Risks to Cybersecurity from Data Localization, Organized by Techniques, Tactics, and Procedures

By Peter Swire, DeBrae Kennedy-Mayo, Drew Bagley, Sven Krasser, Avani Modak & Christoph Bausewein[1]

Table of Contents

I. Introduction

This paper continues the research program begun in "The Risks of Data Localization to Cybersecurity – Organizational Effects" ("Risks"). (Swire 2022). That paper is now available on SSRN, and is in final phases of revision for a peer-reviewed, inter-disciplinary journal. In this new paper, we continue to examine obstacles to cybersecurity that result from "hard" data localization, where transfer of data is prohibited to other countries. We also continue the focus on defensive cybersecurity –risks to the ability of organizations such as corporations and government agencies to identify, protect, detect, respond, and recover in the face of cyber-attacks.

The importance of data localization has risen rapidly in recent years. This paper focuses on examples from the European Union ("EU"), which has taken significant steps toward *de facto* localization of personal data in the wake of the 2020 *Schrems II* decision of the European Court of Justice (*Schrems II* 2020). Among enforcement actions since that decision, the Portuguese data protection authority ordered a government agency to terminate its use of cybersecurity services from U.S.-based cybersecurity company Cloudflare (CNPD 2021). In the Data Act and other proposed legislation, the EU would also impose localization rules for defined categories of both personal and non-personal data (COM/2022/68 2022, Art. 2(1)(af)). Additional localization could result from the proposed European Cybersecurity Certification Scheme for Cloud Services ("EUCS"), discussed in Part V (ENISA 2020).

This paper thus continues to examine the risks of localization rules for personal data, while recognizing that some localization rules may also block categories of non-personal data. As Nigel Cory and Luke Dascoli have documented, the number of data localization measures roughly doubled from 2018 to 2021, including at least 62 countries with 144 restrictions (Cory and Dascoli 2021).

Using an approach based on organizational form, *Risks* provides a new categorization of the risks of data localization on cybersecurity management. We analyzed risks within an organization, across organizations with payment, and across organizations without payment. First, our analysis showed that despite data localization often being used as a proxy for better data protection, such policy would actually threaten an organization's ability to achieve *integrated management of cybersecurity risk*. We analyzed International Standards Organization ("ISO") 27002, as a way to systematically examine the risk of data localization for a widely-used set of cybersecurity management controls. We found that 13 of the 14 ISO 27002 controls, as well as multiple sub-controls, would be negatively affected by localization of personal data. Second, the analysis explained how data localization pervasively limits *provision of cybersecurity-related services by third parties*, a global market of roughly $200 billion currently, with doubling expected within a few years. Put simply, a great variety of cybersecurity services rely on transfers of personal data across borders. Third, data localization threatens non-fee cooperation on cybersecurity defense. Notably, localization undermines *information sharing for cybersecurity purposes*, which policy leaders have emphasized as vital to effective cybersecurity.

This paper supplements *Risks* by organizing the risks to cybersecurity by the techniques, tactics, and procedures ("TTPs") of threat actors and defenders. To categorize the TTPs, we have relied on two authoritative approaches. First, we analyzed types of attacks in the widely-known MITRE ATT&CK Framework, which details high-level adversary tactic categories and individual techniques that adversaries can use within each of the tactic categories. We also examined the technical and organizational measures supported by the European Union Agency on Cybersecurity ("ENISA") and the German TeleTrust IT Security Organization in Germany in their 2019 guidelines on "The State of the Art" for cybersecurity (ENISA and TeleTust 2021). Using these two approaches, we highlight three important tactics defenders use for cybersecurity purposes – (1) threat hunting/threat intelligence; (2) privilege escalation attack/lateral movement; and (3) red teaming/pen testing. The two categorizations result in similar conclusions -- all three of these categories, considered essential to a mature cybersecurity program, would routinely

require the cybersecurity defenders to access types of personal data that would be restricted by current data localization laws and proposals.

Part II of this paper examines the tension between the EU's regulatory requirements for cybersecurity and data protection. Requirements for effective cybersecurity include Article 32 of the General Data Protection Regulation ("GDPR"), Article 13(1) of the EU CER Directive, Article 21(1) of the NIS2 Directive and Article 5 (1)(g) of the EU Cybersecurity Act (EU Directive 2022/2557, 164-198; EU Directive 2022/2555, 80-152; EU Regulation 2019/881, 15-69). Under these and similar laws, organizations in the EU are expected to deploy effective security safeguards appropriate to the risk taking into account the "state of the art" in cybersecurity as outlined by ENISA's guidance (ENISA and TeleTrust 2021). At the same time, data protection laws prohibit the processing of personal data unless it is lawful (Art. 6 and 9 of the GDPR) and adequately protected when transferred out of the EU (Chapter 5 of the GDPR). As defined within the EU, "personal data" is a broad term that includes numerous categories of data routinely used by cybersecurity defenders. For example, IP addresses are provided to a server as an essential part of web communications (*Breyer* 2020).[2] Despite this functionality and ubiquity, IP addresses are included within the scope of "personal data" that EU enforcement actions have found should not be transferred to the U.S. and other non-EU third countries. In recent EU enforcement actions, simply the possible transfer of IP addresses to the U.S. has been the stated basis for data protection authorities to find that Google Analytics is unlawful on EU websites (ADPA 2021; CNIL 2022).

Part III of this paper examines the MITRE ATT&CK Framework and how it organizes relevant aspects of a cybersecurity defense system. The analysis highlights how data localization requirements undermine the three examples of threat intelligence, privilege escalation, and red teaming.

Part IV supplements the effects in Part III by providing a quantitative model illustrating effects of data localization under plausible assumptions. In the model, halving the number of IP addresses available to a defender would more than double the likely time until a new attack is detected.

Part V extends the analysis to the cybersecurity approaches now being considered under the proposed EUCS. The hard data localization in some proposals appear to conflict with the findings of this paper, that hard data localization would undermine defensive measures such as threat intelligence, privilege escalation, and red teaming.

Part VI offers conclusions. The U.S., Europe, and other nations face incessant and sophisticated cyber-attacks. In the face of these threats, imagine that policymakers were considering a law that would degrade threat intelligence, leave systems open to privilege escalation, and bar effective pen testing and other red teaming. Such a proposed law would deserve great skepticism. As documented in this paper's research, however, data localization laws appear to create such risks. This paper adds to the finding in *Risks*, that "until and unless proponents of localization address these concerns, scholars, policymakers, and practitioners have strong reason to expect significant cybersecurity harms from hard localization requirements."

II. The Tension in EU Data Protection Law Between Cybersecurity "State of the Art" and Potential Privacy-based Limits on Processing Personal Data

In this Part we set forth the legal requirements in the European Union for cybersecurity "state of the art," and briefly describe ENISA's guidelines for achieving that "state of the art." We then discuss potential privacy-based limits on processing personal data.

A. The GDPR's Call for State of the Art Cybersecurity

The GDPR, including in its Recital 49, requires that cybersecurity be an integral part of data protection (ENISA and TeleTrust 2021, 9; GDPR Rec. 49). Article 5(1f) of the GDPR sets forth "Principles relating to processing of personal data." Among other requirements, personal data shall be "processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures" (GDPR Rec. 39).

Article 32 of the GDPR addresses the "Security of processing" particularly mandating that the "state of the art" of cybersecurity practices be included in the risk analysis for handling data security (GDPR Art. 32; ENISA and TeleTrust 2021, 9, 87).[3] One commentator has described the "state of the art" as consisting of "measures that are based on proven knowledge, of an advanced technical development, practically suitable, ready and available for technical implementation, but have not necessarily become established in practice yet" (Selzer 2021, 123). Notably, cloud services for global threat analysis are already a common component of many security solutions. Furthermore, approaches such as Threat Intelligence Platforms, at their core, aggregate and share data pertinent to threat detection. This threat-relevant data often includes IP addresses and other information considered "personal data" under EU law (Kime, 2023).

Article 32 of the GDPR adopts a risk-based approach to what measures are appropriate. Data controllers should deploy "appropriate technical and organizational measures." The appropriateness of measures depends both on risks to cybersecurity and to "the rights and freedoms of natural persons" (GDPR Rec. 78).

Where a breach of security occurs, Article 33 of the GDPR requires an organization to notify the competent data protection authorities within 72 hours unless it is unlikely to pose a risk to the fundamental rights and liberties of data subjects. Furthermore, Article 34 of the GDPR requires an organization to notify the individuals themselves where there is likely to be a high risk to their rights and freedoms. Complementary guidance from the EDPB makes clear that "high risk" is defined by the circumstances surrounding the nature of the data, risk mitigation measures in place, and the recipient of the breached data (EDPB 2023).

