

Regulating Artificial Intelligence Proposal for a Global Solution

Olivia J. Erdélyi

School of Law
University of Canterbury
Christchurch 8140, New Zealand
olivia.erdelyi@canterbury.ac.nz

Judy Goldsmith

Department of Computer Science
University of Kentucky
Lexington, KY 40506, USA
goldsmi@cs.uky.edu

Abstract

Given the ubiquity of artificial intelligence (AI) in modern societies, it is clear that individuals, corporations, and countries will be grappling with the legal and ethical issues of its use. As global problems require global solutions, we propose the establishment of an international AI regulatory agency that — drawing on interdisciplinary expertise — could create a unified framework for the regulation of AI technologies and inform the development of AI policies around the world. We urge that such an organization be developed with all deliberate haste, as issues such as cryptocurrencies, personalized political ad hacking, autonomous vehicles and autonomous weaponized agents, are already a reality, affecting international trade, politics, and war.

Introduction

AI is increasingly affecting our lives. Self-driving cars are being released on the roads; we rely on medical diagnosis tools to catch rare diagnoses; product recommendation systems use pattern recognition softwares to analyze our needs and optimize our shopping experience; we use automated surveillance techniques, killer bots, and other weaponized AI technologies to defend our countries; powerful data mining applications allow us to sift through a wealth of information within the shortest period of time; and AI-enabled decision-making systems using predictive analytics are employed in financial services to detect fraud, tax evasion, or money laundering, and in the justice system for predictive policing and sentencing. Undoubtedly, some of these technologies can make life a lot easier, but they also present substantial problems. Sometimes these problems result from the imperfection of AI applications, as when AI systems produce discriminatory biases. Sometimes, issues arise because AI is doing its job far too perfectly, as evidenced by the increasing privacy threat posed by pattern recognition applications (Kosinski and Wang 2017). Some instantiations of AI are ethically questionable (e.g., child-like sex bots (Strikwerda 2017)), potentially dangerous (e.g., autonomous kill-decisions by machines), or raise broader systemic challenges (e.g., labor displacement through AI, impugment of existing ethical, legal, and social paradigms). For a good

overview of the contemporary AI landscape and policy environment around AI, see (Calo 2017).

Do these applications really make human society more efficient, better, or safer? Or is AI rather a looming menace that will ultimately destroy mankind (Kohli 2015)? A lot depends on how we handle the historic opportunity to shape these potent, emerging technologies — a question acquiring increasing practical importance in view of the rapidly growing number of machine-human interactions thanks to recent breakthroughs in AI (in particular in machine learning).

The AI community has long been calling for policy action with respect to AI and criticism is getting louder on the growing legal vacuum in virtually every domain affected by technological advancement (Wadhwa 2014; Breland 2017; Walsh 2017). Policymakers around the world are beginning to address AI policy challenges. Canada, China, Japan, the UK, the US, and the EU have launched ambitious strategies to promote the development and commercialization of AI with a view to maintain sustained economic competitiveness after the inevitable global transition to an AI-driven economy (New 2017; Viola 2017). There are also many academic, and joint public and private sector venues that support governments in promoting AI R&D — examples include the International Association for Artificial Intelligence and Law, the Partnership on AI, the Artificial Intelligence Forum of New Zealand, and SPARC in the EU, not to mention a small number of tech firms' (Amazon, Apple, Baidu, Google, Facebook, IBM, and Microsoft) uncontested hegemony over the AI scene. Courts, faced with first AI-related disputes, contribute to clarifying situations, although some decisions reflect a complete lack of technological expertise (CJEU ; LGH), adding to the confusion and underscoring the need for interdisciplinary cooperation and improving policymakers' expertise on AI (Calo 2017).

In short, both the AI revolution and the challenges it presents to society are very real and it is policymakers' turn to do something about it. However, caution is advised with purely national approaches: Once the emergence of AI technologies is constructed as a problem and the necessity of regulation identified, a number of peculiarities intrinsic to lawmaking should be considered so the newly created legal norms become truly authoritative, that is, accepted as legitimate and institutionalized, as opposed to merely formal or symbolic rules without any impact on normative orientations

and behavior (Hurd 2008).

