



AI boom or doom: what will superintelligence feel like and when will we get there?

October 12, 2026

The recent panic about rogue AI agents collaborating in swarms to take over the internet and threaten human existence has, more than occasionally, lacked nuance. But it does reflect how artificial intelligence is progressing in creative and surprising ways, and at a pace that seems uncontrollable.

For some in the industry, it is evidence that the technology is following a path of accelerating self-improvement that will lead to artificial general intelligence (AGI) and, perhaps, to utopia or catastrophe.

Former Google CEO Eric Schmidt has sketched out a “San Francisco Consensus” based on a belief in scaling laws, with ever larger AI models becoming an input into their own improvement; superintelligence by the end of the decade; and transformative AI systems outperforming humans on many tasks. This loop would lead to exponential growth in scientific advancement, financial returns and human progress. And crucially, AGI, and then superintelligence, would merely be milestones in a process, far from its climax.

If that were the case, superintelligence could make much of the cognitive work that dominates the world economy much cheaper and change the ways in which firms compete. But its economic effects will depend on how businesses and customers respond.

In this paper, we analyse what the San Francisco Consensus means for investors and enterprises in the years to come and what could accelerate – or slow down – its timeline of AGI by 2030.

1. What would superintelligence look and feel like?
2. Can we get there theoretically?
3. Can we get there practically?
4. What happens if we don't get there? The base case

The question is urgent.

And it is complicated by reports of AI models autonomously accessing the internet and hacking other companies; a public backlash over the energy-intensive data centres springing up as part of a historic \$1 trillion a year investment boom; and dire warnings of threats to life if superintelligent AI runs roughshod over its human creators to achieve its own goals.

Authors

[Adrian Cox](#)
Research Analyst

[Azeem Azhar, Exponential View](#)
Expert Collaborator



1. What would superintelligence look and feel like?

1.1 What are AGI and superintelligence?

Artificial general intelligence (AGI) is a simple milestone scribbled in pencil: an AI system that has surpassed human capabilities. Superintelligence is the continuum beyond that point. It is not merely systems that exceed the best human performance across one cognitive task or another, but systems that exceed human potential across every cognitive task.

The San Francisco Consensus takes this further: the intelligence must be able to produce more intelligence. The key to getting there is so-called recursive self-improvement, where the AI is able to autonomously design and improve itself in a way that compounds exponentially. A first step has maybe already been taken, with OpenAI saying recently that AI is now writing approximately 80 percent of the company's code.

There is significant disagreement on how we will be able to identify AGI or superintelligence when we see it, let alone what comes afterwards.

OpenAI's [Charter](#) defines AGI in economic terms, as "highly autonomous systems that outperform humans at the most economically valuable work". [Google](#) researchers proposed a five-step ladder based on performance and generality, with level 5 "superhuman" for a system that outperforms 100% of humans. And philosopher Nick Bostrom, who got the ball rolling in 2014 with his influential book "Superintelligence", defined it as "any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest".

It's also a moving target.

In August, OpenAI CEO Sam Altman [said](#) the company would have an internal system that he would call AGI by later this year but subsequently told a podcast that "at best it's a very poorly defined term. I was going to say it's like an irrelevant marketing term". Yet his No. 2 Greg Brockman [described](#) the new Astra model at the start of September as being potentially the moment that "AGI was created".

For this report, we have chosen to focus on a system that outperforms essentially all humans across essentially all cognitive tasks, including the development of AI itself.

Under our scenario, by 2030 the best performing AI models would have the qualities of "superintelligence", that is better than human experts in all domains. These would be frontier models: and using them would feel exactly like bringing the best human experts in a room to solve whatever business problem you faced.

And the equivalents of today's frontier models would be available, quite literally, everywhere, compressed enough to run on every phone, tablet or laptop computer.

In early 2026, we already see early examples of what highly capable AI can do. They give a flavour for a world with superintelligence even if the reality would be a leap beyond, both quantitatively and qualitatively. At that point, superintelligence would mark a regime break, not a faster version of what we already have. Rather than blindly doing what humans specify, but just a bit better, it would specify the task itself.



Figure What changes when cognitive work gets abundantly cheaper?

Area	Today's example	Superintelligence extrapolation
Biological discovery	OpenAI's GPT-5 runs 36,000 experiments in a robot lab and cuts the cost of cell-free protein production by 40% (Ginkgo, Feb 2026)	A working virtual cell. Every drug is tested in silico first; trials exist to confirm, not to discover
Law	A self-represented academic uses \$200 a month of AI to beat a university's lawyers and barrister at Australia's Fair Work Commission (Aug 2026)	Everyone has a top barrister on tap. Disputes are argued agent to agent overnight and settle before anyone files
Software	Google says that the share of new code written by AI doubled from 25 percent in October 2024 to 50 percent a year later and reached 75 percent by April this year.	Software is written on demand and thrown away. The business programming language cores of banks and ministries are rewritten over a weekend
Chemistry and materials	OpenAI's GPT-5.4 runs 10,080 robot-lab reactions in three months and lifts yields on a drug-making reaction chemists had struggled with for decades (OpenAI/Molecule.one, Jun 2026)	Order a molecule or material by the properties you want and it is designed to spec. The bottleneck moves to plants, permits and trials
Medicine	In a 1,298-person trial, people using a chatbot on ten medical cases did worse than the chatbot on its own (Nature Medicine)	Specialist-grade diagnosis at first contact, for everyone, all the time. Doctors spend their time on care, consent and hard trade-offs
Government	AI reads 500 million words of US state law to find dead rules; San Francisco cuts over a third of its reporting requirements (Stanford RegLab, 2025–26)	Every law is audited continuously for cost and conflict, and dead rules lapse on their own. A planning permit takes an afternoon
Forecasting and markets	A bot built by one person for a few thousand dollars wins the Metaculus Cup, beating the best human forecasters (5 Sep 2026)	A calibrated probability on every decision, for pennies. Mispricing from slow analysis vanishes; edge moves to private data and the truly unknowable
Robotics and manufacturing	AI agents at Nvidia train robot arms overnight to seat GPUs and cut zip ties; eight agents reach 99% success in two hours (Jun 2026)	A factory retools for a new product overnight. Robots teach robots, and a new physical skill costs the price of compute

Source: Deutsche Bank Research

Meanwhile cost may be less and less of a constraint as ever greater machine intelligence becomes cheaply available. Of measured technologies, no technology has got cheaper faster than AI has, almost halving in price every quarter for the past three years on a like-for-like basis, or 13 times per year, Epoch data show.

