

Learning from the laws of the sea, Foucault and regulatory theory: proposing a ‘regulatory harbour’ model for the regulation of social media that serves rather than rules the waves

DR LAURA SCAIFE

Practising Lawyer

It takes time to know the sea, years and those who keep measuring the brief moments of romance against the punishment of mind and body that the sea imposes, soon gravitate towards a sullen hatred which is akin to the hatred of an alien land where none speak your tongue and all customs are strange . . . Accursed be those moments of deadly fear when the ship groans and shudders as the screw races wildly in the hollow of the valleys of the waves, the engines thumping, the anchor chains thundering with a fearful sound . . . and now the cargo shifts, now the ship itself is listing, and still she plunges into the valleys of the waves . . . Have you lived though that ever? . . . How many facets has a gem? The sea has more. (W Kaufman, *The Eternal Sea*)

Abstract

This paper will make a modest case of the regulation of social media. Improved rule-making comes when the means of securing compliance is shaped having regard to the particular problem at hand, rather than by clinging to the notion that rules shape the world. Through a consideration of the regulation of the sea, this paper has served to illustrate that regulation is possible. In the context of social media, where individuals work constantly on their self-identity and perform self-development through expressive activities in public, the use of social media sites can represent what Foucault described as techniques of the self. This paper proposes a theoretical and methodological approach to thinking about the regulation of social media as conceptualised in an iterative and dynamic model that is characterised by the technologisation of human interaction and increasingly transparent ways of living in an age of ‘technologies of the self’. This approach will facilitate a more critical, responsive and iterative awareness of the regulation of expressive content in a way that can grow with the technological state of the art, alive to cultural sensitivities and the use of the tools as a vessel for self-development.

Keywords: social media; regulatory theory; cyberspace regulation; libertarianism; cyberpaternalism; Foucault; internet; waterfall methodology; agile methodology; governance; hybrid modelling

1 Introduction

This paper will make a modest case of the regulation of social media, drawing on the laws of the sea and regulatory theory to propose a ‘regulatory harbour’ model for its regulation. What it takes to make regulation responsive has been extensively researched;¹ latterly a body of work has developed querying the responsiveness of regulation as applied to the internet,² yet the effective regulation of social media remains largely unaddressed,³ characterised as it is by its dynamic quality, spontaneity and reach. Much like the internet, social media as current legal problem is frequently characterised as somehow unique, differing from other seemingly untameable spaces, such as the sea, which have successfully become the subject of regulation. Regulation can be created for anything. Tomorrow, legislators could ban social media in its entirety if they so wished. However, the test of regulation is whether it is equipped to meet the demands of the legal dilemmas which it raises and whether those persons subject to regulation, submit to it. Improved rule-making comes when the means of securing compliance is shaped having regard to the particular problem at hand, rather than by ‘clinging to the notion that rules shape the world’.⁴

My starting point is to consider the vast and continuous expanse of the ocean’s water to be akin to the equally vast extent of global information and social media content. The shape and geography of the oceans have morphed and evolved through the process of plate tectonics, unrecognisably over the millennia, and existed long before any form of regulation or even the first of passage by man-made vessels. However, there have been rules in place to govern participants of sea-transport since about the first millennium BC. Contemporary maritime law is a mixture of ancient doctrines and new laws, both national and international. From oil to tin, diamonds to gravel, metals to fish, the resources of the sea are enormous. The reality of their exploitation grows day by day as technology opens new ways to tap those resources. The point of the comparison is this; despite the changing times and tides, the nature of basic hazards associated with seafaring does not in essentials change, for example, collisions at sea, sinking, smuggling, environmental pollution and piracy. The full list of potential hazards at sea, much like social media (for example, copyright breaches, defamation, criminal speech, breach of privacy rights etc.) is extensive.⁵ My argument is that understanding the development of international maritime law can help prompt a dynamic, dialogue-based model for the regulation of international social media.

1 S Barrett and C Fudge, *Policy and Action* (Methuen 1981); R Baldwin, ‘Why rules don’t work’ (1990) 3(3) *Modern Law Review* 321–37; F Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision Making in Law and in Life* (Oxford University Press 1991); R Baldwin and J Black, ‘Really responsive regulation’ (2008) 71(1) *Modern Law Review* 59–94; R Baldwin and J Black, ‘Really responsive risk-based regulation’ (2010) 32(2) *Law and Policy* 181–213; J Black, *Rules and Regulators* (Oxford Scholarship Online 2012, first published Clarendon 1997); J Black and R Baldwin, ‘When risk-based regulation aims low: approaches and challenges’ (2012) 6(1) *Regulation and Governance* 2–22.

2 Lawrence Lessig, *Code: And Other Laws of Cyberspace*, 2nd edn (ebook, Basic Books 2008) <<http://kentuk.eblib.com/patron/FullRecord.aspx?p=903037>>.

3 There is some work in the field of co-regulation. See Christopher T Marsden, ‘Internet co-regulation and constitutionalism: towards European judicial review’ (2012) 26(2–3) *International Review of Law, Computers and Technology* 211–28, 212. However, there is as of yet no comprehensive attempt to address the overarching regulation of social media in a holistic manner.

4 Baldwin (n 1) 337.

5 This article will not consider these issues individually. For an analysis of the issues as applied to intermediaries see Uta Kohl, ‘The rise and rise of online intermediaries in the governance of the internet and beyond: connectivity intermediaries’ (2012) 26(2–3) *International Review of Law, Computers and Technology* 185–210.

The theoretical context for this paper is provided by the proponents of regulatory theory.⁶ I will argue that such theories are of significant importance in the context of social media communications, given the prevalence, immediacy and internationality of such communication.

The aim of this article is not to examine regulatory theory, or even attempt to apply such a breadth of literature to the technologies and practices characterising the luminous depths of social media platforms powered by Web 2.0.⁷ My aim is to use regulatory theory as a vessel and apply it to social media to formulate a 'really responsive' regulatory approach to a contemporary legal issue.⁸ I draw on software development modelling to propose a 'regulatory harbour' model, for the regulation of social media. This proposed model takes into account the dynamic technological environment, the differential nature of the logics, and the working practices adopted by regulator and regulated to create an iterative model which is built to respond to robust testing and challenges.

The proposed model is as rich and diverse as the oceans: it facilitates tidal waters to flow in and out of protected areas, so that law-abiding vessels which choose to be part of a socially responsible, respectful and self-governing system, based on effective principles and standards, can sail through the harbour entrance as they wish, in accordance with the accepted rules and conventions. The systems and regulation will similarly evolve over time through a process of review, whilst building up customary courses and standards of behaviour based on underlying common principles which should be agreed are desirable of protection and common interest, representing a collaboration between the legislators, developers and the community itself. I will suggest that users and stakeholders could learn to navigate social media with greater aplomb and security, whilst affording technology providers, like those early ship-builders, the necessary latitude to design more sophisticated system architecture and codes, which could again be compared to the evolution of early maritime knowledge, navigational instruments and charts. Through the global adoption and harmonisation of these instruments, explorers will be able to recognise the glorious and expansive opportunities of the social media high seas whilst reducing the risks of being shipwrecked on jagged rocks as they sail through storms, seemingly still waters and violent tempests.

This approach will facilitate a more critical, responsive and iterative approach to the regulation of expressive content, devising a model that can grow with the technological state of the art and that is alive to cultural sensitivities and organisational constraints of both the regulator and the regulatee. It goes further than simply providing a refinement of existing models, demanding an ongoing consideration of the regulatory strategy to be adopted, beginning with principles and problems rather than the regulation itself. The regulation is responsive because it knows the environment which it seeks to *assist* rather than attempting to *command* it, so that it is alive to performance sensitivities and changes in tides, rejecting the notion that regulation is an inherently restrictive force, fostering a regulatory environment in which the internet, the web and social media can exist and have the potential to flourish.

This is only a modest case for regulation because, after advancing the case for the reconceptualisation of social media regulation, much like the oceans, it will not be possible to regulate every aspect of social media, as an omni-directional regulatory system

6 See Baldwin (n 1); Schauer (n 1); Baldwin and Black (2008) (n 1); Baldwin and Black (2010) (n 1); Black and Baldwin (2012) (n 1); Black (2012) (n 1).

7 This paper is confined to a consideration of UK law, and does not consider the role of international law.

8 See generally Baldwin and Black (2008) (n 1).

would rely on all 'social-seafarers' obeying the rules, which is unrealistic. Like the oceans, the content posted on social media platforms has its own uncontrollable weather patterns, currents and tides, which will inevitably produce difficult seafaring conditions, creating the need for safe environments, such as harbours, to be designed and constructed to shelter users from violent tempests so they can engage in commercial, social and competitive activity. Regulation should not attempt or expect pirates or smugglers to choose to be subject to regulation. Rather, rogue groups of people, such as trolls, 'keyboard warriors', criminals or terrorists will not choose to be constrained by international rules and must continue to be pursued by other investigatory and judiciary systems. The conceptual model proposed for the holistic regulation of social media may be valid, but the application of the principal should vary.

2 Social media: community, power and influence

2.1 THE SELF-DETERMINATIVE POWER, INFLUENCE AND ALLURE OF SOCIAL MEDIA

The wide adoption of internet-enabled smartphones has afforded social media ubiquity, raising it to a modern-day religion which many observe devoutly on a daily basis, with its various preachers (finding form as influential bloggers, and celebrity Twitter account holders) and devoted followers. It even enjoys a confessional quality.⁹ Having recently celebrated its 25th birthday,¹⁰ the internet in comparison is still in its historical infancy, yet it represents a powerful medium for the formation of self-identity. In the western technosocialised world, individuals are constantly working on their self-identities in the public realm, performing such self-development through the use of social networking sites which for Sauter 'can be conceptualised as what Foucault described as "*techniques of the self*".¹¹

Social media is not ethereal, it is a man-made construct, building upon the architecture of Web 1.0,¹² yet it is held up as an omnipotent being, becoming for many a significant part of their day-to-day life, afforded a status akin to our basic human needs, quenching our insatiable thirst for knowledge. According to Sauter, users of social media 'figure out how to manage their daily actions and interactions within the context of the complex techno-social hybrid realities they live in, constantly navigating their public appearance and their relation to self and others'.¹³ These globules of information which spill forth, shape and carry our collective and individual identities (for example politics, jurisdiction, legal system, social norms, shared values, religion, history, culture etc.), which are some of the most valuable 'commodities' that we, as human beings possess, are now passing back and forth across the giant uncharted digital oceans that collectively represent social media content generation which takes place on Web 2.0.¹⁴

As much as the then uncharted waters did for the original global explorers and commodity traders that used the oceans to explore new lands and trading routes, the internet allows users to discover new markets, cultures and friends at the click of a smartphone or the click of a 'like' button. For Foucault, practices of the self 'permit

9 T Sauter, T 'What's on your mind? Writing on Facebook as a tool for self-formation' (2014) 16 *New Media and Society* 823–83, 823, 827.

10 <www.w3.org/webat25>

11 Sauter (n 9) 823.

12 L Scaife, *The Handbook of Social Media and the Law*, 1st edn (Informa Law, Routledge 2014) 4–7.

13 Sauter (n 9) 824.

14 See Sauter (n 9) 824: 'On Facebook, people write and update status messages not just as a form of communication, but as a means of shaping understandings of self and establishing normative ways of acting, and sometimes transgressing them.'

individuals to affect by their own means or with the help of others a certain number of operations on their own bodies and souls, thoughts, conduct and a way of being, so as to transform themselves to attain a certain state of happiness, purity, wisdom or immortality'.¹⁵ In a paper written shortly after this exploration, Foucault began to describe these practices as 'technologies of the self'.¹⁶ Social media can be located within this theoretical framework in a very general sense, as they represent a means through which one can undertake 'an exercise of self upon self by which one tries to work out, to transform oneself and to attain a certain mode of being'.¹⁷ Foucault could not have known when he wrote his work that it would so neatly apply to the meteoric rise of Instagram photographs of Avocados,¹⁸ 'selfies'¹⁹ and 'clean eating' trends.²⁰

In the Western techno-socialised world, social networking sites (SNSs) 'can be conceptualised as what has been described as "*techniques of the self*".²¹ As noted by Foucault, self-writing as borne through historical analysis has been a relational practice that leads to an 'intensification of social relations'.²² Van Kreiken suggests that 'as social life becomes denser, more competitive, more highly differentiated and more dependent on a variety of means of indirect communication, visibility beyond ones immediate circle of face-to-face contacts becomes increasingly significant'.²³ Indeed, 'the ease of access to these tools for most in the developed world, and the immediacy with which others respond to, and thus provide guidance to, modern self-writers adds a new facet to the age-old practice of self-formation'.²⁴ However, in making their private lives public 'and revealing themselves though their semi-public online self-writing, users expose themselves to an expanded and less clearly defined readership, "inviting, albeit cautiously, a certain level of surveillance"'.²⁵

The quest for self-reflection and its relationship with water is not new. In Greek mythology, Narcissus, who was known for his beauty, was lured by Nemesis to a still pool in which he saw his own reflection in the water and fell in love with it, not realising it was merely an image. Unable to leave the beauty of his reflection, Narcissus lost his will to live. He stared at his reflection until he died. Modern self-writers are lured by social media's own reflective capacities, searching deeper still beyond the water's reflection to understand more about themselves and the community in which they form their self-identities. Walter Kaufman, in *The Eternal Sea* wrote of its powers: 'the sea knows all moods of man, and she imposes them all on those who venture out in ships: serenity and rage, gravity and vehemence, jocundity and gloom, charity and malice – the sea is like no

15 M Foucault, *Discipline and Punish: The Birth of the Prison*, Alan Sheridan (trans) (Vintage 1977) 18.

16 M Foucault, *Technologies of the Self*, in L H Martin, H Gutman and P H Hutton (eds), *Technologies of the Self: A Seminar with Michael Foucault* (Tavistock 1988).

17 M Foucault, *The Care of the Self: The History of Sexuality*, vol 3 (Random House 1986) 2.

18 'Avocados now the Oprah of Instagram' (*New Zealand Avocado*, 7 August 2015) <www.nzavocado.co.nz/newsroom/Avocados_now_The_Oprah_of_Instagram>.

19 D Paulhus and K Williams, 'The Dark Triad and trait self-objectification as predictors of men's use and self-presentation behaviors on social networking sites' (2015) 76 *Personality and Individual Differences* 161–5.

20 Hadley Freeman 'Green is the new black: the unstoppable rise of the healthy-eating guru' *The Guardian* (London 27 June 2015) <www.theguardian.com/lifeandstyle/2015/jun/27/new-wellness-bloggers-food-drink-hadley-freeman>.

21 Fox and Rooney (n 19).

22 Foucault (n 17) 53.

23 K Van Kreiken, *Celebrity Society* (Routledge 2012) 138.

24 Sauter (n 9) 833.

25 Ibid.

man, and yet she is like humanity boundless as life'.²⁶ These words could have been written in relation to social media platforms and apps, which facilitate such information-sharing and are adopted by users to navigate new worlds, charting progress through both rough and smooth waters, regardless of geographical boundaries. In 2012 complaints to police about alleged crimes linked to the use of Facebook and Twitter increased by 780 per cent in four years, resulting in about 650 people being charged. In the context of the sea, Conrad contends 'if you would know the age of the earth, look at the sea in a storm'.²⁷ For the purposes of this paper, if you would know the infancy of the application of regulation to social media, look to examples of human nature at its worst as played out on this global stage. For instance, the amount of negative 'trolling' aimed at high-profile figures, including female UK Members of Parliament²⁸ and high-profile feminist campaigners²⁹ has led to the formation of the Reclaim the Internet campaign³⁰ and discussions as to what extent the reform of the legislation relating to harassment need to be addressed³¹ in the light of anonymity and the problems presented by jurisdiction. In May 2016 the UK's referendum concerning membership of the EU became an active resource for voicing political views,³² leading to social media platforms becoming a vessel for political road rage.³³ It also led to a rise in reported hate crime.³⁴ It could be argued that such consequences are a risk of engaging with the medium, when one chooses to submit information to the public realm. However, other examples such as contempt of court, incitement to riot during the 2012 London riots, football hooliganism and racist speech highlight that the content played out on social media fall squarely within the remit of the criminal law.³⁵ Recently, the Crown Prosecution Service released figures which show that there were 4908 reports in which Facebook and Twitter were a factor in reported crimes, compared with 556 in 2008.³⁶

One area which throws into sharp focus the limitations of existing regulation is social media's relationship with terrorism. There has been a marked increase in the use of social media to glorify, encourage the commission of and/or provide support for terrorist-related offences, notably the organisation known as the Islamic State of Iraq and Levant (ISIS) which has posted violent video clips of terrorist attacks and beheadings on social

26 W Kaufman, *The Eternal Sea* (Druckerei Fortschritt Erfurt 1970) 9.

27 Ibid 10.

28 See Scaife (n 12) 147–50.

29 Ibid 147–48; *R v Isabella Sorley and John Nimmo*, Westminster Magistrates Court 24 January 2014 (unreported).

30 <www.reclaimtheinternet.com>

31 Scaife (n 12) 182.

32 Vyacheslav Polonski, 'Analysing the social media voices in the UK's EU referendum' (15 May 2016) <<https://medium.com/@slavacm/social-media-voices-in-the-uks-eu-referendum-brexite-or-bremain-what-does-the-internet-say-about-ebb7b27cf0f#gg68vebd0>>.

