TACTICAL TECHNOLOGY COLLECTIVE

Digital Security in Context:

Learning how human rights defenders adopt digital security practices

Becky Kazansky

Table of Contents

Acknowledgements

We wish to express our gratitude to the human rights defenders who made this study possible, and whom we wish to support through this work. This research also owes much to the digital security trainers and facilitators whose work has inspired this project. Thank you to the participants of the project's November 2013 kick-off workshop in Germany, and the April 2015 workshop in the Philippines, in which the research team discussed findings with regional partners and local networks; to Ben Wagner of the Centre for Internet and Human Rights, and Alice Nah and the Centre for Applied Human Rights at the University of York who gave us their support, collaboration and guidance at different stages of this project; to Becky Faith who worked with the research team in editing the report and supporting the final stages of writing, analysis and presentation; and to the colleagues at Tactical Tech who have contributed to this work with their feedback, guidance and support at various stages: Alistair Alexander, Dan O' Clunaigh, Maya Ganesh, Alexandra Hache, Stephanie Hankey, Fieke Jansen, Hannah Smith, Bobby Soriano, Marek Tuszynski, Chris Walker and Carol Waters.

Becky Kazansky[*]

[*]Becky Kazansky was the Lead Programme Researcher at Tactical Tech and conducted this study from November 2013 - August 2015. As of September 2015, she is a PhD researcher at the University of Amsterdam at b.kazansky@uva.nl. Inquiries related to the current study should be addressed to Maya Ganesh at maya@tacticaltech.org. Technical details in this paper relating to academic and practitioner research studies, status of open source software tool development, malware and social media platform policies have not been updated since November 8, 2015.

Setting the Scene

1 Background and Rationale

Tactical Tech has been supporting the effective use of technology in advocacy among communities of human rights defenders (HRDs) around the world since 2003. As information and communication technologies (ICTs) have become central to organising, campaigning and day-to-day communication in human rights work, the human rights sector has concurrently observed the increased use of technologies to monitor communications and harass HRDs. Stories of monitoring and intrusion facilitated through digital surveillance technologies have been relayed to Tactical Tech throughout different engagements over the course of these years. These stories confirm that HRDs are targeted because they challenge powerful interests, expose injustices and make rights claims in repressive environments. These stories also provide evidence that surveillance, intrusion and online harassment can end in physical harm or arbitrary and unjust imprisonment.

Since 2005 Tactical Tech has observed the ways in which digital security practices help to protect HRDs capacity for organising, campaigning and conducting various rights-based work in repressive environments. In response to evolving needs around digital security, Tactical Tech has fostered digital security trainings and produced resources and guides such as Security in-a-Box[1] and Myshadow[2], Tactical Tech has also documented the importance of digital security practices in previous research and writing.[3]

Despite the observed and documented importance of digital security practices in helping HRDs carry out their work, creating effective digital security strategies within human rights work presents a serious challenge. HRDs contend with an ever-shifting landscape of threats and technologies and express a need for support in developing strategies for the secure use of information and communications technologies in their work. In 2013, Tactical Tech underlined the need to for a deeper understanding of human and behavioural factors in privacy and digital security in order to better address the challenges of digital security in human rights work. Executive Director Stephanie Hankey and Associate Daniel O' Clunaigh set out several considerations for rethinking security approaches. In a position paper

[1] Securityinabox.org

[2] Myshadow.org

[3] Notley, T., & Hankey, S. (2013). Human rights defenders and the right to digital privacy and security. In Lannon, J.M., Halpin, E. F. & Hick, S. (Eds.), *Human Rights and Information Communication Technologies: Trends and Consequences of Use.* Hershey, PA: IGI Global

published in the Journal of Human Rights Practice in 2013, they proposed:[4]

> *First, we should not solely rely on fighting problems created by technology with more technology; second, we need to understand the role of behaviour when building capacity; and third, we need to move beyond a tech-centric approach to capacity building, embedding these issues within broader approaches to security.*

These proposed considerations provided the original impetus for the Digital Security in Context study. For further guidance in scoping a set of questions for this study, 15 developers, trainers and intermediaries were brought together in a two-day workshop which took place in November 2013. The workshop allowed the team to collaboratively identify barriers to effective capacity building on digital security. Participants noted the common perception among HRDs that adopting practices taught in trainings would have a negative impact on their work due to the amount of time and energy required to learn to use challenging digital security tools and integrate their use within their workflows. In explaining the difficulties of effective digital security practices, a tool developer described an "overwhelming request of actions" made upon ICT users and HRDs specifically, due to a landscape of constantly shifting digital threats, tools that were inaccessible or difficult to learn to use without assistance, and a lack of sustained support in building digital security practices.

13 semi-structured pilot interviews were concurrently conducted with geographically diverse HRDs.[5] The HRDs offered their perspectives on challenges to the protection of their data and of their overall safety, sharing stories of phone-tapping, targeted malware and hacking by state and non-state actors. The barriers identified by HRDs in pilot interviews largely intersected with those identified by trainers, developers and intermediaries in earlier discussions. HRDs also described an ever-shifting landscape of threats and technologies impacting their work, and confirmed the difficulty of integrating digital security practices within their workflows. One HRD described security as a 'trade-off' because many digital security tools are challenging to implement and lamented the difficulty of getting others to adopt a digital security practice: "They still don't get the trade-off that if you want convenience you can't be safe".[6] Another HRD noted that "there is no perfect security. Things keep on changing. Effective activism is more important"[7].

Despite the common perception that digital security presented a barrier or required a difficult trade-off, it was also common to hear that security is essential to human rights work: "digital security should be integral to my work. If it's not secure then the whole work is not secure".[8] HRDs told the team that integrating digital security practices into their work would aid broader strategic aims. In seeing digital security as an integrated element of their work, HRDs pointed to the importance of framing digital security as an ongoing practice: "It should be a way of life. You don't think so, but it becomes natural. You force yourself at first and then it becomes something you do."[9]

Yet many HRDs interviewed felt ambivalent about how much individual agency they have in issues of security. HRDs explained that in their networks, coalitions and organisations, rarely

[4] Hankey, S. and Ó Clunaigh, D. (2013). Rethinking Risk and Security of Human Rights Defenders in the Digital Age. *Journal of Human Rights Practice*. Vol 5, Issue 3, pp. 535-547.

[5] Interviews took place at a conference and through VOIP

[6] Field Work Interview with Anonymous HRD Site B, #1, 2014

[7] Pilot group discussion, 2013

[8] Pilot interview #2, 2013

[9] Pilot interview #2, 2013

are there dedicated resources available for security infrastructure or training. Instead, responsibilities are shared between colleagues with varied priorities and a limited budget for outside help. Thus, advocacy for security competes for time and attention amidst the many already pressing concerns of day-to-day work. HRDs felt that because "our work is not just security", it was reasonable to recognise a need for outside help in tackling difficult issues; "someone in the background advising me what to do."[10] Yet even when HRDs did find the resources to recruit outside support, they found it difficult to find trusted individuals in their region.

The concerns shared in pilot interviews highlight that HRDs contend with constraints which can make worrying about the privacy and security of ICT use feel like an overwhelming additional burden. HRDs emphasised the fact that capacity-building efforts such as trainings are empowering: "It's partially my responsibility, but we should be empowered and provided with training."[11] Given the importance of capacity building efforts such as digital security trainings and an ever-shifting landscape of threats and technologies, Tactical Tech sought to understand how digital security trainings can best contribute to the efficacy of digital security practices in human rights work. The Digital Security in Context study focuses on answering the following two questions:

- How do human rights defenders adapt their digital security and privacy practices to shifting sociotechnical contexts?
- How well are individual and collective needs met through current models of digital security capacity building efforts?

The study draws on findings from interviews, discussion, and trainings, with a total of 60 people involved in human rights work. Research took place over 18 months. Section one of the report explains the role of digital security strategies in human rights work and locates the study in relation to recent work done by scholars and practitioners to understand and improve upon privacy and digital security practices. Section two outlines the digital security concerns of the three groups which the Tactical Tech team worked with, and investigates how digital security strategies are formed in response to these concerns. Section three outlines specific challenges around digital security tools and practices in relation to the three digital security trainings taking place over this study. Section four examines the factors which help digital security practices to be sustained over time. The report concludes with a set of recommendations and appendix.

The team readily acknowledges that many important issues have remained outside of the scope of this report. Analyses regarding specific tools may become outdated due to the rapid pace of technological change. Rather than attempting to provide definitive descriptions or conclusive findings, the report thus seeks to provide a contextual 'snapshot' and analysis of some of the more salient concerns around digital security in human rights work at the time of writing. The report offers a set of learnings and recommendations which may be useful to other practitioners, and points to future lines of inquiry for others to expand upon.

[10] Pilot Interview with Anonymous HRD#1

[11] Pilot Interview with Anonymous HRD#1

2 Situating Tactical Tech's Practitioner Research

This section of the report locates the Digital Security in Context study in relation to recent work done by practitioners and scholars to understand and improve upon privacy and digital security practices, examining literature from several academic disciplines and from within the human rights sector.

Defining digital security

In this report, the term 'digital security' refers to a set of practices dealing with the confidentiality, integrity and availability of information; practices which help human rights defenders (HRDs) meet their goals and which fit within broader security strategies in human rights work. The term 'digital security tools' in this report is shorthand for the tools present in Tactical Tech's Security in-a-Box[12], and which were introduced by Tactical Tech in the trainings conducted in the study. Tools included are sometimes categorised as 'privacy enhancing technologies', 'circumvention technologies' and elements of 'digital hygiene.' The terms 'privacy' and 'security' may be seen as complementary throughout this report, because in many cases, digital security tools and practices aid users in managing confidentiality and anonymity.

However, 'privacy' is a broad term. Some digital security practices described in this study address elements that might not fall under the umbrella of 'privacy', depending on the definition of the term and the specific needs of an ICT user. For example, the encryption found in individual 'cryptographic access control tools'[13] such as PGP might preserve the confidentiality of message content between two people, but leave various forms of metadata exposed, meaning that people using these tools' together with commercial online platforms and services such as Facebook or Google will likely continue to provide trails of identifying information through their online behaviours. This is not simply a shortcoming of the tool: the information included in PGP 'keys' can help users confirm the authenticity of those seeking to communicate with them, through a cryptographic 'web of trust'.[14] If an ICT user's priority is anonymity, they will have to carefully consider how to maintain it when using a tool such as PGP.

In this study, digital security is understood in relation to people involved in human rights activism, advocacy, and other forms of work that support human rights. Individuals and groups doing this kind of work are referred to as human rights defenders (HRDs) within the guidelines set out in the United Nations in the Declaration of Human Rights Defenders in 1998.[15] Despite the shorthand use of the terms privacy and digital security, privacy and digital security constructs arise amidst contextual considerations and are articulated through relationships between colleagues, friends and peers, as will be demonstrated in this report. For

[12] https://securityinabox.org/en

[13] Balsa, E., Brandimarte, L., Acquisti, A., Diaz, C. & Gürses S. (2014) Spiny CACTOS: OSN Users Attitudes and Perceptions Towards Cryptographic Access Control Tools. San Diego, Internet Society. USEC '14, 23. https://www.internetsociety.org/sites/default/files/02_2-paper.pdf

[14] Definition for 'web of trust' https://en.wikipedia.org/wiki/Web_of_trust

[15] United Nations Declaration of Human Rights Defenders. 1998: http://www.ohchr.org/EN/Issues/SRHRDefenders/Pages/Declaration.aspx and see 'Who is a defender?' for more specifics on the kinds of activism that qualify http://www.ohchr.org/EN/Issues/SRHRDefenders/Pages/Defender.aspx. We include the group of Environmental Rights Defenders we worked with as HRDs. Margaret Sekaggya, the UN's former Special Rapporteur on human rights defenders, has referred to HRDs working on environmental issues as environmental human rights defenders within the broader definition of 'human rights defender': http://www.ohchr.org/EN/NewsEvents/Pages/EnvironmentalHumanRightsDefenders.aspx

as long as sociotechnical contexts and priorities continue to evolve, interpreting what digital security means remains a generative endeavour. In this sense, digital security is best understood *in context*.

Though the focus of this study is on security as it applies to digitally stored and transmitted data, Tactical Tech's experience shows that digital security is connected directly to physical safety and psychosocial well-being, as expanded upon in the 'holistic security' methodology currently in development together with the Center for Victims of Torture and Front Line Defenders. The goal of this methodology is the improved integration of physical and psychosocial elements of security strategies with digital or informational elements.[16] Security is defined in the forthcoming materials on holistic security as 'well-being in action'.

Elements of the holistic framework for security have previously been set out in texts such as the 'Integrated Security Guide'[17] by Jane Barry with Kvinna Till Kvinna, and in the 'New Protection Manual for Human Rights Defenders', by Protection International with Enrique Eguren Fernández and Marie Caraj[18]. These two texts have served as important influences for Tactical Tech's recent work in developing new 'holistic' security methodologies. The 'New Protection Manual' by Protection International and the forthcoming 'Holistic Security' materials frame digital security practices within larger security strategies, identifying three main categories: acceptance, protection, and deterrence. This excerpt from the forthcoming 'Holistic Security Strategy Manual for Human Rights Defenders' explains their breakdown[19]:

- **Acceptance strategies:** acceptance strategies might include running campaigns to build public support for your work or that of human rights defenders generally, or carrying out advocacy to develop positive relationships with local, state, or international authorities, which correspond to their obligations to respect human rights defenders. An acceptance strategy involves engaging with all actors – including allies, adversaries and neutral parties – in order to foster acceptance and ultimately support of your human rights activities in society.

- **Protection strategies:** a protection or self-defence strategy emphasises learning new methods and implementing new practices, or leveraging the strength of your allies to protect yourself and cover the gaps in your existing practices. Examples of practices that fall into this category might include implementing the use of email encryption or stress management practices within the organisation or receiving protective accompaniment or human rights observation during your activities.

- **Deterrence strategies:** a deterrence strategy focuses on raising the costs for your adversaries of carrying out attacks against you or your work.