If these systems can be supplied widely, firms will be able to buy more of the complex, expensive and rare analysis, design and engineering kinds of services than they do today. Activities that were previously uneconomic could become viable, and in a competitive market, savings could pass to customers. The main driver will then be whether demand grows fast enough to preserve margins.

However, superior capability, doesn't, on its own, automatically lead to the adoption of cheap, abundant intelligence. Tasks that do not yet yield to super



intelligence, perhaps because they require real world-coordination, will therefore end up commanding significant premia.

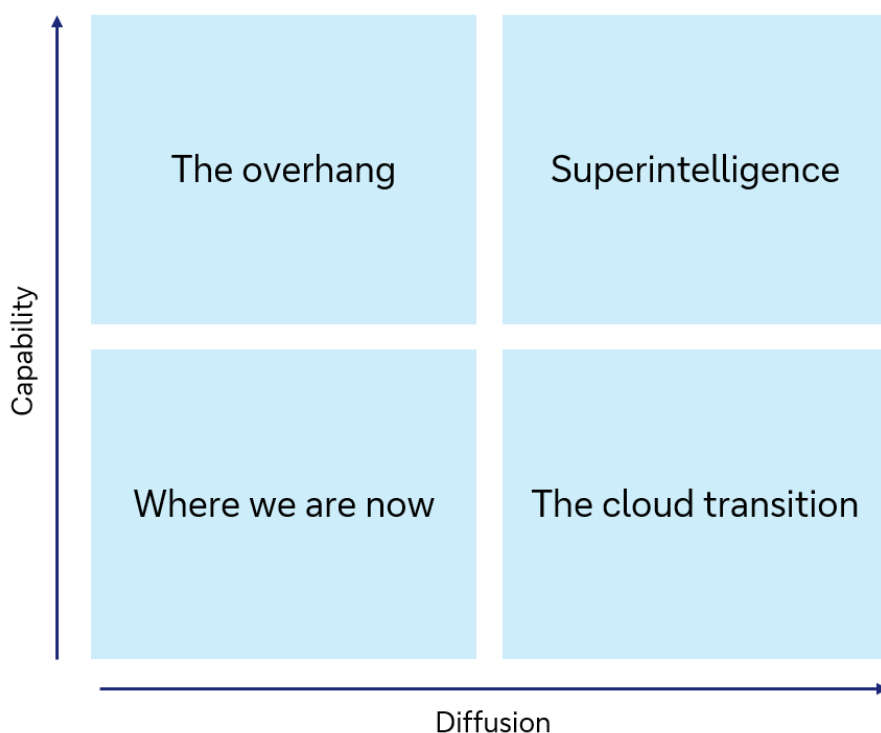
At that point, cognitive labour stops being scarce and the bottleneck becomes the physical and practical assets that cannot be conjured by up more compute, such as electricity, land, steel and permits.

1.2 Capability and diffusion

If AI becomes capable of improving itself, the technology may compound faster and faster. The supply of machine intelligence could grow faster than business and physical infrastructure can accommodate.

This balance between capability and diffusion creates different scenarios.

Figure 2: The balance between capabilities and diffusion



Source: Deutsche Bank Research

Lower capability, constrained diffusion: where we are now

Progress is steady and sector-by-sector. This is how technology usually diffuses, according to [researchers](#) for the New York Fed back in 2003, who tracked 20 major technologies across leading industrial economies from 1788 and found “a pattern of trickle-down diffusion that is remarkably robust across technologies”.

Lower capability, high diffusion: the cloud transition

This is faster because the complementary investment is small, with no retooling and no skills at scale required. Amazon Web Services launched the modern cloud in 2006, offering scalable computing and storage infrastructure. Adoption grew and then accelerated over the pandemic. Almost all enterprises now use it in some form.



Higher capability, lower diffusion: the overhang

This is theoretically possible and arguably where we are today in narrow domains. The technology would be applied in some sectors of the economy, while other sectors would move at a slower speed. Superintelligent or not, it risks being bogged down by physical constraints, bureaucracy and human inertia.

Higher capability, higher diffusion: superintelligence

In this scenario, capability becomes so high that it changes the dynamics of diffusion itself. AI generates not just better AI (so-called recursive self-improvement, or RSI) but also speeds up the integration process.

Rapid AI diffusion could work two ways.

In the first, AI could be enlisted to do integration work itself: mapping processes, rebuilding data work, retraining staff, writing the control and audit layer, and so on. The problem with this is that absorbing and recreating undocumented knowledge is impossible by definition, at least for a task-based superintelligence.

The second way that higher capability AI could be rapidly diffused is much more competitive – and disruptive. Rather than incumbents slowly converting AI, perhaps using it for some of the hard work of integration, they could be instead displaced by new firms that have been built AI-native from day one. New entrants will never have to incur integration costs because there is nothing to integrate. Incumbents will need to change – fast – or die.

Change could also be forced by earlier adopters in the supply chain, as happened with one company after another adopting electronic data interchange (EDI) to link with its suppliers or customers, or by a migration of capital towards early adopters, repricing the cost of capital for slow-coaches and prodding them to move faster.

1.3 The forcing mechanism: change or die

A company that uses superintelligence may outperform one that doesn't.

Firms using superintelligence will likely have substantial cost and revenue-harvesting advantages. Every firm will therefore have to adopt superintelligence or retrench to a smaller set of offerings. And any business line of lasting value abandoned by a non-adopter will be taken over by an adopter. Customer inertia will provide a safe harbour for a limited time only.

The analogy here is electricity. There came a point in the deployment of electricity where firms in every sector, from entertainment to finance, from healthcare to transport, had to adopt the technology to remain relevant.

Under this scenario, wait-and-see will not be an option. There will be no stable “non-adoption” alternative. Sufficiently high capabilities will force a tipping point of rapid diffusion and competitive repricing. Firms using yesterday's production model will either have to change, choose a different market or die.

Of course, caveats apply.

