THEIA FINANCE LABS

OXFORD SUSTAINABLE FINANCE GROUP

UNIVERSITY OF OXFORD

Silicon Boundaries

Understanding, anticipating, and managing the societal instability that can be caused by AI and related technologies

By: Dr Jakob Thomä, Dr Ben Caldecott

Foreword

In September 2025, we developed the concept of Silicon Boundaries. Since then, we have consulted widely, tested and refined the idea through multiple iterations, convened a dedicated stakeholder workshop, and conducted an extensive review of the relevant literature.

Dr Jakob Thomä

Dr Ben Caldecott

This report marks the public launch of the concept as we understand it today. It is not the end of the conversation; it is the beginning. Our aim is to open a joint conversation, not only about the existence of Silicon Boundaries, but about what they are in practice, and, once established, how we can avoid crossing them or, where we have already done so, how we can return to safety.

The intellectual origin of the concept is clear: the idea of planetary boundaries in environmental science. We have spent our careers translating those boundaries into implications for financial markets, the real economy, and policymakers. We draw on that experience here.

We must be clear, however, that there are real differences. Not least because these technologies offer significant beneficial opportunities when deployed well. We aim to highlight those in this report and in our wider work.

There are also genuine differences in how these boundaries operate, the literature that informs them (primarily social science rather than environmental science), the ways in which we contribute to crossing them, and the means available to mitigate the risks.

As with planetary boundaries, our role is to build on and translate the science for decision-makers who have the tools to make societies more resilient: policymakers, investors, banks, companies, and the public.

These risks cannot be ignored, but nor are they inevitable. What we do now will determine whether we cross Silicon Boundaries, and whether there is a way back if we do.

Evidence of Societal Risks

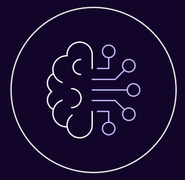
There is growing evidence that a set of technologies powered by compute and silicon create societal risks…

Fig 1 - newspaper clippings and images illustrating societal risks from technology

Silicon Boundaries

Compute-Powered Technologies

While the products and services underpinning these risks are varied in their application and use case, they all have something in common: **they realise their potential via the mobilsation of silicon-based computing power.**



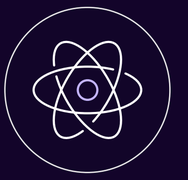
LLMs

Social media

Screen use

Autonomous Agents

Robotics



Quantum computing

Silicon Boundaries

Defining Silicon Boundaries

As compute scales and with it potentially societal risks, we argue that these technologies could cross (and may already have crossed) Silicon Boundaries, temporarily or even permanently, with significant implications for economic, financial and political systems at different scales, from national to global.

Crossing Silicon Boundaries means we cross societal boundaries unsustainably. The concept mirrors the notion of planetary boundaries in environmental and climate science.

Silicon Boundaries are a set of thresholds where the benefits of AI and related technologies are outweighed by their contributions to societal instability. Crossing these thresholds can trigger non-linear change and, in some cases, irreversible tipping points.