33 Rhiannon Lucy Cosslett, 'Family rifts over Brexit: 'I can barely look at my parents' *The Guardian* (London, 27 June 2018) <www.theguardian.com/lifeandstyle/2016/jun/27/brexit-family-rifts-parents-referendum-conflict-betrayal>. Alex Wood, '52% blamed EU for their problems, blame social media for yours' (*The Memo*, 29 June 2016) <www.thememo.com/2016/06/29/brexit-social-media-eu-half-blamed-the-eu-for-their-problems-blame-social-media-for-yours>.

34 Chris York, 'Post-Brexit racism documented on social media' (*Huffington Post*, 2 July 2016) <www.huffingtonpost.co.uk/entry/post-brexit-racism_uk_5777be69e4b073366f0f1d06>. Examples of the postings can be viewed in a Facebook Album entitled 'Worrying Signs' created to document alleged incidents in which people have been targeted with xenophobic comments, accessible via <www.facebook.com/sarah.leblanc.718/media_set?set=a.10101369198638985&type=3&pnref=story>.

35 Scaife (n 12) 177, 182.

36 CPS, 'Violence against Women and Girls Crime Report 2015–2016' <www.cps.gov.uk/publication/cps-violence-against-women-and-girls-crime-report-2015-2016>. Of course, it is debatable whether the number of crimes has increased, or if the 'crimes' are now reported with greater frequency.

media. Such content was available via sites like Archive.org during the infancy of the data revolution. However, these were only accessed by the hardened core who knew where to look for such materials. The new wave social media, like YouTube, Facebook or Twitter, allowed AQI (Al-Qaeda in Iraq) to reach a much wider audience.³⁷ However, for every video seeking to glorify terrorist activity,³⁸ there are many more that, although posted by terrorists or terrorist sympathisers, fall short of inciting terrorist activity in its legal sense, or indeed deal with matters that do not amount to terrorist activity.³⁹ Yasser Arafat, late chair of the Palestine Liberation Organization, captured the issue in a 1974 speech before the UN, declaring '[O]ne man's terrorist is another man's freedom fighter.' Such examples demonstrate that drawing the line between open comment in a community and speech that warrants a regulatory response is not capable of being achieved with the precision of a surgeon's knife, nor draftsmen's pen.

3 Approaches to enforcement and responsive regulation

3.1 DO WE WANT REGULATION?

A critique of governmental and legal regulatory interventions, adopting universalist free speech arguments,⁴⁰ could be advanced that such public entities have no place in responding to individual choices.⁴¹ Examples of speech, as set out above, may be undesirable, but it could be argued that such is the consequence of freedom of choice.⁴²

The concept of a truly free flow of information facilitated by Web 2.0 is on its face a utopian ideal, but this does not accord with the creation of the internet from which the

37 ISIS' first official Twitter account was set up in October 2010 under the name 'al Itisaamm' with the handle @e3tasimo, which encapsulates the notion of maintaining Islamic tradition without deviation. A key figure in the media tide turn for ISIS was a user operating under the handle @reyadiraq. @reyadiraq was steadily posting more gruesome content, departing from doctrinal matters preferred by the official account to the live tweeting of an amputation of an accused thief's hand. See J M Berger, Twitter post 10 January 2014 <<https://twitter.com/intlwire/status/421707838021316609>>. This link has subsequently been dead-linked and flagged for inappropriate content and is no longer generally accessible.

38 With regards to terrorism-related offences specifically, s 1 Terrorism Act 2006 criminalises the 'encouragement of terrorism' which includes making statements that glorify terrorist acts, as well as disseminating terrorist publications under s 2 of the Act, which is punishable by up to seven years' imprisonment

39 The difficulty of content filtering is usefully explored though consideration of the postings of THE radical Muslim cleric, Adnan Awlaki (deceased), who has an enduring active presence on YouTube; with videos covering everything from daily teachings right through to inciting terrorist behaviour. It raises questions as to whether materials about what makes a good marriage, his account of the final moments of the Prophet Muhammad or Awlaki's counsel on the proper diet for a good Muslim should be removed from the internet. These videos do not contravene any particular YouTube standards or law and his postings have also been used in newscasts by CNN and Al Jazeera. Although not directly advocating terrorism in these particular videos, there is evidence that Mr Awlaki's sermons were the gateway materials that lured in individuals who went on to commit violent terrorist activity. An example of this is the 21-year-old British student who confessed to police that she had stabbed a member of the UK Parliament after watching over one hundred hours of Awlaki videos. And, according to a senior source investigating the case 'she was inspired by his sermons, and radicalised by watching them. His message is doing anything, whatever you can.' Quoted in Stephen Timms, 'Attacker guilty of attempted murder' *The Guardian* (London, 2 November 2010) <www.theguardian.com/uk/2010/nov/02/stephen-timms-attacker-guilty>.

40 J Raz, 'Free expression and personal identification' (1991) 11 *Oxford Journal of Legal Studies* 303.

41 Cass R Sunstein, *Flexibility. Regulation: Looking Backward, Looking Forward*, Administrator, Office of Information and Regulatory Affairs Administrative Law Section (American Bar Association, Capital Hilton DC, 10 May 2012) 152; J Perry-Barlow, 'The economy of selling ideas: selling wine without bottles on a global net' (n.d.) <www.eff.org/EconomyOfIdeas.html>.

42 Debora L Spar, *Ruling the Waves: Cycles of Discovery, Chaos, and Wealth from the Compass to the Internet* (Harcourt 2001) 9.

environment for Web 2.0 developed. The internet was a public sector creation, which only subsequently came under private control.⁴³ Relatedly, the regulation though which proprietary, civil and criminal rights are conferred and defended fosters an environment in which the development of the internet can exist. It has been commonly advanced that regulation is a curb on free speech, producing a chilling effect, but this is not the case. For example, without competition law to govern anti-competitive practice and allow for marketing entrants to create platforms and products there would not be a range of new products for consumers to take advantage of and, without intellectual property rights, platform providers could not define and defend their proprietary rights in the software they create, thus monetising their intellectual property. Without legal protection of any kind, according to Sunstein 'all sides would be left with the struggle to show superior force'.⁴⁴ In such an environment only the very strongest player would win the battle for dominance and with such a monopoly could curb privacy and freedom of expression if the market did not allow for freedom of competition. The question is not: should we have regulation? But instead: what sort of regulation should we have?⁴⁵

3.2 WHY RULES DON'T WORK

Why rules fail to work and how they may be more effectively designed to meet the particular regulatory challenge they seek to address has been the subject matter of extensive debate.⁴⁶ The test of the success of regulation is whether it meets the challenges faced by today's regulators.⁴⁷ Black and Baldwin posit that for regulation to be effective:

... regulators have to be responsive not only to the compliance performance of the regulated; to the firms own operative and cognitive frameworks; their attitudinal settings; to the broader institutional environment of regulatory regime; to the different logics of regulatory tools and strategies; to the regimes own performance; and finally to changes in each of these elements.⁴⁸

The criticisms of regulation broadly fall into three groups: the policy or the conceptual; the practical; and the principles. In essence the summation of the critique of all three of these approaches is that they presuppose that top-down regulation is possible, that escalation and de-escalation are possible, and that the regulated will submit to regulation. Assuming always that each 'player' within the regulatory model understands the 'rules of the game'.

43 Sunstein (n 41) 157.

44 Ibid 156.

45 Cass R Sunstein, *Republic.com 2.0*. (Princeton University Press 2007) 160. Murray articulates the question thus: "The starting point for this is to return to the root of the subject and ask the question "what is regulation"?" Andrew Murray, 'Nodes and gravity in virtual space' (2011) 5(2) *Legisprudence* 195–221, 214.

46 Black (2012) (n1); Baldwin (n 1); Christopher Ham and Michael Hill, *The Policy Process in the Modern Capitalist State*, 2nd edn (Harvester Wheatsheaf 1993); Michael Hill (ed), *The Policy Process: A Reader* (Harvester Wheatsheaf 1993).

47 Black and Baldwin (2008) (n 1) 59.

48 Ibid 61.

3.3 THE LIMITATIONS OF LOCATING SOCIAL MEDIA WITHIN EXISTING REGULATORY FRAMEWORKS

The regulation of social media content⁴⁹ has been shoe-horned into a range of existing statutes,⁵⁰ leading some to comment that ‘even where the law is clear, many people have raised concern about the police ability to pursue online crime. And in different cases the police have been criticised for failing to take threats seriously, but also for pursuing cases that shouldn’t be prosecuted’.⁵¹

The Communications Act 2003 demonstrates the limitations of determinative regulation and subsequent attempts to interpret it in the context of social media.⁵² Section 127 which deals with messages of an offensive, obscene or menacing character sent via a public electronic communications network, has been used as a means to seek to regulate offensive speech based on societal standards. Under section 127(1) of the Act, the defendant must be shown to have intended or be aware that the message sent was grossly offensive, indecent or menacing, which can be inferred from the terms of the message or from the defendant’s knowledge of the likely recipient. The knowledge of the recipient’s likely reaction is only relevant when making inferences about the defendant’s intention, and not as to whether the message itself was grossly offensive.⁵³ Thus, the offence is committed by sending the message, there is no requirement that any person sees the message or has to be offended by it.⁵⁴ Such a rule, as noted by Black and

49 This essay considers criminal content and privacy matters. A review of the specific laws relating to civil offences such as defamation, particularly the Defamation Act 2013, is beyond the scope of this essay. For a consideration of the public and private sectors’ approach to the regulation of defamatory content online, notably the website operator’s defence contained in Defamation (Operators of Websites) Regulations SI 2013/3028, see Scaife (n 12), ch 3.

50 For example, in the UK cases involving criminal behaviour can be prosecuted under an umbrella of existing legislation including the Protection from Harassment Act 1997 (see *Review of the Protection from Harassment Act 1997: Improving Protection for Victims of Stalking* (Home Office July 2012) and the CPS’s Stalking and Harassment Guidelines <www.cps.gov.uk/legal-guidance/stalking-and-harassment>), the Malicious Communications Act 1988, the Crime and Disorder Act 1998 (see *R v Cryer* (Unreported) 21 March 2012, Newcastle Magistrates’ Court), the Public Order Act 1986 (see *R v Stacey, Swansea Crown Court* (unreported); *R v Stacey*, Appeal No A20120033, 30 March 2011), the Serious Crime Act 2007 (see *R v Blackshaw* [2011] EWCA Crim 2312; *R v Perry John Sutcliffe-Keenan* [2011] EWCA Crim 2312; *R v Ahmad Pelle*, 25 August 2011, Nottingham Crown Court (unreported); *R v Hollie Bentley*, 29 November 2011, Leeds Crown Court (unreported)), and, the most commonly used provision, s 127(1)(a) Communications Act 2003 (see *DPP v Chambers* [2006] UKHL 40). Doubts have also been expressed as to whether creating a webpage or social network group constitutes ‘sending’ a message (see Policy Memorandum on the Offensive Behaviour at Football and Threatening Communications (Scotland) Bill (2011) at [34]; there are a number of examples where s 127 has been used against internet communications).

51 See n 30.

52 The Communications Act 2003 is only one example of the limitations of the existing law: for further examples both civil and criminal, see Scaife (n 12), part II.

53 See D Ormerod, ‘Telecommunications: sending grossly offensive message by means of public electronic communications network’ [2007] Criminal Law Review 98.

54 In contrast with its predecessor subsections, which require proof of an unlawful purpose and a degree of knowledge, s 127(1) (a) Communications Act 2003 provides no explicit guidance on the state of mind which must be proved against a defendant to establish an offence against the subsection. This inevitably raises the question as to what, if anything, must be proved beyond an intention to send the message. In *DPP v Collins* [2005] EWHC 1308 (Admin); [2006] UKHL 40; [2006] 1 W.L.R. 2223; [2006] 4 All ER 602; (2007) 1 CrAppR 5; (2007) CrimLR 98; *The Times* 21 July 2006, counsel for the DPP, relying by analogy on s 6 (4) of the Public Order Act 1986, suggested that the defendant must intend his words to be grossly offensive to those to whom they relate, or be aware that they may be taken to be so. For further discussion in the context of the facts of *DPP v Collins* [2006] UKHL 40, see his Lordship’s comments at para [10] of the judgment. However, this still does not provide a definitive answer as to when an individual may take gross offence and how the statute is to be interpreted.

Hart in the wider regulatory context, is therefore reliant upon its interpretation by others.^{55&56}

As noted by Black, in making a generalisation (speech should be regulated), the rule-maker has chosen certain qualities which the event possesses (i.e. speech sent via a public electronic communications network) from a range of individual properties (the speech in question is of an offensive, obscene or menacing character) and searches for aspects which are ‘causally relevant to the aim of the rule: the goal which is sought to be achieved or the harm which is sought to be avoided’.⁵⁷ It is here that continuing to apply statutes such as the Communications Act 2003 to social media reveals its inherent weakness as the context in which interactive social media dialogue takes place is quite different to the context in which other communications take place – access is ubiquitous and

55 Black (2012) (n 1) 13; H L A Hart, *The Concept of Law*, 2nd edn (Clarendon 1997) 123.

56 In *DPP v Collins* (n 54) a case pre-dating social media, Lord Bingham considered that the views of Lord Reid in *Sweet v Parsley* fell to be considered in the case as Parliament could not have intended to criminalise the conduct of a person using language which is, for reasons unknown to him, grossly offensive to those to whom it relates, or which may even be thought, however wrongly, to represent a polite or acceptable usage. The issue as to whether offence was meant to be caused and how to prove the intent element of the offence, led their Lordships in *DPP v Collins* to consider cases concerning statutory interpretation for cases where the mental element of the offence is not provided for in the statute. Considering the leading case of *Sweet v Parsley* [1970] AC 132 at [148], their Lordships considered that the court’s duty was to consider the words of the statute in question: if they showed a clear intention to create an absolute offence that is an end of the matter, but such cases are very rare. More frequently, the words of the section usually make it clear that *mens rea* is required in one form or another. However, in a large number of cases there is no clear indication either way. In such cases there is a longstanding presumption that Parliament did not intend to make criminals of persons who were in no way blameworthy in what they did. That means that whenever a section is silent as to *mens rea*, as is the Communications Act 2003, there is a presumption that, in order to give effect to the will of Parliament, the courts must read in words appropriate to require *mens rea*. On the other hand, Lord Reid considered that a culpable state of mind would ordinarily be found where a message was couched in terms showing an intention to insult those to whom the message relates or giving rise to the inference that a risk of doing so must have been recognised by the sender. The court considered that the same would be true where facts known to the sender of a message about an intended recipient render the message peculiarly offensive to that recipient, or likely to be so, whether or not the message in fact reaches the recipient: *DPP v Collins* [2006] UKHL 40 at [10]. The court stressed that individuals are entitled to express their views strongly and that the proper question for determining if s127(1)(a) Communications Act 2003 had been infringed was whether the language used went beyond what could be considered as tolerable in society. It was subsequently affirmed in *DPP v Chambers* [2012] EWHC 2157; 2012 WL 2923016, [38], that the European Court of Human Rights has long adopted the view in its significant body of jurisprudence that an individual’s freedom of expression includes the right to say things or express opinions ‘that offend, shock or disturb the state or any sector of the population’ (*Handyside v United Kingdom* App No 5493/72 A/24 (1976); 1 EHRR 737, at [49]). The UN Special Rapporteur on Freedom of Expression has asserted the same approach, see Frank La Rue, Report of the Human Rights Council Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, A/HRC/17/27, 16 May 2011, at para 2. The report also acknowledges that the internet has created challenges to the right of all individuals to seek, receive and impart information and ideas of all kinds through the internet and considers key trends (see para [37] of the report) and that the intention of s 127(1) Communications Act 2003 was not to create interference with the essential freedoms of speech and expression enshrined in Article 10 of the European Convention on Human Rights (ECHR) (*DPP v Collins* [2006] UKHL 40) at [28].

57 F Schauer (n 1) 17–38.

instantaneous.⁵⁸ Banter, jokes and offensive comments are commonplace and often spontaneous and communications intended for a few may reach millions.⁵⁹

58 As Eady J stated in the civil case of *Smith v ADVFN* [2008] EWHC 1797 (QB) at [14], in relation to comments on an internet bulletin board: '[they are] like contributions to a casual conversation (the analogy sometimes being drawn with people chatting in a bar) which people simply note before moving on; they are often uninhibited, casual and ill thought out; those who participate know this and expect a certain amount of repartee or "give and take".'

