Inventing the Future Generative AI and Medical Publishing – a Collaborative Conversation

Speakers: Jennifer Ghith & Yanshan Wang

Jennifer Ghith (00:00):

Hello, and welcome to InformED, a podcast series where you will hear industry experts share their thought-provoking insights and lessons in the field of medical communications. This series is brought to you by ISMPP, and is generously sponsored by MedThink SciCom.

Jennifer Ghith (00:15):

As a reminder, the views expressed in this recording are those of the individuals and do not necessarily reflect the opinions of ISMPP or the companies or institutions with which they are affiliated. This presentation is for informational purposes only, and is not intended as legal or regulatory advice.

Jennifer Ghith (00:34):

My name is Jenny Ghith, and I'll moderate the podcast discussion today. I am the Omnichannel Strategy and Innovations Lead for the Global Scientific Communications team for Pfizer Oncology. That's a mouthful, but for the purposes of this conversation, what is important to know about me, is that I've been working with AI to help us tame information overload via literature extraction and data mining for the last three-plus years.

Jennifer Ghith (01:02):

I'll also take this opportunity to introduce Dr. Yanshan Wang, the Vice Chair of Research and Assistant Professor at the University of Pittsburgh. His research focuses on AI and its applications in healthcare. Dr. Wang has led a number of NIH-funded projects aimed at developing AI algorithms to extract information from clinical notes and reports, and apply these learnings in many disease areas.

Jennifer Ghith (01:26):

He also actively serves on editorial boards for the likes of the Journal of Healthcare Informatics Research and many others. Dr. Wang, we're lucky to have you, and thank you so much for joining us.

Yanshan Wang (01:39):

Thank you so much for inviting me to this podcast, and thank you for the nice introduction.

Jennifer Ghith (01:45):

And again, we're so happy to have you. And today, we're going to focus on how large language models and generative AI are changing how we communicate, the importance of bridging dialogue amongst our specialties and professions, and how we can learn from AI experts, the tech industry, clinicians, the pharmaceutical industry, and what we might anticipate in the future.

Jennifer Ghith (02:06):

So, I'm just going to jump right in. Dr. Wang, these are interesting times that we're living in, to say the least. What's your perspective on the current generative AI landscape in healthcare and its impact on the communication of scientific information?

Yanshan Wang (02:22):

Thank you, Jenny. This is a great question. So, generative AI has been introduced to academia and industry for actually a long time. It could be traced back to the time when generative probabilistic models proposed in 2000, such as the Latent Dirichlet Allocation or the LDA model, which was designed to generate topics for a given text document.

Yanshan Wang (02:48):

So, these models were applied in healthcare to identify different clinical note topics. Nowadays, the generative AI is becoming a hot topic, again, mainly because of two applications. One is the ChatGPT for generating text, and I think everyone is using that.

Yanshan Wang (03:08):

And since its release in November, 2022 there are millions of users collected in just one month. And the other one is the text to image applications like Midjourney that image can be generated based on input prompt. And notably, ChatGPT has brought a lot of new potentials to healthcare domain.

Yanshan Wang (03:32):

So, just before our conversation today, I just searched the keyword ChatGPT in PubMed, and found 1068 publications. So, remember that this is a very impressive result because ChatGPT was just released nine months ago, because people in healthcare recognized the premise of generative AI within this field.

Yanshan Wang (03:55):

People find that generative AI could summarize, discharge summaries, could support clinical decisions, improve provider-patient communications by translating medical jargons to lay person language - make those administrative more efficient by streamlining those prior authorizations, or streamline those clinical paperwork, et cetera, et cetera.

Yanshan Wang (04:22):

There are a lot of applications right now, but these generative AI applications are not free of risks. One of the prominent challenges associated with generative AI is the so-called hallucinations. These models may generate content that is made up, that could be unreal, or sometimes, those content may have misinformation.

Yanshan Wang (04:46):

And these instances of hallucination may take various forms, ranging from some nonsensical text messages to images with unlikely features. And this hallucination content that appear convincing but in fact, they are fabricated. And this poses a significant concern in the critical domain like healthcare, where that misinformation may have serious outcomes.

Jennifer Ghith (05:20):

It feels like the world is changing under our feet. And it's very difficult to keep up and very difficult to distinguish the signal from the noise, perhaps, in what we're reading and what we're seeing. And the publication space was always so busy anyway. We have so many articles, for example.

Jennifer Ghith (05:38):

I think your PubMed search results are very interesting as well right now. I think that leads to another interesting question. How do you stay up-to-date when a new model is coming out every day when we are seeing so much information and potential misinformation out there?

