Mental Malware

Disinformation and the Erosion of Societal Integrity

Conference for Global Transformation 2022

Version 1.0

Author/SMEs:

Mike Manrod (disinformation SME)

Mary Manrod (Wisdom Conversation and Distinctions SME)

Contents

Contents 2

Introduction 3

Dawn of a New Association 4

Introducing Mental Malware 4

Security Flaws, in People and Systems 5

Addressing Flaws in Information Possessing 7

A Deeper Look: Issues and Errors w/ Human Information Processing 9

Conclusion 10

Sources 11

Introduction

In the (2005) Christmas Special of the science fiction series Dr. Who, in a state of outrage, The Doctor tells the fictional character Harriet Jones that he can bring down her government with a word(s) and after some argument he walks away, pausing to whisper in a reporter's ear, "Don't you think she looks tired?" This sentence starts a web of conversations that ends her career as Prime Minister. While fictional, this example underscores a concept that is well understood by all participating in the Conference for Global Transformation – that is, the tremendous power words have to alter the world around us. While many of us at this conference have made a lifelong study of how a nuanced understanding of language can create a clearing to bring forth new realms of possibility, there have been others across history who attempt to leverage the power of language (old model) to project power and manipulate. Networks of conversations that extend from Niccolò Machiavelli through Benito Mussolini to modern era dictators, form the basis of playbooks for coercive fear-based models for crafting narratives to subvert and control.

While disinformation has been around for millennia, the information age has enabled a 'big bang event' for false narratives, causing the reach to expand with unfathomable force, velocity, and reach. Unfortunately, the models that have been developed since the age of antiquity, and carefully refined over recent centuries, have now found vast new reach due to technical platforms, creating a threat to the very essence of truth and our collective orientation to "what is so" (see also Wilber, 2017).

It is well understood within this community that the most profound breakthroughs into new realms of thinking and acting, are often the product of having a profound experience that shakes us to the core, thus opening a clearing of a magnitude not typically experienced from anything but the most profound realizations. As such, it is connecting with this deep impact that may create a clearing large enough for new realms of thinking at the level of society. We face a societal inflection point that may well determine the context of the next episteme for humanity. Dystopia or Enlightenment, choose.

Dawn of a New Association

This entire journey started for me in 2015 with a flash of insight on a long flight back from Tel Aviv, after teaching a class to people who flew-in from all over the world, on the topic of malware/threat prevention. While my mind had mostly been oriented toward solutions to stop malware, the rise of disinformation in our society had also been on my mind. In particular, the seeming order to it caught my attention: the strange correlation in timing of key events, the nature of the division it seemed to cause, and a possible alignment of objectives between intrusion attempts into information systems (clearly observed) and apparent disinformation campaigns (more subtle, back then). Were these events part of a unified strategy?

On my return journey, after many hours of contemplating over the Atlantic (and at least a few glasses of wine), it dawned on me that the actual infection flows for malware and disinformation are practically identical. For malware and cyber intrusion, you start with a system that has a particular vulnerability (security flaw), followed by an attack exploiting this flaw with words (in programming languages), allowing the attackers ideas (code) to run on the host system – that is, it runs against the programming of the system and the will of the person who owns the system.

Introducing Mental Malware

If we invoke this as a new distinction, we start by saying that malware is really just a collection of "bad ideas" that give the attacker control, using the old model of communication (control/manipulate). Often these "bad ideas" come in strings of text, called shellcode. The most important distinction is that effective malware runs with a great deal of power on the host system (think of it as the technological equivalent shifting to the Father/Child field) – in cyber we call this privilege escalation. It is a phenomenon where the attacker ideas run with such elevated authority, the system thinks these ideas are its own programming (we would say, operating of our own free will on the human side). It is also important to note that for malware to run on a host, a precondition must exist, in the form of a security flaw that makes that system vulnerable to exploitation.

In cybersecurity, we use the term vulnerability to refer to a security flaw in a system, which may allow an attacker to take control and accomplish their objectives. For our purposes here, we will just refer to these as flaws, since the term vulnerability has other meanings in the domain of transformation. The core concept though is that a mistake or flaw in programming, may be used by a hacker to manipulate the way information is handled, producing outcomes not intended by the person who created the software – and, generally against the interests of those using the software.

Likewise, disinformation could be viewed as carefully crafted sets of ideas that exploit a flaw in human thinking to give an attacker the ability to run their ideas with elevated privileges on the host. In these cases – those hosts are people. As with information systems, people have flaws in cognitive processing that make them susceptible to having their thinking exploited. In fact, the closer we look at the commonalities in how malicious communication may impact people and information systems, the more the similarities become obvious. As such, I like to refer to disinformation that meets certain key criteria as **Mental Malware**.

Security Flaws, in People and Systems

Why are there security flaws? Why not just develop things that cannot be hacked? It is easy to imagine the plethora of programming flaws that may exist in your laptop – millions of lines of code in many applications, running on operating systems (e.g., Windows) with millions of lines of code, running on hardware with billions of transistors. If you are a fellow nerd, the calculator is already out, and you are getting unfathomably large numbers as you try to calculate the number of possible interactions these sets of relationships could denote.