59 In *DPP v Collins*, Lord Bingham stated that it was plain from the terms of s 127(1) (a), as of its predecessor sections, that the proscribed act, the *actus reus* of the offence, is the sending of a message of the proscribed character by the defined means. The offence is complete when the message is sent. Thus, it can make no difference that the message is never received, for example because a recorded message is erased before anyone listens to it. On such an approach, criminal liability would turn on an unforeseeable contingency. This is clearly problematic in the social media context as there is a lack of control once a posting is made, as postings will be saved on the profile on which 'User A' posts them (i.e. 'User B's' profile) and on the SNS server. The information may also be stored on A's computer if A used the World Wide Web to make the posting. Analysis of a computer's hard-drive may sometimes reveal the contents of SNS communications, but normally that information would have to be acquired by either accessing the profile itself online, or seeking the information from the relevant SNS which stored it on its server. Even if the poster subsequently deletes the post, if *Collins* is applied, the offence has already been committed regardless of whether the individual subsequently regrets their actions and deletes the posting. For example, in contrast, in *Redmond-Bate v DPP* (Divisional Court, 23 July 1999), prosecuted under the Public Order Act 1986, Sedley LJ emphasised that the mere fact that words were irritating, contentious, unwelcome and provocative was not enough to justify the invocation of the criminal law unless they tended to provoke violence. In a similar vein in *Dehal v CPS* [2005] EWHC 2154 Moses J, referring to s 4A of the Public Order Act 1986, held that, at [5]: 'the criminal law should not be invoked unless and until it is established that the conduct which is the subject of the charge amounts to such a threat to public order as to require the invocation of the criminal as opposed to the civil law.' It is suggested that from cases such as *DPP v Collins* censorship in the context of (racist) hate speech, where, despite unpopularity or otherwise, the speech is outlawed (see Public Order Act 1986, Pt III and in particular s 18; see also Crime and Disorder Act 1998, ss 28(1)(b) and 31(1)(c) in relation to s 5 Public Order Act 1986. See also Art 17 ECHR, discussed in *Norwood v UK* (2005) 40 EHRR SE11; *Glimmerveen and Hagenbeek v Netherlands* (1982) 4 EHRR 260; *DPP v Collins* [2006] UKHL 40; [2006] 1 WLR 2223. In international law, see Art 4 Convention on the Elimination of all Forms of Racial Discrimination 660 UNTS 195, entered into force 4 January 1969; Art 20(2) International Covenant on Civil and Political Rights (ICCPR) 993 UNTS 3, entered into force 3 January 1976. Steve Foster, 'Free speech, insulting words or behaviour and art 10 of the European Convention on Human Rights' (2004) 9(1) Coventry Law Journal 68, 71; F Schauer, *Free Speech: A Philosophical Enquiry* (Cambridge University Press 1982) 3–15, repression may be acceptable when that suppression based on the 'rights of others' because it accords with the principles underlying free speech), pro-civility cannot, however, be an acceptable starting point to determine criminal liability in every case of distasteful speech which does not necessarily amount to hate speech and may merely be distasteful rather than grossly offensive (whatever that may mean, given the varying approaches taken in the case law). The lack of control which users are subsequently able to exert over comments is not fully taken into account by the court on sentencing, which instead focuses its attention on the reaction of those who read the posts (which is the point at which the offence is committed) and use of privacy settings. This is significant because, as noted above under the Communications Act 2003, those who are grossly offended by the message or posts in question need not be the intended recipients (see also *DPP v Collins* above) and raised the issue of how sophisticated account holders' knowledge is of the regulatory landscape of sites which are sent as an extension of the organisation and fulfilment of their social lives which may be perceived by them as 'private'. As noted by Baldwin and Black (2008) (n 1) 80, sometimes regulatory effectiveness is affected due to 'divergences of understanding between the judiciary and the regulated can also prove to be a problem – notably regarding the purposes and the objectives of the regulation at issue. In cases where there are unresolved disagreements on the meaning of compliance, this renders the activity extremely fraught.'

In *DPP v McConnell*,⁶⁰ a pastor was acquitted of grossly offending Muslims in a sermon that he delivered, which was also transmitted over the internet.⁶¹ Pastor McConnell also referred to Allah as a heathen, cruel and demon deity. Because the sermon was transmitted over the internet the defendant was charged under the Communications Act 2003 with having sent a ‘grossly offensive’ message. Consistent with *Chambers*, DJ McNally found that section 127 was not designed to protect people against receipt of unsolicited messages which they may find seriously objectionable. Indeed, in relation to social media the court noted that:

The legislators can be forgiven for not foreseeing the arrival a short time thereafter of Facebook and Twitter the widespread use of which has resulted in an increasing number of prosecutions under section 127(1) and, in turn, a parallel increase in criticism of the section as an interference to the right of freedom of expression. Specifically it has been criticised as a widely drafted law designed: (a) primarily to regulate one to one communications rather than one to many; and (b) to safeguard a public utility built with public money, which is now being applied to a privately owned, publicly accessed, many to many domain.⁶²

Instead, the court found that its purpose was to prohibit the use of a service provided and funded by the public for the benefit of the public for the transmission of communications which contravene the basic standards of society.⁶³

The UN has noted that there has been a large rise in social media ‘crimes’,⁶⁴ with some countries issuing guidance on prosecuting cases involving communications sent via social media.⁶⁵ It was partly as a result of the mixture of decisions arrived at by prosecutors that, in the UK, the Department for Public Prosecutions recognised the inherent difficulties in prosecuting cases involving a social media element and stated that, in order to ensure that Crown Prosecution Service (CPS) decision-making in these difficult cases is clear and consistent, the CPS would issue guidelines on social media cases for prosecutors, designed to assist those tasked with whether criminal charges should be brought in the cases that arise for their consideration.⁶⁶ The result of these

60 *DPP v McConnell* [2016] NI Mag 1

61 In the sermon based on the Book of Timothy, the theme of which was that there was one mediator between God and Man, the Pastor declared that an increasing number of Muslims were ‘putting the Koran’s hatred of Christians and Jews alike into practice’, and the sermon had continued in a similar vein. It should be noted that para [9] of the judgment in *DPP v McConnell* stated: ‘it never entered his head that he would be offending anyone and all he wanted to do was to present the truth. He was completely and totally unaware he had caused offence until he was contacted by Steven Nolan of Radio Ulster a few days later. On being interviewed by Mr Nolan he apologised to any Muslim in Belfast who was offended. It was a sincere apology but he was not apologising for the Gospel. When he said that Islam was satanic, heathen and spawned in hell he was not being gratuitously offensive. He was attacking the doctrine and theology of Islam.’

62 *DPP v McConnell* [2016] NI Mag 1, para [16].

63 *Ibid* [23].

64 In England and Wales, for example, in 2008, there were 556 reports of alleged social media crimes with 46 people charged. In 2012, there were 4908 reports with 653 people charged. See <www.bbc.co.uk/news/uk-20851797>. In Western Asia, a number of recent criminal cases related to internet social media content have also been reported: see <www.bbc.co.uk/news/world-asia-china-16629278>.

65 CPS, Interim Guidelines on Prosecuting Cases Involving Communications Sent via Social Media, Director of Public Prosecutions, 19 December 2012. These Guidelines are discussed in detail below.

66 On drawing up Guidelines for the CPS. Keir Starmer recognised the difficulty in determining whether something was ‘grossly offensive’ when he said: ‘The distinction between offensive and grossly offensive is an important one and not easily made. Context and circumstances are highly relevant and as the European Court of Human Rights observed in the case of *Handyside-v-UK* the right to freedom of expression includes the right to say things or express opinions “that offend, shock or disturb the state or any section of the population”’ quoted in *DPP v McConnell* [2016] NI Mag 1, [22].

considerations arise the CPS's *Guidelines on Prosecuting Cases Involving Communications Sent via Social Media*⁶⁷ (the Guidelines). Guidelines have also been issued by the CPS with regards to stalking and harassment in the social media context.⁶⁸ Whilst it is acknowledged that these measures have sought to finesse the application of existing statutes to the regulation of social media, such innovations still operate in a top-down fashion,⁶⁹ pre-supposing that such a methodology is desirable and can lead to effective and determinative outcomes.⁷⁰ As noted by Baldwin⁷¹ with regard to policy more generally, it is usually 'policy centred' rather than action centred,⁷² which when 'handed down' for implementation, draws an 'unrealistic distinction between policy making and

67 <www.cps.gov.uk/legal-guidance/social-media-guidelines-prosecuting-cases-involving-communications-sent-social-media>

68 CPS, Stalking and Harassment Guidelines <www.cps.gov.uk/legal-guidance/stalking-and-harassment>. The social media guidelines need to be read in conjunction with the Code for Crown Prosecutors. This identifies public interest factors which include whether the offence was motivated by any form of discrimination against the victim, including gender discrimination. For a discussion of the guidelines see Scaife (n 12) 182. The CPS launched a public consultation in March 2016 about the proposed revisions to the updated Social Media Guidelines including consideration of whether a new section on 'Violence against girls and women offences' should be included, to reflect that developments in technology have created a new landscape for controlling, sexually motivated or other forms of interpersonal offending. According to the CPS (n 36): 'the use of the internet, social media platforms, emails, text messages, smartphone apps, spyware and GPS tracking software to humiliate, control and threaten victims is rising' (p. 21). The consultation closed in May 2016 and the finalised guidelines are available at <www.cps.gov.uk/legal-guidance/social-media-guidelines-prosecuting-cases-involving-communications-sent-social-media>.

69 The CPS Guidelines state: 'These guidelines set out the approach that prosecutors should take when making decisions in relation to cases where it is alleged that criminal offences have been committed by the sending of a communication via social media. The guidelines are designed to give *clear advice* to prosecutors who have been asked either for a charging decision or for early advice to the police, as well as in reviewing those cases which have been charged by the police. Adherence to these guidelines will ensure that there is a consistency of approach across the CPS. *These guidelines are primarily concerned with offences that may be committed by reason of the nature or content of a communication sent via social media. Where social media is simply used to facilitate some other substantive offence, prosecutors should proceed under the substantive offence in question.*' (emphasis added) As can be seen from the introductory text, no attention is paid to the unique nature of social media itself, presupposing that existing regulation can be applied to social media so long as the guidance is followed. In *Chambers* [2006] UKHL 40 at [28], however, the court stated: 'the 2003 Act did not create some newly minted interference with the first of President Roosevelt's essential freedoms – freedom of speech and expression. Satirical, or iconoclastic, or rude comment, the expression of unpopular or unfashionable opinion about serious or trivial matters, banter or humour, even if distasteful to some or painful to those subjected to it should and no doubt will continue at their customary level, quite undiminished by this legislation.' See Scaife (n 12) 139–52 for a discussion of the guidelines and 153–5 for a discussion of the weaknesses of the guidelines and how they could be refined.

70 The CPS Guidelines on Prosecuting Cases Involving Communications Sent via Social Media also remind prosecutors that, under the Malicious Communication Act 1988 and Communications Act 2003, the law only applies to communications of a grossly offensive nature. Reiterating the findings of the court in *DPP v Chambers*, the CPS states that this means that a communication has to be more than simply offensive to be contrary to the criminal law. Just because the content expressed in the communication is in bad taste, controversial or unpopular, and may cause offence to individuals or a specific community, this is not in itself sufficient reason to engage the criminal law. As Lord Bingham made clear in *DPP v Collins* [2006] UKHL 40 at [9], there can be no yardstick of gross offensiveness 'otherwise than by the application of reasonably enlightened, but not perfectionist, contemporary standards to the particular message sent in its particular context'. It is suggested that the courts, like their Strasbourg counterparts, need to demonstrate a willingness to take into account the experience of the speaker, e.g. an established broadcaster will be held to a higher standard than a member of the public (*R (Gaunt) v Office of Communications* [2010] EWHC 1756 (QB), [2011] ACD 17, [2010] HRLR 31, [2011] 1 WLR 663) which is not something which is considered in the CPS guidelines.

71 Baldwin (n 1) 332.

72 Barrett (n 1) 13.

implementation',⁷³ which is why the products of policy, such as the Guidelines, are not effective in terms of meeting their desired outcomes.

As exemplified by the Communications Act 2003, in adopting forms of top-down regulation and applying them to newly charted worlds by reference to the comfort of the familiar, there is an inherent danger that only some features of the event become the focus of the rule and 'are then projected onto future events, beyond the particulars which served as the paradigm or archetype for the formation of the generalization'.⁷⁴ This creates a risk of under/over generalisation in regard to matters that may be irrelevant and leads to a risk that future developments may make the rule less relevant.⁷⁵ Such an approach therefore neglects, 'going back to basics', to understand 'the nature of the instrument which is being used, the properties of rules and their inherent limitations, and to see whether we can gain insights from this analysis which would enable us to make better use of rules as a regulatory technique'.⁷⁶ Thus, 'even rules that seem now to be neither under- or over-inclusive with respect to their background justifications retain the prospect of becoming so'.⁷⁷ Black suggests that this can be achieved in three ways: firstly, though a better appreciation and use of different types of rules; secondly, an understanding of the context in which rules operate; and, finally, considering these issues with reference to the nature of the regulatory community and the style of regulation.⁷⁸

Approaching the regulation of emerging technology in a top-down, determinative manner is not a satisfactory starting point for the regulation of social media, or for creating effective regulation more generally. As noted by Black, one of the problems associated with the creation of rules in any context, are 'their tendency to over or under inclusiveness, their indeterminacy, and their interpretation'.⁷⁹ She observes that many of the issues associated with effective regulation stem from the prescriptive nature of rules as 'anticipatory, generalised abstractions' which when 'endowed with legal status are distinctive, authoritative forms of communication'. In the context of the development of disruptive technology, Spar notes that 'although first technology challenges authority for some period of time, but then ironically, [it] seems to invite authority back in'.⁸⁰ The question is whether it should be let back in and with what design. Spar's rule modelling applies well to social media as, although there is some governmental control, it is

73 Ibid Parts 1 and 3; P Knopf H Weidner, 'Formulation and implementation of air quality control programmes' (1982) 10 Policy and Politics 198–285; Hill and Ham (n 46) 101–8.

74 Black (2012) (n 1) 8.

75 Black notes, *ibid*: 'that this mismatch can occur for three reasons. First, as noted, the generalization which is the operative basis of the rule inevitably suppresses properties that may subsequently be relevant or includes properties that may in some cases be irrelevant. Secondly, the causal relationship between the event and the harm/goal is likely to be only an approximate one: the generalization bears simply a probable relationship to the harm sought to be avoided or goal sought to be achieved. Thirdly, even if a perfect causal match between the generalization and the aim of the rule could be achieved, future events may develop in such a way that it ceases to be so.'

76 Ibid.

77 Schauer (n 1) 35.

78 Black (2012) (n 1) 8–10.

79 Ibid 6. The jurisprudential literature on rules is extensive. For legal analyses of rules see in particular Schauer (n 1); W L Twining and D Miers, *How to Do Things with Rules* 5th edn (Law in Context Series, Cambridge University Press 2010); H L A Hart and A Sacks, 'The legal process: basic problems in the making and application of law' (1996) 94(6) Michigan Law Review 1571–95; F Schauer, *Playing by the Rules: A philosophical examination of rule-based decision making in law and in life* (Oxford University Press 1991); D Kennedy, 'Form and substance in private law adjudication' (1976) 89 Harvard Law Review 1685–778; N McCormick, *Legal Reasoning and Legal Theory* (Clarendon Press 1978).

80 Black (2012) (n 1).

minimalist, and in the Foucauldian sense it can therefore remain disciplined from within. The shortcoming of such an approach is that it seeks to herd certain behaviours and/or attributes, to build up a category and/or definition which then forms a basic rule,⁸¹ which is subsequently used as a net which is thrown over a wide variety of circumstances which it was not designed to accommodate. The benefit of adopting a principles-based approach to regulation is that it can 'create a community of interpretation of rules through rules themselves', allowing for the necessary reciprocity between regulator and regulatee. For Black and Baldwin, whilst broad principles-based regulation may be criticised as 'vague', they 'assist with the problem of "honest perplexity" as to the rule's application, and also thwart creatively compliant behaviour'.⁸²

4 Challenges to devising responsive regulation

According to Black and Baldwin, to be really responsive, regulation must respond to attitudinal settings, to the broader regulatory environment, the different logics of regulatory tools and strategies, to the regime's own performance and, finally, to changes in each of these elements.⁸³ At this point, it is worth considering some of the key regulatory challenges social media raises, namely: (i) the dynamic quality of social media content generation; (ii) issues of jurisdiction and sovereignty; and (iii) whether the media through which social media operates defies regulation in its operational context.

4.1 THE DYNAMIC QUALITY OF SOCIAL MEDIA

The dynamic quality of internet content generation has been the subject of debate in terms of its regulation.⁸⁴ Social media concentrates this dynamism into a digital vortex with its speed, permanence and reach.⁸⁵ Whilst social media facilitated by Web 2.0 has the potential to offer vast opportunities, intrepid digital explorers have to discover these new frontiers before they can be shaped, described and regulated. It is little wonder therefore that, in the context of the development of information technology, Barlow declared:

... law adapts by continuous increments and at a pace second only to geology in its stateliness. Technology advances in lunging jerks, like the punctuation of biological evolution grotesquely accelerated. Real world conditions will continue to change at a blinding pace, and the law will get further behind, more profoundly confused.⁸⁶

Barlow's 'permanent mismatch' has even led some scholars to suggest that the internet cannot be regulated. Steinert-Threlkeld sums up the conundrum with remarkable brevity:

81 Black (ibid) at 7 notes that rules are linguistic in nature, and consequently 'how we understand, interpret, and apply rules depends in part on how we understand and interpret language'. Black gives the example of a naughty dog in a restaurant which leads to the formation of a rule which regulates dog's presence in restaurants. The owner in deciding their course of action must consider the dog's characteristics (dog, black, in a restaurant etc.). Black notes that the proprietor could 'consider banning all black things or all things called Rufus'. However, not all things which could be indexed within these categories are necessarily disruptive, these qualities were not the root cause of the disruption. Black therefore posits that the proprietor 'should focus on the fact that Rufus was a dog, and so form a rule, "no dogs allowed"'.
82 Baldwin and Black (2010) (n 1) 218.
83 Baldwin and Black (2008) (n 1) 69.
84 Lessig (n 2); A Murray and C Scott, 'Controlling the new media: hybrid responses to new forms of power' (2002) 65(4) *Modern Law Review* 491–516.
85 Eight of the world's most popular social networks flush out an astonishing amount of content every minute and the social media high seas, double in size every two years. By 2020 it is estimated will reach 44 trillion Gigabytes (See <http://www.emc.com/leadership/digital-universe/2014view/executive-summary.htm>).

