AI AND GLOBAL REGULATION:
CAN MATTHEW SCHERER'S NATIONAL PROPOSAL
BE IMPLEMENTED ON THE INTERNATIONAL LEVEL?

1. INTRODUCTION

The world of technology moves quickly and thus far has waited for no one. Artificial Intelligence (’AI’) is the latest in a series of technological developments which have changed the world and it seems set to play an increasingly important role in our lives. As with many developments which have the potential to inexorably alter how we live, the question of how to control and regulate this technology has become more pressing the more capable AI becomes. On the face of it, it appears to be a particularly difficult area to regulate, if for no other reason than the difficulty in understanding how it actually operates. However, this does not mean regulation is not possible. Whilst there are numerous regulatory avenues to go down, and a number of different proposals, this Article will focus on the proposal by Matthew Scherer, contained in his article “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”. Scherer’s paper focuses on the implementation of a regulatory system at the national level (in his case the USA). This paper will explore whether Scherer’s ideas for regulating AI could work at the international level. As such, the focus will be on whether those ideas could translate to the international level, not necessarily an analysis of the intrinsic values of his ideas, or whether there are preferable alternatives internationally.

2. OUTLINING SCHERER’S PROPOSAL

Legislative

The foundation of Scherer’s proposal is legislation. This legislation would establish an agency which would be responsible for certifying the safety of AI programmes. It would also set out the parameters of the agency’s authority and powers. In Scherer’s proposed national act, AI systems that had not received agency certification would not be banned, however, any company which developed, sold, or operated AI without an agency certification “would be strictly liable for any harm caused by the AI”3. Furthermore, this liability would be joint and several, meaning a plaintiff could recover the full amount of damages from any company that had been involved at any stage in the development, distribution, sale or operation of the uncertified AI. On the other hand, systems which had obtained agency certification would be granted limited tort liability, providing a partial regulatory compliance defence. This would limit tort liability to cases where it could be

---

1 See e.g.; John O. McGinnis, Accelerating AI, 104 NW. U. L. REV. 1253, 1262 (2010); David C. Vladeck, Machines Without Principals: Liability Rules and Artificial Intelligence, 89 WASH. L. REV. 117, 121 (2014)


3 ibid, at 394
established that there had been actual negligence in the design, manufacturing or operation of the system\(^4\).

Scherer’s proposed bill would contain a grandfather clause, presumptively exempting programs which had been in commercial operation for 12 months before the bill’s enactment; however he does recommend that the agency be granted authority to create a separate mechanism with which to review existing AI which potentially presents a risk to the public. He contends that aside from the specific rules he lays out\(^5\), the legislation would grant the agency authority to specify and clarify most of the aspects of the regulatory framework.

**Agency**

In this section, Scherer sets out his vision for the structure and functioning of the agency which the proposed act would create. In his proposal it would have two components: policymaking and certification.

**Policymaking**

The policymaking body would be tasked with defining AI\(^6\), as well as ensuring that the need for future review and potential updating of the definition was laid out. This body would also be responsible for establishing the AI certification process\(^7\). Scherer recommends that a Board of Governors (‘the Board’) be appointed\(^8\), subject to legislative approval. The Board would be responsible for rule making and conducting public hearings on proposed rules and amendments. They would also be responsible for publishing the substantive standards under which any application for certification would be judged.

**Certification**

The primary role of agency staff would be to assess whether AI systems met the substantive standards published by the Board. The agency would have the responsibility of promulgating pre-certification testing rules, with information resulting from this testing required for any application for certification. If the testing was conducted in compliance with agency rules, it would not be subject to strict liability. Amendments to testing requirements could be fast-tracked by the agency, however, these would need to be subsequently ratified.

With respect to the certification process itself, Scherer recommends that the agency have a number of powers, including to limit the scope of certification, for example allowing the AI to only be used in certain circumstances, and to fast track certification for an AI system that has already certified safe to be used in a different context\(^9\). The agency would also disseminate rules to govern licensing and warning notice requirements for AI, for instance specifying that a designer/manufacturer would lose its liability protection if it sold its product without a licensing agreement forbidding any on-sellers from modifying the AI system. This rule would help ensure that the product that ultimately reached the end user was the same one that had received agency certification\(^10\).

