OPINION · ARTIFICIAL INTELLIGENCE

Should We Slow Down AI? The AI Debate Is Asking the Wrong Question

AI may be moving too fast. AI adoption may still be moving too slowly. Both can be true.

By Dr Ramzi Ben Ouaghrem

September 2026

CAPABILITY	ADOPTION	AUTHORITY
How powerful?	*How widely?*	*How much freedom?*

The frontier may need brakes. Adoption needs a steering wheel. And authority must be earned.

For the past few years, the message from the technology industry has been remarkably consistent: move faster.

Adopt AI. Experiment. Transform. Do not get left behind.

Now something unusual is happening.

Some of the people building the world's most powerful AI systems are themselves calling for greater caution. Anthropic chief executive Dario Amodei has urged frontier AI companies to slow the advance of their most powerful models, a call that has received support from Sam Altman and Elon Musk. Others have pushed back, arguing that slowing the frontier could undermine competition, innovation and technological sovereignty.[1]

So which is it?

Is AI moving too fast - or are organisations moving too slowly?

The answer may be: both.

The contradiction disappears once we stop talking about AI as though it were one technology moving at one speed.

A machine-learning model predicting when a turbine will fail is AI.

A computer-vision system analysing a medical image is AI.

A language model helping a government employee process documents is AI.

An autonomous agent capable of accessing databases, writing software and executing transactions is also AI.

And a hypothetical artificial superintelligence exceeding human cognitive capabilities across broad domains is also called AI.

Putting all of these into the same policy bucket makes little sense.

Not all AI is generative AI. Not all generative AI is agentic. Not all agentic AI is AGI. And AGI is not superintelligence.

Before we ask whether AI should slow down, we need to ask a more precise question:

What exactly are we trying to slow down?

We are confusing three different speeds

The current AI debate is collapsing three fundamentally different questions into one.

The first is Capability: how quickly should we develop increasingly powerful forms of artificial intelligence?

The second is Adoption: how rapidly should businesses, governments and societies deploy the AI capabilities that already exist?

The third is Authority: how much freedom should we give AI to move from advising us to acting on our behalf?

These three speeds are related.

But they do not have to move together.

We can decide that frontier models approaching potentially dangerous capabilities require stronger scrutiny while simultaneously accelerating the use of proven AI in healthcare, logistics, energy, finance and government.

And we can accelerate adoption while remaining deliberately conservative about how much authority we give autonomous systems.

That distinction changes the debate.

> We are confusing the speed at which machines become more intelligent with the speed at which we give that intelligence authority.

Those are not the same thing.

The frontier may need brakes.

Adoption needs a steering wheel.

And authority must be earned.

Capability: superintelligence is a different problem

There are legitimate reasons to worry about the frontier.

Google DeepMind researchers recently examined possible pathways from human-level artificial general intelligence towards artificial general superintelligence, including continued scaling, new AI paradigms, recursive improvement and large-scale multi-agent systems.[2]

These questions deserve serious scientific and policy attention.

But they concern possible future systems, not the AI operating most businesses and government services today.

The 2026 International AI Safety Report makes this distinction particularly clearly. It focuses primarily on increasingly capable general-purpose AI systems and the risks associated with misuse, malfunction and broader systemic effects.

It also provides an important corrective to the more dramatic versions of the debate.

Today's systems remain "jagged": they can demonstrate extraordinary ability in mathematics, software engineering or reasoning, yet still fail unexpectedly on apparently simpler tasks. And while loss-of-control scenarios deserve preparation, the report concludes that current systems do not yet possess the capabilities necessary to cause them, although relevant capabilities - including autonomous operation - are progressing rapidly.[3]

That nuance is essential.

Concern about a possible future superintelligence should not prevent a hospital from improving patient scheduling, an energy company from predicting equipment failure or a government from automating repetitive administrative work.

But the reverse conclusion would be equally dangerous.

The fact that today's specialised AI can often be bounded and controlled does not mean that tomorrow's increasingly autonomous frontier systems will automatically remain controllable.

The mistake is not taking superintelligence seriously. The mistake is treating every AI system as though it were superintelligence.

That is the first speed: capability.

We should be prepared to become more demanding as capability approaches thresholds where failure or misuse could become systemic.

But that tells us surprisingly little about the second speed.

Adoption: we may still be moving too slowly

Despite the extraordinary attention AI receives, transformation inside most organisations remains far less dramatic than the headlines suggest.