Let's just say the number of possible relationships that can give rise to a security vulnerability are astronomical – and, finding and resolving all of them is a herculean undertaking at best. Now, imagine the tremendously greater potential of arrangements that could represent flaws in human information processing – Ray Kurzweil suggests Hierarchical Hidden Markov Models with 300 million pattern recognizers with

each holding 100 million neurons assembled into carefully organized hierarchies, just for the prefrontal cortex representing conscious decision making alone. And, there are many very important processes that are not even owned by the prefrontal cortex – moreover, this is before we introduce the complexity of the software / programming languages of people – spoken and written human languages. Finally, every person is a snowflake – whereas all copies of a given version of Microsoft Windows may be virtually the same, the programming of people follows an adaptive process that is nuanced and a little different each time it occurs. Like snowflakes, certain elements are always present (e.g., cold and precipitation or decisions in response to formative events); however, the random aspect makes each instance unique.

This incredible complexity and the nature of the formative processes themselves, give everyone some sort of processing flaws – and, with sufficient knowledge these can be manipulated with great effect (we could say it is effect via affect, for a little bit of fun with words). In a cyber-attack, a phase we refer to as reconnaissance (recon for short), is where information is collected to find the vulnerabilities and plan the attack. Similarly, in an attack using Mental Malware, the first step an attacker may follow is to learn about the demographics of a population being attacked and what cognitive processing flaws may be exploited. In her recent book *This is How They Tell Me the World Ends* Nicole Perlroth (2021) outlines disinformation campaigns launched by Russia against the United States. According to her research, these campaigns started with visits from spies in advance, followed by some small test runs to validate what Mental Malware exploits would be effective.

While the underlying complexity and fundamental nature of the processes that make up a human mind make each one unique, as we understand very clearly in the realm of transformational work, the nature of being human has some common elements. Some examples of common information processing flaws that we all know at CGT include how formative events may make up enduring aspects of a person's personality that seem automatic and cemented, even when it may result in maladaptive behaviors. Other examples may include acting inauthentically, even if it results in ways of being and actioning that rob us of joy and vitality. And, along this line, we must not forget the power of the payoff.

For disinformation campaigns, information processing flaws at the individual level are somewhat limiting for attackers unless they represent a common thread across certain demographics. More useful, are those that represent inherited conversations shared by broad groups of people. For example, if people in Arizona know something about the way California people are or vice versa, this is a widespread inherited conversation people are born into that can be exploited by a set of messaging with targeted delivery to that demographic (keeping it playful). As such, the spies indicated by Nicole Perlroth (2021) visited several key states with different collective conversations, such as California, Texas, and New York (among others). The information from these recon visits was then organized into a comprehensive plan, resulting in communication delivered via multiple channels for maximum effect. A key objective this research phase where collective conversations are mapped with specific flaws of target demographics, is to find out what messages will propagate on their own, once started.

In cyber security, if malware can spread automatically from system to system, we say that the malware is wormable. Worms are particularly devastating, because they can sometimes take on a life of their own once released, often gaining reach far beyond what the author of the malware intended. This concept is important, because most disinformation campaigns are very wormable, if they are thoughtfully crafted. Once a disinformation campaign picks up momentum, if it has a high level of believability for the target demographic and creates a sufficient emotional response, it really does take on a life of its own. Some early disinformation campaigns are no longer being actively targeted, yet they have gained greater adoption over time. As such, we have experienced an implosion of truth, where orientation to an objective reality collapsed in on itself for certain demographics that were targeted aggressively.

Addressing Flaws in Information Processing

If flaws in information processing are what make it possible for disinformation campaigns to work effectively, how can we address these flaws? As a starting point, there are common themes that can give rise to greater understanding and context – for both information systems and people. In the Landmark

Education context, we call these problem-solving algorithms **Distinctions**. For malware and intrusion, we have useful, although less mature frameworks such MITRE ATT&CK – representing a collection of cyber distinctions that help provide a deeper understanding of the mental/technical models of attack and defense. Some of the distinctions for fighting cyber intrusion become uniquely relevant when working to create new models for how to address the problems presented by disinformation.

There is a compelling overlap between what it takes to exploit information systems with malware and what is necessary to compromise human thinking with disinformation. The overlap continues as we consider the fact that it is impractical to discontinue taking-in and processing new information, in either modality – human thought or digital information processing. And, therein lies the conundrum – we must continue to take-in information, make decisions, and live our lives – and, we must rely upon a myriad of information systems to do the same. How do we keep consuming information when much of it is being strategically poisoned and weaponized, without making our mental processes sick by consumption of the cognitive cyanide or prohibitively ineffectual, by shunning all new inputs?

