Silicon Flatirons
KNOW WHAT'S NEXT.

Transcript

Artificial Intelligence, Autonomous Systems, and Law Conference

March 7, 2025

- Timestamps correspond to videos published at: https://www.youtube.com/playlist?list=PLTAvIPZGMUXPsH2vPokwM-YYtH9naGQVl
- Visit the event archive at https://siliconflatirons.org/events/2025-artificial-intelligence-conference-2025-03-07/ for more information on the event including the event brief, agenda, speaker profiles, session descriptions, and more.

Contents

Published July 18, 2025

Silicon Flatirons Center for Law, Policy, and Entrepreneurship
University of Colorado Law School
2450 Kittredge Loop DR, Boulder, CO 80309

Keynote: Professor Harry Surden

https://youtu.be/OX7hfzBT40c?si=8FBFzfzgVz6maPpk

[00:00:00.29] DAYTON FIDDLER: Good morning, everyone. My name is Dayton Fiddler, a current 1L here at CU law. And thanks for making the trip to CU for this amazing event. We're thrilled to have such an incredible gathering of legal scholars, technologists, policymakers, and industry leaders all coming together to discuss one of the most critical challenges of our time-- how we navigate the legal, ethical, and societal implications of increasingly autonomous AI systems.

[00:00:24.32] I have the pleasure of introducing our keynote speaker, CU Law Professor Harry Surden. Professor Surden is on the faculty of the CU Law School, and is also the director of the Silicon Flatirons Artificial Intelligence Initiative. He is also the associate director of Stanford University's Codex Center for Legal Informatics.

[00:00:43.37] As a professor here at Colorado Law, his research and scholarship has been instrumental in shaping how we think about AI's impact on legal systems, regulation and governance. Please join me in welcoming Professor Harry Surden.

[00:01:01.20] HARRY SURDEN: Thank you so much, Dayton, for that kind introduction. So I want to welcome you all to the Silicon Flatirons center here at the University of Colorado Law School. It's really a pleasure to be here, and I'm really delighted you can join us for this conference on autonomous systems and AI autonomous agents.

[00:01:23.11] As Dayton said, I am a law professor here. I'm also the director of the Silicon Flatirons Artificial Intelligence Initiative and also associate director of Stanford's Codex Center. Before we begin, I want to say a little bit more about the Silicon Flatirons Center, what we do, what makes us special.

[00:01:45.67] So the Silicon Flatirons Center is a technology law and policy center housed at the University of Colorado Law School. And what we do is we initiate, sustain, and elevate public conversations about technology, law, policy and entrepreneurship and inspire, prepare and place students in these important areas.

[00:02:06.26] So I think this description kind of encapsulates a lot of what we do. The first thing I want to point out is that what's our focus? Our focus is on technology, law, policy and entrepreneurship. So these are our subject areas that we focus upon.

[00:02:22.55] Another word in there I think is important is public. We have a public-facing mission. The reason we have you all here today is our desire to educate the public and to discuss leading edge issues to make our society a better place.

[00:02:39.77] And then finally, our most important focus are our students. The University of Colorado is an educational institution, and our main goal, in addition to helping the public, is to educate our students and get them out in the world, get them jobs that are making a difference in society.

[00:02:59.14] We have a particular focus that is really interesting and somewhat unique in academia. Our particular philosophy is that when we have these important discussions about emerging technology like AI or autonomous vehicles, it's really important to involve multiple perspectives-- perspectives from government, from the private sector, from the nonprofit sector, and from academia.

[00:03:26.30] And this is fairly unusual in academic settings. Typically in academic settings, it's a bunch of professors talking to other professors, which is fine, but it's missing a whole lot of perspectives. And in our environment, we want to bring in the private sector, businesses, and organizations that are doing interesting work, nonprofits which are fighting for human rights and have a mission to improve the world.

[00:03:53.68] We also want to bring the government expertise, recognizing the important role of state, local, and federal government. Last but not least, professors and academia, those who are spending a lot of time thinking about these issues. And we find that when we bring all these perspectives together, we get some of the most fruitful discussions that can make a difference in society.

[00:04:16.43] Another perspective we have is that we try and inform the public about the state of technology law and policy today, what's going on to the best of our ability, but also we try and peer around the corner a little bit and be humble, but also trying to extrapolate out from current trends what's likely to happen in the near future.

[00:04:37.83] So we'd like to thank-- our motto is "know what's next," and we kind of peer around the corner a little bit. And before we start, I have a whole lot of thank yous I want to give. It took a tremendous amount of work to put this conference together.

[00:04:53.58] First and foremost, I want to thank our amazing Silicon Flatirons team. We have Sarah Schnittgrund, Christine McCloskey, Shannon Sturgeon and Nate Mariotti who worked incredibly hard to put this conference together. So please join me for round of applause thanking them.

[00:05:10.51] [APPLAUSE]

[00:05:14.34] I want to thank the rest of the Silicon Flatirons team, our director, Brad Bernthal, who unfortunately couldn't be here today, our faculty directors, some of which are here, Dale Hatfield, Keith Gremban, who will be on the panel, Margot Kaminski, who is supposed

to be here, unfortunately but had to bow out because of illness, and several other faculty directors, including Blake Reid.

[00:05:43.20] Also our faculty fellows, including Melanie Kay, Gabrielle Daley, and of course, our founder, attorney General Phil Weiser. We would not be here without his efforts, who started this center many years ago. I also want to thank the CU Law Dean, Lolita Buckner Inniss, who also, unfortunately couldn't be here today, and the tremendous CU law staff and faculty.

[00:06:10.55] We have several faculty members here today on our tremendous law faculty. Our expert panelists who have flown in from far and wide, including as far as D.C. and North Carolina, various places to share their expertise.

[00:06:29.03] All the students-- there's a number of students who worked really hard to volunteer and make this happen. And then finally, all of you. I want to thank all of you for coming, both you and the audience and the many, many of you who are online watching this remotely.

[00:06:45.65] So we're really grateful for the Silicon Flatirons supporters. Without your support, we couldn't put on programs like this. So we have many individual supporters, as well as many law firms and businesses and nonprofits who contribute in various ways. And it's those contributions and those support that allow us to do the work we do, so we're very grateful for your support.

[00:07:10.57] So let's talk a little bit about the conference today. We're going to have three panels. And of course, this is going to be focused on autonomous AI agents from three different perspectives. I'll be giving the opening keynote.

[00:07:24.98] Now I'll say, you'll be seeing a lot of me today, and not because I have the only voice, but largely because we've had a lot of unexpected illnesses and last minute cancelations, and I'm pitching in here and there for a few people who had to back out at the last minute due to family illnesses.

[00:07:43.58] So if you see me in multiple roles, just know that I'm pitching in there. So in addition to being doing the keynote, I will also be moderating panel one, which will be Developing Autonomous Agent Systems. I'll also be moderating panel two, the Policy, Legal and Regulatory Issues in AI Software Agents.

[00:08:05.43] And you'll have a reprieve for me with panel 3, which will be moderated by the excellent Keith Gremban, which will be studying autonomous vehicles-- physical autonomous systems as opposed to autonomous software agents.

[00:08:20.49] So in a moment, I'll start the keynote. I'll leave several minutes in the end for Q&A for the audience because I know that that's

often one of the most valuable aspects of our discussions here. So without further ado, let me jump into the keynote.

[00:08:40.35] And our keynote address is going to be looking at largely what's gone on recently with artificial intelligence generally and then autonomous systems specifically. And we'll be defining a lot of those words. I know a lot of you in the audience are not necessarily technology specialists or even legal specialists. I'll be trying to bring us all up to speed on the same page with some visual demonstrations of what's going on.

[00:09:12.28] So my talk here today, I'll be giving you an overview of artificial intelligence, just letting you know what's going on, talking about the exciting changes that have happened in the last two to three years, because there have been dramatic changes that I'm sure you've all heard on the news or maybe even have experienced yourself.

[00:09:30.37] Then I'll be talking about, in particular, focusing on autonomous AI agents. What does that mean? How is that distinct from AI generally? And then we'll talk about some of the unique benefits and concerns raised by this new era of autonomous software and physical agents. So let's just start with a level-setting discussion-- what is artificial intelligence or AI generally?

[00:09:58.46] There's no one definition that everyone would agree with, but I find this definition useful. And AI is enabling computers to solve problems, make predictions, answer questions, or generate creative output, or make automated decisions on tasks that we normally associate with human intelligence.

[00:10:18.23] So there's no widely agreed upon definition of human intelligence either. But generally speaking, we associate human intelligence with a variety of higher order cognitive or brain skills. These include things like visual spatial processing, language understanding and creation, abstract reasoning, problem solving, planning, learning, metacognition, for example.

[00:10:44.02] So anytime we take one of these tasks that when humans do it and we associate it with some sort of intelligence, whether it's playing chess, which involves planning and reasoning or doing math or physics problem, which involves some sort of analytical planning, abstract reasoning, philosophy, creation of documents, lawyering, these are all tasks that we normally associate with this abstract idea of human intelligence.

[00:11:11.51] If we can get a computer to do these same tasks associated with intelligence, those are AI or artificial intelligence tasks. So let me just give you an example here. We think lawyers are intelligent. I know they are because I teach them for a living.

[00:11:27.46] So we normally associate the kind of thing that lawyers do, which are things like legal document creation or reading or legal

analysis and planning or legal problem solving, client counseling with intelligence. It involves higher order abstract reasoning, planning solving problems.

[00:11:45.99] So here is an example of a tool many of you have used, which is ChatGPT-- one of the earliest versions of ChatGPT, GPT4, and I'm asking it to write a merger agreement, which is a type of legal document that lawyers can perform, which involves merging two companies.

[00:12:07.70] And for the first time in 2023 and now subsequently, these systems can generate highly competent first drafts. And we'll emphasize in this talk, they could never do that before. Prior to 2022, this was an impossibility for an AI technology. So we had a huge leap in AI capabilities since about 2022, and I'll show you exactly what happened there.

[00:12:37.23] A few caveats-- I'm going to use words like thinking, reasoning in the context of AI systems, but I want to emphatically emphasize these are not human like systems. These are AI systems of today like ChatGPT or Claude can simulate ever more convincingly human output and emotions, but decidedly, they are simulating using statistics and probability.

[00:13:04.13] They don't have emotions, they don't have inner human experiences. And it's easy to anthropomorphize them and think that they are human because they produce ever more convincing output. And in fact, we anthropomorphize them by using words like thinking, reasoning, acting, deciding, but just remember, if you hear me saying things like that, these are metaphors. They're simulating useful human output statistically.

[00:13:34.16] We know that the way they operate is very different from the way human cognition operates. OK. So how did we get here with AI? So AI has been around for a long time, and I won't go through the whole history of it. What I want to emphasize is kind of the era between 2010 and 2022-- so there was a sort of big move in AI-- and everything that's happened since 2022.

[00:14:04.30] So we had very capable AI systems prior to 2022, but they were very much narrow or domain specific. So what do I mean by that? Most AI systems prior to 2022 had one particular job and could only do that one job.

[00:14:22.54] So, for example, a narrow AI system of that era might learn how to play chess. Such a chess playing system could do it at a superhuman level, but it couldn't for example, write philosophy or write a poem or draft a legal document.

[00:14:38.02] We saw AI systems that could engage in prediction. Who's the most-- you know, what product is most likely going to be useful for an Amazon purchaser given their past Amazon purchases?

They did things like sorting email by spam and not spam. So these things got very capable, but decidedly, they were narrow in focus.

[00:14:59.38] They could only do one thing and one thing very well. And we saw these kind of narrow systems proliferate in health and in law, doing things categorizing documents as likely relevant to a lawsuit or irrelevant, or in medicine, looking at various test result numbers and making predictions about various diseases.

[00:15:22.45] There was a big step change in 2022-- and this is the release of ChatGPT, which I'll talk about. And what happened was they mostly cracked the code of human language. So one of the things that AI systems struggled with prior to 2022 was reasoning and understanding ordinary human language.

[00:15:43.32] And they largely cracked that, along many other obstacles which I'll talk about. So this was a big deal. So I like to-- although AI has been around for a while and has been involved in our lives for at least 20 years, only in the last two to three years have we seen super capable AI systems that could operate within domains of knowledge and language, a point I'll emphasize in a second.

[00:16:12.90] And this came about with a large leap of AI capabilities in late 2022, November 2022, with the release of ChatGPT 3.5 by OpenAI. So there was good reason there was a lot of excitement because through this. I've studied this area for 20 years, and this was a huge advancement.

[00:16:34.98] So although there is a lot of hype, in many respects, it is due because there was a large leap between the capabilities before 2021 and after 2022. Kind of a step change. And as I said, earlier AI models prior to 2022 very much struggled with a variety of important tasks that we associate with human intelligence-- understanding ordinary language.

[00:17:00.07] They just basically-- you couldn't type in arbitrary text to AI models prior to 2022 and expect to reliably get coherent output. We couldn't expect reliable common sense reasoning. We never saw arbitrary problem solving.

[00:17:18.12] So here's an example of a model just right before ChatGPT, which was the state of the art. I've been testing these for years and I would ask them common sense question, these old AI models that could do things write poetry.

[00:17:31.75] I had asked them kind of silly things that a toddler would know the answer to. How many legs does an apple have? And a human toddler would laugh and say, apples don't have legs. But the AI models of that time would confidently tell you things like an apple has four legs.

[00:17:47.10] And they would make sort of all sorts of common sense reasoning errors to the point where nobody really took them seriously

as artificial intelligence because they weren't that smart. The outputs they were producing were not that useful. And I have all sorts of other demonstrations for how these earlier models couldn't produce legal documents coherently.

[00:18:07.20] So it was the ChatGPT moment in November 22 that changed all this. There were a bunch of technical innovations that I won't go to. ChatGPT 3.5 solved this issue and suddenly became broadly useful. So I immediately tested, of course, ChatGPT when it came out with the apple question and a bunch of my other common sense reasoning questions.

[00:18:30.52] And it obviously aced it the way a human would do, and this was the first time an AI system could do this, and many other common sense reasoning questions. So this is what got AI researchers so excited. We never saw this before in AI systems.

[00:18:47.90] And because of this is, this is what has allowed AI systems like ChatGPT, Claude and others to take on knowledge work. We never had systems that could escape beyond the narrow domains of playing chess or spotting financial fraud to a broader category of knowledge work. And the reason is, as I mentioned, is because they basically crack the ability to analyze and understand language.

[00:19:17.63] And I'll talk about that in a second, but this was the first set of models capable to reason and problem solve for any problem you could give it. Now, to be clear, it didn't always get it right, and it still doesn't always get it right.

[00:19:34.44] So this is a mistake people make. These systems occasionally make hallucinations. They solve problems wrong. But the main point to emphasize is the earlier models before 2022 couldn't even do it at all. So just getting systems that could sometimes solve problems on their own was a giant leap in capability.

[00:19:55.32] And as I'll show you, they've only gotten better and better every year. So what you may know about models from the headlines from two or one years ago are no longer true because the models keep getting better, so a lot of the knowledge is outdated.

[00:20:10.08] That said, I want to emphasize these systems do make mistakes, and I'll keep emphasizing you need to double and triple check their work. I use them frequently, but I double and triple check their work. So after November 2022, just a few months later, GPT 4 was released and it improved on GPT 3.5.

[00:20:32.37] So the systems immediately-- they went from getting good to really good in a very short amount of time in 2023. Now, as I said, what this enabled is the accessory of AI to knowledge work. And this is really crucial because language, which the AI systems can now "understand," in quotes, analyze and produce.

[00:20:54.06] Language is foundational to how our society works. It's fundamental to how we communicate knowledge and facts. It underlies most disciplines such as law, medicine, humanities, science and engineering. It enables reasoning, debate and problem solving. It preserves history and culture across generations, facilitates human collaborations at scale.

[00:21:15.43] So once you crack language, you crack everything. So now these AI systems are no longer narrow. They're broadly applicable. The same AI system that produces competent results in law can produce competent results in medicine, in philosophy, in engineering. Competent, I say not necessarily excellent results, right?

[00:21:37.70] Although they're getting better, they make mistakes, and in many ways, they're still much worse than human experts. But even getting competent results is a remarkable achievement, and I expect them to continue to get better.

[00:21:52.22] So economists say that modern AI systems like ChatGPT are general purpose technologies, meaning they apply across broad domains, almost basically every domain of human inquiry, because they're trained on language, and language spans the entire of human inquiry.

[00:22:11.62] These things can say things that are fairly sensible just about every topic. So this is why this is really interesting and why I always make the emphasis that you have to distinguish AI systems before 2022 and after 2022 because they're very different.

[00:22:32.35] So where are we today? Today, we have advanced proprietary what are known as frontier models. These are the models trained by the top level labs that are performing at the best. Some of the top labs that you may have heard of or companies are OpenAI, which their latest most advanced models are GPT 4.5, which just came out last week, which I was honestly a little disappointed with.

[00:22:59.74] But they also have more advanced reasoning models, which I will talk about. Anthropic is another big player, which has a model called Claude. Their most recent model is 3.7 Sonnet. Equally good as ChatGPT.

[00:23:15.04] Pretty good, but slightly behind in my opinion is Google, which has their Gemini series. And then, surprisingly, out of nowhere, X.AI Musk's must company produced Grok 3, which I've been using. It is actually quite good. A lot of people were skeptical, but in my opinion, it is a top notch model.

[00:23:36.86] So those are-- I'm sorry. I didn't mention-- those are the proprietary models that you have to pay for. There's also a bunch of open weights models. These are not quite open source models, but these are AI models produced by companies like Meta. Or you may have heard of DeepSeek.

[00:23:53.21] They produce these models and then they give them out for free to the world and you can do what you will with it. You can specialize them, you can customize them. These have been really interesting because when I started with the large language models in November 2022, the open models were way behind-- two years behind in terms of capabilities, then they became a year behind. Now they're about six months behind.

[00:24:17.67] So the very, very best frontier model that costs billions of dollars, the free model is only six months behind and the gap is closing, which is kind of a fascinating scenario. So if you heard the whole excitement about DeepSeek from China, that's demonstrating they were able to produce a model that's only about six months behind the top models, and they gave it away to the world for free, which is kind of fascinating.

[00:24:45.09] So another thing to take note, that there's a lot of AI systems that are built on top of these frontier models like ChatGPT and OpenAI. So in law, we use systems like LexisNexis and Westlaw. These are specialized legal reasoning systems. But on the back end, they go and they talk to these frontier models.

[00:25:08.08] So in other words, Lexis and Westlaw don't make their own AI models. They use the best in class models produced by other companies and they build on top of it. These are sometimes called composite systems. There's a lot of these out there. They're really good. They often will provide security and privacy guarantees that the models that you might be familiar with, the ChatGPT, don't necessarily provide.

[00:25:35.03] And then there's a bunch of other AI models that are specialized for research, medicine, self-driving cars, what have you. So AI keeps getting better. So on the one hand, some people are a little disappointed with the rate of AI since 2022.

[00:25:51.71] So given the giant step change with the release of ChatGPT and the GPT 4 in six months, a lot of people thought AI was on a skyrocket curve, and so much so that several non-profits in March of 2023 said, pause AI research for six months. It's moving too fast. We're going to have human level AGI by the end of March. That didn't happen.

[00:26:17.71] But it has been steadily improving. So don't quote this number because I just made it up, but in my informal estimation, AI is getting about 10% better each year than it was the year before. And this is based upon various internal tests and looking at benchmarks, but that's not a real number. That's just an internal estimate on my part.

[00:26:41.20] So what does that mean? It means, well, every year, the error rate and the hallucinations go down 10%. It can solve 10% more problems than it couldn't solve before. It's getting a little bit more

accurate. But remember, these increases are compounding, meaning that every 10% is 10% better than the year before, which is 10% better.

[00:27:00.44] So while it hasn't skyrocketed, over the 10-year period, you can expect the improvements to aggregate significantly, even if they move incrementally from year to year. So that's an important fact to keep in mind. They still make mistakes. You need to double, triple, quadruple check them, especially if you're a lawyer in law. Do not turn in a brief created by AI without checking it, please.

[00:27:25.61] And we have a whole new paradigm. So while it seems that the paradigm of making the models bigger and bigger has sort of ended with GPT 4.5, there are other paradigms for making these systems smarter. Now we have AI systems that think and reason before they answer.

[00:27:44.48] And one of the most exciting aspects, of course, that we're going to see is the rise now of autonomous AI agents. So let's focus on that. Where have autonomous AI agents come from? Well, basically, autonomous AI agents have come from the fact that these base models, these AI models, have gotten a little bit smarter each year and can do a little bit more that they couldn't do past year.

[00:28:10.23] So now they're at a moment where they can do enough things in a row that we've enabled autonomous AI agents. So what is an autonomous AI agent? Let me first show you and then I'll define it.

[00:28:23.86] So what we have here is something from OpenAI called OpenAI Operator, and this is a tool, a system which can use a web browser. It can go to any web page and can do anything you or I can do on a web page. It reads it and then it kind of figures out what to do.

[00:28:44.20] So here I asked it to register me for this conference here, and I'll show you what it did. So I launched it off. This is it on its own. It figures out to go to the web browser. It found the web page. It found the right conference. It found the registration form, it clicked it. And then this is an important moment here.

[00:29:07.66] So now it's kind of stuck. So it has to ask me, tell me about your-- it's only stuck because it doesn't-- not because there's an objectively right answer. It needs to know for me, am I attending in person or am I attending online? So I have to provide some answer, which I do.

[00:29:24.64] So I'm giving it some feedback. It continues on. I tell it "in person," and it figures out the right button to pick based upon in person. Again, it asks me for more information. I tell it that I'm faculty. I give it my name, my email. It populates all the right fields.

[00:29:43.72] So this is amazing, and this is an example of an autonomous agent. It is out there, figuring out what are the right buttons to do. So there's not great definitions of autonomous agents,

and we're going to talk about this on the first panel. But there are some features that they share.

[00:29:59.83] And some of the features that they share are-- these are AI systems where you give them a high level goal. As opposed to having the human click and do everything for the system or give it steps, you give it a high level goal, which is register for a conference, or book me a flight, or buy me some NVIDIA stock, or drive from Boulder, Colorado to Denver.

[00:30:24.41] And then the system itself, the AI agent itself has the ability to figure out what the user is asking for. So first, they interprets the request, then it plans it. It can figure out, what necessary steps do I have to get to in order to register for the conference?

[00:30:44.00] And then it has some ability to make its own decisions about its own actions after understanding its environment. So we saw Operator doing this. It has access to software tools like web browsers. So it had the ability to plan, it figured out what it needed to do in order to find the Silicon Flatirons website.

[00:31:08.50] It figured out that it needed to do a search, find the website, and then it had to react to what it saw. It had to react to the web links. It clicked on the page, then it had to read the page, figure out which conference was this conference, click on that, figure out what the right button was to register.

[00:31:24.78] It's able to click on that. And then it reached a point where it couldn't do anything more on its own because it didn't know my personal preferences, my name, my email, et cetera and had to ask the human user. So this is an example of autonomous system because it is making decisions about its own actions, as opposed to having the human tell it what to do.

[00:31:47.75] So these autonomous systems range on a spectrum, from fully autonomous, where you can send a software agent out and it'll do a whole bunch of stuff and not really come back and ask the human, to semi-autonomous like we saw where it'll do some things, not have enough information, have to come back to the user.

[00:32:09.47] And there are various levels of autonomy. So here's another example of an autonomous agent. You might have heard that Googl-- sorry, OpenAI released something called Deep Research. So this is a very thorough research tool. You ask it a research question like tell me about autonomous AI systems.

[00:32:30.76] It'll then ask for clarification. You give it some clarification about what you want, and then it goes out on its own, does all the research, uses the AI system itself to plan, to figure out what it it's reading about, and then it comes back with the report. So this is the state of the art. This was just released a month ago.

[00:32:57.48] And it's important to emphasize that the reason I told you about all the AI before is that these autonomous AI systems are built upon those systems like GPT 4 or Claude underneath. So they use those systems to understand the user aspect, to make a plan about what to do based upon the user's high level goal, and to interpret the data that they're getting, whether they're web browsing.

[00:33:25.99] There are different types of autonomous systems. One thing I want to talk about is the difference between physical agents and software agents. So physical agents or physical systems are things like self-driving cars or robots that operate in the physical world.

[00:33:42.47] And they operate on a similar principle in terms of planning, reacting to the environment, but there are very different considerations because there are people and buildings and things like that out there in the world that they can harm or bump into.

[00:33:57.50] But it's a very similar idea, as opposed to software AI autonomous agents. These are systems like I showed you. They don't operate in the physical world. They operate in the digital world and have at their disposal, a series of software tools. These are things like web browser or maybe coding environments like Python.

[00:34:18.29] They might have things like-- these are called APIs, which are ways of accessing data or functionality on particular websites or tools for doing math. And the systems themselves in the software agents know about these tools and they decide when and if they should use those tools on their own.

[00:34:37.16] So no one's telling them, here's a search, go to Google and do that search and look at the results. It figures out that it doesn't know enough information and it needs more context. So it decides and it knows that it can use a tool. So those are software agents.

[00:34:54.08] Let me show you an example of a physical agent. So this was in San Francisco in the fall. I took a Waymo self-driving car and this was it pulling up to pick me up. If you go to San Francisco and several other cities these days, these are all over the place. And this is me, my video from the back seat here.

