SCTE CABLE-TEC EXPO 22
SEPTEMBER 19-22 • PHILADELPHIA, PA

2022 Fall Technical Forum
SCTE • CABLELABS • NCTA

Privacy Posture 2022

The Intersection of Technology, Policy, Standards, and Security

A Technical Paper prepared for SCTE by

Brian Scriber
Distinguished Technologist & Vice President Security and Privacy Technologies
CableLabs
858 Coal Creek Circle, Louisville, CO 80027

b.scriber@cablelabs.com

Table of Contents

List of Figures

List of Tables

1. Introduction

Privacy Engineering: The Intersection of Technology, Policy, Standards, and Security

While the term "security" is often used interchangeably with "privacy," these two disciplines require different skillsets and process. Privacy focusses on access control and data usage. The toolset for the nascent privacy engineering field requires awareness of the development in the technology sector, awareness of risks and threats to protected data, privacy law in relevant jurisdictions, and different privacy standards, as these standards help set the bar as to what constitutes appropriate privacy protections. Security is focused on data access, integrity, and confidentiality. Organizations holding data need to be aware of the business and regulatory risks associated with this data and shape their internal privacy engineering and privacy policy compliance teams to mitigate these risks. This paper addresses the data security and privacy landscape for 2022 to provide some assistance in assessing your organizations' security and privacy posture.

2. Security and Privacy

Security and privacy get grouped together, it's easy to think of them as the same thing, they are not. Privacy is about the decisions that surround competing claims for access to, modification of, deletion of, and altering the disposition of information (Bambauer, 2013), as well as the legal right, uses, and potential ownership of data. Security, by the other token, is about the confidentiality, the integrity, and the availability of the data being accessible to only those people or system with appropriate identity and credentials.

In the cable ecosystem, a network operator occupies an interesting position, the operator helps to secure the data in transit through technologies like the CableLabs® DOCSIS Security Specification, protecting the confidentiality and integrity of the data through encryption, hashing, and message authentication. At higher levels of the OSI stack, over 92% of US web traffic transiting the internet today uses HTTPS (Google, 2022) and encrypting from endpoint to endpoint.

In many cases, the tools used to deliver security can be used to protect privacy, an example of this is how confidentiality in the above two examples protects from unauthorized users seeing message contents; recall, however, that privacy is about the decisions and competing claims to access information. An encrypted network does not help protect a user from endpoints and services that are going to sell user behavior and identification. The endpoint operator may claim that those observations and data were collected using their hardware, software systems, and algorithms, and how they handle the disposition of those data is their prerogative. The user, on the other hand, may claim that the observations collected about them, and their identification data is personal, and that the expectation was that the interactions made were not reasonably, or lawfully, expected to be shared with others or were used in ways that were not in the interests of the user.

Those differing claims create a new privacy discipline, new technologies and new tools unique to privacy engineering that are not exclusively within the security discipline. This work addresses the current state of how we capture, classify, and protect these aspects and nuance of data.

3. Privacy Enabling Technology

Privacy Engineering, as a discipline, is nascent, as is the supporting ecosystem of products and tooling. Some of the tools in this area begin to show how their application and outcomes drive advancement in data protection and privacy compliance.

3.1. Privacy Compliance

Compliance can depend on the legal jurisdiction, the domain data is held within, and the status of who holds the data (IAPP 2022). A hospital in Tennessee running field trials of new treatments will have different compliance hurdles from a Californian network operator or a direct-to-consumer retailer in Germany, but all of these do have compliance considerations. To address the multitude of combinations of the above, there are new tools being developed and expanded in the privacy sector which are disparate in many ways from the security sector. The Data Privacy Software market size in 2022 is $US3.26B in 2022 but has a CAGR of 40.8% leading to a projected size of $US25.85B by 2029 (Fortune Business Insights, 2022). North America alone will grow from $US682.9M to $US9.3B by 2029 (See Figure 1).

