*Vicki Waye**

REGTECH: A NEW FRONTIER IN LEGAL SCHOLARSHIP

I Introduction

Digital disruption is rolling across a succession of industries, and along the way transforming the social, economic and legal landscapes. The combination of big data, artificial intelligence, cloud computing and robotics is not only disrupting specific legal doctrines and practices,[1] but is also beginning to impact the manner and process of engagement between regulators and the regulated.[2] These changes are being driven as much by growing regulatory intensity and complexity[3] and the escalating costs of regulatory compliance,[4] as by the digital technologies themselves. One commentator has even suggested that we have reached a tipping point, where corporate compliance professionals may soon outnumber

---

*   Professor of Law, University of South Australia; Editor, *Adelaide Law Review* 1989–95.

1   For example, the updating of legal principles to accommodate the use of smart contracts and autonomous vehicles and the adaptation of intellectual property rights like copyright to new technologies.

2   Loretta Michaels and Matthew Homer, 'Regulation and Supervision in a Digital and Inclusive World' in David Lee Kuo Chuen and Robert H Deng (eds), *Handbook of Blockchain, Digital Finance and Inclusion Volume 1: Cryptocurrency, Fintech, Insurtech and Regulation* (Academic Press, 2018) 329, 333–7 [14.3]; Ioannis Anagnostopoulos, 'Fintech And Regtech: Impact on Regulators And Banks' (2018) 100 (November) *Journal of Economics and Business* 7; Nizan Geslevich Packin, 'Regtech, Compliance and Technology Judgment Rule' (2018) 93(1) *Chicago-Kent Law Review* 193, 206–12.

3   See 'National Red Tape Survey', *Australian Chamber of Commerce and Industry* (Web Page, March 2015) <https://www.australianchamber.com.au/publication_taxonomies/red-tape-survey/>. The US National Association of Manufacturers estimates that manufacturers face 297,696 regulatory restrictions on their operations: 'Holding Us Back: Regulation in the US Manufacturing Sector', *National Association of Manufacturers* (Web Page) https://www.nam.org/Data-and-Reports/Reports/Holding-Us-Back--Regulation-of-the-U-S--Manufacturing-Sector/>. See also David Frisch, 'Commercial Law's Complexity' (2011) 18(2) *George Mason Law Review* 245.

4   See Senate Select Committee on Red Tape, Parliament of Australia, *Policy and Process to Limit and Reduce Red Tape* (Final Report, December 2018) 5–6 [2.4], quoting an annual compliance burden of $65 billion.

police officers,[5] placing an unsustainable regulatory burden on participants in highly regulated sectors such as banking and finance.[6]

To cope with the compliance challenge, a large number of automated solutions designed to facilitate regulatory compliance have begun to emerge under the moniker of 'Regtech'. Many of these have arisen within the finance industry alongside Fintech developments[7] that harness digitisation to create new financial products and services and deliver improved efficiencies in existing services.[8] Examples of Regtech include the use of machine learning and biometrics to digitally verify identity and comply with 'know your customer' regulation,[9] and the use of cloud computing and artificial intelligence to manage and secure customer data in order to comply with ePrivacy and consumer data rights regulation.[10] Consequently, Regtech is often character-ised as a spin-off or subset of Fintech.[11] However, while the 'Regtech' label fits neatly into a stable of other 'X' tech monikers associated with financial services and many so-called Regtech developments are surfacing in the finance industry, the use of big data, artificial intelligence, robotics and cloud computing to facilitate and/or monitor compliance is not confined to any particular regulatory domain. In fact, Regtech has been evolving across a number of sectors including policing,[12]

---

[5]   See William S Laufer, 'A Very Special Regulatory Milestone' (2018) 20(2) *University of Pennsylvania Journal of Business Law* 392, 393–4, referring to US data comparing corporate regulatory budgets and local, state and federal law enforcement expendi-tures as well as employment data comparing JP Morgan's employment of compliance professionals and the employment of NYPD officers and FBI agents.

[6]   Tom Butler and Robert Brooks, 'On the Role of Ontology-Based Regtech for Managing Risk and Compliance Reporting in the Age Of Regulation' (2018) 11(1) *Journal of Risk Management in Financial Institutions* 19, 20; Douglas W Arner, Jànos Barberis and Ross P Buckley, *Fintech and Regtech in a Nutshell and the Future in a Sandbox* (Chartered Financial Analyst Institute Research Foundation, 2017) 13.

[7]   Thomas Puschmann, 'Fintech' (2017) 59(1) *Business & Information Systems Engi-neering* 69.

[8]   Dirk A Zetzsche et al, 'Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation' (2017) 23(1) *Fordham Journal of Corporate and Financial Law* 31, 34–5; Veerle A Colaert, 'RegTech as a Response to Regulatory Expansion in the Financial Sector' (Faculty of Law, Katholieke Universiteit Leuven, June 2018) <https://ssrn.com/abstract=2677116>.

[9]   See, eg, 'Know Your Customer (KYC)', *Identity Mind* (Web Page) <https://identity-mindglobal.com/kyc>; 'WebVOI: Three Times Smarter', *InfoTrack* (Web Page, 2019) <https://www.infotrack.com.au/products/innovative-technology/verification-of-identity/web-voi>; 'Onfido: Identity Verification at The Speed of Life', *Onfido* (Web Page) <https://onfido.com>.

[10]   See, eg, 'Trunomi: Switch on Trust', *Trunomi* (Web Page) <https://www.trunomi.com/>; 'NGData: The Intelligent Customer Data Platform', *NGData* (Web Page) <https://www.ngdata.com/platform>.

[11]   See Anagnostopoulos (n 2) 13; United Kingdom Chief Scientific Adviser, *FinTech Futures: The UK as a World Leader in Financial Technologies* (Report, 2015) ch 6.

[12]   Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94(5) *Washington University Law Review* 1109.

forestry management,[13] food and hospitality,[14] healthcare,[15] international trade,[16] and logistics and supply chain management.[17] Parallel to these developments, and to further digitise compliance, attempts are being made to create whole-of-system machine-readable legislation, regulation and policy guidance.[18] Evidently, a revolution is upon us.[19]

By contrast with other analyses that have sought to predict the impact of these revolutionary developments on the regulatory process or upon the relevant industry sector, the purpose of this article is to study the effect of Regtech on future legal scholarship. The 40th anniversary of the *Adelaide Law Review* provides an ideal opportunity not only for looking back and tracing the pathway of legal development over the past 40 odd years but also for considering the future trajectory of legal scholarship in light of phenomenon like the advent of Regtech. While acknowledging the often thin and overlapping divide between different approaches to research, the article embarks on this task by examining how Regtech may affect (a) doctrinal, (b) normative/reform oriented, (c) theoretical, (d) interdisciplinary and (e) empirical forms of legal scholarship. However, before embarking on that analysis, this article commences by examining Regtech and its possibilities.

---

[13] 'Forestry Plantation Regtech Project', *New South Wales Government* (Web Page) <https://www.digital.nsw.gov.au/project-story/forestry-plantation-regtech-project>.

[14] For example, the Southern Nevada Health District uses an app employing geotagging and natural language processing to analyse food consumers' Twitter data and identify sources of food-borne illness: see William D Eggers, David Schatsky and Peter Viechnicki, *AI-Augmented Government: Using Cognitive Technologies to Redesign Public Sector Work* (Report, 2017) 11–12.

[15] See, eg, Halyna Levko, 'Challenge Accomplished: Healthcare Fraud Detection Using Predictive Analytics', *Romexsoft* (Blog Post, 25 April 2017) <https://www.romexsoft.com/blog/healthcare-fraud-detection>.

[16] See, eg, 'Global Trade Management', *Thomson Reuters* (Web Page) <https://tax.thomsonreuters.com/en/onesource/global-trade-management/export-compliance>.

[17] For example, the mining and artificial intelligence analysis of media reports, surveillance footage and mobile phones to support modern slavery compliance: Reuters, 'In the Fight to End Modern Slavery, Machines May Hold the Key', *Voice of America News* (online, 9 August 2017) <https://www.voanews.com/a/modern-slavery-artificial-intelligence/3979776.html>.

