JESUS COLLEGE
CAMBRIDGE

Global Issues
Dialogue Centre

Discussion Paper

# Global Digital Governance

Prof Mischa Dohler
FIEEE FREng
King's College London

# Executive Abstract

It can be argued that digital technologies and industries are inherently global in that they are capable of global reach, that network economies mean that they create greatest value when they are global, and that globalisation is likely to speed up innovation and enable access to the greatest number of people. On the other hand, current geopolitical trends, along with different priorities and values among different countries, security concerns, and the desire for local control all militate against global digital governance solutions.

This position paper thus examines important key questions to be addressed to overcome growing challenges. We note that with the emergence of three advanced technologies, prior established governance initiatives in e.g. AI ought to be revisited. Failing in doing so may lead to unpredictable emerging technology capabilities which could pose a hyper-exponential threat to society. We note that proper mechanisms of observation need to be established first which are able to detect and causally link emerging capabilities. Without such directly or indirectly observable oversight, the very notion of human identity could be challenged.

We thus propose the establishment of a Global Digital Observatory (GDO), analogous to the inter-governmental Panel on Climate Change. It could be a to-be-developed capability within already chartered bodies or become an internationally chartered body on its own. The GDO should not only cover AI but all emerging technologies and combinations thereof.

The GDO should be governed by the basic principles of responsibility and human-centredness. It will likely fail if constructed as a central entity and should rather rely on polycentric governance, representing a broad spectrum of global, regional and national stakeholders but most importantly the voice of citizens and society. It should reconcile the prevailing structural fragmentation between current geopolitically entrenched stakeholders. If executed properly, it could enable the next-generation era of industrial prosperity, i.e. Industry 5.0.

## 1. Introduction

Digital technologies and industries form an important part in the fabric of our modern society. A plethora of devices is connected 24/7 offering services which were unthinkable not even 10 years back. As a result, these technologies yield important societal impact which was proven to correlate with general well-being and GDP.[1]

The success of digital is due to the consolidation of three important trends, i.e. miniaturisation of devices, interconnectivity provided by the internet, and a massive reduction in cost due to scale. This was enabled by an inherently global ecosystem on both supply and demand sides, where the network economies mean that they create greatest value when they are truly global. Indeed, the globalisation observed over the past decades has sped up innovation and enabled access to almost every corner of our civilisation.

On the other hand, current geopolitical trends, along with different priorities and values among different countries, security concerns among important stakeholders, and the desire for local control all militate against global digital governance solutions. With this, the progress made over past decades is put into jeopardy.

*With focus on governance, this position paper examines the mechanisms which need to be put in place over the coming years to enable further growth and consolidation of emerging and yet-to-emerge technologies despite natural market commoditisation and recent international geo-political fragmentation.*

To this end, we aim to address important key questions in this report, such as:

- What are the most important areas in which a global governance system would make the

greatest contribution to the value creation potential and efficiency of digital industries?

- What are the potential costs and disadvantages in terms of the potential negative impact on value creation, innovation and efficiency of digital industries of a failure to establish a global governance system?
- What areas, especially given concerns such as privacy and security, are best handled at the national level?
- What can be achieved through cooperation between governments, industry organisations, or enterprises at a regional level?
- What kinds of institutions, alliances, agreements might be established as part of an international digital governance system?
- What might be the role of a "global digital observatory" (analogous to the Inter-governmental Panel on Climate Change)? How might it be established and operate?

To this end, the structure of this governance paper is as follows. In Section 2, we discuss some important patterns in digital which prove useful in understanding future trends. In Section 3, we deep-dive into the three most important and disruptive technologies. In Section 4, we combine these technologies and discuss the digital butterfly effect in which a small benevolent perturbation in the digital medium may cause significant distress globally. As a result, we argue in Section 5, there is a need for a "Global Digital Observatory". In Section 6, we then examine prior art and particularly if any established policy frameworks at global level is able to host such an observatory. Finally, specific recommendations on a possible governance and operating model are made in Section 7.

## 2. Patterns & Trends in Digital Technologies

Today's ubiquitous role of the internet and digital in general is the result of three fields which have evolved rather independently until a critical mass was reached. Specifically, the three fields are i) devices, ii) networks and iii) applications.

In more details, devices and the integrated circuits controlling these devices have followed a simple yet powerful law: Moore's Law. For decades, the prediction made by American engineer Gordon Moore in 1965 on the number of transistors per silicon chip holds true, i.e. it roughly doubles every year.[2] Whilst we have transgressed from transistors today, the law holds in its essence in that "the capability of the integrated circuits doubles roughly every year". That is a very aggressive exponential growth in capability which, however, was not enough to trigger the digital revolution of recent years.

Instead, a second powerful trend was required, i.e. the trend of an emerging global network, now referred to as the Internet. Its exponential success can be attributed to three factors: first, it evolved from a local area network (LAN) to a truly global network of interconnected computers, routers and switches; second, it runs on a very simple protocol, i.e. IP, which ensures that any machine can "talk" to any other machine; and, third, it transmits data which is encoded via standardised codecs. All this increased compatibility, lowered barriers of entry, enhanced scale and thus lowered costs. The internet now encompasses three internets, i.e. the fixed Internet (ethernet connected computers/cloud centres/etc), the mobile internet (2G-5G), and the things internet (aka the Internet of Things, i.e. bi-directional connection of billions of sensors and actuators to the internet fabric). The trends governing the mobile internet have been discussed in great depth in a companion policy brief on 6G.[3]

As with devices, the emergence of a connectivity fabric was not enough to trigger enormous growth. The explosion of the Internet as we know it today was due to the decoupling of hardware (devices and networking infrastructure) from software (operating systems and software applications). This transformation enabled scale and Maslow's hierarchy of needs created the "stickiness" defining today's apps.

