10 Myths and Facts about Traceability

There are a lot of misconceptions out there about Article 10 of the “Freedom, Responsibility and Transparency on the Internet” bill that was recently approved by the Senate and what “traceability” means. We want to make sure you have all the facts.

**Myth 1:** The Senate’s traceability provision only applies to viral messages. No private, person-to-person conversations will be impacted.

**Fact:** Because WhatsApp cannot predict the future and know which messages will eventually be forwarded, we would have to place a permanent identity stamp, like a fingerprint, on every message sent in order to trace a particular message back to the original sender(s) once that message hits the forward thresholds set forth in the Senate bill (i.e., forwarded by more than five users and reaching 1,000 people). This will fundamentally change the product that more than two billion people around the world use today.

We would have to stamp all the messages you send to close contacts -- friends, colleagues, your doctor, lawyer, or client -- not just in Brazil, but globally. We would also have to stamp the personal messages you send to small groups of family or friends. Once all of these messages are stamped, authorities in Brazil and all over the world will ask WhatsApp to turn over the accounts of the senders, even if the senders condemned the content or shared it to inform or warn others.

To be clear, the vast majority of messages sent over WhatsApp today -- 9 in 10 messages -- are exchanged between just two people and involve no message forwarding. And WhatsApp already restricts message forwarding -- earlier this year we announced a strict limit on messages that have been shared many times (sometimes called “frequently forwarded” messages). This change led to a 70% reduction in the number of frequently forwarded messages across the entire service.

**Myth 2:** Traceability does not break encryption.

**Fact:** End-to-end encryption keeps our online conversations private and secure. It is based on technical guarantees that, if altered by adding traceability, break encryption. For example, encryption guarantees that nobody other than the person you are talking to can know that you sent a particular message. Adding traceability breaks this guarantee. Its primary purpose is to expose who sent what to whom. End-to-end encryption also guarantees that the messaging app you are using does not know what content people are sending. An implementation of traceability that requires private messaging services to provide a one-stop-shop and “tell us everyone who sent this content,” breaks end-to-end encryption, as it forces services to store, and therefore have access to, the content of messages on their servers. That is a big departure from how encrypted messaging services are designed and work today.
As UN Special Rapporteur David Kaye and IACHR Special Rapporteur Edison Lanza wrote in an alert to the Brazilian Mission in Geneva on July 3rd, expressing concerns about the Senate bill: “Encryption of communications provides individuals and groups with an online zone of privacy to hold opinions and exercise freedom of expression without arbitrary or illegal interference and attacks.” By breaking end-to-end encryption, traceability eliminates this critical zone of privacy.

Myth 3: WhatsApp already tracks how many times a message is forwarded. So it’s easy for WhatsApp to know who the original sender is.

Fact: WhatsApp has developed a feature to allow users' devices to identify -- and restrict -- when a particular message has been forwarded many times. This counting of forwards occurs on users' devices (often referred to as the “client side”) and is protected by end-to-end encryption -- only the sender's device and the recipient's device have access to it. In order to preserve encryption, WhatsApp itself does not record how often a given message is forwarded, the content of that message, or who sent or received it.

Myth 4: The Senate’s traceability provision only requires companies to keep users’ data for 15 days, or in some cases a maximum of three months.

Fact: If services are required to stamp the sender and recipient of every message, this information will be stored on users’ devices (client side) and available to everyone who receives the messages (or anyone who gets access to the sender or recipient’s device). There is no way to guarantee that this information will be deleted.

And as mentioned above, an implementation of traceability that instead requires private messaging services to provide a centralized, one-stop-shop for identity stamps on their servers would give services access to the content of messages and break encryption.

Myth 5: Tracing private messages is a reliable way to identify the originator of viral content.

Fact: Traceability is ineffective and highly susceptible to abuse. If you simply download an image and share it, take a screenshot and resend it, or send a news article on WhatsApp that someone emailed you, such a system would register you as the originator of that content. It’s also easy for a user to restart a forwarding chain by copying content within WhatsApp and pasting it into a new thread. In order to see the entire forwarding chain, or to reliably identify the actual originator of a message, Brazilian authorities would have to obtain and trace messages from many users’ phones, or risk having an incomplete picture. Think of it like a tree with many branches - looking at just one branch doesn’t tell you how many other branches there are or where and how all of the other branches originated.

Traceability also puts innocent people at risk because it is easily susceptible to abuse. Bad actors could easily copy and paste messages as explained above for the specific purpose of circumventing
traceability. And relatively sophisticated bad actors could modify a forwarding record. While the potential for abuse already exists in every client-side implementation, the dangers inherent in features like forward counters -- altering a particular message’s forward count -- are significantly different from the dangers inherent in tampering with or altering a message’s origin. If the traceability law passes, this could allow attackers to evade law enforcement at best, or frame innocent people at worst.

Myth 6: Message tracing data cannot be used to surveil people.