Planetary Boundaries *2025*

9 Boundaries assessed, 7 crossed

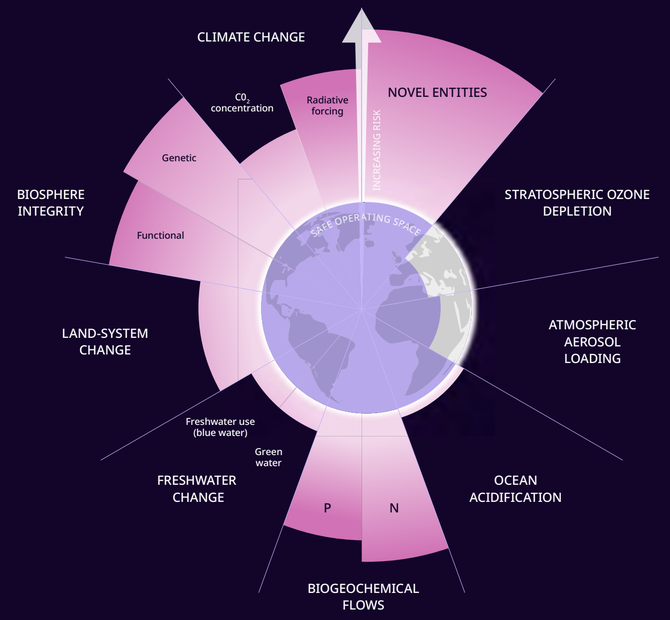


Fig 2 - Planetary Boundaries

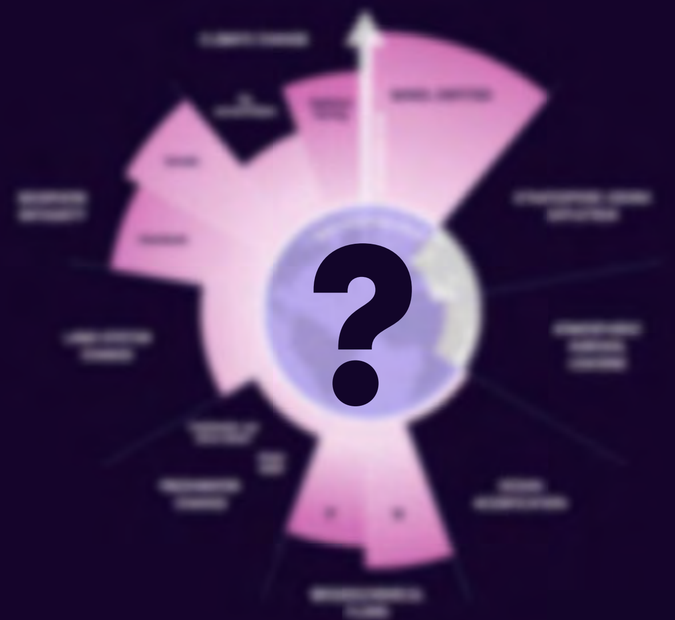


Fig 3 - Silicon Boundaries

Silicon Boundaries

The Silicon Bubble

Like with carbon, there is a certain amount of compute that is planned / forecasted that may become ‘stranded’ if society chooses to avoid breaching ‘Silicon Boundaries’.

A potential Silicon Bubble.