First, whenever the regulation of an issue has externalities that transcend national boundaries — as is the case with AI — differing domestic approaches tend to conflict, raising significant difficulties for those affected by more than one regime. Such problems are then typically perceived as transnational in scope, with the consequence that actors increasingly deem national rules inapt to provide suitable solutions. This discrepancy between the transnational nature of a problem and the national character of the law governing it creates pressures for transnational regulation. Second, transnational legal ordering is characterized by a set of complex, recursive, multi-directional processes, which follow their own logic and crucially affect norms' authority (Halliday and Shaffer 2015). Third, even though the legitimacy of legal norms has predominantly social rather than moral roots, ethical considerations should and will play a pivotal role in shaping the nascent body of law — whether domestic or transnational — concerned with AI.

We therefore hold the view that national efforts to develop AI policies, should, from the very beginning, be coordinated and supported by an international regulatory framework to avoid the risks stemming from the imperfect interaction of fragmented domestic regulatory approaches. Against this background, we propose the establishment of a new intergovernmental organization — which could be named International Artificial Intelligence Organization (IAIO) — to serve as an international forum for discussion and engage in standard setting activities. The IAIO should unite a diverse group of stakeholders from public sector, industry, and academic organizations, whose interdisciplinary expertise can support policymakers in the overwhelming and crucially important task of regulating this novel, immensely complex, and largely uncharted area. We hope that such a wide-scale, in-depth cooperation among all interested stakeholders at this early stage will put national and international policymakers in the position to take proactive action instead of lagging behind technological innovation with potentially devastating implications. Recent turbulences in worldwide financial markets can serve as warning examples of the ramifications of regulators not being ready for what the future might bring. The stakes with AI are even higher. We must get it right the first time.

The paper will proceed as follows: first, we present a brief analysis of transnational normmaking processes, followed by our proposal on an international AI regulatory framework and a short conclusion.

Dynamics of Transnational Lawmaking

In response to economic and cultural globalization, legal, political science, and sociology scholarship have taken manifold attempts to capture processes of various forms of transnational social ordering. Examples include the traditional, purely state-centric legal notion of *international law* with a dichotomous view towards national and international law; *global law*, which refers to legal norms of universal scope while also acknowledging the role of non-state actors in normmaking; *transnational law*, which can have several connotations in reference to norms with a more than

national but less than global purview; the concept *regime theory* developed by international relations scholars, which is likewise state-centric and has a sole focus on international political processes without any regard to the impact of domestic politics or law's normativity; the sociological *world polity theory*, studying the diffusion of legal norms assuming that global conceptual models frame national societies in one dimensional top-down processes; and *transnational or global legal pluralism*, which emphasizes the coexistence of different legal orders and normative contestations among them. Giving a comprehensive overview (including references) of the respective merits and limitations of existing theories, (Shaffer 2010) and (Halliday and Shaffer 2015) introduce a further, socio-legal notion termed *transnational legal order (TLO)*, which builds on these approaches and is defined as a social order of transnational scope consisting of "a collection of formalized legal norms and associated organizations and actors that authoritatively order the understanding and practice of law across national jurisdictions." We explain the determinants influencing transnational legal processes through which legal norms are constructed, conveyed, and institutionalized based on the concept of TLO, owing to its ability to highlight both the legal and institutional aspects justifying the proposed regulatory framework.

Disaggregating the above definition into two parts — (1) formalized legal norms produced solely or partially by some type of transnational legal organization or network, which are (2) aimed at inducing changes within nation-states — we will first give account of the bewildering variety of governance arrangements characterizing modern international relations, and then illuminate the complex ways in which legal norms interact and institutionalize. The terms *international* and *transnational* shall be used interchangeably, referring to norms and institutions spanning national boundaries (whether global or geographically more restricted in scope).

