[Cadence Cordell] All right, and we are recording.

Hello, everyone, in person, online, and in the future. My name is Cadence Cordell. I am a graduate assistant for Scholarly Communication and Publishing, and I am joined today by Sara Benson, our Copyright Librarian, who will - will be speaking online a bit later today, as well as Mary Ton, our Digital Humanities librarian, who is very kindly filling in for me today as our digital representative.

So, if you experience any technical or audio issues online here today, Mary is here to help address those issues. Please send her a message in the chat.

[Mary, voice slightly distorted] One more thing before we go forward, please hide the meeting controls and the participant - or, um, film Sara's view, um, or pin view of Sara on the video.

[Cadence] All right.

[Mary] Don't hide the video panel because we'll need Sara this time.

[Cadence] Okay.

[Mary] Go over to the video panel and if you hover over her - yep. And then minimize it so that it's only one person showing. There you go.

[Cadence] All right. And then hide floating meeting controls?

[Mary] Yep. Slideshow...?

[Cadence] Okay. And now we're ready to go! Today we'll be talking about copyright, text mining, and AI. Um, so we'll be going over the basics of how generative AI works and some of the, uh, limitations it has.

And then Sara will be talking about copyright considerations you need to consider when you are using AI.

So, our goals for today. I covered most of these, uh, just a moment ago. Um, but we'll also be talking about best practices for using AI.

You can also access today slides at go.illinois.edu/AICopyright, um, all one word.

If you are here in-person, you can also use the QR code, um, to get access to these slides.

The link here is case sensitive, so one more time, that's go.illinois.edu, slash capital "A", capital "I", capital "C" for "Copyright".

And I see some people typing, so I'm just going to give them a moment to type that in. All right, I think we're good.

So, how ChatGPT works.

Um, first, we should talk about, what is a GPT? If you didn't know, GPT is actually an acronym. ChatGPT - it's not a fancy part of ChatGPT's name, but the "G" stands for "generative". They are designed to generate new text. This is what a lot of the hype around generative AI has been. Previously, we haven't been able to create things designed to generate new texts like this.

The "P" in GPT stands for "pre-trained". They use machine learning to study large quantities of texts authored by humans in order to understand how language works in a process called language modeling. And I'll go over this in a bit more depth in just a moment.

And lastly, the "T" stands for "transformer". ChatGPT takes the model it has created and generates new text, one word at a time, adding an element of randomness to mimic human creativity.

Um, so to explain this in a bit more depth, let's talk about how an AI like ChatGPT reads a chili recipe.

So as we can see here, we have different quantities, we have lots of ingredients, and we have lots of spices.

So when ChatGPT gets a recipe like this, uh, it starts looking at all the different pieces of the recipe, in this case, all the words, and it starts looking at and processing all of the relationships between those words.

So in this case, it says, okay, I see the word "two" is followed by the word "teaspoons", is followed by the words "hot", "pepper", and "sauce". And I also see the word "one", followed by "teaspoon", followed by "cayenne".

This process of, uh, looking at all of the different relationships between words is so complex that it mimics the connections between neurons in the human brain.

So you may hear large language models like ChatGPT referred to as a "neural network". This is what it's talking about.

As it is looking at all of these relationships, it's also looking at the different contexts for all of these words. What words usually appear right before another word or right after another word?

Um, so for example, as it's looking at our chili recipe, ChatGPT is noticing that there are different kinds of pepper. In this instance, the word "bell" is usually before the word "pepper", and it also sees the word "cayenne" is usually before the word "pepper". It also notices that the word "spicy" is frequently used in relation to both "hot", "pepper", and "cayenne".

It is also paying attention to, uh, how the quantity of an ingredient depends on the type.

So it's noticing that the word "teaspoon" is frequently used before the word "cayenne" and "pepper". But it's not seeing the word "teaspoon" used before something like "bell pepper". It's only seeing a number.

So this is how ChatGPT learns about the multiple contexts a word can be used in.

So - for in this instance, ChatGPT is learning the difference between how to write something where you are cooking with peppers versus growing peppers versus the English colloquialism for peppering your conversation.