Recital 75 of the GDPR defines "Risks to the Rights and Freedoms of Natural Persons" as personal data processing which could lead to physical, material or non-material damage. The Recital applies to damages that specifically can result from cybersecurity incidents, such as identity theft, fraud, financial loss, damage to reputation, and loss of confidentiality. The Recital

also lists other relevant damages, including where data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data.

B. ENISA Guidelines on "State of the Art" for Cybersecurity

The conundrum is even stronger since the European Union Agency for Cybersecurity ("ENISA") in co-operation with the IT Security Association Germany ("TeleTrust") has issued guidelines on the "state of the art" required for appropriate technical and organizational measures ("Guidelines") in 2019, shortly after GDPR went into effect. These Guidelines provide guidance on "What is 'state of the art' in IT security?" (ENISA 2019). In examining 'state of the art,' the Guidelines adopt the approach that "state of the art depends on whether a measure is technically necessary, suitable and appropriate from the perspective of technical practitioners. It can and should be possible to react to more current threats – and especially to the current technical possibilities for attack." (ENISA and TeleTrust 2021, 11).

We offer two observations about these Guidelines. First, the Guidelines provide an explanation of what technical and organizational measures are expected in order to meet the "state of the art" provided by the very EU agency dedicated to achieving a high common level of cybersecurity across the EU. We explore some of these measures in detail below in Part III, as we discuss key Tactics, Techniques, and Procedures that are affected by data localization.

Second, we note an interesting discussion of risk in the Guidelines. As discussed elsewhere in more detail, the European Data Protection Board ("EDPB") has expressed disapproval of a risk-based approach for assessing when transfer of personal data is lawful (Christakis 2020). ENISA, however, explicitly states that appropriate cybersecurity measures should take into account the level of risk to fundamental rights. ENISA says:

> "Article 32 of the GDPR regulates "security of processing" to ensure that, taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as *the risk of varying likelihood and severity for the rights and freedoms of natural persons*, appropriate technical and organizational measures are implemented." (emphasis added) (ENISA 2019).

In forthcoming research, Théodore Christakis is examining in detail whether and how a "risk-based approach" is appropriate under EU law and practice.

C. Potential Privacy-based Limits on Processing Personal Data for Cybersecurity Purposes

Along with the EU expectation for providing state of the art for cybersecurity, there are EU legal authorities that appear to limit achievement of that state of the art. In particular, the GDPR places restrictions on processing of personal data. Both processing and personal data are defined terms in this regulatory scheme.

The GDPR applies broadly to personal data that originates from the EU, as described in Article 3(1) of the GDPR. Under the GDPR, "personal data" means any information relating to an identified or identifiable natural person, pursuant to Article 4 (1) of the GDPR. The European

Commission ("Commission") has explained the broad scope of "personal data" as defined in Article 4 of the GDPR: "Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data" (EC: "What is Personal Data"). As examples, the Commission includes not only obvious identifiers such as name and address, but also more technical identifiers such as: "location data (for example the location data function on a mobile phone); an Internet Protocol (IP) address; a cookie ID; [and] the advertising identifier of your phone" (EC: "What is Personal Data"). Other information available to administrators of company or agency computer systems are also often considered personal data, such as MAC addresses where they are linkable to a personal device (Future of Privacy Forum 2014).

"Processing" is also a broad term, meaning any operation performed on personal data. This includes any access, collection, storage, adaptation or alteration, use, transfer, disclosure by transmission, otherwise making available, or even the erasure or destruction of personal data.

The broad scope of "processing" of "personal data" effectively means that ubiquitous unique, and often-times public, identifiers inherent to modern IT and network infrastructure are regulated by GDPR. Examples from the state of the art for cybersecurity include collection of security telemetry from endpoints, cloud workloads, network email, or threat data from previously siloed security tools across an organization's technology stack for easier and faster investigation, threat hunting, and response. One consequence is that certain processing of personal data may be unlawful even when it would seem necessary and proportionate, such as processing for cybersecurity purposes to protect critical infrastructure, national security, economic purposes, and even the security of an individual's data.

EU data localization has become a more prominent legal risk in the wake of the 2020 *Schrems II* decision of the European Court of Justice, which announced limits on transfer of personal data to third countries (*Schrems II* 2020). Subsequently, the EDPB issued guidance about assessing the laws and practices of the destination country and technical and organizational safeguards (termed "supplementary measures") to ensure adequate protection in the transfer of personal data (EDPB 2021). The EDPB expressed reservations about applying a risk-based approach to such transfers.

These legal developments have raised concerns for organizations using cybersecurity services that are not exclusively delivered within the EU, including hosting, support, engineering, and service. Customer service, IT operations, or a security operations center that "follows the sun," to provide 24/7 support, are examples of services that may be difficult or impossible in practice to provide exclusively from within the EU.

The conundrum is how to proceed when data protection laws designed to limit harms to personal data and to protect personal data have the apparent consequence of increasing harms to personal data (by setting data localization limits on applying the state of the art for cybersecurity). We note that guidance interpreting the data breach rules under the GDPR makes clear that notification is not required where a personal data "breach is unlikely to result in a risk to the rights and freedoms of individuals" (EDPB 2023). Risk is focused on "physical, material or non-material damage" to breach victims, such as "discrimination, identity theft or fraud,

financial loss and damage to reputation." In assessing this, "consideration should also be given to other personal data that may already be available about the data subject." This suggests that in most contexts the data elements used in cybersecurity, such as IP address, MAC address, or email address, are low risk – not requiring a breach notice even when they are seized illegally by hackers and transferred to a third country. On the other hand, under some interpretations processing of these same data elements are considered a violation of the data subject's fundamental rights when they are transferred intentionally, even when being used by an organization for GDPR Article 32 and Recital 49 cybersecurity purposes. When reviewing both the legal requirements in the European Union for cybersecurity "state of the art" and the privacy-based limits on processing personal data, the conclusion that emerges is there has not been full and explicit consideration among EU legal authorities about how overall to achieve both cybersecurity state of the art and also limit use and transfer of many data elements that are required to institute the state of the art (Bagley 2022).[4]

III. Using the MITRE ATT&CK Framework to Develop Themes for Where Localization Creates Obstacles for Defenders

With the current EU data localization approach in mind, we next turn to examination of the risks of hard data localization laws on cybersecurity tactics, techniques, and procedures ("TTP"). NIST defines TTP as

> "The behavior of an actor. A tactic is the highest-level description of this behavior, while techniques give a more detailed description of behavior in the context of a tactic, and procedures an even lower-level, highly detailed description in the context of a technique" (NIST).

In Part III, we first explain the role of the MITRE ATT&CK Framework as a leading approach for assessing TTPs. We explain the methodology for using the Framework to identify which TTPs are most likely to be affected by data localization rules. We explain two themes for such TTPs: the "who and what" of an attack, and the "risks from knowing less than the attacker."

To explain these themes, we provide more detail on threat hunting and privilege escalation as two examples of the "who and what" of an attack, and then use red teaming and pen testing as an example of the "risks from knowing less than the attacker." For each of these, we provide: (a) the anatomy of the approach; (b) types of personal data; (c) alternatives to use of personal data; and (d) the requirements of the ENISA Guidelines applied to each of threat hunting, privilege escalation, and red teaming.

A. The MITRE ATT&CK Framework and TTPs

In researching the risks to TTP, we use the widely-known MITRE ATT&CK framework. MITRE's ATT&CK framework focuses on pre-compromise preparations and post-compromise activities of adversaries (MITRE: Enterprise Matrix). It provides a detailed enumeration of

common adversary behaviors after they have gained access to a system within a network (Strom et al. 2017). MITRE researchers explained that the framework serves "as a method for discovering analytic coverage and defense gaps inside a target network" (ibid, 1). For our purposes, the framework can help pinpoint "defense gaps" resulting from limits on transfer of personal data.

The MITRE Framework relies on examination of "tactics" and "techniques." Tactics describe the reasons why the adversary acts or the goals that the adversary hopes to accomplish (ibid, 12). The tactics discussed in the framework are: (1) Reconnaissance; (2) Resource Development; (3) Initial Access; (4) Execution; (5) Persistence; (6) Privilege Escalation; (7) Defense Evasion; (8) Credential Access; (9) Discovery; (10) Lateral Movement; (11) Collection; (12) Command and Control; (13) Exfiltration; and (14) Impact.