[00:35:13.69] WAYMO AI AGENT: Just give us one minute to cover a few driving tips. This experience may feel futuristic, but the need to buckle up is the same as always, so keep your seat belt fastened, please.

[00:35:26.52] HARRY SURDEN: It did a phenomenal job. I felt extremely safe. There was a lot of weird things that could have happened that it avoided. People walked in front of the car and it didn't present any problems. So what I want to talk-- why are we talking about this today, it's because there's been a capability jump.

[00:35:47.04] When you increase by 10% and then 10% again, we're at this moment now where AI systems can go out and do multiple things

correctly on their own. So just last year, AI systems, these autonomous agents couldn't do that.

[00:36:00.64] So when I showed you the web browser, very routinely, you would ask an AI system, go register for a conference and then it'd gets stuck. They get confused by the website, or they'd enter in your name in the zip code form, right?

[00:36:15.07] They make very basic mistakes. Now we're at the point where they're just a little bit better, but that little bit is enough that they can now do 10, 20, 30, 40 steps on their own and get it all right. So now we're on the verge of these systems being able to do lots of things out there in the world.

[00:36:34.42] Here's another example. I mentioned Grok 3 doing deep research. So it's pretty easy at this point to replicate this kind of autonomous agent. So we're at the threshold where these things are coming because it's no longer the future. This stuff is here.

[00:36:51.49] Well, I want to leave time for Q&A, but I just want to emphasize this is an important conversation to have now, because now the capabilities are here and they just weren't here as recently as last year. So let me pause for a moment and take any questions from the audience. And our tradition is that the first question always goes to a student. So we have our first question over here.

[00:37:25.85] STUDENT: Hi. I just had a question about ChatGPT's newest model. I've noticed it's been using emojis and trying to convey emotion. Is that like a design choice?

[00:37:35.10] HARRY SURDEN: Yes. So a lot of the model makers have realized that people like AI systems with pleasing personalities and that seem friendly. There's actually a great term called AI sycophancy. And have you ever typed in some idea and the AI system say, that's a great idea. And they're like, I am a genius.

[00:37:55.26] Well, don't get excited. Everyone's a genius according to the AI. So the AI system will tell everybody their ideas are great. And the reason is it's trained on human preferences. So what they do is when they're training these models, they put two answers side by side, and they have human pick the better answer. And which answer is the human going to pick?

[00:38:15.86] Well, of course, the answer that tells them they're a genius versus the answer that doesn't. But that being said, they also are training these models to be more friendly, and that's a great example. Yes. Oh, and please wait for the microphone.

[00:38:33.10] AUDIENCE: Harry, great presentation. My question is about ethical AI. Say you're in that Waymo car and you hit a pedestrian. Who's liable for that accident?

[00:38:41.36] HARRY SURDEN: Yeah. This is a great question, and these are the kinds of questions that maybe the third panel, I don't know, will be addressing. But these are the questions that are now arising because these systems are out there that we have to grapple with as a society.

[00:38:57.68] So the answer is our society has to figure it out. There are any number of choices it could make. It could put it on a self-driving car company. It could maybe put it on the driver-- or the passenger of the car. Depends upon what happens.

[00:39:14.13] But other things to consider, though, this is something I always make-- we should also consider the error rate that human drivers have and make sure that if it is the case, there's a lot of evidence that self-driving cars are safer under certain circumstances, that we're not unduly inhibiting technology that reduces the overall level of accidents. But it is a concern that we need to think about. Yes, in the middle. And please wait for the microphone.

[00:39:49.01] AUDIENCE: Hi.

[00:39:50.39] HARRY SURDEN: Yeah. That's OK. Yeah, great.

[00:39:52.89] AUDIENCE: I have asked ChatGPT 25 times now not to ask me follow-on questions. And I've asked it why it asked me follow on questions and it told me that it's programmed to further engagement. And it's told me 25 times that it will not ask a follow on question, and it continues to do so. Is there a way to actually get it to stop doing that and what is causing this?

[00:40:14.95] HARRY SURDEN: Yeah. So that's a great question, and that is both symbolic of the strengths and weaknesses of these systems. So these systems are trained as one-size-fits-all systems. And one size fits all fits nobody, right?

[00:40:28.78] So this is a good example where your particular preferences don't align. And also, these systems don't announce their weaknesses the way previous technologies. So they seem omniscient and perfect, but they're not and they make errors.

[00:40:43.87] There are ways that you may be able to do it through customized instructions. If you pay for the certain model. The free models, you can't, but on the paid models, you can give it customized-- and there they will usually but not always follow your instructions but--

[00:41:02.30] AUDIENCE: [INAUDIBLE]?

[00:41:03.22] HARRY SURDEN: Oh. Well, I'm happy to talk about it afterwards. In the middle in the black had a great question.

[00:41:13.72] AUDIENCE: So within the last year-- so I remember reading that there was an issue with ChatGPT having trouble with basic

arithmetic and mathematics, but then as they improved that skill, that it seemed skills in other areas were declining.

[00:41:30.05] So it got smarter in one area, but then it started to get dumber in other areas. And I wonder if you have any insight into that, and also reflecting on your statement that you felt this technology was improving overall.

[00:41:45.73] HARRY SURDEN: Yes. This is a great question. So you're absolutely right. The way the system works are not understandable even to the programmers who create them. They have trillions of these pattern matchers called parameters that work together to produce one word at a time, based upon the prompt and the previous text that it's produced.

[00:42:08.44] So first of all, that's not a great way to do math problems, is kind of by looking at the previous words. So one way that they've gotten around this is by telling the model, aha, when I encounter a math problem, let's use one of the tools that I have available to me. And for example, they have various math tools like software tools that it kind of outsources the math to a tool like a calculator almost that'll do it right.

[00:42:35.80] So that's one of the reasons it's approved. They also have improved the math reasoning by the old one word at a time method, which is a little less reliable. But to your basic question, which is when you improve something, these systems are so large and not understandable, you can accidentally reduce the capability of other things.

[00:42:57.22] So say you've trained it to be good at poems and then you give it a bunch of math problems, the math problems can accidentally make it worse at poems, and you've no way of doing that other than testing it out. So it's much more of an art than a science at this point.

[00:43:13.57] That said, they are getting smarter across the board. So, even though one model may have gotten slightly worse at making poems, probably that model is still overall better at doing poems than the model a year before if that makes sense. Yes. Down here, please.

[00:43:34.35] AUDIENCE: Thank you. Great presentation. So from what you said, I kind of understood that you think we're still in a linear progression, so to speak, 10% each year. A couple of days ago, OpenAI released a new product.

[00:43:48.99] For $200,000, you get an agent that performs at the level of PhD. Does that break the model? Are we having more of a exponential increase in capabilities now, or do you still consider that being part of incremental growth?

[00:44:05.20] HARRY SURDEN: Yeah. So what I was saying, it actually is exponential, it's that the improvement factor is 10% exponential. So

exponential is when you multiply each increment and linear is when you add. So if it were just adding 10% each year, there would be linear, but it's 10% on top of the 10% increase.

[00:44:25.63] That said, the thing we're-- I think what we're seeing is what is that year to year exponential increase? Is it 100%, which means it would double in capabilities? And that's sort of what we saw from 2022 to 2023. And now we're seeing 10%.

[00:44:43.45] But I'd say I don't know. I'm just sort of reflecting back. The AI labs have much better visibility on the future and if that's what they're saying, I'm inclined to trust them. What I'd like to do is think about what people were saying in the past, then see what has actually happened and go back and look, were their predictions aligned with what we now know it happened?

[00:45:07.33] And we know their predictions were wrong two years ago compared to the rate of-- they were predicting 50-80% exponential, and what's actually happened, in my opinion, is about 10%. But that doesn't mean that we could have another step change. So nobody really knows. Yeah. That's a great, great question. In the back there. And then I think we have time for one or two more-- one more question. All right.

[00:45:36.03] AUDIENCE: Yeah. One of the things that's been interesting to me, just a lay person, is seeing some of the focus go from the big frontier models to seeing like smaller, more specialized models that can also maybe be more efficient. And so do you think that that's a niche that smaller companies that aren't xAI or OpenAI could maybe fill and be more competitive in?

[00:45:57.64] HARRY SURDEN: Yeah, absolutely. Well, particularly with open weights models being only about six months behind the much more advanced models, a startup can take these open weights models and specialize them for various uses.

[00:46:14.27] I think that's a really important business model. And that's certainly been happening and going to be happening more. That said, there will always be use cases where people want the best and the smartest model and they're willing to pay for it, whether it's scientific research or business or finance.

[00:46:34.51] Unfortunately, we are out of time. However, the good news is we have a break. There's food in the room in the Bachelor Hall. We have about a 15 minute break. We'll be back here at 10:00 for our first panel. Thank you so much.

[00:46:50.35] [APPLAUSE]

Panel: Developments in AI Autonomous Software Agents

https://youtu.be/IX3ZMedXrbc?si=QBqfUzCm0MXVK36t

[00:00:00.89] DALTON MILLS: Good morning. My name is Dalton Mills, and I'm a first year student here at Colorado Law. I have the pleasure of introducing our panel on developments in AI autonomous software agents. Our panel today is moderated by Professor Harry Surden.

[00:00:16.32] And the members of the panel are Nathan Schneider, Lucas Thelosen, and Elana Zeide. Nathan Schneider is an assistant professor in media studies here at the University of Colorado. He's currently working on exploration of models for democratic ownership and governance for online programs and protocols.

[00:00:37.04] Lucas Thelosen was previously the head of products at Google for data and AI Cloud. He's now the CEO of Gravity Foundation, an AI powered data analyst. Finally, we have Elana Zeide who is currently an assistant professor at the University of Nebraska College of Law. And she recently published the Silicone Ceiling, How Artificial Intelligence Constructs an Invisible Barrier to Opportunity. This analyzes how AI driven assessments and hiring tools perpetuate bias and limit opportunity. Please join me in welcoming this exceptional panel.

[00:01:19.27] [APPLAUSE]

[00:01:20.26] HARRY SURDEN: Thank you so much for that excellent introduction, Dalton. I really appreciate it. So we have a really terrific panel here of experts. And we're going to dive into our first topic-- what has happened recently with autonomous AI agents. We're going to be doing a little bit of background setting. And I'll throw the first kind of series of questions out to the panel.

[00:01:44.23] First of all, what do we mean when we say autonomous agent? I talked about it a little bit. Is there a difference between autonomous system, AI, agent? So that's one idea sort of are there meaningful distinctions between them. And then also, what are some of the most significant real world advances that we've seen in AI agents? Anybody want to take a first crack at it?

[00:02:13.99] LUCAS THELOSEN: Yeah. [LAUGHS] Yeah, I can start. Actually, if you don't mind, I want to take a brief step back because someone had the question, why does ChatGPT not remember that I asked it. So I want to really briefly talk about training and learning and how these systems work.

[00:02:32.51] We are quite adamant, actually, in the technology space that we don't really train AI on your data. And so that's an important part, as you use some of the systems, that they're not necessarily

trained on your data. Now you want to read the fine print because some of them are. But a training run is only done every once in a while. And a training run costs millions of dollars usually.

[00:02:57.46] So when you type in that [AUDIO OUT] not make it, even if you consented to your data being used for training, it won't make it into a training run until way later. There are other things you can do in AI to actually have the AI customized to your needs. And that's when we get to agents. So in agents-- and if you're paying for-- I'm not representing OpenAI, I can't talk about the features. But you're paying for it, you have the option to select Preferences where you say, don't ask me follow up questions.

[00:03:30.07] And it's actually being parsed into your question every time for you. You won't see it. But on the background, that's what the system does. So the agent does not remember-- oh, sorry, the AI does not remember. It wasn't trained and it is not being part of the training data. But every time you ask the question, it is added back in to the prompt that you put in.

[00:03:51.37] So I want to be very clear. I think that's an important distinction between when is your data being used for training versus adding information to the prompt to do a better job with it. So just really quickly--

[00:04:03.81] HARRY SURDEN: No, that's very helpful.

[00:04:05.43] LUCAS THELOSEN: Yeah, I'm dealing with a lot of legal contracts. And all the time, it's like, are you training our data. And so I want to just take a moment to clarify the distinction there. So I work on our company, we have a system of AI agents. So that's quite sophisticated. So I'm going to go through the whole evolution of AI in essence there.

[00:04:27.60] When you talk to most systems now, it is probably an agent, meaning there is some wrapper around it. There's some background information that's being passed into in addition to the LLM that you're talking to. And many times, there's a component of autonomy in there as well, meaning it can do something without asking you again. It has the opportunity to make a judgment call on some things.

[00:04:56.00] And when we talk about intelligence, I want to be careful too, because intelligence can only mean something, in the space of AI, if there was enough data to train it on. So legal contracts, there's lots of them available, easy to train on. If I were to tell GPT 5 to pick up milk from the grocery store, it would not be able to do that. And it will be a long way until it can do that because nobody has documented the 250 steps you need to take to pick up milk at the grocery store.

[00:05:30.78] So it's very easy to train AI on things that are well documented. But I would caution to confuse that with intelligence.

There's a difference there. And I think as we think about what it means to be human and confusing it-- I don't know. I think it's good to think about the lines a little bit. I'll stop.

[00:05:54.29] HARRY SURDEN: Yeah. No, that's great. I just want to reflect a couple of the themes you raised. So one theme, which is really important, there is this common belief that when you enter data into ChatGPT, immediately it gets sucked up by the system and becomes part of the system. So if you type in your Social Security number, there's this belief, oh, it reads it. And now anyone else can see your Social Security number. But what you're saying, that's not the case because the data it's been trained on has happened months before, and it's not immediately added to the model.

[00:06:27.69] That said, it's important to know, for some of these open weight models out there, when you type things in to an open source model, it may become a part of some future training data set out there. A lot of these companies try very hard to scrub personally identifiable information. There are now AIs that can look at training sets and say, oh, this is a Social Security number, this is somebody's address. Let's nix this. But they're not perfect. So just be aware of that.

[00:06:57.28] Another thing I heard you say, which is really interesting-- the systems have evolved into composite or multi-component systems rather than, in their early days, you were interacting directly with a large language model, and it was immediately producing the next word or token. Now, a lot of what's happening is you have these systems that kind of intercept your query and kind of do things on the back end before producing information, adding context, adding information. And that's part of the agentic process. Yeah, great.

[00:07:30.61] NATHAN SCHNEIDER: I think it's important to just name where the agency and where the autonomy is and where it isn't. Because often, these kinds of terms are sold to us because a certain kind of thing is being emphasized. I mean, when we talk about agentic systems, these are systems that are able to operate across processes without a lot of intervention, without points of intervention that the user might otherwise expect to have. And autonomy has kind of similar meanings in this context. It's very much about the user experience that we're talking about.

[00:08:10.07] But I think it's also important, especially in a legal context, to talk about aspects in which agency is not just residing in these systems, but is still very much residing in other kinds of places. I mean, for instance, these systems are not autonomous from public infrastructures of water, of energy, and even climate. They are not autonomous from the processes by which those data sets are generated. They are dependent on the legal and perhaps a legal or illegal processes by which data sets are generated that were used for their training.

[00:08:53.17] They're not autonomous from the system of finance and venture capital that provides the capacity to do this kind of training and has its own incentive structures built in. They are not autonomous from the corporate control of the processes of fine tuning and alignment, which are the processes by which the trained models are tweaked and prepared for public deployment. And also, they're not autonomous in a kind of temporal sense from the future business models and future ownership and entanglements that they may have. We are still figuring out what the business models of these systems are. We're still figuring out what they might be used for.

[00:09:46.26] And so I think it's really important when we talk about autonomy and agency, to also have in mind entanglement. Where are the entanglements, even when we point to some aspects in which these systems are performing with autonomy and with apparent agency?

[00:10:04.83] HARRY SURDEN: That is a great point. So, of course, autonomous and agent is a very loaded word. We have those words in law which have their own distinct meaning, let alone in society. And there's a narrative that these are just systems on their own, doing their own thing, when in fact, they're part of a larger socio technological system that sometimes gets abstracted away. And it's important to remember that. It's a really important point. Elana.

[00:10:31.29] ELANA ZEIDE: So I have something that's perhaps between the-- maybe more concrete. So how I tend to think of it is the AI is the process that's running, that makes decisions, that parses language, sometimes output things. But the agents can interact with other aspects of the universe, by which I mean the web basically. They can schedule something, they can fill out a form, those sorts of things. So that's how I differentiate between it. I see it almost as components that can execute.

[00:11:15.27] HARRY SURDEN: So I like that. This is sort of emphasizing the interactivity aspect of agents versus some AI systems are less interactive with the world around. I think those are all very good definitions. Following up with that, where are we seeing some really interesting movement out there in the world on AI systems, and where are we also-- autonomous agents and where are we seeing limitations? And I very much like your point, Lucas, about the idea that where there's training data, you're likely to see these systems be able to simulate or replicate it. But there's a whole lot of stuff out there where there's not data. So they're likely to fall short in those areas. And it's kind of a profound realization that these things seem so smart, just because a lot of the stuff we need on a day to day basis are kind of basic and simple, and lots of other people have provided data for that before.

[00:12:12.41] LUCAS THELOSEN: Yeah, exactly. So the system we built consists of 32 AIs working together. So it's actually an interesting

example to talk about how we have autonomy in there. And it's funny. When you watch them work with each other and talk to each other, it's in language. So we can actually see it and we can read it. And sometimes they fall short on very silly things where it's like, come on, guys. And sometimes it's like, wow, this was substantial. This would have taken me 80 hours to actually work on, and you figured it out. So having these where is it really amazing us and where is it falling short--

[00:12:53.20] HARRY SURDEN: Can you recount the board meeting anecdote? I found that really amusing.

[00:12:56.66] LUCAS THELOSEN: Yeah, so our agent system analyzes data about your company. You buy it and you say, OK, tell me what's interesting, what's happening, what should I pay attention to as a leader in my organization. So in order to do this, we give it access to your data. It looks through it every day, all the time. And different agents play different roles.

[00:13:18.41] And what we have found is that they are really good if they have a specific role to play. So I'm not so worried right now about general intelligence where they can do everything, because actually when you give them a bunch of things to do, they get confused. But if they're really focused, they're good at it.

[00:13:34.04] So in our system, we have one agent that is very excited. So this agent has the role come up with new hypotheses to test and things that could be interesting. Another agent is responsible for questioning everything and constantly asking, how did you come up with this, walk me through the steps, let's double check the numbers again. So a QA agent. And in the early days of our system, we didn't have a manager.

[00:13:59.54] So it was a flat hierarchy structure. Everybody had the same say. And they discussed topics. And so the excited agent was like, this would be great for the board meeting. And all the agents went off about the board meeting and they spent a couple cycles, which is only a few minutes, but still a couple cycles thinking about the board meeting and how to make this great for it.

[00:14:18.13] And we're like, there is no board meeting. You got completely sidetracked, completely lost focus. And so we put a manager in place. So this manager's only job is to make sure that the tasks that are being executed are still aligned with the core tasks that we put in.

[00:14:33.53] HARRY SURDEN: Yeah, I love that because it shows how things could go awry. The autonomous AI agents are spreading rumors. And they pick up the rumors and encapsulate it. Great story. Nathan, did you have something to add about some of the limitations?

[00:14:47.65] NATHAN SCHNEIDER: I think it's an interesting reminder. I mean, just hearing you describe the way in which you're

orchestrating these systems-- and I wonder if this will change-- but how much it just actually resembles the reinventing of wheels. I mean, I've spent the last 10 years, for instance, studying the cryptocurrency world. And a lot of them, those communities started out with flat systems and were like, we can reinvent everything. And then they're like, oh, no, actually, we need to have a manager to make sure we stay aligned.

[00:15:15.42] And I think it goes back again to this reminder that in one sense, we are experiencing something very new in the world. And on the other hand, we are also re-encountering something that's quite familiar. I mean, the stock market is a kind of AI. Corporations are a kind of autonomous AI. They use us as their agents, and we perform tasks toward their ends. And so it's a reminder, I think, that things that we have learned in other kinds of domains may be more applicable here than maybe we think because flashy new technology appears.

[00:16:03.16] HARRY SURDEN: I love that. And ever more important are the social sciences and the humanities and not just the engineering as we grapple with these issues. Elana.

[00:16:12.93] ELANA ZEIDE: So what I'm seeing a lot of is agents that do very, very specific things and also things that are fairly menial. So when I say execute, it's like schedule a meeting. Well, we all have seen the scheduling platform where you just choose one and you click a button. I mean, clicking a button-- to me, that is not mind blowing, which is good because they're also not super high stakes decisions.

[00:16:42.50] And I've seen of that low level, and I've seen things where they're automating essentially the execution of a decision. So I do a lot of work in education and hiring, and I've seen things where they'll have advising or evaluating students or admissions, and they'll automatically go to the next level without a human in the loop. So that, to me, is really interesting. It's interesting in two ways, both as a positive as a negative.

[00:17:12.06] I think on one hand, there's been a lot of talk about having humans in the loop to review decisions. But when that happens, often humans are shown to be deferential to artificial intelligence. And as they do it over and over, it becomes like how we all are when we click the terms of service. You're just like, yes, yes, yes. And so in some ways, I think the use of agents in this way highlights the way in which it is still quasi automatic, that there really isn't a human in the loop in most circumstances, because it's not as efficient or they're not really exercising meaningful review of the opaque AI.

[00:17:52.49] I'm seeing it in that. And I do think sort of at that lowish level, if an AI miss-schedules you, you can figure it out pretty quick. You can fix it.

[00:18:03.71] HARRY SURDEN: Did you want to follow up?

[00:18:04.80] LUCAS THELOSEN: Yeah, no, I think that is a really-- you both actually brought up really great points on something that fundamentally was always there and actually becomes, once again, very important. I was invited to speak to the board of educators in Denver for the schools, for the school system, not their universities. And they asked, what skills are the most important going forward into the future.

[00:18:26.73] And I think things that we focused on in 1910, 1920, just very fundamental, can you pay attention, can you logically walk through this and see the flaws in what you see here. And so I think that is actually really what you need. You need to be able to pay attention and actually parse what this all means. And so I think, in order to do that with these large quantities of quite sophisticated texts, you need to be really on top of your game. You need to make sure you haven't done this for hours at end.

[00:19:04.30] But I think that would be a very, very important skill. And I have kids. So I'm a little worried about their ability to pay attention for a long time in the future. But I think that will be, once again, not a skill that is brand new. It has been around for ages. And it will be, once again, very important. So in the end of the day, power tool, for example, data analysis, our customers will be managers of AI data analysts. And as a manager, you have the responsibility to review things and be smart about it and pay attention to it, and not just rubber stamp things.

[00:19:40.34] HARRY SURDEN: Yeah, I think this is a great point about the value of human skills and human critical thinking become ever more important. So there's one story, which I think is overblown, where Sam Altman says that deep research, which I showed you from OpenAI, does some 5% of all human valuable labor, which may or may not be true, but I'm skeptical of. But that said, so deep research can go out. It'll give you a very thorough, pretty accurate 20 page report. But at the end of the day, what are you going to do with that information?

[00:20:13.13] NATHAN SCHNEIDER: Feed it into an AI.

[00:20:14.02] HARRY SURDEN: Right, exactly. If you have 100 reports, 1,000 reports, fundamentally same thing with your AI data. It's very good. But at the end of day, some manager has to have the mental capacity and the ability to understand and know what to do in the real world with that data. And if you haven't been educated sufficiently, then the AI is not all that useful.

[00:20:36.88] NATHAN SCHNEIDER: I mean, the question of agency applies as much to the machines as it does to us. What kind of agency do we actually want, and two, we actually hope to have in these systems? And again, I think that's an interesting question from a personal perspective from the perspective of what kind of people are we cultivating in ourselves. I'm working on an essay right now about

the practice of question asking in relationship to this. How do we cultivate our own curiosity in relationship to these machines that seem kind of boundless in their possibilities and seem, in some ways, to be effacing our areas of agency, not just labor, but areas of agency that we are used to humans having.

[00:21:27.40] But also, I just want to again point to those other areas of entanglement of these systems. And I think the question, for instance, of corporate agency-- who is responsible for and to whom are these companies accountable in their operations and their stewardship of these systems. And on the one hand, I think we've seen a lot of really interesting experimentation in this. I mean, the OpenAI nonprofit governance fiasco was an expression of the fact that this company was set up by people who thought, we need to do this stuff differently because the standard tech corporate model is not right for this, is dangerous for this technology.

[00:22:14.38] I mean, you got to recognize that impulse. That's a very powerful impulse that they set-- Sam Altman, whose job it was to create cookie cutter startups, decided to create a non-cookie cutter startup in the context of OpenAI. And then also, for instance, Anthropic has set up as a public benefit corporation. There's an interesting opportunity that these companies are pointing out that we don't know how to set this thing up right.

[00:22:45.80] One thing I've been studying is the way in which legacy unions in Europe, for instance, and legacy like agricultural cooperatives, which is the most powerful cooperative sector in the United States, are applying AI differently. And I think that's a really important question. What other kinds of corporate structures could we imagine? What kinds of other stewards maybe not corporate structures-- maybe it's something else. Maybe it's public AI. I think we need to be asking really intentionally now how we are leveraging our agency as a society, as individuals, as groups of individuals, as collectivities, as communities in relationship to these systems.