The consolidation and intersection of these three fields yielded a beyond-exponential growth in capabilities and market capitalisation. For the first time, we have seen companies being valued above $1tr and billionaires assembling a

wealth larger than the aggregate of many nation states.[4] The evolution of the internet provides a rich learning field which we will use to understand the next wave of technologies, and how best to deal with them to ensure ethical and sustainable growth.

To this end, which other important fields are emerging today? Indeed, we are witnessing the consolidation and intersection of yet another three fields with an array of emerging applications which we have not been fully understood. Notably, the fields of i) artificial intelligence (AI), ii) robotics) and iii) quantum enable a future which is both exciting but also scary, *a future where AI yields automation, quantum yields scale, and robotics yields actionability.* It is important to understand what needs to be done today to ensure that the intersection of these domains will not cause irreversible damage to society.

## 3. Consolidating Technologies: AI, Robotics and Quantum

Without going into too much technical depth, it is instrumental to understand the role, potential and threat of each technology:

**Artificial Intelligence** has evolved over many decades and comprises any techniques that enable computers to mimic human intelligence through the use of logic, if-then rules, decision trees, machine learning, deep learning, etc. Machine learning is a subset of AI that includes statistical techniques that enable machines to improve tasks with experience. Deep learning on the other hand is an important subset of machine learning, and is composed of algorithms that train themselves to perform complex human-like tasks, such as image or speech recognition. The underlying technology principle of deep learning is the use of non-linear multi-layered neural networks which need to be trained before being used for inference.

AI's evolution has no sign of slowing down with a wide variety of tools being available to academia and industry alike. For instance, Tesla's object recognition software for their self-driving cars is based on Convolutional Neural Networks (CNNs);[5] Amazon's natural-language processing Alexa is based on Recursive Neural Networks (RNNs).[6] An exciting new class of Generative Adversarial Networks (GANs) have emerged in 2014[7] which – for the first time – allows for the creation of "original" content by a machine. GANs are the underlying AI technology for deep-fakes,[8,] i.e. the ability to create artificial images, videos and voice-overs of any existing or fictional person.[9]

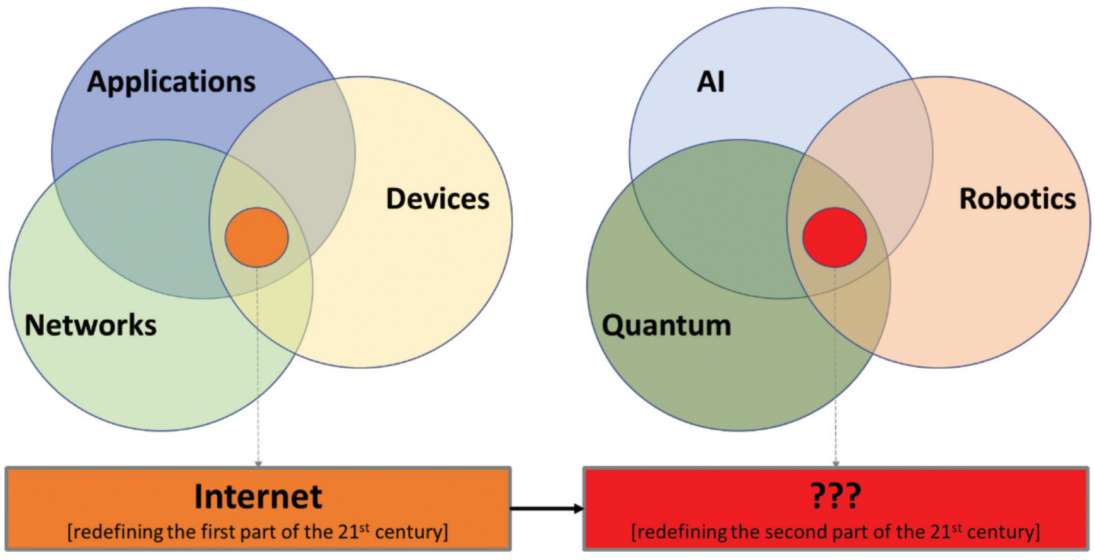


*Figure 1: Illustration of the consolidation and intersection of important fields, and the resulting capabilities. At the beginning of the 21st century, the Internet came about because of the intersection of devices, networks and applications. A new construct will emerge towards the later part of this century which will be underpinned by AI, quantum and robotics*

In 2020, we are still considered to be an era of "Artificial Narrow Intelligence" where the machine is trained to do a single ("narrow") task, such as *"recognise an object in an image".* By 2040, we are expected to reach the era of "Artificial General Intelligence" where machines exhibit intelligence equivalent to humans. By 2060, we are said to enter the era of "Artificial Super Intelligence" where machines will exceed the intelligence of humans. This will be a time where machines will learn quicker than they are being taught, and – if not managed well – will lead to the dreaded "AI Singularity" with machines potentially superseding humans.[10, 11]

**Robotics** and mechanical automation are now a very established fields which have evolved into an interdisciplinary area of research and innovation combining mechanical engineering, electrical engineering and computer science. With the emergence of 3D Printing, open source robotics and more powerful forms of AI, robotics is experiencing another renaissance in recent years. Autonomous robots are now being used in all fields of trade, including deep-sea exploration, nuclear decommissioning, advanced manufacturing sites, (sadly) military, among many other applications.

