PG&E - Marketing & Communications | Teaming Up on AI Models & Data

Good afternoon, everybody. It's a pleasure to be here with all of you. My name is Norma Grubb. And I'm the enterprise AI and data science director for PG&E. What a day. What energy that we are building to change the future using the latest technology and build those technologies to serve the purposes that we have prioritized in a safe way. We're using responsible AI.

I'm going to invite my panel here to join me. We have Brenden Russell, principal manager, technology, strategy at SCE. Welcome, Brenden. Mark Spieler, senior manager, director, global energy industry at NVIDIA. He was previously in a panel. Welcome.

Thank you.

Marc. Jon Eric Thalman, director of risk management and analytics at PG&E. Welcome, Jon Eric. I'm going to share more bios later on. But for now, this is the panel.

The title of our panel is Teaming. up on AI models and data and why we want to talk about this because one of the ways we can accelerate the adoption of AI to fulfill strategic goals is to partner up across utilities, to share use cases, to share the code, the data, and the modeling approach for use cases. And this is what we want to discuss today.

And because one of the goals of the summit is to leave this room energized and with action items that we can move forward really in a very specific way, we are going to ask you to participate in this session. So let's talk first about what are the questions that we are trying to answer a little bit more precisely. Are we sharing? Yes.

The questions are, what are the benefits and risks of sharing data and models across utilities? So we all think and we hear our senior leaders saying, why don't we share the way we do this work? So we don't duplicate. So let's talk about what are the benefits and risks of doing so.

And if the benefits outweigh the risks, what collaboration mechanisms should we consider in order to move forward when the benefits outweigh the risks? What is the problem we are trying to solve? We know as utilities that there are two realities right now.

One is that we have overlapping modeling initiatives. And we have a use case we're going to share, in which three utilities in California are developing solutions to very similar use cases. And there is an opportunity to work together on that.

And also, as data scientists, we know that when we don't have enough data, it's very difficult to model any problem. It's very difficult to build any AI. So if all of us partner up and share data specifically in infrequent events-- those events are very rare, which are the ones that we want to detect. If we can share that data, then we can accelerate the way we build accuracy in our models. And these are the case studies that we want to share today-- computer vision for asset degradation-- PG&E, SCE, and SDG&E. We have independently developed computer vision models to predict degradation of assets.

So the questions are, how can we set up a framework to enhance collaboration and accelerating model development in order to share more and duplicate less the effort and build also knowledge across utilities? This is knowledge and capacity that we are building about how to develop these models and provide these solutions. And how much more accurate could asset degradation predictions be with more data input in the models because we are sharing the data?

Quickly, what we believe is holding us back. There are three types of challenges that we face when we try to share. First of all, we have incompatible tech stacks. Some utilities may use PyTorch, TensorFlow. Some of us have AWS. Some of us have Azure. How do we make sure that we can come up with similar frameworks or model frameworks that are well connected and also modeling assumptions, such as class definitions, the training data characteristics, et cetera. We need to agree on this.

Then data security and privacy. Customer GIS data and more data are highly sensitive. And if not properly managed, then this sensitive data could be at risk. And there is a separate along the California 50-50 rule for safety sharing of data that we need to bear in mind in California.

And we have different legal and regulatory positions across utilities. Different utilities have different regulatory environments and legal positions the lawyer team assume. So how do we agree on those in order to be able to have a framework to work together? So that's why we are asking you to help us think a little bit.

If you can then scan this QR code. And we are asking you. The benefit of expanding collaboration among utilities to share models and data exceed the risks. Do you agree with this? Do you strongly agree? Do you agree? Are you neutral, you disagree, or you strongly disagree?

So please go ahead and scan the code. And whenever we are ready, we are going to switch. Here we are. Strongly agree-- look at where we are moving-- 60% almost.

Agree's growing. In any case, the strongly agree and the agree are pretty massive-- almost 90%, 93%, 95%. We are moving towards, yes, we agree, with different levels of emphasis. Thank you so much. Let's switch please to the presentation.

