

Artificial Intelligence and Legal Disruption: A New Model for Analysis

Hin-Yan Liu,¹ Matthijs Maas,² John Danaher,³

Luisa Scarcella,⁴ Michaela Lexer,⁵ Leonard Van Rompaey⁶

Abstract

Artificial intelligence (AI) is increasingly expected to disrupt the ordinary functioning of society. From how we fight wars or govern society, to how we work and play, and from how we create to how we teach and learn, there is almost no field of human activity which is believed to be entirely immune from the impact of this emerging technology. This poses a multifaceted problem when it comes to designing and understanding regulatory responses to AI. This article aims to: (i) defend the need for a novel conceptual model for understanding the systemic *legal disruption* caused by new technologies such as AI; (ii) to situate this model in relation to preceding debates about the interaction of regulation with new technologies (particularly the ‘cyberlaw’ and ‘robolaw’ debates); and (iii) to set out a detailed model for understanding the legal disruption precipitated by AI, examining both pathways stemming from new affordances that can give rise to a regulatory ‘disruptive moment’, as well as the Legal Development, Displacement or Destruction that can ensue. The article proposes that this model of legal disruption can be broadly generalisable to understanding the legal effects and challenges of other emerging technologies. Thus, while our model of legal disruption is crafted in response to the specific regulatory challenges raised by AI, we believe that, with minor modifications, this model can be usefully deployed to understand the challenges raised by future emerging technologies, and to structure regulatory responses to those challenges.

Introduction

As a widely applicable ‘general purpose technology’⁷ artificial intelligence (AI) has been recognized as having the potential to disrupt many social behaviours, practices and institutions.

¹ Associate Professor and Coordinator, Artificial Intelligence and Legal Disruption Research Group, Faculty of Law, University of Copenhagen. Email: hin-yan.liu@jur.ku.dk

² PhD Fellow, Artificial Intelligence and Legal Disruption Research Group, Faculty of Law, University of Copenhagen. Email: Matthijs.Maas@jur.ku.dk | ORCID ID: 0000-0002-6170-9393.

³ Senior Lecturer, National University of Ireland, Galway. Email: john.danaher@nuigalway.ie.

⁴ PhD Fellow, Department of Tax and Fiscal Law, University of Graz. Email: luisa.scarcella@uni-graz.at

⁵ PhD Fellow, Department of Labour and Social Security Law, University of Graz, Email: michaela.lexer@edu.uni-graz.at

⁶ Industrial Postdoctoral Researcher, Synch Advokat Copenhagen, and the Centre for Private Governance, Faculty of Law, University of Copenhagen. Email: leonard.van.rompaey@jur.ku.dk

⁷ Manuel Trajtenberg, ‘AI as the next GPT: A Political-Economy Perspective’ (National Bureau of Economic Research 2018) Working Paper 24245 <<http://www.nber.org/papers/w24245>> accessed 22 October 2018; Ajay Agrawal, Joshua

Functionally speaking, AI is a portfolio of techniques that enable improvements in the accuracy, speed or scale of machine performance across complex or large (data) environments, with the aim of substituting for, or improving upon, human performance in diverse specific tasks such as decision-making, pattern-recognition and prediction. Accordingly, AI technology raises fundamental questions of power and control across society,⁸ and has been anticipated to challenge almost every sphere of human activity, as society transitions into a ‘digital lifeworld’.⁹ Uses—and concerns—range across diverse sectors, from legal decision-making and policing to health-care,¹⁰ transport,¹¹ and military uses,¹² to name but a few.¹³

Looking beyond the granular problems raised in each domain or sector, what does all this mean for the legal system? New technologies have a long history of challenging existing law.¹⁴ It should,

Gans and Avi Goldfarb (eds), *The Economics of Artificial Intelligence* (2019) pt 1

<<https://www.press.uchicago.edu/ucp/books/book/chicago/E/bo35780726.html>> accessed 12 February 2019 ‘AI as a GPT’.

⁸ Hin-Yan Liu, ‘The Power Structure of Artificial Intelligence’ (2018) 10 *Law, Innovation and Technology* 197.

⁹ See generally Jamie Susskind, *Future Politics: Living Together in a World Transformed by Tech* (Oxford University Press 2018).

¹⁰ The literature is vast; however, on the use of AI in the legal context, see also: Harry Surden, ‘The Ethics of Artificial Intelligence in Law: Basic Questions’ in M Dubber and F Pasquale (eds), *The Oxford Handbook of AI Ethics* (Oxford University Press 2019); Gabriele Buchholtz, ‘Artificial Intelligence and Legal Tech: Challenges to the Rule of Law’ in Thomas Wischmeyer and Timo Rademacher (eds), *Regulating Artificial Intelligence* (Springer International Publishing 2020) <https://doi.org/10.1007/978-3-030-32361-5_8> accessed 6 December 2019. On the use in policing, see for example Annette Vestby and Jonas Vestby, ‘Machine Learning and the Police: Asking the Right Questions’ [2019] *Policing: A Journal of Policy and Practice* <<https://academic.oup.com/policing/advance-article/doi/10.1093/police/paz035/5518992>> accessed 25 June 2019. On the use in healthcare, see: W Nicholson Price, ‘Artificial Intelligence in Health Care: Applications and Legal Issues’ (Social Science Research Network 2017) SSRN Scholarly Paper ID 3078704 <<https://papers.ssrn.com/abstract=3078704>> accessed 23 June 2018; Sarah Jabri, ‘Artificial Intelligence and Healthcare: Products and Procedures’ in Thomas Wischmeyer and Timo Rademacher (eds), *Regulating Artificial Intelligence* (Springer International Publishing 2020) <https://doi.org/10.1007/978-3-030-32361-5_14> accessed 6 December 2019.

¹¹ Jessica S Brodsky, ‘Autonomous Vehicle Regulation: How an Uncertain Legal Landscape May Hit the Brakes on Self-Driving Cars Cyberlaw and Venture Law’ (2016) 31 *Berkeley Technology Law Journal* 851; James M Anderson and others, ‘Autonomous Vehicle Technology: A Guide for Policymakers’ (RAND Corporation 2016) Product Page <https://www.rand.org/pubs/research_reports/RR443-2.html> accessed 17 October 2017.

¹² Cf. Kenneth Anderson and Matthew C Waxman, ‘Debating Autonomous Weapon Systems, Their Ethics, and Their Regulation under International Law’, *The Oxford Handbook of Law, Regulation and Technology* (2017) <<http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199680832.001.0001/oxfordhb-9780199680832-e-33>> accessed 3 January 2019; Dustin A Lewis, Gabriella Blum and Naz K Modirzadeh, ‘War-Algorithm Accountability’ (Harvard Law School Program on International Law and Armed Conflict 2016) Research Briefing <<https://blogs.harvard.edu/pilac/files/2016/09/War-Algorithm-Accountability-Without-Appendices-August-2016.pdf>> accessed 16 September 2016.

¹³ For other general overviews of the issue space, see also Patrick Lin, Keith Abney and George A Bekey, *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT Press 2011); Patrick Lin, Keith Abney and Ryan Jenkins (eds), *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (Oxford University Press 2017); Ryan Calo, ‘Artificial Intelligence Policy: A Primer and Roadmap’ (2017) 51 *UC Davis Law Review* 37; For a survey that focuses on international governance issues, see Allan Dafoe, ‘AI Governance: A Research Agenda’ (Center for the Governance of AI, Future of Humanity Institute 2018) <<https://www.fhi.ox.ac.uk/govaiagenda/>>; For a recent legal angle, see Wolfgang Hoffmann-Riem, ‘Artificial Intelligence as a Challenge for Law and Regulation’ in Thomas Wischmeyer and Timo Rademacher (eds), *Regulating Artificial Intelligence* (Springer International Publishing 2020) <https://doi.org/10.1007/978-3-030-32361-5_1> accessed 6 December 2019.

¹⁴ See for instance Lyria Bennett Moses, ‘Recurring Dilemmas: The Law’s Race to Keep Up With Technological Change’ (2007) 21 *University of New South Wales Faculty of Law Research Series* <<http://www.austlii.edu.au/au/journals/UNSWLRS/2007/21.html>> accessed 3 July 2018; Colin B Picker, ‘A View from

consequently, come as no surprise that AI, given its broad usage, holds the potential to trigger large, and in some cases seismic, shifts across the legal and regulatory landscape.

Given this, we believe that there is a need to present a new, integrated, contextual, and holistic model for understanding the disruptive effects of AI systems on law as such. Such a holistic model can be both illuminating for understanding the challenges posed by AI in specific domains, but it also offers more broadly transferable insights which can also be adapted and extended to understand other forms of technological legal disruption. Indeed, we think there is value to focusing specific attention on the phenomenon of ‘legal disruption’ per se, as opposed to technological disruption that happens to have legal consequences.¹⁵

In the remainder of this article, we justify these claims, and provide a model for addressing legal disruption per se. We do this in four distinct phases. First, we justify the need for a general model of AI-based legal disruption. Second, we situate the need for that model within broader debates about technology and legal disruption, focusing in particular on the lessons we can learn from past debates about ‘cyberlaw’ and ‘roboLaw’. Third, we outline and describe our proposed model in detail. And fourth, we discuss the advantages and limitations of this model. We conclude by reflecting on the future development of the model.

1. The Case for a General Model of ‘Legal Disruption’

We start by answering the ‘why’ question: Why do we need to develop a generalised model of legal disruption? One obvious reason is that having a model helps us to chart, manage and mitigate the disruptive impacts of AI. As outlined in the introduction, AI is expected to have widespread effects on societal practices and values. We may be paralysed in our capacity to address these impacts if we lack a model that allows us to make reasoned judgments about the probability and likely impact of such disruptive possibilities.

One way in which this paralysis could manifest itself is in a regulatory game of ‘whack-a-mole’, whereby both scholarship and regulatory responses are tailored reactively to each challenge raised by AI. This would amount to treating the symptoms of legal disruption instead of addressing the cause. It would overlook the opportunity to develop an overarching regulatory regime capable of guiding law and policy in the face of AI challenges more generally.¹⁶

Consider a concrete example: In recent years, one of the more high-profile substantive debates concerning the legal impact of AI has focused on so-called Autonomous Weapons Systems

40,000 Feet: International Law and the Invisible Hand of Technology’ (2001) 23 *Cardozo Law Review* 151; David D Friedman, ‘Does Technology Require New Law?’ (2001) 71 *Public Policy* 16.

¹⁵ See also Margot E Kaminski, ‘Authorship, Disrupted: AI Authors in Copyright and First Amendment Law’ (2017) 51 *UC Davis Law Review* 589.

¹⁶ Though see Matthew U Scherer, ‘Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies’ [2016] *Harvard Journal of Law & Technology* <<http://jolt.law.harvard.edu/articles/pdf/v29/29HarvJLTech353.pdf>> accessed 5 March 2018; Michael Guihot, Anne F Matthew and Nicolas Suzor, ‘Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence’ [2017] *Vanderbilt Journal of Entertainment & Technology Law* <<https://papers.ssrn.com/abstract=3017004>> accessed 2 July 2018.

(AWS).¹⁷ But many of the basic legal issues surrounding AWS, such as establishing liability for harm, reckoning with and preventing unanticipated accident cascades,¹⁸ considering ‘structural’ effects of the technology’s deployment,¹⁹ and reckoning with the blurring of actor and object, are considerations that are not narrowly applicable to international humanitarian law (IHL), but apply to many manifestations of AI-based technologies. This has already become apparent with the emergence of the debate about the legal repercussions of autonomous vehicles.²⁰ Roger Brownsword has argued that diverse uses of AI, such as autonomous road traffic, autonomous weapons, and profiling, prediction and prevention, rather than being distinct and isolated issue areas, can be productively examined in common, through the lens of how these applications violate or challenge human dignity.²¹

In the siloed view, however, regulatory responses are compartmentalised by, and therefore focus attention on certain activity areas, while overlooking or rendering peripheral other areas of the potential problem (and solution) space.

Consider another example: the case of so-called ‘DeepFakes’. Since emerging in 2017, the Generative Adversarial Networks (GAN)-based technique for generating fake but nearly indistinguishable-from-real content has sparked the imagination (and panic) of the public. Spurred

¹⁷ Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (1 edition, Routledge 2009); PW Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (Penguin Press 2009). Whether the prominence of this issue is because military expenditure in research fuelled rapid technological development and public ‘arms races’, because lethality by algorithm and machine caught civil society attention for being particularly egregious, or because the legal and regulatory issues surrounding military robotics have proven especially visceral and intuitive to grasp, does not matter for our purposes. However, for some examination into the recent ‘arms races’ on AWS, see Justin Haner and Denise Garcia, ‘The Artificial Intelligence Arms Race: Trends and World Leaders in Autonomous Weapons Development’ (2019) 10 *Global Policy* 331; Though for a critical angle on the ‘arms race’ angle, see Heather M Roff, ‘The Frame Problem: The AI “Arms Race” Isn’t One’ (2019) 0 *Bulletin of the Atomic Scientists* 1. On the recent history and developments of civil society movements to ban these technologies, see: Charli Carpenter, *Lost Causes, Agenda Vetting in Global Issue Networks and the Shaping of Human Security* (Cornell University Press 2014) <<http://www.degruyter.com/viewbooktoc/product/487489>> accessed 1 April 2019; Şerif Onur Bahçecik, ‘Civil Society Responds to the AWS: Growing Activist Networks and Shifting Frames’ (2019) 0 *Global Policy* <<https://onlinelibrary.wiley.com/doi/abs/10.1111/1758-5899.12671>> accessed 17 June 2019.

¹⁸ For research into this underexplored angle, see Paul Scharre, ‘Autonomous Weapons and Operational Risk’ (Center for a New American Security 2016) <https://s3.amazonaws.com/files.cnas.org/documents/CNAS_Autonomous-weapons-operational-risk.pdf>; Matthijs M Maas, ‘Regulating for “Normal AI Accidents”—Operational Lessons for the Responsible Governance of AI Deployment’, *Proceedings of the 2018 AAAI / ACM Conference on Artificial Intelligence, Ethics and Society* (Association for the Advancement of Artificial Intelligence 2018) <http://www.aies-conference.com/wp-content/papers/main/AIES_2018_paper_118.pdf>; Stephanie Carvin, ‘Normal Autonomous Accidents: What Happens When Killer Robots Fail?’ (Social Science Research Network 2017) SSRN Scholarly Paper ID 3161446 <<https://papers.ssrn.com/abstract=3161446>> accessed 17 December 2019.

¹⁹ cf. Hin-Yan Liu, ‘Three Types of Structural Discrimination Introduced by Autonomous Vehicles’ (2018) 51 *UC Davis Law Review* 32. See also the discussion of the structural effects of AI technology, and the similarities of such risks across different usage domains--whether military or in self-driving cars--in Remco Zwetsloot and Allan Dafoe, ‘Thinking About Risks From AI: Accidents, Misuse and Structure’ (*Lawfare*, 11 February 2019) <<https://www.lawfareblog.com/thinking-about-risks-ai-accidents-misuse-and-structure>> accessed 12 February 2019.

²⁰ Hin-Yan Liu, ‘Irresponsibilities, Inequalities and Injustice for Autonomous Vehicles’ (2017) 19 *Ethics and Information Technology* 193.

²¹ Roger Brownsword, ‘From Erewhon to AlphaGo: For the Sake of Human Dignity, Should We Destroy the Machines?’ (2017) 9 *Law, Innovation and Technology* 117.

on by high-profile use cases or demonstrations involving fake videos of prominent politicians,²² however, much public debate, academic work and even the incipient regulatory action has framed DeepFakes narrowly in the context of two specific problem domains: a *personal* threat in the form of forged pornography, or a *political* threat to democracy.²³ In consequence, the first legislative efforts to control this technology – a Virginian²⁴ ban on the use of DeepFakes for nonconsensual pornography in summer 2019, and Californian²⁵ and Chinese²⁶ bans in the autumn – focused on these specific challenges.

While certainly an important step, however, such laws have been criticized as being narrowly focused and full of loopholes and shortcomings.²⁷ These shortcomings may derive from problems that are general rather than specific, and have more to do with the underlying reactive domain-focused approach which leads to a range of analytical and practical regulatory shortcomings. First, it may overstate the novelty or ‘added value’ of DeepFake technology, relative to established techniques and approaches (so-called ‘cheap fakes’) for information warfare.²⁸ Second, the emphasis on the use of DeepFakes for the production of fake *visual* media, may underestimate the degree to which political information warfare might find greater use for less-visceral cases, such as the generation of fake text.²⁹ Thirdly, the emphasis on DeepFakes as a particular ‘problem’ requiring specific solutions – such as the Californian and Chinese bans – does not focus sufficient attention on underlying or foundational problems in the way in which digital technology

²² See for instance BuzzFeedVideo, *You Won't Believe What Obama Says In This Video!* (2018) <<https://www.youtube.com/watch?v=cQ54GDm1eL0>> accessed 18 February 2019; BBC, ‘Are You Fooled by This Johnson-Corbyn Video?’ (*BBC News*, 12 November 2019) <<https://www.bbc.com/news/av/technology-50381728/the-fake-video-where-johnson-and-corbyn-endorse-each-other>> accessed 9 December 2019; Bob Schaffer, ‘President Nixon Never Actually Gave This Apollo 11 Disaster Speech. MIT Brought It To Life To Illustrate Power Of Deepfakes’ (*wbur*, 22 November 2019) <<https://www.wbur.org/news/2019/11/22/mit-nixon-deep-fake>> accessed 9 December 2019. For one case in which a political DeepFake video went viral, see Hans von der Burchard, ‘Belgian Socialist Party Circulates “Deep Fake” Donald Trump Video’ (*Politico*, 21 May 2018) <<https://www.politico.eu/article/spa-donald-trump-belgium-paris-climate-agreement-belgian-socialist-party-circulates-deep-fake-trump-video/>> accessed 18 February 2019.

²³ Robert Chesney and Danielle Keats Citron, ‘Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security’ (2019) 107 *California Law Review* <<https://papers.ssrn.com/abstract=3213954>> accessed 17 November 2018.

²⁴ Adi Robertson, ‘Virginia’s “Revenge Porn” Laws Now Officially Cover Deepfakes’ (*The Verge*, 1 July 2019) <<https://www.theverge.com/2019/7/1/20677800/virginia-revenge-porn-deepfakes-nonconsensual-photos-videos-ban-goes-into-effect>> accessed 17 December 2019.

²⁵ ‘AB-730 Elections: Deceptive Audio or Visual Media.’ (4 October 2019) <https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730> accessed 9 December 2019.

²⁶ Yingzhi Yang and Brenda Goh, ‘China Seeks to Root out Fake News and Deepfakes with New Online Content Rules’ *Reuters* (29 November 2019) <<https://www.reuters.com/article/us-china-technology-idUSKBN1Y30VU>> accessed 9 December 2019.