Um, I should note that when it's doing this, it doesn't actually know how to do any of these things. It doesn't know what cooking is, it doesn't know what growing is. It's just learning the different connections between words, so it understands them when it's time to create a text.

So what has ChatGPT read to start creating all of these language models?

Now, this information is proprietary, but we have a pretty good sense of what it has most likely read.

We do know that it's most likely read a lot of, uh, Wikipedia content. It has also read a lot of content on Archive of Our Own, which, if you're not familiar, is a fanfiction website that in particular contains a lot of erotic content. We also know that it has most likely read a lot of Reddit posts.

And I'm going over this not to pass judgment on any of these sources, but to let you know that ChatGPT and other generative AI programs have led - read a wide variety of content, mostly free on the Internet, of varying types of quality.

So getting back to our chili recipe, now that it has a language model to work off of, ChatGPT can start creating its own text, one word at a time, adding in some randomness to create new texts.

So in this case, I could ask ChatGPT to write a vegetarian version of boilermaker chili.

Um, I should note here that one of the key ingredients of boilermaker chili is beer, more on that in a moment.

Um, but ChatGPT gets this question and it says, hm, okay, what do I need to add to create this recipe?

It knows that chili usually includes spices near the beginning of the recipe, and it knows most recipes will start with a number or a quantity of some sort. So this case, it starts with "one half".

It then asks, based on my language model, what are the most likely words to appear after this one? And in this case, it decides to choose the words "teaspoon", "cayenne", and then "pepper".

It then asks itself, based on what I know about cayenne and reviews of spicy foods, what do I need to add? And so it might add a phrase along the lines of "adjust to taste".

The next part of this process is, uh, refining its results.

Um, so you can decide that you don't like the result that ChatGPT has given you, so you can help inform the model to improve its results.

Um, so you can give the output ChatGPT gives you a thumbs up or thumbs down to rate its response, and then you can ask it to modify its output.

So in this case, the recipe for boilermaker chili that ChatGPT produces creates a recipe with lots of beans, and no beer, the one key ingredient for boilermaker chili.

So we can rate this recipe a thumbs down, and then ask it to modify the recipe to include beer and exclude garlic.

So this process of, uh, taking human input to refine its results is known as supervised learning when it comes to generative AI.

All right. Now that we've talked a bit about how generative AI like ChatGPT works, I'm going to do a quick overview of its limitations and biases, and then Sara will talk a bit about our copyright considerations.

So, the limitations I'm going to talk about here are accuracy, limited knowledge, bias, environmental impact, and data privacy.

So first, accuracy and hallucinations.

I like to start with this example by economics professor David Smerdon, who asked ChatGPT to, uh - who asked ChatGPT what the most cited economics paper of all time was. And ChatGPT's response was a theoric - "A Theory of Economic History" by Douglass North and Robert Thomas.

And this is really tricky because this sounds like a really convincing response. The title includes key terms from economics. Both of these authors are real scholars who have collaboratively authored an economic history, but the paper itself does not exist.

And that's getting back to that transformer aspect of ChatGPT. It's not designed to fact check itself or to give you accurate results. It's designed to produce something that sounds plausible.

And this is actually something that you can test for yourself, and I recommend you test for yourself, um, using a prompt like this one.

Um, so depending on your area of expertise, you can ask ChatGPT or another generative AI to, uh, list several journal articles in your area of expertise.

So this is based on what I have experience in. I asked ChatGPT to list five journal articles about female missionaries in the late 19th century, and to include DOIs for each article.

Asking it to include DOIs, um, or something similar like an ISBN, is the really key part here because it's going to be the easiest way to fact check what it gives you.

Um, and very similarly, none of the articles ChatGPT gave me were real articles despite them sounding really convincing. Many of the DOIs were, in fact, real DOIs. Um, they were just

random DOIs that it had in its training set. So this is a really easy way to test that out, um, if you're so inclined.

Before I get to limited knowledge, I'm going to take a quick water break.

[short pause for water break]

Important to hydrate. Alright. Limited knowledge.