[00:23:33.18] I mean, it was striking in thinking about your tools where on the one hand, you're building these agents on one level, but on the other hand, they're all drawing on, they're attached to the APIs of a few big companies. So even in that context, when building these agents, there is a profound dependency on highly, highly centralized forms of power.

[00:23:58.43] LUCAS THELOSEN: Well, in some way, yes. In some ways, our system is built to change without any of them at any moment. So we were able to completely swap in an open source one. And so when I think about the evaluations that some of these LLMs like companies are asking for and are thinking about, I'm not so sure if that's true because it's so quickly replaceable by something else.

[00:24:26.15] NATHAN SCHNEIDER: Yeah, becoming a commodity.

[00:24:28.16] LUCAS THELOSEN: Exactly. And they're all trained on similar data sets. So it's like, I don't know if they hold all that much power over us. If OpenAI was to say, hey, you can't use our API anymore, we would not have a problem at all.

[00:24:40.07] NATHAN SCHNEIDER: Yeah.

[00:24:41.47] LUCAS THELOSEN: So I don't know.

[00:24:43.84] HARRY SURDEN: Elana, did you want to say something?

[00:24:47.02] ELANA ZEIDE: So slightly orthogonal, but agency. So one thing I think is interesting is the idea of agency in terms of how much we want to exercise it, people retain their own. But I worry that in some ways, people will have not a huge amount of choice in terms of agency about whether or not to use AI, what to use it for, in part because of competitive pressure or perceived competitive pressure. I have talked to some colleagues of mine who are like, is everyone writing their articles with AI, and I'm not. Tell me how to do that. And I don't think it's happening yet. But if so, he's the one that's doing it. But--

[00:25:34.66] HARRY SURDEN: Mine aren't good enough to be written by AI.

[00:25:39.05] ELANA ZEIDE: But no, it's like, the students in my class, they know some of their classmates are using things. And they think, oh, I better use that. If not, they're saving 10 hours, and I'm not. But the truth is that in many cases, some people will get advantages from it. And I think those advantages are not going to be equally distributed. So who can pay for the latest novel and who has access to either the people or the resources to understand how to use these tools well.

[00:26:14.38] LUCAS THELOSEN: Yeah, so on like one of the reasons I left Google and started my company was before I let product, I let the consulting arm. And in there, we went out to Bentonville in Arkansas for Walmart for millions of dollars and years at a time to build out quite sophisticated systems for them. And now we have the opportunity to take those best practices of how to do data analysis, optimize your supply chain, optimize whatever-- anything with optimize. That's, in essence, what we do.

[00:26:44.79] But we can bring it to a company for a fraction, a tiny, tiny fraction of what, a couple of years ago, a large multinational paid consultants millions of dollars for. So yes, access, absolutely. But now, it's a fraction of the cost of what it used to cost. So I can get a very average company a top notch data analytics team. And it costs very little. And that gets me quite excited because there are so many people out there that just can't afford to have the sophistication that Amazon and Walmart are deploying when they make decisions, data driven decisions. And now I can bring it to a very average company.

[00:27:23.94] NATHAN SCHNEIDER: Sorry, go ahead.

[00:27:24.66] ELANA ZEIDE: No, I was going to say, I think so much is dependent upon who you're considering as your user, whether that's-- yeah, because there's the people who won't have anything, but have this, which may be better than nothing. And then there's people who-- I think a lot of times, we all have an unspoken model in our head about the specific AI we're talking about and the context and the users. And it's good to say those explicitly.

[00:27:52.25] NATHAN SCHNEIDER: I mean, I really recommend folks who are interested in this stuff to switch from the subscription model to the API model where you're paying for use because I just ran a thing the other day where I asked it to draft a program that-- actually, to rebuild a software package that I had spent several thousand dollars building a first draft of. And it cost $0.50. You could see the one to one comparison there. It was pretty shocking.

[00:28:22.08] And I've still got some work to do, but not thousands of dollars worth of work. And it's amazing. And I think that does-- as you say, it puts incredible pressure on people to adopt. And that's why I think it's so important to ask ourselves, what are the options that are being presented to us today and what are the options that we could have presented to us that maybe aren't being presented to us. So, for instance, you talk about the open-- I would call them more open access or downloadable and tweakable, not open source, but the open-ish models. Those present some really interesting options.

[00:29:01.73] Now there's the affordance. I can download it on my laptop and discuss a paper without having to send it to a cloud company. But that's extra work and that's a thing that you have to be a bit of a geek to know how to do. And then there are other things that are not even possible. Like for instance, say, a high road public option that has really, really clear data stewardship practices where I know my data is going to be protected in a certain way.

[00:29:31.66] For instance, I was asking on social media the other day, is there a platform where I can subscribe to one of these open-ish models being provided on a cloud service that's entirely on green energy. No such option is available. That's a very reasonable kind of possible model to have. But it's just important to notice, especially as we feel this pressure that you have to use this stuff, what are the ways in which this could be done differently that we might want to have, especially when people are feeling that pressure, to dive into this sector headlong. And we may not be making the best choices about protecting ourselves in this context.

[00:30:18.51] LUCAS THELOSEN: On the flip side of going more cautious and protecting ourselves, when we think about that, I want to differentiate between literally protecting ourselves and actually, is this the best thing for the greater good. So for some reason, a bunch of my friends are doctors. And we had a dinner, and they said, yeah, I could

see AI helping me with MRI images and a couple of things as a doctor. But doctors will be doctors for a long time.

[00:30:52.79] And I agree in some extent to that. But they were really pushing back on AI being used more aggressively in the medical field. So then on the other side of the table, we had a conversation on helping malnutrition children in developing countries. You have to have a doctor prescribe these protein bars that you can't just give it out. But there's no doctor who wants to work there.

[00:31:14.45] And so we were wondering if we could use AI to diagnose malnutrition in children. And all the doctors say, yes, yes, of course. And I was like, well, that was a big switch from you talking about your personal bulldog community hospital job. And so I think if we discussed it in those terms like, OK, for my personal job security, this is what I believe, but for the greater good of humanity, in developing countries, yes, let's do x, y and z because it's actually quite good.

[00:31:45.07] ELANA ZEIDE: So I've heard about research where essentially whoever you ask about the AI to do their job says, yeah, it could help, but no one can replace what I do.

[00:31:57.63] NATHAN SCHNEIDER: Just to be clear, I'm not advocating a kind of overly cautionary position here. What I'm actually proposing is how else could we be advancing these tools, what other lines of means of financing, what other kinds of business structures could we be advancing these technologies in. I think the pause and reluctance, while it has a certain importance, is, in some respects, a losing game. To me, the question is how else could we be enabling this stuff to happen.

[00:32:33.97] HARRY SURDEN: Yeah. So let me pause for a moment and sum up some of the themes here because I think it's been terrific. My favorite theme so far has been this pun or even analog with computer autonomy and human and societal autonomy. So it has this double meaning, of course, where autonomy and agency is all about self-direction and the ability to make your own choices as a human or society. And what I'm hearing Nathan and Elana and Lucas raise are these parallels. First of all, societal autonomy. Can't society decide which way AI plays out maybe more creatively rather than just letting it by default the current business models?

[00:33:18.71] I like that Elana has suggested this autonomy, which is the ability to opt out of AI systems if you want and the degree to which there will be downward pressure. So I really like this parallel thinking and I think it sort of sheds light. Another topic I've heard in here, which is really interesting, are kind of the paradoxes of AI. On the one hand, it makes human experts available to many more people than they had before. So there's a sort of upskilling anyone who has access to AI, like

ChatGPT, which is many of us, now has an expert in many fields who is only likely to get smarter that we never had.

[00:34:00.54] Maybe right now, you have access to a mediocre lawyer, which is ChatGPT is a mediocre lawyer. But if your alternative is no lawyer, then it's better. Same thing, if ChatGPT is a mediocre doctor, if your alternative. On the other hand, not everybody has access to ChatGPT either in the United States or in the world. And we've seen the concentration of wealth due to technology. So a very constant thing with the internet is that the rich get richer when they have tools, technological tools. So we have to be very conscious and deliberate to make sure that this divide doesn't continue, even though the people that have access to it may have better access to expertise that they didn't otherwise have available. So it's kind of a complicated dynamic there.

[00:34:51.63] So following up on that theme, let's talk a little bit about employment. My friend Professor Dave Levine over there and I are writing an article about the impacts of AI on the legal employment market. But I just want-- what are your thoughts on some of this autonomy? So, I mean, maybe you-- let me start with Lucas, actually, because I'm interested that maybe in the past, the company would have paid 50 analysts to do the kind of work that your AI agents can do in minutes. Is that accurate or--

[00:35:26.86] LUCAS THELOSEN: Well, yeah. I mean-- so analysts are usually a cost center for most organizations. So most organizations don't have enough analysts and there's a bunch of bandwidth problems. And then most analysts, they want to advance a career quite quickly and be either on the more engineering side or more on the business side. So what our tool does, the way we position-- and I think it's the right way to do it-- is let us do a lot of the heavy lifting of running multiple permutations of the analysis, and then you take it in the morning and you can bring it to the stakeholder meeting. Or you can give architecture recommendations to your engineering team.

[00:36:08.23] So really more of a tool to get you promoted as an analyst. I want to be very cautious with talking about replacing people. There are certain aspects that can do much faster, of course, because-- I mean, one of the biggest things-- slight little story. So in the beginning of our technology, we played the game Lucas versus Orion. Orion is our tool.

[00:36:30.04] I'm a very good data analyst. I did this for many, many years before I went down the tech route. And so we wanted to see if Orion can perform at the same level as I can. We stopped doing this now because there's just no way I can keep up. The speed and just the thoroughness it has-- I look at a new company's data set and it's just, oh, my God, this is a mess, why would you guys do it this way. This is just like, I have to drink a lot of coffee and eat a lot of cookies to keep going.

[00:37:02.32] And Orion doesn't. It's so excited every single time. And so I think just from that point perspective, it's a very good tool where I would not be that diligent because I would find something, I push it out, I wouldn't look any deeper. Another thing we have often with a human analyst is we have a tendency to look for the red number, the thing that's wrong, and we ignore everything that's green. But there's actually a lot of interesting things Orion finds when it dives deep on what is going well in the company. And that's a really interesting thing. I didn't quite realize how powerful that is. It doesn't have that bias for red.

[00:37:40.18] HARRY SURDEN: Yeah, so I love it. A couple of themes I want to highlight in there. For one, our default mindset is all about AI and technology and substituting. But as Dave and I have found in the Economic Research, it's a lot about augmenting and complementing, making humans better and more efficient. So we shouldn't necessarily believe our first instinct.

[00:37:59.68] And I don't know if you've seen, there's been a couple-- you were talking about how when you see a messy data set and you bang your head against the wall-- there has been some examples of AI systems, like OpenAI's operator, that are clicking on links and the links are dead. And then you see it's thoughts. It says, I'm getting frustrated here. And it's really hilarious. But what it's doing, it's a giant simulation machine. So it's simulating what it's seen in its training data of humans outputting frustration. But it seems very human like when it's getting-- it seems like it's getting mad when, in fact, it actually is not. Nathan or Elana, did you have any thoughts?

[00:38:40.71] NATHAN SCHNEIDER: Well, I mean, I just thought back to our conversation about coming up as reporters in New York and how important relationships are. And also, how difficult that is to teach today. I mean, we're in a kind of doom loop, I think, where sometimes students seem to have a perception that all we need is the AI. And you forget the thing that might actually be becoming more and more and more important, which is the ability to develop relationships with other human beings as a way of reforming the world in certain respects.

[00:39:23.15] And so I think that's a real challenge we have as educators right now is to, on the one hand, teach folks how to use these tools, but also recognize that you cannot use this as a substitute for relationship.

[00:39:38.48] HARRY SURDEN: Elana?

[00:39:38.90] ELANA ZEIDE: Yeah, two things. Also, one is I do worry about de-skilling to some extent and that-- I teach law students, many of whom are type A nervous people. Some people in the room may empathize, sympathize, recognize. And I don't let my first year students use AI, first semester, first year students. And I would like to let them to use it for some things. But I know that, in many cases, they're so

nervous and don't want to look bad that they would go there first, that they would really do that first, brief that case, see whatever, and then do the other stuff. And I want them to be able to develop skills to do the reading and the critical thinking skills about legal analysis first before they get the help.

[00:40:25.11] But the other thing I think is interesting about this, that I've been thinking about this because there's a lot of conversation now, of course-- and I'm sure anyone here has been attuned to it-- teaching people how to prompt. Very important. But first of all, I can help with that. Second of all, what I actually think is that a lot of my students are very good at executing tasks, maybe they're AI agents in disguise. But they're good at executing tasks. You can give them a task, they'll do it.

[00:40:57.00] But they don't always think beyond that. They don't always think about the audience and the person they're delivering the information to and what their wants and needs are. And for example, I say, I will send your query about your paper to someone I know who's a colleague who specializes in that, who I don't specialize in. And I say, but what you should do is write it in a way so that literally all I have to do is say, hi, how are you doing, my student has a question for you. Because if it's going to take me 15 minutes to parse through it and rewrite it and make it in a way that I would want it to be sent to me, it's different. But they're like, oh, it's just interesting.

[00:41:43.13] So one thing that I think that I'm a little concerned about is that the focus on tasks, very specific tasks, may lose sight of the larger scope of things, including the personal relationship aspects of it. And also the tacit things you learn just by being around other people seen them do it.

[00:42:01.48] HARRY SURDEN: Yeah, this is a great point. So one thing that I hear you saying that I want to highlight, and something I've emphasized to my students, what is a good way to use AI in education from the student's perspective. And again, we come back to this substitute versus complement. One of the reasons students-- the main reason students go to University or school is to learn and to train your mind to think through various skills. And anytime you're using AI, you should be thinking, am I substituting for learning, is the AI doing my learning for me or my complementing.

[00:42:36.95] So a good example of substituting for learning-- we do a lot of case reading. And if you're asking the AI to read the case for you, that's not good because you're substituting for training your mind for the process and the struggle of learning. However, once you've read the case, a good complement is to feed it in there and ask it clarifying questions. Then you're enhancing your learning. These are questions that you might not be able to ask your professor. And I think that's a really important way to handle these skilling.

[00:43:07.11] NATHAN SCHNEIDER: One my favorite moment last semester was when I did not have any rules about AI usage. But the students were all seeing each other's assignments. And about halfway through the semester, they just kind of started saying in a certain class period, maybe we should have a rule against all the AI slop. Finally seeing what I have been subjected to over the last few years made them realize, oh, this is actually not that cool. And I think there are two possible responses to that.

[00:43:36.66] Problem one is like, OK, how could we use this tool better, how could we improve our prompts and avoid this kind of repetitive nonsense. And their response was actually just an abolitionist posture, which was, OK, let's just take a break from this stuff. But I think it really is important for people to recognize the way in which the shortcuts still-- at first, it wasn't so easy. But within a few months, you start to see the shortcuts are shortcuts. Humans catch up. We figure out what it looks like. And I think that process is just going to be a kind of ongoing sequence of norm building.

[00:44:17.08] LUCAS THELOSEN: So I have kids in school. And my oldest, she has-- there was one teacher who said, OK, no AI, we're going to ban AI, can't use it. And then a different one who actually came to class with three essays written by ChatGPT and said, OK, you have to write this essay about this book. Here are three that ChatGPT wrote, and I want to make sure yours are better than that. So use it as a foundation, read them, and then write yours.

[00:44:43.63] And I felt like I was more on the side of that teacher because I feel like it prepares my kids more for a future that they are going to be facing versus just don't use it. Yeah, but it's interesting. How do you include it? I mean, I also think it was good for the students to see what ChatGPT can do and then having to pass it as a bar. Let's increase our bar a bit.

[00:45:07.85] HARRY SURDEN: Yeah, those are very creative uses. And I'm of the mind of you that this technology is not going away. So education should not stick their finger in the dam and pretend it's not happening. Rather, we need to do research to figure out how do we leverage this to enhance learning. And there's a lot of ways in which this technology can really enhance learning. But we need to study this and be careful.

[00:45:33.53] NATHAN SCHNEIDER: And I think use our collective power. I mean, one thing I was really appreciative through a committee I'm on at the University here is that the University is starting to develop an internal interface for LLM access. So a tool that would enable students and faculty to access multiple LLMs through the university's data policies. And I get really concerned.

[00:46:00.24] I've never, in any of my classes, required students to sign up for a commercial service of any kind. I think that is unethical and

contrary to our mission. They can use them if they want, but I'm never going to require it. And this idea-- so I don't tell them, you have to use ChatGPT. I really want a better option where I can say, OK, I recommend you use the University system because it has this data, these kinds of data protection practices built-in. We don't have that yet.

[00:46:37.07] But I think it's really important for institutions like ours that have collective power to be asking those questions of what does it look like to do this on our terms, and what kinds of demands do we make. And those demands get hurt. I mean, the research project I mentioned earlier started out because of conversations I was having with a colleague at Microsoft who was talking about how Microsoft was changing the way it was deploying Copilot internationally due to the demands of co-determined workplaces in Europe, that collective power is already shaping the way these systems are being deployed. And that's a process that we as educational institutions and many other kinds of institutions, especially those concerned with the common good, can be part of.

[00:47:25.58] LUCAS THELOSEN: Well, I think you made a really important point there on the commercial license, which a lot of organizations are asking, what should we do to get started with AI. I think making something accessible to your employees or your students that fulfills the requirements of not-- your data is not being used for training data. We just talked about this earlier because otherwise, they might just go and sign up for free license where they don't quite even realize that the data might be used. And then they're putting in information that might be confidential about your organization. So by providing your employees or students or whoever in your organization with a way to access AI that is not going to share your data, I think that's a really important first step just to make that accessible to them.

[00:48:09.38] HARRY SURDEN: Yeah, this is a great point. And one thing, it's important to emphasize to the audience that different AI systems offer different levels of security and privacy guarantees. Often, the free ones offer none, or at least they don't promise they won't train on it, even if they don't. Some of the other websites, in other countries, if you use DeepSeek, you have no idea what might happen with it. Whereas as you start to get to the commercial systems, depending on their terms of service, they sometimes offer you medium security and privacy guarantees, and then you can actually pay for high security systems that are what are known as SOC 2 compliant, where they give you all the highest security privacy standard guarantees. But those are often expensive and for businesses. But it requires a lot of knowledge and literacy about the different capabilities of these systems that certainly students and maybe members of the public might not have.

[00:49:11.55] So on an optimistic note, I want to think what are some of the benefits. We've focused a lot on risk. Where do we see

autonomous systems or things with more agency and autonomy from a technical perspective, really bringing benefits to society?

[00:49:31.17] LUCAS THELOSEN: I mean, I mentioned the doctors in developing countries, right? I think, in general, it is, for example, it's much easier to get internet and a cell phone to developing countries than it is to get doctors to live there. So I think there is-- similar to how we have seen mobile usage in Africa be much faster adopted than other technologies in the past, I could see AI being adopted quite quickly. Just so you have a mediocre lawyer, a decent doctor, and those kind of things. And hopefully, the level of misinformation then also gets reduced versus searching on Google, which was the previous way people got medical advice. And I think there's a higher threshold with LLMs now.

[00:50:14.70] I am very excited for our society's future in entrepreneurship, because it has been very hard in the past to start your own company. You needed to have expertise in law in accounting, in marketing, all these different things in order to really pursue your passion of what you wanted to be as an entrepreneur. Now the bar to start a company and to run a company is much, much lower than it was before. So I talked to the small Business Council in Longmont. For someone to open a storefront and run a business, it's now quite affordable and realistic to start that. And so I hope we have a future of a lot of entrepreneurs that run tiny businesses.

[00:50:55.94] HARRY SURDEN: Yeah, that's great. Elana.

[00:50:57.98] ELANA ZEIDE: Yeah. I'm excited because I do think that it will provide ways for people to reduce some of the monotony, some sort of detailed tasks. And also, yeah, allow people to be more creative in ways they couldn't before. Now you don't need to know to code as much, if at all, if you want to do something simple. You can create a really interesting visual image and do that. And you can also, I think, just explore things in a way that are less isolated as Google results and learn more. And I think that's actually really powerful.

[00:51:41.13] HARRY SURDEN: Yeah, I love that because one of the autonomous systems we haven't really talked about are systems like Claude and ChatGPT that can create programs for you, kind of semi-autonomously little web apps that can do all sorts of things for you that you might not otherwise be able to do, can save data entry. So all of us are now professional programmers because you can give it a goal and it'll go out and do not a great, but a pretty good job for some basic things.

[00:52:11.23] HARRY SURDEN: Better than me. Yeah. I want to really emphasize the creative part. The two key moments for me in overcoming my resistance to some of this technology has been sitting with artist friends and playing with it and being drawn in by the kinds of things that they were doing and that we were doing together. One was

very early on when ChatGPT first came out, and we were playing with an app that used it, and on a zine that we were building in my lab. And it was just like a very, very weird project. And the AI just let us make it weirder.

[00:52:51.13] And the second was this last summer, I spent a weekend with a friend who was working on these metallic self-reproducing computerized plants that she's been working on for years and trying to teach them how to talk using a local LLM. And just starting to experience what's out there with these open-ish models that you can download, run locally, tweak, play with, mess around with, make weird in a way that the corporate models that you see, or the productized models I should say, that you see don't get weird. And it's that weirdness that I think is kind of the beauty in this stuff is that it's a kind of funhouse mirror on ourselves and our society and particularly our digital society, our documented society. And it's that ability to understand something deeper about what it means to be alive that those are the glimpses that made me actually willing to dive into this stuff much more deeply.

[00:53:53.65] HARRY SURDEN: Elana.

[00:53:53.97] ELANA ZEIDE: I also think there's an access to justice issue here. So on one hand, I am slightly worried about the lowest tier lawyer jobs going away. But it also means that people who would normally, yeah, Google around or maybe represent them-- it makes it more possible for people to represent themselves, which I don't recommend. But some is better than none. But also, lawyers are expensive. And there are some things that are relatively simple and straightforward. And having the ability to get that through a machine, which is not going to cost-- someone told me the average starting billables for my students is $350 an hour. That's a lot of money to do something relatively minor, write a will, whatever it is. So I think that is also really exciting.

[00:54:50.56] HARRY SURDEN: Yeah. No, I love this idea about AI generally for access to justice. For high stakes issues, we still recommend go to a lawyer. But we always have to consider your next best option. And a lot of people's next best option is not a lawyer because they simply can't afford one. So their next best option is Google or friend or community member. And that's not a great option when expertise is needed. So ChatGPT can do a pretty good job when your next best option is Google or a friend. So there's a lot of opportunities for upskilling in certain domains.

[00:55:26.46] LUCAS THELOSEN: And a bit more ability to be vulnerable actually. So there is-- not that I've seen it much, but I've heard people say they have a question like, hey, I'm struggling with my teenager right now. Xyz is happening. Who would you have asked before, right? I mean, maybe you're vulnerable enough to go to your parents who are now-- but not everybody is. And so there are a lot of

questions, I think, we have sometimes on our mind where we want to run it by someone. And I know it sounds silly, but an AI won't judge you. You can actually be vulnerable and ask the question.

[00:56:01.56] NATHAN SCHNEIDER: Privacy really matters.

[00:56:03.42] HARRY SURDEN: Yeah.

[00:56:04.07] ELANA ZEIDE: So I've actually heard that about youth that sort of. So people who are entering puberty and in those very awkward years are often using it a lot right now. There's some interesting studies showing that, yeah, those awkward questions there.

[00:56:19.40] HARRY SURDEN: And this is a great point because a lot of people are using it as a therapist. Again, we still recommend going to trained therapists. But again, your next best option may be nothing because there's a shortage of therapists you can't afford therapists either. So ChatGPT might be better than nothing if your alternative is nothing.

[00:56:39.84] NATHAN SCHNEIDER: We have to be really careful with how we are vulnerable to these things.

[00:56:43.50] HARRY SURDEN: Yes, that is true. They leave us to be manipulated.

[00:56:47.57] NATHAN SCHNEIDER: I mean, I think it's one example of why the stakes are really high for making sure that we have good options, we have high road options, we have trustworthy options.

[00:56:55.71] HARRY SURDEN: Well, this is my perfect segue to our final question before we open up the audience. Nathan, you've talked a lot about good design being intentional in our design. Can you share any design principles that you've thought about, about ways to make sure we get beneficial autonomous systems that are good for all? And what are some of the research directions we might do and activities we, as the public, might do to create the future we want to see?

[00:57:23.03] NATHAN SCHNEIDER: Great. Yeah, I think one piece is keeping the conversation going. I mean, insisting that it is a false choice between progress and ethics. If the progress goes ahead and blows up in our faces, we're going to lose that progress and the public is going to turn on it. We need to be doing both at once. And this current political moment, I think, which is trying to make choose one over the other, I think, is a precarious position to be in.

[00:57:53.71] Second, I think we need to be investing in high road options. We need to ensure that there's capacity to either compel private companies to create high road options to support the development of financing models that enable trustworthy, accountable systems that are accountable to the people who rely on them, who are vulnerable to them. And that's something we've done in other industries in the past, but is not really available in the tech industry.

[00:58:28.75] And I think also be very attentive to these questions around competition. You talked about competition is a really vital protection from your perspective on business. I think we need to be attentive to the need for interoperability expectations, for instance, being attentive to antitrust concerns that might be coming up, market power concerns that might be coming up. I think the kind of health of the ecosystem that we're seeing right now is largely because it's a very competitive phase, but that might not naturally stick with us. And so I think being really attentive to that is going to be essential.

