**Royal United Services Institute**
for Defence and Security Studies

Occasional Paper

# Balancing End-to-End Encryption and Public Safety

Chamin Herath and Sneha Dawda

# Balancing End-to-End Encryption and Public Safety

Chamin Herath and Sneha Dawda

RUSI Occasional Paper, April 2022


**Royal United Services Institute**
for Defence and Security Studies

**191 years of independent thinking on defence and security**

The Royal United Services Institute (RUSI) is the world's oldest and the UK's leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today's complex challenges.

Since its foundation in 1831, RUSI has relied on its members to support its activities. Together with revenue from research, publications and conferences, RUSI has sustained its political independence for 191 years.



Published in 2022 by the Royal United Services Institute for Defence and Security Studies.




RUSI Occasional Paper, April 2022. ISSN 2397-0286 (Online).

**Royal United Services Institute**
for Defence and Security Studies
Whitehall
London SW1A 2ET
United Kingdom
+44 (0)20 7747 2600
www.rusi.org
RUSI is a registered charity (No. 210639)

# Contents

# Acknowledgements

The authors are grateful to the UK Mission to the EU for funding this research and for their support throughout the research process.

A great deal of thanks must go to the comprehensive RUSI team that helped to guide and shape this paper. Conrad Prince, James Sullivan, Peter Brorsen, Hugh Oberlander, Dina Mansour-Ille, Tom Sayner, Demi Starks and Zenab Hotelwala all provided valuable support and editing.

A huge thank you should also go to the four anonymous peer reviewers who provided vital feedback and advice in broaching such a polarised debate.

A final thank you to the participants of this research, to all those who very kindly gave up their time to participate in interviews in 2021 and 2022. All insights and perspectives have been crucial to this research and have contributed greatly to the authors' findings.

# Executive Summary

OVER THE LAST decade, there has been a significant debate around end-to-end encryption (E2EE) and its implications for public safety. At the forefront of the discourse is a false dichotomy between protecting privacy and ensuring national security. At the extreme ends of this deeply polarised debate are two key arguments. On the privacy side, it is believed that governments and law enforcement agencies desire unrestrained exceptional access to E2EE communications to spy on their citizens. On the security side, it is maintained that obtaining lawful exceptional access is the only way to protect citizens and uphold national security. The debate has reached a deadlock, with both sides perpetuating zero-sum views.

However, experts are calling for a more nuanced conversation about possible solutions to the criminal use of E2EE services. It is vital that a range of views are considered in order to identify the key issues and inform a more productive debate. Through a review of the existing literature and insights from 22 semi-structured interviews, this paper balances the perspectives from a range of relevant stakeholders on the main elements of the E2EE debate and presents some key takeaways in an effort to move away from a crude privacy-versus-security binary.

The paper presents the following key findings:

- There are clear and significant cyber security and privacy benefits to E2EE. Efforts to weaken or restrict its access would be a net loss for all.
- Criminal use of E2EE is a significant risk to public safety and solutions are vital. Yet, it should also be acknowledged that technology is an enabler of criminal and harmful activity and should not be treated as the root cause.
- The possibility of developing technical tools which could assist law enforcement investigations should not be categorically ruled out, but future proposals must be measured against the principles of proportionality, legality and technical robustness.
- Alternative options for law enforcement investigations such as metadata analysis and legal hacking should be considered, but they are not without their drawbacks. Legal hacking could be proportionate but its reliance on software vulnerabilities is largely at odds with strong cyber security. Metadata analysis is promising but more research is needed to determine the extent to which it can be used to aid law enforcement investigations.
- Industry do have a responsibility to make their platforms safer and free from criminal abuse. This requires implementation of safety-by-design principles and the provision of resources for better digital literacy and education. Governments must have oversight over the technical tools developed.
- A more nuanced debate must continue which actively moves away from zero-sum views of absolute privacy versus absolute security, and focuses more on how the risks to public safety can be reduced in proportion with the need to protect citizens' rights and freedoms.

# Introduction

THE DISCOURSE AROUND end-to-end encryption (E2EE) and public safety has reached a stalemate. Absolutist views around protecting privacy versus ensuring national security have polarised the discourse, emphasising two seemingly antagonistic perspectives.

On the 'privacy side', technologists, academics and privacy advocates have stressed the importance of E2EE for safeguarding against a continually evolving cyber threat landscape, mass data harvesting and surveillance by governments and technology companies, as well as for protecting vulnerable and marginalised groups.[1]

On the 'security side', policymakers and law enforcement agencies identify E2EE as a significant challenge to their ability to combat criminal activity, both online and offline.[2] While E2EE applications have been used for several types of criminal activity, including organised crime, arms trafficking, human trafficking and narcotics, policymakers have largely focused on two issues in particular: the use of E2EE by terrorists and violent extremists; and its implication in child sexual abuse online. The latter has dominated recent conversation, and as reports of child sexual abuse online have been steadily increasing over the last decade, policymakers are under significant pressure to develop solutions.[3]

The extent to which a consensus can be reached depends on whether solutions can be developed which can combat criminal activity on E2EE platforms while also ensuring that the cyber security and privacy of civilians is fully protected. Currently, proposed solutions have focused on methods for allowing law enforcement exceptional access to E2EE communications and ensuring E2EE service providers can access data to effectively monitor criminal and harmful activities on their platforms.[4] Both have caused a significant amount of controversy and recent efforts from policymakers to sway public opinion on obtaining access to E2EE data has stirred further controversy and division.[5]

---

1. Internet Society, 'Policy Brief: Encryption', 9 June 2016.
2. Dan Milmo, 'NCA Says End-to-End Encryption Poses Challenge for Law Enforcers on Child Abuse', *The Guardian*, 22 January 2022; James Lewis, Denise Zheng and William Carter, *The Effect of Encryption on Lawful Access to Communications and Data* (Washington, DC: CSIS, 2017).
3. Internet Watch Foundation (IWF), '2020 Trends and Data', <https://annualreport2020.iwf.org.uk/trends>, accessed 10 March 2022.
4. Ian Levy and Crispin Robinson, 'Principles for a More Informed Exceptional Access Debate', *Lawfare*, 29 November 2018; Home Office, 'International Statement: End-to-End Encryption and Public Safety', 11 October 2020, <https://www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety>, accessed 10 March 2022.
5. Dan Milmo, 'Campaign Aims to Stop Facebook Encryption Plans Over Child Abuse Fears', *The Guardian*, 18 January 2022; James Ball, 'Revealed: UK Gov't Plans Publicity Blitz to Undermine

As the debate continues, there have been several calls for a less polarised and more constructive debate.[6] Experts have highlighted the need to find common ground, whereby different views and perspectives from the relevant stakeholder groups are brought together and considered.[7] In accordance with calls for a more balanced and collaborative discussion, this paper collates and outlines different perspectives on the benefits and public safety risks of E2EE.

## Methodology

This paper has been written as part of a four-month independent research project, funded by the UK Mission to the EU. The primary aim of the project was to identify and highlight salient perspectives on the E2EE debate to better inform policymakers, industry and non-technical stakeholders within the UK and the EU. The project aimed to answer the following research question: what are the benefits and the public safety risks of E2EE communications?

In addressing this question, the project inevitably covered potential solutions for countering criminal use of E2EE. This includes proposals for lawful exceptional access to E2EE data, automated on-device scanning of content, metadata analysis and legal hacking.

The data was collected in two ways:

- **Systematised literature review**: The authors conducted a targeted literature review of publicly available sources on E2EE from September 2021 to January 2022. Sources were identified using a selection of keyword search strings on Google, Google Scholar and EBSCO and ranged from journal articles, books and book chapters to legislation and 'grey literature', such as think tank reports and policy documents. While this paper is largely framed within the UK and EU context, this was not strictly applied to the literature review to account for key pieces of research on the discourse as a whole.
- **Semi-structured interviews**: A total of 22 semi-structured interviews were conducted between October 2021 and January 2022 to generate primary source data with respect to the different perspectives on the topic of E2EE. Non-probabilistic purposive sampling was used to identify interviewees with a range of views and from different stakeholder groups, including academia, civil society, government and industry.[8] Interviewees were selected based on their subject matter expertise and experience in the field of encryption and encryption policy. Initial interviewees were identified based on their role and expertise within the UK and EU context, but snowballing identified interviewees outside the European context, specifically from the US.

---

Privacy of Your Chats', *Rolling Stone*, 16 January 2022.

6. Dennis McDonough and Susan Landau, 'Breaking the Encryption Impasse', Carnegie Endowment for International Peace, 16 January 2020.
7. Encryption Working Group, 'Moving the Encryption Policy Conversation Forward', Carnegie Endowment for International Peace, 10 September 2019.
8. The authors interviewed seven government officials, six civil society representatives, four law enforcement professionals, three industry representatives and two academics.

In terms of limitations, there was not an even geographic distribution of interviewees. All interviewees were from Europe and the US. Additionally, while there was explicit effort to get insights from a broad range of stakeholders, the overall sample size of interviewees was small, which could likewise limit the generalisability of the project's findings.

