\*Disclaimer: All transcripts are provided as a resource and are not guaranteed complete accuracy

You're still muted but I was actually I just got a call from my

grad assistant reminding me to um to record so okay I just did that uh

because he knows me that I'd forget so I'm doing well I guess we're are we doing people here we're here with a couple of people in the waiting room oh

gosh the middle okay people are being admitted um hi everybody hi Howard hi Carolyn

um hello yeah I know so um uh I'm doing

well thank you again so much for doing this I'm excited kind of bringing chat

GPT I just heard a great interview with a a warden Professor who's like yes I'm requiring my students to use it it's not

a matter of blocking it has to be part of all education so well I hope that like I want you to talk during that

time too because I it's like at the end it's like two slides it's so funny I mean I was like in this rabbit hole

about it then I like had to pull myself out of it because I started to get real freaked out so I think we can just chat

about you know I feel like a little out of touch sometimes with everything that's happening so I want I'm curious

to see like what your it's a great topic to bring up so let's uh well we'll wait for people to join

I'll probably start about 1202 so hi everybody we'll start in about uh three

or four minutes

well John you know me I see you're on I hear your uh you know I'm constitutionally against ever starting

anything on time but give us one more minute for people to join then we'll start up it's good that you're here John you're on time see you you're an on time

kind of person and you're on mute you know you know the expression if you're if if you're on

time you're late you're late urges if you're on time you're early that's it yeah

it's the kid Eric Meslin he always uh started meetings on time I said you really you should have been more part of

the Bush White House not the not the Clinton White House Lynn White House famously behind yes so

okay well 1201 I'll go ahead and get started so it's my I'll take the liberty of introducing our

speaker uh Dr Chang um who uh has given a version of this

talk in some of our courses I think I've never seen it though because I think it's the course that actually Colin Halverson runs

um and uh I've heard only amazing things about it and she's tweaked it for this setting to be more I think more about

research uh ethics and less about clinical but again I expect this to be an outstanding talk no pressure

Chang is assistant professor at Madison Pediatrics I have I have been honored

and I've learned every time I've come to your mentorship meeting um it's often the mentee is the mentor I

learned from America all the time and it's my pleasure to have her here this is a the next version of our treats

series translational research ethics applied topics that's a series that the bsap program the bioethics program of

the Indiana CTSI has run for a bunch of years he's uh talks are recorded put

online with some supporting documents uh Erica will send me some supporting documents I said after and we'll post it

all uh for researchers the idea is to give an introduction to some of the key topics to think about as you do

research in one of these areas in this case uh artificial intelligence so it is my pleasure to hand this over to Dr.

Cheng and to sit back up and get out of the way I look forward to participating

in the discussion uh afterwards uh Dr. Cheng thank you Dr. Schwartz that was so

nice of you um it's like the warmest introduction I've ever received um I'm really happy to be here and to

give this chat I'm going to share my screen um and start my slideshow

can you all see what you're supposed to see are you seeing my notes nope that's right okay

um so I'm just going to start I'm an assistant professor and then apartment appease as Peter said I do I use a little bit of

machine learning in my research and so it just sort of opened up the door to this whole world of artificial

intelligence I mean no by no means an expert in a topic but I'm going to share with you what I do know is hopefully we

can talk about it and what the ethical implications are of research using these methods and then applying these methods

in your clinical practice so um I'll start by you know just sort of defining what machine learning is and

why it matters and I will say that both AI and machine learning they nest under this field of data science

um where the purpose is to extract insights from data both data science and

machine learning they analyze the really large amounts of data to deliver value to some end user

um so AI describes a machine that's capable of imitating and Performing intelligent human behavior so things

like translation abilities and then machine learning is the largest and most common application of AI and that

describes machines that are taught to learn oops sorry that are taught to learn and to make decisions by examining

the data so these machines they make suggestions they make predictions based on analyzing all these data and they

perform tasks that you know we think like require human intelligence um when companies today are deploying

artificial intelligence they're most likely using machine learning so much so that the two terms are often used

interchangeably but the key difference here is that machine learning and artificial um is really like learning

from the data uh and in ways that we never thought was possible in fact I'm

going to date myself here but this is a I always show this and like some people chuckle and some people look at me like they have no idea what I'm talking about

but there's this movie that came out 20 years ago starring Tom Cruise I know he looks the same right now but this was

from 20 years ago um called Minority Report and the promise behind it was that Tom crew it

was set in the future and some dystopian future and Tom Cruise was a police officer in charge of this like very

special unit that was able to see the future to see who was predicting crime

before the crime was committed and then he would go out and arrest those people to kind of eliminate crime from society

by identifying prior to the crime who was about to commit it and the

thought that like a computer or you know could learn without being explicitly programmed to learn is like this

completely crazy notion right like at the time this was marketed as a sci-fi movie

um and uh it hit Main Street but the reality is that now like these kinds of things are happening every day and

it's no longer science fiction um there was a recent survey from Deloitte that found over two-thirds of

companies I use machine learning to make all sorts of business decisions so I think that you know today the reality is

that um basically every industry has access to these huge amounts of data and they use the data to gain advantage over

their competitors using machine learning so this includes applications to social media the justice system government