86 John Perry Barlow 'The economy of selling ideas': Selling wine without bottles on a global net' available at <<http://www.eff.org/EconomyOfIdeas.html>> accessed 12 July 2016.

'some things never change about governing the Web. Most prominent is its innate ability to resist governance in almost any form.'⁸⁷ Such arguments could equally apply to social media, which arguably concentrates the problem due to its spontaneity, permanence and reach.⁸⁸

4.2 JURISDICTION AND SOVEREIGNTY

In addition to whether a space is regulatable, one must also consider who will regulate it. In the vast majority of circumstances, regulation is not the domain of one sovereign state, but extends to the world, prompting an important question as to how, common with other global issues, regulation can evolve within multi-level governance (state, international conventions etc.). In relation to the internet, libertarians such as Barlow urged governments as the 'weary giants of flesh and still' to accept that this 'home of the mind' was beyond the widely cast nets of such governments' sovereignty:⁸⁹ such arguments could similarly be levied at social media.

For Johnson and Post, 'the rise of an electronic medium that disregards geographical boundaries throws the law into disarray by creating entirely new phenomena that need to become the subject of clear legal rules but that cannot be governed, satisfactorily, by any current territorially based sovereign'.⁹⁰ There is, however, an inherent flaw in this logic, as it is predicated on the assumption that clear and determinative 'rules' are firstly desirable and secondly necessary to create conditions for effective regulatory governance. This is not a satisfactory starting point, given that social media is not unique in its objects and/or that events which form the subject matter of regulation are dynamic and fluid in nature. As noted by Scott and Murray: 'the problem of regulatory arbitrage emerges wherever subjects of regulation have sufficient mobility in their operations or activities that they can choose to be regulated by one regime or another, the effect is to create a form of market regulation within which dissatisfied subjects can "exit" one regime in favour of another'.⁹¹ Furthermore, Murray contends that 'at first glance it appears the regulatory challenge of cyberspace is to be found in the geography of the place, or rather the lack of geography'. He suggests that 'looking at individual case studies though obscures the wider view. The micro and the macro view are very different'.⁹² For example, in essence, until the rise of modern nations, maritime law did not derive its force from territorial sovereigns; instead it represented what was already conceived to be the customary law of the sea. As commerce moved northward and westward, sea codes developed in northern European ports, with important medieval sea codes such as the *Laws of Wisby* (a Baltic port), the *Laws of Hansa Towns* (a Germanic league) and the *Laws of Oleron* (a French island) being developed. These codes each drew inspiration from the *Consolato del Mare*, notably the *Laws of Oleron*, the second great code of maritime regulation. These codes are revered as the three arches upon which rest modern admiralty

87 Ibid fn 81, as quoted in Lessig (n 2) 57.

88 See for instance Murray and Scott's (n 84) observation that 'with much of the literature classical or 'command and control' regulation is held either to be undesirable or unfeasible in the face of new policy changes' 491.

89 Perry Barlow, J 'The economy of selling ideas: Selling wine without bottles on a global net' available at <http://www.eff.org/EconomyOfIdeas.html> accessed 12 July 2016

90 Johnson, D and Post, D 'Law and Borders- The Rise of Law in Cyberspace' Stanford Law Review 48 (1996): p. 1367,1375

91 Murray and Scott (n 84) 494.

92 Murray (n 45) 196. Elaborating on the challenges of such an approach Murray notes, at 197: 'cyber-governance theorists have therefore tended to focus on macro questions of geography and establishment. In this they follow a long established line of enquiry which places the design of the environment centre stage in the regulatory/governance discourse.'

structure (the ‘Three Arches’). According to one historian, the great value of the rules which had been developed for maritime trade lay in the fact that they had been ‘found by practice to be suitable to the needs of a community which knows no national boundaries—the international community of seafarers’.⁹³ The challenge therefore is to find a model suitable for such a dynamic environment whose attempts to impose a degree of artifice work organically with social media’s unique nature, rather than to impose a rigid construct. The goal should be to find a way to marry social media’s naturalistic, metonymic nature with the rules and architecture so that, whilst the rules may remain the ‘organising centre’, they complement the seascape. Rules should not be an edifice whose very infringement of the environment makes them vulnerable.

5 Responsiveness to operational and environmental factors

5.1 THE OPERATIONAL CONTEXT OF REGULATION

This contention that new technologies possess an innate ability to resist governance is an inviting thought, but not one which stands up to scrutiny. Changes of times and tides in technology are not new, nor are they unique to social media. According to Spar, innovative technology (e.g. the compass, the printing press) go through a four-phase cycle of innovation, commercialisation, ‘creative anarchy’ and, finally, rules.⁹⁴ The ‘rules’ phase sees the entry of the regulator into the marketplace and its absorption within the traditional regulatory framework, so that what is seemingly ungovernable is brought under control (much like the regulation of the high seas, the control of airspace etc.). There are, however, limitations, as will be discussed later in this article, in attempting to submit creative spaces to effective regulation.

In the context of social media regulation, it is useful to consider some of the arguments that have been raised in the context of regulating the internet more generally. The original internet made such regulation extremely difficult as originally deployed, as one court put it:

the Internet is wholly insensitive to geographic distinctions. In almost every case, users of the Internet neither know nor care about the physical location of the Internet resources they access. Internet protocols were designed to ignore rather than document geographic location; while computers on the network do have ‘addresses,’ they are logical addresses on the network rather than geographic addresses in real space. The majority of Internet addresses contain no geographic clues and, even where an Internet address provides such a clue, it may be misleading.⁹⁵

The internet was therefore insensitive because it was *designed* to be so. Therefore, the failing of the existing law and the challenge for any proposed regulatory model is that it must be able to adapt to the constraints of the regulatable environment and bring into its modelling the platform architects who design the technological state of the art which is the subject of such regulation, as well as those stakeholders subject to its regulation. The recurring theme is the need to focus on the nature of the regulatable space and how architecture can be designed which better meets its unique demands. The question therefore becomes not one of what makes a rule effective, but what makes for effective

⁹³ Despite the decline in the historical uniformity of early maritime laws these transactions have always been international in nature often implicating several countries, which can lead to an unpredictability for participants when domestic law becomes involved.

⁹⁴ Spar (n 42) 8.

⁹⁵ *American Library Association v Pataki* 969 F Supp 160 (SDNY) 1997, cited in Michael Geist, ‘Cyberlaw 2.0’ (2003) 44 Boston College Law Review 323, 326–7.

architecture and creates an environment in which rules can be conceptualised, nurtured and formed into courses of customary behaviour.

5.2 RESPONSIVENESS TO ENVIRONMENTAL FACTORS

By way of comparison to the sea, since at least the end of the nineteenth century, naval architecture and cargo-handling have changed in significant ways. For example the extensive use of carriers of liquefied natural gas has posed new hazards and questions of liability for oil pollution, calling for the development of new laws that must work alongside the ancient.⁹⁶ In the context of communications technologies, the tension-ridden space inhabited by those caught between innovation and governmental control has been considered by Spar: ‘communication is the sinew of both commerce and politics, the channel through which information – and this power – flows. Ever since God warned Eve to resist the apple, authorities have tried to control information flows.’⁹⁷ The germane point of analysis therefore is perhaps not so much that the internet (and latterly Web 2.0) resists regulation, but that it is ineffectual in practice as the regulation is not observed by those who are regulated.

As noted by Black, ‘increased precision may reduce but can never eliminate the inherent indeterminacy of rules and does not in the end create the understanding which it is trying to replace’.⁹⁸ Therefore, it is necessary to consider how a model could be adopted which can identify, frame and categorise the risks and issues presented by social media which has regard to the complex ‘mosaic of technical, psychological, cultural, social, political, organisational and economic concerns’.⁹⁹ The technical architecture to be adopted must make the threat of visibility inscribed in the subjects, but also consult with them in what essentially began as a community of persons: the users would adjust their behaviours based on community norms and the gaze of fellow users, as well as the regulator.

5.2.1 The development of the regulation of the sea

5.2.1.1 Historical development

Since the seventeenth century, the oceans have been subject to the freedom-of-the-seas doctrine – which essentially limits national rights and jurisdiction over the oceans to a

⁹⁶ Note also that not all of the original principles of maritime law still apply.

⁹⁷ Spar (n 42) 9.

⁹⁸ Black (2012) (n 1) 217.

⁹⁹ Ibid; Black and Baldwin (2012) (n 1) 4. See also Black (2012) (n 1) 216: ‘Rule making has been a continuous activity for the regulators since the regulation’s inception. In the process of rule formation and reformation, four dominant themes in the use of rules can be identified: first, the changing ways in which the tension between certainty and flexibility have been addressed; secondly, the different attempts made to use rule type to distribute discretion and exert control within the rule system; thirdly, the attempts made to develop interpretive communities and so to address the problems of uncertainty, honest perplexity, and creative compliance; and fourthly the use made of rules to reinforce the “self-regulatory” characterization of the system and to define the roles of regulators within that system.’ See also 220–1, the greater the shared understanding of the rule and practices it is addressing, the more the rule maker can rely on tacit understandings as to the aim of the rule and the context in which it operates, the less the need for precision, and the greater the degree to which simple, vague rules can be used. Finally, the idea that a self-regulatory system should be reciprocal, practitioner based, non-governmental, and non-bureaucratic has involved certain assumptions not only as to the particular nature of the regulator/regulated relationship (the regulatory contract), but as to the types of rules which should be adopted, and by which regulator. The emphasis on the self-regulatory characterisation of the system has thus been significant in the use made of rules to define regulatory roles. For further exploration of Black’s work, see Black (2012) (n 1).

narrow belt of sea surrounding a nation's coastline. The remaining vast expanse of the sea was proclaimed to be free to all and belong to none. By the mid-twentieth century there was an impetus to extend national claims over offshore resources due to growing concerns over the toll taken on coastal fish stocks by long-distance fishing fleets, pollution and wastes from transport ships and oil tankers carrying noxious cargoes. Such hazards threatened all forms of sea life. Sovereign nations were competing to maintain a presence across the globe on the surface waters and even under the sea, creating complex claims and growing tensions between coastal nations' rights to these resources. The pressures of long-distance navigation and a seemingly outdated, if not inherently conflicting, freedom-of-the-seas doctrine was threatening to transform the oceans into another arena for conflict and instability.

In 1945, the USA, unilaterally extended US jurisdiction over all natural resources on that nation's continental shelf – oil, gas, minerals – representing a major challenge to the freedom-of-the-seas doctrine. Other nations soon followed suit.¹⁰⁰ As technologies developed, the rich offerings of the oceans, such as transporting goods and extracting natural resources, led to the sea being exploited as never before, activities which were seemingly impossible just a few decades before, such as diamond and tin mining, were now in full swing. This was to say nothing of the fishing, which saw nations flooding the richest fishing waters with their fishing fleets virtually unrestrained as seaboard states attempted to set limits and fishing states contested them.

The tranquillity of the sea was slowly being disrupted by technological breakthroughs, accelerating and multiplying uses. Much like the use of social media, it was a time that held both dangers and promises, risks and hopes. The dangers were numerous: nuclear submarines charting deep waters never before explored; designs for antiballistic missile systems to be placed on the seabed; supertankers ferrying oil from the Middle East to European and other ports, passing through congested straits and leaving behind a trail of oil spills; and rising tensions between nations over conflicting claims to ocean space and resources. The oceans were generating a multitude of claims, counterclaims and sovereignty disputes. The hope was for a more stable order, promoting greater use and better management of ocean resources and generating harmony and goodwill among states that would no longer have to eye each other suspiciously over conflicting claims.

On 1 November 1967, Malta's Ambassador to the UN, Arvid Pardo, asked the nations of the world to open their eyes to a looming conflict that could devastate the oceans, the lifeline of man's very survival. In a speech to the UN General Assembly,¹⁰¹ he spoke of the super-power rivalry that was spreading to the oceans, of the pollution that was poisoning the seas, of the conflicting legal claims and their implications for a stable order and of the rich potential that lay on the seabed. Pardo ended with a call for 'an effective international regime over the seabed and the ocean floor beyond a clearly defined national jurisdiction . . . it is the only alternative by which we can hope to avoid the escalating tension that will be inevitable if the present situation is allowed to continue'.

100 In October 1946, Argentina claimed its shelf and the epicontinental sea above it. Chile and Peru in 1947, and Ecuador in 1950, asserted sovereign rights over a 200-mile zone, hoping thereby to limit the access of distant-water fishing fleets and to control the depletion of fish stocks in their adjacent seas. Soon after the Second World War, Egypt, Ethiopia, Saudi Arabia, Libya, Venezuela and some Eastern European countries laid claim to a 12-mile territorial sea, thus clearly departing from the traditional three-mile limit. Later, the archipelagic nation of Indonesia asserted the right to dominion over the water that separated its 13,000 islands. The Philippines did likewise. In 1970, Canada asserted the right to regulate navigation in an area extending for 100 miles from its shores in order to protect Arctic water against pollution.

101 <www.un.org/Depts/los/convention_agreements/texts/pardo_ga1967.pdf>

Pardo's urging came at a time when many recognised the need for updating the freedom-of-the-seas doctrine to take into account the technological changes that had altered humankind's relationship to the oceans. It set in motion a process that spanned 15 years and saw the creation of the UN Seabed Committee, the signing of a treaty banning nuclear weapons on the seabed, the adoption of the declaration by the General Assembly that all resources of the seabed beyond the limits of national jurisdiction are the common heritage of humankind, and the convening of the Stockholm Conference on the Human Environment. What started as an exercise to regulate the seabed turned into a global diplomatic effort to regulate and write rules for all ocean areas, all uses of the seas and all of its resources. These were some of the factors that led to the convening of the Third UN Conference on the Law of the Sea, to write a comprehensive treaty for the oceans. The Conference was convened in New York in 1973. It ended nine years later with the adoption in 1982 of a constitution for the seas – the UN Convention on the Law of the Sea.

5.2.1.2 *The Convention*

Navigational rights, territorial sea limits, economic jurisdiction, legal status of resources on the seabed beyond the limits of national jurisdiction, passage of ships through narrow straits, conservation and management of living marine resources, protection of the marine environment, a marine research regime and, a unique feature, a binding procedure for settlement of disputes between states – these are among the important features of the treaty. In short, the Convention is an unprecedented attempt by the international community to regulate all aspects of the resources of the sea and uses of the ocean, and thus bring a stable order to humankind's very source of life. The Convention came into force on 16 November 1994, one year after Guyana became the 60th state to adhere to it.

Across the globe, governments have taken steps to bring their extended areas of adjacent ocean within their jurisdiction. They are taking steps to exercise their rights over neighbouring seas, to assess the resources of their waters and on the floor of the continental shelf. The practice of states has in nearly all respects been carried out in a manner consistent with the Convention, particularly after its entry into force and its rapid acceptance by the international community as the basis for all actions dealing with the oceans and the law of the sea.

The Convention retains for naval and merchant ships the right of 'innocent passage' through the territorial seas of a coastal state. This means, for example, that a Japanese ship, picking up oil from Gulf states, would not have to make a 3000-mile detour in order to avoid the territorial sea of Indonesia, provided passage is not detrimental to Indonesia and does not threaten its security or violate its laws.

In addition to their right to enforce any law within their territorial seas, coastal states are also empowered to implement certain rights in an area beyond the territorial sea, extending for 24 nautical miles from their shores, for the purpose of preventing certain violations and enforcing police powers. This area, known as the 'contiguous zone', may be used by a coastguard or its naval equivalent to pursue and, if necessary, arrest and detain suspected drug smugglers, illegal immigrants and customs or tax evaders violating the laws of the coastal state within its territory or the territorial sea.

The Convention also contains a new feature in international law, which is the regime for archipelagic states (states such as the Philippines and Indonesia, which are made up of a group of closely spaced islands). For those states, the territorial sea is a 12-mile zone extending from a line drawn joining the outermost points of the outermost islands of the group that are in close proximity to each other. The waters between the islands are declared archipelagic waters, where ships of all states enjoy the right of innocent passage.

In those waters, states may establish sea lanes and air routes where all ships and aircraft enjoy the right of expeditious and unobstructed passage.

