

# PLUS Podcast: Demystifying AI

## Episode 1

[00:00:00]

**PLUS Staff:** Welcome to this PLUS podcast, Demystifying AI, Episode One. Before we get started, we would like to remind everyone that the information and opinions expressed by our speakers today are their own, and do not necessarily represent the views of their employers, or of PLUS. The contents of these materials may not be relied upon as legal advice. And with that, I'd like to turn it over to Sarah Coutts.

**Sarah Coutts:** Thanks, Tyla, for the introduction. Today, we'll be speaking to Jaymin Kim, Senior Vice President, Emerging Technologies at Marsh, and Dr. Joerg Storm, founder of Digital Storm Weekly, an AI newsletter that is received by over 300,000 readers at companies including Apple, Google, and Microsoft, educating them on AI and tech.

So, this is the first in a series of podcasts covering the ever-present topic of artificial intelligence. While AI is not new, it has exploded into our lives in the last 18 months with the pace of our implementation and the use of [00:01:00] generative AI. In fact, given the acceleration and use of AI, it may not be a stretch to say that Gen AI already features in most news releases, industry papers, client discussions, and strategic planning in your working life. This first podcast is our attempt to give you the basics behind artificial intelligence, provide answers to those questions that maybe you think you should already know the answers to, and of course, give it some insurance context so that you can feel more comfortable talking about it in your workplace and speaking about it with clients and colleagues.

To start off, I'll ask our speakers to give us a summary in their own words of what is AI and how is it developed.

**Joerg Storm:** Yeah, AI refers to the development of systems that can exhibit intelligent behavior like learning, problem solving, and decision making. Its roots, I think back to the 1950s with early pioneers like Alan Turing laying on the groundwork. And over time, [00:02:00] AI has evolved through various approaches. One approach is a rule-based system. Those rely on preprogrammed rules to tackle specific tasks. The second one would be machine

learning, or abbreviated, ML. ML algorithms learn from data, enabling them to improve performance over time.

And the third one would be deep learning, which is a subset of ML using complex neural networks mimicking the human brain for tasks like image recognition and natural language processing.

**Sarah Coutts:** Okay, so while many people may think that AI is a new phenomenon, as Joerg has just mentioned, its roots go way back to the 1950s.

Jaymin, can you give us some more detail about what developments there have been between then and now that has caused AI to explode into our lives recently?

**Jaymin Kim:** Yeah, so AI is a really broad and diverse field that [00:03:00] conceptually speaking, it refers to computational systems that are intended to simulate specific aspects of human intelligence.

We typically think about this broad and diverse field into two subcategories. The first is narrow AI. So, it's worth mentioning that all AI systems that exist so far fall under narrow AI, which refers to computational systems that have been designed to perform specific tasks. In a given limited domain of human intelligence, like text or image, narrow AI can be thought of as the application of various techniques that have existed since the 1950s as you are mentioned across specific limited domains of human intelligence.

In the 1950s, we had more arguably simplistic knowledge-based systems, but over the last recent decades, AI techniques have become significantly more advanced [00:04:00] and have led into the field of deep learning. In deep learning techniques, we see the employment of deep neural networks that, as Joerg mentioned, were inspired by the human brain.

So, when we hear about generative AI models in the media, like ChatGPT, or Gemini, or MidJourney, we're talking about a relatively new subset of narrow AI models. General AI, or Artificial General Intelligence, AGI as it's called, is a theoretical construct that doesn't yet exist and refers to a second subcategory belonging under the broad umbrella of AI.

AGI, or General AI, is what companies like OpenAI hope to create one day, and so represents that North Star for the industry. AGI, in theory, would be systems that broadly reflect human intelligence and beyond, potentially, and aren't necessarily limited to any one specific domain, like text or [00:05:00] image,

and can in theory self-learn into venturing into other domains of human intelligence.

**Sarah Coutts:** Okay, so while we're still trying to get our heads around sort of AI and generative AI, there is something even more complex, albeit at this stage it's theoretical. So, before we delve into maybe the theoretical, let's just stick with say generative AI, which has really been the driver for AI becoming the hot topic or only topic of discussion in the insurance industry, maybe every industry in the last year or so.