Healthcare our decision making including diagnosing things triage financial institutions use it for

online fraud detection retail companies use it for things like personalized marketing and it's not just aspects of

these industries that are using the technology it is the entire industry so Facebook is

a really great example of how ubiquitous machine learning has become because really nearly everything that you engage

with on the platform uses machine learning in some way and I'm going to talk about Facebook a lot today I'm sorry but it's a really good example of

just how machine learning is all around us all the time so Facebook has a facial

recognition system that it uses to just tags and to help your friends recognize each other in photographs it has textual

analysis that takes the words in your posts and use it to guide advertising it's news feed it's friend

recommendations um it uses um machine learning to analyze your age your gender your location your page

likes your interest the mobile data um and your mobile data to profile you into different categories and then show

and then it uses that to show you ads that are specifically targeted towards those categories they have an in-house

translator program to make sure that users all over the world can use the platform and then of course it has this

news feed which delivers prioritizes and even hides stories based on how you enter with there's a chat let me just

make sure I'm not missing anything oh it's just me uh Erica I basically

people were in the waiting room and I did not see them now there now they're all in and I send them a chat to let

them know that you just started again I apologize I'm just talking about Facebook I'm sure everybody knows this already

um but it's just sort of a primer of what of how big data is all over all over the place um

so in a bit you know I'm going to talk about the ethical implications of this in various settings but first I really want to explain I want to take a step

back and explain why these tools were developed in the first place and what really all the fuss was about so why was

there just all this optimism about machine learning um to make high-stakes decisions and

here's one example that I think illustrates it really well it's from employment where there was some really seminal work dating back almost 20 years

some of you are maybe familiar with these studies um where the authors sent resumes to employers that were basically identical

with the only thing that the authors varied um were signals of the applicant's race so in the case of the

paper posted here uh the authors alter the names of the people applying such that the names were you know white

sounding versus more distinctively black names and everything else on the resumes was identical

in this seminal study aimed to determine how the Callback rates uh differed based on the perceived rates of the applicants

meaning were the employers more likely to reach out to the applicant given their race and what they found was

pretty damning right so the Callback rate for applicants with white signing names was 50 higher uh for those with

black sounding names um with everything else being the same and you might think well you know this

study was conducted 20 years ago there's been a huge investment on social awareness about implicit bias that's

taken place since then and so surely like there's been progress and this wouldn't happen now but unfortunately a

somewhat recent meta-analysis that was done in 2017 and found that there hasn't really been a big change in this rate

and I think we can all agree that that's quite discouraging um and so one way to think about how to

deal with this is to try to formalize the decision-making process even more by you know limiting the degree to which

humans are involved in the process so there's this thought that machine learning is going to be this kind of Pinnacle for formal decision making

right like it's going to be it's not going to be us sitting down and like writing out a heuristic about what we

think a good a good employee might do we're going to actually learn what a good employee is from data uh the data

are actually going to come right directly from you know to us and through machine learning we're going to make

sure that our decisions are not discriminatory we're going to automate these decisions and we can eliminate the rule for any sort of subjective

assessment um that often happens when decisions are made by humans

and this was just this real compelling vision and it led to articles like this one in 2015 published in New York Times

where the author said you know hiring could become faster and less expensive and their data could lead recruiters to

more highly skilled people who are better matches for their companies another potential result is the more

diverse workplace the software relies on data to service candidates from a wide variety of places and match their skills

to the job requirements free of human biases lovely thought

and then a few weeks later you know the same author uh published a different story that kind of spelled out with some

of the risks are in fact when we do use data to drive decisions like it doesn't really mean that those decisions are

subtly neutral or objective or fair you know they're actually quite vulnerable to a lot of the biases that we thought

that they might stamp out so this is what we're discovering now right so the quote in the same author a couple weeks

later said but software is not free of human intelligence algorithms are written and maintained by people and

machine learning algorithms adjust what they do based on people's behavior and as a result algorithms can reinforce

human prejudices and so the programs that might show these two articles that again were released within a couple

weeks span is that they illustrate the shift in discussion that we've seen over the past five ten years or so where

there was one this all this enthusiasm and hope that machine learning um and AI algorithm algorithms were

broadly being this like forced to advance civil rights and protect people against discrimination

um but you know increasingly we see that there's a realization that this not this it really isn't an automatic feature of

using this technology that there's many ways that AI can lead to similar problems that human decision making

um of human decision making that it was intended to replace and so you know I think there's like a reason to believe

that this is all still really a possibility um but I think you know despite This Promise there's really there's risks

there's ethical issues that I I want us to talk about today and to be aware of um as you encounter these tools I

mean you know when I first put this talk together I it was all about the drawbacks of AI but you know as sort of

a warning against implementing these algorithms of doing research using these tools but the reality is that AI is here

to stay and it's used especially in healthcare and in research is only going to become more widespread so I want as I

really want people to think about how we can use AI responsibly and I hope that like the information that we talk about

today can help you guide your research um when you're when you're using AI in

your own research and in your clinical practice when your clinical practice is augmented by AI then it can kind of guide your clinical decision making too

um so how does this all play out in healthcare well it might come to no surprise to you but AI is everywhere in

healthcare uh the huge variety of big data that we have access to opens the door to these you know previously

unimaginable applications of data mining for disease prevention diagnosis and treatment

um the thought is that AI and Healthcare can close gaps and access predict clinical outcomes guide Health

