



NEWS, REPORTS

‘Civic Listening’: Political Informants and Citizen Spies, Rebranded

by Allum Bokhari & Oscar Buynevich February 2, 2024

SUMMARY

- The popularity of private messaging channels such as WhatsApp to share news and political opinions is a major obstacle for the censorship industry, which cannot easily monitor such channels.
- These private networks are particularly popular in non-English speaking demographics, [a growing cause of concern for online censors](#).
- In response, top censorship organizations have turned to old-fashioned networks of citizen informants, training them to spy on and report messages in private encrypted chats through online tiplines.
- One of the nonprofits building these tiplines, Meedan, is backed by \$5.7 million in U.S. taxpayer funds via the National Science Foundation (NSF).
- Mass censorship of private messaging platforms has occurred before, most notably when Facebook banned hundreds of thousands of pro-Bolsonaro WhatsApp accounts in Brazil.

Every major 20th-century regime that made policing its citizens’ speech a priority ended up creating vast networks of volunteer political informants to spy on people’s private conversations. Such networks could be found in Imperial Russia, Nazi Germany, and the Soviet Union. At its peak in the late 1980s, the German Stasi reportedly [maintained 1 informer for every 50 citizens](#). The number of KGB informants was estimated to be [in the millions](#).

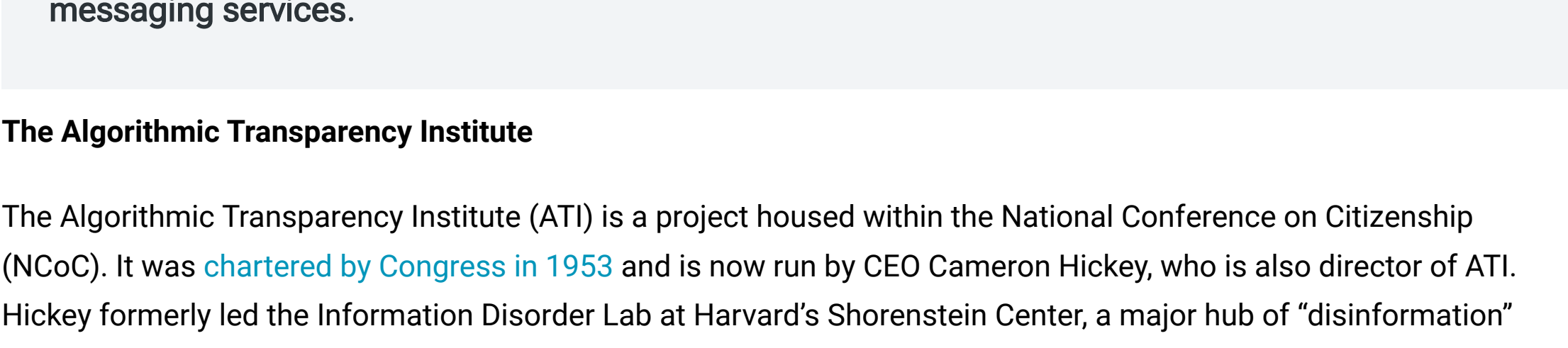
Under the brand of “Civic Listening,” informant networks are now coming to the West. The concept has been embraced by the censorship industry’s network of nonprofits, research institutes and private companies that work to shut down disfavored political speech online. Spurred by concerns about the growth of private discussion groups on platforms like WhatsApp and Telegram, the censorship industry has turned to the cultivation of snitches to gain access to U.S. citizens’ private conversations.

Meedan: Government-Funded Surveillance of Private Chats

One of the organizations at the forefront of creating this snitch network is Meedan, a San Francisco based nonprofit. The organization has received one of the largest federal grants ever awarded for a censorship program: [\\$5.7 million in taxpayer dollars was awarded to it](#) by the National Science Foundation (NSF) for its work to tackle “hate, abuse, and misinformation with minority led partnerships.” Meedan also received a smaller [\(\\$144,850\) grant from the government-funded Open Technology Fund](#) for a “claims and memes database” to monitor “fact-checked claims and debunked visual information from internet repressive countries.”

Meedan’s flagship product is Check, a tool that allows users on private messaging platforms (WhatsApp, Messenger and Telegram are specifically named) to report “misinformation” in private chats through a tipline. Check’s AI-powered bot can then be used to automatically generate a “fact check” that can be shared in closed discussion groups. Any rogue claim reported to the tipline can be permanently stored and grouped with similar claims to build up a database of banned ideas.

