Sadie Witkowski: Ian, you’re part of a union as a Chicago public school teacher, right?

Ian Martin: Yup! I’m part of the Chicago Teachers Union!

SW: I bring it up because we’re going to try a slightly different format for our episode…

Caitlin Parrish 00:23
My name is Caitlin Parrish. I'm a showrunner and writer and a member of the WGA. Most recently, I worked on the shows, foundation and orbital for Apple and Netflix respectively.

SW: For those of you not familiar, WGA stands for the Writers Guild of America. A union of writers in television and film that have been on strike for over 100 days at this point.

Caitlin Parrish 00:47
It includes someone who has done writing. You have to have a writing job. To my knowledge to be a member of The Guild, you have to have performed writing services to a certain degree on a union job to be eligible for guild membership.

IM: I actually have a little experience with this, we went on strike for two weeks back in 2019. It was two weeks of hell, I don't know how they're able to do this for 100 days. What are Caitlin and the rest of the WGA striking over?

SW: Well, there are multiple issues but the big one we’ve been hearing about is something we’ve covered on this podcast.

Caitlin Parrish 03:50
basically the WGA and sag and I'm certainly less informed when it comes to the exact demands of sag are both looking for some very strong language from our corporate partners, ensuring that obsolescence is not on the horizon with AI. For writers, we already have an example of this with, you know, chat GPT being able to generate scripts, they're not on par quality wise, with human writers yet, but looking at how exponentially quickly, technology seems to grow more sophisticated. This seems like something that will happen that could happen and so we're looking for a preemptive strike, with And from the studios ensuring us that humans will not be replaced by AI to generate scripts at any phase of the process.

IM: [said reflectively/thoughtfully] AI-generated scripts… Oh right! Didn’t we do an episode around GPT?

SW: Yup, it was Chat GPT and the recent explosion of AI writing programs. So it turns out that episodes neatly leads us into this current strike and the current impasse between the WGA and the AMPTP - the Association of Motion Pictures and Television Producers. That's basically all the big studios and streamers like Netflix.

Caitlin Parrish 14:29
I think that AI is a very prominent demand because it's very existential demand. And it's one that not just artists are grappling with right now, a lot of people in a lot of industries and a lot of fields are thinking about how and when they might be replaced by machines or computers.
Caitlin Parrish 17:53
I personally welcome our computer overlords at a certain point. But that question of soul has not yet been answered.

IM: I remember when we talked about GPT and how it could potentially upend all sorts of careers. It's wild that we're now starting to see policy fights about this and the future of some creative fields!

SW: I know, right? And it was only at the beginning of this year that we covered that topic!

IM: Well seeing as it's extremely relevant to Caitlin's union's fight, maybe we should go over it again?

SW: That was my thought exactly! But let's not forget introductions. I'm Sadie Witkowski

IM: And I'm Ian Martin

SW: And you're listening to Carry the Two, a podcast from the Institute for Mathematical and Statistical Innovation, AKA IMSI.

IM: This is the podcast where Sadie and I talk about the real world applications of mathematical and statistical research.

SW: Or in this case, we're reviewing our old episode on GPT-3 with a new twist!

IM: Now, let's jump straight into our January episode to cover the basics of chat GPT and similar AI language models.

SW: and don't forget to hang around to the very end to hear a few last words from Caitlin about the WGA's fight!

[musical transition here]
[old script starts here]

Allyson Ettinger: 45:19
It's very easy to be convinced by a model that seems human-like to think this model is going to be as trustworthy as a human. Humans are not all trustworthy, but at least you can trust that they understand language and know how to reason about it.

IM: This sounds like someone who hasn't had to deal with teenagers in a minute cuz let me tell you, they're not all great at explaining their reasoning....

SW: [Laughs] Maybe computers are more trustworthy than teenagers? Anyways, let me introduce Allyson Ettinger. She's a computational linguistics researcher and faculty member in the computer science department here at the University of Chicago.

IM: Seems like she's the perfect person to explain GPT-3 to us!

SW: I thought so, after I asked her about her research field.

AE: 1:21
I do computational linguistics. And for me, what that means is a combination of work in natural language processing, language component of artificial intelligence, where the focus is on looking at the robustness of meaning extraction using methods of analysis and evaluation that are often inspired by areas of cognitive science.

And then on the other side, we also take the insights from what we get out of artificial intelligence, selectively so and apply that to cognitive science questions. So we also do computational cognitive modeling, particularly psycho-linguistic modeling, that is looking at using computational models to study how the brain processes language in real time in humans.