²⁷ Brandie M Nonnecke, ‘Anti-Deepfake Law in California Is Far Too Feeble’ [2019] *Wired* <<https://www.wired.com/story/opinion-californias-anti-deepfake-law-is-far-too-feeble/>> accessed 17 December 2019.

²⁸ Britt Paris and Joan Donovan, ‘DeepFakes and Cheap Fakes: The Manipulation of Audio & Visual Evidence’ (*Data & Society* 2019) <<https://datasociety.net/output/deepfakes-and-cheap-fakes/>>.

²⁹ See also the debate around the social impacts and risks of the OpenAI GPT-2 natural language generation model. Irene Solaiman and others, ‘Release Strategies and the Social Impacts of Language Models’ [2019] arXiv:1908.09203 [cs] <<http://arxiv.org/abs/1908.09203>> accessed 18 November 2019. Although see Aviv Ovadya and Jess Whittlestone, ‘Reducing Malicious Use of Synthetic Media Research: Considerations and Potential Release Practices for Machine Learning’ [2019] arXiv:1907.11274 [cs] <<http://arxiv.org/abs/1907.11274>> accessed 9 September 2019.

increasingly enables granular ‘perception-control’.³⁰ In this way, the narrow focus on (visual) ‘DeepFakes’ primes the debate to seek incremental and dissociated responses to an array of symptomatic issues. In many cases, responses resort too early or easily to narrow, technical solutions such as automatic ways of detecting DeepFake videos³¹ – temporary technical solutions such as automated detection which have at any rate proven easily defeated or bypassed,³² and which may indeed have an inherently limited ‘half-life’ of viability in the face of asymptotically improving fidelity.³³ Such narrow framings which emphasize near-term (technological) solutions may inadvertently but functionally stack the deck against regulation, locking policy responses into arms races which they are not well placed to win.

Similar dynamics play out not just in examples such as DeepFakes, autonomous vehicles, or autonomous weapons systems, but across many other domain areas where AI (and digital technology more broadly) has been intersecting with established legal domains, disciplines, or processes. Exploring the nexus between AI, law, and regulation from pre-established points of departure or the most visceral discrete problem domains is a rather haphazard way to proceed. The problem is five-fold.

First, by following a piecemeal regulatory approach, attention and resources are not necessarily allocated in accordance with the nature or severity of the challenge that an emerging technology poses to the legal and regulatory systems, but are instead prioritized on the basis of a technology’s public profile, hype, and factors that can be easily driven by in-the-lab proofs of concept or by visceral incidents of misuse. As the example of the current use of DeepFakes suggests, these are poor guides to the technology’s actual societal impacts.³⁴ That law and regulation follow where technological development leads, reduces regulatory initiatives very much to mere *responses*, constrained to the paths that are laid down. This responsiveness leaves law and regulation disorganised, as its gaze is focused on the technological capability or artefact, resulting in an incoherent and scattershot approach.³⁵ Furthermore, technological development and deployment are left to proceed in a relatively unguided manner, that relegates regulatory or democratic concerns to an afterthought. If the design or architecture of a system influences or constrains the range of possible behaviours of regulatees,³⁶ this responsive orientation limits the potential efficacy of

³⁰ Susskind (n 9) 142–152.

³¹ To give but one example, see Irene Amerini and others, ‘Deepfake Video Detection through Optical Flow Based CNN’ 3.

³² See for instance Davide Cozzolino and others, ‘SpoC: Spoofing Camera Fingerprints’ [2019] arXiv:1911.12069 [cs, eess] <<http://arxiv.org/abs/1911.12069>> accessed 9 December 2019.

³³ See Alex Engler, ‘Fighting Deepfakes When Detection Fails’ (*Brookings*, 14 November 2019) <<https://www.brookings.edu/research/fighting-deepfakes-when-detection-fails/>> accessed 11 December 2019. For example, recent breakthroughs have further addressed previous flaws and shortcomings in GAN image quality; see Tero Karras and others, ‘Analyzing and Improving the Image Quality of StyleGAN’ [2019] arXiv:1912.04958 [cs, eess, stat] <<http://arxiv.org/abs/1912.04958>> accessed 17 December 2019.

³⁴ Perhaps echoing the legal maxim that ‘hard cases make bad laws’.

³⁵ This suggests such an approach does not work even from a legal coherentist perspective. For discussion of which, see the section on Stage 2 (a) Legal Development (v).

³⁶ Cf. Donella H Meadows, *Thinking in Systems: A Primer* (Diana Wright ed, Chelsea Green Publishing 2008).

regulation, because its range of possible influence is curtailed as a result of the specific policy focus.³⁷

Second, this domain-specific focus treats the symptoms of the problem, rather than the cause. It fragments attention that could otherwise be directed at addressing the core source of the problem. Because they constitute a set of ‘general purpose technologies’ that are readily incorporated into almost any existing activity,³⁸ the breadth of regulatory issues raised by AI span the spectrum of human activities. Likewise, addressing the legal challenges raised by autonomous weapons systems, or autonomous vehicles, to use the examples above, contextualises those challenges within the confines of IHL and the highway code respectively. In other words, the questions raised by robotics and AI in these contexts are not necessarily, or even predominantly, questions triggered by the emerging technologies. Rather, such debates, implicitly or explicitly, seek to maintain ‘legal coherence’,³⁹ and to ensure the non-turbulent continuation of the pre-existing regulatory equilibria in these areas, thereby imposing a unidirectional flow of lessons towards addressing the precise practical problem but not considering the broader nature of the challenges posed. Thus, focusing attention on the introduction of robotics and AI into armed conflict and transportation and political misinformation may teach us lessons in these activity areas, but they rarely shed significant light upon the structural and systemic challenges raised from this tight cluster of emerging technologies.

Third, the lack of extant framework or typology to organise the legal and regulatory approaches to robotics and AI, means that law can, at best, seek to discern patterns in regulatory responses retrospectively. The lack of a framework makes it difficult to distinguish clusters of issues in advance, or to identify the common characteristics that underlie them. For instance, a cross-cutting analysis of DeepFakes within such a broader framework, would anticipate more commonalities with extant work on, for example, computational propaganda generally,⁴⁰ treating these issues together from the start. A different way of putting this is that it is unclear where the source of legal controversy lies. Does the regulatory challenge emanate from the inherent characteristics or capabilities of the technology itself? Does the legal confusion arise from the deployment of that technology into society? Might the shortcomings of the law be due to different equilibria being struck as technological and societal change balance out? Or might the problems lie elsewhere entirely? This jumble of issues makes it difficult to establish a hierarchy that separates challenges according to the breadth of the interface, the severity of the impact, and the duration of the resulting legal and societal turbulence. Gauging whether a particular legal problem poses superficial or structural questions thus becomes obscured or overlooked. But such a determination is critical, because it determines whether legal accommodation will be capable of addressing the problems posed (if these are relatively superficial, or at least not paradigm changing), or conversely if it will

³⁷ Lawrence Lessig, ‘The New Chicago School’ (1998) 27 *The Journal of Legal Studies* 661; Lawrence Lessig, ‘The Law of the Horse: What Cyberlaw Might Teach’ (1999) 113 *Harvard Law Review* 501.

³⁸ Trajtenberg (n 7); Agrawal, Gans and Goldfarb (n 7).

³⁹ On the differences between ‘coherent’, ‘regulatory-instrumentalist’, and ‘technocratic’ regulatory attitudes, see also Roger Brownsword, ‘Law Disrupted, Law Re-Imagined, Law Re-Invented’ [2019] *Technology and Regulation* 10; And generally Roger Brownsword, *Law, Technology and Society: Re-Imagining the Regulatory Environment* (1 edition, Routledge 2019).

⁴⁰ See for instance Gillian Bolsover and Philip Howard, ‘Computational Propaganda and Political Big Data: Moving Toward a More Critical Research Agenda’ (2017) 5 *Big Data* 273; Sam Woolley and Phil Howard, ‘Computational Propaganda Worldwide’ (Oxford Internet Institute 2017) <<http://comprop.oi.ox.ac.uk/2017/06/19/computational-propaganda-worldwide-executive-summary/>> accessed 25 June 2017.

contribute towards entrenching the problem (if these are structural and systemic) because the law by itself would be incapable of addressing the entire problem.

Fourth, this domain-specific focus has a distorting effect upon the debate, both by prioritising regulatory activity along sectoral lines, and by emphasising immediate and near-term practical problems. An example of this is the liability questions arising from AI possessing discretionary autonomy which breaks the chain of causation traditionally relied upon to establish fault and compensation.⁴¹ Such questions loom large over the legal landscape where AI and the law are deemed to collide. Indirect or “second-order” regulatory challenges are thus externalised or resigned to simple ‘ethical considerations’, even though these challenges hold the potential to fundamentally alter the regulatory landscape. Again, the threat of DeepFakes (and computational propaganda generally) might lie less in the possibility that these techniques could increase in quality to such a level that they could get any target public to trust certain political messages (in which case responses aimed at detecting imperfections would help). The deeper threat may be less about manufacturing misplaced trust, but rather about fracturing public knowledge by getting target publics to *distrust* all digital media, as well as affording certain actors the ‘liars dividend’ (the ability to deflect accusations of recorded improper conduct by claiming the material has been faked).⁴² More broadly, the argument has been made that democracies are asymmetrically and intrinsically more vulnerable to certain misinformation attacks than autocracies.⁴³ If this is true, the focus of regulatory responses should lie at this global level – at the broad range of vectors (whether high- or low-tech) that contribute to ‘truth decay’⁴⁴ – rather than on merely patching local vulnerabilities to specific techniques of information warfare. It should look less at the potential impacts of DeepFakes on specific target groups (politicians for example), but rather on tacit fundamentals such as the ‘epistemic backstop’ of society.⁴⁵ Or, to return to an earlier example, traditional liability doctrines structure incentives and disincentives towards certain societal ends. Is it possible that these ends might be achieved more surely and securely through other means (for example through different regulatory modalities),⁴⁶ or through ‘technological management’ of behaviour?⁴⁷

Fifth, and finally, focusing on particular interfaces between the technology and the law presumes stasis in reality, in technological processes, in legal processes, and in societal processes. This is implied in the focus on deriving a straightforward solution once a legal problem has been identified. The reality is very different. Technologies change and develop over time, often rapidly. DeepFakes

⁴¹ Hin-Yan Liu, ‘Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems’ in Nehal Bhuta and others (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (2016).

⁴² Engler (n 33).

⁴³ cf. Henry Farrell and Bruce Schneier, ‘Common-Knowledge Attacks on Democracy’ (Berkman Klein Center 2018) Research Publication 2018–7 <<https://papers.ssrn.com/abstract=3273111>> accessed 13 January 2019.

⁴⁴ Jennifer Kavanagh and Michael D Rich, ‘Truth Decay: A Threat to Policymaking and Democracy’ (RAND 2018) RB-10002-RC <https://www.rand.org/pubs/research_briefs/RB10002.html>.

⁴⁵ As defined by Regina Rini, the epistemic backstop reflects that ‘our reasonable trust in the testimony of others depends, to a surprising extent, on the regulative effects of the ever-present possibility of recordings of the events they testify about.’ Regina Rini, ‘Deepfakes and the Epistemic Backstop’ (2019) 1 <<https://philpapers.org/archive/RINDAT.pdf>>.

⁴⁶ Lessig, ‘The New Chicago School’ (n 37); Lessig, ‘The Law of the Horse’ (n 37).

⁴⁷ Roger Brownsword, ‘In the Year 2061: From Law to Technological Management’ (2015) 7 *Law, Innovation and Technology* 1.

have gone from ‘unknown’ to ‘minor misuse’ to ‘prospective criminal phenomenon’ in under four or five years. In that time, there have been rapid shifts in use cases, as well as important changes in fidelity that have rapidly rendered obsolete assumptions that were once valid. After the publication of early technological proposals for automated DeepFake detection, which relied on DeepFakes (at the time) inability to render teeth or to have subjects blink,⁴⁸ it took only a few months for new techniques to emerge which corrected these imperfections.⁴⁹ As such, the discrete solve-it-once-and-for-all approach fails to track the continuous co-evolution of technological and societal changes, and cripples any resulting legal or regulatory responses. The risk is that initial, domain-focused legal responses may treat a problem as solved as soon as its latest symptom has been addressed. This then closes the door to sustained investigation or more anticipatory regulation that can reckon with longer-term, more cross-sectoral threats. As Lyria Bennett Moses observes, ‘[t]he special relationship between legal problems and technological change can be seen by examining the *timing* of legal problems arising from a combination of technological and behavioural changes’.⁵⁰ Her argument suggests that the legal problems associated with the introduction of new technologies precede widespread social acceptance and adoption of the technology in question, and the social impact that this implies. If nothing else, however, the staggered nature of the legal challenges introduced by new technologies suggests the presence of several stages where new technologies precipitate legal problems, and that each new technology potentially presents a cascade of legal questions.⁵¹

In short, the current domain-specific approach is undermined by being technologically-focused, fragmented in its approach, and isolated in its responses and impact. A general model of the legal disruption induced by AI would provide an organising framework capable of differentiating the significance of the legal questions identified in each domain, and would have the flexibility and dynamism to respond to the shifting legal challenges that this new technology triggers. To put this another way, one reason for favouring a general model is that greater regulatory consistency could be achieved. Our responses would cut across fields of activity, minimising the variation amongst regulatory responses in different sectors. Rather than having one regulatory regime arise for the use of autonomous vehicles, and another to govern the use of autonomous weapons systems and another for the use of algorithmic trading bots, a more consistent and coherent approach that addresses the core problems posed by the technology could be developed.⁵²

⁴⁸ Jonathan Hui, ‘How Deep Learning Fakes Videos (Deepfake) and How to Detect It?’ (*Medium*, 8 October 2019) <https://medium.com/@jonathan_hui/how-deep-learning-fakes-videos-deepfakes-and-how-to-detect-it-c0b50fbf7cb9> accessed 11 December 2019; Yuezun Li, Ming-Ching Chang and Siwei Lyu, ‘In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking’, *2018 IEEE International Workshop on Information Forensics and Security (WIFS)* (IEEE 2018) <<https://ieeexplore.ieee.org/document/8630787/>> accessed 11 December 2019.

⁴⁹ See Engler (n 33).

⁵⁰ Lyria Bennett Moses, ‘Why Have a Theory of Law and Technological Change?’ (2007) 8 *Minnesota Journal of Law, Science & Technology* 589, 599 (Emphasis added).

⁵¹ At minimum, these would include: the new possibilities unlocked at the proof-of-concept stage; the possible and actual legal issues raised by the first wave of limited deployment; the more widespread considerations attendant with mass adoption; and the legal concerns triggered by shifts in society engendered by that technology.

⁵² The question over whether legal ‘coherence’ will remain a primary desideratum for regulators, or whether it may come to be supplemented or even replaced by measures of instrumental effectiveness, is taken up by much of the work of Roger Brownsword. See Roger Brownsword, ‘Law and Technology: Two Modes of Disruption, Three Legal Mind-

Along with these five reasons for developing a general holistic model to respond to the legal disruption caused by a new technology, there is a sixth AI-specific reason for favouring a general model. Some people may argue that the legal challenges posed by AI are to some extent unprecedented: artificial applications or non-human entities which some think should be recognised as legal persons by virtue of their intelligence, capacities or appearance.⁵³ Thus, a large cluster of legal questions posed by AI arises from its apparent liminal status between agents and objects which does not fit easily into the duality enforced by ordinary legal processes. This novelty suggests that AI can serve as a portal through which the legal system can be examined from different perspectives. In other words, while existing law holds the potential to be complete in relation to human beings and human activities, it is currently incapable of being complete with respect to artificial agents and objects, which either carve out new areas that are devoid of regulation, or which reveal gaps or shortcomings in the existing legal order. This suggests that engaging with the general legal challenges raised by AI can serve to teach us new lessons about the law, and help refine legal principles and processes which have hitherto not required precise fine tuning.⁵⁴

The above, of course, assumes that there is something genuinely disruptive about AI. Perhaps this is not true? Some will argue that the hype about the disruptive impact of a new technology is simply just that, and that existing legal doctrine and regulatory processes are largely or entirely able to accommodate or adapt to the new situation.⁵⁵ The iterative nature of the law, for example, has meant that the legal system can be surprisingly resilient to moderate exogenous shocks through gradual adaptation and incremental change.⁵⁶ What staunch adherence to the law often overlooks is not necessarily what is new about a technology or what gaps or inadequacies it might reveal, but how existing legal equilibria might be perturbed by the introduction of new technologies in an indirect, tangential and second-order fashion.⁵⁷ As such, while AI might indeed ‘prove [to be] “exceptional” in the sense of occasioning systematic changes to law, institutions, and the legal

Sets, and the Big Picture of Regulatory Responsibilities’ (2018) 14 *Indian Journal of Law and Technology* 1; Brownsword, ‘Law Disrupted, Law Re-Imagined, Law Re-Invented’ (n 39).

⁵³ We should note here that other work explicitly downplays the significance of ‘intelligence’ as a defining or relevant characteristic of the challenges posed by AI; see Liu, ‘The Power Structure of Artificial Intelligence’ (n 8); Hin-Yan Liu, ‘From the Autonomy Framework towards Networks and Systems Approaches for “Autonomous” Weapons Systems’ (2019) 10 *Journal of International Humanitarian Legal Studies* 89. The point here is that these characteristics of AI still dominate the definitions that are proposed, and thus heavily influence the types of legal challenges that are identified as a result.

⁵⁴ Questions of responsibility are perhaps the best example of this. Before the advent of robotics and AI, the responsible entity has been either individual or collectives of human beings and the main questions revolved around culpability and punishment. Yet, because of the difficulties of attributing responsibility to artificial entities or applications, responsibility gaps have emerged, thereby introducing substantial uncertainty in an area which before did not raise any significant questions. See the example of the responsibility problems raised by autonomous weapons systems in Robert Sparrow, ‘Killer Robots’ (2007) 24 *Journal of Applied Philosophy* 62; Marcus Schulzke, ‘Autonomous Weapons and Distributed Responsibility’ (2013) 26 *Philosophy & Technology* 203; Liu, ‘Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems’ (n 41).

⁵⁵ See for example Martina Kunz and Seán Ó hÉigeartaigh, ‘Artificial Intelligence and Robotization’ in Robin Geiss and Nils Melzer (eds), *Oxford Handbook on the International Law of Global Security* (Oxford University Press 2020) <<https://papers.ssrn.com/abstract=3310421>> accessed 30 January 2019.; Thomas Burri, ‘International Law and Artificial Intelligence’ (2017) 60 *German Yearbook of International Law* 91.

⁵⁶ JB Ruhl, ‘Law’s Complexity: A Primer’ (2007) 24 *Georgia State University Law Review* 885.

⁵⁷ Jack M Balkin, ‘The Path of Robotics Law’ (2015) 6 *California Law Review Circuit* 17.

academy’,⁵⁸ in the end, such effects might occur in more diffuse and subtle ways. Thus, even if AI does not necessarily embody new challenges that sets it apart in a direct and explicit manner, it may be that vigilance is necessary because of unanticipated effects.⁵⁹ This gets overlooked if the legal focus is predominantly tuned into resolving problems posed by AI in discrete domains, which in turn renders the legal order especially vulnerable to the turbulence that AI might trigger.