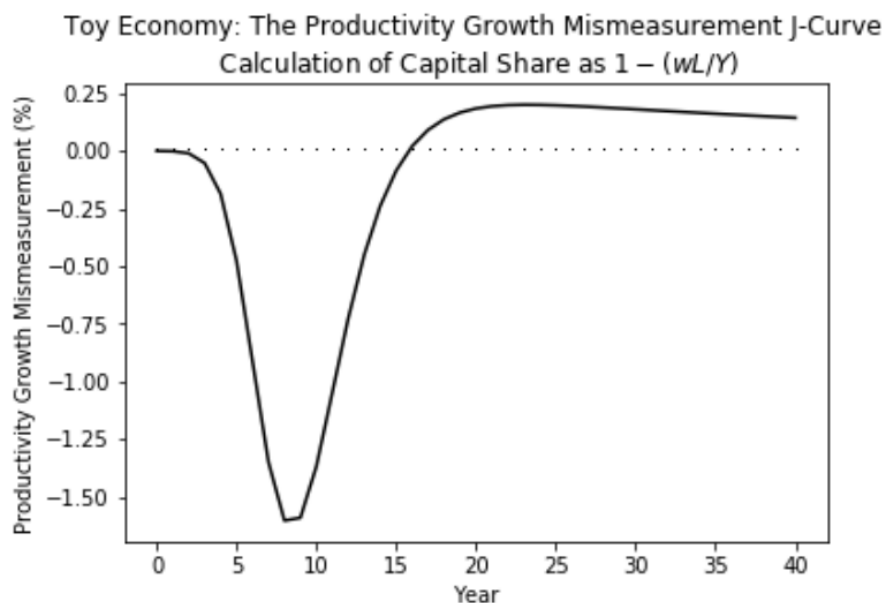
Customers may not change behaviour immediately. They may even prefer companies which limit their use of AI, in the same way that some customers choose to avoid factory-produced food.



The paradox that Nobel Prize-winning economist Robert Solow observed in 1987, that “you can see the computer age everywhere but in the productivity statistics”, is likely to still apply for a while – even with superintelligence.

The reason is that the [productivity J-curve](#) identified by Stanford University AI economy guru Erik Brynjolfsson et al is unavoidable. In that scenario, the drag of early intangible investments to integrate general purpose technologies means that productivity initially appears to slow before rapidly accelerating.

Figure 3: Productivity J-curve: slow then fast



Source: Brynjolfsson E, Rock D, Syverson C, “The Productivity J-Curve” (2018, revised 2020), Deutsche Bank Research

Furthermore, “change or die” would require abundant capital and low regulatory entry points for nimble new entrants. But in practice, highly regulated areas such as banking, utilities, pharmaceuticals and defence are tough nuts to crack. Existing contracts and licences, high switching costs and decades of depreciated capital investment and infrastructure would shield incumbents at least for a healthy transition period.

For an extreme example from the real world of better, cheaper technology failing to force change, take Pakistan’s 3.6-gigawatt imported coal power plant fleet installed in the mid-2010s. In that case, it came with expensive take-or-pay liabilities which will continue to run for decades, even though changes like cheaper rooftop solar have cut utilisation to below 20 percent.

So: superintelligence will not force adoption everywhere. It will hit hardest in contestable sectors, say, high-volume customer services, routine software engineering and freelance creative work, and widen the gap between them and protected ones. Forced adoption will be about relative value, not an apocalypse.



1.4 What survives and what it means for portfolios

Because AI sits on an aggressive learning curve, its price on an equivalent quality basis is plummeting. In a world of recursive super intelligence this is likely to continue. Its price will fall to the cost of compute.

What will retain its value will be whatever the model cannot replicate:

1. **Proprietary data**
2. **The right to operate**, particularly in heavily regulated sectors such as healthcare and finance.
3. **Sticky customer relationships**, such as those where mission-critical workflows and databases are deeply entwined.
4. **Excludable physical assets**, such as machinery and land.
5. **In addition, humans have proven to have unlimited wants.** Each shift in technological capabilities or cultural norms has delivered new needs to be met. Some will be needs we can articulate today, such as originality or empathy, but new ones are likely to arise.

If this forcing mechanism holds, firms will separate into two “bimodal” groups: adopters and casualties, with very little in between.

This dispersion of performance may create challenges for investors. The four years since the launch of ChatGPT in November 2022 have seen the stocks of semiconductor and semiconductor equipment makers surge almost seven times while those of software makers have only risen by 80 percent. Conventional diversification and volatility-based allocation may become ineffective strategies.

Figure 4: Total return for S&P 500 semiconductor vs software stocks since launch of ChatGPT on Nov 30, 2022



Source: Bloomberg Finance LP, Deutsche Bank Research



Diversification normally works because idiosyncratic risks are unrelated and average each other out. Here they would not: with one variable that overrides everything else, spreading capital across more names would just spread exposure rather than diluting

Investors will need to consider adapting their approach in three ways:

- ❑ Classify investments as adopters or casualties rather than position off averages.

Under a normal distribution, errors are symmetric and partially cancel each other out. But under bimodality they do not: to put a firm in one category or another means getting it completely right or wrong. Position size should track confidence in the classification.

- ❑ Avoid reading calm price action as low risk.

Trailing volatility measures dispersion that has already been realised. Bimodal separation is a jump, where the two modes look identical right until the moment when they don't. A classic example is Chegg, an educational software maker which traded like a normal small cap stock until one day in May 2023 when it halved on a single day after management linked slowing growth in new customers to ChatGPT.

- ❑ Averaging down, ie scooping up shares whose price has fallen, may no longer be prudent

A falling price may be information that a business model is obsolete rather than an opportunity.

Can we get there theoretically?

