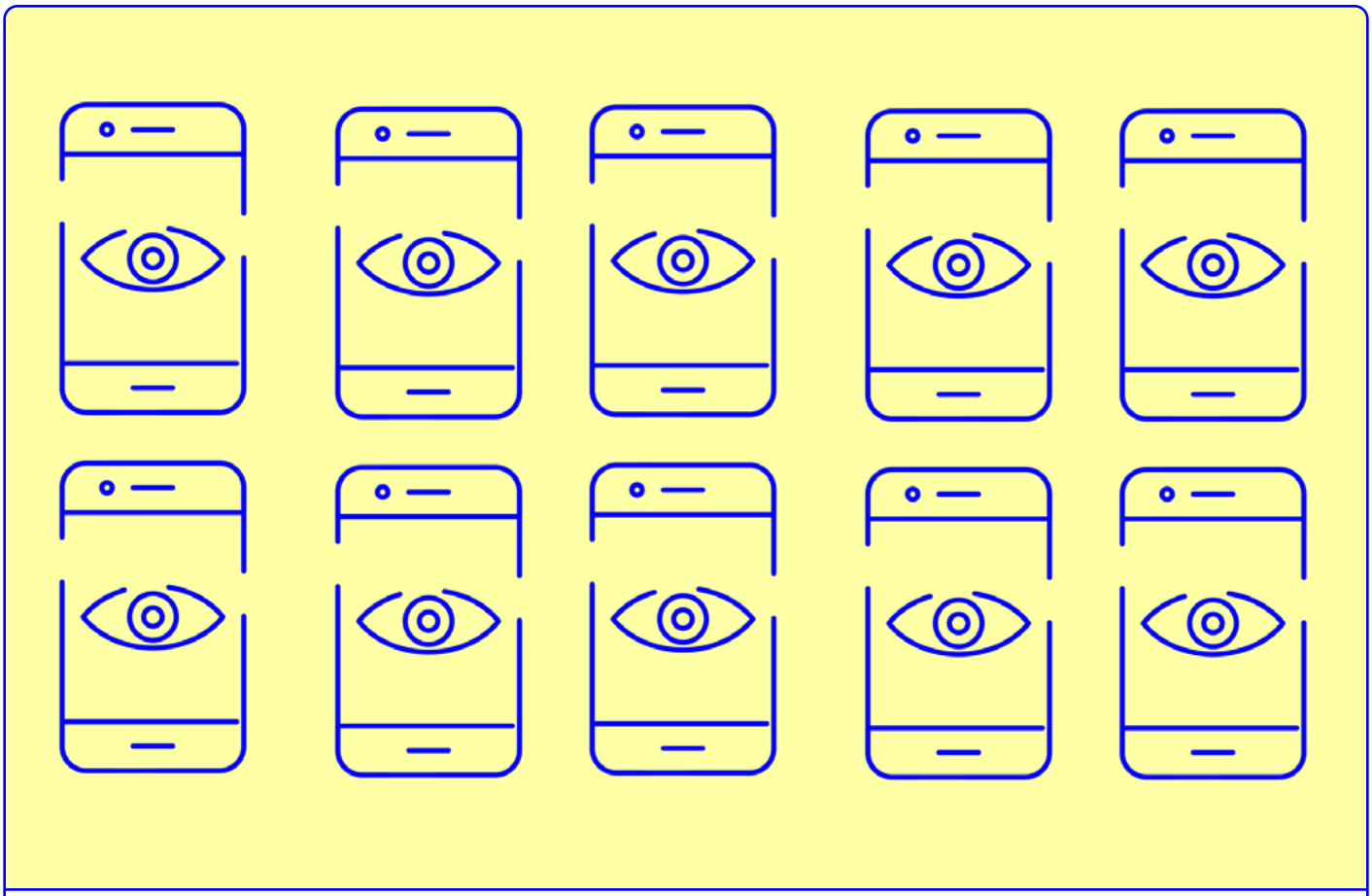
CONVOCATION
RESEARCH + DESIGN

Tech Policy Press



What is Secure?

Analysis of Popular Messaging Apps

JUSTIN HENDRIX, COOPER QUINTIN, CAROLINE SINDERS,
LEILA WYLIE WAGNER, TIM BERNARD, AMI MEHTA

/ JUNE 2023

Authors

JUSTIN HENDRIX is CEO and Editor of *Tech Policy Press*, a nonprofit media venture concerned with issues and ideas at the intersection of technology and democracy, and is an associate research scientist and adjunct professor at NYU Tandon School of Engineering. Previously, he was Executive Director of NYC Media Lab. He spent over a decade at *The Economist* in roles including Vice President, Business Development & Innovation. He holds a BA from the College of William & Mary and an MSc in Technology Commercialization from the University of Texas at Austin.

COOPER QUINTIN is a security researcher and senior public interest technologist with the EFF Threat Lab and Convocation. He is also a Citizen Lab fellow and board member of Open Archive. He has worked on projects including Privacy Badger, Canary Watch, and analysis of state sponsored malware campaigns such as Dark Caracal. Cooper has given talks about security research at prestigious security conferences including Black Hat, DEFCON, Enigma Conference, and ReCon, and has been published or quoted in publications including *The New York Times*, Reuters, NPR, CNN, and Al Jazeera. Cooper has provided security training for activists, nonprofit workers, and vulnerable populations around the world. He previously worked building websites for nonprofits, including Greenpeace, Adbusters, and the Chelsea Manning Support Network.

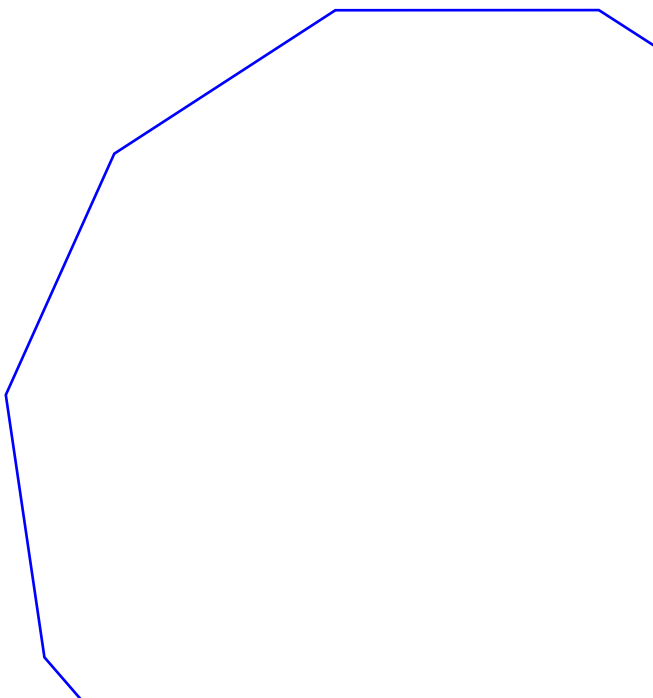
CAROLINE SINDERS is an award winning critical designer, researcher, and artist. She's the founder of human rights, design and technology research and advocacy organization, Convocation Research + Design. For the past few years, she has been examining the intersections of artificial intelligence, intersectional justice, systems design, harm, and politics in digital conversational spaces and technology platforms. She has worked with the United Nations, Amnesty International, IBM Watson, the Wikimedia Foundation, and others. Sinders has held fellowships with the Harvard Kennedy School, Google's PAIR (People and Artificial Intelligence Research group), Ars Electronica's AI Lab, the Weizenbaum Institute, the Mozilla Foundation, Pioneer Works, and others. Her work has been featured in the Tate Exchange in Tate Modern, the Contemporary Art Center of New Orleans, Telematic Media Arts, Victoria and Albert Museum, and MoMA PS1. Sinders holds a Masters from New York University's Interactive Telecommunications Program.

LEILA WYLIE WAGNER is a project management professional who is currently pursuing a Master's degree in Public Administration at Tulane University. Her areas of research include the intersection of technology and disaster response, climate justice in the Gulf South, and reproductive rights and public health. After several years working in disaster response

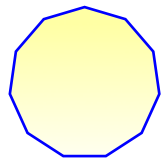
including managing COVID-19 vaccination programs and running logistics for Hurricane Ida response teams, Leila now works with Convocation Research + Design in addition to her day job as a healthcare IT analyst. Leila was born and raised in north Florida and is celebrating her 12th year living in New Orleans, Louisiana.

TIM BERNARD writes at *Tech Policy Press* on issues predominantly related to regulation and trust & safety. He recently completed an MBA at Cornell Tech, focusing on the impact of technology enterprises on society. Previously, he led the content moderation team at Seeking Alpha, and worked in various capacities in the education sector. His prior academic work includes an MA in Talmud and a BA in Philosophy.

AMI MEHTA (she/her) is a creative technologist, researcher, and artist based in Brooklyn. As a Postdoc Fellow at NYU's Interactive Telecommunications Program, Ami is exploring ethnographic approaches to XR design practices. Her research interests include issues related to safety, privacy, self-identity, and representation in online social spaces, virtual worlds, and video games. Ami is a member of the Extended Realities Track at NEW INC, an art, design, and technology incubator led by the New Museum. In the past, Ami has consulted on strategy, planning, and cultural policy for global nonprofit arts and media organizations. She holds a M.P.S. in Interactive Telecommunications from NYU, M.A. in Art History from The Courtauld, and B.A. in History from NYU.



Executive Summary



In a world where privacy and security are increasingly under threat, particularly in countries swept up in a global wave of autocratization and erosion of rights, encrypted messaging apps are an increasingly popular—and necessary—way to share information, organize and engage with one another, and do business. But while the promise of secure messaging is private communications and user control over the spread of personal or group information, the reality is often more complicated, particularly in the age of surveillance capitalism. An overlapping, interconnected set of engineering, design, and system factors, coupled with varied user behaviors and shifting policy environments, have created conditions in which individuals may subvert their own interests or those of their communities while using encrypted messaging apps.

From September 2022 through May 2023, we analyzed popular messaging apps–including Signal, WhatsApp, Telegram, Messages by Google, Apple Messages and Meta's Messenger –across a range of dimensions, including technical security, user experience, how the apps engage with users and developers, and their policies, terms and conditions. We sought to understand how people form mental models of their own individual or group digital security and corresponding threats, ways in which the technical and design decisions that the developers of encrypted messaging apps make can leave users vulnerable, and potential solutions that encompass technical, design, and policy adjustments.

To answer these questions, we adopted principles from frameworks such as Privacy by Design and Design from the Margins. We completed a technical review of selected apps, including a static analysis of source code and a dynamic analysis of network traffic; a detailed user experience and user interaction design analysis; and a comprehensive policy review. We interviewed a range of experts, and conducted field work with at-risk users including abortion rights activists in New Orleans, Louisiana and journalists in Delhi, India.

KEY FINDINGS AND RECOMMENDATIONS INCLUDE:

1. **Users are too often flying blind.** Even those most concerned about privacy rarely have sufficient information to make decisions that are in their own best interest. There is a substantial gap between the promise of encryption and the reality of threats to secure messaging in practice. We encountered various forms of "security folklore" that inform user decisions in place of information grounded in fact, as well as "security nihilism," a debilitating sense among some that there is no way to communicate securely.

2. **An app's cryptographic security doesn't mean it is secure.** Implementation is everything. The failure to implement end-to-end encryption by default, such as on Telegram and Meta's Messenger, illustrate this point. Users may not understand the distinction when presented with confusing options like "secret chat" and "private chat." And few users understand design distinctions, such as different colors for messages in Apple's iMessage and Messages by Google, that are intended to communicate different types of messages (SMS or encrypted,) and thus different levels of security.

3. **Follow Signal's lead and encrypt or don't store metadata.** Signal is the only app that has taken steps to hide users' profiles, contacts, group metadata, and even message sender information. Other developers need to follow Signal's example and hide user metadata by keeping it encrypted with the user's account key and only handling unencrypted versions in secure enclaves.

4. **Let users decide which features should be on or off.** Companies need to allow for any feature that impacts privacy and security to be turned on and off, and to explore and implement more granular settings that allow for users, especially high-risk users, to tailor the service to their needs, including when it comes to disappearing messages, link previews, storing and deleting call logs, and interaction history.

5. **Close technical and design 'loopholes' that betray privacy.** From unencrypted backups of messages and the use of phone numbers as identifiers to flaws in how deleted messages are handled, confusing naming conventions for certain features, and bad user design on some options, there are a range of technical and design issues that the makers of messaging apps need to address urgently.

6. **Beware the bloat.** Especially when it comes to apps that are connected to or are trying to emulate some aspects of social media platforms, including Meta's Messenger, Telegram and increasingly WhatsApp, there is evidence of feature bloat and connections to other apps and services that may create new privacy concerns. The incentives of surveillance capitalism are privacy and safety's worst enemy, particularly when developers deploy deceptive design patterns.

7. **Encryption must be defended.** Governments around the world–including in democracies–are threatening encryption with a range of new regulations and laws that will effectively break the model of apps like Signal and WhatsApp. It is crucial that policymakers, industry voices, and activists that understand the value of encryption speak up in its defense.

We find that the makers of encrypted messaging applications should universally adopt privacy-by-design, and invest more deliberately in working with users from more vulnerable, marginalized, and targeted communities. Encrypted messaging apps are not a magical shield against the most motivated adversaries, but they are a crucial line of defense for billions of users, and are worthy of close scrutiny.

Throughout this report, we refer to the colloquial names of these applications as they are commonly referred to by users, expert stakeholders, and referenced literature.

Messenger (Meta)
Launched as Facebook Chat, 2008; as a standalone app, 2011
Colloquially referred to as Facebook Messenger
~1 billion users
Meta HQ: Menlo Park, CA
~66,000 employees

WhatsApp (Meta)
Launched, 2009 (acquired by Meta, 2014)
~2 billion users
Meta HQ: Menlo Park, CA
~66,000 employees

Signal (Signal Foundation)
Launched, 2014
~40-100 million users
Signal Foundation HQ: Mountain View, CA
<50 employees

Telegram (Pavel and Nikolai Durov)
Launched, 2013
~700 million users
Telegram HQ: Dubai, UAE (parent shell company registered in Tortola, BVI)
<50 employees?

Messages by Google (Alphabet)
Launched as Android Messages, 2009; as a standalone app, 2014
>1 billion app installs
Google HQ: Mountain View, CA
~178,000 employees (Alphabet)

Apple Messages (Apple)
Launched, 2011
Colloquially referred to as iMessage
~1 billion users
Apple HQ: Cupertino, CA
~164,000 employees

Introduction

Around the world, people are increasingly using secure messaging apps—communications apps which completely, or partly, use end-to-end encryption—to share information, organize and engage with one another, and conduct commerce. But while the promise of secure messaging is private communications and user control over the spread of personal or group information, the reality is often more complicated. An overlapping and interconnected set of engineering, design, and system factors, coupled with varied user behaviors, create the conditions for individuals to betray their own interests or the interests of their communities on secure messaging apps.

As part of a program on trustworthy messaging funded by Omidyar Network, we set out to explore how design, technical, and policy choices on secure messaging apps, when combined with user behavior, and adversarial behavior by various parties, may produce malign effects.[1] Our research included interviews with more than 30 experts on privacy, security and technology; multiple individual and group sessions with users concerned about their own security, including groups concerned with abortion rights and access in Louisiana and journalists working in India; a technical analysis of selected secure messaging apps; a design and user experience review of the apps; a consideration of the overall policy environment related to encryption in key jurisdictions; and a review of the terms and conditions regarding secure and encrypted messaging in place at Signal, Meta, Apple, Telegram, and Google.

MOTIVATING THEMES

A range of social, political, and technological trends have combined in recent years to drive user adoption of encrypted messaging apps, and thus make scrutiny of their design and use important.

Technology has long been considered a threat to privacy. In "The Right to Privacy," a seminal *Harvard Law Review* article published in 1890, Samuel Warren and Louis Brandeis warned that "numerous mechanical devices threaten to make good the prediction that 'what is whispered in the closet shall be proclaimed from the house-tops.'"[2] Warren and Brandeis were, at the time, concerned with new communications technologies, such as the phonograph and the newspaper, that made it possible for what had hitherto been private to become public.

More than 130 years later, a range of technologies exist that threaten the privacy of communications between individuals, many far more invasive than anything Warren or Brandeis could have imagined. Most adults have access to mobile devices filled with a variety of sensors that can record and often broadcast simultaneously,[3] while even

1 https://omidyar.com/privatemessaging/

2 https://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html

3 https://www.statista.com/forecasts/1143723/smartphone-users-in-the-world

democratic governments have developed extraordinary tools for the real-time collection and analysis of electronic communications at scale, sometimes beyond constitutional or legal bounds.[4] [5] At the same time, people are more reliant than ever before on the use of mobile devices and applications to communicate. It is in this context that encrypted message apps have grown in popularity, in large part due to the promise that they allow individuals to communicate in text and other media modalities without exposing the information in their messages to third parties.

Of course, people have various motivations for preserving the privacy of their communications. Many are operating under one or more of the different conceptions of privacy identified by legal scholar Daniel Solove in his 2002 essay, "Conceptualizing Privacy," in *California Law Review*, including exercising control over personal information, protecting intimate communications, or simply preserving the right to be let alone.[6] Yet for some, the need to preserve private communication is more severe, including those who face persecution from their fellow citizens or from their government because of their identity, including race or ethnicity, sexuality or sexual orientation; their political preferences or activities; or their role, such as that of journalists and others who produce information. It is these vulnerable and often targeted groups that we are most concerned with here.

Unfortunately, the number of individuals that can be described as belonging to such groups is growing. A global wave of autocratization and erosion of rights has wiped out advances in democracy over the last 35 years, and now more than 72% of the world's population live under autocratic regimes, according to V-Dem Institute's 2023 democracy report.[7] That amounts to more than 5.7 billion people. The decline in democracy and individual rights and freedoms is not merely in countries such as India, now regarded by V-Dem as an electoral autocracy, but is also notable even in democracies such as the United States, which is also in a period of autocratization. Freedom House, another nonprofit organization that measures the health of the democracies, notes that U.S. "democratic institutions have suffered erosion," and in its analysis points to declines in bodily autonomy in the wake of the Supreme Court's ruling that abortion is not a constitutional right, as well as violence provoked by racism and discrimination against LGBTQ+ people.[8]

Given this set of circumstances, it is perhaps no surprise that internet freedom and digital rights are also in decline. In its 2022 Freedom on the

4 https://www.cfr.org/in-brief/how-israels-pegasus-spyware-stoked-surveillance-debate
5 https://www.politico.com/news/2022/12/21/data-brokers-privacy-federal-government-00072600
6 https://www.jstor.org/stable/3481326
7 https://www.v-dem.net/documents/29/V-dem_democracyreport2023_lowres.pdf
8 https://freedomhouse.org/country/united-states/freedom-world/2023

Net report, Freedom House notes that "global internet freedom has declined for a 12th consecutive year," and that even in democracies threats to freedom of expression and unchecked surveillance are major problems.[9] "Too often, rights considerations are disregarded in favor of the misguided belief that more intrusive tools and greater state access to data will necessarily contribute to a safer society," note the report's authors. One of the dimensions that Freedom House charts across dozens of countries is whether governments place restrictions on private and encrypted communication.

These trends are buttressed by the infrastructure and incentives of most major technology firms and a sprawling advertising technology industry, first termed "surveillance capitalism" by author and Harvard Business School professor Shoshana Zuboff in a 2014 essay, "A Digital Declaration."[10] The following year, Harvard communications scholar Vincent Mosco noted the connection between surveillance capitalism and the surveillance state in his book, *To the Cloud: Big Data in a Turbulent World*.[11] The interplay between the two were demonstrated most dramatically, perhaps, by the revelation of the US National Security Agency (NSA) PRISM program by the whistleblower Edward Snowden. PRISM allowed US intelligence agencies to get access to information on foreign individuals from internet firms such as Google, Facebook and Yahoo!.[12]

These revelations preceded the broader 'techlash' that followed the 2016 U.S. election and Brexit referendum in the United Kingdom, sparked by various catalyzing events, including major data and privacy breaches, such as the Cambridge Analytica scandal. A deterioration of trust in tech firms has led to worries that even encrypted communications are not truly secure, that technology firms are willing to provide 'backdoors' to law enforcement, or that they themselves may scan user messages to harvest insights. At the same time, the firms face a policy environment across the world that is increasingly hostile to encryption, as advocates for various schemes to provide law enforcement with the means to identify harmful material on user devices appear poised to advance legislation that threatens end-to-end encrypted communications.