**Structure**

This paper is divided into two chapters. Chapter I presents key concepts pertaining to encryption. It then outlines perspectives on the cyber security and privacy benefits of E2EE and public safety risks of E2EE communications, focusing primarily on its use for child sexual exploitation and abuse (CSEA) and terrorism and violent extremism. Chapter II highlights perspectives on lawful exceptional access, on-device scanning for the purposes of moderating illegal and harmful content, and alternative options of law enforcement investigations. Following this, it examines perspectives on the role and responsibility of industry to make platforms safer. To conclude, the paper summarises findings and looks ahead at potential next steps towards a better discussion around possible technical and policy solutions.

# I. An Overview of End-to-End Encryption

THIS CHAPTER UNPACKS E2EE by outlining how encryption works. Although there is a significant amount of technical literature on the subject, it is beyond the scope of this paper to provide an exhaustive overview of encryption and its applications. As such, it highlights the differences between symmetric and asymmetric encryption, and differentiates between encryption of data at rest and data in transit. Following this, it examines the significant cyber security and privacy benefits afforded by E2EE. It highlights perspectives on the public safety risks of the criminal exploitation of E2EE platforms and services, with a particular focus on child sexual abuse material (CSAM) and terrorism and violent extremism.

## Key Terms, Concepts and Applications

The use of encryption predates computers by at least 2,000 years. Notable applications are Caesar's cipher in Ancient Rome, the Vigenère cipher in the American Civil War and the Enigma machine in the Second World War.[9] In today's digitised environment, its aim is 'to transform a plaintext message (or stored data) into a ciphertext in such a way that the ciphertext reveals little or no information about the original plaintext'.[10] There are three algorithms involved in the development of a modern encryption scheme:

1. **Key-generation algorithm**: This generates the key or keys that will be used to encrypt and/or decrypt data.
2. **Encryption algorithm**: This performs the encryption of data by encoding plaintext or an image into ciphertext.
3. **Decryption algorithm**: This performs the decryption of data by turning the ciphertext back into plaintext.[11]

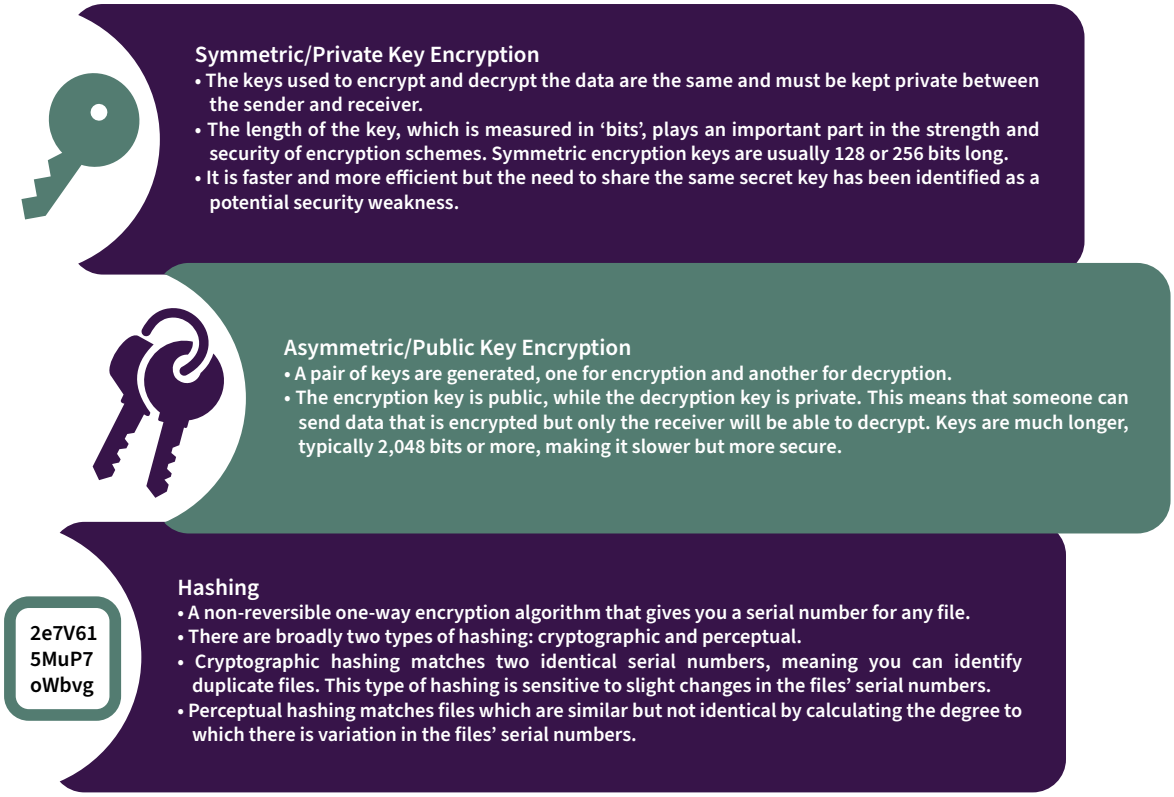
### Types of Encryption

Figure 1 provides a top-level breakdown of the two main types of encryption schemes, as well as hashing, which is not a type of encryption but is often discussed within the context of this debate.

---

9. Jean-Philippe Aumasson, *Serious Cryptography: A Practical Introduction to Modern Encryption* (San Francisco, CA: No Starch Press, 2017), p. 2.
10. National Academies of Sciences, Engineering, and Medicine, *Decrypting the Encryption Debate: A Framework for Decision Makers* (Washington, DC: National Academies Press, 2018).
11. Aumasson, *Serious Cryptography*, p. 2.

**Figure 1:** Encryption Schemes and Hashing




*Sources: National Academies of Sciences, Engineering, and Medicine,* Decrypting the Encryption Debate: A Framework for Decision Makers *(Washington, DC: National Academies Press, 2018); Martin E Hellman, 'An Overview of Public Key Cryptography', originally published in* IEEE Communications Magazine *(Vol. 16, No. 6, November 1978), <https://ee.stanford.edu/~hellman/publications/73.pdf>, accessed 7 March 2022; Vincent Lozupone, 'Analyze Encryption and Public Key Infrastructure (PKI)',* International Journal of Information Management *(Vol. 38, No. 1, 2018), pp. 42–44; Bian Yang, Fan Gu and Xiamu Niu, 'Block Mean Value Based Image Perceptual Hashing', 2006 International Conference on Intelligent Information Hiding and Multimedia (2006), pp. 167–72.*

## Data at Rest Versus Data in Transit

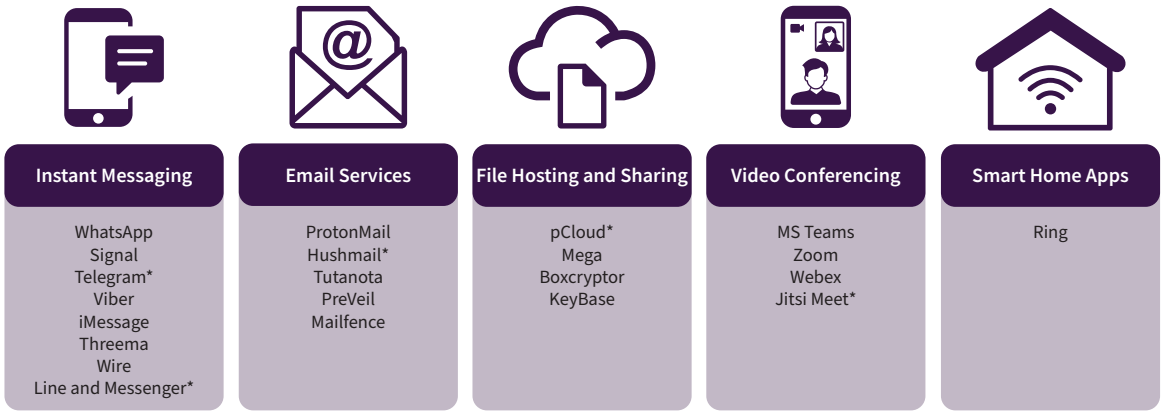
Symmetric encryption, asymmetric encryption or a combination of both can be used to secure two types of data.[12] Encryption of data at rest refers to securing information that is being stored on a computer hard-drive, server or mobile device. Examples include stored file and full-disk encryption and involve device locking using passwords, pin codes, biometrics and Face ID.[13] In contrast, encryption of data in transit refers to securing information that is moving from one place to another,

12. Surveillance Self-Defense, 'What Should I Know About Encryption?', last reviewed 24 November 2018, <https://ssd.eff.org/en/module/what-should-i-know-about-encryption>, accessed 10 March 2022.
13. National Academies of Sciences, Engineering, and Medicine, *Decrypting the Encryption Debate*.