So again, just to Jaymin, I mean, from your perspective, what do we need to know about the difference between generative AI and traditional AI?

**Jaymin Kim:** So generative AI models are differentiated from other forms of AI that for our purposes, all referred to as traditional AI by their ability to employ deep learning models and generate new original content across various [00:06:00] domains of human intelligence.

Previously, we didn't really have AI models that were creating original new content historically thought of as a human capability. From clients in the insurance sector, I often hear the question, "if AI's been around since the 1950s, why has the last few years suddenly become the years of AI?" And I found that in 99% of cases, when folks reach out to me to talk about AI, they in fact want to talk about generative AI.

I think the reason why the world is both excited and fearful of generative AI comes down to three key ways that generative AI differs from preexisting more traditional forms of AI. So, the first is in the ability to generate AI models to create content previously existing, more traditional forms of AI were primarily used for classification or predictive tasks.

Whereas now with generative AI, we [00:07:00] have models that are simulating previously, what was thought of as a human capability to create original new content across text, video, audio and images. The 2nd way that they differ, though, is that whereas traditional AI models learn to predict outcomes based on their input training data.

And then we'd project into what that means for another scenario. Advanced generative AI models actually go beyond just learning the data points in that training data and go so far as to learn the underlying distribution, the statistical distribution, in between all of the data points in that original training dataset.

In so doing, it's inferring new patterns that lie in between all of the data points in that training data set. And that's how these generative AI models are then able to produce new content that is similar to, and yet entirely distinct from, the data that they were initially trained on. [00:08:00] But I think the third way that generative AI differs from traditional AI might be the most important reason why the world can't stop talking about generative AI these days.

I know personally, I can't have a dinner conversation where AI doesn't come up as a topic, and that's how we as humans are interacting with AI. So previously with traditional AI models, I would describe my interactions is more of a one way street. I say do X and I expect it to execute and not talk back to me.

Today, with generative AI models like ChatGPT, my interactions with these models are a two-way street. When I use ChatGPT 4.0, for example, I don't even need to give it a specific set of instructions. I can simply decide to start up a conversation and the model will reply back to me in a very believable, human-like way.

**Sarah Coutts:** Thanks, Jaymin. Admittedly, I haven't quite [00:09:00] gotten to the stage of chatting with ChatGPT myself, but I would love to know the quality of the conversation and how human it is and whether it's coming close to being able to fool us, so to speak. But before we go down that path, let's just consider the implementation of generative AI, as everyone seems to be scrambling to understand it and start using it within their businesses, although they may have some reservations in relation to privacy or security and accuracy.

Joerg, in your view, who is currently using generative AI and who isn't?

**Joerg Storm:** Yeah, good question. I think the AI adoption is really widespread. For example, some tech giants like Google, Amazon, Meta and others leverage AI for optimizing the search algorithm, for product recommendations and also personalized advertising healthcare providers use AI in medical diagnosis, drug discovery, [00:10:00] and also personalized treatment plans.

Finance uses AI to detect fraud, manage risks, and also personalize financial products. And for example, manufacturing uses AI to optimize the production process to predict equipment failure, and also improve product quality. There are some companies which maybe don't use AI yet, and those are, for example, companies which are smaller companies, smaller business, which have limited resources, limited know how which might hinder the AI implementation, or highly regulated industries, for example, like banking regulations, so strict

regulations could delay the AI adoption in sectors like finance or healthcare, and third, but not least, I would say, lack of awareness.

Some businesses, or some owners might not [00:11:00] understand yet the benefits of AI or have misconceptions.

**Sarah Coutts:** Absolutely, because as exciting as it sounds and despite the many benefits it could bring, of course, there is some reluctance to utilize generative AI before fully understanding its risks. Jaymin, Joerg has mentioned large industries like tech and finance and healthcare who are using Gen AI.

However, another point to note is the vast level of engagement the average person has with Gen AI outside of their working lives. Would you agree?