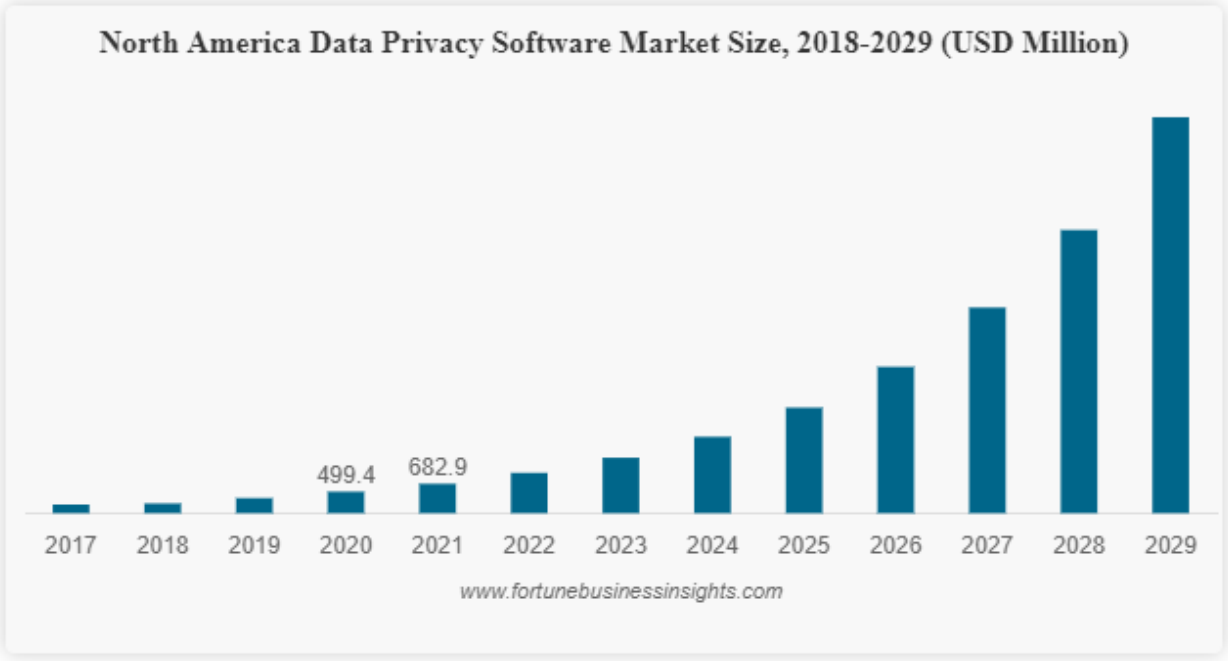

Figure 1 North America Data Privacy Software Market Size (Fortune Business Insights, 2022)

3.2. Technology Solutions

Several solutions exist in this growing Privacy Enabling Technologies (PET) space, many of these can be grouped together, but custom offerings are appearing in the market and in conferences that cannot always be categorized early.

Traditional data masking techniques include obfuscation (signal-to-noise ratio reductions to hide data), anonymizers (hiding identity or limiting cross-session tracking), pseudonymization (replacing protected data with different identifiers), and data minimization to reduce the overall amount of work performed by remaining PETs.

Self-sovereign identity solutions take blockchains, smart contracts, and an ecosystem approach to exchange of information and verifiable rules for how that data is accessed, used, stored, deleted, and shared. Protected data is stored either off-chain, or encrypted on the chain, and the sharing of that data between participants in the ecosystem is managed through the smart contracts of the chain, recording agreements, limiting use, and setting up the rules for data disposition.

Homomorphic encryption and Zero-Knowledge Proofs (ZKP) are mathematical concepts related to encrypted data whereby an operation such as a query can be performed on the ciphertext without decrypting it. In the hospital example from above, consider a study participant who is being evaluated for treatment of a new concern that arose during the study. While the patient's research status is protected (whether they are in the control group or the treatment group), the treating physician may need to consider a pharmaceutical treatment for the new concern. They query the system to find out the answer to a query that includes the patient's research status, and if they have allergies to the potential medication, and if there are drug-interactions with the research study course of treatment. With zero-knowledge proofs and homomorphic encryption, an authorized inquiry can get an answer for situations like the above without exposing the underlying data, but if the query isn't carefully constructed to be only what's necessary, data can be exfiltrated from multiple queries or overly broad queries.