5.2 CONTEMPORARY EXAMPLES OF CHANGES IN TIDES AS TO THE APPROACH TO REGULATION

Murray and Scott observe, with regards to their concept of ‘regulatory arbitrage’, which could be described as a form of jurisdiction forum-shopping by the regulated, that such mischief is addressed for example in Europe by the European Commission issuing harmonised rules requiring all member states to legislate for common standards and services.¹⁰² In Europe, there is an increasing recognition in legislation currently being drafted, which affects the private sector, that regulation can be drafted based on high-level principles which devolve its implementation to those regulated providers who have the capability to deliver technical solutions on a pan-European basis, with extra-territoriality provisions. One such example is the European General Data Protection Regulation (GDPR)¹⁰³ which seeks to regulate information society services¹⁰⁴ at a conceptual level,¹⁰⁵ placing the onus on developers to design appropriate privacy tools by reference to the state of the art in technology which can deliver legally compliant solutions.¹⁰⁶ This sea change in regulatory approach has been conceptualised by the UK Information Commissioner Elizabeth Denman, in the context of privacy regulation, thus: ‘Data protection is a team sport. Effective regulation requires engagement with the public sector, with industry, with civil society and with the public at large.’¹⁰⁷ Indeed, in the context of the UK’s current plan to exit the EU (commonly referred to as ‘Brexit’), the UK after leaving the EU will consider how it can maintain equivalence with European

102 Murray and Scott (n 84) 494.

103 GDPR 2016/679/EU.

104 This includes social media sites, see Scaife (n 12) 16–18

105 Article 25 of the GDPR states that: ‘data controllers must take into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of the GDPR and protect the rights of data subjects’.

106 Recital (78) of the GDPR states: ‘the protection of the rights and freedoms of natural persons with regard to the processing of personal data require that appropriate technical and organisational measures be taken to ensure that the requirements of this Regulation are met. In order to be able to demonstrate compliance with this Regulation, the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default. Such measures could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features. When developing, designing, selecting and using applications, services and products that are based on the processing of personal data or process personal data to fulfil their task, producers of the products, services and applications should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and, with due regard to the state of the art, to make sure that controllers and processors are able to fulfil their data protection obligations. The principles of data protection by design and by default should also be taken into consideration in the context of public tenders.’

107 Information Commissioner’s Office Newsletter August 2016
<<http://ico.msgsfocus.com/q/1AFE2kFCwd/wv>>.

privacy rules¹⁰⁸ in order to offer mutually equivalent laws for the safe trans-border processing of personal data.¹⁰⁹

As well as hard law, industry can also provide end-users with tools to manage their online presence. For example, in December 2014, Facebook sent an update to users promoting its new 'Privacy Basics' service, noting that 'protecting people's information and providing meaningful privacy controls are at the core of everything we do'.¹¹⁰ Yet recently, discussion surrounding Cambridge Analytica and the scope of political engineering through the use of social media data have sparked a myriad of debates as to the ethical parameters of data-sharing and use. Although compliance comes with a cost, which in this model is borne by industry, privacy-enhancing technology is now actively sold as a feature of modern products. By way of example, iPhones offer encryption standards which mean that not even the provider of the phone will be able to decrypt its contents,¹¹¹ and in 2016 WhatsApp and Facebook announced that messages will now be sent with end-to-end encryption.^{112&113} Sunstein has advanced arguments in favour of filtering and feedback technologies as a means of 'nudging compliance',¹¹⁴ founded on the principle that, while we behave irrationally, our irrationality can be corrected – if only the environment acts upon us, nudging us towards the right option.

6 Responsiveness to attitudinal settings and environments

6.1 REGULATORY DEFIANCE

Despite the presence of rules, they are regularly ignored or actively disobeyed.¹¹⁵ We have rules for the use of our highways, yet individuals still run red lights and break the speed limit. Our land and rivers continue to be polluted and emissions levels have been falsified, at cost to the environment. How regulation can be effectively applied in its operational context is not a new conundrum and has been extensively considered in the wider context

108 On 25 May 2018, the GDPR became applicable in all 28 EU member states, repealing the Data Protection Directive (95/46/EC), and thereby changing the substantive scope of regulation of personal data after more than 23 years. In the UK, the GDPR was automatically incorporated into domestic law via the European Communities Act 1972. From 29 March 2019, when the UK is expected to leave the EU, the European Union (Withdrawal) Act 2018 will repeal the European Communities Act 1972 and simultaneously transpose the GDPR onto the statute book, making it domestic legislation in the UK.

109 Areas of data protection regulation, such as international data transfers, applicable supervisory authorities and enforcement of the GDPR that will be affected when the UK's status changes to that of a 'third country' for the purposes of EU law. In her Mansion House speech on 2 March 2018, the UK Prime Minister (Theresa May) stated that a deal on data protection is one of the foundations that must underpin the UK–EU trading relationship and that the UK would therefore seek more than an adequacy arrangement with the EU, in the form of a Treaty. The UK's aim was to achieve a relationship that would provide the Information Commissioner's Office with an 'appropriate ongoing role', including in relation to the operation of the 'one-stop-shop' mechanism for resolving data protection disputes under the GDPR. The government released a presentation on 23 May 2018, which set out this proposal in greater detail see Framework for UK–EU Partnership concerning Data Protection https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/710147/DATA_-_FINAL.pdf.

110 Facebook update, 20 December 2014.

111 See the 'Privacy' section on the Apple website <www.apple.com/uk/privacy/government-information-requests>.

112 'End-to-end encryption' (*WhatsApp Blog* 5 April 2016) <<https://blog.whatsapp.com/10000618/end-to-end-encryption>>.

113 The advantage of encryption is also potentially to buffer the user against the all-pervading surveillance of the Foucauldian model, as discussed below.

114 Sunstein (n 41).

115 Baldwin (n 1) 321.

of how to devise effective regulation.¹¹⁶ These observations map well to the problems facing the effective regulation of social media. Baldwin and Black,¹¹⁷ in devising their approach for the creation of conditions to give effect to really responsive regulation, have considered this issue in the context of enforcing fisheries laws, noting that 'regulates are highly mobile', 'inspection at sea is resource intensive', 'there are many ways to avoid detection' and there are logistical challenges associated with monitoring compliance effectively as it involves a large number of organisations, where 'responsibilities overlap and often the enforcers have no clear set of priorities and outcome objectives to work from'. As a result, the control system relies 'heavily on self-reporting of catches' with extensive under-reporting and mis-declarations of fish landed,¹¹⁸ meaning it is difficult to assess 'off the radar' compliance'.

6.2 RESPONSIVENESS TO INSTITUTIONAL ENVIRONMENTS

Lessig, in his famous treatment of regulation, posited that regulation is not the sole product of law, but also of market and social norms, and is consequently more concerned with high-level choice, values and democracy. The law does not therefore operate directly, but also indirectly through other modalities, such as the technology itself in order to 'regulate to law's own end', leading to direct and indirect regulatory effects.¹¹⁹ Therefore, law can in principle regulate the architecture, and the architecture can regulate the norms, thereby avoiding a situation whereby detailed rules are created which foster a sense of 'distrust' between the rule-maker and its subject.¹²⁰ Both Marsden and Price note that the developer community has traditionally placed great store in self-regulation based upon codes of practice, contractual terms and community standards,¹²¹ with Marsden contending that 'governments have broadly accepted that a more flexible and innovation-friendly model of regulation is required, particularly in view of the rapid growth, complex

116 See n 2.

117 Baldwin and Black (2008) (n 1) 59–60.

118 In response to a Freedom of Information request submitted to the Ministry of Justice (MoJ) on 18 August 2016 by the author requesting figures relating to social media crimes, the MoJ responded that: 'the MoJ Court Proceedings Database holds information on defendants proceeded against, found guilty and sentenced for criminal offences in England and Wales. This database holds information on offences provided by the statutes under which proceedings are brought but not all the specific circumstances of each case. From centrally held data, we are unable to link defendants proceeded against at and found guilty in court back to the original crime reported to the police. Please bear in mind that the offence under which a criminal event is originally reported to the police may differ from the offence for which the defendant is finally convicted. In addition, we are unable to separately identify offences committed through the use of social media from all offences committed via the use of electronic communications. This detailed information is not reported to Justice Statistics Analytical Services due to its size and complexity.' In a follow-up response dated 16 September 2016 the MoJ stated: 'no reports include statistics on offences that involve the use of social media, for the reasons given above. The only information that we have for 2015 is published in Criminal Justice Statistics Quarterly (<https://www.gov.uk/government/statistics/criminal-justice-system-statistics-quarterly-december-2015>) In the outcomes by offence table, there is some data for small groups of offences (for example, 196- dealing with sending grossly offensive messages under the 2003 Communications Act). In addition, there is some information on these crimes in the experimental statistics published on the same web link. However, this information is based around the statute under which the offence is prosecuted and convicted rather than the circumstances in which it was committed.'

119 Lessig (n 2) chs 4 and 5.

120 Black (2012) (n 1) 217.

121 Marsden (n 3) 211–28.

inter-relationships and dynamic changes taking place in [the] Internet'.¹²² Sassen notes that to create a space where users can engage in activities in a manner that is autonomous from states does not exist in non-digital media,¹²³ with Murray and Scott noting in the digital context that 'digitisation of broadcasting and mobile telecommunications create niches for new forms of service provider, shifting power away from both those who own the physical infrastructure of networks and from those who own content'.¹²⁴

Consequently, given Black's observations regarding the context in which law operates and the reciprocal relationship between technology and rules, they must be considered in parallel when proposing a regulatory model suitable for the complex demands presented by social media,¹²⁵ so that a form of regulation can be adopted which can adapt to meet new challenges where non-compliance is concealed and/or new methods of evading

122 Ibid 212. The Director of Public Prosecutions has stated that everyone from tech companies like Twitter and Facebook, to police and the CPS and wider society, needed to do more to tackle the growing scale of threatening and abusive communications online. 'The flip side of that is that the evidence is out there and it is pretty incontrovertible evidence. We have been working with Twitter and Facebook to help us train prosecutors in how to be aware of what evidence is available and talking to them about how they can help us and victims.' As quoted in S Laville, 'Violent crimes against women in England and Wales reach record high' *The Guardian* (London, 6 September 2016) <www.theguardian.com/society/2016/sep/05/violent-crimes-against-women-in-england-and-wales-reach-record-high>.

123 S Sassen, 'Towards a sociology of information technology' (2002) 50(3) *Current Sociology* 365–88.

124 Murray and Scott (n 84) 493.

125 An electronic communications service (ECS) is defined in the UK Communications Act 2003 as 'a service consisting of, or having as its principal feature, the conveyance by means of an electronic communications network of signals, except in so far as it is a content service'. The Information Commissioner's Office, which oversees data protection rights in the UK, has stated that an ECS is any such service that is provided so as to be available for use by members of the public. In terms of the application of ECS to social networking, having regard to the above, it has been suggested by some commentators that SNSs constitute information society services, but not ECSs. This is supported by the fact that retention obligations in the Data Retention (EC Directive) Regulations 2009 do not currently extend to any SNS communications. (Directive 2006/24 on data retention [2006] OJ L105/54). See further, responses of Mr Vernon Coaker to questions posed during the discussion of the Draft Data Retention (EC Directive) Regulations 2009 (Fourth Delegated Legislation Committee, 16 March 2009), and Data Retention Expert Group (Commission Decision 2008/324/EC) on webmail and web-based messaging: DATRET/EXPGRP (2009) 2 -- FINAL - ANNEX- 03 12 2009. For critique of the Directive, see I Brown, 'Communications data retention in an evolving internet' (2011) 19(2) *International Journal of Law and Information Technology* 95–109. Many countries have expressed concern over the Directive and the European Court of Justice is soon due to rule on its legality following a referral from Ireland. See further, opinion of the European Data Protection Supervisor <https://edps.europa.eu/sites/edp/files/publication/11-05-30_evaluation_report_drd_en.pdf>. In *DPP v Chambers*, a ground-breaking element of the Court of Appeal's judgment in this case came from its agreement with the Crown Court's analysis that the internet itself constitutes a public network. The Crown Court has stated ([2006] UKHL 40) at [23]: 'the "Twitter" website although privately owned cannot, as we understand it, operate save through the internet, which is plainly a public electronic network provided for the public and paid for by the public through the various service providers we are all familiar with . . . The internet is widely available to the public and funded by the public and without it facilities such as "Twitter" would not exist. The fact that it is a private company in our view is irrelevant; the mechanism by which it was sent was a public electronic network and within the statutory definition . . . "Twitter", as we all know is widely used by individuals and organisations to disseminate and receive information. In our judgment, it is inconceivable that grossly offensive, indecent, obscene or menacing messages sent in this way would not be potentially unlawful.' In defining the internet thus, the court stated at [24] that the internet 'is plainly a public electronic communications network provided for the public and paid for by the public through the various service providers we are all familiar with' and that 'potential recipients of the message were the public as a whole, consisting of all sections of society'. Judge Jacqueline Davies' reasoning, with which the Court of Appeal agreed, derived from an analysis of the internet's network infrastructure as a series of links. These links cover networks of networks and services linking individuals, service providers, network providers, platform providers and content providers.

detection are devised so that the 'gap between rules and objectives' does not become too wide.^{125a}

6.3 THE PLATFORM PROVIDERS

Barlow's comments with regards to the law's ability to keep pace with change in the internet context support the contention that the destinies of regulation and technology are intertwined.¹²⁶ As rules cannot apply themselves, 'for the rule to be applied in a way which will further the overall aims of the rule maker, then the person applying it has to share the rule maker's interpretation of it'.¹²⁷ Quite simply, the tools that can shape it inform the form of the regulation. As David Brin posits, 'if we admire the Net, should not a burden of proof fall on those who would change the basic assumptions that brought it about in the first place?'¹²⁸

The value of consulting the private sector when determining how to regulate social media is highlighted by reference to the US responses to terrorist content posted online. Even if such materials can be moderated, there is the risk that materials will slip through the (inter)net (given the sheer volume of material that would need to be screened), or that censoring materials may be cast too wide. Social media sites commonly use algorithms to 'flag' accounts that might need to be suspended as there are not the resources or time to conduct a thorough review of all content posted online or every item that attracts a small number of complaints.¹²⁹ In the summer of 2015, a Bill was still put before Congress as part of the Intelligence Authorisation Bill¹³⁰ with provisions similar to the laws that require companies to report child pornography.¹³¹ Such identifiers cannot, however, be assigned en masse to terrorist content. For instance, a news article or video clip used by CNN could end up having the same identifiers as a posting made by an ISIS member (e.g. the presence of an ISIS flag), meaning it is difficult for the algorithm to figure out what is propaganda and what is not.¹³² Whilst it may be argued that regulation is possible, given that copyright breaches committed via social media-enabled platforms have become the subject matter of regulation, one such instance of its successful application of content removal via Twitter and Facebook of Olympics materials, which were subject matter of copyright, represents a compelling example of how market interests have shaped an active response from the platform providers. Matters relating to expressive content and privacy rights are more nuanced, having to absorb constitutional and cultural factors, which such hard-edged law is simply not well placed to accommodate, or indeed adjudicate.

^{125a} Black and Baldwin (2008) (n 1) 81.

¹²⁶ Barlow (n 86).

¹²⁷ Black (2012) (n 1) 218.

¹²⁸ D Brin, *The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom?* (Perseus 1999) 324.

¹²⁹ These algorithms can, however, be exploited through 'flag spamming campaigns' as was evident though their use to push critics of the Russian and Vietnamese governments off Facebook. For example, shortly after their video 'If i wanted America to fail' went viral, Free Market America found themselves kicked off Twitter. See J Hayward, 'If I wanted America to fail' (*Human Events*, 23 April 2012) <www.humanevents.com/2012/04/23/if-i-wanted-america-to-fail>. Companies like Google have sought to grant enhanced flagging privileges to non-governmental 'trusted users' who have the ability to report offensive material and obtain quick action from service providers. Whether this is more of an affront to freedom of expression than member state interference is debatable, but it does demonstrate that with an in providing a degree of latitude there is room for creative solutions to be explored.

¹³⁰ That still has to be approved by the Senate.

¹³¹ Determining what constitutes child pornography is, technically speaking, a simple task as a criminal photograph can be digitally analysed and assigned a unique identifier, which can then be used to detect similar images.

¹³² See nn 37 and 38.

This is not to suggest that the process of comprehensive regulation can merely be achieved though wider and earlier stakeholder engagement; it is instead a far more iterative process.¹³³ By way of comparison, ships started out as basic vessels to carry the goods/thoughts of their passengers in a primitive way.¹³⁴ The passage of time allowed ships to become ever more sophisticated moving from seafarers making use of the sun and the stars (celestial navigation), to using compasses, sextants and telescopes to navigate. Navigation has since evolved with the use of radar, electronic charts, long-range tracking and identification systems and global positioning systems (GPS) technology and so on. Comparing this with social media, platforms and apps were originally quite basic. However, as the algorithms used to generate site content and manage user preferences (notably in relation to privacy) have developed, they are therefore increasingly equipped to deal with complex risks and challenges¹³⁵ and in a position to comment on the practicality of delivering the regulations to which it is proposed they are subject.

There is a note of caution: such systems are often designed by giving notice and choice to customers, which for Lessig can lead to a situation by which ‘code becomes a means by which to transfer decisions from the public realm to the privatised realm . . . [i]t

133 As noted by Murray (n 45) 213–19: ‘As regulators are inclined to rely ever more on gatekeepers as proxies in their attempts to control online activities, it means that our old models for cyber-governance and regulation have become outdated.’ Murray suggests instead a model that applies regulatory modelling to communicative relationships of power to provide a richer model that articulates the regulation spheres of activities.