such as web browsing, instant messaging, emailing and file sharing.[14] Transport Layer Security (TLS) is one commonly used application of encryption for data in transit. TLS encrypts data going from a user to a server. Most web browsers use a version of TLS and while the encryption can be end-to-end between the user and the server, it requires the service provider to decrypt the data at the server end, allowing the service provider to store the unencrypted data.[15]

E2EE is also used for data in transit. While TLS can also provide encryption of data from the client-end to the server-end, within the context of the debate, E2EE refers to zero-access encryption, whereby the service provider only transfers the data between one end-user and another and cannot decrypt the data at any point. For this reason, E2EE is the most secure application of encryption for textual, video and audio communication. As highlighted in a recent report by Tech Against Terrorism, four of the most popular messaging applications implement E2EE by default or as an optional feature.[16] However, as one interviewee suggested, 'details really matter' when thinking about how E2EE is applied platform by platform. While the goal of securing data remains the same, the E2EE applied on, for example, iMessage will differ to that of WhatsApp and other instant messaging platforms.[17]

**Figure 2:** Types of E2EE Platforms




*Source: Tech Against Terrorism, 'Terrorist Use of E2EE: State of Play, Misconceptions, and Mitigation Strategies', 2021, p. 14. This is just a sample of apps which use E2EE. Some of these platforms offer E2EE by default while others provide it as an add-on. Additionally, these categories are not distinct from one another, for example, WhatsApp can be used for messaging, file sharing and video calling.*

14. *Ibid*.
15. Internet Society, 'TLS Basics', <https://www.internetsociety.org/deploy360/tls/basics/>, accessed 10 March 2022.
16. Tech Against Terrorism, 'Terrorist Use of E2EE: State of Play, Misconceptions, and Mitigation Strategies', 2021.
17. Authors' interview with government official 1, 24 November 2021.

## Strong Encryption

Although it is not a distinct application of encryption, 'strong encryption' is a term that is often mentioned within the discourse around E2EE. In its simplest form, strong encryption can be defined as 'an encryption algorithm which cannot be broken within a time frame that would enable the breaker to take advantage of the information that has been encrypted'.[18] There is no single application of strong encryption. As highlighted by one interviewee, it 'can be applied to secure communications, either hop-by-hop and end-to-end, and it can be applied to secure data at rest'.[19] Yet, three high-level characteristics of strong encryption have been identified:

- Use of appropriate and robust algorithms.
- Use of long encryption and/or decryption keys.
- No accidental or deliberate vulnerabilities that could be exploited.[20]

Several governments have expressed their support for strong encryption. Table 1 provides a summary of the statements that have been made.

**Table 1:** Statements of Support for Strong Encryption

| | |
|---|---|
| **UK, US, Australia, New Zealand, Canada, Japan and India** | 'We, the undersigned, support strong encryption, which plays a crucial role in protecting personal data, privacy, intellectual property, trade secrets and cyber security. It also serves a vital purpose in repressive states to protect journalists, human rights defenders and other vulnerable people, as stated in the 2017 resolution of the UN Human Rights Council. Encryption is an existential anchor of trust in the digital world, and we do not support counterproductive and dangerous approaches that would materially weaken or limit security systems'. |
| **EU** | 'The European Union continues to support strong encryption. Encryption is an anchor of confidence in digitalisation and in protection of fundamental rights and should be promoted and developed'. |

*Sources: Home Office, 'International Statement: End-to-End Encryption and Public Safety', 11 October 2020, <https://www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety/international-statement-end-to-end-encryption-and-public-safety-accessible-version#fn:1>, accessed 7 March 2022; Council of the EU, 'Council Resolution on Encryption', 24 November 2020.*

However, responses from interviewees indicated that there are two different perspectives on how the term 'strong encryption' can be used. Several civil society interviewees suggested that to support the use of strong encryption is to support the application of E2EE, which they

18. Darrel Ince (ed.), *A Dictionary of the Internet*, 4th edition (Oxford: Oxford University Press, 2019).
19. Authors' interview with civil society representative 3, 18 November 2021.
20. National Academies of Science, Engineering, and Medicine, *Decrypting the Encryption Debate*.

argue is the most secure application of encryption.[21] Yet, this differs from the government and policymaking perspective. As highlighted in UK government testimony on encryption and lawful access, strong encryption does not refer to a particular application of encryption which is strong, such as E2EE, but a standard of encryption which also 'retains a technical capability to access the content of communications that are already encrypted over that service'.[22] As the term has become integral to governments' perspective on E2EE policy, greater clarity around the meaning of strong encryption, through a more detailed definition, is paramount for a more nuanced and balanced discussion.

## Benefits of E2EE

The benefits of E2EE are significant. Several interviewees highlighted its role as a critical enabler of a secure and resilient society and a hallmark of a liberal democratic approach to the rights and freedoms of citizens online.[23]

### Strong Cyber Security

> *'It's the protection of our digital society … and the protection of our infrastructure'.*[24]

E2EE is the cornerstone of strong cyber security. From securing the communications of public servants to retaining confidence, for example, between doctor and patient through video link, society relies on E2EE to provide a baseline of safety. The core cyber security benefits of E2EE pertain to all applications of encryption. These are highlighted by Alfred J Menezes, Paul C van Oorschot and Scott A Vanstone in their book *Handbook of Applied Cryptography*,[25] and include:

- **Confidentiality**: Ensuring the privacy and security of data by making it unintelligible to those who are not authorised to read it.
- **Integrity**: Ensuring that the data cannot be altered through the insertion, substitution and deletion of content.
- **Authentication**: Ensuring that the sender and receiver can identify where data is coming from and going to (also called 'entity authentication'), and that they know the data

21. Authors' interview with civil society representative 1, 22 October 2021; authors' interview with civil society representative 2, 23 November 2021.
22. Home Office, 'Written Testimony of Chloe Squires, Director National Security, Home Office: Judiciary Committee United States Senate Hearing: Encryption and Lawful Access: Evaluating Benefits and Risks to Public Safety and Privacy: December 10, 2019', updated 23 December 2019.
23. Authors' interview with civil society representative 1, 22 October 2021; authors' interview with civil society representative 3, 18 November 2021; authors' interview with civil society representative 6, 17 January 2022; authors' interview with industry representative 1, 29 November 2021.
24. Authors' interview with civil society representative 3, 18 November 2021.
25. Alfred J Menezes, Paul C van Oorschot and Scott A Vanstone, *Handbook of Applied Cryptography* (Boca Raton, FL: CRC Press, 2020), p. 4.

has not been intercepted and altered by an unauthorised party (also called 'data origin authentication'). With E2EE, authentication is proved through digital signatures.[26]

- **Non-repudiation**: Linked to authentication, this ensures that data has been sent and received by the intended recipient. It is particularly important as it protects against claims, particularly within the context of financial fraud.

E2EE provides enhanced protection for data in transit beyond other applications of encryption. It is widely acknowledged in the information security field as the most secure form and, as a result, there have been calls for E2EE to become a security standard with a widespread rollout on all types of internet communication services.[27] As E2EE does not require decryption at any stage of data transmission, it ensures the confidentiality and integrity of data being sent from one end-user to another. This, in turn, provides strong protection against man-in-the-middle attacks.[28]

On a societal level, increased reliance on technology and the digitisation of critical sectors such as finance and energy mean that protecting the transfer of data has never been more important. Without proper cyber security, critical systems and devices would be vulnerable to malicious cyber activity.[29] Moreover, it has been widely acknowledged that making society more secure starts with protecting individual users online and encouraging greater cyber hygiene and awareness.[30] Regarding E2EE's role in maintaining society-wide security, one interviewee said:

> I think we need to remember that security is infrastructure. Security is the processes, the people, you can't just … build one thing or do one action and make things more secure. It's a holistic approach … taking away the possibility from everyday people and folks in power to have truly secure … communications online.[31]

26. Ilya Sukhodolskiy and Sergey Zapechnikov, 'Analysis of Secure Protocols and Authentication Methods for Messaging', *Procedia Computer Science* (Vol. 169, 2020), pp. 407–11; IBM, 'Digital Signature Overview', last updated 5 March 2021, <https://www.ibm.com/docs/en/b2badv-communication/1.0.0?topic=overview-digital-signature>, accessed 10 March 2022.
27. Alexa Cardenas, 'End-to-End Encryption Gains Momentum as New Security Standard', Venafi, 23 July 2021, <https://www.venafi.com/blog/end-end-encryption-gains-momentum-new-security-standard-encryption-digest-63>, accessed 15 March 2022.
28. Jeff Peters, 'What Is a Man-in-the-Middle Attack: Detection and Prevention Tips', Varonis, 10 August 2020; National Academies of Science, Engineering, and Medicine, *Decrypting the Encryption Debate*, p. 30.
29. Vikas Hassija et al., 'A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures', *IEEE Access* (Vol. 7, 2019).
30. HM Government, 'National Cyber Strategy 2022', updated 7 February 2022.
31. Authors' interview with civil society representative 2, 23 November 2021.