Care assessments inform the delivery of individual tailored interventions classify individuals by Health relevant

characteristics you know this is in this Precision medicine initiative um so I'm just going to talk briefly

about some of the applications of AI and Healthcare that we see so one way that AI is being used is when

prioritizing who should receive care so this was particularly relevant recently during the covid pandemic when you know

acute care resources were really stretched and there was a demand um for ways to identify patients who

might be most in need of the media care and who might benefit most from it um during that time excuse me

epic um developed an index that was implemented in Healthcare Systems across the United States with the goal it was

called a deterioration index and the goal was to help Physicians decide when

to move a patient in or out of intensive care so the thought was that there are only there's you know there's limited number of intensive care space if you

can predict who's going to benefit the most from a from a bed then you can move them out and vice versa so you know all

across the country this score was displayed on the chart of every patient that was admitted to the hospital

um and you could kind of see it in the lower right hand corner there um and uh Hospital Systems used it to

decide who should receive D at the time that it was implemented it had it hadn't been independently

evaluated Epic build the algorithm and without there being a lot of study about it but since

then Studies have shown that there's this one study on the left of uh almost 400 coveted patients admitted to a

Michigan medicine um saw that the index was moderately successful discriminating between low

risk patients and those who are high risk of being transferred to an ICU getting placed on a ventilator or dying

when admitted to the hospital um AI is also being used to share the

expertise and performance of Specialists and to supplement providers who might otherwise lack that expertise or who

might not be available so these um these algorithms are really common um for use uh uh they can read x-ray

images to give diagnoses that would otherwise require like an ophthalmologist or an oncologist or

radiologist so using these programs a general practitioner a technician or

even the patient themselves can reach the same can reach the same conclusion that would normally require specialized

medicine so you can see that this might have really important implications for providing care and under in underserved

areas in rural areas um it's really good for early detection for tumor monitoring Etc so in that way

it sort of increases access to care um AI is really flashy and it's use of

prediction so in this case it's doing things that human providers even really good providers can't do so this is an

example from Google Health where they've developed a program that can predict the onset of acute kidney injury up to a

couple days two days before the injury occurs and so normally in current Medical Practice uh the injury often

isn't noticed until after algorithm you know the thoughts are that

these algorithms performance and I think you know we can

all agree that's a really exciting concept um and then there's this concept of personalized medicine

um so if you think about how AI is being there's how AI is being used to deliver care to cancer patients so cancer

diagnoses can be really complicated both for Physicians who make the decisions and um

um and for patients have to understand sort of the risks and success of different kinds of treatments and so now

people are developing AI models that can help streamline this process by taking information from all different kinds of

sources from blood tests x-ray images genetic information from tissue biopsies

and training a model that can rapidly consult provide a accurate prediction of the

patient's diagnosis about different treatment options that were most likely to succeed about the prognosis and this

you know in turn can Empower Physicians to give their patients a treatment that they know is going to work for them and

and empowers patients to um receive information and you know to create more of a shaking kind of model

um and so these are sort of the types of models that are being deployed to predict Health outcomes and ultimately

improve the delivery of Health Care and you know sounds great right no problems I'm done talking

um you know I think that the the whole idea of this is really uh lovely Ai and

and big data have created like these great opportunities for clinical practice for healthcare delivery for your research

um but I think that there are also some really important questions um and ethical challenges that weren't

discussion so much so that some have suggested that to navigate the increasing complexity of

Life online we have to completely rethink our conceptual Frameworks including our ethical uh um and legal

ones and which ethical medical issues are am I talking about so

um I basically I'm going to talk about the following in terms of uh research and application of the AI so I'm talking

about informed consent data transparency and fairness privacy and safety and

accountability so when I'm talking about informed consent as you know informed consent

enacts respect for research participants and promotes their fundamental ethical interests by disclosing information

about a study's purpose risks procedures including procedures for handling personal samples and data as well as

measures to cope with harms resulting from participation so you know my question to you is how is

this even possible um in an online environment when you're collecting your data um in these big data situations

so I'm going to use this iconic Facebook study as an example sorry again to bring up Facebook um for those of you who I'll just

briefly go over these facts for those of you who might not remember um for one week in 2012 Facebook altered

uh the algorithms it used to determine which status updates occurred in the news feed of nearly 700 000 randomly

selected users which is about one in every 2500 users at the time

um the users had information that appeared on their news feeds altered by Facebook such that you know it was

either mostly positive Stories versus mostly negative stories um the researchers wanted to investigate

whether emotional contagion was possible uh via social networks so in other words you know can a person's mood be

manipulated by the content that they saw on their Facebook feed um importantly the experiment was

conducted without um prior consent from these randomly selected users and the study took place for a week and

then there was a there was a study that was quickly published afterwards in um you know a very

well very high impact journal and got a lot of attention um so my question for you and I was

gonna put a little poll in here but I actually don't know how I know Peter does it all the time but I don't know how but the question I you know that I have for you is what you know was this

ethical um you know what do people think that this wasn't an okay thing oh I don't

want to spoil the surprise um I don't want I mean do people think that this is an okay thing to do and normally I'd

have you all raise your hands and tell me but just think about it for a second like what are the ethical issues behind a study like this

um so the study got a lot of press and there was a lot of positive press and for you know at first there was an

article that said you know one first to conduct painstaking personal interviews with subjects scientists can now sit at

a screen and instantly play with the digital experiences of millions of Internet users this is the future of human subjects

research um but then you know I think as people started to really think about this there became some negative press along with it

especially because the authors themselves you know one was employed at Facebook but the other were research

professors from from universities and so employed by universities with irbs and Gene research training and so there was

a lot of you know why the study um wasn't subject to human you know wasn't human subjects research in the

sense that it should have gone through an IRB what are the issues with informed consent um and that led to the same Journal

