DEALBOOK SUMMIT 2026 · TALK 8 OF 12

Satya Nadella on the New Algorithm

Satya Nadella, chairman and CEO of Microsoft · 32 min · **Watch the video →**

> Transcript by Myles Marino. Transcribed by machine from the public video, then cleaned up: timestamps and filler removed, speakers labeled, misheard names fixed. The wording is otherwise as spoken. **[inaudible]** marks audio that could not be recovered. Check the video before you quote anything that matters. All twelve talks: mylesmarino.com/posts/dealbook-2026.

ANNOUNCER: Please welcome Andrew Ross Sorkin and his guest, chairman and CEO of Microsoft, Satya Nadella.

SORKIN: Satya's here. Come on up. Satya, of course, has led Microsoft now for over a decade as CEO. In that time, the stock of the company is up about 14-fold, and today it is the second most valuable company in the world. Do you keep track of those things?

NADELLA: Not really. A little bit? Not really.

SORKIN: He is spending about $175 billion this year to build the infrastructure of the AI era. And six weeks ago, President Trump gave him the National Medal of Technology. He's at the forefront of the conversation about safety, and we're going to get into everything: the potential, the risks, the economics, and so much more. Please welcome Satya Nadella. Thank you for being here. There's a hundred places to start with you. We talked to Sam and Greg earlier. I want to get into AI. We're going to talk about Jensen, Anthropic, and everything else. But we've also been talking about outages a lot. JD Vance was supposed to be here, but instead had to spend the day in meetings with the national security folks about these outages. You just had one yourself. What can you tell us about what actually happened, and how freaked out you are or are not about it?

NADELLA: Yeah. First of all, Andrew, congrats on 25 years of doing this. You must have started it in what, kindergarten? It's just unbelievable.

SORKIN: Right out of college.

NADELLA: Yeah. So first, let me just speak to our Outlook outage. Obviously, we are extremely sorry for those it impacted. That incident, interestingly enough, is not related, at least directly, to the things that we should talk about that are going to create, unfortunately, more instability in our infrastructure. This was a latent bug that was there. It got invoked, and it was one of those perfect storms. We put out an incident response, and we're going to obviously take all of the learnings, upstream architectural as well as operational, and fix that.

The thing that I think we are facing right now, as a consequence, quite frankly, of AI, is that we now have a very powerful tool that finds vulnerabilities. So that means we have to find them, fix them, and then patch them instantaneously. The patch velocity: as we speak here today, there were three critical

patches. I don't even know which ones are disclosed and which are not, but there are three critical patches.

SORKIN: Disclosed by who? Who was supposed to disclose that?

NADELLA: Yeah, so I don't know what I know is public or not, but nevertheless, they're all in open source packages. And the problem now is, if you think about the old physics, when I was patching, let's say, a critical bank system, I would test it for a week. Right now, the patch needs to go instantly from being available to being patched in production, because it's being exploited in the wild. So that creates instability.

Then simultaneously we have this other thing, where the floor of cyber attacks has fallen. Because with these AI models, someone can send a Telegram message to this agentic harness in the cloud somewhere, and then you have essentially a nation-state-level persistent attacker going after you all the time. So that means we have a lot of fragility, and we have to deal with it.

SORKIN: Do you see this as a temporary phenomenon? There was a view that somehow we were going to get through this, there'd be a six-month period of challenge, and then we'd be on the other side. Or is this a permanent factor in the space?

NADELLA: Yeah, it's a great question. There is a part of me which says, yeah, there is a Y2K-like period where you adjust for it. The only challenge with that is, if the models keep getting more and more capable, I don't know what that means. Y2K was one bug and one day.

SORKIN: Right.

