15

COMPLIANCE AND WHISTLEBLOWING: HOW TECHNOLOGY WILL REPLACE, EMPOWER AND CHANGE WHISTLEBLOWERS

Kieran Pender, Sofya Cherkasova and Anna Yamaoka-Enkerlin[1]

A. INTRODUCTION

15.001 Whistleblowing is not a new phenomenon. Some scholars have traced the concept to Ancient Greece, drawing parallels with the notion of *parrhēsia*, or fearless speech.[2] Lykourgos, an Athenian orator of the mid-300s BC, is

1 The authors acknowledge with thanks the helpful comments of Ashley Savage and Jelena Madir on an earlier draft. Thank are also owed to Maria Shepard and Emma Franklin for their assistance with research for the second edition of this chapter.

2 Michel Focault, *Fearless Speech*, Semiotext(e) (2001); and Alan Chu, In Tradition of Speaking Fearlessly: Locating a Rhetoric of Whistleblowing in the Parrhēsiastic Dialectic, 19 *Advances in the History of Rhetoric* 231 (2016), at 239–48.

reported to have said that 'neither the laws nor judges can bring any results unless someone denounces the wrong doers'.[3] Laws to incentivise whistleblowers are not novel either. In the 7th century, a British king declared that 'if a freeman works during [the Sabbath], he shall forfeit his [profits], and the man who informs against him shall have half the fine, and [the profits] of the labour'.[4] Modern American whistleblower protections, meanwhile, are grounded in the Civil War-era False Claims Act.

15.002 The label whistleblowing, on the other hand, is a more recent invention. The term was popularised in the 1970s by American political activist Ralph Nader, who described it as 'an act of a man or a woman who, believing that the public interest overrides the interest of the organisation he serves, publicly "blows the whistle" if the organisation is involved in corrupt, illegal, fraudulent or harmful activity'.[5] In the following decades, whistleblowing entered the mainstream lexicon. High-profile whistleblowers drew attention to the considerable public interest in their deeds and the adverse consequences they often suffered. Laws were enacted to encourage whistleblowing and protect those who did so – the Public Interest Disclosure Act 1998, a notable early example in the UK – and charitable organisations were established to advocate the whistleblower cause. There remains, though, no universally-accepted definition of a whistleblower or defined criteria of what constitutes whistleblowing.

15.003 Although progress has been slow, by 2020 whistleblowing is beginning to lose the societal stigma often attached to it. Whistleblowers have revealed large-scale tax avoidance (Antoine Deltour and LuxLeaks), widespread data misuse (Christopher Wylie and Cambridge Analytica) and multi-billion pound money laundering (Howard Wilkinson and Danske Bank). The COVID-19 pandemic has accelerated this trend. As governments are fast-tracking relief fund distribution, mass public health procurement, and data collection and surveillance efforts, the strain of an economic downturn is pressuring businesses and individuals to meet goals at all costs. The result: a 'perfect storm for fraud, wrongdoing, and corruption, as well as denunciation and whistleblowing'.[6] Healthcare worker whistleblowers like Dr. Li Wenliang and Dr. Ai Fen were

3 Transparency International, Providing an Alternative to Silence (2013), available at: http://www.transparency.gr/wp-content/uploads/2013/11/WHISTLEBLOWERS_ENGLISH_LOW.pdf, at 13.

4 International Bar Association, Whistleblower Protections: A Guide (2018), available at: https://www.ibanet.org/Conferences/whistleblowing.aspx at 5.

5 Ralph Nader, An Anatomy of Whistle Blowing, in Ralph Nader, Peter J Petkas and Kate Blackwell (eds) *Whistle Blowing* (1972), at vii.

6 Meyr et al., Whistleblowing in a Global Pandemic: Are You Ready? (2020) available at: https://www.mccarthy.ca/en/insights/articles/whistleblowing-global-pandemic-are-you-ready.

among the first to raise the alarm about the gravity of COVID-19.[7] Since then, whistleblowers have exposed unsafe work conditions in hospitals, schools, and businesses, abuses of privacy, government misspending and other wrongdoing. The intimidation, lawsuits and job-loss that many faced as a result has led to renewed calls for global authorities to strengthen whistleblower protections.[8] In recent years Ireland, the Netherlands, France, Italy and Serbia have been among the jurisdictions to pass landmark whistleblower protection regimes – just under 50 countries globally now have specific legal protections for those who blow the whistle. In 2019, the European Union passed a landmark whistleblower protection directive requiring all member states to introduce best practice laws in the coming years.

15.004 But just as societies are beginning to appreciate the significant contributions whistleblowers make to democratic accountability and corporate compliance, the concept of whistleblowing is being transformed by technology. Such disruption is the focus of this chapter. Technology offers much promise to whistleblowers and whistleblower protections, but also many potential pitfalls. A sober analysis is required to determine where technology might add benefit, and where it could prove problematic.

15.005 This chapter has three substantive parts. Section B will begin by detailing how data analytics and artificial intelligence (AI) are becoming 'algorithmic whistleblowers', detecting (or even preventing) misconduct before it can be discovered by human whistleblowers. In other words, how is technology replacing whistleblowers? Section C analyses a range of technological solutions that are helping to empower whistleblowers, by protecting them and giving them new avenues for reporting misconduct. The exciting potential for blockchain to offer anonymity, immutability, resilience, compensation and information escrow in the whistleblowing context will be considered. Section D discusses how technology is changing the nature of whistleblowing, given the modern-day ability to distribute terabytes of information with a few clicks. The implications of technological-driven change, including the increasingly

7 Nie, Jing-Bao, and Carl Elliott. Humiliating Whistle-Blowers: Li Wenliang, the Response to Covid-19, and the Call for a Decent Society. *Journal of Bioethical Inquiry*, 1–5. 25 Aug. 2020, doi:10.1007/s11673-020-09990-x.

8 International Bar Association, Authorities Urged to Protect Whistleblowers during Covid-19 crisis (2020) https://www.ibanet.org/Article/NewDetail.aspx?ArticleUid=76b74307-2379-4f3d-93b1-a8d2b963338b; Transparency International, Governments and corporations need to guarantee safety of covid-19 whistleblowers (2020). https://www.transparency.org/en/press/governments-and-corporations-need-to-guarantee-safety-of-covid-19-whistleblowers;

blurred lines between whistleblowing, leaking and hacking, are reflected upon. Section E concludes.

B. TECHNOLOGY REPLACING THE WHISTLEBLOWER

15.006 We live in the age of information. An unprecedented level of openness is demanded from institutions, and the link between data disclosure, transparency and accountability is often assumed.[9] However, disruptive technology is increasingly becoming necessary to harness the power of this data to detect and minimise misconduct in the public and private sector. As this technology becomes more widespread, it is reducing the need to rely on human whistleblowers. This is positive not only from a compliance and supervisory perspective, but also in minimising the personal, financial and professional toll on would-be whistleblowers.[10] However, the benefits of 'algorithmic whistleblowing' must be weighed against risk, requiring trade-offs between transparency, privacy and accuracy.

1. The 'openness revolution'

15.007 Transparency is a contested concept, but at its heart it 'refers to the notion that information about an individual or organisation's actions can be seen from the outside'.[11] Transparency has become a default policy prescription,[12] often invoked as an essential component of trust and cooperation, as a market efficiency mechanism, as a legitimising procedural tool and, at its broadest, as a value embedded in democracy.

15.008 The ascent of transparency as an institutional norm dovetails with the growing recognition of the value of whistleblowers as 'the primary source of involuntary transparency',[13] reflected in the development of whistleblower protection legislation worldwide.[14]

---

9 L. Carolan, Open data, transparency and accountability: Topic guide (2016) https://assets.publishing.service.gov.uk/media/5857fdcb40f0b60e4a0000d6/OpenDataTA_GSDRC.pdf at 5.

10 Adam Waytz, Why Robots Could be Awesome Whistleblowers (2014), available at: https://www.theatlantic.com/business/archive/2014/10/why-robots-could-be-awesome-whistleblowers/381216/.

11 Matthew S. Mayernik, Open Data: Accountability and Transparency, *Big Data and Society* 1 (2017), available at: https://doi.org/10.1177/2053951717718853, at 1.

12 Aarti Gupta, Transparency Under Scrutiny: Information Disclosure in Global Environmental Governance, 8 *Global Environmental Politics* 1 (2008), at 1.

13 Jennifer Shkabatur, Transparency With(out) Accountability: Open Government in the United States 31 *Policy Review* 89 (2012), available at: https://digitalcommons.law.yale.edu/cgi/, at 113.

