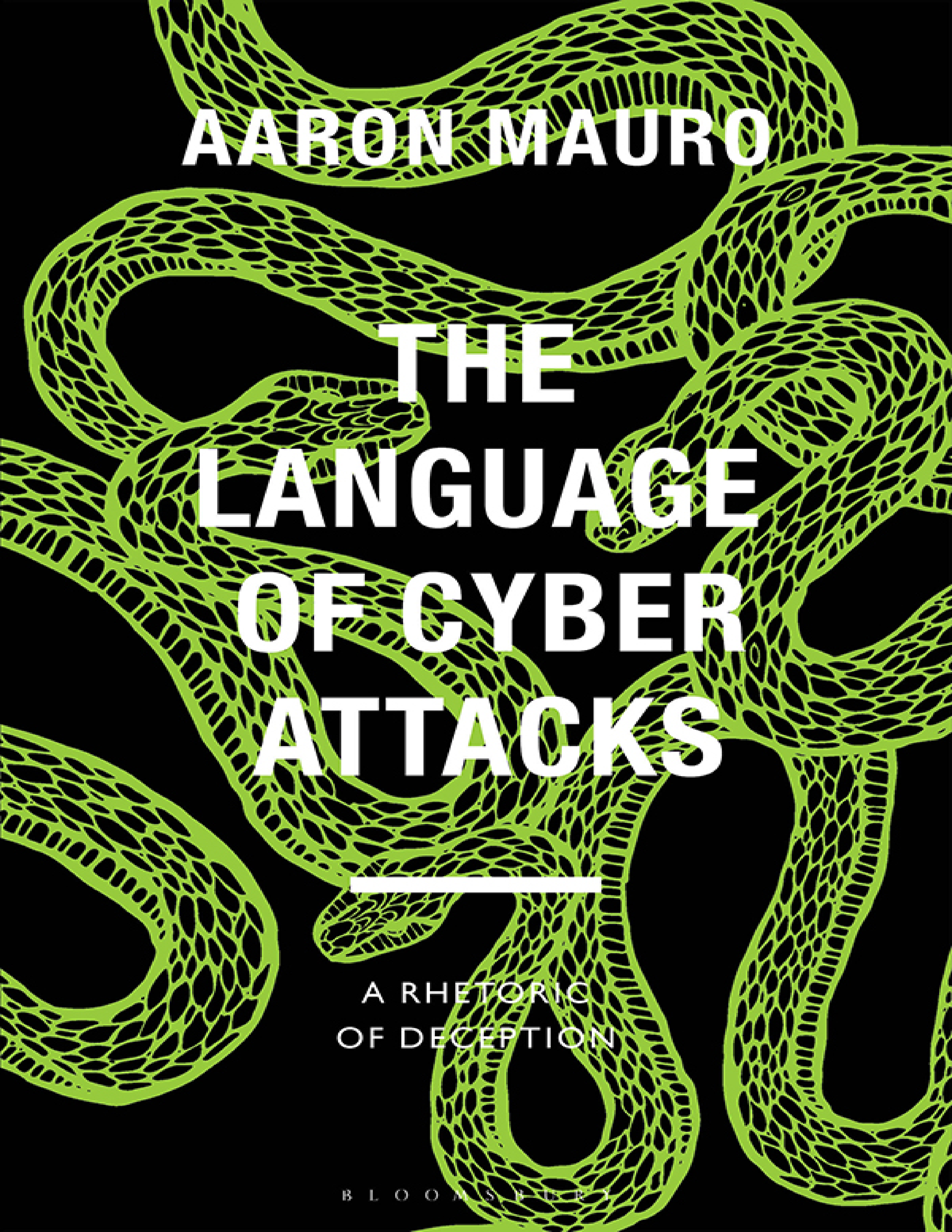
AARON MAURO

THE LANGUAGE OF CYBER ATTACKS

A RHETORIC OF DECEPTION



BLOOMSBURY

THE LANGUAGE OF CYBER ATTACKS

BLOOMSBURY STUDIES IN DIGITAL CULTURES

Series Editors

Anthony Mandal and Jenny Kidd

This series responds to a rapidly changing digital world, one which permeates both our everyday lives and the broader philosophical challenges that accrue in its wake. It is inter- and trans-disciplinary, situated at the meeting points of the digital humanities, digital media and cultural studies, and research into digital ethics.

While the series will tackle the “digital humanities” in its broadest sense, its ambition is to broaden focus beyond areas typically associated with the digital humanities to encompass a range of approaches to the digital, whether these be digital humanities, digital media studies, or digital arts practice.

Titles in the series

The Trouble with Big Data, Jennifer Edmond, Nicola Horsley, Jörg Lehmann and Mike Priddy

Queer Data, Kevin Guyan

Hacking in the Humanities, Aaron Mauro

Investigating Google’s Search Engine, Rosie Graham

Reading Audio Readers, Karl Berglund

Forthcoming titles

Ambient Stories in Practice and Research, Edited by Amy Spencer

Metamodernism and the Postdigital in the Contemporary Novel, Spencer Jordan

Herman Melville and the Digital Humanities, Christopher Ohge and Dennis Mischke

Listening In, Toby Heys, David Jackson and Marsha Courneya

People Like You, Sophie Day, Celia Lury and Helen Ward

For my students

THE LANGUAGE OF CYBER ATTACKS

A Rhetoric of Deception

AARON MAURO

BLOOMSBURY ACADEMIC
LONDON • NEW YORK • OXFORD • NEW DELHI • SYDNEY

CONTENTS

ACKNOWLEDGMENTS

The land that made this book possible is the traditional territory of the Haudenosaunee and Anishinaabe peoples. This land is covered by the Upper Canada Treaties and is within the land protected by the Dish with One Spoon Wampum Agreement. The friendship and generosity of Indigenous people is the foundation that makes possible our security now and in the future.

Inspiration for this book came from so many of my collaborators, colleagues, and friends. Thanks to my dear friends who have checked in on me and encouraged me through the isolation of writing, even through the summer months.

I want to express my gratitude for my whole extended chosen family. Thank you, Shelia and Gurney, for your unflinching support. The little afternoons you gave me were precious, especially knowing that everyone was fed, entertained, and happy. Thank you, John, for all your excellent book recommendations over the years.

I want to offer my gratitude for my colleagues in the Department of Digital Humanities at Brock University. I am grateful for your support and patience in the final months of writing. I could not have taken on the Department Chair at this time without your experience and support in learning the ropes.

Thank you, Clara Suba, for answering so many questions and knowing exactly what to do when I mostly did not. Thanks to Jeremy Leipert, for tirelessly building the relationships with the communities and partners that allow our students to experience the full value of their education. Thanks also to Tom Brown, for immediately knowing answers to technical questions I sometimes struggle to even ask.

Thanks to Jason Hawerliak, Alex Christie, Sarah Stang, and David Hutchison for your trust in me. Your passion and enthusiasm for our students is infectious. It is a privilege to work alongside you.

Thanks to Neta Gordon and Ebru Ustundag for being such inspiring collaborators. It has been a joy to imagine spaces and places that define literary geography.

I want to thank Renata Queiroz Dividino, Ali Emami, Heather Ramey for tackling the thorny problems of online toxicity in all the ways that we need right now. Working with you has reminded me that interdisciplinary collaboration is built on professionalism.

Alison Innes, thank you for so many brilliant conversations. I look forward to many more.

I want to thank Kevin Magee for reading the entire manuscript. Your encouragement and mentorship have meant the world to me this year. I am grateful for the depth of your industry experience, where cybersecurity can draw real insight from the study of the humanities.

Thanks to Laura Horak for allowing me to be a guest on your team. Your work reminds me why security matters.

Thanks to Ray Siemens and the whole INKE Partnership group. Thank you for giving me an opportunity to return to my academic home.

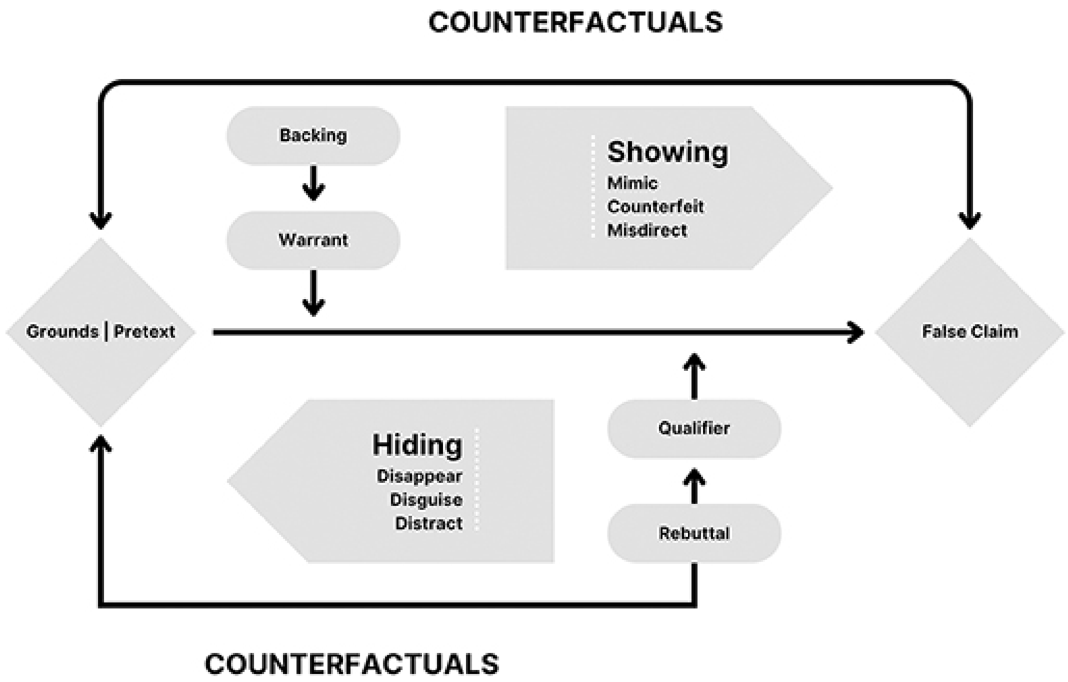
Zameena Hussien, for helping me integrate my past and open windows to the future.

My publisher, Ben Doyle at Bloomsbury, continues to believe in me and in this project. Thank you for being my champion. Thank you, Laura Cope at Bloomsbury, for keeping me on track. Thank you for the confidence that comes with having your support throughout this process.

I want to thank my son for his endlessly surprising imagination and exuberant sense of humour. Your creativity is an inspiration to me, in *every conceivable way*.

To my wife and partner in life, whose support goes beyond anything I could say here. We’ve always grown where we’ve been planted and transplanted over the years. Through it all, I’m so grateful to have found my companion plant, like lavender and roses, that makes life more beautiful by being together.

Last, but not least, I would like to acknowledge two animal friends, Pete and Sparkle, for being my constant companions in this project. These animals have taught me so much more about honesty and being a human than many humans.



Deceptive Security Rhetoric
This diagram depicts the function of "counterfactuals" in a deceitful argumentative structure. The "Grounds | Pretext" for a lure are augmented by selectively "Showing" and "Hiding" information toward a "False Claim" or toward a redefinition of the Grounds | Pretext." The use of "Backing," "Warrant," "Mimic," and "Disguise" collectively describe the dynamics of argument manipulation that recursively uses misdirection and concealment to lead to accepting a "False Claim."

1

Cybersecurity as an Expression of Values

This book describes the language of a cyberattack; it describes the rhetoric of the lures and deceptive tactics that initiate many attacks. The first link in an attack often appears as just another unassuming email in your inbox.

Cyberattacks often begin with a lure. These lures are messages designed to ensnare and capture attention and provoke a compromising action. A phishing lure offers the target a choice, but the "lure of choice" is difficult to resist.[1] There is a risk in taking an action online; the consequences of any interaction must be adequately understood on a technical level, but threat-informed defense increasingly requires a broader cultural awareness and emotional intelligence to identify malicious behavior. In most cases, anonymity enables hostility with impunity. Any little missive might be a missile, but it is a phishing lure that will grant initial access and deliver a malware payload. These pages make the case for a more holistic understanding of the affective, linguistic, and rhetorical roots of many cyberattacks. Perhaps more audaciously, a greater understanding of this complex rhetorical situation may allow users to anticipate threats, mitigate future risks, and build the case for attributing attacks.

In doing so, my aim is to bridge between scholarly and security analysis. My intention is to read security-related events, along with all the data leaks and reporting, to sketch the cultural significance of cybersecurity. This interdisciplinary approach is not without risks of its own. Despite the possibility of falling between disparate disciplines and professions, I remain committed to the increased security that analysts and humanists alike can contribute by interpreting “the language of cyberattacks.” To formalize this humanistic approach, a return to the discipline of rhetoric will be an invaluable resource to make sense of the all too common moments of deception, lies, and manipulation.

This approach combines the linguistic, procedural, affective, and cultural valences of cyberattacks to produce a systematic, interpretive methodology that is better capable of identifying attackers. On the one hand, I make the case for an applied, technical form of humanist cultural analysis, while also making the case for a more philosophically minded, ethically sensitive, and speculatively imaginative form of cybersecurity. The cultural consequences of cybersecurity touch on every aspect of our ubiquitous online lives, wherein a digital proxy of our lived experience resides. If individuals are to protect their communities, histories, stories, and chosen systems of governance, a cultural basis for security literacy must emerge from every aspect of society, and the cybersecurity industry must also work to accurately reflect those values.

Imagine an emotionally intelligent cybersecurity practice adopted throughout society, transforming the culture of work, education, and government. A cultural project on such a scale would demand a societal consensus about significance of cybersecurity for individual citizens. By understanding security at the intersections of broad categories of justice and democracy, security questions animate how countries regulate power and keep peace, how economies secure greater equity and open opportunity, and how global partners manage toxicity in digital and ecological environments.

The humanities might best be applied as a sensible reader of these overlapping territories, developing a new security literacy that looks a lot like literacies of the previous millennium. Close reading and attention paid to understanding things like motive, character, setting, unreliable narratives, incomplete records, and a deep suspicion of the capacity of language to sustain a durable sense of truth, these are just some of the wicked problems faced by humanists in making claims about our shared cultural histories, the evolution of ideas, and the mysteries of human expression.

Cybersecurity has similarly wicked problems. Most significantly among them, the attribution of attackers’ identities. Attribution remains the most dangerously uncertain aspect of incident response and reporting in cybersecurity. The false attribution of an attack, particularly between nation-states, could lead to disastrous consequences for international relations, trade, and global peace. The *attribution problem* also represents the greatest opportunity for humanists to contribute to a truly global project of safety and security online and is the subject of the second chapter of this book. The study of the language of cyberattacks would scour leaks, data dumps, and log files for traces of evidence that might account for criminal motives, political characters, historical setting, corrupted or fallacious information, while tracking the attacker's twists and manipulations.

* * *

So, to begin again, many cyberattacks start with a lure, a message designed to establish trust with a target and compel them to divulge sensitive information or compromise a computer system. These malicious messages are commonly experienced through phishing emails, but a lure can come as an SMS, a phone call, or even an impostor at your door. An attack chain might begin with open-source intelligence (OSINT) gathering on a target. Crafting a carefully worded message is the next stage, paying close attention to what features might trigger a response. These campaigns can be conducted against a group or an individual, or, in the case of a spear phishing campaign, those affiliated with a high-value target institution.

A successful attack often entices the user to click a link or download a malicious payload. This initial compromise may provide enough access to sufficiently escalate privileges. The main attack would come later. The first link in an attack chain is often rooted in deceptive communication tactics; this exploration of the "lure" aims to contribute to education initiatives in defensive cybersecurity. Addressing the human factor of cybersecurity cannot require more than just raising awareness about security issues. The goal of *security rhetoric* is to foster a culture that balances emotional maturity with a realistic understanding of security risks. For effective cybersecurity, a culture of sensible, secure, and safe online behavior must be cultivated at the societal level.

A key concern in cybersecurity is the human user's fallibility and their ability to distinguish authentic messages from deceptive ones. While the so-called human factor in security vulnerabilities is often discussed in cybersecurity, the humanity of the user is often not. Human fallibility and vulnerability are the very things that make us human and good targets. The way we feel, the stories we hold dear, the relationships we maintain are key elements of a security subdomain known as *social engineering*. Usually in the form of a phishing email, social engineering attacks prey on our very human vulnerability to feelings of fear, urgency, authority, and relationships that can be manipulated or be used to curry influence. Anyone can fall for these scams.[2] Users may inadvertently click or download malicious content that leads to serious security compromise at work, at home, and as professional and personal boundaries increasingly blur.

Currently, mitigation strategies are often framed from the attacker's perspective, rather than that of the potential target. The prevailing assumption is that potential targets can identify threats by understanding the attacker's tactics and techniques. This book posits that effectively identifying malicious intent involves a nuanced understanding of style, aesthetics, and emotional cues—elements often overlooked in traditional cybersecurity discourse. Issues of digital literacy intersect with more literary and rhetorical skills of close reading and critical analysis.

This book expands the discussion on the current scope of "cybersecurity awareness" and the training received by employees, students, and the public.[3] The growing focus on security training underscores the role of user education in enhancing overall internet security.[4] New regulations are increasingly mandating that banks, ISPs, and online platforms take responsibility for curbing illegal activities on their networks. Governments are actively taking measures to protect citizens and deter global threat actors from targeting institutions and individuals. The user remains the most vulnerable and unpredictable variable in technical systems. For this reason, interdisciplinary humanities is well positioned to bolster security efforts by fostering a more comprehensive security culture. Having multiple ways to look at a problem helps identify unanticipated problems by observing with a different lens from a new angle. Users are often influenced by emotional and social factors, which are well examined in the humanities through disciplines like philosophy, oral history, gender studies, and literary theory, in an intersectional context.[5]

The ease of a button click often belies the real-world consequences, making the temptation to click an invitation to tangible action.[6] There is a secret, unsaid attraction to being lured. There is, lurking within us all, a desire to see what all the fuss is about. Thoughts and feelings that are glossed over—feelings like curiosity, laziness, fear, and inattention—are explored here as the foundation for a new framework of understanding rhetoric and language in cybersecurity.

Surprisingly, perhaps, this book delves into the emotional labor required to navigate the pervasive deception in our digital communications. There is an emotional toll akin to chronic stress in determining authenticity from deceit in every single push-notification. The incessant demand on our attention lends urgency to the need to find greater balance and boundaries. The emotional tenor of these interpretive moments, wherein we must choose how to interact with technology, will only have a greater impact on our private and public lives over time. In this way, effective security extends to safeguarding our emotional and sensory experiences, paving the way for more resilient and sensible cybersecurity practices.

Closely examining how language influences us is crucial for safeguarding the vulnerable human element in technical systems. There has been some work in thinking about rhetoric and digital technology, which finds itself increasingly between disciplinary homes in the Digital Humanities (DH) and the social sciences, including psychology, criminology, and political science. Alexander Reid issues a call for a kind of "speculative rhetoric" that acknowledges how rhetoric and language have always been working as technology. Language itself can be viewed as a form of technology, augmented by writing tools, thereby introducing non-human actors into the realm of human persuasion. Reid asks for a broadening of how we consider rhetoric today: "what is required is a rethinking of the humanities that accounts for technology and rhetoric in a new way."[7] *Security rhetoric* seeks to explore the interplay of technology and rhetoric, particularly as they manifest in phishing lures—the initial point of compromise in many cyberattacks. Given that specific techniques can quickly become outdated and lose relevance, my focus will be on articulating the underlying thought structures and logical systems that govern cyberattacks.

As a practical and applied contribution to cybersecurity, it is necessary to discuss the *cultural rhetorics* of both the humanities and the cybersecurity industry: the former is an academic discipline as well as a cultural project shared by broad publics, while the latter is a professional, technically oriented industry as well as a practical and applied project shared by citizens, academics, corporations, and various levels of government.

Rhetoric, as the art of persuasion, finds ample application in both academic and professional realms.

Lev Manovich, in his 2002 work *The Language of New Media*, once celebrated what he saw as the “continuing decline” of rhetoric in the rapidly evolving field of new media.[8] After all, to have one’s words described as “mere rhetoric” would be to assume one’s words are empty, hollow, and manipulative. It may appear as a strange choice to turn to rhetoric as a potential antidote to so much deceit and misinformation online. I would like to align my work with key rhetoricians like Jennifer Sano-Franchini, Elizabeth Losh, and Ian Bogost, who are interested in the tactical and practical benefits of working closely with DH. This project can be viewed as an exploration of the cultural rhetorics—defined by Sano-Franchini—of twenty-first-century cybersecurity.

The concept of cultural rhetorics emerges from an earlier discourse on “critical rhetoric” defined by Raymie E. McKerrow in the late 1980s, which sought to draw lines between language and power. McKerrow sought to understand the role of rhetoric in both expressing and sustaining power. He argued that “the discourse of power focuses on the ‘normalization’ of language intended to maintain the status quo.”[9] The proliferation of deceptive online communications, be it cybercriminals phishing for credentials or nation-state disinformation campaigns, serves to maintain the status quo by leveraging appeals to power and authority for manipulation and influence.[10]

In contrast, cultural rhetorics examine how digital cultures and communications are performed in this deeply interconnected way online. Cultural rhetorics aim to articulate a plurality of rhetorical practices. No single university discipline can fully capture the multi-modal complexities of online communication or adequately describe the rhetoric of cyberattacks.[11] The rhetoric of cyberattacks is not static or hard coded. Technical aspects such as platforms, Internet Service Providers (ISPs), and regulatory contexts intersect with discursive elements like philosophy, identity, history, and politics.

Rather than looking at how powerful agents function online to persuade the masses, Sano-Franchini paves the way for a more expansive understanding of cultural rhetoric in digital spaces, which are both heterogeneous and horizontally structured: “cultural rhetorics theorizes how rhetoric and culture are interconnected through a focus on the processes by which language, text, and other discursive practices like performance, embodiment, and materiality create meaning."[12] Rhetoric does not stand outside culture as a dispassionate viewer. Studying rhetorical situations is an integral aspect of cultural participation. In this way, the rise of concerns related to cybersecurity aligns the terms "security" and "rhetoric" by the sheer force of so much public attention.

The term "cultural rhetorics" is useful in the study of deceit and disinformation online, for Sano-Franchini, "because diverse fields of study offer important insights about the relation between culture and knowledge."[13] Cultures themselves offer a heuristic for determining the quality of knowledge that emerges. Understanding the cultural and political context enables analysts to discern what is accepted as truth. Sifting through falsehoods allows us to discover authentic online experiences, fostering cultures of belonging and building bridges between diverse online identities. The emergence of a broader security culture could pave the way for more mindful digital interactions, beyond the frenetic and often toxic atmosphere of social media platforms. Alongside a renewed appreciation for traditional books, intercultural understanding and communication can flourish when misinformation is relegated to the background.

The scale of misinformation and disinformation online requires a broad attention to cultural rhetorics as a kind of literacy that "theorizes collaboration and its implications through the understanding that rhetoric is constellated and built through webs of relationality."[14] A shared sense of security serves as a key binding element for communities within this online web of relationships. Cybersecurity transcends its role as just a technical feature within institutions, corporations, and governments. Online security is integral to the cultural vitality of "networks of people and that live an active, 24-7 life online."[15] A variety of interconnected issues are converging around the need for enhanced rhetorical fluency, which can provide a flexible and robust heuristic for understanding manipulation and influence. In the words of Sano-Franchini, whose work I am extending into the realm of security:

> Some of the key goals of cultural rhetorics scholarship include exposing and disrupting dominant narratives, particularly those that do damage to historically marginalized cultures; building bridges, making connections, and coalitioning for sociopolitical change through teaching, language, and playing with the notion of academic discourse; making space for the work and voices of groups who have traditionally been silenced; and doing the intellectual work of

> renaming, reconceptualizing, and continually resituating the kind of work that rhetoricians can and should do.[16]

Creating space also involves navigating large, often corporate-owned, platforms. Few critics are exploring rhetoric in this way. Both Ian Bogost and Elizabeth Losh have contributed to the new media spaces by refining our understanding of how platforms and infrastructure shape and define rhetorical situations. In the context of gaming, Ian Bogost highlights how platforms and technology work on users in a procedural way: “A theory of procedural rhetoric is needed to make commensurate judgements about the software systems we encounter every day and to allow a more sophisticated procedural authorship with both persuasion and expression as its goal.”[17] Bogost is suggesting that the processes that support so much digital media, including games, work persuasively. Procedures embed ideology and can influence users in unexpected ways by “authoring arguments through processes.”[18] The interplay between infrastructure, hardware, and software creates persuasive arguments that are often implicitly felt. Procedural rhetoric equips individuals to identify, understand, and construct arguments within these technical systems. For Bogost, gaming is a key example of this kind of technologically embedded procedural rhetoric, reliant on elements like gaming consoles and developers.

In her latest book *Selfie Democracy,* Losh chronicles how US politicians use “fantasies of digital connection, access, and participation” to incentivize citizens to participate in democracy with their attention and their votes.[19] Losh identifies a consistent rhetorical threat across the administrations of three recent US presidents: Obama, Trump, and Biden. In doing so, Losh describes how rhetoric works in explicit communication as well as the implicit infrastructures necessary to enable our contemporary communications environment:

> In addition to more traditional channels of communication, rhetoric also operates through algorithms, database structures, and interface design. We live in a culture inundated with obvious messages, but the operations that order, filter, and curate these messages for us are often invisible. Understanding these less explicit techniques of influence and expression is an important part of analyzing contemporary political rhetoric.[20]

The rhetorical patterns Losh identifies across successive administrations are a striking portrait of American political discourse.[21] The categories of “connection,” “transparency,” “participation,” and “access” are as novel as they are familiar to contemporary observers; these themes manifest through procedural rhetorics embedded in “algorithms, database structures, and interface design.”[22]

As an attempt to make sense of how technical systems and multimedia communications work together in a persuasive way, scholars have tried to encapsulate this in the overarching term “digital rhetoric.” Adding a “digital” modifier to rhetoric oversimplifies the complex interplay between infrastructure and content, offering little insight into the phenomena in question. After all, cybersecurity is not just a problem of the internet.

There are security challenges in the way computation is laced throughout our environment, in our phones, our refrigerators, our cars, and our homes. The fabric of our lives, online or not, is a security problem. Of course, larger systems like power plants, water treatment facilities, and mass transit systems are not subject to the rapid replacement of capitalistic planned obsolesce, wherein new products are purchased every few years. *Security rhetoric* serves as the common consensus that animates the need to repair and patch consumer technology while maintaining and updating our longer-lived infrastructure. According to Bogost, the term “digital rhetoric” is unhelpful because it “abstracts the computer as a consideration, focusing on the text and image content a machine might host and the communities of practice in which that content is created and used.”[23] In this way, articulating design pillars and categories that can be broadly applied is more useful than defining new terms to collect theories and merely coining jargon.

The solutions to this problem will start with simply noticing. A refinement of human attention may be the single most sophisticated defense against cyberattacks predicated on deception.

Security Rhetoric

In all this, I propose revisiting the ancient Greek rhetorical concept of “Kairos” through the lens of contemporary cybersecurity. I am not alone in this call for noticing. The late, great David Graeber’s final book *The Dawn of Everything*, co-authored with David Wengrow, sensibly proposes a new history of humanity by looking through and beyond the twinned cultural legacies of colonialism and capitalism that function through ecological destruction. In closing this radical reappraisal of pugnacious “Western” habits, they expose their methodology through this informed moment of noticing:

> We begin this book with a quote which refers to the Greek notion of *kairos* as one of those occasional moments in a society’s history when its frames of reference undergo a shift—a metamorphosis of the fundamental principles and symbols, when the lines between myth and history, science and magic become blurred—and, therefore, real change is possible.[24]

The scale of the project I propose must carry on from this moment wherein change is possible. I argue that *kairos* holds special significance for cybersecurity professionals focused on human vulnerabilities. As Aristotle articulated, *kairos* refers to the precise moment of critical insight that compels action. Rhetoric is as relevant as ever as a reminder of the need to pay attention to how we pay attention. While rhetoric is the art of persuasion, it also involves recognizing and understanding the moments when we are being persuaded.

Defensive cybersecurity is about this kind of awareness and how it is tied to reasoned actions. There is a moment of awareness in cybersecurity awareness training that seeks to show the user how to make the right decision. It may be helpful at this moment to return to that first description of critical rhetoric, wherein McKerrow reminds us that “*influence* is not *causality*.”[25] Simply because someone has influence over us, does not mean that we will do as instructed. Rhetoric is a discipline with a keen attention to agency and the very simple human capacity to choose. The burden of honing this refined attention is not just for individuals. Informed choice is important, but it must also be supported by a cultural consensus that incentivizes critical thought.

There are, after all, so many forces at work on us, especially in contemporary media environments. By integrating a humanistic grasp of rhetoric and language with insights from cybersecurity and social engineering, the fields of user experience and psychology can further enrich this humanistic perspective, bolstering defenses against cyberattacks. This approach also responds, in part, to the call for action I issued in my 2022 book, *Hacking in the Humanities*. I advocated for an activist “gray hat humanities” approach that informs cybersecurity policies and trains future digital citizens.[26] Because of the scale of the challenge of digital literacy and the pace of change, I describe how computational approaches help identify broader stylistic trends and habits of hackers and scammers.

There are several consequences of taking up a security-first approach to scholarship and research, which I describe at length in *Hacking in the Humanities*. Risk assessments can inform project-specific research security policies and guide the behaviors of researchers and research participants much more than is likely the case now. Researchers must begin by threat modeling the hostile online environment into which digital projects and data are contributed. The humanities must also embrace a level of professionalism and technical expertise in building technical systems because the cost of failure could lead to the destruction, poisoning, or disruption of vital, often unique, cultural records. Similarly, the humanities have a long legacy in civil disobedience and activism that must make space for quasi-legal data collection and online protest. In other words, humanists must be able to make things and, when necessary, break things. If the humanities are to participate in developing a confident and capable digital citizenry that can defend against cyberattacks, the cultural project that attends this must not succumb to the edgy, master hacker stereotypes calling to "hack the planet."

In the present book, I would like to focus on a single feature of cybersecurity in our digital lives: the critical moment of noticing deception defines our online lives. From disinformation campaigns and malware to clickbait sales tactics and venture capital funded vaporware, so much of the language online is intended to deceive. Parsing all the lies from is no small task. While phishing detection is largely automated, no system is perfect. The scale of threat information is too vast to grasp by human understanding and simply reading. Given the multimodal and multimedia facets of social engineering, the language of cyberattacks could be termed "the rhetoric of digital deceit" or simply *security rhetoric*. While this book aims to dissect the rhetorical elements of phishing lures by evaluating the evolving, escalating, and proliferating threats to online life, I hope this is one facet of a broader critical approach, shared by both humanists and cybersecurity researchers. In this way, "security rhetoric" is a call to action for university educators to embed cybersecurity training within a comprehensive, digitally focused pedagogy that fortifies institutional integrity through informed, capable, and confident citizens.

After all, online threats have global consequences. The fourth chapter of this book examines the 2022 Russian invasion of Ukraine as a watershed moment in early twenty-first-century geopolitics. Among its many significances is its status as the first large-scale example of coordinated kinetic and cyber warfare. The chapter takes the Conti ransomware data hack and leak from Spring 2022 as a focus of this larger conflict. The so-called Conti Leaks offer a rare opportunity to scrutinize the culture, attitudes, tools, and training manuals of one of the largest cybercriminal organizations operating at the time.

The Conti ransomware crew were among the most feared and ruthless ransomware operators from about 2016 to 2022. They notoriously targeted hospitals during the Covid-19 pandemic, amassing estimated annual ransoms exceeding one hundred million dollars. Ironically, Conti discovered the hard way how a lapse in attention or judgment can compromise their own systems, as they were infiltrated and disrupted by a Ukrainian IT activist. After Conti declared alignment with the Russian government of Vladimir Putin, the ideology of a cybercriminal groups became a key feature of the subsequent hack and leak. There will be many consequences of this event, not least of which will be the precedents set for future conflicts. The weaponization of ransomware attacks to serve military goals through cybercrime means quasi-aligned enemy combatants may increasingly be regarded as legitimate military targets in future cyberwarfare.

Cyberwar is fought on both the open web and protected systems, which are developed and maintained, in large measure, by US corporations. It is an open secret that the National Security Agency (NSA) holds software vulnerabilities in store for clandestine use rather than disclose them to their respective corporations for patching.[27] These so-called zero day exploits are named as such due to the importance of patching systems. Microsoft's Windows operating system is used globally and represents a key attack surface for any nation-state seeking to exfiltrate enemy state or corporate secrets. For this reason, China has moved public sector operations to an open source Ubuntu-based operating system known as Kylin as of 2022.[28] The dynamics of corporate influence create tension among various stakeholders: governments, regulators, users—who are both citizens and consumers—and those concerned with national security in the realms of cybercrime and cyberwarfare.

The uneasy relationship between globalized corporate operations, cybercrime, and nation-state hacking is a theme of security rhetoric because the tactics of clickbait sales are so often repurposed or used as cover for malicious activity. The convenience of clicking a link is a temptation that plays into a corporate drive toward better and better customer service, which suggests that corporate IT playbooks are not enough to cope with this problem. There is charm and attraction in a lure that preys on our moments of indecision that are only increased through the default design logic of "ease of use." The third chapter of the book describes this blurring of lines as a kind of *automatic anxiety* that weaponizes psychological stereotypes to manipulate the emotional states of targets. In this chapter, I describe the work of Christopher Hadnagy, who has had an outsized influence on the field of social engineering and was banned from DEF CON for violating the code of conduct in 2022. This event was equal parts strange and troubling, but it served as critical moment that exposed the false psychological basis of social engineering attacks. Security rhetoric is more productively understood through a speculative mode of thinking that plays with fact and fiction that is perfectly suited to this period defined by disinformation.

A lure is a *counterfactual* experience—an experience that runs counter to the facts of a situation—leading victims to unwittingly assist attackers under the illusion of some perceived benefit. The global digital landscape allows for large risks to be taken with small actions, almost instantaneously.[29] It is feasible to isolate this moment when a victim makes an *invalid conditional inference*, in other words, the moment when the user makes a conclusion about a situation based on an incorrect assessment of the facts and takes a risk. The thought, "If this is an authentic email, this file should be safe to download and open," exemplifies such flawed conditional reasoning. Of course, deductive inferences are necessary for daily life. We must be able to rely on reasonably sound hypotheses for planning and decision-making, without immobilizing ourselves through excessive validation of every choice. The risks we face online are changing so quickly that we might benefit from a deeply attuned sense of authenticity and deceit, of truth telling and lying. Attention to security rhetoric helps intuit the style and feel of deceptive communications.

Security rhetoric derives its counterfactual structure from deceits and tricks, tricks that are inherently future-oriented and speculative in outlook. In this way, I argue that rhetoric and psychology are the twinned interdisciplinary connections for cybersecurity. A malware author might speculate, “*What if* I execute this combination of commands? Could I do something other than the software would allow?” A victim might lament, “*If only* I hadn’t downloaded that file, a victim might bemoan, and run it on my computer? *If only I paid more attention,* I could have avoided that attack.”

These thoughts are deeply emotional in their expression in us, the human operators. The emotional impact of security situations can be profound. Ruth Byrne, in her book *The Rational Imagination,* describes the role that counterfactuals have in our lives: “Most people can make inferences and solve everyday problems; the very existence of a world with science, laws, and technology indicates that at least some individuals pursue rational conclusions. But most people make mistakes.”[30] It is this very human fact that cybercriminals exploit. Given that “most people make mistakes” in their reasoning, cybercriminals can successfully cast a wide net online to manipulate, defraud, and exploit everyday users.

The cumulative effect of frequent errors in counterfactual reasoning can exact a significant emotional toll. This counterfactual experience of failure means so many people are left to ponder, *if only*. Thinking *if only* has a wistful, almost regretful tenor. The retroactive continuity that is possible by imagining about what could have been, *if only I did not click*. The speculative quality of asking, *what if*, implies an open sense of possibility. There is a creative quality to asking, *what if*: what possible future can I create and test with my imagination and my ability to plan logically? In this way, counterfactual thinking can work as both a positive sense of imaginative possibility as well as a tragic, nostalgic sense of regret.

This dual sense of counterfactual thinking serves as a future-oriented thought experiment, allowing for the imagination of both desired and undesirable outcomes. For example, contemplating, “What if I destroy my company systems by running this file?” represents a negative counterfactual scenario. “What if I save my company systems by following the correct reporting procedures for this anti-virus alert?” illustrates a positive, albeit boring, counterfactual scenario. In either case, there is a decision point. There is a moment in which the decision to act has a potential consequence that is positive or negative. Byrne refers to an “action effect,” which often leads individuals to inaction, hoping their indecision won’t be held against them.[31] Counterfactual thinking can lead to inaction, as people may imagine scenarios where not choosing is morally or ethically neutral. “What did they know, and when did they know it" represent some of the most important questions in this use of counterfactual avoidance. Of course, many people fail to act for a range of reasons, including burn out, fear, ignorance, or inability. After all, as Byrne so aptly puts it, "every action implies a corresponding inaction."[32] For instance, the failure to act based on a prospective, imagined scenario resulted in the missed opportunities to detect and thwart the Conti attack on the Irish hospital system in 2021, which resulted in the disruption of health services of an entire country for four months.[33]

Within this system counterfactual thought, the affective quality of cybersecurity joins the victim of cybercrime and their attacker. Both the attacker and the victim engage in counterfactual thinking, sharing a prospective mindset that transcends factual events; the attacker aims to manipulate the situation to their advantage by eliciting a specific response from the victim, while the victim is often deceived into making a choice erroneously. A malicious actor projects an act of deceit that creates a false impression, usually as an unassuming and banal email. The attack is often masked by the camouflage of everyday communication.

After all, email is the most common, least risky, and cost-effective means to launch an attack. Yet, recognizing a phishing email as a security risk requires a significant level of digital literacy. Image rendering, color, button shape, and copy writing are all features of UIs that make up our life online. Security rhetoric includes an understanding of how applications work, running in a browser or an operating system. Security rhetoric also includes a sense of style and voice. Oddly phrased messages, asking us to do something specific for non-specific reasons, are suspicious. Security rhetoric emphasizes proper spelling and grammatical structure, as these are hallmarks of legitimate, professional communication. Crucially, security rhetoric has the ability to intuit and identify the lurking "*what if*" scenarios. *What if* this is a malicious email? *What if* my work computer is compromised? Every decision is counterbalanced by an immediate risk assessment.

Evaluating risk is emotional work we provide, often without compensation from our employers, and this work is significant. We constantly bear the emotional burden of potentially jeopardizing both our workplaces and our digital private lives. Living in fear of regret: *if only Ipaid more attention*. Counterfactual experiences capture the emotional essence of cybersecurity, reflecting the complex feelings we have toward digital technology in the twenty-first century:

> Regret for actions and inactions is associated equally with hot emotions such as anger, disgust, embarrassment, and guilt, but regret for inaction is associated more with wistful emotions such as contemplation, nostalgia, and sentimentality, as well as with despair emotions, such as emptiness, helplessness, and unfulfilledness.[34]

It is for this very human fact that security rhetoric must not be interpreted in a pejorative way. Security rhetoric is not “security theatre.”[35] It is not empty rhetoric or hollow words. Security rhetoric does not appeal to security truisms, tropes, or jargon unless they serve a meaningful purpose. It diverges from the rhetoric of security that emerged during George W. Bush’s war on terror.[36] More than twenty years past 9/11, security is more personal and immediate. Our concerns now focus on physical and digital security and safety, not some racist propaganda about Weapons of Mass Destruction. Caring about our happiness feels more urgent after the Covid-19 pandemic than waiting for terror to strike.

There are real things to protect, every day. Cybersecurity safeguards our various digital lives online, on the one hand. Protecting bank accounts, social media personas, and every digital footprint left shopping, watching, and working on the internet on the other. The emotional toll on victims of cybersecurity incidents, whether at home or work, warrants both compassion and concern. Compassion and concern are opposing cognates of pity and fear. Falling for a phishing email, especially one leading to a major incident, is a twenty-first-century tragedy. There is a cruel tragic irony that cybersecurity incidents are so commonly reported and affect so many people today, but almost anyone who works or lives online can be the next victim.

Security rhetoric must be understood as the ability to understand, identify, and act on credible risks. By grasping the inherent risks of online interactions in various contexts, security rhetoric allows users to identify the technical, social, and cultural influences in social engineering attacks. Security rhetoric must also be practiced as a process, in an ongoing way. Byrne offers evidence to suggest that “people imagine a counterfactual alternative to an inaction only when they have thought about two possibilities for the inaction, and discovered an asymmetry in the goodness of the imagined outcomes for the action and the inaction.”[37] Individuals managing risky situations need both the time and incentive to engage in speculative thought to envision potential attack scenarios. Security rhetoric involves applying imagination and rational thought to identifying the underlying tactics, techniques, and procedures (TTPs) that characterize hostile communications. These TTPs of security rhetoric include the standard appeals to authority, fear, and urgency seen in social engineering attacks. The mitigation steps included within security rhetoric include the emotional work and close reading required to first identify then act on phishing-based attacks.

Spammers manage massive botnets of coordinated compromised systems around the world. These networks consume significant processing and electrical power to send billions of messages, thereby generating millions of dollars for their operators. In the words of Brian Krebs, author of the highly respected and award-winning *Krebs on Security* blog, "these junk email artists earned a few million dollars for their efforts, yet they've forced businesses and consumers to spend hundreds of millions more shoring up digital defenses to fight their daily glut of crimeware."[38] The automation and expansion of these systems are increasing in speed and scale, mirroring broader trends in our digital landscape. The environmental metaphor is appropriate because these botnets extend and grow like a kind of contagion or pollution. Krebs describes this cyclical system in the following way:

> This technological arms race requires the development, production, and distribution of ever-stealthier malware that can evade constantly changing antivirus and anti-spam defenses. Therefore, the hackers at the throttle of these massive botnets that spew plain old spam typically are used to distribute junk email containing new version of the malware that helps spread the contagion.[39]

Krebs is great. Let's continue with him for a moment:

> In addition, spammers often reinvest their earnings from spamming people into building better, stronger, and sneakier malicious software that can bypass antivirus and anti-spam software and firewalls. The spam ecosystem is a constantly evolving technological and sociological crime machine that feeds on itself.[40]

Much like Ouroboros consuming its own tail, spam botnets are a paradox of technological destruction and creation. It feeds itself and grows by its own need to evade, evolve, and expand, all with the goal of maximizing profits. Security rhetoric offers a human-centric heuristic that can intervene before the trap is sprung. Users may be lured by the message for a moment, but influence does not equate to causality.