In the context of these broader dynamics–a global wave of autocratization, the expansion of the surveillance state, the rise of surveillance capitalism, and the dynamics of the broader 'techlash'–encrypted communication tools remain crucial for safeguarding individual freedoms

9 https://freedomhouse.org/report/freedom-net/2022/countering-authoritarian-overhaul-internet

10 https://opencuny.org/pnmarchive/files/2019/01/Zuboff-Digital-Declaration.pdf

11 https://www.routledge.com/To-the-Cloud-Big-Data-in-a-Turbulent-World/Mosco/p/book/9781612056166

12 https://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/

and human rights. As a wide range of individuals and groups around the world contend with degrading political and legal contexts–from journalists, to dissidents, to those seeking reproductive healthcare in jurisdictions where it has been criminalized–secure communication is necessary simply to permit free expression. Those relying on encrypted messaging applications are often putting an enormous amount of trust in them to prevent the compromise of their privacy and security, and thus these apps are worthy of close scrutiny.

LITERATURE REVIEW

Scholars have considered the security and privacy of mobile messaging apps from a variety of perspectives, using the tools of a range of disciplines. And, they have evaluated the importance of encrypted communications in a range of use cases where security and privacy concerns are acute. Most research on encrypted messaging applications focuses on technical aspects of their functionality. A handful of researchers have focused on design and user behavior aspects of messaging apps, including their use by individuals in vulnerable or targeted groups. Notable studies that inform our thinking include:

- In "A Comparison of Chat Applications in Terms of Security and Privacy," published in 2019 in the proceedings of the 18th European Conference on Cyber Warfare and Security, South Africa Council for Scientific and Industry Research researchers Johnny Botha and Carien van 't Wout and University of the Western Cape, South Africa researcher Louise Leenen compared secure messaging apps "based on the built-in security and privacy features of the apps, as well as the location and subsequent accessibility of stored data." The researchers considered the security and privacy features of messaging apps, including whether they are end-to-end encrypted, the ability to send disappearing messages, screenshot detection, the ability to remotely wipe messages, account self-destruct features, and more. They also evaluated the accessibility of stored data. They found that "people choose messaging apps based on different criteria and would hence have different requirements in terms of levels of security," and recommended that users should take care to "apply the correct settings" to maximize their security.[13]
- In the 2022 report "Unbreakable: Designing for Trustworthiness in Private Messaging," researchers at Dalberg used a mixed methods approach to

13 https://researchspace.csir.co.za/dspace/bitstream/handle/10204/11140/Botha_2019.pdf

explore user interaction and potential harms stemming from the use of private messaging apps, with a focus on participants in Colombia, Nigeria, and the U.S. Dalberg found that "most users have built up fairly complex ways of engaging and adapting to risks and concerns as they perceive them," suggesting users may adjust their behaviors based on their understanding of the potential shortcomings of messaging applications, and that since "platform design and governance can enable and exacerbate these harms, platform providers have a responsibility to both understand them and take steps to mitigate them." The report detailed a range of design opportunities to improve messaging apps, including "securing and/or modifying account information," "providing accessible & tailored security & privacy controls," "providing support mechanism & emergency controls," "improving verification & permission mechanisms," and "improving administrative & management tools."[14]

- In the 2022 report "Digital Crime Scenes: The Role of Evidence in the Persecution of LGBTQ People in Egypt, Lebanon, and Tunisia," Harvard researcher Afsaneh Rigot considered how "selfies, sexts, dating app chats, and other common forms of communications have become potential tools for prosecution" of LGBTQ+ people in these countries. In all three countries, Rigot spoke to individuals who had been pursued by law enforcement over purported crimes related to their sexual orientation. The ubiquity of encrypted apps in such investigations is notable: "Notably, 100% of interviews mention evidence taken from WhatsApp chats." While the situation faced by these groups is dire, Rigot sees inspiration in their creativity and inventiveness to try to preserve their own ability to express themselves and communicate safely. "Queer people, like many other marginalized and oppressed communities, are creative, diligent and masters in navigating the risks that befall them," writes Rigot. "Often those that face the highest brunt of laws and structures created to erase them are those who excel at building strategies to survive." The report offers a range of recommendations, including the introduction of double PIN and secret folders, a "self destruct/panic button" on phones, and the dissociation of app accounts from phone numbers.[15]
- In "The Future of Investigative Journalism in an Era of Surveillance and Digital Privacy Erosion," a chapter in the 2019 book *Digital Investigative Journalism: Data, Visual Analytics, and Innovative Methodologies in Innovative Reporting* edited by Oliver Hahn and Florian Stalph, Reuters

14 https://www.designtrustworthymessaging.org/

15 https://cyber.harvard.edu/sites/default/files/2022-03/Digital-Crime-Scenes_Afsaneh-Rigot-2022.pdf

Institute for the Study of Journalism scholar Julie Posetti points to the importance of encryption, among other tools and tactics, in maintaining secure communications with confidential sources. Posetti notes that among more sophisticated investigative journalists, including the ~400 journalists associated with the International Consortium of Investigative Journalists (ICIJ) that produced The Panama Papers, using encrypted communication channels is a standard practice. But it can be complicated, and since "digital security measures designed to protect sources can be unwieldy and time-consuming," it is often "a deterrent to many investigative journalists." Posetti notes that former UN Special Rapporteur for the Promotion and Protection of the Right to Opinion and Freedom of Expression, David Kaye, "has declared encryption an important tool to secure human rights as a range of States seek to limit the availability of encryption and ensure 'backdoor access' to encrypted data."[16]

- In "Digital Privacy for Reproductive Choice in the Post-Roe Era," forthcoming in *New York University Law Review,* Aziz Huq and Rebecca Wexler note that "The overruling of Roe v. Wade unleashed a torrent of regulatory and punitive activity restricting lawful reproductive options." In their "comprehensive accounting of abortion-related digital privacy after Dobbs," the authors stress that "digital privacy for pregnant persons in the United States has suddenly become a tremendously fraught and complex question." They point to the value of end-to-end encryption, and the value of recommendations from "digital civil liberties communities," but suggest the guides produced by such organizations may be least useful to the most vulnerable users, including "the low-income, and often members of marginalized racial and ethnic minorities" who "may not have the time or resources to invest in researching and implementing privacy precautions." The authors propose a mechanism to seek out and guide pregnant persons "through the necessary precautions to securely use digital services and to minimize" any data they may share that could betray the privacy of their healthcare considerations or decisions.[17]
- In "The Pregnancy Panopticon: Abortion Surveillance After Roe," Surveillance Technology Oversight Project researchers Albert Fox Cahn and Eleni Manis note that companies including Meta and Apple "face significant design choices for encrypted services," and that in some cases the companies make security compromises that may

16 https://link.springer.com/chapter/10.1007/978-3-319-97283-1_23

17 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4191990

betray user privacy, such as collecting unnecessary metadata.[18] In the May 2022 report, Cahn and Manis note that "[w]hile many of the country's largest tech companies have been quick to voice support for pregnant employees, few have addressed how their own data is going to be used against abortion seekers," stressing that companies such as Apple, Meta, and Google must do more to minimize data collection and storage, and to deliver on the promise of secure communications that is often made in corporate marketing materials. "Tech giants like Apple, Facebook, and Google must dramatically improve encryption and privacy protections," they write, "ending the mass police surveillance they've allowed to become commonplace."[19]

- In the 2022 paper "Content-Oblivious Trust and Safety Techniques: Results from a Survey of Online Service Providers," Stanford Internet Observatory researcher Riana Pfefferkorn details the results of a survey of trust and safety techniques that included questions related to "content oblivious" trust and safety strategies and content moderation techniques of the sort that would apply to encrypted messaging applications. The survey focused on two particular trust and safety techniques for "content oblivious" sources- metadata and user reports. The survey found that user reporting is widely used for trust and safety, a finding that "undercuts the assumption that at-will content analysis is the only means of fighting abuse in an increasingly end-to-end encrypted world."[20]

- In "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0," Alma Whitten and J.D. Tygar find that "user errors cause or contribute to most computer security failures, yet user interfaces for security still tend to be clumsy, confusing, or near nonexistent." The paper, originally published in the *Proceedings of the 8th USENIX Security Symposium* in 1999, sets out to define "a usability standard for security." For Whitten and Tygar, security software is deemed usable if users "are reliably made aware of the security tasks they need to perform"; "are able to figure out how to successfully perform those tasks"; "don't make dangerous errors"; and "are sufficiently comfortable with the interface to continue to use it." The authors believe better design strategies must "communicate an accurate conceptual model of the security to the user as quickly as possible."[21]

18 Not to be confused with the paper of the same name written by one of the authors of this report. https://www.eff.org/wp/pregnancy-panopticon

19 https://www.stopspying.org/pregnancy-panopticon

20 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3920031

21 https://people.eecs.berkeley.edu/~tygar/papers/Why_Johnny_Cant_Encrypt/OReilly.pdf

- In a series of 2018 posts on the website of the Electronic Frontier Foundation (EFF), authors Nate Cardozo, Gennie Gebhart, and Erica Portnoy "explain what makes different aspects of secure messaging so complex." Based on "months of investigation," the authors lay out "significant tradeoffs" between different secure messaging apps, including complicating factors such as the policies of different applications, the degree to which a user's network is willing to use one app or another, the incorporation of various other features such as cloud backups, and more. They note that "encryption is the easy part," highlighting the importance of overall code quality, user experience, and service availability and reliability as other complicating factors. The authors also suggest that evaluating secure messaging apps should include consideration of how developers respond to technical problems, legal threats, and government and law enforcement.[22] [23] [24] [25]

- In the 2022 report "Design From the Margins: Centering the Most Marginalized and Impacted in Design Process- From Ideation to Production," Afsaneh Rigot asks, "can companies and technologists reframe their features or standards to support the most marginalized users' needs?" Describing a new methodology developed in answer to that question, Rigot suggests how it might be applied in the design of "privacy-preserving tools and security protocols." If designers take into consideration the needs of the most vulnerable users, it should, says Rigot, offer a higher standard of protection to all users. "In an effort to create more protections for decentered users, the collateral byproduct is better privacy protocols" for the product in general.[26]

- In a 2017 study titled "Obstacles to the Adoption of Secure Communications Tools," researchers from the University of Bonn, Germany, University College London, Stanford University and the Electronic Frontier Foundation detail interview findings suggesting that "the adoption of secure communication tools is hindered by fragmented user bases and incompatible tools."[27] The authors queried interview subjects about their threat models and mental models of secure communications. They conclude that "in the long run, if security developers want to develop new paradigms and secure communication tools in a user-centered design process, they need to understand users' goals and preferences."

22 https://www.eff.org/deeplinks/2018/03/secure-messaging-more-secure-mess
23 https://www.eff.org/deeplinks/2018/03/why-we-cant-give-you-recommendation
24 https://www.eff.org/deeplinks/2018/03/thinking-about-what-you-need-secure-messenger
25 https://www.eff.org/deeplinks/2018/03/beyond-implementation-policy-considerations-secure-messengers

26 https://www.belfercenter.org/publication/design-margins
27 https://ieeexplore.ieee.org/abstract/document/7958575

EXPERT PERSPECTIVES

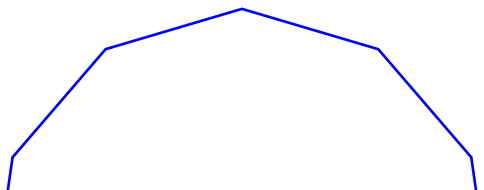
Our exploration of these issues included a range of interviews with independent security experts, technology executives, academics, lawyers, activists, and other representatives from civil society groups concerned with issues related to encrypted communications, privacy, and technology policy. These discussions pointed us to a number of questions and concerns related to the design, engineering, use, and policy environment around encrypted messaging applications.

- **Security is not a binary, but rather a continuum, and a number of factors combine to make applications more or less safe.** These factors include, of course, the technical and cryptographic sophistication of the application, but also crucially its design, its policies (particularly around data retention), and its integration with other services. But these factors, which are those within the control of the makers of the application, are only half of the story: user behavior introduces a range of other factors, from cultural norms to more or less extreme threat scenarios that impact user safety.
- **There is a severe lack of understanding among most users about how to judge their own threat models, how encrypted messaging applications work, and how they should best use the applications in individual and group communications to preserve their privacy.** Efforts to train people are sparse, difficult to scale, and under attended to by both industry and civil society. In the US, events such as the 2020 protests following the murder of George Floyd, as well as acute privacy concerns following the 2022 Supreme Court decision to overturn Roe v. Wade, have created a demand for training. Likewise, in other parts of the world, activists, dissidents, journalists, LGBTQ+ people, and others are in need of training and clear guidance on how best to communicate safely. What is common across jurisdictions is frequent uncertainty about the legality of communication around a given topic; increasingly aggressive use of digital evidence against individuals by law enforcement; and often a desire not to put others at risk.
- **At the same time, it should not fall to the user to scrutinize apps from a technical or design perspective, or read terms and conditions like a lawyer, to know whether a particular application can be used to communicate securely or not.** There is a need for standardization, transparency, and clarity from the makers of encrypted messaging apps to enable users to make better decisions.
- **There is a lack of trust in some applications based on brand associations (WhatsApp) and hearsay (Telegram) that are sometimes incomplete or inaccurate.** Rumors and folkloric

assumptions about the motivations of the owners of particular applications, such as Meta (in the case of WhatsApp) or Telegram, inform user decisions about what applications to use and shape user perceptions of their trustworthiness.

- **There is a class of users who know or suspect they are being surveilled by state actors for whom encrypted applications may offer little protection.** Spyware scandals, such as the widespread use by governments of Pegasus, the powerful phone surveillance system, as well as the physical acquisition of devices and other forms of forced entry into communications, are a real threat for many people who exist just beyond the threshold where taking reasonable precautions with encrypted messaging applications offers substantial defense.

- **There is a likely irreconcilable tension between law enforcement and advocacy organization calls to scan for harmful material, such as child sexual abuse material or terrorism-related content, and the ability to preserve encrypted communications.** While some technical schemes have been proposed, including solutions such as client-side scanning, these are technically fragile and introduce a range of other security and surveillance concerns, when the harms may be better addressed by other means, including better social support systems.



KEY RESEARCH QUESTIONS

Based on our review of relevant literature and expert interviews, we developed a handful of key hypotheses we set out to address over the course of the project:

1. How do people form mental models of their own or their organization's digital security, and threats to it? How does their use of encrypted messaging applications correspond to these threat models?

2. What are the technical and/or design choices app makers have made that lead to or confuse the user into making poor security decisions, or that put the user at risk?

3. How might people end up compromising their own safety and access to quality information, either individually or in groups, due to their behaviors?

4. What is the prevalence of these phenomena, and what are potential technical, design, and policy solutions?

METHODOLOGY

For this study, we focused our attention on Signal, WhatsApp, Telegram, Apple Messages (iMessage), Messages by Google, and Meta's Messenger (Facebook Messenger). Following expert interviews that helped refine our goals, our mixed methods, journalistic approach to this project included four major components:

1. **Fieldwork: Individual and group interviews.** We conducted multiple informal individual and group interviews in two distinct geographies and with two distinct user groups in order to better understand the perceptions, behaviors, and experiences of individuals in vulnerable and targeted communities and professions.

2. **Technical review.** Our methodology for technical analysis of encrypted messaging applications relied on static analysis of source code where available (Signal, Telegram), analysis of publicly available documentation and developer statements, and dynamic analysis of network traffic, where possible, using MITMproxy, an industry standard tool for analyzing encrypted network traffic using HTTPS.[28] This analysis was conducted between February 21st and March 3rd, 2023. While we initially planned to do a protocol level analysis of each app, it became clear that many messaging apps use bespoke communication protocols which thwarted the analysis. We then switched to analyzing publicly available technical documentation for the majority of the technical analysis.

3. **Design review.** The methodology for analyzing the user experience and user interaction design of encrypted messaging applications included conducting a heuristic evaluation, charting the user journey, and diagramming functionality in detail, including evaluating how users initiate conversations, navigate through the app, and adjust their settings, with a focus on how easy and straightforward these processes are. Additionally, we scrutinized how security features are incorporated into the design, and assessed the accessibility and ease of use of the applications. The ultimate goal of the analysis was to understand how these apps achieve the critical balance between providing a satisfying user experience and maintaining robust security measures. We conducted this analysis from March 31st to May 29th, 2023.

4. **Policy review.** We reviewed the privacy statements, terms and conditions for this handful of apps, considering their complexity and where their approaches were the same or different. We also evaluated transparency reporting and other disclosures from the companies, including the channels by which they communicate product and policy updates, and the documentation they provide to users generally. We also considered the broader policy environment in which these applications function, with a particular focus on the US, EU, United Kingdom, and India.

28 https://mitmproxy.org/

Field Work

In order to better understand the perceptions, behaviors, and experiences of individuals in vulnerable and targeted communities and professions, we conducted journalistic fieldwork in two distinct geographical, social, and political contexts: with reproductive health and abortion rights activists in New Orleans, Louisiana; and with journalists in Delhi, India.

In New Orleans, we engaged with individuals with the help of the The Reproductive Justice Action Collective (REJAC), a network of Southern activists based in New Orleans. The reliance of activists concerned with abortion access and reproductive health issues on digital platforms for communication, organization, and advocacy makes digital security a critical component of their operations. Louisiana is one of the states with a strict ban on abortion.[1] Following the Supreme Court's decision to overturn *Roe v. Wade*, there is a wave of legislation in a number of states not just to criminalize abortion, but also to criminalize sharing information or offering assistance to anyone seeking care.[2] In Louisiana, the shifting legal terrain has left even doctors uncertain about what is permitted under the law.[3]

In Delhi, we engaged with journalists concerned about source security and the possibility of government and law enforcement surveillance with the help of the Internet Freedom Foundation, a nonprofit organization concerned with online freedom, privacy and innovation in India.[4] Working as a journalist in India is increasingly risky. Reporters Without Borders, an organization that advocates for press freedom, notes that India "is one the world's most dangerous countries for the media," with journalists facing reprisal from politicians, police, and criminal groups.[5] Physical and digital security are real concerns for journalists, many of whom use encrypted messaging applications for communications with fellow reporters and editors, and with their sources.

Our core objective in these semi-structured discussions was to engage in open dialogue and discussion with individuals in these two groups, in order to better understand their thought processes, current practices, motivations, challenges, and apprehensions concerning digital security, and specifically the use of encrypted communications for individual and group communications. We conducted both individual and group sessions, using a

1 https://liftlouisiana.org/content/faqs-about-abortion
2 https://www.hrw.org/news/2023/04/18/human-rights-crisis-abortion-united-states-after-dobbs
3 https://time.com/6282288/louisiana-abortion-exceptions-confusion-doctors/
4 https://internetfreedom.in/
5 https://rsf.org/en/country/india

semi-structured format to encourage dialogue and exploration of participants' experiences. And, we interviewed the leaders of REJAC and the Internet Freedom Foundation to understand their perspectives on issues related to privacy, surveillance, and the use of encrypted messaging applications in the communities under consideration. We chose these partners based on their common concerns on these issues, and our working relationship with them.

Because of the sensitive nature of some of the concerns and scenarios shared in these discussions, we synthesized and anonymized notes from these sessions. Group discussions were conducted under the Chatham House Rule, which states that "participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed."[6] Key themes related to perceptions and mental models about digital security and surveillance, encrypted messaging applications, and what constitutes safety emerged in both geographies, and in many cases there were commonalities between the groups.

But there were also important differences, which we also took care to note. In general, the security concerns and risks in Delhi were more severe and pervasive than in New Orleans. We heard multiple stories of interactions with law enforcement in our discussions in Delhi, including device seizures.

A range of observations from these discussions are referenced in the below findings. Below are a handful of key overarching themes we encountered from our workshop sessions:

1. In shifting legal and political contexts, individuals are often uncertain what they can say about sensitive topics, whether it may be used against them by authorities, and the extent to which technology companies may be forced to provide information about their communications to law enforcement. There is little understanding of the legality or process for such requests (in either geography).

2. Individuals are in need of more and better information about how to secure their digital communications, but they often find it difficult to find accessible sources of information. Yet at the same time, even individuals who are deeply concerned about privacy and security confess to having never read the terms and conditions of the encrypted messaging applications they use, and to have never visited their transparency reports or other materials.

3. In the place of sufficiently clear and accessible information, individuals rely on word of mouth and "security folklore" to determine what applications and behaviors are safe, and which are not. Often this involves calculations based on the brand or corporate entity behind the

6 https://www.chathamhouse.org/about-us/chatham-house-rule

application, which is reasonable for a user making decisions with limited information. Nevertheless, better sources of information could help ensure individuals make the best decisions possible.

