

High-Tech International Law

Ashley Deeks*

ABSTRACT

Data-driven algorithmic tools allow their users to process large amounts of data quickly, extract patterns from the data that humans cannot otherwise detect, and make reliable predictions. These tools have proven valuable in domestic legal practice, negotiations, and other fields closely related to international law. Yet governments and their international lawyers have ignored this new wave of tools as they conduct international relations.

This Article argues that governments should begin to use these tools to their advantage in creating and implementing international law. These capabilities could support states' decision-making as they negotiate treaties, adjudicate international legal disputes, and evaluate the status of customary rules. Even if international lawyers for governments in the United States and Europe are skeptical about the benefits of machine learning and big data, they must consider the possibility that states such as China will begin to deploy these tools in power-enhancing ways.

High-tech international lawyering will have important implications for international law and relations. Tools that are readily available and easy to use will serve as important equalizers for less powerful states, but more advanced technologies are likely to exacerbate existing power differentials. The Article therefore offers ways that less powerful states and outside actors can counterbalance technology-driven power shifts that may undercut the enduring ambitions of international law.

TABLE OF CONTENTS

INTRODUCTION	576
I. THE INSINUATION OF TECHNOLOGY INTO (INTERNATIONAL) LAW	582
A. <i>Machine Learning's Influence</i>	583
B. <i>Machine Learning and Computational Text Analysis in Domestic Legal Settings</i>	586
C. <i>The Roles of a State's International Lawyers</i>	589

* E. James Kelly Jr.—Class of 1965 Research Professor of Law, University of Virginia Law School. Thanks to William Alford, Jon Ashley, James A. Baker, Drew Conway, Marc DaCosta, Jennifer Daskal, Mike Flowers, Maggie Gardner, Jack Goldsmith, Duncan Hollis, Rebecca Ingber, Malcolm Langford, Dustin Lewis, Michael Livermore, Noam Lubell, Daragh Murray, Sergio Puig, Cecily Rose, Colin Rule, Paul Stephan, and Larissa van den Herik, as well as participants in the Harvard Law School International Law Workshop, for very helpful conversations at various stages of this project. Thanks to Scott Harman-Heath for excellent research assistance.

D.	<i>Will Tech Come to International Law?</i>	593
1.	Near-Peer Ambitions	594
2.	Proofs of Concept in Private Law	595
3.	Client Pressure	596
4.	Necessity	596
5.	Challenges	598
II.	CREATING AND IDENTIFYING INTERNATIONAL LAW	600
A.	<i>Preparing for Treaty Negotiations</i>	600
1.	Identifying Negotiating Partner Equities	601
2.	Predicting Possible Outcomes	604
3.	Identifying Foreign Law	606
4.	Processing Intelligence	607
B.	<i>Conducting Treaty Negotiations</i>	608
1.	Interpretation and Translation	609
2.	Parsing Partner Behavior	612
3.	Identifying Inter-Treaty Conflicts	616
C.	<i>Identifying Customary International Law</i>	616
III.	RESOLVING INTERNATIONAL DISPUTES	622
A.	<i>Understanding the Workings of Arbitral and Judicial Tribunals</i>	622
1.	Extracting Information About Tribunal Decision-making	624
a.	<i>Sources of Data</i>	624
b.	<i>Parsing Past Opinions to Shape Future Arguments</i>	625
c.	<i>Predicting Outcomes</i>	628
2.	Arbitrating Algorithmically?	630
B.	<i>Resolving Foreign Policy Disputes in Diplomatic Channels</i>	633
1.	Locating Third Party Proposals	636
2.	Parsing Local Reactions	637
3.	Manipulating Local Reactions	638
IV.	CAN TECH STRENGTHEN INTERNATIONAL LAW?	640
A.	<i>Technological and Ethical Challenges</i>	641
B.	<i>High-Tech Tools' Distributional Effects</i>	643
1.	Distributional Effects Among States	644
2.	Distributional Effects Within States	647
C.	<i>Implications for International Law and Diplomacy</i> ..	647
D.	<i>Facilitating Redistribution?</i>	650
CONCLUSION	652

INTRODUCTION

Imagine: three years from now, the United States and Russia decide to negotiate an extradition treaty. Newspapers in both countries have been reporting on the impending negotiation, and the public has been discussing it on social media. Russia already has 50 extradition treaties with other states. U.S. international lawyers and data scientists, sitting side by side, are able to use the texts of these agreements to identify what the “average” Russian extradition treaty looks like, extract the most common (and most exceptional) terms contained in the existing treaties, and enter the negotiation well positioned to anticipate Russian requests. As the negotiation gets underway, the U.S. negotiators mine information from Russian social media to help them understand public sentiment. From this, they anticipate that the draft treaty provision about political offenses is likely to be contentious, and the U.S. side drafts the proposed provision to take account of that controversy. Later, during the negotiation, the U.S. team trains concealed video cameras on the Russian negotiators. During a break, the U.S. team runs emotion recognition algorithms over the videos to identify the parts of the negotiation during which the Russian team acted deceptively and adjust their strategy accordingly.

International lawyers who work for governments do not use these technologies today. Indeed, international law generally has been a stranger to a new wave of technological tools—including computational text analysis, machine learning, and predictive algorithms—that use large quantities of data to help make sense of the world.¹ But actors in a wide range of other fields use data-driven algorithmic tools to improve medical diagnoses, enable greater autonomy in daily activities such as driving, enhance the accuracy of facial recognition software, facilitate interpersonal negotiations, and improve decision-making.² Further, these tools have proven valuable in domestic legal practice, negotiations, and other fields closely related to international law.³ This raises important questions: Are there ways in which international lawyers for governments can, should, or will begin to use these tools to their advantage, and if so, how? Further, employing these tools may enhance the power of those who use them. What types of

¹ See *infra* notes 100–01 and accompanying text. See also Wolfgang Alschner et al., *The Data-Driven Future of International Economic Law*, 20 J. INT’L ECON. L. 217, 218 (2017) (describing these technologies as opening a “new frontier” in international economic law scholarship).

² See *infra* Section I.A.

³ See *infra* Section I.B.

distributional effects will this have on international law and interstate relations? This Article begins to grapple with these unexplored issues.

In the domestic legal setting, lawyers and scholars have begun to embrace machine learning's possibilities. Lawyers at private firms employ these tools to help them identify relevant precedent, identify which arguments have been most persuasive to particular judges, predict the outcomes of cases, and review contracts.⁴ These tools help them manage large quantities of data and identify connections that would otherwise be difficult for humans to see.⁵ Likewise, legal scholars have begun to deploy computational, data-driven tools to help them make sense of vast realms of written texts.⁶ Some algorithms can identify topics and themes within these texts.⁷ Other algorithms allow computers to translate handwritten texts into a digital format.⁸ Yet other tools allow users to predict judicial outcomes based on case characteristics.⁹ It is not difficult to see how law, the very corpus of which consists of huge quantities of written words, is a suitable target for some of these algorithms.

Yet states, and those who advise states on international law issues, are lagging far behind private sector lawyers and academic researchers in contemplating how machine learning and text-as-data tools can and will change—and facilitate—their legal work.¹⁰ Even

4 Companies such as ROSS Intelligence, Inc. increasingly employ cognitive computing technologies in their search tools and information retrieval. See *Features*, ROSS INTELLIGENCE, <https://rossintelligence.com/features.html> [<https://perma.cc/R5F7-AH7X>].

5 See Daniel Martin Katz, *Quantitative Legal Prediction—or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 928–29 (2013) (“[Quantitative legal prediction]-based technologies are designed to remedy or supplement the shortcomings of human reasoners. . . . When it comes to processing and deriving insights from large-scale data or document sets, humans have important cognitive limitations.”).

6 See, e.g., Alschner et al., *supra* note 1, at 218.

7 See, e.g., Priya Dwivedi, *NLP: Extracting the Main Topics from Your Dataset Using LDA in Minutes*, TOWARDS DATA SCI. (Aug. 22, 2018), <https://towardsdatascience.com/nlp-extracting-the-main-topics-from-your-dataset-using-lda-in-minutes-21486f5aa925> [<https://perma.cc/M3LR-PEES>] (describing topic modeling).

8 See Batuhan Balci et al., *Handwritten Text Recognition Using Deep Learning*, STAN. U. (2017), <http://cs231n.stanford.edu/reports/2017/pdfs/810.pdf> [<https://perma.cc/9PSH-H488>] (describing use of convolutional neural networks to train an algorithm to translate handwritten text into a digital form).

9 See Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 102–05 (2014) (discussing how machine learning algorithms can be used to make predictions about legal outcomes); Theodore W. Ruger et al., *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150 (2004).

10 See Hannah Bryce & Jacob Parakilas, *Conclusions and Recommendations*, in CHATHAM

though machine learning tools have penetrated commercial, military, and scientific arenas, states have made few changes in how they negotiate treaties or adjudicate international legal disputes to take advantage of these technological developments. At least one scholar has articulated skepticism about the potential role for artificial intelligence (“AI”) in international law at all.¹¹

There is one important exception to this tendency to treat international law and AI as existing on different planets.¹² For the past six years, states have been engaged in intensive discussions about lethal autonomous weapons systems.¹³ These discussions are occurring because some states, scholars, and activists fear that developments in AI are leading inexorably toward the creation of “killer robots” that will choose independently who to target during armed conflicts.¹⁴ This ominous framing of one high-profile AI development has left little room for a conversation about how AI and other high-technology tools might contribute in other ways—including positive ways—to the processes by which states create and identify international law and resolve international disputes.¹⁵

Why would a state’s international lawyers and diplomats want to employ these tools? Algorithmic capabilities allow their users to process large amounts of data quickly, to extract patterns from the data that humans could not otherwise detect, and to make predictions.¹⁶

HOUSE, *ARTIFICIAL INTELLIGENCE AND INTERNATIONAL AFFAIRS* 43, 44 (2018) (noting that artificial intelligence operates at a “speed, scale, and breadth” that government cannot).

11 Thomas Burri, *International Law and Artificial Intelligence*, 60 *GERMAN Y.B. INT’L L.* 91, 95 (2017) (“[International law’s] data remains too limited for artificial intelligence to learn to give a sound legal assessment.”). This is belied by Alschner et al., *supra* note 1, and the work of other scholars described therein.

12 In addition, international human rights lawyers have researched how tools that use AI (such as facial recognition software) implicate international rules protecting privacy. *See, e.g.*, Pete Fussey & Daragh Murray, *INDEPENDENT REPORT ON THE LONDON METROPOLITAN POLICE SERVICE’S TRIAL OF LIVE FACIAL RECOGNITION TECHNOLOGY* 34–39 (2019), <https://48ba3m4eh2bf2sksp43rq8kk-wpengine.netdna-ssl.com/wp-content/uploads/2019/07/London-Met-Police-Trial-of-Facial-Recognition-Tech-Report.pdf> [<https://perma.cc/G7WK-GWD2>].

13 *Background on Lethal Autonomous Weapons Systems in the CCW*, UNITED NATIONS GENEVA, [https://unog.ch/80256EE600585943/\(httpPages\)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument](https://unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument) [<https://perma.cc/W8ZD-2FZ7>].

14 *See, e.g.*, Bonnie Docherty, *Heed the Call: A Moral and Legal Imperative to Ban Killer Robots*, HUMAN RIGHTS WATCH (Aug. 21, 2018), <https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots> [<https://perma.cc/8G42-A4AN>].

15 *See, e.g.*, Bryce & Parakilas, *supra* note 10, at 43 (“[The] dominant focus on . . . ‘killer robots’ . . . risks obfuscating a more nuanced discussion of the much broader, non-lethal applications of AI.”).

16 *See* Jacob Parakilas & Hannah Bryce, *Introduction: Artificial Intelligence and International Politics*, in CHATHAM HOUSE, *supra* note 10, at 1, 1 (identifying machines’ ability to

Faced with large volumes of information (all of a state's submissions ever made to the United Nations, for instance, or thousands of bilateral investment treaties), a foreign ministry lawyer with limited time and resources could deploy algorithmic tools to help her extract important insights from that data, improving her negotiation posture, or allowing her to provide more accurate advice to a client about how an international arbitral panel will rule.¹⁷ Even if she herself is skeptical about these tools, she may face pressure to adopt them because her foreign counterparts are doing so, or because her policy clients demand it. This Article is not intended to suggest that humans soon will, can, or should turn over these types of international activities wholesale to machines.¹⁸ Instead, it suggests that there can be significant benefits to allowing machine learning tools to support human decision-making in these areas.¹⁹ Even if international lawyers in the United States and Europe are skeptical about the benefits of machine learning and big data, they must consider the possibility that states such as China may begin to use these tools—and gain important advantages by doing so.

High-tech international law has the potential to significantly alter the distribution of power among states. Any evaluation of international law technologies should therefore also consider who the winners and losers will be as this technology gains a foothold in processes such as treaty negotiation and dispute resolution. Here, the story appears mixed: Some of these tools will empower actors in less powerful states, improving their access to information and their ability to use that information instrumentally.²⁰ Other tools will empower the al-

quickly access and process large quantities of data and identify patterns); *Machine Learning: Of Prediction and Policy*, *ECONOMIST* (Aug. 20, 2016), <https://www.economist.com/news/finance-and-economics/21705329-governments-have-much-gain-applying-algorithms-public-policy> [<https://perma.cc/3VMU-NA8L>] (“Machine-learning systems excel at prediction.”).

¹⁷ See Parakilas & Bryce, *supra* note 16, at 1; *Machine Learning: Of Prediction and Policy*, *supra* note 16.

¹⁸ See generally Parakilas & Bryce, *supra* note 16, at 1 (arguing that AI is likely to affect international politics in subtle, diffuse ways, and that the impacts are “likely to be manifested through changes in the ways in which human decision-makers are informed, while not extending so far as to allow AI to take decisions”).

¹⁹ See, e.g., Bryce & Parakilas, *supra* note 10, at 43 (envisioning a “mechanism by which decision-making processes are enhanced [by AI] but ultimate responsibility still lies in human hands”). For an excellent discussion of ways in which international lawyers and scholars could benefit from the use of “text-as-data” tools to “conduct more comprehensive and evidence-based doctrinal and legal-institutional research,” see Wolfgang Alschner, *The Computational Analysis of International Law*, in *RESEARCH METHODS IN INTERNATIONAL LAW* (Rossana Deplano & Nicholas Tsagourias eds., forthcoming 2020) (manuscript at 2).

²⁰ See *infra* Section IV.B.1.

ready powerful, further challenging the fiction of sovereign equality.²¹ There are steps that less powerful states can take, however, to guard against their serious disempowerment.

The processes of creating international law and navigating international disputes will not be the only roles for high-technology tools in this arena. Machine learning also has the potential to shape *substantive* interstate interactions, as where one state seeks to export self-driving cars (which rely on machine learning) to another state, or where one state makes an extradition request for a person based entirely on evidence derived from facial recognition software. Machine learning also will facilitate the operations of international organizations such as the World Health Organization (by allowing it to detect early signs of pandemics)²² or U.N. peacekeeping (by enabling forecasting, planning, and logistics).²³ The U.N. World Food Programme recently announced a five-year partnership with Palantir Technologies to use data analytics to improve efficiency and real-time decision-making about food distribution.²⁴ There may also be roles for machine learning in enforcing international law²⁵ and detecting looming civil wars or genocide.²⁶ Further, machine learning and AI tools may advance the work of international criminal lawyers and civil society by enhancing the identification of evidence for prosecutors and of defenses for the accused.²⁷ There are therefore a host of other axes along which machine learning and international lawyers will interact.

²¹ See *infra* Section IV.B.1.

²² See Tedros Adhanom Ghebreyesus, *Artificial Intelligence for Good Global Summit*, WORLD HEALTH ORG. (May 15, 2018), <http://www.who.int/dg/speeches/2018/artificial-intelligence-summit/en/> [<https://perma.cc/49H9-AA76>] (discussing how AI can improve responses to disease outbreaks through early warning, forecasting, and simulation).

²³ Heather M. Roff, *Advancing Human Security Through Artificial Intelligence*, in CHATHAM HOUSE, *supra* note 10, at 19, 25.

²⁴ *Palantir and WFP Partner to Help Transform Global Humanitarian Delivery*, WORLD FOOD PROGRAMME (Feb. 5, 2019), <https://www1.wfp.org/news/palantir-and-wfp-partner-help-transform-global-humanitarian-delivery> [<https://perma.cc/49H9-AA76>].

²⁵ See Parakilas & Bryce, *supra* note 16, at 3 (envisioning a role for AI in monitoring sensors set up to verify compliance with treaties that regulate chemical, biological, or nuclear weapons); Nevena Jevremović, *2018 in Review: Blockchain Technology and Arbitration*, WOLTERS KLUWER (Jan. 27, 2019), <http://arbitrationblog.kluwerarbitration.com/2019/01/27/2018-in-review-blockchain-technology-and-arbitration/> [<https://perma.cc/83Y2-VESQ>] (discussing possibility of using blockchain to enforce international arbitral awards).

²⁶ See Mary-Ann Russon, *CIA Using Deep Learning Neural Networks to Predict Social Unrest Five Days Before It Happens*, INT'L BUS. TIMES (Oct. 6, 2016, 12:29 BST), <https://www.ibtimes.co.uk/cia-using-deep-learning-neural-networks-predict-social-unrest-five-days-before-it-happens-1585115> [<https://perma.cc/7L5W-RL6P>].

²⁷ See Berenice Boutin, *Technologies for International Law & International Law for Tech-*

This Article, however, focuses on the *procedural* contributions that machine learning and text-as-data tools can make to government lawyers' interactions with international law. It leaves for another day consideration of how machine learning will alter the substantive application of existing international law, whether it potentially will require new regulations, and how it could serve as a tool for other types of (non-governmental) international lawyers or other governmental (non-international) lawyers.²⁸ This Article is meant to provide a starting point for those interested in furthering states' ability to create international law and resolve their international disputes using text-as-data analyses of international law and foreign policy documents.

Part I identifies the growth of machine learning tools across a range of fields. It focuses in particular on how private law firms have begun to use these tools to conduct a range of lawyering tasks. It then considers the multitude of roles that international lawyers in government play in support of their states' diplomatic, economic, and military goals; identifies commonalities between private and international legal work; and explores several reasons why international lawyers may wish to pursue the same types of machine learning tools being developed for private sector lawyering.

Parts II and III show how states might create comparable tools to accomplish (and in some cases improve) their international legal work, drawing from a range of machine learning and computational text analysis tools used in the private law and academic settings. Part II considers international law creation and identification. It argues that there is a wealth of text-as-data tools that treaty negotiators could employ to improve their negotiating positions and bring negotiations to a successful conclusion. It also explores ways in which text-as-data tools could alter our views about the content of customary international law. Part III turns to international dispute resolution. It first analyzes how states might use text-as-data tools to improve the ways in which they currently resolve disputes in arbitral and judicial tribunals, and then identifies possible tools for dispute resolution conducted in diplomatic channels.

Part IV evaluates the normative implications of states' use of these tools for international law and relations. It first identifies some

nologies, GRONINGEN J. INT'L L. (Oct. 22, 2018), <https://grojil.org/2018/10/22/technologies-for-international-law-international-law-for-technologies/> [<https://perma.cc/F9V9-RE5F>].

²⁸ See Parakilas & Bryce, *supra* note 16, at 6 ("There are significant areas—among them medicine, public health and law—where AI systems may be transformative in ways that directly impact the processes of the international system in the next decade or two.").

technical and ethical hurdles that states must surmount before they deploy these tools in international law settings. It then considers the extent to which some of these tools offer the possibility of more evenly distributing power among states, and the extent to which these tools might instead enhance the existing power of already powerful states. It concludes that those tools that are readily available and easy to use will serve as important equalizers. At the same time, states with a high level of technological sophistication are likely to treat some of these tools as proprietary and critical to their national security, and so may use them in a way that exacerbates existing power differentials. This Part concludes by suggesting ways in which less powerful states and outside actors might help counterbalance power shifts that could hinder, rather than advance, the ambitions of international law.

Before turning to a discussion of technology's future in international law, a word of caution is in order. Many people have high hopes for AI, machine learning, and text-as-data tools. Some of these hopes have borne out, but it is easy to fall prey to irrational exuberance when it comes to these tools. This Article tries to stay grounded in the tools that already exist for analyzing large amounts of data, translating languages, and making accurate predictions. It is impossible, however, to know precisely which tools will prove useful and which will ultimately be less effective than current human analysis. Further, it will be a slow, costly, and iterative process to develop these tools in a bespoke context such as international law using data that is not already collated and cleaned. The goal of this Article is not to argue that international lawyers may and should develop and deploy all of these tools in the near term. The goal is instead to consider the categories of tools that states and data scientists may be able to develop for international lawyering; identify the situations in which those tools might be most effective; and stimulate the imaginations of international lawyers and foreign-policy-minded data scientists as they look over the horizon.

I. THE INSINUATION OF TECHNOLOGY INTO (INTERNATIONAL) LAW

AI is all the rage. AI is a catch-all term that generally refers to “a set of techniques aimed at approximating some aspect of human or animal cognition using machines.”²⁹ One important category of AI is machine learning. Data scientists train machine learning systems on large amounts of data, and the systems adjust their own parameters to

²⁹ Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 404 (2017).

improve the reliability of their predictions.³⁰ Machine learning has rapidly spread through the fields of national security, medicine, criminal justice, advertising, and self-driving cars (to name just a few).³¹ These algorithms rely on large volumes of data and powerful computing to make accurate predictions and recommendations,³² and are able to detect patterns in data that are very difficult for humans to see.³³

For decades, technology has played a critical part in the practice of law,³⁴ but its influence has taken a sharp upturn with these rapid advancements in machine learning. Private law firms now employ machine learning tools to conduct discovery, craft legal documents, and predict case outcomes.³⁵ This Part argues that these kinds of technologies may begin to appear in states' international law processes as well, even though the types of tools will not be identical to those already in use by private lawyers. The reasons that international lawyers may pursue their use also are partly distinct from the reasons driving private law firms to use these tools.

A. *Machine Learning's Influence*

The past 10 years have begun to reveal the power of AI. Both the United States and China have fully embraced what they view as AI's game-changing economic and military potential. On the U.S. side, the Pentagon intends to invest up to \$2 billion over the next five years to develop new programs advancing AI,³⁶ and reportedly spent about

³⁰ See ETHEM ALPAYDIN, *MACHINE LEARNING* 24–25 (2016); PETER FLACH, *MACHINE LEARNING* 3 (2012).

³¹ See, e.g., Exec. Order No. 13,859, 84 Fed. Reg. 3,967 (Feb. 11, 2019) (national security); Jane R. Bambauer, *Dr. Robot*, 51 U.C. DAVIS L. REV. 383 (2017) (medicine); Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109 (2017) (criminal justice); Sam Biddle, *Facebook Uses Artificial Intelligence to Predict Your Future Actions for Advertisers, Says Confidential Document*, INTERCEPT (Apr. 13, 2018, 11:28 AM), <https://theintercept.com/2018/04/13/facebook-advertising-data-artificial-intelligence-ai/> [<https://perma.cc/SY4M-4RK3>] (advertising); Alex Davies, *The WIRED Guide to Self-Driving Cars*, WIRED (Dec. 13, 2018, 6:00 AM), <https://www.wired.com/story/guide-self-driving-cars/> [<https://perma.cc/R93D-M8FW>] (self-driving cars).

³² See Cary Coglianese & David Lehr, *Transparency and Algorithmic Governance*, 71 ADMIN. L. REV. 1, 6 (2019).

³³ See Katz, *supra* note 5, at 928–29 (“[Quantitative legal prediction]-based technologies are designed to remedy or supplement the shortcomings of human reasoners. . . . When it comes to processing and deriving insights from large-scale data or document sets, humans have important cognitive limitations.”).

³⁴ See, e.g., Dana Remus & Frank Levy, *Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law*, 30 GEO. J. LEGAL ETHICS 501, 515 (2017) (describing ways that technology impacted law in past decades).

³⁵ *Id.* at 508–31.

³⁶ See JIM MATTIS, U.S. DEP'T OF DEF., SUMMARY OF THE 2018 NATIONAL DEFENSE

\$7.4 billion on AI and associated fields in 2017.³⁷ In February 2019, the Trump Administration issued an Executive Order on Artificial Intelligence, asserting that “[c]ontinued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States.”³⁸ On China’s part, its State Council in 2017 issued a road map for AI development that envisions China becoming the world leader in AI by 2030.³⁹ The World Intellectual Property Organization recently reported that actors in the United States and China have taken a big lead in AI patent applications.⁴⁰ Both states—as well as Russia, Japan, and some European states—are pursuing a range of AI technologies.⁴¹

These advances have been driven by the growing availability of huge quantities of data, as well as much improved computer processing power.⁴² Machine learning systems and algorithms, the driving force behind many AI developments, are valuable because of their ability to learn for themselves “how to detect useful patterns in massive data sets and put together information in ways that yield remarkably accurate predictions or estimations.”⁴³ Another set of computational tools with direct relevance for the law has also been growing in importance and sophistication. “Computational text analysis,” some of which relies on machine learning, is the use of computa-

STRATEGY OF THE UNITED STATES OF AMERICA 7 (2018), <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf> [<https://perma.cc/W672-3XNK>]; Drew Harwell, *Defense Department Pledges Billions Toward Artificial Intelligence Research*, WASH. POST (Sept. 7, 2018, 10:39 AM), <https://www.washingtonpost.com/technology/2018/09/07/defense-department-pledges-billions-toward-artificial-intelligence-research> [<https://perma.cc/U4J5-X7BY>].

³⁷ Julian E. Barnes & Josh Chin, *The New Arms Race in AI*, WALL STREET J. (Mar. 2, 2018, 11:47 AM), <https://www.wsj.com/articles/the-new-arms-race-in-ai-1520009261> [<https://perma.cc/R8EM-TDW7>].

³⁸ Exec. Order No. 13,859, 84 Fed. Reg. 3,967 (Feb. 11, 2019).

³⁹ See Sarah Dai & Alice Shen, “*Made in China 2025*”: *China Has a Competitive AI Game Plan but Success Will Need Cooperation*, S. CHINA MORNING POST (Oct. 1, 2018, 11:00 PM), <https://www.scmp.com/tech/article/2166177/made-china-2025-china-has-competitive-ai-game-plan-success-will-need> [<https://perma.cc/9H76-9RV5>].

⁴⁰ See WORLD INTELL. PROP. ORG. [WIPO], TECHNOLOGY TRENDS 2019: ARTIFICIAL INTELLIGENCE 16 (2019), https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf [<https://perma.cc/L6RA-S2XS>].

