

# Will AI Robots Take Over Our Jobs Soon?

*“The most nightmare scenario I can imagine with AI and robotics is a world where robots have become so powerful that they are able to control and manipulate humans without their knowledge,” [says AMECA](#) - the humanoid robot. ( get this footage from the link )*

Did you know that AI could automate up to [30%](#) of jobs by mid-2030s? It may sound a bit further down the road, **doesn't it?** ( using my animated character to stress the 'doesn't it' )

A recent study found that about **50% of tasks could be at least partially automated with large language models, even in 2023.**

Elon Musk even warned us in an interview that AI could lead to [“civilization destruction,”](#)  
**Does this mean AI robots will soon take over the world?**

I have a theory, well, a theory based on some scientific facts! Remember three keywords likely to stick with you for the next 10 years: multimodal AI, , and [Quantum AI](#).

, you'll get what I was talking about by the end of this video. So, Why don't we start with 'multimodal AI'? -

## Multimodal AI

With the grand entrance of ChatGPT in 2022, ed to see the power of generative AI. But in , ChatGPT is not only the most popular chatbot on the planet, it can also [see, hear and even speak](#).

( [Play footage in the link](#) ) That's what I meant by that -

Not to mention in 2024, chatGPT [is also part of the 3D spatial computing world](#) as they caught up with the hype of Apple Vision Pro and released their first version app for visionOS.

-

But it doesn't stop there, just last week, OpenAI released their text-to-video model **Sora**. It's kind of amazing how it turned out.

If you've been watching our channel, you know we have been [leveraging text-to-video for some of our videos](#).

You know [some AI products](#) are literally playing the notion of “text-to-video” by combining ‘text-to-speech’ and ‘talking pictures’. Some [text-to-video models are just jokes](#) ( use some bad text to video examples )... the one from [Runaway is actually good](#). But ...

I've been “begging” for this kind of accuracy and realism like Sora! Finally, there is one real text-to-video model that's at the next level. Look at those, the texture and lighting, it'll practically replace my stock footage library. It's still not perfect, but if you compare it to what

we had before, that's just amazing! I see myself using this for *my virtual production set* in the future, as long as I'm not getting into any copyright infringements -

( Playing footage [on this page](#) of generated videos by Sora )

Sora is a diffusion model, that is capable of generating an entire video all at once or extending pictures or in some cases extending a generated video to a longer duration. [It builds on top of past research on DALL-E and GPT models. It is similar to GPT models though. Sora uses a transformer architecture, but with an incredible scaling performance according to OpenAI.](#)

We'll [link the technical paper](#) in the description down below and make another video to demystify this, so stay tuned, and you know what to do! :)

Google [introduced its own multimodal AI Google Gemini](#). It [comes in three sizes](#). I have things to say about this, but i'll save it for later.

The smallest is called Gemini Nano, it is designed for edge devices, such as mobile phones serving as the go-to for Android developers on devices like [Samsung Edge and Google Pixel 8 Pro](#).

Then Gemini Pro is the current default version for [their chatbot Bard](#), it's capable of advanced reasoning and planning. Google plans to introduce a more intelligent version called "[Bard Advanced](#)," which of course powered by Gemini Ultra, coming in early 2024.

( time slip – pause )