In summary, a general model of AI-induced legal disruption is justified because it (a) ensures a more consistent, dynamic and joined-up approach to the regulatory challenges posed by AI; (b) allows us to appreciate some of the unique and potentially unprecedented effects that AI may have on legal systems that might otherwise be overlooked, and (c) more generally allows for greater vigilance with respect to the unanticipated, second-order effects that might arise from the widespread use of AI. The argument can be bolstered by considering some of the recent debates about the impact of technology on law in a little more detail.

2. Towards a General Model: Lessons from the Cyberlaw and Robolaw Debates

If a general model is required, what form might it take and what pitfalls should be avoided when trying to develop it? We can help to answer this question by considering past debates about the legal disruption caused by new technologies.

Consider, in the first place, the old debate in cyberlaw between Frank Easterbrook and Lawrence Lessig.⁶⁰ This debate revolved around the question of whether a new discipline of ‘cyberlaw’ would be necessary to study the impact of information communication technologies on the legal system. Easterbrook famously deployed the analogy of ‘The Law of the Horse’, claiming it was as preposterous to establish a distinct study of legal problems in cyberspace as it would be to compile a distinct study of all the specialised interactions between horse-related activities and the law. Instead, Easterbrook asserted that the study of general rules is the best way to learn about the law applicable to specialised endeavours, and that law school “courses should be limited to subjects that could illuminate the entire law”.⁶¹ Lessig agreed that the illumination of the entire law should be the purpose of legal education, but argued “that there is an important general point that comes from thinking in particular about how law and cyberspace connect”.⁶² In particular, Lessig argues that it shed light on “the limits on law as a regulator and about the techniques for escaping those limits”.⁶³

⁵⁸ Ryan Calo, ‘Robotics and the Lessons of Cyberlaw’ (2015) 103 *California Law Review* 513 See especially 550-562.

⁵⁹ A different way of putting this claim is that robotics and AI may merely constitute a hazard, but whether the legal order succumbs to the shock of that hazard depends upon the latent vulnerabilities of that system, and the exposure that that system has to the hazard. For an example of this differentiation (in the context of existential risk research) see Hin-Yan Liu, Kristian Cedervall Lauta and Matthijs M Maas, ‘Governing Boring Apocalypses: A New Typology of Existential Vulnerabilities and Exposures for Existential Risk Research’ (2018) 102 *Futures* 6.

⁶⁰ Frank H Easterbrook, ‘Cyberspace and the Law of the Horse’ (1996) 207 *The University of Chicago Legal Forum* 11; Lessig, ‘The Law of the Horse’ (n 38).

⁶¹ Easterbrook (n 60) 207.

⁶² Lessig, ‘The Law of the Horse’ (n 37) 502.

⁶³ *ibid.*

Reconciling these positions requires aligning divergent impulses: on the one hand, legal problem-solving seeks to fill gaps in the law to promote certainty and predictability; on the other hand, there may be much to learn about the structures and systems of the law from those very ambiguities and anomalies in legal principles and processes that are caused by technological disruption. At a general level then, jurists are trained to think in ways that resolve specific questions, in manners that re-establish the existing legal equilibrium. Yet, this impulse also closes off those rare opportunities to examine the law from a different vantage point in light of a new technology. For instance, the perspective offered or provoked by AI reveals many of the tenuous presumptions and precepts upon which much of the law rests. The impulse to quickly close the gaps revealed by the introduction of AI also undermines legal doctrine, insofar as structural or systemic adaptations are not contemplated. Thus, studies at the intersection of AI and the law hold the potential to, as Easterbrook phrased his challenge, ‘illuminate the entire law’.⁶⁴ and ‘identify the limits of the law as regulator and about the techniques for escaping those limits’.⁶⁵

To illustrate this point, consider Lessig’s identification of the four regulatory modalities in his effort to understand the impact of the cyber revolution on the legal system: law, social norms, market, and architecture/code. Identifying these modes of regulation significantly advanced understanding of the strategies for escaping the limits of law by subsuming law under a broader framework of regulatory strategies that could be deployed to the same behavioural end.⁶⁶ This opened up the discussion of regulation beyond the confines of legal discussions that would not have been possible if we limited ourselves to Easterbrook’s traditional approach. Thus, Lessig not only provided an approach capable of illuminating the entire law, but also one that allowed us to view the law in a relational manner – as one regulatory modality among others. This not only determined some of the boundaries and limitations of law, but also strategies for overcoming those limitations. This provides an example of the claim made above: that orthodox responses to a legal system destabilised by technological change close off potential avenues for greater insight and rejuvenation.

Building upon Lessig’s model, Roger Brownsword has argued that the modalities of regulation may not be atemporal and overlapping, as Lessig’s model presumes. There may instead be a distinct developmental arc to regulation that is facilitated by technological change.⁶⁷ This developmental arc moves us away from reliance on traditional rule-based legal regulation, to the gradual deployment of systems of ‘technological management’, as the requisite technologies mature. Brownsword imagines three ideal-type generations of regulators. The first rely solely on normative signals being sent from the regulators to the regulatees; the second use design and (physical or technological) architecture to incentivise compliance with the behavioural norms; the third tries to build regulatory compliance into social systems directly, often by bypassing human behaviour. Brownsword’s model suggests a general push away from ‘law’ as the dominant or primary modality of regulation (even if

⁶⁴ Easterbrook (n 60) 207.

⁶⁵ Lessig, ‘The Law of the Horse’ (n 37) 502.

⁶⁶ Lessig, ‘The New Chicago School’ (n 37); Lessig, ‘The Law of the Horse’ (n 37).

⁶⁷ Roger Brownsword, ‘In the Year 2061: From Law to Technological Management’ (2015) 7 *Law, Innovation and Technology* 1.

this trend may not be readily evident), thereby diminishing the scope and force of law as well as legal procedures for checking the exercise of power.⁶⁸

Observing these subtle structural shifts in the law, in regulatory modalities, and in their particular configuration is one of the under-emphasised opportunities for legal research emerging out of the cyberlaw debate – one that might usefully be built upon in contemporary discussions of AI.

There are also lessons to be learned from recent debates in the field of ‘RoboLaw’ concerning the disruptive impact of robotics technology on the law.⁶⁹

Ryan Calo initiated this debate through his attempt to draw out the lessons learnt from cyberlaw that would be applicable to the regulation of robotics. His claim was that robotics were poised to become the next transformative technology by combining the promiscuity of information with the capacity to inflict physical harm. His overarching aim was to identify the unique features of robotic technologies that make them legally interesting: an approach to regulation that seeks to identify certain “essential qualities” of technologies and draw out the legal implications of these qualities. For the internet, he argued, the relevant essential qualities were *connection*, *community* and *control*; in contrast, he claimed that the essential qualities of robotics were *embodiment*, *emergence* and *social valence*.⁷⁰ Taken in turn, *embodiment* is the physical presence of robots and the physical impact or manipulation made possible by robotics. *Emergence* encompasses the greater latitude of behaviour than mere programming might dictate, raising the question of effective or exclusive control that leads to the prospect of unpredictability and uncertainty by design. Finally, *social valence* invokes the different feelings that robots trigger among those who interact with them. The feeling is that robots constitute a separate ontological category, not quite agent but not quite object either. The overarching question for robotics regulation from this perspective is the extent to which robots are exceptional when compared to other technologies.⁷¹ Thus, Calo’s approach to robotics regulation attempts to distil stable and inherent characteristics that robots possess, which can then be factored into an ‘exceptionalism’ calculus to determine whether or not robots create or reveal exceptional systemic challenges to the existing law or legal order.

Jack Balkin’s response to Calo argued for a more contextualised approach to the regulation of robotics. Balkin suggested that the “essential qualities” approach was misleading, because it implied that the challenges raised by the technologies are independent from the use of those technologies and their mediation of human relations. As such, the essential qualities approach insulates the characteristics of the technology from contextual considerations which likely raise a very different set of challenges than might be predicted by examining the essential qualities in isolation. Furthermore, he argued that Calo’s approach may unjustifiably ossify a set of contingent essential qualities and the challenges they pose, whereas in reality robotics regulation is likely to be an organic process of continual negotiation and renegotiation.

⁶⁸ Liu, ‘The Power Structure of Artificial Intelligence’ (n 8); Hin-Yan Liu, ‘The Digital Disruption of Human Rights Foundations’, *Human Rights, Digital Society and the Law: a Research Companion* (Routledge 2019).

⁶⁹ Calo, ‘Robotics and the Lessons of Cyberlaw’ (n 58); Balkin (n 57).

⁷⁰ Calo, ‘Robotics and the Lessons of Cyberlaw’ (n 58) 532–548.

⁷¹ In Calo’s conception, “a technology is exceptional if it invites a systemic change to laws or legal institutions in order to preserve or rebalance established values.” *ibid* 553.

Balkin proposed that *salience* should instead be the organisational framework for robotics regulation. Regulators should ask: what is foregrounded and what is relegated through the introduction of the technology? This in turn captures two factors: that salience depends upon the use of the technology, including uses that were neither intended nor foreseen; and that technology mediates social interactions, in part by creating new opportunities for power and control.⁷² Thus, salience captures both how people interact with new inventions as well as how people interact with other people through those new inventions. This approach allows us to see past projections of contemporary concerns that are espoused by the essential qualities approach and embraces the dynamic social world that is the scene of regulation.

The salience approach also enables regulatory initiatives to grapple with complex interaction effects between the technology and human society, rather than assuming ‘complex usage’ to be an essential quality of the technology. Interaction effects are by definition dynamic. While similar ideas of unforeseeability and unpredictability are captured within Calo’s concept of emergence, his approach is much more static because it defines emergence as one of the essential properties of robots and not something that arises from social use. Because of this, the salience approach offers benefits over the essential qualities approach, which may be overly restrictive for three reasons.

First, as Balkin observes, that approach projects contemporary concerns into the future. While such projection may be unavoidable, the essential qualities approach ossifies present concerns about the future as those that are paramount in importance, even if the configuration of ‘essential qualities’ might alter over time. He describes how legal scholarship on cyberspace altered over time, as the technology saw different usage. Thus, a scholar writing in 1991, he argues, might well perceive that “the Internet’s “essential” features were its abilities to cross jurisdictional lines at will, to send digital information quickly and cheaply, and to facilitate anonymous communication”.⁷³ By 1999, however, as it became increasingly clear that the degree of online anonymity was overstated and that states could control features of Internet traffic, dominant legal concerns instead turned to filtering and surveillance. Then by 2008, the Internet’s dominant features seemed to be “the ability to lower the cost of organization, to facilitate crowdsourcing and open-source projects, and to undercut professional norms of information production in areas ranging from music to journalism to science.”⁷⁴ Thus, Calo’s account reflected a specific snapshot of the technology rather than a timeless account of the technology’s immutable characteristics.

Second, the list of identified essential qualities, beyond being somewhat arbitrary and contingent to its specific time, may also be forever incomplete.⁷⁵ For instance, early international treaties governing naval warfare included prohibitions on firing on civilian vessels without prior warning, and established a duty to rescue survivors. Rebecca Crootof, however, has discussed how the rise of submarine warfare functionally overturned assumptions about what might be termed as ‘essential qualities’ of naval military technology – namely the presumption of the dominance of surface vessels with spare capacity to accommodate survivors – with the result that customary international

⁷² Liu, ‘The Power Structure of Artificial Intelligence’ (n 8).

⁷³ Balkin (n 57) 47.

⁷⁴ *ibid* 48.

⁷⁵ Matthijs M Maas, ‘Innovation-Proof Governance for Military AI? How I Learned to Stop Worrying and Love the Bot’ (2019) 10 *Journal of International Humanitarian Legal Studies* 129.

law rendered the existing treaty regime ‘jurisprudential space junk’.⁷⁶ The problem with it being impossible to exhaustively consider or list all ‘essential qualities’ is inherent to any dichotomous taxonomy: what methodology is utilised to identify these essential characteristics, and in particular what are the threshold values, and the mechanisms for promoting and demoting characteristics over time? With a necessarily incomplete list of essential qualities, the prospects of designing a robust regulatory framework for AI diminishes.

Third, the essential qualities approach is overly isolating and introspective in approach, which may result in false positives or false negatives. Looking solely at which assumed ‘essential characteristics’ of an emerging technology may become challenging can give clues as to the problems that *could* conceivably arise, but whether those problems actually become a widespread phenomenon depends upon contextual factors and usage.

But how does the essential qualities and salience debate factor into the earlier critiques against adopting a piecemeal approach?⁷⁷ At first glance, it would appear that the essential qualities approach can be read as being more in line with the spirit of an holistic approach insofar as it seeks out the root causes of the disruptive challenge and directs attention and responses to those root causes. This would also suggest that the salience approach is more piecemeal by virtue of its contextual and reaction character. Yet, we would suggest that the opposite conclusion holds true: that the essential qualities approach is, almost by definition, more piecemeal as a direct result of the need to select or foreground certain characteristics above others as those meriting attention and being definitive of the new technology. Since it would be impossible to take into consideration all the characteristics of a new technology, and since it would defeat the purpose of the analysis if the essential criteria were removed, the essential qualities approach is not as holistic as it might appear to be. Conversely, by taking into consideration the broader contexts into which a new technology is deployed, the salience approach is less piecemeal even though it is more contingent – in other words, the context is a core aspect of forging a holistic response.

As should be clear from this, we are more sympathetic to Balkin’s approach, and in particular to his point that embodiment does not constitute an essential quality that sets robotics aside from other manifestations of AI. The core challenge introduced by robotics and AI is not necessarily that non-human entities are the means through which actions are performed, but rather that they can have an influence over the ends that society pursues.

But while our sympathies lie closer to the contextualisation espoused by Balkin’s “salience” criterion, there remain shortcomings with his approach as well. In his own words:

When we consider how a new technology affects law, our focus should not be on what is *essential* about the technology but on what features of social life the technology makes newly *salient*. What problems does a new technology place in the foreground that were previously underemphasized or deemed less important? What

⁷⁶ Rebecca Crotof, ‘Jurisprudential Space Junk: Treaties and New Technologies’ in Chiara Giorgetti and Natalie Klein (eds), *Resolving Conflicts in the Law* (2019) <<https://brill.com/view/book/edcoll/9789004316539/BP000015.xml>> accessed 15 March 2019; See also the discussion in Matthijs M Maas, ‘International Law Does Not Compute: Artificial Intelligence and The Development, Displacement or Destruction of the Global Legal Order’ (2019) 20 *Melbourne Journal of International Law* 29, 36–38.

⁷⁷ We would like to thank Roger Brownsword for the ideas developed in this section.

aspects of human activity or of the human condition does a technological change foreground, emphasize, or problematize? And once technology moves a certain problem or a certain feature of our lives from the background to the foreground of our concern, what are the consequences for human freedom?⁷⁸

From this passage, it becomes clear that salient features of a technology are those that impinge upon latent aspects of social life: in other words that the problems are revealed by the technology, and not necessarily created nor introduced by the technology. Thus, while Balkin recognises the shifts in social relations that are triggered by new technologies, he at least understates the fact that new technologies can create both new problems as well as novel categories of problems.

A related shortcoming of the salience approach to regulation is that there is no metric provided that gauges the relative importance or severity of the features that are elevated or relegated by the technology. As a measure of relevance, the salience approach is capable of sifting through problems that are brought to light or magnified by the technology, but once a problem is deemed to be salient, there is no further hierarchy to organise and prioritise regulatory responses. In a sense, with the salience approach the pendulum has swung to the opposite extreme to that of the essential qualities approach: where the former emphasises relevance and social dynamics, the latter focuses upon ranking inherent features of the technology. This suggests that even the salience approach to robotics regulation remains incomplete, and that it is the first step of a multistage procedure that would structure the regulatory challenges posed by AI and robotics by importance or impact, after a particular feature has been deemed to raise salient considerations.

3. Artificial Intelligence and Legal Disruption

We now turn to our proposed model of AI-induced legal disruption, which we call the ‘Legal Disruption Model’. This model is both a lens to identify which new problems will prove to be the most structurally challenging for regulation, and is also a theoretical attempt to draw on the disruptive potential of AI in order to generate new perspectives upon the workings of law and legal system itself.

An important assumption framing this model is that only AI or the manifestations of AI that are capable of fundamentally displacing certain core legal presumptions, subvert legal principles, or systematically distort the functioning of the legal system will be considered ‘legally disruptive’.⁷⁹ In other words, AI systems and their manifestations must raise structural or systemic challenges to regulation and governance to be included within this model. As such, the legal disruption

⁷⁸ Balkin (n 57) 46–47. Emphasis original.

⁷⁹ Legal disruption, as a general concept or approach, can be applied to anything that is capable of destabilising legal principles or practices. For example, the rise of the Private Military Company, a private corporation providing martial services for profit disrupts the legal mechanisms structuring control and responsibility for the use of force in situations of armed conflict. Hin-Yan Liu, *Law’s Impunity: Responsibility and the Modern Private Military Company* (Hart Publishing 2015). This suggests that the Private Military Company is capable of legal disruption, at least in those legal areas associated with armed conflict and criminal conduct until legal doctrine proves capable of accommodating for its status and activities. In this context, the application of legal disruption to artificial intelligence is just a special and particularly visible case.

framework sets a high threshold for engagement with AI and is aimed only at identifying aspects of technological change which are turbulent and affect the foundations of the legal and regulatory order. Although we present this model as an attempt to address the impact of AI, we believe it has broader utility and could be applied to other debates about disruptive technologies and their impact on the legal system. This will become apparent in our description of the model. We first discuss the model and then outline some of its advantages and disadvantages.

Stage One: The Disruptive Moments

The Legal Disruption Model starts with the ‘disruptive moment’ that is heralded once actors within the legal system perceive the injection of AI systems into society as creating legally relevant problems.⁸⁰ It invokes the idea that the legal and regulatory system is thrown into dis-equilibrium as a result of AI itself, or of the resultant sociotechnical change that is (perceived to be) brought about by its adoption. Implicit in the legal disruption model, then, is the idea of a relatively sharp and identifiable departure from the ordinary legal and regulatory processes currently at play.