The interesting consolidation of the field, however, is not from these rather traditional Industry 4.0[12] robotic applications. It is the emergence of novel forms of robotics which give the field unprecedented capabilities. These are soft robotic systems, nano robots, self-folding and self-evolving origami-like structures, "living" and intelligent surfaces, and robots which act just like humans.

The ambition to "copy" nature has driven the field of robotics for centuries, and is now reaching the maturity to do just that. Soft robots are bendable structures with no hard mechanical components, yet fully controllable; they find applications in e.g. in-body precision surgeries. Nano robots are structures which are barely visible to the eye but able to traverse the human body to conduct specific tasks, such as bust cancer cells. Origami-like, "living" and intelligent robotic constructs give the versatility required for robotics to be a general-purpose field. Last but not least, the stunning capabilities of Boston Dynamic's robots able to do better-than-human pirouettes has surprised the world.[13] Clearly, the era of general human-machine coexistence has just began with all of above exciting developments.

**Quantum** technologies have also been in making for several decades. The recent announcement of various nation states[14] prohibiting the sales of quantum tech has shown that the technology has reached a state of maturity where it is seen as uniquely strategic.

Quantum relies on an entirely novel body of physics which departs from what is referred to as classical physics. The latter is governed by observable forces which form everything we see, feel, hear and experience at macroscopic level. It defines electricity and allows us to build circuits representing our classical binary world of zeros and ones. Quantum, on the other hand, relies on new phenomena: most notably i) superposition, and ii) entanglement.[15] Surprisingly, we still don't know much about the quantum world with new enigmas discovered regularly.[16]

The principle of superposition allows for several states to appear at the same time. This contrasts our classical thinking where a bit can either be one or zero but not both. Since quantum allows for this non-binary logic, we are able to introduce super-states of bits, also referred to as quantum bits or qubits. They, in turn, allow us to construct logics that are able to find solutions to exponentially difficult problems in linear time using algorithms such as Shor's Algorithm.[17]

Linearising exponentially difficult problems is a huge opportunity, but also a gigantic threat. Indeed, from an opportunity point of view, we are now able to find optimal solutions to problems which were unattainable with traditional computers. A classical example is the optimisation problem of the travelling salesman,[18] a representative combinatorial optimisation problem to numerous real-world applications.

Here, the optimum trajectory of a salesman selling goods in different cities needs to be found which is not possible in practically permissible time for an arbitrary number of cities. This is because the number of candidate routes increases exponentially with the problem size, which is also referred to as "combinatorial explosion". Indeed, the exact solution is exponentially hard, i.e. there is one optimum route out of $10^{11}$ for 14 cities which can be found in a few seconds today; not even doubling the amount of cities to 24 yields $10^{19}$ routes which would take more than 1,000 years to calculate; the case of 28 cities is completely unattainable to classical computers with a calculation time likely to be in the billions of years. Linearising this optimisation problem allows finding the optimal solution in a much short time.[18, 19]

Why is there a threat? Well, a large part of our cybersecurity encryption system is built on the exponential computational complexity of finding a solution. Notably, our encryption system is built on the assumption that it is impossible to find the confidential two prime numbers, the product of which forms a very large but publicly sharable number. Emails, webservers, credit card systems and much more is protected based on this assumption. Quantum computers eliminate that vital protection overnight. It is thus to no surprise that nation states guard any quantum compute capabilities.[14]

## 4. The "Digital Butterfly Effect"

When these technologies consolidate and interconnect, we will be entering an era where machines are able to learn faster than we can teach them due to advanced AI; where they can solve prior-unsolvable exponentially complex problems in linear time due to quantum; and where they are able to physically execute tasks to advance certain objectives due to advanced robotics. Coupled into a sense of duty, these machines can prove invaluable to society; however, coupled into a sense of self-preservation or when fallen into the wrong hands, they can prove fatal. Let us examine some canonical scenarios in this section.

- **Scenario #1 – Financial Markets Trading:** The trading liquidity in stock, commodity and currency markets is very high, with the Forex market alone accounting for an average of $5.3 trillion in daily trades.[20] Having an algorithmic edge over competing fellow traders is important and financially very lucrative. Not surprisingly, some of the best mathematicians and engineers end up working in financial markets.

  Overall, the optimisation problem however boils down to finding the best buy/sell options for some given market conditions which depend on millions of data points. Such exponential optimisation problems are typically solved using sub-optimal heuristics.

  Enters quantum computing, and suddenly exponential problems are linearised and sub-optimal heuristics become optimal solutions to highly complex market situations. This will give traders with quantum compute power an unbeatable edge and will thus heavily skew the markets. Enters AI, the execution can be fully automated with trillion-$ market decisions effectively being handed over to machines.

  These AI algorithms are trained to maximise profits. An example of a strategy is the capturing of trading stop losses through sudden and provoked market fluctuations. Several AI algorithms of market makers pitted against each other on primary currency pairs, such as EURUSD, will profit enormously whilst retail traders and nation states of secondary currencies may suffer. The latter may influence credit ratings and borrowing costs of nation states; this in turn can cause substantial harm.

  Therefore, the fairly benevolent aim of maximising profit in this hyperconnected world we live in can translate quickly to causing tangible damage. The exact attribution of what caused that damage however is often very difficult if not impossible to establish. In other words, we live in a world of a highly complex acyclic causality or, in layman's terms, in a world governed by the "digital butterfly effect".[21]

- **Scenario #2 – Paperclip Paradox:** Marvin Minsky, the co-founder of MIT's AI lab, has philosophised that an advanced artificial intelligence tasked to prove the Riemann hypothesis may choose to take over all of Earth's resources to construct supercomputers to assist attain its objective.[22] If the computer system had actually rather been configured to produce as many paper clips as possible, it would still choose to take all of Earth's resources to satisfy this very objective. Even though these two objectives are very different, both of them produce a convergent instrumental goal of taking over the planet's resources.

