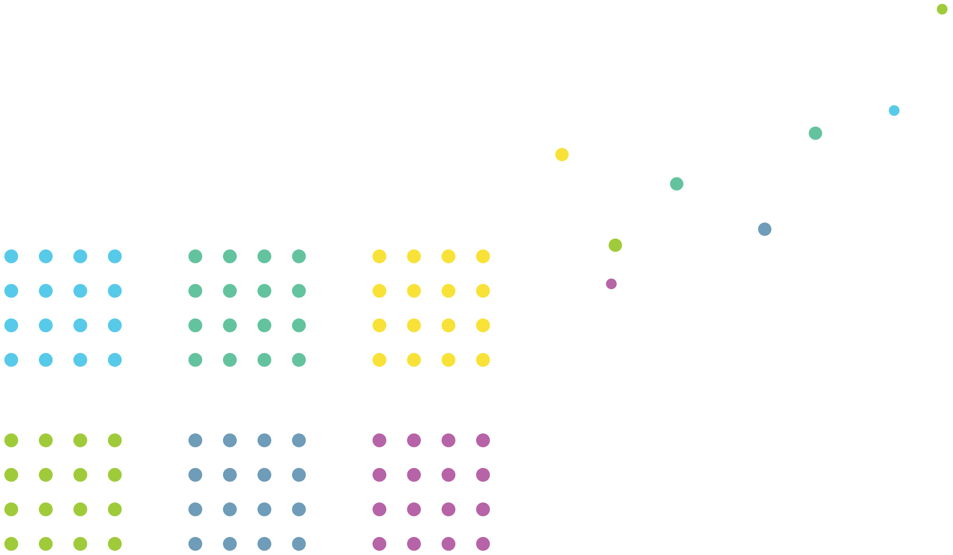
SEPTEMBER 2017

GOVERNING MACHINE LEARNING

Exploring the Intersection Between Machine Learning, Law, and Regulation



This white paper summarizes a discussion that took place in May 2017 amongst top academics, lawyers, and data scientists focused on the intersection of regulation and machine learning. While the following document is meant to provide an overview of that discussion, it does not represent the views of individual participants, their affiliated institutions, or the sponsoring organizations.

Information Society Project
Yale Law School

IMMUTA

EXECUTIVE SUMMARY

On May 16, 2017, the Information Society Project at Yale Law School and Immuta gathered top academics, lawyers, and data scientists for a discussion on governing machine learning (ML). Participants discussed issues relevant to structuring a basic framework for how governments, tech companies, and society writ large might think about governing ML—an area of technology that's transforming the way that key decisions are made, from the justice system to medicine, finance, and self-driving cars. This white paper was prepared to capture shared insights from this discussion.

Key takeaways from the discussion are as follows:

1. ML systems differ from other complex software systems in ways that pose new governance issues: Several features of ML models make governance more challenging than other complex software systems. Such features include the inherent difficulty in understanding many ML systems before they are deployed; the corresponding (though separate) problem of explaining how such models have made any given decision after they have been deployed; the potential demand for new applications of liability; and the possibility that such systems challenge traditional models of agency by performing complex tasks traditionally reserved for humans.

2. Attempting to regulate ML, or regulating too soon, could hinder the technology's development: Regulating ML, as with other forms of technology, may serve to stifle innovation and distort markets due to the nascent state of ML. As such, great care and caution should accompany any new proposed regulations or guidelines meant to govern ML.

3. New case law around ML will emerge, regardless of whether ML is formally governed under new regulations: Where ML is not regulated in any sector, case law will develop, as judges will make determinations regarding liability in the nearly inevitable lawsuits for accidents associated with ML. This fact should be acknowledged when deciding whether to pursue new or additional regulations.

4. Every proposed method of governing ML carries significant risks: Participants suggested a number of methods of governing ML. Each method, however, contained significant downsides that elicited objections from various members of the group.