A moment of critical awareness regarding the message can make all the difference. These moments will only increase in the coming years.

The recent revolution in generative Artificial Intelligence (AI), particularly in Large Language Models (LLMs), is accelerating the speed, size, and complexity of various issues related to social engineering. The November 2022 release of ChatGPT by OpenAI has upended many assumptions about the potential and scope of computer systems' natural language understanding, just as the general public is becoming aware of generative AI. It did not take long for hackers to use these systems to generate useable malware in seconds and work around controls with so-called "prompt injections."[41] In the days after the release of GPT-3, as the world grappled with the consequences of LLMs, Bruce Schneier and Barath Raghavan wrote an opinion piece in *Wired* that summed up the mood of the moment. "Brace Yourself for a Tidal Wave of ChatGPT Email Scams" summed up the ways that LLMs will supercharge an already enormous problem.[42] Their thesis is clear: the automation of spam messages through AI will exacerbate an already significant problem. This will be problematic precisely because LLMs are just so good at sounding human. In 2021, Lily Hay Newman was already flagging the issue with the headline, "AI Wrote Better Phishing Emails than Humans in Recent Test."[43] AI will pose a public safety risk if cybercriminals find a way to commodify the technology, similar to how ransomware has commercialized the criminal use of encryption.[44] It will be increasingly difficult to tell the difference between what is a bot and what is not.[45]

Phishing is already an enormous problem that anyone with an email account can confirm.[46] It is such a prosaic problem that it threatens being overlooked for the flashy new technology.[47] It is a problem with deep roots in the growth of the internet itself. The noise of so many spam messages that go sent by the *billions* each and every day has a transformative impact on the content of the web. What if those manipulative messages that seek to influence our behavior train AI? What bias would be trained into that system? Finn Brunton described the "global spam machine" as representing a "shadow history of the internet."[48] Spam serves as the dark underbelly of what is often hailed as humanity's greatest invention, rooted in deception and criminality. Brunton reminds us of the way this "inexplicable irritant" has shaped our language:

> The word “spam” means very different things to different people at different times. It is a noun, collective and singular, as “this spam” can mean “all these messages I’ve received” or “this particular email.” It is a verb, as in “they spam me,” and an adjective, as in “this is spammy.” It refers to many varieties of exploitation, malfeasance, and bad behavior, and pam terminology has branded out into specific subdomains, from “phishing spam” and “419 spam” to splogs, linkfarms, floodbots, content farms. […] But spam begins to make sense only when we get specific and separate out the different types, motives, actors, and groups.[49]

Of course, attributing these messages to individual senders is challenging, which is why security and software companies publish yearly reports on the state of the problem. According to a report from Radicati, email remains the most ubiquitous and stable form of communication for business and consumer communications, with a daily global average of 333.3 billion messages sent in 2022.[50] Estimates suggest that 56.5 percent of these emails are unwanted. Email marketing companies indicate that over a third of these messages are advertising-related. 31.7 percent of spam messages are malicious with the goal of uploading malware. Just 2.5 percent are targeted scams, of which 73 percent are seeking identity theft fraud schemes. These seemingly solid numbers are of course estimates.[51] The finer details of geography are sometimes lost in viewing a planetary problem.[52]

Research Culture

Spam represents only a facet of the broader issue of malicious messages, which can range from national security concerns to influencing the tone and style of online discourse. The quality of our conversations influences how we interact with our neighbors and shapes internal public debates on various national issues. Moreover, the internet serves as a platform for numerous legitimate and enriching cultural dialogues. The scholarly conversation held within research culture is a significant venue for these public conversations because of how universities guide policy discussions and the development of new technologies and ideas. This style of public debate is essential for a stable and engaged democracy, and the internet provides an apt medium for disseminating information and facilitating such conversations. I advocate for a comprehensive, scholarly-focused initiative aimed at understanding the implications of cybersecurity across all academic disciplines. Scholarly discourse can demystify this rapidly evolving, technical issue for students, researchers, and the public alike.

In the face of so much trolling and keyboard culture warriors, academic institutions continue to serve as reliable platforms for forging durable and authentic consensus. Distinguishing satire and parody from disinformation and misinformation is often an insurmountable challenge in a society that values free expression. Numerous studies grapple with defining disinformation and misinformation due to the complex interplay of irony, parody, satire, trolling, and propaganda.[53] In liberal democracies with free speech protections, any legal definition of misinformation is likely to face immediate challenges.[54] After all, there is money to be made by peddling inflammatory ideas. Alphabet, one of the world's largest corporations, readily enables the monetization of conspiracy theories on YouTube.[55] The YouTube comment section is a bastion of free speech and online toxicity in equal measures. Misinformation is also locally targeted, which undermines global solidarity to confront a problem on this scale. Dog whistle calls to violence that resonate in one part of the world will not be understood or identified by people outside that cultural context, further hampering content moderation efforts.[56]

The Covid-19 pandemic offered a unique opportunity to refine the local dissemination and amplification of disinformation.[57] Disinformation can be strategically aimed at specific demographic groups across national and regional boundaries.[58] Ideologically driven groups can also be collected under a shared banner, such as science denialism, which exploits scientific debate as a political wedge.[59] Certain platforms can become disinformation hotbeds due to factors like end-to-end encryption, inadequate content moderation, or regional and national biases.[60] National governments also use disinformation against their own populations as well as an attempt to disrupt political coherence in competing nations.[61]

How we know things about ourselves, each other, and the world around us represents the foundation of our values as members of complex communities of aligned interests. These shared ways of knowing shape and define cultures. The languages we use, the stories we tell, the histories we keep, and the beliefs to which we assign value represent the daily, practical basis for belonging.

These are global, human problems. According to We Are Social, a transnational social media design company, more than half the world are urban and internet-connected: The global population has grown well beyond 8 billion people in 2023, and 57 percent of the world lives in urban areas; 68 percent of the global population uses mobile phones, a total of 5.44 billion people; 64.4 percent of the world's population is online and is increasing at a rate of more than 1.9 percent each year.[62] Increasingly, the online world of social media platforms, institutional websites, and peer-to-peer networking represent the communications media of humanity. It is also a huge opportunity for cybercriminals, with an increasing number of targets each year. Each of these individuals has access to varying degrees and types of education, technology, legal protections, as well as internet connectivity.

Long-term cultural questions loom large in cybersecurity as a result. As our human cultural legacy increasingly transitions online, it is time to ask: what are we protecting? What ideas deserve to be preserved through protection? What knowledge do we privilege through our protection? What and who do we leave outside digital safeguards? Of course, we can flip this set of questions that animate this book: What gets attacked? What online cultures do not deserve the protection? What online cultures deserve protection but are neglected? What ideas and values do we sequester and destroy without adequate security? What ideas proliferate through a lack of protection? The answers to these questions and the degree of their importance will change over time. They are, however, questions that must be continually asked and agreed upon.

Digital Humanities archivists—who are actively engaged in galleries, libraries, archives, and museums—address these questions and work toward building a consensus around them. The protection of our languages, stories, histories, and beliefs extends beyond the physical artifacts and objects housed in these cultural preservation centers. The urgency to safeguard digital cultural artifacts and the professionals who curate them is escalating, and failure is not an option when our cultural legacies are at risk.

Security practices are a largely implicit feature of community values and belonging, but protecting digital cultures from online threats requires a strict, formal attention to what and how we protect digital cultural assets. A shared understanding of the relative risks associated with the people, things, and data needing protection is an important expression of values and shared interests. This mutual awareness of threats not only enhances our security posture but also opens the door for broader collaborations, fostering a refined sense of solidarity through the cultivation of security cultures.

Academic culture accretes community values in a similar manner, though the pace of change now runs in step with the rapid evolution of digital technology. Such a broad opening critical frame is warranted given the ubiquitous and emergent quality of digital culture and the internet in general. The economic imperative to protect bank records and corporate secrets often eclipses the cultural and historical importance of safeguarding digital archives. Lisa Gitelman in *Always Already New* describes this historical logic this way: “The history of emergent media, in other words, is partly the history of history, of what (and who) gets preserved—written down, printed up, recorded, filmed, taped, or scanned—and why.”[63] Although digital archiving practices and research data management processes are well established in many disciplines, the evolving and emergent quality of cybersecurity risks require a holistic and continual reappraisal of the research life-cycle over the long term. This reassessment should encompass both technical systems and infrastructure as well as human factors.[64]

Digital culture can be readily characterized by its emergent qualities, embracing iterative processes that recombine and repurpose existing materials while incrementally adding wholly new ideas. Consider how memes evolve, merge, and grow through modification and reuse. Similarly, academic cultures and the knowledge they generate are shaped by peer review, citation, and scholarly discourse. There are many practical, workaday activities necessary to undertake, manage, and support successful scholarly work. Open and social scholarship is predicated on access to research materials, yet openness must be counterbalanced by security considerations. In March 2021, the Government of Canada highlighted the importance of research security policies, emphasizing the need to “integrate national security considerations into the evaluation and funding of research partnerships,” thereby making research security a national strategic priority. Increasingly, a measure of success must include the security of research data as well as researchers themselves. As institutional risk to cybersecurity incidents continues to grow, security best practices are now a simple reality of development and operating a scholarly project or research institution.

The internet’s fundamental role in numerous research practices renders the core infrastructure of knowledge production, storage, and dissemination susceptible to risk and vulnerability in the event of a cybersecurity incident. *Threat modelling* and *risk assessments* must become a condition of both funding and ongoing safe operations of complex, public facing research groups. With the increasing threats posed by criminal and nation-state threat actors, which includes critical research and development infrastructure, security culture must also become a key facet of sharing what we know and validating the integrity of our knowledge systems and infrastructure. While archivists have long managed this kind of risk, the social and engaged quality of scholarship today broadens researcher attack surfaces beyond secure data storage.[65]

Data integrity can be defined for our purposes here as digital archiving protocols like LOCKSS (Lots of Copies Keep Stuff Safe) or Research Data Management (RDM) processes described by the formerly titled Portage Network (now covered under the four pillars of The Digital Research Alliance of Canada).[66] Software assurance is the practice of ensuring software is free from vulnerability or defect that might disrupt research practices anywhere along the research life-cycle. Both proprietary and open-source software present unique challenges in verifying the software supply chain and managing complex dependencies that could potentially invalidate, corrupt, or otherwise compromise research outputs.[67] Data integrity and software assurance are just two categories within the purview of security practices that will inform fundamental considerations in our research methods. From this perspective, operational security becomes a guiding principle in all research practices.

When taken seriously, security practices require the incorporation of core principles such as data integrity and software assurance into our commitment to collaborative ethics, openness and transparency, preservation, and physical safety.[68] While cybersecurity provides the impetus for integrating formal threat assessments in our research cultures, physical security must also account for the security of researchers, research participants, physical artifacts, and the built environment. The breadth of the risks facing researchers must match the urgency of the problem. It is possible to make a specific claim beyond the figurative hesitancy of my title: **Security culture is an expression of values**. Researchers must be ready to reflect on disciplinary values and how best to enact those values in securing our work. By augmenting current lab-based research methods with a robust security-first research methods, it is possible to express the ethics at work in digital projects. As a humanities-based scholar of digital media, aligning security best practices with humanistic values is not easy or simple.

Evolving Values

Managing and deploying infrastructure is fundamentally about enforcing authenticated access. It is necessary to describe people, places, and things as assets, risks, and threats. A reductive or simplistic assessment of assets, risks, and threats only serves to increase our risk. As such, suspicion and paranoia are recast as the work of critique within a holistic security practice. Suspicion and paranoia are not generally regarded as productive critical approaches in the humanities or any other research culture for that matter![69] Scholarly curiosity is a source of new discovery and insight, whereas suspicion and paranoia are often associated with a type of prejudice or pre-judgment.

For example, a realistic accounting of nation-state level risks may be regarded with incredulity or subject to charges of cultural insensitivity or outright racism.[70] Anticipating threats requires forethought and speculation about potential threats, wherein national political forces set the regulatory and legal frameworks.[71] By contrast, a hypothesis might emerge to align some of these attitudinal differences complement each other through the mechanisms of discourse, discussion, and debate. Bridging the gap between security and research cultures has already become an existential question for the validity and access to useful research.[72] If productive points of overlap can be found between security best practices and digital research practices, scholarship will remain available and have greater impacts over time, while security researchers benefit by better understanding how their systems and behaviors impact online cultures.

A robust security-oriented research culture must also be capable of protecting knowledge stakeholders across the research life cycle. Research participants must be given reasonable and informed consent regarding the security of their contributions; researchers of all ranks must be assured of physical security in the face of increasing politicization of disciplines across the university sector, and citizens must retain access to publicly funded research regardless of external forces, such as science denialism, political extremism, religious fundamentalism, or criminal profiteering. In such a context, our research security cultures must *enact our values* in a way that conveys the importance of the people, objects, and ideas that must come together to produce some new insight or make the next great discovery.

The first articulation of this value-driven security approach, described as a “goals and values alignment,” comes from Eugene Spafford in 2019.[73] The alignment of goals and values in security emerges from a holistic notion of trust, wherein researchers align the priorities of developers and users across the supply chain of both software *and* ideas. Spafford articulates this in his remarks at a gathering of Purdue University’s CERIAS research group: “When we’re going to say we’re going to trust something in the system, we have to be sure we understand what trust means to each of us as individuals, who have values and goals about the things we want to do and the information we have.”[74] When trust is extended through technical, social, and cultural values to achieve a shared goal, a security culture is formed.

Security cultures can be deeply affirming and valuable relationships, but security cultures can also seek to divide us against them. Spafford also describes how our research institutions have goals and values that may be out of alignment with our own as researchers. Community values define who and what gains trust and further defines what and how we protect. When extrapolated further, within the domain of the vast amount of research occurring globally online, the goals and values of our institutions, governments, colleagues, departments, funding agencies, and others will not necessarily align with our own priorities as individual research groups.

Researchers cannot defend against all adversaries across the entire research lifecycle, and new threats will emerge from both technological and political change. New autocratic governments may supplant democratic ones, resulting in a hostile environment literally overnight. The guerrilla archiving of public climate data that occurred in the wake of the election of Donald Trump in 2016 is an example of defensive measures precipitated by a change in government.[75] Set within the tumultuous first decades of the twenty-first century, researchers must also remember that security best practices are not a durable methodology in themselves because they are mutable by necessity, predicated on a complex network of technological, political, and cultural realities.

The security of research infrastructure and assets may run counter to the proposed ethics espoused by humanistic research in general. What is to be done if our ethics are out of line with a necessarily restrictive security posture? What if our security procedures are out of step with the content and intent of our research? We can first identify areas where security practices do, in fact, align well with the methodological norms of our disciplines. In this case, DH has long represented a widening of the methodological scope of the humanities, which has forced a decades-long revaluation of the motivations and intentions of humanistic inquiry. As a discipline, DH has been voracious in its re-evaluation of key questions in academic work—including peer review, credit allocation, labor rights, positionality, learning through failure, prototyping, and other process-based experiential learning—which of course still says nothing of computation, the internet, or any number of multimedia approaches that are expanding what and how humanities scholars make discoveries and add to the public discourse on human culture in general.

The scope and scale of DH as a scholarly project is large and evolving, which is why security best practices must now be added to this list of procedural, technical, and organizational approaches. It may be productive to have a provisional list of values that are likely to conflict with security practices. These points listed below are thematic consistencies throughout the chapters of this book and serve to highlight just a few of the most salient points of friction between possible assumptions of researchers and the assumptions of security professionals. The values of many academic researchers might be summarized, as an emphasis on openness, transparency, collaboration, and sharing among others:

> *Threat Modelling and Risk Assessments:* Security best practices must become a condition of both funding and ongoing safe operations of complex, public facing research groups. The human factors of research security must protect against the misuse of project systems, while also protecting the privacy and physical security of researchers, students, and research participants. Engaging in a threat modeling process that is ongoing and evolving would require an inventory of assets and a risk assessment related to the protection of researchers and research materials, including software and data.

> *Validating Open-Source*: Understanding the limits of the Open-Source movement in security is critical for the operational security of public facing digital research projects. While there is no security through obscurity, open-source tooling matches well with a scholarly predilection toward open access. However, an unexamined use of open-source tooling may occur because of scholarly values of openness. Validating our software supply chain may represent an enormous cost to research projects, but threat assessment processes may prove that such protections are necessary. Validating research tooling may need to become a community effort among researchers using shared tooling, which would require new approaches to reporting and collaboration on security-related findings.

> *Breaking Things*: DH researchers, at least since Jerome McGann's *Radiant Textuality*, have been interested in "making things."[76] Researchers may need to be involved in *breaking things* as a mean of enacting a form of civil disobedience for digital citizens. Encrypting, destroying, and moving data are all important mechanisms to manage risk from this perspective. If data poses a risk to individuals, it may be necessary to simply delete it. However, we may also be compelled to download and share data without permission. Hack and leak operations may be the only means to preserve incriminating data that risks being destroyed by those eager to cover their tracks. Researchers may find new allies with activist archivists, like Archive Team or Distributed Denial of Secrets.[77] Breaking things will require a new form of collegial solidarity to support and protect researchers engaged in quasi-legal data collection in the course of their research.

> *Failure Is Not an Option*: There is little room for learning through failure in a security-oriented research culture. Defensive security is inherently asymmetrical. Defenders must be completely successful in their defensive measures, all the time. An attacker need only succeed once to compromise a digital research project. A project may then be taken off-line, deleted, poisoned, or otherwise subverted for political or financial reasons. Shawn Graham's *Failing Gloriously and Other Essays* is an excellent history of this embrace of failure in humanities research.[78] Failing, however gloriously, may not be an option for many researchers in the increasingly hostile threat environment emerging online. Failure must instead be replaced by a more responsive framework that includes identifying risks, protecting systems, detecting breaches, responding effectively, and recovering operations.[79] New opportunities for collaboration may exist among researchers interested in validating security measures through penetration testing.

These four points of overlap between research methods and security practices demonstrate the opportunities and the urgency of this project. The final chapter will return to these values to describe a means to move beyond a threat-informed security training. Cultural criticism is often a highly engaged, activist research practice. Public facing research is also increasingly visible to antagonistic forces, who are hostile toward institutions of higher education.

These processes, procedures, and protocols enact a security rhetoric that must be identified and understood in a defensive posture, but security rhetoric must also be an expression of values. Finally, I turn to Meg Worley's essay, "The Rhetoric of Disruption," which reminds us that, by ignoring rhetoric, "we claim that it cannot affect us, but we are surrounded by evidence to the contrary."[80] She describes how the constructive social rhetoric around can serve as a means of coercion and control.[81] Along with many generous suggestions for researchers, particularly in DH, she reminds researchers that "we must give serious thought to the role emotion should play in the realm of academic discourse."[82] In this way, the ability to identify emotional tenor found in the rhetoric of cyberattacks is a key defensive tool.

University-based researchers are facing increasingly sophisticated, often automated, security threats.[83] Cybersecurity best practices for researchers represent a highly technical set of challenges, which are added to an already broad set of expertise areas. This is a significant responsibility, and the fact remains that the processes and policies followed by researchers and research participants are critical in mitigating a range of risks and vulnerabilities. The security culture that is precipitated by social scholarship would be transparent and participatory. Through a participatory style of security policy adoption and training, researchers at every level can make better choices that impact the security of research projects. Because security threats evolve quickly, institutions with deeply integrated security policies will be able to improve situational awareness and protect researchers, data, and institutional infrastructure.

As digital technology rapidly evolves, researchers face the challenge of securing their work and protecting stakeholders across the research life cycle. Security practices, encompassing data integrity, software assurance, and authenticated access, must align with the values of openness, transparency, and collaboration that underpin scholarly research. The increasing sophistication of cyber threats, coupled with the politicization of research and the rise of autocratic governments, necessitates a holistic approach to security that considers both cyber and physical dimensions. Research security culture must emphasize the importance of trust, informed consent, and ethical collaboration. Security practices should be user-centric, participatory, and transparent, reflecting the values and goals of researchers and participants. By integrating security best practices into research methods, scholars can mitigate risks, protect research assets, and ensure the continued availability and impact of scholarship in a world where threats are constantly evolving.

To bolster the security of research practices, project goals should now incorporate a comprehensive framework that addresses the security of individuals involved in research, the assets generated or utilized, and the supporting software and technical infrastructure. These research security practices should be continuously measured and assessed at each stage of the research life cycle. All stakeholders in knowledge creation—ranging from researchers and participants to institutions and end-users—should be well versed in the security protocols followed throughout the creation, preservation, and dissemination of research, particularly when it is publicly funded. To capture performance validation and verification of these practices, appropriate tools will be needed to securely implement new knowledge environments. These tools must reflect the values, goals, and needs of researchers and other participants, thereby fostering a security culture that aligns with the broader objectives of scholarly inquiry.

The next chapter delves into the intersection of cybersecurity, psychology, and rhetoric, arguing for a human-centered approach to understanding online threats. It critiques the militaristic language often used in cybersecurity and suggests that such terminology can be limiting. Instead, the chapter proposes that cybersecurity should be seen as a form of "emotional labor," requiring not just technical skills but also a deep understanding of human psychology and emotion. Drawing on theories from the humanities and social sciences, including the work of Aristotle, Burke, and Toulmin, it explores how language and emotion can be manipulated to deceive and harm. It also proposes the need for an intersectional approach to cybersecurity, recognizing that threats can be experienced differently based on social and demographic factors. The chapter concludes by advocating for a more nuanced, emotionally intelligent approach to cybersecurity, one that takes into account the complex interplay of language, emotion, and social context.

Notes

[1] Nicola J. Bown, Daniel Read, and Barbara Summers, "The Lure of Choice," *Journal of Behavioral Decision Making* 16 (2003): 298.

[2] Dan Gooden, "I'm a Security Reporter and Got Fooled by a Blatant Phish," *Ars Technica*, November 8, 2022, https://arstechnica.com/information-technology/2022/08/im-a-security-reporter-and-got-fooled-by-a-blatant-phish/; Dan Gooden, "Behold, A Password Phishing Site that Can Trick Even Savvy Users," *Ars Technica*, March 21, 2022, https://arstechnica.com/information-technology/2022/03/behold-a-password-phishing-site-that-can-trick-even-savvy-users/; Dan Goodin, "Ongoing Phishing Campaign Can Hack You Even When You're Protected with MFA," *Ars Technica*, December 7, 2022, https://arstechnica.com/information-technology/2022/07/microsoft-details-phishing-campaign-that-can-

hijack-mfa-protected-accounts/. Dan Goodin has done a fantastic job reporting phishing attacks over the years, which is difficult and important because phishing can be boring and banal.

[3] Cybersecurity professionals have been celebrating "Cybersecurity Awareness Month" in October for 20 years! See https://staysafeonline.org/programs/events/20-years-of-cybersecurity-awareness-month-kick-off-event/.

[4] A standard citation at this point would be to Christopher Hadnagy, *Social Engineering: The Science of Human Hacking* (Indianapolis, IN: Wiley & Sons, 2018). However, this book takes issue with the ethics and academic premise on which his work rests as well as the broader influence Hadnagy has had on social engineering and cybersecurity more broadly.

[5] Respectively, see McKenzie Wark, *A Hacker Manifesto* (Cambridge: Harvard University Press, 2004); Stephen Melville (ed.), *The Lure of the Object* (Williamstown, MA: Clark Studies in the Visual Arts, 2004); Marcus J. Carey, and Jennifer Jin, *Tribe of Hackers: Security Leaders* (Indianapolis, IN: Wiley & Sons, 2021); John Paul Ricco, *The Logic of the Lure* (Chicago: University of Chicago Press, 2002); Guy Barefoot, *Trash Cinema: The Lure of the Low* (New York: Wallflower, 2012).

[6] Bruce Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World* (New York: W. W. Norton and Company, 2018).

[7] Alexander Reid, "Digital Humanities Now and the Possibilities of a Speculative Digital Rhetoric," in *Rhetoric and the Digital Humanities*. Jim Ridolfo and William Hart-Davidson (eds.) (Chicago: University of Chicago Press, 2015): 16.

[8] Lev Manovich has since walked back some of these triumphal claims against rhetoric, at least implicitly in his more recent *Cultural Analytics* (Cambridge, MA: MIT Press, 2020), which returns to a need to notice "types, structures, or patterns" through data science techniques.

[9] Raymie E. McKerrow, “Critical Rhetoric: Theory and Praxis,” *Communication Monographs* 56 (1989): 100.

[10] See also Krista Ratcliffe, *Rhetorical Listening: Identification, Gender, Whiteness* (Carbondale: Southern Illinois University Press, 2005).

[11] I will remove the plural form in my articulation of security rhetoric and the rhetoric of cybersecurity lest I sound like a snake, with so many trailing consonant “s” sounds. I hope the plurality of contexts, media, relations, and situations are assumed.

[12] Jennifer Sano-Franchini, “Cultural Rhetorics and the Digital Humanities: Toward Cultural Reflexivity in Digital Making,” in *Rhetoric and the Digital Humanities*. Jim Ridolfo and William Hart-Davidson (eds.) (Chicago: University of Chicago Press, 2015): 52.

[13] Ibid.

[14] Ibid., 61.

[15] I evoke Matthew Kirschenbaum’s “What Is Digital Humanities and What’s It Doing in English Departments?” as a link to my scholarly, disciplinary home and to gesture to the breadth of consequences that emerge from our always online life. The cultural consequences of cybersecurity may emerge from a research context, interested in preserving digital scholarly objects, but the cultural consequences of participating online every day have something to do with Cynthia Selfe’s opening epigraph to Kirschenbaum’s famous essay. As DH matures, “the rhetoric of technopower” appears to be an increasingly significant feature of our work as scholars as well as the concerns of digital citizens the world over. Kirschenbaum’s essays are available, open access: https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities.

[16] Ibid., 53.

[17] Ian Bogost, *Persuasive Games: The Expressive Power of Video Games* (Cambridge, MA: MIT Press, 2007), 29.

[18] Ibid., 29.

[19] Elizabeth Losh, *Selfie Democracy: The New Digital Politics of Disruption and Insurrection* (Cambridge, MA: MIT Press, 2022), xi.

[20] Ibid., xx.

[21] See https://www.americanrhetoric.com/.

[22] Losh, *Selfie Democracy,* xx.

[23] Bogost, *Persuasive Games*, 25. I follow Bogost's skepticism about this term's usefulness, especially as it has been used to describe questions related to writing, composition, and the web. See Douglas Eyman, *Digital Rhetoric: Theory, Method, Practice* (Ann Arbor, MI: University of Michigan Press, 2015).

[24] David Graeber, and David Wengrow, *The Dawn of Everything: A New History of Humanity* (New York: Signal, 2021), 524.

[25] McKerrow, "Critical Rhetoric," 106. Emphasis in the original.

[26] Aaron Mauro, *Hacking in the Humanities: Cybersecurity, Speculative Fiction, and Navigating a Digital Future* (New York: Bloomsbury Publishing, 2022), 185. See also Domenico Fiormonte, "Digital Humanities and the Geopolitics of Knowledge," *Digital Studies* 7 (2017), https://www.digitalstudies.org/article/id/7313/.

[27] Liam Tung, "NSA: Our Zero Days Put You at Risk, but We Do What We Like with Them," *ZDNet,* March 13, 2014, https://www.zdnet.com/article/nsa-our-zero-days-put-you-at-risk-but-we-do-what-we-like-with-them/.

[28] Tobias Mann, "China Rallies Support for Kylin Linux in War on Windows," *The Register,* July 3, 2022, https://www.theregister.com/2022/07/03/china_openkylin/.

[29] Bruce Schneier's *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World* (New York: Norton & Co., 2018) surveys the range of potential attacks and their disastrous consequences.

[30] Ruth Byrne, *The Rational Imagination: How People Create Alternatives to Reality* (Cambridge, MA: MIT Press, 2005), 16–17.

31 Ibid., 46.

32 Ibid., 56.

33 See Chapter 4.

34 Byrne, *The Rational Imagination*, 57.

35 In the wake of the September 11, 2001 terrorist attacks in New York, security in airports increased. Bruce Schneier described these largely ineffective security protocols as "security theatre" that sought to make us feel better in the face of so much fear but did little to increase overall safety. See Bruce Schneier, *Beyond Fear* (New York: Copernicus Books, 2003), 37–8.

36 Rosa Brooks, *How Everything Became War and the Military Became Everything: Tales from the Pentagon* (New York: Simon and Schuster, 2017); Yaseen Noorani, "The Rhetoric of Security," *CR* 5 no. 1 (2005): 12–41.

37 Byrne, *The Rational Imagination*, 62.

38 Brian Krebs, *Spam Nation: The Insight Story of Organized Cybercrime —From Global Epidemic to Your Front Door* (Naperville, IL: Sourcebooks, 2014), 131. See https://krebsonsecurity.com/ for more.

39 Krebs, *Spam Nation*, 4.

40 Ibid.

41 For an early description of AI-generated malware with LLMs, see Dan Goodin, "ChatGPT Is Enabling Script Kiddies to Write Functional Malware," *Ars Technica*, June 1, 2023, https://arstechnica.com/information-technology/2023/01/chatgpt-is-enabling-script-kiddies-to-write-functional-malware/. For an excellent description of "prompt injection attacks," see Matt Burgess, "The Security Hole at the Heart of ChatGPT and Bing," *Wired*, May 25, 2023, https://www.wired.com/story/chatgpt-prompt-injection-attack-security/.

[42] Bruce Schneier, and Barath Raghavan, “Brace Yourself for a Tidal Wave of ChatGPT Email Scams,” *Wired*, April 4, 2023, https://www.wired.com/story/large-language-model-phishing-scams/.

[43] Lily Hay Newman, “AI Wrote Better Phishing Emails than Humans in a Recent Test,” *Wired*, August 7, 2021, https://www.wired.com/story/ai-phishing-emails/.

[44] See https://www.cisa.gov/stopransomware for more information about the almost ubiquitous ransomware problem on the early twenty-first century internet.

[45] See “Four Truths about Bots,” *Twitter Common Thread Blog*, September 21, 2021, https://blog.twitter.com/common-thread/en/topics/stories/2021/four-truths-about-bots.

[46] As a sketch of the threat faced by phishing, see the following: Jess Weatherbed, “A Huge Phishing Campaign Has Targeted Over 130 Companies, Affecting Twilio and Signal,” *The Verge*, August 26, 2022, https://www.theverge.com/2022/8/26/23323036/phishing-scam-campaign-twilio-hack-companies; Alex Hern, “Pentagon Leak Traced to Video Game Chat Group Users Arguing Over War in Ukraine,” *The Guardian*, April 11, 2023, https://www.theguardian.com/world/2023/apr/11/pentagon-leak-traced-to-video-game-chat-group-users-arguing-over-war-in-ukraine; @Serpent, “NEW PHISHING SCAM Already $650,000 stolen from a single individual and it’s going to happen to a lot more people. This is how it happened,” *Twitter*, April 17, 2022, https://twitter.com/serpent/status/1515545806857990149?s=12&t=5nys4sgMXlCS9T-ONklcLQ; Thomas Claburn, “This Browser-in-Browser Attack Is Perfect for Phishing,” *The Register*, March 18, 2022, https://www.theregister.com/2022/03/18/browser_in_browser_phishing/; @tokyoneon, “How to Phish for User Passwords with Powershell,” *Black Hills Infosec Blog*, July 27, 2021, https://www.blackhillsinfosec.com/how-to-phish-for-user-passwords-with-powershell/; Brian Krebs, “How 1-Time Passcodes Became a Corporate Liability,” *Krebs on Security*, August 30, 2022, https://krebsonsecurity.com/2022/08/how-1-time-passcodes-became-a-

corporate-liability/; Dan Goodin, “Hackers Can Infect >100 Lenovo Models with Unremovable Malware. Are You Patched?” *Ars Technica*, April 19, 2022, https://arstechnica.com/information-technology/2022/04/bugs-in-100-lenovo-models-fixed-to-prevent-unremovable-infections/; Dan Goodin, “Discovery of New UEFI Rootkit Exposes an Ugly Truth: The Attacks Are Invisible to Us,” *Ars Technica*, June 26, 2022, https://arstechnica.com/information-technology/2022/07/researchers-unpack-unkillable-uefi-rootkit-that-survives-os-reinstalls/.

[47] Bill Toulas, “Phishing Attacks Abusing SaaS Platforms See a massive 1,100% Growth,” *Beeping Computer*, August 23, 2022, https://www.bleepingcomputer.com/news/security/phishing-attacks-abusing-saas-platforms-see-a-massive-1-100-percent-growth/; See also https://www.microsoft.com/en-us/wdsi/threats.

[48] Finn Brunton, *Spam: A Shadow History of the Internet* (Boston: MIT Press, 2013), xiii.

[49] Ibid., xiv.

[50] Radicati Group, “Email Statistics Report, 2018–2022,” *Radicati.com*, January 2018, https://www.radicati.com/wp/wp-content/uploads/2018/01/Email_Statistics_Report,_2018-2022_Executive_Summary.pdf.

[51] See the “Spam Statistics” report from dataprot.net, which compiles reports from Mailmodo, Oberlo, and Talos Intelligence: https://dataprot.net/statistics/spam-statistics/.

[52] The planetary scale of spam also relates to the ecological impact of so many unwanted messages. Mike Berners-Lee, brother to Sir Tim Berners-Lee, inventor of the internet—is a carbon footprinting researcher who has estimated the global carbon cost associated with spam messages as 0.03 grams of CO2 per spam email picked up by a spam filter. A quick bit of math on the back of napkin produces staggering sums of carbon for messages that are rarely read. See Mike Berners-Lee, *How Bad Are Bananas? The Carbon Footprint of Everything* (London: Profile Books, 2020), 16.

53 Jason Cabañes, C.W. Anderson, and Jonathan Corpus Ong, “Fake News and Scandal,” *The Routledge Companion to Media and Scandal* 88 (2019): 1–18.

54 Ronan Ó Fathaigh, Natali Helberger, and Naomi Appelman, “The Perils of Legally Defining Disinformation,” *Internet Policy Review* 10 no. 4 (2021), https://doi.org/10.14763/2021.4.1584.

55 Cameron Ballard, Ian Goldstein, Pulak Mehta, Genesis Smothers, Kejsi Take, Victoria Zhong, Rachel Greenstadt, Tobias Lauinger, and Damon McCoy, “Conspiracy Brokers: Understanding the Monetization of YouTube Conspiracy Theories,” *WWW’22: Proceedings of the ACM Web Conference* (2022): 2707–18. Available: https://doi.org/10.1145/3485447.3512142.

56 Jialun Aaron Jiang, Morgan Klaus Scheuerman, Casy Fliesler, and Jed R. Brubaker, “Understanding International Perceptions of the Severity of Harmful Content Online,” *PLOSOne* (2021). Available: https://doi.org/10.1371/journal.pone.0256762.

57 Sharifa Sultana, and Susan R. Fussell, “Dissemination, Situated Fact-Checking, and Social Effects of Misinformation Among Rural Bangladeshi Villagers During the Covid-19 Pandemic,” *Proceedings of the ACM on Human-Computer Interaction* 5 (2021): 1–34. Available: https://doi.org/10.1145/3479580.

58 “Gender-Based Disinformation: Advancing Our Understanding and Response,” *EU DisinfoLab*, October 20, 2021, https://www.disinfo.eu/publications/gender-based-disinformation-advancing-our-understanding-and-response/.

59 Richard Horton, “Offline: Science and the Breakdown of Trust,” *The Lancet* (2020). Available: https://doi.org/10.1016/S0140-6736(20)32064-X.

60 Rama Adithya Varanasi, Joyojeet Pal, and Aditya Vashistha, “Accost, Accede, or Amplify: Attitudes Towards Covid-19 Misinformation on WhatsApp in India,” *CHI22: Proceedings of the 2022 Conference on Human Factors in Computing Systems* (2022): 1–17. Available: https://doi.org/10.1145/3491102.3517588; Syeda Zainab Akbar, Anmol

Panda, and Joyojeet Pal, “Political Hazard: Misinformation in the 2019 Indian General Election Campaign,” *South Asian History and Culture* 13 no. 3 (2022). Available: https://doi.org/10.1080/19472498.2022.2095596; Chinmayi Arun, “On WhatsApp, Rumours, Lynchings, and the Indian Government,” *Economic and Political Weekly* 54 no. 6 (2019): 1–10. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3336127.

[61] Neelanjan Sircar, “Disinformation: A New Type of State-Sponsored Violence,” *The India Form*, September 15, 2021, https://www.theindiaforum.in/article/disinformation-new-type-state-sponsored-violence; Demitry Chernobrov, and Emma L. Briant, “Competing Propagandas: How the United States and Russia Represent Mutual Propaganda Activities,” *Politics* 42 no. 3 (2022): 393–409. Available: https://doi.org/10.1177/0263395720966171; Rong-Ching Chang, Chun-Ming Lai, Kai-Lai Chang, and Chu-Hsing Lin, “Dataset of Propaganda Techniques of the State-Sponsored Information Operation of the People’s Republic of China,” *Arxiv CS* (2021). Available: https://arxiv.org/abs/2106.07544.

[62] See “The Changing World of Digital in 2023,” *We Are Social*, January 26, 2023, https://wearesocial.com/ca-en/blog/2023/01/the-changing-world-of-digital-in-2023-2/.

[63] Lisa Gitelman, *Always Already New: Media, History, and the Data of Culture* (Cambridge, MA: MIT Press, 2006), 26.

[64] Also see my Aaron Mauro, *Hacking in the Humanities: Cybersecurity, Speculative Fiction, and Navigating a Digital Future* (London: Bloomsbury Publishing, 2022) for a more fulsome treatment of these issues.

[65] For example, see the United Nations and US National Archives frameworks for digital preservation: https://archives.un.org/content/managing-risk and https://www.archives.gov/preservation/digital-preservation/risk.

[66] Respectively, see https://www.lockss.org/ and https://alliancecan.ca/en/services/research-data-management.

[67] Proprietary software is not inherently more secure, since security through obscurity is simply not a viable means to assure software reliability. Open-source software faces supply-chain attacks in the open, particularly in software repositories like NPM, PyPI, RubyGems, and others. Alex Birsan has coined the term “Dependency Confusion” in a 2021 blog post that has done well to describe the scope of the problem. See Alex Birsan, “Dependency Confusion: How I Hacked Into Apple, Microsoft and Dozens of Other Companies,” *Medium*, February 9, 2021, https://medium.com/@alex.birsan/dependency-confusion-4a5d60fec610.

[68] The following section appeared previously in the journal *Pop!*. Please see Aaron Mauro, “Security Culture as an Expression of Values.” *Pop!* Public. Open. Participatory. no. 5 (2023), https://doi.org/10.54590/pop.2023.003.

[69] Corey Nachreiner, “The Perfect InfoSec Mindset: Paranoia + Skepticism,” *Dark Reading*, July 29, 2014, https://www.darkreading.com/operations/the-perfect-infosec-mindset-paranoia-skepticism.

[70] Take, for example, the case of Drs. Xiangguo Qiu and Keding Cheng, who were stripped of their security clearance at Canada’s National Microbiology Lab in July of 2019. While shrouded in secrecy and conflicting reports, CBC reporting implied national security and commercial secrets as possible concerns for their dismissal. See Karen Pauls and Kimberly Ivany, “Mystery Around 2 Fired Scientists Points to Larger Issues at Canada’s High-Security Lab, Former Colleagues Say,” *CBC*, July 8, 2021, https://www.cbc.ca/news/canada/manitoba/nml-scientists-speak-out-1.6090188. Elaine Dewar’s *On the Origin of the Deadliest Pandemic in 100 Years: An Investigation* (Windsor, ON: Biblioasis, 2021) goes into much more detail about the complex issues present in expelling foreign nationals during a national emergency.

[71] The US Patriot Act is an excellent example of the legal contexts that may contract research security cultures in the humanities because of the sweeping surveillance powers it affords the US government.

[72] Consider the security reviews required to collaborate with researchers in Communist held China. The Canadian Federal Government, through Innovation, Science and Economic Development Canada, issued new guidelines and procedures, *National Security Guidelines for Research Partnerships*, which describe conditions necessary to receive government support: https://science.gc.ca/site/science/en/safeguarding-your-research/guidelines-and-tools-implement-research-security/national-security-guidelines-research-partnerships. These guidelines have been drafted under principles such as academic freedom; institutional autonomy; freedom of expression; equity, diversity, and inclusion; research in the public interest; transparency; integrity; and collaboration. These principles match well with academic values in Canadian institutions, with the notable omission of decolonization. However, these principles do not easily match with a risk assessments and threat models that must assume hostile actions by some potential collaborators.

[73] Spafford's presentation was posted on August 30, 2019 as part of Purdue University's CERIAS research group (https://www.cerias.purdue.edu/).

[74] Eugene Spafford, "Rethinking Cyber Security" *YouTube*, August 30, 2019, https://www.youtube.com/watch?v=MI6pq4zIBx0: 26:49.

[75] See Morgan Currie, and Britt S. Paris, "How the 'guerrilla archivists' Saved History—and Are Doing It Again Under Trump," *The Conversation*, February 27, 2017, https://theconversation.com/how-the-guerrilla-archivists-saved-history-and-are-doing-it-again-under-trump-72346.

[76] Jerome McGann, *Radiant Textuality: Literary Studies After the World Wide Web* (New York: Palgrave MacMillan, 2001), 19.

[77] See respectively https://wiki.archiveteam.org/ and https://ddosecrets.com/wiki/Distributed_Denial_of_Secrets.

[78] Shawn Graham, *Failing Gloriously and Other Essays* (Grand Forks: The Digital Press at the University of North Dakota, 2019). Available for free here: https://thedigitalpress.org/failing-gloriously/.

[79] See the NIST Cybersecurity Framework: https://www.nist.gov/cyberframework.

