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

uh is that right are you guys seeing a a screen that says data ethics de-identification

okay guys hi everybody welcome to this um treats talk uh this this is the next installment of

the translational research ethics applied topic series that is run by the uh

bioethics and subject advocacy program of the Indiana CTSI um we do these talks

uh let's say i guess before we do these talks monthly uh roughly monthly about key topics they're

meant to be a short introduction to a an issue in translational research

ethics and regulation uh they're meant to be sort of used to people who come across this topic

have questions and want a quick intro it's great to have you here and actually we love having a big group and the zoom

era of coronavirus has allowed now some bigger groups than we used to get so we get some good discussion

you get to ask the questions that may be on the minds of others who might be watching the video so we do record this and then we post it

on our website for the bioethics and subject advocacy program and we also post it on the IU center for

bioethics that way if a researcher or somebody wants to know more about the ethics and regulation of data

related to de-identification they can come to our site watch the video have

the benefit of your questions on our discussion and the benefit of some associated readings that I'll be mentioning today

and that we'll post on the site once we're done I'm peter schwartz I'm the director of

the center for bioethics and I'm the director of the bioethics and subject advocacy program at the Indiana CTSI

and it's my pleasure to be here with you today thank you for being here this is the topic we mentioned we really

do focus on making this applicable making it 20 to 30 minutes of didactic of lecture and then a

conversation so please uh feel free to yell at your questions as we go i might ask you to wait until

I'm done but I might not we might be able to just take a question as we go especially if it's a point of clarification or

something I seem to have skipped or gotten wrong I'm very thankful by the way to TJ who's the postdoc at the iu for

bioethics and to my regulatory gurus Amy waltz and Chris Caldwell who always

helped me make my make my comments in ethics actually somewhat correct about the

regulation uh regulation of research and ethics of research of course are closely related and uh we want to be

informative and helpful in both areas although myself I'm trained as an ethicist and as a

researcher rather than really a purely regulatory person so I go to Amy who's the previous head of the IU IRB

and works in ora and uh to Chris who's the regulatory knowledge support person at CTSI for a

bunch of years and Chris is a wonderful source so if you come back to us for consultation after this feel free but

also um Chris is a great resource okay so this

talk about de-identification i will tell you has been bouncing around for a few years some of the people on this call may have

even seen my previous versions of this I did it a few years ago three or four or five

years ago i had a version of this I'm gonna go back a slide

I did that three or four or five years ago and that was in the context of the previous uh common rule so before the revision of

the common rule i had a whole talk I did about this I did it at BSAP I even did it at other places um

and I've always enjoyed that talk and then the revised common rule came out the revisions to that in 2017 and so

i really updated this talk for today and i hope it's helpful it'll replace any previous guidance we've given

and part of the developments that have led to changes in this area are just the burgeoning role of data and

digital uh features of our life and so i don't have to tell anybody any of this this is

too obvious for saying but i'll say it so we of course live our life online now

doing our email doing our shopping we go to the library online we carry our phones with all those apps

that are of course picking up information about all of us we go on social media like Facebook and Instagram and all of

those our health data is now getting to be all electronic we've made it through to the electronic

health record age we pay through electronics and data is collected there

we of course now more than ever before know that education goes on electronically through e-learning for

our children and our in many cases our college students and our and our graduate

students and now research is done largely online we download our pdfs everything

is data and of course we've all seen the somewhat positive but often

concerning sides of all this that we see the monetization of data we see the ways

in which sharing data like this like crazy can lead to compromised privacy and we see of course

uh the spread of misinformation and the danger of permanent connection and everybody

having a bullhorn some people say things that are just wrong and misleading as well as other things

we lead then to problems with national security and of course election difficulties i will say here at

IU we're very interested in the ethics of data Fred Kate the vice president for research leads a project

here over the last few couple years called ethics values and technology developing character for digital world

it's funded by the Lilly endowment to try to figure out how best to equip students and faculty

to address ethical challenges presented by digital technologies so I'd say we are absolutely being um

seeped in the ethical questions involving data now today I'm going to talk about data and human subjects research especially

looking at de-identification I want to go all the way back to what we try to do in the ethics and regulation of human

subject research and I'm summarizing it very inadequately here as four basic bullet

points we want to avoid harming the participants you may have noticed the bioethics program of the CTSI which

is hosting this talk is called the research subjects advocacy program as well so we

are supposed to be advocating for our participants and protecting them we should be respecting them as well we

respect persons we do that by thinking about their autonomy their informed consent and

and other ways we want to treat people fairly through justice we want to advance science and medicine

