



## AI GLOBAL GOVERNANCE – WHAT ARE WE AIMING FOR?

---

### CONCLUSIONS AND RECOMMENDATIONS

- A. Effective AI global governance is urgently needed and a campaign to establish an AI Global Governance Regime Complex is necessary.
  - B. Further thought, debate and agreement is required regarding the design of an appropriate Regime Complex.
  - C. Recommendations regarding such a design are:
    - 1. To be effective, the Regime Complex should be: inclusive, anticipatory, responsive and agile, reflexive, open to differentiated cooperation, transparent, enforceable where necessary and holistic and comprehensive, whilst avoiding duplication (See Table 2 for more detail).
    - 2. The intention is that the Regime Complex will play a major role in the future of humanity. Existing institutions should only be used as a short-term expedient, or in a way that is peripheral to the central regime.
    - 3. The Regime Complex should be designed to complement existing legislation, such as that which relates to human rights or the sale of goods and should focus on critical issues such as the protection of governability and human self-determination.
    - 4. A decision needs to be made as to whether the AI Regime Complex forms part of a larger regime embracing digital and/or disruptive technologies, or if it is to stand in its own right within the UN System.
  - D. The process leading to the Summit for the Future in 2023 is a major opportunity for the furtherance of the proposed campaign / AI Global Governance.
    - 1. Leading up to the Summit, every effort should be made to progress the thinking on these issues, feed comments into the process where possible, and seek to reach sufficient consensus to secure a constructive way forward.
    - 2. To achieve the above, the different AI Governance initiatives need to work together, mindful of the vanity of small differences.
- 



## 1. INTRODUCTION

The arguments for and against centralisation regarding AI governance have been clearly set out by Cihon et al<sup>1</sup> and an AI global governance regime is not without its downsides. There is urgent pressure however for an AI global governance regime and in the long term such a regime will be essential. A strong case has been made for the establishment of an AI global governance regime<sup>2,3,4,5</sup> and for this paper it is assumed that one will be established in the coming years.

The end destination however is not well defined. What should such a regime look like, how should it act and how broad should the scope of its powers be?

It has been suggested that AI Global Governance should be “Effective, Timely and Global”<sup>6</sup>. The central purpose of this paper is to unbundle the question of effectiveness. Section 2 seeks to identify the key characteristics of a global AI governance regime, and Section 3 sets out its framework options. Later, in Section 4 and Section 5, the issue of timeliness is addressed. Given the subject of this paper, the ‘global’ nature of governance is taken as read.

Whilst it may be premature to seek to identify every aspect of such a regime, it is not too early to give thought to the nature of its key elements.

## 2. THE NECESSARY CHARACTERISTICS OF AN AI GLOBAL GOVERNANCE REGIME COMPLEX.

There are a variety of regime models that have been put forward, some of which are discussed in Section 3 below. A number of actors have sought to identify the desirable features of an AI regime, including Kemp and colleagues<sup>7</sup>, Wendell Wallach and Gary Marchant<sup>8</sup>, Pekka Ala-Pietila and Nathalie Smuha<sup>9</sup>, The World Economic Forum / Deloitte<sup>10</sup> and the UN Technology Envoy. Each set of criteria is reviewed below.

### *Kemp et al*

Luke Kemp, along with colleagues from the Cambridge Centre for the Study of Existential Risk and the Oxford based Future of Humanity Institute, puts forward four key elements, originally identified by Stilgoe<sup>11</sup> et al as the key tenets of Responsible Research and Innovation, necessary for the design of a UN AI Regime<sup>12</sup>. The four elements are listed below.

<sup>1</sup> Cihon, P., Maas, M. and Kemp, L. “Fragmentation and the Future: Investigating Architecture for International AI Governance”, *Global Policy*, 11(5), 2020.

<sup>2</sup> Erdelyi, O., and Goldsmith, J., “Regulating Artificial Intelligence: Proposal for a Global Solution”. 2018 *AAAI/ACM Conference on AI, Ethics, and Society (AI/ES '18), February 2--3, 2018, New Orleans, LA, USA*. 2018.

<sup>3</sup> Jelinek, T., Wallach, W. & Kerimi, D. “Policy brief: the creation of a G20 coordinating committee for the governance of artificial intelligence”. *AI Ethics* 1(141-150), 2021. <https://doi.org/10.1007/s43681-020-00019-y>

<sup>4</sup> Ala-Pietila, P. and Smuha, N. ‘A Framework for Artificial Intelligence and its Governance’ *Reflections on Artificial Intelligence for Humanity 2021*.

<sup>5</sup> Club de Madrid and Boston Global Forum, “Fundamental Rights in AI and Digital Societies: Towards an International Accord” 2021. <https://aidigitalrights.com/wp-content/uploads/2021/09/Club-de-Madrid-Boston-Global-Forum-Governance-Issues-Paper-Workshop-September-2021.pdf>

<sup>6</sup> Whitfield, R. et al., “Effective, Timely and Global – The Urgent Need for Good Global Governance of AI”, *One World Trust*, 2020, <https://www.wfm-igp.org/publication/effective-timely-and-global-the-urgent-need-for-good-global-governance-of-ai/>

<sup>7</sup> Kemp, L., Cihon, P., Maas, M. et al, “UN High-level Panel on Digital Cooperation: A Proposal for International AI Governance”, 2019, <https://www.cser.ac.uk/news/advice-un-high-level-panel-digital-cooperation/>

<sup>8</sup> Wallach, W., Marchant, G. “Toward the Agile and Comprehensive International Governance of AI and Robotics”, *Proceedings of the IEEE*, 107(3), 2019 <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8662741>

<sup>9</sup> Ala-Pietila and Smuha, “Artificial Intelligence and its Governance”.

<sup>10</sup> Deloitte and the World Economic Forum, “Harnessing Fourth Industrial Revolution Technologies in a COVID 19 World 2021”, *Global Technology Governance Report*, 2021, <https://www.weforum.org/reports/global-technology-governance-report-2021>

<sup>11</sup> Stilgoe, J., Owen, R., and Macnaghten, P. “Developing a Framework for Responsible Innovation”, *Research Policy*, 42(9), 2013: 1568-1580 <https://www.sciencedirect.com/science/article/pii/S0048733133000930>

<sup>12</sup> Kemp et al, “A Proposal for International AI Governance”.

- a. Inclusive: this term is widely used in the context of AI Governance. All stakeholder groups and marginalised groups should be included. The regime should be sensitive to power imbalances. Open participation in discussions of different options and supporting arguments should be encouraged<sup>13</sup>. Kemp suggests that all parties should be involved, at least to some extent, in voting decisions. Implicitly, the concept of inclusivity implies that authoritarian countries such as China or Russia should be included in the regime, (although in practice that is not always the case).
- b. Anticipatory: anticipate fast-progressing AI technologies and impacts (with super-intelligence being the ultimate example, where anticipation will be essential).
- c. Responsive: responsive to the evolving technology and its uses.
- d. Reflexive: critically reviewing and updating its own policy principles.

### **Wallach and Marchant**

Two key criteria that Wendell Wallach and Gary Marchant argue for are Agility and Comprehensive-ness<sup>14</sup>.

They express concern with regard to the “pacing problem”, namely the mismatch between the accelerating pace at which emerging technologies such as AI are being developed on the one hand, and the slowing down of traditional institutions of legislation, regulation, and judicial review on the other. To avoid an inadequate governance structure requires agility<sup>15</sup>.

They identify the safety risks that AI presents in a variety of contexts, ranging from autonomous vehicles to financial algorithms, but also the “concerns relating to privacy, autonomy, enhancement, bias, fairness, justice, relationships to others, unemployment, national security, and

existential risk”. They also observe “AI, robotics, and other emerging technologies develop in an international context, often making national regulation disadvantageous, inept, or incomplete”. A comprehensive solution is needed to overcome these risks.