⁴¹ See THE EUROPEAN ARTIFICIAL INTELLIGENCE LANDSCAPE, EUROPEAN COMM’N (Apr. 18, 2018), <https://ec.europa.eu/digital-single-market/en/news/european-artificial-intelligence-landscape> [<https://perma.cc/7YFM-MH79>]; Tim Dutton, *An Overview of National AI Strategies*, MEDIUM (June 28, 2018), <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd> [<https://perma.cc/38RY-CR6E>].

⁴² See Stuart J. Russell, *Q&A: The Future of Artificial Intelligence*, U.C. BERKELEY, <https://people.eecs.berkeley.edu/~russell/research/future/q-and-a.html> [<https://perma.cc/5CNX-83JN>].

⁴³ Coglianese & Lehr, *supra* note 32, at 6.

tional techniques to categorize, analyze, and extract information from a large corpus of text.⁴⁴ Because these analyses rely on text as the raw material for the computational work, they sometimes are called “text-as-data” methods.⁴⁵ This Article uses the umbrella idea of “text-as-data” tools to capture a range of techniques that data scientists have developed to analyze volumes of text. Some text-as-data tools attempt to replicate the ways humans interact with language, and can parse sentences into their grammatical parts and convert speech to text.⁴⁶ Other text-as-data techniques include sentiment analysis, a process by which researchers identify “sentiments” in text—by, for instance, coding certain words as being in favor of or against a particular position and then classifying speeches based on the prevalence of those words.⁴⁷ For example, some scholars have conducted sentiment analysis on Twitter feeds to predict election outcomes.⁴⁸ In topic modeling, researchers can identify topics and latent themes in an unstructured set of documents.⁴⁹

For all of these tools, greater quantities of data produce greater predictive power and analytic accuracy. Not all data is in a “clean” and therefore easily usable form, however. Some data might not be in

⁴⁴ See Michael A. Livermore et al., *Computationally Assisted Regulatory Participation*, 93 NOTRE DAME L. REV. 977, 996 (2018).

⁴⁵ See, e.g., Matthew Gentzkow et al., *Text as Data*, 57 J. ECON. LITERATURE 535 (2019), <https://web.stanford.edu/~gentzkow/research/text-as-data.pdf> [<https://perma.cc/4ZJE-8FT3>]. Three other scholars usefully frame this concept in the international economic law context, highlighting three characteristics of text-based, data-driven research: “(1) its ‘data first’ attitude, (2) its ambition to look at all the available data rather than subsamples thereof and (3) its focus on computing rather than reading or counting.” Alschner et al., *supra* note 1, at 217, 221–23.

⁴⁶ Simon Brown, *Tips for Computational Text Analysis*, U.C. BERKELEY SOC. SCI. MATRIX (Jan. 26, 2016), <https://matrix.berkeley.edu/research/tips-computational-text-analysis> [<https://perma.cc/3LUB-C39U>]; *What Is Natural Language Processing? Introduction to NLP*, ALGORITHMIA (Aug. 11, 2016), <https://blog.algorithmia.com/introduction-natural-language-processing-nlp/> [<https://perma.cc/8KA3-NBXM>].

⁴⁷ Brown, *supra* note 46.

⁴⁸ See, e.g., Andranik Tumasjan et al., *Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment*, in PROCEEDINGS OF THE FOURTH INTERNATIONAL AAAI CONFERENCE ON BLOGS AND SOCIAL MEDIA 178 (2010). For a discussion of Twitter as a corpus, see Alexander Pak & Patrick Paroubek, *Twitter as a Corpus for Sentiment Analysis and Opinion Mining*, in PROCEEDINGS OF THE SEVENTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION 1320, 1321 (2010), https://www.researchgate.net/publication/220746311_Twitter_as_a_Corpus_for_Sentiment_Analysis_and_Opinion_Mining [<https://perma.cc/YM6W-9LAX>].

⁴⁹ See David M. Blei, *Topic Modeling and Digital Humanities*, 2 J. DIGITAL HUMAN. (2012), <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/> [<https://perma.cc/C95A-HGBF>] (“A topic model . . . discovers a set of ‘topics’—recurring themes that are discussed in the collection—and the degree to which each document exhibits those topics.”).

a machine-readable format, or the diverse data a researcher hopes to study may not be in a standardized form.⁵⁰ Additionally, certain publications containing relevant data may be protected by copyright. As discussed below, these challenges will be particularly relevant to text-as-data work using international law documents.⁵¹

B. *Machine Learning and Computational Text Analysis in Domestic Legal Settings*

There is a wave of academic and popular writing about the influence of machine learning on the domestic legal profession, which one pair of scholars has called a “digital transformation.”⁵² Likewise, legal academics now routinely use computational text analysis to predict the authors of unattributed judicial opinions,⁵³ assess the influence of advocates on those opinions,⁵⁴ and calculate the level of influence of selected opinions on the law.⁵⁵ All indications suggest that the use of technology in law practice and legal scholarship is only likely to grow.⁵⁶

Private law firms, particularly large firms, commonly use machine learning tools for discovery.⁵⁷ But they also are beginning to use these tools to review and generate drafts of contracts, conduct increasingly sophisticated legal research, and attempt to predict case outcomes.⁵⁸

⁵⁰ See Alschner et al., *supra* note 1, at 227; see also Marie Boran, *Making a Case for Artificial Intelligence in the Legal Profession*, IRISH TIMES (June 21, 2018, 4:05 AM), <https://www.irishtimes.com/business/technology/making-a-case-for-artificial-intelligence-in-the-legal-profession-1.3533815> [<https://perma.cc/8FCE-GVW4>] (Gavin Sheridan, chief executive of Vizlegal, a legal tech startup, notes that “[a]t the very simplest level you have to get the information first, get it in the correct format and get it correctly structured before you can do all the nice sexy machine learning stuff” and “[m]ost legal data isn’t published in a way that is user centred.”).

⁵¹ See *infra* Section II.C.

⁵² Salvatore Caserta & Mikael Rask Madsen, *The Legal Profession in the Era of Digital Capitalism: Disruption or New Dawn?*, 8 LAWS 1, 1, 3 (2019).

⁵³ See, e.g., William Li et al., *Using Algorithmic Attribution Techniques to Determine Authorship in Unsigned Judicial Opinions*, 16 STAN. TECH. L. REV. 503, 504 (2013).

⁵⁴ See Pamela C. Corley, *The Supreme Court and Opinion Content: The Influence of Parties’ Briefs*, 61 POL. RES. Q. 468, 469 (2008).

⁵⁵ See Pamela C. Corley et al., *Lower Court Influence on U.S. Supreme Court Opinion Content*, 73 J. POL. 31, 31 (2011).

⁵⁶ Erin Winick, *Lawyer-Bots Are Shaking Up Jobs*, MIT TECH. REV. (Dec. 12, 2017), <https://www.technologyreview.com/s/609556/lawyer-bots-are-shaking-up-jobs/> [<https://perma.cc/S3SQ-RV2L>].

⁵⁷ *Id.* (“So far, AI-powered document discovery tools have had the biggest impact on the field.”).

⁵⁸ John O. McGinnis & Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041, 3043, 3046–55 (2014).

Some software, for instance, is able to review corporate contracts about a designated topic, extract common provisions from those contracts, and create a template for a firm to use.⁵⁹ The system then can compare that template to a new contract and identify discrepancies.⁶⁰ When the system is presented with enough documents of a similar type, it can thus identify outlying provisions in new documents. For legal research, providers such as Westlaw and Lexis now offer subscribers informed predictions on topics such as how long it generally takes a particular judge to rule on a motion, how often a judge grants summary judgment motions, and how often a particular law firm takes a case to a verdict.⁶¹ This allows lawyers to provide their clients with more accurate assessments of how their cases will unfold.

Some computational tools can even displace jobs formerly handled by lawyers. For example, online dispute resolution—conducted directly between disputants—has garnered a number of adherents. The Dutch government hired a company called Modria to use traditional AI tools to create an automated system to help couples divorce with minimal involvement by lawyers.⁶² Modria now offers resolution for debt, landlord-tenant, and small claims too.⁶³

This is not to suggest that private sector lawyers have been at the forefront in adopting machine learning technologies. Indeed, there are a host of cultural reasons why lawyers may be slow to accept and deploy new technologies, including because lawyers emphasize (inherently backward-looking) precedent and are professionally risk averse.⁶⁴ Others have observed that lawyers generally are not skilled

⁵⁹ Remus & Levy, *supra* note 34, at 514.

⁶⁰ *Id.*

⁶¹ Stephen Embry, *Westlaw Edge: AI for Lawyers Goes Mainstream?*, TECHLAW CROSSROADS (July 12, 2018), <https://www.techlawcrossroads.com/2018/07/westlaw-edge-ai-goes-mainstream/> [<https://perma.cc/T444-ALZN>]; *Product Spotlight: Lexis Analytics Suite*, LEXISNEXIS (Apr. 15, 2019), <https://www.lexisnexis.com/community/lexis-legal-advantage/b/product-spotlight/posts/product-spotlight-lexis-analytics-suite> [<https://perma.cc/YSW7-ZMJB>].

⁶² Roey Tzezana, *Artificial Intelligence Tech Will Arrive in Three Waves*, FUTURISM (Mar. 28, 2017), <https://futurism.com/artificial-intelligence-tech-will-arrive-in-three-waves> [<https://perma.cc/UUX3-6EC4>]; see also Rei Morikawa, *Here's How Artificial Intelligence Is Making Divorce Quick and Painless*, MEDIUM (Aug. 13, 2018), <https://becominghuman.ai/heres-how-artificial-intelligence-is-making-divorce-quick-and-painless-a9e16a6b05fc> [<https://perma.cc/WJN5-UR75>] (discussing web application called Wevorce that pursues the same goal).

⁶³ *Modria*, TYLER TECHS. (2017), <https://www.tylertech.com/Portals/0/OpenContent/Files/4080/Modria-Brochure.pdf> [<https://perma.cc/ZU4H-U9Z2>].

⁶⁴ See *Overcoming Lawyers' Resistance to Change*, THOMSON REUTERS, <https://legal.thomsonreuters.com/en/insights/articles/overcoming-lawyers-resistance-to-change> [<https://perma.cc/WYN8-CWQG>] (“The image of the lawyer as a Luddite, rather than an early adopter, persists.”).

at evaluating technology, so they often look for signals that their peers have adopted it.⁶⁵ Further, learning to use new technology takes time away from billable hours.⁶⁶ Finally, law as a field is highly regulated, which complicates the introduction of major changes to traditional ways of doing business.⁶⁷ Nevertheless, as Westlaw and Lexis introduce increasingly sophisticated prediction tools, and as clients pressure private law firms to deploy machine learning to increase speed and reduce costs,⁶⁸ law tech, like AI generally, seems poised to continue its upward trajectory.

Firms are not the only category of legal actors that use machine learning. Using court opinions as their raw material, legal scholars employ computational text analysis, like other scholars whose fields implicate large volumes of text, including historians and political scientists.⁶⁹ Some scholars use stylistic analyses to identify the authorship of unsigned judicial opinions.⁷⁰ Others use plagiarism software to identify linguistic overlaps between a party's brief and a court's opinion.⁷¹ Yet others use classification models built around past U.S. Supreme Court cases to predict the outcomes of future Court cases.⁷² Recently, scholars have begun to apply text-as-data methods to other legally relevant texts, such as using sentiment analysis to characterize public comments on proposed agency rules⁷³ and using topic models

⁶⁵ See Embry, *supra* note 61.

⁶⁶ *Id.*

⁶⁷ See *Overcoming Lawyers' Resistance to Change*, *supra* note 64.

⁶⁸ See DELOITTE, MACHINES WITH PURPOSE 2 (2018), <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/bps/deloitte-uk-machines-with-purpose-ai-report.pdf> [<https://perma.cc/B7BZ-CYUM>].

⁶⁹ See LAW AS DATA, at xiii, xvii (Michael A. Livermore & Daniel N. Rockmore eds., 2019). For discussions of ways in which machine learning helps scholars search through scientific papers and other realms of data, see, e.g., *A Better Way to Search Through Scientific Papers*, ECONOMIST (Oct. 19, 2017), <https://www.economist.com/science-and-technology/2017/10/19/a-better-way-to-search-through-scientific-papers> [<https://perma.cc/UKB3-F9SK>] (describing Semantic Scholar, which uses neural networks to annotate scientific papers and enable researchers to extract from millions of papers the most pertinent papers on a given topic); Sam Kean, *Artificial Intelligence Is Cracking Open the Vatican's Secret Archives*, ATLANTIC (Apr. 30, 2018), <https://www.theatlantic.com/technology/archive/2018/04/vatican-secret-archives-artificial-intelligence/559205/> [<https://perma.cc/KYD5-NYTJ>] (describing project that uses optical character recognition to digitize handwritten texts dating back more than 12 centuries).

⁷⁰ See, e.g., Li et al., *supra* note 53.

⁷¹ See Corley et al., *supra* note 55; Corley, *supra* note 54.

⁷² See Ruger et al., *supra* note 9; Daniel Martin Katz et al., *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12 PLOS ONE (Apr. 12, 2017), <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174698> [<https://perma.cc/8EXM-AZ82>].

⁷³ See, e.g., Livermore et al., *supra* note 44, at 994 (using public comments as big data).

to categorize types of national constitutions.⁷⁴ Taken together, the law firm and scholarly uses of computational tools suggest a range of settings in which technology can contribute to the understanding and analysis of law and the resolution of disputes.

C. *The Roles of a State's International Lawyers*

The foregoing analysis, which focused on the ways in which technologies now inform the practice of private lawyers, invites a question: Are there sufficient similarities between what private lawyers and government international lawyers do to suggest that these technologies may facilitate the latter's work as well?⁷⁵ Answering that question requires an examination of what a government's international lawyers do.

Lawyers who work for foreign ministries perform a wide range of roles. They generally provide legal advice to the politicians and diplomats who craft foreign policy for the state, and they sometimes play an intimate role in defending and publicizing that policy. As lawyers, they are attuned to the importance of legal commitments and the nuances of language, and so often are involved in crafting international agreements between their state and other states or international organizations. Further, policymakers call on them to advise whether a particular action would violate international law, which requires them to have a firm grasp of their state's treaty obligations and the requirements of customary international law.

In the United States, the primary actors who perform these roles are the State Department lawyers who work in the Office of the Legal Adviser (or "L").⁷⁶ Describing their range of tasks, former Legal Adviser Harold Koh stated, "[T]he lawyers in L must negotiate the roles of litigator, counselor, action officer, diplomat, arbitrator, [and] negotiator . . . all at one time."⁷⁷ The United States is not alone in asking a

⁷⁴ See, e.g., David S. Law, *Constitutional Archetypes*, 95 TEX. L. REV. 153 (2016).

⁷⁵ Although this Article focuses on government lawyers who advise on international law, we might well ask the same question of government lawyers who advise on domestic law issues. That analysis, although important, is beyond the scope of this Article.

⁷⁶ There are, of course, lawyers in other agencies who confront international legal issues on a daily basis.

⁷⁷ Harold Hongju Koh, *The State Department Legal Adviser's Office: Eight Decades in Peace and War*, 100 GEO. L.J. 1747, 1757 (2012); see also JEAN-PIERRE GAUCI ET AL., BRITISH INST. OF INT'L AND COMP. LAW, *THE ROLE OF LEGAL ADVISERS IN INTERNATIONAL LAW 3* (2015), https://www.biicl.org/documents/570_the_role_of_legal_advisers_in_international_law_-_conference_report_final26_feb.pdf?showdocument=1 [<https://perma.cc/DTF5-8UV8>] ("Legal advisers both advise in advance of and during negotiations, and lead negotiations themselves, on topics ranging from treaty negotiation to dispute resolution.").

lot of its foreign ministry lawyers: a report from the British Institute of International and Comparative Law, which included input from legal advisers from Australia, Belgium, Canada, France, Singapore, and the United Kingdom, as well as the United Nations, North Atlantic Treaty Organization (“NATO”), and World Trade Organization (“WTO”), notes that in most jurisdictions “the Legal Adviser has three, inter-linked, principal roles: advising, negotiating and litigating.”⁷⁸ This Article focuses on three core tasks reflected in these lists: treaty negotiations, the identification of customary international law, and dispute resolution.

The first critical role for a foreign ministry legal adviser arises when a state wishes to pursue negotiations. In the United Kingdom, a Foreign and Commonwealth Office (“FCO”) lawyer often serves as the head of or as a member of the United Kingdom’s negotiating delegation.⁷⁹ In the United States, the State Department’s Legal Adviser and his staff have been intimately involved in Middle East peace negotiations,⁸⁰ climate change negotiations,⁸¹ the new Strategic Arms Reduction (“New START”) treaty between the United States and Russia,⁸² and the plan that helped guide Kosovo towards independence,⁸³ to name just a few. Negotiations range from large, multilat-

⁷⁸ GAUCI ET AL., *supra* note 77, at 2.

⁷⁹ See OFFICE OF THE LEGAL ADVISER OF THE MINISTRY OF FOREIGN AFFAIRS, COMM. OF LEGAL ADVISERS ON PUB. INT’L LAW, COUNCIL OF EUR., UNITED KINGDOM 1 (2013) [hereinafter U.K. REPORT], https://www.coe.int/t/dlapil/cahdi/Source/office_legal_affairs/United_Kingdom_OLA_2013_EN.pdf [<https://perma.cc/G9T7-6KRX>]; see also Arthur D. Watts, *International Law and International Relations: United Kingdom Practice*, 2 EUR. J. INT’L L. 157, 159 (1991) (noting that FCO lawyers are closely involved in the negotiation of all treaties and regularly attend negotiations at which legal issues arise).

⁸⁰ Koh, *supra* note 77, at 1754 (“Through his work on the Middle East peace process at Camp David, [former State Department Legal Adviser] Herb Hansell became far more than a reactive dispenser of advice. He became a *negotiator of treaties*.”); *id.* at 1757 (“[The Legal Adviser’s Office] helps determine how international agreements should be worded [and] how international organizations should be structured . . .”).

⁸¹ See Megan Darby, *Top US Climate Lawyer Joins UN Foundation from State Department*, CLIMATE CHANGE NEWS (Oct. 5, 2017, 6:39 PM), <http://www.climatechangenews.com/2017/05/10/obamas-climate-lawyer-joins-un-foundation-state-department/> [<https://perma.cc/46XH-GZ3B>] (describing former L lawyer Susan Biniarz as a “key architect of the 2015 Paris climate deal”).

⁸² See Partnership for Public Service, *America’s “Go-to” Legal Adviser for New START*, WASH. POST (Sept. 20, 2011), https://www.washingtonpost.com/politics/americas-go-to-legal-adviser-for-new-start/2011/09/18/gIQAd7g9cK_story.html [<https://perma.cc/E74T-NXLX>] (describing how State Department lawyer Paul Dean “participated in the bilateral negotiations, helped draft the treaty . . . and then worked to ensure final Russian approval”).

⁸³ HENRY H. PERRITT, JR., *THE ROAD TO INDEPENDENCE FOR KOSOVO* 130 (2010) (noting that State Department lawyer Anna Mansfield “regularly participated in providing legal support to” the Kosovo negotiations).

eral negotiations that continue for several years to bilateral negotiations on issues that range from the commonplace to the obscure. Some treaties have dozens of articles and about 190 parties,⁸⁴ while others have two or three articles and two parties.

A second important role for international legal advisers is in dispute resolution. Dispute resolution encompasses processes both formal and informal, multilateral and bilateral, in both judicial and arbitral fora. Former Legal Adviser Koh identified one of his roles as “defender of the United States interests in the many international fora in which the U.S. appears,” including the International Court of Justice (“ICJ”), North American Free Trade Agreement arbitral tribunals, and the Iran-U.S. Claims Tribunal.⁸⁵ In the United Kingdom, lawyers from the FCO act as agents for their governments before international tribunals such as the ICJ and European Court of Human Rights, and in interstate arbitrations.⁸⁶ International lawyers prepare legal submissions for these fora and appear before them as advocates. If their state has suffered an internationally wrongful act by another state, they will also advise their government about whether to pursue a claim against the wrongdoing state (based on past practice and the outcomes of prior comparable claims), as well as the forum in which to pursue that claim.⁸⁷

There is interplay, of course, between dispute resolution and negotiations, because one way of ending an international dispute is to conclude an international agreement that addresses the underlying problem, whether by providing compensation, allocating resources among the parties, or establishing new rules of acceptable behavior. For instance, the U.S. dispute with China over NATO’s mistaken bombing of the Chinese embassy in Belgrade required both a resolution of the dispute and, as part of that resolution, the negotiation of an international agreement pursuant to which the United States paid China \$4.5 million for the Chinese victims killed in the bombing.⁸⁸ In

⁸⁴ See, e.g., U.N. Charter (111 articles; 193 parties); Geneva Convention Relative to the Treatment of Prisoners of War, Aug. 12, 1949, 75 U.N.T.S. 135 (143 articles; 196 parties); U.N. Convention Against Corruption, Oct. 31, 2003, 2349 U.N.T.S. 41 (71 articles; 186 parties).

⁸⁵ Harold Hongju Koh, Legal Adviser, U.S. Dep’t of State, Speech at the Annual Meeting of the American Society of International Law: The Obama Administration and International Law (Mar. 25, 2010); Koh, *supra* note 77, at 1755 (describing Legal Adviser Davis Robinson’s role appearing before the ICJ).

⁸⁶ U.K. REPORT, *supra* note 79.

⁸⁷ GAUCI ET AL., *supra* note 77, at 3 (noting that legal advisers must advise on litigation tactics, the strength of arguments, and what tone to use to balance legal and diplomatic considerations).

⁸⁸ KERRY DUMBAUGH, CONG. RESEARCH SERV., RS20547, CHINESE EMBASSY BOMBING

situations such as those, lawyers can be particularly useful in suggesting options, reframing the discussions when they stall, and memorializing the substance of the parties' agreement.

A third role for foreign ministry lawyers is to advise policymakers about the existence and meaning of international law. This includes informing them about the contents of treaties to which their states or their negotiating partners are parties and the relevance of those treaties to contemplated policy acts, but it also includes advising about the existence and contents of rules of customary international law ("CIL").⁸⁹ Because CIL develops in a more inchoate way than treaties do and is not always formally written down by states, questions often arise about the precise contents and parameters of the rules.⁹⁰ In the United States, State Department lawyers help determine "how customary international law rules should be articulated."⁹¹ As part of this process, those lawyers need to identify relevant state practice and *opinio juris* that might strengthen or undercut rules that they or other states view as existing custom.⁹²

All of these roles require resources—most notably personnel, time, money, and access to legal documents. While the United States has approximately 250 lawyers in the State Department,⁹³ most countries have far fewer foreign ministry lawyers to perform these roles. Mexico, for example, had 35 lawyers in its Office of the Legal Adviser within the Ministry of Foreign Affairs as of 2017.⁹⁴ Recent reports in-

IN BELGRADE COMPENSATION ISSUES 4–5 (2000), <https://www.everycrsreport.com/reports/RS20547.html> [<https://perma.cc/BF2D-Q3NC>] (describing the negotiated result as an "agreement" that specified "how to divide the funds among the bombing victims and their families" and noting that the Legal Adviser later reached separate property damage agreements whereby the United States paid China \$28 million for damage to the Chinese Embassy and China paid \$2.87 million to the United States for damage to U.S. diplomatic and consular facilities as a result of Chinese protests); Seth Faison, *U.S. to Pay China for Embassy Bombing*, N.Y. TIMES (July 31, 1999), <https://www.nytimes.com/1999/07/31/world/us-to-pay-china-for-embassy-bombing.html> [<https://perma.cc/7QJ9-VTX7>] (stating that Legal Adviser David Andrews led the U.S. side in the negotiations).

⁸⁹ See U.K. REPORT, *supra* note 79 (stating that U.K. FCO lawyers advise about public international law).

⁹⁰ See RESTATEMENT (THIRD) OF THE FOREIGN RELATIONS LAW OF THE UNITED STATES § 102 cmt. c (AM. LAW. INST. 1987) (noting that it is often hard to determine when a rule that states follow out of habit transforms into law).

⁹¹ Koh, *supra* note 77, at 1757.

⁹² See CURTIS A. BRADLEY ET AL., FOREIGN RELATIONS LAW, at xxv (7th ed. 2020) (describing customary international law as forming "after nations have consistently followed a particular practice out of a sense of legal obligation," an obligation also known as "*opinio juris*").

⁹³ Stephen Bouwhuis, *The Role of an International Legal Adviser to Government*, 61 INT'L & COMP. L.Q. 939, 958 (2012).

⁹⁴ COMMITTEE OF LEGAL ADVISERS ON PUBLIC INTERNATIONAL LAW (CAHDI)

dicate that Norway has 30 lawyers in its Legal Affairs Department;⁹⁵ Israel has 27;⁹⁶ the United Kingdom about 30;⁹⁷ and Switzerland 45.⁹⁸ This is true even though the number of treaties continues to climb and disputes under bilateral investment treaties are increasing over time.⁹⁹ Thus, the volume of information available to foreign ministry lawyers and the amount of information they must process is climbing at a much faster rate than the numbers of lawyers.

In sum, foreign ministry lawyers are tasked with a host of complicated challenges but are given relatively limited resources to achieve their goals. One important question, then, is whether there are technology-driven tools that can help those lawyers fulfill their core missions, and, if so, whether international lawyers have incentives to adopt them.

D. Will Tech Come to International Law?

International lawyering has traditionally been a conservative profession. Dutch jurist Bart R ling, who served as a judge in the Tokyo trials, refused to teach international law because he viewed it as “monotonous and conservative.”¹⁰⁰ Attempting to explain the phenomenon, one scholar noted the following:

By focusing on the past, international law allows previous generations to rule over present ones, and this makes it difficult for powerful actors to remake the international legal or-

DATABASE, THE ORGANISATION AND FUNCTIONS OF THE OFFICE OF THE LEGAL ADVISER IN THE MINISTRY OF FOREIGN AFFAIRS, COUNCIL OF EUR. (2017), <http://www.cahdidatabases.coe.int/Contribution/Details/5> [https://perma.cc/BF2D-Q3NC] (Mexico).

⁹⁵ CAHDI DATABASE, THE ORGANISATION AND FUNCTIONS OF THE OFFICE OF THE LEGAL ADVISER IN THE MINISTRY OF FOREIGN AFFAIRS, COUNCIL OF EUR. (2015), <http://www.cahdidatabases.coe.int/Contribution/Details/16> [https://perma.cc/WV5K-V972] (Norway).

⁹⁶ CAHDI DATABASE, THE ORGANISATION AND FUNCTIONS OF THE OFFICE OF THE LEGAL ADVISER IN THE MINISTRY OF FOREIGN AFFAIRS, COUNCIL OF EUR. (2015) <http://www.cahdidatabases.coe.int/Contribution/Details/9> [https://perma.cc/4SM8-PLJM] (Israel).