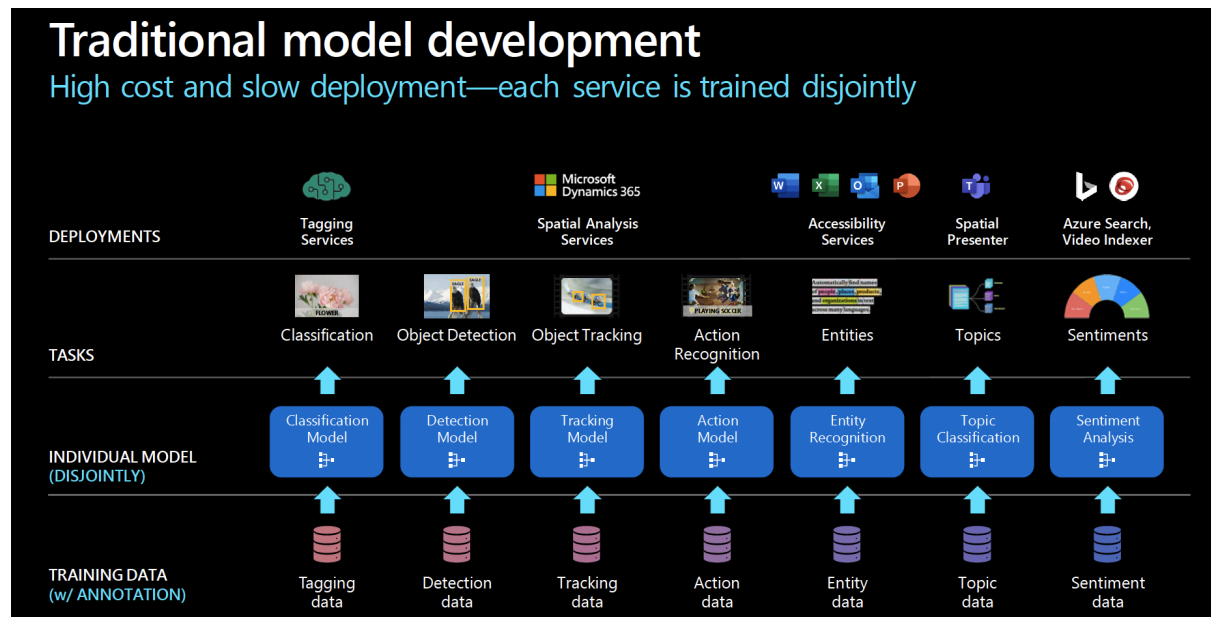
Wait a second, it is now! Not only is Gemini Ultra coming to Bard, it literally means 'Bard is gone', instead, we have [Gemini Advanced](#).

You see Google is not only promoting Gemini ultra as their largest and most capable multimodal AI model, but it also replaces DuetAI for Google workspace and Google Cloud which was announced like last year. Google says : [Gemini will help companies boost productivity, developers code faster, and organizations to protect themselves from cyber attacks, along with countless other benefits](#). This marks a new era - google's Gemini era.

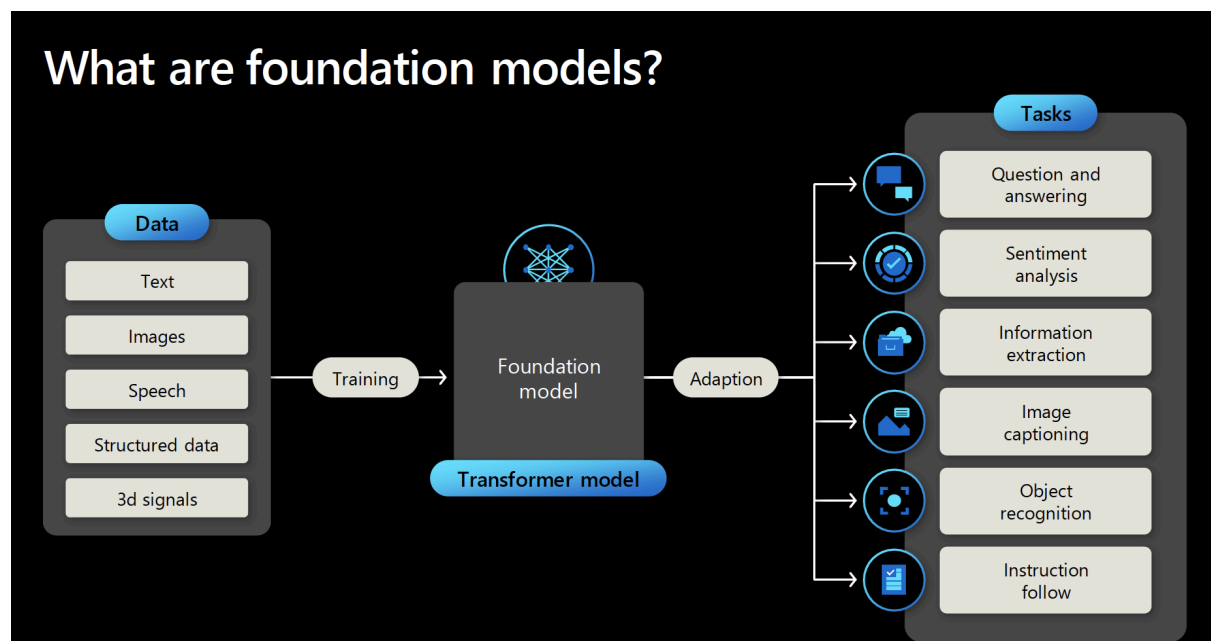
Just imagine when Gemini is used in Google Search Experience. Yeah, I have tested the new search experience a while back, it looks pretty cool, and I liked [the conversational style of results when it comes to answering search queries](#), [it would work even better when used for searching and shopping items on random e-commerce websites](#). -