From Meedan’s [website](#):



Meedan is explicit about its goal of encouraging others to monitor private conversations for “misinformation.”

From Meedan’s [blog](#):

Misinformation is destabilizing elections, slowing pandemic recovery, entrenching climate change denial, and creating civil unrest and violence. **Today’s internet has four billion publishers distributing memes and reels through open networks and closed messaging platforms.** Much of this content is enriching communities, but the underlying infrastructure and business model that power it make it fertile soil to sow division, discord, and ignorance.

Meedan’s team is actively countering this challenge by building open-source software that enables local partners to organize and scale collaborative efforts to discover and address that harmful misinformation, **particularly in closed messaging spaces** during critical public and civic moments. Our community impact programs run in over 35 countries across the Larger World and Check has become the global leader in fact-checking and annotation software for collaborative fact-checking projects, helping journalists respond to misinformation across the open web and in **closed messaging services**.

The Algorithmic Transparency Institute

The Algorithmic Transparency Institute (ATI) is a project housed within the National Conference on Citizenship (NCoC). It was [chartered by Congress in 1953](#) and is now run by CEO Cameron Hickey, who is also director of ATI. Hickey formerly led the Information Disorder Lab at Harvard’s Shorenstein Center, a major hub of “disinformation” research at Harvard. Under Hickey, the ATI has embraced the term “Civic Listening” to whitewash the practice of spying on private conversations in closed messaging platforms.

ATI’s sub-branch dedicated to this mission, the “Civic Listening Corps,” openly states its purpose: to train volunteers to report disfavored speech back to the censorship industry and/or content moderation departments at tech companies.

From ATI’s [web page for the Civic Listening Corps](#):

Instead of waiting for platforms and legislators to act, **we must build capacity within our communities to establish resilience to the impact of misinformation now.** We must cultivate a shared sense of civic responsibility for understanding and identifying these harmful and misleading messages.

The Civic Listening Institute (CLC) is a civic engagement program created by the Algorithmic Transparency Institute (ATI). The CLC is **a volunteer network of individuals trained to monitor for, critically evaluate, and report misinformation on diverse topics** central to our civic life: voting, elections, public health, civil rights, and other important issues.

The ATI also maintains a database of reported information, called “Junkipedia,” which it describes as “digital public infrastructure for civic listening.”

From ATI’s [web page for Junkipedia](#):

Junkipedia is **digital public infrastructure for civic listening**. Understanding how problematic content such as misinformation, hate speech, or junk news impact society requires shared tools to identify and archive that content. Junkipedia is a technology platform that enables manual and automated collection of data from across the spectrum of digital communication platforms including open social media, fringe networks, and **closed messaging apps**. The system is powered by the active engagement of a large and diverse network of civic-minded stakeholders from the civil rights, journalism, and academic communities.

In a 2021 zoom presentation, ATI director Cameron Hickey explained his vision for “civic listening” — a vast, society-wide network of snitches that report information (even information in private group chats) to the censorship industry, which can then pressure social media platforms into taking it down.

Transcript:

We’ve created a set of tip lines. There’s an SMS phone number you can text any link to. There’s an email address. There’s a WhatsApp number you can forward content you find on WhatsApp, and there’s a web form that you can access.

The next time you’re scrolling through TikTok or scrolling through Instagram or **browsing a private group on WhatsApp and you see something that you think is problematic, you can click the share button or the forward button and send it to this tipline.** And then the folks at AltaMed, the folks at ATI and all the other organizations partnering to address mis- and dis-information will then have one more data point to go on to better understand what’s happening.

Why do we do this? We do it for a couple of reasons. **First and foremost, we’re flagging content because we can escalate that content.** We can say ‘there are 18 pieces of content that are claiming X around voting or around voter registration, and folks like Emma and their organization take it directly to the platforms. Or Stanford Internet Observatory will take it and escalate it to all of the different social media platforms. The second is so that we can get situational awareness about what’s happening so that we can develop counter-messaging, what we call inoculation messaging. We need to inoculate against problematic narratives that are spreading on social media.