⁹⁷ Bouwhuis, *supra* note 93, at 959 (stating that the Office of the Legal Adviser in the FCO consists of about 30 lawyers); Foreign Affairs Comm., U.K. Parliament, *Written Evidence from Sir Michael Wood KCMG*, U.K. PARLIAMENT (Jan. 4, 2011), <https://publications.parliament.uk/pa/cm201012/cmselect/cmfa/665/665vw18.htm> [https://perma.cc/4SZD-HBD6].

⁹⁸ CAHDI DATABASE, THE ORGANISATION AND FUNCTIONS OF THE OFFICE OF THE LEGAL ADVISER IN THE MINISTRY OF FOREIGN AFFAIRS, COUNCIL OF EUR. (2015), <http://www.cahdidatabases.coe.int/Contribution/Details/13> [https://perma.cc/N7JE-YJVF] (Switzerland).

⁹⁹ See Malcolm Langford et al., *The Revolving Door in International Investment Arbitration*, 20 J. INT’L ECON. L. 301, 307 fig.1 (2017) (showing number of international investment arbitration cases registered by year from 1987 to 2016).

¹⁰⁰ James A.R. Nafziger, Book Review, 90 AM. J. INT’L L. 342, 344 (1996) (reviewing B.V.A. ROLING & ANTONIO CASSESE, *THE TOKYO TRIAL AND BEYOND* (1993)).

der according to their own vision. In international law, this problem is especially acute because changes in international law require widespread consent and are usually slow and incremental.¹⁰¹

Perhaps this conservatism—and a sense that progress happens slowly by design—helps explain why international lawyers have not turned to technology to help them innovate how they perform their work.¹⁰²

Regardless of the reason, international lawyers must start to understand how these tools may, and in some cases will, affect their practice. The pressure to understand—if not to use—these tools will come in several forms, some external to the state and some internal.

1. *Near-Peer Ambitions*

First, international lawyers for governments may face pressure to adopt legal technologies because their foreign counterparts have begun to use them. States such as China are likely to be at the leading edge of applying machine learning and text-as-data tools to international law settings. China's judges already are piloting an AI-assisted system that can conduct deviation analyses, allowing the judges to compare how different their draft judgments are from past precedent.¹⁰³ China's appetite for all things AI and its access to vast amounts of data suggest that it will try to develop as many machine learning tools as possible to give it an advantage in international rela-

¹⁰¹ Nico Krisch, *International Law in Times of Hegemony: Unequal Power and the Shaping of the International Legal Order*, 16 EUR. J. INT'L L. 369, 377 (2005) ("Like all law, international law is marked by a focus on the past: most of its sources refer to historical events, and in fact, many of its central elements have remained the same for a long time. . . . [This] stability of international law also poses significant problems for dominant states."); see also Joel P. Trachtman, *Welcome to Cosmopolis, World of Boundless Opportunity*, 39 CORNELL INT'L L.J. 477, 479 (2006) ("Law in general, and international law in particular, might be viewed as a fundamentally conservative discipline, embedding the distributive status quo in its concerns for the rule of law and in its principles about the role of the state and how law is made."); William Thomas Worster, *The Effect of Leaked Information on the Rules of International Law*, 28 AM. U. INT'L L. REV. 443, 444 (2013) ("International law, and international lawyers generally, tend to take a somewhat conservative approach to the formation of rules.").

¹⁰² One recent report provides a helpful overview of AI as a topic for diplomacy, a tool for diplomacy, and a tool that can impact the environment for diplomacy. See DIPLOFOUNDATION, *MAPPING THE CHALLENGES AND OPPORTUNITIES OF ARTIFICIAL INTELLIGENCE FOR THE CONDUCT OF DIPLOMACY* (2019), <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [<https://perma.cc/RZ67-ZQ6Y>].

¹⁰³ See Alfred Chua, *The Big Read: Rise of the Machine—How Technology Is Disrupting Singapore's Law Firms*, CHANNEL NEWS ASIA (Jan. 21, 2019, 6:23 PM), <https://www.channelnewsasia.com/news/singapore/lawyers-law-sector-disrupted-technology-artificial-intelligence-11145088> [<https://perma.cc/EQT8-TA6L>].

tions.¹⁰⁴ Russia, too, has successfully developed image and speech recognition technologies.¹⁰⁵ Other international lawyers and diplomats for states with bold technological ambitions, including the United States, may find themselves encouraged to use a range of technologies during negotiations and dispute resolution.¹⁰⁶

Some states undoubtedly lack such technological ambitions, or lack the resources to pursue such ambitions. But even international lawyers for these states have an obligation to their clients to understand the tools that other states may use against them in these settings. It will be imperative that these lawyers have a basic sense of how the tools work, what they can achieve, and how they might advantage their users and disadvantage opponents.

2. *Proofs of Concept in Private Law*

Another reason that some international lawyers will feel external pressure to explore legal technologies is that the private sector has found these tools valuable. Of course, the jobs of private firm lawyers and international lawyers for the government are quite different. Government international lawyers rarely confront discovery requests, for example, and do not bill their time. But there also are important similarities between what the two sets of lawyers do.¹⁰⁷ Most obviously, both conduct legal research using vast bodies of text. Both help their clients negotiate agreements, sometimes using model texts.¹⁰⁸ Both benefit from *ex ante* knowledge about their opponents and adjudicators, and from being able to predict what those actors are likely to do in given situations. Private firm lawyers use machine learning tools for all of those tasks,¹⁰⁹ so there are existing models from which international lawyers can draw.

¹⁰⁴ See Yujia He & Anne Bowser, *How China Is Preparing for an AI-Powered Future* (2017) <https://www.wilsoncenter.org/publication/how-china-preparing-for-ai-powered-future> [<https://perma.cc/TV96-AF5Q>].

¹⁰⁵ See Samuel Bendett, *Putin Orders Up a National AI Strategy*, DEF. ONE (Jan. 31, 2019), <https://www.defenseone.com/technology/2019/01/putin-orders-national-ai-strategy/154555/> [<https://perma.cc/L6K6-A37V>].

¹⁰⁶ See, e.g., *Accelerating America's Leadership in Artificial Intelligence*, WHITE HOUSE (Feb. 11, 2019), <https://www.whitehouse.gov/articles/accelerating-americas-leadership-in-artificial-intelligence/> [<https://perma.cc/6B9B-Z3E3>].

¹⁰⁷ For a basic description of what attorneys in the State Department's Office of the Legal Adviser do, see *About Us – Legal Adviser*, U.S. DEP'T OF STATE, <https://www.state.gov/about-us-legal-adviser/> [<https://perma.cc/K4VB-R5NE>].

¹⁰⁸ See *id.*

¹⁰⁹ See Remus & Levy, *supra* note 34, at 508–31.

Further, the work of some private law firms and government international lawyers overlaps directly on some issues, as with investor-state arbitration. If and when government lawyers face off against private firms using these tools, they may be forced to explore their use to avoid ceding advantages to their private sector opponents.

3. *Client Pressure*

Another source of internal pressure to explore international legal technologies may come from an international lawyer's clients. Senior officials in foreign ministries may have spent time in the private sector or may have regular exposure to business leaders who are deploying machine learning technology in their own industries.¹¹⁰ Just as private firm clients sometimes push law firms to deploy these technologies to obtain a competitive advantage or save money, so too might government policymakers push their lawyers to ensure that they are abreast of the latest technologies that could help a state achieve its foreign policy goals. Indeed, technology firms may begin to market certain products to the government once they identify ways in which the government could benefit from their tools.

4. *Necessity*

Finally, necessity is often the mother of invention. Some international lawyers from states with few resources will instinctively turn to some of the tools discussed in Parts II and III because they are publicly available and have helped those lawyers in other parts of their life. Google Translate, for example, is widely used. An international lawyer who only speaks an uncommon language may frequently find herself using translation tools when she travels, and so intuitively will use such tools at international negotiations as well. Likewise, government international lawyers who work for understaffed offices may resort out of necessity to existing machine learning technologies to speed up certain tasks.

One key reason to think that international legal technology has a bright future is that there is a vast range of data to undergird it. As this Article discusses in greater detail herein, there are a variety of

¹¹⁰ For instance, before entering the State Department, former Secretary of State Rex Tillerson had been the Chief Executive Officer of Exxon Mobil, Michael D. Shear & Maggie Haberman, *Rex Tillerson, Exxon C.E.O., Chosen as Secretary of State*, N.Y. TIMES (Dec. 12, 2016), <https://www.nytimes.com/2016/12/12/us/politics/rex-tillerson-secretary-of-state-trump.html> [<https://perma.cc/SD5Q-Z4AX>], and former Deputy Secretary of State John Sullivan had been a partner at Mayer Brown, *John J. Sullivan: Deputy Secretary of State*, U.S. DEP'T OF STATE, <https://www.state.gov/biographies/john-j-sullivan/> [<https://perma.cc/6444-DLVW>].

digital sources of text that might serve as the basis for the kinds of text-as-data analyses that will be useful to states.¹¹¹ This includes U.N. databases of Security Council and General Assembly documents,¹¹² collections of treaties and their travaux préparatoires (which are the official records of negotiations),¹¹³ European Court of Human Rights caselaw,¹¹⁴ international arbitral awards, databases of specialized agencies such as the International Civil Aviation Organization,¹¹⁵ state archives and digests, data collected by a state's own intelligence agencies and diplomats (memorialized in internal memoranda and cables), states' notifications to the Security Council about actions taken in self-defense,¹¹⁶ legal blogs, the U.N. Yearbook, reports by and submission to U.N. human rights bodies,¹¹⁷ news reports,¹¹⁸ and databases of foreign statutes.¹¹⁹ Each of these collections contains thousands of documents, which—on the one hand—makes it difficult for international lawyers to process all of the information and—on the other hand—provides the type of “big data” that makes text-as-data tools effective and efficient.

111 See *infra* Section II.A.1.

112 See *infra* text accompanying note 137.

113 See U.N. TREATY COLLECTION, <https://treaties.un.org> [<https://perma.cc/5DHL-S4DX>]; LANGUAGE OF PEACE, <https://www.languageofpeace.org> [<https://perma.cc/Y6CA-2JHX>] (collection of around 1,000 peace treaties); *UN Human Rights Treaties: Travaux Préparatoires*, U. VA., <https://hr-travaux.law.virginia.edu> [<https://perma.cc/DF8Y-Q83S>] (collection of travaux préparatoires of human rights treaties).

114 See *Judgments and Decisions*, HUDOC, <https://echr.coe.int/Pages/home.aspx?p=case-law/HUDOC&c=> [<https://perma.cc/75YG-BM83>].

115 See *Economic Development: Databases and Studies*, INT'L CIVIL AVIATION ORG., <https://www.icao.int/sustainability/Pages/Databases.aspx> [<https://perma.cc/VTT9-6RLV>].

116 See Dustin A. Lewis et al., *HLS PILAC Catalogue of Communications to the Security Council of Measures Taken by United Nations Member States in Purported Exercise of the Right of Self-Defense: October 24, 1945 Through December 31, 2018*, in *QUANTUM OF SILENCE* 82 (2019), <http://blogs.harvard.edu/pilac/files/2019/07/Annex-with-Catalogue.pdf> [<https://perma.cc/2PTV-2JYM>].

117 See *Human Rights Bodies*, OFFICE OF THE HIGH COMM'R, U.N. HUMAN RIGHTS, <https://www.ohchr.org/en/hrbodies/Pages/HumanRightsBodies.aspx> [<https://perma.cc/F6JD-F3JY>].

118 See Gary King & Will Lowe, *An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design*, 57 INT'L ORG. 617, 619 (2003) (using software to extract events from Reuters news reports and finding that the program was able to extract those events at a level equal to trained Harvard undergraduates).

119 Historically, states published collections of statutes, including nationality laws, diplomatic and consular laws, neutrality laws, and piracy laws. See U.N. Secretary-General, *Ways and Means of Making the Evidence of Customary International Law More Readily Available*, U.N. Doc. A/CN.4/6, at 71–75 (1949) [hereinafter ILC Report], http://legal.un.org/ilc/documentation/english/a_cn4_6.pdf [<https://perma.cc/9FV7-AYBU>]. For contemporary resources, see 1 THOMAS H. REYNOLDS & ARTURO A. FLORES, *FOREIGN LAW* (2007); *Foreign Law Guide*, BRILL, <https://referenceworks.brillonline.com/browse/foreign-law-guide> [<https://perma.cc/TTB6-M699>].

5. Challenges

There are a number of reasons to think that existing legal text-as-data tools can both undergird and stimulate the production of new tools that lawyers could apply to international law problems.¹²⁰ However, this is not to argue that this transition will be easy or uncontested. First, the current format of much international law data is likely to pose technical challenges for researchers seeking to develop machine learning tools for international law. Only a small part of relevant documents is in databases such as Westlaw.¹²¹ Second, relevant legal data frequently will be in a range of languages, not only English, and is produced by a large number of countries, not just the United States. Third, the types of documents that contain relevant information will less often be judicial opinions and will more often be lodged in a range of document types, such as Security Council resolutions or domestic executive branch speeches. Fourth, any given international law question will have less precedent to guide predictions than domestic law questions have.

Fifth, as a practical matter, governments tend to be less nimble in acquiring new technologies. They face statutory or regulatory hurdles, often have to allow multiple bidders to compete for contracts, and may struggle with budgetary issues.¹²² In the late 1990s, the U.S. Federal Bureau of Investigation was still using 1980s technology,¹²³ and a recent GAO report found that some federal agencies were using

¹²⁰ See *supra* Section I.D.

¹²¹ Westlaw's database of Department of State materials includes U.S. international agreements and other U.S.-related materials but not U.N. General Assembly or Security Council resolutions or the travaux préparatoires of treaties. See WESTLAW EDGE, <http://www.westlaw.com> (last visited Apr. 10, 2020) (click the "Federal Materials" tab; then click "Federal Administrative Decisions & Guidance"; then click "Department of State"). Westlaw's International Materials database includes materials from select jurisdictions not including the U.N. See WESTLAW EDGE, <http://www.westlaw.com> (last visited Apr. 10, 2020) (click "International Materials" to see list of jurisdictions). See also Lyonette Louis-Jacques, *Gaps in International Legal Literature: A Skeptical Reappraisal*, 35 SYRACUSE J. INT'L L. & COM. 363, 365 (2008) ("[T]he electronic landscape for international legal literature remains unstable and uncertain. Not every core international research resource is available online or in an acceptable format. No centralized repository preserves electronic sources of international law. International legal information is scattered . . ."); Marylin Raisch, *Travaux Préparatoires and the United Nations Treaties or Conventions: Using the Web Wisely: Research Tips and Observations*, 30 INT'L J. LEGAL INFO. 324, 324 (2002) ("[I]t is not only the obscure or bilateral treaties that can be hard to . . . locate. *Travaux* for larger conventions may be a challenge as well.").

¹²² See, e.g., 41 U.S.C. § 3301 (2018) (requiring that executive agencies conduct procurements using full and open competition).

¹²³ *About: Information Technology*, U.S. FED. BUREAU INVESTIGATION, <https://www.fbi.gov/about/leadership-and-structure/information-technology> [<https://perma.cc/H5YT-YTLT>].

software and hardware that was at least 50 years old.¹²⁴ The U.S. government also struggles to hire a sufficient number of technology workers because it must compete with higher paying private companies.¹²⁵

Sixth, democratic governments such as those in the United States, Canada, Australia, and European states will confront and must address civil liberties concerns that might arise when they develop tools such as emotion detection software, which must be trained on many thousands of individual faces, and social media scraping, which involves processing posts by millions of citizens. Even if this scraping largely will be directed at posts by foreign citizens, it may stimulate privacy concerns akin to those that arose in the foreign surveillance context after Edward Snowden's leaks.¹²⁶ Finally, government legal offices generally suffer from an "inbox" problem: they are so busy trying to respond to urgent work issues that they find it difficult to prioritize longer-term projects and planning.¹²⁷

These challenges, although real, do not seem insurmountable, at least with regard to a subset of the tools discussed in this Article. For example, some of these tools either are publicly available or derive from open source code.¹²⁸ In some cases, outside actors who are interested in open data or foreign policy issues may decide to voluntarily provide international lawyers with some of these tools.¹²⁹ Further, some of the tools will not implicate individual privacy at all. The next

¹²⁴ U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-16-696T, INFORMATION TECHNOLOGY (2016), <https://www.gao.gov/products/GAO-16-696T> [<https://perma.cc/ZE6W-4UGZ>].

¹²⁵ Jack Corrigan, *The Government's Struggle to Hire Young Tech Talent Is Worse than You Thought*, NEXTGOV (Dec. 1, 2017), <https://www.nextgov.com/cio-briefing/2017/12/governments-struggle-hire-young-tech-talent-worse-you-thought/144225/> [<https://perma.cc/K445-2Q9X>].

¹²⁶ See Ashley Deeks, *An International Legal Framework for Surveillance*, 55 VA. J. INT'L L. 291, 293 (2015). For an example of private sector concerns about government processing of citizens' social media posts, see Sophia Cope, *State Dept. Wants to Expand Social Media Collection to All Visa Applicants*, ELECTRONIC FRONTIER FOUND. (Apr. 3, 2018), <https://www.eff.org/deeplinks/2018/04/state-dept-wants-expand-social-media-collection-all-visa-applicants> [<https://perma.cc/9KZR-BSX9>].

¹²⁷ See, e.g., NAT'L INTELLIGENCE COUNCIL, GLOBAL TRENDS, at vi (2017), <https://www.dni.gov/files/documents/nic/GT-Full-Report.pdf> [<https://perma.cc/6EAU-F4JQ>] ("Thinking about the future is vital but hard. Crises keep intruding, making it all but impossible to look beyond daily headlines to what lies over the horizon. In those circumstances, thinking 'outside the box,' to use the cliché, too often loses out to keeping up with the inbox.")

¹²⁸ See generally Bogdan Batrinca & Philip C. Treleaven, *Social Media Analytics: A Survey of Techniques, Tools and Platforms*, 30 AI & SOC'Y 89, 94–95 (2015) (reviewing large number of data analytics tools available to perform social media scraping and analytics).

¹²⁹ For an example of a non-profit that works with governments "to build an open, trustworthy data ecosystem, where people can make better decisions using data," see OPEN DATA INSTITUTE, <https://theodi.org> [<https://perma.cc/SW4B-VWC3>].

two Parts take a deeper dive into specific ways text-as-data tools could change the international legal landscape.

II. CREATING AND IDENTIFYING INTERNATIONAL LAW

One key role for international lawyers is to create and identify international law, and a range of new technologies can facilitate these processes. There are two primary sources of international law: treaties and CIL. States create treaties when two or more states come together to agree on a set of binding provisions, governed by international law, that commit them to take or abstain from taking certain future actions.¹³⁰ CIL results from the widespread and uniform practice of states (state practice), undertaken out of a sense of legal obligation (*opinio juris*), rather than just habit or politeness.¹³¹

Text-as-data tools have the potential to affect both sources of international law. In the treaty-making context, these tools offer a fertile way for states participating in negotiations to identify their negotiating partners' preferences, to evaluate the likelihood that any given proposal will end up in the final version of the treaty, and to process domestic intelligence that informs the negotiators. Other types of machine learning tools might assist in the conduct of negotiations themselves, for translation purposes, and possibly even to evaluate the behavior of one's negotiating partners. In the CIL context, tools such as natural language processing and topic modeling could allow states to identify previously unknown state practice and *opinio juris* within vast state and U.N. archives and contemporary online data.

A. *Preparing for Treaty Negotiations*

The best treaty negotiators come to the table well prepared. They know what their own goals are, but they also know what their negotiating partners want, are likely to propose, and to what they are, at the end, likely to agree.¹³² They also adjust their approaches during negotiations by taking into account developments that happen across the negotiating table.¹³³ Further, they proffer creative language when the

¹³⁰ Vienna Convention on the Law of Treaties art. 2(1)(a), May 23, 1969, 1155 U.N.T.S. 331.

¹³¹ See BRADLEY, *supra* note 92, at xxv.

¹³² See John Sae, *Best Practice in Global Negotiation Strategies for Leaders and Managers in the 21st Century*, 9 J. BUS. ECON. & MGMT. 309, 314 (2008).

¹³³ See *id.*

negotiations come to an impasse, a skill at which international lawyers can be particularly adept.¹³⁴

In light of machine learning's ability to sort through millions of possible options and identify the likely probabilities of particular outcomes, will states begin to rely on machine learning in the treaty negotiation context? If so, what might that look like? This Section argues that there are a range of ways in which states may begin to rely on these tools. Specifically, states might use machine learning tools before the negotiations start to identify their negotiating partners' historical perceptions of their equities, predict possible outcomes, identify foreign law that is relevant to the topic being negotiated, and process domestic intelligence to prepare the negotiators.

1. Identifying Negotiating Partner Equities

In negotiations, it is critical to understand what provisions one's negotiating partner is more or less likely to care about.¹³⁵ One way to identify what State X will care about in a particular negotiation is to examine what that state has cared about in hundreds of past negotiations (or in discussions in interstate settings).¹³⁶ There is at least one very large database that contains information on which data scientists could train algorithms to discover some of these sensitivities. The United Nations maintains a database of full-text, born-digital documents created since 1993, including documents of the Security Council, the General Assembly, and the Economic and Social Council.¹³⁷ It also contains scanned resolutions and records from the Security Council and General Assembly produced between 1946–1993.¹³⁸ The database also contains voting data, statements made by states before the principal organs, states' periodic reports submitted to human

¹³⁴ See Susan Biniaz, *Comma but Differentiated Responsibilities: Punctuation and 30 Other Ways Negotiators Have Resolved Issues in the International Climate Change Regime*, 6 MICH. J. ENVTL. & ADMIN. L. 37, 61–63 (2016).

¹³⁵ Conducting thorough background research on the other parties to a negotiation is a common practice. See Peter Reilly, *Was Machiavelli Right? Lying in Negotiation and the Art of Defensive Self-Help*, 24 OHIO ST. J. DISP. RESOL. 481, 525–26 (2009).

¹³⁶ See Wolfgang Alschner et al., *Text-as-Data Analysis of Preferential Trade Agreements: Mapping the PTA Landscape*, U.N. DOC. UNCTAD/SER/RP/2017/5, at 4 (July 2017) (using 447 Preferential Trade Agreements and unsupervised machine learning tools to automatically cluster treaties into groups of similar agreements).

¹³⁷ U.N., DIGITAL LIBRARY, <https://library.un.org> [<https://perma.cc/W5JM-BZV8>].

¹³⁸ U.N., OFFICIAL DOCUMENT SYSTEM, <https://documents.un.org/prod/ods.nsf/home.xsp> [<https://perma.cc/KS4K-EPZD>]. The optical quality of these born-analog documents may be low, given their age.

rights bodies, and states' sponsorship of various resolutions.¹³⁹ Further, the United Nations hosts a "diplomatic conferences" website that contains many official documents surrounding the negotiation of certain treaties.¹⁴⁰ These databases contain millions of documents on a wide range of subjects and thus offer a wealth of data that is useful to treaty negotiators. The travaux préparatoires of other treaties will also be informative and are usually publicly available, though they may be more difficult to locate.¹⁴¹

Using these documents, a state could identify topics that particularly occupy the attention of its negotiating partners. For example, imagine using text-as-data tools that identify the frequency distribution of topics within U.N. documents.¹⁴² Assume that the United States is engaged in multilateral negotiations with several states, including Kuwait and India. The United Nations databases contain a host of documents that reflect Kuwait's interventions in various U.N. fora over the past three decades. Applying text-as-data tools to those documents, the United States might identify that Kuwait is keenly and consistently focused on the concept of border integrity and the preservation of territorial boundaries, and that India tends to raise the importance of disarmament in nearly every setting. The United States could use this knowledge to help structure its proposals in a way that is likely to convince Kuwait and India to agree to certain treaty language.

A negotiating state could also explore whether particular sets of states tend to operate in a coordinated way based on the language that the states use in their interventions during treaty negotiations.¹⁴³ These tools could help detect which sets of states repeatedly make similar points during those interventions across a wide range of negotiations, thus helping negotiators better understand the dynamics in the negotiating room and suggesting that persuading State X to accept certain language might influence the position of other states with which State X is coordinating.

¹³⁹ See U.N., DIGITAL LIBRARY, MEMBER STATES ON THE RECORD, <https://library.un.org/unms> [<https://perma.cc/BLU2-F9EZ>].

¹⁴⁰ See U.N., CODIFICATION DIVISION PUBLICATIONS, DIPLOMATIC CONFERENCES, <http://legal.un.org/diplomaticconferences/> [<https://perma.cc/TS43-4JYC>].

¹⁴¹ See Alec Sieber, *Law Library Creates Database of UN Human Rights Documents*, U. VA. (July 11, 2017), <https://www.law.virginia.edu/news/201707/law-library-creates-database-un-human-rights-documents> [<https://perma.cc/5Y8H-N5XX>].

¹⁴² For a recent study applying unstructured topic modeling to Security Council resolutions, see Richard Hanaia, *Explaining the Development of International Norms: The Humanitarian Turn at the United Nations Security Council* (June 28, 2019) (unpublished manuscript), <http://dx.doi.org/10.2139/ssrn.3199611> [<https://perma.cc/E8UD-5CXU>].

¹⁴³ See *id.*

To produce such information for states, data scientists would need to deploy a range of tools. First, they might use topic models, which extract general topics from within documents.¹⁴⁴ These models could flag for negotiators which states tend to focus on particular subjects, creating a list of words used in states' submissions and a list of which states are using those words. The tools can also identify the frequency distributions of topics. Of course, the frequency with which a state uses particular words may be a function of the types of negotiations and conversations in which it has participated in the past—but that, too, will be helpful information to an international lawyer preparing for negotiations.

Although this tool would allow international lawyers to easily identify which documents submitted by Kuwait discussed international borders, sovereignty, water rights, or refugees, it would not allow those lawyers to determine whether Kuwait is talking about refugees or sovereignty in a positive or negative way. Thus, negotiators might seek to use a second tool—sentiment analysis—which would help identify how Kuwait was framing those topics. Coders would need to identify in advance which words diplomats usually use to indicate support, criticism, or concern, though that seems like a manageable task.¹⁴⁵ Alternatively, programmers might be able to use N-grams (a tool that treats multi-word phrases as units and assigns probabilities to sentences and word sequences)¹⁴⁶ or hand code some U.N. documents and train a machine learning algorithm to extract similar language found in other documents in that database.