um editorial board publishing an expression of concern and correction where they said questions have been

raised about the principles of informed consent and the opportunity to opt out in connection with research in this paper

um so you know despite publishing the paper a month later the editor the

editors expressed concern about these this informed consent issue and you know going back to whether or

not this is ethical the question becomes had Facebook been given permission by an independent IRB to perform the research

and the answer is no because Facebook is a private company they're not governed

by the common rule they're not required to get an IRB and then what about informed consent you might ask well you

know there are terms of agreement which is on this part on the right this little part where you click on their um terms

and use which is the circle part if you click on that and you read it there's a line in there that says that Facebook may use the information we receive about

you for internal operations including troubleshooting data analysis Testing Research and service improvement so does

that you know do you think that applies to existing user data in other words you know allowing Facebook to observe your

data um or does that mean that Facebook can manipulate the data that the users see you know are either of those things okay

are both of those things okay um and I think this brings up another

important issue with big data analysis is that this thought of like repurposed data so in this example you know

Facebook again they alter the home pages of millions of users to include this I voted box with a get out and vote

message um and then they put pictures of um friends who said they had already voted

so nobody you know opted into this like this is just something that they wanted to see what happened that they

manipulated these screens that they could increase voter turnout and then they found they did this to 60 plus

million people um and they found that it did generate an additional half a million votes you

know according to their methods of analysis um and again you know they said that

this was sort of under the same data use agreement that everybody clicks on in order to use the platform that gave that

gave Facebook permission to do this kind of this kind of feed of page manipulation and news feed manipulation

so I think you know when you

um but one of the big features of Big Data analysis is like when you are using existing data sets to yield new

information you know a human would find it nearly impossible to search through tens of thousands of medical records to

discover these new patterns and novel insights but an AI algorithm can do that you know in a moment

um and so this might be really beneficial in some contexts but

um but informed consent like do you think that informed consent needs to be obtained if the original consent is no

longer applicable so for somebody who can sense you know somebody who consented to use their postal code may

wish to withdraw consent if they learn that the data are then going to be used to just you know to determine somebody else's insurance premiums but a person

or a person might initially consent to having Facebook you know their likes publicly viewable on the page but they

might not consent to having a third party you know use that to see to manipulate voter turnout

um or to figure out you know to predict her sexual orientation or to use those predictions to Target ads towards them

for most users that kind of data sharing likely goes beyond what you think you're

doing when you're consenting to using the platform so when I turn on the location services on my phone I just want to check the weather you know I

don't think about I'm not I don't think that I'm consenting towards some future use of the data that I'm putting into my

you know that I'm typing into my phone and I think that's a really important question about who's governor who's making the rules for this and who's

governing that and the answer right now is you know nobody really um

another ethical issue with this type of research relates to transparency and fairness or bias I think researchers I

think researchers have an obligation to describe their findings and to understand the outputs of their data models

um but you know when I prefer a logistic regression I can very clearly articulate my thought process and my conceptual

model behind that and what the output means but AI is not transparent

by Design and so we can be led to believe that we understand something but we can't really understand what's going

on and the larger and more complex a model is the harder it is to understand even though its performance is generally

better that's the whole point of AI and I think this is a problem because we are adopting technologies that we don't

really understand like on one hand um you know you can't understand the data

and that's probably bad but on the other hand you know how is this different from other Technologies do we use like our ubiquitous use of computers or we don't

really understand how our car Works etc. like is that okay is that a problem um you know with AI many of the

algorithms are produced by private companies and there isn't any regulation guiding in the field which contributes to like a real lack of transparency

regarding how the models are produced there are you know best practices for validation but these practices aren't

regulated either and they're likely to be really domain specific and because many systems are these black boxes the

reasons why the decisions are made the way they are are not easily accessed or

understood by humans and therefore they're difficult to go back and and probe and ask questions

so what I you know what I mean by this in other words like with a traditional statistical analysis on the top box

right we have control over what we're putting in the model and how we measure things and how we code things and so we

can tell the model to ignore or to not ignore certain features so for helping build a model guide like hiring

decisions we might program the model to weigh certain characteristics like prior work experience more heavily than others

and then ignore you know and also tell the program to ignore certain features like race

and this means that the results that we're seeing you know we're very confident are not biased or influenced

by these factors but with a machine learning model we're not programming the model in the same way we're not supposed

to know like how the machine machine's thinking all on its own we're not supposed to know how it's doing that and so we can't really determine how these

factors are considered um and I just want to pause for a second

and think about this notion of bias because when people talk about you know fairness and ethics and bias and machine learning uh especially when you're

coming from a technical background it can be really confusing you know what this term actually means so for those of you that are statisticians

um you know when you hear the word bias you might think about the following things so the idea first that the way

that you're collecting the data might be biased in the sense that you're not collecting a sample that is representative of the population like

selection bias um you could also be talking about bias of an estimator meaning like the predictions that you're making are kind

of off the mark that they're imprecise um but in the case of machine learning I'm really talking about something that

we would think of as inductive bias which is the way um we're trying to focus on a specific

pattern of the data or what we're trying to learn um that is not what we're that is unjust

or unfair so the differences that are created um in the from the model are

in themselves biased in the sense that they're not they're not fair you know I think that the purpose of machine

learning you know is to differentiate right it's to determine who's at risk for something uh which requires you to

categorize into some sort so thinking through bias today I'm really referring to the unjust and unfair biases for

differentiation and when we say that like some particular feature about a person for example their race or their

gender that isn't really relevant to the decision at hand is embedded into the model and the highlighted in the model

and I think this is another interesting issue because there are circumstances when teachers like gender or your race

might be predictive of certain outcomes but we've decided you know by society

and by law that these things shouldn't be considered because you know we've deemed them to be morally irrelevant

even if they are predictive so in a traditional statistical way you would not put those in the model but AI

doesn't know that and so AI and so AI will

and one way we see the problems with how this plays out um the selection advice from um so you know