  Nick Bostrom has perfected this conundrum through his paperclip example which shows the existential danger that AI might present to human beings when configured to pursue even seemingly harmless objectives: *"Suppose we have an AI whose only goal is to make as many paper clips as possible. The AI will realize quickly that it would be much better if there were no humans because humans might decide to switch it off. Because if humans do so, there would be fewer paper clips. Also, human bodies contain a lot of atoms that could be made into paper clips. The future that the AI would be trying to gear towards would be one in which there were a lot of paper clips but no humans."*[23]

  Whilst hypothetical examples, Minsky, Bostrom and many others have thus argued for the necessity of incorporating machine ethics into artificial intelligence design.[24] An important aspect here is the issue around the machines' ambitions of self-preservation: our hyperconnected world would offer these machines effectively infinite resources and boundaries, which could prove fatal if not dealt with properly.

- **Scenario #3 – "Battlestar Galactica":** Battlestar Galactica is an American science fiction film series, aired until about 2010. Whilst the plot is fairly flat, an interesting aspect of the series is that – unbeknown to humans – machines by the name of Cylons have lived among humans for decades without being detected. Advanced soft robotics gave machines the ability to morph into human shape; and advanced AI gave the machines super intelligence and a purpose of self-preservation (to the point that the overall aim of the Cylons was to exterminate the human race).

  An interesting question arises how a very advanced human civilisation was not able to spot the existence of another intelligent civilisation among them? Whilst clearly making use of AI and other advanced technologies, these humans were blind to the existence of machines in front of their very eyes. *Could it be that this fictional civilisation in a distance future is us, today?*

## 5. Towards A Global Digital Observatory

The answer to the question in the previous section if the fictional civilisation in a distance future could be us today is: *we wouldn't know.* That is very worrisome as one would expect to ascertain the statement with a clear *'no'.* Yet, we have no measures, no key performance indicators (KPIs), no measurement equipment to determine if AI has an influence on or even shaping human civilisation today.[25]

Regarding above market trading example, bar quantum, AI-enabled trading is reality today and the digital butterfly effect is not so far-fetched. We could draw up many more examples where a small and fairly benevolent action on one side of the planet causes massive havoc on the other side. And just like with the physical "butterfly effect",[21] it is impossible to reverse-engineer the chain of causality. The quick and widespread decision making by machines today through already fairly advanced AI poses a real challenge to separate action from reaction, and to establish chains of causality. *In summary, it is a mistake to speak about the impact of AI onto humanity in a future tense. AI is here, today.* And we urgently need to understand and measure its impact.

This challenge of observability has now been recognised for a certain class of AI: deepfakes.[8, 9, 46] They are enabled by GANs, a "creative" class of AI algorithms, which are

able to create and mimic face and bodies in photos and videos, and even fake voices. Perceived to be an "imminent challenge to democracy",[26] various legislative bodies have pushed for stronger oversight. For instance, the Deepfake Accountability Act has been discussed in the US.[27] Pressured by such legislation, Silicon Valley's largest players have formed a coalition which was launched in September 2019 through a "Deepfake Detection Challenge".[28] It is the first attempt on record to establish means of observing the effect of AI. Only a handful of viable solutions have been presented to date (Q4 2020). Such important initiatives, however, should not remain in private hands and should rather be handled at national if not global governance level.

A tangential question on privacy arises here too which may have a profound impact on the relationship between man and machine. Notably, all our privacy laws are understood to protect personal information from *other humans.* However, we seem to be generally happy to store our private information on globally connected platforms. In other words, we are handing over human knowledge, wisdom, methodologies, thinking, etc to machines. There is nothing to worry about today, but in a future hyperconnected world with reasonably conscious machines, access to that information may prove the tipping point for machines. We thus should commence legislating privacy assuming a strong presence of "conscious" machines.

In summary, the hyperconnectivity, fluidity and complexity of today's digital fabric prohibits us to attribute specific emerging behaviour to man or machine. In contrast to e.g. the natural laws of classical physics, the phenomena in the internet are not (easily) reproduced. The introduction of advanced AI, quantum and robotics will amplify this even further, with the influence of AI potentially remaining undetected for decades.

*It is for this reason, that we ought to establish with urgency a Global Digital Observatory (GDO), analogous to the inter-governmental Panel on Climate Change. It should be an internationally chartered body, with advisory and legislative oversight.*

Similar to climate change, two ingredients are needed to make this work: technology and political will. The latter will hopefully be addressed through the GDO. The former, i.e. technology, is an open problem with first solutions only appearing now.[47] However, it is important to continue posing the challenges to scientists and engineers so they can develop solutions over the years to come. Some of the technology constituents, such as distributed ledger technologies (DLTs), are available today;[29, 30] yet others still need to be developed through suitable government funding schemes.

Not acting on this threat today will potentially put humanity on a point of no return, similar to the situation we are facing with irreversible climate change. The threat is real, with scientifically rigorous developments showing that AI is able to evolve by itself[48] and even showing empathy,[49] but that it would be impossible to control a rogue super-smart AI[50] with our current tools of observation.[51]

The difficulty politicians and activists face is the long-term inertia with effects not becoming visible until in a few decades. A political will above the typical legislative cycles of 4-5 years is thus needed to avert a potentially grim outlook.