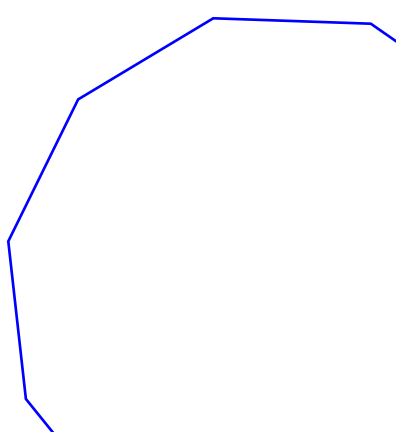
4. For the most sensitive communications, individuals often revert to in-person or non-digital forms of communication for maximum security. Though it was not possible to verify the circumstances, we encountered multiple stories of individuals believing metadata related to their calls or communications on encrypted messaging applications were intercepted or flagged. Regardless of the circumstances, these incidents point to the difficulty encrypted messaging apps must face in maintaining user trust in the context of increasingly aggressive state and law enforcement activity.

5. Individuals are particularly wary of applications that are connected to other services and products, such as Google or Meta services and apps. And, they are concerned about potential security compromises when new features are added to encrypted applications.

6. There are some workflow downsides to utilizing features such as disappearing messages. People often need to refer back to information, or to have records for their own internal review or processes.

7. We encountered various concerns, especially among journalists, about the shifting policy landscape with regard to encryption, the possibility of client-side scanning being introduced on devices, and questions about whether the benefits of encrypted messaging applications would be nullified.

These observations underscore the reality that no individual threat models are the same, since no two users and no two situations are the same. In this research we present suggestions for improvement based on a synthesis of different threat models. It is important for the reader to consider their own threat model when considering our recommendations.



The Interaction Between Privacy, Security, and Design

Design is a broad term. Here we use it to define the entire lifecycle of a product, which includes software development, product planning, user experience, user research, testing and QA, and launch. The products we analyzed have complex technology stacks and designed interfaces developed with consideration given to how individuals and communities will interact together, separately, and orthogonally. We use design here intentionally to refer to **design and technology**, with an emphasis on how a product combines cryptography, backend and frontend engineering, the design affordances of the production including graphic design, user experience, user interface design, and copywriting–how features are named and even descriptions provided to the users to explain what a feature does–along with myriad other product-related considerations.

In our research, we analyzed the apps specifically and holistically, by mapping nearly every feature and functionality that was visible to the user in the mobile application. Our analysis considered the user journey; granular UI choices; how or if cryptography was presented and described to the user; the 'tech' stack of particular apps; the language or 'content strategy' within the apps themselves; and how all of these specific design choices come together within a product that a user experiences. Like puzzle pieces, these specific elements of design impact how a user understands and builds mental models of what the app does and what actions it allows.

Design is a facilitator and narrator of a user's experience within an app. Similarly, design has a politics to it: as much as design can explain, elevate and illuminate information or functionality to a user, it can also obfuscate and confuse what an app does and how safe it is, with potentially dire consequences.

The products we analyzed are built using a variety of methodologies, from human centered design[1], flat design[2], material design[3], and Privacy by Design.[4]

Our approach to this analysis and research, especially as regards our suggested improvements, is to analyze what features and choices are presented to the user, how features and choices are described, and what, technically, is happening "underneath the hood". We then look at these design decisions through the lenses of security best practices and two design methodologies, Design from the Margins and Privacy by Design.

Design from the Margins, created by Afsaneh Rigot, "centers the most impacted and marginalized users from ideation to production, pushes the notion that not only is this something that can and must be done, but also that it is highly beneficial for all users and companies. For this to happen, consumer interest conversations need to be framed outside the 'biggest use case' scenarios and United States and European Union-centrisms and refocused on the cases often left in the margins: the de-centered cases."[5]

1 https://www.ideo.com/post/design-kit
2 Apple's preferred design methodology. See https://developer.apple.com/design/human-interface-guidelines/ and https://developer.apple.com/design/
3 The preferred design methodology for Google. See https://material.io/design
4 https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf
5 https://www.belfercenter.org/publication/design-margins

This approach prioritizes the most marginalized users' concerns, reflecting their lived experiences, rather than focusing primarily on the needs and interests of the 'common user' so often centered in the design of products produced by big tech companies that seek scale.

Privacy by Design, developed by Dr. Ann Cavoukian in the 1990s,[6] heavily centers privacy in the development of a product, service or software. It has seven principles:

1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default
3. Privacy Embedded into Design
4. Full Functionality-Positive-Sum, not Zero-Sum
5. End-to-End Security-Lifecycle Protection
6. Visibility and Transparency
7. Respect for User Privacy [7]

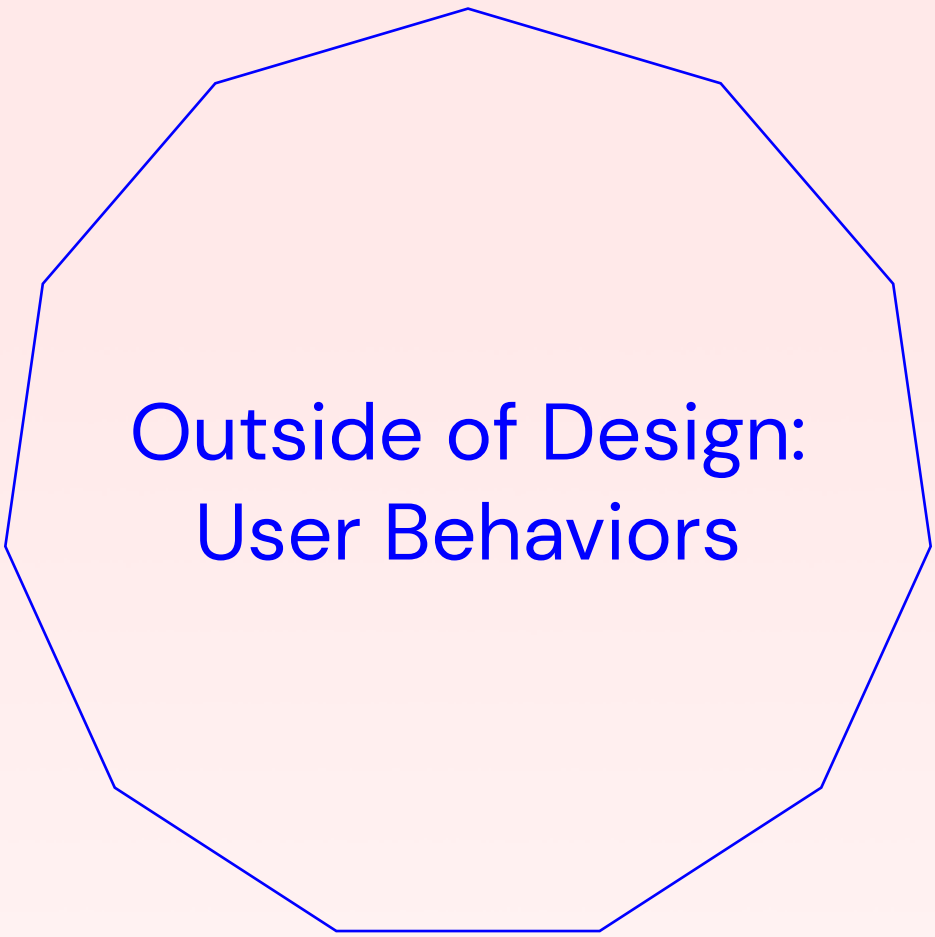
Design from the Margins and Privacy by Design are frameworks that, if utilized by technology companies, could help produce safer technology that works better for the most at-risk users as well as the most privileged users, while protecting their digital rights, personal information, and privacy. A shift towards these two methodologies in modern software design would have material impacts on users by producing software that keeps these users safer, treating their needs as paramount rather than as 'edge cases'. High risk users have real, urgent, and specific concerns for their safety, and their adversaries are often powerful state actors. An additional challenge is that the design of much of the modern, corporate software on which billions of users rely reflects the interests of companies operating on the model of surveillance capitalism, and is thus at odds with the privacy, and therefore safety, of marginalized users.

We incorporated threat modeling as an auditing, brainstorming, and analysis tool by plotting out how each research finding and recommendation could potentially cause harm, and how it could be utilized to harm, unintentionally and intentionally, from large-scale to small-scale threats, including how that harm could be asymmetrical depending upon a user's background, location, and demographics. This paradigm is part of a digital rights framework and built out from users' experiences, needs and accounts of real world events. We crafted personas to test the recommendations against, such as 'the most privileged' (often referred to as the 'most common'); users facing different kinds of online harassment;

6 https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf
7 https://iab.org/wp-content/IAB-uploads/2011/03/fred_carter.pdf

and high risk users. In developing these profiles, we drew on real users the team has encountered in prior work, such as activists and journalists in Majority World countries facing a variety of threats from state level actors to adversaries in their own communities, or even families.

In each section, we offer suggested improvements that focus on policy, design and/or technology changes that the developers of encrypted messaging applications could implement now. These recommendations are not and cannot be perfect solutions; some of the issues users face are systemic and span geographies, regulatory blocs, culture, and language. With that in mind, some of our recommendations follow the approach of harm reduction, not letting the perfect be the enemy of the good. Small changes can create helpful friction to minimize harmful behaviors and outcomes. An example is our suggestions related to screenshot notifications or disabling screenshots: while it is not technically possible to stop someone from using another device to photograph a message, these interventions can make it that much harder for a bad actor to harass or harm individuals, or for an individual to accidentally or unnecessarily betray their own interests.



Outside of Design: User Behaviors

Our field work suggests that user perceptions of the trustworthiness of any given secure messaging app is largely influenced by design. Design is not just what's in an app, service, or platform, it is also how users interpret and engage with that app and incorporate it into their daily lives and activities. With privacy- and security-focused apps, like secure messaging, how users understand or misunderstand safety related to those apps is incredibly important. In our qualitative interviews and workshops, we observed a variety of types of user behavior that directly impact user safety.

Folk security, or **security folklore**, is a mental model of risk in which a user employs different kinds of privacy and security protocols that were communicated through friends, family, or coworkers, often without knowing why these protocols are believed to be effective.[1] Sometimes these actions are helpful, and sometimes these actions are extra precautions that are extreme in nature but do not necessarily meet the user's threat model, and while these actions technically make the user safer, it does not have a substantive impact on the user's safety.

Users we interviewed mentioned being afraid of state actors such as local government or police, or being afraid of an overreaching surveillance state. One user mentioned that they would place their phone into a freezer or microwave when having sensitive in-person meetings. Other users mentioned using mic-jamming apparatuses when attending protests.[2] These were actions they were encouraged to do by others to help mitigate surveillance.

To note, these specific users are based in a "Global North" or minority world context, where the majority of local police departments likely lack such capabilities, and do not have a history of engaging in malware-based surveillance. However, these users' fears are very much justified; they live in a state where abortion laws are constantly changing,[3] and there is limited knowledge as to what forms of advice are legal to provide. In this new legal landscape, users have difficulty realistically modeling the threats that they face from technology developed under the incentives of surveillance capitalism,[4] particularly given headlines that call into question promises made by large technology companies about the handling of sensitive data.[5]

Conversely, another theme we noticed was **security nihilism**, in which individuals are so overwhelmed by either information or by extreme threats (real or perceived) that engaging in good safety and privacy best practices seems pointless. Users questioned that if the state can seize a phone or install malware at any time, can messaging through Signal or using disappearing messages actually help? The authors, who are based within the minority world/Global North, acknowledge this painful reality. But, just as any amount of friction from privacy

1 https://cups.cs.cmu.edu/soups/2010/proceedings/a11_Walsh.pdf

2 There is reason to question the efficacy of these devices, but beyond that, in the myriad ways local police departments surveil protests it has to our knowledge never been demonstrated that a local police department used malware to listen to people's microphones in real time.

3 https://www.nytimes.com/2022/07/13/technology/personaltech/abortion-privacy-roe-surveillance.html

4 https://www.npr.org/2022/07/02/1109565803/data-privacy-abortion-roe-apps

5 https://www.washingtonpost.com/technology/2023/05/09/google-privacy-abortion-data/

features can be helpful, even if an activist or journalist is being monitored by the state, using tools like Signal and features such as disappearing messages make it that much harder for the state to surveil those individuals. There remains a risk of backfire, however, such as in the case of a phone seizure, if merely having such an app on one's device is regarded by the state as incriminating.

Other misconceptions we noticed were the idea that **software is a silver bullet**, and that any encrypted app must mean that the users are 'safe'. There was a disconnect in understanding where the safety of an app stops, and where a user's offline decision making starts. For example, using Signal without disappearing messages protects messages from dragnet surveillance, but if a user's phone is seized and unlocked, the messages are still there. While Signal is end-to-end encrypted, the encryption alone does not guarantee privacy.

Relatedly, we noticed that **group behaviors and norms can harm safety**, such as with the creation of messaging groups in Signal that are so large that users are not sure of the identity of every user in the group, or even how many users there are in total. Subjects reported that some users will share sensitive information about themselves in such groups. Interviewees also expressed hesitation about sharing in groups, but had difficulty expressing why this felt particularly unsafe.

These findings intertwined with users **hearing from others that Signal is better, but not knowing why**. Users recounted that they heard from friends of friends that Signal was a safer and better option, but none could articulate why exactly. There was not any language within the Signal app that could explain to them why Signal is a safer or better option than other messaging apps.[6]

In group discussions in Louisiana, we were able to unpack how encryption works, what features may make Signal safer, and what other actions these individuals can do to be safe while using end-to-end encrypted messaging apps. A large part of this thorny issue is recognizing the geographical location that these users exist in and the users' backgrounds.[7] They are not primarily technology experts—most work in fields other than technology—and they live in a US state that lacks a significant number of technology advocacy groups, widespread technology literacy and educational programs, and local news organizations are not necessarily technology-focused.[8] Civil society organizations, a press that

---

6 It's important to note that even if the Signal App had explained in plain language why it is safer than Whatsapp, Telegram and other messaging apps, that these particular users might not have believed Signal.

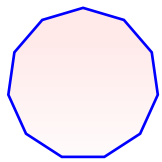
7 "An estimated 460,000 Louisiana adults aged 18 to 64 do not have basic computer skills (ConnectLA 2023). This means that 1 in 10 Louisianians struggle to participate in online commerce, banking, telehealth services, educational opportunities, and many other facets of life that require digital skills to participate." https://connect.louisiana.gov/media/fc2pdo4y/la-draft-digital-equity-plan.pdf

8 Some local journalism groups focus on surveillance, but few reporters are diving into technology related harms or have the ability to analyze and assess technology related investigations.

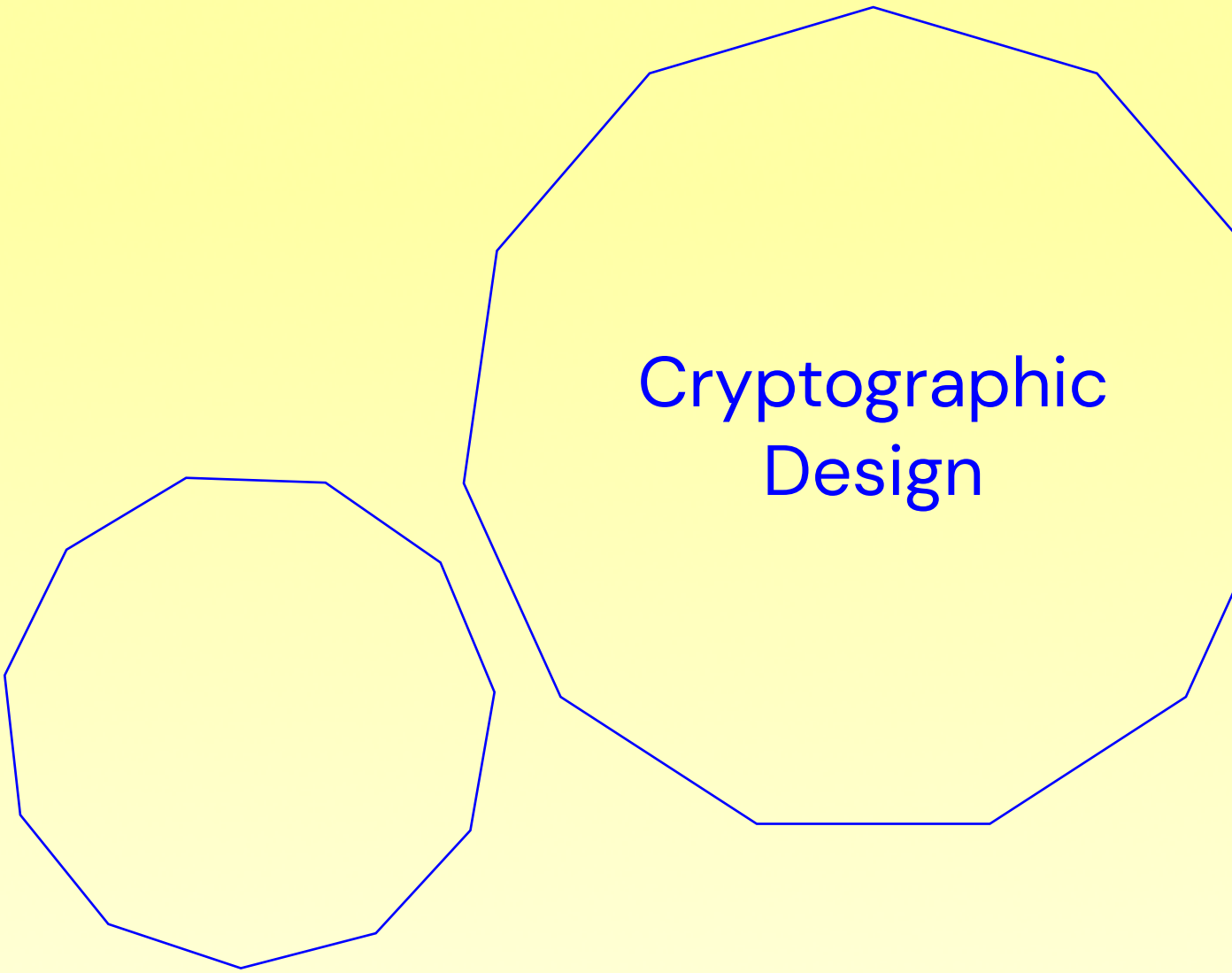
understands technology, and investment by a city and state in technology related funding, including educational and digital literacy programs, are a non-obvious but necessary infrastructure to make issues related to privacy, safety, and technology understandable to all audiences.

Among some of the users we interviewed in Louisiana, we also found an overarching distrust of **WhatsApp, because it is owned by Facebook (now Meta)**. Users mentioned recent media attention around Meta handing over two users' messages in Facebook Messenger in response to a Nebraska law enforcement warrant, and other cases where police used texts in prosecutions related to abortion.[9] [10] Users were unclear as to how different WhatsApp is from Facebook Messenger, if it is secure, and if it should be trusted. The connection between Facebook and WhatsApp makes it seem unsafe to users even if WhatsApp is a safe option, given their threat model.

One user also recounted being physically attacked and having their device seized by local police, while others expressed fear of this happening to them. This kind of physical harm is often called "**rubber-hose attack**."[11] [12] In cryptography, this refers to an attack that doesn't involve cryptanalysis or clever mathematical tricks, but instead refers to extracting sought after passcodes, encryption keys or other information from a human, proverbially by beating them with a rubber hose. In practice this may include police interrogation, arresting someone and stealing their unlocked phone, or other forms of violence. Some interviewees shared techniques for avoiding these types of attacks including: using disappearing messages with a short timer, having contacts saved under code names, leaving their devices at home in high risk situations, having an alternate device to use in high risk situations such as a protest, and turning off biometric unlocking features. These tactics would, of course, have varying degrees of usefulness depending on one's threat model, and some of them may not be attainable for all users for practical reasons, e.g., the expense of a second phone.



9 https://www.npr.org/2022/08/12/1117092169/nebraska-cops-used-facebook-messages-to-investigate-an-alleged-illegal-abortion
10 https://www.washingtonpost.com/technology/2022/07/03/abortion-data-privacy-prosecution/
11 https://xkcd.com/538/
12 Discussed originally in Bruce Schneier's 1996 book Applied Cryptography, https://www.schneier.com/books/applied-cryptography/



Cryptographic Design

Our technical research focused primarily on sending messages as a core feature of the applications we reviewed, with a specific focus on Signal, WhatsApp, Telegram, Apple Messages, Messages by Google and Meta's Messenger. Cryptographic design and implementation are a necessary metric of evaluation, including design cues and acknowledgments of said cryptography within the products. To emphasize, we did not perform analysis of any of the cryptographic libraries used by any of the applications we reviewed, but we did conduct a literature review.