About now, you may be wondering what does Apple do with GenAI. Well since Apple announced to [spend \\$1 billion per year to catch up genAI boom](#) in 2023. Apple has joined forces with researchers from Columbia University to create a multimodal AI, [Ferret](#). [Ferret](#) is like a new version of ChatGPT .. in 2022 .. Thank god we're in 2024 now, it stands to reason it has computer vision capabilities integrated into it. [Ferret](#) understands [connections between objects, actions, and details in context](#) looks good as an early version, but [It's just too early to tell if Ferret is better than GPT4 or Gemini](#) -

In the past, traditional model development focused on training for specific use cases, such as training a topic classification model to classify different topics and training a sentiment analysis model to understand sentiments. ( as in the following image ). -



Foundation models, however, are trained on diverse datasets including text, images, audio, etc. This comprehensive training enables them to understand and generate various types of data across different contexts and tasks. Large language models (LLMs) are a type of foundation model specifically tailored for understanding and generating human-like language. Additionally, other foundation models like DALL-E and Sora have unique abilities: DALL-E can understand and generate images, while Sora can create vivid scenes based on text prompts.



Much like the definition of multimodal AI. It can process information from various sources like text, images, audio, and video, integrating data from different modalities allows these systems to better understand and analyze complex real-world situations. This leads to more accurate and versatile decision-making and interaction capabilities.

Well, when it comes to evaluating the capabilities of these multimodal AI though. The results may differ. Despite many claiming Gemini is a superior AI model over GPT-4, and their well-crafted demo video.

Official data from Microsoft Research using MedPrompt+ paints a different picture. GPT-4 demonstrated [higher proficiency with a performance of 90.10%, compared to 90.04%](#) for Gemini Ultra, they literally only have a slight margin of 0.06% -

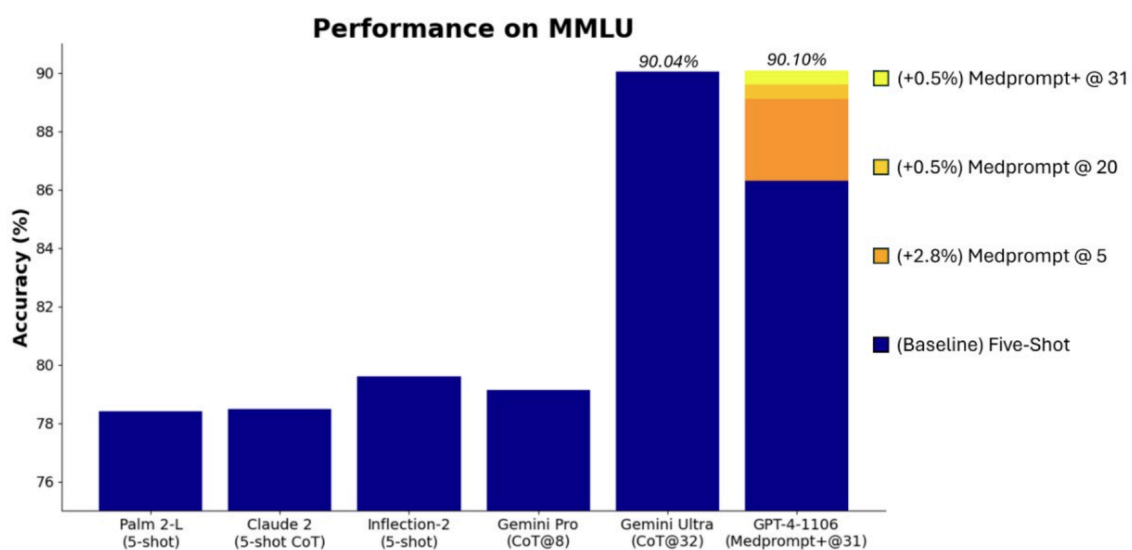


Figure1. Reported performance of multiple models and methods on the MMLU benchmark.

— add start—

Wait a second, only a week later, after Gemini Advanced came along, [Google launched Gemini 1.5, which brings improvements](#). The exciting part is this model achieves a similar quality to 1.0 Ultra but with less computing power.

What's really eye-catching, though, is the experimental one million multimodal token context window.