So, it takes a lot of computational power for a generative AI like ChatGPT to create new language models. So it doesn't do it every time it generates a new response.

It takes in new information every so often, and that also allows it to generate faster responses for you. However, it means that it hasn't, uh, intaken new information at a certain date.

So for example, this was a prompt I ran a couple months ago, back when the Olympics were just wrapping up. If you are not familiar, Raygun is a breakdancer who performed at the Paris Olympics this past summer, who...had a very memorable performance, is one way of putting it, and it got a lot of attention online. It went viral, but ChatGPT doesn't know about it.

So this prompt took a bit of trial and error, just trying to get ChatGPT to tell me, "I don't know about this". Um, but eventually, I asked ChatGPT to summarize the controversy surrounding Raygun's performance at the 2024 Olympics, and it told me, I don't have specific information on a controversy surrounding Raygun's performance because my knowledge only extends up to August 2024.

So make sure to keep this in mind. If you are, uh, using the free chat version of ChatGPT, it's not going to give you up-to-date responses, and it is more likely to hallucinate something.

I should note here that some generative AI companies are trying to address this issue, and they're incorporating realtime searches into premium versions of, uh, their algorithms to address this issue.

However, this doesn't fix its issue with hallucinations. It just means ChatGPT is less likely to give you out-of-date information.

Next, bias and stereotypes.

This is a really well-documented issue with all kinds of generative AI, from, uh, language generators to image generators, um, but they are very susceptible to recognizing patterns in their sources, in particular, stereotypes, bias, anything along those lines, and then reproducing them in the material it generates.

Um, this seems to be something inherent to these kinds of software and these algorithms. Um, AI companies are working to address it.

But please keep in mind that if you are asking ChatGPT to give you information, particularly anything about a minority population or group of people, please keep in mind that its answer is very likely biased by its source material.

Next, AI's environmental impact. This is being discussed, um, more and more. It's still not frequently discussed in these conversations about AI, but generative AI requires massive amounts of power and water to run. Um, and in some places, power systems are struggling to keep up with this demand.

A search or a query with an AI language model is estimated to cost up to ten times as much as a traditional search in terms of its power demands.

Um, and many AI companies are looking to invest in data centers, power resources, and the like to keep their, um...to keep their engines up and running.

So again, this is something to keep in mind. It's not something that we can necessarily address here.

Last but not least, before we talk a bit more about copyright, AI and data privacy.

Generative AI tools can retain the information you give them to help improve the transformer, and data retention policies are not always clearly articulated.

It can be a violation of ethics to share sensitive information with these tools.

So if you're working with sensitive information of any kind, especially if you are in the social sciences, if you are working with assessment data like surveys, interviews, anything with participant information, you most likely should not use ChatGPT.

It could absolutely retain sensitive information about participant names, contact information, and the like.

Um, but this also applies to your personal use.

Um, for example, I know people like to use AI to analyze their resumes and find ways to improve them. If you do that, please make sure to remove your contact information, your name, your email, your address, from those resumes because ChatGPT and other generative AI programs can keep that information to train their models.

And I believe there have been instances where people's sensitive information have been produced in, uh...ChatGPTs, like, uh, results that it gives to people.

All right. I'm going to mute myself and I'm going to turn it over to Sara Benson, our copyright librarian, to talk about AI and copyright.

[Sara] All right. Hello, everybody. Um, hopefully you can hear me okay in the room.

I'm going to have you, Cadence, if you can change the slides for me, that would be wonderful. Um...

Okay. So what is copyright?

Copyright is all around us. I usually ask a quick question, a poll, how many of us own a copyright? It's a trick question. We all do. We all own multiple copyrights. Because it's really easy to get a copyright.

All you need is a minimally creative, um, or original work that is fixed in a tangible medium of expression.

And so, fixed just means you've taken a photograph. So all the photographs on your phone, you own the copyright on those.

If you wrote down something or you typed it into the computer, you also own the copyright on those things.

So copyright is very easy to get and it lasts a very long time. Today, the, um, bundle of rights that you get with copyright lasts the life of the author plus 70 years after death.

