Gary Marchant
Center for Law, Science & Innovation, Arizona State University

Introduction


In this and other recent pronouncements, Musk is calling for artificial intelligence (AI) to be regulated by traditional regulation, just as we regulate foods, drugs, aircraft and cars. Putting aside the quibble that food, drugs, aircraft and cars are each regulated very differently, these calls for regulation seem to envision one or more federal regulatory agencies adopting binding regulations to ensure the safety of AI. Musk is not alone in calling for “regulation” of AI, and some serious AI scholars and policymakers have likewise called for regulation of AI using traditional governmental regulatory approaches.

But these calls for regulation raise the questions of what aspects of AI should be regulated, how they should be regulated, and by who? The reality is that at best there will be some sporadic piecemeal traditional regulation of AI over the next few years, notwithstanding the increasing deployment and application of AI in a growing range of applications and industry sectors. In the interim at least, this “governance gap” for AI will mostly be filled by so-called “soft law” (see Part I, supra). These “soft law” mechanisms include various types of instruments that set forth substantive expectations but are not directly enforceable by government, and include
approaches such as professional guidelines, private standards, codes of conduct, and best practices. A number of such soft law approaches have already been proposed or are being implemented for AI (see Part II, supra). While soft law has some serious deficiencies, such as lack of enforceability, there are additional strategies that can help maximize the effectiveness of this second-best approach to governance (see Part III, supra). For example, the lack of enforceability problem can be solved at least in part by various types of indirect enforcement by entities such as insurance companies, journal publishers, grant funders, and even governmental enforcement programs against unfair or deceptive business practices. Another problem, the lack of coordination between a potentially large number of overlapping and perhaps even inconsistent soft law programs, is to create what has been described as a Governance Coordinating Committee to help serve a coordinating function.

The Unsuitability of Traditional Regulation for AI

While some piecemeal regulation of specific AI applications and risks using traditional regulatory approaches may be feasible and even called for, AI has many of the characteristics of other emerging technologies that make them refractory to comprehensive regulatory solutions. For example, AI involves applications that cross multiple industries, government agency jurisdictions, and stakeholder groups, making a coordinated regulatory response difficulty. In addition, AI raises a wide range of issues and concerns that go beyond traditional regulatory agency focus on health, safety and environmental risks. Indeed, many risks created by AI are not within any existing regulatory agency’s jurisdiction, including concerns such as technological unemployment, human-machine relationships, biased algorithms, and existential risks from future super-intelligence.
Moreover, the pace of development of AI far exceeds the capability of any traditional regulatory system to keep up, a challenge known as the “pacing problem” that affects many emerging technologies. The risks, benefits and trajectories of AI are all highly uncertain, again making traditional preemptory regulatory decision-making difficult. And finally, national governments are reluctant to impede innovation in an emerging technology by preemptory regulation in an era of intense international competition.

For these reasons, it is safe to say there will be no comprehensive traditional regulation of AI for some time, except perhaps if some disaster occurs that triggers a drastic and no doubt poorly-matched regulatory response. Again, there may be slivers of the overall AI enterprise that are amenable to traditional regulatory responses, and these should certainly be pursued. But these isolated regulatory advances will be insufficient alone to deal with the safety, ethical, military, and existential risks posed by AI. Something more will be needed.

That something more that will be needed to fill the governance gap for AI will, at least in the short term, be within the category of “soft law.” Soft law are instruments that set substantive expectations that are not directly enforceable by government. They can include private standards, voluntary programs, professional guidelines, codes of conduct, best practices, principles, public-private partnerships and certification programs. Soft law can even include what Wendell Wallach and I refer to as “process soft law” approaches such as coding machine ethics into AI systems or creating oversight systems within a corporate Board of Directors. These types of measures are inherently imperfect, precisely because they are not directly enforceable.
This core weakness results in many other limitations, such that participation is incomplete, with the “good guys” complying and the “bad guys” not. These soft law measures are sometimes used as “whitewashing” (or “greenwashing”) to make it look like a problem is being addressed when it really is not. And soft law measures are often expressed in vague, general language that is hard to measure compliance with. Finally, soft law measures generally do not provide the same reassurance to the public as traditional government regulation that the problems presented by a new technology are being adequately managed. This public reassurance effect is an important secondary function of regulation.