[18] See, eg, the Australian Government's utilisation of regulation as a platform to digitise regulation tools and services: 'Regulation as a Platform', *CSIRO and Data61* (Web Page) <https://www.data61.csiro.au/en/Our-Work/Future-Cities/Optimising-service-delivery/RaaP>. See also the New Zealand Government's approach to digitising regulation by developing government policies and rules in machine consumable format: *Better Rules for Government Discovery* (Report, March 2018) <https://www.digital.govt.nz/dmsdocument/95-better-rules-for-government-discovery-report/html>.

[19] KPMG, *There's a Revolution Coming: Embracing the Challenge of RegTech 3.0* (Report, 2018) <https://home.kpmg/uk/en/home/insights/2018/09/regtech-revolution-coming.html>.

II Regtech

There is no universally accepted definition of 'Regtech'. The term appears to have been first used in a United Kingdom Treasury Budget Report from 2015,[20] which was later elaborated in a review published by the United Kingdom's Office for Science entitled, 'Fintech Futures: The UK as a World Leader in Financial Technologies'.[21] Others use terms such as 'algorithmic regulation',[22] 'smart regulation',[23] 'dynamic/ adaptive regulation'[24] and — when applied to supervisory agencies like monetary authorities — 'Suptech'.[25] The common thread among these terms is the use of digital technologies to facilitate the delivery of regulatory outcomes more effectively and cheaply than traditional, non-digital forms of regulation. These traditional types of regulation are currently beset by a pastiche of methodologies and taxonomies, resource-intensive manual processing, lengthy reporting lags, and significant gaps in regulatory coverage.[26] Regtech employs tools that involve the collation and analysis of big data, natural language processing, the linking of analytics with machine learning, the application of distributed ledger technology, and the automation of advanced algorithmic processes to expedite and improve compliance and regulation.[27]

---

[20] Her Majesty's Treasury, *Budget 2015* (Report, 18 March 2015) 53 [1.204], 98 [2.272] <https://www.gov.uk/government/publications/budget-2015-documents>.

[21] United Kingdom Chief Scientific Adviser (n 11).

[22] See Karen Yeung, 'Algorithmic Regulation: A Critical Interrogation' (2018) 12(4) *Regulation & Government* 505; Lena Ulbricht, 'When Big Data Meets Securitization: Algorithmic Regulation with Passenger Name Records' (2018) 3(2) *European Journal for Security Research* 139, 145–6; Thomas Riis and Sebastian Felix Schwemer, 'Leaving the European Safe Harbor, Sailing towards Algorithmic Content Regulation' (Research Paper No 2019–64, Faculty of Law, University of Copenhagen, 7 December 2018).

[23] See Zetzsche et al (n 8).

[24] See Lawrence Baxter, 'Adaptive Financial Regulation and Regtech: A Concept Article on Realistic Protection for Victims of Bank Failures' (2016) 66(3) *Duke Law Journal* 567.

[25] Suptech or supervisory technology refers to the use of technology by supervisory agencies to support their supervisory functions: Basel Committee on Banking Supervision, *Sound Practices: Implications of Fintech Developments for Banks and Bank Supervisors* (Report, February 2018) 35 <https://www.bis.org/bcbs/publ/d431.htm>.

[26] For the United Kingdom's approach, see 'Call for Input: Supporting the Development and Adoption of Regtech', *Financial Conduct Authority* (Web Page, 6 April 2016) <https://www.fca.org.uk/news/news-stories/call-input-supporting-development-and-adoption-regtech>. For Australia's approach, see Australian Securities and Investment Commission, *ASIC's Innovation Hub and Our Approach to Regulatory Technology* (Report 523, 26 May 2017) 18–19 [72] <https://asic.gov.au/regulatory-resources/find-a-document/reports/rep-523-asic-s-innovation-hub-and-our-approach-to-regulatory-technology>.

[27] Anagnostopoulos (n 2) 14.

It forms part of the broader digital transformation of government that has evolved over the last 20 years, and which is culminating in more integrated and transactional governance via the cyber world.[28] Through these developments, governments worldwide[29] — including that of Australia[30] — plan to use digital technologies to deliver more responsive policy, a simpler and more integrated public service built around the needs of users, proactive and personalised delivery of information, advice and assistance, and more effective risk-based regulation. According to the Organisation for Economic Co-operation and Development ('OECD'), this steady assimilation of digital technologies into government offers 'opportunities for more collaborative and participatory relationships'[31] between government, citizens and businesses. As a subset of the digital transformation of government, it is envisaged that the widespread adoption of Regtech will result in seamless end-to-end approvals processes across government, more dynamic interaction with regulators, and the embedding of compliance by design.[32]

The Regtech field is largely populated by a growing number of private businesses, some of which have been established for many years. However, many of these businesses have emerged since the global financial crisis as regulatory requirements around transparency and accountability of business have strengthened, the volume of digitised data has vastly increased, and technological capabilities have matured.[33]

28　Tomasz Janowski, 'Digital Government Evolution: From Transformation to Contextualization' (2015) 32(3) *Government Information Quarterly* 221; Luis F Luna-Reyes and J Ramon Gil-Garcia, 'Digital Government Transformation and Internet Portals: The Co-Evolution of Technology, Organizations and Governance' (2014) 31(4) *Government Information Quarterly* 545.

29　See, eg, the evolution of e-government internationally: Allan Brown, Jerry Fishenden and Mark Thompson, *Digitizing Government: Understanding and Implementing New Digital Models* (Palgrave Macmillan, 2014) ch 1. See also the incorporation of digital government reviews of various OECD countries: 'Digital Government', *Organisation for Economic Co-operation and Development* (Web Page) <http://www.oecd.org/gov/digital-government>.

30　'Digital Transformation Strategy', *Digital Transformation Agency* (Web Page, 2018) <https://www.dta.gov.au/our-projects/strategies/digital-transformation-strategy>. The Australian Government has established a Digital Transformation Agency to drive these planned transformations: see 'About Us', *Digital Transformation Agency* (Web Page) <https://www.dta.gov.au/about-us>.

31　Public Governance and Territorial Development Directorate, Organisation for Economic Co-operation and Development, *Recommendation of the Council on Digital Government Strategies* (Report, 2014) 2 <http://www.oecd.org/gov/digital-government/recommendation-on-digital-government-strategies.htm>.

32　Australian Productivity Commission, *Shifting the Dial: 5 Year Productivity Review*: *Regulation in the Digital Age* (Supporting Paper No 13, 3 August 2017). See also Karen Yeung, '"Hypernudge": Big Data as a Mode of Regulation by Design' (2017) 20(1) *Information, Communication & Society* 118.

33　Douglas W Arner, Jànos Barberis and Ross P Buckley, 'RegTech: Building a Better Financial System' in David Lee Kuo Chen and Robert H Deng (eds), *Handbook of Blockchain, Digital Finance, and Inclusion* (Academic Press, 2017) 359, 363;

These firms service the compliance and risk management needs of larger business organisations like banks. A recent survey conducted by Deloitte of 306 known private Regtech firms in the finance sector found that 40% of these were focussed on supporting compliance, 25% on identity management and control, 15% on risk management, and 9% on transaction monitoring.[34]

To date, financial regulators have remained on the periphery, but nonetheless supportive, of Regtech developments. However, they are increasingly demonstrating a willingness to harness Regtech themselves to enhance the efficiency and efficacy of regulatory policy formulation as well as their supervision and enforcement activities. Some, like the Monetary Authority of Singapore, are already using Regtech to detect financial market manipulation.[35] Similarly, the United States' Securities and Exchange Commission is currently using big data techniques such as topic modelling and cluster analysis to identify outlier behaviour and spot potential investment adviser misconduct.[36] For the future, the United Kingdom's Financial Conduct Authority has expressed interest in using Regtech for regulatory policy modelling and impact analysis, using real-time data to engage in more accurate risk assessment to facilitate better targeted regulatory activity and to automate some regulatory processes.[37] Meanwhile, the Australian Securities and Investment Commission ('ASIC') is conducting several trials of Regtech: a cognitive tool that analyses the webpages of service providers in the self-managed superannuation fund sector; machine learning to conduct document analysis for forensic purposes; monitoring of social media trends; and market and graph analysis tools to identify connections between business entities.[38]

Wendy L Currie, Daniel P Gozman and Jonathan JM Seddon, 'Dialectic Tensions in The Financial Markets: A Longitudinal Study of Pre- and Post-Crisis Regulatory Technology' (2019) 33(4) *Journal of Information Technology* 304. For a discussion of problems associated with maintaining a number of proprietary, domain-specific automated compliance checking in the construction industry: see also Nawari O Nawari, 'A Generalized Adaptive Framework (GAF) for Automating Code Compliance Checking' (2019) 9(4) *Buildings* 86, [3].