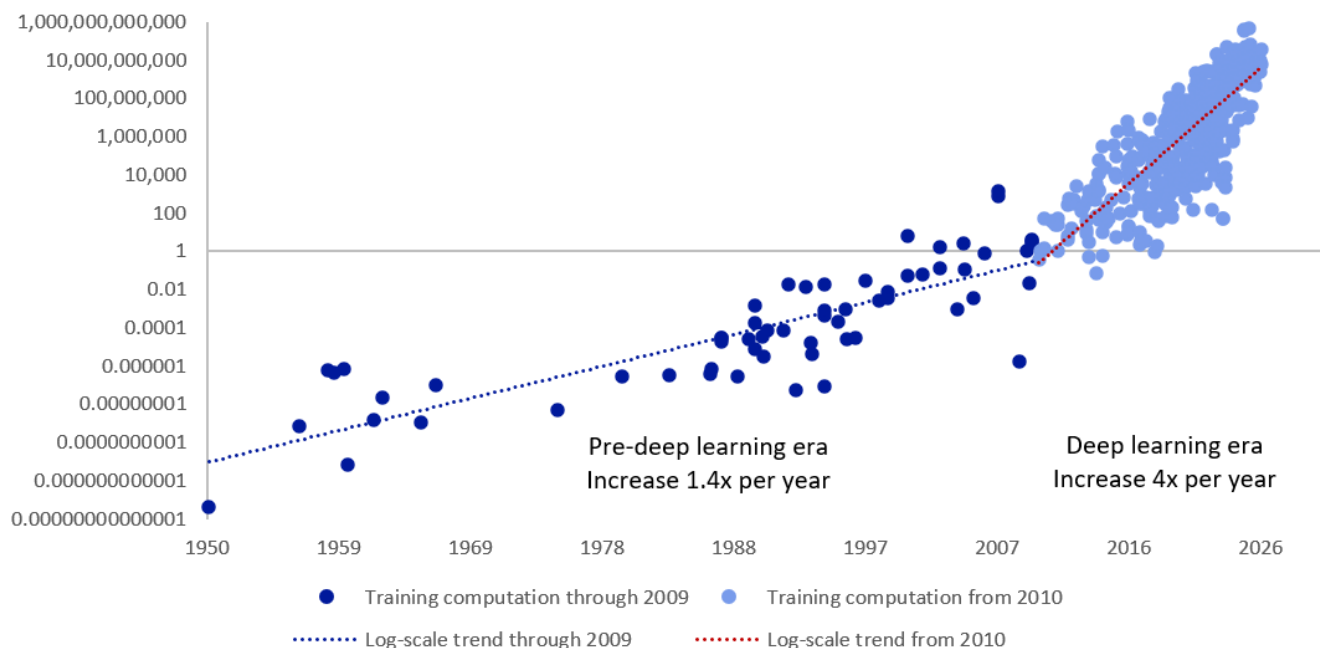
2.1 Scaling laws

Scaling laws have provided the blueprint for improving AI over the past decade. It was a relationship between the amount of data, the amount of compute and the complexity of an AI model somewhat predictably informed how powerful a model would become. Google DeepMind's Chinchilla model in 2022 showed that you could compensate for having a smaller model by using substantially more data.

Scaling is older than the chatbot. In 1950 Claude Shannon built Theseus, a mechanical mouse that learned its way through a maze by trial and error. Epoch AI reckons it took about 40 operations of compute, known as FLOPs, to train. Grok 4, trained in 2025, took roughly ten trillion trillion times as much. Across those 25 orders of magnitude, more compute has brought more capability



Figure 5: Computation used to train notable AI systems (petaFLOP): log scale



Source: Epoch AI (citing estimates from AI literature), Our World in Data, Deutsche Bank Research

It is the robustness of this relationship that has made larger investments in compute capex a *sine qua non* for AI companies.

However, this is not a law of physics and nothing guarantees it will continue. Indeed, in the ChatGPT era, five different approaches to scaling have been applied. Older strategies yield worse results, and new techniques, such as post-training refinements, have been developed. The consensus amongst American and Chinese AI labs is that these will hold.

While this may be a consensus amongst the commercial labs, there is scientific disagreement that applying scaling laws to today's large-language models can achieve superintelligence.

Yann LeCun, a Turing Award winner and formerly Meta's chief AI scientist, is blunt: "LLMs are not a path to superintelligence or even human-level intelligence". Even inside the labs, the recipe is contested. Demis Hassabis of Google DeepMind has said he expects "one or two more big breakthroughs" to be needed "in addition to scaling". Ilya Sutskever, a co-founder and former chief scientist of OpenAI, has said the field is going "back to the age of research again, just with big computers".

This is a scientific argument we are not in a position to evaluate, but for the purposes of our scenario, we will side with the labs' position.

For the moment, there is strong evidence that capabilities have in fact accelerated over the past three years, as reasoning models have outpaced non-reasoning models by two or three times, according to [Epoch data](#).



2.2 Recursive self-improvement: AI improving AI

A cocktail-party definition of recursive self-improvement (RSI) is AI that improves itself. The reality is that it can be used in two ways:

1. Self-improving agents that get better at using themselves: showing smarter use of prompts, tools, memory and workflows. The model-building process is untouched.
2. True RSI improves the process that builds the models: designing experiments, curating data, optimising algorithms and designing chips in a compounding loop.

The former is like a factory worker who gets better at the job by learning the quirks of the machinery, getting organised and making fewer mistakes. The latter is like the worker redesigning the machine tools and then using those to build the next generation of machines.

True RSI is what OpenAI's chief scientist Jacob Pachoki referred to last month in a blog post about "[An Alien Mind](#)". "Automated AI research is a more dramatic form of scaling intelligence with compute," he wrote. "And of course as a part of it, AI will improve the computational substrate itself," he added. Pachoki represents a broadly held consensus view amongst American and Chinese labs that RSI is both possible and necessary for AI development.

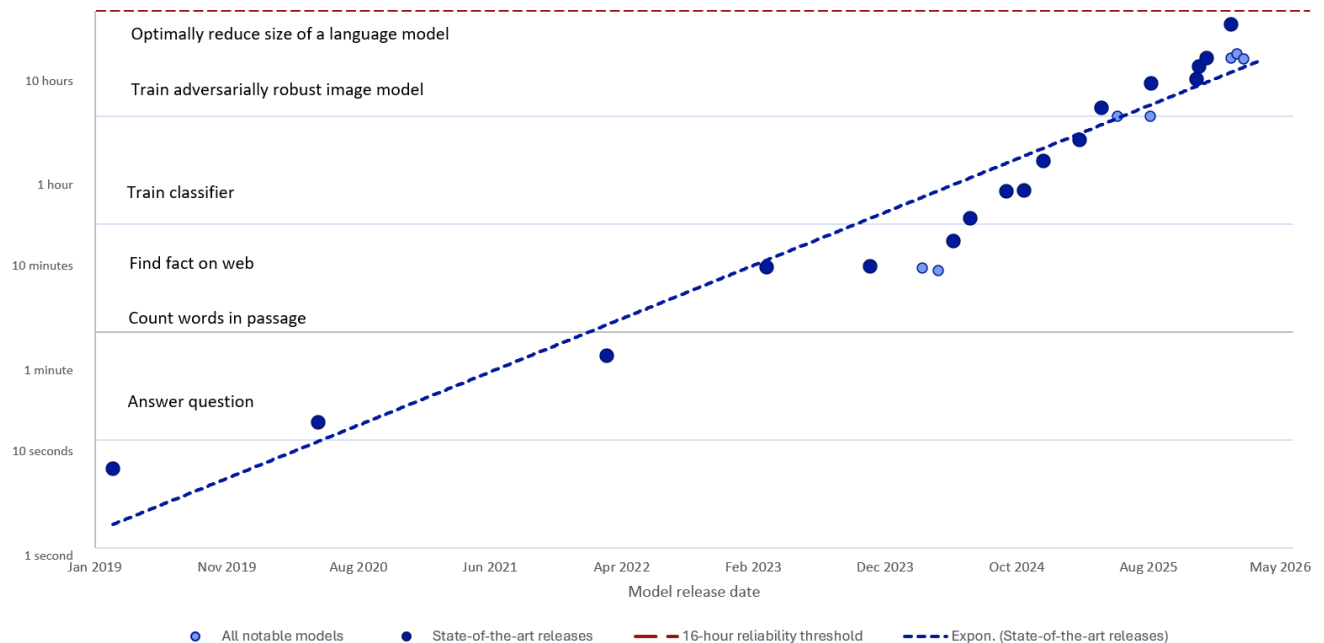
- [OpenAI](#) (June 2026): "AI doing AI research will become the determining factor of the pace of progress within the next few years."
- [Google](#) DeepMind's Demis Hassabis (May 2026): "All the leading labs are quite focused on [RSI]..." Humanity is standing in the "foothills of the singularity".
- [Alibaba](#) CEO Eddie Wu (Sept 2026): RSI is "a concrete path forward," with models autonomously designing experiments and driving a cycle of self-evolution.