So the first question is: how does this disruptive moment come about, and how can we identify it when it does? In this context, it might be analytically productive to draw on the specific concept of ‘affordances’ to explain the process of disruption. An outline of this argument, which we will unpack in detail below, is that new social challenges often result when new technologies are translated into new affordances, i.e. new possibilities for behaviour or action in our personal and social environments.⁸¹ Disruptive moments can come about when new technologies generate or otherwise reveal or unlock new affordances; when such affordances are actualized (they are recognized and acted upon); and when the resulting behaviour is deemed a problem or a hazard by the legal system.⁸²

While it may be tempting to deploy the concept of affordances in a manner that directly connects AI applications to the new possibilities for new behaviours, this is only one part of the overall contribution that the affordances approach might present. A new technology like AI may indeed enable behaviours or outcomes that have been desirable, yet hitherto impossible. Yet, there are at least two other ways that the concept of affordances may be connected to the disruptive moment. In

⁸⁰ One could also compare this with Bruce Ackerman’s similar (though distinct) concept of a ‘constitutional moment’; see Bruce Ackerman, *We the People, Volume 1: Foundations* (Reprint edition, Belknap Press: An Imprint of Harvard University Press 1993); Bruce Ackerman, *We the People: Volume 2: Transformations* (Belknap Press: An Imprint of Harvard University Press 2000); And for a critique, see Michael J Klarman, ‘Constitutional Fact/Constitutional Fiction: A Critique of Bruce Ackerman’s Theory of Constitutional Moments’ (1992) 44 *Stanford Law Review* 759.

⁸¹ Bennett Moses (n 50) 591: (“When lawyers claim to be interested in issues surrounding law and technology, it is usually related to questions of how the law ought to relate to activities, entities, and relationships made possible by a new technology. As technology changes, we can do things, make things, and form connections that were not previously practicable.”).

⁸² The relationship may be more subtle than this, because new kinds of responses can be facilitated by these new technologies. This is a distinction that we elaborate upon below in relation to intrinsic and extrinsic versions of legal disruption.

the section that follows, we first set out an adapted version of the three different types of affordances and then discuss the implications this has in the lead up to the disruptive moment.

Legal Disruption Arising from New Affordances

The concept of ‘affordance’ was introduced by James Gibson to make simultaneous reference to *all* action possibilities available to an animal in its environment:⁸³ ‘the *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill’.⁸⁴ Thus, the concept of affordances initially mapped onto the potentialities of the environment for an agent. The concept of affordances was subsequently adapted and popularised by Don Norman in the context of (technological) artefacts:

The term affordance refers to the relationship between a physical object and a person (or for that matter, any interacting agent, whether animal or human, or even machines and robots). An affordance is a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used.⁸⁵

In addition to emphasising the relational dimension of affordances, Norman also further distinguished *affordances* from *signifiers*: ‘Affordances determine what actions are possible. Signifiers communicate where the action should take place. We need both.’⁸⁶ For Norman, separating out signifiers from affordances remedied the question of recognition because ‘[t]o Gibson, an affordance exists whether or not anyone ever notices it. To the designer, if affordances are not known, then they might as well not exist. In other words, the designer is primarily concerned with *perceived* affordances; the perception is critical’.⁸⁷

This refinement becomes especially pertinent for our model of legal disruption because the range of possible actions open to an agent in the designed virtual, or the virtually-mediated, realm hinges upon those opportunities that the agent actually registers as possibilities – recognition that hinges upon successful signifiers. This is important when the designed or engineered world effectively becomes synthetic⁸⁸ or artefactual,⁸⁹ and differs from the physical world with respect to our habitual and intuitive sense of the possibilities and limitations on our behaviour. Furthermore, if artificial intelligence does in fact invoke the artefactual, either in the virtual worlds it creates or in the mediation between us and the worlds that we perceive and inhabit, there is the further consideration that the artefact is created expressly to serve human interests and ends.⁹⁰ If this distinction holds, the question of affordances and signifiers become a central arena of contestation because they precede, and therefore occlude, other forms of regulation or behavioural influences. This why Norman

⁸³ Most famously in James J Gibson, ‘The Ecological Approach to the Visual Perception of Pictures’ (1978) 11 Leonardo 227.

⁸⁴ James J Gibson, *The Ecological Approach To Visual Perception* (Erlbaum 1986) 127. Note that Gibson was a psychologist specialising in visual perception, hence the reference to ‘animal’ and ‘environment’.

⁸⁵ Donald A Norman, *The Design of Everyday Things* (MIT Press 2013) 11.

⁸⁶ *ibid* 13.

⁸⁷ Donald A Norman, *Living with Complexity* (MIT Press 2016) 228. Emphasis original.

⁸⁸ Christopher J Preston, *The Synthetic Age: Outdesigning Evolution, Resurrecting Species, and Reengineering Our World* (MIT Press 2019).

⁸⁹ Keekok Lee, *The Natural and the Artefactual: The Implications of Deep Science and Deep Technology for Environmental Philosophy* (Lexington Books 1999).

⁹⁰ *ibid*.

demanding that ‘Designers are still responsible for ensuring that the objects and systems they design have the proper affordances, but if they are not noticed or perceived, then they will probably fail to achieve their purpose. The designer must communicate the range of actions through signifiers. Signifiers are critical to effective communication’.⁹¹

The crucial difference between *design* and *disruption* in the context of affordances is thus the perception and realisation of what might be done. In the former, design serves a particular purpose, aim, or objective, and the role of affordances and attendant signifiers is to facilitate people to achieve the ends that are served by the designed artefact. In the latter, the prospect for disruption inheres in the potential absence of signifiers and in the broad, yet unperceived or unrealised, affordances that may lurk within a particular artefact or system. Thus, disruptive moments can arise from sudden or lurching discoveries of affordances that were hitherto hidden or opaque, or, at any rate, lacking effective signifiers. Put differently, it appears as though disruption invokes a more raw Gibsonian understanding of affordances, which exist whether or not they are recognised, with the disruptive moment taking place when a hitherto hidden affordance is suddenly perceived and acted upon. In this line of thinking, the disruptive moment need not (only) inhere in the invention of a particular technology: rather, disruption (also) arises when new uses for that technology are recognised and deployed. This reaffirms the cascades of disruptive effects that follow from the invention of revolutionary technologies: as a ‘general purpose technology’,⁹² the vast space of potential use cases for AI unfold gradually, yet in a potentially disruptive manner, within this conceptual model of affordances. This downplays the role of the technology itself in the process of legal disruption, and instead shifts the focus to the social, political, and cultural dimensions that spark crises for the legal order.

To explore this relationship between affordances and legal disruption, we appeal to an additional distinction that Vlad Glăveanu has previously drawn between three types of affordances: (a) “uninvented affordances”, connoting the current lack of a technological capacity; (b) “unperceived affordances”, implying the lack of awareness of a new affordance; and (c) “unexploited affordances” which involve normative resistance against pursuing particular forms of behaviour which are otherwise possible and realisable.⁹³ In the context of AI, examples of these three types of affordances could be (a) ‘mind uploading’ or ‘whole brain emulation’; (b) ‘zero-day’ security vulnerabilities or back routes into certain systems, for instance the susceptibility of a specific AI architecture to particular ‘adversarial input’;⁹⁴ (c) the use of AI systems for optimizing the

⁹¹ Norman (n 87) 229–230.

⁹² Trajtenberg (n 7); Agrawal, Gans and Goldfarb (n 7) pt 1 ‘AI as a GPT’.

⁹³ Vlad P Glăveanu, ‘What Can Be Done with an Egg? Creativity, Material Objects, and the Theory of Affordances’ (2012) 46 *The Journal of Creative Behavior* 192, 196–199. We particularly refer the reader to the visualization on page 197 and we slightly altered the original categories for relevance and consistency in this paper.

⁹⁴ For obvious reasons, it is hard for us to speculate, in advance, about examples of ‘unperceived affordances’ offered by AI systems. Recent work on the susceptibility of AI systems to ‘adversarial input’, however, suggests we might expect many such systems to have such vulnerabilities that allow exploitation. Ian J Goodfellow and others, ‘Attacking Machine Learning with Adversarial Examples’ (*OpenAI*, 16 February 2017) <<https://openai.com/blog/adversarial-example-research/>> accessed 18 February 2017.

effectiveness of torture respectively.⁹⁵ Which of these types of affordances are the most relevant from the perspective of legal disruption?

In the most straightforward and direct sense, the introduction of a new technology overcomes the ‘uninvented affordances’ obstacle because that new technology makes achievable certain possibilities of action that are already sought by individuals and society. Put into the context of the above debate between essential qualities and salience, ‘uninvented affordances’ appear to fall in line with the former because the focus is placed upon what is technologically novel and possible, thereby also aligning this approach with those that treat technology as the regulatory target.⁹⁶ For instance: the desire towards implementing systems of surveillance that make social life completely ‘transparent’, ‘legible’ and controllable or profitable, is surely not new. Indeed this has been one that various actors have long harboured, for strategic or political reasons.⁹⁷ Yet the introduction of many sensor technologies, coupled with big data and AI analysis, may have rendered widespread systems of such pervasive surveillance and ‘algocracy’ (much more) possible.⁹⁸ As the past lack of technical capacity has been overcome, such situations often appear as though the new technology has unlocked or revealed new types of behaviour in a direct manner. It may even be treated as the technological disruption of society, whereby a technological application or artefact is conceived of as an independent external hazard that impacts and destabilises societal equilibria. For example, a stated aim in relation to IHL is to reduce unnecessary and superfluous suffering in the conduct of armed hostilities. The prospect of introducing autonomous weapons systems, from the perspective of their advocates, is that superior compliance with both the principles of distinction and proportionality at the heart of IHL is *now made possible* by the military applications of AI. Thus, the presently-desirable objectives in IHL align with the promised capabilities of AI in the battlefield, making it appear as though AI has overcome the ‘uninvented affordance’ obstacle. Yet, this potentially leads to the destabilisation of the framework that regulates armed conflict in that legal objections to autonomous weapons systems that are grounded in IHL become neutralised, perhaps to the point where the deployment of such weapons systems become morally virtuous or obligatory.⁹⁹

This first category of affordance, however, might overshadow the other two categories of affordances, because the presence or absence of a new technology is only a part of the equation of legal disruption. It is therefore critical that the model of legal disruption is not overly blinkered to take into account only, or primarily, disruptive moments arising from uninvented affordances. Different types of disruptive moments can also arise when “unperceived affordance” and

⁹⁵ However, see, unfortunately, Amanda McAllister, ‘Stranger than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture’ [2018] Minnesota Law Review 47.

⁹⁶ Gary Marchant, Braden Allenby and Joseph Herkert (eds), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (Springer 2011).

⁹⁷ On the long historical drive, by (especially modernist) states or elites, towards rendering societies ‘legible’, see JC Scott, *Seeing Like a State - How Certain Schemes to Improve the Human Condition Have Failed*. (Yale University Press 1998).

⁹⁸ Cf. Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1 edition, PublicAffairs 2019); John Danaher, ‘The Threat of Algocracy: Reality, Resistance and Accommodation’ (2016) 29 *Philosophy & Technology* 245.

⁹⁹ Liu, ‘From the Autonomy Framework towards Networks and Systems Approaches for “Autonomous” Weapons Systems’ (n 53).

“unexploited affordance” obstacles are overcome: in other words, when new possibilities for behaviour or action are *recognised*, and separately, are *exploited*. Such disruptive moments are unlike instances of what might appear as technological disruption because the switch from an unperceived affordance to a recognised possibility for behaviour, or from an unexploited affordance to an exploited possibility for action can be much more abrupt. This is because there need not be a manifest technological innovation to serve as a regulatory focal point, and also because such switches can be temporally and causally disconnected from a given technological innovation, unlike that which transformed an unperceived affordance into a technically-enabled possibility for behaviour.

Continuing the example of military applications of AI, the build up to a disruptive moment in the shift from an *unperceived* affordance to a recognised possibility for behaviour precipitating a disruptive moment could involve the realisation that autonomous weapons systems could be deployed in the battlefield. This raises the familiar questions surrounding control and responsibility,¹⁰⁰ but also questions around structural strategic pressures created by the very appearance of this technology.¹⁰¹ These potentially disruptive moments need not address actual scenarios, in other words the unexploited affordance need not be overcome, because some legal challenges can be projected out from plausibly hypothetical use cases. Indeed, attempts at regulation can take place in anticipatory fashion in relation to the unrecognised affordances barriers. In some cases, ‘preventative frameworks’ have proven able to anticipate and ban the use of imminent technologies such as exploding bullets or blinding lasers.¹⁰² In other cases, such scholarship prejudged the actual state (or even viability) of the technology in question in a way that now seems odd: consider, for example, the 1960s international legal scholarship proposing a ‘Center of the Earth Treaty’; the 1970s proposals to regulate weather control technology; or the 1982 Deep Sea Bed Mining Provisions in the 1982 Law of the Sea.¹⁰³

It is thus worth noting that these are ‘just’ social problems: new technologies and the capacities that these enable may be the trigger or the catalyst, but the new technology cannot be considered as the root cause of these unrecognised affordances. As the disruptive moments that flow from unrecognised affordances are centred upon recognition and social change, these can be very sudden and abrupt as they cannot be directly inferred from the development or deployment of a new technology qua technology. From this perspective, the examples of the seemingly odd pre-emptive regulations drafted to meet new technologies that never actually materialised raised just above might appear to be prudent preparations. Insofar as unrecognised affordances can precipitate legally disruptive moments once the recognition hurdle is cleared, a rational hedge against legal disruption would be to prepare the legal system against the range of challenges that might reasonably be

¹⁰⁰ Bhuta Nehal and others (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press 2016).

¹⁰¹ cf. Zwetsloot and Dafoe (n 19).

¹⁰² Denise Garcia, ‘Future Arms, Technologies, and International Law: Preventive Security Governance’ (2016) 1 *European Journal of International Security* 94; Denise Garcia, ‘Lethal Artificial Intelligence and Change: The Future of International Peace and Security’ (2018) 20 *International Studies Review* 334; Myriam Dunn Cavelty, Sophie-Charlotte Fischer and Thierry Balzacq, “‘Killer Robots’ and Preventative Arms Control”, *Routledge Handbook of Security Studies* (Routledge 2016).

¹⁰³ Picker (n 14) 184–187; For the case of envisioned weather control technology, see Edith Brown Weiss, ‘International Responses to Weather Modification’ (1975) 29 *International Organization* 805.

foreseen in the more obvious and observable development of new technologies. Of course, the realities of a treaty-based response might counterintuitively render the legal system more susceptible to legal disruption from unrecognised affordances by ossifying a set of pre-identified legal challenges and thus eroding the adaptability and responsiveness of legal responses more broadly to meet those new challenges. It would seem then that preventative frameworks have their virtues in response to unrecognised affordances, but these must necessarily retain a high degree of flexibility and adaptability to function appropriately for the heavily contextual nature of such challenges.

Finally, there are the disruptive moments that flow from breaking the unexploited affordances obstacle, in our case where new uses for AI are deployed. For example, the actual deployment of military AI applications in the battlefield will likely precipitate legally disruptive moments that may be different from those that were identified and discussed prior to the watershed event of actual usage. While such examples are harder to pin down, we can draw upon a recent proposal to connect AI with the US nuclear force in response to the challenges of attack-time compression that threatens to destabilise present deterrence strategy.¹⁰⁴ This presents an AI use case that has overcome the invented and unrecognised affordance hurdles, but which at present is limited by the unexploited affordance barrier due to important outstanding concerns.¹⁰⁵ If and when this final barrier against exploiting these new possibilities is broken, we can expect a new and independent set of disruptive moments to flow from this reversal.

In practice, both the unrecognised and unexploited affordance hurdles often trail behind technological innovation,¹⁰⁶ and we present a sequential way of considering the affordances approach to the disruptive moment below. This trailing dissociates the legal, regulatory of governance disruption that is observed, from the original technological innovation that set the affordance cascades into motion. The point raised here from a legal disruption perspective is an important one: legally disruptive moments need not be neatly coupled to the technologies that are deemed to precipitate them. While legally disruptive effects may be hinged to new technologies when overcoming uninvented affordance obstacles, we can also expect a cascade of other less obvious forms of legal disruption to follow as a consequence of new affordances that subsequently break through the unperceived and unexploited affordances obstacles (and to which the new technology does not necessarily have a direct bearing), leading to new behaviours and generating new social challenges.

¹⁰⁴ Adam Lowther and Curtis McGiffin, ‘America Needs a “Dead Hand”’ (*War on the Rocks*, 16 August 2019) <<https://warontherocks.com/2019/08/america-needs-a-dead-hand/>> accessed 2 September 2019.

¹⁰⁵ Matt Field, ‘Strangelove Redux: US Experts Propose Having AI Control Nuclear Weapons’ (*Bulletin of the Atomic Scientists*, 30 August 2019) <<https://thebulletin.org/2019/08/strangelove-redux-us-experts-propose-having-ai-control-nuclear-weapons/>> accessed 2 September 2019; Pavel Sharikov, ‘Artificial Intelligence, Cyberattack, and Nuclear Weapons—A Dangerous Combination’ (2018) 74 *Bulletin of the Atomic Scientists* 368; Shahar Avin and SM Amadae, ‘Autonomy and Machine Learning at the Interface of Nuclear Weapons, Computers and People’ in V Boulanin (ed), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk* (Stockholm International Peace Research Institute 2019) <<https://www.repository.cam.ac.uk/handle/1810/297703>> accessed 16 October 2019; Although see for recent developments: Michael T Klare, ‘“Skynet” Revisited: The Dangerous Allure of Nuclear Command Automation’ (2020) 50 *Arms Control Today*; Washington 10; James Johnson, ‘Delegating Strategic Decision-Making to Machines: Dr. Strangelove Redux?’ (2020) 0 *Journal of Strategic Studies* 1.

¹⁰⁶ This has been under-appreciated in the context of disruptive moments, because this process may proceed at different rates in different spheres of activity. See for instance: Erik Brynjolfsson and Andrew McAfee, *The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies* (WW Norton & Company 2014) 72–88.

It is especially worth noting here that the nature and capabilities of a new technology are most relevant in terms of uninvented affordances, because the plausibility of forecasting the kinds of new behaviours that can be associated with the configuration of that new technology might still be retained. When it comes to the unrecognised and the unexploited affordance obstacles, however, the connection between forms of legal disruption and the characteristics of the new technology at play both becomes highly complex and continues to unfold over time. This is because the range of recognised and exploited uses of a new technology can depart radically from its designed or intended use. Furthermore, because realisation breakthroughs depend upon recognition and exploitation of (new, technologically-underpinned) possibilities, these can take place more suddenly, at a faster pace, and move in non-intuitive directions. This suggests a contextual dimension to legal disruption in that there will be differential regulatory impacts flowing from a particular technology among different jurisdictions and legal orders.

Breaking down a disruptive moment along these lines suggests that there will be different types of disruptive moments precipitated by new and emerging technologies, and that the manner in which these moments are constituted, and the ways in which they drive legal disruption, will differ depending on which type of affordance is being operationalised.

This affordances approach to the legal disruption model suggests that there are at least three potential types of disruptive moment that each correspond to the realisation of each type of affordance. In such a context, it becomes clear that the technologically-induced legal disruption that follows as a consequence of overcoming uninvented affordance obstacles is only the most direct, quickest, and readily-perceptible, type. This suggests that additional vigilance is required to identify and address those disruptive moments further down the line that arise when unrecognised and unexploited types of affordances become operationalised. What this means in terms of the legal disruption model is that it appears quite plausible, and indeed quite probable, for the same technological innovation to trigger a series of disruptive moments for law, regulation and governance. Following this logic, these disruptive moments will involve inflection points for change over time, but these will occur in response to different underlying reasons and can therefore appear to be unrelated because of spacing in time and effect.