Techniques are more detailed than tactics and describe the actions that the adversary takes to accomplish their tactics. The ATT&CK framework analyzes these techniques from both the offensive and defensive points of view (ibid, 12; CrowdStrike: IOA v. IOC).[5] Version 13 of ATT&CK for Enterprise includes the aforementioned 14 tactics and 196 techniques (MITRE 2023). Instead of being a theoretical taxonomy that seeks to categorize every possible category of attack, the techniques have been based empirically on observed intrusions. MITRE sought to "address the need for additional details while remaining grounded in observed and plausible adversary behavior" (Strom et al. 2017, 9). To illustrate the application to current types of attacks, rather than a general taxonomy, some of the techniques are specific to widely-used software, such as "Windows Remote Management" or "DLL side loading" (referring to dynamic link libraries in the Windows operating system) (Cybereason Global SOC Team 2022).

Although the methodology for use of the MITRE ATT&CK framework evolved during the course of research for this article, the combined team (i.e., the current co-authors from Georgia Tech and CrowdStrike) utilized the MITRE ATT&CK framework as a starting point for assessing the TTPs that are most affected by the localization limits.

B. Two Key Themes for TTPs Affected by Localization

Based on the combined team's assessment, we present two themes. First, localization laws can disrupt the defenders' ability to determine "The Who and the What" of an attack. The basic idea is that details about "who" is attacking often requires access to personal data. Similarly, as an attacker moves through a defender's system, there are often account names or other personal data in tracking "what" the attacker does in the system. Put another way, the telemetry and other data used by defenders may often include personal data (or other protected data), in ways that create obstacles for defenders if that data is not available due to hard localization. As discussed further below, threat hunting and privilege escalation are two important defensive measures that are likely to be especially hard hit by limits on data transfer.

Second, localization laws can result in "Risks From Knowing Less Than the Attacker." An essential part of good cyber defense is for the defenders to test the system through "red teaming," including penetration ("pen") testing. Pen testing involves the defense hiring "white hat" attackers to find as many vulnerabilities and configuration issues as possible, exploit them, and determine risk levels (Talamantes). Red teaming is a more general approach, for the defender to identify physical, hardware, software, and human vulnerabilities. In addition to pen testing, red team skills include social engineering, threat intelligence, and reverse engineering (Coursera 2022).

Data localization laws would appear to present serious obstacles to pen testing and other red teaming. The intuition is that attackers will be willing to break the law, to seek out and transfer personal data across national borders. Defenders, by contrast, must follow the law. If defenders hire pen testers, those testers would not be able to probe for vulnerabilities that would require learning account information and other personal data, notably where the data is stored in a different country. Since a large fraction of cyber-attacks involve crossing national borders, the defenders would systematically be able to test and learn about vulnerabilities in their own systems less well than the attackers.

1. Impacts to Cybersecurity: "The Who and the What" of an Attack

Notably, cyber attacks today often do not involve the use of malware, instead leveraging the use of legitimate credentials increasingly obtained from "access brokers" (CrowdStrike 2023b). This means defenders must prioritize detecting the behavior of the adversary in a victim's system. To highlight the complexities of the need to use personal data in "state of the art" cybersecurity, we discuss two examples that go to "the who and the what" of an attack – threat hunting and privilege escalation. The basic idea is that details about who is attacking often requires access to personal data. Similarly, as an attacker moves through a defender's system, there are often account names or other personal data in tracking "what" the attacker does in the system. Put another way, the telemetry and other data used by defenders may often include personal data (or other protected data), in ways that create obstacles for defenders. Data may not be available due to hard localization or encumbered data may be inadvertently obtained in the course of pen testing, exposing the defenders to legal risks.

a. Threat Hunting.[6]

Threat hunting is the practice of proactively searching for cyber threats in a company's environment that have bypassed the endpoint security defenses. Threat hunting is critical to addressing advanced persistent threats (APTs). Threat hunting works with the assumption that the attacker is already in an organization's system. Steps include: hypothesis-driven investigation, investigation using tactical threat intelligence to catalog known Indicators of Compromise (IOCs) or Indicators of Attack (IOAs), and advanced analytics and machine learning investigations (Taschler 2023; CrowdStrike 2023a; Baker 2023; CrowdStrike 2022a).

i. Anatomy of Approach. The process for threat hunting involves three steps: the trigger, the investigation, and the resolution. The trigger points the defender to a specific system or area of the network for further investigation. The investigation involves using tools to determine

whether the potential compromise of the system is malicious or benign. The resolution involves communicating the intelligence related to the malicious activity to operations and security teams so they can respond to the incident and mitigate the threats (Taschler 2023).

ii. Types of personal data. An attacker generally creates digital footprints that may include personal data. These may exist in logs generated in the operating system or telemetry data captured by cybersecurity technologies such as the EDR (ENISA and TeleTrust 2021, Sec. 3.3.22). Examples of personal data may include usernames, file names, and IP addresses.

For hypothesis-driven investigations, a large pool of crowdsourced attack data gives insight into attackers' latest tactic, techniques, and procedures (Taschler 2023). This crowdsourcing often includes personal data.

A large and increasing fraction of attacks do not use malware, so defenders doing threat hunting routinely rely on examining details of Indicators of Compromise (IOCs) or Indicators of Attack (IOAs), which may include personal data, as shown for instance in documentation of detection of attacks by an APT (CrowdStrike: IOA v. IOC).

For advanced analytics and machine learning, defenders look for irregularities across an array of telemetry. These defenses use queries and automation to extract leads and then have a skilled human defender identify the signs of attacker activity (Taschler 2023).

iii. Alternatives to Use of Personal Data. The effectiveness of threat hunting would likely be decreased due to the limited IOCs and IOAs from countries with data localization. Data localization would imply that protected data could likely not leave these countries but would not necessarily prohibit data from non-data localization countries from entering. It is certainly possible that these non-data localization countries may be unwilling to share information with countries that have data localization in place. Countries or regions with larger populations, such as the EU and India, may be less affected than countries with smaller populations, but the extent of such difference deserves additional empirical investigation.

The non-sharing of information could lead to a situation where an attack could be successful region to region, instead of cybersecurity specialists being able to defend against the attack worldwide once it had appeared in one region. In addition, the smaller regional datasets complicate building a proper baseline of normal behavior for an organization. That impedes human threat hunters, but it especially hinders the creation of machine learning-based threat detectors, which require large and diverse datasets to be trained properly. This in turn raises the costs of operating security infrastructure at scale and increases the costs of security incidents where time is of the essence.

iv. Threat hunting and the ENISA Guidelines. The ENISA Guidelines on "state of the art" contemplate extensive efforts to conduct threat hunting, and threat intelligence more broadly. Notably, the guidelines include "Technical Measure - 3.2.27 Threat intelligence," which provides:

> Tactical Cyber Threat Intelligence includes malware analysis and the import of individual, static, and behavioral threat indicators into defensive IT security solutions. Operational Cyber Threat Intelligence is used to improve knowledge about an attacker, his skills, infrastructure and attack tactics to implement more targeted cybersecurity measures such as proactive threat hunting. Strategic Threat Intelligence enables a better understanding of the current threat situation (threat assessment) (ENISA and TeleTrust 2021).

Other relevant parts of the ENISA Guidelines are "Technical Measure - 3.2.22 - Endpoint Detection and Response" and "Technical Measure - 3.2.24 – Attack Detection and Analysis" (ibid).

b. Privilege Escalation Attacks

The next example focuses on a privilege escalation attack using spear phishing (Falcon OverWatch Team 2021; CrowdStrike 2022b; CrowdStrike 2023b). Spear phishing often is part of a "privilege escalation" attack – an attack designed to gain more access than authorized by the system. Gaining privilege enables "lateral movement" by the attacker, so that the attacker can move from the account originally compromised by phishing to other parts of the system of interest to the attacker (CrowdStrike 2023b).

Defending against spear phishing implicates "the who and the what" of an attacker's activities, such as identifying accounts and tracing an attack through a system, including by use of individuals' credentials. The concern would be if the adversary can move unobstructed while the defender is legally prohibited from seeing unique identifiers within an organization that are stored on networked endpoints in multiple countries.

i. Anatomy of Approach. Phishing is a term for emails or other communications that are designed to trick a user into believing they should provide a password, account number, or other information. The user then typically provides that information to a website controlled by the attacker. Spear phishing is a phishing attack that is tailored to the individual user, for example, when an email appears to be from the user's boss instructing the user to provide information.

In the White Paper entitled "Finding Cyber Threats with ATT&CK-Based Analytics," MITRE describes a hypothetical campaign involving spear phishing that we incorporate here to emphasize the impact of personal data in the scenario (Strom et al. 2017, 2).