IM: Ok, so now that we know what she does, what can she tell us about GPT-3?

AE: 5:11
So GPT three, is a recent, very large example of a paradigm that has been dominant in natural language processing for a few years now. And this is the use of so called pre trained language models. And so what these models are, is a type of model. That is, it's a deep neural network model. So it falls within this category of deep learning that folks may or may not have heard this, this buzzword. And pre trained language models are

trained in a particular way, where the way that they learn everything that they know about language, is by learning how to predict words in context.

SW: And I know this all sounds crazy technical or like something from science fiction, but we interface with these kinds of models all the time.

AE: 3:57
anytime that you are interacting using language to interact with an artificial system, like Google Translate, or Google search, or Alexa, anything that you're using language to interface with an artificial system that could roughly fall under natural language processing.

IM: Are we going to annoy our listeners if I say, “hey google!” or “hey alexa”?

SW: [laughs] Don’t do that! And anyways, we aren’t really going to focus on these examples of NLP, that’s natural language processing.

IM: Why not?

SW: Well, when you get into NLPs that use audio commands, you’re adding a whole new level of complexity. Not only does it have to understand language construction, but it’s also having to deal with acoustics and accents and audio waveforms. For this episode, we really want to just stick to the programming required to understand language like GPT-3.

IM: And besides, neither of them use GPT-3, right?

SW: Well, their software is proprietary so they *could* have some elements of GPT-3 in there. But it’s not the basis of their function, no.

IM: Then let’s go back to… NLP? That was the natural language processing idea?

AE: 3:18
So when I refer to NLP, I'm specifically referring to the subfield of artificial intelligence that is working on trying to better engineer language processing capabilities in AI. And so what we're trying to do there is design models that can try to process the things that humans say to those models, the things that humans type to those models, and be able to process that and produce outputs, and some sort of semblance of understanding that will allow those models to do downstream tasks in a way that is effective. And that reflects the way that humans would respond to language.

SW: So we’re kind of using GPT-3 as the case study to understand more about this class of AIs in general and how they learn to write and respond to prompts we give them.

IM: How *do* they learn?

AE: 7:54
Honestly, there is a ton that we still don't know about exactly what it is GPT three has learned how much it is just sort of producing stuff that it managed to memorize. Because we it's very difficult to know exactly what it has seen and how similar the things it produces are to the things it has seen. So this is an ongoing challenge in the field. Because it's less interesting to us if it's just sort of repeating parroting things that it has seen before.

IM: So it has to already get a bunch of information in order to guess at a best response to you?

SW: Yeah, that's the idea in broad strokes. NLPs are pre-trained on a bunch of sentences and use that information to create distributions to make predictions.

AE: 5:53
So if I say something like, "he caught the past and scored a touchdown, there was nothing he loved more than a good game of blank", you can, if you're familiar with American football, which, obviously not everyone is, but if you are, then you can guess very quickly that what should go there is football, because it's a very constrained type of context. And so you can use information from that context to predict what should come next. But in order to make that prediction, you need to know a lot of stuff about language. And so this is a type of learning signal that these models have been able to take advantage of. And this has become an incredibly popular and an incredibly effective way of training these deep neural network models to learn language, basically just teaching them how to predict stuff in context. And they learn things like sentence structure, and things like okay, things that are going to occur near drink are probably going to be liquid like things, these are the types of information that intuitively you can expect to learn on the basis of prediction in context. And so this is the basic principle driving these language models. And so GPT three is an example of one of those such models. But it has just been scaled up massively. And it's been trained on a ton, a ton, a ton of data. And so basically, with more data, and more size, in terms of the parameters of these models, you tend to see continually increasing performance. And so GPT three is an example of that paradigm. And it's been a particularly successful example.

IM: So in a simple sense, GPT-3 is literally just guessing the next word based on what it's most often seen before.

SW: That's what a basic NLP would do, yes. GPT-3 is more advanced than that. Both because it has a TON of data that it's learning from and because it's making more wide ranging predictions than just what word comes after another.

AE: 8:22

it's very clear that it is able to generalize, it's not strictly producing things that it has memorized. There are ways to test that. But But exactly the balance of sort of memorization versus intelligent looking generalization is still very much up in the air.

IM: Ok, so GPT-3 is learning through experiences, aka the data we give it. But how is a computer able to generalize?