[80] Meg Worley, “The Rhetoric of Disruption: What Are We Doing Here?,” in *Disrupting the Digital Humanities*. Dorothy Kim and Jesse Stommel (eds.) (Brooklyn, NY: Punctum Books, 2018): 72.

[81] Ibid., 68–9.

[82] Ibid., 73.

[83] Mauro, *Hacking in the Humanities*, 2022.

2

The Attribution Problem

In the classic 1983 computer science textbook, *Computers in Society*, Nancy Stern and Robert A. Stern explore the societal consequences of computing before the widespread adoption of personal computing and the internet. Nancy Stern's background in history and the liberal arts enabled her to find the "balance between technical concepts and humanistic issues."[1] This blend of technical and humanistic perspectives is articulated in the now common phrase "computer literacy." They sought to bridge *The Two Cultures* identified by C. P. Snow so long ago in 1957. As a philosopher, Snow proposed two separate spheres of intellectual endeavor defined by a "scientific, technological orientation" and an opposing "humanistic or liberal arts focus." Already in *Computers and Society* computing exposed the need to "bridge the ever-widening communication gap" and "move us closer to a single culture."[2]

The rise of Science and Technology Studies (STS) in numerous universities highlights the growing importance of interdisciplinary thinking. Scholarly societies have long aimed to bridge these disciplines; the Society for the History of Technology was founded in the United States in 1958, and the Society for the Social Studies of Science came into existence in 1975. The social and cultural consequences of technology are all around us now, and ChatGPT and other Large Language Models (LLMs) like it are just the latest in a series of disruptive technologies driven by computing, including the web, smartphones, and social media. For over forty years, the fundamental technological imaginary at work is the awareness that the “only way that intelligence can be demonstrated is with the use of some method of communication. Moreover, the ability to understand key ingredient of intelligence, requires one to understand language.”[3] If we consider Nancy Stern’s insights on language and intelligence, it becomes clear that many computing challenges can be framed as problems in natural language understanding.

The current crisis in disinformation and online manipulation is an excellent application of automated content moderation powered by natural language understanding. Recently, there has been a growing call for incorporating a humanist perspective into security intelligence and combating disinformation. Defense and security experts are increasingly citing the humanities as a tool for fostering critical, close reading skills. One such expert notes, “The reason we are having trouble getting on top of disinformation is because we are mislabelling it and therefore misunderstanding the phenomena. We are not dealing with simply wrong information. We are dealing with weaponized information in story form.”[4] To effectively address disinformation campaigns, a more nuanced understanding of the relationship between narrative and national security is essential:

> There is nothing inherently seductive about information, whether true or false. That is why we use storytelling. Stories play a special role in human cognition. Our brains are receptive to stories, especially stories about ourselves, or stories that we see ourselves in, or project ourselves into. And we are particularly receptive to stories that speak to our preferred identities especially when we feel our identity is under threat and the story gives us a way out—a way to respond to the threat.[5]

Facts alone are not a natural antidote to falsehoods, and more information is not necessarily the solution to disinformation. Influence is shaped by the rhetorical context in which communication occurs. There is no guarantee that an audience will consume, comprehend, and choose to act on information in an ethical and safe manner. The task of identifying key features of online content is further complicated by the sheer volume and variability of available information.

Kenneth Burke, the eminent twentieth-century American rhetorician, prioritized *identification* over persuasion, diverging from the classic Aristotelian focus. According to Burke, miscommunication is not just a technical issue but a moral one; he believed that our human nature—what he describes as a postlapsarian, tragically fallen state—hinders effective communication, driving us to seek *identification* with others. In *The*

Rhetoric of Motives, Burke envisions a utopian state of perfect understanding, stating, “In pure identification there would be no strife. Likewise, there would be no strife in absolute separateness, since opponents can join battle only through a mediatory ground that makes their communication possible, thus providing the first condition necessary for their interchange of blows.”[6] For Burke, our proximate misidentification literally leads to conflict. In a Burkean sense, deceit, manipulation, and lies are a symptom of misinformation, disinformation, and propaganda. Jumping to the realm of cyberwarfare, it is this weaponization of deceit that triggers malicious access and attacks.[7] The escalation of Russian disinformation prior to their invasion allowed Ukrainian defense intelligence to anticipate the move months in advance.[8]

There are now many classic works in cybersecurity that might be productively paired with the likes of Burke, and Thomas Rid’s *Cyber War Will Not Take Place* (2013) is one that requires thorough reappraisal in the wake of the Russian invasion of Ukraine.[9] Rid’s central argument, declared in the title of his book, has proven remarkably durable. Even in the face of the Russian invasion of Ukraine in 2022, where cyberattacks played a significant role, a cyber war did assault the country’s cultural institutions, but it was also a hot war with shelling, tanks, and a long grinding retreat. I will offer a more complete treatment of this period in Chapter 4, where the disinformation campaign, satellite hacking, and continued attacks on Ukrainian utilities were a significant feature of the early stages of the invasion.[10] One enduring insight from Rid’s 2013 analysis, which still holds true, is that cyberattacks generally represent a less violent form of conflict, minimizing damage to property and loss of human life.

Whereas traditional warfare and espionage often put soldiers and spies in harm’s way—sometimes at the expense of numerous other lives, both military and civilian—cyberwarfare now allows for the extraction or destruction of sensitive information without the need for elaborate spy craft or the dramatic tactics often depicted in film and television. Thomas Rid argues that because political sabotage, espionage, and subversion can be executed through technical means, “cyber attacks help to diminish rather than accentuate political violence.”[11] Identifying the perpetrators of a cyberattack is extremely difficult to achieve for a number of reasons, not least of which is that the very basic protocols of the internet—namely, the

Transmission Control Protocol and the Internet Protocol, collectively known as TCP/IP—were designed as open standards. Idealistic academics rolling out the early ARPANET simply did not consider security questions and relied on the goodwill and trust of known colleagues. Today, we still live with these early design decisions that allow numerous methods for routing traffic through low-level proxy servers, including TOR, allowing attackers to effectively conceal their location and true identity.

In a 2009 Citizen Lab report, the so-called GhostNet attacks were attributed to China. These attacks targeted high-value diplomatic, political, economic, and military entitles across 103 countries, including the private offices of the Dalai Lama and other Tibetan interests.[12] This report serves as a reminder to Thomas Rid about the complexities of the *attribution problem*. The problem that Burke might simply understand as *identification*. The authors of the report clearly articulate this ongoing challenge for cybersecurity defenders.:

> Hand-in-hand with the problem of attribution is the difficulty of identifying motivating factors behind a cyber attack. Many perpetrators of Internet-based attacks and exploits are individuals whose motivation can vary from a simple profit motive through to fear of prosecution or *strong emotional feelings*, including religious belief and nationalism. Many cyber attacks and exploits which seem to benefit states may be the work of third-party actors operating under a variety of motivations. This makes it difficult to separate the motivation of the individual from the potential motives of the party on whose behalf the attacks have occurred, or a prospective client to which the perpetrator is trying to market his or her wares. In either case, the challenge of identifying perpetrators and understanding their motives gives state actors convenient plausible deniability and the ability to officially distance themselves from attacks.[13]

Attribution in cybersecurity is a multifaceted process that requires a balanced attention to technical, social, and political “layers” of evidence. Like human intelligence gathering, the classic motivations for hostile cyberattacks include Money, Ideology, Compromise|Coercion, and Ego (MICE).[14] In signals intelligence, technical indicators of compromise may include IP addresses, known proxies, unusual network traffic, unfamiliar applications in the system, atypical activity from administrator accounts, an increased number of incorrect login attempts, unusual database or file read activity, a large volume of compressed files being sent over the network, and unexpected changes to system settings. These technical indicators often manifest as anomalous spikes in behavior. The remaining MICE motivations for hostile actions are generally rooted in social and cultural factors.

Social indicators of compromise may include access reported from atypical geographic locations or during unusual hours, which could also imply access from a different time zone. Suspicious behavior among staff could indicate either malicious intent or unprofessional conduct, often linked to poor social cohesion and low team morale. Associating technical resources and behaviors with known users is crucial for maintaining situational awareness of individual actions. “The [social] layer is very hard,” Rid reflects, “if not impossible, to penetrate.”[15] It is for this reason that attribution is arguably the most challenging aspect of cybersecurity.

Cyberattacks often function as malicious imitation games, where hostile agents use potential, yet innocent, third parties as decoys to mislead targets about the true source of an attack. Both intention and opportunity contribute to the fog of war in this digital landscape. Accurate attribution is difficult and often necessitates technical analysis of various factors such as source code, timing of operations, IP addresses, and other activity markers. The humanities can play a vital role in honing the human critical faculties needed to discern intention in online communication, to augment Security Information and Event Management (SIEM) tools. It is human sensibility that remains capable of interpreting both language and the multicultural historical nuances that underlie the assumptions and attitudes manifest in cyberattacks. After questions related to how an attack is conducted, questions quickly turn to the who and the why.

The political layer of attribution is most commonly employed to ascribe attacks to nation-state actors and ideologically motivated protests. In Citizen Lab’s GhostNet report, the authors recognize the gravity of accusing a nation of hostile actions, especially when “the most obvious explanation, and certainly the one in which the circumstantial evidence tilts the strongest.”[16] Simply because the list of affected systems include high-value political and economic targets that align with the foreign and defense policies of a hostile nation does not automatically implicate that nation as the attacker. Apparent ideological motivations can easily serve as a smokescreen to divert attention away from the true perpetrator.

There may be a third actor exploiting tensions between nations as cover. A series of uneasy questions arise: Who stands to benefit from escalating political tensions? Who gains financially, politically, or socially? If the political attack aims to distort public perceptions, who is the intended audience for these manipulations, whether domestic or foreign? The motivation behind the attack may be confined to a small group or individual, whether inside, aligned with, or outside of national-level cybersecurity agencies. Complicating matters further, the perceived threat actors seemingly responsible for an attack may actually be the result of cherry-picked, incomplete, or random data masquerading as evidence. Attribution is the fog of cyberwar.

Rid contends that "the attribution problem" lies at the "root of cyber security" and is fundamentally a "political rather than a technical problem."[17] Regardless of the method of intelligence gathering—be it human intelligence collected in person, signals intelligence gathered online, or open-source intelligence from publicly accessible websites—there is an inherent limitation to attribution. A solid grasp of the status of evidence and standards of proof is essential for an effective response. Coordinated human effort is required to reorganize an IT department's workweek for tasks such as system patching and password resets; those actions may constitute offensive measures, sometimes referred to in the US defense agencies as "defending forward," which would reorder global political and economic relations.[18] Defending forward uses offensive cyber capabilities to disrupt threat actors in advance, but defending forward requires identifying hostile actors prior to an attack. Establishing and maintaining access to threat actor systems requires an evolving, progressive scan that anticipates future hostile actions. "Attribution is," as Rid concludes, "always a call of judgement."[19] Rid's *one weird trick* is human judgement.[20] There can be, it must be said, grave consequences to what amounts to a judgment call.

A false attribution could invite diplomatic repercussions or covert retaliation. The standard of proof required for an accurate attribution simply cannot meet the threshold of legal judgment, as establishing fact "beyond a reasonable doubt" is challenging when doubt is almost always reasonable.[21] Counterfactual thinking, the ability to ask "what if" and envision alternative scenarios, can further complicate the attribution problem by introducing cascading imaginative possibilities. Logs and code are also insufficient as stable forms of evidence for attribution. Lawrence Lessig's first coining of "code is law" was not intended for the twenty-first-century media environment.[22] Just because a program or system permits certain behavior does not mean it condones it. Lessig's waxing about liberty has more to do with the freedom of corporations than individuals.[23] Lessig's commentary on liberty is more concerned with corporate freedom than individual rights, a perspective that resonates with cryptocurrency enthusiasts.[24] Code is not definitive proof; it can be copied, spoofed, and reverse-engineered. While Rid emphasizes the importance of sound critical judgment, he may have overlooked other disciplines that could help narrow the uncertainty inherent in the attribution problem. The technical, social, and political layers of this problem are well-understood within the existing security establishment, but the rise of disinformation and misinformation on social media over the past decade suggests an additional cultural layer in need of exploration.

The attribution problem is a central issue animating security rhetoric. It represents the murky shallows of the internet where both criminals and government agents thrive on deception. Tracing the threads to identify the involved threat actors and accounting for the possibility of false attribution is often an insurmountable challenge. However, the level of certainty tends to increase with the severity of the attack. Large-scale attacks are more noticeable, and the absence of covert operations often points to an obvious perpetrator, as was the case with the GhostNet attacks. This relationship between the severity of an attack and the likelihood of accurate attribution can be summarized as follows: "*the attribution problem is a function of an attack's severity*."[25] In other words, a larger and more overt attack is often more easily attributable. Conversely, with smaller-scale and more covert attacks, the risk of falsely attributing the action to a hostile but innocent party is heightened, especially when known tensions can serve as a cover for a third hostile actor.

While the rhetoric surrounding security is multifaceted, encompassing layers of technical, social, political, and now cultural elements, these layers aim to compromise vulnerable individuals, demographics, and institutions. In cybersecurity, a vulnerability can be described as a point of ingress that allows an attacker to cross a trust boundary and subvert authentication. Trust is not just a concept; it is a feeling often articulated within groups and between individuals. The National Institute of Standards and Technology (NIST), the author of the Cybersecurity Framework, defines a vulnerability more succinctly as: "Weakness in an information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat source."[26] In this context, the NIST Cybersecurity Framework helps apply security rhetoric most critically to the work of identifying threats. Cultural vulnerability might be best described in an intersectional framework that pays careful attention to how overlapping inequities and systemic vulnerabilities make some targets more susceptible to cyberattacks.

Affective Cybersecurity

In classic Burkean rhetoric, human connection and understanding are achieved through identifying with others via direct human communication. When we communicate, we not only rationally understand meaning but also gauge its authenticity based on similar past experiences. Sometimes communications just feel wrong. This process of identification is deeply validating and meaningful, serving as the foundation for cultural belonging. However, when words are designed to disguise harm, the manner in which cultures are shared and reinforced—through a reciprocal exchange of shared language and codes over common systems and platforms—becomes suspect. This corrupts the exchange of meaning, introducing cynicism and uncertainty, and ultimately violating trust.

The term "intersectionality" was coined by Kimberlé W. Crenshaw in 1989 to describe how multiple, intersecting identities contribute to complex forms of discrimination. Published by the Columbia Law School's Faculty Scholarship journal, her seminal work "Demarginalizing the Intersection of Race and Sex" critiqued how white feminist theory perpetuates repressive systems that fail to capture the experiences of Black women. According to Crenshaw, the intersection of race and gender creates unique experiences of injustice within a legal system designed to protect white women and their chastity. This system often fails to extend even these imperfect protections to Black women, which is made worse by "the historical fact that the protection of white female sexuality was often the pretext for terrorizing the Black community," like, Crenshaw reminds us, in the case of Emmitt Till.[27]

In her work *Digital Black Feminism*, Catherine Knight Steele describes the evolution of the term intersectionality in this way: "the use of *intersectionality* has traversed far from its original meaning. It has become a catchall term used by many new to Black feminist thought to signify multiple identities or different perspectives, to signal the inclusion of women of color, or as a descriptor of the ways that everyone has competing points of privilege."[28] There is a risk of erasing Black feminist origins of the term through its use in articulating the affective qualities of security rhetoric, so any invocation of the term must acknowledge the origins of this method of understanding power and oppression from Black feminist theoretical, political, and emotional work. In this way, security rhetoric works in solidarity with intersectional cybersecurity. It is for this reason that Steele calls back to Patricia Hill Collins's coining of the "matrix of domination" outlined in *Black Feminist Thought* (2009). This term possesses a "rhetorical utility" in the phrase, wherein the "'matrix of domination' resists the appropriation, misuse, and memeification of intersectionality in popular culture. The phrase requires speakers to attend to unequal power distribution and white male supremacy."[29] There are emerging intersectional approaches to cybersecurity discussed in a moment.

Intersectional approaches offer valuable insights into how social, cultural, and political forces exert influence through disparities in power and entrenched systems of privilege, even in cyberspace. The backlash against Critical Race Theory (CRT) in the United States serves as a reaction to this widening understanding of injustice and complicity. This reaction opposes both single-issue politics and a narrow interpretation of social relationships. The resistance to viewing race as a deeply ingrained, systemic issue has sparked protests nationwide and has become a shorthand for conservative efforts to influence public education.[30] "It is a way of seeing, attending to, accounting for, tracing and analysing the ways that race is produced," Crenshaw explained in an interview, "the ways that racial inequality is facilitated, and the ways that our history has created these inequalities that now can be almost effortlessly reproduced unless we attend to the existence of these inequalities."[31]

Injustice can be understood logically, but these grinding inequalities are also deeply felt. Increasingly, US conservatism seems to be targeting this basic sense of empathy and compassion. Recently, "social-emotional learning" has become a political battleground over education. Social Emotional Learning (SEL) promotes emotional maturity in young people and higher executive function.[32] Organizations like Parents Defending Education rail against the "harmful agendas" of long-standing fixtures of American education like Title IX, Pride, CRT, and SEL.[33] They fight the "IndoctriNation" of young people by proclaiming, "there are problems everywhere … but there are also people fighting back."[34] Make no mistake, emotional maturity invites compassion and care, which leads quite logically to policies that promote greater equity and safety. This affective antipathy finds root in political conversations and are exported globally through networks of conservative influence.[35] The partisan politicization of emotional life appears in primary as readily as higher education.

In the months following the election of the forty-fifth president of the United States, Lisa Feldman Barrett published *How Emotions Are Made*, which made the case that chronic stress, often resembling a persistent form of PTSD, is detrimental to one's health.[36] Summarizing her well-received book in a very short *New York Times* article dated July 14, 2017, she wrote, "Words can have a powerful effect on your nervous system. Certain types of adversity, even those involving no physical contact, can make you sick, alter your brain—even kill neurons—and shorten your life." According to Barrett, as the body responds to stress, inflammatory cytokines can trigger autoimmune disorders and accelerate the aging process by shortening the telomeres at the ends of your chromosomes.[37] In retrospect, considering the Covid-19 pandemic, this alignment of emotional and bodily harm should resonate strongly with anyone who has endured prolonged lockdowns and years of chronic stress.

In times of crisis, emotions often come to the forefront of our experiences and shape our reality. Barrett explains how language and emotions shape our sense of community and security. Shared words and concepts to describe feeling and experience are the very foundation for society, which she describes as a "collective intentionality." It is a feature of human society because we can understand language as a representation of another person's feelings: "Emotions are social reality. We construct instances of emotion in exactly the same manner as colors, falling trees, and money: using a conceptual system that is realized within the brain's wiring." Language not only constructs an emotional reality but also has the potential to cause harm through collective intentionality predicated on shunning and exclusion. This causal relationship between violent, coercive language becomes evident on a macro scale, particularly in the realms of social engineering and surveillance capitalism.

But, in the days following the 2016 US election, some people were quick to defend the First Amendment rights of those in power.[38] This stance perhaps masks an ongoing debate about whether harmful speech should be categorized as a form of violence, a debate complicated by a general public perception that free speech is unequivocally good. Jonathan Haidt and Greg Lukianoff reduced the work of Barrett to a syllogism because they rejected the intersectional nature of chronic stress, which does indeed cause physical harm and shorten lifespans.[39] Opting for a strictly causal burden of proof to demonstrate harm, Haidt and Lukianoff stop short of labeling anyone as overly sensitive. They argue that the presence of right-wing provocateurs on college campuses is justifiable, dismissing them as mere passing stressors. According to this view, these provocateurs are akin to fleeting political weather patterns and should not be confused with the broader cultural climate—suggesting that one should leave the metaphorical parka at home to avoid being labeled a snowflake.

In her chapter on emotions and the law, Barrett outlines the ramifications of a simplistic understanding of emotion, especially when such emotional states are invoked to justify violent behavior. She critiques the "classical view of emotion," which is rooted in essentialist perspectives on human experience. According to this view, humans possess both a rational mind and an emotional mind. Not only does this dualistic framework serve as a sexist and racist shorthand, but it also provides a convenient excuse for police officers who claim to kill out of fear. Barrett argues, "The science of emotion is a convenient flashlight for illuminating some of the law's long-held assumptions about human nature—assumptions that we now know are not respected by the architecture of the human brain. People don't have a rational side and an emotional side, with the former regulating the latter."[40] Acknowledging the unity of feeling and meaning grants a view to the potential harm manifests through toxic content online. The popular understanding of emotion is a political lightening rod because this antiquated dualism that divides the rational and emotional mind justifies these intersectional harms.

The memeification of politics and propaganda preys first on emotion then on reason. In *Meme Wars*, Joan Donovan, Emily Dreyfuss and Brian Friedberg have described this as moving "from the wires to the weeds," where online activities move into public space: "This recursive cycle is a meme war."[41] In such wars, the aim is not so much to establish coherence as it is to occupy space. The correlation between language and harm is expressed in the correlation between memetics and warfare. Meme wars often rely on emotional immaturity, using anger and hostility to obscure care and compassion. The tragic nadir of a historical period marked by social and cultural division and extremism—from the Tea Party and #Gamergate to Dylan Roof and Kyle Rittenhouse—was the January 6 attack on the Capitol. “Meme Wars” cites the Trump speech that incited the violence:

> If you don’t fight like hell, you’re not going to have a country anymore […] So we’re going to walk down Pennsylvania Avenue—I love Pennsylvania Avenue—and we’re going to the Capitol, and we’re going to try and give our Republicans—the weak ones, because the strong ones don’t need any of our help—we’re going to try and give them the kind of pride and boldness that they need to take back our country.[42]

Senator Lindsey Graham later defended the speech, effectively condoning a toxic culture that shirks responsibility for inciting violence. He justified this by stating that “every politician has used the word ‘fight,’ ‘fight hard,’” as if to use the aggressive tone of political rhetoric as a shield.[43] The world has grown accustomed to conflict-oriented language that prioritizes fighting as a solution to problems. Astute readers will also notice the use of the first-person plural “we,” which implicitly pits supporters against an unnamed “other” group. This is a classic example of dog-whistle politics. Trump’s supporters hear what they want to hear, a sentiment amplified by social media backchannels that frame any opponent of “us” as “them.” This culture of conflict is counterbalanced by appeals to love, boldness, assistance, pride, and strength, all aimed at some amorphous need to “take back our country.” The political climate serves as a telling symptom of a broader normalization of violent language that condones, valorizes, and ultimately authorizes violent political actions.

Some might dismiss this as mere moralizing—pearl-clutching over the not-nice words. However, the metaphorical frog is boiled by ignoring the gradually rising temperature of rhetoric until it reaches a deadly boiling point. The “with us or against us” mentality of the Bush era has evolved into a surge of domestic terrorism. The chronic stress of so much feeling—of so much terror—has seemingly triggered the political equivalent of an autoimmune response, where the body politic begins attacking itself for lack of an external threat. A 2023 report from the US Government Accountability Office highlighted a dramatic increase in domestic terrorism cases, rising from 93 in 2015 to 449 in 2021. The majority of these cases are racially motivated and include any “criminal acts dangerous to human life on US soil that appear intended to coerce a civilian population or influence or affect the conduct of government.”[44] Racialized domestic terrorism can be heard in the dog whistle of “take our country” back. The normalization of vilifying opponents in Conservative politics serves as a form of coercive white supremacy, paving the way for violent actions to follow violent language.

Gabor Maté’s *The Myth of Normal: Trauma, Illness and Healing in a Toxic Culture* explores how toxic environments in work, family, and politics shorten lives through traumatic responses to chronic stress and the relentless pressures of online life. “Politicians make policy,” says Maté, “and policy creates or cements the very cultural conditions we know are antithetical to our health.”[45] Maté frames two intersectionally related social determinants of health, race and class, as a trauma sustained through a politically acceptable domestic terrorism. “As the Black American writer Ta-Nehisi Coates tersely asserts, ‘Race is the child of racism, not the father.’ In other words, the very concept of race emerges from the distorted imagination of the racist. Though racism’s impacts are real, in phycological or genetic terms race does not exist.”[46] The normalization of racist language and ideas leads to tangible harm, fostering societal conditions that enable terrorism and inflict profound damage on individuals.

Bob Altemeyer’s work may offer a way to synthesize these recent examples that highlight the psychological aspects of historical events. As a psychologist, Altemeyer focused on defining the psychological traits of Right-Wing Authoritarians (RWAs). These traits often point to emotional immaturity and are characterized by zealous ethnocentrism, a fear of the broader world, a high degree of self-righteousness, aggressiveness, dogmatic and contradictory belief systems, a dependency on social reinforcement of those beliefs, and susceptibility to manipulation.[47] Altemeyer describes RWA attitudes through systematic questionnaires conducted over decades in his 1998 book, *The Authoritarian Spectre:*

> High RWAs, accordingly, can be easily frightened, which makes them vulnerable to precisely the kind of overstated, emotional, and dangerous assertions a demagogue would make. So how hard would it be for a sufficiently unscrupulous, power hungry, *real* agitator to turn authoritarians’ general anti-Semitism into the Nuremberg Laws? Or for a washed-up, unprincipled senator from Wisconsin to turn Cold War fears into a life-crushing, four-year witch hunt? Or to get [High RWAs] to join a “posse” after any vulnerable group today?[48]

The psychological fervor of McCarthyism, as evoked by Altemeyer, resonates easily with contemporary authoritarian slogans such as “Muslim ban,” “build the wall,” “Southern White House,” “very fine people on both sides," and "fight like hell" from January 6, 2021 onward. Phrases like these serve as dog whistles, signaling a sense of belonging through the social reinforcement of beliefs, often accompanied by highly emotional and self-righteous sentiments.

Layers of cultural context overlap and intersect in ways that are not immediately apparent from overused Trumpian slogans. Dog whistle politics means these terms are deeply implicit and are often only understood by the ideologically initiated. Connecting the erosion of social emotional learning curriculum to the memeification of politics and simplistic views of free speech are not always possible in real time, but the attribution of these cultural forces is the challenge posed by security rhetoric.

The failure to recognize intersections of security rhetoric as harmful is symptomatic of a diminished capacity for Burkean identification. Words can cause harm on the level of the individual and within broader society, particularly when many people dismiss, underestimate, or remain unaware of the damaging language around them. Security rhetoric plays a critical role in both identifying harmful systems as well as in designing policies that protect and care for vulnerable people and the systems they inhabit. Security policies often perpetuate structural issues related to access to support and education, leaving vulnerable those who lack the systems and infrastructure for online protection. While it seems intuitive to allocate appropriate training and resources for corporate, governmental, or military personnel, there is less emphasis on protecting communities that face systemic and intersectional inequities due to factors like race, class, disability, gender, sexuality, and others.

By invoking the qualities of feeling, emotion, and affect in a security context, I am working to align the interdisciplinary field of "affect theory," which joins neuropsychology with the humanities. This approach seeks to understand the aesthetic, historical, and philosophical implications of recent advancements in psychology to help redefine social engineering. Two seminal works in this field—Brian Massumi's *Parables for the Virtual: Movement, Affect, Sensation* and William E. Connolly's *Neuropolitics: Thinking, Culture, Speed*—explore the interplay between language, individual experience of feeling, social emotion, and non-conscious, pre-personal affect.[49] While I concur with Massumi's assertion that there is a "growing feeling within media, literary, and art theory that affect is central to an understanding of our information and image-based late capitalist culture,"[50] I take Ruth Leys's critique of affect theory seriously in her essay "The Turn to Affect" in *Critical Inquiry* from 2011. Leys carefully correlates the emotional schemas used by *affect theory* with the growing understanding from experimental neuropsychology advances with fMRI (functional Magnetic Resonance Imaging) and Positron Emission Tomography (PET) scanning. In summary, Leys argues that the ontological foundations of affect theory are increasingly at odds with contemporary neuropsychological understanding. The science no longer supports the theory humanists have built atop once cutting-edge research.

A similar disconnection occurred in the heyday of critical theory in the 1980s and 1990s, when Freudian and Lacanian psychoanalysis, employed as literary tools, diverged significantly from the clinical practice of actually treating patients. Even as Massumi and Connolly invoked the empirical rigor of scientific observation in their analyses of cultural objects and events, rapid advancements in the sciences soon rendered these humanistic theoretical frameworks obsolete. It is important to celebrate this obsolescence because "affect theory" proposed a dangerous carve out for feeling, as something independent of rationale thought. It relegates feeling and emotion as an autonomic response system rather than a sensible and valuable way of identifying hostile ideology. Leys cautions against the notion that affect should "be viewed as independent of, and in an important sense prior to, ideology."[51] In the next chapter, it will be important to explore how this theoretical form of psychology has found footing in cybersecurity practice in the form of social engineering and the dubious claims of their most well-known proponents, Christopher J. Hadnagy.

For my part, the psychoanalytic view of affect has become a tool to trigger ideologically driven acts that are malicious, manipulative, or cruel. When regarded at the scale of cyberattacks initiated by online lures, affect transcends Massumi's notion of a "virtual point of view" and becomes a tangible, global cybersecurity attack vector.[52] The fact that these affective phishing schemes lead us to unwittingly compromise our own interests doesn't necessarily render them "subconscious." We are more vulnerable when our behavior can be almost automatically triggered. Enhancing our digital literacy will empower us to approach affective media with a discerning mindset. Our collective digital defense hinges on our ability to calmly disregard malicious messages.

It is perhaps because we are so readily manipulated by “our information and image-based late capitalist culture” that Massumi argues that affect is simply beyond our conscious understanding. However, this cannot be accurate, as words do cause observable social and emotional harm both collectively and individually. It is emotionally taxing at every level to be online often. This is exacerbated by a reactive media environment that not only encourages trolling but also desensitizes us to violent rhetoric and actions. Disinformation and misinformation coexist with reliable and trusted sources, making the media landscape a polluted well that we distrust yet are compelled to consume to avoid disconnection and cultural exile. The toxicity of online culture does not preclude our ability to understand it and to make meaningful interventions for our safety. A practical grasp of the manipulative potential of affect, emotion, and feeling can be incorporated into a broader, human-centered cybersecurity practice. The intentions behind cyberattacks are discernible, and the ways we are lured by affective appeals are part of our everyday experience. It is within this intersectional framework that the current discussion circles back to the issue of phishing and lures in general.

Estimative Uncertainty

Rather than focusing solely on valuable assets, an intersectional approach takes into account how socially defined categories such as gender, race, ethnicity, disability, sexual orientation, and class intersect to compound an individual’s experience of threats. In March 2022, Marissa Conway and Nehmat Kaur authored *The Intersectionality and Cybersecurity Toolkit*, published by the Centre for Feminist Foreign Policy in the UK. Along with a regularly updated online portal, the project works to “reconceptualise cybersecurity’s purpose as protecting people.”[53] Traditional cybersecurity policy and practices often employ a risk management approach that prioritizes data, software, and physical assets like corporate products, government secrets, and intellectual property.

However, an intersectional approach makes cybersecurity personal and immediate. It becomes about my security. It shifts the focus to individual security, emphasizing that risks are specific to one’s occupation, relationships, and habits. In this way, intersectional approach to security is a practical and realistic approach to online threats. An intersectional approach reorients security practice to protect *assets*, which include people and spaces that gain intrinsic value by serving human needs:

> Human security disrupts traditional and mainstream ideas about security as state focused, but an intersectional lens is necessary to further point to how modern social, political, and economic systems often function to prevent marginalised people from feeling safe and secure and how policy has the power to exacerbate or reconcile this.[54]

This perspective is influenced by an activist approach outlined in the *Holistic Security Handbook*, published by the Tactical Technology Collective and informed by the UK Government's National Cyber Strategy, published in 2022.[55] Challenging the mainstream cybersecurity focus on state-sponsored attacks against corporate and military targets, an intersectional approach to cybersecurity argues that everyday citizens should be equipped with a broad awareness that security culture is an expression of societal values; moreover, individuals must be prepared to intuitively recognize hostile intentions in their email and social media feeds. Security is a collective endeavor, requiring that anyone with access to potentially threatened systems be capable of critical thinking when encountering malicious messages.

In his seminal 1964 report on the "estimative uncertainty" of language in CIA reporting, Sherman Kent discussed the inherent challenges of intelligence analysis. Kent, a Yale University history professor during the Second World War who later served with the CIA throughout the Cold War, is often hailed as the "father of intelligence analysis" and the "master of historical methods and salty quips."[56] He should be considered a pioneer in applying humanities-based approaches to security analysis. Kent characterized the difficulty of security analysis as the task of defining "something which is knowable in terms of the human understanding but not precisely known by the [person] who is talking about it."[57] He famously summarized these challenges through an "odds table," which aimed to quantify the relative certainty and probability of an event occurring. Kent attributed the problems related to linguistic ambiguity to the "two cultures" of poets and mathematicians:

> What slowed me up in the first instance was the firm and reasoned resistance of some of my colleagues. Quite figuratively I am going to call them the "poets"—as opposed to the "mathematicians"—in my circle of associates, and if the term conveys a modicum of disapprobation on my part, that is what I want it to do. Their attitude toward the problem of

> communication seems to be fundamentally defeatist. They appear to believe the most a writer can achieve when working in a speculative area of human affairs is communication in only the broadest general sense. If he gets the wrong message across or no message at all-well, that is life.[58]

Kent's concept of "estimative uncertainty" serves to delineate both "a statement of indisputable fact" as well as "something which is knowable in terms of human understanding but not precisely known" by the analyst.[59] Kent outlines a continuum of probability in his estimative language model, ranging from certainty to a "general area of possibility."[60] A cybersecurity incident may be a historical fact, in that we have evidence that it is occurring through many indicators of compromise.[61] There remains, however, the bias of anthropocentric aesthetic judgments about the scope, scale, and attribution of the event. The emotional and affective experience of an attack could potentially compromise the accuracy of this analysis, which is why Kent introduced the term "estimative uncertainty."[62]

It is an elegantly concise term that captures the essence of measuring what is unknown. Like Kent, I believe "there is a point which the poets can make with telling effect." That is to say that conveying the reality of a situation may necessitate language that imbues a technical or technocratic description with emotional significance. Security rhetoric would demand that understanding the context of lures—the malicious messages that we so often experience as phishing emails in our inbox—is a fundamental literacy in the twenty-first century. However, like all forms of literacy, the ability to read for threats is only half the equation. True literacy also demands a balanced proficiency and judgment in articulating the emotional urgency of a cyberattack. Journalists covering cybersecurity often grapple with this balance, striving to meld emotional urgency, which spurs attention and action, with technical accuracy, which ensures honest risk assessment. This equilibrium is crucial for both local and global contexts, as the emotional weight of a threat often varies with its proximity.

Kent concludes with a rather poetic refrain in his conclusion, exposing perhaps his predilection to refined prose. "Let us isolate and seize upon exactly the thing that needs estimating," he urges his CIA colleagues. "Let us endeavor to make clear to the reader that the passage in question is of critical importance the gut estimate, as we call it among ourselves," he continues, underscoring the emotional weight of the report. Finally, Kent's conclusion echoes Rid's own uneasy stance on the attribution problem: "Let the judgement be unmistakable and let it be unmistakably ours."[63] It appears that even within the CIA, security analysis is a deeply personal endeavor.

The judgment of the analyst remains the final arbiter in a low-information environment characterized by imperfect records and messy data. This refined sense of judgment must be applied specifically and resist cascading conclusions. The expression of analytical judgment must also carry enough affective, poetic clarity to make the urgency of a critical situation known. Finally, the analyst must accept responsibility for the use of their judgment and the decisions that emerge from those interpretations, whether the outcomes are good or bad. It seems that the aesthetic opposition between the mathematicians and the poets joins quite amicably around a clear understanding of a specific reality and an accurate description of the urgency of a judgment call. A key feature of Kent's "words of estimative probability" is that they also serve as a gauge of responsible judgment, if they later prove to be accurate reflections of reality.

Online threat environments are almost always imperfect, yet defenders are required to respond in real-time. Words of estimative probability serve as a heuristic technique, offering an approximate understanding and a sufficient response in a crisis. However, this model of argumentation can perpetuate prejudice, stereotypes, and generalizations, reinforcing the biases of those in positions of judgment. The word *heuristic* is derived from the Ancient Greek, *heurisko*, which means to find or discover. This sense of discovery and responsibility encapsulates the humanistic approach to thinking and argumentation that lies at the heart of cybersecurity analysis. The analyst's role is inherently exploratory, generating new insights through synthesis of desperate sources. Heuristic techniques are employed across various disciplines, including psychology, computer science, law, philosophy, and rhetoric—the study of argument.

The history of rhetoric is a long one that can still be productively approached through the likes of Aristotle, Kenneth Burke, and Stephen E. Toulmin.[64] Aristotle defined rhetoric as "an ability, in each particular case, to see the available means of persuasion," which still accurately describes the twenty-first century lure.[65] Burke was ready to extend this persuasion as the "exploitation of opinion" including "verbal deception."[66] There is little definite or assured in these definitions. In making an argument and seeking to sway an audience to understand and accept it as true, rhetoric works only, if ever, by way of probability. In his 1958 work, *The Uses of Argument*, Toulmin presented a model of rhetoric that acknowledges the imperfections and fluidity of situations and relationships as the basis for meaning and truth, as they emerge through the act of communication. Emphasizing the human capacity for judgment, Toulmin used detailed, almost narrative-like scenarios to illustrate how arguments are essentially imperfect, practical expressions of "generalized jurisprudence" and logic that will resonate with a reasonable proportion of discerning readers.[67]

A pivotal set of questions again come to the foreground: whose judgment is privileged and valued? What decisions emerge? What systems are supported? Whose assumptions matter? For Toulmin's part, rhetoric is bound by many academic disciplines, so his answer to these questions would have been hedged. His model is anti-idealist, rejecting Platonic formal logic, and accommodates both "field-dependent" arguments, which are context-specific, and "field-invariant" arguments, which are not.[68]

Toulmin's model of rhetorical argument is both practical and applied, making it particularly useful for understanding the logic behind online lures. He described the process of determining the truth value of a statement as navigating a "labyrinth of probability."[69] These hedged arguments are made in a "guarded" way that insures the person making the argument "against some of the consequences of failure."[70] There is something problematic for Toulmin about the "probabilification" of speech.[71] While Toulmin frequently employs legal and mathematical frameworks in *The Uses of Argument,* he also makes extensive use of poetic metaphors and analogies.[72] Toulmin admits that "extra-scientific use of the term 'probability' may well harbour gross fallacies also."[73] He resists "an excessive respect for mathematics" by acknowledging that any statement cannot possibly be made "beyond all reach of possible future amendment."[74] If absolute certainty is the goal then we will only seek those things we can speak about absolutely. Toulmin argues that the "trustworthiness" of the speaker is an inadequate measure due to the limitations of individual perspectives and inherent biases. Instead of framing lies and untruths in terms of intentional deceit, Toulmin addresses them through the language of probability—a perspective that gains added relevance given recent political discourse on falsehoods.[75]

Prior to Toulmin's applied logic, rhetoricians worked in abstract logical proofs. These formal logical expressions describe the propositions with deductive reasoning; logicians seek inferences that are deductively valid by following logical premises to be either true or false. Take for instance the Aristotelian syllogism, "Socrates is a man. All men are mortal. Socrates is mortal." This logical expression can be described in the following way: A is B; All B are C; A is C. George Boole's complete acceptance of Aristotle's logic produced the mathematical logic for computing and has shaped the whole of the twentieth century and beyond.[76] Toulmin would have us tell a story and find evidence to support our claims to it.

The Toulmin model is a staple in first-year university composition courses and employs a six-part method of argumentation. It consists of three core components that can be further augmented by three additional elements. The core components function much like a legal argument. The core components of the Toulmin model move much like a legal argument. As Joan Karbach, whose essay has educated generations of undergraduates, explains: "(1) a person makes a claim, then (2) gives grounds to support that claim, and (3) backs the grounds with a warrant."[77] The claim serves as the central assertion or purpose of the argument. It relies on the grounds, which provide the specific evidence or context validating the claim. The warrant then establishes a link between the grounds and the claim, grounding the argument in what is generally accepted as true or false. This warrant ensures that the grounds are based on truth, thereby validating the claim.

Let's pretend a house is on fire (*claim*). There is smoke pouring out of the windows (*grounds*). In the phrase "where there is smoke, there is fire" would serve as the *warrant* to prove my *claim* using a basic truth. Now, if you are a sufficiently imaginative individual, it is possible to describe situations where smoke does not indicate fire. Of the three remaining components of the Toulmin model, the *backing* establishes the reliability and relevance of the *warrant*. A *qualifier* may be needed to modify the claim. For instance, a counterfactual re-rendering of the scenario can be described in any number of ways: if this house is in a movie studio, surrounded by a film crew, the smoke is likely produced with special effects, which would modify the claim to something like, "my house appears to be on fire." Of course, accidents happen, even on a movie set. A *rebuttal* would modify the *qualifier* and allow for a pretend fire, with the appearance of verisimilitude, to then become a real fire if the smoke machine should burst into flames! This is a thought experiment, so none of this is really true anyway. It could be true, in a fake kind of way. Through a counterfactual churn, the appearance of a real fire could be a simulated fire that becomes a real fire that is an imaginary fire.

The Toulmin model hinges on the speculative, imaginative abilities of the observer, who exercises judgment based on experience. By using observable facts (*grounds*) to understand the scope and context of an event (*claim*), in language or otherwise, the observer can assess the general truthfulness of the evidence (*warrant* and *backing*) to balance an understanding of specific conditions (*qualifier*) or exceptional circumstances (*rebuttal*). While this model is valuable in constructing complete arguments, it is a remarkable heuristic tool for assessing the *truthiness* of any statement or situation. One of the model's most remarkable features is its insistence that those making an argument consider the questions and concerns of their audience. As a heuristic, the model allows readers to grasp the basis of a claim and its relative durability and completeness. For an argument to be successful using the Toulmin model, a significant act of empathy and understanding of the audience is required. If any components of the model are missing, it serves as an immediate red flag for the audience, necessitating further inquiry. This is particularly crucial in low-information environments, which are rife with deceptive potential.