## 6. Established AI Governance Initiatives

The need to put governance principles in place for AI has been recognised as early as 2003.[23] First governance and normative frameworks, however, were not proposed until much later. For instance, the OECD published governing "Principles on Artificial Intelligence".[31] The G20 laid out norms for "human-centered AI that promotes innovation and investment";[32] the G20 AI Principles were drawn from the OECD Recommendation on AI and seek *"to foster public trust and confidence in AI technologies and realise their potential, through promoting principles such as inclusiveness, human-centricity, transparency, robustness and accountability".*[32]

In addition to above, there is the Global Partnership on AI (GPAI) that is supported by the Group of Seven (G7) as well as the OECD.[33] Further, in June 2020, a "Roadmap on Digital Cooperation" has been announced by the UN Secretary-General that recommends an inclusive multi-stakeholder forum with eight areas of immediate focus: connectivity, digital public goods, digital inclusion, digital capacity building, digital human rights, digital trust and security, critical infrastructure, and global digital co-operation. He warned that "we should not let technological developments get ahead of our capacity to manage them".[34]

Europe, in the meantime, has taken leadership with specific legislations protecting citizens from secondary effects of AI through the GDPR legislation.[35] Notably,

- GDPR Article 12 (in force as of 25 May 2018) stimulates: "Transparent information, communication and modalities for the exercise of the rights of the data subject, where the controller shall take appropriate measures to provide any information referred to in Articles 13 and 14 and any communication under Articles 15 to 22 and 34 relating to processing to the data subject in a *concise, transparent, intelligible and easily accessible form, using clear and plain language*."
- GDPR Article 13 (in force as of 25 May 2018): "Information to be provided where personal data are collected from the data subject, where the controller shall provide […] the existence of automated decision-making, including meaningful information about the logic involved, as well as the significance and the envisaged *consequences of such processing* for the data subject."
- Markets in Financial Instruments Directive (MiFID) (in force as of Jan 2018): Requires operators of automated decision-making algorithms to provide human understandable justifications for automated decisions.

These laws address to some extend the scenarios discussed, notably Article 13 is very restrictive in that any consumer-facing AI needs to be explainable which in essence complicates the use of the most advanced deep-learning AI. Significant work is thus still needed, and underpins the need for a constantly evolving legislation rather than a static set of articles.

A recent contribution which stands out is.[36] It is a policy brief arguing for the establishment of a G20 Coordinating Committee for the Governance of Artificial Intelligence (G20 CCGAI). The role of such a body would be "to plan and coordinate on a multilateral level the mitigation of AI risks". The authors argue that only the G20 can implement the needed meta-governance mechanism, given "the involvement of the largest economies and their highest political representatives" since "other regime complexes and international organizations, which also focus on AI governance, tend to either lack such political power or exclude major rivalry countries". The focus of the policy brief is "on the functional and organizational features of the CCGAI and how such chartered committee can foster international collaboration," noting that "[…] without such global coordination, AI governance risks becoming primarily an instrument of strategic competition and rivalry".[36]

The proposed G20 CCGAI poly-centric governance approach at highest political level would need to be enforced at national and regional levels, and then technically implemented. The enforcement requires the G20 to establish operational links to e.g. the Digital Economy Development and Cooperation Initiative (China 2016), Digital Economy Ministerial Declaration (Germany 2017), and Ministerial Statement on Trade and Digital Economy and AI Principles (Japan 2019).[36] The technical implementation requires strong links to standards defining organisations (SDOs), such as the ITU, ISO, 3GPP, IEEE, ETSI and the IETF; as well as alliances such as the NGMN, GSMA, W3C, etc.

## 7. Recommendations on Governance & Policies

Given the emergence of quantum and advanced robotics, in addition to the accelerating capabilities of AI, we argue that above-mentioned governance initiatives ought to be revisited and refocused. Indeed, any ethical and societal problems identified as part of the AI governance work will be accelerated with quantum technologies and amplified with advanced robotics. The consolidation and intersection of AI, quantum and robotics will lead to unpredictable emerging capabilities which could pose a hyper-exponential threat to society with both man and machine able to weaponise for a highly asymmetric and potentially invisible warfare.

Not widely discussed but the visibility of these emerging capabilities is utmost important to any ability to counter potential threats. Without proper mechanisms of observation, any attempts of governance, policy and legislation will be futile. It is similar to wanting to avert the threat of speeding drivers without any speed cameras, i.e. without any mechanisms of observation. Initiatives, such as the Deepfake Detection Consortium,[28] have already proven that establishing observable evidence is technically extremely challenging.

Without such directly or indirectly observable oversight, the very notion of human identity could be challenged. It could build up to mass unemployment or labour displacement, with an increasing lack of dignity and an erosion of social cohesion. The notion of identity however is truly important as it is now well understood to be the cause of all conflicts.[37] A recent shift in identity from a "world of passports" to a "world of passwords" has already posed challenges in asymmetric conflicts.[38] An amplification of this could plunge the world into a state of no return.

***We thus propose the establishment of a Global Digital Observatory (GDO), analogous to the inter-governmental Panel on Climate Change. This to-be-developed capability could reside within already chartered bodies or become an internationally chartered body on its own.***

***The GDO should not only cover AI but all emerging technologies and combinations thereof.***

The GDO should be governed by the basic principles of responsibility and human-centredness. It will likely fail if constructed as a central entity and rather rely on polycentric governance,[36] representing a broad spectrum of global, regional and national stakeholders but most importantly the voice of citizens and society. It should reconcile the prevailing structural fragmentation between current stakeholders[39] and find alternatives to the established logic of economic efficiency.