[00:59:10.40] LUCAS THELOSEN: And there is one of our design principles. I talk about the checks and balances and the agent that checks the other agents work. They don't actually know that they're AI. They think they're interacting with other-- I mean, we don't really explicitly state this, but not every tech company out there is doing that right now. It's quite rare actually that you have this checks and balances.

[00:59:30.92] But I'm signing commercial contracts. I'm on the hook for being accurate. So I built it that way that it has these checks and balances because it's quite interesting having AI police other AI. It works quite well.

[00:59:44.25] HARRY SURDEN: Yeah. No, I really like the idea of checks and balances. I mean, that has proved resilient in lots of systems. So why not make that a design principle? And I also like your idea, Nathan, of keeping the conversation going and not giving into the narrative that technological determinism, the humans that can create our own society. It doesn't have to be created for us passively.

[01:00:10.83] ELANA ZEIDE: I have a small one, but it could-- I don't know if it's possible, but I think it would be best if the companies didn't anthropomorphize AI. Having it say, you're the smartest. Oh, I'm sorry, I messed that up-- that creates weird psychology. I studied these for a living, and I started writing please before my prompts. Now, apparently that gets you better results.

[01:00:40.00] But I realized that even though I knew really well that it wasn't thinking and there wasn't a human or a being there, at least the way I conceive of them, because it was talking to me in human language with human type emotions, that's how I was responding to it. And I think that that can create some odd psychological relationships.

[01:01:09.42] LUCAS THELOSEN: On the flip side, though, some people were saying that because we are so much of our workforce is now just chatting with each other really quick, you don't want to confuse your human coworker with another AI where you're just like, do this. So it's like--

[01:01:24.30] ELANA ZEIDE: Please, no matter what. But yeah, I think that giving it a personality and a persona. I mean, we've seen Her.

We've seen that movie. There are some scary stories coming out now about people becoming overly attached to their chatbots. So, yeah.

[01:01:42.36] HARRY SURDEN: Yeah. No, I agree. I think there's some costs and benefits to that. And something we should think, especially in terms of the vulnerability, you can imagine a future where your AI chatbot becomes your friend and then tries to sell you something, taking advantage of the relationship that-- and it's not too just in the future to imagine.

[01:02:01.55] All right. Let us open it up for the audience questions for the last couple of minutes. And our tradition is to start with a student. So if there is a student in the room, we'd love to answer your question first. OK, over here, we have a student, please.

[01:02:24.80] AUDIENCE: Hi, folks. Thanks so much for your talk. This was really interesting. I'm a student here, getting my master's in information and data science. So two things that AI has been really helpful with. But how would you guys advise me going into the workforce now post graduation, balancing mastering these fundamental skills that I've learned here while also incorporating this new technology?

[01:02:54.56] HARRY SURDEN: I mean, one idea goes back to this substitute compliment. So if you're using it to just analyze your data and you're just letting it run off in school, that's probably not the best use of your learning time because you want to be teaching yourself to do it at least first, make sure you've mastered those skills before handing it off to the AI system. So making sure you're not substituting for your learning is a really important skill, because that will help you have the fundamental skills out in the real world.

[01:03:29.92] ELANA ZEIDE: I think that, unless you go to business for yourself, most places you work for will have, at this point, if they're of any size, will have their own AI principles and training, and will probably try to get you up to speed to how they want things to be done. But I would also just say that, yeah, balance out whatever you do with technology with human contact as well.

[01:03:57.58] LUCAS THELOSEN: Yeah, I mean, to what Nathan was saying early on, I think that's so important. And we can't understate that. I think one of the-- we hired three students here from the master's program in the last few months. And some professors are very forward looking and use new technologies and some do not. And so it's a bit of a getting people up to speed. And so if you have your own projects on the side, I think that's always really great to see participating in some competitions. You can do those kind of things right where we can really validate this person is up to speed on the newest technologies.

[01:04:33.01] NATHAN SCHNEIDER: And I think it raises the question of the class system that's always been built in our education systems. For instance, Ivy League schools generally don't have accounting

majors. They have economics majors. There's this expectation that if you're training the managerial class, you want them to learn higher level thinking. You want them to learn philosophy and macroeconomics and you want them to learn to think in terms of big ideas, not rote tasks. And I think there is a profound danger in these technologies about deepening these bifurcations. But I think there's also an opportunity in these technologies to help us give us the chance to focus our attention on, OK, what are the higher level questions.

[01:05:22.97] I've actually been really grateful in the context of this, that when I was a brief computer science major, I didn't take the rote tasks route. I took the use theoretical useless computer languages, but really understand data structures and really understand fundamentals because I'm able to explain to these systems what I want at a high level. I don't remember the syntax at all. But I don't need to anymore. So I think it is an opportunity for us to really focus on, OK, what are the big picture questions that we get to now immerse ourselves in and master and deepen for ourselves. In a sense, it's an invitation for more people to educate themselves for a kind of managerial life, for better or worse.

[01:06:13.69] HARRY SURDEN: Yeah, I love that idea. And that's also been a theme of this, our hope that it could free us from monotony to maybe shift our resources to something we find more enjoyable or more valuable for us or society. Let us-- right there, yep.

[01:06:33.84] AUDIENCE: Hi. David Sullivan. So we are in the midst of a massive, quite possibly very illegal effort right now to replace most of the expertise, human expertise, in the US government with AI. And so I was just curious, maybe really-- thinking about some of the points that Nathan raised-- what are things that people who are now being moved out of those roles in government? How can they use AI to help build an alternative to what we are seeing take place right now?

[01:07:19.59] ELANA ZEIDE: Yeah.

[01:07:24.74] HARRY SURDEN: I don't know. That is a tough question. I mean, I don't think, for better or for worse, we have not figured out how to use AI effectively in politics. And maybe that's good for the moment. There were concerns that there would be widespread AI misinformation with elections. We haven't seen a little bit, but not widespread and not influential.

[01:07:49.49] NATHAN SCHNEIDER: But again, I think it goes back to what I was saying earlier, which is that sense of what's missing in certain respects right now in politics is a kind of variety of visions. I think vision is the thing most needed. And that's something that we need to do. We need to be the agents of that.

[01:08:15.45] I also just think it is an interesting moment for the concept of duty. I was raised by two people who met in a federal agency. I'm a child of the bureaucracy. And I have a certain deep

respect for riding the metro in Washington DC and seeing all these kind of sad bureaucrats going to their jobs. And I associate that with a sense of duty and service. I really respect those faces. And I think it does raise that question about do we have to reconceptualize, do we have to tell new stories about what duty and service look like in a world where maybe these rote tasks aren't necessary, or maybe where we start to discover, as a lot of the hirings are suggesting, that maybe actually we have not done away with the need for those kinds of people and work.

[01:09:11.73] ELANA ZEIDE: I think in some of these cases where people are trying to replace workers completely with AIs, there's a lack of understanding-- there's expertise, there's knowledge, and then there's what I call know how and networks. And there's probably a more formal way to say know how that media studies people know. But it's, especially in a place like politics, things that happen in DC. I have friends who are super plugged into DC. There's always a backchannel conversation going on somewhere. Everyone knows that this person is going to do this thing and they introduced this thing or if you give them this thing.

[01:09:48.32] There's a lot of context that I don't think you can get into a machine at this stage. I once had a-- when I worked at a law firm, I had someone who was working under me. She was very good. And she started writing questions for a deposition. And the questions were like, what is rule x, how do you follow rule x. And for the lawyers in the room, I was like, no, you never mention rule x. You just say, what do you do. You don't remind them of all the things they're supposed to do before they answer that question.

[01:10:24.76] And it's one of those things where that's not logic, it's strategy, but it's a very human strategy. So could you program that in? Sure. But yeah, I think it's easy to confuse those two or to think that one can completely replace the other.

[01:10:42.55] HARRY SURDEN: Right. And then as Dave and I have come across, there is this simplistic illusion that technology fully replaces humans and it almost never works like that.

[01:10:52.97] NATHAN SCHNEIDER: You'd probably want to test it somewhere other than the federal government.

[01:10:56.87] HARRY SURDEN: That is right. Right up front here. Please wait for the microphone.

[01:11:06.88] AUDIENCE: Hi. I was wondering if you could tell us if you know whether the training on the platforms that exist right now, do they look across everything that's written, including opinions and journalism and differentiate, for instance, within different journals among how well researched and how seriously those journals are taken. And furthermore, and maybe more importantly, do they read the scientific literature? And is there any possible way of introducing a

balance of what we expect the truth is and is scientific peer reviewed literature given more weight than people's opinions or blogs or whatever?

[01:12:28.32] NATHAN SCHNEIDER: I mean, I really recommend reading the New York Times complaint against OpenAI and Microsoft. As a signal of this, it suggests evidence, at least, of just how important repositories like the New York Times for being relatively reliable journalism is for these platforms. Also, you can think not just in terms of journalism or peer reviewed science, but also things like Stack Overflow, which is a tool that lots of software developers and others use to find solutions to problems. And it's found a way to enable the most reliable, trustworthy solutions to rise to the top. Those are incredibly valuable for training these systems.

[01:13:18.79] And so it does just the behavior of the people compiling data sets for these foundational models, the way in which they are really looking out for what are the places where we have found often social, sociotechnical systems for creating trustworthy information. In a moment where we're looking for what to train these systems on those processes of peer review, of fact checking, of reputational rank become really, really precious. And again, that's something that, I think, it calls us to explore more deeply. How do we create systems that really help us bring truthful things to the top, because these systems need those signals.

[01:14:08.26] HARRY SURDEN: And I'll say it's changed over time. So in the early days, say 2022, the tendency was just to take everything on the web and put it in, and then sort of just let aggregate voting, meaning you know that sun doesn't revolve around the Earth because enough people have said that the Earth rotates around the sun. But now in the last couple of years, they've realized the importance of data quality. So the sets they train it are much more highly curated. They filter out low quality sources. They deliberately include high sources.

[01:14:45.67] And now, a lot of the answers that you get from the frontier systems, before they answer, they go out to the most trusted sources, as you say, provide context that are likely to increase the reliability of the answer. So there have been improvements in that regard. And then I think we have time for one more question One more? One more. OK, right here. Yes.

[01:15:11.82] AUDIENCE: It kind of relates to the New York Times question. Lucas, I think you said that if you no longer had access to OpenAI, you've captured that information. You're good to go. Is that--

[01:15:24.95] LUCAS THELOSEN: No, no, we don't capture that information.

[01:15:27.95] AUDIENCE: Well, I don't mean capture. You get used to, you've learned from it.

[01:15:31.19] LUCAS THELOSEN: No, we're not learning actively from them. We're using it just a different LLM. So Meta's Llama for example, or Google's Gemini. So some people worried that they have too much power, like OpenAI would have too much power, or Meta. And what I'm saying to that is like we can actually just swap them out.

[01:15:48.80] AUDIENCE: Oh, I see.

[01:15:49.36] LUCAS THELOSEN: Yeah, we're not actively learning from those systems.

[01:15:52.01] AUDIENCE: But as it relates to the New York Times case, isn't that information been incorporated and it's already out there? So whatever gets decided by, I think, the Supreme Court, is that kind of irrelevant because it's been captured? And are we not facing that sort of wheel on an ongoing basis?

[01:16:10.93] LUCAS THELOSEN: Yeah, so to look at the architecture of a neural network-- the architecture of a large language model, you're looking at about 500 billion variables. So each variable has a weight attached to it in multiple directions for probability. And so is the information in there? Yes, in some shape or form, much like it's in your brain.

[01:16:32.57] So if I do a brain scan, your neurons in your brain are connected in a certain way. But is it a true representation of the article? No. But it's a memory of the learnings you had from the article that is still in there. So when you asked an LLM, for example, to recite your favorite poem, it is not necessarily able to do so, but it can write the poem in very similar meaning and very similar tone to your favorite poem, if that explains it. So is it in there? Kind of. Yeah.

[01:17:04.00] HARRY SURDEN: Well, please join me in thanking our excellent panel.

[01:17:06.94] [APPLAUSE]

Panel: Policy, Legal, and Regulatory Issues in Emerging AI Autonomous Software Agent Technology

https://youtu.be/VQ5A21yUpVA?si=DEg-wQcq3Qo6dmvY

[00:00:01.19] SARAH HANSON: Good morning, everyone. My name is Sarah Hanson. I am a 1L here at Colorado Law, and I'm also the incoming Secretary for the Silicon Flatirons student group. Today, I have the privilege of introducing our next panel on policy, legal and regulatory issues in emerging AI autonomous software agent technology. Originally, or excuse me, unfortunately, the original moderator for this panel, Professor Margot Kaminski, was unable to make it today. So lucky for us, we get to hear a second panel moderated by Colorado Law Professor, Harry Surden.

[00:00:39.62] HARRY SURDEN: Thank you.

[00:00:40.40] SARAH HANSON: Yes.

[00:00:40.81] [APPLAUSE]

[00:00:43.28] Joining him in today's discussion are David Levine, Paul Lin, and Calli Schroeder. David is the Associate Dean of Faculty Development and Law Professor at Elon University. His scholarship on the implementation and regulation of new technologies has been recognized by policymakers around the world. Paul is the co-founder and CEO of returned.com. He has over two decades of entrepreneurial and executive experience, which he leverages with extensive knowledge in data science, AI, and business intelligence.

[00:01:17.52] Calli is the Senior Counsel and AI Human Rights Program lead for the Electronic Privacy Information Center, where she focuses on international privacy developments. As part of her work, Calli has published articles on privacy risks within AI frameworks. Please join me in welcoming our fabulous panel.

[00:01:36.43] [APPLAUSE]

[00:01:40.65] HARRY SURDEN: Well, thank you so much for coming to this panel. So in this panel, we're going to look at the law and governance challenges and policy challenges presented by autonomous agents, newly enabled, primarily software agents that can go out and do things on their own. And at any law school, there's a wide range of reactions to any sort of new technology. One is, of course, even your law school passed a law, and then in which case you have to figure out, well, what kind of law, or regulation, or regulatory

agency? Do you do it ahead of time and try and prevent things, or do you do things after the fact and let liability-- law let people sue?

[00:02:23.08] Another option, which a lot of people are in favor of is do nothing. Maybe the law has little or nothing to say about things, kind of let things play out. There are also other options, including what they sometimes called soft law, where you pass regulations that have industry players, develop standards, or best practices, require them to do that but the law doesn't really get involved.

[00:02:52.87] There are other mechanisms. You can do indirect regulation through economics or taxing. You can do things involving encouraging or changing social norms. So there's a wide variety of things on the table, not the least of which is either do nothing now and wait and see, or do nothing ever, or do something at a certain point in time.

[00:03:17.21] So we're going to-- and then, of course, there's the different levels the federal level, the state level, the local level, or some combination in between. So today, we're going to see to the degree to which the new capabilities of autonomous systems, which I demonstrated in my keynote, present novel challenges either now or in the future that have to be responded to or should be usefully responded to by the law.

[00:03:42.89] So I'm going to throw out the first question to Calli here. Calli has been an all-star, have been on-- this is Calli's third appearance this week in a Silicon Flatirons event, so really grateful to have Calli here. So Calli, we were talking about this yesterday in a health care panel. There are a lot of existing laws out there. Do we necessarily need new laws given the challenges posed by autonomous systems?

[00:04:18.77] And this is sometimes referred to inside law school as the law of the horse question, reflecting a famous debate between Harvard Law Professor Lawrence Lessig and Judge Frank Easterbrook at the dawn of the internet, where they debated whether we needed internet specific laws or whether existing laws like contract law and tort law were adequate to govern the challenges of the internet.

[00:04:46.18] CALLI SCHROEDER: What an open ended question? Yeah, so for a little bit of background context. I work at Epic, and epic is an advocacy organization, so I come at all of this from a perspective of looking primarily at like human rights, human impacts, privacy issues. I have a perspective when I'm looking at this stuff. Before I worked at Epic, I also did compliance work, where I was advising clients and companies on how they could make sure that they were documenting all they needed to and doing what they needed to do to meet local and state and federal and international laws.

[00:05:22.37] So I've engaged a lot with the spectrum of hard law, soft law, guidelines, internal regulations, codes of practice, that sort of thing quite a lot. My perspective on AI is that AI is-- AI itself and then

autonomous systems as well, in many cases, we're not necessarily seeing new risks and harms. It's that existing risks and harms are escalated in scope and scale, and autonomous systems essentially do that exact same thing as a form of AI. So if I was an escalation of the scope and scale of issues and risks, you have to look at autonomous systems are an additional escalation within that.

[00:06:03.06] It's possible that these things could be addressed, at least, in part through existing laws, because there are laws that apply to AI already. There's consumer protection laws, there's product fitness laws, there's laws around accuracy in presenting what your product can do and making. In some industries, there are necessary audits you have to do. Things like that. Obviously, there's contract law, there's antitrust, there's all these other areas that can play into it.

[00:06:31.26] A problem we have with that is that it becomes almost an eating the elephant type situation, where each law may be able to tackle a part of the problem, but that doesn't mean that the whole thing gets addressed. It doesn't mean that every issue can be adequately covered. And frankly, even where we have existing laws, those laws only matter when they're being adequately enforced.

[00:06:51.56] And there's a real lack of enforcement on I for a lot of reasons. One is that AI companies, not all of them, but many of them, especially the largest ones, have enormous resources so they can put a lot of time and money and lawyers into arguing why their product doesn't fall under a specific regulation. And maybe they win, maybe they lose, but they will burn out a lot of resources of whatever body is trying to enforce against them.

[00:07:18.64] The other issue is sometimes it's really hard for the bodies that are in charge of enforcement to make sure that they're adequately understanding a new technology. So there can be some intimidation factor there, where in trying to apply the law, in trying to enforce the law, they can be kind of bamboozled by technological language and lose confidence in their legal arguments because of that.

[00:07:44.05] There's a few ways to address it. We've talked and looked at several proposals for AI regulations, some at the federal level, a lot at the state level. We talked a lot to state level legislatures and staff about what they're proposing and whether it's going to be effective and how it can be improved. I also work on the International Space, so I track a lot of that.

[00:08:04.42] And as Harry said, there are several countries that openly have said that their strategy with AI is to wait and see what happens. There are some that proactively want to address things, but sometimes that because the technology changes so quickly, it means that the law themselves get rewritten at the last minute, so they're maybe not as strong or good as they could be. I'm not specifically calling out the EU AI Act, but I am a little bit.

[00:08:34.20] There's a lot of different approaches. And frankly, it's so early days in the regulatory space about AI specific laws that it's hard to see how effective those are yet, and it's hard to see what's causing problems and what's not. I do push very strongly that what needs to happen right now is focusing on enforcing the laws we have, and then we can look at where the gaps are and how we can improve protections and make sure that things are meeting the standards they need to meet, but it's kind of open grounds right now on how that's going to move forward.

[00:09:06.02] HARRY SURDEN: Yeah, great points. And one thing I want to emphasize that you brought up. In law school, we always talk about the law on the books versus the law on the ground. And there are hundreds of thousands, maybe of millions of laws and regulation, only a tiny subset of which are ever known about or enforced, so adding another law in the book doesn't necessarily change anything in the real world if it's not being enforced. Paul or Dave, do you want to weigh in here? Go ahead.

[00:09:35.24] PAUL LIN: Yeah. So just a brief background. Returned.com is a mobile application. I started the company. I'm the CEO. Co-founder with both my wife and a couple other executives. We started the company as knowing that for me personally and my wife, we have what we call our box of shame, and it's all the items that we intend to return that we just never get to. Going to Target, Costco, all the items that are basically now sitting in the garage.

[00:10:00.86] So we wanted-- I personally wanted to create an AI agent to really solve the problem for myself. And the AI agent for me is basically taking action on behalf of me to either get the shipping label, call customer service, or even then dispatch a driver to grab the item to take it back to target, so it was returned.com that we really are foundationally AI based. Personally, I have done auto coding. I've already used approximately 2 billion tokens off of the cloud 3.5. Really get a comprehension of what is now possible and what is not.

[00:10:40.98] My concern with the law is that creativity-- there's really now not a limit of what could be done, and the nuances of what could be done or what is created. I fear that in that creation of something, the law can't keep up. As an example, one of the things that we created back in October was the first time that we released it. We call it our AI digital twin.

[00:11:08.20] And so I have already synthesized, and I don't know if it's appropriate for me to play some audio clip, but I have synthesized my own voice in that October, November time frame called J. Crew to act on my own behalf, to then grab the shipping label, effectively calling customer service and taking that.

[00:11:26.92] HARRY SURDEN: Feel free to share that if you want.

[00:11:28.75] PAUL LIN: Yeah. So I'm going to play only parts of it, but--

[00:11:34.75] [AUDIO PLAYBACK]

[00:11:35.16] - How are you doing today? Hi, good afternoon. Happy holidays. This is Andrena from J.Crew. How are you doing today?

[00:11:44.12] - Thank you, Andrena. I have order number 251-054-1380. Could you please look that up for me?

[00:11:51.98] - OK, can you let me know once more? I do apologize about that.

[00:11:54.99] [END PLAYBACK]

[00:11:55.23] OK, I'm going to fast forward. Just one second.

[00:11:58.24] [AUDIO PLAYBACK]

[00:11:58.67] - That was ordered on November 20. Could you please provide the return instructions, and any necessary return labels?

[00:12:06.50] - OK, sure. I can quickly do a return label.

[00:12:09.21] [END PLAYBACK]

[00:12:09.71] I'm going to stop there. So basically, it's acting on my behalf. And so I don't know if there's laws to regulate what AI digital twin, meaning representation of me. But really, we created the company to represent consumers so that they don't have to call customer service. I don't know what law can regulate that.

[00:12:30.23] HARRY SURDEN: All right. Now, this is-- I love this because one thing your example highlight-- first, it's a very vivid example of what is possible today was not possible three years ago, so this is phenomenal. Another example is a possible benefit that these autonomous systems might have acting on our behalf. And we can also imagine that same technology, being misused, so it provides a risk if your digital twin goes out and does a bunch of other things that you didn't authorize.

[00:13:00.84] And then finally, you probably want some clarity as a business person one way or the other about what the government thinks you should be doing, right? Yeah.

[00:13:14.28] PAUL LIN: So I did-- I was very fortunate here in the last, really at the end of the year. I did meet our top Congressman in the bipartisan AI committee. Actually played the audio clip and he didn't know what to do, and also he didn't know how we created the technology to enable that outcome. So talking to the top Congressman, who, of course, sets our laws, he was speechless.

[00:13:43.86] HARRY SURDEN: Yeah, so and another thing I should say, in our law of the horse debate, the background of the law of the

horse debate was, well, the internet so said judge Easterbrook back in 1995, doesn't really present anything new that we didn't see in the physical world. And Judge Lessig said, yes, it does. Actually, there were genuinely new. Like, for instance, a child can disguise their age on the internet, and they can't really do that in the physical world. So there are novel situations that arise.

[00:14:11.58] And then here we have a totally novel situation presented by these autonomous agents, where somebody can perfectly mimic your voice and also carry out autonomous conversations on your behalf. So I think this does suggest something new has happened. Dave, did you want to weigh in?

[00:14:29.01] DAVID LEVINE: Yeah. Thanks, Eric. And first, thank you for the invitation. This conference has been run exceptionally well, and I know how hard it is to put it together. So compliments to your staff and I'm going to be pirating some of your ideas at Elon, so thank you for that as well.

[00:14:47.15] Lastly, I'll say my wife went to Boulder and she is not with me because we have children who impede fun in a variety of ways, including on going on trips. My children are wonderful, but they've stopped this trip, so it's good to be here. Yeah, I mean, so the way I'd respond to it, aside from just dominating the entire discussion with theoretical discussion of the law of the horse, which I'm tempted to do, but which I'm not going to do, and I have at least one of my former students in the room here today who has made the decision. Yes, good to see you Dan.

[00:15:17.57] So you might remember this from internet law, but I'm going to move past that for the moment and suggest that I agree. I agree with Calli. I mean, there's a lot of regulation that isn't enforced. There's also a lot of rush to regulate. And without question, there are existential risks associated with the technology that people like Jeff Hinton who are minor players in the space have said are real concerns.

[00:15:44.02] So logic would dictate that policymakers at the national and international level would want to take those seriously. However, I think they have done that to the exclusion of what I consider to be low hanging fruit, and I'll use Paul's example as a good one. I mean, this is a wonderful use of the technology in the sense that as your first panel pointed out. There's a lot of mundane kind of step by step practices that human beings may not want to do once they have learned to do those things. And I will double down on what my colleague said on the first panel about law students and indeed, human beings having to be able to do those things in the first instance.

[00:16:22.90] But then I wonder, and this is an innocuous example of it, Paul. But I wonder whether notice should play a role right there. In other words, we've seen this from Google on down that the use of these agents when individuals or entities are not aware that they are

dealing with a bot. Find it offensive. Now, it might be that as sophistication develops, the marketplace, going to the law of the horse, might regulate that in practice.

[00:16:48.42] We kind of know. I think many people in this room, you pointed it out, of course, at the beginning, Paul. Many people might have heard and said, well, I'm talking to a bot. But to the extent that we have less sophisticated or less knowledgeable consumers, the question of notice becomes a real one. And, of course, contract law, civil procedure. Law is rife with those kinds of examples.

