The National Academies of
Sciences, Engineering, and Medicine

Roundtable on:

Accelerating Climate Progress with AI: From Science to Action Workshop

January 14, 2026

Proceedings by:

CASET Associates, Ltd.
caset@caset.net

TABLE OF CONTENTS

P R O C E E D I N G S (9:00 a.m.)

Agenda Item: Welcome & Reflections from Day 1

DR. SAIN: - some of the things that we have talked about the importance of human expertise has come up again. Bees and elephants. I think that's going to be a common theme for the day. The importance of trustworthiness. I'm just kind of scanning through here. More human expertise. Ethical guidance. Decision context. Cannot hide an elephant behind a bee. Yeah, fancy new tools don't matter if people can't or don't want to use them. That was sort of another theme on what we need to do to encourage adoption.

Any other comments from the room? Yes, sir. There is a microphone right behind you if that's okay.

DR. KOVACH: Good morning, everyone. My name is Jon. I'm the director of the UCI Science Project. I work in the School of Education. Welcome to the campus, and welcome to Southern California in this wonderful weather.

I just want to highlight that I don't see education or TK-12, TK-16 education to be mentioned anywhere here, especially bringing up the idea of trust within communities or within systems. Working with teachers and working with students is one way to, I think, to start to build that trust around this technology and around these ideas.

Sitting in the room yesterday, I just saw so many connections with ways that this can leave the university, leave the room, and get into the communities where you work at, hopefully particularly marginalized communities where students can use AI to solve problems that they're interested in and that they want to be involved with, with their community.

I don't know if you know this, but California, in grades 1 through 12, we must specifically address climate change, not just cause and effect, but also adaptation and mitigation strategies.

So, working with teachers, I think, is a great way for a lot of you to maybe address those trust issues, that idea of getting these ideas to students at an early age, where there are a lot of creativity and diverse and divergent thinking. So, thank you.

DR. SAIN: Thank you for that perspective. Anybody else? Okay, can we go back to the deck? Okay, so a little bit about the agenda for today. Yesterday was very much focused on very specific areas, trying to get into the weeds a little bit. Today we'll take a step back, look for commonalities, and ways to step forward.

I think I'm going to now turn it over to Adrienne, who will introduce our first two keynotes.

Agenda Item: Keynote Presentation – The Trust Deficit: How AI Can Bridge or Widen the Divide in Climate Progress

DR. WOOTTEN: Good morning, everyone. If I didn't get to meet you yesterday, I'm Adrienne Wootten. I'm a research scientist at the University of Oklahoma and one of the members of the Workshop Planning Committee. It's an honor and a pleasure to be on the Planning Committee, to have helped organize, and to be both introducing our keynote speakers this morning and moderating the question and answer that will follow.

Each of our two keynote speakers this morning will speak for about 20 minutes, and then we'll have time, 10 minutes, for question and answer, thereabouts. For folks in the room, please go on back to the microphones in the back so that folks online watching can hear you. For folks online, through the Slido application, you can ask questions, but you can also upvote questions, and during the Q&A, we'll go back and forth between in-room and online, and April will read out the online questions so folks in the room can know what folks were saying.

I had a lot of thoughts on yesterday, too, so before we get there, I want to add, and we'll get to introduction through this, I promise. It was really fascinating to see the breadth of tools and the breadth of discussion, particularly around the issue of trust. And, of course, you can tell our first speaker is going to talk about that in a lot more depth, in particular today, because we get into a lot more with the interactions with people and policymaking and the general public alongside all the other issues that we discussed yesterday.

And so, with that in mind, it is my pleasure now to introduce our first speaker. Our first speaker, Mr. Kieran White, is a multimedia journalist with nearly 20 years of experience. Hails just from the U.K. Here he's a member of the National Union of Journalists and the International Federation of Journalists.

He began his career in the creative industries, traveling globally and doing quite a lot of photography work, a body of work that has resulted in hundreds of published features and credits across international outlets.

In 2020, in particular, he pivoted towards investigative journalism on sociopolitical impacts, particularly with COVID-19 and open-source intelligence and examining how groups used emerging technologies and psychology to recruit folks in the restrictive environments that happened under the pandemic.

Today, he has enhanced that work with his current project. He's called Cool Your Bytes, very cheeky title, a project synthesizing his creative roots in media and looking at the societal risks of AI in a critical manner, particularly with an eye towards the public and how they perceive it, and issues of trust in it and the good, the bad, and the ugly. I'm embellishing a little bit in all of this.

So, it is my pleasure. Please let us welcome Mr. Kieran White.

MR. WHITE: Good morning. Now, I wanted to start off with a little title slide, but I didn't want to be outdone by David yesterday. So, absolutely no bearing on the presentation to come, but I didn't want to be outdone. So, we switched things up a little bit. We have, rather than an elephant with bee wings, we have a bee with an elephant trunk. So, that was just a bit of a palate cleanser for everybody this morning.

Now, last night I was having dinner, and we were talking about various different things. And one of the points someone brought up was that I think it was today or tomorrow is the 20-year anniversary of quite a big event. And if I'm correct, it was a thermoplastic explosion that happened in Siberia. Now, you'll have to forgive me. I pulled a lot of archive.org footage, but I wanted to include it in the talk this morning. And I promise it'll loop back around. But I thought we'd test how good the caffeine has hit the system this morning.

I thought I would play some of the footage that I found and see if anyone remembers the event from the footage. So, just for context, this was in Siberia about 20 years ago, as I said, about today or tomorrow. I can't quite remember exactly when it was, but there wasn't really much footage I could find from archive.org. But as I think we can all agree, great resource for finding bits of footage.

Show of hands. Does anyone recognize this event or have anything that is sticking in the mind? Anybody? No. Okay. Well, I was cruel to you. I was cruel this morning. The caffeine's working because you're all on the ball. So, that was a great little experiment.

But it does work because I want to talk about how trust is built. Why do we trust? And so, indulge me for a second. Pretend you did believe this footage. Why did you trust what I presented? Because if you don't play into the game, then the whole presentation flops.

The thing is, there are lots of ingredients that go into the process of trust. And even for myself that has done a lot of work in open source intelligence, looking at footage, analyzing media, I think the people in the room here that engage with a lot of this, especially generative AI on a day to day basis, is becoming progressively harder to trust what we see.

So, this is my blow-by-blow of the potential ingredients that go into trust. So, the psychology of trust. Now, context. Context is a very important ingredient here. One of the most important. And so, especially when it comes to generative AI, which is the public's main touch point with a lot of AI at the moment, is what is it that we are seeing? What is the context in which we're seeing it?

A lot of that goes into whether people trust it or not. Then, personal experience, whether the media they consume. So, for example, with this footage that I showed you, whether they've had a personal experience of that event, as to whether it triggers something within their memory, and that is able to build that connection there.

And then familiarity. Familiarity, a little bit like personal experience. It goes a long way to building the believability and hijacking the mind into giving more credence to what we're seeing.

And then visually compelling. Is the media visually compelling that we're watching or consuming? Now, for my part, this one was interesting as I was trying to put this together because a lot of it was coming out very plastic. A lot of anyone that's dabbled with generative AI knows that it has a very specific look about it.

So, what I went and done was I put a lot of filters over the footage and tried to down res, export it at a lower resolution. And we'll see this a lot with a lot of footage that's coming out of war zones. A lot of it is hidden under the guise of cell phone footage. So, we know from the start that there are that there are that narrative that's being built there.

The narrative. So, as I as I laid this out to you, I was talking about how I had dinner. We were talking about the event. And so, it's building that story, that package that allows us to see the footage or see the picture and be able to believe it a lot more.

And then the source. Where is it coming from? And this one is quite an important one in terms of what we're discussing here at the event today and yesterday, which is that obviously generative AI is not the sum total of what is going to be used when it's deployed in the climate sciences, but it's one of those ones where the source of the media, specifically for generative AI, it can be a lot of different things.

For example, does it come from an individual's favorite influencer? Does it become come from a trusted source or does it come from a family relative? And whether it's shared on social media, a family relative that shares something, there is already that built-in trust there.

Now, when it comes to the source, we can then loop this back into what we're discussing over this two-day event. When it comes to data, do we trust the outputs from the AI systems if we trust the source? So, if it's coming from systems that we've built that we have an inherent trust with, there is a likelihood that we may not always give it the due diligence of checking over the outputs there because we trust that the source is legitimate.

So, the problem we have is that there is a great trust deficit across the board, especially with the public perception in AI. And I thought this was really interesting. Pew Research Center put out this study, where you can see the results for yourself, that 50 percent of Americans say they're more concerned and excited about the increased use of AI in daily life. And this is up from 37 percent in 2021. But 10 percent are more excited than concerned.

I think that the excitement is not necessarily a negative or a positive. It is showing that there is a level of anticipation about the new technology. I think this is really important to bear in mind going forward because ultimately, the public is the end user of everything that we're doing.

So, this was also an interesting result that came out of their study. And it was showing that 29 percent of U.S. adults say it will make people better at solving problems. But in a larger share, 38 percent say AI will make people worse at solving problems. So, it's very interesting with this that when you contrast this with the previous slide, there is definitely an anticipation. But people are trying to rationalize it against their use of AI. And as I said previously, most of the public's touchpoint is with generative AI. So, not really conceptualizing how this can be used, especially in the domains that we're talking about over the past two days.

And then we have this one, where most Americans say it's extremely or very important to be able to tell if pictures, videos, and text were made by AI or people. And then 50 percent - and I think this is a concerning one -

that 53 percent of Americans are not too or not at all confident that they can detect if something is made by AI versus a person.

And so, with that, coming back to the topic of trust, the public is very cognizant of AI and the fact that it can trick you, but they are relying on the fact that the data that's being put out there or the information that's being put out there can be trusted. And so, it's important that individuals like ourselves and in the climate sciences, if we're going to be leveraging this technology, that it is something that we're ensuring that we can trust, because the public are finding it very hard to trust the outputs as it is or be able to discern what is real.

And then we have this one here, which again is interesting. So, majority say AI should play at least a small role. And I wanted a good news story here, and I thought that you'd enjoy this one, that 74 percent said that it should be used in forecasting the weather. So, the other ones are a little bit doom and gloom, but this one should put a little bit of a smile on people's faces.

Now, this one I thought was also interesting. So, it feeds into what we've previously been talking about. This is the KPMG study. And it's very interesting that they go through the different points and different topics here.

And especially in AI and trust and acceptance, where they say three in five or 61 percent are wary about trusting AI systems. And then 67 percent report low to moderate acceptance of AI.

Now, I'm not going to read all of these, but it's key to point out that there is also a divide in who they trust to develop AI, which is why it's so important that at source we can prove that it's auditable. And I think that's going to be the main buzzword here, making sure that the AI systems are deployed, are auditable, and anything that is used is transparent to have public buy-in with this.

And then responsible AI. 97 percent strongly endorse the principles of trustworthy AI, and three in four would be more willing to trust AI systems when assurance mechanisms are in place. And 71 percent expect AI to be regulated.

So, going back to what we were talking about yesterday with the examples from Europe, there is a lot of good regulation coming in with AI, and it's what the public is expecting. And that, in itself, regulation, provides a certain layer of trust, a perception of trust to the public.

And in terms of how people are understanding AI, half of the respondents feel they don't understand AI when or how it's used. This goes back to the transparency of explaining to the end user how it's being deployed and in which ways it's being used and when it's being used.

I think this is a good news point to take away from this, that 85 percent want to know more about this. There is a hunger. It is not something that people want to ignore altogether. There is an interest there. There is an appetite to learn more.

So, what is trust? We spoke about this yesterday, and I'd like to thank Deb and Kathy for their contributions yesterday in helping me finalize my points around this topic. But it was really interesting. So, trust is not just one static thing. It is different depending on who we're talking to. So, the public will have one conception of trust. Science in general will have a different definition of how they measure trust. And communities will have a different definition of trust as well.

The key takeaways from this were collaboration, ensuring that the data is able to be collaborative. Communication, making sure that the points are communicated. As we saw from the previous slide as well, there is a hunger to know more. So, communication is key. And understanding. People want to understand how this is used. So, being able to talk directly to the end user, or in the context here, the output, which is the public, goes a long way to reestablish the trust. As we said, there is a natural trust deficit here right out of the gate.

I put this together. I hope it makes sense, but I'll walk you through it. So, this is what I'm calling the legitimization pipeline. So, here we see the information pipeline, so to speak. So, we have source data. Now we can look at this in two ways. We could say that this is the climate science in general, science in general, or even we can go even further back to the input data that's being used to build these systems. So, this is the really, really important step in this pipeline.

And then the interpretation stage. Now this is where individuals like me in the media, relying on the source data to be correct, or even legislators, we operate in the same domain here. We rely on effective communication from source data or science to be able to translate that and communicate effectively and in an accessible way. Because ultimately, through budget cuts in newsrooms or plenty of stuff that legislators have on their plate, they lack the time and the resources, but also the technical understanding a lot of the time, to be able to really crunch a lot of this very technical jargon.

We rely on the fact that at source it is correct. So, we rely on the previous step to have done their due diligence before it reaches us. And ultimately we are the ones that legitimize the data.

And then the public consumption. Everything goes back to the public, no matter what industry you're in. Arguably, at some point, the endpoint is the public. And so, it's very easy to get wrapped up in this little bubble and talking amongst ourselves, but it's so important to realize that everything we do is for that end user, is for public consumption. So, ensuring that our work has been done early on in the stage to ensure that the data is correct, and then being able to communicate. Those are the two things that are going to ultimately win back trust with the public.

Now if we consider stakeholders and what we need to think about in terms of how this will affect them. This is another thing that Kathy very interestingly put forward, that data comes in many forms. It is not just scientific data. It is community data. It is cultural data. And a lot of this goes back to that whole thing of communication. It's all very well having the data output. But how do we translate that to the stakeholders? To the public? To the people that are making the decisions?

And so, having those two synthesized together, one cannot really work without the other. And then leveraging existing experience of AI. Angel made this point yesterday in her presentation about the generative AI and the boom in generative AI. And as we experienced previously in a couple of slide, the public's main touch point at the moment with AI is generative AI. A lot of people think that AI is just creating memes to put on Facebook. But obviously it's a lot more than that.

And so, I think it's really powerful to look at what the public know, meet them where they're at, and leverage the 85 percent of people that were saying they were excited and wanted to know more, and be able to synthesize those two things together. And that way you can through transparency and proving that it is auditable, and showing that you are willing to explain stuff, that you're able to, again, hopefully build back that trust.

So, something that I always believed is that trust is earned, not given. And ultimately, it's all very well being part of these great big institutions, but the problem is I think the days of people having default respect for the big institutions, just because there are provenance with them or well known, I think they're coming to an end, if not over, at least for the public perception.

And so, it's not good enough just to be able to say we are from X organization or we are from X body. It is trying to show exactly what you're doing.

And so, these I think are the main takeaways here. Transparency. Making sure that everything that we're doing is transparent and people can trace it back to the source. Auditability. Making sure that we're not seeing AI as an oracle. Making sure that we are ultimately still in control. And I think it was spoken about yesterday about making sure that it doesn't become a black box. Understanding what is going on under the hood as much as we can do.

Public education. And again, one of the main points I'm driving with this, is there is such a great hunger from the public to learn more about this technology. So, it would be foolish to not leverage that. So, educating the public. Understanding how AI works in their life, how it can work for them, and how, within the climate sciences, it can be used to help them and give them better life opportunities, and solve a lot of problems that are pressing for them.

And then compelling narratives. A lot of this again goes back to communication. How are we wrapping this up? We've had some great discussions over the past two days, but how do we package this in a digestible way that is able to translate to someone from the Midwest who has no idea about most of this stuff, but understands how, at root, this all will affect them down the line.

And so, building those narratives out to make it accessible and relatable to them, again, is something that is able to build trust back up, rather than feeling that they are separate from what we're doing here today, because ultimately it is a process that every step of the way, everybody is part of.

And so, if you would like to read more about what I'm doing currently, about focus on AI from an individual liberty point of view, civil liberties, data, and making it more accessible to people, so it breaks down things in a digestible way, you can scan the QR code. If you have any questions you can reach out via email. I think most people have Substack nowadays.

Thank you so much. I'd like to thank the National Academies for the invite of coming here, and speakers from yesterday as well. So, thank you.

(Applause)

DR. WOOTTEN: All righty. We have time for questions. I know I have a question or two. If you have a question, go ahead up to the microphone if you're in the room, so folks online can hear you. I think we have a couple of people who might be moving that way. I don't know if we have any questions on Slido. Kathy, do you want to go first?

DR. BOOMER: Kathy Boomer with the Foundation for Food and Agriculture Research. A fantastic conversation. It's really struck. One of the things that came out yesterday was uncertainty, and underlying the process of consultation is a willingness to embrace uncertainty, and that's actually brought out explicitly in the adaptive management literature.

And one of the struggles that I've continued to experience is the tension with building a compelling narrative and wanting to put uncertainty under the rug, and yet seeing that as a really exciting opportunity to work together and solve problems together. So, I don't know if it's a comment, but also a question for guidance on that challenge.

MR. WHITE: I think, yeah, thank you very much for the question, and, yeah, thank you again for allowing me to pinch some points from your talk yesterday, but I think that's the thing. I think ultimately we need to meet people where they're at and understand that, again, people are very, very interested about this technology.

And there is fear, but it's the fear of the unknown. And so, being able to relate things back to how things will affect them in their lives, I think, goes a long way. There are many different ways that we can take anything, not even this topic, but being able to make it digestible and relatable to people makes them feel like they've got buy-in to that situation, so they can psychologically put themselves in that position. And I think that's the best that we can really do.

And then just proof is in the pudding, so to speak. Because as we said, straight out of the gate, unfortunately, there is a trust deficit with science in general, experts, the state. People are very fatigued. And so, with that, we have to build back trust. We can't start off from an expectation that people are already there. It's not a case of there is trust to lose. There is trust to gain.

And so, I think that's the biggest thing, meeting people where they're at and making it relatable.

DR. WOOTTEN: Can we get one question from Slido?

MS. MELVIN: Sure. Studies from MIT show statistically significant negative impacts of cognitive ability after relatively short three- to six-month usage of

AI. How can academia temper AI usage so the smartest among us don't get dulled from AI dependence?

MR. WHITE: I think I remember this study. I think it was, if I'm correct, it was the study where they got participants to do an essay. Yeah, and they'd done the scan of the - yeah, I have got a memory. No AI was used in recalling that information.

But, yeah, I think because there are a couple of different aspects there that we need to think about in terms of one that not many people talk about is that developmental pipeline of individuals, and in terms of developing reason and logic. And a lot of kids that are coming into using AI at the moment, they're what we call digital natives. They were born with iPads in their hands. And so, it's very interesting to us that remember a time before the internet.

And so, it's very interesting because we're trying to put ourselves in that situation, but there is a lot of stuff that education is trying to adapt very quickly, whether it's deploying tools to look for AI use in essay writing and different ways. Students are crafty. They'll get around. They'll use it.

But to the central point there, I don't know, is the thing. I think we're in the early days still, and it's something that is unfortunately, I think it's a whack-a-mole game. When things come up, you have to address it where we're at, but in terms of the fatigue element of this, trying to enforce and tell people to write stuff by hand, do the analog stuff because it's just as important, but then also reinforcing that AI is - and this is a central point - AI is a tool. It is not something that we're meant to outsource all of our responsibilities to. It's not something that is meant to do everything for us. It is a tool. We should be in command of it.

So, understanding when to use it and how to use it and when it's appropriate to use it, I think that's the best way. I hope I didn't ramble there, and I got to the point.

DR. WOOTTEN: All righty. I think we have a question in the room.

DR. ALESSI: Hi. My name is Marc Alessi. I'm from the Union of Concerned Scientists. I just want to say thank you for this excellent presentation because I feel like it's a wake-up call for scientists who struggle to communicate and build trust.

I was really interested by one of your slides where you showed the source data, being the scientist, and then having the government media in the middle to disseminate that information. I'm curious what your thoughts are if people or the public live in an authoritarian regime where media and government may disseminate wrong information knowingly.

I'm just curious if you have thoughts on that. Like should the scientist directly speak to the public?

MR. WHITE: I've got many, but unfortunately, not much time. I think, yeah, that is a vital point. It's controversial to say possibly, but I don't think it's limited to authoritarian regimes. There is a necessity, whatever side of the political spectrum one sits on.

I think we can all agree that the COVID years showed the politicization of science, and scientists were used as human shields for a lot of bad policy, and this, unfortunately, objective truth matters, but perception is sometimes more powerful. And so, public perception of what they were consuming through media and the way they were seeing policy enacted, that was enough. Whatever was actually happening within the various administrations, it's secondary. The objective reality is secondary to the perception.

And so, when it comes to things like this, there is a lot of political point scoring on being able - at that translation stage. We also have to trust - again, talking about that legitimization pipeline, going back to that, it's understanding that it's a domino effect, and the end point is the public. Everything we do is for the public. And it's very easy in our own little bubbles here to lose sight of that, but they are the end point, and if they can't trust that us in the media and the legislative is doing their job properly and we can't trust the source data, whether it be the scientists or the tools they're using, yeah, unfortunately that legitimization pipeline fails and it ruptures.