Turning to the first element of our TLO definition, both the norms and the institutions issuing them come in a diverse array of forms. Norms are contained in various formal texts of softer or harder legal character such as treaties, conventions, codes, model laws, standards, administrative rules and guidelines, and judicial judgments. International institutionalization displays a similar diversity featuring public intergovernmental (also called international) organizations (IGOs or IOs) and private non-governmental organizations (NGOs) of varying levels of formalization (Klabbers 2015; Schermers and Blokker 2011; Vabulas and Snidal 2013). The widespread use of both hard and soft legalization in international governance begs the questions of what *hard* and *soft law* are and what drives actors' choices between disparate legal and institutional settings.

Note that the existing literature is divided on what constitutes hard and soft law. Some authors concentrate on legal rules' binding quality either in binary terms or along a continuum between fully binding legal instruments and purely political, non-binding arrangements, while others focus on their ability to impact behavior. For a good overview see (Guzman and Meyer 2010; Shaffer and Pollack 2010; Abbott and Snidal 2000). Because it includes both the legal norms and the institutional arrangements responsible for

their development within the scope of its analysis, the most interesting definition for our purposes is the one adopted by Abbott and Snidal. They distinguish hard from soft law along three dimensions, namely (1) the extent of rules' precision, (2) the degree of legal obligation they establish, and (3) whether or not they delegate authority to a third-party decision-maker for interpreting and implementing the law. Hard law refers to legally binding obligations that are either precise or can be made such by adjudication or further clarifying regulation, and that empower a third party to oversee their interpretation and enforcement. Soft law, on the other hand, embodies legal instruments that exhibit some degree of softness along any of these three dimensions.

(Guzman and Meyer 2010) and (Abbott and Snidal 2000) provide a very instructive comparison of the relative advantages and disadvantages of hard and soft legalization and the various factors that determine actors' preferences towards different types of international governance arrangements.

Hard legalization is typically characterized by a coherent, established, and formalized institutional and procedural framework to ensure smooth implementation, elaboration, and enforcement of commitments. These arrangements are generally perceived as legitimate, resulting in a concomitant enhanced compliance-pull, and backed up by international law that provides international actors readily available mechanisms (e.g., for recognition or enforcement) to order their relations. The combination of these factors enhances the credibility of commitments by constraining opportunistic behavior and increasing the costs of renegeing; reduces post-contracting transaction costs by restricting/constraining attempts to alter the status quo by way of frequent renegotiation, persuasion, or coercive behavior; allows parties to pursue political strategies through legal rather than political channels at low political cost; and solves problems of incomplete contracting by vesting an administrative or judicial institution with power for interpreting and clarifying rules intentionally left imprecise in anticipation of unforeseeable future contingencies. Yet, hard legalization comes at certain costs: it restricts actors' behavioral freedom, entails potentially severe sovereignty implications, and is less effective in accommodating diversity or adapting to changing circumstances by reason of its relative rigidity.

Thus, in many instances, softer forms of legalization, which offer some of hard law's perks yet alleviate its intrinsic disadvantages through their flexible, more or less informal cooperation mechanisms, may better serve parties' purposes. By relaxing the level of formality along one or more of the dimensions precision, obligation, and delegation, soft legalization minimizes initial contracting costs and facilitates speedy conclusion of agreements. Bargaining problems become less pronounced, negotiation and drafting requires less scrutiny, and there is no need for potentially challenging approval and ratification processes. Thanks to soft legal commitments' malleable cooperation frameworks, parties retain more control over the overall design and organization of their cooperation, incur lower sovereignty costs, and have an easily adjustable system at their disposal to deal with change and uncertainty. Soft law also has a way of evening out power asymmetries by securing and perpetuating power-

ful actors' interests at lower sovereignty costs, while at the same time shielding the weak from their pressure. Furthermore, soft law is the only directly available instrument to non-state actors for ordering their interactions. Due to their conciliatory properties, softer forms of legalization leave actors time to acquire sufficient information and expertise to gradually test and develop solutions to problems, encouraging collective learning processes and ever deeper cooperation between them — benefits that plentifully compensate soft law's central weakness: diminished compliance pull.

In conclusion, the choice between harder and softer types of legalization involves a context-dependent tradeoff, which actors should carefully consider on a case-by-case basis. (Vabulas and Snidal 2013) describe the pros and cons of institutional formalization and the tradeoffs actors face when moving along a broad spectrum of intergovernmental organizational formality — especially between formal and informal intergovernmental institutions (FIGOs and IIGOs) — in an analogous fashion.