Yanshan Wang (06:00):

Yeah, exactly. There's almost a new model coming out every day and staying up-to-date with latest generative AI models and the technologies could be really challenging. But I can talk a little bit about some of the strategies I use to stay updated. I hope it is helpful.

Yanshan Wang (06:20):

So, I'm a natural language processing AKA, NLP expert, and specifically NLP applications in the medical domain. To ensure that I stay up-to-date with new technologies, I actively engage with two academic communities.

Yanshan Wang (06:39):

The first one is the ACL community. So, ACL stands for Association of Computational Linguistics. It is NLP community primary composed of computer scientists specializing in NLP. Going to these conferences and reading most recent articles from ACL make me aware of the most recent challenges and the solutions in NLP, and the so-called large language models, which are the methodology behind ChatGPT.

Yanshan Wang (07:14):

And also, make me rethink about how these advancements in the general domain can be effectively utilized in the field of medicine. And there are several conferences in the ACL community. For example, there is a general ACL conference. There is a North America version of ACL, NACL conference happening either in Canada or the United States.

Yanshan Wang (07:43):

And there are also NLP focused on evaluation of NLP and a dataset. So, there's several conferences regarding ACL in the ACL community.

Yanshan Wang (07:56):

The second community is the medical informatics community like AMIA (American Medical Informatics Association), where healthcare professionals that apply NLP, apply AI in the data science, they usually go to the AMIA conferences. And there are three major AMIA conferences.

Yanshan Wang (08:18):

One is the AMIA Annual Symposium and that usually happens in November. And there is another one Informatics Summit in the spring. And then there's another one, Clinical Informatics Conference in the summer. So, those three are major AMIA conferences, I usually attend.

Yanshan Wang (08:35):

And I'm also very active in AMIA and in the AMIA community. And I currently lead the annual AMIA NLP Working Group, and also participate in a lot of leadership committees. So, participating in AMIA conferences and getting engaged with those activities and discussing with healthcare practitioners like physicians, clinicians or healthcare administrators, allows me to grasp the real-world problems in medicine, as well as people's concerns regarding the using generative AI models in their practice.

Yanshan Wang (09:16):

And I'm also very engaged with social media like LinkedIn and Twitter, where I actually learn a lot from your posts, Jenny, about generative AI.

Jennifer Ghith (09:28):

I started using it myself to find communities. So I think social media is a way for us to share information and share learnings with each other and bridge dialogue amongst our respective communities. I think it's really helpful for us to seek out others and seek out experts in the fields because these aren't necessarily communities or congresses that we all are historically interacting with. Would you agree?

Yanshan Wang (10:00):

Yeah, yeah, exactly. I think those social media platforms allow us to connect with our peers and connect with people who have a common interest in the area. And another thing I'm using those social media, is that it allows me to interact not only with academic people, because also, ACL conference, AMIA conference, most are academic people who attend those.

Yanshan Wang (10:26):

So, those social media platforms allow me to connect with industry experts like you and learn what does the industry needs and how these technologies with developing academia can be transferred to industry. I think now is really actually an era of team science.

Yanshan Wang (10:47):

I believe we really couldn't go far if we go alone by ourselves. If we just go alone by ourselves in academia, I don't think we will go very far or we will go very far. But applying those technologies in real-world scenario is important to collaborate with people from different disciplines, from industry, from medicine, from informatics, and then collaboratively navigate the landscape of generative AI in advancing the medicine.

Jennifer Ghith (11:20):

I think that's an excellent point. I think we should always look for ways to bridge dialogue amongst experts like yourself, amongst our industry, amongst clinicians. And I think that'll only help us to do more and to serve patients ultimately. We're also seeing a lot of noise and movement in the publications industry, of course. And the journals also seem to be issuing point of view statements on AI and the use of generative AI.

Jennifer Ghith (11:54):

We're also seeing some major journals launch sister journals in AI. That seems to be one other way that hopefully, we're going to see more bridging of dialogue and conversation across industry. Is there anything else we could do to help further and bridge that dialogue, do you think?

Yanshan Wang (12:14):

Yeah, I think there are several ways we could actually do better to bridge the dialogue among different experts. For example, we could organize more interdisciplinary conferences, workshops and seminars where experts from AI, computer science, clinical practice, and pharmaceutical companies can share insights, research findings, and also discuss challenges.

Yanshan Wang (12:41):

I think we didn't do a good job currently organizing those cross-discipline conferences. As I mentioned earlier, at the ACL conference, there are very, very few people from medicine who go to those ACL conferences. The majority of people going to ACL are computer scientists.