5. Recommended next steps: Creating an effective governance framework for ML will require a clear definition of what constitutes ML, precise goals, and an understanding of the benefits and drawbacks of each governance approach. As such, we recommend that future efforts attempt to: (1) further specify the features of ML systems that create new governance challenges, so that any future efforts can be properly scoped to ML in particular, rather than complex software systems as a whole; (2) set forth clear and specific goals for any new governance framework; and (3) clearly delineate the costs and benefits created by each governance approach so that tradeoffs can be fully understood and weighed appropriately.

OVERVIEW

On May 16, 2017, the Information Society Project at Yale Law School and Immuta gathered top academics, lawyers, and data scientists for a discussion on governing ML. The goal of the discussion was to shed light on one of the most vexing legal and technological questions we currently confront: how should we—as data scientists, lawyers, and regulators—think about governing the types of ML that are playing an increasingly prominent role in our lives?

This white paper was prepared to capture shared insights from this discussion. A list of participants has been included at the end of this document.

DISCUSSION GUIDE

Participants broke down the discussion into one overarching question, with three component questions, set forth below and addressed separately in Sections I-III of this white paper.

Overarching Question: How should governments, tech companies, and society writ large think about governing ML?

Component 1 — The Why Question: Why do ML models seem to pose new governance problems? What end state, exactly, are we hoping to achieve (or avoid)?

Component 2 — The Who (and When) Question: Who specifically should be governing ML? Congress? Regulatory agencies? Courts? Contracts? And when should this responsibility be assigned?

Component 3 — The How Question: What does this regulatory and policy framework look like in practice, in whole or in part? And how do we get from the present to that desired end state?

A NOTE ABOUT TERMINOLOGY

Talk of "artificial intelligence," or AI, and its impact on technology is pervasive—and frequently inexact. As a result, we chose to focus on ML and one of the techniques most commonly associated with ML, the training and deployment of deep neural networks. Beyond talk of such neural networks, and the general tradeoffs associated with model choice, more specific ML model types did not play a significant role in our discussion. It should also be noted that we did not focus in-depth on the difference between models that are fully trained once deployed and those that continue to train (and therefore to change) after deployment.

I. THE WHY QUESTION

Why do ML models seem to pose new governance problems? What end state, exactly, are we hoping to achieve (or avoid)?

Participants set forth four broad reasons why ML models challenge standard governance constraints, as follows:

- Understandability issues: Central to the value of ML is its ability to extract complex patterns from large sets of data—patterns that are frequently incomprehensible to the human mind. Because of this complexity, it can be difficult to understand how ML models work even before they are deployed. As such, the ability to understand the models' inner workings will be both difficult and frequently limited due to technical reasons (and to intellectual property issues as well). Participants reached a consensus that the inability to fully grasp how ML models were working, even before their deployment, forms a governance challenge that makes ML unique.

- Explainability issues: There was also consensus that the issue of understandability is distinct from the issue of explainability—that is, understanding why and how a model has reached a particular decision after the fact is a separate (and sometimes technically intractable) challenge from attempting to understand a model before it has been deployed. We might know that dozens of weighted features led to one particular decision, for example, but we still might not be able to explain in aggregate why that decision was made or what changes could alter that decision in the future. Whether, and when, post hoc explanations might be demanded of ML models was set forth as a separate issue. Some participants alluded to early efforts in the European Union on this subject, as manifested in the General Data Protection Regulation, effective May 2018.

- Liability issues: Some participants suggested that ML poses liability challenges that are distinct from traditional software systems. There was, however, debate on whether this assertion was accurate. On the one hand, for example, it will usually be clear who designed an ML model, and the data sets on which it was originally trained. Imposing liability upon an ML model's creators does not therefore immediately constitute a new legal challenge. On the other hand, ML models may change and evolve throughout their deployment and use of data; to the extent that such systems might adapt over time, assigning liability between those who built the models and those who use and maintain them may pose new challenges.