14 International Bar Association, *supra* note 4.

15.009 Meanwhile, regulators and advocates alike are increasingly pursuing another transparency frontier which has, along with whistleblowing, grown from the historical right to information movement: access to data.[15] Open data proponents support the disclosure of data in a way that allows it to be freely used, modified and shared by anyone for any purpose. The OECD identifies open data as a 'key public good' and a powerful tool in the fight against the abuse of power.[16] Efforts are being made to open up government data sets which include public officials' directories, budgets, public procurement, political financing, voting records and land registries.[17]

15.010 The 'openness revolution' is also marching into the private sector.[18] For example, the movement pushing for a global public database featuring country-by-country reporting (CBCR)[19] on the economic activity and tax contributions of multinational corporations achieved a breakthrough in 2017, when the European Commission voted for the second time in favour of public CBCR by multinationals.[20] Another example is the call for beneficial ownership reporting.[21] In 2016, the UK became the first country to publish the identity of those who benefit from, own and control companies;[22] and in April 2020, 18 countries were found to have public beneficial ownership registers.[23] Meanwhile, navigating the regulatory complexity that followed the Global Financial Crisis has 'inevitably required greater granularity, precision and frequency in data reporting, aggregation, and analysis' from corporations,

---

15 Katleen Janssen, Open Government Data: Right to Information 2.0 or its Rollback Version? 8 ICRI Research Paper (2012), available at: https://ssrn.com/abstract=2152566 at 4–8.

16 OECD, Compendium of Good Practices on the use of Open Data for Anti-corruption (2017), available at: http://www.oecd.org/gov/digital-government/g20-oecd-compendium.pdf.

17 World Wide Web Foundation and Transparency International, *Connecting the Dots: Building a Case for Open Data to Fight Corruption* (2017), available at: http://webfoundation.org/docs/2017/04/2017_OpenDataConnectingDots_EN-6.pdf.

18 The Openness Revolution (2014), *The Economist,* available at: www.economist.com/business/2014/12/11/the-openness-revolution.

19 Since 2002, CBCR has become the extractive industry standard in more than countries, and has since spread to the financial institutions; Alex Cobham et al., What Do They Pay? Towards a Public Database to Account for the Economic Activities and Tax Contributions of Multinational Corporations (2017), available at: datafortaxjustice.net/what-do-they-pay/#extractive-industries-data.

20 Financial Transparency Coalition, Letting the Public In (2015), available at: https://financialtransparency.org/wp-content/uploads/2016/09/OpenData_fullpaper.pdf.

21 According to the World Bank, up to 70 per cent of cases of financial misconduct involve anonymous companies. Open Ownership: *Ending anonymous company ownership*, available at: https://openownership.org/.

22 Jonathan Grey and Timothy Glyn Davies, Fighting Phantom Firms in the UK: From Opening Up Datasets to Reshaping Data Infrastructures*?* (2015), available at: doi:10.2139/ssrn.2610937.

23 U4 Anti-Corruption Resource Centre and Transparency International, Beneficial ownership registers: Progress to date, available at: https://knowledgehub.transparency.org/assets/uploads/helpdesk/Beneficial-ownership-registers_2020_PR.pdf annex 1, at 17.

financial institutions and supervisory authorities alike.[24] Transparency and efficiency in taxation systems are likely to become all the more important after the COVID-19 crisis.[25]

15.011 In summary, the demand for open data coupled with an increase in data intensive regulation is adding unprecedented dimensionality to high-volume, high-velocity and high-variety information assets, also known as big data.[26] At the outset, this might appear to be an unreservedly good development from transparency and anti-corruption perspectives, reducing reliance on human whistleblowers. However, there are at least three potential limitations.

2. Challenges

15.012 First, more data does not necessarily mean more transparency. Jonathan Fox distinguishes between two kinds of transparency.[27] Opaque transparency involves 'the dissemination of information that does not reveal how institutions actually behave in practice, whether in terms of how they make decisions, or the results of their actions', while clear transparency 'sheds light on institutional behaviour permit[ting] interested parties to pursue strategies of constructive change'. Despite a tendency to equate more data with more transparency, clear transparency necessitates not *data* per se, but the ability to extract relevant *information* about the entity in question from that data.[28] Even putting aside data quality issues,[29] the sheer volume of potentially available data[30] and a dearth of data literacy[31] among the general population makes actualisation of the average citizen as auditor doubtful.

---

24 Douglas W. Arner et al., FinTech, RegTech, and the Reconceptualization of Financial Regulation 37(3) *Northwestern Journal of International Law and Business* (2017), available at: https://scholarlycommons.law.northwestern.edu/njilb/vol37/iss3/2, at 388.

25 Bob can der Made, European Union: The Revival of Public CbCR Amid New interest in ESG Transparency (2020), available at: https://www.internationaltaxreview.com/article/b1kzwk564ld7k3/european-union-the-revival-of-public-cbcr-amid-new-interest-in-esg-transparency.

26 Doug Laney, 3D Management: Controlling Data Volume, Velocity, and Variety, *Gartner* (2001), available at: https://blogs.gartner.com/doug-laney/files/2012/01/.

27 Jonathan Fox, The Uncertain Relationship between Transparency and Accountability, 663 *Development in Practice* (2007), available at: https://doi.org/10.1080/09614520701469955, at 667.

28 Catharina Lindstedt and Daniel Naurin, Transparency is not Enough: Making Transparency Effective in Reducing Corruption, *International Political Science Review* (2010), available at: https://journals.sagepub.com/doi/abs/10.1177/0192512110377602, at 302.

29 Open Knowledge International Blog: *Open Data Quality – the Next Shift in Open Data?* (2017), available at: https://blog.okfn.org/2017/05/31/open-data-quality-the-next-shift-in-open-data/.

30 Of all data existing in 2018, 90 per cent was created in 2016-2018, amounting to 2.5 quintillion bytes of data created per day. See Domo, Data Never Sleeps 5.0 (2018), available at: https://www.domo.com/learn/data-never-sleeps-5.

31 Annika Woolf et al., Creating an Understanding of Data Literacy for a Data-driven Society, 12 *Journal of Community Informatics* (2016), available at: http://oro.open.ac.uk/47779/, at 10.

15.013 Second, the push for transparency has resulted in a fragmented web of financial regulations, contributing to ever-increasing compliance costs. The rate of new regulation led one analyst to suggest that 'much like Moore's law in the field of computing there is a "Regulatory Law" that means the operational burden of controlling regulations will double every few years'.[32] Third, regulators are under considerable pressure to effectively supervise with limited resources, even as technology is enabling innovative difficult-to-trace methods for abusing power. For now, reliance on whistleblowers persists. Disruptive technology might provide opportunities for these concerns to be addressed.

3. Opportunity

15.014 The use of information and communications technology, including AI, to spot patterns and make predictions from vast amounts of data is far from new. What has changed is the unprecedented availability of data coupled with the exponential growth of computing power, leading to a dramatic increase in the rate of technological progress.[33] The result has been the delivery of insights from big data necessary to achieve clear transparency, via algorithmic whistleblowing, coupled with the use of automation to eliminate opportunities for corruption.

15.015 AI has been taken to include machines that exhibit aspects of human intelligence like problem solving, making predictions, identifying objects and analysing language.[34] Machine learning is one subset of AI. Supervised machine learning algorithms are 'trained' through the processing of labelled samples of training data by a learning algorithm, before the algorithm is presented with unlabelled test data. Typical applications include the prediction of a label (classification) or a continuous value (regression). Unsupervised learning involves tasks like clustering and dimensionality reduction, in order to 'learn the inherent structure of our data without using explicitly-provided labels'.[35]

15.016 Deep learning is an approach to machine learning which departs from the statistics-based methods that ground the solutions previously described. Deep

---

32 Tom Groenfeldt, Taming The High Costs Of Compliance With Tech (2018), available at: www.forbes.com/sites/tomgroenfeldt/2018/03/22/taming-the-high-costs-of-compliance-with-tech/#3f7d5285d3f7.

33 Tom Simonite, How can AI keep Accelerating after Moore's Law (2017), available at: https://www.technologyreview.com/s/607917/how-ai-can-keep-accelerating-after-moores-law/.

34 While beyond the scope of this chapter, the definition and meaning of 'artificial intelligence' is fiercely contested. Shane Legg and Marcus Hutter, A Collection of Definitions of Artificial Intelligence, 157 *Frontiers in Artificial Intelligence Appl.* 17 (2007), available at: https://arxiv.org/pdf/0706.3639.pdf.

35 Devin Soni, Supervised vs. Unsupervised Learning, Towards Data Science (2018), available at: science.com/supervised-vs-unsupervised-learning-14f68e32ea8d.

learning algorithms learn via layers of artificial neural networks imitating the biological structure and functions of the brain.[36] Whereas the performance of trained machine learning algorithms will at some point reach a plateau, the ability of deep neural networks to replicate real world systems has no such theoretical ceiling.[37] Most promising is deep learning's superior potential to discover structures within otherwise unstructured, unlabelled data – the format of most data in the world.

15.017 A combination of these innovations, among others, is responsible for the displacement of whistleblowers: an under-appreciated consequence of the emergent RegTech and Suptech fields. As described in more detail in Chapter 12, RegTech refers to technological applications which enhance the ability to meet regulatory requirements. Some of the most advanced RegTech solutions are being applied by financial institutions to automate KYC and AML compliance, analysing customer transactions for anomalies and vastly decreasing the number of false positives which are costly to investigate.[38]

15.018 Within companies, RegTech can prevent and detect asset misappropriations, corrupt schemes and financial statement fraud, which would be otherwise indiscernible to a human analyst.[39] AI is also enabling auditors to analyse data and detect connections between e-mails, pdf documents, expense reporting, social media profiles, criminal record checks, work hour reports, registered attempts to access restricted work areas and more.[40] This could even reveal behavioural insights so that 'companies can identify individuals who might pose a higher risk to business'.[41]

15.019 It follows that in some jurisdictions there may be an incentive to adopt RegTech to minimise the risk of corporate criminal convictions. For instance, section 7 of the UK's Bribery Act 2010 – one of the strictest examples of international anti-bribery legislation – makes the failure of an organisation to prevent bribery an offence. However, it is a defence under section 7(2) for the organisation to 'show that [it] had in place adequate procedures designed

36 Snezana Agatonovic-Kustrin and Roderic Beresford, Basic Concepts of Artificial Neural Network (ANN) Modeling and its Application in Pharmaceutical Research, 22 *J Pharm Biomed Anal* 171 (2000) available at: https://doi.org/10.1016/S0731-7085(99)00272-1, at 718–22.