Notwithstanding these significant limitations, soft law has become a necessary and inevitable component of the governance framework for virtually all emerging technologies, including AI. Traditional regulatory systems cannot cope with the rapid pace, diverse applications, heterogeneous risks and concerns, and inherent uncertainties of emerging technologies. So although soft law measures are a second best solution, they are often the only game in town, at least initially. It recalls the quote attributed to Winston Churchill that “democracy is the worst form of government, except for all the others.”

Soft law has important advantages that explain its growing popularity and gap filling role. Soft law instruments can be adopted and revised relatively quickly, without having to go through the traditional bureaucratic rulemaking process of government. It is possible to experiment with several different soft law approaches simultaneously, indeed sometimes creating a problem of a proliferation of inconsistent private standards and other soft law instruments. They can sometimes create a cooperative rather than adversarial relationship among stakeholders. They are not bound by limited agency delegations of authority, and so can address any
and all concerns raised by a technology. And because they are not adopted by a formal legal authority, they are not restricted to a specific legal jurisdiction, but can have international application.

**Existing AI Soft Law Examples**

We are already seeing the rapid infusion of soft law initiatives and proposals into the AI governance space. Indeed, the likely first ever governance proposal for AI (at that time focused on robotics) was Isaac Asimov’s three laws of robotics first published in 1942. These “laws” were actually a form of soft law as they had no formal legal authority. More recently, an early entry into the AI soft law landscape was a “robot ethics charter” that the government of South Korea initiated in 2007, even though no final version of the ethics charter has ever been posted online.

**Institute of Electric and Electronic Engineers (IEEE)**

Perhaps the most comprehensive soft law initiative for AI was launched in 2016 by the IEEE, one of the world’s largest standard-setting and professional engineering societies. This initiative, entitled “The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, is intended to “ensure every stakeholder involved in the design and development of autonomous and intelligent systems is educated, trained, and empowered to prioritize ethical considerations so that these technologies are advanced for the benefit of humanity.” The Initiative has two intended outputs. The first is a guide known as Ethically Aligned Design, which has now been published as draft versions I and II for public comments. Version II is a document that exceeds 250 pages and that addresses over 120 policy, legal and ethical issues associated with AI, with recommendations assembled from more than
250 expert participants. It seeks to “advance a public discussion about how we can establish ethical and social implementations for intelligent and autonomous systems and technologies, aligning them to defined values and ethical principles that prioritize human well-being in a given cultural context, inspire the creation of Standards (IEEE P7000™ series and beyond) and associated certification programs, [and] facilitate the emergence of national and global policies that align with these principles.” The final version of Ethically Aligned Design is scheduled to be published in 2019.

The second and even more relevant activity by the Initiative is to produce a series of IEEE standards addressing governance and ethical aspects of AI. The IEEE has given official approval to create the following standards, with standard-setting committees now established to develop each standard:

- IEEE P7000™ – Model Process for Addressing Ethical Concerns During System Design
- IEEE P7001™ – Transparency of Autonomous Systems
- IEEE P7002™ – Data Privacy Process
- IEEE P7003™ – Algorithmic Bias Considerations
- IEEE P7004™ – Standard on Child and Student Data Governance
- IEEE P7005™ – Standard for Transparent Employer Data Governance
- IEEE P7006™ – Standard for Personal Data Artificial Intelligence (AI) Agent
- IEEE P7009™ – Standard for Fail-Safe Design of Autonomous and Semi-
Autonomous Systems
IEEE P7010™ – Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems
IEEE P7011™ – Standard for the Process of Identifying and Rating the Trustworthiness of News Sources
IEEE P7012™ – Standard for Machine Readable Personal Privacy Terms
IEEE P7013™ – Inclusion and Application Standards for Automated Facial Analysis Technology

These fourteen AI standards are scheduled to be finalized by the end of 2021, and will provide a broad set of governance requirements relating to the governance of AI. For example, the chair of the working group developing standard IEEE P7006 on personal AI agents has recently written that the standard is being developed to provide “a principled and ethical basis for the development of a personal AI agent that will enable trusted access to personal data and increased human agency, as well as to articulate how data, access and permission can be granted to government, commercial or other actors and allow for technical flexibility, transparency and informed consensus for individuals.”