through our research that's a key component of ethics and regulation as well

so as we think about these principles and doing research that involves the

data that's collected from human subjects we actually come across some very interesting topics which i will not have

the time this is not the place to go into great depth on them but these questions we will

touch on even today in this short introduction to these questions about ethics and regulation and the use of de-identified data is

what is harm what does it mean to harm a person is that a bodily harm are there dignitary

harms many of the issues about information raise the question of if you harm a person by violating

their privacy what would count as a harm what are the sort of harms that we in our commitment to research

subjects uh need to really focus on in that first bullet point under ethics and regulation

how do we take advantage of sorry take care of data to avoid harm to our participants in a

way that still allows that last bullet point of advancing science and medicine it's also some fascinating i think very deep

issues about what's the value of autonomy and choice and how far do those spread what do those involve

so you could say that actually one of the issues in terms of thinking about data both in the health sciences uh and also

beyond that and everything other area of digital life that i mentioned that we have there's a value to

respecting our autonomy and giving us choice about how our data is used and while that's very attractive in many

ways we should be in control there's also questions about what sort of choices people make and then how you

balance that again against these goals of advancing science and medicine could could people's choices that we

allow them to have that we maybe encourage or require them to make could that interfere with advances in science and medicine

and if so would that be an unfortunate outcome and one which would not be consistent with ethics would we be over

prioritizing or misinterpreting the nature of autonomy and choice by prioritizing that so the guidance in

the area of human service research and data mostly comes from hipaa the health information portability and

accountability act of 1996 and then of course as for all human subjects research the common rule

and now the revised common rule which was released in 2017.

some key concepts we're going to define up front are the idea of private information the idea of personally identifiable

information which is a HIPAA concept mostly emphasizing HIPAA also individually identifiable

information which is the term used in the common rule and protected health information or phi which is another hipaa term or i

just link it to HIPAA really so private information i'm not going to say much about this here we'll come back

to it later which is that some kind of information is private suggesting to you by the way that some information about you is not

private right this is actually an interesting thought right here at the beginning which is that

private information is actually not all the information about you so information that's been private information for the purposes of

research is information has been provided for specific purposes by an individual

where the individual can reasonably expect it will not be made public for example if you tell your doctor

something that's going to go into your medical record it also includes information about behavior that occurs in a context

in which individual can reasonably expect that no observation or recording is taking place so some settings where you might be

recorded or viewed are in public where your information is no longer private

we'll get back to that a little later to move on to the next concept the idea

of individually identifiable information is that information is individually identifiable when it can be linked

to a specific individual by the investigator either directly or indirectly through a coding system

so personally identifiable information is any data about a person that could potentially identify them

such as these sort of things name address driver's license number date of birth etc

now we move from individually identifiable information to a category of that a subcategory which is protected health information so

you're individually identified by information may be any of this information separate from your health

separate from anything having to do with your health but once you get into personally identifiable information you may get to specifically protected

health information which is individually identifiable and which relates to

your past present or future physical or mental health or condition getting certain kinds of health care

paying for your health care or um uh which i and and i should mention again an

individual identifiable which does actually uh identify you so it is actually a subcategory

of personal identifiable information so um uh so those are the key concepts i

wanted you to hear about uh for the concepts of data and identifiability that we'll be using now

so now let's think about human subjects research it's always good to start to remember

that research is any systematic investigation including development testing evaluation

which is designed to develop or contribute to contribute sorry to develop or

contribute to generalizable knowledge so that's going to be the area we're looking at is research

and we consider some research to be human subjects research if it involves identifiable private

information so that will if you're doing research and you're using identifiable private information

now it counts as human subjects research or again more standardly you could say it's human subjects research if you are

doing an intervention or interaction with living individuals where interaction is communication or

interpersonal contact an intervention is some sort of procedure where data is gathered so we know when we're doing human

subjects research it's talking to or physically interacting with people or

using their information using their information separately so uh

if you are doing human subject if you are doing research and you want to know whether you're

doing human subjects research um uh and you're just using data

your first question should be if that data is identifiable is it identifiable and private um so

uh uh if yes then you're doing human service research as seen by the definition on the previous page

and if not i.e the data is not private it was not collecting that private setting we talked about or it's not

identifiable then you are not doing human subjects research phi

uh HIPAA does not apply because you're not actually doing human

subjects research at all actually here i will ask my any point although I just say people ask questions

what do I know I've got a couple of my good um in uh commentators and uh

helpers here from like um if Amy or I don't know if Chris is here actually they're welcome to jump in here

but the idea is think about the information you're using if you're just using information your first question is whether that's private