37 Ian Goodfellow et al., *Deep Learning*, The MIT Press (2016), at 197.

38 See generally: Financial Stability Board, The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions (2020) https://www.fsb.org/wp-content/uploads/P091020.pdf.

39 Institute of International Finance, Deploying Regtech Against Financial Crime (2017), available at: https://www.iif.com/system/files/32370132_aml_final_id.pdf.

40 Ibid.

41 Ibid.

to prevent' such conduct. It is not far-fetched to imagine that a company's demonstration of reliance on a RegTech solution to combat bribery could be sufficient to succeed under section 7(2).

15.020 RegTech's counterpart, SupTech, is revolutionising the capabilities of supervisory agencies. Instead of periodically collecting aggregated data in reporting templates, 'data pull' approaches source data directly from the operating systems of regulated institutions at intervals ranging from 24 hours to 15 minutes.[42] This allows for the real-time monitoring of transactions, minimising reporting errors, and removing the opportunity for financial misstatements – even allowing automatic incorporation of changes in regulatory requirements into the technological reporting protocols. 'Data-input' approaches, on the other hand, involve reporting institutions submitting data that are encoded into a human- and machine-readable format that use standardised electronic taxonomies, or 'tags', and sending it to a central database in an unaggregated form. The US Securities and Exchange Commission (the SEC), for example, has since 2009 required public reporting in XBRL format.[43] This data is then fed to its corporate issuer risk assessment dashboard, which analyses the reports to detect traces of fraud. On June 30th, 2020, the US Federal Deposit Insurance Corporation (FDIC) launched a competition in which 20 technology providers pitched prototypes aimed at further improving the ease and effectiveness of timely and granular data use by supervisory institutions. At the time of writing, the pitches are under consideration. The EU's adoption of the Single Electronic Format in January 2020 has also opened up opportunity for innovation in this area.

15.021 Data analysis and visualisation is another important focus of SupTech solutions. For example, the Bank of International Settlements is using data analytics to assess the impact of COVID-19 on local labour markets in the US and Europe in real time using Google Trends Data.[44] The Bank of England's Policy Response Tracker is using webscraping and natural language processing to keep a dedicated policy response tracker dashboard up to date on rapidly changing COVID-related monetary or fiscal measures of 50 countries and counting, taking the burden off of supervisors of international UK-based

42 The National Bank of Rwanda (BNR) was one of the first financial institutions to implement this. Bank for International Settlements, Innovative Technology in Financial Supervision (Suptech) – the experience of early users (2018), available at: https://www.bis.org/fsi/publ/insights9.htm, at 6.

43 Business Reporting Language, or XBRL, is the international data standard for international business reporting: Marc. D Joffe, Open Data for Financial Reporting, *Data Foundation* (2017), available at: https://www.datafoundation.org/xbrl-report-2017/.

44 Ibid., n 44

firms.[45] The Bank of Italy is combining suspicious activity reports with natural language processing analysis of press reviews to detect money laundering. The UK's FCA, meanwhile, has trained algorithms to model normal trading behaviour and automatically report signs of insider trading.[46] These are revelations which otherwise might only have been brought to light by a human whistleblower, if at all.

15.022 Although not typically conceptualised in this context, arguably SupTech can also be used to refer to the use of technology by governments to detect and deter systemic risks within their own agencies and to assist with the supervisory activities of the public and independent watchdog organisations.[47] One example is ProZorro, the much-lauded Ukrainian public e-procurement system, which is being enhanced by AI to identify procurement violations and tenders with a high risk of corruption, including COVID-19 related purchases.[48] Unlike orthodox risk management systems, the indicators are not pre-set beforehand and there is no exhaustive list.[49] In Colombia, the comptroller general's Oceano programme – another fraud detection analytics platform that mines public procurement documents – recently blew the whistle on suspicious links between companies and politicians related to emergency health spending.[50] On 9 December 2020, in conjunction with the UN's Anti-Corruption Day, Microsoft unveiled its Anti-Corruption Technology and Solutions (ACTS) which will reportedly work with governments and other organisations to leverage Microsoft's AI, cloud computing, and data visualisation technologies to 'aggregate and analyse...enormous datasets in the cloud, ferreting out corruption from the shadows where it lives, and even preventing corruption before it happens'.[51] The stated aim is to 'help governments innovate' and eventually 'bring the most promising solutions to the broadest possible audi-

45 Ibid., at 59.

46 Bank for International Settlements, *supra* note 42.

47 Global Witness, Three Ways the UK's Register of the Real Owners of Companies Is Already Proving Its Worth (2018), available at: https://www.globalwitness.org/en/blog/three-ways-uks-register-real-owners-companies-already-proving-its-worth/.

48 Transparency International, Where do we go from here to stop the pandemic? (2020), available at: https://www.transparency.org/en/news/where-do-we-go-from-here-to-stop-the-pandemic.

49 Transparency International Ukraine, Dozorro Artificial Intelligence to Find Violations in ProZorro: How it Works (2018), available at: https://ti-ukraine.org/en/news/dozorro-artificial-intelligence-to-find-violations-in-prozorro-how-it-works/.

50 World Economic Forum, Why data is Latin America's best weapon against COVID-19 corruption (2020), available at: https://www.weforum.org/agenda/2020/08/why-data-is-latin-americas-best-weapon-in-the-fight-against-covid-19-corruption/.

51 Microsoft, Microsoft launches Anti-Corruption Technology and Solutions (2020), available at: https://blogs.microsoft.com/on-the-issues/2020/12/09/microsoft-anti-corruption-technology-solutions-acts /.

ence,' which so far has included partnering with the IDB Transparency fund to bring transparency to the use of COVID-19 stimulus funds.[52]

4. Obstacles

15.023 Various obstacles to RegTech and SupTech adoption remind us that technological solutions are no panacea for eliminating the challenges faced by whistleblowers. As described in more detail in Chapter 12, these include regulatory and legislative barriers to knowledge sharing (such as data protection and localisation laws), legacy IT systems, the lack of integrated data taxonomies and the limited room for financial institutions to innovate while maintaining compliance.[53] There are also at least three potential ethical challenges arising from the proliferation of technology in the present context that merit further examination.

a. Privacy

15.024 The first obstacle is the balance between transparency and privacy. As Fox quips: 'One person's transparency is another's surveillance.'[54] A survey conducted by Ernst & Young revealed a 'tension between opinions about what channels companies should monitor and the types of surveillance that their employees consider a violation of privacy'.[55] The GDPR, as the global legal standard for data protection and privacy, imposes various duties on data controllers and data processors, including obligations to declare a lawful basis for data collection and processing, and limitations on the export of personal data outside the EU. Special attention to these provisions should be paid by organisations that are effectively outsourcing their RegTech and SupTech compliance solutions to third parties.

---

52 Note, however, that Microsoft itself was allegedly implicated in a bribery scheme in Hungary, paying the US SEC 25 million to settle the investigation in 2018. Kyle Wiggers, Microsoft launches effort to fight corruption with AI and other emerging technologies (2020), available at: https://venturebeat.com/2020/12/09/microsoft-launches-effort-to-fight-corruption-with-ai-and-other-emerging-technologies/ .

53 Institute of International Finance, *supra* note 39.

54 Fox, *supra* note 27; and Privacy International, Fintech: Privacy and Identity in the New Data-Intensive Financial Sector (2017), available at: https://privacyinternational.org/sites/default/files/2017-12/Fintech%20report.pdf.

55 For example, around 65 per cent of respondents felt that e-mail and phone-call monitoring was a violation of privacy. See *EY Reporting*, What should be Monitored? (2017) available at: https://www.ey.com/Publication/vwLUAssets/, at 9.

b. Bias

15.025 The second obstacle concerns the risks of error, bias and the threat of algorithmic discrimination.[56] Machine learning algorithms will learn from and perpetuate distortions in training data. Moreover, inherently algorithms are optimised to achieve particular goals, which can lead to biased decision making. RegTech and SupTech are not immune. For example, fraud detection algorithms have been shown to be biased against certain ethnic minorities, immigrants and even against men.[57] While extensive technical research is being done on identifying and correcting bias in algorithms, others are advocating algorithmic impact assessments and even making a business out of algorithmic auditing.[58] ORCAA, one such consultancy, assesses the quality of training data, testing the algorithms' design, implementation, execution and ethical consequences, and offers training in algorithmic auditing. The resultant seal is 'like an organic sticker for algorithms', on the basis that 'the food we eat has quality certifications. Why shouldn't the algorithms that shape our world?'[59] In the meantime, the question arises: how much less biased than a human does an algorithm have to be before we are willing to let it loose on the work of whistleblowers?

c. Black box

15.026 Finally, the third obstacle is that machine learning algorithms tend to be 'opaque in the sense that … rarely does one have any concrete sense of how or why a particular classification has been arrived at from inputs'.[60] This is known as the explainability or black box problem, and it is particularly acute in deep learning. Some argue that even the technologically-increased accuracy of decisions does not compensate for the inability to explain the weighting of decision-making factors and essentially fails to respect a subject's dignity,[61] offending one's 'right to an explanation',[62] and raising Kafkaeque concerns

56 Solon Barocas and Andrew Selbst, Big Data's Disparate Impact, 104 *California Law Review* 671 (2016), available at: http://www.californialawreview.org/wp-content/uploads/2016/06/2Barocas-Selbst.pdf.