These three analyses show that, in general, actors opt for hard law/higher institutional formality when they (1) wish to enter into a binding commitment in issue areas subject to a high degree of consensus, because violations are hard to detect, or parties wish to signalize their intention to engage in sincere cooperation; (2) are willing to accept sovereignty costs stemming from delegating decision-making authority to a central body in order to establish stronger collective oversight over issue areas where the probability of violations is high and monitoring and enforcement is important; (3) put more value on collective control of information, for instance, to unveil violations and increase peer pressure to induce universal compliance; (4) aim for lower long-term transaction costs to effectively tackle recurring or clear-cut issues in standard operating procedures; (5) intend to set up a sophisticated centralized administration to provide legitimacy and stability for supporting complex work processes such as the design and elaboration of norms, coordination involving multiple parties, or judiciary and/or enforcement procedures; (6) are faced with the task of managing routine problems, which is more easily done with established administrative and implementing systems.

Conversely, soft law/lower institutional formality is preferable when actors (1) want to maintain flexibility to deal with uncertainty, distribution problems, diversity, and changing circumstances; (2) prefer to preserve state autonomy and avoid sovereignty intrusions because welfare gains of cooperation outweigh the potential for defection and opportunism so that agreements are self-enforcing once any focal point for discussions has been established, or when external effects elicited by domestic actions are negligible; (3) insist on avoiding formal transparency mechanisms to maintain closer control of information typically among a more homogeneous group; (4) need lower initial contracting costs to speed up negotiations to be able to act fast (e.g., in crisis situations) or because hard law is not available for lack of consensus; (5) find that minimalistic administrative functions are sufficient to support their purposes; (6) must manage high uncertainty (e.g., in initial stages of cooperation or in new/complex issue area) and want to allow

Hard Law/High Institutional Formality	Soft Law/Low Institutional Formality
binding commitment	flexible cooperation arrangements
delegation/high sovereignty costs	state autonomy/low sovereignty costs
collective control of information	close control of information
low long-term transaction costs	low initial contracting costs
complex centralized administration	minimalistic administrative functions
routine management	crisis/uncertainty management

Table 1: Tradeoffs in legalization/institutional formality.

themselves time for coordination and establishing common ground without making strong commitments.

Sometimes soft law eventually paves the way towards harder forms of legalization and cooperation becomes increasingly formalized, but in many cases soft legalization and institutional informality have their own justification. In practice, both highly institutionalized FIGOs, such as the United Nations (UN) or World Trade Organization (WTO), IIGOs allowing for laxer cooperation, like the Basel Committee on Banking Supervision (BCBS), private NGOs, for instance the International Chamber of Commerce (ICC), and hybrid forms can be fairly successful and instrumental actors in international lawmaking.

Table 1 gives an overview of the above outlined six trade-offs actors have to weigh when choosing their desired level of legalization/institutional formalization.

Moving on to the second part of our TLO definition, transnational legal norms directly or indirectly pursue the ultimate goal to induce shifts in countries' policies and individuals' normative preferences through various formal or informal channels. This generates convoluted, recursive cycles of international lawmaking processes across diverse transnational and national fora, until norms eventually settle and institutionalize (Shaffer 2010). (Halliday and Shaffer 2015) note that transnational normmaking may encounter difficulties in the following situations: First, actors may find themselves caught up in diagnostic struggles over the framing of problems, which favors particular alliances and antagonisms, supporting diagnoses reflecting the respective interests of these groupings. Second, domestic implementation of transnationally agreed rules is frequently thwarted and a new cycle of lawmaking is triggered by parties who are influential at the national level, but not represented or unsuccessful in international negotiations and therefore refuse to accept such norms as legitimate — a situation referred to as *actor mismatch*. Third, in their endeavor to reach widely accepted compromises, parties often resort to vague language or leave delicate issues unresolved in their agreements. The resulting ambiguity and built-in contradictions of transnational norms open avenues for nationally fragmented, likely conflicting implementation, again calling for further transnational lawmaking to eliminate related problems.