SW: That's where we get into another one of our jargon words, or concepts in this case: deep learning

IM: Oooooo, sounds…. Deep?

SW: I don't know why, but my brain really wanted to make a reference to The Boys character just now, the Deep.

IM: You just did make the reference… But stay on track! What is deep learning?

AE: 21:22
So deep learning very simply just refers to two models that are deep neural networks. So they're just neural network models with lots of layers.

IM: And layers are…?

AE: 21:32
layers are basically, matrix computations with other nonlinearities, and things like that. So basically, the more layers, but you can also think of it in sort of graphical form.

IM: Sadie, you know that cutting to Allyson's explanation didn't make any sense to me!

SW: I know, sorry I couldn't help myself. I've found the easiest way to understand the layers necessary for deep learning is to compare them to the cells in our visual system. Want me to walk you through it?

IM: Finally putting that psychology degree to work?

SW: You betcha!

[background music for this explanation]

SW: So you probably know that our visual system doesn't work like a camera snapshot

IM: Sure

SW: To start with, the cells in our eyes, the rods & cones that is, really only pick up on the basics. For the cones, we have 3 different cones that respond to red, blue, or green light. So a red cone will only fire when it's hit with a red wavelength. And same for the blue and green cones.

IM: Is that why old tube televisions were RBG?

SW: [laughs] probably? I don't actually know. But to simplify, we're going to skip over the optic chiasm and the tons of pathways and hop straight to the visual cortex. This chunk of brain, located at the back of your skull, processes the visual information in a series of layers labeled V1 to V5.

V1 acts as a kind of sorting area from the retinas, the eyes and passes signals along to V2. V2 then passes the information to more *layers*, namely V3, V4, and V5 while also sending a feedback loop signal to V1.

V4 handles mostly processing color while V5 focuses on motion and there are a bunch of other bits involved. But I hope that from this description, you get how the brain's layers are needed to process the complexity of a visual scene.

IM: So are these literally layers stacked at the back of my brain like a layered cake or something?

SW: Heh, you're making me hungry with that analogy. But yes, the layers are literally stacked on top of each other, with connections both feeding forward to the next layer of cells and feeding back to the earlier cells to provide a kind of feedback. By passing the visual information through these layers, we develop the whole visual scene.

IM: So bring this back to deep learning…

SW: Deep learning relies on this same process of layers where information is passed from one level to the next, with slight tweaks made each time and feedback layers that are used to correct the input and adjust the system. The more layers you add, the more complex calculations you can do on the material.

AE: 23:00
in early layers of processing as there's like this transformation, this transformation, this transformation, the representations, the types of information that are being represented at these different layers, is different, it seems like roughly more or less gets sort of more

complex, more high level more abstract as you proceed through the layers. And so if you have like a sentence, like the cat went to the store, you may have in early representations have just sort of mostly just representations of the properties of those individual words, maybe their syntactic properties, and then later you start to get other things like semantic dependencies between those things. You know, if you have the cat went to the store, and then it meowed, right, right, the, it refers back to the cat. And this is something that models and humans need to be able to compute in order to know what it means in that sentence. So this is a type of thing that may, we may expect to get represented sort of later, as those earlier sort of lower level types of things are represented in earlier layers of these of these deep neural network models.

IM: So how many layers does GPT-3 have?

SW: It has 96 layers and 175 billion nodes.

IM: Daaaaang

SW: Yeah, and while we can tell how many layers there are, because we built it that way, it isn’t always clear how GPT-3 is generalizing based on the interactions between those layers or how it updates based on feedback.

AE: 9:53
So we know exactly what the sort of architecture of the model is that's completely transparent to the people who designed the model. But exactly how the learning of the model plays out. That's very, very opaque. That's something that folks are still very much trying to figure out even with smaller models that have more open access.
10:40
we know what all of the pieces are going into the learning process. But what happens is in that learning process, and what ends up being represented, and what sensitivity is happening, what strategies the model develops, in order to map between its inputs, and its outputs, those are the types of things that are not nearly as clear.

IM: So would this be like knowing if a neuron is in layer 2 or 3 of the visual cortex, but not knowing which other layers it's connected to?

SW: More like knowing the neurons, but not knowing how they tweak the information as they pass it along to another layer.

AE: 11:16
We know what the outputs are, we know what types of computations the models able to do in order to learn the functions that it's going to learn between the inputs and the outputs. But yeah, exactly how it's going to choose to weight those different inputs and produce that giant function that is going to map between the inputs and the outputs. That's something that's not as clear.