They highlight the key issue of enforceability, typically associated with hard law. Marchant argues strongly in favour of soft law, at least in the early days. While not enforceable, soft law mechanisms facilitate agility and responsiveness – so enforceability is nuanced.

They also state the need for new institutions and methods that are both more reflexive and more inclusive.

### **Ala-Pietila and Smuha**

Pekka Ala-Pietila, Head of the European Commission High-Level Expert Group on AI (HLEG on AI), and Nathalie Smuha, from KU Leuven, identify seven elements that can help guide the organisation of both existing and new global cooperation initiatives<sup>16</sup>. These elements, largely self-explanatory, are listed below.

1. Balancing the need for swift action, a holistic approach and attention to context-specificity. [This sentence embraces three elements, namely 1A = swift action, 1B = a holistic approach, 1C = attention to context specificity].
2. Clarifying the rules of engagement.
3. Building on existing cooperation structures.
4. Developing a network of networks.
5. Maintaining openness to differentiated cooperation. [This relates to the practice developed within the European Union, whereby some countries can agree to cooperate more closely on certain issues about which other countries might feel more cautious, at least initially.]
6. Securing an inclusive and transparent way of working, mindful of power imbalances. [This sentence embraces two elements, namely 6A = Inclusivity, 6B = Transparency].

<sup>13</sup> Dodds, F. *Stakeholder Democracy – Represented Democracy in a Time of Fear* (Routledge, 2019)

<sup>14</sup> Wallach and Marchant, “International Governance of AI and Robotics”.

<sup>15</sup> Cihon, Maas and Kemp., “Fragmentation and the Future”.

<sup>16</sup> Ala-Pietila and Smuha, “Artificial Intelligence and its Governance”.

7. Establishing a feedback loop and preparing for the future [This again embraces two elements, namely 7A: Establishing a feedback loop – reflexivity, 7B = Preparing for the future – anticipatory.]

Whilst some of these elements describe desirable features of a future regime, others describe how to go about establishing such a regime in the first place. In particular, items 1A, 1C, 2, 3, and 4 all relate to the process of reaching an AI Global Governance regime, rather than the design of that regime per se. The issue of process is addressed in later sections of this paper.

**Deloitte / WEF:**

Deloitte has produced a report for the World Economic Forum. The Global Technology Governance Report 2021<sup>17</sup> reviews a wide range of Technology Governance initiatives from around the world and produces a list of the most desirable cross-cutting features. These features are then used to assess a number of existing governance frameworks. While the features are not intended as a list of desirable institutional qualities, they are worth reviewing. The features are:

- a. Agile: This is a relevant institutional quality.
- b. Fit for purpose: This is relevant for a critique of an existing institution but is tautologous as a design criterion.
- c. Globally relevant: The initiatives assessed by the report were typically not global and therefore this was a relevant criterion for the study, but not a design criterion for a global institution.
- d. Inclusive: This is a relevant institutional quality.
- e. Innovative: Another term that is useful as a descriptive observation but not as a design criterion.
- f. Evidence based: The concept of evidence-based policy making is widely supported but in the realm of AI there are potential dangers.
- g. Produced outcomes: This relates to the past and is not a design criterion.
- h. Currently live: Not relevant.

<sup>17</sup> Deloitte and WEF, “Harnessing Fourth Industrial Revolution Technologies in a COVID 19 World”.

From the above list of innovative governance framework criteria, the two features that are relevant to the current assessment are Agility and Inclusiveness. A third feature, the concept of evidence-based policy making is discussed below.

**UN Envoy on Technology**

UN Assistant Secretary General Maria Francesca Spatolisano, Officer in Charge, Office of the UN Envoy on Technology has spoken of the need for inclusive, responsive and effective coordination structures<sup>18</sup>. Effective is the term this paper is seeking to unpack, with inclusive and responsive emerging as key institutional qualities.

**Analysis**

**The Basic Institutional Qualities Sought in a Global AI Regime: Initial summary**

Source	S/K	WW	A-P	WEF	ET
Inclusive	Y	Y	6A	Y	Y
Anticipatory	Y		7B		
Responsive and Agile	Y	Y		Y	Y
Reflexive	Y	Y	7A		
Open to Differentiated Cooperation			5		
Transparent (see below)			6B		
Evidence Based (see below)				Y?	
Enforceable (see below)		Y?			
Holistic and Comprehensive (see below)		Y	1B		

(Table 1)

<sup>18</sup> Carnegie Council for Ethics in Artificial Intelligence, “ICGAI, Catalysing Cooperation: Working Together across AI Governance Initiatives” *Carnegie Council*, <https://www.carnegiecouncil.org/studio/multimedia/20210324-icgai-catalyzing-cooperation-artificial-intelligence-governance>

### *Reflections on Table 1*

The first five characteristics listed in the above table are relatively uncontroversial. While some might conflict with others, they are each desirable in their own right. However, the last four characteristics listed require further reflection:

- a. **Transparent:** This was only mentioned by Ala-Pietila and Smuha, but others such as Algorithm Watch see it as key. There is wide-spread support for the idea that trustworthiness is necessary<sup>19</sup>. Trust requires transparency. Thus, transparency should be an essential characteristic.
- b. **Evidence based:** There is much debate over the extent to which regulation should be ex-post or ex-ante. The EU argues that it should be a combination of both. In extremis, with a superintelligence, by the time the evidence has been gathered it is too late. Thus, the idea that a regime model should be ‘evidence-based’ is not necessarily recommendable.
- c. **Enforceable:** The need for law to be enforceable depends upon its significance. Soft law is not strictly enforceable, but norm setting and peer pressure can have a marked impact. Hard law is enforceable through the courts, but only in so far as the relevant parties are bound by that law’s jurisdiction. For something to be enforceable, (and to have the power to make it happen as would be essential with Advanced AI) a UN Charter review would be required as a minimum. In reality, the formation of a World Federation would likely be necessary to ensure full compliance.
- d. **Holistic and comprehensive:** At first sight, the requirement to be holistic and comprehensive seems obvious. But it does depend upon how the characteristic is interpreted. Paul Nemitz, Principal Advisor to the European Commission, argues in favour of a different approach, suggesting that the proposed agreement be stripped

down to the absolute minimum, as a way to avoid rehearsing all of the relevant human rights and consumer legislation in terms of AI<sup>20</sup>. He suggests focussing on technical aspects of AI, such as whether a system does what it is supposed to do<sup>21</sup>. Nemitz argues that this does not imply any backing down on human rights or democracy. Rather, he suggests that areas of serious disagreement should be addressed using existing human rights law. While this argument is controversial, it appears to offer significant potential. Nemitz’ thinking continues to evolve. His view as of September 2021 is that any agreement should be limited to the protection of governability and of human self-determination<sup>22,23</sup>. The essence of this argument is that the unnecessary duplication of legislation should be avoided where possible.

Embracing the reflections above, it would appear that the Regime Complex, as proposed within this paper, should have eight basic institutional qualities. These qualities are set out in Table 2 below.

---

<sup>19</sup> Jain. S., Luthra. M., Sharma. S. and Fatima. M, “Trustworthiness of Artificial Intelligence,” *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020, pp. 907-912, doi: 10.1109/ICACCS48705.2020.9074237.

---

<sup>20</sup> Nemitz, P, “Remaking the World – Toward an Age of Enlightenment” *Fundamentals of International Law: AI and Digital*, 2021a,

<https://bostonglobalforum.org/publications/remaking-the-world-the-age-of-global-enlightenment-2/>

<sup>21</sup> Nemitz, “Remaking the World”.

