Internet Censorship in the United Kingdom: National Schemes and European Norms

TJ McIntyre[1]

This is a pre-print of a chapter to be published in Lilian Edwards (ed), *Law, Policy and the Internet* (forthcoming, Hart Publishing, 2018)

Introduction

The United Kingdom (UK) has been at the vanguard of online censorship in democracies from the beginning of the modern internet.[2] Since the mid-1990s the government has developed distinctive patterns of regulation – targeting intermediaries, using the bully pulpit to promote 'voluntary' self-regulation, and promoting automated censorship tools such as web blocking – which have been influential internationally but raise significant issues of legitimacy, transparency and accountability.[3] This chapter examines this UK experience in light of the European Convention on Human Rights (ECHR) and EU law, arguing that in key regards current censorship practices fail to meet European standards.

There is already an extensive literature in this field: authors such as Akdeniz, Edwards, and Laidlaw have examined the fundamental rights implications of UK policy from a number of legal and regulatory perspectives.[4] The current chapter builds on that work in two main ways. First, it assesses emerging censorship practices in the area of terrorist material and extreme pornography. Second, it considers how recent EU legislation and ECtHR case law might constrain the freedom of the UK government and force a move towards different models of censorship.

The chapter starts by outlining the regulatory context. It then takes three case studies – Child Abuse Material (CAM), terrorist material, and pornography/extreme pornography under the Digital Economy Act 2017 – and traces how censorship has evolved from one context to the next. These systems are then evaluated against the standards set by Articles 6 and 10 ECHR, the Open Internet Regulation[5] and

---

[1] Lecturer in Law, University College Dublin. This chapter draws on material from 'Internet blocking law and governance in the United Kingdom: an examination of the Cleanfeed system' (University of Edinburgh, 2014) and 'Content, control and cyberspace: The end of Internet regulatory forbearance in the United Kingdom?', a paper presented at 'Governance of New Technologies', SCRIPT, Edinburgh, 29-31 March 2009. Disclosure: the author is chair of civil liberties group Digital Rights Ireland.

[2] The term 'censorship' is a loaded one, but it is used here neutrally as a catch-all to describe state actions which aim to prevent the distribution or viewing of types of material. Censorship in this sense is narrower than 'content regulation' – it refers to schemes which aim to suppress certain material entirely, rather than merely regulating aspects such as how it is published, whether children can see it, or whether it meets the terms of use of a particular service.

[3] See e.g. Ben Wagner, *Global Free Expression - Governing the Boundaries of Internet Content*, Law, Governance and Technology Series (Cham, 2016), chap. 4.

[4] See e.g. Yaman Akdeniz, *Internet Child Pornography and the Law: National and International Responses* (Aldershot, 2008), chap. 9; Lilian Edwards, 'Pornography, Censorship and the Internet', in *Law and the Internet*, ed. by Lilian Edwards and Charlotte Waelde, 3rd edn (Oxford, 2009); Emily Laidlaw, 'The Responsibilities of Free Speech Regulators: An Analysis of the Internet Watch Foundation', *International Journal of Law and Information Technology*, 20/4 (2012), 312.

[5] Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 laying down measures concerning open internet access and amending Directive 2002/22/EC on universal service and users'

the Directives on Sexual Abuse of Children[6] and on Combating Terrorism.[7] The chapter concludes by considering what lessons we can learn from the UK experience.

PART 1: UK CENSORSHIP SCHEMES

1. UK regulatory context

UK government policy towards internet content has been very different from that applied to media such as television, cinema, and video games. In those areas the norm has been sector-specific legislation overseen by statutory regulators.[8] For the internet, however, successive governments have opted for 'legislative forbearance': application of the general law rather than sector-specific legislation, overseen by industry self-regulation rather than statutory bodies.[9] Until recently, internet content regulation in the UK has largely been carried out through a patchwork of government promoted self-regulatory systems. To summarise the most important examples:

- The Internet Watch Foundation (IWF) has operated a notice and takedown system for CAM since 1996, a URL blocking list[10] since 2004 and more recently a range of other responses including worldwide proactive searches[11] and an image hash list[12] to enable intermediaries to detect and block uploads of files.
- Under the Mobile Operators' Code of Practice, mobile operators have age-rated and blocked certain content accessed on mobile phones since 2004, using a framework developed by the British Board of Film Classification (BBFC).[13]
- Since 2008 several filtering software companies have blocked webpages which police identify as illegally terror-related, under a confidential agreement with the Home Office.[14]
- Since 2010 the Counter Terrorism Internet Referral Unit (CTIRU) has notified sites such as Facebook and YouTube of material which it deems to be illegally terror-related, for voluntary takedown as violating their terms of use.[15]

---

rights relating to electronic communications networks and services and Regulation (EU) No 531/2012 on roaming on public mobile communications networks within the Union.

[6] Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA.

[7] Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA.

[8] See e.g. Geoffrey Robertson and Andrew Nicol, *Media Law*, 4th edn (London, 2002), chap. 15 and 16.

[9] David Currie, 'The Ofcom Annual Lecture 2008', 2008 <http://www.ofcom.org.uk/media/speeches/2008/10/annual_lecture> [accessed 21 January 2009].

[10] See e.g. T.J. McIntyre, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems', in *Research Handbook on Governance of the Internet*, ed. by Ian Brown (Cheltenham, 2013).

[11] Tony Prosser, 'United Kingdom: New Proactive Approach to Seek Out Child Pornography', *IRIS Legal Observations of the European Audiovisual Observatory*, 2013 <http://merlin.obs.coe.int/iris/2013/8/article22.en.html> [accessed 21 July 2017].

[12] 'Image Hash List', *IWF* <https://www.iwf.org.uk/our-services/image-hash-list>.

[13] Unless a subscriber verifies that they are an adult. See 'Codes of Practice', *Mobile UK* <http://www.mobileuk.org/codes-of-practice.html>; Christopher Marsden, *Internet Co-Regulation: European Law, Regulatory Governance and Legitimacy in Cyberspace* (Cambridge, 2011), 139–46.

[14] Jane Fae, 'The Internet Censorship Programme You're Not Allowed to Know About', *politics.co.uk*, 2014 <http://www.politics.co.uk/comment-analysis/2014/03/27/the-internet-censorship-programme-you-re-not-allowed-to-know> [accessed 30 March 2016].

[15] ACPO, 'CTIRU Factsheet', 2010 <http://www.acpo.police.uk/documents/TAM/CTRIU%20factsheet.pdf> [accessed 10 March 2015].

- Since 2013 all major fixed line ISPs have presented subscribers with an 'unavoidable choice' whether to activate 'family friendly filters', blocking access to content unsuitable for children.[16] The main providers of public wifi – covering over 90% of the market – have also agreed to impose these filters on all users of their public networks.[17]
- Since 2014 the .uk registry Nominet has policed new registrations to prevent use of domain names which appear to 'promote sexual offences'.[18]

There has been little co-regulation or statutory regulation. The most significant is the regulation of on-demand audiovisual media services ('TV-like' services) by the Authority for Television on Demand (ATVOD) and Ofcom from 2009 onwards.[19] It is telling that this was not a domestic initiative but was forced on a reluctant government by the Audiovisual Media Services Directive.[20] As Petley notes, historically 'British governments generally do not like to appear to be playing the censor and are far happier when they can instigate apparently "self-regulatory" systems in which they play a key role, albeit very much behind the scenes'.[21]

Why has self-regulation been so dominant? In part the UK is simply reflecting a global consensus on the expediency of this approach; in 2001 the Cabinet Office adopted the principle that internet self-regulation 'generally provide[s] a more rapid and flexible means of responding to changing market needs, and achieving international consensus, than is possible through legislation'.[22] However, there are also significant domestic factors at play.

Since the early years of the Thatcher government, trends such as privatisation, outsourcing and deregulation have led to a 'contracting state' – 'contracting' in the senses of shrinking, contracting out functions to the private sector and also using contracts as a tool of governance.[23] The result has been a 'post-regulatory state' in which more emphasis is placed on approaches such as self-regulation, soft

[16] Ofcom, 'Ofcom Report on Internet Safety Measures: Strategies of Parental Protection for Children Online', 2015, 3–6 <http://stakeholders.ofcom.org.uk/binaries/internet/fourth_internet_safety_report.pdf> [accessed 3 March 2016].
[17] 'The Internet and Pornography: Prime Minister Calls for Action', *GOV.UK*, 2013 <https://www.gov.uk/government/speeches/the-internet-and-pornography-prime-minister-calls-for-action> [accessed 22 July 2013].
[18] 'Nominet Formalises Approach to Tackling Criminal Activity on .Uk Domains', *Nominet*, 2014 <http://www.nominet.org.uk/news/latest/nominet-formalises-approach-tackling-criminal-activity-uk-domains> [accessed 21 August 2014].
[19] Initially by ATVOD under a co-regulatory agreement with Ofcom, until Ofcom took the function fully in-house in 2015. See Jenny Metzdorf, 'The Implementation of the Audiovisual Media Services Directive by National Regulatory Authorities: National Responses to Regulatory Challenges', *JIPITEC*, 5/2 (2014), 88–104; 'Ofcom Brings Regulation of 'Video-on-Demand' In-House', *Ofcom*, 2015 <https://www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2015/1520333>.
[20] Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services.
[21] Julian Petley, 'The Regulation of Pornography on Video-on-Demand in the United Kingdom', *Porn Studies*, 1/3 (2014), 294.
[22] Office of the e-Envoy, 'E-Policy Principles: A Policymakers Guide to the Internet', 2001 <http://tna.europarchive.org/20050311005439/http://www.cabinetoffice.gov.uk/regulation/ria-guidance/documents/pdf/epolicy.pdf> [accessed 3 November 2011].
[23] See e.g. Jody Freeman, 'The Contracting State', *Florida State University Law Review*, 28 (2000), 155.

law and non-state rulemaking.[24] Against this background the preference for self-regulation online is not an outlier but exemplifies a wider UK tendency.[25]

Structural conditions have also played a part.[26] A highly centralised government with dominance over parliament means that the state can make credible threats of legislation and prosecution unless internet firms implement government policy.[27] This has been helped by the concentrated market for internet access, which enables the government to achieve outcomes covering almost all subscribers by dealing with a small number of ISPs.[28]

There are, however, some developments pointing to a greater use of statutory regulation in future. The Conservative/Liberal Democrat coalition of 2010-15 continued to use self-regulatory tactics, but was also marked by an increased use of legislation, including a contentious widening of the online material censored by ATVOD/Ofcom.[29] Continuing this trend, the Digital Economy Act 2017 has put in place a new statutory scheme requiring pornography websites to introduce age verification, with blocking of sites which don't apply age verification or which host extreme pornography.

As described later in this chapter, developing EU and ECHR norms on freedom of expression and net neutrality[30] will also increasingly constrain self-regulatory controls on internet content and may force greater use of legislation – assuming, of course, that these standards continue to apply given the on-going uncertainty about the effects of Brexit and possible withdrawal from the ECHR.[31]

2. Child abuse material (CAM) and the Internet Watch Foundation (IWF)

The best known UK initiative in this area is the IWF, which was established by the internet industry in 1996 following government pressure on internet service providers (ISPs), including a threat of prosecution should self-regulation not be put in place.[32] It is a charitable body which has been funded by industry and the EU through successive Safer Internet Programmes. It describes its remit as being 'to remove child sexual abuse content hosted anywhere in the world [and] non-photographic child

---

[24] Colin Scott, 'Regulation in the Age of Governance: The Rise of the Post-Regulatory State', in *The Politics of Regulation: Examining Regulatory Institutions and Instruments in the Age of Governance*, ed. by Jacint Jordana and David Levi-Faur (Cheltenham, 2004).

[25] Richard Collins, 'Networks, Markets and Hierarchies: Governance and Regulation of the UK Internet', *Parliamentary Affairs*, 59/2 (2006), 314.

[26] For a comparative empirical analysis of structural factors shaping internet censorship see Patrick Theiner, Yana Breindl and Andreas Busch, 'Internet Blocking Regulations: A Comparative Analysis of 21 Liberal Democracies' (presented at the U4 Cluster Conference "Governance of a Contemporary Multilateral Institutional Architecture", Groningen, 2015).

[27] Wagner, *Global Free Expression - Governing the Boundaries of Internet Content*, 80.

[28] For example, in 2015 four firms (counting BT and EE as one entity following the agreed merger) held 92% of the retail fixed broadband market: Ofcom, 'Communications Market Report 2016', 2016, 149 <https://www.ofcom.org.uk/__data/assets/pdf_file/0024/26826/cmr_uk_2016.pdf> [accessed 3 March 2017].

[29] Daithi Mac Sithigh, 'Computers and the Coalition: Legislation on Law and Information Technology, 2010-2015', *SCRIPTed*, 12 (2015), 141; Daniel Haley, 'Dirty Laws: A Critique of the Audiovisual Media Services Regulations 2014 and Section 63 of the Criminal Justice and Immigration Act 2008', *Cardozo Journal of Law and Gender*, 22 (2015), 493.

[30] Swiss Institute of Comparative Law, *Comparative Study on Blocking, Filtering and Take-Down of Illegal Internet Content* (2017), 24–25.

[31] See e.g. Tobias Lock and others, *Brexit and the British Bill of Rights* (Edinburgh, 6 February 2017) <https://papers.ssrn.com/abstract=2913566>.

[32] DTI, 'DTI Press Release P/96/636', 1996 <http://www.mit.edu/activities/safe/cases/demon/minister-statement.txt> [accessed 28 February 2009]; John Carr, 'A Brief History of Child Safety Online: Child Abuse Images on the Internet', in *Online Risk to Children: Impact, Protection and Prevention*, ed. by Jon Brown (2017).

sexual abuse images hosted in the UK'[33] and the ways in which it implements this remit have expanded significantly over the two decades it has been in existence.

Notice and takedown

The IWF began by establishing an internet hotline to receive and adjudicate on complaints from the public.[34] This operates using a notice and takedown model which is mandatory for members under the Code of Practice; if the IWF issues a takedown notice then members must remove items which they host.[35]

Initially the IWF dealt mainly with complaints about individual Usenet posts. However in 2003, following pressure from the Home Office, the IWF controversially widened its policy to require members to remove entire Usenet newsgroups if it found that the newsgroups 'regularly contained' CAM, or which had 'names judged to support or condone paedophilic activity' – as critics pointed out, suppressing newsgroups in which the overwhelming majority of content was perfectly legal.[36]

Web blocking

This focus on newsgroups became less effective as CAM began to spread from Usenet to websites. These were mostly hosted outside the UK and for that reason beyond the reach of the IWF.[37] From 2001 onwards, ISPs came under pressure from the Home Office to introduce network level blocking of child pornography on the web, including a threat to introduce legislation unless the entire industry 'voluntarily' did so.[38] Though some smaller ISPs have refused to implement blocking, all major ISPs have acquiesced and in 2009 the Home Office eventually shelved plans to legislate when an Ofcom survey confirmed that over 98% of home connections were subject to blocking.[39]

The scheme developed by the IWF – popularly but inaccurately known as 'Cleanfeed'[40] – provides members with a list of URLs to block; unlike compliance with takedown notices, however, blocking is not a condition of membership and members have discretion as to whether and how they will block.