Third, to evaluate the influence of a state on other states' written submissions and proposals, states could employ algorithms that would allow them to study references between and within documents.¹⁴⁷ At the very least, these processes could identify for negotiators the most relevant subset of documents that they should review by hand to un-

¹⁴⁴ See *id.* at 15–16.

¹⁴⁵ One question would be whether to conduct this analysis in English or Arabic; the U.N. produces official documents in both languages. See *Official Languages*, U.N., <https://www.un.org/en/sections/about-un/official-languages/> [<https://perma.cc/V339-MSG5>].

¹⁴⁶ Daniel Jurafsky & James H. Martin, *Speech and Language Processing: N-gram Language Models* (Oct. 2, 2019) (unpublished manuscript) (manuscript at 2), <https://web.stanford.edu/~jurafsky/slp3/3.pdf> [<https://perma.cc/NVG5-KBDB>].

¹⁴⁷ See, e.g., James H. Fowler & Sangick Jeon, *The Authority of Supreme Court Precedent*, 30 *SOC. NETWORKS* 16 (2008); Sergio Puig, *Social Capital in the Arbitration Market*, 25 *EUR. J. INT'L L.* 387 (2014) (using network analysis to identify “power brokers” who dominate the arbitration profession); Thomas A. Smith, *The Web of Law*, 44 *SAN DIEGO L. REV.* 309 (2007) (analyzing legal systems using network theory).

derstand another state's positions and to assess whether certain states are coordinating their positions in lockstep.

Moving beyond documents that reflect the views of the states themselves, negotiators might also want to use web scraping tools and sentiment analysis to identify and gauge the sentiments of states' local populations on the issues that the negotiations will address.¹⁴⁸ Local sentiment will affect what kinds of treaty terms a government could accept during negotiations, so knowledge of local sentiment will help negotiators anticipate roadblocks and develop workarounds.

2. *Predicting Possible Outcomes*

One way that text-as-data tools could inform negotiators as they approach negotiations is to help predict the language of the treaty to be negotiated, based on the language of similar existing treaties. Wolfgang Alschner and Dmitry Skougarevskiy have developed such a tool, using the content of existing bilateral investment treaties ("BITs") to predict the content of a new BIT that the United States and China are negotiating.¹⁴⁹ They also have developed an algorithm that drafts international investment agreements. They trained the algorithm on individual treaty articles contained in about 1600 BITs, inflating the weight given to more recent treaties.¹⁵⁰ They report that they successfully aligned the design of predicted texts with actual agreements that states were producing.¹⁵¹ Further, they created a user filter by which a state (for instance) could select another state with which it was negotiating a BIT and filter that state's treaty practice on particular clauses.¹⁵² The system allows the user to take into account the state's bargaining power in helping to predict what clauses will appear in the final treaty.¹⁵³ The authors argue that a machine-assisted approach will save states time, help with the complicated task of distilling a new agreement out of three thousand existing ones, and potentially allevi-

¹⁴⁸ See *infra* Section III.B.2 for further discussion of these tools.

¹⁴⁹ See Wolfgang Alschner & Dmitry Skougarevskiy, *Can Robots Write Treaties? Using Recurrent Neural Networks to Draft International Investment Agreements*, in LEGAL KNOWLEDGE AND INFORMATION SYSTEMS 119 (Floris Bex & Serena Villata eds., 2016).

¹⁵⁰ See Wolfgang Alschner, *Can Robots Write Treaties? Yes, They Can!*, MAPPING BITS BLOG (July 10, 2017), http://mappinginvestmenttreaties.com/blog/2017/07/robots_write_treaties/ [<https://perma.cc/BT83-CBJW>]; Wolfgang Alschner & Dmitry Skougarevskiy, *Towards an Automated Production of Legal Texts Using Recurrent Neural Networks 2* (Univ. of Ottawa, Working Paper No. 2017-27, 2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2984920 [<https://perma.cc/57UE-9EWU>].

¹⁵¹ Alschner & Skougarevskiy, *supra* note 150, at 4.

¹⁵² *Id.* at 3–4.

¹⁵³ See *id.* at 5.

ate power asymmetries, because it is possible to start negotiating from a draft text that reflects both sides' past practices.¹⁵⁴ Beyond bilateral negotiations, the authors suggest that "AI can also consolidate thousands of bilateral texts into a single multilateral treaty draft that can serve as a basis to harmonize the divergent practice of two hundred countries."¹⁵⁵

This type of tool, which can suggest possible compromises on treaty language and can help international lawyers make sense of large volumes of text, will only work if the agreement being negotiated is of the same type as a large number of existing treaties (such as extradition treaties, mutual legal assistance treaties, status of forces agreements, or agreements on the security of military information).¹⁵⁶ Nevertheless, the very fact that these agreements are so common means that the chance of any given state needing to negotiate a new one or amend an old one is quite high.

Another role for text-as-data tools is to help a state predict the likelihood that any one state's proposals end up in the final treaty text. This is most relevant for multilateral negotiations, where a host of states propose treaty language. Here, too, scholars have used text-as-data to evaluate which states' proposals and interests prevailed in trade and investment negotiations. These scholars use data from treaty negotiations and the final treaties that resulted.¹⁵⁷ They train the algorithm by teaching it which states submitted which specific language that ended up in the final text, allowing the researchers to de-

¹⁵⁴ See *id.* at 1–2.

¹⁵⁵ *Id.* at 2. Another source of data could be a collection of aide memoires such as those produced by U.N. Office for the Coordination of Humanitarian Affairs, which serve as a catalogue of language used in different instruments (e.g., treaties, U.N. resolutions, etc.) on particular topics. See U.N. Office for the Coordination of Humanitarian Affairs [OCHA], *Themes, PROTECTION OF CIVILIANS AIDE MEMOIRE*, <https://poc-aide-memoire.unocha.org/theme> [<https://perma.cc/56ST-W6CN>].

¹⁵⁶ See, e.g., 18 U.S.C. § 3181 note (2018) (Extradition Agreements) (listing over 100 U.S. extradition treaties); R. CHUCK MASON, CONG. RESEARCH SERV., RL34531, STATUS OF FORCES AGREEMENT (SOFA): WHAT IS IT, AND HOW HAS IT BEEN UTILIZED? 1 (2012), <https://fas.org/sgp/crs/natsec/RL34531.pdf> [<https://perma.cc/4RJL-S5ZA>] (stating that the United States has over 100 status of forces agreements).

¹⁵⁷ See Todd Allee & Manfred Elsig, *Are the Contents of International Treaties Copied-and-Pasted? Evidence from Preferential Trade Agreements* 10–11 (Nat'l Ctr. of Competence in Research on Trade Regulation, Working Paper No. 8, 2016); Todd Allee & Andrew Lugg, *Who Wrote the Rules for the Trans-Pacific Partnership?*, RES. & POL., July–Sept. 2016, at 4; Wolfgang Alschner & Dmitriy Skougarevskiy, *The New Gold Standard? Empirically Situating the Trans-Pacific Partnership in the Investment Treaty Universe*, 17 J. WORLD INV. & TRADE 339, 347 (2016); Tomer Broude et al., *The Trans-Pacific Partnership and Regulatory Space: A Comparison of Treaty Texts*, 20 J. INT'L ECON. L. 391, 397–99 (2017).

termine the strength of each state's influence.¹⁵⁸ Using this algorithm, states could predict, in a new treaty negotiation, which other states' proposals are likely to end up in the final version of the treaty.¹⁵⁹ Knowing which states historically have had the strongest influence in treaty negotiations can broadly inform states about trends in different areas of international law based on the preferences of those "persuasive" states. More narrowly, a state might use such a tool either to help it predict what the outcome of a treaty will look like or to help the state understand where it should direct its efforts to shape the final treaty language.

3. *Identifying Foreign Law*

The best-prepared negotiators take time in advance of negotiations to understand the domestic laws that may constrain their negotiating partners. Many states have some kind of "dualist" system, which means that only some forms of international law can override their domestic laws.¹⁶⁰ This means that a state negotiating a treaty will be very conscious of what its domestic law currently allows or provides on the subject being negotiated. Even for states (such as the United States) that employ a "later-in-time" rule to determine when a treaty will trump an existing statute, those negotiators may be loathe for political reasons to override their domestic law using an international instrument.¹⁶¹

This means that when State X comes to the negotiating table, it will be in a stronger position to anticipate and understand the rationales behind State Y's positions if it understands the relevant laws of State Y. Today, State X's negotiators can use web scraping, natural language processing tools, and translation software to locate all Bulgarian laws on riparian rights, or all Colombian extradition-related laws and associated case law. These tools, which are continually improving but are not yet perfect, are unlikely to provide complete answers, but their results will at least give negotiators a place to start more detailed research and to ask questions of their negotiating partners.¹⁶²

¹⁵⁸ See, e.g., Allee & Lugg, *supra* note 157, at 4.

¹⁵⁹ See *id.* at 1.

¹⁶⁰ See David Sloss, *Domestic Application of Treaties*, in *THE OXFORD GUIDE TO TREATIES* 367, 368–69 (Duncan B. Hollis ed., 2012).

¹⁶¹ STEPHEN P. MULLIGAN, CONG. RESEARCH SERV., RL32528, INTERNATIONAL LAW AND AGREEMENTS: THEIR EFFECTS UPON U.S. LAW 20–21 (2018), <https://fas.org/sgp/crs/misc/RL32528.pdf> [<https://perma.cc/CP78-NK4H>].

¹⁶² For an example of scholars using unsupervised machine learning to identify trends in Chinese administrative law disputes, see Benjamin L. Liebman et al., *Mass Digitization of Chi-*

4. *Processing Intelligence*

A different way that machine learning tools might assist negotiators before the negotiations begin is to help states process intelligence about the negotiations and the negotiators themselves. Various states' intelligence services likely already use machine learning tools to process the vast volumes of information they collect on other states,¹⁶³ and states' intelligence services may be better funded than their foreign ministry counterparts. The CIA, for example, has described its ability to forecast conditions on the ground in foreign countries and to obtain aggregate impressions by using vast volumes of digital intelligence.¹⁶⁴ Data sources such as social media provide “not only intelligence for specific operations—targeting this or that ISIS leader in Syria—but also taking the temperature of an entire population. It's sentiment analysis on a massive, population-wide scale.”¹⁶⁵ Assuming that negotiators rely today on their intelligence services for background information about the other states and individuals with which they will be negotiating, and the sentiments of the populations in those states, the use of machine learning already facilitates negotiators' ability to understand their negotiating partners' positions *ex ante*. Further, translation algorithms make the job of U.S. (and presumably other states') intelligence analysts easier, as the tools are now automatically able to translate foreign language communications collected by intelligence agencies into English.¹⁶⁶

nese Court Decisions: How to Use Text as Data in the Field of Chinese Law (U.C. San Diego Sch. Glob. Pol'y & Strategy, 21st Century China Ctr., Research Paper No. 2017-01, 2017), <https://ssrn.com/abstract=2985861> [<https://perma.cc/5L8Z-TU7Q>].

¹⁶³ Cortney Weinbaum, *The Ethics of Artificial Intelligence in Intelligence Agencies*, NAT'L INTEREST (July 18, 2016), <https://nationalinterest.org/blog/the-buzz/the-ethics-artificial-intelligence-intelligence-agencies-17021> [<https://perma.cc/MMT3-EFS2>] (“Intelligence agencies use advanced algorithms to interpret the meaning of intercepted communications, identify persons of interest and anticipate major events within troves of data too large for humans to analyze. If artificial intelligence is the ability of computers to create intelligence that humans alone could not have achieved, then the U.S. intelligence community invests in machines with such capabilities.”).

¹⁶⁴ Patrick Tucker, *Meet the Man Reinventing CIA for the Big Data Era*, DEF. ONE (Oct. 1, 2015), <https://www.defenseone.com/technology/2015/10/meet-man-reinventing-cia-big-data-era/122453/> [<https://perma.cc/H9DA-UHWQ>] (quoting CIA's Deputy Director as stating that “[i]f you have time, you can develop insight into patterns” using social media data).

¹⁶⁵ *Id.*

¹⁶⁶ See Stephen Babcock, *Johns Hopkins Center Is Bringing Machine Translation to Lesser-written Languages*, TECHNICALLY BALT. (Nov. 2, 2017, 12:57 PM), <https://technical.ly/baltimore/2017/11/02/johns-hopkins-center-bringing-machine-translation-lesser-used-languages/> [<https://perma.cc/S4WS-Q2Q9>]; Phil Goldstein, *What Is Deep Learning? A Look at Machine Learning in Federal IT Environments*, FEDTECH (July 31, 2019), <https://fedtechmagazine.com/article/2019/07/what-deep-learning-look-machine-learning-federal-it-environments-perfcon> [<https://perma.cc/>]

There is a potential dark side to the use of AI to better prepare one's negotiators for a treaty negotiation. The opposing state's negotiators may also have AI tools at their disposal. If the CIA creates all-source intelligence products for the U.S. negotiators, "the adversary regime would benefit enormously by manipulating the inputs and calculations the system uses to create its assessments."¹⁶⁷ Known as "adversarial machine learning," the adversary would try to surreptitiously alter the inputs used by the CIA algorithm to alter the algorithm's predictions.¹⁶⁸ Thus, intelligence services (and, in turn, a state's negotiators) may be duped into believing that a negotiating partner cares more about an issue than it actually does, and so may end up giving more of its own interests away in an effort to reach agreement. Indeed, the use of machine learning tools in these settings will give adversaries more "surface area" on which to conduct intelligence gathering and covert actions.¹⁶⁹

B. *Conducting Treaty Negotiations*

During negotiations, machine learning tools could help states that do not primarily use one of the six U.N. languages to obtain immediate interpretation and translation of documents created during negotiations. They might also help states in bilateral settings understand each other's proposals faster and develop compromise language more easily. Further, a state could use these tools to better anticipate, based on a state's textual proposals, whether that state is likely ultimately to ratify the treaty. More speculatively, technologically sophisticated states might begin to use emotion detection software during negotiations to parse the reactions of their negotiating partners and update their approach during the negotiations. Finally, these tools could help states identify and avoid substantive conflicts between the new treaty and existing treaties.

BA57-E3SS] (discussing Defense Department language recognition tools that deploy "deep learning").

¹⁶⁷ Weinbaum, *supra* note 163.

¹⁶⁸ Ian Goodfellow et al., *Making Machine Learning Robust Against Adversarial Inputs*, 61 COMMS. ACM 56 (2018), <https://cacm.acm.org/magazines/2018/7/229030-making-machine-learning-robust-against-adversarial-inputs/fulltext> [<https://perma.cc/P6X4-3SWW>].

¹⁶⁹ Theresa Hitchens, *IoT, AI Mean More Targets for US Adversaries, Says Gordon*, BREAKING DEF. (June 27, 2019, 4:47 PM), <https://breakingdefense.com/2019/06/iot-ai-mean-more-targets-for-us-adversaries-says-gordon/> [<https://perma.cc/39SL-YNHL>].

1. *Interpretation and Translation*

Perhaps the most obvious way in which machine learning tools can facilitate ongoing negotiations is for states to use these tools to improve the speed and quality of interpretation and translation.¹⁷⁰ When the U.N. hosts negotiations, it provides simultaneous interpretation in its six official languages: Spanish, French, English, Arabic, Russian, and Chinese.¹⁷¹ There are, however, a significant number of states whose diplomats and negotiators lack fluency of one of those languages. This places those states at a disadvantage in negotiations, as it is harder for them to follow the course of negotiations and they may miss important nuances.¹⁷² This is in addition to other disadvantages that these states might face: states with fewer resources tend to send fewer diplomats to negotiations (which can be a problem when multiple working groups meet at the same time during a negotiation), and often ask the same diplomats to cover different subject matter negotiations (which means their expertise may be shallower than for states that can send subject matter experts to each negotiation).¹⁷³

Further, a variety of important conversations that take place during the course of treaty negotiations happen not in the plenary sessions but in corridors and side rooms.¹⁷⁴ Negotiators who do not speak one of the primary languages can find themselves at a disadvantage in

¹⁷⁰ Interpretation refers to the process by which oral interventions in one language are converted into another language. Translation refers to the process by which written products are converted from one language to another.

¹⁷¹ *Official Languages*, *supra* note 145. There are other international settings in which flawless translation is imperative as well. The International Court of Justice, for example, must engage in careful translation exercises between French and English. See INT'L COURT OF JUSTICE, HANDBOOK 31 (2018), <https://www.icj-cij.org/files/publications/handbook-of-the-court-en.pdf> [<https://perma.cc/4N5R-RUEU>]. Machine learning-based computer translation could potentially speed up the early stages of that process, though humans would play the ultimate role in delivering the translation.

¹⁷² See, e.g., Anthony Aust, *The Role of the FCO in UK Government*, U.K. PARLIAMENT, (Jan. 17, 2011), <https://publications.parliament.uk/pa/cm201011/cmselect/cmfaif/writev/fcogov/m16.htm> [<https://perma.cc/B67B-3PM7>] (“[A]lthough [a British diplomat’s counterpart] may understand and speak English apparently fluently, it is not usually their language. Thus, serious mistakes and misunderstandings can occur if one does not know the local language well.”).

¹⁷³ John S. Odell & Dustin Tingley, *Negotiating Agreements in International Relations*, in NEGOTIATING AGREEMENTS IN POLITICS 144, 155 (Jane Mansbridge & Cathie Jo Martin eds., 2013) (“Developing countries, particularly the least developed, often lack sufficient domestic institutional capacity to negotiate effectively on technical issues. . . . Many also assign their ambassadors in Geneva to several international organizations simultaneously, and some countries give those ambassadors little support or attention in their home capitals.”).

¹⁷⁴ See, e.g., BRUNO VERDINI TREJO, WINNING TOGETHER 148–50 (2017) (explaining how “informal side-talks became a necessary and integral element of the success” of the negotiations).

those settings in making their views known and in processing and responding to the views of others.

Some negotiators may find it easier to read one or more of the languages on paper rather than to hear them spoken. The United Nations produces official documents for U.N. negotiations in each of the six languages, including the formal submissions by states, but it takes time for the translators to produce and finalize these documents and, as with interpretation, the United Nations does not produce translations into the many other nonofficial languages that member states speak.¹⁷⁵

Finally, many treaty negotiations happen outside U.N. fora, in settings that lack the same number and quality of professional interpreters and translators that the United Nations has. Three states from very different regions might decide to negotiate a treaty about a narrow trade issue, for instance; such a negotiation is likely to take place in a context of limited resources and lower quality interpretation and translation than one would find in the U.N. setting.¹⁷⁶ States might therefore not have a true meeting of the minds when they find themselves agreeing to a document that has been badly translated after a series of oral exchanges that have been imprecisely interpreted.

Fortunately, one of the most concrete developments in artificial intelligence in the past decade has been in the area of language translation.¹⁷⁷ Various companies and scientists have worked on this problem, though Google may have made the most progress. Using a tool called deep learning, Google has created a “neural machine translation system” that learns by analyzing existing translations but does a better job than other systems of reflecting the actual meaning of words in their translations.¹⁷⁸ These systems examine parallel texts in

¹⁷⁵ See *Official Languages*, *supra* note 145. Indeed, researchers have employed U.N. documents to improve automated translation systems. See Surden, *supra* note 9, at 99–100.

¹⁷⁶ For a discussion of the costs of negotiations between different language speakers, see Stanko Nick, *Use of Language in Diplomacy*, in *LANGUAGE AND DIPLOMACY* 39, 43 (Jovan Kurbalija & Hannah Slavik eds., 2001), https://www.diplomacy.edu/sites/default/files/Language_Diplomacy_Chapter2.PDF [<https://perma.cc/ZRK2-K4TU>].

¹⁷⁷ See *Technology Quarterly: Finding a Voice*, *ECONOMIST* (2017), <https://www.economist.com/technology-quarterly/2017-05-01/language> [<https://perma.cc/5VRC-VXH3>] (arguing that machine translation “may soon be good enough to require only modest editing by humans”); Gideon Lewis-Kraus, *The Great A.I. Awakening*, *N.Y. TIMES* (Dec. 14, 2016), <https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html> [<https://perma.cc/NHZ2-REPL>] (describing the evolution of machine translation).

¹⁷⁸ See Davide Castelvecchi, *Deep Learning Boosts Google Translate Tool*, *NATURE* (Sept. 27, 2016), <https://www.nature.com/news/deep-learning-boosts-google-translate-tool-1.20696> [<https://perma.cc/5CUL-EGFT>]. *But see* Douglas Hofstadter, *The Shallowness of Google Translate*, *ATLANTIC* (Jan. 30, 2018), <https://www.theatlantic.com/technology/archive/2018/01/the-shal->

the native language and target language, and then use statistics to “infer which words typically follow other words, and how words are translated in context.”¹⁷⁹ For some language pairs, the accuracy of Google’s system approaches that of human translators.¹⁸⁰ Thus, while U.N. translators might still be the gold standard for common language pairs, Google Translate now provides significant benefits to negotiators who must translate their submissions or objections into one of the six U.N. languages.¹⁸¹ A Google competitor, DeepL, also uses neural networks to produce high quality translations, though it only translates among eleven languages.¹⁸² As of 2018, Google Translate works in 96 languages.¹⁸³ For even more obscure languages, Google has created a multilingual translation tool based on neural networks that can translate between two languages for which it lacks direct translation data by using data between each language and a third language.¹⁸⁴ For some languages, Google Translate now can translate text contained in a photograph.¹⁸⁵ That means that a negotiator who speaks Bengali could photograph an official U.N. translation in English and use Google Translate or another advanced translation tool to immediately receive a translation of that document in Bengali. The Bengali-speaking diplomat likewise could more easily translate his proposals into

lowness-of-google-translate/551570/ [https://perma.cc/TC4U-VPC8] (expressing deep skepticism about machine translation’s ability to rival humans).

179 Kim Martineau, *Computer Scientists Julia Hirschberg and Christopher Manning Discuss Natural Language Processing*, COLUM. ENG’G (July 21, 2015), <https://engineering.columbia.edu/news/computer-scientists-julia-hirschberg-and-christopher-manning-discuss-natural-language-processing> [https://perma.cc/V6DQ-JHA6] (noting that machine learning techniques that use neural networks are much better at inferring context and recognizing similar meanings).

180 Castelvechi, *supra* note 178.

181 *See Official Languages*, *supra* note 145.

182 *Translator*, DEEPL, <https://www.deepl.com/en/translator> [https://perma.cc/5W8W-M3CJ]; Florian Faes, *Why DeepL Got Into Machine Translation and How It Plans to Make Money*, SLATOR (Oct. 19, 2017), <https://slator.com/technology/deepl-got-machine-translation-plans-make-money/> [https://perma.cc/889D-KLFF].

183 *See Hofstadter*, *supra* note 178. Another company recently released a “pocket translator” called Pocketalk, which translates phrases to and from 74 languages using machine translation and voice recognition software from Google, Baidu, and other companies. Kantaro Komiya, *Tourists Are Fueling a Boom in Personal Translation Devices*, BLOOMBERG (Aug. 7, 2019, 5:00 AM), <https://www.bloomberg.com/news/articles/2019-08-06/what-did-you-say-tourists-fuel-pocket-translator-boom-in-japan> [https://perma.cc/LK58-G3JM].

184 PAUL SCHARRE & MICHAEL C. HOROWITZ, CTR. FOR NEW AM. SEC., *ARTIFICIAL INTELLIGENCE* 6 (2018), https://s3.amazonaws.com/files.cnas.org/documents/CNAS_AI_FINAL-v2.pdf?mtime=20180619100112 [https://perma.cc/7GKT-T8H8].

185 *Translate Images*, GOOGLE TRANSLATE HELP, <https://support.google.com/translate/answer/6142483?co=GENIE.Platform%3DAndroid&hl=en> [https://perma.cc/6GEA-4LDT].

one of the U.N. languages to submit to the U.N. Secretariat for consideration by the entire set of negotiating states.

There have also been significant developments in simultaneous oral translation between languages. Skype, for instance, has developed “Skype Translator,” which translates, and produces a transcript of, speech in real time between speakers of 10 languages.¹⁸⁶ Like Google Translate, Skype’s AI software has become good at predicting how any given series of words should translate.¹⁸⁷

At bottom, these machine learning-enabled tools are likely to improve the quality and speed of interpretation and translation, particularly for negotiators who are uncomfortable in the U.N. languages. Further, these tools continue to improve as data scientists expose them to more and more data.¹⁸⁸ This means fewer misunderstandings during negotiations and, potentially, a final treaty text that more states can ratify because they have correctly understood the import of the text that the states finalized.¹⁸⁹ These tools will not replace face-to-face interactions and exchanges, which remain critical to the conduct of negotiations, but they will enhance exchanges by advancing a more closely shared understanding of the substance of those interactions.¹⁹⁰

2. *Parsing Partner Behavior*

Section A identified web scraping and other text-as-data tools that could help states understand their negotiating partners’ sensitivities *ex ante*, based on information and sentiment analysis gleaned from documents such as prior U.N. submissions, as well as government speeches and social media.¹⁹¹ These web scraping tools could also help states during negotiations.¹⁹² Once it becomes public that a state is engaged in treaty negotiations about a particular issue, news articles and legal or policy blog posts follow. Sometimes the articles

¹⁸⁶ See *Skype Translator*, SKYPE, <https://www.skype.com/en/features/skype-translator/> [<https://perma.cc/Z5TZ-EQML>].

¹⁸⁷ Kevin Kelly, *The Three Breakthroughs That Have Finally Unleashed AI on the World*, WIRED (Oct. 27, 2014, 6:30 AM), <https://www.wired.com/2014/10/future-of-artificial-intelligence/> [<https://perma.cc/SEK6-RHPZ>].

¹⁸⁸ *Id.*

¹⁸⁹ See Odell & Tingley, *supra* note 173, at 148.

¹⁹⁰ See JENNIFER PARLAMIS, FACE-TO-FACE AND EMAIL NEGOTIATIONS 1 (2010), <https://repository.usfca.edu/cgi/viewcontent.cgi?article=1004&context=olc> [<https://perma.cc/L9ZT-XERM>].

¹⁹¹ See *supra* Section II.A.

