Security considerations when engaging sensitive sources

by Łukasz Król (Internews) and Harlo Holmes (Freedom of the Press Foundation)

We took the utmost care to make sure that this document accurately reflects the most up-to-date knowledge about technology and threat models. Still, in our field, technologies and security models continue to change rapidly. If you feel that something in this piece is out of date or have any further suggestions on how we could improve it, please drop us a line at gtp-jsf@internews.org

This document was last reviewed on 2022-05-24

Introduction

If you are a journalist working on a high-level project or investigation, then it's possible that an adversary, like a government or corporation, would take additional steps to track you or your sources. The attacks and the scenarios that we describe below are unlikely to happen in most investigations. These recommendations are targeted towards individuals working with sensitive sources on high-level projects that uncover government or corporate misdeeds. We have drafted a few additional steps you should keep in mind if you will be communicating with sensitive sources. Following the below steps will make it much harder for a dedicated adversary to figure out who your sources are.

When drafting this list, we talked to many journalists and experts in Central Europe, and conducted desk research as well. We base our recommendations on case studies we observed from the US, Central Europe, and other regions. This list is designed to be a series of potential, often unconfirmed risks. We are sharing them with you not because we expect you to be on the lookout for these risks at every moment of your reporting. We share them so that you can identify potential causes of sophisticated leaks and adjust your procedures accordingly if you feel it's necessary to do so. A small change in procedures could play a huge part in helping to protect the most sensitive sources even further.

Concrete steps you can take to protect yourself and your sensitive sources

Sanitize all emails and documents

Emails and documents that are shared with you can contain malware or tiny electronic trackers. The malware could infect your devices[1], while the trackers could record every time when the email or document has been opened, potentially exposing some information about your network or helping to de-anonymize a source.

[1] There have been many cases over the past years in which journalists received malware-loaded documents, usually from an attacker who pretends to be a source. For just one example, see https://threatpost.com/eff-activists-journalists-hit-by-targeted-malware-attack/103712/

Fortunately, there are some tools or workflows you use which limit much of the impact that malware or trackers embedded in a document could have. You can **use Dangerzone[2] to first turn any documents you work with into a safe PDF file**. Alternatively, you could also **only open the documents on a computer that's not connected to the internet**. When it comes to emails, **do not load or forward any images from emails**, as email trackers will almost always take the form of images. (This can be done by blocking "remote content", depending on your email client.)[3]

Don't forget that many organizations track outgoing emails and all online activity and could easily catch a source who writes to a journalist from their organizational account or device. If you can avoid it, do not contact sources on any of their organizational accounts, devices, or services.

Many printers also add printer dots - or invisible fingerprints - to every document that was printed. This could theoretically allow an investigator to look at a printed document or a high-quality scan thereof and figure out which printer it was printed on. This allows law enforcement, or even private companies to sometimes figure out who printed it.[4] If you receive a printed document or high-quality scan, it might be a good idea to **fully re-create the document before publishing;** either by retyping it or summarizing its contents.

If you choose to publish an original document in full, without rewriting it, make sure that you first remove all the metadata on it (including information about tracked changes, if any). One easy way to do so might be to open the document in Google Drive and then print it as a PDF. Once you have done so, inspect the document properties in your PDF reader (here's how to do so in Adobe Acrobat and macOS Preview) and in your file explorer (Windows and macOS) to make sure that there's no sensitive metadata remaining.

Be aware that images can be geolocated

Assume that, given enough time, a talented or well-resourced adversary like the police or a public or private intelligence service can figure out the location of any photo. This could allow a source to be de-anonymized.