The reader will note that encryption practices are identified as an element of protection strategies; however, privacy and digital security concerns also surface prominently within

[16] For more information on the holistic security methodology see: https://tacticaltech.org/holistic-security and the forthcoming guide: the Holistic Security Strategy Manual for Human Rights Defenders

[17] Barry, J. with Kvinna Till Kvinna. (2011). Integrated Security: the Manual. http://www.integratedsecuritymanual.org/

[18] The new protection manual for Human Rights Defenders (2009) Protection International Research and Training Unit Research and text by Enrique Eguren Fernández and Marie Caraj Protection International *http://protectioninternational.org/publication/new-protection-manual-for-human-rights-defenders-3rd-edition/*

[19] More about these strategies can also be found in the 'New Protection Manual for Human Rights Defenders' by Protection International: http://protectioninternational.org/publication/new-protection-manual-for-human-rights-defenders-3rd-edition/

acceptance strategies, as may be seen in a campaign like the one mentioned in the above categorisation. Working on campaigns almost invariably includes the use of platforms and tools with little built-in protection for users. Campaigners may chose to incorporate practices to preserve the confidentiality necessary to their planning process, or to protect the identity of certain vulnerable individuals. Additionally, many digital security strategies may also be seen as deterrence strategies, in that the use of encryption can be framed as a way to make carrying out attacks against HRDs more expensive for an adversary. This underlines the fact that digital security considerations may come into play throughout all three kinds of strategies listed above.

Aside from the texts mentioned above, peer organisations have published a number of theoretical and practical resources seeking to distil the complex topic of digital security into actionable frameworks for educators, trainers, technologists and HRDs. Some organisations have focused more on physical or psychosocial security in their practical support and materials on security. For example, the 'Workbook on Security' by Front Line Defenders,[20] and the older 'Rukus Society: Security Culture for Activists' guides[21] examine security as a whole, highlighting certain digital elements within. Other organisations have focused more on IT and digital security. The Electronic Frontier Foundation's Surveillance Self-Defense online resource[22] focuses on fostering 'safer online communications,' while Tactical Tech and Front Line Defenders' Security in-a-box has traditionally focused on digital elements of security, with efforts made to include some physical security considerations.[23]

3 Configuring the User: Perspectives from Academic Research

In research on the security of information systems, scholars have tried to explain why security persists as a problem in the design and use of information systems despite the implementation of policies, protocols, and tools to ensure positive security outcomes. This line of questioning has resulted in an increased focus in computer science and human computer interaction (HCI) on understanding the 'human factors' of security. In the 1999 study 'Users are Not the Enemy'[24] authors Anne Adams and Angela Sasse counter a common, implicit assumption that ICT users work at cross-purposes with engineers and developers. Adams and Sasse highlight the many difficulties users face in working to follow rigid communication protocols and policies. The 2005 paper 'Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0' by Alma Whitten and J.D. Tygar'[25]notes that "user errors cause or contribute to most computer security failures, yet user interfaces for security still tend to be clumsy, confusing, or non-existent". Whitten and Tygar propose a definition of 'usability' with respect to digital security tools:

> *Usability necessarily has different meanings in different contexts. For some, efficiency may be a priority, for others, learnability, for still others, flexibility. In a security context, our priorities must be whatever is needed in order for the security to be used effectively.*

[20] Front Line Defenders (2011) Workbook on Security: Practical Steps For Human Rights Defenders At Risk: https://www.frontlinedefenders.org/files/workbook_eng.pdf

[21] Rukus Society, with Bell, J. & Spalding, D. (Publishing date unknown) Security Culture for Activists http://www.ruckus.org/downloads/RuckusSecurityCultureForActivists.pdf

[22] Electronic Frontier Foundation. Surveillance Self Defence https://ssd.eff.org/

[23] Tactical Technology Collective. Security in-a-Box: tools and tactics for digital security https://securityinabox.org/en

[24] Adams, A. & Sasse, M. A. (1999). *Users are not the enemy*. Commun. ACM 42, 12, 40–46

[25] Whitten, A. & Tygar, J.D. (2005) Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0 679-702

In a 2014 article by entitled 'Why Doesn't Jane Protect Her Privacy?'[26] authors Karen Renaud, Melanie Volkamer and Arne Renkema-Padmos offer a different interpretation of 'usability,' arguing that human factors which affect security outcomes arise outside the confines of the interface, from "incomplete threat models, misaligned incentives, and a general absence of understanding of the…architecture".

Despite a greater emphasis on human factors and the growth of human computer interaction scholarship within computer science, some literature on human factors continues to frame security problems as a function of users' 'human error' rather than designer bias or rigid systems. Non-expert lay knowledge is presented as the main barrier to good security. Computer scientists Jim Blythe and L. Jean Camp, in their work on understanding 'cognitive bias' and 'mental models' – different ways people understand systems – theorise that security problems would finally be solved if computer scientists were to manage to shift 'lay' mental models to become more like 'expert' mental models.[27] In blaming security problems on the incompatibility of experts and lay users, the computer scientists continue a tendency to 'configure the user'[28]; to specify a narrow domain of acceptable user behaviours and to subsequently blame users for failing to fit within it. This same tendency might be witnessed in the common saying that people are the 'weakest link in the security chain'.[29]

The tendency to 'configure the user' also comes up in an article entitled 'Stories as Informal Lessons about Security', by Emilee Rader, Rick Walsh and Brandon Brooks, who, through a qualitative approach using semi-structured interviews, examine the way awareness of digital security threats and security incidents spreads through stories among 'non-expert 'users of 'home computers'.[30] The findings in the study are rich in that the researchers are able to highlight the social dimensions of responses to digital security threats, yet their conclusions prioritise the problems of 'folk models' among 'home computer' users instead of investigating how models might be indicative of broader social processes. In doing so, they risk ignoring important structural issues contributing to problems of security.

A 2014 study on mental models entitled '"My Data Just Goes Everywhere": User Mental Models of the Internet and Implications for Privacy and Security'[31] problematises some of the assumptions implicit in previous work on mental models. Authors Ruogu Kang and Laura Dabbish, et al note that while mental models do vary between experts and non-experts – in line with previous research – they "did not find a direct relationship between people's technical background and the actions they took to control their privacy or increase their security online",[32] concluding that technical education alone was not predictive of the extent to which people were able to respond to security concerns. Instead, they found that the kind of responses they saw were owing to personal experiences with security incidents.

---

[26] Renaud, K., Volkamer, M. & Renkema-Padmos, A.(2014) Why Doesn't Jane Protect Her Privacy? In De Cristofaro, E. & Murdoch, S.J. (Eds) *Privacy Enhancing Technologies* presented at the 14th International Symposium, PETS 2014, Amsterdam, The Netherlands, July 16-18, 2014. Proceedings

[27] Blythe, J. & Camp, L.J. (2012). "Implementing Mental Models", Semantic Computing and Security, an IEEE Symposium on Security and Privacy (SP) Workshop in San Francisco, CA, May 24, 2012.

[28] Woolgar, S. (1990) Configuring the user: the case of usability trials *The Sociological Review* Vol. 38, Issue S1, pp 58–99

[29] One such instance of weakest link framing can be found here: https://www.techdirt.com/articles/20120810/18401819991/humans-still-weakest-link-security-chain.shtml

[30] Rader, E. Wash, R. Brooks, (2012). Stories as Informal Lessons about Security. Brandon Symposium on Usable Privacy and Security (SOUPS),Washington DC, US July 11-13, 2012.

[31] Kang, R., Dabbish, L., Fruchter, N. & Kiesler, S. (2015) "My Data Just Goes Everywhere": User Mental Models of the Internet and Implications for Privacy and Security. Paper presented at the Symposium on Usable Privacy and Security, (SOUPS) Ottowa, Canada, 22-24 July.

[32] *Ibid.*

To gain a better understanding of the use of ICTs as a practice situated within broader processes, it is useful to examine 'sociotechnical' approaches to the study of information systems. Broadly, sociotechnical approaches see the social and technical dimensions of ICT use as being equally important, interdependent, or mutually constitutive. The term 'sociotechnical' originates from research and design work done at the Tavistock Institute of Human Relations in London in the 1940s, which was aimed at improving working conditions by prioritising the 'human' factors of systems. In their survey, entitled 'Sociotechnical approaches to the study of Information Systems',[33] Steve Sawyer and Mohammad Hossein Jarrahi note that the school's interventionist and activist approach – situated between research and design – influenced later developments in Scandinavian participatory design and design research methods used to conduct 'action research' on the use of information systems.

Sociotechnical approaches counteract 'technological determinism' – the notion that technologies have predetermined impacts upon people.[34] Some lineages are more 'socially constructivist' – such as Social Construction of Technology theories (SCOT), which posit that the 'construction' of technologies ultimately takes shape through the way people adapt them to their contexts. Conversely, sociologist Anthony Giddens' Structuration Theory and the information systems specific Adaptive Structuration Theory[35] hypothesise that technological structures and people mutually constitute one another through ongoing interaction. Scholars use the notion of this ongoing interaction between 'structures' and 'people' or, alternately, 'human' and 'non-human' actors[36] to explain how practices form. Susan Leigh Star points to some of the kinds of structural constraints that affect practices involving the use of ICTs:

> *The computers may work fine, but the electricity is dirty or lacking. Old floppy disks do not fit new drives, and new disks are expensive. Local phone calls are not always free. New browsers are faster, but more memory hungry.*[37]

Wanda Orlikowski writes that practices are "...recurrent, materially bounded and situated action engaged in by members of a community."[38] Similarly, HCI scholar Philip E. Agre writes that practices are "the ensemble of embodied routines that a particular community of people has evolved for doing particular things in a particular place."[39] Positioning her work closely to Social Shaping of Technology theories, researcher Dana Boyd uses ethnographic methods to better understand how information and privacy practices among teens take shape and evolve over time. In her PhD dissertation 'Taken Out of Context: American Teen Sociality in Networked Publics', [40]Boyd shows that the information practices of teens change as technologies and group priorities evolve. Boyd finds that contrary to the popular opinion that teenagers do not care much about privacy in their practices online, teens negotiate a complex 'set of dynamics'

---

[33] Sawyer, S. and Jarrahi, M. H. (2013). Sociotechnical approaches to the study of Information Systems in Munindar P. Singh (Ed.) *Handbook of Computing*, Chapman and Hall/CRC

[34] A definition of 'technological determinism': https://en.wikipedia.org/wiki/Technological_determinism

[35] Desanctis, G. & Poole, M.P. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science.* Vol. 5, (2) pp. 121-147.

[36] As most prominently featured in the Science and Technology Studies scholarship of Bruno Latour through his development of 'actor-network theory'

[37] Leigh Star, S. (1999). The Ethnography of Infrastructure *American Behavioural Scientist* November. 43: pp. 377-391.

[38] Orlikowski, W.J. (2002). Knowing in practice: Enacting a collective capability in distributed organizing. *Organization Science* Vol.13 (3) pp. 249-273. excerpted from Sawyer, S & Jarrahi, M. H. (2013). Sociotechnical approaches to the study of Information Systems in Munindar P. Singh (Ed.) *Handbook of Computing*, Chapman and Hall/CRC

[39] Agre, P.E. (2001) Changing Places: Contexts of Awareness in Computing University of California, Los Angeles *Human Computer Interaction*, Vol. 16, pp. 177–192 Lawrence Erlbaum Associates, Inc.

[40] Boyd, D. (2008). *Taken Out of Context: American Teen Sociality in Networked Publics* PhD dissertation. University of California, Berkeley. pp. 3.

including invisible audiences, collapsed contexts and the blurring of the public and private due to larger shifting 'structures forces' enacted upon them. Boyd writes:

> *A technology's value is shaped by its social construction—how designers create it and how people use it, interpret it, and reconfigure it. It is not an outcome of the technology alone or its potential.*

Helen Nissenbaum's work also grapples with the notion of changing social contexts in relation to privacy practices. Nissenbaum defines 'social context' as the way goals and aims are shaped through social considerations and processes.[41] She focuses on understanding practices around privacy and information exposure through considerations specific to particular 'information flows.' Nissenbaum suggests tracking damaging changes to societal privacy norms through people's reactions to unexpected exposures of data, which she explores through the notion of 'contextual integrity'.[42]

In the 2004 study 'Security in the Wild', scholars Paul Dourish and Rebecca E. Grinter et al examine notions of context in relation to the digital security practices of users. The authors study the 'non-use' of digital security tools such as PGP by using qualitative, ethnographic research methods that allowed them to "understand their experience of security as they (users) encounter it".[43] The authors found that users had a number of protective strategies that might not have been classified as security practices within some of the narrow definitions assigned by computer scientists. For example, users they spoke to obscured the content of messages by referring to contextually relevant information without ever explicitly stating the subject of the email. These findings highlight that digital security can't necessarily be defined or quantified through a narrow focus on particular behaviours.

4 Prioritising Context and Practice in Practitioner Work

Though scholars taking a sociotechnical approach to the research of information systems offer many empirically grounded theories applicable to the study of privacy practices, the body of research highlighted in this review does not offer up concrete examples for how to specifically understand digital security practices with respect to human rights work. This review thus highlights a few recent contributions from non-governmental organisations and academic institutions documenting the privacy and security concerns of specific groups of people, in relation to human rights frameworks and practices.

- Human Rights Watch released a report documenting the experiences of individuals targeted by the use of invasive surveillance technologies in Ethiopia, providing an unusually in-depth look at life under the eye of the state.[44]
- Citizen Lab, at the Monk School of Global Affairs at the University of Toronto published findings from four years of research on the digital security vulnerabilities of the Tibetan

[41] Nissenbaum, H., (2015). Respect for context as a benchmark for privacy online: what it is and isn't. In Roessler, B. & Mokrosinska, D. (Eds) *Social Dimensions of Privacy: Interdisciplinary Perspectives* Cambridge University Press

[42] Nissenbaum, H. (2010) *Privacy in Context. Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

[43] Dourish, P. R., Grinter, E., Delgado de la Flor, J. & Joseph, M. (2004). Security in the wild: User strategies for managing security as an everyday, practical problem *Personal and Ubiquitous Computing*, Vol. 8 Issue 6 391-401

[44] Human Rights Watch (2014)."*They Know Everything We Do" Telecom and Internet Surveillance in Ethiopia* https://www.hrw.org/report/2014/03/25/they-know-everything-we-do/telecom-and-internet-surveillance-ethiopia

community, devoting a sizeable section of the report to analysing civil society responses to intrusion and monitoring by the Chinese government.[45] The section on civil society actor responses highlights the importance of digital security trainings and hands-on work to meet the demands of a rapidly shifting landscape of threats.