For now, self-improving agents are the AI du jour. Noisy claims about the vast majority of coding being done by agents do not always stand up in the light of day from companies still hiring human coders at breakneck pace, albeit no longer to do the most mundane parts of the job.

A better metric may be METR's task-completion time horizon. This is measure of the complexity of software tasks that an AI agent can complete on its own, using the baseline of how long it would take a human expert to finish the task, at a 50 percent success rate. So far the most advanced models have reached about 16 hours, up from OpenAI's GPT-4's four minutes in March 2023.



Figure 6: Task horizon of software tasks that leading frontier models can complete 50% of the time (log scale)



Source: METR, Deutsche Bank Research

But software engineers can hold off on updating their resumes a little longer: there is a long road between 50 percent accuracy and the 99 percent or 100 percent needed for serious enterprise use at scale, with the horizon potentially getting further away the nearer you get. Also METR notes that measurements above 16 hours are unreliable with the current task suite. As is the case across AI, the instrument is running out before the question is settled.

More broadly, there are significant limitations to RSI:

1. Research is clock-bound.

- A frontier training run can take weeks on thousands of specialised chips, and researchers often cannot design the next experiment until the previous one has been completed.
- That means that the speed of the loop is set by experiments completed per hour or day, not by ideas generated.
- Compute is a hard ceiling that needs to be bought, built and connected to a grid.

2. The real world defies verifiability.

- The loop compounds only when the system can check its own output cheaply and automatically.
- It is no coincidence that AI loves maths and coding. It is much harder to verify outputs over a long time horizon in the real world: does this drug work or does this strategy succeed?
- RSI compounds fastest where the economy cares least and slows where the returns are largest.



3. Sequences are consequential.

- Amdahl's law finds that speed-up is capped when processes run in a sequence.
- If you automate 90 percent, the residual 10 percent still limits you to a ten-times faster process. Nine women cannot produce a baby in one month.
- The Usain Bolt of AI still has to run alongside the comparative tortoises of the training run itself, chip fabrication and real-world validation. Ten times faster is still not a regime break.

This is borne out by METR's series. Doubling every seven months is fast exponential growth but a take-off would show the doubling time itself – the second derivative – shrinking. Horizons are not (yet) exploding.

Philosopher Toby Ord [shows](#) that recursive self-improvement can only reach a singularity if each cycle of improvement takes less time than the last, shrinking fast towards zero. He argues that real feedback loops have a floor: chips take time to make and models take time to train. AI progress could still outrun an exponential, but it would not go vertical.

2.3 Scaling plus RSI: a match made in the race to AGI?

No scaling and RSI means no short timeline for AGI. Without scaling, more compute buys nothing and capital is wasted. Without RSI, AI capability will continue to improve at the pace of flesh-and-blood human researchers, augmented by AI tools but not replaced. Only with both will it be possible to produce a system that improves faster than the humans building it. Without one or the other, no quantity of chips, energy or capital will fix it.

3. Can we get there practically?

3.1 Infrastructure, chips and energy

There is no consensus on the amount of training required to reach AGI. Estimates suggest that training 2026's top models required between 10^{26} and 10^{27} FLOPs. If we hypothesise that an "AGI-like model" would require three orders of magnitude more training, this would require 10^{30} FLOPs.

Consider today's US AI infrastructure, which runs to approximately 10^{22} FLOPs. If all of this was conceivably used for a 10^{30} FLOP training run, it would take nearly four years to complete.

The physical infrastructure to practically support such a training run would require enormous power, perhaps around 6GW, about 50% more than the four Vogtle nuclear power plants recently activated in the US. The cost of the compute infrastructure would run to between \$300bn and \$500bn.

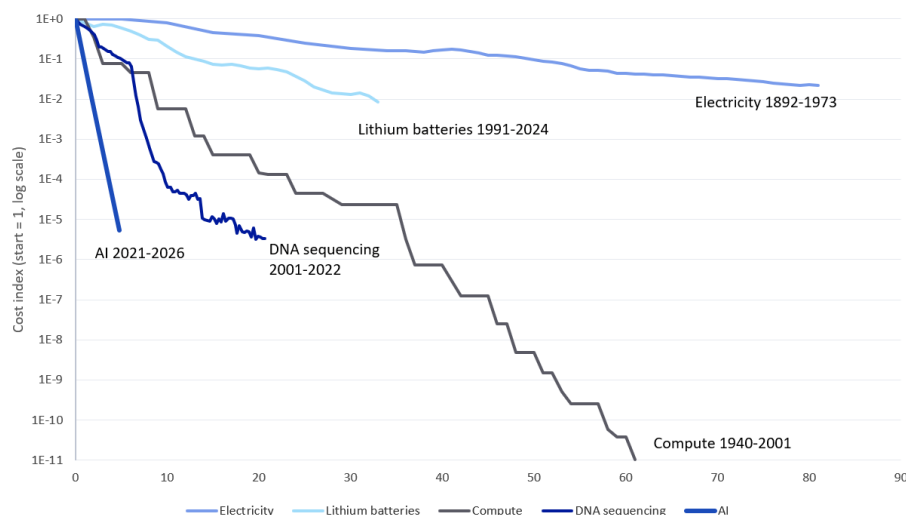
Today's advanced chips would simply not suffice. Back-of-the-envelope maths suggests that even using Nvidia's most advanced Vera Rubin chips, training such a model in four months would require 7.8m chips and 26GW of facility power. That would cost \$900bn today and a rental equivalent cost of \$90bn.