---

[33] Internet Watch Foundation, 'Our Remit and Vision', *IWF* <https://www.iwf.org.uk/what-we-do/why-we-exist/our-remit-and-vision> [accessed 7 August 2017] The IWF had previously (reluctantly) taken on functions in respect of criminally obscene adult content, extreme pornography, and hate speech; however by mid-2017 it had succeeded in divesting itself of these responsibilities.

[34] 'Our History', *INHOPE* <http://www.inhope.org/gns/who-we-are/Ourhistory.aspx> [accessed 14 February 2018].

[35] Internet Watch Foundation, 'FC Code of Practice', *IWF* <http://www.iwf.org.uk/what-we-do/who-we-are/governance/funding-council/fc-code-of-practice> [accessed 15 February 2018].

[36] Akdeniz, *Internet Child Pornography and the Law*, 256–57; Wendy Grossman, 'IWF: What Are You Looking At?', *The Independent*, 25 March 2002 <http://www.independent.co.uk/news/business/analysis-and-features/iwf-what-are-you-looking-at-655425.html>.

[37] John Carr, *Child Abuse, Child Pornography and the Internet* (London, 2004), 18–19 <http://www.make-it-safe.net/esp/pdf/Child_pornography_internet_Carr2004.pdf> [accessed 12 September 2009].

[38] Sean Hargrave, 'Surfing with a Safety Net', *The Guardian*, 29 June 2006 <http://www.guardian.co.uk/technology/2006/jun/29/guardianweeklytechnologysection> [accessed 1 May 2009]; Jane Merrick, 'Internet Providers Face Child Porn Crackdown', *The Independent*, 6 September 2009 <http://www.independent.co.uk/news/uk/crime/internet-providers-face-child-porn-crackdown-1782530.html> [accessed 7 September 2009].

[39] Chris Williams, 'Home Office Backs down on Net Censorship Laws', *The Register*, 2009 <http://www.theregister.co.uk/2009/10/16/home_office_iwf_legislation/> [accessed 16 October 2009].

[40] For the history of the Cleanfeed name see the appendix to T.J. McIntyre, 'Internet Blocking Law and Governance in the United Kingdom: An Examination of the Cleanfeed System' (University of Edinburgh, 2013) <https://www.era.lib.ed.ac.uk/handle/1842/17971>.

The list is updated twice daily and contains about 500 URLs at any time.[41] The list is limited to 'child sexual abuse images', by which the IWF means indecent images of children prohibited by the Protection of Children Act 1978.[42] This does not include all forms of 'child pornography' prohibited by UK law – non-photographic images of children ('virtual child pornography') have not been included by the IWF, due to fears that their inclusion would flood the blocking system and undermine popular support.[43]

Blocking – transparency and remedies

Decisions to add a URL to the list are taken in almost all cases without notice to the site involved.[44] The URL List itself is not publicly available. Nor is there a requirement of notice at the point of blocking. Although the IWF recommends that members use block pages[45] to tell users that a page has been blocked, ISPs are free to disregard that recommendation and many have used fake error messages instead.[46] Similarly, while the IWF recommends that ISPs block only the exact URL on the list, a number of ISPs have used the cheaper and cruder methods of IP address blocking or DNS poisoning instead – resulting in blocking of unrelated pages or even completely unrelated sites in some cases.[47]

Following recommendations in a 2014 human rights audit[48], the IWF has introduced a new appeal system which can be invoked by any person who believes that a URL has been wrongfully assessed.[49] This provides for internal review, review by police and ultimately a formal appeal to an independent inspector (a retired High Court judge); however, this appeal mechanism has yet to be used.[50]

---

[41] 'URL List Policy', *IWF* <http://www.iwf.org.uk/become-a-member/services-for-members/url-list/url-list-policy> [accessed 7 August 2017].
[42] Internet Watch Foundation, 'FAQs Regarding the IWF's Facilitation of the Blocking Initiative', 2011 <http://www.iwf.org.uk/services/blocking/blocking-faqs> [accessed 5 September 2011].
[43] Internet Watch Foundation, 'Board Minutes 29 September 2009', *Internet Watch Foundation*, 2009 <http://www.iwf.org.uk/corporate/page.215.617.htm> [accessed 4 February 2010].
[44] In exceptional cases the IWF may seek takedown at source first. See Internet Watch Foundation, 'URL List Policies Procedures and Processes', *IWF*, 2015 <https://www.iwf.org.uk/sites/default/files/inline-files/URL%20List%20Policies%20Procedures%20and%20Processes%2011.8.15.pdf> [accessed 14 February 2018].
[45] Also known as 'stop pages' or 'splash pages'.
[46] 'Administrators' Noticeboard/2008 IWF Action', *Wikipedia*, 2008 <http://en.wikipedia.org/wiki/Wikipedia:Administrators%27_noticeboard/2008_IWF_action> [accessed 22 December 2008].
[47] Sebastien Lahtinen, 'Be Unlimited Causes Stir in Effort of Blocking Child Abuse Images', *Thinkbroadband.Com*, 2007 <http://www.thinkbroadband.com/news/3235-be-unlimited-causes-stir-in-effort-of-blocking-child-abuse-images.html> [accessed 14 March 2009]; Joe McNamee, 'Blocking of Innocent Websites by O2 Ireland', *EDRi: European Digital Rights*, 2010 <http://www.edri.org/edrigram/number8.14/o2-blocking-websites-ireland> [accessed 2 June 2013].
[48] Ken Macdonald, 'A Human Rights Audit of the Internet Watch Foundation', 2014, 24 <https://www.iwf.org.uk/sites/default/files/inline-files/Human_Rights_Audit_web.pdf> [accessed 16 July 2017].
[49] Internet Watch Foundation, 'Content Assessment Appeal Process', *IWF* <http://iwf.org.uk/content-assessment-appeal-process> [accessed 7 August 2017]; Internet Watch Foundation, 'Content Assessment Appeal Flow Chart Process', *IWF* <https://www.iwf.org.uk/sites/default/files/inline-files/Content%20assessment%20appeal%20flow%20chart%20process.pdf> [accessed 14 February 2018].
[50] Mark Hedley and others, 'Independent Inspection Report 2017', 2017 <https://www.iwf.org.uk/sites/default/files/inline-files/Approved%20amended%20INTERNET%20WATCH%20FOUNDATION%20Hotline%20report%20-%202017.pdf>.

Blocking as a surveillance tool

Should web blocking be used to identify and prosecute users who attempt to view CAM? At the outset, the industry opposed this. When the IWF established the URL List it did so in close partnership with BT, which was the first ISP to introduce blocking. When news of that blocking system emerged, BT responded to criticism by stressing that it did not log the IP addresses of computers which had attempted to reach a blocked URL.[51] There had been some police interest in using the system to identify users – this was made impossible, however, by a deliberate design choice on the part of BT to prevent the system becoming a means of surveillance.

In 2006, however, South West Grid for Learning (SWGfL) began to use the URL List explicitly as a surveillance tool.[52] This was done as a pilot project in 2,500 schools in the South West, under Home Office supervision and with the permission of the IWF.[53] Under this system, attempts to visit blocked URLs were automatically reported to police. In the decade from 2006 to 2016 this resulted in 12 reports being generated; of these, two were false positives, one resulted in a criminal charge, one in a caution and for the remainder there was either insufficient evidence to bring charges or it was not in the public interest to do so.

Following this pilot, the system has been commercialised and can now be bought by schools in the UK as a self-contained network device under the ICAlert brand name.[54] As well as the IWF URL List, the device also uses the CTIRU blacklist (discussed further in the next section of this chapter) and automatically reports to police any attempts to visit URLs on either list. The Metropolitan Police has expressed an interest in rolling the device out to all schools across the UK.[55]

This use of the IWF URL List illustrates issues of function creep and convergence of censorship and surveillance which recur throughout this chapter – in this case, a system initially developed to protect users against inadvertent access to CAM has been repurposed to presumptively criminalise such users. The school setting complicates the assessment of proportionality, as staff and students will have a reduced expectation of privacy in their internet use at school. However, it is hard to see how a system which monitored the internet use of all students and staff at 2,500 schools for a decade but resulted in only one criminal charge could be said to be proportionate, much less one that should be replicated nationwide.

Proactive searching

---

[51] Malcolm Hutty, 'Cleanfeed: The Facts', *LINX Public Affairs*, 2004 <https://publicaffairs.linx.net/news/?p=154> [accessed 15 January 2010].
[52] South West Grid for Learning, 'ICAlert Pilot Project', 2016 <http://swgfl.org.uk/magazine/ICAlert-launches-to-safeguard-schools-against-onli/PilotProjectReport.aspx> [accessed 25 July 2017].
[53] Internet Watch Foundation, 'Board Minutes 16 January 2007', 2007 <http://www.iwf.org.uk/accountability/governance/board-minutes/2007-board-minutes/16-january-2007> [accessed 15 February 2011].
[54] Christ Heal, 'ICAlert: Safeguard Schools against Online Child Abuse', *SWGfL*, 2017 <http://swgfl.org.uk/magazine/ICAlert-launches-to-safeguard-schools-against-onli> [accessed 7 August 2017].
[55] 'ICAlert Related Agreements with Police and Other Organisations - a Freedom of Information Request to South West Grid for Learning Trust', *What Do They Know?*, 2017 <https://www.whatdotheyknow.com/request/icalert_related_agreements_with?post_redirect=1> [accessed 25 July 2017].

In 2013 the Prime Minister, David Cameron, launched an initiative on internet safety and children[56] including a 'cyber-summit' focusing specifically on CAM.[57] As part of this, the government asked the IWF to begin proactively searching the web for material to take down and block and the IWF has developed that role since 2014 (with the help of significantly increased industry funding).[58]

Proactive search addresses a long-standing criticism of the IWF – that by relying on *ad hoc* reports from the public it works in a reactive and haphazard way. A proactive approach has the potential to significantly increase the amount of material taken down and the number of URLs blocked and perhaps, therefore, the overall effectiveness of the system in its expressed aim of limiting inadvertent and casual access to CAM. However this change is a significant move towards a policing role, particularly as the IWF will be sharing intelligence with law enforcement, and risks taking the IWF further away from its core functions.[59] The IWF has taken on this new role cautiously: it limits proactive searches to the public web, and has not implemented proactive searches of peer to peer services. This reflects concern that peer to peer searches would involve more intrusive surveillance aimed at individual users with significantly greater privacy implications.[60]

Image Hash List

Another result of the Cameron initiative was the creation of an IWF Image Hash List containing hashes of CAM, which serve as 'digital fingerprints' and which hosting providers can use to automatically block uploads of identical or nearly identical images.[61] This contains in excess of 200,000 unique hashes and incorporates hashes from the government Child Abuse Image Database (CAID) project, which compiles images seized by UK police.[62] It uses Microsoft PhotoDNA signatures as well as MD5 hashes, and so can be used to detect either exact matches of images or modified variants – albeit at some risk of false positives in the latter case.[63] The list is already used by Facebook, Google, Microsoft, Twitter and Yahoo to screen images, and the IWF is seeking to roll it out to other firms also.[64]

---

[56] David Cameron, 'The Internet and Pornography' (presented at the NSPCC, London, 2013) <https://www.gov.uk/government/speeches/the-internet-and-pornography-prime-minister-calls-for-action> [accessed 9 December 2013].

[57] Samuel Gibbs, 'UK's Top Tech Executives Meet for Summit against Online Child Abuse', *The Guardian*, 18 November 2013 <http://www.theguardian.com/technology/2013/nov/18/uk-top-tech-executives-online-child-abuse> [accessed 9 December 2013].

[58] Internet Watch Foundation, 'IWF Ready to Step up the Fight against Online Child Sexual Abuse Content', 2013 <http://www.iwf.org.uk/about-iwf/news/post/360-iwf-ready-to-step-up-the-fight-against-online-child-sexual-abuse-content> [accessed 2 July 2013]; LINX, 'IWF to "Proactively" Search for Illegal Content', *LINX Public Affairs*, 2013 <https://publicaffairs.linx.net/news/?p=9861> [accessed 9 December 2013].

[59] Internet Watch Foundation, 'Police' <https://www.iwf.org.uk/partnerships/police> [accessed 10 December 2013].

[60] See the discussion of this point in Macdonald, 'A Human Rights Audit of the Internet Watch Foundation', sec. 8.3.

[61] Liat Clark, 'Child Sexual Abuse Hash Lists Shared with Internet Giants', *WIRED UK*, 2015 <http://www.wired.co.uk/article/iwf-hash-lists-child-abuse-images> [accessed 15 February 2018].

[62] 'Child Abuse Image Database', *GOV.UK*, 2015 <https://www.gov.uk/government/publications/child-abuse-image-database> [accessed 25 July 2017].

[63] A Microsoft researcher has stated that the false positive rate will be approximately one in two billion images. Riva Richmond, 'Facebook's New Way to Combat Child Pornography', *New York Times*, 2011 <http://gadgetwise.blogs.nytimes.com/2011/05/19/facebook-to-combat-child-porn-using-microsofts-technology/> [accessed 20 July 2011].

[64] Internet Watch Foundation, 'Tech Breakthrough Announced on the 20th Anniversary of IWF's First Child Sexual Abuse Imagery Report', *IWF*, 2016 </news/tech-breakthrough-announced-on-20th-anniversary-of-iwfs-first-child-sexual-abuse-imagery> [accessed 7 August 2017].

The IWF has described this technology as a 'game-changer' and the description has some force: hash matching offers new opportunities for censorship at scale.[65] Hash matching systems could completely change the dynamic of regulation, enabling automated detection and takedown of CAM which is already available, and restricting new distribution by blocking images at the point of upload. The technology will allow for an image to be categorised once and once only, eliminating a repetitive aspect of the workload of IWF analysts and allowing for more time to be devoted to new images and proactive searching.

