Citizen Lab, Munk School of Global Affairs, University of Toronto

MUNK SCHOOL OF GLOBAL AFFAIRS

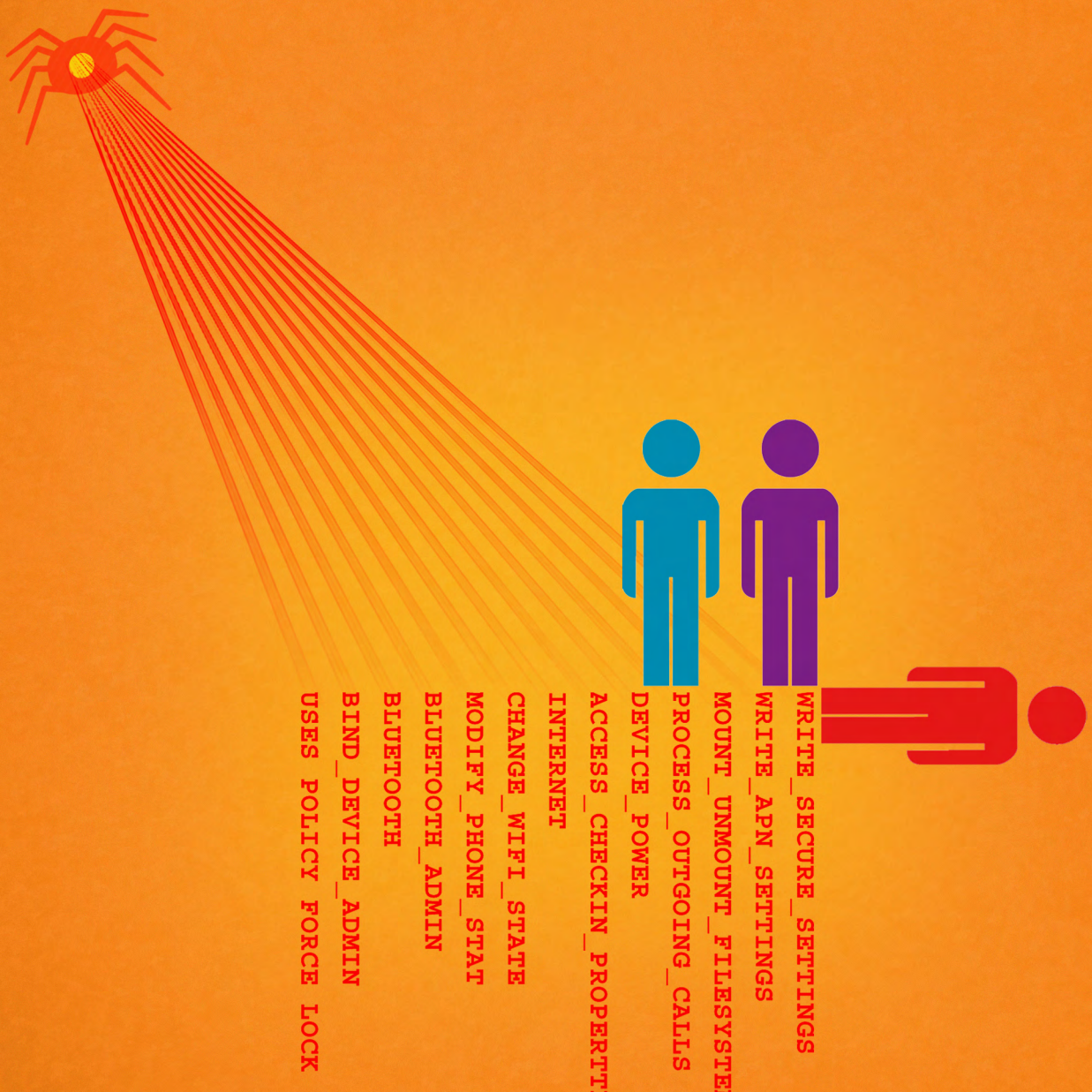
UNIVERSITY OF TORONTO

COMMUNITIES @ RISK

Targeted Digital Threats Against Civil Society

November 11, 2014

https://targetedthreats.net/



Acknowledgments

Communities @ Risk: Targeted Digital Threats Against Civil Society is a product of a collective effort conducted over a four year period.

The research and writing of this report was undertaken by Masashi Crete-Nishihata, Jakub Dalek, Ronald Deibert, Seth Hardy, Katharine Kleemola, Sarah McKune, Irene Poetranto, John Scott-Railton, Adam Senft, Byron Sonne, and Greg Wiseman.

In addition, we would like to thank Christopher Davis, Brandon Dixon, Phillipa Gill, Claudio Guarnieri, Bill Marczak, Morgan Marquis-Boire, Justin Wong, and Nart Villeneuve for their input and expertise, Jane Gowan for the report layout and cover design, and Andrew Hilts for the website design.

We thank VirusTotal for providing us with an intelligence subscription and to the developers of free information security tools, including Cuckoo Sandbox, URL Query, and PassiveTotal.

The Citizen Lab's research into targeted digital threats is supported by a grant from the John D. and Catherine T. MacArthur Foundation.

We would especially like to thank the 10 civil society organizations that participated in the project and made this study possible.

EXECUTIVE SUMMARY

> [Technology is] this funny thing where it's a life line, and then...maybe your ticket to jail."
>
> –Tibet Group 1

A civil society organization that works on China-related social justice issues receives an email from a program officer at one of its funders. She suggests that they review details of an upcoming meeting.

A Tibetan webmaster receives an email continuing a thread with several collaborators about developing a poster for an upcoming campaign.

These messages look like the mundane email traffic of everyday work. Most of us would not think twice about opening them. However, these two emails, and many more like them, are carefully crafted digital attacks.

These attackers had artfully composed the messages using private information that only the recipient and that person's group of contacts would know, suggesting that at some point someone else in their organization or community had already likely been compromised.

The emails contain file attachments implanted with malware that exploits vulnerabilities in programs like Adobe's PDF Reader and Microsoft Office, enabling distant attackers access to computers at the offices of the civil society organization, or the living room of the webmaster. The attackers then turn the computer into an ideal spying device: they take files, record keystrokes, and turn on the webcam and microphone. All of this monitoring begins with a seemingly inconsequential behavior: double clicking a benign-looking attachment.

These are not isolated incidents. The emails are real examples of an epidemic of targeted malware that is becoming a reality for human rights groups, journalists, and activists under threat from determined actors. Targeted attacks like these are organized in campaigns that persistently attempt to compromise systems and gain access to networks over long periods of time, while remaining undetected. They are custom-designed for specific targets and are conducted by highly motivated attackers. The objective is to extract information from compromised systems and monitor user activity.

Attacks like these are best understood as a form of espionage.

Abstract

Communities @ Risk: Targeted Digital Threats Against Civil Society is the culminating report of a multi-year, multi-group study on targeted digital threats. We define targeted digital threats as persistent attempts to compromise and infiltrate the networked devices and infrastructure of specific individuals, groups, organizations, and communities.

The study involved 10 civil society organizations (CSOs) that shared suspicious emails, network traffic, and other data with Citizen Lab researchers who conducted detailed, confidential analysis over a four year period. Citizen Lab researchers also paid site visits to the participating CSOs and interviewed them about their perceptions and the impacts of the digital attacks on their operations. Data from both the technical and contextual aspects of the research informs the report's **Key Findings**:

- **In the digital realm, CSOs face the same threats as the private sector and government, while equipped with far fewer resources to secure themselves.**
- **Counterintuitively, technical sophistication of malware used in these attacks is low, but the level of social engineering employed is high.**
- **Digital attacks against CSOs are persistent, adapting to targets in order to maintain access over time and across platforms.**
- **Targeted digital threats undermine CSOs' core communications and missions in a significant way, sometimes as a nuisance or resource drain, more seriously as a major risk to individual safety.**
- **Targeted digital threats extend the "reach" of the state (or other threat actors) beyond borders and into "safe havens."**

Remediation of the problem will require major efforts among several stakeholders, from the foundations that fund civil society, to the private sector, to governments.

