

Using AI for Support of Thematic Analysis

Welcome, everyone.

Uh, my name is Dr. Smith. I have an arsell.

I have the opportunity to serve as one of the leads of the research methodology group, and we are so excited to partner with the AI Research Group to conduct this webinar, uh, using AI for support of thematic analysis.

We have an amazing, uh, scholar, researcher, colleague, and collaborator, Dr.

Jim Rice, to share his knowledge related to this topic.

So, uh, without further ado, I'm gonna turn it over to Dr. Rice.

Thank you so much for being here.

Thank you so much, Dr. Smith.

Um, as, uh, Dr. Smith said, my name is, uh, Dr. Jim Rice.

I'm a, uh, faculty member at the University of Phoenix here.

I, um, a a chair in the College of Doctoral Studies.

I'm also a senior research fellow of the C-E-I-T-R, and a member of the ai, uh, research group.

So, uh, this is a topic that is near and dear to my heart.

I've been, uh, working with AI since I started my writing my first lisp code in 1985.

So, I've been around the AI platforms for a while.

Um, but it's, uh, a rapidly evolving, uh, landscape of tools and capabilities out there.

And, uh, it's always great to, uh, try and keep up with what, uh, the latest and greatest is, particularly within research.

Um, uh, this is a great tool.

Uh, AI in general is a great tool for us to help, uh, accelerate our research efforts if we use them correctly.

And that's, that's part of what we're gonna talk a little bit about tonight.

So, uh, in qualitative research, I think we all know that, uh, thematic research or thematic analysis is a, a common tool that we use in many of the, uh, dissertation designs for qualitative research.

And, um, most of us are used to either, um, doing that thematic analysis by hand, doing the coding by hand, the thematic analysis by hand, or using some of the traditional tools like, uh, Envivo, uh, for doing that work.

AI is presented a new capability or a new set of tools that we can consider, um, and have actually been, um, over the last year or two.

There have been quite a few papers that have been written on how to leverage some of the ais to do, uh, thematic research, and we'll

talk about that in a little bit.

Um, but I wanted to start out by just, uh, letting people know the concept of artificial intelligence has been around for a while.

Um, it actually goes back to, uh, the four hundreds, but the, the most, uh, extensive writings were done by descartes in, uh, 1641 or in 1642, uh, when he really started a posit.

What is, what is intelligence? And is it possible to create artificially created intelligence? Now, descartes is, um, position at the time, and I think would still be the position.

It's actually my position as well, um, that artificial intelligence is not equivalent to human intelligence because, uh, artificially created intelligence isn't a result of agency, uh, from the machine or consciousness of the machine.

Uh, but it does provide a, a fascinating tool for us to use.

And it's a capability that, as it becomes more mature in the technology space, uh, continues to be, uh, an area of debate among philosophers.

And folks that know me, know that I've got a, um, a master's degree in, uh, pastoral ministry and theology from the seminary.

So I love some of the philosophical and, uh, uh, metaphysical conversations as we talk about, uh, ai.

But that's not our purpose for tonight.

Um, our purpose for tonight is really to talk about how to use AI as a tool, as a collaborator when we're doing thematic analysis.

Now, just, uh, by way of, of transparency here, the examples that I've got in this presentation today, were built using chat, GPT and Gemini.

These are just two different tools.

Um, I've used these processes that I'm gonna describe tonight for thematic analysis on, uh, chat, GPT, Claude, copilot, Gemini, and Perplexity.

And there's even some new tools that are coming out that are specifically designed for thematic analysis.

Um, but, um, just wanted to, um, let people know the University of Phoenix doesn't specifically recommend or endorse any specific AI or ML tools or platforms.

Um, but that doesn't mean we can't use them.

Uh, we can certainly, uh, use these tools to help us in our research efforts.

Um, just wanted to set the stage a little bit.

You know, when I talk about using AI for doing, um, uh, thematic analysis, I'm talking about using AI as a collaborator.

I'm not using it as an alternative.

I'm not using it as an excuse for not understanding how to do thematic analysis.

Uh, it is just that it's a tool, it's that we can use to, to collaborate with, but it's not an author.

It is not a substitute for understanding how to do thematic analysis.

And I think that'll become more, uh, clear, uh, why I say that as we go through these, uh, uh, through this presentation tonight, it's also very important, um, if you are using these tools that you figure out how to avoid the bias that can creep into these AI ML tools.

Just like a, a human researcher, um, AI tools all have biases that are based on the body of information that they, uh, use to build their models.

And one of the things that's important for us as researchers to do is to try and negate that, uh, bias that exists through good discipline, just as we do when we are doing analysis of our own data, um, and our own collections, uh, in our own way.

We have to figure out how to negate our own biases.

And so when we're using ai, we have to figure out how to, uh, negate our biases as well.

And we'll talk about a, a technique for how to do that.

Um, but I also wanted to, as we get started here, just, um, put a, a note out there that it's important for us all as we're using these AI tools to understand what our ethical boundaries and guideposts are.

Every different publisher, um, whether it's Elsevier, Elsevier, or Sage or Emerald, all have policies and standards, um, uh, about the ethical use of generative AI and AI ML platforms, what you can do with them and what you can't do with them.