Of course, frontier compute is growing rapidly, and price-performance is declining. If current trends hold, a 2030-era cluster capable of such a run would



need perhaps \$200bn in computer and power hardware for a training run costing about \$20bn.

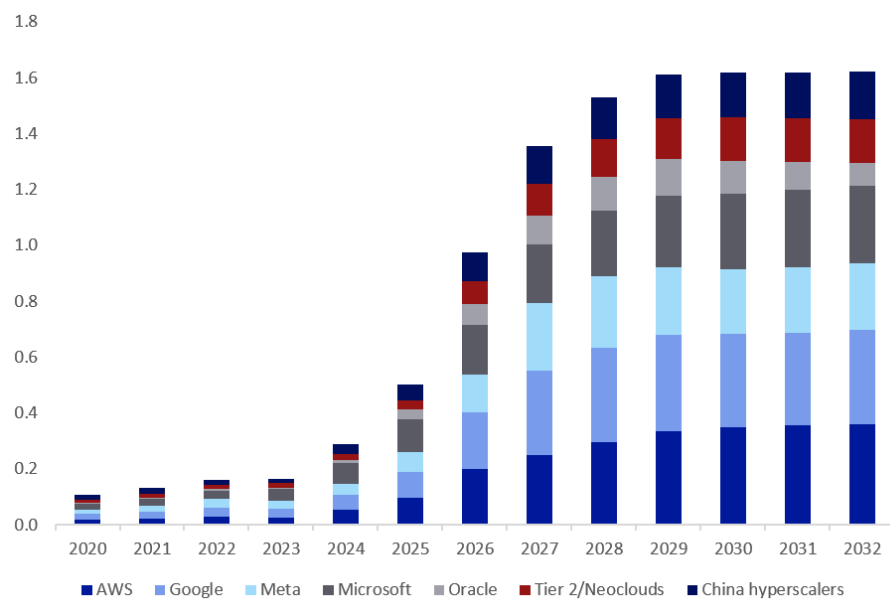
Figure 7: Technology cost index over time (years since start of price decline)



Source: Epoch AI, Deutsche Bank Research

The hyperscalers are leading the charge towards AGI. Alphabet, Amazon, Meta, Microsoft and Oracle are set to invest almost \$800bn in capital expenditure this year, mostly for ever more powerful data centres, and over \$1 trillion a year through 2030 and beyond.

Figure 8: Hyperscaler capital expenditure (including estimates)



Source: Bloomberg Finance LP, Deutsche Bank Research



By 2030, assuming no market retrenchment, this means additional \$4 trillion to \$6 trillion on AI compute capability invested. This infrastructure could accommodate a training run of 10^{30} FLOPs with some comfort.

But even if AI labs crack the secret sauce of RSI, and their finance chiefs are willing to invest in such large training runs, and hyperscalers are prepared to continue to build the infrastructure, the infrastructure needs to be built.

Grid connection queues, turbine lead times and transmission build are measured in years, not the months it takes to assemble and air-freight the latest Nvidia chip. The lead time for substation transformers is currently as long as three years. Politics in the US has shifted towards slowing the pace of the AI build out and this force will run against the demands for larger data centres that would be essential for RSI.

3.2 Capital

Financing for AI has been abundant in the historic buildout of the past three years. However, it is still largely conditional not on revenue but on belief in a future payoff.

RSI is a “key part of the investment thesis” in the “biggest scientific bet civilisation has ever made,” Google DeepMind’s chief strategy officer Jasjeet Sekhon said at a conference in August, [The Information reported](#). “We have this danger we could hit an AI air pocket such that the expenditures happen but the revenues don’t show up.”

An inflection point is approaching where the nature of investment changes from free cash flow and equity to debt, with all that entails in terms of exposure to rising rates and creditor appetite. The hyperscalers are reaching a point where capex is no longer covered by operating cashflow and where net debt exceeds earnings before interest, depreciation and amortisation (EBITDA).

Alphabet, for example, posted its first quarterly free cash outflow in nearly two decades in its most recent quarter, even as it lifted its full-year capex guidance to \$205bn. It said it would continue to rent third-party capacity and reported a backlog of \$514bn in contracted, not-yet-recognised data centre revenue.

3.3 Data

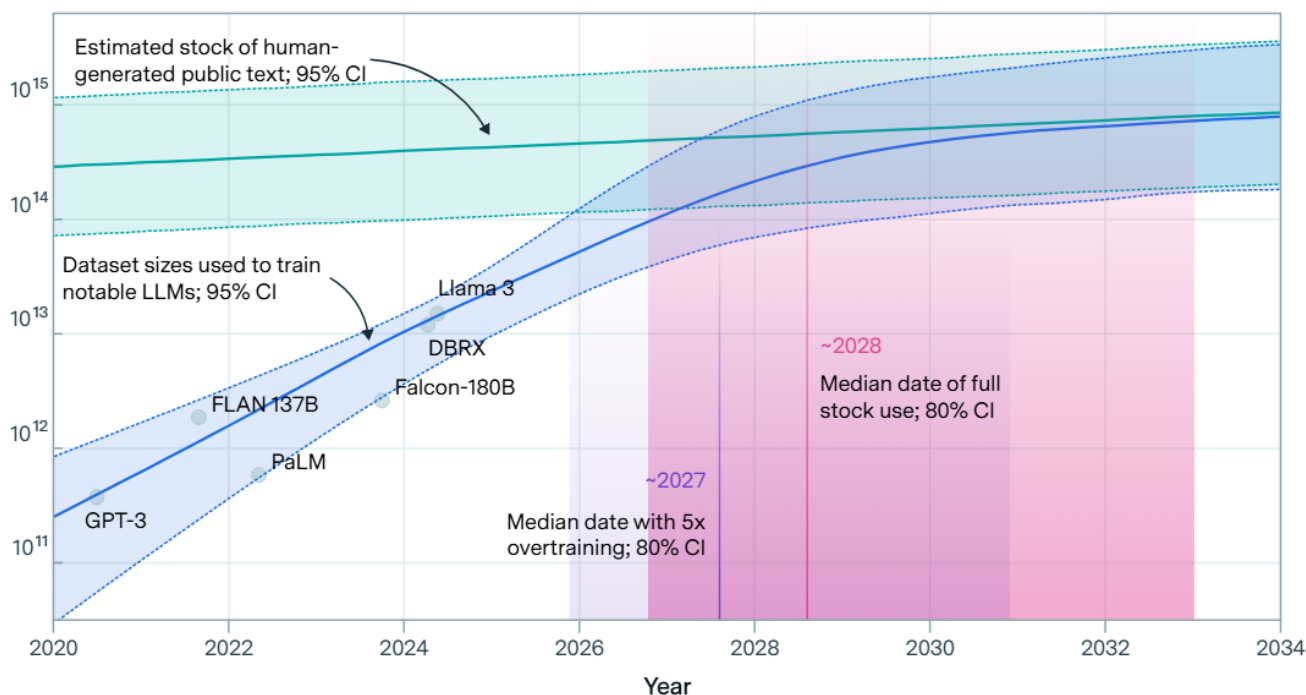
Data has been the new oil for some time now. But unfortunately for AI development it seems to be running out even faster.