[34]   'RegTech Universe', *Deloitte* (Web Page, 30 May 2019) <https://www2.deloitte.com/lu/en/pages/technology/articles/regtech-companies-compliance.html#>. An initial survey of Regtech firms was conducted in 2017 and has been continuously updated. The figures outlined in the text are current as of 27 May 2019.

[35]   Alice Shen, 'MAS Uses Machine Learning to Spot Market Manipulation', *Central Banking* (online, 5 April 2019) <https://www.centralbanking.com/fintech/regtech-suptech/4129566/mas-uses-machine-learning-to-spot-market-manipulation>.

[36]   Scott W Bauguess, 'The Role of Big Data, Machine Learning and AI in Assessing Risks: A Regulatory Perspective' (Keynote Address, OpRisk North America, 21 June 2017) <https://www.sec.gov/news/speech/bauguess-big-data-ai>.

[37]   Financial Conduct Authority, *FCA Research* (Agenda, April 2019) <https://www.fca.org.uk/publications/corporate-documents/fca-research-agenda>.

[38]   ASIC (n 26) 21 [90].

Outside of the financial services sector, code-based rules that automatically execute underlying decision-making have already begun transforming other areas of public regulation and enforcement. Police and security forces have been at the forefront of these initiatives. For instance, police in the United Kingdom are reportedly using algorithmic data technologies to support operational intelligence gathering and analysis, to predict where offences are likely to take place and to undertake risk analysis of targeted individuals.[39] Similar activities are now mainstream in many larger police forces in the United States, which mine big data to correlate offending with geospatial information, social networks, drivers' licenses and commercial transactions.[40] Data analytics have also been used to predict the risk of recidivism for the purpose of informing parole and sentencing determinations.[41] In other fields, algorithmic regulation has also been deployed (with mixed success) to assess the quality of healthcare and education under mandated quality assurance frameworks.[42] It is also deployed by online platforms like Google and Youtube to protect the copyright of digitised work and automatically detect and remove infringing materials from the platforms.[43] Thus, while the term Regtech originated in the financial services sector, it falls within a broader genre of digital regulation.

Regtech seeks to computationally align the business processes of regulated entities with the rules set out in the legal language of regulation. A proof of concept developed for the construction industry, which involved aligning Building Information Modelling ('BIM') software[44] with a framework for automating building code compliance checking, illustrates how Regtech can work.[45] The creators' purpose was to establish a process that marries building code requirements such as minimum

---

[39]   Marion Oswald et al, 'Algorithmic Risk Assessment Policing Models: Lessons from The Durham Hart Model and "Experimental" Proportionality' (2018) 27(2) *Information & Communications Technology Law* 223.

[40]   Andrew Guthrie Ferguson, 'Predictive Policing and Reasonable Suspicion' (2012) 62(2) *Emory Law Journal* 259, 277.

[41]   Ric Simmons, 'Big Data and Procedural Justice: Legitimizing Algorithms in the Criminal Justice System' (2017) 15(2) *Ohio State Journal of Criminal Law* 573.
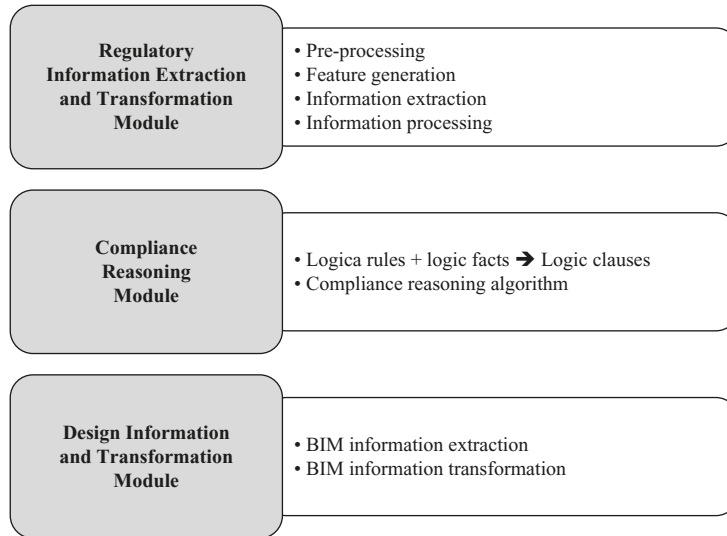
[42]   Alex Griffiths et al, 'Algorithmic Regulation' (Discussion Paper No 85, Centre for Analysis and Risk Regulation, September 2017) 19–24. See Kate Carnell and Ron Paterson, *Review of National Aged Care Quality Regulatory Processe*s (Report, October 2017) xi [3.i]. The Review recommended the formulation and piloting of an algorithm for performance benchmarking of aged care facilities in Australia.

[43]   Dan L Burk, 'Algorithmic Fair Use' (2019) 86(2) *University of Chicago Law Review* 283, 284.

[44]   According to Autodesk, BIM 'is an intelligent 3D model-based process that gives architecture, engineering, and construction (AEC) professionals the insight and tools to more efficiently plan, design, construct, and manage buildings and infrastructure': 'What is BIM', *Autodesk* (Web Page, 2019) <https://www.autodesk.com/solutions/bim>.

[45]   Nawari (n 33); Jiansong Zhang and Nora M El-Gohary, 'Integrating Semantic NLP and Logic Reasoning into a Unified System for Fully-Automated Code Checking' (2017) 73 (January) *Automation in Construction* 45.

room area, ceiling height, energy efficiency, evacuation and safety, and construction quality with elements of the three-dimensional modelling of the physical and functional characteristics of buildings afforded by BIM. The proof of concept comprises the following components:[46]

Regulatory Information Extraction and Transformation Module	• Pre-processing • Feature generation • Information extraction • Information processing
Compliance Reasoning Module	• Logica rules + logic facts ➔ Logic clauses • Compliance reasoning algorithm
Design Information and Transformation Module	• BIM information extraction • BIM information transformation

The first component requires pre-processing of the raw text of building codes using natural language processing techniques, followed by feature generation to develop a set of syntactic and semantic features of the text. During this process, a building ontology is utilised to generate semantic meaning so that words and terms found in the text can be matched to various categories within the ontology. An example might be a category that references walls. Such a category might include retaining walls, exterior walls, load bearing walls, fire walls and so on. Information extraction processes then isolate instances of these categories from the building code text, which are then transformed into logic rules. These logic rules then generate compliance checking logic clauses.

The BIM material is subjected to similar processing and, along with the transformed regulatory information, is fed into a compliance reasoning system comprised of functional inbuilt logic clauses and the utilisation of an artificial intelligence application and B-Prolog reasoner to create compliance checking reports.[47] By this means, Regtech provides a bridge between the instruments used by architects and engineers to design and measure building performance and the standards used by regulators to ensure that buildings meet public policy objectives related to safety and sustainability.

As a result of these capabilities, a number of commentators have predicted that Regtech will not only enhance regulatory efficiency and efficacy, but that it also

---

46   Zhang and El-Gohary (n 45) 48–52.

47   Ibid 51–2.

foreshadows a paradigm shift in the nature of regulation. Instead of being a reactive principle-based approach that produces standardised regulatory responses, certain commentators believe Regtech will introduce a more proactive and adaptive, insight-driven regulation that not only monitors and responds to non-compliance but also provides an enabling environment for business and consumers.[48] These predictions, therefore, assume that regulators will have broad-based access to the data generated by regulated entities through automated real-time reporting processes, and that regulators will also be able to monitor, evaluate and respond to this reporting in real-time. However, while some progress has been made in formulating the types of systems necessary to support these developments in discrete fields such as mandatory reporting of child abuse and medication error,[49] widespread adoption remains a long way off.

