Global Transparency Initiative as a framework for increasing trust in cyberspace

Anastasiya Kazakova, Public Affairs Manager, Kaspersky (Anastasiya.Kazakova@kaspersky.com)

Abstract

Cyberspace has made a significant path from something futuristic to an everyday reality that has brought with it both opportunities and risks, including cyber threats and cyber-enabled crime, privacy concerns, digital divides across communities. These challenges, as a consequence of a crisis of trust in cyberspace, showed how vulnerable we are and how our common wellbeing now depends on solutions we may take to ensure that cyberspace remains open, stable and secure. This paper seeks to find a solution to fix the crisis of trust and to enhance confidence in ICTs and digital technologies. The analysis presented in the paper relies on the view that there are three dimensions of the mistrust in cyberspace where each of them includes different actors and concerns. As a result, the paper suggests the Global Transparency Initiative (GTI) - a solution based on three years of practical experience of Kaspersky to enhance confidence in ICTs, promote transparency in cyberspace, and engage a wider community to work together on building cyberspace a more trusted place.

Introduction

The digital transformation of certain sectors of economies and society is transforming into a digital transformation of every aspect of our lives, and this fact brings both benefits and drawbacks. The use of information communication technologies (ICTs) creates more automation and, as a result, lower costs, higher speeds, better functionality, and greater comfort and opportunities. However, further far-reaching expansion of ICTs into social and economic areas opens up new risks and issues, including cyberthreats and cyber-enabled crime, privacy concerns, and digital divides across communities.

As cyber-insecurity has become a growing problem worldwide (Ruhl et al. 2020) and with further awareness of possible negative implications for economies, society and world security from digitalization, different solutions have been created to address these issues: from multilateral processes and dialogues (at global, regional and national levels), to new regulatory and legislative practices to address harmful use of ICTs (such as measures to ensure critical infrastructure protection, incident notification and incident reporting, establishing computer emergency response teams (CERTs) and competent national authorities for cybersecurity, developing threat information sharing across industries and sectors, etc.). Multilateral cyber-initiatives are abundant with different purposes and targets in mind, but their number is increasing together with a growing number of different aspects in cyberspace to ensure its security and stability. Some of the initiatives are aimed at addressing development of norms; others – capacity-building or confidence-building.

The most prominent global efforts on ensuring cybersecurity and peaceful use of ICTs take place under the auspices of the UN General Assembly's First Committee. Initially, it was the UN Group of Governmental Experts (GGE) on Developments in the Field of Information and Telecommunications in the Context of International Security. The UN GGE produced three consensus reports (2010, 2013, and 2015) and, as a key outcome, 11 non-binding norms on responsible state behavior, confidence-building measures (CBMs), capacity-building measures to ensure international peace and security, and a consensus agreement that international law applies in cyberspace. The attempts to go into detail on how international law applies led to a failure in 2017 to produce a consensus report and collapse of the UN GGE. While diplomatic efforts failed to increase the security and stability of cyberspace, the world continued experiencing further cyber-incidents in 2018 such as WannaCry and NotPetya that were wake-up calls indicating that cyberattacks are becoming more targeted and disruptive, while the threat landscape and threat actors are becoming more sophisticated, and sometimes – political and strategic.

These events indicate a further gradual erosion of trust in cyberspace and an urgent need to fix this to ensure international security and peace. After a two-year break, new attempts are currently being made within two parallel processes: the next round of the UN Group of Governmental Experts (GGE), which traditionally

comprises a limited number of states (now – 25), and the Open Ended Working Group (OEWG), which, for the first time, accepts participation of any interested state. Other international organizations that seek to ensure the use of ICTs create more benefits than threats include the Shanghai Cooperation Organization, the G7, and the G20.

Despite these efforts, it would not be correct to assume that cyberspace is becoming a more trusted place. While consultations on norms for responsible behavior in cyberspace continue, the COVID-19 pandemic underlined further existing challenges to society, economies and world security associated with the use of ICTs and digital technologies. Looking at the threat landscape results of Q1 2020 (Advanced threat predictions, 2020), it is evident that the COVID-19 topic has been actively used by cybercriminals: from phishing, scamming and ransomware, to more targeted, state-sponsored attacks putting healthcare sectors, healthcare organizations (e.g., the attack on the World Health Organization (Reuters, 2020), and medical facilities that conduct innovative research under greater pressure and threat.

The trust issue in cyberspace is further aggravated by a growing interest in and development of emerging technologies, including artificial intelligence, 5G, the Internet of Things (IoT), quantum computing, etc. While the Pre-Draft report (both versions) of the UN OEWG expresses views of states on the importance of a technology-neutral approach, and that "it is the misuse of technologies, not the technologies themselves" (OEWG Second Pre-Draft, 2020) that causes the concern, the key question remains open: how to ensure that ICTs and digital technologies are trustworthy, human-rights based, safe and sustainable and promote peace (Roadmap for Digital Cooperation 2020), and are not used for malicious or military use?

The research paper is therefore guided by the overarching question, "What solution serves as a measure enhancing confidence in ICTs and digital technologies and therefore fixes a crisis of trust in cyberspace?". As we see a trust issue reflected in three dimensions – mistrust among states; mistrust between state and non-state actors; and mistrust of users to ICTs and digital technologies – the research paper is also guided by three supporting questions where each question focuses on each dimension and all of them provide a more granular view on the main overarching question.

As a private sector entity, we have experienced all three dimensions of the trust issue in our work, and early on realized that cybersecurity as the industry critically relies on trust and requires it. A parallel could be done with the work of doctors: without trust, the mission to cure people is hardly achievable. Therefore, to fix a trust issue and address the main research question we applied constructivism as a theory to the development of our practical steps.