<sup>22</sup> Nemitz, P. “Policy Lab: Fundamental Rights In AI & Digital Societies - Towards An International Accord”, *Club De Madrid and Boston Global Forum Policy Lab: Fundamental Rights in AI and Digital Societies – Day 1*, 2021b,

<http://www.clubmadrid.org/policy-lab-fundamental-rights-in-ai-digital-societies-towards-an-international-accord/>

<sup>23</sup> Nemitz, “Fundamental Rights In AI & Digital Societies”.



### Institutional quality summary

Quality	Description
Inclusive	The regime should ensure that the voices of all stakeholders are heard, irrespective of power imbalance. All countries should be invited to participate, as well as all stakeholder groups.
Anticipatory	The regime should seek to anticipate future developments in AI.
Responsive and Agile	The regime should respond promptly to new events and should move rapidly to take the necessary action to maintain good governance of AI.
Reflexive	The regime should critically review and update its policy principles in a regular and timely manner, and in response to new events.
Open to Differentiated Corporation	There are likely to be issues that some countries feel cautiously about. Initially, at least, countries should be free to cooperate more closely with some countries than they do with others.
Transparent	The actions of the regime should be transparent, with a focus on building trust between stakeholders, between countries, and within the wider community.
Enforceable	The effectiveness of the enforcement needs to be broadly in line with the degree of seriousness of the specific issue.
Holistic / Comprehensive	The regime should ensure that its governance embraces all aspects of AI and all AI systems from initial conception to implementation and operation, subject to the avoidance of any duplication of legislation.

(Table 2)

Note that the above description relates to an AI Regime Complex, but would apply equally to a regime complex with a broader technological scope.

These eight characteristics are not all mutually independent. The main inter-relationships are described in Table 3 below.

X = Characteristics making flexibility more difficult  
Y = Characteristics demanding flexibility  
Z = Independent characteristics.

### The Inter-Relationship Between Basic Institutional Qualities

Source	X	Y	Z
<b>Basic Institutional Quality</b>			
Inclusive	*		
Anticipatory		*	
Responsive and Agile		*	
Reflexive			*
Open to Differentiated Cooperation			*
Transparent			*
Enforceable	*		
Holistic / Comprehensive	*		

(Table 3)

When designing the regime, the aim would be to:

- a. Optimise the trade-off between X and Y, seeking to achieve both sets of goals fully in time.
- b. Maximise the independent characteristics (Z) from the outset.

To conclude, the aim would be to:

- a. Optimise the trade-off between being responsive, agile, and anticipatory on the one hand, and inclusive, enforceable and being efficiently holistic and comprehensive on the other hand. In time, the aim would be to achieve both sets of goals.
- b. Ensure that the Regime Complex is reflexive and transparent, and that it is open to differentiated competition from the outset.
- c. Use AI to help manage the Regime Complex. The proposed Regime Complex would be difficult to achieve. At the outset, AI may be of some help. In later years, we can expect that AI will help a great deal.

### 3. GOVERNANCE FRAMEWORK OPTIONS

The above discussion sets out the key attributes required of a new AI Regime Complex. But, what else can one say about such a framework? What about the institutional structure? What should the relationship be between the different levels? What should the techno-logical breadth of scope be? Should the approach be to build on existing institutions or should it start afresh?

Each of these questions is discussed below.

#### *Institutional structure*

Alter and Raustiala argue that in practice, the governance relating to any major subject will not be a single element nor simply a regime, but rather a Regime Complex<sup>24</sup>, that is to say, an array of partially overlapping and non-hierarchical institutions that includes more than one international agreement or authority. This arrangement reflects what they assess to be a rise in international regime complexity. The point is well made, yet at the heart of their Regime Complex is still a core Regime, which needs to be well structured.

Various institutional models of a basic regime have been proposed, often based upon the institutional structure of an existing regime from another discipline.

#### 1. Single institutional solutions:

The two main categories of a single institutional model are a convention and an agency. UNI Global Union has called for the establishment of a Global Convention on the Ethical use, Development and Deployment of Artificial Intelligence, Algorithms and Big Data<sup>25</sup>. Zlatko Lagumdžija, a former Prime Minister of Bosnia Herzegovina, has called for the creation of an International Artificial Intelligence

<sup>24</sup> Alter, K. and Raustiala, K. "The Rise of International Regime Complexity", *Annual Review of Law and Social Sciences, UCLA School of Law, Public Law Research Paper*, 17-47, 2018, SSRN: <https://ssrn.com/abstract=3085043>

<sup>25</sup> UNI Global Union, "UNI Global Union calls for the establishment of a global convention on ethical artificial intelligence", <https://www.uniglobalunion.org/news/uni-global-union-calls-establishment-a-global-convention-ethical-artificial-intelligence>

Agency (IAIA), modelled on the International Atomic Energy Agency<sup>26</sup>. He makes this proposal against the background of a broader plan to establish an Artificial Intelligence International Accord (AIIA). Erdelyi and Goldsmith propose that an International Artificial Intelligence Organisation be established, initially as an informal organisation that can then evolve, becoming more formal and elaborate as agreement is reached on the best way forward<sup>27</sup>.

#### 2. Multi-institutional solutions / generators:

Others have proposed more complex solutions, modelled upon examples such as the United Nations Environment Programme (UNEP)<sup>28</sup>, the United Nations Convention on the Law of the Sea (UNCLOS)<sup>29</sup> or the United Nations Framework Convention on Climate Change (UNFCCC)<sup>30,31</sup>.

Kemp et al propose the creation of a Coordinator and Catalyser of International AI Law, modelled on the way in which UNEP has synchronised international agreements on the environment, while also facilitating new ones<sup>32</sup>. Kemp also suggests "new institutions could be brought together under an umbrella, as the World Trade Organisation (WTO) has done for Trade Agreements"<sup>33</sup>.

The specific components of the Kemp et al proposal included:

- a. A Coordinator and Catalyser of International AI Law
- b. An Intergovernmental Panel on AI (IPAI)

<sup>26</sup> Lagumdžija, Z. "Imagining a New World: AI World Society" *Remaking the World – Toward and Age of Enlightenment - Boston Global Forum*, <https://bostonglobalforum.org/publications/remaking-the-world-the-age-of-global-enlightenment-2/>

<sup>27</sup> Erdelyi and Goldsmith, "Regulating Artificial Intelligence".

<sup>28</sup> Kemp, Cihon and Maas et al., "A Proposal for International AI Governance"

<sup>29</sup> Nemitz, P. "Remaking the World".

<sup>30</sup> Whitfield, R. et al., "The Urgent Need for Good Global Governance of AI".

<sup>31</sup> Zilman, J. "A History of Climate Activities World", *Meteorological Organisation Bulletin*, 2009

<https://public.wmo.int/en/bulletin/history-climate-activities>

<sup>32</sup> Kemp, Cihon and Maas et al., "A Proposal for International AI Governance"

<sup>33</sup> *Ibid*

- c. A UN AI Research Organisation (UNAIRO)<sup>34</sup>.

Paul Nemitz proposes the use of the United Nations Convention on the Law of the Sea (UNCLOS) as an appropriate model<sup>35</sup>.

Whitfield et al use the United Nations Framework Convention on Climate Change (UNFCCC) as a model and propose:

- a. A UN Framework Convention on AI (or a broader digital / technological scope).
- b. A Protocol on AI.