Introduction

The Internet and other digital technologies are a transformative, disruptive force throughout society, impacting governments, businesses, and CSOs. For the latter, social mobilization, advocacy, policy engagement, and internal operations and management are now deeply intertwined with the same mass market communications platforms most of us use daily, from instant messaging applications to Twitter. CSOs manage, often with few resources, to accomplish remarkable effects thanks to these technologies, leading some to predict a worldwide flourishing of rights, democracy, and individual empowerment.

While the positive outcomes for free expression and access to information are evident, we are only now beginning to get a handle on the new risks that digital technologies introduce. Among those risks, arguably the most well reported on and widely discussed have been those related to documents leaked to the press by Edward Snowden (See "The Snowden Disclosures"). The documents show in detail how the US National Security Agency (NSA) and its "Five Eyes" allies have been able, with considerable effort and resources, to exploit the Internet and other digital technologies as tools of mass surveillance for national security and foreign policy aims. Separate from the Snowden leaks, there have also been a growing number of case studies and reports of journalists or human rights defenders being targeted by governments with malicious software (malware) or even commercial spyware. Through this reporting, a more nuanced understanding of the risks associated with the Internet and digital technologies is developing among CSOs and the foundations that fund them. Secure tools, trainings, and other forms of support are a burgeoning field. Individuals working in areas at risk are beginning to understand that those very same technologies that provide liberating means of communication and organization can also be sources of insecurity. However, much remains to be done, particularly in the area of systematic, evidence-based research of targeted digital threats.

For the past 10 years, Citizen Lab has researched the use and impact of digital technologies within civil society, focusing in particular on their unintended consequences as potential sources of insecurity or threat. Our aim is to apply a systematic, mixed methods approach to this research, combining technical and social sciences with field research. We also consider principles of international human rights law as an important touchstone for our research, for at least two reasons. First, some CSOs and individuals appear to be targeted as a direct result of their human rights-related investigations and advocacy. Second, the use of digital attacks undermine such actors' internationally-recognized human rights, including freedom of expression and right to privacy.

In 2009, we were part of a team that published the first open academic study documenting a major global cyber espionage campaign involving compromised computers in dozens of high-value targets around the world. The resulting report, entitled *Tracking GhostNet: Investigating a Cyber Espionage Network*, was followed up a year later with *Shadows in the Cloud: Investigating Cyber Espionage 2.0*. Both were organized as case studies starting with Tibetan groups as study subjects. Both generated unexpected and quite sensational findings concerning a range of other governments and businesses whose computers we found compromised by the same groups targeting our Tibetan study subjects. We hypothesized that these types of targeted digital attacks were likely not uncommon, and were affecting more than the few organizations we were studying.

Following these foundational case studies, we embarked on a plan to develop a multi-year, multi-group study on targeted digital threats. Our aim was to apply principles from comparative methods in other academic disciplines to the study of targeted digital threats against CSOs. **We define "targeted digital threats" as persistent attempts to compromise and infiltrate the networked devices and infrastructure of specific individuals, groups, organizations and communities.** Targeted digital threats are not widespread compromises of networked devices that affect

Civil society is feeling the heat around targeted attacks and surveillance and I think it's affecting the public sphere and our ability to feel comfortable communicating in what used to be understood as a free and open medium."

—Rights Group 2

[At the time of the 2008 Tibetan uprising] you could make all the noise you want in DC or in Seattle or in Paris, but when it came to actual Tibetans organizing on the ground inside...there was nothing... they had no knowledge, no capability... We saw...a generation of activists taken out because of the inability to support them safely..."

—Tibet Group 1

individuals or groups in an undifferentiated fashion. They are not the typical spam or financial fraud that one may encounter more or less randomly across the Internet—the equivalent of a digital "flu." Rather, they are focused on *specific targets*, they *persist* over a period of time, and they are motivated by *political objectives*.

Outside of this study, Citizen Lab and its collaborators have engaged in parallel research projects on targeted digital threats against CSOs. These projects include several pathbreaking reports on commercial spyware (see "The Market for Lawful Intercept") and targeted digital threats in and around the Syrian armed conflict (see "Syria and Targeted Digital Threats"). While we report primarily on the findings of the formal study in this document, lessons learned from those other projects inform our analysis. Together this body of work moves us towards mapping the targeted digital threat tactics and approaches of governments and other actors around the world and documents how these capabilities are used against CSOs.

Reflecting on the sum total of all of our various research projects on targeted digital threats, we observe that there are at least three distinct models that characterize the capacities and tactics of actors carrying out targeted digital attacks:

1. NATIONAL IN-HOUSE DEVELOPMENT AND OPERATIONS (APT)

In the first model, threat actors have capabilities and resources to develop their own customized malware and conduct wide scale operations. This level of capacity requires significant time and investment to develop, and is generally restricted to well-resourced actors like states. However, these kinds of operations can also be achieved through "cyber militia" groups that receive direct or tacit government support. Within the security industry, this approach is frequently referred to as the "Advanced Persistent Threat" (APT). At the high end of this model is the NSA's Tailored Access Operations (TAO) group that reportedly has a large and highly trained staff with significant budgets for developing and implementing targeted digital attacks. While not as advanced as the NSA, China-based threat actors have been developing custom malware and carrying out extensive campaigns for the last 15 years. The attacks we document in this study are prime examples of the work of threat actors within this model.

CHINA AND TARGETED DIGITAL THREATS

Public reports on malware campaigns originating from or related to China go back over a decade. In the past five years, the number of reports on these activities has exploded with high profile compromises documented against governments around the world and a large number of industries, including companies like Google, RSA, and Boeing. The United States has been particularly vocal on the threat these attacks pose to national security and commerce. The Commission on the Theft of American Intellectual Property claims intellectual property theft against the US is primarily orchestrated by China through cyber espionage and accounts for losses of up to 300 billion dollars a year. General Keith Alexander, former Director of the National Security Agency and Commander of United States Cyber Command, has called the theft of US intellectual property through cyber espionage the "greatest transfer of wealth in history."

While governments and businesses are often highlighted as victims, malware attacks against ethnic minority groups in China including Tibetans and Uighurs, and religious groups such as Falun Gong, go back to at least 2002, and possibly earlier.

Claims of attribution surrounding these attacks abound, with some analysts making direct connections to the Chinese government and military, and others drawing links to the Chinese hacker underground or universities. Conclusive proof that a targeted attack is the work of a state-sponsored attacker is often elusive. Regardless of how connected the Chinese government is to these attack campaigns, the years of documentation around these operations show that there are well-resourced and persistent threat actors originating from China.

2. RE-PURPOSED CRIMEWARE

The second model is best represented by campaigns conducted by parties involved in the Syrian civil war. These attacks primarily rely on basic Remote Access Trojans (RATs) that are circulated among hobbyists and criminals, but which we have found are deployed for political reasons and—in the case of Syria—in the context of armed conflict. This approach blurs the worlds of cybercrime and espionage, and is forged out of necessity and, to some degree, a kind of "do-it-yourself" mentality. These kinds of operations can be conducted by state actors and / or groups that may be directly sponsored, encouraged, or tacitly accepted by states. Early China-related operations followed this approach, but over time have become more organized and mature. A similar type of maturation process could occur in other contexts.

SYRIA AND TARGETED DIGITAL THREATS

In January 2012, it was becoming clear to Syrian opposition groups that something was going on with their computers. Suspicious messages and social media postings directed them to download documents and programs purporting to contain useful information. Troublingly, some of the files were sent from the accounts of individuals detained by the regime. Early analysis by the opposition led to the conclusion that they were being targeted by malware attacks. Researchers from a number of groups, including Citizen Lab, began investigating and were eventually able to develop compelling evidence linking the attacks to the Assad regime.

Our research on Syria consists both of close work with targeted groups to identify and track malware campaigns, and the use of signatures and other techniques to identify malware in the wild. Taken together, these methods provide a useful but necessarily incomplete picture of the targeting, as attackers regularly refine their techniques, thus reducing the likelihood of being observed.

