Open Letter from Security and Privacy Researchers in relation to the Online Safety Bill

We are a group of established researchers and scientists working in the fields of information security and cryptography. In this letter, we wish to highlight alarming misunderstandings and misconceptions around the Online Safety Bill and its interaction with the privacy and security technologies that our daily online interactions and communication rely on. We understand that this is a critical time for the Online Safety Bill, as it is being discussed in the House of Lords before being returned to the Commons this summer. In brief, our concern is that surveillance technologies are deployed in the spirit of providing online safety. This act undermines privacy guarantees and, indeed, safety online.

The nature of our work means that we are deeply invested in online safety, privacy, and security. Our role is to research and build practical technological solutions that underpin this safety, be they cryptographic protections to keep private messages private or other information security defence mechanisms to keep data and everything that depends on it out of reach from adversaries. These adversaries include, for example, bored teenagers, competitors, abusive intimate partners/family members, criminal enterprises, and nation states.

It is for this reason that we have welcomed the wide-spread adoption of end-to-end encryption over the last 10 years, a development, we recall, that was at least partially motivated by revelations of extensive digital surveillance by nation-state actors.

It is also for this reason that we are now alarmed by the proposal to technologically enable the routine monitoring of personal, business and civil society online communications. This monitoring is motivated and is being proposed to prevent the dissemination of child sexual exploitation and abuse (CSEA) content. We cannot speak to the relative merit of this step in preventing harm to children in our professional capacities. What we can confirm with authority is: Firstly, such monitoring is categorically incompatible with maintaining today's (and internationally adopted) online communication protocols that offer privacy guarantees similar to face-to-face conversations. Secondly, attempts to sidestep this contradiction are doomed to fail on the technological and likely societal level.

Technology is not a magic wand.

Access to protected private messages and images can be attempted in two ways: while in transit and protected by cryptography or before/after transit on the involved clients.

Cryptography. There is no technological solution to the contradiction inherent in both keeping information confidential from third parties and sharing that same information with third parties. Giving the State the technological means to access every private message and image implies that any actor with access to the relevant monitoring facilities will have the same access. Such actors include future governments with looser definitions of prohibited content, civil servants and police officers across different departments and forces, and any adversary who compromises the monitoring infrastructure. We note that such compromises are not merely an abstract possibility but eventualities to prepare for, when keeping in mind recent high-profile breaches at the national security level, e.g. of US and UK security services. The history of “no one but us” cryptographic backdoors is a history of failures, from the Clipper chip to DualEC. All technological solutions being put forward share that they give a third party access to private speech, messages and images under some criteria defined by that third party.

Client-side scanning. A popular deus ex machina is the idea to scan content on everybody’s devices before it is encrypted in transit. This would amount to placing a mandatory, always-on automatic wiretap in every device to scan for prohibited content. This idea of a “police officer in your pocket” has the immediate technological problem that it must both be able to accurately detect and reveal the targeted content and not detect and reveal content that is not targeted, even assuming a precise agreement on what ought to be targeted.

These proposals for client-side scanning come in two variants. One is to detect known images of abuse held in a database maintained by an authority. These technologies have been shown to have several issues. Foremost, research has shown that client-side scanning does not robustly achieve its primary objective, i.e. detect known prohibited content. Moreover, it has been recently shown that these algorithms can be repurposed to add hidden secondary capabilities (e.g. facial recognition of target individuals) to client-side scanning, covertly enabling surveillance.

Second, there are also more far-reaching proposals to mass-deploy AI models to scan messages for previously unseen but prohibited content relating to CSEA. However, sufficiently reliable solutions for detecting CSEA content do not exist. This lack of reliability here can have grave consequences as a false positive hit means potentially sharing private, intimate or sensitive messages or images with third parties, like private-company vetters, law enforcement and anyone with access to the monitoring infrastructure. This may in itself constitute exploitation and abuse of those whose messages are being disclosed. Finally, the more far-reaching task and implied necessary flexibility of such AI models also makes it easier to repurpose them and to expand their scope, by compromise or policy change.

We note that in the event of the Online Safety Bill passing and an Ofcom order being issued, several international communication providers indicated that they will refuse to comply with such an order to compromise the security and privacy of their customers and would leave the UK market. This would leave UK residents in a vulnerable situation, having to adopt compromised and weak solutions for online interactions.