**The courts**

Under Scherer’s national proposal, the legislative framework would require the courts to adjudicate individual tort claims for any harm caused by an AI system, within the liability framework set out in the legislation. Part of this will entail allocating responsibility along the production line for cases involving uncertified AI. Scherer believes it is likely that parties will dispute whether the version of the AI system which is at the centre of the issue was one which had been agency certified, or that there will be a dispute regarding at what point any modifications took the AI outside of that which had received certification. He argues that in such cases, the court would undertake a pre-trial hearing to assess whether the product was a certified version of the system at the time of the harm caused, and if not, decipher at which point the product deviated from the certified version. Any modification point would then become the dividing line between defendants who enjoy limited liability, and those defendants who would be subject to strict liability.

3. **COULD THESE IDEAS BE IMPLEMENTED ON THE INTERNATIONAL LEVEL?**

**Legislative**

researchers being subjected to strict liability, *supra* note 2, at 395

\(^4\) *ibid*

\(^5\) *ibid*, at 394-395

\(^6\) Any definition would require legislative ratification, as this would play a major role in determining the scope of the Agency’s jurisdiction, *supra* note 2, at 395

\(^7\) Including establishing exemptions for AI research to allow it to be conducted in certain environments without the
It would seem that the most prudent means with which to translate the legislative pillar of Scherer’s idea from the national to the international, would be by way of a treaty, often referred to as a convention. One of the main sources of international law\(^1\), treaties are a hard law instrument, meaning that the legal obligations laid down within them are binding on the States which ratify them. Any State which is a party to an international treaty must ensure that their own domestic law and practices are consistent with the requirements of that treaty\(^2\). This has the effect of bringing international rules into the national sphere.

The UN would seem to be the model to follow in this regard. Self-styled as the “world’s only truly universal global organisation”, it claims it has become the “foremost forum to address issues that transcend national boundaries and cannot be resolved by any one country acting alone\(^3\). Notable international treaties which could provide useful guidance for the potential international regulation of AI include the United Nations Convention on the Law of the Sea\(^4\) (‘UNCLOS’) and the United Nations Framework Convention on Climate Change\(^5\) (‘UNFCCC’).

In Scherer’s national proposal, the legislative act would establish an agency responsible for certifying the safety of AI systems, whilst also setting the parameters of the agency’s authority and powers. A UN Convention on AI could legislate along the same lines. The Law of the Sea for example establishes the International Seabed Authority\(^6\). It also grants it the power to establish any regional centres/offices as it deems necessary for the exercise of its functions\(^7\). The Convention expressly confers powers and functions upon this Authority, as well as providing it with implicit incidental powers it needs to adequately exercise its powers and carry out its functions\(^8\). Much in the same way, Scherer’s proposed legislation would grant an agency authority of its own to refine and clarify aspects of the regulatory framework. Any AI agency could also be granted the powers to review existing AI systems which could potentially present a risk to the public; again, this will be dependent on the scope of authority the agency is granted by the convention.

The foundation of any effective AI regulation rests on agreeing a definition. It is key for any successful regulatory system to set out what it is they are actually regulating. Scherer does not propose a formal definition, but argues that any definition should be reviewed periodically by the agency, and amended if necessary to take account of any changes in the industry\(^9\). Treaties can, and frequently are amended over time, so there should be no issue with taking this approach to defining the topic.

An area which promises to be a highly contentious aspect of AI regulation is that of liability. As outlined above, Scherer’s approach would be to implement a certification system, which would provide limited tort liability to successful applicants and result in their having a partial regulatory compliance defence. Any AI system which had not received agency certification however would be strictly liable for any harm that the AI caused. Whilst there can and will be many disputes over what the optimal approach to take in this regard is, there would seem to be nothing at the international level which would forbid the introduction of the liability regime outlined by Scherer.

Similarly, there is no reason why a grandfather clause of the type recommended by Scherer could not be included. The recent European Commission proposals on AI for example contain similar caveats about AI systems already on the market\(^10\).

**Agency**


\(^{16}\) Supra note 14, Section 4, Article 156

\(^{17}\) Supra note 14, Article 156(5)

\(^{18}\) Supra note 14, Article 157(2)

\(^{19}\) Supra note 2, at 396

Scherer envisages his national agency to have two components, policymaking and certification. This structure, and how it is intended to function, is something that can be replicated on the international level. One option to follow would be for any treaty to create a UN Specialised Agency which deals directly with AI issues. A specialised agency would operate under the umbrella of the UN, but they would be autonomous, with their own governing body, procedural rules and membership. These agencies can play a policy and standard setting role. If we look at, for example, the International Civil Aviation Organisation (‘ICAO’), this UN Specialised Agency helps shape the principles of international air transport policy and standardisation. The ICAO Council for example adopts standards and recommended practices concerning various aspects of air travel, which is the type of policymaking influence that Scherer advocates his agency should have.