And these policies not only talk about the AI Is, but also the AI enabled, um, platforms like Grammarly, et cetera.

And then the publishers aren't unique in that the University of Phoenix also has policies around those.

So I encourage everybody to go look at the policies, um, for the university, but also if you're working on articles for publication, make sure that you go look at the, uh, policies and the guidelines for the author, for the, uh, publishers that you're in, the journals that you're looking to publish in before you put a lot of effort into a, uh, a research project, only to have it rejected because of the tools you chose us.

Um, let's talk a little bit about what an AI is.

Um, an AI ML platform is a tool that does a really, really good job of taking in volumes of data and looking for patterns in that data.

So even when you're having a, a chat or a conversation or asking an AI to generate pictures, it's drawing, drawing upon its repository of information to make a best guess at how it should respond.

In many ways, it looks and behaves a lot like, uh, any statistics model you may have.

And I put these two graphics up here, uh, the, the scatter plots up here for a reason.

Um, sometimes the data that we're presenting to the AI very clearly has, um, an outcome or a pattern within it that can be modeled and

responded to, and andis that, um, are prevent or presenting that response, um, are doing their best to predict what the, the, uh, that answer should be.

Sometimes the data is very scattered and not very, um, not very well formulated.

And the AI will still give you an answer.

Um, just like a statistics, uh, a correlation coefficient would, would give you an answer, but its confidence in that answers can be relatively low.

The reason I put this up here is, this is an important thing to understand whenever you're using an AI ml, because whenever you're using it to identify patterns in the data, it's important to understand how the AI is determining its confidence in the data.

And, uh, simply is, most of the time, it's just a matter of asking it what his confidence is.

We'll talk about that a little bit more in a minute, but you're gonna hear me talk about confidence levels within ai, and people are gonna think they're in their statistics class, and it really isn't.

You're, this is, this is what the AI does.

It looks for patterns in the data and does a prediction on what the response should be.

If it's confidence is low, it's still gonna give you an answer.

It's what we sometimes refer to as the imagination of an, of the ai.

It'll just make up an answer if it doesn't know the, the correct pattern.

Um, so like I said, it's AI's tools.

Um, they are not an alternative to understanding how to do thematic analysis.

And that's our, our topic for tonight is thematic analysis.

We still have to understand how to do thematic analysis in order to have a, an intelligent conversation with the AI to get the, the types of results that we're looking for.

Thematic analysis has been pretty well studied.

Um, uh, and I think most people that have done thematic analysis have seen this circle many times.

This was this, this, um, order of analysis, this gathering of data, coding of the data, looking for emergent themes, defining the themes and reporting.

The results was proposed by Ron and Clark back in 2006.

But since then, there's been quite a bit of, uh, research, both on thematic analysis, how to create, um, flexibility in the use of thematic analysis for different research designs.

And also, as you look in the list of articles on the right there, and I, I'm, I'm not expecting you to read them tonight, you'll have this handout after the presentation you can use as a resource.

But you'll also see in here, we're starting to see, um, publications and articles specifically on how to use ai, mls and generative AI effectively to do the thematic analysis.

And the process I'm using tonight, uh, is really built, built based off of, uh, an article that was written, uh, and published last year, uh, by, uh, turbo, uh, on how to use chat DPT for thematic analysis.

The analysis, um, form that I'm using tonight is, is geared towards, uh, doing delphis.

But the, the basic concepts are the same for any thematic analysis.

So understand what a thematic analysis is, understand these steps, uh, because it's gonna be important, understanding why we're doing the analysis the way we are.

So what do we do first? Um, anytime you're doing a thematic analysis, the first steps are the same, regardless of whether you're doing, uh, doing your analysis by hand, you're doing it with envivo, you're doing it with an ai.

First thing you have to do is collect your data, uh, make sure you've got well-defined research questions that you've recruited your participants, you've collected the data, you've transcribed it, you've validated it for accuracy.

If you don't have clean data, um, you're not gonna get a clean analysis, whether you're doing it by hand or you're doing it by, um, with an AI assist.

Um, bad data will result in bad, bad results.

So make sure you've got quality data, you've got sufficient quantities of data to make sure that you can identify those patterns precisely.

Um, also, make sure, and even if you're using an ai, do not skip this step.

Uh, make sure you read through your data, read all of the data, all the transcripts, read it multiple times, jot some notes down, become familiar with it, because, um, as you'll see in some of the examples tonight, you may get results back from the AI that are low confidence or that you disagree with.

And that's okay because the AI is a collaborator, and it's good to have a collaborator, a second set of eyes, if you will, in the AI tool that are giving you things to look at and think about in the analysis.

And you may determine that it's something that you agree with, something that you don't agree with, but it's, it's a, uh, piece of feedback that you, you can use to enrich your research.

Um, this step is particularly important.

Um, it's, it's, uh, essentially if you are using NVivo, but it's even more important if you are using ai.

Make sure that you've got your data in a good digital form that you've masked the participant or business names, uh, for confidentiality, that you organize your data clearly so your participants are well identified, um, and that the data sources are, are aware.

As we look for patterns in the data, and we look for trends in the data, it's important to understand how many, uh, participants are

contributing to the emergence of those themes.