The lessons we draw from the Syrian case align with those from the formal study. Namely, attackers exploit the pervasive, constant use of mass market communications tools by the opposition, just as do threat actors that target participants in our study. Additionally, we consistently observe sophisticated social engineering and well-informed targeting, rather than a high degree of technical sophistication.

3. COMMERCIAL SPYWARE

The third approach relies on the procurement of commercial "lawful intercept" products and services that provide actors with turnkey surveillance solutions. Companies like FinFisher and Hacking Team are actively marketing surveillance suites to governments, law enforcement, and intelligence agencies. Actors that do not have an in-house capacity to develop and operate targeted digital attacks can now buy these capabilities directly from these companies. The high cost of these products and the claim by vendors that sales are restricted to government clients make this primarily a state-centric route, although it is conceivable that non-state actors could be clients too. Citizen Lab research has identified troubling evidence that these products and services are ending up in the hands of regimes that are using these powerful tools to actively target civil society.

THE MARKET FOR LAWFUL INTERCEPT

In parallel to work on targeted civil society groups, Citizen Lab researchers have conducted extensive investigation into the global proliferation of so-called "lawful interception" malware sold exclusively to governments. These tools allow governmental purchasers the ability to gain remote entry into, and monitor the computers and phones of their targets.

Research published by Citizen Lab as well as other investigative groups has demonstrated that some governments and security services abuse these tools by hacking political opponents and human rights groups, both domestically and in other jurisdictions. Despite the potential for abuse, the market for these tools is largely unregulated, which has helped the governmental customer base grow, and likely led to substantial profits for developers. Our scanning has enabled us to develop global lists of suspected government users of this technology. Meanwhile, our close work with victims in the United States, the United Kingdom, the United Arab Emirates, and Bahrain has helped us document the abusive use of these tools against human rights defenders, journalists, and civil society groups.

Similar to the threat actors featured in this study, the attack tools require effective social engineering campaigns to gain entry to targeted organizations. This requirement also results in the creation and transmission of links and files that can be collected and analyzed by researchers. However, we have seen commercial tools that provide more advanced network injection functionality. While these tools can be technically identified, they are challenging to systematically detect.

The findings in this report primarily reflect our research on campaigns that have followed the first model—specifically how China-based threat actors are targeting CSOs. *Readers should be aware, therefore, of the limitations of our research*. We note that our research into targeted threats within the formal study is largely "China facing." Most of the participating CSOs have missions dedicated to human rights in the context of China, with Tibetan Groups representing the lion's share of the study. Our technical data in particular is based primarily on data shared with us by Tibetan organizations, given our existing contacts in the community and prior research for the *Tracking GhostNet* and *Shadows in the Cloud* reports, and is thus skewed toward consideration of attacks against Tibetans. The findings reflect, in other words, one model of threat actor: the model emanating from China (which may or may not be state-linked). While these considerations are specific to the Groups (attackers and CSOs) we analyzed in the four-year formal study, we believe the findings are generalizable to other contexts.

The report combines two major sections:

1. The Executive Summary (which you are reading now) provides detail on how the study was organized and why we feel it is important to read, a high-level overview of the key findings of the research, and considerations about next steps for several stakeholder communities in responding to targeted digital threats.
2. The Extended Analysis explains our methodology, and examines the detailed data we gathered during the study period. It is the evidentiary basis for the claims we make in the Executive Summary, and will likely be of interest to a more specialized audience (although we hope everyone will read it).

STUDY BY THE NUMBERS

- Participating Groups: 10 CSOs
- Study Duration: 2010-2014
- Number of Emails Submitted: 817
- Malicious Payloads Identified: 2,814
- CVEs Identified: 24
- 0-days Identified: 1
- Malware Families Identified: 44

The entire report is written in language that assumes little or no prior knowledge of malware, computer network attacks, or other technical details. We have attempted to define key terms and concepts along the way, and for those requiring some help navigating technical terms, we have included a detailed glossary.

Accompanying the Executive Summary and the Extended Analysis, we are publishing other resources for the research community, including indicators of compromise, and a guide for further reading. Links to all of the documents can be found at targetedthreats.net.

Framework and Methods

Today, many are aware that individuals, communities, and CSOs face digital risks. It is common to hear warnings not to open unsolicited email attachments or to log onto untrusted WiFi networks. The number and frequency of reports about cyber espionage campaigns or major data breaches contributes to this growing awareness. However, documented evidence of the precise nature of digital risks facing CSOs remains scattered and mostly anecdotal. The situation might be likened to growing awareness around the health impacts of smoking prior to the publication of formal epidemiological studies undertaken by the Royal College of Physicians of London in 1962. We intuitively sense that there is a problem, common sense supports it, but there is a lack of empirical evidence based on formal, systematic research.

FORMAL, COMPARATIVE ANALYSIS

This report is the first of its kind of which we are aware to take a *formal, systematic, and comparative approach to the study of targeted digital threats against civil society.*

The genesis of the project was our desire to build upon the focused case studies undertaken in *Tracking Ghostnet* and *Shadows in the Cloud*. While those studies involved technical forensics, interviews, and field research with the consent of several Tibet-related organizations, we felt a follow-on could be broader in scope, more rigorous, and self-consciously styled on formal comparative research methods common to the social and natural sciences. We aimed to better formalize the study of targeted digital threats, and begin making the topic a legitimate area of inquiry for academic research. At the same time, we were aware of a growing body of research in the private sector and within government on targeted digital threats, most of which is focused on attacks against industry or government agencies. Attacks on CSOs, on the other hand, have gone under-reported despite the fact that they are frequently included in the very same major cyber espionage campaigns investigated by those groups. We sought to rectify this imbalance.

KEY RESEARCH QUESTIONS

- Are civil society organizations selectively targeted for digital compromise? What commonalities exist across campaigns against civil society and entities in other sectors?
- What are the origins of targeted attacks against civil society organizations? Is attribution feasible? What, if any, conclusions regarding attribution can be drawn from the evidence obtained?
- What methods are employed to compromise civil society organizations? Have those methods changed over time?
- How technically advanced is malware used against civil society organizations, and how does it operate?
- How sophisticated are the social engineering techniques used against civil society actors?
- What level of knowledge or awareness do attackers exhibit concerning their targets?
- How do civil society organizations perceive and respond to targeted digital threats?
- What is the impact of digital threats on civil society groups?
- How can civil society best protect itself?
- What are the next steps for future research?

Our research plan had several components: First, we aimed to enrol a range of study groups to participate in the project over a *significant period of time*. Our feeling was that a manageable but larger sample size would give us a better indication of the range of digital threats. Doing so required a fairly substantial investment in outreach and engagement, in part to explain the framework of the study to interested groups, but also to assuage any concerns they might have about the risks of participating in the study. We needed to gain their trust. Our engagement included a public call on our website, email announcements, and outreach to individual organizations.

We were able to enroll 10 CSOs from a variety of sectors, eight of which are groups concentrated around rights issues related to China and Tibet. The concentration is the result of both our previous connections to some of the groups in prior research, and the fact that these communities have been targeted by digital attacks for at least 10 years. Many have a high level of awareness of digital threats, and also have a strong interest in being included in and supporting the aims of our study. To help control

(partially) for selection bias, we also enrolled two large human rights organizations focused on issues in multiple countries. (In subsequent phases of the project we expect to broaden our sample groups in size and focus.) We worked with the groups over a four-year period, observing their behavior, analyzing the malware samples we received from them, and interviewing them concerning their perceptions, practices, and common concerns. To protect the participants' confidentiality, we refer to them throughout the report in generic terms (e.g., Tibet Group 2, China Group 1, Rights Group 2).

FULL ETHICAL REVIEW

The second component of our research plan was that we applied for and underwent a *full ethical review of the study*, and received approval from the University of Toronto's research ethics office for our plan of research. Going through the ethical review was important for the following reasons.

First, much of the information shared with us is of a highly sensitive and confidential nature. It may reference detailed internal and strategic matters, or personally identifiable information that may put individuals at risk. Moreover, the fact of actual or potential compromise is itself a sensitive topic. Groups may be concerned that disclosing information about an attack could reveal vulnerabilities and encourage further intrusions. They may also fear that disclosure could subject them to embarrassment with communities served, funders, or the public at large. (On the other hand, some groups may see targeting as a sign that their work is significant enough to warrant such attention.) It is therefore essential to ensure that the rights of participants in the study are protected, and that they retain control over the use of confidential information and data.