To support the negotiation and implementation of these agreements, Whitfield et al propose the establishment of several new bodies, namely:

- a. An Intergovernmental Panel on AI, possibly building upon the Global Partnership on AI. This would provide technical support analogous to that provided by the IPCC to the UNFCCC.
- b. An AI Global Authority, empowered with powers of monitoring and inspection. This would support the work of the UN Framework Convention on AI.
- c. A supervisory body with democratic input, as exists in other treaty-based institutions<sup>36</sup>.

The specific components of the Regime Complex need to be considered carefully in the light of decisions within other fields discussed in this paper.

### ***Relationship between different levels***

A regulatory system can be characterised as having three distinct elements, namely a legislative source of legitimacy, an agency for policy-making and certification, and courts for adjudication<sup>37</sup>. Matthew Scherer sets out a coherent proposal as to how an effective AI regulatory system could work at the

national level<sup>38</sup>. He acknowledges that a global governance system would be most appropriate, but at the time of writing, considers it unrealistic. This model has potential at the global level, but this would require the application of international laws to businesses and their activities: the adjudication of tort law claims and the provision of individual remedies currently remain an issue<sup>39</sup>.

### ***Technological breadth of scope***

The question of the scope of the technology that the Regime / Regime Complex should cover is an open discussion. Both the High-Level Panel on Digital Cooperation, in its report *The Age of Digital Interdependence*<sup>40</sup>, and the UN Secretary General, in his *Roadmap for Digital Cooperation*<sup>41</sup> produced in response to the High-Level Panel report, see the need for greater cooperation in the digital world, without being explicit in terms of the ultimate extent or structure of any ensuing global governance.

A key issue concerning the breadth of any new regime is about overall organisational effectiveness. Looking back, the UN and other forms of global governance have grown organically, with a plethora of different types of organisation including agencies, programmes, funds, commissions and departments. The multitude of different conventions and COPs makes it very difficult for small countries to adequately represent their views. There has therefore been a trend towards the rationalisation of different Conventions<sup>42</sup>. For example, the Multilateral

<sup>34</sup> *Ibid*

<sup>35</sup> Nemitz, P. "Remaking the World".

<sup>36</sup> Whitfield, R. et al., "The Urgent Need for Good Global Governance of AI".

<sup>37</sup> Scherer, M. "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies", *Harvard Journal of Law & Technology* 29(2), 2016: 394  
[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=26097](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=26097)

77

<sup>38</sup> *Ibid*

<sup>39</sup> Sharkey, P. (2021a), "Artificial Intelligence and Global Regulation: Can Matthew Scherer's National Proposal be Implemented on the International Level?", to be published on [www.oneworldtrust.org](http://www.oneworldtrust.org) website in late 2021.

<sup>40</sup> United Nations, "The Age of Digital Interdependence", *United Nations Secretary General's High Level Panel on Digital Cooperation*, 2020, <https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf>

<sup>41</sup> United Nations, "Road map for digital cooperation: implementation of the recommendations of the Secretary General's High-level Panel on Digital Cooperation", 2020, <https://www.un.org/en/content/digital-cooperation-roadmap/>

<sup>42</sup> Oberthur, S. "Clustering of Multilateral Environmental Agreements: Potentials and Limitations", *UNU Archive*, 2002, <https://archive.unu.edu/inter-linkages/docs/IEG/Oberthur.pdf>



Environmental Agreements (MEAs) could be grouped together in terms of Water MEAs, Land MEAs, Atmosphere MEAs – and could ultimately all be consolidated into one coherent Environmental Regime.

Other factors include:

- a. The future significance of the issue concerned: in this case, AI will only become more important and indeed could become dominant, suggesting the need for a regime dedicated solely to AI.
- b. The degree of urgency: the broader the scope, the more elaborate the organisation and the longer its establishment is likely to take.
- c. Simplicity of establishment: The easier an institution is to establish, the quicker it can be operational. For instance, a dedicated AI Regime would be easier to establish than a broader, all-embracing digital regime complex.

The four main options would appear to be:

1. AI alone: this would be the most straightforward option, and would reflect the urgency of the need for AI regulation.
2. AI and other disruptive technologies: this depends upon the breadth of scope of the other disruptive technologies. It should not need to take longer than setting up a regime for AI alone if the other ‘disruptive technologies’ do not present problems until some time in the future. In this case, a regime framework can be established now, and a designated slot within the framework, such as a protocol dedicated to new technologies, can be established in due course as the need should arise. If a technology such as gene editing is included however, the design of an appropriate regime structure could be much more complex.
3. Embracing all digital: due to the very close links between different aspects of digital technology, there is clearly an argument for establishing a regime that embraces all

digital<sup>43</sup>. Since some governance structures already exist however, this would be a more complex route to follow and would take longer to establish.

4. Disruptive digital technologies: a subset of options (2) and (3), this could offer a practical compromise.

Currently, at a regional level, the main momentum is for a regime for AI alone (1 above). For the moment however, it is too early to tell whether this momentum will continue.

#### ***Adaptation or creation***

Another axis is whether the model is based upon:

- a. Extending the role of (a combination of) existing institutions: such an approach would speed up the creation of a global AI governance framework<sup>44</sup>.
- b. Creating entirely new institutions: such an approach would probably take longer to establish.
- c. An amalgam of the two approaches.

Whilst Ala-Pietila and Smuha argue in favour of the use of existing structures to establish effective AI global governance as soon as possible<sup>45</sup>, this would appear to ignore the significance of artificial intelligence today and the dramatically greater role that it is destined to play in the future. The governance of an increasingly dominant aspect of life on earth needs a dedicated regime. In the interim, valuable progress could be made through the use of existing institutions.

#### ***Conclusions***

There will be many aspects of an AI global governance framework, several of which are discussed above. Initial conclusions suggest that:

---

<sup>43</sup> Pentland, A. “Our Digital Future – from the Internet to the Interledger”, *UN at 100*, 2021, <https://ide.mit.edu/wp-content/uploads/2021/08/Our-Digital-Future-UN-2045.pdf>

<sup>44</sup> Ala-Pietila and Smuha, “Artificial Intelligence and its Governance”, 256/7

<sup>45</sup> *Ibid*, 256

- a. One should envisage the establishment of a Regime Complex.
- b. Whether that Regime Complex should be limited to AI or should have a broader technical scope needs further consideration. Current trends suggest that the regime would (at least initially) be limited to artificial intelligence (namely an AI Regime Complex or AIRC).
- c. Whilst the regime could be adapted from existing institutions, the ever-increasing significance of the subject matter strongly supports the argument that the main elements of the regime should be created from scratch

Gates has argued that the owners of robots should be taxed<sup>48</sup>. As the EU takes the initiative on AI governance, the US appears to be adjusting its position. The US argues in favour of innovation, whilst Europe insists that its approach will not stifle innovation<sup>49</sup>. Furthermore, the criteria that Kemp proposes are based upon a set of key criteria previously proposed by Stilgoe for responsible research and innovation, suggesting that responsible innovation can be achieved through suitable design criteria for the regime. There will, no doubt, be lively discussions on the issue, but it would seem that the concept of AI global governance is no longer unthinkable in the US.

#### 4. BIGGEST BLOCKERS

There have been a number of blockers preventing the development of AI Global Governance. Quite apart from needing to resolve all the questions rehearsed above, there are a number of geopolitical issues to consider: the US has been resistant, arguing that existing technology is too immature; some countries wish to establish their own AI strategies first before considering broader governance issues; and there is no agreement over the involvement of China, Russia, and other authoritarian states.

##### *The US and Big Tech*

An argument put forward by the AI sector, and by the US, has been that it is too early to regulate AI<sup>46</sup>. Certainly, it would be undesirable to set AI in concrete at today's level of development. There is a great deal of potential for AI to evolve. Positive evolution should be encouraged, but some monitoring and control is also appropriate, given the power of the technology that has already been revealed.