identifiable information and only if it is are you doing human subjects research

um okay so now that you've now that so

so the way so one way which you would not be doing human-centered research would be if your research if your data was not identifiable and

the key concept here I'm going to try to lead you through the idea of coded information and de-identified information

so code information is for all of us researchers when we when we record information about a person

we often use coding to identify to link to that data to keep the name

identifiable information away but we keep a key to uh to translate uh from

that code uh to the person so we can re-identify the data easily if we need to and that is code

information is not de-identified because this key exists and we hold it and we can easily

re-identify the information ourselves then it's not de-identified information so code information counts as

identifiable and therefore if you're using it in research your research does count as human subjects

if you wish to avoid uh the burdens and the dangers of human

subjects research then you can get involved in trying to de-identify your data which is again our focus for today's

talk um and the privacy rule uh provides two ways to do that there's one

that almost everybody uses there's a second rule so the first way uh the one that i

i don't see much from people who come to talk to us is a formal determination by a qualified expert who looks at the data

and confirms that it cannot be used to identify individuals

and that is a way to to know that you're not using identifiable information and you therefore are not doing

human service research based on that use of that data and you can go forward the hippo rule

realized that that that's a expensive and complex process

and they provided a safe harbor approach where you can remove specified individual identifiers and

and de-identify data that way I will point out to the second part of this which is once you've eliminated the identifiable individual

identifiers you also must be able to say that there's an absence of other knowledge that would allow them to use

the remaining data to identify the individuals but as long as this is the absence of that knowledge and you remove these individual

identifiers then you can consider this data to be de-identified

so i had some there's some noise outside my room okay so the 18 identifiers

that need to be removed to make data de-identified are these um I can highlight a couple of them

um name of course would make data identifiable um

geographical subdivisions oh I see I can't do that in this setting geographical subdivisions like a street

address would make information identifiable dates related to individuals such as

specific dates of their admission and discharge from a hospital or their age in certain categories where

there are relatively few people of that age uh telephone numbers fax numbers email addresses social security etc.

if your data set does not include any of these uh pieces of information then you

consider that data set to be de-identified again given there's not some other way that you know of you can go from that

data back to the people involved now as i said coded information does not

count as de-identified and so if you hold a code that allows you to take uh some codes assigned to the data

back to the individuals by yourself then you are not dealing with the identified information the rules do allow that one entity like an IU health

or Regenstrief can assign a code to data that will allow it to be re-identified

and hold that code themselves as well if they if they release that data to somebody else they release it to

me to use in my research that data is de-identified for the purpose of my research um even though they could re-identify

that data as needed if they wanted to um but as long as I don't hold that code and there's no way for me to identify those

individuals through some other trick like using the code to um to like it's like the code is some

you know way of just coding the person's last name and I could figure that out and figure out who the data is about

as long as there's no trick like that you can de-identify data and it can be de-identified to me even

though to IU health or Regenstrief it would not be because they hold the key to take that code back to the

individuals so I'm going to pause here and talk

about why this step matters again it's given a lot of prominence in the regulation i mentioned right i mentioned back there

that if you are doing research on data you aren't involving any kinds of interaction or intervention with people

then the number one question for you about whether you're doing human service research is whether that data is identifiable or not

assuming it's private and we can see why it is seen is seen to be

so important again thinking back on our key principles of human subjects research

ethics so in terms of harm if a data set is de-identified

then presumably and again I'll put that a little bit in uh I'll say presumably the harm to

individuals that might come from using that data is minimized or eliminated

right so uh if it's information about you know an individual doing something

they would be embarrassed about um de-identifying that data so that

the researcher has no way of knowing who that person was that did that embarrassing thing

should protect me as the person who that data is about so the first step of de-identification

seems to be the value is of protection of our participants getting that issue of harm that if you could be harmed

by violations of your privacy or confidentiality if somebody might learn something about

you that would hurt you in some way if that data is de-identified that should protect me

okay the second thing which is a little more interesting and a little more sort of it raises more

questions we'll get to at the end is whether when you've de-identified my data and you've gone ahead and used

that data towards research there's something about that de-identify the idea identification process that

sort of eliminates the need for me to say yes to my using that to the researcher

using that data as long as it was data that still had my name on it you might think my autonomy and need for

informed consent should control the use of that data just like i control

the use of my body for research just as I would only I would need to give informed consent

in order to be involved in a research project so identifiable information unless there's

certain exemptions which I'll talk about in a minute would be a case where i would need to give my consent using my data

once you've de-identified that you've created a sort of break between the individual and the data that