IM: Wait, so humans wrote the model of 96 layers and how they connect, but we still don't know how it's learning? Allyson says it's opaque, which feels weird because we know all the ingredients going into it.

SW: Well, it's an incredibly complex model. Like, did you know we have a map of all 302 neurons in the C. elegans worm? It's one of the most basic animal models that we use for neuroscience research. But even with mapping all the neurons and knowing how they all connect, we still aren't able to recreate all its behavior. GPT-3 is similar to that.

IM: So, it's kind of like if someone gave me a list of ingredients for a cake and the temperature setting for the oven, but I still don't know the steps to get to my final outcome- in this case, tasty chocolate cake.

SW: That's not too far off as an analogy. And just a reminder, the communication between these layers doesn't just spit out a good or bad sentence. Rather, it's using these transformations between layers, via weights, to make probability predictions for the next word in a sentence.

AE: 25:40
There are a lot of different types of parameters in, in these models. And all of those are going to define, you know, how information gets communicated from layer to layer between the input in the output. And so it's going to be a little different from a 'that was a good sentence, that was a bad sentence', because it's going to be more like, How close was the model's probability for the next word to the actual word, right, that actually occurred in the text?

IM: So to recap, the deep learning part basically means layers that are passing information to each other and doing computations as it moves between the layers?

SW: Yup

IM: and the NLP, natural language processing, is using this computing technique to make predictions about what word comes next in a sentence.

SW: More or less. It makes predictions about possible words and selects the one it thinks will fit the best.

AE: 26:50
the model is going to be predicting probability distributions, really across the entire vocabulary at every position. And so you could give the model that that sentence that I said before, the he got the pass and scored a touchdown, there was nothing he loved more than a good game of this is taken from an existing psycholinguistic experiment is

just one of my examples that I have, on the top of my head, there was nothing he loved more than a good game of blank, and you could just what the model is going to output is something that you that can be interpreted as a probability distribution over all possible words in the vocabulary. And then you can just look and say, okay, it assigns the highest within that probability distribution, the word that receives the highest probability is say, football. And, and so it which, for instance, with the BERT model, which was the one that I was testing with those sentences from which that sentence was drawn. In the case of BERT, I seem to recall that it had a really, really confident prediction. In that case, it was like point seven, something like that. Whereas in other cases, you may have the most confident the highest prediction be like point O O two or something like this. Or lower, you may have have it much more distributed, because there's a lot less certainty with respect to that prediction. But yeah, it's going to be a probability distribution over that vocabulary.

SW: And based on the probabilities that it identifies, it selects a word that it then gets feedback on and can update the distribution probabilities. Let’s look at an example sentence like…. Before pouring the Earl Grey, she put sugar in her ‘blank.’

AE: 26:18
the model thought that, that, that iguanas should really be the next word here. But actually, it was teacup. And so we really need to update our distributions here. And so you're gonna, that's going to then be back propagated through the model, and it's going to be updating the various parameters that contributed to that prediction.

IM: Man, 96 layers and I don’t know how many computations between them, just to get you a percent likelihood that you should say teacup instead of iguana… I guess this explains why the human brain has so many neurons and connections huh?

SW: Well, while programs like GPT-3 are super impressive for their complexity and ability to produce meaningful-sounding language…. That’s probably not how an organic brain works.

AE: 8:42
the way GPT three learns probably is quite different from the way that a human learns language. And so the analogy goes, at least to some extent, well, because the more humans have been exposed to the more they can sort of draw on that information. And similarly with GPT three, but humans in terms of the basics of learning language in the first place, probably learned in a very, very different way than GPT-Three does, because humans, quite simply because humans don't seem to learn language purely by predicting.

SW: I should mention that Allyson’s original area of research was cognitive neuroscience, just like mine! So her research does a lot of work comparing between

NLPs like GPT-3 and human cognition. There are some similarities, but they definitely don't work in the same way.

IM: Oooo, this is reminding me of Ben's episode from season 1 on language learning!

SW: Yeah, I think this kind of work is a good companion piece to some of the ideas he shared. Listeners can find that episode in the show notes if they're curious.

IM: Ok, so I actually do feel like I have a bit of a grasp on how GPT-3 and NLP systems work. But, it's still a lot to digest. Can we take a break?