Given the nascent level of Regtech development, whether these predictions come to pass or whether the Regtech sceptics will be vindicated constitutes fertile ground for legal scholarship. The potential for Regtech to deliver better regulation is clearly enticing, but normative and practical questions remain, associated with matters such as: the absence of semantic interoperability between the plethora of proprietary Regtech solutions;[50] the variable capacity of organisations and regulators to effectively process and act on the data and insights they ascertain through Regtech;[51] the opacity of bias in algorithmic decision-making;[52] and the ability to manipulate data for illegitimate ends.[53] Consequently, Regtech is already attracting the attention of legal scholars and as this article proposes, will have a significant impact on future legal scholarship. The next section of this article therefore proffers views on the potential impact of Regtech across established methods of legal scholarship: doctrinal; normative; theoretical; interdisciplinary; and empirical forms of research.

---

[48]   World Government Summit, *Regtech for Regulators: Re-Architect the System for Better Regulation* (Report, February 2018); Arner, Barberis and Buckley (n 33).

[49]   Mohammad Badiul Islam and Guido Governatori, 'Rulers: A Rule-Based Architecture for Decision Support Systems' (2018) 26(4) *Artificial Intelligence and the Law* 314.

[50]   Tom Butler and Leona O'Brien, 'Understanding Regtech for Digital Regulatory Compliance' in Theo Lynn et al (eds), *Disrupting Finance: FinTech and Strategy in the 21st Century* (Palgrave MacMillan, 2019) 85.

[51]   Sarah Giest, 'Big Data for Policymaking: Fad or Fasttrack?' (2017) 50(3) *Policy Sciences* 367, 370–1.

[52]   Robert Brauneis and Ellen P Goodman, 'Algorithmic Transparency for the Smart City' (2018) 21(1) *Yale Journal of Law & Technology* 103; Rory Van Loo, 'Rise of the Digital Regulator' (2017) 66(6) *Duke Law Journal* 1267, 1309, 1322; Wayne A Logan and Andrew Guthrie Ferguson, 'Policing Criminal Justice Data' (2016) 101(2) *Minnesota Law Review* 541, 545; Marijn Janssen and George Kuk, 'The Challenges and Limits of Big Data Algorithms in Technocratic Governance' (2016) 33(3) *Government Information Quarterly* 371, 374–5.

[53]   Packin (n 2) 214.

A *Doctrinal*

Doctrinal research encompassing an exposition of the rules and principles of positive law[54] has been characterised as the 'core of legal scholarship',[55] and thus described as the quintessential method that distinguishes legal scholarship from the research praxis of other disciplines.[56] Doctrinal research comprises the systematic inter-pretation and analysis of legal materials such as legislation, regulation, case law, regulatory guidance, soft law and authoritative legal texts. The essential features of this form of scholarship include: (a) the internal perspective of its practitioners who address judges, lawyers and legal policy makers in the same manner and using the same terms which they themselves use to formulate the law;[57] (b) its focus on synthesis contingent upon the characterisation of the relationship between statutory material, judgments, regulatory guidance and soft law as a rational, albeit multi-layered system;[58] (c) due to the law's specialised idiom, its characterisation of that legal system as a self-referential system autonomous of other disciplines;[59] and (d) its emphasis on the elucidation of the ongoing evolution of the current law.[60]

With its major aims of interpretation and understanding of legal text, doctrinal legal research thus has analogical connections with hermeneutic disciplines such as history, philosophy, theology, and literature.[61] Like these disciplines, doctrinal legal research seeks to provide a connection between legal text and those who need to determine its meaning.

---

[54]   Jan M Smits, 'What is Legal Doctrine? On the Aims and Methods of Legal-Dogmatic Research' in Rob van Gestel, Hans W Micklitz and Edward L Rubin (eds), *Rethinking Legal Scholarship: A Transatlantic Dialogue* (Cambridge University Press, 2017) 207.

[55]   Susan Bartie, 'The Lingering Core of Legal Scholarship' (2010) 30(3) *Legal Studies* 345.

[56]   Terry Hutchinson and Nigel Duncan, 'Defining and Describing What We Do: Doctrinal Legal Research' (2012) 17(1) *Deakin Law Review* 83, 85.

[57]   Matyas Bodig, 'Legal Doctrinal Scholarship and Interdisciplinary Engagement' (2015) 8(2) *Erasmus Law Review* 43, 46. See also Edward L Rubin, 'The Practice and Discourse of Legal Scholarship' (1988) 86(8) *Michigan Law Review* 1835, 1847–8. Bodig argues that 'the most distinctive feature of standard legal scholarship is its prescriptive voice, its consciously declared desire to improve the performance of legal decisionmakers' and that the point of most doctrinal legal research is 'to remon-strate with the judge [or legislator] for the conclusion [or policy] reached and for the rationale adopted'.

[58]   Bodig (n 57) 45; Gunther Teubner, 'Substantive and Reflexive Elements in Modern Law' (1983) 17(2) *Law & Society Review* 239, 240.

[59]   Gunther Teubner, 'And God Laughed — Indeterminacy, Self-Reference and Paradox in Law' (2011) 12(1) *German Law Journal* 376, 377.

[60]   Smits (n 54) 207–9.

[61]   Mathias M Siems and Daithi Mac Sithigh, 'Mapping Legal Research' (2012) 71(3) *Cambridge Law Journal* 651, 654–5.

In some respects, the rationales of traditional doctrinal research and aspects of Regtech mirror each other. Traditional doctrinal research aims to navigate the legal system's complex thicket of multiple intersecting laws, each comprised of varying degrees of ambiguity, so as to provide greater clarity and hence greater certainty for the law and its subjects. It does so using organising principles that reflect historical pathways of legal development and classification as well as the development of new frameworks of analysis derived from emerging areas of the law.[62] Likewise, one of the functions of Regtech is to produce regulatory intelligence that collates and analyses large swathes of complex regulatory data to produce reliable and comprehensible information about the regulatory environment that can be used by compliance professionals to make better informed decisions. Nonetheless, compared with traditional doctrinal research, Regtech uses methods that differ in scale, complexity and approach.

When engaged in traditional doctrinal legal research, lawyers and legal scholars currently use a variety of online legal databases that provide up to date access to primary legal sources, precedents and informed commentaries. Some of these databases are jurisdiction-specific while others provide access to material across a variety of jurisdictions. Commonly, doctrinal legal research commences with a scenario, problem, or area of exploration that leads researchers to produce a number of potential research terms. These research terms are partly based on the factual features of the relevant matter and also draw upon the expertise and experience of the researcher. Once formulated, the research terms are put through the relevant online database search engine, and, in turn, generate research results categorised according to source — the case law, legislation and regulation of various jurisdictions, as well as secondary legal materials. Following analysis for the relevance and significance of these initial results, further searches may be undertaken to determine whether the initial results have been applied consistently or inconsistently to the scenario, problem or area of exploration, whether they have been extended to analogous situations, whether related search terms yield different results, or whether exceptions and caveats apply. Generally, the process of online legal research is iterative and non-linear. Depending on the subject matter of the research and the experience and expertise of the researcher, it is also likely to be time-consuming.

Doctrinal legal research does not end with the generation of results from these legal databases. Legal information retrieval is then followed by a synthesis and analysis of

---

62   Felix Frankfurter, 'The Conditions For, and the Aims and Methods of, Legal Research' (1930) 15(2) *Iowa Law Review* 129, 134: '[R]esearch requires the poetic quality of imagination that sees significance and relation where others are indifferent or find unrelatedness; the synthetic quality of fusing items theretofore in isolation; above all the prophetic quality of piercing the future, by knowing what questions to put and what direction to give to inquiry'. The Council of Australian Law Deans identified doctrinal research as 'key to understanding the mystique of the legal system's simultaneous achievement of constancy and change, especially in the growth and development of the common law': Council of Australian Law Deans, *Statement on the Nature of Legal Research* (Report, May and October 2005) 3 <https://cald.asn.au/resources/>.

the results using techniques of deductive logic and inductive reasoning to produce an opinion on the meaning of the law and its likely application.

