give you a sense, there is really -- there is a race for the big tech company to figuring out where they can build the data center, where they can get a lot of power, as cheap as possible as quick as possible, where I think we have the data to be able to inform where would be a better location for a data center, so then we rely as less as possible on fossil fuel.

So we are doing a lot of work, a lot of methodological work, a lot of data scraping, that basically give us information about data centers, what type of, what's the power capacity based on their own characteristics, how much electricity they consume, which power plants are providing the electricity, what is the carbon footprint of the centers, and now we are at the point that from the emission, we are also submitting the air pollution, where their pollution is traveling, and how that impact health.

So to just give you some statistics about in terms of the balancing authority and the states, you can see which are the geographic areas in the United States. Clearly, we know that Virginia is the capital in terms of the largest number of data centers, the most electricity, and the largest carbon footprint.

It's also interesting to see across the different balancing authorities. So balancing authority is basically where the power demand and production is balanced within geographic region. So on top you see that the United States is articulated to different geographic regions their balancing authority. But the interesting thing is that actually the balancing authority rely more or less have different formats of how much is relying on renewable sources. So for example, the PGM, the first three are the ones that are more carbon intensive, and that's where actually we should have less construction of data center and having more construction of data center in the area where actually the fuel mix relies more heavily on renewable source of energy.

So just to give you some statistics, the total CO_2 emission for last year only from the hyperscalar data centers, which are the ones that are supporting AI, is 52.69M, which is basically amount to the entire CO_2 emission of the entire United States aviation industry. So this is becoming substantial at the point that it's comparable to an entire sector.

It's right now 1 percent of all of the U.S. carbon emission. They have been increasing five times since 2018. Virginia is the largest one, which has a significant state contribution. Actually I should say that following Virginia, Ohio is actually another one that is extremely large.

So the other interesting thing to consider is that it's not only how much carbon is, but also how intensive in terms of carbon. So what I mean by it is how much carbon there is for unit of electricity that is required. You can see the geographical area in the United States where actually there is more relying on coal for unit of electricity.

So I think I'm just going to stop here. I think some of you might say, okay, so now are we getting better or worse? Well, of course, I don't have the answer for that, but I thought to just put a fun fact where I calculate that on the one hand from the utopia part, so you should think that the Environmental Protection Agency 2024 regulatory analysis showed that if we were to pass the new standard for coal, we will cut about 55 million metric tons per year in CO_2 emission.

On the other hand, with the explosion of AI, actually we are estimating that AI is generating an additional 52 million metric tons in CO_2. So as of right now, I think we are at really exactly I would say a washout, because AI in passing regulation we're going to save 52 million in CO_2 emission, but then the explosion of AI is going to cause more the same amount of CO_2 emission. Of course these are just some toy examples. I think that we will see, but it's important to think about how we're using AI responsibly for solving climate adaptation questions, and at the same time, really being aware of the tremendous amount of carbon footprint that computation for AI can cause.

I'm going to stop here, and I'm happy to address any question you might have.

(Applause)

DR. SAIN: Thanks, Francesca. I'm sure there's probably some questions out there. Maybe I'll just start with one. I thought the concept of a foundation model that you described in the first half of your talk is really interesting. But uncertainty estimates are really crucial, and you even showed error bars on one of your earlier studies. Can you get uncertainty estimates from that foundation model?

DR. DOMINICI: Oh, yes. Well, no, I mean, Steve, you're 100 percent right. I think yes. Yes, but what type of -- I am using my statistician hat now for a moment. So I do think that we need to break the process into phases. So number one, you train these models a massive amount of data, and so the training phase, there are no error bars.

I do think that then when we are talking about downstream tasks, then yes. Error bars are possible, and again, that's work in progress, and error bars will be needed for impacting policy. So this foundation model will have zero utility if we cannot, I would say, in a satisfactory way to quantify the uncertainty.

Having said so, I'm honest from a statistician point of view, it's not easy, it's not easy because it's very -- what I would say, there is a tremendous amount of temptation to underestimate the uncertainty, right? Especially in the context of causal inference. But on the other hand, I'm optimistic, because -- but that's where also the computation -- because I think we can re-quantify this downstream task under different assumptions, but then don't forget every time we are thinking about a different iteration of a bootstrap, for example, in this context could mean a tremendous amount of CO_2 in the atmosphere.