57 Adeesh Goel, Algorithmic Bias: Challenges and Solutions (2017), available at: https://mse238blog.stanford.edu/2017/08/adeesh/algorithmic-bias-challenges-and-solutions/.

58 FAT/ML: Principles for Accountable Algorithms and a Social Impact Statement for Algorithms, available at: http://www.fatml.org/resources/principles-for-accountable-algorithms.

59 Katharine Schwabe: This logo is like an organic sticker for algorithms (2018), available at: https://www.fastcompany.com/90172734/this-logo-is-like-an-organic-sticker-for-algorithms-that-arent-evil.

60 Jenna Burrell, How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms, *Big Data and Society* (2016), available at: https://journals.sagepub.com/doi/pdf/10.1177/2053951715622512, at 1.

61 Jeremy Waldron, How Law Protects Dignity, 71 *Cambridge Law Journal* 200 (2012), at 210.

62 Reuben Binns, Max Van Kleek, et al., 'It's Reducing a Human Being to a Percentage'; Perceptions of Justice in Algorithmic Decisions (2018), available at: https://doi.org/10.1145/3173574.3173951.

for fair trial standards.[63] There are a growing number of researchers, business leaders and policy makers who are developing both technical solutions to explainable AI (XAI) and corporate civil regulation for the development of ethical AI.[64] On the other hand, some counter that to the degree that AI becomes 'explainable', bad actors may be able to adjust their behaviour to 'game' the system.[65] As it stands, the paradox is that the increase in transparency in the sense of information disclosure is dependent on an opaque mechanism. Minimising harm to whistleblowers requires a trade-off with the potential of harm to the subjects of inexplicable algorithmic outputs.

15.027 In the age of information, reliance on disruptive technology is essential to achieving clear transparency and accountability. RegTech and SupTech promise to minimise compliance costs and revolutionise supervision and detect and report misconduct more effectively and efficiently than human whistleblowers. Moreover, algorithms cannot be personally victimised or directly retaliated against. Algorithmic whistleblowing is therefore a positive development insofar as it reduces reliance on whistleblowers who often endure significant personal and professional costs, despite the public interest in their reporting. However, the technology that is enabling these changes also poses new ethical dilemmas, involving the balancing of privacy against transparency and accuracy against dignity. Despite the inherent difficulties, eventual adoption of RegTech and SupTech is likely inevitable, which means that whistleblowing may be just the latest human endeavour to be taken over by machines.

C. TECHNOLOGY EMPOWERING THE WHISTLEBLOWER

15.028 The barriers to blowing the whistle are widely known and have been extensively analysed, the foremost being the fear of retaliation.[66] The difficulty of ensuring confidentiality and, in some cases, the anonymity of the whistleblower therefore looms large. A related problem is the utilisation of trusted channels of reporting, which must be secure and effective. Technological applications, designed to facilitate the whistleblowing process, offer potential

63 Council of Europe, Algorithms and Human Rights (2017), available at: http://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5.

64 See the 2017 Asilomar principles, which has 1,273 AI/Robotics researchers as signatories. The Future of Life Institute, The Asilomar AI Principles (2017), available at: https://futureoflife.org/ai-principles/?submitted=1&cn-reloaded=1#confirmation.

65 Paul B. de Laat, Algorithmic Decision-making Based On Machine Learning from Big Data: Can Transparency Restore Accountability? *P.B. Philos. Technol.* 17 (2017), available at: https://link.springer.com/content/pdf/10.1007%2Fs13347-017-0293-z.pdf.

66 International Bar Association, *supra* note 4.

solutions.[67] As the whistleblowing technology has developed and diversified (partly in response to the growing body of regulations imposing whistleblowing standards on organisations), two trends have become apparent: integration and specialisation. First, more whistleblowing technology providers are offering integrated platforms for whistleblowing that include text, e-mail, hotlines, mobile apps, increasing companies' case management efficiency by making use of the cross-benefits of different technologies.[68] The second trend is the breadth of the specialised, 'professionalised' whistleblowing platforms that have emerged in in areas like journalism,[69] sport,[70] securities trading[71] and, most prominently as of late, healthcare.[72] This has allowed for the creation of industry-wide cross-border accessible protection mechanisms, which are tailored to the needs of the given professionals.

15.029 This section aims to examine four technologies which are running in parallel to these trends to empower the whistleblower: hotline services and web portals, mobile apps and blockchain.[73]

67 For a more specific discussion of the potential of technology for addressing harassment, see Emma Franklin and Kieran Pender, Innovation-led cultural change: can technology effectively address workplace harassment? (2020), available at: https://www.ibanet.org/Document/Default.aspx?DocumentUid=4c00afd9-53e7-4ad6-8db0-c663c2f23f45.

68 For example, Got Ethics A/S, available at: https://www.gotethics.com/en/; Whispli, available at: https://www.whispli.com/; NAVEX Global's Whistleblower Hotline, available at: https://www.navexglobal.com/en-us/products/hotline-reporting-and-intake; Your Call, available at: https://www.whistleblowing.com.au/solutions/?gclid=Cj0KCQiA8dH-BRD_ARIsAC24umYuPBhBvQqxswSWcg4F1hO9nsXDf3Tnz1kiIEK6sJtxw3tTMVzKsRYaAnFYEALw_wcB; EQS Integrity Line, available at: https://www.eqs.com/en-us/compliance-solutions/integrity-line/#features.

69 For example, SecureDrop, available at: https://securedrop.org/; Digital Whistleblowing Fund, available at: https://www.hermescenter.org/supporting-diverse-initiatives-in-europe-through-the-use-of-secure-whistleblowing-platforms/. For a comprehensive review of whistleblowing digital platforms used by journalists see Philip di Salvo, *Digital Whistleblowing Platforms in Journalism. Encrypting Leaks* (2020) Palgrave Macmillan at 63–89.

70 For example, SportsLeaks, available at: https://www.sportsleaks.com/; World Anti-doping Agency's SpeakUp, available at: https://speakup.wada-ama.org/WebPages/Public/FrontPages/Default.aspx; International Olympic Committee's Hotline, available at: https://ioc.integrityline.org/. See generally, United Nations Office on Drugs and Crime, Reporting Mechanisms in Sport: A Practical Guide for Development and Implementation (2019), available at: https://stillmedab.olympic.org/media/Document%20Library/OlympicOrg/IOC/What-We-Do/Protecting-Clean-Athletes/Competition-manipulation/IOC-UNODC-Reporting-Mechanisms-in-Sport-ebook.pdf.

71 For example, SEC Whistleblowing program, available at: https://www.sec.gov/whistleblower.

72 For example, NixWhistle proposed to create 'CoronaSpeak', encouraging people to report positive COVID-19 cases, available at: https://www.nixwhistle.com/.

73 Social media has become a popular channel for online whistleblowing. Analysis of this phenomenon is outside the scope of the present chapter; however, for some analysis of the role of social media for whistleblowers see: H. Latan, C.J. Chiappetta Jabbour, andLA.B.opes de Sousa Jabbour, Social Media as a Form of Virtual Whistleblowing: Empirical Evidence for Elements of the Diamond Model. *J Bus Ethics* (2020). https://doi.org/10.1007/s10551-020-04598-y

1. Hotline services and web portals

15.030 The oldest and most widely used technical applications are hotline services.[74] They offer anonymity and increased accessibility, but at the same time, it is impossible to share documents, expensive to maintain qualified operators, who are able to work across different languages and time zones, and hard to establish further contact unless the whistleblower calls again.[75]

15.031 A potentially more effective technology, which increasingly replaces hotlines and direct reporting, are web portals. They give the whistleblower the benefit of remaining anonymous by creating an account with a random username, through which the whistleblower can submit a report from anywhere in the world in a variety of languages, attach any type of document, communicate with an investigative authority and track the progress of the disclosure.[76] Additionally, an automated whistleblowing system allows companies to compile statistical data to analyse problematic areas using technologies described in Section B above.