I would say is lurking there in the thoughts of how we do this regulation that once it's

been de-identified now we don't need to go back and ask the individual because it's no longer their

data it's not like a part of them it's been split from them again i'm happy to talk about in the

questions and answers and i've got another section here where we'll talk about that from a different angle but i've always been fascinated by that

side of it the idea that when you de-identify some data are you sort of separated from the person where it came from

uh i won't talk about these two aspects of this

so now we've considered de-identified data i feel in the sake of completeness i should

mention that if you don't de-identify data if it is private and identifiable then

there are ways you can use that data that are not subjected to the most rigorous

protections of human subjects research this is exempt research so if you're using data

in certain ways you can get an exemption from IRB review and this is really the big things that

the revised common rule changed and to put a point on it the revised

common rule loosened up the ways that you can use personally

identifiable information in certain contexts let me talk about those contexts

and this will again get at the thoughts of the relation between the person and the data and what sort of harms and

dangers we have to think about when we use data that comes from people for an exemption

there are two key aspects you have to think about and I'm again once we post this video I'm happy to post some reading that's very good

about this as well um the first question is how it was collected and the second is what its

characteristics are i'm going to separate those here so the method of collection is specified

in the in the rules about exemption there are three categories of exemption where i'd like to focus on

uh category two is educational tests surveys interviews and observations of public behavior

this is data that's been collected through one of those processes by the way in the change of the common rule they allow now auditory and video

recording um during those sort of assessments and that can be again under an exam type of research category three is

benign behavioral interventions uh with collecting information from adult subjects and benign behavioral

interventions are low risk and low invasiveness non-invasive and category four are cases where the

method of collection is just that the data is there it's collected from a data set not from an intervention or an

interaction with a person so that's the key these three categories are types of

exempt research and then if you are doing those types of research and you are using information from you know to

collect it in those ways now you have to look at what kind of information it is if the information is

de-identified then you will qualify for the exemption and now we know how to de-identify information

and this is the major change in the common rule if the information is identifiable

you can qualify for an exemption of the research from human subjects protections as long

as two things one or two things apply first

a disclosure of the participants responses outside of the research would not reasonably place the subjects at risk of

criminal or civil liability or be damaging to their financial standing employability educational advancement or reputation

again it's low risk data the data which you're collecting which is again identifiable the risk of it being

disclosed eyes low or the IRB can do a limited review showing

that that data will be protected very well with protection of privacy and maintaining of confidentiality

if you can assert either of those things it's an ore between them either of them or perhaps both either of them you can

get an exemption under categories two or three now category four is the one where you're not collecting data from people

at all you're collecting it from a data set or you're using

some biospecimens so this applies to secondary research on identifiable private information or

identifiable biospecimens and here you can qualify for the exemption if again the information is

de-identified or they're publicly available so they're not private if you think about that

not private anymore you could say or the analysis is on behalf of a federal agency or department that's governed by

federal law it doesn't apply to most of the research we see here or

the information is phi protected health information and its uses regulated by hipaa

as assumedly a research activity so here you go you can you're using a secondary use of data

and as long as it's phi and you're following the rules of hipaa you can get the exemption what are the

rules of HIPAA that apply here this is my last sort of big content big content slide and again we're happy

to go into greater depth both in questions and answers and if you'd like to talk to us if you'd like to call us up for a consult uh the key

aspects of hipaa that apply here um are that if it's personal identifiable information you want to use it for

research you can get an authorization from the individuals where it came from or you can ask for a waiver of

authorization you can ask to use that data without going back to the people who gave it and that would be if it satisfies three

requirements um again it must involve no more than minimal risk to privacy

and there must be an adequate plan to protect the identifiable the identifiers and they must be

destroyed as soon as possible they must not be re reused or disclosed

to other entities so if they're very strict plan for protecting that data given its personal

identifiable it's being used for research and second you must be able to claim and convince the IRB

um or privacy board that the research cannot be carried out um in a reasonable way practicably can't

be um you know done in a you know this this word practicable is an

interesting word it was coined in the research ethics world in the regular regulatory world

the idea is that maybe it's sort of theoretically possible but it's not actually possible to do the research

without access to that protected health information and it's again impracticable it's not

practicable to go back to each of the people and get the and get their authorization so it's impossible to do it

so if you can claim that it's key to the research and that it's really basically impossible to go back

and get consent um and you will take very good care of that of that research of that data and

really minimize the risk to participants uh following those rules then you can get a waiver based on

that so in all these ways these three categories for

exempt research you can use identifiable information and still be exempt from irb review