Hickey’s use of “inoculation” refers to what the censorship industry calls “pre-bunking” — the censorship industry’s plan to brainwash the public at scale by feeding them weakened versions of disfavored narratives before they encounter the real thing, biasing them ahead of time. [As a previous FFO report highlighted](#), Google has been working on deploying this technique online.

In another part of the presentation, Hickey admits that the goal of his censorship squad is to target not just false information, but also “covert hate” and arguments based on “logical fallacies.” The two examples he shares are Tucker Carlson’s discussion of “great replacement” theory, and Sen. Rand Paul (R-KY)’s contention that those who support vaccine passports without also backing voter ID are hypocritical.

Transcript:

We see a lot of hate speech spreading overtly and covertly on social media. It’s critical to recognize that this is part of the problem. This isn’t a separate problem, and we need to identify and deal with it just the same as we do when something is actually false. **Especially when it’s not something that contains that overt hate speech that’s easily taken down by a platform or whatever. It’s when we get into this more nuanced area like this example around replacement theory gets to be really problematic.**

This is a critical area that we see a lot of challenges with, which is content that isn’t easily fact checked at all. Content that contains logical fallacies. An argument or a structure that’s rhetorical in nature, that it’s easy to buy into, but in fact it’s not founded on something that’s logical. It’s really easy to make a statement like **“If you think voter ID is racist but a vaccine passport is just fine, you need some serious help thinking through public policy.”** That is a logical fallacy argument, right? It is not legitimate to make a comparison between those two things. However, people share it, it resonates with them on a base level because it matches their political orientation, so they re-share it and other people get confused by it. We see this over and over and over again on social media.

Hickey does not bother to explain to attendees why Carlson’s discussion of “replacement theory” is “covert hate,” or how Sen. Paul’s statement is a logical fallacy. Agreement from the attendees appears to be assumed.

Election Interference At Home and Abroad

Meedan and ATI are not concerned with just any type of information. They zero in on speech surrounding specific topics, particularly COVID-19 (in 2020-2022) and elections — not just U.S. elections, but elections in foreign countries too. Tiplines and informant networks have already been used in foreign elections as well as the 2022 U.S. midterms monitor and report information on WhatsApp.

The results of these efforts were positively reported in an article by Meedan employees for Harvard Kennedy School’s “Misinformation Review,” which presented the resilience of platforms like WhatsApp to centralized surveillance as a problem to be solved — by volunteer informants sending information to “tiplines” like the ones built by Meedan.

Via [Misinformation Review](#):

Platforms such as WhatsApp that offer end-to-end encrypted messaging face challenges in applying existing content moderation methodologies. **End-to-end encryption does not allow the platform owner to view content. Rather, only the sender and recipients have access to the content — unless it is flagged by a receiving user (Elkind et al., 2021).** Even though WhatsApp is extremely popular, used by over 2 billion users all over the world, **there is currently no large-scale way to understand and debunk misinformation spreading on the platform.** Given the real-life consequences of misinformation (Arun, 2019) and the increasing number of end-to-end encrypted platforms, developing tools to understand and uncover misinformation on these platforms is a pressing concern.

While our paper is, to the best of our knowledge, the first peer-reviewed study on WhatsApp tiplines, **tiplines are quite common in practice. WhatsApp, for instance, currently lists 54 fact-checking organizations with accounts on its platform.**¹ Other efforts include the Comprova project² and FactsFirstPH,³ an initiative of over 100 organizations uniting around the 2022 Philippine presidential election. Tiplines are similar to features on platforms such as Twitter and Facebook that allow users to flag potential misinformation for review, but tiplines are operated by third parties and can provide instantaneous results for already fact-checked claims (Kazemi et al., 2021).

In 2020, Meedan [boasted about the results](#) of its own WhatsApp tiplines in multiple countries:

Our collaboration with five global fact-checkers – BOOM, India Today, Africa Check and AFP (Brazil and India), in phase I of the project was key to further customizing Check to meet the needs and demands of fact-checking misinformation in **closed messaging networks**.

In the second phase, five news partners – Factly, Vishvas News, Estadão Verifica – an initiative of Estadão; Agência Lupa, and Webqoof – the fact-check wing of The Quint joined the project.