Constructivism claims that the main vehicles for transformation are not shifts in the balance of power, but ideational shifts - "shared ideas, expectations and beliefs about appropriate behavior that give the world structure, order, and stability"(Katzenstein et al. 1999). Having been inspired by these constructivist conclusions, our first practical steps led us to acknowledge the key finding that transparency should be widely adopted by both state and non-state actors as a fundamental norm for building trust in cyberspace and keeping it stable, secure and open. We went further and framed a transparency norm into a set of practical measures we called the Global Transparency Initiative (GTI). The GTI is a conceptual framework and set of actions to serve as measures enhancing confidence and trust in ICTs. This framework includes (1) data care measures; (2) transparency centers for executive briefing and source code reviews; (3) third-party assessments to confirm the trustworthiness of engineering practices; (4) a vulnerability management program and ethical principles for responsible vulnerability disclosure; and (5) a cyber capacity building program. The paper, therefore, presents the results of practical project work and experience of Kaspersky.

This paper continues with further analysis outlined in some works before and provides findings based on three years of practical experience. With regard to previous works, this paper further elaborates on the need to establish norms for good behavior in cyberspace, including transparency and confidence-building measures (Hitchens & Gallagher, 2019; Sabbah, 2018). The paper also provides more evidence supporting the view of researchers about the role of private entities and non-state actors in shaping the behavioral standards that new regulation needs to take into account (Hurel & Lobato, 2018) and to create a more favorable stable regulatory environment (Pawlak & Biersteker, 2019). The paper also addresses the need to develop evidentiary standards and norms that will underpin the future resilience of ICTs, and the negotiation and establishment of new norms and institutions that should govern the use and misuse of ICTs (Dunn Cavelty & Wenger, 2020).

The paper will begin by analyzing the above-mentioned three dimensions of mistrust in cyberspace, thus supporting the importance of the main overarching research question. Secondly, this paper will elaborate

on conceptualization of confidence-building measures. In order to do so, the conceptualization will rely on a short discussion of the concept of CBMs in academic literature as well as use of the term in adopted documents on multilateral global processes such as the UN GGE, UN OEWG and OSCE. Finally, an answer to the research question will be provided based on three years of project work by a private sector entity – a leading global cybersecurity company operating in many countries and regions and thus facing the trust issue in different territories.

Three dimensions of mistrust in cyberspace

In this paper we support the view on the existing mistrust in cyberspace and the fact this creates risks to international security and peace and, therefore, should be fixed. To be more precise, we see a crisis of trust in three dimensions: mistrust among states; mistrust between state and non-state actors; and mistrust of users to ICTs and digital technologies which leads to growing inconfidence of users in cyberspace. We will analyze each dimension in detail below and provide with arguments that support our views.

The increasing mistrust between states is accepted due to several factors. First, publicly available reports on cyber threat landscape indicate a further increase in cyber operations, including state-sponsored attacks. At Kaspersky, we investigate and monitor more than 300 Advanced Persistent Threat (APT) operations (Advanced threat predictions, 2020) that to the groups - often state sponsored or well-funded in other ways - that are responsible for launching such precision attacks. The APT trends report for Q1 2020 reveals that threats grow in sophistication and become more targeted, diversifying under the influence of external political factors, including tensions around trade routes between Asia and Europe. These attacks include a growth in political espionage as governments seek to secure their interests at home and abroad. The COVID-19 agenda has been also widely exploited (APT trends report, 2020) by threat actors targeting medical institutions that conduct innovative research - at times like this it is easy to assume that cures or tests relating to coronavirus could be a target priority for intelligence organizations of an affected countries. The investigation of the attack at the WHO shows that the malicious web infrastructure used in this particular attack has been also used to target other healthcare organisations in the same weeks. The clear and very precise modus operandi could indicate possible motivations behind threat actors and in case of attacks at medical research institutions there is a probability of state-sponsored activities.

The growing mistrust between states in cyberspace is also understood due to the fact that cyberspace has been accepted as a new field for interstate military conflict in some states' doctrines and public documents. In particular, the Heads of State and Government participating in the meeting of the North Atlantic Council in Brussels on 11-12 July 2018 issued the 'Brussels Summit Declaration' where cyberspace is called as a 'domain of operations' and countries stated that they are 'determined to employ the full range of capabilities, including cyber, to deter, defend against, and to counter the full spectrum of cyber threats, including those conducted as part of a hybrid campaign' (NATO, 2018).

Rapidly developing cyber-offense military capabilities by some countries (Dorfman & Deppisch, 2019) increase risks to the international security in cyberspace, while the fact that not all countries have resources for developing these capabilities and not all countries are transparent about this (only few countries have publicly accepted that they develop cyber military capabilities and tools) indicate increasing threats and make cyberspace a less secure place. There is no consensus either among states, but some of them, in particular, Australia, the United Kingdom, France and Denmark stressed at the first substantive session of the UN OEWG in September 2019 that they recognize the legitimate right of countries to develop offensive cyber capabilities (Stadnik, 2019). The representative of Australia has also stated that "Australia is looking to militarize cyberspace because many countries have developed, and many more are in the process of developing these capabilities" (Stigherrian, 2019).

The situation also seems to pose additional risks to the security and stability in cyberspace due to different views of states on the principle of sovereignty in cyberspace and on what constitutes a cyber military operation. Though only few states have publicly shared their views on these aspects, clear differences in their approaches are already seen. The UK Attorney General Jeremy Wright in his 23 May 2018 speech (UK Government, 2018) states that certain types of cyber operations - which fall below the use-of-force threshold and do not constitute an intervention into the internal affairs of another state - are not prohibited under international law. France and the Netherlands have opposite views supporting the existence of a rule of territorial sovereignty in cyberspace. In particular, the Netherlands state that "respect for the sovereignty of other countries is an obligation in its own right, the violation of which may in turn constitute an

internationally wrongful act" (Netherlands Government, 2019). According to France, "any cyber attack against French digital systems or any effects produced on French territory by digital means by a State organ [or otherwise attributable to a State] constitutes a breach of sovereignty" (French Government, 2019).