Thus, our analysis relies in part on published formal analysis by cryptographers[1][2] and public documentation from the companies.[3][4][5][6][7][8] We also undertook traffic analyses of apps in use, as well as inspection of source code using techniques described in the methods section at the beginning of this report.

There are two separate types of encryption that are generally thought of when talking about secure messaging. One is transport layer encryption, which is encryption between the user nd the server; this is the type of encryption that an HTTPS connection uses, for example. Chats that only use transport layer encryption, such as normal Facebook messages, can be theoretically read by anyone with access to the application servers, which makes such systems a frequent target of law enforcement requests. The other main type of encryption discussed is end-to-end encryption, where the message content is encrypted between the sender and all recipients of a message, meaning that the only place a message can be read is on a sender or receiver's device. Transport layer encryption and end-to-end encryption are both important and work in tandem to protect the integrity and confidentiality of messages and metadata from eavesdroppers. In the apps we tested, we found that all traffic was encrypted at the transport layer (excluding SMS messages sent by iMessage and Google Messages) using TLS (Signal, iMessage), QUIC (Google Messages, Facebook Messenger), or custom protocols (Telegram uses MTproto, WhatsApp uses the Noise protocol). When discussing cryptography in this report, we are generally talking about end-to-end encryption.

At a high level, the generally accepted opinion on the cryptographic libraries used for end-to-end encryption in the messaging apps we reviewed is that the state of the art for cryptography in messaging is the Signal Protocol[9] and forks of it[10] (as used by Signal, Whatsapp[11], Facebook Messenger, and Google Messages). Currently, there are no major criticisms of the Signal Protocol. The last formal analysis of the Signal Protocol found no issues.[12] Several security vulnerabilities were found in the Signal application in 2017 which have since been patched.[13]

1 https://blog.cryptographyengineering.com/2016/03/21/attack-of-week-apple-imessage/
2 https://www.researchgate.net/publication/346702021_Automated_Symbolic_Verification_of_Telegram's_MTProto_20
3 https://support.apple.com/guide/security/how-imessage-sends-and-receives-messages-sec70e68c949/web
4 https://www.gstatic.com/messages/papers/messages_e2ee.pdf
5 https://www.documentcloud.org/documents/2806301-WhatsApp-Security-Whitepaper-1
6 https://core.telegram.org/mtproto
7 https://about.fb.com/wp-content/uploads/2016/07/messenger-secret-conversations-technical-whitepaper.pdf
8 This literature review is consistent with common practices for determining the security properties of a given application.

9 https://eprint.iacr.org/2016/1013.pdf
10 https://eprint.iacr.org/2019/436.pdf
11 Whatsapp uses the noise protocol, a fork of the signal protocol http://www.noiseprotocol.org/
12 https://eprint.iacr.org/2016/1013.pdf
13 https://conference.hitb.org/hitbsecconf2017ams/materials/D2T1%20-%20Markus%20Vervier%20-%20Hunting%20for%20Vulnerabilities%20in%20Signal.pdf

iMessage and Telegram are the only messengers we reviewed which use bespoke cryptographic protocols not based on the Signal protocol.[14] For iMessage, some concerns have been raised about the security tradeoffs it presents and the fact that it uses a centralized key server, which presents a vector for Apple, or a government working with Apple, to carry out a man-in-the-middle attack.[15] This issue is slightly mitigated by Apple's introduction of key verification in iMessage, which we will discuss later. The last cryptographic vulnerabilities in iMessage were found, and presumably patched, in 2016.[16]

Telegram uses a bespoke protocol called MTProto for transport encryption and end-to-end encryption. Vulnerabilities were found in MTproto as of 2021,[17] which Telegram claims to have fixed in official clients. However, the same researchers issued several other caveats about MTproto and especially third-party Telegram apps. Overall we consider MTproto, and therefore Telegram, to be the weakest cryptographic protocol of the apps we studied.

SUGGESTED IMPROVEMENTS

Telegram should switch to a more modern and well-tested secure chat protocol.

THE DEVIL IS IN THE DETAILS

Even though a cryptographic library is mathematically secure, there is no guarantee that the application has implemented it securely.[18] Implementation bugs can have serious consequences for the secrecy of encrypted messages such as revealing metadata, re-ordering or replaying of messages (as seen above in Telegram), to revealing the unencrypted text of messages. To clarify, we did not perform any analysis of the implementation of these cryptographic libraries within these apps and make no claims that any of them are implemented securely beyond publicly available analysis, as our project has focused primarily on barriers and frictions to secure and safe messaging.

14 See Conclusions and Suggestions for Future Research for more on the centrality of the Signal protocol.
15 https://blog.cryptographyengineering.com/2013/06/26/can-apple-read-your-imessages/
16 https://blog.cryptographyengineering.com/2016/03/21/attack-of-week-apple-imessage/
17 https://mtpsym.github.io/
18 https://mjg59.dreamwidth.org/62598.html

One main concern in regards to the cryptography of the analyzed applications is that applications such as Telegram and Facebook Messenger don't use end-to-end encryption by default for all messages. These apps present the user with unnecessary frictions that cause privacy and security risks by having the user take extra steps to start an encrypted chat and continue an encrypted chat once it is started instead of going back to the plaintext chat. (Meta has committed to make end-to-end encryption the default for Facebook Messenger, and says it plans to roll this out fully in 2023.[19] [20])

Another concern with these two apps is that they use the terms "secret message" (Telegram) or "private chat" (Facebook Messenger), to describe end-to-end encrypted messages, which we will discuss more later. Telegram further confuses users by naming normal chats "Cloud Encryption." Cloud encryption, in this particular case, refers to encryption at the transport layer, and Telegram is still able to decrypt and turn over the contents of Cloud encrypted chats, though they promise they will only do so in "verified cases of terrorism."[21]

Other concerns have been raised about Telegram as well,[22] chief among them that Telegram does not offer end-to-end encryption for Group chats or Channels, and that Telegram has been caught secretly turning over data to the German police.[23]

SUGGESTED IMPROVEMENTS

- Telegram and Facebook Messenger chats should be end-to-end encrypted by default.
- Telegram and Facebook Messenger should make it clear that encrypted chats are encrypted and not use confusing language such as "secret chat" and "private chat."
- Telegram should not use "Cloud Encryption" to refer to chats which are not end-to-end encrypted.
- Telegram should enable end-to-end encryption for groups.

Messages by Google and iMessage encrypt messages opportunistically, meaning they can send unencrypted messages

19 https://www.facebook.com/notes/2420600258234172/
20 https://about.fb.com/news/2023/01/expanding-features-for-end-to-end-encryption-on-messenger/
21 https://www.rferl.org/a/telegram-ceo-defends-new-privacy-policy-says-user-data-still-safe/29458179.html

22 https://www.eff.org/deeplinks/2022/03/telegram-harm-reduction-users-russia-and-ukraine
23 https://www.bitdefender.com/blog/hotforsecurity/der-spiegel-says-telegram-gave-user-data-to-german-police-in-fight-against-terrorism-child-abuse/

but will default to using encryption if all parties in the conversation support encryption. This is not necessarily a bad idea in theory; Signal relied on this in the past to gain traction amongst users, upgrading chats from SMS to encrypted chats on Android until it switched to only supporting encrypted chats in 2022.[24] However, in practice encrypting messages opportunistically can cause users confusion and frustration, and it can be a vector for 'downgrade attacks' or accidental downgrades, where the target is forced to switch to a lower level of security, resulting in users unwittingly sending unencrypted messages.

Much has been said about the fact that Apple and Google do not interoperate well when it comes to encrypted messaging and the resulting frustration it causes users.[25] [26] [27] Google and Apple will need to agree on a standard interoperable messaging format for their default messengers to be compatible with each other or risk driving more users to less secure alternatives such as Telegram, or more secure alternatives such as Whatsapp and Signal.

In our user interviews, we observed that a majority of the iMessage users we spoke with did not realize that blue chats in iMessage signaled that the messages were encrypted. Users didn't seem to think of iMessage as a particularly more secure option, and most didn't realize it used any type of end-to-end encryption.

Google's implementation of encrypted chat suffers from problems as well. Messages by Google uses a new transport protocol[28] that the company and mobile carriers are pushing as a standard to replace SMS (regular text messages), which is called RCS.[29] The end-to-end encrypted version of RCS uses the Signal Protocol and is generally referred to as Encrypted RCS. Google has been attempting to push RCS as a standard that Apple could integrate into iMessage so that Android phones and iPhones could interoperate.[30]

However, currently Messages by Google falls short of iMessage in several key ways. First, Google requires multiple steps to turn on the ability to initiate an encrypted chat, which both parties must do for encrypted messaging to work.[31] Additionally, RCS relies on mobile carriers to support it, though RCS support is quickly gaining traction worldwide. However, there are still laws with Encrypted RCS, even if Google was able to get every Android user using it by default. Even if the mobile carrier supports RCS they can still choose to block encrypted messages.

24 https://signal.org/blog/sms-removal-android/
25 https://www.androidpolice.com/google-rcs-messaging-feud-apple-imessage/
26 https://www.droid-life.com/2022/09/08/dont-like-the-green-bubble-buy-an-iphone/
27 https://www.androidauthority.com/green-bubble-phenomenon-1021350/

28 https://www.bbc.com/news/technology-43836504
29 Rich Communication Services
30 https://9to5mac.com/2023/05/10/google-apple-rcs-support-iphone/
31 https://www.androidcentral.com/how-enable-and-use-end-end-encryption-google-messages-app

Also, the current standard for Google RCS cannot hide any metadata, since messages are routed through servers controlled by the phone company (possibly an RCS server hosted on behalf of the phone company on Google Cloud). RCS by its nature does not hide metadata such as: phone numbers of senders and recipients; timestamps of the messages; IP addresses or other connection information; sender and recipient's mobile carriers; protocol headers, such as User-Agent strings which may contain device manufacturers and models; whether the message has an attachment; the URL on the content server where the attachment is stored; and the approximated size of messages.[32] Another potential problem with RCS is that phone companies could easily choose to block encrypted RCS for all users, or even just a specific user, thereby forcing them to downgrade to an insecure messaging protocol or not be able to communicate.

A final problem is that Messages by Google currently offers very little visual distinction between an SMS message, an unencrypted RCS message, and encrypted RCS. Google's implementation of RCS is admittedly far better than standard SMS, since the contents of messages can't be read by the phone company and it is possible to cryptographically verify the recipient. Regardless, Messages by Google shouldn't be relied on by users in need of a high degree of privacy and security, especially where the potential adversary is the state, law enforcement, or the phone company. However, RCS is still in its infancy and should see significant improvements in the coming years.

SUGGESTED IMPROVEMENTS

- Google should turn on RCS message support by default in Messages.
- Google should take steps to hide more metadata in encrypted RCS.
- Apple should support RCS in iMessage
- Google should make unencrypted messages more visually apparent.
- Apple could be more clear about the fact that blue chats are end-to-end encrypted while green chats are not within the UI and their messaging.
- Both Apple and Google could improve overall user safety by agreeing on an interoperable standard for encrypted messages. (Apple could keep iMessage with extra features, while still supporting encrypted RCS in a green chat bubble.)

32 https://www.gstatic.com/messages/papers/messages_e2ee.pdf

KEY VERIFICATION

Key verification, the process by which chat participants can mathematically verify that they are talking to the person they think they are talking to and there is not an eavesdropper or attacker-in-the-middle, is an important concept in cryptography and an important element of security in encrypted messaging. One significant drawback of key verification is that many users don't understand what its purpose is, why it's important, or even that it exists in the first place. There certainly exists room for improvement in how key verification is presented and communicated to users. The Signal protocol has made some strides in this endeavor, but there is still a long way to go.

Regardless, key verification is important for the safety of high risk users and should be considered a key feature of any "secure" chat program. All of the apps reviewed have the ability to verify account keys. Apps that use the Signal Protocol call this feature "safety numbers," allowing users to mark each other's keys as verified and warning users if the safety numbers of a contact are changed (such as when that contact gets a new phone).

Apple only introduced this feature as of 2023, calling it "Contact Key Verification."[33] The feature is not enabled by default, and has not yet been rolled out to all users. Telegram calls this feature "identicons," and lets users verify keys by comparing two images.[34] Images are by their nature harder to accurately compare than numbers, and this could lead to attackers being able to generate a key where the image is substantially similar visually, though no such attack has been demonstrated. Telegram also does not appear to notify users when a chat key or "identicon" changes.

SUGGESTED IMPROVEMENTS

- Telegram should switch to a harder to spoof key verification system.
- Apps should standardize on the terms it uses for key verification.
- Apple should enable key verification by default.
- More research should be done to determine better ways to communicate and present key verification to users.

33 https://www.apple.com/newsroom/2022/12/apple-advances-user-security-with-powerful-new-data-protections/

34 https://telegram.org/img/key_image.jpg

WHAT ISN'T ENCRYPTED

Even though messages between users are end-to-end encrypted, that doesn't mean that all data is end-to-end encrypted. In all the messaging apps we reviewed except for Signal, some metadata appears to be stored in a way that can be retrieved by app operators, and thus law enforcement agencies that seek it with a warrant. These types of data can include:

- Group metadata, including group membership
- Users' profiles
- Users' contacts and social graphs
- Users' IP addresses
- Source and destination of messages

While some of these apps will take steps to hide some of this information, in our analysis Signal is the only app that has taken steps to hide users' profiles, contacts, group metadata, and even message sender information by taking advantage of new techniques such as storing user information encrypted with the users' account key and only handling unencrypted data within a "secure enclave" (a part of the computer where not even the owner of the computer can see what data is being computed on) so that that information cannot be recovered without the users' personal key derived from their password.[35] [36]

SUGGESTED IMPROVEMENTS

Other messaging apps should borrow techniques from Signal and take steps to hide user metadata by keeping it encrypted with the user's account key and only handling unencrypted versions in secure enclaves.

NAMING MATTERS: CONFRONTING INCONSISTENCIES AND MISLABELS

While later sections in this report dive into this finding more in depth, it's important to highlight how labels, descriptions, and names deeply matter for cryptographic and security design, particularly in highlighting how users' build mental models and understand or misunderstand how security functions within the apps they are using.

For example, Telegram and Facebook Messenger label their encrypted chat functions as 'secret messages', while Telegram labels 'cloud encryption' as encryption. The discrepancy in labeling one encryption-based functionality explicitly but not the other can create user confusion. But, additionally, it is unclear what 'secret' means and

35 Many attacks on secure enclaves have been demonstrated; however it is a useful friction to keep information safe in the event of data breaches, law enforcement requests, and malicious employees.
36 https://signal.org/blog/private-contact-discovery/

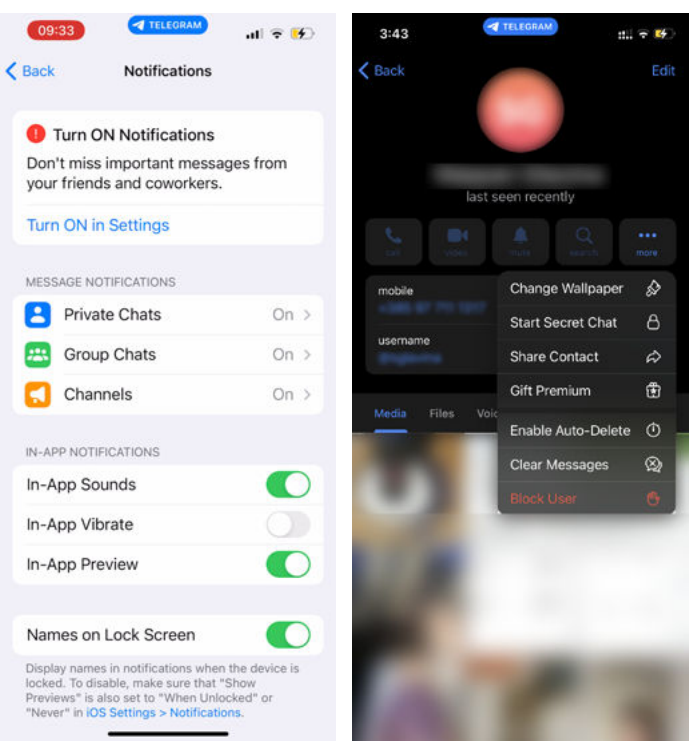
if it is linguistically best aligned with or describes safety, security, and encryption. In our field work users expressed similar confusion about what 'secret message' meant, and if it meant the app was safe and secure to use.

Additionally, in our design analysis of Telegram, we found that Telegram inconsistently labels UI within the app itself. For example, in Settings under 'Data Storage', Telegram has a label for 'private messaging,' but private messaging was not a feature we could locate. The logical assumption is that 'private messaging' and 'secret messaging' are the same or similar features. But such naming inconsistency can have real, downstream effects in terms of impacting user safety, and sow further confusion in users' mental models of Telegram, its features and its security and privacy functionality.

And, we found that two factor authentication is identified by different terms in each of these apps, leading users to confusion about whether apps have 2FA capabilities and whether they have been turned on. Signal calls this feature "registration lock," WhatsApp, Google, and Telegram call it "two-step verification," and iMessage and Facebook Messenger call it "two factor authentication." This inconsistency in naming is unnecessary and can make it harder for users to reason about security.

SUGGESTED IMPROVEMENTS

The developers and designers of encrypted messaging applications should agree on standard terms whenever possible.



An example of Telegram's "secret chat," showing settings with "private chat" and engaging with "secret chat" from a user's profile.

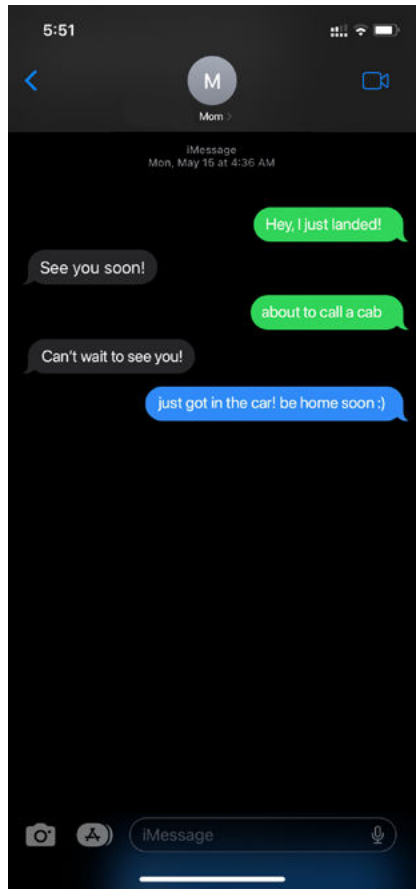
User Experience Design

User experience design refers to the layout, interaction design, and product features of an app or piece of software, including some guiding methodologies that determine why certain features are prioritized and utilized over others.

FORM FOLLOWS FUNCTION

Form follows function, or the idea that design should reflect what it does, is a popular design truism and principle,[1] and is reflected in design researcher Don Norman's description of affordances in his Six Design Principles.[2] The importance of an app's design and its seamless integration of "cues that allow people to figure out what an object does without any instructions" cannot be overstated, especially when considering its impact on security design and encryption. If a user does not adequately understand what a feature does and how that relates to their particular threat model, the user can be inadvertently exposed or led to make choices that compromise their safety.

For example, end-to-end encrypted messages in iMessage are blue and unencrypted SMS messages are green, but they are not labeled as such in the interface. In some cases, a singular message thread between two individuals will jump between 'blue' and 'green' messages. The difference between the two may remain unclear to users beyond vague annoyance or confusion at not having all of the same features available in a green message, such as stickers and reactions. In particular, the safety differences between blue and green messages are never made clear to users in the interface, nor is it clear to users how to ensure that encrypted conversations on iMessage aren't downgraded to unencrypted SMS messages.[3] In this example, the 'form' of the app is obscuring the functionality, and potentially creating user confusion that can compromise user safety.



An example of iMessage switching between unencrypted and encrypted messages in one text conversation.

1 https://medium.com/@AprilHQ/why-good-design-works-from-form-to-function-and-back-again-595885e7d257

2 https://crm.org/articles/the-godfather-of-ux-don-norman-user-centered-design

3 There is a setting in iMessage which allows users to send messages as SMS when iMessage service is not available but the security implications of this choice are not made clear to users who may not understand the safety and privacy differences between iMessage and SMS messages.