So, the spoiler alert is that most things on the Internet are copyrighted.

So you don't have to have, you know, a notice of copyright on a page, for instance, the copyright is automatic. You don't have to have registered the copyright with the Copyright Office. Again, it is automatic.

So that means that what - you know, most AI is scraping the Internet, right, because they need large, large amounts of data in order to, um, train these neural networks that Cadence, um, so well described.

So some key concepts for AI include fair use and transformative fair use, as well as authorship.

So, we were - we're going to talk about some of those things, because normally, um, AI producers, they don't have permission from every single author.

Let's say, if they're using Reddit, for instance, I mean Reddit, the authors are thousands and thousands and millions probably of people, right, who post on Reddit.

And so they're not going to go to every single person and say, "Hey, can we use your post on Reddit?" That would be a folly, right?

So then they assert, um, fair use or transformative use to try to justify what they've done. Okay, next slide.

So what is fair use? So fair use is not a defense to copyright, but rather it's a limitation on the rights of the copyright owner. The difference is important, because what it means is that you still can get sued for violating copyright, um, and assert fair use, but you would have to answer that lawsuit in court. Okay?

So it is important always, when asserting fair use, to do some sort of risk assessment. What is the risk of being sued for that?

Um, what is kind of important to note is that there have been cases about text and data mining, and the court has found in, say, the HathiTrust case, that text and data mining are quintessential fair uses.

So, in other words, if you're using the work, not for the manner of which it is intended, in that case, it was books, many, many, many books that were digitized. You're not reading the books, you're not using them for reading, but you're using them more for, um, text mining or just understanding, you know, what terms are in the book, um, then that is an okay use.

And again, it's a transformative use, and the court has, erm, defined transformative use as, uh, one that alters the original work with a new expression, meaning, or message.

Again, it's this idea that you're not using the work for the manner in which it is - was intended, which is to read a book or to read Reddit or to engage with Reddit, even.

But if you're doing it for, say, AI learning processes, then your use could be transformative in that you're using it to train AI. So, um, that would be the argument, that it is a transformative use.

This has yet to be decided though, and - by courts, and they are the ones who generally, um, decide fair use.

The - the fair use statute is in the Copyright Act, but then case by case, the courts have to decide whether a particular instance is a fair use or not. Okay, next slide.

So again, this example of data mining with HathiTrust, the HathiTrust library inputs entire digital copies of books.

These are physical books that our library also participates in this. We digitize the book, and then, um - for the copyrighted protected works, we cannot see the entire book, but we can search for how many times a specific term is used in that work.

So for instance, if you are looking for a book that discusses anaphylactic shock, you would search for that term and it would tell you, either it comes up zero times in this book, or maybe 20 times on page 99 and 30 times on page something else, right, and that would tell you if that book is relevant to what you were doing.

And the court found that this was quintessentially a transformative use because the purpose of the original books was for communication and the purpose of the HathiTrust term numbers was for research.

And then this was further extended in the Google Books case, where the court found that, not only was it a - a term number used or a term, um, in the page numbers, but a - a snippet of the work was included to show you the context that was going on when you looked for that particular term.

Again, you cannot read the whole book because there are just small amounts. But it was extended even further than in HathiTrust. Next slide.

So what is the difference though between HathiTrust and Google Books, for instance, and what the cases that are pending right now about generative AI.

So in HathiTrust and Google Books, the corpus that was trained was based on physical books, and physical books have no licensing, right?

I mean, you can put, like, a shrink wrap or a notice or like, something in writing on the - on the front page that says, this book is only for, you know, professors' use or what have you.

But those are not contracts because they're not signed by anyone. They're not agreed to.

And so the difference here is that on the web, we have terms of service and we have terms of use that are arguably enforceable contracts. And this is debatable.

Uh, courts kind of, um, have tried to decide if some of them are enforceable or not. Um, if they're in, you know, boilerplate - plate language that's really hidden on the back page of many, many websites, it may not be enforceable, but - there, there may be enforceable contracts, and - and contracts trump any kind of defense or limitation we have on copyright authors' rights.