- Article 19 published a report tying the digital security practices of activists in Iran to their later arrests,[46] providing a section of recommendations for the improvement of practices in relation to the most common forms of threats.
- In 2014, Human Rights Watch released a report on the responses of investigative journalists to surveillance in the United States,[47] and the United Nations Educational, Scientific and Cultural Organisation released a survey of digital security threats and capacity building efforts being done to address journalists needs around the world.[48]
- Groups have also undertaken 'action-orientated' research to provide immediately useful feedback for their practice-based work. In 2013, the Open Technology Institute released a report evaluating current digital literacy programs in the United States, on the basis of the extent to which they prepare 'marginal users' who had only recently gained access to the internet to deal with privacy issues.[49] The study was done in conjunction with a collaborative effort to create a privacy literacy tool.
- The Internews Center for Innovation & Learning, with Engine Room, released a study identifying gaps in current approaches to digital security 'training of trainer' (ToT) programmes in order to inform the efforts of the LevelUp team in building its digital security training resource.[50] The report recommended that the sector develop more resources covering digital security threats and responses and provide more opportunities for mentorship for new trainers. The LevelUp resource proved important in Tactical Tech's research on digital security practices. The project's collection of learning modules developed by a number of digital security trainers in the sector[51] helped the team develop a set of research activities to use in conjunction with the trainings undertaken over the course of the Digital Security in Context study.
- An evaluative report by Amnesty International documents learnings from the development of the Panic Button application, a tool designed to support HRDs in reaching out to trusted peers in critical situations.[52] The report provides a unique glimpse into the development of a tool in relation to the contextual priorities of HRDs. While there is other work evaluating the effectiveness of security and protection programmes, much of it has not been released due to concerns around the sensitivity of

[45] The Citizen Lab, Munk School of Global Affairs, University of Toronto (2014) Communities @ Risk. Targeted Digital Threats Against Civil Society. https://targetedthreats.net/

[46]Article 19 (2015). Country Report: Computer Crimes in Iran- Risky Online Behaviour. https://www.article19.org/resources.php/resource/38039/en/country-report:-computer-crimes-in-iran--risky-online-behaviour

[47]Human Rights Watch (2014) With Liberty to Monitor All https://www.hrw.org/report/2014/07/28/liberty-monitor-all/how-large-scale-us-surveillance-harming-journalism-law-and

[48]Henrichsen, J.R., Betz, M. & Lisosky, J.M. (2014) Building Digital Safety for Journalism: A survey of selected issues. UNESCO Publishing.

[49]Gangadharan, S.P. with the Open Technology Institute. (2013) Joining the Surveillance Society? New Internet Users in an Age of Tracking https://www.newamerica.org/oti/joining-the-surveillance-society/

[50] The Engine Room with Internews Center for Innovation & Learning (2013). Training Digital Security Trainers: A Preliminary Review of Methods, Needs and Challenges https://www.internews.org/sites/default/files/resources/InternewsWPDigitalSecurity_2013-11-29.pdf

[51]These included current or former Tactical Tech trainers.

[52] Amnesty International (2015). Security is Community. Lessons from the Panic Button experience. https://www.amnesty.org/en/documents/act10/2133/2015/en/

the information. The tensions around determining what is safe to release publicly inform the research approach of this study.[53]Extensive discussion on this is found in the Research Approach and Methods section of the report.

5 Practice-based Research Foundations

This review has highlighted the importance of understanding privacy and security within situated practices and specific sociotechnical contexts, and noted the need to account for considerations specific to human rights in understanding digital security concerns. The report now expands upon these considerations in relation to the development of the research approach particular to the Digital Security in Context Study.

The approach taken in this study grounds itself first and foremost in practitioner approaches to capacity building particular to Tactical Tech, which the report now introduces. In workshops and trainings and while co-creating tool-kits and learning materials with communities of HRDs, Tactical Tech has emphasised that capacity builders must engage with privacy and security issues in a contextually appropriate way, meaning that the work must forefront the needs and priorities of specific communities at a particular place and time. Tactical Tech develops strategies and advises on tool choices through a participant-driven process which foregrounds concerns and priorities particular to the context of individuals and groups. Capacity building efforts are guided by a 'context analysis' undertaken together with HRDs to assess the political, social and technological factors affecting rights-based work at a particular time and place.[54]

Tactical Tech's forthcoming Holistic Security Manual for human rights defenders describes the context analysis as a process of "deliberately learning more about our surroundings in order to identify and analyse the threats we face to our well-being in action".[55] This process helps to bridge knowledge of relevant concerns and threats between training facilitators and participants, and helps facilitators create an agenda to address the priorities particular to a group, providing a roadmap for effective digital security strategies. Concerns and priorities raised throughout the process serve as the basis for hands-on work with appropriate digital security tools. [56]

The context analysis can include a number of different activities focusing on facilitating the identification of relevant concerns or 'threats'.[57] In one hands-on exercise that helps surface these concerns, participants are asked to draw an 'actor map' of the most important influences and uses of technology in life and work, including allies, opponents and neutral parties.

[53] The team presented a paper documenting some of these tensions at the hotPETs workshop at the Privacy Enhancing Technologies meeting in July, 2015. https://petsymposium.org/2015/papers/ganesh-activists-hotpets2015.pdf

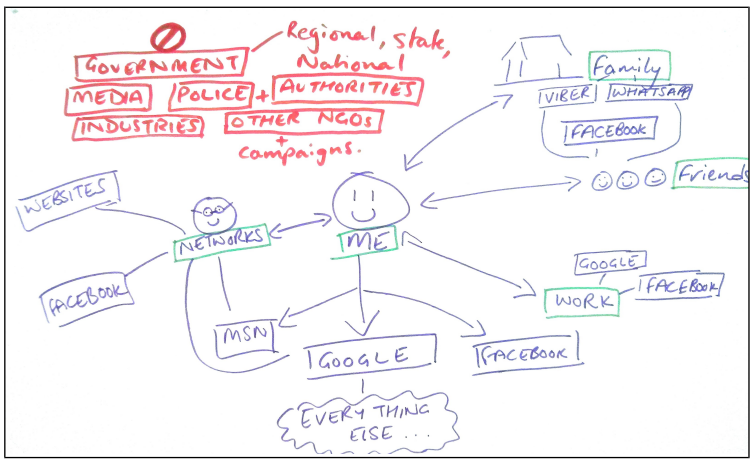
[54] Tactical Tech formerly called its series of community-specific digital security guides 'Security in Context'. The guides have since been renamed the Security in-a-Box Community in Focus Guides, but the methodology and philosophy behind these guides served as a primary inspiration for the title and approach of this study.

[55] For an in-depth guide to the context analysis process, please see the forthcoming Holistic Security Strategy Manual for Human Rights Defenders by the Center for Victims of Torture and Tactical Technology Collective, 2015.

[56] The context analysis is also an important element in the preparation and due diligence that goes into the planning of trainings, such as interviews with potential participants, and other pre-assessment work done in advance of interventions.

[57] In some trainings, a facilitator leads this exercise, eliciting concerns and representing them visually in one map representative of the group. In trainings conducted throughout this project, we had participants each draw their own maps and discuss their differences and similarities. Sometimes it is framed as a 'stakeholder analysis'. For more descriptions see https://www.level-up.cc/

Illustration 1. Fig.1 Actor map



Another activity typically undertaken in Tactical Tech's context analysis is an exercise called 'information mapping'. The information mapping exercise aims to help participants understand the potential consequences of their information handling practices, by thinking about what kinds of sensitive or valuable information they store and share throughout their workflows. Participants are asked to think about all of the locations and devices where they store their information, to brainstorm the ways it may be compromised or lost, and to gauge the relative difficulty of replacing it. These questions are used to help participants parse through an overwhelming number of potential scenarios and concerns. An 'information mapping' scenario is included in the appendix of this report.

Activities found throughout the 'context analysis' specific to this study are reconfigured throughout a number of related practices and activities. Some digital security training organisations refer to various elements within a process of contextualisation as 'tailoring', 'need finding', 'conducting a risk assessment' or 'threat modelling'. 'Threat modelling' itself concerns a varied constellation of practices. Developers working on digital security tools might use threat modelling processes to mitigate against the effects of potential 'attack scenarios',[58]while threat modelling exercises used within digital security trainings by civil society organisations centre more broadly around the day-to-day information handling routines of 'end users.' The Surveillance Self Defence guide, by the Electronic Frontier Foundation, defines threat modelling as 'a way of narrowly thinking about the sorts of protection you want for your data'. The resource guides readers through a series of questions to help them narrow their priorities, asking “What do you want to protect? Who do you want to protect it from? How likely is it that you will need to protect it? How bad are the consequences if you fail? How much trouble are you willing to go through in order to try to prevent those?”[59]

Several of Tactical Tech's Security in-a-Box resources feature a 'risk assessment' grid to aid users in identifying threats, 'capacities' and 'vulnerabilities' in order to assess their preparedness in the face of potential threats.[60]The guide notes that “you can think of your risk as an interplay of the threats you face, your vulnerabilities, and the capacities you have”. The

[58] An example of a threat model to guide the development of a tool can be seen here: https://conorsch.github.io/securedrop/technical_information/threat_model/

[59] The Electronic Frontier Foundation's introduction of the 'threat modelling' concept can be found here: https://ssd.eff.org/en/module/introduction-threat-modeling

[60] An example of a Risk Assessment framework used by Tactical Tech can be seen here https://securityinabox.org/en/eco-rights-africa/security-risk

New Protection Manual by Protection International features a formula, also designed to facilitate a more accurate understanding of relevant threats. It similarly focuses on the identification of assets, capacities and vulnerabilities. The focus of the formula is on calculating a quantified measure of risk, where 'risk = threats x vulnerabilities/capacities'. The Integrated Security Manual by Jane Barry and Kvinna Til Kvinna pinpoints a 'threshold of acceptable risk'[61] to help HRDs gauge what level of danger they deem acceptable to expose themselves to in their work.

In practice, all of these interrelated frameworks involve discussion between trainers/facilitators and participants in order to highlight the most important concerns and priorities of participants and to identify patterns or indicators of likely threats to human rights work. Though many of the activities conducted in this study also aimed at creating a more accurate picture of potential dangers, this report does not frequently refer to 'risk assessments' and 'threat modelling' because this terminology was not explicitly used within the three trainings conducted in the study. Out of concern for the level of abstraction that can figure into discussions focused on risks and threats, the focus in these trainings was instead placed on looking at and understanding concerns and priorities through the kinds of 'context analysis' activities described above. These are described in more detail in the Appendix 1 of this report.

[61] The 'Threshold of Risk' can be found in this section of the Integrated Security Manual: http://www.integratedsecuritymanual.org/exercise/defining-the-threshold-of-acceptable-risk

Research Approach and Methods

The Digital Security in Context study was undertaken in order to aid Tactical Tech's practitioner work. Tactical Tech's digital security trainings aim to facilitate the effective use of communications technologies in a particular context through a participatory approach. Participants are treated as experts, playing an equal role in the creation of the agenda and implementation of an intervention. This study was structured to fit within Tactical Tech's overall approach to capacity building, taking further inspiration and guidance from action-oriented[62] and participatory research methodologies.[63] The study used qualitative methods throughout.

The Digital Security in Context study took place over 18 months. Trainings and research were conducted by a team consisting of a lead trainer and a researcher-trainer, with the additional participation of several colleagues in the planning and implementation of the overall project. The choice to use an embedded researcher-trainer was taken in order to ensure a trust on the part of the participants and to allow the researcher a depth of insight which can only be achieved from the position of a trainer. The research was conducted in two phases: a formative pilot and a field work period. Over the pilot and field work periods, the team engaged with a total of 60 interviewees, training, and workshop participants. For a period of five months during the pilot and planning phase, the team surveyed literature on human factors in security and practitioner research on surveillance, privacy, and digital security. Some of the literature cited in this survey is discussed in the section above.

The project was designed to answer the following questions:

- How do human rights defenders adapt their digital security and privacy practices within shifting sociotechnical contexts?
- How well are individual and collective needs met through current models of digital security capacity building efforts?

1 Research Implementation

Pilot

In the pilot period, the researcher-trainer conducted interviews with 13 geographically diverse

[62] Baskerville, R.L. (1999). Investigating information systems with action research *Communications of AIS* Vol. 2, Article 19

[63] Genat, B. (2009). Building emergent situated knowledges in participatory action research *Action Research* 7: 101

HRDs who had previously participated in digital security trainings with Tactical Tech. These interviews were conducted around several conferences and through VOIP. A workshop was held by a larger team of Tactical Tech colleagues in Germany, together with 15 geographically diverse participants from across the sector working to build capacity on human rights and digital security. The discussions and interviews from the pilot period helped the project team refine its research questions.[64]

Field Work

In the second stage of the research work, field work was conducted in three different countries with three different groups; an environmental rights organisation, an ICT4D and human rights network and a women's and LGBTQI rights network. After conducting due diligence and setting up agreements with three groups who were interested in participating, the team planned semi-structured interviews, two five-day trainings, and one two-day training with participants from these groups. A total of 32 participants was engaged in trainings across the three groups.

After the three training engagements were complete, the researcher-trainer distributed a short email survey inquiring into which material participants found most and least useful from the trainings, and which digital security practices they continued to follow. The researcher-trainer then conducted follow-up discussions with point people from the trainings after interims of two and four months, in order to discuss answers from the survey, main takeaways from the training experience, and to plan potential follow-up opportunities. Follow-up engagements through workshops and training of trainer opportunities were planned with several participants. Thus far, five of the 32 total participants attended further trainings and gatherings, and further opportunities have been discussed for helping to strengthen local advocacy.

Sampling

Research into digital security practices is intensive and sensitive. The 'sampling' for the field work was shaped by central factors around safety and the extent to which research activities would be of mutual benefit to both the researchers and the participating groups. Though the team spoke to several HRDs who believed they might be at active risk in their contexts in the pilot interviews, the team felt that it would be inadvisable to do field work in actively unstable contexts. Thus, the team followed the below criteria in planning in-country research with groups:

1. The team worked with groups already involved in an ongoing relationship with Tactical Tech, and who had expressed a motivation to learn about digital security practices.

2. The team worked with groups not perceived to be under any current direct physical threats, following a situational analysis.

These selection criteria, along with the small 'sample size', mean that findings from this study are highly contextual, however, the team saw these constraints as a necessary prerequisite to being able to do the research, rather than as a 'limitation'.

[64] Conclusions from the pilot can be found in the first section of this report, 'Background and Rationale'.

2 Sample

Environmental rights organisation

The environmental rights organisation which the team worked with expressed deep concern over the shifting political climate of their country, observing that rights issues were taking a back seat to environmentally harmful development priorities.

I saw that the government was systematic in quelling dissent, and that this was opening up avenues for surveillance and quelling resistance. We are at a juncture where the engine is going to be driven by economic interests. Rights have to take a back seat.[65]

The organisation observed the increased scrutiny of civil society. New rules restricting funding were accompanied by political discourse framing foreign sources of funding as cause for suspicion of the work. Other organisations were pressured or forced to scale down operations and lay off workers as they lost crucial sources of funding. These limitations on funding and the atmosphere of scrutiny contributed to a so-called 'shrinking space' for civil society. [66]

The organisation also expressed concerns over the accelerated deployment of mass and targeted surveillance systems in their country, worrying equally about 'data integration' by companies and stories of targeted threats, such as incidents of phone tapping. This concern intersected with concerns about the increased scrutiny of civil society:

In the recent context of NGOs coming under the scanner, I won't put it past this government to creep up on your data if they can access it.[67]

A top priority for the organisation was to ensure the security and confidentiality of materials shared between themselves and partner organisations. The organisation felt that together with the importance of implementing certain practices themselves, it would be equally important to know how to advocate for and spread fundamental security practices throughout their broader network.