But, yeah, there is going to be a lot of misinformation put out there, intentionally, or accidentally, through poor communication or, again, poor understanding of that. And so, again, transparency, auditability of the data, but also communication and being able to wrap that up again in a compelling narrative. So, yeah, thank you very much.

DR. WOOTTEN: Okay, I think we have time for one more very brief question. I see a second one on Slido. I think that one on the farmers would be interesting here.

MS. MELVIN: Sure. So, this is, I'm referring a little bit back to yesterday's conversation. Farmers are hard to fool. We heard that farmers need to know that the neighboring farmers believe the same thing. How do we put this trusting your neighbor into practice online? Might this be more in personal messaging or testimonials?

MR. WHITE: Yeah, online is a battlefield. This is the thing. That's a whole different talk on the use of social media and algorithms and things like this, but I think there is a lot to be said for decentralization of a lot of this stuff, especially in communities.

Centralizing some of the aspects, especially what we're talking about over the past couple of days, is important to be able to get arguably some of the bigger tasks dealt with, but on the community level, I think being able to decentralize a lot of this communication and being able to talk to one another and, again, sharing that knowledge, understanding that someone might understand AI a little bit better. Oh, I used AI to track when to put the crops down, or I was able to track the weather to be able to know that, actually, this year it's not going to be the best for yields.

But that's the thing. I think ultimately, going back to what we were talking about with source, if you don't trust the state, if you don't trust the scientists and the experts, and you don't trust the media, you'll trust your friend, you know. And so, I think understanding and leveraging that is also important. There is trust to be found. Don't worry. No one is rocking back and forth and doesn't trust anybody. Even the most cynical person, there will be someone they go to that they trust.

And once this gets deployed further and further in day-to-day life, everyone will have an experience with it, and that will boost confidence with using the technology, but also in the data, because that will just promote more understanding as well. So, demystify things.

DR. WOOTTEN: Awesome. I know there were four people who wanted to ask questions, but that's what the break is for. So, I will direct you to find Kieran during the break. Let's thank him one more time.

(Applause)

MR. WHITE: Thank you so much.

Agenda Item - Keynote Presentation – Earth Observation for Climate Action

DR. WOOTTEN: Now, I think Dan is our virtual keynote speaker today, so hopefully we can bring him up. Dan Hammer is our next speaker and is the co-founder of LGND, a venture-backed startup building AI infrastructure for Earth observation data, and co-founder of Ode, a design technology agency for environmental applications.

He is also a fellow at Renaissance Philanthropy, supporting open-source AI for climate and nature. Dan's previously served as chief data scientist at the World Resources Institute, where he co-founded Global Forest Watch, and is senior advisor at Google X. He was also a senior policy advisor in the Obama White House and received multiple awards here, which I'm not going to read them all out, but that's why Dan's bio is also on the workshop website. Dan, please take it away.

DR. HAMMER: Thank you for having me. I really appreciate the chance to talk to you guys about this. I'll start with a demo here, because it's just easier to understand some of the new capabilities by just looking at it and watching it happen. But the basic idea is that it's now possible to search Earth observation like you would search the internet. It's not just the availability of the data, like data that is collected from satellites, aerial, airplanes, or drones, but actually search the semantics, the information that's otherwise stranded in the pixels and the raw material, because of inventions that had nothing to do with us, the transformer, but just applying it directly to this new modality.

This has some pretty significant implications for applications in both the public and private sectors. And so, the focus of my talk today is the role of the social sector in the guidance and governance of these large Earth observation models, and specifically focusing on the gap between what is not covered by either the public or the private sector, specifically the things that the public sector can't do and the private sector won't or probably even shouldn't do. So, with that, I'll show a particular instantiation or a particular front end that we can now do with this technology.

A couple of years ago, along with my co-founder Bruno, Bruno was the builder of the Microsoft Planetary Computer, a couple of years ago in the open as a nonprofit project, we started Clay, and Clay was an initiative to train a large Earth observation model, the P and the T of GPT, but trained on satellite imagery, open public satellite imagery, rather than text. And since then, there have been dozens, hundreds of these large Earth observation models released, and all of them are pretty difficult to use.

But simultaneously over the past couple of years, the infrastructure in order to work with the output of these models has come into being, and we can use it and pull together specific interfaces like this one.

So, for this particular demo, what we did is we indexed all of the USGS aerial imagery for search. This is one meter resolution imagery over the past, I think it's like seven years or so, annual imagery from the USGS, mainly for farmers, for agricultural purposes. But what you can do now is you can type in something like, find solar farms in Texas. This is not just querying pre-labeled information, things that we've already seen before. This is going each time you call it in a live way, it goes and finds solar Panelists directly from the imagery itself.

For this demo, we just returned the top hundred results, but this type of project would have taken, I don't know, a week to do, to train some custom convolutional neural net, specifically trained on solar Panelists in order to inventory solar farms in Texas, but now it's really quick. What that allows is 10,000 times reduction in cost and an increase in speed in order to more naturally engage with the information in overhead imagery from various sensors, not just RGB, but from dozens of spectral bands that are collected all the time from commercial vehicles as well as public sector satellites.

The most important part of this has been just a much greater access or level of access to the information inside of these pixels, and a lot of them are patterns or those subtle and implicit patterns that are just difficult. They're literally buried in stacks of satellite imagery to be able to find, for example, not just where solar

Panelists are, but where they were built. Show me new solar Panels. I preloaded this since it takes 20 seconds to do, as well as a couple of other examples - but preloaded imagery of these are locations where there was not solar Panelists recently and now there are

These are those types of interactions with information that already exists. It's already been collected by other people, by Google and commercial imagery providers in the public sector, but it just hasn't been indexed for search. And now, using these transformer models, it can be. So, for example, this particular one, this is some area in Texas where even the base map imagery hasn't caught up yet, but apparently here between 2020 and 2022, there was a solar farm developed in this particular plot of land.

But this wasn't trained on any particular concept. This is on the fly. So, for example, in this particular case, suppose that I was just interested in the sub-categorization or looking for a very specific way in which solar Panelists were developed. So, not those that were developed in bare Earth areas, but those that replaced forests. So, I can say just show me things like this, but not something like this, and you end up with a more refined search.

Again, each time the variable cost of doing this in the way in which we used to do Earth observation analysis was to retrain a model, but now the idea of being able to point and click in order to find similar locations and then subsequently monitor and measure those locations, it's just way more accessible. So, it doesn't require some expert in Earth observation and coding, but one of those non-coding analysts that knows exactly what they're looking for, what they need to measure, and can do so with a point and click interface, and that's what AI has afforded us. It's the accessibility of the information that's within this otherwise pretty esoteric and inaccessible data.

Just to show the value of this for the long tail of applications, I wanted to know where these concentrated animal feeding operations are in the country. In this particular case, I did it in Alabama. It's weird that we just don't have that data set, but what's more is you can find the new ones, the construction of where pigs, chickens, and cows are raised in a particular state, very, very quickly. Or aquaculture in Virginia. Or where timber is being aggregated and processed, these logs in Oregon. Or new farm building development in Oklahoma, and even those specific places where there has been new farm development, where there is heavy machinery present, for example.

This has now created this new paradigm for interacting with this particular data source, which we've collectively, even in just in the public sector, spent tens of billions of dollars on this data set, and it's just wildly underutilized. And so, these models, these large Earth observation models, and the subsequent infrastructure to make the outputs more useful and usable, allow for all sorts of new applications and considerations for policy. How to balance, for example, privacy and efficacy. Or to develop trust. Or the best way in which to deploy governance of these large Earth observation models in the service of public, in the public interest.

These models are in their early stages. They're nascent. Consider GPT 1.5 or 2 or something like that, early days. But they're getting better fast, and the infrastructure and downstream algorithms in order to use the outputs are getting better, even faster. And it's about to hit the world in a particular way, and we have the benefit of hindsight, and some of the lessons learned potentially from the release of other AI modalities, like language, in order to understand what would have to be in place in order for us to see the responsible rollout of large Earth observation models, specifically in the public interest.

That is what has given rise to this project through Renaissance Philanthropy. Now it's a project of Renaissance Philanthropy, which is a project in order to establish the benchmarks, develop policy, as well as actually release some of the models that are a realization of those benchmarks and policies specifically developed to encourage the other large AI labs to develop their architectures and models in a way that increases the likelihood that this would be used in the public interest.

So, we took our motivation from this. This is GDPval. It's released by OpenAI and I love this project, and for those interested, it's worth going through and understanding the improvement of GPT 5.1 relative to 5.2, which is the current release, and the way in which these benchmarks have actually mattered to move billions, if not tens of billions of dollars of development and training toward benchmarks, towards evaluations about how we're getting better and how we assess progress.

GDPval, what they've done is they've taken nine industries, like real estate or healthcare, and then included 44 occupations or users bucketed within those industries, and each one of them has certain tasks where large language models and the OpenAI suite here can help those specific real-world people and the tasks that they need to perform, and that's how they're assessing progress is they have expert graders in each one of those occupations that continuously try a certain set of prompts and then grade the responses from ChatGPT, for example.

So, this is a prompt for a manufacturing engineer, an order clerk, or a producer, as they call it, and they assess the output. And then there are automated graders in order to further remove the human from the loop. But the basic idea is that the way in which these new models are being assessed is based on these specific OpenAI-determined real-world use cases.

And so, one of the things that we're thinking about right now is, if we were to determine those industries for large Earth observation models, how would we define what good looks like in a way that's consistent with public utility? So, it might be that one of the industries would be education, or for the K through 12 students in California. It could be that it's health. So, to identify large waste aggregation sites near rivers before different plastics enter the ocean. It could be for wildfires or investigative journalism.

The way in which we would want to see that the industries that aren't natural for targets for Earth observation, which is by and large driven by defense and intelligence, and has been since its inception, basically. It was meant for reconnaissance. And that's where the private sector is primarily focused, as we define these benchmarks for these large Earth observation models to look on the extensive margin, new topics of development.

So, one of the primary outputs are these benchmarks in order to, not reign in the industry, but offer a set of carrots or incentives and scoreboards in order to assess the progress across these AI labs, which are doing some just amazing work right now, but founded in real world tasks that are consistent with public interest.

There is also the intensive margin, and this gets to some of the issues that were brought up in previous conversations. These are the policy outputs or recommendations for within a particular task. The ways in which we collectively would want the public sector to show up in order to regulate the models and the use of those models.

So, these are policy memos that might deal with the openness of the source data or the privacy considerations, or the ability to maintain trust, which is the interaction of the model and the source data.

And then finally, it is aggregating donated compute, open compute, in order to build and train these models out in the open, in the biggest possible way. Until we see evidence to the contrary, there are, I guess we would call it, economies of scale, associated with model training right now, in the sense that it is much better to train one large model than 10 smaller models, or at least it will yield much more effective model development.

But if we're also assessing those models and that model architecture against these real world tasks and the way in which you can develop multiple downstream applications on top of those models, the architecture matters for developer use. And so, training very large models in a way that is consistent with the real-world application, that is the third prong of this social sector activity here.

I love the public sector, and I'm currently in the private sector, and there is value in limitations to both. For the public sector, it brings trust and scale in a way that the private sector just can't necessarily do or trust in scale, but that is not consistent with speed, and we've seen that from the public sector. It's very difficult to keep up with the weekly cycles that are coming out of the development of these models.

And so, it needs a little help here, and that's part of the idea of having a mission-driven nonprofit enter the space.

And then, for the private sector, it is rewarded for things that people are willing to pay for, not what's valuable, but things that people are willing to pay for. And so, this Clay project is meant specifically to sit in the middle of it. And frankly, my PhD is in economics, and this is one of the first times where it's actually coming in useful, which is basically, it's trying to deploy basic concepts in public finance to this model development and digital infrastructure. I think that's really just the early innovation of this, of this nonprofit project.

The last thing I'll say is, there was a comment from the audience before that that inspired this closing remark, which is we've been at this for a really long time and have been poking around the margins of the potential of Earth observation in the long tail of applications, not just defense and intelligence.

And one of them was, five years ago, I was working on a project to help bring Earth observation into schools in order to find illegal gold mines in the Amazon. This was through a project called the Earthrise Media. And it took a lot of very, very smart developers and teachers and motivated students, some of the best, I think, in the country, in order to successfully find these gold mines in the Amazon.

And what I've been struck by in the last couple of years is the ability to provide, even fairly young students, with the ability to find things that would take an investigative journalist being paid for a month to find; we can now find much, much faster, even in a matter of sometimes minutes, where it took weeks before, and has lowered the barrier to entry. And that offers some really exciting opportunities, but it also offers some pretty hairy questions that we hope to approach through this nonprofit effort.

So, I'll leave it there for now. Thank you.

(Applause)

DR. WOOTTEN: Thank you very much, Dan. We do have time now for questions. I started in the room last time, so let's start with one on Slido if we have one, April.

MS. MELVIN: Sure. Can you comment on how often the satellite data that you're using as the inputs for the queries is refreshed for a region? How real time is it? Are you doing the inference every time there is a query?

DR. HAMMER: Well, I'll start with the last question first. And so, in the sense that we run inference whenever a new image is available, we basically create what's known as an embedding for any new streaming data that we have access to. As far as what imagery is available, the public sector, specifically the European Space Agency at this point - and having worked at NASA and clearly a very big NASA fan, it sort of pains me to say that the Europeans are kicking our butts right now in terms of Earth observation - but the Sentinel 1 and 2 missions, they release information for every point on Earth. I would say on the order of, I think it's like once, every five days at 10 to 15 meter resolution, effectively. That's across, it's like the 15 bands, including synthetic aperture radar, near infrared.

And so, we index that imagery immediately upon receiving it, which allows you to search for changes. We don't have to retrain the model each time we just apply that model. There are, however, a lot of commercial providers. And so, there's that old sort of adage in this space that there's spatial resolution. So, what you can see from satellites, can you see windshields of cars or just farm buildings, basically? So, there's spatial resolution, temporal resolution, and how quickly it's refreshed, ranging from on demand to every 2 days to every 10 days to once a year. And then there is cost, and you get to pick two. So, spatial resolution, temporal resolution, and cost, you pick two.

So, the free stuff is available at a baseline of like, let's say, around once a week at 15-meter resolution. If you want higher resolution, you have to pay for it. And if you want it more updated than that, you have to pay for it. It is really important that a baseline of imagery remains in the public sector and built by the public sector. So, it's not just a line item that can be crossed off at the next appropriations consideration for this to be truly in the public interest and with an eye towards equity and application.

So, there's no simple answer there, except for the second part of the question, which is that we run inference whenever we get an image.

DR. WOOTEN: Thank you. Let's go to another one over here in the room.

DR. RAO: Thanks, Dan. I'm glad you mentioned about embedding. Douglas Rao, Senior Research Scholar at NC State University. My question is, who is responsible for creating those embeddings? I think you're getting to that question early on. So, you are sort of doing embeddings on your end for the data that's available from different federal agencies.

But in the sphere of the push and pull between the public and private sectors, or maybe there's in between organizations from foundations who should be doing those embeddings, especially when we're thinking about the multimodality of different data. You're talking about visible and then the SAR data, but then there's much more variety of those satellite observations that can be indexed and in those embeddings.

And when you say you don't retrain model, right now your Clay model is probably based on the Sentinel and some other channels. And so, when you get to other wavelengths, you probably still want to include those modality in your model itself. So, that may still require different set of models for that as well. So, that's more broad question.

DR. HAMMER: Yeah, great question. So, there are two questions in there. Again, I'll start with the last one first, which is that Clay was trained on Sentinel 1 and 2. And I think that's 18 or 19 bands, as well as a digital elevation model. So, I think that's, I don't know, 18 bands now in the latest release.

The way in which it was architected, and this is really important, is that you don't require all 18 bands in order to run inference or to create the embeddings. So, as long as the subset of bands, let's say red, green, blue, and near infrared are included in the original set of bands, you don't need all of those 18 bands in order to create a Clay embedding, which has proven probably the most valuable part of this model, because it allows for you to pull in sensors that it's never seen before, including the high resolution aerial imagery that I just showed from USGS. It's the Clay embedding, but it does not have access. That aerial imagery does not collect synthetic aperture radar information. And that's a really important point for the breadth of use in the interest of the public sector is the way in which those models are architected.

To get to the first question about who should create and release those embeddings, for me, it could go either way: the private sector or the public sector. In full disclosure, I'm working at Legend, which is that seed stage startup, which is basically the value proposition is to be able to, once any model, not just the Clay model, is out there, the ability to very easily - where Legend handles the creation, storage, and serving of embeddings. It is like a critical value proposition. How to create them at very, very low cost, about a hundred times lower cost than just running the models out of the box, how to store them effectively, and then how to serve them up specifically for downstream algorithms to sort through and fine-tune models on top of the embeddings.

And then one of the reasons why that isn't in the public sector, although maybe it should be eventually, if there really is a standard bearer model in the public sector where you would want to create the embeddings, and we all agree that this model is sufficient for the long tail of applications, and it's in the public benefit to bring the imagery, this open imagery further to application, because the raw data is more upstream and requires more processing.

But there is this one model that we all agree should be the standard bearer. It probably does make sense for the embeddings to be created once and distributed specifically by NASA or the European Space Agency. But this industry is so much in flux right now that I don't think it makes sense for that to be in the public sector, because as soon as it's created, it's probably outdated.

And so, right now, the embeddings are created, stored, and served by the public sector, whereas the models can be out in the open. Nonprofit NASA has one of the best models out there called Prithvi right now. But the handoff to the private sector remains the imagery itself, which is an incredibly valuable public good.

DR. WOOTEN: Okay. I think we're past time. So, I'm sorry to have to cut folks off up there and on Slido, but I know Dan's contact information is available, I think, in the program, but if not, we can get that for folks. Let's thank Dan one more time.

(Applause)

DR. HAMMER: Thank you.

DR. WOOTEN: And I believe now we have a break until around about 10:50 Pacific time, and then we'll proceed into our next Panel session on the common challenges, if I recall correctly. Thank you all.

(Break)

Agenda Item: Addressing Common Opportunities & Challenges to Accelerate Action

DR. HOLMES: Well, thanks very much for returning. The sun was almost too beautiful to bring everybody back, but we've got a wonderful Panel. And I was talking with the Panelists outside in the sun, talking about how we're going to try and really think about addressing common opportunities and challenges to accelerate action using artificial intelligence.

My name is John Holmes. I am at the National Academies. I have worked mainly in the area of energy and climate, looking at decarbonization and air quality for a long time. If you're at the Academies for any length of time, you move around on topics, and you get to interact with some of the most wonderful experts, as I have on this Panel today.

To set this up, this is the first group that really is about the Earth systems modeling group, the people that use large-scale climate models and think at the larger scale. Yesterday, we heard a lot about water resources. Use of AI within urban environment and agriculture. Today, and with this group, we're with a group of people that spent a lot of time on the large-scale models that are really important because we use that information when we downscale into individual topics like agriculture and water resources.

I'm going to go in order here. Karen McKinnon is going to give some opening remarks. Karen is at University of California, Los Angeles. What I've heard from her background as statistician and someone who is very interested in heat and in things like that. Douglas Rao will follow up, someone that is very interested in not only large-scale Earth systems modeling, but also thinking about the data. And then we'll end with Monica Morris, who's at NCAR and is not a social scientist. She is a philosophy of science person, which I have learned is different and is an important difference.

The reason I'm glomming onto Monica's background is I'm very interested in the history of science and the history of climate change, and thinking about how we've addressed climate change through the centuries. And so, I always like to look back on, like a book, the Earth is modified by human interactions from 1870s. And one of the lines of George Marsh was within a comparatively short amount of time, we will have gathered information to fully understand humans' impact on the Earth.

Now, I think that was overly optimistic. We're still gathering that information. But what's being shown by this event is the huge amount of information that we have to utilize to understand the impacts of climate. And I want to introduce Karen to start us off.

DR. MCKINNON: Wonderful. Thank you so much, John. And also thank you to all the organizers for organizing this fabulous workshop. It has been so fun to be involved and also for the chance to speak as part of the workshop. So, yeah, as was mentioned by John, this Panel we realized is maybe the first one where we're all in some way or another climate scientists, which I think is an important context for the way that we are going to be thinking and talking about AI. And then it's going to be coming from our background working with climate data.

I primarily associate as a climate scientist and a statistician on the side. And most of how I interact and what I'm excited about AI is about AI coming in as a tool that can allow us to replace our system models or complement our system models, or improve them in some way.

And just for those who might not work in the climate space, our system model is basically a large-scale numerical model of different components of the Earth system, the atmosphere, the ocean, the land surface, and the cryosphere. And traditionally, these have been very large computationally intensive models, large supercomputing centers like NCAR, and provided up until now, basically a lot of the forward-looking climate information we have, as well as a lot of our understanding about climate and climate change.