Partnership on AI

Another significant “soft law” player in the AI field is the Partnership on AI. This Partnership was originally started by the big players in the AI space such as Google, Microsoft, Facebook, IBM, Apple and Amazon, but has expanded to include a wide variety of companies, think tanks, academic AI organizations, professional societies, and charitable groups such as the ACLU, Amnesty International, UNICEF and Human Rights Watch. One of the stated goals of the Partnership is to develop and share best practices for AI which includes: “Support research, discussions, identification,
sharing, and recommendation of best practices in the research, development, testing, and fielding of AI technologies. Address such areas as fairness and inclusivity, explanation and transparency, security and privacy, values and ethics, collaboration between people and AI systems, interoperability of systems, and of the trustworthiness, reliability, containment, safety, and robustness of the technology.”

The Partnership on AI has published a set of “Tenets” that include:

“We are committed to open research and dialogue on the ethical, social, economic, and legal implications of AI....

We believe that AI research and development efforts need to be actively engaged with and accountable to a broad range of stakeholders....

We will work to maximize the benefits and address the potential challenges of AI technologies, by: Working to protect the privacy and security of individuals....Working to ensure that AI research and engineering communities remain socially responsible, sensitive, and engaged directly with the potential influences of AI technologies on wider society....Ensuring that AI research and technology is robust, reliable, trustworthy, and operates within secure constraints....Opposing development and use of AI technologies that would violate international conventions or human rights, and promoting safeguards and technologies that do no harm.

We believe that it is important for the operation of AI systems to be understandable and interpretable by people, for purposes of explaining the technology.”

It remains to be seen if and how the Partnership will advance beyond these general tenets to produce more specific best practices and guidelines for responsible AI
The Future of Life Institute convened a meeting of many leading AI practitioners and experts at the Asilomar conference center in 2017, which is the home of the famous Asilomar Conference on Recombinant DNA held in 1975 which pioneered the soft law governance of technology by agreeing on a set of voluntary guidelines for genetic engineering research. At the 2017 Asilomar conference, the participants agreed on 23 principles to guide AI research and applications. These principles include “Failure Transparency” (“If an AI system causes harm, it should be possible to ascertain why.”); “Responsibility” (“Designers and builders of advanced AI systems are stakeholders in the moral implications of their use, misuse, and actions, with a responsibility and opportunity to shape those implications.”) and “Value Alignment” (“Highly autonomous AI systems should be designed so that their goals and behaviors can be assured to align with human values throughout their operation.”).

Industry groups have adopted their own soft law instruments for AI. For example, the Information Technology Industry Council (ITI) has developed its own set of AI principles. For example, these principles include a commitment to “recognize our responsibility to integrate principles into the design of AI technologies, beyond compliance with existing laws…. As an industry, it is our responsibility to recognize potentials for use and misuse, the implications of such actions, and the responsibility and opportunity to take steps to avoid the reasonably predictable misuse of this technology by committing to ethics by design.” The statement of principles, itself a form of soft law governance, also states a commitment to soft law principles: “We promote the development of global voluntary, industry-led, consensus-based
standards and best practices. We encourage international collaboration in such activities to help accelerate adoption, promote competition, and enable the cost-effective introduction of AI technologies.”

Company-Specific Soft Law Initiatives

Some individual companies have also adopted their own statement of principles or guidelines for AI. For example, in June 2018 Google’s CEO Sundar Pichai announced a set of seven principles that Google will follow in its AI activities. Other major AI companies such as Microsoft and IBM have also announced their own AI principles that will guide their conduct.

Governmental AI Soft Law Initiatives

Governments have also supported the use of soft law methods to govern AI. The EU Commission published its strategy paper on AI on April 25, 2018. Contrary to what many members of the European Parliament had hoped for and requested, the Commission did not propose any new regulatory measures for AI at this time. Rather, it committed to develop a set of draft guidelines by the end of 2018. In December 2018, the Commission published a “Coordinated Action Plan on AI” that set forth the Commission’s objectives and plans for an EU-wide strategy on AI. However, the Commission did note that “[w]hile self-regulation can provide a first set of benchmarks against which emerging applications and outcomes can be assessed, public authorities must ensure that the regulatory frameworks for developing and using of AI technologies are in line with these values and fundamental rights. The Commission will monitor developments and, if necessary, review existing legal frameworks to better adapt them to specific challenges, in
particular to ensure the respect of the Union’s basic values and fundamental rights.”