Let's say that you receive a photo of graffiti on the wall of an abandoned supermarket in a small rural village. It turns out that the graffiti is a few months old, there have only been three days of snow in the village this year, and the village only receives a handful of visitors a month. Even a hotel room or office might have identifying characteristics that are not immediately obvious. Because of this, it might be possible to narrow down who took the photo and when.[5] Sources might not always be aware of the risks of publishing such photos, so take a **moment to think whether photos sources take could in any way be used to figure out their identity or location.**

Visual cues aren't the only way a source can be identified. By default, any digital image or video will contain metadata, which provides the machine-specific context around how a piece of media is created on a device. Metadata (such as timestamps, make and model of the devices used, and sometimes even geo-coordinates) are automatically inscribed into any media file, and can be used to add more detail to when, where, and by

---

[2] Dangerzone is a useful tool but is better suited for more technical users. We are hoping to see better documentation and workflows for it soon. In the meantime, see here for a quick introduction: https://tech.firstlook.media/dangerzone-working-with-suspicious-documents-without-getting-hacked

[3] Every email client will have a slightly different mechanism for blocking remote content, so just search for the name of the client and "block remote content".

[4] There's been at least one recent case study of a media organization that accidentally published a document with printer dots: https://www.eff.org/deeplinks/2017/06/printer-tracking-dots-back-news

[5] In general, we know very little about the geolocation done by governments, corporations, or security services. If you want a good introduction to how geolocation works and how journalists do it, check out the work of Bellingcat

whom an image was taken. **Always remove metadata from photos before publishing**, by either using tools like Exiftool to scrub the original data, or by submitting screen captures of the original media from your own computer to prevent your source's metadata from leaking out publicly**.**

Think about canary traps and whether they affect your decision on what to publish

Some companies and government institutions will secretly create many different versions of an email or a document. Every version might contain subtly different spelling, punctuation, spacing, or other features, which means that every person receives a slightly different email or document. That way, the moment it is leaked and published, the institution can easily figure out who was responsible for the leak.

Canary traps are usually a well-kept secret. Governments or corporations don't like to tell others if or when they use them.[6] **Talk to your newsroom about what your policy should be on publishing documents you receive from sources in full**. Some newsrooms will treat any non-public document as off-the-record. Others will fully rewrite the document but not publish the original. Others yet are not willing to rewrite or publish the original in full, only describing it or paraphrasing its conclusions.

Remember that much social media data is public

If a source asks a journalist to follow them on Twitter for a DM or comments on an Instagram post to get their attention, that will be public. If a journalist suddenly follows a bunch of new people, for example because they are starting to research a new topic, then those new accounts followed are public. Finally, when people create social media accounts, their first followers often consist of people whom they know in real life. Twitter, for example, no longer displays followers in chronological order, but it's still possible to figure this out with a little time and resources.

Anybody closely tracking a journalist, for example through a social media bot, might notice the new accounts and predict their new investigative or reporting interests. It's rare for journalists to be directly tracked on social media in such a way, and only sophisticated adversaries will likely go through the effort of doing so. Still, if you are for example working on a story about a corporation and communicating with sensitive whistleblowers, it might be worth keeping this in mind. You could, for example, use private lists[7] instead of public follows on Twitter: that way, you can still see the accounts' updates, but they aren't publicly linked to you.

It's also a good idea to have a way in which a source can contact you without leaving a metadata trail behind - you could for example publicly list a Signal number for your newsroom and ask sensitive sources to write to you there.[8] You could also ask for information through metadata-resistant portals like SecureDrop or GlobaLeaks, but they can be difficult to set up and maintain.

When working with social media, be aware of all those risks and be ready to create an alternative account that's not tied to your name if necessary.

---

[6] There seems to be at least one case in Poland where canary traps were used to catch a leaker in Poland, in combination with other investigative practices. See this piece (in Polish) for more details: https://www.tvp.info/356359/piotr-kownacki-ujawnil-raport-abw

[7] Twitter uses both public and private lists. When you add someone to a public list, they receive an alert about it. Not only that, but such lists can also be publicly viewable and searchable. Make sure to only use private lists for sensitive research.

[8] The Freedom of the Press Foundation also published a great guide on how to run a second Signal account.

Always verify safety numbers on Signal and WhatsApp

Skilled attackers have sometimes been able to take over people's phone numbers and could, by extension, take control of their WhatsApp or Signal accounts. Because of this, **turn on two-factor authentication** on your Signal and WhatsApp accounts and encourage sensitive sources to do so as well. This adds a second password to Signal and WhatsApp; even if someone cloned your SIM card or stole your number somehow, they would not be able to easily take control of your account without this password.