As independent information security and cryptography researchers, we build technologies that keep people safe online. It is in this capacity that we see the need to stress that the safety provided by these essential technologies is now under threat in the Online Safety Bill.

Signatories

1. Dr Ruba Abu-Salma (King's College London)
2. Professor Martin Albrecht (King’s College London)
3. Dr Ihsen Alouani (Queen's University Belfast)
4. Dr Myrto Arapinis (University of Edinburgh)
5. Professor David Aspinall (University of Edinburgh)
6. Dr Rishiraj Bhattacharyya (University of Birmingham)
7. Dr Elizabeth Black (King's College London)
8. Professor Eerke Boiten (De Montfort University)
9. Professor Ioana Boureanu (University of Surrey)
10. Professor Ian Brown (Centre for Technology and Society, Fundação Getulio Vargas)
11. Professor William (Bill) Buchanan OBE (Edinburgh Napier University)
12. Professor Lorenzo Cavallaro (University College London)
13. Professor Liqun Chen (University of Surrey)
14. Dr Partha Das Chowdhury (University of Bristol)
15. Dr Michele Ciampi (University of Edinburgh)
16. Professor Jon Crowcroft (University of Cambridge)
17. Dr Santanu Dash (Royal Holloway, University of London)
18. Dr Benjamin Dowling (University of Sheffield)
19. Dr François Dupressoir (University of Bristol)
20. Dr Tariq Elahi (University of Edinburgh)
21. Professor Flavio Garcia (University of Birmingham)
22. Dr Essam Ghadafi (Newcastle University)
23. Professor Thomas Gross (Newcastle University)
24. Honorary Professor Jens Groth (University College London)
25. Professor Hamed Haddadi (Imperial College London)
26. Professor Julio Hernandez-Castro (University of Kent)
27. Dr Darren Hurley-Smith (Royal Holloway, University of London)
28. Dr Rikke Jensen (Royal Holloway, University of London)
29. Professor Adam Joinson (University of Bath)
30. Dr Saqib A. Kakvi (Royal Holloway, University of London)
31. Professor Elham Kashefi (University of Edinburgh)
32. Dr Narges Khakpour (Newcastle University)
33. Dr Markulf Kohlweiss (University of Edinburgh)
34. Professor Michael Levi (Cardiff University)
35. Professor Shujun Li (University of Kent)
36. Dr Nora Ni Loideain (Information Law & Policy Centre, IALS, University of London)
37. Dr Sergio Maffeis (Imperial College London)
38. Dr Bernardo Magri (University of Manchester)
39. Professor Keith Martin (Royal Holloway, University of London)
40. Dr Chloe Martindale (University of Bristol)
41. Professor Corinne May-Chahal (University of Lancaster)
42. Dr Maryam Mehrnezhad (Royal Holloway University of London)
43. Professor Sarah Meiklejohn (University College London)
44. Dr Charles Morisset (Newcastle University)
45. Professor Steven Murdoch (University College London)
46. Professor Sean Murphy (Royal Holloway, University of London)

47. Professor Shishir Nagaraja (Newcastle University)
48. Dr Daniel Page (University of Bristol)
49. Professor Kenneth Paterson (ETH Zürich, Switzerland)
50. Dr Paul Patras (University of Edinburgh)
51. Dr Claudia Peersman (University of Bristol)
52. Dr Christophe Petit (University of Birmingham)
53. Dr Fabio Pierazzi (King's College London)
54. Dr Elizabeth Quaglia (Royal Holloway, University of London)
55. Dr Ciara Rafferty (Queen's University Belfast)
56. Professor Awais Rashid (University of Bristol)
57. Professor Kasper Rasmussen (University of Oxford)
58. Professor Mark D. Ryan (University of Birmingham)
59. Professor Nishanth Sastry (University of Surrey)
60. Professor Steve Schneider (University of Surrey)
61. Dr Siamak Shahandashti (University of York)
62. Professor Nigel Smart (KU Leuven, Belgium)
63. Dr Nicholas Spooner (University of Warwick)
64. Professor Jose Such (King's College London)
65. Professor Luca Viganò (King's College London)
66. Dr Petros Wallden (University of Edinburgh)
67. Dr Christian Weinert (Royal Holloway, University of London)
68. Professor Alan Woodward (University of Surrey)

For inquiries please contact:
Professor Martin Albrecht <martin.albrecht@kcl.ac.uk>
Professor Hamed Haddadi <h.haddadi@imperial.ac.uk>