Stanford eCorner

Principled Engineering: AI and Drug Development [Entire Talk]

09-03-2022

URL:
https://stvp.stanford.edu/blog/videos/principled-engineering-ai-and-drug-development-entire-talk

Russ Altman is the Kenneth Fong Professor of Bioengineering, Genetics, Medicine, Biomedical Data Science and (by courtesy) Computer Science at Stanford University. His primary research interests are in the application of computing to problems relevant to medicine. Kim Branson is a senior vice president and Global Head of Artificial Intelligence and Machine Learning at GSK, where he leads the GSK.ai team, a group of nearly 100 machine learning researchers and engineers who are pioneering the application of AI to drug discovery and development. In this conversation with Stanford adjunct lecturer and STVP Director of Principled Entrepreneurship Jack Fuchs, Altman and Branson discuss how ethics and principles can shape innovation at the intersection of AI and drug development.

Transcript

(sci-fi music) - Welcome to winter quarters final 00:00:13,430 Entrepreneurial Thought Leaders Series live stream.. We've teamed up with the Dean's office at Stanford School of Engineering to create a special ETL session focused on Stanford's thought leadership in how ethics and principles can shape innovation.. Today is the first, in what we plan to be a series of events, that will engage both School of Engineering alumni, and Stanford students in vital conversations about the role of ethics in engineering and innovation across disciplines.. I'm Jack Fuchs, adjunct lecturer at Stanford and Director of Principal Entrepreneurship at STVP.. If you're familiar with STVP, you know that we strive to equip every student with the tools to brave ethical complexity.. We believe that if people and organizations have well articulated principles, they will make better decisions.. In my own teaching, we take students on a journey where they develop their own values and principles they will bring with them in their careers.. They will then help instill those principles in their organizations, helping better navigate difficult decisions.. For today's session, we're thrilled to welcome Dr.. Russ Altman and Dr..

Kim Branson for a conversation about the ethical issues at play at the intersection of artificial intelligence and drug discovery.. Russ is the Kenneth Thong Professor of Bioengineering, Genetics, Medicine, Biomedical Data Science, and Computer Science at Stanford.. And is a past chair of Stanford's bioengineering department.. His primary research interests are in the application of computing to problems relevant to medicine.. He also holds a Sirius XM radio show and podcast entitled "The Future Of Everything." We will put a link to this podcast in the chat.. Kim is a senior vice president and global head of artificial intelligence and machine learning at Glaxo Smith Klein, GSK.. He leads the gsk.ai team, a global organization of nearly 100 machine learning researchers and engineers, who are pioneering the application of AI to drug discovery and development.. And since Russ and Kim know each other, in part through their joint development of the GSK AI/Stanford ethics fellowship, which we'll focus on later in the conversation.. But first Russ and Kim, let's begin looking at how issues of ethics and principles first emerged in your careers.. Russ, can you begin by sharing some thoughts on how you developed the values and principles that your work in bioengineering? - Yeah..

Yes.. 00:02:55,196 And thanks very much Jack, for having us.. Really looking forward to this.. So you listed all of those departmental affiliations.. And the one I wanna start out with, because it really was my first, was the Department of Medicine.. So I grew up at the School of Medicine and I did my residency there and I was hired on the faculty in the Department of Medicine.. And of course, I think no one is surprised, that bioethics or bio medical ethics has always been a thing.. Certainly, for most of the 20th century, as it became clear that it was important to put some guardrails in place because doctors previously had been doing some very questionable things and really entire societies with some of the Nazi experiments that came out and the Tuskegee incident, in which African Americans were ill treated by the system, by the US government.. This has always been present in medicine.. So it was a part and parcel of my training..

Usually however, focused on kind of two obvious settings.. The first setting would be the individual patient and the ethics of making decisions as a physician or a clinician provider for a patient who's sitting in front of you, who's depending on you for assistance and needs you to behave ethically.. And then of course the second setting would be clinical trials, which need to be always designed to be ethical.. And that, for example, means you can't have a placebo arm versus a potential treatment if there's already a treatment available.. That would be unethical to have half of your patients not even receive standard of

care.. So just that one example kind of shows you that we had to learn how to think through even the very process of gathering basic medical knowledge and certainly it's application to patients.. So if I may, I think it'll be useful for our conversation to just quickly remind people, many people might know this, about one of the most useful frameworks for ethical reasoning.. And let me acknowledge that there are many frameworks and CONT had at a framework, continuum deontological reasoning, consequentialism and utilitarianism.. But the one that has had the most purchase in medicine is based on something called the Belmont Report.. And there are four principles that we always use in biomedical ethics..

So the first principle is beneficence.. You should be doing an endeavor to do good for the patient.. And that's very important.. So you can't be doing something that's null for the patient.. It has to be in support, or I'm talking about the patient, but that's in any situation.. You should be trying to do good.. So that's beneficence.. Non maleficence means you should not be doing harm.. You should not be intentionally, now bad things can happen in the course of research and in the course of clinical care, but you are trying to not have bad things happen.. And that's non maleficence..

And so when you're evaluating a situation, you'll say, "Am I trying to do good? Am I avoiding doing bad?" Then the two other ones, and these are a little bit more continuum.. Those first two are a little bit more utilitarian for those of you who know about that, but it's not important.. Justice is number three.. Justice is basically, is it fair? Are these rules that can be applied impartially across an entire population and everyone is getting a fair shake? So you look for injustice because that's a sign, potentially, of an unethical treatment or an unethical clinical trial.. And finally autonomy.. Is whatever's happening being done in such a way that the person being affected, the clinical trial subject or the patient, has autonomy, I shouldn't use that word, has control of their own fate.. Anytime you're taking away control of someone's fate or taking control away from them entirely, that's a red flag for ethics.. So whenever we get a situation, we look at beneficence, non maleficence, justice, and autonomy.. That's what the Belmont Report taught us.. And I think it's a very good framework that, to be honest, I even use in other settings, like my everyday life and certainly in engineering, which maybe we'll talk about later..