ITC4D and human rights network

In the country where this ICT for Development (ICT4D) and human rights network works, internet access is still uneven across the country, and *'digital security is a new concept'*.[68] Until recently, there were more independently owned ISPs, but rapid economic development has lead to a consolidation of local infrastructure. Now there are only a few telecommunications companies available to consumers for mobile and internet access. At the same time, the country is seeing high levels of investment in ICTs from a neighbouring country well known for its use of surveillance and censorship technologies. The consolidation of the telecommunications sector and the foreign investment in infrastructure were cited as a cause for concern among HRDs, who worried that a regional neighbour provides free internet access in order to gain access to the data generated by a growing number of internet users in the country.

[65] Field Work Interview with anonymous HRD site A #1, 2014.
[66] http://www.un.org/press/en/2014/gashc4112.doc.htm
[67] Field Work Interview with anonymous HRD Site A #1
[68] Field Work Interview with anonymous HRD Site B #1, 2014

The ICT4D and human rights network described a shifting perspective among civil society regarding the role of ICTs in democratic participation. Until recently, resources available to civil society engagement with ICTs went towards the promotion of social media platforms for different kinds of activism and civil society engagement, without much attention turned towards information management strategies or digital security: "Facebook is used to mobilise on issues, with little literacy of the risks. Youth are very active on Facebook."[69] HRDs believed that the rising availability and popularity of Facebook is seen as a threat to stability by the government, raising fears of censorship:

The government fears the information age. Information is shared on this social media site much faster than in the mainstream media. People now try to criticize the government using Facebook. There is worry among everyone using Facebook, young and old generations, that it will be blocked.[70]

A pending cyber crime law served as a catalyst for discussions between civil society groups (CSOs) on their ability to conduct their work freely. Though the recently introduced cyber crime law is framed around a stated goal of online safety, it is described by civil society as a means to criminalise rights-based work. The network observed that neighbouring regions appeared to follow each other's example in applying new cyber crime laws, giving civil society an idea of what might happen locally.

The network told the team that discussions of digital security could risk intimidating those who had only recently begun to enjoy the use of ICTs. Thus it was seen as crucial that digital security be framed as an enabling practice rather than as a barrier to the use of ICTs.

WOMEN'S AND LGBTQI RIGHTS NETWORK

As with the ICT4D and human rights network, the women's and LGBTQI rights network which the team worked with is based in a country with limited internet access, but which is experiencing rapid rates of economic development. The network observed widespread excitement among the public about the use of Facebook. Interviewees similarly noted that a neighbouring country known for its censorship and surveillance measures controls much of the local internet infrastructure. Because of this development, the network expressed concerns over how the use of different platforms and services might expose sensitive information to this neighbouring government. The network noted that the government of their country has recently begun to devote more resources to combating cyber crime, causing HRDs in the network to worry that this will have a harmful rather than helpful effect on the digital security and privacy of individuals: "National security is one thing: individual digital security is another."[71]

The women's and LGBTQI rights network relied on a public presence online and offline to mobilise, recruit volunteers and gain acceptance for women's rights issues. As the political context changed, it saw advocates working on women's and LGBTQI rights issues become scapegoats for politicians pushing for a return to traditional family values. The network experienced repeated waves of online harassment and attacks in the street by organisations they believed had connections with the government. When public organisational materials and personal photographs on Facebook were used in misinformation campaigns against the

[69] Field Work Interview with anonymous HRD Site B #1, 2014

[70] Field Work Interview with anonymous HRD Site B #4, 2014

[71] Field Work Interview with anonymous HRD Site C #6, 2014

network, its members saw a need to shift their information handling practices.

They used my photos, they exposed my personal contacts, and we couldn't reach the creator of the video because it was done in an anonymous way. We asked ourselves, what are the next steps we need to do?[72]

Though many members of the women's an LBGTQI rights network had done extensive work building a security strategy, including learning how to navigate social media settings and build strong passwords, many had concerns over the fact that no amount of vigilance would allow them to keep up with Facebook's changing policies.

METHODS

The study used qualitative approach[73] based on existing organisational conventions and practices, taking into account priorities around safety, learning, and reciprocity.[74] The research activities, which are listed fully in Appendix 1 of this report, were planned jointly and distributed between the lead trainer and researcher-trainer in order to ensure these priorities were continuously met. The researcher-trainer conducted semi-structured interviews, observation and in-training discussion and followed up with questionnaires, follow-up interviews and opportunities for further learning.

The semi-structured interviews complemented questions frequently touched upon in digital security training pre-assessment questionnaires typically used to prepare for trainings, but which allowed for a free discussion of concepts and encouraged the kinds of stories that questionnaires don't allow for. Interview themes broadly covered security awareness and practices. Discussions were held on how concerns over digital security initially arose and how HRDs decided to first try implementing digital security practices. Interviews covered past experiences with digital security tools, with regards to frustrations and perceived benefits in their use. The interviews also inquired into participants' learning priorities and offered pertinent information to the lead trainer in order to help shape agendas for the trainings.[75]

Following interviews, the team facilitated trainings taking place over several days. The discussion and participatory activities undertaken during trainings were based upon an existing repertoire of exercises developed and used by trainers which were reconfigured in a few instances to encourage more discussion and reflection. Activities were based largely on existing modules in the LevelUp resource for trainers.[76] Activities were chosen if they were deemed constructive to the training. The team chose to forgo particular activities if they did not suit the learning styles of participants or took away from the adult pedagogy-based rhythm of the training.[77]

Over the course of the training, the researcher-trainer took notes on the discussions taking place in a notebook. Notes were later transferred to the computer, collated and analysed together with the interviews and other materials. After the conclusion of the trainings, follow-up was conducted with the distribution of short open-ended surveys and VOIP discussions. These materials, together with semi-structured interviews and notes from the training, were coded in an open,

[72] Field Work Interview with anonymous HRD Site C #2, 2014

[73] Baskerville, R.L. (1999) op.cit

[74] Maiter, S., Simich, L., Jacobson, N. & Wise, J. (2008). Reciprocity: An ethic for community-based participatory action research *Action Research* 6: 305

[75] The final agendas were determined though the context analysis and participatory evaluations conducted throughout the training

[76] The LevelUp resource can be found here: https://www.level-up.cc/

[77] For more on andragogy, please see: https://www.level-up.cc/resources-for-trainers/pedagogical-resources/adult-learners

inductive fashion[78] in order to allow themes to emerge. It should be acknowledged, however, that the themes emerging through the coding process were undoubtedly shaped by the interventionist nature of this study, which already had specific goals and priorities guiding its course.

3 Operationalising Ethics in Practitioner Research

Security concerns were foremost in the design and implementation of this study. The research design hinged upon the safety of groups but equally on the necessity for it to be inclusive in methods and aims. The team saw the extent of inclusiveness of the research and the security and safety of participants and materials as interdependent considerations in structuring the research approach. The team aimed to set standards which met those required in humanitarian and human rights documentation work. The Do No Harm framework by CDA Collaborative Learning Projects[79] was found to be particularly informative. Do No Harm in this study began with the acknowledgement that both research and training interventions alter contexts and that both the short and long-term implications of an intervention need to be considered.

Despite the rich history of human rights and ICT4D (ICT for development) field research and an emergent body of empirical research into the impact of surveillance on human rights, there are no standardised best practices for conducting this kind of research in a way that protects the research participants. While in academic institutions, the Institutional Review Board plays a deciding role in shaping research involving human subjects, questions around whether to keep identities anonymous and how to plan research to align with operational security concerns remain open.

In both practitioner work and academic scholarship, some research showcases the names of HRDs and advocates to draw attention to their cause, which can have important benefits. Highlighting names through published research and stories can attract attention to human rights abuses in the media and assist in main-streaming rights efforts and drum up support for organisations. Using legal names in order to document the harms of surveillance can also strengthen its status as evidence in court. An approach relying on visibility can thus constitute a protection strategy in its own right, if the HRDs and advocates in question have requested this visibility. However, even with the inclusion of a risk assessment, neither human rights organisations nor individual HRDs can necessarily anticipate the unintended consequences of such exposure.

This study adopted a policy of anonymisation of identities and locations because many of the HRDs the team worked with rely on discretion and anonymity to do work in challenging environments. Knowing ahead of time that anonymisation would be necessary, the team planned the format of the research findings with an eye to preserving contextual details, while at the same time abstracting, aggregating and removing potential identifiers. Names were stripped from research documents and stored in a separate encrypted document. Organisation titles and countries were taken out of research materials before dissemination outside the immediate team. Specific details of secure and sensitive fieldwork are included in Appendix 2.

[78] Saldana, J. (2009). *The Coding Manual for Qualitative Researchers*. Los Angeles, CA: SAGE

[79] A description of the Do No Harm framework by the CDA Collaborative: http://www.cdacollaborative.org/programs/do-no-harm/key-principles-in-do-no-harm-and-conflict-sensitivity/ .

4 Benefits of this Research Approach

In light of the constraints placed on the Digital Security in Context study, its important to note what the team perceived to be its benefits:

- The team worked with groups rather than isolated individuals. This allowed the team to focus on fostering practices among people with an expressed need to communicate and share information securely and to strategise effective forms of advocacy for digital security within larger organisations and networks.
- The interviews conducted with participants ahead of each training aided the team in its preparation as facilitators. Though trainers and facilitators at Tactical Tech commonly use pre-assessment surveys in preparation for trainings, the additional interviews allowed training participants to describe experiences in a more open, conversational format.
- During the training, the team conducted a set of participatory assessment and evaluation activities. These activities allowed the team to follow a stronger and more dynamic training agenda and helped to solve a recurring evaluation problem in digital security trainings, where end-of-training evaluations tend to provide short answers and reflect a 'gratitude bias' towards the trainer.
- The team was able to conduct one follow-up training, a type of intervention which trainers often point to as a missing key to building and sustaining digital security practices, but which is often left under-resourced or unplanned. Participants reported that the second training provided a crucial opportunity to review concepts and tools introduced in the first training. Additionally, the second training both strengthened skills and provided the space for participants to brainstorm how to communicate concepts and skills outside of the training group.

Results and Findings: Shifting Landscapes of Evolving Threats, Technologies and Responses

> *I don't have the background knowledge to protect our website from hacking, because I know that nowadays technology changes from day to day.*[80]

In this study, the team paired trainings with a set of complementary research activities in order to carry out action-orientated research. Based on three workshops, and 60 interviews, the key themes that emerged from an analysis of the data that are discussed in this section. The results indicate that the three groups Tactical Tech worked with - an environmental rights organisation, an ICT for Development (ICT4D) and human rights network and a women's and LGBTQI rights network- are facing unique challenges, there are also shared issues of concern. This section presents result and an analysis of factors influencing how these three groups learned about digital security practices.

Though the three groups faced different challenges, all groups experienced the effects of new administrative barriers and laws restricting for the work of civil society organisations. Two groups witnessed a high level of economic development which lead to a consolidation of ICT infrastructure, together with new regional geopolitical alignments threatening more surveillance and censorship. In both countries, participants pointed out that 'the government still relies on people' for its monitoring and censorship, but that governments were working in close alignment with powerful regional neighbours to acquire more advanced systems that would allow them to automatically monitor and censor traffic. These two groups also saw their governments as preoccupied by new cyber policies which they felt powerless to shape as civil society actors.

The groups also dealt with a common dependency on unstable, commercial platforms which do little to protect user privacy. Sometimes HRD's chose commercial platforms such as Google and Facebook because they believed that their governments did not have access to the technology necessary to monitor these platforms. However, if governments were to move to exercise sovereignty through 'data localisation' efforts, these platforms would become more vulnerable to monitoring and surveillance by local intelligence agencies. The three groups found it difficult to adjust their strategies in response to potential threats emerging from shrinking civil society spaces, changing infrastructures, increased levels of surveillance and

[80] Field Work Interview with anonymous HRD site B #4 2014

censorship and unreliable ICTs and online platforms.

The next parts of this section describe the contextual factors affecting HRDs use and maintenance of security practices. These are presented as four main themes.

- The first is a direct outcome of cross-cultural contexts of technology transfer: language differences as a barrier to learning about and applying security practices.
- There are challenges in using open source privacy enhancing tools recommended by Tactical Tech and similar organisations in the skills-transfer and training process.
- Security emerges when practised in a collective; a case study from one of the research sites is discussed here to show the offline and online components of this, and exposes limitations in the construction of security and privacy as problems of the individual.
- The long-term integration of security practices within organisations and networks is dependent on a number of factors that must be factored in to the uptake of digital security strategies.

These findings inform recommendations to communities of practice engaged in and supporting digital security learning and uptake for HRDs.

Country names are left out in our descriptions in line with the de-identified and anonymised approach used to present the findings.

1 Responding to a Shifting Landscape

Among the HRDs engaged in this research, awareness of surveillance, privacy, and digital security often arose from stories of security incidents spreading through peer groups, organisations and networks. Knowledge of security incidents spread through stories and narratives originating from media sources: "I have been quite worried about our website after we heard about the hackers who tried to hack the government website."[81] Stories also spread through family, friends and colleagues: "I met someone who was talking about how her phone got tapped."[82] Stories served as warnings of potential risks: "I'd heard about threats somewhere else. Then I got hacked."[83] With the heightened awareness raised by these kinds of stories, suspicions of phone tapping and other forms of surveillance and intrusion were common. Many suspicions were later confirmed through concrete incidents or through investigations undertaken by the HRDs.

The HRDs spoken to in pilot interviews described being targeted by government malware, online harassment, email hacking, mobile phone interception and confiscation of their devices. Threats were also commonly targeted at colleagues and peers in order to obtain information about organisations and networks, or in order to hit personal 'weak points.' Sensitive information obtained through the use of intrusive malware and phone tapping was later used to blackmail the families of two HRDs in order to pressure them into ceasing their work. Hacked social media accounts exposed information about group actions, which lead to the

[81] Field Work Interview with anonymous HRD Site B #7, 2014
[82] Field Work Interview with anonymous HRD Site A #1, 2014
[83] Field Work Interview with anonymous HRD Site A #1, 2014

disruption of a protest.

