For the record: self-deleting messaging systems and compliance with public law duties

Joe Tomlinson* and Cassandra Somers-Joce+

In the last couple of years it has become clear that instant messaging technologies which have the capacity to automatically delete messages, either immediately or after a specified period of time, are used within the UK government. In a debate in the House of Commons in September 2019, Caroline Lucas MP reported allegations that the rationale for the constitutionally controversial prorogation of Parliament had been articulated "not through the official channels of Government emails and memos, but by personal email, WhatsApp and "burner" phones."[1] In February 2021 it was reported that there had been extensive WhatsApp messages between Matt Hancock MP, the then Secretary of State for Health, and Alex Bourne, who secured a government contract for the supply of medical devices.[2] In May 2021, it was reported that David Cameron, the former Prime Minister, had sent numerous WhatsApp messages to the Chancellor of the Exchequer and civil servants concerning Greensill Capital, where Mr Cameron was an advisor, had options on shares, and which collapsed soon after.[3] In June 2021 the Department for Digital, Culture, Media and Sport confirmed that their officials use self-deleting instant messaging systems.[4] Soon after, a Cabinet Office policy that positively encourages the use of self-deleting instant messaging emerged.[5] Such developments make clear

* Senior Lecturer in Public Law, University of York. We are grateful to Simon Lovitt for research assistance.
+ Researcher, University of York.
[1] HC Deb 9th September 2019, vol 664, col 552.
[2] BBC News 'Coronavirus: Medical regulator investigates £30m Covid contract firm' (BBC News 21st Feb 2021), <https://www.bbc.co.uk/news/uk-politics-56145492> (accessed 12-11-2021).
[3] House of Commons Treasury Committee, 'Treasury Committee Oral evidence: Lessons from Greensill Capital,' HC 151, Thursday 13th May 2021, at pages 9, 13, 17, 23 and 24. <https://committees.parliament.uk/oralevidence/2163/pdf/> accessed 02-12-2021.
[4] Haroon Siddique, 'UK government admits ministers can use self-deleting messages' (The Guardian 13th Jun 2021) <https://www.theguardian.com/politics/2021/jun/13/uk-government-admits-ministers-can-use-self-deleting-messages> accessed 16-11-2021.
[5] Cabinet Office 'Information and Records Retention & Destruction Policy' (Undated), 13.

the reality of the increasing deployment of such technologies within government and that they can be part of important decision-making processes. However, the extent to which they are being used, where they are being used, and how they are being used remains much less clear.

These developments may appear inevitable and unsurprising: officials in 2021, like everyone else in society, exist at a time of profound transformation in communication technologies and the use of what are now everyday technologies in their work may be expected. Moreover, efficient and secure internal communications are an important part of good government. Many complex decisions with multiple considerations and implications are taken every day, and often without the luxury of time. It is right that government seeks to capitalise on new technologies to strengthen its competencies in this respect. Nevertheless, the use of self-deleting instant messaging technology raises an important question of public law as, in the adoption and use of such technology, public officials must be mindful of their particular legal duties and ensure they are complied with in practice, including that there are sufficient structures in place to ensure compliance with those legal obligations. The central question at this juncture is whether there has been compliance with those duties or if incremental and fragmented implementation of these systems has lapsed, at least in places, into illegality.[6]

At the outset, it is important to clarify the features of the technology in question and the nature of the challenge it presents. Arguably the most popular messaging platform with self-deleting instant messaging functionality is WhatsApp. Owned by U.S. technology giant Meta, the system is mostly used on smartphones but is also available on other types of device. It allows users to message each other directly or form messaging groups. Messages can be in a

[6] Similar questions are arising in other jurisdictions, see *e.g.* Daxton Stewart, 'Killer Apps: Vanishing Messages, Encrypted Communications, and Challenges to Freedom of Information Laws When Public Officials go Dark' (2019) 10(1) Case Western Reserve Journal of Law, Technology and the Internet 1.

variety of mediums, including written text, photos, videos, voice notes, and emojis. Messages are end-to-end encrypted, so that the communication is entirely private between the person sending the message and the person(s) receiving it. WhatsApp has a "disappearing messages" option which allows users to have messages deleted automatically and completely after a set time period. When this deletion occurs, no users can access the messages and nor can the platform itself. It is, however, possible for a user to take a screenshot of a message and store it as a separate image before it disappears. There are multiple messaging platforms with this seld-deleting functionality. For some platforms it is optional and for some it is mandatory.[7] It is also highly likely that the availability of this functionality in instant messaging platforms will expand in the future. The central rationale for the self-deleting function in platforms created by private companies is "privacy by design."[8] That is to say, users generally prefer their messages to be private, such that platforms which provide self-deleting functionality may enjoy a competitive advantage. It may also be generally considered to represent more ethical data practice to do so. The essential, first-order concern with the use of such functionality in messaging systems used in the public sector, however, is that what was intended to enhance the privacy of citizens may result in the destruction or undermining of the official record. It is important to be clear on this point; the initial concern that this technology triggers pertains to the maintenance of the public record, not the disclosure of those records. There are separate and established processes and principles relevant to when and how public records may be disclosed. Those mechanisms risk being undermined if there is a failure to preserve the record in the first instance, but the issues are conceptually distinct.