Yet, even with the luxury of access to online legal databases that provide ease of searching through keyword-text concordance, some contend that the increasing volume and complexity of the law is making it more difficult for lawyers and legal scholars to undertake doctrinal legal research in a cost-effective and comprehensively accurate manner.[63] Using big data techniques and machine learning, Regtech applications can perform legal research tasks much faster, more accurately and at lower cost than human scholars and lawyers.[64] Indeed, there are many examples which illustrate the efficiency advantages of Regtech technologies. One such experiment conducted by LawGeex found that when reviewing contractual documents, artificial intelligence performed the review on average at 94% accuracy within 26 seconds, whereas experienced lawyers on average performed at 85% accuracy within 92 minutes.[65] An earlier experiment involving an algorithm based on a collection of 584 decisions was able to predict the outcomes of European Court of Justice decisions with 79% accuracy.[66] Meanwhile, in a 2018 experiment carried out between the Commonwealth Bank (Australia), ING (Netherlands), and the United Kingdom's Financial Conduct Authority, artificial intelligence was applied to the 1.5 million paragraphs of the European *Markets in Financial Instruments Directive II* to turn it into actionable compliance with 95% accuracy under two weeks, whereas it would normally take the Commonwealth Bank's compliance team six months to complete this task.[67]

To reduce the amount of time lawyers and legal scholars spend collecting and collating the results of legal research, publishers like LexisNexis and Thomson Reuters are already connecting their existing online databases with 'legal analytics' tools to discover insights such as connections between judges, parties and courts that might have previously been extremely time-consuming to find. This allows users to determine which cases are more influential than others, to quickly isolate the judicial treatment of words and phrases in visually attractive formats, to track and compare

---

63   Kenneth A Grady, 'Mining Legal Data: Collecting and Analyzing 21st Century Gold' (2017) 20(7) *Journal of Internet Law* 9.

64   Richard Susskind, *Tomorrow's Lawyers: An Introduction to Your Future* (Oxford University Press, 2nd ed, 2017) 186–7.

65   See Kyree Leary, 'The Verdict Is In: AI Outperforms Human Lawyers in Reviewing Legal Documents', *Futurism* (Blog Post, 27 February 2018) <https://futurism.com/ai-contracts-lawyers-lawgeex>.

66   Marie Boran, 'Artificial Intelligence Judges Court Cases with 79% Accuracy', *The Irish Times* (online, 27 October 2016) <https://www.irishtimes.com/business/technology/artificial-intelligence-judges-court-cases-with-79-accuracy-1.2842492>.

67   George Nott, 'CBA Reveals Promising Regtech AI Results, Proceeds with Caution', *Computerworld* (online, 7 May 2018) <https://www.computerworld.com.au/article/640828/cba-reveals-promising-regtech-ai-results-proceeds-caution/>.

regulatory changes over time, and to provide more responsive suggestions to legal queries.[68]

Beyond enhancing the efficiency of legal research tasks, Regtech may be used to undertake regulatory tracking and policy modelling to simulate the likely impact of various initiatives upon existing regulation as well as upon the relationship between the regulated and various regulators.[69] The application of artificial intelligence to large-scale contract review by international law firm, Pinsent Masons, to determine risks associated with Brexit is an example of such an approach.[70] London-based Regtech CUBE, a platform that provides business with advice regarding the impact of regulatory change on their internal policies and procedures, is another example.[71] To enrich regulatory impact analysis, it has also been suggested that Regtech might be applied to predict the importance of legislative initiatives to particular regulated segments of business.[72] Not only can Regtech identify the relevance of legislative text for certain areas or products, but it can also scale the importance of that relevance to particular firms.

One important way that Regtech adopts a different approach to traditional doctrinal legal research can be seen in its efforts to undertake advanced forms of content and taxonomic analysis of regulator, judicial, legislative and treaty artefacts of decision-making.[73] Using topic modelling techniques, analysts can evaluate a vast number of artefacts to ascertain the underlying factors that determine decision-making. By this means, for example, it is possible to identify the critical factors that will determine whether a court will undertake corporate veil piercing or impose corporate successor liability based on analysis of a dataset comprised of many thousands of judicial decisions.[74] Interestingly, in the former case, the relevant researchers found that many of the factors cited by the judiciary as of relevance to

---

68   See, eg, 'High Court Analyser', *LexisNexis* (Web Page, 2019) <https://www.
      lexisnexis.com.au/en/products-and-services/legal-analytics/high-court-analyser>;
      'Westlaw Edge', *Thomson Reuters* (Web Page) <https://legal.thomsonreuters.com/en/
      products/westlaw/edge>.

69   United Kingdom Chief Scientific Adviser (n 11) 49.

70   Pinsent Masons, 'Pinsent Masons Rolls Out AI for Brexit Challenge' (Media Release,
      7 November 2016) <https://www.pinsentmasons.com/en/media/press-releases/2016/
      pinsent-masons-rolls-out-ai-for-brexit-challenge/>.

71   'Solution', *CUBE* (Web Page) <https://www.cube.global/solution>.

72   Adedayo Banwo, 'Artificial Intelligence and Financial Service: Regulatory Tracking
      and Change Management' (2018) 10(4) *Journal of Securities Operation and Custody*
      354, 357.

73   Wolfgang Alschner, Julia Seiermann and Dmitriy Skougarevskiy, 'Text-as-Data
      Analysis of Preferential Trade Agreements: Mapping the PTA Landscape' (Working
      Paper No 2017–13, Centre for Trade and Economic Integration, 10 July 2017); Frank
      Fagan, 'Big Data Legal Scholarship: Toward a Research Program and Practitioner's
      Guide' (2016) 20(1) *Virginia Journal of Law & Technology* 1.

74   Fagan (n 73) 12.

corporate veil piercing, such as undercapitalisation, were simply not applied.[75] There are several advantages of this form of research over traditional doctrinal research: its scalability (a far greater volume of legal material can be analysed over a shorter time frame); its relative completeness (datasets are not limited by the capacity of researchers to physically collate and analyse them); its reduction of variable selection bias; and its relative independence from the subjectivity of researchers who are guided by the variables they wish to investigate. By applying these techniques, it is thus possible to better understand the unstated motivations and reasoning of regulators and lawmakers.

Nevertheless, while it may provide insights into regulatory trends that are otherwise difficult to ascertain, many doubt the capacity of Regtech to engage with the 'creative evolution of the legal system'.[76] They argue that big data and machine learning mine the past and make predictions based on averages without accounting for the 'lower-level variance or heterogeneity that makes legal systems adaptive and dynamic'.[77] As Searle has observed, these technologies are not designed to understand the meaning of the legal data they process.[78] Big data and machine learning methods conform to the precise instructions of their programmers[79] and largely deal with routine, structured data.[80] Therefore, they may find it difficult to take account of creative agents who continuously adapt their behaviour to take full advantage of the law and unforeseen changes in their respective environments. Big data and machine learning cannot effectively deal with legal ambiguity,[81] discern the difference between bright-line rules and fuzzy principles, or create novel approaches to legal situations. Moreover, these technologies are blind to the inherent value of the rule of law and other political and constitutional values that underpin legal development. Rather than being steeped in incremental adaptation anchored by reference to established but open-ended legal principles designed to do justice, Regtech is devoid of values.

---

75   Jonathan Macey and Joshua Mitts, 'Finding Order in the Morass: The Three Real Justifications for Piercing the Corporate Veil' (2014) 100(1) *Cornell Law Review* 99, 102.

76   Caryn Devins et al, 'The Law and Big Data' (2017) 27(2) *Cornell Journal of Law & Policy* 357, 359. See generally Dana Remus and Frank S Levy, 'Can Robots be Lawyers? Computers, Lawyers and the Practice of Law' (2017) 30(3) *Georgetown Journal of Legal Ethics* 501, 511.

77   Devins et al (n 76).

78   John Searle, 'Can Computers Think?' in David J Chalmers (ed), *Philosophy of Mind: Classical and Contemporary Readings* (Oxford University Press, 2002) 669, 671. Searle notes that digital technologies are built on syntax rather than semantics.

79   Saul Levmore and Frank Fagan, 'The Impact of Artificial Intelligence on Rules, Standards and Discretion' (2019) 93 *Southern California Law Review* (forthcoming) 11.

80   Remus and Levy (n 76) 509.

81   The CBA made a finding that the accuracy of AI analysis of legal material decreased from 95% to 63% when dealing with 'vague' parts of regulation: Nott (n 67). See also Kevin D Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* (Cambridge University Press, 2017) 45.