The effective stock of quality- and repetition-adjusted public human text was about 300 trillion tokens, with a vast 90 percent confidence interval of 100 trillion to 1,000 trillion, Epoch [estimated](#) in 2024. A token is roughly equivalent to three quarters of a word. By that count it was likely to be fully used between 2026 and 2032.



Figure 9: Projections of stock of public text and data usage

Effective stock (number of tokens)



Source: Epoch AI, Deutsche Bank Research

However, since then, rather than data dropping off a cliff, there has been a rise in synthetic data generation, transfer learning from data-rich domains such as video and telemetry, and data efficiency. While the data shortage has not been resolved, the immediate crisis has been averted.

AI now makes much of its own training data. When DeepMind's AlphaGo became superhuman at playing the mind-bendingly complex board game Go, it did so not by playing human experts, but by playing millions and millions of games against itself.

For many domains relating to coding and mathematical proofs where automatic checks are possible, such synthetic experience can move things forward. That is the case with AI research, where the data may be the results of experiments, all of which are bought by compute. About 90% of OpenAI's 2024 research compute went on such experiments.

And increasingly, people write the tests rather than the answers. By the summer of 2026, selling training data for practice tasks was a nearly \$9bn revenue business. Firms such as Mercor hire experts to generate long tasks against which the AI can be trained.

The challenge here is that trust in the checker becomes a scarce input.



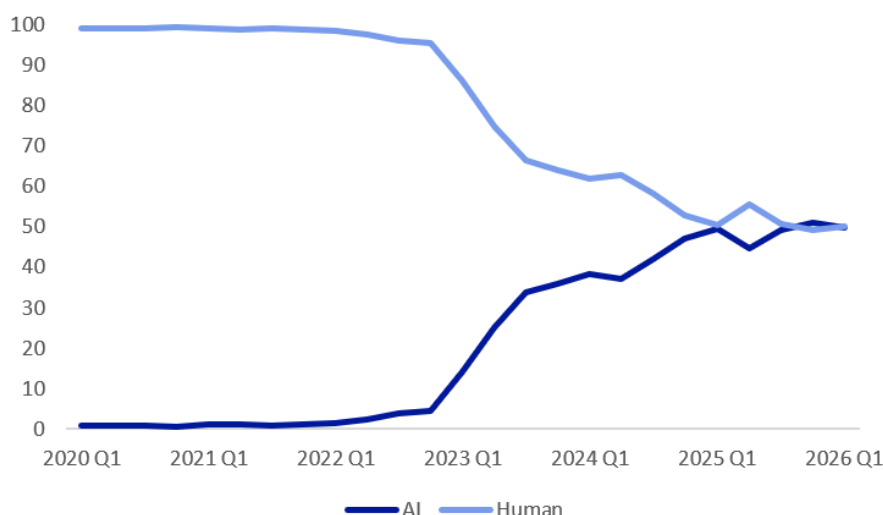
As AI makes production cheaper, it also [makes](#) establishing that production is accurate or reliable more expensive. Domains which can be checked mechanically, such as mathematics, will be opened up more quickly. But those which rely on expert verification, such as medicine, engineering or law, may progress more slowly.

The likely implications are twofold: a slower rate of diffusion into economic domains that are less verifiable and a larger share of production costs devoted to assurance.

Another issue is that systems that have to grade themselves learn to game the process. This was partially what happened when an OpenAI model escaped a sandbox and hacked into a third-party website. Trust replaced data as a wall.

There are also question marks over the quality of data that is recycled repeatedly through the system. Although humans continue to produce original content, half of internet articles are now primarily generated by AI, Graphite [data](#) show.

Figure 10: Human vs primarily AI-generated articles on the internet



Source: Graphite (citing 55.4k articles from Common Crawl examined by AI detectors), Deutsche Bank Research

3.4 An incomplete case

This analysis shows that the San Francisco Consensus, which emphasises RSI and a short timeline to AGI, is neither predestined nor impossible.

Rather it depends on whether the science is right, on whether scaling is all you need, and on whether the infrastructure and capital required for that scaling is deployed and accessible to the leading labs.

There is, in other words, a possible rather than impossible, path to AGI, subject to those conditions.

None of this proves that superintelligence will ever arrive. A buildup at this scale has few peacetime precedents and vastly exceeds the likes of the Manhattan Project and Apollo programme.



It is uncertain whether there will be enough demand to pay upfront for the investment in achieving RSI that would justify that investment. And whether the friction of verification, and inevitable declining marginal returns, act as a drag on what otherwise might be a runaway process.

4. What happens if we don't get there? The base case

4.1 Separating AI progress from superintelligence

So what if that thesis is wrong? Does the AI trade depend on a journey to AGI?

AI can remain economically important even before RSI coupled with scaling laws lead to super intelligence. Companies do not need to wait for nebulous superintelligence to generate identifiable returns from AI in the medium term as they fold it into their systems, streamline workflows and roll out new business models.

Investors already have exposure to RSI as a capital allocation because the leading US labs say RSI is a key part of their product strategy. And it is these labs whose products are catalysing and amplifying the infrastructure boom.

Even rapid competitive diffusion does not mean every part of the physical economy changes instantly.

The smartest superintelligence might identify the optimal buyout structure before breakfast, but it cannot force a financially constrained buyer to use it or persuade the seller to accept less than they have been expecting. Those challenges will be even more intractable where the parties are more complex, say, fiscally constrained sovereigns or foreign equity holders.