Within the US, there are many who are sensitive to the risks of uncontrolled AI. Big Tech companies, such as Facebook and Microsoft, have indicated their support for regulation. Mark Zuckerberg has called for the regulation of four areas, namely elections, harmful content, privacy and data portability<sup>47</sup>. Bill

##### *States' own strategies*

It has been suggested that global governance cannot be contemplated at present, given the fact that nation states are in the process of determining their own strategies. It could be argued that before countries determine their own strategies, and dig trenches to defend those strategies, we should aim to establish a global governance framework. However, at the time of writing, all major AI countries have developed their national strategies, and the OECD AI portal has noted over 700 national AI policies and strategies covering over 60 countries globally<sup>50</sup>. Whilst some nation states are yet to develop an AI strategy, that fact alone is unlikely to prevent the initiation of discussions around AI global governance.

[https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f\\_story.html](https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html)

<sup>48</sup> Medhora, R. "The Need for AI Governance", *Centre for International Governance Innovation (CIGI)*, 2018, <https://cpr.unu.edu/publications/articles/ai-global-governance-three-paths-towards-a-global-governance-of-artificial-intelligence.html>

<sup>49</sup> Kaili, E. "Policy Lab: Fundamental Rights In AI & Digital Societies - Towards An International Accord", *Club De Madrid and Boston Global Forum Policy Lab: Fundamental Rights in AI and Digital Societies – Day 1*, 2021,

<http://www.clubmadrid.org/policy-lab-fundamental-rights-in-ai-digital-societies-towards-an-international-accord/>

<sup>50</sup> OECD, "National AI policies & strategies", *OECD. AI Policy Observatory*, <https://oecd.ai/en/dashboards>

<sup>46</sup> Kharpal, A., "AI is in its infancy and its too early to graduate it, Intel CEO Brian Krzanich says", *CNBC*, 2017, <https://www.cnbc.com/2017/11/07/ai-infancy-and-too-early-to-regulate-intel-ceo-brian-krzanich-says.html>

<sup>47</sup> Zuckerberg, M., "The Internet needs new rules. Let's start in these four areas", *The Washington Post*, 2019,

### *China, Russia and other authoritarian states*

In the past couple of years there has been a surge of interest in AI governance, with several major initiatives running at regional and global levels. Arguably, the key perceived blocker relates to the geo-political situation and the inclusion of China, Russia, and other authoritarian states in a significant new global regime. Some argue that with the level of tension obtaining between these countries and the West, it would be inappropriate to seek to negotiate a new convention with them. On the other hand, China is unquestionably a major player in tech. Alibaba hosts twice as much e-commerce activity as Amazon, and Tencent runs the world's most popular super-app, with 1.2 billion users<sup>51</sup>. Today, 73 Chinese digital firms are worth over \$10 billion. Of China's 160 unicorns, half are in fields such as AI, big data and robotics<sup>52</sup>. Against this background, an AI global governance regime without China's involvement would appear somewhat meaningless.

On the other hand, there are serious concerns about the actions of both China and Russia in various spheres. Territorial issues, such as the South China Seas can perhaps be dealt with through naval manoeuvres and other measures as now. A key concern however relates to human rights: AI has close links with potential human rights abuses in fields such as facial recognition and surveillance; the UN Secretary General states in his Digital Roadmap that AI cooperation should be human rights based<sup>53</sup>; the EC proposal for Harmonised Rules on Artificial Intelligence<sup>54</sup> is closely bound up with human rights; and Council of Europe texts are bound up with issues relating to both human rights and democracy<sup>55</sup>.

<sup>51</sup>The Economist, August 14 2021, "Xi Jinping's assault on tech will change China's trajectory", *Leaders*, 2021  
[://www.economist.com/leaders/2021/08/14/xi-jinpings-assault-on-tech-will-change-chinas-trajectory](https://www.economist.com/leaders/2021/08/14/xi-jinpings-assault-on-tech-will-change-chinas-trajectory)

<sup>52</sup> *ibid*

<sup>53</sup> United Nations, "Road map for digital cooperation".

<sup>54</sup> European Commission, "Proposal for a Regulation of the European Parliament and of the European Council: Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts", *EUR-Lex*, 2021,  
<https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>

<sup>55</sup> Council of Europe, *CAHAI Council of Europe Ad hoc Committee on Artificial Intelligence*,  
<https://www.coe.int/en/web/artificial-intelligence/cahai>

Bing Song, Director of the Berggruen Institute China Centre, argues that one can consider three issue levels regarding AI governance<sup>56</sup>. At the most fundamental level (e.g. existential risk from advanced intelligence), China and the rest of the world are in agreement. At the surface level, China agrees with other countries that there is a need for controls and standards in order to ensure safety and effective world trade. But the problem is at the middle level, where there are different values, essentially around human rights and democracy. We are told that the Chinese Government would like a global AI regime<sup>57</sup>, presumably because it would allow their AI businesses to flourish. There are certain areas in which constructive engagement with China is particularly important. On climate change, for example, there is evidence that the US and China are prepared to cooperate<sup>58</sup>. The extent to which parties are willing to compromise in the area of AI and human rights remains to be seen.

One approach could be to establish a Framework Convention, addressing only those aspects of AI governance where there is a close alignment of views, and creating a Forum which could be used to debate the more contentious areas. Pending resolution of areas of disagreement, other aspects of AI governance could be provided through regional agreements.

Nemitz's approach, outlined in Section 2 above, provides a very different approach. This approach appears to have real merit, and thus, it requires further discussion. Nemitz's approach is linked to the Regime Complex argument. He argues that where AI is being used in a way that conflicts with human rights law, then that case should be pursued under human rights law, rather than new AI law. Human rights law would become a part of the AI Regime Complex rather than being duplicated within the AI Regime Convention itself.

<sup>56</sup> Mialhe, N. and Bing.s., "Cross-Looks – AI, Lost in Translation?", *Politico AI Summit 2020*,  
<https://www.politico.eu/wp-content/uploads/2020/03/POLITICO-AI-SUMMIT-2020-Program-3003.pdf>

<sup>57</sup> *Ibid*

<sup>58</sup> BBC News, "COP26: China and US agree to boost climate cooperation",  
<https://www.bbc.co.uk/news/science-environment-59238869>

## 5. A WAY FORWARD

Ala-Pietila and Smuha perceive an urgent need for global AI governance. They highlight the need for swift action, for the building upon existing cooperation structures, and for the development of a network of networks<sup>59</sup>. Whilst these issues are not characteristics of an AI regime per se, they reflect the urgency of the need for effective governance of some aspects of AI.

There is a need however to create an AI Global Governance Regime Complex that will successfully perform its role in the medium and (especially in the) long term as well. Such a regime needs to be built upon sound foundations. Its design should be developed with care.

All of this begs the question as to what an appropriate pathway towards achieving global governance of AI should look like. Should there be an interim step before seeking to establish such a regime? Whilst a detailed study of these next steps is beyond the scope of this paper, it is worth noting some of the existing initiatives in this area. These include initiatives by groups of governments, by groups of stakeholders and experts, and by the UN Secretary General.