So now, I'm going to invite my panel here to share the expertise and insights so we can keep building knowledge on how we go about doing this. I'm going to share a short bio from each one of them so you get to know them better. Brendan Russell is the principal manager of grid innovation at Southern California Edison and has over 20 years of utility industry experience.

In his tenure as CE, Brenden has served in multiple leadership roles, including overseeing the technology and business strategy function, SCE's grid modernization program, and has filled engineering and managerial roles in support of advancing SCE's real-time grid operations, championing the operationalization of new grid technologies in support of SCE's long-term net zero vision. Prior to his relocation to the United States in 2011, he worked for Endeavor Energy in Australia, where he gained extensive knowledge and experience in utility operations, supporting various functions, including operations, engineering, system planning, and construction.

Brenden graduated with honors with a bachelor's degree in electrical engineer with a specialization in power systems. He is currently a licensed professional engineer in the state of California. He also holds a master of business administration from the University of California, Los Angeles, UCLA. Thank you for being with us.

Marc Spieler is a recognized leader in the technology and energy sectors, currently serving as senior managing director for the global energy industry at NVIDIA. Named to Business Insider's 2024 powerhouse in AI list and serving on the advisory board of Grid Forward, Marc has been instrumental in transforming how energy companies use accelerated computing and AI to tackle industry challenges across oil and gas, renewables, power generation, and utilities.

At NVIDIA, he drives applications in high performance computing, deep learning, and AI, focusing on enhancing decision making and optimizing outcomes for energy companies. His expertise in leveraging technology to address complex industry challenges is backed by an MBA from rice business and advanced degrees in organizational leadership and marketing. Welcome, Marc.

Jon Eric Thalman, director, risk and data analytics, is responsible for developing second life models that represent risks and potential risk reduction due to work plants spatial across the electric transmission and distribution systems. These models currently provide insights that guide the prioritization of electric system work plans for inspections, maintenance work, and asset replacement to vegetation management and system hardening. Jon Eric joined PG&E in 1996 as a transmission planner and has led teams with responsibilities in both transmission and distribution assets, strategy and planning, as well as regulatory strategy, vegetation management, and corporate strategy.

Jon Eric holds a bachelor of electrical engineering degree from California State Polytechnic University Pomona and a master of science degree in electric power engineering from Rensselaer Polytechnic Institute. He is a licensed professional electrical engineer in the state of California. What a wonderful panel to go deeper in these questions.

So let's start the conversation. I'm going to start with you, Brenden. Let's look at the art of the possible, pretending for a moment that we are able to resolve these key risks and we are able to move forward. What will this data and model sharing future look like? What is the vision if it's successful? What are some of the representative use cases that might highlight this future?

Great question. So I'm going to take a step back as well to answer that question. I think in our industry in particular, like for many, many years, we have a lot of, you could say, silos of excellence, actually. And what we need is actually the opposite.

We have a lot of data silos of excellence. And a lot of that's been inherent in the way utilities have been able to plan for the grid, from a predictable load standpoint, and then obviously dispatchable supply. But as we get into the world where we're going to have to forecast how the supply is going to operate on the system, demand is going to obviously need to become a lot more flexible, it's going to require us to share data at a volume at a scale that probably we haven't ever done before.

And so I think if you were to go out 10, 20 years look at our end state, where we're in net zero, we're going to have to require not just I think data collaboration across utilities and sharing but I think I would even say beyond the utility as well as we think about interfacing and orchestrating customer resources at the edge of the system to be able to unlock that dynamic supply. And so I think at the end state, having a world where you can actually orchestrate all these resources is going to require a lot of data sharing, not just for planning purposes but also in the operational theater as well, where a lot of that is going to have to happen real time.

And I want to put this out there as well. It's not like the electric industry hasn't done this before. If you look at the transmission system today, actually, data is shared between utilities, between the California ISO. It's actually shared on a real time basis today to actually manage the reliability of the system.

And it's just not planning as well. Like, we share models of our transmission system to be able to plan the bulk electric system. And I would preface this. There is a model out there. It's been done before.