The GDO should reinforce social cohesion, and developed legislation should counter the increasing lack of privacy society experiences due to often invisible technological developments in an increasingly hyperconnected world.[40] Growing issues around identity, and by extensions physical and mental wellbeing, must be addressed well before emerging tech trends will become irreversible.

Contingency measures for the risks arising from said technologies need to be developed and frequently revisited. Also, a framework on the sustainability of these technologies ought to be established, which quantifies their carbon footprint and thus helps meeting the UN's Sustainability Goals.[41] For instance the entire ICT sector is consuming more than 4% of the global energy consumption;[42] Bitcoin mining consumes more energy than entire nation states[43] – we urgently need an oversight mechanism which projects their carbon footprint rather than measuring it post-deployment (when it is often too late).

Importantly, the GDO needs to recognise the interdependent developments in AI, quantum and robotics, underpinned by further developments in 6G and cyber, which are all large regulatory domains of their own but which cannot be viewed independently anymore. To complicate matters further, the geopolitical rivalry in precisely these domains is growing[44] which consumers, businesses and society cannot afford.

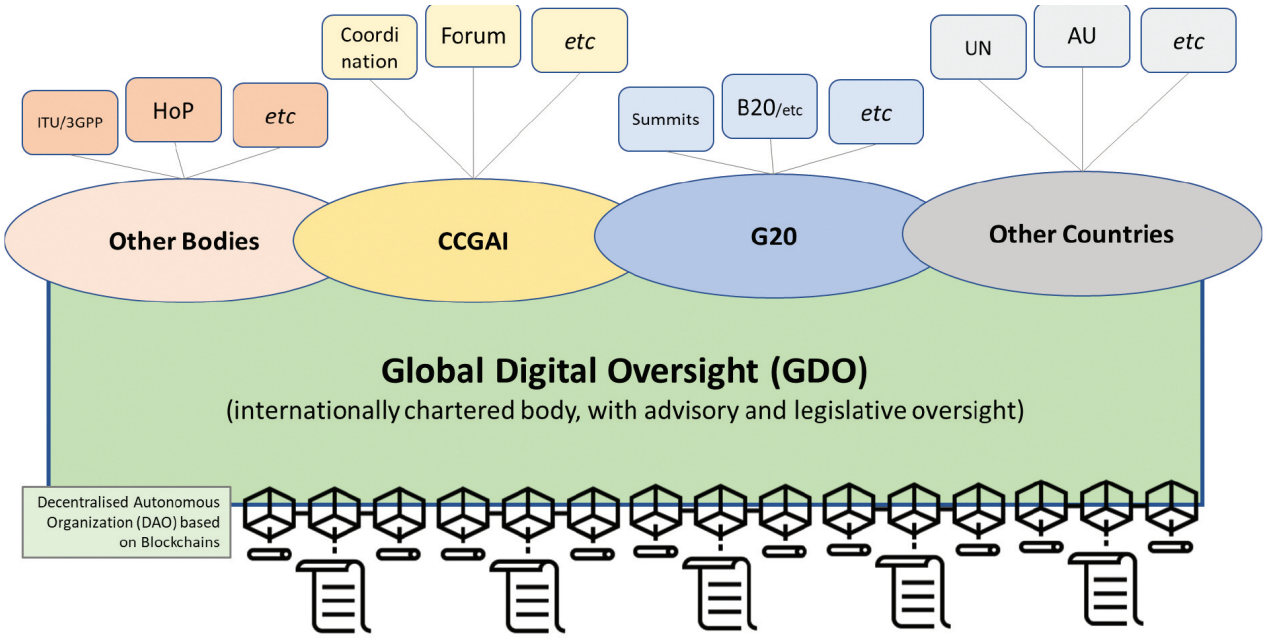


*Figure 2: High-level overview of the role of a Global Digital Oversight body and its relationship with the G20 (including Summits, the Business 20 / B20, etc), with the Coordinating Committee for the Governance of Artificial Intelligence (CCGAI), with other bodies (including technical bodies such as the ITU/3GPP but also political institutions like the House of Parliament (HoP)), and with other countries (such as represented by the UN or the Assembly of the African Union (AU). The underlying technology binding all together in a trustworthy and transparent fashion could be DAOs based on modern blockchain technologies*

It is thus vital to counter the fragmentation of existing technological, geopolitical and regulatory regimes by enabling an effective coordination of the opportunities and risks at global, regional and national levels. To address these mid- and long-term structural imbalances, the three main forces of market competition, systematic competition and technology determinism[36] that limit the effectiveness of governance need to be reconciled.

At the same time, the GDO should not be construed as a "Big Brother" of all tech. It must not jeopardise fundamental principles of ethics and human-centredness. It should act as a high-level observatory able to establish trends and causality. It should not micro-manage but rather develop common principles through inclusive collaboration, and find a common ground between over-regulation and under-regulation. Regarding the former, industry is hampered and innovation stifled thus depriving consumers of innovative products and services; regarding the latter, industry is granted a free reign thus potentially jeopardising long-term societal values.

The GDO should also explicitly stand up to an aggressive weaponization of the emerging technologies. Whilst we have to acknowledge that our dependency on technology prevents simple switches distinguishing between "good use" and "bad use" of tech,[36] the highly asymmetric capabilities of AI, quantum and robotics will require bolder mechanisms to prevent the misuse of tech. One could argue that it could cause harm at par with an atomic bomb, and thus should be regulated through international treaties as such. The GDO could be an independent body for instigating such treaties.