PRIVACY BY DEFAULT VERSUS PRIVACY AS AN AFTERTHOUGHT

While apps like Telegram and Facebook Messenger technically have more 'privacy features,' their security settings are not as robust Signal's. We explore this more in the Architectural Design section, but Signal, utilizing a privacy by design ethos, tracks no user data and thus has fewer features in which to track or monitor users' behaviors and data. Even Signal's Stories, a pro-social feature, are not stored and no data is gathered on usage. While there is a minimalism to modern app design which can be helpful in avoiding users' cognitive overload with features, Signal offers minimalism along with granularity in its security features; almost every feature can be turned off, adjusted, or customized. For example, Signal's disappearing messages feature is the only product to offer customization in terms of time, showing users granular options going down to five seconds. Telegram shows "One Month, One Week, 24 hours" for its 'auto-delete' messages option, with customization options that only serve to **elongate** the time span for disappearing messages, spanning from one day to one year. WhatsApp has many fewer customizable features, including for disappearing messages. It only offers 24 hours as the lowest time duration. Facebook Messenger ranges from 5, 10, and 30 seconds to 1, 5, 10, 15, and 30 minutes to 1, 6, and 12 hours to 1 day, with no customizable options.

SUGGESTED IMPROVEMENTS

Platforms should continue to prioritize privacy features and to implement more granular customization options, along with shorter and customizable disappearing message times. The default choices should, at a minimum, allow for a variety of selections, from a few seconds to a few minutes to a few hours. These defaults could be similar to Signal and Facebook Messenger's deleted message time range settings.

GRANULARITY AND USER AGENCY IS A SAFETY FEATURE

Granularity which allows for user agency is itself a necessary safety feature that allows a user to be able to tailor a setting for their particular use case or threat model. Signal, even with fewer product features, still offers more customization, and granularity within its settings, such as custom disappearing message times, while not overloading its app with confusing user journeys and unnecessary pathways. Telegram does offer a wide variety of privacy enhancing features, but a user must dig deep into the privacy and security settings to find these features and then turn them on.

Features that Could Compromise Safety

In our technical review and design review, we found a variety of features that could potentially compromise users' safety.

SUGGESTED IMPROVEMENTS

Companies need to allow for any feature that impacts privacy and security to be turned on and off, and to explore and implement more granular settings that allow for users, especially high risk users, to tailor app functionality to their needs.

THE STORING AND DELETING OF CALL LOGS AND INTERACTION HISTORY

There was no consistent way to delete call logs across apps, and no straightforward way to completely remove them from the apps themselves. When a user places a call via a messaging app installed on an iPhone, that phone still shows individual phone calls in the device call history. Call logs still appear in their Signal app, even if the records of calls themselves are not stored by Signal on their server. This does not mean that the call is compromised, it just means that it is visually stored in the user's local phone history.

On one hand, this makes sense if we follow the design ethos that 'design should reflect what it does.' Showing phone calls placed as a history of interaction in Signal is a form of truthful design. While Signal has the ability to turn off call logs from appearing in the phone's call history, Signal still stores a history of interaction within the UI of the app (again, this information is not stored as data on a server, it's just visually shown within a user's app). The only way to remove this interaction history is to delete the chat session and start a new one.

In contrast, apps like WhatsApp and Telegram easily allow for the user to delete their own call history within the app, but to our knowledge, the calls are still present in the device's call log. This is an acute example of design not showing what it's doing; as the actual call logs are still preswent in the phone; the call history is only really deleted from the visual interface on the device, which could lead to unintentional security compromises for the user. Users can manually delete call logs from their device, but WhatsApp and Telegram

should illustrate or explain this two-step process for users. Relatedly, there is not a way to mask or hide a WhatsApp or Telegram call in the device's call log; Signal does allow for this masking.

Another issue we found related to **deleting and removing messages**. Signal, WhatsApp, iMessage and Telegram offer ways to delete individual messages from both devices for a short period of time after they are sent. Facebook Messenger offers the ability to delete a message if the recipient hasn't seen it yet. Telegram is the only app to offer the ability to delete and clear all messages for both users without deleting the chat thread itself. Signal allows for the entire chat thread to be deleted, resulting in every message and interaction history to be removed, but only on one user's side. The only way to remove every message in a Signal or WhatsApp thread is to either have both sides delete the chat, or to have turned on disappearing messages so the messages will automatically be deleted from all devices.

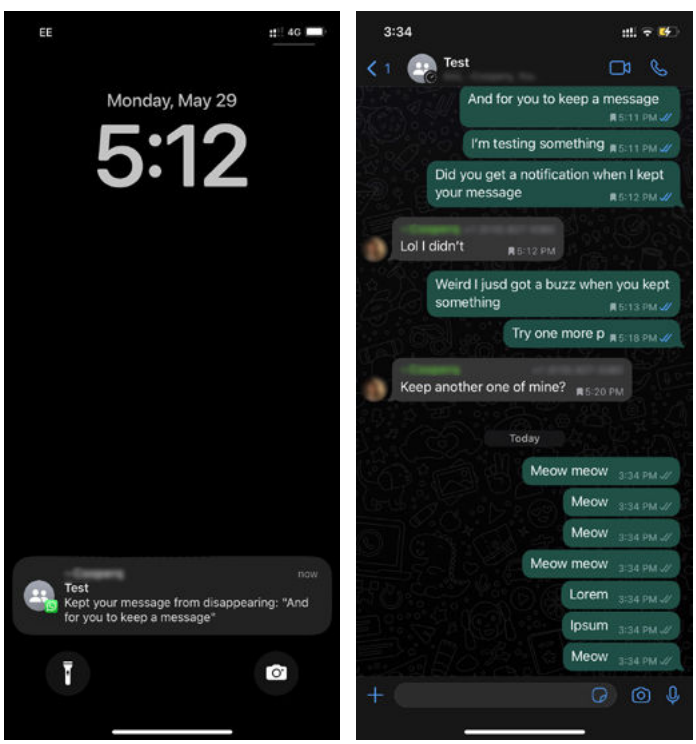
Lastly, a potential vector for harm we documented is **WhatsApp's 'keep' messages feature**. Recently, WhatsApp launched a new feature called "keep" that can only be accessed when disappearing messages are turned on. "Keep" allows a message to be 'kept' from being deleted in both one-to-one and group messages. Any user can keep any other user's message, and it seems any user can 'unkeep' anyone's kept message.[4] WhatsApp shows which users have kept which messages, and this is visible to every group member.

Part of the potential harm of this feature is that there is little feedback given to the user, at least as observed in two user testing sessions with users based in the UK, the US, and India. In the first test, only once was one of the testers notified with a notification banner of a 'kept' message; every other notification was a haptic vibration which was identical to the general notifications' vibration. In the second test, two out of three users were occasionally and inconsistently notified via a notification banner, but only when the app was closed. If the testers had WhatsApp open, there was no notification banner. Notification banners only appeared on mobile; the users received no visual notifications on the desktop client. Generally, there were no other visual acknowledgements, pop-ups or notifications about a message being 'kept'; the only indication is a small gray bookmark icon within the kept message itself, along with a haptic vibration.

While the 'keep' option may be useful in some social interactions, especially given how easily it can be undone, it

4 A note of caution: we tested this twice, in three different countries. The first test was with two US-based users and one user located in India on May 9, 2023 across three iPhones and two MacBook Pros. The second test was on May 29, 2023, with two US-based users and one user located in the UK, across three iPhones and two MacBook Pros. As mentioned above, the notifications were inconsistent, which could be the result of a bugs in this new feature.

does pose some security risks. Given that the shortest disappearing message WhatsApp allows for is 24 hours, and that some WhatsApp groups can be extremely large, the lack of system feedback or notifications around 'keep' messages makes it a risky form of design, particularly with the inconsistency and occasional lack of a notification banner. While it may be true that, in a group chat where not all participants are known and trusted, there are other avenues for bad actors to subvert disappearing messages, such as simply taking screenshots, the lack of a clear notification may encourage security-compromising behavior.



The image on the left is an example of WhatsApp notification banner related to 'keep messages'; the image on the right shows a group conversation where some messages have been 'kept' and other messages have not.

SUGGESTED IMPROVEMENTS

Developers should make it easier to quickly delete entire message threads and interaction histories between users and in groups across all apps. WhatsApp, Facebook Messenger, and Telegram should allow for phone calls to not appear in a device's call log, similar to a functionality that Signal allows. WhatsApp should implement (and turn on by default) banner and visual notification settings for "keep" messages across mobile and desktop.

PHONE NUMBERS AS USER IDENTIFIERS

Phone numbers are considered sensitive, personally identifying information, but almost every app we reviewed treats them slightly differently. For example, iMessage users can message via an Apple ID tied to an email and/or phone number; in some cases, just the email is presented to other users when messaging, but choosing which one is exposed isn't always easy for users to figure out. At present, only Telegram offers the option to hide one's phone number in favor of a username. While Signal and WhatsApp currently lack this feature, both appear

to have plans to implement it soon.[5] [6] Facebook Messenger operates without using a phone number, even though its sign up process negatively nudges users to share that information via a dark pattern we discuss below (however, users can then go and delete that phone number and/or hide it from other users). Messages by Google currently only supports messaging users with a phone number. Overall, it seems that developers have heard requests (Signal especially) from activists to allow messaging identifiers other than phone numbers. This important change should be implemented urgently.

WHEN FEATURES AND SECURITY ARE IN CONFLICT

Some features that are core parts of the user experience in messaging apps have security or privacy implications. Often these features can leak metadata but some can reveal message contents to unintended recipients or even leave messages available for recovery by law enforcement.

LINK PREVIEWS

Link previews are one such vector of information leakage. Link previews work by fetching the URL a user is sending to others and parsing some HTML tags which contain an image and description of the website the user is linking to. Link previews can be generated either on the device (Signal, WhatsApp, iMessage) or over a proxy service (Facebook, Google, Telegram.) Some privacy issues in generating link previews have been discussed in prior work, but those do not precisely apply to any of the apps we tested and other issues, such as having link previews generated by a proxy server, were observed.[7]

There are other important privacy and security issues raised by link previews. For apps where link previews are generated from the sender's device, the owner of the website whose URL was linked can see the specific IP address of the person that is sending the link, and what app they are sending it from. If that URL is unique, such as when there is a tracking beacon included in the URL, the owner of a website could determine the IP addresses of who that link is being shared with on that app as well and start to make social graphs of a targeted Signal or Whatsapp user. But using a proxy to fetch the data, as Telegram and Facebook do, is not much of an improvement. In this case the proxy operator–namely Facebook Messenger, Messages by Google, or Telegram–will know exactly what links a user is sending, breaking one of the core principles of encrypted messaging. Nor do proxies solve the problem of mapping a social graph, since the contacts a user shared the URL with will also presumably click

5 https://community.signalusers.org/t/usernames-in-signal/9157/1004
6 https://www.engadget.com/whatsapp-may-soon-introduce-usernames-105558183.html

7 https://www.mysk.blog/2020/10/25/link-previews/

on it. Thus, the web operator could still start to piece together who is sharing items with whom, and the only thing the proxy will protect is the IP address of the user that initially shared the link, though presumably the web operator will have access to that information from visitor logs in any case. Where it is safe, using a trusted VPN could help mitigate this concern, if all members of a group do so.

One core issue here is tracking parameters on URLs: if a user sends a link to a very popular tweet, for instance, that link is still unique because of the `t` parameter at the end of the URL.[8] Stripping similar tracking parameters on even a few popular websites would go a long way towards protecting users who choose to leave link previews enabled.

Visiting a URL and sharing a URL may also be different under the eyes of the law, in that one may display intent. Putting this into context, consider a hypothetical scenario: if the state of Texas were to order a website that contains instructions for self-managed abortions to turn over its traffic logs, Texas law enforcement might then look for any Texas IP addresses that were sharing content over Signal, use these logs to then locate or subpoena the owner of said IP address, and then charge the owner of that IP address with sharing illegal content, whereas simply visiting that content might not be indicative of a crime.

8 For instance, https://twitter.com/dril/status/1593408939651043328?s=46&t=7H3nU0Y_-E4ofNNFwxhkaw

SUGGESTED IMPROVEMENTS

- We recommend that high risk users disable link previews immediately, especially if using an app in a country where that app is forbidden (such as Signal in Iran).
- WhatsApp should provide an option to disable link previews, as Signal already does.
- For encrypted messages, link previews should be off by default and potential risks should be better communicated to users consistently and over a sufficient period of time to ensure that as many users as possible understand the risks. This could and should be communicated via UI design, along with warning labels (potentially as pop ups or interstitials), and blog posts.
- Apps should endeavor to strip tracking parameters in the client for as many popular websites as possible. This is not a perfect solution, but it would help mitigate the problem.

GIF SEARCH

GIF searches are another potential avenue of harm. GIF searches in WhatsApp and Facebook use Tenor, a service run by Google. These create a record of the IP address of the searcher, what is being searched for, and presumably what app was used for the

search.[9] Google knowing what a user is searching for in WhatsApp is probably not a high risk in many users' threat models, but it is an important risk vector for some users to be aware of. Signal avoids this problem by running a transparent proxy for GIF searches in Giphy. This ensures that Signal knows who a user is but not what the user is searching for, and that Giphy knows what the user is searching for but not who the user is.[10]

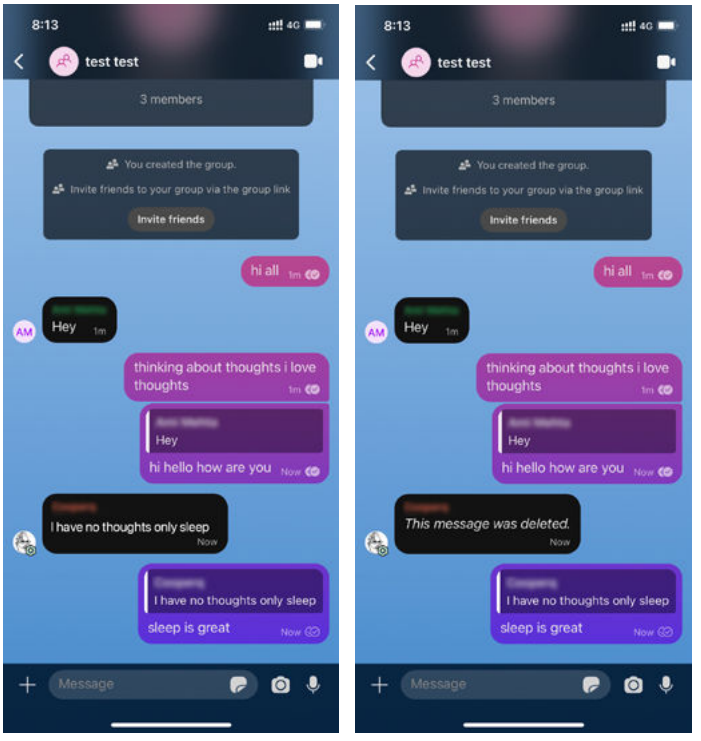
Suggested improvements: Apps should proxy GIF searches in a privacy preserving way, possibly using a transparent proxy to decouple the user's identity and what they are searching for.

SUGGESTED IMPROVEMENTS

Signal and WhatsApp should delete quoted text in replies to a message when the original message has been deleted.

QUOTE REPLY TO DELETED MESSAGE

Alexis Hancock, a security researcher at EFF, recently discovered that in both Signal and WhatsApp, if a message has been replied to by quoting the original message, then the message text is still visible in the quote even when the original message text has been deleted.[11] This behavior may subvert users' expectations that when a message is deleted "for everyone" that all copies of it are deleted. A quote message that preserves the text of a user's message after it has been deleted could be a severe privacy risk, such as for someone in an abusive relationship whose partner has access to their phone.



In Signal, a deleted message still appears in a quoted message.

9 https://policies.google.com/privacy?hl=en
10 https://signal.org/blog/signal-and-giphy-update/
11 https://www.eff.org/deeplinks/005/how-do-different-encrypted-messaging-apps-treat-deleted-messages

BACKUPS

Unencrypted backups of messages are an extremely high risk feature. According to a leaked FBI document, law enforcement agents have gained access to the contents of encrypted chats via unencrypted backups of iMessage and WhatsApp stored on iCloud and Google cloud.[12] Although iMessage and WhatsApp backups are not encrypted by default,[13] users can and should turn on encryption for their backups in the WhatsApp settings and for iMessage by enabling advanced device protection. WhatsApp is also extremely persistent about nudging users to turn on backups while not explaining the associated risks. Telegram and Signal avoid this problem by not backing up secret chats in Telegram's case and only having encrypted and locally stored backups for Signal.

SUGGESTED IMPROVEMENTS

WhatsApp should encrypt backups by default. iMessage should warn users if backups are enabled but advanced device protection is not turned on.

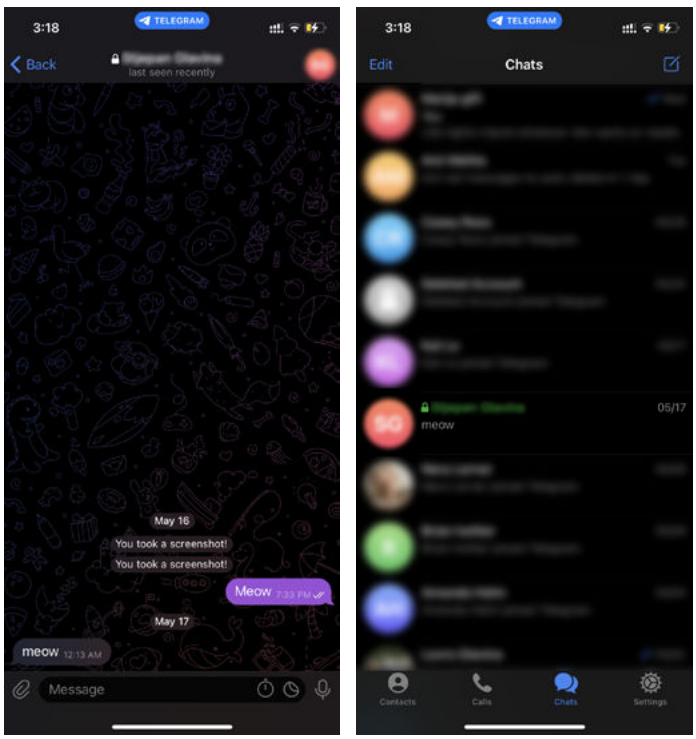
12 https://twitter.com/PropOTP/status/1466159714610212868/photo/1
13 https://www.forbes.com/sites/zakdoffman/2021/12/04/apple-iphone-ipad-mac-icloud-warning-as-dangerous-settings-exposed/?sh=4e6ebd4e7e72

DOES "SECRET" MEAN SECURE?

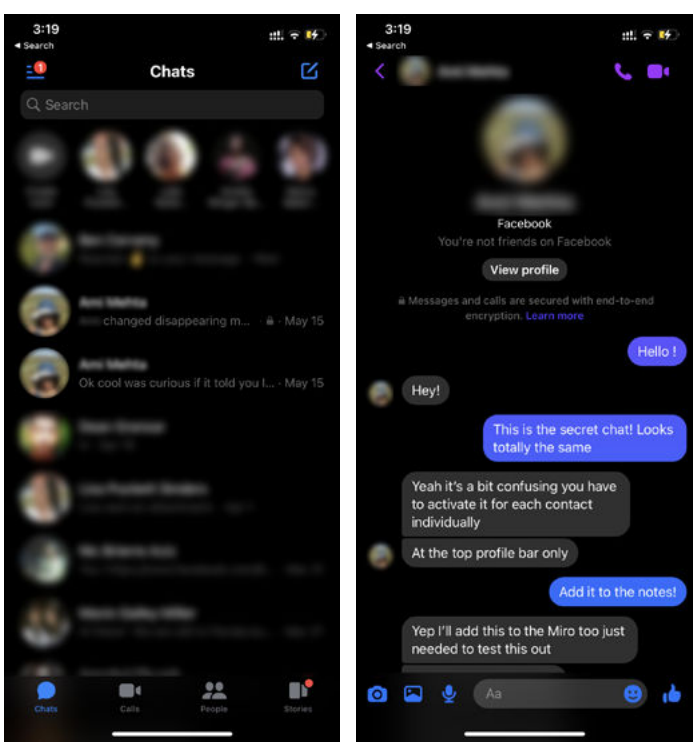
Encrypted messages for Telegram and end-to-end encrypted messages for Facebook Messenger were turned off and can only be accessed via 'secret message' chats. "Secret message" chats pose a variety of harms. In regards to UI, "secret messages" look nearly identical to regular chat messages. "Secret messages" have to be specially accessed; via selecting a user's profile and then selecting 'secret message' for both Facebook Messenger and Telegram. Relatedly, the word 'secret' does not adequately present 'secure', 'safety', 'privacy' or 'security'; this misnaming could be misleading and confusing for customers, especially given Meta's push to advertise privacy and security in its platforms. Having the most secure options turned off by default and named 'secret' creates unnecessary friction and a lack of safety for users. Additionally, it raises the potential that users might have thought they were engaging in 'safer' messaging by using these two apps, when in fact they were not.