Inspired by (Shaffer 2010), we now describe the recursive processes of international lawmaking, which encompass mutual interactions both vertically between transnational and domestic venues, as well as horizontally among various TLOs. Vertically, transnational norms impact states. This impact can encompass the whole or parts of the state

(*location of change*), it may occur in a slow, progressive process or abruptly owing to unexpected circumstances (*timing of change*), and across five interrelated dimensions. The most obvious aspect of state change is the dynamic evolution of domestic legal systems elicited by the formal national enactment of transnational law. Formal enactment may or may not have a substantial effect on rules' practical implementation depending on the extent to which the transnationally induced change is viewed as legitimate. In more subtle ways, however, these primary legal changes set much broader systemic transformations in motion with potentially heavy social repercussions. For one thing, they continuously reshape established governance models by altering the allocation of functions between the state, the market, and other forms of social ordering, at times prompting more state intervention giving birth to new public and public-private hybrid agencies, while other times propelling deregulatory tendencies resulting in a retreat of state administration and simultaneously engaging self-regulation by the private sector. Moreover, transnational legal processes are often responsible for revamping states' institutional architecture, shifting power between different branches of government and upsetting the division of responsibilities among existing state institutions, sometimes giving rise to new additions to the institutional landscape. It is not hard to see that domestic systems may starkly differ, and such fragmentation often entails devastating consequences in issue areas with cross-border effects. These legal, governance, and institutional changes directly affect individuals by reconfiguring markets for professional expertise, which, in turn, feeds back into the adaptation of governance models by, e.g., a move towards more technocratic forms of governance. This highlights an important, yet admittedly somewhat elusive point, namely that not only institutions but also individuals — acting as conduits facilitating the diffusion of transnational norms — play a crucial role in domestic and transnational lawmaking. The fifth domain of state change concerns the modifications in patterns of association and mechanisms of accountability across various national and international sites of governance, which ultimately shape individuals' legal culture and consciousness, as well as their expectations towards the state, triggering new processes of state change where these views conflict with the prevailing state of affairs.

The extent, location, and timing of state change hinges on three clusters of factors pertaining to the TLO's nature, its relation to the receiving state, and the receiving state's peculiarities. First, TLOs are generally better received if perceived legitimate, i.e., norms are adopted by respected actors with preferably similar interests, in a fair (especially non-coercive) process, and they effectively tackle designated target problems. Myriads of international and national, state and non-state actors interact in complementary or conflicting ways in shaping every aspect of transnational lawmaking, seeking to legitimize rules that serve their purposes and delegitimize those that run against them, and powerful players typically dictate the outcome of such struggles. TLOs are more likely to have real behavioral impact if they consist of accepted, clear, and well-understood norms. As discussed above, binding hard law does not necessarily score better

in this respect. In a large part, TLOs' coherence is a function of the quality and quantity of their horizontal interaction, and can be threatened where significantly overlapping TLOs interact in an antagonistic rather than complementary fashion (Shaffer and Pollack 2010).

As far as TLOs' relation to the receiving state is concerned, powerful actors sometimes resort to coercive measures to impose their will on weaker countries. However, because coercion irrevocably destroys norms' legitimacy, changes forced on states in this manner are at best symbolic and short-lived before they are successfully blocked at the stage of domestic implementation. Another essential prerequisite for the sustainability of transnationally triggered change is the support of intermediaries, who link transnational and national lawmaking processes and are deeply familiar with the interests of both sides. Whether government representatives, industry specialists, academics, social movement leaders, or professionals employed with various public or private organizations on the national and international platform, these intermediaries are instrumental in coordinating communication, easing tensions, and conveying norms between the national and transnational levels.

Finally, the single most important condition for transnational legal norms' national acceptance is their conformity with the target country's existing cultural and institutional settings and pursued reform initiatives. This strongly depends on the receiving country's prevailing power configurations, institutional capacities, path dependencies, and cultural disposition, and tends to decrease as the distance between the transnational and national contexts and interests and/or the extent of state change increases.