**Jaymin Kim:** Yes, so technically anyone with access to the internet and a computer or mobile device can use generative AI. I think this is a really important point, because for the first time in history, the average retail customer, folks like you and I, are interacting with machines that are seemingly and believably human like in what they're able to do. I'd also mention that this is also one of those [00:12:00] rare moments in history where technology became retail grade first, before it became commercial grade.

Meaning generative AI made its way into the hands of the retail individual before corporations and organizations had a chance to figure out how to use the technology and what it means for the enterprise. I think the speed of adoption of generative AI technology has been nothing short of remarkable.

An interesting data point is that in its first month, ChatGPT acquired more than 50 million monthly active users as a point of comparison. It took TikTok 9 months approximately to achieve the same. And at the enterprise level, while it is still a nascent technology, generative AI is being used by early mover companies across numerous industries.

If we dive a little bit deeper into what kinds of organizations are using generative [00:13:00] AI in financial services, we're seeing the likes of Bloomberg, GPC. Which conducts sentiment analysis on financial data to generative AI enabled chatbots that can equip wealth managers to better serve their clients by more efficiently drawing on massive volumes of company proprietary data and the health care sector.

We're seeing Gen AI systems help surgeons and doctors more quickly answer medical questions, and they could have done previously by going on dense

medical texts. They're also using generative AI systems to help with generating more advanced and more precise differential diagnoses to aid doctors and surgeons and making the final recommendation and the entertainment sector receiving companies use advanced models to create end to end video scripts, commercializing them in production, all the way to creating original music and retail generative is being tested as drive through operators.

Real estate and the hospitality is yet another sector where Gen [00:14:00] AI is helping virtualize staging for properties and also hyper personalized travel itineraries for the end customer.

**Sarah Coutts:** Thanks. So, as we thought, Gen AI take up by normal average everyday people, as we said, has been nothing short of phenomenal, however, it is also commercial industries who are diving headfirst into implementing it into their businesses. And we hear that it can provide some great commercial benefits and increased profitability. Joerg, can you flag some of the key opportunities that Gen AI offers commercial entities?

**Joerg Storm:** Yeah, sure. For example, in the innovation area that AI can automate tasks analyze vast data sets, and also accelerate scientific breakthroughs, for example, in finding new treatments. Commercial benefits is another area where AI can help. I can optimize operations personalized customer experience, as we heard before, and also develop new products [00:15:00] and completely new services.

Another opportunity I think which we get with AI is increased profitability because AI can reduce costs, improve efficiency, and also generate completely new revenue streams.

**Sarah Coutts:** Thanks. Many potential benefits there, but for many, they are somewhat overshadowed by the potential risks both known and unknown.

Jaymin, can you provide some insight into what you are doing at Marsh to help clients to identify these risks, particularly from an insurance perspective?

**Jaymin Kim:** So, at Marsh we've built a risk framework specific to generative AI to help our clients assess what's net new here, if anything, from a risk and insurance perspective, as we've discussed already, AI is not new, and really, the focus is generative AI. This is important as a question to address, because I find that with new technologies, like generative AI, sometimes there can [00:16:00] be a lot of new buzzwords and scary sounding terms. But I think as an insurance industry, we need to look under the hood and separate out what's real and what's

hype. Often, I see that people think, if there's a new technology with new buzzwords and terms around it, that this must mean there's completely new risks and therefore new insurance solutions that need to come into play, new insurance solutions that are needed. That's not necessarily true, although it could be true.

And when it comes to generative AI, what we're finding is that many of the risks associated with Gen AI are actually just extensions of existing and familiar risk categories, like data, privacy and copyright infringement or technological error, technology misuse to really understand the risks that come with developing and deploying generative AI.

I think it's important to recognize that there are some key [00:17:00] components that create what generative AI systems are and then some corresponding nuances to how we need to assess existing risk categories and the impact of generative AI on those at risk categories. So, the first set of risks lie around data.

And when I think about generative AI, I think there's three sets of important data that organizations need to be thinking about and corresponding risks. The first is around the training data set. So, Gen AI systems are typically trained on just massive volumes of big data, and here it's important for organizations to know that the quality and the scale of that original training data set can have a significant impact on how the model performs.