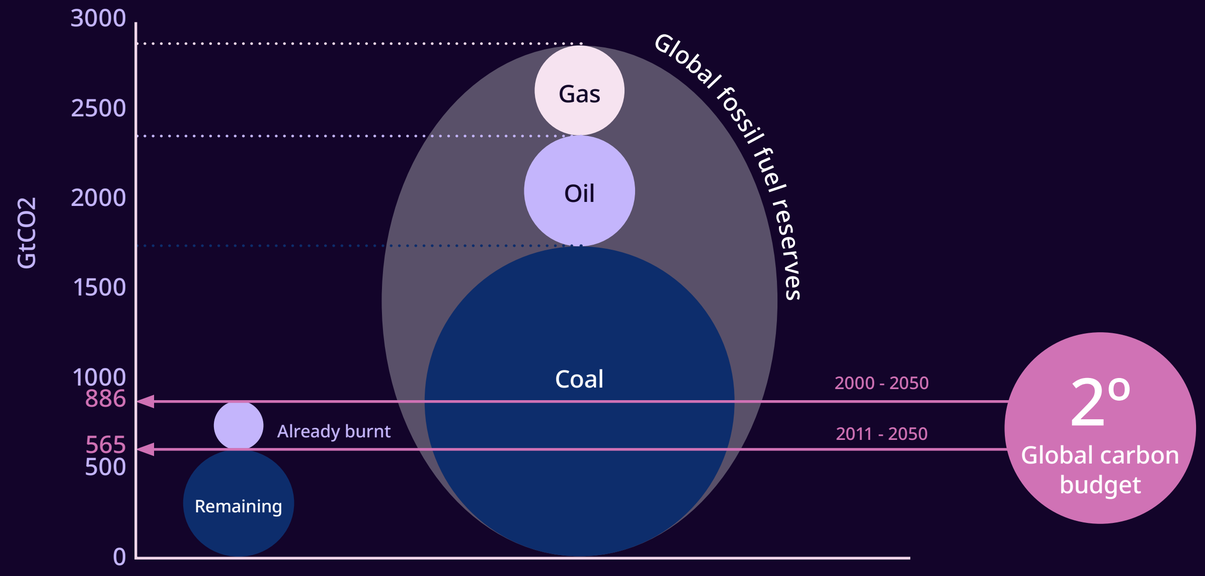


Fig 4 - Compute planned by ‘Silicon’-powered companies

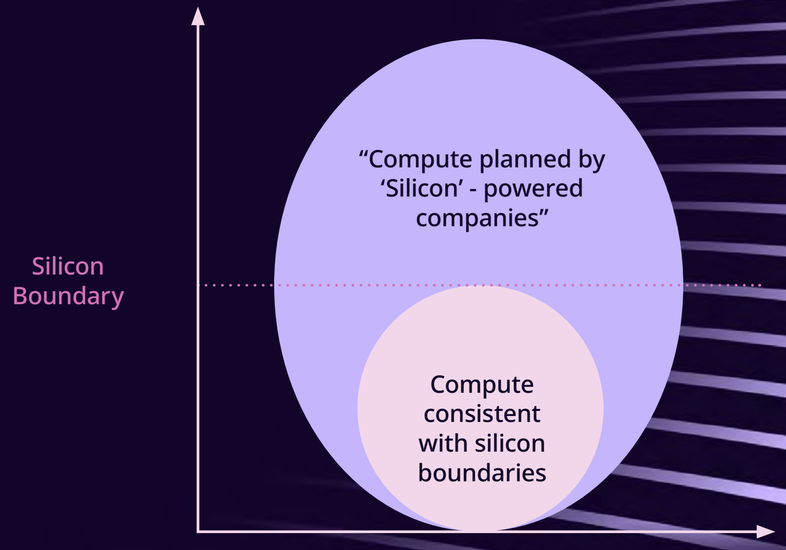


Fig 5 - Compute consistent with Silicon boundaries

Silicon Boundaries

Differences from Carbon

Unlike with carbon, the science is not settled as to the relationship between compute and Silicon Boundaries.

In addition, the relationship between the volume of compute and impacts is much less linear, driven (unlike with carbon) not just by the levels but also the nature of deployment, as well as the societal response.

Moreover, there is meaningful research on the potential benefits of compute, both today and in the long run.

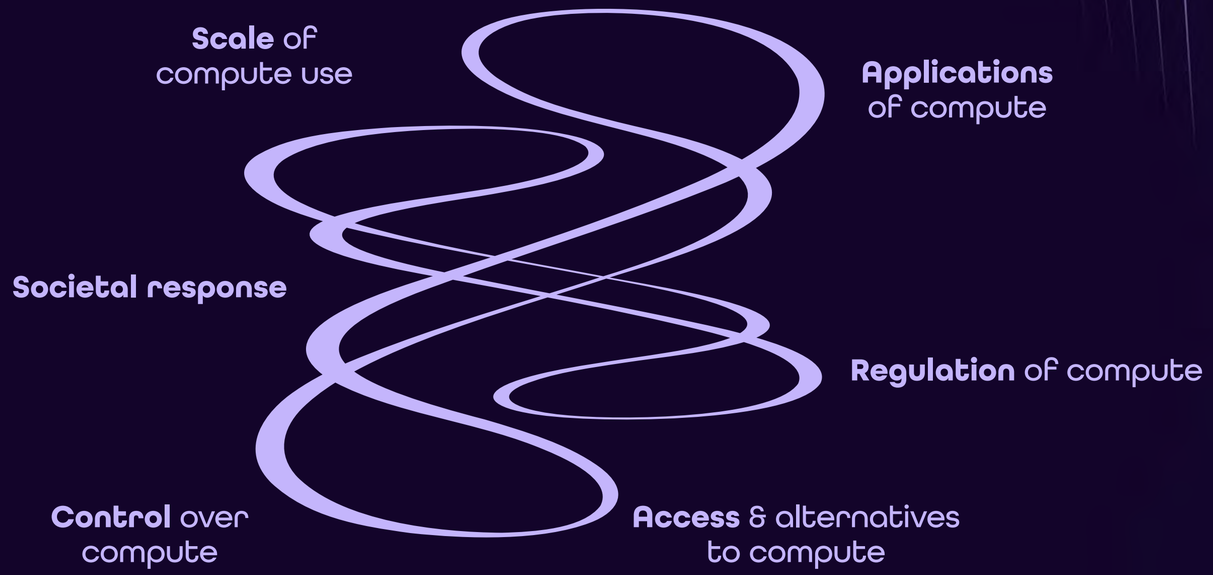


Fig 6 - Factors

Differences from Carbon

There are two crucial differences between ‘carbon’ and ‘silicon’:

1

We do not know at what level the aggregate costs of the Silicon economy (measured in e.g. welfare, GDP, ecosystem health) outweigh the benefits.

2

The Silicon Boundaries at which the costs outweigh the benefits are different for the ‘personal use’ and ‘societal use’ of the Silicon economy. Thus, limiting compute for ‘personal consumption’ may be desirable even as expanding compute for societal benefits in e.g. healthcare remains viable.

Social Science Foundations

Planetary boundaries were developed on the basis of an environmental science that teaches us what ecosystems – and humans relying on them – need to survive and thrive, and the threats that our actions represent to these ecosystems.

Our work in Silicon Boundaries will be developed on the basis of insights across the social sciences that teach us what societies – and humans relying on them – need to survive and thrive, and the threats that our actions represent to these societies.