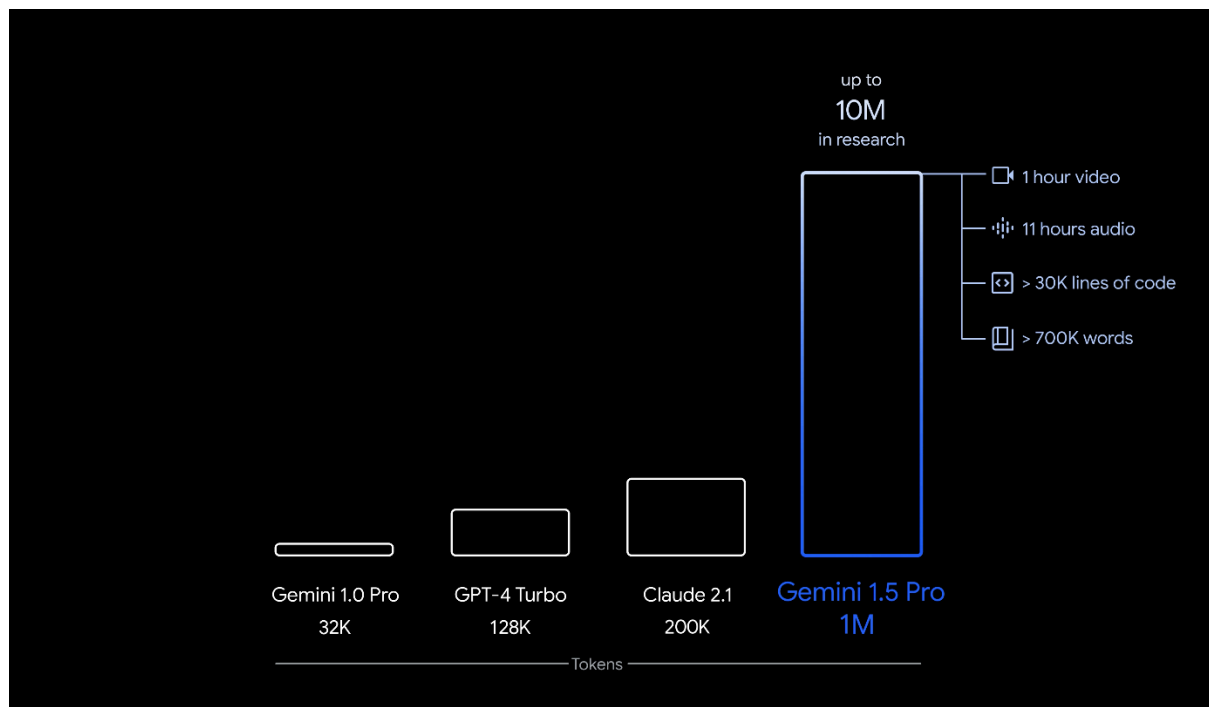
What does "one million multimodal tokens" even mean, though?

Firstly we should talk about: What is a long context window?

Context windows allow AI models to retain lots of information while they work by utilizing extensive context during their computations. It's like when you forget somebody's name during a conversation or rush to write down a phone number you just heard.

And the 1M multimodal tokens , on the other hand, mean :  
1 hour video

11 hours audio  
>30K lines of code  
And > 700K words



So basically, the type of input that we can use ranges from a simple prompt with a structured sentence of a few words to the next level, for example:

- [A 400 pages of PDF](#)
- [a 44-minute movie](#)
- and [even 100K lines of code](#).

Also, the output is pretty amazing, I would say, It is a model that's genuinely needed, and we can indeed get what we need from it.

If you want to learn AI beyond those buzzwords, you may [check out this video on our channel](#), which helps you understand those overwhelming concepts in a simple manner with real-life examples.

## Robotics

With the power of multimodal AI, the next natural step is to give the machine some sort of physical form. That's where robotics comes into play. -

You may have seen some robots who're doing quite good dance moves to classic R&B. Or some [humanoid Robot who can do backflips](#) and even [the wildest doggie robot on the](#)

[internet](#) . They all come from from Boston Dynamic AI Institute in Cambridge, Massachusetts.

The one's doing backflips, called [Atlas](#). With [stereo vision and various sensors](#), Atlas can grab and move things around and since it has the ability to [sense obstacles on its own](#), is good at navigating tricky landscapes. What's impressive is that it can pick itself back up and stay balanced if it gets knocked over, so it can potentially reach somewhere that can't be reached by humans.-

In the world of robots, there's a big question pops up: should we design robots to resemble humans or focus on their specialization for tasks beyond human capabilities ?