AI is really adding this new and super exciting tool. And we saw a little bit of a preview on that with the discussion from the Allen Institute yesterday about emulators. So, that's the space that I am working in.

I want to first talk about the opportunities side of this Panel. So, why did I decide to start spending my time thinking about AI? I think about what I can do for climate studies, and also just climate impacts and climate action down the line. So, there are three things that come to mind for me as to why I'm super excited about the potential that AI provides in climate.

The first one is what I think of as democratizing access to climate information. So, as I already mentioned previously, historically, climate models are these big models and these big supercomputing centers. You couldn't really download a climate model on your laptop and run some experiment you wanted to do. With AI models, you basically can actually do that. And again, the Allen Institute has been a huge leader on this. I've been super impressed by what they've done.

In AI, the calculus is a little bit different, in that you do need a big center, lots of GPUs, whatever. Well, not a center, you can use the cloud, but you do need a good amount of computing resources to train an AI model. But once that AI model is trained - and again, I'm largely thinking of basically emulators of bigger system models - the prediction or inference part tends to be relatively cheap. And as I learned a month ago, you can just download the Allen Institute model on your laptop, and even if you don't have any GPUs, you can still run it. It's not super fast without GPUs, but it works. And that's amazing.

That's something that was literally something that is not possible within our system model that now has become possible with these AI emulators. And so, to me, that's number one big deal that people can basically run climate models, think about how the climate model looks at their local scale, and do that on their laptop. And don't worry, there are going to be challenges coming up. So, we're going to be optimistic to start. There are caveats in all of these.

Okay, the other thing that I'm super jazzed about with the AI models is how they make things that in traditional statistics were really hard, they make them easy. And so, that's awesome because, again, things we couldn't really do before we can now do.

So, one example that comes to mind for me is what I think about is data fusion. This is basically producing some AI product that combines information between different sources of data. So, let me give two examples. The first is, let's say, that you want to make an emulator of the real world. This is actually my main interest. I don't like to emulate climate models, I kind of want to make an emulator of the actual world. But we don't have that long of a record of climate data from the real world.

What you can do is something called transfer learning, which is where you basically train your model first on tons of climate model simulations, and then you transfer learn, or basically fine-tune some of your weights with the observations. And I've seen in my own experience and other folks have as well, that this is actually a really effective way to get most of the way there with something that's not quite right, and then get all the way there with something that's more correct. And that's something that, again, in traditional statistics, it's unclear if there is even an analog to how you do that.

Similarly, we often have different types of data modalities. So, we might have satellite data that has these gridded or average estimates of some quantity. We might have station data that's at a point, and you want to maybe combine those. And that's something that becomes a lot easier with AI methods than traditional statistics.

The third opportunity, which I think we'll talk about more broadly, is the idea that AI - and this links to the other two points as well - it can be used to take a large scale climate information and really make it meaningful, useful, and actionable at a local level.

So, right now, there are lots of people, companies, city managers, water managers who might reach out to someone like me and say, what can I expect for my location specifically, maybe it's in the mountains or in a valley or some complex topography. Who knows? It doesn't matter. A single location. And I want to know what type of climate I could experience in the future with climate change.

Right now, our only tool to do that is climate models, and climate models were not really developed to produce this local information. They don't tend to do that well, and that's by design. They were designed to get large-scale structures of the climate system, not tell you about the future of climate in your city. But people mostly are using them for this purpose because there is not another option.

I think that AI emulators and models provide the option that you can train them on a climate model. You can fine-tune them with observations. You can fine-tune them with local information and provide that at a local scale that ideally is trustable, which gets us to our challenges.

I have three challenges I want to bring up. They're not the only ones. And I think that these actually go more broadly beyond just climate emulation to other parts of AI as well. The first came up yesterday as uncertainty quantification. AI models, it just hasn't been developed. There are just not established ways to do what is really standard in statistics, which is really standard in science. We just don't have it in AI. There are some methods to UQ. They can take a lot of compute in some cases, but we just need to move things forward with uncertainty quantification in order for us to use AI output in an actionable way for people to make decisions. We need to know how uncertain our predictions are.

The second is transparency. So, it was kind of funny yesterday in some of the talks; it was like a small model listed that was a million parameters. So, someone who still does a lot of linear regression, a million parameters is a lot. And it's funny that that's now small. Anyways, these models are big. They have a lot of parameters. They can be black boxes. And again, I think if we want this information to be actionable and useful, we need to really understand what's happening.

Some people have done this in terms of, say, explainable AI. In tandem with that, I think that we can maybe make sure that we trust our model predictions in terms of them doing the thing that we as scientists know is correct. So, interrogate them, their physical consistency, and in a more traditional scientific way to make sure that when we're, say, emulating a system with an AI model, it's still more or less doing what we expect and what's consistent with our knowledge of the system. And that can be in tandem with methods like explainable AI that have been developed more generally.

And then the last challenge I want to bring up is this question in terms of, I'm going to broadly maybe put it in the idea of out-of-sample predictions, but more generally as cases where we are either intentionally or unintentionally using an AI model to predict something that hasn't happened yet.

So, the case that probably comes up most clearly in my field is people who want to use AI for climate change prediction. That is an out-of-sample problem. In every future year from now on, we'll see CO_2 emissions or concentrations higher than what we've seen before. We're constantly changing our system right now. So, anytime you're doing, say, climate change prediction, this is an out-of-sample problem. There are a lot of other examples as well.

AI models are data-driven. So, if we're going to do out-of-sample, we need to think carefully about incorporating physical information, or - not exactly sure what the solution is, to be honest - but we need to be aware of these out-of-sample problems concerned with AI models, and we need to be very clear about quantifying the uncertainty if possible and doing everything out-of-sample.

So, I will stop there and pass it over to Douglas.

DR. RAO: Thanks. I think, Karen, we need to be unapologetic to be a statistician. I also got my first degree in statistics. And in the early 2000s, when I was doing my predictions. At that time it was for the stock market, so I gave my mom false hope that I may make money at some point.

I got lured into geoscience with my graduate school advisors looking at, I can do field works outside and using machine learning to do land use and cover classifications on crops and analysis of the ecosystems, which I was really excited about doing field work. So, I switched into geoscience.

With that, I started doing a lot of data fusion work in my thesis, combining the station network measurements and satellite data to hopefully improve the value-added temperature information for global and regional climate-monitoring impact studies. So, my whole progression started using machine learning very early on, and before the neural network was even invented, which is in the early 2010s.

So, I think I'm really appreciating what Mariela mentioned yesterday, that different types of machine learning and AI. and then before the deep learning, in the 2010s, that's all those traditional machine learning methods being used. And in my current work, I do a lot with NOAA through the Cooperative Institute for Satellite Earth System Studies. NOAA has a center for artificial intelligence really trying to understand how AI can be used to accelerate a lot of NOAA's missions, to support those missions, including weather forecasting and disseminating that information, as well as use those satellite data for public goods.

And with the opportunities, I think yesterday is a good demonstration of how the innovation in AI can lead to impact downstream for different use cases, which, first of all, can be the improved understanding of the large-scale patterns, like what Karen has mentioned earlier about emulations and other things. And the other is that knowledge of computation, which I think from the keynote yesterday from Angel Hsu from the UNC, Chapel Hill, really looking at how we use the generative AI tools to translate complex scientific information into things that people can digest and to support some decision making process with trusted validation or evaluation by the experts.

And the other one is also how can we use AI as a tool to improve information services that can be used either for bridging the last mile gaps between what we can offer in Earth science data and information to what things people actually needs to their daily decision making process, like for city planning or for other things, and disseminating those alerts systems about those natural hazards and other things to support them for decision making.

To get there, I think data is one of the most important parts, especially integrating the data from different domains. And one important thing I was listening to yesterday was that different Panels are looking at how to integrate data from large-scale data from large-scale models and cell observations to the things that are relevant to those specific use cases, let it be for the underground water, let it be for the agricultural informations. Whether it's for the labels or for the specific crop types or the information, whether can we use those AI tools to support the farmers making decisions in their farms or in their specific grounds.

And then for the urban planning one, I was having a conversation with Mariela today, thinking about the privacy of those street view image. And so, how can we link larger-scale information to the sensitive information that can potentially have limitations on how we want to use it. And so, this is an exciting opportunity, but also there are a lot challenge come together, which how can we integrate the data across scales, both temporal, spatial, and also across domains. And that's the challenge of that data integration

I think I was talking with John early on this morning about what the future data infrastructure should look like if we want to support developing those innovative solutions. There are different data infrastructure funds, federal or public-funded agencies, and then there are also a lot of those data institutions from private sectors, from nonprofits, and from other parts. And what are the future of the data infrastructure looking like to support those integration of different type data to enable those applications.

And I'm so glad that one of the comments made early on today is about workforce and education. So, I think to get there, we also need a lot of those upscaling and the literacy about AI and data in general, not just for the future workforce, but also for the current workforce. We have been using traditional set of tools, and now AI becomes a new tool that came along and just changed almost everything, at least in the past two or three years, after ChatGPT has become quite popular in the daily news media discussions.

So, how can we, first of all, upscale the current workforce and then also thinking about how can we change the educational accountants to support the future workforce to become AI literate and data literate? And I think one of the phrases I learned last year is that we need to teach people about those ballooning tests, like what's actually just BS, and with all those claims. And so, having the literacy for people to make those differentiations about the right way and the good ways of using AI to support different tasks.

I'm going to stop here. I know there are a lot of other topics we're going to talk later on, so I'm going to pass it to Monica.

DR. MORRISON: Okay, thank you. I'm Monica Ainhorn Morrison, I work as a program specialist at NSF NCAR, and I have a split appointment between the Climate and Global Dynamics Laboratory and the Computational Information Systems Lab.

And as John noted, I guess my position is somewhat unusual as I come to climate science through philosophy of science and ethics. So, my doctorate is in philosophy with an emphasis on epistemology and ethics of climate science, where I largely focus on trying to understand Earth system models, their development trajectories, how they're used, their limitations, how we can communicate that, et cetera.

I'm currently working on ideas related to AI in the context of thinking about responsible and reliable AI practices when it comes to integrating these technologies into our Earth system modeling endeavors. And I do that in the context of helping out with the construction of NSF NCAR's AI Roadmap, and also just thinking about, in general, strategies for bringing AI into the development of things like our Community Earth system model, which is housed out of NCAR.

And just to point out, I've already noted this term once and saying that I specialize in epistemology and ethics. I can't get away from this word. It's one of my favorite. I've defined it here, and I'm inevitably going to use it again. But when I say epistemic, I simply mean relating to knowledge or knowledge-generating processes.

My philosophical background largely shapes how I think about the integration of AI into our climate modeling endeavors, especially when we're thinking about how AI in modeling can be used to accelerate actionable climate research, actionable climate science, and informed decision-making.

And what I see is the central component of future opportunities and challenges for bringing AI into climate science generally, but specifically with respect to modeling, informing modeling, or helping with modeling, is the management of epistemic and ethical risks, which are an inevitable part of modeling, but also the application of modeling to informing decision making.

And ultimately, what I'm going to argue is that it comes down to developing infrastructure to tackle these sorts of risks. So, as we're developing technological infrastructure and we're accelerating model development, we also need to have epistemic and ethical infrastructure so that we can have reliable products going out there and they can be responsibly purposed by people that might use them in high-risk or high-consequence decisions.

So, let me just go to my next slide, which will be the last one. So, as AI allows us to generate climate information at unprecedented scales and resolutions, the central challenge, I think, is bringing intelligibility to that endeavor. And this is not just looking at whether or not we understand what the model is doing, but as Karen pointed out, understanding whether or not the model is acting in ways that is consistent with our well-established physical theories.

This also means having an understanding of what assumptions are embedded in our modeling. This means having an understanding of what the uncertainties are, having an understanding and knowing what tendencies for error might be and what failure modes look like, documenting those, and understanding what that means for the adequacy of these models and their products for use in different contexts, as well as what the ethical risks of their deployment might be, especially those risks that have material consequences for vulnerable and marginalized populations.

And so, intelligibility is really the foundation of being able to manage epistemic and ethical risk, which I'm going to talk about that risk in terms of three central things. The first is representational fitness, general fitness for purpose in terms of it being relational and ethical, and then lastly, thinking about responsible integration. So, how we take these theoretical ideas and operationalize them to provide practical guidance, which is essentially what we need to do in the development of that, as I noted, epistemic and ethical infrastructure that needs to accompany the technological infrastructure and development.

So, climate data sets, which we're using to inform all of our model development, are not exact representations or descriptions of the reality out there. The representations, which are shaped by who's constructing them, who's putting the observational instruments together, what their purposes are, what methods they use, how that data is post-processed, and then interpreted.

And climate data is really interesting because it's what we call phenomenon-agnostic. So, this was introduced in a paper by Lloyd et al in 2024. Unlike biological data, the data is not an exact description of some object that exists in the world. What we do is we construct definitions of phenomena like atmospheric rivers based on our purposes and then interpret the data in terms of its description of a state space and various properties of that state space, and then say, oh, there is an atmospheric river in the data. And so, it's open to representational interpretation, if you will.

And these data sets can adequately or inadequately capture the things that we care about and the things that communities actually experience. And so, making sure that the data that we're using to train and validate our models is representationally fit for those model purposes is incredibly important.

To give you an example of places we might need to do a little bit more work, we have, as a big concern for adaptation and resilience, things like high-impact, low-probability events. Extremes. They're called HILPs. A lot of the data that we have out there fails to capture these events, so that models training will pick up on this phenomena and represent it when they're applied to questions. And we're using reanalysis data sets like AR5. This is one of the big ones that people are using to train various sorts of emulators and models. And these have notorious biases and gaps with respect to the ability to represent these high-impact precipitation and heat events.

We really need to be interrogating the data and understanding its representational fitness, especially its limitations, as we go about training these models and being transparent and communicating honestly about those limitations.

We also, I think, need to develop and extend the way that we're evaluating models to adopt a more holistic fitness for purpose way that we're evaluating models as opposed to our traditional methods, whereby we evaluate model skill according to whether or not it's consistent with the observations that we have.

Empirical adequacy, even when it comes to our physical models, is a necessary but not a sufficient condition for determining that a model can actually perform in the way necessary, as even with our physical models, we see instances in which we get agreement with observations because of underlying compensating errors. This is something that can probably happen in our AI systems and that we need to attend to, which is why, as Karen noted, we need to interrogate to actually understand whether or not we have adequate representation of processes that give rise to that emergent behavior where we see consistency with observation.

And extending fitness for purpose beyond this epistemic fitness, we also need to think about ethical fitness. So, ethical fitness will lead us to ask questions like, what are the potential errors that might arise from the system, whose risks are those, and what vulnerabilities do they have? So, that we're not just looking at whether or not we have agreement of the observation, but with respect to the uncertainties and the potential for errors we see, we're also trying to make a judgment as to whether or not that's consistent with the ethical values that we hold and the interests of the communities that might be using this information hold.

And the last thing I want to say is that once we have things like these holistic fitness frameworks, we have things or we're developing methods for uncertainty quantification, we have standards and protocols for evaluating models beyond just considerations of empirical accuracy. We need to really standardize and systematize that and make it a central component of what we're building as we democratize access to these systems.

If we can't responsibly deploy AI, if we don't know what it's telling us, we don't know the risks associated with its uncertainties, its errors, and its limitations, and we don't communicate those so that downstream users of the models or of the outputs of the models can make informed judgments about whether or not, given those epistemic and ethical risks, that information is fit for use in their decision context.

I just went over a bunch of challenges and wasn't really much of an optimist, but I present this as an opportunity because I think it's a collaborative opportunity. We can all bring this as something that needs attention alongside technological development, but it's also maybe more of a challenge to build this epistemic and ethical infrastructure, mainly because it requires exercising caution with our development and our innovation.

It requires oversight. It requires governance structures. And it maybe requires a certain tempering of our excitement at these technologies that might put us in a position to deploy them without rigorous consideration of limitations they might have.

Thank you.

(Applause)

DR. HOLMES: I could hardly wait for her to sit down because I really wanted to start the Q&A. This is a wonderful topic. And what is nice about this is that we want to think about a little bit more about this use of AI and the use of these systems and models in decision-making.

And I want to throw out - and we were talking about it beforehand - the subnational context. Because even though we might think that climate change at a national level is very politicized, and it is, many states, including those that may be one color or another color, are extremely interested in understanding climate change impacts on their locales. Subnational climate assessment cuts across many different political spectrums because of the realities on the ground of the changing climate. You don't have to say what's causing it to understand that there is something going on.

So, we're going to move into moving this into a little bit of what's generally needed for adaptation and resilience when you think about these topics of use of AI, use of the data, use of emulators, or whatever. And who's going to go first? Let's go to Karen first.

DR. MCKINNON: Okay, sounds good. Yeah, and as you might remember, this is one thing that I mentioned as an opportunity, and I want to mention both what I think is a near-term, meaty opportunity, and then also something that's still a challenge for this.

Okay, so imagine that you're a city planner, water manager, or something, and you need to make some plans for the next five or ten years about what's going to happen, what you need to manage in terms of heat, precipitation, and climate conditions.

We talk a lot about climate change. So, for like heat in particular, we can expect that we're going to go towards warmer temperatures over that period. But on top of climate change, we have this other side that's a major focus of my research, which in climate we call internal variability. But this is basically just all of the random noise that happens in the climate system that is not related to climate change, but can still cause major high-impact extremes. I've never heard HILP, by the way. I'm not sure I like it because it's so hard to say, but high-impact, low-probability extremes. These things would happen without climate change. If we didn't have climate change, we would still have high-impact extremes that would affect us.

When I think about the information that we need for adaptation, we need to quantify both of those components. What is climate change doing at a local scale? And then what could happen in terms of sampling of basically weather, but on longer timescales, if you're thinking about the annual temperatures or short timescales, if you're thinking about a discrete extreme.

On that internal variability side, for me personally, this is where I'm most excited about the potential for AI to basically expand our historical record of the types of high-impact extremes that you could get, or the types of climatic conditions you could get by, again, basically creating an observationally-based emulator that we can understand. Not only did we see what we saw in the last 70 years of that historical record, but this is a more complete sampling of the extremes that we could see.

And so, the extreme that, for a lot of the climate community, was a real eye opener recently, was the 2021 Pacific Northwest Heatwave. For folks who don't remember this or weren't aware of it, it was the temperatures and the typically moderate Pacific Northwest were basically record-smashing. So, just to give one number that always sticks in my head is Portland, Oregon, which typically has a moderate cool summer, reached 115 degrees Fahrenheit, which is not a temperature you usually think about Portland. There were temperature extremes in Washington, and actually, it was also centered more in British Columbia and Canada.

So, this was an extremely extreme event. It took people a while to even understand how it happened. Put a probability above zero on this event. But what a number of analyses, including, I'll admit, my own, showed over time, is that, yes, there certainly was a climate change component of the heatwave. That's not surprising to anyone, I don't think. But of course, the intensity of the heatwave was also just due to sampling a previously unsampled, very, very high-impact extreme.

And so, if we had known in advance or thought about in advance what the full envelope of things we could see, even without climate change, that would help us realize that we could get these very extreme events in the Pacific Northwest. And maybe their heat planning team needed to be aware that it's not likely, but yes, you could reach 115 degrees Fahrenheit in Portland, and perhaps you can reach that again. So, this emulation that's producing more data to sample the extremes, I think, is something that's super exciting and promising with AI.

Projecting climate change with AI is a lot harder because all of a sudden, we're doing this out-of-sample thing. We're trying to say we actually haven't seen climate change yet. And so, typically, how we've done those projections is with physical-based models where the projection is driven by the physics. And in AI models, these are data-driven models; they perform best in sample, which is well known.

I think one major potential with projecting future climate change at a local scale is that if we first take our physical-based models, we assume that they give us this future projection. And then basically downscale, by which I mean move course resolution information that our climate models provide to a finer spatial resolution, and that ideally gives us a better estimate of what's going to happen at these local scales that we need to adapt on.

But I think as a community, we need to be aware that the projection is an out-of-sample problem, and that we need to continually make sure that we are as correct as we can expect to be, given sampling of internal variability going forward, because we are going to keep getting data every year. Unfortunately, it will probably include some surprises, and we need to be very active in incorporating that information, assessing our models, and knowing if we're being successful, especially if we want people to use that information on the ground.

DR. HOLMES: Beautiful. And that that Pacific Northwest heat wave changed residential consumption of electricity and technologies in a fundamental way, and it also probably changed the electricity systems' assessment of what peak demands are going to look like on their system. So, these types of events have fundamental changes on our decarbonization and our technology.

Douglas, do you want to -

DR. RAO: Right. I think I'm going to use another hot topic that people are talking about, which is digital twins, which people started using digital twin in different ways in development. And I think there is opportunity for digital twin, and AI embedding in the digital twin development processes to enable a lot of things related to adaptation resilience.