- Thank you, Russ.. 00:07:17,760 And turning to Kim, could you share with us some of the values and principles you've brought with you to your work with AI and ML and drug discovery and perhaps how they fit into the values and principles of GSK as an organization? - Sure.. I mean like Russ, I also, 00:07:34,240 you know, wandered past of med school, but most of my training has been in science.. And interestingly, in science, you don't often have a lot of the same sort of exposure of ethics that we do in medicine and that, and maybe that's something we need to sort of revisit.. I've worked in building sort of predictive systems in medicine and healthcare for a long time.. And one of the things that, as a core principle I've always had, is that we have data.. These sorts of things gives us tremendous potential to make things better, but you have to look very carefully at the system you're putting in place and it's unintended consequences and the feedback loops you can have.. So coming to GSK and establishing the AI group.. GSK is a 300 year old company, right.. It's been around for a long time..

It's extremely trusted brand.. We know how to make medicines.. You know, when you see that little logo on the box, it's gonna be safe, efficacious.. You know, and it's a really great thing.. Now we're building this new AI group.. We're gonna start using data and building new things.. We need to make sure that we build medicines for all people.. We wanna make sure algorithms, the algorithm products that we're building, also affect all people in the same way.. So it's very important for us to have those sort of same types of principles and think about them.. And so we are now in a world where a lot of the way we discover and develop medicines is now data driven..

And it's all machine learning, it's all feedback loops.. So we realize that anything we build, anything we put in place gets used, people start to base decisions upon it and use that again can lead to a feedback loop.. So for us, we wanna be very careful to understand what that is and to think through the consequence of that, and really for us, it's a choice of what problems we do work on.. And what we don't work on right, is an ethical decision.. But a key thing else of this is that we can't decide not to do something because we can't make it like, you know, it's also unethical not to solve a problem for someone, if you can't solve it for everybody, right.. So how do you wrestle with these types of things? And I think, and that's something we know there are biases in literature, you know, who's represented version databases and things like that.. You know, and so all those types of questions, they actually come into play in a very realistic fashion, as we start to think through what we do.. So we have an idea of building software for each one of our assets, right, how to use it in a setting.. Now, the best way we generate data for that is in clinical trials.. However, the people that come into clinical trials, there are many biases to who can participate in a clinical trial or not..

So how do you actually, what efforts do you go to expand that access? And there are many questions along that.. So all those types of things, when you start to think about this, you suddenly realize that you need medical framework for thinking through these things.. Now, GSK as a company has amazing ethical frameworks for lots of other types of areas, but this is a new thing, right.. And so this is where, this is really led to actually the idea of, chatting with Russ about like, we need a sense of practical engineering ethics for these types of things.. These aren't abstract things that we're gonna think about gray goose scenarios and what ifs.. These are real things that are happening now that we actually have to make decisions about.. - That's delightful to hear Kim 00:10:33,600 'cause we teach students that companies should develop a broad set of principles that need to be communicated, that need to be instilled in an organization.. And that will come into conflict as you just described.. And it's the wrestling with those principles within the context of a decision that provides the power in decision making.. You just described one in particular..

And I'm gonna try to use that as an example.. Kim or Russ, actually, then you can also respond.. Do you have any recommendations from your experience at GSK of how best to incorporate those principles into the decision making at the organization? Or examples where GSK has wrestled with, you know, maybe even specifically that those principles of, hey, if we can't solve it for everyone, we shouldn't solve it for anyone.. Well, that's clearly not right, right.. Right, it's unethical as you say to not solve something for some, if you can, but like where do you draw the line? And how, you know, where is the bias and how is the bias and where do you go from there? So I guess, you know, that those kinds of questions to you about whether there's tips for people about how to instill this throughout the organization or examples in situations of how you've had to wrestle with it.. And then Russ, you can as well comment on that.. - Sure.. I mean, 00:12:01,230 I think one of the things we have is sort of a checklist.. When we're starting to consider to build a product.. So us as an organization, we only try and do large impactful things, right? We're not solving someone, I wanna transfer this word document to an Excel spreadsheet and build a thing for that..

We're building big things to impact patients or help discover targets and things like that.. And they have large term, long term commitments in the company.. We discover a target, we do something, we're committing money to it.. There's an opportunity cost.. There's all these things that happen to that.. So these are big problems we're working on.. So the first we look at is, well, first of all, let's look at the data you are using, right.. And you can't it there and say, "Well, this is the data I've got.. This is data I could find." The first question is, is your data adequately representative, right.. Is it got biases, if not..

And you know, are there other sources where you could find more data, right? And that's a very interesting question.. And then you ask like, "Well, it's sometimes it's not okay to say I just couldn't find it.. It's not out there." You say, "Well, should we gather it," right? Should we just go and actively gather the data? And that's actually pretty new to most people that come into machine link, particularly at school.. What do you mean like? Like, well, we'll just buy them all.. We'll generate more.. We can maybe run a study.. How do we do this? And we can look at the cost of doing that, right.. Now, if it's obviously, there are lot, you know, there are time bounds and large amounts of money to this sort of stuff.. And it maybe infeasible to things depending on what you want to do.. But you have to make an attempt and we have to have understood that..

And we pre specify, what biases we know about that.. And we made reasonable attempts to address all of them.. And I think that's before you even write a single line of code, right.. Just to think through what that is.. And then you think the other thing you think about is the what if scenario.. I have the model, it's in production.. So what do people do with it? What's my intended use, right? And you can write that.. Well, that's easy.. That's what I wanna build a model for.. And then what are my unintended uses? If someone else had access to this or what would users, how else could they use this in different settings? And that's where, you know, I worked with systems at the past that could predict people's probability of getting disease in a time, given their past medical history..