The Sequential View of Legally Disruptive Moments

A different way of approaching the relationship between affordances and legal disruption is to consider the cascade in a sequential manner. As this is merely another way of organising the ideas set out in the preceding section, this will be presented in the abstract to avoid repetition of illustrative examples.

A new technology first overcomes the “uninvented affordance” obstacle, thereby introducing new capacities to realise pre-existing desires that were hitherto unachievable. The removal of such barriers to implementing desired realities is itself a potential trigger for legal disruption.

Then the second potential wave of legal disruption occurs when the “unrecognised affordance” barrier succumbs by people recognising what they can do with a new technology. This sets the scene for hypothetical and theoretical legal disruption arising from possible or plausible scenarios. At this stage, discussions can be opened up as to what types of legal disruption may arise as a consequence of new and emerging technologies in anticipation of their introduction into society.

Legal disruption can take place even absent the actual or meaningful introduction of a technology into society, insofar as thinking and debating about the mere prospect reveals latent inconsistencies or inappropriateness in the existing legal system.

A third potential wave of disruption may take place when the “unexploited affordance” obstacle is surmounted as new technologies are actually exploited and applied into society. While this last stage provides for evolving real-world situations that can be grappled with from a legal disruption perspective, this is also the stage where unintuitive and non-linear impacts occur which are often overlooked.

AI is thus creating new affordances that parlay into new possibilities for behaviour, either by complementing and assisting human behaviours, or by automating and displacing them. This creates new social problems because our current legal and regulatory frameworks depend centrally on human agents as the bearers of rights and duties. We discuss the dynamics of this in more detail below when discussing the second stage of our framework.

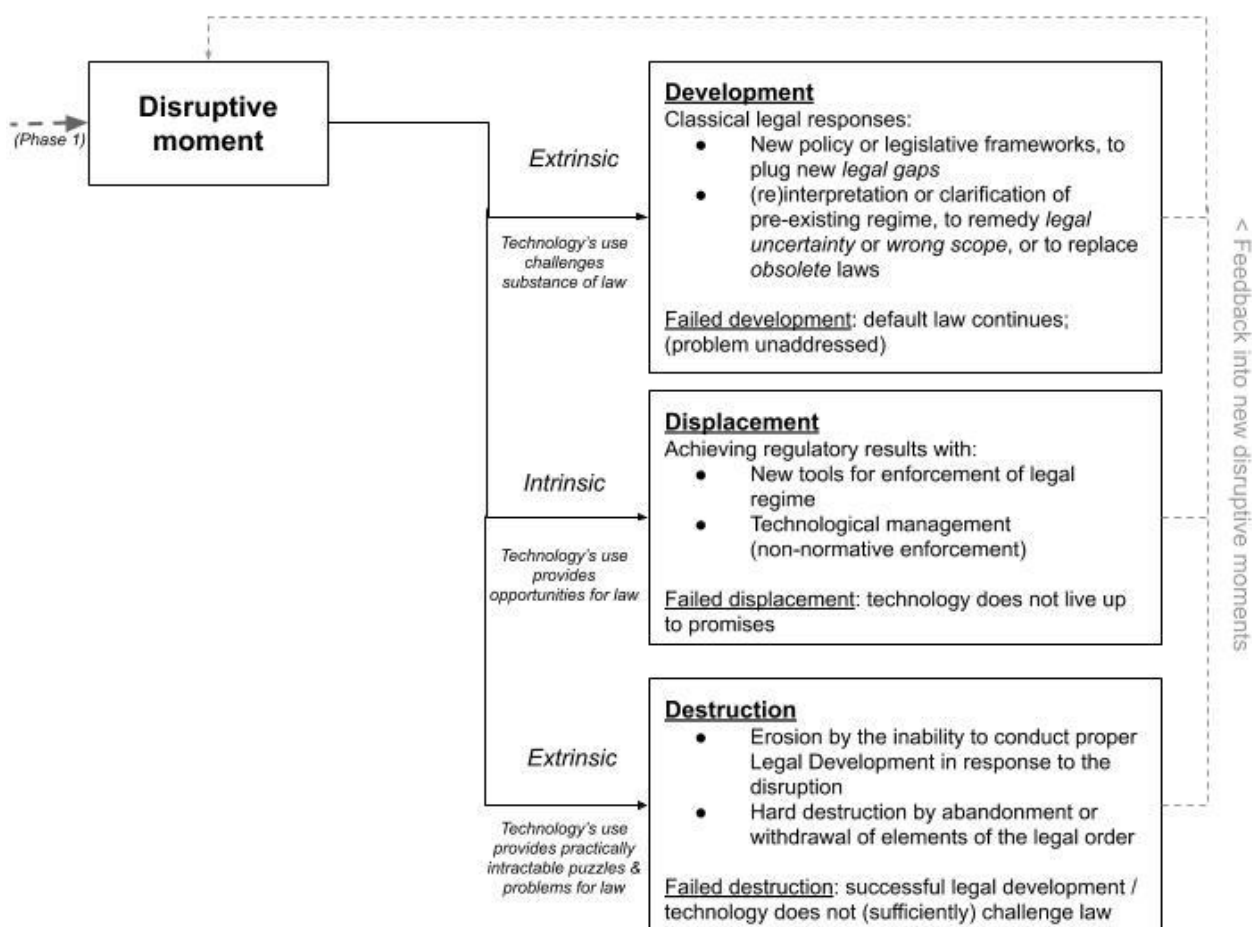
Reflections on Disruptive Moments

In addition to the above, there is a distinction to be drawn between disruption that is *intrinsic* or *extrinsic* to the legal system. Intrinsic disruption arises when the legal system itself encourages the use of a disruptive technology – in some cases to strengthen the enforcement or functioning of law itself (see the discussion below on Displacement) – in a way that creates problems for the legal system down the line. For example, the widespread use of facial recognition technology may initially be motivated by a desire to better enforce the existing criminal law, but may then disrupt or undermine the existing law due to its second order effects on both human behaviour and other legal norms. Extrinsic disruption arises when the technology is used for reasons independent of the law and the law must react to its social and legal consequences. For example, the desire to develop and use autonomous vehicles is not primarily motivated by the desire to better comply with the highway code, but it has consequences for that code (and other areas of law) and so the legal system must adapt and respond to its development.

Two further questions are worth asking about our proposed model of the disruptive moment: (i) what makes it *disruptive*? and (ii) is it really a *moment*? The answer to both questions is somewhat similar. The model idealises and abstracts from the real world, as all good models do. The underlying reality maybe fuzzier and more complex than the model presumes. Thus, what makes a moment disruptive is never going to be precisely definable. It is always going to be partly dependent on both the context and the people affected. In broad outline, we can say a moment is ‘disruptive’ when the existing sets of applicable legal norms no longer seem to work (i.e. there are gaps or omissions or anachronisms that are cast into the light by the new behaviours enabled by the affordances of the new technologies). Relatedly, there is unlikely to be a single sharp ‘moment’ of disruption. There may be a single moment in time when people realise that the existing legal order is no longer fit for purpose, for example this may have happened historically with the advent of nuclear weapons. But there may also be an ongoing and iterative process of disruption that only become obvious once a certain critical mass of disequilibrium has been reached. The model, as we envisage it, embraces both possibilities.

Stage Two: Responding to a Disruptive Moment: Development, Displacement or Destruction

The second stage of the model focuses on regulatory responses or reactions to the initial disruptive moment. As regulators are made aware of (or begin to perceive) the new affordances created or revealed by a given technology, how could they react? How do they react? And what effects can these particular reactions have on specific laws, or on legal processes, or on the legal system as a whole? As illustrated in Figure 1, legal responses to a disruptive moment can follow three general pathways: ‘Development’; ‘Displacement’; and ‘Destruction’.¹⁰⁷ While these pathways are illustrated as separate and mutually-exclusive trajectories for the sake of clarity in setting out ideal-type categories, there is little that prevents fragmented responses to a particular disruptive moment to follow some combination of these pathways.¹⁰⁸



¹⁰⁷ This follows and expands further upon a set of categories discussed in Maas, ‘International Law Does Not Compute’ (n 76).

¹⁰⁸ While this is elaborated upon in greater detail below, to provide just one example here for clarity: drone technology may simultaneously: create the need for legal *Development* to clarify questions of liability, privacy, and no-fly zones; create opportunities for legal *Displacement* by strengthening the monitoring and surveillance abilities of policing departments; create opportunities for legal *Destruction* by improving the ability of e.g. gangs to smuggle drugs past traditional avenues. How these three trends interact--and the degree to which these developments may amplify or cancel out one another--is unclear and complex.

Figure 1. Responses to the Disruptive Moment

Turning to the three primary channels of the framework, a relatively sharp divide can be drawn between Legal Development on the one hand, and Displacement or Destruction on the other. This division pivots upon whether the disruptive moment signals a recommitment to the dominant regulatory paradigm, or whether it precipitates a significant departure from the status *ex ante*, leading to reliance upon other modalities of regulation (Displacement), towards other goals and objectives than those sought through the dominant regulatory paradigm, or even towards a decline or rupture of (components of) the extant legal framework (Destruction). In other words, Development seeks to retain, through rebalance, the core features and aims of the contemporary legal order while both Displacement and Destruction seek to achieve regulatory objectives through different means and/or pursue different purposes than those sought by law and regulation at present. Let's now consider this in more detail.¹⁰⁹

(a) Legal Development

Taking Legal Development first, this category of response implies legal actions to address the disruptive moment from within the orthodox legal paradigm. Development therefore is an incremental response that pushes legal principles and processes towards accommodating, and thereby diffusing, the disruptive moment. For instance, Matthijs Maas has recently discussed ways in which international law might, in principle, be capable of Legal Development to accommodate many types of sociotechnical legal change produced by AI.¹¹⁰ Drawing on the taxonomy developed by Lyria Bennett Moses,¹¹¹ and focusing on the context of international law, Maas discusses four types of changes wrought by AI which might spur the need for Legal Development. Specifically, AI might create new legal gaps that require entirely new rules; it might lead to legal uncertainty; it might lead to existing laws now having a wrong scope of application; or it might lead to obsolescence of the existing legal provisions for a range of reasons.¹¹² Finally, the legal system might fail to reckon with the disruptive moment, resulting in a nominal continuation of the 'default' state of the law, with the risk that underlying issues continue to go unaddressed.

¹⁰⁹ Throughout the following sections, we will draw on a host of examples, both actual and speculative, to illustrate the argument. These are not meant to be exhaustive, and indeed many other cases or situations could be discussed.

¹¹⁰ Maas, 'International Law Does Not Compute' (n 76).

¹¹¹ Bennett Moses (n 14) 4–5.

¹¹² Maas, 'International Law Does Not Compute' (n 76) 39–55. Compare also however the excellent analysis, by Rebecca Crootof, of the ways in which new technology (specifically new weapons in the context of international law) fosters legal disruption. Rebecca Crootof, 'Regulating New Weapons Technology' in Eric Talbot Jensen and Ronald TP Alcalá (eds), *The Impact of Emerging Technologies on the Law of Armed Conflict* (2019) <<https://papers.ssrn.com/abstract=3195980>> accessed 15 March 2019. She buckets such legal disruption into five categories: 'Significantly Changing How Law is Created or Used', 'Highlighting Existing Legal Ambiguity'; 'Introducing New Uncertainty'; 'Undermining Foundational Assumptions'; 'Overlapping Categories'. Note, the latter four of these categories map loosely onto the distinct categories in our analysis of Legal Development, whereas 'Significantly Changing How Law is Created or Used' maps to our discussion of Legal Displacement. We argue our typology also subsumes (or expands on) previous frameworks, such as that offered by Friedman (n 14) 71. (arguing that technological change affects law "(1) by altering the cost of violating and enforcing existing legal rules; (2) by altering the underlying facts that justify legal rules; and (3) by changing the underlying facts implicitly assumed by the law, making existing legal concepts and categories obsolete, even meaningless.").

i. New legal gaps

In the first place, new AI applications may simply create the need for *new sui generis* rules to deal with the legal lacunae opened up (or revealed) by the technology's use. This may be expressed in the need for new regulation to deal with completely novel situations or forms of conduct enabled by this technology, or entirely new classes of technology, which do not seem to fit within any existing legal regime. For instance, AI could enable new, ethically challenging or politically or strategically destabilizing forms of international conduct.¹¹³ To be sure, in many such cases, such as with the deployment of fully autonomous weapons,¹¹⁴ these developments might still be seen as capable of plausible accommodation within existing frameworks, such as the modular Additional Protocols of the Convention on Certain Conventional Weapons.¹¹⁵ In other cases, however, the situation is sufficiently distinct that they might create the need for entirely new treaty regimes, institutions, or other global governance instruments.¹¹⁶ This might be the case, for instance, for facial recognition systems that could enable widespread global surveillance,¹¹⁷ or military AI analysis systems capable of tracking rival nuclear assets (submarines and mobile launchers) in ways that threaten the stability of nuclear deterrence.¹¹⁸ Of course, the question as to whether such envisioned Legal Development is subsequently politically viable is an entirely distinct one.¹¹⁹ The key situation here, however, is that the existing legal instruments have undergone a legally disruptive moment, recognizing a new technology (or its specific usage) to clearly fall out of scope of existing instruments.

ii. Legal uncertainty

In the second place, AI applications may create *legal uncertainty* over how, or even whether, existing laws apply to certain new affordances – forms of conduct, types of entities, or possible relationships – which are created or enabled by their use. As Bennett Moses has argued, technology can spark a need for legal rules when there is uncertainty over how to legally classify the new entities, behaviour or relationships it enables. This is because: there may be no adequate classification rubric that exists; because the new behaviour fits into more than one existing category

¹¹³ cf. also Dafoe (n 13).

¹¹⁴ cf. Nehal and others (n 100). However, for a critique of the focus on 'autonomy' in many of these debates as currently conceived, see Hin-Yan Liu, Léonard Van Rompaey and Matthijs M Maas, 'Editorial: Beyond Killer Robots: Networked Artificial Intelligence Systems Disrupting the Battlefield?' (2019) 10 *Journal of International Humanitarian Legal Studies* 77; Liu, 'From the Autonomy Framework towards Networks and Systems Approaches for "Autonomous" Weapons Systems' (n 53); Léonard Van Rompaey, 'Shifting from Autonomous Weapons to Military Networks' (2019) 10 *Journal of International Humanitarian Legal Studies* 111.

¹¹⁵ Maas, 'Innovation-Proof Governance for Military AI?' (n 75).

¹¹⁶ However, for the contrary argument that many challenges created by AI might well be accommodated within existing regimes of international law, see also Kunz and Ó hÉigeartaigh (n 55).

¹¹⁷ On arguments that favour a ban on such systems, see also Luke Stark, 'Facial Recognition Is the Plutonium of AI' (2019) 25 *XRDS* 50; Woodrow Hartzog, 'Facial Recognition Is the Perfect Tool for Oppression' (*Medium*, 2 August 2018) <<https://medium.com/s/story/facial-recognition-is-the-perfect-tool-for-oppression-bc2a08f0fe66>> accessed 23 April 2019.

¹¹⁸ cf. Michael C Horowitz, 'When Speed Kills: Lethal Autonomous Weapon Systems, Deterrence and Stability' (2019) 42 *Journal of Strategic Studies* 764; Michael C Horowitz, Paul Scharre and Alexander Velez-Green, 'A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence' [2019] arXiv:1912.05291 [cs] <<http://arxiv.org/abs/1912.05291>> accessed 18 December 2019; Avin and Amadae (n 105); Edward Geist and Andrew J Lohn, 'How Might Artificial Intelligence Affect the Risk of Nuclear War?' (RAND 2018) <<https://www.rand.org/pubs/perspectives/PE296.html>>; Sharikov (n 105).

¹¹⁹ Critically, if it is not, this situation--where a key problem remains unaddressed because of political gridlock or failure--can result in 'soft legal destruction', as discussed below, under section (c.ii).

and becomes subject to conflicting rules; or because an existing legal category is blurred.¹²⁰ AI systems also combine a number of features that may actively blur existing legal categories.¹²¹ For instance, it has been suggested that some AI systems, given their autonomy and unpredictability, might effectively occupy a liminal position between agent and object,¹²² blurring key categorizations that are at the heart of settling questions of responsibility and liability around AI. When a given disruptive moment manifests in legal uncertainty of this type, it usually creates the need for the clarification or sharpening of existing rules and their definitions and conditions, whether by legislative or judicial action. Once again, the question of whether there is sufficient conceptual flexibility that the previous concept can be recovered, or political willingness to carry out such Legal Development, is a distinct question.

iii. Wrong scope

In the third place, new technologies such as certain AI applications can create problems regarding the *scope* of existing laws, which are suddenly held to be over- or under-inclusive of the new affordances or applications associated with the AI system. One example of over-inclusivity might be found in the prospective case for achieving functional ‘legal personhood’ for certain algorithms, by exploiting loopholes or lacunae in existing regulation. Shawn Bayern and others have argued that loopholes in existing US company law allows for the incorporation of a limited liability company, whose operating agreement places it under the control of an AI system.¹²³ Thomas Burri has argued that, if such an entity were established in the EU, the internal market principle of the mutual recognition of national legal personality would imply that all other EU member states would be forced to recognize the personhood of such an entity.¹²⁴ Provided they were upheld in court,¹²⁵ such ploys would reveal inadequacies or loopholes in the existing state of the law – gaps which, if left unplugged, could be exploited or misused by certain actors.¹²⁶ Conversely, new AI systems could also reveal the ‘under-inclusivity’ of the law, for instance: when existing privacy regulation is not held to extend to the production of incriminating DeepFake-generated media; when abuse aimed at very human-like social robots provokes widespread societal outrage but is not covered under anti-

¹²⁰ Bennett Moses (n 14).

¹²¹ cf. Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Springer Berlin Heidelberg 2018) 64–80.

¹²² Or, in the specific case of military AI, may blur the distinction, in International Humanitarian Law, between ‘weapon’ and ‘agent’. cf. Hin-Yan Liu, ‘Categorization and Legality of Autonomous and Remote Weapons Systems’ (2012) 94 *International Review of the Red Cross* 627.

¹²³ Shawn Bayern, ‘The Implications of Modern Business–Entity Law for the Regulation of Autonomous Systems’ (2016) 7 *European Journal of Risk Regulation* 297; Shawn Bayern and others, ‘Company Law and Autonomous Systems: A Blueprint for Lawyers, Entrepreneurs, and Regulators’ (2017) 9 *Hastings Science and Technology Law Journal* 135.

¹²⁴ Thomas Burri, ‘Free Movement of Algorithms: Artificially Intelligent Persons Conquer the European Union’s Internal Market’ in Woodrow Barfield and Ugo Pagallo (eds), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar 2017) <<https://papers.ssrn.com/abstract=3010233>> accessed 13 September 2018.