uh so to conclude i want to i want to um think again about the ethics side of this why

do we talked a little bit about why the use of de-identified information would be considered to be

not even human subjects research and therefore not protected by the usual

codes of research ethics and regulation why now that we're considering identifiable

information and therefore clearly human subjects research why do we put this in a special category called an exemption

why do we say don't worry about this IRBs don't worry about this maybe research ethicists this is so safe

we're not going to apply the usual standards we'll even give waivers for informed consent like we talked about so

uh if you look again at these four principles the first one seems to be the one about

harm again these principles we've seen even though the data is identifiable

the risk from using that data in the research is low and it's being it's being minimized

in the ways we've described of course like in categories two and three educational tests or these benign

behavioral interventions there's low risk from the interaction there's low risk and identified data

this is not where we need to focus our principles of ethics or our our efforts of our regulators and

that's actually one of the big changes of the new common rules that they've expanded this idea of exempt research

saying let's really focus on research where there's a significant risk to participants

um and that's why now identifiable data can be used in these ways

i was going to say this last little bullet point here is that when we imagine that de-identifying data

would completely protect the participants from any sort of negative interactions or negative effects of them having given that data to the research

project that is of course an oversimplification as some of you may know that de-identified data put in data sets

and analyzed in various clever ways can be re-identified even in the absence of that key

so basically following the rules of HIPAA and you produce a de-identified data set there are some

very clever people way more clever than I am and much more evil than I am who can basically take

that data and walk it back and identify the people involved this involves both genetic data um also

non-genetic data even things like your credit cards that can figure out who you are and i can refer you to TJ

Kasperbauer our postdoc who will refer you to a massive literature about this actually TJ and I just published a paper

in hastings center report a bioethics journal about thinking about this danger of re-identification so you might say that

the new version of the of the common rule that has i would say diluted the

distinction between de-identified and low-risk identified data is actually

is actually right that they've basically seen that this magic bullet of de-identifying data

is something it's helpful for reducing risk but it's not a be-all end-all in terms of reducing risk

but that does get me back to that other question which is do people have a right to control how their data is used in the future and

this has to do beyond harms that's to do with the idea that look a researcher the data you have there

about my life or my response to your survey or my health

is that about me and i should control it it's my data we speak that way this is my data

on the other side of course there are those who say no that data about a person is not yours

unless we choose to make it so and it's very different anyway than having a person participate in some

other direct way in research or some activity like receiving a medication um we do absolutely respect

people's choice about being involved say physically in research but do we need to also respect their desires about how their data is used

perhaps years later we could ask them we could ask them look do you want it to be used in the

Alzheimer's research can we use it in Alzheimer's research can we use it for heart disease can use it for mental health

and they might say yes you can use it for Alzheimer's you can use it for heart disease but you can use it for mental health research

now on the one hand again we'd say yes that's your prerogative it's your data and so let's ask you what you want

[Music] the other hand there's a part of me and i think probably part of some of the people listening to this talk

who would say that is really not a very useful way to ask people what they want

you're basically just asking for these peculiar priorities but they might prioritize certain diseases or whether

it's for the use of their data they can certainly choose which research projects to participate in physically

based on whatever priorities they have but directing a control and use of their data later

based on these preferences may not really rise to the level of something we wish

to ethically defend well that's a debate anyway of course here we're balancing

individual autonomy and individual wish for control versus the progress of science and

knowledge of medicine i mentioned earlier this idea that when you de-identify data you somehow rip it away from the individual it's no

longer their data even if you were in that first category saying it's my data once you de-identify

it you could say well it's not yours anymore it's nothing to do with you um i will tell you as an emphasis i can

think of the other side of that people say well sure you take away my identifiers it's still data that came from me and so

if i have a strong preference on how you'll use that i'd like to enforce that even if you de-identify it

i will say there's been a tremendous amount of empirical research on this asking people what they want when do

they want to be asked for approval for the use of their data and

when do they not generally i'll summarize you know decades of research and thousands of research subjects the

general take-home message on that is people do want to be asked they prefer to be asked waivers in some

ways are not what they want they want to be asked if you can use their data for research the other side of it though is they will

say yes they want to be asked so that they can say yes I'm getting better getting back to ideas

autonomy and freedom in in data which leads us to my final slide so you

can always um get a consultation on this this is a complex area um again i've been sort of updating this

talk over the years and I think I'm still rough around the edges I was getting emails from Amy Waltz yesterday when I

was coming up against my last questions so please feel free to call us um we have a phone number it's on our

webpage we have an email address you can just email me at phschwar@iu.edu you can always find that easily