So to be seen, but I think you immediately point a finger on something that I think is extremely challenging, that we need to take very seriously.

DR. SAIN: Fair enough. We are getting close to the end here. I know it's been a long day and everybody is pretty tired, but are there any questions in the audience? Karen?

DR. MCKINNON: Hi, Karen McKinnon, I'm over at UCLA. I am going to follow up Steve's question with another annoying statistics question. So causality is of course the other part of this, and especially if you want to argue that in your earlier kind of more traditional analysis you said if you move the standard from 12 to 9, you'll save this many lives. That's a causal statement. So how are you thinking about causality in the context of this foundation model, and what still needs to be overcome or solved to use it in a causal way?

DR. DOMINICI: I mean, I guess not surprising, that is another really important challenge. So I think that there is work now on -- let me first step back by saying that there is a lot of work right now on causal AI, and the idea of causal AI again, it goes to first you train the model, and there, there is not causality. You're just learning about the multidimensional relationship between the variables.

And then in the downstream task, you can start asking what-if questions. Having said so, the ability to -- I think you can use the same reasoning and causal inference in terms of whether or not you are getting balance with respect to the covariate, whether or not you have a measure to measure confounding. I think you can use at that point the same amount, the same type of principle that are valid and they're used in traditional causal inference to do downstream tasks after you have pretrained a foundation model.

This is still ongoing work. It's very early work. The degree to which you can put guardrails in terms of whether or not you can be confident that this correlation versus causation, that is all theory in causal inference, that's not been developed yet. So again, that's another area where we think we are now like starting to do step-by-step, but the degree to which we might trust, it's still up in the air, right?

So we don't know yet, to be honest with you, but I do think that it's definitely good opportunity to go to the bottom in terms of the theory and try than just not taking an opportunity to learn something that we haven't learned before.

DR. SAIN: I know we're getting close, but I kind of want to ask this Slido question, and it goes to some other data issues that we've talked about earlier today. How are you handling data gaps that might lead to model biases that further exacerbate existing injustices? For example, under sampling of heat extremes in marginalized communities or underestimates of exposure to PM2.5 also in certain areas? You hadn't mentioned the FAIR principles somewhere in your talk.

DR. DOMINICI: I think this is all of these questions, which are absolutely valid and talks about, really pointed out the fact that we are absolutely not the point where we are saying let's throw away everything we have done up to now and let's have the foundation model solving the problem, right? We are by far not there.

But I think all of these questions underline the need to think about foundation modeling agent in a very responsible manner, and I think that we are at the point where we should ask more questions than answer, and I think the issues of environmental justice, I mean, absolutely. I think we all know, for example, that if you train AI on electronic medical record where most of the studies have done on white people, you can't make conclusion on treatment on Black Americans that have not contributed to the data to the model, right?

So I do think that the most important thing is to go slow and to go responsibly and try to better understand which question these technologies and these innovations can make us better, and which questions we might led us to more confusion and bias. So that is all an area of research, that's why we're having this conversation, and that's why I am devoting a lot of my time to think about in a very balanced way and not saying, okay, we have a foundation model that's going to give us all of the answer, and problem solved.

DR. SAIN: Okay, thank you, Francesca. That was a wonderful talk, thank you.

(Applause)

So, Francesca, you can't quite leave yet. We do have a poll question here, and it was a quiz to see if we were following, but I'm not sure I saw it in the talk. The poll question there on Slido, let's see. How many gallons of water per day does a typical hyper-scalar consume for cooling?

DR. DOMINICI: I didn't give you the answer. Now you have to guess.

DR. SAIN: Well, we have guessed. The results are coming in about, well, we just switched: 500,000 gallons per day has got some momentum now at over 60 percent. Are we on track?

DR. DOMINICI: Yes.

DR. SAIN: All right, 500,000.

DR. DOMINICI: Yes, 500,000 per day.

DR. SAIN: Per day, all right. Okay, thank you again, Francesca.

DR. DOMINICI: Take care. Thank you for your attention. Bye, bye.

(Applause)

Agenda Item: Day 1 Wrap-Up

DR. SAIN: All right, so we're very close to being done here. We were going to just quickly wrap up with a few perspectives on the day. I have a really long list, and I'm like looking at the clock. So I think the one thing that I can't quite get out of my head is bees and elephants, I'm sorry. Especially the picture of the elephant with the wings. I'm going to -- that's going to be with me for a long, long time.