Always watch out for the notification that somebody's safety number has changed (Signal and WhatsApp). If you see this warning, then something has happened to the other person's Signal or WhatsApp account - they might have simply reinstalled the app or someone might have taken over their account. If the account has been taken over, the adversary would not have access to any messages you shared in the past but would receive messages you send from now on. It's good to **have a policy on what to do when a sensitive contact's safety number changes**: you could, for example, call them or ask them to send a selfie to make sure that it's still the same person who is behind the number and account.

Signal always warns you when safety numbers change. WhatsApp doesn't do so by default, but you can ask it to warn you every time.

Don't forget that Signal and WhatsApp accounts are also tied to phone numbers. In some countries, SIM registration is compulsory, so a phone number will always be tied to a person in a government database. Additionally, contacts on your phone (and on the phones of your sources) are likely to be synced to iCloud or Google accounts. If a list of contacts were ever revealed (by hacking into someone's account, for instance), it could be problematic if a source had the phone number of a publicly known journalist among their contacts. Signal does not store phone numbers but contact details can get leaked in many other ways, for example through another app that synchronizes your contacts or someone looking through your phone or that of a source. If you're worried about this, look at other options for a number that is unlinked from your identity, for example by purchasing **an additional SIM card from a country with no compulsory registration**.

Remember that this list isn't exhaustive and that sources can be surveilled in other ways

If you've followed all the recommendations we've listed above, then you will have already taken huge steps to making sure that your sensitive sources are more secure.

Still, there are other ways in which sources can be caught. They could have been captured on CCTV as they were getting documents. Additionally, offices with electronic door access usually track who entered and when.

Governments have access to mobile phone location data. If you are going to meet a very sensitive source and want to keep secret the fact that the two of you met, it's best for both sides to **leave mobile phones behind at home or in the office**. That way, anyone surveilling you through mobile networks will assume that you are still at home or in your office. If both you and the source suddenly switch off your mobile phones, it's far more likely to arouse suspicion.

Leakers are often caught not just using one method, but through many different data crumbs, for example a mix of canary traps, CCTV footage, mobile phone data, and others.[9] If you are working with very sensitive sources, it's a good idea to discuss your security strategy with your editor, have a plan in place for what happens if information about a source does get leaked, and maybe have steady legal support as well.

---

[9] As evidenced by the case study we mentioned earlier (in Polish): https://www.tvp.info/356359/piotr-kownacki-ujawnil-raport-abw

Many corporations and governments can also use document access controls and keep detailed logs of who opened which document or email at what time, when they logged into computers, and the like.

Summary of Security Considerations when Engaging Sensitive Sources

- If you're opening documents from unknown sources, use something like Dangerzone to turn them into safe PDFs
- Block your email program or webmail from downloading remote content
- Be careful when republishing documents in full. Either fully retype it or, if you suspect a canary trap, paraphrase or summarize before publishing. Talk to your editors and newsroom about what the best and most reasonable strategy could be
- Assume that every photo can be geolocated. Do not publish photos which, if geolocated, could endanger sources
- Always remove metadata from photos and documents before publishing
- Remember that social media data can be much more public than you think. Give sources a way to contact you without leaving a metadata trail - for example, by publishing a Signal number for your newsroom
- On Signal and WhatsApp: enable two-factor authentication, have a policy on what to do when a contact's safety number changes. Turn on automatic safety number warnings on WhatsApp
- Consider getting a second, anonymous SIM card. When writing to sensitive sources on Signal or WhatsApp, use this anonymous number and encourage them to have one, too
- If you are worried about being tracked through mobile phone networks, it's better to leave phones at the office before meeting a source

Conclusion

Made it through this list? Congrats! Even if you follow some of the steps above, you will already have done a lot to make yourself and your sources safer.

As usual, if you have more questions, get in touch with us and best of luck!!

Other recommended resources to check out

The CPJ guide to protecting confidential sources

With special thanks to Ashley Fowler, Jon Camfield, Jan Cibulka, Martin Shelton, and Martijn Grooten for their invaluable feedback and comments