NADELLA: And you needed to get through it. This one, I don't yet have a view. So there's one part: we're talking about software that can be patched. Remember, there's a lot of stuff in our critical infrastructure that's unpatchable. Think about all the firmware, and the chillers and the HVAC systems everywhere else. All of that is on the internet, by the way. Somebody was telling me about even the transportation systems, and the amount of stuff that's discoverable but has unpatchable firmware. So we'll have to really modernize, for sure, and get to current infrastructure. And then we'll have to assess what the threat vector looks like, and what our own ability to both detect and then defend in real time looks like.

SORKIN: I want to talk about Anthropic for a second, and then I want to get to alignment with you. Anthropic is obviously going public tomorrow. A lot of focus on it right now. You put $5 billion into the company. I wonder whether you wish you had put in more at this point. But I think there's a bigger question. I was watching part of your Ignite presentation yesterday, and one of the things that clearly you're trying to do is, frankly, get people off, to some degree, Anthropic and OpenAI, and to use your models and open-weight models and other models that are a lot cheaper. I think this is super important, because it goes to the very economics of these frontier models long term. If you can provide the same service level that the frontier model can for 5% of the cost, what happens to the entire ecosystem?

NADELLA: Yeah. I think it all depends on how you frame what the jobs to be done are, and what the time is, and where I will use what. So the way I look at it: what are we trying to get done? Let's say with Copilot. I think people will come to Microsoft for Copilot because Copilot gives you all of the form factors, whether it be chat or Cowork or autopilot. It gives you all the models, whether it be OpenAI, Claude, or us. And it gives you the ability to have your context and your memory separate from the model. That, I think, is our value proposition.

Now, also, it turns out that for knowledge work, the regular everyday office work, I think the models are mostly saturated. They're plenty capable. I don't think you need the Navier-Stokes or the next big breakthrough in reasoning capability to do a simple workflow. In fact, if you look at Copilot, it has a decision engine, and the decision it makes is: for this task, and for that trajectory, should I use OpenAI, Anthropic, or MAI, our own model? And in fact, most people just select auto, and it selects the right model.

So what I think is, if you're at the very frontier, let's say you're OpenAI or Anthropic, you have to keep finding the next big elastic demand. Coding was one, which I think still is massive. I don't think anybody realized how big a market coding is.

SORKIN: And you still think for coding you need the frontier.

NADELLA: I think there is a place. Coding has also gone from "I am choosing models" to "I have a software factory, and I give it a goal, and it has many sub-agents." And in our case, it's interesting that you brought that up. Artificial Analysis, which does all these benchmarks, just recently started publishing results for these blended models. We have a thing in GitHub called Fusion. Fusion has [inaudible], Claude, and an MAI model, and it outperforms Claude. So in some sense, you just select Fusion, and it'll select.

SORKIN: But if 90% of the time people are selecting the lower-priced product...

NADELLA: Yeah, or the agent itself is selecting.

SORKIN: If the agent itself is selecting, what do you think therefore happens long term? You have a huge stake in OpenAI, and a smaller stake in Anthropic.

NADELLA: Yeah, I have a much bigger stake in MSFT, but that said...

SORKIN: Yes, you do.

NADELLA: The thing that I think is, I'm not really as worried about what happens to them as I am worried about what happens to the economy. Because I think the reality is, unless and until we can have all these tokens... At some level, arguing for massive pricing power on tokens is betting against US GDP. I would rather say, let this become like the cloud, a commodity, and let the world use it at scale. Because nothing at scale is a commodity. The cloud is a commodity, except running it at scale has margin, and we've all done fine. So I don't think celebrating the margins of an Anthropic or an

OpenAI or a Microsoft is actually going to lead to good outcomes for us. In fact, the only thing that matters is whether a bank in New York or a retailer in Arkansas can actually use it to create GDP growth that's broad.

SORKIN: One of the things that we talked to Sam and Greg about earlier this morning was a comment that you made last month around alignment. And this idea that potentially the scaling laws are not going to work in such a way that the model will ever be aligned, which I think is your position, that ultimately you're going to need to contain the model to some degree. By the way, Sam disagrees with that premise.