Direct experiences led the HRDs spoken to in pilot interviews to seek ways to protect themselves, their families and their organisations. Most introduced security practices into workflows, including changing passwords more frequently and teaching others to do so as well. Some HRDs devised counter-surveillance strategies and physical security protocols, installing better lighting systems around offices, monitoring social media, devising alternate online identities and making contacts within local police departments for protection. Responsive measures were undertaken with consideration for peers, networks, and organisations. HRDs underlined that privacy and security issues are understood in relation to friends, families, organisations, and networks. “Privacy is not only about ourselves but the people connected to us”.[84]

The HRDs spoken to in pilot interviews sought trusted outside support and advice on digital security from peers and colleagues who had experienced similar incidents or from digital security trainers and IT experts. Trainings served as a forum for discussion and learning. Strategies undertaken independently by HRDs were honed, augmented and modified through work with facilitators specialising in protection work. Strategies learned in trainings were applied in several important instances in order to strengthen both physical and digital security. However, almost no HRDs spoken to in pilot interviews made use of tools for encrypted communications because they were deemed 'challenging' to integrate into workflows.

Experiences shared in the 13 pilot interviews demonstrate that privacy and digital security concerns and responses emerge through relationships among peers and colleagues. The following case, relating to the women's and LGBTQI network which the team worked with in the study suggests that security practices evolve through a collective negotiation of priorities, particularly due to a shifting understanding of what constitutes effective protection, and the particular constraints of available technologies and resources.

Security Emerges Through the Collective

Some time ago, the women's and LGBTQI rights network were subjected to a series of attacks following public actions organised through Facebook. After first being attacked verbally online and then at protests in the street, the head of the organisation sought support:

That was the time that we saw that not only did we need to think about physical security but also online security. We didn't have knowledge and didn't trust any IT people in our surroundings. We didn't know of any female IT specialist that could help us, so we were obliged to ask someone from outside and pay him. It was very challenging to trust anyone.[85]

Unable to find a trusted specialist, the organisation implemented basic digital security practices and created a new office-wide physical security policy. The organisation also began sharing information on threatening circumstances with other local organisations working on parallel issues. This information sharing process led to the creation of a collaborative security strategy within a network of organisations.

We started exchanging information and started doing things that are obvious, like changing passwords, thinking about what to put online and what not to put online, how to deal with

[84] Field Work Interview with anonymous HRD site B #4, 2014

[85] Field Work Interview with anonymous HRD site C # 13, 2014

comments on Facebook, how to monitor ourselves more. At first we wouldn't monitor social media and now we do that 24 hours a day almost.[86]

The network also set up a secret Facebook group in order to document security incidents such as online harassment and attacks at protests. Tracking incidents together enabled the network to find common patterns in the behaviours of their harassers. The similar tone and language of the 'online' and 'offline' harassment allowed the network to establish links between these behaviours and trace responsibility to a certain set of actors.

One HRD who spearheaded this effort described first hearing about the idea of documenting security incidents in a security training about a year before the attacks.[87] At the time, the HRD couldn't find any immediately relevant application for this advice, but the attacks changed his understanding of protection:

> *Before the attacks I felt it was more important for people not to worry than to know about threats [made] on Facebook. I was keeping secret from everybody when I was receiving threats on Facebook. I used to delete them. I thought it was protection if they didn't know.*[88]

While the use of Facebook compromised privacy and security, the use of the platform also facilitated information sharing among peers, which aided the development of a collaborative protection strategy. As one HRD reported: "Security becomes more real when it is about my colleagues."[89] Another HRD noted:

> *Much of our understanding of safety is about reaching out to other people instead of relying on ourselves. We need to be able to rely on community.*[90]

As this case demonstrates, the evolving prioritisation of security led the LGBTQI network to make use of those imperfect tools and platforms which proved easiest to implement in light of the constraints particular to its workflows. Interviewees from the network shared that tactics learned in trainings aided security strategies, building on what HRDs shared in pilot interviews about the positive role of trainings in their human rights work. However, for both the LGBTQI network and pilot interviewees, it proved difficult to use 'challenging tools' such as those offering encrypted communications. This report addresses the particular challenges of using different digital security tools in the next section.

2 Introducing Digital Security Tools and Practices

Choosing tools to meet priorities

This section of the report expands upon the challenges of building practices making use of digital security tools featured in trainings and in the Security in-a-Box resources.

Tactical Tech's digital security capacity building efforts are not concerned with insisting on the

[86] Field Work Interview with anonymous HRD site C #4, 2014
[87] This security training was not conducted by Tactical Tech.
[88] Field Work Interview with anonymous HRD site C #14, 2014
[89] Field Work Interview with anonymous HRD site C #3, 2014
[90] Field Work Interview with anonymous HRD site C #7, 2014

adoption of particular tools. The reasons for this are philosophical and practical: any form of communication can be compromised and thus become a reason for concern, but this does not mean that each potential 'risk' of compromise should be of concern in a given situation. Because the landscape of threats and digital security tools changes so rapidly, it would be a poor use of resources to devote limited training time to covering a prescribed set of tools. Insisting on certain tools would also be out of line with the adult education methods applied in trainings,[91] which respect the 'expert knowledge' and contextually driven priorities of participants.

The desired focus in Tactical Tech digital security trainings and resources is on surfacing security concerns and addressing them through work to develop contextually driven digital security practices. Facilitators then work with participants to determine what a meaningful integration of digital security tools within their work would look like. These steps are taken as part of the context analysis process which focuses on surfacing priorities and concerns particular to the situation.[92]

Tactical Tech does however follow a particular set of criteria in the kinds of tools it introduces to groups in hands-on work and in online resources. Tactical Tech fosters the use of actively maintained, trusted and preferably free/libre/open-source software (FLOSS) tools and platforms,[93] because these tools are seen to offer a strong value proposition in an online environment lacking in protection for individual users. Through their federated and decentralised infrastructure, FLOSS tools remove the need to 'trust' a company to abide by the privacy and data collection policies they claim to follow. FLOSS 'end-to-end' encryption such as GPG and OTR leaves encryption keys with users rather than having them stored in company servers. Additionally, FLOSS tools are created within a community of developers who share an imperative to make the source code available for review. The transparency of source code provides a way to verify that a tool lives up to the claims made by its developers.

However, not all FLOSS tools are robust in their code and infrastructure and many are withdrawn from development with limited notice to users.[94] Additionally, the openness of the tool architecture doesn't guarantee that code has been recently or thoroughly reviewed; and even if it has, exploitation or tampering is always possible, as was demonstrated with the revelation of the NSA 'BULLRUN'' program[95] and with the discovery of the 'heartbleed' Open SSL bug.[96] Furthermore, being able to see the code is generally not of direct, tangible benefit to the vast majority of 'end-users' including HRDs, as verifying the properties of code requires the involvement of developers and security researchers.

Additionally, cryptographic access control tools such as those using OTR and GPG often leave metadata exposed, meaning that a trail of data is still left for opponents of human rights work or commercial third parties even if the content of messages is shielded from view. Neither can these tools fully protect against 'advanced persistent threats' – attacks

[91] https://www.level-up.cc/resources-for-trainers/pedagogical-resources/adult-learners

[92] For a description of a context analysis, please see the first section of this report, Background and Rationale, under Practice-based research foundations A description of exercises done to surface and highlight the role of communications infrastructure in enabling communications can be found in the Appendix. One exercise involved having participants 'draw the internet' in order to surface personal experiences using technology in their work. Drawings and descriptions by four participants are included.

[93] Tool selection criteria can be found here https://securityinabox.org/en/about

[94] An example of such a tool can be seen in the Truecrypt controversies of 2014. https://en.wikipedia.org/wiki/TrueCrypt

[95] http://www.theguardian.com/world/interactive/2013/sep/05/nsa-project-bullrun-classification-guide

[96] http://heartbleed.com

and social engineering strategies which allow intruders with sufficient resources to bypass encryption.

Finally, for HRDs – and for ICT 'users' generally – the use of tools can mean abandoning the popular platforms where audiences, peer groups and allied organisations plan actions and showcase campaigns. Taking the above inadequacies into consideration, Tactical Tech's Security in-a-Box tool-kit primarily recommends FLOSS tools that have been tested over time through a feedback cycle between trainers and participants; that are maintained by active communities; and that have had their source code actively reviewed by information security experts.[97]

The following subsections focus on participant experiences with a particular set of digital security practices involving anti-virus, chat, social media, password management, email and data storage, because these were consistently flagged for their importance in trainings conducted within this research. Though this report touches on usability issues, the focus is not on the interface but upon the evolution in tool development and practices and the kind of confusion this can create among trainers, developers and users. Of note, it was not possible for the team to cover every new development or problem within the ecosystem of tools and practices described within the scope of this report and it is likely that some of this information will soon be outdated due to the rapid pace of technological change.

PREVENTING AND REACTING TO MALWARE

A good example of a constantly shifting and evolving 'threat vector' can be seen in the spread of malware. The problems of prevention and harm mitigation are challenging for everyone from civil society actors to industry anti-virus and anti-malware vendors and the forensic analysts who track and 'reverse engineer' them. Common anti-virus and anti-malware programs are unable to keep up with the evolution of some of the most intrusive forms of malware of concern to the human rights community, and it is difficult even for forensic analysts to correctly attribute the origin of 'advanced persistent threats'.[98]

At the level of common end-user software used to detect malware infections, participants in the three trainings in this study often arrived to the training with some form of anti-virus program already installed, but didn't realise that the virus definitions that enable an anti-virus program to identify infections had not been updated in a long time. This was down to the fact that they expected the software program to automatically run updates. Without updating the anti-virus, HRDs were left exposed despite taking steps to protect themselves.

There was also a common confusion about the intent behind common malware attacks. A substantial number of participants in the three trainings conducted in this study believed that malware is deployed with the intention of destroying files. Yet the concerns that have drawn the most attention in the human rights community of late come about from the way targeted malware has been used to spy on and extract sensitive information from HRDs rather than to destroy files or computers.

'Phishing' and 'spear fishing' techniques have included sending emails with fake conference

[97] Tool selection criteria can be found here https://securityinabox.org/en/about

[98] The 'attribution problem' is explained in this article: http://www.csoonline.com/article/2881469/malware-cybercrime/whodunit-in-cybercrime-attribution-is-not-easy.html

invitations[99] in order to manipulate HRDs into opening attachments or clicking on false URLs and inadvertently executing malware programmes. Malware uncovered in recent years includes Remote Administration Tools[100](RATs) containing key-loggers to record a users' keystrokes, screen capture and image control capabilities to observe users through the computer's webcam, and the ability to access system files. With an invasive spying tool like a RAT, protective measures such as encryption are easily subverted and the integrity of the machine may remain compromised despite re-installation of the operating system. For HRDs, responses to malware issues can quickly become a resource issue, as many cannot afford to discard machines even if they suspect an infection.

Due to rapidly changing threat vectors for malware attacks and the difficulty of detecting infections, educational resources and trainings have emphasised the importance of prevention through best practices in avoiding a malware infection or have identified 'indicators of compromise'.[101] A digital security training might focus on proper installation and updating of virus and anti-malware software and offer pointers for exercising caution in opening email attachments and links. Users might be urged to confirm with email senders that they intended to send an attachment or to check the URL of links to identify suspicious wording which are indicative of an attempt to conduct a man-in-the-middle attack. But as The Citizen Lab's Communities @ Risk report notes:

> *Threat actors are highly motivated and will likely adapt their tactics as users change their behaviours. For example, it is possible that if every user in a particular community began to avoid opening attachments, attackers would move on to vectors such as watering hole attacks or attacks on cloud-based document platforms.*[102]

Human rights advocates are exploring different political and legal mechanisms to deal with the proliferation of the use of targeted malware,[103] but on a practical level, infrastructure to deal with these kinds of threats is nascent.[104]

Policies, identities and harm mitigation in the use of social media platforms

Along with the creation and management of strong, complex passwords, the most common practice to continue beyond trainings was more frequent checking of privacy settings in social networking platforms such as Facebook, with participants taking actions to minimise the disclosure of sensitive information. One participant related that after participating in a training, she was more sensitive to the visibility of personal content on Facebook and in her responsibility in revealing the identities of friends: “Before the training I tended to tag people

[99] Galperin, Eva, for the Electronic Frontier Foundation. January 19, 2014. Vietnamese Malware Gets Personal. https://www.eff.org/de/deeplinks/2014/01/vietnamese-malware-gets-personal

[100] Definition of RAT: https://en.wikipedia.org/wiki/Remote_administration_software

[101] Resources covering the prevention of malware infections: https://securityinabox.org/en/guide/malware and https://ssd.eff.org/en/module/how-do-i-protect-myself-against-malware and https://digitaldefenders.org/digitalfirstaid/#section-malware

[102] The Citizen Lab. 2014. Extended analysis: civil society perspectives and responses, from Communities @ Risk: Targeted Digital Threats Against Civil Society https://targetedthreats.net/media/2.3%20Extended%20Analysis-CivilSociety.pdf

[103] A coalition by civil society organisations calling for export controls of targeted malware http://www.globalcause.net/

[104] The Detekt tool, developed to help identify FinFisher infections and released in November 2014 https://resistsurveillance.org/

on Facebook because I thought well they are my friends and they are in a group photo".[105]

While this was a positive finding in light of Facebook's importance to the groups in their work, vigilance over settings didn't guarantee safety and privacy. Facebook's privacy settings were difficult to manage even for those who checked their settings regularly, due to a lack of adequate disclosure regarding Facebook's constantly updated Terms of Service. Unexpected changes to privacy settings meant sensitive content was exposed to larger circles than intended. This exposure affected the safety and security of the women's and LGBTQI rights network. Harassers exploited the exposure of these photos, using them to fuel spurious misinformation campaigns.

Criticism of Facebook's constantly shifting policies recently lead the company to create a feature Facebook calls the 'Privacy Check-up,' which provides a streamlined set of privacy controls for users. While this addresses some of the issues regarding information shared between users – what's called their 'social privacy'[106] – users are still unable to prevent Facebook selling their data and making it available to third parties such as intelligence agencies, credit rating agencies and other institutions. Thus, 'instrumental privacy' remains illusive.