The benefit of conducting this thought experiment with Scherer’s work, is that multiple options can be looked at, but there is no need to necessarily fully replicate an existing structure. Rather, multiple ideas can be considered and supplement one another. For example, as alluded to above, Section 4 of the UNCLOS outlines the establishment and structure of the International Seabed Authority. This offers another potential avenue for replicating Scherer’s idea of an agency which regulates the certification of AI, for many of the reasons discussed above. To add to this, and better replicate Scherer’s ideas about how an agency should be set up, we can use something like this as a framework, and then take the idea of for example The European Commission proposals on AI, which advocate that a governance system be set up at both the Union and national level. Much like Scherer, this proposal establishes an AI Board (‘the Board’); in Scherer’s proposal this is populated by persons appointed by the Executive Branch, whilst in the European proposal, it is composed of representatives from Member States and the European Commission. Even though both proposals envisage slightly different approaches that the Board should take, they both see the Board as playing a role in implementing the regulations. There is no reason therefore that any Board could not be more closely aligned to Scherer’s view of how it should function, rather than how the European Commission envisages it to function.

One area in the European Proposal which could assist in the effective implementation of Scherer’s ideas on the international level is found in Article 59 - which provides that “national competent authorities shall be established/designated by each Member State for the purpose of ensuring the application and implementation of this Regulation.” It would make sense for there to be a two-tier approach to global regulation, that is, an overarching international regulatory agency, supported by national agencies who would supervise the application and implementation of the regulations at the national level. This aspect of the agency would fulfil the certification role envisaged by Scherer.

Indeed, the advantage of this would be that if there was harmony with standards amongst UN Member States, gaining adequate accreditation in one country under the rules of the Convention could and perhaps should result in accreditation across all Member State countries. This would possibly allay some fears about the difficulties of global bureaucracy and help ensure that AI certification could be an expeditious process.

As to the question of whether such a set up would unduly impact on a State’s national sovereignty, some argue that international organisations operating at the regional level do not necessarily threaten the current conception of Statehood. De Brabandere argues that as long as the international organisations are seen to derive legitimacy from the powers that member States have conferred to them, and they exercise their powers in conformity with those granted to them, the position of States can be reinforced, rather than weakened. As discussed above, these agencies would be created by legislation (in the form of a treaty), which would set the parameters of their power. The act of acceding to such a treaty would indicate a particular State’s endorsement of such a structure. This act would be both an exercise of the

24 Supra note 14, Article 156(1)
25 Scherer’s views on the responsibilities of the Board outlined in the ‘Policymaking’ paragraph, at page 2, above
26 Supra note 20, Article 59
State’s sovereignty and simultaneously a conferral of legitimacy on the agency.

**The courts**

Under Scherer’s national proposal, the legislative framework would require the courts to adjudicate individual tort claims for any harm caused by the AI system in question, as set out by the legislative framework. This will also involve adjudicating on questions of responsibility along the production line for cases which involve uncertified AI. Legal systems and indeed individual countries can and will approach questions of tort law differently, however, it is mostly concerned with imposing liability and awarding damages for wrongful acts or an infringement of rights. For example, in the US the purpose of the system is “to deter people and companies from injuring others, and to compensate injured parties”\(^{28}\), in the UK “tort law is concerned with civil wrongs”\(^{29}\) and both the imposition of liability and remedies for the wrongs suffered.

This seems to be the most difficult part of Scherer’s proposal to translate to the international level, mainly because generally tort focuses on wrongs done by individuals to individuals, or by, for example, companies to individuals. International law however tends to focus on the interaction between Nation States or between States and their citizens. The key question in this regards is how an international agreement ratified by States, that is, a treaty, would become binding on companies or private persons. The U.N. Human Rights Council (“UNHRC”) has made an attempt to introduce such a mechanism, establishing an open-ended working group in 2014 to develop “an international legally binding instrument to regulate, in international human rights law, the activities of transnational corporations and other business enterprises”\(^{30}\). This proposed treaty would hold corporations directly liable under international law for violating human rights\(^{31}\).