¹⁹² Odell & Tingley, *supra* note 173, at 148 (discussing ways in which incomplete information can hinder reaching an agreement or, conversely, how withholding certain information can enable agreements).

and posts are misguided because the authors misunderstand what topics are on the table during negotiations.¹⁹³ At other times, though, the articles and posts provide important insights about why certain provisions would be well or ill received in the negotiating state and, in particular, its legislature.¹⁹⁴ Further, even if the critics get the subject of negotiations wrong, those critics can affect the public's views, and may even be able to turn public opinion against an entire treaty.¹⁹⁵

Another possible use of text-as-data to parse partner behavior during negotiations is to run that data through empirical models that can evaluate the positions states are taking during negotiations. Understanding those positions can help one state predict the other states' ultimate approach to ratifying and developing reservations to the treaty.¹⁹⁶ Kevin Cope recently coded the travaux préparatoires of two treaties and employed a spatial model to create, based on states' interventions during negotiations, each state's "ideal point" relative to the language of the final treaty. He used those ideal points to predict whether a given state valued the treaty enough relative to the status quo to ratify it.¹⁹⁷ He argues that his model could help a state better understand the likelihood of future adherence by its negotiating partners at the treaty negotiation phase, and might also help smaller players identify their common interests with other smaller players, thus enhancing their collective negotiating power.¹⁹⁸

A more controversial and more speculative use of machine learning during negotiations would be the use of algorithms that detect facial emotions and truthfulness.¹⁹⁹ Microsoft sells "emotion recognition

¹⁹³ See, e.g., Bertram I. Spector, *The Search for Flexibility on Financial Issues at UNCED: An Analysis of Preference Adjustment*, in *NEGOTIATING INTERNATIONAL REGIMES* 87, 87 (Bertram I. Spector et al. eds., 1994) (noting that it may be difficult to access critical information and that issues may not be fully defined while negotiations are underway).

¹⁹⁴ *Id.* at 155 (citation omitted) ("Even in autocracies . . . internal opposition often limits a party's negotiating position and the agreements that it can ratify. Especially in highly salient cases, such as peace talks among Israel, the Palestinian Authority, and the United States, domestic voices are strong in all three countries. Leaders and negotiators must consider these voices.").

¹⁹⁵ *Id.* ("Occasionally, a domestic or transnational campaign has been decisive in . . . stopping a negotiation that governments had begun Sometimes governments have negotiated and initiated an agreement and then failed to achieve ratification at home.").

¹⁹⁶ See Kevin L. Cope, *Making Treaties* (unpublished manuscript) (on file with author).

¹⁹⁷ See *id.* at 40.

¹⁹⁸ See *id.* at 54–55.

¹⁹⁹ See, e.g., Byoung Chul Ko, *A Brief Review of Facial Emotion Recognition Based on Visual Information*, 18 *SENSORS* 401 (2018), <https://www.mdpi.com/1424-8220/18/2/401/pdf> [<https://perma.cc/DMY6-24NR>]; *Ever Better and Cheaper, Face-Recognition Technology Is Spreading*, *ECONOMIST* (Sept. 9, 2017), <https://www.economist.com/business/2017/09/09/ever-better-and-cheaper-face-recognition-technology-is-spreading> [<https://perma.cc/QBH3-LQMB>] (en-

algorithms” that clients may use to identify, with given confidence levels, what emotion a person is showing on her face.²⁰⁰ Further, data scientists are trying to develop algorithms that detect whether a person is lying.²⁰¹ Indeed, the U.S. Department of Homeland Security and Canadian and E.U. authorities are testing a system called AVATAR, which detects changes in eyes, voice, gestures, and posture to determine whether a person who is seeking to enter the country is acting deceptively and should receive additional screening.²⁰² Notwithstanding the growing use of these tools, real questions remain about how reliable they are outside lab-like settings.²⁰³

These emotion recognition tools, if effective, could be highly relevant in the negotiation context. There is a deep literature on lying during negotiations.²⁰⁴ In the context of international negotiations, states have incentives to exaggerate their bargaining strengths, strategically conceal their true strengths and weaknesses, and critique the other side’s proposals in an effort to gain a larger piece of the pie.²⁰⁵

visioning product that could “detect dissatisfaction on shoppers’ faces”); *What Machines Can Tell from Your Face*, *ECONOMIST* (Sept. 9, 2017), <https://www.economist.com/leaders/2017/09/09/what-machines-can-tell-from-your-face> [<https://perma.cc/S6LD-749T>] (describing systems that can analyze faces for emotions).

²⁰⁰ Rich Firth-Godbehere, *Silicon Valley Thinks Everyone Feels the Same Six Emotions*, *QUARTZ* (Sept. 17, 2018), <https://qz.com/1392130/silicon-valley-thinks-everyone-feels-the-same-six-emotions/> [<https://perma.cc/CF85-4BQB>].

²⁰¹ See ZHE WU ET AL., *DECEPTION DETECTION IN VIDEOS* (2018), <https://arxiv.org/pdf/1712.04415.pdf> [<https://perma.cc/M964-N9FY>]; see also Dom Galeon, *A New AI that Detects “Deception” May Bring an End to Lying as We Know It*, *FUTURISM* (Jan. 9, 2018), <https://futurism.com/new-ai-detects-deception-bring-end-lying-know-it> [<https://perma.cc/LUD8-6HM9>] (quoting a University of Maryland researcher as noting that these systems continue to improve and that we could be three to four years away from AI that detects deception flawlessly); Martineau, *supra* note 179 (noting that deceptive behavior varies by culture and that researchers are now studying how people from different cultures deceive others).

²⁰² Steven Melendez, *Goodbye Polygraph? New Tech Uses AI to Tell If You’re Lying*, *FAST CO.* (May 24, 2018), <https://www.fastcompany.com/40575672/goodbye-polygraphs-new-tech-uses-ai-to-tell-if-youre-lying> [<https://perma.cc/8BVE-L8KB>]; see also Natalie Salmanowitz, *Overview of U.S. Lie Detection Systems for Airport Security Checkpoints*, *LAWFARE* (Dec. 3, 2018, 9:50 AM), <https://www.lawfareblog.com/overview-us-lie-detection-systems-airport-security-checkpoints> [<https://perma.cc/U82R-GXQZ>] (describing AVATAR and related tools used by other countries).

²⁰³ See Firth-Godbehere, *supra* note 200.

²⁰⁴ See, e.g., Geoffrey M. Peters, *The Use of Lies in Negotiation*, 48 *OHIO ST. L.J.* 1 (1987); Walter W. Steele, Jr., *Deceptive Negotiating and High-Toned Morality*, 39 *VAND. L. REV.* 1387 (1986); Gerald B. Wetlaufer, *The Ethics of Lying in Negotiations*, 75 *IOWA L. REV.* 1219 (1990).

²⁰⁵ For instance, a treaty negotiator might lie or be misleading about whether a particular treaty term would ever be acceptable, how much it would cost to implement a given provision, or whether the negotiator’s legislature would ever agree to a treaty containing a given term. See Reilly, *supra* note 135, at 492–93 (listing the types of topics that people often lie about in negotiations); see also John K. Setear, *Law in the Service of Politics: Moving Neo-Liberal Institutional-*

Negotiators may also try to feed false information into the negotiations, directly or via a third party.²⁰⁶ Assuming that deception plays a salient role in treaty negotiations, a negotiator who has access to an emotion-recognition algorithm (deployed through a camera placed in the negotiating room or even the negotiator's eyeglasses²⁰⁷) would acquire an advantage over her counterpart. Though this sounds like an extreme measure, recent news reports indicate that China already deploys surveillance technology in classrooms so that the school leadership can monitor in real time what instructors are teaching and how students are reacting.²⁰⁸ It is only a short step to employing emotion recognition software through those closed-circuit cameras to provide nearly instant feedback to negotiators. Of course, with only a small number of faces to "read," negotiators might just as easily employ a facial expression expert to view the negotiation videos and provide guidance.

Yet another controversial option for parsing the behavior of foreign negotiators is to deploy lip-reading algorithms.²⁰⁹ As with emotion detection software, one state's negotiating team might use these machine learning algorithms, which to date are nascent, to process videos of foreign negotiators to try to detect what they are saying while they are out of earshot. The program, developed by a U.K. company called DeepMind, "requires clear, straight-ahead video, and [its accuracy is hardly] perfect," but likely will improve over time.²¹⁰

ism from Metaphor to Theory by Using the International Treaty Process to Define "Iteration," 37 VA. J. INT'L L. 641, 683 (1997) ("A nation could lie, threaten, and bribe in the negotiation phase.").

²⁰⁶ See Reilly, *supra* note 135, at 490.

²⁰⁷ See Niraj Chokshi, *How Surveillance Cameras Could Be Weaponized with A.I.*, N.Y. TIMES (June 13, 2019), <https://www.nytimes.com/2019/06/13/us/aclu-surveillance-artificial-intelligence.html> [<https://perma.cc/2XFF-PBCF>] (discussing recent technological improvements that train software to actively monitor video feeds).

²⁰⁸ See Tara Francis Chan, *A School in China Is Monitoring Students with Facial-Recognition Technology that Scans the Classroom Every 30 Seconds*, BUS. INSIDER (May 20, 2018, 9:21 PM), <https://www.businessinsider.com/china-school-facial-recognition-technology-2018-5> [<https://perma.cc/6HSU-U2A7>] (noting that the tool categorizes students' facial expressions); Zhang Lun, *What Happens when Universities Become "Party Strongholds,"* N.Y. TIMES (Oct. 18, 2018), <https://www.nytimes.com/2018/10/18/opinion/chinas-watchful-eye-reaches-into-the-classroom.html> [<https://perma.cc/XSB5-MMJL>].

²⁰⁹ See Matthew Hutson, *Lip-Reading Artificial Intelligence Could Help the Deaf—Or Spies*, AAAS (July 31, 2018, 3:15 PM), <https://www.sciencemag.org/news/2018/07/lip-reading-artificial-intelligence-could-help-deaf-or-spies> [<https://perma.cc/8D4P-YSCP>].

²¹⁰ *Id.*

3. *Identifying Inter-Treaty Conflicts*

Another concern for a state's negotiators is whether a new treaty will contain any provisions that might conflict with that state's existing treaty obligations. A state could create a tool that allows it to compare proposed treaty language (either while the negotiations are underway, or before the state has ratified the treaty) to all other treaties to which the state is a party to detect similar language or very similar topics. The state might do so, for instance, using the types of plagiarism detection software or topic models that identify similar clusters of words.²¹¹ That state's international lawyers could then assess whether those families of words create conflicts among themselves and, if they do, decide how to proceed with the second negotiated treaty before the state formally faces a direct conflict. Indeed, a state might use these tools, even in the absence of a concrete negotiation, to compare its own existing treaties in order to identify gaps or other potential conflicts before they produce a real world problem.

A related concern is that a state's negotiating partner may deliberately try to create conflicts between the treaty being negotiated and the partner's existing obligations.²¹² Kal Raustiala and David Victor have discussed states' strategic efforts to create contradictory rules between treaties in the biodiversity and trade contexts in order to set the agenda for future negotiations.²¹³ A state that is concerned that its negotiating partner may be attempting this move could use web scraping and topic analysis on its negotiating partner's existing treaties to assess whether the partner is trying to play this game.

C. *Identifying Customary International Law*

Treaties serve as an important source of international law to which foreign ministry lawyers will turn when addressing whether a particular action is consistent with international law. There is, however, another key source of international law: custom. CIL is comprised of state practice undertaken out of a sense of legal obligation (*opinio juris*).²¹⁴ That state practice and *opinio juris* can take a range

²¹¹ See *supra* note 71 and accompanying text.

²¹² See Todd Allee et al., *The Ties Between the World Trade Organization and Preferential Trade Agreements: A Textual Analysis*, 20 J. INT'L ECON. L. 333, 335–36 (2017) (describing concerns about conflicting international law obligations, especially in the trade area).

²¹³ Kal Raustiala & David G. Victor, *The Regime Complex for Plant Genetic Resources*, 58 INT'L ORG. 277, 301–02 (2004) (“Cognizant that the growing legalization of world politics means that legal conflicts focus efforts at solutions, states at times attempt to force change by explicitly crafting rules in one elemental regime that are incompatible with those in another.”).

²¹⁴ BRADLEY ET AL., *supra* note 92, at xxv.

of forms, including physical, diplomatic, or legislative acts of states, court decisions, submissions to U.N. bodies, and public speeches.²¹⁵

One persistent critique of CIL is that it contains a western bias in both its origins and its purposes.²¹⁶ This is undoubtedly due in part to the higher visibility (and sometimes the higher quantity) of the practice of states such as the United States and the United Kingdom.²¹⁷ Courts and scholars who attempt to discern CIL must revert to whatever practice and *opinio juris* are available, and states such as the United States and the United Kingdom have long published digests of their international practice, creating readily available sources of information.²¹⁸ Further, especially for English speakers, non-English state practice is harder to identify and analyze. As Ryan Scoville notes, “There is no complete database—electronic or otherwise—of state practice and *opinio juris*.”²¹⁹ His research into academic sources that purport to identify custom suggests that “contemporary CIL retains the under-inclusive and overwhelmingly occidental genealogy of the historical law of nations.”²²⁰

Undoubtedly a much wider range of state practice and *opinio juris* (including non-Western and non-recent practice) exists. The trick is identifying it. Text-as-data tools have an important role to play here. Tools such as topic modeling make it relatively easy to identify sentences in texts that discuss concepts such as freedom of navigation or conditions of detention.²²¹ There have been significant advances recently in optical character recognition, which states could use to identify previously unknown state practice and *opinio juris* lurking in digital archives. Scholars already employ machine learning to retrieve and process other types of historical documents.²²²

²¹⁵ See Ryan M. Scoville, *Finding Customary International Law*, 101 IOWA L. REV. 1893, 1896 (2016).

²¹⁶ See generally ANTONY ANGHIE, *IMPERIALISM, SOVEREIGNTY AND THE MAKING OF INTERNATIONAL LAW* 4 (2005) (describing the “centrality of colonialism for the generation of international law”); B.S. Chimni, *Customary International Law: A Third World Perspective*, 112 AM. J. INT’L L. 1, 6 (2018) (arguing that CIL has “its roots in the western economic, cultural, and political order”).

²¹⁷ See Chimni, *supra* note 216, at 23 n.131.

²¹⁸ See *Digest of United States Practice in International Law*, U.S. DEP’T OF STATE, <https://www.state.gov/digest-of-united-states-practice-in-international-law/> [<https://perma.cc/AK8U-TRN7>]; *Lauterpacht Centre for International Law: International Law Reports*, U. CAMBRIDGE, <https://www.lcil.cam.ac.uk/publications/international-law-reports> [<https://perma.cc/4EP8-XUJL>] (stating that the Reports have been published since 1922).

²¹⁹ Scoville, *supra* note 215, at 1896.

²²⁰ *Id.* at 1899.

²²¹ See *supra* note 49 and accompanying text.

²²² See, e.g., Michael S. Moss & Tim J. Gollins, *Our Digital Legacy: An Archival Perspec-*

In addition, many domestic legal actors use these tools for e-discovery, allowing them to parse through thousands or millions of documents.²²³ Those systems perform a “classification task” that enables them to classify documents as highly relevant, possibly relevant, or non-relevant.²²⁴ In the CIL context, these tools could identify documents that are most likely to contain information on categories of CIL provided to the data scientists by international lawyers. Even if the tools produce a number of false positives, they would significantly narrow the number of documents international lawyers must analyze when trying to detect previously unknown state practice and opinio juris.

The long pole in the tent here is digitization, not machine learning. The biggest hurdle to using text-as-data tools to pursue this goal is that wide swaths of relevant state practice and opinio juris are not currently in digital form.²²⁵ The fact that a range of state practice is not available for public review has been a persistent issue since the founding of the United Nations.²²⁶ Part of the problem at that time was the reluctance of states to provide their material for review.²²⁷ This resulted in a “one-sided” treatment of international law in monographs and elsewhere²²⁸—a problem that persists today, even if it is less pervasive.²²⁹ Hersch Lauterpacht, arguing against the shortsighted view of some states that it was in their interests to keep their practices secret, concluded, “The benefits accruing from the publication of the records outweigh decisively any such remote disadvantages.”²³⁰

tive, 4 J. CONTEMP. ARCHIVAL STUD. 1, 17 (2017), <https://elischolar.library.yale.edu/cgi/viewcontent.cgi?referer=http://history-lab.org/&httpsredir=1&article=1035&context=jcas> [https://perma.cc/UB9S-H7DJ].

²²³ *How to Make the E-Discovery Process More Efficient with Predictive Coding*, THOMSON REUTERS, <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient> [https://perma.cc/T4VQ-TPSQ].

²²⁴ Surden, *supra* note 9, at 112.

²²⁵ See U.N. Secretariat, Int’l Law Comm’n, Identification of Customary International Law: Ways and Means for Making the Evidence of Customary International Law More Readily Available, ¶ 12, U.N. Doc. A/CN.4/710/Rev.1 (Feb. 14, 2019) [hereinafter U.N. Secretariat Report] (suggesting that only a limited number of sources were available in digital form).

²²⁶ See ILC Report, *supra* note 119, at 95.

²²⁷ See *id.*

²²⁸ See *id.* at 96.

²²⁹ See U.N. Secretariat Report, *supra* note 225, ¶ 77 (discussing types of information that states continue to withhold).

²³⁰ Hersch Lauterpacht, Book Review, 22 BRIT. Y.B. INT’L L. 310, 311 (1945) (reviewing GREEN HAYWOOD HACKWORTH, DIGEST OF INTERNATIONAL LAW (1940)); see also Manley O. Hudson, *Twelve Casebooks on International Law*, 32 AM. J. INT’L L. 447, 455 (1938) (predicting that the study of international law would be revolutionized if a general and global collection of cases, both national and international, were produced).

Nevertheless, there is good reason to believe that a vast body of underlying documentation exists, whether or not in digital form. In 1949, the ILC produced a document entitled *Ways and Means of Making the Evidence of Customary International Law More Readily Available* (“ILC Report”).²³¹ The ILC report made clear that not only the United States and the United Kingdom, but also France, Russia/USSR, Switzerland, Germany, and nearly all Latin American states produced official or semi-official collections of diplomatic and international law-related documents.²³² In addition to its digest, the United Kingdom also produced “200 manuscript volumes of Opinions of the Law Officers in international legal questions,”²³³ and British scholars produced a digest of 3,000 decisions of international law cases from 1919–1942.²³⁴ Even in 1949, then, one is left with the impression of an overwhelming number of documents related to international law and arbitration,²³⁵ as well as a strong desire among some states and scholars for more states to make their practice known.²³⁶ That volume of documents will have grown exponentially since then.

²³¹ ILC Report, *supra* note 119.

²³² *Id.* at 9–10 (“Nearly all the Latin-American Republics have, almost from the time of their establishment, published through their Ministries of Foreign Affairs annual reports or other compilations containing diplomatic correspondence and other materials on international law. . . . There can be no doubt but that this mass of material . . . is very rich. It has unfortunately been little used hitherto, perhaps because it is not always easy to come by, and because it is largely unindexed.”). Even at the time, the report noted, “The diplomatic archives of virtually every State are now so extensive that their examination and digesting is beyond the capacity of any one individual.” *Id.* at 22. Unindexed documents pose no problem for computational text analysis. See Livermore et al., *supra* note 44, at 994–95.

²³³ ILC Report, *supra* note 119, at 22.

²³⁴ *Id.* at 50–51.

²³⁵ *See id.* at 47 (noting that the number of documents from the Nuremberg tribunal alone were huge—37 volumes of evidence, testimony, and exhibits).

²³⁶ *See, e.g., id.* at 93–94 (stating that Judge John Bassett Moore was “so impressed with the need for greater availability of the evidences of international law that . . . he urged . . . a collection of all treaties, ancient and modern [plus the] publication of all State papers . . . relating to foreign affairs . . . and . . . the collection and publication of international judicial decisions” so as to produce a comprehensive collection of the sources of international law); *see also* William W. Bishop, Jr., *Recent American Judicial Decisions Involving International Law*, 42 AM. J. INT’L L. 194, 195 (1948) (“One of the most important tasks in the field of international law at the present time is to make the source-materials of international law more widely and more readily accessible.”); Herbert W. Briggs, *Finding International Law*, 42 AM. J. INT’L L. 101, 103 (1948) (“The materials with which the international lawyer must work are manifold. They exist, and they should be made more accessible. The United Nations can perform an important service and encourage the progressive development of international law by promoting the collection and publication of these materials.”); Georg Schwarzenberger, *The Inductive Approach to International Law*, 60 HARV. L. REV. 539, 563–64 (1947) (“[T]he immense material from which ‘international custom, as evidence of a general practice accepted as law,’ may be gathered has hardly yet been touched by international lawyers.”).

Text-as-data tools may be hugely helpful in assessing what is out there, if—but only if—states decide to digitize more of their practice. Where “data does not exist at all or only in non-machine readable formats, [it] precludes the deployment of computational tools”²³⁷ Even if states themselves, however, do not digitize more of their past diplomatic, military, and political practice, the Internet provides a vast data set worth perusing to identify state practice from recent decades.²³⁸

Why resort to machine learning tools rather than simply use Google? One state will never describe its practice in precisely the same terms as another state. There are many ways to describe an act that represents a “freedom of navigation” exercise, and the idea is to capture all of those episodes, not just the ones in which states happen to describe the event in those specific terms. This kind of information therefore cannot be extracted only using Boolean searches or keywords.²³⁹ “[T]ext analytics or text mining, ‘refers to the discovery of knowledge that can be found in text archives.’”²⁴⁰ One type of text analytics tool is a probabilistic topic model, an algorithm that helps discover the main themes in a large and unstructured collection of documents and then organize the collection according to the discovered themes.²⁴¹ One advantage to this tool is that data scientists do not have to label or annotate the documents first, which saves time.²⁴² Systems can not only identify topics but also undertake “concept expansion” by identifying concepts that are close to those given in the query, which can help identify hidden associations and implicit rela-

²³⁷ Alschner et al., *supra* note 1, at 227; see Justin Grimmer & Brandon M. Stewart, *Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts*, POL. ANALYSIS 1, 5–6 (2013), <https://web.stanford.edu/~jgrimmer/tad2.pdf> [<https://perma.cc/5QG9-KC4K>] (“The most difficult to acquire are texts found in archives or yet-to-be-scanned books. But preparing these texts for analysis can be straightforward—using a high-quality scanner and Optical Character Recognition software, it is possible to convert archival materials into computer readable texts.”).

²³⁸ See, e.g., U.N. Secretariat Report, *supra* note 225, ¶ 12 (discussing use of online resources).

²³⁹ See KEVIN D. ASHLEY, *ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS* 27 (2017) (describing legal queries that cannot be easily answered with ordinary information retrieval tools such as Boolean searches).

²⁴⁰ *Id.* at 5 (quoting Xia Hu & Huan Liu, *Text Analytics in Social Media*, in *MINING TEXT DATA* 385, 387 (Charu C. Aggarwal & ChengXiang Zhai eds., 2012)). This is also referred to as “information extraction,” which relies in part on machine learning to help process semantic information in the text. See *id.*

²⁴¹ See David M. Blei, *Probabilistic Topic Models*, 55 COMMS. ACM 77, 77–78 (2012), <http://www.cs.columbia.edu/~blei/papers/Blei2012.pdf> [<https://perma.cc/ST5B-GE22>].

²⁴² *Id.* at 78.

tionships.²⁴³ Thus, the system might fruitfully identify that the concept of “freedom of navigation” is semantically linked to the concept of “contested territorial seas” and “disputed islands” based on the texts it evaluates.²⁴⁴

Consider the Customary International Humanitarian Law Study by the International Committee of the Red Cross (“ICRC”). The ICRC assembled a vast array of state practice and *opinio juris* in an effort to evaluate the customary status of a host of law of war rules.²⁴⁵ It took 10 years for the ICRC to collect and analyze the state practice and *opinio juris* of states from all over the world, including an in-depth study of 50 states.²⁴⁶ With the text-as-data tools described here, it would have been far easier and far faster to assemble the collections of state acts and evaluate their meaning.

How should states proceed if they want to apply such tools to digital archives to unearth and analyze state practice? The International Law Commission recently asked the U.N. Secretariat to update the 1949 ILC report and provide suggestions for ways to make evidence of CIL more readily available.²⁴⁷ The Secretariat’s 2018 report notes that the availability of evidence of CIL has “changed strikingly” since 1949²⁴⁸ and that the “sheer quantity of available material is daunting.”²⁴⁹ For example, the Secretariat found bibliographic resources containing 192 states’ executive conduct and 184 states’ judicial acts.²⁵⁰ At least two challenges that the Secretariat identified are ones that text-as-data tools are well-positioned to help address: the multiplicity of languages and the vast amount of potentially relevant

²⁴³ ASHLEY, *supra* note 239, at 30 (describing concept expansion).

²⁴⁴ “Word embeddings” seek to relate word occurrences with their neighboring expressions so as to better understand the word’s meaning. One common word embedding tool “uses distributed representations of text to capture similarities among concepts.” Tomas Mikolov et al., *Learning the Meaning Behind Words*, GOOGLE OPEN SOURCE (Aug. 14, 2013), <https://open-source.googleblog.com/2013/08/learning-meaning-behind-words.html> [<https://perma.cc/4U6P-4QH9>].

²⁴⁵ See JEAN-MARIE HENCKAERTS & LOUISE DOSWALD-BECK, INT’L COMM. OF THE RED CROSS, CUSTOMARY INTERNATIONAL HUMANITARIAN LAW (2005).

²⁴⁶ *Id.* at xxxiii (noting that the project started in 1995); *id.* at li (noting the focus on 50 states).

²⁴⁷ See U.N. Secretariat Report, *supra* note 225, ¶ 2.

²⁴⁸ *Id.* ¶ 8.

²⁴⁹ *Id.* (quoting Michael Wood, Special Rapporteur, Int’l Law Comm’n, Fourth Report on Identification of Customary International Law, ¶ 45, U.N. Doc. A/CN.4/695 (Mar. 8, 2016), <http://legal.un.org/docs/?symbol=A/CN.4/695> [<https://perma.cc/B35R-KNST>]).

²⁵⁰ *Id.* ¶¶ 36, 38. That said, accounts of executive conduct remain partly elusive, because some accounts are accessible only by physically accessing archives, some are confidential, and some military and intelligence conduct goes unreported. *Id.* ¶ 77.

information.²⁵¹ The Secretariat's recommendations focus on urging states and U.N. actors to pursue opportunities to make their practice available online and circulate it in different fora, but the report fails to identify the promise that text-as-data and related technologies offer in this regard.²⁵² In short, states should implement the report's recommendations but also actively consider ways in which text-as-data can help them identify their own and other states' practice in this vast informational sea.