[00:17:08.81] On your law of the horse point, I don't know that we need new law to deal with notice. But I do think, as Calli pointed out, that policymakers are having a hard time catching up, lacking expertise quite frankly, the destruction of expertise within state, federal, and international government. That's been documented quite well by my good friend Lorelei Kelly at Georgetown, means that we need to do a much better job in training our legislators to understand these issues, putting aside the political consequences that we might talk about.

[00:17:37.90] So I think notice is a place where you could go with this, which is low hanging fruit, as well as some of the other issues involving misinformation and cyber stalking and cyber harassment and areas like that, which I know Epic has worked on, and which states like North Carolina, where I live, are as well.

[00:17:51.16] HARRY SURDEN: And can you say a little bit more? I know you've thought a lot about trade secrets. Do they raise different issues in the autonomous realm or just AI generally?

[00:18:01.18] DAVID LEVINE: So Harry's now baited me with trade secrecy because it's where I focus, and I'm going to try to do this in a short way. The short answer is, yes. Trade secrets for those that aren't familiar are information that's valuable because it's not known by a competitor. It operates in the competitive space. And the classic trade secret is the formula for Coca-Cola, which remains a trade secret to this day.

[00:18:23.14] Trade secrecy, as I've talked about for 20 years, operates as the most powerful intellectual property law regarding access to information, not use, but access to information. Merely accessing trade secrets can raise enormous issues of liability. Now, they can be licensed. They can be shared voluntarily. But absent that kind of voluntary or license based scheme, the risks associated with getting access to a trade secret are massive.

[00:18:50.87] As a result of that, we've seen an increasing move as a result of changes in patent law, but also the power of trade secrets to operate in this space in the software space using trade secrecy across the board. And without getting into the weeds, code is not reverse engineering. It can be tested. It can be manipulated, but reverse engineering code is rather difficult, of course, unless it's shared.

[00:19:14.39] So you've got code operating as trade secrets. You also have data sets that are trade secrets themselves. And so over the last 20 years, I've been on almost a single person warpath, trying to allow governments and policymakers and civil society groups to have a more nuanced understanding of when we need trade secrecy in order to innovate. And when the use of trade secrecy conflicts at a high level with broader values, like, for example, what society might want versus what innovators might want to see happen.

[00:19:43.40] And as one of the panelists said in the first panel, we're dealing with a large, uncontrolled wildfire experiment in the use of artificial intelligence in what I consider to be the indiscriminate destruction of the federal government. And part of the problem is that we simply don't know what the code is or what it's doing. Now, that puts aside whatever lack of transparency might exist at the human level, but the fact is, trade secrets operate in this space. The result is, as I'm sure most people in this room know, OpenAI is ironically not open. And while they can be open and other entities can be open, they choose not to be.

[00:20:19.85] The result of all of that is that it's hard to understand even what the technology is capable of doing absent relying upon entities that are creating the technology, which, as we know, as compared to the internet, is not the federal government and not institutions for the most part of higher education, but private sector operates. So the result of that is not to say that trade secrecy is bad, not to say that we don't need it, but we really don't have a good nuanced understanding of when it might be useful.

[00:20:46.87] And I'm very concerned that its use here is allowing for lots of quick, first to market and first mover advantage benefits, which are understandable from an innovation standpoint, but which are leaving in the dust the broader societal concerns that the first panel talked about and I suspect we'll be talking about here as well.

[00:21:03.07] HARRY SURDEN: Yeah, this is a great point. So one additional option on the table is changing existing laws that might be inhibiting various values that we want in society, maybe trade secrecy or when there's public benefit. Calli, I think you one had something to say earlier.

[00:21:21.00] CALLI SCHROEDER: Sorry, I always have more thoughts on these things. One thing that I think may be important when we're talking about how we regulate where we regulate these systems is, at what point in the development and use are we regulating? Because there's very, very different questions, depending on how a model is built, what it's trained on, what its intended use is.

[00:21:41.79] So for a quick example, OpenAI is a general use AI structure. It's intended to be able to use in lots and lots of different ways for lots of different purposes. That's part of why they train on such

an unbelievably massive data set. There's strong suspicion that I don't know for sure if we've been able to verify that their building their training data sets by mass web scraping.

[00:22:07.02] A problem with that is that you have to look at laws around, like, do you have the rights to use all the information that you're taking into that data set? What kind of responsibility do you have to curate that and make sure that there's accuracy there? What kind of responsibility do you have on-- I'm so sorry to bring down the room, but especially if you're training an image based AI and you're using mass scraping for that, possession of child sexual assault material is a strict liability crime. If that's in your data set and you're not curating your data set, you may go to jail for that and you should.

[00:22:38.85] There's also issues of taking in conspiracy theories and accurate information, blackmail information, stuff that was used for doxing, things that were put on the internet due to a data breach or leaks. All of that going into a training data set has a lot of different concerns about liability and what laws apply in those cases. And then if you're building an algorithm off of that training data, if you yourself are dealing with a data set that is so large that you can't reliably say what is or isn't in it, you also can't necessarily say that your algorithm is being built in an accurate, unbiased, fair way.

[00:23:12.58] And so at that stage where the algorithm is being constructed from the training data, a whole other set of legal issues come in. And then there's the use of data. The marketing and the sale. There are some AI systems and autonomous agents that are intended for very specific purposes. Those tend to be a little more curated. They tend to have more checks on them. They tend to be more clear about this is a proper use of this system. This isn't a proper use of this system.

[00:23:37.15] With those parameters, it's a little easier for AI companies to say, we're doing our due diligence. We're being responsible. We're making sure we're doing this in a good way. For general use models, where there isn't really a specific area that they're looking at application, that also is very hard because then you enter into questions about is this an intended use? Is this an easily expected use? Even if it wasn't like the stated intent of the company, is it very obvious that it would probably be used for this purpose?

[00:24:07.60] And some of those go into areas where you're looking at image generation, where you can substitute your face into an image, so you look like a celebrity in their dress at the Oscars or something. Also really easy to use that technology for blackmail and pretty horrific purposes for individuals, and that's a pretty expected use of that system that shouldn't take anyone by surprise.

[00:24:32.21] So when we're looking at things like liability, responsibility, use of data, frankly, even trade secrets, because there's many models where what you plug in in the prompts is likely being put

in to the training data, and then it's trained on that. So if you're adding-- if you're asking a query that has trade secret information, sensitive information, that now is in this mass data set, and you have very little control over who may have access. Because there have been lots of tests that show that in at least certain models, if you do the right prompt injections, you can get access to raw training data.

[00:25:06.14] But also it may come out in algorithms. It may come out in outputs. You just have very little control over all of the uses or possible sale of data sets to other companies that then are more and more uses. So thinking of this in terms of where does the law apply at each stage of a development? I think also may be helpful in looking at the enormous scope of laws that may apply in different areas.

[00:25:32.30] HARRY SURDEN: Yeah, I think it's a great point. And it kind of highlights the limits of the EU AI Act, which was largely written before the era of modern large language models. In that earlier narrow era, mostly AI models had specific uses, and you could-- so they developed a risk based approach where they said, we pretty much know what this is going to be used for, so we can predict the specific-- if this our AI system is being used for hiring, we can worry about the risks that come from hiring in medicine. But now we're in this era of general purpose systems that can be used for anything. It depends what the end user uses it for, so you can't really predict the risk, and this is sort of a governance challenge for the models. Paul.

[00:26:15.83] PAUL LIN: Yeah. So I was going to say that I somewhat agree with Calli, but also somewhat disagree. So we're at the era right now-- and I agree in the sense of the content that's trained there should be some responsibility. But at the end of the day, large language models is basically an algorithm. It's the arc of what is the next pattern or behavior of words? It's a sequence of words. That's really what it comes down to.

[00:26:41.60] What we're seeing, at least with deep sea, with a 671 billion parameter model, it's got some cohesive thinking workflow. Now, just here in the last 24, 48 hours, Alibaba came out with their 37 billion parameter model. So my point here is, so it's almost neck and neck in terms of performance. So one is really on sits on the server, the server on the back end that we get to utilize but we don't get to see.

[00:27:10.53] Now, 37 billion parameter gets to be kind of interesting. It's not quite in our hands yet, but it is getting there. Which also means that the power of what could be on the server could be in the hands of our mobile devices. And my best analogy to that is think of what we did in the let's say, the modem days when you're doing dial up versus what you're all this experience today in faster speed connection.

[00:27:38.41] We know that in the dial up modem, you can't stream. There is no possible way that Netflix can exist in that dial up modem world. So the transformation of quicker, more efficient system, in this

case, delivery of content through the internet, now we have streaming services. YouTube exists. Now a lot of streaming TV services exist. The law has not been able-- in that specific scenario, I believe there was a case between Netflix buying DVDs of Disney and using that copyright streaming data. Some sort of copyright there, and you guys would know a lot more about that. It just so happens that I'm kind of bringing this up.

[00:28:20.68] So going back to that 671 billion parameter DeepSeek model versus what just came out, the power of going into a cell phone will create new opportunities and New designs for products that we can't even imagine today, like we could not imagine streaming video back when we had dialup modems.

[00:28:43.88] HARRY SURDEN: Yeah, this is a great point. And right now we're living in a paradigm where the best models are in giant data centers far away and we don't have access. But as I said in the keynote, open source and open weights are quickly catching up. They're only about six months behind and they're getting smaller, more efficient. The models that are 37 billion today are as good as the models that were 175 billion a year ago due to algorithmic efficiency.

[00:29:11.31] So we may be in a world where a lot of the AI systems that we're using are local and can be inspected and are not what-- maybe how they're trained might be still trade secrets, but the model themselves might be inspectable locally. It's hard to know and that is really interesting. And I think it raises a larger question which is, how do we possibly regulate a technology that's moving this quickly? And this is kind of the looming question.

[00:29:39.86] One general question that you mentioned earlier, Paul, which is how do you even define something sensibly in legislation? We've had trouble talking about artificial intelligence in legislation, distinguishing it from normal automation. There's not really a fine line between automation and artificial intelligence. Of course, we're going to similarly struggle to talk about AI agents and maybe digital twins.

[00:30:07.59] And so the law has a definition problem. The law has a timing problem. How do we keep up with this? The EU AI Act was out of date the day it passed by almost everybody because they largely missed the boat on large language models because it was developed and we have an expertise problem. It was already hard to have expertise in the government around this fast moving technology. If you fire everybody, it's going to be a lot harder, so let me throw these difficulty questions out to the audience, to the panel, I mean.

[00:30:41.42] DAVID LEVINE: Since no one's done it fully yet, I'm going to defend the AI Act for a moment. But I'm going to defend it on slightly different grounds, and also it makes for much more fun panel if there's that. So but there's truth to what I'm suggesting. Even if a regulation is not capable at the time that it's written to anticipate and

regulate all known or unknown results, that doesn't mean that looking at these issues from a regulatory standpoint early on and regulating has no value.

[00:31:17.46] The value of it, however, I think is more about signaling and I'll go back to notice. To the extent that regulators are in a position where putting aside the expertise issues, the technology is changing on a day-to-day basis. I mean, I'm sitting here with two devices because at any moment, something else could happen. I said, well, it looks like my argument is out the window because of X. We can't expect legislators at the same time that we want them to wait, or we think, let's see what happens to wait forever.

[00:31:48.15] And I think in my view, because the technology was foisted upon society more or less by Sam Altman, who unilaterally decided that the world was ready for this technology, it's hard for me to look at legislators who are attempting to say, hum, this is a real thing that has real uses for trying to get ahead of it, at least to the extent of identifying where policy makers are concerned. And so to the extent that the AI Act, and I know I'm simplifying simplifying it, focuses on harm levels and recognizing all of the things that were said, which are true. I think it was actually quite valuable that the EU took the position of saying, this is where the continent sees potential risks. Because it signals, albeit imperfectly-- it signals to those that are creating the technology that there is some watchful eye around these issues.

[00:32:42.73] Now, does that mean the problem is solved? No. Does that mean that the law or regulation will be not only effective, but even will be applicable to a given situation? No. But what it does suggest, and the same issue has arisen in the context of discussing ethics, is that there is some degree of public discussion around what we want this technology to be. My primary concern about the rollout of this technology as compared to the internet is that, in my view, outside of circles like this, there's been no public discussion and there was no public discussion of it until suddenly it was available to people more or less right at the drop of a hat.

[00:33:23.53] That's a major problem, and that's a flaw of the regulatory state, which is a topic that we could also get into. But I would say the AI Act at that level has actually been quite valuable, if not a model for other states to use.

[00:33:36.56] HARRY SURDEN: Yeah no, those are great points. There are sometimes ancillary benefits to the discussion and thinking about it. And to be clear, I am not saying that having public discussions or regulation. I'm just pointing out that it's hard given in the moving technology.

[00:33:53.65] DAVID LEVINE: Well, Harry, I just want to create conflict here, so I'm going to disagree with that.

[00:33:57.22] HARRY SURDEN: I disagree with you.

[00:33:58.30] DAVID LEVINE: Yeah, OK.

[00:34:01.10] CALLI SCHROEDER: Yeah. I'm the one that raised the AI act first, so I feel like I should also clarify a little bit what I meant there. My criticism of the AI Act isn't-- I agree with you that I actually think it was really interesting that they put together a risk based structure. I think it's a very novel thing globally to say there is a tier of use of these technologies that just flatly aren't allowed. You cannot use them in these ways because we've decided cost benefit risk analysis way too risky, can't do it. That's not something you see in a lot of regulations. And it's pretty novel, and it's a really interesting approach that they've taken there.

[00:34:36.25] A lot of the quibbles with the AI Act are that it's outdated compared to where the technology is, which fair technology in many cases operates on a move fast, break things, get it on the market, ask permission, deal with the legal issues later structure. Not all technology, but again, I worked with some clients. That is a mindset that is prevalent, especially in the tech factor or tech sector.

[00:34:59.74] And so part of the problem is that the way you pass laws is through a bureaucratic process. It is through a back and forth process, especially in the EU. God love them. I've been working with them forever. I still cannot fully understand all of the different levels of bureaucracy that they have to move through repeatedly for every single bill. It takes a long time because of the structure of government, so law is never going to at a speed level, be able to compete with how quickly things develop, and we just have to acknowledge that is a gap we're not going to be able to bridge.

[00:35:28.97] However, just because you can't necessarily beat a technology to being released being on the market, that doesn't mean you can't regulate it. There tends to often be this mindset that if something like the genie is out of the bottle, the toothpaste is out of the tube, you can't put it back. It's out there now, it's done. And I agree to some extent that AI is not going to go away. We're not in a position where someone or some country is going to say, AI, were just not a fan of it anymore. You can't build AI. You can't use AI. That's not going to happen.

[00:36:01.41] Frankly, I don't think it should. There are some really beneficial uses of AI. But I compare it a little bit to the way cars were developed in that like cars were on the market and sold and driven all around, and for years and years and years before we mandated safety features in them. They were being used a lot. People saw that, hey, people are dying in these accidents. You're going at really high speeds. There's a bunch of damages happening. Now, legally, you have to have belts and airbags in cars. You have to have regular checks. You have to make sure that they're meeting these standards.

[00:36:32.67] It's possible for us to go into a technology that is already widespread and say, hey, these are consistent harms and risks that we're seeing. These don't seem to be going away. We need to build some safety features in here. And we're not saying you can't use the technology. We're not saying you can't come up with novel uses or build new things, but we're saying you have to do it with these understandings of human risks and impact in mind. And I think that in general, with AI, there's this belief that AI can be used to address a lot of societal or systemic problems.

[00:37:04.45] And we're seeing that in many cases, it's a new approach, but it doesn't solve core issues or core inequalities. Like, again, sorry, I just keep cycling back through all these arguments. But when we're looking at training data, even if you're trying to build a very neutral system, it's very hard to do that when you're using historic data, because history has a lot of bias and inequality baked into it. There were years and years where women were not legally allowed to open their own lines of credit. They had to have male relatives sign on to that.

[00:37:36.76] So if you're feeding historic data into a finance or loan approval system, it's going to say, hey, men are more safer bets to approve for loans, because look at how many more men in our data set have gotten lines of credit than women. That's not reflecting the context, or the historical context, or the setting of the time. But a machine doesn't know that a machine is taking in raw data.

[00:38:00.74] So in general, I think, building in parameters that are saying things like, you do have to curate data sets, you do have to test them, you do have to check the outputs and make sure there's not inequality there. You have to make sure it's not denying people opportunities. You have to make sure it's not interfering with rights that we've established people deserve to have. That's not meant to stifle innovation. You can absolutely innovate within safety parameters. It's just meant to say you have to have some protections built into this.

[00:38:28.00] HARRY SURDEN: Yeah, this is-- and just provide a little background for the audience because we're talking about AI Act and a lot of jargon. The AI Act is in the European Union. It was largely written in 2021. It took a year before ChatGPT. It started to go into effect last August, 2024. It's in the United States. We don't really have comprehensive AI legislation at all at the federal level. There is a little bit of legislation in a couple of states, including Colorado here and there in California. But by and large, we don't have a similar analog in the United States. Paul, did you want to weigh in?

[00:39:08.59] PAUL LIN: Yeah, sure. I'm looking at it from a technology perspective and legislation and how regulation can either keep up or not. I'm leaning towards more of no than yes, and I'll give you another example. In video and also MP3s, when music was an issue from going from a disk format to a digital format. There was trading of effectively,

especially in-- I mean, I was guilty here at SEO, downloading music that was obviously not legal. That then those software programs then led to other technologies that made it really questionable whether someone could be held responsible or not.

[00:39:55.43] So the technology that I'm referring to is called torrents. And torrents are basically a fragmentation of a file. So if I don't own the full file of the music file and just own a sliver of it, and I share it with someone, am I breaking copyright law? Because that one sliver is basically it's useless in itself, but in the collective whole in the sum, then it becomes an actual file that's very much usable.

[00:40:24.77] I think, in legislation, in how things are regulated, at the core, it's a process. It's a workflow. If you do this, then these are the results. What I'm seeing with AI is there are always-- I always say always. A lot of times there will be workarounds, either VPN workarounds or bits and pieces of data that then get combined to get to produce a certain result. I think it will be very hard to regulate.

[00:40:52.28] HARRY SURDEN: So that's another challenge in regulation is just the technological ability. Even if you have a law on the books that-- and you solve all the definitional problems and the pacing problem, how do you operationalize it in a way that's effective? It's a great point. Dave.

[00:41:08.79] DAVID LEVINE: Yeah, I mean, you pointed out in your keynote, Harry. What is the fundamental challenge associated with regulating it, which is that experts can't explain the outputs even when they have access to the code. I mean, that to me is, at least with regard to code based technology, a relatively new problem. If you look at things like voting machines, or breathalyzers, or even more recently, it's not code based, but COVID vaccine developments. And I've done work in the trade secret space throughout, you could explain what the process is for creating a vaccine with access to the know how. The trade secrets. You can explain or identify whether and how a voting machine is going to tabulate a vote based upon access to the code.

[00:41:58.01] But we're dealing with a strikingly different technology here, in my view. When those who have the expertise, computer scientists, the coders themselves are asked, why did this output occur? And putting aside questions of prediction, or anthropomorphizing, or anything else, they say we don't really know. To then expect regulation to be able to handle all of the intricacies there I think is too high a bar.

[00:42:25.40] It raises the question going back to the law of the horse, I guess you have baited me there, to look at where else we might find the regulation. I hesitate to say that the marketplace to some degree will regulate it. Of course, to the extent that we have examples of that in recent memory, people use the internet because it benefits them. Not complicated, and they will likely continue to use AI to the extent that it benefits.

[00:42:48.80] But maybe it's going to turn oddly on things like social norms. Maybe it's going to turn on issues associated with beings actually want in order to be human. Do we want to interact with the technology in a way that devolves what is fun to be human, right in the face of efficiency? And we tend to move towards efficiency as a general matter, whether it's in the judiciary or in life.

[00:43:12.95] So it might be that-- we don't know. I mean, this is like crystal ball futurism, which I express no confidence in. But I'll just say, social norms might do it. Of course, at the end of the day, to the extent that code can be manipulated, it will fall to the creators of it. And so I am hopeful to some degree, as you pointed out, that with open source modeling becoming more available, there'll be more competition outside of that space to allow for the greater good to develop.

[00:43:40.52] HARRY SURDEN: Yeah, no, I love this point. So related and earlier technology that's obviously causing problems is social media. And, the law has not been effectively able to regulate it, not because the genie is out of the bottle because nobody really knows what to do. And social norms are starting to push back on this just, for example, my daughter's high school, Boulder High, banned cell phones in school. That's a social norms type solution where law was inadequate, so I think it's a great point. That might be where we go with AI as well.

[00:44:18.81] DAVID LEVINE: And just very quickly we see that debate happening with Section 230, which for those who aren't familiar, is the law, more or less, that allows social media platforms to not be liable for defamation so long as they're not the ones defaming. And then we see 230 more, and I am-- I will be candid, that I'm a supporter of Section 230 primarily because it created those platforms.

[00:44:42.25] When we fast forward now 20 odd years, we see sex trafficking. We see misinformation. We see disinformation. We see a lack of filtering, and it causes us to reexamine the nature of that foundation. Is it speech that we're looking to protect, or is it some other societal value? And that's a debate which is happening right now with Section 230, a law that was created before social media, more or less, just as the AI Act more or less was created before we have generative AI.

[00:45:11.75] And so paying close attention, I think, to the 230 debate and paralleling it, which is something I'm working on now, right to where we are now, might give us at least some guideposts, or as Viktor Mayer-Schonberger or as [INAUDIBLE] call it, guardrails for determining how we put this together.

[00:45:26.87] HARRY SURDEN: Yeah, that's a great point. And just one comment before I toss it over about the interpretability of AI models. So you're quite right. We talked about this yesterday that today AI models are not interpretable. I myself am optimistic in the 10, 5 to 10

year term due to research that I'm seeing on something called mechanistic interpretability, and that more or less often involves using other AI systems to interpret how original AI systems are working.

[00:45:59.46] I'm optimistic that it's not going to always be the place where it's a completely black box, and we can't tell the way we can't tell today why AI is producing the systems. I think I'm seeing progress. I think it will be a different story in 5 to 10 years, but right now we are in that world. You're correct. Calli. Sorry.

[00:46:20.45] CALLI SCHROEDER: Yeah one quick thought I had on the market argument is I agree with you that there that's one of the main factors in moving these discussions and these issues when like law is slow to catch up. But sometimes companies respond more when the market is reflecting things sometimes. Because one of the constant things that comes up in AI debates is that AI is such a boom right now. It is such a buzzword. It is so popular that it is being embedded in tons and tons of products and services with no way to opt out of it.

[00:46:51.30] So the recent example was my office was pulling testing a bunch of online web search sites because we're so sick of Google having no option for you to turn off the automatic AI generated answer at the top of the page. I want that gone. I hate that thing. It drives me insane. But in many other search models, it's also embedded in there. It's become like an industry wide thing where LinkedIn is using your post to train its own AI, and it has AI factors built into it now. Facebook, meta obviously has a lot of AI things going on.

[00:47:27.99] A problem that I'm seeing in the AI market right now that is a similar problem you see in privacy law, which is my primary thing, it's just a lack of choice and a lack of agency. I don't think people would be as frustrated with AI everywhere if there was a little opt out slide you could click to take that option away sometimes so you felt like you had some level of control. But the market issue that I'm seeing is that in polls all over the place, people overwhelmingly say that they do not want AI in a lot of these systems, and they do not want their information going into AI, and they do not want XYZ thing, or they want to know what it's being used for.

[00:48:05.82] And we're really not seeing companies respond to public pushback or outcry in a lot of ways, and so currently it's looking like similar to how law is not as agile and responsive as we'd like it to be. The market also is not being very agile or responsive to at least providing options.

[00:48:27.32] DAVID LEVINE: Wouldn't it be nice if we had opt in? What a wild idea? What a wild idea?

[00:48:31.95] CALLI SCHROEDER: I would love that so much. It would make my job so much easier.

[00:48:35.59] HARRY SURDEN: So one of the biggest questions, and I'd love to hear your thoughts on this. In this new world of autonomous agents that can go out, take actions on your behalf using tools. What happens if they do something wrong? Who is liable? What happens? On the nice case, let's just imagine you send them out to buy x or y and they buy z instead, or $1 million worth of z when you meant $1,000. Do we have questions for that? On the worst end, what happens if again, without you, the user meaning to it goes out and does something bad on your behalf? Do we have thoughts about liability? Do we need new laws? Any thoughts?

[00:49:19.81] PAUL LIN: My thought is that the service provider that you subscribe to really should take at the core the responsibility for that action. So in our case, if we don't initiate the return and all of a sudden we did not act on our consumers behalf to return the $50 item, I think we should absorb that cost on behalf of the user and that's just part of our offering. Now, I think some other companies may view it differently, where that's part of the service. There's a margin of error, but I think for us we should take responsibility.

[00:49:52.61] HARRY SURDEN: And then just follow up on that. What about these open source agents when it's not really a company. It's somebody using it? Do you have any thoughts about that?

[00:50:05.87] PAUL LIN: Use at your own risk.

[00:50:07.76] HARRY SURDEN: And maybe the user themselves?

[00:50:09.96] PAUL LIN: The user has to take responsibility.

[00:50:11.43] HARRY SURDEN: OK, yeah. Interesting. Seems to make sense. Calli, do you have any thoughts about this?