### *Current initiatives leading to hard or soft law*

In the last few years international groupings have started negotiating international AI governance hard and soft law. Such initiatives include:

- a. **European Union:** The European Commission's Proposal for a Regulation laying down harmonised rules on AI (Artificial Intelligence Act) and amending certain Union legislative acts is currently the subject of negotiation between the Council of Ministers and the European Parliament<sup>60</sup>. This has been seen by some as establishing the way forward for other countries such as the BRICS countries<sup>61</sup>.
- b. **Council of Europe:** The Ad Hoc Committee on Artificial Intelligence (CAHAI) is examining the feasibility and potential elements of a legal framework for the development, design and application of artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law<sup>62</sup>.
- c. **UNESCO:** UNESCO has embarked on a two-year process to elaborate the first global standard-setting instrument on the ethics of artificial intelligence in the form of a Recommendation, following the decision of UNESCO's General Conference at its 40th session in November 2019. Currently, there is a focus on engaging with an intergovernmental process, and on negotiating the draft text of the Recommendation for possible adoption by UNESCO's General Conference at its 41st session in late 2021<sup>63</sup>.
- d. **OECD:** The OECD Principles on Artificial Intelligence<sup>64</sup> promote artificial intelligence that is innovative and trustworthy and that respects human rights and democratic values. They were adopted in May 2019 by OECD member countries when they approved the OECD Council Recommendation on Artificial Intelligence and are the first such principles that governments have signed up to<sup>65</sup>. Some see a nascent polycentric and fragmented AI regime gravitating around the OECD, which they consider to hold

<sup>59</sup> Ala-Pietila and Smuha, "Artificial Intelligence and its Governance".

<sup>60</sup> European Commission, "Proposal for a Regulation of the European Parliament and of the European Council"

<sup>61</sup> Cyman, D., Gromova, E., and Juchnevicius

. E., "Regulation of Artificial Intelligence", *BRICS Law Journal* 86–115, 2021,

DOI:10.21684/2412-2343-2021-8-1-86-115

<sup>62</sup> Council of Europe, "CAHAI Council of Europe Ad hoc Committee on Artificial Intelligence".

<sup>63</sup> UNESCO, "Elaboration of a Recommendation on the ethics of artificial intelligence", *UNESCO Artificial Intelligence*,

<https://en.unesco.org/artificial-intelligence/ethics>

<sup>64</sup> OECD, "Recommendation of the Council on Artificial Intelligence", *OECD Legal Instruments*, 2019,

<https://legalinstruments.oecd.org/en/instruments/OECD-LE-GAL-0449>

<sup>65</sup> OECD, "What are the OECD Principles on AI?", *The OECD AI Principles*,

<https://www.oecd.org/digital/artificial-intelligence/ai-principles/>

considerable epistemic authority and norm-setting power<sup>66</sup>.

Whilst the EU proposal would provide hard law for the EU, and UNESCO would potentially provide a global set of principles as soft law, there are no current proposals for any AI hard law that is global. A number of organisations have produced databases and portals to facilitate access to and monitoring of the various AI soft law and governance initiatives<sup>67,68,69</sup>.

### ***Current and proposed groupings***

There are relatively few NGOs active in the sector (in contrast for instance to the number of NGOs in the environmental sector at the time of development of Environmental Governance). There are however, various groups of experts, stakeholders and Governments who are not necessarily negotiating, but are debating AI governance with a view to sharing viewpoints and seeking to move towards a common position. Such groups and processes (actual and proposed) include:

- a. Wallach, Marchant, Kasperson et al: a group of experts has gradually developed, supported by such organisations as the Hastings Centre, with leading roles played by Wendell Wallach, Gary Marchant and Anja Kasperson. Starting in the early part of the last decade, various vehicles have been used to progress towards AI governance such as Building Agile Governance for AI and Robotics. The way forward has been articulated in a number of forms, such as a Governance Coordinating Committees<sup>70</sup>, a

G20 Initiative Coordinating Committee for AI Governance (CCGAI)<sup>71</sup> and a Global Governance Network for AI<sup>72</sup>. The proposed International Congress on the Governance of AI suffered from COVID delay and eventually became virtual, with the group now operating through the Artificial Intelligence and Equality Initiative (AIEI) of the Carnegie Institute of Ethics in Foreign Affairs. Wallach and Kasperson are concerned that the Ethics of AI is failing<sup>73</sup>.

- b. The Future Society: a not-for-profit think tank with a broad network, has played a formative role in facilitating the development of AI governance initiatives around the world.
- c. Global Alliance for Digital Governance / Boston Global Forum (BGF): a group of academics and politicians led by former Governor Michael Dukakis and Nguyen Anh Tuan. More political than the Wallach group, the BGF have collaborated with the Club de Madrid and other security organisations to form an international network committed to the global governance of AI. The group has drafted an outline Artificial Intelligence International Accord (AIIA)<sup>74</sup> to provide a background framework against which AI Global Governance can be negotiated. In September 2021 they launched a Global Alliance for Digital Governance.
- d. The Forum for Cooperation on AI: established by the Brookings Institute and CEPS, the Forum is exploring the opportunities for international cooperation on AI. It was initially set up as a trans-Atlantic body, bringing together high-level officials

<sup>66</sup> Schmitt, L. "Mapping global AI governance: a nascent regime in a fragmented landscape". *AI Ethics*, 2021. <https://doi.org/10.1007/s43681-021-00083-y>

<sup>67</sup> Gutierrez, C., Ignacio, C., Kearl, A., Marchant, G., Carden, A and Hoffner, K, "Preliminary Results of a Global Database on the Soft Law Governance of Artificial Intelligence", *IEEE / ITU International Conference on Artificial Intelligence for Good*, 2020. <https://ssrn.com/abstract=3756939>.

<sup>68</sup> OECD. *OECD AI Policy Observatory*, <https://oecd.ai/en/>

<sup>69</sup> Nesta, *AI Governance Database*, <https://www.nesta.org.uk/data-visualisation-and-interactive/ai-governance-database/>

<sup>70</sup> Marchant, G and Wallach, W., "Coordinating Technology Governance." *Issues in Science and Technology* 31(4), 2015

<sup>71</sup> Jelinek, T., Wallach, W. & Kerimi, D., "a G20 coordinating committee for the governance of artificial intelligence".

<sup>72</sup> The Carnegie Council for Ethics in International Affairs and the World Technology Network, *The International Congress for the Governance of AI*, <https://www.icgai.org>

<sup>73</sup> Kaspersen, A and Wallach, W., "Why Are We Failing at the Ethics of AI?" *Artificial Intelligence & Equality Initiative*, 2021, <https://www.carnegieaie.org/blog/why-are-we-failing-at-the-ethics-of-ai/>

<sup>74</sup> Choucri, N. "Framework for Artificial Intelligence International Accord", *Boston Global Forum*, 2021, <https://bostonglobalforum.org/news-and-events/news/framework-for-artificial-intelligence-international-accord/>



from the EU, the US, the UK and Canada, together with stakeholders from industry and AI experts from academia. More recently, it has expanded to include Australia, Japan and Singapore<sup>75</sup>.

- e. AI Governance Forum: a recent suggestion is that an AI Governance Forum be formally established, modelled on the Internet Governance Forum equivalent<sup>76</sup>. Whilst it would be essential to focus such a Forum on the goal of timely AI global governance, the community would benefit from a more formalised structure, bringing the relevant stakeholders together to address the main areas of contention.

### *The UN Secretary General's initiative*

The UN Secretary General, in his September 2021 *Our Common Agenda*<sup>77</sup> report (OCA), places great emphasis on listening to youth and ensuring that the global system places due emphasis on the interests of future generations. Specific measures proposed in this regard include the establishment of a Futures Laboratory; the appointment of a Special Envoy for Future Generations; the tasking of the Special Envoy to investigate the use of the Trusteeship Council as a means of representing Future Generations; a Declaration for Future Generations designed to build on the UNESCO Declaration of the Responsibility of the Present Generations to Future Generations<sup>78</sup> by specifying duties for succeeding generations and developing a mechanism to share good practices and monitor how governance systems address long-term challenges; a five-yearly Strategic Foresight and

Global Risk Report; and last but not least, a Summit of the Future in 2023.