Second, research on targeted threats—particularly when it involves technical research—may lead to unexpected results that raise ethical dilemmas. Technical investigations into malware may uncover command-and-control interfaces used by attackers, or repositories of sensitive data culled from targets. Researchers may even find themselves with the ability to issue commands directly to compromised computers. Research might also lead back to those suspected of orchestrating the attacks, or to previously unidentified targets, presenting questions of notification to law enforcement or victims, and whether to publicly disclose the suspected culprit(s). A pre-existing framework to guide an appropriate response is beneficial when encountering such circumstances.

The research protocol we submitted to the University of Toronto Office of Research

Ethics includes sections on conflicts of interest, research rationale and methods, participants, recruitment, possible risks and benefits, the consent process, and confidentiality. Additionally, Citizen Lab enters into a formal written consent agreement including confidentiality provisions with each participant, follows up with an oral explanation of the parameters of the study, and provides participants an opportunity to ask questions and discuss details before enrolling. Moving forward, we intend on further engaging the ethical questions around this type of research as a topic of study itself.

MIXED METHODS APPROACH

One of the distinguishing features of the report is the combination of methods employed to undertake the research. Citizen Lab has employed this *mixed methods* approach for several years, and it is particularly well suited to the challenges of our study.

Technical analysis

At the core of the analysis is the technical data we collected from the groups, including malware samples and network traffic. We analyzed malware samples using static and dynamic analysis tools as well as manual analysis to extract information on exploits, malware functionality, malware family, command-and-control infrastructure, and other properties of the malware code. By examining patterns in malware families, development cycles, shared infrastructure, and social engineering tactics, we identified relationships between attacks and, where possible, linked them to known malware campaigns and threat actors.

Field research, site visits, and interviews

We interviewed the study subjects in a semi-structured process. These sessions were, where practical, recorded, transcribed, and analyzed by researchers. We undertook site visits to all but one of the study groups to get a better sense of their on-the-ground experience. The input of the study groups gave us unique insight into their perceptions of targeted digital threats, their capacity to deal with them effectively, and the impact these kinds of threats can have on their daily operations. We identified emergent themes from the interviews that provide insights into how CSOs perceive targeted digital threats. A full overview of these themes is available in the Extended Analysis.

Contextual analysis

In addition to technical analysis and interviews, we conducted legal, social, and political analysis, including research of contextual details and social engineering particular to the attacks—such as timing, language employed, topic flagged in the email text, relevant

political climate, etc. This analysis relies on the background information provided by participating organizations, open source intelligence, and the expertise of our team in international law and human rights, and regional and country-specific history and politics.

We also engaged with civil society, including human rights groups and digital defenders reporting experiences of targeted digital attacks; exchanged information with individuals in the security community (including our technical advisory board); and liaised with other groups undertaking similar research.

OPEN DATA, OPEN METHODS, OPEN TOOLS

We have attempted to share as much as possible with the wider research community the data, the tools, and the methods we gathered, used, and developed during the course of this study. Accompanying the release of this report are datasets we are making openly available, including YARA signatures of malware families, MD5 hashes of samples, and lists of command-and-control servers. The intention of this data release is to help security researchers and network administrators identify and defend against the threats analyzed in our dataset. Additionally, we developed the Targeted Threat Index, a metric to characterize and quantify the social engineering and technical sophistication of targeted threats and assess their relative risk. We hope this metric will be adopted by other researchers and applied to other datasets. During the project, we developed a web-based malware repository "The SHARK" for managing our dataset. Although many similar platforms exist and are used by the security industry, most of them that we evaluated are either proprietary and / or were not suited to our specific requirements. We are in the process of planning a new version of The SHARK that will encompass what we have learned from developing our internal system, and it will be released open source. By openly publishing datasets, methods, and tools we hope to encourage and assist other researchers who may be interested in targeted digital threats and begin the process of building on accumulated knowledge.

Key Findings

We have identified several important findings regarding targeted digital threats based on data obtained through the formal study, our ad-hoc research, and our own long-term experience. These findings inform our recommendations to stakeholders. (An additional list of specific, more detailed findings is included in the Extended Analysis.)

1. In the digital realm, CSOs face the same threats as the private sector and government, while equipped with far fewer resources to secure themselves.

In recent years, a growing number of high profile security industry reports have cast a spotlight on targeted digital attacks against Fortune 500 companies and government agencies. These reports have received broad press coverage, triggered major public policy debates, and brought about government action. One of the main findings from the technical investigations is that some of the groups participating in our study are targeted by the same *threat actors* using the *same techniques*, *tools*, and *infrastructure* as those highlighted in industry reports. The inset "Campaigns Targeting CSOs and Government/Industry" provides a snapshot of the connections.

CAMPAIGNS TARGETING CSOs AND GOVERNMENT/INDUSTRY

Through cluster analysis that groups attacks by common malware, development patterns, shared infrastructure, social engineering tactics, and other indicators, we have identified ten distinct attack clusters of which four have clear connections to campaigns that target government and private industry. These findings echo previous reports going back to at least 2008 (e.g. Tracking GhostNet) that have also shown threat actors targeting governments, private industry, and CSOs.

- APT1 (Reported by Mandiant)
  - » Targeted 141 organizations from 20 industry sectors
  - » Targeted Tibet Group 1, compromised Rights Group 1
- DTL Campaigns (Reported by FireEye)
  - » Targeted government and 11 industry sectors
  - » Targeted Tibet Groups 1, 2, 3, and 4
- NetTraveler (Reported by Kaspersky)
  - » Targeted 350 organizations from NGOs, government, and industry
  - » Known to target Tibetan and Uyghur CSOs
  - » Targeted Tibet Groups 1, 2, 3, 4, and 5; China Group 3
- PlugX Campaigns (Reported by TrendMicro, AlienVault)
  - » Targeted companies in Asia and US, and Tibetan CSOs
  - » Targeted China Groups 1 and 2; Tibet Groups 1 and 2

For example, we found evidence that the prolific threat actor known as "APT1"—also referred to as "Comment Crew" or "Byzantine Candor," which is known to have compromised numerous government entities and Fortune 500 companies—targeted Tibet Group 1 and significantly compromised Rights Group 1. The malware we examined incorporated much of the same code and used one of the same command-and-control servers as the APT1 attacks previously documented by security firm Mandiant.

Evidence of this type of cross-targeting is not coincidental. It shows that the actors behind campaigns like APT1 and others like them, place the same strategic value on penetrat-

ing CSOs as they do on companies and governments. Yet while digital intrusions against private sector or government actors have resulted in high-profile media coverage, criminal investigations, and increasingly forceful national policy responses, *few avenues for escalation exist when those same intrusions are directed against CSOs*. Governments that cooperate closely with, and speak loudly on behalf of industry actors concerning intellectual property theft, have not taken the same approach to protection of domestic civil society, which involves the far thornier issues of right to privacy and freedom of expression.

For example, while the US government has taken a strong political stance on Chinese cyber espionage against US companies—even filing a criminal indictment against members of the Chinese military for alleged hacking—we have not seen the US Attorney General demand an end to the persistent attacks of US-based NGOs that work on China-related human rights issues, despite the threats to life and liberty that could result. The political capital such a move would require, particularly in the aftermath of the Snowden disclosures, is perhaps considered too great by many governments; even those with active Internet freedom policy agendas have not fully addressed the question of cyber espionage against civil society.

Meanwhile, CSOs are hard-pressed to resolve matters themselves. CSOs reported to us an understanding of some of the digital risks they face, but in many cases noted a lack of capacity and resources to dedicate to the problem. With rare exceptions, they typically do not have the funding to hire technical security experts, or the opportunity to engage with government on digital defence or overall policy in a manner that protects their security and confidentiality needs. Some barely have dedicated IT staff, let alone staff that can handle APTs. Even if CSOs are able to undertake basic remediation after an attack, they are unlikely to be able to conduct the technical investigation and training necessary to fully understand and mitigate future threats.