SW: Sure! And when we get back, I want to talk about the implications of this work. What happens when we assume more intelligence from GPT-3 and some of the dangers that can create?

[Break]

[ad - Not another politics podcast]

[break ends]

[Old script starts here (cut the 2nd GPT interview)]

SW: ~~Well,~~ even with these really sophisticated NLPs relying on deep learning and machine learning, they still aren't true artificial intelligences. They fail in really specific ways, showing that we shouldn't just trust the program to always make the correct choice.

AE: 46:11
maybe Artificial General Intelligence is actually solved with this latest model. But usually the pattern is we then find the cracks, we find the brittleness we find the types of superficial heuristics that they were using.

IM: What does she mean by brittleness? I mean, I've been pretty impressed with what I've learned about them.

SW: Sure, a lot of these programs look really impressive at first. But dig around a bit and you'll find it's not providing information like a human would.

AE: 44:11

To me, I would say that the main risk, and the main source of frequent misunderstanding is simply overestimating the actual intelligence of these models, right. And this is, this is a very easy thing to do. But it sort of happens repeatedly that people often will sort of take at face value, what looks like intelligent behavior, but which folks like myself and other other folks who do the type of research that I do, we tend to take those things more with a grain of salt. And we tend to be less surprised when subsequently results come out and say, Actually, we were overestimating the capabilities are actually brittle in this way or that way.

IM: So GPT-3 is just really good at faking it?

SW: In a sense, yes.

AE: 38:29
Often, when we see really striking results like this. Once we have a chance to analyze it a little bit more, we find, oh, actually, there were some sort of heuristics and things like that, that they were able to use or some cheap ways that they were able to cheat. So this is sort of hot off the presses, really recent, interesting results, and I suspect that we will continue to find ways in which we should, you know, sort of temper our expectations and interpretations of this

SW: So GPT-3 and other similar natural language processing programs are just filling in the blanks based on the data they're trained on. In some ways, it reminds me of horoscopes. They sound really special and unique to our signs, but are just some general fluffy predictions.

IM: Aries would say that! I, as a Cancer on the other hand, would say…. Yeah ok that makes sense.

AE: 40:30
if you look carefully enough, and test carefully enough with respect to its outputs, what you're gonna find that not only does it not have domain expertise, but it just kind of probably completely lacks any type of common sense knowledge, it probably doesn't even have a lot of the sort of key levels of understanding of what it's even producing. It's just producing stuff, that's high probability, you know, and that ends up producing really impressive outputs, given how much data it has seen, and how much it has managed to be able to represent in terms of abstractions in all of those different layers.

IM: In some ways, this actually makes me feel better about GPT-3. At least I know that our jobs are safe, for the time being.

AE: 39:46
Is it coming for our jobs? No, I think there's a lot of very clear evidence that ultimately, it is going to make some pretty striking mistakes that humans would never make

SW: Exactly, if anything, GPT-3 is like a really sophisticated tool. We embraced the help of computers when it came to storing information or writing new documents, but it’s ultimately just a tool.

IM: And it’s a tool that we have to train, right?

AE: 41:40
So in terms of how much human input it needs, I think it's going to vary like there are ways that you can set it up to do particular tasks. I mean, as it is, its whole job is just to produce high probability text, basically, and to predict what words should come next. And so you need to do additional things to sort of prompt it to do one task versus another.

SW: Right, and this leads me to the other big point that Allyson made. Just like a tool can be abused or misused, the same goes for GPT-3.

AE: 42:05
It's probably inadvisable, almost certainly an advisable to trust, a model like GPT, three to do anything that actually matters

IM: Based on our conversation so far, I’m guessing Allyson isn’t saying this because the robots are coming for our jobs.

SW: Nope. The problem is when these kinds of programs seem to over deliver and are taken as gospel. When you implement GPT-3 without any oversight and just take its responses at face-value, that can be really dangerous.

AE: 43:05
these are models that learn on the basis of the data that they receive. Right. And so you can't blame them for producing things that resemble the data that they're trained on. And so, but you absolutely, we should absolutely hold ourselves accountable with respect to how much we trust models that we know work that way. And, you know, end up deciding to deploy them for public use.

IM: So, junk training in, junk responses out?

SW: Yeah.

IM: So you could also say that the AI will reflect the biases of the people who programmed it?

SW: Yeah. This is actually something we saw with the Microsoft chatbot on twitter a few years back.

IM: Oh no, what happened there?

