

POLICY BRIEF NO. 1

# ACHIEVING TRUST IN EU CYBERSECURITY

## Trust and trustworthiness

How important is cybersecurity for online trust? And how can trust in cybersecurity be achieved? We shall briefly outline the gist of many philosophical, psychological and sociological views on trust. Then, we shall explore its implications for cybersecurity.

### Theories of trust

Trust is often associated with conditions of risk, uncertainty and vulnerability. Vulnerability derives from reliance on others in conditions of imperfect information, in particular about others' motivations and/or behavior. Scholars have defined trust to involve essentially moral commitments (Baier 1986; Nickel 2007), or a foreseeable convergence of interests (Hardin 1992). Others yet require mutually reinforcing motivations, e.g. trustworthy people are intrinsically motivated to be reliable by the trust placed in them (Pettit 1995). Trust is often contextual: I may trust my accountant to file my tax returns but not as a political adviser.

Well-placed or reasonable trust is trust towards trustworthy trustees: trustees deserving of trust. Well-placed trust is socially beneficial. Conversely, trusting non-trustworthy agents is typi-

cally harmful. Trust is not only easy to gain, but even harder to re-gain after it has been lost.

Trust has rational and non-rational aspects. The non-rational aspect is a kind of optimism about other people's motivations (Jones 1996). Some people trust others about whom they have little information, even complete strangers. This allows the emergence of socially beneficial forms of cooperation, which in the absence of such attitudes could not be built (Ostrom 2000). These conditions are more easily achieved in small groups where non-cooperative behaviour can be detected (Ostrom 2000). Thus, achieving social cooperation solely based on trust in complex societies or in the international arena may not be feasible (Hardin 2009).

### Online trust

In the early days of the internet, it seemed that online trust could not be built, due to internet anonymity (Pettit 2004): as the famous vignette said, "on the web, no one knows that you are the dog". In the current internet 2.0, this is no longer true. Online feedback and digital social scoring systems are all based on semi-stable and semi-verifiable internet identities (Et-

zioni 2017). These solutions have allowed online markets (e.g. Ebay and Amazon) and interactions in the sharing economy (e.g. AirBnB and BlaBlaCar) to emerge.

### Interpersonal and institutional trust

An agent or organization who proves to be reliable only because it fears being brought to court is not trustworthy. Many forms of trust and trustworthiness take place in the absence of direct economic incentives or legal sanctions. Yet, in high stake exchanges among complete strangers, especially if one-shot, strong interpersonal trust is hard to achieve, because it is too risky. So, designing an environment, and system of reputation and incentives, that enables and promotes trust is not an easy task. Effectively enforced legal sanctions against dishonest behaviour may favour trust-based relations, by mitigating the risks associated with trusting strangers (Hardin 1992).

Yet, trust in legal sanctions presupposes trust in the institutions that enforce these norms.

## Cybersecurity and trust

### Cybersecurity is necessary for trust

Before we consider trust in cybersecurity, it is worth showing that cybersecurity is necessary for trust. The classical goals of cybersecurity: confidentiality, integrity and availability, are all prerequisites of trust in the digital domain. From the data protection point of view, integrity is defined as the property that data and services cannot be modified in an unauthorized or undetected manner. Thus, integrity is an essential element creating and maintaining trust in information. Data integrity, for example, is essential to assess ICT services and to justify the trust placed in companies providing good services. The integrity and availability of information is also necessary to verify the digital identity of the people with whom one interacts online. If anyone can appropriate someone's digital identity online, reputation can be misused and no online identity is trustworthy any longer.

Moreover, mutual trust is characterized by confidentiality. Trust enables the sharing of sensitive information. Those who trust make themselves vulnerable to those whom they trust, because the sensitive information they share could be used against them. This can only be reasonable with a trusted party. Hence, cybersecurity is needed for online trust: in cybersecurity, confidentiality means that data and services cannot be accessed by unauthorized entities. Cybersecurity reduces the risk of sensitive information becoming accessible to non-trustworthy agents. In the absence of cybersecurity, it would not be reasonable to share sensitive information online so many online transactions would not take place.

### Trust is necessary for cybersecurity

While trust depends on cybersecurity, cybersecurity also depends on trust. Because trust and cybersecurity are interdependent, there may be both virtuous and vicious cycles. The case study of ethical hacking illustrates a virtuous cycle, in which trust enables cybersecurity, which promotes higher levels of online trust. The case study of zero-day exploits presents a vicious cycle, where initial mistrust weakens cybersecurity, which then further undermines trust.

### Ethical hacking

Ethical, or 'white-hat' hackers are defined here as hackers who pursue legal goals such as testing information system security against malicious attacks. Therein, they may use similar techniques like so-called malicious, or 'black-hat' hackers do, yet white-hat hackers would not criminally exploit ICT vulnerabilities for their own benefit. Trustworthy white-hat hackers exist and they are employed by many companies in order to identify their vulnerabilities. This form of security research is only possible because some companies have been initially willing to trust some white-hat hackers. The choice of initial, less risk-averse, firms benefits also other, more risk-averse firms. It is more reasonable to trust white-hat hackers with a demonstrable reputation, guaranteed by other firms. So, companies who trust each other have ways to identify trustworthy ethical hackers.

The recommendation of a trusted company provides a basis for the belief in the trustworthiness of the white-hat hacker. In addition to that, an environment in which, for a hacker it pays to have a good reputation, also enables trust because of its incentives.