If digital attacks of CSOs continue to spread unchecked, we risk the gradual erosion of many of the core institutions of a vibrant democratic society: NGOs, foundations, independent journalists, activists, and others—all of which have experienced and continue to experience targeted threats. The shared experience of targeted digital threats among civil society, the private sector, and government could lend itself to sharing of threat information and coordination of prevention and defense. Moreover, all three sectors stand to benefit from a comprehensive financially- and politically-supported bulwark against targeted digital threats; given the diversity of attacker targets, zero tolerance for and investigation of today's attacks against CSOs may help prevent tomorrow's attacks against a major company or government institution.

2. Counterintuitively, technical sophistication of malware used in these attacks is low, but the level of social engineering employed is high.

Targeted attacks against CSOs are rarely examples of 'technical wizardry.' Throughout the course of our study, we found that attacks frequently employed technically unsophisticated malware (relative, that is, to malware used by financially motivated cyber criminals and commercial lawful intrusion kits), some of which have been widely reported on for years. (Our Targeted Threat Index provides a detailed analysis of how we measure and rank sophistication.) Similarly, the majority of exploits we observed are for vulnerabilities that have been patched for long periods of time. In four years of documenting attacks using over 22 different exploits (CVEs) we observed only one zero-day exploit, suggesting that attackers targeting CSOs rarely see the necessity of using zero-days against what could be considered "soft targets." This is not to suggest, however, that digital threats against CSOs never utilize advanced malware or zero-day exploits. We have encountered more technically sophisticated malware outside of the study, and in particular in our research of the commercial spyware used against CSOs.

Still, attackers appear to employ malware that is only as technically advanced as it needs to be to generate results, investing fewer resources to rely on known exploits so long as their targets remain susceptible to them for behavioral reasons. This approach works because key factors determining whether a compromise occurs are typically behavioral rather than technical in nature: whether the user triggers the exploit by choosing to open a malicious file or click on a malicious link; and whether the user has kept software fully up-to-date with all security patches that would render known exploits ineffective, which requires current licensing of the software (not possible for pirated copies sometimes utilized by under-resourced CSOs and activists). Once the compromise occurs, basic malware is no less dangerous than more advanced malware—even unsophisticated exploits can permit installation of RATs providing the ability to search for and exfiltrate files and contacts, activate a device's video and audio recording, and log keystrokes.

We just have never had the time to do any forensics on what we assume are like denial of service [attacks]...the load on the server starts to rapidly increase, and [there are several] IPs that are very suspicious and the only thing that we can do is mitigate, fix, and just move on."

— Rights Group 2

At the same time, congruent with a lack of emphasis on technical sophistication, *we find greater sophistication around the social engineering employed in attacks against CSOs*. Social engineering is an attacker's method of crafting the delivery vector for the malware—typically

an email—in a manner designed to entice recipients to open the infected payload. Attackers often "spoof" the sender identity to appear as someone the target already knows and trusts; reference timely and target-specific issues and events; repurpose real content taken from other sources of interest to the target; or attempt to exploit the emotions of the target by addressing sensitive, provocative, or inflammatory subjects. Good social engineering thus requires some knowledge of a target's contacts, areas of interest, and current priorities or activities. It is likely that attackers conduct some form of preliminary reconnaissance or otherwise "study up" on their targets to develop their social engineering, perhaps drawing on social media and other open source information, or leveraging information or credentials gleaned from existing access to the systems of others within the target's circle of trust (what might be called "collateral compromise"). *Thus attackers appear to invest primarily in knowing their targets, rather than creating or purchasing advanced malware*.

The importance of social engineering relative to technical sophistication raises two major issues to consider in addressing targeted digital threats. First, on the positive side, our findings suggest that in many instances behavioral modifications and sensitivity to commonly relied-upon social engineering techniques may reduce the susceptibility of CSOs to targeted attacks. User education and awareness campaigns within communities at risk may help CSOs and others contend with evolving threats and adaptive techniques, especially if known risky behaviors (e.g. opening attachments or clicking on links from unverified sources) are widely communicated and understood.

Second, significant negative impacts flow from attackers' reliance on social engineering that require a systemic response. Constant use of socially engineered emails as "bait" creates an environment in which it is increasingly difficult to authenticate genuine content and digital trust is eroded. At the same time, for many CSOs, responsiveness to digital communications and use of digital platforms are essential for the conduct of their work. For example, a malicious email that appears to be from an important sender, such as a funder, will likely be opened if that funder has not agreed upon secure means of contact with the CSO in advance. *Coordinated, standardized measures for encryption and authentication among civil society actors and those entities with whom they are in regular contact (e.g., funders, journalists, and government officials) should be seen as a critical priority* (see the "Next Steps" section below for elaboration).

3. Digital attacks against CSOs are persistent, adapting to targets in order to maintain access over time and across platforms.

The attacks against CSOs that we analyzed in our formal study, and that we have observed in our other research projects, are persistent and evolve in response to defences or changes in target communications behavior. By "persistent" we mean that the intrusion is designed to take place over a substantial period of time, avoiding detection, gathering and exfiltrating data, and preserving an attacker's ability to issue a variety of commands to the infected system. They require a non-trivial investment of time and resource from a threat actor, in order to surreptitiously acquire access, monitor the infected system, search for and select data of interest, and maintain a low profile throughout the compromise. While the technical tools permitting such action run the gamut in sophistication and cost, ongoing human involvement—*what amounts to a commitment to the target*—is apparent. In the case of CSOs, that work on sensitive human rights issues and generally do not possess financial assets that would entice an attacker, it is highly probable that the motivation behind such an intrustion is political.

We have also found that attackers are *adaptive*, modifying or designing attacks for use against specific software (including mobile applications) and hardware to reflect new and emerging methods of communication among their targets. Attackers exhibit an evolving awareness of civil society technical trends and defenses, which is reflected in their attack techniques. As a general practice, attackers make improvements to the malware they employ. For numerous malware samples in our study we observed several versions of the malware appearing over time, showing evidence of technical improvements. Adaptations, however, go well beyond malware maintenance. They encompass a wide range of responses to new platforms and behaviors.

As civil society actors have leveraged new technologies to advance their goals, attackers have done so as well, designing social engineering strategies and malware around the technical platforms that have become popular with their targets. For example, Citizen Lab and other researchers have documented a rise in Mac and mobile malware. While the majority of the malware we observed in our study targeted Windows operating systems, we also observed malware designed for OS X and Android. Indeed, Mac malware is increasingly paired with Windows malware, allowing attackers to compromise the target's computer without concern over which platform is used. In one instance, Tibet Group 1 and Citizen Lab tested the responsiveness of the attacker(s) behind one particularly well-crafted spoofing attempt by replying to the email, stating that the attached (malicious) Excel file could not be opened on the recipient's Mac. Within four days, the attacker diligently followed up with a new file containing Mac malware.

Study participants have also experienced ongoing targeting of the various cloud-based and popular communications programs on which they rely. CSOs have flagged Skype, Twitter, Gmail, and mobile devices and applications as vectors of confirmed malicious activity. The latter in particular present what Tibet Group 1 terms "a whole new battleground." For example, in late 2012 Tibetan community members began discussing alternative applications to mobile messaging application WeChat—owned by Chinese company Tencent—following concerns raised about its security. As part of this effort, an information security expert within the community sent an Android application package (APK) file for the alternative KakaoTalk messaging application to a private contact. Shortly thereafter, attackers circulated a maliciously repackaged version of the KakaoTalk file to Tibetan targets. The file was implanted with malware that added system permissions allowing attackers to collect user contacts, SMS message history, and cellular network location. The attackers were able to acquire the email and file because the original recipient's account was compromised.

4. Targeted digital threats undermine CSOs' core communications and missions in a significant way, sometimes as a nuisance or resource drain,more seriously as a major risk to individual safety.