I think for those people who are not aware of digital twin, especially the recent development in the European side, it's the destination areas where they're trying to do is to have a digital replica of the Earth systems that can support different use cases, including climate adaptation, which is one of the use cases, and extreme weather planning.

And so, in different ways, and going to what Karen was mentioning with AI, that can really enable a lot of rapid simulations on those emulations about different phenomena that can allow people to play those what-if scenarios. So, what if we did this type of adaptation or this type of evacuation strategies and how things will be impacting. And so, this is one technology that can be useful if designed in the right way.

And I think maybe others in the room, they can talk more about digital twin for the agriculture systems. We're thinking about what is a large-scale patterns of extreme weathers and how it impacts on the agriculture and different systems, and can we plan it better to support the resilient agriculture and for the farmers.

And I think there is growing interest in the twins at local and regional scales, and especially thinking about urban planning and urban use cases. I was talking with some of the city planners in the U.S., and they have been using digital twin technology for things related to the emergency responses, as well as smoke and air quality related to wildfires, for which you can do really rapid modeling to see how the smokes move across part of the city and then thinking about how should you issue some of the alerts to the residents and thinking about for their health protection and other activities that should not be done outside.

If you can do accurate modeling of those and predicting what it might be, that leads to the adaptation and all, also the mitigation of risks for the community as well.

DR. MORRISON: I agree with both of those and I'm going to piggyback off of what Karen said about the potential and what I think is the large potential to utilize emulators, so better sample the space of probability so that we can get a better account of what these high-impact, low-probability events might look like.

Currently, the computational cost of developing large ensembles with our coupled Earth system models that are physically based is beyond what we can achieve. So, this is a really promising avenue. And even if we're doing it with our course resolution models, this is helpful because that information and that ability to better sample the space of possibility of what future extremes will look like can be used to inform things like the development of story lines, which are really good because they integrate well into localized context, which is one of the problems that we have with course resolution or these coupled Earth system models that do not admit themselves to treatment of things that are happening at local scales.

That's where I think there is promise with something that Douglas said, where we can take models from our CMIP, coupled model intercomparison models that are coarser resolution, and integrate that into other forms of data to better contextualize information from these longer-term projections.

And then I think the other thing that's promising, maybe not something that admits itself or looks like it's highly relevant to adaptation resilience, would be using AI to improve our physical models. So, we talk a lot about the development of AI models that can take the place of physical models, but I really think that hybrid modeling is a promising avenue forward.

Some of what's happening at Columbia with LEAP project, which a couple of people are associated with, who are sitting in this room, is to replace parameterizations, which are representations of small-scale processes that are not resolved by our Earth system models, with machine learning models. These are able to not only increase the computational efficiency of the models but also, if we have poorly understood processes in our models, they're able to increase the fidelity of those representations to the extent that they can be informed by observations.

I also want to point out the utility of utilizing AI to do things like perturb parameter ensembles, which help us to better constrain our models so that we can understand them and we can get them to perform, especially in that parameter, that physics space, in ways that are more consistent with what we think processes in the outside world look like. And they give us a better understanding of our models because we don't only need intelligibility of our AI models, we can increase intelligibility of our physical models. And AI is actually something that we can leverage in order to do that, especially with these large Earth system models that are highly complex, millions of lines of code, hundreds of parameterizations across multiple components.

DR. HOLMES: So, continuing along this local and subnational decision-making utilization - I'm going to start with you, Monica, just warning you - what do we need to do to establish, maintain trust? What are issues associated with maladaptation for use in decision-making? These are big issues, especially when you're considering local and state context, where the expertise isn't amongst a whole bunch of people that have been involved in the development of many of these tools.

DR. MORRISON: Yeah, so, can I just refer you back to my presentation? I think, ultimately, it comes down to what I noted earlier. We need to understand these tools that we're developing. Ideally, in a utopia with a U in an ideal world, we would be co-developing these tools, and we would have tools that are tailored and fit for the purposes to which they're being applied.

Unfortunately, we do not live in a utopia. So, we have to find other proactive ways to prevent maladaptation, where we're dealing with a limited number of tools that are not fit for all purposes. And so, I think trustworthiness comes down to being brutally honest about where these tools have limitations, where you can't apply them, where they can be applied but you need to exercise caution, et cetera.

I think that comes down to understanding the fitness of purpose and effectively communicating that fitness. Not only in terms of what questions is this suitable for answering, but we also know that there are going to be uncertainties or potential for error. There are biases that might lead to overestimation of certain phenomena and underestimation of certain phenomena. Being able to also look at the implications of those uncertainties and errors, as I noted, in terms of what those ethical implications are.

So, what are the risks of an already vulnerable community to using this model when it might underestimate exposure to PM 2.5? What are the risks of using this model when it was trained on data that might not adequately represent the heterogeneity of an urban environment and miss certain aspects of heat that are due to lack of tree cover or tree cover in certain areas.

That takes a lot of resources to interrogate models to be able to develop that information that's necessary to be able to say something about model fitness and communicate it, but we're dedicating a lot of resources to the development of these tools, and so, I think that we can allocate some of those to doing this epistemic ethical work too.

DR. HOLMES: Douglas, I'm going to pick up on a couple of points that you said about education and data. What do you think we need to do when we're bringing these models down into the more community, local level, to not require PhDs or anything like that? But how are we going to train the people that are trying to use these models in a decision-making context?

DR. RAO: That's a hard question. I work in a university setting, but I think overall thinking about, first of all, they should have a basic understanding of how the AI model is developed. It doesn't necessarily mean they need to know the whole infrastructure of the model itself, but understanding the impact of the training data itself leads into the outcomes of the model itself.

So, a lot of the developments online right now, a lot of online AI literacy training to give people the basic understanding of what AI is, and then how we should understand to evaluate those tools. And that doesn't require to take an entire degree in AI or computer science to do that. And one of them, I couldn't remember the exact name, but from Finland, they did an online course that is just basically websites that everyone can go through the basic steps.

And then the other thing is to have those domain-specific examples of how those tools developed and how people can play around with it. I think the trust comes when you interact with those tools in your daily lives. And so, a lot of new technology, like those computational notebooks, allowing you to play around with some basic settings of those models and tools to understand when you change certain part of the model itself, how it performs, and react to those changes. And if you change the input data, what it may look like in a difference with different input data, and then what if you add additional noise into the data.

So, through NOAA Center for Artificial Intelligence, what we have been doing is to translate those existing AI deployments in the operational settings into those educational materials that can be used to let you understand, oh, this is now something new that has been used in the operational setting, and this is how it's being developed, and then you can play with it either in your educational setting in the classrooms, or you can just play with it on your own computer to look at how those have been developed and how those are evaluated. That gives people the basic understanding, so they can have better trust in it.

DR. HOLMES: Monica mentioned storylines and how important those are, and maybe exposing more of the public to their current experiences with simple things like machine learning that they already experience every day if they drive in traffic and wait for timed lights or things like that. That may help getting people - because I think what you're talking about is bringing the K through 12, K through 16, better education, but I think there are the general public.

Karen, any thoughts on this?

DR. MCKINNON: Yeah, that was a wonderful answer. And I think I'll look up your notebooks and use them in my class, actually, they sound like they're a wonderful resource.

Yeah, the one thing that came to mind, just in terms of the trustworthiness and the education, is that a lot of these tools are at a research level now. So, whether or not they're ready for prime time, I guess, is debatable, but they're also publicly available, so if someone wants to use them, they're out there.

The one thing I wonder is if we, as scientists developing these tools, just need to be a little bit more honest than we traditionally have been about both their strengths and their drawbacks. When you write a paper, you're mostly talking about how your method is awesome. Maybe you have a few paragraphs in the discussion about how there are a few caveats, but really, you want the viewers to feel it's awesome, so you focus on that.

But it's like what Monica was alluding to, what if on the website, you just have this really clear list of caveats, of like, in these regions of the world, it doesn't seem to work that well, you haven't validated these regions of the world, this variable, or all of these things.

I can see that especially being true for these emulators, that again, in the way I'm thinking about them, they're really full global models. So, you might be producing global-scale information, but you haven't really had time or resources to do the validation at certain spatial scales or certain regions, which is totally understandable, as Monica mentioned. These things take a lot of time and resources.

So, yeah, I think the only thing I'd add is just maybe a call for more honesty on really where we are with these tools and some of the drawbacks or risks that can come with using them.

DR. HOLMES: I know I took us a little bit off and I also wanted to spend as much time with these folks as possible, but I'd like to get a couple of questions, one or two in the room, if you could go to the mics. And the first person that gets to the mic gets to ask the first question. How about that? Just a little bit of Hunger Game type things. Oh, there we go. Oh, come on. Some people are shy. Go in the blue shirt. Go.

DR. HAMMER: Hi, everybody. Thanks for this great Panel. So, one thing I was curious to hear your thoughts about. Normally the paradigm in the climate modeling community has been these models were developed and then we had a decade to sit with them and explore them and look at the ins and outs of what they do and they don't do well. AI and other disciplines has changed that. Now we get new things every week, changes, adjustments, not very well document or explored.

So, I was curious to hear how you think about building frameworks or how we can adapt to this new space where we have new global models coming out every month and how, as a community, we deal with that from the stance of knowing what these are and can do.

DR. RAO: I can start. I think that the driving factor of those rapid development is about benchmarking. Everyone just get on to do the specific things they want to be for those for those models based on benchmark.

I know people in this room have been working on benchmark for climate modeling for quite a while. And so, thinking about what, as a climate modeling community, should we set out as a more comprehensive way to benchmark different model developments and then including domain-specific things we want to understand in those benchmarking.

I think some people in this room are working on a new framework of benchmarking climate modeling, not necessarily just for AI models, but also thinking about physics-based models as well. And I think CMIP, the Coupled Model Intercomparison Project, there is the benchmarking task team, looking at what are the basic metrics all the models should be evaluated against. So, that should be applied to other AI-based models as well. I don't think it should be anything different than the traditional models.

DR. MORRISON: I'm going to fully agree that I think having benchmarking protocols and standards is really important. I would just say that I believe that the Sema Model Benchmarking Task Team is working on this, as are other benchmarking efforts such as ILAMB, which looks at benchmarking the land models that we have as components to these coupled Earth system models.

But going beyond just looking at, again, the empirical adequacy, and also trying to understand the way in which uncertainties and errors and observations might impact our ability to benchmark and get the type of information we need about model fitness via benchmarking. And so, we need a more holistic treatment of what benchmarking is to include protocols for looking under the hood of our models and not just evaluating their skill via these empirical metrics.

I would say this is why we need to build up some of that epistemic infrastructure, to go beyond even what we're doing in our physical modeling system or endeavors to have more holistic treatment of the way we're evaluating both our physical and AI models.

DR. HOLMES: I'd love to go to Slido now I'm going to try and cram in as many questions as possible. We're both going to do as many questions and short answers. So, give us a Slido question.

MS. MELVIN: In the public and private sectors, have you seen a deliberate, careful implementation of AI, or have you seen agencies pushing out AI applications quickly because it works, despite its effectiveness?

DR. HOLMES: Karen, we're going to start with you.

DR. MCKINNON: I don't have an answer to this question.

DR. RAO: I think there are really deliberate ways of thinking about how those can be implemented. I've worked with a lot of colleagues, and Gemini has been included in the Google Workspace that NOAA is doing. And so, there are a lot of the pilot experiment to see how things are working and then also looking at, how can we evaluate the efficacy of that for people's daily workflows.

I think people are thinking very deliberately and then with some more detailed information about this is pilots and then see how this works. And then if it doesn't work, maybe when you think about either improve the training or updating how those can be used, and then providing guidance on the specific use as well.

DR. HOLMES: Let's go to the next question up there, please.

DR. BERTETTI: Thank you. This has been an excellent Panel. My question is for Monica. So, the premise of the climate modeling community, which I'm a part of, is that if we build a really accurate representation of the Earth's climate, it will be useful. And that's a very slow process. But it is a choice. And would you recommend that the climate community, of which I'm a part, reevaluate the way that we allocate resources? Because essentially we treat this as a science problem that might end up being useful. And should we take stock of that presumption?

DR. MORRISON: That's an excellent question. And hopefully I don't upset too many climate modelers with my answer.

First, I think that the idea - and maybe you're alluding to this one model to rule them all idea that's come in, where if we pool a bunch of resources into having the model that most accurately represents all the processes, then we'll have a model that can be useful for all purposes sort of thing. I just think that that is a bad assumption because in any model development, whether we're talking about physical models or we're talking about AI, model developers are faced with a myriad of decisions. And you have to make difficult choices about what you want to represent and how you want to represent things.

This, in AI modeling, might come down to making decisions about what data sets you want to choose because you're never going to have the vast amount of data sets, at the quality necessary, that represent all the things you care about. So, there are all these trade-off decisions. So, I don't think you're ever going to have a model that is like the one that you describe.

I think we would be much better situated having an ecosystem of models that are tailored to different sorts of purposes and intentionally developed down those fitness-for-purpose pathways. And that can be a mix of AI models and physical models.

But I do think that this idea of a CERN-like modeling endeavor that we all engage in, where we funnel billions of dollars into the development of this really sophisticated hybrid physical AI model, is just based on a misunderstanding of what model development actually looks like in practice.

DR. HOLMES: Karen, any thoughts?

DR. MCKINNON: I think it's overall an excellent answer. I don't know if this is a direct answer to the question, but I do wonder if it is worth rethinking about what is the magnitude of resources that we want to continue to put into traditional climate modeling versus other approaches, which is paired with the question of, what is the outcome that we expect or want from all of this?

It's a complicated situation now because we're, of course, in a very anti-science, anti-climate political structure. And so, I'm very hesitant to feel like I'm saying we should take resources away, but I don't mean that. I just mean we allocate with clear eyes towards, is there a societal goal, or just a pure scientific goal, which is fine, too. But like, what is that goal? And are the resources currently allocated appropriately to reach that goal at the end of the day?

DR. HOLMES: Hey, we're really cranking through questions up here. Maybe we'll have time for two more. I don't know. I know I'm going to take us a little bit over, but go ahead.

DR. LUERS: Great. Thanks so much. Amy Luers. Great Panel discussion. I'm going to squeeze in two very short questions. One is for Monica.

You presented these risks, which was really helpful in terms of thinking about responsible AI And you did say this, but I want to just push you a little bit that you presented risks without really solutions. But if you think about this in a practical way, I work at Microsoft, and we do have different tasks that we have for responsible AI and different things, and for trying to do that for energy as well, and various different things. And one of the challenges, of course, is how do you do that in a way that's really operational and scalable?

I want to push you to - I totally get the risks. I totally understand that, and highlighting them is valuable. I also heard there is a lot of opportunity here. So, how can we get to a place where there is a structure? I'm going to give the other question, too, so you don't kick me off before the other question.

The second question is for Karen, and maybe others have thoughts on this. When I first started engaging with the AI community, now eight, ten years ago, the one thing that I always said was, but it's not going to help us with a climate problem because that's about a world we don't know. You can't put these data in. But since Generative AI has expanded, the one area that I've found really interesting, having had a history in climate risk management and resilience work, is, wow, how do we use this with physical and AI, more traditional AI, and generative AI, to be able to begin to do scenarios in a way that is forward looking, combining the issues of the physical with these different tools, to be able to give us much more powerful adaptation at different scales.

I just wondered if that was something that is being explored. It seems like it's a really rich area of research and of practice.

DR. MORRISON: I'll try to be as quick as possible. I think it actually needs to start with some conceptual work because there are a lot of different people talking about things like trustworthiness and reliability and risk. And we don't necessarily have a shared conceptual landscape or definitions that we can operationalize to describe the sorts of things that we might worry about.

There are sources of these things in data, there are sources in models and algorithms. And so, I think a lot of conceptual work needs to be done, and it needs to be collaborative because some of that conceptual work happens, but it's to some extent fragmented.

I'm also, just to be clear, thinking about this in terms of those risks that come from models that are either hybrid or system models or other forms of models. So, I'm being very specific about what the strategy might look like for AI in that context.

I think once we have that, though, that there is actually potential to develop tools that we can make open source as we make our models and data open source, that people can apply to interrogate these sorts of risks according to their own parameters. I think there is potential for AI to help us assess the risks, even though it also presents us with a lot of these risks.

I really think the starting point is to co-develop a framework by which we can actually identify sources of these things that present ethical and epistemic risks, and to understand what they look like in different contexts in which AI might be used, and then to work with people to co-design tools so that we can interrogate them better and scale that interrogation.

And then I think the other part of it is that we need translators. Because that output that you might be getting from these interrogations about where data might have gaps, or there might be underrepresentation or overrepresentation of phenomena because you have biased data, these sorts of things, it needs translation because it's difficult to understand what the implications of these sorts of errors or these sorts of epistemic risks especially might be.

And those translators, I think, going back to what Karen said, could be really helpful for constructing websites, where we say, look, here's this tool we've made available. Here's how people have interrogated it. Here's the things that they found about what its failure modes look like, what its error looks like, and what its uncertainties are. Somebody's translated that to say, oh, here's a potential ethical risk of applying this in certain contexts. And resources need to be allocated to the development of that infrastructure, not just the development of the models themselves or the AI technologies. So, I'll just reiterate that.

It can look something like, oh, here's a nutrition label, or even if we design something like AI readiness levels for our various models and technologies that are not only based on technical readiness but also epistemic and ethical readiness. I think we have gotten away from thinking about those components and technological innovation, development, and deployment with the AI hype.

DR. HOLMES: Karen, do you want to start on the second?

DR. MCKINNON: Yeah, so I think the question is mostly focused on this out-of-sample thing, which is, can we actually use this for climate change projection, which is one caution I gave before.

From where I sit right now, the answer is I don't think we totally know. So, we know that these models are initially, for instance, data-driven. That's the concern with the out-of-sample approach. But at the same time, when I think about AI emulators, like what the Allen Institute has, they include, for instance, CO_2. So, you could think, okay, the model learns what CO_2 does to the atmosphere. We just give it future projections of CO_2, and it does the right thing.

But then people have looked at the model, and it seems to not quite do the right thing with CO_2. So, again, you take that physical intuition, you analyze the model, and you say, oh, actually, this doesn't seem to be doing what we physically expect, and therefore maybe we don't trust it. So, I would put it in the category of we have work to do, but maybe we'll get there. I'm not saying we can't. I think we just have to have work to do.

I'll just very briefly say, I think the other component on adaptation is I always like to push people about how much real high-quality climate information do you need? Like, I live in Los Angeles. Like everywhere else, Los Angeles will get hotter. This is not surprising. We don't know maybe exactly how much this is going to. So, if you're a city planner and you're talking about you look at the spatial structure of the city, there are certain parts that are hotter than other parts, if you want to make a heat plan, I'd just start with places that are already hot are going to get hotter, and that's true everywhere, but it matters more if you're already hot.

So, on the way other side away from AI is also just a question of adaptation of, how far can we go, or how much value add are those really, really high-quality, fine, detailed climate projections versus, how far can we go from our common sense, but very, very well-founded understanding of the climate system?

So, I'll grab you at lunch. I'd love to talk to you more.

DR. HOLMES: I hate to say it, but this is a good place to end because it really links where we are with these large-scale models with the on-the-ground reality that we're all going to be facing.

As Karen said, it's going to get hotter, and the places that are already hot are probably the ones that we want to think about in terms of mitigation, adaptation, things like that. I don't want to say this is not too hard, but there are some elements that we can really take. We need models to help us understand the spatial-temporal distribution, but there is already truth in the knowledge of our city planners and our local residents, and things like that.

We're now going to do one poll, Opportunities and Challenges. In your view, which of these areas is a top priority for fostering accelerated climate action using AI in the near term, creating a path to democratization of information and modeling platforms, addressing uncertainties in climate information, developed using AI models, addressing uncertainties in the usability or trust of AI tools, integrating available data across disciplines for use in AI models, or others. You've got 40 seconds.

It's pretty clear, addressing uncertainties in the usability and trust of AI tools that can inform decision-making. Is information fit for purpose? Nice way for us to end. Thank you.

(Applause)

(Luncheon break)

AFTERNOON SESSION

Agenda Item - Enhancing Cross-Sectoral Partnerships

DR. HAMILL: My name is Tom Hamill. I am the director for the numerical weather prediction at the Weather Company. Think of the Weather Channel as an affiliated brand, and perhaps you have our cell phone app.

I joined the Weather Company close to four years ago now after a large part of my career in NOAA, where I worked on global ensemble predictions. The closest I came to actual climate prediction was working on seasonal forecasts. So, I do not pretend to be a climate expert, but I hope I have relevant expertise in cross-sectoral partnerships here.

Let's talk about generating climate predictions and projections. If you go back 10 or 15 years ago now, these are mostly performed by governments or government-funded institutions, such as the National Center for Atmospheric Research or NOAA's Geophysical Fluid Dynamics Lab.