Yanshan Wang (13:00):

And for AMIA, there are also very few people in computer science and ACL going there. So, I think we could actually do better. Also, both venues have very, very few presences from the pharmaceutical industry, to be honest.

Jennifer Ghith (13:22):

I mean, I've noticed it too, in trying to bridge out and attend some of these other meetings that I don't typically go to. I think there's certainly more work to be done. I think we have every reason to be optimistic. Right now, we're at some sort of inflection point that is capturing the attention of the broader community.

Jennifer Ghith (13:39):

So, given all that, what is your view of the future? What partnerships and sorts of breakthroughs are you most excited about in the field?

Yanshan Wang (13:50):

I think first of all, the future of generative AI from medicine is bright and there's no doubt. And I'm particularly excited about the potential solutions, about several challenges we face right now when we apply generative AI in medicine.

Yanshan Wang (14:09):

I think the first one is of hallucination a problem I just mentioned earlier, because we have to, or we must address or at least control the hallucination problem or those misinformation problem of the current generative AI models, and make those models free of misinformation, or at least we need to control them before we apply them in practice.

Yanshan Wang (14:36):

And another challenge is reasoning. A lot of the time you get a response from let's say ChatGPT or those large language models, but you don't know what the reason is behind generating this response. So, this

reasoning challenge also needs to be addressed, and we need to make those models' answers much more transparent in terms of reasoning process.

Yanshan Wang (15:03):

And if we really want to see the vast applications of generative AI healthcare, those two challenges are the first two challenges we have to solve. And my lab is actually also actively working to address this challenge. A recent pre-print article from my lab, we propose a model called the ChatGPT Care, that integrates clinical guidelines into ChatGPT to improve the hallucination and the reasoning of the original ChatGPT.

Yanshan Wang (15:37):

We see some improvement, but I think it's still not enough if we really want to use that in practice. So, there is a lot more research work that needs to be done.

Yanshan Wang (15:50):

And the other thing is related to evaluation and evaluation standards for generative AI models. So, the conventional matrix used to evaluate machine learning models like precision we call F1 Score, or there's some conventional metrics to evaluate the generated response, like a BLEU Score. It is not the optimal matrix to evaluate generative AI models. So, I think we need to develop a new matrix for evaluating those generative AI models and to evaluate those models as well.

Yanshan Wang (16:17):

And the third one I think is the discussion I'm really excited about, is around ethical principles before applying generative AI in practice. So, for example, generative AI makes the health disparity vast because of the model bias.

Yanshan Wang (16:49):

And however, those ethical principles for generative AI in healthcare have been understudied. And a lot of healthcare decision makers often fail to consider the significance of those ethical principles before applying them in practice.

Yanshan Wang (17:07):

And we have to work together with different people to address those problems. And these challenges are real. They are real when we apply those models. But I'm also excited about solving these challenges together through collaborations and partnerships.

Jennifer Ghith (17:26):

I think it's a bit of a double-edged sword because what we're seeing is the ability to ask questions and get responses that are really understandable, particularly in domains like clarity, for example. But the question is, are these biases going to be exacerbated for people who don't have access to these systems? What happens to them?

Jennifer Ghith (17:46):

But then, I think you're also talking about the data that we ingest into these systems. If it is inherently biased, then the outputs will be biased as well. So, we definitely have a lot of work ahead of us. I think,

again, these are very interesting times. We are at such an inflection point in the technology. And I personally look very much forward to continuing our conversations together because I learn a lot every time I speak with you.

Jennifer Ghith (18:14):

So, thank you.

Yanshan Wang (18:15):

Thank you, Jenny.

Jennifer Ghith (18:16):

And thank you so much for this conversation. We are out of time for today.

Jennifer Ghith (18:20):

For other ISMPP resources on AI, please listen to our earlier podcast chatting more about ChatGPT and generative AI, and visit the ISMPP University and e-learning site for our webinars: The Rise of ChatGPT and Generative AI and How AI can Increase Accessibility. Links can be found in the written transcript of this episode. https://www.pathlms.com/ismpp/courses/46881

Jennifer Ghith (18:43):

Thanks for listening to InformED for medical communications professionals. Please take a minute to subscribe to the show on your favorite podcast app, inform your colleagues, and rate our show highly if you liked what you heard today.

Jennifer Ghith (18:55):

We hope you'll also join us at an upcoming ISMPP U Webinar or even consider becoming a member of our association. Just go to ismpp.org to learn more. I'm Jenny Ghith.