With more collaborations and focussed efforts we have been able to scale up distribution of fact-checks. As of November 2020, all ten tiplines interacted with over 600,000 unique users and distributed over 15,000 fact-checks across India, Brazil, South Africa, Kenya and Nigeria. Within 5 months of launching the Phase I tiplines, Check had distributed around 5,800 fact-check reports.

And, in 2022, [Meedan deployed its WhatsApp fact-checking apparatus to target Spanish speaking voters](#) in the U.S. midterms, with an eye on 2024:

This midterm election cycle is fueling misinformation on closed messaging apps like WhatsApp — especially in non-English languages. The goal of the 2022 project is to help stabilize the information ecosystem at scale as Americans head to the polls.

“The 2022 U.S. midterms initiative will boost Spanish-language audiences with high-quality information and test a model for a higher-profile collaborative reporting project in 2024,” said Pierre Forcoli-Conti, Meedan’s Director of Product.

As highlighted by FFO’s last report, monitoring of WhatsApp and other private messaging apps is [a growing point of focus](#) for elements of the censorship industry that are focused on non-English speaking demographics, owing to the popularity of WhatsApp in those communities as a means of sharing news and opinions.

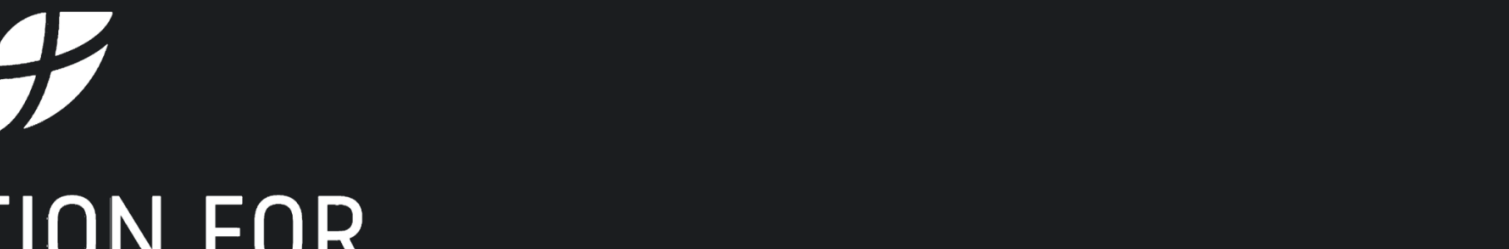
The goal of the censorship industry is clear — the infiltration of private messaging groups for the purpose of surveillance in the first instance and censorship in the second. This is made plain by Cameron Hickey’s explanation that the purpose of collecting tips from volunteers is to “escalate” cases of disfavored speech to the platforms.

This can result in private messaging groups and users being censored. One of the most prominent examples of this is in Brazil, where Facebook was [pressured into shutting down hundreds of thousands of WhatsApp accounts](#) that were used to share material in support of populist former president Jair Bolsonaro. Whether a discussion is private or not, platform owners can still ban users if information in those chats somehow leaks out. The censorship industry knows this, which is why they want to create a culture of informants — a culture of “Civic Listening.”



Allum Bokhari & Oscar Buynevich

Share this post:



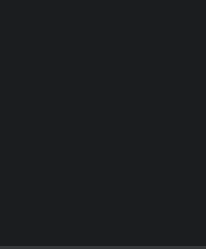
Previous:

The Censorship Industry’s Plan to Censor Latin...

Next:

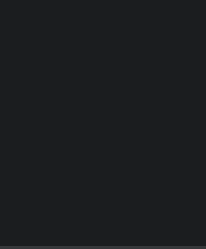
Protection Racket: Why The Censorship Indust...

Recent posts



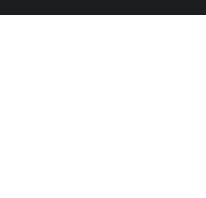
‘Trusted Messengers’: Volunteer Censors Trained to Control Your...

February 8, 2024



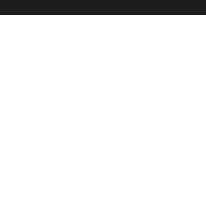
Protection Racket: Why The Censorship Industry Needs to...

February 5, 2024



‘Civic Listening’: Political Informants and Citizen Spies, Rebranded

February 2, 2024



The Censorship Industry’s Plan to Censor Latino Communities

January 28, 2024



FOUNDATION FOR
FREEDOM ONLINE

Defending digital liberties