Similarly, the UK House of Lords issued a detailed report on AI earlier in April 2018 and likewise recommended an ethical code of conduct for AI rather than any traditional “hard” regulation. The report cited testimony on “the possible detrimental effect of premature regulation” such as that “the pace of change in technology means that overly prescriptive or specific legislation struggles to keep pace and can almost be out of date by the time it is enacted” and that lessons from regulating previous technologies suggested that a “strict and detailed legal requirements approach is unhelpful”. Based on such testimony, the House of Lords therefore concluded that “[b]lanket AI-specific regulation, at this stage, would be inappropriate.”

Instead, the House of Lords recommended a soft law strategy at least in the interim: “We recommend that a cross-sector ethical code of conduct, or ‘AI code’, suitable for implementation across public and private sector organisations which are developing or adopting AI, be drawn up and promoted … with a degree of urgency…. Such a code should include the need to have considered the establishment of ethical advisory boards in companies or organisations which are developing, or using, AI in their work. In time, the AI code could provide the basis for statutory regulation, if and when this is determined to be necessary.”

Evaluation and Moving Forward

A variety of entities from the government, industry and the non-government sectors
have proposed or adopted soft law initiatives for the governance of AI. These soft law instruments include private standards, best practices, codes of conduct, principles and voluntary guidelines. They are in various states of development and implementation, and individually and collectively provide some initial guidance for the governance of AI. However, they suffer from major limitations. One prevalent problem is the generality of most of the provisions in these instruments. To some degree, this vagueness is inevitable and necessary, given the broad range of AI applications and the rapid pace and uncertain trajectory of its future, making precise requirements difficult if not impossible. Indeed, this is the very reason why the technology is primarily being governed by soft law rather than traditional hard law approaches at this time.

Two other limitations of the current matrix of soft law programs are however more amenable to progress and improvement. First, the unenforceability of these soft law provisions is the Achilles’ heel of soft law approaches generally. There is no assurance or requirement that all, or even any, AI developers and users comply with the soft law recommendations. However, there are a number of mechanisms that can be used to indirectly enforce these soft law provisions. Any entity with a supervisory role can adopt and monitor compliance with one or more AI soft law programs. For example, a corporation could create a committee of its Directors or a free-standing ethics committee and task it with ensuring compliance with the appropriate guidelines or codes of conduct adopted by or agreed to by that company. Universities could use the existing chain of authority, such as through department heads and deans, to require compliance with specified soft law AI provisions as part of the annual evaluation of faculty and staff. Or universities could create new, or expand the jurisdiction of existing, research oversight committees such as the Institutional Biosafety Committee to ensure adherence with specified AI
soft law provisions.

Other actors could also play an important role in indirect enforcement of AI soft law programs. Certification bodies could create certification programs to certify that a company or other entity is adhering to a particular set of guidelines or principles. Business partners could require certification with applicable AI soft law programs as a condition of doing business with that company. Insurers could require the implementation of appropriate AI risk management programs as a condition of liability coverage, just as some did with nanotechnology. Granting agencies could condition funding on compliance with specified AI guidelines or codes of conduct. Professional journals could require compliance with certain best practices or guidelines as a condition of publication.

Even more legal quasi-enforcement approaches could be pursued. The Federal Trade Commission (FTC), under its general authority to take enforcement actions against deceptive and unfair business practices, could take enforcement action against a company that publicly commits to comply with a certain code of conduct or best practices but then fails to live up to its commitment. Private standards, especially those adopted by well-known standard setting bodies such as the IEEE, could be used to set a standard of care in tort law, and a company’s failure to adhere to such standards, even though they are voluntary, could be evidence of failure to use reasonable care in a product liability or personal injury lawsuit.

Soft law measures result in experience and field testing that can provide learning for subsequent traditional regulation. Indeed, soft law can sometimes be seen as a transitionary phase of governance that gradually “hardens” into traditional government regulation. We may already be starting to see this hardening process
of soft law in the AI space – for example, the State of California recently adopted legislation “expressing support” for the Asilomar AI Principles.\textsuperscript{37}

Second, the confusing proliferation of different AI soft law programs and proposals creates confusion and overlap with regard to AI governance. It is hard for an actor in the AI space to assess and comply with all these different soft law requirements. Where do these various soft law programs overlap and duplicate each other? Where do they contradict each other? What gaps are not addressed by any of the existing soft law proposals? Some type of coordination is needed.