Facebook's 'real names' policy – which requires the use of 'authentic names' in profiles – also caused substantial problems for the women's and LGBTQI rights network. Initially the network had publicly viewable profiles and used legal names. This allowed their profiles to serve as easy points of contact for women seeking them out for support. However, after experiencing a wave of harassment and threats of violence, it became crucial for members of the network to be able to use pseudonyms. Being able to control how names appeared online was an important element of safety. In response to fierce criticism[107] over the 'real names' policy, Facebook clarified its stance on the policy in early 2015, saying the company was not requiring legal names but 'authentic' names conforming to 'offline' identities, but did not change their overall policy. In a promising development in July 2015, a German regulator ordered Facebook to allow pseudonyms on the ground that the policy violates the right to privacy.[108]

Despite concerns over Facebook in the three trainings conducted in this study, no participants spoke of completely leaving Facebook in response to threats they faced. This fact supports the findings of the HCI study 'Limiting, Leaving, and (re)Lapsing: An Exploration of Facebook Non-Use Practices and Experiences,' which found that users often express a desire to stop using particular ICTs due to a number of possible concerns but feel unable to cease use altogether because of social dependencies.[109] Whilst Tactical Tech does promote alternatives to Facebook such as Diaspora[110] and Crabgrass,[111] participants felt 'locked in' to Facebook because their audiences and communities continued to use it.

---

[105]Field Work Interview with anonymous HRD site B #4

[106]De Wolf, Ralf. Rob Heyman, Jo Pierson. Privacy by Design through social requirements analysis of social network sites from a user perspective. www.cosic.esat.kuleuven.be/spion VUB, IBBT-SMIT, Brussels, Belgium SPION1/EMSOC2

[107] A comprehensive history of controversy around the 'real names policy' https://en.wikipedia.org/wiki/Facebook_real-name_policy_controversy

[108]http://www.theguardian.com/technology/2015/jul/29/germany-fights-facebook-over-real-names-policy

[109]Baumer, E. Adams, P., Knovanskaya V., Liao, T.C., Smith, M. E., Sosik, V.S.& Williams, K. Limiting, Leaving, and (re)Lapsing: An Exploration of Facebook Non-Use Practices and Experiences (2013) CHI 2013, Paris, France, April 27 – May 2, 2013.

[110]The Diaspora social network site: https://diasporafoundation.org/

[111]The Crabgrass organising platform: https://crabgrass.riseup.net/

PASSWORD MANAGEMENT THROUGH SOFTWARE-SUPPORTED STRATEGIES

Along with the increased sensitivity to social media privacy settings, the most common practice maintained two months after the conclusion of the three trainings was password management through Keepass and KeepassX password management software. However, though password practices are often considered a basic skill by digital security trainers, best practice in password choice and management – either through software or manual means – was not considered to be intuitive by many training participants.

A sizeable majority of participants displayed an awareness of the vulnerability of passwords generally, but were not aware of different options available to remedy this vulnerability. Most participants relied on memory to recall passwords, but described a difficulty in remembering more than a few passwords at any one time. They thus relied on a small set of passwords with slight differences, which were used across different devices and services.

Difficulties around passwords were magnified by complex requirements that passwords must be long, complex and unique and be changed regularly due to the fact that passwords are easily compromised. For this reason, in trainings conducted in this study, emphasis was placed on the fact that passwords can be made memorable and personal, and a set of activities was conducted to strategise how to make them so. In one exercise, participants had to to come up with reusable codes that allow them to generate any number of unique passwords without the aid of any technology.

The team introduced Keepass and KeepassX password management software as an option allowing people to delegate their memory of passwords to a program and store them in one secure place. Using this software could potentially make it easier for users to follow recommendations such as having long, complex pass phrases that incorporate different kinds of characters, or using a unique password for each service. However, though password management software promises to aid strong password practices, the function of the software often promoted feelings of distrust. Participants described how easy it was to forget to save the password database into which they stored newly generated passwords. If they forgot to 'click save', then when participants closed the database, they lost all their new passwords. Often participants forgot where they stored their database files or accidentally deleted them. If a participant did not also have a consistent practice around 'backing up' their information, then they were likely to lose access to many of their password-protected accounts in the event that they lost their password data.

In terms of relative difficulty of use, the task of setting up password management software across different systems and devices was seen by some participants to be as much of a barrier as attempting to implement PGP/GPG email encryption – the most common tool exemplified for its lack of usability among digital security tools. Participants were often unclear where they should download the appropriate version of the password management software for different devices, how they could install it and how passwords stored on one device would be transferred to another. Some users were concerned about giving away the fact that they stored all their passwords in one file. Others expected to be able to log in directly to services by opening up their password database. Taking the extra step of copying and pasting passwords from the database into the login screens of different services was seen as an unwelcome step; however, when participants realised they would no longer have to worry about remembering passwords, it was seen as worthwhile.

Despite the fact that password management could be as complicated in practice for some as GPG, it was the most common practice to persist. One reason for this may be that the mechanics of password management does not require the participation of others. Of course, if a user is sharing access to a database or communications platform within a group, it quickly becomes important for every member to create strong passwords, lest some form of group communications is compromised. Thus, even if the mechanics of the tool are dependent on one individual, broader practices require interaction with others.

CHAT APPLICATIONS: SIFTING THROUGH FALSE CLAIMS, COMMUNICATING SECURELY IN GROUPS

A number of the new chat applications have marketed themselves with strong claims about the privacy and digital security protected and enabled through their use. Though some claims have since been proven false, that kind of information did not reach participants consistently. A good example of misinformation regarding the relative benefits of tools concerns the tool Snapchat. Participants were not aware that regulators accused Snapchat developers of making false promises regarding the deletion of photos sent through the application.[112]

Confusion over the security of chat applications affected trainers and developers equally. For a period of time in 2014, trainers in the human rights sector were hopeful about the merits of an open source application called Surespot. Surespot allows a user with one device to maintain several different accounts at the same time and to easily delete them. Account names are not required to correspond to 'real' or legal names and are not publicly indexed. This kind of infrastructure helps avoid privacy concerns arising in other applications. At the end of July 2015, however, rumours began to circulate in the security community that Surespot was found to have been potentially compromised. Surespot developers responded with a blog post stating that rumours were unfounded and that the tool remained strong and secure.[113]

These sorts of confusions being quite common, participants expressed the need for a vetting system to get beyond marketing claims with regards to chat applications such as Snapchat, Threema, Wickr, Firechat, Line, Telegram and WhatsApp. A couple of projects have taken on this problem of late: the Open Integrity Index has been developing criteria to understand the merits of tools and platforms more broadly,[114] and the Electronic Frontier Foundation has compiled a comparative chart on the relative merits of an array of chat applications.[115]

In terms of priorities with regard to chat applications, many participants wished for more privacy-protecting options for chat within groups. Though a number of new mobile applications have been released in the last two years which allow for seamless 'many-to-many' communications, strong 'end-to-end' encryption within group chat remains out of reach in many mobile applications. Since the time at which trainings in this study were conducted, the Signal[116] application by open source developer Open Whisper Systems has become a notable exception. The application does offer encrypted group chat, and while previously only available for iOS, is now also available for Android.

---

[112] Rushe, D. (8 May 2014). Regulators reprimand Snapchat over false claims about messaging service *The Guardian* http://www.theguardian.com/technology/2014/may/08/snapchat-ftc-false-claims-messaging-service

[113] Surespot. (24 September 2015) Don't Believe the Hype http://surespotencryptedmessenger.blogspot.com

[114] Open Integrity Index tool vetting project: https://openintegrity.org/en

[115] Electronic Frontier Foundation Secure Messaging Scorecard https://www.eff.org/secure-messaging-scorecard

[116] https://whispersystems.org/blog/signal/

Multi-party encrypted chat was however not a possibility with desktop chat applications such as Jitsi and Pidgin, which are recommended in Security in-a-Box resources. These applications use Off the Record Messaging, an encryption protocol which was built for one-to-one communication. Recently, developers have been working on a 'multi-party' version of OTR (mpOTR) but have found it non-trivial to change the specification to allow for multi-party encryption. Efforts to develop mpOTR have been lead by original OTR developer Ian Goldberg,[117] Cryptocat,[118] and through the work of the group Equalite under the banner of the np1sec Project. A white paper outlines the current approach.[119] Open source developer Open Whisper Systems says that encrypted 'multi-party' group chat using a different encryption protocol is on the roadmap for integration with the popular tool WhatsApp.[120]

The dearth of 'multi-party encrypted group chat options might explain the enthusiasm encountered when introducing a FLOSS platform called Jitsi Meet in trainings.[121] This platform, which works off of Jitsi's VideoBridge,[122] enables users to carry out encrypted individual and group chat, share documents and conduct video conferences. It also has the advantage of running in the browser, so it does not require an installation process. However, the use of any tool comes with a slew of caveats. At the time of writing, Jitsi Meet is only available for use in two browsers, and its relative security – as with any other tool claiming to be security-protecting or privacy-enhancing – continues to be a subject of debate.

ENCRYPTED EMAIL: STRONG CRYPTOGRAPHY, PERSISTENT PROBLEMS WITH USE

The use of email encryption, either through Pretty Good Privacy (PGP) and the open source implementation Gnu Privacy Guard (GPG) and the Enigmail plugin or through 'portable' applications such as GPG4usb, was consistently cited as the most difficult tool to incorporate within practices.[123] Most participants cited concepts around key management[124] to be very challenging to grasp. “I find (GPG) encryption hard. I try to use the key that the trainer trained me to use but it doesn’t work all the time.”

With the unclear distinction between the function of public and private keys, users exchanging GPG keys with friends and colleagues mistakenly attached and emailed their public keys together with their private keys, an action that would nullify the potential security benefits of using GPG while potentially further exposing sensitive information. Understanding the role of all the components necessary to use GPG with email clients such as Thunderbird was another contributing factor to feelings of difficulty.

The combination of difficulty with concepts around key management, problems with mail clients and a lack of people to use GPG with often led participants to abandon GPG use altogether. The majority of participants which participated in the followed up phase after the the training reported not continuing to use PGP/GPG among other things. This leaves a major vulnerability when using email. GPG offered an interesting instance where the expectations of

[117] Goldberg, I., Van Gundy M.D. & Ustaoglu, B. & Chen, H. (2009). Multi-party Off-the-Record Messaging at the 16th ACM conference on Computer and Communications Security, November 9–13, 2009, Chicago, Illinois, USA
[118] Cryptocat Github repository for mpOTR project plan https://github.com/cryptocat/cryptocat/wiki/mpOTR-Project-Plan
[119] Wiki for equalit.e Np1sec https://learn.equalit.ie/wiki/Np1sec
[120] https://whispersystems.org/blog/whatsapp/
[121] Jitsi Meet: https://meet.jit.si/
[122] Jitsi Videobridge: https://jitsi.org/Projects/JitsiVideobridge
[123] Field Work Interview with anonymous HRD cite B #3
[124] Security in-a-Box Guide for using GPG https://securityinabox.org/en/guide/thunderbird/windows

ease and difficulty between trainers, developers and users did not match. For example, while the portable tool gpg4usb is sometimes seen as a simpler option by trainers than setting up GPG with an email client, participants experienced problems when they tried to migrate from gpg4usb to a full mail client like Mozilla's Thunderbird.[125] In this sense, it was a similar problem to that experienced with password management software where worry over having to migrate information to other devices and interfaces became a major barrier to use.

PGP usability concerns have persisted for years. PGP's creator himself Phil Zimmermann recently admitted to not being able to reply to an encrypted email because he didn't know how to install PGP on his MacBook.[126]Mailvelope[127] and Mailpile[128] are examples of recent, experimental projects to simplify the experience of using PGP. Companies such as Google and Yahoo have also begun development of more user-friendly email encryption tools to be used within their platforms and services.[129] Critics note that email encryption using the infrastructure of commercial services guarantees that user metadata will continue to be captured and processed by third parties. Meanwhile, the open source GPG is struggling due to the difficult funding environment for FLOSS tools. GPG has been maintained for years by Werner Koch, a single individual with limited human and financial resources. Recently Koch received an influx of funding thanks to a crowd-funding campaign[130] beneficial in the short-term but inadequate for long-term infrastructure on its own.

Mediating meanings: language barriers and the importance of interpretation

While the above sections have focused on practices in relation to specific tools, the report will now turn to a consideration which the interviews and trainings conducted within this study have shown to affect learning about and implementing digital security practices more broadly. In dealing with the concerns arising within challenging contexts, HRDs engaged in this study needed to understand digital security and privacy through a predominantly English-language based lexicon. This issue around language added to the overall challenges of creating constructive and appropriate digital security strategies, becoming an issue both in tool use and in spaces for learning and discussion, such as trainings.

Two groups which the team worked with were only able to access English-language versions of certain digital security tools despite predominantly conducting their work in other languages. Though Security in-a-Box offers translated instructions for tool installation and implementation in 13 languages, many translation gaps remain with regards to different elements of the tools. The importance of translation with respect to the development of digital security practices was addressed in an interview conducted with a trainer and translator who has pushed for contextually appropriate translations of tools in recent years. By not making translation a priority, the trainer and translator felt they were contributing to the demise of their language: “I want to keep my language alive. It’s important to me. If you cannot translate it then you have no more language”.[131]

[125] Mozilla Thunderbird email client https://www.mozilla.org/de/thunderbird/
[126] Franceschi-Bicchierai, L. (September 2, 2015) Even the Inventor of PGP Doesn’t Use PGP http://motherboard.vice.com/read/even-the-inventor-of-pgp-doesnt-use-pgp
[127] Mailvelope: https://www.mailvelope.com
[128] Mailpile: https://www.mailpile.is/
[129] Google End to End https://github.com/google/end-to-end
[130] GPG crowdfunding campaign: https://lists.gnupg.org/pipermail/gnupg-announce/2013q4/000338.html
[131] Pilot Interview #15, 2013

However, many participants and trainers interviewed explained that the translation of tool-related resources and elements within tool interfaces doesn't guarantee the cultural legibility of tools and concepts. Training participants explained that translation efforts often fail to capture the correct, contextually appropriate or desirable words in their respective languages. One translator told the team that with regards to localisation efforts aiming to translate tools, "the challenge is not just translating, because there are certain words where there is no one-to-one meaning". In describing the difference between a direct translation and a meaningful contextualisation or interpretation of digital security content, participants and trainers who had encountered linguistic issues told the team that prioritising local, contextual meanings would mean exploring culturally relevant metaphors to describe human relationships to networked technologies.[132]

Participants emphasised the importance of drawing on local meanings relating to technological concepts throughout the cycle of tool development and within discussion of concepts around privacy and digital security. In the trainings conducted over the course of this study, the team held discussions to address cultural and linguistic issues arising over discussion of digital security and privacy and the use of digital security tools. In one group, the team learned that there was no appropriate analogy for the word 'protection', as the term for protection in the local language had a negative connotation.[133] Fittingly, the term 'encryption' translated to mean 'hard to understand'. The local word for 'surveillance' was largely not recognised by two groups. Participants from the group decided the appropriate word for surveillance would be 'monitoring', but this was not a commonly used word, and that when the word was used, monitoring related 'to people but not technology.'