Against this strong intergenerational equity background, OCA focusses on Global Public Goods and the Global Commons. In particular, the report proposes to reclaim the digital commons. Building on the Road-map for Digital Cooperation<sup>79</sup>, it suggests that “the United Nations, governments, the private sector, and civil society could come together as a multi-stakeholder digital technology track in preparation for a Summit of the Future to agree on a Global Digital Compact.”

In his June 2020 Digital Roadmap, the UN Secretary General proposed to establish “a multistakeholder advisory body on global artificial intelligence cooperation”<sup>80</sup>. It was felt that bringing the global community together to discuss common issues in an open and loose framework could provide an effective bridge to a more formal process establishing AI Global Governance. The OCA builds upon this commitment, suggesting that “the Compact could also promote regulation of artificial intelligence to ensure that this is aligned with global values”. The fact that the report uses the word “could”, referring to a proposal that was first published 15 months earlier, is disappointing. Within the context of Global Public Goods, one could question whether the term fits AI. However, it has been argued that whilst there is strong national competition, as evidenced by the national AI strategy reports published, cooperation might be more appropriate<sup>81</sup>. Thus, whether or not the term fits is less important than action being taken to establish effective global governance.

OCA provides additional back-up in the form of a High-level Advisory Board (HLAB) made up of former Heads of State and/or Government. The HLAB will be asked to “identify global public goods and potentially other areas of common interest where governance improvements are most needed” and to “propose options for how this could be achieved.”

<sup>75</sup>Kerry.C., Meltzer, J., Renda, A., Engler, A. and Fanni, R., “Strengthening international cooperation on AI”, *Brookings*, 2021,

<https://www.brookings.edu/research/strengthening-international-cooperation-on-ai/>

<sup>76</sup>Sharkey, P. “How the IGF has worked so far and improvements that could be made. Are there lessons for AI?”, *One World Trust*, 2021b,

<https://www.oneworldtrust.org/blogs>

<sup>77</sup> UN Secretary General, “Our Common Agenda”, United Nations, 2021,

<https://www.un.org/en/content/common-agenda-report/>

<sup>78</sup> United Nations Educational, Scientific and Cultural Organization, “Declaration on the Responsibilities of the Present Generations Towards Future Generations”, UNESCO, 1997,

[http://portal.unesco.org/en/ev.php-URL\\_ID=13178&URL\\_DO=DO\\_PRINTPAGE&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=13178&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html)

<sup>79</sup>United Nations, “Road map for digital cooperation”.

<sup>80</sup> *Ibid*

<sup>81</sup> Ramnath, N. “AI as a global public good, a digitised world, and why philosophy matters”, *Founding Fuel*, 2019, <https://www.foundingfuel.com/column/this-week-in-disruptive-tech/ai-as-a-global-public-good-a-digitised-world-and-why-philosophy-matters/>

These recommendations are meant to feed into the Summit of the Future in 2023.

Both the development of a Global Digital Compact and the HLAB process will also provide an opportunity to discuss whether AI should be included in the Digital Compact or have its own regime.

It is important to bear in mind that a major point of concern regarding AI global governance relates to the risks to future human generations potentially posed by AI, both in relation to the control problem<sup>82</sup> and issues of human dignity and self-determination. Threats to self-determination are already being faced. While these immediate issues need to be urgently addressed, the framing of “looking to the future” is still highly appropriate.

Crucially, the proposed steps and timetable leading to a Summit of the Future provide the AI governance community with a major opportunity to make the case for launching a formal AI Global Governance process, a process that could begin by scheduling a World Conference on AI<sup>83</sup>.

## 5. CONCLUSIONS AND RECOMMENDATIONS

It is clear that AI will become an increasingly dominant force in life on earth over the coming years and decades. There are aspects of AI that require urgent global action, such as Lethal Autonomous Robots, deepfakes, and attention driven algorithms. Yet, the need for effective global governance of AI will only grow.

This paper has sought to explore the key characteristics of a global regime to address such matters. Whilst there is a limit to the degree of detail of any particular plan<sup>84</sup>, there is a growing consensus over those characteristics. However, the institutional design necessary to achieve such characteristics needs further thought and debate, particularly the basic

architecture and scope of an AI global governance regime

The Summit of the Future, and the process leading up to it, offers a real opportunity for AI Global Governance.

A full set of conclusions and recommendations are set out at the beginning of this paper.

---

---

<sup>82</sup> Russell, S. *Human Compatible – AI and the Problem of Control*, (Allen Lane, 2019).

<sup>83</sup> Whitfield et al. “The Urgent Need for Good Global Governance of AI”, 61.

<sup>84</sup> Dafoe, A. “AI Governance: Opportunity and Theory of Impact”, *Effective Altruism Forum*, <https://forum.effectivealtruism.org/posts/42reWndoTEhFqu6T8/ai-governance-opportunity-and-theory-of-impact>

## BIBLIOGRAPHY

Ala-Pietila, P. and Smuha, N. 'A Framework for Artificial Intelligence and its Governance' *Reflections on Artificial Intelligence for Humanity* 2021.

Alter, K. and Raustiala, K. "The Rise of International Regime Complexity", *Annual Review of Law and Social Sciences*, 2018, *UCLA School of Law, Public Law Research Paper*, 17-47, 2018, SSRN: <https://ssrn.com/abstract=3085043>

BBC News, "COP26: China and US agree to boost climate cooperation", <https://www.bbc.co.uk/news/science-environment-59238869>

Carnegie Council for Ethics in Artificial Intelligence, "ICGAI, Catalysing Cooperation: Working Together across AI Governance Initiatives" *Carnegie Council*, <https://www.carnegiecouncil.org/studio/multimedia/20210324-icgai-catalyzing-cooperation-artificial-intelligence-governance>

Choucri, N. "Framework for Artificial Intelligence International Accord", *Boston Global Forum*, 2021, <https://bostonglobalforum.org/news-and-events/news/framework-for-artificial-intelligence-international-accord/>

Cihon, P., Maas, M. and Kemp, L., "Fragmentation and the Future: Investigating Architecture for International AI Governance", *Global Policy*, 111 (5), 2020.

Club de Madrid and Boston Global Forum, "Fundamental Rights in AI and Digital Societies: Towards an International Accord" 2021. <https://aidigitalrights.com/wp-content/uploads/2021/09/Club-de-Madrid-Boston-Global-Forum-Governance-Issues-Paper-Workshop-September-2021.pdf>

Council of Europe, *CAHAI Council of Europe Ad hoc Committee on Artificial Intelligence*, <https://www.coe.int/en/web/artificial-intelligence/cahai>

Dafoe, A. "AI Governance: Opportunity and Theory of Impact", Effective Altruism Forum, <https://forum.effectivealtruism.org/posts/42reWndoTEhFqu6T8/ai-governance-opportunity-and-theory-of-impact>

Deloitte and the World Economic Forum, "Harnessing Fourth Industrial Revolution Technologies in a COVID 19 World 2021", *Global Technology Governance Report*

<https://www.weforum.org/reports/global-technology-governance-report-2021>

Dodds, F. *Stakeholder Democracy – Represented Democracy in a Time of Fear* (Routledge, 2019)

European Commission, "Proposal for a Regulation of the European Parliament and of the European Council: Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts", *EUR-Lex*, 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>

Erdelyi, O., and Goldsmith, J., "Regulating Artificial Intelligence: Proposal for a Global Solution". *2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18), February 2--3, 2018, New Orleans, LA, USA*, 2018, doi/10.1145/3278721.3278731.