Wendell Wallach and I have proposed such a coordinating entity, which we have called a Governance Coordinating Committee (GCC).\textsuperscript{38} This entity would not seek to duplicate or supplant the many organizations working on developing governance approaches to AI, but rather would provide a coordinating function much like an orchestra conductor in ensuring all the various players were connected with each other and aware of and responsive to each other’s proposals, while also identifying gaps and inconsistencies in existing programs. In a forthcoming publication, we describe the functions of the GCC to include the following coordination functions:

- \textit{Information clearinghouse}, by collecting and reporting in one place all significant programs, proposals, ideas or initiatives for governing AI;
- \textit{Monitoring and Analysis}, such as identifying gaps, overlaps, and inconsistencies with respect to existing and proposed governance programs;
- \textit{Early Warning System}, by noting emerging issues or problems that are not addressed or covered by existing governance programs;
- \textit{Evaluation Program}, which scores various governance programs and efforts for their metrics and compliance with stated goals.
- \textit{Stakeholder Forum}, by providing a space for stake-holders to meet and discuss
governance ideas and issues and to produce recommendations, reports, and roadmaps;

- Credible Intermediary, serving as a trusted “go-to” source for the media, the public, scholars and stakeholders to obtain information about AI and its governance;
- Convener for Solutions, by convening interested stakeholders on specific issues to meet and try to forge a negotiated partnership program for addressing unaddressed problems or governance needs.  

There are many unanswered questions about how a GCC would function. Who would fund it? Who would be its employees and how would they be selected? What would be its administrative structure? What would be its precise functions and charter? How would stakeholders interact with the GCC? How would the GCC achieve and maintain its credibility as an “honest broker”? Initiatives are currently underway to explore such questions in the context of planning an international conference to discuss and possibly create a global GCC for AI governance.

Conclusion

Soft law measures are very imperfect governance tools because of their lack of enforceability and accountability, as well as often being written in very general and self-serving language. Yet, for a rapidly developing and expansive technology like AI, comprehensive regulation by governments is not feasible, at least in the short term with at best piecemeal regulatory enactments possible. Accordingly, soft law will be the default approach for most AI governance at the present time. For that reason, there is a need to explore ways to indirectly enforce and coordinate the proliferation of soft law measures that have already been proposed or enacted for
“Soft Law” Governance of Artificial Intelligence

by: Gary Marchant

AI.

Gary Marchant, Regent’s Professor of Law and Director of the Center for Law, Science and Innovation, Arizona State University Law School


6. In fact, what Churchill actually said on the floor of the House of Commons was: “No one pretends that democracy is perfect or all-wise. Indeed it has been said that democracy is the worst form of Government except for all those other forms that have been tried from time to time....” House of Commons, 11 November 1947, quoted in Winston Churchill & Richard Langworth, Churchill by Himself: The Definitive Collection of Quotations 574 (2008).

7. Indeed, there has been such a proliferation of soft law programs and proposals for AI that the following examples provide just a sampling and not a comprehensive
Asimov’s three laws are: (1) A robot may not injure a human being or, through inaction, allow a human being to come to harm; (2) A robot must obey the orders given it by human beings except where such orders would conflict with the First Law; and (3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws. These laws were first published in the 1942 short story *Roundabout*, which was published in Asimov’s 1950 collection *I, Robot.*


12. Id. at ii.


17. Future of Life Institute, Asilomar AI Principles,
“Soft Law” Governance of Artificial Intelligence

by: Gary Marchant

18. Id.
20. Id. at 3.
21. Id. at 5.
26. European Parliament resolution with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)); European Economic and Social Committee opinion on AI (INT/806-EESC-2016-05369-00-00-AC-TRA).
“Soft Law” Governance of Artificial Intelligence

by: Gary Marchant

31. Id. at 113.
32. Id. at 116.
33. Id. at 125.
35. Id.
39. Wallach & Marchant, supra note 5.