So therefore, if a contract says, you may not exercise fair use or you cannot data mine this, then we may not be able to assert fair use if that contract is held to be enforceable. Next slide.

So there are many lawsuits pending at the time, as we speak.

One of them is Authors Guild versus OpenAI. And, um, the issue here is that not only do we have them scraping all of these websites that maybe is a violation of the terms of use, which would be a violation of contract. We also have them spitting out the data that they used.

And so they've copied a bunch of the stuff, um, which could be a violation of copyright and or contract law. But then we also have, um, the outputs. And if the output is too similar to the input, then they're going to say that there's an infringement of copyright.

So also, in this case, they've seen that the training sets that were used maybe pirated copies, right?

We know that textbooks live on the web all over the place, right, and different things that are in copyright that shouldn't be on the Internet often are. And if your ChatGPT is not, um, discerning whether this is a lawful copy or not, they could be copying that as well.

Um, and I did see a case where the New York Times was, um, claiming that their work was being copied and showing the New York Times article versus the trained, um, output, and it was almost identical.

And so unfortunately, if you train on too little data, then your outputs are going to be very similar, right, to the train data because you don't have enough corpus.

So ironically, what some copyright librarians have pointed out is that the more data you've used, even - even if it's sketchy in terms of legality, even if you don't know if it's a fair use, the less likely you are to infringe on the output. So that's - that's kind of an interesting conundrum. You're better off scraping a bunch of stuff than just taking a small corpus.

Um, the other thing that I've heard this defined as, is - and - and I kind of love this term is, the so-called Snoopy problem.

Because if you train on a corpus that has Snoopy, right, the character, the lovable Peanuts character. And you keep asking for an image of a particular beagle with a friend who's a bird, who lives on a dog house. I mean, at some point it's going to spit out Snoopy, because that's what you've described, and it's doing its job because that's what it should do. Unfortunately, that's also copyright infringement.

So, um, there are - there are ways even on a well-trained system to get it to spit out something that's copyright protected. Okay. Next slide.

Okay, so copyright summary. Um...Generally, you - you can likely make a fair use claim to train your AI models. Again, this has not been tested in court yet, so it's still early days.

Most folks agree that it's a pretty good copyright claim under transformative use, but again, we have contractual issues there.

Um, now, you may be able to claim copyright on your prompts that you put into, say, the - the Midjourney or ChatGPT, if they're sufficiently creative.

Now, if I just say, give me a picture of an apple, that's not going to be it. But what if I say, give me a picture of an apple covered in holes that look like dinosaur eggs with little gems in the center and spitting out fire, or something - I mean, just something super weird, right?

Maybe I could claim - I don't know why I want to claim that, but maybe I could.

Um, you also might be able to claim copyright on modifications that you make.

So there's a very famous case, Zarya of the Dawn, where a graphic designer put together, uh - a graphic novel using AI-based images.

But they did enough to put them together in an arrangement and changed them enough and added words and added, you know, a story, that they said, okay, you can't own these images alone, but you can maybe own what you've modified them or the arrangement of them.

It's a lot like the way the Copyright Office treats data. You cannot copyright pure facts, but the arrangement of facts, you can have a small copyright over.

Um, so on the negative side, um, there may be contractual issues involved in scraping the web, and that is what a lot of these cases are about.

Um, and, you know, we have a lot of folks who want to use say, library databases to train AI. And the problem is that many of our library databases uh, do not allow licensing for AI use.

And so it's a question that we need to then bring to the publishers and say, can we - can we use this?

We may have, um, kept our fair use rights. And if they haven't then added any other restrictions, it may be an okay thing to do, but they also have restrictions on how many articles you can download at a given time, et cetera, et cetera, and how many you can keep and how long you can keep them.

So there are so many ways that the - the licensing can really interfere with the process.

Um, generally, you cannot claim copyright on text and images that you generated with AI. Again, you may be able to claim modifications if they're significant and also, um, arrangement compilation.