Over the past six months we have on the basis of internal and external workshops and interviews identified a range of concepts that link to ‘silicon boundaries’ and over 30 potential boundaries.

Preliminary Framework

On the basis of the analysis described above, we have identified a preliminary set of nine categories of Silicon Boundaries and over thirty individual boundaries across these categories.

The Silicon Boundaries 0.1 framework will be further developed, fine-tuned and evolve as the science and the world evolves, and on the basis of expanding our engagement around this work.

Nine Boundary Categories

The v0.1 boundary framework identifies nine categories of Silicon Boundaries and over thirty individual boundaries across these categories. The boundaries span information integrity, social cohesion, democratic integrity, economic participation, physical and mental health, safety and security, financial stability, rights and consent, and environmental sustainability. Definitions and thresholds for each boundary are under preparation as part of the v1 output.

Individual Boundaries by Category

1

INFORMATION INTEGRITY & SHARED REALITY

Disinformation and synthetic media; erosion of epistemic commons; algorithmic amplification of falsehood; decline of trusted intermediaries; loss of shared factual basis for public discourse; large-scale infiltration of social and information systems by AI agents operating under fabricated human identities, degrading the assumption that one is interacting with humans and eroding the trust infrastructure on which markets, politics, and public discourse depend.

2

SOCIAL COHESION & CONNECTEDNESS

Screen-mediated displacement of in-person interaction; algorithmic polarisation and filter bubbles; youth mental health impacts; erosion of community institutions; loneliness and social fragmentation.

3

POLITICAL STABILITY

Computational propaganda; micro-targeted manipulation; AI-generated political content; “loss of human control” scenarios related to AI tools that can e.g. conduct information operations, manipulate markets, constituting novel forms of power that do not map onto existing categories of state, firm, or social movement.

4

ECONOMIC PARTICIPATION & DISTRIBUTIONAL STABILITY

Labour displacement and stranded skills; winner-takes-all market concentration; algorithmic wage-setting and surveillance; barriers to economic mobility; widening inequality between adopters and non-adopters.

5

PHYSICAL & MENTAL HEALTH

Addictive design and compulsive use; youth developmental harms; sedentary behaviour from screen displacement; diagnostic and treatment errors from AI in healthcare; erosion of clinician autonomy.

6

SAFETY, SECURITY, & THE STATE'S MONOPOLY ON VIOLENCE

Autonomous weapons and lethal decision-making; AI-enabled cyber-attacks; dual-use capabilities in biological and chemical domains; erosion of state capacity for law enforcement; non-state actors accessing advanced capabilities; AI agent proliferation beyond the capacity of human institutions to monitor, govern, or contain.

7

FINANCIAL STABILITY & CAPITAL MISALLOCATION

Overinvestment in compute infrastructure relative to societal returns; algorithmic herding and flash dynamics; concentration of systemic risk in a small number of technology firms.

8

RIGHTS, CONSENT, & HUMAN DIGNITY

Mass surveillance and erosion of privacy; biometric data collection without informed consent; algorithmic discrimination; deepfake exploitation; loss of meaningful human agency in consequential decisions; deployment of AI systems in architectures that are effectively irrevocable, including blockchain-embedded agents, decentralised autonomous organisations, and distributed systems with no single point of deactivation, such that society loses the capacity to reverse course, recall, or correct these systems even when harms become apparent.

9

ENVIRONMENTAL SUSTAINABILITY

Energy and water consumption of data centres; embodied emissions in semiconductor supply chains; electronic waste; rebound effects where efficiency gains drive increased demand; resource extraction for hardware production.

Key Questions

Many questions remain across the Silicon Boundaries, here are some:

What are the economic costs of crossing Silicon Boundaries, and how do they compare with the potential economic benefits of the technologies themselves?	Categories 4, 5, 7
What levels of screen time and the digitisation of social & public life are consistent with population-level mental and physical health?	Category 5
Where are the thresholds at which technology-mediated displacement of human connection, community, and consent undermines social functioning?	Categories 2, 8
Can these technologies push their own “Silicon” boundaries outward, or will they erode the societal conditions needed to deploy them safely?	Cross cutting
Do silicon boundaries differ across societies and populations, and what does that mean for how they are governed?	Cross cutting

Urgency Despite Uncertainty

The fact that the science is not settled on Silicon Boundaries does not mean these boundaries do not exist.

Nor does it mean we can afford to ‘wait and see’ to discover whether they are being crossed. The stakes are too high.