And unless our data is well organized, um, the, the tools we have won't necessarily know whether those insights are coming from one participant, two participants, or 20 participants.

So it's important that we make sure we've got things, um, clearly organized and well labeled.

This is true of any thematic analysis.

So this is a foundational step.

Um, and it's true no matter what to make sure we've got good quality data.

Now we start to sneak into, you know, what does it mean to organize that? And this is, these are just some screenshots from the data that I'm gonna use for the examples tonight.

Um, I've got four fictional, um, and they're just fictional, um, interviews that are put together, um, on the topic of, um, retail workers and their motivations.

Um, in my examples tonight, I've got four different interviewees.

Um, they're all, their interviews are all maintained in separate files.

All their names have been masked, uh, with, with labels.

Um, and so when I actually go to feed this to the ai, um, for analysis, it's gonna be very under.

It's gonna be very clear to the tool how many participants I have, and that they've been labeled consistently.

The other thing that's, um, I found important, and it took me a few, few missteps to figure out, I make sure that myself as the interviewer, or the interviewer in general, uh, that's operating off the instrument for asking the questions and interviewing the participant, is consistently labeled as the interviewer, um, so that the AI or in vivo can filter that out appropriately.

Um, this is how, this is the trick that I highly recommend to try and remove some of the bias that can creep into AI generated analysis.

Um, I started doing this, um, first with, uh, with chat GPT, and then with some of the later AI mls.

And I actually did this more recently with the new AI enabled InVivo.

'cause the AI enabled InVivo was introducing some potential, um, issues with bias, uh, in the way its ais were generated.

Um, so, um, wanted to make sure that as part of the data set we're loading into it, we're putting in our, our, um, conceptual framework theories into this.

So in my particular case, I'm putting an agency theory in here, a contingency theory and a human capital theory to evaluate my data for retail workers.

And if you look at this, this, this, um, the theoretical context looks just like, um, the context you would have in chapter one when you're talking about the theories and how to align those theories with, um, your research questions.

So, um, this is not something that is commonly described and when you talk about thematic analysis, but this is, um, a, a trick that I've been using to make sure I'm getting the similar outcomes from all the different AI platforms that I work with.

They're not always, they're not always identical, but I've removed the bias out of them.

And actually, they're, the findings and, and analysis wind up being very, very similar regardless of the platform I'm coming from.

So, be focused on your theories, make sure you've got a good, well-founded critical description, uh, good citations for those, um, and from quality subject matter experts, uh, in the field.

So how do you feed this into the ai? Well, um, it's actually fairly simple, and it's really based on just making sure your questions are answered or asked, um, appropriately.

Uh, most of the AI will just let you drag and drop all the files that you've accumulated.

So you can see here in my example, with chat GPT-4, oh, I just dropped my four interviews and my three theories, and I dropped them into the tool.

And then I asked the AI to, to analyze that data with a very specific question.

And this is where it's important to recognize how to, how we're doing thematic analysis.

The first step in thematic analysis is coding the data.

So what I asked the AI to do was look at my interview data, look at my three theories that I am using for coding, and I ask it to propose initial theories, uh, or, I'm sorry, propose initial codes, uh, for the data.

And this, this allows me to take a look at the coding, uh, that seems to be appropriate, um, in the patterns of the data for my interviews.

Now, this is because I've got a fairly small sample.

The coatings tend to be pretty broad, but, um, it, I think it's good for the, uh, purposes of the presentation.

And I, this question I can use with any set of data, you know, I can take this question and just copy and paste it with any collection of files, and I'll get the same kind of output.

What does that look like? So, um, in the case of chat, GPT, it gave me a set of six, I'm sorry, seven codes that I, you could apply in various places within the various transcripts.

And it gave me a definition of that code.

How is it using that code? And I identifying text that aligns with that code.

It even gives me examples from, uh, the transcripts of where it applied that code, or where it would recommend applying that code.

And it even tells me what theories, because I asked it to, and it gave me which theories, um, contributed to the identification of that code for that data.

So I've got seven, seven codes that, um, now can be used to take a look at the, uh, and, and provide a lens for looking at my interviews.

Now, there's only four interviews in this particular sample, so you don't see lots of, uh, uh, references.

In fact, the two references I gave were from the same interview, uh, which kind of makes that that code a little bit suspect for me when I see only one, uh, one participant in that code.

Um, but it'll give me, uh, in this particular case, it gave me this information for all 70 of these codes.

Here's the next step. Remember, whenever you're doing a, um, a thematic analysis, you do that initial coding, the next thing you do is go through and look to see how can I combine codes? What is the level of impact? How often did I use that code? So you can start to reduce the number of codes you're using or consolidate them so you can have a higher impact theme emerge from the data.

The way I do this with the AI is I ask it to first identify the confidence level of that, that code.

Remember, the AI will give me an answer regardless of what its confidence level is of that particular answer it gave me, or in this case, the theme.

So I asked it, what is the confidence level of your, of your response, and what, what this particular AI did that chat GPT-4 oh said.

It's got four of those codes that had a high confidence in it, had two that had a medium confidence and one that had a low confidence.

Um, so I'm going, Hmm, there's, there's some opportunity there for refining the coding that I want to use.