And why was this? It was because these models at the time were conventional models, models that codified physical relationships like $F = MA$, $p = \rho R T$, rather than being data-driven. And as such, especially implemented as coupled models where we're bringing in the ocean state, the sea ice state, aerosols, and so forth, they were extraordinarily computationally expensive to run, and it took thousands, literally thousands of years of cumulative scientists' time and lots and lots of data and computational resources to develop them. And so, this was really an investment that was too big for even the big players like a Google or Microsoft to be thinking about.

Now, starting around 2020, there has been, of course, an acceleration in deep learning for weather prediction and climate prediction. And these are conceptually different. They're driven by data. At the start, it was ERA5 re-analyses. Now it's broadening out. And though early experiments started at the European Center, really a lot of the early advances were made by private companies: Google, DeepMind, NVIDIA, Microsoft, other companies in China. And these were, again, trained on, typically on ERA5 re-analyses. While that training was computationally expensive to do, the inference, the predictions that could be generated after you have a trained model, are maybe a thousand times less expensive to generate.

Now, that's one revolution, the revolution of radically lower computational expense. Another revolution has been what we call in this field, foundation models. These models are not just trained on one data source like ERA5, but they may bring in CMIP data or global forecasts from other models or regional forecasts. And by being trained on that diversity of models, that makes them more generally applicable to new situations here.

So, if you're interested, for example, in a climate forecast for an emissions scenario that wasn't part of the suite of CMIP5 tests, let's say, well, that may be possible with a deep learning model in a way that's so much more computationally tractable than it is with a conventional model.

The barriers are also substantially lower for new players getting into this. And so, it started with a Google and a Microsoft, but now it's smaller companies like the one that I belong to that are becoming players in this game. And it's even smaller companies like a WindBorne Systems, with a staff that number roughly a dozen or two dozen, that are being involved in deep learning weather prediction and climate prediction.

This makes this more of a cross-sectoral whole-community enterprise. It's not just concentrated in government agencies. And we're growing our discipline by having folks like data scientists join us, not just traditional atmospheric scientists.

This growth factor has been particularly strong in the private sector, and the federal government wants to catch up. Now, how do they catch up? Well, it's those partnerships that I think are really key. And the nice thing about this particular field is that there are mutual interests in partnership, both on the government side and on the side of, be it academia or industry, here.

The government tends to have the data, the re-analyses, or the observation collections that are necessary for a data-driven approach. The advanced techniques, traditionally, have now been generated by these private companies. So, there is a natural basis for this collaboration.

Two years ago, I attended a workshop related to weather, not climate prediction, talking about deep learning for numerical weather prediction. And there were scientists, not just from NCAR and from other government agencies, the Navy, and so forth. But Google was there, NVIDIA was there, Microsoft was there, Weather Company, and more. And there was real energy that was in that meeting because we realized we needed each other. The government had the resources of data, the verification software, and for that matter, even the stakeholder relationships that people cared about. And the advanced techniques were being developed in private industry.

I remember the Environmental Modeling Center - this is the one that generates the numerical weather prediction guidance for NOAA - their director, Brian Gross, said that he'd been 40 years in government and he'd never been to such an intellectually-engaging and productive workshop as that workshop. So, I think that's an example, one example of how the community is actually thirsty now for this collaboration. And I think we can build on that.

Each sector has something valuable to contribute, consistent with their roles and responsibilities. And I don't mean to neglect tensions. I think, for example, one tension is going to be the tension about intellectual property. A company may want to hold close to the vest the stuff that they've developed. But by and large, we've been lucky to have players like Google and Microsoft who've shared their early tech. And I think that even those companies have gained by getting a lot of feedback on their techniques that they would not have gotten otherwise.

And so, the relatively astonishing rate of progress of deep learning for weather and climate prediction is due to this culture of openness, and let's build on that moving forward. So, thank you for a chance to give some early remarks.

(Applause)

DR. LUERS: Paul, you want to go?

DR. BERTETTI: Thank you very much, Amy. I'm Paul Bertetti. I'm the senior director of the Aquifer Science and Modeling Group at Edwards Aquifer Authority, which is a regional groundwater management agency. And we're not a water supplier or water provider. We regulate and manage the permitting system that allows people to withdraw water from the aquifer.

We're unique in the state of Texas in that we have legislative mandates, not only for the total amount of pumping that's allowed based on this permitting system, but also, we have restrictions on pumping and mandated reductions based on conditions when we go into drought periods. And that's, like I said, unique in the state of Texas and limited to our group.

The aquifer is extensive. We have a large amount of permits. The total amount of permitted water is like 572,000-acre feet. And that's each year. So, it's a significant amount of water potentially.

One of the things I wanted to emphasize here is that we're a regional organization. So, we've been talking a lot about these big picture global concepts and weather and data. We're a little group in South Central Texas, but we need specific information to do the things that we have to accomplish in the timeframe that we need to accomplish them.

And so, we developed some partnerships with a bunch of people. And so, whether it's my staff, Hakan, Changbin, Logan, or Adrian at the Climate Adaptation Science Center, Debaditya, and his post-docs at University of Texas, San Antonio, which provide us with our AI modeling expertise. So, this merge of climate modeling. We have a couple of modelers in-house that help facilitate that. All those need to work together to facilitate our goals.

Just to get you oriented, again, we're in South Central Texas, San Antonio. The aquifer system spans about 180 miles across from west to east, about 300 kilometers, across a better part of about eight counties. The aquifer system is important in that it's a karst-based system. So, that means it's a limestone aquifer that has solution enhancement: caves, sinkholes, and enhanced solution fractures. So, the recharge mechanism primarily is from water in streams crossing over the outcrop, or the recharge zone on this map, and then losing that water into the aquifer quite rapidly.

And then that's enhanced by faulting. We're at the edge of the Edwards Plateau, and the Balcones Fault Zone helps bury those rocks into the ground, so there are these fast pathways to several hundred feet underground. That water is recharged in the recharge zone and essentially moves south, and then mostly west to east, to discharge at major springs, Comal Springs in New Braunfels, and San Marcos Springs in San Marcos. Comal Springs is the largest spring system in Texas and one of the largest in the southwest United States, averaging historically something like 300 cubic feet per second discharge. And those springs have endangered and protected species.

As a result of lawsuits in the late '80s and early '90s, the legislature in Texas formed the Edwards Aquifer Authority to manage water levels to govern for minimum spring flows to protect these species. So, the aquifer is extensive, tens of millions of acre feet in terms of volume, but we're only managing the pressure, essentially the pressure enough to discharge water out of those springs, and it's a minimalist approach. Our management room is really based on pressure to have discharge.

The aquifer's been studied for six, seven decades. It has an extensive process of physics-based models. The problem is that we have very big limitations. For instance, the recharge model for the aquifer is based on empirical measurements of stream flow over the year. So, there is no precipitation infiltration model component because it's at specific locations due to specific spring flow. So, we have no way of projecting what aquifer recharge might be, even if we did have the climate information in the future. So, that's part of the issues of what we're trying to do and what we have to deal with at the authority.

So, why have we engaged in the process of artificial intelligence, machine learning modeling, and climate science? Well, assessment of sustainability is important because of development pressures, water use pressures. So, just because we have a cap on pumping doesn't mean that the population is not going to increase. We're also under pressure from adjacent aquifer systems that likely contribute to ours but are not incorporated in our management structure.

And the assessment of future climates is really important for renewal of our incidental take permit, which requires us to have a climate description and influence of climate on our permit period, which we expect to be about 30 years. And that renews ideally in 2028.

And then we also wanted to develop to see if we can use these AI models to offer alternatives to our process-based or physics-based typical groundwater models. So, can we short-circuit it, get precipitation, and directly get spring flow and water levels in terms of that process?

We're a small agency. We have limited or zero climate expertise. And we have fairly modest resources based on that, but we do have some budget for research. But collaboration is a big part of our strategic plan goals. So, we're a little fish down there, but for our small bucket, we're the big fish in terms of technical expertise. So, all the adjacent groundwater management districts, which might have a general manager and administrative person, we're the only ones that have a technical staff. So, they rely on us to get that information.

A good example is, email today was from Hill Country Alliance, and they want geological expertise to describe, how might these tower footings for a power line to Fort Stockton - proposed to put into power a data center - how might those footings impact the karst structure spring flow in the system? So, that's our supporting network that we have involved.

So, we have several applications to date. Most of our success has been with tree-based ensembles. We used to emulate these physics-based groundwater models. They've been extended to do some other things. We didn't have this recharge model. We have historical recharges calculated by the U.S. Geological Survey. So, what we were able to do is have essentially an ensemble model approach to mimic that recharge, tie it to precipitation and temperature, link that with future model projections so we can get a recharge model in the future. Now that recharge is used to input into our groundwater physics-based model to produce spring flows in 30, 35 years into the future.

Other stuff we've done is short-term predictions for modeling. Can we improve our management strategy by anticipating low water levels and drought conditions so that we can have some voluntary suspensions of pumping or maybe incentivize suspensions of pumping sooner than our mandated restrictions that we have now?

The other thing is, can we evaluate conservation measures? One of the things that Dr. Chakraborty showed yesterday is, how do you know if your managed reductions have a real impact? And so, using a counterfactual approach, we were able to show that, hey, with and without those measures makes a big difference in terms of long-term expectations for water levels in the spring flow in the system.

We've done some neural network modeling as well, but those have been less successful primarily because we have significant data limitations. So, imagine trying to estimate stream flow from precipitation events where 70 percent of your time, those streams are dry. They're very ephemeral. So, the data processing is very difficult, and those models are less successful in that environment.

We've used some generative AI, but in a very limited basis. So, can we identify sensitivity analyses? Can we help summarize a process, some of these process-based modeling results? Can we use that to help formulate other scenarios? That's the limit of what we've done so far with that. Like I said, most of our information is machine-learning oriented. And that's kind of the issue: can we convince our board and then convince Fish and Wildlife that what we're producing is adequate?

And so, what I would say right now is that we've learned a lot from the application of these AI models, but none of them are really used directly in terms of helping us make our technical case for spring flows and long-term system behavior, because they still are inadequate in that sense.

Finally, the critical needs really are, we need access to climate expertise. We need access to computing resources. So, one of the issues originally with downscaling the climate data for a region, we had to purchase extra computer nodes at University of Oklahoma to process that data in the timeframe that we needed to have it processed. So, that's another disconnect here, is we need information not on some unconstrained timeframe, but we had about 18 months to solve that recharge model problem.

Expertise in refining the models with UTSA and others have been good, but we can't do any of that without the system guidance and expertise. AI models produce some interesting things, but they may not be realistic at all. And so, there are a lot of acknowledgement that has to go on there.

The challenges are really time and cost. Who's going to pay for providing these partnerships, pulling them together, paying for people's time? The other thing is the time aspect. We worked on these collaborations since 2021 to be able to implement them. So, we have a five-year process to get that. We have two years' worth of modeling and effort. It's not an overnight thing to develop these at all. And so, one of the issues is, who's going to take charge of organizing the collaboration. That takes time and effort in mind. You have to have someone separately defined for that.

And then the last thing is, well, can you have external funding and survival of these organizations that you depend on for these resources? If CASC loses all their money and those staff go away, well, where's my climate support going to come from after I've established these collaborations over a long period of time? So, that's particularly critical. As Tom mentioned, the government has all this data that it's collecting. Well, what if those processes for collecting that large-scale data go away? You need to have those infrastructure commitments in the background to do that.

So, that's a summary of what we're doing and why we're there. I appreciate your time.

(Applause)

DR. WIRZ: Hello, everyone. My name is Chris Wirz. I am an assistant professor in Risk and Science Communication at the University of Illinois, Urbana-Champaign. And I am honored to be here. So, thank you for the invite. I appreciate all of the thought that's gone into planning this and the efforts and logistics and the great conversations of a full notebook, full head of a lot of ideas. And I'm going to try to represent and call back to some things we've heard so far around partnerships.

Before I get into that, I want to talk a little bit about my background. So, I study communication, which I have appreciated that pretty much every Panel has touched on in some way, so that feels really good. And I've also studied a lot around trustworthiness along with Anne and others, really in the dimension space of AI and what that means for a user. And so, I often think about partnerships in terms of not just the people using it and developing it, but also across different industries, like we've heard a lot of today, but also interdisciplinary partnerships. What does it mean to really collaborate across pretty different fields, and what that looks like.

And so, I'm going to start with an idea. I think of I got this as a fortune cookie after working on a convergence project, which is, if you want to go fast, go alone, and if you want to go far, go together. I think when we have those partnerships and we get in these spaces, it can be really hard and it can be really sticky and it can be frustrating. And it could be like, I could just have done this faster by myself. We probably all felt that.

We're like, why did I do it with this person? Why did I bring them along in the early days? But remembering that that is an investment, that these partnerships and relationships are big investments that we're making because they let us go further than people could on their own.

Challenging those different perspectives and those sorts of, whether they be cross-sector or cross-discipline or even just different worldviews or different ways of knowing, can really get us to a farther end goal. So, those are a few things that I'm going to frame some of my thoughts around today.

My work is focused on a lot of the weather scale. So, Anne and I and others have worked with National Weather Service forecasters, thinking about, how do those folks use AI? Are they ready for it as an industry? What are the factors that are important for them to assess whether or not to use a tool, or whether they trust these different things, and what it might mean for their decision-making?

We've also worked with Department of Transportation officials. So, thinking about how those folks might use different AI-generated products in their decision-making and what the stakes are there, as well as some coastal folks who make a lot of different decisions around natural resource management and supply chains.

A few themes that you see across those sorts of different groups of people as they start to navigate the use of AI and important partnerships come down to trust, which we've talked about before. I'm going to sneak in a paper as a framework that Anne and I worked a lot on. It had a lot of deep discussions about, what do we mean by trustworthiness, and how can we make this useful.

And we reviewed a lot of different social science literature because there is a lot of deep theoretical understanding about how people trust other people or other organizations. There is also a whole rich literature that exists around, how do people trust automation. So, things like autopilots and milking machines and things that have gone back to the '80s that people have been thinking about those relationships with.

And then there is a field where Anne and I come from, too, which is thinking about risk, where we often think about trust in the people who manage or are dealing with the risk. So, do you trust your government officials? Do you trust scientists? Do you trust developers of things around a risk or hazard? And when you look across all this literature, a few big things that we really understand about trust as we start to navigate these partnerships.

The first one is that it is a relational concept. Trust is something that's between me and someone else, or you, an individual, and another thing or person. So, I have to decide, do I trust Anne for something? Do I trust Paul? And it's going to exist between the truster and trustee. And on that other point, though, that means trustworthiness is going to be something that's also relational. And it's going to be the evaluative part of, how much do I trust Paul? To what extent and in what context?

And so, the implicit point of this is that there is no trustworthiness score for Paul. You don't look under his shoe, and it's a 99 out of 100 trustworthiness generally, because it's really in the eye of the beholder. We call that perceptual. And it's also going to depend on context. So, trust for what? Do I trust you to give great remarks? Yes, awesome. Like, maybe do I trust you to be a babysitter for me? It's like, oh, do you even have kids? What does it look like? I might have different factors for that. And so, really understand those two points. It's relational in the eye of the beholder and then context dependent.

We spend a lot of time as scientists thinking about the developmental context. Is it developed? Does it have good data? Do we have data availability? Are we using the cutting-edge best practices to develop this?

There are other contexts that are really important, especially in these partnerships, the use context. How is this being deployed? Is it making a decision? Is it replacing a person? Is it augmenting their decision-making? What type of decision is even being made? Do we think that that's a trustworthy thing to have done? And so, really spending time and understanding that use context is really key for different types of partnerships and relationships, as well as, how is it governed. Who is watching over this? What does this look like over time? Do we have checks and balances and standards for these things?

I think there are some really good examples of people in this room that have thought about what it means to really, in their partnerships and their work, respect that use context. So, just talking with Mariela about work of understanding, okay, I need to communicate that these are predictions. Because I know if I give my user something and they see somewhere identified maybe as a light post or isn't a light post, they're not going to trust the rest of it. So, you're managing those relationships and understanding the importance of what we call like a signal event. If something goes wrong, it could all break down.

Other examples, too, are thinking about Katherine, like actionable beats perfect. So, you have a perfect model. If it comes too late, it's not helpful. It's not useful for a partner. So, having something that's in the right time and hits the right need is really key. Emma talked about this, too, with her decision support platforms with John Deere. If you do research in this platform, it's going to be, we'll incentivize that because people have limited bandwidth. They can't go to many different tools and different websites that aren't available all the time. They make decisions with one tool. If we can figure out ways to have collaborations where research is put into the tool they're already looking at, that's really powerful. It has high-impact.

And thinking, too, about these bees and elephants discussion. So, do we have your generic elephant versus these specific bees? Asking yourself, who has the time and the money and the power to get themselves a bee? And who is forced to use an elephant? Is it, they're using an elephant because they want to or because it's all they have time and access for? Which gets to the point when we think about partnerships and relationships, capabilities and capacities are different things. You might even have the capabilities to use models, but you're constrained in the time that you have. Or you don't have the resources to do it, or there are operational mandates that dictate what you're doing. And so, making sure we're not just trying to increase the capabilities of people, knowing they have real constraints in their decision-making, what those capacities are. We're making sure that our science navigates both of those things in tandem.

And so, thinking of stepping back from those user relationships and great examples we've seen here, there are also those interdisciplinary collaborations that we've talked a little bit about here. We can work across sectors that could be helpful. And examples from my work, and just also experiences, these are things we often say we celebrate and we really reward and we want to do it, but I think a lot of us know that we don't always reward these things in reality. We make a super interdisciplinary student, but where do they get a job? They still have to find a discipline in some way. And how do we really point to maybe the longer timelines that it may take to produce papers and to get tenure or whatever you may have for your metrics in your job?

I think if we can get over that activation barrier, as an example, again, another project - I keep pointing to Anne, we're on a lot of these same projects together - we're like, we've used social science methods to try to speed up in some ways and scale the training, the hand labeling of data so it increases both the reproducibility and replicability of the data set in ways that are also going to be helpful for building trust.

So, if I can say, hey, we make decisions about what a snowy road or a wet road looks like, it's 10 pages, even if you don't agree with us, you know exactly how we made these decisions. Our whole team is statistically reliable in how we made those decisions. You can check it out here, but then we have ten of us who can all label these roads for training our model. It's something that can actually maybe add a little bit of a boost if we do it better.

I think too, like we go to cross-sector partnerships, if we can find those mutually beneficial goals, those opportunities that both Tom and Paul talked about, where we need each other and we can rely on each other, and we can cover some of our gaps, I think will be essential, especially as you feel some resource constraint. Where are our goals aligned? How can we negotiate those upfront? Make sure we go further together.

I'll leave it there.

(Applause)

DR. LUERS: Wonderful, thanks so much. Great diversity of perspectives on this. I am going to open it up pretty soon, so do start lining up if there are questions, or add them in, because it would be great to get thoughts from everybody here and online.

I'm going to kick us off by reflecting on what each of you has shared. You talked about partnerships, you talked about how you use AI, and how you engage with AI. But I want to push you a little bit further because obviously, the work of science to climate action inherently has partnerships in it. Partnerships, boundary organizations, boundary people who work in these different spaces. You need expertise, you need funding. I'm thinking about the challenges and the needs that have been highlighted here.

I really want to push for each of you to say what is different about these partnerships compared to a decade ago. And maybe there are not that much or maybe there are a lot. What is different? What is a challenge, and what is an opportunity? Each one of those could be, I'm assuming in some ways a dissertation, I'm not asking for that. But just really trying to get some ideas out there to say, is there anything really different about this moment, or is it incremental, these big changes that are happening now? Or is there really something fundamentally different?

I will start with you, Tom.

DR. HAMILL: Okay, I kind of talked to this in my remarks. How are we different from ten years ago in numerical weather and climate prediction? It was primarily at least the generation of the raw data.

DR. LUERS: Just to clarify, it's not how we're different in more modeling, it's how we're different in partnerships, specifically. Just wanted to clarify.

DR. HAMILL: Okay, good. So, basically, where I was going with that is with climate modeling, being so intensive of human and computational resources, that data was largely generated in the public sector, which had the resources. So, now we're switching to something where it is not necessarily the data being generated in the public sector.

Now, there still is, per the other speakers, a cascade of relationships before it gets to actual stakeholders that are using and making these decisions. And I don't suppose that that may have changed that radically in the last ten years, but certainly on the production of the underlying data, I think that that has changed pretty fundamentally.

DR. BERTETTI: Well, from a partnership standpoint, I think the primary difference is, in the past, if we needed some technical support and a collaboration and a contractor, we could find a single source or the best capability. Whereas in the current environment, we need to bring in multiple entities. Typically, there is not the do-all experience level that's available.

And so, one of the things is that the AI component offers a tremendous amount of opportunity potentially for us to explore a number of things. For instance, we're trying to utilize land management techniques to enhance the sustainability and water quality, and the content of the aquifer system. And we're trying to collect data on really small amounts, but how can we scale that up? Now, how do we do that over a large range? Well, there are some potential uses for these AI modeling approaches that might allow us to do that more effectively than if we used a process-based surface water hydraulic model.