Really good for planning care, really good for doing pre implementations, things like that, all great stuff.. Also really good factorial methods for insurance companies and pricing healthcare and things like that.. Possibly good stuff, maybe less good stuff, depending on the regulatory environment.. So you have to also think of the intended and unintended uses that you can, when you put something into production and how it changed the world around them.. Those are two anchoring principles.. What goes into the machine, willing thing, and what comes out of it from the other side.. And we have checklists and things that help people think through that, right.. And a lot it's about thinking through that and to make sure, okay, tell us what your original intent with your model is, how you wanna make the world a better place.. And then let's work out the best way that we can practically do to attack the aim.. And also, how do we monitor, that's the third thing, how do we monitor this, doing that in practice, right? So you can't just build these things and like off it goes, now go work another problem..

Someone has to watch it, and log its results, and review it.. And that's something that also is kind of new, that you don't get during grad school and other types of industries as well.. I'll leave Russ chime in there.. - Yeah.. So, I think of, 00:15:02,420 you know, I'm an academic, I'm a professor, and I think of our role as upstream of what, you know, we're sending our folks into Kim's environment and to other environments.. So when you think about that, there's a couple of things that we have to do.. First, we have to have some formal training in ethics for engineers.. And so, I'll just briefly tell you the story of the bioengineering undergraduate major is pretty new and it was actually kicked off when I was chair.. And there's a process.. I'm happy to report that Stanford has a pretty serious process for vetting new majors, where you get comments from your colleagues who are outside of the department..

And by far, and this was in the late to early, you know, 2000 to 2010 timeframe.. So more than 10 years ago.. By far the most common and strongest piece of advice we got from every committee that looked at this major was you must train these students in ethics.. And their logic was pretty straight line and short.. It was, you are giving them unbelievable power tools.. And this was, even before CRISPR, kind of came out where now we can literally, but they knew where bioengineer was going and they said, "You're giving amazing power tools to these folks.. They have to have a compass for navigating this." And so as a side story, I drew the straw.. It wasn't short, I'm happy to have drawn the straw for teaching.. And so, for the last 10 plus years with David Magnus, a colleague in the Stanford Center for Biomedical Ethics, we've been teaching a class on bioengineering ethics.. It's different from medical ethics because it's more upstream in terms of technology, development, and whatnot..

So and as many people know, our colleagues in computer science have kicked off a class in ethics.. The Stanford's current

strategic plan has ethics all over it.. And I think appropriately.. And so we, knew this was coming.. And so all of our undergraduates at least, get a pretty good exposure for a whole quarter on these issues.. And they write about it.. I should say that we have 30 or 40 undergraduates each year, and we're getting 200 people in that class.. So there's a lot of people who are not bio engineers, who are taking that class.. And I'd be happy to talk about it more, but that was the first thing that we need to do for Kim, basically.. We need to send Kim employees who have this vocabulary..

On that same theme though, we have to practice it in our labs.. I think PIs, who host graduate students, and undergraduates and postdocs, we have to model this ethical behavior.. And Kim said it earlier, why are we working on problem X and not problem Y? I mean, a bad answer would be there's more money in X than there is in Y.. A good answer would be X is a more pressing societal problem, or is part of a very complex, and a little bit of progress in X would make the world a better place in very real ways.. So, and this, let me be very honest.. This has not been a routine way that me and my colleagues choose research projects.. I'm not saying that we're mercenaries, but I don't think people have always been intentional about the ethical framework about, and their choice of research projects, but I'm beginning to see it.. And I'm excited about it because I think that will also model for the students and the trainees before they go out into their, you know, various exciting careers, that this stuff needs to be embedded in your everyday decision making.. It's not just something, this is a famous saying, "You don't just sprinkle ethics on top of a project." Just like a plane starts with the wing, and the fuselage, a project has to start with the scientific question and the ethical framework of that question.. It doesn't have to be, you know, a tome..

But it has to be one of the dimensions.. And so those are the two ways that we're trying to set people up so that when GSK hires them, they're pretty clueful about all this.. Work in progress, not claiming that we're done.. - Well, and also leads to a question Kim, 00:19:11,830 kind of a practical question.. You know, Russ described, well, you know, a bad answer would be, we work on this because there's a lot of money to be made.. You know, a good answer is we do it because it's solving an important societal problem, I'm paraphrasing.. You know, Kim, I mean, is there a penalty for having good ethics? If you, as a company, if GSK follows that advice, do they then wind up underperforming relative to competition? And you know, how does GSK think about that? And how do you think about that? - I mean, for us, 00:19:47,553 it comes down to really longevity and trust, right.. Like, you know, you need to think we, this as I said before, 300 years of history.. You wanna continue on, right.. Do you wanna be the guy who breaks the chain and ruins that 300 year brand? I don't think you do..

And there's a reason for that.. You have to think very carefully about it.. And I think there's some things you can, you don't have to pick up every dollar that's out there.. And there are some things that are the correct things to work on.. And we kind of fundamentally believe that ethical science, right, is better science.. And the reason why it it's better science is those types of questions force you to engage more.. Rather, this is the data I've got.. Forces to engage more, forces you to understand how to collect the data, why they've structures to get the data.. It helps you involved in the patient's lives.. It helps you understand the systems that have that data in a real practical sense, which means you probably build a better solution, right..

A better solution, right, we would fundamentally will win out in the market.. It's probably more robust as well, from a machine learning perspective that so, and it can be used more.. So it is, fundamentally, comes down to trust, right? Making the attempt to build the best thing you can for everybody.. And in the process of doing that, you engage more people.. You also make it a flywheel effect, right? You're solving problems, maybe understanding why people don't come to trial, maybe you can fix something to get the data.. Now you have more people in trials.. And we actually can come up with better medicines, right, more representation.. So there is a flywheel effect of doing that properly.. For us it's not, as Russ talked about sprinkling ethics about it, it's not about having the ethics police, right.. So too often, you see people trying to bring ethicists on top of the team as a regulatory thing or something else coming after the fact or things like that..