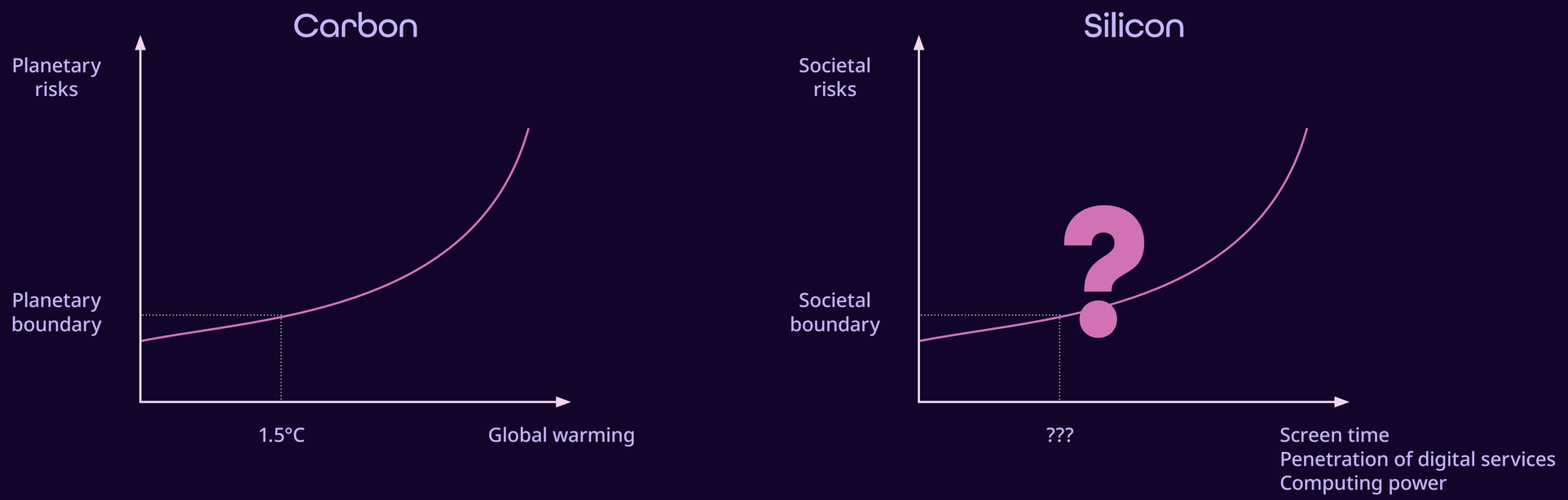


Fig 7 - Diagram comparing Carbon and Silicon

Tipping Points

Not only because of the costs, but also because, as with climate, crossing Silicon Boundaries may lead to irreversible tipping points from which society cannot recover.

Potential Impacts of Crossing Silicon Tipping Points

SOCIAL

Digitisation drives loneliness and reduces physical and mental well-being. Loneliness increases mortality risk by 26%: the equivalent health effect of smoking 15 cigarettes a day. A Silicon Tipping Point could thus have the equivalent health effect of turning the majority of the population into chain smokers.

POLITICAL

The media landscape is driving political polarization that undermines the foundation of liberal democracies. The Liberal Democracy Index has already seen its biggest decline since WWII. The Silicon Tipping Point could drive the collapse of liberal democracies in an environment of increased political violence.

ECONOMIC

Silicon could create similar monopolies to the “Seven Sisters” oil companies (controlling at one point roughly 85% of oil reserves), except with vastly greater control over economic activity. Silicon investment constrains capital for other productive investment. For labour and for countries, economic losers multiply.

ENVIRONMENTAL

Silicon resets demand projections for natural resources, reinforcing geopolitical imbalances between producers and consumers, and of course generating dramatic climate and environmental costs. Silicon could contribute to ecosystem collapse and climate tipping points.

Note: the four impact categories above (Social, Political, Economic, Environmental) cover only a subset of the nine boundary domains identified in the v0.1 framework. The boundaries also extend to information integrity and shared reality (Category 1), physical and mental health (Category 5), safety, security, and the state’s monopoly on violence (Category 6), financial stability and capital misallocation (Category 7), and rights, consent, and human dignity (Category 8). Tipping point dynamics in these domains are under analysis.

Three Potential Outcomes

Transition RISKS

We ‘pop’ the Silicon Bubble, with associated impacts on the valuation of the tech industry and the global economy. These ‘transition risks’ would suggest structural misallocation beyond potential short-term misallocation / mispricing as part of a cyclical growth dynamic.