[7] For instance, as discussed in Agnieszka McPeak, 'Self-Destruct Apps: Spoliation by Design?' [2018] *Akron Law Review* 633.
[8] Ira S Rubinstein, 'Regulating Privacy by Design' (2011) 26(3) *Berkeley Technology Law Journal* 1409.

In terms of the relevant legal obligations, there are essentially three frameworks that regulate this issue. The first is the Public Records Act 1958. Under Section 1(1) of that Act, the Secretary of State is responsible for supervising the care and preservation of public records. Section 10 and Schedule 1 define the scope of "public records" to include "records of, or held in, any department of Her Majesty's Government in the United Kingdom" and this extends "not only to written records but records conveying information by any means whatsoever." Section 3 of the Act sets out duties for public officials to make arrangements for the selection and preservation of those records which ought to be preserved, in line with the guidance of the Keeper of Public Records (which has statutory effect). This guidance, The National Archives' Record Collection Policy, further clarifies what sort of material will constitute a "record."[9] It provides that public records can exist in any format, including digital formats: records may be in "any medium, including social media channels and they may have originated in private email accounts, not only in the government's own systems." The policy further provides that the National Archives will seek to "collect and preserve public records which document" the "principal policies and actions of the UK central government and English and Welsh Governments." This includes "records illustrative of the process of developing government policy and legislation," "records which detail changes in the strategic functions and obligations of the UK and English and Welsh Governments," and "records relating to the review and evaluation of policy." The Public Records Act is therefore relevant in this context as the obligations to preserve records may well require the preservation of instant messages, which may be undermined by automatic deletion functions.

[9] National Archives, *Record Collection Policy* (November 2012).

The second relevant legal framework is the Freedom of Information Act 2000, which creates a right, qualified by various provisions, to information held by public authorities.[10] There is therefore a corresponding duty placed on public authorities to provide such information when a request is made, and public authorities are given detailed guidance on how to manage information for this purpose in a Code of Practice produced under Section 46.[11] Section 77 of the Act creates a criminal offence in circumstances where an individual "alters, defaces, blocks, erases, destroys or conceals any record held by the public authority, with the intention of preventing the disclosure by that authority of all, or any part, of the information." The Act is relevant to self-deleting messages in government as it is entirely possible that a legitimate request may be made within the parameters of the Act for information that is stored as part of an instant message but is subsequently automatically deleted. Failure to preserve messages due to the operation of automatic deletion technology has the potential to cut across the scheme of the Act in general, and in the manner that sections 46 and 77 in particular were patently concerned with preventing.

The third framework is the common law duty of candour. A public authority defendant in judicial review proceedings has a duty "to co-operate and to make candid disclosure by way of affidavit of the relevant facts and (so far as they are not apparent from contemporaneous documents which have been disclosed) the reasoning behind the decision challenged."[12] The central underlying idea is that, in the context of judicial review, a public authority's aim is to assist the court in its role of ensuring the lawfulness of the decision under challenge, rather than to conduct litigation with a "win at all costs" attitude. Under the duty, a public authority

[10] As regards information concerning environmental matters, see The Environmental Information Regulations 2004, SI 2004/3391.
[11] *Code of Practice on the Management of Records issued under section 46 the Freedom of Information Act 2000* (2021).
[12] *Belize Alliance of Conservation Non-Government Organisations v Department of the Environment* [2004] UKPC 6 [86].

must fairly and fully disclose all relevant information, including that any information that may be adverse to its own position.[13] The duty applies as soon as a public body is aware that someone is likely to challenge a decision that affects them. It then applies "to every stage of the proceedings including letters of response under the pre-action protocol, summary grounds of resistance, detailed grounds of resistance, witness statements and counsel's written and oral submissions."[14] Importantly for present purposes, it is not just a duty to disclose documentary evidence. Instead, the duty revolves around information – in the broader sense – that may be relevant, and it cannot be assumed that documents alone will suffice to ensure it is discharged. A variety of adverse consequences may arise where the duty of candour is breached by a defendant public authority. For instance, a lack of candour may allow the court to draw adverse inferences of fact.[15] The relevance of the duty of candour to self-deleting instant messaging technology results from the fact that it is well within the realm of reasonable possibilities that messages sent via a platform with self-deleting functionality enabled ought to be disclosed in the course of a judicial review.[16] In this capacity, a self-deleting function developed to embed "privacy by design" has the potential to become a function for "spoliation by design."[17]