In our example, consider an employee in the EU, working for a company that also operates in the U.S., who falls victim to a spear phishing attack. The employee downloads an executable, which downloads a second stage Remote Access Tool (RAT) payload, giving a remote operator access to the victim computer as well as an initial access point into the network. The adversary then uses tools already available on the compromised computer to learn more about the victim's system and the surrounding network (ibid, 2-4).

ii. Types of Personal Data. In defending against the privilege escalation attack, numerous steps of the attack could have personal data at issue. First, the attacker uses spear phishing to gain access, compromising one person's account (CrowdStrike 2022b). Next, the attacker exploits the credentials of the victim's account to explore the network and achieve lateral movement (CrowdStrike 2023b). The attacker's goal may be to extract particular high-value documents or to remain in the system as an advanced persistent threat (APT) (CrowdStrike 2023a).[7] For example, to achieve persistence, the attacker might create a fake account under an assumed identity, which will contain types of personal data from the defender's vantage point (MITRE: Create Account).

More generally, the attacker regularly creates digital footprints that may include personal data. Such footprints may be left, for instance, in: (i) operating system generated logs or; (ii) telemetry data captured by cybersecurity technologies such as ENISA-endorsed endpoint detection and response (EDR) (ENISA and TeleTrust 2021, Sec. 3.3.22). Such data may contain usernames, file names, IP addresses, and other sorts of personal data or non-personal but protected data.

For the defender, the MITRE framework examines both detection and defense. Approaches to detection of spear phishing include reviewing application log content, network traffic flow, and network traffic content (MITRE: Internal Spearphishing). Network traffic content has numerous possible approaches such as monitoring and analyzing traffic patterns. These patterns may include gratuitous or anomalous traffic patterns, anomalous syntax, anomalies in use of files that do not normally initiate connections for respective protocols); and monitoring network traffic for requests and/or downloads of container images (MITRE: Network Traffic).

Because successful spear phishing results in the attacker having access to a valid account, defenders may still be attempting to detect the attack after the attacker has entered the system (MITRE: APT28). The defender monitors logon sessions looking for suspicious activity. Such activity may include: one account logged into multiple systems simultaneously; multiple accounts logged into the same machine simultaneously; and accounts logged in at odd times or outside of normal business hours (MITRE: Valid Accounts). With regard to the types of data monitored, these monitoring activities would review data such as device ID, time of day, day of week, and geolocation. The patterns from the new logins would then be compared to the historical information to determine if deviation exists (ibid; MITRE: CAR-2013-10-001). This routine defender monitoring would apparently be degraded if the defender could no longer log and review parts of the system that happen to be in a different country. In addition, as discussed in *Risks*, cybersecurity services may be provided remotely, such as when a company operating in the EU staffs its cybersecurity team outside of the EU, hires a vendor who does so, or manages or relies upon global infrastructure.

iii. Alternatives to Use of Personal Data. When attempting to detect or defend against spear phishing, the user account suspected to be compromised is generally one of a person affiliated with the targeted organization, often at a high level. Much of the data reviewed during an investigation into such a cyber incident would be personal information, if actually associated

with the person rather than a system account, such as email address, IP address, and geolocation information. Certainly, looking at the historical data of the victim of the attack and building a profile of the victim's habits (and location) over time would be reviewing personal data. In a company that is international, data localization may be particularly impactful. If the person spear phished works in the EU, and the company is based in the U.S., the EU division of the IT department likely could not send information concerning the spear phishing attack to the main IT department in the U.S.

In some instances, it may be possible to create defenses against phishkits consistent with hard data localization. If detection of phishing operated, for instance, on hashed or encrypted versions of personal data, some defensive operations may still operate successfully. Such approaches, however, have not been widely deployed to date and may well be technically infeasible.

iv. Privilege Escalation and the ENISA Guidelines. Among other possibly relevant measures, the ENISA Guidelines on "state of the art" include two technical measures directly relevant to defending against privilege escalation.

First is Technical Measure - 3.2.22 - Endpoint Detection and Response: "Ideally, detections are correlated and the technique and tactics (including tools used such as malware, trojans, PowerShell scripting) and the attacker's target are displayed (exfiltration of data, setting up a backdoor, lateral movement within the organization, rights escalation, etc.) are displayed" (ENISA and TeleTrust 2021).

Second is Technical Measure - 3.2.29 Monitoring of Directory Services and Identity-Based Segmentation: "Which IT security threat(s) is the measure used against? Misuse of privileged accounts and escalation of authorization" (ENISA and TeleTrust 2021).

2. Impacts to Cybersecurity: Risks from Knowing Less Than the Attacker[8]

An essential part of good cyber defense is for the defenders to test the system through "red teaming," including penetration ("pen") testing. Pen testing may be more familiar to readers. It involves the defense hiring "white hat" attackers to find as many vulnerabilities and configuration issues as possible, exploit them, and determine risk levels (Talamantes). Red teaming is a more general approach, for the defender to identify physical, hardware, software, and human vulnerabilities. In addition to pen testing, red team skills include social engineering, threat intelligence, and reverse engineering (Coursera 2022).

a. Pen Testing and Other Red Teaming

Data localization laws would appear to present serious obstacles to pen testing and other red teaming. The intuition is that attackers will be willing to break the law, to seek out and transfer personal data across national borders. By contrast, defenders are obliged to comply with the law. If defenders hire pen testers, those testers would not be able to probe for vulnerabilities that would require learning account information and other personal data, notably where the data is stored in a different country. Since a large fraction of cyber-attacks involve crossing national

borders, the defenders would systematically be able to test and learn about vulnerabilities in their own systems less well than the attackers.

i. Anatomy of Approach. Red teaming identifies the risk and susceptibility of attack against key business information assets. The red team effectively simulates the techniques, tactics, and procedures of genuine threat actors, in a controlled manner, and with authorization from the defending organization. The red team assesses the organization's ability to detect, respond, and prevent sophisticated and targeted threats. The red team engages closely with the internal cybersecurity team, including the incident response team, to provide meaningful mitigation and comprehensive post-assessment debriefing (bsi).

Pen testing views the organization through the eyes of a bad actor, seeking to discover cybersecurity vulnerabilities. An effective penetration tester may identify where a hacker might target, how they would attack, how the defenses would fare, and the possible magnitude of a breach. At the conclusion of testing, pen testers generate detailed reports, including examples of successful attacks, screenshots, methodology, and remediation recommendations (Coursera 2022).

ii. **Types of Personal Data.** In order to understand the extent to which pen testing and other red team activities might be affected due to hard data localization, we carried out an academic exercise where we identified what personal data might be needed by defenders to emulate an APT or to detect if any adversarial techniques are currently being employed within an organization. In order to do a systematic and exhaustive study, we used the MITRE ATT&CK framework to help us walk through the techniques employed in each of the 14 tactics and see how many of those would be impacted if personal data were to be removed. Here, we analyzed the techniques for each tactic to deduce what personal data would be needed to detect and defend against them. Based on the nature of information, we identified that the personal data leveraged here would comprise of one or many of the following: IP addresses, email addresses, domains, social media profiles, digital certificates, access tokens, etc.

Our analysis led us to the conclusion that 13 out of the 14 techniques (all tactics except 'Execution') would be negatively impacted by removal of personal data. By impacted, we mean that it would hinder information sharing for cyber defense purposes.

iii. Alternatives to Use of Personal Data. Thus far, we have not identified effective alternative strategies for conducting pen testing and red teaming, in the absence of the ability to see specific identifying information concerning individual accounts, file names, IP addresses, log entries, and other information that a pen test would routinely access. In essence, cybersecurity is naturally reliant upon the very protocols and identifiers inherent to modern computing.

iv. Red Teaming, Pen Testing, and the ENISA Guidelines. Numerous legal regimes have recommended or mandated penetration testing as an essential component of an organization's overall cybersecurity program. Article 32 of GDPR states that companies must regularly test, assess, and evaluate the effectiveness of technical and organizational measures that

ensure the security of data. ISO 27001 provides that information about technical vulnerabilities "shall be obtained in a timely fashion" and remediated to address the associated risk (isms.online; ibid, Standard A.12.6.1). Requirement 11 of PCI DSS specifically mandates the performance of regular penetration testing for service providers and large merchants (ERMProtect). SOC 2, in CC4.1 and CC7.1, has specific requirements that mention penetration testing and vulnerability management for auditors to review (AICPA) (Fowler). There are specific provisions concerning pen testing in financial services regimes, including under FINRA, SWIFT, and the New York state law governing financial institutions (FINRA; NYAG; SWIFT).

More specifically, ENISA has also noted the effectiveness of penetration testing and related techniques. The ENISA Guidelines include Technical Measure – 3.3.2.12 Technical System Audits: Technical system audits (inspections at the network, system and application level). Such audits "must be performed regularly by or on behalf of the organisation. These are typically carried out as penetration tests or web checks." It adds: "For a comprehensive IS penetration test, in addition to the technical audit, vulnerabilities in the IT systems tested are rooted out through technical investigations using special security tools, among other things. In doing so, the testers access the IT systems to be inspected on site under supervision by the administrators" (ENISA and TeleTrust 2021, 77).