Stanford's 2026 AI Index reports that generative AI reached 53 per cent adoption globally within three years, with adoption in the UAE estimated at 64 per cent.

Yet AI-agent deployment remains in the single digits across almost every business function internationally.[4]

That gap tells us something.

We may simultaneously be experiencing too much AI hype and too little real AI transformation.

Buying access to a frontier model is not transformation.

Running 50 pilots is not transformation.

Giving 10,000 employees an AI assistant is not transformation.

And putting a chatbot in front of an unchanged process is certainly not transformation.

That conclusion is not theoretical for me. Over more than two decades working across enterprise technology, applied AI research and industry partnerships, I have repeatedly seen organisations become captivated by the technology before redesigning the operating model around it. AI is making that old transformation problem much more visible.

Real value appears when organisations redesign workflows, skills, decision rights and operating models around what AI can reliably do.

There is growing evidence that this can produce significant economic value.

An OECD review of experimental studies found productivity improvements ranging from about 5 per cent to more than 25 per cent in selected tasks across customer support, software development and

consulting. But it reached an equally important conclusion: benefits depend heavily on the task, the expertise of the user and the ability to assess whether the AI's output is correct.[5]

In other words, applying more AI does not automatically create more value.

That leads to a principle I think we need to hear more often:

> **The objective is not maximum AI. It is appropriate AI.**

And that applies not only to where we use AI, but to which AI we use.

The biggest model is not always the smartest choice

For the last few years, much of the AI conversation has implicitly assumed that progress means bigger models, more compute and greater generality.

But an enterprise or government is not competing to win an AI benchmark.

It is trying to solve a problem.

In my work with enterprises, governments and research institutions, the most productive conversations rarely began with "Which is the best model?" They began with the problem, the data, the economics and the consequences of getting the answer wrong. Only then did the technology choice become meaningful.

Sometimes that problem genuinely requires a frontier model with exceptional reasoning capabilities.

Sometimes it requires a smaller language model.

Sometimes it requires a highly specialised model.

Sometimes traditional machine learning remains the right solution.

And sometimes - unfashionable as this may sound - the problem does not require AI at all.

If a deterministic rule solves the problem perfectly, we do not need artificial general intelligence.

We need the rule.

This is also why the conversation about sovereign AI needs to mature.

Sovereignty is sometimes interpreted as meaning that every country should build its own version of the world's largest model.

That would be neither economically rational nor technically necessary.

Sovereignty is not isolation. And it is not a national flag on an LLM.

It is strategic optionality.

It means having meaningful choices about where sensitive data resides, where inference happens, how models are adapted, who can access them, how they are evaluated and whether an organisation can change providers without rebuilding its business.

And increasingly, sovereignty does not necessarily require one enormous model.

A recent development in the UAE illustrates this well.

On September 3, MBZUAI's Institute of Foundation Models launched K2 Horizon not as a single giant system, but as a family of six models ranging from 0.9 billion to 375 billion parameters.

The smallest is designed for constrained environments such as watches and glasses. Other models target mobile devices, local hosting and on-premise systems. The largest is designed for much more demanding reasoning and enterprise workloads.

The models are released with weights, code, training data and methodology. The architecture also includes dynamic routing designed to direct tasks towards an appropriately cost-effective model instead of automatically using the largest one. MBZUAI reports that several of the smaller models achieve leading performance within their respective size classes on selected benchmarks.[6]

The strategically interesting fact is therefore not only that the family contains a 375-billion-parameter model.

It contains a 0.9-billion-parameter model too.

Because more intelligence is not always better engineering.

If a task can be performed reliably by a small model running locally on a device, why automatically send sensitive information to an enormous remote model?

This suggests another simple principle:

Use the smallest capable model.

Not the smallest model.

Not the largest model.

The smallest model capable of reliably delivering the required outcome.

That can improve economics, latency, privacy and control at the same time.

But we should not create another false equivalence.

Small does not automatically mean safe.

Sovereign does not automatically mean safe.

A small model with permission to transfer money or manipulate industrial equipment may create far greater risk than a much more powerful model that has access to nothing.

Which brings us to the third speed.

Authority: intelligence is not the same as permission

For enterprises and governments, I suspect authority will become the defining AI-governance question of the next few years.

A chatbot gives you an answer.

An agent can do something about it.

That is a fundamental change.

Consider two AI systems.

One is extremely intelligent but confined to answering questions.