You can't deal with like, it has to be done at the same time, right.. So you have to think very carefully about the ethical culture you create with your machine learning people.. And that's something we're very keen to have, right.. Is the right kind of people this.. So this is why we have responsible AI people and a VP of Ethics and Policy around these things.. 'Cause policy and ethics are actually kind of intertwined, right? That's why we have regulators, right.. What are the regulators doing, right? They're actually enforcing a lot of the principles that Russ actually stated earlier, right.. That's why we have regulators in society and things like that.. So there is a constant dialogue between that.. And I think, so this is something we've thought very carefully about..

And so we have, there are fully fledged part of the team, but they're also technical people as well.. And that's a very key thing.. I think that's part of that two cultures thing is something we're trying to sort of change with this new fellowship.. - And let's ask this question to Russ first 00:22:21,413 and then to Kim.. But Kim just mentioned something that one of the attendees has a question in the chat.. For prospective founders, looking to build in the intersection of ML/AI and health, do you think governmental regulators and interventions are a reason to worry about progress happening too slowly? And is that a big enough deterrent not to get into the area? And that's the specific question, but let's generalize a bit to the role and the interaction between and among regulators in industry and how that plays out.. How you think that plays out? And we'll go to Russ first, 'cause Kim you just answered, but I definitely want to go to you afterwards 'cause it touches on what you just described.. - Yeah.. Yeah.. 00:23:12,150 That is a great question because it's so, it seems so obvious that regulation could only be a thing that slows things down and kind of torpedoes some of your best efforts..

And let me say that I've had conversations with CEOs of extremely famously, extremely large tech companies in the Silicon valley, who will say, we will not engage.. Now, this is old information.. This was 10 years ago.. But 10 years ago they would uniformly say to me, "We will not engage in a project within our company if we see the FDA anywhere near it." And I just wanna say that the world has changed and the FDA has changed.. So first two sentences on my credentials on this.. I am the co-PI of an FDA funded Center of Excellence on regulatory science, where we have collaborations between FDA scientists and UCSF and Stanford scientists in areas that are critical for the FDA to understand.. And you will not be surprised to learn that digital health and AI is one of the areas that they come to us a lot.. There are several of these centers around the company, but around the country, but you not be surprised to learn that we're the one who gets a lot of the interest for AI and ML in health.. And we've had several projects with them.. The FDA scientists are extremely interested and caring about bringing these technologies to patients..

That's their entire professional is to protect the health of the American public while allowing drugs, diagnostics, and that includes therapeutics, including AI therapeutics.. And many of you will know there's some have been approved.. And so it seems scary, but I have seen over and over again, that even for a start-up there is a very well understood way to have a pre-submission meeting with the FDA, where you describe the technology that you're developing.. You describe everything you can about it.. It's kind of a confidential meeting.. So this is not on the public record.. And then they give you an assessment of what their questions would likely be.. They do that at the meeting, but they also go back think, and talk, and then send you a letter with these concerns so that you don't have too much of a moving target.. Now they can't guarantee, because things happen.. But they give a best effort of what you would need to do to demonstrate the safety or ethicacy of this, whatever this tool is..

And I've talked to startup people and it starts out scary.. They usually wait too long for that meeting, because it kinda reminds me of a PhD student who doesn't want to do their defense because everything is not finished or even their quals.. And I tell them, just get in front of your committee and tell them where you are.. It will be useful.. And it's the same thing for founders.. They're so worried about the FDA interaction that they actually delay it too long.. You're allowed to have more than one of these meetings.. You don't have to have all the answers at the first meeting.. And it's extremely unlikely that you say something that torpedoes your entire effort because the FDA is not out to torpedo your effort.. They're out to set expectations..

So I hope that wasn't too long, but I'm actually bullish that the FDA is learning how to spell AI, understands there's a tsunami of things coming to them, and it is staffing up or else getting collaborative help like from our center to make sure that they can, you know, obey all the laws about those response times for various submissions.. So I wouldn't be afraid of the FDA.. And in fact, I would engage with them ASAP so that you can see that they're just real scientists who are just trying to protect the public and help you get your product out the door.. And I'd be interested to hear what Kim's experience, that's usually for startups.. I wonder if GSK has the same experience as a 300 year old big drug company.. - Oh, and Kim also broaden to how 00:27:24,080 GSK thinks about that regulators in GSK each involved in this emerging set of principles.. - Yeah.. I mean, 00:27:31,800 I've had experience from the startup side and also from the company side.. And I think the first thing I learned from startup stuff is, yeah, we wait too long, but turns out the regulators are people too.. And all of them are really passionate scientists..

And I think it's really important to understand like, okay, we actually have regulation laws in lots of industries.. There's regulation running a food truck.. You know, there's regulation if you wanna sell your product, this is the people you have to talk to.. But actually engage them early on and having a dialogue.. And they're quite open to understanding how to do things.. A lot of it is, they're seeking understand.. And they'll tell you why this particular regulation, why they think like this, and you've gonna have a dialogue, whether that's a case or not.. And it's actually one of those things you're sort of not taught.. You kind of have this thing, the FDA is like the IRS.. It's only a bad thing, right..

Very different department, right.. They're there for a very different reason and you can engage them.. And I think the other thing is that I think what people fail to understand is that regulatory environment is a constant dialogue between industry, right, and the regulator.. So, you know, from the GSK perspective, we could have conversations with them.. I'm like, we're helping develop.. Like what is good machine learning practice, right.. For building medicines, right.. Should the same thing of how we do update softwares for a pacemaker, be the same thing of our updating software that I'm trying to do a diagnostic, you know, for a competition pathology algorithm, right.. And actually how even should the regulator validate work, right? So previously in drug discovery, run a clinical trial, I get all my primary data.. We do some statistical analysis..