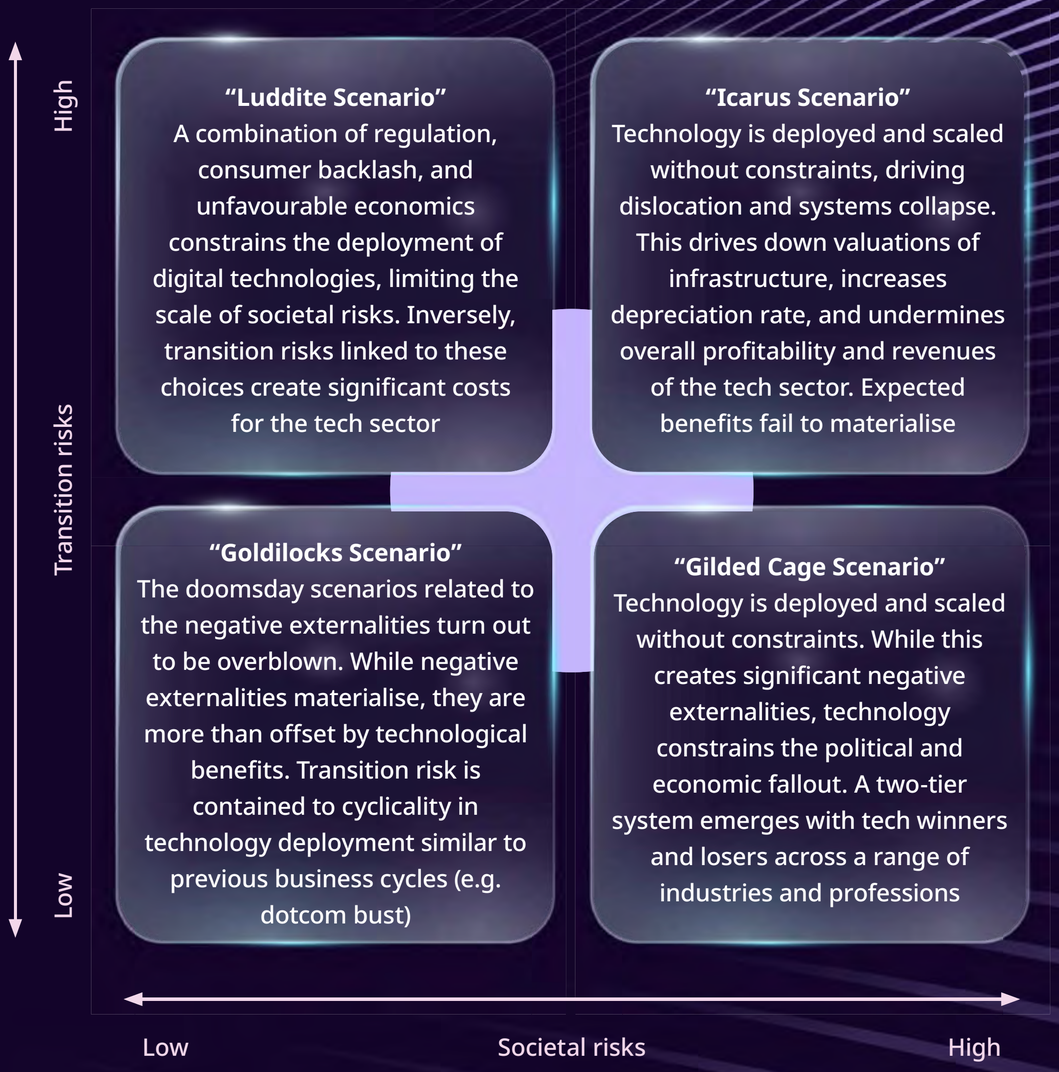
Societal RISKS

We breach Silicon Boundaries and trigger associated negative social, political, economic, and financial impacts.

...or technological progress resolves the tension.

NGFS Scenario Framework

The NGFS scenario framework also applies to silicon risks, with potential for high transition - high societal risks; low transition – low societal risks and dynamics in-between



Transition Risks if the Silicon Bubble Bursts

Potential **transition** risks if the Silicon Bubble bursts...

The market capitalisation of the technology sector is two to three times higher than the market capitalisation of the ‘carbon bubble’ companies at the time of the original Carbon Tracker publication.

We estimate that based on valuation, financial markets expect these companies to capture more than two-thirds of all stock market earnings growth over the next 5 years.

While there is a wide range of data centre capex estimates, Morgan Stanley estimates $2.9 trillion in data centre capex for 2025–2028.

There is no meaningful way to estimate costs related to litigation risks, but they can be existential threats to business models.

Societal Risks if the Silicon Bubble Bursts

Potential **societal** risks if the Silicon Bubble bursts...