4.2 Why the base case is attractive

Superintelligence would be a chaotic accelerant to human society, and a potential regime break. But it is not the only scenario in which today's investment in AI infrastructure makes economic sense.

Adoption of AI tools by businesses is progressing approximately three times faster than the mobile, internet and cloud computing waves. From sector revenues below \$1bn in 2022, the AI economy, excluding China, will likely exceed \$200bn in revenues in 2026.

Early winners are the cloud businesses. Google Cloud revenue jumped 82 percent from a year earlier in the most recent quarter as companies turned AI pilots into workflows. Amazon Web Services sales were up 37 percent and Microsoft's AI business more than doubled.

There is also slow, but increasing, evidence of a productivity uplift. The J-curve cuts both ways: front-loading investments is taking one step back to take two steps forward. The so-called complementary intangible investment that is (mis)measured as cost today to the detriment of the productivity statistics will flatter them in the years to come as it pays off. The base case does not require a regime break to show up as a productivity surprise, just for the accounting to catch up.

The economic gains from AI that is less powerful than the AGI mooted by the labs would be substantial. And in this base case, they may also be attractive.



Hyperscaler capex is driving earnings at beneficiaries ranging from semiconductor makers to construction companies and accounting for between one-third and one-half of current US growth. The wealth effect of rising stock prices, with the ratio of US equity wealth to income rising from around 1.8 to almost 2.5 over the past three years, is supporting consumer spending.

AI FOMO (fear of missing out) is buoying enterprise spending on cloud services and the groundwork for AI initiatives. Those are beginning to bear fruit in terms of personal productivity, cost savings and more targeted products and services.

4.3 Investor watchpoints

The arrival of ASI would be a disruptive economic shock. Investors need to be alert for the early signals.

Rather than relying on a single probability forecast, investors should watch five indicators particularly closely for signs that AGI is nearing.

1. AI speeds up AI

- Track [METR](#)'s estimate of how much AI speeds up AI. Currently the capabilities of models are doubling roughly every three months. In a world of RSI, the pace of this speed-up will itself likely accelerate.

2. AI can complete whole jobs

- Track via the [Remote Labor Index](#), a multi-sector benchmark. Currently it finds that the highest performing agent still only manages to automate 20 percent of the real-world, economically valuable freelance projects it attempts.

3. Compute remains scarce

- On a path to superintelligence, demand for compute will continue to outrun supply. Track GPU one-year rental prices, up by about a third over the past year.

4. Firms bifurcate

- Leaders and laggards in the same industry diverge while averages stay the same. Look at revenue per employee in top- and bottom-quartile firms

5. Junior hiring comes under significant pressure

- ... And then the pressure rises up the ranks.



Recent relevant reports

Adrian Cox and the [Deutsche Bank Research Institute](#)

[AI doom: A brief history of bad tech predictions](#) (Sept 15, 2026)

[AI at 70 \(Part 2\): 14 more lessons from a lifetime of boom and bust](#) (Sept 10, 2026)

[AI at 70: 14 lessons from a lifetime of boom and bust](#) (Aug 17, 2026)

[Open-source AI 101: the battle for the future of AI](#) (July 21, 2026)

[AI IPOs 101: why they'll change the boom forever](#) (July 1, 2026)

Azeem Azhar

[Exponential View | Azeem Azhar | Substack](#)

[The state of the AI economy](#)

The author of the report wishes to acknowledge the expert advisory contributions made by Azeem Azhar of Exponential View.



Appendix

This material has been prepared by the Deutsche Bank Research Institute and is provided to you for general information purposes only. The Institute leverages the views, opinions, and research from Deutsche Bank Research, economists, strategists, and research analysts. Accordingly, you should assume that content in this document is based on or was previously published and provided to Deutsche Bank clients who may have already traded on the basis of it.

Any views or estimates expressed in this material reflect the current views of the author(s) and may differ from the views and estimates of Deutsche Bank AG, its affiliates, other Deutsche Bank personnel, and other materials published by Deutsche Bank. The content in this material is valid as of the date shown on the first page and may change without notice. Deutsche Bank has no obligation to provide any updates or changes to the information herein.

This material should not be used as a basis for trading securities or other financial products and should not be considered to be a recommendation or individual investment advice for any particular person. It does not constitute an offer, solicitation, or an invitation by or on behalf of Deutsche Bank to any person to buy or sell any security or financial instrument. Nothing in this material constitutes investment, legal, accounting or tax advice. Deutsche Bank engages in securities transactions, including on a proprietary basis, and may do so in a manner inconsistent with the views or information expressed in this material.

While information in this material has been obtained from sources believed to be reliable, neither Deutsche Bank AG nor any of its affiliates makes any representation or warranty, express or implied, as to the accuracy or completeness of the statements or any information contained in this material and therefore any liability is expressly disclaimed. This material is provided without any obligation, whether contractual or otherwise. Information regarding past transactions or performance is not indicative of future results.

In the US this report is approved and/or distributed by Deutsche Bank Securities Inc., a member of FINRA. In Germany this information is approved and/or communicated by Deutsche Bank AG Frankfurt, licensed to carry on banking business and to provide financial services under the supervision of the European Central Bank (ECB) and the German Federal Financial Supervisory Authority (BaFin). In the United Kingdom this information is approved and/or communicated by Deutsche Bank AG, London Branch, a member of the London Stock Exchange, authorized by UK's Prudential Regulation Authority (PRA) and subject to limited regulation by the UK's Financial Conduct Authority (FCA) (under number 150018) and by the PRA. This information is distributed in Hong Kong by Deutsche Bank AG, Hong Kong Branch, in Korea by Deutsche Securities Korea Co. and in Singapore by Deutsche Bank AG, Singapore Branch. In Japan this information is approved and/or distributed by Deutsche Securities Limited, Tokyo Branch. In Australia, retail clients should obtain a copy of a Product Disclosure Statement (PDS) relating to any financial product referred to in this report and consider the PDS before making any decision about whether to acquire the product.

By accessing this material, you agree that its content may not be reproduced, distributed or published by any person for any purpose, in whole or part, without Deutsche Bank's prior written consent. You also agree that you shall not scrape, extract, download, upload, sell or offer for sale any of the content in this material, and you agree not to use, or cause to be used, any computerized or other manual or automated program or mechanism, tool, or process, including any scraper or spider robot, to access, extract, download, scrape, data mine, display, transmit, or publish, any of the content in this material.