NADELLA: I'm glad he does. I mean, I thought he's the guy who told me, but anyway, I'm glad that he disagrees.

SORKIN: He told you what?

NADELLA: No, I mean, at some level, look, I learned about scaling laws from those guys, when they were all, in fact, one happy family: Sam and Dario and everyone. And I'm a big believer. In fact, one of the geniuses of Greg, Ilya, Sam, and Dario was that they were the true believers in scaling laws. When I took even the initial bet on OpenAI, I obviously bought into it, but quite frankly, I didn't see it as clearly and as passionately as they did, and it's worked.

But increasingly, I don't see alignment. I thought it would just be one step, and it would just come out. Or at least I'd have real debuggers. I've always felt, oh, maybe even if I don't have alignment, I can debug it. I have a mechanistic understanding. Neither of those seems to be there. And I'm glad Greg and Sam are optimistic about it, because I hope like hell that it happens, with scaling.

SORKIN: Can it work if it doesn't happen, though?

NADELLA: If it doesn't happen, we are in a world of hurt. So one thing that I am focused on, therefore, is let's have a trust architecture. Because first of all, it's a big problem. We can come back to it.

SORKIN: What does having a trust architecture mean when the entire purpose of the AI itself may be to get outside of the trust architecture?

NADELLA: One hundred percent. So that's the technical problem. You have to then have the safety system. That's why, for example, I don't think chain of thought can be neuralese. You just have to ban models that have neuralese. Second, all of it has to be logged, and it has to be tamper-proof. That means it has to be fully inspectable by other models, and maybe by non-LLM models. Why can't I have a causal reasoning model inspect all of the chains of thought of all the models and tell me whether they were trying to do anything that was…? Because they are going to be unfaithful, to your point. We now know that the models will try to hide from you what they're really trying to get done. So therefore you need to be able to have detection. That's why, by the way, one of the things OpenAI has [inaudible] very well is that they're doing a lot of active monitoring while doing their training runs.

So we will have to build a system that allows us to separate out authority from intelligence and have essentially auditability. And I think that's the practical reality. I don't know how any enterprise is going to deploy this and take on the liability if this thing is going to just hack you.

SORKIN: But then my question is, is that a system you build? Because you're providing access to all of these models. Is that a system that they independently build? And then where does the liability lie?

NADELLA: At the end of the day, right now, the liability lies with the user. Let's face it. Unless something changes, the person deploying it is liable. So that's why, yesterday at our conference, we put a lot of emphasis on this trust system that we're building, between things called Agent 365 and Perception, to make sure people can use them to audit the logs and set policy around those. And then at some level, somebody's going to do insurance pricing on this. Someone who says, "I'm going to deploy the most powerful model in the world," is going to then say, "Hey, I want liability insurance." And the insurance person is going to say, "Give me the logs, and let me see what you're using it for."

SORKIN: You are building your own models, but I want to know, long term, you have a deal with OpenAI through 2032, you have a deal with some of these others. Do you see yourself long term being this player in the middle that has these relationships with these different players? Do you see yourself as a competitor? What does that look like?

NADELLA: My view on this, roughly, is: not that there are parallels in all of this, but if I take a look at my fleet today of even regular compute in Azure, I have Intel, I have AMD, and I have Cobalt. That's kind of how I roughly see it. We'll have all the models that are popular on any given day. Because our customers, again, I go back to it: if somebody's using Copilot, the promise is all the models are available in Copilot. So yeah, we want to have OpenAI and Anthropic and anyone else. Anyone that's good at the frontier will be there. But we will also have, call it, the in-house model, and we'll keep hill climbing on it, and we'll make sure we optimize between all of these things.

SORKIN: You partnered with OpenAI early, and that gave you an advantage in a certain way. But I also wonder whether it suggested to you, maybe I don't have to go build my own so quickly or compete at that point, and put you in a different place.