## Strong Privacy and Data Protection

*'People want to be able to have a private conversation. It's really no more straightforward than that'.*[32]

In addition to the cyber security benefits, E2EE protects the user's right to private communications and their personal data. The growth in popularity of E2EE was a direct result of the Snowden revelations in 2013, which sparked fears about global surveillance and the mass collection of personal data by intelligence agencies around the world.[33] Many have also highlighted the mass data breach by Facebook and subsequent exploitation by Cambridge Analytica in 2016 as an equally important turning point in emphasising the need for privacy-enhancing technologies.[34] As demonstrated by these two examples (both of which occurred in liberal democracies), citizens' right to privacy can be eroded if there is a lack of proper safeguards.[35]

E2EE is regarded as critical for protecting freedom of speech.[36] According to Freedom House's most recent report on freedom online, freedom of speech 'is under unprecedented strain' and it was reported that approximately 75% of all internet users were from countries where people had been imprisoned for posting online content of a political, religious or social nature.[37] In 2018, the UN special rapporteur on freedom of expression suggested that companies should provide E2EE, either by default or through an opt-in function, to fulfil their 'responsibility to safeguard freedom of expression'.[38] This is particularly important in the context of protecting vulnerable and marginalised groups, journalists and the free press from censorship by illiberal and authoritarian regimes.[39] In such regimes, the right to privacy for journalists, activists and opposition politicians protects their life. Governments have gone to extreme lengths to infringe upon this right, including buying and using spyware, which has sparked a global market for

32. *Ibid.*
33. Ciaran Martin, 'End-to-End Encryption: The (Fruitless?) Search for a Compromise', Blavatnik School of Government, University of Oxford, November 2021, <https://www.bsg.ox.ac.uk/research/publications/end-end-encryption-fruitless-search-compromise>, accessed 10 March 2022.
34. Steven Song, 'Keeping Private Messages Private: End-to-End Encryption on Social Media', IPTF, 29 April 2020, <https://bciptf.org/2020/04/keeping-private-messages-private/>, accessed 10 March 2022.
35. See Article 8 of the European Convention on Human Rights (1953) on the right to privacy, p. 11, <https://www.echr.coe.int/documents/convention_eng.pdf>, accessed 17 March 2022.
36. See Article 10 of the European Convention on Human Rights (1953) on the right to freedom of expression, p. 12, <https://www.echr.coe.int/documents/convention_eng.pdf>, accessed 17 March 2022.
37. Freedom House, 'Freedom on the Net 2021: The Global Drive to Control Big Tech', October 2021.
38. UN Office of the High Commissioner for Human Rights, 'Mandate of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression', June 2018.
39. Vulnerable and marginalised groups include LGBTQ+ people, members of ethnic minorities, and children, men and women who are the victims of domestic and sexual abuse. Journalists also rely on E2EE for securely communicating with sources about sensitive topics.

spyware companies,[40] with the Israeli NSO Group's Pegasus spyware implicated in the murder of Jamal Khashoggi.[41]

Additionally, many have emphasised the necessity of E2EE services to fully protect personal data. As highlighted by the UK's independent authority on data protection, the Information Commissioner's Office:

> Real-life circumstances where the lack of E2EE has exposed people to harm include: children having their pictures accessed or location tracked, access to medical data, collection of data for fraud and misuse, and the acquisition of sensitive data as part of broader data collection processes.[42]

It also ensures that companies are properly adhering to relevant data protection legislation.[43] In the UK and EU context, this pertains to the Data Protection Act and the General Data Protection Regulation respectively, the latter of which states that internet companies must 'protect [consumers] rights and freedoms … [and] implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk'.[44]

# Public Safety Risks of E2EE

> *'End-to-end encryption offers criminals the ability to communicate and … drive forward their agendas in a way that law enforcement and partners increasingly struggle to deal with'.*[45]

Despite its widespread benefits, several public safety risks associated with the use of E2EE have been highlighted, all of which pertain to the criminal exploitation of E2EE platforms and services. These risks are examined below, first by looking at criminal use of E2EE more broadly, then with a specific focus on two areas of particular concern: CSAM; and terrorism and violent extremism.

**Criminal Use of E2EE**

There are a host of platforms, both with and without E2EE, that are being exploited by criminals, but their use of different E2EE services is a recognised problem for law enforcement. The most

40. Joseph Menn and Joel Schectman, 'U.S. Lawmakers Call for Sanctions Against Israel's NSO, Other Spyware Firms', *Reuters*, 15 December 2021.
41. Corin Faife, 'New Analysis Further Links Pegasus Spyware to Jamal Khashoggi Murder', *The Verge*, 21 December 2021.
42. Information Commissioner's Office, 'A Framework for Analysing End to End Encryption in an Online Safety Context', 2 November 2021.
43. In the EU, this is set in accordance with the General Data Protection Regulation (GDPR) (2016). In the UK, this is set in accordance with the Data Protection Act (2018).
44. See GDPR, 'Article 32: Security of Processing'.
45. Authors' interview with law enforcement professional 3, 16 November 2021.

significant issue for law enforcement is that they cannot easily access the content that is being shared, because they cannot rely on service providers to give them this access.[46]

There are two broad types of E2EE services being used for criminal purposes: mainstream E2EE platforms; and criminally dedicated secure communications (CDSCs). As highlighted in a report by Europol, 'criminals are not only misusing [the] encryption possibilities of mainstream platforms but also using dedicated communication channels providing end-to-end encryption'.[47] Most notable of all known CDSCs was EncroChat, a communications provider which sold burner phones with an integrated E2EE messaging service. Although not strictly developed for criminal purposes, after a joint law enforcement assault on the service, it has been implicated in a significant amount of organised criminal activity.[48] However, according to one interviewee from law enforcement, the environment of E2EE platforms is becoming more complex: 'there are a plethora of platforms now. It's a fragmented market completely. You might have somebody who puts part of the message on a CDSC device and then moves to an over-the-top service like Signal or WhatsApp'.[49] In a recent trend, E2EE services have started merging with privacy-focused cryptocurrencies. For example, Signal are integrating MobileCoin into their platforms to allow for E2EE payments between users.[50] Private payments will inevitably enable some degree of criminality, which law enforcement officers will find difficult to investigate due to E2EE.

## Child Sexual Exploitation and Abuse Material

> *'There's no bad application of encryption … There's bad usage by several actors with specific services that are encrypted'.*[51]

In recent years, the discussion around the risks of E2EE has been focused on the spread of CSAM online. Several well-known child protection organisations such as the National Society for the Prevention of Cruelty to Children (NSPCC) in the UK, as well as WeProtect and the National Centre for Missing & Exploited Children (NCMEC) in the US, have published reports on the proliferation of this content online, all emphasising the harm of not being able to track CSEA in E2EE environments.[52] According to the NSPCC, 'should widespread E2EE be implemented as expected, it will have significant detrimental effects on children's online safety

46. Europol and Eurojust, 'Third Report of the Observatory Function on Encryption', June 2021.
47. *Ibid*., p. 9.
48. For more on EncroChat, see Europol and Eurojust, 'Third Report of the Observatory Function on Encryption', June 2021. Other examples of criminally dedicated secure communications platforms include Sky ECC, PhantomSecure, IronChat and Ennetcom, all of which have been taken down.
49. Authors' interview with law enforcement professional 1, 4 November 2021.
50. Andy Greenberg, 'Signal's Cryptocurrency Feature Has Gone Worldwide', *Wired*, 6 January 2022.
51. Authors' interview with civil society representative 6, 17 January 2022.
52. NSPCC, 'End-to-End Encryption: Understanding the Impacts for Child Safety Online', April 2021; NCMEC, 'NCMEC's Statement Regarding End-to-End Encryption', 3 October 2019; WeProtect Global Alliance, 'Global Threat Assessment 2021'.

and remove platforms' ability to proactively identify harm within direct communications'.[53] In non-E2EE environments, service providers can use content moderation tools such as PhotoDNA, which uses perceptual hashing to match images with other hashes of CSAM. However, because PhotoDNA and other perceptual hashing tools operate on the server side, they cannot be used in E2EE environments.[54]

The drive to address concerns over the use of E2EE in CSEA is motivated by recent figures which show that there has been a significant increase in reports of CSEA online. There has been a 15-fold increase in the amount of CSAM online compared to 2011.[55] In 2016, 90% of all child sexual abuse URLs were hosted in Europe (including Russia and Turkey), which was a shift from North America.[56] This has been steadily increasing for the last five years and has spurred responses from the UK, the EU and its member states.[57] Much of the recent conversation around CSAM has been focused on Meta and its plans to apply E2EE by default on its platforms Messenger and Instagram.[58] Several mainstream social media platforms make up the bulk of all reports concerning the sexual exploitation, abuse and enticement of children, with Meta (and its associated platforms) making over 20 million referrals in 2020 (out of a total of 21.4 million).[59] Moreover, with the sheer number of reports coming from Meta's applications, it has been argued that the rollout of E2EE on all of their communication channels would significantly impact the number of reports received by the NCMEC in the future.[60]

**Case Study 1: David Wilson, Messenger**

Although not an example of a perpetrator using E2EE, the case of serial sex offender and paedophile David Wilson has been used to outline the risks of applying default E2EE to Messenger. Wilson, a 36-year-old man, was found guilty of committing 96 sex offences against 52 children, all of whom were under the age of 15. While he operated across several social media platforms, such as Snapchat and Instagram, and used multiple prepaid phones for SMS messaging, he primarily used Facebook and Messenger to create fake accounts of teenage girls and approach his victims to communicate with them respectively. His offences ranged from blackmailing victims into sending explicit images to forcing them to commit sexually exploitative acts.