15.032 Web portals can be established within an organisation or outsourced to a third-party company. A notable example of the latter is Business Keeper AG. It provides a certified secure and private external whistleblowing portal for companies in 197 countries, including administrative bodies such as the Austrian Central Department of public prosecution of economic crimes and corruption and the German Federal Financial Supervisory Authority.[77]

15.033 The perceived effectiveness of web portals has prompted the emergence of companies offering open-source whistleblowing software, providing any organisation with tools to create their own whistleblowing web portal. Two

---

74 An interesting finding in one of the reported surveys is that the effectiveness of this technology increases when it is branded as a 'helpline', rather than a 'hotline'. See Stephen R. Stubben, Evidence on the Use and Efficacy of Internal Whistleblowing Systems, 58 *Journal of Accounting Research* (2020), available at: https://onlinelibrary.wiley.com/doi/10.1111/1475-679X.12303.

75 John Wilson, Whistleblowing: What are the Most Effective Speak-up Channels? (2017), available at: http://in-houseblog.practicallaw.com/whistleblowing-what-are-the-most-effective-speak-up-channels/. For more analysis of the hotlines' effectiveness, see e.g., Eugene Soltes, Paper Versus Practice: A Field Investigation of Integrity Hotlines, 58 *Journal of Accounting Research* (2020), available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-679X.12302; Stephen R. Stubben, Evidence on the Use and Efficacy of Internal Whistleblowing Systems, 58 *Journal of Accounting Research* (2020), available at: https://onlinelibrary.wiley.com/doi/10.1111/1475-679X.12303.

76 Mostafa Hussien and Toshiyuki Yamanaka, Whistleblowing at Work. Can ICT Encourage Whistleblowing?, 27 *Joho Chishiki Gakkaishi* 150 (2017), available at: https://www.jstage.jst.go.jp/article/jsik/27/2/27_2017_017/_article/-char/en at 151.

77 Business Keeper AG, available at: https://www.business-keeper.com/en/whistleblowing-system/references.html.

prominent examples of such software are GlobaLeaks and SecureDrop.[78] GlobaLeaks is an Italian-based software that allows any organisation to set up their own website, guaranteeing the security level necessary for a whistleblowing platform.[79] It has been used widely by private companies, non-governmental organisations and government institutions.[80] Further, GlobaLeaks offers an accessible solution for the countries to implement the recently adopted EU Whistleblowing Directive through a project called 'Expanding Anonymous Tipping', which is already operating in 11 EU member states and allowing for both the safe blowing of a whistle and assisting companies with compliant next steps.[81]

15.034 SecureDrop is 'an open-source submission system that organisations can install to securely accept documents from anonymous sources'.[82] Employing the Tor network, it acts as an intermediary between whistleblowers and journalists by allowing the former to download documents to a server and contact journalists using SecureDrop messages.[83] Since its launch in 2013, prominent news organisations such as *The New York Times* and *The Guardian* have used SecureDrop to solicit information.

15.035 Wikileaks is another web portal, infamous for its major role in publishing millions of leaked documents, including the Iraq War Logs and Hillary Clinton's emails. Wikileaks has garnered significant controversy and the division between public interest whistleblowing and politically-motivated leaking is contested – an increasingly blurred distinction considered further below.

15.036 Such web portals have been required to adopt certain technological measures to guarantee the security and anonymity necessary for whistleblowing. The majority employ Tor, a 'group of volunteer-operated servers' that constitute a distributed anonymous network. Servers of this network are connected via virtual tunnels, concealing the path of a user's traffic, ensuring privacy and preventing tracking. Whistleblowing portals reliant on Tor adopt its 'onion

78 For a further review of the features of GlobaLeaks and SecureDrop see Matthew Jenkins, Overview of whistleblowing software, *U4 Helpdesk Answer* (2020), available at: https://www.u4.no/publications/overview-of-whistleblowing-software.pdf.

79 GlobaLeaks, available at: https://www.globaleaks.org/about-us#our-vision.

80 Europe will begin to protect whistleblowers; institutions and firms must prepare (03 March 2020), available at: https://eatproject.eu/europe-will-begin-to-protect-whistleblowers-institutions-and-firms-must-prepare/.

81 See Expanding Anonymous Tipping website: https://eatproject.eu/.

82 SecureDrop, available at: https://securedrop.org/.

83 Amy Davidson Sorkin, Introducing StrongBox (2013), available at: https://www.newyorker.com/news/amy-davidson/introducing-strongbox.

service' in order to publish a website without revealing its location. This service operates by using random 'rendezvous-points' where a client can go to access the website, using a public key and an onion address, without revealing their identity.[84]

2. Mobile apps

15.037 A relatively recent promising development is the emergence of whistleblowing mobile apps. These channels, which are increasingly being offered by producers of whistleblowing solutions, combine the advantages of web portals with an accessibility of mobile phones, creating a universal platform that can be employed almost anywhere. For example, Whispli offers a mobile app, available in 60 languages, which provides a secure, user-friendly and accessible way to blow the whistle, taking use of such features as a QR-code for easier reporting.[85] Similarly, Expolink's app 'SpeakingUp' allows for the submission of an encrypted report from anywhere in the world in multiple languages.[86]

15.038 The use of mobile apps has been especially prominent on governmental levels in countries with less developed technological infrastructure. Most recently, in September 2020, the Zimbabwe Anti-Corruption Commission introduced an anti-corruption whistleblowing app.[87] The app is expected to assist in anti-corruption matters by allowing citizens to report 'safely, loudly and visibly', submitting evidence necessary to launch anti-corruption investigations. The added benefit, relevant in times of COVID-19 pandemic, is the minimum social contact required for reporting.[88] Similar apps, intended to provide a secure and private outlet for blowing the whistle, were launched in Nigeria ('Wahala Dey'),[89] in Abu Dhabi ('Inform the Prosecution')[90] and in India, exclusively for members of political party Makkal Needhi Maiam in order to flag issues caused by party members ('Maiam Whistle').[91]

---

84 Tor Project, Tor: Onion Service Protocol, available at: https://www.torproject.org/docs/onion-services.

85 Whispli, available at: https://www.whispli.com/whistleblower-platform-features/#.

86 Expolink's SpeakingUp mobile app, available at: https://www.expolink.co.uk/whistleblowing-hotline/mobile-app-2/.

87 See ZACC Launches Anti-Corruption Whistleblower Mobile App, New Zimbabwe (09 September 2020), available at: https://www.newzimbabwe.com/zacc-launches-anti-corruption-whistleblower-mobile-app/.

88 Ibid.

89 Wahala Dey, available at: https://icpc.gov.ng/.

90 Wam, Abu Dhabi Launches New Whistleblower App (2018), available at: https://www.khaleejtimes.com/nation/abu-dhabi-launches-new-whistleblower-app.

91 Dharani Thangavelu, Kamal Haasan Launches Party App to Focus on Key Issues (2018), available at: https://www.livemint.com/Politics/.

3. Blockchain

15.039 Perhaps the most promising technological development for whistleblowers – albeit the least actualised – is blockchain, which when applied to whistleblowing can offer several important advantages.

a. Anonymity

15.040 Given the ease of tracking the identity of whistleblowers when communicating online,[92] blockchain is a potential solution which can strike the balance between the need for anonymity and the importance of an investigative authority being able to contact the whistleblower for further details. A project called WhistleAI is currently working towards realising this potential by combining the benefits of blockchain, crowdsourcing and AI. To ensure anonymity their platform relies on zero-knowledge protocols, which entails splitting information into fragmented pieces before sending it to the nodes for verification. This ascertains the protection of whistleblower identity while allowing the members of the network to verify the correspondence of the whistleblower's allegation with the information provided in their report.[93]

b. Immutability

15.041 The second advantage of blockchain is its immutability. Data, once uploaded on a blockchain-based platform, cannot be deleted or tampered with as it is aggregated into interconnected blocks. This prevents employers or organisations implicated by the disclosure from concealing the whistleblowing report. An additional function of this platform may be public time stamping, which allows whistleblowers to aggregate data for a period of time, before deciding whether to publish the materials or not.[94] Time stamping and immutability of data would mean information could be used in future court proceedings without concern for the veracity of evidence.[95]

c. Resilience

15.042 Relying on blockchain for whistleblowing would drastically increase the resilience of the whistleblowing platform. Unlike website-based platforms, block-

92 Owen Bowcott, Whistleblowers Endangered in Digital Age, Says Lawyers' Report (2017), available at: https://www.theguardian.com/media/2017/feb/22/.

93 WhistleAI, available at: https://www.whistleai.io/WhistleAI.pdf.

94 Shafi Goldwasser and Sunoo Park, Public Accountability vs. Secret Laws: Can They Coexist? A Cryptographic Proposal (2017), available at: https://eprint.iacr.org/2018/664.pdf, at 4.