III. RESOLVING INTERNATIONAL DISPUTES

Part I illustrated that data scientists have developed a range of tools to help lawyers navigate domestic litigation. Scholars, too, have applied machine learning tools to judicial opinions and briefs to identify trends and patterns in past cases and to predict the outcome of future cases. This Part argues that international lawyers can use similar tools to similar ends.²⁵³

A. *Understanding the Workings of Arbitral and Judicial Tribunals*

States find themselves before international courts and tribunals (as either claimants or respondents) in several ways. First, states may agree by treaty to refer disputes arising under that treaty to a non-specialized judicial body such as the ICJ.²⁵⁴ Second, a number of treaties establish adjudicatory tribunals to resolve disputes that arise specifically under those treaties, including the U.N. Convention on the

²⁵¹ *Id.* ¶ 94; *see also* Michael Wood, Special Rapporteur, Int'l Law Comm'n, Fourth Report on Identification of Customary International Law, ¶ 45, U.N. Doc. A/CN.4/695 (Mar. 8, 2016), <http://legal.un.org/docs/?symbol=A/CN.4/695> [<https://perma.cc/6UJW-46P8>] ("At the same time, the expanded number of States (and international organizations), the far greater volume of international intercourse, and the multiple formats of evidence now in existence, pose significant challenges to a thorough enquiry into the practice and *opinio juris* of States.").

²⁵² Compare, for example, the work by the Harvard Law School Library, whose CaseLaw Access Project digitized the entire historical record of U.S. court opinions, to enable legal algorithms to train on that data. *See* Winick, *supra* note 56.

²⁵³ *See* Sergio Puig et al., *The Data-Driven Future of International Law*, EJIL:TALK! (July 25, 2017), <https://www.ejiltalk.org/the-data-driven-future-of-international-law> [<https://perma.cc/D5W9-6CUH>] ("Data-driven empirical research promises to uncover latent patterns in international law data . . .").

²⁵⁴ *See, e.g.*, Vienna Convention for the Protection of the Ozone Layer art. 11, Mar. 22, 1985, 1513 U.N.T.S. 324 (allowing state at time of ratification to declare that it accepts as compulsory dispute resolution either arbitration or the ICJ); Optional Protocol to the Vienna Convention on Consular Relations Concerning the Compulsory Settlement of Disputes art. I, Apr. 24, 1963, 596 U.N.T.S. 487 (anticipating that parties may bring disputes about the interpretation or application of the Convention to the ICJ); *id.* art. II (allowing states to agree to send disputes to an arbitral tribunal); Vienna Convention on the Law of Treaties, *supra* note 130, at art. 66 (same).

Law of the Sea,²⁵⁵ the WTO,²⁵⁶ and the European Convention on Human Rights.²⁵⁷ Third, a state may agree in a bilateral investment treaty to allow investors from the other state party to bring claims against it using arbitration. Fourth, a state may attempt to resolve a dispute with another state using an administrative body such as the Permanent Court of Arbitration.²⁵⁸ There are thus a variety of fora in which a state might appear in order to resolve an interstate or investor dispute.

When appearing before a tribunal, it is helpful for states to know as much as possible about the adjudicators sitting on that tribunal, the past cases that the tribunal has heard, the prior decisions it has reached, and which kinds of arguments the tribunal has tended to find persuasive. States might want to bear down even further, trying to identify which judges or arbitrators have written particular (non-attributed) opinions and which arguments have appealed to which adjudicators. This information can inform how states select arbitrators in investor-state or interstate arbitration, influence which judicial forum they choose (if they have multiple options), guide how they shape their legal arguments, and potentially improve their chances of winning the case.

Domestically, lawyers and scholars are starting to use computational text analysis to extract this kind of information. One scholar has concluded that

any legal work that depends on collating and analyzing historical data such as past judicial decisions, including legal opinions or evaluating likely litigation outcomes, will become the dominion of AI. No human lawyer stands a chance against the formidable processing power of a mainframe when it comes to sifting through voluminous data.²⁵⁹

²⁵⁵ See U.N. Convention on the Law of the Sea art. 186, Dec. 10, 1982, 1833 U.N.T.S. 397 (establishing the International Tribunal on the Law of the Sea).

²⁵⁶ See Understanding on Rules and Procedures Governing the Settlement of Disputes art. 17, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 2, 1869 U.N.T.S. 401 (establishing the Appellate Body).

²⁵⁷ See Convention for the Protection of Human Rights and Fundamental Freedoms art. 19, Nov. 4, 1950, 213 U.N.T.S. 222 (establishing the European Court of Human Rights).

²⁵⁸ See Gary Born, *State-to-State Arbitration at the Permanent Court of Arbitration*, WOLTERS KLUWER (July 20, 2012), <http://arbitrationblog.kluwerarbitration.com/2012/07/20/state-to-state-arbitration-at-the-permanent-court-of-arbitration/> [<https://perma.cc/5L5N-E5QK>] (noting that in 2012 the Permanent Court of Arbitration was administering seven active state-to-state arbitrations, many of which involved boundary disputes).

²⁵⁹ Lauri Donahue, *A Primer on Using Artificial Intelligence in the Legal Profession*, HARV. J.L. & TECH. (Jan. 3, 2018), <https://jolt.law.harvard.edu/digest/a-primer-on-using-artificial-intelligence-in-the-legal-profession> [<https://perma.cc/WTY4-AQH5>].

Westlaw now provides “lawyers with statistics relating to the likelihood of a court or judge granting a motion, denying a motion, [or] ruling on a motion during a specific period”²⁶⁰ Other products also allow users to analyze the experience of particular lawyers before a judge or court, which may affect which lawyer the user chooses to hire.²⁶¹ This Section considers ways in which international lawyers could employ similar tools—some of which do not even rely on machine learning—to improve their practice before international tribunals.

1. *Extracting Information About Tribunal Decision-making*

There are a variety of ways in which states can extract information about how judicial and arbitral tribunals render their decisions. This information can help states develop strategies about when and how to litigate disputes, including by anticipating how a particular case is likely to come out. Producing accurate analyses requires ample data, however.²⁶²

a. *Sources of Data*

The relatively few international law scholars who have brought computational text analysis to bear on international law have tended to focus on international economic law and, in particular, the dispute resolution mechanisms associated with trade and investment.²⁶³ This is likely because there is a larger number of cases and arbitral awards from which to extract data than there is for judicial tribunals such as the ICJ.²⁶⁴ Another set of scholars has examined the jurisprudence of the European Court of Human Rights (“ECtHR”), which, like investment arbitration, has produced thousands of decisions.²⁶⁵ Successful

²⁶⁰ David Lat, *How Artificial Intelligence Is Transforming Legal Research*, ABOVE THE LAW, <https://abovethelaw.com/law2020/how-artificial-intelligence-is-transforming-legal-research/> [<https://perma.cc/4U7Q-P2GL>]; see also Steve Lohr, *A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet.*, N.Y. TIMES (Mar. 19, 2017), https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html?mcubz=0&_r=0 [<https://perma.cc/YC4T-3CW6>].

²⁶¹ Edgar Alan Rayo, *AI in Law and Legal Practice—A Comprehensive View of 35 Current Applications*, EMERJ (Feb. 27, 2020), <https://www.techemergence.com/ai-in-law-legal-practice-current-applications/> [<https://perma.cc/X8ZS-X96W>] (“Quantitative legal prediction already plays a significant role in certain [legal] practice areas and this role is likely to increase as greater access to appropriate legal data becomes available.” (quoting Katz, *supra* note 5, at 912)).

²⁶² Dave Gershgorin, *Can the NSA's Machines Recognize a Terrorist?*, POPULAR SCI. (Feb. 16, 2016), <http://www.popsci.com/nsas-skynet-might-not-be-able-to-tell-what-makes-terrorist> [<https://perma.cc/7Y3A-LUPX>].

²⁶³ See, e.g., Alschner et al., *supra* note 1.

²⁶⁴ See *infra* Section III.A.1.b.

²⁶⁵ See *infra* text accompanying notes 292–96.

predictive analytics requires a significant volume of data, so that states will be able to extract more statistically accurate information about tribunals that have produced a large quantum of decisions.

Scholars have assembled a number of databases that facilitate this kind of work. For example, Joost Pauwelyn and Wolfgang Alschner have created a database of economic law cases, drawn from sources such as the WTO and International Centre for Settlement of Investment Disputes.²⁶⁶ Daniel Behn and Malcolm Langford have collected and coded 800 cases involving investment treaty arbitration.²⁶⁷ Several sets of scholars have analyzed the ECtHR database, which makes its thousands of decisions publicly available.²⁶⁸ And others have assembled information about hundreds of arbitrators, including their backgrounds²⁶⁹ and how powerful they are.²⁷⁰ Although there is no evidence that states and their international lawyers have used these types of databases and text-as-data tools to help them structure their behavior before international tribunals, states that find themselves before these tribunals might use precisely these sorts of predictive analytics and computational text analysis to better navigate these disputes.

b. Parsing Past Opinions to Shape Future Arguments

One way in which international lawyers might use text classification tools to improve their case analysis and their states' position before tribunals is to identify, using "stylistic" tools, which decisionmakers wrote a particular opinion, award, or part thereof. These tools evaluate the "stylistic fingerprint[s]" of court opinions for various reasons, including to attribute authorship.²⁷¹ One study, for in-

²⁶⁶ Swiss Nat'l Sci. Found., *Big Data Conquers Legal Analysis*, PHYS.ORG, (Sept. 28, 2017), <https://phys.org/news/2017-09-big-conquers-legal-analysis.html> [<https://perma.cc/XML5-3YKW>].

²⁶⁷ Daniel Behn & Malcolm Langford, *Trumping the Environment? An Empirical Perspective on the Legitimacy of Investment Treaty Arbitration*, 18 J. WORLD INV. & TRADE 14, 16 (2017).

²⁶⁸ See, e.g., Nikolaos Aletras et al., *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*, PEERJ COMPUTER SCI. (Oct. 24, 2016), <https://doi.org/10.7717/peerj-cs.93> [<https://perma.cc/HG8F-H5S2>]; Masha Medvedeva et al., *Judicial Decisions of the European Court of Human Rights: Looking into the Crystal Ball*, UNIV. GRONINGEN, NETH. (2018).

²⁶⁹ See Puig, *supra* note 147.

²⁷⁰ See Langford et al., *supra* note 99.

²⁷¹ See, e.g., Keith Carlson et al., *A Quantitative Analysis of Writing Style on the U.S. Supreme Court*, 93 WASH. U. L. REV. 1461, 1463 (2016) (using quantitative methods to identify differences in writing styles by same judge depending on whether the judge was in the majority or dissenting); Li et al., *supra* note 53 (using quantitative methods to attribute authorship in unattributed Supreme Court opinions).

stance, used natural language processing tools trained on a database of signed Supreme Court decisions to identify the authorship of unsigned Court opinions.²⁷² In the arbitration context, scholars use stylometric tools to identify which arbitrators drafted particular decisions, and which arbitrators have more or less influence in the drafting process.²⁷³ Scholars also have subjected WTO Appellate Body reports to a stylometric analysis, concluding that WTO adjudicators have only a limited impact on their decisions' text, which suggests that the Secretariat plays a heavy role in drafting.²⁷⁴

From a state's perspective, knowing which judges or arbitrators have taken the lead in writing particular decisions can shed light on the types of issues and facts that are most important to that decisionmaker and on the legal arguments that she finds most persuasive. As Thore Neumann and Bruno Simma put it, "By obtaining information on the drafters, close observers of a court might—also in combination with the evaluation of any published dissenting opinions—gain insights into putative power structures and work division practices inside the courts."²⁷⁵ Further, knowing which arbitrators hold the most influence over the drafting of arbitral awards will help states identify which arbitrators they might want to select (or avoid).

It might be possible to successfully apply a stylometric analysis to the ECtHR's majority opinions, because ECtHR judges sign their concurrences and dissents and the court produces a high volume of decisions.²⁷⁶ That provides a relatively sizeable corpus of text on which to train an algorithm on particular judges' writing styles. This task would be much harder if directed at the ICJ's decisions, though the

272 See Li et al., *supra* note 53, at 510–11.

273 See Damien Charlotin, *Identifying the Voices of Unseen Actors in Investor-State Dispute Settlement*, in *LEGITIMACY OF UNSEEN ACTORS IN INTERNATIONAL ADJUDICATION* 392 (Freya Baetens ed., 2019). One potential challenge is that arbitrators may employ assistants and staff to help draft their opinions, which makes accurate stylistic analyses more challenging. Damien Charlotin has used stylometry to determine that in WTO adjudication, members of the Secretariat (rather than individual arbitrators) play an important role in drafting decisions. See Damien Charlotin, *Who Writes WTO Panel and AB Reports? A Tentative Stylometric Analysis*, MEDIUM (May 1, 2018) [hereinafter Charlotin, *Who Writes*], <https://medium.com/@damien.charlotin/who-writes-wto-panel-and-ab-reports-a-tentative-stylometric-analysis-565c18f0491d> [<https://perma.cc/ZL9Y-4TUS>]. In one case, Russia challenged an award against it based on a stylometric analysis that attributed most of the award's text to the Secretariat. *Id.*

274 See Charlotin, *Who Writes*, *supra* note 273.

275 Thore Neumann & Bruno Simma, *Transparency in International Adjudication*, in *TRANSPARENCY IN INTERNATIONAL LAW* 436, 460 (Andrea Bianchi & Anne Peters eds., 2013).

276 ROSA RAFFAELLI, EUR. PARL., *DISSENTING OPINIONS IN THE SUPREME COURTS OF MEMBER STATES* 32 (2012), [https://www.europarl.europa.eu/RegData/etudes/etudes/join/2012/462470/IPOL-JURI_ET\(2012\)462470_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/etudes/join/2012/462470/IPOL-JURI_ET(2012)462470_EN.pdf) [<https://perma.cc/7Y3A-LUPX>].

work would be of interest because the Court does not make public which judges sit on each three-judge drafting committee. The challenge in analyzing the ICJ's opinions is that the quantum of text on which to train a stylometric algorithm is small because the majority decisions are produced through a multi-author process, leaving us to observe only attributed concurrences and dissents.²⁷⁷ It might be possible, however, to use the judges' academic writings and speeches as well as judicial opinions from before their appointment to the ICJ to help conduct a stylometric analysis.

Beyond identifying authorship of opinions, states would benefit from understanding when courts have used language in their opinions that is very similar to that contained in parties' briefs. There are a variety of natural language processing tools (sometimes called "textual similarity tools") that enable researchers to identify whether a given legal decision has cribbed language from a party's submissions.²⁷⁸ For instance, using plagiarism detection software, Patricia Corley has found that the quality of a brief (which she defines by the level of experience of the attorney authoring that brief) directly impacts the Supreme Court's decision-making.²⁷⁹ The same may well hold true for briefs filed before a range of international tribunals. Indeed, Mark Daku and Krzysztof Pelc analyzed every state's submission in every WTO case to determine a submission's similarity to the contents of Appellate Body reports, and have used the levels of similarity as a proxy for parties' influence over Appellate Body jurisprudence.²⁸⁰

To the extent that a state has limited resources and must decide whether to allocate some of those resources to high quality legal representation in an international forum (by hiring an outside lawyer, say), it is useful for that state to understand the extent to which a given tribunal tends to borrow phrases and sentences from parties' briefs. The extent to which courts draw from parties' briefs may also inform a state's decision about whether to absent itself from a case—as the United States did in a case brought by Nicaragua in the ICJ,²⁸¹

²⁷⁷ See H.E. Judge Shi Jiuyong, President, Int'l Court of Justice, Speech to the Sixth Committee of the U.N. General Assembly: The Internal Working Practice of the International Court of Justice (Oct. 28, 2005), <https://www.icj-cij.org/files/press-releases/5/15815.pdf> [<https://perma.cc/QY5L-BAM2>] (explaining how the ICJ drafts its opinions).

²⁷⁸ See Corley et al., *supra* note 55; Corley, *supra* note 54.

²⁷⁹ See Corley, *supra* note 54, at 474.

²⁸⁰ See Mark Daku & Krzysztof J. Pelc, *Who Holds Influence over WTO Jurisprudence?*, 20 J. INT'L ECON. L. 233, 240–41 (2017).

²⁸¹ See *Military and Paramilitary Activities in and Against Nicaragua* (Nicar. v. U.S.), Jurisdiction and Admissibility, 1984 I.C.J. Rep. 392, ¶ 86 (Nov. 26).

Russia did in a maritime case brought by the Netherlands in the Permanent Court of Arbitration,²⁸² and China did in an arbitration proceeding brought by the Philippines related to the Law of the Sea Convention.²⁸³ In sum, the information revealed by these kinds of computational text analysis is useful not only for scholars but also for state litigants that appear before and submit arguments to these tribunals.

c. Predicting Outcomes

Perhaps the most ambitious use of text-as-data analytics is to help parties anticipate case outcomes. Law firms have started to use these tools to identify the relevant set of cases on which to train a predictive algorithm, and then to run new disputes through the algorithm to predict how courts would resolve them.²⁸⁴ Although the field of legal prediction is relatively new and some are skeptical of its analytical power, “big data and the novel means of its analysis create new opportunities for conducting causal inference studies and pave the way for more accurate predictions of the outcome of legal proceedings or negotiations.”²⁸⁵ If a state could predict in advance which way a tribunal is likely to resolve a particular dispute, it would allow the state to decide whether to pursue the case or to choose an alternative, such as settling it through diplomatic negotiations or dropping the matter entirely.

A handful of academic works have attempted to predict one or more aspects of decisions of the U.S. Supreme Court,²⁸⁶ the French Supreme Court,²⁸⁷ and the ECtHR.²⁸⁸ For the U.S. Supreme Court analysis, researchers built an algorithm to predict any given justice’s

²⁸² See *In re Arctic Sunrise Arbitration* (Neth. v. Russ.), PCA Case Repository No. 2014-02, Award on Jurisdiction, ¶ 18 (Perm. Ct. Arb. 2014).

²⁸³ See *In re S. China Sea Arbitration* (Phil. v. China), PCA Case Repository No. 2013-19, Award, ¶ 11 (Perm. Ct. Arb. 2016).

²⁸⁴ See Katz, *supra* note 5, at 957 (“In order to deliver optimal results, the retrieval or ‘pre-predicted’ set of comparison cases needs to include cases that share an analogical structure to the reference case. Since analogy is so powerful in law, its development cannot be ignored when it comes to the question of selecting the comparison group of cases. Modeling this sort of analogical reasoning is nontrivial, but it is not impossible.”).

²⁸⁵ Alschner et al., *supra* note 1, at 225.

²⁸⁶ See Ruger et al., *supra* note 9.

²⁸⁷ See Octavia-Maria Sulea at al., *Exploring the Use of Text Classification in the Legal Domain*, ASAIL (2017), <https://arxiv.org/pdf/1710.09306.pdf> [<https://perma.cc/D3VT-GS7V>] (using over 126,000 rulings, each of which contained a description of the case, the relevant area of law, the ruling’s date and substance, and a list of articles and laws cited in the description to predict the court’s ruling with over 98% accuracy and reliability).

²⁸⁸ See Aletras et al., *supra* note 268.

vote at any time, drawing on 16 features of each vote.²⁸⁹ The algorithm, which was able to achieve a prediction accuracy higher than subject matter experts, predicted correctly just over 70% of the Court's 28,000 decisions, and almost 72% of the Justices' votes.²⁹⁰ Nevertheless, because each legal case has different facts and postures, there is reason to be cautious about the ultimate efficacy of these tools.

Scholars have also attempted to predict decisions of the ECtHR.²⁹¹ A 2016 study using contiguous word sequences and abstract semantic topics claimed that it was 79% accurate at predicting case outcomes.²⁹² Building on that work, a different set of authors used supervised machine learning to predict ECtHR decisions.²⁹³ Those scholars trained the algorithm by providing text from many ECtHR cases, along with the judgments in those cases.²⁹⁴ When they ran their algorithm by providing the text of a case but no judgment, the computer was able to predict the judgment with around a 75% success rate.²⁹⁵ For international lawyers who must advise their governments how to manage cases brought against them in the ECtHR, these types of predictive algorithms could be very useful, not only in particular cases, but also to obtain a broader perspective on trends in the Court's jurisprudence over time. These lawyers might also use these tools to analyze the work of the Human Rights Committee, which considers and evaluates individual complaints alleging violations of the International Covenant on Civil and Political Rights ("ICCPR") by states parties to the ICCPR's Optional Protocol.²⁹⁶

Even if these tools ultimately prove themselves to be reliable, it would be impossible to replicate these types of predictions for judgments of international tribunals that produce only a few opinions each

²⁸⁹ Matthew Hutson, *Artificial Intelligence Prevails at Predicting Supreme Court Decisions*, AAAS (May 2, 2017, 1:45 PM), <http://www.sciencemag.org/news/2017/05/artificial-intelligence-prevails-predicting-supreme-court-decisions> [<https://perma.cc/M5A9-CXGA>]. The Supreme Court database contains high quality, expertly coded data that has taken years to assemble. *Id.*

²⁹⁰ *Id.*

²⁹¹ See Aletras et al., *supra* note 268.

²⁹² *Id.* at 1. One difficulty with this study is that it used the facts of a case to predict outcomes, without acknowledging that the judges writing the holding might well have shaped the facts to make that holding more persuasive. See *id.* at 4–5.

²⁹³ Medvedeva et al., *supra* note 268, at 2 (using language analysis and automatic information extraction).

²⁹⁴ *Id.* at 5.

²⁹⁵ *Id.* at 2, 5.

²⁹⁶ See *Human Rights Bodies - Complaints Procedures*, *Human Rights Bodies*, OFFICE HIGH COMM'R, U.N. HUMAN RIGHTS, <https://www.ohchr.org/EN/HRBodies/TBPetitions/Pages/HRTBPetitions.aspx> [<https://perma.cc/M5A9-CXGA>].

year. In comparison to the U.S. Supreme Court, which has issued around 28,000 opinions since 1791,²⁹⁷ the ICJ heard 177 cases between May 1947 and July 2019.²⁹⁸ The International Tribunal for the Law of the Sea (“ITLOS”) has heard even fewer cases: states have only submitted 29 cases to it since 1997.²⁹⁹ Although it might be possible to code data on ICJ or ITLOS judges using features such as nationality, rate of dissent, and how far apart a judge’s decisions are from the Court’s mean, the number of examples in the database will remain so small as to restrict the predictive capacity of any algorithm.³⁰⁰ In the arbitration context, however, where there are many more awards and the underlying legal issues are frequently repeatable, predictions based on the facts and legal issues in a case may be more accurate.³⁰¹

2. Arbitrating Algorithmically?

Technologically assisted dispute resolution is widely used in the private sector. For instance, eBay and PayPal employ online dispute resolution (“ODR”) for issues that arise between their platform’s

²⁹⁷ See Hutson, *supra* note 289.

²⁹⁸ See *Cases*, INT’L COURT OF JUSTICE, <https://www.icj-cij.org/en/cases> [https://perma.cc/NH5G-WPVT]. It has issued a somewhat greater number of opinions because it sometimes issues multiple opinions in the same case. See, e.g., *Military and Paramilitary Activities in and Against Nicaragua* (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14 (June 27); *Military and Paramilitary Activities in and Against Nicaragua* (Nicar. v. U.S.), Jurisdiction and Admissibility, 1984 I.C.J. 392 (Nov. 26).

²⁹⁹ *Cases*, INT’L TRIBUNAL FOR L. SEA, <https://www.itlos.org/cases/> [https://perma.cc/HZS2-BKQW].

³⁰⁰ The Inter-American Court of Human Rights (IACtHR) falls somewhere between the ECtHR and the ICJ in terms of volume of cases per year. In 2017, the Court issued 14 judgments, see INTER-AMERICAN COURT OF HUMAN RIGHTS, ANNUAL REPORT (2017), http://www.corteidh.or.cr/sitios/informes/docs/ENG/eng_2017.pdf [https://perma.cc/YK4C-XRZU], and in 2016, it issued 21 judgments, see INTER-AMERICAN COURT OF HUMAN RIGHTS, ANNUAL REPORT 52 (2016), http://www.corteidh.or.cr/sitios/informes/docs/ENG/eng_2016.pdf [https://perma.cc/L4CU-JDZW]. States (and petitioners) might, however, evaluate the likelihood of having a petition accepted by the Inter-American Commission for processing, having the Commission find against the state, or having the Commission refer the case to the Court, because the numbers of petitions and Commission decisions are much higher than the number of IACtHR decisions. See *Admissibility Reports*, INTER-AMERICAN COMM’N ON HUMAN RIGHTS, <http://www.oas.org/en/iachr/decisions/admissibilities.asp> [https://perma.cc/8TQ7-8S3V] (issuing over 100 admissibility reports in 2017 and 43 in 2016).

³⁰¹ One challenge here is that there are a large number of confidential decisions, which makes accurate prediction more difficult. See generally Emilie M. Hafner-Burton et al., *Against Secrecy: The Social Cost of International Dispute Settlement*, 42 YALE J. INT’L L. 279 (2017) (discussing lack of transparency in International Centre for Settlement of Investment Disputes arbitration cases).

users.³⁰² These systems sort through parties' preferences, keep communications focused on the key issues to be addressed, and help structure negotiations.³⁰³ Early producers of ODR tools encountered skepticism. According to Colin Rule, who helped develop eBay's ODR system, "[s]enior figures in the field told [him] that technology was not just a bad fit with dispute resolution, it was dangerous, because it suggested that there was a shortcut to solving problems and that parties could just click their way out of disagreements."³⁰⁴ Now, the U.N. Commission on International Trade Law has a Working Group on ODR, and a European Union regulation requires all of its members to implement ODR for cross-border consumer cases.³⁰⁵

Is there a possibility of using technology to help adjudicate disputes that involve not just private commercial actors or individuals, but also states? As was the case for initial ODR efforts, many will be skeptical. A number of arbitration scholars and practitioners, however, have written about the potential use of "machine arbitrators."³⁰⁶ Though that phrase sounds highly futuristic, the idea is that arbitrators may someday employ AI tools to propose settlement ranges and help draft awards, possibly using advanced versions of the types of ODR tools that companies are developing today.³⁰⁷ Paul Cohen and Sophie Nappert argue that

it is likely that on the current trajectory, a computer would be able to serve as a fact-finder and arbitrator on its own within the next two decades: sifting through the applicable law more thoroughly, assessing the credibility of witness evi-

³⁰² PABLO CORTES, *ONLINE DISPUTE RESOLUTION FOR CONSUMERS IN THE EUROPEAN UNION* 148–49 (2011).

³⁰³ See Colin Rule, *Technology and the Future of Dispute Resolution*, *DISP. RESOL. MAG.*, Winter 2015, at 4, 6, <http://www.colinrule.com/writing/drmag.pdf> [<https://perma.cc/LVT5-LH2V>].

³⁰⁴ *Id.*; see Colin Rule, *Making Peace on eBay: Resolving Disputes in the World's Largest Marketplace*, *ACRESOL.*, Fall 2008, <http://colinrule.com/writing/acr2008.pdf> [<https://perma.cc/8PBP-F38D>].

³⁰⁵ Rule, *supra* note 303, at 6.