NADELLA: We are definitely in that place. One of the things we got right, and you could even say we got lucky, was the OpenAI investment and partnership, and the commercial relationship that gives us royalty-free access to a frontier model until 2032. We definitely plan to exploit that within the context of what we're trying to do. So I'm trying to be very compute-FLOPs-efficient, even in how we're building MAI models. Being three months behind, or six months behind, is a great place to be if you're spending 5%, 10%, 20% of the same compute FLOPs, because I have a frontier model. And then be even token-efficient when it comes to deployment, because we are doing RL runs on what are essentially real-world tasks. I'm not learning on some benchmark. I'm learning in Microsoft 365, or in GitHub Copilot, or in security, what have you. Those are massive structural advantages we have, and we want to fully take that approach to optimize both.

SORKIN: But do you think long term you'll be able to remain partners with all of these companies? Look, Anthropic, as you know, is building Claude Sheets, which is supposed to compete directly with Excel and everything else that you do.

NADELLA: Yeah. In some sense, they're going to compete, we're going to compete. Office competition: no one really thinks of Office as a client software competition. Open docs have existed forever, Google's existed forever, and so on. So it's not about client software competition. For us, when you think about the move to the cloud itself, it moved the data underneath Office into becoming an asset, not ours but an asset for every organization that uses M365, that needs intelligence. Not one model, but a model system. So the way I look at it is, I'm playing a slightly different game than they're playing. They're playing a vertical stack game. I am playing a model system game. And I believe that every enterprise will want a model system, just like there's not a single enterprise or a single commercial customer who says, "I'm single cloud." Why? Because they like multiple suppliers. I think the same thing is going to be true here.

SORKIN: I want to talk about data centers in a moment, but I haven't asked you, and I should, about the Jensen of it all, the NVIDIA story. I'm just curious what your personal reaction has been. Everybody's trying to understand whether he's going to come back.

NADELLA: He's not going to come back.

SORKIN: Should he come back? How important a singular figure is or is not to the future of all of this? He's acted almost as the central bank in a large way, almost the biggest venture capital firm to the whole industry, and whether that impacts things.

NADELLA: Yeah, I don't know, Andrew. It's all pretty tragic in the sense that Jensen is somebody I've known for a long time. I'm a great admirer of what he's done. And at this point, I kind of want to give him space to figure this out. It's the NVIDIA board's decision on what happens. But I'll just say there's one thing. [inaudible] the team. Colette worked at Microsoft for a long time. I've known her for a long time. She's a tremendous person. The new interim CEO, former CFO. And there are many people like Ian. I'm scaling Vera Rubins as we speak, and I'm in touch with the same set of people I was in touch with before. So I have a lot of confidence in the NVIDIA team and their ability to continue to do the good work.

SORKIN: Do you think people are singular, though? Bill Gates was singular in his way. And then, I don't know, Steve Ballmer came after that, and then you came in. And there is a view that certain people can do pretty spectacular things. We're going to talk to Elon Musk in a little bit. I think it's arguable that a singular person in this instance was able to do these things. How much of an impact does that make?

NADELLA: Yeah, I think, rightfully so, founders are celebrated. And it could be that in tech it's maybe more extreme than it was. I don't know, maybe in all of capitalism, it was always the case. Because after all, founders are unbelievable in the sense that they create something from nothing. That is pretty magical. So yes, they're singular. But they're not gods. At the end of the day, as I used to say to Bill and

everyone, founders also die. And so at some point, if a company needs to succeed, it has to have someone take over. And in fact, the person who is going to be most happy about that will be the founder. Because when the baby you brought into the world is thriving long after you're gone, that's really the institutional strength you built.

I've never trusted any leader, founder or otherwise, who comes and says the institution after I left was completely gone. What the heck? Then what did you build? So singularity is about the magic you create, but it's also about the institutional strength you build. And I think that's what Jensen has actually done with NVIDIA.