Differential privacy is another tool based on mathematics. In hospital examples, like the one above, the concern is that medical data will be revealed, but if any given datum or small collection of data were to have a sufficient probability of being false, or untrue, then the release of information would not be damaging to an individual and the aggregate dataset would still be useful for use cases like that of the hospital field trial. For example, modifying data with an acceptable amount of noise to reduce the probability of identifying an individual but still providing a sufficiently reliable result to a query. For example, a query made with an exact age may be used to identify an individual. Adding noise to the data so an age range is created lessens the probability of identifying an individual but still returns clinically relevant information, such as adding sufficient noise to the age data so the query returns results for individuals aged from 34 to 38.

Federated Learning, also Privacy Enhanced Federated Learning (PEFL) is a subset of machine learning and artificial intelligence for using multiple datasets without sharing data (M. Hao, 2020). The data are distributed across independent local hosts and algorithms are trained for each host or dataset and for contributing to aggregate observation. This approach allows for data to be neither centralized, nor homogenous in its structure or distribution. What is shared are the resultant weights and biases for neural networks which can then be used in aggregate without risk of sharing private data used to obtain the trained network.

Secure Multi-Party Computation (SMPC or MPC) is based on the idea of several parties or systems who work together to solve problems using disparate stores of data, not unlike Federated

Learning, but in SMPC the objectives, rules, and model parameters are shared using finite field cryptographic tooling.

3.3. Intersection of Tools Supporting Technology and Compliance

There exist several different hybridized versions of technical solutions and approaches from Technology Solutions, some integrating cryptography in various steps for additional protection from accidental exposure or malicious actors. There also exist several tools for enterprise compliance with privacy regulation, these help to automate the auditing of datastores, databases, files, as well as system interfaces and communication tools for the storage or transit of data that is likely to be protected. The identification of that likelihood can range from simple rules to machine learning based on general training or custom training for specific jurisdictions and industries. The newest entrants into the tooling market are development operations tools (DevOps and PrivOps) where engineers using build tools for custom software development have steps where fields likely to have protected data are identified, in part to help train the privacy engineers as well as identify potential areas inside the software for additional protection.

4. Risks and Threats to Protected Data

Risks resulting in data exposure are different from the risks of exposure. Several risks can result in exposure, these range from accidental, to intentional, and even established markets for this protected information.

4.1. Risks Resulting in Exposure

Risks resulting in exposure include accidental exposure (e.g., misplaced documents), endpoint sale of data, online tracking (e.g., browser fingerprinting, cookies, etc.) intentional aggregators and profile building, linked data sets building identifiable collections (unintentional or otherwise), active fraud (these often have colloquial terms like “social engineering”, “phishing” and “spear-phishing”), real-time markets for data (whether sold or shared), insider threats, data loss or deletion (leading to misidentification or misclassification which can result in as many concerns as incorrect data), intentional exfiltration, technical vectors (insufficient security, unpatched systems, trojans, compromised websites, mobile and personal devices, removable media, poor configuration management, and access to local or cloud servers), as well as several other security compromises from various actors with different intentions (hactivism, cybercriminal networks, disgruntled employees, etc.).

Some of these risks may be more likely than others, and some risks may have smaller or more targeted datasets. The granularity and compartmentalization of the protected assets can limit the scope of a breach and each of the above risks has its own set of potential mitigation steps. This is an area where there is a significant overlap with security, however, some of the risks to exposure, above exist outside that intersectionality.

Risks where the entity or system on the other side of the endpoint has access to protected data (as defined by the laws of the prevailing jurisdiction) don’t necessarily have a security solution;

these include the intentional collection and sharing of data by the operators of that endpoint. Browser fingerprinting, activity monitoring (e.g., how quickly or slowly the user scrolls past different content, or when and where a mouse hovers, or what videos are played and when they are stopped), markers from previous websites, these are all aspects of how even unintentional data is collected. When that data is then shared with larger networks, when linked (or collections of unlinked data) are sold in real-time bidding markets (ICCL, 2022), or when it is retained and associated with future transactional records, individuals and corporations lose control over data that may be sensitive or protected.

Empirical evidence on the public understanding of the scale and scope of what is tracked seems to indicate a significant underestimation of how much data on individuals is collected each day. The Irish Council on Civil Liberties (ICCL) published a breakdown by European and US locality showing this data being sold in the greater than $US117B Real-Time Bidding (RTB) network market, shown in Figure 2.