Toulmin's acceptance of relative truth value and probability does not wholly concede the foundation of understanding in logical formula and algorithmic proofs. In the words of Mark Weinstein in a recent study of Toulmin, "Mathematical logic construes truth as a univocal property of statements. This obscures the complexity with which truth functions in extra-mathematical contexts."[78] There are moments in which maths becomes mere tools for argument rather than accepted for their flat, numerical correctness. This notion of "metamathematics" reasserts that the rhetorical situation generalized jurisprudence in language as the fundamental space where truth or deception can be ascertained. Mathematical proofs can become mere fodder for the imaginative gymnastics found in misinformation. Climate deniers and election conspiracy theorists are well acquainted with the logical contortions required to twist data to fit alternative facts.

Lying to the Nazis

Our capacity to imagine possible future events and empathize with the emotional states of others enables us to lie, deceive, and cheat. While deception is fundamentally human, lying is often closely associated with moral decay and societal decline. Discussions about the role of lying and deception gained prominence during the early days of the internet in the 1990s, a period also marked by the 24-hour news cycle and the impeachment of President Bill Clinton. Philosophical debates on this subject are rare, given the discipline's slow-moving nature and focus on extensive arguments. Sissela Bok's *Lying: Moral Choice in Public and Private Life* has stood as a seminal work on the subject. Originally published in 1978 and reissued in 1989 and 1999, the book serves as a moral touchstone over several decades, arguing that lying is generally unacceptable and can only be justified in very specific circumstances to reduce harm. While this book stood as moral bulwark decrying the indecency of deception, Bok's optimism and good faith in humanity appear to be, shall we say, unreasonable by twenty-first-century standards. She concludes and adequately summaries *Lying* in the following words: "Trust and integrity are precious resources, easily squandered, hard to regain. They can thrive only on a foundation of respect for veracity."[79] For decades, Bok's book stood as the moral high-water mark, which argued that lying is bad and can only be justified in a very narrow set of circumstances to reduce harm. Lying to an enemy agent in a war is acceptable for instance, but outside of open hostilities and imminent physical harm, Bok always prescribes honesty as the best policy.

The rhetorical situation requires an understanding of hostile intent of those involved, which is often marked by deceit and lying. It is important to remember that the reissues of *Lying* came at a time when "most Americans took the President to have acted wrongly" and openly lied when President Clinton said he did not "have sex with that woman."[80] The prefaces of various editions of *Lying* stand as markers for Bok's own shifting view of popular culture, a view that echoes casual observations about the decline of decorum. Within a section entitled the "Rules of the Game," Bok warns against lying in almost all cases, and even when she does condone lying to reduce harm in cases of "self-defense," she recommends that an immediate truthful declaration must be made as soon as the threat of harm passes. She asserts, “No matter how hostile or dangerous a person, dealing with them honestly will always be preferable to deceit.”[81] Bok also cautions against a Machiavellian expansion of “the net of ‘enemyhood’,” advocating for honesty even when dealing with antagonistic forces.[82]

Bok’s perspective on threats and harm from potential “enemies” appears to be at odds with standard cybersecurity principles, which often involve threat modeling and risk assessments associated with any credible hostile threat actor. While Bok’s framework is geared more toward intimate, person-to-person interactions, she cautions against “paranoia” that pre-emptively identifies threats, labeling it as a potential pitfall.[83] Again, this ethical stance contrasts sharply with cybersecurity practices, which often require a certain level of pre-emptive caution—essentially, a form of prejudice—to identify potential bad actors and mitigate risks. Bok suggests that in openly declared conflicts, such as wars, deception is generally expected and often tacitly accepted, even if not explicitly consented to. In her words, if “the designation of a foe is open, as in a declaration of war,” says Bok, “deception is likely to be expected on all sides. While it can hardly be said to be *consented* to, it is at least known and often acquiesced in.”[84] This notion aligns with Thomas Rid’s argument that the problem of attribution becomes less significant as the severity of an attack increases. In such high-stakes situations, the ethics of deceit become a complex issue, ultimately requiring nuanced judgment.

While Bok champions the virtues of honesty and transparency, David Nyberg presents a more nuanced view, shall we say, arguing that deception is not only inevitable but also essential in human interactions. Truth-telling is not only morally overrated for Nyberg, but he also argues that life is more interesting with a little deception: “We all value the truth and yet we are *all* ordinary human deceivers; we neither want to know all the truth nor tell it all. Deception is not so much a plague as it is part of the atmosphere that sustains life.”[85] Lying is just part of the social lubricant that allows for humans to “escape confrontations” without constantly coming to blows over someone’s perception of honesty. “We humans are active, creative mammals who can represent what exists as if it did not, and what doesn’t exist as if it did” says Nyberg of the speculative, sometimes deceptive human condition: “And we do this easily and routinely.”[86] In contrast to Bok, who believes that truth-telling is the default ethical position and that lying requires justification, Nyberg suggests that both truth and deception are self-motivated acts. He challenges the notion that truth is inherently virtuous and unmotivated, arguing instead that we should focus on justifying lies rather than merely finding excuses for them.[87]

Lying is something deeply human. Telling a lie might help us feel better about ourselves or avoid hurting others. Lies might help us protect the innocent from the harsh full truths about our often-ugly world. It could even be said that there is a simple human honesty about the ways we lie to each other. Deception intends something else. It is a trap that is sprung. Luring requires planned deceit, a premeditated intent to trick. Nyberg offers a definition that is useful for cybersecurity: “Deception is the shrewd and sober art of ‘showing and hiding’ which is meant to control what is and is not perceived, assumed, or understood.”[88] In the augmented Toulmin model of communication, showing and hiding accounts for the counterfactual churn that is possible online. The pretext for a false claim shows an erroneous warrant by mimicking or counterfeiting factual sources and logic. Misdirection can emphasize convenient facts that support the false claim, while hiding legitimate qualifiers to a false pretext, as well as the associated rebuttals, returns the logical cycle to the original pretext. Hiding works to tamper with sound judgment and disguising or distracting from legitimate criticism.

With a reliable heuristic tool in the Toulmin model, it is possible to assess rhetorically the language of malicious messages online. In what remains, it is necessary to devise a rationale for deception that will help better understand the logic and motives of online threat actors. With an oversized emphasis on nation-state and institutional scale attacks, the mental model of a malicious message must be customized to account for digital citizens engaging in public spaces online. The art of *showing* and *hiding* helps augment the Toulmin model, by assessing the ways that a rational heuristic can be co-opted in a cycle of distraction and misdirection. By integrating Nyberg’s *logic of deception*, it is possible to reprocess this rhetorical argument through *showing* and *hiding.*

A security rhetoric interpretation of a house fire serves as a compelling metaphor for how attention can be manipulated, especially in high-stakes situations. The emotional urgency of a fire demands focused attention, often to the exclusion of other important details. Fire narrows an ability to notice, causing our attention to nearly all other features to *disappear.* This disappearing attention causes all other details to “become invisible either under cover or by blending in with the background.”[89] The *backing* used in this example, “where there is smoke, generally there is fire,” uses obvious truths to *disguise* events and consume rational resources, which causes the event to “become unrecognizable by adding to or modifying your characteristics.”[90] Finally, the fire and smoke *distracts* because fire is urgent and emotional, which allows a threat actor “to escape notice by creating uncertainty in the perceiver.”[91] This distraction creates a fertile ground for threat actors to operate unnoticed, capitalizing on the emotional chaos to sow uncertainty. Therefore, it’s crucial to question what might be hidden or obscured when our attention is seized. How do features of our environment *disappear?* How does an obvious truth *disguise* deeper relationships and vulnerabilities? How does urgency *distract* our emotional awareness of a situation? Understanding these dynamics is essential for a wide range of individuals who are often targeted online, including journalists, activists, whistle-blowers, human rights defenders, and academics.

What does the rhetoric of the house fire show? The adage “where there is smoke, there is fire” implies a high degree of certainty that might *mimic* a scientific, causally assured relationship. To mimic the credibility or authority of a socially respected position can add an assurance to *qualifier* statements. To anticipate and disrupt an anticipated *rebuttal*, a *counterfeit* worldview works to “invent a fake reality,” where it is possible for the fire to be set by a convenient scapegoat that further masks the situation. Conspiratorial thinking is amenable to *misdirection* efforts that seek “to emphasize an alternative to your real interest.”[92] This serves to misdirect attention by emphasizing an alternative issue, masking the real situation at hand. For instance, in the event of a house fire, an attacker could divert your attention away from the immediate crisis and toward a fabricated issue, such as questioning the competence of the local fire service.

The act of selectively showing and hiding information can manipulate both our beliefs and feelings. An attacker may aim to instill a “false belief,” perpetuate an existing misconception, or even introduce new information that disrupts our understanding of a situation.[93] An attacker may wish a target to acquire a false belief, continue believing a lie, stop believing something true (or false), or inject some new information that makes it impossible, even if it is obvious or desirable, to believe something. Feeling and believing are enmeshed. Belief comes easy when believing feels good. For this reason, Nyberg argues that sometimes we must go against our own principles to do what is morally right, as in the example of "lying to Nazis": "Feelings are vital to moral conduct, but they are not everything. Sometimes they must be subordinated. The same may be said of moral principles. Sometimes we must rise above principle to do the right thing."[94] This suggests that a rigid belief system can make us more susceptible to manipulation. The key is to exercise judgment, which should be informed by both rational thought and emotional awareness. In doing so, we can navigate complex moral landscapes, rejecting the excuse of "just following orders" and making ethical decisions that protect the vulnerable.

Security rhetoric focuses on the art of discerning what is shown and what is hidden to identify deception. These lures, often seen in phishing emails, aim to deceive users into self-harm by assisting an attacker. However, the concept of luring extends beyond emails to misinformation and disinformation across various platforms. These deceptive tactics mimic, misdirect, and counterfeit legitimate public discourse, serving to lure users into adopting or rejecting politically charged beliefs. By understanding the grounds and warrants that support these deceptive claims, as well as the tactics used to disguise their true intent, we can better protect ourselves and our communities.

Security rhetoric aims to dissect those fleeting moments of indecision—those potential lapses in judgment—that are often exploited by malicious messages, whether in emails or on other online platforms. These moments are not just minor inconveniences; they pose a significant risk to individuals and organizations alike. The economic repercussions of cyberattacks, such as ransomware, are now so severe that they are turning into election issues in various countries. In this context, digital literacy and user education are not just individual responsibilities but national-level strategic priorities. They serve as a crucial line of defense against both cybercriminals and nation-state hackers, safeguarding the integrity of critical infrastructure and protecting the security and privacy of citizens.

The shift toward applied humanistic research demands a new level of public engagement, political activism, and cultural relevance that translates academic insights into actionable outcomes. As researchers increasingly focus on public-facing digital environments, integrating security planning into everyday research practices becomes non-negotiable. Given that much of our cultural heritage and public discourse now reside in digital spaces, understanding the broad cultural implications of security practices is crucial. This includes taking steps to mitigate the risks posed by disinformation and misinformation. However, this focus on security will inevitably clash with the humanistic ethos of openness and transparency, creating a complex landscape that researchers must navigate carefully. By actively engaging with these challenges, applied humanistic research can play a pivotal role in shaping a national culture of security. This, in turn, can lead to more informed discussions about the ethical considerations and attitudes necessary for achieving comprehensive digital literacy.

This security culture hardens against manipulation and starts with the language we use. The humanities must go beyond the war metaphors of defenses and attackers. We need to re-imagine the motivations of attackers within an intersectional matrix of those in power and those without power. Cybersecurity is an asymmetrical relationship that grants hostile actors with limited resources the ability to attack global superpowers and influential citizens internationally. It can be daunting to consider global risks. Emotion and feeling are part of this work. As this chapter has already argued, cybersecurity work is "emotional work."[95] In her 2009 article, "Toiling in the Field of Emotion," Harriet Fraad defines emotional work in the following way:

> Emotional labor is the expenditure of time, effort and energy utilizing brain and muscle to understand and fulfil emotional needs. By emotional needs, I mean the human needs for feeling wanted, appreciated, loved and cared for. Individuals' emotional needs are often unspoken or unknown/unconscious. Emotional labor often occurs together with physical labor (producing physical goods or services), but emotional labor differs from physical labor by aiming to produce the specific feelings of being wanted, appreciated, loved and/or cared for. Of course like all powerful forces, emotional labor may be used to undermine others or frustrate their emotional needs as well as help them.[96]

While Fraad does not describe the security aspect of emotional labor in her article, the following chapter will take up the status of emotional work as a praxis of care in digital space. Digital citizens must be possessed of enough emotional maturity to resist being manipulated by triggering, offensive, insensitive content. Triggering is the short fuse that attackers seek to manipulate. A mock emergency masked by manners to distract from the deception, tricking us to click.

By moving away from war metaphors, we can focus on building a culture of care and generosity predicated on shared responsibility to one another, rather than one of constant conflict and defense. This could involve creating educational programs that not only teach the technical aspects of cybersecurity but also address the emotional and psychological factors that make people vulnerable to manipulation in the first place. It could mean designing systems that are not just secure but also user-friendly, reducing the emotional burden on end users. It could also involve public awareness campaigns that aim to reduce the stigma associated with falling for a cyber scam, encouraging a more open dialogue about these common vulnerabilities.

Notes

[1] Nancy Stern, and Robert A. Stern, *Computers in Society* (Prentice-Hall: Englewood Cliffs, NJ, 1983), xxi.

[2] Ibid., 4–5.

[3] Ibid., 340.

[4] Ajit Maan, "Disinformation is not the problem and information is not the solution," *Homeland Security Today*, February 6, 2021, https://www.hstoday.us/subject-matter-areas/counterterrorism/disinformation-is-not-the-problem-and-information-is-not-the-solution/. I have silently corrected typos in this quote from the original. I flag these errors as evidence that this article was not written by an AI, and it appears as though there is an increasing need for depth in reflections on, as the section is titled, "narrative and national security."

[5] Ibid.

[6] Kenneth Burke, *A Rhetoric of Motives* (Berkeley, CA: University of California Press, 1969), 25.

[7] An understanding of rhetorical concepts enables deceit for Burke, by bringing "rhetoric to the edge of cunning." It takes a thoughtful

identification of rhetoric to temper the manipulative capacity of rhetoric. See Burke, *A Rhetoric of Motives*, 36.

[8] Howard Altman, “Russia Preparing to Attack Ukraine by Late January: Ukraine Defence Intelligence Agency Chief,” *Military Times*, November 20, 2021, https://www.militarytimes.com/flashpoints/2021/11/20/russia-preparing-to-attack-ukraine-by-late-january-ukraine-defense-intelligence-agency-chief/.

[9] Ciaran Martin, “Cyber Realism in a Time of War,” *Lawfare Blog*, March 2, 2022, https://www.lawfaremedia.org/article/cyber-realism-time-war; See also Thomas Rid, “Cyber War Will Not Take Place,” *The Journal of Strategic Studies* 35 no. 1 (2012): 5–32.

[10] David Cattler, and Daniel Black, “The Myth of the Missing Cyberwar,” *Foreign Affairs*, April 6, 2022, https://www.foreignaffairs.com/articles/ukraine/2022-04-06/myth-missing-cyberwar.

[11] Ibid., xiv.

[12] Ron Deibert, and Rafal Rohozinksi, “Tracking GhostNet: Investigating a Cyber Espionage Network,” *CitizenLab*, March 29, 2009, https://citizenlab.ca/wp-content/uploads/2017/05/ghostnet.pdf.

[13] Ibid., 12. Emphasis mine.

[14] Randy Burkett, “Rethinking an Old Approach: An Alternative Framework for Agent Recruitment: From MICE to RASCLS,” *Studies in Intelligence* 57 no. 1 (2013): 7–17. Available: https://www.cia.gov/resources/csi/studies-in-intelligence/volume-57-no-1/an-alternative-framework-for-agent-recruitment-from-mice-to-rascls/.

[15] Thomas Rid, *Cyber War Will Not Take Place* (London: Hurst and Company, 2013), 145.

[16] Ibid., 48.

[17] Ibid., xvi.

[18] Scott Graber, “Defend Forward: Adapting Offense and Defense Strategy to Cyberspace,” *Yale Cyber Forum*, July 20, 2021, https://www.cyber.forum.yale.edu/blog/2021/7/20/defend-forward-adapting-offense-and-defense-strategy-to-cyberspace; see also https://media.defense.gov/2018/Sep/18/2002041658/-1/-1/1/CYBER_STRATEGY_SUMMARY_FINAL.PDF.

[19] Ibid., 162.

[20] See Alex Kaufman, “Prepare to Be Shocked! What Happens When You Actually Click on One of Those ‘One Weird Trick’ ads,” *Slate*, July 30, 2013, https://slate.com/business/2013/07/how-one-weird-trick-conquered-the-internet-what-happens-when-you-click-on-those-omnipresent-ads.html.

[21] Rid, *Cyber War*, 153.

[22] Lawrence Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999) argues that cyberspace and Silicon Valley corporations do not require direct regulation because code is a self-regulating system, too complex for government oversight. With the failure of large social media platforms to moderate misinformation, despite the rise of LLMs, direct regulation of Silicon Valley is no longer regarded as hindering innovation. See also Olga V. Mack, “‘Code Is Law’: Should Software Developers Protect Our Freedoms?” *Above the Law*, August 12, 2019, https://abovethelaw.com/2019/08/code-is-law-should-software-developers-protect-our-freedoms/.

[23] Lawrence Lessig, “Code Is Law: On Liberty in Cyberspace,” *Harvard Magazine*, January 2000, https://www.harvardmagazine.com/2000/01/code-is-law-html.

[24] “Code Is Law and the Quest for Justice,” *Ethereum Classic Blog*, September 9, 2016, https://ethereumclassic.org/blog/2016-09-09-code-is-law.

[25] Rid, *Cyber War*, 160.

[26] See the NIST Computer Security Resource Center glossary for more definitions: https://csrc.nist.gov/glossary/term/vulnerability.

[27] Kimberlé W. Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," *Faculty Scholarship: Columbia Law School* 139 (1989): 159–60.

[28] Catherine Knight Steele, *Digital Black Feminism* (New York: New York University Press, 2021), 53.

[29] Ibid., 53.

[30] Jacey Fortin, "What Is Critical Race Theory? A Brief History Explained," *New York Times*, June 26, 2023, https://www.nytimes.com/article/what-is-critical-race-theory.html.

[31] Ibid.

[32] Fabiola Cineas, "Conservatives' War on Emotions in the Classroom," *Vox*, February 22, 2023, https://www.vox.com/the-highlight/23584837/social-emotional-learning-conservative-culture-war-in-schools.

[33] See https://defendinged.org/.

[34] See https://defendinged.org/map/ for a map of Conservatives fighting contemporary education policy.

[35] See the site, *Conservative Internet*, as a measure of influence of conservative social medias like Gab, Godtube, Censored.tv, and Bitchute, among others: https://cinternet.org/2021/july/conservative-youtubers. See the video from Innuendo Studios as a breakdown of how conservative social media "red pills" or "radicalizes a normie" otherwise rational individuals: https://www.youtube.com/watch?v=P55t6eryY3g.

[36] Lisa Feldman Barrett, *How Emotions Are Made: The Secret Life of the Brain* (New York: Mariner, 2018).

[37] Lisa Feldman Barrett, “When Is Speech Violence?” *Gray Matter: New York Times*, July 14, 2017, https://www.nytimes.com/2017/07/14/opinion/sunday/when-is-speech-violence.html.

[38] Trump’s lies were part of a larger cultural trend of dishonesty: https://www.psypost.org/2022/02/study-suggests-trumps-false-tweets-were-mostly-intentional-lies-not-accidents-62627.

[39] Jonathan Haidt, and Greg Lukianoff, “Controversial Speeches on Campus Are Not Violence,” *The Atlantic*, July 2017, https://www.theatlantic.com/education/archive/2017/07/why-its-a-bad-idea-to-tell-students-words-are-violence/533970/.

[40] Barrett, *How Emotions Are Made*, 251.

[41] Joan Donovan, Emily Dreyfuss, and Brian Friedberg, *Meme Wars* (New York: Bloomsbury, 2022), 20–1.

[42] Donovan et al., *Meme Wars*, 303–4.

[43] Ibid., 317.

[44] See https://www.gao.gov/blog/rising-threat-domestic-terrorism-u.s.-and-federal-efforts-combat-it. See also the US Government Accountability Office report “Domestic Terrorism: Further Actions Needed to Strengthen FBI and DHS Collaboration to Counter Threats,” *gao.gov*, February 22, 2023, https://www.gao.gov/products/gao-23-104720.

[45] Gabor Maté, *The Myth of Normal: Trauma, Illness and Healing in a Toxic Culture* (New York: Alfred A. Knopf, 2022), 345.

[46] Ibid., 314.

[47] For a simple breakdown of Altemeyer’s findings, on the occasion of the election of President Trump in 2016, see “A Word from Dr. Bob Altemeyer on Donald Trump and Authoritarian Followers,” *Daily Kos*, March 2, 2016, https://www.dailykos.com/stories/2016/3/2/1494504/-A-word-from-Dr-Bob-Altemeyer-on-Donald-Trump-and-Authoritarian-Followers. Altemeyer has continued to track Trump

supporters after his retirement from the University of Manitoba, on his site https://theauthoritarians.org/. His final comprehensive book can be found, for free, here: https://theauthoritarians.org/options-for-getting-the-book/.

48 Bob Altemeyer, *The Authoritarian Specter* (Cambridge, MA: Harvard University Press, 1996), 101.

49 Brian Massumi, *Parables for the Virtual: Movement, Affect, Sensation* (Durham, NC: Duke University Press, 2002); William E. Connolly, *Neuropolitics: Thinking, Culture, Speed* (Minneapolis: University of Minnesota Press, 2002).

50 Ibid., 29.

51 Ruth Leys, "The Turn to Affect: A Critique," *Critical Inquiry* 37 no. 3 (2011): 437.

52 Massumi, *Parables for the Virtual*, 38.

53 Marissa Conway, and Nehmat Kaur, "The Intersectionality and Cybersecurity Toolkit," *The Centre for Feminist Foreign Policy,* March 2022, https://centreforfeministforeignpolicy.org/feminist-peace-and-security/, 6. See also the following link to access the Intersectionality and Cybersecurity Resource Dashboard: https://start.me/p/RM18la/intersectionality-and-cybersecurity-resource-dashb.

54 Conway, and Kaur, *"The Intersectionality and Cybersecurity Toolkit,"* 10.

55 Craig Higson Smith, Daniel Ó Cluanaigh, Ali G. Ravi, and Peter Steudtner, *The Holistic Security Manual: A Strategy Manual for Human Rights Defenders* (Berlin: Tactical Technology Collective, 2016). See the following link to access the UK Government's National Cyber Strategy, 2022: https://www.gov.uk/government/publications/national-cyber-strategy-2022.

56 See "The Sherman Kent Center for Intelligence Analysis: Occasional Papers" available on Archive.org:

https://web.archive.org/web/20070612215517/https://www.cia.gov/library/kent-center-occasional-papers/vol1no5.htm.

[57] See Sherman Kent's "Words of Estimative Probability" on the CIA historical review program site: https://www.cia.gov/static/0aae8f84700a256abf63f7aad73b0a7d/Words-of-Estimative-Probability.pdf.

[58] Ibid.

[59] Ibid.

[60] Ibid.

[61] There are other attempts to define probability in language, such as the UK's Defense Intelligence "probability yardstick," which remains indebted to Kent. See https://www.gov.uk/government/news/defence-intelligence-communicating-probability.

[62] Words of estimative probability admit the aesthetic quality of analysis that is predicated on human cognition and perception. By systematizing the rhetoric of cyberattacks, I want to propose something like a Bell test for threat detection in malicious messages.

[63] Ibid.

[64] Aristotle, *On Rhetoric: A Theory of Civic Discourse*, 2nd ed. Trans. George A. Kennedy (New York: Oxford University Press, 2007); Burke, *A Rhetoric of Motives*; Stephen Toulmin, *Uses of Argument, updated edition* (New York: Cambridge University Press, 2003).

[65] Aristotle, I.1.2.

[66] Burke, *A Rhetoric of Motives*, 64.

[67] Stephen E. Toulmin, *The Uses of Argument* (Cambridge: Cambridge University Press, 2003), 7.

[68] Ibid., 15.

[69] Ibid., 57.

[70] Ibid., 46.

[71] Ibid., 50.

[72] In this way, I work to augment and extend Rita Felski's 2008, *Uses of Literature*, which offers a "tentative taxonomy" (132) to describe the usefulness of literature in the twenty-first century. While I see no convincing argument that literature is somehow special or different from other forms of narrative or cultural expression, I share the "problem of justification" (4) for a humanistic training and outlook, particularly as it pertains to specialized training and higher-education. Humanistic critique, honed on the complex linguistic and historical contexts of literature, has a role in training cybersecurity analysts and citizens because of the speculative, often intersectional, descriptions of human experience. This is an argument that I extend in my previous book, *Hacking in the Humanities* (2022). The problem the humanities have been struggling with is the prospect that human cultural artifacts might now be the ineffable and beautiful windows into the soul of humanity. Literary study in particular has done its job of dethroning religious texts as a source of direct inspiration for human affairs. All human creative output now can serve to gain peeks into the motivations of human misery and logic. Now data drive our world and our sense of self. There is nothing outside the text, said Jacques Derrida someplace or other. At the dawn of AI, we, humanity, learn that our written language might not be that important after all. All of human creative and intellectual output has amounted to a training set for some new kind of intelligence. It is for this reason that the humanities, maybe humankind altogether, must practice an applied form of humanities. An applied humanities would seek to identify, define, and actually solve problems. The humanities needs to do more an admire problems. We need to contribute to solving them.

[73] Ibid., 85.

[74] Ibid., 64, 56.

[75] Zachary Jonathan Jacobson, "Many Are Worried about the Return of the 'Big Lie.' They're Worried about the Wrong Thing," *The Washington Post*, May 21, 2018, https://www.washingtonpost.com/news/made-by-

history/wp/2018/05/21/many-are-worried-about-the-return-of-the-big-lie-theyre-worried-about-the-wrong-thing/.

76 John Corcoran, “Aristotle’s ‘Prior Analytics’ and Boole’s ‘Laws of Thought,’” *History and Philosophy of Logic* 24 (2003): 261–88.

77 Joan Karbach, “Using Toulmin’s Model of Argumentation,” *Journal of Teaching Writing* 6 no. 1 (1987): 81.

78 Mark Weinstein, “A Metamathematical Extension of the Toulmin Agenda,” in *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation.* David Hitchcock and Bart Verheij (eds.) (New York: Springer, 2006): 51.

79 Sissela Bok, *Lying: Moral Choice in Public and Private Life* (New York: Vintage, 1978), 249. See also *Secrets: On the Ethics of Concealment and Revelation* (New York: Vintage, 1983).

80 Ibid., xvi.

81 Ibid., 141, 144.

82 Ibid., 136.

83 Ibid., 139.

84 Ibid., 144.

85 David Nyberg’s *The Varnished Truth: Truth Telling and Deceiving in Ordinary Life* (Chicago: University of Chicago Press, 1993), 25.

86 Ibid., 12.

87 Ibid., 21.

88 Ibid., 67.

89 Ibid., 67.

90 Ibid., 67.

91 Ibid., 69.

92 Ibid., 73.

[93] Ibid., 74.

[94] Ibid., 57.

[95] Harriet Fraad, “Toiling in the Field of Emotion,” *The Journal of Psychohistory* 35 no. 3 (2008): 270–86. The term “emotional labor” was coined by sociologist Arlie Russell Hochschild in *The Managed Heart: Commercialization of Human Feeling* (1983). Despite the popular conversations about “emotional labor” occurring on social media and blogs, the scholarly uptake of Fraad’s work has been limited to Marxist-feminist analysis of the parenting, domestic labor, and mental health. The broad applicability of emotional labor should not diminish or further mask this largely unacknowledged, gendered labor. By moving away from the war metaphors within discourses around cybersecurity, the personal and the private are centered as the space of mental, psychic vulnerability. It also positions cybersecurity professionals as caregivers, teachers, and emergency responders. Care and generosity become the animating forces of IT rather than combat and battle.

[96] Ibid., 270.

3

Automatic Anxiety

Approximately 14.5 billion spam emails are sent every day, with an average cost to a mid-sized company of $1.6 million in managing fraud.[1] These seemingly innocuous emails account for the rise in ransomware attacks that have reshaped cybersecurity, foreign policy, and individual online behavior.[2] The phishing email is the most prevalent and effective attack vector, including the scale of the messages sent, the range of attacks that begin with a successful phish, as well as the effectiveness in attaining initial access.

We've all seen it: a seemingly legitimate email from the government, maybe tax officials or other law-enforcement, tries to intimidate a user into following instructions. Click a link. Download a file. Enable macros.[3] Some messages may adopt a more personal tone. A message may appear to be from a friend-of-a-friend or colleague asking for a simple favor or alerting you to a billing issue with your local utility. They just need a bit more information. What if your account has been compromised? Enter your password to check. Or, you just need to confirm that you've won a prize in a contest! Make a quick account and the prize is yours!

These lures lead to nothing good. A lure entraps or ensnares. In the natural world, hunters lure animals to bait, trap, or hook them. A lure catches the unsuspecting and the unaware. Similarly, in the digital realm, lures catch the unsuspecting and the inattentive. Whether it's street-market hucksters or online marketers, the goal is to pique curiosity and entice action. Gaining initial access relies on both technical and rhetorical strategies. Luring is about concealing intentions and seeking privileged access to our digital lives. Software vendors sell unexciting software for Privileged Access Management (PAM) tools because so much depends on a password.[4]

Despite our current dependence on passwords, the era of the password is rapidly ending. Tech giants like Apple, Microsoft, Amazon, and Google are collaboratively working towards a "passwordless" future. The move toward passwordless authentication started sometime in 2013 and began getting noticed in late 2018 and early 2019, when tech reporters started making headlines like "When can we finally get rid of passwords?" or "Why Passwords Might (Finally) Go Away."[5] Passwords are a huge part of our daily experience, especially if you can't remember them.[6] Martin Paul Eve says it more poignantly: "Passwords are crucial to our lives. They regulate our finances, protect our communications and prove who we are to others. They are powerful words."[7]

Powerful as they are, passwords are going the way of the dodo. The FIDO (Fast IDentity Online) Alliance has been advocating for "simpler, stronger authentication" with "phishing-resistant, hardware-bound authentication mechanisms" since its inception in 2013.[8] This corporate-driven shift toward a passwordless future makes the transition all but inevitable.[9] Often touted as "the industry's answer to the password problem," public key cryptography is poised to replace easily memorable passwords.[10] Gone are the days of creating passphrases like "correct horse battery staple."[11] Even when bolstered by Multi-Factor Authentication (MFA) tokens, passwords remain vulnerable to phishing attacks.[12] While there are phishing-resistant MFA standards, they serve as a stopgap rather than a long-term solution.[13] It's safe to say that MFA is not a popular solution anyway.[14]

To attain the highest rating in the Digital Identity Guidelines set forth by NIST (National Institute of Standards and Technology), the FIDO Alliance is leveraging the widespread availability of smartphones and computers equipped with bio-authentication hardware, such as facial and fingerprint recognition.[15] In 2015, the World Wide Web Consortium (W3C) introduced the open WebAuthn specification to standardize web-integrated authentication across all major browsers.[16] Utilizing this near-ubiquitous hardware and open standards, FIDO has succeeded in building consensus around a low-cost, hardware-based, instantly deployable, phishing-resistant authentication mechanism. This mechanism meets the stringent requirements of NIST's ALL3-level smart cards.[17] If widely adopted, this would represent a significant achievement given the complexities involved in implementing any internet-wide standard—especially one that replaces something as personal and private as passwords.

The potency of these "powerful words" will transition to machine-readable, bio-authenticated, and hardware-bound forms. Instead of relying on memorized information for identity authentication, systems will recognize unique biometrics such as fingerprints, facial scans, or even more specialized methods like retinal or palm vein scans. Employment, education, banking, digital wallets, and extensive digital libraries of personal information will all be accessible via FIDO-enabled devices in this brave near future. While FIDO credentials are designed to "survive device loss," this feature introduces a significant vulnerability: attackers can exploit the authentication recovery process through social engineering.[18] Brian Krebs has interviewed multiple security researchers who have raised concerns about this account recovery mechanism, including Nicholas Weaver, who characterized it as "a really hard problem to do security and already one of the biggest weaknesses in our current system."[19]

The limitations of hardware-bound authentication, backed by corporate-owned, industry-led cloud services, quickly circle back to the social dimensions of security. This centralizes the power to authenticate or de-authenticate individuals within these corporations. Implementing FIDO would effectively end password sharing for subscription-based services such as content-driven streaming platforms. Patterned on the individual for individual use, these authentication services further entrench the concept of individualism in the security landscape. Participation in this new security environment would require individuals to sacrifice a biometric feature, such as a facial scan or fingerprint. Facial data is poised to become the most valuable form of authentication, raising the likelihood of public face-scanning becoming more prevalent. Offloading the responsibility of security onto individuals in the pursuit of mass convenience poses significant risks.[20]

What if a passwordless future allows authorities to withhold access to finances, medical services, or even your own home with the flick of a switch? "Smart locked" apartments could allow landlords to instantly evict tenants if they miss a rent payment.[21] Internet-connected vehicles could lock out owners and autonomously return to the dealerships if payments are overdue.[22] While digital identity currently emphasizes real-name policies and consistent online identities, governmental participation in this system, as a future arbiter of trust by validating citizens still remains uncertain. What is clear, however, is that biometrics will increasingly serve as the cornerstone of identity authentication. Given that we have a finite number of biometric features suitable for authentication, these markers will accompany us for our entire lives and stored perpetually on corporate servers thereafter.

It's easy to imagine how insurance companies will grant preferred rates in exchange for comprehensive access to your lifestyle choices. Every drink of alcohol consumed, and every cigarette smoked, could be documented through facial scans. Insurance providers would have a detailed record of your habits and impulses, all linked to your authentication history. Governments could offer tax incentives for healthy choices, a benefit more easily accessible to those with wealth and privilege.

Cyberinsurance is already a product for businesses, often categorized under the umbrella term "Internet Security Liability."[23] Like all insurance markets, premiums rise and coverage narrows as risks escalate. The global landscape of cyber warfare and rampant cybercrime has led to policy exclusions for offensive cyber operations and unpatched ransomware vulnerabilities.[24] Cyberinsurance may become increasingly mainstream as a means of securing peace of mind online. In his book, *Threat Modeling*, Adam Shostack discusses how adversaries target a "normal person" and trick them into harming themselves.[25] The doldrums of daily communication and the scale of digital messaging disguise the danger of these lures. The anxiety associated with this ubiquitous, low-level risk is taxing for normal people and cybersecurity experts alike. Security rhetoric aims to bridge the gap between everyday users and security experts, offering an in-depth look at the current state of online security risks and dissecting the linguistic and stylistic elements of malicious messages.

Anyone with an email account will benefit from the current sophistication of commercial spam-filtering software.[26] These systems employ a multi-faceted approach, including blocklists of known malicious senders, validation of email header data, and content analysis. Email content may avoid shorted URLs or look for marketing language or explicit language that could entice a user to click a malicious link. Machine learning can detect a range of features, trained on huge collections of known phishing emails. Early research framed the struggle between phishers and security teams as doing "battle in user interface space," where being redirected to a fraudulent website represented the sum of all harms. Factors such as a "Lack of computer system knowledge" and limited attention were thought to exacerbate the problem, alongside the "visual deception tricks" that made one website easily mimic another.[27] Early browser plugins aimed to counter "semantic attacks" that "exploit human vulnerabilities."[28] Even at this nascent stage, researchers recognized the inherent asymmetry of the problem: attackers can effortlessly generate and disseminate phishing emails, while users need only slip up once to be compromised.[29] Successful phishing templates can be refined by evaluating relative improvements, campaign by campaign.

Gamification was offered as a solution to training users to identify and avoid phishing attacks.[30] As machine learning gained prominence and increased applicability, it enabled the development of classifiers capable of improving the predictive performance of filters by training on real data.[31] Data-hungry approaches sought to classify the malicious web pages as well as the email contents to determine if the destination was legitimate in advance.[32] Conversely, engineers who conducted "phishing studies" on human subjects found that social science research on user behavior is highly variable and prone to error.[33] The success rate of correctly identifying a phishing email was found to be context-specific, largely influenced by how users categorize emails and assign trust, often referred to as the "expectancy effect."[34] To address this, tool development has focused on incorporating user interface add-ons that provide timely warnings, thereby supplementing users' decision-making abilities.[35]

In the interceding years, phishing attacks only escalated.[36] Identity theft remained the main concern in the years before ransomware attacks emerged into the public consciousness with WannaCry in 2017.[37] There was a growing awareness of the economic toll of phishing attacks and the risk faced in "e-commerce" as services moved online in a significant way.[38] The field of phishing research has expanded within computer science to encompass heuristic logic, deny-and-allow listing, and machine learning techniques. As phishing became part of a broader umbrella of *social engineering* attacks, disciplinary breadth of academics and commercial researchers included "social psychology, economics, distributed systems, machine learning, human-computer interaction, and public policy."[39]

Psychology has been the disciplinary foundation for understanding phishing and online deception, particularly since Christopher Hadnagy popularized *social engineering*.[40] However, it's becoming increasingly evident that the psychological impacts of these activities are evolving in ways that traditional psychology may struggle to anticipate or adapt to quickly. Social engineering's historical focus on deception may actually limit its effectiveness in comprehending the nuances of online manipulation and influence. This challenge is highlighted in a recent comprehensive study titled *Deception in the Digital Age*, where the limitations of psychology in this context become glaringly apparent:

> Predicting future technologies that could provide platforms for misinformation and misdirection is a difficult task. What is more certain is that whatever technologies emerge, there will likely be opportunities to alter information or perceptions along the way. This suggests that it would be a good idea to initiate research that can provide us with some ideas and theoretical knowledge about the effects of such deceptions, along with developing methods to detect possible deception as well as providing some careful thought into the possibility of deception countermeasures.[41]

In answer to this call for deception countermeasures, this chapter will first summarize the rhetorical trends and habits of phishing campaigns from the cybersecurity industry and government reports.[42] Social engineering seeks to blend neuropsychology with social science psychology principles to formulate a theory to describe influence and manipulation.[43] A thorough analysis of the origins and cultural underpinnings of these social engineering principles is essential to ensure their global applicability in the detection and attribution of cyberattacks.

Equal Opportunity Victimizer

Many definitions of social engineering exist, but IBM's definition encapsulates the human vulnerability at the core of technical systems: "Social engineering attacks manipulate people into sharing information they shouldn't share, downloading software they shouldn't download, visiting websites they shouldn't visit, sending money to criminals, or making other mistakes that compromise their personal or organizational security."[44] There is something transgressive about social engineering attacks; a social engineer tries curry trust and rapport with a target to manipulate or influence them to perform a compromising action. Since authentication relies on something unique that users possess (a key), know (a password), or inherently are (a fingerprint), social engineering often targets vulnerable users to divulge that unique thing. However, a vulnerable user may also take compromising actions, such as inappropriately approving a MFA request, downloading unauthorized software, or clicking on a malicious link. For example, so-called MFA-fatigue attacks bombard users with authentication requests to irritate them into approving a login, possibly mistaking it for a system glitch.

Hadnagy has been a pivotal figure in shaping the concept of social engineering for cybersecurity professionals. He even secured a foreword from Steve "Woz" Wozniak, the beloved co-founder of Apple Computers, for his book *Social Engineering: The Science of Human Hacking* (2018).[45] Hadnagy is the author of four other books, including *Social Engineering: The Art of Human Hacking* (2010), *Unmasking the Social Engineer* (2014), *Phishing Dark Waters*(2016), and *Human Hacking* (2021). In 2016, Hadnagy and Michele Fincher defined *phishing* as "the practice of sending e-mails that appear to be from reputable sources with the goal of influencing or gaining personal information."[46] Today, phishing has become a term familiar to the general public, thanks in part to government public service announcements about identity theft and extortion. *Phishing Dark Waters* was published before the term was generally known, but phishing was coined in an earlier period. According to *Ollmann, Gunter* in 2011, the term originates in the hacker magazine *2600*:

> The word "phishing" originally comes from the analogy that early Internet criminals used email lures to "phish" for passwords and financial data from a sea of Internet users. The use of "ph" in the terminology is partly lost in the annals of time, but most likely linked to popular hacker naming conventions such as "Phreaks" which traces back to early hackers who were involved in "phreaking"—the hacking of telephone systems.[47]

Any effective rhetoric requires a lexicon, and in this context, phishing stands as the single most important term.[48] Phishing has long passed into the realm of common knowledge. The potency and impact of sending malicious messages are such that any major global event quickly triggers a thematic shift in spam messages. These messages exploit topics that are top of mind for the general public—events like January 6, the conflict in Ukraine, or the Covid-19 pandemic have all been leveraged by scammers.

Cybersecurity professionals using social engineering methods are likely well acquainted with Hadnagy's contributions. Despite the number and prominence of his books, it is critical for security professionals and social scientists working in online influence and manipulation to build a new foundation for this important aspect of security practice. Hadnagy's conceptualization of social engineering is rooted in ethically questionable practices, pop-psychology, and selectively cited neuroscience.[49] He diminishes the role of emotions and feelings, dismissing them as irrational elements to be mastered in oneself and exploited in others. I'll have more to say about that later, but the critique of Hadnagy emerges out of a series of events that cast his methodology and work into stark relief against his own political beliefs.

At the 2021 Idaho Falls BSides Cybersecurity Conference, Hadnagy decided to organize his talk around "Cancel Culture Facts." You can see where this is going. To illustrate the ease with which "cancel culture" can be triggered, he drew an unusual analogy involving a garden hose, describing it as "a very long black thing."[50] It was oddly specific and deeply weird. He said that being accidentally scared of a hose means that "cancel culture" would feel the need to punish garden hose companies because snakes are scary.[51] The analogy is perplexing and only makes sense if one assumes that "cancel culture" is not motivated by a moral imperative to address injustice and inequality. The moment serves to reveal Hadnagy's perspective that those who protest systemic bias are themselves misguided and somewhat unintelligent.