Today, there is various ongoing lawsuits around alleged copyright infringement, and the courts are still deciding whether it's legal to include copyrighted data in training Gen AI models [00:18:00] without necessarily having sole appropriate permission from respective authors. So, copyright infringement risks come to play when we think about developing and deploying Gen AI models.

But in the meantime, if you have data publicly available online as an organization, and that data is proprietary or is copyrighted, there is a risk that your copyrighted data may then be used to train Gen AI models. And so, organizations need to be thinking about what kinds of risk mitigation controls they can put in place in order to mitigate against such risks. One potential solution concerns data encryption. Second, there is a prompt data, which is the data that the end user puts into that search box looking screen where we kickstart our interaction with the Gen AI model.

It's important to understand here that the data that end users put into that box can lead to various privacy and security concerns, but it will depend on the



instance of the generative [00:19:00] AI model we're talking about. For example, the risks that come with using the publicly available version of ChatGPT are not quite the same as using an enterprise specific instance that has its own set of data privacy and security controls wrapped around it.

And then third, there's the output data itself coming back to the whole point of generative AI models, which is to create original new content. A lot of media headlines today are talking about how generative AI models are lying to us. And what they really mean is that these models are hallucinating or confabulating. Hallucinations refer to just when a Gen AI model spits out error or nonsense. And this tends to happen when the Gen AI system is confronted with a scenario that isn't grounded in its training data. Currently there is no foolproof mitigation against hallucinations.

And so, if your organization is using Gen AI models and customer facing settings, [00:20:00] hallucinations can mean that your organization is presented with various liabilities, including reputational damage. There's also the risk associated with Gen AI models potentially putting out discriminatory or biased outputs.

And even if your Gen AI model comes with default safety settings that are in fact jailbreaking methods where end users intentionally try to make the model deviate from the safeguards that have been built into said model. In other cases, the human end user might inadvertently prompt the Gen AI model in a way that gets the model to reveal confidential data that the company may not have wanted the model to reveal. Beyond the data related risks, though, there is a model itself with its underlying architectures and algorithms, and then there's a training methodology by which the Gen AI model will take in and quote, unquote, learn the training data set. Here, it's [00:21:00] important to know that these models are based on deep learning techniques and inherently are complex and nonlinear.

Gen AI models come with a lack of 100 percent explainability, where the experts that create these models themselves are not always able to explain to us why you get X, Y, Z outputs based on 1, 2, 3 inputs and the model design. And so sometimes what we see happen is General AI models developing emergent capabilities.

Which is when a Gen AI model starts showcasing the capabilities that the experts didn't intend or program the model to have. So, one example of this might be a Gen AI model that's been trained on massive volumes of English text based data.



And one day we realized that the same model is also able to write and review programming languages. Then, of course, there's the fact that GenAI relies on underlying software and hardware infrastructure, [00:22:00] which means that companies that are adopting generative AI for enterprise use cases will be susceptible to various counterparty risks.

For example, relying on the consistent functioning of cloud infrastructure on which typically the massive volumes of training data sets are stored. I'll just take a brief step back and comment that, every step of the way Gen AI technology comes with a potential of errors and various risks, simply because this is technology that humans have created for use by other humans.

And of course, as we all know, humans are biased and fallible from curating the training data to developing algorithms, to creating model guardrails, assessing model performance. There's a risk that something might go wrong every step of the way, but there is a silver lining, which is that because the humans are still in charge, there are various risk management controls ranging from technical to people and process [00:23:00] controls that we can implement proactively in order to mitigate against potential risks and to also drive clarity regarding the model's likelihood to be safe and secure.

**Sarah Coutts:** Thanks, Jaymin. And that's really helpful explanation of the number of the key risks with Gen AI. And one of which stands out, of course, is that the technology is created by humans and therefore can be biased and fallible like humans. Also good to know the humans are still in control, of course. Joerg, you mentioned this as a key risk as well.

Do you have some more comments in that regard?