CSOs may experience a range of impacts resulting from targeted digital threats. *In the most serious cases, staff or individuals with whom they are in contact may experience physical intimidation, abuse, detention, or imprisonment by authorities that stems in whole or in part from surreptitiously monitored communications.* Although digital surveillance may not be the proximate cause of this harm, it provides the authorities with an opportunity or rationale that may not otherwise have existed to take such action—for example, when digital evidence reflecting opposition to government policy serves as the primary basis for sentencing an individual for subversion or subjecting them to torture. In environments where mere contact with a CSO may heighten scrutiny of an individual, when digital records reflecting such contact are stolen, physical harm is a real possibility.

> *It's like when you do all this work to secure peoples' systems from surveillance and trying to help [them] avoid Chinese authorities monitoring them, and then everybody installs WeChat on their phones—so it's like, 'forget your laptop, forget the desktop, forget all of it—you've just perhaps given them complete access with this...' Everything we do is undermined overnight by this app that everyone is using... [Tibetans] adopt this stuff super fast. Especially when it's free...because it just facilitates community."*
>
> —Tibet Group 1

The psychosocial impact of targeted digital threats on

CSOs and activists is also significant and requires further attention. Staff of CSOs where intrustions are suspected or discovered report a variety of psychosocial effects, including a sense of violation, a state of fear of physical harm to themselves or loved ones, and chillings effects on their speech and use of technology. Groups particularly hard hit over the years have reported a loss of morale or "malware fatigue" (feeling like the threat has existed forever and cannot be escaped), which can lead to feelings of resignation and to abandoning security practises. There are reputational consequences to digital threats as well. Despite their ubiquity, exposure still carries a stigma in certain contexts, as a targeted group may be unfairly perceived as somehow to blame for "allowing" an intrusion or serving as the "bait" in a spoofed email or other attack vector. Or, a malicious email may circulate damaging misinformation about a person or entity. These impacts affect not only the will and ability of CSOs to carry out their missions and properly prevent or remediate a digital compromise, but also staff health and retention, and adoption of important digital platforms over the long term.

The most common impacts of targeted digital threats are the financial burdens of preventing or remediating intrusions, and undermined organizational efficiency—the "nuisance value" of the intrusion. CSOs, often on tight budgets, can easily incur extensive security costs. Security assessments, remediation, secure communications infrastructure, and technically proficient staff are all expensive, and typically priced for a commercial market, not struggling nonprofits. In addition, CSOs may need to spend considerable staff time identifying and notifying people whose communications were exposed to the attackers. This effort may sap the capacity of CSOs to conduct their primary human rights-focused missions.

Finally, one critical impact unique to targeted digital threats is their potential to wholly degrade the communications of CSOs and, as demonstrated by the Tibetan experience, entire communities. An essential element of civil society work is communication with the constituents served and with those entities CSOs wish to reach through their advocacy. Communication is a crucial factor in conducting research, obtaining important information on topics of concern, and disseminating messages concerning such topics. Targeted digital threats exploit the importance of communication to CSOs, undermine their methods of reaching constitu-

> *We were in the middle of a Skype conversation and we could hear screenshot sounds over Skype. Both of our computers were compromised and we had to clean up... It wasn't totally the end of the world, although it felt horrible...like a huge invasion... I think it sort of paralyzed us emotionally... for a few days."*
>
> —Tibet Group 2

ents and audiences, and create a climate of fear and lack of trust. It is possible that the goal of certain targeted digital threats may ultimately be to make communication more troublesome or raise the costs of communication for civil society.

5. Targeted digital threats extend the "reach" of the state (or other threat actors) beyond borders and into "safe havens."

Just as technology allows diaspora, exile movements, and international human rights groups to extend their reach and have greater connections with each other and the communities they are trying to support inside countries of concern, it also allows threat actors to do the same—with malicious intent. The China and Tibet Groups in our study are all advocating for issues from outside of mainland China. Rights Groups 1 and 2 are hubs that support regional offices spread around the world. Groups tend to perceive (quite reasonably) contacts, offices, staff, and associates closest to the adversary as the most at risk and their communication links between these entities as the most sensitive.

What might be easily overlooked, however, is the extent to which digital espionage also provides threat actors a means of leverage over individuals and groups that are located beyond the physical reach of repressive regimes. Individuals within these communities often faced persecution in their home countries, and established themselves elsewhere to seek refuge from violations of their human rights by the state. *In the universe of targeted digital threats, no such safe havens exist*. Attackers target individuals and groups outside of their jurisdictions to track those inside who have connections abroad, and / or to monitor activist movements and organization in the diaspora. Our research has shown that exiled journalists and human rights workers who have become naturalized citizens or refugees in democratic countries have had their computers and mobile devices compromised, their communications monitored, and their movements tracked—*as if they were still in the country from which they fled*. For those who assume that leaving a repressive country for one where civil liberties are protected solves the risks around persecution, targeted digital threats reopen the issue. Even those individuals who have never lived in the country and were born abroad can be drawn into the tentacles of a far-off regime as a consequence of their political advocacy.

> *Files were literally disappearing from our server... We don't know how much was actually taken... They were clearly letting us know that the files were gone... It took about a week of rebuilding and diagnosing everything... We had to order new servers, we had to write everything and then we had to reload everything..."*
>
> —China Group 1

THE SNOWDEN DISCLOSURES

Beginning in June 2013 onwards, a stream of highly classified documents leaked by former United States National Security Agency (NSA) contractor Edward Snowden has provided the public with an unprecedented view of the highly classified capabilities of the world's most powerful signals intelligence agency, the NSA, and its allies in the United Kingdom (GCHQ), Canada (CSEC), Australia (ASD), and New Zealand (GCSB). They show an extraordinary effort across every layer of the global telecommunications infrastructure, from the code to satellites and everything in-between, to infiltrate, collect, and even subvert or destroy data that passes through it. The impacts of these disclosures, many of which are too early to discern, are far-reaching, and have generated intense debates about the proper balance between security and privacy.

With respect to our study, at least two considerations stem from the disclosures. First, we did not encounter in our research any concrete evidence of NSA or allied malware attacks or espionage campaigns. Unlike those that we documented and which are generally assumed to originate in some manner from China, any analogous operations undertaken by the NSA and its allies would likely be very difficult for us to discern given the high level of their sophistication and the steps undertaken to obfuscate their attribution. It is important to be clear that our lack of material evidence of such attacks does not mean that they did not or will not happen; indeed, Edward Snowden himself testified to the Council of Europe that "The NSA has specifically targeted either leaders or staff members in a number of civil and non-governmental organizations...including domestically within the borders of the United States." The additional evidence provided by the disclosures may, over time, help inform future research into any such digital attacks, and we certainly intend to take them into consideration in subsequent Citizen Lab research.

The second consideration concerns perceptions of risk. Whereas prior to Snowden's disclosures vague concerns about widespread digital spying were voiced by a minority and sometimes trivialized, afterwards those concerns have been given real substance and credibility, and are now increasingly seen as a practical risk that requires some kind of remediation. After Snowden, there are now many more organizations offering security tools and trainings from which CSOs can benefit.

As one of our study participants stated: "I don't think it was until those attacks manifested at the end point of the user laptop that people really cared... [B]ecause that is visible for users in places that they understand—again your email, your Twitter account—even if they don't understand the implications, the connections; how your email is the gateway for most of your life, they now see it as something real and personal... [T]he paranoia is not for crazies anymore." —Rights Group 2

Responses and Next Steps

Our research into targeted digital threats provides a window into a troubling set of problems affecting CSOs, and thus by extension the health of civil society networks worldwide. We have identified some urgent considerations and next steps that should be taken to address the problems and begin the process of crafting effective solutions. Some of these points were raised by Citizen Lab study participants themselves. It is important to emphasize that solutions will require the involvement of multiple stakeholders, and there is no one single solution, technological or otherwise, that will stem the harm to CSOs from targeted digital threats. Accordingly, the following section puts forth considerations for next steps across multiple sectors.