From a fundamental rights perspective, however, hash matching presents new threats to privacy. Since 2004, similar hash value systems have been used in the United States[66] in a way which goes beyond scanning publicly available materials and scans private emails and files also.[67] Microsoft, for example, scans all files on its OneDrive service[68] using PhotoDNA, Google scans all emails sent through Gmail[69], and there have been numerous cases of individuals arrested as a result.[70] This form of indiscriminate surveillance is disproportionate in itself and is also prone to function creep – being readily co-opted for other types of content. There is a risk that scanning of private communications and files is being normalised in the context of CAM as a prelude to its use in other areas.

Public law status and governance

The IWF is neither a statutory body nor publicly funded. However it came into being at the behest of the state, was reorganised in 2000 on the basis of a review[71] carried out for central government, and carries out its functions on the basis of a memorandum of understanding between the Association of Chief Police Officers (ACPO) and the Crown Prosecution Service (CPS).[72] Significantly, in that document the IWF is described by ACPO and the CPS as carrying out functions for the state:

> The IWF is… supported by UK law enforcement and CPS and *works in partnership with the Government* to provide a 'hotline' for individuals or organisations to report potentially illegal

[65] 'Hash List Could Be Game-Changer in the Global Fight against Child Sexual Abuse Images Online', *IWF*, 2015 <http://www.iwf.org.uk/news/hash-list-could-be-game-changer-global-fight-against-child-sexual-abuse-images-online> [accessed 15 February 2018].

[66] It is not clear whether hash matching is being used against the private emails and files of European users in the same way; if so, this will raise significant questions regarding the application of data protection law.

[67] See e.g. T.J. McIntyre, 'Blocking Child Pornography on the Internet: European Union Developments', *International Review of Law, Computers & Technology*, 24/3 (2010), sec. 5.6.

[68] Microsoft, 'About Our Practices and Your Data', *Microsoft & Data Law* <https://blogs.microsoft.com/datalaw/our-practices/> [accessed 15 February 2018]; Leo Kelion, 'Microsoft Alerts Police to Child Porn', *BBC News*, 2014 <http://www.bbc.com/news/technology-28682686> [accessed 7 August 2014].

[69] Rich McCormick, 'Google Scans Everyone's Email for Child Porn, and It Just Got a Man Arrested', *The Verge*, 2014 <https://www.theverge.com/2014/8/5/5970141/how-google-scans-your-gmail-for-child-porn> [accessed 15 February 2018].

[70] See e.g. Samuel Gibbs, 'Microsoft Tip Led Police to Arrest Man over Child Abuse Images', *The Guardian*, 2014 <http://www.theguardian.com/technology/2014/aug/07/microsoft-tip-police-child-abuse-images-paedophile> [accessed 15 February 2018].

[71] KPMG Peat Marwick and Denton Hall, *Review of the Internet Watch Foundation* (London, February 1999).

[72] Crown Prosecution Service and Association of Chief Police Officers, 'Memorandum of Understanding Between Crown Prosecution Service (CPS) and the Association of Chief Police Officers (ACPO) Concerning Section 46 Sexual Offences Act 2003', 2014 <https://www.iwf.org.uk/sites/default/files/inline-files/CPS%20ACPO%20S46%20MoU%202014%202.pdf> [accessed 14 February 2018].

> content and to seek out illegal content online. *It then assesses and judges the material on behalf of UK law enforcement agencies*.[73]

At an operational level we see this also in the close working arrangements between the IWF and the police. UK police forces direct complaints of online illegality in the first instance to the IWF, and in assessing those complaints IWF analysts rely on training from police and apply standards of legality which reflect the criminal law.[74] In 2010 the IWF formalised this relationship further by entering into a 'Service Level Agreement' with ACPO which 'provide[s] a protocol for the management of investigations into criminal content' hosted in the UK and to guide the interactions between the IWF and its law enforcement partners.[75]

This close integration into the work of the police means that the IWF is in many ways operating as a *de facto* public body, and to its credit it has recognised this point. In 2001, the IWF accepted that it is 'acting in a quasi-regulatory role on matters of great public interest' so that it should commit to higher standards of governance and transparency.[76] Also in 2001 the IWF board declared itself to be bound by the Human Rights Act 1998, stating that:

> The IWF accepts the principles of the European Convention on Human Rights and undertakes to be governed subject to the Human Rights Act on the basis that it should be treated as a public body.[77]

This is an important concession which opens the door to litigants who wish to challenge its decisions. As Walden puts it, '[t]he IWF, in accepting that it is subject to the Human Rights Act, is essentially telling a future court that they would be susceptible to judicial review'.[78]

That said, the concession of public body status is at best only a partial response to criticism that the IWF exercises considerable power in a way which is largely unchecked.[79] While judicial review will allow challenges to particular acts of the IWF, it still leaves the IWF outside the scope of other public law oversight mechanisms such as the Freedom of Information Act 2000.[80] As Cane has pointed out, 'judge made administrative law is only the tiny tip of a huge iceberg of norms by which the performance of public functions is framed, influenced, guided, and regulated.'[81]

Another limitation is that the concession applies only to the IWF, but the systems it has established include many other entities whose actions will not be subject to judicial review. For example, an ISP may choose to use a technology which overblocks, to provide a deceptive error message to users

---

[73] Emphasis added.
[74] Internet Watch Foundation, 'FAQs Regarding the IWF's Facilitation of the Blocking Initiative'.
[75] Internet Watch Foundation and Association of Chief Police Officers, 'Service Level Agreement between the Association of Chief Police Officers (ACPO) and the Internet Watch Foundation (IWF)', 2010 <http://www.acpo.police.uk/documents/crime/2010/201010CRIIWF01.pdf>.
[76] Roger Darlington, 'Chairing The Internet Watch Foundation', *Roger Darlington's Homepage* <http://www.rogerdarlington.co.uk/iwf.html> [accessed 21 July 2009].
[77] See Akdeniz, *Internet Child Pornography and the Law*, 264.
[78] Ian Walden, 'The Future of Freedom of Speech' (presented at the SCL 6th Annual Policy Forum: 'The New Shape of European Internet Regulation', London, 2011) <http://www.scl.org/files/scl_policy_forum_2011/The_Future_of_Freedom_of_Speech_-_Professor_Ian_Waldren.mp3> [accessed 20 March 2013].
[79] See e.g. Laidlaw, 'The Responsibilities of Free Speech Regulators', 312.
[80] Daithi Mac Sithigh, 'Datafin to Virgin Killer: Self-Regulation and Public Law', 2009 <http://ssrn.com/paper=1374846> [accessed 4 May 2009].
[81] Peter Cane, *Administrative Law*, 4th edn (Oxford, 2004), 8.

instead of a stop page, or to monitor users' private emails against the IWF Hash List. These actions by ISPs will have significant effects in their own right, but being outside the hands of the IWF would escape judicial scrutiny.

3. Terrorist material

Background

Since the internet became a mainstream technology there has been concern that it can be used to promote terrorism.[82] Initially the focus was on its use as an organisation tool for terrorist groups, but in the 2000s the dominant narrative – particularly in relation to jihadi terrorism – shifted to one of 'radicalisation' and use of the internet to indoctrinate and recruit.[83] The main concern now is that the internet can foster a decentralised type of terrorism: 'autonomous radicalisation' in which individuals or small groups are influenced by online material but otherwise act to plan and carry out an attack on their own initiative.[84]

The result has been a strong emphasis in UK and European counter-terrorist policy on internet censorship, though this has not gone unchallenged. The concept of radicalisation itself has been condemned as simplistic and reductionist.[85] Stevens and Neumann have been critical of the focus on censorship and technical solutions, describing strategies which rely on reducing the availability of content online as necessarily 'crude, expensive and counterproductive'.[86] Most fundamentally, while it is plausible that internet content may promote violence there is little evidence as to whether or how it actually does so.[87]

Whatever the merits of internet censorship as a counter-terrorism strategy, it is now firmly cemented in the UK and Europe. Following the 7 July 2005 London bombings the UK adopted an offence of 'encouragement of terrorism' which criminalises not just direct incitement to violence but also speech which might indirectly promote terrorism.[88] This fits into a wider trend by which European states have adopted laws which criminalise 'glorification', 'apology for', 'encouragement' and 'public promotion' of terrorism.[89] At a European level this has been matched in 2005 by the Council of Europe Convention on the Prevention of Terrorism which requires states to criminalise 'public provocation to commit a

---

[82] Ian O. Lesser and others, *Countering the New Terrorism* (1999), 66.

[83] Johnny Ryan, *Countering Militant Islamist Radicalisation on the Internet: A User Driven Strategy to Recover the Web* (Dublin, 2007), chap. 1.

[84] The term lone wolf is often used but is misleading as in many cases attackers have some contact with or support from the terrorist group. See e.g. Raffaello Pantucci, *A Typology of Lone Wolves: Preliminary Analysis of Lone Islamist Terrorists* (London, 2011) <http://www.academia.edu/download/24801864/1302002992icsrpaper_atypologyoflonewolves_pantucci.pdf> [accessed 7 March 2017].

[85] Arun Kundnani, 'Radicalisation: The Journey of a Concept', in *Counter-Radicalisation: Critical Perspectives*, ed. by Christopher Baker-Beall, Charlotte Heath-Kelly, and Lee Jarvis (2014).

[86] Tim Stevens and Peter R. Neumann, *Countering Online Radicalisation: A Strategy for Action* (London, 2009), 1 <http://www.icsr.info/news/attachments/1236768445ICSROnlineRadicalisationReport.pdf> [accessed 11 March 2009].

[87] Maura Conway, 'Determining the Role of the Internet in Violent Extremism and Terrorism: Six Suggestions for Progressing Research', *Studies in Conflict & Terrorism*, 40/1 (2016), 78.

[88] Terrorism Act 2006, ss. 1 and 2 cover publication and dissemination of statements likely to be understood as encouraging terrorism.

[89] David Banisar, *Speaking of Terror* (Strasbourg, 2008), 20 <http://www.coe.int/t/dghl/standardsetting/media/Doc/SpeakingOfTerror_en.pdf> [accessed 12 May 2009].

terrorist offence'[90] and most recently by the Directive on Combating Terrorism.[91] The Directive, agreed in March 2017, requires Member States to criminalise 'public provocation' including 'glorification of terrorist acts'[92] and 'to ensure the prompt removal of online content' which amounts to public provocation.[93] It also permits Member States to 'take measures to block access to such content' where removal is not feasible.[94]

These glorification offences have been controversial due to their vague terms, remoteness from any completed offence, and fears that they will chill legitimate discussion.[95] There is an extensive literature assessing the extent to which they are compatible with freedom of expression.[96] In this section, however, we focus not on the content of these offences but rather the procedures by which they are policed in the UK. How, in practice, is this type of speech suppressed and what implications does this have?

Takedown notices under section 3 of the Terrorism Act 2006 and voluntary takedowns

Section 3 of the Terrorism Act 2006 provides a mechanism for police to serve notices on service providers that material is 'unlawfully terrorism-related' and requiring that the material be removed. Failure to do so means that the service provider will be treated as having endorsed it, giving rise to possible criminal liability.

The section was criticised at the time as effectively delegating decisions on legality to police, with no judicial oversight, but was defended by the government on the basis that this was necessary to avoid delay. Despite that claim, it has never been used.[97] Instead, it has been deliberately kept in reserve as a fall-back in case 'negotiations with industry falter'.[98] Guidance from the Home Office requires police to use voluntary arrangements where possible, on the basis that these give:

> greater flexibility to discuss how to ensure the material is removed, how further publications can be prevented, whether there is other similar material, whether evidence identifying those involved in publishing the material can be obtained, whether the accounts responsible for the publications can be terminated, timescales, etc.[99]

This emphasis on voluntary cooperation is significant. As with child abuse images, a reliance on informal cooperation with industry is seen by the state as more effective and lending itself to

---

[90] Article 5.
[91] Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA.
[92] Article 5.
[93] Article 21(1).
[94] Article 21(2).
[95] Ben Emmerson, *Report of the Special Rapporteur on the Promotion and Protection of Human Rights and Fundamental Freedoms While Countering Terrorism, A/HRC/31/65* (New York, NY, 22 February 2016), 14–16.
[96] See e.g. S. Chehani Ekaratne, 'Redundant Restriction: The UK's Offense of Glorifying Terrorism', *Harvard Human Rights Journal*, 23 (2010), 205.
[97] Home Office, 'Memorandum to the Home Affairs Committee: Post-Legislative Scrutiny of the Terrorism Act 2006' (2011), 6–7 <http://www.official-documents.gov.uk/document/cm81/8186/8186.pdf> [accessed 20 May 2013].
[98] Home Office, 'Memorandum to the Home Affairs Committee: Post-Legislative Scrutiny of the Terrorism Act 2006', 6–7.
[99] Home Office, 'Guidance on Notices Issued under Section 3 of the Terrorism Act 2006', *WhatDoTheyKnow*, 2010, 6 <http://www.whatdotheyknow.com/request/implementation_of_terrorism_act> [accessed 30 July 2011].

outcomes such as preventing republication and terminating user accounts which are not provided for by law.

The takedown function is administered by the specialist police Counter Terrorism Internet Referral Unit (CTIRU), established in 2010 by ACPO.[100] The CTIRU has followed the model of the IWF in relation to child abuse images by encouraging individuals to make anonymous reports about online material. The CTIRU prioritises English language websites, and has massively increased the numbers of takedown requests it issues: from 1,527 in 2010 and 2011 to 121,151 in 2016 – over 2,000 per week.[101]

This is an opaque process, but a recent glimpse behind the curtain suggests that it is of variable quality: in evidence before the House of Commons Home Affairs Committee Google has said that the accuracy rate of requests from the CTIRU in relation to YouTube is only 80%.[102] It is remarkable that one in every five takedown requests from a specialist police unit is rejected on the basis that the video in question is neither illegal nor contrary to YouTube's own 'community standards'. The implications are unappealing: either the CTIRU and/or YouTube are at fault in making inconsistent decisions, the assessment of this material inherently involves an undesirable level of subjectivity, or perhaps all of the above.