I mentioned the selection bias from a statistical sense might create you know some sort of normative bias and I think

this is really well Illustrated in how AI is used in criminal justice specifically when arrest information is

used to train models for predictive policing so judges and courts have started to rely on machine learning to

guide their sentencing decisions currently courts and Corrections departments around the U.S they use algorithms to determine a defense risk

which ranges from the probability that the individual will commit another crime to you know the likelihood that they'll

appear for his or her court date and then these algorithm outputs are informed decision making about bail

about sentencing about parole you know so you know questions who do you actually incarcerate who do you keep

pending trial how long you know what do you set their bail at and I think again there are a lot of benefits here in

theory so if a computer could accurately predict which defendants were more likely to commit a crime the criminal

justice system might be fairer it might be more selective about who's incarcerated and for how long

um but I really what I really want like what's really important to understand here is that arrest records which are

the data being used to make these models they're not a perfect representation of

the incidence of crime and Society just like you know medical record data are not a perfect representation of who's

sick and who you know who's healthy and who's not um I think police records you know

somebody else said and I have this quote written down here that police records are really you know quote a complex interaction between criminality policing

strategy and community policing relations so you know you're those records are not really capturing

crime as it occurs everywhere it's capturing crime where the police happen to see it so it's the type of crimes

that people are likely to report given the relationship with the police and if we build a model using these data we

won't necessarily build a model that predicts crime wherever it might occur we're going to build a model that

predicts crime where the police have been able to observe it in the past and this contributes to this vicious

cycle where you build a model and you suggest that you deploy police to a particular area and that those

predictions are confirmed because you actually do observe crime there but you've also directed your resources to a

particular area that gives you less opportunity to view crime in other places so you never get disconfirmatory

evidence from your model and the predictions keep getting confirmed over time because the model never has a chance to learn from other data does

that make sense this kind of vicious cycle this is like a huge problem in Ai and it occurs you know all in all types

of decision making where disadvantaged groups are you know are typically judged by these algorithms more harshly than

other groups and that can have devastating effects for these populations over time

little power um for example this often leads to bias in healthcare delivery

um so you know these models as I said they often replicate the biases and blind spots of the humans who create

them so the algorithms are you know built on these biased data are the homages data sets that don't really

reflect the patient population at large if you think about Hospital you know when you're using big data to build a

prediction to research to build a model for health care the biases that are embedded in the

healthcare data itself are going to be replicated in your model so who you know is the data coming from the IU Health

System really reflective of who is ill or are they reflective do they are they

do they have embedded biases of the number from other things um you could think that that's also true

if data is primarily collected from like an economic Medical Institution the AI systems that you'll produce are

going to learn a lot less about communities that don't visit academic Medical Systems and then they'll be less

likely to treat those patients as effectively um

and then I think this leads to a lot of racial bias in the system as well so these algorithms you know have been

criticized for propagating race-based medicine uh so the racial differences

found in a lot of data sets you know the disparities that we see between black and white patients are most often

effects of sort of racism or you know the experience of being black rather than being black itself things like

toxic stress poverty the physiological consequences from that and so adjusting for race is not going to do anything to

address those root causes of disparity instead if the adjustments deter clinicians from offering Clinical

Services to those specific patients then they risk sort of baking inequity into the system and the Machine learning

model is not going to know that you know we think that the computer is going to take away the biases but they're really embedding the bias is back into the

system so in this study there was a clinical algorithm that many hospitals were using to decide which patients

needed care um so this model was being used to predict intensive care for patients who

had complex medical needs so when this algorithm this had been going on

for a while and then researchers decided to evaluate it and they found that black patients um had to be deemed a lot

sicker than white patients to be recommended for the same Care by the algorithm and this happened because the

algorithm had been trained using patient data on prior Healthcare spending which reflects the history in which black

patients spend less on their health care not because they have fewer you know not because they have less need but because

they have fewer resources and of course the machine doesn't know this so they returned a higher risk score for white

patients because they said oh white patients are spending more money in healthcare they must be sicker leading more referrals to White patients for

specialty services and that sort of perpetuated this both you know the spending discrepancies that we saw and the race biased in healthcare

uh another set of ethical issues Rises around privacy so the requirement of

large data sets creates incentives for developers to collect data from many patients

um so some patients might be concerned that this collection might violate their privacy lawsuits have been filed Based

on data sharing which we enlarge Health Systems and AI developers um you know I think it's also

interesting to think about how AI can implicate privacy in another way um when it predicts information about

patients even though the algorithm never received that information so you know this is kind of the goal of healthcare

AI is prediction so for instance you know an AI system might be able to identify that a person has Parkinson's

disease based on the trembling you know they're hand trembling over their computer mouse so even if that person

never revealed that information to anyone or didn't even realize it themselves you know patients might

consider this to be a violation of their privacy especially if AI um especially if the AI interface

through some data use agreement was that information was available to the third party such as to a bank or a life

insurance company um Healthcare privacy regulations are in

place for HIPAA but they don't cover tech companies so HIPAA only protects patient Health Data when it comes from

organizations that provide the health care services um so green has gone has gone black is

that just me or other people seeing that do you are you showing slides still something has happened I can see them I

can see them as well I couldn't see them I'm going to be quiet now well it's um it's a screen that says AI