The ENISA Guidelines also include Operational Measure – 3.36 Management of Information Security Risk. It states: "Hardly less difficult is the estimation of probabilities of occurrence. It is advisable to use as many external and internal sources of information as possible. The former includes CVE[22] lists, vendor information, CERT services (e.g., from the BSI), and the latter include the evaluation of information security incidents, penetration tests, audits or awareness measures" (ibid, 89).

IV. Modeling Quantitative Effects of Data Localization

We next present a model for estimating, in one setting, the quantitative effects of data localization. We first explain reasons why data localization would likely increase the time needed for defenders to spot a new attack. We then provide a quantitative model that indicates that the time for detection will more than double if the Internet is partitioned in half due to localization rules.

A. Data Localization and the Speed of Detection Matters

As described earlier in the paper, data localization would result in a reduction in observable telemetry. Telemetry in this setting is "data collected from a network environment that can be analyzed to monitor the health and performance, availability, and security of the network and its components, allowing network administrators to respond quickly and resolve network issues in real-time" (BlackBerry). Examples, among many others, include data from Intrusion Detection and Intrusion Prevention systems, and netflows into and out of a system (Lebovitz 2021). Telemetry also includes data about execution, file transfers, configurations, and other observable activity on an endpoint (Karantzas 2021).

Prior research has shown a strong relationship between reducing observable telemetry and the speed of detecting and responding to attacks. Network Telescopes have been previously used to detect Distributed Denial of Service (DDoS) attacks and Internet worm propagation patterns (Moore et al. 2001; Moore et al. 2002). The size of the Network Telescope, i.e. the number of IP addresses used for observation, has a significant impact on the telescope's efficacy (Moore et al. 2004). Similar network effects leveraging global insights and observations have also been used to stop Unsolicited Bulk Email (UBE, "spam") and Business Email Compromise (BEC) (Tang et al. 2008). Beyond network detection, endpoint detection and response (EDR) provides a core set of cybersecurity telemetry used by defenders to is detect adversary activity on devices, including computers, virtual machines, and cloud containers (Karantzas 2021).

Reducing the scope of monitoring can reduce the efficacy of cybersecurity in various ways. First, the activities of an adversary can occur outside of the monitored footprint, slowing detection. Second, the fidelity of monitored quantities can decline. The prevalence of a measured quantity contributes to multivariate analysis using approaches such as machine learning (ML); therefore, reducing quantity reduces the accuracy of inferences about cybersecurity risks. Third, a smaller monitored footprint results in a smaller dataset that can be used for ML training. A smaller dataset especially hobbles modern deep learning-based approaches that require large amounts of data to establish baselines.

Far from achieving state of the art cybersecurity, the reduction in observable telemetry from data localization would likely cause delays in detecting and responding to attacks. Speed is vital to detecting and containing the adversary. Industry data shows the "breakout time," the time until an adversary moves laterally after initial access, averages about one and a half hours (CrowdStrike 2022c). Once the adversary moves laterally, the attack is harder to contain. In light of the high costs from a data breach, moving quickly is vital to limiting damage (IBM 2022). Recognizing the importance of speed in mitigating cybersecurity risks, the U.S. Cyberspace Solarium Commission noted that:

> A company's ability to rapidly, detect, investigate, and remediate network intrusions is a useful indicator of the maturity of its security operations, in its ability both to defend against cyberattacks and to mitigate the types of cybersecurity risks that could harm its business operations and financial conditions (Cyberspace Solarium Commission 2020).

The need for speed is also derived from regulatory requirements such as GDPR's 72-hour breach notification requirement. Consequently, how quickly a defender is ability to collect, analyze, and leverage security-related telemetry data is a key component of modern cybersecurity.

B. A Model for Reconnaissance and Initial Access

To provide a quantitative assessment of the impact of localization requirements on the detection of Advanced Persistent Threats (APTs), we present a high-level model of adversary scanning behavior during reconnaissance.

We assume an adversary is scanning a list of 100 million IP addresses (N) for a zero-day vulnerability. We further assume that we need to observe the adversary communicate with a

vulnerable system to distinguish it from a common scan. Furthermore, a cloud-based protection platform is protecting K=100,000 of these systems. These systems may be endpoints that run a sensor software, which communicates and coordinates defenses using a global centralized cloud platform. We successfully detect the campaign when the first protected system is contacted by the adversary. From that point forward, mitigations across the population of vulnerable systems can be taken. Hence, the faster we can react, the better.

Localization requirements may force the operator of the centralized cloud to segment the monitored footprint into several isolated domains. As a result, each segmented cloud would have fewer sensor-protected systems available to detect the campaign. For the sake of this analysis, we assume that we segment the cloud into two domains of equal size with $K/2$ systems each.

We assume the adversary scans at a rate r of 60 probes per hour (i.e., one per minute). This rate is based on the adversary's strategy of evading volume-based detection – detection would become easier if the adversary used a higher rate of probes.

To model the problem, we use a hypergeometric distribution, i.e. drawing from a total population N with K instances allowing for detection. The probability that after $n=rt$ probes we achieved k detections is given by:

$$\Pr(X = k) = \frac{\binom{n}{k}\binom{N-n}{K-k}}{\binom{N}{K}}$$

The probability that at time t we achieved more than zero detects is given by:

$$\Pr(X > 0) = 1 - \Pr(X = 0) = 1 - \frac{\binom{n}{0}\binom{N-n}{K}}{\binom{N}{K}} = 1 - \frac{\binom{N-rt}{K}}{\binom{N}{K}}$$

Figure 1 graphs the results of our model. The solid line shows the probability of detecting the attack where the defense can use the full set of sensors. In this scenario, achieving a minimum 80% probability of protection requires 27 hours of observation. The dashed line shows the probability of detecting the attack where localization enables the defense to see only half ($K/2$) of the sensors. To achieve the same 80% probability, it would now take 55 hours.
In summary, using the sort of plausible, simple model seen previously in the literature, reducing the number of sensors in half would result in average detection time for an attack taking more than twice as long.

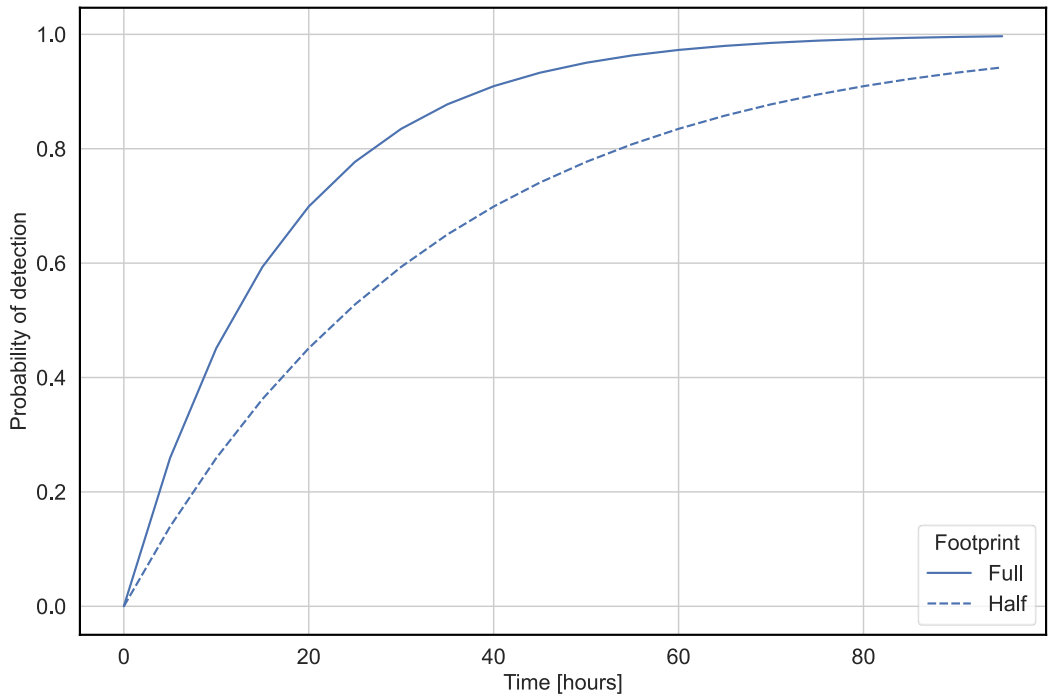

Figure 1: probability of a sensor monitoring a target chosen by the adversary vs time in hours

V. Assessing European Cybersecurity Certification Regimes Requiring Localization

We next turn to recent measures and proposals in the EU to require data localization, justified in the name of improving cybersecurity. We discuss the certification known as SecNumCloud adopted in France, as well as proposals by ENISA and Italian authorities. In light of the multiple and significant risks to cybersecurity from localization, discussed in *Risks* and this paper, it is logical that such measures and proposals should at a minimum consider the risks to cybersecurity from localization, along with consideration of claimed benefits.