SUGGESTED IMPROVEMENTS

Both Telegram and Facebook Messenger should make all messages encrypted by default, and at the very least ensure the naming of features matches up with what the feature does. Encrypted messages aren't 'secrets,' per se, and naming conventions should reflect that.



The image on the left illustrates Telegram's secret chat message; the image on the right illustrates what a secret message looks like in Telegram's messages home screen.



The image on the right illustrates what a secret message looks like in Facebook Messenger's messages home screen; the image on the left illustrates Facebook Messenger's secret chat message.

FEATURES THAT FACILITATE PRIVACY

There are a variety of features that facilitate privacy and help mitigate harm; some of these features, luckily, are universal across apps, like the ability to mute and block contacts, groups, etc. Some of the features mentioned here are not intended to be perfect solutions, but rather, are small choices that can be helpful in enabling users to enhance their safety. Safety is a spectrum, not a binary state, and may change quickly depending upon a a range of factors. A variety of features can help facilitate privacy, and function as necessary building blocks; for example, disabling the ability to take screenshots or sending a notification that a screenshot has been taken does not stop someone using another device to photograph the message, but it does add friction or alert the user to a potential threat. These design features will not necessarily stop a bad actor, but any feedback that lets a user know something has happened gives that user time to react, make a plan, or take action, and thus such features can be useful in mitigating harm.

Key privacy enhancing features we identified in the process of design analysis include:

- Disappearing messages with custom times on Signal and Telegram.
- Signal's overall granularity with regard to the ability to turn on and off all features, including standard features like link preview, stickers, etc.

- Facebook Messenger's notifications of screenshots being taken.[14]
- WhatsApp's functionality for thwarting or disabling screenshots for some messages.

SUGGESTED IMPROVEMENTS

Developers should allow for customizable disappearing message times; increase granularity for all product features that impact security and privacy; allow for relevant features to be disabled; universal notifications of screenshots; and the ability to disable screenshots. Lastly, in the Signal app, the 'delete for me' option sits above 'delete for everyone' in the option menu to delete messages. Users we interviewed alluded to accidentally hitting 'delete for me' when meaning to select 'delete for everyone.' We recommend flipping the order of these two buttons.

14 App developers may push back on this as being not perfect since there are other ways a screenshot could be taken, such as with another camera. Nevertheless, these notifications present a useful friction which discourages bad behavior.

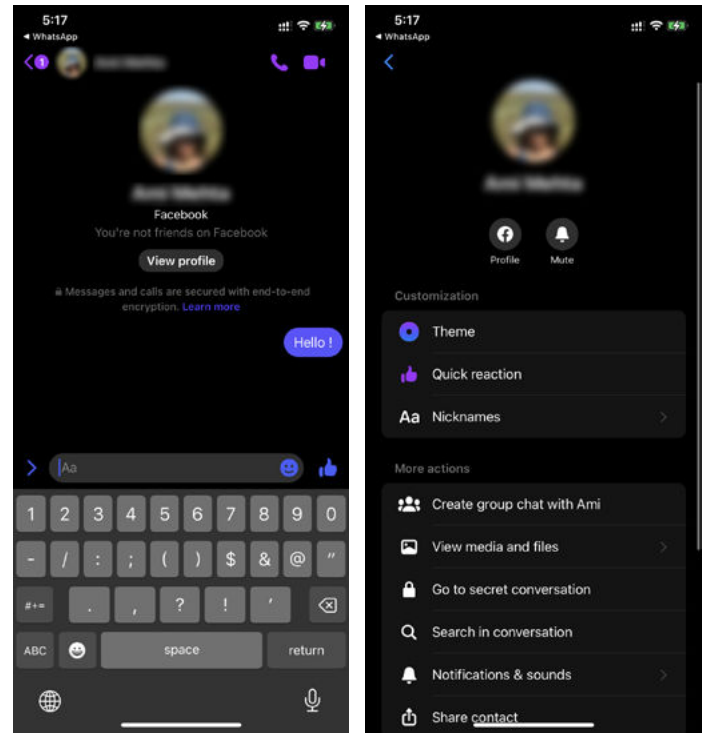
DAMAGING, DECEPTIVE AND EXPLOITATIVE DESIGN IN FACEBOOK MESSENGER

Dark patterns or 'damaging design' are design patterns that unintentionally or intentionally confuse, manipulate, deceive or exploit users into making decisions they normally would not make.[15]

(While there are legitimate objections to the phrase 'dark patterns,'[16] it has entered the cultural and regulatory lexicon,[17] including in recent regulation such as the Digital Services Act and Digital Markets Act. The United States[18] and Germany[19] have passed specific legislation to curb their harmful effects. These particular design harms exist across all technology domains and can be found in many products, such as in subscription cancellation flows,[20] cookie banners,[21] account sign up patterns,[22] and even in privacy and security related design.[23])

15 https://www.stiftung-nv.de/en/publication/dark-patterns-regulating-digital-design
16 https://medium.com/@carolinesinders/whats-in-a-name-unpacking-dark-patterns-versus-deceptive-design-e96068627ec4
17 https://medium.com/@carolinesinders/whats-in-a-name-unpacking-dark-patterns-versus-deceptive-design-e96068627ec4
18 https://www.theverge.com/2021/3/16/22333506/california-bans-dark-patterns-opt-out-selling-data
19 https://www.lexology.com/library/detail.aspx?g=612f8a61-5225-43f8-ace3-2a02d5278c6e
20 https://pudding.cool/2023/05/dark-patterns/
21 https://dl.acm.org/doi/fullHtml/10.1145/3411764.3445779
22 https://medium.com/@daily.design/dark-ux-pattern-forced-registration-564d43feefe8
23 https://www.reuters.com/legal/legalindustry/dark-patterns-new-frontier-privacy-regulation-2021-07-29/

The decision to name Telegram and Facebook Messenger secure chats as "secret chat", in conjunction with marketing messages around privacy from the companies that may mislead users, the many steps needed to enact secret chats, and the UI that makes it difficult to distinguish a secret chat from a regular chat, all could be regarded as a dark pattern. This type of confusing design is described by the European Data Protection Board (EDPB) as 'fickle' or 'left in the dark', where the language and interface design are not clear, and/or are inconsistent.[24] The naming of what a security feature does and how to access it is very important. In this case, 'secret messages' are buried and are not default options, with no way to turn on encryption for all messages. Instead, a user must know what a 'secret' message is and that it is buried underneath an individual's profile, often hidden under an additional menu. If Facebook Messenger completes its transition towards end-to-end encryption as default, this issue may be obviated on that platform, pending how it is implemented.

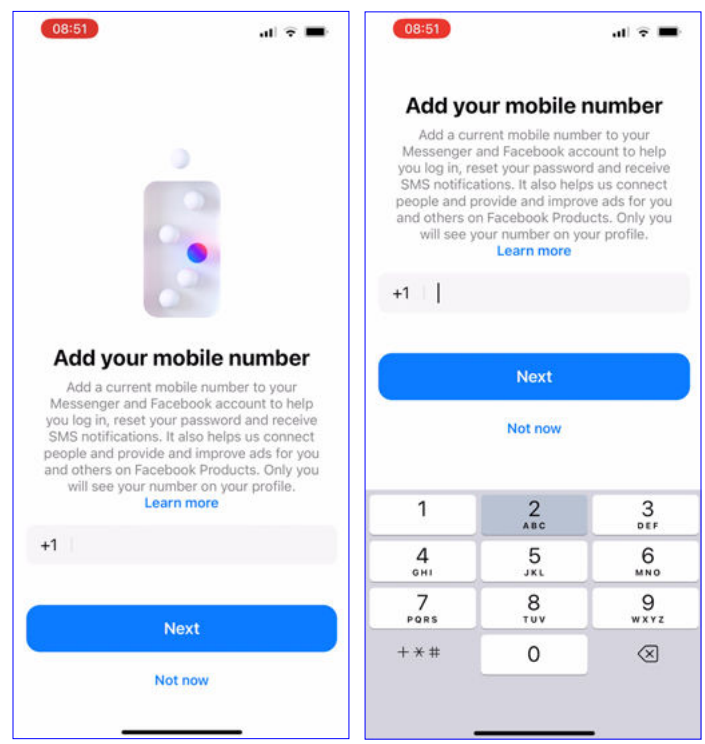


The steps to activate Facebook Messenger's secret chat. A user has to click on another user's profile (image on the left) to then select 'secret conversation' from a list of actions (image on the right).

The other potential dark pattern we identified is within the sign up process for Facebook Messenger. The app requests a user's personal phone number, when the number is not needed to use the app. This is an example of what the Norwegian Consumer Council describes

24 According to the EDPB, "fickle" means the design of the interface is inconsistent and not clear, making it hard for the user to navigate the different data protection control tools and to understand the purpose of the processing. The following two dark pattern types fall into this category: "lacking hierarchy" and "decontextualising." "Left in the dark" means an interface is designed in a way to hide information or data protection control tools or to leave users unsure of how their data is processed and what kind of control they might have over it regarding the exercise of their rights. The following three dark pattern types fall into this category: "language discontinuity," "conflicting information" and "ambiguous wording or information." https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-32022-dark-patterns-social-media_en

as a 'deceptive click flow',[25] in which an app asks for more data than is necessary and represents the request as one legitimately needed to use the app. Additionally, how the request is visually rendered reveals another potential dark pattern, sometimes referred to as 'choice disparity', or 'interface interference'[26] in which a choice is grayed out,[27] or one choice is favored over another. In this example, Facebook Messenger emphasizes the 'next' button as a blue button, with the 'not now' link rendered as floating blue text against a white background. Nowhere in the previous screens or following screens does the app clarify that a user's phone number is not needed to use the app. Selecting 'not now' allows the user to move forward in the process without submitting their number, but this is never clarified for the user. Instead, the design and copy in the app seems to imply a user's phone number is needed.



Screenshots from Facebook Messenger's sign up and account activation flow; Facebook Messenger is requesting a phone number but does not specify that a phone number is not needed to use Facebook Messenger.

25 The "deceptive click through" in that context is "the click-flow when setting up an Android device pushes users into enabling 'Location History' without being aware of it. This contradicts legal obligations to ask for informed and freely given consent" (Norwegian Consumer Council's 2018 report,"Every Step You Take." https://storage02.forbrukerradet.no/media/2018/11/27-11-18-every-step-you-take.pdf. While the cited example is Android specific, it highlights an important design function in which the app asks for unnecessary information but presents it as necessary, potentially violating data protection law in various European countries.

26 From Gray, Kou, Battles and Hoggatt's "The (Dark) Patterns Side of UX Design." Interface interference is defined as "any manipulation of the user interface that privileges specific actions over others, thereby confusing the user or limiting discoverability of important action possibilities (cf., false or hidden affordances). Interface interference manifests as numerous individual visual and interactive deceptions, and is thus our most involved strategy with three subtypes: hidden information, preselection, and aesthetic manipulation." https://www.researchgate.net/publication/322916969_The_Dark_Patterns_Side_of_UX_Design

27 The U.S. Federal Trade Commission paper "Bring Dark Patterns to Light" describes this similar pattern of choice architecture in which the company "highlight[ing] a choice that results in more information collection, while greying out the option that enables consumers to limit such practices" is a form of "Design Elements that Obscure or Subvert Privacy Choices." While this option is not 'greyed out', it is de-emphasized, which is similar to the action of 'greying' out informational choices. https://www.ftc.gov/reports/bringing-dark-patterns-light

Architectural Design: Social Network Features and Surveillance Capitalism UI

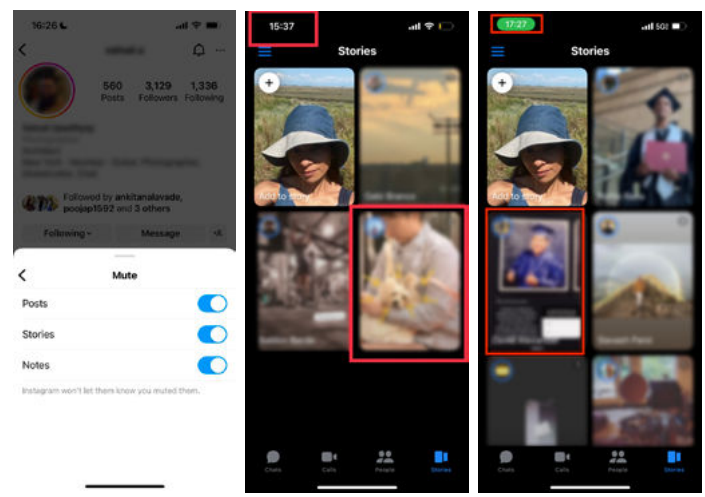
We use ‘architectural design’ metaphorically and descriptively in this section to refer to a variety of factors that underpin a messaging app’s feature load, technology stack, and product identity, along with the infrastructural or overall architectural design. Conceptually, this differs from the ‘user experience design’ section, which focused on specific user experience and user interface product design choices. The concept of architectural design affords a broader analysis of the apps, highlighting tensions such as how surveillance capitalism has impacted app design and how certain messaging apps aim to be much ‘larger’ than just a messaging app, even functioning akin to a social media network or platform.

SOCIAL NETWORK OR APP? WHY TELEGRAM AND FACEBOOK MESSENGER'S ARCHITECTURE IS SO BLOATED

We found that both Telegram and Signal possess nearly double the number of features in WhatsApp, iMessage, and Signal. At first blush, this could be interpreted as a good thing; after all, aren't more features a sign of more user agency? Additionally, some of those features are specific and nuanced privacy-related features.

However, the sprawling size and scale of both Telegram and Facebook Messenger ultimately raise safety concerns about these apps. This number of features may make sense for a social network that has different services and surfaces, such as a timeline, groups or communities, as well as group messaging and private messaging; but for messenger services, these choices may confuse or overwhelm users, especially with almost all default settings on both apps being set to the most public, and not the most private. Both Telegram and Facebook Messenger allow for large community groups to engage and message, not unlike WhatsApp, however Telegram makes it easy to find new users and join new groups via its Global Search function. This kind of ease of searching, joining, and interacting with strangers in large scale groups is also akin to social network and streaming platform design, like Discord or Twitch.

Facebook Messenger has multiple sharing settings that allow the app to connect across multiple Meta products; these settings have the ability to make Facebook Messenger the underpinning, main-communication app for a user's Facebook *and* Instagram profiles. Meaning, Facebook Messenger is designed to allow a user to communicate with friends and other users across both platforms. While this feature 'off' by default, it helps explain the relatively large size of Facebook Messenger, and suggests that it is an architectural tool to link multiple apps and profiles with interconnected settings, giving Facebook Messenger the ability to post stories, chat with communities, access Facebook Marketplace chats, and create rooms.



Instagram stories appear in Facebook Messenger from accounts this user previously muted on Instagram.

The connection to so many additional services raises privacy safety concerns. One of our researchers discovered during testing that a profile they had muted on Instagram kept reappearing in the Facebook Messenger stories, regardless of their Instagram settings. By using Facebook Messenger as a 'connector' between these two distinct apps, Instagram and Facebook Messenger itself, it appears that individual app settings are not universal and thus users are exposed to content and accounts they might have muted, blocked, or filtered out on the other platform.

SURVEILLANCE CAPITALISM AND PRIVACY BY DESIGN: CAN THEY CO-EXIST?

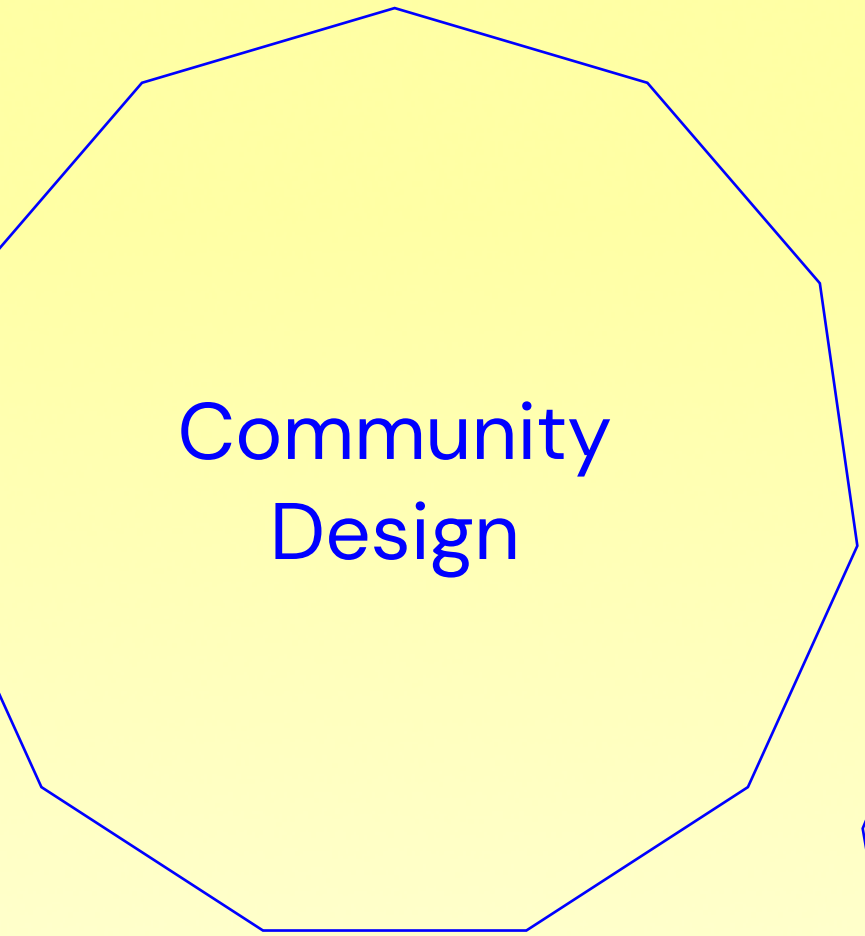
Feature bloat and the interconnected nature of apps like Facebook Messenger are arguably driven by the incentives of surveillance capitalism, which necessitate the centralization and maximization of data collection. Especially in the context of secure messaging, unnecessary features have the result of making potential security threats harder to recognize and understand. Many of the 'standard' features of applications that thrive under surveillance capitalism, such as knowing what a user is doing and when and where they are doing it, can put users at risk, particularly the most vulnerable users.

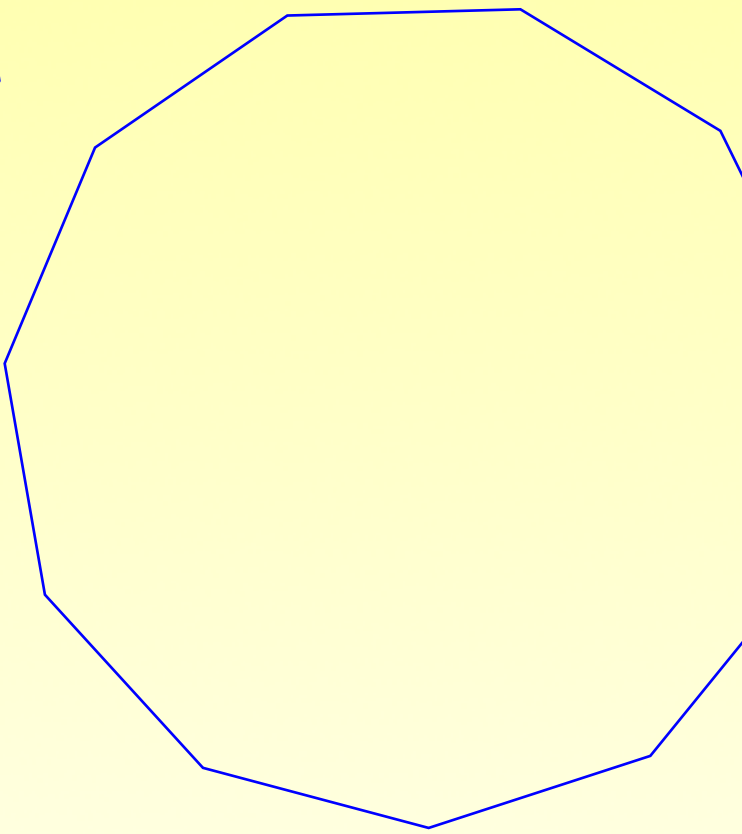
Over years of providing security training and support to online harassment victims, members of our research team have observed that often people do not realize how exposed they are when using certain apps. Features like 'location tracking' may be turned on, or a feature like 'when you were last online' can send a dangerous cue to persistent harassers or stalkers. Facebook Messenger affords this ability to see who is online, while Telegram offers the ability to see who is nearby, even by pulling from a user's contacts. Such features can result in real world harm.