I then asked that to, uh, uh, take a look at what are, what would be the emergent themes if I asked that to only focus on the high confidence codes that more than half of my interviews contributed to.

Um, and so it narrowed it down to three different themes that emerged from those codes, uh, those high confidence codes.

And these are the, the three themes that emerged, uh, from that coding.

So you can see I've gotten with a fairly rich set of data, even though it's only four participants, they're 10 minute interviews, um, a fairly rich set of data.

I very quickly got to, uh, a set of, of emergent themes that are relevant to that particular data set that I can have a reasonably high confidence interval or very high confidence in.

I still, as a, as a researcher, I go back and I reread the data, uh, with each of these steps to make sure I'm concurring with what the AI is saying, or if there's something that's emerging that, um, I didn't think of.

It's, um, it's interesting to see where Theis will find a theme that I didn't, through my own lens, my own experiences didn't identify, but, uh, AI found it through these pattern recognition techniques.

Most of the time I agree with them, I'll, I'll concur, but occasionally I'll find one that I go, yeah, that's just that, it just, its imagination made

that up.

And I'll just discard that one.

Um, the human still rules in these.

Um, so in addition to identifying the theme, I ask you, go, go ahead and tell me what are the, you know, describe that theme.

Give me, uh, examples of quotes from the participants that support those themes.

And again, tell me what the theories are that support the emergence of that theme.

So that, as I'm going back and aligning my themes with my research questions, 'cause remember, every one of my theories is associated with a research question or two.

Um, I can then tie these themes back to the research questions that are the core of the research that we're doing.

Um, and then again, I ask it to summarize it so that, uh, I can see what those, uh, themes are the major findings were and what the theories were that contributed to that.

So you can see I've got contribution from all three of my theories.

I've got three emergent themes that, um, I can support with quotes from my, uh, from my data.

This gets me through, uh, chapter four pretty quickly, um, which is, which is fascinating.

You notice here, I didn't ask the AI to write anything for me.

What I asked it to do is to help me analyze the data, give me new insights into the data that I might have missed, um, by my limited experience or my, um, my personal understanding or, or biases related to the, um, population.

So this is helping me remove my biases, uh, relative to that population's also helping me explore additional insights that I might not have otherwise had.

The AI hasn't written anything for me here, but it's helped me with the analysis.

So when I go into the process of writing these, uh, writing up this, this findings, I've got a good basis and I've good, good references and citations from the, from the, the data that I can use to support what I'm, uh, finding.

I'm also willing to, and I've been playing around more with this, allowing the AI to help me identify some potential, um, recommendations for these findings.

Uh, and I do this specifically because I'm using the theories, um, to, uh, to support each of the recommendations that I'm doing.

So you can see here's a re set of recommendations for, um, the theme one.

These are specific actions that I've asked the retail manager to that, uh, that the retail manager may take based on the results.

And so it, it looked at the theories that I use agency theory and human capital theory in this particular case, and it made some

potential recommendations for actions that retail manager could take.

I very seldom take these literally, uh, but I like to take these and go back and look at the objectives that I had for my research.

And if you remember, the purpose and objective section of chapter one we're, we're I looking for particular actions that a, uh, a stakeholder or a practitioner may take, uh, based on the findings of the research? And I look to see if any of these, uh, if I missed anything that that makes sense, or more specifically, if any of the recommendations of the eyes, um, uh, proposing for any of my themes lines up well with me, the objectives.

I identified my proposal, uh, again, uh, a summary of those.

Now, um, you notice all those examples I had were for Jet GBT.

I just wanted to, um, show you something that's important to understand.

Here's doing exactly the same data, which exactly the same questions using Gemini, which is a different ai, and this, and Gemini happens to have a much higher bar, uh, for, uh, confidence intervals before it will propose an answer.

Gemini has a smaller data set and a higher confidence interval threshold.

So it's going to come up with different answers.

It's always fascinating to me to, to use Gemini or one of the other.

Um, I, I always do use more than one, uh, platform right now to do this.

Uh, set of size is like having two or three collaborators I'm working with on the analysis.

You notice in this particular case, it only came up with two high confidence codes, not four.

And again, that's for two reasons.

One, it's threshold, it's confidence interval threshold for providing responses higher and chat GPT.

But also I have a very small sample.

I've only got four participants in my sample here.

And so, getting a high confidence interval with a small sample population, Gemini's gonna be pickier about that than chat GPT would be.

Um, and so if I were to continue to sample at a fifth participant, a sixth participant, get 15 participants, what you would start to see is the findings from Gemini and chat GPT would start to look more and more alike, um, because the confidence interval will tighten up on that data set.

Um, so using more than one, uh, AI ml as a collaborator can give you some interesting insight, both on the data, the, the findings of the data, the patterns in the data, but also in your confidence in the, in the data.

When I asked, um, uh, the AI in this particular case for the themes, you notice, I ask it not to look for saturation.

I asked to look at, for, I asked it to look for, um, consensus or over 50% participation.

If I were to ask it for, um, saturation, both of these ais would probably have much lower confidence intervals in their findings.

So it depends on what you're using for, uh, your sampling technique and what you're using for determination of your sample size based on your design.