134 Scaife (n 12) 4–7.

135 In June 2011, the Special Rapporteur, together with the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights of the Organisation of American States, the Representative on Freedom of the Media of the Organisation for Security and Cooperation in Europe, and the Special Rapporteur on Freedom of Expression of the African Commission on Human and Peoples’ Rights, issued a joint declaration establishing guidelines to protect freedom of expression on the internet, reinforcing the need to protect freedom of expression even in the online context <www.osce.org/fom/99558?download=true>. On the issue of censorship, the mandatory blocking of websites is an extreme actions that may only be justified in accordance with international standards. The rapporteurs stated that content-filtering systems that cannot be controlled users, imposed by governments or commercial providers, are also actions that are incompatible with freedom of expression. It is noted that in relation to the criminal law this has caused significant difficulties in the UK regarding when filtering and content removal should occur. The report considered this issue in some depth, noting that, while the state (which is party to human rights treaties) has an obligation to establish criminal law and systems sufficient to deter and respond to attacks on individuals (See, for example, ECtHR App No 23452/94. 28 October 1998, in which the court stated that the right to life (Art 2(1) ECHR) included the obligation to put in place ‘effective criminal law provisions to deter the commission of offences against the person backed up by law enforcement machinery for the prevention, suppression and sanctioning of breaches of such provisions’), it must not go so far as to deny individual rights by its criminalisation of particular conduct (United Nations Commission on Narcotic Drugs, and Commission on Crime Prevention and Criminal Justice, 2010: Drug control, crime prevention and criminal justice: A Human Rights perspective. Note by the Executive Director. E/CN.7/2010/CRP.6 – E/CN.15/2010/CRP.1. 3 March 2010). In order to undertake an impact assessment, states must therefore assess criminal provisions on a ‘right-by-right’ basis. By approaching an analysis of the provisions in this way, it is possible to test whether its contents infringe a range of individual rights – such as the right not to be subjected to arbitrary or unlawful interference with privacy, family, home or correspondence (Art 17 ICCPR), the right to freedom of thought, conscience and religion (Art 18 ICCPR), or the right of peaceful assembly (Art 21 ICCPR).

is a way to convert political rights into market commodities',¹³⁶ for example, non-profit organisations such as Code for America which then co-opt the state – under the guise of encouraging talented hackers to tackle civic problems.

Furthermore, there is also a distinction to be drawn between the *trust* placed in private companies as opposed to the government, with research indicating that the government is more trusted.¹³⁷ More recently, concern has been expressed regarding the government's use of data,¹³⁸ particularly in terms of recent profiling and data leaks,¹³⁹ compounded by

136 Lessig (n 2) ch 11 and 159–63. The old adage that nothing in life is free very much applies to 'free online services'. There is a value trade-off garnered from the information you provide to companies offering online services. This is because data relating to your buying habits and social/lifestyle preferences is a valuable trade for big organisations being employed to different degrees to do everything from marketing to helping to determine credit scores and insurance price (see J Leber, 'How wireless carriers are monetizing your movements' (*MIT Technology Review Website*, 12 April 2013) <www.technologyreview.com/s/513016/how-wireless-carriers-are-monetizing-your-movements>) creating 'a detailed composite of the consumer's life' (*Data Brokers: A Call for Transparency and Accountability* (Federal Trade Commission May 2014) <www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>). Such site providers disclose this information in their terms of service, for instance, with the aim of generating advertising revenue, for example Google's online terms of service state 'Our automated systems analyze your content (including emails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection. This analysis occurs as the content is sent, received, and when <[https://policies.google.com/privacy/google-partners?hl=en](http://policies.google.com/privacy/google-partners?hl=en)>. Sources can include a user's IP address, Google and YouTube profiles. Social media sites are engaged in the trading of data, according to Facebook's Data Policy <www.facebook.com/help/111814505650678> it 'shares' information about users 'within the family of companies that are part of Facebook' to 'facilitate, support and integrate their activities': see <www.facebook.com/about/privacy/update>. There are currently 10 companies listed in the family, including WhatsApp, Instagram and Atlas. Facebook's Audience Network programme provides app developers with aggregated data to target their ads. 'Facebook Services' are also covered by this Data Policy and include services such as 'Audience Insights'. This service is designed to provide businesses with information about the 'geography, demographics and purchasing behaviour and more' (See <www.facebook.com/business/news/audience-insights>). In March 2016, social networking forum Reddit removed a section from its website used to tacitly inform users that it had never received a certain type of US government surveillance request <www.reuters.com/article/us-usa-cyber-reddit-idUSKCN0WX2YF>. Reddit also deleted a paragraph usually included in its transparency report known as a 'warrant canary' to signal to users that it had not been subject to so-called national security letters, which are used by the FBI to conduct electronic surveillance without the need for court approval. This 'Big Brother' society has caused privacy watchdogs to more closely scrutinise the practices of big businesses. Indeed, in April 2015, the Information Commissioner's Office, which oversees data protection issues in the UK, launched an investigation into UK firms sharing pension, medical and financial data (see the announcement on the Information Commissioner's Office website <<https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2015/04/ico-to-make-enquiries-about-sale-of-pension-data>>). Access to services is often a condition of signing up, which means that the consent in practice may not be freely given in a world where such services are the lifeblood of modern life, as expressed by one commentator: 'It's not reasonable to tell people that if they don't like the data collection, they shouldn't email, shop online, use Facebook or have a cell phone. I can't imagine students getting through school anymore without Internet search or Wikipedia, much less finding a job afterwards. These are the tools of modern life.' B Schneier, *Data and Goliath*, 1st edn (W W Norton 2015) ch 4.

137 See 2.27(a) above, last bullet point, and Ipsos MORI: ESRC/ONS; Deloitte; Eurobarometer. Within the US government at least, there may also be some differentiation; see Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values* (May 2014), in which law enforcement and intelligence agencies were ranked low in terms of public trust.

138 Polling was conducted by Ipsos MORI for the *Evening Standard* in October 2014. See 'Public backs curbs on police seeing phone records of journalists' *Evening Standard* (London, 21 October 2014) as excerpted in D Anderson, *A Question of Trust* (Report of the Investigatory Powers Review 2015) 33–4; Ipsos MORI: RSS; 13% had high trust in the British government compared to 46% with low trust.

139 Anderson (n 138).

the ‘Snowden effect’.¹⁴⁰ A number of studies have suggested that most people had already assumed that the type of action alleged in the Snowden documents was undertaken;¹⁴¹ however, this has not overturned some research indicating low levels of trust in the UK government to use people’s data appropriately,¹⁴² with many holding the opinion that neither government nor private companies can now keep their data completely secure.¹⁴³

Baldwin notes that, as such interests are accommodated, the broad principles can become less accessible as they seek to accommodate private sector interests and the contemporary politics of policy-making¹⁴⁴ which leads to a jarring as between participation and effectiveness.¹⁴⁵ However, the private sector in recent years has increasingly become a backstop for state interference with privacy rights. Microsoft has successfully resisted the US government’s demand for emails of a non-US citizen that the company has stored in a data centre located in Ireland.¹⁴⁶ Similarly, Apple has refused requests by the Federal Bureau of Investigation (FBI) to reveal information and/or change code on its phones which would allow the authorities to unlock phones which are cryptographically protected.¹⁴⁷ Of course, not all platform providers are so scrupulous. For example, platforms such as Telegram and Surespot, which are popularly used by terrorist groups like ISIS, have been accused of harbouring terrorists, at the expense of risking state security and the protection of citizens in favour of commercial gain.¹⁴⁸

A situation must not, however, be created by which the price of co-operation is so high that it erodes the original purpose for which the regulation was adopted. Nevertheless, this is not to say that co-regulation is not possible, merely that the interests of third-party providers and the remaining need to maintain some sort of coercion must be considered. In the context of fisheries,^{148a} note that regulation ‘may lead fishermen to act in a way in which they regard as unnatural’. A potential solution to this for such ‘maverick offenders’ is to encourage general buy-in from the industry or make access to

140 Ibid.

141 See TNS-BMRB Polling 23–27 January 2014, as excepted in Anderson (n 138).

142 Anderson (n 138). 13% of those polled had high trust in the British Government compared to 46% with low trust: *ibid.*

143 Ibid.

144 Baldwin (n 1) 334.

145 Foucault (n 15) 278

146 ‘Microsoft suit is latest tech clash with US over privacy’ (*Tempo.co*, 1 April 2016) <<http://en.tempco.co/read/news/2016/04/16/310763182/Microsoft-Suit-is-Latest-Tech-Clash-with-US-over-Privacy>>.

147 Gary Fagan, ‘In the matter of the search of an Apple iPhone seized during the execution of a search warrant on a black Lexus IS300, California License plate 35KGD203’ (*Ars Technica*, 4 March 2016) <www.clearinghouse.net/detail.php?id=15497ext>.

148 Surespot <www.surespot.me> and Telegram <<https://telegram.org>> are popular with jihadis, due to their ability to facilitate the encryption of messages so that only the intended recipient can read it. Tantalisingly, Surespot declares on its website: ‘we don’t know or share anything about you . . . Surespot is about taking back your right to privacy and it is made free to provide unrestricted access for everyone.’ The website explains Surespot’s encryption by using the example of a postcard that anyone who touches can read: ‘Typically you do not send information like a credit card number or your pin number or an intimate thought using the postcard format. Today this is what sending an email or a text message or an instant message or a picture is like. The message is the postcard which travels along many hops until it reaches its destination. At every one of these “hops” the message could potentially be read’.

148a Black and Baldwin (2008) (n 1) 82.

the market conditional on submitting to being regulated.¹⁴⁹ Inevitably, there will always be detection challenges, which is why coercive compliance measures such as the issuing of fines and undertakings must still be factored into a model to catch those who will only respond to punitive measures. It is therefore suggested that any proposed regulatory model must be backed up with stringent financial and operational sanctions such as fine and enforcement regimes, as well as rights for the state to audit information society service providers.¹⁵⁰

6.4 RESPONSIVENESS TO THE BEHAVIOUR, ATTITUDE AND CULTURE OF THE COMMUNITIES WHICH REGULATION MUST SERVE

The media has been described as an 'extension of man',¹⁵¹ acting as a vessel through which a discovery process of giving and receiving information makes individuals more visible to other users. As noted by Bucher, 'the regime of visibility associated with Web 2.0 connects to the notion of empowerment, as it has greatly expanded the social field of becoming recognised as a subject with a voice'.¹⁵² For Foucault, the 'the turning of real lives into writing . . . functions as a procedure of objectification and subjugation'.¹⁵³ This raises the question of to what extent platforms can design and build sufficient architecture to successfully 'police' their platforms when faced with an endless tide of content creation. Modern-day King Canute YouTube has stated that, for every single minute of the 24 hours in a day, it receives an average of 35 hours of video from millions of contributors.¹⁵⁴ Essentially, this boils down to a quite simple concept: sites cannot pre-screen material prior to uploading and active policing is neither possible nor practical.

Prisons and other such institutions are essentially places of 'constructed visibility'.¹⁵⁵ The sixteenth century onward saw the emergence of disciplinary societies in which a small few controlled the many. Jeremy Bentham's panopticon¹⁵⁶ was a powerful representation of the regulatory force of power, strengthened though certain architectural compositions,¹⁵⁷ and even the Disney World is based on a model in which visitors are 'controlled by an architecture in which nearly every aspect of the design has a disciplinary function'.¹⁵⁸ Although long pre-dating the digital content, Foucault operated from the proposition that it is not just what is seen (or posted) that is important

149 Black (2012) (n 1) 217–18: 'The propensity for detailed rules to fail in this manner has been recognized, at least to an extent. In particular, the moves towards the creation of interpretive communities can be seen as attempts to find ways to escape the need for precise and complex rules. Three different aspects of this development can be identified. The first is the use of rule type, the second a change, in the level of rhetoric at least, of regulatory approach, and the third an increased emphasis on the regulatee's own attitude, education, training, and competence. Each is aimed at changing the internal attitude of (p.218) the regulated to the regulation; displacing the need for control by building up understandings within the regulatory system as to what the regulatee is meant to do, how it is meant to act'.

150 Scaife (n 12) 16.

151 M McLuhan, *Understanding Media: The Extensions of Man* (MIT Press 1964).

152 Raz (n 4); T Bucher, 'Want to be on the top? Algorithmic power and the threat of invisibility on Facebook' (2012) 14(7) *New Media and Society*, 1164–80, 1165.

153 Foucault (n 15) 290.

154 Statistics, YouTube Press Office. Accessible via <<https://www.youtube.com/yt/press/statistics.html>>. Accessed January 2016.

155 Foucault (n 15) 171.

156 J Bentham, *Proposal for a New and Less Expensive mode of Employing and Reforming Convicts* (London 1798).

157 Bucher (n 152) 1170.

158 Murray and Scott (n 84) 500. Clifford D Shearing and Philip Stenning, 'From the Panopticon to Disney World: the development of discipline' in A N Doob and E L Greenspan (eds), *Perspectives in Criminal Law: Essays in Honour of John LJ Edwards* (Canada Law Book 1985) 335–49 195–231.

in terms of understanding its visibility, but how it is constructed (by the regulated community). Therefore, social media platforms could be described as paradigms of constructed virtual visibility. As noted by Bucher, with regards to the regulation of social media by reference to the Foucauldian notion of visibility 'the medium specificity of the architectural organization of visibility i.e. what can be seen and heard, to a large extent became a question of software',¹⁵⁹ highlighting the symbiosis between law and technology, which leads to the need for a cyclical approach to the formation of effective principles-based regulation.¹⁶⁰

Lessig imagines a world where there would be a dynamic dialogue between regulator and regulated.¹⁶¹ Here the panoptic model may offer some assistance: 'The major effect of the panopticon: is to induce in the inmate a state of conscious and permanent visibility that assures the automatic functioning of power.'¹⁶² The uncertainty associated with the spectral gaze of the regulator means that users adjust their behaviour to act as if they are indeed being permanently observed. Discipline in this context makes the principle the subject of its own subjugation,¹⁶³ in which the subjects are responsible for their own behaviour's discipline, therefore making individuals, as it is 'the specific technique of power that regards individuals as both objects and instruments of its exercise',¹⁶⁴ imposing particular conduct on a particular human multiplicity.¹⁶⁵ For Bucher, Foucault developed this notion of disciplinary power in 'order to account for the duality of power and subjectification',¹⁶⁶ thus the trained subject becomes the principle of their own regulatory conduct through hierarchical observation, normalising judgement and examination.¹⁶⁷

One of the most common representations of Foucault's paradigm shift is the circle of desks, which decentralise power, creating parity. 'In the first instance', Foucault writes, 'discipline proceeds from the distribution of individuals in space . . . and sometimes requires enclosure, the specification of a place heterogeneous to all others and closed in upon itself.'¹⁶⁸ The circle of desks encloses it to a 'collaborative composition cell . . . transforming the confused, useless or dangerous multitude into ordered multiplicities'.¹⁶⁹ Such 'cells' of desks can be compared to groups, pages liked, Facebook newsfeed content, each individual's specific Facebook friends, Twitter followers, liking hashtags etc. According to Foucault, the challenge for disciplinary spaces in this context is that it must

159 Bucher (n 152) 1166.

160 Murray and Scott (n 84) note, at 502: 'Law in Lessig's terms is merely the constraint placed upon the individual. Accordingly hierarchical control provides both a better label and a substantively enriched conception of this modality.'

161 See, for example, Lessig's concept of the 'pathetic dot' (n 2) 122: 'That someone is regulated is represented by this (pathetic dot) – a creature (you or me) subject to different regulations that might have the effect of constraining (or enabling) the dot's behaviour.' For a discussion of the 'pathetic dot' and the development of dynamic forms of cybernetic and nodal governance models see Murray (n 45) 203–7.

162 Foucault (n 15) 201.

163 Ibid 203.

164 Ibid 170.

165 G Deleuze, *Foucault* (Continuum 2006) 27.

166 Foucault (n 15) 212.

167 Sunstein (n 41) in the broader context of the requirements for the issuance of 'new rules' emphasises five key points for the success of such measures: 1 public participation; 2 advance consultation 3 simplification; 4 harmonisation; and 5 Qualification.