FOR CIVIL SOCIETY GROUPS

CSOs are in the midst of what is likely to be a protracted contest for security, rights, and openness in the digital realm. Digital security considerations must inform their actions. At the same time, digital security is not (and should not be) their number one priority. Efforts to integrate digital security solutions with their operations must be aligned with their core mission. Initiatives to address digital security within these organizations must account for the organization's purpose, needs, and constraints. Nevertheless, there are fundamental actions CSOs can take to empower themselves. These actions complement, but do not replace, necessary technical and financial investments.

Document incidents

A relatively straightforward (but often overlooked) aspect of addressing targeted digital threats is the documentation of incidents by CSOs. Understandably, when experiencing a digital compromise CSOs may direct their attention exclusively to remediating the problem and recovering any lost material. Documenting the details of precisely what happened, and preserving attack vectors, malware, or compromised devices for analysis and digital forensics, are likely far down the list of priorities. Yet this step could significantly enhance individual and collective knowledge of targeted digital threats, as well as

the ability of CSOs to prepare for future attacks or seek justice for past ones.

One of the main challenges in researching targeted digital threats, particularly around civil society compromises, is the lack of concrete data regarding the problem. Basic facts regarding the number of incidents experienced by CSOs, the nature and timing of the incidents, the suspected vector of compromise (e.g. malicious email, drive-by download), the individuals involved, and the impact of the attack are rarely kept in a systematic fashion. CSOs are often quite capable, however, of keeping such documentation, as many do this already for physical incidents, and the associated burdens are relatively low—primarily an investment of a small amount of staff time. Maintaining copies of the malware itself or imaging infected machines for digital forensics presents additional complexity, but is something that could also be considered by CSOs that have sufficient technical support.

Benefits of standardized documentation of targeted digital threats could include: better understanding among CSO leadership and funders of the current digital risks to the organization, and the areas requiring additional resources (funds, training, tools) or change in practices; preservation of evidence that may be essential to legal claims, or other advocacy CSOs may wish to pursue; and, establishment of a repository of targeted threat data which, if shared, would inform a variety of investigations by researchers, activists, and others into evolving digital threat patterns, trends, and potential solutions.

Share knowledge and coordinate

As this Citizen Lab study and much other research has shown, no entity or individual is immune from digital threats. A large number of CSOs have already gone through the process of discovering, mitigating, and recovering from a compromise; those who have not are increasingly aware of digital risks and the need to prepare for them. Given the common experience of civil society actors in confronting and responding to this problem, a *collective approach to digital threats* may yield greater benefits than attempting to tackle these threats in isolation.

First and foremost, CSOs will know they are not alone or to blame in their experience; second, details of attacks can be shared so that others have more knowledge of

> *[I]t all comes back to public awareness, education... If we could break it down for people in a way that they understand... and paint the bigger picture, it has an impact. Tibetans are probably more apt to listen than other communities because the stakes are so high...it's just about the time and resources to stay on top of [the risks]."*
>
> —Tibet Group 1

current threats and can take preventive measures; and third, successful responses to the problem—perhaps emanating from entirely different communities or areas of interest—can be studied and adapted by other CSOs. We found that the Tibetan CSOs in particular have made significant strides in raising awareness of digital risks and encouraging digital hygiene through their adoption of a collective, community-based approach to the issue, including development of educational resources that provide security information reflecting Tibetan culture.

CSOs should also consider involving funders in collective efforts, communicating with them regularly about security issues and incidents. Digital security is an area in which funders can play an important role (see our specific recommendations to funders below), given their high-level vantage of the cross-sectoral trends affecting CSOs with which they work, and their capacity to bring resources to bear on the problem. They possess a unique ability to coordinate among CSOs. Funders, however, are not always aware of the digital security challenges faced by particular grantees. Sharing information regarding digital threats with funders can help illustrate areas of need, cultivate support, and disseminate learning across a significant portion of civil society.

Encourage a culture of digital security awareness

We frequently heard from CSOs staff were reluctant to confront such a complex problem as targeted digital threats, and that responsibility for digital security was often assumed to be siloed with the few individuals associated with the CSO who had technical expertise. One of our key takeaways from the study, however, is that individual human behavior is a critical facet of exposure. To address this problem, CSOs should gradually *build a culture* in which all staff, regardless of technical background, feel some responsibility for their own digital hygiene. While staff need not become technical experts, CSOs should attempt to raise the awareness of every staff member, from executive directors to interns—groups are only as strong as their weakest link—so that they can spot issues, reduce vulnerabilities, know where to go for further help, and educate others. Of course, there is no way to anticipate and warn against every form of digital threat; new technologies and new threats

> *What I tell [staff] is 'use your common sense, if you see something...suspicious, do not open it, or at least forward it to somebody to reveal it before you do anything about it.' But sometimes, I know it's probably difficult. I do not have any clever way to tell them, 'If when you see this, don't open it.' It's sometimes very difficult, [there are] just too many varieties of attacks."*
>
> —China Group 1

emerge constantly, outpacing security awareness. In such an environment, it is important for CSOs to *develop a framework for critical thinking and informed decision-making* by their staff about digital threats, not tethered to any specific application, device, attack vector, or situation.

FOR FUNDERS

Funders are uniquely positioned within the civil society landscape to contend with targeted digital threats. Both funders themselves, as well as the grantees they support, are at risk for politically-motivated digital compromise. Funders thus have at least two core responsibilities related to targeted digital threats:

1. To help their grantees implement better security.
2. To secure themselves (thereby also preventing collateral compromise).

This dual responsibility presents a number of challenges and opportunities to grantmakers, who are, we suspect, still engaged in an internal learning process about their own digital security.

Securing grantees

Many of the CSOs discussed in this report are dependent to some degree on external funding. Grantmakers are no stranger to finding strategies to help their grantees manage risk. For the past decade or more, funders with grantees operating in high risk areas (e.g., human rights organizations) have placed higher priority on supporting the *physical* security of grantee organizations. However, as this report makes clear, physical and *digital* security are increasingly interconnected. Indeed, in the current climate of risk, *a lack of attention to digital security can erode the gains made by investing in grantees' physical security.*

Funders are in a unique position to develop programs and funding lines that could help grantees make measurable improvements in their organizational security. However, funders face challenges around getting this right. *Program staff, while often aware of the specific political and physical threats to grantees, are generally unequipped to evaluate digital threats to grantees because of the kind of technical expertise it requires*. This extends to their ability to evaluate the quality and appropriateness of security solutions offered by third parties. This issue becomes particularly sensitive in the context of funding digital security training and related services for CSOs that face direct threats from state and non-state actors; uninformed training, advice, and/or flawed products could have grave repercussions. Moreover, an occasional training is not an acceptable substitute for serious support with digital security. Funders should

ensure that the security programs they do support are not sporadic, piecemeal, or focused only on encouraging changes in behavior without comparable investments in well-developed technology, policies, and practices that can help organizations attain and maintain better security.

Funders need to identify and build robust practices around implementing and measuring the success of any programming around digital security for *entire* program portfolios. However, many Monitoring & Evaluation consultancies, at time of writing, have limited specialized expertise in digital security. Identifying the right partners will be important if digital security is to be given a sustainable place in grant portfolios. Some funders may also seek to make large grants directly to digital security support organizations. We encourage funders to conduct extensive diligence, and seek expert advice, before doing so.

We think funders may have an untapped internal capacity here: *the CTO and other technical staff may be able to help evaluate potential providers of security services, or help to choose consultants who will do so.*

Funders are also well-placed to gather critical data regarding targeted digital threats experienced by their grantees, as they can do so in an aggregated and anonymized fashion. Such aggregation represents a promising avenue for collaboration between funders, and a mechanism to provide concrete data—with identifying information carefully anonymized—about the civil society experience of targeted digital threats and the results of funder efforts to support security.

Funder: secure thyself

Funders constantly handle sensitive information about their grantees, and may be well aware of the ways in which it could be used against them. Most would be horrified to discover that their handling of confidential material *brought risk to their most sensitive grantees*. However, in some cases this type of compromise is clearly happening. While outside the scope of data collected in our report, there appears to be an epidemic of compromises against Western NGOs with international

> *Certainly there is generally... a whole lot of cluelessness on the part of funders... Part of organizational development should be [technological skills and the utilization of] modern tools. Part of that needs to be how do you do so prudently and safely? I think that a lot of the funding models are very difficult for funding sustainable technology programs."*
>
> —Rights Group 1

portfolios. Some of these come from the same threat actors we have tracked in this report.