ENISA is currently considering the European Union Cybersecurity Certification Scheme on Cloud Services (EUCS). The EU Cybersecurity Act of 2019 called on ENISA to assist with the preparation of 'candidate cybersecurity certification schemes.' ENISA launched a public consultation in 2020 and proposed its first draft of the EUCS in 2021 (ENISA: Certification). The task for the EUCS is to provide a voluntary, EU-wide framework for the certification of the cybersecurity of cloud services. The certification is supposed to counter fragmentation between the EU member states, while facilitating trade and understanding of security features by harmonizing the security of cloud services with EU regulations, international standards, best industrial practices, and existing certifications in EU Member States. Although the certification would be generally voluntary, the high assurance level is expected to become mandatory for the

essential services listed under the EU Network and Information Security 2 (NIS2) Directive (EU Directive 2022/2555, 80-152).

ENISA has considered basing the EUCS on the cybersecurity certification known as SecNumCloud, developed by France's national cybersecurity agency, ANSSI, in 2016 (ANSSI). As updated in 2022, SecNumCloud has two related localization requirements (Prime Minister of France 2021a). First, it requires defined cloud services and other organizations to prohibit data and system access from organizations located outside the EU. This requirement requires data to be stored locally and use only local support and technical staff (Cory 2021). Second, it requires cloud providers to be "immune to any extra-EU regulation," with strict limits on foreign ownership and representation on a company's board of directors (Propp 2022; Prime Minister of France 2021b; Cory 2021).

Several member states have opposed this approach, which would prohibit Software as a Service and cloud services generally that store data outside of the EU (Bertuzzi 2021). The U.S. government has raised concerns about possible violation of international trade agreements (Propp 2022). Other EUCS' critics have described "limited transparency and lack of stakeholder engagement" in ENISA's drafting process, and say ENISA should focus instead on "the actual technicalities of cybersecurity" rather than base cloud service provision on national origin (Digital Europe 2022; Cory 2023).

Going beyond certifications, Italy considered but later rejected a draft presidential decree with strict localization rules for cybersecurity services. As originally drafted, Italy would have implemented the 2017 European Network and Information Security Directive (NIS I) to set requirements on functions and services covered by its National Cybersecurity Perimeter legislation (EU Directive 2022/2555, 80-152). The original requirement would have effectively meant that organizations deemed part of Italy's cybersecurity perimeter could only adopt the cybersecurity technologies and practices endorsed by ENISA if the requisite infrastructure and workforce was solely in Italy. According to FAQs issued by the Italian National Cybersecurity Agency (Agenzia per la Cybersicurezza Nazionale), the proposed localization constraints were aimed at facilitating interactions with the supervisory authority in case of incidents having an impact on national security. The stated goals, among others, were to verify the implementation of security measures through physical on-site inspections, as well as verification and assessment of possible causes of incidents.

As discussed in this paper, such localization proposals run counter to the cybersecurity "state of the art," as set forth by ENISA and other government agencies. Moving forward, our research suggests that ENISA should, at a minimum, consider the risks to cybersecurity from localization before adopting localization measures (Digital Europe 2022). In addition, transfers of data for cybersecurity purposes can be fostered within the ongoing political dialogue supporting Data Free Flow with Trust (DFFT), the approach proposed by former Japanese Prime Minister Abe at the World Economic Forum (Davos Conference) in January 2019 (WEF 2020; Arasasingham and Goodman 2013). The DFFT concept promotes the international free flow of

data useful for addressing business and social issues while ensuring trust in privacy, security and intellectual property rights. Consistent with the DFFT approach, the Organisation for Economic Co-operation and Development (OECD) in 2022 published its Declaration on Government Access to Personal Data Held by Private Sector Entities. This Declaration announced common principles for government access to data held by the private sector, to safeguard privacy when accessing personal data for national security and law enforcement reasons.

VI. Conclusion

In our first paper, *Risks*, we provided a framework for understanding the risks of data localization on cybersecurity, based on effects within an organization, across organizations for payment, and across organizations without payment. This paper complements that analysis, with greater focus on technical measures, for the techniques, tactics, and procedures of threat actors and defenders. We have used the ENISA Guidelines and the MITRE ATT&CK Framework as authoritative approaches for cataloguing relevant TTPs. In this paper, we have focused on the example of data localization in the European Union, but similar analysis would apply to any countries contemplating such a localization regime that restricts data transfers.

We have used examples to highlight two themes for when data localization laws appear to pose particularly severe obstacles to cybersecurity. The first theme concerns "the who and the what" of attackers. Threat hunting and threat intelligence are core activities for defenders, but they involve analysis of identifying information, including account names, IP addresses, and many other types of potentially personal data. Our other example concerns privilege escalation, where attackers seek to move laterally in an organization to reach their objectives. As the spear phishing example illustrates, organizations analyze telemetry and information of many sorts, to detect initial intrusions and follow clues across the organization to detect and then respond to APTs and other intruders (OECD 2022).[9]

The second theme concerns pen testing and other forms of red teaming. Put simply, there are risks where defenders know less than attackers. Yet data localization laws block pen testing and other forms of red teaming whenever the probe moves from part of the organization (in one country) to another part of the organization (in another country). The same analysis applies if the red teaming applies to the increasingly important portion of cybersecurity focused on supply chain. Organizations today often purchase services and infrastructure from other organizations, in ways that implicate the purchaser's cybersecurity if there are vulnerabilities in the supply chain. Effective red teaming today includes a comprehensive approach to an organization's risks, including from vendors. Thus, even where a company operates only in one jurisdiction, there are often dependencies on other jurisdictions. Even though pen testing is expected or required for many organizations, data localization laws thus put at risk the effectiveness of such pen testing.

As we continue this research, we welcome comments and suggestions about other ways that data localization laws may affect risks to defenders' TTPs. For now, we close with three implications of the research.

First and most generally, we recommend that cybersecurity experts and government agencies examine the risks detailed in this paper. For instance, before ENISA takes any action to localize cybersecurity services, we believe it important for ENISA to consider how any proposal would impact the state of the art mandated by ENISA, for activities such as threat hunting, preventing escalation of privileges, and red teaming/pen testing. Our research to date has not discovered any such analysis by ENISA. Relatedly, we have not thus far seen discussion by ENISA of how Article 32 of GDPR and ENISA's state-of-the-art requirements can be achieved consistent with the strict sort of localization that data protection regulators have supported in recent enforcement cases.

Second, where policymakers decide in favor of data localization, we urge consideration of creating cybersecurity exceptions. Such exceptions might be relatively general, such as use of personal data for cybersecurity purposes. Alternatively, exceptions could be more targeted, such as permitting use of personal data for pen testing, incident response, and other specific purposes.

Third, the risks to cybersecurity from localization – including effects on individuals, corporations, and national security – should be analyzed together with any claimed benefits. The claimed benefits of localization may include less lawful access by governments and other actors who seek data held outside of the country. Empirically, it is far from clear whether risk systematically increases with data transfer, or that most types of data shared for cybersecurity purposes would actually be of any interest to other governments. Whatever the actual risks from transferring data, it seems irrational to use data localization as a proxy, or even pillar, for data protection and to focus only on possible benefits from restricting data flows while ignoring known, likely, and apparently substantial risks to cybersecurity. This has the unintended effect of disincentivizing the adoption of practical, EU-endorsed cybersecurity best practices. In sum, until and unless proponents of localization address these concerns, scholars, policymakers, and practitioners have strong reason to expect significant cybersecurity harms from hard localization requirements.

Notes

[1] The statements in this document are solely by the authors and should not be attributed to the Cross-Border Data Forum, CrowdStrike, or any client. For research support on this project, the authors thank the Center for International Business and Education at Georgia Tech, the Cross-Border Data Forum, the Georgia Tech Scheller College of Business, the Georgia Tech School of Cybersecurity and Privacy, and Microsoft. The authors thank Nathan Lemay for his substantial initial research contributions to this paper.

[2] After (*Breyer*, 2020), static IP addresses could fall in the scope of personal data within the meaning of Directive 95/46/EC, as far as they provide sufficient information on the history of a user and make it possible to identify him.

[3] GDPR is based on "a risk-based approach in terms of its protection objectives."

[4] See (Bagley, 2022). "Four Takeaways as the European Union's General Data Protection Regulation (GDPR) Turns 4." *Security Senses*, May 26. https://securitysenses.com/posts/four-takeaways-european-unions-general-data-protection-regulation-gdpr-turns-4. By design, the state of the art is not a static requirement, as cybersecurity risks evolve rapidly.