Finally, a lack of communication and established institutional dialogue among states also indicate the crisis of trust between states in cyberspace. Only for the first time, suggestions for establishing a regular institutional dialogue and national points of contact (PoC) for diplomatic, policy, legal and technical exchanges, as well as incident reporting and response appeared in the both Pre-draft versions of the UN OEWG (OEWG Initial and Second Pre-Drafts, 2020). However, following the views of some national delegations that have been submitted to the first Pre-draft, it is clear that some states do not support these suggestions, while view the idea of a regular institutional dialogue as "premature"(US Comments, 2020).

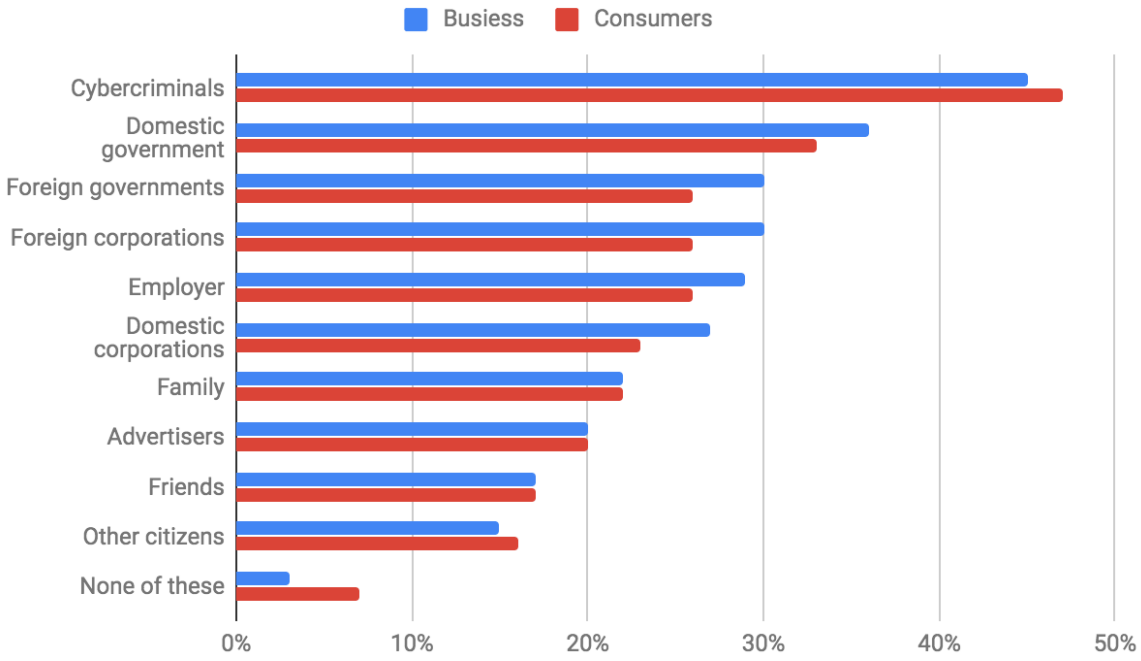
The second dimension of a crisis of trust in cyberspace is expressed through a lack of trust between some state actors and private sector entities. Growing in complexity and sophistication, modern technologies and algorithms raise the challenge of trustworthiness and trust and make decisions-makers to think about the necessity to ensure that algorithms and ICTs, including emerging technologies, serve for the good and do not pose a risk to social and national security. The further dependence on ICTs - which was amplified during the pandemic crisis - and further dependence on private sector's technologies for both critical and non-critical functions pose significant challenges of understanding, managing and security ICTs. Vulnerabilities in ICTs and applications could be created simply because vulnerabilities are inevitable due to the fact any human work may have flaws. However, exploitation of these vulnerabilities by criminal actors in products of even trusted entities may result in the destruction or disabling of critical functions and create significant financial, physical, and political harm (Kuehn & McConnell, 2020).

This mistrust to private sectors' ICTs is further created by the fact that some vulnerabilities could be secretly created as hidden functionalities or backdoors in software or hardware parts or could be purchased as zero-day exploits for covert intelligence or law enforcement needs (Joyce, 2016). In the context of growing political tensions between some states, governments may fear of creating or exploiting vulnerabilities ICT supply chains by adversary foreign government. As a result, the fact of mistrust to private sector entities and their ICT products due to growing complexity, sophistication, global distribution and growing reliance and dependency on them lead to bans and restrictions of technologies or vendors based based upon alleged threats to national security from foreign ICTs (Kuehn & McConnell, 2020). These measures as a consequence of a crisis of trust leads to fragmentation in cyberspace by hampering innovation and endangering long-term global economic development and growth.

The third and final dimension of the crisis of trust in cyberspace is reflected through users' decreasing confidence in cyberspace based on privacy-related concerns. The fact that ICTs could be used as tools for influence and manipulation (OEWG, 2020) and associated with this fact incidents (such as Cambridge Analytica scandal in 2018) brought many sensitive issues to the fore. They are protection of personal data, privacy, location of data storage and data processing, availability of data. Again with technologies and applications, growing in complexity and sophistication, users have started thinking about how the technologies and ICTs work and whether they can be trusted. This leads to growing feeling among users that they do not have enough information and understanding to manage and use these technologies and ICTs safely as well as that they lack full control over their devices and, therefore, technology (Auxier et al., 2019).