There are already criticisms of the UNHRC proposals, with questions still abounding as to how such an aim would be achieved as a matter of law, whether corporations would be expected to become signatories to the treaty, or if somehow the treaty would automatically bind corporations which fell within its scope\(^{32}\). There has for example already been a rejection of proposals to extend the jurisdiction of the International Criminal Court to legal persons at the Rome Conference\(^{33}\). A detailed analysis of whether such a proposal would be possible under international law is beyond the scope of this article, and whilst it may eventually prove to be an insurmountable hurdle, the fact that the idea has been floated means that it is worth considering how such an idea, if successfully passed, could be applied to AI regulation.

The most recent iteration of these proposals, which take the form of a revised third draft issued by the Office of the United Nations High Commissioner for Human Rights (“UNHCHR”) (hereafter ‘the third draft’), states that the proposed legally binding instrument would apply to all business activities\(^{34}\) and provides that “States Parties shall regulate effectively the activities of all business enterprises within their territory, jurisdiction, or otherwise under their control, including transnational corporations”\(^{35}\). There are potential difficulties here, as even though a proposed agency could be brought into being via State passed legislation, Scherer envisages the agency running as an “independent administrative entity”\(^{36}\). The third draft further recommends that States Parties shall take all appropriate legal and policy measures to ensure that the aforementioned business enterprises “respect internationally recognised human rights and prevent and mitigate human rights abuses throughout their business activities and relationships”\(^{37}\). It would likely be unfeasible to grant any agency operating in an independent capacity anywhere near this level of power, and if the State were to take the lead regulatory role themselves, then we have strayed away from a key component of Scherer’s proposals.

---


31 ibid


33 ibid


35 ibid, at Article 6.1

36 Supra note 2, at 396

37 Supra note 34, at Article 6.2
It is still worth looking at the rest of the third draft, and indeed other current conventions to see what other potentially useful ideas or proposals could come into play in the AI regulatory space. The third draft, for example, does provide for access to remedies. The States Parties would provide their courts and State-based non-judicial mechanisms to enable victims to have access to remedy and justice. It also addresses questions of legal liability, stating that States Parties must ensure their domestic law provides for “a comprehensive and adequate system of local liability of legal and natural persons conducting business activities within their territory”.

This proposal is in line with Scherer, in that he too suggests that a system of legal liability be passed in domestic law, the remaining issue being that of applying it to legal and natural persons, which as mentioned above, is unresolved and likely to be a controversial topic.

**Jurisdiction**

The ease with which AI can permeate national borders can lead to some difficult issues of jurisdiction. The UNHRC proposals do address this issue, with Article 9 stating that jurisdiction shall vest in the courts of the State where:

(a) The incident occurred and/or produced effects; or
(b) An act or omission contributing to the incident occurred; or
(c) The legal/natural persons alleged to have committed an act or omission causing or contributing to such an incident in the context of business activities, including those of a transnational character are domiciled; or
(d) The victim is a national of or is domiciled in that State.

**Resolution of disputes**

The UNFCCC and UNCLOS can provide ideas for how to approach the question of resolving disputes arising from tort claims. Both treaties include articles which provide for the settlement of disputes required. Arguably UNCLOS provides a better framework for transposing Scherer’s tort liability conceptions to the international level. Section 5 of UNCLOS outlines the settlement of disputes and advisory opinions, establishing the Seabed Disputes Chamber and setting its jurisdiction. This potentially is very important, as it outlines the different types of disputes that the Chamber has jurisdiction over, including who is involved (i.e. in this case States Parties) and the disputes it covers (for example acts or omissions which directly affect the legitimate interests of a party).

In any proposed AI regulation, this general framework could be followed, with the provisions tailored to provide for the settlement of disputes involving individuals and companies. This could help ensure that a framework could be built that would allow for the provision of remedies to wronged parties under Scherer’s liability framework. This of course would be dependent on the treaty being applicable to companies, which as discussed above, is far from certain.

Article 287 of UNCLOS provides for a State to have a choice of procedure for the settlement of disputes concerning the interpretation or application of the Convention. Four options are given:

(a) The Seabed Disputes Chamber of the International Tribunal for the Law of the Sea
(b) The International Court of Justice
(c) An arbitral tribunal
(d) A special arbitral tribunal for a number of particular cases which are specified in a separate Annex.