53. NSPCC, 'End-to-End Encryption', p. 10.
54. Hany Farid, 'An Overview of Perceptual Hashing', *Journal of Online Trust and Safety* (Vol. 1, No. 1, 2021).
55. Sarah Marsh, 'Fifteen Times More Child Sexual Abuse Material Found Online Than 10 Years Ago', *The Guardian*, 13 November 2021.
56. IWF, '2020 Trends and Data – Geographical Hosting', <https://annualreport2020.iwf.org.uk/trends/international/geographic>, accessed 10 March 2022.
57. Mar Negreiro, 'Curbing the Surge in Online Child Abuse: The Dual Role of Digital Technology in Fighting and Facilitating its Proliferation', Briefing, European Parliament, November 2020.
58. Meta also owns WhatsApp, which has applied E2EE automatically since 2016.
59. NCMEC, '2020 Reports by Electronic Service Providers (ESP)', 2021.
60. Matt Burgess, 'Police Caught One of the Web's Most Dangerous Paedophiles. Then Everything Went Dark', *Wired*, 12 May 2021.

Following his arrest and subsequent imprisonment in 2021, the NCA stated that Meta's ability to flag content on its platform was vital for building a case. If E2EE was to be applied by default on Messenger – without adequate provisions to identify and flag illegal content – it could allow offenders like Wilson to freely exploit the platform. In this regard, the case has been used to highlight the challenges of applying E2EE by default on instant messaging services, especially those linked with discovery social networking platforms.

*Source: Matt Burgess, 'Police Caught One of the World's Most Dangerous Paedophiles',* Wired*, 12 May 2021.*

**Case Study 2: Anonymous, WhatsApp**

While there are few detailed case studies which exemplify the extent of CSAM on E2EE platforms, anonymised cases do appear in research on the topic. For example, in research by Pieter Hartel and Rolf van Wegberg on the impact of E2EE on the outcomes of 3,241 publicly available court cases in the Netherlands, the authors identify one case where the perpetrator used WhatsApp to disseminate sexually exploitative images of a 15 year old on a group chat. This, combined with the fact that WhatsApp made over 400,000 reports to the NCMEC in 2020, indicates that cases like these are not isolated.

*Sources: Pieter Hartel and Rolf van Wegberg, 'Going Dark? Analysing the Impact of End-to-End Encryption on the Outcome of Dutch Criminal Court Cases', ArXiv, 13 April 2021; Will Cathcart, 'We've worked hard to ban and report people who traffic in it based on appropriate measures, like making it easy for people to report when it's shared. We reported more than 400,000 cases to NCMEC last year from @WhatsApp, all without breaking encryption', Twitter, <https://twitter.com/wcathcart/status/1423701475595755524>, 6 August 2021.*

## Terrorism and Violent Extremism

The use of E2EE platforms by terrorists and violent extremists has also been highlighted as a significant risk to public safety, and has been one of the central points in the discussion around lawful exceptional access to E2EE communications over the last five years. As highlighted by recent research, E2EE services are part of larger ecosystems of social media platforms, instant messaging apps and websites used by terrorists and violent extremists.[61] The privacy afforded by E2EE communications 'make[s] it a preferred feature for operational purposes, including internal communications and logistical coordination, as well as the sharing of material for online

61. Maura Conway, Ryan Scrivens and Logan Macnair, 'Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends', *ICCT Policy Brief* (October 2019); Stephane J Baele, Lewys Brace and Travis G Coan, 'Uncovering the Far-Right Online Ecosystem: An Analytical Framework and Research Agenda', *Studies in Conflict & Terrorism* (December 2020); Bennett Clifford and Helen Powell, 'Encrypted Extremism: Inside the English-Speaking Islamic State Ecosystem on Telegram', Program on Extremism, George Washington University, June 2019.

training'.[62] Platforms which offer E2EE such as Telegram, Element, Rocket.Chat and Hoop are particularly popular among terrorist and extremist online communities.[63] Additionally, recent research into the online behaviour of terrorists inspired by the Islamic State in the US found that the use of E2EE platforms has become more popular over time.[64]

**Case Study 3: Khalid Masood, WhatsApp and iMessage**

In March 2017, Khalid Masood, a 52-year-old man, conducted an Islamist-inspired terror attack in central London. Masood hired a vehicle and drove it across Westminster Bridge, mounting the pavement to hit oncoming pedestrians before exiting the vehicle and stabbing a police officer. He killed six people and injured 29.

While incident reports suggested that Masood had planned and acted alone, it was found that minutes before he had conducted the attack, he had sent a PDF document entitled 'Jihad' to a large number of his contacts on WhatsApp and iMessage, both of which were – and still are – E2EE by default. As such, the case has been used to exemplify the use of E2EE by a terrorist for sharing material which cannot be moderated by platforms and monitored by law enforcement.

*Sources: Max Hill,* The Westminster Bridge Terrorist Attack *(London: The Stationery Office, 2018); Chief Coroner, 'Inquest Arising From the Deaths in the Westminster Terror Attack of 22 March 2017', 2018;* BBC News*, 'WhatsApp Must Not Be a "Place for Terrorists to Hide"', 26 March 2017.*

62. Tech Against Terrorism, 'Terrorist Use of E2EE', p. 42.
63. *Ibid*.
64. Joe Whittaker, 'The Online Behaviours of Islamic State Terrorists in the United States', *Criminology and Public Policy* (Vol. 20, 2021), pp. 177–203.

# II. Countering Criminal Use of E2EE

CHAPTER I HIGHLIGHTED that there are clear and significant societal benefits from E2EE, but the overarching risk to public safety pertains to the use of E2EE platforms to commit and enable crime and cause harm. To explore how a balance can be struck between these benefits and risks, this chapter examines proposed solutions to counter the criminal exploitation of E2EE. First, it highlights perspectives on lawful exceptional access to E2EE and automated scanning on E2EE platforms. It then analyses alternative options for countering criminal exploitation of E2EE, industry's role in developing safe platforms and summarises key findings.

## Lawful Exceptional Access

Lawful exceptional access for the purposes of catching and prosecuting criminals has been deemed essential by governments and law enforcement agencies.[65] Recently, the UK government stated that they believe 'it is possible to develop a lawful, exceptional access solution which would not disproportionately increase cyber security risk or undermine individuals' privacy'.[66] While this has yet to be evidenced, there have been two proposed methods of exceptional access to E2EE communications that have been publicly discussed. First is key escrow, which requires social media platforms and telecommunications providers to 'retain a copy of keys necessary to decrypt information with a trusted third party who would turn over keys to law enforcement upon proper legal authorization'.[67] The most notorious example of a proposed key escrow system was the 'Clipper Chip' from the US government.[68]

However, experts have outlined a multitude of issues with key escrow which could compromise citizens' security and privacy. For instance, Harold Abelson and colleagues highlighted that several proposed key escrow systems rely on modifying E2EE to make it less secure for users.[69] They use the example of prohibiting the forward secrecy function used by many E2EE platforms, which ensures enhanced privacy by creating a new session key for each message or transfer of data.[70] Other concerns around key escrow relate to the exploitation of vulnerabilities in the software developed to grant access to a trusted escrow organisation by cyber criminals and

65. Council of the EU, 'Council Resolution on Encryption', 24 November 2020.
66. Home Office, 'Factsheet: Encryption', 5 November 2019, <https://homeofficemedia.blog.gov.uk/2019/11/05/factsheet-encryption/>, accessed 10 March 2022. See also Home Office, 'Written Testimony of Chloe Squires'.
67. Harold Abelson et al., 'Keys Under Doormats: Mandating Insecurity by Requiring Government Access to All Data and Communications', *Journal of Cybersecurity* (Vol. 1, No. 1, 2015), p. 70.
68. Harold Abelson et al., 'The Risks of Key Recovery, Key Escrow, and Trusted Third-Party Encryption', *World Wide Web Journal* (Vol. 2, No. 3, 1997), pp. 241–57.
69. Abelson et al., 'Keys Under Doormats'.
70. DongGook Park, Colin Boyd and Sang-Jae Moon, 'Forward Secrecy and its Application to Future Mobile Communications Security', *PKC 2000* (2000), pp. 433–45.