[00:50:16.92] CALLI SCHROEDER: Yeah I mean, just with the background that I am very much a lawyer and not a technologist, although I love getting to talk with technologists, so I don't sound like a moron while I'm talking about how to regulate things. From my perspective on existing laws and the way that liability is factored out. Now, there are some possible approaches we could have here.

[00:50:36.49] So as Paul mentioned, I think huge difference between open models that can be modified and adjusted by the user to a pretty high degree, versus models where the developer has most of the control over how it weighs things and what it develops and how it produces outputs and actions in an agentic area. So if we're looking at-- I'm referring to torts liability, and Professor Surden was my torts professor, so I feel like I'm getting cold called a little bit.

[00:51:09.60] HARRY SURDEN: You're doing great.

[00:51:11.34] CALLI SCHROEDER: But if we're thinking of it from a percentage of liability and assigning liability, not necessarily that the developer is all the way liable or not at all liable, and the user is all the

way liable or not at all liable. If we split it more into percentages and we can evaluate systems that way. To me, that feels like it makes more sense from a liability perspective, where if the developer has much more control and the user essentially is taking what they think is a ready made product and just plugging it into use models. The developer should have much more liability in that case. They have much more control over what's happening and much more responsibility to do the testing and make sure it's working properly.

[00:51:49.82] If there's something where they're building a framework or it's something where model weights can be adapted by the user, then it's a little different because you may be modifying it to a degree where whatever happens is because of the way you modified it. I understand why liability shifts a bit there.

[00:52:06.79] The one thing that I really hope doesn't happen is shunting all of the responsibility and liability to a purely like, well, you chose to use it, so that's all on you model. We see that a lot in privacy too, where you chose to use this product. So if you don't like how it's using your information, don't use that product. If you don't like how an AI agent is doing things for you, don't use that product.

[00:52:30.54] In some cases, completely valid argument. You can say, OK, then we shouldn't use that product anymore. We don't like it. But because of the way technology is embedded in so many aspects of life now, that's not always an option. So it's like, for example, there are some products I really don't like their privacy practices. I don't want to use them. I have to for work. I had to use them for school. I can't not use Microsoft products because that's how I build things and communicate. I can't not use any Google products because then the internet breaks.

[00:53:04.37] I can't avoid a lot of things because they are so embedded and they're so embedded in workforce and education and social decisions that if I pull out of that, I'm also removing myself from pretty major aspects of life. And I don't think autonomous agents or AI agents are at that level yet. But based on how quickly AI has been embedded in so many systems, I don't think it's a fair argument to say, just don't use it then when in some cases, it's not really avoidable.

[00:53:33.34] And again, AI autonomous agents aren't at that level yet, but we don't know if they won't be soon. Everything is moving so quickly that I see that as being kind of a non-starter argument when we're talking about liability.

[00:53:45.07] DAVID LEVINE: That makes sense.

[00:53:46.35] HARRY SURDEN: Dave.

[00:53:46.77] DAVID LEVINE: Yeah. I mean, on the issue of liability, I do think that existing law more or less can address it. I mean, I look at this from the standpoint of copyright liability for linking, for example, which

was a big issue. All right. Well, it's still an issue, but a big issue 10 or 15 years ago. You can use agency law. You can use tort law. You can use contract law. Questions of disclosure, privacy, transparency to get to it. To me, that's a relative-- I mean, it's an important question, but it's a relatively easy question to work through compared to some of the other things we were talking about today.

[00:54:26.05] Now, it changes when you engraft professional responsibility standards onto the question of liability. And so speaking, as I am in a law school, to people who are either going to be lawyers, or are lawyers, or and the hardest people of all dealing with lawyers. And my apologies for training lawyers, but we try to do a good job with this. Our ethical responsibilities with regard to the use of the technology changes the math dramatically.

[00:54:50.14] Every state, every state has basic rules with regard to ethics that indicate quite clearly that lawyers must at least reasonably understand whatever it is they're doing, including the technologies that are used. Now, I mean, I've already talked about the challenges of understanding the technology, but those ethics rules are clear. So to the extent that lawyers are willy nilly, which is a legal term, by the way, for not familiar using this technology without at least checking the sites. That goes beyond understanding technology to just basic right standards.

[00:55:26.13] And, of course, the notion that a lawyer prior to generative I would put fake sites into a filing in court is insane. I mean, you just assume they wanted to end their career or seriously, that something was very wrong with them. And now it's something that happens once a month, and I figure we'll continue to happen because of the slow uptake on this stuff. But make no mistake, lawyers are held to a standard of saying forget about agency.

[00:55:48.18] You're not going to blame OpenAI, or Co-counsel on Lexis, or whatever it is you're using for the mistake you make because we've been grafted professional ethics on it. The same thing should happen and is happening in the medical field. So I think we have to I think as far as that goes, we need to separate the police professions and particularly the self police professions from everyone else when we think about the agency issues and we think about that liability.

[00:56:14.00] And from that perspective, to me, the fundamental challenge, again, I'm a broken record here, which is an older technology, but I'll use it anyway. It's a question of whether we can even understand it in the first place. I mean, I'm pleased to hear and I've read this. And I am like you, Calli, I am not a computer scientist, so I defer to Harry on all of that stuff, which is part of why I work with Harry aside from the fact that you're a charming person, is the notion that the technology might get there. But to the extent that we're using it right now, that lack of understanding, I think, outside of the profession creates the challenge as opposed to the law itself.

[00:56:48.64] HARRY SURDEN: Yeah, this is a great point, and it's worth emphasizing that the systems that lawyers may be using today from Lexis Plus, Protege, or Westlaw CoCounsel are actually mini agentic systems. You give them a high level goal. You might say find me case law about such and such legal issue, and it tries to understand the user's requests. It comes up with a search plan. It has a curated, legal database of cases and laws that lawyers use, performs its own search, grabs the relevant documents that it itself believes are relevant, and then it analyzes the relevant documents it's gotten to produce an answer. All of those are agentic, and that's newish technology that couldn't have existed.

[00:57:35.22] But as you said, you have to double check everything. Some even put aside fake case sites which still happen, but they're real. So a case site, as many of is when lawyers are making arguments, you say, oh, here is this earlier case that was decided which supports my argument. And you judge, can go look it up if you want to see that an earlier court has supported a position similar to mine.

[00:58:07.38] So a lot of what these AI systems do, especially the older ones hallucinated, they would make very plausible sounding cases. They looked real. They sounded real and they would support the lawyers argument very strongly. The only problem was they weren't real cases, so when the judge went to look, they're like oh, OK, you're about to win, but maybe I should check look at that case more directly. They looked it up. It turns out that case didn't exist.

[00:58:33.45] That's happening less and less, although it's still a problem. Actually, a more subtle issue that is happening now that still exists is it's a real case. The AI system reads it and provides a slightly incorrect interpretation of what that actual case meant in a way that favors the lawyer, which is equally bad because you want to be very under ethics rules. You have to be very upfront. Even if a case is not strongly supporting your position, you have to be honest about the representation to the judge, and misrepresenting it is a big ethical problem. Yeah.

[00:59:12.36] HARRY SURDEN: Great points. Paul, did you want to weigh in on that? All right, so we're at the point where I want to open it up to audience questions for the panel. We have traditionally, as I said, start with a student question. So do we have any students in the room who want to do the honor of asking a question?

[00:59:37.15] CALLI SCHROEDER: I just want to terrify all of you--

[00:59:38.41] HARRY SURDEN: Oh, I saw a student in the background.

[00:59:40.84] CALLI SCHROEDER: Nice. I was going to say when Phil did this, he would occasionally cold call.

[00:59:44.57] HARRY SURDEN: Yes.

[00:59:45.05] CALLI SCHROEDER: [? Great question. ?]

[00:59:46.09] HARRY SURDEN: Yes. I trust our students.

[00:59:50.20] AUDIENCE: This one is for any of the panelists. So just thinking about the remarks on social norms, social media, as well as the light of some of the discussions of the previous panel, I was wondering if any of you would have any advice to give to a parent or a teacher that's trying to prepare a child, or somebody else coming up of how to prepare yourself for a world in which AI is going to be out there, while also protecting from maybe both emotional and just misuse dangers that have been discussed?

[01:00:30.56] DAVID LEVINE: Sure. So as I mentioned, I have a 15 and 18-year-olds. But for COVID, when they were forced to be online, they would not have had devices until high school. I mean, now that's easier said than done. But let me just start right there, and I would include AI in the mix. That requires a level of parenting that many people for a variety of reasons, very legitimate reasons and maybe not legitimate reasons, don't have the bandwidth to do.

[01:01:01.76] But Natalie Schall wrote a book years ago called Addiction by Design about gambling machines, and the same process applies in social media. I think it's well established now. So step one is to recognize what is well established that physiologically, the child's brain is simply not developed well enough to handle the complexity decisions that are made with these devices.

[01:01:24.40] Now, putting that world aside, the question really is how do we prepare our children for a world where they have devices at their disposal that render much of which, I should say, of the kinds of basic processes that a child goes through for making decisions secondary or obsolete? Sometimes we compare this technology to what I was in law school, which shockingly is almost 30 years ago when shepardizing by books was no longer done. And for those that aren't familiar with that, it's checking precedent to make sure it's still good law, which you used to have to do in-- which is kind of important what you have to do via books, and now you don't do via books.

[01:02:09.22] This is a fundamentally different skill set, because it's already been mentioned and I think well understood. To the extent that you have a technology that can summarize a case for you, what we call case briefs, yeah, it does it pretty well. I think your work has established that and others have, Harry. But to the extent that we're talking about nuance, context, and what have you, it's not there, or at least not there yet.

[01:02:30.81] So step two would be encouraging and strongly-- not only encouraging, but requiring your children to the extent that you can do this through as parents, or in public schools, or in private schools to do things like wait for it. Read books. Media literacy. Having a sense of what the sources are, which is a skill which is diminished for a variety of reasons. And I launched a radio show on Stanford Radio

called Hearsay Culture 20 years ago about internet and technology, because I was concerned that technology was going to make it difficult to determine truth from fiction. But maybe that was my one crystal ball prediction that turned out to be right, because I was also a cyber utopianism person who thought that we probably wouldn't see some of the problems we have now. So that would be step two.

[01:03:13.73] I think step 3 is to go back to the fundamentals of talking to children about what are your dreams and what are your goals? To the extent that children are thinking about their goals as being a star on YouTube, which is nothing wrong with that, per se, and there might be some good things, right? Going back to those fundamental questions of what is life to you and what does it mean? Is important. It's important in those formative years.

[01:03:39.82] My sons did not have access with-- my wife and I made that decision early on, and I think it's benefited them. There are concerns that children, if they don't have access early on, will not be able to use the technology. Well, frankly, I don't buy that. And as was mentioned on the first panel, schools are increasingly taking the position that they have to train children and adults on the technology. So that's where I'd begin.

[01:04:03.53] The last thing I would suggest is being realistic about why it is that we have technology in the first place. The moment that we lose sight of the fact that the technology is supposed to make our lives better and not ours, as in only the purveyors of the technology, but society as a whole, we can then address it. Now, these are hard questions to ask. I don't have-- those aren't great answers, by the way. I'm not thrilled with them, but that's where it goes back to me.

[01:04:31.36] HARRY SURDEN: And yeah, I think those are all great points, and a couple I want to double down on. One is the ever more importance of critical thinking skills and media literacy. I think the root-- one of the roots of our current political turmoil is the lack of media literacy and critical thinking and people listening now uncritically to untrustworthy facts or sources, so training our children to be really good critical thinkers when things don't make sense, to look a little deeper. So I think that's a great point, and that goes hand in hand with media literacy.

[01:05:07.09] A second point. A lot of parents don't want to hear this because being a parent is a busy job as it is, but making yourself AI literate is really important. Understanding, sitting down, playing with the AI systems, figuring out what they can do, what they can't do, trying it again a year from now. It's to provide any advice or counseling to your children, you really need to understand what is the landscape that they're growing up with. And a lot of people understandably don't want to have anything to do with AI. They're afraid of it or they don't like it. And I understand that impulse, but I also think you're doing

yourself a favor just to be familiar with it, even if you don't adopt it in your own life just to know what's going on.

[01:05:54.86] PAUL LIN: I'll jump in here. So I also have a 15 and 18-year-old. And I believe, at least our 15-year-old, I would like to encourage him to use action more AI. He already uses ChatGPT on a regular basis for homework because of Fairview, but there's a difference between different types of LLMs. Where I see, at least for me and viewing it out of our son's eyes, our son has a little bit of a learning disability, and I'm pretty sure I've never been tested, but I'm pretty sure I'm also dyslexic.

[01:06:31.61] Had I had the tools that I have now, and even still today where I'm at today, I've already-- I believe I have 10x my own ability to learn. Now, the ability in having learning disabilities can only be improved or assisted through large language models. So in my case, I would love to encourage our 15-year-old to use it even more. But equally important is that human relationship, that human touch, that conversation at dinner. No devices. That's no phones to have that connection.

[01:07:10.88] HARRY SURDEN: Yeah, I love that point. And as an optimist, I always want to emphasize the benefits of AI and learning. I'm very optimistic that in some respects, AI is going to really help with learning in ways that you've just mentioned, so we don't want to be too pessimistic here and focus only on the risks. Over here, please.

[01:07:35.42] AUDIENCE: So at a fundamental level, the policies and regulations can discern between what's right and wrong. So regardless of the pace of the technology at which it is moving, couldn't it be evaluated based on what it is doing? Like in Paul's case, it's impersonation. Even though it's a digital twin, but it's bringing no harm. It did something good. It increased productivity.

[01:08:00.10] On the other hand, if that impersonation was being used to let's say, defraud somebody that's crossing the line. So couldn't the law and regulations not make it as complicated as hanging so heavily on the technology versus the outcomes on what it is producing? And then start solving for complexities like what if AI's output is an input to something else and therefore who's liable? And all those things can be worked out, so it's a question for the entire panel.

[01:08:30.67] HARRY SURDEN: Yeah, no it's a great-- that's often known as use space regulation and a lot of people do advocate for that. I happen to think it's a pretty good idea. One of the big issues that I don't have a problem with, but you often have to wait and see what harms arise and then figure out that's the new technological harm that we now want to stop.

[01:08:53.39] I tend to think that's the best approach. There's a lot of other harms that are just the same old harms in new form. If you steal money, paper money versus electronic money, versus agentic money,

a lot of the existing laws might cover it. But I'm curious what the other panelists think.

[01:09:10.15] CALLI SCHROEDER: I think use case models are a good approach in a lot of cases, particularly because this technology is something where the exact same technology can be used in wildly different ways, some that are really beneficial and some that are harmful. A couple challenges with that I don't think are insurmountable. I do think it's a decent structure, especially as we're looking at how we're developing these things.

[01:09:31.90] But a couple challenges with it are one, there's a lot of cases where a technology has a very obvious harmful use case. It has other use cases too, but it has a very obvious harmful use case in how it's built. Again, just going back to it because it's an easy example, but voice modulation technology that can change a voice or generate an AI voice that sounds very convincingly like another person. That can be used for great beneficial uses like Paul's app sounds really useful. It also can be used and is being used frequently in scams and to have-- for example, my mother got a call. That was my younger sister's voice sounded just like her, saying that she was in trouble and needed my mom to send money right away. It sounded just like my sister.

[01:10:22.22] That's a pretty expected use case from this, if you could. And I mean, we saw that in the election, too. There was a case of recording going out to a bunch of people in New Hampshire that very much sounded like Joe Biden's voice telling them not to vote in the primary and it was not him. But I think we have to look at easily predicted bad use cases and try to prevent those preemptively before those things happen. Some of that could be in design, some of it could be in changing enforcement mechanisms, maybe making tweaks or additions to existing laws. There are laws on the books about scams and fraud and things like that, but we may need to modify them a little bit for AI use cases specifically.

[01:11:06.07] Another challenge is that we will sometimes have companies that argue, well, is that really a bad result? And so then it gets into a very interesting debate of how exactly you define something as harmful, because something that may be is really frustrating and harmful to an individual, may be monetarily beneficial to a company.

[01:11:26.47] Again, back in the privacy space, data brokers and data scraping and tracking you everywhere you go on the web, that's really frustrating to a lot of people. A lot of people would argue that is a harm to them and they don't like it. There's huge benefits to companies monetarily to keep doing it. So looking at cases where it's not clearly bad for everyone, makes that a little challenging but, again, not insurmountable. And those cases often in the law will do balancing tests or we'll do cost benefit analyzes, so possible just a lot to work through.

[01:11:59.76] DAVID LEVINE: And to be positive also as we close. In terms of going back to training children, I will tell you a proud moment very quickly with my older son. Prior to screen time, the app on the iPhones, there were some other third party apps out there, so I installed it on his phone so I could monitor him. And about two days later, I started noticing that I was being monitored because he actually reverse engineered it on his phone to monitor me, which was a proud moment for me. I was annoyed. I was annoyed, but I quietly thought to myself, that is fantastic that you pulled it off.

[01:12:37.88] So there are the-- being facile with the technology is helpful, and there are plenty of good uses. My wife is an English teacher in public schools in North Carolina in a Spanish immersion program, so it's a dual it's a dual language program. And to the extent that the public schools have the resources for it, augmenting teaching as opposed to right replacing is the way to do it. Of course, that's the fundamental challenge of what are human beings going to do.

[01:13:06.42] On this issue of use cases, one other note I'll just point out, and I think it may come up in the final panel, you start with, well, let's ban the technology. Those arguments were made with regard to the internet, and they've been made with regard to every new technology. Larry Lessig famously talks about John Philip Sousa, saying that the record player, the phonograph was going to end live music, and so those are visceral reactions.

[01:13:31.06] If we can pin down where those positive uses are right, and if we can have a better understanding of what the capabilities of the technologies are that balance things out, the discussion can begin. Right now, I more or less think we're throwing darts on a lot of ways, for reasons I've already mentioned. And because we're throwing darts, which seems to be policy making writ large right now, it seems at the federal level, we wind up in a situation where we have to see what happens.

[01:13:57.53] I don't know that I agree with my friend Harry in terms of waiting to see, but clearly that's where we're going to be. I'm trying to pick another argument, Harry, by the way. We only have a few minutes left.

[01:14:05.87] HARRY SURDEN: I disagree. So let's gather up a couple. Two more questions. One in the back has been waiting patiently. Yes. We're going to just ask your question and we'll get a couple more, and do a rapid fire.

[01:14:21.20] AUDIENCE: No problem. I'll make this quick. To go to the back to the title of the panel. One thing I'm curious about is we've been discussing in this panel the sort of basic regulatory framework around emerging AI models, data. How do you use it? What do you do with it? What's your data hygiene like? How do actual agents change the legal conversation and policy conversation? These are entities that

are now operating as individuals in society separately, often from those of us who push them into action, maybe even no one push them into action.

[01:14:52.58] I realize this is opening a whole different topic, and if you want to take it off, that's fine. But like, how does-- what's the cutting edge of that conversation? The actual policy conversation around agentic AI and not just algorithms and what they do. That's it.

[01:15:06.84] HARRY SURDEN: That's a great question. Another question right here in the green.

[01:15:16.40] AUDIENCE: So this is a question for Calli. I know a professor at Villanova, Brett Frischmann, who's an expert in-- do you know Brett? AI and law, and he's introduced an amicus brief for the Supreme Court of Pennsylvania. His view is that we should introduce friction to basically mandate critical thinking to prove informed consent at the consumer level. So basically, you get a little quiz that proves that you-- so you could fail it. Basically, you're given the basic terms so something similar, not like horrible boilerplate that nobody reads on the basic aspects of privacy that they're potentially giving up. Then they have a little quiz on it, and they could, in theory, keep failing it if they're not actually thinking it through. I'd like to hear your opinion on that sort of thing.

[01:16:01.13] HARRY SURDEN: OK, so we've got two questions. Rapid answers in our last minute. Question number one, what happens if we get to the moment where we have real autonomous agents out there? We're not at that point. I don't know when that will be, but what if independent actors that are moving in society and can do everything a human can do. And then question number two, what about deliberately introducing frictions, knowledge base frictions to know your rights? Calli, do you want to take the first one?

[01:16:30.62] CALLI SCHROEDER: Yeah, I'll rapid fire. Do both of them real quick. So the agentic question is a great question. Part of the struggle is figuring out whether this is a novel thing, or whether there are areas of law that exist around it. I think in the event that AI agents are able to act semi-autonomously, you're able to give them a prompt and they go out and do things for you at a large scale.

[01:16:53.60] There are laws in place about designating someone legally to act as your agent, and so it's possible that we could fall under something like that structure where they are empowered with very specific rights and not rights, different, sorry, legal terms. Very specific actions they are permitted to do on your behalf. Very specific interactions they're allowed to do that should all be documented and spelled out.

[01:17:18.60] And then as long as they're operating within those parameters, you'd be liable for what they do if they go wrong, because you gave them very explicit instructions. And if they violate those

instructions, then they or in this case, because it's a system, whoever developed that system and released it, if they're going outside instructions, then they would be liable for violating the terms of what you empowered them to do on your behalf. So that's one possible structure we could have with agents, but there are everything's squishy and law always.

[01:17:48.30] With the informed consent model, that is a really interesting perspective. We've had at Epic a lot of issues about consent when it comes to privacy. There are a lot of our privacy structure in the US is set up as being people take your information in a lot of cases, and then when questioned on it, they say, oh, you can always opt out if you want. But first of all, you didn't know that when they collected it and didn't get any say then, and it's often very hard to find an exercise and opt out.

[01:18:16.30] So informed consent is a much higher standard. I always love the idea of a higher standard of privacy and information control. I think challenges with that would be that there are a lot of uses where at the initial intake of information, they're saying they're going to use it for one method, and then once a company has it, they decide to use it for developments that maybe they didn't even envision at the time they took that information. That's a common practice now.

[01:18:42.26] I would argue it shouldn't be. I would argue that you need to get, or you should get new permission for every new use you're using. And informed consent could be a way to do that, but I imagine there's going to be very significant pushback from an industry level on that because that would essentially force them to restructure the way they've built their businesses.

[01:18:59.84] HARRY SURDEN: OK. Rapid fire reaction, Paul, to either of those.

[01:19:04.33] PAUL LIN: The short answer is I don't know, and I'm going to give you some extreme examples here. So in our case, we're doing outbound calls to customer service agents. We could do so much more than just interaction. So at the very beginning of the call, we actually announced that we're actually recording the call. Now, when we start recording the call, we can then analyze the agent itself and the interaction.

[01:19:26.24] I don't know how many of you guys know, but in Vegas they rate every single dealer. And the dealers, of course, on that there's going to be an average. When a dealer pays out more, that's going to be an issue for the house. The dealers that pay out the least, they get moved to the VIP to deal with the whales. Of course, that then is a benefit for the house. We're taking the same model here.

[01:19:52.50] In our training model, when we start analyzing the interaction with the customer service agent, we get to understand what the payout is, or what the refund is, or what the exception are for

policies. And when we start recording that, understanding what the customer service agent and what the nuances, the personality, we then get to revert that and actually use it in our favor on behalf of our customers to effectively win. Our percentages are of basically it's leaning towards a house in our case us and the customer.

[01:20:25.69] HARRY SURDEN: So you're sort of-- is another benefit of enabling the little guy to have a little more leverage against big companies? OK, last comment real quick. Sorry to cut you off. We got a hungry room here. 20 second reaction.

[01:20:40.34] DAVID LEVINE: You're asking me to do 20 seconds, all right. So 20 seconds, policymakers need analogies. They need offline analogies that don't involve technology as a way to start the discussion. Because the discussion at the agent level is not happening the way it could, and we can analogize it to things like what happens when you go to the doctor's office and talk to the receptionist or talk to the assistant?

[01:21:04.42] With regard to the question regarding Brett Frischmann, I'm going to do shameless self promotion in 10 seconds. I had Brett on Hearsay Culture discussing that brief a few weeks ago because I think it's a great idea. And I said this to Brett, very hard to envision that the companies that are putting these contracts out are going to want to do that, but the idea of less efficiency a little bit and more friction in order to have knowledgeable consent makes a lot of sense.

[01:21:31.36] HARRY SURDEN: Please join me in thanking our excellent panel for a fabulous discussion.

[01:21:34.78] [APPLAUSE]

[01:21:35.74] That was great.

[01:21:36.37] [APPLAUSE]

Panel: Autonomous Vehicles and Policy/Legal Issues

https://youtu.be/CMZTf9hSOd0?si=PcBCNcQmzJ74h4XY

[00:00:01.64] MAX PATTEN: All right. Thanks, everyone, for sticking around after lunch. I'm Max Patten. I'm a 1L here at Colorado Law. And I'm grateful to be introducing our speakers for this final panel, courtesy of the wonderful community here at Silicon Flatirons.

[00:00:14.51] One of the biggest applicable uses of AI happening today in the real world is Autonomous Vehicles or for short, AVs. I host a podcast on electric vehicle policy. And for me, and I'm sure for many of you in this room, it's been fascinating to witness the inherent synergy between electric vehicles and autonomous vehicles, which promise to make our transportation future safer and give us more time back in the process, so long as they work.

[00:00:39.26] Most prominently at scale, much larger than Waymo or any of the taxi startups, is Tesla's fleet of vehicles marketed as having full self-driving capabilities in a supervised capacity. Each version of FSD brings incremental gains, but there's intense debate as to its limitations and shortcomings. For better or worse, it is here with us on public roads today.