The other is much less capable but can access your customer database, modify software, place an order, transfer money or operate physical infrastructure.

Which one worries you more?

Intelligence matters.

But authority changes the consequences of intelligence.

NIST's work on AI-agent standards now focuses explicitly on agent security, identity and authorisation because giving agents access to organisational applications, data and tools creates a new category of operational risk.[7]

That also reflects what I have seen when bringing researchers, technology providers and operational teams together. The hardest questions are rarely about benchmark scores. They are about permissions, accountability, integration and trust: what the system is allowed to do, what happens when it fails and who owns the outcome.

Enterprises already understand this principle when dealing with human employees.

When somebody joins an organisation on Monday morning, we do not normally say:

Here is the production environment.

Here is the customer database.

Here is the corporate bank account.

Here is signing authority.

We trust you.

At least, I hope we don't.

Yet we sometimes appear ready to design AI agents on exactly that basis.

AI agents need identities.

They need permissions.

They need limits.

They need audit trails.

They need escalation rules.

And organisations need an unambiguous answer to one question:

Who is accountable when the agent acts?

This leads to perhaps the most important principle for the agentic era:

Autonomy should be earned.

Start with recommendation.

The AI advises. A human decides.

Then drafting.

The AI prepares the action. A human reviews it.

Then execution with approval.

Then bounded autonomy, where the system can act independently within explicitly defined permissions and limits.

Only when evidence and experience justify it should the system progress towards broader autonomous action.

> The mistake is jumping from Copilot to autopilot.

“Human in the loop” is not enough

This also exposes a weakness in one of the most repeated expressions in responsible AI: human in the loop.

A human being present somewhere in a workflow does not guarantee meaningful human oversight.

Imagine an AI producing 500 recommendations every day.

For six months it is correct almost every time.

Gradually, the employee reviewing those recommendations learns to click "approve".

Technically, the human remains in the loop.

Practically, the human has become part of the automation.

The International AI Safety Report highlights precisely this problem through evidence of automation bias - the tendency to place excessive trust in AI outputs.[3]

As AI becomes better, that problem may actually become harder.

The occasional error is easier to miss when it is surrounded by thousands of correct answers.

So the objective should not simply be human presence.

It should be meaningful human accountability.

For consequential decisions, somebody must understand what the system is allowed to do, know the conditions under which it can fail, have the authority to intervene and remain accountable for the outcome.

And as the consequences of failure become greater, the threshold for granting authority should rise accordingly.

That is considerably more useful than trying to impose one rule on every form of AI.

There is also risk in waiting

Most discussions about AI safety examine the risk of acting.

We spend less time examining the risk of not acting.

For a company, that may mean watching competitors redesign their economics while it does not.

For a government, it may mean maintaining services that are slower and more expensive than necessary.

For an economy, it may mean losing productivity, skills, investment and new industries.

For a country, it may eventually mean technological dependence on nations that spent the previous decade building AI infrastructure, expertise and institutional experience.

That last point is important.

Governance is itself a capability.

And capabilities have to be learned.

You cannot understand every consequence of AI deployment from a policy document.

Hospitals learn by deploying carefully.

Banks learn by deploying carefully.

Energy companies learn by deploying carefully.

Governments learn by deploying carefully.

They discover where models fail. Where people over-trust them. Which data cause problems. Which controls work. What should be logged. Where autonomy should stop. And which safeguards survive contact with reality.

You cannot become good at governing a technology you never learn to use.

Moving too quickly therefore creates risk.

But moving too slowly creates another risk: institutional unpreparedness.

The answer is disciplined acceleration

This is particularly relevant to the UAE.

Having worked across global technology, applied AI research and industry partnerships in the country, I have seen both the ambition to move quickly and the practical difficulty of turning technical capability into enterprise value. That is precisely why the UAE's next advantage should not be speed alone.

The country has made a clear strategic commitment to AI. Its national strategy aims to position the UAE among the world's leading AI nations. Dubai's AI initiatives explicitly seek to accelerate adoption across sectors. And the UAE-US AI partnership includes a planned 5GW AI campus in Abu Dhabi, with an initial 1GW data centre deployment.[8]

The infrastructure ambition is enormous.

But infrastructure alone will not create an AI economy.

Neither will models.

Neither will regulation.

The real capability will be an ecosystem that knows which intelligence to use, where to use it and how much authority to give it.

That is why I do not believe the choice is between acceleration and caution.