Fact: Traceability seeks to force private companies to collect and store who-said-what and who-shared-what for billions of messages sent each day. Requiring platforms to collect more data than they need solely for the purpose of turning it over to law enforcement runs counter to international norms, directly contradicts Marco Civil and the Lei Geral de Proteção de Dados’ data minimization principles, and raises the specter of mass surveillance. As the Brazilian Bar Association says, the draft law “would require the creation of a mass storage system for communications on a permanent and continuous basis, admitting a breach of confidentiality.”

This massive data collection also makes messaging platforms inherently less secure by opening up more avenues for hacking. Further, WhatsApp is intentionally designed to be a data-light app that can work for any user, regardless of how much internet bandwidth they can afford. Traceability would increase data usage on our system to levels that could slow down the app significantly.

And, given that Brazilian users send and receive messages to and from people all around the world, traceability contradicts international data protection laws and threatens the privacy of communications well beyond Brazil’s borders. WhatsApp would need to trace messages from Brazilian and non-Brazilian users alike. Limiting the law to “Brazilians” or “Brazilian numbers” would mean losing visibility into any forwarding chain that includes a non-Brazilian, and giving criminals the ability to use a foreign number to easily circumvent the system.

Make no mistake: traceability will make Brazil a true outlier on the global stage - no democracy in the world mandates private message tracing.

Myth 7: Traceability does not put users at legal risk for forwarding content.

Fact: Traceability strips people of the presumption of innocence and inverts law enforcement investigations. Private companies would be asked to turn over the names of users who have sent or shared something, even when they did not create it, shared it out of concern or to check its accuracy, and are not being investigated for a crime. This means innocent people could go to jail. As Freedom of Expression Special Rapporteurs Kaye and Lanza explain in the same alert referenced above, “The provision even reverses the burden of proof. The individual must demonstrate that by sharing certain content he or she did not intend to misinform. This gives State and non-State actors a legal tool to criminalize journalists or opponents for political reasons or in order to generate a chilling effect among
those who disseminate information that is intended to be concealed."

**Myth 8:** Tracing private messages is the only way Brazilian authorities can prosecute people spreading malicious content on WhatsApp.

**Fact:** WhatsApp respects the important work Brazilian law enforcement does to keep people safe. The information we provide complements additional law enforcement sources, such as direct access to devices or cloud storage to secure messages. We regularly assist Brazilian authorities in criminal investigations with a dedicated team that reviews, validates and responds to law enforcement requests. While all messages remain end-to-end encrypted, we support valid requests by providing the information available to us, consistent with applicable law and policy. We also have a team devoted to assisting law enforcement 24/7 with emergencies involving imminent harm or risk of death or serious physical injury. We consistently receive feedback from Brazilian law enforcement that our responses to requests help solve crimes and bring people to justice.

**Myth 9:** WhatsApp currently allows users to send mass and broadcast messages.

**Fact:** WhatsApp has taken numerous measures to prohibit mass messaging including banning users and limiting forwards.

- First, our product is designed and used for private, personal conversations: nine in ten WhatsApp messages are sent from just one person to another.
- Second, unlike some messaging apps, WhatsApp restricts group and broadcast list size. On average, the majority of WhatsApp groups in Brazil have fewer than ten people.
- Third, WhatsApp is one of the few platforms that actively and intentionally constrains sharing and virality. We recently announced a strict limit on messages that have been shared more than five times (known as frequently-forwarded messages). Now people can only forward these messages to one chat at a time. This change has led to a 70% reduction in the number of highly-forwarded messages across the entire service.
- Fourth WhatsApp gives users control over who can add them to a WhatsApp group -- this prevents people from being added to groups without their consent.
- Fifth, WhatsApp is testing an innovative new feature in partnership with Google that will allow users to check the accuracy of highly forwarded messages shared in groups.
- WhatsApp also uses advanced machine learning to proactively scan for accounts engaged in automated, inauthentic, or mass messaging behavior. These technologies ban 2 million accounts from WhatsApp per month.
- And WhatsApp has filed lawsuits in Brazil against companies that operate paid bulk messaging services and engage in sending automated messages.

Instead of weakening privacy for millions of WhatsApp users through traceability and opening up new avenues for abuse, we can address viral misinformation more effectively through the types of measures
described above. WhatsApp is actively considering additional product changes, with input from civil society, government, and based on user research.

**Myth 10:** WhatsApp is the main medium through which misinformation campaigns are disseminated.

**Fact:** Misinformation is a societal problem that exists with every form of communication and across services and platforms, as well as in-person conversations, letters, and phone calls. The ability to reach a mass audience on social media platforms far exceeds the purpose and reach of messages sent on WhatsApp.

WhatsApp is designed and used for personal conversations. We enforce strict limits on mass messaging and group communication. Calling WhatsApp the “main medium” for misinformation lacks substantiation. Knowing the main medium for misinformation would require a comprehensive study that objectively quantifies misinformation across all communications platforms. That analysis does not exist today. While we should continue research in this area, what we do know and can substantiate is that forcing services like WhatsApp to trace private messages will not eradicate misinformation and will likely not be effective in identifying the originator of messages; but it *will* undermine privacy and free expression for millions of Brazilians.