In 2018, the independent study, commissioned by Kaspersky with the data analysis undertaken by Applied Marketing Research, surveyed 600 mid-sized companies with IT security professionals as well as 6000 consumers with security software installed on their devices, split equally across France, Germany, Italy, Spain, the UK and the US (Kaspersky, 2018).. The study revealed that people and organizations do not fully trust anyone when it comes to their data, including personally identifiable data, available online. Other results include:

- 65%—78% of businesses and 54%—80% of consumers told that they are worried about the provider accessing their private data, opinions, location or online behavior and sharing this information with foreign entities;
- Both business and consumers expressed their desire to protect their online data from their own national government (consumers 33%, businesses 36%), with keeping it out of the hands of foreign governments and companies coming next (consumers 26%, business 30%).



The share of businesses and consumers surveyed who worry about each of these groups accessing their personal information online - data analysis: Applied Marketing Research Inc. for Kaspersky, 2018

The data presented above thus provides evidence that technologies and software are considered as a black box for many users - both businesses and consumers. Due to growing ICTs' and their algorithms' complexity, users do not know how they work, what is going on inside, how data is collected and how data is stored and whether strong safeguards are applied for data protection. This 'black box' nature of modern technologies intensifies growing mistrust to ICTs and companies which produce them.

We have provided arguments that support our view on the crisis of trust in cyberspace which is reflected through three dimensions. One of the prominent efforts to address the lack of trust is through confidence-building measures (CBMs), which the papers analyzes in the next section to provide the brief overview on what is perceived under CBMs in the academic literature and in the modern political context with regard to cyberspace.

Confidence-building measures: conceptualization

Confidence-building measures (CBMs) in international relations are perceived as processes central to preventing unwanted and accidental conflict, and helping to lead participants to satisfactory resolutions thereof. CBMs are important for building trust among parties through correcting misperceptions, breaking down barriers to communication, and facilitating the achievement of common and shared ideas. CBMs have been acknowledged and advocated by the UN as a means for dispelling mistrust and stabilizing tensions between parties. The concept of CBMs was born during the Cold War to increase transparency among countries in the context of armed conflict and thus reduce the risk of nuclear attacks. In 1981, the UN published a comprehensive study on CBMs to provide guidelines to governments for developing and implementing CBMs for greater peace and security(UN Study, 1981). However, the first attempts to conceptualize CBMs had been taken earlier within the Vienna Talks on the Mutual Reduction of Forces and Armaments and Associated Measures in Central Europe, which led to establishing the Conference on Security and Cooperation in Europe (CSCE), ongoing since the adoption of the 1975 Helsinki Act. The Stockholm document, adopted in September 1986 as a result of several rounds of meetings of the CSCE, produced the first security agreement among the 35 participating states and established that they adopt politically binding, and verifiable confidence-building measures.

In the military context, CBMs started to include non-military instruments (political, economic, societal, cultural) to build trust among parties. The Organization for Security and Co-operation in Europe (OSCE)

is considered to play a key role in conceptualizing CBMs and producing guidelines for the design and implementation of CBMs.

In the academic literature, CBMs are analyzed in terms of their objectives, types, possible actors involved, and practical lessons for their development. Several types of actors are outlined: negotiators, decision-makers, and wider constituencies (Mason & Siegfried, 2013). Several types of CBMs are also identified: military or security; political; economic and environmental; humanitarian, social, and cultural CBMs. CBMs and their development can also be bottom-up or top-down (Kemp, 2011), but they can be most effective when designed symmetrically and when the respective agreements are adopted by all the parties at the same time (Mason & Siegfried, 2013).

With regard to cyberspace, CBMs usually presume efforts for transparency, cooperation, and stability (Ziolkowski, 2013), and the difference among them lies in the target of these measures and expected outcomes. Transparency measures, for instance, aim to enhance mutual understanding about military capabilities and activities that states have. Cooperation presumes states' willingness to engage in collaborative efforts such as joint exercises, exchange of information, experts, and development of consensus-based terminology and definitions. Stability measures are focused on enhancing predictability and managing parties' expectations around military activities.

Another classification of CBMs that can be established for building trust in cyberspace includes measures on collaboration, crisis management, restraint, and engagement (Healey et al., 2014). Collaboration aims at building best practices and cooperative efforts together. Crisis management focuses on establishing communication channels for information sharing and dialogue. Restraint implies agreements between states aimed at preventing escalation of crises or incidents in cyberspace. Finally, engagement presumes the use of non-state organizations to establish norms and standards, including technical guidelines for the security of cyberspace.

In the modern multilateral political context, CBMs have been adopted by the UN GGE and OEWG processes as a pillar of the framework on responsible behavior in cyberspace. While the UN OEWG currently discusses its pre-drafts and refers to the UN GGE reports with no chance yet of producing its own additional CBMs, the UN GGE managed to produce non-binding voluntary CBMs in its 2013 and 2015 reports (UN GGE, 2013 & 2015). Among them, the GGE recommends that states carry out the following:

- identify appropriate points of contact at the policy, technical, legal and diplomatic levels to address ICT-related requests;
- develop and support mechanisms and processes for bilateral, regional, subregional and multilateral consultations;
- encourage, on a voluntary basis, transparency at the bilateral, subregional, regional and multilateral levels, including voluntary sharing of national views and information on national and transnational threats, vulnerabilities, best practices, CBMs developed in other forums, strategies and policies relevant to ICT security;
- share, on a voluntary basis, national views of categories of infrastructure that states consider critical, as well as national efforts to protect them – including sharing information on national laws and policies for the protection of data and ICT-enabled infrastructure;
- establish a repository of national laws and policies for the protection of data and ICT-enabled infrastructure;
- adopt voluntary national arrangements to classify ICT incidents in terms of scale and seriousness of incidents to facilitate the exchange of information on incidents;
- strengthen cooperation between relevant agencies to address ICT security incidents, and develop additional technical, legal and diplomatic mechanisms;
- establish a national computer emergency response team (CERT) and/or cybersecurity incident response team, as well as expand and support cooperative practices in such teams, such as information exchange about vulnerabilities, attack patterns, and best practices for mitigating attacks;

- cooperate, in a manner consistent with national and international law, with requests from other states in investigating ICT-related crime or the use of ICTs for terrorist purposes.