HIPAA and then it has um pictures of a smartphone and it has this suicide algorithm Again by Facebook

so Facebook created this in 2017 they rolled out the suicide detection algorithm in an effort like a reliable

effort to promote suicide awareness and prevention so the um the system uses AI

to gather data from your posts and then to predict your mental state and your propensity to commit suicide

abuse case for AI and Healthcare um but you know benevolent intent aside the fact remains that Facebook is

gathering data and storing Mental Health Data and doing it without your explicit consent you know you consent when you

sign the user or use agreement in the past whenever you sign up for Facebook 10 years ago but that doesn't necessarily mean you're consenting for

the specific use of your data um and you know if you like and nobody

really knows what they're going to do with the data now beyond their stated purpose now that they have it so in

principle like you can imagine that Facebook gathers this data or step data from the app on your smartphone Your

Health Data from my watch you know you know my fitness tracker and then they buy Healthcare data from another company

they match the two then they have Healthcare data that's matched to names and they can start selling Advertising

based on you know based on that or they could sell that data to others and I think that just like opens up a real

scary sort of Preston and then you know anecdotally I was um at a conference and I was out to dinner with somebody who

has four teenagers and she was talking about this app on

her phone and now I can't remember what it was called but it was something really creepy and anyway the app tracks you know what her what her teenagers are

texting what they um what they're Googling and so that she knows like and

then it'll send her alerts if the algorithm thinks that her kids are about ready to go about drinking if they're thinking about having sex if they've

done something illegal and then she gets these alerts on her phone it's called like Mama Bear or something and you know I mean the thought is that like that

kind of use of big data to me is a little um Big Brother a little creepy

um okay safety and accountability is another ethical issue the last ethical

issue I think we're going to talk about um and the you know the point here is that AI systems can sometimes be wrong and that patient injury or other kind of

problems can result from that so you might remember this story it's not a Facebook story but you might remember the story of 2018 when there was a

self-driving Uber who killed a pedestrian in Arizona um and this happened because the researchers who programmed this

algorithm they only recognized they programmed it so it only recognized pedestrians when they were crossing a

crosswalk not when they were crossing in the middle of the street probably because they'd like developed this in California where everybody abides my

crossbar Clause so um you know this seems like a real simple thing right like one that's really easy to predict like sometimes

people cross the middle of the street but the program didn't have that information and so it failed and

somebody died um Healthcare has so many situations like this where human behavior is hard to

predict right so like did your patient take only half of the blood pressure medication because it makes them feel dizzy and then they felt bad they didn't

tell you about it um does the patient take like herbal supplements and then just kind of forgot to mention them do they secretly smoke

but they now they might not want their partner to find out is their mom in the room so they don't want to tell you that they're having sex so all of these sort

of complexities make it even more likely that the machines are going to make mistakes so if an AI system recommends

the wrong drug for a patient or it fails to notice a tumor on a scan or allocate

a hospital Bell to or an allocates a hospital bed to one patient over another who eventually dies

um you know who who's responsible when AI makes these mistakes is it you who made the

algorithm is it the clinician who is implementing the algorithm is it the machine itself you know and I think that

like what's real important here is that um the responsible party is still unclear

we don't know there's no resolution to that question and then finally you know monitoring

intervening during data collection as a researcher you have an ethical obligation to monitor the emergence of

Adverse Events and to potentially intervene um to mitigate negative outcomes so how is this feasible

um in the world of Big Data you know what's the trade-off here between participant safety and privacy and how

does this affect data validity these are all really important questions that people don't yet have the answers to

um okay so in the last couple minutes I said I would talk about chat GPT

um which is hot news right now and I was I was telling Peter before I started talking but I like started to do a deep

dive and chat gbt and then I had to really pull myself out of it because I started to get real panicked about the

future um but basically Chad gbt is an AI program that uh people use for

generating dialogue so it's a chat bot um with a language based like neural processing model that

um is able to you know find you're able to fine-tune it for for conversational

manner so the main feature is that it generates responses like a human would provide like in a text box so

um like people could you think that people could use it for like a virtual assistant or something like it doesn't come across this machine it really comes

across as somebody that you're actually chatting with um people think for research purposes

you know you can use it to write code you can use it for research dissemination to write articles for translating for writing stories and this

has gotten so much Buzz over the past couple weeks that people have said stuff like chatGPT is one of those rare

moments in technology where you see a glimmer of how everything is going to be going forward I mean that's intense

right like this is the future of how everything is going to be um I went on chatGPT because I wanted

to use it to write this presentation I was like oh maybe I'll just you know kind of chat it up but in the system

had crashed so I will say that I am not sure uh I can't speak from experience

but I know Peter has a story about his experience with chapter I wasn't allowed to use it but

um people are like all um and it proposed correct options including the correct one so this was

verified by a cardiologist who was surprised so you know the saw the chat jpg might be really good in in clinical

care application and also for research dissemination so

can chat GPT write your research papers for you is that okay you know how do you

feel about how do you feel about that I think that that's something that I do want us to spend you know we have almost