SW: So I should say this wasn’t GPT-3.
But, a lot of the data that chatbot was trained on came from the worst elements of the internet and the chatbot ended up reflecting that. It started saying racist, jerk things. As is the case with any machine learning, we have to be aware of the kinds of inputs they're getting and not just immediately trust the outputs.

IM: This is honestly starting to sound like a metaphor for human relationships. We aren't always aware of the things that formed people, and why they might be acting like a jerk. But anyway... at the end of the day, GPT-3 is super impressive and is probably helping advance our understanding of language and all sorts of neat stuff, but we shouldn’t be trusting it with anything big like online counseling any time soon?

SW: I think that’s a succinct way to view it. And you’re definitely right that studying GPT-3 and similarly trained deep learning programs is helping advance our knowledge in all sorts of fields.

AE: 48:14
for us, one of the highest priorities at the moment is continuing to better understand what is it that these models are learning, where they're brittle, and where they're robust? These are really very critical questions to answer in order to, to fill in these gaps that we've been talking about where we say, you know, how much can we really trust these models? What is it that we can trust them to do? And where can we expect them to fail? So that's the type of thing that we've been prioritizing

IM: So basically, they’ve built the model and shown that it works. Now, they want to understand the strengths and weaknesses.

SW: And to understand why and how those strengths and weaknesses exist. Science and research is often just the questions: how does it do that? Why does this work or not work:

AE: 48:37
And then the next question, as we identify, you know, what these limitations are? A critical question is, how far is the language modeling paradigm going to take us? So this basic paradigm that we've been talking about this whole time that's based on word prediction and context, based on what we've been seeing, there are reasons to think that there will ultimately be a ceiling in terms of what we can achieve with respect to human like language understanding, using this type of learning paradigm? We don't know this for sure. But there are reasons to think that it's reasonably likely that this will not be the final thing that will get us to human-like language understanding. And so then the question is, what is it that we need? Can we take other approaches to complement the language understanding and combine those things, since humans also are doing this type of predictive statistical processing and things like that? This is absolutely, you know, we don't necessarily want to lose all the things that we have gained that are good using language modeling. But there seems to be, that there's a good chance, we're also going to need to complement that with things that allow us to do more systematic

language understanding. And so what that's going to need to be that's really the big open question.

IM: No wonder Allyson works on both machine learning and human language learning. Sounds like we'll need a combination of both to continue advancing artificial intelligence.

SW: Yeah. GPT-3 is far from perfect, but researchers are hard at work on a fourth generation, GPT-4. And it'll take findings from both human cognition and computer science…

[musical transition]

IM: Listen to us back then, don't we sound so young?

SW: Ian [laughs], that was only from January… so like, 8 months ago.

IM: Look, I like to think I've grown a lot since then. [indignant]

SW: [laughs] Ok, but looking back, a lot of the concerns with GPT that Allyson brought up are directly applicable to the current WGA and SAG strike.

IM: SAG is the screen actors guild. That one I knew!

Caitlin Parrish 05:06
right now we're basically saying, Hey, could you please ensure that we not become obsolete meat sacks because of robots. And the studios are saying, we really don't want to ensure you have that I believe that our request was met with no, we're not going to do that. But I'll tell you what, we'll meet once a year to talk about your AI concerns. And
05:40
this seems to be a crossroads, and a do or die moment. If we don't get it now, it's unlikely that we will get it in the next three years. And the technology might be completely out of the barn by then.

IM: So the unions right now aren't worried about GPT-3 writing the next season of Abbott Elementary?

SW: No, but they're worried that GPT4 or 5 or 6 will be entrusted to write the whole thing. Or even worse…

Caitlin Parrish 06:11
Have GPT vomit out a script and then have a writer come in at a significantly reduced fee to give it that human touch.

IM: Oh yeah, I'm sure Caitlin didn't go into the creative field of screenwriter to have to play nanny to a chat bot.

SW: She wants to tell her stories, not fix dumb computer mistakes!

IM: And I'm just thinking about what inspires humans to create. It's often this driving need to tell a story or share an idea.

SW: Right, using stories to explore themes and try to make some kind of moral or philosophical argument - that's a huge part of what writers do. But AIs….

Caitlin Parrish 10:15
its form without ethics or intention, which is also very scary when you consider how important representation is, and how much artists should be held to account in terms of what they decide to put out into the world, you can't really hold the machine ethically accountable for something they put out in the world. And that makes for a very muddy situation.