¹²⁵ A step of which some are sceptical; see for instance Matthew U Scherer, ‘Of Wild Beasts and Digital Analogues: The Legal Status of Autonomous Systems’ (Social Science Research Network 2018) SSRN Scholarly Paper ID 3223174 <<https://papers.ssrn.com/abstract=3223174>> accessed 19 November 2018.

¹²⁶ For a discussion of some of the challenges and criminal opportunities which such an ‘algorithmic entities’ construction might introduce, see Lynn M LoPucki, ‘Algorithmic Entities’ [2017] UCLA School of Law, Law-Econ Research Paper <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2954173> accessed 19 May 2017.

cruelty laws;¹²⁷ or if new sex robots that simulate non-consensual sex are introduced, and likewise held to evade existing standards.¹²⁸ When technologies create a perceived problem of the scope of existing legislation, this creates a need for Legal Development, to extrapolate and reaffirm existing lines between legal categories so as to explicitly include or exclude the new behaviour, entities or relationships.

iv. Obsolescence

Fourthly, AI applications can lead to functional *legal obsolescence*, where existing law is rendered unfit for its originally intended purposes. This can occur for one of three subsidiary reasons: because the technology (a) renders once regulated behaviour obsolete; (b) undercuts certain justifying assumptions; (c) makes existing laws no longer cost-effective to enforce.

In the first case, obsolescence might result because (a) formerly common behaviour that was subject to regulation has been superseded or rendered obsolete in practice as a result of new technology. To be sure, it is not necessarily the case that such laws are entirely without effect. After all, the precedent they set out might on occasion still provide the key legal metaphors that shape subsequent legal interpretations,¹²⁹ especially in common law contexts. Nonetheless, these laws are (almost) never invoked anymore because the entity or behaviour that was the object of regulation has become so rare. Are there clear cases of such legal obsolescence? One might consider the status of past ‘dead letter’ laws, such as rules on the management of telegraph infrastructure.¹³⁰ More speculatively, in the context of international law, one might consider an extreme extrapolation of the current military trend towards ‘remote warfare’; if AI-steered combat platforms come to widely replace human soldiers on the battlespace, this might marginalise or render functionally moot certain principles of IHL dictating the treatment of prisoners of war.¹³¹ On reflection, this category of legal obsolescence may be relatively rare because, while technological progress creates new affordances, it less often results in the wholesale disappearance of certain types of behaviour.¹³² In the second place, one can debate to what extent this first type of legal obsolescence truly poses a problem for the legal system. While in some cases, antiquated and virtually uninvoked ‘dead letter’

¹²⁷ cf. Kate Darling, ‘Extending Legal Rights to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects’ [2012] SSRN Electronic Journal <http://robots.law.miami.edu/wp-content/uploads/2012/04/Darling_Extending-Legal-Rights-to-Social-Robots-v2.pdf> accessed 7 January 2019.

¹²⁸ See generally John Danaher, ‘Robotic Rape and Robotic Child Sexual Abuse: Should They Be Criminalised?’ (2017) 11 Criminal Law and Philosophy 71. And for proposal and discussion around a presumed ‘consent module’, see Anco Peeters and Pim Haselager, ‘Designing Virtuous Sex Robots’ [2019] International Journal of Social Robotics <<https://www.readcube.com/articles/10.1007/s12369-019-00592-1>> accessed 8 October 2019.

¹²⁹ For a discussion of this in the context of AI, see Ryan Calo, ‘Robots as Legal Metaphors’ (2016) 30 Harvard Journal of Law and Technology 209; See generally Gregory N Mandel, ‘Legal Evolution in Response to Technological Change’ [2017] The Oxford Handbook of Law, Regulation and Technology 233–234 <<http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199680832.001.0001/oxfordhb-9780199680832-e-45>> accessed 26 September 2018.

¹³⁰ On the other hand, at least within the common law system, cases dealing with telegraphs or other outmoded forms of communication may nonetheless have continued relevance. In such cases, these laws might be only partially obsolescent.

¹³¹ cf. Maas, ‘International Law Does Not Compute’ (n 76) 42.

¹³² Of course, it is also an open question whether such ‘dead letter’ obsolescence truly requires categorical cessation or obsolescence of certain behaviour, or whether ‘virtual’ disappearance suffices.

laws do rear their head,¹³³ in most cases the legal disruption imposed on a legal system by having ‘superfluous but uninvoked’ laws seems problematic, if at all, only from a legal coherentist perspective.¹³⁴ This type of legal obsolescence, however, can still pose problems for the law (even from a regulatory instrumentalist perspective) if obsolete statutes provide (misleading) judicial metaphors; or if they grow into ‘jurisprudential space junk’ that muddies the waters.¹³⁵

Secondly, and more problematically for the law, legal obsolescence can also occur because (b) one or more basic *justifying assumptions* that underlie the original historical introduction and specific formulation of a particular law are no longer valid. For example, one might argue that the (admittedly aspirational) human right to work, as enshrined in the Universal Declaration of Human Rights and the International Covenant on Economic, Social and Cultural Rights, is premised on the assumption that societies will continue to need (and be capable of providing) employment for a large fraction of their population. As AI systems continue to outstrip human performance in more domains,¹³⁶ and as a large population is rendered structurally unemployable, this change may render this right to work functionally obsolete, with repercussions to the large parts of the global legal regime constructed by the International Labour Organization,¹³⁷ creating the need for legal development that assures and secures new, different notions of human meaning and productivity in a post-work era.¹³⁸

Finally, legal obsolescence (and the accordant need for legal development) can occur (c) because a law is no longer *cost-effective to enforce*. For instance, scholars have raised the possibility that ‘DeepFakes’ might adversely affect the epistemological foundations of both domestic courts and also international human rights investigations.¹³⁹ Practical difficulties introduced by new technologies can often slow the enforcement of older legal frameworks to new spaces: the difficulty of attributing attacks in cyberspace has been held as one (if not the only) hurdle to effectively regulating cyberattacks.¹⁴⁰ In some cases, Legal Development can seek to address this variant of obsolescence by simply seeking to restrict the innovation that undercuts the cost-effectiveness of enforcement.¹⁴¹ In other cases, however, the technology has many additional legitimate uses, or has already been sufficiently integrated in the economic and social fabric, that banning it wholesale,

¹³³ cf. in the context of blasphemous libel, Jeremy Patrick, ‘Not Dead, Just Sleeping: Canada’s Prohibition on Blasphemous Libel as a Case Study in Obsolete Legislation’ (2008) 41 U.B.C. Law Review 193.

¹³⁴ cf. the discussion in Graham McBain, ‘Abolishing Some Obsolete Common Law Crimes’ (2009) 20 King’s Law Journal 89.

¹³⁵ Crootof, ‘Jurisprudential Space Junk’ (n 76).

¹³⁶ Katja Grace and others, ‘When Will AI Exceed Human Performance? Evidence from AI Experts’ (2018) 62 Journal of Artificial Intelligence Research 729.

¹³⁷ Maas, ‘International Law Does Not Compute’ (n 76) 43.

¹³⁸ See for example John Danaher, *Automation and Utopia: Human Flourishing in a World without Work* (Harvard University Press 2019).

¹³⁹ Steven Livingston and Mathias Risse, ‘The Future Impact of Artificial Intelligence on Humans and Human Rights’ (2019) 33 Ethics & International Affairs 141, 144.

¹⁴⁰ Michael J Glennon, ‘The Dark Future of International Cybersecurity Regulation’ (2013) 6 Journal of National Security Law & Policy 563.

¹⁴¹ Consider the N.Y.P.D.’s 2019 appeal to Google, demanding that the company remove the locations of police drunk-driving checkpoints from its Waze navigation app. Michael Gold, ‘Google and Waze Must Stop Sharing Drunken-Driving Checkpoints, New York Police Demand’ *The New York Times* (6 February 2019) <<https://www.nytimes.com/2019/02/06/nyregion/waze-nypd-location.html>> accessed 4 December 2019.

simply because certain uses of it undercut the enforcement of certain laws, is not tenable.¹⁴² In such cases, regulation might seek not to restrict, but rather to channel the processes of innovation or of its deployment, to ensure that they proceed along pathways that ensure the innovation is still susceptible to cost-effective regulation and enforcement.¹⁴³ In the cases of AI in particular, where distinct techniques have such immensely broad applications, we may therefore expect that legal development in response to obsolescence will not often take the form of reactive bans or restrictions to the circumventing technology (effectively an attempt to reaffirm the status ex ante). Rather, Legal Development strategies will pursue the implementation of new, substitute vectors of regulation: in some cases in the form of new, differently specified regulation; in other cases ‘coded in’ directly through the use of new technology.¹⁴⁴

In summary, where a technology drives legal obsolescence, the disruptive moment can spark a drive towards Legal Development, as it creates a realisation (or perception) amongst lawmakers that (a) certain existing laws are now ‘dead letter laws’ which can be struck from the books without much consequence; (b) that the law is no longer serving its original, intended purpose because one of its key assumptions having changed such that it may need replacement by another law; or (c) certain laws need a reformulation, in a way that is more cost-effective (or possible) to enforce.

While the above discussion of Legal Development – to respond to the creation of new legal gaps, legal uncertainty, wrong scope, or legal obsolescence – draws on examples that are specific to AI, this tracks the ways other technologies have challenged law, and shows how extant tools and sources of law might accommodate these disruptive changes in principle. For instance: legislative action or judicial review can both develop existing law or create new bodies of law (if not always in time) to cover the completely new applications; clarify the status or definition of concepts, jurisdiction, or the applicability of laws to a technology; close loopholes; repeal now-unnecessary laws, or substitute new regulation for now-unjustifiable or unenforceable laws.

v. Default and Other Feedback Dynamics

It is important to recognize, however, that even if the use of a new technology creates a ‘disruptive moment’, this does not mean that Legal Development is bound to follow, let alone succeed. In the ‘default’ scenario, the impetus towards Legal Development fails and the legal system remains in its

¹⁴² In a sense, this could be understood as a version of the well-established ‘Collingridge Dilemma’: “The root of the manifest difficulties with which the control of technology are beset is that our technical competence vastly exceeds our understanding of the social effects which follow from its exercise. For this reason, the social consequences of a technology cannot be predicted early in the life of a technology. By the time undesirable consequences are discovered, however, the technology is often so much part of the whole economic and social fabric that its control is extremely difficult.” David Collingridge, *The Social Control of Technology* (Palgrave Macmillan 1981) 11.

¹⁴³ For instance, Lessig notes how, when the transition to digital telephone networks rendered traditional governmental wiretapping difficult (i.e. less cost-effective), the US government responded by passing regulation that simply required telephone companies to select network architectures that facilitated wiretapping. Lessig, ‘The Law of the Horse’ (n 37) 531. Likewise, when faced with computer piracy of intellectual property, the response was not to ban internet communications (or computers) entirely, but instead to harness the possibilities of the new technology itself--the architectural possibilities of ‘code’--to constitute new regulatory frameworks that rendered such crimes ‘impossible’. *ibid* 523–528.

¹⁴⁴ Compare, for instance, Youtube’s Content ID algorithm, which automatically processes copyright control cases on the video sharing platform. Leron Solomon, ‘Fair Users or Content Abusers: The Automatic Flagging of Non-Infringing Videos by Content ID on Youtube Note’ (2015) 44 Hofstra Law Review [i]. See also the discussion on Displacement, in the next section.

pre-existing status quo. The continuation of the default state can come about for a range of reasons including neglect, lobbying, or deadlock. This needs to be distinguished from the persistence of a default state where a disruptive moment has been erroneously identified. In other words, the default state can also persist in situations where there is actually no legally disruptive moment arising from a given application of a new technology. Lyria Bennett Moses, for instance, once argued that “traffic rules continue to apply to cars with electric windows, and no sane person would seek to challenge these laws as inapplicable or write an article calling for the law to be clarified.”¹⁴⁵ In practice many technological developments that in retrospect turn out modest, may initially be treated as comprehensively disruptive.¹⁴⁶

While we discuss this below, it should be noted that a regression to ‘default’ can also occur in the aftermath of Legal Displacement,¹⁴⁷ whereby decisions to enforce extant normative law shift the problem-portfolio *beyond the realm of law*, thereby leaving the principles and processes at the legal level unchanged. If we are dealing with a genuinely disruptive technology (as opposed to one that is merely anticipated or imagined) the likelihood is that the continuation of the default regulatory response will increase the tensions between the system and reality, triggering subsequent disruptive moments, or feeding directly into soft legal destruction (erosion),¹⁴⁸ as elements of the law are revealed as inadequate at responding to the technology’s challenge. This suggests that regression to default is neither a likely resting nor terminal state for the legal disruption framework, when the disruption does not exist purely in the imagination.

Furthermore, there are clear feedback cycles of development in relation to the disruptive moment. Development is clearly an important response to the disruptive moment. Yet, developmental responses may become divergent and subsequently fragment in the sense that domain-specific developments proceed along trajectories that are not coordinated or harmonised.¹⁴⁹ The coherence of the legal system then becomes diluted, undermining the robustness and resilience of the legal order to confront subsequent disruptive moments. Furthermore, the resulting instability may then channel responses to later disruptive moments towards Displacement and Destruction, decreasing the relevance of the law itself. Development may also have more directly counterproductive effects that feed back into disruption because attempts to accommodate the disruptive moment may open up latent inconsistencies in legal doctrine. Common questions of responsibility and liability relating to AI agents arising from their liminal position between agent and object, for example, could boomerang and force re-evaluations of what these legal concepts entail in relation to human beings. Insofar as foundational legal doctrines are excavated, this can render the legal order more susceptible to other disruptive moments arising from other sectors.

Finally, because Legal Development as a response does not seek to radically alter or reconceive the legal order, its focus on the level of the law directs efforts towards reinforcing the legal order. While this introspective orientation may be necessary, it may detract from simultaneous Displacement and Destruction processes at play. In other words, while the legal order is being shored up against the shock of a disruptive moment, both its reach and relevance have diminished

¹⁴⁵ Bennett Moses (n 50) 596.

¹⁴⁶ These points will be explored more in section 3.4 on ‘A User-Guide to the Legal Disruption Model’.

¹⁴⁷ As discussed in more detail in the next section.

¹⁴⁸ As discussed in section 2(c)(i), on ‘erosion’.

¹⁴⁹ As discussed above, under section 1.

because the disruptive moment has set into motion responses that displace the law and initiated steps to depart from the rationale underpinning the law. The warning here is not only that it is quite possible to win the legal battle, but lose the regulatory war: it is also possible to appear victorious without realising the Pyrrhic nature of that victory.

(b) Legal Displacement

In the Legal Displacement response category, the problem encapsulated in the disruptive moment is not that new technological innovations create affordances that challenge the substance of existing law (as under Legal Development), but rather that they create new affordances *for* the regulators producing and enforcing the law. Accordingly, while it may appear as a less orthodox case, legal displacement too can be captured under the three-stage model of legal disruption as a result of new affordances. That is, new technologies (especially AI) have a plethora of applications in policing, administration, judicial decision-making, and the business of government writ large. In doing so they may open up the possibility of achieving certain regulatory results without relying as much on the traditional institutional machinery of the law.

The introduction of new technology challenges existing *practices* of law because it creates new procedural or instrumental opportunities for legal actors.¹⁵⁰ This is often, but not always, the result of an intrinsic disruptive moment and may initially complement, but gradually ‘displace’ (rather than substantively challenge) certain existing laws in a functional manner. There are two subcategories of Legal Displacement by AI. First, *legal automation* draws on the use of AI in the production, adjudication or enforcement of ‘normative’ law. Secondly, *legal replacement* involves the use of AI to shift the configuration of regulatory modalities in which a legal system is embedded. This enables the progressive ‘nudging’ or even full-scale non-normative ‘technological management’ of regulatee behaviour.¹⁵¹

i. Automation

AI systems can be used to automate a range of routine legal tasks. Drawing on adequate databases, machine learning systems can even learn to predict the outcomes of legal disputes, such as judicial decisions of the European Court of Human Rights.¹⁵² There are many extensive and valid critiques of the prospects for legal automation given the limits of the current generation of machine learning approaches,¹⁵³ as well as critiques identifying the particular pitfalls of overtrust and interface design

¹⁵⁰ Which, of course, can in turn lead to new disruptive moments that result in Legal Development--as in the case of the famous *State v. Loomis* case, where the use of the COMPAS algorithm in assessing risk of recidivism, sparked debate over the admissibility of judges relying on such opaque systems. ‘*State v. Loomis: Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing.*’ (2017) 130 *Harvard Law Review* 1530.

¹⁵¹ This section draws on, and extends the discussion in Maas, ‘*International Law Does Not Compute*’ (n 76) 44–50.

¹⁵² Nikolaos Aletras and others, ‘*Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*’ (2016) 2 *PeerJ Computer Science* e93; Masha Medvedeva, Michel Vols and Martijn Wieling, ‘*Judicial Decisions of the European Court of Human Rights: Looking Into the Crystal Ball*’ 24.

¹⁵³ See for instance: Frank Pasquale and Glyn Cashwell, ‘*Four Futures of Legal Automation*’ (2015) 26 *UCLA Law Review Discourse* 23; Frank A Pasquale, ‘*A Rule of Persons, Not Machines: The Limits of Legal Automation*’ (2019) 87 *George Washington Law Review* 1; Mireille Hildebrandt, ‘*Law As Computation in the Era of Artificial Legal Intelligence. Speaking Law to the Power of Statistics*’ (Social Science Research Network 2017) SSRN Scholarly Paper ID 2983045 <<https://papers.ssrn.com/abstract=2983045>> accessed 9 July 2018.

errors that plague hybrid human and AI ‘cyborg justice’ arrangements.¹⁵⁴ Others have cautioned that the automation of law enforcement systems might close the loop on inefficiency and indeterminacy, which are important safeguards against the perfect enforcement of laws that were drafted with an implicit assumption of lenience.¹⁵⁵ Nonetheless, barring a major backlash,¹⁵⁶ it seems likely that many of these technologies will see continued development and deployment, supporting a general shift towards ‘technocratic’ regulatory attitudes.¹⁵⁷

The use of technologies in support of existing legal structures and processes, offers the first face of a (intrinsic) disruptive moment resulting in Legal Displacement. This could take the form of the automated monitoring of citizen compliance with existing legal norms. For example, a police force could use automated speed cameras, facial recognition technology and automated identity-tracking to locate people who have breached road traffic laws. Punishment could also be automated through something akin to the social credit system, such as the set of systems that have reportedly been undergoing trials in China.¹⁵⁸ As technology advances, Legal Displacement by AI could even extend beyond the prediction or adjudication of legal questions, moving instead towards the production of new normative laws (which can then be automatically monitored and enforced). This would be the case in the hypothetical case of a ‘legal singularity’, which envisions a system that can predict the outcome of legal cases based on their details. Such a system would be capable of tailoring ‘micro-directives’ to each situation and each regulatee, collapsing the traditional distinction between legal rules and standards.¹⁵⁹ This could have far-reaching implications for how we conceptualize the essential features of legality.¹⁶⁰ Even such legal automation, however, still

¹⁵⁴ Rebecca Crotoft, ‘“Cyborg Justice” and the Risk of Technological-Legal Lock-In’ [2019] Columbia Law Review Forum <<https://papers.ssrn.com/abstract=3464724>> accessed 18 November 2019.