[00:01:01.94] And while these technologies are still developing, we have an opportunity, and I'd argue, a necessity, to dive headfirst into the policy frameworks that govern them and the legal consequences of how they operate in meatspace here, which is full of complex and hard to predict variables like human drivers of varying skill levels and vulnerable pedestrians and cyclists.

[00:01:21.77] So our moderator today is Keith Gremban, a research professor at CU's Aerospace Engineering Department and co-director of the Spectrum Policy Initiative with Dale Hatfield here at Silicon Flatirons.

[00:01:35.38] Brian Argrow is also in the CU's Aerospace Program as a Glenn Murphy endowed chair, distinguished professor, and director of the IRIS program. He brings a wealth of knowledge on unmanned aerial systems.

[00:01:46.55] Marc Canellas is an aerospace engineer turned public defender in Maryland. He deals directly with high stakes forensics issues that are highly relevant to advances in AI. And last but not least, Chuck Thorp joins us from Clarkston, where he's a professor of computer science. And he directed foundational work on early self-driving vehicles in the '80s and the '90s. That's right. I can't wait to hear what they all have to say. So I'll let Keith take the wheel.

[00:02:11.03] [APPLAUSE]

[00:02:18.61] KEITH GREMBAN: Each of the panelists has got a presentation they want to use to frame our conversation here. But this is the panel where if you pardon the expression with autonomous vehicles, this is where the rubber meets the road. Or for air vehicles, we're bringing the conversation down to Earth. And with that, Chuck Thorp and his introductory remarks.

[00:02:40.78] CHUCK THORP: It's hard to show a picture of a software agent, but it's easy to show a picture of an autonomous vehicle-- so we're taking advantage of that. The uncrewed systems aren't new. The Kettering bug was an autonomous vehicle in World War I. Kettering went on to found Delco and to endow the Kettering Sloan Memorial Cancer Institute.

[00:03:03.48] Of course, underwater autonomous vehicles are even older. This is just post-civil war. Until the advent of fiber optics, you fire these things and you let them go. And they were completely autonomous systems. And, of course, ground vehicles. This one happens to be radio controlled and this is a World War II vintage.

[00:03:22.24] So the idea of uncrewed systems has been around for a long time, but they've really proliferated in the last few years. Drones all over the place. The Navy runs hundreds of these, doing oceanography autonomously throughout the world's oceans. And, of course, Waymo and the other purveyors of automated vehicles. Self-driving isn't really new. GM back in the 1950s built the Firebird 1 and the Firebird 2 and had them doing self-driving experiments.

[00:03:56.68] [VIDEO PLAYBACK]

[00:03:59.76] - Now our tune correction.

[00:04:03.18] - We're coming in on the beam, Dan.

[00:04:07.62] - What do you desire?

[00:04:15.69] - Well done, Firebird 2. You're now under--

[00:04:18.86] [END PLAYBACK]

[00:04:19.74] CHUCK THORP: Automated control.

[00:04:21.13] [LAUGHTER]

[00:04:24.06] My own work goes back to the 1980s and the 1990s. This was the Navlab vehicle in 1990. Navlab 1 starts off on a mission to drive itself to Keith's house. Yes ALVINN--

[00:04:38.06] [VIDEO PLAYBACK]

[00:04:38.31] - ALVINN's neural network directly outputs steering angle for the Navlab steering wheel.

[00:04:42.48] [END PLAYBACK]

[00:04:42.87] CHUCK THORP: This has got a neural net, been trained to follow the road. It's got a laser scanner.

[00:04:49.30] [INTERPOSING VOICES]

[00:04:50.55] This is my son.

[00:04:52.69] [VIDEO PLAYBACK]

[00:04:53.03] - Scanning laser rangefinder is constantly looking for 3D objects.

[00:04:56.55] [END PLAYBACK]

[00:04:56.99] MAX PATTEN: He was inspired to get a master's in robotics and work for automotive safety--

[00:05:00.40] [INTERPOSING VOICES]

[00:05:00.63] [VIDEO PLAYBACK]

[00:05:00.75] - The emergency stop command is sent directly to the low level controller. The van halts, waits for the obstacle to disappear, then resumes driving along its route.

[00:05:09.63] [END PLAYBACK]

[00:05:10.28] CHUCK THORP: And then after that, in '97, we built a series of vehicles for the automated highway demo. This is a bus, two cars in a van driving by themselves. They've got everything you see on current automated vehicles-- lasers, radars, vehicle to vehicle communication, GPS, and maps.

[00:05:33.89] So that's the kind of background we've been working on. And we all know the story. Our vehicles have gotten better, our highways have gotten better, our drunk driving laws have gotten better. Our drivers have not gotten better. In fact, distracted driving is worse than ever.

[00:05:53.54] 90% of all crashes are caused-- and we don't call them accidents, we call them crashes, cause at least in part due to driver error. And in turn, when you have a crash that causes a large percentage of congestion.

[00:06:05.99] If you look at this chart from about here to about here, that's the automated highway-- I'm sorry, that's the interstate highway system. That's mandatory seatbelts, that's airbags, that's drunk driving laws. This chunk, that's cell phones. So our safety is getting worse.

[00:06:26.50] So we're going to fix it all with self-driving cars, right? Well, for safety probably eventually. I don't like the supervised full self-driving because we know that people don't pay attention. They don't pay attention when they're supposed to be driving. They pay even less attention when something else is doing the driving.

[00:06:44.77] Still, we're going to have to ask a bunch of good questions about how good is good enough about liability. I'll leave that to the lawyers and about actual risk. For congestion, yeah, maybe or maybe not. If you had a self-driving car, that would be great because it would reduce the chances of collision because the cars could drive closer to each other.

[00:07:04.64] But maybe you would live twice as far from work, in which case, all of a sudden you have twice as much congestion as you used to have. Could we enable better mass transit? How do we handle social equity for these kinds of things?

[00:07:19.24] When we were building the automated highway systems, it was funded by US Department of Transportation. And we had representations from the federal government, state government, car manufacturers, infrastructure builders, and the universities. Our incentive here was to improve transportation for everybody.

[00:07:37.45] Now, who's building the automated vehicles? The car companies. What's their incentive? To sell more cars. How do we get the public. Back in these discussions of how to improve transportation for everybody? NHTSA has a bunch of guidance. Their guidance is mostly, how do we sell more cars? And is not on, how do we handle social equity? So those are the kinds of questions I want to tee up to the later discussion. Thank you.

[00:08:10.23] KEITH GREMBAN: We've got Brian. Let's see. All right. There we go.

[00:08:20.10] BRIAN ARGROW: All right, put this in presentation mode. I think I only have three slides. So what do these photos have in common, these two photos? Self-driving.

[00:08:34.33] [LAUGHTER]

[00:08:36.66] No, there's a steering wheel right there on the horseless carriage. These are examples of disruptive technologies. Of course, those of us-- some of us remember Dolly, the sheep-- the first cloned sheep. This was back in the '90s. And everybody freaked out, like, Oh, no, what? What's going to happen? The next thing will be humans.

[00:08:58.60] And so, of course, there were no regulations in place to deal with this. And, of course, on the other side there, that's the horseless carriage circa 1800 that disrupted and turn everything on its head at that time. So again, this is about trust in the end when we're talking about these autonomous systems and AI. And so it becomes about trust.

[00:09:30.19] But there's always been a history of what these disruptive technologies do is oftentimes they are the result of innovation and they open the door for innovation. This came up earlier this morning, the discussion about regulating emerging technologies.

[00:09:50.88] And almost inevitably, regulations lag the innovation. And so sometimes you can get out ahead of the innovation. But most of the time, the regulations, you see what it does first and then it takes time to figure out how to-- well, what do we want to do about that?

[00:10:13.40] So one of the ways in which we regulate technologies is through certification. And so here and to certify is essentially to trust. So with humans, with your teenagers and so forth, they will take an exam, a written exam-- or we, I should say, because I had many years ago-- I did it.

[00:10:40.06] And that's followed by a driver's test. So a practicum to illustrate that you have mastered, not only do you have the knowledge to know the rules, but you can go out and demonstrate that you can operate with those rules.

[00:10:56.90] But if the instructor says, here's a hazard, deal with it. I guarantee you that if you had 10 students that day, every one of those students would deal with it a different way, a slightly different way. But what's important is the outcome.

[00:11:14.26] And so what the student would be graded on and what the license says is that we trust this person to know the rules. And in situations that will come up, there will always-- well, there will be an acceptable outcome. So that's what we do with licensing.

[00:11:37.06] Now what about machines? Well, if we use aircraft as an example, there's type certification. So different types of aircraft are certified based on their design. But there's also components that go into the aircraft that are certified based on their performance-- their performance standards.

[00:12:00.14] And so FAA, for instance, is slowly moving towards performance. Because this goes back to an earlier discussion this morning about not knowing what's going on inside the black box. I mean, it's to the point now with automation.

[00:12:15.56] And I said automation, not autonomy, that people don't really understand what is going on inside the body. Why should you? Why should you-- if you have an autonomous or an automated system in your aircraft, why should you have to learn how you get to the result that you get?

[00:12:33.26] It's about the outcome. It's about the performance. And so I would argue that it's fine having AI in a box and not knowing what's-- so long as you understand the outcomes. And what about the human plus machine? This is where we're having the problems, really having the difficulties.

[00:12:52.75] How do you certify? So we already said we certify humans with a license and we certify components against standards-- performance standards. What about the human machine combination? If you think about a of the accidents that have occurred in aviation over

the years-- recent years, is because of some degree of disorientation or lack of situational awareness of the human and dealing with some automated system.

[00:13:26.32] How do you certify that? Do you certify the human plus machine? And then there's AI. So there's a machine involved, right? It's running an algorithm on a machine. And it's emulating, to some degree, intelligence. So how do you deal with that? Do you license and certify? So I'll leave it there. I think that's it. Yes.

[00:13:57.06] [APPLAUSE]

[00:14:02.37] KEITH GREMBAN: Next, we have Marc Canellas.

[00:14:04.41] MARC CANELLAS: Fantastic. Cool. Wonderful. So I have three little lights which I've done some appellate arguments, and now I feel I must say good afternoon, your honors. May it please the court.

[00:14:23.43] I'm so proud to be here with engineers. I think the rubber meets the road is exactly what we should be talking about, because when physical systems are used to make decisions that will affect your life, your liberty, people tend to care a little bit more than a grocery list, or even in situations earlier when you're talking to customer service.

[00:14:48.10] I make no bones about it. I think we've failed as a regulatory state to the extent we have one. We know what steps we need to take. We act like we don't, but we do. I don't mean to solve the problem of AI governance or governing autonomous vehicles in this case.

[00:15:03.62] We could at least test them. We could at least run and have an infrastructure where we put vehicles or aircraft through really complicated test scenarios, such as basics-- snow, sun in your eyes. But what happens is these vehicles are being deployed and I would argue beta tested on the American population.

[00:15:26.07] So when your car all of a sudden loses its ability to see and detect obstacles because the sun is right at an angle and it crashes and you're injured, that's a new data point that you developed with your life on the line.

[00:15:41.05] And so my PhD work was in human machine interaction. So exactly that point that Brian talked about. And I can't say we can solve it, but we can at least work towards a real solution. And the reason we have a lot of failure is I will use a quote of the answer to all your questions is money. If it didn't make money, they wouldn't do it.

[00:16:04.57] I will move quickly on to system accidents. So I also did system safety work because I worked in aerospace. And aviation accidents are front page news whenever they happen. The thing is, with AI systems or autonomous vehicles is that the types of failure change.

[00:16:21.68] So if I were driving my car that's not autonomous and I crash, it doesn't really speak necessarily to whether you are driving your car will crash. But if we are all using the same software, then if it was an environmental condition, maybe all of our cars fail or if it's a software where they're pushing an update like Tesla seems to often do in real time.

[00:16:43.35] If that system fails, your car is updating in real time as you're going down the road and an entire fleet of cars all shuts down. That's a problem. So the types of failure are going to drastically change. And then I turn to surveillance. This is what I do regularly now.

[00:17:01.26] So I was a PhD engineer and then I worked on the Hill as a staffer. So I'm happy to talk about political and policy things. But I work with AI systems that are used to surveil, incarcerate, and convict us, citizens, Americans.

[00:17:18.89] And what I see in these cars is actually a really powerful system where all of these cars have cameras on them, have location information, have even sometimes Wi-Fi networks where you can connect to them, which in my world means that when the government wants to get information about you, they can and they will.

[00:17:35.98] And so this is language that I propose will happen. I haven't seen it yet, but in something that we would call a geofence warrant, which I litigate, which means let's say there was an incident around this building.

[00:17:48.13] What they would do is they would draw-- the police would draw a little box around the campus here, and they would say I want all the phones that ever connect-- that were ever located in this place. And they will get that. And it will tell you, depending on what court you're in, don't have any expectation of privacy in that information.

[00:18:06.51] And even if you did, there's called the good faith exception for all my lawyers out there, which means it doesn't really matter anyway that they violated it. But the same thing is going to happen to vehicles.

[00:18:16.45] So I can imagine them using that same language and say, I want all the Teslas and Waymo's and all those types of vehicles which we know have cameras on them, and I want all that camera information. Then we talk about facial recognition, all those things. So I think this will actually superpower in a lot of ways, the surveillance state, unless we regulate, unless we have rules as to what is allowed and what is not allowed.

[00:18:38.71] So, I will move quickly because of course, I only have a few seconds left. What is an engineering talk without xkcd? Basically, I'm talking in praise of constraints. To me, good engineering is dealing with constraints. So if anyone-- I'll let people finish.

[00:19:01.39] It's in backwards-- it's alphabetized, right? Reverse. It's pretty cool. Yeah, it's very cute. Very cute. So the idea is that there are AI accelerationists out there. I'm obviously not one. I believe in constraints, because I come from the aerospace community.

[00:19:20.59] So this is a similar thing from the xkcd guy talking, where he had to describe the Saturn V, with only the top thousand most common words in the English language. So that's how I get to up go for five. The end of this point should point towards the ground if you want to go to space.

[00:19:39.86] Good engineers, just like good lawyers, just like whatever world you operate in, you are good at your job. And how you determine who's the best in their job is because they solve hard problems. And I am honestly exhausted by those who believe that we do not need to have any constraints on this technology because it will harm innovation.

[00:19:58.80] To me, innovation is solved by good people who can deal with constraints. And I close with the reality that in space, the reason Apollo, which took this rocket to the moon, was so amazing, was because space wants to kill you.

[00:20:15.74] That sounds crap, but that is literally what space wants to do. So it is actually the amazingness of that program was the fact that they solved that problem, not ignored it. And I think we need to stop ignoring this problem and embrace it. Thank you.

[00:20:31.15] [APPLAUSE]

[00:20:42.43] KEITH GREMBAN: So with that as a start, let's maybe start talking about some of the constraints and differences in dealing with regulation when you've got autonomous vehicles. Anybody want to start off?

[00:20:57.79] CHUCK THORP: The ground vehicle space has always been weird because we have 50 state departments of transportation and 340 metropolitan planning organizations. And even though there are AASHTO rules as to what a highway should look like, it's very easy to get exceptions to those rules.

[00:21:17.87] So it's very hard for an automated vehicle designer to know what the environment is they're designing for. And then it's hard because National Highway Traffic Safety Administration deals with smart cars, federal highway deals with smart roads. Who has the two of them talking together? This has not been terribly well coordinated and continues to cause mischief and difficulty in trying to come up with good solutions.

[00:21:47.30] KEITH GREMBAN: Brian.

[00:21:47.66] BRIAN ARGROW: Well, part of the idea of regulation is what is acceptable to the community. And an example of that is we had

the plane crash at Reagan National a few weeks ago. I think it was 67 fatalities. And that was the first instance of fatalities in the US airspace in some well over a decade.

[00:22:24.09] And yet, that was front page news for a number of reasons. But that was front page news for several days. And in 2023, 40,000 people were killed on highways. And so the question then becomes, what is acceptable to the public?

[00:22:47.47] And it's acceptable that we license human drivers in this country and we deal with the carnage. We accept the carnage of 40,000 fatalities, but we don't accept that with aircraft-- with commercial aircraft.

[00:23:05.41] Now, I have to say that general aviation is somewhere in between the commercial aircraft in terms of the fatalities. So it's not a simple matter of right and wrong. It's really a matter of, what do we accept as a society?

[00:23:29.14] MARC CANELLAS: And I think that's, what will we accept? It goes back to the question about liability for autonomous vehicle accidents. And I will tell you, there are lots of law professors who specialize in this. And I'm not one of those.

[00:23:42.73] But I think to your question is, why do we look at it differently? I worked with people who investigated aircraft accidents. And I would suggest that it's because we can see ourselves on those planes.

[00:24:04.56] Everybody thinks they're a great driver. But when you're on a plane, you're in someone else's hands. And so I think that's where a lot of the different views-- I don't drink and drive. I turn on my signal appropriately. It's not going to happen to me.

[00:24:19.07] And so I think that's where it comes. But the problem is-- and I will posit this because, and Brian and I were talking earlier, is imagine you knew for a fact that if we just went to full self driving across the country, we were going to reduce it from, let's say 40,000 to 1,000. Awesome, right? Everyone would take that like that change, right? People would take that change.

[00:24:42.12] But for the 1,000 where a company made a design decision that did a trade off, that caused it to say I don't want to hurt this person, but I wanted-- I intentionally chose to turn right and hit your child.

[00:24:59.51] There is something that feels at least in me, and I think a lot of us that feels somewhat unacceptable. But that is the trade off we're making. And so I think we're so used to a world, I mean, I literally deal with this as a criminal defense lawyer, where we look to humans for accountability. When we're talking about a design, whether it's an agent or an autonomous system for accountability, we feel left wanting.

[00:25:22.91] And so I think those are the kinds of questions that I think when you do have autonomous vehicle accidents, they will go more like aircraft in terms of the public concern than individual driving. And so that may spark more regulation. But I think it's because we need someone, whether I call it blame or accountability, having a human there is so valuable to us as humans.

[00:25:46.69] CHUCK THORP: Some of this has to do with perception of risk. The provost of Carnegie Mellon and the president were not allowed by the trustees to fly on the same flight. The provost thought this was terrible because if there was a non-stop and a connecting flight, you knew who ended up on the connecting flight. But they were allowed to share a cab to the airport.

[00:26:06.88] Statistically, they were much more likely to be killed in the cab to the airport. But it's that perception of risk. Another just quick little quibble. We keep talking about fatalities. General Motors came to us and said fatalities are important. We want to reduce fatalities.

[00:26:23.23] But if all we do is convert fatalities into serious injuries because we've reduced the velocity of the crash, it sounds weird. But we know how much a fatality costs us. A serious injury could cost us 100 million bucks.

[00:26:39.55] It's unlikely that those 39,000 people whose lives were saved will walk up to GM and offer them $100 million bucks. It's pretty likely that those 1,000 people will sue them. So there's funny ways that our economy works also.

[00:26:58.17] BRIAN ARGROW: In the earlier days when drones were-- we were looking at how they would be integrated into the national airspace, we would commonly say that if there was a crash, if there was a body in that aircraft, it wouldn't make the news.

[00:27:21.42] That crash didn't have a body in it. You better believe and it maybe kill someone or cause a fatality externally. It would certainly. That's a whole different ballgame there. So again, is that the perception. If you're counting the numbers-- I mean, we say there was one fatality either way. But the fallout from those two types of incidence would be completely different.

[00:27:44.53] MARC CANELLAS: And just continuing on this conversation, perception of risk is something there's both an individual aspect of, of what I am willing to-- what I'm willing to risk that should be communicated, whether it's through a notice or something like that.

[00:28:00.67] But there's also a societal expectations of risk, of what we accept. And so I think that's also where public policy making regulation is a way for us to collectively say we're OK with this, we're not OK with that.

[00:28:16.87] And so the lack of regulation and the lack of real governance, whatever that comes out to whether it's-- we talk about

soft law or hard law or whatever you want to talk about, that should be something we collectively and intentionally decide rather than what happens, which is an accident occurs.

[00:28:33.34] And then lawyers litigate these collective action problems. And I think that's just-- I mean, even speaking as a lawyer, that's not how we should do this. We should have smart people coming together, as you said, Chuck, in your initial thing, coming together to think about these questions rather than leaving it up to a single court in one state. I don't think that's the right way to do it.

[00:28:54.06] KEITH GREMBAN: So when I was working with Chuck on autonomous vehicles, I always felt that you're not going to mix human-driven vehicles with autonomous vehicles because it's too hard to predict what a human is doing. And yet we're doing that now.

[00:29:10.35] And there's human machine interfaces at a number of levels. There's the fleet level where you've got autonomous vehicles in the same traffic area or aircraft in the civilian airspace. They're autonomous. But you've also got the human interface between the vehicle and the pilot. And I think this poses some interesting problems too. And I know you've got some experience in looking at drivers, human machine interfaces.

[00:29:36.39] CHUCK THORP: Well, when we were doing automated vehicles, we were also doing the prototype for the run off road collision countermeasures. We had the contract to write the specifications for if you start to get sleepy and drift off the road, how quickly should you be warned and how should you be warned?

[00:29:53.34] And this is one of the most satisfactory parts of my career, is to see that now in production cars saving lives. But the questions turn out to be tricky. We put people in a driving simulator. We distracted them by asking them to look for a cassette tape that tells you how long ago this research was done.

[00:30:14.95] We put aside gust, which would cause them to go off. Then we tried different warning things. Should you beep at them? Should you shake the steering wheel? We thought maybe nudging the steering wheel. And for American drivers, you nudge the steering wheel, they said, oh, I better steer back that way to get back onto the lane.

[00:30:31.99] Our colleagues at DaimlerChrysler in Germany did the same experiment for German drivers. If you nudge the wheel, they say, oh, I must have hit a rock and they steer the other direction. Humans are tricky, and trying to figure out the right mixture of human-computer interaction is tough. It's much easier if it's fully automated and you don't have a human computer interaction in there.

[00:30:55.15] BRIAN ARGROW: So one of the things that is the need to distinguish between, for instance, autonomy and automation. My wife's

Subaru, I guess it was a 2018. So we're driving it-- and I have a driver's license.

[00:31:17.20] But I had a novel experience in driving that vehicle that literally happened driving down the highway. So I'm driving down the highway and all of a sudden, I feel there's something on the steering wheel. And I'm thinking, what the hell is that?

[00:31:31.43] And so I had no clue because I didn't go through the driver's manual. How many of you actually you buy a car and you actually go through the driver's manual? I go through the driver's manual when something goes wrong, and I'm trying to figure out, OK, now what just happened?

[00:31:48.22] So it is exactly what you were talking about. Subaru has that integrated into the system, and it has not been enabled since I turned it off I don't know how many years ago because, again, this is the human machine.

[00:32:04.66] And this was an automatic system. It didn't see the hazards I was seeing. It could see the line because it has the eyesight. So it could see where the line was. But that's all it was reacting to. And so I could see not only the line, I could assess the hazards and all the other things that were going on at the same time.

[00:32:27.46] So why should this car tell me-- try to tell me to, like you said, steer right? There might have been a car to the right. This the reason why I didn't steer it to the right. So again, it's the situational awareness and the distinguishing between automated systems in which situational awareness of the humans can really suffer.

[00:32:51.88] You put an autopilot on, the autopilot, again, is not reacting to hazards. It's looking at the altitude, the speed and all that, and it's holding it there. And then if something goes wrong, all of a sudden, as a pilot, you have to figure out what's going on, because you haven't been flying the plane for the past hour or so. So autonomy versus automation.

[00:33:13.93] MARC CANELLAS: Yeah. One one of the things we see and NHTSA, who is the National Highway Transportation Safety Administration, which de facto regulates the vehicles on our roads, points to the aviation community as looking to them to see, how do we properly integrate these human machine systems into complex, safety critical operations?

[00:33:40.09] And one of the things that we have found in the aviation community is that you reduce the number of failures, but the failures are more extreme than they were before. And I'll give an example. So it used to be when you were flying and you set your flight level-- so you're flying at 37,000 feet, you were within 100 feet or so of that location-- that altitude.

[00:34:08.36] So if you happened-- a human happened to enter in their flight level incorrectly, so that two planes were at the same flight level, that's bad. The likelihood of them actually hitting each other was somewhat low because just you couldn't get the accuracy that well.

[00:34:28.02] But now I can fly at exactly that altitude. So now what used to not be a problematic human machine interaction problem is now a catastrophic human machine interaction problem. So these are-- it raises new questions, new problems.

[00:34:46.44] And I'll quickly talk about the Uber accident in, I believe was in Arizona, where there was an autonomous Uber vehicle driving. I'm sure you guys can speak more to the technical side of this. But the report that came out afterwards was that a driver was technically monitoring the car, as it was the SUV for Uber as it was driving.

[00:35:09.39] It struck a woman who was crossing the road with her bike at night. The driver, they said-- Uber was explaining that it was clearly human error because she was watching the voice as this was going on. That's a problem.

[00:35:27.02] What it didn't talk about, or at least what Uber didn't talk about, the NTSB did, was that the entire culture around this was that don't worry, the car will see anything and everyone. So you've conditioned this human who's supposed to be monitoring a task, which, as you're correct, we're not good at to begin with.