A final point on legal frameworks must be made at this juncture. The relevant frameworks set out above include statutory-backed guidance policies, which elaborate in detail on the duties of public authorities to maintain the record and standard practice on how to operationalise those duties. It is also clear that different bodies are now creating their own guidance, or at least adjusting their extant policies. This in turn raises the issue of the effect of

[13] Treasury Solicitor's Department, *Guidance on Discharging the Duty of Candour and Disclosure in Judicial Review Proceedings* (January 2010), 7.
[14] Ibid, 4.
[15] *R v Lancashire County Council, ex parte Huddleston* [1986] 2 All ER 941, 947; *R (Quark Fishing Ltd) v The Secretary of State for Foreign and Commonwealth Affairs* [2002] EWCA Civ 1409 [50].
[16] *R (Good Law Project Ltd) v Secretary of State for Health and Social Care* [2021] EWHC 2595 (TCC).
[17] Agnieszka McPeak, 'Self-Destruct Apps: Spoliation by Design?' [2018] *Akron Law Review* 633.

policy on public law duties. The adoption of a policy creates a public law duty to comply with that policy, unless there is a good reason for departing from it.[18] In this context, it is also worth noting that the two key policies—The National Archives' Record Collection Policy and the Section 46 Code of Practice—are made pursuant to statutory functions and, in the case of the former, written by a body with specialist expertise. As regards obligations as to the content of any policy, the Supreme Court recently considered this matter in *R (A) v Secretary of State for the Home Department* [19] and *R (BF (Eritrea)) v Secretary of State for the Home Department*.[20] The restated position is that a policy will be unlawful if it misdirects the government as to its legal obligations or creates a risk of impeding access to justice.[21] In the former situation, there must be a:

> [C]omparison of what the relevant law requires and what a policy statement says regarding what a person should do. If the policy directs them to act in a way which contradicts the law it is unlawful... The test does not depend on a statistical analysis of the extent to which relevant actors might or might not fail to comply with their legal obligations.[22]

In a case where a policy risks impeding access to justice, the approach differs insofar as "it is legitimate to have regard to evidence regarding its likely impact and the court has to make an overall evaluative assessment whether this legal standard is met or not (and statistics might have a part to play in making such an assessment)."[23] In relation to policies in the context of persevering public records, it is strongly arguable that policies fall into the latter category,

[18] Lord Wilson in *Mandalia v Secretary of State for the Home Department* [2015] 1 WLR 4546 [29-31]; *R (Lee-Hirons) v Secretary of State for Justice* [2017] AC 52 [50] (Lord Reed), [17] (Lord Wilson); *R (Hemmati) v Secretary of State for the Home Department* [2019] 3 WLR 1156 [50] [69] (Lord Kitchin).
[19] [2021] UKSC 37.
[20] [2021] UKSC 38.
[21] *R (A) v Secretary of State for the Home Department* [84].
[22] *R (A) v Secretary of State for the Home Department* [41].
[23] *R (A) v Secretary of State for the Home Department* [80].

given the relevance of the public record to potential legal challenges to government decision-making and the duty of candour in judicial review.

It is crystal clear that messages sent via platforms that have self-deleting functions, such as WhatsApp, may constitute both public records under section 3 of the Public Records Act 1958, information held by a public authority for the purposes of the Freedom of Information Act 2000, and material that may be relevant to disclose in the course of judicial review proceedings under the duty of candour. Though not all such messages will meet the threshold where they will need to be preserved, it is similarly clear that some exchanges do meet this threshold. For instance, there are multiple examples in the public domain of WhatsApp messages where the critical elements of the COVID-19 management strategy is discussed between ministers and senior advisors. The pressing question of law, therefore, is whether policies and practices are complying with these duties to preserve the public record. The apparent fragmentation in the implementation of self-deleting messaging systems in government and the divergence in approach to how policies and guidance have been adjusted makes this a thorny matter that will require unpicking in particular instances. However, there are three points that can be made at a general level, based on the current approach and available evidence, as to the potential for public law failures in this context.