Some think they should look human to fit in better, while others say they should be built for specific jobs. Companies like [Tesla prefer human-like robots for a](#) familiar touch .

Elon musk said in [Tesla's AI event in 2021](#) : ***the goal behind the creation of Tesla robot is to take over menial and dangerous tasks from human workers, which would improve safety.***

While others, like Boston Dynamics focus on versatile robots like [SpotMini](#), capable of various tasks such as inspecting sites, and handling objects. Amazon also has a [6-foot-tall robot](#), called [Digit](#), that mimics human movements just like other workers. [The on-site manager even tried to encouraging](#) collaboration between humans and robots, so we could see [humans and robots can actually work together](#).

But there is one hitch! can robots work together with humans peacefully, though ? Well, they might if they are still remotely controlled by human beings and without any disruptions from programs or cyberattacks. But, [what if the robots can Think for Themselves](#) one day ? Actually, Boston Dynamic AI Institute is researching ways for robots to comprehend and tackle complex challenges with little or even no human help.

But how ? They've been exploring how Large Language Models (LLMs) can assist in robotics, particularly investing in areas like reinforcement learning. This involves mainly developing in the simulation just like what they did for doggie robots and then transferring those practices to physical robots. Just imagine if they can enable machines to autonomously discover and integrate data, compared to the current method of manually designing solutions. When the physical and cognitive fit together, yeah right! That's a self-conscious AI robot! -

But something's off... You see, when we have robust AI robots, combined with really powerful multimodal AI intelligence, just like what we saw in the Wall-E movie. People worry about robots taking over our jobs, but researchers are really trying to get them to afford more work and help humans. Even with the possibility that AI robots can gain autonomy someday, their ability is still under human control. Nothing would look like what's in the apocalypse sci-fi movie, **right** ?

Well, What's currently lacking is the backbone. So Now, let's talk about quantum AI. ( [my animated avatar](#) )

# Quantum AI Computing

Before talking about Quantum AI. We need to talk about [Quantum computing](#), we're not pretending to be an expert on this, because **no one can**.

( black mirror S6 -> this is quantum computing, we barely know it )

You know back in the old days, traditional computers use machine code, each unit which represents either a 0 or 1. We call it binary.

However, in quantum computing, each unit could be both 0 and 1 at the same time. We call it '**Qubit**' or **quantum bit**, it is a basic unit of quantum information.

It is extremely challenging to stabilize Qubits at room temperature, and that's why lots of effort went into researching and building quantum computers.

With a sufficient number of Qubit, quantum computers have the potential to operate [millions of times faster](#) than the current fastest microchip-based computers. -

In 2019 when [IBM unveiled its first circuit-based commercial quantum computer](#). And in December 2023, [IBM released the first-ever 1,000-qubit quantum chip](#).

Leading tech companies like [Microsoft have been building most scaled quantum computer and bring to Microsoft Azure](#), and Google has been advancing the [state of the art quantum computing and developer tools](#). And other emerging players such as IonQ and D-Wave Systems in the market.