The CTIRU has been influential in a European context, and in July 2015 Europol launched the European Union Internet Referral Unit (EU IRU) which, as the name suggests, was largely modelled on the CTIRU. The revised Europol Regulation provides for this by conferring power on Europol to refer internet content 'to the online service providers concerned for their voluntary consideration of the compatibility of the referred internet content with their own terms and conditions'.[103] As with the CTIRU, the voluntary nature of the scheme is being relied upon to insulate it from a requirement for procedural safeguards or judicial oversight. Europol has asserted that:

> Referral activities will not constitute an enforceable act, thus the decision and related implementation of the referral is taken under full responsibility and accountability of the concerned service provider… The overall process of the EU IRU reporting terrorist and extremist online content to the online service provider is no different from any citizen flagging content for removal by the respective online service provider.[104]

Automating takedown and preventing uploads

While European, as distinct from national, measures are beyond the scope of this chapter, it is worth noting that the UK has been active in promoting the use of hash value matching and other automated

---

[100] ACPO, 'CTIRU Factsheet'.
[101] '250,000th Piece of Online Extremist/Terrorist Material to Be Removed', *Metropolitan Police*, 2016 <http://news.met.police.uk/news/250000th-piece-of-online-extremist-slash-terrorist-material-to-be-removed-208698> [accessed 15 March 2017].
[102] House of Commons Home Affairs Committee, 'Oral Evidence: Hate Crime and Its Violent Consequences', 2017 <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/home-affairs-committee/hate-crime-and-its-violent-consequences/oral/48836.pdf>.
[103] Regulation (EU) 2016/794 of the European Parliament and of the Council of 11 May 2016 on the European Union Agency for Law Enforcement Cooperation (Europol) and replacing and repealing Council Decisions 2009/371/JHA, 2009/934/JHA, 2009/935/JHA, 2009/936/JHA and 2009/968/JHA, Article 4(1)(m). Emphasis added.
[104] Chris Jones, 'Policing the Internet: From Terrorism and Extremism to "Content Used by Traffickers to Attract Migrants and Refugees"', *Statewatch*, 2016, 2 <http://www.statewatch.org/analyses/no-290-policing-internet.pdf> [accessed 20 April 2016].

censorship tools through the EU Internet Forum, drawing on the experience of the IWF in relation to CAM. The Minister for Internet Safety & Security, Baroness Shields, describes the aim as being:

> to improve solutions that classify the language of extremism, automate the identification and removal of dangerous content at scale, and create tools that better tackle automated bots and other techniques that support these propaganda machines.[105]

The main result of this initiative has been a December 2016 agreement to establish an industry database of hash values for images and videos, with a view to pre-emptively blocking uploads.[106] Facebook, Microsoft, Twitter, and YouTube are the initial partners, but the intention is to make the database available to other firms also. Rather than assessing legality, the firms will add material to the database on the basis that they are 'hashes of the most extreme and egregious terrorist images and videos we have removed from our services — content most likely to violate all of our respective companies' content policies'.[107] Regarding procedures, the firms have stated that:

> each company will independently determine what image and video hashes to contribute to the shared database. No personally identifiable information will be shared, and matching content will not be automatically removed. Each company will continue to apply its own policies and definitions of terrorist content when deciding whether to remove content when a match to a shared hash is found. And each company will continue to apply its practice of transparency and review for any government requests, as well as retain its own appeal process for removal decisions and grievances.[108]

This leaves a number of questions unanswered. Will the database be used to scan private messages? Could it be used to detect and report users to law enforcement if they email or upload particular material? Both of these developments have already taken place in the context of CAM, where blocking and hash matching have both rapidly become prosecution tools, and it would not be surprising if this database were to evolve in the same way.

Indeed, the database may already be used for these purposes. A June 2017 story in the *Guardian* indicates that Facebook has established a 'counter-terrorism unit' which has the power to 'carry out investigations into user accounts if they are suspect of having links to terrorist groups'.[109] Moderators are given '[f]ull account access… to any profile once it has been flagged by algorithms looking for certain names, images or hashtags associated with terrorist groups' and may then 'access the individual's private messages, see who they are talking to and what they are saying, and view where they have been'. In a remarkable passage, the story describes how this results in users routinely being referred to police:

---

[105] Joanna Shields, 'Countering Online Radicalisation and Extremism' (George Washington University Centre for Cyber and Homeland Security, 2017) <https://www.gov.uk/government/speeches/countering-online-radicalisation-and-extremism-baroness-shields-speech> [accessed 6 August 2017].

[106] European Commission, 'EU Internet Forum: A Major Step Forward in Curbing Terrorist Content on the Internet', 2016 <http://europa.eu/rapid/press-release_IP-16-4328_en.htm> [accessed 14 March 2017].

[107] Twitter, 'Partnering to Help Curb the Spread of Terrorist Content Online', *Twitter Blogs*, 2016 <https://blog.twitter.com/2016/partnering-to-help-curb-the-spread-of-terrorist-content-online> [accessed 5 December 2016].

[108] Twitter, 'Partnering to Help Curb the Spread of Terrorist Content Online'.

[109] Olivia Solon, 'Counter-Terrorism Was Never Meant to Be Silicon Valley's Job. Is That Why It's Failing?', *The Guardian*, 29 June 2017, section Technology <https://www.theguardian.com/technology/2017/jun/29/silicon-valley-counter-terrorism-facebook-twitter-youtube-google?CMP=share_btn_tw> [accessed 2 July 2017].

> The team's highest priority is to identify 'traveling fighters' for Isis and Al-Qaida. Someone would be categorized as such if their profile has content that's sympathetic to extremism and if they had, for example, visited Raqqa in Syria before traveling back to Europe. When a traveling fighter is identified – which according to one insider takes place at least once a day – the account is escalated to an internal Facebook team that decides whether to pass information to law enforcement.

It would take another chapter to discuss all the issues presented by this form of privatised policing (not least the data protection issues involved in profiling users and routinely accessing private messages in this way) but the story is significant for the way in which it again highlights the intersection between censorship and surveillance. As the CJEU recognised in *Scarlet (Extended) v. SABAM*[110]*,* filtering systems which scan communications inherently threaten the fundamental rights to privacy and data protection as well as freedom of expression.

Blocking

Enlisting the IWF?

In addition to takedown, the UK government has long sought ISP level blocking of terrorist webpages. In 2008 and again in 2011 the Home Office floated the idea that terrorist material should be incorporated into the IWF list so that ISPs would block access in the same way as they do for CAM using the Cleanfeed system.[111] The 2015 *Counter-Extremism Strategy* again addressed ISP blocking, but taking a different tone – talking about 'learning from' the IWF rather than directly using its structures.[112] This change in emphasis is significant. The IWF has resisted expanding its remit to new categories of material, largely due to fears of reputational damage and increased costs which might undermine its child protection role.[113] It seems that the government has, for the time being, abandoned any intention to enlist the IWF in relation to terrorist material.

CTIRU blacklist

In a parallel programme, however, the UK government has proceeded with its own URL List system, compiling a blacklist of terrorist-related URLs which it makes available under direct agreements with companies which provide filtering software.[114] From November 2008 to February 2011 the Labour government ran this as a pilot project; following a pause for evaluation it restarted in July 2011 and has been in place since then.

---

[110] Judgment of 24 November 2011, *Scarlet (Extended) v SABAM,* C-70/10, EU:C:2011:771.
[111] Helene Mulholland, 'Government to Stamp down on Terror "grooming" Websites', *The Guardian*, 17 January 2008 <http://www.guardian.co.uk/politics/2008/jan/17/uksecurity.terrorism> [accessed 24 November 2008]; Chris Williams, 'Jacqui's Jihad on Web Extremism Flops', *The Register*, 2009 <http://www.theregister.co.uk/2009/02/13/jacqui_smith_web_extremism/> [accessed 13 February 2009]; Home Office, *Prevent Strategy* (London, 2011), 79; Jane Fae, 'Gov and ISPs Clash over Informal Policing of Net', *The Register*, 2011 <http://www.theregister.co.uk/2011/03/16/self_regulation/> [accessed 18 March 2011].
[112] Home Office, *Counter-Extremism Strategy* (London, 2015), 25.
[113] For example, in relation to extreme pornography and non-photographic images see Internet Watch Foundation, 'Board Minutes 25 November 2008', *Internet Watch Foundation*, 2008 <http://iwf.org.uk/corporate/page.200.htm> [accessed 25 January 2009]; Internet Watch Foundation, 'Board Minutes 29 September 2009'.
[114] See generally Fae, 'The Internet Censorship Programme You're Not Allowed to Know About'.

There is little transparency about the system, but if we piece together material in the public domain and responses to several Freedom of Information Act requests[115] it appears to operate as follows:

- The CTIRU identifies URLs relating to material which breaches the Terrorism Acts 2000 and 2006 but which is hosted overseas and cannot be taken down.
- That material is (in some cases at least) considered for legality by specialist prosecutors in the Crown Prosecution Service (CPS).[116]
- The URLs assessed as relating to illegal material are passed under a confidentiality agreement to filtering companies who restrict access to those URLs in their software.[117]
- There is no notification of websites, opportunity to make representations about the blacklisting or after the fact, nor appeal mechanism against a decision to blacklist.
- There is no information as to whether or how the list is checked periodically to determine whether URLs should still be blocked.
- Individuals who attempt to view a blacklisted URL will see a stop screen indicating that it has been blocked, but not that the CTIRU was responsible for blacklisting it.
- As of 2013, approximately 1,000 URLs were on the list.[118]

The blacklist was rolled out initially to companies who supply filtering products across 'the public estate' – including locations such as schools, colleges and libraries. In December 2013 the Report of the Prime Minister's Task Force on Extremism indicated an intention to use the blacklist more widely, saying that the government would 'work with internet companies to restrict access to terrorist material online which is hosted overseas but illegal under UK law'.[119]

Following pressure from the Prime Minister, in November 2014 the major UK ISPs (BT, TalkTalk, Virgin Media and Sky) agreed to adopt the blacklist.[120] Unlike the IWF URL List, however, the blacklist is currently only used in their 'family friendly' internet connections – it is not used for connections where the subscriber has chosen to disable parental filters. Nevertheless, this still covers a significant portion of UK internet users: depending on the ISP between 6% of subscribers (BT) and 30-40% of subscribers (Sky) have these network level filters active.[121]

---

[115] 'Filtering of Terrorist Material - a Freedom of Information Request to Home Office', *WhatDoTheyKnow*, 2010 <https://www.whatdotheyknow.com/request/filtering_of_terrorist_material> [accessed 5 April 2013]; 'Filtering of Terrorist Material - a Freedom of Information Request to Home Office', *WhatDoTheyKnow*, 2011 <http://www.whatdotheyknow.com/request/filtering_of_terrorist_material_2> [accessed 16 October 2011]; 'Current Status of Terrorist Internet Filtering - a Freedom of Information Request to Home Office', *WhatDoTheyKnow*, 2013 <https://www.whatdotheyknow.com/request/current_status_of_terrorist_inte> [accessed 30 March 2016].

[116] 'CPS Role in Internet Filtering of Terrorist Material - a Freedom of Information Request to Crown Prosecution Service', *WhatDoTheyKnow*, 2013 <https://www.whatdotheyknow.com/request/cps_role_in_internet_filtering_o> [accessed 30 March 2016].

[117] 'Current Status of Terrorist Internet Filtering - a Freedom of Information Request to Home Office'.

[118] HC Deb 19 June 2013, vol. 564, col 687W.

[119] Prime Minister's Task Force on Tackling Radicalisation and Extremism, *Tackling Extremism in the UK* (London, 2013), 3 <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263181/ETF_FINAL.pdf> [accessed 6 December 2013].

[120] Alex Hern, 'ISPs Criticised over Deal to Filter Extremist Material Online', *The Guardian*, 14 November 2014, section Technology <http://www.theguardian.com/technology/2014/nov/14/isps-filter-extremist-material-internet> [accessed 18 July 2017].

[121] Ofcom, 'Ofcom Report on Internet Safety Measures: Strategies of Parental Protection for Children Online', 25–28.

The Home Office has stated that there is no statutory underpinning for the blacklist, saying that it is done 'on a voluntary basis'.[122] The only formal basis appears to be contracts with the firms which receive the blacklist, which require them to keep the list confidential.[123] The Home Office has refused to release copies of these contracts, and in particular has refused to confirm how the contracts deal with liability for wrongful blocking – though the context suggests that the Home Office has agreed to indemnify the firms against any liability.[124]

The criteria and procedures used for blocking are set out by internal guidelines. These are not publicly available, but have been partially disclosed in response to a Freedom of Information Act request.[125] The test applied is whether the publication of material breaches section 57 or 58 of the Terrorism Act 2000 (information useful in commission of act of terrorism; e.g. bomb-making manuals) or sections 1 or 2 of the Terrorism Act 2006 (e.g. encouragement to commit acts of terrorism and glorification). Guidance for the decision makers includes the question 'Does the context of material demonstrate that it is aimed primarily at an extremist audience?', explaining that '[f]or instance, a video on a mainstream media web site is likely to be aimed at a different audience to one in an extremist chat forum'.

This guidance illustrates that the decision to block involves a complex assessment of whether material is 'likely to be understood' as encouraging terrorism, including the context in which it is published and its target audience – the same video may be blocked if it appears on an extremist chat forum but not if it appears on a mainstream news site – making it particularly inappropriate that the decision should be carried out by the executive in secret, without any notification or appeal mechanism.

The decision to blacklist has no direct legal effect. However, it is significant that the filtering companies do not exercise any discretion in deciding whether to block the URLs provided by the CTIRU: unlike the referral of videos to YouTube, for example, there is no independent assessment of compliance with terms of use to distance the censorship from the state. In this context, an official and effectively binding determination that particular material amounts to a criminal offence certainly appears to be a direct 'interference by public authority' with the right to freedom of expression so as to ground an action under section 7 of the Human Rights Act 1998.

This remedy is, however, largely theoretical given that the individual will have no notification that the Home Office has blacklisted a particular URL, and the filtering companies are contractually precluded from revealing this fact. In these circumstances, there is an argument to be made that the secret nature of the system also means that there is no 'adequate and effective remedy' as required by Article 13 ECHR.

4. Pornography and extreme pornography under the Digital Economy Act 2017

Background

Part 3 of the Digital Economy Act 2017 creates a general obligation to introduce age controls on online pornography as well as wide-ranging statutory powers to block pornographic material including – but

[122] 'Filtering of Terrorist Material - a Freedom of Information Request to Home Office'.
[123] 'Filtering of Terrorist Material - a Freedom of Information Request to Home Office'.
[124] 'Filtering of Terrorist Material - a Freedom of Information Request to Home Office'.
[125] 'Filtering of Terrorist Material - a Freedom of Information Request to Home Office'.

not limited to – extreme pornography.[126] As such it is a significant departure from the self-regulatory models used to date and also a break from Western European norms in providing for extensive blocking of legal material.[127] By setting out to regulate the behaviour of websites based outside the UK it is also an ambitious attempt to compel firms to abide by national law even though they may have no connection with the UK.