Again, though, the challenges from my perspective are always, do we have enough money and time to bring those together? And are the organizations that are expert going to be interested in our little problem? So, if you have academics working very hard on their research and their component, it's hard to publish your results for little Edwards Aquifer system, but that's exactly what we need. We're not world changers. We need applications to that sense. And although there may be some larger applications, it's really important.

So, I think that's the main difference from what I see.

DR. LUERS: Chris?

DR. WIRZ: Yeah, this relates to something Paul was saying, but I think the speed of development and the attention that some of these things got could be distracting. And I think this is something Libby Barnes has said, was how do we find those big questions that we can do from a science perspective, without trying to compete in the industry space? And I think that we saw that on big teams, you've spun up, you've got all these interdisciplinary collaborations, you have the relationships, and then, okay, the science has changed. Now this type of thing is there, and we're reacting, trying to follow industry, but we're not thinking, what are those big science questions that we can ask really well, and that's really grounded in our disciplines and our perspective.

I think it's easier to lose sight of what the really big science questions are and to react to what's happening at a more, some of these industry or flashier levels.

I think also from a different type of partnership is the automation user pairing team, what that looks like. We've had a lot of these questions for ages, like since the '70s: what is the role of the human and what is the role of the machine? But we navigate those now at a much faster pace, and there is institutional pressure around; we should be using this, we should be using AI, or we're outdated. And so, you have a little bit more of like, not fear management, but like different expectations and external pressures that you then have to navigate in your science and in your partnership with other people, because they're feeling pressured to use this or not, and to move quickly and be agile. It's a lot more management around some of those spaces.

Do you want us to do the next two as well? You had?

DR. LUERS: Yeah, if you had it.

DR. WIRZ: I think, as far as like a challenge, and we talk about this a lot, the context matters. Small-scale context is really important. Each city, each person, each need is really different. But also, we can't have a social scientist, a policy expert, a climate modeler, an urban planner, all in every community, every city. So, we do have to scale at some point. So, how do we really represent that context and things that really matter, but also start to scale the bigger level? I think that's a big challenge.

I think the opportunities are there, if we can look at how we're developing knowledge and start to make some of these more generalizable frameworks, especially in social science, where maybe we're not as much entrenched in these like really actionable spaces, there is a lot of room to make new science, to figure out how can we scale those things and answer some of those, or address some of those paradoxes.

DR. LUERS: Great, thanks. Let me throw out some reflections on what I've seen in the context of thinking about partnerships and see if any of those resonate. And if not, maybe it would spark other things that do.

I've heard a little bit of this, but what I think is different about partnerships now in this space is, for one thing, the issue of trust is different. Just the dynamics of, there is definitely an issue of, do people trust. There has been a lot of research on that, and this was an issue when I was a PhD student of, do they trust the projections? How do you get them to trust the projections and so forth? There is a difference in the trust, I think, now because of AI. Really, even if AI was used in the other ones in a different way, but today there is a difference.

I also think that the limitations, the skills ability, and the diversity of where you can connect in the context of AI partnerships are limited by skills in a way that weren't before. Because there was no assumption that you would necessarily need the skills in terms of some of the information. It was not machine-based, that collaboration. So, I find that there is a skilling - you had raised capacity and capabilities, and that access. I think that's related to it.

But then a third piece, maybe I would say that as an opportunity, is that in some ways, when you do have those others overlapped, you have a broader group of partners that you can engage in.

And so, those were some of the things that I would see as different from my experience in terms of thinking about the opportunities and challenges in this space now, really thinking about when you're partnering around AI and climate action.

I just wondered, does any of that resonate with any of you, or does it spark any other thoughts about other types of opportunities and challenges you might be facing in your work?

Okay, great. Tom?

DR. HAMILL: Yeah, your question about a broader group of partners being possible was one that resonated with me. I think one of the hardest things when you're working on a particular problem is finding others who might have relevant expertise or interest in this.

I was thinking as Paul was talking, okay, he's working in a basin that's got some fairly unique geology. But one of the cool things about AI is you could say, here's about the geology of my basin. Tell me other locations that have similar geologies. Oh, there is somebody that could be a potential partner that I can line up in a basin that's got similar geology. AI, the chatbots, and such really facilitate us finding new partners that we might have really struggled to find five years ago.

DR. LUERS: Yeah, I hadn't thought about it in that way. It's a different way, but that's really interesting. Others, Chris, or Paul?

While they're thinking about this and answering it, please go to the microphones if you have questions and start adding them online.

DR. WIRZ: Yeah, so I think this might be a data scarcity issue because my record doesn't go back super far. I do feel like the space that trust fills in all of our brains seems bigger than it was before. So, we give a lot of - and some of this is media rhetoric. We're really grabbing onto the crisis of trust and all these issues and what this looks like. Is it really confidence? Is it trust? Is it institutional problems? Is it in our science specifically?

If you look at some of these arguments, too, around everyone is against science now, have an argument. Let's have a vaccine argument. People are going to throw science from each side. It may not be science we think is rigorous, but a lot of these things, people are still engaging some science argument for us. So, it's got a role in there. People are still using it. But the way that we think about it and the way that it comes up pretty much every scientific meeting, it does feel new. And so, it's a little bit of managing our own I think expectations, understandings, and concerns around trust. And then the specific trust in the context of our science, I think, is a little bit different.

And also, to your point about re-skilling and the opportunities. I think it does change the game a bit, where someone, maybe their whole science career is based on knowing a coding language really well and being really strong at that. Then maybe it is not the most useful skill. Generally, it's not going to be enough to make a career, but still has this fundamental knowledge is really useful versus someone who maybe had a lot of really good interesting ideas but not the technical language, who can ask these questions.

But there are also issues with each of those things. So, maybe you are making some of those leaps. You can use a tool. You can ask Gemini to make all your code, but are you going to recognize when things start to fall apart?

And so, it's like, yes, new leaps, but there are also new slipping points. How to manage that, I don't think we have a good idea for.

DR. BERTETTI: Yeah, my perspective is a little bit different. As a scientist, you're working on particular topics and collaborating over a long period of time with like-minded or like-focused individuals or groups, or entities. In this sense, we're trying to expand into areas that we don't know very much about, but we know might really help. It's like the discussion with Monica earlier, well, we're not reading peer-reviewed science on the best climate model approach and integration with AI. So, it's one of those things where it's easy to say, well, just go out and look up some people and collaborate, but it's a time issue.

It's great to be involved in a consortium where you can discuss your research opportunities over a long period of time. But typically, what I need is somebody we can discuss with and identify a set of projects and make some progress toward it in a fashion that's responsive for our budget and our timeframe and to answer the questions that we need to do. So, there is always this overarching thing about what's happening on this side. And there are a lot of other pieces that we're trying to juggle with, six or seven people on the science staff.

So, I think from my standpoint, it'd be great to have some way of establishing being able to integrate into established collaborative environments where it's accepting things like, here's what we have to say, what can people offer? Can we get some direction on how to go and what to do with it?

DR. LUERS: Great. I think there is a question up here.

DR. MORRISON: Monica Morrison from NCAR. This question is maybe a little bit more explicitly directed towards Chris, but I welcome responses from the entire Panel. So, seems like a big part of developing trust in the context of partnerships or collaborations more generally is to understand the interests and values that are brought into a context and to develop some shared understanding of what those are, and maybe navigate any differences that exist.

These things are often incredibly implicit. So, I'm curious just because I think AI might actually exacerbate the implicitness or the lack of transparency around values, especially in thinking about the way in which we might actually bring those values into the tools that we're introducing into these contexts.

So, my question is just if you have any thoughts about how to bring these things to make them more explicit, how to have conversations around these, the importance of these things in the context of developing trust, and how we go about navigating the diversity of values that people bring to collaborative contexts and partnerships.

DR. WIRZ: Oh boy, yeah.

DR. MORRISON: Yeah, that was a lot, sorry.

DR. WIRZ: No, it's good, and it's important stuff to think about. How to make them more explicit is a challenge, but I think we also struggle - and I'm sure you agree with this - as scientists to bring in a really clear, articulate understanding of what our values, goals, and motivations are in all of these spaces. Because there are some times when you're bringing a model that you just really hope someone loves and really hopes that they know that you put a lot of time into it, and it's like really important, which is an important relationship and way to manage, but are you open to the idea that maybe it's not useful for this user? And what do you do if it's not? Are we okay walking away?

Or if you really want to help someone, are you okay changing the science that you would do? Are you thinking or doing your science in a different way? And that's a big question. And so, thinking if these are questions that are hard to grapple with as scientists, and this may not be something you're ready for, maybe the user conversation is not one you're ready for, which is a hard point to think. So, making sure we're aware, how are we communicating and why, and with what goals?

If we can walk through some of those things pretty explicitly before going into those relationships, I think it can help a lot. Even if some of the answers is I'm not really sure yet, that you can put that forward, and someone knows, I'm really hoping to see if this specific model has an application, and here's roughly where the bounds are, as opposed to you also don't want to be too open and like, I'm going to solve all your problems with the tool and someone gets their hopes up that maybe this really tough decision could be removed from them by a product that you probably can't actually make.

Having those things be as explicit conversations or even putting some of those things in an initial template for a user discussion, how we start to have those relationships, I think, could be really powerful. And maybe that's the way as an initial conversation would be perfect. But a lot of work to do there.

DR. LUERS: Feedback on that? So, we have one question online.

MS. MELVIN: So, this one is also geared towards Chris. Risk communication is a complex topic. Should we train science media journalists on risk communication, focusing on detailed aspects of data gaps, embellishments, transparency of models, and auditability of models?

DR. WIRZ: Well, science journalism is - back when I was an undergrad, we were talking about how there was a handful of science journalists left, and that's like ten years ago. So, thinking about, what is a science journalist now? It's any journalist who's going to have to cover more and more things for less and less pay. And it's not a particularly great job.

And so, thinking about, how do you support and empower freelance writing to be a good job, I think is a bigger first question because you can't train an industry that doesn't exist. And so, how do we support journalism, how do we support good quality writing, I think, is a more important question. And that often comes through organizations who tend to do a lot of writing the press briefs and a lot of the stories that are getting out there and understanding our shifting media system.

You're also not going to be rewarded for doing it well. You're going to be rewarded for sensationalism because we live in a click-based system, and you can put the vegetables out for everyone to eat every time, but people are going to grab the donuts at the end of the table because that's the headlines we want to read.

And so, it's like, partially, yes, it'd be great if we had a system in which we had an empowered science journalism industry that was well supported. So, thinking about how we can play a better role in supporting media is probably the first question, like the first support, because it is really complicated. And then thinking about what skills are needed for that is a little bit tangential.

So, picking out the wallpaper for a house that isn't going to be built system, unfortunately. I think what's more important is how we get that house built, unfortunately.

DR. BOOMER: Kathy Boomer with the Foundation for Food and Agriculture Research. I still appreciate the diversity of Panelists in this session, including yours, Amy, which takes me to a question, maybe you could also respond to.

So, two topics that have come up in this workshop are trust and uncertainty, but we haven't really talked about how they're connected. And I wondered if you could share thoughts also as to how this might shape your views on partnership needs and partnership opportunities.

DR. LUERS: Can I just clarify the question a bit? So, coming back to my point of what's different now. So, you're saying that uncertainty issues are of a different nature now, with an increasing focus of AI. And trust, we already have that conversation. And in that context, what does that mean? Specifically around that point, is that what - am I accurate in -

DR. BOOMER: So, we're talking about AI for climate impact and action together, and what are the barriers? We've identified trust and uncertainty. But how do those two together influence your thoughts about where we need to go together in the future? How can we work together more effectively?

DR. LUERS: Who wants it? Tom, do you want to respond to that first?

DR. HAMILL: I think others are more expert on this subject.

DR. LUERS: Okay, Paul.

DR. BERTETTI: I do have an opinion about communicating that risk and uncertainty. I'm not sure that it's a lot different from communicating the risk and uncertainty from another number of models. So, in my past life I worked on the Yucca Mountain Project, and so that was a risk-based regulatory approach. It's probabilistic. And so, you've got to make sure that you clearly communicate to the stakeholders that are involved, what the model is doing and why you think it's doing it.

A big part of that trust is your relationship with those stakeholders and those decision-makers at the time, they have to understand what you're presenting, you have to do it in a way that makes it clear, and you're able to answer the questions that they have.

I know that there are a lot of uncertainty associated with the black box component of AI, but if you're using those and you're going to present that as a way to make a management decision, it's got to have the same characterization and robustness that you would present any set of models or any set of assumptions to so that people can do that.

And so, without speaking directly to the climate science component, I think a good example is this recharge model effort. It was very successful. We were able to use this AI model to do something that had been hanging over our heads for years and hadn't been done previously. Well, who believes that that's going to work? Well, we provided how the model performed, what the uncertainties were, how that looked with respect to each basin. And so, I think we're able to communicate, hey, it has uncertainties, it's performing what we can, and it's a way to do that.

So, that piece to me is not much different than any of our other approaches.

DR. LUERS: So, since you've posed it to me as well, I'll take the opportunity to share some thoughts. It's probably not the answer that you expected, but something that I think a lot about, especially working at Microsoft, we talk a lot about the frontier firm where there are going to be a lot of agents working under teams as individual workers. And that's already emerging.

And so, what I think a lot about is, how do you deal with trust and uncertainty in a world where you have frontier firms and frontier NGOs, in other words, teams that are agents? And these are individual entities that work as members of your team. They're part of the discussion. They influence how people think and the decisions that are made.

And so, the question, if there had been time that I was going to ask here, was, in five years, what would be different? We said, what is different now? And I think that that component of how - at least I grew up in my academic world of thinking about how decisions are made, how do we partner and do bridging organizations? My head blows up when I try to think about, well, what does this mean for all of those things that we learned about how to do that.

I think it's scary, but I do think there is a part of me, if we think about it and start doing it now, there is an opportunity. This field is moving so quickly that I think it's going to - saying at lunch that policy is not catching up with the technology. I'm not sure that our integrated science community is catching up with the systems change in society. And I think, how do we get the science community to get ahead of that, in terms of what are the trust and uncertainty issues in that new frontier world?

Anyway, more than maybe you - I know that there are -

DR. WIRZ: Can I add a little bit to that?

DR. LUERS: Yeah, or do we have time? Okay, go ahead.

DR. WIRZ: Because I think one more part of the equation - I love what you're saying - is the vulnerability. So, you think about trust as, in the presence of uncertainty, the willingness to be vulnerable to another group or entity. And so, if we can figure out how to communicate what those uncertainties are, then we don't need trust if there are no uncertainties. If I'm in a partnership, what are the vulnerabilities for you? And we as scientists, channeling my Bernie Brown here, we have a lot of armor against vulnerability. We caveat everything. We put a lot of big, dense words around being wrong.

The idea is, if I'm working with a partner, it's like, where could this go wrong for you if we don't come through on this? If you're an emergency manager, what's the worst-case scenario for you? We didn't even raise the potential issue that there was going to be a flood. So, we need to have not been caught behind.

So, how do I know what you're vulnerable to and what I'm vulnerable to, and how we can start to address those things to mitigate where those problems could be, is really strong.

DR. LUERS: Thanks. I think we do have time for one more question.

DR. WOOTEN: Adrienne Wooten, University of Oklahoma. One of the, I wanted to piggyback up from what Paul said to ask this question because we've run into this too, is boundary organizations, be it the CASCs, or the RESAs, or Climate Hubs or any number of others, is the issue of capacity much the same as what Paul has talked about with some of the smaller water management agencies. There are only so many of us to go around to help all of these different stakeholders.

At the same time, we've talked about expertise not being replaced by AI. And so, my question, a two-parter actually, but for time, if you only want to answer one, I understand. Could AI be used to help folks find those partners that they need effectively in the sea of the huge number of people of different expertise that they could go to? And would that be useful for different stakeholder groups, but also what might be the pitfalls of doing that, like maybe finding absolutely the wrong person and thinking that's great, and now you suddenly have to start all over.

DR. BERTETTI: I think that's, as Tom mentioned, an intriguing aspect of using that approach to mine the database and get the right entities. There is always the uncertainty where the partner with whom you're working is not going to be able to deliver. It's one of the risks that we have to take every time we have a contract.

And so, universities, a graduate student may decide to leave in the middle of the project or something. And so, to me, more uncertain is identifying someone who's not properly vetted. It's not just, can't do the work, but they're truly not qualified. So, you still have to somehow, like we've talked about with the data uncertainty and the models themselves, how do you ensure that you have the Angi's list of technical experts available?

DR. LUERS: I am going to go to the question after. I guess we have more time than I expected.

DR. RAO: I just wanted to go back to Paul's response earlier about it shouldn't be anything different than not for AI-related partnerships, and like for risk communication. And I think now people have formed their own opinion about AI already. And so, I would say that requires a different way of thinking about it because they already have their own opinion about AI, based on things in the news, and then different instances have happened in the past. So, I think there are some unique things we should consider to try to overcome some people's formed opinion already about AI as a tool. Clearly, we want it to be neutral for different tools.

DR. BERTETTI: I think communicating risk and uncertainty is consistent with how you would have done it in the past. But I agree, there is always this preconception based on what people have read and what they understand. And so, I'll just go back to the nuclear power and nuclear waste issue. You have to overcome this inherent fear that no matter what happens, it's going to be incredibly dangerous and not achievable.

And so, in the AI sense, it's a similar thing. You've got to demonstrate over and over that you have a handle on the uncertainty, that you can communicate that risk. And then you're doing the things that people want to see to build their trust. And so, whether it's any contentious issue like that that goes up has all these preconceptions.

I guess what's different is there is a whole, really responsive social media aspect that maybe isn't out there before. And so, maybe that whack-a-mole is infinite, and you can't really deal with that. But consistency is key in that delivery of that message.

DR. LUERS: Tom, yeah, great, thanks.

DR. HAMILL: Yeah, it's interesting. Paul is very concerned about, can partners live up to their side of the bargain? In some sense, when I interact with a chatbot like Gemini, that partner doesn't have that problem. So, I have found in my recent science explorations that Gemini is becoming the smart colleague that, in a small company, I don't really have to talk through some of these issues.

Now, you have to accept the limitations of these. They're no smarter than all the corpus of data that's gone into training them. But that's a pretty substantial corpus of data. So, recognizing those limitations, I've found that Gemini is a new partner that I didn't quite appreciate a few months ago, to being as vital to my research as it is.

DR. LUERS: We'll go with your question just to bring -

DR. NAKALEMBE: Thank you. Catherine Nakalembe, University of Maryland. I work with a lot of partners from university to farmers to extension aid and to people in the UN, the World Bank, across NASA programs, et cetera. But one of the things that ends up happening is that in the name of partnerships, you can have a lot of participants in an initiative. But what ends up happening is everybody wants to use their tool or their method. And it's like this out-of-this-world competition.

So, imagine we are working in the education space, and you go to a grade school, and you like to teach kids how to work with a chatbot. Google wants to use Gemini. There is this idea of like, if there are so many tools, there is such a saturation that you almost need a decision tool for the tools. And my other thinking about tiny LLMs, it's almost like you need an LLM to figure out what LLM to use.

And so, I'm always grappling with this problem of how do you collaborate to build towards one goal, because I can choose any one of those tools, but as long as I'm clear about my mission and I'm continuing forward, I'm continuing forward for the partner who might be in ministry that just wants to figure out where we're going to put irrigation infrastructure or how much to pay for insurance.

But if Catherine's deep-learning workflow predicts better, and then Emma's other deep-learning workflow is a little bit better, it becomes very complicated to work with a decision maker. So, I was wondering whether it would become more useful to have our chatbot for dealing with chatbots for the decision maker.

Like the tools just end up in your face. How do you decide what's best, and how do we remove experimentation and learning and separate it from the decision process? If it's just experimental and you're iterating every minute, a decision maker is not going to go with that because the decision might change with the next F1 score.

So, I would like to hear from the Panel, what are your thoughts around this? Because everybody wants to collaborate, but everybody wants to bring their tool.

DR. WIRZ: I agree with you. I don't bring a tool to the table. So, I think of a different perspective, and you do want to almost have like the Google search, like this is an ad option available for folks, making it clear, like, we do have a vested interest in you taking this product on. And I think that that goes back to setting clear goals and expectations.

If you're hiding that under, that we're trying to get this into your whole school system, so you subscribe, and in five years we will increase the price, that's got to be there. But if you're willing to fund different research, how do we as people from the public sector make the agreements on what is right?

So, if you're going to do this, you have to have a 20-year price agreement, or you need to do these types of things to raise the bar for what it means to be a partner in that space. And if everyone's willing to meet that bar, then we can have options, and maybe we do need a meta system, but maybe take advantage of that opportunity space a little bit on our own, which we're not always good at doing as scientists.

DR. LUERS: I'll jump in on that. I'm thinking back, I don't know, was it 10 or 15 years ago, working on the National Climate Assessment. And one of the things I proposed then was that we should do an assessment of all the climate adaptation tools out there because I think it's not that different a situation. It was at the moment where everybody was doing a new climate adaptation tool, make this decision, make this decision. And the National Climate Assessment traditionally was to look at the scientific literature.