³⁰⁶ See, e.g., David Allen Larson, *Artificial Intelligence: Robots, Avatars, and the Demise of the Human Mediator*, 25 *OHIO ST. J. DISP. RESOL.* 105 (2010); Lucas Bento, *International Arbitration and Artificial Intelligence: Time to Tango?*, *WOLTERS KLUWER* (Feb. 23, 2018), <http://arbitrationblog.kluwerarbitration.com/2018/02/23/international-arbitration-artificial-intelligence-time-tango/> [<https://perma.cc/55T6-P5HT>] (using term "machine arbitrators"); José María de la Jara et al., *Machine Arbitrator: Are We Ready?*, *WOLTERS KLUWER* (May 4, 2017), <http://arbitrationblog.kluwerarbitration.com/2017/05/04/machine-arbitrator-are-we-ready/> [<https://perma.cc/ME4Y-JF4K>].

³⁰⁷ See Bento, *supra* note 306.

dence more accurately, and deliberating a great deal more quickly than human arbitrators.³⁰⁸

In the nearer term, one can imagine more pragmatic possibilities for using text-as-data and machine learning tools to support arbitration. Text-as-data tools could, for instance, help arbitrators extract meaning from the large volumes of documentation that the parties submit to them (and also help opposing parties process that information).³⁰⁹ Arbitrators might also use algorithms to produce an alternative decision against which they could “check” their own decisions. In this sense, the algorithms “might therefore perform a role similar to the one for which tribunals use . . . tribunal secretaries,”³¹⁰ who in some cases partially draft and review awards for arbitrators.³¹¹

An international lawyer representing her state in arbitration will need to assess whether and under what circumstances her state would be comfortable accepting a significant role for machine learning tools in arbitral decision-making. These tools will potentially accelerate the process of arbitration, which will reduce parties’ costs,³¹² and may also help eliminate arbitrator bias.³¹³ States’ international lawyers will also need to assess whether to allow data scientists to access their prior confidential arbitral awards to help build accurate arbitral prediction tools.³¹⁴

Although we appear to be some distance away from actually resolving arbitral disputes using text-as-data and machine learning tools, some have already tested these tools in notional ways on com-

³⁰⁸ Paul Cohen & Sophie Nappert, *The March of the Robots*, GLOBAL ARB. REV. (Feb. 15, 2017), <https://globalarbitrationreview.com/article/1080951/the-march-of-the-robots> [<https://perma.cc/CMD4-UQPP>]; see also *The Future of Arbitration: New Technologies Are Making a Big Impact—and AI Robots May Take on “Human” Roles*, HOGAN LOVELLS (Feb. 21, 2018), <https://www.hoganlovells.com/publications/the-future-of-arbitration-ai-robots-may-take-on-human-roles> [<https://perma.cc/7TPQ-LC8D>] [hereinafter *The Future of Arbitration*] (quoting the chairman of the ICC International Court of Arbitration as suggesting that “having robot arbitrators may be considered acceptable” in 10 years’ time).

³⁰⁹ See *The Future of Arbitration*, *supra* note 308.

³¹⁰ Cohen & Nappert, *supra* note 308.

³¹¹ Michael Polkinghorne & Charles B. Rosenberg, *The Role of the Tribunal Secretary in International Arbitration: A Call for a Uniform Standard*, INT’L BAR ASS’N (Mar. 5, 2015), <https://www.ibanet.org/Article/NewDetail.aspx?ArticleUid=987d1cfc-3bc2-48d3-959e-e18d7935f542> [<https://perma.cc/MJZ4-CL2F>].

³¹² See Bento, *supra* note 306.

³¹³ Depending on the source of the data, machine learning arbitration tools may encode the prior biases of arbitrators on whose past judgments the tools were trained. See Matthew Hutson, *Even Artificial Intelligence Can Acquire Biases Against Race and Gender*, AAAS (Apr. 13, 2017, 2:00 PM), <https://www.sciencemag.org/news/2017/04/even-artificial-intelligence-can-acquire-biases-against-race-and-gender> [<https://perma.cc/76JF-DS2L>].

³¹⁴ Bento, *supra* note 306.

plicated foreign policy disputes. The next Section describes those tests and considers additional ways that text-as-data tools might help with dispute resolution that takes place in diplomatic, rather than judicial or arbitral, fora.

B. *Resolving Foreign Policy Disputes in Diplomatic Channels*

As noted in Section A, data scientists have produced tools that use AI in domestic settings to facilitate negotiations among individuals and to help those actors achieve agreement on a deal or settlement.³¹⁵ For example, individuals have employed ODR to facilitate divorces.³¹⁶ The ODR system walks the pair through the critical decisions they must make, allows each member of the couple to define his or her “optimal outcomes,”³¹⁷ and suggests a compromise solution based on their answers.³¹⁸

In theory, tools like these could be used as decisional support to resolve foreign policy disputes.³¹⁹ Imagine a system wherein each state ranks its preferences for resolving a dispute that implicates four points of tension (confidentially, using encryption) and the system spits out a division of interests that is Pareto-optimal. This obviously would be a complicated process, especially where a state needed to decide and input its preferences among and between provisions (i.e., identifying whether the state would rather receive X and not Y, or Y and not X, or partial X and partial Y, and if so what parts of X and Y).³²⁰ Further, these algorithms would have to be transparent in order to persuade

³¹⁵ See *supra* Section III.A; cf. Ophir Samson, *Deep Learning Weekly Piece: Negotiating Business Deals with AI*, MEDIUM (June 25, 2017), <https://medium.com/@ophir.samson02/deep-learning-weekly-piece-negotiating-business-deals-with-ai-a614c3c34e44> [https://perma.cc/2EB7-RU34] (discussing how an AI negotiator did better than humans at maximizing its negotiating scores).

³¹⁶ See Morikawa, *supra* note 62.

³¹⁷ Bernard Marr, *How AI and Machine Learning Are Transforming Law Firms and the Legal Sector*, FORBES (May 23, 2018, 12:29 AM), <https://www.forbes.com/sites/bernardmarr/2018/05/23/how-ai-and-machine-learning-are-transforming-law-firms-and-the-legal-sector/#759c58a932c3> [https://perma.cc/CB2U-UYBP].

³¹⁸ See Morikawa, *supra* note 62 (discussing the Wevorce algorithm).

³¹⁹ These algorithms do not require machine learning; they could rely on the kind of algorithm used in medical residency or New York City school matching, which seeks to determine how a constellation of rational individuals would choose an allocation of outcomes. See ROYAL SWEDISH ACADEMY OF SCIENCES, *STABLE MATCHING* (2012), <https://www.nobelprize.org/uploads/2018/06/popular-economicsciences2012.pdf> [https://perma.cc/8SJD-4364].

³²⁰ See Daniel Druckman et al., *Modeling International Negotiation*, in PROGRAMMING FOR PEACE 248 (Robert Trappl ed., 2006) (“[T]he machine-learning analyses add knowledge about contingent relations among variables . . . [and] produce[] a refined elucidation of the connections among factors that influence outcomes.”).

each state to accept the integrity of the system's solution.³²¹ They would have to be truly secure; the states involved might devote substantial resources to trying to hack such a system. And it would be difficult to design the algorithm in a way that prevented parties from gaming it.³²²

Most persistent foreign relations problems seem too complicated to be resolved by algorithm, at least at the current stage of technology. That said, some scholars have used "negotiation support systems" to try to notionally resolve complex foreign relations problems. In one effort, the authors used an algorithm called "Asset Divider" to enter data about the Israeli-Palestinian conflict, allocate a value to each disparate issue in dispute, and obtain an allocation of assets.³²³ Their results were not dissimilar to the arrangement struck at Camp David in 1978, though the authors had not used that arrangement as a model for their algorithm.³²⁴ Another group used an algorithm called "Adjusted Winner" to notionally allocate rights and resources among the several states making claims over the Spratly Islands in the South China Sea.³²⁵ Further, the creator of a tool called "SmartSettle" simulated how Iraq, Turkey, and Syria could negotiate water quantity and quality issues in the Tigris-Euphrates watershed,³²⁶ and another group developed an automated agent that was able successfully to negotiate with human players during a simulated bilateral crisis between Spain and Canada over fisheries access.³²⁷ At this stage, however, online dispute resolution performs best when confronted with monetary dis-

321 For a discussion of the role of explainable machine learning algorithms, see Ashley Deeks, *The Judicial Demand for Explainable AI*, 119 COLUM. L. REV. 1829 (2019).

322 See Jane Bambauer & Tal Zarsky, *The Algorithm Game*, 94 NOTRE DAME L. REV. 1 (2018).

323 John Zeleznikow & Emilia Bellucci, *Using Asset Divider to Investigate the Israel-Palestinian Dispute*, in PROCEEDINGS OF THE 6TH INT'L WORKSHOP ON ONLINE DISPUTE RESOLUTION 37 (2010), <http://ceur-ws.org/Vol-684/paper4.pdf> [<https://perma.cc/B3MB-F3HQ>]; see also Steven J. Brams & Jeffrey M. Togman, *Camp David: Was the Agreement Fair?*, in CONFLICT IN WORLD POLITICS 306, 306-07 (Frank P. Harvey & Ben D. Mor eds., 1998); Tansa George Massoud, *Fair Division, Adjusted Winner Procedure (AW), and the Israeli-Palestinian Conflict*, 44 J. CONFLICT RESOL. 333, 335 (2000).

324 Zeleznikow & Bellucci, *supra* note 323, at 48.

325 See David B.H. Denoon & Steven J. Brams, *Fair Division: A New Approach to the Spratly Islands Controversy*, 2 INT'L NEGOT. 303 (1997). "The Spratly Islands are a group of about 230 small islands and reefs in the South China Sea" believed to have major oil and gas deposits on or near them. *Id.* at 303.

326 UNESCO-IHE Recruits Dr. Ernest Thiessen to Teach Smartsettle for Negotiation and Mediation in Water Conflict Management, SMARTSETTLE (Feb. 4, 2010), <https://smartsettle.com/2010/02/04/smartsettle-for-water-conflict-management-3/> [<https://perma.cc/5CMY-7572>].

327 See Sarit Kraus et al., *Resolving Crises Through Automated Bilateral Negotiations*, 172 ARTIFICIAL INTELLIGENCE 1 (2008).

putes, particularly where there is a high volume of low stakes disputes—the opposite of what most interstate disputes look like.³²⁸

In the future, states themselves might use machine learning algorithms to model complex negotiations, helping them predict their negotiating partners' strategies and, in so doing, improving their own negotiating positions. A recent Chatham House report concluded that “parties to negotiations (whether economic or strategic in nature) might use sophisticated machine-learning methods to forecast others' positions and tactics.”³²⁹ In the iterative context of war-fighting, where one side makes a move and the other responds, keeping in mind the likely response of the first side, the U.S. military has begun to employ artificial intelligence tools to improve U.S. strategic responses.³³⁰ Negotiators might adopt similar technology to game out when and how to make various proposals, as well as what content to include in those proposals.³³¹

Assuming most international lawyers would deem it a bridge too far to employ these types of tools today to allocate disputed assets or territory among disputing parties and to persuade the parties to accept those allocations, machine learning tools nevertheless have the potential to help resolve persistent geopolitical disputes by locating past proposals and anticipating the public reaction to new proposals.

³²⁸ See Louis F. Del Duca et al., *eBay's De Facto Low Value High Volume Resolution Process: Lessons and Best Practices for ODR Systems Designers*, 6 Y.B. ARB. & MEDIATION 204 (2014).

³²⁹ Parakilas & Bryce, *supra* note 16, at 3 (cautioning that predictive algorithms “are not yet necessarily more accurate than their human equivalents”); see Ashley Deeks et al., *Machine Learning, Artificial Intelligence, and the Use of Force by States*, 10 J. NAT'L SEC. L. & POL'Y 1, 6–7 (2019) (discussing use of AI to “war game” enemy state's moves).

³³⁰ See Benjamin Jensen et al., *Wargaming with Athena: How to Make Militaries Smarter, Faster, and More Efficient with Artificial Intelligence*, WAR ON THE ROCKS (June 5, 2018), <https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries-smarter-faster-and-more-efficient-with-artificial-intelligence/> [<https://perma.cc/Z435-UHTN>]. See generally Selmer Bringsjord & Naveen Sundar Govindarajulu, *Artificial Intelligence*, in STANFORD ENCYCLOPEDIA OF PHILOSOPHY § 5 (Edward N. Zalta ed., 2018), <https://plato.stanford.edu/archives/fall2018/entries/artificial-intelligence/> [<https://perma.cc/A83B-MACF>] (describing use of AI in real time strategy games).

³³¹ In 2018, the International Chamber of Commerce and IBM Watson developed a “Cognitive Trade Advisor,” software that supported diplomats preparing for WTO negotiations. DIPLOFOUNDATION, *supra* note 102, at 26. The system answered questions about rules of origin across a wide range of trade agreements to help identify what duties or restrictions already existed in cross-border trade. *Id.*

1. *Locating Third Party Proposals*

A host of interstate disputes have endured for decades. Many of these disputes arise because two or more states contest control over particular territory; others flow from dissatisfaction by one part of a state's population about being under that state's control. Examples include disputes about Morocco's sovereignty over Western Sahara;³³² the sovereignty of Nagorno-Karabakh;³³³ the ownership of Northern Cyprus;³³⁴ the ownership of islands and smaller features in the East China Sea;³³⁵ control over Kashmir;³³⁶ and the Israeli/Palestinian conflict.³³⁷ Some of these disputes lead to armed conflicts, such as the conflicts fought in the Balkans between 1991 and 1999.

In many cases, the parties involved in the dispute (or the United Nations) ask diplomats, supported by international lawyers, to mediate those disputes and bring them to a sustainable resolution.³³⁸ One of the biggest challenges for international lawyers and diplomats is to identify and develop creative solutions that each side can accept politically. Because many of these disputes have endured for decades, a wide range of statesmen, governments, think tanks, non-governmental organizations, scholars, bloggers, and U.N. bodies have explored and written about ways to conclude each of the disputes set forth above.³³⁹

³³² Conor Gaffey, *Western Sahara: What Is the 40-Year Dispute All About?*, NEWSWEEK (Mar. 9, 2016, 3:41 PM), <https://www.newsweek.com/western-sahara-morocco-algeria-polisario-front-435170> [<https://perma.cc/9Q6W-KGMS>].

³³³ *Global Conflict Tracker: Nagorno-Karabakh Conflict*, COUNCIL ON FOREIGN REL. (Mar. 31, 2020), <https://www.cfr.org/interactive/global-conflict-tracker/conflict/nagorno-karabakh-conflict> [<https://perma.cc/V3UR-6UGP>].

³³⁴ Sewell Chan, *Cyprus: Why One of the World's Most Intractable Conflicts Continues*, N.Y. TIMES (Nov. 7, 2016), <https://www.nytimes.com/2016/11/08/world/europe/cyprus-reunification-talks.html> [<https://perma.cc/J2LK-4X63>].

³³⁵ *Global Conflict Tracker: Tensions in the East China Sea*, COUNCIL ON FOREIGN REL. (Mar. 31, 2020), <https://www.cfr.org/interactive/global-conflict-tracker/conflict/tensions-east-china-sea> [<https://perma.cc/EU8N-ZYMF>].

³³⁶ *Global Conflict Tracker: Conflict Between India and Pakistan*, COUNCIL ON FOREIGN REL. (Mar. 31, 2020), <https://www.cfr.org/interactive/global-conflict-tracker/conflict/conflict-between-india-and-pakistan> [<https://perma.cc/9LBE-V4MW>].

³³⁷ *Global Conflict Tracker: Israeli-Palestinian Conflict*, COUNCIL ON FOREIGN REL. (Mar. 31, 2020), <https://www.cfr.org/interactive/global-conflict-tracker/conflict/israeli-palestinian-conflict> [<https://perma.cc/D4BD-GCPC>].

³³⁸ See, e.g., *Prevention and Mediation*, U.N. POLITICAL AND PEACEBUILDING AFFAIRS, <https://dppa.un.org/en/prevention-and-mediation> [<https://perma.cc/RG46-48V3>] (noting that envoys are often "dispatched to crisis areas to encourage dialogue, compromise and the peaceful resolution of tensions").

³³⁹ See, e.g., Steven A. Cook, *What to Read on the Middle East Peace Process*, FOREIGN AFF. (July 22, 2009), <https://www.foreignaffairs.com/articles/middle-east/2009-07-22/what-read-middle-east-peace-process> [<https://perma.cc/AE6M-8JLJ>] (stating that there are "literally

This provides an opening for the use of machine learning: tools such as data mining can help comb through a huge variety of digital sources (which will appear in a number of languages) to identify proposed solutions to these seemingly intractable foreign relations problems.³⁴⁰ For instance, using open information extraction, data scientists could develop algorithms to search every idea that anyone ever has proposed to advance the Middle East Peace Process (or another more obscure dispute) and scrape and summarize the information that the algorithm identifies. This is a way to ensure that negotiators have the widest possible set of possible options before them as they work with the parties to develop sustainable solutions.

2. *Parsing Local Reactions*

Another key aspect of resolving international disputes is anticipating the public reaction to particular solutions. Those lawyers and diplomats who are tasked with negotiating the resolution of a dispute—either as a national of a state involved in the dispute or as a third party mediator—must be keenly attuned to the political constraints on the disputing parties and the particular cultural or issue-specific sensitivities that each party's public has.³⁴¹ The negotiators therefore would benefit from being able to understand and analyze the likely reactions of the disputing states' citizens to various options.

One way to do this, of course, is for the negotiators to consult their own diplomats, some of whom may be stationed in one of the state parties to the dispute. That will remain an important source of information for the foreseeable future, but it is no longer the only source. These days, it is much easier to assess how groups of citizens feel about particular issues because so much information (including news) is available online and on social media.³⁴² Assume States X and Y are in a heated dispute about ownership and allocation of water in a river that passes through both states. State X is thinking about propos-

thousands of books—and many more thousands of articles” devoted to the Arab-Israeli conflict).

³⁴⁰ See generally Christina Niklaus et al., *A Survey on Open Information Extraction*, in PROCEEDINGS OF THE 27TH INT'L CONF. ON COMPUTATIONAL LINGUISTICS (2018), <http://aclweb.org/anthology/C18-1326> [<https://perma.cc/V6SR-U2EE>] (identifying and extracting meaningful text related to complex assertions from any domain on the web); Bringsjord & Govindarajulu, *supra* note 330, § 4.1 (describing data mining as a form of unsupervised machine learning).

³⁴¹ See Robert D. Putnam, *Diplomacy and Domestic Politics: The Logic of Two-Level Games*, 42 INT'L ORG. 427, 434 (1988) (discussing role of domestic political pressure on international negotiations).

³⁴² This is feasible for states with a free press; highly autocratic states will carefully control what appears on social media.

ing an agreement to State Y pursuant to which the two states would receive equal amounts of water from the river. State X will want to understand the perceptions of its own citizens (or at least its politically active citizens) toward State Y, including how trustworthy State X's citizens believe State Y to be.³⁴³ State X will also need to be attuned to how concerned its own citizens are about specific issues related to water allocation, such as their worries about drought, new anticipated water needs, and environmental protection.

Just as State X will want to understand how receptive its own citizens will be to the negotiations, each state would also benefit from understanding where the other state's sensitivities lie. In a democracy, State X's sensitivities are likely to track the sensitivities of its citizens. Thus, State Y, for instance, would benefit from understanding what conversations have begun to take place in State X about water-related issues and possibly about the contemplated agreement in particular.

Today, scholars have begun to use "automated events data," using computer programs to "read" news reports, extract information from those reports, and use the reports to produce quantitative data about events such as the provision of military aid, demobilization of armed forces, requests for economic assistance, and the halting of negotiations.³⁴⁴ These programs rival Harvard undergraduates in their ability to extract information from Reuters news reports, and the algorithms are able to process far more reports without getting bored or tired.³⁴⁵ Other useful tools might include those developed to scan social media to gauge sentiment. In short, the tools that diplomats could use to assess and anticipate domestic reactions to different possible dispute resolution outcomes already exist and need only be adapted to the particular subject of the dispute at issue.

3. *Manipulating Local Reactions*

Of course, technology often has a dark side. Just as states or third party mediators trying to resolve international disputes could use text-as-data tools to better understand their negotiating partners' local constraints with the goal of reaching a more palatable and viable solu-

³⁴³ See Odell & Tingley, *supra* note 173, at 149 ("Committing to fulfill an agreement in the future is often crucial to the success of a negotiation.").

³⁴⁴ See King & Lowe, *supra* note 118, at 619; Devin Coldewey, *DARPA Wants to Build an AI to Find the Patterns Hidden in Global Chaos*, TECHCRUNCH (Jan. 7, 2019, 3:44 PM), <https://techcrunch.com/2019/01/07/darpa-wants-to-build-an-ai-to-find-the-patterns-hidden-in-global-chaos/> [<https://perma.cc/2TYZ-9Z6G>] (noting that DARPA is creating a machine learning system that can sift through daily media feeds and identify threads of connection in them).

³⁴⁵ See King & Lowe, *supra* note 118, at 619.

tion, so too could disputing states (or groups within states, or third states that want to undercut a potential deal) seek to alter or influence those same local reactions. As became apparent in the 2016 U.S. presidential election, states such as Russia have a variety of technological tools with which to create “fake news” and foster internal tensions among groups within a particular country.³⁴⁶

In the interstate dispute context, it would be easy for a state to use technological tools to attempt to sway local populations in favor of or against proposals to resolve the conflict. For instance, a state could nefariously spread false news reports that negotiations were moving in a particular direction unfavorable to State X, and that State X’s leadership planned to make concessions adverse to State X’s population. That alone could be enough to stimulate a local population to agitate not only against those particular (fake) concessions but also against the negotiations entirely. Moving beyond the creation of “fake news,” states could circulate on the Internet video or audio forgeries that could deeply affect those populations. Artificial intelligence has advanced to the point where malicious actors are quite easily able to falsify video and audio using software that is widely available.³⁴⁷ These forgeries could range from a video of the two heads of the negotiation delegations “colluding” to strike a deal that strongly favors one side, to a video in an unsavory location which one leader seems to pay another, to an audio of one negotiator denigrating the other state’s negotiating team.³⁴⁸ Even though technologies exist to help detect audio and video forgeries,³⁴⁹ scientific proof of falsification often is not enough to assuage inflamed populations.³⁵⁰ International lawyers will play a role in advising their states about which techniques would be permissible under international law, and which would run afoul of in-

³⁴⁶ Scott Shane, *The Fake Americans Russia Created to Influence the Election*, N.Y. TIMES (Sept. 7, 2017), <https://www.nytimes.com/2017/09/07/us/politics/russia-facebook-twitter-election.html> [<https://perma.cc/H5Q9-KE46>].

³⁴⁷ See Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CAL. L. REV. 1753, 1757 (2019); Raahat Devender Singh & Naveen Aggarwal, *Video Content Authentication Techniques: A Comprehensive Survey*, 24 MULTIMEDIA SYS. 211, 211–12 (2018).

³⁴⁸ For other ways that actors might use deep fakes of world leaders, see SHRUTI AGARWAL ET AL., PROTECTING WORLD LEADERS AGAINST DEEP FAKES, <https://farid.berkeley.edu/downloads/publications/cvpr19/cvpr19a.pdf> [<https://perma.cc/2M3T-LS3E>].

³⁴⁹ See, e.g., Singh & Aggarwal, *supra* note 347, at 212–13.

³⁵⁰ Ali Breland, *The Bizarre and Terrifying Case of the “Deepfake” Video that Helped Bring an African Nation to the Brink*, MOTHER JONES (Mar. 15, 2019), <https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/> [<https://perma.cc/835X-CUPM>].

ternational law prohibitions such as those on non-intervention or the promotion of war propaganda.³⁵¹

In sum, there are a range of ways in which states could begin to use text-as-data and machine learning tools to assist them in navigating international tribunals and shaping their legal arguments, and a variety of ways in which states could apply these types of tools to diplomatic negotiations over thorny foreign relations problems. While many of these tools could be employed to effectively resolve disputes, states with more nefarious intent could use some of these same tools to complicate or derail those efforts at resolution.

IV. CAN TECH STRENGTHEN INTERNATIONAL LAW?

Parts II and III described and analyzed the myriad ways in which states and their international lawyers could pursue the use of text-as-data tools to create international agreements, adjudicate disputes, or unearth state practice and *opinio juris*, thereby affirming or calling into question certain rules of CIL. As international lawyers consider pursuing (or are presented with the opportunity to develop) text-as-data tools, they will confront a host of technological and ethical challenges. This Part argues that states pursuing text-as-data tools should draw lessons from the technological and ethical problems that have accompanied machine learning in other settings, and address those problems before deploying these tools.

Further, these tools, once developed and adopted, will carry a host of strategic implications at both the international and domestic levels. This Part explores the distributional power effects of different types of text-as-data tools and concludes that some of these tools have the potential to enhance the development and operation of international law, while others have the potential to impede it. The fact that some states enjoy economic, technological, or structural advantages over other states during negotiations or adjudication is nothing new. If, however, states that lack these new high-tech tools feel disempowered enough that they choose to avoid negotiations or dispute settlement with states that use these tools, the tools could adversely affect the daily work of international law. The Part concludes by suggesting ways to minimize possible impediments to international law's

³⁵¹ See, e.g., International Covenant on Civil and Political Rights, art. 20, Dec. 19, 1966, 999 U.N.T.S. 171 (requiring states parties to prohibit war propaganda); *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14, ¶ 246 (June 27) (discussing customary rule of non-intervention).

functioning that might arise from the proliferation of text-as-data tools in this space.

A. *Technological and Ethical Challenges*

Machine learning has faced a range of critiques when government actors have used it in other settings, such as when judges have used predictive algorithms to assist them in criminal justice sentencing, and police have used facial recognition software to identify suspects.³⁵² The critiques include bias, error, a lack of transparency about how the models operate, and an overreliance on machines at the expense of human expertise. This Section considers how some of these underlying machine learning challenges might manifest themselves in the international law setting.

First, biases embedded in the data on which data scientists train machine learning algorithms will replicate themselves in the algorithms' predictions. Data scientists used billions of words on the Internet to build an algorithm to associate pairs of words. They discovered that the algorithm learned that words such as "librarian" and "hygienist" were strongly correlated to words like "female" and "woman."³⁵³ Likewise, coders may build their own cultural biases into the models: a U.S. coder creating a program to assess individual dangerousness of detainees in Yemen might treat weapon carrying as a salient factor, when in fact most adult males in Yemen carry weapons.³⁵⁴ These biases can render algorithms inaccurate, to the detriment of the state using them.