We give the FDA, they check our homework, say we get the same conclusion.. There's a debate.. I'm glossing over this in horrible terms.. And Russ is smiling, but that's kind of the process, right.. And off they go.. For machine learning algorithm, why doesn't the regulator have their own independent set of data that I don't have, right.. It's an API.. Why don't they decide they're gonna call at any time? They can continuously monitor something in production if they wanted to, right.. We do have monitoring in production right now in the pharmaceutical industries, adverse drug reports, right.. There is a parallel for these types of things, right..

So what you have to do is have a conversation with the regulator.. You need to talk to 'em about this sort of stuff saying, "Hey, we think this is how you could do it," right.. And also explain to them the types of people they need to be hiring or

training, right.. You know, they need software engineers, right.. They've got really great statisticians at the FDA and people like that because they've needed to have them to understand new trial design.. So it's an evolving concept.. And I would say the more you engage with them and tell them what you want to do and how to do something, right.. You know, you get to know people there, right.. And you can actually, as a small 20 person company, or a giant pharmaceutical company, have these conversations with them as well, right.. And I think that's a really important thing is to engage in that dialogue and the regulatory environments are there for a reason..

Now it may not evolve in law as fast as you want, right.. But this is where I think that typically there is always a path to work with them, to do something, right.. Because as Russ said, they are highly motivated, right.. To try and change healthcare, right.. Because the other thing is if they just say no to everything, right.. And I know Russ probably (unintelligible) there was a famous regulator who said no to everything, that never allowed a single drug in, in their entire career.. But that also, that doesn't work, right.. And that leads to a backlash against that sort of thing, right.. And then once that happens, we lose.. We end up having building things aren't as safe, right..

So there is that interplay.. So I think it's very important to engage, involve, and realize it's an evolving thing and you can push on it, the laws aren't static, right.. You can campaign for change and give really strong reasoning for the doing that.. And it's an educational component as well.. And you'll learn from them., right.. - Thank you.. Both of you.. 00:31:10,630 And Kim, another question from the audience, Kim said, we only do big things.. Address the ROI of not looking at rare diseases when a corporation doesn't see enough market share, or market size is probably what they mean, for a big return or even a return.. And let me make it a little bit harder..

You know, this is one of those wrestling with the financial return versus solving an important societal problem.. Sometimes the answer to that is to charge an insane amount for the drug in order to make the market size big enough.. So how do you wrestle with that at GSK? - Yeah.. When I was saying that we only do big things, 00:31:54,732 I was talking about like the objectives, the AI department.. We'd only work on a few things that we think are very impactful, right.. Have thousands of impact within GSK.. But if you come down to the landscape, so in many ways the cost to run a trial and things like that can only get so cheap, right.. No matter what you do, whether you're doing it for one patient or a thousand.. You know, things like that.. There just some sitting, right? So those are some just the economic realities of running trials at multiple centers and all the sort of overhead that goes with things..

So in order to offset that there are various types of incentives, right, and tax structures, Orphan Drug Acts, things like that, Rare Disease Acts.. And what you actually have is a set of companies that have built therapeutics for those rare diseases as single Mendelian genetic disorders and things like that.. And they are typically smaller companies, right.. So they have lower overhead.. Maybe they've done a longer research phase.. And then they sort of maybe shut that down and go to production to conserve their capital and do that sort of thing.. But that's the way we've typically seen those sorts of things being addressed.. Now there's a lot of those companies and in many ways, a lot of those easy to build a medicine, rare diseases are getting done.. And you're seeing some interesting things, right.. There are, you know, and it's not that they're always done by small companies..

Sometimes they're done by very, very big companies as well, right.. But the way they get around that is through the various sort of incentive structures of patent life cycle and pricing and things like that, right.. And expedited review.. The standards are sometimes that we apply to that for a rare disease, let's say you're giving a medicine to a small population for which there is nothing else, right.. And maybe it's 50 or 100 patients, different ethical risk reward, risk harm ratios are used compared to, I'm gonna give this statin to most people aged over 50.. And they're gonna take it for 20 years, right.. Very, very different risk pool.. Let's be very careful about that.. That all comes to regulatory sort of stuff.. So it's not something that like, you know, oh, we only do things if we can make a billion dollars out of it, right..

And I think the other thing to remember is, to be honest, the age of the billion dollar blockbuster drugs is going away, right.. There are very few of those left defined anymore where it's the one medicine that works amazingly well in all compass, right.. There's always competition in these different types of things as well.. So what it actually is now is the way of finding cheaply and effectively the patients for which your drug is 15 times X more effective than in another than another population.. So differentiation now is the name of the game, right.. And that's the trend, right.. Rare diseases are really nice cause they're a very well differentiated population cause they are rare and you know who has them perhaps, right.. It's the discovery thing.. So there's a lot of things that go into play for that decision.. - Great..

And shifting gears a little bit to Russ, 00:34:35,600 'cause I wanna make sure and we'll also ask him this question too.. I'm thrilled with this, with the partnership that you two have been working on and Stanford faculties on the forefront of ethical thought leadership applied to technology and often in collaboration with industry.. And this gsk.ai Stanford ethics fellowship is a case study.. And I mentioned it briefly at the outset.. It's a new post-doctoral fellowship that allows researchers to study ethical considerations at the intersection of AI and machine learning and drug discovery.. So first Russ, could you briefly touch on what led to this kind of post doctoral fellowship and how it was designed? - Yes.. So I don't know if this came up in the introductions, 00:35:23,090 but Kim as a youth spent time at Stanford.. And so I know Kim and we meet every now and then, and we talk about stuff.. And so I'm gonna have to give a lot of credit to Kim because I had my jaw drop one day when a representative of a major pharmaceutical company, who I also knew when he was a postdoc, came up to me and really said, I

don't wanna say drop it in my lap.. But he said, "We have a pretty clear idea, Russ." He made the arguments that you heard earlier about why this was absolutely critical to GSK..