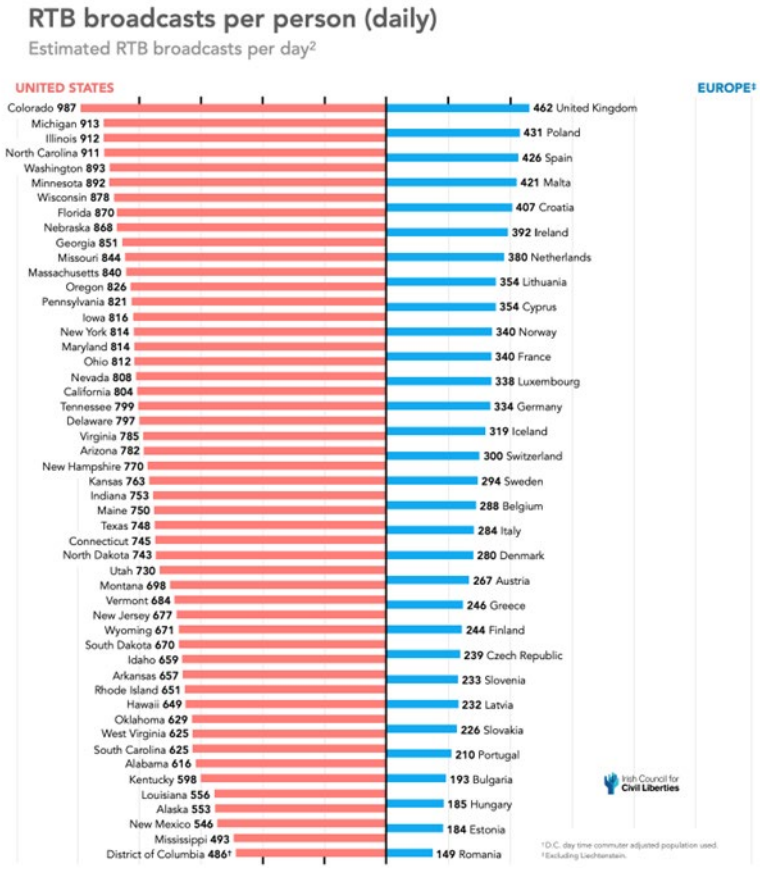

Figure 2 Real-Time Bidding Broadcasts Per Day (ICCL, 2022)

4.2. Risks of Data Exposure

Risks of exposure, what happens after the data is in the hands of another party, focuses on two primary aspects, the use of that data by malicious actors, and the regulatory and legal response.

Operational risk is not addressed here, because this paper focuses upon data exposure, not necessarily data obfuscation, modification, or deletion as is common in ransomware attacks.

4.2.1. Malicious Actors

Buyers in the markets described in Risks Resulting in Exposure include sales and marketing organizations, but the buyers also include malicious actors who seek to monetize the data collected. This can be accomplished through resale (there are underground marketplaces specializing in the sale of personal information), through direct ransom for the data (pay to have it returned and presumably not used or resold), and through extortion and the threat of public release of the data.

Ransomware used to be primarily about encryption and extortion payments to return to normal operations, but ransomware attacks that exfiltrated data increased from 22% of cases in Q2 2020 to a massive 81% in Q2 2021 (~270% YOY increase) (Schein 2022). These actors have recognized that threats of exposure can be part of criminal extortion, and some of these actors, when stymied by corporate victims that are reluctant to pay, have turned to threats (and action) of direct engagement with the individual people whose data is the subject of the breach.

4.2.2. Regulatory, Reputational and Legal Response

Those organizations who find that data in their stores has been compromised also find that there may now be both legal responsibilities and regulatory notifications required. Some jurisdictions require notification of real or potential breaches of protected data; this can be based upon the size of the company, the size of the breach, the content of the breach, and the industry within which the company may play. The legal impact may also involve public or private civil responses from those who suffered harm through the breach; depending on several factors, including the jurisdiction, this could involve significant damages.

In addition to the legal requirements on notification, there may be regulatory responses that could involve significant fines. In the European Union, significant violations of the General Data Protection Regulation (GDPR) can result in fines up to 4% of the corporation's gross annual revenue for each offense. Notifications also carry with them the reputational risk to the organization, potentially signaling to the market that they may be ineffective or careless with protecting this sensitive data. Starting with the probability of breach based on controls and tooling; then comparing that to the legal, regulatory, and reputational price organizations may pay in the event of a breach, an organization can balance risk and plan for mitigation investments.