Broad estimates of mental health costs already equate to ~2% of US GDP. This number could easily double under pessimistic scenarios

At global level, digital services will have a marginal impact on climate (<0.1°C warming), but local impacts (e.g. pollution, water use, electricity costs) may be significant and such impacts may scale as AI scales

The impact of political instability may reduce global growth by 10%, using historical literature as a reference point

Unemployment rates driven by AI, especially in the short run, could in a pessimistic scenario increase by 5–10%.

The Goldilocks Possibility

Unlike with climate, there is still an option for a ‘Goldilocks Scenario’ where compute scales in line with societal welfare.

It does not seem reasonable, however, to assume such a scenario will simply materialise without the collective efforts of governments, companies, consumers, citizens, and financial markets.

Technology-Optimist Lens

There is a technology-optimist lens through which the ‘Silicon Bubble’ narrative can be interpreted: one where artificial intelligence and other technological innovation mitigates negative externalities and perhaps even offsets them. In such a world, there is no Silicon Bubble. We call this the “Goldilocks Scenario”. While there is now broad scientific agreement that the negative externalities from carbon outweigh the benefits, the future of digital technology and AI is not yet written.

If technology delivers on its promise for progress and innovation, there may be short-term cyclical bubbles, but no systemic bubble, as negative externalities are mitigated by the very technologies that created them.

If technology does not deliver on that promise, the social tipping points will in all likelihood be at least equivalent to the planetary tipping points triggered by climate change, in terms of societal impact.

The Key Question

How can we build a future that stays within Silicon Boundaries while minimising the societal risks from a Silicon Bubble and maximising the social benefits of technology?

Societal Conditions for Responsible Technology

Like with any technological revolution before, the digital revolution will likely bring a new set of social, political, economic, and financial norms, each of which create new risks and opportunities. Delivering a ‘responsible technology’ future that respects Silicon Boundaries will not just depend on the technology itself, but the societal conditions in which this technology will or won’t flourish.

The real world recreational offer & cultural norms

Digital safeguards & boundaries

Economic opportunity

Robust political systems

Silicon Boundaries

Investor Roadmap: Carbon vs Silicon

To deliver on these conditions, investors face a more complicated roadmap than with carbon.

Carbon

GHG emissions cause...

global warming

Finance & engage on emissions reductions...

..to mitigate global warming

Silicon

Silicon Boundaries

Next steps for research

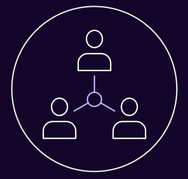
Research on quantifying societal boundaries

Develop scenarios, data, and models that measure their impact on companies & portfolios

Understand synergies between the Silicon Bubble and other sustainability themes

Next steps for investors

Implement scenario analysis to understand various impacts under ‘Silicon scenarios’



Ramp up capabilities for engagement with policymakers and companies

Develop portfolio allocation strategies that reflect silicon risks and opportunities and meaningfully differentiate between Silicon companies

Assumptions

The conclusions in this note are informed by the following assumptions.

1) Compute and the technologies powered by it (e.g. social media, LLMs, digital applications) have radically changed the structure of society.
2) These changes, like any other historical economic transformations, have enormous potential for good, but also represent significant societal risks.
3) As with climate change, society operates within a set of ‘safe’ boundaries that ensure the delivery of societal or ‘public’ goods (e.g. welfare, civil liberties, economic prosperity, environment).
4) The risks arising from silicon technologies and their applications may cause us to cross these boundaries.
5) We introduce in this note the concept of Silicon Boundaries: the point at which the scale and nature of compute powered by silicon leads us to cross societal boundaries.
6) We argue the dynamic identified here has a number of key similarities to the notion of ‘planetary boundaries’ in particular in the context of climate change and more specifically the ideas around a ‘carbon bubble’, for which we see similarities to a potential Silicon Bubble.
7) Preventing the crossing of ‘Silicon Boundaries’ requires one of two outcomes to materialise: a) We constrain compute either through regulation or consumer choice in a way that prevents us crossing these boundaries; b) compute develops a set of technological solutions that allow us to ‘shift’ the ‘Silicon Boundaries’ upwards in terms of the point at where they trigger systemic societal risks.
8) Both of these options are preferable to a future where we cross Silicon Boundaries. And so we cannot leave this up to chance.
9) Like with climate, there are four potential scenario narratives that describe the future in this space across the level of ‘transition’ and ‘societal’ (for climate: physical) risks.

Bibliography:

Abramson, B., Boerma, J. & Tsyvinski, A. (2024) ‘Macroeconomics of mental health’, NBER Working Paper No. 32354, National Bureau of Economic Research.

Angus Crawford (2023) ‘Molly Russell: Tech firms still failing after teenager’s death, says father’. BBC News.