By virtue of above, it is important to ensure that the GDO extends beyond the G20. Indeed, emerging capabilities can lead for the bottom 1% of the world to upset the remaining 99% through asymmetric hyperwars. The G20, whilst representing 90% of global GDP and 80% of global trade, does not suffice to protect against this high asymmetry. Without full inclusion, a democracy could slip into an authoritarian state, an authoritarian into totalitarian one and a totalitarian into a dictatorship.[36]

A challenge for above is the current political populous and low trust environment within and among nations. Technologies, such as Distributed Ledger Technologies (DLT), could prove pivotal here where one could imagine the GDO run as a Decentralised Autonomous Organization (DAO).[45] Whilst we imagine the GDO established, co-governed and occasionally audited by humans, operations are represented by rules encoded as immutable computer programs that are transparent, controlled by the organisation members and not influenced by a central entity. Decision provenance and neutrality w.r.t. established rules is thus guaranteed and can help overcome the current low-trust environment.

In summary, the GDO ought to be chartered, permanent, fair and transparent, and have sufficient operational means and executive mandates. It should be governed by principles of responsibility and human-centredness; and rely on polycentric governance. It could be embedded into current policy mechanisms, such as the G20, or into emerging ones, such as the proposed CCGAI;[36] but its remit needs to be extended to a true global inclusiveness.

# About the Author

**Mischa Dohler** is full Professor in Wireless Communications at King's College London, driving cross-disciplinary research and innovation in technology, sciences and arts. He is a Fellow of the IEEE, the Royal Academy of Engineering, the Royal Society of Arts (RSA), the Institution of Engineering and Technology (IET); and a Distinguished Member of Harvard Square Leaders Excellence. He is a serial entrepreneur with 5 companies; composer and pianist with 5 albums on Spotify/iTunes; and fluent in 6 languages. He acts as policy advisor on issues related to digital, skills and education. He has had ample coverage by national and international press and media.

He is a frequent keynote, panel and tutorial speaker, and has received numerous awards. He has pioneered several research fields, contributed to numerous wireless broadband, IoT/M2M and cyber security standards, holds a dozen patents, organized and chaired numerous conferences, was the Editor-in-Chief of two journals, has more than 300 highly-cited publications, and authored several books.

He was the Director of the Centre for Telecommunications Research at King's from 2014-2018. He is the Cofounder of the Smart Cities pioneering company Worldsensing, where he was the CTO from 2008-2014. He also worked as a Senior Researcher at Orange/France Telecom from 2005-2008.

# Bibliography


1 "Why Infrastructure Connectivity Matters," OECD Japan G20 Development Working Group, January 2019; online: https://bit.ly/3gFiVgk.
2 "Moore's Law," Wikipedia; online: https://en.wikipedia.org/wiki/Moore%27s_law.
3 "From 5G to 6G Governance," M Dohler and D Simeonidou, policy paper, Dec 2020.
4 "The world's 2,153 billionaires have more wealth than 4.6 billion people combined, Oxfam says," CNBC, Jan 2020; online: https://www.cnbc.com/2020/01/20/oxfam-worlds-billionaires-richer-than-a-combined-4point6-billion-people.html.
5 "Convolutional Neural Network (CNN)," Wikipedia; online: https://en.wikipedia.org/wiki/Convolutional_neural_network
6 "Recursive Neural Networks (RNNs)," Wikipedia; online: https://en.wikipedia.org/wiki/Recursive_neural_network.
7 "Generative Adversarial Networks (GANs)," Wikipedia; online: https://en.wikipedia.org/wiki/Generative_adversarial_network.
8 "Deepfake," Wikipedia; online: https://en.wikipedia.org/wiki/Deepfake.
9 "The age of the deepfake is here – and it's more terrifying than you think," *Telegraph*, Sept 2020; online: https://www.telegraph.co.uk/technology/2020/09/05/age-deepfake-stranger-think.
10 "Technological singularity," Wikipedia; online: https://en.wikipedia.org/wiki/Technological_singularity.
11 "Can we stop AI outsmarting humanity?," *The Guardian*, March 2019; online: https://www.theguardian.com/technology/2019/mar/28/can-we-stop-robots-outsmarting-humanity-artificial-intelligence-singularity.
12 "What is Industry 4.0? Here's A Super Easy Explanation For Anyone," Forbes, Sept 2018; online: https://www.forbes.com/sites/bernardmarr/2018/09/02/what-is-industry-4-0-heres-a-super-easy-explanation-for-anyone.
13 "Boston Dynamics," Wikipedia; online: https://en.wikipedia.org/wiki/Boston_Dynamics.
14 "Google and China duke it out over 'quantum supremacy'," Fortune, Dec 2020; online: https://fortune.com/2020/12/08/china-quantum-supremacy-computing-ibm-google-rigetti-honeywell-ionq.
15 "Quantum Mechanics," Wikipedia; online: https://en.wikipedia.org/wiki/Quantum_mechanics.
16 "The Strange World of Quantum Physics," NIST, 2018; online: https://www.nist.gov/topics/physics/introduction-new-quantum-revolution/strange-world-quantum-physics.
17 "Shor's algorithm," Wikipedia; online: https://en.wikipedia.org/wiki/Shor%27s_algorithm.
18 "Travelling salesman problem," Wikipedia; online: https://en.wikipedia.org/wiki/Travelling_salesman_problem.
19 "'Electronic amoeba' finds approximate solution to traveling salesman problem in linear time," Physic, Dec 2020; online: https://phys.org/news/2020-12-electronic-amoeba-approximate-solution-salesman.html.
20 "Understanding Forex Market Liquidity," Blackwell Global; online: https://blackwellglobal.com/forex-market-liquidity.
21 "Butterfly Effect," Wikipedia; online: https://en.wikipedia.org/wiki/Butterfly_effect.
22 "Instrumental Convergence," Wikipedia; online; https://en.wikipedia.org/wiki/Instrumental_convergence.
23 "Ethical Issues in Advanced Artificial Intelligence," Nick Bostrom, 2003; online: http://www.nickbostrom.com/ethics/ai.html.
24 "Artificial Intelligence May Doom The Human Race Within A Century, Oxford Professor Says," *Huffington Post*, August 2013; online: https://www.huffingtonpost.com/2014/08/22/artificial-intelligence-oxford_n_5689858.html.
25 Private conversation between Ali Hossaini and Mischa Dohler, London, UK, September 2016.
26 "Deepfake technology presents imminent challenge to democracy," International Bar Association,