and we're happy to work through your data ethics topics data ethics is one of the focus areas for the center for bioethics

okay so I'm gonna stop there I'm gonna minimize this enormous slide on my

screen um and I will I'll keep sharing my screen in case somebody wants

me to go back to a certain slide um but I'll open it up for questions and conversation

please [Music]

and uh there's no need to raise your hand you can just yell out your question or your concern or your uh

comment I see Matt is off mute he's come to the top of my screen

I'm thinking Matt and I were just talking about these topics two nights ago I haven't even read this talk yet matt

so forgive me if I've changed what I'm saying I just was in in in the research

on patients preferences about research what do you think they interpret

research as do they think it says academic centers doing research and publishing it for the good of everybody

do they think of it as a pharmaceutical company doing research and maybe profiting from it

do they think of it as something that could be sold exclusively for the use of one

particular company to charge them for it later yeah what do you think they think

yeah so data that's out there suggests that they are of course and again uh just for everybody to know

matt is a is a senior advisor at lilly so he asks this question being at one of these you know the

evil uh empire so uh he wants to know if the evil empire is seen as such by by research participants well look

actually so research participants and lay people do have a more positive view of

non-profits and academic centers collecting their data and using it then then lily uh they're easily instructed though

they're various studies actually where they've tried to instruct people and said you know lily's not you know lily might be a

private company but it's also an essential part of developing and bringing to market things like vaccines for the covenant virus

or medications and so people are open to that uh it's not the most hard and fast rule

between what an academic would do and what a lily would do there's something else i like talking about in this area which i

think about when people bring me questions about this and that is what about collecting data for other purposes what if

you know i wanted to collect samples or data from people and spray it on the walls an art project do people want how do they

feel about that is that something where they are more concerned about the possible use or they're less concerned and again is

that something where they would need to you know give their approval in this talk i've a few times said

that we wanted to think about what is research and what's not but of course any use of a person's data

even something as seemingly different as artistic you know is again something they might want to

control they might have various feelings about um a photograph of them to be used in an

art project in that case there are legal restrictions even ethical ones about you know using somebody's likeness

for a project and i think in the case of art I'm going to say and outside the case of research the idea of de-identifying

would be very powerful that while I can see why I would be against using my picture in an artwork if i didn't

approve of the artwork um if it was something that didn't look anything like me just because it came from me

would seem to be a strange way to control it I'm not sure i fully answered your question matt but um i think people do have relatively

um uh they have sometimes strong opinions but they're not very uh well worked out ones and so they

are changeable with some discussion and reflection as with most opinions

I jump in for two quick things and I apologize I miss most of your talk in terms of the art stuff a photographer

friend of mine said that it's perfectly okay to take a picture of anyone

in a public setting and use it for anything other than an ad for a particular product so if you

walk out of your office I could take your picture and plaster it on trump tower in Chicago yeah apparently

you'd have no recourse to that yeah Howard you missed the very not a problem do you ask I actually love

this question I did actually put in I actually have somebody know where to put this slide let me get it up here

um can I make it bigger I'll make it big again so a private information is special we

think i see um that basically it has to be in a private setting but you're right Howard you may know that actually before

i became a academic I actually worked at a newspaper for a couple years as a photographer um and I was it was my love my life was

was photography at that time and when as a photographer a newspaper you can take a picture of

people on the street um as an artistic photographer or even for a newspaper and you can publish that without

approval about the individual signing a form actually I'm not sure on the law on that maybe that's changed since when I was a

photographer but the idea is a public setting is very different than taking a picture inside a person's house or climbing on their

window and photographing inside their house even if you can view it from the street this is actually law that I think part

of the lawyers on the call can correct me on but right Howard it's very important whether something private or not now health information we think of as

private and that's often the case I will say something funny

in research I've done we've had the question of what is what is a health information versus

private non-health information so if i answer a survey standing on the street in front of an iu

health building that may be private information if my name's assigned to it but you could say it's not health information I'm just

standing in front of the IU   
Health building or if I'm standing in the lobby has that become health information

because I was in the lobby of IU health I was on my way to see my doctor when they asked me my survey and this

comes up these kind of lines and distinctions between private and public and between health and non-health information

really does come up Howard there are also so many paradoxes in this

area um and some

I'll give you a paradox from an animal survey from the animal protection stuff

I gave up doing anything with animals many years ago partly because I had to fill out about 20 pages

of information to use a right to sacrifice a rat to get a bit of tissue whereas i can

go into a hardware store and buy rat poison mouse you know mouse traps whatever i