So again, using the A IML is not an excuse for not understanding how the research tenets work, how thematic analysis works, how sampling works, um, but it is a great tool to help accelerate this analysis based on those insights and knowledge we have.

So this is the Gemini finding.

Um, you notice that it, it looked focused much more on the social influence and intrinsic motivations of my participants rather than, um, payroll.

So this leaned more heavily on contingency and human capital than the, than chat GPT did.

Um, mostly because, and basically I agree with both of these.

I think they both had valid value for this small sample analysis.

Uh, basically what this told me when I looked at the two, two is that I needed more interviews before I could get these two, uh, collaborators to, to come to agreement or come to con to come to some level of a concurrence in what they're finding.

Um, so again, just to, you know, some takeaways from this presentation.

One, um, the, as a tool, uh, and like any tool, you wanna make sure you understand the guidelines, the requirements, the policies of your institution before you use any of them.

Uh, if you misuse the tools or you don't use them correctly, you're running a serious risk of having your, um, your, your work rejected, uh, particularly from the journals today.

But even within the university, your, your chair's gonna probably, uh, c**k an eye at if you're not sampling correctly, if you're doing an incomplete, uh, analysis, even using the ai.

So make sure you understand the requirements of the institution for use for citing, uh, and for application.

Also, I encourage everybody, you look at the eyes as a collaborator, um, it's a, it'll provide you some wonderful insight and a unique set of eyes on your data, but don't trust it implicitly.

Um, trust it, uh, but always go back and verify it based on the actual data and your own analysis of the, the findings.

Again, it'll give you some interesting things to look at that you might not have otherwise looked at.

It will help remove your, your biases from the analysis by giving you that fresh set of eyes, but don't trust it implicitly.

Always make sure that you're understanding, uh, what it's, what its findings are saying based on the data.

And again, AI may return different answers.

It doesn't mean those answers are wrong, it just means they're different.

And using multiple AI platforms as collaborators is a great way to get additional insight into your data.

But always remember, you're the principle investigator, you're the principle researcher, and so you're the one responsible for the conclusion.

So don't trust AI implicitly without, uh, uh, verifying and understanding what it's doing, uh, or what its findings look like.

Um, just a little plug, um, for the, uh, AI Research Center, if the folks aren't aware of the AI research group at the university, um, I encourage you to go take a look at it, join it.

Um, there's a, a great, uh, array of resources, materials, uh, et cetera that is, um, available to researchers out there.

Um, and if you're new to ai, uh, or if you're a veteran of ai, um, there's gonna be a lot of great information out there for you to use and improve your usage of these tools.

So give it a, give it a look.

Um, and with that, I'm gonna go take a look at some of the questions I see in the chat.

Somebody ask a good question about costs.

Um, and that's a good question.

Different AI platforms, um, are free or have a for cost option.

Um, I use the basic membership with, uh, chat GPT, so I pay like \$20 a year for it.

Um, but, um, there's a lot of great free tools out there.

There's tools for a fee.

Um, just be aware, uh, when you start using, most of them have a free entry point and the free entry point.

Um, usually the biggest limitation on the free entry point is whether you can upload attachments or whether you have to load the, the text, the, the data as raw text.

But, um, uh, even the free ones will do a great job with, uh, particularly the smaller data sets.

Um, Josephine, did that answer your question? You can, you can take yourself off of mute.

Hi, Dr. Rice. Thank you.

I totally forgot I could unmute myself, but yes, this is very helpful. Thank you very much.

Okay. Um, what other questions do we have out there? Let me look up in the two, Dr. Rice, there is another question in the chat, and then Gina has her hand up as well.

Okay. So, um, Sherry asked, I know you might cover this, but I don't wanna forget what is the citation expectation? Is there an a PA format or are we using ULP student requirement, et cetera? Uh, well, it depends on what you're doing.

If for the dissertation or for classroom, uh, post, there is, uh, there are specific guidelines.

Um, so they've updated the, uh, student code of conduct, and there's also a specific, um, university, uh, policy and requirement for citation and use of A IML, uh, particularly for content creation.

Um, so I would encourage you to always stay up to date on what those are.

There's links to those on the AI research group.

If you Google, uh, AI policy and University of Phoenix, um, it'll pop right up.

It'll be the top one on the list, uh, 'cause it's so commonly looked up.

Um, so yes, there is one.

Um, and there's specific guidelines about when you're required to include it as a citation in your classroom papers or in your, um, dissertation work.

Also know that every publisher out there, Elsevier, sage, uh, Emerald, I can run down the list, all have, um, their own guidelines.

Uh, and so don't assume the university guidelines are, are gonna be applied everywhere.

Some of the publishers are very restrictive and they, they will reject anything that has any usage of ai.

Others are much more, um, flexible and have specific usage guidelines.

So I was, uh, reading, reading the latest elver, uh, today actually, it, it just got republished.

Um, and they talk about AI as an embedded tool like the AI that's embedded in, um, uh, Grammarly.

Um, they don't require that you cite Grammarly, but if you are using, uh, chat GPT, you do have to cite that.

So they give very specific examples for different tools and for general usage.

Um, Gina, what, what can I ask Answer for you? Yes, hi there. Um, so thank you for the presentation.

It's very helpful. I've been, uh, wondering if, uh, my use of AI here and there was okay to do.