hostile states – particularly as the very knowledge that it is possible could attract the attention of those with the skills to exploit it.[71] There is also a worry that if these tools are developed and implemented, they will be copied and shared by illiberal and authoritarian states to suppress marginalised groups and compromise privacy and freedom of speech. For example, as one interviewee from civil society stated: 'digital technologies and digital knowledge has [sic] the tendency to be pervasive ... what is now in the knowledge and capabilities of a democratic nation will be the knowledge and the capabilities of an undemocratic nation in the future'.[72] Similarly, the question of who or which entities would be responsible for monitoring the secure and proportionate use of escrowed keys has also been raised, especially within the sharing of E2EE data between different countries.[73]

Second is the proposal to 'silently add a law enforcement participant to [an E2EE] group chat or call'.[74] Although the proposed solution would not break the encryption, critics still highlighted significant challenges to cyber security and privacy. These challenges echoed that of key escrow, namely the possible introduction of vulnerabilities, its potential for abuse and the modification of E2EE security protocols such as user authentication.[75] Yet, accompanying this proposal was a set of principles for a more informed debate around lawful access.[76] The suggested principles, which emphasised the importance of proportionality, transparency and legal oversight when developing exceptional access tools, have been praised as a good foundation for future policy development.[77] As one interviewee stated: 'I think the principles are great, but they have to be properly applied'.[78]

Some have proposed possible mitigations to the risks of lawful access. 'The lawful intercept system and lawful acquisition systems are quite grand structures which are set up ... and we've got lots of oversight from judges', suggested one interviewee.[79] However, judicial oversight is not the sole answer to providing adequate and effective safeguards against abuse.[80] One

71. *Ibid*.
72. Authors' interview with civil society representative 3, 18 November 2021.
73. Abelson et al., 'Keys Under Doormats'.
74. Levy and Robinson, 'Principles for a More Informed Exceptional Access Debate'.
75. Sharon Bradford Franklin and Andi Wilson Thompson, 'Open Letter to GCHQ on the Threats Posed by the Ghost Proposal', *Lawfare*, 30 May 2019; Internet Society, 'Fact Sheet: Ghost Proposals', 2020.
76. Levy and Robinson, 'Principles for a More Informed Exceptional Access Debate'.
77. Susan Landau, 'Exceptional Access: The Devil is in the Details', *Lawfare*, 26 December 2018.
78. Authors' interview with academic 2, 17 December 2021.
79. Authors' interview with law enforcement professional 3, 16 November 2021. See also Home Office, 'Interception of Communications: Draft Code of Practice', February 2017; Council of the EU, 'Lawful Interception – Strengthening EU Cooperation', 13 October 2020.
80. Kathryn Wilson, 'The European Convention on Human Rights and the Investigatory Powers Tribunal: Rationalising a Law Unto Itself?', *Trinity College Law Review* (Vol. 23, 2020); Paul F Scott, 'Hybrid Institutions in the National Security Constitution: The Case of the Commissioners', *Legal Studies* (Vol. 39, No. 3, 2019).

interviewee suggested that the framework outlined in the UK–US data access agreement could protect against the abuse of lawful access mechanisms by undemocratic states.[81] As a pursuant of the US Clarifying Lawful Overseas Use of Data (CLOUD) Act, the agreement allows for the access of data to enhance law enforcement investigations 'as long as both countries and each request meet a certain criteria'.[82] The agreement does not discuss encrypted data but it could provide a framework for ensuring encrypted data that has been lawfully accessed can only be provided to states which adhere to liberal and democratic values.[83] However, limitations to use of current CLOUD Act data access agreements have been highlighted in the literature. For instance, currently the only parties to the act are the US, the UK and Australia, meaning its scope is limited to those specific countries. Furthermore, it has been suggested that its application within an EU context could be difficult due to conflicting EU laws.[84] Nevertheless, as these limitations are specific to existing CLOUD Act data access agreements, these do not rule out its use as a framework for future agreements which could address the sharing of lawfully accessed encrypted data while also preventing access to authoritarian regimes.

## Automated On-Device Scanning

On-device scanning, which is developed and issued by the technology provider, has been identified as a potential tool to access and moderate content on E2EE platforms. It refers to the scanning of data at rest on a user's phone to detect harmful materials such as CSAM or terrorist-related content. As outlined by Abelson and colleagues:

> Instead of weakening encryption or providing law enforcement with backdoor keys to decrypt communications, [it] would enable on-device analysis of data in the clear. If targeted information were detected, its existence and, potentially, its source, would be revealed to the agencies; otherwise, little or no information would leave the client device.[85]

81. Authors' interview with government official 1, 24 November 2021; authors' interview with government official 2, 15 December 2021.
82. Home Office, 'Written Testimony of Chloe Squires', p. 13. For more on the UK–US data access agreement, see HM Government, *Agreement Between the Government of the United Kingdom of Great Britain and Northern Ireland and the Government of the United States of America on the Access to Electronic Data for the Purpose of Countering Serious Crime*, CP 178 (London: The Stationery Office, 2019). For more on the US CLOUD Act, see US Department of Justice, 'Promoting Public Safety, Privacy, and the Rule of Law Around the World: The Purpose and Impact of the CLOUD Act', white paper, April 2019.
83. Home Office, 'Written Testimony of Chloe Squires'.
84. Theodore Christakis, '21 Thoughts and Questions About the UK/US CLOUD Act Agreement: (and an Explanation of How it Works – with Charts)', European Law Blog, October 2019; Theodore Christakis and Fabien Terpan, 'EU–US Negotiations on Law Enforcement Access to Data: Divergences, Challenges and EU Law Procedures and Options', *International Data Privacy Law* (Vol. 11, No. 2, 2021), pp. 81–106.
85. Hal Abelson et al., 'Bugs in Our Pockets: The Risks of Client-Side Scanning', *ArXiv*, 15 October 2021.

Automated solutions for on-device scanning vary in method. Some solutions could seek to scan plaintext 'problematic messages' using pre-written classifiers on apps which the receiver of the message can flag as spam.[86] No data would be sent externally to a third party until the receiver has flagged the message. However, this would be reliant on self-policing and technical issues such as battery life and processing power.

Another example of automation in client-side scanning is image scanning, such as the automated client-side scanning tool NeuralHash, which was developed by Apple to detect CSAM uploads from users' phones without compromising E2EE.[87] NeuralHash uses perceptual hashing 'to detect a hash match without revealing what the image is or alerting the user'.[88] Apple's on-device scanning would scan to see if any image hashes on the device match the hashes of known CSAM. Apple then uses threshold secret sharing, which enables the decryption of the content only if the hashes match and it crosses a threshold of instances. After decryption of the material, Apple would manually review the case and refer to law enforcement agencies if it is deemed CSAM.

However, Apple received major backlash for its plans to deploy its on-device scanning tool, to the extent that dozens of major organisations and experts signed a joint letter to CEO Tim Cook denouncing such technology as rife for manipulation.[89] Many of the risks and concerns highlighted by technologists about Apple's tool can be applied to automated on-device scanning systems more broadly. Much like with the proposals for lawful exceptional access, one major risk is their exploitation by cyber criminals and hostile states for the purposes of mass surveillance. As pointed out by Paul Rosenzweig, system architecture could play a significant role. For example, if a scanning system is employed on a device's operating system, rather than a specific E2EE application, 'it would be more broadly invasive and would require greater permissions with deeper access to information on a device'.[90] If security vulnerabilities in the device's operating system were to be exploited, this would give an attacker the ability to scan any data on that device, both in transit and at rest. Moreover, a reliance on AI or automation to make decisions and verify images that infer an offence can be equally as dangerous. False positives are possible and would have a detrimental impact on those being falsely accused.

In the case of finding common ground for the assessment of such tools, interviewees highlighted three criteria to assess tools against. The first is **proportionality**. As one interviewee stated,

86. Seny Kamara et al., 'Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems', Center for Democracy & Technology, 2021, p. 26.
87. Paul Rosenzweig, 'The Apple Client-Side Scanning System', *Lawfare*, 24 August 2021; Zack Whittaker, 'Apple Confirms it Will Begin Scanning iCloud Photos for Child Abuse Images', *TechCrunch*, 5 August 2021.
88. For more information on perceptual hashing, see Figure 1. See also Whittaker, 'Apple Confirms it Will Begin Scanning iCloud Photos for Child Abuse Images'. For more information on Apple's NeuralHash, see Apple, 'CSAM Detection: Technical Summary', August 2021.
89. Electronic Frontier Foundation, 'Coalition Letter to Apple CEO Tim Cook', 19 August 2021.
90. Paul Rosenzweig, 'The Law and Policy of Client-Side Scanning', *Lawfare*, 20 August 2020.

'CSAM is a huge threat, but not everyone on the planet is exchanging [this content] every day … [so] would we take a platform with billions of users and make them less secure? That would be disproportionate. It's always this question of proportionality'.[91] Ultimately, is the intervention (on-device scanning) to prevent crime (CSAM) justified by the scale of the crime, or does the intervention potentially open up avenues for other crimes?