**Joerg Storm:** Yes, sure. Healthcare organizations and the data privacy issues also topics AI has besides those was other risk. The human apprehension, the fear of losses lack of control and also ethical concerns can create a resistance of employees towards AI adoption. Another fear would be of job losses due to the automation [00:24:00] through AI. It could display some workforces, also require retraining and adaption of the workforce. And I think this was mentioned before, discriminatory algorithms AI algorithms, which have been trained on biased data can also perpetuate discrimination.

This we saw already that something like this happened.

**Sarah Coutts:** Absolutely. I think some of those are really key areas of concern for many people in workforces right now, particularly as there have been

numerous articles about potential mass job losses, which could occur even in the near future due to the implementation of Gen AI.

Now, while there's no assurances around the job situation, governments, of course, have recognized the need for regulation and have been collaborating on how to do that. The EU arguably led the way in formalizing AI specific regulations with the EU's Artificial Intelligence Act, which came into force last year.

Now, [00:25:00] notably, it takes a principle rather than rules-based approach to regulating AI. Joerg, what is your view on this first tranche of AI regulation?

**Joerg Storm:** This regulation, like the European Union Artificial Intelligence Act, aims to ensure responsible development and the use of AI. And I think there are two sides.

On the one hand, it can help AI and promote public trust, transparency, and also support the ethical development of AI. On the other hand, it can hinder AI due to strict regulations, and this could slow down the innovation and also stifle development in certain regions. So, the ultimate impact, I think, will depend on how regulations are implemented and how well they are balanced the innovation with the safety aspect.

**Sarah Coutts:** Yeah, absolutely. Jaymin, what's your view on the impact of these emerging regulatory frameworks?

**Jaymin Kim:** So, we are certainly [00:26:00] in nascent waters with the emerging regulatory environment. And I think one thing we can say safely is that we expect new regulations to emerge specific to AI in the coming years.

And this may mean new reporting requirements or expenses for some organizations. Arguably, the European Union's act is the most comprehensive AI specific regulation to date and we'll see how the implementation and execution and enforcement of the act pans out, but while we're in these nascent regulatory waters, I believe that contractual liability will increasingly play an important role.

While we wait to see how regulations are applied and enforced at the end of the day, some stakeholder will stand to financially benefit from developing and or deploying these generative AI models. And depending on whether your organization [00:27:00] is. The developer of models, or maybe a third party systems integrator, perhaps, or the end user of an off the shelf Gen AI model, or

your organization might be potentially further fine tuning a pre-trained model for your enterprise specific use case, regardless of who you are in that.

In that value chain, it's going to be important to be very clear about where you stand in that broader stakeholder ecosystem and what your responsibilities are going to be not just in terms of what you are providing to your customers, but also how you're interacting with vendors and potential downstream and upstream impacts of belonging to the interconnected AI ecosystem.

**Sarah Coutts:** Thanks. Thanks, Jaymin. Time is almost up on our discussion today already, which of course has been so insightful. But before we go, [00:28:00] we have some time for horizon scanning Joerg, you first. What are your thoughts on the next big developments on AI?

**Joerg Storm:** There are some new areas upcoming. One is, for example, the area of explainable AI making eye models much more transparent and understand the loss of a human.

This is, at the moment, not always the case. The other option would be the norm of a computing designing computer chips inspired by human brains for fast and more efficiently AI processing. And third option would be human AI collaboration. Exploring how humans in a I can work together effectively, and they're already some products out here in the market, which show that.

And I think Damon has previously mentioned the topic artificial general intelligence abbreviated AGI and AGI is also known as strong AI and [00:29:00] this refers to machines with human level intelligent. While this is still in the distant future. Some say it's 3 to 10 years. I'm not sure about this.

The advancement in reasoning and planning could definitely pave the way for AGI and there are a lot of opportunities, and one opportunity would be, for example, that AGI could solve this problem. complex problems. AGI could tackle challenges like climate change or pandemics. AGI could support scientific breakthroughs by accelerating scientific discovery and innovation.

On the other hand, AGI has quite some risks. Worst case would be really existential threat to human mankind. And some experts warned already that uncontrolled AGI could pose a threat to humanity. And also, job displacement would be a major threat as AGI could [00:30:00] potentially automate almost all of our jobs.