IM: [big sigh] Yeah, that's a good point. [small pause] But wait a minute, earlier we mentioned that SAG, that's the actors, have joined the strike. How do they come into this?

SW: Well the actors are more concerned about AI creating digital versions of their faces and bodies, which are then owned by the studios in perpetuity.

Caitlin Parrish 12:56
Both demands are about making sure that our respective creative tools and instruments cannot be misappropriated and cannot be used without us. In a writers case, it's their words, it's their ideas. And in an actor's case, it's their face, their body, their voice, the emotional instrument that they bring to art. And but it just in both cases, it's it's it's saying we are not replaceable. And please don't treat us as though we are and please don't actively try to make our lives worse in the next few decades because it's more cost effective.

IM: I just wish we lived in a society that valued people and didn't just view them as a tool to enrich themselves. Yes it might save money, but what's the real cost of treating people as disposable. And on top of that, I'm sure that executives at these corporations are pulling in massive salaries, why not cut from that?

SW: Couldn't have said it better myself. And while I was talking to Caitlin about how language models like GPT worked, she brought up this really good point.

Caitlin Parrish 07:28
isn't the basis for AI generated scripts and images effectively plagiarism because it has learned from texts and images that other people have generated that humans have made to which they should have a copyright?

SW: We don’t think about holding copyrights over our tweets or chat logs on social media sites, but maybe we should?

Caitlin Parrish 07:44
so that regurgitation that recycling is coming from materials that humans have generated. So it shouldn't, in my opinion, count as wholly original art coming from Ai.

IM: That is a fascinating way to look at that. And aren't some writers suing Chat GPT over this same issue? Weren’t you ranting about AI companies scraping a fanfiction website the other day? Is this the same idea?

SW: Yes! Archive of Our Own, shorted to AO3, had to put out an official statement about ways they’re trying to stop auto data scraping from the site. Because it’s such a massive repository of free, written stories, that can easily be used to train AI models.

IM: Sounds like this issue around AI isn’t going to end with the WGA and SAG strikes.

SW: Definitely not.

IM: So let’s get back to the topic at hand.

Caitlin Parrish 20:02
I think it's important to keep repeating a story that came out in Deadline a few weeks ago that said a studio source said they wanted to make sure that writers and actors lost their homes and their apartments. And I just want to say that, that made it very personal. For us. It is impossible for it to not be personal for us. These people are saying that they want my three year old son to starve and be cold. There are a lot of writers who are very close to that situation already. If not already, in that situation. There are a lot of actors who are facing the same experience. And we are here to fight and we are here to stick it out because we are used to living on a knife's edge, and we are used to hustling and we are used to side gigs, I just want to say and say as many times as humanly possible, the cruelty is the point. They want it to hurt. And they're not being subtle about it. They're saying the quiet part loud.

SW: So if you want to support the strikes, you can donate to the Entertainment Community Fund. We’ll have a link in the show notes.

IM: And what if you are but a humble high school teacher with not a lot of funds to spare?

Caitlin Parrish 21:28
~~Although~~ if you happen to be based in New York and Los Angeles, we absolutely welcome you on the picket line. We've got plenty of extra signs and sunscreen and cold water, please come down and join us. We would love to have you we would love to feel your solidarity.

[Music starts]

SW: Don't forget to check out our show notes in the podcast description for more on natural language processing research and how to support the current WGA and SAG strikes.

IM: And if you like the show, give us a review on apple podcast or spotify or wherever you listen. By rating and reviewing the show, you really help us spread the word about Carry the Two so that other listeners can discover us.

SW: And for more on the math research being shared at IMSI, be sure to check us out online at our homepage: IMSI dot institute. We're also on twitter at IMSI underscore institute, as well as instagram at IMSI dot institute! That's IMSI, spelled I M S I.

IM: And do you have a burning math question? Maybe you have an idea for a story on how mathematics and statistics connect with the world around us. Send us an email with your idea!

SW: You can send your feedback, ideas, and more to sadiewit AT IMSI dot institute. That's S A D I E W I T at I M S I dot institute.

IM: We'd also like to thank our audio engineer, Tyler Damme for his production on the show. And music is from Blue Dot Sessions.

SW: Lastly, Carry the Two is made possible by the Institute for Mathematical and Statistical Innovation, located on the gorgeous campus of the University of Chicago. We are supported by the National Science Foundation and the University of Chicago.