The second principle to assess against is the **legality** of the content. In the case of terrorist or 'harmful' content, the line is much blurrier than with CSAM.[92] One interviewee highlighted that often harmful or terrorist-related content is vital for the media to present the news and related facts around an event. The legality around the content may be in question because it may be shared across forums and online groups that venerate such activity, while equally being used in news sources as evidence of a harmful act.[93] Included within this principle of legality should be the ability to prove criminal intent. How could a potential tool or system identify criminal intent? As there are inconsistencies as to what is classified as a criminal act, are there effective provisions for a fair investigation? How would systems handle the issues around jurisdiction?

> [For terrorist content] there's no black and white in terms of what content is legal and what content is illegal … it differs from jurisdiction to jurisdiction. Filtering content for a grey area of what's legal and illegal is really difficult to do.[94]

The third principle is **robustness**, which determines the practical extent to which scanning can be open to security and privacy exploitation by malicious actors. The reality of such a tool is that illiberal governments and authoritarian regimes will seek to use it for surveillance. As one academic noted, 'Go and see if you can come up with a solution that actually works and doesn't cause mass surveillance and is equitable … If so, then let's see if it works at scale'.[95]

The key to any on-device scanning tool being deployed is that a conversation must happen in civil society before design and deployment. Governments and technology suppliers can do this in a number of ways: extensive consultation with a cross-section of society; in-depth user research through testing and focus groups; and transparent reports on plans for experts to feed back on.

Nevertheless, the need for open-mindedness on behalf of experts and technologists was also emphasised. One interviewee highlighted how the response to Apple's on-device scanning tool was unhelpful, and not conducive to the development of solutions to mitigate legitimate concerns and risks.[96]

91. Authors' interview with civil society representative 6, 17 January 2022.
92. Authors' interview with civil society representative 1, 22 October 2021.
93. Authors' interview with government official 1, 24 November 2021.
94. Authors' interview with civil society representative 1, 22 October 2021.
95. Authors' interview with academic 2, 17 December 2021.
96. Authors' interview with government official 1, 24 November 2021.

# Alternative Options for Law Enforcement Investigations

While lawful access and content moderation through on-device scanning are central to the suite of options law enforcement agencies may argue for, other options that avoid compromising E2EE are worth considering. One interviewee argued that 'anything that would target some specific platforms that are putting in place end-to-end encryption is not going to solve the problem' and went on to note 'if a criminal wants to commit a crime and wants to use an encrypted format, he or she would be able to do it. They would not need WhatsApp or Telegram or Facebook … they will use something else'.[97] The reality of the situation is there will likely always be E2EE communications, whether it is through mainstream service providers, criminally dedicated applications or simply through the Tor browser (which, as one interviewee highlighted, was created by the US government[98]). As a result, law enforcement agencies should consider what other investigative data-driven tools are at their disposal that can fill the gap of evidence needed.

## Metadata Analysis

'Metadata' is an umbrella term which refers to non-content data made up of information about a user's behaviours and communications online, 'such as sender and receiver identification, IP address, basic subscriber information, date, time, and location data'.[99] More specifically, there are two types of metadata that are generated from data in transit (see Table 2).

**Table 2:** Types of Metadata

| | |
|---|---|
| **Network-level metadata** | Summary information about which devices are communicating; time and duration of communication; where the devices are; and the volume of data in transit. |
| **Application-level metadata** | More detailed information about the sender and receiver of data. For example, in an email, this would be the email addresses that were used and the subject line. |

*Source: Sophie Stalla-Bourdillon, Evangelia Papadaki and Tim Chown, 'Metadata, Traffic Data, Communications Data, Service Use Information … What Is the Difference? Does the Difference Matter? An Interdisciplinary View from the UK', in Serge Gutwirth and Ronald Leenes (eds),* Data Protection on the Move *(Dordrecht: Springer, 2016).*

Relatively little is known about the metadata that is generated from E2EE communications, other than that it is unencrypted. This is partly attributed to a lack of transparency from service

97. Authors' interview with civil society representative 6, 17 January 2022.
98. Authors' interview with academic 2, 17 December 2021.
99. Authors' interview with civil society representative 2, 23 November 2021.

providers.[100] Nevertheless, under warrant, law enforcement can request that companies provide this data for the purpose of an investigation. However, perspectives on whether metadata offers an alternative option for law enforcement were largely mixed.

Proponents of metadata analysis suggest there is a lot for law enforcement to work with. As highlighted by one interviewee, 'you get a lot of information from metadata. You don't know what's in the message … but you can do a lot via the quantity of data and the synchronicity of data travelling over the internet, and you can draw a lot of conclusions from that as well'.[101] Although the exact types and range of metadata collected will vary by platform, experts have suggested that, in some cases, metadata is more important than obtaining content as it can provide context.[102] Moreover, despite the knowledge that the usefulness of metadata analysis can be varied, some believe it is a good compromise for better cyber security. 'I'm not saying that every investigation could be equally easily solved by utilising this data', suggested one interviewee, 'but I'm sure that many could be helped with this data and maybe it takes a little bit more effort. But it's possible. And I think this effort is worth it if it helps us protect something that makes our society more secure'.[103]

However, some describe the inability to see content as a significant obstacle for law enforcement. According to the NSPCC, the analysis of online behavioural patterns through metadata is 'highly limited in [its] ability to detect, assess and respond to harm. Regardless of other uses, metadata indicators would not solve the problem of adult offenders trading CSAM with each other'.[104] Furthermore, as one law enforcement interviewee stated: 'There is nothing like having content because content allows you to understand what's going on, whereas geolocation and metadata just allows you to understand where a device might be, and you might not even have the individual with the device'.[105]

Yet, the prevailing argument against relying solely on metadata is its inability to provide intent in a court of law. In this regard, one interviewee suggested: 'You can do all sorts of things with metadata, and you can get to a point that says there is probably something weird going on. But you can't present "probably something weird" to a judge and do anything, not without getting into some really nasty second-order effects'.[106]

100. Wafa Ben-Hassine and Anamitra Deb, 'Metadata from Encrypted Messages Can Keep People Safe', *Wired*, 12 November 2021.
101. Authors' interview with industry representative 3, 9 November 2021.
102. Tech Against Terrorism, 'Terrorist Use of E2EE'.
103. Authors' interview with civil society representative 5, 24 November 2021.
104. NSPCC, 'End-to-End Encryption'.
105. Authors' interview with law enforcement professional 1, 4 November 2021.
106. Authors' interview with government official 1, 24 November 2021.

While the benefits and challenges of using metadata have been debated, some have highlighted the need for greater transparency on behalf of law enforcement. As one interviewee posited: 'What is the number of cases where metadata has been insufficient? What is the size of the real problem? Give us data about that. You hardly ever see that data coming up'.[107] This emphasises a key point about building an evidence base around the usefulness of metadata from E2EE platforms, explicitly within the context of law enforcement investigations.

### Legal Hacking and Joint Law Enforcement Operations

Legal hacking, also called 'equipment interference', refers to the hacking of devices for the purposes of criminal investigations.[108] This typically involves the exploitation of software vulnerabilities to gain access to a device. It would not grant access to data in transit, but if the device is breached, E2EE communications could be intercepted.[109] Some have suggested that under strict oversight mechanisms, law enforcement should be able to gain access to encrypted data by their own capabilities without mandating an exceptional access tool. Regarding the use of legal hacking over a lawful exceptional access mechanism, one interviewee stated, 'the exploit of vulnerabilities for access to communication is fair game because it's not systematically undermining the architecture of the system … it's not systematic because you are using vulnerabilities to get in'.[110] Moreover, in countries including but not limited to the UK, Germany, France, Italy and the Netherlands, there is existing legislation that permits the use of legal hacking for criminal investigations and national security purposes.[111]

As mentioned in the previous chapter, there have been several successful joint law enforcement operations on criminally dedicated E2EE platforms, all of which used existing vulnerabilities to access data. Additionally, there have been cases where law enforcement used similarly innovative techniques to get around the lack of oversight from E2EE platforms. On the topic of the FBI-operated E2EE platform Anom, one interview stated: 'It was a very creative method to lure criminals to use what they thought was just your average encrypted service. And it wasn't breaking encryption. It was really thinking outside of the box and thinking across borders. And that's what we need more of'.[112] However, from a cyber security perspective, there are several challenges with relying entirely on software vulnerabilities. Despite the presence of strict frameworks for legal hacking, such as the UK's Equipment Interference Code

107. Authors' interview with civil society representative 3, 18 November 2021.
108. Home Office, 'Equipment Interference', March 2018.
109. Carlos Liguori, 'Exploring Lawful Hacking as a Possible Answer to the "Going Dark" Debate', *Michigan Technology Law Review* (Vol. 26, No. 2, 2020).
110. Authors' interview with civil society representative 3, 18 November 2021.
111. European Parliament, 'Legal Frameworks for Hacking by Law Enforcement: Identification, Evaluation and Comparison of Practices', 2017.
112. Authors' interview with civil society representative 1, 22 October 2021; Europol, '800 Criminals Arrested in Biggest Ever Law Enforcement Operation Against Encrypted Communication', 8 June 2021, <https://www.europol.europa.eu/media-press/newsroom/news/800-criminals-arrested-in-biggest-ever-law-enforcement-operation-against-encrypted-communication>, accessed 10 March 2022.