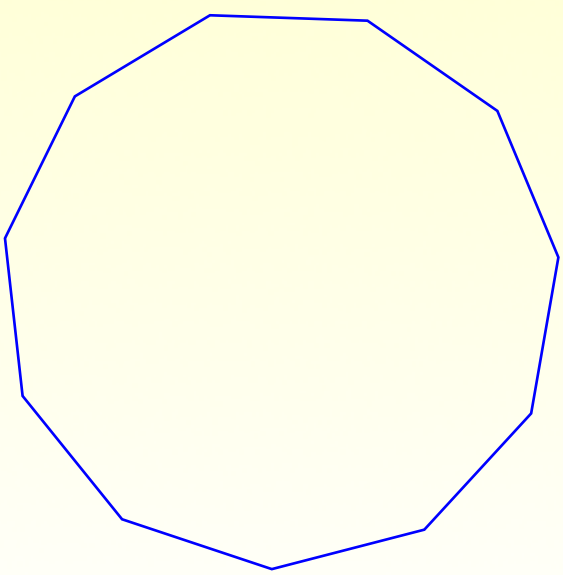
Some of these features are reminiscent of the early social web, like Foursquare check-ins, and Facebook's more granular and immediate updates. The 'last seen online' feature, for example, is a direct descendent of AOL Instant Messenger, in a time predating smartphones, when users could only access messaging apps and email via a desktop. Perhaps 'last seen online' was a useful feature over two decades ago, but it should now be obsolete. While it is commendable that Telegram and Facebook Messenger allow for some of these features to be turned off, the creation and inclusion of these features alone demands analysis, since it would appear that these features are not needed to improve messaging and create potential safety liabilities.

Signal, Whatsapp, and iMessage lack all of the above features because they are first and foremost messaging applications, and are not primarily underpinning social media platforms. But WhatsApp does appear to be starting to dip a toe in the social media space with status updates and Communities; these new features could increase the amount of bloat in this app, as well, increasing the attack surface and making users less safe. Because Signal is a privacy focused app backed by a non-profit, it says it is committed to not creating features that ultimately track and measure user engagement in the same way as these other, for-profit products. Even Signal's new social-media feature, Stories, is still end-to-end encrypted. It is simple enough for users posting Stories to understand who is seeing them, and, importantly, it is a feature that can be turned off. The design of Stories is a good example of how to add social features without diminishing privacy or security.

Community Design







We uncovered interrelated themes of 'community design': how software, apps and services interact with their user communities, including open source contributors to Signal, and how the developers of the applications interact with or update individual users.

We considered a range of 'community design' components for each application, including:

- Analyzing how or if high risk users can contact the app, service, platform or company;
- Transparency reports: if the app, service, platform, or company releases such reports, where they are located, and how detailed they are;
- General documentation: what is the available documentation and how easy is it to find, access, and surface for all different kinds of users;
- How the app generally updates or interacts with users: such as through in-app messaging, emails, blog posts, or social media outreach;
- Vulnerability disclosures process and/or bug bounties: whether these exist for an app, service, platform or company; and if so, where can users contribute or find the disclosures and/or bounties.

HIGH RISK USERS CONTACTING THE COMPANIES, APPS, SERVICES OR PLATFORMS

We found that there is not a specialized way within any of these apps for high risk users to communicate with the app, service, platform or company beyond message support centers, help pages, and other general methods available to all users.[1] Some advice to users is offered in very general terms- for instance, searching "India" in WhatsApp's "Privacy, Security and Safety" pages does yield an entire page of recommendations for what users can do to be safe in India. But there is no method for a user to contact WhatsApp itself in the country.

TRANSPARENCY REPORTS

From our analysis, Signal is the only app to release the full text of government requests it receives and share them via a page labeled "Government Requests."[2] WhatsApp and Facebook Messenger's transparency report is rolled into Meta's general transparency report[3] and was linked to via the FAQ page. Similar to Meta, iMessage's overall transparency

1 Except possibly through certain NGOs, though this method of communication will not be obvious to many high risk users, and it requires them to know the NGO exists in the first place.
2 https://signal.org/bigbrother/
3 https://transparency.fb.com/data/government-data-requests/?source=https%3A%2F%2Ftransparency.facebook.com%2Fgovernment-data-requests

report is rolled into the general Apple Transparency Report,[4] which currently only goes through December 2021. Both Apple and Meta share high level information online such as how many government requests they received in a particular country, the high level kind and nature of the request, if the request involved data, etc. Meta provides much more information in their online transparency report than Apple; to get the same level of information, we had to download a PDF of Apple's transparency report. However, neither disclose explicitly what they turned over exactly in these individual cases. The best information we have about what each app turns over to law enforcement comes from an FBI training document[5] which was released to the nonprofit watchdog group Property of the People through the FOIA process in 2021.

GENERAL DOCUMENTATION

Every app we analyzed has its own social media account, save for iMessage. Signal and WhatsApp have blogs linked via their websites, while Telegram has a blog-like section on its website entitled 'Recent News'; these pages serve as updates on recent features, announcements or important updates. However, the app with the most documentation, and spaces for users to interact, is Signal. Signal shares its GitHub, where its developers post real time code updates and where contributors can contribute to Signal's code base, or just follow along on Signal's progress.[6] Signal also has a community forum on its website where any user can sign in and post inquiries or questions they have, and see updates and announcements from Signal.[7]

HOW THE APP UPDATES OR INTERACTS WITH USERS

Overall, an app's documentation, at times, bleeds into how the app updates, interacts, or communicates with users. As stated above, all the apps have individual social media accounts, save for iMessage, and use those social media profiles along with their blogs or news sections on their websites to update users. Companies like Meta and Apple also use press releases to update users on major changes, new features, or new functionality. However, Telegram is the only app, at the moment, to have a special channel in its own app to communicate with users on updates and new features; occasionally, this same channel will nudge users to turn on certain features, like two factor authentication.

4 https://www.apple.com/legal/transparency/
5 https://propertyofthepeople.org/document-detail/?doc-id=21114562

6 While "just look at the commit logs on GitHub" is a standard in open source projects to generate transparency amongst the community, this is not a great way of updating users about new features, since most users are unlikely to know what GitHub is or how to use it to read commit logs. This therefore creates barriers to legible and accessible transparency, especially for users who do not have a technical background.
7 https://community.signalusers.org/

VULNERABILITY DISCLOSURES PROCESS AND BUG BOUNTIES

Having a vulnerability disclosure process is considered a security baseline for any modern software company, especially one making software for security and privacy. Bug bounties have also become an almost de-facto standard for incentivizing security researchers to help find and report security problems.

Signal has a dedicated email to report vulnerabilities[8] and a decent track record of disclosing security issues,[9] but no bug bounty program. Meta,[10] Apple.[11] Telegram,[12] and Google[13] all have robust bug bounty and vulnerability disclosure programs. However, Telegram has recently failed to respond to or address some potentially severe vulnerabilities.[14] [15]

SUGGESTED IMPROVEMENTS

Community design is important in ensuring users understand new features, or new flaws and security risks, and what to do in response to those findings and updates. Ideally, users should be able to find this information easily within the app without having to read a company blog, or GitHub commit logs. Users need to know in real time about pertinent information from the app, service, company, or platform; for some users, this is a matter of physical safety.

We recommend making transparency reports easier to find within each messaging app itself, and on the specific product's main webpage. iMessage, WhatsApp, and Facebook Messenger need to release their own transparency reports, independent of their parent companies. Signal should be more proactive in its public communications, including plans for new features and updates. And, all encrypted messaging apps should prioritize the development of fast and secure ways for vulnerable and high risk users to communicate with the app, service, company or platform.

8 https://support.signal.org/hc/en-us/articles/360007320791-How-can-I-report-a-security-vulnerability-
9 https://support.signal.org/hc/en-us/articles/4850133017242-Twilio-Incident-What-Signal-Users-Need-to-Know-
10 https://www.whatsapp.com/security/advisories?lang=en_US
11 https://security.apple.com/bounty/categories/
12 https://core.telegram.org/bug-bounty
13 https://bughunters.google.com/about/rules/6625378258649088/google-and-alphabet-vulnerability-reward-program-vrp-rules
14 https://danrevah.github.io/2023/05/15/CVE-2023-26818-Bypass-TCC-with-Telegram/
15 https://twitter.com/telegram/status/1658492425315430401

Technological Threats to the Promise of Encryption

CLIENT-SIDE SCANNING

A variety of voices keen to identify harmful material such as child sexual abuse material or terrorist message content are enthusiastic about the potential of client-side scanning, a technology that generally refers to mechanisms that can scan content for offending material on a user's device, only alerting security agencies or content moderators if there is a match. But experts recognize that not only are these systems fragile, they create serious privacy and security risks, ranging from false positives to backdoor access to messages before they are encrypted, opening the possibility of unsanctioned monitoring of content and conversations.[1][2]

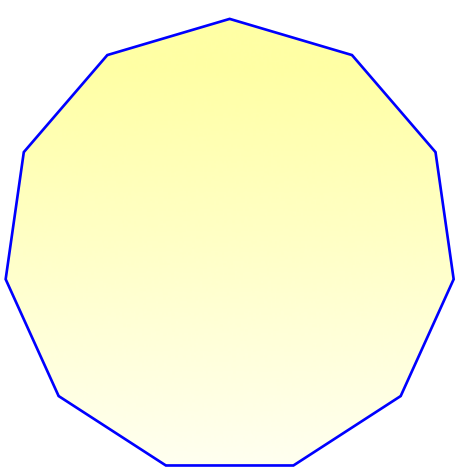
QUANTUM COMPUTING

Quantum computing and post-quantum cryptography are both fields which are steadily advancing but which are still far from being relevant concerns for the majority of industry and for encrypted messaging. The best known quantum computers currently have around 433 quibits (quantum bits) while the best known quantum algorithm for breaking classical encryption, "Shor's algorithm" would require a computer with at least 3072 total quibits to execute, though some cryptographers argue it would possibly require millions of quibits.[3][4][5] Expert cryptographers we interviewed for this report agreed that quantum computers shouldn't be a problem for at least 10-25 years and were generally unconcerned about quantum cryptanalysis. One cryptographer interviewed asked, "Do you think the messages you're sending now will still need to be secret in ten years when quantum computers come? If so, you have already lost." Another cryptographer we interviewed pointed out that, "We still should be working on [quantum], but we also need to recognize that a couple of the proposed post quantum things have already fallen to classical attacks. You know, you may be secure against quantum computers, but not against ordinary computers. And, as we developed all of the modern crypto that we have, we made many, many mistakes. And [quantum cryptographers] haven't made any of those mistakes. So they're all there to be relearned."

1 https://blog.cryptographyengineering.com/2019/12/08/on-client-side-media-scanning/
2 https://arxiv.org/pdf/2110.07450
3 https://newatlas.com/computers/ibm-osprey-worlds-most-powerful-quantum-computer/
4 https://www.math.stonybrook.edu/~tony/whatsnew/may07/quantumI.html
5 https://arstechnica.com/information-technology/2023/01/fear-not-rsa-encryption-wont-fall-to-quantum-computing-anytime-soon/

SPYWARE/PEGASUS

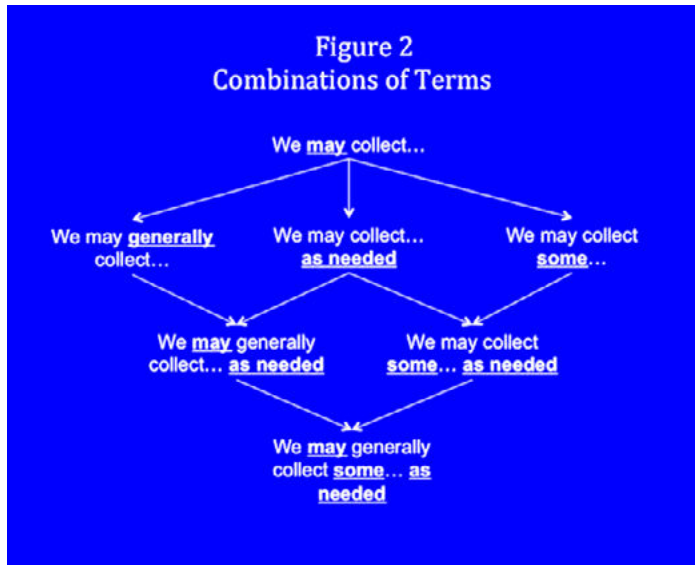
Despite the many regulations against encryption on the horizon, more and more of our everyday communications are encrypted. This change in the landscape has made dragnet surveillance of communications much more difficult, meaning that governments are relying more on technologies like NSO Group's Pegasus and other spyware to get information that they used to be able to get with dragnet surveillance. Unfortunately, advances in spyware and malware set up a battle that security and privacy folks must continue to fight in years to come. But this is not a problem that secure messaging apps can solve, nor should they be expected to.



Privacy Policy, Terms and Conditions Review

As part of our design review, we reviewed the Terms of Service and Privacy Policies, as well as associated documentation, for select messaging applications. Of course, the primary incentive for those writing these documents is to protect the organization providing the service from lawsuits and regulatory action, rather than to inform users about the services. That is not to deny that some of these platforms appear to truly care about their users' privacy, or that they have worked hard to make their legal documents understandable to non-lawyers. But this purpose is evident in the ambiguity or vagueness that a close reading of these documents reveals: terms like "may," "including," "generally," often in combination, serve to obfuscate any definite description of the actual practices of the service.[1][2]

The Terms also include warranty disclaimers that sometimes specifically mention "security" as one of the things being disclaimed. Even if a user has no intention of suing a messaging platform for breach of privacy, this text might still give pause.



Reidenberg et al, "Ambiguity in Privacy Policies and the Impact of Regulation," Journal of Legal Studies (June 2016).[3]

An additional factor that complicates comparing the policies of the various services is that the Terms of Service and Privacy Policies all contain certain required information, but much of the style, format, and level of detail differs from platform to platform. This is especially acute for Apple and Google's primary documents, which cover all (Google) or a large range (Apple) of their product offerings, making it hard to know

1 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2715164
2 https://www.cs.cmu.edu/~breaux/publications/jbhatia-re16.pdf
3 https://www.journals.uchicago.edu/doi/full/10.1086/688669

which specifics are relevant to the messaging platform of concern. Google and WhatsApp have separate policies for Europe and the rest of the world. This means that some information (such as a particular type of data use) may be revealed in European policies to conform with local regulations, but not in the other policy, even though it would apply to users there too.

Accordingly, the following information comes with the caveat that it is based only on reading the relevant documents, and is a summary of what they clearly reveal about the services, rather than a comprehensive overview.

APP TAKE-AWAYS

WhatsApp has very detailed Terms of Service and Privacy Policies (separate documents for the European Economic Area / UK and for the rest of the world), as well as extensive Help Center documentation explaining encryption and policy. Their content prohibitions are also the most extensive. All communication in WhatsApp is end-to-end encrypted, so the content of messaging and calls is essentially protected. Profile information is not encrypted and is immediately viewable by any other user who has your phone number.

Some information, such as phone number and data about how the service is used, is retained by WhatsApp and used for a wide variety of purposes (slightly more restricted in the EEA/UK), including some sharing with other Meta companies (e.g. Facebook), as well as providing metrics for business customers. This means that, for users of Meta's products, Meta may have a lot of information connected to the WhatsApp identity, even though it does not keep or share logs of who/when you message or call. WhatsApp appears to be the only service in the group that may use data for marketing purposes.

Signal has a short and, for the most part, easy-to-digest Terms of Service and Privacy Policy. All messaging and calling on the service is end-to-end encrypted, as is profile and account information, and address books that are shared for finding contacts are also encrypted so that Signal can process your request without decrypting data. The minimum of data is retained by Signal to function effectively and safely.

Signal has the only useful transparency report, which includes the full documentation of government requests for data and Signal's response (often assisted by the ACLU) and includes appeals to remove gag orders so that the requests can be shared.[4]

4 https://signal.org/bigbrother/

Messages by Google: Google's Terms of Service and Privacy Policies are written to be digestible and understandable for the general public, but they are for the whole of Google's ecosystem. Of note, they claim a commitment to (and have the resources to) carefully review government requests for user data and may not comply unless all legal requirements are met, though it is worth noting that Google, Apple, and other large companies have been fooled into giving out personal information by responding to fake legal requests in the past.[5] A few documents are available on the Messages support site that explain privacy and encryption on the app.

Google collects a lot of information about users for a wide variety of purposes across the Google ecosystem. Information from the Messages app in particular is not used for advertising, though as an integrated Android app, Messages is linked to the user's Google identity. Messages is a primary texting app (for Android only) that only defaults to end-to-end encryption when all parties in the conversation are (1) using a recent version of Messages and (2) turn on RCS chats. Otherwise messages sent will be on the unencrypted SMS protocol.

Telegram's Terms of Service mostly deal with their paid premium service. Their Privacy Policy reveals that the different types of chat within the Telegram service have different security and privacy regimes. Most importantly, only "secret chats," voice, and video calls are end-to-end encrypted. "Cloud chats" are encrypted at the transport and storage layer (with the content and the encryption key in separate data centers in different jurisdictions), and public groups or channels are accessible by anyone. Also noteworthy: bots may be present in some groups and can share all content posted in that group with their developer. Even though Telegram does

5 https://www.bloomberg.com/news/articles/2022-04-26/tech-giants-duped-by-forged-requests-in-sexual-extortion-scheme

not have the same volume of user data as Google, Meta, and Apple, it does retain metadata including users' IP address, which, with normal cell phone usage, is typically sufficient to pin down the user's location and real identity (a VPN can mitigate this, though cell phone network providers may also disclose location data based on phone number).

Telegram has the least restrictions on content that can be posted—promoting violence and illegal pornography are specifically prohibited on publicly viewable channels only. Currently based in Dubai (with a legal presence in the British Virgin Islands), Telegram also claims to only share user information with law enforcement on receipt of a court order identifying the user as a terror suspect, and that this has never happened; their transparency report is empty for this reason. However *Der Spiegel* reports that Telegram has handed over some user information relating to terrorist activity and child abuse to the German government.[6]

6 https://www.spiegel.de/netzwelt/apps/telegram-gibt-nutzerdaten-an-das-bundeskriminalamt-a-0e4d3fcb-8081-4b87-b062-db412bbc294b

Apple Messages and iMessage:
Apple's Software Licence Agreement and Privacy Policy are quite detailed, and, like Google's, not specific to their messaging service, but an additional document from the legal section of the Apple website gives a fuller explanation of privacy on the app. It is important to note that the iOS app is "Messages," which functions as a primary texting app. When all parties in a messaging conversation are using Apple Messages, the conversation is on the "iMessage" system, which is end-to-end encrypted. When any participant in the call is not using Apple Messages, the conversation will be on the unsecure SMS protocol. Messages is integrated with iOS and so will link to traditional phone calls (not encrypted) and FaceTime voice or video calling (encrypted).

As with Google, Apple retains a lot of information about its users, and uses it for various purposes. One point of concern is that Apple does not classify IP addresses as personally-identifiable information (which gets a higher level of privacy protection) unless mandated to do so by the local jurisdiction. Apple also adds the catch-all term "public importance" to reasons that they may disclose user data. A possible benefit is that, unlike the other apps reviewed here, an email address can be used instead of a phone number to set up the Apple ID that iMessage uses, enabling someone to set up an Apple device and message without disclosing their phone number.

Apple Messages + iMessage

GOOD FOR USERS WHO

are iPhone users who do not want a separate app and only require encryption when communicating with contacts who all use the Apple ecosystem. May be useful for some individuals who do not want to disclose their email address.

Messenger from Meta follows Meta's Terms of Service and Privacy Policy that applies across the Facebook and Instagram products. They do specify, however, that message content (even through unencrypted chat) is not used for ad targeting. Otherwise, message content and metadata may be used for a wide variety of purposes "subject to applicable law."

Meta Messenger

GOOD FOR USERS WHO

need to contact people who they only know through Facebook (or potentially Instagram or Oculus) and will be careful to enable end-to-end encryption before discussing any potentially sensitive topic.