The background to the 2017 Act is a longstanding public concern (many have said moral panic[128]) about the availability of pornography to children. This has developed against a wider background where puritanical views of the right have been matched by a swing on the left from libertine to censorious, leaving pornography with few friends on either side of the UK political spectrum.[129] The result has been cross party consensus on a need for greater controls, initially expressed by pressuring ISPs and wifi providers to adopt family friendly filters, followed by an aggressive expansion of the implementation of the AVMS Directive[130], and culminating in the 2017 Act which implements the 2015 Conservative Party manifesto pledge: '[W]e will stop children's exposure to harmful sexualised content online, by requiring age verification for access to all sites containing pornographic material'.[131]

Age verification

The main aim of Part 3 of the 2017 Act is to reduce availability of pornography to children by requiring commercial sites to implement age verification. It applies to any person, anywhere in the world, regardless of whether they target the UK, establishing a general rule that they shall not 'make pornographic material available on the internet to persons in the United Kingdom on a commercial basis' unless they ensure that 'the material is not normally accessible by persons under the age of 18'.[132] The restriction is not limited to websites: it prohibits any form of 'making available on the internet' which appears to include apps and peer to peer services also.[133] 'Pornographic material' is defined in some detail, but broadly speaking means material 'produced solely or principally for the purposes of sexual arousal'.[134]

Enforcement

Enforcement is the responsibility of an age-verification regulator to be designated by the Secretary of State and confirmed by both Houses of Parliament.[135] In February 2018 the Department for Culture,

[126] Though age verification was previously in place for the narrow category of TV-like material regulated by the Audio-visual Media Services Directive, and statutory blocking powers in the context of intellectual property infringement.
[127] Putting the UK in the company of the Russian Federation and Turkey: Swiss Institute of Comparative Law, *Comparative Study on Blocking, Filtering and Take-Down of Illegal Internet Content*, 14.
[128] Jim Greer, 'Children and Internet Pornography: A Moral Panic, a Salvation for Censors and Trojan Horse for Government Colonisation of the Digital Frontier', in *Revisiting Moral Panics*, ed. by Viviene E. Cree, Gary Clapton, and Mark Smith (Bristol, 2015).
[129] Jerry Barnett, *Porn Panic! Sex and Censorship in the UK* (2016), chap. 5.
[130] Petley, 'The Regulation of Pornography on Video-on-Demand in the United Kingdom', 276–300.
[131] 'The Conservative Party Manifesto 2015', 2015, 35.
[132] Section 14(1).
[133] Neil Brown, 'Age Verification and Online Pornography under the DEA 2017: Whose Fine Is It Anyway?', 2017 <https://www.scl.org/articles/3920-age-verification-and-online-pornography-under-the-dea-2017-whose-fine-is-it-anyway> [accessed 3 July 2017].
[134] See section 15.
[135] Digital Economy Act 2017, sections 16 and 17.

Media and Sport designated the BBFC – a private company, albeit one that is already embedded in a number of co-regulatory schemes for UK media regulation.[136]

The age-verification regulator can impose financial penalties (up to £250,000 or five per cent of 'qualifying turnover') or seek an injunction against a person who fails to comply with the age verification requirement.[137] The regulator also has two powers to indirectly enforce age-verification, designed with overseas websites in mind.

Blocking

First, the regulator is empowered to block sites who do not age-verify, by giving a notice to an ISP requiring them to 'prevent persons in the United Kingdom from being able to access the offending material'.[138] The ISP is then obliged to take either 'steps specified in the notice' or 'arrangements that appear to the [ISP] to be appropriate' to implement the blocking – a formula which gives the regulator power to specify the blocking technology to be used.

The section explicitly permits over-blocking by specifying that the blocking measures may 'have the effect of preventing persons in the United Kingdom from being able to access material other than the offending material'.[139] The implication is that the regulator does not have to block at the level of the individual URL but may block, for example, all content hosted on a particular site. This will be significant given the move towards websites offering secure (HTTPS) connections: blocking of individual pages becomes impossible where a website is served in this way, leaving blocking of the full site as the all or nothing option.[140]

There is no requirement of proportionality regarding either the decision to block or the manner of its implementation (such as the collateral damage it might cause, or the cost it might impose on the ISP). The regulator may, but is not required to, specify that a blocking notice (stop page) be used to tell individuals why the material is not available.[141]

Notices to payment and ancillary service providers

The second indirect enforcement power permits the regulator to issue a notice to payment service providers (such as credit card processors) and ancillary service providers (others providing services to the site, including advertisers and, crucially, platforms such as Twitter[142]), specifying that a person has failed to comply with the age verification duty.[143] The Act does not require these providers to take any

[136] Department for Digital, Culture, Media & Sport, 'Notice of Designation of Age-Verification Regulator', 2018 <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/683354/BBFC_Designation_Notice.pdf> [accessed 16 March 2018].
[137] Digital Economy Act 2017, section 19.
[138] Section 23.
[139] Section 23(3).
[140] *Internet Society Perspectives on Internet Content Blocking* (2017), 15 <https://www.internetsociety.org/resources/doc/2017/internet-content-blocking/> [accessed 15 February 2018].
[141] Section 23(4).
[142] The government position was set out in the Lords as being that 'services, including Twitter, can be classified by regulators as ancillary service providers where they are enabling or facilitating the making available of pornographic or prohibited material'. See Brown, 'Age Verification and Online Pornography under the DEA 2017'.
[143] Section 21.

specific action on foot of such a notice[144]; but the expectation is that they will choose to cut-off dealings with sites which are 'named and shamed'[145] – in effect, a form of state-directed boycott, which is intended to have extraterritorial effect by targeting firms such as Visa, MasterCard and PayPal. Again, there is no requirement that this be a proportionate response.

Procedures

The procedures to be followed by the regulator differ as between the fining, blocking and notice of non-compliance powers, without any obvious justification. In the case of fines the regulator must give the person to be fined prior notice and an opportunity to be heard.[146] In the case of blocking, the regulator must give the non-complying person prior notice but, surprisingly, is not required to give an opportunity to be heard.[147] For a notice of non-compliance to payment service providers and ancillary service providers, only notification after the fact is required[148] – remarkably so, given the reputational and business damage such a notice is intended to cause. The regulator must make arrangements for appeals against these enforcement decisions, by a person who is 'sufficiently independent of' the regulator.[149]

Each of these enforcement powers is discretionary. Section 26(1) provides that the regulator may choose to target 'persons who make pornographic material… available on the internet on a commercial basis *to a large number of persons* [or] *generate a large amount of turnover* by doing so'.[150] This gives considerable freedom as to how enforcement will take place, and the BBFC has indicated that it plans to aim at a relatively small number of high profile sites, those most visited by children, and then 'moving down the list'.[151]

Significantly (in light of *Scarlet (Extended) v. SABAM*[152] and the potential cost of implementing blocking) the ISP as well as the non-complying person may appeal against a blocking notice.[153] However there is no provision for appeal by individual web users, notwithstanding that their Article 10 rights will also be implicated by a decision to block.[154]

Privacy

Privacy was one of the most contentious issues during the passage of the 2017 Act, with fears that age verification requirements would mean 'databases of the UK's porn habits' as well as collections of

---

[144] Though it may be that a notice might put a platform or hosting provider at risk of accessorial liability if they failed to act, on the basis that they would have actual knowledge of publication of obscene material or extreme pornography to persons within the UK. See e.g. Sam Frances, 'Digital Economy Bill Contains Mandatory Age-Verification, but Avoids Blocking', *LINX Public Affairs*, 2016 <https://www.linx.net/public-affairs/digital-economy-bill-contains-mandatory-age-verification-but-avoids-blockin> [accessed 22 July 2017]; Brown, 'Age Verification and Online Pornography under the DEA 2017'.

[145] Burkhard Schafer, 'No Data Protection Please, We Are British—Privacy, Porn and Prurience in the Digital Economy Bill', *Datenschutz Und Datensicherheit-DuD*, 41/6 (2017), 357.

[146] Section 19(3).

[147] Section 23(10).

[148] Section 21(3).

[149] Section 16(5).

[150] Emphasis added.

[151] David Austin, chief executive of the BBFC, in parliamentary testimony on 11 October 2016, discussed in Brown, 'Age Verification and Online Pornography under the DEA 2017'.

[152] Judgment of 24 November 2011, *Scarlet (Extended) v SABAM,* C-70/10, EU:C:2011:771.

[153] Section 16(6).

[154] See in particular *Cengiz and others v Turkey*, Application nos. 48226/10 and 14027/11, judgment of 1 December 2015.

identity documents open to misuse.[155] Despite this, Part 3 of the Act does not contain the terms 'privacy' or 'data protection'. The age verification regulator is obliged to publish guidance as to the types of age verification systems it will treat as complying with the Act – but there is no provision that the regulator have regard to the need to protect privacy in doing so and (despite some obfuscation on this point from government) the regulator has no power to insist on any privacy requirements. The scheme of Part 3 will force the regulator to recognise systems such as credit card authentication as effective notwithstanding that the systems may collect details of individuals' identity as well as their age.[156]

The response of the government on the privacy issue has been to leave the issue to the Information Commissioner's Office (ICO) as a matter of general data protection law. The Department for Culture, Media & Sport has published draft guidance to the regulator which provides that the regulator's guidance should include information about the data protection obligations of pornography providers and age verification services – but does not provide for any concrete requirements on providers. Instead, the draft guidance includes a bare statement that the regulator 'should be satisfied that due consideration has been given to data protection legislation' and should 'inform the ICO where concerns arise'.[157]

The privacy risks are significant, particularly in light of recent large-scale data breaches. Data on individuals' sexual preferences is sensitive by definition, and especially so in the case of sexual minorities, but this scheme is likely to force users to link their identity to that information. In this, it highlights the recurring theme that internet censorship schemes, which often requires monitoring of users' activity, must be assessed for their impact on privacy as well as freedom of expression.

Extreme pornography

The 2017 Act also provides for the age verification regulator to take enforcement action against sites containing 'extreme pornography', which is a relatively recent concept in the UK.[158] Until 2008, violent pornography was dealt with by the law in the same way as other forms of pornography (excluding CAM), with publication, but not simple possession, of 'obscene articles' prohibited by the Obscene Publications Act 1959.[159] The expansion of the law can be traced to the 2003 strangulation of Jane Longhurst by a man obsessed with violent pornography, including images of sexual asphyxiation and

---

[155] See e.g. Jamie Grierson, 'Privacy Campaigners Criticise UK Plan for Age Checks on Porn Websites', *The Guardian*, 17 July 2017 <http://www.theguardian.com/technology/2017/jul/17/age-checks-introduced-porn-websites-uk-credit-card-details> [accessed 4 August 2017].

[156] Jim Killock, 'A Database of the UK's Porn Habits. What Could Possibly Go Wrong?', *Open Rights Group*, 2016 <https://www.openrightsgroup.org/blog/2016/a-database-of-the-uks-porn-habits-what-could-possibly-go-wrong> [accessed 5 August 2017]; Jim Killock, 'Fig Leafs for Privacy in Age Verification', *Open Rights Group*, 2016 <https://www.openrightsgroup.org/blog/2016/fig-leafs-for-privacy-in-age-verification> [accessed 5 August 2017].

[157] Department for Digital, Culture, Media & Sport, 'Draft Guidance to the Age Verification Regulator', 2017 <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/600733/Draft_Guidance_to_the_Age_Verification_Regulator_March_2017__1_.pdf> [accessed 5 August 2017].

[158] See generally Alexandros K. Antoniou and Dimitris Akrivos, *The Rise of Extreme Porn: Legal and Criminological Perspectives on Extreme Pornography in England and Wales* (London, 2017).

[159] Section 2(1) of the Obscene Publications Act 1959, as amended by the Obscene Publications Act 1964. Section 1(1) defines 'obscene' in the following terms: 'an article shall be deemed to be obscene if its effect... is taken as a whole, such as to tend to deprave or corrupt persons who are likely, having regard to all relevant circumstances, to read, see or hear the matter contained or embodied in it'.

rape.[160] Following her death, a campaign to criminalise the possession of violent pornography received considerable media and political support, resulting in the Labour government criminalising possession of 'extreme pornographic images' (such as depictions of bestiality, necrophilia, and acts which threaten serious harm or death) in 2008.[161] In 2015 the Conservative/Liberal Democrat coalition widened the scope of the offence to include images of rape also.[162]

While the merits of the extreme pornography offence are beyond the scope of this chapter, it should be noted that it has been extremely controversial from the outset, with many arguing that it is overbroad and largely motivated by unfounded views of a causal relationship between fantasy and sexual violence.[163] Murray points out that it criminalises material featuring consenting adults, making it a crime to video something which it is legal to do and creating a risk that 'police will use [the offence] as a Trojan Horse to regulate the underlying activity'[164] while Jackman suggests that it has been abused as a 'consolation prize' offence where a defendant has their phone or computer seized on an unrelated matter in respect of which they are acquitted or face no charges.[165]

Powers of the age-verification regulator over extreme pornography

As already mentioned, the focus of Part 3 of the Digital Economy Act 2017 is to prevent children from accessing pornography. The provisions in relation to extreme pornography were an afterthought: as originally introduced, the Bill did not address extreme pornography at all. The legislative history is complex, but these provisions are watered down from an earlier government amendment to the Bill which aimed to block any 'prohibited material' – i.e. pornographic material which would be refused classification by the BBFC.[166] Following opposition in the House of Lords, the government replaced that provision with a narrower one limited to extreme pornography, acknowledging that 'as this measure is about protecting children, we do not want to create a new threshold for what adults can or cannot see'.[167]

The 2017 Act therefore creates additional censorship powers in respect of 'extreme pornographic material'.[168] These piggyback on the age-verification enforcement powers, and in respect of extreme pornography the regulator has the same powers to require ISPs to block, and to issue notices of non-compliance to payment service providers and ancillary service providers, subject to the same notice

[160] Andrew Murray, 'The Reclassification of Extreme Pornographic Images', *Modern Law Review*, 72/1 (2008), 73–90.
[161] Criminal Justice and Immigration Act 2008, section 63.
[162] Criminal Justice and Courts Act 2015, section 37.
[163] See in particular Nick Cowen, *Nothing to Hide: The Case against the Ban on Extreme Pornography* (2016) <https://www.adamsmith.org/s/Nicholas-Cowen-Nothing-to-hide-FINAL.pdf> [accessed 22 February 2017].
[164] Murray, 'The Reclassification of Extreme Pornographic Images', 90.
[165] Cowen, *Nothing to Hide: The Case against the Ban on Extreme Pornography*, 2. Contrary to some claims that the offence would be largely symbolic, prosecutions for extreme pornography are increasingly common and in most cases will be based on possession of material downloaded from the internet. In 2009-10 (the first full year of the extreme pornography offence being in force) 270 prosecutions were brought, rising to 1,740 in 2015-16: Crown Prosecution Service, 'Violence against Women and Girls Crime Report 2015-16', 2016, 89 <http://www.cps.gov.uk/publications/docs/cps_vawg_report_2016.pdf> [accessed 20 July 2017].
[166] This would have included material depicting acts such as face sitting, watersports, and fisting. See e.g. Damien Gayle, 'UK to Censor Online Videos of "non-Conventional" Sex Acts', *The Guardian*, 23 November 2016, section Technology <https://www.theguardian.com/technology/2016/nov/23/censor-non-conventional-sex-acts-online-internet-pornography?CMP=twt_a-technology_b-gdntech> [accessed 23 November 2016].
[167] Lord Ashton of Hyde, HL Deb 20 March 2017, vol 782.
[168] Defined in section 22 by reference to the offence in section 63(7) or (7A) of the Criminal Justice and Immigration Act 2008.

and appeal mechanisms. Significantly, these powers apply against any person 'making extreme pornographic material available on the internet to persons in the United Kingdom', whether or not they are doing so on a commercial basis – something which greatly increases the possible scope of blocking. As with the age-verification powers there is no requirement of proportionality in enforcement – on the face of it, the Act would permit an entire domain to be blocked if necessary to prevent access to a single image – raising the same questions as to whether these powers are compatible with Article 10 ECHR.