First, it is clear, based on what is in the public domain, that there is some basis to suspect there has been a degree of failure to comply with record maintenance policies that are in place as a result of the use of certain messaging technologies, particularly as regards the Code of Practice under Section 46 of the Freedom of Information Act and the National Archives' Record Collection Policy under the Public Records Act. Aside from the examples set out at the start of this article, multiple authoritative sources have given public commentary on their

concerns about this possibility. The Information Commissioner recently launched a formal investigation into the use of private correspondence channels at the Department for Health and Social Care and alluded to the issue of self-deleting messages:

> To be clear, the use of private correspondence channels does not in itself break freedom of information or data protection rules. But my worry is that information in private email accounts or messaging services is forgotten, overlooked, autodeleted or otherwise not available when a freedom of information request is later made. This frustrates the freedom of information process, and puts at risk the preservation of official records of decision making.[24]

Similarly, in October 2020, Bodley's Librarian and President of Digital Preservation Coalition, Richard Ovenden, wrote to *The Financial Times* expressing concern about the use of messaging services with automatic deletion capability within government:

> The mode of communication in government has already shifted to the digital realm, and the use of such technologies should be a matter of concern for all members of the public whatever their political persuasion. They include services like Snapchat and Signal, where messages auto-erase, being designed originally for teenagers who did not wish to have their private messages hanging around on their phones to be discovered by parents. Today, the systems of recalcitrant youths have been adopted by senior government officials and politicians.[25]

---

[24] Elizabeth Denham, 'ICO launches investigation into the use of private correspondence channels at the Department of Health and Social Care' (*Information Commissioner's Office*, 6 July 2021) <https://ico.org.uk/about-the-ico/news-and-events/blog-ico-launches-investigation-into-the-use-of-private-correspondence-channels/> **(**accessed 12-11-2021**).** On the linked but similar issue of the use of private messaging systems, see: Joey Senat, 'Whose Business Is It: Is Public Business Conducted on Officials' Personal Electronic Devices Subject to State Open Records Laws? (2014) 19 Communication Law and Policy 293.

[25] Richard Ovenden, 'Ephemeral messages remove scrutiny from government' (*The Financial Times*, October 13 2020) <https://www.ft.com/content/d7f10eb2-895e-4a7e-9522-2d365e7a205b> accessed 02-12-2021. See also: Richard Ovenden, *Burning the Books: A History of Knowledge Under Attack* (John Murray Press 2021).

There are clear grounds for suspecting that there is variable compliance, without good reason, with existing policies on record keeping.

Second, some approaches to adjusting guidance in response to the use of self-deleting instant messaging may call into question the legality of the content of these changes. The most prominent example of this risk thus far is the Cabinet Office policy which mandates the use of self-deleting messages.[26] That policy explains that "instant messaging is provided to all staff and should be used in preference to email for routine communications where there is no need to retain a record of the communication." The policy explains that, where possible, "instant messages history in individual and group chats must be switched off and should not be retained once a session is finished." The policy also recognises that the "[c]ontents of instant messaging are subject to FOI and Data Protection searches and the Public Records Act." On retention, however, the recommended practice is practically onerous and cumbersome:

> If the content of an instant message is required for the record or as an audit trail, a note for the record should be created and the message content saved in that. For example, written up in an email or in a document created in a word processor which is itself saved into the relevant drive.

As set out above, the likely correct approach to determine the legality of this guidance is to have regard to "evidence regarding its likely impact" and then undertake "an overall evaluative assessment." While there is no evidence in the public domain as to the extent this procedure is used, it is highly likely that busy officials—particularly senior officials and ministers who are likely to be central to the most important government decisions and be extremely short on time—will find this process impractical on a routine basis. An important question is therefore

[26] Cabinet Office, *Information and Records Retention & Destruction Policy* (Undated).

whether the kind of adjustment the Cabinet Office policy has included can withstand close legal scrutiny, and there is the certainly the potential that it may not do. There could well be similar provisions in guidance, not yet in the public domain, elsewhere in government.

Third, in the instance where extant policies of specific public bodies have not been adjusted and there is evidence of messages being automatically deleted, there is a question about whether those policies have lapsed into being so unclear that they are insufficient to fulfil legal obligations. It is possible that the lack of particular guidance vis-à-vis automatic deletion could be deemed to be an unlawful flaw within the policy, in that it permits the automatic deletion of messages before such time that consideration of preservation can be considered, as is required under the Public Records Act, the Freedom of Information Act, and potentially the duty of candour.

The overall picture of the extent to which each of these issues of legality arise across government remains hazy but, based on the limited amount we know about internal government practice at present, there is an apparent and serious risk of the government lapsing into illegality when adopting self-deleting messaging technologies. The solution is relatively straightforward. Compliance with public law duties in this context does not require banning completely the use of self-deleting messaging systems, preserving every WhatsApp message sent or otherwise placing unduly onerous duties on public authorities. But it does mean ensuring that such platforms are used in a way that is sensitive to the state's duties to maintain a public record. Put simply, there needs to be a clear, effective, and enforced policy that are in line with legal frameworks and ensures that the public record is preserved. It seems that the government may still have some work to do to get its house in order in this respect.