He took time during his talk to describe an "un-balanced" view in "cancel culture" as well as an "intolerance" to differing points of view.[52] He took time to coin an acronym to define his vision of "cancel culture": DEAD, stands for "Deindividuation" as a kind of group think; "Emotions" because he assumes heightened emotional states in groups lead to illogical thought; "Acceptability" of acts changes in groups; and finally "Diffusion" of responsibility through group actions. In his concluding remarks, Hadnagy suggests some "takeaways" such as "don't let the poor emotional control of others make you show lack of control" and "work on empathy and self-awareness."[53] His lack of self-awareness in advocating self-awareness in this moment is astonishing. There is a breath-taking kind of self-chastising irony dripping in these words.

There are consequences for social engineering that are more important than having another discussion about "cancel culture." Hadnagy's influence on the cybersecurity field's understanding of social engineering places undue emphasis on manipulating emotional states, including "the use of authority and gender differences in compliance."[54] He advocates for the exploitation of social and gender inequalities as manipulative tools. While citing his own work, *Unmasking the Social Engineer* (2014), he describes the psychological principles of decision-making as "amygdala hijacking," wherein the fight, flight, and freeze survival instincts are used as a method to control "an automated brain process" that "human hacking" exploits like a computer.[55]

Hadnagy advises those who find themselves targeted by "cancel culture," as he sees it, "look for opportunities to be humble."[56] A few months after making these comments, Hadnagy was banned from DEF CON in 2022 for violating the event's Code of Conduct.[57] Being banned from a hacker convention like DEF CON is no small feat, given that attendees routinely engage in quasi-legal hacking activities during the event, targeting everything from WiFi networks to charging ports and vending machines: maybe avoid that OS update while you're at the conference. However, the conference's Code of Conduct is straightforward and primarily aims to prohibit harassment: "Harassment includes deliberate intimidation and targeting individuals in a manner that makes them feel uncomfortable, unwelcome, or afraid."[58] The conference issued a succinct yet unambiguous statement regarding Hadnagy's behavior at DEF CON:

1. We received multiple CoC violation reports about a DEF CON Village leader, Chris Hadnagy of the SE (Social Engineering) Village. After conversations with the reporting parties and Chris, we are confident the severity of the transgressions merits a ban from DEF CON.
2. We have also taken the rare action to disband the DEF CON Group DCG414. Code of Conduct violations by the group's primary Point

> of Contact and subsequent mishandling of the event left us without confidence in the group's leadership.[59]

There is a real trend toward making cybersecurity culture welcoming to newcomers.[60] Hadnagy claims on his personal blog that he has "not heard the exact allegations"; he claims that he has "no desire for the accuser's identity to be made public," and he does not "know what the accusations are," while claiming that "DEF CON has informed me that the allegations are NOT related to sexual misconduct."[61]

Alyssa Miller summarized the sentiments of this period by saying, "it's sad to see people that were trusted as leaders in our security community continue to be involved as bad actors in these situations." She went on to explain that "These are people that others count on, who have position that gives them a certain power dynamic in the industry."[62] Hadnagy opted to take legal action against DEF CON over the ban, but his lawsuit was dismissed due to a jurisdictional technicality.[63] So there it is: he got upset about "cancel culture" and then "got cancelled."

None of this would be surprising to readers of Hadnagy's books. His work is largely grounded on pop-psychology and hacker edginess.[64] In *Social Engineering* (2018), he uses a brutal image of an emaciated and abused dog in a cage with a figure description that says, "How does this make you feel?"[65] He defines social engineering as "not politically correct" and suggests that this "truth can be hard for many people to swallow, but it's real."[66] He claims to acknowledge that some of his ideas are distasteful, but he's just being honest. Don't shoot the messenger, right? But he is a teacher and industry leader, with some distasteful ideas.

Hadnagy further advises social engineers—whom he trains through his books, his formerly prominent role at DEF CON, and his consulting firm—to exploit "the fact that gender bias, racial bias, age bias, and status bias (as well as combinations of those biases) exist."[67] What if one of his students is directly affected by these biases? He says nothing about that. Such guidance on manipulating social relationships based on these biases reveals much about the perspectives that hold power and privilege in society. Hadnagy's viewpoints have served as the ethical framework widely taught within the cybersecurity community until recently. It all comes off as a little cringey in retrospect. At one point, he admits to feeling "like some kind of mind-reading superhero."[68]

A cornerstone of social engineering is the construction of a pretext—a fabricated scenario that intentionally melds truth and deception. Hadnagy recommends "that your pretext should be based on facts, emotions, and knowledge that you already possess or can easily fake."[69] After gathering sufficient information online, the social engineer targets an individual to influence. This is often done by exploiting innate human tendencies such as the inclination for "reciprocity," the response to a sense of "obligation," the acceptance of ideas under the guise of "concession," the creation of urgency through the notion of "scarcity," and the appeal to "authority."[70]

Interestingly, Hadnagy employs these very techniques in his "Official Statement" released after his expulsion of DEF CON. He opens his statement by directly addressing his intended audience: "As many of *you* are aware, DEF CON recently announced that *they* were banning me from their conferences."[71] In this construction, readers are linguistically positioned as the second-person "you," distinct from the third-person "they" that refers to DEF CON organizers. Immediately, readers of this document are implicitly asked to *concede*, through syntactic relationship, that we are not with DEF CON and by extension not with the decision to ban.

In the next sentence, Hadnagy claims a larger moral *authority* by speaking not just for himself but also on behalf of his family and co-workers: "It's important to me, for my family and coworkers that I share with you the following thoughts."[72] Notably, he refrains from mounting a defense or argument; instead, he opts to merely share *thoughts*. A defensive argument could be construed as an implicit admission of guilt, and Hadnagy advises against defending oneself in emotionally charged situations.[73] It's evident that he is adhering to his own previously outlined playbook.

Next, Hadnagy claims a that there is a *scarcity* of facts: "Much of the speculation on social media seems to assume sexual misconduct."[74] He appears to downplay the allegations as mere speculation and assumptions. He then couples this claimed scarcity of facts with an appeal to authority: "While I still have not heard the exact allegations, a person closely affiliated with DEF CON has informed me that the allegations are NOT related to sexual misconduct."[75] While he claims not to know the "exact allegations," it seems implausible that DEF CON organizers would not communicate the nature of the complaint against him. The purported scarcity of facts serves as a ruse to inject the *authority* of “a person closely affiliated with DEF CON.” The identity of this person is withheld, as is their relationship to the event. Playing a *scarcity* of fact against an assumed *authority* intentionally re-orient the allegations away from allegations of sexual misconduct. DEF CON organizers decided that they had enough information to ban a long-standing and then highly regarded Social Engineering Village leader.

Continuing to employ his own social engineering techniques to navigate the allegations, Hadnagy appeals to the principle of reciprocity to demonstrate his goodwill: “Others have assumed that I want to know who the accusers are. I have no desire for the accuser’s identity to be made public. I understand the sensitivity of that information and I would do the same regarding protecting their identity.” These “others” are another strawman, a fictional persona that serves as a convenient foil for his own arguments. He claims that he has no desire for his accuser’s identity to be made public, which is different from making their identity known to Hadnagy privately. There is a false *reciprocity* in this claim because he is in no position to “do the same” because he has already been publicly accused! The sense that he would do the same, if the shoe were on the other foot, is misleading because we are discussing these real allegations and not some imaginary, future misdeed committed against Hadnagy. By using a counterfactual “what if” scenario, Hadnagy is imagining himself as the victim rather than demonstrating concern for his accusers.

Beyond the veneer of false reciprocity and counterfactual scenarios, Hadnagy has expressed a clear stance against anonymous complaints. Speaking at the 2021 Idaho Falls BSides Cybersecurity Conference, where he just wanted to share some his views on “cancel culture facts,” he explicitly says that the Social Engineering Village at DEF CON does not allow “anonymous complaints” of any kind, except those of “a violent nature.” He further rationalized this by saying that allowing anonymous complaints “empowers” anyone to make any complaint they want.[76] It’s strange that he is worried about empowering others, especially empowering people to make complaints. Again, it is an oddly specific thing to say.

Returning to his “Official Statement” on the DEF CON ban, it is noteworthy that Hadnagy emphasizes his own moral *obligation*—and by extension, ours—to avoid causing further harm, specifically to him: “I am not trying to feed the negativity that surrounds this situation," he writes. "Instead, my focus and attention are on my family, my employees, and others connected to me who have been hurt by this announcement."[77] By framing it this way, Hadnagy claims the moral high ground, purporting to defend not just himself but also his family, employees, and "others connected to me who have been hurt by this announcement." Put simply, he centers harm against himself, while diminishing the harm implied in the allegations, and uses his friends and family as a shield.

What is more sacrosanct than one's family? The concept of family carries significant social weight, often invoking a sense of shared social *authority* and familiar *obligation*. Family is based on mutual *reciprocity* and *concessions* made with ones we love to fulfill roles and expectations. It serves as a stable identity anchor for many and is deeply rooted in private, personal values. Because family is a universal concept that resonates across different social structures, it becomes an ideal tool for manipulation. In his chapter entitled "I Can See What You Didn't Say" in *Social Engineering* (2018), Hadnagy even uses several images of his own family to discuss how to "read nonverbal communication."[78] The inclusion of these intimate, personal photographs, taken in the heart of his home—his family kitchen—serves to misdirect critical readings. After all, what kind of terrible person would critique a work that so openly shares such private and personal moments?

In Hadnagy's conception of social engineering, emotions serve merely as an interface to be manipulated, a perspective that arises from a narrowly defined set of cultural values. Let's be explicit about what those values are: they often align with cis straight white conceptions of family. This limited set of values also limits defensive social engineering. A narrow point of view assumes that attackers will share these same cultural values. When security culture stands as an expression of values, it can also serve as an expression of simplistic demographic majority views and stereotypes. While existing social engineering frameworks may be effective within populations where these majority viewpoints hold sway, they fall short of providing a resilient and adaptable defense against social engineering over the long term. A progressive, evolving mental model of manipulation should instead center on empathetic capacity associated with emotional maturity and an observational mindset.

The diversity of human identity is a strength to be embraced. Groups become harder to target when they do not neatly fit within demographic boxes. While family is a universal concept, it's important to recognize that the term can encompass both families of origin and chosen families. Unique families limit the opportunities of scammers to assume our relationships. However, it's crucial to note that the expectations within these intimate relationships are often shaped more by societal norms than by the individuals involved. Returning to *The Managed Heart*, the seminal book that first defined the concept of emotional labor, it is interesting to see how the dynamics of any relationship have been described as complex and intersectional for a long time:

> The family is often considered a "relief zone" away from the pressures of work, a place where one is free to be oneself. It may indeed be a refuge from the emotion work required on the job, but it quietly imposes emotional obligations of its own. […] Any bond like the one between parent and child is subject to ambivalence and the rules that contain it. The child loves and hates the parent, and the parent loves and hates the child. But cultural rules in each case prescribe acceptable mixes of feelings.[79]

As long as the discipline of social engineering draws core examples from folksy stereotypes—such as those related to family, pets, and work—critical cybersecurity education will be static and unable to adapt. The cultural rules that guide social relationships are constantly being negotiated, adapted, and augmented. While progressive attitudes embrace this dynamism, conservative viewpoints cling to the past in predictable ways.

Hadnagy's notion that social engineering is an "equal opportunity victimizer" may hold some truth, but it is not true the way that he intends. According to him, anyone can fall prey to a social engineering attack, not just the "dumb humans."[80] It is an imperative to hone a more confident style of social engineering education that is both realistic and blunt. An attention to security rhetoric is suspicious of requests of reciprocity, obedience through obligation, concessions to bullies, scarcity mindsets, and unearned authority. These are negative definitions that describe how an attention to security rhetoric helps identify hostile social engineering. What if these negative definitions were described in a positive sense?

A counterfactual approach to social engineering would shift the focus from attackers to defenders, emphasizing the attitudes necessary to guard against malicious messages. In this way, those attuned to security rhetoric will react critically to the demands of someone in an authority position;

they will question the societal roles and expectations that impose arbitrary obligations to gain obedience; they challenge and confront bullies when they can, or they will ask for help if they are threatened; and those aware of the security needs of vulnerable people will generously offer assistance and advice; those concerned with security rhetoric understand that "security is a team sport" that must be played at a community level, wherein a consensus-driven security culture permeates education and work contexts. Those attuned to security rhetoric will be confident in their ability to learn independently and make decisions collaboratively. They will also use what power they possess to help and support those struggling to assert their own agency in social and technical systems. These are actively engaged digital citizens who are aware of the laws and regulations governing online services and recent security issues. They approach the digital landscape with a confident skepticism, finding humor in the tactics employed by scammers.

Ethical Phish

If your workplace begins referring to employees as family, you know that they are either not paying you enough, cutting your benefits, or conducting a phishing simulation. The ethics of ethical phishing has long been an issue within institutions with many managed accounts. Universities, corporations, and governments often test their employees by asking (or forcing) them to participate in "phishing simulations." A phishing simulation allows IT security to test users' ability to detect deceit, like email or any other cloud service with messaging—by attempting to tricking them into clicking, downloading, or otherwise engage in risky behavior that would result in a point of compromise. The simulation is generally designed to test security policies and practices, as well as train users to reduce susceptibility to legitimate attacks. There is a broad consensus among vendors of phishing simulation services about some basic ground rules: avoid playing on an employees' trust in the organization, including a lure predicated on bonus pay-outs, losing work benefits, termination notices, loss of personal possessions in the workplace, or personal company-protected data leaks; avoid threatening an employees' homes or personal sense of security;[81] avoid embarrassing or shaming employees; and avoid punishing employees for making a mistake during a simulation.[82] It is more than a little concerning that any of these ethical practices need to be spelled out, but here we are.

To conduct an ethical phishing simulation, several best practices are recommended. First, a pre-launch campaign should be initiated to clearly define the purpose and scope of the phishing simulation, establish lines of communication for the duration of the test, and ensure that managers are equipped to support their teams. Second, the simulation should be designed with the mental well-being of participants in mind, considering the potential for heightened anxiety or panic. Lastly, it is crucial that the results of the simulation are communicated transparently to employees, and that any subsequent training is framed as a supportive measure rather than a punitive action.[83]

There are many vendors for these services and all large cloud service providers offer documentation on how they will run a successful phishing simulation. While this is a common security service offering that helps managers feel as though their employees are well trained, the effectiveness of human subject experimentation as a testing and training exercise is not clear. There is good research suggesting that phishing simulations do very little to reduce susceptibility to phishing in the workplace, specifically in the university context, in actual terms.[84] The authors of a study exploring the susceptibility to phishing in the university-based workplace report "67 percent of employees who respond to simulated phishing attacks are repeat victims and therefore likely to respond to phishing emails more than once."[85]

The authors of this study used many of the terms associated with Hadnagy's central tenets of manipulation and influence defined under social engineering, including authority, urgency, reciprocity, and scarcity.[86] The authors of this study went on to discover that urgency and authority cues "contribute to increased susceptibility with a workplace setting."[87] Perhaps most surprising, this study concludes that "these findings provide novel evidence that such concepts also apply within a workplace domain where employees have previously received a base level of cyber security training."[88] This suggests that the simulations may not be as effective in changing long-term behavior as one might hope. It raises questions about whether the simulations are designed in a way that truly educates employees about the risks and signs of phishing, or whether they simply induce a temporary state of heightened awareness that fades over time.

The effectiveness of phishing simulations is questionable, even when employees have had cybersecurity awareness training. The situation only deteriorates further when considering specifically targeted spear phishing attacks that contain authentic *legitimacy cues*: "This suggests that individuals may be unaware of their vulnerability to the influence techniques commonly contained within spear phishing emails, representing a gap in current knowledge."[89] There is even a genre of online writing that chronicles how brave online writers will engage with phishers by replying to these the message, but this is not recommended.[90] Quite often, the reality of these scams is banal and mechanical than the dramatic insights offered from credible social science research.

Employees in universities and other workplaces remain susceptible to attacks even after cybersecurity awareness training. Most employees who are susceptible to social engineering attacks repeat their mistakes. Authority and urgency remain the most effective affective tools to manipulate and influence, especially in a work context. The hierarchical relationships defined by the "org-chart" suggest that top-down authority is a systematic vulnerability for organizations. Workplaces that prioritize speed and obedience in correspondence will be more vulnerable. Fast-paced, email-heavy workloads centralize daily communications, alongside urgent decision-making and spending authorizations.

The readily available Open Source Intelligence (OSINT) information for most institutions—whether it is a public event, social media, or company blog post—means that spear phishing attacks can possess a real sense of authenticity with ease. More concerning is the evidence that existing heuristics are insufficient for making adequate gains in identifying phishing emails. The use of phishing simulations is either too limited or harmful to employees to foster greater critical capacity within an institution. If cybersecurity is a team sport, harassing and tricking your employees may not be the best approach.[91] Any ethical limits set on phishing attacks only serve to highlight to legitimate attackers' points of urgency and authority to exploit: "Your position has now been terminated due to budgetary constraints. Click here to receive your severance package"; "Due to your exceptional performance, your annual bonus has been increased. The details of this compensation package are attached to this email." While appeals to authority and urgency may continue to be effective in social engineering attacks, building a resilient security culture does not justify using hostile tactics. As I will argue here, an increasing a sense of care and empathy, paired with a realistic threat model, may be the best approach to the problem of social engineering attacks.

In the 1960s, social psychologist Stanley Milgram ran a now famous series of experiments to test obedience to authority at Yale University.[92] These are the experiments that asked participants to follow orders and deliver bogus electric shocks to a victim. Though the experimental subject delivering the (fake) electric shocks cannot see the victim, they can hear the (fake) screams of pain until, after a lethal dose of electricity is (falsely) administered, the victim falls ominously silent. The findings are resonant today because the experiment highlighted a very human susceptibility to authority and susceptibility to "just follow orders," even when actions cause horrible pain or death to others. The consequences resonate in atrocities and crimes against humanity from the Holocaust to Abu Ghraib.

Alan J. Kimmel's opinion piece for the *British Psychological Society*, titled "Deception in Psychological Research—a necessary evil?" summarizes years of research into deceptive practices in research ethics. He discusses Milgram's legacy and how deception played a central role in psychological experimentation for decades but raised deep ethical concerns for the treatment of participants in these studies. The extreme stress and guilt resulting from the belief that they had harmed or killed innocent people should have, in retrospect, concluded the study. Defenders of deceptive tactics in social science research claim that these "technical illusions" are a "necessary evil" to increase the impact of an "experimental situation."[93] Kimmel concludes by reflecting on how "the days during which deception was used more out of convention than necessity and accepted without comment are long past." Furthermore, cybersecurity is an industry that trades in defending against and testing for deception, but the effectiveness of phishing simulations should be informed by this history. Moreover, the techniques used to exploit a population cannot simply be repackaged as a training exercise. Hostile techniques used in an attack cannot simply be inverted and repurposed as a defensive framework.

The stress and anxiety of deception, when placed on research participants, have been regarded as a largely unethical practice that requires very close monitoring by research ethics boards in research intensive universities and funding agencies. Social media and deceptive communications in email may be regarded as an uncontrolled, live experiment that causes stress and anxiety at a much broader scale. Since the disclosures of Frances Haugen at Meta, the public has become much more aware of how platforms like Instagram can be harmful to adolescent mental health in particular.[94] Facebook has the power to sway elections and ferment political discord, even stoke coup attempts.[95] We can understand the effect of these systems on humans, but what is the effect of so much distrust on the systems themselves.

There is, however, an increasing awareness of the role of emotion and affect in the human dimension of cybersecurity. Human error results in most cybersecurity incidents. The UK Information Commissioner’s Office claims that up to 90 percent of attacks were human error in 2021.[96] Cybersecurity writers and blogs are beginning to define “emotional intelligence” and “emotional awareness” as a step toward mitigating these human errors.[97] Recommendations are emerging around an “emotionally intelligent cybersecurity strategy,” with an emphasis on optimism, empathy, and calm.[98] These are vague, unauthored blog posts—industry-focused blogs designed to drive Search Engine Optimization (SEO) for a company website—without anything approaching a “cybersecurity strategy.” Still, maybe there is something to all this industry churn?

Symptom Language

There is a growing awareness that emotional intelligence and empathy have a real and applied value in a security context. Often termed “soft skills,” these qualities have gained prominence, particularly as the Covid-19 pandemic has prompted a reassessment of the value of technical skills and emotional intelligence. There is a sense that the Covid-19 pandemic spurred a “rebalancing between technical capability and emotional intelligence,” since working from home demanded increased independence and an increasing emphasis on human qualities that sometimes get measured as an Emotional Quotient, or EQ.[99] There is real value in assuming a position of optimism in the face of overwhelming deceit online, empathy with users and potential attackers to understand the threat environment, and calm in the face of false urgency and the stress of obedience commanded from a position of authority.

After all, it can be traumatic to be the victim of a cybersecurity breach, particularly if you are responsible for disrupting your workplace. The rippling psychological effects of a security breach situation result in negative affective stress responses. In a 2021 study of a 150 cybersecurity breach victims, researchers working out of the UK recommended that "emotional support systems" must be put in place after a security incident to "avoid negative long-term psychological consequences."[100] Increasingly, the role of emotional intelligence in a cyberattack is regarded as a strength in the wake of breach, to help manage the immediate pressures of the response and recovery phases.[101]

Emotional intelligence may be a growing issue for Chief Information Security Officers (CISOs) as well, who are increasingly facing the legal consequences of their decisions during an attack. In 2022, the US District Court in San Francisco convicted Joe Sullivan, the former CISO of Uber for concealing and not reporting the attack. He later obstructed a federal investigation into the attack by the Federal Trade Commission by concealing a new breach.[102] Public disclosures made after the conviction captured the sense of panic in phrases like, "This can't get out," "We need to keep this tightly controlled," or "This may also play very badly."[103] The concern for public relations is an emotional concern that privileges public perception over the security of user data, the use of company funds to pay ransoms, the need to cover up $100,000 payment to criminals, or any other decision made as a result of impaired judgment from public embarrassment or shame. Notably, as of July 2023, the US Securities and Exchange Commission (SEC) has implemented rules for companies to disclose cyberattacks in four days, saying "companies and investors alike, however, would benefit if this disclosure were made in a more consistent, comparable, and decision-useful way."[104]

On October 30, 2023, the SEC charged the CISO of SolarWinds with fraud and failures related to internal quality control and reporting measures.[105] Timothy Brown, the SolarWinds CISO, is being personally charged for misleading investigators. The Orion software contained vulnerabilities related to critical system data and privileged access that were, according to the complaint, inappropriate for a product with national security level significance. The errors are related to management and processes rather than an exotic exploit from Russian hackers. Another startling finding from the charges is the specific charges leveled against an individual responsible.

Regulators and industry leaders are pressing for accountability. This new period of accountability and consequences adequately defends against social engineering attacks by studying the tactics used in phishing schemes. Basic cybersecurity awareness training is insufficient for mitigating susceptibility to online manipulation and influence. Currently, emotional intelligence is often seen as a tool for managing stress during active cyberattacks. Psychological tests used for measuring emotional responses, through word associations in particular, "are probably not suitable for measuring emotions in the cybersecurity context."[106] There remains a consistent set of priorities that manifest in a deeply emotional register. These feelings include a sense of uncertainty, lack of awareness, societal harm, responsibility, and violating laws.[107] These are all deeply emotional aspects of a cyberattack. An integrated approach that blends technical literacy with the emotional maturity needed to feel calm and balanced under pressure, without minimizing the urgency of a situation.

An affective cybersecurity framework must be grounded in principles that are free from the influence of regressive notions of emotion, which often pit strong feelings against logic or reason. Such an affective model of cybersecurity should neither denigrate nor elevate emotions based on outdated social norms. As Sara Ahmed articulates in her seminal work, *The Cultural Politics of Emotion*, there exists a hierarchy that privileges certain thoughts and feelings. Ahmed notes, "The hierarchy between emotion and thought/reason gets displaced, of course, into a hierarchy between emotions: some emotions are 'elevated' as signs of cultivation, whilst others remain 'lower' as signs of weakness."[108] Security rhetoric should focus on identifying instances where emotions are valued over others and question the assumptions and points of view that are being validated in those moments.

An understanding and awareness of our strong feelings must be regarded as a litmus test for detecting influence and deception. It is necessary to integrate both the personal experience of emotions and their commercial applications. Emotional labor must be devoted to balancing this emotional load. Although Sigmund Freud's theories have been discredited in modern psychological practice, his influence on popular perceptions of emotion persists into the twenty-first century. At this moment, recall Hadnagy's analogy of "cancel culture" as confusing a visceral fear of black snakes. This specific appeal to the fear of snakes in the defense of theoretical practices in practiced in public is precisely the rhetoric Freud used to "cure" his first patient, known then as "Anna O.," with Josif Breuer. Like so much snake oil, it seems that those who claim to have mastered the human mind appeal to pseudo-Biblical rhetoric to lend (simulated) urgency and (mock) credibility to their theories.

Freud's theory of anxiety, as described in *Inhibitions, Symptoms, and Anxiety* (1926), claims that emotion serves as a "signal function"—the anxious core of the psychoanalytic theory of affect—signals danger. According to Freud, an overload of undischarged emotion spurs action and survival as a kind of "automatic anxiety."[109] Hadnagy's "amygdala hijacking" is just the twenty-first-century updated cyber-metaphor of Freud's automatic, mechanical functioning of the mind. For both Freud and Hadnagy, involuntary and automatic behaviors are a key feature of their vision of the psyche. These involuntary responses are predicated on a neat division between rational and emotional faculties, which was discredited in the last chapter. However, common sense reminds us that few feelings really signal an emergency response. It is important to notice when certain feelings are being elevated and why. Suspicion, inquisitiveness, and curiosity should not be confused for automatic anxiety or amygdala hijacking. Additionally, the emotional pressures experienced during a crisis are not intrinsically irrational; they are external factors that require social patterns to regulate emotional responses in culturally acceptable ways.

It may be significant for those who have internalized the Freudian model of mind to know that much of his founding research has been wholly debunked and marked as more or less a fraud. Todd Dufrensne, in *Killing Freud* (2003), revisits and updates many of the early criticisms of Freud, often with hilarious critical dexterity. Dufrensne refers to Mikkel Borch-Jacobsen's *Remembering Anna O.: A Century of Mystification* (1996), which argues that Freud's first patient, Anna O. (real name Bertha Pappenheim), was not actually suffering from hysteria.[110] According to Borch-Jacobsen, hysteria is a construct created to describe symptoms and largely performed by patients. The first etiology of hysteria with Anna/Bertha served as the basis for well-known and popularly accepted psychological truisms such as the unconscious, the talking cure, the repression of libidinal desires, and traumatic experiences.[111]

Through a study of contemporary criticisms of Freud, many scholars have concluded that "Pappenheim suffered from 'nothing' except the 'desire to communicate psychic distress in physical terms.'"[112] Due to the social restrictions placed on an educated, middle-class, Jewish woman living in Vienna, her behavior is better understood as merely, in Dufrensne's words, the "symptom language of the hysteric."[113] Pappendheim was coached by her doctors, and she performed symptoms for want of legitimate help. After all, Anna/Bertha was a real person living with all the stresses and pressures of being a well-known social worker and feminist at this time. Borch-Jacobsen recalls that Anna/Bertha's "cure" arrived by finding the "root of her symptoms" wherein "she relived the night when, literally paralyzed and struck mute by fright, she saw a black snake slithering toward him with the intention of biting him"; in this way, says Freud, "the whole illness was brought to a close."[114] The snake serves as a familiar refrain for those wishing to substantiate theories of mind with nothing more than a myth-making imagination and a predilection to Biblical imagery. With Dufrensne and Borch-Jacobsen distributing the patricidal burden of dismissing Freudian psychoanalysis, we are free to consider alternative models unencumbered by the libidinal, self-destructive allure of the subconscious.

An integrated techno-social framework would empower individuals to make decisions that lead to a secure future. Social engineering has always involved emotional labor. In *The Managed Heart*, Hochschild explores how the commercialization of emotion shapes the attitudes and experiences of both workers and consumers, requiring constant rebalancing and adjustment. "As workers," Hochschild notes, "the more seriously *social engineering* affects our behavior and our feelings, the more intensely we must address a new ambiguity about who is direct them."[115] The question becomes: Whose voice does a worker use? Do they speak for the company or for themselves? Many social media accounts that are connected to the individual's professional life leave the tag, "opinions my own" to disentangle this rhetorical situation that may blend professional communications with personal expression. "As customers," says Hochschild regarding the other side of commercialization of affect, "the greater our awareness of social engineering, the more effort we put into distinguishing between gestures of real personal feelings and gestures of company policy."[116] In either case, as customer or worker, we are experiencing the symptom language of capitalism in social relationships. We perform because our lives and livelihoods depend on it. Our emotional labor is just another human resource to be exploited in the production of value.

This focus on emotional labor is difficult to dismiss as merely a rant fit for r/antiwork, given that it reflects the daily reality for many workers.[117] As working individuals, we try to correct for that "social engineering of feeling" and "mentally subtract" the commercially motivated expression from a sense of humanity that is sincerely felt.[118] However, the lens of emotion labor privileges work. What protections and frameworks exist for the unemployed, the underemployed, the retired, or those working outside the traditional economy? Or for those whose work is in the home? The next innovations on the web and online security must be designed for equity and justice to include those not directly participating in the economy, especially as AI-driven automation takes hold.

We need an affective heuristic or testing protocol to mentally subtract the deceit inherent in manipulation and influence, allowing us to objectively assess the intentions behind lures. The best way to dismantle the psychological tooling of social engineers is to "mentally subtract" the influences of urgency and authority in the workplace they exploit. By acknowledging that emotional labor is exploited in our work, we can return to our objective feelings and avoid the malicious messages that imperial our institutions by regaining our autonomy. While *objective feelings* may sound like an oxymoron, Hochschild asks that we embrace and understand feeling toward a holistic awareness of events:

> The word *objective*, according to the *Random House Dictionary*, means "free from personal feelings." Yet ironically, we need feeling in order to reflect on the external or "objective" world. Taking feelings into account as clues then correcting for them may be our best short at objectivity.[119]

Social engineering attacks succeed because we allow ourselves to be exploited first by our employers, colleagues, and customers. By being immediately amenable, helpful, and demure, we inadvertently cover up the power imbalance and diminish our own agency. This situation also lends legitimacy and power to socially engineered attacks by reinforcing outdated roles that define professional relationships. It is worth restating that Hadnagy advises social engineers to exploit “the use of authority and gender differences in compliance”![120] Security culture should not perpetuate these harmful systems of bias and violence. By advocating for a more respectful work environment based on evenly distributed agency, respect, and honesty, we are more likely to identify a malicious phishing message when it appears in our inbox.

Perhaps the most cunning aspect of this approach is the need to be cautious when attackers seem to cede authority. In a chapter entitled “I Know How to Make You Like Me,” Hadnagy advocates for “ego suspension” on the part of the attacker. For example, Hadnagy recounts how he suspended his own ego and allowed an environmental organization he calls “Oil Hater” to lead. He recollects how he suspended his “own ego and allowed OilHater to be ‘the boss.’”[121] By weaponizing agency and “validating others,” Hadnagy is unintentionally pointing to the long-term solution to manipulation and influence in phishing campaigns.

In this case, Hadnagy’s description of validating others would be unlikely to trick someone possessed of an authentic sense of agency and responsibility: “Simply put,” Hadnagy says, “validation is agreeing with, complimenting, or endorsing someone else’s statements, decisions, or choices. When someone feels validated by you, their brain releases dopamine and oxytocin, which in turn enables you to create feelings of trust and rapport.”[122] Honestly, how depraved is your work culture if you get a release of dopamine because someone agrees with you? In a toxic work culture, validation feels good and builds rapport because emotional work is hierarchically defined by the org-chart.[123] In a healthy work culture that shares responsibility, dictatorial demands on another’s time come less often and appear more suspicious.

Witnessing when and how feelings matter is crucial. Emotional work is work, but a failure to grasp the significance of emotional work is also an indicator of other attitudes. In returning to Harriet Fraad, in “Toiling in the Field of Emotion,” she recounts a significant piece of the “psychohistory” of the twentieth century. In the wake of the Second World War, German academics collected under the name “The Frankfurt School” worked to understand the roots of right-wing authoritarianism. In an effort to understand and prevent the horrors of the Holocaust, they found that a majority, “about 60 percent of the vast populations they studied, passively embraced right-wing authoritarianism."[124] Another 20 percent of the population was actively and "sadistically engaged in worshiping authority."[125] The remaining 20 percent are opposed to authoritarianism:

> They found that fearful conditions such as Pre-World War II Germany's rapid inflation and the bombing of the Reichstag could precipitate the ambivalent 60 percent majority to blindly obey Hitler and the Nazi Party and condemn those who did not obey. Anyone who questioned was attacked as unpatriotic, weak or poisonous such as women, homosexuals, Communists, Socialists, Jews, etc.[126]

This sketch is hauntingly familiar. After all, the findings of the Frankfurt School, which distributes attitudes to authoritarianism, are still resonant in today's right-wing politics. Let's continue with Fraad's line of thought on George W. Bush and the post 9/11 period:

> We note that the steady decline in U.S. male wages since 1970, combined with the bombing of the World Trade Center, seem to have had the same effect in the United States. After the trauma of the World Trade Center (WTC) bombings combined with the severe economic losses suffered by male workers and their families, Americans transformed their perceptions of an unpopular president who stole votes and was selected by the court rather than elected by a majority. All those who questioned were condemned as unpatriotic, weak, terrorist sympathizers. Right wing forces that condemn feminism, homosexuality, and foreigners burgeoned.[127]

The symptom language of capitalism, sexism, and homophobia must be transformed to reject the cultural context that enables attacks. These attacks prey on hierarchies and rigid social roles. Late-stage capitalism and the authority of middle management serve as legitimate cultural attack vectors, especially in an increasingly exploited and weary working class.[128]

These engrained automatic anxieties function as proxies for false authority, designed to maintain hierarchical social structures, justify wage gaps, and perpetuate inequities. Those who manage and manipulate others' labor in dehumanizing ways set the stage for a worker to satisfy an attackers request as just another request in a series of seemingly pointless decisions. There is a broad sense of security found in being able to identify authentic emotional maturity, rather than rehearsing toxic work culture.[129] Setting meaningful boundaries in work communications and flattening organizational hierarchies to empower individuals can be an important part of the solution. With emotional boundaries in place, it becomes easier to resist urgent appeals to authority. "Can you send me $1,000 with your company credit card?" Hell, no. "Please click this link to find your raise amount." Nope. "I'm a Nigerian Prince!" That's a hard pass.

Notes

[1] Emily Bauer, “15 Outrageous Email Spam Statistics that Still Ring True in 2018,” *Propeller Blog*, February 1, 2018, https://www.propellercrm.com/blog/email-spam-statistics; Dashlane, “Phishing Statistics: What Every Business Needs to Know,” *Dashlane Blog*, January 17, 2018, https://blog.dashlane.com/phishing-statistics/.

[2] Kemba Walden, “The Growing Threat of Ransomware,” *Microsoft Corporate Blog*, July 20, 2021, https://blogs.microsoft.com/on-the-issues/2021/07/20/the-growing-threat-of-ransomware/.

[3] Microsoft Office files like Excel and Word can run Visual Basic Applications (VBA) called macros, which is an easy and common way to deliver malware. In July 2023, Microsoft claimed they would disable macros by default, a major win for security. See https://learn.microsoft.com/en-us/deployoffice/security/internet-macros-blocked.

[4] Names like ARCON, BeyondTrust, Centrify, and Symantec sell Software as a Service applications that privileged access management processes in 2023.

[5] Jon Porter, “When Can We Finally Get Rid of Passwords?” *The Verge*, April 24, 2019, https://www.theverge.com/2019/4/24/18514225/passwords-fido2-authentication-webauthn-security-key-cybersecurity-online-browser-web; Ben Dickson, “Why Passwords Might (Finally) Go Away,” *PCMag*, October 31, 2018, https://www.pcmag.com/opinions/why-passwords-might-finally-go-away.

[6] Martin Paul Eve, *Password* (New York: Bloomsbury, 2016); Brian Lennon, *Passwords: Philology, Security, Authentication* (Cambridge, MA: Belknap Press, 2018).

[7] Eve, *Password*, 9.

[8] “How FIDO Addresses a Full Range of Use Cases,” *FIDO Press Release*, March 2022, https://media.fidoalliance.org/wp-

content/uploads/2022/03/How-FIDO-Addresses-a-Full-Range-of-Use-Cases.pdf.

[9] Google, Microsoft, Apple, Amazon, and others have expressed their commitment to the new standard in authentication. See Sampath Srinivas, "One Step Closer to a Passwordless Future," *Google Blog*, May 5, 2022, https://blog.google/technology/safety-security/one-step-closer-to-a-passwordless-future/; Alex Simons, "Expansion of FIDO Standard and New Updates for Microsoft Passwordless Solutions," *Microsoft Tech Community Blog*, May 5, 2022, https://blog.google/technology/safety-security/one-step-closer-to-a-passwordless-future/; "AWS IAM Now Supports FIDO2 for Multi-Factor Authentication in AWS GovCloud (US) Regions," June 23, 2023, https://aws.amazon.com/about-aws/whats-new/2023/06/aws-iam-fido2-multi-authentication-govcloud-regions/; and "Apple, Google, and Microsoft Commit to Expanded Support for FIDO Standard to Accelerate Availability of Passwordless Sign-ins," *Apple Newsroom*, May 5, 2022, https://www.apple.com/newsroom/2022/05/apple-google-and-microsoft-commit-to-expanded-support-for-fido-standard/.

[10] See https://fidoalliance.org/fido2/.

[11] Randall Munroe, "Password Strength," *xkcd: A Webcomic of Romance Sarcasm, Math, and Language*, https://xkcd.com/936/. https://xkcd.com/936/.

[12] Dan Goodin, "Ongoing Phishing Campaign Can Hack You Even When You're Protected with MFA," *Ars Technica*, July 12, 2022, https://arstechnica.com/information-technology/2022/07/microsoft-details-phishing-campaign-that-can-hijack-mfa-protected-accounts/; Dan Goodin, "Microsoft Takes Pains to Obscure Role in 0-Days that Caused Email Breach," *Ars Technica*, July 14, 2023, https://arstechnica.com/security/2023/07/microsoft-takes-pains-to-obscure-role-in-0-days-that-caused-email-breach/.

[13] Password recycling can expose multiple accounts. Additionally, incrementing integers in response to IT password reset requests are easy to brute force in a credential stuffing campaign. These are daily,

immediate, and urgent risks online today. Using password managers to generate long, unique passphrases centralizes vulnerability around a single platform, like in the case with the LastPass breaches in 2022. See Lawrence Abrams, “LastPass Developer Systems Hacked to Steal Source Code,” *Beeping Computer*, August 25, 2022, https://www.bleepingcomputer.com/news/security/lastpass-developer-systems-hacked-to-steal-source-code/.

[14] According to Kyle Lady, Senior Information Security Engineer at Duo, 28 percent of people use a second factor of authentication in 2017. 2FA usage has climbed to 78 percent in 2021, but the vast majority (85 percent) of these are the phishable SMS-based authentication. See Kyle Lady, “State of the Auth: Experiences and Perceptions of Multi-Factor Authentication,” *Duo Labs Blog*, November 7, 2017, https://duo.com/blog/state-of-the-auth-experiences-and-perceptions-of-multi-factor-authentication; Chrysta Cherrie, “The 2021 State of the Auth Report: 2FA Climbs, While Password Managers and Biometrics Trend,” *Duo Labs Blog*, September 14, 2021, https://duo.com/blog/the-2021-state-of-the-auth-report-2fa-climbs-password-managers-biometrics-trend.

[15] See https://fidoalliance.org/.

[16] See https://www.w3.org/TR/webauthn/. WebAuthn also hosts a suite of online tools to help demonstrate the technology and provide detailed documentation for developers. See https://webauthn.guide/ for a simplified documentation. See also https://webauthn.io/ (designed by Duo Labs) to see a demo working in a typical browser context. FIDO has its own documentation to implement their own WebAuthn-based standard called CTAP (Client to Authentication Protocol). See https://fidoalliance.org/specs/fido-v2.0-ps-20190130/fido-client-to-authenticator-protocol-v2.0-ps-20190130.html.

[17] See https://pages.nist.gov/800-63-3/sp800-63-3.html.

[18] “How FIDO Addresses a Full Range of Use Cases,” *FIDO Press Release*, March 2022, https://media.fidoalliance.org/wp-content/uploads/2022/03/How-FIDO-Addresses-a-Full-Range-of-Use-Cases.pdf.

[19] Brian Krebs, “Your Phone May Soon Replace Many of Your Passwords,” *Krebs on Security,* May 7, 2022, https://krebsonsecurity.com/2022/05/your-phone-may-soon-replace-many-of-your-passwords/.

[20] Tim Wu, “The Tyranny of Convenience,” *New York Times, Opinion,* February 16, 2018, https://www.nytimes.com/2018/02/16/opinion/sunday/tyranny-convenience.html.

[21] Jonathan Margolis, “Smart Homes Are a Dystopian Nightmare,” *Financial Times*, April 25, 2018, https://www.ft.com/content/bf3ba564-4715-11e8-8c77-ff51caedcde6.

[22] Peter Valdes-Dapena, “The Car of the Future? Ford Applies for a Patent on Car that Can Automatically Repossess Itself,” *CNN Business*, March 3, 2023, https://www.cnn.com/2023/03/03/business/ford-repossessing-car-patent/index.html.

[23] Josephine Wolff, “A Brief History of Cyberinsurance,” *Slate*, August 30, 2022, https://slate.com/technology/2022/08/cyberinsurance-history-regulation.html.

[24] Daniel Woods, “The Insurance Industry and Offensive Cyber Operations: Slow and Steady Wins the Race?” *Offensive Cyber Working Group*, July 11, 2022, https://offensivecyber.org/2022/07/11/the-insurance-industry-and-offensive-cyber-operations-slow-and-steady-wins-the-race/; Patrick Davison, “Cyber War and Cyber Operations Exclusion Clauses,” *Lloyd’s Market Association Bulletin*, November 25, 2021, https://www.lmalloyds.com/LMA/News/LMA_bulletins/LMA_Bulletins/LMA21-042-PD.aspx.

[25] Adam Shostack, *Threat Modeling: Designing for Security* (Indianapolis, IN: Wiley & Sons, 2014), 395.