The OSCE took a similar approach, and first adopted, in 2013, 11 CBMs on information sharing, national legislation, and the development of contact points (OSCE, 2013). In 2016, five additional CBMs were indicated, including: national reporting of vulnerabilities, cooperation on protection of national and transnational critical infrastructure, and the development of protected channels of communication for preventing risks of conflict stemming from the use of ICTs (OSCE, 2016). It should be noted that the OSCE, as a regional body, provides more practical steps to build trust among states and enhance cooperation in cyberspace in comparison to the UN GGE, which does not provide clear guidelines on implementation of CBMs, but rather outlines directions for work.

Analyzing the existing efforts made both at the UN and OSCE, it is evident that a large among of work has been already done – key issues such as information sharing, enhancing communication, protecting critical infrastructure, reducing risk of conflict due to cooperating on threats and incident response have been addressed equally at both forums. However, at the same time, we observe that in both cases CBMs target inter-state relations only, and do not clearly address cooperation with non-state actors – especially when it comes to critical infrastructure protection, where in many states such infrastructure is owned, managed or operated by the private sector (OEWG Initial Pre-Draft, 2020).

An explanation of this could be the fact that armed conflicts are a matter of inter-state relations only, and since CBMs historically appeared as instruments of de-escalation and building trust in the event of armed conflicts, they continue addressing inter-state cooperation and state behavior mostly. We see a serious drawback in this approach as that which could not fully address challenges of the modern context – meaning cyberspace, where many activities are shaped and driven by non-state actors and private sector entities, in particular. What is more, due to this limitation, the existing CBMs, as outlined in the 2015 GGE report or OSCE frameworks, seem incapable of addressing the crisis of trust and its three dimensions as being discussed in this paper.

This view is supported by recommendations in the academic literature; in particular, by the belief in the necessity of inviting input and increasing participation of other types of stakeholders to transform CBMs as principles, plus transforming high-level political recommendations into concrete actions and arrangements (Hitchens & Gallagher, 2019). Though supporting the existing consensus-based agreements on CBMs, this paper aims nonetheless to fill the gap and thus find a solution – a confidence-building measure – that would directly address the key research issue: the crisis of trust in cyberspace and its three dimensions.

Global Transparency Initiative: a framework to build trust in cyberspace

Having acknowledged all three dimensions of the crisis of trust in cyberspace and, at the same time, having been affected by them (Zetter, 2019), we at Kaspersky started thinking of a possible solution to rebuild trust in the industry as we soon realized that other companies and actors might be affected by the crisis of trust as well. Our understanding was that the solution should significantly impact the industry in which we operate and create a new mindset – a new approach, since cyberspace as it existed in 2017 (the date when we started working on a possible solution) was largely non-transparent, and that lack of transparency aggravated the crisis. As discussed above, a lack of transparency was observed among state actors operating in cyberspace, including the rationale behind their actions, motivations and intentions (the first dimension); over technologies and algorithms both corporate and consumer customers use (primarily, questions of logic, functionality, and data processing applied in these technologies and algorithms).

As a private sector entity and global cybersecurity company we clearly realized that it is beyond our capacity and expertise to make any impact on the first dimension (mistrust between state actors in cyberspace) and enhance trust and transparency on this track. Our actions as well as the actions of any other non-state actor may only have an indirect effect raising awareness of state actors only. Being powerless on this track, we believed that we could directly improve trust and transparency on other two tracks (dimensions).

Our understanding was that the main vehicle for most any transformation are ideational shifts – “shared ideas, expectations and beliefs about appropriate behavior that give the world structure, order, and

stability". This was further supported by the view that "people act toward an object, including other actors, on the basis of the meanings that the objects have for them" and "meaning arises out of social interaction" (Wendt, 1994). Relying on these views, we focused on changing the meaning that actors might have about technologies, ICTs, cybersecurity products and, more broadly, about cyberspace. The process of change implied efforts to reinterpret or frame this meaning so the new frame would need to resonate with broader public understanding and to be adopted as new ways "of talking about it" (Finnemore & Sikkink, 1998). In constructing these new frames, we were about to acknowledge the existing meaning to create alternative versions of "both appropriateness and interest" (Finnemore & Sikkink, 1998).

In search of a solution to the research question, we therefore began communicating with both corporate and consumer customers to try and gain a comprehensive understanding of the *meanings* they understood. A series of dialogues with regulatory bodies in Europe, large enterprise customers, and partners allowed us to realize two fundamental concerns that all those actors had. These included concerns about: (1) trustworthiness of ICT products, applications (in particular, fears about possible vulnerabilities and hidden functionalities that, if exploited or coerced by threat actors, might lead to significant harm and unwanted consequences); and (2) data management-related processes (in particular, greater clarity over data location, data safeguards and data protection measures applied were sought). As an outcome, we faced the assumption based on our practical experience which states that "if you do not understand the technology, you cannot trust it". Having realized the problem, our task was then to change the rule and create a new rule: "trust, but verify – applying a set of clear evidence-based measures", which would, first of all, address the two fundamental concerns mentioned above. The new *meaning* should therefore be as follows: "we trust because of greater transparency and clarity about what is going on in cyberspace, how ICTs and technologies work, and what intentions and motivations both state and non-state actors follow in cyberspace".