It would move the conversation forward to have a better understanding of the scale of compromises that have already taken place among major funding organizations. Funders should also consider the responsibilities they have to their grantees and partners concerning disclosure of breaches. Grantmakers are not the first sector to deal with this issue, and the emerging consensus is that balanced transparency (to the extent that is practical and does not further compromise confidentiality and security) can be positive.

The bottom line

Grantmakers occupy a critical position and are among the few who can directly put their resources where their concerns are, and help effectuate change at scale. If funders care about the continued success of their programming and the security of their grantees, attention to digital security needs to occupy a proportionate part of their activities. Funders need to make sure they are also secure: they should not be the party adding risk to their relationships with grantees. They should be part of the solution.

FOR COMPANIES

The technology sector benefits tremendously from the association of their tools with positive social change. In the past several years, we have observed that social media and technology companies are often publicly thrilled when their products turn up in use during periods of dramatic social change. We think this enthusiasm reflects a genuine, ambitious idealism about the transformative power of new technology. Yet we also observe that many of these platforms, especially when used without adequate precaution, serve as efficient vectors of attack against CSOs.

If the technology sector indeed has a strong commitment to a possible role for its projects in the civic realm, which includes civil society, this will entail some responsibilities, including:

- Understanding how CSOs make use of your services
- Tracking and mitigating specific threats to CSOs using your platforms
- Being transparent about appropriate uses and potential risks of a platform in specific contexts.

Know your users

CSOs make creative and sometimes unexpected use of social media and communications platforms for all manner of strategic and tactical purposes. Yet these uses create new vulnerabilities and avenues of attack. As many of these attacks are highly targeted, they can fall below the radar of security teams who are looking for attacks with effects on a very large user base. Targeted digital threats often constitute an exceptionally high risk against a small number of users. Tracking these attacks, in our experience, often requires a close understanding of the practices and operations of targeted groups, and some communication with the groups about the threats they face.

We are aware of some companies that, often discretely, are developing the capacity to engage with civil society while managing reputational risks and allocation of staff time. In practice, this sometimes means careful collaboration with intermediary and partner organizations, including funders, policy groups, academics, and others that can help technology companies navigate the challenges.

Think creatively about flexible licensing

We have observed that CSOs are often stuck in least-common-denominator models of organizational security because they are unable to standardize software products and devices across their organization. This lack of software adaptability makes it hard to create a security policy or consistent set of security guidelines within an organization. Moreover, many of the tools and software used by CSOs are often counterfeit or expired. This situation leads to a practice of avoiding software and operating system updates. In still other cases, CSOs use free versions of tools and packages that offer lower levels of security than for-purchase versions.

A number of companies have shown that it is possible to provide reduced cost and free licenses to bonafide CSOs. Many companies have the resources to make this kind of commitment. However, we note that for these programs to be effective, the low-cost or free versions must not skimp on security features.

> *"The market is offering tools with decreasing cost that simplify the use of targeted attacks against anyone in civil society by informal actors. So I think that it will be very, very important for us beyond the human rights movement to understand the role that the private sector is playing here, because we're in a moment we could impact. But in general I think that's one of the things that is missed the most."*
>
> —Rights Group 2

Time for pro bono?

We believe digital security is a right, regardless of ability to pay for services. We are in desperate need of ways to address CSOs digital insecurity that take into account the limits of CSOs' and grantmakers' resources. CSOs need sustained, expert technical services and consultation tailored to their specific needs and risk profiles. This kind of service, typically provided in-house or by consulting companies, is considered essential by the corporate and public sector, but the cost of quality service is out of reach for the vast majority of CSOs, with the exception of a few, very well-resourced ones based in North America and Europe.

The tech sector has substantial resources that can be tapped. As the sector professionalizes, as with law and medicine, it is time to examine pro bono models of support for digital security assistance to civil society. There are many people in the sector with extensive experience providing this kind of services and assistance, whether in incident response, the development of security policies, or assistance managing security services. Many developers and technologists would likely find it rewarding and meaningful to contribute time and resources to CSOs with the support and approval of their employers.

As a first step, we encourage technology companies to consult staff and management to ascertain interest in pro bono programs, and begin thinking through the other benefits, but also reputational risks and how they might be mitigated.

FOR RESEARCH

We see the *Communities @ Risk* report as the culmination of only the first phase of our formal study, and in subsequent phases of the project we intend to make adjustments to the research design and the scope of the project. In particular, we would like to expand the number and type of participants enrolled in the study. As mentioned above, the first phase of the research involved eight groups out of 10 whose activities and/or orientation were "China" or "Tibet facing," and only two that were globally oriented. Moving forward, we will look to enrol groups from other regions and countries (e.g., Latin America, sub-Saharan Africa, Southeast Asia, and the Middle East), and include groups that are working in sectors not presently covered by our study groups, such as the environment, LGBT rights, or extractive resources. We also intend to undertake a dedicated research effort regarding journalists at risk—a particularly salient sub-category of civil society for which there are preliminary indications of targeting and uneven security practices.

One of the challenges of researching targeted digital threats is that evidence of the harms around such threats are often disparate, unconnected, and/or incomplete. For example, a member of a CSO may be detained because of a compromised device, but have no idea of the nature or even existence of the compromise in the first place. Likewise, a researcher in the computer security industry may have evidence of victims of a targeted computer espionage campaign, but have no channel to notify the victims of the breach before it is too late. Meanwhile, a public interest group may want to launch litigation around a specific case of targeted digital threats, but lack data preserved in a fashion that is useful as evidence in a court of law. Moving forward, we hope to encourage a better means of collecting, archiving, and organizing data among relevant stakeholders so that stories of harm are better documented and understood.

We also hope to improve upon the methods and tools available to research groups outside of Citizen Lab and among the wider communities of which we are a part. There were tools developed by other researchers, including affiliates and colleagues of Citizen Lab, from which we benefited tremendously, including Cuckoo Sandbox and Viper, a binary management and analysis framework for security researchers developed by Citizen Lab collaborator Claudio Guarnieri. Unfortunately, many other tools, methods, threat intelligence platforms, and repositories of data used by security researchers are proprietary and / or prohibitively expensive for researchers to employ. In the next phase of our research, we intend to explore and further develop tools, methods, and platforms for open data sharing.

Finally, throughout the course of our study we sought to publish timely reports on our website while also working on a separate track to publish in peer-reviewed journals and conferences. Our mixed methods approach to targeted digital threats fits uneasily into any one academic discipline and finding venues for publication of formal research of this sort is a challenge. However, we successfully published several papers in major academic conferences, held two annual summer institutes on mixed methods research on information controls, and helped develop a new fellowship program on interdisciplinary research, in which Citizen Lab will become one of several host organizations supporting a community of researchers working in this area. We hope to build upon this success moving forward, and contribute to a growing community of academic research around targeted digital threats. In the long run, a robust community of researchers producing rigorous, evidence-based, and impartial research on digital risks will offer a powerful form of support for civil society networks.

FOR GOVERNMENTS

While we recognize that governments have complex agendas and competing interests when it comes to cybersecurity, long-term solutions to targeted digital threats will require government involvement. Current debates concerning government reform naturally have emphasized mass surveillance, but the question of targeted cyber espionage and digital attacks against civil society actors also merits further consideration. This is not an issue on which governments should get a free pass by simply asserting espionage is an established feature of state intelligence; the ability to distinguish among targets in cyberspace and treat legitimate civil society actors as off-limits for such conduct is essential. States that support the right to privacy and freedom of expression online should take steps to raise the profile of targeted digital threats against civil society in their domestic policy and diplomacy, *treating the matter as of equal priority to their defense of the private sector*. Moreover, governments should take urgent action to reign in—and avoid driving the growth of—the increasingly dangerous and largely unregulated market for commercial spyware. In all of these efforts, it will be essential for government to engage with civil society in meaningful dialogue to inform appropriate solutions.