Effectiveness

Even on its own terms Part 3 seems likely to have little effect in preventing access to pornography. It does nothing, for example, to prevent access to non-commercial material, and the BBFC have already indicated that they will be targeting only a small number of high profile sites. Ample alternatives will be available. Quite apart from those structural problems, at best blocking is likely to be only a minor obstacle to users who have become increasingly familiar with circumvention tactics such as virtual private networks (VPNs) and the use of alternative DNS providers. In 2011 Ofcom concluded (in the context of file-sharing) that '[f]or all blocking methods circumvention by site operators and internet users is technically possible and would be relatively straightforward by determined users'.[169] As Schafer points out, it is difficult to see why blocking should be expected to work any better in this context.[170]

Part 2: IDENTIFYING EUROPEAN LEGAL STANDARDS AND ANALYSING THE CASE STUDIES

1. Fundamental rights scrutiny

To what extent do our three case studies meet fundamental rights standards? In this section we summarise the main requirements for internet censorship under the ECHR and identify some areas where each system appears to fall short.

a. Article 10 ECHR

'Prescribed by law'

Article 10 ECHR requires that interferences with freedom of expression be prescribed by law. The best-known treatment of this concept was given in *Sunday Times v. United Kingdom*[171] where the ECtHR held that in addition to requiring a legal basis it also imposes requirements regarding the quality of the law. First, 'the law must be adequately accessible: the citizen must be able to have an indication that is adequate in the circumstances of the legal rules applicable to a given case'. Secondly, 'a norm cannot be regarded as a 'law' unless it is formulated with sufficient precision to enable the citizen to regulate his conduct: he must be able – if need be with appropriate advice – to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail'.[172]

The *Sunday Times* approach has been supplemented in *Ekin Association v. France*[173] which held that in relation to prior restraints 'a legal framework is required, ensuring both tight control over the scope

[169] Ofcom, *'Site Blocking' to Reduce Online Copyright Infringement: A Review of Sections 17 and 18 of the Digital Economy Act* (London, 2011), 5–6.
[170] Schafer, 'No Data Protection Please, We Are British—Privacy, Porn and Prurience in the Digital Economy Bill', 355.
[171] Series A No 30, (1979-80) 2 EHRR 245.
[172] Paras. 47 and 49.
[173] Application no. 39288/98, judgment of 17 July 2001.

of bans and effective judicial review to prevent any abuse of power'.[174] In that case a French law which gave the Minister of the Interior a wide-ranging power to ban foreign publications by administrative action was held to be contrary to Article 10. Central to this finding were the facts that bans took place prior to any hearing, the criteria for bans could produce results which were 'surprising' or 'verge on the arbitrary', while the only judicial review available was limited in its scope (it did not provide for a full review of the merits) and was not automatic but required the publisher to apply to the courts.[175] Consequently, the ECtHR took the view that the judicial review procedures in place provided 'insufficient guarantees [against] abuse'.

The decision in *Ekin Association* has been applied to the internet in *Yıldırım v. Turkey*[176], in which the ECtHR considered for the first time the question of internet blocking. *Yıldırım* concerned a decision of a Turkish court which issued an order blocking access to the entirety of the Google Sites service in an attempt to prevent access to a single site critical of Atatürk. The court had initially issued an order which was limited to the offending website. That order was sent to the state Telecommunications and Information Technology Directorate (TİB) for execution. The TİB however lacked the technical capability to block this particular site and advised the court that it would be necessary to block all material hosted on the domain sites.google.com. The court varied the order accordingly.[177]

The blocking order therefore blocked a vast number of unrelated sites, including one belonging to a Turkish PhD student who found himself unable to access his own site. He claimed that this measure breached his right to freedom to hold opinions and to receive and impart information and ideas under Article 10.

The ECtHR found at the outset that this overblocking constituted an interference with his rights notwithstanding that his site was collateral damage from an order intended to target a third party.[178] Consequently the ECtHR considered whether the measure could be said to be 'prescribed by law'.

Turkish law provided that a court could order the blocking of access to 'Internet publications where there are sufficient grounds to suspect that their content is such as to amount to... offences'[179] and specified the eight classes of offences for which such orders could be issued.[180] The ECtHR nevertheless found that the blocking was not prescribed by law in that the law had 'no provision for a wholescale blocking of access such as that ordered in the present case' and did not expressly permit 'the blocking of an entire Internet domain like Google Sites which allows the exchange of ideas and information'.[181] The ECtHR was also critical of the role of the TİB as an administrative body in widening the blocking order, stating that 'the TİB could request the extension of the scope of a blocking order even though no proceedings had been brought against the website or domain in question and no real need for wholesale blocking had been established'.[182]

The ECtHR then referred to *Ekin Association* in reiterating the need for 'tight control' and 'effective judicial review' in the case of prior restraints, and found that these elements were missing. In relation to judicial review of prior restraints the Court held that this required 'a weighing-up of the competing

[174] Para. 58.
[175] Paras. 58-65.
[176] Application no. 3111/10, judgment of 18 December 2012.
[177] Paras. 8 to 12.
[178] Paras. 53 to 55.
[179] Para. 61.
[180] Para. 15.
[181] Para. 62.
[182] Para. 63.

interests at stake... designed to strike a balance between them' and also 'a framework establishing precise and specific rules regarding the application of preventive restrictions on freedom of expression'. Both were absent in Turkey, where national law did not provide for any balancing test and where the domestic courts had simply acted on the recommendation of the TİB without considering the proportionality of the blocking measure and its collateral impact on internet users. Consequently the ECtHR held that the measure 'did not satisfy the foreseeability requirement under the Convention and did not afford the applicant the degree of protection to which he was entitled by the rule of law in a democratic society'.[183]

Yıldırım is a judgment of fundamental importance for internet censorship, and particularly for its treatment of overblocking. By identifying internet blocking as a form of prior restraint the ECtHR subjects blocking to a particularly stringent set of criteria for legality as well as proportionality. Not only blocking, but also any collateral damage in the form of overblocking must be explicitly authorised by law, and that law must in turn build in safeguards to ensure that the extent of the blocking is both necessary and proportionate. Otherwise, the collateral damage caused by a blocking order will not be 'prescribed by law'.[184]

Applying the reasoning in *Ekin Association* and *Yıldırım* to the CTIRU and IWF blocking schemes, it seems clear that both fall at the first hurdle: in each case there is a system giving rise to prior restraint with no legal basis, no criteria for assessing proportionality of the interference with Article 10 rights, and no judicial review. It should be stressed that in both *Ekin Association* and *Yıldırım* there was underlying legislation, albeit that it was lacking in some regards – by contrast, the CTIRU and IWF blocking practices have no legal basis whatsoever. The 'voluntary' nature of those blocking systems would not appear to change this assessment – in each case the ISPs and filtering companies cannot look behind the determination that the relevant URLs contain illegal content. By making determinations of legality which are directly applied by the filtering firms, the state is going beyond mere support for voluntary action carried out by the firms themselves.

There are also question marks over the legality of Part 3 of the 2017 Act when measured against these standards. The power of the age-verification regulator to make blocking orders without any requirement of proportionality would not seem to be 'prescribed by law' insofar as it permits prior restraints and overblocking without any right to be heard or 'a framework establishing precise and specific rules regarding the application of preventive restrictions on freedom of expression'.

Necessity, proportionality and overblocking

For restrictions on freedom of expression to meet the requirements of Article 10(2) ECHR they must also be 'necessary in a democratic society'. This involves a proportionality test which considers whether there is a 'pressing social need'[185] for the interference and whether the interference goes further than is 'proportionate to the legitimate aim pursued'.[186] As part of this, we must ask whether an outcome could be achieved by less restrictive means than those actually used.[187] In the context of

[183] Para. 67.
[184] The ECtHR has since reiterated this approach on extremely similar facts (application of the same law to block the entirety of YouTube) in *Cengiz and others v Turkey*, application nos. 48226/10 and 14027/11, judgment of 1 December 2015.
[185] *Handyside,* para. 48.
[186] *Handyside,* para. 49.
[187] See e.g. *Campbell v United Kingdom*, Application no. 13590/88, judgment of 25 March 1992, where the Court held that the routine reading of correspondence from lawyers to prisoners could not be justified where

online censorship this test requires us to consider whether there is overblocking of unrelated content and whether more granular blocking might be used.[188]

There is as yet no consensus on what degree of collateral damage will make a restriction disproportionate under Article 10(2). In an offline context we can look to comparators such as *Ürper and others v. Turkey*[189] where the ECtHR held that orders banning the future publication of entire periodicals went beyond any restraint which might be necessary in a democratic society and therefore amounted to impermissible censorship. By analogy it seems unlikely that filtering systems which block at the level of an entire website or domain would be acceptable where there is significant legitimate content on that site or domain.[190] This is supported by the decision in *Yıldırım* where the ECtHR appeared to take the view that the blocking – in addition to lacking a legal basis - was in any event disproportionate as it 'produced arbitrary effects and could not be said to have been aimed solely at blocking access to the offending website, since it consisted in the wholesale blocking of all the sites hosted by Google Sites'.[191]

Under the Charter of Fundamental Rights, the judgment of the CJEU in *UPC v. Constantin*[192] sets a similarly stringent standard by holding that 'measures adopted by the internet service provider [to implement a blocking injunction] must be strictly targeted, in the sense that they must serve to bring an end to a third party's infringement of copyright or of a related right but without thereby affecting internet users who are using the provider's services in order to lawfully access information. Failing that, the provider's interference in the freedom of information of those users would be unjustified in the light of the objective pursued.'[193] What this means for the technical implementation of blocking measures remains to be seen, but the language of 'strict targeting' would seem to require, at a minimum, that the most granular form of blocking available should be used.

At a domestic level, the judgment of Arnold J. in *Newzbin2* has held that an order requiring blocking of an entire site was proportionate, notwithstanding that not all of the content on that site would infringe copyright, on the basis that non-infringing use of the site was *de minimis*.[194] Later intellectual

---

the legitimate needs of the prison authorities could be achieved by the less invasive method of opening (but not reading) letters suspected of containing contraband.

[188] Cost of implementation may also be a factor: see *Twentieth Century Fox v British Telecommunications* [2011] EWHC 1981 (Ch); *Twentieth Century Fox v British Telecommunications (No.2)* [2011] EWHC 2714 (Ch) and *EMI Records Ltd & Ors v British Sky Broadcasting Ltd & Ors* [2013] EWHC 379 (Ch). In those cases BT has estimated its costs in complying with blocking injunctions at £5,000 for the initial implementation and £100 for each injunction notified to it thereafter.

[189] Application nos. 55036/07, 55564/07, 1228/08, 1478/08, 4086/08, 6302/08 and 7200/08, *Ürper and Others v Turkey*, judgment of 20 October 2009, para. 44.

[190] Compare the analysis in General Comment 34: 'Any restrictions on the operation of websites, blogs or any other internet-based, electronic or other such information dissemination system, including systems to support such communication, such as internet service providers or search engines, are only permissible to the extent that they are compatible with paragraph 3. Permissible restrictions generally should be content-specific; generic bans on the operation of certain sites and systems are not compatible with paragraph 3.' United Nations Human Rights Committee, *General Comment No. 34 - Article 19: Freedoms of Opinion and Expression*, 12 September 2011, para. 43 <http://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>.

[191] Para. 68. Though compare *Akdeniz v Turkey*, application no. 20877/10, decision of 11 March 2014, which held that the applicant was not a 'victim' of a block ordered against an entire music service where he could access the same music by other means and services (para. 25). This reasoning would suggest that a deliberate degree of overblocking might be permissible if the same material remains available elsewhere.

[192] Judgment of 27 March 2014, *UPC Telekabel Wien* v *Constantin*, C-314/12,EU:C:2014:192.

[193] Para. 56.

[194] *Twentieth Century Fox v British Telecommunications* [2011] EWHC 1981 (Ch) para. 186.

property blocking cases have not elaborated on this; while the courts have developed safeguards against *accidental* overblocking, they do not appear to have assessed whether *deliberate* overblocking is permissible, and if so to what degree.[195]

This issue of overblocking will be important in the context of the powers of the age-verification regulator under Part 3 of the Digital Economy Act 2017. The regulator is expressly given power to overblock, and from the outset the government has promoted the legislation on the implicit assumption that whole sites will be blocked.[196] That may be appropriate for sites such as PornTube.com which are dedicated to pornographic content. But it will be problematic in relation to sites which mix pornography with other material – for example, a video streaming site which mixes pornography with mainstream content. The judgments in *Constantin* and *Newzbin2* suggests that blocking of such sites in their entirety would not be permissible, at any rate if the non-pornographic content is more than merely *de minimis.*

b. Article 6 ECHR

Is a decision by a public body to block a particular URL or to require takedown of particular material a 'determination of… civil rights and obligations' within the meaning of Article 6 ECHR so as to trigger the entitlement to a 'fair and public hearing' before 'an independent and impartial tribunal established by law'?