So, the development of AGI needs really careful consideration of both those potential benefits I mentioned and also all those risks.

**Sarah Coutts:** Thanks, Joerg. Great opportunities, but yeah, despite the risks that some may have about Gen AI many do believe that it can hold the key, of course, to solving some of the world's most complex problems.

I should say AGI, such as climate change. Jamin, on the horizon scanning front, I know that you've recently written an article about responsible AI as the question of how we balance innovation with protecting the public interest is a recurring theme. So, a key issue is how we make sure that we continue to distinguish, or be able to distinguish, between human and AI content and interactions.

So, in a final few words, are you able to explain what's being done to address this?

**Jaymin Kim:** Sure, sure. So, while there's no unified movement responsible [00:31:00] AI generally refers to developing and deploying AI in line with principles like fairness and transparency, security, privacy and reliability. And I would say that emerging regulations across most jurisdictions broadly reflect these responsible AI principles. One top of mind AI risk. For many organizations today, I know, revolves around the theme of transparency how to distinguish, for example, between human and AI content and between human to human versus human to AI interactions. One best practice that I think organizations can adopt here is to proactively inform their customers when the organization is deploying AI as opposed to a human.

For direct customer facing interactions, according to this year's World Economic Forum's global risks report, AI generated misinformation and disinformation was identified as the most significant risk that organizations face today as a countermeasure to [00:32:00] misinformation and disinformation. There has been considerable regulatory focus on content authentication technologies.

For example, in the US. The executive order on artificial intelligence calls for the need to make it easy for Americans to know that the communications they received from their government are authentic. And in the EU, the AI act specifies the need to detect and label AI generated content to enable the public to distinguish AI generated content effectively.

We don't actually have foolproof content authentication technologies yet, and this is particularly true in the domain of AI generated text. My view is that even if we were to develop foolproof content authentication tech one day, our conception of what AI risk management controls can do actually needs to change.

One example as to why I think this, it pertains to startups that are already providing AI platforms where humans can basically [00:33:00] toggle among character traits and features to build their ideal friend, or relationship from, with the ideal girlfriend, to the tutor, to companions to ward off loneliness among senior citizens.

And when I think about these emerging AI use cases, even if we know that we're dealing with an AI, over time, we may not care and create our interactions with AI as though we're interacting with humans, and in fact form emotional human-esque relationships with AI. This will, over time, likely pave the way for AI to become a new vector of manipulation, misinformation, and disinformation, because we're entrusting our emotions over to AI.

The bottom line is that even if we were to follow every principle and check off everything that's laid out in responsible AI frameworks, there is no end state at which we can or should declare AI systems as [00:34:00] trustworthy because AI risks aren't just about AI. In fact, they're equally, if not more, about human risks.

And so, what that means is we need to change our conception of AI risk management being sufficient. It's necessary, but the most we can do is thereafter consciously monitor for and manage risks at all times on an ongoing manner. Finally, I do think that we need to take a broader view of how various technologies are converging.

Beyond just AI, to understand the bigger picture of AI and the risks it could pose. Personally, I see AI as just one piece of the bigger technological picture that's changing the risk landscape for organizations and retail individuals alike. And to comprehensively address AI risks that are emerging on the horizon, we need to think longer term about what AI is likely to look in the future.

Considering the convergence already happening across various [00:35:00] other technologies, including augmented reality, haptics, holograms, and brain computer interface technology. Although perhaps this might be a broader topic for another time.

**Sarah Coutts:** Well it might well be, Jaymin. I can imagine we could need quite a few more podcasts to properly delve into all those technologies of the future.

But great to hear about and both of you have given us so much food for thought. For now though, I would like to thank Jaymin and Joerg for such a great introduction to AI for all of our listeners. And as I said, this is one in a series of podcasts on AI and our next podcast we will be joined by underwriters in the London market who can provide some insight into how they assess risks posed by AI and what products are being developed to address those risks.

Thank you to PLUS and thank you to all of our listeners.

**PLUS Staff:** Thank you for listening to this PLUS podcast. If you have ideas for a future PLUS podcast, please [00:36:00] complete the content idea form on the PLUS website.