And, um, yeah. When you create something that is too close to the original, AKA, the Snoopy Problem, you're going to run into copyright infringement issues because if the data that it spits out is just too close to the original, then, um, that's the test for infringement.

So that should be, um, a good summary. So we're back to you Cadence.

[Mary Ton] Um, this is where Mary jumps in.

[Sara] Oh, we're back - we're into Mary's part!

[Mary] And I switched chat moderation to Cadence. Okay. So a few best practices with these copyright concepts in mind.

Uh, we like to emphasize that AI, at its best, improves accessibility. It lowers barriers by making it easier to iterate and, um, and to work through a bunch of ideas very quickly.

It also promotes - AI can be used to promote critical thinking and creativity, that we emphasize the human in the loop as part of the process.

So the acronym that we are using to summarize best practices, we encourage people to use AI with POWER.

And those letters stand for "P" for "pause", "O" for "orchestrate", "W" for "write", "E" for "engage" and "R" for "review".

So a - a closer look at each of these things.

So when we're talking about pausing, we encourage people to pause before they use AI to check any policies that may be in place in your department or your profession.

We also encourage you to evaluate privacy of your content. So, are you working with sensitive data? If so, most AI tools are not rated to protect - to offer the data privacy that you need in order to work with - with sensitive data. UIUC currently has a list of approved AI tools.

I think in light of copyright, some other things that you might want to consider as you're pausing before you use AI is, is this something - is this an academic article that I got from a library database?

If so, there might be some things in the terms of service that limit how you use that. So you might not be able to take the article that you got from interlibrary loan or from one of our databases and feed that into an AI tool to summarize.

So always double check the policies and the privacy.

You might also, especially if you're considering publishing your work and using an AI writing tool as part of your publication, check your journal's policies around AI uses and how you declare that. And if they allow it.

Orchestrate. So AI is one of many tools that we use as part of our research process. Those tools include things like Word to build our prose. It might include a - a finding aid in an archive.

We're constantly navigating between analog and digital tools when we do our research. And so we're encouraging researchers to orchestrate all of the tools that you use to be, um, thoughtful about your practice and to choose the right tool for the job.

Um, so this might mean not using an AI tool to look for scholarly sources because it has the tendency to hallucinate, or, uh, it might also be avoiding using an AI tool to summarize because in your discipline, it's expected that you're doing that work.

Um, and also, just be mindful to choose the best tool for the task. Some - some tools are better than others.

When we're thinking about orchestrate, we're also - especially if you're a visual artist, we're also thinking about AI as part of a broader suite of tools that you could use to create images.

So remember what Sara was saying about modification, you might be navigating between an image generation tool like Midjourney, and a tool like Adobe Photoshop to manipulate and change the image. The more that you change that image, the stronger your copyright claim becomes.

Third, write your documentation as you go, your future self will thank you.

So what we're seeing across publishers is that most publishers are requiring a - some sort of statement or documentation of AI usage.

In the social sciences and sciences, this includes a, um, section in your methods paper or an addendum to your article that provides these sorts of things.

So it includes the user, which tool you used, the date and time, because these tools evolve, the prompt that you used, a copy of the generated text, and sections of your writing that contain AI-generated text.

Presumably, this is to help publishers do two things. First, it's to help reviewers assess the quality of your work. So a reviewer might be looking more closely at AI-generated sections than the human-authored ones.

And second, this is also demonstrating the degree of human intervention. Remember, if it's something that's completely generated by AI, you may or may not be able to make a strong claim for copyright.

So, demonstrating and documenting your use and really showcasing your role in the process, is going to help make a stronger case that you are the copyright owner of the material.

And then finally, engage with the process. So this is related to showing the human in the loop, the human intervention in this part of the conversation, your control over the - the tool when you're generating text, and your modifications.

So, use interaction, iteration, and conversation to refine your own ideas. The more that you intervene, the stronger your copyright case becomes, or copyright claim becomes.

And then finally, review.

As Cadence was describing earlier, there are issues with AI-generated outputs. So be sure that you're reviewing the outputs for the audience.