[5] Indicators of Compromise (IOCs) include file hashes, IP addresses, and domain names. An important distinction between a technique in ATT&CK and an IOC is that many of the ATT&CK techniques are legitimate system functions that can be used for malicious purposes, whereas an IOC deployed as an intrusion detection mechanism is typically an indication of an action known to be caused by or under the influence of an adversary.

[6] The authors thank Nathan Lemay for his early suggestion to focus on the threat hunting/threat intelligence example.

[7] Commenters have asked us whether personal information about the attacker would also be covered by data localization laws. To date, we are not aware of any data localization law that would ban transfers generally but allow transfers to detect criminal cyber-attacks. We note, however, the analogy to the "hacker trespasser" provision in Section 217 of the USA-PATRIOT Act. https://www.justice.gov/archive/ll/subs/add_myths.htm#s217. Under that provision, the owner of the computer system may request law enforcement assistance to monitor trespassers in the system, without violating otherwise-applicable wiretap laws that would prohibit providing the information to law enforcement. In both settings, it would seem perverse to protect the hacker/trespasser's personal data from the system owner and law enforcement seeking to counter the criminal intrusion.

[8] The co-author Avani Modak played the leading role on researching red teaming and pen testing.

[9] Another unintended consequence of data localization is that attackers and their wrongful activity may be protected by the data protection regimes, making it harder to detect their activity. Attackers thus may seek to locate their activity hubs (either actual or appear to be located) within countries or regions with strict data localization mandates.

Notes On Contributors

Peter Swire is J.Z. Liang Chair, Georgia Tech School of Cybersecurity and Privacy, and Professor of Law & Ethics, Georgia Tech Scheller College of Business. He is Research Director of the Cross-Border Data Forum and senior counsel with Alston & Bird LLP.

DeBrae Kennedy-Mayo is a faculty member in the Georgia Tech Scheller College of Business and a Senior Fellow, the Cross-Border Data Forum.

Drew Bagley is Vice President & Counsel, Privacy and Cyber-Policy, CrowdStrike and an adjunct faculty member, American University School of Public Affairs.

Sven Krasser is Senior Vice President & Chief Scientist, CrowdStrike.

Avani Modak is a Masters in Cybersecurity graduate of the Georgia Tech School of Cybersecurity and Privacy.

Christoph Bausewein is Assistant General Counsel, Data Protection & Policy, CrowdStrike.

References

AICPA. 2017 Trust Services Criteria for Security, Availability, Processing Integrity, Confidentiality, and Privacy, (March 2020)

https://us.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/trust-services-criteria-2020.pdf

ANNSI. "European Secure Cloud – A New Label for Cloud Service Providers." https://www.ssi.gouv.fr/en/actualite/european-secure-cloud-a-new-label-for-cloud-service-providers/.

Arasasingham, Aidan and Goodman, Matthew. 2013. "Operationalizing Data Free Flow with Trust (DFFT)." *Center for Strategic and International Studies*, April 13. https://www.csis.org/analysis/operationalizing-data-free-flow-trust-dfft.

Austrian Data Protection Authority (ADPA) decision, as issued on 22 December 2021. https://noyb.eu/sites/default/files/2022-01/E-DSB%20-%20Google%20Analytics_DE_bk_0.pdf (original German), https://noyb.eu/sites/default/files/2022-01/E-DSB%20-%20Google%20Analytics_EN_bk.pdf (unofficial English translation)

Bagley, Drew. 2022. "Four Takeaways as the European Union's General Data Protection Regulation (GDPR) Turns 4." *Security Senses*, May 26. https://securitysenses.com/posts/four-takeaways-european-unions-general-data-protection-regulation-gdpr-turns-4.

Baker, Kurt. 2023. "What is Cyber Threat Intelligence?" *CrowdStrike*, March 23. https://www.crowdstrike.com/cybersecurity-101/threat-intelligence.

Bertuzzi, Luca. 2021. "Germany calls for political discussion on EU's cloud certification scheme." *Euractiv*, September 21. https://www.euractiv.com/section/cybersecurity/news/germany-calls-for-political-discussion-on-eus-cloud-certification-scheme/.

BlackBerry. "Telemetry for Cybersecurity." https://www.blackberry.com/us/en/solutions/endpoint-security/extended-detection-and-response/telemetry.

bsi. "What is red teaming and what are the benefits to my business?" https://www.bsigroup.com/en-IE/Blog/digital-trust--blog/what-is-red-teaming-and-the-benefits-to-organizations/#:~:text=A%20red%20team%20assessment%20is,the%20business%20into%20the%20future.

Christakis, Théodore. 2020. "'European Digital Sovereignty': Successfully Navigating Between the 'Brussels Effect' and Europe's Quest for Strategic Autonomy." December 7. doi: 10.2139.

Cory, Nigel and Dascoli, Luke. 2021. "How Barriers to Cross-Border Data Flows are Spreading Globally, What They Cost, and How to Address Them." *Information Technology & Innovation Foundation*, July 19. https://itif.org/publications/2021/07/19/how-barriers-cross-border-data-flows-are-spreading-globally-what-they-cost/.

Cory, Nigel. 2021. "'Sovereignty Requirements' in French—and Potentially EU—Cybersecurity Regulations: The Latest Barrier to Data Flows, Digital Trade, and Digital Cooperation Among Likeminded Partners." *Information Technology & Innovation Foundation*, December 10. https://itif.org/publications/2021/12/10/sovereignty-requirements-france-and-potentially-eu-cybersecurity/.

Cory, Nigel. 2023. "Europe's Cloud Security Regime Should Focus on Technology, Not Nationality." *Information Technology & Innovation Foundation*, March 27. https://itif.org/publications/2023/03/27/europes-cloud-security-regime-should-focus-on-technology-not-nationality/.

Coursera. 2022 "Red Team vs. Blue Team in Cybersecurity." November 1. https://www.coursera.org/articles/red-team-vs-blue-team.

CrowdStrike, "Indicators of Attack (IOA) v. Indicators of Compromise (IOC)." https://www.crowdstrike.com/resources/white-papers/indicators-attack-vs-indicators-compromise/.

CrowdStrike. 2022c. "Global Threat Report 2022." https://www.crowdstrike.com/global-threat-report/.

CrowdStrike. 2022s. “IOA VS IOC.” October 5. https://www.crowdstrike.com/cybersecurity-101/indicators-of-compromise/ioa-vs-ioc/.

CrowdStrike. 2022b. “What is Privilege Escalation.” June 3. https://www.crowdstrike.com/cybersecurity-101/privilegeescalation/#:~:text=A%20privilege%20escalation%20attack%20is,operating%20systems%20or%20web%20applications.

CrowdStrike. 2023a. “Advanced Persistent Threat (APT).” February 28. https://www.crowdstrike.com/cybersecurity-101/advanced-persistent-threat-apt/.

CrowdStrike. 2023b. “Lateral Movement.” April 17. https://www.crowdstrike.com/cybersecurity-101/lateral-movement/.

CrowdStrike. 2023c. “Global Threat Report 2023.” https://www.crowdstrike.com/global-threat-report/.

Cybereason Global SOC Team. 2022. “Threat Analysis Report: DLL Side-Loading Widely (Ab)Used,” *cybereason*, October 26. https://www.cybereason.com/blog/threat-analysis-report-dll-side-loading-widely-abused.

Cyberspace Solarium Commission, Cyberspace Solarium Commission Report, March 2020, p. 83, https://www.solarium.gov/report

Declaration on Government Access to Personal Data Held by Private Sector Entities. 2022. *Organisation for Economic Cooperation and Development (OECD)*. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0487.

Digital Europe. 2022. “Joint Letter on ‘sovereignty requirements’ in candidate European Cybersecurity Certification Scheme for Cloud Services.” June 16. https://digital-europe-website-v1.s3.fr-par.scw.cloud/uploads/2022/06/DIGITALEUROPE_Joint-letter-on-%E2%80%98sovereignty-requirements-in-candidate-EUCS.pdf.

Directive 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive) , OJ L 333, 27.12.2022, p. 80-152.

Directive 2022/2557 of the European Parliament and of the Council of 14 December 2022 on the resilience of critical entities and repealing Council Directive 2008/114/EC, OJ L 333, 27.12.2022, p. 164-198.

Discreto del Presidente del Consiglio dei Ministri 14 Aprile 2021. https://www.gazzettaufficiale.it/eli/id/2021/06/11/21G00089/sg, n. 81.