The Convention stipulates the location of the seat of the arbitral tribunal, which is something any AI convention can easily replicate. Furthermore, Article 188(2)(c) outlines that “in the absence of a provision in the contract on the arbitration procedure to be applied in the dispute, the arbitration shall be conducted in accordance with the UNCITRAL Arbitration Rules or such other arbitration rules as may be prescribed in the rules, regulations and procedures of the Authority, unless the parties to the dispute otherwise agree.” This should cover any disputes surrounding the rules under which the arbitration should be conducted and so provide a strong framework that would ensure worldwide accountability for AI systems which caused harm.

Any court/tribunal under this section would also have the ability, in consultation with the parties, to select a number of scientific or technical experts to assist the court/tribunal, in a non-voting capacity. The inclusion of a clause such as this would be extremely helpful for an area as complicated as AI. All decisions

---

38 ibid, at Article 7.1
39 ibid, at Article 8.1
40 In the case of the UNFCCC, Article 14, supra note 15; and for the UNCLOS, Article 186, supra note 14
41 Supra note 14, at Article 186
42 ibid, at Article 187
43 ibid, at Article 187(c)(ii)
44 Supra note 14, constituted in accordance with Annex VII, which sets out the parameters of the Arbitration
45 ibid, see Annex VIII
46 ibid, in this case Hamburg, Annex VI, Article 1(2)
47 ibid, at Article 188(2)(c)
48 ibid, at Article 289
by a court/tribunal under this section would be final and the parties would be bound to comply with the resolution\textsuperscript{49}.

A broad structure such as this, which could be refined to specifically address the key challenges and areas that AI regulation could present, would seem to provide a strong basis for translating Scherer’s ideas on judicial adjudication of tort claims to the international level. In his proposal, a key aspect of the courts involvement is their fact-finding nature and their ability to look at the chain of events and apportion blame. The Annex in the UNCLOS which outlines the details on arbitration, states that the parties to the dispute shall facilitate the tribunal’s work, providing it with all relevant documents, facilities and information\textsuperscript{50} and enabling it to call all necessary witnesses or experts and receive their evidence\textsuperscript{51}. This would allow the arbitral panel to adequately fulfil the role envisaged for the courts. When one considers that the tribunal can also make a binding resolution apportioning both blame and damages, then it would seem to be able to act as an adequate substitute for the judiciary.

An added benefit of this approach is the fact that disputes can be sent to the International Court of Justice ['ICJ']. The potential for AI to be involved in criminal activity is high, and already debates around the use of Lethal Autonomous Robots ('LARs') are prevalent in the public sphere. Having the option of recourse to the ICJ would help to address the myriad of potential usages and mis-uses of AI that could occur.

4. CONCLUSION

One of the difficulties in offering a regulatory solution to the problems AI poses is that we have never had to regulate anything akin to AI before. It may be that a novel technology requires a novel approach to regulation. What this might look like would no doubt be subject to much debate. If however, one were to stick with more traditional regulatory measures, there are a lot of aspects of Scherer’s proposals which \textit{prima facie} appear well suited to transposition to the international level. The UN model seems the most appropriate one to follow, with a treaty signed by State Parties ensuring that the regulations would be brought into domestic law. It could also bring into existence and set the parameters of a specialised agency, with seemingly no reason why the agency could not take on the form envisaged by Scherer. The major problems with translating Scherer’s ideas to the international level are found in the third strand of his proposal, namely the adjudication of tort claims and the provision of individual remedies. The major difficulty here is that international law generally applies only between States, or between a State and their citizens, not between companies and individuals. Whilst the UNHRC proposals for introducing a legally binding instrument to regulate the activities of companies may prove to be a game-changer in this respect, they are still merely proposals, and ones which many people consider to be unworkable. For now, taking things as they stand, it must be said that Scherer’s proposals are ultimately unworkable at the international level, due to the nature of the application of international treaties. If this however were to change in the future, it would seem that his ideas are well placed to offer a solution to the problem of how to regulate AI on a worldwide basis.

\textsuperscript{49} ibid, at Article 296
\textsuperscript{50} ibid, at Annex VII, Article 6(a)
\textsuperscript{51} ibid, at Annex VII, Article 6(b)
BIBLIOGRAPHY

Treaties


Articles


Other


United Nations General Assembly (UNGA)
Elaboration of an international legally binding instrument on transnational corporations and other business enterprises with respect to human rights (2014) UN Doc A/HRC/RES/26/9