As a solution serving as a measure enhancing confidence in ICTs and digital technologies and building trust in cyberspace, we developed the Global Transparency Initiative (GTI) – a framework and set of clear practical measures to enhance transparency in cyberspace and trustworthiness of ICTs and technologies (Kaspersky, 2017). The GTI includes:

1. Data care measures: relocation of data processing and data storage to Switzerland, a country with a long history of neutrality, today – good ICT infrastructure and connectivity, and a robust approach to data protection regulation.
2. Transparency Centers with a specially designed three-layer approach to executive briefings and external examinations of software development and business processes as well as source code reviews (Kaspersky, 2018). At the Transparency Centers in Zurich and Madrid we also provide access to the source code compilation process to compile a new product out of source code and then compare the product with publicly available modules. This procedure allows to ensure the absence of any hidden functionality in the source code.
3. Third-party assessments (security audits) to confirm the security and reliability of engineering practices and data services applied to the functionality of ICT products. At Kaspersky, we specifically chose (1) SOC 2 reporting for confirming the security of the release and development process of security updates (Kaspersky, 2019); and (2) ISO 27001 certification for data services (Kaspersky, 2020).
4. Vulnerability Management Program and Ethical Principles for Responsible Vulnerability Disclosure (RVD): this program is implemented through launching our Bug Bounty Program with awards to security researchers for finding the most critical security flaws in products (a clear scope is provided beforehand). The Ethical Principles for RVD provide clarity and transparency on vulnerability handling and vulnerability mitigation as well as communicate the company's approach to RVD in case of multi-party coordination (Ethical Principles, 2020).
5. Cyber Capacity Building Program: dedicated product security training that aims to help government organizations, academia and other companies develop the necessary skills and mechanisms against supply chain risks (Kaspersky, 2020). The idea behind the Program was born after several visits of our customers and partners to Kaspersky Transparency Centers. Only a small percentage of those visitors had the necessary technical background to benefit from the full range of options provided at the Transparency Centers. Our Program is designed therefore with a

train-the-trainer approach to help different actors develop necessary knowledge for assessing product security and trustworthiness of ICT products, and this knowledge can be passed and shared further within the ICT community, thus enhancing the security of the broader ICT ecosystem.

The first measure was designed to address the concern over the data management-related processes, while measures such as Transparency Centers, third-party assessments and the Vulnerability Management Program were aimed at enhancing product security and, at the same time, providing clear measures to our customers and partners to verify and assess the level of security.

More than three years since the project began, we can state that the GTI as a measure for enhancing confidence in technologies and for building trust in cyberspace has started demonstrating its effectiveness. For instance, our Transparency Centers helped us build a transparent and clear channel for communication with our customers, partners and regulatory bodies. Among other actors in the ICT community, the GTI has also been widely supported: in particular, the initiative has been flagged as supporting principle 6 of the Paris Call for Trust and Security in Cyberspace (Paris Call, 2019). The GTI has also been recognized by some experts as a "commitment to rules, norms and principles" for security in cyberspace (CFR, 2020) as well as a framework that "serves as a de facto downstream supply chain assurance vehicle, allowing the company to demonstrate the absence of hidden functions in its products to its customers and regulators in national markets" (Demidov & Persi Paoli, 2020).

Conclusion

Cyberspace has come a long way from being something esoteric and futuristic to an everyday reality that has brought with it both opportunities and risks – including cyberthreats and cyber-enabled crime, privacy concerns, and digital divides across communities. These challenges, as a consequence of a crisis of trust in cyberspace, have shown how vulnerable we are and how our common wellbeing now depends on solutions we may take to ensure that cyberspace remains open, stable and secure.

With cyber having started out as something computer and technology-centric, then, as one more avenue for military conflict and as a tool for influence and manipulation, now is the time for making everything with the 'cyber' prefix – in particular, cyberspace and cybersecurity – trusted concepts, where trust is created and maintained through transparency of algorithms and ICT solutions, and transparency of intentions and actions of both state and non-state actors in cyberspace.

In this paper, we sought to explore a possible solution to enhancing confidence in ICTs and digital technologies as well as building trust in cyberspace by addressing the dimensions of the crisis of trust that the paper discusses. As a private sector entity that faced all of three dimensions of mistrust in cyberspace, we first started analyzing the concept of confidence-building measures to choose some among the already existing CBMs. The search was focused on analyzing the concept of CBMs in the academic literature as well as the use of the term CBM in adopted documents on multilateral global processes such as the UN GGE, UN OEWG and OSCE.

This work allowed us soon to realize that all existing CBMs for cyberspace did not presume participation from non-state actors, and mainly addressed inter-state relations, though cyberspace is accepted as comprising of various actors – both state and non-state, including the private sector, technical community and civil society. Therefore, we started working on the development of our own CBM, which would engage a wider community to participate and contribute to enhancing trust in cyberspace.

In line with the constructivist paradigm, we internally outlined a course of actions and determined that the first step should include collecting evidence for a comprehensive understanding of the crisis of trust. This understanding would then allow us to develop a granular approach to the problem: to realize causes, associated factors that either aggravate or improve the problem, and hear what possible solutions could be theoretically satisfying. As a result, we identified two fundamental concerns that undermine trust and security in cyberspace, and these concerns were aligned with our view of three dimensions of the crisis of trust. It led to the development of the Global Transparency Initiative (GTI) - a framework and set of clear and practical measures aimed at fixing mistrust in cyberspace.

Though the GTI is a result of the sole experience of Kaspersky, we nonetheless hope to contribute to the expansion of scope in the debate on norms and CBMs in cyberspace with our practical three-year project

work. There is extensive literature on norm development as a process of inter-state relations, and in this light this paper continues with further analysis and contributes with findings from the private sector. Future research could further aim at exploring the spectrum of other non-state actors that work on related trust issues. The interaction of state and non-state actors, the exchange of views and norms, CBMs developed by both sides, as well as efforts to work together for trust and security in cyberspace should also be considered for future research.

Acknowledgements

The paper is based on the three-year project team work of Kaspersky colleagues - Anton Shingarev and Igor Kumagin as well as other colleagues who were among authors of the very idea of the Global Transparency Initiative as a framework and set of practical measures enhancing confidence in and trust in modern technologies and broadly cyberspace. This result would not be possible without their ideas, insights, and committed work.