This concludes our analysis of transnational legal ordering, highlighting the main factors instrumental in determining transnational legal norms' efficiency in influencing the behavior of states and their various constituencies. We now turn to our proposal on a consistent international AI regulatory framework.

Proposal for a New International Artificial Intelligence Organization

International institutions are the prevalent vehicles of international cooperation in our interconnected world. When a critical mass of states and/or non-state actors feel that transnational cooperation is necessary to solve a problem that is impossible to tackle by isolated national measures, they establish a new IGO or NGO for that particular purpose. Based on legal and international relations definitions in circulation — see (Klabbers 2015; Schermers and Blokker 2011; Vabulas and Snidal 2013) — we define an IGO as a formal entity (1) established by an international agreement governed by international law; (2) with at least three (sometimes two) members — typically states but increasingly also IGOs; and (3) having at least one organ with a will distinct from that of its members. While FIGOs' organizational purpose is laid down in a binding international agreement such as a treaty or a formal legal act of another IGO, their membership is clearly defined in the founding legal act, and they have a permanent and significant institutionalization in

place, IIGOs operate based on an explicitly shared, but informal expectation about purpose, their membership is not always clear, as members are explicitly associated but only by non-legal mutual acknowledgment, and they do not possess any significant institutionalization. Conversely to IGOs, NGOs are not created by treaty and are therefore governed by national rather than international law, and their membership is made up of non-state actors.

Given the severity and global nature of AI's anticipated impact on humanity, we expect it to join the long line of issue areas requiring interstate cooperation, raising the question of establishing an IGO at some point in the future. Against this background, we propose the creation of the IAIO as a new IGO, which could initially serve as a focal point of policy debates on AI-related matters and — given sufficient international support — acquire increasing role in their regulation over time. We start by determining the degree of desired institutional formalization by examining, in turn, the six above elaborated tradeoffs in relation to AI.

Binding commitment vs. flexible cooperation arrangements: As pointed out earlier, AI will fundamentally transform human society worldwide. Since this process of transformation is likely to be inescapable for any single state, states will probably wish to cooperate sincerely. Also, violations will be difficult to detect as keeping pace with technological innovation will require considerable technical expertise and capacities, presumably exceeding especially weaker countries' capabilities and evoking severe power asymmetries. While, apart from this latter circumstance, these factors speak for hard legal commitments, it must be kept in mind that AI research and AI-human interactions are relatively young phenomena, meaning that we are not even able to grasp the spectrum and extent of the impending changes, let alone the dimension of the problems they will raise. Many AI instantiations encroach on our most basic rights, pose an existential threat, or bring up profound ethical questions, not to mention that they will utterly and completely upset our legal system. So, we are looking at heated debates among radically diverse parties over a variety of uncertain issues, which may change in rapid and currently unimaginable ways — conditions that, based on past experience, do not exactly favor international consensus. Therefore, we need all the flexibility we can get to acquire familiarity with the issues at hand, sort out differences, and establish common ground, before we can contemplate drawing up a more binding framework for cooperation.

Delegation/high sovereignty costs vs. state autonomy/low sovereignty costs: Weaponized AI technologies and certain data mining practices are clearly relevant for national security. As this is a sensitive issue area involving particularly high sovereignty costs, at least initially, states will show reluctance to give up and delegate decision-making authority to the IAIO. On the long run, however, powerful collective oversight and enforcement mechanisms will probably be indispensable in order to curb incentives for violations and opportunistic behavior, which should otherwise be high in light of the major shifts in international power constellations triggered by changes in countries' competitive positions. Also, domestic AI policies will produce significant externalities,

affecting other countries. Based on this analysis, it is hard to escape the conclusion that a highly institutionalized organization with binding legislative, dispute resolution, and enforcement authority would be better suited as new international AI regulator. Nevertheless, the political reality remains that until sufficient clarity is reached on the IAIO's precise purpose, membership, the issues to regulate, and the broad directions to follow, international consensus supporting such a high degree of institutionalization is off the table.