5. Technology Policy and Privacy

The disparate technology policy and regulation across multiple jurisdictions makes this discipline difficult to standardize. These policies have an impact on the technology and are critical to understand for businesses who are engaged in commerce in these jurisdictions. This

paper is not intended as legal advice, please check with your own legal counsel as this may be out of date or incomplete.

5.1. United States

The United States does not have comprehensive national preemptive consumer privacy policy, and although talks on this front continue as of this paper's submission, rapid changes in that status are not anticipated. This said, there are several privacy legislative initiatives, some of which are laws, that cover specific sectors and actors. Examples of areas where laws exist in the US are in health privacy, finance, and protecting children. Other actions have an impact on privacy including cybersecurity, trade, and restrictions on governmental actions. Existing state legislation enacted over the last couple years can make for difficult corporate navigation of privacy rulemaking.

The growth of privacy initiatives around legislative action in the United States does show tremendous interest and inertia; this topic appears to be growing, but the resultant proposals are not consistent, and due to that non-standard deployment, privacy compliance is also growing in expense and potential for errors across jurisdictions. The growth of state privacy legislation can be seen in **Error! Reference source not found.** (IAPP 2022).

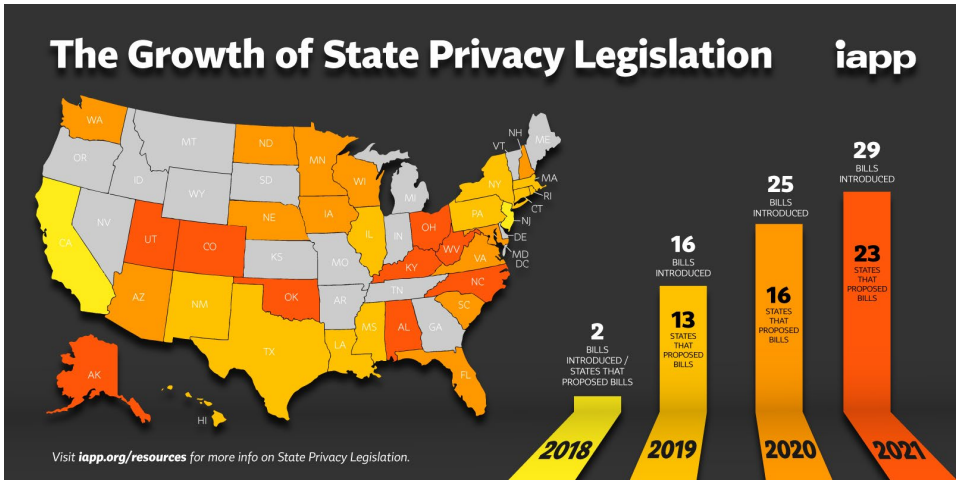

Figure 3 Growth of State Privacy Regulation

Currently, five states have enacted laws addressing private data and enterprise responsibilities, California (California Consumer Privacy Act & California Privacy Rights Act), Colorado (Colorado Privacy Act), Virginia (Virginia Consumer Data Protection Act), Utah (Utah Consumer Privacy Act), and Connecticut (Connecticut Data Privacy Act). Additional bills were considered during this 2022 legislative session, visible in Figure 4 (IAPP 2022).

The five states that have laws on the books, have some commonalities, but some dramatic differences. The consumer right to access, rectify, delete, and restrict records exists in all five, as do the business responsibilities for opt-in as the default, transparency, and limits on processing based on purpose of the data. Additionally, enterprises cannot discriminate against consumers who are exercising their rights.