The lack of one-to-one meaning for certain words and the need to re-interpret digital security concepts lead a trainer and translator spoken to in the pilot interviews to take the word 'firewall' and devise a new metaphor which would convey similar ideas:

> *For several years I was looking for someone to translate the world ' firewall'. When we translated fire and wall, people would ask what is the meaning of that? A wall of fire? Instead I made a metaphor focused on referring to a place that keeps us secure.*[134]

[132] For an examination of the role of interpretation in communicating technological concepts, see: Sun, H.(2009) "Designing for a dialogic view of interpretation in cross-cultural IT design." In *Internationalization, Design and Global Development*, pp. 108-116. Springer Berlin Heidelberg.

[133] Field Work group discussion, 2014

[134] Pilot interview #15, 2013

Results and Analysis: Integrating Digital Security Practices in Human Rights Workflows

This section of the report considers the 'long view' of digital security practices in human rights work, looking at processes observed by HRDs as being enabling factors for the continuation of digital security practices beyond initial digital security trainings.

1 Sustaining Digital Security Practices

Questionnaires and interviews conducted in the months after trainings showed that in addition to a greater general 'awareness' and concern over digital security issues, the most common tool-based practice continued after an initial training was password management through Keepass and KeepassX. There was also a higher sensitivity to privacy settings in social networking platforms such as Facebook. Not only were password practices and social media settings perceived to be the easiest practice to maintain at an individual level, but these skills also spread easily through groups.

However, digital security practices which involved coordination and communication between two or more people were much less likely to be continued: “If you have a security tool but they don't, you cannot communicate anything”.[135] All interviewees from pilot interviews and trainings expressed the view that a key barrier to adoption of digital security tools was having no one to use them with outside the training setting: “Even though I know and learn, I have no one to practice with. When you don’t have people to practice, you forget it.”[136]

Participants in all three groups which the team worked with thus felt that ‘difficult tools' could only be used with other training participants: “with some friends who have maybe been in the training I use encrypted emails. Only people with security training can understand it”.[137] In addition, participants told the team that the people they felt the most need to communicate securely with were not present in the training. In order to be able to use many of the tools taught in 'end-user' trainings, former training participants underlined the need to serve as ad-hoc trainers: “Unless I train others, I can’t use things from this

[135] Field Work Interview with anonymous HRD Cite B #4, 2014

[136] Field Work Interview with anonymous HRD Cite B #4, 2014

[137] Field Work Interview with anonymous HRD Cite B #3, 2014

training".'[138]

Many former training participants thus went on to serve as ad-hoc trainers and advocates in order to implement digital security practices learned in trainings. During the time that this research was carried out, several HRDs who began as training participants were trained as trainers or began to conduct informal workshops and trainings independently. Participants would often identify one person in their network or organisation who they saw as the most important 'agent of change' in kicking-off efforts to develop digital security strategies and promoting a change in practices. One important advocate for digital security was the executive director of an organisation, who after attending a short awareness-raising workshop immediately shared password practices and chat encryption with staff members and made an effort to integrate learning about digital security into different issues and topics at events and clinics.

After one of the trainings taking place during the Digital Security in Context study,[139] four out of 15 former participants self-organised workshops and skill shares with a focus on password practices and social media with colleagues and friends. One former training participant and student went on to train 30 fellow students at a local university. A development worker said he incorporated digital security into his usual repertoire in trainings focused on using social media effectively. Two former participants provided trainings to staff in their non-governmental organisation on passwords, along with 'email security' and 'Facebook security.'

Though this kind of advocacy proved crucial to the spread of practices within peer groups and organisations, former participants' experiences showed it was a challenge to serve as advocates for digital security practices. Former participants encountered difficulties in convincing others of their importance. As an example, an executive director who spent substantial energy into advocating for security practices with peer organisations found themselves being accused of undue paranoia: "A lot of people thought maybe we were being self-obsessed in thinking that the government cares what we're doing".[140]

Additionally, advocates shouldered what they felt to be a problematic extent of responsibility to drive organisational change without the aid of additional support and follow-up. Colleagues, collaborators and networks expected advocates to return home from digital security trainings ready to share and implement their new skills. It was also the case that several former participants who advocated for digital security did not always have the authority or local support to implement changes in IT or communications policies: "At my workplace, I'm administered by IT people and they won't allow staff to install any programs on their computers".[141] A lack of communication with IT staff or a divergent perspective on security matters became a real barrier to follow-up work.

With regards to specific tools and practices, former participants found it difficult to train peers and colleagues on tools and practices incorporating encryption: "It's a bit hard for me to advocate others to use it because with encryption you need time".[142] Difficulties with spreading encryption practices meant that while, for example, social media settings available

[138] Field Work Interview with anonymous HRD Cite B #2, 2014

[139] This training occurred before the start of the study. The researcher was present to interview participants afterwards; and for a subsequent follow-up training

[140] Field Work Interview with anonymous HRD Cite A #1, 2014

[141] Field Work interview with anonymous HRD Cite B #3, 2014

[142] Field Work Interview with anonymous HRD Cite B #10, 2014

within Facebook were often changed, some of the more technically robust privacy and security measures available for use with social media (such as tools using the Off the Record Messaging (OTR) protocol)[143] went unaddressed.

When former participants tried to introduce the use of OTR encryption with chat applications like Jitsi and Pidgin into organisations and networks, they often encountered reluctance from fellow staff due to their complex installation process. Unsurprisingly, encrypting email with PGP/GPG was perceived to be the most difficult practice to continue and spread, also due to issues around the complexity of installation as well as with key management. In contrast, participants told the team that when they introduced tools that did not require multiple installation steps, such as Jitsi Meet, these practices were quickly adopted for group communications and were integrated into workflows.

Former participants added that a strong understanding of conceptual elements underpinning digital security tools and ICT infrastructure was necessary in order to successfully advocate for and demonstrate the use of tools to peers and colleagues. Though web browser and social media account settings are characterised by some digital security guides and resources to be a 'basic' element of ICT use, the sheer number of settings and the frequency with which companies change policies relating to privacy and personal data were seen by former participants to be overwhelmingly confusing in aggregate. Former participants relayed that in order to convince people to change privacy settings in web browsers and social media platforms, they would need to understand the technical and political factors guiding their change. Though the digital security trainings in the Digital Security in Context study devoted much of the training time to developing a conceptual understanding of digital security and ICT infrastructure, participants desired more opportunity for follow-up discussion.

2 Supporting Security Integration Through Follow-up

While digital security training participants came away from trainings ready to advocate for the importance of digital security considerations with peers and colleagues, many had lingering concerns and questions regarding topics covered in the trainings, suggesting that participants would benefit from ongoing learning and engagement. To address a well-documented need for support beyond one-off trainings, one follow-up training was organised as a part of the Digital Security in Context study. In a feedback process conducted through discussion at the end of the follow-up training and in later surveys and engagements, former digital security training participants reported that the follow-up training helped to solidify skills and aided the process of knowledge transfer amongst their colleagues and friends.

In a discussion held at the conclusion of the follow-up training, participants reported that the follow-up provided an important chance to review concepts introduced in the first training: “For the second training I can listen more and ask more and make it clear for me”.[144] The training also provided an opportunity to learn new tools which may not yet have been available at the time of the first training, or which there was not enough time to cover in the first training. Additionally, the training provided a chance for participants to clarify questions about challenging, complicated practices and concepts: “How to use PGP has become clearer”.[145] The

[143] At the time of writing it was still possible to use OTR with Facebook

[144] Field Work Group discussion, 2014

[145] Field Work Group discussion, 2014

solidification of skills was tied by participants to an increased ability to contextualise the need for the use of certain digital security tools. Participants told the team that contextualising tool use through further practice and reflection allowed them to better prioritise security in the course of day-to-day work. "Now we know the concept and how it fits into our workflow, which means that now we know why and understand more about how the application works".[146]

For those participants who wanted to spread skills, skill developed and practised in the follow-up training increased their confidence to share their learnings with others: "This helps us to keep sharing what we have learned".[147] The second training allowed participants to focus more on how to communicate ideas: "In the second training I think about how to convince people. I focus on the way I make explanations".[148] After the second training, participants felt more confident that they would be able to share what they described as more difficult practices with others, such as encryption:

> *I feel that I'm able to share more advanced tools like PGP.*[149] *After our (first) training we did one session to introduce friends and colleagues to basic tools but we did not go into detail and we did not go into PGP. I think this time maybe we will expose them to PGP and call on them to practice it."*[150]

It should be noted that the follow-up training discussed in this report was done four months after an initial training, which was not part of this study. The team believes that future follow-up opportunities would be strongest if planned from the very beginning of the training process, as this would enable organisers and facilitators to better manage expectations, stagger the pace of learning, and develop a more dynamic agenda. Despite these caveats, conducting this follow-up provided a unique opportunity to evaluate outcomes from the first training, and showed itself to be highly beneficial to the team's and the participants' learning. Outcomes from the follow-up training lead the team to conclude that training programmes stand to benefit from building follow-up trainings into the overall structure of digital security learning interventions.

3 What the Integration of Practices Looks Like: HRD Perspectives

With many digital security practices perceived as being complex to implement, the involvement of peers colleagues was seen by training participants as an essential enabler of their successful integration within workflows. The perceived importance of advocating for and spreading digital security skills among peers and colleagues led the team to reserve between half a day and a day at the end of the five-day trainings conducted in the Digital Security in Context study to strategise effective advocacy and learning. Training participants were asked to highlight the most important new concepts and practices learned in the training and to brainstorm how these could be fostered within their organisations and networks. This discussion led to the creation of draft communication and information handling policies which were shared and discussed in training follow-ups with the team. After the conclusion of the three trainings done over the course of the study, five former participants and several colleagues were trained as advocates and trainers, continuing their engagement with digital

[146] Field Work Group discussion, 2014
[147] Field Work Group discussion, 2014
[148] Field Work Group discussion, 2014
[149] Field Work Group discussion, 2014
[150] Field Work Group discussion, 2014

security issues.

Pilot interviews with HRDs and later interviews with former training participants highlighted a number of stories of progressive engagement with digital security within human rights work. HRDs who now consider themselves trainers often began as training participants and were later trained as trainers, or went on to advocate informally around privacy and digital security through learning sessions with family, friends, community and colleagues. Some went on to engage in capacity building activities in their communities; to dedicate increased resources to digital security in their own organisations; and to increase the depth and scope of individual skills. These stories of engagement with digital security highlight the significant role that long-term support plays in the continuation of digital security practices and their incorporation within human rights work. The report highlights several of these stories below:

- A HRD and her organisation faced attacks which led her to create a collective security strategy together with other groups. After participating in a short awareness-raising workshop, she began to teach staff how to manage social media settings and make stronger passwords. Since her initial experience, she has continued to improve her skills and advocate for her local community to increase its capacity on digital security issues. Digital security practices have fed into broader security strategies implemented by the group in response to threats.

- A HRD believed at the start of one training that there was nothing she could do in response to increasing 'data integration' by companies like Google or governmental surveillance programs: “I don’t see any practical way around data integration”. After the training, the HRD's view was different. She expressed the motivation to be trained as a trainer in order to show others that practical measures are possible:

  “Technology develops every day, so what I worry about is the new thing that I don't know. But I feel more secure than before, because at least I trust myself to know what to click and what not to click on. I feel more confident now”.

- Over the last decade, a group who began their engagement with digital security as training participants watched their country acquire new surveillance capabilities. Over the course of these years, the group increased its focus on digital security, weaving 'practical' discussions throughout human rights and information policy projects and leading trainings of their own. The group's engagement with digital security aided them in reaching out to stakeholders outside of civil society, which created a broader base of support for their work. The group underlined that policy and practical work helped inform one another and strengthen the overall mission of the organisation.

- A former training participant-turned-trainer saw the effect of sustained support on his own digital security practices. After taking part in his first training as a participant, he didn't continue to use many of the tools introduced, but when he needed to serve as a training organiser some time later, it gave him a chance to review what he'd learned and motivated him to continue. “That was when I started using it (tools) consistently.” He observed the impact of sustained support and continued engagement on the continuation of practices:

  “You need to keep reinforcing it and exposing people to trainings, especially since things are changing rapidly”.

- Another former training participant-turned-trainer described his experience of aiding the process of integration of digital security practices among participants. The trainer felt that integration of practices was successful when HRDs were able to foster practices within their groups: "After a while, some of the staff started using the tool. They started using it together and now they have someone to turn to for help. We must understand the importance of support and try to keep giving it". The trainer emphasised that support must be sustained beyond the training space:

  "We think that after the training it's done, but it's not. It's just the beginning and we have to persist in that role".

Conclusion and Recommendations

This report introduced the practitioner research approach developed to conduct the Digital Security in Context study and situated Tactical Tech's questions around the efficacy of capacity building efforts within recent debates on privacy and digital security. The report then examined a shifting landscape of threats, technologies, and responses relative to human rights work in different contexts; looked at how privacy and digital security practices emerge; explored challenges with regards to common digital security practices; and demonstrated how privacy and digital security practices are integrated within human rights workflows.

The Digital Security in Context study found that digital security practices are shaped and restricted by quickly shifting threats, the dependencies of using commercial platforms, the challenging elements of FLOSS tool use, and a priority to communicate within groups of peers and colleagues. Though many digital security tools were seen as being challenging to incorporate within practices, HRDs worked hard to spread digital security practices and advocate for their importance among peers and colleagues.

Throughout the research process, the team saw that digital security trainings served as a crucial site for the articulation of digital security concerns, for building an understanding of digital security concepts, and for strategising how to integrate new digital security practices within workflows. As a place for discussion, trainings provided a space to corroborate security incidents and to strategise how to use ICTs and digital security tools in light of their critical shortcomings.

Trainings brought the social dynamics involved in the development of digital security practices to the fore. HRDs participating in trainings expressed that in addition to the challenges of a shifting landscape of threats and technologies, a crucial barrier to building effective digital security practices resulted from the fact that there was no one to practice or use digital security tools with outside the training. In instances where HRDs were able to spread digital security practices and establish collective security strategies among peers and colleagues, their work was strengthened and overall level of protection increased. Follow-up support was seen to strengthen the ability to share skills, build practices, and advocate for the importance of security. These findings suggest that the presence or absence of enabling social structures around the establishment of digital security practices is a critical determinant in their long-term integration within human rights work.