But as we think about the distribution network, we're going to have to do this at volume and at a scale we probably have never done before. And so as I think about that end state, I think-- and my team does. We're thinking about that. What do we need in place to do that?

And I think as we think about AI specifically around the models and the data, we're going to have to think about different types of techniques that can be adopted today. Actually, Marc and I were talking about this before, this whole idea of transfer learning, federated learning, not having everyone build their own models. But how can we actually build upon the models and share the models and then retrain the models and then share that back out?

And so as we think about the challenges we have today around data privacy or data security, we need to think through the technical apparatus that could actually facilitate doing this at scale. And so I think there have been other industries that have done this. And we can definitely learn from that.

But I think it's an exciting opportunity as well for us to come together and do that. And ultimately, we also need to bring in the right industry stakeholders and partners who can really drive standardization as well across the industry. And so as I think about going back to my example around what we do on the transmission side today, there's standardized protocols that have been developed. But it requires really the industry to come together, define those, and then push that forward as a way to really enable the grid of the future. So yeah.

Excellent. Thank you so much. Data-- I like the concept of the style of excellence and also the federated model as an implementation practice. It also is recommended by the governance community as a way to save the privacy of the data. So in the same line generic, what would be the benefits of getting this right in the line of work of mitigating the risk and moving forward in accelerating the collaboration? What do you think the benefit will be in getting it right?

That's a good question as we look at, what is our goal? What do we want to do? There's obviously power in the models, a predictive power. There's obviously great insights. And as Brenden mentioned, as we look at other industries that have had gone through this adoption, there's definitely a maturity curve that seems to take around a decade that's a little sobering.

If we look at energy procurement within the utility industry-- energy procurement, capacity planning, even Nuclear Regulatory Commission, with modeling that they've used as part of their regulatory construct, it takes almost a decade to go through that. So things we can learn from them and some of the risks along that maturity path is that the improvement that's going on. Data is improving. The understanding of the models is improving but within the users and the utility, within the regulatory, within the public, as you're trying to explain why you're making decisions based upon the models, whether it's proving compliance or improving the performance or efficiency of programs.

So one of the risks here is that along that maturity curve, we are subject to-- I'll call it armchair quarterbacking. Hey, you did this three years ago. You've got a better model. That new model proves that your old decision was not optimal. So there is a very real tension there as we go along this improvement that we're judged by our past mistakes, even though that is the exact path we have to go over to get there. And that risk trickles down through the desire to share more, to be more transparent. All that pressure is placed upon that. So that's one of the risks I would like to highlight.

Yeah, absolutely. AI is powered by data science. And data science has science. There is a science methodology that we need to follow. And that requires time and data and experimentation highly. So it will take some time. Thank you for bringing that aspect.

Marc, going back to you, why can't utilities learn from other industries, particularly regulated industries like ours that have implemented successful data model sharing efforts?

So this is a really exciting time to be in energy. And I think compared to other industries, our ability to collaborate should be easy. And what I mean by that is utilities are regulated monopolies of sorts, where you have an area. And your customers come to you for services. And you're not necessarily competing against others.

But your goals are the same goals as other utilities-- reliability, resiliency, safety, affordability, all of those things that all utilities are striving to be So why wouldn't you want to collaborate? When I think of other regulated industries, I think of health care is probably one of my favorite examples of industries that are regulated.

They can't share customer data or patient data. They have a common goal to wipe out disease or solve cancer or do other things. They're a little bit more competitive because, obviously, you have a choice as to where you go for services. So people are competing for patients and things like that.

But to me, it serves as the best example of something that we should lean into. They have a common objective. How do we solve cancer? How do we cure cancer? But we can't share data.

So a university, a research center, a hospital, all have pieces of the puzzle. And they all have different examples or anomalies that they see. But how do they train models and use AI to solve those problems when they only have a piece of the puzzle? But by using federated learning and tools out there, they're able to start with a model, take that model in, and train it using their data without letting their data leave their facility or sharing proprietary information or confidential information.