I was actually on the Committee for a sustained assessment, like how would we do assessments differently? And my focus was, well, there are all these tools, and they're all coming out, and they're not peer-reviewed. How do we assess them? And there are going to be so much more of those.

I don't think that what you highlight is necessarily a new problem, and maybe has a new character.

Maybe there is a bigger group of people who are using them, but I actually think the challenges between whether you use ChatGPT, Gemini, or Claude are not as dramatic as we had when you use these different climate adaptation pieces. I bet if you did a comparison, the implications would be - but I guess my point is, I think that's a problem overall. That would be my take on that.

I know we're running out of time. So, that's a topic for discussion maybe that we can have on the break. Thanks so much and give us a big thank you.

(Applause)

Agenda Item: Accelerating Climate Action with AI - A Path Forward

DR. DAGON: Okay, good afternoon, everyone. I think we'll go ahead and get started. Thank you so much for sticking with us for this last session of the day and of the workshop. My name is Katie Dagon. I'm a scientist at the National Center for Atmospheric Research, NSF-NCAR. My background is a climate scientist and a climate modeler.

I just want to thank everyone who's been involved with this workshop. I've learned a lot. I was also on the Workshop Planning Committee, so I'm honored that so many of our speakers that we were excited about were able to come and join us in person, online, and really bring a lot of excellent, diverse perspectives. And so, I also want to thank the Roundtable and the National Academies staff for their leadership in putting this event together.

Our task with this Panel is to really think about the path forward after our discussions at this workshop. So, we're interested in synthesizing and reflecting on the workshop while also bringing some new ideas that I don't think we've heard about yet. So, I think you'll get that sense from the Panelists' remarks.

What I'm going to do is just briefly introduce our three Panelists - we have two in-person and one online - and then let them share their introductory remarks. We'll go in the order on the screen here, and then open it up for discussion. As we've seen throughout this workshop, I think there is a lot of interest in questions and discussion topics, so we'll make sure to, again, incorporate our online folks and our in-person folks, and hopefully have lots of time for discussion.

I'm really excited to introduce our Panel today. Our first speaker will be Alexis Hoffman. She is the Senior Manager of Climate and Data Science at Jupiter Intelligence. I've known Alexis for a couple of years now, and I've always been struck by her passion and enthusiasm for these topics, both on a personal and a professional level. And I think she'll allude to that in her remarks. I'm really looking forward to hearing from her.

Our second Panelist will be Elizabeth Barnes, who is Professor of Computing and Data Science and Earth and Environment at Boston University. Libby's research and her group has really led the forefront of work in this area of AI for climate science and modeling, just an incredible breadth and depth of work here. And I'm always excited to read a new paper coming out from her group.

Our third Panelist is Marc Alessi, who's a Science Fellow in the Climate and Energy Program at the Union of Concerned Scientists. I think I first encountered Marc through his work with the 100-Hour Weather and Climate Livestream, which was an amazing effort to advocate for sustained funding in weather and climate research. And so, I've always appreciated his passion for advocacy and science.

So, with that, I will turn it over to Alexis to get us started.

DR. HOFFMAN: First, I want to say thank you for the invitation and the opportunity to sit on this Panel today and also to participate in this workshop. I've learned a ton, and I felt really lucky to hear other people articulate thoughts that I've had, but I hadn't had the chance to articulate them yet. So, this has been really helpful in putting words to my thoughts.

I'm going to keep most of my responses relatively brief because one, I'm between you guys and going home. Two, I want to keep things light and energetic, keep everybody awake at the end of the day. And also, I think, I was talking to Karen about this earlier, but I think a lot of what has been said already is incredible, and I agree with it so wholeheartedly, and so many of the points I've made so far. So, I don't know if there is too much that I can add besides like excitedly agreeing with a lot of the statements thus far, but I'm going to try. I'll try to add some novel aspects here.

Katie mentioned I am a senior manager of climate data science at Jupiter Intelligence. I called this my industry postdoc. Anyone who's talked to me the last 36 hours, I've been at Jupiter now for seven and a half years. So, I've been there for a while, and it's because we keep encountering new challenges and I keep learning new things.

I don't think Steve has talked a little about what Jupiter Intelligence is, so I'm going to give you a quick overview about what the company does, just in case you were curious, and if not, it's really brief, I promise.

Jupiter Intelligence does climate risk analytics. Typically, I'd flip to a slide that is developed by a marketing team, but this is supposed to be a short introduction, so I'm not going to do that. So, Jupiter just turns climate science into actionable data for organizations looking to strengthen their resilience and delivers it as a service. That's the TLDR, people who aren't familiar with what climate risk analytics are. And if you are curious, I'll give you one quick detail.

We have one primary flagship product, which is Climate Score Global. It's a global product that provides physical risk data globally for three unique scenarios for quintennial periods every five years up to 2100. Most of our customers are concerned with that 10 to 30-year time horizon, which is consistent with investment horizons. But we often encourage them to use data from the end of the century, just giving them an intuition about climate signals, just because not all of them have that going into the discussions.

This product is where the majority of our machine learning takes place, and the majority of the AI, also machine learning, at Jupiter is really just in a machine learning or regression framework. It's not necessarily generative, although we do have some generative AI models in-house that we use.

I also want to say something about that we've talked a lot about listening to your stakeholders and people actually using your data. Jupiter has sort of advanced with the market, I would say. When we first started, our primary product was very like hyper-local data for cities or locations or specific locations, geographies. And this product, which is our flagship product, started as like an Excel spreadsheet originally, and it was an afterthought. And then it turned out that people who had multiple assets around the world were really interested in trying to like prioritize where their climate risks were. And so, this product turned into our primary product.

Then people started asking the question, well, which physical hazards in which locations should we care about? And so, in that conversation, we're like, okay, well, maybe we need to add on a damage and loss model, like an economic impacts component to all of this. So, giving a physical risk of say two feet of flood at like five of your assets, which ones are going to experience the biggest losses? And maybe you should target your adaptation or your resilient strategies in those locations.

Then they were like, well, then what do we do with that? Is there going to be a positive return on that investment? And so, after this economic impact development, which is always ongoing, by the way, we started adding an adaptation module. So, if I invest $10 million into flood-proofing most of my assets along rivers, is that a good investment? And if it is a good investment, when am I likely to see that good investment down the line? And so, I just wanted to note that if we don't listen to the feedback of the market and the people actually using our data, we wouldn't be a company still. So, have to move with the tides.

I mentioned this very briefly, but Jupiter is doing some generative AI. We worked on a stochastic weather generator for some of our industrial partners. Callie McNicholas presented on some of that at AMS last year. It was entitled a *Multivariate Probabilistic Weather Generator Using A Denoising Diffusion Model*. So, you can always look that up on AMS or email me after, and I can send it to you. It was really good work.

The other thing that we're doing is building event-based systems for climate risk. I think Karen mentioned earlier using AI to expand the historical record to look at extremes that are consistent with the historical climate. And so, that's something that we're considering as well. That's pretty much all I got on Jupiter.

One of the other reasons I think I was asked to do this Panel is because I've dabbled in a lot of the aspects that have been discussed on Day 1. I have experience in climate modeling and crop modeling. I've fought with water use data on a personal basis. And I also am very interested in urban planning, and I'll get to that in one second.

But my background, I did a master's and a PhD with Chris Forest at Penn State, focused on using CESM, to estimate model sensitivities, actually focused on dust and mineral aerosols. If anyone's excited about dust, I always love talking about it. My PhD focused on using diagnostic crop models to understand the impact of different variables on crop yields. And in that process, I started talking the language of a data scientist, talking about data fusion and machine learning. And that's when I ended up at Jupiter after that.

I'm also on a Committee for the Open Environmental Information Services with AMS. And we talk a ton about this tension between public and private data. We try to publish on it. And so, I'm always interested in what the incentives are for people to collaborate with respect to data, from that aspect.

And then lastly, I have a lot of aspects that my academic journey touches on, a lot of things that people have talked about today. If anyone's talked to me in the last 36 hours, you probably know that I moved back to the area. I grew up in Pacific Palisades. I moved back to the area about four years ago. And so, I'm going to put on my civilian hat right now, not my Jupiter hat, and talk briefly about some of the stuff that I've encountered in the Palisades fire and the recovery efforts there.

So, before the fire, I will note that I participated in a group called Resilient Palisades. It's an environmental advocacy group. I focus mostly on water sustainability and on native plants, how to grease the wheels to install gray water systems, because there are so many different places to look. So, we focus a little on that. But I joined that group as the data person. Obviously, that's part of my job. And I was thinking that we could incentivize people in the Palisades to choose good water solutions or choose sustainably, if I could show material impacts and actually show changes in their water usage statistics.

Turns out, water usage statistics for Los Angeles are incredibly challenging to find. At first, I was really excited to maybe do this on a neighborhood-by-neighborhood level. That data doesn't exist. So, I scaled it up to the zip code level. And I was like, okay, maybe that exists. That also is not easily accessible, although it could technically exist. But it would be a lot of effort on the people making the data, since it's not publicly available.

Anyway, my colleague and I on that group, we spent a two-year rabbit hole trying to access water use data. And in the process, I became very familiar with some of the California rules and privacy concerns, and it was a whole thing. So, I'm very familiar with water use statistics as well. And then, obviously, last year, almost exactly a year ago, a lot of the work for Resilient Palisades pivoted to recovery efforts. And I tangentially engage in the Palisades Recovery Coalition, which is spearheaded by Maryam Zar and Rob Lempert. And Rob Lempert is actually associated with Rand, and a colleague from way back in my Penn State days.

Using the term climate action has made me think a lot about what the Palisades. And the urban planning conversation yesterday was, I was like, oh my gosh, the Palisades ultimately has this opportunity to start fresh.

It's really exciting, but so many people are focused emotionally on rebuilding quickly and trying to get back to some sense of normalcy that I feel like we're losing our chance to almost start fresh. Although I will note that Resilient Palisades is out there, literally pulling out invasive species, because we have a chance to start fresh.

I will pause on that, but I'm going to try to wear both hats, like my industry representative hat, but also my civilian and Palisadian hat, for anyone who wants to ask those questions.

DR. DAGON: Thank you, Alexis.

(Applause)

DR. DAGON: Okay, let's turn to Libby online.

DR. BARNES: I just have one slide here on closing thoughts. I've gotten to participate online here in most of this two-day meeting. It's been a lot of fun. And just to give my perspective, it is that of a climate dynamicist who really works, tries to work at the intersection of climate and AI. And I will add, also a professor, and I will allude to that at the end of my comments, that actually, what we should be teaching students is something that we're talking about every day right now, which has changed a lot even in the last two years, given the big jump in AI usage across the country.

I decided to distill my summary of thoughts from the last two days, as well as add a few things that maybe I haven't heard as much about. So, the first and foremost, actually, is, I think in my view, we need to move beyond discussions of this catch-all AI. It's like talking about regression or math to me at this point. There are so many different kinds.

And as we heard from David yesterday, the first speaker, the elephants, and the bees, which we've heard so many times now, but this idea that there are so many different uses. There are the simple, there are the complex, there are the generative. And I actually think at this point, the field as a whole of sciences in general, we can start getting more specific. And I think we need to get more specific, or we're going to start confusing ourselves as well as all of those listening in.

So, my number two is I see an immediate threat - I know we're supposed to be looking forward, but I can't help but talk about some of the challenges here - this immediate threat for this AI slop. The hype and the rapid pace of these untested products - and I don't mean that they didn't have a testing set. I mean untested in the sense that people are building new models every day, pushing out an RMSE on a testing set, and saying, now go use it.

I think, given that rapid pace, it means we may irreparably damage trust. And we've heard that quite a bit the last two days, which I've found really good that we're talking about it. This idea of fool me once, shame on you; fool me twice, shame on me.

I do get nervous that if the scientific community jumps too fast, even though we're always blamed for being too slow, if we jump too fast and we get it wrong, when it really, really matters, we are going to put ourselves much further back because of this break in trust. And so, this is something that I think quite a bit about.

Number three, and this is just a little anecdote, but I started working in explainable AI for climate science over five years ago. And at the time, I was thinking, oh, this is going to be the next big thing. Everybody's going to want to do explainability. And actually, yes, we talk about it as scientists, but I thought it would be a bigger deal. And instead, the LLM exploded, and people, for a while, stopped caring about why it's doing what it's doing? It's just pretty amazing that it can do what it does at all.

I think I'm starting to see, and we've heard it the last two days, this renewed emphasis on how important explainability and interpretability really are. This feeds back into number two, that understanding what we are doing and building that trust requires we understand how these tools are built as well as why they are there.

I also think we're going to start asking more questions related to how much explainability is enough for action. Not just do we need to explain it, but what is that level? Because, and I actually think this is where I have found a lot of use from people in philosophy of science, folks in the social sciences, really thinking about what it means to be explainable to one person is different from another, and is different to the people making decisions. And this is something we're going to keep talking about.

And then finally, I guess I did the opposite. I didn't end on a positive note, but the thing that most concerns me right now is how society will obtain and process climate information by LLMs. That actually, we saw some very nice slides earlier, I believe, I think it was today, about this process of who's consolidating, who's summarizing the data for the next step, who could summarize that data for the next step, et cetera.

In the past, I've largely, largely trusted that middle person. Maybe it was, maybe it was the scientific writer, which Chris was talking about earlier, or some other scientist who loves the communication aspect. And now, it's these LLMs. Their ability and the way they distill things into these superficial blurbs, which we heard about yesterday, and disinformation. So, one is this, just it's wrong, and the other, it's been simplified too far. I think this is going to end up to the climate community, personally, to actually be our greatest challenge, not collaboration, not the technical aspects, but actually how people are adjusting the information that we're putting out in the first place.

And just to end here, I mentioned I'm a professor, I teach many graduate students, teaching undergraduate students, and the way they interact with these LLMs is very, very different than the way my colleagues interact. And I think moving forward, there is a huge part of the faculty wondering, what should we be teaching these students? How should we be teaching these students? And are we preparing them to be those technical experts that I think others were talking about in the last Panel?

And right now, as faculty, it's unclear what we're supposed to be doing to prepare them for this next step, this next world they're about to enter. So, I'll end there, thank you.

DR. DAGON: Thanks Libby.

(Applause)

DR. DAGON: Okay, Marc, we welcome your remarks.

DR. ALESSI: First of all, I just want to thank the Planning Committee for such an excellent workshop. I thought it was really interesting to have so many different perspectives and professions all talking about the same topic of climate science and AI.

What I'm going to go through now is my takeaway or my takeaways from the last two days. And also, I just want to introduce some other things that I'm thinking about that will be very important for climate science and AI. And I am using a catch-all AI that Libby just said not to use.

So, for me, something that was talked about over the last two days was this black box that is AI. We don't really know what's going on under the hood. Mariela mentioned it on her Panel. We talked about explainable AI from Adam, Debaditya, and Karen. And I thought this was really important because for me, what I think is the next step for climate science and AI is building trust. But at the same time, I think the process of building trust, understanding our models better, and using them better, also allows us to return to the fun of being an atmospheric scientist.

Like me, if you've been obsessed with the weather since you were five years old, you're really interested in how the atmosphere and ocean system actually works. AI is very cool because it can make a prediction based on their relationships. Okay, whatever. But explainable AI is really interesting because it can reveal how the AI model made those relationships or made that prediction given those relationships. And I think the next step is proving the causality and dynamical mechanisms in these identified relationships.

xAI is good; it reveals those relationships, but it's still statistical. It's not really like we know for sure that this causes this process. And so, causality, also something mentioned by David, Debaditya, Francesca, and Karen over the last two days. And I'm actually, as an early career scientist and someone in not a permanent job, I'm going to plug my own research at the bottom of the screen here, which does actually do something like this, where we use an AI model to basically make a prediction on a relationship it learned. We used xAI to be like, okay, where is the AI model looking to make this relationship?

And then, for me, I am an atmospheric scientist. I want to know why is the AI model making this prediction? And so, what we do is we perturb an atmospheric climate model inspired by what the xAI says is causing the AI model to make its prediction. And in this way, we use the atmospheric GCM or climate model to prove this causal dynamic connection. So, yeah, I think for me, it's like we're killing two birds with one stone. We are also, we're building trust, we're understanding, we're collaborating, we're communicating our research. And at the same time, building trust allows us to also be atmospheric scientists again.

So, something I also wanted to talk about that wasn't really brought up the last two days, Catherine did briefly mention financing in the face of extreme weather and climate disasters. But in the international space, some work that I'm involved in is discussions around how climate science and AI are going to affect this international policy and law space.

So, starting with the United Nations Loss and Damage Fund, this is basically a fund that wealthier global North countries pay into, which is operationalized by the

World Bank. And when a global South nation is affected by an extreme weather climate disaster, this fund helps pay for the loss and damages. And basically, we're in the process right now of developing this fund. So, that's one piece of what's going on in the international space.

Another piece is climate litigation. People, communities, and countries are starting to sue fossil fuel companies for climate damages because, of course, we know that some fossil fuel companies knew about climate change in the '80s and tried to cover it up and spread disinformation.

And so, the reason I bring up these two very important topics in the international space, and Karen briefly mentioned this on her Panel, is that there is this huge importance in climate attribution studies. This is research that basically says how much of this extreme weather climate event was caused by anthropogenic climate change or fossil fuel emissions, not just internal variability.

These studies are very important to inform the Loss and Damage Fund. They're very important to inform climate litigation because they help us and the courts, they help the United Nations make funding decisions on how countries can receive funding in the face of these extreme weather climate disasters that are exacerbated by climate change.

And at the same time, I just put at the bottom here, addressing data gaps in the global south. The people that will be affected the most by climate change are those vulnerable and living in the global south. And it's hard to do attribution studies in these locations because we have a gap in the long-term historical record. So, AI is important in this space because it will help lead to more rapid climate attribution studies with emulators being part of the equation, and also addressing data gaps like folks were talking about extending our historical record to inform those attribution studies.

I also wanted to talk about the elephant in the room. And by elephant, I don't mean generative AI. I literally mean there is a large elephant in the room, and no one is talking about it. And that is the fact that science is currently, especially our science, under attack in the United States. And this is very important for us as climate scientists to think about and talk about because we all depend on the research, the computer models, the supercomputers done by NOAA research, by NCAR, all of these very important agencies.

It's really important for us to communicate how we use the models, how we're improving people's daily lives. On the one hand, I think it is true that we as scientists have received taxpayer money for the last few decades, and we've done very little work in communicating how our research impacts everyone's day-to-day lives. And a silver lining of the past year, I think, is that all of us are waking up and being like, all right, we need to be able to communicate our work, or else we're going to lose the ability to do climate science and AI and advance the field further.

I also wrote, privatization of our work, what if we lost NCAR's supercomputer? What if we lost access to the NCAR's research data archive? These are all just important questions that the community should think about if things like this do happen.

I'm going to try to end on a more positive note, and that, to me is, again, the silver lining is it's a wake-up call for all of us to learn how to better communicate our research, and at the same time, that will build trust and collaboration and understanding of our work.

So, I think there is a ton of opportunity, and I'm excited to be in the field right now for sure. Thank you.

DR. DAGON: Thanks, Marc. And definitely, I want to encourage people to start thinking about their questions since I really want to get to the audience questions. I think to start, I'd like to try and wrap in some of the other perspectives we've heard from our Panel and from other Panels, and try to maybe tailor a question on the fly to each of you.

I'd like to return to this theme of constructing a narrative, which has come up in other Panels and how we can meet people where they are and connect back to the trustworthiness issue in AI for climate.

If I can start with you, Alexis, I'm wondering if you can elaborate a little bit more on the water use data example that you spoke about, and if you had to construct an ideal narrative for how to interact with communities in that space and encourage trust and open things up for partnerships, maybe you could elaborate a little bit more on that.

DR. HOFFMAN: Sure. I will say that a lot of the water use work was in the last three years, and obviously, my attention for the last year has been on more rebuilding efforts. And so, I'd say that some of the progress that we had made on the water use data stalled as a result of the fire.

But I guess some of the trustworthiness has to do with your connections. It's all about networking and who you know. And so, I basically started with all the people on the group with me and building out some of the partnerships there. My colleague had worked at Rand, and he knew somebody at the Santa Monica Water District, and then the person at the Santa Monica Water District knew somebody at the Metropolitan Water District, and just crafting the story around that.

When we first started this journey on trying to find water use data, it was really just to motivate a single community. And then it turned into us like trying to write a state bill to get water use data easily accessible for different communities, both from a climate justice perspective and training models to maybe make actual community-level decisions. That probably gets more to your question. Sorry for the ramble in the beginning.

At the end of the day, it turned into an availability issue. That type of data should be available. The story is that it should not be obscured behind - I think we were about to submit a FOIA request, a Freedom of

Information Act request, to get the data. It shouldn't be that hard to get something that you pay money for. You have to pay your bill.