95 Wolfie Zhao, Chinese Supreme Court Admitted Blockchain Evidence as Legally Binding (2018), available at: https://www.coindesk.com/chinas-supreme-court-recognizes-blockchain-evidence-as-legally-binding/.

chain would not be susceptible to DDOS attacks or disruption of a domain name and there would be no need to change the hosting servers.[96]

d. Compensation

15.043 A unique and arguably controversial feature of blockchain is its ability to offer compensation to the whistleblower through smart contracts. This mechanism would not only offer the whistleblower confidence of their identity's security through blockchain, but also provide them with adequate compensation through the use of cryptocurrency, which could be automatically transferred to their account once the 'leaked' data is verified and the appropriate conditions for reward are satisfied. This idea has been implemented in WhistleAI, where a privacy coin named WISL is used both for compensating whistleblowers and for incentivising crowdsourcing participants that allow the platform to continue operating.[97]

e. Escrow

15.044 Another possible advantage of blockchain is its application as an information escrow. This can be done through a smart contract, programmed to release information only if certain conditions are satisfied. For example, Callisto, initially designed to combat sexual harassment on college campuses, forwards the reported misconduct only when there are at least two complaints about the same perpetrator.[98] Such technology, combined with the other benefits of blockchain, would help eliminate the 'first-mover disadvantage' and lessen the likelihood of retribution.[99]

f. Concerns

15.045 There are, of course, potential disadvantages to blockchain's adoption in the present context. First, employing a distributed network means that all users of a blockchain could have the sensitive data on their nodes (computers), potentially exposing them to liability in some jurisdictions.[100] One way in which this

96 Yochai Benkler, A Free Irresponsible Press: Wikileaks and the Battle Over the Soul of the Networked Fourth Estate, 46 *Harv. C.R.-C.L. L. Rev.* 311 (2011), available at: http://benkler.org/Benkler_Wikileaks_current.pdf, at 3.

97 WhistleAI, *supra* note 93.

98 Callisto project, available at: https://www.projectcallisto.org/.

99 Ian Ayres and Cait Unkovic, Information Escrows, 111 *Michigan Law Review* 145 (2012), available at: https://repository.law.umich.edu/mlr/vol111/iss2/1, at 3; and Carsten Tams, Can 'Allegation Escrows' Remedy the Underreporting of Sexual Harassment? (2017), available at: http://www.fcpablog.com/blog/2017/11/20/carsten-tams-can-allegation-escrows-remedy-the-underreportin.html.

100 Roman Matzutt et al., A Quantitative Analysis of the Impact of Arbitrary Blockchain Content on Bitcoin, (2018), available at: https://www.martinhenze.de/wp-content/papercite-data/pdf/mhh+18.pdf, at 7.

risk may be mitigated is through the use of Enigma secret contracts.[101] These smart contracts use 'secure computation' technologies to compute over the encrypted data, thus concealing sensitive information contained in the report from the other members of the network, but retaining their ability to validate the transactions.[102] Secret contracts, thereby, offer privacy to the whistleblowers, hiding their identity and mitigating the risk of retaliation. A second potential drawback of blockchain is that there has to be an established mechanism of incentives to continue mining and, consequently, verifying the information. Employing cryptocurrencies for this purpose might be one of the solutions, for example, as demonstrated by the WhistleAI project.

g. Application

15.046 Today there are very few blockchain-based whistleblowing platforms. One of the only examples is NixWhistle, a whistleblowing platform, built on an open-source blockchain technology 'Corda'.[103] It operates by assigning to a user one of three roles: whistleblower, investigator and reviewer. Such clear division of roles, coupled with the benefits of the blockchain technology, has allowed NixWhistle to ensure the anonymity of the whistleblower, integrity of the reported data at every stage of the process and adherence to the role-based access to information.

15.047 Another broadly analogous platform, Darkleaks, is essentially a black market for information (such as trade secrets and source codes). The leaked document is presented in the form of several segments, each one hashed with different Bitcoin addresses.[104] While Darkleaks might not represent a whistleblowing platform in a traditional sense, it demonstrates the technological possibilities offered by blockchain in this context.

15.048 The growing use of technology to empower whistleblowers coincides with heightened appreciation of the role of whistleblowers in society. These are welcome developments, albeit they need to be coupled with mechanisms to ensure that new technologies actually protect whistleblowers. Emergent technology, particularly blockchain, promises not only to facilitate whistleblowing but also to secure the veracity of information passed on, empowering whistleblowers to a potentially unprecedented degree.

---

101 Enigma Protoco, Overview, available at: https://enigma.co/protocol/.

102 Guy Zyskind, Defining Secret Contracts (2018), available at: https://blog.enigma.co/defining-secret-contracts-f40ddee67ef2.

103 NixWhistle, available at: https://www.nixwhistle.com/.

104 Mellisa Tolentino, Darkleaks, a Haven for Whistleblowers and Pirates (2015), available at: https://siliconangle.com/2015/02/04/darkleaks-a-haven-for-whistleblowers-and-pirates/.

D. TECHNOLOGY CHANGING THE WHISTLEBLOWER

15.049 Some 50 years ago, Daniel Ellsberg spent 18 months meticulously copying page after page of incriminating materials to reveal the Pentagon Papers.[105] It took only one memory card and several clicks for Edward Snowden to expose gigabytes of data to the masses.[106] Indeed, the most notorious of recent whistleblowing incidents have taken the form of massive dumps of information to web-sources: Snowden's 2013 disclosure was 60GB in size; Antoine Deltour's 2014 Luxleaks were 4GB; while the 2016 Panama Papers included 2.6TB of information.[107] Moreover, all of them contained massive collections of documents containing a multitude of revelations of which the whistleblowers themselves may have been unaware.

15.050 By radically affecting the sheer amount of information that can be disclosed and the means by which to do so, technology is altering the nature of contemporary whistleblowing by blurring the lines between 'whistleblowers', 'hackers' and 'leakers'. The implications of these developments must be confronted if whistleblower protections worldwide are to be fit for the future.

15.051 Whistleblowing, leaking and hacking do not, in public discourse or as a matter of law, have universally accepted definitions, and often these concepts are conflated. However, 'whistleblowing' has been largely defined as disclosing information about a wrongdoing,[108] while 'leaking' is usually understood as revealing confidential information without official authorisation.[109] As whistleblowing takes the form of increasingly large leaks, these concepts are eliding, Savage suggests that 'unauthorised disclosures made to the public are likely to be considered whistleblowing where there is public interest value in the information disclosed'.[110]

---

105 Daniel Ellsberg, *Secrets: A Memoir of Vietnam and the Pentagon Papers* (2003), at 301.

106 Richard J. Aldrich and Christopher R. Moran, 'Delayed Disclosure': National Security, Whistle-Blowers and the Nature of Secrecy, *Political Studies* (2018), available at: https://journals.sagepub.com/doi/abs/10.1177/0032321718764990, at 7.

107 Suelette Dreyfus, Chelsea Manning and the Rise of 'big data' Whistleblowing in the Digital Age (2018), available at: https://theconversation.com/chelsea-manning-and-the-rise-of-big-data-whistleblowing-in-the-digital-age-102479.

108 OECD, Committing to Effective Whistleblower Protection (2016), available at: https://www.oecd.org/daf/anti-bribery/Committing-to-Effective-Whistleblower-Protection-Highlights.pdf, at 18.

109 Ashley Savage, Whistleblowers for Change: The Social and Economic Costs and Benefits of Leaking and Whistleblowing (2018), available at: https://www.opensocietyfoundations.org/sites/default/files/20181120-whistleblowers-for-change-report.pdf, at 7.

110 Ibid.

15.052 The third dimension complicating this emergent dynamic is the rise of hackers fulfilling a whistleblower-like function. This is illustrated by the 'rise of cybersecurity whistleblowers',[111] who have been thrust into the limelight following incidents such as the WannaCry ransomware attack, the Equifax breach and the Cambridge Analytica Facebook data breach. This trend has only been underscored during the COVID-19 pandemic. Cybersecurity whistleblowers can include a range of actors from non-technical company employees unintentionally made aware of security flaws or unreported data breaches, to 'ethical' or 'white-hat' hackers, cybersecurity professionals or hobbyists who conduct solicited or unsolicited hacks and disclose any discovered vulnerabilities to the public ('full disclosure') or to the company ('coordinated/responsible disclosure'), without exploiting those flaws.[112] For example, since revelations about Russian interference in the 2016 US election emerged, a group of ethical hackers has turned their attention to election security, even spending their own money to buy electronic voting machines for study. [113]Many companies are increasingly investing in the services of ethical hackers, as evidenced not only by the demand for employees and contractors, but also by the growth of 'hacking as an industry', which includes crowdsourced cybersecurity and the burgeoning freelance 'bug-hunter' bounty market, under which companies are amending their terms of service to include standardised safe harbour provisions for good-faith security research[114] and incentivising ethical hackers to hunt for vulnerabilities by paying for discovery.[115]

15.053 Difficulties arise when, in light of cybersecurity revelations, a company responds by suppressing the information to save the remedial expense or conceal a past incident. In these situations, those who make vulnerability disclosures risk reprisals in the form of intimidation, job termination and criminal or civil suits.[116] Outdated laws in this area, such as the US's Computer Fraud and Abuse Act 1984 (CFAA) and the UK's Computer Misuse Act 1990, mean that ethical hackers may be prosecuted relatively easily. On 30 November 2020, the US Supreme Court heard arguments in *Van Buren v United States*

111 Dallas Hammer and Evan Bundschuh, The Rise of Cybersecurity Whistleblowing (2016), available at: https://wp.nyu.edu/compliance_enforcement/2016/12/29/the-rise-of-cybersecurity-whistleblowing/.

112 National Cyber Security Centre, Coordinated Vulnerability Disclosure: The Guideline (2018), available at: https://www.ncsc.nl/english/current-topics/responsible-disclosure-guideline.html.

113 Chris O'Brien, How ethical hackers are trying to protect the 2020 U.S. elections (2020), available at: https://venturebeat.com/2020/10/23/how-ethical-hackers-protect-2020-u-s-elections/.