It was a good point, right, of finding balance between these largescale models and smaller, more specific and focused models, and trying to find that balance I think is super important.

I have others. I have a list, if you want me to just read my list. But does anybody else have something they'd like to point out that was sort of stuck in their mind from the day?

I can pick on people.

DR. LEE: Two points. Arthur Lee again, retired from Chevron. I am taken by two points really. One is data. You really still need lots of data, good data. That's critical.

Number two is AI needs human expertise, domain expertise. That's been emphasized over and over again in these talks that I've learned a lot from today, and of course, that adds to also the ethical guidelines, both for researchers as well as for public consumption of data. I think that's what will increase that level of trustworthiness.

Maybe even a third point, a half of a point, is that farmers are very smart people. You cannot fool them.

DR. SAIN: I spent a few years working in digital agricultural, and I was amazed on a daily basis how smart and capable farmers are. And you touched on a couple of points I also had, the importance of not just having lots of data, but having good data, the right data, may be even more important than quantity.

I saw somebody up there, yes.

DR. HARPER: Anna Harper, University of Georgia. I'm part of the roundtable, so when we were discussing the ideas for like conceptualizing what has become this workshop, we talked -- one thing that we talked about is workforce development and training, and this is relevant to something that kept coming up today that if we're trying to accelerate climate action with AI, each question and challenge that we have with climate change requires a different like customized approach and modeling approach. So we can't just have like a one size fits all approach that's going to solve all the problems, unfortunately.

But so this means we really need to train the next generation of scientists to be able to handle this.

DR. SAIN: Absolutely. I see a hand.

DR. FURTADO: Hello, Jason Furtado from University of Oklahoma. So kind of jumping off of Anna's point, I think with some of the training on this, I think in science in general, but especially in this area, is that the successful scientist also has to be an effective communicator, and that includes all levels, not just about the model itself, but also the results, the uncertainty, the applicability, the transparency, the data, et cetera. So there's a communication element that has to go into our training and to our workforce development that in my opinion is probably lacking currently in a traditional science kind of foundational classes.

DR. SAIN: Yeah, excellent points, and we're seeing a ton of points from the audience coming up here. I don't think I'm going to read all of these. There's a ton of them.

Anybody else? I could read more from my list. Well, another thing that struck me is my background is in applied statistics, but I consider myself very much an applied statistician, data scientist. So I've been sort of thinking about a lot about collaborative work and I think I love working in a collaborative space. Collaboration to me drives innovation and statistical methodology, and that collaboration drives machine learning methodology, but one thing we heard was even though we're trying to push the frontiers of AI and machine learning, sometimes going back to the basics to solve the scientific question is the best approach, and that's sort of something again, workforce development is sort of to impress upon people is sometimes the simpler model is maybe even an older model that doesn't quite have the whizbang feel of some of the newer stuff, so is really, really important. I thought that was really an important thing here.

I also liked many of the characterizations of trustworthiness. I liked one of the ag speakers kind of talked about this, too, and it maybe even rethinking the trustworthiness as really about adoption, and so, yeah, we could talk about getting these models better and better, but then there's this huge gap before they can actually be used, and how do we sort of close that gap?

I have a couple of others. A lot of people are reading the great things off the line.

I think we're at time. I'm just rambling. Does anybody else, one last thing? There's one up there, all right. Don't let me end this.

PARTICIPANT: I was just struck today or in the back of my mind is we've spent a lot of time talking about trustworthiness and trust and I think one of the things that's become apparent to me is that looks different in the different contexts that we're using AI, but then I was also thinking today what we haven't mentioned is sort of the societal moment that we're at with science and trusting science, and I think is at somewhat of a low right now, and so that's going to provide this additional impediment to adoption of AI, I think, and we're going to need to work even harder on the trust issue to overcome that.

DR. SAIN: That's an excellent point.

With that, come back tomorrow. That's really important. We're not done. Tomorrow is a little bit of a different feel as we start kind of putting these together with sessions on maybe sort of thinking about what the commonalities are amongst things we saw today, and then starting to look more forward. How do we move this along over the next few years or so.

Thank you very much. I appreciate everybody. Appreciate everybody online, and yes, we'll see everybody tomorrow.

(Whereupon, the meeting was adjourned.)