Consequently, while Regtech can considerably enhance legal information retrieval and analysis by deriving insights which volume and complexity render difficult for humans to achieve, at this stage in its development, Regtech can only work in tandem with, rather than completely replace, doctrinal scholarly analysis.[82] To date there still appears to be a significant gap between legal information retrieval and automated legal argumentation. Regtech itself does not understand the purposes driving regulation or the teleological concepts that bind particular provisions into a system of regulation.[83] On the other hand, high quality doctrinal research requires an ability to address indeterminacy, to creatively synthesise legal materials into a compelling narrative, to analogise and draw connections between areas of law, and to differentiate between important and unimportant precedent. Although we might expect that doctrinal research will become a more instrument-enabled form of scholarship, at this point Regtech falls short of these higher order tasks.

Nonetheless, this could change. Research is currently underway to exploit the ability of neural networks to generate legal argumentation and findings, and to employ argumentation mining and analysis to assist with producing doctrinal exegesis[84] that will eventually lead to fully automated interpretation and implementation within business systems. Once it becomes possible to fully translate regulation into machine-readable form, it will then be possible to ensure alignment between the interpreted provisions and compliance technology that will enable automated reporting and auditing of regulated activities.[85] If that eventuates, it does not appear to be a giant leap for algorithmic programs utilised in these tasks to learn more about legal reasoning and its potential scope of application.[86] As a result, the central importance of doctrinal research to legal scholarship may well fall away or at the very least become far more automated and data-driven, rather than expert-driven.

B *Normative/Reform-Oriented*

With the aim of explaining and critiquing the efficacy of current law, doctrinal research is often associated with legal positivism,[87] and therefore typically contrasted with normative analysis which focuses not upon the law as is, but upon what the law

---

82   Tania Sourdin, 'Judge v Robot: Artificial Intelligence and Judicial Decision-Making' (2018) 41(4) *University of New South Wales Law Journal* 1114; Tom Butler, Leona O'Brien and Marcello Ceci, 'Beyond the Hype of AI: A Smart Approach to Unpacking Regulations' (2017) 36(10) *Banking & Financial Services Policy Report* 1; Remus and Levy (n 76).

83   Ashley (n 81) 97.

84   Karl L Branting, 'Artificial Intelligence and the Law from a Research Perspective' (2018) 14(3) *Scitech Lawyer* 32; Ashley (n 81) ch 10.

85   Butler and O'Brien (n 50) 97.

86   Ashley (n 81) 383, applying Wittgenstein's view that meaning lies in use.

87   Mathias M Siems and Daithí Mac Síthigh, 'Why Do We Do What We Do? in Rob van Gestel, Hans-W Micklitz and Edward L Rubin (eds), *Rethinking Legal Scholarship: A Transatlantic Dialogue* (Cambridge University Press, 2017) 31, 68–71 [1.4.2.2].

should be.[88] However, in reality, doctrinal analysis and normative prescription are inextricably connected, either covertly by mute acceptance of existing law, or overtly through critique and arguments in favour of the reform of legal institutional settings, legal policy implementation, and/or legal decision-making. At the very least, doctrinal coherence and acceptability are critical to determining whether policies will be successfully translated into law and faithfully implemented by legal decision-makers.[89] More generally, it is hard to imagine modern legal research scholarship without an analysis of the values inherent in the law which, by definition, promulgates policy choices and behavioural norms.[90] Arguably, legal rules and principles are only capable of intelligible doctrinal exposition in light of their underlying aims and objectives which, in turn, unavoidably incorporate social, economic or political ends.

However, it seems likely that Regtech will help to drive a shift away from manual forms of doctrinal legal research toward more empirically-based normative analysis. By providing greater insight into the efficacy of regulator activity, it will become easier to analyse the efficiency and fairness of the law and regulation. From an efficacy perspective, there may be many hidden biases at work embedded in regulatory text or regulatory behaviours that operate to undermine policy objectives. Using interconnected data sets comprised of over 380,000 judicial decisions, hundreds of judicial biographical records, speech pattern analysis of judicial decision-making, and information regarding judicial professional networks, for example, researchers have demonstrated how judicial analytics can be used to assess the psychological, political and economic factors that can influence judicial decisions.[91] One startling finding, is that judges are more lenient toward asylum seekers before lunch and towards the end of the day than at other times when orders are made.[92] Other research has shown that judges' thinking about law and economics is strongly predictive of harshness in sentencing.[93] In each case, the researchers' aim was to produce analysis for the purpose of generating more even-handed decision-making, de-biasing the law. The same techniques can be applied to regulatory activity.

Exploring the normative aspects of regulation will become even more critical as compliance by design is gradually embedded between businesses and regulators. Currently, Regtech is only capable of 'providing partial reasoning and modelling

---

[88]   Joshua B Fischman, 'Reuniting Is and Ought in Empirical Legal Scholarship' (2013) 162(1) *University of Pennsylvania Law Review* 117, 119–20; Richard A Posner, 'The Present Situation in Legal Scholarship' (1980) 90(5) *Yale Law Journal* 1113, 1119.

[89]   Bodig (n 57) 46.

[90]   Smits (n 54) 215.

[91]   Daniel L Chen, 'Judicial Analytics and The Great Transformation of American Law' (2019) 27(1) *Artificial Intelligence and the Law* 15.

[92]   Ibid 27–8.

[93]   Elliott Ash, Daniel L Chen and Suresh Naidu, 'Ideas Have Consequences: The Impact of Law and Economics on American Justice' (Working Paper, National Bureau of Economic Research, 2 November 2017) <https://users.nber.org/~dlchen/papers/Ideas_Have_Consequences.pdf>.

support' for the digital expression of legal norms.[94] As a result of its lack of conceptual connection to the legal domain, insofar as multi-layered and open-ended legal rules are concerned, Regtech produces conflicting and inaccurate compliance verification. Accordingly, 'adopting formalisms that are not conceptually grounded in legal practice creates a framework that is unreliable, and not suitable to be used in real-life applications'.[95] An example of the types of normative problems that can arise in Regtech when dealing with something as relatively straightforward as 'know your customer' identification and verification, relates to the impact that such procedures may have on financial inclusion, particularly for people from socially disadvantaged communities.[96] Whether a financial institution ought to incorporate financial inclusion measures into its 'know your customer' procedures is an ethical and social question that simply cannot be answered by current natural language processing applications. If Regtech is to become a practical reality, new modelling languages must be built in conjunction with specialist legal scholars who can articulate the normative dimensions of legal text and assist computational experts to express these faithfully.

Otherwise from a normative perspective, there is clearly great scope to research and analyse the legal, ethical and governance aspects of Regtech. Worldwide legal scholars, governments, businesses and regulators are discussing what kinds of legal frameworks may be required to govern the use of Regtech technologies especially those related to big data, machine learning and automated decision-making.[97] There are a myriad of concerns: rights of due process, the embedding of conscious and unconscious bias in decision-making, the potentially poor quality of algorithmic decision-making, overbearing surveillance, and the lack of democratic accountability.[98]

C *Interdisciplinary*

The decline of legal scholarship as a purely autonomous discipline has coincided with the falling popularity of doctrinal legal research among legal scholars.[99] Doctrinal

---

[94]   Mustafa Hashimi and Guido Governatori, 'Norms Modeling Constructs of Business Process Compliance Management Frameworks: A Conceptual Evaluation' (2018) 26(3) *Artificial Intelligence and the Law* 251, 295.

[95]   Ibid 298.

[96]   Pompeu Casanovas, Jorge Gonzàlez-Conejero and Louis De Koker, 'Legal Compliance by Design (LCbD) and through Design (LCtD): A Preliminary Survey' (Conference Paper, Workshop on Technologies for Regulatory Compliance Conference, 13 December 2017) 46.

[97]   Corrine Cath, 'Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges' (2018) 376 (2133) *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 1*; Yeung (n 22).

[98]   Yeung (n 22).

[99]   Richard A Posner, 'The Decline of Law as an Autonomous Discipline: 1962–1987' (1987) 100(4) *Harvard Law Review* 76. Posner charts and applauds the shift toward interdisciplinary scholarship.

research in law nowadays is frequently accompanied by research about the socio-eco-nomic impact of law, extra-legal factors affecting legal actors and regulators, how law controls access to essential services, and how law allocates power and responsibility across society. As a result, interdisciplinary research methods are also inextricably linked with normative reasoning.