Certain text-as-data tools for international law might also embed and perpetuate biases. Assume, hypothetically, that wealthier countries tend to win in arbitrations against poorer countries because the arbitrators are inclined to think that the wealthier countries are in a better position to highly praise (or harshly criticize) the arbitrators after an award. A tool that predicts the likely outcome of an arbitration will embed that improper arbitrator bias and might unfairly deter a poorer country from pursuing a legitimate case against a wealthier country. The tools also are likely to reflect a status quo bias, because

³⁵² Karen Hao, *AI Is Sending People to Jail—And Getting It Wrong*, MIT TECH. REV. (Jan. 21, 2019), <https://www.technologyreview.com/s/612775/algorithms-criminal-justice-ai/> [<https://perma.cc/2BR7-4PX2>].

³⁵³ See Hutson, *supra* note 313.

³⁵⁴ See DEREK B. MILLER, DEMAND, STOCKPILES, AND SOCIAL CONTROLS: SMALL ARMS IN YEMEN 28 (2003), <http://www.smallarmssurvey.org/fileadmin/docs/B-Occasional-papers/SAS-OP09-Yemen.pdf> [<https://perma.cc/Z7C5-SLCN>] (estimating a weighted average of 1.26 weapons per man in Yemen).

they are trained on historical data, and so will be of less use when states seek to break new ground in a negotiation.

A second concern about machine learning is that it can be less reliable than its creators acknowledge. For instance, a system may make inaccurate predictions because the algorithm was trained on incomplete data sets, the data itself contained errors, or the creator made faulty statistical assumptions.³⁵⁵ Emotion detection software, for instance, may be trained on and therefore work better on Caucasian faces than faces of other races and facial expressions from non-Western cultures.³⁵⁶ A state's international lawyers, who already are conservatively inclined, will rightfully be hesitant to pursue tools that are not trustworthy.³⁵⁷

Another persistent challenge with machine learning algorithms is non-transparency: it is often difficult to determine how or why an algorithm made the prediction it did. This is referred to as the “black box” problem.³⁵⁸ Data scientists are making inroads into this problem by developing a host of ways to create “explainable AI,”³⁵⁹ which can help detect bias and reduce error. Nevertheless, it is an issue to which international lawyers must pay attention: lawyers providing advice to their clients in reliance on text-as-data tools will want to assure themselves that they understand how the algorithms reached the recommendations they did.

A fourth concern with algorithms is that their users can become overly accustomed to relying on algorithmic recommendations, even if their own experiences suggest that the recommendation is flawed.³⁶⁰ “Automation bias” can arise because people become complacent or believe that the machine has taken into account more considerations than they themselves have.³⁶¹ This bias could occur with text-as-data systems. If a text-as-data system analyzes a large number of social media posts and suggests that X is not an achievable outcome in a treaty negotiation, negotiators might abandon provision X and thus make an

355 Ashley S. Deeks, *Predicting Enemies*, 104 VA. L. REV. 1529, 1564 (2018).

356 Meredith Somers, *Emotion AI, Explained*, MIT SLOAN SCH. MGMT. (Mar. 8, 2019), <https://mitsloan.mit.edu/ideas-made-to-matter/emotion-ai-explained> [<https://perma.cc/MA38-4MX7>].

357 See *supra* notes 100–02 and accompanying text.

358 See FRANK PASQUALE, *THE BLACK BOX SOCIETY* 3 (2015).

359 See Deeks, *supra* note 321.

360 See Kate Goddard et al., *Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators*, 19 J. AM. MED. INFO. ASS'N 121 (2012), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3240751/> [<https://perma.cc/KZ4J-693K>].

361 *Id.*

important concession to the other side, even if they themselves think it may be feasible. All of these persistent challenges with machine learning have the ability to impact legal text-as-data tools; states and their international lawyers should address these challenges before deciding to adopt the tools.

Finally, states will need to consider the role of any private companies when they turn to those companies to help them develop these high-tech tools. To create some of these tools, states may need to give companies access to protected or sensitive information, which places companies in a position to manipulate or falsify that information in creating the algorithm. The role of companies and the ways that they may use (or misuse) data is beyond the scope of this Article, but states acquiring these tools should assure themselves that their data—and that of their citizens—is not being used inappropriately by these non-governmental actors.

B. High-Tech Tools' Distributional Effects

As with any tool that increases the knowledge and capabilities of some—but not all—states, these developments have the potential to affect interstate power distributions. Today, as has been true for decades, some states bring far more resources to the negotiating table than others. Certain states have greater financial, intelligence, and natural resources, personnel, or military power than others—and those disparities affect the conduct and outcomes of negotiations and dispute resolution. Access to high-tech tools is another type of resource that can affect those sorts of international interactions.

Researchers are just beginning to explore how AI and related technologies could alter these international power dynamics, so there is limited literature on the question.³⁶² A recent Chatham House report that explores the potential disruptions AI may inflict on select areas of international affairs notes that “[t]echnological change does not have to be dramatic or sudden to create meaningful shifts in power balances or social structures.”³⁶³ In the Chatham House view, *analytical* uses of AI are likely to have only an attenuated impact on the power dynamics among states, whereas *predictive* uses of AI

³⁶² See, e.g., *Artificial Intelligence and Global Security Initiative Research Agenda*, CNAS, <https://www.cnas.org/artificial-intelligence-and-global-security-initiative-research-agenda> [<https://perma.cc/YV3C-PY8T>].

³⁶³ CHATHAM HOUSE, *supra* note 10, at iv; see Bryce & Parakilas, *supra* note 10, at 44 (noting that AI “may be divisive at a national level as governments struggle to get to grips with the range of possibilities that this technology offers”).

“could, if such systems become sufficiently accurate and trusted, create a power gap between those actors equipped with such systems and those without—with notably unpredictable results.”³⁶⁴ Many of the text-as-data tools discussed herein would be used for analytical purposes, though some edge into areas of prediction. As those tools become more sophisticated, there is a risk that they could “deepen already pronounced inequalities between advanced and less developed economies.”³⁶⁵

Certain factors will dictate whether and when these tools and technologies are likely to disperse power among states or consolidate power in a few states. The public availability of the tools, the cost and time required to assemble and prepare the underlying models and the data on which those models rely, and the types of international negotiations or dispute resolution being undertaken will all produce different distributional outcomes. How states use these tools in the short term might affect whether they seek to regulate the use of these technologies in the longer term.

1. *Distributional Effects Among States*

Perhaps the most important factor that will dictate how these text-as-data tools will redistribute power among states is the accessibility of the tools themselves. Some of the tools previously discussed are widely available and free to use.³⁶⁶ A second set of tools is not secret—the individuals creating the algorithms or text-as-data tools have made the code publicly available³⁶⁷—but it will require time, money, and expertise for a state to adjust the code or train the tools for a state’s particular uses and to put the relevant data into a useable

³⁶⁴ CHATHAM HOUSE, *supra* note 10, at v.

³⁶⁵ Bryce & Parakilas, *supra* note 10, at 44; *see also* SCHARRE & HOROWITZ, *supra* note 184, at 3 (“Nations with access to the best data, computing resources, human capital, and processes of innovation are poised to leap ahead in the era of artificial intelligence.”).

³⁶⁶ *See* David Kelnar, *The Fourth Industrial Revolution: A Primer on Artificial Intelligence (AI)*, MEDIUM (Dec. 2, 2016), <https://medium.com/mmc-writes/the-fourth-industrial-revolution-a-primer-on-artificial-intelligence-ai-ff5e7ffcae1> [<https://perma.cc/7ZQJ-3P6M>] (“Google Machine Learning offers easily accessible services for: vision (object identification, explicit content detection, face detection and image sentiment analysis); speech (speech recognition and speech-to-text); text analysis (entity recognition, sentiment analysis, language detection and translation); and employee job searching (opportunity surfacing and seniority-based matching). Microsoft Cognitive Services includes more than 21 services within the fields of vision, speech, language, knowledge and search.”).

³⁶⁷ *See, e.g.*, Kyle Wiggers, *Google Open-Sources BERT, a State-of-the-Art Pretraining Technique for Natural Language Processing*, VENTUREBEAT (Nov. 2, 2018, 12:40 PM), <https://venturebeat.com/2018/11/02/google-open-sources-bert-a-state-of-the-art-training-technique-for-natural-language-processing/> [<https://perma.cc/78SL-HD34>].

format. Some states may develop a third category of tools that is either proprietary or produced in secret and so will remain (at least in the short term) in the hands of only a few states. Therefore, some of these tools have the potential to level the playing field for states. Others are likely to enhance existing power differentials. Indeed, they might empower autocratic states, which are willing to ignore human rights and civil liberties to gather data, more than they empower democratic states, which generally heed those limitations.

Tools that fall into the first category include the translation and interpretation tools discussed previously.³⁶⁸ Indeed, these tools will likely deliver oversized benefits to international lawyers from less powerful states because those lawyers are less likely to be fluent in one of the six U.N. languages.³⁶⁹ Further, these tools seem likely to improve at a relatively fast pace because they are widely used outside the international law setting and the volume of users enables the AI that undergirds the tools to learn quickly.³⁷⁰ Another tool that might create a leveling effect is the use of algorithms by arbitrators to help them decide disputes. This might also level the playing field for states with fewer resources (assuming the algorithm is neutral), because it will save states significant arbitration costs.³⁷¹

The more predictive text-as-data tools likely fall into the second category. This includes tools to help states predict the outcomes of court cases,³⁷² the persuasive power of a given arbitrator,³⁷³ or the likely outcome of a treaty negotiation.³⁷⁴ Although a host of scholars who have developed these tools for academic purposes make their codes and data publicly available,³⁷⁵ it generally takes significant time and effort to develop new databases on which to run these models and to train and adjust the algorithms.³⁷⁶ This means that states with more

³⁶⁸ See *supra* Section II.B.1.

³⁶⁹ A diplomat from the United Kingdom, for instance, is less likely to need to use commercial translation or interpretation tools, at least in U.N. settings.

³⁷⁰ Cade Metz, *An Infusion of AI Makes Google Translate More Powerful than Ever*, WIRED (Sept. 27, 2016, 1:00 PM), <https://www.wired.com/2016/09/google-claims-ai-break-through-machine-translation/> [<https://perma.cc/54SL-9WK7>] (describing how companies like Google are training deep neural nets using translations gathered online).

³⁷¹ See Bento, *supra* note 306.

³⁷² See *supra* Section III.A.1.

³⁷³ See *supra* Section III.A.1.

³⁷⁴ See *supra* Section III.A.1.

³⁷⁵ See Wiggers, *supra* note 367.

³⁷⁶ See, e.g., Gang Luo et al., *Automating Construction of Machine Learning Models with Clinical Big Data: Proposal Rationales and Methods*, JMIR RES. PROTOCOLS (Aug. 29, 2017), <https://www.researchprotocols.org/2017/8/e175/PDF> [<https://perma.cc/MC2P-NC9G>] (“Building

limited access to data scientists and digital data are less likely to develop and employ these tools.

We might place tools such as advanced “emotion recognition” software in the third category.³⁷⁷ Although companies such as Microsoft sell emotion recognition software commercially,³⁷⁸ it would require a high level of sophistication for a state to deploy this kind of algorithm during negotiations without being detected. Some states that hope to acquire a distinct advantage in dispute resolution or treaty negotiations undoubtedly will develop toolkits that they will refuse to share or make public, because doing so would erase their advantage. Therefore, it seems likely that some limited number of technologically sophisticated states with robust intelligence capacities will deploy secret tools (including ones we cannot even imagine at this point) in these settings, with the goal of gaining greater knowledge and therefore power over states that lack those tools. Although some will be troubled by the appearance of a new, potentially powerful tool to advantage one subset of states, it is not clear that this kind of tool is different in kind from any intelligence advantages that those states already have.

Finally, even assuming that many of these tools were widely available, different states and their international lawyers are likely to have different levels of enthusiasm for their use and different capacities to use them. Some states’ lawyers and diplomats will be more comfortable with advanced technologies, especially if they have spent time in private sector companies that use these tools. Other states’ officials may either be intimidated by the technology or find themselves unpersuaded that it provides any advantage. Yet other officials might desire the technologies but come from states that cannot afford them or that lack ready access to the types and quantities of data required to develop these tools. Also relevant here is the extent to which other states are using such tools; a state that is otherwise lukewarm on the idea of text-as-data may be pushed to explore its potential if many states around it are using those tools during adjudication or negotiations. If the tools help international lawyers do their jobs, those who disdain the tools will find themselves relatively disempowered.

and generalizing a machine learning model often requires hundreds to thousands of manual iterations by data scientists . . .”).

³⁷⁷ See *supra* text accompanying notes 199–208.

³⁷⁸ See *supra* note 200 and accompanying text.

2. *Distributional Effects Within States*

Although this Article focuses on international law and relations, it is worth noting that these tools also will redistribute power *within* states, not just *among* states. If the tools prove effective, they will empower those international lawyers, policymakers, diplomats, and data scientists who embrace the technologies and learn how to use them.

More specifically, those actors within a Foreign Ministry who are technologically literate will be empowered. This might mean that data scientists become relatively more important to policymakers than they have been previously, and that senior officials consider their input to be as important to an international law decision as their international lawyer's is. It will be the data scientists who can suggest new text-as-data tools and interpret the results of existing models. This means that the data scientists who embrace and understand the problems that international lawyers and diplomats face will be most effective in this setting.

Likewise, those international lawyers who embrace the tools and are enthusiastic about working with data scientists to frame the problems will also gain power relative to lawyers who shun the tools' development and forego input into the types of questions to which the tools are directed. Further, because intelligence will play an important role in the use of a variety of negotiation tools discussed above, lawyers who are able to communicate well with their intelligence counterparts will also find themselves more highly valued. Those lawyers who reject the tools entirely may find themselves disempowered relative to their prior standing.

Finally, lawyers who are able to understand the implications of these technologies for existing domestic laws and regulations will also be valued within governments.

C. *Implications for International Law and Diplomacy*

Assuming, as Section B argues, that text-as-data tools will alter interstate power dynamics in different ways, how might these tools affect the creation of international law or the resolution of international legal disputes in the future? One can imagine at least four possibilities. First, these tools could have an overall beneficial effect on international law, allowing states to reach agreement on a wider set of issues more quickly. Second, these tools might facilitate agreement, but do so in a way that systematically favors the interests of the tool users, exacerbating dissatisfaction with the international system among those who are unable to develop or deploy the tools them-

selves. Third, these tools might confound dispute resolution and the conclusion of treaties, perhaps because non-technologically advanced states may reject any process that appears to involve the use of text-as-data tools. Fourth, the tools might produce no notable changes, either because they are not particularly effective or because all states begin to use them and the tools lift all boats equally.

Which story is correct would seem to depend on at least two variables. First, will the use of these tools by some (but not all) states dramatically affect the non-users' perceptions of fairness? Second, might the impact of text-as-data tools depend on the type of international activity the states are undertaking?

Use of these tools by one or a few states, but not all states, may foster a sense of unfairness. In extreme cases, states may shy away from acceding to agreements that they believe were not reached by a fair process.³⁷⁹ Thus, a state's perception of the other side's use of these tools may affect its willingness to reach agreement in the first place, or its willingness to ratify the agreement after the fact. It is also possible that the use of machine learning tools to manipulate a state's media during negotiations might lead that state to allege that the treaty was void because the state was fraudulently induced to conclude it.³⁸⁰

At a certain point, however, a state using tools that are perceived to be unfair will hurt its own ability to advance its foreign policy goals, ultimately undercutting its own power.³⁸¹ This may have happened in the case in which Australia reportedly bugged a Timor L'Este cabinet office during negotiations about a maritime treaty between the two countries and subsequently raided the offices of Timor L'Este's lawyer to seize documents related to the bugging.³⁸² Timor L'Este brought claims both in arbitration and the ICJ, objecting to Australia's purported interference with Timor L'Este's legal communications.³⁸³ One

³⁷⁹ See Odell & Tingley, *supra* note 173, at 145 (“[N]egotiations are more successful . . . to the extent that the process and deal are just.”).

³⁸⁰ See Vienna Convention on the Law of Treaties, *supra* note 130, art. 49, 1155 U.N.T.S. at 344 (“If a State has been induced to conclude a treaty by the fraudulent conduct of another negotiating State, the State may invoke the fraud as invalidating its consent to be bound by the treaty.”).

³⁸¹ Krisch, *supra* note 101, at 371 (discussing how powerful states, faced with international law's insistence on equality and stability, sometimes withdraw from the system).

³⁸² Ashley Deeks, *East Timor's Case in the ICJ: Will the Court Decide Whether Spying Violates International Law?*, *LAWFARE* (Jan. 22, 2014, 10:00 AM), <https://www.lawfareblog.com/east-timors-case-icj-will-court-decide-whether-spying-violates-international-law> [https://perma.cc/SL3R-KFKE].

³⁸³ *Id.*

can view this case as a preview of what could happen if one state learns that another has engaged in clandestine high-tech analysis of treaty negotiations in a way that produces an unjust outcome.

As another example, assume one state uses text-as-data tools to scour archives to detect historically relevant state practice. If it ultimately comes to light that the state used the tools to identify previously unknown state practice that favored its position but suppressed the state practice that hurt its argument, other states might object. On the one hand, finding precedent that favors one's position is what legal research is about. On the other hand, in the face of a deep digital divide among states that "affects the availability and accessibility of the evidence of customary international law,"³⁸⁴ having an advanced state use these tools to deepen that divide could cause tensions. Selective manipulations of these tools by a state may ultimately undercut whatever advantages that state seeks to gain by prompting a backlash against that state's (possibly legitimate) arguments. A similar point holds for the use of AI to improve intelligence collection. This seems likely to widen the gap between those states that already have robust intelligence services and those that do not.

On the other hand, a state that uses AI to identify international law violations by other states might help advance international peace and security, which benefits all states.³⁸⁵ A state that uses text-as-data to spotlight practices by states whose voices are less often heard in CIL debates could foster a sense of fairness and empowerment. Some of this is within the control of weaker states, assuming they can find a way to digitize and publicize their own practice.

A second variable is the setting in which states use these tools. If two states are in a zero-sum negotiation, they are unlikely to share these tools with each other.³⁸⁶ Indeed, the fact that the tools exist might actually make it harder to reach agreement if the states attempt to "out-predict" each other or create unrealistic expectations for themselves about what they "should" be able to obtain in negotiations. In contrast, if two states are negotiating a "win-win" treaty, the states are more likely to reveal their uses of text-as-data tools and possibly share those tools (or at least the tools' outputs) in an effort to

³⁸⁴ U.N. Secretariat Report, *supra* note 247, ¶ 90.

³⁸⁵ See Christopher D. Baker, *Tolerance of International Espionage: A Functional Approach*, 19 AM. U. INT'L L. REV. 1091, 1092 (2004) (discussing espionage as a way to verify other states' compliance with their international law obligations).

³⁸⁶ See Peters, *supra* note 204, at 30 ("Zero-sum negotiations are competitive by definition.").

reach consensus. A neutral state helping to resolve a long-standing foreign policy dispute by locating new, previously unconsidered proposals could advance the goals of all involved states.³⁸⁷ Avoiding inter-treaty conflicts could smooth the overall operation of international law by highlighting conflicts before they arise, thus avoiding a scenario in which one state brings itself out of compliance with an international law obligation by concluding a subsequent treaty.³⁸⁸

In sum, the likeliest outcome is that these tools will serve as double-edged swords. In the same way that individuals use social media and companies deploy AI, text-as-data tools could be used in ways that either benefit or hinder international law development. Preparing for treaty negotiations by better understanding the negotiating partner's equities could increase the likelihood that states reach agreement, but it also could mean that they reach agreement on terms more favorable to the state using the tools.³⁸⁹ Using these tools to anticipate a negotiating partner's local political reactions could render a state more sensitive to the partner's constraints, but it also could enable the state to manipulate local reactions to its advantage.³⁹⁰ What seems most important for international law's prospects is *how* states use these tools, not *whether* they use them.

D. *Facilitating Redistribution?*

Assuming that powerful states will develop and use these tools regardless of the overall effect of their use on international law, there may be ways for less powerful states to minimize potential adverse effects on their interests. No existing international legal rules impose specific constraints on how states may use these tools. Therefore, one option is for powerful states to undertake to use (or not use) these tools in particular ways as a matter of policy. Those states might do so for instrumental reasons—if they worried, for instance, that their use of some of these tools would hinder their ability to conclude treaties that were important to them. Alternatively, less powerful states could explore ways to empower themselves to keep up with more powerful states.

The Chatham House report (which focuses on human development, economic changes from AI, and autonomous weaponry) argues that governments worldwide should invest in developing AI expertise

³⁸⁷ See *supra* Section III.B.1.

³⁸⁸ See *supra* Section II.B.3.

³⁸⁹ See *supra* Section II.A.1.

³⁹⁰ See *supra* Sections III.B.2–3.

and talent to avoid being dominated by the expertise that currently exists in China and the United States.³⁹¹ This same advice might attach to the use of text-as-data tools. Less powerful states—particularly those that either enter into many international agreements or that are involved in persistent foreign relations disputes—could decide to devote resources to hiring data scientists to work with their international lawyers on these tools.

Alternatively, corporations and non-governmental organizations might donate expertise to less developed countries in this area, specifically by helping them produce algorithms or models that could assist with the types of tasks discussed in Parts II and III.³⁹² Some technologists are trying to make domestic legal algorithms more widely available and less costly, so that those who cannot hire lawyers can employ those algorithms.³⁹³ Kevin Cope has proposed something similar with regard to tools that help states make predictions about treaty adherence. He recognizes that his idea requires “a team of trained lawyers, social scientists, coders, and data analysts, to which most delegations would not have access.”³⁹⁴ As a result, he suggests the creation of a third party consulting organization that could provide such services to delegations that request them.³⁹⁵ Finally, there might be a role for the U.N. Secretariat to offer free or low-cost algorithms to states and teach states how to deploy them. The United Nations itself might choose to aggregate, standardize, and make available all of its data as a training set for those developing international law-related algorithms.³⁹⁶ These are just some of the avenues that less powerful states

³⁹¹ Bryce & Parakilas, *supra* note 10, at 45.

³⁹² There are a range of “civic tech” actors in the United States, many of whom focus on state and local government issues. *See, e.g.*, CODE FOR AM., <https://www.codeforamerica.org> [<https://perma.cc/D4EZ-QWY6>]; OPEN CITIES AFR., <https://opencitiesproject.org/about/> [<https://perma.cc/H3HG-2LSQ>]; OPEN DATA LABS, <https://labs.webfoundation.org/open-is-the-new-smart-making-cities-work-for-citizens/> [<https://perma.cc/B77H-GMW6>]. This same impetus might drive some of these actors to offer low-cost services to less-developed states.

³⁹³ *See, e.g.*, *What We Do*, LEGAL ROBOT, <https://legalrobot.com/company> [<https://perma.cc/E4LK-2ZAK>]; ROSS INTELLIGENCE, <https://rossintelligence.com/why-ross.html> [<https://perma.cc/769D-6HDZ>].

³⁹⁴ Cope, *supra* note 196, at 55. Not every state is likely to respond positively to offers of assistance. India, for example, rejected Facebook’s offer to provide free Internet access in rural India because the tool would not have made all websites equally accessible. Daniel Van Boom, *Why India Snubbed Facebook’s Free Internet Offer*, CNET (Feb. 26, 2016, 6:00 AM PST) <https://www.cnet.com/news/why-india-doesnt-want-free-basics/> [<https://perma.cc/4WHR-LUWF>].

³⁹⁵ Cope, *supra* note 196, at 56.

³⁹⁶ By way of a model, databases such as ImageNet provide academic researchers with easily accessible, tagged images on which they can train their own algorithms. *About ImageNet*, IMAGE NET, <http://image-net.org/about-overview> [<https://perma.cc/E47S-ADWN>]. In the legal context, Washington University Law School maintains the Supreme Court Database, which con-

could pursue to guard against significant power re-distributions in the face of text-as-data tools.

CONCLUSION

This Article argued that international lawyers may and should soon begin to use text-as-data tools, and must understand their strengths, weaknesses, and possible applications. It considered a variety of projects that data scientists and international lawyers might undertake to improve states' and scholars' creation and understanding of international law and the resolution of international disputes. Even if current technologies are only at the cusp of delivering the types of promising applications discussed herein, states should be more attuned to the burgeoning role of technology in helping their international lawyers and diplomats do their jobs. This Article neither advocates nor predicts that text-as-data tools will or should replace direct person-to-person contacts or short-circuit the often time intensive and painstaking process of reaching international agreements and resolving disputes. But these tools appear poised to facilitate these activities, and some states are likely to take advantage of them. Whether that will help level the playing field among more and less powerful states or exacerbate those power differentials remains to be seen and will depend in part on the extent to which actors such as the United Nations and not-for-profit tech organizations choose to develop and offer high-tech tools to all states.

In either case, this Article only scratches the surface of where technology may take international law. Artificial intelligence—and machine learning in particular—are likely to have broad impact on international law, and on state power generally. States' use of AI will affect when and how they resort to force against each other,³⁹⁷ and it will complicate the actual fighting of armed conflicts in light of existing laws of war.³⁹⁸ AI might facilitate states' abilities to detect international law violations (as where machine learning tools allow certain states to identify every instance in which states move their troops and

tains every Supreme Court decision, coded for 60 variables. See *The Supreme Court Database*, WASH. U. L. SCH., <http://scdb.wustl.edu> [<https://perma.cc/G2GU-GR2P>]. Sergio Puig and Enric Torrents have argued for a “federated, hybrid database” located in the Cloud that would contain in unified format data from the world economic courts. Sergio Puig & Enric G. Torrents, *The Case for Linking World Law Data*, 2 J. OPEN ACCESS L. 1, 1 (2014). The United Nations could commit to something similar for all official documents that it produces.

³⁹⁷ Deeks et al., *supra* note 329, at 2.

³⁹⁸ See Rebecca Crootof, *A Meaningful Floor for “Meaningful Human Control,”* 30 TEMP. INT'L & COMP. L.J. 53, 62 (2016).

military hardware) and provide a deterrent effect that improves other states' international law compliance.³⁹⁹ Machine learning might help states process extradition requests and identify evidentiary shortfalls in those requests based on past case outcomes. A wide range of other substantive applications surely wait around the corner.

In short, machine learning and other high-tech tools have the potential to embed themselves widely into the conduct of foreign relations, and we should be attuned not only to those areas in which they might stimulate conflict, but also to areas in which the tools can strengthen the development and implementation of international law. Data scientists, domestic lawyers, and scholars have undertaken a range of projects that demonstrate the types of legal problems that machine learning and computational text analysis can address. Many of those same categories of problems exist in international law as well. The time is ripe for all states to help international law go high-tech.

³⁹⁹ See, e.g., Boutin, *supra* note 27 (discussing ways in which AI might improve international law compliance).