'Civil rights and obligations'

Early jurisprudence gave the concept of 'civil rights and obligations' a restrictive interpretation applying to private law obligations only, leaving most public law matters outside its scope.[197] More recently, however, the trend in the judgments of the ECtHR has been to widen the scope of the concept to ensure greater protection for individuals.[198] Consequently, while there is no authority expressly on this point, there is a strong case to be made that censorship decisions made by national authorities would fall within Article 6[199] – either on the basis that freedom of expression must be regarded as a 'civil right'[200] or else on the basis that the result of a blocking decision is decisive for private rights and obligations[201] by interfering with the commercial operation of a site.[202] The Council of Europe *Recommendation on Internet Filtering* supports the view that Article 6 applies by stating that content blocking should only take place if:

[195] See the discussion in Jaani Riordan, *The Liability of Internet Intermediaries* (2016), 492–93.
[196] 'New Blocking Powers to Protect Children Online', *GOV.UK*, 2016 <https://www.gov.uk/government/news/new-blocking-powers-to-protect-children-online> [accessed 21 November 2016].
[197] See generally David Harris and others, *Law of the European Convention on Human Rights*, 2nd edn (Oxford, 2009), chap. 6.
[198] See e.g. *Ferrazzini v Italy*, application no. 44759/98, judgment of 12 July 2001.
[199] Though compare Ola Johan Settem, *Applications of the 'Fair Hearing' Norm in ECHR Article 6(1) to Civil Proceedings* (Cham, 2016), sec. 4.1.3 who argues that the French text and caselaw require a dispute between parties for Article 6 to apply.
[200] Compare *Reisz v Germany,* application no. 32013/96, decision of 20 October 1997.
[201] *Ringeisen v Austria*, application no. 2614/65, judgment of 16 July 1971.
[202] Martin Husovec, 'Injunctions against Innocent Third Parties: The Case of Website Blocking', *JIPITEC*, 4 (2013), 123.

> the filtering concerns specific and clearly identifiable content, a competent national authority has taken a decision on its illegality and the decision can be reviewed by an independent and impartial tribunal or regulatory body, in accordance with the requirements of Article 6.[203]

Requirements imposed by Article 6

Assuming that Article 6 applies in the context of internet censorship decisions, it will trigger the requirements identified in *Belilos v. Switzerland* that such decisions must be made by a tribunal which 'determin[es] matters within its competence on the basis of rules of law and after proceedings conducted in a prescribed manner' with safeguards of 'independence, in particular of the executive; impartiality; duration of its members' terms of office; [and] guarantees afforded by its procedure'.[204]

This is not a requirement for an independent tribunal at the initial stage in every case.[205] In the interests of 'flexibility and efficiency'[206] the ECtHR will allow decisions which affect fundamental rights to be made administratively at first instance, provided that there is 'subsequent control by a judicial body that has full jurisdiction and does provide the guarantees of Article 6(1)'.[207] The case law on this point, however, makes it clear that 'full jurisdiction' requires the existence of an appeal on the merits: a mere review of legality is insufficient.[208] Consequently, judicial review in the narrow sense of domestic law would not be sufficient – for a state blocking system to be compatible with Article 6 a statutory appeal mechanism would have to be established.

Some form of notification – either before or after an initial decision is made – will also be required by Article 6 which establishes a right to an adversarial trial to include 'the opportunity for the parties to a civil or criminal trial to have knowledge of and comment on all evidence adduced or observations filed with a view to influencing the Court's decision'.[209] This applies *a fortiori* where one party has not been notified and is therefore entirely unaware of the existence of proceedings. Indeed, quite apart from Article 6 the failure to provide any notice of a censorship decision seems likely to provide the basis for judicial review on traditional procedural grounds.[210]

Who should be notified? National internet censorship procedures sometimes provide for notice to an intermediary – such as an ISP, host or social network – but not to the person responsible for the content. However, the logic of Article 10 as a right which extends to both publishers and intermediaries suggests that in all cases notification should be made to the person responsible for the material as well as the intermediary, to the extent that this is practicable.[211] In the case of blocking,

---

[203] Committee of Ministers of the Council of Europe, 'Recommendation CM/Rec(2008)6 of the Committee of Ministers to Member States on Measures to Promote the Respect for Freedom of Expression and Information with Regard to Internet Filters', 2008 <https://wcd.coe.int/ViewDoc.jsp?Ref=CM/Rec(2008)6>.
[204] Application no. 10328/83, judgment of 29 April 1988, para. 64.
[205] Harris and others, *Law of the European Convention on Human Rights*, 228–32.
[206] *Le Compte, Van Leuven and De Meyere v Belgium,* application nos. 6878/75 and 7238/75, judgment of 23 June 1981, para. 51.
[207] *Albert and Le Compte v Belgium*, application nos. 7299/75 and 7496/76, judgment of 24 October 1983, para. 29.
[208] Harris and others, *Law of the European Convention on Human Rights*, 228–32.
[209] *Vermeulen v Belgium*, application no. 19075/91, judgment of 22 January 1996, para. 33.
[210] Particularly following the decision in *Osborn v The Parole Board* [2013] UKSC 61 which stressed that human rights protection is not a distinct area of law limited to ECHR obligations but permeates the legal system as a whole.
[211] See the discussion in Poorna Mysoor, 'A Pirate Too Needs to Be Heard: Procedural Compromises in Online Copyright Infringement Cases in the UK', *Laws*, 3/3 (2014), 559–60.

this should also include a blocking notice (stop page) which indicates to the intended viewer why material was blocked and how to challenge that decision.

Applying Article 6 to our case studies, it again seems clear that the IWF and CTIRU blocking systems would fall at the first hurdle – in each case there is neither any notice nor possibility of appeal to an independent tribunal.[212]

The procedures under Part 3 of the Digital Economy Act are better, as one would expect from a statutory regime, and do provide for an independent appeal mechanism. Nevertheless, some of those procedures are still of questionable legality. As we have noted already, there is no entitlement to be heard before the regulator makes a blocking order against a site, while a notice of non-compliance to payment service providers and ancillary service providers can be issued without any prior notification to the site affected. In each case there is a likelihood of significant economic harm to a site, as well as an impact on Article 10 rights, making it surprising that prior notice and a right to be heard are not the norm.

2. EU law standards

In parallel with the ECHR, EU law establishes a number of standards which may limit censorship by Member States. This section identifies the most important EU instruments applying to these case studies and considers the ways in which they may impose obligations going further than the ECHR.

Directive on Sexual Abuse of Children and Directive on Combating Terrorism

Internet censorship has been addressed at EU level in the context of both CAM[213] and public provocation to commit a terrorist offence, and the result in both the 2011 Directive on Sexual Abuse of Children[214] and the 2017 Directive on Combating Terrorism[215] is a compromise as to how such censorship should be implemented: under each Directive Member States *must* take measures to ensure the prompt removal of material hosted in their territory; *must* endeavour to obtain the removal of such material hosted outside their territory; and *may* take measures to block access to such material.[216]

---

[212] In France, by comparison, administrative blocking (of CAM and now terrorist material) has been upheld by the Conseil d'État where the law provides for oversight of the blocking system by the CNIL and the possibility of a judicial remedy against blocking. See e.g. Marion Lacaze, 'Latest Developments in the Repression and Prevention of Terrorism under French Criminal Law', *Montesquieu Law Review*, 3, 2015, 2; Winston Maxwell, 'CNIL's New Role: Overseeing Website Blocking', *Lexology*, 2016 <https://www.lexology.com/library/detail.aspx?g=2632d9f2-27bd-4485-9092-33f326fc473d> [accessed 16 March 2018].

[213] For background see McIntyre, 'Blocking Child Pornography on the Internet', 209.

[214] Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA.

[215] Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA.

[216] In the case of child pornography, approximately half of all Member States have introduced blocking. Of these, half have done so on a legislative basis, half on a voluntary basis. European Commission, 'Report from the Commission to the European Parliament and the Council Assessing the Implementation of the Measures Referred to in Article 25 of Directive 2011/93/EU of 13 December 2011 on Combating the Sexual Abuse and Sexual Exploitation of Children and Child Pornography', 2016, COM(2016) 872 final <http://www.europarl.europa.eu/RegData/docs_autres_institutions/commission_europeenne/com/2016/0872/COM_COM(2016)0872_EN.pdf> [accessed 6 July 2017].

At first glance, each Directive appears to impose significant safeguards, including a requirement of judicial redress. Article 25(2) of the 2011 Directive requires these in the case of blocking measures only:

> Member States may take measures to block access to web pages containing or disseminating child pornography towards the Internet users within their territory. These measures must be set by transparent procedures and provide adequate safeguards, in particular to ensure that the restriction is limited to what is necessary and proportionate, and that users are informed of the reason for the restriction. Those safeguards shall also include the possibility of judicial redress.

Article 21(3) of the 2017 Directive goes further, specifying safeguards which must apply to both removal and blocking:

> Measures of removal and blocking must be set following transparent procedures and provide adequate safeguards, in particular to ensure that those measures are limited to what is necessary and proportionate and that users are informed of the reason for those measures. Safeguards relating to removal or blocking shall also include the possibility of judicial redress.

In each case, however, the effect of these safeguards is limited by the fact that they apply only to measures taken by Member States themselves. The Recitals to both Directives expand on this by excluding 'voluntary' action by the internet industry, and Member State support for such action. Recital 47 of the 2011 Directive provides that:

> The measures undertaken by Member States in accordance with this Directive in order to remove or, where appropriate, block websites containing child pornography could be based on various types of public action, such as legislative, non-legislative, judicial or other. In that context, this Directive is *without prejudice to voluntary action taken by the Internet industry to prevent the misuse of its services or to any support for such action by Member States*. Whichever basis for action or method is chosen, Member States should ensure that it provides an adequate level of legal certainty and predictability to users and service providers.[217]

Recital 22 of the 2017 Directive goes further again by specifically excluding situations where Member States themselves detect and flag terrorist content:

> The measures undertaken by Member States in accordance with this Directive in order to remove online content constituting a public provocation to commit a terrorist offence or, where this is not feasible, block access to such content could be based on public action, such as legislative, non-legislative or judicial action. In that context, *this Directive is without prejudice to voluntary action taken by the internet industry to prevent the misuse of its services or to any support for such action by Member States, such as detecting and flagging terrorist content.* Whichever basis for action or method is chosen, Member States should ensure that it provides an adequate level of legal certainty and predictability for users and service providers and the possibility of judicial redress in accordance with national law. Any such measures must take account of the rights of the end users and comply with existing legal and judicial procedures and the Charter of Fundamental Rights of the European Union.[218]

[217] Emphasis added.
[218] Emphasis added.

The remaining language in each Recital about ensuring 'legal certainty', 'predictability' and 'end user rights' is aspirational – using the term 'should' rather than 'must' – and reflects the position of Member States such as the UK which already had self-regulatory systems and were unwilling to agree to rules which would have required them to put in place a domestic legislative basis. The result is that these safeguards are of little practical significance. Any removal/blocking measures taken directly by public bodies would be subject to the equivalent requirements of Articles 10 and 6 ECHR in any event[219], while voluntary/self-regulatory measures are beyond their scope.[220] For these systems the language in the recitals provides a veneer of legality, but not the reality.

The EU Open Internet Regulation

The developing legal framework on net neutrality offers more potential for applying fundamental rights to internet censorship, even in relation to genuinely voluntary and self-regulatory actions. While net neutrality has often been narrowly framed as an issue of economics, innovation policy and consumer choice[221], more recent work highlights how it promotes freedom of expression and privacy in communications and in this sense it has the potential to disrupt internet censorship implemented through ISPs.[222]

At an EU level, net neutrality provisions have been in force since April 2016 under the Open Internet Regulation.[223] The key provision for blocking is Article 3(3) which provides that:

> Providers of internet access services… shall not block… specific content... except as necessary, and only for as long as necessary, in order to… comply with Union legislative acts, or national legislation that complies with Union law, to which the provider of internet access services is subject, or with measures that comply with Union law giving effect to such Union legislative acts or national legislation, including with orders by courts or public authorities vested with relevant powers.

Recital 13 elaborates on this by explaining that:

> The requirement to comply with Union law relates, inter alia, to the compliance with the requirements of the Charter of Fundamental Rights of the European Union ('the Charter') in relation to limitations on the exercise of fundamental rights and freedoms. [A]ny measures liable to restrict those fundamental rights or freedoms are only to be imposed if they are

---

[219] Joe McNamee and Maryant Fernández Pérez, 'Net Neutrality: An Analysis of the European Union's Trialogue Compromise', in *Net Neutrality Compendium*, ed. by Luca Belli and Primavera De Filippi (Cham, 2016), 187.
[220] In a 2016 report on the implementation of the 2011 Directive the Commission noted, but did not criticise, the fact that the UK does not offer judicial redress but merely a non-statutory appeals process in relation to IWF decisions. European Commission, 'Report from the Commission to the European Parliament and the Council Assessing the Implementation of the Measures Referred to in Article 25 of Directive 2011/93/EU of 13 December 2011 on Combating the Sexual Abuse and Sexual Exploitation of Children and Child Pornography', 11.
[221] Christopher Marsden, 'Network Neutrality: A Research Guide', in *Research Handbook on Governance of the Internet*, ed. by Ian Brown (Cheltenham, 2013).
[222] See e.g. Milton Mueller, 'Net Neutrality as Global Principle for Internet Governance' (2007) <http://www.internetgovernance.org/wordpress/wp-content/uploads/NetNeutralityGlobalPrinciple.pdf> [accessed 12 July 2013]; Christopher Marsden, *Net Neutrality: Towards a Co-Regulatory Solution* (London, 2010), 18–19, 105–31.
[223] Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 laying down measures concerning open internet access and amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services and Regulation (EU) No 531/2012 on roaming on public mobile communications networks within the Union.

> appropriate, proportionate and necessary within a democratic society, and if their implementation is subject to adequate procedural safeguards in conformity with the European Convention for the Protection of Human Rights and Fundamental Freedoms, including its provisions on effective judicial protection and due process.

These provisions – by requiring that blocking be carried out on the basis of legislative acts, and must include effective judicial protection and due process – sound the death knell for purely self-regulatory blocking systems by providers of internet access services.[224] Article 10(3) of the Regulation confirms this by identifying self-regulatory schemes as violating Article 3 and requiring that existing schemes be brought to an end by 31 December 2016.[225]

Impact on the IWF URL List and family-friendly filters

This prohibition on self-regulatory censorship is particularly significant in the UK context where it precludes both main forms of ISP censorship – use of the IWF URL List and family-friendly filters. This does not seem to have been fully appreciated by the UK government as an inevitable consequence of the Regulation – in October 2015 Prime Minister David Cameron told the House of Commons that he 'spluttered over his cornflakes' on reading a Daily Mail story stating that the newly adopted Regulation would prevent ISPs from using family-friendly filters.[226] Cameron went on to promise legislation to permit such filters, and this was eventually adopted as section 104(1) of the Digital Economy Act 2017 which provides:

> A provider of an internet access service to an end-user may prevent or restrict access on the service to information, content, applications or services, for child protection or other purposes, if the action is in accordance with the terms on which the end-user uses the service.