¹⁵⁵ Woodrow Hartzog and others, ‘Inefficiently Automated Law Enforcement’ (2016) 2015 Michigan State Law Review 1763.

¹⁵⁶ A real possibility, to be sure. Recent years have seen a ‘techlash’ to the conduct of digital technology companies in a range of sectors. cf. Mark Scott, ‘In 2019, the “Techlash” Will Go from Strength to Strength’ (*POLITICO*, 30 December 2018) <<https://www.politico.eu/article/tech-predictions-2019-facebook-techlash-europe-united-states-data-misinformation-fake-news/>> accessed 9 December 2019. Likewise, cities such as San Francisco have begun to ban sidewalk delivery robots (2017) and facial recognition (2019); see Matt Simon, ‘San Francisco Just Put the Brakes on Delivery Robots’ [2017] *Wired* <<https://www.wired.com/story/san-francisco-just-put-the-brakes-on-delivery-robots/>> accessed 9 December 2019; Kate Conger, Richard Fausset and Serge F Kovaleski, ‘San Francisco Bans Facial Recognition Technology’ *The New York Times* (14 May 2019) <<https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html>> accessed 28 June 2019. Some have even suggested that, if mishandled, AI technology may experience a similar backlash and subsequent strict regulatory response as GMO crops in the EU. Cf. Andrew Grotto, ‘Genetically Modified Organisms: A Precautionary Tale For AI Governance’ [2019] *AI Pulse* <<https://aipulse.org/genetically-modified-organisms-a-precautionary-tale-for-ai-governance-2/>> accessed 26 February 2019.

¹⁵⁷ cf. Brownsword, ‘Law Disrupted, Law Re-Imagined, Law Re-Invented’ (n 39).

¹⁵⁸ Rogier Creemers, ‘China’s Social Credit System: An Evolving Practice of Control’ (Social Science Research Network 2018) SSRN Scholarly Paper ID 3175792 <<https://papers.ssrn.com/abstract=3175792>> accessed 16 May 2018.

¹⁵⁹ Benjamin Alarie, ‘The Path of the Law: Towards Legal Singularity’ (2016) 66 *University of Toronto Law Journal* 443; Benjamin Alarie, Anthony Niblett and Albert H Yoon, ‘Law in the Future’ [2016] *University of Toronto Law Journal* <<https://www.utpjournals.press/doi/abs/10.3138/UTLJ.4005>> accessed 28 January 2019; Anthony J Casey and Anthony Niblett, ‘Self-Driving Laws’ (2016) 66 *University of Toronto Law Journal* 429; Anthony J Casey and Anthony Niblett, ‘The Death of Rules and Standards’ (2017) 92 *Indiana Law Journal* 1401.

¹⁶⁰ See the discussion by Brian Sheppard, on how machine learning systems in legislation and adjudication, and specifically the tradeoffs they may force between intelligibility and results, may lead us to reject concepts of legality that require critical officials (Hart), reason-based tests of legitimacy (Raz), or justification for coercion (Dworkin).

constitutes only an incomplete or soft form of Legal Displacement, since it involves the continued existence (and indeed, the expected cybernetic strengthening) of a legal-normative order.

ii. Replacement

More fundamentally, AI could facilitate or drive a shift towards novel ‘modes’ of achieving societally desired regulatory outcomes, which no longer rely on explicitly articulated normative laws at all. It is helpful to think of this in terms of Lessig’s regulatory modalities of law, social norms, markets, and architecture.¹⁶¹ Although lawyers may think it natural to assume the best response to AI-driven legal disruption is to develop and reform the legal system, this may not always be possible or desirable. But if the legal modality of regulation falls short or fails to pull its weight, the regulatory outcome derived from the sum total configuration (or ‘mix’) of regulatory modalities will be affected. If Legal Development is slow and unwieldy, and if the technology is truly disruptive, other regulatory modalities such as the architecture modality of regulation, may be called upon to take up the slack and address the challenge. This can happen intentionally, where there is a concerted effort to reconfigure the constellation of regulatory modalities to make up for the legal shortfall, but more problematically this can occur also through uncoordinated responses. While both scenarios represent legal displacement, those that arise through the latter path are less perceptible and more difficult to subsequently overturn.

In addition to this, AI systems themselves may open up new possible modes of regulation. Among Lessig’s regulatory modalities, AI taps most readily into the market and architecture or code alternatives (converging with technological management¹⁶²) to the law. Thus, prominent examples of Legal Displacement may involve regulators deploying market or architectural modalities where once they appealed to legal mechanisms to pursue their objectives. The resulting systems of AI-mediated ‘technological management’ (or less dramatically, nudge choice architectures) of human behaviour would be ‘non-normative’, since they no longer explicitly invoke or appeal to a legal norm with which the regulatee should align their behaviour (or face penalties). Rather, the use of such AI systems might simply present a technologically-shaped environment which renders certain choice options impossible – or at least imperceptible. This raises the spectre of AI systems being deployed in projects to ‘hypernudge’¹⁶³ citizens, potentially building towards an ‘algocracy’.¹⁶⁴

In sum, Legal Displacement is one possible outcome of legal disruption. Here the law is disrupted, not by extrinsic technology-enabled affordances that challenge its substance or doctrine (as under Legal Development). Rather, the technology in question creates new affordances for governments, hence the disruptive moment here is often intrinsic to the regulator. In some cases, this may nominally strengthen the efficacy of legal systems, but in others it may lead to the iterative replacement of those very legal systems. In return, the use of technology in the service of Legal Displacement often generates strong feedback effects that feed into new legal disruptive moments.

Brian Sheppard, ‘Warming up to Inscrutability: How Technology Could Challenge Our Concept of Law’ (2018) 68 *University of Toronto Law Journal* 36.

¹⁶¹ Lessig, ‘The New Chicago School’ (n 37); Lessig, ‘The Law of the Horse’ (n 37).

¹⁶² Roger Brownsword, ‘Technological Management and the Rule of Law’ (2016) 8 *Law, Innovation and Technology* 100.

¹⁶³ Karen Yeung, ‘“Hypernudge”: Big Data as a Mode of Regulation by Design’ (2017) 20 *Information, Communication & Society* 118.

¹⁶⁴ Danaher, ‘The Threat of Algocracy’ (n 98).

This happens once regulatees (and in some cases, other branches of government) become cognizant of, and challenge, the legality of these new tools. Just as the introduction of phone wiretapping, fingerprint analysis and DNA tests not only directly changed the processes of the law and its evidence requirements, but also led to indirect new Legal Developments,¹⁶⁵ the use of AI systems in facial recognition, drone surveillance platforms, predictive policing, and judicial decision-making are already generating new disruptive moments as societies and lawmakers begin to identify them as sites for the constitution of new Legal Developments.

(c) Legal Destruction

Legal Destruction constitutes the final pathway of the legal disruption model and is deeply tied up with the practical foundations and limits of legal systems. Under this category, the response to a disruptive moment ends in the sustained failure of regulatory initiatives. This can occur in the face of intractable practical or political hurdles involved in the regulation of the new technology, rendering legal development impossible or impracticable. Alternatively, Legal Destruction can take place as a result of a decline in the underlying conditions that support the legal order. Legal Destruction, insofar as it challenges the status quo of the existing legal system, is therefore most often the result of an extrinsic disruptive moment, but in residual instances can also arise from an intrinsic one with unanticipated results. It can be subdivided into soft and hard versions: ‘erosion’ and ‘decline’.

i. Erosion

Our earlier treatment of Legal Development covered different ways in which AI technology could give rise to new affordances (new entities, relationships, or especially behaviours) across different domains. This can result in disruptive moments resulting in regulators changing, clarifying or updating the legal system to adapt it to these challenges (Legal Development). This, however, raises a practical question: to what extent can such Legal Developments be carried out? What are the constraints on these initiatives? And what are the results if no avenue of legal development is sufficiently effective? Or politically viable?

This yields the first version of Legal Destruction: a ‘soft erosion’ of specific existing legal frameworks, because the particular affordances created by the use of AI have certain features, which render effective legal development practically or politically difficult if not impossible. They become ‘unsolvable puzzles’ for law by creating situations where the need for legal adaptation is clear – a disruptive moment is sparked – but regulators are unable to carry through (meaningfully effective) changes to legislation in response.

Why does this challenge occur? Is it a plausible risk? In principle, a legal system ought to be capable of being reconfigured to accommodate nearly any external challenge that can be meaningfully captured within human natural language. Given complete freedom to alter or redefine disrupted doctrinal categories, it seems that there would be few technology-driven legal challenges that could not, in principle, be resolved by Legal Development. For example, problems over the liminal status of AI systems between ‘agents’ and ‘objects’, or over their unpredictability, could be

¹⁶⁵ Mandel (n 129).

defined away by introducing wholly novel legal categories, principles or standards. Such a solution would be compatible with a regulatory-instrumentalist perspective, if not a legal-coherentist one.¹⁶⁶

In practice, however, matters are obviously more constrained. New legislation or case law that seeks to grapple with new technology is often to a large degree path-dependent: it is constrained by precedent; by its reliance on legal metaphors that invoke past cases;¹⁶⁷ and by the practical need to accord with a certain degree of established ‘folk understanding’.¹⁶⁸ These factors conceptually limit which avenues of Legal Development regulators are able, seek, or even consider, to pursue in response to a new disruptive moment.

In addition to these conceptual limits, however, certain technologies can prove recalcitrant to Legal Development for political reasons. For instance, they might create a range of affordances that empower some stakeholders much more than others. Or alternatively, their future trajectory and implications might be very unclear, inhibiting the regulator’s articulation of a disruptive moment and resultant effective regulatory action at an early stage, until the technology has already become firmly embedded at a later stage, and many parties have established stakes, and it is hard to dislodge.¹⁶⁹ The technology might map onto open political fault-lines; or simply be easily disseminated. As Lessig noted in the context of cyberspace, some designs of cyberspace made on-line behaviour more ‘regulatable’, whereas others made it less so.¹⁷⁰

While these problems can sometimes be manageable on the domestic level,¹⁷¹ regulatory resistance can become a particularly acute problem in the global arena, where some weapons technologies, it has been argued, are more ‘regulation-resistant’ than others.¹⁷² For instance, take cyber-weapons: in spite of calls for global regulation of cyberwar,¹⁷³ it has been argued that traditional arms control regimes are unlikely to be successfully transferred to the realm of cyberspace, because of the differences in how the ‘weapons’ in question are used (specifically the difficulty of attribution), and the relative difficulty of monitoring compliance.¹⁷⁴ In such cases, the reason that there are no global cyberwarfare conventions is not that no problem is perceived to exist (a lack of a disruptive

¹⁶⁶ That is, it would be incompatible with a legal-coherentist approach that emphasises the coherence of new legislation with extant legislation. Alternatively, one could consider a complete paradigm shift within the orientation and conceptual underpinnings of a legal system, that would functionally resolve all the conceptual unclarity as well as practical challenges posed by a new technology. This would be an extreme legal ‘phase transition’, in the systems-theory based framework explored by Ruhl: Ruhl (n 56); JB Ruhl, ‘Managing Systemic Risk in Legal Systems’ (2014) 89 *Indiana Law Journal* 559. In such a case, legal coherentists might accept the new legal equilibrium so long as it were self-consistent.

¹⁶⁷ See Calo, ‘Robots as Legal Metaphors’ (n 129); Mandel (n 129).

¹⁶⁸ Pierre Schlag, ‘Spam Jurisprudence, Air Law, and the Rank Anxiety of Nothing Happening (A Report on the State of the Art)’ [2009] *The Georgetown Law Journal* 803, 821.

¹⁶⁹ The so-called ‘Collingridge Dilemma’. Collingridge (n 142).

¹⁷⁰ Lessig, ‘The Law of the Horse’ (n 37) 533–534.

¹⁷¹ For instance, in Lessig’s example, the government could regulate the design of cyberspace in ways that facilitated its own ability to regulate the internet.

¹⁷² Sean Watts, ‘Regulation-Tolerant Weapons, Regulation-Resistant Weapons and the Law of War’ (2015) 91 *International Law Studies* 83.

¹⁷³ Mette Eilstrup-Sangiovanni, ‘Why the World Needs an International Cyberwar Convention’ (2018) 31 *Philosophy & Technology* 379.

¹⁷⁴ Erica D Borghard and Shawn W Lonergan, ‘Why Are There No Cyber Arms Control Agreements?’ (*Council on Foreign Relations*, 16 January 2018) <<https://www.cfr.org/blog/why-are-there-no-cyber-arms-control-agreements>> accessed 22 January 2018; Glennon (n 140).

moment),¹⁷⁵ or that the existing governance regimes are perceived to be entirely adequate and up to the task. Instead, it is either conceptual difficulties of classification, or more saliently the underlying political gridlock,¹⁷⁶ which prevents the resolution of this visible, technology-precipitated problem.¹⁷⁷

Even at a domestic level, legal erosion can occur in a staggered or multi-level process.¹⁷⁸ There may be cases where the emergence of a new technology (whether GMOs, file-sharing, reproductive cloning, for example), or certain high-profile visceral incidents spark public concern and drive rapid calls for regulators to impose bans. Regulators may then decide that a full-scale prohibition is indeed the appropriate course of action. If the prohibition—or the sanctions invoked to enforce it, or the indirect societal consequences of its enforcement—is eventually judged to be deeply unpopular, however, and there is widespread non-compliance amongst regulatees, then regulators may find themselves forced to redraw regulatory lines in a less strict or constraining way.¹⁷⁹ This would then kick-start a second phase of regulation, aimed not at full-scale prohibition but rather restricted access—with rules crafted in ways that reflect the most pressing concerns. This may eventually merge into a third phase of regulation, as restrictions become increasingly watered down, and affordances much easier to act on. The law becomes so permissive that it simply reflects and endorses prevailing public opinion. In such cases, the relevant legal frameworks are perhaps not visibly ‘eroded’; but that is only because they have yielded the contested ground, and have been left stretched so far as to be almost nominal.¹⁸⁰

Thus, while some affordances created by AI may lend themselves more easily to Legal Development in principle, it is plausible that many uses of the technology will also prove highly resistant to many of the established tools of regulation. Along with definitional problems around the technology, there are also practical ones: Matthew Scherer has argued that AI research and development processes are often ‘Discreet, Diffuse, Discrete, and Opaque’,¹⁸¹ and that this creates practical problems for effective national regulation.¹⁸² While these types of Legal Destruction are not a categorical threat to regulation, such technologies may leave semi-permanent ‘holes’ in legal systems. The continuing dysfunction, contestation, or rapid obsolescence of any early regulatory

¹⁷⁵ Although some have argued that the risk of cyberwar is in fact overblown. Thomas Rid, *Cyber War Will Not Take Place* (1 edition, Oxford University Press 2013).

¹⁷⁶ On the topic of global governance ‘gridlock’ generally, see also Thomas Hale and David Held, *Beyond Gridlock* (Polity Press 2017).

¹⁷⁷ One might argue that, if the barrier that is preventing adequate Legal Development is not conceptual, but simply political, that this means this is not so much a disruption to law and regulation, and rather more a disruption to broader governance. However, while the source of the problem’s intractability may be political, its effects can still carry over onto the legitimacy of the prevailing legal instruments.

¹⁷⁸ We thank Roger Brownsword for suggesting the distinct scenario explored in this paragraph.

¹⁷⁹ One might, indeed, call this the ‘Prohibition-problem’, in memory of the ill-fated US regulatory response, during the eponymous ‘Prohibition-era’, which attempted to restrict the ‘technology’ of widely available alcohol.

¹⁸⁰ Finally, in this context it is important to clarify that legal erosion does not necessarily need to start from a regulatory prohibition or restriction. In principle, it could emerge even from relatively permissive or standard-setting frameworks, if these were simply not followed or taken up by (open-source) developers of the technology, or citizens and consumers. However, because failed bans create the most extreme and (publicly visible) stresses and failures of the law, they create the most legal friction if they are not upheld, and as such create the strongest driver of erosion.

¹⁸¹ Scherer (n 16) 369.

¹⁸² For a discussion of ‘soft’ legal destruction of components of international law, see also Maas, ‘International Law Does Not Compute’ (n 76).

initiatives – and the resulting inability of regulators to adequately address certain pressing societal issues through Legal Development – may come to undercut the perceived efficacy and legitimacy of regulators in the eyes of their public.

ii. Decline

Finally, and more speculatively, there is a ‘hard’ version of the Legal Destruction argument which, while rarer, is more foundationally erosive to law insofar as the affordances introduced by a new technology potentially threaten the scaffolding of the rule of (normative) law itself by increasing avenues for resistance, evasion or contestation. Just as the early internet’s opportunities for anonymous trans-border communication were once (though only for a time) held to be beyond the jurisdiction of states, and to enable the organization of social resistance, so new technologies may afford different parties either greater ability or willingness to evade or even outright defy the rule of law. For instance, some have discussed how cryptography’s potential to resist governmental surveillance makes it an inherently political tool, one that ‘rearranges power [as] it configures who can do what, from what’.¹⁸³ Likewise, the use and proliferation of 3D-printed guns might hypothetically challenge (if only in a very limited way) the state monopoly on the use of force. Likewise, AI has been anticipated to enable the scaling up of existing categories of crime, as well as the introduction of entirely novel categories of misuse.¹⁸⁴ For instance, the use of AI in the context of DeepFakes and computational propaganda might help various non-state actors further contest the epistemic authority of states. Furthermore, the generation of anti-facial recognition ‘adversarial patches’ can help some evade (or even spoof and hijack) automatic video surveillance systems in the context of Smart City surveillance grids.¹⁸⁵ The prospects for such legal destruction are uncertain, of course. In many (if not all) areas, new technologies appear to disproportionately aid and abet (state) power and surveillance capability,¹⁸⁶ and – as the discussion of Legal Displacement indicates – AI appears no exception here. Nonetheless, insofar as new AI technologies afford more parties the possibility to evade, neuter or contest the modalities of regulation, it poses a more general challenge to the authority, legitimacy or efficacy of law in general. This is hardly without precedent: in Lessig’s analysis, the rise of private-sector code on cyberspace saw the dispersal of functional regulatory authority (by architecture) to the private sector, with all that that entailed.¹⁸⁷ Likewise, hard Legal Destruction may connote the dispersal of regulatory authority, in the broad sense that certain technologies may yield an increased ability, in a broader range of actors (beyond

¹⁸³ cf. Phillip Rogaway, ‘The Moral Character of Cryptographic Work’ (2015) 1 <<https://web.cs.ucdavis.edu/~rogaway/papers/moral.html>>.