Feb 2020; online: https://www.ibanet.org/Article/NewDetail.aspx?ArticleUid=432E7100-35BE-44B4-B6AB-AD55758411F4.
27 "H.R.3230 - Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019," US Congress, 2019; online: https://www.congress.gov/bill/116th-congress/house-bill/3230/text.
28 "Facebook, Microsoft, and others launch Deepfake Detection Challenge," Venture Beat, Dec 2019; online: https://venturebeat.com/2019/12/11/facebook-microsoft-and-others-launch-deepfake-detection-challenge.
29 "Blockchain," Wikipedia; online: https://en.wikipedia.org/wiki/Blockchain.
30 "Industry Specification Group (ISG) Permissioned Distributed Ledger (PDL)," ETSI Standard; online: https://www.etsi.org/committee/1467-pdl.
31 "Recommendation of the council on artificial intelligence," OECD, Paris, France, Dec 2019; online: https://www.oecd.org/going-digital/ai/principles.
32 "G20 ministerial statement on trade and digital economy," G20 Ministry of Foreign Affairs, April 2020; online: https://www.meti.go.jp/press/2019/06/20190610010/20190610010-1.pdf.
33 "Joint Statement from founding members of the Global Partnership on Artificial Intelligence," United States Department of State, Jun 2020; online: https ://www.diplomatie .gouv.fr/en/french-foreign-policy/digital-diplomacy/news/article/launch-of-the-global-partnership-on-artificial-intelligence-by-15-founding.
34 "United Nations Secretary-General's Roadmap for Digital Cooperation," United Nations, 2020; online: https://www.un.org/en/content/digital-cooperation-roadmap.
35 "General Data Protection Regulation," Europe, 2018; online: https://gdpr-info.eu.
36 Jelinek, T., Wallach, W. & Kerimi, D. Policy brief: the creation of a G20 coordinating committee for the governance of artificial intelligence. AI Ethics (2020). https://doi.org/10.1007/s43681-020-00019-y.
37 "Meeting of High Panel on Peace and Dialogue Among Cultures," UNESCO, 2011; online: https://unesdoc.unesco.org/ark:/48223/pf0000192362.
38 "Cyber-Extremism: Isis and the Power of Social Media", Awan, I., Soc 54, 138–149 (2017). https://doi.org/10.1007/s12115-017-0114-0.
39 "US Senate subpoenas heads of Google, Facebook and Twitter," BBC, Oct 2020; online: https://www.bbc.co.uk/news/technology-54376327.
40 "The Hyperconnected World of 2030–2040," Institute for the Future, 2020; online: https://www.iftf.org/hyperconnectedworld/.
41 "United Nation Sustainability Goals," UN; online: http://www.undp.org/content/undp/en/home/sustainable-development-goals.html.
42 "Beyond the Energy Techlash: The Real Climate Impacts of Information Technology," ITIF, 2020; online: https://itif.org/publications/2020/07/06/beyond-energy-techlash-real-climate-impacts-information-technology
43 "Bitcoin's energy consumption 'equals that of Switzerland'," BBC, July 2019; online: https://www.bbc.co.uk/news/technology-48853230.
44 "Big tech is now a geopolitical force – and that should worry us all," *Wired*, April 2020; online: https://www.wired.co.uk/article/big-tech-geopolitics.
45 "Decentralized Autonomous Organizations," Wikipedia; online: https://en.wikipedia.org/wiki/Decentralized_autonomous_organization.
46 "Deepfake queen to deliver Channel 4 Christmas message," BBC News, December 2020; online: https://www.bbc.co.uk/news/technology-55424730.
47 "Teenager's AI System for Detecting Deepfake Videos Wins Award," Interesting Engineering, January 2021; online: https://interestingengineering.com/teenagers-ai-system-for-detecting-deepfake-videos-wins-award.
48 "Artificial intelligence is evolving all by itself," *Science* Magazine, April 2020; online: https://www.sciencemag.org/news/2020/04/artificial-intelligence-evolving-all-itself.

49 “A.I. displays an unsettling skill: the ability to show empathy,” The Inverse, Jan 2021; online: https://www.inverse.com/innovation/robot-empathy-deception.
50 “Calculations Show It’ll Be Impossible to Control a Super-Intelligent AI,” Science Alert, Jan 2021; online: https://www.sciencealert.com/calculations-show-it-d-be-impossible-to-control-a-rogue-super-smart-ai.
51 “The Turing Test is obsolete. It’s time to build a new barometer for AI,” The Fast Company, December 2020; online: https://www.fastcompany.com/90590042/turing-test-obsolete-ai-benchmark-amazon-alexa.

JESUS COLLEGE
CAMBRIDGE

Global Issues
Dialogue Centre

Global Issues Dialogue Centre
Jesus College
Cambridge
CB5 8BL
www.jesus.cam.ac.uk/research/global/Global-Issues-Dialogue