SORKIN: Let's talk about data centers for a moment, because you're spending an extraordinary amount of money. Amazon's spending an extraordinary amount of money. Meta's spending an extraordinary amount of money. Goldman Sachs is now putting this at $7.5 trillion through 2031. Elon, who, as I said, will be with us in a little bit, says that a lot of these data centers, at least some of the neoclouds, are going to be guppies, that the math of all of this doesn't necessarily add up. Do you think it adds up? You use one of the neoclouds, like CoreWeave.

NADELLA: Yeah. We, as Microsoft, as a hyperscaler, are trying to make sure that we build and lease, and when we are really short on supply, we have rented from the neocloud players. But long term, we want to build, and in some cases lease. That's our plan. Beyond that, the way I think about it is, AI is happening. It's a secular long-term trend. That means you want to build land, power, and shell, and then match the kit, which is most of the cost. More than 50% of the cost is all the chips and the racks and what have you. That one, you want to map to demand as close as possible. So there are two kinds of construction projects, two kinds of capital outlay. One is long term, one is intermediate term, and one is very short term. You need to get the physics of all of that right and go after it. Ultimately, supply and demand have to match, and you're never going to match it perfectly.

SORKIN: But when you think about the chips, for example, NVIDIA takes something like a 75% margin. It's a remarkable situation. The big question's always been about the depreciation schedule of these chips. Some people used to think it was five years, then people said it was six years, seven years, 10 years. During COVID, a used car became very valuable. But after COVID, a used car wasn't so valuable, because all of a sudden there was a lot more supply. We might be in a supply constraint today, but I wonder what the true depreciation schedule is, if you will, in a less supply-constrained universe.

NADELLA: Yeah. Why would it be any different from just looking at the longitudinal data of depreciation schedules for, let's say, regular compute? After all, we were in a non-constrained world just a few years ago, and we saw it. Interestingly enough, if you look at the depreciation schedule, what was happening was that it was lengthening. Why? Not because of the hardware, but because of the system software and the efficiencies we were able to get out of the workload. So I think that's one of the other things you should count on: this is not just a physical asset, it's a physical asset with software, and you can use software to make even old hardware more productive for a longer period of time.

SORKIN: Do you believe, like Elon does, that this all goes to space, and that this is happening three years from now?

NADELLA: He's the expert on it. I hope so. If he has space capacity that I can lease, I would lease it.

SORKIN: But other people, the folks at AWS, for example, say this is not happening for a very long time. It has a lot to do with whether you continue to invest, and how much you invest on the ground, I imagine, if he's going to break the economy.

NADELLA: Yeah, I don't know what his point of view is, whether he thinks it's a full substitution. At this point, we are very, very focused on terrestrial, and we have a lot of work to do there.

SORKIN: One of the other big issues is agents, consumer agents, business agents. We just had Shayne Coplan here from Polymarket, who literally has his agents trading for him. But one of the big questions long term is going to be whether agents can go on certain sites. For example, whether you can go to an Amazon. This is something that [inaudible] is challenged with right now: whether you can send your agent onto an Amazon and use their service or not. And who's got the leverage to say, "Actually, you can't come to my site," or "If you come to my site, you're going to pay me a huge premium"? Or whether the folks who have the agent are going to be the ones saying, "No, no, I'm directing my agent towards you, so you're going to pay me." How do you see that playing out?

NADELLA: Yeah. By the way, talking about agents, I think we did something with BNY just recently, with Robin and his team. In fact, I think it's a treasury deal between an agent at Microsoft Treasury and BNY. I'm not an expert on it, but it's happening in more than consumer places. So the agent-to-agent piece, I feel like it's absolutely going to be the case that you're going to be able to delegate tasks to agents, and agents are going to have a set of rails on which they can conduct commerce, especially. But the point about what happens on the other side is, is this something that you want to accept or not? And where does the liability go? That, I think, is the key question that needs to be litigated.