Gutierrez, C., Ignacio, C., Kearl, A., Marchant, G., Carden, A and Hoffner, K, "Preliminary Results of a Global Database on the Soft Law Governance of Artificial Intelligence", *IEEE / ITU International Conference on Artificial Intelligence for Good*, 2020. <https://ssrn.com/abstract=3756939>

The Carnegie Council for Ethics in International Affairs and the World Technology Network, *The International Congress for the Governance of AI*, <https://www.icgai.org>

Jain, S., Luthra, M., Sharma, S. and Fatima, M, "Trustworthiness of Artificial Intelligence," *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020, pp. 907-912, doi: 10.1109/ICACCS48705.2020.9074237.

Jelinek, T., Wallach, W. & Kerimi, D, "Policy brief: the creation of a G20 coordinating committee for the governance of artificial intelligence". *AI Ethics* 1, 141-150, 2021.

Kaspersen, A and Wallach, W., "Why Are We Failing at the Ethics of AI?" *Artificial Intelligence & Equality Initiative*, 2021, <https://www.carnegieaie.org/blog/why-are-we-failing-at-the-ethics-of-ai/>

Kharpal, A., "AI is in its infancy and its too early to graduate it, Intel CEO Brian Krzanich says" *CNBC*, 2017, <https://www.cnbcm.com/2017/11/07/ai-infancy-and-too-early-to-regulate-intel-ceo-brian-krzanich-says.html>

Kemp, L., Cihon, P., Maas, M. et al, *UN High-level Panel on Digital Cooperation: A Proposal for*

- International AI Governance*, 2019, <https://www.cser.ac.uk/news/advice-un-high-level-panel-digital-cooperation/>
- Kerry, C., Meltzer, J., Renda, A., Engler, A. and Fanni, R., “Strengthening international cooperation on AI”, Brookings, 2021, <https://www.brookings.edu/research/strengthening-international-cooperation-on-ai/>
- Lagumdzija, Z. “Imagining a New World: AI World Society”, *Remaking the World – Toward and Age of Enlightenment - Boston Global Forum*, <https://bostonglobalforum.org/publications/remaking-the-world-the-age-of-global-enlightenment-2/>
- Medhora, R. “The Need for AI Governance”, *Centre for International Governance Innovation (CIGI)*, 2018, <https://cpr.unu.edu/publications/articles/ai-global-governance-three-paths-towards-a-global-governance-of-artificial-intelligence.html>
- Nemitz, P., “Remaking the World – Toward and Age of Enlightenment” *Fundamentals of International Law: AI and Digital*, 2021a, <https://bostonglobalforum.org/publications/remaking-the-world-the-age-of-global-enlightenment-2/>
- Nemitz, P. “Policy Lab: Fundamental Rights In AI & Digital Societies - Towards An International Accord”, *Club De Madrid and Boston Global Forum Policy Lab: Fundamental Rights in AI and Digital Societies – Day 1*, 2021b, <http://www.clubmadrid.org/policy-lab-fundamental-rights-in-ai-digital-societies-towards-an-international-accord/>
- Nesta, *AI Governance Database*, <https://www.nesta.org.uk/data-visualisation-and-interactive/ai-governance-database/>
- Oberthur, S. “Clustering of Multilateral Environmental Agreements: Potentials and Limitations”, *UNU Archive*, 2002, <https://archive.unu.edu/inter-linkages/docs/IEG/Oberthur.pdf>
- OECD, “National AI policies & strategies”, *OECD AI Policy Observatory*, <https://oecd.ai/en/dashboards>
- OECD, “Recommendation of the Council on Artificial Intelligence”, *OECD Legal Instruments*, 2019, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- OECD. *OECD AI Policy Observatory*, <https://oecd.ai/en/>
- OECD, “What are the OECD Principles on AI?”, *The OECD AI Principles*, <https://www.oecd.org/digital/artificial-intelligence/ai-principles/>
- Pentland, A. “Our Digital Future – from the Internet to the Interledger”, *UN at 100*, 2021, <https://ide.mit.edu/wp-content/uploads/2021/08/Our-Digital-Future-UN-2045.pdf>
- Ramnath, N. “AI as a global public good, a digitised world, and why philosophy matters”, *Founding Fuel*, 2019, <https://www.foundingfuel.com/column/this-week-in-disruptive-tech/ai-as-a-global-public-good-a-digitised-world-and-why-philosophy-matters/>
- Russell, S. *Human Compatible – AI and the Problem of Control*, (Allen Lane, 2019).
- Sharkey, P., “Artificial Intelligence and Global Regulation: Can Matthew Scherer’s National Proposal be Implemented on the International Level?”, to be published on [www.oneworldtrust.org](http://www.oneworldtrust.org) website in late 2021a.
- Sharkey, P. “How the IGF has worked so far and improvements that could be made. Are there lessons for AI?”, One World Trust, 2021b, <https://www.oneworldtrust.org/blogs>
- Scherer, M. “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”, *Harvard Journal of Law & Technology* 29(2), 2016: 394 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2609777](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609777)
- Schmitt, L. “Mapping global AI governance: a nascent regime in a fragmented landscape”. *AI Ethics*, 2021. <https://doi.org/10.1007/s43681-021-00083-y>
- Stilgoe, J., Owen, R., and Macnaghten, P. “Developing a Framework for Responsible Innovation”, *Research Policy*, 42(9), 2013: 1568-1580 <https://www.sciencedirect.com/science/article/pii/S048733313000930>
- The Economist, August 14 2021, “Xi Jinping’s assault on tech will change China’s trajectory”, Leaders, 2021 [://www.economist.com/leaders/2021/08/14/xi-jinpings-assault-on-tech-will-change-chinas-trajectory](https://www.economist.com/leaders/2021/08/14/xi-jinpings-assault-on-tech-will-change-chinas-trajectory)

UN Secretary General, “Our Common Agenda”, United Nations, 2021,  
<https://www.un.org/en/content/common-agenda-report/>

[https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f\\_story.html](https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html)

United Nations, “The Age of Digital Interdependence”, *United Nations Secretary General's High Level Panel on Digital Cooperation*, 2020,  
<https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf>

UNI Global Union, “UNI Global Union calls for the establishment of a global convention on ethical artificial intelligence”,  
<https://www.uniglobalunion.org/news/uni-global-union-calls-establishment-a-global-convention-ethical-artificial-intelligence>

Wallach, W., Marchant, G. “Toward the Agile and Comprehensive International Governance of AI and Robotics”, *Proceedings of the IEEE*, 107(3), 2019  
<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8662741>

UNESCO, “Elaboration of a Recommendation on the ethics of artificial intelligence”, *UNESCO Artificial Intelligence*,  
<https://en.unesco.org/artificial-intelligence/ethics>

United Nations, “Road map for digital cooperation: implementation of the recommendations of the Secretary General’s High-level Panel on Digital Cooperation”, 2020,  
<https://www.un.org/en/content/digital-cooperation-roadmap/>

United Nations Educational, Scientific and Cultural Organization, “Declaration on the Responsibilities of the Present Generations Towards Future Generations”, UNESCO, 1997,  
[http://portal.unesco.org/en/ev.php-URL\\_ID=13178&URL\\_DO=DO\\_PRINTPAGE&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=13178&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html)

Whitfield, R. et al., *Effective, Timely and Global – The Urgent Need for Good Global Governance of AI*, 2020,  
<https://www.wfm-igp.org/publication/effective-timely-and-global-the-urgent-need-for-good-global-governance-of-ai/>

Zilman, J. “A History of Climate Activities World”, *Meteorological Organisation Bulletin*, 2009  
<https://public.wmo.int/en/bulletin/history-climate-activities>

Zuckerburg, M. “The Internet needs new rules. Let’s start in these four areas”, *The Washington Post*, 2019,