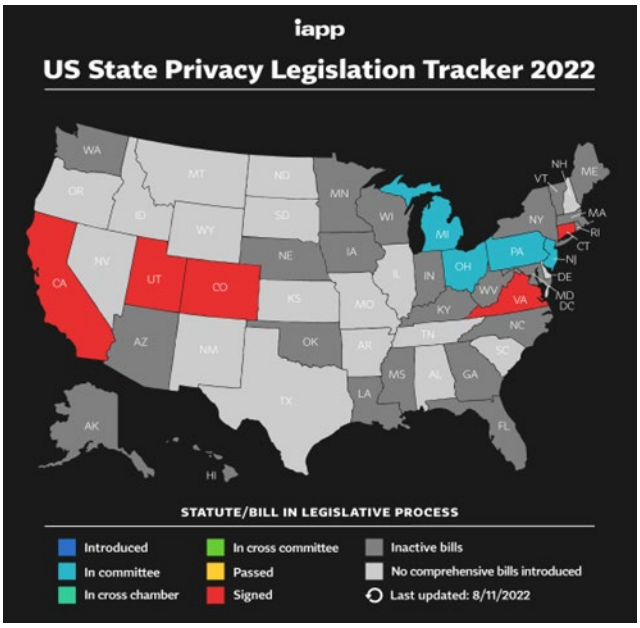

Figure 4 US State Privacy Legislation Tracker 2022

Utah does not have restrictions on automated decision-making based on protected data while the other four states do restrict this (California added this in the CPRA). In Colorado, Connecticut, and Virginia, additional rules are in place for risk assessments and those three states rely upon the Attorney General to prosecute cases. California is currently the only state in the United States with a private right of action for privacy violations. Detailed rulemaking is still taking place in these states, so the fine-grained details of these regulations are still subject to some level of change.

5.2. International

The European Union and eighteen other countries (excluding the United States) have comprehensive consumer privacy legislation and regulation, these include Argentina, Armenia, Australia, Benin Republic, Brazil, Canada's Personal Information Protection and Electronic Documents Act (PIPEDA), China, Colombia, European Union, Hong Kong, Israel, Kenya, New Zealand, Nigeria, Philippines, Singapore, South Africa, South Korea, and Turkey.

As in the United States the consumer right to access, rectify, delete, export, and age-based opt-in options are predominantly consistent across these. The default opt-in or opt-out as well as the consumer right not to be subject to automated decision-making seem to be inconsistently applied or available. Business obligations in these markets often include transparency, purpose limitations for data retention, data minimization, security requirements (somewhat varied), record keeping and breach notification requirements. Requirements for data protection officers, international data transfer restrictions, preemption, and sector-specific regulation are inconsistently across these jurisdictions. The EU, Australia, and a few others require privacy by design raising questions about enforcement interpretation with older systems and processes.

6. Privacy Standards Community

Tool development for inconsistent requirements like those listed in Section 5.2 is difficult due to the lack of economies of scale across ecosystems, in some cases this is made even more difficult by conflicting expectations (e.g., opt-in versus opt-out). The lack of mature, consistent, international standards and due to the variability in technology options, network operators and enterprises are left buying privacy compliance tools that cannot interoperate effectively with other tools in the space, where one tool can't complete all the tasks, and where incumbent vendors and contracting rules inhibit innovation.

For enterprises and operators who are looking for standards that support operations and enable innovation, Table 1 is meant to serve as an incomplete guide:

Table 1 Standard Development Organizations and Initiatives

Organization	Document/Specification/Initiative	Notes
ISO/IEC	Privacy Information Management Systems Scheme (PIMS Scheme) 27701 & 27702 (ISO 2019)	
National Institute of Standards and Technology (NIST)	Privacy Framework	1.0 (Jan 16, 2020) United States Dept. of Commerce
International Association of Privacy Professionals (IAPP)	www.iapp.org	Policy-neutral information privacy organization. Certification and professional credentialing
World Wide Web Consortium (W3C)	https://www.w3.org/Privacy/IG/	Public-interest non-profit web-focused privacy standards group.
3rd Generation Partnership Project (3GPP)	https://www.3gpp.org/	Mobile broadband standards development organization. Consumer data privacy is a working group topic for R18 in 2022.
Connectivity Standards Alliance (CSA/Matter)	https://csa-iot.org/all-solutions/matter/	Consumer IoT standards development organization.
Institute for Electrical and Electronic Engineers (IEEE)	https://digitalprivacy.ieee.org/	Privacy collaboration, policy, and research for individual private needs online

Future Directions Digital Privacy Initiative		
Wireless Broadband Alliance (WBA)	https://wballiance.com/wi-fi-imsi-privacy-protection/	Permanent IMSI privacy protection initiative
Multi-Party Computation Alliance (MPC Alliance)	https://www.mpcalliance.org/	Standards and advocacy group focused on the adoption of MPC technology