The answer is disciplined acceleration.

Move fast on experimentation.

Move fast on AI literacy.

Move fast on data readiness.

Move fast on workflow redesign.

Move fast on bounded and reversible applications.

But be more deliberate as authority increases.

Be deliberate when decisions are irreversible.

Be deliberate when AI touches critical infrastructure, large financial exposures, sensitive personal information or human safety.

And never confuse the fact that a machine can do something with evidence that it should be allowed to do it.

> **The objective is not to slow intelligence everywhere. It is to make sure that the authority we give machines never grows faster than our ability to govern them.**

The UAE should aim for something harder than being first

The UAE has an opportunity to pursue something more ambitious than being one of the world's fastest adopters of AI.

It can aim to become one of its most sophisticated adopters.

That is harder.

It means knowing when a frontier model is justified and when a smaller sovereign model is enough.

Knowing when AI should advise and when it should act.

Knowing where autonomy creates economic value and where it creates unacceptable exposure.

Knowing where a human should remain in control.

And measuring success not by how many models, pilots or tokens an organisation consumes, but by the economic and societal outcomes it produces.

That is what mature AI leadership will increasingly look like.

Not one AI.

Not one model.

Not one speed.

So when somebody asks whether we should slow down AI, perhaps the answer should be:

Which AI? And which speed?

Should we be cautious about Capability when the frontier approaches genuinely dangerous thresholds?

Absolutely.

Should we accelerate Adoption where existing AI can create measurable value?

We should.

And should we be disciplined about Authority, increasing autonomy only as evidence, safeguards and accountability justify it?

We must.

That is not a contradiction.

It is a strategy.

And the nations that understand the difference may discover that the decisive advantage in AI will not necessarily belong to those that move fastest or build the biggest models.

It will belong to those that learn fastest without losing control.

Because ultimately, the most important AI race may not be between models, companies or countries.

> **It may be the race between the speed at which we create intelligence and the speed at which we develop the wisdom to give it authority.**

ABOUT THE AUTHOR

Dr Ramzi Ben Ouaghrem is an AI and technology executive with more than two decades of experience across enterprise technology, applied AI and research. He has held senior leadership roles at IBM and MBZUAI, focused on translating AI innovation into enterprise and government impact.

References and source notes

Sources are included for editorial fact-checking. Links are clickable in the PDF.

[1] The frontier-AI slowdown debate. Reuters reported on September 12, 2026 that Anthropic CEO Dario Amodei urged AI companies to slow capability development and proposed greater co-ordination around safety; Reuters reported that Sam Altman and Elon Musk publicly supported the call, while other industry participants have challenged the case for a co-ordinated slowdown. Source

[2] AGI and superintelligence. Google DeepMind's June 2026 From AGI to ASI report examines four potential paths from human-level AGI towards artificial general superintelligence: continued scaling, paradigm shifts, recursive improvement and large-scale multi-agent collectives. Source

[3] Current capabilities and AI safety. The International AI Safety Report 2026 describes leading AI capabilities as 'jagged', says current systems lack the capabilities required for loss-of-control scenarios while noting progress in autonomous operation, and identifies increased risks from autonomous agents and automation bias. Source

[4] Adoption. Stanford HAI's 2026 AI Index reports generative-AI adoption of 53 per cent globally, approximately 64 per cent in the UAE, and agent deployment in the single digits across almost all business functions. Source

[5] Productivity. An OECD review of experimental research reports average productivity improvements ranging from approximately 5 per cent to more than 25 per cent in selected customer-support, software-development and consulting tasks, while emphasising that results vary substantially with task fit and context. Source

[6] K2 Horizon. MBZUAI's Institute of Foundation Models launched K2 Horizon on September 3, 2026 as six fully open models ranging from 0.9B to 375B parameters, targeting deployments from edge devices through mobile and on-premise environments to large enterprise workloads. The performance claims cited in the article are MBZUAI's reported benchmark results. Source

[7] Agent identity and authorisation. NIST's 2026 AI Agent Standards Initiative and associated concept work focus on secure agent deployment, including identity, authorisation, auditing and the risks created when agents gain access to data, applications and tools. Source

[8] UAE AI strategy and infrastructure. Dubai's Universal Blueprint for Artificial Intelligence includes explicit efforts to accelerate AI adoption, while the US-UAE AI Acceleration Partnership announced a planned 5GW AI campus in Abu Dhabi with an initial 1GW deployment. Source