And now I know that as long as I, you know, I'm using it correctly, uh, which I am, uh, that it's okay.

Um, so my question though is, uh, going kind of back in time for me regarding IRB stuff.

So when you're getting I permission, would it be appropriate, like for future research, when I do future research, is it appropriate to say that I will be, or is it necessary to say that I'll be using AI tools to supplement, um, you know, my data analysis? Or is that, and then

second part, I'm sorry. Yeah, so lemme answer the IRB question first.

IRB is a, is an application for, um, data collection from a population.

Mm-hmm. And so you're typically not gonna be using the AI to collect data, so it's not generally relevant to, um, analyze for IRB, but it is gonna be relevant to your QRM application.

And so you're gonna wanna make sure you're working very closely with your URM and your usage of AI to make sure that, uh, you're in agreement for, um, how you're using it.

Um, and you're gonna have to be very explicit in chapter three about, um, your application of it, uh, so that when you go into, uh, QRM that you've got that, uh, uh, everybody's on the same page. Mm-hmm.

And so regarding then, privacy and confidentiality and all these kind of things, um, I wanted to try what you just showed at one point, but I was too scared to upload an interview of my participant into like chat GPT I think it was chat GPT because it's private, so obviously their name is not there.

I have pseudonyms from my participants. Mm-hmm.

But it's still their information and mm-hmm.

Once it goes into the chat, GPT, it's there forever. Right.

As a body of like an entity of information that can be found. Right. I mean, Um, well the, this can get to be a really long conversation, but few things.

One, you wanna make sure that, um, any usage of external tools is included in your informed consent.

Mm-hmm. Um, so that your participants are aware of that.

Um, the same is true if you're planning on doing a publication, uh, based on your research in a journal after your dissertation, that should be included in your informed consent.

Um, so make sure you're looking at the, the template from the IRB on the informed consent and make sure that you've got all the usages that you and potential exposures of data clearly identified in there.

So you notice one of the things that they called out here before I expose any of the data to the AI mls, I, I masked it.

Um, so there was no identification of individuals in that data.

There's, you know, still the possibility of, uh, accidental exposure to a population.

Mm-hmm. But one of the things I, I do, and I didn't call it out here, is I always clear the, uh, cache within the AI platform before and after I use it with any data like that to make sure that it's not preserved within the platform.

Oh, good. Okay. Good. I like that.

And I have one more little, tiny little thing. Sure.

Um, I, in my journey also discovered another tool called Site S-C-I-T-E.

Mm-hmm. And I like it quite a bit because it allows us to find peer review journals and specify the search.

And it's just, I, you know, not that our library is not awesome 'cause it is, but Site provides, it's just a different type of a search engine and it chat t will give you a thousand sources, but none of them will be peer reviewed.

None of them will even be, they'll just be there.

And I, you know, I tried using it in chapter two and it didn't really work out 'cause they don't exist.

But this site, uh, site was very helpful.

So I just wanna throw that out there.

Yeah. Well, thank you.

Thank you for the pointer to the tool.

I haven't used it, so I'll, I'll go take a look at it.

And that's what I love doing about these things.

I learned something new every time we Can do a workshop together.

'cause this is my jam. I just need to finish my dissertation, that's all. Yeah.

Well, and um, also remember that within the library there's an advanced search function.

So you can limit your searches to peer review journals or full text journals.

So you can constrain it to make sure what's coming back, uh, is peer reviewed.

And if you're doing Google Scholar searches, you can also query there to say that you want only peer reviewed and only want, um, pull text.

Right. And, uh, it will help you get to the right kind of sources.

Um, so there's lots of great tools out there, but thank you for the reference to site.

Thank you. Yeah. Jim, Mr. Lane.

Dr. Lane.

Hey, Dr. Rice, thank you so much for this, uh, uh, appreciate all of this that, you know, you're, you're a treasure.

Uh, so, uh, I heard what you said about the, the, the, you know, the cost, uh, fine.

Uh, you used, uh, an example, um, of four interviews, which I understand why you use only four.

How robust are these things? And even the one that you pay for 20 bucks a year or whatever, could they accommodate 15 or 20 interviews or, or guess the question is How many? I just, I just uploaded one.

Well, it depends on how big the files are 'cause it's the size of the files, but I just did one with over 900 interviews in it.

Oh ha. Okay.

Yeah, it's, especially when you start to do, um, uh, and, and Dr. Smith is gonna do one this summer on, um, using AI MLS for literature reviews.

Um, it'll take, most of Theis will take just about any format.

So you can actually use it for doing thematic analysis or, um, synthesis of multiple journal articles.

So I just did one with, uh, 30 different journal articles.

Um, so they'll, they can handle some fairly substantial data sets.

Wow. Amazing. Thank you. I appreciate that Dr. Smith, you're probably watching these questions better than I am.

So yes, I see a question from Dr. Johnson.

So she asked Dr. Rice, can you speak to how long the AI tools keep our interactions with them? If we want to refer back to the responses, is there a way to archive them? I'm particularly interested in perplexity in this regard.

Well, I can't speak to perplexity 'cause I haven't done enough with perplexity just yet.

So I, I don't have a lot of firsthand experience, so my apologies for that.