References

1. Auxier, Brooke, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. "Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information". Pew Research Center report, November 15, 2019. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

2. CFR. 2020. "From Multilateral to Multistakeholder? New Developments in UN Processes on Cybersecurity." Council on Foreign Relations, January 27, 2020. https://www.cfr.org/blog/multilateral-multistakeholder-new-developments-un-processes-cybersecurity.

3. Demidov, Oleg and Giacomo Persi Paoli. 2020. "Supply Chain Security in the Cyber Age: Sector Trends, Current Threats and Multi-Stakeholder Responses." United Nations Institute for Disarmament Research (UNIDIR), 2020.https://unidir.org/publication/supply-chain-security-cyber-age-sector-trends-current-threats-and-multi-stakeholder.

4. Dorfman, Zach and Breanne Deppisch. 2019. "The Rise of the Rest: Maturing Cyber Threats Beyond the Big Four". The ASPEN Institute Cyber and Technology Program. November. https://www.aspeninstitute.org/programs/cybersecurity-technology-program/threat-assessment-2019/.

5. Dunn Cavelty, Myriam and Andreas Wenger. 2020. Cyber security meets security politics: Complex technology, fragmented politics, and networked science. Contemporary Security Policy, 41:1, 5-32, DOI:10.1080/13523260.2019.1678855.

6. Ethical Principles for RVD. 2020. "Ethical Principles for Responsible Vulnerability Disclosure." Kaspersky website, May 18, 2020. https://www.kaspersky.com/blog/vulnerability-disclosure-ethics/35581/.

7. Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." International Organization. Vol.52, No.4 (Autumn 1998), (pp. 887-917). https://www.jstor.org/stable/2601361.

8. French Government. 2019. "International Law Applied to Operations in Cyberspace". Ministry of the Armies, February 12, 2018. https://www.defense.gouv.fr/content/download/567648/9770527/file/international+law+applied+to+operations+in+cyberspace.pdf.

9. "Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security." A/68/98*. General Assembly of the United Nations. June 24, 2013. http://www.un.org/ga/search/view_doc.asp?symbol=A/68/98.

10. "Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security." A/70/174*. General Assembly of the United Nations. July 22, 2015. http://www.un.org/ga/search/view_doc.asp?symbol=A/70/174.

11. "Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security." A/RES/72/266. General Assembly of the United Nations. January 2, 2019. https://www.un.org/disarmament/group-of-governmental-experts/#:~:text=In%20GA%20resolution%2073%2F266,the%20General%20Assembly%20in%202021.

12. Healey, Jason, John C. Mallery, Klara Tothova Jordan, and Nathaniel V. Youd. 2014."Confidence-Building Measures in Cyberspace: A Multistakeholder Apporach for Stability and Security." Atlantic Council of the United State, November, 2014. https://www.files.ethz.ch/isn/185487/Confidence-Building_Measures_in_Cyberspace.pdf.

13. Hitchens, Theresa, and Nancy W. Gallagher. 2019. Building confidence in the cyberspace: a path to multilateral progress. Journal of Cyber Policy, 4:1, 4-21, DOI:10.1080/23738871.2019.1599032.

14. Hurel, Louise M. and Luisa Lobato. 2018. Unpacking Cybernorms: Private Companies as Norms Entrepreneurs. Journal of Cyber Policy, 4:1, 4-21. DOI:10.1080/23738871.2019.1599032.

15. Joyce, Rob. 2016, "Disrupting Nation State Hackers." USENIX Enigma, San Francisco, January 27, 2016. https://www.usenix.org/node/194636.

16. Kaspersky. 2017. "Global Transparency Initiative." Kaspersky website. https://www.kaspersky.com/transparency-center.

17. Kaspersky. 2018. "The boundaries of trust: Privacy and protection in cyberspace". Kaspersky, November 12, 2018. https://www.kaspersky.com/blog/the-boundaries-of-trust/

18. Kaspersky. 2018. "Transparency Center". Kaspersky website. https://www.kaspersky.com/transparency-center-offices.

19. Kaspersky. 2019. "Kaspersky receives SOC 2 audit." Kaspersky website.https://www.kaspersky.com/about/compliance-soc2.

20. Kaspersky. 2020. "Kaspersky Security Bulletin 2019. Advanced threat predictions for 2020." Kaspersky, November 20. https://securelist.com/advanced-threat-predictions-for-2020/95055/.

21. Kaspersky. 2020. "Kaspersky receives certification for its data security systems." Kaspersky website. https://www.kaspersky.com/about/iso-27001.

22. Kaspersky. 2020. "APT trends report Q1 2020." Kaspersky, April 30. https://securelist.com/apt-trends-report-q1-2020/96826/.

23. Kaspersky. 2020. "Cyber Capacity Building Program." Kaspersky website.https://www.kaspersky.com/capacity-building.

24. Katzenstein, Peter J., Robert Owen Keohane, and Stephen D. Krasner. 1999. *Exploration and Contestation in the Study of World Politics*. MIT Press. Cambridge, Massachusetts and London, England. https://mitpress.mit.edu/books/exploration-and-contestation-study-world-politics.

25. Kemp, Walter. 2011. "From Confidence Tricks to Confidence Building: Resolving Conflict in the OSCE Area." International Peace Institute, May 4, 2011. http://reliefweb.int/sites/reliefweb.int/files/resources/Full_Report_504.pdf.

26. Kuehn, Andreas and Bruce McConnell. 2020. "Weathering TechNationalism: A Security and Trustworthiness Framework to Manage Cyber Supply Chain Risk." EastWest Institute. https://www.eastwest.ngo/sites/default/files/ideas-files/weathering-technationalism.pdf.

27. Mason, Simon J. A. and Matthias Siegfried. 2013. "Confidence Building Measures (CBMs) in Peace Processes." In, Managing Peace Processes: Process related questions. A handbook for AU practitioners, Volume 1, African Union and the Centre for Humanitarian Dialogue, 2013 (pp. 57-77).