He knew that we had some presence in bioethics and biomedical ethics at the medical school and at the engineering school, which is very important because this is a, you know, this is a bridging kind of issue because of the AI, as well as the medicine being involved.. And so really Kim's challenge to those Stanford faculty was how could this work and how fast can we get it up running? Because you know, I work for a big company Russ and we don't have the same time scales that you sometimes have.. I need it yesterday.. And so we right away started talking about how it could run.. Fortunately, the Stanford Center of Biomedical Ethics has a long history of training fellows, especially fellows with technical backgrounds.. So that fell into place very nicely.. We kind of wrote down some of the scenarios that somebody might research.. We worked out all the intellectual property to make sure that, you know, very importantly, this has to not look like we're shills for JSK.. And there has to be kind of academic independence.. And yet GSK is paying for a lot of the training..

And so we have to make sure that the outputs are of at least of relevance and of interest.. So working through those conversations was fun and challenging, but not killer.. And then we announced this program and we're actually recruiting for it actively.. They will have full academic freedom to focus in the area.. I think that, let me just say that the big new idea here is ethics in clinical trials for drug development is well established.. But what Kim said to us that really got us excited is we wanna focus on the very early stage discovery when you're doing genome-wide association with people's genomes, when you're look at cellular data about how diseases.. So this is very early, when it's not about patients, it's more about those issues I was discussing before, the choice of problem, the choice of whose cell line are we gonna discover, make discoveries on, which population genome sequences.. So I'll let Kim chime in on this.. But that was the genesis of the project.. And now we're gonna have people come usually post PhD, spend a couple of years with faculty mentorship and kind of a collaboration with relevant GSK scientists, and then try to make traction on some of these tough problems..

- Great.. And Kim, how did you navigate the waters at GSK 00:38:24,513 to make this happen? - I think Russ's description 00:38:31,000 is fairly accurate.. I think one of the things we realized and coming back to the way we're discovering drugs now is it's a lot more data driven, right.. So it is genome (indistinct) association studies.. It is, in that case, gives you a hint of which genes you do.. You might do the functional genomics now, right.. So this is CRSPR for other technologies.. We have induced pluripotent stem cell lines, right.. Which donors are we taking those from, right? You know, is it just one line or a panel of donors, right? And all those things when you put together, right.. Think of, you know, I've done the GWAS in this population, if a minority group isn't represented there, right..

And maybe their gene, there's a different (indistinct) for that.. And we don't discover it.. And you realize something when you chain all these things together, they have a reinforcing effect.. And those are things that come out the other ends as target.. We're prosecuting, making medicine, do it running to a trial.. And you realize all these things individually like, oh, well I'm not using stuff for clinical patients so it doesn't apply.. But actually when you step up a layer, look at this as a systems level type approach, you realize like, oh yes, it really matters what goes into the machine where and how we use the data and how we make the decisions at each gate as things flow through.. And it was also driven by, you know, when you have a conversation to like, and we have a code of ethics.. Everyone's like, oh yeah, sure.. I read it..

Good tick, Kim.. I'm like, no, no.. It means something.. Like it's not just, you know, you have to tick a thing compliance, right.. Cause we're used to doing with regulated patient data and things like that.. The problem that we observed is that a lot of the people in ethics literature are not technical and not computational.. So typically, the way they've addressed problems or talked about things are either so remote, they were like, I don't know what to deal with that.. Like an AGI type problem.. Or just like, you know, just not relevant, right.. And not practical..

They're like, "Okay, that I see that.. What should I do? Help me?" So the idea was like, well, actually what we need to do is we really need to do research in ethics, right.. So we have things right now.. And like, there are whole things like, for instance, there's a concept called a data cold chain.. So a cold chain is something I have so I can move, which we all know about, like, you know, vaccines across the country.. And I can do things, keep refrigerated and safe.. So I built a really great algorithm, does computation pathology, and it can predict various outcomes and things like that.. If you are in a country that doesn't have the data infrastructure to do that, you can't benefit from it, right.. So what does that mean? Should we be investing in the data cold chain for those companies? Should we be advising sort of stuff? Or should we just say, well, you don't have it, but you know, it works great if you're living in North America, right.. Or the UK or something like that..

And so you realize that there are a lot of issues around this sort of thing.. But there's also research about how these things are used.. And like, because we don't know, these are all new, we need to actually start doing the research now.. And rather than then like just doing it and then saying, "Oh, that was bad.. Now we write some policies about it." Like, actually let's get ahead of the game and try to do that.. So let's sort of learn from the previous approach that happened in medicine, right.. As Russ said, all those sort of things that happened previously.. We can start doing research about that.. What are the implications? So that's idea of like, let's create a fellowship that also has people that are technically trained, right.. That will learn these types of things..

So that really, and to be honest, I would say a fraction of the AI community does simply ignore a lot of the ethic type stuff,

right.. Like it's not as bad as it used to be.. It's definitely changing.. People are realizing that.. But for a long time, they're like, they just kind of ignored it, right.. It was like, no, I've got stay the out.. I've done this sort of right.. I'm focusing on my facial recognition task.. It's really awesome.. My model's 10% better than this other guys, look at my new architecture..

Like, yeah, it's a cool problem you're focusing on, but it lives in a wider context.. Have you thought about that? And now people are thinking like, "Oh yeah, let's probably a thing, " right.. So we want to do the same thing now, right.. In our particular niche, right.. And we think it's fundamentally gonna be about research driven to lead practical solutions, right.. - Very good.. 00:42:23,970 And I'm trying to do as many of the audience questions as possible.. I still have a few that will probably get to, but on the other side, but from the audience, there's always a lot of hype in the media about the use of AI and drug discovery and development.. What's the true state of the situation? What are the areas where AI and ML does and does not work is the way they or let's say, where is it more promising? Where is it potentially over-hyped? - I'll start out.. 00:42:53,810 This is what Kim does for a living so that I definitely don't wanna follow him on this question..