don't fill out 20 pages of stuff there you know the people have data in these

large databases and if they're being done for research there are a lot of protections when google scrapes up all of your data

there are no protections at the moment or Facebook or anyone else so you know we deal with these things in

very um idiosyncratic ways and

one other thought just to get to get a reaction um there have been cases in research

where as best where people have drawn their consent so

narrowly that they essentially can say we're not

allowed to share our data with anyone else which in some cases seems like a

strategy to keep working on their data and use other people's data but not ever share their own

fascinating although Howard I'll do the second one first if they're not refusing to give you their data you should say well fine make a de-identified data set

and you can share that because that's what that's the beauty of de-identification right that basically argued they couldn't ah well I like

their clever then I would say stop sharing your data with them um uh but right i hear you're saying it

can be a as all games can be gamed all rules can be used for good or for bad

uh it sounds like that might be misused about the question about collecting data using for other purposes there's quite an enormous area of law here that i know

almost nothing about and many of the same ethical issues apply you know do you have control over your data

what is wrong with google taking your data and making billions off of it is it um is it all covered by that

disclaimer that that that disclosure that you didn't even read

lawyers will tell you whether it is they can tell you whether they legally can do it of course there's an area of great

change in the law both in Europe now in California different places now developing law to change the rules for

who can you use your data um and of course lurking behind all that debate now i think is

ethical questions about whose data is it um do you know what you're doing when you give it away

is it a fair trade as with all property um in our society and so i do think this

is a little corner of it the corner is my head around is only within research be right Howard there's about as many

different questions about the use of data and ethics as there are areas where it's used and as i discussed at the beginning

as you didn't see um I mentioned that data is everywhere it's now everywhere other questions or

thoughts questions of clarification Howard has some very high-end questions there

and matt did too anybody with a basic question you'd like other people to hear if i was unclear about something

or i left something ambiguous please um let me know now so we can get this on the recording for those of you watching this later

I had not a basic question I'm sorry you can't see me my internet keeps coming in and out but the

when you mentioned that you didn't want to ethically defend the a la carte notion

of kind of um of values i was wondering if you want to

talk a little bit about that specifically what I'm interested in is why specifically you said ethically

about that obviously there's a lot of practical issues uh that could result in like a stifling of science because of

the burdens that that would impose but I'm wondering if like really specifically the ethics thing you could

talk about yeah do you mean this slide colin the last slide here yes

one of my favorite slides uh and those who hang around at the center for buyers heard me go into this too much and Jane who i see is on the call also i

think she and i spoke about five years ago whatever it was Jane that summer when you came on we're doing some work on

data and privacy and i just think it's a fascinating thought of um

you know the control of your data you know your data is a fascinating phrase that

it packs in so much you know that that it's mine like my arm is my arm like my house that

i bought is my house you know the word yours stretched the breaking point when using those three

those three settings they did possession um there's relation between them they're all maybe integrally related to me my house

where i live my arm which is part of my body my data which is about me um

i i just think we need to unpack in the ethics world the notions of the relation between the

person and the data again which I'm going to avoid the word your the data which is collected

from or about that person right my genetic code gee it's mine like my arm is however i

never would have known my code if there wasn't this incredible area of genetic science that can take a blood sample and

sequence my data ah but it's my data because you got it from my blood vessel i don't know you know

um there may be other things that my participation allows which are not mine at all

it's not my football game just because I played in it so again I'm sort of throwing out some

some just a lot of un connected thoughts there perhaps to some ears

but i'm trying to say is that these questions about what is yours about your data what should matter to

you about your data should it matter to you so one of the thought experiments some people raised is wouldn't it be a violation of ethics if

i'm a person who's against cloning but then you went and cloned me from some leftover blood sample

i can sort of see that one but what if you actually took data from my health record and figured out how to do cloning

something i don't approve of should i have controlled my health record and been able to say

yes you can do cloning research on my health record or no i don't want you doing cloning research on my health record

when i say this thing about peculiar priorities which is just to say your individual priorities i guess i

guess what i'm trying to do is to undermine or show the other side of the idea that individuals

should be asked to make choices about how their data is used or that they should be seen as having

the right to have to have control over that these are great open questions in the world of

ethics very much of the day because we now have these tools and ways of doing things with data

that we just could not have imagined about oh i don't know two years ago or certainly not 10 years ago Colin I'm

sure that totally avoided your question but you're welcome to ask another one unless somebody else has one

no actually that did answer it I thought in a way that I wasn't expecting so I thought that was great thank you

but you're right con the simplest way to consider what I was saying is as a as a as a question of balancing ethics