Collective control of information vs. close control of information: History shows that states are generally cautious about sharing information on fate-changing technologies, which speaks for close control of information with respect to AI. However, if and when we manage to gather consensus for hard legal commitments (e.g., treaty on certain AI applications), we will probably need to be more forthcoming with certain information to ensure compliance with those instruments. This is again a strong argument in favor of starting cooperation on AI regulation in a softer institutional framework and using soft law instruments, although a move towards harder legalization seems to be desirable over time.

Low long-term transaction costs vs. low initial contracting costs: Given that international discussions on AI are just beginning and powerful states will likely have divergent preferences with respect to the regulation of this high-impact field, and also owing to the difficulties discussed in the context of previous tradeoffs, the possibility of reaching a workable international consensus is rather remote on the short term. Yet crucially, swift regulatory response is imperative to prevent proliferating unregulated AI applications from causing social harm and to ensure that the opportunity presented by the rise of AI is harvested to humanity's benefit rather than detriment. This all speaks for lowering initial contracting costs with soft legalization and low institutional formalization, whereas the idea of setting up a more robust regulatory framework with standard operating procedures has merit only at a later stage, in possession of sufficient expertise and political consensus to better assess the implications of various policy options and formulate informed policy recommendations.

Complex centralized administration vs. minimalistic administrative functions: Similar considerations apply as far as the level of administrative sophistication of the IAIO is concerned. In the initial stage of determining the purpose of the organization, its membership, the issues that need to be regulated, and the backbone of its regulatory agenda, less is probably more. Later, with perhaps binding legal instruments governing selected aspects of AI for a wide membership, work will get more complex, requiring stronger oversight, dispute resolution, and enforcement mechanisms as well as more powerful bureaucratic functions to service them.

Routine management vs. crisis/uncertainty management: In view of AI's novelty, extreme complexity, unforeseeable evolution, and the controversies it is expected to elicit among a very heterogeneous circle of members, we are up against managing an extraordinarily uncertain issue area. Consequently, we need time and soft legalization's flexibility to establish commonly shared ideas, interests, cooperation mechanisms, and solutions, which can then form the basis of more

formalized cooperation arrangements in the future.

In summary, at least initially, the IAIO should start out as an IIGO displaying a relatively low level of institutional formality and using soft law instruments, such as non-binding recommendations, guidelines, and standards, to support national policymakers in the conception and design of AI-related regulatory policies. Its interim goal should be to galvanize international cooperation in this domain as early as possible, before states develop their own, diverging policies, which may be hard to rescind without political damage. Whether the international community wishes to move towards more formalized cooperation at some point in the future, remains to be seen. Sometimes informality turns out to be the key to an organization's success (see the Bank of International Settlements especially during its initial years of operation and World War II, or the BCBS and the different G-Groups at present (Borio and Toniolo 2006)), initially informal arrangements may turn into formal frameworks of cooperation (as with the General Agreement on Tariffs and Trade's (GATT) gradual transformation into the WTO (Abbott and Snidal 2000)), and there are examples for remarkably successful, sustained, complementary, and mutually beneficial cooperation between several organizations of varying institutional formality in the same issue area (as between the IMF and various G-Groups in financial regulation or the Australia Group (AG), an IIGO, and the Organization for the Prohibition of Chemical Weapons (OPCW), a FIGO, in the regulation of chemical and biological weapons (Vabulas and Snidal 2013)).

Our excursus in the domain of international lawmaking shows that beyond the optimal level of legalization and institutional formality, the IAIO must fulfill a number of more subtle requirements to be perceived as a *fair and legitimate* regulator. While leaving the elaboration of details to future research, we would like to stress the importance of including an interdisciplinary mix of experts (with, e.g., AI, legal, political, and ethics background) in the initial deliberations related to the IAIO's establishment, *modus operandi*, and regulatory agenda, as well as conducting regular, large-scale consultation processes with a diverse spectrum of interested stakeholders from public sector, industry, and academia, to ensure due consideration of all relevant perspectives.

Conclusion

Given the intensifying worldwide activism in AI regulation and AI's anticipated substantial and global impact on human society, we propose a consistent international regulatory framework — with a new IGO, the IAIO, as its focal point — to streamline and coordinate national policymaking efforts. Learning from past experience in other regulatory fields, our objective is to offer a viable framework for international regulatory cooperation in the issue area of AI to avoid the development of nationally fragmented AI policies, which may lead to international tensions. Should our proposal find sufficient support in the international community, more concrete steps towards setting up the here advocated regulatory framework, and regulatory policies on specific AI issues can be elaborated.