Is this something that is appropriate for your audience? Does it meet audience expectations? Is this accurate? Do I have other sources that I can use to verify the citations and - and bibliography? Bias, what is this generated text excluding? And what might I need to do further research about?

And then finally, looking for articles and scholarly sources from our library resources.

Not, um...each tool is pulling from a very limited set of resources, and no one tool will provide you a complete view of all the resources that we have access to through our institution.

So, um, again, navigating different tools at different parts of the process to make sure that you are showcasing your role in the process and also mitigating any of the issues that come up with- when you are - that come with using AI.

So these are just a few of my favorite activities when I'm writing.

I like to write the first and last sentence of a paragraph and ask AI to generate the middle, and then will edit the suggestions.

I also tend to write bullet points of key evidence and ask AI to turn it into full sentences.

And finally, I - I will admit, I have - my former self was not as good at using a citation management tool. Please use Zotero and Endnote. It's much more sensible.

But when I've had legacy citations, AKA, the stuff I threw in my dissertation that's in MLA, and now I need it to be in APA. ChatGPT, Microsoft Copilot are pretty good at changing citations from one style to another.

Just make sure that you double check the DOIs because these chat bots are - will hallucinate a DOI that either doesn't exist or links to a completely unrelated article just because they recognize the pattern in citations that usually they are followed by a DOI.

Okay. Running through campus-wide resources before we turn this over to question and answer.

Institutional conversations. So our Generative AI Community of Expertise, now called the Solutions Hub, was charged by Provost Coleman and CIO Martin to aggregate capabilities and corresponding resources to empower responsible adoption of generative AI at the University of Illinois Urbana-Champaign.

So this means that our institutional momentum is devoted to encouraging and supporting thoughtful and ethical uses of AI. So this means that our access to AI tools and AI-powered tools is only going to increase.

Um, but as a humanist, who is skeptical about these technologies, I am deeply grateful to our colleague, Celenia Graves, who made sure that the humanities were well-represented in the center of the Solutions Hub working groups.

AI tool access. So we have Chat Illinois, which enables you to upload PDF, PowerPoint Word, Excel, and then interact with those documents through a GPT-powered chat.

It used to require an open API access key, which means a charge with it, but you can now request a free account by going to the website.

Microsoft Copilot offers a secure chat environment with real time Internet search. It's not HIPAA-compliant, but it is more data secure than the free version of ChatGPT.

It's available with a Microsoft login to faculty and most staff. It hasn't been rolled out across the university yet, but it is available to students through GitHub.

Finally, so we have campus resources. Again, um, the Generative AI Solutions Hub, formerly the Center of Expertise. CITL's Innovation Studio offers access to popular AI tools like the image generation tool Midjourney, and I think some subscription-based versions of ChatGPT. And of course, the Writer's Workshop. If you struggle with your writing and you want a human to help, they offer in-person and online consultations.

We've also created a generative AI LibGuide that talks about some of the major topics. Things - It summarizes what we talked about with legal issues today, identifies some of the ethical considerations, includes a little bit more for instructors, and also has contact information for library experts for you to ask questions of.

We also have the "AI in this Course" Canvas module. This was designed to help instructors communicate their expectations around the use of AI. So thinking about human authorship and that criteria as part of copyright and really emphasizing that in the ways that AI is being engaged with in a course.

More library workshops. We have...I think we have a few more AI workshops coming up this semester. If not, there will be more next semester.

Those are posted to our AI Savvy Researcher calendar, and this recording and many more are going to be available on the DH recordings on Media Space.

Uh - an old slide with apologies.

Okay. So I want to emphasize AI as a tool and not a solution. That this is something that we really want to focus on as a tool for human creativity and communication. Not necessarily something that can replace or just be used wholesale. You have to review, you have to engage. You have to be the driver in that process.

We are excited to field questions. So if you feel comfortable, we'll - we'll turn things over to Q&A in just a moment. But if you think of a question after the - the workshop is over, please feel free to contact either myself or Sarah Benson.

We really appreciate you joining us synchronously today and online in the future. And with that, I'm going to turn off the recording and open the floor to the questions. Thank you!