POLICY REVIEW

Even democratic governments, supposedly champions of free expression and privacy, are in the regular habit of proposing policies that undermine end-to-end encryption.[7] Such policies are typically driven by a perceived need to maintain national security, combat illegal activities, or protect public safety. Such policies often take the form of laws or regulations requiring technology companies to provide "backdoors" into their encryption protocols or applications. But these measures fundamentally weaken the security offered by end-to-end encryption, exposing users to potential threats including from government overreach as well as bad actors. A review of current policy proposals (as of June 2023) in the UK, US, EU, and India give a sense of the policy landscape in four important jurisdictions.

UK

Relevant legislation: Online Safety Bill

Danger to E2EE: Imposes a general monitoring obligation on messaging platforms, especially with regard to CSAM and terrorist content.[8] This could be server-side scanning, which would break E2EE, or, (as the government appears to be suggesting at the time of publication), client-side. However, client-side scanning would also almost certainly generate a large number of false positives, thus undermining the promise of E2EE. Ofcom, the regulator that would be charged with enforcing the Online Safety Bill, would be entitled to insist on the use of as-yet-undetermined third-party technology providers, introducing a potential further risk to privacy.

Outlook: Civil liberties organizations have spoken out about this risk to privacy, and executives from WhatsApp and Signal (among others) have made clear their opposition to this aspect of the legislation, threatening to refuse to comply and either withdraw access in the UK or even assist with the availability of proxy servers to evade the law. A few of the amendments being considered in the Lords attempt to release platforms from monitoring private communications or breaking encryption.[9] However, some child safety groups favor monitoring, and neither the government nor the opposition wish to be seen as lenient on terrorism and child safety. Despite broad support for the Bill, it has grown to hundreds of pages in length, with over 500 amendments proposed in the Lords, and risks collapsing under its own weight.

7 https://www.nytimes.com/2023/06/13/opinion/encryption-messaging-privacy-signal-whatsapp.html
8 https://bills.parliament.uk/bills/3137
9 https://bills.parliament.uk/bills/3137/stages/17371/amendments/94592
https://bills.parliament.uk/bills/3137/stages/17371/amendments/94693
https://bills.parliament.uk/bills/3137/stages/17371/amendments/95044

US

Relevant legislation: EARN IT Act, STOP CSAM Act

Danger to E2EE: The EARN IT Act removes the Sec. 230 liability shield from online platforms with regard to child sexual abuse material (CSAM). Though the act specifies that providing an end-to-end encrypted service cannot be an "independent basis" for platform liability, it could be a contributing factor to liability. It therefore may dramatically increase liability risk for platforms who offer end-to-end encrypted communication services and encourage them to, minimally, introduce client-side scanning. State governments may respond to the bill's passage by outright requiring the scanning of all content.[10]

The STOP CSAM Act creates additional criminal and civil offenses for platforms that host CSAM or facilitate exploitation-related conduct (distributing CSAM is already illegal), with a corresponding carve-out to Sec. 230, leading, again, to increased liability risk for providers of end-to-end encrypted communications.[11]

Outlook: Opposition to previous versions of the EARN IT Act has already derailed it in two previous sessions of Congress, and civil liberties organizations have restated their position this time. The tech industry also has substantial lobbying power in Washington.

With the low chances of any proposed federal bill related to tech becoming law in recent years, the prospects of these two pieces of legislation are uncertain at best.

EU

Relevant legislation: Digital Markets Act, Child Sexual Abuse Regulation

Danger to E2EE: The Digital Markets Act (DMA), which focuses on encouraging fair competition by imposing certain requirements on the largest "gatekeeper" platforms, includes a mandate for interoperability in messaging systems.[12] This would apply to the platforms operated by Meta, Google, and Apple. Some security experts (and the head of WhatsApp) state that E2EE is not consistent with the interoperability requirement.[13]

The proposed Child Sexual Abuse Regulation would include messaging systems amongst in-scope platforms.[14] A regulator would have the power to run their own detection protocols and impose technological solutions if the self-assessments and mitigation strategies submitted by platforms were deemed inadequate. Though general monitoring of all private content is currently prohibited in the EU, it is anticipated that either E2EE would be broken or client-side scanning introduced under the regulation as drafted.

10 https://www.blackburn.senate.gov/services/files/77265173-A51C-4500-814E-729FCC27D9E0
11 https://www.congress.gov/bill/118th-congress/senate-bill/1199
12 http://data.europa.eu/eli/reg/2022/1925
13 https://www.platformer.news/p/three-ways-the-european-union-might
14 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2022:209:FIN

Outlook: The DMA has been adopted and its requirements will start to come into force next year (the law sets out a multi-year roadmap for text, group chat, and voice/video). Some experts have claimed that interoperability is possible without breaking E2EE. It remains to be seen how the platforms and regulators will implement this requirement, though the EU commitment to privacy gives hope that the regulators will work with in-scope companies to surmount technical and legal hurdles and continue to provide E2EE services.

The Child Sexual Abuse Regulation also contains an emphasis on preserving privacy as well as multiple safeguards against excessive regulator demands. Furthermore it is several stages away from becoming law and given European priorities, may have protections for E2EE written into it later in the process—at least one set of MEP amendments has been proposed to do so.[15] The European Council has solicited and received feedback on this question from member states, who have expressed a wide range of views on protecting E2EE, as well as notable uncertainty regarding the technological proposals for monitoring encrypted messaging.[16]

INDIA

Relevant legislation: Indian Telecommunication Bill 2022

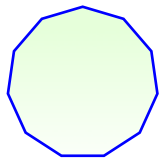
Danger to E2EE: The draft bill requires telecommunication services, including online messaging platforms, to be licensed in order to operate in India.[17] It also mandates that these services share the identity of senders with the recipients of the communications, which would force messaging platforms to disallow anonymous communication. Most concerning, it allows for the blocking or interception of communications for "public emergency" or "public safety," as determined by the Central or State Government officials, and even the temporary takeover or suspension of a service, with no review or oversight requirements.

Outlook: Public comments included requests to distinguish internet services from telecoms. A news report expected a new draft to be published earlier this year, clarifying which requirements apply to which service, though this does not appear to have happened yet.[18] The bill is set to be considered by Parliament this summer. Though the Indian government has been tending towards authoritarianism, there is a measure of pressure being applied by the judiciary, with the Supreme Court ruling

15 https://www.svenja-hahn.eu/post/surveillance-does-not-protect-children-chatcontrol-must-be-prevented-svenja-hahn-fdp
16 https://techpolicy.press/divergence-and-uncertainty-in-responses-to-eu-regulation-on-child-sexual-abuse/

17 https://dot.gov.in/relatedlinks/indian-telecommunication-bill-2022
18 https://www.outlookindia.com/business/dot-likely-to-introduce-a-revised-draft-of-telecommunication-bill-report-news-243561

that the Indian constitution includes rights to privacy and to conducting correspondence and business online. (India has existing legislation requiring user traceability and permitting the interception of online messages, but the E2EE platforms continue to operate and challenges—including from WhatsApp—are pending at the Supreme Court.) Here too, Signal has said that it will not comply with a demand to break E2EE.



Recommendations

In addition to the various suggestions for improvement and company and app-specific recommendations above, we provide the following recommendations more broadly to users, encrypted messaging app developers, and policymakers. Encrypted messaging apps are a vital means of communication, and their functionality and integrity is important to private and secure discourse.

Addressing the challenges to providing reliable encrypted messaging applications at scale is an immense undertaking, and we acknowledge this fact. There exists no straightforward remedy for improving the usability, accessibility, user understanding, and security of encrypted messaging apps for the wide variety of user threat models. With the above in mind, we present the following recommendations.

USERS

- **Users need to consider their threat model** before they communicate on sensitive issues, and recognize that no means of digital communication is 100% safe. While for most users who are not already targeted by authorities, encrypted messaging apps offer a safe and efficient means to communicate, they are not impossible to compromise. State-sponsored actors or highly skilled hackers may have resources and capabilities that surpass typical defenses and protections on most devices.
- **Establishing group norms for users communicating on encrypted messaging apps is a proactive measure** that fosters respect, promotes a culture of privacy, and maintains the integrity of the group. Such norms guide the type and scope of information shared, encouraging cautious behavior around sensitive topics, and discouraging the spread of potentially harmful or intrusive content. They also serve to remind everyone that while encryption provides a substantial layer of security, encrypted apps are not infallible, and personal responsibility is paramount. Group solidarity necessitates collective adherence to norms and protocols that reduce risk, thereby protecting both the group and individual members from inadvertent exposure of personal or sensitive information.
- **Users must be particularly diligent about settings on encrypted messaging apps**. Settings should be adjusted to situational needs, and users should consider enabling or disabling specific features based on their needs. It is particularly important to check settings after software updates, or when new features are added to an application.
- **Users should default to using disappearing messages on encrypted messaging apps**, and as short as possible a time frame. We also recommend disabling link previews in apps that permit this feature to be turned off.
- **For users under extreme circumstances** such as imminent arrest or likely targeting with advanced spyware, or where use of secure apps such as Signal is criminalized, we recommend taking more substantial precautions, such as having a separate device for sensitive conversations, leaving your device at home if you are likely to interact with law enforcement,

having conversations in person, or having a trusted person to delete or recover your accounts in the event of your arrest.

DEVELOPERS

- **In general, the makers of encrypted messaging applications should adopt privacy-by-design approaches, and invest more deliberately in working with users from more vulnerable, marginalized and targeted communities.** Designing from the margins is the approach that is necessary to maximize safety, privacy, and security for the most users, particularly those in the most dangerous circumstances.

- **Designers of encrypted messaging applications should consider user-friendly ways to enable users to easily adjust settings based on their own degree of security concern.** While it would be best if all settings default to the most secure, conservative position, a "slider" interface might allow users to automatically configure app settings on a spectrum from "most secure/least features" to "least secure/most features." Such a feature would streamline the user experience and allow users the ability to tailor their experience, fostering trust. Likewise, a "kill switch" to erase the application at the touch of a button would be valuable to users in extreme situations.

- **The makers of encrypted messaging applications should enable encryption by default and also ensure that account details and metadata are also encrypted.** While metadata such as the time of message exchanges, what participants are involved, or even user location data may not reveal the content of a conversation, they can still be used to make substantial inferences about a user's behavior that could expose them to harm.

- **Makers of encrypted messaging apps must recognize the danger in feature creep and scrutinize the necessity of every function.** App developers should think hard about the necessity of features like call logs, and wherever possible avoid feature over-proliferation and integration with other services that may create new potential security risks.

- **Companies and organizations developing encrypted messaging applications should collaborate on industry standards for terms and conditions and transparency reporting.** Harmonizing critical aspects of how platforms operate and report will promote accountability and bolster the overall integrity of the encrypted messaging space.

- **Likewise, the makers of encrypted messaging applications must stand united against government efforts to weaken or break encryption.** A unified stance is necessary, particularly in the current policy environment in democracies around the world.

POLICYMAKERS

- **Policymakers must affirm the value of privacy and its importance to maintaining democratic values.** While there are legitimate security and safety concerns that are in conflict with the right to privacy, policymakers must recognize the magnitude of harms that follow from weakening or breaking encryption for all.
- **Policymakers and regulators in jurisdictions that support encryption need to support those in jurisdictions where it is threatened.** This is particularly important in the United States, since should technology firms be forced to compromise in the U.S. it will have dramatic implications elsewhere in the world, enabling repressive regimes.
- **Policymakers should advance measures related to transparency and standardization of reporting**, as well as measures that require terms and conditions to be written in plain language.
- **Consumer protection agencies should be on the lookout for deceptive design and false claims** about the degree of security conferred by messaging applications, alongside social and digital media applications generally.

Conclusion and Suggestions for Future Research

Privacy, security and safety are ultimately collective issues that can only be solved by collaboration across society. In the course of this project, we identified a range of questions and issues that are worthy of future research and development. Here, we offer five potential directions for consideration.

1. Governments and foundations should invest in a privacy-preserving public interest technology ecosystem.

The extraordinary value of Signal as a privacy-centric, public-interest technology serves as a compelling testament to the potential of non-profit-driven models in creating a safer and more equitable digital ecosystem. Governments and foundations should invest significantly in the development of similar public-interest technologies to foster post-surveillance capitalism communication networks. This entails not only providing substantial funding for research and development in privacy-preserving technologies, data sovereignty, and digital rights, but also promoting open-source models, innovative public-private partnerships, and regulatory frameworks that prioritize user privacy and digital equity.

Furthermore, investing in education and public awareness about the importance of digital privacy, and how these technologies serve this purpose, is crucial. Building a robust public interest technology ecosystem, in harmony with platforms like Signal, offers an opportunity to redefine digital spaces as more user-focused, democratic, and secure, countering the dominance of profit-driven surveillance capitalism.

2. Governments and industry should partner to develop standard evaluation criteria and methodologies for messaging application transparency, design, and technical functionality.

Standard criteria and methodologies for transparency can help encourage accountability, interoperability, and public trust. Transparent evaluation criteria will help ensure that app developers are held accountable for the privacy and security promises they make to users. This not only should include encryption standards but also design decisions and metadata handling practices. A key goal should be to develop common language and signifiers to increase user literacy on privacy and security.

3. Industry and civil society should develop specific design and technical standards for operating in authoritarian contexts.

Developing rubrics and frameworks for design and technical standards for encrypted messaging apps and other digital media products to operate specifically in authoritarian contexts could be vital for safeguarding fundamental human rights and enabling secure communication in environments where free expression may be threatened or criminalized. Such standards can help protect activists, journalists, and dissidents by protecting their identities and ensuring the confidentiality of their communications. Moreover, such standards may include features or technical innovations that help users circumvent censorship or surveillance.

4. Designers and engineers should future-proof private messaging by looking at future generations of technology and the extent to which there are specific questions they raise that have bearing on private and secure communication.

Further research into understanding tomorrow's technologies is essential to future-proofing private and secure communication. As technology evolves, so too does the complexity and sophistication of threats to privacy. Proactive efforts to evaluate emerging technologies can help anticipate potential vulnerabilities and design security measures to counter them.

Some future technologies, such as quantum computing, may fundamentally reshape the privacy landscape. Quantum computers could eventually break many current encryption algorithms, necessitating a complete re-evaluation of security and privacy protocols.

5. Technologists should develop tooling for mixed-method and interdisciplinary investigations, such as reverse engineering tools, that would help independent researchers assess the security of encrypted services and the privacy of platforms.

The development of investigative tooling such as reverse engineering tools can be instrumental in assessing the security of encrypted services and the privacy of platforms. These tools allow researchers to examine how encrypted messaging apps and other such services function, helping identify potential vulnerabilities or flaws in the encryption protocols or data handling practices. Without such capabilities, it is challenging to evaluate the security and privacy claims of these services.

Such tools and methods could contribute significantly to the field of independent auditing and third-party verification of marketing claims from companies such as Meta, Apple, Google, Telegram and more, helping to build trust, transparency and accountability in these areas.

By investing in the development of such tooling, we can ensure the resilience of encrypted services in the future.

6. The Signal protocol is responsible for protecting the privacy of billions of people and deserves further investment, study, and analysis.

The Signal protocol has taken over the world of encrypted messaging. It is being used in every encrypted message app we looked at with the exception of iMessage and Telegram. Currently more than two billion users are relying on the Signal protocol to protect the privacy of their communications, and if Meta implements encrypted chat by default in Facebook Messenger there could soon be billions more. As far as we know the Signal protocol is cryptographically secure, but currently all the eggs are in this one proverbial basket. An as yet unknown flaw in the Signal protocol could compromise the safety of billions of users. More research should be done to further confirm its safety.

In addition to the above, researchers should continue to study phenomena that occur on encrypted messaging applications, such as the spread of harmful material, disinformation, abuse and harassment, and to look for new ways to address content moderation challenges that preserve the promise of security and privacy on encrypted applications.[1] From the advance of authoritarianism and erosion of rights to the rise of artificial intelligence, the preservation of secure, private communications is an urgent challenge that must be met by a broad and committed community of researchers, designers, technologists, policymakers, and engaged users.

1 https://arxiv.org/abs/2303.03979

Funding Statement

This research and publication of this report was supported by a program at Omidyar Network focused on safe, private, and trustworthy messaging.

Thanks and Acknowledgements

We would like to thank Afsaneh Rigot, Ahmed Razek, Albert Fox Cahn, Alex Leavitt, Anushka Jain, Apar Gupta, Caitlin Seely George, Chris Gilliard, Chris Kaiser, Dhevy Sivaprakasam, Eye on Surveillance, Ian Goldberg, Jillian York, Jon Bell, Jon Callas, Kate Bertash, Konstantinos Komaitis, Matt Mitchell, Riana Pfefferkorn, Rose Jackson, Sarah Aoun, Wai Phyo Mint, and Zelly Martin Geurink for providing input for this report. Thanks to our expert reviewers: Georgia Bullen, Martin Shelton, Jon Callas, Erica Portnoy, Michal Luria and Dhanaraj Thakur.

To Wafa Ben-Hassine and Emma Leiken, we appreciate your support and guidance throughout this endeavor.

We would also like to express our sincere gratitude to REJAC; Prateek Waghre, Tejasi Panjiar, Anushka Jain and Ashlesh Biradar at the Internet Freedom Foundation; and the numerous individuals who participated in interviews and group discussions, sharing their thoughts and experiences regarding privacy and security. Thank you to Roxy Zeiher for the graphic design of this report.

Cooper would like to thank his family, who put up with him being gone for several weeks during this project. He would also like to thank his colleagues at EFF and other friends who wish to remain anonymous for their guidance and support.

Appendix
Suggested Improvements

CRYPTOGRAPHIC DESIGN

- The developers and designers of encrypted messaging applications should agree to use standardized terms whenever possible.
- Telegram should switch to a more modern and well-tested secure chat protocol.
- Google should turn on RCS message support by default in Messages.
- Google should take steps to hide more metadata in encrypted RCS.
- Google should make unencrypted messages more visually apparent.
- Apple should support RCS in iMessage,and make unencrypted messages more visually apparent.
- Apple could be more clear about the fact that blue chats are end-to-end encrypted while green chats are not within the UI and their messaging.
- Both Apple and Google could improve overall user safety by agreeing on an interoperable standard for encrypted messages. (Apple could keep iMessage with extra features, while still supporting encrypted RCS in a green chat bubble.)
- Telegram should switch to a key verification system that is harder to spoof.
- Apps should standardize on the terms it uses for key verification.
- Apple should enable key verification by default.
- More research should be done to determine better ways to communicate and present key verification to users.
- Other messaging apps should borrow techniques from Signal and take steps to hide user metadata by keeping it encrypted with the user's account key and only handling unencrypted versions in secure enclaves.
- Mobile app makers should converge on standardized terms for two factor authentication, encrypted messages, and key verification.

USER EXPERIENCE DESIGN

- Developers should make it easier to quickly delete entire message threads and interaction histories between users and in groups across all apps. WhatsApp, Facebook Messenger, and Telegram should allow for phone calls to not appear in a device's call log, similar to a functionality that Signal allows. WhatsApp should implement (and turn on by default) banner and visual notification settings for "keep" messages across mobile and desktop.

- We recommend that high risk users disable link previews immediately, especially if using an app in a country where that app is forbidden (such as Signal in Iran).
- WhatsApp should give users a way to turn off link previews, as Signal already does.
- For encrypted messages, link previews should be off by default and potential risks should be better communicated to users consistently and over a sufficient period of time to ensure that as many users as possible understand the risks. This could and should be communicated via UI design, along with warning labels (potentially as pop ups or interstitials), and blog posts.
- Apps should endeavor to strip tracking parameters in the client for as many popular websites as possible.
- Apps should proxy GIF searches in a privacy preserving way, possibly using a transparent proxy to decouple the user's identity and what they are searching for.
- Signal and Whatsapp should delete quoted text in replies to a message when the original message has been deleted.
- WhatsApp should encrypt backups by default. iMessage should warn users if backups are enabled but advanced device protection is not turned on.
- Both Telegram and Facebook Messenger should make all messages encrypted by default, and at the very least ensure the naming of features matches up with what the feature does. Encrypted messages aren't 'secrets,' per se, and naming conventions should reflect that.
- Developers should allow for customizable disappearing message times; increase granularity for all product features; allow for relevant features to be disabled; universal notifications of screenshots; and the ability to disable screenshots. Lastly, in the Signal app, the 'delete for me' option sits above 'delete for everyone' in the option menu to delete messages. Users we interviewed alluded to accidentally hitting 'delete for me' when meaning to select 'delete for everyone.' We recommend flipping the order of these two buttons.

Thank You.