By being able to connect many more data points regarding law, behaviour and socio-economic phenomena, there appears to be little doubt that the advent of Regtech will advance the capacity of legal scholars to successfully undertake inter-disciplinary research.[100] First, as Regtech methods analyse huge volumes of data from interconnected datasets without the constraint of a theoretical or disciplinary frame, it is much more likely that the patterns and correlations that emerge from Regtech analysis will require interdisciplinary expertise to understand and address. Second, the use of Regtech data related to the same complex phenomena using the same analytical methods will likely lend itself to interdisciplinary collaboration. The work referred to earlier analysing how non-legal factors influence judicial decision-making demonstrates how legal scholars may need to draw upon a variety of different disciplines and interact with scholars from different backgrounds to make the most of the predictive analysis generated by big data. Among other things, the judicial analytics undertaken by the researchers draw upon psychological theories related to gamblers' fallacy and implicit egoism, political theories regarding the impact of the electoral cycle, and linguistic analysis of judicial speech patterns.

Furthermore, there are many aspects of compliance that cannot be fully captured quantitatively. Qualitative research methods used by social scientists are likely to be required to supplement and contextualise quantitative findings so as to identify whether a regulated business not only complies with the letter of regulation but also fosters an ethical and socially responsible culture.[101] There are an array of regulatory responses available to regulated entities ranging from minimal compliance that merely attempts to pass regulator review to a robust level of compliance that is morally defensible. Regtech alone cannot determine where in that range a business ought to sit. Alternatively, unless a business' culture effectively supports algorithmic regulation, there are significant risks that Regtech may be manipulated in a manner that encourages lip-service to, or outright evasion of, regulatory obligations.[102]

Interdisciplinary research will also be essential for the development and critical evaluation of computational legal reasoning modelling and its implementation. As noted above, there is a great deal of work to be done in specifying how regulatory requirements can be expressed in computational language that will facilitate automated analysis and eventually lead to automatically executed regulatory compliance. To ensure that compliance does not simply become an automated box ticking exercise but that it is fully understood and assimilated into business governance and processes,

---

[100]   Nicola Lettieri et al, 'Ex Machina: Analytical Platforms, Law and the Challenges of Computational Legal Science' (2018) 10(5) *Future Internet* 37, [6.4].

[101]   Packin (n 2) 214.

[102]   Ibid.

this work will require an integrated approach by legal scholars, lawyers, business information systems experts, governance and ethics advisers, computational experts, and organisational management experts. Compliance does not simply execute actions based on text. It necessarily incorporates conceptual considerations drawn from several sources that are not just data points but are informed by (sometimes conflicting) policies which of themselves draw on the expertise of many disciplines.[103] Although currently the computer science, engineering, and business and economics disciplines are the major leaders in interdisciplinary research projects that critically explore big data and machine learning,[104] it is anticipated that as Regtech matures, the legal discipline will begin to play a more significant role in these projects.

D *Theoretical*

Theory building and theory testing are well-known scholarly activities. Theory building attempts to advance knowledge by explaining the relationship between cause and effect in a new or different light. By contrast, theory testing involves the evaluation of a theory or theoretical model against empirical evidence that either supports or refutes the theory.[105] Theory building thus comprises the construction of an explanation of experiences and phenomena, whereas theory testing involves the collection and analysis of data about those experiences and phenomena.

Legal scholarship is replete with examples of theory building derived from a mix of interdisciplinary knowledge and traditional doctrinal research. Examples include the relational contract theories of Stewart Macaulay and Ian Macneil,[106] Guido Calabresi and Douglas Melamed's theories about liability rights and property rights,[107] Catherine McKinnon's feminist jurisprudence,[108] and John Braithwaite's theory of responsive regulation.[109] The value of these and other legally related theories lies in their capacity to explain and predict legal development and the way in which the law

---

[103] Casanovas, Gonzàlez-Conejero and De Koker (n 96).

[104] Jiming Hu and Yin Zhang, 'Discovering the Interdisciplinary Nature of Big Data Research Through Social Network Analysis and Visualization' (2017) 112(1) *Scientometrics* 91.

[105] Jason A Colquitt and Cindy P Zapata-Phelan, 'Trends in Theory Building and Theory Testing: A Five-Decade Study of the "Academy of Management Journal"' (2007) 50(6) *Academy of Management Journal* 1281, 1283–5.

[106] Stewart Macaulay, *Law and the Balance of Power: The Automobile Manufacturers and Their Dealers* (Russell Sage Foundation, 1966); Ian R Macneil, *The New Social Contract: An Inquiry into Modern Contractual Relations* (Yale University Press, 1980).

[107] Guido Calabresi and A Douglas Melamed, 'Property Rules, Liability Rules, and Inalienability: One View of the Cathedral' (1972) 85(6) *Harvard Law Review* 1089.

[108] Catharine A MacKinnon, 'Feminism, Marxism, Method, and the State: Toward Feminist Jurisprudence' (1983) 8(4) *Signs* 635.

[109] John Braithwaite, *Restorative Justice and Responsive Regulation* (Oxford University Press, 2002).

shapes behaviour. They can be used to explain existing law or to justify a future legal intervention.

Theory building comprises an iterative process of domain choice, data collection, description, concept specification, premise identification, synthesis of thematic patterns, and correlative and causal analysis.[110] However, as many of these functions can be performed using big data and machine learning, it has been suggested that automated data mining will lead to new discoveries rendering theory building obsolete.[111] By parsing thousands if not millions of legal materials, it is envisaged that Regtech techniques will be able to produce predictions and generate hypotheses of how laws will impact behaviour and vice versa bereft of any theory. Big data and machine learning can do this using probabilistic determinism.[112]

It is true that Regtech can engage in predictive analysis based on probabilities inherent in past data, categorise data into groups and subgroups, make connections and map the relationship between data points, undertake similarity and divergence measures, and even simulate how parties may behave under certain parameters. Neverthe-less, despite those capabilities, it does have epistemological limitations.[113] The first relates to the nature of the legal hypotheses that might be generated by predictive analytics. The types of hypotheses that might currently be generated by Regtech are not comparable to grand theories about the law like Calabresi and Melamed's theory of liability or Braithwaite's theory of responsive regulation. Hypotheses generated by Regtech applications are currently constrained by their capacity to reduce the relevant legal text to logic rules that reflect technical terms of quite specific appli-cation. For example, if we return to the example concerning the interconnection between BIM and legal text postulated early on in this article, we can see that simply applying closed technical legal terms in a building code to building design is unlikely to generate any theory about regulation or regulatory behaviour or even any insight as to the ability of the building code to foster safe and sustainable building practices. In other words, Regtech alone cannot extrapolate beyond its limited logic rule domain.

The second problem relates to the material produced by Regtech. As we have noted earlier, Regtech can generate correlations between data and reveal patterns that are

---

[110]   Anne Galletta, *Mastering the Semi-Structured Interview and Beyond: From Research Design to Analysis and Publication* (New York University Press, 2017) ch 5.

[111]   Ahmad Elragal and Ralf Klischewski, 'Theory-Driven or Process-Driven Prediction? Epistemological Challenges of Big Data Analytics' (2017) 4(1) *Journal of Big Data* 19; Fulvio Mazzocchi, 'Could Big Data be the End of Theory in Science? A Few Remarks on the Epistemology of Data-Driven Science (2015) 16(10) *European Molecular Biology Organization Report* 1250; Janet Chan and Lyria Bennett Moses, 'Is Big Data Challenging Criminology?' (2016) 20(1) *Theoretical Criminology* 21, 28–31.

[112]   Probabilistic determinism aims to establish a relationship between cause and effect using the tools of probability theory. See further Zoubin Ghahramani, 'Probabilistic Machine Learning and Artificial Intelligence' (2015) 521(7553) *Nature* 452.

[113]   Elragal and Klischewski (n 111).

not otherwise ascertainable. However, correlations can be arbitrary and are meaning-less without the theoretical inquiry as to why they may have arisen.