between the good of advancing science and the good of letting people control how their data

is used so you could say well that's the question peter you're not questioning whether it's good to control your data you're just questioning how to trade off

that good against the good of making advances in science but you're right con i'm going farther I'm asking whether it's even a good

to give people control over quote-unquote their data so hey um peter it's Jane

hi I have a question that's kind well or an unformulated thought that's kind of

in conversation with what colin was talking about which is that um one of one of the things that

will make um research human subjects research is whether there's interaction or

intervention or some kind of engagement with the quote unquote patient

um that implicates their care and as we get into the world of really

really big data and the way that like say pharmaceutical companies are interacting simultaneously

with both Indiana university and Facebook and gathering large amounts of data that

they then triangulate for purposes not just of understanding how to sort of like market to

Indianapolis but then to understand how to market to one specific individual like via Facebook or via this

the automatic search you know filling on google and then to sort of direct your

behavior in a very specific way right and so I think

one of the things that we're on the cusp of or maybe that we're in the

you know we should start to be thinking about is how this traditional notion of an interaction with a patient or an

interaction with a research subject is changing [Music] as companies are getting really

smart about targeted advertising and targeted um sort of engagement with

very small individuals using just kind of algorithms and things like that so they pull your data and then they use it to talk

specifically to you even though they may not even know they might i mean like the connection is

sort of still broken perhaps a little bit but maybe not enough for increasingly not enough for me to even

be very comfortable with it and I'm not a privacy hawk you know I mean like I kind of don't care I know you're not yeah I know Jane um

I got nothing to say about that it's complex when you're asking I can ask you to repeat it you could but I could just say this

question about how you're using my data to communicate with me and maybe to target a message to me

of course goes way beyond medical research that's the standard done on

Facebook right they want to know how to how to sell me something by looking at what I'm looking at or google all this all this part of the world I

don't fully understand um but you're right you could say I'm

taking a question this way jane my entire talk and the ethics of data

and human subjects research has been about the front end how you're collecting that data how

you're holding it but you're raising a point about how it's then used back on you

maybe even directly back on you the only thing i'll say about that is if it's de-identified

and not re-identified but it's truly de-identified then i guess that block back to you has been blocked right so if

google collects information about what i like and then wants to sell me a car based on it they can't do that if they've lost

the connection to me they might draw general principles about you know middle-aged men living in

Indianapolis and then sell a car to me through that but that's at least through this you know more general group that I'm in

rather than really me so you can see the power of de-identification even in the case you're raising jane where the deputy identification block is

now actually cutting that link where the path back to where the data came from

has been blocked um but again I'm talking about this the research area where i guess

in general in research we're not planning to go back to the individual who gave us the data in order to treat

them but Howard will not give me an example where we are so I guess in my

top my head i can't think of a case like that i suspect we could think of one but jane your point's very well

taken i will say jane's also been uh working on a project we've been looking at about the use of

um of uh well we think a lot I would say just through the center and both will work with jane and others

about the use of clinical material in later research that draws on all these topics what do you do with

with data that comes out of clinical research

lab you've it's time for one more question for those who have been holding back until now

you can still vote guys I mean you can still ask a question

I'm gonna throw out one last comment then just to really rock the boat here you may remember i know i mentioned here

this thing about how you can get a waiver for secondary use actually i don't know if we have any um

human subjects regulatory people there but um where's the regulatory right category

four here it is so you get a waiver if you really can't go back i also i often wonder

maybe this is gaming the system kind of like what Howard was talking a different kind of gaming the system Howard was mentioning

um which is this question of whether you could collect data uh for the purposes of clinical work

knowing you'll later use it for research and knowing that you at that point will be able to apply for the waiver

um if you suspect you will later use it for research i think the waiver probably is a little bit of a

of a of a misuse I would say and unfortunately it's actually a very hot topic right now in

research because we have clinical work where we know there'll be later research we don't

consent people for the use of their data for research knowing we can come back later and ask for the waiver

and i'm not sure that's really the best way to do things if you're serious about giving people the choice

and basically the option to avoid needing to ask for waivers that's a cutting-edge issue in research

ethics which means I've really stepped beyond the treats talk once again i'll take you to our last slide which is if

everything I've said made no sense or if just it made some sense and you've got another question please contact us

I'm easy to find I'm Peter Schwartz at the IU Center for Bioethics um and this will be posted along with

some linked readings and resources uh soon on our on our treats

page and I will mention we have a new bioethics website that will have a very nice treats page for these talks

that's it for me I'm going to end this unless we've got any final truly final comments

okay great guys thanks so much for being here take care