The chain of trust expands further. Trust between the company and the ethical hacker makes companies more secure, and therefore more trustworthy for their clients. The overall legal frameworks of European data protection and, more broadly, all national and international laws fostering cybersecurity contribute to online cybersecurity via institutional trust. Consumers may receive assurance from the fact that companies that do not protect cybersecurity can be sanctioned, assuming that firms will rationally act to avoid sanctions. Companies that comply with obligations to implement sufficient organizational and technical measures fostering a high level of cybersecurity are, therefore, more trustworthy. Data protection laws requiring transparent communication to data subjects in case of data breaches also provide an incentive for companies to adopt higher levels of cybersecurity. Companies have greater incentives to avoid data breaches if they cannot (legally) hide this fact to their clients.

### **Distrust weakens cybersecurity: The use of zero-day exploits**

Zero-day vulnerabilities are weaknesses in software which have not yet been identified before release. The exploitation of zero days is a kind of weapon, as it can disrupt computers and networks, as well as they can give unauthorized access to relevant information. Governments may make use of zero days for foreign intelligence activities by buying and deploying them in order to attack or to spy on other countries or individual opponents. In this context, it can be said that the motivation to buy zero-day exploits for cyber-espionage against national entities usually derives from a failure of mutual trust between states. This is a self-reinforcing distrust mechanism, as the engagement in a zero-day market in turn undermines the trustworthiness of governments, in more than one way.

For example, if each government seeks for vulnerabilities in the systems of other countries, this has the effect that in the long-run each country will be less secure. This causes not only the danger of such vulnerabilities falling into the hands of criminals, but can also be seen as an arms race scenario between nation states. In such a situation, no country can afford to stay inactive due to the fear that other countries will gain an advantage that can be used against them. Thus, the search for 'cyber-vulnerabilities' of the other countries makes relationships of trust among countries impossible. With mistrust as a baseline, each state rightly assumes that such an offensive strategy is favorable in contrast to purely defensive ones in order to gain the upper hand in foreign espionage or cyber-sabotage activities. Moreover, a national government that is shown to be vulnerable to exploits can also appear less trustworthy towards other countries. When confidential communications of diplomats and politicians are revealed or compromised, this may cause serious disruptions in interstate relations. Furthermore, governments making use of exploits will most likely be feared, rather than trusted, by their citizens. After all, if governments have these weapons, their citizens may also fall prey of cyber-espionage and sabotage. Acquiring zero-day exploits may appear strategically beneficial, even obligatory in an unfavourable context of low international trust. Yet, the most likely result of developing or purchasing these tools is a race to the bottom with respect to trust.

### **Conclusions**

In conclusion, trust and cybersecurity are mutually supportive. Trust is beneficial (when well-placed), fragile and hard to rebuild. Individual strategies (by persons, companies, and states) that may appear rational and effective to protect national security should always be scrutinized to assess their effects on trust, or they risk to back-fire. In commercial contexts, trust may be achieved with or without strong oversight and legal sanctions, and also with a balanced mix of legal sanctions and reputation systems.

## Where more info can be found

### References:

- Baier, Annette. 1986. "Trust and Antitrust." *Ethics* 96 (2): 231–60.
- Etzioni, Amitai. 2017. "Cyber Trust." *Journal of Business Ethics*, July. <https://doi.org/10.1007/s10551-017-3627-y>.
- Hardin, Russell. 1996. "Trustworthiness." *Ethics* 107 (1): 26–42.
- Jones, Karen. 1996. "Trust as an Affective Attitude." *Ethics* 107 (1): 4–25.
- Nickel, Philip J. 2007. "Trust and Obligation-Ascription." *Ethical Theory and Moral Practice* 10 (3): 309–19. <https://doi.org/10.1007/s10677-007-9069-3>.
- Ostrom, Elinor. 2000. "Collective Action and the Evolution of Social Norms." *The Journal of Economic Perspectives* 14 (3): 137–58.
- Pettit, Philip. 1995. "The Cunning of Trust." *Philosophy & Public Affairs* 24 (3): 202–25.
- 2004. "Trust, Reliance and the Internet." *Analyse & Kritik* 26 (1): 108–121.

This Policy Brief is based on the research work done by the CANVAS project (Constructing an Alliance for Value-driven Cybersecurity). Detailed reports of this work have been published in four main White Papers:

1. Cybersecurity and Ethics
2. Cybersecurity and Law
3. Attitudes and Opinions Regarding Cybersecurity
4. Technological Challenges in Cybersecurity

**All White Papers can be found on our website, along with** all of our (downloadable and printable) Policy Briefs, short online explanations of the key cybersecurity issues, and commented literature lists for further reading:

[canvas-project.eu](https://canvas-project.eu)

Moreover, you can find even more CANVAS project material on our website:



**CANVAS Reference Curriculum**  
(integrating the value perspective into cybersecurity training and education)



**CANVAS MOOC**  
(Massive Open Online Course)



**Open Access Book**  
"The Ethics of Cybersecurity"



Co-funded by the Horizon 2020 programme of the European Union

The CANVAS project (Constructing an Alliance for Value-driven Cybersecurity) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 700540. This work was supported (in part) by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0052-1. The opinions expressed and arguments employed therein do not necessarily reflect the official views of the Swiss Government.

#### Objective of CANVAS:

To bring together stakeholders from key areas of the European Digital Agenda to approach the challenge how cybersecurity can be aligned with European values and fundamental rights.

#### Partners:

The CANVAS Consortium consists of 11 partners (9 academic institutions and 2 partners outside academia) located in 7 European countries.

#### Version and date of publication:

Version 2.0, October 2019

#### Funding:

1.57 Mio. €, of which 1 Mio. € is funded by the European Commission and the remaining part emerges from the Swiss State Secretariat for Education, Research and Innovation.

#### Project coordination and contact:

PD Dr. sc. ETH Markus Christen, University of Zurich (UZH), Digital Society Initiative, Rämistrasse 66, 8001 Zürich

#### Project duration:

September 2016 – October 2019