28. NATO. 2018. "Brussels Summit Declaration Issued by the Heads of State and Government participating in the meeting of the North Atlantic Council in Brussels 11-12 July 2018." North Atlantic Treaty Organization, July 11, 2018. https://www.nato.int/cps/en/natohq/official_texts_156624.htm.

29. Netherlands Government. 2019. "Letter of 5 July 2019 from the Minister of Foreign Affairs to the President of the House of Representatives on the international legal order in cyberspace." Ministry of Foreign Affairs, July 5, 2019. https://www.government.nl/ministries/ministry-of-foreign-

affairs/documents/parliamentary-documents/2019/09/26/letter-to-the-parliament-on-the-international-legal-order-in-cyberspace.

30. "Open-ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security." A/RES/73/27. General Assembly of the United Nations. December 11, 2018. https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/73/27.

31. OSCE.2013. "Permanent Council Decision 1106." Organization for Security and Cooperation in Europe, December 3. http://www.osce.org/pc/109168.

32. OSCE.2016. "Permanent Council Decision 1202." Organization for Security and Cooperation in Europe, March 10. http://www.osce.org/pc/227281.

33. OEWG Initial Pre-Draft. 2020. "Initial "Pre-draft" of the report of the OEWG on Developments in the Field of Information and Telecommunications in the Context of International Security." UN Open-ended Working Group, March 16, 2020. https://unoda-web.s3.amazonaws.com/wp-content/uploads/2020/03/200311-Pre-Draft-OEWG-ICT.pdf.

34. OEWG Second Pre-Draft. 2020. "Second "Pre-draft" of the report of the OEWG on Developments in the Field of Information and Telecommunications in the Context of International Security." UN Open-ended Working Group, May 27, 2020. https://front.un-arm.org/wp-content/uploads/2020/05/200527-oewg-ict-revised-pre-draft.pdf.

35. Pawlak, Patryk and Thomas Biersteker. 2019. "Guardian of the Galaxy: EU Cyber Sanctions and Norms in Cyberspace." EU Institute for Security Studies. DOI:10.2815/672270.

36. Paris Call. 2019. "Paris Call for Trust and Security in Cyberspace: 9 Principles." Paris Call website. https://pariscall.international/en/.

37. Ruhl, Christian, Duncan Hollis, Wyatt Hoffman, and Tim Maurer. 2020. "Cyberspace and Geopolitics: Assessing Global Cybersecurity Norm Processes and a Crossroads." Carnegie Endowment for International Peace. https://carnegieendowment.org/files/Cyberspace_and_Geopolitics.pdf.

38. Roadmap for Digital Cooperation.2020. Report of the Secretary-General, United Nations, June, 2020. https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf.

39. Sabbah, Cedric. 2018. "Pressing pause: A new approach for international cybersecurity norm development." In T. Minárik, R. Jakschis, L. Lindström (Eds.), 10th international conference on cyber conflict CyCon X: Maximising effects (pp. 263-281). Tallinn: CCDCOE.

40. Satter, Raphael, Jack Stubbs, and Christopher Bind. "Exclusive: Elite hackers target WHO as coronavirus cyberattacks spike." Reuters, March 23, 2020. https://www.reuters.com/article/us-health-coronavirus-who-hack-exclusive/exclusive-elite-hackers-target-who-as-coronavirus-cyberattacks-spike-idUSKBN21A3BN.

41. Stadnik, Ilona. 2019. "2nd Meeting of the first substantive session of the Open-Ended Working Group (OEWG)." Report. Geneva Internet Platform, September 9, 2019. https://dig.watch/resources/2nd-meeting-first-substantive-session-open-ended-working-group-oewg.

42. Steinberg, Gerald M. 2003. "Confidence Building and the Concept of Spillover in Mediterranean Conflicts." Organization for Security and Cooperation in Europe, October 7, 2003. https://www.osce.org/files/f/documents/7/9/12564.pdf.

43. Stigherrian, 2019. "Australia to keep playing the UN cyberspace norms game." ZDnet, April 16, 2019. https://www.zdnet.com/article/australia-to-keep-playing-the-un-cyberspace-norms-game/.

44. UK Government. 2018. "Cyber and International Law in the 21st Century." Speech by Attorney General Jeremy Wright QC MP. UK Government, May 23, 2018.

https://www.gov.uk/government/speeches/cyber-and-international-law-in-the-21st-century.

44. UN Study. 1981. "Comprehensive Study on Confidence-Building Measures." United Nations General Assembly studies in the field of disarmament, 1981. https://www.un-ilibrary.org/disarmament/comprehensive-study-on-confidence-building-measures_c35ca164-en.

45. US Comments. 2020. “United States Comments on the Chair’s Pre-draft of the Report of the UN Open Ended Working Group (OEWG).” UN Open-ended Working Group, April, 2020. https://front.un-arm.org/wp-content/uploads/2020/04/oewg-pre-draft-usg-comments-4-6-2020.pdf.

46. Wendt, Alexander. 1992. “"Anarchy Is What States Make of It: The Social Construction of Power Politics." International Organization, Vol. 46, No.2 (Spring, 1992), (pp. 391-425). https://www.jstor.org/stable/2706858.

47. Zetter, Kim. 2019. “Exclusive: How a Russian firm helped catch an alleged NSA data thief.” Politico, September 1, 2019. https://www.politico.com/story/2019/01/09/russia-kaspersky-lab-nsa-cybersecurity-1089131.

48. Ziolkowski, Katharina. 2013. “Confidence Building Measures for Cyberspace – Legal Implications.” Tallinn: CCDCOE (pp. 30-32), 2013. https://ccdcoe.org/uploads/2018/10/CBMs.pdf.