¹⁸⁴ Miles Brundage and others, ‘The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation’ [2018] arXiv:1802.07228 [cs] <<http://arxiv.org/abs/1802.07228>> accessed 21 February 2018; Thomas C King and others, ‘Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions’ [2019] Science and Engineering Ethics <<https://doi.org/10.1007/s11948-018-00081-0>> accessed 21 February 2019; Keith J Hayward and Matthijs M Maas, ‘Artificial Intelligence and Crime: A Primer for Criminologists’ [2020] Crime, Media, Culture.

¹⁸⁵ cf. Simen Thys, Wiebe Van Ranst and Toon Goedemé, ‘Fooling Automated Surveillance Cameras: Adversarial Patches to Attack Person Detection’ [2019] arXiv:1904.08653 [cs] <<http://arxiv.org/abs/1904.08653>> accessed 24 April 2019; Kaidi Xu and others, ‘Evading Real-Time Person Detectors by Adversarial T-Shirt’ [2019] arXiv:1910.11099 [cs] <<http://arxiv.org/abs/1910.11099>> accessed 25 November 2019.

¹⁸⁶ cf. Susskind (n 9).

¹⁸⁷ Lessig, ‘The Law of the Horse’ (n 37) 533–536.

the traditional regulators), to begin to influence, shape and affect the modalities of regulation that constrain not just other regulatees, but also the regulators themselves.

4. A User-Guide to the Legal Disruption Model: Reflections

We have discussed why a comprehensive model of Legal Disruption would be needed, both in the context of law's relation with new technology in general, but also in the specific case of the full-spectrum societal changes that AI technology appears set to unleash. We then discussed in detail both Stage I (potential pathways towards a 'disruptive moment') and Stage II (the regulatory responses to a disruptive moment). Before discussing some of the specific benefits and drawbacks of the model, it is important to highlight that there are two distinct attitudes or mindsets in which the Legal Disruption Model can be deployed.

In the first, more moderate case, the model can be used *descriptively*: to assess, chart, and interpret past and present legal disruptive moments that were (or are) precipitated by socio-technological change, and which were (or are) explicitly recognised as such by a legal system. Under this analysis, one could explore the pathway by which a given specific usage of AI technology¹⁸⁸ was conceived, developed, explored, implemented, and thereby gave rise to certain new affordances which were subsequently and critically recognized ('the Disruptive Moment') by legal scholars and authorities as requiring Legal Development, offering opportunities for Legal Displacement, or threatening Legal Destruction. To give but one example, many commentators argue that responses to cyberwarfare requires a new global treaty framework having recognised it as a truly global problem. This is a clear case of a technologically-driven problem being recognized as exceeding the efficacy of existing law. This creates a need for Legal Development. Yet, the cyberwarfare example also underscores where political and operational considerations can make progress on such necessary Legal Development difficult or even intractable, in this case where the failure of the envisioned Legal Development feeds back into a second-stage Disruptive Moment, but this time leaning responses towards outcomes within Legal Destruction.

Using this model descriptively is valuable because it allows one to see the interactions of one disruptive moment with another, and the interaction effects between Legal Development, Displacement, and Destruction. Implicitly, such usage of the model would take an 'intersubjective' (or even positivist) stance on the question of what 'merits' treatment as legal disruption.¹⁸⁹ Simply, this is whatever a legal system has already proclaimed, or at least conceded, is a problem that exceeds the bounds of their (current) response. Such a mindset does not seek to contest perceptions. This approach has its merits: it enables one to work productively on problems that are already recognized and highlighted in legal scholarship, and does not seek to arbitrate or second-guess legal responses to what will and will not constitute a legally challenging technology. It does, however, risk being a 'thin' reading that limits the analyses to retrospective or presently-identified legal

¹⁸⁸ Of course, one could also extend or apply this framework to analyzing the legal disruption precipitated by any past technology. For instance, to assess the legal development precipitated by the railway in the 19th century; or wiretapping; or the internet. We leave this to future work.

¹⁸⁹ Compare here also the work on 'securitization theory', as developed by the 'Copenhagen School' of security studies, which likewise conceives of 'security threats' as being intersubjectively constituted through performative speech acts. Barry Buzan, Ole Wæver and Jaap De Wilde, *Security: A New Framework for Analysis* (UK ed edition, Lynne Rienner Publishers 1997).

challenges posed by technology. It is, furthermore, a retrospective analysis that precisely reflects the weakness of much of the current fragmented legal responses to technology. This is because it is only able to reckon with the efficacy or effects of regulatory responses that arose once the societal disruption precipitated by a new technology was severe enough that the fact of a disruptive moment became self-evident, thereby failing to moderate the turbulence that society experiences in the transitional phases.

In contrast, a ‘thick’ application of our model would take a more *predictive* stance. Here, one is able to examine both the validity of hyperbolic claims that certain technologies create a state of legal disruption, or exception that requires radical legal overhaul (or the relinquishment of certain values). Conversely, this predictive stance enables the identification and examination ongoing or anticipated socio-technological *shocks* in society which ought to give rise to a disruptive moment, even though scholars and regulators are not (yet) treating it as such. Such a normative analysis, of course, requires grounding. There are several ways to do so, and we do not take a strong position on which is more meaningful, legitimate, or warranted.

In the first place, one can evaluate (the lack of) an explicit disruptive moment by reference to the stated goals of the regulators themselves. We can then interrogate the validity of a disruptive moment. For example, do DeepFakes really pose a threat to democracy? Do trolley problems really represent a plausible or common failure mode of self-driving cars? Moreover, we can interrogate the efficacy of the proffered legal responses. Do these actually correspond to the underlying problematic (technological) affordances? For instance, the growing unemployment witnessed in many countries might be driven by technologies of automation, but might be framed by certain regulators as being driven by migration instead. In such cases, Legal Development emphasizing the introduction of new migration bans, or rollbacks on certain individual rights, would simply not connect much to the underlying causal factors of automation. In such cases, the legal disruption model can highlight the limited efficacy of a response, and reaffirm that if the stated goal of the regulator is to address the wave in unemployment, then there is a need for a *different* legal disruptive moment (focused on Legal Development to respond to growing structural unemployment as a result of automation).

In the second place, such a normative analysis might be grounded by appeal to an objective ‘pre-competitive’ criteria. Compare, for instance, Roger Brownsword’s work on how various uses of AI (in personal recommender systems, killer robots and biased decision making) may be understood to affect the fundamental ‘commons conditions’ of human agency or human life – the conditions necessary to have and express any higher-level value or normative disagreements in the first place.¹⁹⁰ In such cases, one could argue that certain new technologies are on a plausible trajectory that will threaten to exceed the bounds of the existing legal system eventually (for example by radically challenging legal concepts; or by offering far-reaching opportunities for wholesale Legal Displacement or even contestation and Legal Destruction). The normative argument here would be that an early response to what presently appears to be a small threat might be justified lest it eventually snowballs into some truly catastrophic threat. Even if the challenges or risks created by these technologies are not yet severe enough to warrant treatment as a truly disruptive moment, or its severity has not yet been recognized as justifying inclusion within the legal disruption model, its

¹⁹⁰ cf. Brownsword, ‘From Erehwon to AlphaGo’ (n 21).

trajectory suggests more cautious treatment. This could perhaps take the form of a series of smaller, precautionary steps of Legal Development in order to forestall an accumulation of ‘tensions’, the locking-in of path-dependent decisions, and far more broad, pervasive, and irreversible disruptive moments down the road.

Finally, a normative analysis might simply be grounded in the usual set of first-order political or normative principles. This can take place by opening up opportunities to ‘reveal new rights’,¹⁹¹ or to identify and contest areas where certain new affordances emerge which a prevailing legal system might perceive as being unproblematic. An example of this might be the ability of AI technology to empower scrutiny and surveillance writ large. Indeed such a system might even be complicit in implementing such effects, as part of a move towards Legal Displacement, but which could be considered problematic from a wide range of normative or political principles.

These are some basic reflections on how one could approach and implement this model, and what implications this has for the types of analysis one can undertake, and the (normative or epistemic) precommitments this entails. We do not, however, seek to settle debates over which of these mindsets or attitudes is preferable here. Keeping this distinction in mind, we now turn to discuss some of the benefits and drawbacks of the model more broadly.

5. Advantages and Drawbacks of the Proposed Model

From the perspective of how to approach socio-technologically driven change, a strong advantage of the Legal Disruption Model is that law, policy, regulation and governance approaches are placed front and centre. In foregrounding the regulatory framework, this relegates the importance of the intrinsic characteristics of the technology under discussion. As such, this approach clearly follows Balkin and rejects the path taken by Calo. Taking such a stance is important because AI and the claims made around its capacities can be extremely distracting for legal scholars and decision-makers.¹⁹² Furthermore, the legal disruption approach is not limited to examining the legal impact emanating directly from the properties of AI—for example the questions of liability arising from the liminal position that AI occupies between the categories of agent/person and patient/object—but it is also capable of taking into account ‘second order’ effects that impact upon the regulatory configuration. While Balkin situates such effects primarily in social relationships, legal disruptions can also cascade from changes in political, economic or normative-cultural shifts that AI applications trigger. The prospect for widespread technological unemployment, for example, might constitute an example of this: studying the intrinsic characteristics or capacities of AI from a legal perspective does not appear to further legal or policy responses capable of addressing the challenge of mass-unemployment. The legal disruption perspective, however, would address the socio-economic upheaval as regulatory turbulence and engage with the complexity of such a problem.

Placing law and regulation front and centre through the legal disruption framework also acts as a counterbalance to concerns that ‘technologists’, and their advocates who are positioned to gain materially from the development of emerging technologies, dominate in social commentaries and ethical discussions in relation to those developments. At the same time, ethicists and jurists can

¹⁹¹ Jack Parker and David Danks, ‘How Technological Advances Can Reveal Rights’ (2019) <http://www.aies-conference.com/wp-content/papers/main/AIES-19_paper_129.pdf>.

¹⁹² Mandel (n 129).

often be overly sceptical of fast-paced technological change or at least keep their distance from engaging with such developments which are perceived to be beyond their separate magisteria. Thus, the conceptual framing of legal disruption provides an additional benefit: it becomes a clearly *legal* undertaking. This appears trite, but there has been a strong temptation for legal scholars to take on the role of technology critics, journalists, and even futurists, as they became distracted by the hype around emerging technologies such as AI. The initial ‘robolaw’ framing failed to sufficiently emphasise the legal dimension, perhaps because it merely juxtaposed the law alongside technological developments.

In the context of the Easterbrook-Lessig debate in cyberlaw, discussed above, the legal disruption framework provides a suitably high threshold for the legal consideration of AI and so maximises the possibility that something fruitful might be learnt from this consideration about the whole law. This is because legal questions that can be accommodated in existing legal doctrine are sufficiently routine to trigger relatively reflexive thinking and response processes. The requirement for legal disruption, however, *defamiliarises* and thereby facilitates a re-evaluation of why the law is structured in such a way as to have inconsistencies and anomalies when AI is factored into the system. In this sense, AI is merely the vehicle which overturns automatic thinking processes in response to legal questions by forcing us to articulate justifications for positions that would have otherwise been taken for granted and be subject to wide consensus.

An additional benefit of utilising legal disruption as a conceptual framework is that it orientates regulatory attention towards a dynamic target. Whereas Balkin criticised Calo for unnecessarily fixing a set of characteristics as being legally-relevant, Balkin’s salience approach appears to be more static than initially evident because it is focused upon the social impact of robotics without factoring in the mechanisms that alter this through time. In an important sense, the legal disruption framework addresses sociotechnical change, and indeed change more generally, over time because it focuses on both the disruptive properties of the technology and how this affects social-normative relations.¹⁹³ The disruptive changes covered by this framework must be both structurally significant and trigger turbulence. In these senses, the legal disruption framework establishes additional criteria that direct attention at the sources of the challenges introduced by AI. As legal and policy responses to challenges posed by artificial intelligences are overcome or otherwise settled, those issues lose their disruptive effect and thereby fall out of the ambit of the model. As such, legal disruption constantly filters out those questions that no longer offer the alternative vantage point that might lead to new understandings of the law.

Related to this, the legal disruption approach allows us to deploy AI as a critical mirror to the legal and regulatory system. This is because AI forces us to reevaluate a great number of social constructs and knowledge that were just taken for granted as functional, and requires that we enter in an exercise of self-assessment and criticism.¹⁹⁴ It offers a rare chance to step outside of contemporary legal processes, principles and presumptions, to stress-test their continuing efficacy, validity and viability, so as to review and if necessary revise them. From a separate vantage point, it

¹⁹³ Bennett Moses (n 50).

¹⁹⁴ This mirror might only reveal ‘latent’ ambiguities which have persisted in the law all along, but have only become perceptible through the prism of AI, for example, Calo, ‘Robotics and the Lessons of Cyberlaw’ (n 58) 550–553; Lawrence Lessig, *Code: And Other Laws of Cyberspace, Version 2.0* (2nd Revised ed. edition, Basic Books 2006) 25–26 <<http://codev2.cc/download+remix/Lessig-Codev2.pdf>>.

is also possible to see any flaws or inconsistencies more clearly. It also provides a rare opportunity to improve the law, by updating its doctrine to reflect the contemporary scientific paradigm and future realities. In other words, by applying the legal disruption framework we can improve the law and not just learn more about its dynamics: by examining the ways in which its doctrines might breakdown when confronted with AI we are given the chance to make it more resilient against other extraneous shocks (for example, the shocks emanating from other emerging technologies such as neuroscience and nanotechnology). The lens of legal disruption, then, is really at the core of ‘illuminating the entire law’ from a different perspective, and not just the disruptive impacts of a particular set of technologies. It is a general framework for understanding the process of legal disruption that can be applied to specific, technology-induced, disruptions.

Finally, the legal disruption model offers three main practical benefits. The focal nature of the disruptive moment gives rise to three general regulatory responses that both streamlines and separates those responses. First, this saves users from having to reinvent the wheel every time legal and regulatory questions are raised anew by AI (and by other emerging technologies); but it also expands the range of vision for those framing such responses as these need not be mutually exclusive, but rather may be nested at different levels of analysis. Second, in setting out the flows and feedbacks of possible regulatory responses, the legal disruption framework illuminates the types of consequences that might flow from the disruptive moment, the kinds of regulatory responses that might be pursued, and the possible consequences of adopting those responses. And third, the legal disruption framework identifies subtle yet perilous trajectories that regulatory responses might take which are more difficult or even impossible to recover from: such warning lights suggest that greater vigilance is necessary under certain conditions that would have been overlooked absent this framework.

Despite its many advantages, the legal disruption framework has its limitations. The legal disruption framework, like Lessig’s regulatory modalities, is atemporal. The lack of trajectory constitutes a significant shortcoming that limits the legal disruption framework to being merely responsive to external developments. While the legal disruption framework identifies categories and sub-categories for responses to the disruptive moment, the model as it currently stands does not indicate the factors of conditions which prejudice certain types of responses over others. Presumably, the configuration of the preceding factors and its interaction with the characteristics of the legal and regulatory order would suggest that certain categories of responses are more likely or more useful than others.

Questions also arise as to how exactly we identify a disruptive moment. In setting out the framework we have commented on what happens to generate a disruptive moment in the abstract (i.e. some change – real, perceived, or anticipated – in behaviour that puts pressure on the existing legal framework), but more detail will be required if we are to identify disruptive moments in real time. What are the characteristics that signal or distinguish one? Are there features in the preceding factors that prejudice towards certain general category outcomes over others?

Finally, there are lingering questions about the more advanced dynamics that might support or encourage certain forms of legal disruption. How does Legal Development, Legal Displacement, or Legal Destruction feedback into social behaviour and generate the potential for future disruptions? How do social values change in response to technological change? Would this moderate any resulting legal disruption? These unaddressed questions provide plenty of room for future

elaboration and refinement of the model. We do not feel, however, that they undermine its utility in its current form.

6. Concluding Thoughts for the Artificial Intelligence and Legal Disruption Model

In this paper, we have developed a new model to understand the ways in which new technologies drive legal disruption in the doctrine, processes, practices, and even viability of legal systems, and we clarified this model through an application to the legal disruptions driven by AI systems. In section I, we first defended the need for this new conceptual model. In section II, we situated this model in relation to preceding debates concerning the interaction of legal systems with new technologies, in particular the ‘cyberlaw’ and ‘roboLaw’ debates; and to set out a detailed model for understanding the legal disruption precipitated by AI, both in terms of the factors and developments that feed into a ‘disruptive moment’, and in terms of the different types of responses (Development, Displacement, or Destruction) that can result, as well as how to use such a model of legal disruption.

The study of artificial intelligence and legal disruption should not be viewed as its own distinct legal subject. It is, rather, a vantage point from which we can scrutinise existing legal presumptions, principles and processes. This holds the promise of being able to refine and reinvigorate the law, rather than just the scramble to fill in the *lacunae* that are wrenched open or revealed by the prospect of artificial intelligence. Engaging critical reflection about the roles and functions that the law should fulfil, not merely how *lex lata* actually functions, encourages responses in more expansive terms about the prospects and pitfalls of the regulatory enterprise. By employing the legal disruption framework we can avoid Easterbrook’s Law of the Horse fallacy, even if there are certain developments which do in fact spawn entirely new legal areas (corporate law and European Union law come to mind) which are now not considered contentious.

As artificial intelligence is often claimed to be an exponential technology,¹⁹⁵ and law progresses incrementally in a linear fashion, there is bound to be a point at which the exponential take off crosses the straight line if these assumptions hold. Everything to the left of this intersection, where AI is below the line, is where hype about the technology does not quite live up to expectations and is generally disappointing in terms of functioning and capability. To the right of this intersection, however, the previously dull technology takes on a surprising and startling tone as it rapidly outpaces both predictions about its capacities and collective abilities to contextualise, accommodate or situate it. It is widely claimed that we are now nearing this intersection. If these claims hold up, the law is one of the institutions that stands to be shocked by the rapid progression and incorporation of AI into society. If this is right, then it is important to start projecting forward in an attempt to minimise the gap between exponential technologies and linear expectations. The legal disruption framework we have presented does exactly this. Furthermore, even if these claims turn out to be misguided, thinking through such transformations sheds different light upon the legal enterprise which hopes to illuminate the entire law.

¹⁹⁵ Brynjolfsson and McAfee (n 106).

Acknowledgements & Disclaimer:

The authors would like to thank students of the Autumn 2018, Autumn 2019, and Spring 2020 ‘Artificial Intelligence and Legal Disruption’ classes for commentary and engagement on early version of this model; and attendees and participants at the TILTING2019 conference for their input on developing critical aspects of the research; Agnieszka Mlicka for facilitating an engaging ‘Visual Agency’ workshop that enabled the authors to further develop the ideas presented here; Andrew Mazibrada for valuable comments and inputs during this workshop; and Roger Brownsword for encouragement and support of the research dimension. Any remaining errors are our own. No conflicts of interest are identified.