But I think about this and I do some kind of consulting and I've seen some, so here's what I see.. First of all, for the very, very large genomic data sets, AI is pretty much mandatory to pull out the signals of which genetic variants, especially in combination, are correlating with the phenotypes, that is to say the diseases of interest so some of the big data is almost big enough to impress the folks who really deal with really big data.. You know, the Google Facebook, Twitter people are in a stratosphere.. We are below that, but it's pretty respectable big data at the genome.. So that's the first thing.. Second of all, everybody knows that AI is good at detecting patterns.. And this can be very useful in looking at electronic medical records for finding a bunch of patients whose disease looks similar.. A lot of diseases are really waste bins of like all different kinds of people with slightly different diseases.. And when you do a clinical trial, of course your drug is only gonna work on 20% of the people because 20% of them actually have a version of the disease where that drug is relevant and 80% might not even have the disease, or they just have a totally different form.. The AI systems are very good at finding patients who are kind of looking very similar, along all available dimensions..

And that's also very valuable.. Now jumping very molecular we've all heard about AlphaFold 2.. AlphaFold 2 is the program out of Deep Mind that was able to predict a three-dimensional structure of proteins.. I'll just remind people who haven't taken biology since high school, that those three dimensional proteins are typically what we call the targets of a drug.. A drug is often a small molecule that binds one of these proteins and modulates its function to help the patient have a, you know, have their disease go away or get better.. So going from just some protein, three dimensional structures to all of the structures at least almost all at a reasonable level of accuracy opens up our ability to think about in a very rational way, new drugs to interact with proteins that we weren't able to think about previously, because we didn't have the structure.. And so that whole three-dimensional structure and molecular understanding of drug action is about to be revolutionized.. I mean, it's happening right now.. I can't tell you how quickly engineering students read those papers, walked into my office and said, "I wanna work on the spin out effects of that discovery." It took about a month for them to figure out.. And so, I'm sure Kim has other examples, but right now those are the areas where I'm seeing a lot of excitement and a lot of positive results even impacting successful drug launches..

- Yeah.. I think drug discovery development 00:46:03,960 is a big, big set of fields, right.. They intersect there.. So there are, as Russ said, there's a lot of things that are happening in, you know, the early discovery phase.. So certainly, you know, yes, GWAS.. Now GWAS itself is a pretty old technique, right.. Multiple sequential, independent hypothesis.. - You better define GWAS.. 00:46:19,570 - Genetic wide association studies, right.. 00:46:20,510 So we think like what, that's still how we do things..

Except we're now building models that look on raw DNA sequence and predict open and close chromatin.. And these are stacked in Coda models, right? For doing these sorts of things that explain how things can change across different cell types.. We frequently do these very large function genome screens where I can do every single gene up and down, by the pair wise type stuff, or these very poor CRISPR things.. So these methods like perturb seat that generate very large data points.. Like, I think last year we probably generated, I don't know 25 million data points as a feedback loop for a single ML model we're building.. So there's something, we actually generate data just to feed into our algorithm.. So the algorithm becomes a discovery tool.. So that probably gives you an indication of how important it is.. There's been amazing advantage, obviously, in computer vision.. Well, that now applies to what we're doing with cells..

So the phenotypes, I can track things over time.. Adding time to everything is key for biology, right.. Cause these are dynamical systems that change.. I can now do single cell RNA 6.. I can take single cell and I can do genome sequencing and RNA sequencing of that.. Right, and I can do that over time.. I can look at cell morphologies.. I can look at protein expression as well, functional characterization.. So now we have all this multimodal data, right.. It's really difficult to integrate..

So one of the great methods and particularly around neuro networks, the sort of stuff is transfer learning and multi modal integration.. So all these things come into about not so much, like, I'm talking mostly about what is the title? What is the thing you should make the medicine about? Right.. So even in doing that, we have an AI system that we use for even here's where I wanna drive my cellular model to, my clinical translation model too, right.. What is the best target for doing that? Now I can just do one gene at tie in 20,000 things.. And I'm using CRISPR to modulate it rather than making a small molecule tool.. That's something super new.. Or I can actually be a bit more smart and I can turn everything sequential learning problem..

Take some data, generate some data, have a multimodal feedback where I wanna try and make the thing look like this thing, express the same genes.. And it maybe make this protein at a higher level, right.. That's gonna be a really good drug target..

That helps me discover something.. And then from that, it's the design, right.. So it could be a small marker, like Russ is talking, but it equally well could be an antibody, right.. And now we use, you know, Gaussian processes and all sorts of beta optimization to optimize the antibody sequence, right.. For various properties.. Not only just binding, but aggregation, (indistinct), stability, 'cause it has to last for a long time, right.. So it starts to feed into the manufacture aspects.. It goes into the recruitment aspects, as well.. And so we see, you know, you have this lab in the loop of all your models, but we also have the clinic in the loop, right.. So the Russ's earlier point, a lot of diseases are sort of, you know, made up by Victorian men in top hats, right? They're all symptomology, right..

But now we can measure things on such a fine scale.. So we can get a population of people who have disease and observe them and start to characterize things what's going on.. Like let's be Parkinson's, right.. Some people progress really slowly.. Some people progress really fast, right.. Are they all just the same Parkinson's or is there a different path of physiological process under there? We wanna sort of dissect that out because that tells us something about things.. So we see a lot of impact across all the sorts of domains, even in manufacturing and scale, in selecting patients, operating trials, right.. We have a lot more sensors and variables now, that we can use in trials as well.. So even our scheduled clinical assessments is changing.. So it is one of those things that it's not just one revolutionary thing that you put in the sequence of the person and here's the molecule you make, right..

But it's across all these little changes, across the entire big pipeline will speed up the whole thing, right.. And I think for me, the biggest impact to something we discovered was if you have targets that have genetic validation and functional validation, they about twice as likely to become really successful medicines.. Even if you left the rest of the whole machine unchanged, right.. So it shows impact of working on the right things in that pipeline.. So there'll be lots of change happening you know, across industry, right.. - Thank you.. 00:50:21,140 The are the criteria that have emerged in the broader AI ML world of trust, security, fairness as being principle, you know, principles that people evaluate technology against.. Are there adaptations of that framework or other vectors that are particularly suited to AI and ML as applied to drug discovery? - Those are good ones.. 00:50:51,730 And you know, in general, you have to be thinking about those.. There are a lot of detailed issues that, I'll keep this short 'cause I can think we have a lot of questions..