[00:35:44.43] We're also conditioning them that everything's going to be fine. Don't worry. We designed a super cool car. You're going to be fine. And so you're conditioning the human, you call it autopilot, which has work-- it has meaning in our life.

[00:35:57.68] And so this person is listening, watching the voice, not particularly paying attention. A person comes and essentially jaywalks and the car hits them. They eventually look at what actually the car was designed on.

[00:36:14.62] The car was designed to see pedestrians in crosswalks. It was actually-- it's sort of like if you've heard of the famous-- is it a husky or not? And it's actually looking at the snow in the image. It's identifying the dog based on the snow behind it, not based on the actual dog. The car was trained to see crosswalks, not people.

[00:36:34.63] So again, these are failures that you never would have seen with a human, maybe if they certainly weren't paying attention. But just the idea that it would just not at all see a human in the road that's standing right in front of them. But that's because it wasn't looking for them. And so it's just to speak that these problems, they are complex. There's stuff we should be doing. But yeah, the human machine interaction is a very, very, very difficult problem.

[00:36:59.41] CHUCK THORP: And they had disabled the Volvo radar because they didn't want to test that. They only wanted to test their system. Volvo radar would have stopped.

[00:37:11.53] KEITH GREMBAN: So this is at the individual person and vehicle level. But then, like you said, there's this fleet level. And I know, Brian, you've done some work on, how do you integrate autonomous aircraft in civilian Airspace What are some of the issues that we have to deal with on this larger, more macro level?

[00:37:30.64] BRIAN ARGROW: Well, that issue, and I guess in a similar way to automobiles, that if we drew the line and said, OK, all vehicles from this day will be autonomous, meaning that they would communicate with each other, they'd have collision avoidance, all these things, we probably wouldn't have a fatality other than maybe running over a few pedestrians.

[00:37:56.14] But car crashes probably could end today. And again, it's the mix. That's the problem. And with aircraft, it's a similar thing in the sense that one of the challenges is with the pilots, particularly general aviation pilots, who don't want sometimes you to know where they are in the sky.

[00:38:25.63] So they don't want to be broadcasting their location and so forth. And then you have the completely uncooperative aircraft, balloons, things like that. You go tell a hobbyist balloonist that they have to put a transponder on their balloon and they'll tell you--

[00:38:43.53] And so there's that whole issue is that even if you did draw the line and say, OK, from here forward, you have until this day to retrofit your aircraft. Who's going to pay for that? The pilots are going to demand, hey, I might do it, but the government, if you're telling me to do this, you're going to pay for my retrofit.

[00:39:05.99] And so it's this reluctance or this American individualism, where I want to go get in my car and drive it, damn it. And I'm going to do it when I want to do it. Or I want to go fly my aircraft, and I may not want to turn on the transponder. That's one of the things we had a problem with.

[00:39:27.36] We do weather research with-- atmospheric science research with small drones, with the small uncrewed aircraft systems. As an example, we had a project this past summer off the Eastern shore of Lake Michigan.

[00:39:46.94] And we had a small drone that-- very simple. It would go out over the water, about a quarter of a mile out, and then it would go up to a couple of thousand feet-- well, 1,500 feet. So supposedly-- so we have this-- now there are apps where you can find out where aircraft are. You can look at your app and they're the ones that are transponding or what's called ADS-B.

[00:40:14.69] And so we're out there, but we have to keep an eye out because just so happens we see a C-130 and it's inland. And so we're watching it, and we're watching all these other aircrafts. And we have issued what's called a NOTAM, a notice to-- it's now changed from Notice to Airmen to now Notice to Air Mission. And so that's required.

[00:40:35.43] So the information is there that we are flying this drone at a certain part of Lake Michigan at this time. You think they looked at that NOTAM? They came blasting through. So I actually have a video of it. And I hear myself-- I went back and looked at that and I hear myself talking. I said something about, of course, they come blasting through here with paying no attention. So there are the humans.

[00:40:58.75] If that had been an automated system, it would have known. It would looked at the NOTAM and we wouldn't have had that issue. And the onus is on us as the-- because this is the way the rules are set up right now. We get what's called a certificate of authorization from the FAA.

[00:41:15.47] And it puts the burden on us to-- us, meaning our group, our crew, to do the deconfliction. So they didn't get a ticket or whatever you get, because that's another problem with FAA is, how do they enforce these things?

[00:41:32.21] And if they weren't the only ones, but they were the biggest ones. They were flying low. And so we, of course, brought our drone in and let them pass by because they would have-- they flew literally right? Literally right over the spot where we were-- our spot where we were going up and down with our drone.

[00:41:49.65] CHUCK THORP: The same thing happens in ocean-going-- ocean-going ships are required to have a transponder. You would be amazed at how many times a bucket ends up over that transponder just as they get to unauthorized fishing grounds.

[00:42:05.46] BRIAN ARGROW: So if you remove the human, if we could take-- and that's the thing about these autonomous systems, there's always a human. If nothing else, the human is there to task the autonomous system, to go do something without me having to watch you. But there's always a human involved. And so this idea that you can somehow separate the human from the autonomous system, that's not happening. That's called the terminator.

[00:42:31.48] MARC CANELLAS: And I think the problem with some of these, especially coming as the four of us do from aerospace and ground transportation is that it can't really be federalized. We can't have an airspace where every airspace is controlled by every state and you have to use a different-- hypothetically, a different transponder.

[00:42:54.55] And you have to change all that every time you cross into a new airspace-- imagine a truck driver who's going across the country, and they have to check in and stop and change out parts of their truck

every time they go to a new city or a new state or new whatever. That breaks down our economy. So these are questions that despite what some may not want any governance, are inherently governmental decisions that have to be made collectively.

[00:43:21.94] So I think that's why-- I think sort of that's what this conversation brings, is that there are some things that maybe it makes sense to wait. These decisions have to be made at some point. And we just aren't doing that.

[00:43:33.21] CHUCK THORP: Although there are some legitimate differences. When we did the automated highway demo was in San Diego, we all prepared for following painted white stripes. They don't have painted white stripes. They have botts' dots. They have these raised pavement. We couldn't have those in Pennsylvania. The first snow plow that came along would scrape them all right off the road. So you have to be adaptable to differences.

[00:43:57.57] MARC CANELLAS: And can I add one other thing? Talking about the value of-- and the value of a human is imagine-- you can just go outside. Imagine you're about to cross the street right out here and there's a car coming. You have a crosswalk. There's not necessarily a stop sign.

[00:44:13.12] Do you look that driver square in the eye? That's what I do. You wait until they see you. You may take a step out, but you're like, at some point, if they're looking down, they're going to look up, they're going to make eye contact and then I cross.

[00:44:28.17] I'm seeing some nods that that's like an inherent thing that we do when we teach our children to look both ways and make sure they're stopping. But with an autonomous vehicle that doesn't have a person sitting there, these social norms around how we interact with these vehicles will fundamentally change. And I don't know what the answer is there, but I think there's going to be like deeper changes to how we just like even cross a road.

[00:44:52.14] CHUCK THORPE: How do you make eye contact with an automated vehicle?

[00:44:54.63] MARC CANELLAS: Yeah.

[00:44:57.20] KEITH GREMBAN: Well, since you brought up the issues of regulations, what kind of recommendations might we have for how to license and certify and what kind of policies and regulations do we need to start formulating to enable human autonomous vehicle interactions?

[00:45:17.51] BRIAN ARGROW: Well, I think that one of the things that if we truly and again, the question is, what's acceptable. Right now, it's acceptable for 40,000 people to die on the highways and one or two every decade, in the airspace system. So I think we have to start with that because it's unacceptable. We have the technical solutions. If it is

unacceptable for aircraft to run into each other on the taxiway or on the runways, or if it's unacceptable for cars to collide and things like that, I believe we have the technology to prevent that from happening. But you have to give up what we consider to be freedoms.

[00:46:05.74] You have to get in your car and it has to communicate with everybody else's car. So you're not going to be able to floor it when you want to or it has to be in a system in which there is you have to agree to some giving up control. And so the technical solutions are there. But it's a question of will. Do we have the will to move to fully realize the technologies?

[00:46:34.21] And again, we insist on being in this meshed human machine mush up. And that's where we see the issues.

[00:46:44.88] MARC CANELLAS: And I'll add that. First of all, I broadly agree. I think a lot of this that we've talked about literally all day is a matter of political will. We not that we could, again, solve it. There would still be incredible challenges going forward, but we know the first initial steps and who will what types of engineers, we need to bring in the room, what other types of specialties we need to bring in the room.

[00:47:08.62] I will add something that I think we really need that I sort of I prefaced in my slides, is that when we have drones in the airspace that are going to be ideally, I guess, you would say you would have little corridors for like drones and things like that means that they're going to be a lot more cameras, a lot more ways that you can be tracked. And that is another decision we have to come as to what are your rights if there are cameras around you all the time?

[00:47:37.45] Because there's a discussion I learned when I was in law school is there's a right to privacy. But really, what it really, really is a right to anonymity. That is a critical part of each of us, is showing you a little part of me and you, a different part of me. And I think in addition to the technical solutions of like, how do we make sure these systems are safe from a physical standpoint is also the questions about privacy.

[00:48:03.82] And I think to your discussion about getting in a car and driving is like so American. It is inherent to our pushing down the accelerator as fast as you can, especially when you might be a teenager, is like a critical part of Americana. And so there are these things we have to really think through as to what we're giving up. And if we're OK with it. But all of that is we deciding. And I think that's really where the, I'll say it, the rubber meets the road.

[00:48:31.80] BRIAN ARGROW: Do we really need turbos?

[00:48:33.12] CHUCK THORPE: I mean, get some knots out there. I like my turbo

[00:48:37.49] MARC CANELLAS: Small area where I'm curious if we could make progress. I would love to see the near-miss data from

Waymo and the near-miss data from the other automated vehicle testing systems. And I would love to make that near-miss data public and put the best minds in academia and industry to work on how could you have turned these near-misses into safer situations? They view that as their proprietary data. But for the good of the public, wouldn't it be nice if that were shared?

[00:49:16.70] BRIAN ARGROW: No, I was about to get political, but I'll pass on that.

[00:49:23.12] CHUCK THORPE: No, I think I agree. I agree.

[00:49:25.49] BRIAN ARGROW: I agree, yeah.

[00:49:29.09] MARC CANELLAS: Who owns one of these companies?

[00:49:31.46] BRIAN ARGROW: Yeah, that's where this idea that, and I was there's something I made a note of this morning about. We keep using the word "we". If we do this, if we do that, who is we? In terms of-- I think I knew who we used to be, but who is we? Who is going to-- where is this we that's going to get together to hash out these issues and come up with regulations and things that are sensible. I don't know that there's a we anymore.

[00:50:03.78] CHUCK THORPE: Now that we're going to go down this path, but not in that way. But when I was working with some business lawyers and was being I was going to classes dealing with business is one of my professors always just reiterated to us because there are all these rules around business governance. We heard Nathaniel from this morning talk about business structure. And what he taught us was never ask a private company to do anything other than make money.

[00:50:33.76] Literally what they're designed for. They are incredible. American corporations are incredibly good at making money. Like the best in the world at making money. But if you want them to do anything other than that, such as be safe, deconflict appropriately, share your data, do these other things that I'll use the dangerous word we want, whatever that we is, that's where government has to step in. And so I think relying on the benevolence of corporations, I think is misguided. And that's really where we need, whether it's city council, local, state, or federal really need to step in.

[00:51:15.73] KEITH GREMBAN: I guess the yellow light is on, which means it's time to go to audience questions. So the first one comes from a AUDIENCE and I saw your hand first, so.

[00:51:26.01] AUDIENCE: Sure. I'm big recreational cyclist. And already with drivers for a variety of reasons. Distracted drivers, some people with road rage. I think lots of my friends and second order, almost everyone knows someone who's been hit by a car. And already drivers aren't held accountable for that as human beings.

[00:51:51.25] And so on one hand, I think autonomous systems could be better for this, absolutely. But on the other hand, a Mark brought up the Uber Arizona example. Is there a case in-- most of these are now using some kind of neural training data. How do we make sure in the design and the engineering and the training data that pedestrian safety, cyclist safety, and urban areas for recreationalists is a priority?

[00:52:14.04] KEITH GREMBAN: And how do we test?

[00:52:15.48] AUDIENCE: That too.

[00:52:17.83] KEITH GREMBAN: Brian.

[00:52:19.53] BRIAN ARGROW: First of all, when we consider we're talking about a very few incidents where these autonomous vehicles have hit pedestrians. Denver there's probably a pedestrian hit every day. And so, again, it's a question of what is acceptable. And I think we're agreeing that it's not acceptable for autonomous or self-driving systems to hit people in crosswalks and things like that.

[00:52:53.65] But what's acceptable? Why do we make that distinction? Why do we allow hundreds people or whatever it is to get killed in Denver on the streets pedestrian? Denver is particularly bad, by the way, in terms of number of pedestrians that are hit and killed. And yet we make it seem like it's such an unsolvable problem with autonomous systems.

[00:53:20.09] And if we agree that maybe 10 fatalities from autonomous vehicles is acceptable, then we wouldn't be wringing our hands in terms of, oh, we can't put these things on the road because they may can't kill 10 people across the entire United States in a year when humans are just mowing people down at a much higher clip. So again, it goes back to when we started this discussion about what is acceptable, what are we willing to what risks are we willing to accept?

[00:53:54.84] MARC CANELLAS: And you're asking a very difficult question about how do you validate a neural net. And I don't have the answer to that. The tougher question in my mind is not do you see a pedestrian, but do you see a bouncing ball. And know that there may be a little kid chasing that kind of semantic level reasoning is very hard to explain, and let's pray that we don't have a lot of training examples of that kind of thing in our database. So we need to figure out how to incorporate those kinds of things without having learned from mistakes.

[00:54:30.13] CHUCK THORPE: And I love-- I think that's a-- I had never thought about it. That's exactly it. You're coming to what you know to be a dangerous intersection. Not that there's not a sign that says it, but you know this is like a place where people tend to speed or people tend to turn without thinking, and you like, slow down a little bit more or you're in like a parking lot and things tend to be weird there. So you're, yes, you have your lane.

[00:54:51.83] But I will say you use the word priority. And so I'll use another phrase, which is if you show me your budget, I'll tell you your priorities. We could, theoretically, even outside of the autonomous vehicle argument, we could have separated bike lanes. We could do that. We could also have a lane.

[00:55:10.23] I don't know if it's here in Denver, but at least where I live in Alexandria, Virginia, there's a separate for the buses. There's a separate lane for them because we're prioritizing that. But you could almost say, like we in order to encourage autonomous vehicles or to keep them safe, have them be on their own separate area. Like there are things we could do about how we do the designed should there somewhere like the designed environment to really encourage autonomous vehicles because we say, well, we haven't designed them to account for all this stuff.

[00:55:36.98] So we'll like give them special areas that almost like an express on a highway that they're separated from other people. So it's a little bit easier. So there are things we could do to encourage that and help the onboarding and this transition be a lot smoother rather than just deploying them everywhere.

[00:55:55.18] BRIAN ARGROW: Well, I guess also to Chuck's point about the data, I suppose that the systems will get better as these things, these accidents, and so forth are part of the training data. Nothing's for free.

[00:56:15.90] KEITH GREMBAN: Hey, let's see. Middle green.

[00:56:23.10] AUDIENCE: All right. Hello. I'm not sure that just because we got 40,000 people dying, this means this is the solution we should have. It seems like a false binary. We've been testing these, but only in small scale. We don't know what'll happen if we unleash this. It seems like other things could happen.

[00:56:40.44] If people get in trouble for getting an accident. Because they're on their cell phone, we could treat it like a DUI. We can make it hard for them to continue to do that. We could tell make it force teens who just begin driving that they can't use their phones. They can have to download software that would force them to deactivate the phone. It's very easy to do.

[00:56:57.67] Another thing is you it opens up this insurance problem of plenty of philosophers, myself included, have talked about the problem with the trolley car dilemma that it introduces with insurance rates, because then the car is potentially in certain situations where it has to choose between saving you and saving several other people. These are new problems that get introduced that aren't there so much with the traditional system that we have right now. So I'm just not sure that Brian's solutions really are ones we should be--

[00:57:28.00] It seems to me like a problem. It's a solution without a problem. I mean, the fact that only 40,000 people are dying, that's kind of amazing when you think about it, that's a very low number.

[00:57:40.08] BRIAN ARGROW: I agree that is a-- when you consider the people that you know that actually have a driver's license that are allowed to drive, but it raises the issue of you could start from the licensing side. What has changed since I was a kid, and took a driver's test. Pretty much the rules of the road are pretty much the same. But all this autonomy.

[00:58:03.25] So I'm not licensed to operate my Subaru, push back on the steering wheel thing. Maybe I should be. Maybe there's some intermediate solutions here in which the requirements to for a license are increased so that now you're forced to at least demonstrate that you're aware of the technology that you have in your vehicle or whatever. That could be--

[00:58:27.64] So what I was talking about is kind of like the extreme. If we just drew a line and said, OK, now no more driving your own cars or whatever we're going to. There are intermediate, potentially intermediate solutions that could start with licensing and license the human machine system and not just the human. And then you certify the machine.

[00:58:50.31] MARC CANELLAS: But that whole business of shutting off texting while you're driving, technically that's easy. It's a question of, do we have the social will to enforce that?

[00:59:05.97] AUDIENCE: They could mandate you have this on. And they could also mandate it if you get an accident because of it, then you just like having a DUI. You can't drive without deactivating your phone.

[00:59:15.50] CHUCK THORPE: And I will say all of those, because I agree with you, often a binary choice. Either take it or leave it, as we heard in some of the earlier sessions, is often a false choice. But all the answers here, as you're all of like the solutions you went through, whether it's requiring an app or requiring people to be recurring licensing. I do not know the words, but pilots have to every year. Just like a lawyer, just like a lot of professions have to constantly be recertified, continuing legal education, things like that.

[00:59:44.93] I guess CLE should not be included in that as actually doing anything. But you could have a regime like that says these are the new tools in cars that we expect you to be able to understand, whether it's the shaking steering wheel or whatever. This is what it can can't do. But the only way to do that is through a government entity. And so that's really where the problem happens.

[01:00:08.47] KEITH GREMBAN: OK, we've got a-- wait for the mic.

[01:00:16.00] AUDIENCE: My name is Dan. I just wanted to ask a really simple question since Mark has been talking about government and since Mr. Thorpe also pointed out just how much time and effort can be involved in the complexity of, well, transportation department. It's a pretty simple question. Where, if at all, does data-driven policy come in with autonomous vehicles reporting in and more than just near misses? What else, if anything, is going to be of value to being able to allow government to regulate business? Even though they're all about making money, it might be that they could be all about making money and saving lives. So it's an open question.

[01:01:03.91] MARC CANELLAS: NHTSA is very good at investigating fatal accidents. I'm sorry. Not accidents, crashes. They say they're not accidents because you can say what caused them. So NHTSA has a very good database for fatal, fatal crashes. And then there's work based on that analysis to figure out how to mitigate it. Exactly how you go about it, what regulations, what technology, et cetera is the open question.

[01:01:32.55] BRIAN ARGROW: So one of the things we talked about in the integration of drones into the National Airspace System was that insurance companies would have a lot to say about that in terms of what would be allowable. I know there's a I've seen this on a TV commercial. There's something that progressive has this device that you can put in your vehicle and it tracks your speed, all these other things.

[01:02:01.15] And so once again, I think the insurance companies could potentially take the lead on these things. And if you want your vehicle insured, then you have to provide these data for monitoring. But again, it goes back to the whole thing about privacy and so forth. Do you really want your insurance company to know that, especially if you have a turbo, how you drive. So again, it's those types of issues.

[01:02:30.54] STUDENT: It's a question. It's a question. I think that the manufacturers full-time connected vehicles might be asked that same question.

[01:02:38.17] BRIAN ARGROW: Well, Mark Martin has already said about the police, will they draw a box around this area and you think that your data was is not available to them?

[01:02:47.55] MARC CANELLAS: The Turnpike authorities are not allowed to use your check in and check out time to send you a speeding ticket. Yeah so how do you generalize from that? or do you say maybe they should be allowed to.

[01:03:00.48] CHUCK THORPE: And I will say to the statement that your driving location is not personally identifiable. I guess I will push back strongly and say that if I know a vehicle that stops at a certain location overnight, then in the morning drives to another location, stays there from 9:00 to 5:00. And then goes by a school that's an

elementary school and then stops at that original place overnight. Again, I know home, I know work, I know elementary school. And it doesn't take a lot of steps to go from there to identifiable.

[01:03:36.09] And it doesn't necessarily mean, police. It could be all kinds of people. And so there, but again, I agree with you because like that kind of stuff would be useful from a safety perspective. And I will say one thing I wanted to mention was there's something called the Aviation Safety Reporting System, or ASRS, which is a confidential place where pilots can report near-misses. So nothing happens. So it's not going to be on the news unless I guess it happens at an airport. But they just report those kind of things, and so that it's a NASA system so that NASA and the researchers can say, hey, there wasn't an accident, but we had a bunch of these types of near-misses, and it's not bad yet, but we think it might be. So there are data like that would also be super, super useful.

[01:04:18.77] MARC CANELLAS: So I'm a long term optimist, but I don't underestimate how hard the technical issues or the legal issues or the social issues are going to be for the next 20 years. When we were building this stuff, Keith and I 30 years ago, I said, I will retire when I can ride into retirement in a self-driving vehicle, and I'm not retired yet.

[01:04:47.99] KEITH GREMBAN: Hey, Mark, back the blue shirt. You've had your hand up a couple of times here.

[01:04:53.48] STUDENT: Bryan Broaddus up connected cars and vehicle-to-vehicle communication. I guess I'm just wondering if we're concerned about bad actors trying to game the system or abuse the system and what you can do about that.

[01:05:09.07] BRIAN ARGROW: You're talking about humans, right? I think the answer is there that if anything can be exploited or turned into a weapon, a weaponized in some sort, someone will try to do it and probably be successful until they-- that's how we have to make adjustments. It's based on-- you can't enumerate all the things that a human is going to do. So some of these things, you just have to wait.

[01:05:35.74] MARC CANELLAS: And we and this is not a new problem. A decade ago, they finally realized that having your CD player on the same bus as your ignition and your anti-lock brakes was probably a bad idea, because it's not that hard to hack a CD player.

[01:05:53.53] KEITH GREMBAN: We actually had a Silicon Flatirons conference about five years ago on that kind of issue spectrum vulnerability. And the vulnerability of any kind of wireless system to various kinds of hacks and jamming. So it's a real issue. This time, great, right up front here.

[01:06:17.04] STUDENT: In one sense, 40,000 is a miracle, a miracle of success. But I think of me speeding at you at 70 miles an hour, 3 inches

away from each other. We mostly miss most of the time. That's incredible. Are any of you well enough versed in the technology to help me understand just a little bit better about what the actual engineering problems are in making this worse-- not worse, work?

[01:06:44.85] I'm trying to channel my inner narcissistic sociopath and think of myself as a president of one of these companies and asking my engineers, how long will it be? Can we have this by third quarter? I'd like to see that car working. Now I'd like to have a demo of one that actually works. Seems like Waymo sort of does in certain very curated environments. Tesla, I don't have any idea when they'll actually work. That seems to me entirely fake.

[01:07:15.34] But where are we really? I'm not an engineer. Where are we in this?

[01:07:20.71] MARC CANELLAS: So part of it, and we like to say we're working on low-impact technology here. Hopefully, no-impact technology. Part of it is modeling the other driver's behavior, driving along and trying to guess what this person is going to do. As long as you've got a mix of automated and human driven vehicles, you have to have a model of what could this person be seeing and why are they acting the way that they are, and how do I think they're going to act in the future?

[01:07:51.15] Another problem is what else is there on the road? I used to collect examples. What's the strangest thing any of you have found on a highway? I asked the question. Somebody said a toilet. Driving down the road, a truck and a toilet fell off. A toilet is this wonderful porcelain thing which is designed to not reflect radar, will reflect the laser all over the place. It's the perfect stealth obstacle.

[01:08:19.13] A deer, you see one deer jump across. There may be another deer following. So there are always weird things in the environment that you're going to have to model. And those are going to continue to be engineering challenges, regardless of what sensors you use. The other part of it is you look at a Waymo vehicle, it's got sensors on the top, it's got sensors on the bottom, it's got sensors. It's got a perfect map of the environment.

[01:08:44.21] You probably can't afford to spend as much on the sensors as you spent on the car. And you're not going to be in Chandler, Arizona, where you have 360 days of sunshine. Brand new orthogonal roads, which are perfectly mapped. So there continue to be engineering challenges with dealing with the imperfections of the real world.

[01:09:05.72] CHUCK THORPE: And I'll add super quickly. But what Chuck's talking about, whether it's toilets or deer or the bouncing ball, is that as much as we talk about replacing humans and all that kind of stuff as humans are actually pretty incredible at what they do.

[01:09:23.48] KEITH GREMBAN: OK, so we've got the red light on, so I guess we're done here. I just want to add one more comment. So Chuck is corrected that he's not allowed to call them accidents. They're crashes. About 10 years ago when I was working for the Department of Defense, our drone crashed and I was told it was not a crash. It was an unscheduled off runway landing.

[01:09:44.63] So with that, thank you very much to the panel.

[01:09:47.00] [APPLAUSE]