Furthermore, Regtech does not operate in a jurisprudential vacuum. Regtech identifies and applies the meaning of legal text expressed as formal logic rules based upon previous findings and the judgement of legal experts. Data, including regulatory data produced by Regtech, is therefore not neutral, but the product of the particular technology and platforms used to create it as well as the organisational, legal and social practices of those that compile it.[114] Consequently, Regtech-produced data will inevitably be shaped by the purposes of those seeking to utilise it. Accordingly, understanding the potential for human bias and framing within Regtech is essential for any critical inquiry of the data it produces.

Ultimately, Regtech is a tool that can be used to make compliance more responsive and more efficient. It is also a tool that can produce data that informs but does not replace theoretical inquiry about regulation and regulatory behaviour. While it may display a preference towards inductive methods over deductive reasoning, in practice, it will not replace the need to develop causal explanations of how the law works or understandings of the law in context.

In any event, the idea that theory building is dead is not sustainable. At the very least one would expect that Regtech would inform theories about regulation itself. A good example is the work of Karen Yeung.[115] Yeung has developed a taxonomy of algorithmic regulation that explores how regulation operates as a means of social ordering. Yeung's taxonomy considers algorithmic regulation as part of the following forms: outcome-based regulation; data-driven performance management; pre-emptive risk-based regulation; actuarial justice; and surveillant-driven social sorting. Yeung explores the drivers of algorithmic regulation and postulates upon its impact on democracy and decision-making accountability.

E *Empirical Legal Scholarship*

The growth in volume and impact of empirical research has been described as the most important development in legal scholarship in the last 25 years.[116] However, like many other fields of legal scholarship, there is no universally accepted definition of this category of legal research. Some confine empirical legal scholarship to

---

114   Rob Kitchin, 'Big Data, New Epistemologies and Paradigm Shifts' (2014) 1(1) *Big Data & Society* 1, 4–5.

115   Yeung (n 22).

116   Robert L Fischman and Lydia Barbash-Riley, 'Empirical Environmental Scholar-ship' (2018) 44(4) *Ecology Law Quarterly* 767, 768. See also Shari Seidman Diamond and Pam Mueller, 'Empirical Legal Scholarship in Law Reviews' (2010) 6(1) *Annual Review of Law and Social Science* 581; Michael Heise, 'An Empirical Analysis of Empirical Legal Scholarship Production 1990–2009' [2011] (5) *University of Illinois Law Review* 1739. Heise's article documents the growth in empirical legal scholarship appearing in law reviews.

quantitative studies that test falsifiable hypotheses, whereas others take a broader view proffering a definition more consistent with the concept of systematic observation and analysis.[117] Most agree that empirical legal scholarship draws largely from the methods and theories of the social sciences.[118]

Applying big data and machine learning to legal materials is another means of engaging in empirical research. As outlined previously, by turning regulation into data, it is possible to see hidden patterns and correlations between legal materials and to develop hypotheses derived from those correlations and patterns.

Regtech promises to enhance empirical legal research in a number of ways. First, the availability of data is a major problem in undertaking empirical work. Although researchers may be able to freely access primary materials such as cases, legislation, regulation, regulatory guidance and some commentaries, data on regulatory efficacy and behaviour is much more difficult to obtain. Outside of the field of criminology, statistics on compliance, the regulatory process, and regulatory enforcement are limited. If data is not publicly available, researchers are required to identify and collect it themselves largely through qualitative methods. Consequently, much data sits in researchers' own curated datasets or in proprietary datasets that are expensive for other individual researchers to access, and which are therefore difficult to validate.

While regulators publish annual reports that outline their regulatory activities, these annual reports normally do not provide researchers with the type of granularity that raw data provides, nor are the reports intended to constitute research data. They are written to fulfil public accountability responsibilities. Recent research apropos ASIC enforcement activity undertaken by Ramsay and Webster illustrates the current regulatory data deficit.[119] Despite the fact that the researchers were engaged in ground breaking work documenting how ASIC discharges its regulatory mandate, their work was considerably hampered. This was due to the absence of precise identification of the actual misconduct activity being studied, the absence of reference to specific legislative provisions being enforced, and the failure to discriminate between enforcement outcomes where multiple offences or multiple offenders were involved. Other problems identified by the researchers included issues related to multiple and/ or incorrect reporting of enforcement action, guilty pleas, and imposition of civil and criminal penalties concerning the same offenders.[120] Most of all, the researchers were limited in only being able to access data about enforcement from the regulator. No data was available to them regarding the effect of ASIC's enforcement on the regulated parties or more generally upon compliance across the marketplace. In fact,

---

117   Diamond and Mueller (n 116) 582–3; Lee Epstein and Gary King, 'The Rules of Inference' (2002) 69(1) *University of Chicago Law Review* 1, 2–3.

118   Deborah R Hensler and Matthew A Gasperetti, 'The Role of Empirical Legal Studies in Legal Scholarship, Legal Education and Policy Making: A US Perspective' in Gestel, Micklitz and Rubin (n 87) 450.

119   Ian Ramsay and Miranda Webster, 'ASIC Enforcement Outcomes: Trends and Analysis' (2017) 35(5) *Company and Securities Law Journal* 289.

120   Ibid 292–4.

any information about how a particular regulated business or segment of business might elect to embed compliance into their business operations and processes was distinctly lacking. As a result, we have only limited understanding of the efficacy of ASIC's regulatory role.

By embedding automated reporting and real-time monitoring using artificial intelligence, Regtech will produce much more information about regulation than is currently available to researchers and in a machine-readable form which will be easier to mine and analyse. Given that there has been a global trend toward open data in government,[121] which will make commercially sensitive information related to regulation publicly accessible in de-identified form, there appears to be a rich harvest of potential material available. Assuming that Regtech data will be made available, legal researchers with an interest in how regulation works can thus progress to more dynamic, data-rich research settings that, in turn, will enable the development of more innovative, more complex and more advanced models of regulatory behaviour.[122]

Additionally, theories about regulatory behaviour can be more effectively tested using data modelling and simulation modelling applications associated with big data and machine learning unavailable to researchers using traditional research methods. Research looking at discriminatory pricing for mortgage products that set about to 'stress-test' whether algorithmic pricing would produce less discriminatory outcomes than human-generated pricing constitutes a good illustration of this type of virtual experimentation. Interestingly, in this instance, the researchers found that simply restricting an algorithm from using discriminatory information such as race, national origin, religion, or disability status only satisfied the letter of legal requirements and would not necessarily result in non-discriminatory practices in financial services.[123]

All in all, Regtech is likely to be a boon to legal scholars seeking to test theory empirically and develop normative positions regarding the rationale and efficacy of regulatory activity.

III CONCLUSION

Regtech is bound to have a significant impact on legal scholarship. To the extent that doctrinal legal research can be automated, Regtech will enhance research efficiency and will signal a reduction in the human effort required for its production. Regtech will be able to generate predictive hypotheses about the regulatory process and

---

[121] Yannis Charalabidis et al, *The World of Open Data: Concepts, Methods, Tools and Experiences*, (Springer, 2018) 1–2 [1.1]. See also Australian Government, *Australian Government Public Data* (Policy Statement, 7 December 2015) <https://www.dta.gov. au/help-and-advice/guides-and-tools/requirements-australian-government-websites/ open-data>.

[122] Kitchin (n 114) 10.

[123] Talia B Gillis and Jann L Spiess, 'Big Data and Discrimination' (2019) 86(2) *University of Chicago Law Review* 459.

impact, and will also be able to generate underlying materials that help support the hypotheses more quickly and accurately than humans. However, reducing large and complex regulatory language to formal logic that can be used to automate compliance poses considerable practical difficulties. To date, Regtech is limited in its capacity to deal with ambiguous legal text and doubts have been raised about its capacity to draw analogies across legal areas and to synthesise findings creatively into a compelling narrative that fully captures regulatory complexity. Consequently, doctrinal research will still be a significant, albeit diminishing, field of legal scholarship.

Regtech will likely increase our capacity to develop and test theories and models about regulation and regulatory behaviour. It will also lead to greater degrees of collaboration between legal scholars and scholars from other disciplines as the correlations and patterns produced by Regtech will require interdisciplinary expertise to interpret and analyse. As a result, as more data is generated by Regtech and more is revealed about regulation and regulatory behaviour, Regtech will also increase confidence in normative aspects of legal scholarship.