[26] See https://www.mailmodo.com/guides/spam-filters/.

[27] Rachna Dhamija, J.D. Tygar, and Marti Hearst, “Why Phishing Works,” *CHI 2006 Proceedings in Security* (2006): 581–90.

[28] Yue Zhang, Serge Egelman, Lorrie Cranor, and Jason Hong, “Phinding Phish: Evaluating Anti-Phishing Tools,” *CMU Center for Compuational Analysis of Social and Organizational Systems*, 2006, http://www.casos.cs.cmu.edu/publications/papers/ndss-phish-tools-final.pdf.

[29] Ian Fette, Norman Sadeh, and Anthony Tomasic, “Learning to Detect Phishing Emails,” *WWW ’07: Proceedings of the 16th International Conference on World Wide Web* (2007): 649–56. Available: https://doi.org/10.1145/1242572.1242660.

[30] Nalin Asanka Gamagedara Archchilage, and Steve Love, “A Game Design Framework for Avoiding Phishing Attacks,” *Computers in Human Behavior* 29 (2013): 706–14.

[31] Neda Abdelhamid, Aladdin Ayesh, and Fadi Thabtah, “Phishing Detection Based Associative Classification Data Mining,” *Expert Systems with Applications* 41 (2014): 5948–59.

[32] Mahmood Moghimi, and Ali Yazdian Varjani, “New Rule-Based Phishing Detection Method,” *Expert Systems with Applications* 53 (2016): 231–42.

[33] Kathryn Parsons, Agata McCormac, Malcolm Pattinson, Marcus Butavicius, and Cate Jarram, “The Design of Phishing Studies: Challenges for Researchers,” *Computers and Security* 52 (2015): 194–206; Steve Sheng, Brad Wardman, Gary Warner, Lorrie Faith Cranor, Jason Hong, and Chengshan Zhang, “An Empirical Analysis of Phishing Blacklists,” *CEAS 2009—Sixth Conference on Email and Anti-Spam* (2009): https://www.uab.edu/cas/thecenter/images/Documents/An-Empirical-Analysis-of-Phishing-Blacklists.pdf.

[34] Ibid., 203.

[35] Melanie Volkamer, Karen Renaud, Benjamin Reinheimer, and Alexandra Kunz, “User Experiences of TORPEDO: Tooltip-poweRed Phishing Email DetectiOn,” *Computers and Security* 71 (2017): 100–13.

36 B.B. Gupta, Aakanksha Tewari, Ankit Kumar Jain, and Dharma P. Agrawal, “Fighting Against Phishing Attacks: State of the Art and Future Challenges,” *Neural Computing Applications Forum* 28 (2017): 3629–54.

37 Mahmoud Khonji, Youssef Iraqi, and Andrew Jones, “Phishing Detection: A Literature Survey,” *IEEE Communications Surveys and Tutorials* 15 no. 4 (2013): 2091–121.

38 Rami M. Mohammad, Fadi Thabtah, and Lee McCluskey, “Tutorial and Critical Analysis of Phishing Websites Methods,” *Computer Science Review* 17 (2015): 1–24.

39 Jason Hong, “The State of Phishing Attacks,” *Communications of the ACM* 55 no. 1 (2012): 74–81.

40 In a collection of interviews on various cybersecurity experts, *Tribe of Hackers*, Hadnagy is described as having “established the world’s first social engineering penetration testing framework at www.social-engineer.org,” as well as having “created the first hands-on social engineering training course and certification, Advanced Practical Social Engineering, attended by law enforcement, military, and private-sector professionals.” Much of the interview is concerned with how he can encourage his employees to be “loyal to me.” See Marcus J. Carey, and Jennifer Jin, *Tribe of Hackers: Security Leaders* (Indianapolis, IN: John Wiley and Sons Inc., 2020), 147–52.

41 Cameron H. Malin, Terry Gudaitis, Thomas J. Holt, and Max Kilger, *Deception in the Digital Age: Exploiting and Defending Human Targets Through Computer-Mediated Communications* (London: Academic Press, 2017), 250.

42 There are many classic texts from this field, including Craig A. Shue et al., “Spamology: A Study of Spam Origins,” *6th Conference on Email and Anti-Spam*, July 16, 2009, https://www.ornl.gov/publication/spamology-study-spam-origins; there are also historical databases, granting evidence in the study of the evolution of phishing, for example http://untroubled.org/spam/ and https://www.projecthoneypot.org.

[43] I am not entertaining any fMRI studies, such as Simon Baron-Cohen at Cambridge University's Autism Research Centre, who claims that fMRIs can give us a view of the human mind. Andrew Scull, noted medical historian and author of *Psychiatry and Its Discontents* (Oakland, CA: University of California Press, 2019), who describes this and other research of its ilk as "nothing more than pseudoscience, a modern version of phrenology" (21). Furthermore, Scull attributes contemporary mainstream psychiatry repentance of Freud as resulting in the wholesale embrace of "the magic potions of modern psychopharmacology" (83). Having sold their "souls to Big Pharma," modern psychiatric practice was further corrupted by "burying Freud," wherein "Biobabble replaced psychobabble" (83). Barrett confirms this skepticism, in *How Emotions Are Made*, through her own laboratory studies: "Overall we found that *no brain region contained the fingerprint for any single emotion*" (22, emphasis in original).

[44] See https://www.ibm.com/topics/social-engineering.

[45] Christopher Hadnagy, *Social Engineering: The Science of Human Hacking* (Indianapolis, IN: John Wiley and Sons Inc., 2018), xviii.

[46] Christopher Hadnagy, and Michele Fincher, *Phishing Dark Waters: The Offensive and Defensive Sides of Malicious E-mails* (Indianapolis, IN: John Wiley and Sons Inc., 2015), 2.

[47] Gunter Ollmann, "The Phishing Guide: Understanding and Preventing Phishing Attacks," *Technical Info*, January 31, 2011, http://www.technicalinfo.net/papers/Phishing.html. See also https://2600.com/.

[48] The Jargon File was the first collection of the terminology associated with the earliest hacker phraseology: https://www.catb.org/jargon/. Founded by Eric Raymond, author of *The Cathedraland the Bizarre* (2001) as well as *The New Hacker Dictionary* (1996), the Jargon File is now preserved on several sites including a historical collection and a GitHub repo. There is an informal effort to preserve this material free from Raymond's influence. His personal web page, which he appends to his author profile on Amazon.com, has a running head that describes the contents as "Armed and Dangerous: Sex, software, politics, and

firearms. Life's simple pleasures …" (See esr.ibiblio.org.) The site contains a blog post-dated September 4, 2020 entitled "Kyle Rittenhouse and the militia obligation" that celebrates the death of three Communists. (See http://esr.ibiblio.org/?p=8752.) Raymond is hardly the foundation for a sophisticated understanding of rhetoric, but it is an important reminder of the libertarian, often hostile, language of American style freedom that can influence such communities.

[49] It should be said that Hadnagy's principles of social engineering appear in as a more complete and compelling form in Randy Burkett's "Rethinking an Old Approach: An Alternative Framework for Agent Recruitment: From MICE to RASCLS," published by the CIA's *Studies in Intelligence*, wherein the old Cold War perspective that money, ideology, coercion, and ego (MICE) is replaced by reciprocation, authority, scarcity, commitment, liking, and social proof (RASCLS).

[50] It is important to emphasize that the term "cancel culture" is Hadnagy's own phrasing. While much has been said about "cancel culture," I personally do not believe it to be a genuine phenomenon. Those who decry "cancel culture" often seem to be responding to widespread criticism of their own regressive views. To employ a rhetorical term, "cancel culture" serves as nothing more than a strawman argument.

[51] See Hadnagy's Idaho Falls BSides Cybersecurity Conference presentation from October 2, 2021 on YouTube: https://www.youtube.com/watch?v=YJx-559_rBI&t=16987s. He cites this presentation as part of his defense against being banned from DEF CON.

[52] Ibid.

[53] Ibid.

[54] Hadnagy, and Fincher, *Phishing*, 6.

[55] Ibid., 44, 48.

[56] Ibid.

[57] See https://defcon.org/html/links/dc-code-of-conduct.html.

[58] Ibid.

[59] See DEF CON 29 “transparency report” from February 9, 2022: https://defcon.org/html/links/dc-transparency.html.

[60] Shaun Nichols, “DEF CON Bans Social Engineering Expert Chris Hadnagy,” *Tech Target Security*, February 10, 2022, https://www.techtarget.com/searchsecurity/news/252513274/DEF-CON-bans-social-engineering-expert-Chris-Hadnagy.

[61] Chris Hadnagy, “Chris Hadnagy’s Official Statement,” *Social Engineer Blog*, February 25, 2022, https://www.social-engineer.org/general-blog/chris-hadnagys-official-statement/.

[62] Nichols, “DEF CON Bans,” *Tech Target Security*, 2022. Alyssa Miller was very visible during this period on Twitter, where her handle is @AlyssaM_InfoSec. A moment-to-moment breakdown of the events at DEF CON and the later resignations of the Cleveland BSides executive can be found on her Twitter feed. She uses the same handle on Mastodon’s infosec.exchange. She is the author of the *Cybersecurity Career Guide* (2022).

[63] See Hadnagy v. Moss proceedings: https://www.courtlistener.com/docket/64866230/hadnagy-v-moss/.

[64] In his interview in *Tribe of Hackers* (2020), Hadnagy claims that his favorite books include Amy Cuddy, *Presence: Bringing Your Boldest Self to Your Biggest Challenges* (New York: Little Brown Spark, 2015)is; Joe Navarro, *What Every Body Is Saying: An Ex-FBI Agent’s Guide to Speed-Reading People* (New York: William Morrow Paperbacks, 2008); Robin Dreeke, *It’s Not all About Me: The Top Ten Techniques for Building Quick Rapport with Anyone* (Robin K. Dreeke, Amazon, 2011); and Paul Ekman, *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication* (New York: Holt Paperbacks, 2007). From this list, Ekman’s studies of emotion date to the 1960s, and the work of psychologist Silvan S. Tomkins, who was Ekman’s mentor. He practices a style of emotional analysis described as “basic emotion method studies.” Despite the long-standing tradition of using facial expressions to measure and identify emotion, these

studies have been debunked because of the indirect relationship between expression and subjective feeling. Facial electromyography (EMG) studies attack electrodes to facial muscles to gain a finer measurement of facial response, which is of course itself an indirect measure of feeling. Lisa Feldman Barrett describes the limitations in the basic emotion method and the desire to grasp a universal measure of emotion from facial or bodily cues. See Barrett, “The Search for Emotion’s ‘Fingerprints’,” in *How Emotions Are Made: The Secret Life of the Brain* (New York: Mariner Books, 2017): 1–24.

[65] Hadnagy, *Social Engineering, 6.*

[66] Ibid., 6.

[67] Ibid., 6.

[68] Ibid., 28.

[69] Ibid., 87.

[70] Ibid., 125, 128, 131, 134, and 137.

[71] Hadnagy, “Chris Hadnagy’s Official Statement.” Emphasis mine.

[72] Ibid.

[73] Hadnagy advises in his Idaho Falls BSides Cybersecurity Conference talk that a defense in the face of “cancel culture” is counterproductive. See https://www.youtube.com/live/YJx-559_rBI?feature=share&t=18445.

[74] Hadnagy, “Chris Hadnagy’s Official Statement.”

[75] Ibid.

[76] See https://www.youtube.com/live/YJx-559_rBI?feature=share&t=18556.

[77] Hadnagy, “Chris Hadnagy’s Official Statement.”

[78] Hadnagy, *Social Engineering*, 184.

79 Arlie Russell Hochschild, *The Managed Heart: Commercialization of Human Feeling* (Berkeley, CA: University of California Press, 2012), 57.

80 Hadnagy, *Social Engineering*, 224.

81 See “5 Questions to Decide If Your Phishing Simulation Training Is Ethically Sound,” *Living Security Blog*, August 25, 2021, https://www.livingsecurity.com/blog/questions-to-decide-if-your-phishing-simulation-training-is-ethically-sound.

82 Ian Muscat, “Keeping Phishing Tests Ethical,” *Phish Deck Blog*, August 10, 2021, https://www.phishdeck.com/blog/keeping-phishing-tests-ethical/.

83 Conor Mckenna, “The Ethics of Ethical Phishing: How to Ethically Phish Your Employees,” *The Security Company Blog*, July 10, 2019, https://thesecuritycompany.com/the-insider/the-ethics-of-ethical-phishing/.

84 Emma J. Williams, Joanne Hinds, and Adam N. Joinson, “Exploring Susceptibility to Phishing in the Workplace,” *International Journal of Human-Computer Studies* 120 (2018): 1–13. Available: https://doi.org/10.1016/j.ijhcs.2018.06.004.

85 Ibid., 1.

86 Ibid., 2.

87 Ibid., 9.

88 Ibid., 9.

89 Ibid., 9–10.

90 Brittni Devlin, “Explaining an Impossible Situation,” *Make Use Of*, July 8, 2021, https://www.makeuseof.com/what-happened-responded-to-phishing-email/.

91 Tom Pendergast, in the opinion article, “Is All Fair in Simulated Phishing?” (*CSO Online*, November 16, 2017) has argued that ethical simulated phishing is like taking a vaccine to inoculate against social

engineering attacks. The analogy is appropriate if only to highlight the controversial nature of such tactics. See https://www.csoonline.com/article/563615/is-all-fair-in-simulated-phishing.html.

92 Miligram’s study was originally published in an article called “Behavioral Study of Obedience,” *theJournal of Abnormal and Social Psychology* 67 no. 4 (1963). In 1974, he published the book length *Obedience to Authority: An Experimental View* (New York: Harper Collins).

93 Allan J. Kimmel, “Deception in Psychological Research—A Necessary Evil?” *The British Psychological Society,* August 26, 2011, https://www.bps.org.uk/psychologist/deception-psychological-research-necessary-evil.

94 Jeff Horwitz, “The Facebook Files,” *The Wall Street Journal,* October 5, 2021, https://www.wsj.com/articles/the-facebook-files-11631713039.

95 Robert M. Bond, Christopher J. Fariss, Jason J. Jones et al., “A 61-Million-Person Experiment in Social Influence and Political Mobilization,” *Nature* 489 (2012): 295–8. https://doi.org/10.1038/nature 11421.

96 Michael Hill, “90 percent of UK Data Breaches Due to Human Error in 2019,” *Info Security Magazine,* February 6, 2020, https://www.infosecurity-magazine.com/news/90-data-breaches-human-error/?_gl=1*1whjau8*_gcl_au*OTEwMTA1NTY0LjE2ODkxNzU5MTM.

97 Nadja El Fertasi, “Three Ways to Leverage Emotional Intelligence and Minimize Cyber Risk Through Human Vulnerability,” *Global Cyber Alliance Blog,* November 10, 2020, https://www.globalcyberalliance.org/three-ways-to-leverage-emotional-intelligence-and-minimize-cyber-risk-through-human-vulnerability/.

98 “Is Your Cybersecurity Strategy Emotional Intelligent?” *App Guard Blog,* August 13, 2019, https://www.appguard.us/blog/is-your-

cybersecurity-strategy-emotionally-intelligent/.

[99] Alex Scroxton, "Emotional Intelligence, Empathy Increasingly Valued in CISOs," *Computer Weekly*, 17 February 2021, https://www.computerweekly.com/news/252496468/Emotional-intelligence-empathy-increasingly-valued-in-CISOs.

[100] Sanja Budimir, Johnny R.J. Fontaine, and Etienne B. Roesch, "Emotional Experiences of Cybersecurity Breach Victims," *Cyberpsychology, Behavior and Social Networking* 24 no. 9 (2021): 612–16. Available: https://doi.org/10.1089%2Fcyber.2020.0525.

[101] "Reimagining Cybersecurity: The Crucial Role of Emotional Intelligence," *Thrive with EQ Blog*, March 22, 2023, https://www.thrivewitheq.com/blog/what-is-a-cyber-attack.

[102] Mark Rasch, "Former Uber CISO's Conviction Affirmed by Trial Court," *Security Boulevard Blog*, January 19, 2023, https://securityboulevard.com/2023/01/former-uber-ciso-conviction-affirmed-by-trial-court/.

[103] Ibid.

[104] "SEC Adopts Rules on Cybersecurity Risk Management, Strategy, Governance, and Incident Disclosure by Public Companies," *U.S. SEC Press Release*, July 27, 2023, https://www.sec.gov/news/press-release/2023-139.

[105] See the SEC press release from October 30, 2023, "SEC Charges SolarWinds and Chief Information Security Officer with Fraud, Internal Control Failures." Available here: https://www.sec.gov/news/press-release/2023-227.

[106] Karen Renaud, Verena Zimmerman, Tim Schürmann, and Carlos Böhm, "Exploring Cybersecurity-related Emotions and Finding that they are Challenging to Measure," *Humanities and Social Sciences Communications* 8 (2021): 1–17. Available: https://doi.org/10.1057/s41599-021-00746-5.

[107] Ibid., 9.

[108] Sara Ahmed, *The Cultural Politics of Emotion* (New York: Routledge, 2015), 3. Ahmed is instructive regarding the role of fear in particular: “When a feeling becomes an instrument or a technique it is not that something is created from nothing. But something is being created from something: a wavering impression of nervousness can strengthen its hold when we are given a face *to be nervous about*. To track how feelings cohere as or in bodies, we need to pay attention to the conversion points between good and bad feelings” (227).

[109] Freud describes an “automatic reaction of anxiety” that enters the conscious ego from the trauma experienced by the id at birth. Future “danger-situations” will run under “an automatic influence.” See Sigmund Freud, *Inhibitions, Symptoms and Anxiety,* Alix Strachey (trans.) (London: Hogarth Press, 1949), 114, 138.

[110] Mikkel Borch-Jacobsen, *Remembering Anna O.: A Century of Mystification* (New York: Routledge, 1996).

[111] Todd Dufresne, *Killing Freud: Twentieth Century Culture and the Death of Psychoanalysis* (New York: Continuum, 2003), 23–4.

[112] Ibid., 17.

[113] Ibid.

[114] Borch-Jacobsen, *Remembering Anna O.*, 19–20.

[115] Hochschild, *The Managed Heart*, 36. Emphasis mine.

[116] Ibid.

[117] See https://www.reddit.com/r/antiwork/ or https://www.reddit.com/r/LateStageCapitalism/.

[118] Ibid.

[119] Ibid., 34–5.

[120] Hadnagy and Fincher, *Phishing*, 6.

[121] Hadnagy, *Social Engineering*, 113.

[122] Ibid., 114.

[123] For a definition of a toxic work environment, see Nedra Glover Tawwab, “Work,” in *Set Boundaries, Find Peace: A Guide to Reclaiming Yourself* (New York: Tarcher Perlgee, 2021): 219–32.

[124] Harriet Fraad, “Toiling in the Field of Emotion,” *The Journal of Psychohistory* 35 no. 3 (2008): 279.

[125] Ibid.

[126] Ibid.

[127] Ibid.

[128] See Daniel Gaztambide, “Capitalism Hits Home: An Interview with Harriet Fraad,” *Public Seminar*, July 23, 2020, https://publicseminar.org/essays/capitalism-hits-home-an-interview-with-harriet-fraad/. Fraad describes the exploitative labor practices in the United States as an “adversarial relationship” that manipulates family structures to perpetuate hierarchies across the economy.

[129] See Lindsay C. Gibson, “How to Identify Emotionally Mature People,” in *Adult Children of Emotionally Immature Parents: How to Heal from Distant, Rejecting, or Self-Involved Parents* (Oakland, CA: Harbinger Publications, 2015): 177–96.

4

Conti Leaks

On February 24, 2022, the Russian military invaded neighboring Ukraine, in the largest military attack in Europe since the Second World War.[1] The following day, the Russia-based Conti cybercrime group declared "full support" for the government of President Vladimir Putin.[2] Conti, which has been operational since at least 2019, already had "documented ties with Russian intelligence apparatus."[3] Conti's "warning" message was published on their blog and read as follows:

> As a response to Western warmongering and American threats to use cyber warfare against the citizens of Russian Federation, the Conti Team is officially announcing that we will use our full capacity to deliver retaliatory measures in case the Western warmongers attempt to target critical infrastructure in Russia or Russian-speaking region of the world. We do not ally with any government and we condemn the ongoing war. However, since the West is known to wage its wars primarily by targeting civilians, we will use our resources in order to strike back if the well being and safety of peaceful citizens will be at stake due to American cyber aggression. [sic]

On Sunday, February 27, 2022, an anonymous Ukrainian national living in Kyiv leaked several years' worth of chat logs and sensitive information belonging to the Conti ransomware crew. The following day, more logs and more data were released along with the message, "Glory to Ukraine!"[4] and "fuck russian inviders!"[5] The Twitter account @ContiLeaks shared links to raw data and screenshots of ransomware source code, chat logs, and other internal documents relating to the criminal organization's operations. The "Conti Leaks" affair granted an unprecedented view into the day-to-day operations and "business" culture—as well as the technical documentation related to tactics, techniques, and procedures (TTPs)—within one of the most notoriously ruthless and capable cybercrime organizations in the world. Early analysis of the timestamped chat logs revealed the often mundane, workaday practices of cybercriminals.[6] "Conti doesn't work weekends" was a headline that matched their routine 9 to 5 working hours, Moscow time.[7]

In the days and weeks following the leak, a vivid portrait of a global cybercrime organization emerged, juxtaposed with a less distinct portrait of a single individual, hacking back against Russia-aligned ransomware gangs from the middle of a war zone.[8] In an interview published on March 30, 2022, the still-anonymous hacker spoke to Sean Lyngaas at CNN, using the pseudonym, Danylo. Amid Russian artillery and missiles attacks on the Ukrainian capital, Danylo struggled to locate friends and family. He explained his actions in the interview, stating, "I cannot shoot anything, but I can fight with a keyboard and mouse."[9]

The FBI contacted Danylo shortly after making the initial leaks, requesting him to cease his activities. In global espionage, the immediate benefits of actions like exposing Conti's tooling and internal communications must be weighed against the future advantages of maintaining a foothold in an enemy network. There are also risks in making criminal software tools openly available on the web. The value of such access became evident when the US Department of State offered $10 million rewards for the arrest and conviction of Conti members in May 2022, following the public disclosures.[10] Ironically, the FBI lost its access when the information became publicly available. As Conti reorganized, the group altered its tactics and tools, leading to the creation of numerous malware variants as other threat actors repurposed the leaked code.[11] While Danylo saw an opportunity to fight an invading army, in a matter of weeks, Conti could be seen regrouping under other brand names.[12] The emotional toll of having one's home invaded and loved ones endangered makes Danylo's motivations understandable. However, the specifics of how he initially gained access to Conti remain unclear. He claims to have monitored the group for an extended period, stating, "You need to catch them when they make a mistake. I was in the right place at the right time. I was monitoring them."[13]

As a case study in security rhetoric, the Conti Leaks episode is significant for illustrating national-level backing of cybercrime as a form of warfare. It also demonstrates the asymmetrical relationship between attackers and defenders from both a hostile and a defensive posture. Conti's operational security showcases how the human element in cybersecurity can be both an asset and a liability for attackers. The leaks confirmed exact details about how Conti manages attacks against hundreds of targets for instance. Hostile actors need only succeed once against a target to gain a foothold, while defenders must always succeed across complex institutional contexts. This asymmetry applies equally against large cybercriminal organizations as well as their targets. As a single individual with an IT background, Danylo was able to thoroughly own a criminal organization worth more than $100 million a year in revenue.[14] This asymmetry is a key feature of security rhetoric, a dramatic irony of the twenty-first century.[15]

The Conti Leaks affair also provides insight into the effectiveness of political and historical analysis in the attribution problem of security rhetoric. Security rhetoric is dependent on the interwoven cultural features, including political history. For instance, the political significance of Danylo's pseudonym is a feature of the attribution problem, which may help locate his motivations and his relationship to Conti itself. Daniel of Galicia, or King Danylo Romanovych, was the much embattled first King of Ruthenia, who ruled over an area that included much of present-day Ukraine.[16] He began his career as an exile with his mother, following the death of his father, Roman II Mstyslavich, at the age of four. The year of his exile was 1205, but by 1245 he had reconstituted his father's lands and successfully repelled the combined forces of Polish, Hungarian, and Mongol invaders. By 1246, he was summoned by the Golden Horde at Sarai, where he was forced to pledge fealty to the Mongol khan Batu over a cup of fermented goat milk. Danylo conceded his overlordship and protected his people from further invasion or occupation. Concurrently, he re-established alliances with the Kingdoms of Poland and Hungary and secured Papal recognition to shore up support with neighboring powers. Despite later Mongol invasions, Danylo was able to maintain and grow the territory of his father while providing stable conditions for his people's well-being. So, the question is simple: why would this embattled thirteenth-century leader carry significance a twenty-first century Ukrainian hacker?

In Danylo, we have the figure of an astute tactician willing to use verbal pledges of fealty for strategic gains, even if it means submitting to an enemy in words more than actions. To what degree is the pseudonym a significant indicator of the hacker’s intentions? It is a judgment call, but in coming to this question, we have helped inform an opinion. Drawing analogies to contemporary stories, one could see a parallel in feigning allegiance to invaders for the sake of security. While a double-cross is intriguing, the emotional and political gravity of war suggests that Danylo’s choice may reflect a deeper historical resonance. Given the culturally genocidal nature of Putin’s invasion, resilient figures like Daniel of Galicia are affirming military motivators drawn from history.[17]

In this context, I find it hard to accept Danylo’s claim of being “in the right place at the right time.” He may have been monitoring Conti, but it is much more likely that he was working for Conti as a low-level operator. When the war started, Danylo chose sides and used his existing access to hack back. Brian Krebs cited a statement from Alex Holden, the Ukrainian-born founder of Milwaukee-based cyber intelligence firm Hold Security, who asserted that the leaker was not a former Conti affiliate. In Krebs’s reporting he goes on to quote a statement from Holden, “The person releasing this is a Ukrainian and a patriot. […] He’s seeing that Conti is supporting Russia in its invasion of Ukraine, and this is his way to stop them in his mind at least.”[18] The timeline of events further suggests that he turned against Conti the moment they swore allegiance to Putin’s Russia and used his legitimate access to Conti tooling and resources to defend himself and his nation.

Long before Danylo’s hack and leak, there was an extensive history of efforts to defend against and disrupt Conti. The leak contained years of logs, internal software tooling, as well as functioning malware such as versions of Emotet, TrickBot, and BazarLoader.[19] Conti had already been highly politicized, notably for targeting hospitals during the Covid-19 pandemic. Krebs’s reporting again demonstrated how the gaps in the chat logs “roughly correspond to times when Conti’s IT infrastructure was dismantled and/or infiltrated by security researchers, private companies, law enforcement, and national intelligence agencies.”[20] For example, in late-September 2020, the US National Security Agency (NSA) seized control over the TrickBot botnet, “a malware crime machine that has infected millions of computers and is often used to spread ransomware.”[21] TrickBot uses a cryptographically generated DNS when separated from its Command and Control (C2) infrastructure.[22] After the NSA disabled

TrickBot's recovery mechanism, the C2 servers were unable to reconnect individual machines to the pre-configured DNS. Krebs was able correlate this moment in the leaked chat logs, where one of the top Conti leaders, Hof, complained, "Sorry, but this is fucked up. I don't know how to get them back."[23]

It took Conti several weeks to rebuild its botnet, but by late October 2020, the group's network of infected systems included a purported 428 medical facilities. On October 26, a Conti manager wrote, "Fuck the clinics in the USA this week. There will be panic. 428 hospitals."[24] Krebs also observed that on October 28, the FBI and the US Department of Homeland Security (DHS) issued a warning about an "imminent cybercrime threat to U.S. hospitals and healthcare providers."[25] Similar gaps in the chat logs coincided with other disruptions described by Krebs, including efforts to disrupt another malware strain, Emotet.[26] It is at this moment that the recent trends in ransomware operations can be understood through a criminal business operation, wherein the motivations of ideology and money come together as a "cybercrime-as-service" model.[27]

Cybercrime case studies often adhere to a structure that is conducive to humanist approaches in security rhetoric. This includes historicization and political contextualization, theoretical framing, close reading of facts, prospective analysis, as well as criticism and policy recommendations. Mounting a thesis or argument, with a weight of evidence as support, is how the humanities slowly builds well-supported opinion from suspicion and speculation. These methodological approaches can be organized into an argumentative framework that encompasses both technical analysis and policy responses. A detailed account of the attack and response strategy should be integrated with an impact analysis that considers lessons learned, explores potential future actions, and includes ongoing risk assessment.

The NIST Cybersecurity Framework can serve as an analytical scaffold for structuring humanistic analysis.[28] Within this scenario, attack, response, impact, and discussion framework, it is necessary to account for the human factor in security rhetoric. Vulnerabilities and risks are context-dependent and may evolve over time. By studying the Conti Leaks episode as a discrete event with a finite set of data, we can identify features that might be applicable to future attacks. Given the urgency of a twenty-first-century conflict in Europe, insights gleaned from such an analysis could potentially serve as indicators of future conflicts with China.

Perpetually Responding to Emergencies

Conti's primarily targeted service industry (37 percent) and manufacturing (25 percent).[29] Company size did not significantly influence target selection: 24 percent of targets were small companies with less than 100 employees, 39 percent of targets were medium-sized companies with 100–500 employees, and 37 percent were companies with more than 500 employees.[30] The inclination to target larger companies likely correlates with a broader attack surface and greater financial resources, suggesting both increased opportunity and reward. The majority of these targets were US-based. Notably, healthcare represented just 4 percent of their victims. On November 2, 2020, the US Cybersecurity and Infrastructure Security Agency (CISA) issued a cybersecurity advisory warning of a threat targeting US hospitals that may disrupt service and steal data.[31] In early 2021, they claimed to have attacked 428 hospitals with ransomware. When Danylo made the leak on Twitter, he described part of his motivations in doing so: "more sanctions! they destroy hospitals, and a lot of ppl died! even some of my friends !"[32] The compounding impact of perpetrating a cyberattack against a hospital during a global pandemic did draw significant attention. The intersection of multiple factors is a distinct feature of the Conti Leaks affair that produces distinct effects related to increased risk, vulnerabilities, harms, and uncertainty.

An analysis of the security rhetoric of this situation must consider the intersecting realms of warfare and cybercrime, as the intersecting feature that drew attention back to Conti. Chat platforms like Jabber and Rocket overlap with social media platforms like Twitter and Telegram, wherein the dark web world of ransomware stumbles into public view. These are contextual features of the attack that helped persuade Danylo to burn Conti's tooling in such spectacular fashion. A contemporaneous Check Point report on Conti from the time described Conti as a "normal start-up, sort of."[33] The report included a network graph mapping Conti managers to other roles and Conti operators. User handles like "stern," "mango," "buza," "professor," and "hof" managed a network of over 100 affiliated operators. These operators were accountable to the managers through a system of penalties and incentives. Errors like losing access or absenteeism resulted in fines of $100, while bonuses of “+50% of salary” were granted, in one case, to the operators responsible for “pulling the project in such difficult conditions.”[34] Surprisingly, Conti also maintained three physical office spaces, a departure from the remote nature of most cybercrime operators.

Initial reports on Conti Leaks were simply surprised to see so much information.[35] Subsequent commentary frequently repackaged original reporting from Krebs.[36] Over time, information would emerge from the tech news, including indicators of compromise as well as phishing email templates.[37] The cybersecurity community is committed to learning from events like this. Will Thomas, also known as BushidoToken, is a threat intelligence researcher who is committed to archiving the ephemeral remains of significant events like the Conti Leaks: “I prefer to just do threat research on my own time (which I affectionately call internet dumpster diving because it often involves public sandbox submissions) and collaborate with other analysts.”[38] Thomas’s internet dumpster diving has produced a remarkable collection of “Lessons from the Conti Leaks.”[39] Thomas highlights the sophistication of the reconnaissance efforts, which is his area of specialization. These investigations revealed that Conti ran an “OSINT Team” that used commercially available tools and even employed in-person Human Intelligence (HUMIT) gathering techniques. Conti would pose as marketing or salespeople, gathering details and information about managers and executives to better understand how a company operates for later exploitation. Thomas notes that “Conti ransomware attacks often begin via a phishing email” and that email marketing tools are used to manage these campaigns. According to Thomas, “what sets Conti apart from the rest of their peers in the cybercrime ecosystem is that members of this ransomware group are innovators and quick to leverage newly disclosed techniques.”[40] Working at the cutting edge of technology, also means falling for the hype of new technology, like cryptocurrency. With a cybercrime empire as large as Conti, making $100 million per year with more than $3.6 million in salaries to operators, they had a lot of digital currency to manage.

In the leaked materials, Thomas also notes that Conti was in the early stages of developing several cryptocurrency platforms.[41] Laundering that amount of cryptocurrency would require exactly this kind of innovation, with the added benefit of hiding a few ill-conceived cryptocurrency platforms within the sea of breathless crypto-bros waiting for the next rug pull.[42] But, unlike so many crypto-dreams, Conti isn't going anywhere soon. Remarkably, "Conti has seemingly recovered from the leaks and," as Thomas puts it, "might be at the 'too big to fail' stage of operations."[43] Assuming that Conti has deep TTP and new innovative capabilities, "the invisible part of the iceberg is still yet to unravel."[44] While ego and arrogance fueled Conti's ascent, those same traits led them to publicly support Putin's government and the war in Ukraine. Being innovative as ever, Conti appears to have distributed its operations across several concurrent operations. These operations are branded with names like HelloKitty, AvosLocker, Hive, BlackCat, BlackByte, Gold Ulrick, Diavol, Karakurt, Royal, and more recently Akira.[45] The notoriety and branding of Conti appear to be, in retrospect, one of its greatest weaknesses. So, while the Conti Leaks affair is a significant look at cybercriminal tradecraft, the comparison to the "Panama Papers of Ransomware" is perfectly appropriate because little will change as a result of the disclosures.[46]

One of the more prominent Conti attacks targeted the Irish Health Service Executive (HSE) on May 14, 2021.[47] Healthcare services across the country were severely disrupted, with 4,000 locations, fifty-four acute care hospitals, 70,000 devices, as well as 130,000 staff. In response, the HSE was forced to shut down all IT systems and disconnect the National Healthcare Network from the internet in an attempt to contain and assess the impact of the attack. Despite receiving a decryption key from Conti, it took over four months for the HSE to fully recover. The details of the event are described in a Post Incident Review (PIR) report that should be regarded as a canonical example of the kind of transparency and knowledge sharing needed in the wake of an attack.[48] It is this kind of reporting that will avoid the "never again, whoops" model of recovery and relapse.

The attack originated on March 18, 2021, through malware delivered via a phishing email containing a malicious Microsoft Excel file.[49] It took the Conti operators eight weeks to "detonate" the ransomware package on 14 May. During this three-month interval, the Conti operators compromised accounts with higher, administrative privileges. They compromised a "significant number of servers, exfiltrating data and moving laterally to statutory and voluntary hospitals."[50] Although HSE IT detected signs of Conti's activities between the initial compromise and the ransomware activation, these alerts were not acted upon, and no efforts were made to identify or remove the threat actors.

A significant shortfall in the HSE's cybersecurity measures was the lack of specialized staff, with only fifteen full-time IT professionals dedicated to security. Subsequent recommendations also included the hiring of a Chief Information Security Officer (CISO).[51] The PIR report describes a dire situation that was only recoverable because the Conti operators, fearing nation-state retribution for attacking a hospital during the Covid-19 pandemic. "Without the decryption key," explains the authors of the PIR, "it is unknown whether systems could have been recovered fully or how long it would have taken to recover systems from backups, but it is highly likely that the recovery timeframe would have been considerably longer."[52] The awareness of the risks and vulnerabilities associated with cyberattacks, particularly in vulnerable infrastructure like healthcare, invites a *counterfactual* reflection on what could have happened in this emergency.

A PIR seeks to establish the facts, but in identifying ways to improve and learn from these events requires anticipating events that did not happen but could have. While this speculative mode is *counter* to the *facts*, they establish a future-oriented planning process that is capable of mitigating the next attack. Recognizing that the incident's impact could have been "significantly greater, with far more sever clinical impact" is both sobering and invites a speculative approach: What if attackers targeted "specific devices" keeping people alive? What if Conti destroyed "data at scale" and left the nation without access to medical records? What if the malware had "auto-propagation and persistence capabilities" and spread to other critical, lifesaving infrastructure? What if the "Covid-19 vaccination system" was affected? The unspoken answer to these questions is the potential loss of life, which remains the most critical metric when assessing the impact of such attacks.[53]

These counterfactual questions focus on the negative aspects, acknowledging that the attackers "used relatively well-known techniques and software to execute their attack." Positive counterfactual statements might have included questions related to the human factors of the successful attack, including that "alerts were generated by antivirus software on key systems in the days leading up to the attack, which were passed to the cybersecurity team."[54] These alerts were received by an overstretched IT security team without the adequate expertise to handle an attack in the midst of the Covid-19 crisis: "As a result, opportunities to prevent the crisis were missed."[55] While the report stops short of stating that additional IT support and a robust mitigation process could have prevented the attack, there is an underlying tone of regret in criticizing the victim. Such criticism is often seen as impolite and tinged with the shame of failure. Security rhetoric aims to navigate this delicate issue by fostering a sense of care for attack victims and a concern for preventing future harm.

At the time, the world was grappling with the Covid-19 pandemic, implementing social distancing measures and striving to support overwhelmed healthcare systems. Attacking hospitals under these circumstances would be ill-advised and likely provoke strong public and governmental reactions. Conti likely provided the HSE with the decryption keys to avoid further attention of Five Eyes security establishment, which disrupted their operations previously. So, why did Conti choose to seek escalated privileges and later detonate the ransomware payload? Did Conti have a failure in their chain of command or were they unaware of what system they were in? It is improbable that the Conti operators did not know the environment they were seeking higher levels of access, which suggests a disconnection between Conti's managers and their lower-level operators. The chat logs from the days leading up to the detonation of the malware on May 14 suggest that there was some disagreement or failures to respond from Conti's managers to the operator level. There are no communications in the leaked Jabber logs on either March 18, the date of the initial compromise, or May 14, the date of the malware detonation. On May 16, the Conti operator "Salamandra" messaged "Stern," a mid-level manager responsible for working with the lower-level operators: "I'm not in Russia and we have a flood. there is a suspicion that tomorrow there will be no light and the Internet. I warn you about this in advance. I hope that everything will be fine, but it may turn out that I will be without communication for 72 hours. please do not punish." The difficulties of managing a global cybercriminal organization have a range of complications it seems, since working conditions are so varied. With the grave consequences of a ransomware attack on a national hospital network, Conti operators were motivated by punishments of just $100 dollars.

Echoing the *zeitgeist*, the employees at Conti and the HSE were similarly overworked. Among the key recommendations, the HSE PIR, in the "human factors and cultural contributions" section, related most keenly to the initial point of compromise, the missed opportunities to prevent the crisis, and the overall emotional toll on medical staff facing the Covid-19 crisis: "Chronic stress without recovery, depletes energy reserves, leads to burnout and ultimately compromises the crisis response capability."[56] The PIR report describes "a prevalent theme in interviews is that the staff are 'perpetually responding to emergencies.'"[57] The "culture of preparedness" in medical settings is related to inclement weather, natural disaster, or terrorist attack that would suddenly, with little warning, trigger a sharp increase in medical resource needs. These major emergencies must now include cyberattack, both inside and outside the medical system.

The public seldom hears about the potential harm of multiple "black swan" events occurring simultaneously. The backdrop of a chronically stressed and toxic digital environment adds additional noise to the stress faced by employees. The "cultural contributions" described in the PIR are complementary of HSE staff for their professionalism in the face of chronic emergencies. HSE staff are heroes and deserve accolades for managing these crises. But should they have to battle both a pandemic and a cyberattack at the same time? A cybersecurity emergency response cannot be, as the report describes, "fuelled by staff members' 'can-do attitude.'"[58] A sense of confidence is certainly a feature of security rhetoric, but it should not be the only response mechanism. While the report recommends instilling "a culture of preparedness," the broader plans related to "comprehensive training" must build a deeply embedded security culture in all operations and mitigate the employee burnout that contributes to being deceived by a phishing email in the first place.

быстрый старт хакера.txt

In another striking similarity, employee training is a constant feature for both attacker and victim. The staff of the HSE in Ireland lamented how, in the midst of multiple crises, a "significant amount of time was spent onboarding and integrating third parties, particularly education them on the intricacies of the health sector."[59] In the midst of a double crisis, existing

HSE staff had to train consultants. These must have been dark days indeed. In contrast, Conti had a text file. “быстрый старт хакера.txt” is Russian for “hacker’s quickstart guide.txt.”[60] This document was a part of the Conti training manuals for new employees.[61] Conti had to bring on many new employees, particularly after dismissing lower-level workers to maintain their own, albeit illicit, operational security. Many of them were paid between $1000 and $2000 per month for their efforts. Despite the criminal context, workplace grievances were common among low-level workers, who complained of long hours, delayed payments, and poor communication from higher-ups. But Conti had training manuals to help. The *hacker’s quickstart guide* reports on a range of techniques, including IoT devices, RDP, and Active Directory representing the preferred initial points of compromise.

Conti’s training materials not only expect recruits to read security reports about their activities but also instill a sense of skepticism in their security culture. The “quickstart guide” warns that security news and analysts may distort reports about Conti’s activities to mislead them: “Sooner or later, any hacker reads analyses of his art in information security articles. And he is surprised to find that important information is missing, and insignificant information is protruding.” They describe security analysts reporting on Conti behavior as exaggerating minor points and ignoring “important details”: “they know everything, but pretend not to—to use it against you.” The authors of the guide use a Russian proverb to describe their adversaries: “не боги горшки обжигают—аналитики действительно могут упустить важные детали,” which translates to “It’s not the gods who burn the pots—analysts can really miss important details.” These training documents aim to highlight the role of disinformation while also acknowledging that their adversaries, primarily the US cybersecurity industry, are not infallible. The contradictions within Conti’s training materials reflect contradictions elsewhere.