114 Disclose.io, Safe, Simple, Standardized Vulnerability Disclosure (2020) available at: https://disclose.io/.

115 Bugcrowd, A New Decade in Crowdsourced Security (2020) available at: https://www.bugcrowd.com/blog/3-major-security-priorities-in-the-covid-19-era/.

116 Ibid.

– a seminal case in cybersecurity law.[117] The case centres on the interpretation Section 1030(a)(2) of the CFAA, which prohibits obtaining information from a protected computer by intentionally accessing a protected computer without authorisation or by exceeding authorised access. In amicus briefs filed in the case, cybersecurity researchers argued that an overbroad interpretation of this provision would be catastrophic for cybersecurity, as researchers and ethical hackers whose actions may involve violating company's terms of service face further risk of criminal and civil penalties. This should concern us all. Since the COVID-19 pandemic, malicious actors have exploited the ensuing chaos and the move to remote working, leading to a seven-fold increase in ransomware attacks,[118] and the World Health Organization reports a five-fold increase in cybersecurity attacks.[119] Cybersecurity whistleblowers have proven themselves to be an essential aspect of compliance reporting, risk minimisation, and defence of the technological infrastructure that underpins virtually every aspect of modern society.[120]

15.054 The consequences of the convergence of whistleblowing, leaking and hacking can be broadly divided into three groups: political, ethical and legal implications.

1. Political implications

15.055 Whistleblowing, leaking and hacking carry different connotations. In particular, whistleblowing, though controversial, is widely recognised as having positive attributes as a valuable compliance and accountability tool. Leaking and hacking, on the other hand, arguably carry negative, criminal connotations.[121] The conflation of these three phenomena gives rise to the risk of strategically-framed policy decisions to justify undermining whistleblower protections. For instance, while campaigning, Obama expressed support for whistleblowers as one of 'the best source(s) of information about … abuse

117 At the time of writing, the outcome of the case is unknown. Amit Yoran, The Future Of Cybersecurity Law Hinges On The Supreme Court, available at: https://www.forbes.com/sites/amityoran1/2020/11/16/the-future-of-cybersecurity-law-hinges-on-the-supreme-court/?sh=6ee680c8528a; for commentary and case updates, visit: https://www.scotusblog.com/case-files/cases/van-buren-v-united-states/.

118 University of South Florida, Research Shows a 715% Increase in Ransomware Attacks in 2020 (2020), available at: https://cyberflorida.org/covid/bitfender/.

119 World Health Organization, WHO reports fivefold increase in cyber attacks, urges vigilance (2020), available at: https://www.who.int/news/item/23-04-2020-who-reports-fivefold-increase-in-cyber-attacks-urges-vigilance.

120 Madrea Matwyshyn, Hacking Speech: Informational Speech and the First Amendment, 107 *North Werstern U. L. Rev.* 795 (2013), available at: https://scholarlycommons.law.northwestern.edu/nulr/vol107/iss2/10/.

121 Will Ma et al., Whistleblower or Leaker? Examining the Portrayal and Characterization of Edward Snowden in USA, UK, and HK Posts, in Ma et al. (eds) *New Media, Knowledge Practices and Multiliteracies* (2014), available at: https://www.researchgate.net/publication/265121066_Whistleblower_or_Leaker_Examining_the_Portrayal_and_Characterization_of_Edward_Snowden_in_USA_UK_and_HK_Posts.

in government'. However, his administration proceeded to engage in an unprecedented 'hunt for leakers', and eight individuals were prosecuted under the Espionage Act 1917, more than all previous governments combined.[122] In Australia, meanwhile, the federal government imposed severe potential criminal liability for public servant 'leakers', before implementing whistleblower protection reform. One employment lawyer noted: 'Cracking down on unauthorised disclosure before a robust and effective program for authorised disclosure of official information is illogical at best, and shows a barely disguised contempt for whistleblowers at worst.'[123] These are but two manifestations of a wider issue – the increased politicisation of the distinction between leakers, whistleblowers and hackers in the digital age. This strategy is becoming more prevalent as big data leaks, and the hacks that may enable them, have the potential to do more damage than ever before. As Shkbatur writes, 'the war against leaks can therefore be understood as a response against the new whistleblowing reality created by the Internet'.[124]

2. Ethical implications

15.056 The debate surrounding the justifiability and desirability of whistleblowing, in particular regarding the balance between collateral damage to individuals and the public's right to know, is not new. However, the nexus between leaking, hacking and whistleblowing exacerbates these issues. As new-age whistleblowers take the form of custodians of huge amounts of data, they arguably have even greater ethical burdens to discharge.

15.057 The lack of responsible practices for publishing data leaks has already had consequences. For example, in 2016 Wikileaks revealed 300,000 e-mails dubbed the 'Erdogan e-mails'. While subsequent investigation of these e-mails did not yield significant evidence of wrongdoing, sensitive personal information, including current phone numbers, citizenship IDs, addresses and political party affiliations of millions of women, was released.[125] Such data dumps have

122 Mary-Rose Papandrea, Leaker Traitor Whistleblower Spy: National Security Leaks and the First Amendment, 94 *B.U. L. Rev* 449 (2014), available at: https://www.bu.edu/bulawreview/files/2014/05/PAPANDREA.pdf, at 451.

123 John Wilson, Whistleblowing isn't Dobbing. It Supports our Democracy (2018), available at: https://www.smh.com.au/public-service/whistleblowing-isnt-dobbing-it-supports-our-democracy-20180203-h0t6cv.html.

124 Shkabatur, *supra* note 13, at 116.

125 Zeynep Tufekci, WikiLeaks Put Women in Turkey in Danger, for No Reason, *The Huffington Post* (2017).

also revealed the personal information of rape victims and even the identities of several gay men in Saudi Arabia, where homosexuality is illegal.[126]

15.058 These dilemmas demand a balancing test and – perhaps more than ever before – there are no easy answers. On the one hand, excessive whistleblower self-censorship is undesirable, in particular from a transparency perspective. On the other hand, there is a need to minimise harm to those who, in the absence of basic responsible data practices, may be unnecessarily compromised as a side-effect of holding the powerful to account.[127]

3. Legal implications

15.059 Finally, changes in the nature and methods of whistleblowing are bound to affect the drafting of new whistleblower protection laws that are being adopted across the world, as well as the assessment of the extent to which existing whistleblowing laws are fit for purpose. While there are many legal issues that might arise, only a few relating to some of the key elements of whistleblower protection schemes will be highlighted below.

15.060 First, as a result of these changes, the whistleblower, the regulator, any whistleblowing service provider, and the entity accused of misconduct will commonly be based in different jurisdictions. This puts pressure on jurisdictional inconsistencies between the kinds of disclosure which will attract the protection of whistleblower legislation. There is a range of sources of law where whistleblower protections may be found, including bespoke legislation, sectoral laws and laws specifically aimed at the public service. This can result in legal loopholes which may deter potential cybersecurity whistleblowers who are unsure whether they would be protected. Other laws take a more expansive approach, capturing for instance disclosures in the 'public interest' or those which disclose 'abuse of laws'.[128]

15.061 For example, the law in the US is unclear about cybersecurity whistleblowing, as there is no federal statute that directly addresses it. Instead, protection must

126 Nicky Woolf, WikiLeaks Posted Medical Files of Rape Victims and Children, Investigation Finds (2016), available at: https://www.theguardian.com/media/2016/aug/23/wikileaks-posts-sensitive-medical-information-saudi-arabia.

127 Alix Dunn, Responsible Data Leaks and Whistleblowing (2016), available at: https://www.theengineroom.org/responsible-data-leaks-and-whistleblowing/?fbclid=IwAR0_8kIpLnGmSaCRZHChpbiegf0dWKd1XqXNdm7gSGGKdrIak1Vjc4sNUk.

128 OECD, G20: Study of Whistleblower Protection Frameworks, Compendium of Best Practices and Guiding Principles for Legislation (2012), available at: https://star.worldbank.org/document/study-whistleblower-protection-frameworks-compendium-best-practices-and-guiding-principles, at 6.

be read down from various existing federal or state laws.[129] The SEC has been taking a proactive approach to whistleblowing, and cybersecurity whistleblowing in particular. In 2011, the SEC's Division of Corporate Finance called for the disclosure of cybersecurity incidences materially relevant to a company's operations as a part of regular reporting requirements under the federal securities regulation.[130] In 2018, the SEC reiterated this call and offered further interpretive guidance.[131] However, this is not legally binding, giving rise to a 'grey area' with respect to whether cybersecurity whistleblowers can take advantage of the robust protection under the SEC's whistleblower protection programmes and the Dodd-Frank Act.[132] Meanwhile, potential cybersecurity whistleblowers who work on entities not regulated by these federal statutes are left to fend off potential criminal liability for their actions.[133] Even if *Van Buren* is decided in favour of petitioners, for years security researchers have reported a chilling effect on their work, admitting that 'facing legal action is just one of those things where it's just not worth it anymore'.[134] By contrast, Article 1 of the European Union Directive on Whistleblowing 'lays down common minimum standards for the protection of persons reporting (on) unlawful activities or abuse of law' and specifically includes 'protection of privacy and personal data, and security of network and information systems.'[135] While the inclusion of cybersecurity whistleblowers is a positive development, there remains a lack of legal clarity over intersection with criminal laws directed at not dissimilar conduct.[136]

15.062 Some question whether even broader whistleblower laws should capture the reporting of ethical or immoral conduct, especially where 'these tread the fine line between illegality and morality'.[137] On the one hand, it is argued that extending protections to the disclosure of ethical or moral concerns may

129 Alexis Ronicker, Cybersecurity Whistleblower Protections (2017) available at: https://www.kmblegal.com/sites/default/files/cybersecurity-whistleblower-protection-guide.pdf.