7. Conclusion

Privacy technology is advancing on several fronts, technology, policy, standards, and the discipline's overlap with security. Entirely new disciplines like Privacy Engineering are being developed (Carnegie Mellon 2022), nascent tools are being brought to the market, standards are catching up, and legal, compliance and regulatory frameworks are being established and renovated. The objective of this work was to provide the reader the tools necessary to begin their own assessment of the posture of the reader's own organization, to understand the different aspects of the space, better evaluate the nuance and differences between privacy and security and begin to learn to discern where and how privacy technologies should be applied within their influence.

Abbreviations

CAGR	Compound Annual Growth Rate
CCPA	California Consumer Privacy Act
CDPA	Connecticut Data Privacy Act
CPA	Colorado Privacy Act
CPRA	California Privacy Rights Act
DevOps	Development Operations
DNS	Domain Name System
DOCSIS	Data Over Cable Service Interface Specification
GDPR	General Data Protection Regulation
HTTPS	HyperText Transfer Protocol Secure
ICCL	Irish Council for Civil Liberties
MPC	Multi-Party Computation
PIMS Scheme	Privacy Information Management Systems Scheme
PIPEDA	Personal Information Protection and Electronic Documents Act
PrivOps	Privacy Operations
RTB	Real-Time Bidding
SCTE	Society of Cable Telecommunications Engineers
SMPC	Secure Multi-Party Computation
UCPA	Utah Consumer Privacy Act
VCDPA	Virginia Consumer Data Protection Act
W3C	World Wide Web Consortium
ZKP	Zero-Knowledge Proofs

Bibliography & References

Bambauer, Derek E. "Privacy versus security." *J. Crim. L. & Criminology* 103 (2013): 667.

Google, 2022: Transparency Report on Network Traffic, HTTPS encryption on the web: https://transparencyreport.google.com/https/overview?hl=en

M. Walker, B. Scriber, K. Shockey, D. Slagle "Have Your Privacy Cake and Eat It Too: How New Technologies Look to Protect Consumer Privacy While Promoting Innovation." *TPRC47: The 47th Research Conference on Communication, Information and Internet Policy*. 2019.

Fortune Business Insights, Report ID: FBI105420, Data Privacy Software Market Size, Share & COVID-19 Impact Analysis, By Deployment (On-premises and Cloud), By Application (Compliance Management, Risk Management, Reporting & Analytics, and Others), By Organization Size (Small & Medium Enterprises (SMEs) and Large Enterprises), By Industry (BFSI, IT and Telecommunication, Government, Manufacturing, Retail, Healthcare, and Others), and Regional Forecast, 2022-2029

Dilmegani, Cem, Information Security Privacy Enhancing Technologies and Use Cases, https://research.aimultiple.com/privacy-enhancing-technologies/ *AI Multiple* (2022)

M. Hao, H. Li, X. Luo, G. Xu, H. Yang and S. Liu, "Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6532-6542, Oct. 2020, doi: 10.1109/TII.2019.2945367.

O. Goldreich, S. Micali, A. Wigderson "*How to Play any Mental Game or A Completeness Theorem for Protocols with Honest Majority*". STOC 1987: 218-229

ICCL (Irish Council for Civil Liberties), *Note on scale of Real-Time Bidding data broadcasts*, https://www.iccl.ie/wp-content/uploads/2022/05/Mass-data-breach-of-Europe-and-US-data-1.pdf, 2022

M. Schein, 16 Mar 2022, Marsh McLennan, National Co-Chair Cyber Center of Excellence

IAPP International Association of Privacy Professionals), 2022 Resources: https://iapp.org/resources/article/us-state-privacy-legislation-tracker/ , https://iapp.org/resources/article/the-growth-of-state-privacy-legislation-infographic/ , https://iapp.org/media/pdf/resource_center/global_comprehensive_privacy_law_mapping.pdf

ISO, 2019, https://www.iso.org/news/ref2419.html

NIST, Privacy Framework 1.0, https://www.nist.gov/privacy-framework/privacy-framework

Carnegie Mellon, Privacy Engineering, 2022, https://privacy.cs.cmu.edu/