And then those weights are sent back. And they're able to strengthen those models and then redistribute those models out to the edge. We do this with our cell phones, with autocorrect and stuff like that. We just don't think about it. But we don't want our data and our text messages being read by people. But to understand that 85% of the time this was corrected to this helps the model for the next person.

Same with health care. And in this industry, we should be doing this. My question is, why wouldn't we be doing this? I live in Texas. And we lost energy for about three or four days at my house, sometimes over a week for a lot of people during burrow or also during the winter storm that froze so many things.

But if you ask yourself could ERCOT or could CenterPoint have trained a model to predict it, got to be honest, they probably didn't have enough freezing data to create a very accurate model. But their friends up North probably did. And this is where I go back to is, is these anomalies that help you train for the unexpected has to be done by sharing models, sharing data. But moving data is expensive. And storing data is expensive. So how do we collaborate on those things?

I think it's really important to understand that the job of a utility is to not have failures. So when I ask you, how many failures have you had, chances are it's not a lot in certain areas because you do everything you can to prevent it. So when you're training a model for a failure, you don't have a large amount of failure data if you're good at what you're doing.

So it's somewhat counterproductive. And therefore, that's when you talk about a little data versus a lot of data and how we can expand that by reaching out and collaborating across the industry, where we might have certain use cases that some people might have only seen a few times or actually experienced or have done the root cause analysis on or certain pieces of equipment that maybe they don't have a lot of or any of or no expertise anymore. But somebody else does.

And if you train models to understand, to be able to predict, to be able to read a manual and be able to talk about what that is, it could take you a week to try to find that in your own system. But a model trained by somebody who has that makes it available to you. So the collaboration shouldn't be a fear.

I really, truly believe that this industry, more than any, should be able to do it and do it with the easy stuff first, things like predictive maintenance, customer support, different back office functions, operational things that don't cause high risk until your users buy into the validity of the models. And I tell you. We're an engineering company. So we're first principles-based.

But even with some of our energy models around solar prediction and climate prediction and all of those, I tell you the AI based physics-ML models are as good or better in some cases than traditional full physics simulation. And until you see it a bunch of times, you wouldn't believe it as an engineer. But it's getting really good.

Thank you so much. Excellent insights from experience based on the benefits of this and how we need to collaborate together. This question is for Brenden and Jon Eric, either one of you who wants to pick it up first. What pitfalls would need to be avoided for this to be successful, focusing on the pitfalls that are potential?

I can maybe start. And then go ahead, Jon, if you want to add to it. Some of, I think, the key learnings our teams have come back like-- because we've been doing this for a couple of years. Actually, I think wildfire was sort of a really good catalyst to start looking at machine learning models to help us with some of the understanding the risks on the system.

I think some of the big pitfalls-- no one's perfect. We're all learning as we go through this journey as well. But some of the big pitfalls is sort of like unlocking a lot of the data silos that we have internally. You don't know it until you start to match data from different diversity of different data sets to actually start training your models how siloed some of that is.

And a lot of it requires good sort of data architecture, data management, data governance to be actually implemented at the enterprise level. Because the way we ran the grid traditionally, sort of going back to what I was talking about before, we got away with what we had because it was good enough. And everything was predictable from the demand side.

And we had the certainty on dispatch of supply going back to first principles. But now, as we become-- our infrastructure is impacted by these climate events happening on the system, going what Marc was just talking about. We really need to bring in a lot more diverse data sets.

And some of it is like, actually, we don't have the data. We might have to start acquiring data from different sources. Or we might have to expand our census strategy to be able to get the right data sets in to train the models.

And so I think over time, it's a learning experience. You have to have a culture internally that is driven by learning and learning from one iteration of the model to the next to be able to improve the accuracy and get to the outcomes that you want to get to. And so I think some of the pitfalls, like data management, even data quality, do we have the right?

What is our data quality? And feeding bad data quality into the model, it's like garbage in, garbage out. And so could result into bad predictions or false predictions, et cetera. And so foundationally, I think just the emergence of AI is sort of just emphasized the role of data within the utility and how there needs to be stronger governance around that. Jon.