130 Jennifer M. Pacella, The Cybersecurity Threat: Compliance and the Role of Whistleblowers, 11 Brook. J. Corp. Fin. & Com L. (2016), available at: https://brooklynworks.brooklaw.edu/bjcfcl/vol11/iss1/3/ at 50.

131 Securities and Exchange Commission, Statement and Guidance on Public Company Cybersecurity Disclosures (2018) available at: https://www.sec.gov/rules/interp/2018/33-10459.pdf.

132 Pacella, *supra* note 130 .

133 Ronicker, *supra* note 129.

134 Zach Whittaker, Lawsuits Threaten Infosec Research—Just When we Need it Most (2018), available at: https://www.zdnet.com/article/chilling-effect-lawsuits-threaten-security-research-need-it-most/.

135 Article 2(1)(a)(x) Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law.

136 Centre for European Policy Studies, Software Vulnerability Disclosure in Europe: Technology, Policies and Legal Challenges (2018), available at: https://www.ceps.eu/publications/software-vulnerability-disclosure-europe-technology-policies-and-legal-challenges, at 42.

137 Savage, *supra* note 109, at 16.

introduce too much subjectivity and uncertainty as to the scope of protections.[138] However, as Savage points out, 'if the definition is restricted to the law, and the law is restrictive in scope, this may deter individuals from raising concerns and provide justifications to not protect those that do'.[139] Arguably, the restrictiveness of laws may become increasingly problematic for potential whistleblowers as technology rapidly advances, widening the gap between ethically problematic practices and illegality as the law struggles to keep up.

15.063 This gap and the question of whether disclosures of 'immoral' conduct should entitle whistleblowers to protection may, for instance, be directly relevant to the joint initiative launched by a whistleblower non-profit 'the Signals Network' in collaboration with international media groups. With a combined audience of 46 million people, this consortium is encouraging whistleblowers who believe that corporations are 'misusing' big data to come forward.[140] Consider the GDPR, which took six years to come into force – although it has been hailed as setting the global standard, some argue that its primary focus on individual privacy rights and the protection of personally identifiable data is already outdated. This on the basis that it fails to 'account for the actual technological landscape unfolding before us', where the scale of big data analysis is such that many of the most powerful applications and risks of harm are directed not at individuals, but at groups.[141] Imagine a data scientist at a start-up who, encouraged by the Signals Network Initiative, came forward disclosing that his/her company was engaged in what he viewed as the use of big data in a way which posed harm to a group.[142] Given the relative underdevelopment of law relating to group privacy harm, it is unlikely to be clear whether the practice disclosed is *unlawful* or 'merely' immoral, and therefore whether whistleblower protections apply. As the law struggles to keep up with technological development and the ethical issues it raises, there will be increased pressure on developing whistleblowing legislation to adopt an expansive approach to the breadth of protection, while intensifying the debate over whether to extend the scope of protection schemes to the reporting of immoral or unethical conduct.

---

138 International Bar Association, *supra* note 4.

139 Savage, *supra* note 109, at 17.

140 The Signals Network, Global Investigation on the Misuse of Big Data (2018) available at: https://thesignalsnetwork.org/press-release/.

141 Linnet Taylor et al., Group Privacy: New Challenges of Data Technologies, available at: https://linnettaylor.files.wordpress.com/2017/01/groupprivacy.pdf, at 3.

142 See, e.g., Christopher Wylie, Why I Broke the Facebook Data Story and What Should Happen Now (2018), available at: https://www.theguardian.com/uk-news/2018/apr/07/christopher-wylie-why-i-broke-the-facebook-data-story-and-what-should-happen-now.

15.064 Another contentious area in the development of whistleblower protection regulation is the relevance, if any, of motive. Many definitions include a 'good faith' requirement.[143] By contrast, the EU's whistleblower protection directive only requires that the individual has 'reasonable grounds'[144] to believe that the wrongdoing disclosed falls within the scope of the regime. To many, this is a welcome development, as the good faith requirement can lead to unnecessary scrutiny of whistleblower's motivations, rather than focus on the wrongdoing itself.[145] However, whether notions of good faith can or should be abandoned in the context of the changing nature of whistleblowers, is open to question. The question of motive is highly relevant in the context of distinguishing leaking and whistleblowing. In the public discourse, 'whistleblowers' are sometimes regarded as defenders of public interest, while 'leakers' are portrayed as selfish, jealous or overtly competitive.[146] Some also argue that the motive considerations can affect the state authority's decision on the merits of the case, even when the motive *de jure* is not relevant.[147]

15.065 Notions of good faith and the relevance of motive are also inherent in the concept of a whistleblower-as-hacker. Countries where the public prosecutor can exercise discretion in pursuing cases throw this into sharp relief. Article 2 of the EU's Cybercrime Directive and Article 3 of the Cybercrime Convention lay down provisions regarding illegal access to information systems. However, the notion of 'ethical hacking' does not exist in the criminal law. In order to decide whether or not prosecution would be in the public interest, prosecutors are relying on assessments of security researchers' 'bona fides' and motives to distinguish 'white-hat' and 'black-hat' hackers.[148] It could be argued that the distinction should instead be drawn by examining the proportionality of the hacker's actions: whether they did more than was necessary to expose the breach. However, this determination is not free from difficulty, and for an ethical hacker this may be impossible to predict in advance. This suggests that an examination of motives and good faith will, in the context of new-age whistleblowers, continue to be relevant.

143 International Bar Association, *supra* note 4.

144 Article 6(1)(a) Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law.

145 Ibid.

146 Leo Wolinski, Leaking, Whistleblowing and the Truth – An Expert's Guide (2017), available at: https://www.sacbee.com/opinion/california-forum/article162331793.html.

147 Papandrea, *supra* note 122, at 38.

148 Centre for European Policy Studies, *supra* note 136.

15.066 Finally, the changing whistleblower is challenging policymakers to confront the cross-border reality of whistleblowing in 2020.[149] Globalisation is both a consequence and reason for whistleblowing, leaking and hacking. It is a consequence in the sense that the online platforms, discussed above, emerged as a way to allow the disclosure of large amounts of information on a safe platform that would be freely available to all. The distinguishing feature of these platforms is that they lack a territorial connection to one given state as they typically rely on a complicated web of interconnected servers, situated in different parts of the world. It is also a reason: In light of a steady rise of cross-border commerce and financial interactions, leaking and hacking have become ways to bring about transparency and accountability as whistleblowers traditionally have, addressing the demand for information about international transactions, especially in the tax-related area.[150] The technologically enabled opportunity for cross-jurisdictional whistleblowing is heightening the urgency of addressing the challenges that come along with it, including legal uncertainty over which jurisdiction's laws apply (which is crucial where there are significant differences in the level of protection offered and issues to do with conflict of laws).[151] Greater cooperation among regulators and law enforcement is needed, as well as, where possible, the harmonisation of laws. The EU Whistleblowing Directive is a positive step in this direction.

15.067 The lines between whistleblowing, leaking and hacking are becoming blurred in the age of information, changing the nature and methods of whistleblowing. Protection of the whistleblowers of the future will depend not only on addressing the political and ethical implications of these developments, but also a worldwide effort to confront the need for protections responsive to the globalised nature of modern whistleblowing.

E. CONCLUSION

15.068 Whistleblowing has changed considerably since the ancient Athenians hailed the important function undertaken by those who drew public attention to private wrongdoing. The most dramatic developments have occurred in the

149 Ashley Savage, Embracing the Challenges and the Opportunities of Cross-jurisdictional Whistleblowing (2018), available at: http://www.oecd.org/corruption/integrity-forum/academic-papers/Savage.pdf.

150 Shu-Yi Oei and Diane M. Ring, Leak-driven Law, 65 *UCLA Law Review* (2018), available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2918550, at 14.

151 Ashley Savage and Richard Hyde, Whistleblowing Without Borders: The Risks and Rewards of Transnational Whistleblowing Networks, in David Lewis and Wim Vandekerckhove (eds) *Developments in Whistleblowing Research* (2015), available at: http://www.track.unodc.org/Academia/Documents/151110 IWRN ebook 2015.pdf.

past half-century, as whistleblowing went mainstream and regulatory protections became increasingly commonplace. Yet, greater societal support for whistleblowers is now coinciding with technological disruption. FinTech, RegTech, SupTech and other innovations have the potential to replace, empower and change whistleblowing, whistleblowers and whistleblower protections. Predicting the future is a notoriously fraught exercise. It is hoped this chapter has provided a helpful summary of recent and forthcoming developments, alongside thought-provoking speculation about 'whistleblowing 2.0'.