Acknowledgments

Part of the work was done while the first author was affiliated with the School of Economic Disciplines at University of Siegen. The second author was partially supported by NSF grant 1646887.

References

- Abbott, K. W., and Snidal, D. 2000. Hard and soft law in international governance. *International Organization* 54(3):421–456.
- Borio, C., and Toniolo, G. 2006. One hundred and thirty years of central bank cooperation: a bis perspective. BIS Working Papers, No. 197.
- Breland, A. 2017. Elon Musk: We need to regulate ai before 'it's too late'. The Hill, July 17, 2017, <http://thehill.com/policy/technology/342345-elon-musk-we-need-to-regulate-ai-before-its-too-late>.
- Calo, R. 2017. Artificial intelligence policy: A primer and roadmap. August 8, 2017, Available at SSRN: <https://ssrn.com/abstract=3015350>.
- CJEU. Court of Justice of the European Union, Judgment of 8 September 2016, GS Media v Sanoma Media Netherlands and Others, Case C-160/15, ECLI:EU:C:2016:644.
- Guzman, A. T., and Meyer, T. L. 2010. International soft law. *Journal of Legal Analysis* 2(1):171–225.
- Halliday, T., and Shaffer, G. C. 2015. Transnational legal orders. In Halliday, T., and Shaffer, G. C., eds., *Transnational Legal Orders*. Cambridge University Press. 3–72.
- Hurd, I. 2008. Theories and tests of international authority. In Cronin, B., and Hurd, I., eds., *The UN Security Council and the Politics of International Authority*. Routledge. 23–39.
- Klabbers, J. 2015. *An Introduction to International Organizations Law*. Cambridge University Press.
- Kohli, S. 2015. Bill Gates joins Elon Musk and Stephen Hawking in saying artificial intelligence is scary. QUARTZ, January 29, 2015, <https://qz.com/335768/bill-gates-joins-elon-musk-and-stephen-hawking-in-saying-artificial-intelligence-is-scary/>.
- Kosinski, M., and Wang, Y. 2017. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. PsyArXiv.
- LGH. Landgericht Hamburg, Az. 310 O 402/16.
- New, J. 2017. How governments are preparing for artificial intelligence. Center for Data Innovation, August 18, 2017, <https://www.datainnovation.org/2017/08/how-governments-are-preparing-for-artificial-intelligence/>.
- Schermers, H. G., and Blokker, N. M. 2011. *International Institutional Law: Unity Within Diversity*. Martinus Nijhoff Publishers.
- Shaffer, G. C., and Pollack, M. A. 2010. Hard vs. soft law: Alternatives, complements and antagonists in international governance. *Minnesota Law Review* 94(3):706–799.
- Shaffer, G. C. 2010. Transnational legal process and state change: Opportunities and constraints. *Minnesota Legal Studies Research Paper* (10-28):1–53.
- Strikwerda, L. 2017. Legal and moral implications of child sex robots. In Danaher, J., and McArthur, N., eds., *Robot Sex*. MIT Press.
- Vabulas, F., and Snidal, D. 2013. Organization without delegation: Informal intergovernmental organizations (iigos) and the spectrum of intergovernmental arrangements. *The Review of International Organizations* 8(2):193–220.
- Viola, R. 2017. The future of robotics and artificial intelligence in europe. European Commission, Digital Single Market Blog Post, February 16, 2017, <https://ec.europa.eu/digital-single-market/en/blog/future-robotics-and-artificial-intelligence-europe>.
- Wadhwa, V. 2014. Laws and ethics can't keep pace with technology. *Massachusetts Institute of Technology: Technology Review* 15.
- Walsh, T. 2017. EU parliament: Consultation on robotics and artificial intelligence - summary and contributions. <http://thefutureofai.blogspot.de/2017/10/eu-parliament-consultation-on-robotics.html>.