With a heightened focus on disinformation in their training materials, Conti is training for both offensive and defensive social engineering tactics in practice. The “quickstart guide” places an emphasis on technical vulnerabilities like those found in Active Directory and RDP, which were exploited heavily during the Covid-19 pandemic and the shift to work from home practices.[62] In an analysis of the “Network Landscape” for potential attack, the guide reveals the mindset of Conti’s leadership.[63] The authors of the “quickstart guide” claim that “any (almost) modern network can be hacked” because of a few key features, which including “network redundancy,” the priority of “convenience over safety,” as well as the “human factor” or “человеческим фактором.”

The authors of the guide note that network redundancy and the prioritization of convenience make their targets vulnerable, poetically stating, “the prison is safe, but it is very difficult to do something in it.” In other words, defenders often sacrifice security for convenience. The “quickstart guide” also emphasizes that the commercial sector’s focus on profits leaves no room for even the slightest slowdown in operations. Thus, while Conti aims for speed, the group’s “quickstart guide” also recognizes the importance of thorough and patient research:

> Social engineering requires knowledge about personalities. Everything is important: phone numbers, place of residence, dog’s name, hometown, favorite color, favorite band, hobby. Of particular importance: your candidate’s personal network of contacts, especially business contacts. The structure of organizations reflects the structure of society. Moving from one person to another through a network of contacts, you can change the entry point within one network, or open new networks. To collect information, they are used as a means of reconnaissance OSINT, and information about contacts found in previously open networks (Outlook address books, correspondence, etc.)

Conti operated much like a conventional corporation, complete with HR and customer service departments. They elevated ransomware to a service by streamlining the payment and recovery process for their targets. In doing so, they built a brand based on credibility and fairness. Conti represented a credible threat capable of doing real harm, while also dealing fairly with victims in restoring their systems. They had customer service representatives that helped victims manage Bitcoin wallets and use decryption keys to restore their encrypted systems. Social engineering blends well with marketing, wherein “KYC” stands for know your customer. OSINT appears to be the ultimate form of customer service, knowing perhaps even how much cybersecurity insurance coverage your victim can access! Knowing your customer in a ransomware attack requires personal attention to apply leverage to executive boards and other leaders as well as to serve them well after payment. Ultimately, paying the ransom becomes a viable option if it allows for quicker and more efficient system restoration than relying on backups. Like any business decision, the choice to pay off a criminal organization must offer a return on investment.

The greatest paradox for a criminal organization like Conti lies in the need to foster trust among recruits trained in deception. Low-level operators, responsible for initial system compromises, must report their activities to senior members. There is real value in obtaining initial compromise and that access can be readily sold on forums that cybercriminals frequent.[64] Conti's management faces the challenge of motivating these operators, who are paid modestly, work long hours on monotonous tasks, and must adhere to strict documentation protocols. The organization expects these criminals to follow business processes from home, all while under the constant threat of financial penalties and termination for suspected security breaches. It appears that a toxic work culture and grueling hours are universal, even for criminal operations. Crime pays, but the hours are terrible.

In their "Developer Guide" ("наставление разработчику"), Conti strikes a delicate balance between conventional professional developer documentation and the subversive ethos of hacker culture. It is a tricky balance to strike in training documentation. The "Developer Guide" blends the language of corporate tech with anarchist rhetoric, merging strict procedural guidelines with illicit hacking activities. The procedural requirements for rigorous technical processes are blended with fraudulent hacking of valuable institutions. The lines between business and crime get a little blurry. For instance, they describe their "management principles" as borrowing "flexibility from Agile," while also retaining "the lack of hierarchy in anarchist self-organization."[65] They describe a horizontal quasi-egalitarian "network organization of working groups" with "direct communication" between any group or individual operator. Furthermore, they adopt practices from the free software movement, such as version control and error tracking, aligning them with Agile development principles.[66]

These standard software development practices in the Conti training manuals often return to their more florid, anarchist platitudes: "Remember: LIFE is a self-organization of chaotic matter. Chaos is characterized by self-organization on energy flows. We are this CHAOS. LIFE itself will correct us when we start to stagnate." The New Age references to "energy flows" is bit surprising but serves a broader cultural cohesion within Conti. This quasi-philosophy appears to be a custom brand of atheistic spiritualism animated by the contradictory facets of organized chaos. Conti's brand of hacker anarchism seems to blend trust and deceit in equal measure. While it is unclear how much importance Conti operators attach to this eclectic philosophy, it may have been sufficiently compelling to help maintain technical processes and manage reporting behavior among new recruits.

The training materials include detailed descriptions of criminal TTPs, all while displaying a listless disdain for their targets. These philosophical musings on life and chaos segue into a section on “coding principles.” The focus here seems to be on naming conventions for variables, functions, and classes: “If you give a good and appropriate name to a variable or function or class, then everything else will almost automatically be fine. Otherwise, almost automatically everything else will be bad.” The Conti managers fear poor naming conventions because these features are difficult to refactor in deployment on a victim’s system. Variable names are malware fingerprints that Conti managers would like to avoid: “The wrong choice will grow in the code forever!” they warn. They give plenty of good advice on commenting code and limiting dependencies on extraneous software packages. Sprinkled within the pragmatic stylistic features of good malicious code, they blend their New Age anarchist spiritualism with corporate reporting and professional training. It seems the aim is to cater to two distinct needs in their new recruits: those seeking meaning and belonging, and those looking for work and education. A generous reading could claim that the documents manage to strike an uneasy balance between criminality and community.

This delicate balance of trust is upheld through a mix of financial incentives and fostering a culture of compliance. The contradiction lies in maintaining a “cool” image while also dedicating long hours to meticulous code documentation and commenting.[67] Some of these tools come packaged with Kali Linux, which the Conti training manuals recommend. It includes very well-known packages like Metasploit, Burp Suite, Core Impact, Powershell Empire, and Pupy. Conti also uses standard commercial penetration testing tools, for which Cobalt Strike has a special place. Cobalt Strike claims to work with vetted partners in issuing their software suite, which is designed to deploy and automate all but the most current and exotic exploits. Russian operators like Conti have a long track record of setting up shell companies to pay for quasi-legitimate licenses or simply pay a security professional working in IT for a shared license.

In a file called “Anonymity for Paranoids” (“Анонимность для параноиков”), the Conti managers discuss the paradox of using easy-to-spot tools and the risks they entail. They recognize that efforts to obfuscate user profiles and activities can ironically make them more conspicuous: “The task is not to hide (it still won’t work), but to merge with the crowd. So disabling webrtc, Javascript, Flash, etc. just draw more attention to yourself. It is necessary NOT to DISABLE, but to REPLACE what allows you to be detected.” Conti managers recommend merging with regular traffic to disguise malicious behavior. Since Kali Linux is not commonly used as a daily operating system, any digital footprints left behind could serve as an indicator of compromise. Again, from “Anonymity for Paranoids,” the author describes this problem by stating,

> About Kali and other operating systems for hackers. Here is a group of people (Hackers) that needs to be traced. Technically, this problem is difficult to solve. It’s easier to play on human weakness (laziness) and gather everyone together by providing a properly advertised, convenient, ready-made and popular solution. I think the idea is clear. I advise you to use Debian or build something of your own.

Kali Linux can function like a honeypot for hackers because it is so much easier to use, but it leaves traces indicative of the OS.[68] Kali plays on an attackers’ own human weakness and laziness by providing a ready-made solution that is easy to trace for US governmental operators with access to the telemetry provided by Google, Microsoft, and others.

In a document titled “Safety Engineering,” Conti managers offer several recommendations to their low-level operators who opt to build their own Debian-based hacking platforms. These guidelines emphasize the use of strong passwords and the importance of regular backups stored on separate media. The managers also advocate for comprehensive encryption, covering not just the working partition but also backups and payloads: “Copy them to a separate medium, store it separately. It is advisable to store the media with backups in another room. Backup media must be encrypted.” Interestingly, these managers expect a high level of diligence and consistency from their employees, who have chosen a path of illegal activity for income.

The contradictions in Conti’s operations partly stem from the technological context in which they operate. They work with Microsoft’s Windows OS, as both their own tooling as well as their target. Since Windows telemetry is automatically reported back to Microsoft and the US intelligence community, poorly configured systems can quickly become liabilities. In the training manuals, Conti managers both admire and fear US technology giants. They describe Google Analytics and Google Chrome as US government spyware, but they also routinely use Google's Virus Total to check any file that is not a plain text: "DO NOT OPEN other people's.doc,.docx,.xls,.pdf,,rtf and so on files on your personal machine," they emphatically instruct their new operators in the "Safety Engineering" document: "any files that interpreted by a Turing-complete finite automaton (which almost all programs are). These formats are used to inject viruses. If you have to—a virtual machine and virustotal to help." Virus Total was purchased from the Spanish security company Hispasec Sistemas in 2018 by Google and now operates as the subsidiary, Chronicle Security. Virus Total is a quick and easy way to check for known exploits embedded in these common file types, particularly macros in.docx and.xls files. As it happens, these short Visual Basic scripts can be executed in common Microsoft files, which was the most common initial payload in Conti phishing attempts.

Even when scripting malware, Conti adheres to a style guide outlined in their "Developer Guide" document. They advise their malware authors that "the best code is the one that doesn't exist (don't write too much)." They suggest that good code should be less than a kilobyte to avoid detection from anti-virus (AV) programs, particularly Windows Defender. To stay ahead of Windows updates, Conti refreshes their code as frequently as every four hours, recognizing that the risk of losing access to compromised systems is a constant threat. So, while much of their guidance regarding passwords, backups, documentation, and commenting all represent good developer practices, they also have cause to encourage their developers to use non-standard file names like "1.rar" or "c.rar" to avoid AV detection, which is described in the "Code Requirements" document ("оформление кода и сборок").

When it comes to initial access techniques, they are less particular. In the whimsically titled "The spirit of the old school" ("дух старой школы"), the Conti managers remind their operators, "Any hacks leading to the goal are used. The end always justifies the means." For instance, "The goto statement is allowed and widely used" in the "Code Requirements." The "goto" statement in C has few restrictions regarding transferring control to the location of a specified statement label. Using goto statements are quick and easy, but they also break the scope of execution. A hack using "goto" would be considered a basic hack: “There are no forbidden tricks, there are only ineffective ones. Good code is one that solves a problem.” While carefully crafted 0k exploits might be recognized as a goal, they also understand the only good hack is a working hack: “Stylistically beautiful, but not problem-solving code is not needed. We need a code that solves the problem, because the solved problem is money, yours and mine.” This rare acknowledgment of financial motives in Conti’s leaked documentation quickly pivots back to the topic of social engineering.

They coach their operators on how to conduct a social engineering attack in the “qickstart guide” cited above, but there is a real risk to social engineering malicious actors when organized in such large groups like Conti. The Conti managers remind their operators to manage their online personas: “Use different personalities for different activities.” They explain in the “quickstart guide”: “A unique nickname on the Internet is very bad. Now there are automatic commercial scanners that pierce a person by nickname or photo. Use different nicknames in social networks, at work, in other activities.” They suggest limiting use of social networks altogether, but if it is needed to use pseudonyms and false photos. Conti operators must “learn to invent names and biographies” as a regular course of their work. There is a creative, almost theatrical, quality to being a cybercriminal for Conti.

The contradictions and duplicity present with these employees affect every aspect of their lives. There is a deep sense of countervailing perspectives in their work, between visions of Russia and the West. There is a sense of using, admiring, and emulating “Western,” US-dominated technology and corporate behavior that is inversely mirrored by exploiting, rejecting, and enacting anarchic appeals to justify chaotic criminal behaviors. The best hack is no hack at all. In this way, “the ideal backdoor is legitimate access.” Gaining a legitimate password with high-level access captures this mirrored perspective: “One password for everything, and the presence of a user on many nodes is the main help to a hacker,” they explain in the “quickstart guide.”

This inverted, mirrored perspective of the sociotechnical and cultural contexts of threat actors like Conti allows us to reimagine the potential risks implicit in security services. The very popular “Have I Been Pwned?” tool cuts both ways. Conti includes this tool in their OSINT resources because it can be useful for attackers. While the site claims to “check if your email or phone is in a data breach," it can also be used by attackers to determine if a target's email or phone is already available publicly in an existing data breach. If an individual is a high value target with privileged access, searching their public email on Have I Been Pwned? may mean other valuable information may have been already disclosed. Have I Been Pwned could get you owned by pointing an attacker to the latest dark web data dump containing your passwords or other personally identifying information (PII) needed to phish or harass a target.

Not all attacks are precipitated by vast OSINT efforts to gather PII. Information collection and analysis is still a time-consuming process, though there are lots of platforms that make accessing PII much easier. In casting the widest net possible in a phishing campaign, general and vague communication ensures that the many people will identify with the message and believe it to be legitimate. For example, the Conti Leak contained many phishing email templates, which operators shared and discussed the relative merits. They often shared screen captures of test emails with managers to test the legibility of phishing campaign messages. Here is an example template shared by "Lemur" on October 15, 2021, formatted in the original JSON:

```
{"ts": "2021-10-15T13:25:39.156774",

"from": "lemur@q3mcco35auwcstmt.onion",

"to": "terry@q3mcco35auwcstmt.onion",

"body": "{Greetings|Hello|Good day to you|Good arternoon}{!|,|}

{We are|We're} {writing|messaging|mailing} {to you|you} {regarding|concerning} your
{transaction|payment|transfer|money transfer} TRANSFER NUMBER for AMOUNT for your
order is {processed|completed|approved} and {received|collected|accepted}.

{We have|We've} {submitted|sent|scheduled} your {order|purchase|purchase order|online order}
for the {delivery|shipping and delivery|transfer}: {it will|this will} {take|require}
{about|approximately} {3-5|4-5|five} days. Please {check your|look at your|inspect your|check
the} {information and|details and} {payslip|bank check|receipt} in the {attached file|file
attached}.

{Thank you for|Thanks for} {your business|your order|your interest}.

THEMES:

{Invoice|Given invoice|Bill} INVOICE NUMBER for {order|purchase|purchase order|online
order} ORDER NUMBER is {approved|given approval|affirmed|covered|paid}

ATTACH:
```

> ord_details
>
> purch_info
>
> invoice_details
>
> transact_info
>
> ord_documentation
>
> order_summary”}[69]

There are a few interesting features to discuss in this example, including just how boring it is. The phishing template is an example of how ransomware has weaponized banal customer service experiences by disappearing into a steady stream of soulless customer service. These false corporate expressions of kindness and care are just as empty as the authentic customer service email they emulate. The phishing email template is written with “or” expressions in curly braces, which means this template can produce a few million different emails to help fool spam filters. It also preys on the feeling of surprise in receiving an unexpected invoice and feeling fear for being on the hook for something the target never ordered. There is a deflection of suspicion because the first inclination is to assume the fraud has already happened. It also deflects attention from phishing to ecommerce fraud. With widening global income inequality and inflationary pressures, chances are people will be keen to protect what money they have.[70]

The user is misdirected to the attached file because of the feigned threat to their finances. User receives a “nudge” to open a file and trigger a malware package. Targets have the freedom to choose, but they are made to feel as though there is no choice. The weaponization of consumer culture fits well with the cybercrime-as-service model taken up by Conti. Richard H. Thaler and Cass R. Sunstein’s *Nudge* (2008) describes a “libertarian paternalism” that defines so many of our economic decisions.[71] It should be noted that the authors couch this term in a lot of conditions because of its many terrible associations: “Libertarian paternalism is a relatively weak, soft, and nonintrusive type of paternalism because choices are not blocked, fenced off, or significantly burdened.”[72] Citizens are free to choose in capitalist economies, but they must submit to paternalistic influence of corporations and marketers. Conti operators are all too aware of this free-market compromise made in so many liberal democracies. Reading *Nudge* through the lens of ransomware operators validates the relationship between marketing and social engineering, since similarly “small and apparently insignificant details can have major impacts on people’s behavior.”[73]

The argumentative nudges in these messages are carried in the User Interface (UI) design as well. These manipulative design patterns, which are sometimes regarded as “dark patterns.”[74] Jon Yablonski, in the *Laws of UX*, describes these psychological principles with computing metaphors. Good user experience often makes use of psychological principles, intentionally or not, like Hick’s Law, which asserts that “increasing the number of choices available logarithmically increases decision time.”[75] Speed kills when it comes to deception. So, despite the sheer number of different emails Lemur’s template can produce, the amount of information and number of choices is very limited. The time it takes to decide, even an erroneously, increases with the number and complexity of choices. The increased reaction time is related to the cognitive load placed on the user.[76] This design principle can be co-opted for malicious ends. The UX of phishing emails is more akin to the style of good malware: it’s almost like it isn’t even there. Good malware needs to be small, less than a kilobyte maybe, to avoid AV detection. The cognitive load of a malicious message needs to be very light to avoid thoughtful reflection. The cognitive dissonance of being falsely told you owe money, from a trusted and safe source, elicits the risky choice.

These small details are described in UX as a “choice architecture” or “experience architecture.” Designers are all too aware that the way that UIs are designed can have significant impacts on behavior. In building spam emails, the details of a UI can have significant impacts on the choices made. UX designers are thinking more about the role of rhetoric and how visceral emotions are evoked in digital interfaces. UX researchers are subtly sensitive to how people experience a UI logically and emotionally. For instance, the affective quality of Apple’s design language has a subtly nuanced minimalism, which blends hardware and software seamlessly. UX researchers study the components of Apple design in remarkably effective ways: “The focus of this experience was not just completeness or correctness; rather, it was on delight, aesthetics and visual appeal, and comfort. This is the rhetorical appeal of *pathos,* persuading by appealing to users’ emotions, making them the major focus of product use (or at least making them appear to be).”[77] The implicit message of the UI is often felt before it is understood. For Apple, completeness gives a sense of comfort that focuses directly on the user. User-centered design implies that UIs are made for us, the user. Everything is designed to suit us, built to accommodate our every need. We can trust it because it is made for us.

User-centered design has been the cornerstone of human-computer interaction (HCI) processes for decades. The result is a sense that UIs are designed to serve us, rather than trick or fool us. UIs are so good that there is little resistance and little cause to pause and reflect. Thaler and Sunstein emphasize many of the tenants of UI design, including the need to layer complexity, provide timely feedback, give meaningful incentives, offer sensible defaults, and expect user error.[78] Error-tolerant designs tend to give users a moment to recover, but the counterfactual reading of UI best practices would suggest that effective and intuitive UIs can lead to poor decisions. The alluring familiarity of a corporate logo—one that is associated with a lot of enjoyment, pleasure, satisfaction, and status—may *nudge* a user into a false sense of certainty.

Typography is another key feature of good, intuitive UIs. The fonts or typography chosen for a UI is often the most overlooked component of a design, but typography conveys a tremendous amount of affective information. Fonts are felt. Robert Bringhurst, in *The Elements of Typographic Style*, begins his classic study of typography with some “principles” of the “grand design” of letter forms and layout: “In a world rife with unsolicited messages, typography must often draw attention to itself before it will be read. Yet, in order to be read, it must relinquish the attention it has drawn.”[79] Typography, it seems, is not to be trusted so readily. Typography works on us in a very intuitive way that elides its true intention:

> The same alphabets and page designs can be used for a biography of Mohandas Gandhi and for a manual on the use and deployment of biological weapons. Writing can be used both for love letters and for hate mail, and love letters themselves can be used for manipulation and extortion as well as to bring delight to body and soul.[80]

Bringhurst’s writing is more beautiful, but it describes the function of such crude communications as phishing emails. Phishing emails would appear more conspicuous without capable forging of corporate or institutional design language. Included in the Conti Leaks are many screen-captures of various phishing emails for large companies.

The simple, clean sans serif of Apple, paired with a pleasing grey body text and the punch of blue to color the button follows the design language of a legitimate email. The tight left-aligned text with generous negative space feels familiar. The call-to-action text on the button is lacking specificity and the final prompt, “Provide all the related information asap,” has a tone of desperation that is unbecoming of Apple corporate communications. Similarly, the look and more importantly the feel of the Amazon gift card offer are perfectly designed, right down to the color and center-aligned text. The need to download an application to enter a prize draw is definitely suspicious, but it appears that the unscrupulous behavior is on Amazon’s part. Again, the idea that Amazon has a “Lottery App” that describes the draw as a “raffle” is, again, rather suspicious.

Parasitic on the Target

Conti was only a proxy target for Danylo, with no direct military impact. He had the ability and opportunity to act, but Conti was only ever an indirect target for the Russian forces on the ground. The cyberattacks that were coordinated with the invasion had nothing to do with Conti directly. Even after the leaks, there is little direct evidence that Conti worked for Russian intelligence.[81] It is also clear that “the global cybersecurity community,” comprised of Western aligned corporations and national defense organizations, has failed to impose sufficient costs on attackers like Conti. Patrick Gray of the Risky.biz podcast would call it “releasing the hounds.”[82] Citizens may not wait for governments to act, especially if every security agency is playing the long game. Institutionalization of enemy assets means that US-based agencies are incentivized to maintain access to enemy systems rather than destroy them. Disruption campaigns are psychological rather than sweeping, like a hack and leak operation.

Disruption tactics must be balanced with the military capabilities of the host country as well. The Russian invasion is a feature of *peaking-power syndrome:* the tendency for rising states to become more aggressive as they become more fearful of impending decline.[83] When nations come to terms with the limits of growth and greatness espoused in so much patriotic rhetoric, political leaders have a choice to develop inwardly or lash out, aggressively. Putin’s brand of conservative authoritarianism has sought to extend a very narrow vision of Russian history that is both militarily expansionist and historically revisionist. Internet censorship within Russia has been increasing in the lead-up to the invasion, with a view to controlling the political narrative from within and outside national boundaries.[84]

In 2021, NATO defined its responses to cyberattack in an ad-hoc basis: "Allies also recognised that the impact of significant malicious cumulative cyber activities might in certain circumstances be considered an armed attack that could lead the North Atlantic Council to invoke Article 5 of the North Atlantic Treaty, on a case-by-case basis."[85] The next World War may be declared on a case-by-case basis, as a terrible judgment call. How "significant"? What "cumulative" limit? Which circumstances?[86] There appear to be thresholds to cross, but the lines are blurred. The attribution of malicious intent must balance the motivations and overlapping concerns of cybercriminals and hacktivists, corporations and nations, and defense and intelligence.

In September 2022, Google's Threat Analysis Group (TAG) observed "financially motivated threat actors targeting Ukraine."[87] Repurposing Conti tooling under the larger banner of Wizard Spider, the overarching Russian cybercrime organization. Interestingly, this retaliatory attack included phishing emails impersonating the National Cyber Police of Ukraine. This report opens with mention of the war in Ukraine and concludes that these "activities are representative of blurring lines between financially motivated and government backed groups in Eastern Europe, illustrating a trend of threat actors changing their targeting to align with regional geopolitical interests."[88] Dan Goodin highlighted how TAG and IBM Security X-Force noted "an unprecedented shift as the group had not previously targeted Ukraine."[89] The lines between financially motivated criminals and military conscripts blur, while being indicative of the kind of old style Soviet thinking that drove the invasion from the outset.

In the early days of the war, Ukraine's supporters formed a volunteer "IT Army" to hack and harass Russian government sites and conduct OSINT collection.[90] Ukraine's own Ministry of Digital Transformation described the ensuing cyberattacks as the "First World Cyber War."[91] President Volodymyr Zelensky even credited the IT Army with helping defend digital infrastructure so Ukrainian citizens could continue to access services and receive pay check, helping keep basic features of life.[92] On the one hand, cybercriminal proxies, through their resources and tooling, are conscripted into service of the state and they even pay their own salaries from the proceeds of their crimes. On the other hand, a volunteer militia of IT professionals moonlight as cybersecurity soldiers and live alongside events online. The first cyberwar was fought by two volunteer armies. Cyberwar in Ukraine was also remarkable because it was the first use of offensive and defensive positions in a kinetic war on the ground. The attacks were technically interesting, but the strategic, military effects were limited.

In a white paper published by the Government of Canada's Centre for Cyber Security, their analysts determined that "the scope and severity of the cyber operations related to the Russian invasion of Ukraine has almost certainly been more sophisticated and widespread than has been reported in open sources."[93] The Canadian report includes a list of notable Russian and Russia-linked cyber activities that are directly coordinated with kinetic attacks.[94] Traditional military attacks in Odessa, Sumy, Vinnytsia, Dnipro, and Kyiv were all aligned with cyberoperations between February 14 and May 16, 2022. Most ambitiously, Russian attackers sought to disrupt communications in what would become known as the "Viasat hack."[95] On February 24, in an apparent attempt to disrupt military communications in Ukraine, the attack disabled thousands of broadband modems. There were many Distributed Denial of Service (DDoS) attacks that disrupted government, military, and banking sites across the country.[96] Named "HermeticWiper" by ESET analysts, Ukrainian digital infrastructure also saw attacks using wiper malware that sought to destroy systems by using a false ransomware attack as a distraction.[97] There was even an attempt by Sandworm to disrupt the Ukrainian power grid again, in an apparent attempt to repeat the startling 2016 attack.[98] Ukrainian defenses have been remarkably resilient during this period. In retrospect, the largely successful cyber defense of Ukraine before and during the invasion was due to US efforts to bolster defenses and Ukraine's existing experience in repelling Russian cyberattacks.[99] Tanks amassing on the border was interpreted as an opportunity to harden digital infrastructure and precipitated the first wave of attacks.

Cyberwar is a reality in the twenty-first century. There are numerous features of this period that may be a sign of things to come. It was notable, though not unprecedented, to see the Pentagon's Ukraine War plans leaked on a Minecraft Discord server before spreading globally on Telegram and Twitter.[100] In this case, radicalized American military personnel, with privileged access, served enemy forces by being indoctrinated by Russian disinformation. Existing technologies are weaponized as tools of spy craft and warfighting. The previous evolution of military technology was much the same. The first uses of the airplane in war were very effective for surveillance and supporting forward forces in their first appearance in conflict. Like biplane pilots dropping grenades by hand in the First World War, the effects were largely psychological. The sense of fear and confusion was matched by a realization that this tactic was just the beginning of something much larger. For instance, this was the first war in which drones became more than surveillance tools but were also used to actively drop weapons on ramshackle Russian armor.

The sophistication of current weapons emerges from this fundamental technique. From tactical nuclear bombs to laser guided munitions, these systems evolve in much the same way. The relative success of these attacks, as a measure of achieving strategic aims, must be balanced against their novelty.[101] A concerning feature of cyberwar is that noncombatants can be followed and harassed beyond the battlefield. Phishing attacks targeted Ukrainian refugees as they fled for safety. Spear phishing Ukrainian armed service members and their families were conducted to distract and disrupt operations on the ground.[102] Another concerning feature of this cyberwar was how other government-backed actors from China, Iran, and North Korea used Russian hostilities as cover for their own attacks.[103] There is a parasitic quality that strikes a nation beset by cyberattacks, where other hostile actors can find convenient disguise amidst the flow of malicious activity.

Coordinated cyberwar requires years of preparation by slowly gaining and escalating access, by loading and updating malware, and finding and exfiltrating data. Preparing the battle space in cyberwar is a constant process to stay embedded, anonymous, and retain access. Once a cyberwar starts, all bets are off. The fog of war becomes cover for hostile threat actors around the globe. The Russian cyberwar lacked longevity and resilient logistical planning, much like their ground war. Future cyberwars will not be able to rely on enemy disorganization as an adequate defense. As China learns about NATO strategy and alliances, it benefits by honing its own expansionist goals toward Taiwan. In March 2019, the US assessment of risk from adversaries like Russia and China was remarkably prescient. "I kind of look at Russia as the hurricane. It comes in fast and hard," Rob Joyce, NSA's senior cybersecurity adviser and former White House cybersecurity coordinator, told reporters. China, on the other hand, "is climate change: long, slow, pervasive."[104] The guise of hacktivism masks hostilities from ideologically aligned, well-supported threat actors, who find haven or direct support from host governments.[105] Rather than resting on assurances that the cyberwar will not take place, a global cyber–Cold War will represent a persistent and hostile pre-cyberwar. The lessons learned in Ukraine will serve as the basis for future conflicts.

If cyberattacks represent a growing channel for nation-state expressions of force and power, then Thomas Rid's conclusion that "cyber offenses represent an attack on violence itself" is also an admission that nations will increasingly project power at the intersection of culture and politics.[106] He argues that bits and bytes will replace bombs and bullets. Rid builds his argument on Carl von Clausewitz's concept of war, defined in the following way: "War is a mere continuation of politics by other means."[107] The mediated expression of violence in cyberwar means that it is regarded as a feature of sabotage, espionage, or subversion in warfighting, rather than the use of kinetic force causing death and destruction directly:

> Code, quite simply, doesn't come with its own explosive charge. Code-cased destruction is therefore *parasitic on the target.* Even the most sophisticated cyber attack can only physically harm a human being by unleashing the violent potential that is embedded in that targeted system. This could be a traffic control system, causing trains or planes to crash; a power plant that may explode or emit radiation; a dam that may break and cause a devastating flash flood; a pipeline that may blow up; hospital life support systems that may collapse in emergency situations; or even a pacemaker implanted in a heart patient that could be disrupted by exploiting vulnerabilities in its software. Yet so far, no such scenario has ever happened in reality.[108]

There are findings that support increased mortality in healthcare settings that experience a cyberattack.[109] The causal relationship between disease and cause of death is mediated. Rid's logic is limited by his founding definition of war that defines violence as a zero-sum. Cyberwar is a continuation of war by other means. Coordinated cyber and kinetic attacks serve to increase the psychological damage of warfare. Assuming that kinetic violence will be supplanted by cyberattacks also assumes that military victories cannot be coordinated with cyber operations. Most cyberattacks that violate national sovereignty remain below the threshold for the use of force or armed attack. These breaches are generally used for espionage, political advantage, and international statecraft, with the most damaging attacks undermining trust and confidence in social, political, and economic institutions. Determining the limits of this threshold is a tricky proposition at present.

In the July 2023 communique from the Vilnius Summit, NATO members spoke directly about Russian hybrid warfighting and appeared to mention quasi-aligned cybercriminal groups: "Russia has intensified its hybrid actions against NATO Allies and partners, including through proxies."[110] Point 66 of this wartime communique describes how "Russia's war of aggression against Ukraine has highlighted the extent to which cyber is a feature of modern conflict." While the mutual defense of NATO members will be judged on a "case-by-case basis," it is the mention of a "single or cumulative set of malicious cyber activities" that lends still greater uncertainty. A chilling set of questions emerges:

When does a cyberattack serve as a pretext for kinetic warfare? When does mediated violence become real violence? When are bytes answered with bombs?[111]

If the answer is never, governments need to protect citizens by imposing real costs on cybercriminals and nation-state supported threat actors. A case-by-case basis suggests there is a threshold, wherein political, military, or technical damage precipitates a proportional military response. A single cyberattack may result in a simpler equation of damage and proportionality. If this damage is measured in a cumulative way, the historical accounting of cyberattacks represents a significant feature of a national conversation needed to justify and rationalize the use of force. Measuring the psychological toll of a cyber-Cold War may prove difficult, but it will be felt. Given the lack of transparent accounting of cybersecurity incidents in business and government, the example provided by the Irish HSE will become a significant feature of public debate and decision making, especially when invoking the mutual defense of NATO Allies.

As a final gesture, it may be helpful to return to an early philosopher of digital spaces and cyberspace. In 1984, Donna Haraway colorfully described "modern war" as a "cyborg orgy, coded by C3I, command-control-communication-intelligence."[112] In the 1980s, these words might have been more easily disregarded. Today, the orgiastic blurring of lines is oddly familiar in the first example of cyberwar. Russian attacks follow no international rules of war. Attacking hospital IT that leads to children's deaths is less visibly destructive than bombing a day-care, though the results are the same.[113] Rules of engagement are more useful in binding allies together than in constraining adversaries.[114] Future cyberwars may not be fought by humans at all. Artificial Intelligence like recent military-trained Large Language Models will increase strategic instability by further blurring lines between threat actors.[115] There is no shortage of AI doomers: some have claimed LLMs are alive.[116] Luminaries like Geoffry Hinton and Yoshua Bengio warned of human extinction and the need for humans to adapt to a new technological paradigm.[117] Others like Timnit Gebru, Emily M. Bender, Angelina McMillian-Major, and Margaret Mitchell argued that AI must be regulated by governments and developed responsibly by corporations.[118] Without a human-in-the-loop, can the actions of a rogue AI be soft-pedaled and explained away? Can an AI commit an act of war? Who is responsible when an AI commits a war crime?

I hope these are questions for another day in the far future. Maybe by asking *what if* we can avoid the regret of asking *if only.*

Notes

[1] The Russian invasion of Ukraine must of course be set in the context of the 2008 Russian invasion of the Republic of Georgia as well as the 2014 annexation of Crimea.

[2] "Conti Ransomware," *Cybersecurity and Infrastructure Security Agency Alert*, March 9, 2022, https://www.cisa.gov/news-events/alerts/2021/09/22/conti-ransomware.

[3] Christopher Bing, "Russia-based Ransomware Group Conti Issues Warning to Kremlin Foes," *Reuters*, February 25, 2022, https://www.reuters.com/technology/russia-based-ransomware-group-conti-issues-warning-kremlin-foes-2022-02-25/. For a broader portrait of the ransomware problem and Conti's position in it, see Danny Palmer, "Ransomware: These Two Gangs are Behind Half of All Attacks," *ZDNet*, April 14, 2022,

https://www.zdnet.com/article/ransomware-these-two-gangs-are-behind-half-of-all-attacks/.

[4] conti leaks (@ContiLeaks), "Glory for Ukraine!" Twitter, February 28, 2022, 5:26 PM, https://twitter.com/ContiLeaks/status/1498424225790582785. The Conti Leaks were made available on the well-known anonfiles.com. In late August 2023, the Anon Files modes shut down the site and put the domain up for sale. Though it is not at all surprising that an anonymous file storage site would be rampantly abused, the moderators explain in their farewell message: "After trying endlessly for two years to run a file sharing site with user anonymity we have been tired of handling the extreme volumes of people abusing it and the headaches it has created for us."

[5] conti leaks (@ContiLeaks), "fuck Russian inviders!" Twitter, February 28, 2022, 5:26 PM, https://twitter.com/ContiLeaks/status/1498424329956212737.

[6] Matt Burgess, "The Workaday Life of the World's Most Dangerous Ransomware Gang," *Wired*, March 16, 2022, https://www.wired.com/story/conti-leaks-ransomware-work-life/.

[7] "Five Things We learned from the Conti Chat Logs," *Reliaquest*, April 5, 2022, https://www.reliaquest.com/blog/five-things-we-learned-from-the-conti-chat-logs/.

[8] For a complete list of cyber intendents in Ukraine, please see the National Security Archive, hosted by the George Washington University in Washington, D.C.: https://nsarchive.gwu.edu/project/ukraine-cyber-project.

[9] Sean Lyngaas, "'I can fight with a keyboard': How One Ukrainian IT Specialist Exposed a Notorious Russian Ransomware Gang," *CNN*, March 30, 2022, https://www.cnn.com/2022/03/30/politics/ukraine-hack-russian-ransomware-gang/index.html.

[10] Ned Price, "Reward Offers for Information to Bring Coni Ransomware Variant Co-Conspirators to Justice," *U.S. Department of State Press Statement*, May 6, 2022, https://www.state.gov/reward-offers-for-

information-to-bring-conti-ransomware-variant-co-conspirators-to-justice/.

[11] Eduardo Ovalle and Franscesco Figurelli, “Lockbit Leak, Research Opportunities on Tools Leaked from TAs,” *Secure List Blog by Kaspersky,* August 25, 2023, https://securelist.com/lockbit-ransomware-builder-analysis/110370/.

[12] Lawrence Abrams, “Conti Ransomware Finally Shuts Down Data Leak, Negotiation Sites,” *Beeping Computer*, June 24, 2022, https://www.bleepingcomputer.com/news/security/conti-ransomware-finally-shuts-down-data-leak-negotiation-sites/.

[13] Ibid.

[14] Brian Krebs, “Conti Ransomware Group Diaries, Part I: Evasion,” *Krebs on Security*, March 1, 2022, https://krebsonsecurity.com/2022/03/conti-ransomware-group-diaries-part-i-evasion/.

[15] Tragic or dramatic irony is asymmetrical by disparities of awareness between a targeted actor and a hostile observer. In the Aristotelian articulation of tragic irony, pity and fear are evoked within the audience because they know the mistakes made, even as the observed character is failing to identify it. In this scenario, disparities of awareness occur against spectators and actors alike.

[16] Pavlo Hrytsak, “Danylo Romanovych,” *Internet Encyclopedia of Ukraine*,https://www.encyclopediaofukraine.com/display.asp?linkpath=pages%5CD%5CA%5CDanyloRomanovych.htm.

[17] Timothy Snyder, “Ivan Ilyin, Putin’s Philosopher of Russian Fascism,” *New York Review of Books*, March 16, 2018, https://www.nybooks.com/online/2018/03/16/ivan-ilyin-putins-philosopher-of-russian-fascism/.

[18] Krebs, “Conti Ransomware Group Diaries, Part I.”

[19] Report on Trickbot: https://securityintelligence.com/posts/trickbot-gang-doubles-down-enterprise-infection/ and https://www.bleepingcomputer.com/news/security/trickbot-teams-up-

with-shatak-phishers-for-conti-ransomware-attacks/ Emotet 101: https://news.sophos.com/en-us/2019/03/05/emotet-101-stage-1-the-spam-lure/ BazarCall method or call-back phishing: https://www.bleepingcomputer.com/news/security/ransomware-gangs-move-to-callback-social-engineering-attacks/ used by conti.

[20] Krebs, “Conti Ransomware Group Diaries, Part I.”

[21] Ibid. See also “TrickBot Malware,” *Cybersecurity and Infrastructure Security Agency Advisory*, May 20, 2021, https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-076a.

[22] Assaf Dahan, Lior Rochberger, Eli Salem, Mary Zhao, Niv Yona, Omer Yampel, and Matt Hart, “Dropping Anchor: From a TrickBot Infection to the Discovery of the Anchor Malware,” *Cybereason Blog*, December 11, 2019, https://www.cybereason.com/blog/research/dropping-anchor-from-a-trickbot-infection-to-the-discovery-of-the-anchor-malware.

[23] Krebs, “Conti Ransomware Group Diaries, Part I.”

[24] Ibid.

[25] Brian Krebs, “FBI, DHS, HHS Warn of Imminent, Credible Ransomware Threat Against U.S. Hospitals,” *Krebs on Security*, October 28, 2020, https://krebsonsecurity.com/2020/10/fbi-dhs-hhs-warn-of-imminent-credible-ransomware-threat-against-u-s-hospitals/.

[26] Brian Krebs, “International Action Targets Emotet Crimeware,” *Krebs on Security*, January 27, 2021, https://krebsonsecurity.com/2021/01/international-action-targets-emotet-crimeware/. See also “Emotet Malware,” *Cybersecurity and Infrastructure Security Agency Advisory,* October 24, 2020, https://www.cisa.gov/news-events/cybersecurity-advisories/aa20-280a.

[27] Ibid. There are many examples of cybercrime software as services online that are designed to send phishing emails, by-pass MFA, or even automate phishing on specific platforms. For instance, the ESET blog *We Live Security* published an analysis of Telekopye Telegram bot that weaponizes online marketplaces. See Radek Jizba, “Telekopye: Hunting Mammoths Using Telegram Bot,” *We Live Security by ESET,*

August 24, 2023, https://www.welivesecurity.com/en/eset-research/telekopye-hunting-mammoths-using-telegram-bot/.

[28] See https://www.nist.gov/cyberframework.

[29] Vedere Labs, “Analysis of Conti Leaks,” *Forescout*, March 11, 2022, https://www.forescout.com/resources/analysis-of-conti-leaks/.

[30] Ibid.

[31] See https://www.cisa.gov/news-events/cybersecurity-advisories/aa20-302a.

[32] conti leaks (@ContiLeaks), “more sanctions! they destroy hospitals, and a lot of ppl died! even some of my friends!” *Twitter*, March 1, 2022, 5:59 PM, https://twitter.com/ContiLeaks/status/1498613833363034112.

[33] “Leaks of Conti Ransomware Group Paint Picture of a Surprisingly Normal Tech Start-Up … Sort Of,” *Check Point*, March 10, 2022, https://research.checkpoint.com/2022/leaks-of-conti-ransomware-group-paint-picture-of-a-surprisingly-normal-tech-start-up-sort-of/.

[34] Ibid.

[35] A timeline was drawn up to make sense of the leaks posted to Twitter by Malware Bytes. See “The Conti Ransomware Leaks” *Malware Bytes Labs Blog*, March 1, 2022, https://www.malwarebytes.com/blog/threat-intelligence/2022/03/the-conti-ransomware-leaks.

[36] The cybersecurity industry blogs rapidly become an echo chamber. This echo chamber must not be regarded as a negative interpretation of security rhetoric. See, for example, the consistency of insights: “An In-Depth Look at Conti’s Leaked Log Chats,” *Avertium*, March 22, 2022, https://explore.avertium.com/resource/in-depth-look-at-contis-leaked-log-chats; and Allan Liska, “5 Things We Learned from the Conti Leaks,” *Ransomware.org Blog*, March 25, 2022, https://ransomware.org/blog/5-things-we-learned-from-the-conti-leaks/; Bryce Webster-Jacobsen, “What the Conti Ransomware Group Data Leak Tells Us,” *Dark Reading*, March 24, 2022,

https://www.darkreading.com/attacks-breaches/what-the-conti-ransomware-group-data-leak-tells-us.

[37] "Conti Ransomware Group Internal Chats Leaked over Russia-Ukraine Conflict," *Rapid7 Blog*, March 1, 2022, https://www.rapid7.com/blog/post/2022/03/01/conti-ransomware-group-internal-chats-leaked-over-russia-ukraine-conflict/.

[38] "Interview with Security Researcher, Speaker Will Thomas (@BushidoToken)," *Security Noob*, July 18, 2022, https://thesecuritynoob.com/interviews/interview-with-security-researcher-speaker-will-thomas-bushidotoken/.

[39] Will Thomas, "Lessons from the Conti Leaks," *BushidoToken Threat Intel*, April 17, 2022, https://blog.bushidotoken.net/2022/04/lessons-from-conti-leaks.html.

[40] Ibid.