1 Improve Training Design

Prioritise relationships: strengthen existing relationships and foster new social connections

All HRDs engaged with in pilot interviews and trainings in the Digital Security in Context study consistently expressed the view that a key barrier to adoption of digital security practices is the fact that there is often no one to use digital security tools with. Participants in all groups complained that 'difficult tools' could only be used with other digital security training participants, and that the people that they most needed to use these tools with were often not present in the training: "If you have a security tool but they don't, you cannot communicate anything".[151] The fact that key people were not present in trainings meant that participants needed to serve as trainers and advocates in order to put new practices to use. Yet getting others to use these tools was a challenge due to their complexity: "It's a bit hard for me to advocate others to use it because with encryption you need time".[152] Training organisers and facilitators might consider how to address these challenges within the training planning process. While a number of trainings and workshops conducted in the field still bring together individuals with no prior relationships, prioritising the selection of participants who already have an expressed need to communicate with one another could facilitate more effective reinforcement of practices learned in trainings.

Prioritise sustained learning

HRDs engaged with over the course of the Digital Security in Context study emphasised the fact that there was a need for more and different opportunities to integrate digital security practices within the workflows of groups, organisations, networks and movements. While a first training was seen to serve as an effective introduction to digital security concepts and practices, a second, training or other follow-up events, provided the space and time to solidify skills, to strategise security within and among groups, and to support the development of security advocates. The actual forms of these follow-up events can be discussed with digital security trainers who would help plot HRDs' learning paths.

In a discussion held at the conclusion of the follow-up training taking place during the Digital Security in Context study, participants reported that the follow-up training provided an important chance to review concepts introduced in the first training. Participants said that they began to understand the context behind the tools better the second time around. The follow-up training also provided an opportunity to learn new tools which may not yet have been available at the time of the first training, or which there was not enough time to cover in the first training. The follow-up training also helped clarify questions about the use of challenging tools such as GPG. For those who wanted to spread digital security skills, the honing of skills made possible through the follow-up training increased the confidence to share skills with others, with participants saying that they felt more confident that they would be able to share what they had learned after the second training.

The team noted the importance of this planning 'follow-up' from the beginning stages of work with participants. In the follow-up training, which was organised after the conclusion of the

[151] Field Work Interview with anonymous HRD Cite B #4, 2014

[152] Field Work Interview with anonymous HRD Cite B #10, 2014

first training, the team saw a set of diverging interests emerge due to the pressure to make the most out of an unexpected opportunity. Some participants wanted more training on tools, while other participants wanted to be trained as trainers. Participants worried that these goals were incompatible. The expressed desire to learn how to teach their peers and colleagues among participants with varying degrees of experience points to a need to foster different levels of training and incorporating some 'training of trainer' (ToT) elements, acknowledging the role of advocates, skill sharing and the need for a longer-term planning process.

Support advocates working to spread digital security practices in their communities

During the time-frame of the Digital Security in Context study, several HRDs who began as training participants were trained as trainers or began to conduct informal workshops and trainings of their own, acting as community advocates and sharing the skills they had learned in the trainings. After one training,[153] four out of 15 former participants self-organised workshops and skill shares with a focus on password practices and social media with colleagues and friends. One former training participant and student went on to train 30 fellow students at a local university. A development worker said he incorporated digital security into his usual repertoire in trainings focused on using social media effectively. Two former participants provided trainings to staff in their non-governmental organisation on passwords, along with 'email security' and 'Facebook security.' It was clear that advocacy for digital security done by former participants was likely to have a real impact upon the continuation of digital security practices and development of collective security strategies. The team thus believes it is crucial to address the importance of nurturing those individuals who advocate for digital security in their communities and networks.

Discussion with training participants highlighted that an important element in supporting advocates of digital security pertains to how the role of an 'advocate' is conceived. Participants reported that the trainer title carries particular expectations which may not align with the need to spread skills among people with different areas of expertise; this is especially risky when related to highly technical topics or the potential to spread incorrect or dangerous advice. The reality of how individuals now professionally known as 'trainers' graduated to such roles from having been participants supports the need to address the sometimes binary distinction made between 'trainers' and 'participants', adding another definition for those individuals who, after being trained, begin to raise awareness and share skills in their own communities. One former participant suggested a rethinking of the professional 'trainer' title to something more inclusive of people who juggle many different kinds of roles in their professional work.

Explore inclusive strategies for evaluation

Because situations change so quickly and security practices arise to meet emergent priorities, it can be very challenging to establish a baseline upon which to evaluate the effectiveness of trainings. As shown in the case entitled 'Security Emerges Through the Collective', the benefits of trainings come into play over a long period of time after trainings, and 'indicators' by which to judge the efficacy of digital security practices are likely to change as the goals of groups evolve. These challenges indicate that where changes to evaluation methods are needed:

[153] This training occurred before the start of the study. The researcher was present to interview participants afterwards; and for a subsequent follow-up training.

- Any activity tracking the efficacy of practices should be guided by indicators dictated by participants.
- Any data collected on participant practices should be done through carefully delineated guidelines and protocols prioritising the privacy and security of the information.
- In order to avoid the 'gratitude bias' at end of a training and to aid in the overall tailoring and contextualisation of the training approach, facilitators and trainers might work to evaluate outcomes through participatory activities woven throughout the training rather than reserved for the end.
- As demonstrated in this study's findings, the beneficial practices and strategies learned in digital security trainings became apparent long after their initial conclusion. Funders and organisations planning evaluation programmes should develop a strategy that allows the long-term benefits of capacity building work to be accounted for.

2 Customise Tool Development and Adoption

Contextualise privacy and digital security concepts

In dealing with the concerns arising within challenging contexts, HRDs in this study needed to understand digital security issues through a predominantly English language-based lexicon, which added to the overall challenges of creating constructive and appropriate digital security strategies. This language issue became an issue both in tool use and in spaces for learning and discussion, such as trainings. Two groups worked with were not native English speakers, but were only able to access English-language versions of the tools. A lack of resources in the local language resulted in HRDs struggling to understand installation instructions in tools. Though Security in-a-Box offers translated instructions in 16 languages, many gaps remain within tools.

Many participants and trainers the team spoke to noted that language translation doesn't guarantee that meanings are culturally legible. One translator told the team that with regards to localisation efforts aiming to translate tools, “the challenge is not just translating, because there are certain words where there is no one-to-one meaning”. In describing the difference between a mere translation and a meaningful contextualisation or interpretation of digital security content, participants and trainers who had encountered language issues told the team that prioritising local, contextual meanings should involve exploring culturally relevant metaphors.

Foster tool vetting mechanisms to keep up with a shifting landscape

The ever-shifting nature of threats and technologies means that HRDs who want to build strong digital security practices have to keep up with an enormous number of confusing developments with regards to the relative merits of different tools. In making decisions about which communications and digital security tools to use, participants noted that it was difficult to find good information regarding the relative privacy and security offered by different services and tools. Participants expressed the need for a vetting system to verify tool and platform privacy and security claims. A couple of projects have taken on this problem of late: the Open Integrity Index has been developing criteria to understand

the merits of tools and platforms more broadly,[154] and the Electronic Frontier Foundation has compiled a comparative chart on the relative merits of an array of chat applications.[155] The magnitude of the problem suggests a need to support and expand such initiatives within a wide variety of resources produced by different organisations.

The Digital Security in Context study began as an inquiry into the factors affecting the uptake of digital security practices by human rights defenders. Results indicate why and how training methodologies can be strengthened and deepened by being more responsive to the shifting landscapes of threats HRDs face in their local contexts. Yet, this may just be the start, and there may possibly be more creative responses to learning, and sharing, practices amongst HRD communities. Given the paucity of research in this area,Tactical Tech hopes that this is the start of approaches that integrate more reflective and evaluative practices into digital security training.

[154] https://openintegrity.org/en

[155] See EFF resource on chat applications https://www.eff.org/secure-messaging-scorecard

Appendices

Appendix 1 : Research methods

Context Analysis Activities

Semi-structured interviews

The researcher-trainer conducted semi-structured interviews covering' key concerns in relation to existing security practices. These interviews aided in the 'tailoring' of training and workshop agendas.

Participatory Actor Map

The participatory actor map is a variation of the actor map described in the first section of this report, under 'Practice-based Research Foundations'. The team asked each participant taking part in the activity to draw a picture representing the most important influences in their lives, including allies, neutral parties and potential adversaries and the interconnections between them.

'Draw the internet'

Participants were asked to draw how they imagine the internet. The framing was kept completely open ended. Instead of pointing to 'right' and 'wrong' answers, the team asked participants to highlight their personal experiences using the internet in their work. When all participants were finished, the drawings were placed in a gallery and discussed. The team used this discussion as the basis for an input module focusing on a communications infrastructure. The team would take cues and language from these drawings and feed them into a drawn diagram, filling in gaps in knowledge that surfaced through the discussion.

Participants drew clouds, globes, fishing nets and nodes. Some participants highlighted elements of ownership and control in regards to its effect on the freedom of users: "However how freely you use its still cut off by company".[156] One participant wanted to highlight the abstract notions of space and distance particular to the experience of using the internet: "I see the internet like a sky- it stretches across the world but at the same time it feels very far away. It's fascinating to look and very pretty and you can work on it forever but that doesn't mean you understand how it works". (see Fig. 1)

One participant used the cloud metaphor to draw attention to its unknown parts: "the dark clouds are spaces people don't know much about, like threats"(see Fig. 2) while another included threatening actors and structures in narrating their understanding of internet infrastructure*:* "I crossed out Google because it is one of my big concern areas". (see Fig. 3) A fourth participant described the internet as a tool which enables community building, but which also opens up new avenues of threats:

> *My work related to the internet is to help community people who are fighting for their rights… Seeking data from different networks and connecting them. But there are guys also watching and tracking us, which is my concern about the internet.* (see Fig. 4)

[156] Field Work group discussion, 2014

Information mapping

The information mapping exercise is designed to help participants understand where important or sensitive information is stored throughout the devices and services they use in their work, and may form the basis for the creation of data backup and storage strategies to keep important or sensitive information secure. Participants were asked to draw a grid. In this grid, participants wrote all of the locations and devices where they stored information. Participants were asked to brainstorm all of the 'pieces of information' stored in these different locations and devices and to separate them according to their relative sensitivity and value. This activity was done in self-selected groups.

The scenario below traces some of the different elements of information handling within a specific workflow. Questions stemming from this scenario aid in the creation of a roadmap for an effective, appropriate digital security strategy.

Workflow Scenario: Planning, collecting, storing, and sharing information before, during and after a trip to collect human rights testimonies:

1. The researcher plans site visits via email and mobile phone.
2. The researcher travels with a recorder, computer and notebook to collect testimonies.
3. The researcher records interviews with their digital recording device and takes notes in their notebook.
4. The researcher transcribes the interviews in a word document on their computer.
5. The researcher saves the file to a shared folder hosted through a service called Dropbox.

Potential Considerations:

- The Dropbox folder is accessible to staff members in the organisation and a local copy of it is saved to each of their computers.
- Some staff are based in areas where rule of law is considered relatively strong, their work experiences a degree of acceptance and the organisation considers itself to be physically safe. Other staff believe themselves to be under threat of an office break-in due to a volatile political environment.
- The folder is potentially subject to mass surveillance by companies and governments conducting wide sweeps of data collected and stored through cloud services such as Dropbox.

This scenario highlights the variety of potential concerns that may surface through one particular workflow. Parsing through these concerns is an important element in the process of devising an effective, contextual digital security strategy.

Questions arising from this scenario:

- How sensitive is the information stored in each device/location?
- Does a spare copy of this information exist in a back-up location?
- How might the loss or unintended exposure of sensitive field notes or raw audio files

through theft or device damage impact the work of the organisation? For example, what if one individual staff member accidentally deletes the files from the organisational Dropbox account? What if this information makes its way into the hands of opponents or antagonistic actors?

ASSESSMENT ACTIVITIES

Pre-assessment survey

The team used a four-part pre-assessment surveys assessing skills, learning styles, and participant priorities as previously developed by the lead trainer.

In-training assessment

Several exercises allowed participants to flag issues, concerns and takeaways at the end of every day or at the mid-points of trainings. Below are two examples of exercises in this vein.

- Flower Petal

  The flower petal is a daily evaluation activity in which participants cut out flower petals from construction paper and assemble a flower over the course of five days. Each petal is added at the conclusion of each day, with reflections from participants on the activities of the past day. The flower petal activity was adapted from an activity previously developed by a peer organisation.

- Plus/Delta

  Training participants chose post-it notes in two different colours. One colour is used designated for 'pluses'– good, useful, inspiring elements of the training. The second colour is for 'Deltas' – frustrations or unmet priorities. In the study, these post-its were anonymously placed on a wall and grouped by participants into relevant categories. The trainer then used these to modify the training agenda. This activity was done at the conclusion of each day. The Plus/Delta exercise is common in workshop facilitation methods and has been used by Tactical Tech trainers for several years.

- Paper Spectrogram

  The paper spectrogram is a variation on a physical and interactive spectrogram activity where people move around the room and pick points along a physical Likert scale[157] to express varying levels of agreement or disagreement with statements made by a facilitator. The facilitator might state that 'cats are better than dogs'. In order to indicate agreement or disagreement with such a statement, people move to position themselves physically along the designated scale. Usually this activity is followed by a group discussion about the views exchanged by participants in response to the facilitator's statement, or participants propose their own statements. The paper spectrogram – used in two of the trainings in this study – instead places the Likert scale within the confines of a piece of paper, and participants anonymously plot points on the scale with markers or stickers.[158]

[157] https://en.wikipedia.org/wiki/Likert_scale

Appendix 2 : Operationalising Security and Privacy in Research

In the course of this study, the following guidelines were used to ensure minimal communication traces during pre-planning of interviews and trainings.

- Prior to trainings, contact was maintained with one main point person, in order to avoid exposing the social graph of the community. This point person organised interviews and spread information about the training within their organisations, networks, and communities, rather than publicly advertising it.
- Discussions with organisers and participants were carried out using VOIP Skype alternatives.
- Details about the training and research context were kept confidential to the team and a small circle of colleagues.

During trainings, the following guidelines were followed to establish a 'safe space':

- Participants were asked not to take photographs of other participants or training materials without first obtaining consent.
- The lead trainer, researcher-trainer, organisers, and participants agreed to keep the identities of participants confidential to the training group.

Follow-up was conducted with consideration of the relative security of communications:

- The team used VOIP alternatives to Skype and encrypted lines of communication, where appropriate.
- The team stored documents containing identifying information from the field work in encrypted folders with backed-up encrypted versions stored at a separate location.
- Interviews were anonymised upon transcription.
- During the report writing process, the team stripped content of geographic detail and other unique identifiers.
- The team tested the strength of these de-identification efforts with trusted readers outside the immediate research group.

[158] An explanation of the spectrogram can be found in Aspiration Tech's facilitation resource: http://facilitation.aspirationtech.org/index.php?title=Facilitation:Spectrogram