Um, for, uh, Gemini, for chat, GBT and for, um, copilot, um, they will retain your information until you ask it to forget.

So you can ask, um, the AI to forget a particular session or clear its contents completely.

It'll retain it as long as, um, you allow it to keep it there.

Um, and if you're doing a complex analysis with a whole series of questions, and if you go back and look at the slides here, you notice I was asking questions about the data set that I only loaded once.

Um, it'll retain that for the duration of the session.

It'll retain it for days.

You can go back and select a prior session and ask additional questions.

I, as a, as a routine, I don't leave legacy sessions up there.

Um, I always ask it to clear the sessions when I'm, when I know I'm done with them.

Um, but, um, it, it depends a little bit on the, uh, AI platform that you're using.

So, um, I would, I would always point you to the, um, guidelines and the, the practices of the AI that you're using on, uh, data retention and data destruction.

But they all have them. Um, every one of the ais have a data retention, data destruction policy and, and how to remove your content from their platform.

But I'm sorry, I can't, I can't speak specifically to propensity.

I haven't looked at it, I haven't spent enough time with you yet.

Dr. Rice, there is a question from Dr. Baron.

If you ran your example again today, would you get the same answers? Um, if I ran it again in the same model.

And so every one of the AI mls, um, have model versions.

The model version is a, um, the databases that it's using as it's experienced, it's applying for doing the data analysis.

So the example I ran today in chat, GBT was in version, was in Model 4.0.

If I were to run that in an earlier version in 3.0, it would get a different answer.

But if I were to, um, based on what I did today, cleared, cleared prior history and ran it in 4.0, I would get the same answer.

Um, that's generally true of, of most of the ais.

Um, so always be aware of the version or the model that the AI is using because as a refresher, push out new models and you'll see all kinds of different models that pop up.

Even with a chat, I think there's six or eight different active models right now.

You can select a model and it'll give you different answers based on the data repository, it's retaining.

Uh, so while I use the example of two different, um, uh, a imls, I, I did an example with Chad, GBT 4.0 and, um, Gemini two point, uh, version two, I believe.

Um, I get different answers in those two.

If I were to do a chat GPT on 4.0 and 3.0, um, I would get different answers from the different models.

Um, again, the, it's a second set of eyes.

I, I view each model as a different personality.

Um, and every AI has a different personality.

It brings its own perspective and experience and, and set of eyes to the problem.

And I treat them as collaborators, not as definitive answers that, uh, I can lean on.

Uh, again, I'm the, I'm the principal investigator, so it's my decision whether that's the answer or not.

Does that make sense? Yes. There is a follow up question. Mm-hmm. Okay.

Um, oh, so, um, Ms. Gray asks, I'm finishing up my chapter two.

How often are you running this webinar? It would be a great refresher.

So I can answer that really quickly.

We, I don't know where we're gonna run it again, but we have recorded it so you can access the recording for this webinar, um, definitely in the research methodology group and most likely also in the AI research group.

So you will have access to that.

And the PowerPoint there is a follow up question from Dr. Johnson.

I'm more concerned about keeping the information from an interaction that I am in deleting it.

I guess the best approach then is to copy and paste the information into a Word document if I want to refer to it again later.

Oh, absolutely. Yeah.

It, um, copy and paste it if, um, you know, most of the imls have local clients as well, so you can run it in a browser or you can run it in a local client, just like you do with Microsoft Teams.

You can run local client, run it on the web.

The local clients get more likely to be, and, and basically most of them are pretty persistent to you, clear them.

But, um, if you're concerned about losing, just like any other data, you want to back it up and, you know, just copy and paste it into a Word document.

It's a good, good way to retain it. I do that. So, Hi, Jim.

Hello. Ooh, what, what can I answer for you? Yeah, I just wanted to follow up.

I, I, um, coincidentally I attended a, um, an, an AI workshop hosted by Sage Publishers today.

And I wanna tell you that your presentation is as good as, or better than that hosted by Sage.

But, and, and these guys were on a kind of a, a more general level.

They, they didn't go specifically and show examples, uh, as you have done so in their defense.

But, um, couple of points that made that, that are the same as yours was, one, interrogate the data by revising the prompts and, and beginning more general prompts and then being more specific, which I think is exactly exactly what you said.

Mm-hmm. Uh, interrogate the data from different ways, using different questions.

Um, uh, also they, they advocated using different platforms to, to, to, uh, for the same purpose.

So exactly what you said.

And they also emphasize that, that, that we are the researcher.

This is a tool. Um, it's not a ghost writer, it's a tool.

It's an aid, uh, just like the library is. Um, mm-hmm.

Uh, and, and, you know, the same thing to check the results now that we're talking about doing, doing like lit reviews and research, obviously check to be sure that, um, the sources are, are, uh, valid, right.

And then, uh, also check the publisher guidelines because it's, it's a, it's a, it's an emerging field.

So, uh, I just wanna say you hit all, all the key points and, um, you, you, you did a great job. Thank you.

Well, thank you. I appreciate that's, it's, that's a high bar to hit, so I appreciate that feedback.

It's nice to hear, um, you hit it. Yeah, Yeah. Speaking about the, the refined prompts.