Section 104(1) is quite a remarkable piece of drafting. It goes well beyond the IWF URL List or family-friendly filters and permits any form of blocking, for child protection or any 'other purposes', without any legislative basis or user consent, provided only that it is 'in accordance with the terms on which the end-user uses the service'. Since terms of use are set by ISPs on a take it or leave it basis, this guts the Open Internet Regulation by allowing ISPs to dictate their own blocking policies.

A narrowly tailored provision confirming that ISPs can block material at the request of subscribers would have been permissible: despite some uncertainty in this area, the Commission has taken the position that filters blocking pornography can be used by ISPs if they respect 'the basic principle in the new Regulation that end-users — in this case, parents of minors — should be able to choose freely the content, applications and services to which they wish to have access'.[227] As is, however, the section is in clear breach of EU law by displacing Article 3(3) and purporting to permit blocking without the judicial protection and procedural safeguards required by the Regulation.

In ordinary circumstances, one might have expected this section to be challenged on these grounds eventually. In light of Brexit, however, it seems likely that the section will survive until the UK leaves

---

[224] Swiss Institute of Comparative Law, *Comparative Study on Blocking, Filtering and Take-Down of Illegal Internet Content*, 24–25.

[225] Article 10(3) provides that 'Member States may maintain until 31 December 2016 national measures, including self-regulatory schemes, in place before 29 November 2015 that do not comply with Article 3(2) or (3). Member States concerned shall notify those measures to the Commission by 30 April 2016.'

[226] Andy Phippen, *Children's Online Behaviour and Safety* (London, 2017), 26.

[227] 'Answer to a Written Question - Impact of Telecoms Regulation upon Opt-in Porn Filters Online', 2016, E-014433/2015 <http://www.europarl.europa.eu/sides/getAllAnswers.do?reference=E-2015-014433&language=EN> [accessed 1 August 2017].

the EU. The Open Internet Regulation will be important in governing internet censorship elsewhere in the EU – in Germany, for example, the telecoms regulator has already taken steps to prohibit ISP blocking without legal basis.[228] In the UK, however, this aspect of the Regulation has effectively been negated.

EU Consumer law

A common complaint about self-regulatory censorship measures is that they may evade both vertical and horizontal accountability norms – vertical in that the actions of private firms cannot be attributed to the state, and horizontal in that a firm's terms of use will invariably be drafted so widely as to prevent a user from having a remedy in cases of abuse.[229] Might it be possible to address this as a consumer rights issue?[230]

In March 2017 the European Commission and a number of European consumer authorities adopted a common position which challenges the censorship practices of Facebook, Google and Twitter as part of a wider review of their compliance with consumer law.[231] The core argument is that provisions which permit removal of content on vague and discretionary grounds, without any form of notice or redress, constitute unfair terms within the meaning of the Unfair Contract Terms Directive.[232] For example, while each firm already has appeals in place for certain issues, such as account suspensions, only Twitter allows appeals against removal of individual items.[233] According to the common position:

> The criteria on the basis of which social media operators can refuse to display or remove content generated by the consumer cannot remain general or unspecified. Standard terms and conditions should contain a sufficiently detailed indication of the main grounds on which content can be removed and possibly of how consumers are informed and can appeal to the removal of content… This however, should not prevent social media operators from providing in their standard terms very clearly that user generated content can be removed without notice, when this is needed to stop rapidly illegal conducts [*sic*].[234]

In short, the consumer authorities have demanded that the firms put in place detailed rules regarding what content is permissible, along with notification prior to removal of content (except in cases of urgency) and some form of appeal mechanism after the fact – procedures which have the potential to

---

[228] Bundesnetzagentur, 'Net Neutrality in Germany Annual Report 2016/2017', 2017, 11.
[229] See e.g. Laura DeNardis and Andrea Hackl, 'Internet Governance by Social Media Platforms', *Telecommunications Policy*, 39/9 (2015), 761–70.
[230] For a general analysis of the application of consumer law to social media terms of use see Ellen Wauters, Eva Lievens and Peggy Valcke, 'Towards a Better Protection of Social Media Users: A Legal Perspective on the Terms of Use of Social Networking Sites', *International Journal of Law and Information Technology*, 22/3 (2014), 254–94.
[231] European Commission, 'The European Commission and Member States Consumer Authorities Ask Social Media Companies to Comply with EU Consumer Rules', 2017 <http://europa.eu/rapid/press-release_IP-17-631_en.htm> [accessed 25 July 2017]; 'EU Increases Pressure on Facebook, Google and Twitter over User Terms', *Reuters*, 25 July 2017 <http://www.reuters.com/article/us-socialmedia-eu-consumers-idUSKBN1A92D4>.
[232] Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts, specifically as a term within Annex 1.m which prohibits 'Terms which have the object or effect of: giving the seller or supplier the right to determine whether the goods or services supplied are in conformity with the contract, or giving him the exclusive right to interpret any term of the contract.'
[233] 'How to Appeal', *Online Censorship* <https://onlinecensorship.org/resources/how-to-appeal> [accessed 31 July 2017].
[234] European Commission, 'The European Commission and Member States Consumer Authorities Ask Social Media Companies to Comply with EU Consumer Rules'.

address some of the criticisms around otherwise opaque systems such as the CTIRU takedown function.

How have the firms responded? As of February 2018, Google has accepted these criticisms and agreed to clarify the grounds on which content can be removed, introduce notification before removal, and put in place a general appeal procedure against removal.[235] Facebook and Twitter, however, have resisted these demands – in particular, neither is willing to notify users before content is removed, Facebook has refused to introduce appeals against content removal, and Twitter has refused to modify terms of use which give it unfettered discretion to remove content.

The application of consumer law in this area has potential, but is still very much at a preliminary stage. It remains to be seen how hard national consumer authorities will push and whether the courts will uphold this use of the Unfair Contract Terms Directive. The possible impact is also limited by the common position of the consumer protection authorities which is limited to procedural, rather than substantive, guarantees of freedom of expression. Terms of use which ban lawful speech would not be affected, provided that they are clearly set out and provide minimum procedural fairness. At best, consumer law might help to promote transparency as to *what* and *how* social media firms censor – but it will not itself require these private spaces to be governed as though they were public forums.

Part 3: CONCLUSION

It is difficult to offer an overall assessment of internet censorship in the UK. There is no single model of censorship but rather a patchwork of individual schemes, and the three case studies chosen for this chapter differ greatly in relation to core elements such as their legal basis, procedural safeguards and accountability mechanisms. Nevertheless, these case studies do share some common themes and illustrate several wider issues.

At the outset, we can identify a direct evolution of UK policy from CAM to terrorist material. While the government has not succeeded in its aim of adding terrorist material to the IWF URL List it has nevertheless drawn heavily on tactics previously developed in the context of CAM: crowd sourcing the process of detection by encouraging users to report content they consider to be illegal, matching images by using hash value databases to automate the process of identification and takedown and relying on voluntary cooperation with industry rather than statutory processes.

Unfortunately, in borrowing from the censorship of CAM the government has neglected the strengths of the IWF system while replicating its failings. While the IWF has often been criticised as operating a form of privatised censorship[236], the IWF has nevertheless demonstrated a level of transparency in its activities which far outweighs that of the Home Office in relation to the CTIRU's functions. The IWF has also shown much greater willingness to examine its own practices against human rights standards; most notably in 2014 by commissioning Ken Macdonald QC (Lord Macdonald) to carry out a human rights audit prompted largely by criticisms expressed by an academic author in a journal article.[237]

---

[235] European Commission, 'Social Media Companies Need to Do More to Fully Comply with EU Consumer Rules', 2018 <http://europa.eu/rapid/press-release_IP-18-761_en.htm> [accessed 15 February 2018].
[236] Edwards, 'Pornography, Censorship and the Internet', 657–58.
[237] Laidlaw, 'The Responsibilities of Free Speech Regulators', 312; Internet Watch Foundation, 'Board Minutes 16 October 2012', 2012 <https://www.iwf.org.uk/assets/media/accountability/board/Minutes%2016%20October%202012%20Web.pdf> [accessed 12 June 2013]; Macdonald, 'A Human Rights Audit of the Internet Watch Foundation'.

The experience in relation to CAM and terrorism also illustrates a dangerous trend towards convergence of censorship and surveillance. Both the IWF and CTIRU blocking lists are being used, with the cooperation of the state, to identify individual internet users visiting blocked web pages, echoing the US experience of the PhotoDNA system being used extensively to scan private emails and files. In each case this is a form of indiscriminate surveillance, taking place in close coordination with the state but without any prior suspicion or legal basis. This is an under-examined area: the bulk of the literature on this point comes from the United States where privacy rights against private firms are much weaker[238] and there is little written on this from a European perspective.[239] However it is difficult to see that these systems would survive legal challenge when fully assessed against Article 8 ECHR, the General Data Protection Regulation[240] and the standards elaborated by the CJEU in *Schrems*.[241] (A particularly disturbing development in this context is the Commission recommendation of 1 March 2018 on tackling illegal content online which recommends that service providers 'in the context of their activities for the removal… of access to illegal content' should be obliged to report any 'evidence of alleged serious criminal offences' to police.[242] If implemented, this would have the effect of unifying censorship and surveillance across all service providers, requiring user information to be handed over without any legal process.)

This overlap between censorship and privacy recurs in a slightly different way in the context of age verification under the Digital Economy Act 2017. Age verification systems have the potential to record and expose sensitive information about the sexual preferences of individuals, and it is disappointing that the Act does not take any steps to minimise these risks.

More generally, all three case studies indicate that the application of fundamental rights standards to UK internet censorship is still at a nascent stage. This is, perhaps, unsurprising in relation to the IWF and CTIRU schemes where the lack of a legal obligation to filter or takedown material feeds into a traditional view that 'voluntary' measures are not subject to fundamental rights standards. It is more surprising to see that the Digital Economy Act 2017 still does not take full account of ECHR obligations in relation to even basic matters such as the right to be heard before a blocking order is made, or the requirement of proportionality in blocking.

Turning to the EU law issues which arise from these case studies, this chapter highlights a tension between different legal instruments. On one side, there is an EU framework which can be used to ensure freedom of expression rights against internet intermediaries – even in schemes where there is no state involvement. The Open Internet Regulation requires a legislative framework for blocking measures; similarly, the Commission and national consumer authorities have made creative use of the Unfair Contract Terms Directive in a way which will ensure greater procedural protections for users against removal of their content. On the other side, the Directive on Sexual Abuse of Children and

---

[238] See e.g. Alexandra L. Mitter, 'Deputizing Internet Service Providers: How the Government Avoids Fourth Amendment Protections', *New York University Annual Survey of American Law*, 67 (2011), 235.

[239] See Christina Angelopoulos and others, *Study of Fundamental Rights Limitations for Online Enforcement through Self-Regulation* (Amsterdam, 2016), sec. 4 <https://openaccess.leidenuniv.nl/bitstream/handle/1887/45869/IVIRStudyOnlineenforcementthroughself-regulation.pdf?sequence=1> [accessed 12 July 2017].

[240] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

[241] Judgment of 6 October 2015, *Maximillian Schrems v Data Protection Commissioner*, C-362/14, ECLI:EU:C:2015:650.

[242] European Commission, 'Recommendation on Measures to Effectively Tackle Illegal Content Online', 2018, C(2018) 1177 final, para. 24.

Directive on Combating Terrorism have been drafted with a view to facilitating self-regulation, and both envisage state involvement in voluntary blocking and takedown schemes without the need for a legislative basis or procedural safeguards.

This tension reflects a wider trend in Brussels where the Commission has increasingly moved towards weakening intermediary immunities and encouraging or requiring greater policing by internet intermediaries.[243] Most recently we see this in the Commission communication of 28 September 2017[244] and recommendation of 1 March 2018[245] on tackling illegal content online, which call for 'an enhanced responsibility of online platforms', to include proactive searches for potentially illegal content, automated hash value matching for preventing uploads and removal of content, and systems of 'trusted flaggers' to report and remove allegedly illegal content more easily – with these measures to take place on a voluntary basis. As with the EU IRU, this approach follows the informal UK model rather than national schemes such as the German Network Enforcement Law (the Netzwerkdurchsetzungsgesetz/'NetzDG') which provide for legislative frameworks.[246] It is ironic that the UK model appears to influence EU policy here even as the UK itself heads for the door.

Speaking of Brexit brings us to the somewhat dispiriting conclusion that fundamental rights protections in this area – as in other areas – will be substantially weakened by the UK's departure from the EU.[247] The most visible loss will be the Charter of Fundamental Rights, but the Open Internet Regulation will also be a significant casualty which is unlikely to be replaced by comparable domestic legislation. In addition, Brexit will have the effect of dis-entrenching the ECHR, making it possible for the Conservative Party to proceed with its reported plan of campaigning in the next election on a commitment to leave.[248] UK governments have not been notable for their respect for fundamental rights online; one wonders whether how they might behave if freed entirely from oversight by either the Strasbourg or Luxembourg courts.

---

[243] Luca Belli and Cristiana Sappa, 'The Intermediary Conundrum: Cyber-Regulators, Cyber-Police or Both?', *JIPITEC*, 8/3 (2017) <http://www.jipitec.eu/issues/jipitec-8-3-2017/4620>.

[244] European Commission, 'Tackling Illegal Content Online: Towards an Enhanced Responsibility of Online Platforms', 2017, COM(2017) 555 final.

[245] European Commission, 'Recommendation on Measures to Effectively Tackle Illegal Content Online'.

[246] For an initial analysis of the NetzDG see Aleksandra Kuczerawy, 'Phantom Safeguards? Analysis of the German Law on Hate Speech NetzDG', *CITIP Blog*, 2017 <https://www.law.kuleuven.be/citip/blog/phantom-safeguards-analysis-of-the-german-law-on-hate-speech-netzdg/> [accessed 16 February 2018].

[247] See e.g. Tobias Lock, *Human Rights Law in the UK after Brexit*, Edinburgh School of Law Research Paper No. 2017/17 (29 September 2017) <https://papers.ssrn.com/abstract=3046554> [accessed 16 February 2018].

[248] Christopher Hope, 'Theresa May to Fight 2020 Election on Plans to Take Britain out of European Convention on Human Rights after Brexit Is Completed', *The Telegraph*, 28 December 2016 <http://www.telegraph.co.uk/news/2016/12/28/theresa-may-fight-2020-election-plans-take-britain-european/> [accessed 17 February 2018].