168 Foucault (n 15) 141.

169 P Heilker, 'Discipline and punish and process and paradigms (or Foucault, visibility, (dis)empowerment, and the construction of composition studies)' (1995) 27(2) *Composition Studies* 4–13; Foucault (n 15) 148.

be a 'filter, a mechanism that pins down and partitions; it must provide a hold over this whole mobile, swarming mass'.¹⁷⁰

By viewing regulation in 'cells' (in the 'collaborative unit' sense, rather than a prison room), self-discipline and interdisciplinary activity can help to 'organise' the space. The perfect disciplinary space, Foucault writes, would be constructed to 'permit an internal articulated and detailed control – to render it visible to those who are inside it', thus transforming them 'to act on those it shelters, to provide a hold on their conduct, to carry the effects of power right to them, to make it possible to know them, to alter them' through 'the calculation of openings, of filled and empty spaces, passages and transparencies'.¹⁷¹

The value of such modelling is that, in this context, disciplinary power 'has its principle not so much in a person as in a certain concerted distribution of bodies, surfaces, lights and gazes'¹⁷² and thus 'anyone can come and exercise . . . the functions of surveillance'.¹⁷³ In the disciplinary space, therefore, 'it does not matter who exercises power . . . any individual, taken almost at random, can operate the machine', indeed 'the more numerous those . . . temporary observers are, the greater the risk of the [subject] being surprised and the greater his anxious awareness of his being observed'.¹⁷⁴ If an appropriate technological architecture can be adopted which creates an environment for the mechanics of observation, 'the ability to penetrate into men's behaviours; knowledge follows the advances of power, discovering new objects of knowledge over all the surfaces on which power is exercised'.¹⁷⁵ The advantage of this is that it can substitute for a 'power that is manifested through the brilliance of those who exercise it, a power that insidiously objectifies those on whom it is applied; to form a body of knowledge about these individuals, rather than to deploy the ostentatious signs of sovereignty'.¹⁷⁶

As modern day drug smugglers have recently used improvised submarines and semi-submersible vessels to transport their illegal narcotics to their wider distribution channels, the 'Dark Net' will always provide a three-dimensional space for criminal activity to take place undetected. Regulation cannot go beneath social media's surface to the dark sea bed and seek to regulate it completely; instead, it must provide a principles-based code for the safe and transparent passage of users on the surface.

As complete regulation is not possible, instead, the sites could rely on a self-policing community to police content through reporting materials that infringe the YouTube community guidelines.¹⁷⁷ It is suggested that the use of voluntary lifeboat persons

170 Foucault (n 17) 144.

171 Foucault (n 15) 148, 204, 72.

172 Ibid 148, 204, 202.

173 Ibid 148, 204, 207.

174 Ibid 148, 202.

175 Ibid 148, 204.

176 Ibid 220.

177 The community guidelines can be accessed via <www.youtube.co.uk/t/community_guidelines>. In the case of terrorism-related material, objections could fall in the categories 'violent or repulsive conduct', including subcategories for 'physical attack' or 'promotes terrorism'. Recognising that there are shades of grey and not all videos may be specifically calling individuals to commit terrorist activity, content can also be reported if it amounts to 'hateful or abusive content' that 'promotes hatred or violence'.

remains a viable solution for the regulation of social media which does not amount to legally regulated infringement of the principles.¹⁷⁸

In a multi-regulatory environment, and recognising the unique challenges of social media, the social media community of users¹⁷⁹ itself should also share a responsibility to make use of the technologies such as privacy settings, content blockers, report buttons and flagging of breaches of community standards in order to take personal responsibility for the content posted online and how it is viewed, shared and engaged with by other users.¹⁸⁰ Bucher makes an interesting observation with regards to punishment, which goes beyond the traditional notion of 'regulatory censure'. Not conforming with the rules and architecture is punishable in the sense that it will reduce the users visibility on the platform, thereby limiting their participation in the discourse which takes place online either through the sites removing content or other users choosing to block that user, un-

178 For example just a few days after the terrorist attacks in Paris in 2015, Anonymous announced #OpParis 'this is only the beginning, ISIS. We will hunt you, take down your sites, accounts, emails and expose you . . . You will be treated like a virus and we are the cure', stated a masked crusader in a video posted on YouTube. 'We are Anonymous. We are legion. We do not forgive. We do not forget. ISIS, it is too late to expect us' and 'Anonymous from all over the world will hunt you down. You should know that we will find you and we will not let you go. We will launch the biggest operation ever against you. Expect massive cyber-attacks. War is declared. Get prepared. The French people are stronger than you and will come out of this atrocity even stronger.' As a result of the Oklahoma shootings, a hacker by the name of WauchulaGhost hacked an ISIS account going under the Twitter handle @gi_h_a_d35, changing the profile picture to a gay flag, accompanied by the following message: 'Hello World. It's time I share with you a little secret . . . I'm Gay and I'm Proud!! #GayPride #OrlandoWillNotBeForgotten !!! #GhostOfNoNation.' WauchulaGhost has also been tweeting out IP addresses, phone numbers and other contact information for other hackers to get busy with. WauchulaGhost published a statement regarding the motivation for his actions: 'I did it for the lives lost in Orlando. Daesh [ISIS] have been spreading and praising the attack, so I thought I would defend those that were lost. The taking of innocent lives will not be tolerated. Our actions are directed at Jihadist extremists. Many of our own [group of hackers] are Muslim and we respect all religions that do not take innocent lives.' R Lee, 'Anonymous hacks ISIS's Twitter, makes it as fabulously gay as humanly possible' (*Techly.com*, 16 June 2016) <www.techly.com.au/2016/06/16/anonymous-hacks-isis-twitter-makes-it-as-fabulously-gay-as-humanly-possible>. As well as blocking accounts, Anonymous has also launched cyber-attacks on Turkey, accusing the country's leaders of supporting ISIS. In one operation, Anonymous claimed to have brought down 40,000 websites across Turkey by attacking the country's 'root servers' and threatened to sabotage servers of Turkey's airports, banks, military services and government facilities if they failed to stop aiding ISIS. Sabotage has also been a fruitful avenue of hactivism, for example the Ghost Security Group (an Anonymous affiliated group) has plunged into the Dark Web to find ISIS recruitment hubs and donation pages, replacing them with ads for Prozac and Viagra: M Stainer, 'Ghost Sec, Anonymous affiliate, hacks ISIS site on deep web with Viagra, Prozac ad' *Washington Times* (Washington DC, 26 November 2015).

179 See Scaife (n 12) 3–7 for a chronology of the development of social media communities. See M Bakardjeva and G Gaden, 'Web 2.0 technologies of the self' (2012) 25(3) *Philosophy and Technology* 408–11, for an analysis of the literature relating to the ways in which the self is constituted through peer groups and social medias wider circle of 'friends'.

180 Murray (n 45) 215–20 provides an excellent analysis of the second regulatory state. Explaining how 'collective decisions and communities form, as well as the complexity of regulatory control systems', he cites Burris et al's explanation of nodal governance: 'Take any group of people living in the same place or on some other basis identifying themselves as a group for at least some important purposes. We will refer to this as a "collectivity". The things these people do create outcomes over space and time. These outcomes are not necessarily the result of their intentional activities or of their activities alone. Outcomes are produced by the complex interaction of what people do, how they relate to one another, the institutions, technologies and mentalities they deploy, their biological equipment and the conditions and stimuli from the larger physical and social environment in which they operate. Any given collectivity can be understood to be an "outcome generating system" (OGS). By this we mean that the products of the operation of the collectivity across space and time are not exogenous but organic to the collectivity, reflect the conditions of the OGS, and in turn influence the further development of the system over time. Individual collectivities constitute larger and even more complex systems. A system of collectivities is simply a larger OGS.' S Burris, P Drahos and C Shearing, 'Nodal governance' (2005) 30 *Australian Journal of Legal Philosophy* 30, 37.

follow or report them. As asserted by Foucault, 'what is specific to the disciplinary penalty is not observance, that which does not measure up to the rule, that departs from it'.¹⁸¹ The 'goal' is not to create a perfectly regulated space – indeed, this is not possible, rules will inevitably be circumvented or broken – it is instead to create a regulatory environment in which those users 'manning the world's ships' are alert, motivated, educated, trained and qualified to the proper standards and in fact possess the skills necessary to engage with social media in a manner consistent with the regulatory principles.¹⁸²

7 Reconceptualising the regulation of social media

7.1 DEVISING A METHODOLOGICAL APPROACH TO THE REGULATION OF SOCIAL MEDIA BY COMPARISON TO SOFTWARE DEVELOPMENT MODELLING

As with all emerging technologies, software engineering has faced technological challenges, heavily influenced by 'hardware-oriented development approaches on software development'.¹⁸³ In response to this in the 1960s and 1970s, industry considered how best to control the development of complex software projects. The technique developed¹⁸⁴ systems that relied heavily on upfront process. This method came to be known as the waterfall approach due to its common graphic representation. In its most basic representation, the waterfall model has the main blocks as shown in Figure 1.

181 Foucault (n 15) 148, 204, 202, 178. In September 2015, with the support of stakeholders, the CPS ran its first social media campaign on this issue. The #ConsentIs campaign subsequently won the 2016 UK Public Sector Communications Award, recognising its impact in provoking discussion and understanding of the issue. The campaign was launched with the support of a number of stakeholders, including Rape Crisis, Survivors Manchester, End Violence against Women, National Union of Students and White Ribbon Campaign. The CPS worked closely with these groups beforehand, interviewing them for three #ConsentIs videos which discussed what consent is, myths and stereotypes about consent, and where consent fits into wider society. The public was asked to engage on Twitter using the hashtag #ConsentIs. On its first day of launch, the hashtag was trending on Twitter and over the next two-and-a-half weeks it had a potential reach of almost 12 million accounts, potentially being seen over 45 million times. In addition, over 4000 tweets were sent using the hashtag, and a variety of people became involved in the discussion, including local police forces, MPs, local councils, prominent charities, media commentators and members of the public. CPS (n 36) 54. It is interesting to note that in July 2016, the CPS won the 2016 UK Public Sector Communications Award for its campaign in the category 'Low Budget Campaign of the Year', highlighting that regulation does not necessarily mean that large budgets are required to assist with compliance.

182 Black (2012) (n 1). See n 149 above.

183 Steve Palmquist, Mary Ann Lapham, Suzanne Garcia-Miller, Timothy Chick and Ipek Ozkaya, *Parallel Worlds: Agile and Waterfall Differences and Similarities* (Software Engineering Institute 2013).

184 Winston Royce is often credited with being the father of the waterfall methodology as set out in his paper 'Managing the development of large software systems' (1970) 26 Proceedings of IEEE WESCON 328–88. Royce never envisioned his paper as a solution to complex software development problems. 'He was always a proponent of iterative, incremental, evolutionary development. His paper described the waterfall as the simplest description, but that it would not work for all but the most straightforward projects. The rest of his paper describes [iterative practices] within the context of the 60s/70s government contracting models (a serious set of constraints).' See C Larman and V R Basili, 'Iterative and incremental developments. a brief history' (2003) 36(6) Computer 47–56.

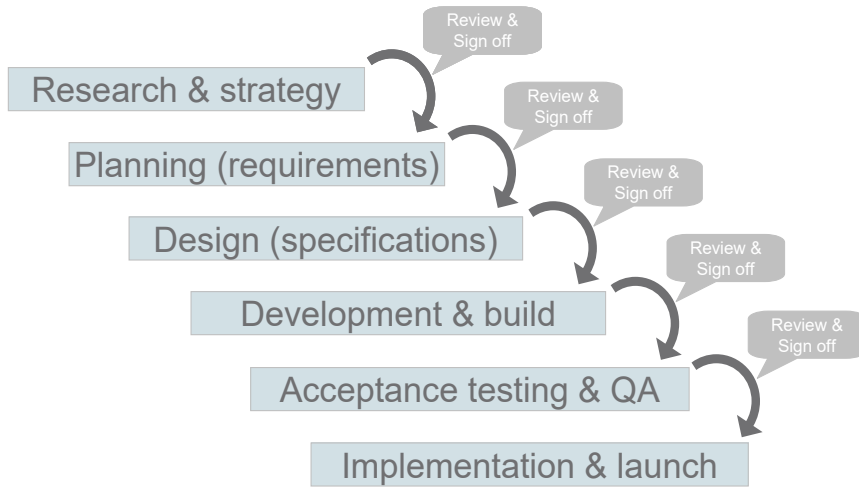


Figure 1

As can be seen from the Figure 1, the process is sequential, and all requirements must be known upfront, with the analysis and design undertaken proceeding on the assumption that those assumptions will remain unchanging, which leads to a rigid, document-intensive and review-intensive approach. As noted by Palmquist et al, such an approach does not pay heed to the customer, who is most involved in integrating the requirements. All of the core features of the product need to be locked down upfront and all aspects of the design and decision-making must be documented before work can commence to ensure that the architecture is adequate.¹⁸⁵

Recognising the pitfalls of the waterfall approach, a more iterative approach to modelling began to develop, called the ‘Agile method’. Agile is an umbrella for a collection of methods and/or approaches that share common characteristics and/or goals. One definition for Agile offered by Lapham is ‘an iterative and incremental (evolutionary) approach to software development which is performed in a highly collaborative manner by self-organising teams within an effective governance framework with “just enough” ceremony that produces high quality software in a cost effective and timely manner which meets the changing needs of its stakeholders’.¹⁸⁶ Agile shares some commonalities with the waterfall methodology, insofar as requirements are characterised upfront. However, they are assumed to be changing, which allows for processes to evolve in short iterations where work is analysed, tested and created in ‘huddles’ which feed into and off one another. This can

185 Palmquist et al (n 183) 5. More iterative approaches to waterfall have been developed. See Larman and Basili on Walker Royce, Dr Royce’s son, said this of his father and the paper: ‘System Require Software Require Analysis Program Design Coding Test Operations’ CMU/SEI-2013-TN-021 79: ‘He was always a proponent of iterative, incremental, evolutionary development. His paper described the waterfall as the simplest description, but that it would not work for all but the most straightforward projects. The rest of his paper describes [iterative practices] within the context of the 60s/70s government contracting models (a serious set of constraints).’ (Larman and Basili (n 184), as quoted by Palmquist et al (n 183). A consideration of such modelling is beyond the scope of this article but is considered in depth by Palmquist et al. An important point here is that Agile is a disciplined planning process, including understanding requirement dependencies, potential groupings and infrastructure needs. Agile planning also includes other technical practices such as configuration management, testing, and the like as part of this disciplined planning perspective.

186 Mary Ann Lapham, Ray Williams, Charles (Bud) Hammons, Daniel Burton and Alfred Schenker, *Considerations for Using Agile in DoD Acquisition* CMU/SEI-2010-TN-002 (Software Engineering Institute, Carnegie Mellon University 2010) <www.sei.cmu.edu/library/abstracts/reports/10tn002.cfm>.

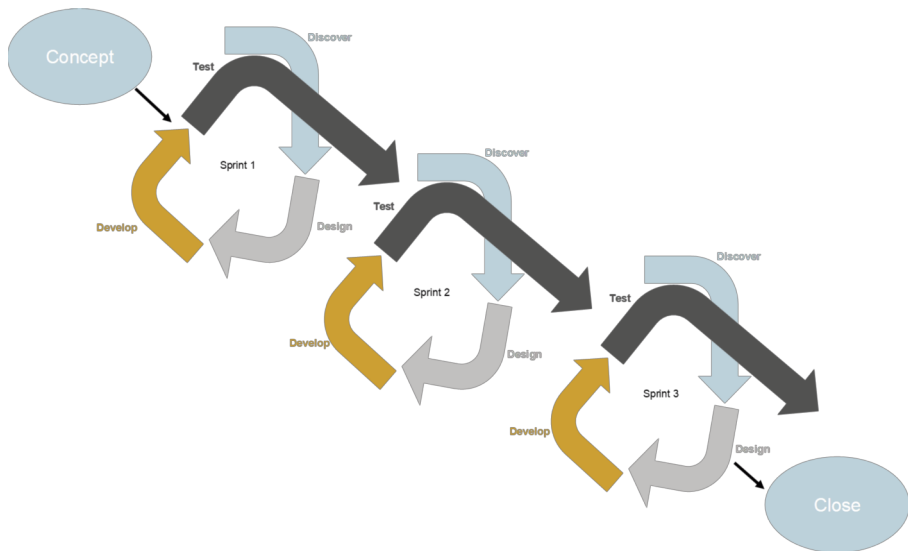


Figure 2

be usefully compared to Foucault's circle of desks, which decentralise power, creating parity. Crucially, there is stakeholder involvement in the process at each iteration, so that the documentation develops only as it is needed and tailored to fit the project.

This type of iterative modelling can be useful if located within constitutional dialogue theory. The employment of a 'constitutional dialogue', in 'constitutional dialogue theory',¹⁸⁷ tends to refer to *equality* between dialogic partners, or at least a (temporary) ignoring of power relations, moving from hierarchy to heterarchy as a basic structure within which the disciplining of public power takes place.¹⁸⁸ The dialogue also enjoys a deliberative quality.¹⁸⁹ Although participants still have agendas and interests to defend, they are fused into their relationships with their dialogic partners and the agendas and interests those partners may have, even where such interests conflict.¹⁹⁰ For example, in the content of international dialogues, Jackson considers the *transnational nature of the process* leading to the incorporation of certain human rights norms into national constitutions.¹⁹¹ According to Meuwese and Snel, the value of such dialogue is that the constitutional actors do not need to acknowledge that they are engaged in a dialogue.¹⁹² They suggest that the instances of dialogue may be called 'constitutional' 'because of their *aspiration* to contribute to, implement or even shape the basic norms that govern the

187 E.g. B M Bakker, 'Blogs as Constitutional dialogue: rekindling the dialogic promise?' (2008) 63 New York University Annual Survey of American Law 215–67; G Dor, 'Constitutional dialogues in action: Canadian and Israeli experiences in comparative perspectives' (2000) 11 Indiana International and Comparative Law Review 1–36.

188 R S Allan, 'Constitutional dialogue and the justification of judicial review' (2003) 23 Oxford Journal of Legal Studies 563–84.