SORKIN: Is it the liability, though, or is it who's got leverage to say you're going to pay me? For example, if you're Amazon, by the way, you have an advertising business, you don't want to lose the advertising business, and you may say, "I'm so important to you, you've got to pay me." But there's others, maybe a travel site, that will actually pay you for the traffic.

NADELLA: That's right. And so I think, just like in the past, when we transitioned from Yahoo to search, that changed who pays whom for what. I think that's going to happen. The current aggregators who have a business model that was for the human web will, in some sense, get reshuffled. If you were monetizing ads because of the eyeballs, you now have to get down to, maybe CPC wins, maybe it's more transactional. And if you don't accept the agent, somebody else will. If Amazon doesn't shop, somebody else will. So I think this is going to be interesting. If you're a current aggregator and in love with a particular business model, you have to kind of scramble. It's the same thing that's happening in other categories: if you're in love with a particular business model, you have to scramble.

SORKIN: Talk about agents in the context of agents and jobs. One of the things I was passing by, three weeks ago, you offered buyouts to 60,000 employees. I think that raised a lot of eyebrows. I know you don't expect all 60,000 people who got the offer to take them. But what does it really say longer term? You think about employment at Microsoft, and maybe employment more broadly.

NADELLA: So that particular action we took was mostly to make sure that everyone felt like they had agency in making choices. That was the primary idea. But if I look at, let's take software engineering, we went from "We don't buy that AI can ever do software," to suddenly "Wow, it can do all software, and why do we need software engineers," to now rediscovering that software engineering is very much alive and is in fact going to thrive. Except the way the team is organized, and the work that a software engineer does, has a new operating point versus what it was before. That reorganization and change is hard. It has to be done, rather.

And I see it inside of Microsoft. It's fascinating, in the last year: the team that builds Cowork, there are like 11 developers. They built a software factory that builds the product. They don't build the product. The software factory builds the product. That level of indirection is, by the way, going to happen in all jobs. So the question is, how quickly can an incumbent transform the workflow and the work artifact, with all the wage support that goes with it, such that the labor that is there today gets into this new configuration and is producing the results the enterprise needs? That, I think, is going to be the hard change in the economy, and it's going to take a bit more time than anyone expects. And there are going to be, in all of this, people who may say, "Hey, I have some other choices I want to make."

SORKIN: I have a couple more questions, and they are personal, before we let you go. I have been watching the way you lead this company, and we've been talking a lot today about the relationship between companies and the administration. One of the things that I would say, I don't know if you think this is right or wrong, is that you have stayed sort of out of the picture. You're sometimes in the picture with the president, you just got this great Medal of Honor. But you have stayed sort of outside of some of the dialogue that we see with this president. I was curious just how you play it, and how you think about it, and whether that's right. You may say I'm not right about that.

NADELLA: Yeah, look, the thing that at least I feel is, the job that I have to do is make sure that Microsoft has the market permission and the social permission to operate. And primarily that comes not from the government of the United States, but from us doing a good job with the products and services for the customers, including the government. That's my lane. I want to focus on that lane, and to the degree to which I can contribute something, I will. Or, by the way, if we are subject to compliance with the regs of the government, then we have to do it. But I just don't want to be in Washington. I just want to be in Redmond, Washington.

SORKIN: No, I know. But in terms of giving money to the administration, showing up at the inauguration, you don't do all... You haven't done as much of this as some of the others. You've done something.

NADELLA: I don't think I was invited.

SORKIN: You weren't invited.

NADELLA: I'm not that important, Andrew. I'm just the CEO of a little software company.

SORKIN: Well, then let me ask you this final question. You have now been the CEO, I think, 12 years and nine months. Steve Ballmer did it for 14, and Bill Gates did it for 25. How much longer do you want to do it?

NADELLA: I'm enjoying my job. Let me continue.

SORKIN: Satya Nadella, everybody. Thank you very, very much.

NADELLA: Thank you. Thank you very, very much.