[41] Laundering cryptocurrency requires a platform called a "mixer" or "blender" that combines legitimate and ill-gotten digital coins. Criminals are able to withdraw laundered coins that are no longer connected to, in this case, ransomware payments. See Rosie Perper, "Are Crypto Mixers Legal?" *Coin Desk Blog*, May 11, 2023, https://www.coindesk.com/learn/are-crypto-mixers-legal/.

[42] Matt Burgess, "Conti Leak: A Ransomware Gang's Chats Expose Its Crypto Plans," *Wired*, March 17, 2022, https://www.wired.com/story/conti-ransomware-crypto-payments/. See also Kristy Moreland, "What Is a Rug Pull?" *Ledger Security Blog*, May 11, 2023, https://www.ledger.com/academy/what-is-a-rug-pull.

[43] Thomas, "Lessons from the Conti Leaks."

[44] "An Overview on Conti Ransomware Leaks: Is This the End for Conti?," *SOCRadar Blog*, March 18, 2022, https://socradar.io/an-overview-on-conti-ransomware-leaks-is-this-the-end-for-conti/.

[45] Lawrence Abrams, "Conti Ransomware Shuts Down Operation, Rebrands into Smaller Units," *Beeping Computer*, May 19, 2022, https://www.bleepingcomputer.com/news/security/conti-ransomware-

shuts-down-operation-rebrands-into-smaller-units/; Ravie Lakshmanan, “Gold Ulrick Hackers Still in Action Despite Massive Conti Ransomware Leak,” *The Hacker News*, April 26, 2022, https://thehackernews.com/2022/04/gold-ulrick-hackers-still-in-action.html; Lawrence Abrams, “New Royal Ransomware Emerges in Multi-Million Dollar Attacks,” *Beeping Computer*, September 29, 2022, https://www.bleepingcomputer.com/news/security/new-royal-ransomware-emerges-in-multi-million-dollar-attacks/; Steven Campbell, Akshay Suthar, Connor Belfiore, and Arctic Wolf Labs Team, “Conti and Akira: Chained Together,” *Wolf Labs Blog*, July 26, 2023, https://arcticwolf.com/resources/blog/conti-and-akira-chained-together/.

[46] John Fokker, and Jambul Tologonov, “Conti Leaks: Examining the Panama Papers of Ransomware,” *Trellix Blog*, March 31, 2022, https://www.trellix.com/blogs/research/conti-leaks-examining-the-panama-papers-of-ransomware/.

[47] Brian Krebs, “Inside Ireland’s Public Healthcare Ransomware Scare,” *Krebs on Security*, December 13, 2021, https://krebsonsecurity.com/2021/12/inside-irelands-public-healthcare-ransomware-scare/.

[48] “Conti Cyber Attack on the HSE: Independent Post Incident Review.” *PricewaterhouseCoopers*, December 3, 2021, https://www.hse.ie/eng/services/news/media/pressrel/hse-publishes-independent-report-on-conti-cyber-attack.html.

[49] Ibid., 2.

[50] Ibid.

[51] Ibid., 38.

[52] Ibid., 4.

[53] Ibid., 4.

[54] Ibid., 38.

[55] Ibid.

[56] Ibid., 76.

[57] Ibid.

[58] Ibid.

[59] Ibid., 79.

[60] These leaked documents were downloaded as they happened on Twitter. The majority of the leaked documents are in Russian. The chat logs appear as JSON files, while the documentation and manuals appear in plain text. They were translated with a Python library called “deep translator” with access to Google Translate, Microsoft Translator, as well as ChatGPT Translator. See Nidhal Baccouri, “deep-translator,” *GitHub*, July 1, 2023, https://github.com/nidhaloff/deep-translator.

[61] Labs, “Analysis of Conti Leaks.”

[62] Thomas Roccia, “Cybercriminals Actively Exploiting RDP to Target Remote Organizations,” *McAfee Blog*, May 6, 2020, https://www.mcafee.com/blogs/other-blogs/mcafee-labs/cybercriminals-actively-exploiting-rdp-to-target-remote-organizations/.

[63] Here is the section “Network Landscape” of the *Hacker’s Quickstart Guide* in Russian:

ЛАНДШАФТЫ СЕТЕЙ

Любую (почти) современную сеть можно взломать.

Это обусловлено:

- избыточностью сетей: наличие множества сервисов, разных точек входа в одну и ту же сеть;

- приоритетом удобства над безопасностью: тюрьма безопасна, но в ней очень трудно что-то делать;

- человеческим фактором: ошибки конфигурации, социальная инженерия.

Второй пункт усиливается ощутимой реакцией прибыли на малейшие замедления оборачиваемости в коммерческом секторе,

так что без перехода на военные рельсы сети капитализма всегда будут дырявые)

https://habr.com/ru/company/selectel/blog/576482/

Если в найденных точках входа нет *известных* уязвимостей, это значит лишь, что

- нужно искать другие точки входа;
- нужно искать уязвимости самому (если уж очень нужно попасть в сеть);
- нужно искать человека;
- нужно искать другую цель с той же информацией.

Хорошо защищают сети те организации, которым государство выставляет требования по защите.

Это необязательно военные сети или режимные учреждения: если вы храните у себя ФИО и личные данные клиентов,

то вы обязаны провести мероприятия по их защите. Защита же коммерческой тайны – “ваши” проблемы.

Ценность цели часто обратно пропорциональна её защите.

В военных сетях могут оказаться списки кальсонов за 196х год (что полезно для военных аналитиков),

а в слабо защищенной коммерческой сети, или на личном ноутбуке, могут быть важнейшие фарм/ИТ/инженерные разработки.

Но это не всегда так.

Есть суперкрепости внутри плохо защищенных сетей, взять которые может либо разносторонняя команда, либо хакер экстра-класса.

Такие сценарии обычны в Standoff’ах (Hack The Box etc) для белых шляп.

Если вы не хакер экстра-класса, используйте смекалку (уроните сервер и словите в сети/ закейлогьте пароль, отдайте задачу на аутсорс, итд)

[64] I will refrain from listing examples here, but they are easy to find.

[65] Agile Software Development is a set of practices and principles used in software development. In 2001, seventeen developers authored a *Manifesto for Agile Software Development*, which is available at https://agilemanifesto.org/. Today, Agile has grown into a complex development philosophy, with a business logic that suits the needs of developers

[66] The free software movement seeks to guarantee the freedom of users to both modify and share computer code. Emerging directly from the hacker culture of the 1970s and 1980s, the Free Software Foundation was founded by Richard Stallman in 1985. Learn more about the philosophy animating the free software movement here: https://www.gnu.org/philosophy/philosophy.html.

[67] A complete list of recommended tools in the training manual include the following: “1. Metasploit Framework (MSF) (+ armitage GUI)—the largest selection of sploits and modules 2. Core Impact (+impacket python)—the most convenient features for penetration testing (of the minuses—only Windows) 3. Powershell Empire—pure powershell framework with all the consequences 4. Posh2c 5. Koadik—these two are exotic, that is, they have less detectable traffic 6. Cobalt Strike—expandable https://www.cobaltstrike.com/downloads/csmanual43.pdf 7. Burp Suite—web oriented, very popular 8. Pupy—RAT (Remote Administration Tool) in Python, difficult for AB due to the fact that it is not the usual AB native code, and not native to Windows (and AMSI) scripting language https://github.com/n1nj4sec/pupyhttps://ptestmethod.readthedocs.io/en/latest/pupy.html and it comes with an injector https://github.com/infodox/python-dll-injection.”

[68] See https://www.kali.org/. There are other penetration testing oriented OS distributions, like Parrot OS: https://www.parrotsec.org/. Conti appears to prefer dragons over parrots.

[69] Please note that I have used “deep-translator” to translate the passages in Russian, which appear in all caps in the original and the version present here.

[70] See the World Bank’s 2021 report, “The Changing Wealth of Nations 2021: Managing Assets for the Future.” Available http://hdl.handle.net/10986/36400.

[71] Richard H. Thaler, and Cass R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (New York: Penguin, 2008), 4–5.

[72] Ibid., 5.

[73] Ibid., 3.

[74] Ben D. Sawyer, and Peter A. Hancock, “Hacking the Human: The Prevalence Paradox in Cybersecurity1,” *Human Factors* 60 no. 5 (2018): 597–609. There have even bipartisan efforts to ban “dark patterns” in law as the “Deceptive Experiences To Online Users Reduction (DETOUR) Act.” See Mark R. Warner, “Warner, Fischer

Lead Bipartisan Reintroduction of Legislation to Ban Manipulative 'Dark Patterns,'" *Mark R. Warner Press Release*, July 28, 2023, https://www.warner.senate.gov/public/index.cfm/2023/7/warner-fischer-lead-bipartisan-reintroduction-of-legislation-to-ban-manipulative-dark-patterns.

[75] Jon Yablonski, *Laws of UX: Using Psychology to Design Better Products and Services* (Sebastopol, CA: O'Reilly Media, 2020), 24.

[76] Yablonski, *Laws of UX*, 25.

[77] Roger Grice, "Experience Architecture: Drawing Principles from Life," in *Rhetoric and Experience Architecture*. Liza Potts and Michael J. Salvo (eds.) (Anderson, SC: Parlor Press, 2017): 49.

[78] Thaler and Susstein, *Nudge*, 83–102.

[79] Robert Bringhurst, *The Elements of Typographic Style, fourth edition* (Vancouver: Hartley and Marks Publishers, 2019), 17.

[80] Ibid., 20.

[81] Matt Burgess, "Leaked Ransomware Docs Show Conti Helping Putin from the Shadows," *Wired*, March 18, 2022, https://www.wired.co.uk/article/conti-ransomware-russia.

[82] The Risky.biz podcast published a feature interview called "How Sandworm prepared Ukraine for a cyber war" on August 21, 2023. It featured an interview with the Security Service of Ukraine (SBU), Head of the Department of Cyber and Information Security, Illia Vitiuk. This interview is the most up to date, public portrait of the ongoing cyberware between Russia and Ukraine. See https://risky.biz/illiavitiuk/.

[83] Michael Beckley, "The Peril of Peaking Powers: Economic Slowdowns and Implications for China's Next Decade," *International Security* 48 no. 1 (2023): 7–46. For example, the International Memorial project, first founded in Russia in 1992, has been liquidated by the Russian Supreme Court for its work in preserving the human rights abuses of the USSR and the current Russian government under Putin. Because this project runs counter to the Russian government's cultural narrative

and latest imperial ambitions, it was expedient for the corrupt government and courts to destroy the project within Russia. Please see https://www.memo.ru/en-us/.

84 Matt Burgess, "Russia Is Quietly Ramping Up Its Internet Censorship Machine," *Wired*, July 25, 2022, https://www.wired.com/story/russia-internet-censorship-splinternet/.

85 See "Cyber Defense," *NATO, What We Do*, August 3, 2023, https://www.nato.int/cps/en/natohq/topics_78170.htm.

86 Please see my review of Andrii Demartino's *False Mirrors* and my short article "Ukrainian cultural artifacts are at risk during the Russian invasion, but digitizing them may offer some protection," respectively: https://www.tandfonline.com/doi/full/10.1080/00085006.2022.2136841 and https://theconversation.com/ukrainian-cultural-artifacts-are-at-risk-during-the-russian-invasion-but-digitizing-them-may-offer-some-protection-185673.

87 Pierrre-Marc Bureau, "Initial Access Broker Repurposing Techniques in Targeted Attacks Against Ukraine," *Google TAG*, September 7, 2022, https://blog.google/threat-analysis-group/initial-access-broker-repurposing-techniques-in-targeted-attacks-against-ukraine/.

88 Ibid.

89 Dan Goodin, "Ukraine Is Under Attack by Hacking Tools Repurposed from Conti Cybercrime Group," *Ars Technica*, September 7, 2022, https://arstechnica.com/information-technology/2022/09/hackers-with-conti-cybercrime-group-are-repurposing-tools-for-attacks-on-ukraine/; Ole Villadsen, Charlotte Hammond, and Kat Weinberger, "Unprecedented Shift: The Trickbot Group Is Systematically Attack Ukraine," *IBM Security Intelligence Blog*, July 7, 2022, https://securityintelligence.com/posts/trickbot-group-systematically-attacking-ukraine/.

90 Matt Burgess, "Ukraine's Volunteer 'IT Army' Is Hacking in Uncharted Territory," *Wired*, February 27, 2022, https://www.wired.com/story/ukraine-it-army-russia-war-cyberattacks-ddos/.

[91] Mykhailo Fedorov, "The First World Cyber War. The first IT Army in the world. 270K of angry IT-warriors of cyber frontline. Rutube shutdown. AI tech & identification of war criminals. And many more cases to disclose after the victory. You are free to join, by the way." *Twitter*, May 26, 2022, 7:50 am, https://twitter.com/FedorovMykhailo/status/1529792057442717696.

[92] Geoffery Cain, "Volodymyr Zelensky on War, Technology, and the Future of Ukraine" *Wired*, June 2, 2022, https://www.wired.com/story/volodymyr-zelensky-q-and-a-ukraine-war-technology/.

[93] "Cyber Threat Activity Related to the Russian Invasion of Ukraine," *Canadian Centre for Cyber Security,* May 3, 2023, https://www.cyber.gc.ca/sites/default/files/cyber-threat-activity-associated-russian-invasion-ukraine-e.pdf.

[94] With Microsoft's telemetry network, it is little surprise that they were among the first to report on malicious behavior. See Tom Burt, "Malware Attacks Targeting Ukraine Government," *Microsoft Blog,* January 15, 2022, https://blogs.microsoft.com/on-the-issues/2022/01/15/mstic-malware-cyberattacks-ukraine-government/; see also Microsoft's security blog for coverage of the so-called Cadet Blizzard threat actor during this period: https://www.microsoft.com/en-us/security/blog/2022/01/15/destructive-malware-targeting-ukrainian-organizations/.

[95] Matt Burgess, "Viasat Satellite Hack Spills Beyond Russia-Ukraine War," *Wired*, March 23, 2022, https://www.wired.com/story/viasat-internet-hack-ukraine-russia/.

[96] Yuras Kurmanau, and Frank Bajak, "Cyberattacks Knock Out Sites of Ukrainian Army, Major Banks," *APNews*, February 15, 2022, https://apnews.com/article/russia-ukraine-technology-business-europe-russia-e791990f60841b599f664c34f58403de.

[97] See "HermeticWiper: New Data-Wiping Malware Hits Ukraine," *We Live Security Blog*, February 24, 2022, https://www.welivesecurity.com/2022/02/24/hermeticwiper-new-data-

wiping-malware-hits-ukraine/; Juan Andrés Guerrero-Saade, “HermeticWiper, New Destructive Malware Used in Cyber Attacks on Ukraine,” *Sentinel Labs Blog*, February 23, 2022, https://www.sentinelone.com/labs/hermetic-wiper-ukraine-under-attack/.

[98] Andy Greenberg, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin’s Most Dangerous Hackers* (New York: Knopf Doubleday, 2019).

[99] Mehul Srivastava, madhumita Murgia, and Hannah Murphy, “*The Secret US. Mission to Bolster Ukraine’s Cyber Defenses Ahead of Russia’s Invasion,” Ars Technica*, March 9, 2022, https://arstechnica.com/information-technology/2022/03/the-secret-us-mission-to-bolster-ukraines-cyber-defences-ahead-of-russias-invasion/. This article was first published in the *Financial Times*, which is paywalled.

[100] Matthew Gault, “Pentagon’s Ukraine War Plans Leaked on Minecraft Discord before Telegram and Twitter,” *Motherboard*, April 7, 2023, https://www.vice.com/en/article/pkadnb/pentagons-ukraine-war-plans-leaked-on-minecraft-discord-before-telegram-and-twitter.

[101] David Cattler, and Daniel Black, “The Myth of the Missing Cyberwar,” *Foreign Affairs*, April 6, 2022, https://www.foreignaffairs.com/articles/ukraine/2022-04-06/myth-missing-cyberwar.

[102] Sergiu Gatlan, “Phishing Attacks Target Countries Aiding Ukrainian Refugees,” *Beeping Computer,* March 2, 2022, https://www.bleepingcomputer.com/news/security/phishing-attacks-target-countries-aiding-ukrainian-refugees/.

[103] Sergiu Gatlan, “Russian Phishing Attacks Target NATO, European Military,” *Beeping Computer,* March 30, 2022, https://www.bleepingcomputer.com/news/security/google-russian-phishing-attacks-target-nato-european-military/.

[104] Bastien Inzaurralde, “The Cybersecurity 202: U.S. officials: It’s China Hacking that Keeps Us up at Night,” *Washington Post*, March 6, 2019,

https://www.washingtonpost.com/news/powerpost/paloma/the-cybersecurity-202/2019/03/06/the-cybersecurity-202-u-s-officials-it-s-china-hacking-that-keeps-us-up-at-night/5c7ec07f1b326b2d177d5fd3/.

105 A.J. Vicens, "Hacking Group KittenSec Claims to 'pwn anything we see' to Expose Corruption," *Cyberscoop*, August 24, 2023, https://cyberscoop.com/kittensec-hacktivism-corruption/.

106 Thomas Rid, *Cyber War Will Not Take Place* (London: Hurst and Company, 2013), 166.

107 Ibid., 2.

108 Ibid., 13. Emphasis in the original.

109 In a 2021 Censinet report that surveyed nearly 600 health delivery organizations, nearly one quarter of survey participants reported an increase in mortality rates as a result of ransomware. See https://www.censinet.com/ponemon-report-covid-impact-ransomware.

110 See https://www.nato.int/cps/en/natohq/official_texts_217320.htm. Please note that NATO will hold a Cyber Defence Conference in Berlin, in November of 2023. This will be the first of its kind meeting for NATO Allies.

111 Jeremy Straub, "Defining, Evaluating, Preparing for and Responding to a Cyber Pearl Harbor," *Technology in Society* 65 (May 2021): 1–10.

112 Donna J. Haraway, "A Cyborg Manifesto," in *Manifestly Haraway* (Minneapolis: University of Minnesota Press, 2016): 6.

113 See "Ukraine: Cluster Munitions Kill Child and Two Other Civilians Taking Shelter at a Preschool," *Amnesty International*, February 27, 2022, https://www.amnesty.org/en/latest/news/2022/02/ukraine-cluster-munitions-kill-child-and-two-other-civilians-taking-shelter-at-a-preschool/.

114 Nathaniel Fick, and Jami Miscik, "Confronting Reality in Cyberspace: Foreign Policy for a Fragmented Internet," *Council on Foreign Relations Report*, July 2022, https://www.cfr.org/task-force-report/confronting-reality-in-cyberspace/findings.

115 Miachel E. O'Hanlon, "The Role of AI in Future Warfare," *Brookings Institute*, 29, November 2018, https://www.brookings.edu/articles/ai-and-future-warfare/.

116 Nitasha Tiku, "The Google Engineer Who Thinks the Company's AI Has Come to Life," *The Washington Post*, June 11, 2022, https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/.

117 See "Pause Giant AI Experiments: An Open Letter," *Future of Life Institute*, March 22, 2023, futureoflife.org/open-letter/pause-giant-ai-experiments/; "AI Extinction Statement Press Release," *Center for AI Safety Press Release*, May 30, 2023, https://www.safe.ai/press-release.

118 Timnit Gebru, Emily M. Bender, Angelina McMillian-Major, and Margaret Mitchell, "Statement from the Listed Authors of Stochastic Parrots on the 'AI Pause' Letter," *Distributed AI Research Institute Press Release*, March 31, 2023, https://www.dair-institute.org/blog/letter-statement-March2023/. See also Isabella Struckman and Sofie Kupiec, "Why They're Worried: Examining Experts' Motivations for Signing the 'Pause Letter'" *Arxiv CS*, June 19, 2023, https://arxiv.org/abs/2306.00891v2.

Coda: Beyond Threat Informed Security

This book begins by describing a broad cultural project that takes cybersecurity as a critical domain in ensuring the safety of cultural objects and culture workers. I define *security rhetoric* as a kind of critical noticing, under the aegis of the ancient Greek concept of *kairos*. Security rhetoric builds broader understanding about digital culture by describing mechanisms and history of deceit and lying, wherein reasoned and empathetic judgment can be applied to thwart attacks. It is the very moment, when this sound judgment can be made through a prospective moment, that asks, "what if?" These counterfactual moments are speculative and benefit from an imaginative understanding of the rhetorical situation. Perhaps most significantly, affective security considers the local consequences on a community to which the user is a member. A security culture that places value on this moment of noticing might be productively used to augment the current approach to online security by moving beyond the war metaphors and beyond threat informed security as mere "cybersecurity awareness."

Currently, the cybersecurity industry privileges knowledge about threats rather than an understanding of defenders. Defenders are regarded as assets to protect or human resources who are a vulnerability to mitigate. Given the scale of cybersecurity failures described in this book alone, there are grounds to claim that threat informed defensive security is simply not working. In order to starve cybercriminal organizations of the access and opportunity to profit from deception, a culture of security must build cybersecurity training as a first principle of educational curriculum, work practices, and even research endeavors. Security rhetoric pays close attention to tactics, techniques, and procedures of attackers, but it also values the agency, psychology, and feelings of defenders.

Threat informed defense is an excellent business model to maintain constant and ongoing demand from institutional clients for threat intelligence, evaluation, and software licenses, but it will never deliver a transformative change in the security landscape, especially with the emergence of increasingly advanced AI systems. The kind of resilience needed to thwart attacks by sophisticated actors like Conti on medical institutions during a global crisis must be collectively earned through cultural consensus. The existing rationale holds that, if individual users are sufficiently informed about tactics, techniques, and procedures of attackers, those same users, usually an employee, will be able to identify and thwart attacks in real time. Those in workplaces deemed significant enough for daily updates may be better protected, but there is no moat to shelter networked environments exposed to the global internet and software supply chains. The literacy needed to reliably achieve this moment of noticing cannot simply be refined by a stream of security alerts. A broad-based security literacy must emerge from a consensus about a shared culture that values the security of others.

A threat-centered defense focuses attention on the aggressor for identification and responses. Mitre, the company behind the Mitre ATT&CK Framework for naming and describing security threats, defines threat-informed cybersecurity programs “as one that successfully incorporates threat information into its regular security practices, and thereby enhances both its tactical and strategic defensive capabilities.”[1] Threat-informed defense includes three pillars of response such as the collection and analysis of cyber threat intelligence, a range of technical and policy-driven defensive measures, as well as a system of testing and evaluation. Implicit in this definition is how these approaches work to secure the people, procedures, and technology involved. The kind of broadly held security literacy I describe requires an engrained cultural awareness about how technical systems shape our lives online and elsewhere. There must be a deeply felt and intuitive *knowledge* about security that comes from the ability to understand and act on the threats in our online environment. Defenders need space and time to develop boundaries. Harangued by endless, cascading emergencies narrows the window of tolerance and limits the sound judgment of defenders. The emotional tenor of work and the media must be dialed back to give people time to think critically.

If security rhetoric is a durable approach for increased understanding of security situations, a foundational understanding must build on simple and irreducible concepts. There is an opportunity to articulate some general principles that emerge from a humanities-focused approach to cybersecurity: what are the first principles of security rhetoric? The notion of first principles is once again classical in origin. Aristotle's *Physics* sets the very condition in which knowledge (*epistêmê*) is possible, which seems like a good place to conclude.[2] If knowledge and understanding are graspable by human minds, there must be some first observations that ground all future insights. Aristotle preferred to believe that first principles are possible through dedicated observation, which allows for the first proposition of hypotheses.

This book also began by proposing a somewhat dangerous interdisciplinary project that joins Digital Humanities (DH) and cybersecurity practices. If you are reading this sentence, it suggests that the discipline of DH and interdisciplinarity cybersecurity is not so anathema to your sensibilities that you have slammed it shut! In the opening chapter, I make several recommendations for my colleagues in the universities with public facing research practices, where exposure to the internet and political ideology foregrounds the urgent need for research security practices. I recommend developing a *threat model* for research that includes an understanding of research assets and *risk assessments* that also include risks to research participants. Because so much research aligns with open access and transparency, there is a need to *validate open source software*. The software supply chain is a complex problem to solve for many researchers but the longevity and durability of public facing software, tools, and websites require an awareness of this potential vulnerability. Also, *activist scholars* engaged in data collection and archiving will require support from institutions as research methods expand. These first principles include a need for technical excellence because the cultural artifacts housed in DH projects represent the shared cultural memory of humanity. The stakes could not be higher, and *failure is not an option*, which is why a return to first principles will ensure some measure of durability in these claims.

A template for the future of security rhetoric lies with the Open Source Intelligence (OSINT) community, led in many ways by Bellingcat and its founder Eliot Higgins. Having reported on everything from Syrian and Yemeni Civil Wars as well as ongoing reporting on the Russo-Ukrainian War from 2016 to present, Bellingcat has stood as a template for a new type of investigative journalism that is predicated on open data collection and fact-checking official sources. They have investigated war crimes perpetrated in Syria with the use of chemical and cluster munitions by President Bashar al-Assad. They have reported on the well-reported poisoning of Alexei Navalny by the Russian Federal Security Service (FSB) with the military nerve agent Novichok in 2020. The 2022 documentary file Navalny brought further attention to clandestine Russian assassination attempts, like those conducted against Sergei and Yulia Skripal in the UK in 2018.

These are just a few of the high-profile reports conducted by Bellingcat investigators in their short decade long history. During this time, the core team of just eighteen staffers have also developed the tradecraft that makes use of a strange paradox of the disinformation age: "in this age of online disinformation," Higgins reflects in the opening of *We are Bellingcat,* "facts are easier to come by than ever."[3] In this 2021 book, Higgins demonstrates how effective "online open-source investigation" can be in pursuing criminal behavior of the powerful by exposing evidence of wrongdoing and demanding accountability on a global stage.[4]

Bellingcat is developing a "new methodology" for citizen open-source investigators to track and describe, what Higgins calls, the "Counterfactual Community."[5] Building on the work of the Syrian Archive,[6] which catalogues evidence related to ongoing war crimes, this new methodology seeks to archive conflicts as they happen and make them durable and searchable repositories of fact: "Eventually, you could study, say, demonstrations in oppressive regimes around the world, and check what munitions are used by riot police, thereby figuring out who is selling the weaponry that supports despots."[7] Picking up openly available information on the web holds those in power accountable through nothing more than the transparent display of information. Social media provided the access and opportunity to deploy citizen-driven open source investigations. With each new media form to emerge, new opportunities and new threats emerge.

Systems like FraudGPT and WormGPT are just the first examples of criminal LLMs and will not be the last.[8]

The opportunities and threats posed by AI systems represent the next turn. Any response to automated trolls or deepfakes will require vulnerable communities to become informed to prepare and respond. In the words of Higgins, "the uninformed give this technology powers beyond its current capabilities."[9] The challenges posed by multimodal AI systems capable of generating multilingual text, image, video, and speech capable of deceiving human audiences are evolving at such a pace that it is a change to stay fully informed, even for experts in a particular aspect of AI. If the current revolution in these AI systems is to be harnessed productively for this new methodology, it could work to automatically describe video of ongoing war crimes or detect online toxicity for content moderation.[10] The ability to monitor malicious behavior in real time could inform United Nations resolutions and the re-emergence of Peacekeeping as an approach to mitigating humanitarian disasters and cyberattacks alike. In this way, security culture remains an expression of values. What and who we protect represent judgment calls about the people and the cultures that matter.

Notes

[1] Clement Skorupka, and Lindsley Boiney, "Cyber Operations Rapid Assessment (CORA): A Guide to Best Practices for Threat-Informed Cyber Security Operations," *Mitre*, February 25, 2016, https://www.mitre.org/news-insights/publication/cyber-operations-rapid-assessment-cora-guide-best-practices-threat.

[2] There are authors in both DH and cybersecurity who think in terms of first principles. John Unsworth and Rick Howard are likely known to readers hailing from respective camps: See John Unsworth, "Scholarly Primitives: what methods do humanities researchers have in common, and how might our tools reflect this?" originally delivered at the Symposium on Humanities Computing: formal methods, experimental practice sponsored by King's College, London, May 13, 2000. Available at https://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html and Rick Howard, *Cybersecurity First Principles: A*

Reboot of Strategy and Tactics (Hoboken NJ: John Wiley & Sons, Inc., 2023).

3 Eliot Higgins, *We Are Bellingcat: Global Crime, Online Sleuths, and the Bold Future of News* (New York: Bloomsbury Publishing, 2021), 3.

4 Ibid., 7.

5 Ibid., 204, 137.

6 See https://syrianarchive.org/ and https://mnemonic.org/.

7 Higgins, *We are Bellingcat*, 207.

8 Daniel Kelley, "WormGPT—The Generative AI Tool Cybercriminals Are Using to Launch Business Email Compromise Attacks," *Slashnext,*July 13, 2023, https://slashnext.com/blog/wormgpt-the-generative-ai-tool-cybercriminals-are-using-to-launch-business-email-compromise-attacks/ and Matt Burgess, "Criminals Have Created Their Own ChatGPT Clones," *Wired*, August 7, 2023, https://www.wired.com/story/chatgpt-scams-fraudgpt-wormgpt-crime/.

9 Ibid.

10 See https://www.perspectiveapi.com/.

SELECTED BIBLIOGRAPHY

Ahmed, Sara. *The Cultural Politics of Emotion, second edition.* New York: Routledge, 2015.

Altemeyer, Bob. *The Authoritarian Specter*. Cambridge, MA: Harvard University Press, 1996.

Barrett, Lisa Feldman. *How Emotions Are Made: The Secret Life of the Brain*. New York: Mariner, 2018.

Bogost, Ian. *Persuasive Games: The Expressive Power of Video Games*. Cambridge, MA: MIT Press, 2007.

Bok, Sissela. *Lying: Moral Choice in Public and Private Life*. New York: Vintage, 1989.

Borch-Jacobsen, Mikkel. *Remembering Anna O.: A Century of Mystification*. New York: Routledge, 1996.

Bowlby, John. *Attachment and Loss, Volume I: Attachment*. New York: Penguin, 1969.

Bown, Nicola J., Daniel Read and Barbara Summers. “The Lure of Choice,” *Journal of Behavioral Decision Making* 16 (2003): 297–308.

Brooks, Rosa. *How Everything Became War and the Military Became Everything: Tales from the Pentagon*. New York: Simon and Schuster, 2017.

Brunton, Finn. *Spam: A Shadow History of the Internet*. Boston: MIT Press, 2013.

Burke, Kenneth. *A Rhetoric of Motives*. Berkeley, CA: University of California Press, 1969.

Burkett, Randy. “Rethinking an Old Approach: An Alternative Framework for Agent Recruitment: From MICE to RASCLS,” *Studies in Intelligence* 57 no. 1 (2013): 7–17.

Byrne, Ruth. *The Rational Imagination: How People Create Alternatives to Reality*. Cambridge, MA: MIT Press, 2005.

Carey, Marcus J. and Jennifer Jin. *Tribe of Hackers: Security Leaders*. Indianapolis, IN: John Wiley and Sons Inc., 2020.

Connolly, William E. *Neuropolitics: Thinking, Culture, Speed*. Minneapolis: University of Minnesota Press, 2002.

"Conti Cyber Attack on the HSE: Independent Post Incident Review." *PricewaterhouseCoopers*, December 3, 2021, https://www.hse.ie/eng/services/news/media/pressrel/hse-publishes-independent-report-on-conti-cyber-attack.html.

Crenshaw, Kimberlé W. *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*. Faculty Scholarship: Columbia Law School, 1989. Available at: https://scholarship.law.columbia.edu/faculty_scholarship/3007.

Deibert, Ron and Rafal Rohozinksi. "Tracking GhostNet: Investigating a Cyber Espionage Network." *Citizen Lab*, March 29, 2009, https://citizenlab.ca/wp-content/uploads/2017/05/ghostnet.pdf.

Demartino, Andrii. *False Mirrors: The Weaponization of Social Media in Russia's Operation to Annex Crimea*. Stuttgart: Ibidem, 2021.

Donovan, Joan, Emily Dreyfuss and Brian Friedberg. *Meme Wars: The Untold Story of the Online Battles Upending Democracy in America*. New York; London: Bloomsbury Publishing, 2022.

Dufresne, Todd. *Killing Freud: Twentieth Century Culture and the Death of Psychoanalysis*. New York: Continuum, 2003.

Eyman, Douglas. *Digital Rhetoric: Theory, Method, Practice*. Ann Arbor, MI: University of Michigan Press, 2015.

Felski, Rita. *Uses of Literature*. Malden, MA: Blackwell Publishing, 2008.

Fraad, Harriet. "Toiling in the Field of Emotion," *The Journal of Psychohistory* 35 no. 3 (2008): 270–86.

Freud, Sigmund. *Inhibitions, Symptoms and Anxiety.* Alix Strachey, trans. London: Hogarth Press, 1949.

Graeber, David and David Wengrow. *The Dawn of Everything: A New History of Humanity*. New York: Signal, 2021.

Greenberg, Andy. *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*. New York: Knopf Doubleday, 2019.

Hadnagy, Christopher. *Social Engineering: The Art of Human Hacking*. Indianapolis, IN: John Wiley and Sons Inc., 2011.

Hadnagy, Christopher. *Social Engineering: The Science of Human Hacking*. Indianapolis, IN: John Wiley and Sons Inc., 2018.

Hadnagy, Christopher and Michele Fincher. *Phishing Dark Waters: The Offensive and Defensive Sides of Malicious E-mails*. Indianapolis, IN: John Wiley and Sons Inc., 2015.

Higgins, Eliot. *We are Bellingcat: Global Crime, Online Sleuths, and the Bold Future of News*. New York: Bloomsbury Publishing, 2021.

Hochschild, Arlie Russell. *The Managed Heart: Commercialization of Human Feeling*. Berkeley, CA: University of California Press, 2012.

Howard, Rick. *Cybersecurity First Principles: A Reboot of Strategy and Tactics*. Hoboken, NJ: John Wiley & Sons, Inc., 2023.

Krebs, Brian. "Conti Ransomware Group Diaries, Part I: Evasion." *Krebs on Security*, March 1, 2022, https://krebsonsecurity.com/2022/03/conti-ransomware-group-diaries-part-i-evasion/.

Krebs, Brian. "Conti Ransomware Group Diaries, Part II: The Office." *Krebs on Security*, March 2, 2022, https://krebsonsecurity.com/2022/03/conti-ransomware-group-diaries-part-ii-the-office/.

Krebs, Brian. "Conti Ransomware Group Diaries, Part III: Weaponry." *Krebs on Security*, March 4, 2022, https://krebsonsecurity.com/2022/03/conti-ransomware-group-diaries-part-iii-weaponry/.

Krebs, Brian. "Conti Ransomware Group Diaries, Part IV: Cyptocrime." *Krebs on Security*, March 7, 2022, https://krebsonsecurity.com/2022/03/conti-ransomware-group-diaries-part-iv-cryptocrime/.

Krebs, Brian. *Spam Nation: The Insight Story of Organized Cybercrime—From Global Epidemic to Your Front Door*. Naperville, IL: Sourcebooks, 2014.

Lessig, Lawrence. *Code and Other Laws of Cyberspace*. New York: Basic Books, 1999.

Leys, Ruth. "The Turn to Affect: A Critique," *Critical Inquiry* 37 no. 3 (2011): 434–72.

Losh, Elizabeth. *Selfie Democracy: The New Digital Politics of Disruption and Insurrection*. Cambridge, MA: MIT Press, 2022.

Losh, Elizabeth. *Virtualpolitik: An Electronic History of Government Media-Making in a Time of War, Scandal, Disaster, Miscommunication, and Mistakes*. Cambridge, MA: MIT Press, 2009.

Ludlow, Peter (ed.). *High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace*. Cambridge, MA: MIT Press, 1996.

Lyngaas, Sean. "'I Can Fight with a Keyboard': How One Ukrainian IT Specialist Exposed a Notorious Russian Ransomware Gang." *CNN*, March 30, 2022, https://www.cnn.com/2022/03/30/politics/ukraine-hack-russian-ransomware-gang/index.html.

Manovich, Lev. *Cultural Analytics*. Cambridge, MA: MIT Press, 2020.

Massumi, Brian. *Parables for the Virtual: Movement, Affect, Sensation*. Durham: University of North Carolina Press, 2002.

Maté, Gabor. *The Myth of Normal: Trauma, Illness and Healing in a Toxic Culture*. New York: Alfred A. Knopf, 2022.

Mauro, Aaron. *Hacking in the Humanities: Cybersecurity, Speculative Fiction, and Navigating a Digital Future*. New York: Bloomsbury Publishing, 2022.

McKerrow, Raymie E. "Critical Rhetoric: Theory and Praxis," *Communication Monographs* 56 (1989): 91–111.

Miller, Alyssa. *Cybersecurity Career Guide*. New York: Manning Publications, 2022.

Nyberg, David. *The Varnished Truth: Truth Telling and Deceiving in Ordinary Life*. Chicago: University of Chicago Press, 1992.

Potts, Liza and Michael J. Salvo (eds.). *Rhetoric and Experience Architecture*. Anderson, SC: Parlor Press, 2017.

Ratcliffe, Krista. *Rhetorical Listening: Identification, Gender, Whiteness*. Carbondale: Southern Illinois University Press, 2005.

Renaud, Karen, Verena Zimmermann, Tim Schürmann and Carlos Böhm. "Exploring Cybersecurity-Related Emotions and Finding That They Are Challenging to Measure," *Humanities and Social Sciences Communications* 8 (2021). https://doi.org/10.1057/s41599-021-00746-5.

Rid, Thomas. *Cyber War Will Not Take Place*. London: Hurst and Company, 2013.

Rid, Thomas. "Cyber War Will Not Take Place," *The Journal of Strategic Studies* 35 no. 1 (2012): 5–32.

Ridolfo, Jim and William Hart-Davidson (eds.). *Rhetoric and the Digital Humanities*. Chicago: University of Chicago Press, 2015.

Sano-Franchini, Jennifer. “Cultural Rhetorics and the Digital Humanities: Toward Cultural Reflexivity in Digital Making,” in *Rhetoric and the Digital Humanities*. Jim Ridolfo and William Hart-Davidson (eds.). Chicago: University of Chicago Press, 2015: 49–64.

Sano-Franchini, Jennifer. “Feminist Rhetorics and Interaction Design: Facilitating Socially Responsible Design,” in *Rhetoric and Experience Architecture*. Liza Potts and Michael J. Salvo (eds.). Anderson: NC, Parlor Press, 2017: 84–108.

Sawyer, Ben D. and Peter A. Hancock, “Hacking the Human: The Prevalence Paradox in Cybersecurity,” *Human Factors* 60 no. 5 (2018): 597–609.

Schneier, Bruce. *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World*. New York: W. W. Norton and Company, 2018.

Scull, Andrew. *Psychiatry and Its Discontents*. Oakland, CA: University of California Press, 2019.

Smith, Craig Higson, Daniel Ó Cluanaigh, Ali G. Ravi and Peter Steudtner. *The Holistic Security Manual: A Strategy Manual for Human Rights Defenders*. Berlin: Tactical Technology Collective, 2016. Available at: https://holistic-security.tacticaltech.org/.

Steele, Catherine Knight. *Digital Black Feminism*. New York: New York University Press, 2021.

Stern, Nancy and Robert A. Stern. *Computers in Society*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1983.

Tawwab, Nedra Glover. *Set Boundaries, Find Peace: A Guide to Reclaiming Yourself*. New York: Tarcher Perigree, 2021.

Thaler, Richard H. and Cass R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York: Penguin, 2009.

Thomas, Will. “Lessons from the Conti Leaks.” *BushidoToken Threat Intel*, April 17, 2022, https://blog.bushidotoken.net/2022/04/lessons-from-conti-leaks.html.

Toulmin, Stephen E. *The Uses of Argument, updated edition.* Cambridge: Cambridge University Press, 2003.

Williams, Emma J., Joanne Hinds and Adam N. Joinson. “Exploring Susceptibility to Phishing in the Workplace,” *International Journal of Human-Computer Studies* 120 (2018): 1–13. Available at: https://doi.org/10.1016/j.ijhcs.2018.06.004.

Worley, Meg. “The Rhetoric of Disruption: What Are We Doing Here?,” in *Disrupting the Digital Humanities*. Dorothy Kim and Jesse Stommel (eds.). Brooklyn, NY: Punctum Books, 2018: 61–78.

Yablonski, Jon. *Laws of UX: Using Psychology to Design Better Products and Services*. Sebastopol, CA: O’Reilly Media, 2020.

INDEX

BLOOMSBURY ACADEMIC
Bloomsbury Publishing Plc
50 Bedford Square, London, WC1B 3DP, UK
1385 Broadway, New York, NY 10018, USA
29 Earlsfort Terrace, Dublin 2, Ireland

BLOOMSBURY, BLOOMSBURY ACADEMIC and the Diana logo are trademarks of Bloomsbury Publishing Plc

First published in Great Britain 2024

This electronic edition published in 2024 by Bloomsbury Publishing Plc

Cover design by Rebecca Heselton
Cover image: rapapazzi/ Adobe Stock

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Mauro, Aaron, author.
Title: The language of cyber attacks : a rhetoric of deception / Aaron Mauro.
Description: London ; New York : Bloomsbury Academic, 2024. | Series: Bloomsbury studies in digital cultures ; vol 6 | Includes bibliographical references and index.
Identifiers: LCCN 2024005683 (print) | LCCN 2024005684 (ebook) | ISBN 9781350354678 (PB) | ISBN 9781350354685 (HB) | ISBN 9781350354692 (ePDF) | ISBN 9781350354708 (eBook)
Subjects: LCSH: Computer crimes. | Online manipulation. | Rhetoric and psychology. | Disinformation.
Classification: LCC HV6773 .M3833 2024 (print) | LCC HV6773 (ebook) | DDC 005.8/7014–dc23/eng/20240412
LC record available at https://lccn.loc.gov/2024005683
LC ebook record available at https://lccn.loc.gov/2024005684

ISBN: HB: 978-1-3503-5468-5

PB: 978-1-3503-5467-8
ePDF: 978-1-3503-5469-2
eBook: 978-1-3503-5470-8

Series: Bloomsbury Studies in Digital Cultures

Typeset by Integra Software Services Pvt. Ltd.