Those are all good points. I think prompts me to think about also there's the experience of producing AI models for use. And right now, it's enjoying a good popularity that everyone's very excited about it. And it can have the pitfall of being a shiny object, if you will.

And you have this model. And wow, it has this really great output. And we've got all these great examples where it's working well. But behind that, you have to remember-- and the pitfall here is that the application of the model is as important as the development of the model.

The user of the model, it can't just be, oh, there's this cool thing. Let's use it. You need to understand how it was developed, what the purpose of it for. And the thing we always remind ourselves what PG&E is models are for insights and not for answers.

Because the minute your user is like, oh, there's the answer, I'm off and running, there's a higher chance you're going to make a misapplication of the model. And the challenge with the misapplication of the model, obviously, is then people lose confidence in it. And then you've got a problem where you're trying to prove the value of it.

And then people aren't paying attention to it. And we lose the value. So it's important to develop the output of the model, to educate the clients, the people who are using the model, so that they understand how to glean insights from it so they can make better decisions. And that's whether it's, like I said before, referring to just like a compliance assessment or risk assessment or driving work or where to do work or safety of the system.

The other thing I want to add is, as Brenden mentioned, it's very dependent on data. And historically, in the utility industry, data was for two purposes. One, to run the grid safely. So we needed to know the status of the grid to keep the lights on.

And second was to generate an accurate bill for our customers. Past that, uses of data have, of course, have continued to evolve. But traditionally, that's what utility data was for.

Now, we want all sorts of other applications. And we go to the same data sets and think, well, I need these other columns. I'd like something else. It'd be great if the troubleshooter would tell me what failed instead of just focusing on getting the lights back on. I agree that that's their top priority. But I'd love to know a little bit more about that thing you threw in the trash that broke. So those are some of the pitfalls we-- as we're on this, like I said, maturity travel here.

And from--

Maybe just add to that, I think that's key that iterative feedback sort of agile process to leverage the output of the models but sort of inquire what additional data could actually improve it. And so I think your data strategy and your acquisition strategy is going to evolve over time. We're just going to have to accept that.

But we need to have that culture internally with your workforce to be inquisitive and drive for those answers. But I think that's sort of a realization that our teams are coming to as well at the same time.

Yeah, data always at the back end of everything. And not using the model for a purpose the model wasn't intended to, it's one of the best practices in governance. We create what is called model cards. And the model cards should explain what the model is intended to. Like an owner's manual, you should read your card and know what's exactly the intent of the prediction of the model. Excellent points.

Turning into tools, what tools could we put in place to mitigate those risks or worst case scenarios? I think we talked a little bit about data governance, this concept of documenting the purpose of the model. Which other tools-- and I opened just this for the three panelists. Marc, you talked about federated learning if you want to expand a little bit more as well on that.

Sure. So there's lots of tools out there in this space. And there's proprietary tools and open tools. And I think as an industry, we will develop additional tools that are industry-specific. And I think that's a big key.

So from a federated learning perspective, there are tools that run on certain infrastructure and certain clouds. There's open federated learning tools that run across clouds or on any cloud or on prem. I think choosing the tools that allow you the most amount of flexibility is always great.

But once again, you're not a data science community. You're not a software development community. Your job is to run energy grids very effectively. So you want as much as necessary, as little as possible, meaning, how do you solve the problem with as much effort as necessary but as little effort as needed? And so finding those tools are going to be pretty important.

I think what we're going to find as we collaborate across the industry-- and I think this work is already underway. And I think there's organizations out there like EPRI and Grid Forward and Grid Connect and all of these that are trying to work across utilities to solve problems that will allow for this collaboration.

I think what we'll start to see is industry-specific models that are going to be foundational models that are quite a bit smaller than what in a trillion parameter model is out there today. A utility doesn't need to know how to bake a cake. They don't need to know how to deliver flowers. They don't need to do these things.