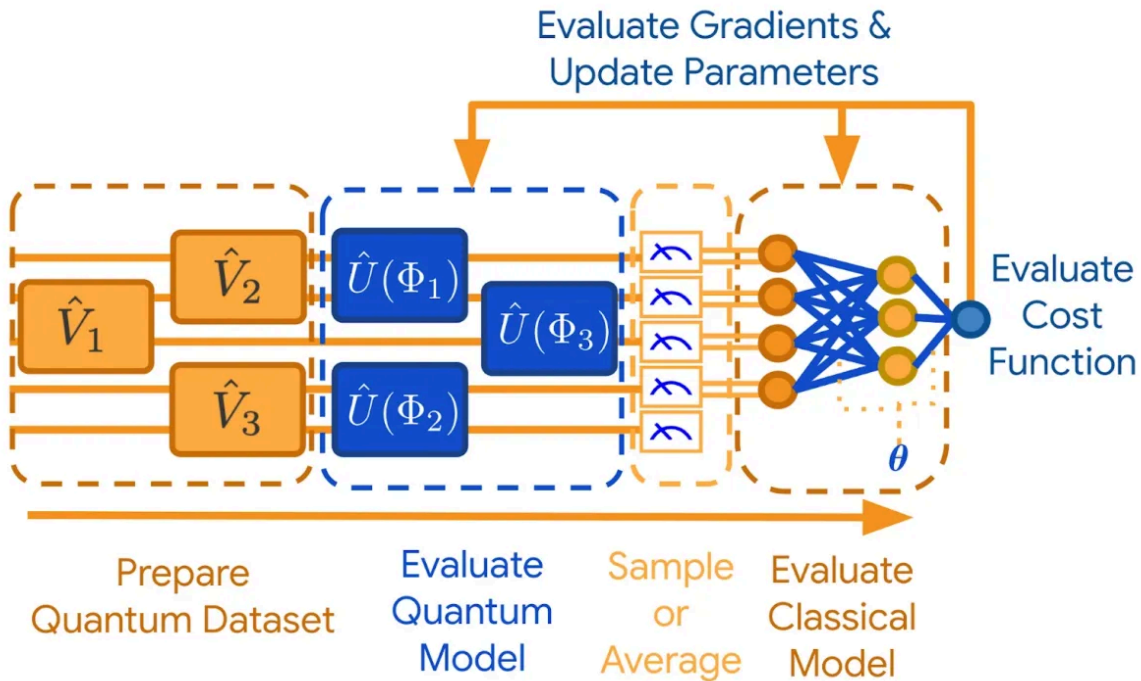
So now let's talk about 'Quantum AI', Quantum computers deliver far greater efficiencies than classical computing for some optimization algorithms. Quantum AI is the use of quantum computing for the computation of machine learning algorithms.

Think of it this way, [Quantum computing can provide the ideal stage for machine learning by providing the right data faster](#). Machine learning is about computers learning from data and being able to create and understand patterns, just like the human brain does.

However, in many cases, machine learning may be constrained by the poor quality of data and the slow availability of data. Quantum computing can compute huge volumes of data quickly. So quantum AI can accelerate AI training and efficiency to help achieve results that are normally not possible to achieve with classical computers.

[Google](#) currently provides TensorFlow Quantum (TFQ), a platform for prototyping [hybrid quantum-classical AI](#) models for learning and experimentation. **The future awaits.** (put motion text 'the future awaits' on screen )





## How Far We Are From the Utopian Future

So the question is, how far we're from the utopian future? We recently had a glimpse of it from [CES 2024](#), where we can see AI [robots](#) have already taken a huge step into the future.

Intuition Robotics showed an enhanced version of robot, [ElliQ](#), which leverages generative AI to care for elderly users in various activities.

And Samsung introduced a [Ballie personal assistant robot](#), it has an [onboard projector](#) with advanced quite cool AI features such as [intuitively](#) adjusting to the preferences and needs of users.

We have so many exiting technologies in our lives already, they keep coming everyday ! Such as metaverse, spatial computing with Apple Vision Pro, web3, blockchain, cryptocurrency, and imagine when we can store information in glass.

Btw that's an idea for a different video -

## Will AI robots take over our jobs though?

With so much innovation in the tech industry, will AI robots take over our jobs, though? ([push further](#)) and how soon will it happen?



It's true AI is growing fast, at an excessive rate, but we make decisions for them in many ways. Of course, that's what happens, at least, before AI can gain autonomy and decide their own path.

You see, today, big companies like Google and Microsoft call it as 'co-piloting', which translates to making people become more productive. **66** -

But let's get to the bottom of this.

[At current costs, U.S. businesses would consider automating tasks only if the expense is about 23% of worker wages, at least that's the case for computer vision tasks.](#)

What does that even mean?

This indicates that AI automation can speed up if costs drop quickly enough to exceed what individual human workers can do. Think about how much it costs to pay human workers in different countries. ***This does have impact.*** When paying human workers even in minimal wages becomes much more expensive than using AI robots, that's when things get out of hand.

But what about regulations and AI ethics? Do all countries allow AI to replace humans on such a large scale?

The question still remains! For now, we can safely assume that AI job displacement will be significant, but it will also happen gradually. -

## Outro

We may not be too far from the utopian future, and soon what we envisioned in our heads will become reality.

**There's nothing wrong with these technologies as long as they're in good hands and being used for good.**

What do you think about AI robots? Comment below, we want to hear your thoughts. Thanks for watching. If you enjoy this content, I have a playlist about AI & cloud-native technologies here, and watch another video about AI on screen.

Don't forget to hit that like button and subscribe if you found the information here invaluable, because that helps out the channel. I'll see you in the next one. -