EDPB. 2021. “Recommendations 01/2020 on measures that supplement transfer tools.” June 18. https://edpb.europa.eu/system/files/2021-06/edpb_recommendations_202001vo.2.0_supplementarymeasurestransferstools_en.pdf.

ENISA. 2019. “What is ‘State of the Art” in Cybersecurity?” February 7. https://www.enisa.europa.eu/news/enisa-news/what-is-state-of-the-art-in-it-security.

ENISA and TeleTrust. 2021. “IT Security Act (Germany) and EU General Data Protection Regulation: Guideline ‘State of the Art’ Technical and Organisational Measures.” https://www.teletrust.de/fileadmin/user_upload/2021-09_TeleTrusT_Guideline_State_of_the_art_in_IT_security_EN.pdf.

ENISA. “The European Union through ENISA is developing EU cybersecurity certification which provides evidence of compliance to a given level of trust.” https://www.enisa.europa.eu/topics/certification.

ERMProtect Staff. “Penetration Testing for Compliance.” *ERMProtect*. https://ermprotect.com/blog/penetration-testing-for-compliance/.

European Commission. 2022. “Proposal for a Regulation of the European Parliament and of the Council of 23 February 2022 on harmonised rules on fair access to and use of data (Data Act).” Brussels, COM (2022) 68 final.

European Commission. “What is Personal Data.” https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_en.

European Data Protection Board (EDPB). 2023. “Guidelines 9/2022 on personal data breach notification under GDPR version 2.0.” March 28. https://edpb.europa.eu/system/files/2023-04/edpb_guidelines_202209_personal_data_breach_notification_v2.0_en.pdf.

European Union Agency for Cybersecurity (ENISA). 2020. “Cloud Certification Scheme: Building Trusted Cloud Services Across Europe.” December 22. https://www.enisa.europa.eu/news/enisa-news/cloud-certification-scheme.

Falcon OverWatch Team. 2021. “Nowhere to Hide: 2021 Threat Hunting Report.” *CrowdStrike*. https://go.crowdstrike.com/rs/281-OBQ-266/images/Report2021ThreatHunting.pdf, p. 33.

Financial Industry Regulatory Authority (FINRA). 2018. “Report on Selected Cybersecurity Practices – 2018.” December. https://www.finra.org/sites/default/files/Cybersecurity_Report_2018.pdf.

Fowler, Adam. “SOC 2 CC4: Common Criteria related to Monitoring Activities.” *Design Compliance and Security*. https://www.designcs.net/soc-2-assessments-common-criteria-related-to-monitoring-activities/.

French Commission nationale de l'informatique et des libertés (CNIL) decision, as issued on 10 February 2022. https://www.cnil.fr/sites/default/files/atoms/files/decision_ordering_to_comply_anonymised_-_google_analytics.pdf

Future of Privacy Forum. 2014. “MAC Addresses and De-Identification.” March 27. https://fpf.org/blog/mac-addresses-and-de-identification/.

Government of France: Office of the Prime Minister. 2021a. “Circular No. 6282-SC of July 5, 2021 relating to the doctrine for the use of cloud computing by the State.” July 5. https://www.legifrance.gouv.fr/circulaire/id/45205?page=1&pageSize=10&query=*&searchField=ALL&searchType=ALL&sortValue=PUBLI_DATE_DESC&tab_selection=circ&typePagination=DEFAULT.

Government of France: Office of the Prime Minister. 2021b. “SecNumCloud.” https://www.ssi.gouv.fr/uploads/2021/10/anssi-referentiel_exigences-secnumcloud-v3.2.a.pdf.

IBM. 2022. “Cost of a Breach Report 2022.” https://www.ibm.com/resources/cost-data-breach-report-2022.

In Case C-311/18, Data Protection Commissioner v. Facebook Ireland Ltd, Maximillian Schrems (16 July 2020), Court of Justice of the European Union.

In Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland (19 October 2020), Court of Justice of the European Union.

isms.online. “ISO 27001 Annex A.12.1.” https://www.isms.online/iso-27001/annex-a-12-operations-security/.

Karantzas G, Patsakis C. An Empirical Assessment of Endpoint Detection and Response Systems against Advanced Persistent Threats Attack Vectors. *Journal of Cybersecurity and Privacy*. 2021; 1(3):387-421. https://doi.org/10.3390/jcp1030021

Kime, Chad. 2023. “Top 7 Threat Intelligence Platforms.” *eSecurity Planet*, February 10. https://www.esecurityplanet.com/products/threat-intelligence-platforms/.

Lebovitz, Gregory. 2021. “Network security threat detection – Comparison of analytics methods.” *Google Cloud*, September 16. https://cloud.google.com/blog/products/networking/when-to-use-5-telemetry-types-in-security-threat-monitoring.

MITRE. “APT28.” https://attack.mitre.org/groups/G0007/.

MITRE. “Create Account.” https://attack.mitre.org/techniques/T1136/.

MITRE. “Enterprise Matrix.” https://attack.mitre.org/matrices/enterprise/.

MITRE. “Internal Spearphishing.” https://attack.mitre.org/techniques/T1534/.

MITRE. “Network Traffic.” https://attack.mitre.org/datasources/DS0029/#Network%20Traffic%20Content.

MITRE. “Updates April 2023.” https://attack.mitre.org/resources/updates/updates-april-2023.

MITRE. “Valid Accounts.” https://attack.mitre.org/techniques/T1078/.

Moore, David, Colleen, Shannon, and Brown, Jeffery. 2002. "Code-Red: a case study on the spread and victims of an Internet worm." *Internet Measurement Workshop (IMW)*. https://catalog.caida.org/paper/2002_codered.

Moore, David et al. 2004. "Network Telescopes: Technical Report." *Cooperative Association for Internet Data Analysis (CAIDA)*. https://catalog.caida.org/paper/2004_tr_2004_04.

Moore, David, Voelker, Gregory and Stefan Savage. 2001. "Inferring Internet Denial-of-Service Activity." *USENIX Security Symposium*. https://catalog.caida.org/paper/2001_backscatter.

Moulinos, Konstantinos and Pauna, Adrian. 2013. “Good practice framework for an EU ICS testing coordination capability.” *ENISA*, December. https://icscsi.org/library/Documents/Best_Practices/ENISA%20-%20Good%20Practices%20for%20an%20EU%20ICS%20Testing%20Coordination%20Capability.pdf.

National Institute of Standards and Technology (NIST). “tactics, techniques, and procedures.” https://csrc.nist.gov/glossary/term/tactics_techniques_and_procedures.

New York State Attorney General (NYAG). “Report a Data Breach.” https://ag.ny.gov/internet/data-breach.

Portugal National Data Protection Commission. 2021. “CNPD Suspends Flows to the US.” April 27. https://www.cnpd.pt/comunicacao-publica/noticias/censos-2021-cnpd-suspende-fluxos-para-os-eua/.

Propp, Kenneth. 2022. “Cybersecurity Regulation Takes a Sovereign Turn.” *European Law Blog*, September 12. https://europeanlawblog.eu/2022/09/12/european-cybersecurity-regulation-takes-a-sovereign-turn/.

Regulation (EU) 2016/679 of the European Parliament of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (2016) OJ L 199, Rec. 49.

Regulation 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act), OJ L 151, 7.6.2019, p. 15-69 .

Selzer, Annika. 2021. “The Appropriateness of Technical and Organisational Measures under Article 32 GDPR.” *European Data Protection Law Review*, 2021(1): 120-128. doi:10.21552/edpl/2021/1/16.

Strom, Blake et al. 2017. “Finding Cyber Threats with ATT&CK-Based Analytics.” *MITRE*, June 22. https://www.mitre.org/publications/technical-papers/finding-cyber-threats-with-attck-based-analytics.

Swire, Peter & Kennedy-Mayo, DeBrae. 2022. “The Risks of Data Localization on Cybersecurity – Organizational Effects. SSRN, June 22. https://ssrn.com/abstract=4030905.

Talamantes, Jeremiah. “Penetration Testing vs. Red Teaming: What’s the Difference?” *Red Team Secure*. https://www.redteamsecure.com/blog/penetration-testing-vs-red-teaming.

Tang, Yuchun et al. 2008. “Support Vector Machines and Random Forests Modeling for Spam Senders Behavior Analysis.” *Proceedings of the IEEE Global Communications Conference*. doi: 10.1109.

Taschler, Scott. 2023. “What is Cyber Threat Hunting?” CrowdStrike, April 17. https://www.crowdstrike.com/cybersecurity-101/threat-hunting/.

World Economic Forum (WEF). 2020. “Data Free Flow with Trust (DFFT): Paths towards Free and Trusted Data Flows.” June 10. https://www.weforum.org/whitepapers/data-free-flow-with-trust-dfft-paths-towards-free-and-trusted-data-flows.