They need to understand power systems. They need to understand customer behavior as it relates to utilities. They need to understand the specific terms that this industry needs.

And I think what you're going to find-- and we talked about this a little bit in your opening-- is we could have 50 utilities training their own models or fine tuning or creating vector databases and stuff. But eventually, that will get costly for each of those utilities. The difference is is how do we build industry-specific models that we can invest in and then build submodels off of that for grid planning or for capital budgeting or for regulatory approvals for customer call centers or whatever.

I think you're going to see small language models evolve from companies that will then take those open models that were trained on more data and create company-specific models that are more tuned or use other technologies to solve those problems. But I think this is where the industry is going. As opposed to one company putting five people, five data scientists, or software developers on a project, you might have five companies that each put one. Now, all of a sudden, your cost is 20% of what it was before. You can solve more problems.

One thing to think about-- and this is why AI is, I think, blowing up as much as it is. There are more problems in this world to solve than people to solve it. And so I think to your point, Jon Eric, it's not a matter of answering the question or replacing people. It's a matter of augmenting them and allowing them to do more work faster.

I often use the analogy when somebody created-- and I don't know who it is-- created the calculator, it wasn't the end of mathematicians. I think mathematicians might have been a little afraid. Why are people needing me if they've got a calculator? But what it really did was allow them to solve much harder problems much faster.

And there is no shortage of hard problems facing this industry, especially as we add a tremendous amount of DERs. We've got all kinds of regulatory things. We've got to increase the amount of generation. We've got no shortage of problems. We will have a shortage of people to solve all those problems.

And so leveraging that AI and those open tools-- because not everything can be solved in a data center. Some things need to be solved at the edge. Some things need to be solved in a facility. If the data needs to be tied to that facility, data sovereignty issues, depending on which country you're in or which state or whatever, you need to have tools that can work anywhere. And so open models, open tools coming up with consistency that will work across all utilities, across all clouds, across all on-prem locations, I think will make us much more efficient in solving this problem.

Wonderful. Thank you, Marc. Just for the sake of time, I'm going to move us on to a new poll. So whenever we are ready, this is a poll. So as I mentioned before, we want to leave this room with some action items and some ways to start collaborating to build this next generation future, this AI and power future.

So the question is, which of the following collaboration mechanisms, if any, will be exceptionally helpful in spurring this collaboration? Here, we have the first answer. So we are talking about regular cross industry meetings, data sharing platforms-- oh OK-- model sharing across utilities-- there are some strong opinions there-- and open models-- having open models that we can share code. It seems like we are in between data sharing platforms and model sharing across utilities seems to be the two winners. Still giving people some time to keep voting. Thank you for giving us your voice.

There you go. It seems like we are between data sharing platforms and having a data sharing platform and model sharing across utilities are competing head to head. So let's keep the conversation moving. We have a few more minutes only. So this question is for Brenden and Jon Eric. How do you think about the relative feasibility of these collaboration mechanisms that our audience were voting on?

Are there other that are missing? How do you think about these mechanisms on a maturity curve on how might we accelerate progress across this maturity curve of collaboration? Any insights about these mechanisms that you can give us?

I'll share a couple of thoughts on it. If you look at other models and other efforts that are in the utility space, at least that have gone through this maturity curve, there was a lot of trial and error. There was a lot of learnings.

And I think what I've observed looking at that is that depending on how those learnings or opportunities are viewed and incorporated across the utility, the interveners, the regulators, the people producing either the models or the equipment. How they see that and how quickly they use that feedback to improvement is a big gauge into how fast we move along this maturity curve.

There also seems to be-- there's a key point, I think, where you move from each utility kind of have their own boutique approach to something to an industry standard starts to become established. And it can come by a vendor saying, OK, I'm bringing all this together. And I'm offering a product that's now cheaper for you than setting up your own shop or through industry groups like IEEE or something, where we start to establish standards.

But there's a key inflection points, I think, on that maturity curve, where those commonalities are agreed upon. And there's multiple ways to get to those. But I think those are key milestones for us if we're thinking about the maturity, is how do we get to those key points where we all start to agree and coalesce around, OK, that's the way it works and then let's go to the next thing.