of Practice, interviewees pointed out that an over-reliance on exploiting vulnerabilities would be at odds with vulnerability disclosure processes.[113] This includes the processes where, if a software vulnerability is identified, it is reported to the service provider.[114] Furthermore, it has been highlighted that a reliance on exploiting software vulnerabilities could create and proliferate 'a shady marketplace of vulnerabilities and exploit chains that would be available to anyone with the cash'.[115] This has been exemplified by the widespread use of NSO Group's Pegasus software for surveillance, which harvested data from E2EE messaging services such as WhatsApp and iMessage.[116]

## Industry's Responsibility to Make Platforms Safer

As with any technological intervention that involves users, the private sector must be involved as the providers and operators of technology. Moreover, there has been significant emphasis across government and civil society that companies have a responsibility to combat criminals exploiting their platforms. As stated by one government interviewee: 'We're asking for platforms to be responsible in how they roll stuff out and helping government and law enforcement manage their public safety risk that they engender from the services they provide'.[117] Several governments have or are in the process of passing legislation which compels industry to deal with the spread of illegal and harmful content, particularly CSAM.[118] As one interviewee suggested, 'Companies have a responsibility to make sure that their platforms are a force for good in society'.[119]

However, interviewees suggested that the extent to which industry should be given autonomous decision-making power to develop solutions is important to consider, not just for the individual user, but also for government. In this regard, governments and law enforcement agencies must choose carefully what responsibility and power is handed to industry to investigate and gather data itself: 'governments will still have to work very closely with private sector tech companies … as long as there is a very robust legal framework for actually operating in this space'.[120] Furthermore,

113. Authors' interview with government official 1, 24 November 2021.
114. NCSC, 'Vulnerability Disclosure Toolkit', <https://www.ncsc.gov.uk/information/vulnerability-disclosure-toolkit>, accessed 23 March 2022.
115. Levy and Robinson, 'Principles for a More Informed Exceptional Access Debate'.
116. Stephanie Kirchgaessner, 'How NSO Became the Company Whose Software Can Spy on the World', *The Guardian*, 23 July 2021.
117. Authors' interview with government official 1, 24 November 2021.
118. Department for Digital, Culture, Media & Sport (DCMS), 'Draft Online Safety Bill', 12 May 2021, <https://www.gov.uk/government/publications/draft-online-safety-bill>, accessed 10 March 2022; Council of the EU, 'Encryption: Council Adopts Resolution on Security Through Encryption and Security Despite Encryption', press release, 14 December 2020, <https://www.consilium.europa.eu/en/press/press-releases/2020/12/14/encryption-council-adopts-resolution-on-security-through-encryption-and-security-despite-encryption/>, accessed 10 March 2022.
119. Authors' interview with civil society representative 3, 18 November 2021.
120. Authors' interview with civil society representative 1, 22 October 2021.

if companies are responsible for reducing criminal activity and illegal content online, what can they do to make their platforms safer and aid law enforcement efforts to catch criminals who abuse their platforms?

Many have highlighted the importance of developing and updating online spaces to have 'safety by design'.[121] In short, this means building platforms which prevent and reduce exposure to illegal and harmful content. Examples include building in age-verification software at the platform level and user reporting at the user or end-point level.[122] On the latter, one interviewee suggested user-empowered reporting to help law enforcement,[123] which involves giving a user the option to alert the platform of any unwanted content. User reporting in E2EE environments is called 'message franking' and is used by several platforms.[124] Within the context of terrorist use of E2EE, recent work by organisations such as the Global Internet Forum to Counter Terrorism and Tech Against Terrorism has provided an important step forward when thinking about industry collaboration and transparency when developing automated tools.[125] Moreover, interviewees also highlighted the wider role that industry can play in educating people about how to act safely online: 'companies could provide some resources to science education programmes that are teaching people how to interact with digital media, how to protect themselves from disinformation and manipulation'.[126]

However, it was still maintained by several government and law enforcement interviewees that companies themselves should have oversight of the content on their platforms to fulfil their duty of care to users. Although it remains to be seen how this can be achieved while protecting citizens' security and privacy, one potential route lies in continued innovation from industry and technologists. One example of this is homomorphic encryption, but this is currently not functional.[127]

121. DCMS, 'Principles of Safer Online Platform Design', 29 June 2021.
122. DCMS, 'The UK Safety Tech Sector: 2021 Analysis', last updated 28 May 2021.
123. Authors' interview with civil society representative 2, 23 November 2021.
124. Jonathan Mayer, 'Content Moderation for End-to-End Encrypted Messaging', 6 October 2019.
125. Global Internet Forum to Counter Terrorism (GIFCT), 'GIFCT Technical Approaches Working Group Gap: Analysis and Recommendations for Deploying Technical Solutions to Tackle the Terrorist Use of the Internet', July 2021; GIFCT, 'Transparency Recommendations for GIFCT', June 2021.
126. Authors' interview with civil society representative 6, 17 January 2022.
127. Authors' interview with civil society representative 3, 18 November 2021. For more information on homomorphic encryption, see Anastasios Arampatzis, 'Crime Detection: Preserving Privacy Through Homomorphic Encryption', Venafi, 4 February 2020, <https://www.venafi.com/blog/crime-detection-preserving-privacy-through-homomorphic-encryption>, accessed 15 March 2022.

# Conclusion

THIS PAPER HAS examined the benefits and public safety risks of E2EE. In doing so, it addressed proposed solutions to allow for access to E2EE data, examining perspectives from different stakeholder groups. Although it was beyond the scope of this paper to provide concrete solutions, it does highlight key takeaways for future discussion.

- **The benefits of E2EE are significant**. All interviewees acknowledged the significant cyber security and privacy benefits of E2EE. As outlined in the GDPR, companies are required to ensure user data is protected to the best of their ability. As the most secure application of encryption of data in transit, the implementation of E2EE ensures that this is fulfilled.
- **The public safety risks are a societal issue**. The use of E2EE as an enabler for criminal activity represents a significant challenge, not just for law enforcement but for society as a whole. Fears around CSEA are warranted and solutions to the spread of CSAM are necessary, but it will not address the root of the issue. Tackling this must be a collective effort from all stakeholders, not just policymakers.
- **A more nuanced discussion is needed**. Although the risks from criminal use do not mandate the weakening or restricting of E2EE, they do highlight the need to move away from the unhelpful zero-sum privacy-versus-security debate to a more nuanced discussion about how criminal use of E2EE can be addressed in proportion with the need to protect citizens' rights and freedoms.
- **The current proposed solutions are fraught with issues**. Existing proposals for lawful exceptional access have failed to account for the risks from poor design and implementation. Critics suggest that if these tools were employed at scale and without effective oversight mechanisms, they could easily be exploited by cyber criminals and authoritarian governments as tools for surveillance. A recent shift towards on-device scanning tools has moved the discussion away from intercepting data in transit to data at rest, but similar issues persist. Interviewees broadly agreed that any technical solutions must be guided by principles of proportionality, legality and robustness.
- **Weigh up the costs and benefits of future options**. The costs and benefits of proposed tools and solutions must be assessed against the principles of proportionality, legality and technical robustness. Alternative options such as metadata analysis and legal hacking are useful to consider when thinking about how to deal with the problem, yet both have their drawbacks – mainly the inability to prove criminal intent through metadata and the risk of relying on undisclosed vulnerabilities for legal hacking. As there is a mixed narrative around metadata, more access to data for research is vital to openly assess its utility in criminal investigations involving E2EE platforms.
- **Industry is responsible but oversight is key**. Interviewees from all stakeholder groups acknowledged that industry is pivotal for providing the tools and resources to address the issue, both through the implementation of safety-by-design principles and educating users on how to act online. While their role in developing potential solutions, technical

or otherwise, is likely to be key, governments and wider society must still have oversight of what they do.

Although collaboration in the online safety space already exists through forums such as the EU Internet Forum and the Safer Internet Forum, a consensus needs to be reached by policymakers, proponents of technical measures and those who argue that access to E2EE communications is not possible without severe implications for privacy. One way of doing this is to ensure the technical community is involved in the development and assessment of capabilities and tools that are being considered. Encouraging a stronger relationship through collaborative research with experts is key for building trust and ensuring tools are properly tested and critiqued before any technical measures are implemented.

# About the Authors

**Chamin Herath** is a Research Analyst in cyber threats and cyber security at RUSI. His research interests include cyber risks to critical sectors, policymaking around emerging technologies, online harms and violent extremists' use of the internet.

**Sneha Dawda** is a Research Fellow in RUSI's Cyber Security research programme. She specialises in national cyber security strategies, internet governance, critical national infrastructure vulnerabilities and cybercrime.