189 Staffan Himmelroos, 'Discourse quality in deliberative citizen forums – a comparison of four deliberative mini-publics' (2017) 13(1) Journal of Public Deliberation, Article 3.

190 M Claes and M de Visser, 'Are you networked yet? On dialogues in European judicial networks' (2012) 8(2) Utrecht Law Review 100–114, 105.

191 V C Jackson, 'Constitutional dialogue and human dignity: states and transnational constitutional discourse' (2004) 65 Montana Law Review 15–40.

192 A Meuwese and M Snel, 'Constitutional dialogue: an overview' 9(2) Utrecht Law Review 123–40.

actions of public entities and their exercise of coercive power in particular'.¹⁹³ This is not to presuppose that the model need be a codified 'dialogical mechanisms' as is present in some constitutions (such as Article 33 of the Canadian Charter). Indeed, whilst this may be suitable for the courts, technology providers and policy makers, it would not be appropriate for the regulated community of users: the act can still be constitutional as an *institutionalised practice*.¹⁹⁴ An example of a simple Agile model is shown in Figure 2.

7.2 An iterative and dynamic model for the regulation of social media

The great value of the rules which had been developed for maritime trade lay in the fact that they had been found by practice to be suitable to the needs of a community which knows no national boundaries. Just like the Law of the Sea and of the *Consolato del Mare* (which represented compilations of comprehensive rules for all maritime subjects, covering everything from ownership of vessels, the duties and responsibilities of the masters or captains thereof, duties of seamen and their wages, freight, salvage, jettison, average contribution, and the like, enjoying an authority far beyond the port from which it hailed), a principles-based approach can be adopted to create an environment in which law, users and technology can work together to create a co-regulated principles-based space which has the necessary flexibility to adapt to the development of technology, whilst respecting core principles of human rights. Taking these principles as a base, international regulation can be developed which has the necessary flexibility to take into account the state of the art and sustain a rigorous regulatory environment, with a better and more effective outcome-focused approach despite being applied, rather than artificially seeking to apply detailed rules prescribing how outcomes must be achieved,¹⁹⁵ instead creating a system which can adapt based on performance assessments and modification of the approach adopted.¹⁹⁶ It is suggested that an Agile methodology could be applied to the regulation of social media.

It has been noted by several authors **Bucher (2014)**¹⁹⁷ that, in the Foucauldian sense, the diagrammatic representation of regulatory architecture highlights the 'distribution of individuals in relation to one another, of hierarchal organisation, of dispositions of centres and channels of power'¹⁹⁸ and 'it is the diagram of a mechanism of power reduced to its ideal form'¹⁹⁹ representing 'a forceful framework for understanding visibility, to highlight the 'distribution of individuals in relation to one another, of hierarchical organisation, of dispositions of centres and channels of power'.²⁰⁰ However, whilst principles-based regulation informs the technological architecture that powers social media to make things more visible, what is clear is that something more flexible is required to regulate social media effectively, something that moves away from an outcomes-based focus of a social media panopticon and more towards overcoming and working with the limitations of its architecture to find a new blueprint. Crucially, therefore, the model must depart from a purely panoptic model to allow for

... a sequel of implicitly or explicitly shaped communications back and forth between two or more actors characterized by the absence of a dominant actor –

193 N Walker, 'The idea of constitutional pluralism' (2002) 65(3) *Modern Law Review* 317–59.

194 Muwese and Snel (n 192) 131.

195 Black (2012) (n 1). See n 149 above

196 Black and Baldwin (2012) (n 1) 140.

197 Bucher (n 152); Sauter (n 9); Kohl (n 5); Spar (n 42); Bucher (n 152).

198 Foucault (n 15) 205.

199 Bucher (n 152).

200 Ibid.

or at least by a bracketing of dominance – with the shared intention of improving the practice of interpreting, reviewing, writing or amending constitutions.²⁰¹

8 The regulatory harbour model

Drawing upon the theoretical perspectives explored in this essay and the way in which the Laws of the Sea were created, this paper will now propose a framework through which social media can be regulated, by offering a holistic and iterative approach to the design of the regulatory architecture.²⁰²

Firstly, core principles such as freedom of expression, the state's ability to protect its citizens and privacy concerns (the 'arches of social media') could be adopted as the principles envisaged which must be designed and/or drafted into the vision for regulation. At this point regulation can be drafted that is sufficiently broad to protect principles, by reference to the state of the art of technology, backed by sanctions. These high-level requirements will by their nature be relatively coarse, setting out the overall scope of the principle to be respected. To recognise the inherent broadness of a principles-based approach, and as noted above,²⁰³ the developer community has a role to play in considering how to deliver the outcomes, which has the necessary breadth to find a closer fit between achieving their business objectives and meeting regulatory requirements.²⁰⁴ This can be achieved though the parties agreeing an overarching roadmap, which, however incomplete, has the capacity to evolve. High-level outputs can be agreed by putting the principles into the context of the technological environment in which they operate, which, over time, should lead to a situation by which the discourse allows the parties to understand one another and communicate more effectively.²⁰⁵

Engagement with the technological community may, of course, lead to concerns that solutions are inconsistent and that there is a lack of transparency in decision-making, as well as costs for the state regulators in terms of creating continuity. It is suggested that to counter this risk and foster a consistent approach to regulation, 'peer panels'²⁰⁶ comprised of the main social media platform providers and policy-makers could be adopted to foster the development of common language about risk and to facilitate learning.²⁰⁷ The solution can then be finessed to balance the government's priorities and human rights obligations and understand industries' risk (such as cost, complexity of delivery etc.) in order to resolve the construction, deployment and operation of the

201 Muwese and Snel (n 192) 126.

202 For the purposes of this modelling, the notion of rules, is adopted in the sense of norms as described by Murray and Scott, namely 'the preferred meaning of the word norm is as the generic term of standards, guidelines, and legal and non-legal rules' (n 84) 503. See also P Drahos and J Braithwaite, *Global Business Regulation* (Cambridge University Press 2000) 20.

203 Black (2012) (n 1) 218: 'the rule to be applied in a way which will further the overall aims of the rule maker, then the person applying it has to share the rule maker's interpretation of it'.

204 For Black (2012) (n 1) 219: 'part of the aim behind the Principles was to lift the issue of compliance from compliance officers up into the boardroom, and to help senior executives to see the moral wood for the technical trees'.

205 For example, see Murray and Scott's (n 84) control methodology of: elemental, hierarchical, community-based, competition and design-based control(s), Figure 1 at 504.

206 Black (2010) (n 1); Black (2012) (n 1); Black and Baldwin (2012) (n 1).

207 See Baldwin and Black (2008) (n 1) 71: 'the third element of really responsive regulation is responsiveness to the logics of different regulatory strategies and tools. Different regulatory strategies can have different logics. They embody, or at least place emphasis on, different under-standings of the nature of behaviours or an institutional environment, and in turn have different pre-conditions for effectiveness (namely that the institutional environment of behaviours conforms to those foundational understandings)'.

proposed model. The roadmap can then be broken down into a series of releases, with the high overarching architectural requirements taking precedence over the successive iterations of the technological solutions used to achieve the overall legal principle. When regulation is approached in this manner, the model adopted can take into account the institutional environments in which the regulators act,²⁰⁸ whilst factoring in the view of the regulated without having the need to resort to a radical remodelling of the regulatory approach which would not have the necessary flexibility to take into account the degree of institutionalism that is present in regulation and which cannot be easily divested. The goal should be iterative evolution not prescriptive arbitrary revolution.

As the delivery of any one given principle will involve the consideration of different legal issues (e.g. criminal liability, privacy infringement, defamation, trolling etc.),²⁰⁹ reflecting the reality that there will always be a mixture of detailed rules and principles – ‘rules will never be able to be dispensed with. But they should in a number of areas be able to become less prescriptive’²¹⁰ – the releases can be broken down into sprints which can be considered in manageable blocks. During each iteration the regulator cannot add more requirements to the iteration, but it can clarify its requirements. With the platform providers feeding back on envisaged problems in the process, it is suggested that this is a crucial part of the modelling, given that regulation is ‘not self-executing’ and ‘depends on willingness to engage with the process’,²¹¹ as those to whom the rules apply will be less willing to comply if they are not informed or engaged with as to what compliance will mean and/or involve.²¹²

Once the parties are satisfied, the policy/principle can be released, in a way that has allowed the platform providers to consider the architecture issues in a timely manner, prior to the implementation of the regulation. There is, of course, an important role to be played in terms of the delivery to the wider social media community (such as companies, individual users etc.). It is therefore suggested that, when the draft legislation is prepared, information campaigns, guidance notes, training and awareness, and stakeholder consultation could be undertaken with interested parties to improve the design and implementation of the principles. It is suggested that the outcome of such consultations would be best reflected through the adoption of detailed guidance notes issued by an appropriate centralised regulator which reflect cultural, political and community standards. The advantage of adopting such a model is that the process need never be considered ‘complete’ and, if there is sufficient friction in any one of the given sprints to cause the creation of a reverse current when the water attempts to flow past an obstacle (known in marine terms as a ‘back eddy’), there will always be an inlet in the preceding harbour which can accommodate the water.²¹³ Therefore unlike the traditional

208 Ibid 70.

209 It is beyond the scope of this article to consider these discrete areas of regulation. For an analysis of the issues as applied to intermediaries see Kohl (n 5).

210 Black (2012) (n 1) 220.

211 Black and Baldwin (2012) (n 1) 141.

212 Baldwin (n 1) 336.

213 Michael Hill, ‘The policy/implementation distinction: a quest for rational control’ in Barrett and Fudge (n 1) has argued that ‘the policy-implementation distinction will tend to be more clearly drawn when, at some point in the chain, one of the parties asserts a right to prescribe goals for those at subsequent points in the chain . . . where one party regards itself as having a legitimate right to pay down ‘a policy’ for others and where it has powers to do so. It is suggested that by averring from this narrowness of perspective the harbour model will allow for enough stakeholder engagement to ensure that policy is not prescriptively delivered in a top-down fashion. For Baldwin and Black (2008) (n 1) 72: ‘performance sensitivity rests on the regimes ability both to assess its performance in the light of its objectives and to modify its tools and strategies accordingly’.

drafting model, which would lead to a situation whereby an intolerable strain would be placed on the legislation seeking to contain the issue, there is always a pressure-release system which can accommodate such backward currents.²¹⁴ The model for social media which has been suggested embraces this complex matrix, placing it at its heart, accepting the tension and allowing ways for it to ebb and flow, whilst always accepting that there will be an undercurrent in the water. In this way, a framework can be offered which can respond to attitudinal setting, institutionalisation and the private sector's interests, creating sensitive models which respond to change.²¹⁵

Inevitably, there will be situations where content falls short of an infringement of the legal principles and/or involves individual users or involves matters of a trivial and/or low speech value nature.²¹⁶ Whilst in the overall context they are low risk,²¹⁷ they create disharmony within the community. Much like the seas, it is considered that this is best dealt with by fellow seafarers, by adopting a Foucauldian model of visibility, such as 'trusted users' who will be best placed to engage with such matters within their 'community cells'. In certain circumstances, a more formal approach may be required and, in such scenarios, harbourmasters or coastguard authorities (courts, policy makers etc.) could be utilised. However this should not be the first (choice) 'port of call'. The Harbourmasters should be used sparingly for enforcing the regulations of a particular harbour or port, in order to ensure the safety of navigation, the security of the harbour and the correct operation of the port facilities. coastguard agencies (policymakers, prosecution services etc) could be responsible for issuing local 'safety information' as they do for seafarers. The outcome of this process can thereby lead to a process by which appropriate community standards, acting as a contract between the user and the platform provider, can be adopted that have community stakeholder engagement. In this way the 'community of persons' can adjust its behaviours based on community norms and the gaze of fellow users, as well as the regulator and the platform provider.

A representation of the proposed model is shown at Figure 3:

214 See Baldwin (n 1) 334–5: 'The message here is that a more realistic approach to rule choice can be adopted, that it is a mistake to assume that a rule will emerge with the same degree of rigour and as in the same form as the first draft . . . trade offs should be made in the real not the ideal world'. Baldwin also notes Eugene Bardach's metaphor of the 'game' (E Bardach, *The Implementation Game* (MIT Press 1977) 6) in which groups struggle to control the administrative process; the process for Bardach 'directs us to look at the players, what they regard as the states, their strategies and tactics, their resources for playing, the rule so of play . . . the nature of the communications (or lack of them) among the players and the degree of uncertainty surrounding possible outcomes'.

215 Black and Baldwin (2008) (n 1) 72: 'Responsive regulation requires escalation up a punitive scale that crosses logics; smart regulation theory encounters such issues in a more complex framework (and to a degree addresses complementarities and inconsistencies of approach).'

216 For a comprehensive case history of the hierarchy of speech see Scaife (n 12) 145.

217 Black and Baldwin (2012) (n 1) deal comprehensively with the regulation of 'low risk' regulatory activities, it is beyond the scope of this essay to fully consider these issues.

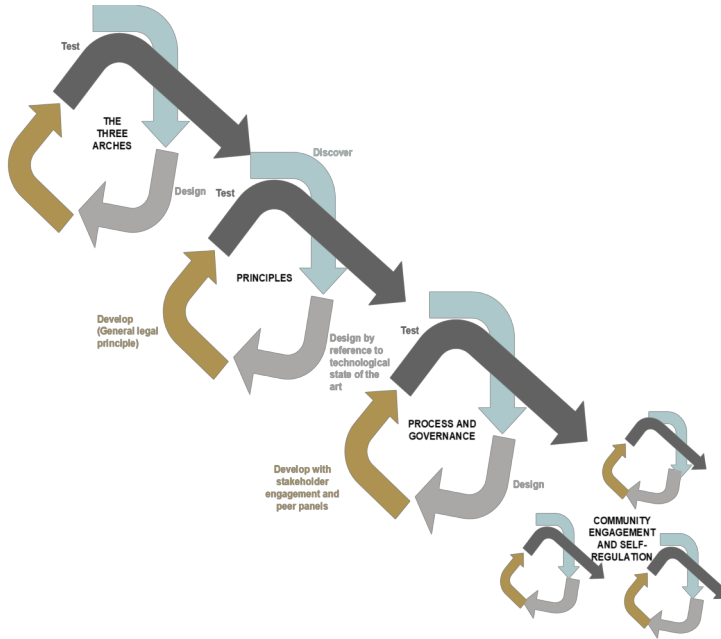


Figure 3

9 Conclusion

Improved rule-making comes when the means of securing compliance is shaped having regard to the particular problem at hand, rather than by ‘clinging to the notion that rules shape the world’²¹⁸ is possible. Through a consideration of the regulation of the sea, this paper has served to illustrate that regulation is possible. It has proposed a theoretical and methodological approach to thinking about the regulation of social media as conceptualised in an iterative and dynamic model that is characterised by the technologisation of human interaction and increasingly transparent ways of living in an age of ‘technologies of the self’. This approach will facilitate a more critical, responsive and iterative awareness of the regulation of expressive content in a way that can grow with the technological state of the art, alive to cultural sensitivities and the use of the tools as a vessel for self-development.

Building upon the work of Murray and Scott,²¹⁹ the model adopted seeks to take into account in its design how norms are obeyed, which leads to the formation of rules whilst acknowledging the role of challenging and encouraging ‘honest perplexity’ as to the rule’s application, so as not to thwart creatively compliant behaviour. It acknowledges the historical context and limitations of this unique medium and avoids a technologically deterministic or meta-narrative approach to regulation of multiple actors. With regards to the comparisons between ‘unregulatable spaces’ such as the sea, care much also be taken between overstating the translatability of such older practices of rule formation. There is a fine line to be drawn in terms of providing a model that has just enough architecture to be used as a blueprint for continuous refinement, and over-engineering the suggested

²¹⁸ Baldwin, (n 1) 337.

²¹⁹ Black (2012) (n 1).

model so that it becomes the very sort of panoptic model which it is seeking to transgress in the search for really responsive regulation.

Whilst such modelling may be viewed as eclectic and broad, with regards to enforcement and compliance the approach has served to demonstrate that there are several ways in which to achieve the overall principles which are deemed worthy of protection, each naturally with its own strengths and weaknesses. Indeed, the challenges of 'such novel governance mechanisms are how to deploy them in such a way that they are perceived legitimate'. The legitimacy of democratic government is linked to the processes of representation and open decision-making.²²⁰ The approach taken towards regulation will naturally vary according to context. Similar to theoretical approaches adopted with regards to really responsive regulation,²²¹ the model and underlying analysis put forward in this article go beyond prescriptive approaches so as to offer the regulator and the regulated a framework for evaluating the relative merit of different approaches and players' involvement and allows for the adoption of innovative combinations of regulatory logic, allowing optimal responses to be developed by the whole community, with the scope for feedback by each player in the community. In this regard, it goes further than simply providing a refinement of existing models, demanding an ongoing consideration of the regulatory strategy to be adopted, beginning with principles and problems rather than the regulation itself. The regulation is responsive because it knows the environment that it seeks to assist rather than attempting to command it, and is responsive to performance sensitivities and changes in tides – waving, not drowning.

220 Ibid 516.

221 Black and Baldwin (2008) (n 1) 94.