And then I think your, your, uh, comment triggered something for me.

You can ask these prompts all at the same time.

Um, and you notice I didn't do that.

I, I asked incremental prompts as I went, because I always wanna leave myself the latitude to refine something.

So, for example, if there's a list of seven, uh, codes out there, and this one presented seven codes to me, and one of those is a low priority, and I don't even want it considered in any future analysis, my future prompts would be, uh, would say, do the thematic analysis.

Uh, but please exclude this, this code.

Um, so, you know, doing the incremental validation as you're presenting the prompts is important also, if you ask it all the questions at the same time.

So if I were to say, um, you know, present me with the codes and give me a summary table, it may only give me the summary table and skip the whole intermediary, um, set of presentation of the data, which robs me of the opportunity to go validate the analysis.

So I, I encourage short but very precise questions, um, that you plan on building on.

'cause again, one of the things that's unique about an AI versus a, like a Google search engine or a Bing or a Yahoo, is that it remembers that session.

And so I, you know, I only uploaded the data at one time.

It did an analysis on the initial coding one time, and I can refer to that as knowledge that it has accumulated within that session.

Um, so I can ask follow up questions and clarifying questions.

And that's an important part of doing this process, uh, effectively.

'cause that allows you to validate the, the responses between each of the questions.

But, so, uh, great feedback, Dr. Lane. Thank you, Dr. Johnson.

Sorry, I have a lot of questions today.

Um, I know that we're talking about data analysis, but we have a couple minutes, so I'm wondering if I could ask you about data collection.

Mm-hmm. Is there an ai, uh, tool that would be helpful in like a content analysis of, um, a topic that we're looking for on social media? Is there a way to use AI in essence to mine data like, um, like from Twitter or Not that I've found that I trust? Um, they, they are out there. Um, okay.

But I would not trust 'em.

I, you know, I, there's not a, I haven't found a good way to remove the bias from the, the web scrapers.

Um, so that, that, and that's just my personal opinion, right? Um, I'm sure there's people that would argue with me about that, but there are some tools out there that do, um, you can point them to, um, sets of websites and have it, uh, evaluate content.

You have to very, very careful when you do that, that you've got permission from those websites to do that data collection.

Um, if you probably remember from the election a couple years ago, that was one of the AI companies out of California was going and mining data from, um, I think it was YouTube, uh, YouTube and Facebook, and, uh, in violation of the terms of use of those platforms.

So you in a lot of legal trouble very quickly.

So again, for, um, if you're going to do a content analysis, I would personally still go collect it by hand rather than using the, the minors.

So we're just looking for, uh, tweaks that, uh, discuss our topic.

This is what, this is not a student project, it's a faculty project.

And, and, uh, you know, perplexity, I had a long conver, I had a long interaction with perplexity yesterday, be like, got to become, you know, my second best friend or something.

But I, it, it says that it can't do that.

But it gave me a lot of excellent ideas for keywords that we hadn't thought of before for us to search for, you know, because we're just using the publicly available information.

And so it was able to give us additional keywords.

And, and then also it gave us, um, individuals and organizations who likely would be discussing this topic on Twitter so we could just go search for their tweets.

I wasn't expecting that.

I thought that was interesting. That was, Yeah. No, I, I think what perplexity is doing in that regard is, is what it should be doing, which is leaving the, um, approval to do that search up to you.

It's not going and doing it on your behalf because it doesn't know your, um, access or your right to search.

Um, as a researcher, we have to get, we, we always have to get our per forms.

We have to get permission to mine public sites and social media sites before we go do that.

Um, and the, the AI that, that are capable of doing that today, that don't understand the terms of use of those platforms or have a good way to get them consistently.

And so they, they run the risk of putting you, uh, in legal jeopardy, uh, by mining data that you don't that in an app in inappropriate way.

Hmm. That make sense? Yes. Um, so, so like we have IRB approval to collect this data from Okay.

From Twitter, and we're only collecting it from public tweets.

So, And that's fine. You're, you're being consistent with the terms of use.

You know, that IRB knows that, but perplexity doesn't know what your limits are.

Okay. That's, but its why also Says, it also says that, it also says at this point, it doesn't have the capability to do that.

So I thought that Well, it, It can, yeah.

It doesn't know what your boundaries are.

Right. I gotcha. Okay. Yeah. Okay. Thank you.

Sounds like a, sounds like a fun project.

It is.

Did I see another hand, Papa? I think that was Dr. Johnson. Okay.

Um, are there any other questions? All right. We're right at the top of the hour. Um, how did that, How did we end it? Right at the right on time? I never end right on time. It's amazing.

Um, please join me in thanking Dr. Weis.

There's tons of accolades in the chat, um, for your, um, amazing presentation.

Um, I also have put the link for the feedback survey so that you can share your, uh, thoughts about the presentation and additional presentations that will be helpful for you.

Um, I put in the chat the link to the AI research group, the website so you can see all of the resources there and join our team site.

Uh, and also shared information about where this information will be, right.

So it's gonna be posted most definitely in the research methodology group, both the PowerPoint and the recording, and most likely in the AI research group as well.

So we are so grateful to Dr. Rice for sharing his knowledge on this topic, and thankful to all of you for being here.