Excellent. Brenden. Go on.

Yeah, I agree with everything Jon Eric just outlined. I think as well, compared to a lot of technologies in the past, this is moving at a pace and a speed that is much faster than what we've seen historically, and I-- which is good as well. Because if you look at where we are as an inflection point in the industry, electrification is moving at a very rapid speed.

And so the two actually align very well together. And it's not like this hasn't been done before in the industry. I gave some examples at the beginning of what we've done on the transmission side, which I would say is a lot more mature. It's not AI model-specific. But it's models and data sharing.

And so yeah, I would say, especially in here, the more we can get to open standards, I think this is just going to be a key accelerator for AI in the industry. And so we should leverage the institutions that are out there today. So we think about EPRI or IEEE or the key institutions that have been fundamental to moving the industry over many, many decades.

And so we should leverage the partnerships we have there to be able to do that. And going back, we shouldn't do this on our own. We should accelerate innovation through data sharing, model sharing as much as we can. And all these mechanisms are fantastic. It's just a matter of getting everyone to the table and doing that.

And talking about that precisely, we're proposing to leverage the concept of a playbook. Many utilities, including Bejjani, we use playbooks in order to accelerate results and set the rules on how we want to interact and how we're going to get things done. So again, this question is for Brenden and Jon Eric.

We are so early in evaluating what data model sharing will look like. If there were a generally agreed collaboration playbook, what would that playbook include? If we can put chapters, if we can put a list of things that the playbook should have, what would you toss there?

Wow, that's a long list. I'm probably going to go too tech here on this. I think that when you're explaining models in general, whether it's a fellow collaborator, you need to be clear about the assumptions you're making. You need to be clear about the data, how it was collected, where it came from because we've all made this mistake. We go in. And we start using data. And then we realize that column heading didn't mean what we thought it meant know.

Yes.

So giving the history of the data and the background of it, it's important to the extent you can. I mean, sometimes you haven't done as thorough a job maybe as you would like on exploring a range of algorithms because we learn from our mistakes as much as we learn from our successes. And so if you're putting together documentation on what you've done-- and you say, these are all the algorithms we tried. And this is why they didn't work as well as this one did.

And maybe it was limited by our assumption or the data that we had on hand. I was going to say one thing that Brenden mentioned is sometimes the best learning after first producing a model is you know what data you'd like to have for the next time. And so that kind of documentation, that allows others to leapfrog over each other as we collaborate together. So that'd be some key headings I look for.

Excellent. One minute for a final thought before I go to the next poll.

I'm sure it could be a big list as well. I think one thing we might want to do as an industry as well-- and going back to what Marc said around health care, I think we should also look at other industries and what they've done. Other industries that are a lot more mature in this area have similar characteristics or constraints and see what they've done and actually incorporate that into our learning here.

But I agree. I mean, data is so key to this. And as we think about things like federated learning, then you're like, if someone's training a model here, there's going to have to be some rigor around how that was trained. And obviously, getting to that sort of world building trust across the whole ecosystem is going to be key. And so I think that ultimately will drive that playbook and what we're going to need out of it.

Excellent. Well, we want to end this asking you a question again. We have another poll. And the poll is-- oh, sorry. Oh, here it is. Let's do in the right order. If you can please provide the name of your company or organization, scan this QR code, and provide your information so we can be in touch in the future, that will be appreciated.

And the last poll, we want to ask you the question of, do you want to be part of this journey? Do you want to be part of this playbook and start working together for a better future, where we can leverage these technologies in a responsible way but also accelerating results?

And here is the question. The development of a collaboration playbook could be a good first step towards data and model sharing. If you are interested in participating, please share your contact information with us through this QR code. And I will leave this in case you want to check your time to please be part of this.

Thank you so much, panel, for this wonderful, insightful conversation. Very much appreciated. And to all of you for coming to join us to this today. Thank you.

[APPLAUSE]