New paradigms and conversations shall be invented to bring power and context to the dawning episteme, we now call the Information Age. Of course, what history calls this new episteme, is entirely dependent upon what paradigm becomes predominant as the new context for being human. Assuming we are looking backward from a future where possibility wins the day and integrity is a cornerstone of society, what was the process we went through to arrive at such an extraordinary destination? The short and philosophical answer is that God only knows… the long and practical answer is that it should be the purpose and calling of many of us in this era, to strive toward solutions and answers.

New ways of being and thinking call for new and appropriate actions, which in turn can bring forth the dawn of a new human context. Simply put, when the invented conversations emerging from possibility are large; perhaps, the power of disinformation and the narratives of tyranny will become smaller. The future of the next several millennia hangs in the balance, with our best hope being the ability to drop another weight on the side of possibility – by saying something new. Or better yet, by saying many things

newly, repainting the world in a vibrant paint, using the broad spectrum of colors of possibility. As we have all experienced in our own lives, we are at a phase where humanity has a predictable almost-certain future – and, it is upon us to say something new.

It is upon us to raise our voice for truth and authenticity; to shout out for human connection and intimacy; to explode forth with enough love and compassion to envelope the wounds and misgivings of those who seek to destroy. As the esteemed Sandy Robins has indicated, we may accept and embrace what is already in existence, absorb it into our paradigm and then invent something new that pulls forth what exists into a newly created realm. Now is the time to accept what is so and speak <mindfully> about what is to come.

A Deeper Look: Issues and Errors w/ Human Information Processing

What is it that makes human beings so susceptible to disinformation? While we could speculate about a variety of factors, such as how Dunbar's Number orients our trust models to small groups and personal relationships and makes us distinctly unprepared for large social groups (e.g., 100k followers). Then there is how Solomon Asch showed that people may conform to the perspectives expressed by their peer group under certain conditions, there are even broader models of human decision making that should clarify observations. The body of research from the works of Kahneman and Tversky, provide a comprehensive narrative, outlining a series of fundamental flaws in human information processing, distinguished in the form of biases and heuristics.

There is also the dimension of how human attention shifts, based upon progressive development along a spectrum of maturity. Abraham Maslow drew distinctions related to the progressive nature of how levels of thought advance along the Hierarchy of Needs. A variety of researchers introducing Spiral Dynamics, which culminated in the tremendous work of Ken Wilber (1995, 1999) who introduced the multi-dimensional planes of waking up and growing up.

Finally, there is the line of thought that brings us to this conference, building upon the brilliant and revolutionary ideas of Werner Ehard, steadfast leadership of Harry Rosenberg and work of multitudes here at CGT. While many of the aforementioned geniuses that have focused on cognitive and social sciences have mostly focused on theory without application, we are all here because of a belief that theory alone is useless. What humanity needs to take the average level of thought to a tipping point (Gladwell, 2000) that brings forth a new realm of thinking and acting at the level of world, is a seamless alignment of both that will elevate what is good, true, and useful (Socrates, 399 BC). Socrates would give a pejorative glare upon the modern phenomenon of disinformation, since it is bad, false, and useless to the general utility of society. Side note – in this case, I mean bad in the objective and literal sense of people suffering and dying, not the philosophical application of bad we deal with in our work in transformation.

Conclusion

While much of the world is focused on the nuclear existential threat of Russia in the war on Ukraine (have another possible whitepaper on that topic) – and, a well-informed minority is focused on the related threat of cyberwar, there is another realm emerging that is quite neglected. What good does it do to have a dramatic superiority militarily and economically, if we collapse from within under the tremendous weight of [believed] lies, causing division and infighting?

It is time to say something new. The time has arrived to expand the conversation of completion, possibility, and enrollment. Now is the time to reach out, to connect, and to help reduce the number of cognitive information processing flaws that provide fertile ground for the seeds of disinformation to germinate. It is also the time to think of how to make strides toward upgrading the paradigm of what it means to be human, before other forces begin to downgrade humanity. It is a cliché of contested origin that there is a curse, "May you live in interesting times." Personally, I rage against that curse with the invented future that it is interesting times where all new paradigms are born. We live in interesting times, and we should seize the opportunity to elevate what it means to be human.

Sources

Davies, R. (2005). "Dr Who: The Christmas Invasion"

BBC One.

Gladwell, M. (2002). The Tipping Point: How Little Things Can Make a Big Difference.

Bay Back Books.

Kahneman, D. (2011). Thinking Fast and Slow

Farrar, Straus and Giroux.

Kurzweil, R. (2013). How to Create a Mind.

Penguin.

Machiavelli, N. (1469-1527). "The Prince."

(various publishers).

Perlroth, N. (2021). This is How They Tell Me the World Ends: The Cyber Weapons Arms Race.

Bloomsburry Publishing.

Wilber, K. (2017). Trump and a Post-Truth World.

Shambala.

Wilber, K. (1995). Sex, Ecology, Spirituality: The Spirit of Evolution.

Shambala.

Wilber, K. (2000). A Theory of Everything.

Shambala.