Bertelsmann Stiftung (2025) ‘EuroStack: A European alternative for digital sovereignty’.

Caldecott, B., Clark, A., Koskelo, K., Mulholland, E. & Hickey, C. (2021) ‘Stranded assets: Environmental drivers, societal challenges, and supervisory responses’, Annual Review of Environment and Resources, 46, pp. 417–447.

Carbon Tracker Initiative (2013) ‘Unburnable carbon 2013: Wasted capital and stranded assets’.

Chatterjee, R. (2025) ‘Their teenage sons died by suicide. Now, they are sounding an alarm about AI chatbots’, NPR.

Coase, R.H. (1960) ‘The problem of social cost’, Journal of Law and Economics, 3, pp. 1–44.

de Groot, O.J., Bozzoli, C. & Alamir, A. (2022) ‘The global economic burden of violent conflict’, Journal of Peace Research.

De Agostini, L. & Giovanardi, M. (2024) ‘AI and Global Security: From Early Warning to AI-Assisted Diplomacy’, Centre for International Governance Innovation (CIGI).

Devi, K.A. & Singh, S.K. (2023) ‘The hazards of excessive screen time: Impacts on physical and mental health’, Journal of Education and Health Promotion.

Deniz, F. (2025) ‘The impact of AI and machine learning on conflict prevention’, Trends Research.

Dietz, S., Rising, J., Stoerk, T. & Wagner, G. (2021) ‘Economic impacts of tipping points in the climate system’, Proceedings of the National Academy of Sciences of the United States of America, 118(34).

Dorsett, R. (2013) ‘The effects of the Troubles on GDP in Northern Ireland’, European Journal of Political Economy, 29, pp. 119–139.

Healthcare in Europe (2025) ‘AI helps GPs detect lung cancer 4 months earlier’.

Jung, C. & Srinivasa Desikan, B. (2024) ‘Up to 8 million UK jobs at risk from AI unless government acts, finds IPPR’, Institute for Public Policy Research.

Max Planck Society (2025) ‘Digital media – A threat to democracy?’

Metz, C. (2025) ‘Anthropic Agrees to Pay $1.5 Billion to Settle Lawsuit With Book Authors’, The New York Times.

Mitre, J. & Predd, J.B. (2025) ‘Artificial General Intelligence’s Five Hard National Security Problems’, RAND Corporation.

Network for Greening the Financial System (NGFS) (2024) ‘NGFS Climate Scenarios Technical Documentation’.

Office of the Surgeon General (2023) ‘Our epidemic of loneliness and isolation: The U.S. Surgeon General’s advisory on the healing effects of social connection and community’, U.S. Department of Health & Human Services.

Our World in Data (2024) ‘Liberal Democracy Index.

Petracca, L. (2025) ‘Is AI to blame for the lousy job market?’, Investopedia.

Planetary Health Check (2025) ‘The Planetary Health Check – Our planet’s vital signs are flashing red’.

Reuters (2025) ‘AI and gravity-defying US GDP’.

Rockström, J. et al. (2009) ‘A safe operating space for humanity’, Nature, 461, pp. 472–475.

Rodríguez Salamanca, L. (2025) ‘Meta brought AI to rural Colombia. Now students are failing exams’, Rest of World.

Sevilla, J., Besiroglu, T., Cottier, B., You, J., Roldán, E., Villalobos, P. & Erdil, E. (2024) ‘Can AI scaling continue through 2030?’, Epoch AI Blog.

Telegrafi (2025) ‘Artificial intelligence weapons are changing the war in Ukraine’.

Tidy, J. (2025) ‘Nursery hackers threaten to publish more children’s profiles online’. BBC News.

Vlamis, K. (2025) ‘Anthropic cofounders say the likelihood of AI replacing human jobs is so high that they needed to warn the world about it’, Business Insider.

Wang, Q., Li, Y. & Li, R. (2025) ‘Integrating artificial intelligence in energy transition: A comprehensive review’, Energy Strategy Reviews, 57.

Silicon Boundaries

About Silicon Boundaries

Silicon Boundaries is a new venture by Theia Finance Labs, developed in partnership with the Oxford Sustainable Finance Group, focused on understanding the risks from the ‘silicon economy’ to public welfare and the achievement of climate goals.

Contact:

Theia Finance Labs
(Legal name: 2° Investing Initiative Deutschland e.V.)

c/o WeWork

Neue Schönhauser Allee 3-5, 10119 Berlin

Germany

Jakob@theiafinance.org

THEIA FINANCE LABS

OXFORD SUSTAINABLE FINANCE GROUP

UNIVERSITY OF OXFORD