20 minutes to talk about all of these issues um about you know how you balance what

researchers want versus sort of the ethical risks behind using these this technology these are not going away

um it's part of research life now and I think that it's important that we have these kinds of chats because I know it's

it's like as it evolves it's hard to embed it into sort of the curriculum that we have so I have you know ethical

questions up here but I don't know that we really need them um because I think

that there's a lot that we can just kind of chat about on our own so I will stop talking and stop my screen share

um uh stop sure

thank you very much everybody for patience

oh Peter you're muted as usual so I'm going to want to field your questions and conversations so go at it everybody I

assume we've got many thoughts um Peter I actually want to hear your

chat jbt story because I like said that I would talk about it and then it was only like 30 seconds of my presentation so I kind of let me hear what you have

to say well I'm happy to so I just my teenage son uh my 10th grader showed it

to me about two months ago a month ago and basically he'd written into it a uh the prompt he'd had for an English essay

which was like discuss gender and Power in this obscure short story they had read and I thought oh you can't do that

I mean Google's amazing but it can't it wrote a four paragraph you know solid B

I mean it made me a little repetitive little like formulaic but you know I I'm a teacher it was a B so I'm like oh

that's a fluke so I put in my favorite which was um uh you know could discuss physician

assistant suicide using a utilitarian framework classic thing I would do in one of my classes not bad B minus C plus

you know passing grade for sure a little repetitive a little simplistic but really very good it's not that it's

cutting and pasting right it's actually an AI apparently it's going it's finding data and it's putting it together into a paragraph as you said

Erica you know it looks like a person and of course this is what we've been expecting all this time and I will say about all the things you raised and

about this discussion I feel like you show a Minority Report I always go to the Terminator right it's like the

The Singularity that that machines take Consciousness and it's the end of the world and in some ways we can't talk

about that that's not a medical ethics topic but in other ways it's the it's the topic here which is when

machines can think what do we do and the question you're asking here which I'm sure if you put a lot of thoughts about is as we're doing research now you know

Arnold Schwarzenegger is still an old man in California not a uh cyborg from the future

um we can um we have we can use AI in our research we're involved in

actually moving AI forward how do we handle the ethics I think you did a great job with an enormous topic of

showing some of the issues I'd be very interested if you want to talk about chat gbt go ahead um you want to talk about uh the

research stuff that that Erica focused on that's great too so go ahead um everybody and I see John's come on video

which usually means he's ready to talk but maybe not you know also to listen

um thank you that was that was very interesting thank you very much um I'm and I'm completely

unknowing about a lot of the AI stuff uh it was interesting to hear about your

talk about the um the ethical issues in the predictive use of AI

I remember I was I don't know it was decades ago I was in Massachusetts in the Massachusetts Department of Juvenile

Justice whatever it was called wanted to create a super max for the juveniles that were locked up in order they said

to capture the up to five percent that might commit might try to commit suicide

or hurt themselves during their incarceration and at that time using just computers somebody I knew said that

in order to capture that five percent they would have had to put somewhere around 72 percent of the kids in that

super Max and so what's interesting is just to see the advancement of the predictive stuff in terms of the ability

to see greater day more intense and more detailed dangers of the predictive model

yeah I think

using AI in my own stuff so I'm using it for prediction as well so I'm taking data from uh kids birth certificates and

their you know first couple their year the first two years of life their well-child visits and then some

information from their neighborhood and using that in a machine learning model to predict their reset their risk

for obesity when they're older because like with obesity it's one of those things that is really hard to shed the

weight once you have it but so the thought is move upstream and predict who predicts who's going to have the weight

before the weight comes on then you know perhaps we can we can intervene and move the needle a little bit but you know

it's really hard because how do you communicate that risk score to a parent like what do you say to a parent of a

six month old they're going to ask well why you know what is it about my child that that

um is going to make them more at risk for obesity and the question the answer will be well I don't know something in the model that's happening

behind the scenes that I don't know about because you know it's thinking all by itself you know and then this other question it's going to show

disproportionate views based on race is that right is that okay like it's not about a person's race that makes them

have more likelihood for obesity it's all these other things and so there's you know it's not a perfect you know I

think people think the machine's going to do it better and I don't think that's true I think that in our in our application of AI it's more about like

how can we use that to augment what we know you know like let's not let the machine think for us let's think with it

maybe so you literally were just using birth

certificate and what other data uh well child visit and then uh Geo like

census geodata from their neighborhood up to what age

24 months wow

I mean I'm just thinking of all the I'm a sociologist so I think of all the of the more social aspects that contribute

to obesity right that would not be captured in that correct I see

hand raised uh you're muted

you're muted yeah he's working sorry about that I think doing it on an iPad

and it's a little slower to unmute or do these other things you pointed out a lot of the issues in bias and some of them

are really unconscious some of them are hard to deal with uh I I'd make a quick

prediction about what you just said which is the zip code would do a reasonable job of predicting obesity and

the problem there is all sociological um you know the stores around there the

uh what's available to you in shops etc. it differs greatly in minority

communities for example or any underserved or under under-resourced communities and that's always a big

problem um you know the the I've been involved actually and and I

submitted some information to the FDA actually because a group was

a company was wanting approval for its AI thing

which came out as a device for the FDA by the way to predict who is likely to become an

opioid abuser or opioid addicted if you treat their pain with opioids

and a few of us saw this and were at a much earlier stage there was this person

who has a company published this a few years earlier and it was based on totally garbage science you threw in a

bunch of these things that they thought were genes that were relevant and they threw it in with a