- Agency issues: Participants also suggested that as ML models are adopted in increasingly sensitive environments, we might rely on ML in a manner similar to the way we rely on other humans. To that extent, we may be granting a new type of "agency" to such models that we don't traditionally grant to other complex forms of software. This in itself, it was suggested, might spur us to think about governing ML in new ways.

II. THE WHO (AND WHEN) QUESTION

Who specifically should be governing ML? Congress? Regulatory agencies? Courts? Contracts? And when should this responsibility be assigned?

This question led to vigorous debate, as there was broad recognition of the fact that there is no accepted baseline for what such regulation would look like in practice. Would the ML models themselves be regulated? Their weighting? The data sets they are trained upon? And what would differentiate such attempts from existing regulations surrounding a product's safety and soundness?

Noting the problems associated with regulatory overlap in non-ML contexts, one participant posited that the latest state-of-the-art tractor fell under the regulatory jurisdiction of over a dozen agencies in the U.S. alone. As such, some participants raised the possibility that ML might benefit from a single regulatory body setting standards and clarifying liability, but this proposal was met with a significant amount of skepticism. Others argued against the idea of standardizing the way that ML might be regulated. Some, for example, argued that any new form of regulation or regulatory authority over ML would stifle innovation of a nascent industry, attempt to solve for problems that haven't yet arisen, and potentially create barriers to entry for new entrants.

Participants reached consensus, however, around the notion that timing is critical, and even those who favored an increased regulatory burden on ML argued that it would be a mistake to attempt to regulate ML too soon, or to determine in advance what bodies are best equipped to draft or enforce those regulations. Significant dangers could exist, for example, in attempting to legally constrain the field of ML before it is fully developed in practice. As such, it was agreed that the best and most prudent approach to governing ML would necessarily take timing as a major factor.

In waiting for the use of ML to mature, however, participants also acknowledged the likelihood that new rules may be generated by default in the form of case law. As ML is deployed, for example, damages will likely arise, and courts will create new case law on the matter. Some thought this outcome could be avoided, while others believed this was an intelligent "hands-off" approach (described in further detail below), insofar as it would allow real-world damages, rather than hypothetical ones, to drive the rules surrounding ML.

Overall, the group found both merits and drawbacks associated with every possible regulatory approach. No consensus was reached on this subject.

III. THE HOW QUESTION

What does this regulatory and policy framework look like in practice, in whole or in part? And how do we get from the present to that desired end-state?

Given the lack of consensus on the questions addressed in Section II, participants varied widely in their ideas—and their enthusiasm—for proposed methods to govern ML, or whether such methods were necessary at all. A variety of approaches were suggested, as summarized below:

- A hands-off approach: One course would allow the market to shape ML, or what one participant deemed a “wait and see” or an “innovation-focused” approach. This approach would involve a mix of current tort liability and contract-based governance mechanisms, allowing individuals to determine liability on a case-by-case basis, and applying to ML the legal framework currently applied to complex software systems.

- Defining regulatory standards: In areas where ML might either need to be standardized or might have particularly sensitive or important use cases (such as in aviation or medicine), some argued that clear regulatory requirements could be applied to the development, testing, and deployment of ML by sector-specific regulators that already have regulatory authority over products or services that would use ML technologies. Such requirements could be enforced by one or multiple regulatory bodies and could encompass the creation of ML models, post hoc auditing, or both.

- Professional associations for data scientists: ML might also be governed through codes of conduct enforced by professional associations, much like regulated professions such as architects, doctors, and lawyers. Although the roles and experiences of working data scientists vary significantly in practice today, such associations were suggested by some as a potential middle ground between vigorous regulatory oversight and the hands-off governance approach. It should be noted, however, that this approach could lead to significant new licensing burdens on ML practitioners and academics, which many participants deemed unreasonable and likely to stifle innovation. To allay this burden, professional organizations might also create voluntary accreditations, similar to those that exist in the privacy and cybersecurity fields, which, while not mandatory, frequently generate preferential treatment by hiring organizations.