There are a lot of detailed issues that are a little bit different in biology, maybe.. There is responsibility because there's liability in medical devices and drugs there.. The chain of responsibility for decisions has to be somewhat transparent or a bunch of lawyers will make it transparent.. So that's on the mind for the AI system.. Explainability, which has come up in all areas of AI.. It's controversial because there's a difference between the explainability you need when you're convincing people that it works, and that might be different from the explainability you need once it becomes a routine part of practice.. One thing that I wanna stress, and this is really relevant for all AI systems is we've talked already a lot in the last 53 minutes about the importance of data to inform our AI systems.. Once we start deploying AI systems, then the AI systems will be affecting the data that we collect.. And it's gonna be a very complicated situation to unravel why we're seeing a new set of biases.. So we have all the old biases, and then we're gonna have a new set of biases, which are the biases that came from using an AI system in the first place, that didn't used to be there..

And I'm very A concerned and B interested in research that starts to untangle, how do you up data an AI system, when part of the data that you're training it with, is the previous version of that same system? Very complicated.. - Yeah.. There are some lessons from that 00:52:41,930 in the sort of the people who build financial trading algorithms.. Cause once they put it into it, it changes the market, right.. And so there are some things to learn from that, but I think to the point about, sort of interoperability, a lot of that is sort of robustness and reliability concerns that people are trying to address, right.. Because if you ask people, you know, we all interface with algorithms in our everyday lives and things like that, or technologies, right.. How does an LED work, right? How does this model in front of me work? Can I just explain the physics of that? Not all, but it works, I'm sure.. And it has a minimal impact to me so I don't really need to know.. So a lot of the time, where we want that is we want to make sure it's robust and reliable.. Like why does it make it make a decision with inputs and lots of time people get care about interpretability a lot more when there's a potential for harm, right? And it all comes back to ensuring that it is a robust system..

And we know some of AI systems, right? Particularly some of the visual ones, I flick a few pixels and things like that.. And I turn a banana into an owl or something like that, right.. Well, what happens if I've got my computation pathology algorithm, which is classifying someone's tumor stoma boundaries and the types of properties, it's slightly out of focus, right.. Or I've got folder tissue.. How do I show that's robust, reliable, and safe? Right.. And you know, certainly the pathologist look at things.. He may have off days.. Maybe he doesn't look at all the slide and things like that.. There's noise with that, but we can give it to two people.. Right, and we know there's noise in the system..

So there's many different ways to think about that.. And I think these are some of the areas that need a lot of research, right.. And are going to be critical to get this.. Because we are sort of taking something into it and we're gonna have to run things in parallel with the old system, right.. It's not gonna be something we're just gonna cut out over overnight, right.. Because we'll need the observation.. But then we also know that like sepsis prediction algorithms that like the ones that I

guess are probably made by various EHR manufacturers, right.. If you look at what's happened now, and you look at their performance dips, their performance decays over time.. Cause medicine is ECS static.. We get better at understanding, recognizing sepsis and things like that..

So that, so the signal strains, and maybe that's only the harder patients, right, to detect, right.. So there's a lot of things that go into this thing that, as I said, once you put them in into practice.. And I think that's something that we'll have to consider as well, particularly as we gather other data about people, right.. - All right.. Let's try to get one 00:54:58,940 or two more questions in the last three minutes here.. Someone asks about penalties for working unethical problems, about whether there's a biology tax in that if you work on AI and ML for biotech, you're gonna get paid less.. Does that disparity bother you? And how do you think about addressing that? - So lemme address that because 00:55:20,360 I have some specific experience in this.. It can definitely be the case.. There is a class of machine learning employees who I would call quite mercenary.. And, you know, I've had a startup and we had AI people..

And I talked to the chief, the CEO, and he was very clear.. He said, "Russ, this is a tax.. We can't pay as much as Facebook and, you know, Google, but it's not that hard to address because our mission is very compelling." So if you're in front of one of these mercenary people, it is not hard to figure out that they are just going to the highest bidder and you do not hire them.. - Yep.. 00:56:00,520 - You find somebody who understands 00:56:02,700 that the mission of curing cancer or making new drugs is a worthwhile mission.. And they can, and also of course, they'll still be able to live and take care of their family, but some of that big bonus that they might get, they're getting it because they're working on the help system for the help system.. And okay, God bless you.. But maybe you wanna work on drugs that are gonna cure Alzheimer's.. So you have to do that at recruitment time is the advice that I've seen work because you can compete simply on monetary reimbursement.. But we do have a couple of knobs we can turn because it's a pretty uplifting mission..

- Okay, and just in the interest of time, 00:56:43,710 30 seconds each just on, where do you see the principles that we've all been talking about today evolving, you know, we talk about the importance of principles evolving here.. Just a quick, quick couple sentences each.. - Go ahead, Kim.. 00:56:59,323 - I think we're gonna look at the time where 00:57:01,690 we sort of rolled out all these recommended systems and stuff in society to be the equivalent of sort of Victorian chimney sweeps and like, you know, carcinogenic, certain cat compounds.. Like I can't believe they used to do that.. So I think that it's gonna become just a part of life.. If we start to think about these sort of things, it's a new technologies that's emerging, right? So I think will bake into everything.. - Russ.. 00:57:22,470 - I think it's gonna be about 00:57:24,950 understanding governance and embedding responsibility throughout the organization.. The ethical theories need to be expanded, to think about distributed responsibility..

They're not very good at that right now.. And they have to be expanded to understand how governance decisions have a direct impact on at the ethical decision making.. (sci-fi music)..