badly chosen population and they come out with this machine learning answer

we did an exhaustive test on this using you know five different types of machine learning language because they wouldn't

reveal what their black box was with all sorts of things and what we found is it predicts race very well

and that's all it predicts it has zero predictive power about opioid use of user versus not and yet this was going

in front of the FDA for approval so obviously a number of us tried to have

are still trying to stop that so there's you know this conscious bias there's uh just badly designed

experiments um you know this is a real tangle to

deal with and one quick comment on a totally different issue in terms of the Facebook

using your data yeah you pose it as if their experiment was their only experiment on you but in fact

their Baseline algorithm is a huge experiment on you also and probably not

much different than in ethical um uh I I don't know in the general

ethics of their General Baseline model even before they tweak it you know when I started uh my faculty

appointment there was an associate professor or she had just I don't know she was another professor in my division

and she left right after I started because she got a job at Facebook and she was a social behavioral PhD real you

know real strong research um analytic skills and I just remember thinking like why in the world would she

go work at Facebook that doesn't make any sense at all and you know now of course it all connects to me because I

just didn't realize the except I'm not I'm not a social media person so I didn't realize the extent to which um they were doing it and I think your

um predictive you know your garbage in garbage out raises a lot of really really critical ethical issues about

machine learning in this space because a lot of times you know there isn't any

regulation over the development of these models as academics we are in you know what the work that I'm doing we have

these highly curated academic data sets we are able to sit down and really think through things and to really you know

make the data sort of pretty before we put it into the model but the other you know third-party institutions that are

not regulated at all that are being hired to do these things we have no idea you know who's validating those models

if they're being validated what's going on and then I think when you deliver the results and you're like oh this is machine learning it's AI it's a computer

everybody you know gets kind of caught up in like the sexiness of it all and sort of doesn't ask these questions because they don't know that they should

Christopher sorry I see your hand Christopher oh no thanks for thanks for taking the

time so I'm Chris Newman I'm with the Radiology at Riley Hospital and um so I think one ethical issue I would

like to think through in terms of these issues is uh the ability of AI to allow

these new predictive models um that will often expedite the our ability to our expedite

the gap between our ability to predict risks but also intervene on those risks and so I'll give you an example I think

the history of medicine is replete with examples where we were able to predict risk and then we started doing interventions without actually having

evidence for a lot of those interventions and those outcomes and so even one example from the code pandemic is you know paxillin was approved for

non-vaccinated patients yet we readily use it in vaccinated patients even though the only randomized controlled

trial being conducted at this point is the panoramic study in UK and we're still awaiting those results and so this

is an example where we we look at patients with back who are vaccinated who have risk factors now we're

intervening using something that actually hasn't really been well tested and so there's a lag between the ability

to predict risk factors um and the ability to actually put interventions that make a real

difference in that have evidence-based outcomes attached to their those and so

do you worry that our ability to make these risk factor predictions will exacerbate that problem now in

terms of basically having a longer delay of actually testing many of these interventions and their outcomes even

though we know a lot more about the risk factors

anybody have any thoughts about that I don't know that I I do off the gate I have to really think about that

um I like it as a comment I know we've made some other hands but I like as a comment Crest because predicting risk

because we know in the risk prediction field the smoltifying arrest is only valuable if you've got an intervention which helps right it doesn't hurt and so

all these things are really important to say well you know oh great we can figure out who's at high risk for

suicide or oh great repair who's at high risk for obesity that will be great and valuable if there's an intervention

which actually helps in addition to other issues being raised by you know what if you had a great risk stratification model that actually

relied on Race exactly and also with the AI as um who you've said the Black Box

we don't even know if it's using race although it sounds like Howard uh and his colleagues were looking at that and

figured out that actually was basically just using race it's basically yeah I mean Peter like unfortunately in our

society race usually is powerful for differentiation of risk right and like

and but that is not you know that is about so much more than you know so I

will say yes all the AI models are likely using race to predict risk

um is that the right thing to do I mean we've decided has decided that it's morally irrelevant but does that mean

that it's not like clinically relevant question of questions now we don't say

the race of the patient at the top we say at the beginning so you know African-American 40 year old female we

don't right stop anymore your interaction with the patient

shouldn't that shouldn't matter right but if you're looking at a risk score for who should you know who should be in intensive care

is that important especially save a life especially if looking at it

okay well are we at the end but quickly to the point you made about that though you get the Vicious Cycle of all the

socioeconomic things that cause the differences in in what we call Race and that's a bad word also but

self-described race is probably the clearer one and and these models build

it in again so some of these things may predict better but they're predicting better not based on things that that not

based on the reality of what's happening to the person they're based on a societal discrimination right

um and then you asked about the regulatory institutions I you know I think that they of the Facebook uh

professors I they were I think they're from UCLA I think they got their hands slapped but you know I they didn't

technically do anything wrong

well I really appreciate everybody's well I mean there's another second but I appreciate everybody's attention this is a lot it's always a

lot of fun yeah great enough of course um It's A continuing conversation um we always say at the end of these

talks that if you are doing research involving AI you want to talk to somebody come talk to the bioethics program of ctsi we call ourselves

bsap um and Erica we will I'm sure help you in all of her copious free time

um uh to we always help to help to think things through because doing this research and helping medicine become better is of course our goal but then as

uh Professor Chang has shown us uh it's just a tremendously complex area raising

issues as it maybe can help us address so again thanks everybody and thanks especially Erica for doing this thanks

thank you so much everybody has a great afternoon