- Voluntary best practices: No current data science organization advocates clear, commonly accepted best practices in data governance to avoid the many types of harms that can arise from deploying ML without sufficient care. Non-binding professional standards may help to further the nascent field of data science and to lay the groundwork for best practices that might also be applied in legal settings when determining reasonable levels of care.

CONCLUSION

The wide range of perspectives among participants did not lend itself to concrete conclusions about how best to approach the governance challenges created by ML. As such, the differing viewpoints we encountered might foreshadow those likely to arise if lawmakers or regulators seek to engage the broader data science community on the subject.

There was, however, a general awareness on the part of participants that creating new regulatory burdens without sufficient thought, or at too early a time in the development of ML, could negatively impact the potential benefits of the technology. Creating a framework for how to think about these issues was therefore deemed critical.

As such, we recommend that any future framework for governing ML take the following three steps into account:

1. Define ML with precision: Any efforts to govern ML must be built upon a clear delineation of what differentiates ML from other complex software systems. Participants outlined some of the challenges ML poses in broad terms, set forth in Section I above. Future governance efforts, however, will require further specifying what, exactly, differentiates ML systems and their associated governance problems from more traditional types of complex software.

2. Set clear goals: In order to successfully govern the creation and deployment of ML, clear goals—specifying what such efforts are attempting to achieve and what they are attempting to avoid—will be needed. Such goals could include mandating levels of explainability in certain contexts, preventing specific types of bias in ML systems, or specifying what types of models, training data, or new data sets can be used for what purposes. Outcomes to avoid, on the other hand, may include stifling innovation or creating a patchwork of country-specific regulations that hinder the global deployment of ML systems.

3. Examine tradeoffs and benefits of each approach: Any attempts to govern ML will alter its deployment and, potentially, its long-term utility. Limitations on the use of specific models due to explainability concerns, for example, could have significant drawbacks in terms of the accuracy of such models—drawbacks that could cause literal harm in the case of autonomous vehicles or in medical environments. Such drawbacks illustrate the serious tradeoffs associated with governing ML and should therefore be fully understood when weighing the benefits of any new governance approach.

As one academic remarked to one data scientist, “You might not be interested in regulation, but regulation is interested in you.” The importance of these issues, in other words, may become apparent at a faster pace, and with broader implications, than many in the technology and legal communities expect.

LIST OF PARTICIPANTS

JACK BALKIN

Director, Information Society Project Knight Professor of Constitutional Law and the First Amendment, Yale Law School

KIEL BRENNAN-MARQUEZ

Visiting Fellow, Information Society Project
Research Fellow, Information Law Institute, New York University

ANDREW BURT

Chief Privacy Officer & Legal Engineer, Immuta
Visiting Fellow, Information Society Project

MATTHEW CARROLL

Chief Executive Officer, Immuta

REBECCA CROOTOF

Executive Director, Information Society Project
Research Scholar and Lecturer in Law, Yale Law School

JIM FENTON

Consultant and Researcher
Former Chief Security Officer, OneID

AMY GERSHKOFF

Chief Data Officer, Ancestry

ADAM GOLODNER

Senior Counsel, Arnold & Porter Kaye Scholer

DAN JACOBSON

Associate, Arnold & Porter Kaye Scholer

NANCY PERKINS

Counsel, Arnold & Porter Kaye Scholer

KANISHK PRIYADARSHI

Data Science Executive

DAVID ROBINSON

Principal, Upturn
Visiting Fellow, Information Society Project

MIKE ROMERI

Managing Director, Accenture Analytics

ANDREW SELBST

Visiting Fellow, Information Society Project
Visiting Researcher, Georgetown University Law Center

RAMESH SUBRAMANIAN

Visiting Fellow, Information Society Project

GABRIEL FERRUCCI

Professor of Computer Information Systems, School of Business, Quinnipiac University

STEVE TADELIS

Professor, Haas School of Business, University of California Berkeley
VP Economics & Market Design, Amazon

STEVE TOUW

Chief Technology Officer, Immuta

DAVID WOLF

Regional Managing Director, Accenture Digital