So, I think the narrative there in that particular use case was more about climate justice and accessibility to make community-level decisions.

DR. DAGON: Yeah, thanks for that. Libby, if I can turn to you online, you brought up a lot of really interesting points in your slide. Maybe I'll pick up on the explainability point because that was something we were thinking about discussing more with this Panel.

If I can get you to think more about this question of how much is enough, either maybe using examples from your own work and your group's work in success stories or pitfalls with explainability, but just to get us to think a little bit more for the audience perspective as well. Like, if we're trying to incorporate explainability or interrogation or any way you want to try to push these AI models to understand them better, can you speak to this question of how much is enough?

DR. BARNES: Yes, sure. When we started this, it was the black box to me. I'll start with how it was to me. And I said, well, if I could learn how it made the prediction, then maybe I could learn some science. So, it really started with, I love science, and I want to learn more. And so, I'd start to give talks on this, and people would be like, well, they've raised the bar. And I'd say, you've shown me a blob. And that blob might be important in your map, but are you really explaining anything? It sounds to me like you had to put on your climate dynamics hat to do a bunch of legwork to actually turn that blob into anything that makes sense.

I've been thinking about this. That was back in 2020 when someone said that to me, and I think they're exactly right. Is that really the AI that's explainable, or is it still the climate scientists sitting here doing a bunch of that? As scientists, that's fine, actually, that's what we do, we're trained to do that, but we're really talking about action here and users. The end users are not trained to do that.

And so, when we say how much is enough, or Chris, who was just in the last Panel is always saying explainable to whom, I think that's so important, and it's taken me years to really ingest that. That type of explainability, the type of explainability that word means in the computer science world, is very rarely, in my experience, actually explainable to the users at the end.

So, then we say what's next. This is jargon, but that's where we often talk about this phrase, interpretability, or a model that's built by design to be understood every step of the way. And even then, it's not understood every step of the way; it's maybe understood this part, and that part, but the part in the middle is still a blob, but we're working towards that.

The thing I have found with working with end users in one or two of my projects is that, as much as you can relate it back to how they would have thought through the problem without AI, the better they are at understanding that explanation in the first place.

You asked for an example, Katie. I'll give you one and then leave it. But you've heard of analog forecasting, which is, I like to argue, one of the oldest types of forecasting there is, like the Farmer's Almanac. Today looks like that one day back 15 years ago, and the next day it was really rainy, so maybe tomorrow will be really rainy too. This is something humans, like we just do this regularly, no AI required.

And so, something we've done with our end users of one of our prediction products is really try to do the prediction as much with that thinking, that very basic analog forecasting approach, but just inject AI where it makes sense, where we need that extra help. And we've explained to them, there are these other products out there that may give you better skill, but we can explain these steps to you for what we're doing. And at least in our examples, they've always preferred the one they understood to the one that had the highest skill.

And so, I think in this, the context of this Panel, that might be something moving forward, where the pendulum has swung, and we might swing back a bit to trying to incorporate our old ways with these new ways, and try to get the best of both.

That's just one example, but I think there is so much research and work to be done here.

DR. DAGON: Great, yeah, thank you for that. So, for Marc, I'd like to pick up on what you said about returning to the fun of being an atmospheric scientist or whatever it is we do in our day-to-day jobs. And so, I'm curious if you were constructing a narrative around trying to think about meeting people where they are, and public trust and AI for climate.

If you could do that within the lens of either returning to the fun, to the big science questions, or if you want to speak to how that relates to the education component as an early career scientist, like what you would have liked to see in your training. I'll leave it up to you, how you want to elaborate on that.

DR. ALESSI: Yeah, I think that's a really good question. I definitely feel like - and Libby was talking about this just now too - I think returning to the science is so exciting because we want to understand our system. And I think you're right, Katie, it affects how we design a narrative to communicate our science and AI to the public.

I think, yeah, definitely it could help drastically, which is why I was trying to say, both things are happening at the same time. We are being our sciencey selves, which is fun. While at the same time, it's also like helping to build trust, I think, because you can design the narrative, you can be curious and you can show people that, oh, you're so interested in this and you want to understand the system, you want to be able to give them better predictions for farming or urban planning or things like this.

I think with AI, it's hard to talk about that with our communities, just because it is a black box, it's hard to understand, but when you can make the dynamical mechanism connection to what you're trying to explain, I think it makes the story so much easier.

DR. DAGON: Great point. Let's open it up to questions. We'll start with a question in the room, and then it looks like we've got a few questions on Slido. So, Douglas, go ahead.

DR. RAO: Dougla Rao, NC State University. I think the comments Alexis made about the Palisades fire remind me, because I live in Asheville, North Carolina, and in 2024, Hurricane Helene went through, and the entire community lost power, utility, and water for more than two months. And so, thinking about the resilience adaptation that really requires a really integration of different types of data, especially how AI can be able to help.

I think the questions for the entire Panel, ideal world, what would be data's ecosystem or data infrastructure can support from your own perspective, either it's from the university research settings, or it's from the industry, about the risk considerations or from, I think, scientists have been advocating for publicly funded research and those like data to support a lot of this you're making.

So, what would be an ideal data ecosystem for supporting those AIs to accelerate those climate progresses?

DR. DAGON: You want to start, Alexis?

DR. HOFFMAN: I do. When you say a data ecosystem and data to make decisions, I think of it from a commercial perspective. An insurer might want the best risk maps available. Somebody in the California state government or the city government might want information relating to insurance risk, or insurance typically uses catastrophe modeling, which is not necessarily forward-looking in a robust way.

But then, if you think about it from a climate perspective or maybe a local government, they want to look into the future. They want to know what the next ten years are like. So, and then if you look at it from a homeowner's perspective, I've been having a lot of conversations around people who want to know if they should rebuild. They want that information at their level.

And so, as far as a data ecosystem, I think there are a ton of users. So, I don't even know how to answer your question in a way because so many different people want access to similar data, but if we were talking about my parents, for instance, they don't know how to go online and use a nice little user interface to get the climate information about their parcel.

I think I might pause there and let somebody else jump in, and while I think a little bit more about it, just because I think there are so many facets to your question. I'll pause while I think about it.

DR. DAGON: Libby, do you want to say anything about the ideal data ecosystem question?

DR. BARNES: Sure, so I'm not going to design the whole thing on the spot here. I think something that needs to change from how we've done data in the past. In the past, most scientists collecting data and putting data together have assumed other scientists will use that unless it was put together for a specific reason.

I think now data is the currency, and because of that, everybody wants access to everyone's data. And what that means is we need to - this is going to be a boring answer, Katie - but the metadata side, really describing, what was this data collected for? What are its downsides? How should it not be used? How was it intended to be used when it was put together in the first place?

I think that's something that I was a bit complacent about over the years. And I'm realizing now that we need to clean up and put a lot more time and energy and emphasis on that piece because that data will be used, and it's our job to make sure it's used properly. I need to rephrase that. It's our job to do as much work as we can so the people that want to use it will use it properly.

DR. DAGON: Yeah, absolutely. I think that resonates a lot with what we were talking about this morning. Marc, do you want to add anything to that question? Or there are also a couple of specific questions for you in the chat, so do you want to move on to those?

I see this question popped up. Thank you for mentioning the elephant in the room, climate science is being attacked. How should we respond to climate change doubters and deniers? Can we find ways to win them over? Maybe this relates to broader work that UCS is thinking about when working with politicians or different funding agencies. Do you want to speak to that at all?

DR. ALESSI: I will try to. I will first say, if you're familiar with the work at Yale that asked these kinds of questions, do you believe in climate change, how concerned are you, to the general public, I think they do it like every couple of years. It's like over 70 percent of Americans are like, yeah, climate change is real, we see it, we know what's happening.

So, to me, the issue right now isn't really trying to convince the general public, just in general, about climate change. I think it's more like the people who are representing us and our government, what do they think? Who is funding their campaigns? Are there any bad actors?

What I was trying to hint at yesterday. These are all very important.

I think as scientists, and I was trying to also mention this when I was starting, we used to be able to take a back seat, and that's not the case anymore. Like if you are sitting in this room and you haven't yet called your representatives to advocate for NCAR, you're making a mistake. I will also say that there is significant bipartisan support for the work of NSF, of NCAR, of NOAA; they know that what we do affects everyone in the United States. I think it's just, electing is important, voting, and just advocating your work, I think, is important.

DR. DAGON: Thank you, appreciate that. Let's turn back to the room. I see some folks up there. Let's get another question in the room, Anna, go ahead.

DR. HARPER: Yeah, hi, Anna Harper, University of Georgia. And I'm going to come back to a point that Libby brought up about education. You all have some background in atmospheric science, and so, thinking about how our field developed and how it was taught, it happened over probably a few decades. The first general circulation models were developed in the '60s. And now, if you look at any atmospheric science program across the US, there is a standard suite of courses that are offered.

The students that are coming out now, and we as well, have to operate in this quickly evolving world. And we don't have decades to decide what the new programs should look like. So, I want to get your thoughts on that. And the thing I think is we can't - I guess it comes back to something Chris said - we maybe don't have time to do this on our own; we need to collaborate on these educational initiatives.

I'm wondering what you all have seen as far as approaches to do this, and maybe thoughts of ways that education can work together to do this.

DR. DAGON: Yeah, maybe we'll start with you, Libby, since I think this came up in your remarks.

DR. BARNES: Yeah, sure. This has been a longstanding issue for the past few years of educators saying, how am I supposed to educate this totally new tool when I didn't have a chance to learn it, and all my students are scrambling?

One thing that's given me great hope is that many people in STEM fields are already data scientists anyway. We just didn't call it that until the last five to 10 years. And so, I've actually found it, in my experience, far easier to teach the basics of AI rather than the physics side of things. I don't think our curriculum has to completely change to really start to incorporate the AI component. And ultimately, we've heard in the last two days that domain science is still important. That's not going away.

And so, I personally would say, we need to stick to what you know, and let's add. And I actually think many students are ready for that just with a few courses.

I will say something in terms of your collaboration comments that has also been positive for me is that the private sector, which really up until now has been taking over training these huge models, when I talk to them, they all say we don't want to train our staff in these topics. We don't want to train them in the domain or the AI side. We are so thankful to the universities for doing this part for us, so they can hit the ground running when they join our team.

And so, with that in mind, we've been talking about, like Marc brought up, privatizing our science and that sort of thing. I actually feel some hope there that education may even just be that much more important to make sure the right people are getting trained in the domain and the AI together to really make a difference out there in the world. So, that's my maybe positive sense on that.

DR. DAGON: Marc or Alexis, do you want to add to that?

DR. ALESSI: I just want to say full disclosure that Libby is actually the person who taught me most of what I know about machine learning and AI. And something that has really stuck with me that I remember her saying, in probably the first class I took with her, was that AI or machine learning is just another tool in our toolbox. We use climate models. We have all these different statistical approaches. Now we have AI, it's just an additional tool. We're not replacing anything in our curriculum with AI. It's just something we can use to develop new hypotheses to understand new dynamical mechanisms in our climate system. So, I think that is really important.

Something else that also I think sticks with me is there are this paper that Kerry Emanuel and David Randall wrote about the weather and climate schism in our community, where a lot of meteorologists don't work with climate scientists and vice versa. And I think that they really stressed it's important for education in the future to make sure all students take a climate dynamics class, but also a synoptic meteorology class, so that they can really fully understand how the two fields are connected.

And I think that is so critical for domain expertise as AI continues to grow in our field.

DR. DAGON: Great, thanks. I think next we'll move to an online question that is primarily for Alexis, but I think we can hear from everyone on this. So, being on the front lines now at Palisades, you mentioned water data access was not available at the scale you needed, and you actually had to advocate legislation as part of that process. So, can you share any communication lessons that came up or any thoughts on communication when moving from this data space to advocacy and interacting with the legislative process?

DR. HOFFMAN: I think at the end of the day, the communication lesson that I took away was, in conversations, you want to develop a narrative, but maybe try to keep emotion out of it, I'll say. You want to be, as a matter of fact, especially when it comes to writing bills, you wanted to keep it very clean, I guess, and precise. But then almost the converse is true when you're trying to garner support from the community and then you want to share emotional stories and things that garner reactions and I made some of Kieran's points earlier, like it's reactive and click-baity, and so you wanted to get buy-in and get people to advocate for your bill but there was a lot of pushback for certain people.

One of the downsides to doing this in the Palisades prior to the fire, it might have used a lot of data, so people didn't want to support that bill. Local representatives didn't want to support that because they probably would have highlighted them as an outlier in using too much water relative to the rest of the state, which was interesting. So, you kind of have to understand all of the perspectives in the conversation, I guess.

I think that's probably the biggest lesson. Again, a very personal lesson. I feel like I should talk more about other climate-related things, but I'm glad you guys are interested in some of the Palisades stuff.

DR. DAGON: Well, yeah, on that personal note - well, first I want to check to see if we have any more questions in the room. There are a few folks standing in the back now. We have about seven minutes left, and I think I want to end with asking each of you to discuss your personal takeaways because I think that's actually very compelling to hear folks' personal perspectives of how attending this workshop, participating in the discussions, might influence your work going forward and your engagement with this topic.

We could also combine this with thinking about the greatest near-term opportunities, but I think that's come up in some of our conversations already, and I'd be really interested in your personal takeaways.

I'd also be interested to hear how we keep these types of conversations going with this unique audience, where we might not always interact with these folks at the conferences and meetings that we normally go to. So, if you have ideas for how we can keep the momentum going, I think that's something that all of us would be interested in hearing.

So, let's see, why don't we start with you, Marc?

DR. ALESSI: So, I was thinking about your last question: how do we keep these conversations going? And I think this is why this workshop was so amazing is because a lot of us did come here in person and it's small enough so that we can make connections and I think just that networking and connection building and understand what people are doing, again, across all these different professions that to me is building the conversations and that will, I think, continue past this workshop. So, I think that answers that.

And then how will this influence my work? I think I really need to start thinking more about building trust.

I think we've talked about the trustworthiness of our AI model so much, and I really just want to be sure that I'm not doing something wrong in my work, and I want to be sure that I'm communicating it correctly with the right confidence to our stakeholders.

When we're talking about our research, perhaps to people who are involved in a climate litigation lawsuit, people will start to use our research on climate attribution, for example, and I think that's where it really impacts how society functions and how we react going forward.

DR. HOFFMAN: I'll jump in, I won't have a long-winded answer, or maybe like piecemeal. But the first thing I wanted to say is what struck me over the last couple or last 36 hours is - in a conversation around something as complex as AI and climate, both of which are incredibly complex, especially when we put them together - it almost was the human aspect that has emerged. And talking about AI and complicated issues, we come back to the fact that we need to talk to each other. All of this trust starts with the human relationship and building relationships with customers, end users, or decision makers, whatever you want to call them. It starts with human-to-human contact and relationships.

From scientific training, you don't really consider it that often in a weird way. And so, focusing maybe on moving forward in science communication and even advocacy around that. I think somebody spoke earlier on the fact that there were no science journalists. That's wild to me, I think. And I think that there is a gap there in taking what we do, communicating it in the right way to the right people, and just meeting people where they are.

To have that precipitate out of this conversation and mixture around climate and AI is wild to me.

DR. DAGON: Great, yeah, thank you. Libby, let's turn to you, thinking about any personal takeaways, how this might influence your work going forward, or how we could continue these conversations.

DR. BARNES: I guess the word I'm leaving this meeting with is perception, and that how people perceive the AI, what it outputs, how they perceive the information they're getting from these elements about climate information is really the key here. Not about us scientists, if we know what the truth is, whatever that means, but really, how people perceive that information.

For me, in terms of a personal takeaway, Angel's talk yesterday was amazing and has made me rethink a lot of things. I have been concerned about climate information being processed about LLMs. It made me feel hopeful that there were smart people thinking about this problem, but I actually think we all need to be thinking about this problem in our own way.

And so, that's the thing I'm going to be mulling over in the weeks to come.

DR. DAGON: Great, well, I think we'll end there. I just want to thank the Panel again for their participation and thank you all for your engagement.

(Applause)

Agenda Item: Meeting Reflections and Take Aways

DR. SAIN: I think we're going to do a quick wrap here. We do have two questions in Slido. What general area did you learn the most about during this workshop? Importance of trustworthiness, and I'm paraphrasing as well. How both AI and more traditional models along with observational data will be necessary. Different approaches of how AI can be used in climate science and climate action, and AI used in climate science and climate action is different from ChatGPT.

Sorry, I didn't really mean to stumble on that one when I read it, but I did. We have a couple more. What is the greatest challenge you heard come out of the workshop? And the last one, what is the greatest opportunity you heard come out of the workshop?

Importance of trustworthiness is up at the top. Different approaches. And then the third was how both models will be important. I was hoping that one would jump up a little bit. But trustworthiness is definitely holding on there at number one. Well, that's interesting. Trustworthiness, data. Maybe that's two. Trustworthiness and data. Okay, I think we get the picture.

What's the next one? Oh, nice, nice. Collaboration. Well, I have an even longer list of thoughts from today's meeting, but I'm not sure we want to see that. But does anybody have any thoughts as we walk out the door?

DR. MORRISON: One thing that maybe has happened in this workshop is that we have many different connotations of trust. And so, I think it just might be worth making a distinction between trustworthiness as Chris defined it when he was sitting on the Panel as something that is relational and is an element of a human's perception of an instrument or another person. And something like reliability, which - and here comes my favorite word - is an epistemic feature of an instrument that is to some extent independent of people's subjective perception of what might give them trust in an instrument.

I think both are equally important, but the distinction is worth noting because what makes something reliable is not necessarily synonymous with what makes it trustworthy, because humans are fallible and we can get things wrong and place confidence in the wrong sorts of properties of models or data. So, just worth noting.

DR. SAIN: I have it on my to-do list to learn how to say the word epistemic. Now I just have to learn what it actually means. Arthur, do you have another?

ARTHUR: Yeah, very quick one. I'm retired, but I really enjoyed this workshop. Thank all the planners and the Roundtable Committee members, a fantastic job, fantastic job.

(Applause)

ARTHUR: So, I just hope to see the proceedings of this, maybe sometime, whenever April can get to it.

DR. SAIN: They'll be out. I just don't know what a timeline is.

ARTHUR: Thank you.

MS. MELVIN: We'll aim to release the proceedings late spring to early summer.

DR. SAIN: Okay, do we have another? All right, Alexis.

DR. HOFFMAN: I was just going to say, I completely forgot that I had slides, by the way. And yesterday I actually used Otter AI to transcribe everybody's Panel discussions and keynote presentations. And I made a word cloud, which is hilarious. And I'm not going to ask Jasmine to pull it up, but I want you to pull it up because it was really funny, and it was interesting because when I asked it to pull out main themes, it actually had climate action and collaboration and data and trust. Yeah, there it is. Yeah, so thank you, thank you for that.

These aren't proceedings at all, obviously, but I thought it was a fun way to highlight some of the stuff that we've been talking about over the previous couple of days. But also highlighting the fact that there is some work still to do, and then crafting the narrative around this.

DR. SAIN: Very much. All right. Do I have another one up there?

DR. SAMARAS: Hi, all. Costa Samara from Carnegie Mellon. Just thank you to the organizers and all the Panelists and then the Academies for a wonderful workshop. I wanted to offer a parting thought on the things that we heard today around the challenge of epistemic crisis in science in general.

And so, AI accelerates a underlying knowledge and information asymmetry that already existed. And so, we in this broader community have an opportunity to course correct and enable a broader discourse between science and the public in a way that gets better outcomes.

How do we operationalize that with AI and climate? I think that this is a great opportunity for this community to engage with the broader AI ethics and computing side of the research and policy that grapples with a lot of these issues, without the climate part there. It seems like we are talking to ourselves and they are talking to themselves, and we should have a broader collaboration with the social scientists, political scientists, and ethicists that are grappling with the rest of trust, democracy, information, and participation, really with a very strange absence of us as climate researchers and scientists.

So, maybe this gives us a way to talk to ourselves, but also find new communities to talk with. Thanks.

DR. SAIN: Thanks. Yeah, that's a great call out. Thank you.

(Applause)

DR. SAIN: Anne, I saw you going up.

ANNE: Yeah, I couldn't have asked for a better lead in from Costa. That was really fantastic. I want to thank you all again. Echo everybody's thanks for all the work that the Committee put into this workshop. And you guys did a fantastic job of pulling together some really amazing speakers. I wish we all were as dynamic and charismatic as Angel, but I think there is a lot of room for more collaboration with social science.

I also want to say that I personally was impressed with the number of partnerships and opportunities that we saw in this workshop. Thank you.

(Applause)

DR. SAIN: Thank you. Anybody else? I think we're done. Thank you. Thank you so much. And I will echo again, having been in any number of meetings with the staff over the last few months, I am so impressed with how they operate, how they work, how capable they are. And there is absolutely no way we could have pulled this off without them. So, thank you very much.

(Applause)

DR. SAIN: Thank you to all who attended. And have a good one.

(Whereupon the Roundtable was adjourned at 3:30 p.m.)